

PENGANTAR  
**ROBOT  
VISI**



PENGANTAR ROBOT VISI

DAHLAN ABDULLAH, CUT ITA ERLIANA, FADLI SYAH



DAHLAN ABDULLAH  
CUT ITA ERLIANA  
FADLI SYAH



PENGANTAR  
**ROBOT  
VISI**

# PENGANTAR ROBOT VISI

**Dahlan Abdullah  
Cut Ita Erliana  
Fadlisyah**

**Diterbitkan Oleh:**



## **PENGANTAR ROBOT VISI**

**Oleh: DAHLAN ABDULLAH, CUT ITA ERLIANA, FADLI SYAH**

Hak Cipta © 2017 pada Penulis

Layout & Cover Design : *M. Rizki, S.Kom.I*

Pracetak dan Produksi : *CV. Sefa Bumi Persada*

*Hak Cipta dilindungi undang-undang.*

*Dilarang memperbanyak atau memindahkan sebagian atau seluruh isi buku ini dalam bentuk apapun, baik secara elektronik maupun mekanis, termasuk memfotokopi, merekam atau dengan sistem penyimpanan lainnya, tanpa izin tertulis dari Penulis*

Penerbit:

**SEFA BUMI PERSADA**

Jl. B. Aceh – Medan, Alue Awe - Lhokseumawe

email: sefabumipersada@gmail.com

*Telp. 085260363550*

*Printed in Lhokseumawe, 2017*

*Pengantar Robot Visi*

**ISBN 978-602-6960-35-1**

1. Hal. 91 : 18 x 25,5 cm

I. Judul

# KATA PENGANTAR

Dengan mengucapkan puji dan syukur kehadiran Allah SWT, di mana atas rahmat dan karunia-Nya Penulis telah dapat menyelesaikan buku yang berjudul "*PENGANTAR ROBOT VISI*".

Rasa terima kasih penulis ucapkan juga untuk Bapak Prof. Dr. Apridar, SE, M.Si selaku rektor Universitas Malikussaleh, yang telah memotivasi penulis untuk terus menyelesaikan karya ini, istriku Cut Ita Erliana, ST, MT, Bapak Fadlisyah, S.Si, MT dan pihak-pihak yang telah memberi bantuan dan bimbingan sehingga penulis dapat menyelesaikan naskah Robot Visi ini, yang tidak dapat disebutkan satu per satu.

Adapun materi-materi yang dibahas di dalam buku ini mencakup: Kehidupan awal robot, Robot interaktif, Sensor komputasi, Persepsi, VH construction, Pendeteksian gerakan tubuh, Visual Servoing, Estimasi gerakan 3D dari citra stereo, dan Sistem *Catadioptric*.

Penulis hanya mengucapkan selamat membaca.

Penulis

# DAFTAR ISI

Kata Pengantar	
Daftar Isi	
BAB 1 – “KEHIDUPAN” AWAL ROBOT	1
1.1 Sejarah robot secara industrial	2
BAB 2 – STUDI AWAL INTERAKSI MANUSIA-ROBOT PEER-TO-PEER	7
2.1 Pengenalan	7
2.1.1 Proyek HRI peer-to-peer	7
2.1.2 Sistem operasi HRI	8
2.2 Evaluasi awal	9
BAB 3 – ROBOT INTERAKTIF SECARA SOSIAL	13
3.1 Pengenalan	13
3.1.1 Sejarah robot sosial	13
3.1.2 Robot sosial dan penanaman nilai-nilai sosial : konsep dan definisi	15
3.1.3 Robot interaktif secara sosial	15
3.1.4 Robot interaktif secara sosial, kenapa ?	16
3.2 Metode	17
3.2.1 Pendekatan desain	17
3.2.1.1 Biologically inspired	18
3.2.1.2 Rancangan fungsional	19
3.2.2 Berbagai isu desain	20
3.2.3 Embodiment	21
3.2.3.1 Morfologi	22
3.2.3.2 Pertimbangan desain	22
3.2.3.3 Anthropomorphic	23
3.2.3.4 Zoomorphic	24
3.2.3.5 Karikatur	25
3.2.3.6 Fungsional	25
3.2.4 Emosi	27
3.2.4.1 Emosi buatan	27
3.2.4.2 Emosi sebagai mekanisme terkontrol	28
3.2.4.3 Speech	29
3.2.4.4 Ekspresi face	30
BAB 4 – SENSOR KOMPUTASI UNTUK “TRACKING” OBJEK	33
4.1 Pendahuluan	33
4.2 Implementasi	35
BAB 5 – ROBOT : PERSEPSI	37
5.1 Pendahuluan	37

---

5.2 Melihat segalanya	39
5.2.1 Sensor jangkauan <i>right-stripe</i>	39
5.2.2 Omnikamera aliran optis	40
5.3 Mendeteksi pergerakan dan pergerakan lanjutan objek	41
<b>BAB 6 – VISUAL HULL CONSTRUCTION</b>	<b>45</b>
6.1 Pendahuluan	45
6.2 Latar belakang	46
6.2.1 Skenario SFS	46
6.2.2 Visual Hull	47
6.3 Visual Hull Construction	48
6.3.1 Komputasi visual hull dengan irisan batas kerucut	48
6.3.2 Konstruksi dari pengaproksimasian visual hull dengan pengklasifikasian voxel	49
6.4 Visual hull construction dari <i>silhouette</i> “noise” pada waktu nyata	49
6.4.1 Kecepatan dan keakuratan metode visual hull construction berbasis voxel	49
6.4.2 Algoritma fast visual hull construction berbasis voxel menggunakan Sparse Pixel Occupancy Test (SPOT)	53
<b>BAB 7 – PENDETEKSIAN GERAKAN TUBUH</b>	<b>55</b>
7.1 Pendahuluan	55
7.2 Pemodelan probabilitas <i>single-frame</i>	56
7.2.1 Pemodelan jaringan Markov	56
7.2.2 Pembelajaran model jaringan Markov	57
7.2.3 Melibatkan pose dalam model jaringan Markov	59
7.3 Pemodelan probabilitas <i>multi-frame</i>	60
7.3.1 Batasan temporal	60
7.3.2 Inferensi	61
<b>BAB 8 – PENGANTAR PENGOLAHAN CITRA VIDEO (VIP) – PENDEKATAN DALAM ROBOT VISI</b>	<b>65</b>
8.1 Pendahuluan	65
8.2 Pengolahan citra pada mobil robot	66
8.2.1 Komputer visi dan pengolahan citra	66
8.2.2 Robot visi	67
<b>BAB 9 – VISUAL SERVOING</b>	<b>69</b>
9.1 Pendahuluan	69
9.2 Pemodelan formasi citra <i>catadioptric</i> sentral	70
9.3 Tracking dan kontrol	73
9.4 Matriks Jacobian <i>catadioptric</i>	74
<b>BAB 10 – ESTIMASI PERGERAKAN 3D DARI CITRA STEREO – FORMULASI DISKRIT DAN DIFERENSIAL</b>	<b>77</b>
10.1 Pendahuluan	77

10.2	Estimasi pergerakan	78
10.3	Pendekatan diferensial	78
10.4	Formulasi diskrit	81
BAB 11 – ESTIMASI POSE UNTUK SISTEM <i>CATADIOPTIC</i> SENTRAL : PENDEKATAN ANALITIS		83
11.1	Pendahuluan	83
11.2	Formasi citra <i>catadioptric</i>	83
11.3	Estimasi pose	86
11.3.1	Estimasi rotasi	86
11.3.2	Estimasi translasi	88
11.4	Eksperimen	90
11.5	Kesimpulan	91
Lampiran		
Daftar Pustaka		

---

---

# Bab 1

## “KEHIDUPAN” AWAL ROBOT

*One machine can do the work of a hundred ordinary men, but no machine can do the work of one extraordinary man*

(Elbert Hubbard)

Kata *robot* berasal dari bahasa Czech, *robota*, yang dapat diartikan sebagai kerja. Kamus Webster mendefinisikan robot sebagai “an automatic device that performs functions ordinarily ascribed to human beings,” yang dapat kita artikan sebagai suatu perangkat atau alat yang bekerja secara otomatis yang mampu melakukan aktifitas - aktifitas menyerupai manusia. Berdasarkan definisi di atas maka peralatan mesin cuci biasa dapat kita kategorikan sebagai sebuah robot, maka definisi ini tidak dapat mengkategorikan apa-apa yang termasuk dalam kategori robot. Jadi kita perlukan suatu definisi yang lebih tepat dalam mendeskripsikan robot sebagai sebuah mesin industri, dan untuk itu, definisi yang diberikan oleh Robot Institute of America, “A robot is a *reprogrammable multi-functional* manipulator designed to move materials, parts, tools, or specialized devices, through variable programmed motions for the performance of a variety of tasks.” Definisi yang lain adalah, “A robot is reprogrammable general - purpose manipulator with external sensors that can perform various assembly tasks.” Berdasarkan definisi terakhir ini, sebuah robot harus memiliki syarat “intelligence,” yang mana secara normal dapat kita katakan syarat “intelligence” tersebut sangat berkaitan erat dengan algoritma pemrograman komputer untuk aplikasi kontrol dan sistem sensor.

Kata *robot* sendiri dikenalkan di dalam bahasa Inggris pada tahun 1921 oleh seorang artis, Karel Capek dalam perannya pada drama satire, R.U.R. (Rossum’s Universal Robots). Di dalam drama tersebut, robot-robot berperan sebagai suatu mesin yang menyerupai fisik manusia, dan memiliki kemampuan bekerja yang tidak kenal lelah. Pada awalnya, robot dipabrikasi untuk meningkatkan produksi perusahaan dengan menggantikan peran manusia pekerja, tetapi pada selang waktu yang tidak lama kemudian, robot berbalik arah melawan penciptanya, dan

menghancurkan ras kehidupan manusia. Terlepas dari kekuatiran yang disampaikan di dalam drama tersebut melalui sindiran peran, permainan drama yang dilakukan Capek telah membuat kesatuan persepsi di kalangan umum manusia, bahwa, robot merupakan mesin yang menyerupai manusia yang dilengkapi suatu kecerdasan dan personalitas individual. Pandangan umum ini, makin mengakar dan menjadi persepsi baku di kalangan manusia pada umumnya, setelah film robot Jerman, *Metropolis*, pada tahun 1926 dibuat, dan dipertontonkan di New York World's Fair pada tahun 1939, yang ceritanya diperankan oleh robot Electro yang dapat berjalan sebagaimana manusia dan ditemani anjingnya Sparko. Dan mungkin juga kita masih ingat sebuah film ruang angkasa *Star Wars* dengan robotnya C3PO, yang dirilis tahun 1977. Tetapi jika kita bandingkan robot-robot imajinatif yang hadir di beberapa film tersebut, dengan robot-robot industri yang sebenarnya, maka akan timbul pertanyaan pada diri kita, "Apakah sudah secanggih itukah robot - robot yang telah dikeluarkan oleh industri dunia ?." Di dalam film "*Artificial Intelligence*," yang dirilis beberapa tahun yang lalu, menceritakan bagaimana robot "anak kecil" yang dilengkapi emosional seperti manusia, merindukan peran ibu yang dapat menyayangi dan memperlakukan dirinya (robot "anak kecil") sebagaimana anak manusia lainnya yang normal. Apakah telah sejauh inilah kontribusi yang telah diberikan oleh ilmuwan - ilmuwan robot dunia terhadap robot ciptaannya, dan seberapa mungkinkah ini dapat direalisasikan ?.

### 1.1 Sejarah robot secara industrial

Pekerjaan awal yang telah membuka pintu industri robot hari ini dijejaki oleh suatu periode pada saat sedang berlangsungnya perang dunia ke - II. Dimulai oleh suatu program riset pada akhir tahun 1940 di Oak Ridge dan Argonne National Laboratories, mengenai pengembangan beberapa manipulator mekanis terkontrol untuk penanganan materi-materi radioaktif. Sistem ini merupakan jenis "master - slave" yang dirancang untuk menghasilkan suatu pergerakan tangan dan lengan yang dapat dipercaya yang dilakukan oleh seorang operator, dalam hal ini manusia. Manipulator "master" tersebut dikendalikan oleh seorang user melalui sederetan pergerakan, sementara manipulator "slave" menduplikasi unit "master" sedekat mungkin. Kemudian, manipulator mekanis terkontrol tersebut diberikan suatu timbal balik gaya, untuk mengkopling secara mekanis pergerakan dari unit "master" dan "slave" dengan tujuan operator dapat merasakan suatu gaya seakan - akan mereka bekerja mengembangkan di antara manipulator "slave" dan lingkungannya. Di pertengahan tahun 1950 kopling mekanis telah tergantikan oleh tenaga hidrolis dan elektrik, yang penerapannya pada beberapa manipulator seperti, General Electric's Handyman dan, Minotaur I yang diciptakan oleh General Mills.

Proyek manipulator master - slave secara cepat diikuti oleh perkembangan sistem - sistem berkaitan yang lebih mampu bekerja secara otomatis, dan beroperasi secara repetitif. Dalam pertengahan tahun 1950, George C. Devol merancang suatu perangkat yang dia sebut "*programmed articulated transfer device*," sebagai sebuah

manipulator yang memiliki operasi yang dapat diprogram (dirubah) dan diikuti oleh sebarisan pergerakan yang dapat ditentukan dengan memberikan beberapa instruksi di dalam program pada sistem tersebut. Pengembangan lebih jauh dari konsep tersebut, dilakukan oleh Devol dan Joseph F. Engelberger, yang telah berjasa membuka sejarah awal robot industri, yang diperkenalkan oleh Unimation Inc. pada tahun 1959. kunci dari perangkat ini adalah peran sebuah komputer yang melakukan sharing dengan manipulator untuk menghasilkan suatu mesin yang dapat berpikir atau mengambil beberapa keputusan secara otomatis menghadapi berbagai tugas atau kerja yang variatif. Tidak seperti mesin automasi biasa, robot tersebut dapat diprogram dan dirancang ulang dengan biaya yang relatif rendah, untuk melakukan tugas - tugas sesuai kebutuhan manufaktur.

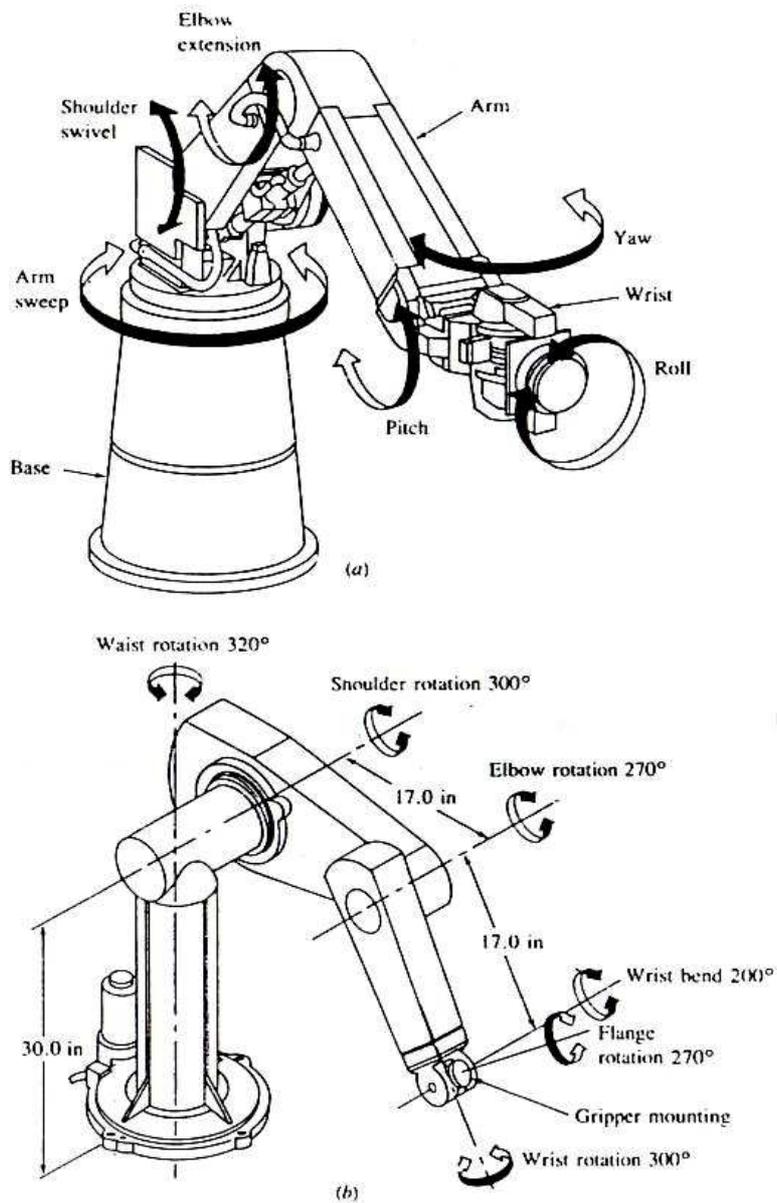
Pada tahun 1960-an, kefleksibelan mesin (robot) mampu ditingkatkan secara signifikan dengan mengintegrasikan suatu sistem sensor pada mesin (robot). Awal dari dekade ini, H. A. Ernst (1962) mengemukakan suatu proyek pengembangan sebuah *computer-controlled mechanical hand* dengan sebuah sensor tactile, yang disebut MH - I, suatu perangkat yang dapat “merasakan” beban, dan memanfaatkan informasi tersebut untuk mengontrol tangan (*hand-machine*) supaya beban dapat ditumpuk tanpa bantuan dari operator. Dan proyek ini merupakan contoh pertama dari sebuah robot yang dapat mengadaptasikan sifat-sifat atau prilakunya di dalam suatu lingkungan yang tidak terstruktur. Sistem manipulasinya, terdiri dari sebuah manipulator ANL model - 8 dengan 6 derajat kebebasan yang dikendalikan oleh sebuah komputer TX-O melalui suatu perangkat antarmuka (*interface*). Program riset ini kemudian menjadi bagian dari proyek MAC, dengan penambahan sebuah kamera televisi terhadap manipulator sebagai awal dari riset “*machine perception*.” Di dalam periode yang sama pula, Tomovic dan Boni (1962) telah mengembangkan suatu prototipe tangan manusia yang dilengkapi sebuah sensor tekanan yang berfungsi menyeleksi keputusan pada berbagai objek dan mensuplai sinyal-sinyal input sebagai timbal baliknya terhadap sebuah motor untuk penginialisasi satu dari dua pola-pola pegangannya. Pertama-tama tangan melakukan kontak dengan objek, informasi sepadan mengenai berat dan ukuran objek yang diperolehnya selanjutnya dikirim ke komputer oleh beberapa elemen yang sensitif terhadap tekanan. Pada tahun 1963, American Machine and Foundry Company (AMF) memperkenalkan robot komersilnya, VERSATRAN. Dimulai pada tahun yang bersamaan pula, berbagai rancangan lengan sebagai manipulator telah dikembangkan, seperti lengan Roehampton dan, lengan Edinburgh.

Di akhir tahun 1960-an, McCarthy (1968) dan beberapa temannya di Stanford Artificial Intelligence Laboratory telah memberitahukan tentang pengembangan komputer dengan tangan, mata, dan telinga. Mereka mendemonstrasikan sebuah sistem yang mengenal pesan-pesan suara, “melihat” balok-balok yang disebarkan di atas meja, dan mampu memanipulasi kesemuanya sesuai dengan instruksi - instruksi yang telah diinputkan. Selama periode ini, Pieper (1968) telah mengkaji permasalahan kinematis terhadap *computer-controlled manipulator*, sementara Kahn dan Roth

(1971) melakukan penganalisaan terhadap aspek dinamis dan kontrol dari suatu robot lengan menggunakan "bang-bang control."

Pada saat bersamaan, negara Jepang mulai tertarik untuk memanfaatkan kepotensialan robot, dan ini dibuktikan pada awal tahun 1968, perusahaan Kawasaki yang bergerak dalam bidang industri alat-alat berat melakukan negoisasi untuk memperoleh lisensi dari Unimation untuk pengembangan robotnya. Sesuatu yang spektakuler di dalam perkembangan industri robot, adalah dengan diciptakannya sebuah "walking truk" untuk kebutuhan tentara U.S oleh General Electric Company pada tahun 1969. Pada tahun yang sama, robot lengan Boston diciptakan, dan tahun berikutnya robot lengan Stanford dikembangkan, dimana robot tersebut dilengkapi dengan sebuah kamera dan komputer kontroler. Selalu saja, proyek pengembangan robot yang serius mengawali proyeknya dengan pengembangan robot lengan yang biasanya digunakan sebagai manipulator robot. Robot lengan Stanford juga memiliki kemampuan yang lebih otomatis, dan ini dibuktikan oleh kemampuan robot tersebut yang mampu membentuk suatu tumpukan balok secara otomatis dengan strategi yang lebih variatif. Pada tahun 1974, Cincinnati Milacron mengenalkan robot industri pertamanya (berbasis computer-controlled) yang disebut "The Tomorrow Tool,"  $T^3$ , yang mampu mengangkat lebih dari 100 lb dan memindahkan objek - objek tersebut tepat pada jalur-jalur yang telah disediakan.

Selama era 1970-an, riset-riset yang dilakukan terfokus hanya pada sistem sensor eksternal untuk memfasilitasi operasi-operasi yang bersifat manipulatif. Di Stanford, Bolles dan Paul (1973) memanfaatkan visual dan force feedback, untuk mendemonstrasikan robot lengan Stanford yang dikoneksikan dengan komputer PDP - 10 untuk kegunaan perakitan pompa air otomatis. Pada waktu yang bersamaan, Will dan Grossman (1975) pada perusahaan IBM mengembangkan suatu manipulator berbasis computer-controlled dengan sentuhan dan sensor untuk melakukan perakitan mekanis 20 bagian dari mesin ketik biasa. Di laboratorium Draper, Nevins dan tim kerjanya (1974) memperkenalkan suatu teknik pensensoran berbasis compliance. Pekerjaan ini selanjutnya dikembangkan menjadi instrumentasi sebuah perangkat compliance pasif yang disebut juga RCC, *Remote Center Compliance*. Bejczy (1974) di Jet Propulsion Laboratory mengimplementasikan sebuah komputer berbasis teknik *torque-control* untuk proyek eksplorasi ruang angkasa. Semenjak itu, berbagai variasi metode kontrol telah diajukan untuk pemanfaatan manipulator mekanis.



Gambar 1.1 (a) Robot lengan  $T^3$  Milacron Cincinnati, (b) robot lengan PUMA 560 series



Gambar 1.2 Robot astrobiologis, Zoë, digunakan dalam proyek Atacama

# Bab 2

## STUDI AWAL INTERAKSI MANUSIA - ROBOT PEER-TO- PEER

### 2.1 Pengenalan

#### 2.1.1 Proyek HRI peer-to-peer

Semenjak february 2004, NASA telah merencanakan "Vision for Space Exploration," VSE, dengan tujuan utamanya untuk mengimplementasikan sebuah program jangka panjang yang serius untuk penjelajahan luar angkasa dan solar sistem. Misi awalnya adalah mendaratkan robot penjelajah pada tahun 2008 ke bulan dan membuka kemungkinan bagi manusia mengeksplorasi planet Mars. Untuk meraih tujuan tersebut, maka diperlukan suatu kolaborasi antara manusia dan robot agar tugas-tugas misi dasar tercapai (perakitan struktur, inspeksi, dan perawatan).

Sasaran dari proyek "interaksi manusia dan robot secara peer-to-peer," (P2P HRI), adalah untuk membangun suatu teknik atau pendekatan HRI yang memungkinkan robot dan manusia bekerja bersama secara efektif. Adalah tidak mungkin kemampuan robot akan setara dengan manusia, jadi istilah peer-to-peer disini lebih menekankan suatu ide tentang manusia dan robot untuk bekerja sebagai partner, dengan kata lain, peran keduanya harus seimbang mungkin di dalam menjalankan suatu misi yang telah diprogramkan.

Pada bab ini kita akan memfokuskan pada perancangan tim manusia - robot yang efektif dalam melaksanakan eksplorasi ruang angkasa dengan membangun suatu mekanisme interaksi secara alami, mengurangi beban kerja yang berkaitan dengan kontrol robot, dan memaksimalkan kerja dimana robot dan manusia mampu menyelesaikan pekerjaan secara bersama-sama. Penambahan lain, kita akan membahas suatu pendekatan yang memungkinkan tim manusia- robot mengoperasikan "ground control" secara tidak terikat.

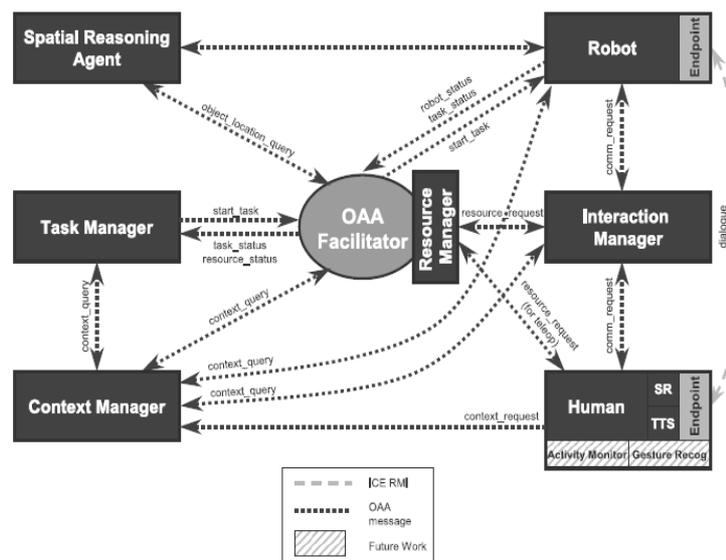
Ada tiga komponen utama di dalam pendekatan yang kita bahas di sini, pertama, kita akan membentuk suatu framework interaksi yang disebut juga "Human-Robot Interaction Operating System" (HRI/OS). HRI/OS dirancang untuk memungkinkan manusia dan robot mengkombinasikan dialog dan pemecahan masalah. Kedua, kita memanfaatkan arsitektur kognitif komputasional untuk memodelkan sifat atau perilaku manusia dan membuat hubungan "saling memahami" satu sama lain antara robot dan manusia. Terakhir, kita akan membuat

suatu prosedur evaluasi menggunakan tim manusia- robot, lingkungan analog, dan beberapa metrik HRI secara kuantitatif.

### 2.1.2 Sistem operasi HRI

Dalam menata robot dan manusia agar bekerja secara efektif bersama, maka antara keduanya, baik robot maupun manusia, harus mampu berkoordinasi mengenai keseluruhan kemampuan, tujuan, dan hal-hal yang akan dicapainya. HRI/OS merupakan suatu sistem agen yang memberikan koordinasi dan pelayanan interaksi untuk kelompok kerja yang terdiri dari manusia dan robot. Fitur utama dari HRI/OS adalah untuk memungkinkan robot menanyakan pertanyaan kepada manusia, supaya robot dapat memperoleh dukungan untuk penyelesaian tugas-tugas kognisi dan persepsi.

Tujuan perancangan HRI/OS tiada lain adalah untuk mendukung performance dari tugas - tugas operasional. Tugas-tugas operasional untuk eksplorasi ruang angkasa meliputi antara lain : konstruksi, inspeksi dan perawatan struktur, dan transportasi dan pengumpulan sumber daya *in-situ*. Versi HRI/OS yang aktual diimplementasikan dengan memanfaatkan OAA (*Open Agent Architecture*) dan mesin komunikasi internet.



Gambar 2.1 HRI/OS sebagai suatu sistem berbasis agen

Komponen-komponen utama dari HRI/OS (lihat gambar 2.1) adalah :

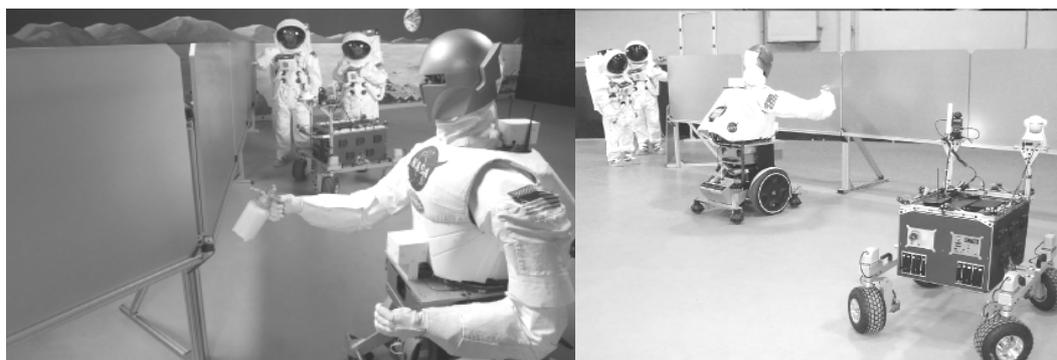
- 1) *Task Manager* : *Task Manager* (TM) merupakan unit yang bertanggung jawab terhadap pengkoordinasian dan manajemen eksekusi tugas-tugas operasional.

- 2) *Resource Manager* : *Resource Manager* (RM) merupakan unit yang bertanggung jawab terhadap penjejakan secara kontinu, jadi ketika suatu agen menanyakan hal-hal yang berkenaan dengan tugas-tugas yang akan dilakukan, maka RM akan mengeluarkan skala prioritas dari berbagai pertanyaan yang dilontarkan oleh agen-agen tersebut.
- 3) *Interaction Agent* : *Interaction Agent* (IM) merupakan bagian yang bertugas mengkoordinasikan dialog yang terjadi diantara berbagai agen.
- 4) *Context Manager* : *Context Manager* (CM) merupakan bagian yang bertanggung jawab dalam memfasilitasi “kesadaran” yang situasional. Kongkritnya CM menjaga alur pemrosesan dari suatu sistem yang sedang bekerja, seperti, status tugas dan pengekseskuan, aktivitas agen, dialog agen, dan lain-lain.
- 5) *Spatial Reasoning Agent* : *Spatial Reasoning Agent* (SRA) merupakan unit yang menangani keambiguitasan dialog antara manusia dan robot. Implementasi aktual dari SRA harus mampu menanggulangi *frame of reference* dari ambiguitas yang meliputi berbagai referensi : *ego-*, *addresse-*, *object-*, dan *exocentric* menggunakan suatu pemodelan kognitif komputasional.

## 2.2 Evaluasi awal

Fong dan kawan-kawan pada bulan November 2005 melakukan simulasi pengembangan proyek (*seam wealding* dan inspeksi) dengan melibatkan tim robot-manusia. *Seam wealding* merupakan tugas-tugas yang dibutuhkan untuk pembangunan dan perawatan berbagai variasi struktur dari permukaan planet.

Tim manusia-robot yang dibentuk terdiri dari 5 anggota yang bekerja secara paralel (lihat gambar 2.2).



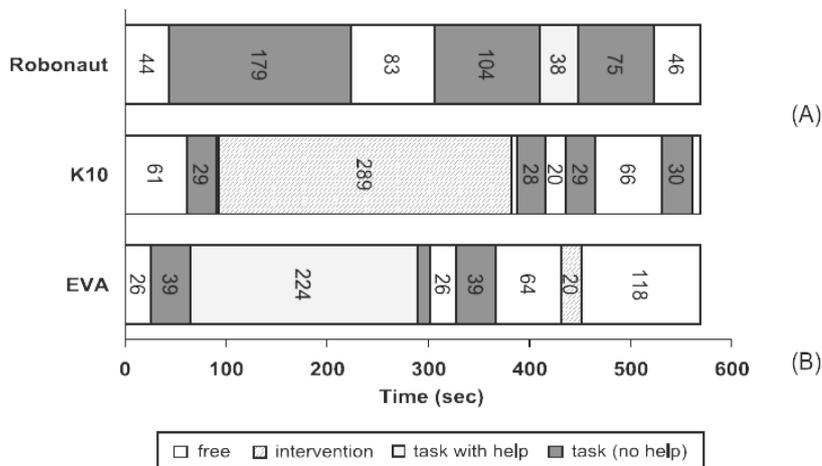
Gambar 2.2 Tim robot-manusia yang bekerja pada simulasi proyek konstruksi. Dua astronot menempatkan panel-panel pada bingkainya, dan robot mobil K10 melakukan pemeriksaan hasil dari pengelasan.

Pada setiap percobaan, tim bekerja mendapatkan panel-panel yang disediakan pada depot penyimpanan, menempatkan panel-panel tersebut ke dalam bingkai, mengelas panel-panel yang berdekatan, dan memeriksa hasil pengelasan. Dua orang yang memakai seragam ruang angkasa yang berperan sebagai astronot pada EVA (*Extravehicular Activity*, "EVA" is work performed by an astronaut outside of a spacecraft) juga membantu melakukan tugas pengambilan dan pemasangan panel-panel, yang diperlihatkan pada gambar 2.2. Robot *Robonaut Humanoid* yang dibuat oleh Johnson Space Centre bertugas sebagai pengelas menggunakan proses simulasi *seam weld*. Robot penjelajah K10 yang dibuat oleh Ames Research Centre bertanggung jawab terhadap pemeriksaan hasil dari lapisan-lapisan yang dilas. Dan personil tambahan, IVA (*Intravehicular Activity*, "IVA" is work performed inside a pressurized module--spacecraft, habitat, etc) mendukung proyek konstruksi dari dalam maket habitat.

Untuk kebutuhan *baseline*, Fong mengumpulkan keseluruhan data rincian waktu yang dihabiskan oleh masing-masing agen untuk melaksanakan tugas-tugas primitif, mengambil dan memasang panel (astronot), pemindahan antara masing-masing panel (keseluruhan), pengelasan (astronot dan robonaut), pemeriksaan hasil pengelasan (K10). Adapun rincian waktu yang diperlukan oleh masing-masing agen yang disajikan di dalam bentuk tabel dapat dilihat pada tabel berikut :

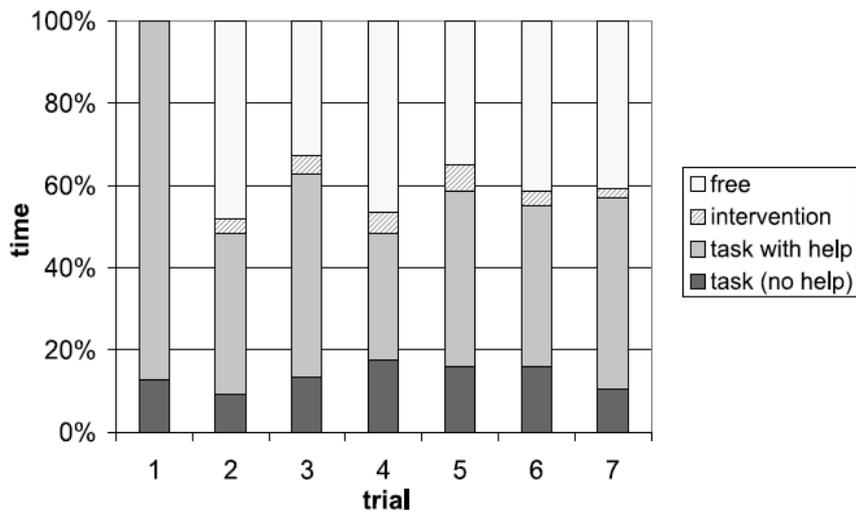
BASELINE DATA		
Agent	Primitive task	Mean time (sec)
EVA astronauts	move between panels	3.9
	weld seam	16.5
	retrieve panel	13.5
	assemble panel	7.7
K10	inspect seam	11.3
	move between panels	12.6
Robonaut	weld seam	66.1
	move between panels	27.4

Untuk mendukung studi ini, Fong melakukan sederetan tes yang meliputi 7 (tujuh) pengujian, penempatan 6 (enam) panel, pengelasan, dan pemeriksaan yang dilakukan oleh suatu tim yang terdiri dari robot dan manusia. Hasil pengujian yang diperoleh cukup memuaskan, dimana pengujian 2 hingga 7 berhasil mencapai target. Gambar 2.3 menunjukkan deretan waktu dari pengujian 6, yang diklasifikasikan menjadi 4 (empat) kategori aktifitas yang meliputi, waktu luang, waktu pelaksanaan kerja, waktu pelaksanaan kerja dengan bantuan, dan intervensi.

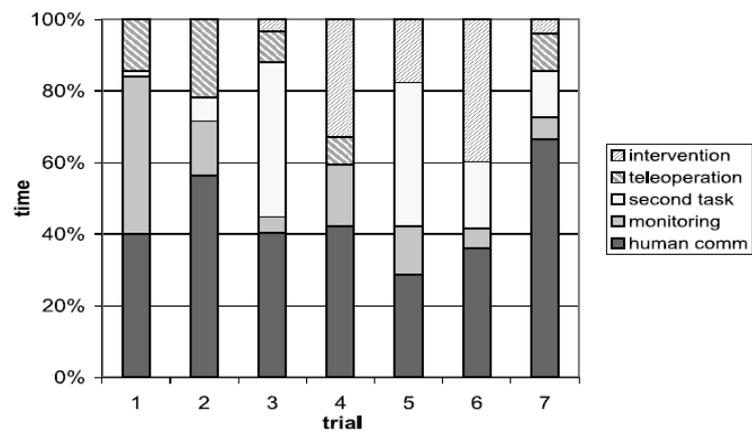


Gambar 2.3 Deretan waktu yang merepresentasikan aktifitas pengelasan (welding) yang dilakukan oleh tim manusia-robot

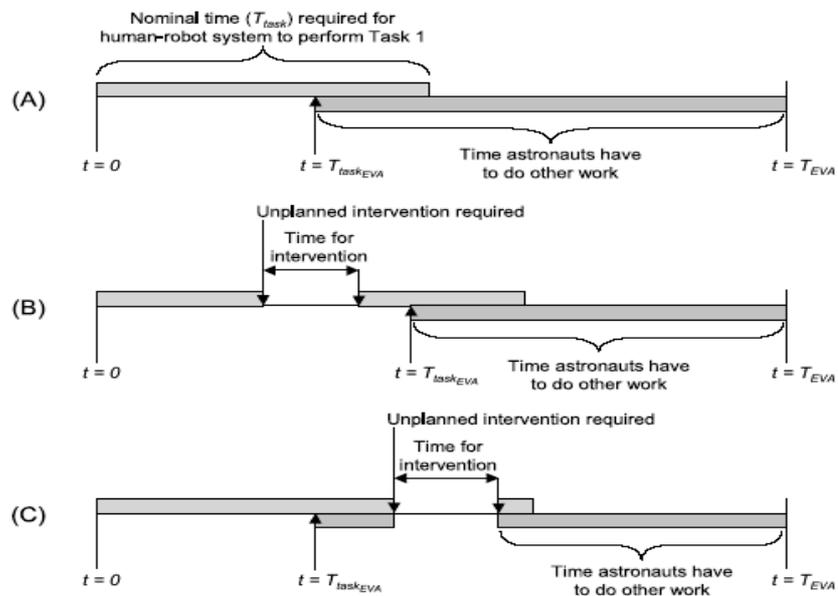
Prosentase waktu yang digunakan untuk melakukan berbagai kategori aktifitas dapat dilihat pada gambar 2.4.



Gambar 2.4 Prosentase waktu yang digunakan untuk melakukan berbagai kategori aktifitas (astronot EVA)



Gambar 2.5 Prosentase waktu yang digunakan untuk melakukan berbagai aktifitas (astronot IVA)



Gambar 2.6 Pengaruh (efek) intervensi pada deret waktu EVA, (A) tidak ada intervensi, (B) intervensi pada saat robot dan manusia sedang melakukan kerja, dan (C) intervensi sesudah manusia melakukan kerja selanjutnya

# Bab 3

## ROBOT INTERAKTIF SECARA SOSIAL

*“Jika merasa kecewa, saya akan bernyanyi. Jika merasa sedih, saya akan tertawa, jika merasakan sakit, saya akan menggandakan kerja”*

(Og. Mandino)

### 3.1 Pengenalan

#### 3.1.1 Sejarah robot sosial

Perkembangan sebuah robot selalu diinspirasi oleh keadaan biologis dari manusia, dan pada saat ini, para ilmuwan mulai menaruh perhatian terhadap berbagai kemungkinan penginteraksian robot dengan robot lainnya maupun penginteraksian robot dengan manusia. Gambar 3.1 memperlihatkan beberapa robot tortoise yang diciptakan oleh Walter pada akhir tahun 1940-an. Maksud dari berbagai lampu yang terpasang di depan tubuh robot yang memancarkan photoaxis positif, adalah sebagai indikasi bahwa robot-robot telah melakukan interaksi dengan cara-cara sosial yang umum, walaupun tanpa komunikasi eksplisit atau rekognisi mutual.



Gambar 3.1 Pelopor robot sosial, tortoise Walter, Elmer, dan Ersie melakukan suatu tarian disekitar robot yang lainnya

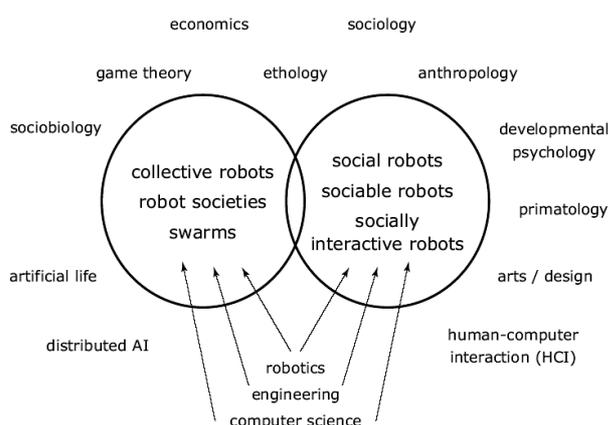
Para peneliti telah memulai mengaplikasikan prinsip-prinsip komunikasi tidak langsung diantara masing-masing individu melalui suatu pemodifikasian yang dirancang pada lingkungan bersama, diistilahkan *stigmergy*, untuk memperoleh berbagai ciri, sifat, atau tingkah laku robot secara kolektif. Stigmergy pertama sekali digunakan oleh Grassé untuk menjelaskan bagaimana komunitas semut sosial mampu menghasilkan berbagai pola tingkah laku dan struktur fisik yang kompleks secara kolektif, walaupun masing-masing individu sering muncul sebagai “pekerja tunggal.”

Deneubourg dan tim kerjanya merintis uji coba pertamanya yang berkaitan dengan stigmergy dengan memanfaatkan simulasi robot semut, pada awal tahun 1990-an. Semenjak itu, muncullah beberapa ilmuwan yang mengembangkan berbagai robot kolektif sebagai suatu model untuk mempelajari tingkah laku serangga secara sosial.

Apakah definisi robot sosial tersebut ?, Dautenhahn dan Billard memberi definisi sebagai berikut,

*“Social robots are embodied agents that are part of a heterogeneous group: a society of robots or humans. They are able to recognize each other and engage in social interactions, they possess histories (perceive and interpret the world in terms of their own experience), and they explicitly communicate with and learn from each other”*

Untuk mengembangkan sebuah robot sosial yang bersifat individual dibutuhkan suatu teknik atau model khusus yang berbeda dengan pengembangan robot sosial yang bersifat kolektif. Faktor-faktor yang kita perhatikan di dalam pengembangan robot tersebut antara lain : peniruan dan pembelajaran secara sosial, komunikasi dengan menggunakan bahasa alami ataupun berupa gerak dan isyarat, emosi, dan rekognisi interaksi.



Gambar 3.2 Berbagai bidang yang memberikan kontribusi terhadap pengembangan robot baik kolektif maupun robot sosial

### 3.1.2 Robot sosial dan penanaman nilai-nilai sosial : konsep dan definisi

Robot pada masyarakat terindividual telah menunjukkan jangkauan sifat dan tingkah laku yang luas, hal ini dapat kita simpulkan dengan mengabaikan keberadaan robot-robot lainnya, manusia, atau keduanya pada lingkungan tersebut. Breazeal mendefinisikan empat kelas dari robot sosial berdasarkan keadaan (1) sebagaimana kemampuan robot dapat mendukung model sosial yang dikaitkan terhadap robot sosial tersebut, dan (2) kompleksitas skenario interaksi yang dapat didukung.

- *Socially evocative*. Robots that rely on the human tendency to anthropomorphize and capitalize on feelings evoked when humans nurture, care, or involved with their “creation”.
- *Social interface*. Robots that provide a “natural” interface by employing human-like social cues and communication modalities. Social behavior is only modeled at the interface, which usually results in shallow models of social cognition.
- *Socially receptive*. Robots that are socially passive but that can benefit from interaction (e.g. learning skills by imitation). Deeper models of human social competencies are required than with social interface robots.
- *Sociable*. Robots that pro-actively engage with humans in order to satisfy internal social aims (drives, emotions, etc). These robots require deep models of social cognition.

Untuk melengkapi keempat poin di atas, di sini dapat kita tambahkan tiga kelas sebagai berikut :

- *Socially situated*. Robots that are surrounded by a social environment that they perceive and react to. Socially situated robots must be able to distinguish between other social agents and various objects in the environment.
- *Socially embedded*. Robots that are: (a) situated in a social environment and interact with other agents and humans; (b) structurally coupled with their social environment; and (c) at least partially aware of human interactional structures (e.g., turntaking).
- *Socially intelligent*. Robots that show aspects of human style social intelligence, based on deep models of human cognition and social competence.

### 3.1.3 Robot interaktif secara sosial

Di sini fokus kita dibatasi hanya pada konsep peer-to-peer, interaksi antara robot dan manusia. Secara spesifik kita mendeskripsikan robot yang mampu melakukan berbagai karakteristik sosial manusia, seperti :

- Ekspresi atau emosi persepsi
- Berkomunikasi menggunakan dialog *high-level*

- Mempelajari atau mengenal berbagai model dari agen-agen yang lain
- Membangun atau memelihara hubungan sosial
- Memanfaatkan petunjuk atau isyarat alam (tatapan, isyarat gerak, dan lain-lain)
- Menunjukkan personalitas dan karakteristik yang berbeda
- Dapat mempelajari atau mengembangkan kompetensi sosial

Robot interaktif secara sosial dapat dimanfaatkan untuk berbagai kegunaan, yang antara lain : sebagai platform riset, permainan, alat bantu pendidikan di sekolah atau universitas, atau juga sebagai pertolongan therapy. Robot interaktif secara sosial beroperasi sebagai teman, pembantu atau asisten, yang berarti berbagai robot harus mampu menunjukkan tingkatan tertentu atau derajat keadaptasian atau fleksibilitas untuk mengendalikan berbagai interaksi pada jangkauan manusia yang lebih tersebar.

#### 3.1.4 Robot interaktif secara sosial, kenapa ?

Robot interaktif sosial sangat dibutuhkan atau penting pada wilayah-wilayah dimana robot harus mampu menunjukkan kemampuannya (skill) untuk menyelesaikan berbagai tugas tertentu, atau rancangan awal robot tersebut sebagai mesin dengan fungsi utamanya untuk berinteraksi secara sosial. B. Fogg, pada publikasi ilmiahnya yang berjudul "*Introduction : persuasive technologies, Communications of the ACM 42*" memaparkan tentang berbagai robot yang mampu mempengaruhi tingkah laku, perasaan, bahkan sikap dari manusia. Untuk contoh kasus yang lain dimana robot berperan sebagai mediator terhadap proses interaksi dari manusia dengan manusia, sebagai *autism therapy*, dipaparkan di dalam karya ilmiahnya I. Werry, dan kawan-kawan yang berjudul "*Can social interaction skills be taught by a social agent? The role of a robotic mediator in autism therapy*" (*Proceedings of the Fourth International Conference on Cognitive Technology, 2001*). E. Paulos dan J. Canny, dalam "*Designing personal teleembodiment, Autonomous Robots,*" (2001), memperkenalkan sebuah robot sebagai avatar, dimana fungsi robot sebagai representasi dari, atau representatif untuk, manusia, sebagai contoh jika sebuah robot digunakan di dalam komunikasi jarak jauh, maka robot akan membutuhkan sikap-sikap (*act*) sosial untuk menyampaikan berbagai informasi secara efektif.

Untuk berbagai skenario saat ini, maka perancangan ditekankan pada pengembangan robot yang mampu meningkatkan kemampuan penginteraksiannya setiap saat secara efektif dengan manusia, agar tercapainya suatu sasaran yang diharapkan dengan tepat. Sebagai contoh, robot *pet* yang menemani seorang anak kecil dengan dibekali sifat-sifat kekanak-kanakannya (robot) yang harus selalu ditingkatkan (sifat kekanak-kanakannya) agar anak yang dijaga dapat selalu tertarik (berinteraksi) terhadap robot tersebut. Pengembangan

pembelajaran kemampuan-kemampuan sosial inilah yang menjadi fokus atau perhatian utama untuk robot epigenetik.

Beberapa peneliti mengadopsi dan menyederhanakan berbagai model tingkah laku sosial dalam mewujudkan perancangan robot sosial yang interaktif. Dan untuk kegunaan tersebut, maka pengembangan robot-robot yang mempunyai berbagai karakteristik, ide, pikiran, skill, dan tujuan secara sosial, mampu menunjukkan empati dan memberikan pengertian yang benar telah menjadi suatu prototipe baku bagi perancangannya. Berbagai disiplin keilmuan seperti kehidupan buatan, psikologi, dan lain-lain telah sangat mempengaruhi pengembangan robot sosial.

Walaupun robot-robot interaktif secara sosial telah sukses dikembangkan dan banyak digunakan dalam berbagai aplikasi, permasalahan yang terkait dengan keefektifitasannya masih tetap menjadi suatu hal yang terus dikaji (ditingkatkan).

## 3.2 Metode

### 3.2.1 Pendekatan desain

Manusia selalu dijadikan pedoman di dalam suatu interaksi, oleh karena itu, jika teknologi mengikuti pedoman interaksi sosial tersebut, maka akan dihasilkan suatu interaksi yang dapat dinikmati, menyenangkan, dan hadirnya nilai-nilai perasaan yang mampu memaknai interaksi tersebut. Dikarenakan hal-hal atau maksud-maksud yang telah disebutkan di atas, maka banyak peneliti melakukan eksplorasi pada wilayah perancangan untuk robot-robot "anthropomorphic" atau "zoomorphic." Untuk berbagai alasan tersebut, maka idealnya sebuah robot wajib atau harus memiliki *faces*, *speech recognition*, *lipreading skills*, dan berbagai fitur-fitur dengan kapasitas tertentu yang dapat mendukung terjadinya suatu interaksi sosial yang menyerupai interaksi sosial antara manusia dengan manusia layaknya.

Berdasarkan perspektif perancangan, kita dapat mengklasifikasikan bagaimanakah robot interaktif sosial dibentuk dengan dua cara utama. Dengan pendekatan pertama, "*biologically inspired*", perancang mencoba untuk menciptakan sebuah robot yang dapat mensimulasikan secara internal, berbagai bentuk kecerdasan sosial yang ditemukan pada makhluk hidup (manusia). Pendekatan kedua, "*functionally designed*", tujuannya adalah untuk membangun sebuah robot yang seakan-akan muncul sebagai sesuatu yang dapat dikatakan cerdas secara sosial, walaupun perancangan internalnya tidak memiliki basis pada sains.

Bagaimanapun robot memiliki keterbatasan perseptual, kognitif, dan berbagai kemampuan sifat, dibandingkan dengan apa-apa yang ada pada manusia. Oleh karena itu, ketidakseimbangan interaksi diantara robot dengan manusia akan selalu terjadi pada masa-masa mendatang. Pada akhir dari sub bab

materi ini, kita dapat simpulkan bahwa perbedaan metode perancangan mengindikasikan perbedaan kriteria evaluasi dan keberhasilan untuk berbagai robot yang berbeda pula.

### 3.2.1.1 Biologically inspired

*The “biologically inspired” approach, designers try to create robots that internally simulate, or mimic, the social behavior or intelligence found in living creatures.*

Perancangan “*biologically inspired*” didasari oleh berbagai teori yang diambil dari sains sosial dan alam, termasuk anthropology, sains kognitif, psikologi perkembangan, ethology, sosiologi, struktur interaksi, dan teori pikiran. Teori *speaking*, teori ini digunakan untuk mengarahkan perancangan robot kognitif, tingkah laku, motivasional (pengendalian dan emosi), sistem motor dan sistem perseptual.

Ada dua pendapat utama untuk memaparkan inspirasi yang diperoleh dari sistem biologis tersebut. Pertama, beberapa peneliti menyatakan bahwa alamiah merupakan model yang paling baik untuk berbagai aktivitas yang menyerupai kehidupan sebenarnya. Secara spesifik, mereka menghipotesakan bahwa agar robot dapat dipahami oleh manusia, maka robot harus memiliki perwujudan kealamiah, mampu berinteraksi sebagaimana manusia atau makhluk hidup berinteraksi dengan sesamanya atau alam sebagai lingkungannya, dan juga mampu mempersepsikan sesuatu benda seperti idealnya cara-cara yang dipersepsikan manusia atau makhluk hidup terhadap benda yang sama. Kedua, rasionalitas dari inspirasi biologis telah memungkinkan kita untuk secara langsung menguji, dan menseleksi berbagai teori sains yang menjadi pondasi perancangan.

Beberapa teori yang sering digunakan di dalam perancangan *biologically inspired* adalah :

- ☞ *Ethology*. Ethology refers to the observational study of animals in their natural setting. Ethology can serve as a basis for design because it describes the types of activity (comfort-seeking, play, etc.) a robot needs to exhibit in order to appear life-like. Ethology is also useful for addressing a range of behavioral issues such as concurrency, motivation, and instinct.
- ☞ *Structure of interaction*. Analysis of interactional structure (such as instruction, cooperation, etc.) can help focus design of perception and cognition systems by identifying key interaction patterns. Dautenhahn, Ogden and Quick use explicit representations of interactional structure to design “interaction aware” robots. Dialogue models, such as turn-taking in conversation, can also be used in design as in.
- ☞ *Theory of mind*. Theory of mind refers to those social skills that allow humans to correctly attribute beliefs, goals, perceptions, feelings, and

desires to themselves and others. One of the critical precursors to these skills is joint (or shared) attention: the ability to selectively attend to an object of mutual interest. Joint attention can aid design, by providing guidelines for recognizing and producing social behaviors such as gaze direction, pointing gestures, etc.

- ☞ *Developmental psychology.* Developmental psychology has been cited as effective mechanism for creating robots that engage in natural social exchanges. As an example, the design of Kismet's "synthetic nervous system", particularly the perceptual and behavioral aspects, is heavily inspired by the social development of human infants. Additionally, theories of child cognitive development, such as Vygotsky's "child in society", can offer a framework for constructing robot architecture and social interaction design.

### 3.2.1.2 Rancangan fungsional

*With the "functionally designed" approach, the objective is to design a robot that outwardly appears to be socially intelligent, even if the internal design does not have a basis in science or nature.*

Pendekatan ini mengambil asumsi bahwa jika kita akan menciptakan sebuah agen sosial buatan yang berkesan yang dikendalikan oleh kepercayaan dan hasrat (keinginan yang memuaskan), maka kita tidak harus terpaku (atau selalu mengerti) akan kebutuhan pemahaman bagaimana otak atau pikiran bekerja secara nyata, tetapi kita sudah cukup mendeskripsikan mekanisme bagaimana manusia menghadapi dan memahami makhluk-makhluk cerdas secara sosial di dalam kehidupannya sehari-hari.

Beberapa motivasi untuk perancangan fungsional robot adalah sebagai berikut :

- ☞ Robot bisa saja memiliki kemampuan sosial yang dangkal, dan secara khusus ini dapat kita katakan benar untuk berbagai permasalahan sosial yang membutuhkan penyelesaian dengan memakan waktu yang singkat atau interaksi yang diperlukan untuk melaksanakan suatu tugas memiliki keterbatasan secara kualitas.
- ☞ Robot mungkin saja memiliki keterbatasan dalam hal : pelaksanaan, kemampuan berinteraksi, bahkan dalam pengambilan keputusan yang tepat pada situasi-situasi yang mendesak di dalam lingkungan tugasnya.
- ☞ Keterbatasan dalam hal berekspresi secara sosial, dapat membantu meningkatkan kemampuan dan kegunaan robot.
- ☞ Perancangan buatan dapat mendesak proses penginteraksian. Contoh, kita dapat mengamati, bagaimana, kebanyakan video game dan permainan elektronik yang mampu mengalihkan perhatian kita.

Ada tiga teknik yang biasanya digunakan di dalam perancangan fungsional :

- ☞ Perancangan *Human-Computer Interaction* (HCI), robot-robot dikembangkan dengan berbagai teknik HCI yang meliputi pemodelan kognitif, inkuiri kontekstual, evaluasi heuristik, dan ujicoba user secara empiris.
- ☞ Rekayasa sistem, rekayasa sistem meliputi pengembangan *top-down* suatu fungsional sistem dan berbagai kebutuhan fisik dari sekumpulan tujuan-tujuan awal. Kegunaannya adalah untuk mengorganisasikan sekumpulan informasi dan pengetahuan yang diperoleh untuk memfasilitasi dan mengontrol perencanaan, pengembangan, dan berbagai operasi pada sistem.
- ☞ Perancangan iteratif, perancangan iteratif atau perancangan sekuensial merupakan suatu proses peninjauan ulang suatu perancangan melalui sederetan pengujian dan siklus “perancangan kembali.” Secara khusus dapat kita katakan perancangan iteratif digunakan untuk mengamati berbagai error atau kondisi “*failure*” dari desain atau untuk meningkatkan basis informasi dari suatu evaluasi.

### 3.2.2 Berbagai isu desain

Keseluruhan sistem robot, baik itu kategori interaktif secara sosial ataupun tidak, harus mengamati sejumlah permasalahan perancangan secara umum, dan ini meliputi kognisi (perencanaan, dan pengambilan keputusan), persepsi (navigasi, sensoran lingkungan), aksi (mobilitas, manipulasi), interaksi manusia-robot (antarmuka pemakai, *feedback display*), dan arsitektur (kontrol, elektromekanikal, sistem).

- ☞ *Human-oriented perception.* A socially interactive robot must proficiently perceive and interpret human activity and behavior. This includes detecting and recognizing gestures, monitoring and classifying activity, discerning intent and social cues, and measuring the human’s feedback.
- ☞ *Natural human-robot interaction.* Humans and robots should communicate as peers who know each other well, such as musicians playing a duet. To achieve this, the robot must manifest believable behavior: it must establish appropriate social expectations, it must regulate social interaction (using dialogue and action), and it must follow social convention and norms.
- ☞ *Readable social cues.* A socially interactive robot must send signals to the human in order to: (1) provide feedback of its internal state; and (2) allow human to interact in a facile, transparent manner. Because robots are constructed, they have limited channels and capacity for emotional expression. These include facial expression, body and pointer gesturing, and vocalization (both speech and sound).

- ☞ *Real-time performance.* Socially interactive robots must operate at human interaction rates. Thus, a robot needs to simultaneously exhibit competent behavior, convey attention and intentionality, and handle social interaction.



Gambar 3.3 Sony Aibo ERS-110 (bagian atas) and K-Team Khepera (bagian bawah)

### 3.2.3 Embodiment

Tubuh secara biologis tersusun atau terbentuk menyesuaikan keadaan lingkungan dimana tubuh tersebut menetap. “Embodiment” memainkan peranan yang sangat penting dalam mewujudkan emosi dan kognisi. Dan secara khusus, perancangan aktuator dan sensor sering secara bersama-sama dirangkaikan terhadap bagaimanakah proses syaraf bekerja.

*Embodiment* dapat kita definisikan sebagai “that which establishes a basis for structural coupling by creating the potential for mutual perturbation between system and environment.” Oleh karenanya “embodiment” dirancang untuk hubungan timbal-balik antara sistem dan lingkungannya, dan lebih-lebih bila robot dapat mengacaukan lingkungannya atau dikacaukan oleh lingkungannya, dan hal ini bukan berarti keseluruhan robot harus memiliki tubuh secara fisik.

Semua robot memiliki wujud (*embodied*), tetapi beberapa robot lebih memiliki wujud daripada lainnya. Lihat gambar 3.3, perhatikan perbedaan-perbedaan yang terdapat pada robot Aibo (Sony) dibandingkan dengan robot Khepera (K-Team), Aibo memiliki kira-kira 20 aktuator (joints across mouth, heads, ears, tails, and legs) dan berbagai sensor (touch, sound, vision and proprioceptive), dibandingkan dengan Khepera yang hanya memiliki 2 aktuator (independent wheel control) dan sebuah sensor (an array of infrared proximity sensors). Karena Aibo memiliki lebih banyak channel-channel *perceptory* dan bandwidth pada disposalnya maka kita dapat mengasumsikan secara hampir tepat bahwa Aibo lebih “embodied” dibandingkan Khepera.

### 3.2.3.1 Morfologi

Bentuk dan struktur robot merupakan bagian perancangan yang sangat penting, dikarenakan bentuk dan struktur mendirikan suatu ekspektasi sosial pada lingkungan dimana robot tersebut berinteraksi. Tampilan fisik dapat kita katakan sebagai bias dari suatu interaksi. Robot yang menyerupai seekor anjing akan diperlakukan berbeda daripada sesuatu yang berupa “*anthropomorphic*.” Lebih-lebih, kedekatan struktur atau bentuk dari robot yang dirancang hingga menyerupai bentuk atau struktur dari manusia, maka akan memberikan pengaruh yang besar pada aksesibilitas, desirabilitas, dan keekspresifan pada robot tersebut.

### 3.2.3.2 Pertimbangan desain

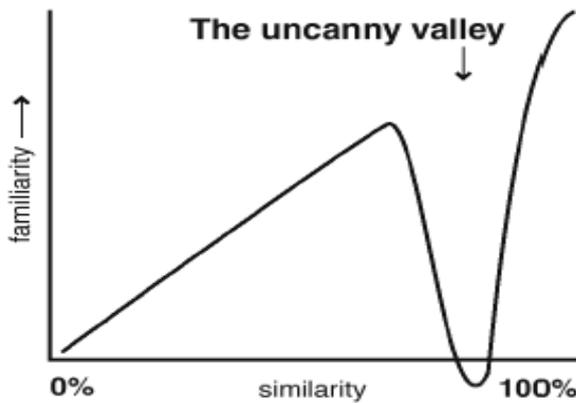
Pada saat industri akan memulai merancang suatu bentuk atau struktur sebuah robot, maka pertimbangan desain harus diperhitungkan terlebih dahulu. Pertama-tama, bentuk atau struktur robot harus sesuai dengan fungsi robot tersebut. Jika robot dibuat untuk kebutuhan melaksanakan berbagai tugas yang akan diperintahkan padanya, maka rancangan bentuknya harus memberi kesan “*product-ness*” agar user merasa nyaman dalam memanfaatkan robot.



Gambar 3.4 “Leonardo,” tampilan robot yang menyerupai makhluk hidup



Gambar 3.5 CERO (dari Severinson-Eklund 2002)



Gambar 3.6 Mori's "Uncanny Valley"

Sebagai penutup, dapat kita katakan bahwa jika sebuah robot ingin berperan seakan-akan sebagai makhluk yang benar nyata dan hidup, maka robot harus memiliki suatu tingkatan pengetahuan yang sesuai (interaksi/adaptasi, *performance*) dengan makhluk hidup yang sebenarnya. Mashiro Mori memaparkan bahwa tingkatan progress perubahan dari makhluk yang "tidak nyata" menjadi makhluk yang mendekati "nyata," adalah tidak linier, secara grafis, terdapat "Uncanny Valley" seperti yang ditunjukkan pada gambar 3.6 - Mori's "Uncanny Valley."

Kita klasifikasikan robot sosial berdasarkan wujudnya menjadi empat kategori : *anthropomorphic*, *zoomorphic*, *caricatured*, dan *functional*.

### 3.2.3.3 Anthropomorphic

Anthropomorphism, berasal dari bahasa Yunani, "anthropos" mempunyai arti "orang," dan "morphe" mempunyai arti "bentuk" atau "struktur."

Anthropomorphism adalah kedekatan sesuatu benda terhadap segala atribut karakteristik dari manusia, yang bertujuan, dengan kedekatan (struktur) tersebut, adalah untuk membantu merasionalkan segala tindakan yang merupakan bagian aplikasi dari proses penginteraksian antara robot dengan partnernya (robot/manusia). Paradigma Anthropomorphism telah berhasil secara luas diterapkan sebagai suatu standar yang tepat untuk memperkaya karakteristik fungsi dan segala tindakan pada robot sosial.

Peran dari antropomorphism adalah sebagai fungsi mekanisme (untuk perancangan, untuk penginterpretasian tingkah-laku, dan lain-lain) melalui suatu penginteraksian sosial yang mungkin difasilitasi. Oleh karenanya, peran ideal dari antropomorphism adalah untuk menghadirkan suatu keseimbangan ilusi dan fungsi yang tepat, atau dengan kata lain meyakinkan manusia bahwa robot dapat melakukan sesuatu yang riil dengan meminimalkan error yang mungkin, supaya tercapainya interaksi manusia-robot seperti yang diharapkan.

#### 3.2.3.4 Zoomorphic

Pertambahan jumlah robot baik itu robot yang bersifat *entertainment*, personal, atau permainan, tiada lain adalah kreasi dari para perancangnya untuk mengimitasi objek kehidupan yang riil. Untuk wujud dari robot-robot tersebut, maka wujud zoomorphic sangat penting dimanfaatkan untuk merealisasikan hubungan manusia-robot yang lebih interaktif.

Perancangan yang umum hampir selalu diinspirasi oleh keadaan hewan-hewan peliharaan yang bersifat rumahan, seperti yang telah diprakarsai oleh Sony dengan Aibo (gambar 3.3), robot Necoro (Omron) (Gambar 3.7) dengan tujuan penciptaanya sebagai partner rumahan, Leornado (gambar 3.4), dan lian-lain.



a



b

Gambar 3.7 (a) I-Cybie (Tiger Electronics), (b) Necoro (Omron)

### 3.2.3.5 Karikatur

Para animator telah lama mendemonstrasikan bahwa suatu karakter tidak selalu harus muncul secara nyata untuk dapat dipercayai oleh yang melihatnya. Karikatur, untuk contoh biasanya diungkapkan dengan melebih-lebihkan fitur-fitur objek yang sebenarnya dengan tujuan tiada lain hanya sebagai efek komik semata.

Scheef, dan kawan-kawan, telah mendiskusikan bagaimana teknik-teknik dari animasi dan *cartooning* tradisional dapat digunakan untuk perancangan robot interaktif sosial. Schulte, juga memaparkan hal yang berkaitan, bagaimana wajah manusia yang telah dibentuk dalam sebuah karikatur dapat memberikan suatu "focal-point" yang mampu mengambil alih perhatian manusia. Severinson-Eklund, 2002, memaparkan kegunaan dari sebuah karakter mekanis yang berukuran kecil, "CERO", sebagai sebuah robot yang representatif dan interaktif (gambar 3.5).

### 3.2.3.6 Fungsional

Sebagian besar peneliti berargumentasi bahwa perancangan wujud robot adalah langkah awal yang sangat penting sebelum dilakukan pembentukan struktur keseluruhan dari robot secara utuh, dan wujud robot harus merefleksikan fungsi atau tugas dari robot yang akan dirancang. Pilihan dan perancangan fitur-fitur robot secara murni harus berpedoman pada tujuan-tujuan operasional yang telah ditentukan.

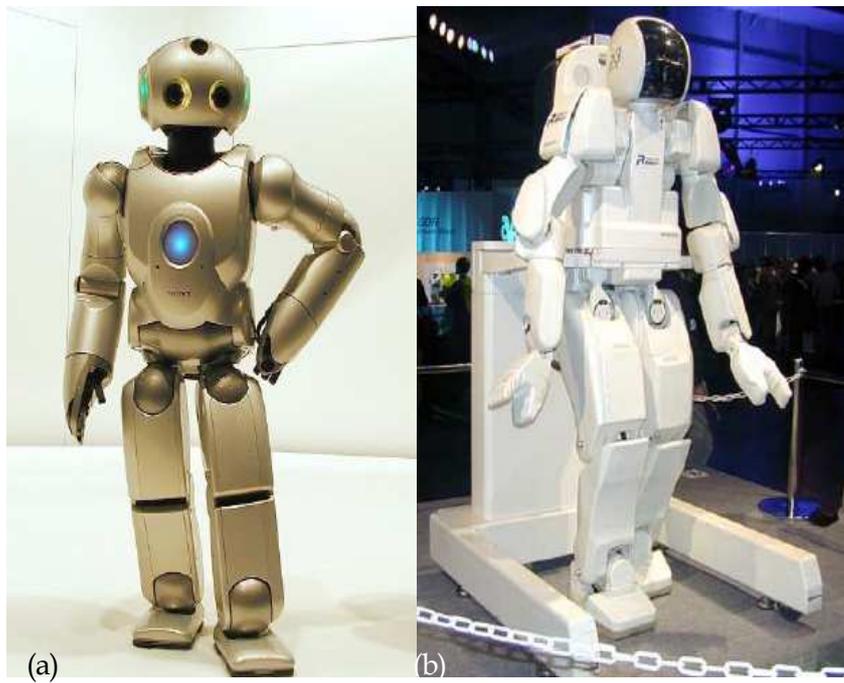
Robot yang dirancang untuk "Health Care", misalkan, diperlukan untuk membantu para pasien yang sudah manula, atau tidak mampu berjalan, memindahkan para pasien, dan lain-lain. Dan untuk itu, fitur-fitur lain yang ada di luar robot sangat diperlukan, seperti "ruang cargo" dan "handle bar."



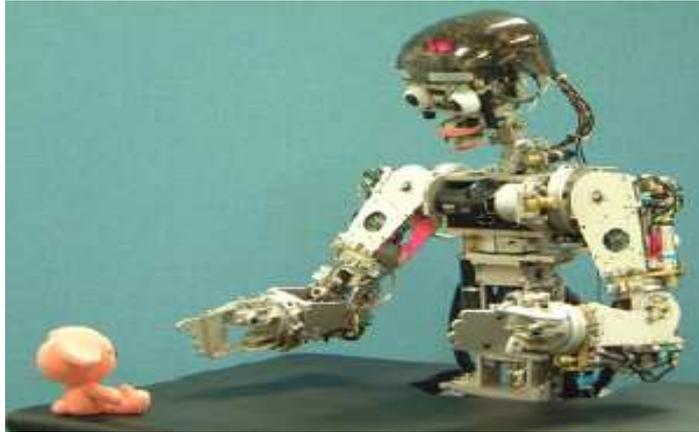
Gambar 3.8 Me and My Shadow  
(MGA Entertainment)



Gambar 3.8 Aibo ERS-200 (top) and ERS-300 series (Sony)



Gambar 3.10 (a) SDR - 4X (Sony), (b) P3 (Honda)



Gambar 3.11 Infanoid

### 3.2.4 Emosi

Emosi memainkan peranan penting di dalam tingkah laku, sifat, dan aktifitas manusia, seperti komunikasi dan interaksi. Emosi sangat mempengaruhi proses kognitif manusia seperti dalam pengambilan keputusan, dan penyelesaian masalah, dan merupakan fenomena yang sangat kompleks yang secara erat berkaitan di dalam segala aktivitas sosial, dan lebih-lebih emosi secara psikologis dipengaruhi oleh keadaan fisiologis atau dapat dikatakan tergantung oleh rupa, wujud, atau struktur dari fisik manusia.

Ada tiga teori utama yang digunakan untuk mendeskripsikan emosi, pendekatan pertama memaparkan emosi dalam konteks kategori diskrit, misalkan, kebahagiaan. Pendekatan kedua mengkarakteristikan menggunakan skala kontinu atau dimensi basis, seperti, *arousal* dan *valence*. Pendekatan ketiga, teori komponential yang mengkombinasikan dimensi dan kategori diskrit.

#### 3.2.4.1 Emosi buatan (Artificial Emotions)

Emosi buatan diintegrasikan ke dalam perancangan robot sosial dikarenakan berbagai alasan atau sebab. Fungsi utama dari emosi adalah untuk memfasilitasi interaksi manusia-robot yang dapat dipercaya, dan memberikan timbal balik kepada user, seperti pengindikasian state internal robot, "maksud dan tujuan." Emosi buatan dapat berperan sebagai sebuah mekanisme kontrol, mengendalikan perbuatan atau gerak tingkah laku, dan merefleksikan bagaimana robot dipengaruhi oleh, beradaptasi terhadap, berbagai faktor yang variatif (berbeda) setiap saat. Dua pandangan tentang pengontrolan emosi yang populer dipaparkan oleh L. Cañamero dan P. Petta, eds., *Grounding Emotions in Adaptive Systems*. Volumes I and II. Special Issue of *Cybernetics and Systems: An*

International Journal, 32 (5) and 32(6) (2001), dan L. Cañamero, *Designing emotions for activity selection in autonomous agents*, in R. Trapp, P. Petta, S. Payr, eds., *Emotions in Humans and Artifacts*, Cambridge, MIT Press, 2003.

Berbagai arsitektur telah diproposalkan untuk pemodelan (perancangan) emosi buatan, antara lain : C. Breazeal, *Designing sociable robots*, Cambridge, MIT Press, 2002, L. Cañamero, *Modeling motivations and emotions as a basis for Intelligent behavior*, in: W. Johnson, ed., *Proceedings of the International Conference on Autonomous Agents*, 1997, C. Elliot, *The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System*. Ph.D. thesis, The Institute for the Learning Sciences Technical Report No. 32, Northwestern University, 1992, E. Hudlicka, *Increasing SIA architecture realism by modeling and adapting to affect and personality*. In K. Dautenhahn et al., eds., *Socially Intelligent Agents: Creating Relationships with Computers and Robots*, Kluwer, 2002, W. Reilly, *Believable social and emotional agents*, Ph.D. Thesis, Computer Science, Carnegie Mellon University, 1996, J. Velasquez, *Modeling emotions and other motivations in synthetic agents*, in: *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, 1997, J. Velasquez, *A computational framework for emotion-based control*, in: *Proceedings of the workshop on grounding emotions in adaptive systems*, Conference on Simulation of Adaptive Behavior, 1998. Beberapa arsitektur yang diajukan menyandarkan pada berbagai teori emosional, atau secara khusus, dapat kita katakan teori tentang bagaimanakah emosi didefinisikan dan dihasilkan. Dan arsitektur yang lain dirancang dengan suatu cara *ad-hoc*, untuk pendapat terakhir didukung oleh Nourbakhsh dan kawan-kawan, yang merincikan suatu sistem berbasis mesin state fuzzy, yang dibangun oleh sederetan evaluasi formatif dan siklus perancangan.

*In terms of expression, some robots are only capable of displaying emotion in a limited way, such as individually actuated lips or flashing lights (usually LEDs). Other robots have many active degrees of freedom and can thus provide richer movement and gestures. Kismet, for example, has controllable eyebrows, ears, eyeballs, eyelids, a mouth with two lips and a pan/tilt neck.*

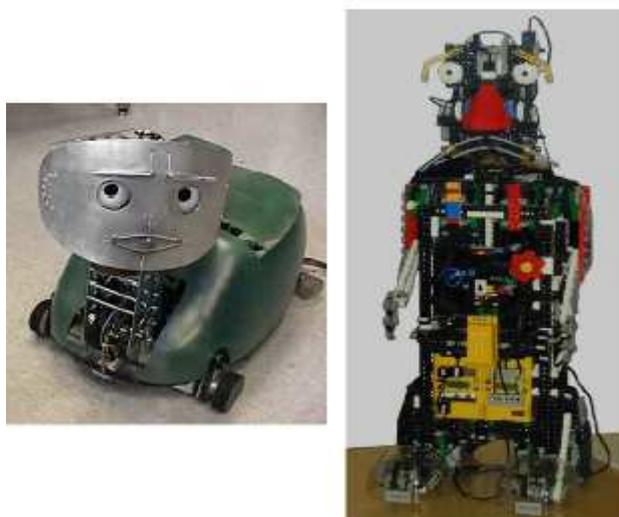
#### 3.2.4.2 Emosi sebagai mekanisme terkontrol

Emosi dapat digunakan sebagai *presedens* kontrol diantara berbagai mode tingkah laku yang berbeda, mengkoordinasikan perencanaan, dan sebagai pemacu pembelajaran dan adaptasi terhadap lingkungan tugasnya, lebih-lebih jika lingkungan yang dihadapinya sulit untuk diprediksi. Untuk itu berbagai pendekatan dirumuskan dengan mengadopsikan kondisi atau keadaan emosi dari berbagai hewan yang menggunakan instingnya dalam menghadapi lingkungannya, misalkan insting pada hewan dalam melarikan diri dari keadaan yang berbahaya, mencari makanan, dan lain-lain.

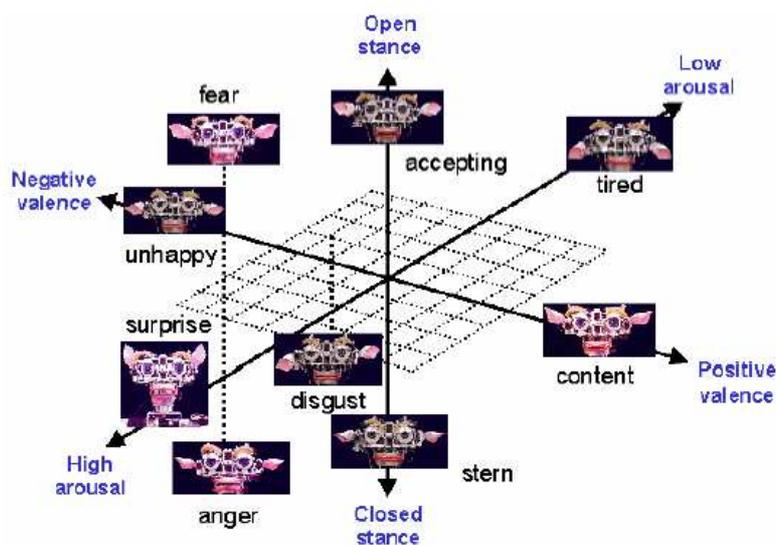
Beberapa peneliti telah melakukan investigasi terhadap manfaat emosi di dalam interaksi manusia-robot. Suzuki telah mendeskripsikan suatu arsitektur yang memungkinkan interaksi memberi dampak perubahan terhadap "state" emosi dari robot dan pemodifikasian respon gerak dari robot tersebut. Breazeal juga mendiskusikan bagaimana emosi dapat mempengaruhi operasi *Kismet's motivational system* dan bagaimana pengaruhnya terhadap proses penginteraksian manusia-robot. Dan Nourbakhsh mendiskusikan bagaimana perubahan "mood" dapat memacu perubahan sikap yang berbeda pada Sage (a museum tour robot).

### 3.2.4.3 Speech

Speech merupakan metode yang sangat efektif untuk mengkomunikasikan emosi. Berbagai parameter utama yang mengendalikan muatan emosi antara lain : "loudness, pitch (level, variation, range), dan prosody." Murray dan Arnott berpendapat bahwa efek vokal atau efek dari suara disebabkan oleh emosi khusus yang tetap konsisten selama berlangsungnya komunikasi diantara berbagai pembicara, dan dengan minimnya perbedaan yang terjadi diantaranya .



Gambar 3.12 Actuated faces: Sparky (left) and Felix (right)



Gambar 3.13 berbagai variasi emosi yang ditunjukkan oleh Kismet

Kualitas pensintesaan speech secara signifikan lebih rendah bila dibandingkan pensintesaan ekspresi face ataupun pensintesaan bahasa tubuh. Cahn dan kawan-kawan telah mendeskripsikan suatu sistem untuk pemetaan kualitas emosional menjadi suatu setting pensintesaan speech, dan termasuk juga artikulasi, pitch, dan kualitas suara.

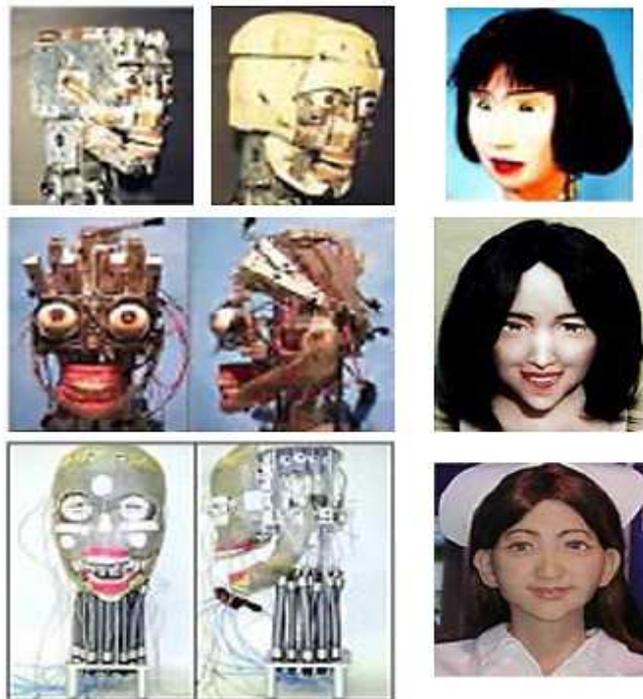
#### 3.2.4.4 Ekspresi face

Face manusia memiliki banyak fungsi, dia dapat menampilkan suatu motivasi, yang akan sangat membantu membuat segala tingkah laku atau perbuatan mudah diprediksikan dan difahami oleh yang lain. Selain itu juga berfungsi sebagai suplemen komunikasi verbal dengan memberikan suatu sinyal sikap pembicara terhadap informasi yang dibicarakan. Ekspresi dan isyarat face juga mengkomunikasikan informasi, seperti sikap kita dengan mengangkat bahu yang mewakili kata-kata "I Don't Know," dan lain-lain.

Ekspresi face pada robot, tidak mungkin sekompleks ekspresi yang terjadi pada manusia, hal ini disebabkan oleh keterbatasan perancangan mekatronik dan perancangan sistem kontrolnya. Untuk contoh, pada perancangan sistem yang sudah dilakukan, ekspresi marah pada robot sangat bertolak belakang dengan ekspresi marah pada makhluk hidup, pada robot ekspresi dan perubahan ekspresi selalu terjadi cepat dan tiba-tiba, dan pada manusia (contoh kasus) berbagai fitur face yang mendukung ekspresi secara frekuensi adalah : *Mouth*

(lips), cheeks, eyes, eyebrows dan forehead. Kebanyakan sistem ekspresi “emosi face” robot merujuk kepada Ekman dan Frieser’s FACS system.

Lihat gambar 3.12, dua face robot yang sangat sederhana yang muncul pada Sparky dan Felix. Face sparky mempunyai 4-DOF, (eyebrows, eyelids, and lips), yang melukiskan suatu himpunan diskrit, emosi dasar. Felix merupakan robot yang dibuat dengan menggunakan LEGO Mindstorms™ robotic construction kit. Seperti Sparky, face Felix juga memiliki 4-DOF, (two eyebrows, two lips), dirancang untuk menampilkan enam keadaan ekspresi, anger, sadness, fear, happiness, surprise, neutral, dan perpaduan keenam emosi tersebut.



Gambar 3.14 Face robot “Saya” (Science University of Tokyo)

Face Kismet memiliki 15 aktuator, dan sering secara serempak bekerja untuk menampilkan suatu ekspresi emosi yang spesifik, (lihat gambar 3.13). sedangkan ekspresi face Kismet dihasilkan menggunakan suatu teknik berbasis interpolasi pada komponensial 3-D “affect space”, (arousal, valence, and stance). Kismet mampu menampilkan ekspresi yang merupakan pemetaan dari keadaan : marah, ketidakpercayaan, takut/kuatir, kebahagiaan, kesedihan, dan terkejut.

Face robot yang lebih reliistik atau mendekati face alami adalah apa yang dirancang di Universitas Sains Tokyo, (gambar 3.14), saya. Secara eksplisit, berbagai face tersebut dirancang untuk menampilkan face robot yang mendekati

face realistik atau susunan face manusia, adanya gigi, rambut, dan lapisan kulit silikon. Dibawah kulit, sejumlah sistem kontrol dicadangkan untuk melakukan berbagai ekspresi, dan pergerakan face layaknya seorang manusia.

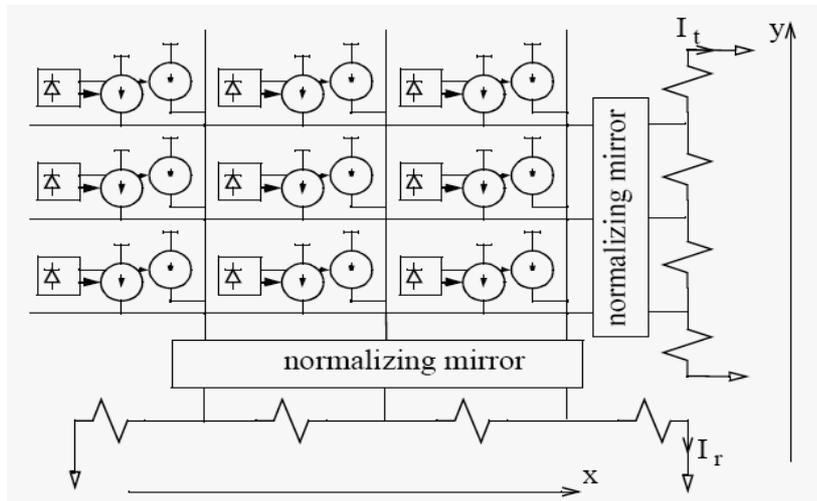
# Bab 4

## SENSOR KOMPUTASI UNTUK “TRACKING” OBJEK

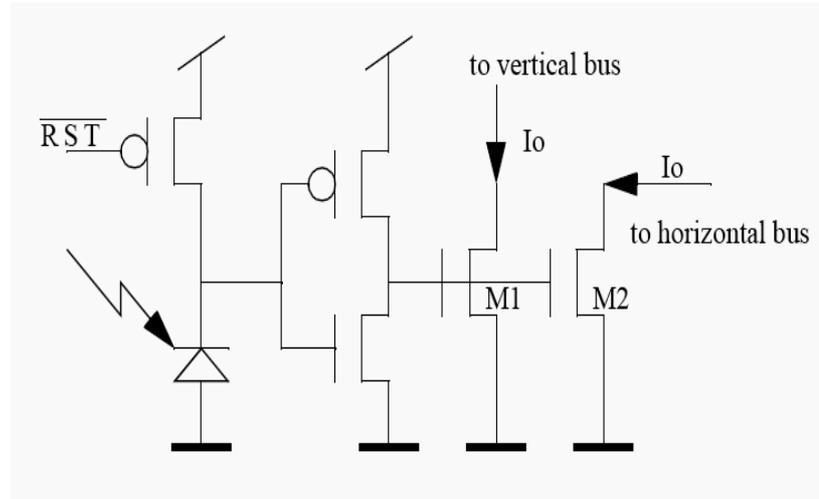
*“An analog VLSI computational vision sensor detects an optical image of multiple intensity-coded objects. As the image is being sensed, the sensor computes and reports the area and the  $x,y$  location of the object on three analog pins. A  $43 \times 43$  cell prototype is implemented in  $2\mu$  CMOS technology. Each cell/pixel is  $47 \times 47 \mu\text{m}$  with a 30% fill factor.”*

### 4.1 Pendahuluan

Pada bab ini kita akan mendeskripsikan sebuah sensor visi komputasional VLSI analog untuk kegunaan perhitungan daerah dan sentroid dari berbagai objek intensitas terkode. Daerah dan lokasi  $(x, y)$  dari objek dibaca secara langsung pada tiga pin analog, sebagai suatu citra hasil sensoran, operasi *low-latency* melakukan penyesuaian untuk aplikasi tracking seperti tracking untuk objek mata atau kepala, dan sebuah prototipe sel  $43 \times 43$  diimplementasikan pada teknologi CMOS  $2\mu$ , dimana masing-masing sel/pixel berukuran  $47 \mu\text{m} \times 47 \mu\text{m}$  dengan *fill-factor* sebesar 30%.

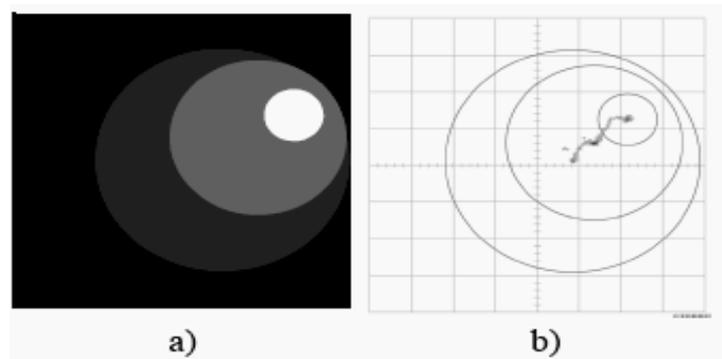


Gambar 4.1 Blok diagram sensor



Gambar 4.2 Diagram sel/pixel

Secara tradisional, *position sensitive detectors* (PSD) menghitung atau mengkalkulasi posisi sebuah objek sebagai sentroid dari suatu distribusi intensitas spatial objek tersebut. Untuk pensegmentasian, pendekatan-pendekatan konservatif lebih menekankan kinerjanya hanya pada citra yang memiliki objek dengan latar yang bertolak-belakang tingkat keabuannya dengan tingkat keabuan objek yang akan ditracking, dan untuk citra-citra yang memiliki susunan objek dan background yang lebih kompleks, maka diperlukan suatu pendekatan pensegmentasian yang lebih handal. Untuk citra yang memiliki susunan warna yang lebih kompleks, Standley memperkenalkan suatu metode pensegmentasian objek dari backgroundnya dengan mengurangi nilai-nilai treshold yang didefinisikan berdasarkan suatu keadaan citra tertentu dari sebuah *photocurrent* pada masing-masing pixel dalam citra yang akan disegmentasikan.



Gambar 4.3 (a) Pola yang diujicobakan, (b) evolusi sentroid

Sebagai perangkat yang harus beroperasi dalam suatu region subthreshold, maka nilai treshold yang terdefinisikan akan menelusuri dan membandingkan berbagai nilai pixel untuk setiap pixel pada citra., dan hal ini bisa menyebabkan terjadinya keadaan “corrupt” pada citra.

Kehandalan dari chip Standley adalah kemampuan chip tersebut memfasilitasi proses pensegmentasian hingga pada berbagai citra yang berukuran besar dan diliputi oleh mayoritas noise. Pada prinsipnya, chip Standley mampu mendeteksi secara kumulatif berbagai objek intensitas terkode dengan memangkas tingkat tresholdnya.

## 4.2 Implementasi

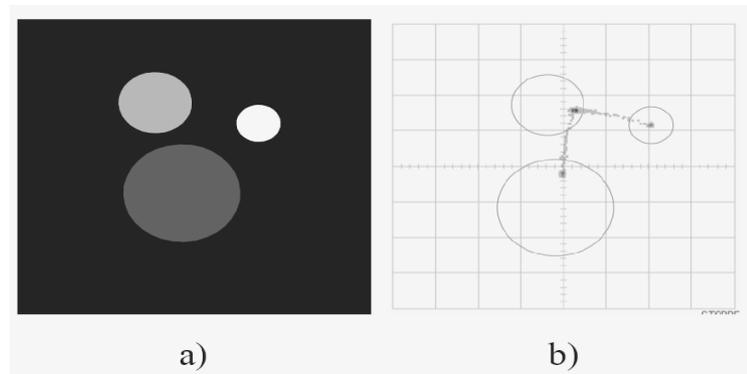
Citra optis yang diperoleh dari pemandangan diluar alat, disensor secara langsung pada chip, dan selanjutnya chip secara kontinu mensegmentasikan objek dari backgroundnya menggunakan paradigma pengolahan “intensity-to-time” dan menyimpan hasilnya dalam bentuk citra biner. Pada awalnya keseluruhan nilai-nilai pixel pada citra biner diset menjadi nol, dan sesudah itu tak lama kemudian, chip melakukan proses selanjutnya hingga terbentuknya sebuah objek yang tidak nol dalam citra biner. Dari citra biner yang telah terbentuk terakhir, chip secara kontinu mengkalkulasikan daerah dan sentroid dari penyebaran berbagai pixel yang tidak nol, menggunakan metode pemroyeksian dan “two 1D resistive grids.”

Blok diagram dari chip yang dirancang dapat dilihat pada gambar 4.1, dan sirkuit pixelnya ditunjukkan pada gambar 4.2. pada awal proses, seluruh sel diset bernilai nol, dan selanjutnya diagram sel memulai prosesnya mengintegrasikan “photo-charge.” Setelah keseluruhan “photo-charge” dikumpulkan di dalam elemen pensensoran, proses berikutnya inverter berkerja menghidupkan dua konstanta sumber M1 dan M2, yang mana kedua konstanta tersebut merepresentasikan keadaan bit-bit biner dari citra. Masing-masing dari kedua konstanta tersebut dijumlahkan menjadi bus vertikal dan bus horisontal. Proses ini merupakan suatu metode pemroyeksian yang mendekomposisikan permasalahan 2D menjadi dua buah distribusi 1D. Jumlah total saat ini  $I_T$  pada bus vertikal atau horisontal (salah satunya) merepresentasikan jumlah total sel-sel yang diswitch dalam keadaan *on*. Sentroid distribusi horisontal dan vertikal merepresentasikan sentroid sebuah objek (x, y). Nilai kuantitas yang dapat kita tentukan sebagai,

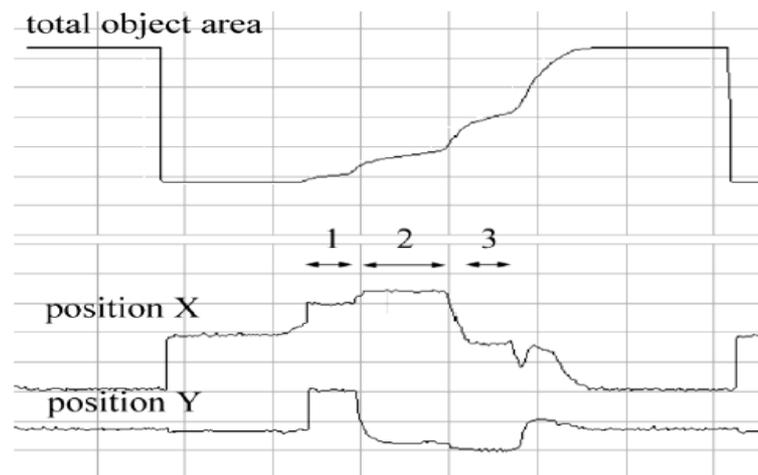
$$\bar{x} = (N - 1) \frac{I_x}{I_T} + 1$$

$$\bar{y} = (M - 1) \frac{I_y}{I_T} + 1$$

Uji coba berbagai pola yang dilakukan adalah dengan memilih dua sampel pola disk, *overlapping* dan *non-overlapping*, dengan berbagai intensitas yang variatif (lihat gambar 4.3 dan gambar 4.4).



Gambar 4.4 (a) Pola yang diujicobakan, (b) evolusi sentroid



Gambar 4.5 Berbagai bentuk gelombang yang dihasilkan untuk pengujian berbagai pola pada gambar 4.3 dan gambar 4.4

Gambar 4.5 menunjukkan bentuk gelombang untuk ketiga sampel disk, yang dimonitor menggunakan sebuah scope digital.

# Bab 5

## ROBOT: PERSEPSI

*“The weakest link in many mobile robots is perception. In order to build robots that are reliable and dependable and safe, we need to build robots that can see. Perception is becoming a solved problem for certain constrained environments. But for robots working outdoors, and at high speeds, and in close proximity to people, perception is still incomplete. Our robots need to see objects; to detect motion; and to detect which of those objects are people. In the current state of the art, this requires multiple sensors and multiple means of interpretation”*

### 5.1 Pendahuluan

Ada beberapa alasan sebuah mobile-robot tidak dapat kita handalkan : error pada mekanik, kegagalan aktuator, kegagalan pada perencanaan, kerusakan sistem komputer, dan lain-lain. Kebanyakan kekurangan-kekurangan tersebut terjadi pada lingkungan : pengoperasian AGV di pabrik, robot pengantar surat di kantor dan robot pengantar obat di rumah sakit, sistem ConneXXion yang mengantar jemput orang dari tempat pemberhentian bus ke kantor, robot eksperimen yang bertugas di museum dan stasiun kereta api. Tetapi saat ini, mesin-mesin robot tersebut dapat lebih terjamin dan tidak perlu lagi dikuatirkan terhadap error atau kegagalan yang biasa terjadi, peluang untuk terjadi error atau kegagalan tersebut semakin kecil. Tetapi untuk sistem ConneXXion, orang-orang dapat memilih, kendaraan tanpa supir atau bus konvensional dengan kapasitas terbatas yang hanya mampu menjemput penumpang yang mana saja yang pertama tiba di perhentian bus, permasalahan disini lebih menekankan istilah “pengadaptasian” saja. Permasalahan terakhir tersebut, untuk sistem ConneXXion, menjadi kabar yang tidak baik bagi perusahaan yang bergerak di bidang perancangan robot untuk aplikasi kendaraan atraktif, tetapi kabar yang baik untuk para pengkaji atau peneliti di bidang ilmu robot.

Ada dua karakteristik yang harus diintegrasikan, agar aplikasi di atas dapat berlangsung sukses seperti yang diharapkan :

1. Lingkungan harus terkondisikan, tujuannya agar robot dapat dengan mudah mendeteksi segala rintangan-rintangan yang muncul setiap saat, dan

2. Robot harus bergerak dengan pelan, dengan gerak robot yang tidak begitu cepat, maka akan memudahkan robot untuk membuat keputusan yang sederhana ketika rintangan-rintangan tersebut muncul. Pada prinsipnya rintangan dapat dengan mudah dideteksi, bila, rintangan tersebut juga mudah disensor dan disegmentasikan oleh robot.

Kajian kita dalam bab ini adalah persepsi. Pada riset yang lebih luas, wilayah kerja lingkungan yang diteliti akan lebih luas juga atau tidak terbatas pada lingkungan-lingkungan yang sudah paten (rintangan-rintangan yang mungkin munculpun sudah dikondisikan, pencahayaan tidak begitu variatif, kondisi cuaca yang tetap, dan lain-lain).

	Video	Stereo Video	Omni Camera	Optical Flow	Motion Detector	Proximity Sensor	Sonar	Short Range Radar	Long Range Radar	Laser Line Striper	Laser Scanner	Lane Tracker
Detect parked object in lane ahead												
Recognize lane markings												●
Detect object in front of it	●	●	●	●		●	●	●	●	●	●	●
Classify object as in its lane				●					●	●	●	●
Determine if object is stationary	●	●	●	●			●	●	●	●	●	
Detect a bicycle approaching from right												
Detect object to front and right	●	●	●	●		●	●	●	●	●	●	
Determine velocity / trajectory	●	●	●	●				●	●	●	●	
Find bicycle when occluded	●	●	●	●						●	●	
Detect an object (stationary pedestrian)												
Recognize curb boundary										●	●	
Detect object to the right side	●	●	●	●		●	●	●		●	●	
Classify object on the sidewalk				●						●	●	
Detect an object (moving pedestrian)												
Recognize curb boundary										●	●	
Detect object to the right side	●	●	●	●		●	●	●		●	●	
Determine velocity / trajectory	●	●	●	●				●		●	●	
Classify object as on sidewalk				●						●	●	
Detect oncoming vehicles												
Recognize lane markings												●
Detect object on front and left	●	●	●	●		●	●	●	●	●	●	
Classify object as not in its lane				●					●	●	●	●
Determine that object is moving				●			●	●	●	●	●	
Detect vehicle in same lane behind it												●
Recognize lane markings												●
Detect object to the rear	●	●	●	●		●	●	●	●	●	●	
Classify object as in its lane				●					●	●	●	●
Determine that object is moving with the flow of traffic				●			●	●	●	●	●	
<b>Total number of functions supported by sensor type</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>17</b>	<b>0</b>	<b>6</b>	<b>9</b>	<b>11</b>	<b>11</b>	<b>19</b>	<b>19</b>	<b>6</b>

Gambar 5.1 Fungsi vs berbagai sensor untuk skenario berkendara di kota

## 5.2 Melihat segalanya

Thorpe, dkk, di dalam karya ilmiahnya yang berjudul “Dependable Perceptions for Robots” memaparkan tentang pengembangan suatu sistem praktis untuk kegunaan visi robot, dengan fokus mereka pada keadaan berkendara di kota yang ramai dengan penduduk (padat).

Pertama-tama, Thorpe dan kawan-kawannya mengumpulkan data-data tentang berbagai keadaan kecelakaan di kota, dan melakukan wawancara tanya-jawab dengan berbagai supir bus yang pernah mengalami kecelakaan jalan raya. Hasil dari analisis data yang dikumpulkan tersebut menjadi acuan Thorpe untuk merancang suatu sistem robot dengan berbagai sensor yang dapat mendeteksi segala keadaan di jalan raya untuk menghindari terjadinya hal-hal yang tidak diinginkan.

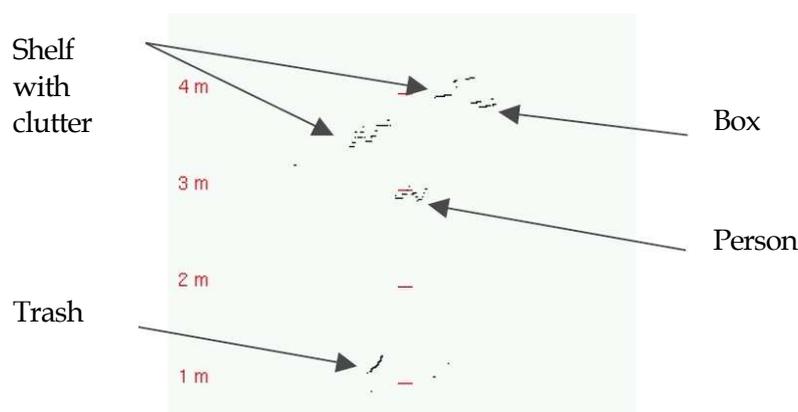
Perbandingan berbagai sensor diperlihatkan pada gambar 5.1, berdasarkan tabel pada gambar tersebut, dapat kita simpulkan bahwa tidak ada sensor tunggal yang dapat melakukan pendeteksian terhadap berbagai keadaan objek yang berada di sekitar bus, maka untuk mengatasi hal-hal keterbatasan sensor tunggal tersebut, kita akan memerlukan radar dan ladar dari vendor yang komersil, dan mengembangkan suatu alat sensor jangkauan *right-stripe* dan, alat pendeteksi objek aliran optis.

### 5.2.1 Sensor jangkauan *right-stripe*

Prinsip kerja sensor jangkauan *right-stripe* telah diketahui sangat baik, cahaya yang berupa suatu laser difokuskan pada satu arah dan diteruskan ke suatu unit untuk menghasilkan suatu bidang cahaya. Kamera video ditempatkan pada suatu interval jarak dari laser dan bertugas mengamati berbagai cahaya yang muncul beririsan dengan objek. Untuk contoh visual dapat dilihat pada gambar 5.2.



Gambar 5.2 Kiri, penglihatan kamera tanpa suppressi background, tanda panah menunjukkan penempatan laser yang dapat dilihat, dan kotak dibelakang diposisikan pada jarak 4m dari kamera. Kanan, penglihatan kamera dengan suppressi background, tanda panah menunjukkan jalur-jalur laser yang beririsan atau bersinggungan dengan kotak



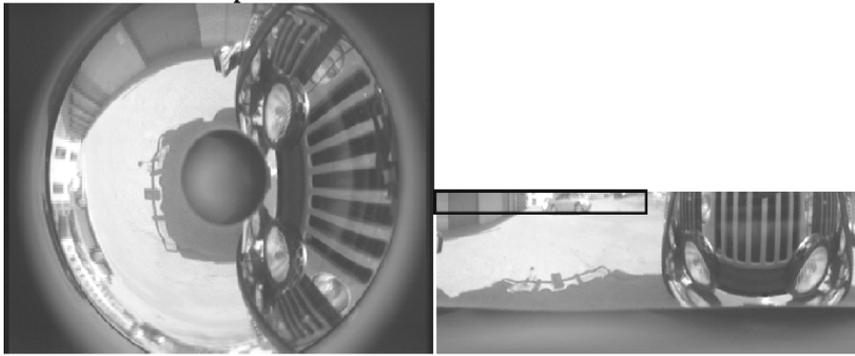
Gambar 5.3 Ekstraksi garis pada koordinat dunia x-y

Ada berbagai permasalahan dengan sensor jangkauan *right-stripe* dan diantaranya adalah sulitnya untuk sensor bekerja secara efektif di siang hari yang begitu cerah, dikarenakan laser yang sulit mendeteksi suatu objek pada citra yang memiliki background yang cerah, dan untuk permasalahan ini Thorpe menggunakan beberapa pendekatan paralel untuk membuat laser lebih mudah melakukan pendeteksian :

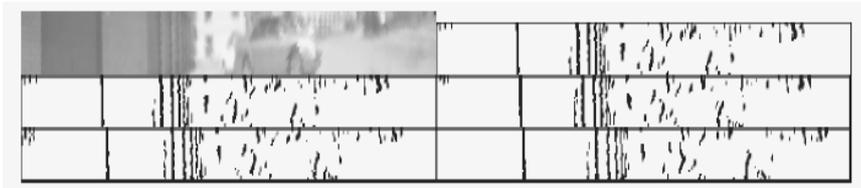
- ☞ *Filtering*: a bandpass filter that closely matches the laser is put over the camera
- ☞ *Shuttering*: the laser is fired in a pulse of a few microseconds, and the camera's electronic shutter is opened for the same interval
- ☞ *Image subtraction*: For a stationary scene, it is easy to collect two images, one with the laser on and one with it off, and subtract images. For a moving scene, implementing this process requires image-splitting optics or image registration.

## 5.2.2 Omnikamera aliran optis

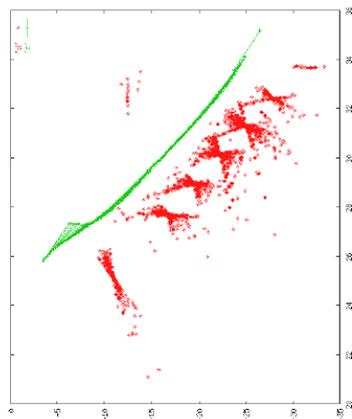
Sistem sensor kedua yang dibuat oleh Thorpe (untuk kendaraan robotnya) adalah aliran optis berbasis pensensoran omnikamera. Perpindahan atau pergerakan kendaraan secara umum dibatasi oleh suatu bidang pensensoran, dan pada kasus omnikamera, citra yang ditangkap akan berada pada kedudukan horisontal. Untuk menyelesaikan permasalahan kedudukan horisontal citra tersebut, maka kita memanfaatkan operator edge-detection Sobel untuk mendeteksi keseluruhan tepi-tepi citra yang vertikal, dan selanjutnya menjejaki keseluruhan tepi yang diperoleh dari satu citra ke citra yang lain untuk objek yang sama. Secara visual, dapat dilihat pada gambar 5.4.



Gambar 5.4 Kiri, citra dari omnikamera. Kendaraan dapat dilihat pada bagian kanan citra. Kanan, citra yang telah diperbaiki, dikonversikan menjadi koordinat selindris, kotak pada bagian atas menunjukkan porsi yang diharapkan, posisi citra sesuai seakan-akan tepat berhadapan dengan kendaraan kita.



Gambar 5.5 Citra orisinal, dan urutan dari lima citra tepi



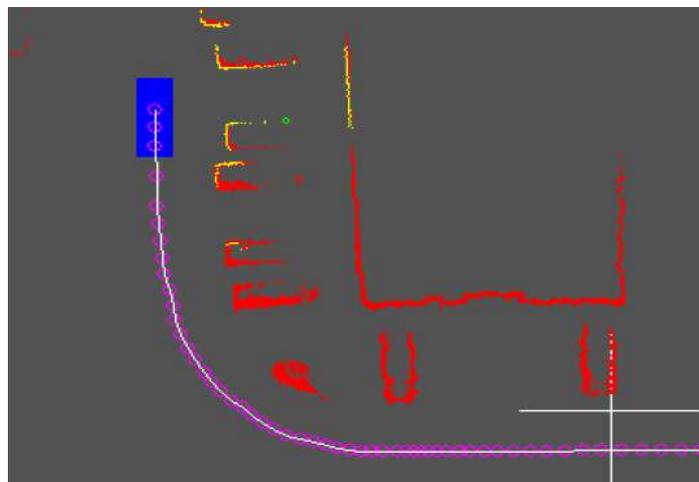
Gambar 5.6 Peta yang dihasilkan dari aliran optis, berbagai segmen garis menunjukkan posisi kendaraan yang telah dikalkulasi dan yang muncul pada setiap frame. Masing-masing posisi objek dikalkulasikan sebagai irisan terdekat dari lima garis penglihatan dari penjejakan optis. Langkah berikutnya adalah melakukan filter pada posisi yang terkalkulasi pada berbagai pengamatan, untuk mengurangi sebaran titik pada lokasi rata-rata

### 5.3 Mendeteksi pergerakan dan pergerakan lanjutan objek

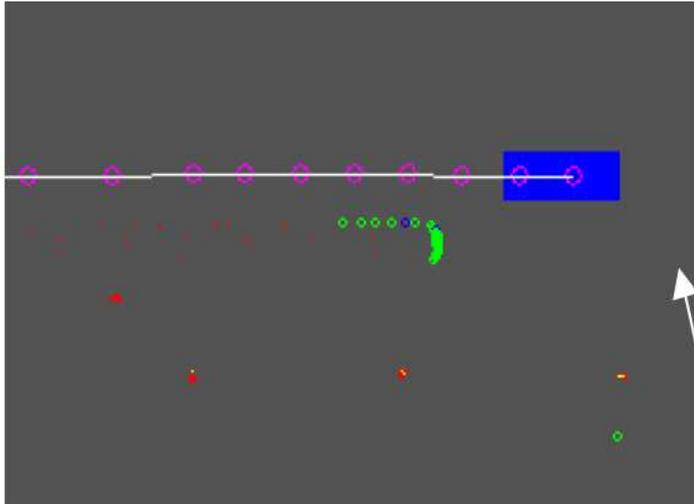
Pendeteksian objek sangat diperlukan tetapi tidak cukup hanya pendeteksian saja. Pertanyaan yang sungguh ingin kita jawab disini bukan dimanakah posisi objek tersebut sekarang?, tetapi kemanakah objek tersebut akan selanjutnya bergerak?.

Untuk itu, Thorpe menggunakan pengolahan ladar untuk menjelajahi dan memprediksi pergerakan objek serta mengklasifikasikan objek secara terbatas. Dalam uji cobanya, Thorpe dan kawan-kawan menggunakan sebuah ladar pada kendaraan NavLab 8 yang dirancangnya, lalu kendaraan tersebut dikendarai mengelilingi kampus CMU dengan melalui berbagai jalan yang terdapat disekitarnya untuk mengumpulkan data *range*. Untuk masing-masing frame data *range*, Thorpe pertamanya melakukan pensegmentasian peta *range* menjadi objek terkait. Untuk masing-masing objek, Thorpe melakukan perbandingan keadaan objek dengan daftar pergerakan objek yang sedang berlangsung, dan memisahkan keseluruhan titik-titik yang berhubungan dengan pergerakan objek tersebut. Untuk langkah utama kedua, Thorpe dan kawan-kawan memanfaatkan algoritma ICP (Iterative Closest Points) untuk mendaftarkan titik-titik yang bukan kategori “non-moving” dengan peta lokal yang dibentuk dari hasil pendeteksian sebelumnya, untuk menghitung posisi relatif kendaraan. Untuk memprediksi bentuk (keadaan struktur) dari kendaraan tersebut, Thorpe menggunakan filter Kalman.

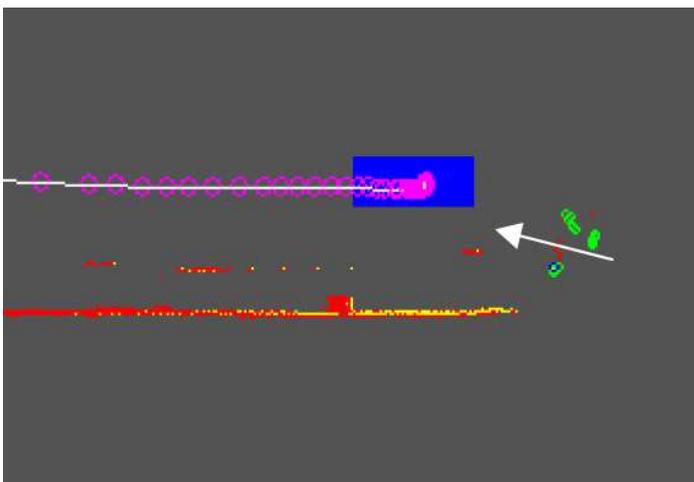
Setelah dilakukan peregistrasian pada dua frame yang ditemukan, maka kita gunakan suatu algoritma pendeteksian pergerakan objek dan pose objek yang telah dikalkulasikan untuk memisahkan sebarang objek yang bergerak dari objek stasionernya. Dari hasil *scanning* sebelumnya, telah diketahui yang mana daerah-daerah yang seharusnya tidak diduduki, dan jika ketahu ada suatu objek yang mendiami atau melalui daerah tersebut, maka ini berarti telah terjadinya suatu perpindahan. Hasil dari pelokalisasi dan pemetaan relatif objek, ditunjukkan pada gambar 5.7.



Gambar 5.7 Lokalisasi dan pemetaan



Gambar 5.8 Pendeteksian kendaraan, titik-titik yang ditunjukkan oleh tanda panah memiliki bentuk dan perpindahan dari suatu mobil yang bergerak



Gambar 5.9 Pendeteksian pejalan kaki, sejumlah titik-titik yang ditandai oleh tanda panah memiliki bentuk dan karakteristik seorang pejalan kaki



Gambar 5.10 Pendeteksian pejalan kaki

# Bab 6

## VISUAL HULL CONSTRUCTION

*“Visual Hull (VH) construction is a popular method of shape estimation. The method, also known as Shape from Silhouette (SFS), approximates shape of an object from multiple silhouette images by constructing an upper bound of the shape called the Visual Hull. SFS is used in many applications such as non-invasive 3D object digitization, 3D object recognition and more recently human motion tracking and analysis. Though SFS is straightforward to implement and easy to use, it has several limitations. Most existing SFS methods are too slow for real-time applications and the estimated shape is sensitive to silhouette noise and camera calibration errors. Moreover, VH is only a conservative approximation of the actual shape of the object and the approximation can be very coarse when there are only a few cameras.”*

### 6.1 Pendahuluan

Pengestimasi bentuk objek 3D telah menjadi suatu topik yang utama dan aktif di dalam sub-bidang komputer visi. Berbagai pendekatan algoritmis untuk pengestimasi telah banyak diajukan dalam dua dekade terakhir ini yang diantaranya seperti yang diungkapkan oleh L. G. Roberts. dalam penelitiannya *Machine perception of three-dimensional solids*, Pada tahun 1977, H. H. Baker. *Depth from Edge- and Intensity-Based Stereo*, pada tahun 1981, W. N. Martin dan J. K. Aggarwal, *Volumetric descriptions of objects from multiple views*, 1983, P. Giblin dan R. Weiss, *Reconstruction of surface from profiles*, (In *Proceedings of Image Understanding Workshop*), 1987, C. Tomasi dan T. Kanade, *Shape and motion from image streams under orthography*, 1992, O. Faugeras, *Three-Dimensional Computer Vision : A Geometric Viewpoint*, MIT Press, 1993, P. Rander, P.J. Narayanan, dan T. Kanade, *Virtualized reality : Constructing time-varying virtual worlds from real world events*, (In *Proceedings of IEEE Visualization '97*), 1997, S. Seitz dan C.R. Dyer. *Photorealistic scene reconstruction by voxel coloring*, (In *Proceedings of CVPR'97*), 1997, E. Boyer dan M. Berger. *3D surface reconstruction using occluding contours*, 1997, dan K.N. Kutulakos dan S. Seitz. *A theory of shape by space carving*, 2000. Diantara keseluruhan pendekatan-pendekatan yang telah disebutkan di atas, ada suatu pendekatan yang disebut “Visual Hull (VH) construction” atau juga sering diistilahkan *Shape From Silhouettes (SFS)*, suatu metode yang mengestimasi bentuk dari sebuah objek berdasarkan keadaan bayangan gelap objek tersebut. Metode SFS ini diperkenalkan oleh Baumgart dalam disertasi Ph.D-nya yang

berjudul *Geometric Modelling for Computer Vision*, Universitas Stanford, pada tahun 1974. Pada saat ini, berbagai variasi dari SFS, representasi VH dan aplikasinya telah banyak muncul di dalam berbagai karya ilmiah atau riset yang diantaranya, Aggarwal telah menggunakan voxel sebagai suatu representasi dari VH, Potmesil memanfaatkan struktur data oktree untuk mempercepat konstruksi dari VH, De Bonet dan Viola memperluas ide rekonstruksi voxel yang diaplikasikan terhadap objek transparan sebagai penerapan konsep Roxels, Buehler menggunakan VH sebagai basis geometris penafsiran citra, dan lain-lain. Dengan jumlah riset yang memanfaatkan pendekatan SFS tersebut, maka, (dapat kita katakan) SFS merupakan metode standar dan populer yang digunakan untuk pengestimasi bentuk objek.

Estimasi bentuk menggunakan SFS memiliki beberapa keunggulan yang diantaranya adalah *silhouette* mudah dibaca dan diperoleh, secara khusus dalam lingkungan yang *indoor* (terkondisikan/di dalam ruangan) dimana posisi bayangan hanya mengalami sedikit perubahan (pergeseran) dan kamera dalam keadaan statis. Implementasi dari SFS juga masih relatif terkondisikan dibandingkan dengan berbagai metode lainnya seperti metode *multi-baseline stereo* dan metode *space carving*. Visual Hull terkonstruksi dari SFS merupakan batas-batas atas dari sebuah objek, dan hal ini dapat kita katakan sebagai sifat konservatif yang melekat untuk berbagai aplikasi seperti menghindari rintangan di dalam manipulasi robot dan analisa visibilitas dalam navigasi, yang mana lebih memerlukan batas-batas atas dari objek daripada batas-batas dalamnya.

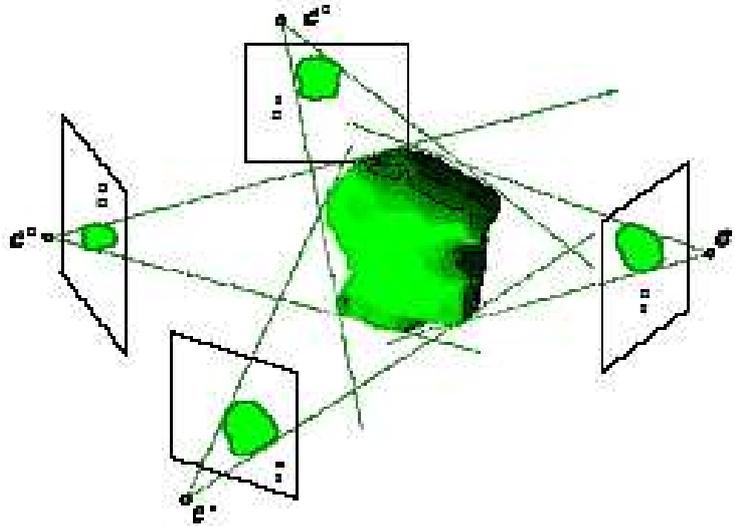
SFS memiliki keterbatasan dalam berbagai hal seperti, pengonsumsi waktu yang boros untuk berbagai aplikasi pada waktu nyata, sensitif terhadap error pada pengestraksian *silhouette* dan kalibrasi kamera, sehingga kita sulit untuk menghindari berbagai noise yang terjadi pada citra yang diambil, lebih-lebih, VH yang diperoleh dari SFS hanya mengaproksimasi bentuk aktual dari sebuah objek, sehingga hasil aproksimasi akan begitu kasar bila fasilitas kamera yang digunakan untuk pengambilan gambar berjumlah sedikit. Tujuan dari kajian ini adalah menginvestigasi berbagai permasalahan yang sering muncul dalam pengaproksimasi menggunakan pendekatan SFS, dan memberikan solusi yang tepat untuk mengatasi permasalahan tersebut.

## 6.2 Latar belakang

### 6.2.1 Skenario SFS

Anggap ada  $K$  kamera yang diposisikan disekitar objek 3D yang tetap,  $O$ , anggap  $\{S_j^k; k = 1, \dots, K\}$  merupakan sekumpulan citra *silhouette* dari objek  $O$  yang diperoleh dari  $K$  kamera pada waktu  $t_j$ . Asumsikan kamera dikalibrasikan dengan  $\prod^k ( ) : R^3 \rightarrow R^2$  dan  $C^k$  menjadi fungsi proyeksi perspektif untuk kamera

tersebut, dengan  $k$  adalah pusat dari kamera. Maka  $p = \prod^k(P)$  adalah koordinat-koordinat 2D dari sebuah titik 3D,  $P$ , dalam citra ke- $k$ . Penambahan notasi yang diperlukan,  $\prod^k(A)$  merupakan representasi proyeksi sebuah volume  $A$  pada bidang citra dari kamera  $k$ . Contoh skenario dapat dilihat pada gambar 6.1.



Gambar 6.1 Contoh skenario SFS, sebuah objek  $O$  dikelilingi oleh empat kamera, citra *silhouette* pada waktu  $t_1$ , dan pusat kamera direpresentasikan oleh  $\{S_1^k\}$  dan  $\{C^k\}$

### 6.2.2 Visual Hull

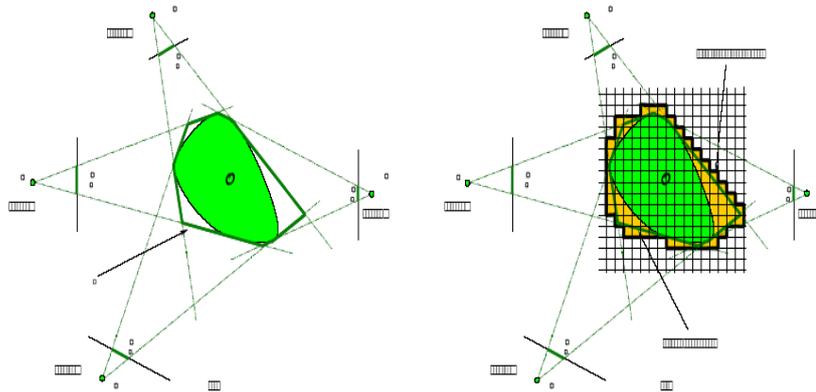
Asumsikan, diberikan sekumpulan  $K$  citra *silhouette*  $\{S_j^k\}$  dan fungsi proyeksi  $\{\prod^k\}$ . Sebuah volume  $A$  dikatakan terdefiniskan dengan jelas, jika dan hanya jika, proyeksi  $A$  pada bidang citra ke- $k$  bertepatan secara mutlak dengan citra *silhouette*  $S_j^k$  untuk semua  $k \in \{1, \dots, K\}$  atau dalam bentuk formulasi dinotasikan sebagai  $\prod^k(A) = S_j^k$ . Jika ada paling sedikit satu entitas saja yang mendefinisikan citra *silhouette* secara tepat, maka dapat kita katakan himpunan citra *silhouette* adalah konsisten, sebaliknya kita katakan tidak konsisten. Secara normal kita hanya dapat memperoleh citra *silhouette* hanya dari objek yang konsisten, dengan mengabaikan error pada kalibrasi kamera atau keadaan noise yang terdapat dari citra.

Visual hull  $H_j$  dengan kecenderungan terhadap sekumpulan citra *silhouette* yang konsisten didefinisikan sebagai sebuah volume terbesar yang mungkin yang mendefinisikan secara tepat  $\{S_j^k, \forall k = 1, \dots, K\}$

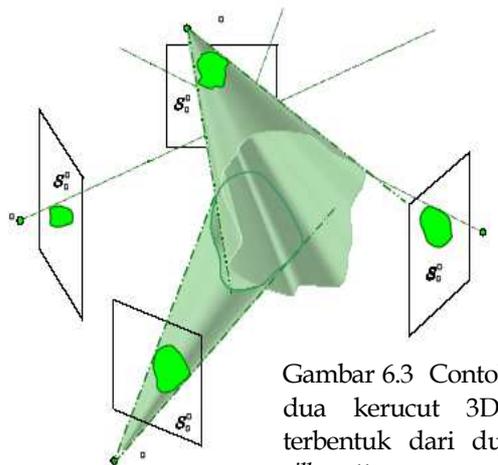
### 6.3 Visual hull construction

#### 6.3.1 Komputasi visual hull dengan irisan batas kerucut

Untuk sekumpulan citra *silhouette*, yang tetap, pengkonstruksian visual hull adalah dengan menginterseksi keseluruhan batas-batas kerucut yang terbentuk dari berbagai citra *silhouette*,  $S_j^k$ , dengan pusat kamera yang berkorespondensi,  $C^k$ . Visual hull hasil terdiri dari jalur-jalur permukaan seperti yang diperlihatkan pada gambar 6.2.



Gambar 6.2 Sebuah contoh pengkonstruksian (2D) visual hull  $H_1$  dari sekumpulan *silhouette*  $S_1^k$  dan pusat kamera  $C^k$ , (a) dengan interseksi langsung, dan (b) pengaproksimasian berbasis voxel, daerah yang berbayang merepresentasikan aproksimasi visual hull



Gambar 6.3 Contoh irisan dua kerucut 3D yang terbentuk dari dua citra *silhouette*

Walaupun tepat digunakan untuk objek 2D, pada 3D cara yang sama tidak begitu praktis untuk digunakan, dikarenakan kesulitan untuk mengekspresikan jalur-jalur permukaan secara efektif, yang diperoleh dari penginterseksian beberapa kerucut dengan memanfaatkan berbagai atribut geometri. Selain itu komputasional yang kompleks dan ketidakstabilan numerik juga menjadi alasan kenapa metode yang sama tidak digunakan pada bidang 3D (lihat gambar 6.3).

### 6.3.2 Konstruksi dari pengaproksimasian visual hull dengan pengklasifikasian voxel

Para peneliti telah mengajukan suatu cara lain yang lebih efektif untuk merepresentasikan dan mengkonstruksi visual hull 3D. berbagai cara atau metode tersebut sejauh ini lebih bersifat diskrit dan berbasiskan pada voxel. Versi sederhana yang diajukan diringkas pada algoritma berikut :

```

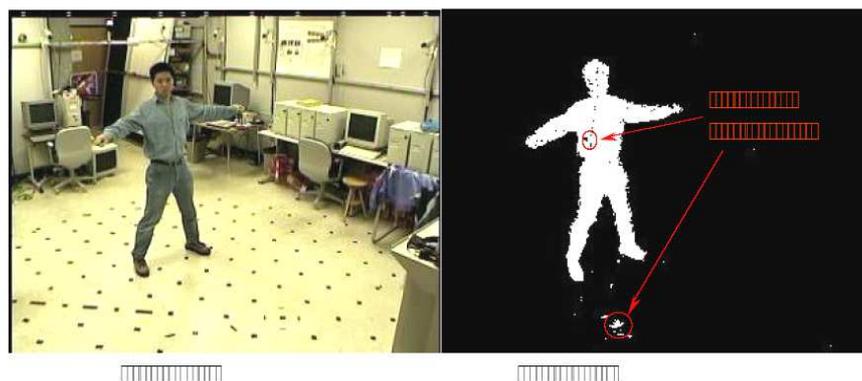
Algorithm 1 : Standard Voxel-based Shape from Silhouette

1. Divide the space of interest into  $N \times N \times N$  discrete voxels  $v_n, n = 1, \dots, N^3$ .
2. Initialize all the  $N^3$  voxels as inside voxels.
3. Perform :
   For  $n = 1$  to  $N^3$  do {
     For  $k = 1$  to  $K$  do {
       (a) Project the voxel  $v_n$  with  $\Pi^k(\cdot)$  onto the  $k^{th}$  image plane;
       (b) If the projected area  $\Pi^k(v_n)$  lies totally outside  $\mathcal{S}_j^k$ 
           then classify  $v_n$  as outside voxel;
     }
   }
4. The approximate visual hull  $\hat{\mathcal{H}}_j$  is represented by the union of all the inside voxels.
    
```

## 6.4 Visual hull construction dari *silhouette* "noise" pada waktu nyata

### 6.4.1 Kecepatan dan keakuratan metode visual hull construction berbasis voxel

Kecepatan metode SFS berbasis voxel yang dipaparkan di dalam algoritma 1 sangat tergantung pada dua prosedur, (1) pemroyeksian voxel (langkah 3(a)) dan, (2) pengujian tumpang tindih *silhouette* (langkah 3(b)). Kedua prosedur dan juga kualitas dari citra *silhouette* tersebut sangat mempengaruhi keakuratan penaksiran dari visual hull yang akan dicapai, terlebih-lebih citra *silhouette* selalu memiliki noise.



Gambar 6.4 Sebuah contoh citra “foreground” dan *silhouette* -nya yang diperoleh dengan pengekstraksian menggunakan metode substraksi “background” pada waktu nyata. Terlihat pada gambar ada beberapa penandaan yang salah pada *silhouette* -nya yang diakibatkan oleh keadaan noise dari citra aslinya

Gambar 6.4 menunjukkan beberapa kesalahan penandaan pada citra *silhouette* yang diakibatkan oleh keadaan noise pada citra aslinya. Walaupun keadaan ini bisa dinormalisasikan dengan menggunakan beberapa operasi morfologi dan analisa komponen terkait, tetapi tetap proses dari algoritma tersebut mengkonsumsi waktu yang begitu lambat untuk digunakan pada aplikasi-aplikasi waktu nyata. Dan dikarenakan oleh hal-hal yang demikian yang telah disebutkan, maka pada bagian sub bab ini, kami akan membahas suatu pendekatan yang lebih berorientasi pada pengimplementasian VH secara cepat dan tentunya akurat.

Untuk mengkararakteristik kualitas citra *silhouette*, maka kita lakukan suatu formulasi sebagai : andaikan  $\xi$  sebagai suatu probabilitas selama proses pengekstraksian, dimana pixel-pixel yang bukan *silhouette* diidentifikasi salah sebagai pixel-pixel *silhouette*, sebaliknya anggap  $\eta$  merepresentasikan probabilitas pixel-pixel *silhouette* yang diidentifikasi salah sebagai pixel-pixel yang bukan *silhouette*. Asumsikan pula keadaan noise tidak meliputi diantara pixel-pixel tersebut, maka ada dua cara untuk menentukan  $\xi$  dan  $\eta$ : secara teoritis, dengan mengasumsikan fungsi distribusi probabilitas untuk error atau secara eksperimen dengan melakukan pemeriksaan pada setiap pixel-pixel yang ditandai salah di dalam sekumpulan besar citra *silhouette*.

Untuk menemukan suatu voxel terproyeksi  $\prod^k(v_n)$  adalah dengan memproyeksikan kedelapan vertex  $v_n$  terhadap bidang citra ke- $k$  dan mengkomputasi *convex hull* dari berbagai titik pemroyeksian. Anggap ada  $Z$  pixel di dalam *convex hull*, maka untuk mengimplementasikan *silhouette* yang tumpang tindih adalah dengan memeriksa keseluruhan  $Z$  pixel, dan jika sedikitnya ada  $Z_e$  dari

keseluruhan  $Z$  yang terletak di dalam citra *silhouette*  $S_j^k$ , maka voxel dapat dikatakan berada di dalam  $S_j^k$ . Pengecualian untuk citra *silhouette* yang mengandung noise, implementasi di atas akan menyebabkan misklasifikasi voxel. Ada dua jenis dari misklasifikasi voxel : *False Acceptance* (FA) yang berarti sebuah voxel yang berada di luar diklasifikasikan berada di dalam dan, *False Rejection* (FR) yang berarti sebuah voxel yang berada di dalam diklasifikasikan berada di luar. Probabilitas  $P(FR)$  dan  $P(FA)$  sangat tergantung pada  $\xi, \eta, Z$  dan  $Z_\epsilon$ .

- *False Acceptance* (FA)

$$P(FA) = \left[ \sum_{i=Z_\epsilon}^Z \binom{Z}{i} \xi^i (1-\xi)^{Z-i} \right]^K$$

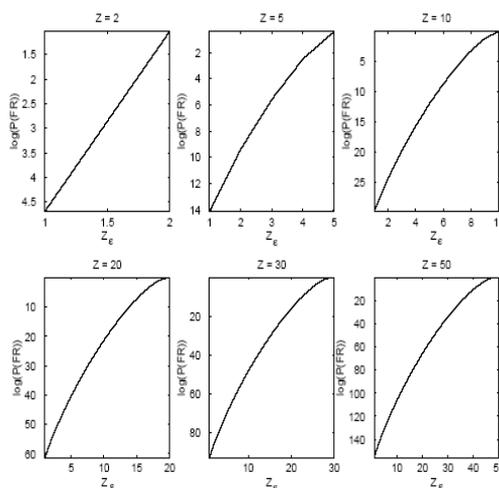
- *False Rejection*

$$P(FR) = p \sum_{j=0}^{K-1} (1-p)^j$$

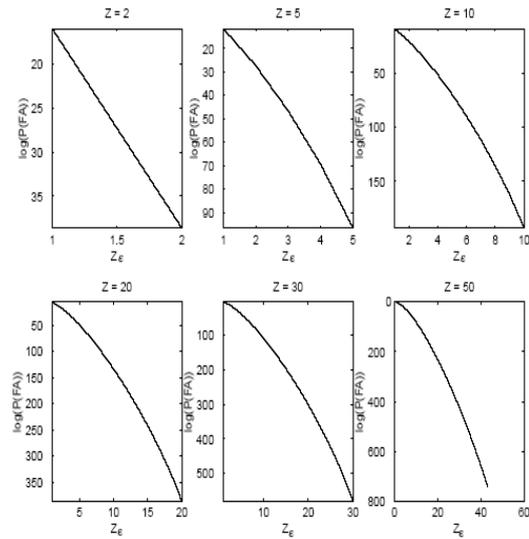
dimana

$$p = \sum_{i=Z-Z_\epsilon+1}^Z \binom{Z}{i} \eta^i (1-\eta)^{Z-i}$$

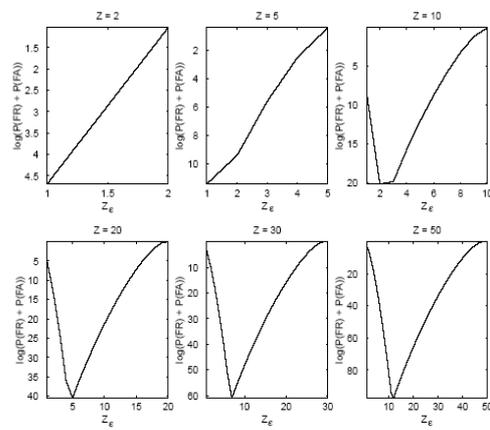
Eksponensial  $K$  pada persamaan di atas sangat berkaitan dengan kenyataan bahwa voxel yang berada di luar diklasifikasikan secara salah sebagai voxel yang berada di dalam pada citra  $K$ . Sedangkan  $p$  merupakan kemungkinan terjadinya pengklasifikasian yang salah terhadap sebuah voxel yang berada di dalam yang diidentifikasi sebagai sebuah voxel yang berada di luar. Penambahan nilai  $j$  didasari oleh suatu alasan bahwa sebuah voxel yang berada di dalam yang diklasifikasikan secara salah sebagai voxel yang berada di luar, hanya mungkin terjadi sekali kesalahan pengklasifikasian di dalam sebuah citra, sedangkan pada sisa citra selanjutnya akan melalui pengujian berikutnya.



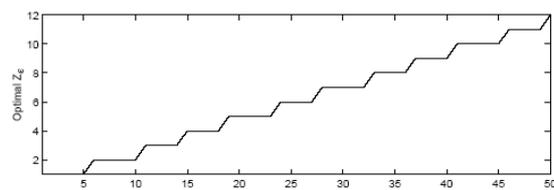
(a) Graph of  $\log(P(FR))$  against  $Z_\epsilon$  for different  $Z$



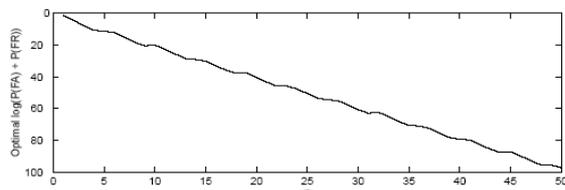
(b) Graph of  $\log(P(FA))$  against  $Z_\epsilon$  for different Z



(c) Graph of  $\log(P(FR) + P(FA))$  against  $Z_\epsilon$  for different Z

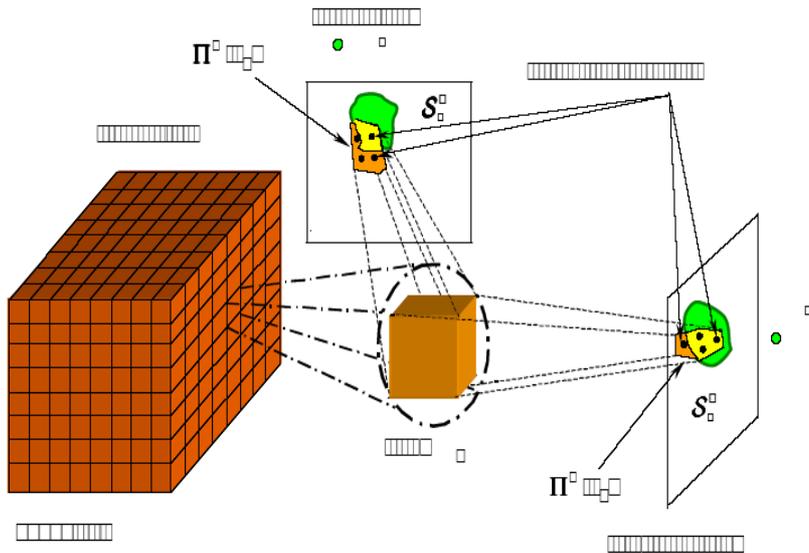


(d) Graph of Optimal  $Z_\epsilon$  against Z



(e) Graph of Optimal  $\log(P(FA) + P(FR))$  against Z

Gambar 6.5 (a)  $\log(P(FR))$  vs  $Z_\epsilon$ , (b)  $\log(P(FA))$  vs  $Z_\epsilon$ , (c)  $\log(P(FA) + P(FR))$  vs  $Z_\epsilon$ , (d) optimal Z vs  $Z_\epsilon$ , (e) optimal  $\log(P(FA) + P(FR))$  vs Z



Gambar 6.6 Modifikasi *visual hull construction* berbasis voxel menggunakan SPOT

#### 6.4.2 Algoritma fast visual hull construction berbasis voxel menggunakan Sparse Pixel Occupancy Test (SPOT)

Untuk mengaplikasikan implementasi langsung dari prosedur (1) proyeksi voxel, dan (2) pengujian *silhouette* yang tumpang tindih, maka kita memerlukan  $N^3Z$  jumlah memori dan  $N^3Z$  jumlah pengujian untuk sebuah voxel yang memiliki volume  $N \times N \times N$ . Dikarenakan alasan keperluan memori dan jumlah pengujian yang besar yang menyebabkan operasi begitu lambat dan mahal, maka, diperlukan suatu program yang handal yang lebih murah di dalam pengoperasiannya dan dengan kebutuhan memori yang lebih sedikit. Pendekatan tersebut disebut juga *Sparse Pixel Occupancy Test* atau disingkat SPOT. Ide dari pendekatan SPOT diilustrasikan pada gambar 6.6.

Ada dua keuntungan dalam menggunakan SPOT, pertama, waktu yang diperlukan  $\frac{Z}{Q}$  lebih cepat dibandingkan pendekatan langsung, kedua, memori yang diperlukan  $\frac{Z}{Q}$  lebih hemat daripada pendekatan biasa. Kendala yang dihadapi adalah seiring dengan pengurangan jumlah pengujian yang dioperasikan, maka akan menimbulkan peluang misklasifikasi yang lebih besar. Pertanyaan yang tersisa adalah bagaimanakah kita memilih  $Q$  dan  $Q_\epsilon$  secara tepat?.

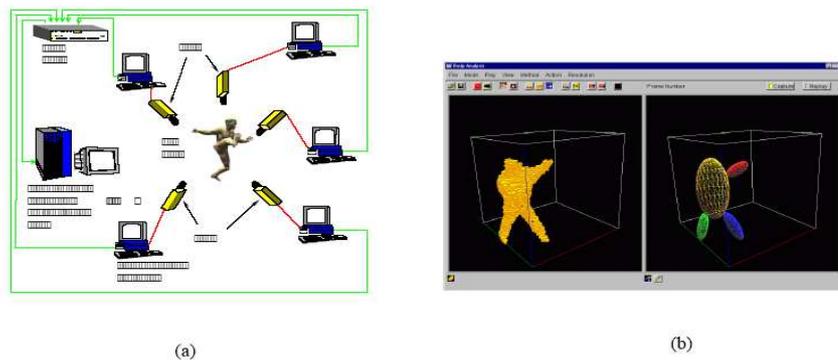
Kita asumsikan pixel-pixel adalah saling lepas satu sama lainnya dan  $Q$  dipilih secara acak, ide penentuan  $Q$  berada di luar  $Z$  secara teoritis adalah ekivalen terhadap pengujian suatu area proyeksi dengan  $Q$  pixel. Dengan kata lain, formulasi

$$P(FA) = \left[ \sum_{i=\xi, \xi}^Z \binom{Z}{i} \xi^i (1-\xi)^{Z-i} \right]^K$$

dan

$$P(FR) = p \sum_{j=0}^{K-1} (1-p)^j$$

dan grafik pada gambar 6.5 adalah valid untuk penganalisaan SPOT, dengan syarat kita harus menggantikan peran  $Z$  dengan  $Q$ . pemilihan  $Q$  semata-mata hanya mempertimbangkan keadaan kecepatan, memori, dan akurasi. Di dalam aplikasi waktu nyata kata-kata cepat lebih dibutuhkan daripada kualitas, sehingga nilai  $Q$  yang begitu rendah hanya bermaksud untuk meningkatkan keakuratan kecepatan.



Gambar 6.7 (a) Sistem arsitektur dari sistem rekonstruksi pemodelan pergerakan manusia untuk aplikasi waktu nyata, (b) Screen-shot user interface

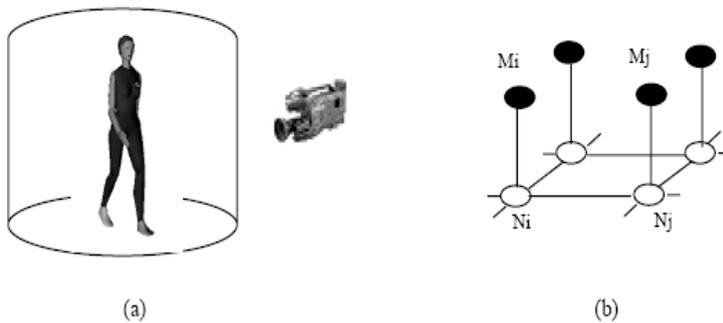
# Bab 7

## PENDETEKSIAN GERAKAN TUBUH

*“We present an algorithm for automatic inference of human upper body motion in a natural scene. A graph model is proposed for inferring human upper body motion, and motion inference is posed as a mapping problem between state nodes in the graph model and features in image patches. A multiple-frame inference algorithm is proposed to combine both structural and temporal constraints in human upper body motion. Belief propagation and dynamic programming algorithms are utilized for Bayesian inference in this graph. We also present a method for capturing constraints of human body configuration under different view angles. The algorithm is applied in a prototype system that can automatically detect upper body motion from videos, without manual initialization of body parts.”*

### 7.1 Pendahuluan

Di dalam interface perseptual manusia-mesin, manusia melakukan komunikasi dengan komputer menggunakan bahasa tubuh atau berbagai pola gerak tertentu. Kebanyakan aktivitas peneliti lebih memfokuskan pada rekognisi dan tracking dari pergerakan dan pola gerak khusus tubuh manusia. Selain daripada rekognisi aktivitas rutin manusia, seperti berjalan misalkan, pola isyarat tubuh bagian atas manusia juga memainkan peranan penting dalam komunikasi, dan juga sebagai alat ukur yang penting untuk pemahaman intensi dan interaksi. Dalam bab ini, kita akan membahas tentang suatu sistem pendeteksiian gerakan atau isyarat tubuh bagian atas manusia secara otomatis.



Gambar 7.1 (a) Pendeteksiian gerakan manusia dan analisis gerak isyarat, (b) suatu pemodelan jaringan Markov dengan state-state tertentu dan node pengukuran

## 7.2 Pemodelan probabilitas *single-frame*

Dalam wilayah kajian grafik dan animasi, konfigurasi tubuh manusia didefinisikan sebagai suatu translasi global dan rotasi *root*, yang secara normal diseleksi sebagai pangkalan dengan sudut relatif antara sendi-sendi tubuh terhadap posisi *root*-nya. Dikarenakan kajian kita batasi hanya pada bagian atas tubuh saja, maka dapat kita definisikan *root* sebagai *shoulder-girdle*. Posisi dari *shoulder-girdle* ditaksir atau diperkirakan berdasarkan posisi dari *face* (menggunakan sistem pendeteksian *face*), dan asumsi pose.

Tujuan kita di sini adalah untuk memperoleh posisi dari sendi-sendi tubuh berdasarkan citra yang diamati. Pada bingkai kerja Bayesian, diberikan fitur citra  $x$ , konfigurasi tubuh sebagai  $h$  dapat diperkirakan sebagai :

$$h = \arg \max_h P(h | x)$$

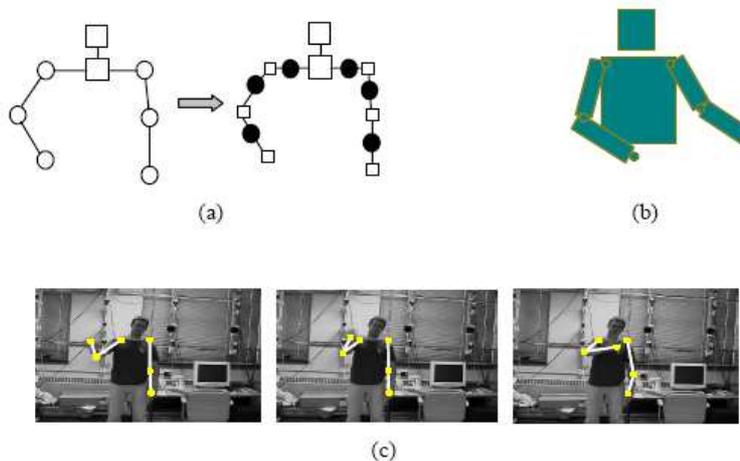
dimana

$$P(h | x) = cP(x | h)P(h)$$

pada kajian bab ini, konfigurasi tubuh  $h$  terdiri dari berbagai posisi dari sendi-sendi tubuh  $(1, 2, \dots, N)$  yang didenotasikan sebagai  $(s_1, s_2, \dots, s_n)$ .  $P(h)$  merupakan fungsi probabilitas *a priori* dari konfigurasi tubuh bagian atas  $h$ .  $P(x | h)$  merupakan fungsi kemungkinan, dan  $P(h | x)$  merupakan probabilitas posterior.

### 7.2.1 Pemodelan jaringan Markov

Model jaringan Markov ditunjukkan pada gambar 7.2.



Gambar 7.2 (a) Pemodelan jaringan Markov untuk tubuh bagian atas, bulatan putih merupakan node *state*, bulatan yang hitam merupakan node pengukuran, (b) pemodelan manusia kartun dan (c) hasil pendeteksian pergerakan tubuh bagian atas

Di dalam model jaringan Markov, bulatan putih atau bulatan kosong merepresentasikan posisi 2D dari sendi-sendi tubuh. Di dalam jaringan Markov masing-masing node *state* terhubung dengan node pengukuran yang digambar sebagai bulatan terisi berwarna hitam (gambar 7.2 (a)). Koneksi diantara node-node dalam jaringan Markov mengindikasikan ketergantungan statistik.

Berdasarkan definisi node state, maka tidak ada satupun citra yang diamati yang berkorespondensi terhadap mereka. Untuk menyelesaikan permasalahan ini, kita dapat memanfaatkan bagian-bagian tubuh yang terletak diantara sendi-sendi sebagai suatu alat pengukuran. Seperti yang ditunjukkan pada gambar 7.2(a), maka dapat kita definisikan ruas-ruas citra yang berkorespondensi terhadap *wrist*, *elbow*, dan *shoulder joint* sebagai *lower arm*, *upper arm*, dan *shoulder girdle* secara berturut-turut. Selanjutnya ruas-ruas citra tersebut dimodelkan dalam bentuk manusia kartun, seperti yang ditunjukkan pada gambar 7.2(b).

Kita denotasikan state pada node  $i$  sebagai  $s_i$ , dan pengamatan pada node pengukuran terkait  $M_i$  sebagai  $x_i$ . Seperti telah dijelaskan sebelumnya,  $s_i$  berkaitan dengan posisi sendi. Kita hitung  $x_i$  dengan mengakumulasikan petunjuk atau isyarat gerakan dalam ruas citra. Fungsi  $P_i(x_i | s_i)$  direpresentasikan oleh berbagai koneksi antara node state  $N_i$  dan node pengukuran  $M_i$  di dalam model jaringan Markov.

Secara praktis,  $P_i(x_i | s_i)$  dihitung dengan mengakumulasikan petunjuk atau isyarat gerakan dalam ruas citra yang tidak terkait dengan suatu fungsi probabilitas. Maka, kita perlu mengkonversikan pengukuran "energy" menjadi pengukuran probabilitas. Untuk suatu *body joint*  $i$ , fungsi kemungkinan  $P_i(x_i | s_i)$  dihitung sebagai

$$P_i(x_i | s_i) = C \exp(-E)$$

dimana  $E$  merupakan fungsi "energy" yang dihitung sebagai

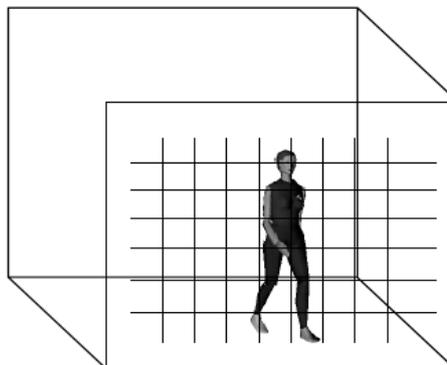
$$E = \xi(s_i, x_i)$$

$E$  berkaitan erat dengan akumulasi isyarat gerakan dalam ruas-ruas citra, yang ditentukan oleh posisi *body joint*  $s_i$ .  $C$  merupakan koefisien normalisasi.

### 7.2.2 Pembelajaran model jaringan Markov

Jaringan Markov mendekomposisi permasalahan inferensi Bayesian menjadi state-state lokal dan pengukuran tertentu. Dalam jaringan Markov, kita memanfaatkan pengukuran probabilitas lokal seperti  $P_i(x_i | s_i)$ ,  $P_j(x_j | s_j)$  dan  $P_{ij}(x_{ij} | s_{ij})$  untuk menyimpulkan posisi sendi tubuh (*body joint*)  $(s_1, s_2, \dots, s_N)$ .

Penambahan lain, untuk menghitung fungsi kemungkinan  $P_i(x_i | s_i)$  maka kita terlebih dahulu mengestimasi  $P_{ij}(x_{ij} | s_{ij})$ , probabilitas *a priori* untuk berbagai posisi sendi tubuh, atau konfigurasi baku tubuh manusia.



Gambar 7.3 Pembelajaran probabilitas *joint* pada berbagai titik sampel dalam ruang citra 2D

Kita gunakan metode pembelajaran “supervise” untuk mengestimasi  $P_{ij}(s_i | s_j)$  di antara sendi  $i$  dan  $j$ . Secara keseluruhan kita dapat mengambil sampel ruang citra 2D seperti yang ditunjukkan pada gambar 7.3. Sebelum proses pembelajaran masing-masing pasangan posisi sampel  $(s_i, s_j)$  harus diberikan probabilitas yang sama :

$$P_{ij}(s_i, s_j) = 1.0 / N_s * N_s$$

dimana  $N_s$  merupakan jumlah posisi sampel. Selanjutnya kita dapat menggunakan strategi *bootstrapping* untuk mengestimasi  $P_{ij}(s_i | s_j)$ .

Untuk masing-masing pasangan posisi *body joint*  $(s_i, s_j)$  yang bukan konfigurasi tetapan tubuh manusia yang valid, dapat dilakukan pengurangan probabilitas dengan :

$$P_{ij}(s_i | s_j) = \frac{P_{ij}(s_i, s_j)}{T}$$

dimana  $T$  merupakan suatu konstanta,  $T > 1$ . Jika  $s_i$  dan/atau  $s_j$  terletak diantara posisi sampel, secara sederhana kita dapat menghubungkan mereka ke koordinat sampel yang paling dekat, dan selanjutnya  $P_{ij}(s_i | s_j)$  dihitung sebagai :

$$P_{ij}(s_i | s_j) = \frac{P_{ij}(s_i | s_j)}{\sum_{s_i, s_j} P_{ij}(s_i | s_j)}$$

yang juga sering diistilahkan sebagai probabilitas renormalisasi.

### 7.2.3 Melibatkan pose dalam model jaringan Markov

Fleksibilitas sikap konfigurasi tubuh manusia merupakan permasalahan yang sangat rumit. Di bawah berbagai sikap tubuh yang berbeda, maka nilai  $P_{ij}(s_i | s_j)$  akan berbeda pula.

Anggap sikap tubuh (*body pose*) sebagai  $u$ , maka persamaan

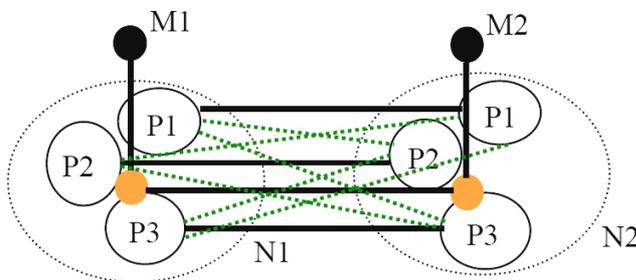
$$P(h | x) = cP(x | h)P(h)$$

akan menjadi

$$P(h, u | x) = cP(x | h, u)P(h | u)P(u)$$

Berbagai pose yang berbeda,  $u$ , akan menimbulkan distribusi probabilitas  $h$  yang berbeda pula. Secara umum, untuk pose yang berbeda  $K$ ,  $P(u)$  dapat direpresentasikan sebagai suatu bobot campuran fungsi basis  $K$  :

$$P(h) = \sum_{k=1}^K w_k \cdot \alpha_k(h)$$



Gambar 7.4 Koneksi di antara state-state dengan berbagai pose  $P1, P2$ , dan  $P3$ .  $N1, N2$  dan  $M1, M2$  merupakan node state dan node pengukuran secara berturut-turut

$w_k$  merupakan bobot campuran, yang merepresentasikan probabilitas untuk masing-masing sikap (pose)  $k$  dan  $\alpha_k(h)$  merupakan distribusi probabilitas  $h$  di bawah pose  $u = k$ .

Untuk mendefinisikan pose yang berbeda dalam model jaringan Markov yang sama, kita dapat memilah state-state dalam masing-masing node menjadi kelompok-kelompok dengan ciri pose tertentu. Kemudian kita definisikan koneksi antara bobot sebagai berikut, untuk state  $s_i$  dan  $s_j$  dengan pose yang sama, maka  $P_{ij}(s_i^{P_m}, s_j^{P_n})$  didefinisikan berdasarkan suatu penaksiran, sedangkan untuk state  $s_i$  dan  $s_j$  dengan pose yang berbeda, maka

$$P_{ij}(s_i^{P_m}, s_j^{P_n}) = 0, \quad m \neq n$$

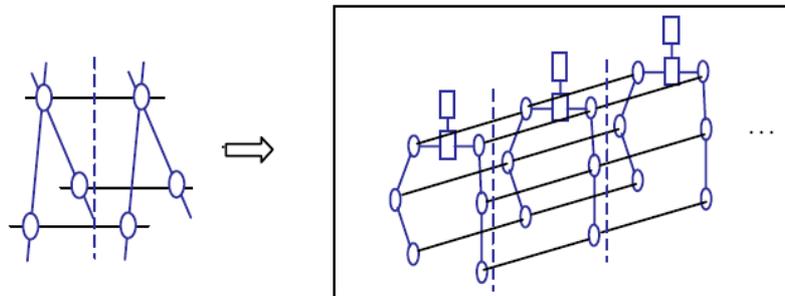
dimana  $s_i^{P_m}$  merepresentasikan state dengan postur  $P_m$ . Sedangkan definisi dari  $P_{ij}(s_i^{P_m}, s_j^{P_n})$  diilustrasikan pada gambar 7.4.

### 7.3 Pemodelan probabilitas *multiple-frame*

#### 7.3.1 Batasan temporal

Batasan temporal dapat ditambahkan ke dalam model jaringan Markov dengan menghubungkan node-node state yang berkorespondensi dengan berbagai sendi tubuh yang sama di dalam frame-frame yang berkaitan, dan probabilitas *joint*-nya telah didefinisikan terlebih dahulu. Anggap  $s_i^t$  sebagai sebuah state dari node  $i$  pada frame  $t$ , dan  $s_i^{t+1}$  merupakan sebuah state dari node  $i$  pada frame  $t+1$ , maka probabilitas *joint*-nya didefinisikan sebagai :

$$P_i^{t,t+1}(s_i^t, s_i^{t+1}) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{|s_i^t - s_i^{t+1}|^2}{\sigma^2} \right\}$$



Gambar 7.5 Penambahan batasan temporal pada model jaringan Markov

$P_i^{t,t+1}(s_i^t, s_i^{t+1})$  juga dikatakan sebagai fungsi penentuan jarak Gaussian antara  $s_i^t$  dan  $s_i^{t+1}$ . Parameter kovarians  $\sigma$  dapat ditentukan secara empiris.

7.3.2 Inferensi

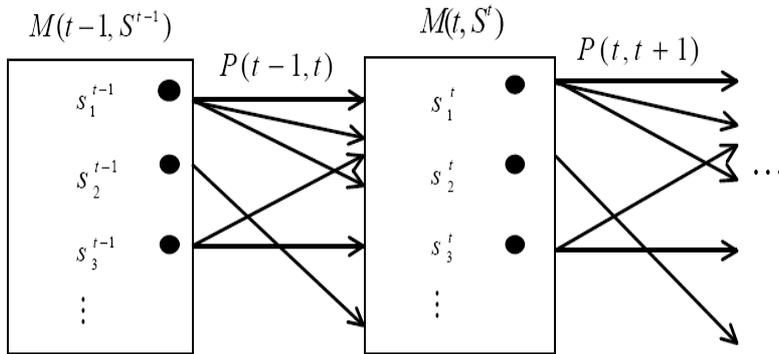
Pada sub materi ini, kita akan memaparkan algoritma inferensi yang diimplementasikan di dalam model jaringan Markov untuk aplikasi multi-frame. Untuk itu, maka kita memiliki dua pendekatan : *Belief Propagation* (BP), dan *Dynamic Programming* (DP) yang diambil untuk menyelesaikan permasalahan inferensi Bayesian di dalam jaringan.

*Belief Propagation* merupakan suatu algoritma iteratif yang berdasarkan *message passing*, dan untuk menyimpulkan state-state tersembunyi di dalam jaringan Markov, maka kita memerlukan iterasi dasar sebagai berikut :

$$m_{ij}(s_j) = \alpha \sum_{x_i} P_{ij}(s_i, s_j) P_i(x_i | s_i) \prod_{k \in N(i) \setminus j} m_{ki}(s_i)$$

$$b_i(s_i) = \alpha P_i(x_i | s_i) \prod_{k \in N(i)} m_{ki}(s_i)$$

dimana  $m_{ij}$  berhubungan dengan *message* yang dikirim dari node  $i$  ke node  $j$ , dan  $b_i$  merupakan *belief*,  $\alpha$  merupakan konstanta normalisasi,  $N(i) \setminus j$  bermaksud untuk semua node yang bertetangga dengan node  $i$ . Semua *message*  $m_{ij}(s_j)$  harus terlebih dahulu diinisialisasi dengan nilai 1 (satu) sebelum proses iterasi dilakukan.



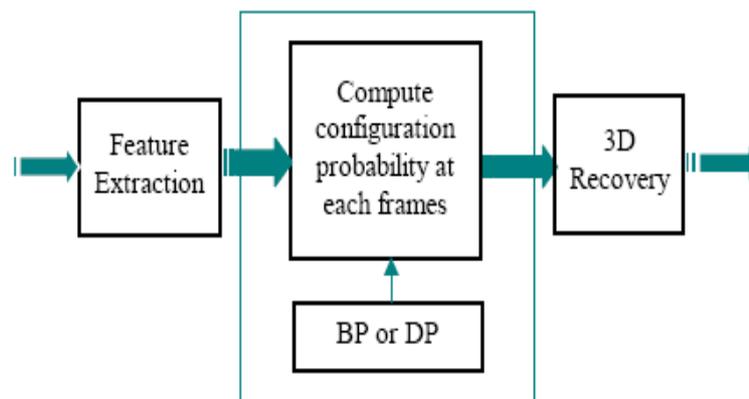
Gambar 7.6 Aproksimasi dengan rantai Markov dalam model jaringan Markov untuk aplikasi *multiple-frame*, sebagai pendeteksi gerakan tubuh bagian atas manusia

*Dynamic Programming*, untuk mengaplikasikan *Dynamic Programming* terlebih dahulu kita harus mengaproksimasi model jaringan Markov seperti yang diilustrasikan pada gambar 7.5, sebagai suatu mega-state  $M(t, S^t)$  seperti yang diilustrasikan pada gambar 7.6.  $S^t$  merepresentasikan suatu kombinasi berbagai posisi *body joint*,  $S^t = (s_1^t, s_2^t, \dots, s_N^t)$  pada frame ke- $t$ . Selanjutnya kita menghitung

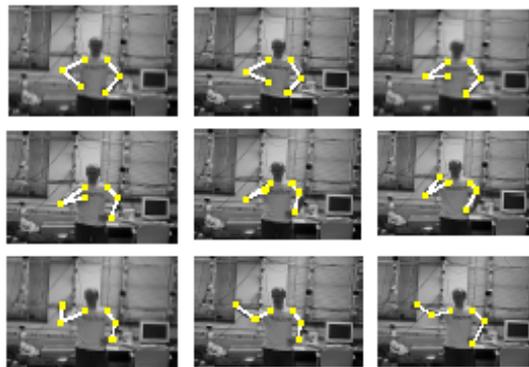
fungsi kemungkinan masing-masing mega-state,  $P(x|M(t, S^t))$ , dan berbagai probabilitas transisi di antara berbagai mega-state tersebut. Kita definisikan fungsi probabilitas Gaussian untuk jarak rata-rata antara seluruh posisi *body joint* di dalam frame  $t$  dan  $t + 1$  sebagai berikut :

$$P(M(t, S^t), M(t+1, S^{t+1})) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(S^t - S^{t+1})^T (S^t - S^{t+1})}{N\sigma^2}\right)$$

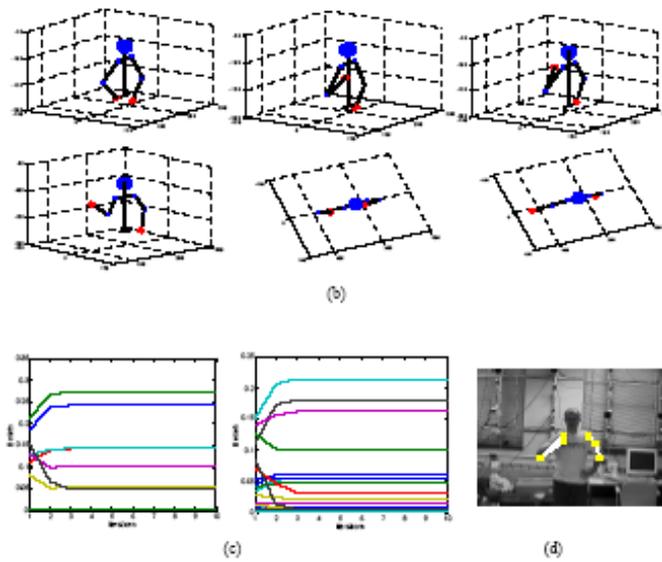
$$P(x|M(t, S^t)) = P(x_1, x_2, \dots, x_N | s_1^t, s_2^t, \dots, s_N^t)$$



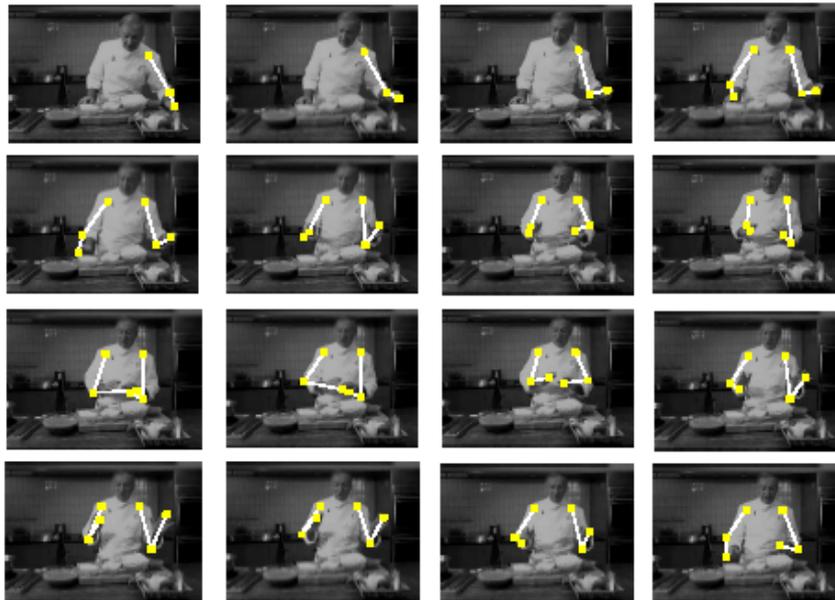
Gambar 7.7 Upper body motion detection system



(a)



Gambar 7.8 (a) Hasil inferensi pergerakan, (b) “stick figure” 3D yang telah direcover, (c) konvergensi *belief* dari 2 node state dengan 9 dan 88 kandidat state secara berturut-turut, dan (d) estimasi optimal *single-frame*, yang dikoreksi dengan batasan *multi-frame* pada (a)



Gambar 7.9 Hasil inferensi pergerakan, ada 2 error yang muncul pada baris ketiga yang disebabkan oleh jarak pergelangan tangan yang berdekatan dalam gambar, sehingga membingungkan pemrograman



Gambar 7.10 Error yang disebabkan oleh oklusi

# Bab 8

## PENGANTAR PENGOLAHAN CITRA VIDEO (VIP), SUATU PENDEKATAN DALAM ROBOT-VISION

*“For robot perception, video cameras are very valuable sensors, but the computer vision methods applied to extract information from camera images are usually computationally expensive. Integrating computer vision methods into a robot control architecture requires to balance exploitation of camera images with the need to preserve reactivity and robustness. We claim that better software support is needed in order to facilitate and simplify the application of computer vision and image processing methods on autonomous mobile robots. In particular, such support must address a simplified specification of image processing architectures, control and synchronization issues of image processing steps, and the integration of the image processing machinery into the overall robot control architecture. This paper introduces the video image processing (VIP) framework, a software framework for multithreaded control flow modeling in robot vision.”*

### 8.1 Pendahuluan

Persepsi terhadap lingkungan merupakan komponen kunci di dalam robot mobile otomatis, dengan memanfaatkan berbagai sensor untuk kegunaan ekstraksi berbagai informasi yang relevan. Untuk mendukung persepsi, maka sangat diperlukan satu atau lebih kamera video yang dipasangkan pada struktur robot mobile tersebut. Pada saat ini, kamera video merupakan perangkat yang murah dan mudah untuk diperoleh dan dapat menjadi sumber berbagai informasi yang lumayan, sementara metode yang diimplementasikan terhadap kamera video tersebut untuk mengekstraksi berbagai informasi yang diinginkan mungkin akan memakan biaya yang cukup mahal. Mengintegrasikan metode-metode yang tercakup di dalam wilayah komputer visi ke dalam arsitektur robot kontrol mungkin akan memerlukan perawatan yang besar dan usaha yang substansial dari para pemrogram sistem tersebut, secara spesial jika persepsi diimplementasikan di dalam lingkungan yang

lebih dinamis yang memerlukan tingkatan dan reaktivitas yang memadai. Tujuan dari bab ini adalah menceritakan proses atau tahapan yang diperlukan untuk mengimplementasikan berbagai pendekatan atau metode komputer visi ataupun pengolahan citra dalam fungsi persepsi pada robot mobile otomatis.

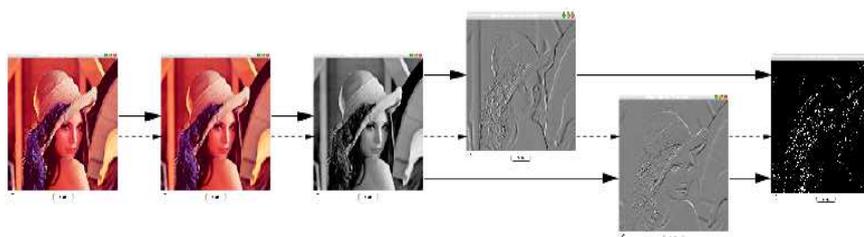
## 8.2 Pengolahan citra pada robot mobile

Sistem visi pada robot mobile otomatis harus menyatukan kebutuhan dan permintaan dua disiplin yang paling menantang :

- i. Komputer visi dan pengolahan citra
- ii. Robotika dan sistem tersembunyi (embedded system)

### 8.2.1 Komputer visi dan pengolahan citra

Konsep dasar komputer visi adalah sebagai aplikasi berbagai operator terhadap data citra, seperti operasi logika dan operasi aritmetika, konversi warna, fungsi morfologi (erosi dan dilatasi), fungsi filter, transformasi linier (Fourier atau transformasi wavelet), dan lain-lain.



Gambar 8.1 Citra lena yang melalui beberapa tahapan pengolahan citra, original, blurred, grayed, convolved, dan edge-detection

Operasi transformasi sering membutuhkan lebih daripada satu citra input untuk menghasilkan sebuah citra yang baru, untuk contoh pada pendeteksi edge Canny, (Canny, J. F., 1986), pendeteksi edge Canny membutuhkan dua buah citra yang diperolehnya dari hasil konvolusi dengan menggunakan operator Sobel vertikal dan horizontal, secara berturut-turut. Berbagai operator yang lain, seperti pada aliran optis yang memerlukan sederetan citra input yang diperolehnya dalam berbagai waktu yang berbeda (Horn, B. K. P. & Schunck, B. G., 1980).

Secara prinsip, suatu operasi citra dapat dipandang sebagai suatu pemetaan dari satu atau lebih citra menjadi sebuah citra output. Sedangkan hasil dari operasi pemetaan citra tidak selalu harus berupa sebuah citra yang lain, tetapi mungkin saja berupa bentuk lain dari struktur data pada citra pertama yang dipetakan, seperti histogram warna, atau berbagai operasi statistik lainnya.

Di dalam kasus ini, maka definisi suatu filter dapat diperluas menjadi suatu pemetaan satu atau lebih citra input menjadi sebuah citra yang baru, atau satu atau lebih nilai-nilai pengklasifikasian yang diasosiasikan terhadap citra.

### **8.2.2 Robot visi**

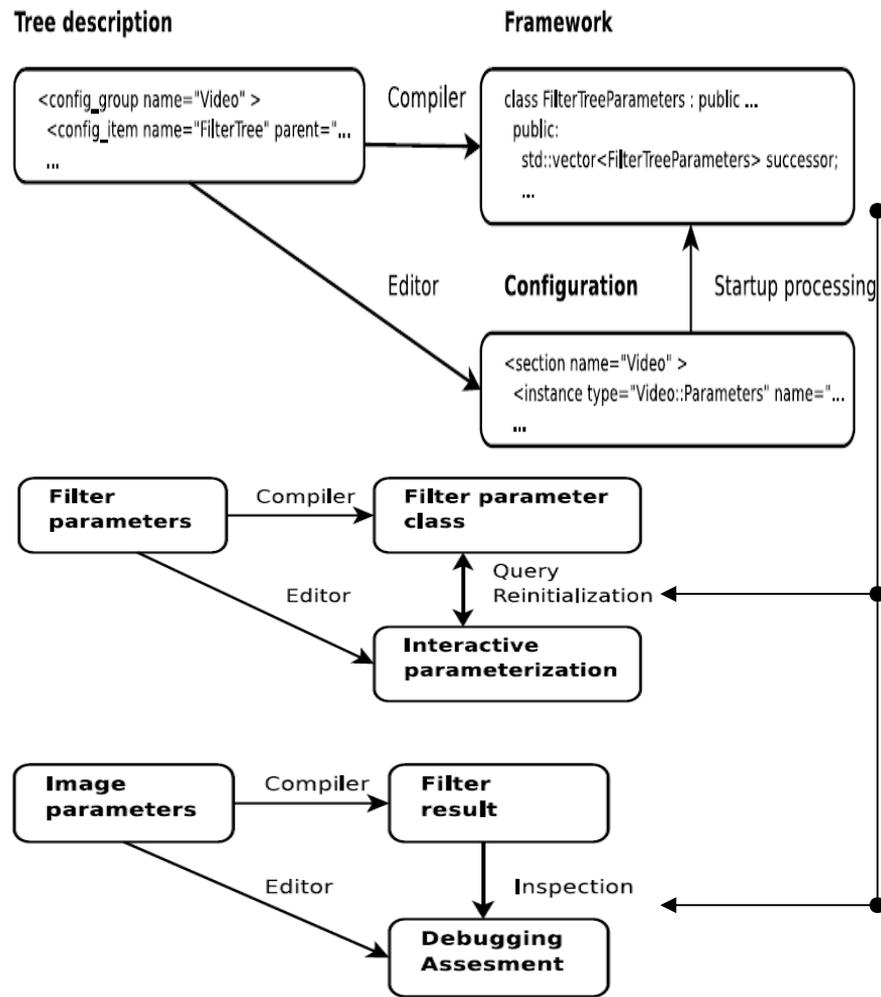
Melakukan berbagai operasi yang telah dipaparkan sebelumnya, dengan data yang berbentuk sebarisan citra video yang disuplai oleh satu atau lebih kamera yang terdapat pada struktur robot mobile otomatis yang selanjutnya diproses menjadi satu citra tunggal dengan informasi atau interpretasi citra yang dimaksudkan.

*Timeliness Constraints.* Robot terlebih dahulu harus disituasikan di dalam dunia fisik. Untuk berbagai tugas seperti pendeteksian rintangan dan *tracking* objek, operasi pengolahan citra harus dikalkulasikan sesering mungkin per detik dan lebih baiknya pada *full frame rate*, yang dalam hal ini sebesar 30Hz.

*Fixed Frame Rate Image Streams.* Citra yang baru biasanya tiba pada suatu *fixed frame rate*. Informasi yang diperoleh mungkin saja secara cepat berkurang dalam suatu lingkungan dinamis, dan hal ini menyisakan tugas untuk sensor sebagai pentrigger model-model evaluasi yang diperlukan.

*Parameterization.* Melakukan parameterisasi untuk satu atau lebih operator pengolahan citra yang dilibatkan, untuk menghasilkan citra dengan kualitas yang diharapkan. Contoh, lebar filter Gauss, jumlah *bucket* untuk orientasi histogram. Kalibrasi dan optimasi parameter juga merupakan bagian yang penting dari pengembangan proses, dan juga, pengkonfigurasi grafik filter yang harus diubah secara frekuensi selama proses pengembangan, yang secara signifikan difasilitasi oleh lingkungan yang fleksibel dan *configurable* untuk sistem visi robot.

*Development Model.* Robot visi dilakukan pada *image stream* secara langsung, yang direkam oleh suatu platform robot yang dapat dipindah-pindahkan, dan dialamatkan di dalam satu atau lebih model pengembangan.



Gambar 8.2 Aplikasi parameter dan konfigurasi *management toolkit* untuk sistem robot visi

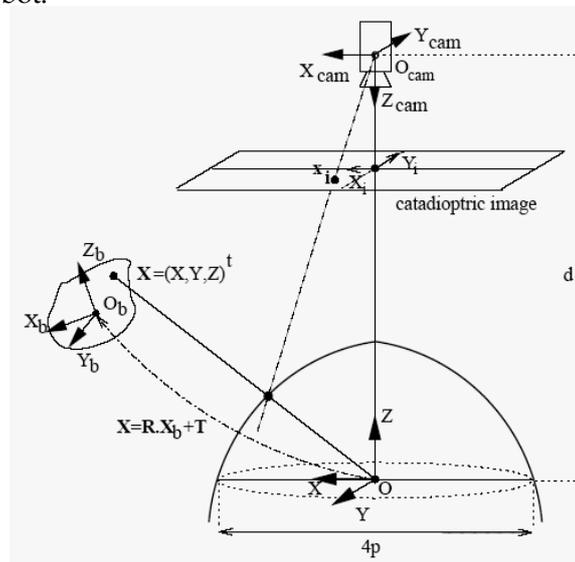
# Bab 9

## VISUAL SERVOING

### 9.1 Pendahuluan

Mesin visi memberikan suatu pengukuran berbagai keadaan dunia tanpa perlu melalui suatu "kontak fisik", memperluas kemampuan dan jangkauan robot untuk beroperasi lebih optimal di dalam wilayah dan keadaan kerjanya, yang mungkin tidak terkontrol secara akurat. Aplikasi *visual-servoing* dapat menjadi salah satu solusi untuk menjangkau segalanya dengan kemampuan penglihatan yang lebih luas. Jadi tujuan dari bab ini adalah membahas secara mendetail konsep dan pengimplementasian dari aplikasi *visual-servoing*, keuntungan, dan hambatan dari penerapannya.

Satu cara yang paling efektif untuk memperluas bidang pandang sebuah kamera adalah dengan memanfaatkan fungsi dari cermin (mirror). Pendekatan-pendekatan yang umum telah digunakan, mengkombinasikan satu atau lebih cermin dengan sistem imaging konvensional yang diistilahkan sebagai formasi citra *catadioptric*. Sistem sentral *catadioptric* memiliki beberapa kelebihan untuk penerapannya di kebanyakan aplikasi, dikarenakan sistem tersebut mengkombinasikan dua fitur penting, (1) pusat proyeksi tunggal, dan (2) jangkauan bidang pandangan. Penerapan *visual servoing* pada sensor sering muncul dalam aplikasi navigasi robot.



Gambar 9.1 Central Catadioptric Projection

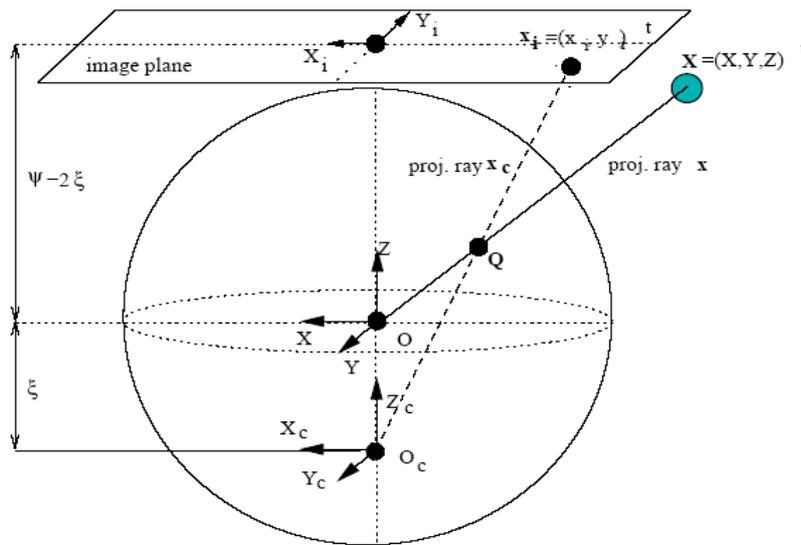
Untuk sistem *catadioptric* sentral generik, kita perkenalkan suatu matriks Jacobian  $J$  yang merupakan penurunan dari fungsi pemetaan *catadioptric* sentral yang dibahas pada penelitian yang dilakukan oleh C. Geyer dan K. Daniilidis, "A Unifying Theory for Central Panoramic System and Practical Implications" pada tahun 2000. Menurut teori tersebut, imaging *catadioptric* sentral dapat dimodelkan oleh suatu fungsi generik  $f_1$  dengan tipe sensor dan bentuk cermin dideskripsikan oleh parameter  $\xi$ . Untuk kasus-kasus tertentu seperti pada kamera konvensional, parameter  $\xi$  diberikan nilai null. Oleh karena itu, dengan mengasumsikan  $\xi = 0$ , matriks Jacobian umum  $J_g$  dapat diturunkan menjadi matriks  $J_p$ , seperti yang dipaparkan oleh B. Espiau, F. Chaumette, dan P. Rives, dalam "A New Approach to Visual Servoing in Robotics" pada tahun 1992. Lebih-lebih matriks Jacobian dapat didekomposisi dalam hasil produk dua buah matriks  $J_c$  dan  $J_p$  ( $J_g = J_c \cdot J_p$ ), di mana  $J_c$  merupakan sebuah matriks  $2 \times 2$ .

## 9.2 Pemodelan formasi citra *catadioptric* sentral

Realisasi visi omni-direksional *catadioptric* mengkombinasikan permukaan reflektif dan berbagai lensa. Baker di dalam penelitiannya telah membangun berbagai kelas sistem *catadioptric* dan menjelaskan tentang batasan titik pandang tetap. Batasan titik pandang tetap dirumuskan untuk memastikan keadaan sensor-sensor visual supaya membatasi pengukurannya hanya pada intensitas cahaya yang melalui sebuah titik tunggal dalam ruang 3D. Suatu pusat proyeksi yang unik merupakan kondisi yang memadai untuk menghasilkan citra perpektif yang benar secara geometris, dan juga sebagai eksistensi sifat-sifat geometri epipolar terhadap sensor-sensor yang bergerak dan struktur pandang yang lebih bebas. Sistem *catadioptric* sentral dapat dibangun dengan mengkombinasikan sebuah cermin parabolis dengan kamera orthografis atau hiperbolis, dan cermin planar atau eliptis dengan kamera perspektif.

	Mirror Surface	$\xi$	$\psi$
Parabolic	$\sqrt{X^2 + Y^2 + Z^2} = 2p - Z$	1	$1 + 2p$
Hyperbolic	$\frac{(Z - \frac{d}{2})^2}{(\frac{1}{2}(\sqrt{d^2 + 4p^2} - 2p))^2} - \frac{X^2 + Y^2}{p(\sqrt{d^2 + 4p^2} - 2p)} = 1$	$\frac{d}{\sqrt{d^2 + 4p^2}}$	$\frac{d + 2p}{\sqrt{d^2 + 4p^2}}$
Elliptical	$\frac{(Z - \frac{d}{2})^2}{(\frac{1}{2}(\sqrt{d^2 + 4p^2} + 2p))^2} + \frac{X^2 + Y^2}{p(\sqrt{d^2 + 4p^2} + 2p)} = 1$	$\frac{d}{\sqrt{d^2 + 4p^2}}$	$\frac{d - 2p}{\sqrt{d^2 + 4p^2}}$
Planar	$Z = \frac{d}{2}$	0	1

Tabel 9.1 Kolom 1 : permukaan reflektif untuk berbagai kasus pencitraan panorama terpusat, kolom 2 dan 3 : berturut-turut, parameter  $\xi$  dan  $\psi$  untuk pemodelan *catadioptric* sentral yang umum



Gambar 9.2 Pemodelan formasi pencitraan *catadioptric* sentral

Gambar 9.2 merupakan skema sistem *catadioptric* yang mengkombinasikan suatu permukaan reflektif hiperbolis dengan sebuah perspektif kamera. Pertimbangkan sistem koordinat  $R$  dan  $R_{cam}$  yang diasosiasikan secara berturut-turut terhadap kamera perspektif dan cermin, maka sumbu hiperbola akan koinciden dengan sumbu  $Z$  dari  $R$ , dan titik-titik foci-nya berkoinciden dengan  $O$  dan  $O_{cam}$ , secara berturut-turut merupakan titik-titik origin dari  $R$  dan  $R_{cam}$ . Latus rectum dari permukaan hiperbolis adalah  $4p$  dan jarak di antara titik-titik foci sebesar  $d$ . Sinar cahaya yang insiden dengan  $O$ , titik focal dalam, akan direfleksikan menjadi titik focal luar,  $O_{cam}$ . Jika pusat proyeksi dari kamera perspektif koinciden terhadap  $O_{cam}$ , maka sinar cahaya yang berhasil ditangkap akan melalui fokus dalam permukaan hiperbolis. Titik pandangan efektif dari citra yang diperoleh adalah  $O$  dan unik. Citra *catadioptric* eliptikal diperoleh dengan mengkombinasikan sebuah cermin elips dengan sebuah kamera perspektif dengan cara yang sama. Dalam situasi parabolis, sebuah cermin parabolis ditempatkan sebagai sumbu  $Z$ , dengan finite-fokus-riil-nya berkoinciden dengan  $O$ . Sinar cahaya yang berinsiden dengan  $O$  direfleksikan menjadi sinar yang paralel terhadap sumbu  $Z$  yang ditangkap oleh kamera orthografis dengan bidang citra yang tegak lurus terhadap sumbu  $Z$ . Titik pandangan efektif dari citra yang diperoleh adalah  $O$  dan unik. Pusat pemroyeksian efektif berada di belakang cermin dalam wilayah garis yang tegak lurus yang membentang melewati pusat kamera, dan jaraknya ke pusat kamera adalah dua kali jarak antara cermin planar dengan kamera.

Asumsikan suatu titik 3D,  $\mathbf{X} = (X, Y, Z)^t$ , diproyeksikan menjadi  $\mathbf{x}_i = (x_i, y_i)^t$  di dalam bidang citra *catadioptric*. Maka kita dapat menunjukkan pemroyeksian panorama sentralnya adalah isomorfis terhadap suatu pemetaan dari sebuah sphere ke suatu bidang. Pertimbangkan skema yang ditunjukkan pada gambar 9.2, dengan suatu unit sphere yang berpusat di dalam titik pandang efektif  $O$ , titik  $O_c$  dengan koordinat  $(0, 0, \xi)^t$  dan bidang  $Z = \psi - 2\xi$  yang ortogonal terhadap sumbu  $Z$ . Keduanya,  $\xi$  dan  $\psi$  merupakan fungsi parameter cermin  $d$  dan  $p$ . Sinar proyeksi  $\mathbf{x}$  yang melalui  $\mathbf{X}$  yang mengiris permukaan speris dalam  $Q = \left( \frac{x}{p}, \frac{y}{p}, \frac{z}{p} \right)^t$  dengan  $p = \sqrt{X^2 + Y^2 + Z^2}$ . Sinar proyeksi kedua dapat didefinisikan sebagai gabungan interseksi  $Q$  dengan  $O_c$ . Interseksi  $\mathbf{x}_i$  dari sinar perspektif  $\mathbf{x}_c$  dengan bidang  $Z = \psi - 2\xi$  merupakan citra *catadioptric* dari titik asal  $\mathbf{X}$ . Pusat pencitraan *catadioptric* dapat dimodelkan dengan memproyeksikan pemandangan dalam permukaan speris dan kemudian memproyeksikan kembali titik-titik tersebut dalam bidang citra dari pusat proyeksi baru  $O_c$ . Jika permukaan reflektif merupakan parabola, maka  $\xi = 1$ , dan bentuk dari pemroyeksian berikutnya akan berupa pemroyeksian grafik stereo. Untuk cermin hiperbola dan elips, pusat reproyeksi  $O_c$  berada di dalam sphere pada sumbu negatif  $Z$ . Cermin planar merupakan suatu kasus pengecualian, di mana pusat pencitraan *catadioptric* dengan  $\xi = 0$  dan  $O_c$  berkoinsiden dengan titik pandang efektif  $O$ . Catat bahwa, sebuah sensor *catadioptric* dengan cermin planar ekuivalen terhadap sebuah kamera perpektif konvensional dengan inversi tanda pada sumbu  $Y$ .

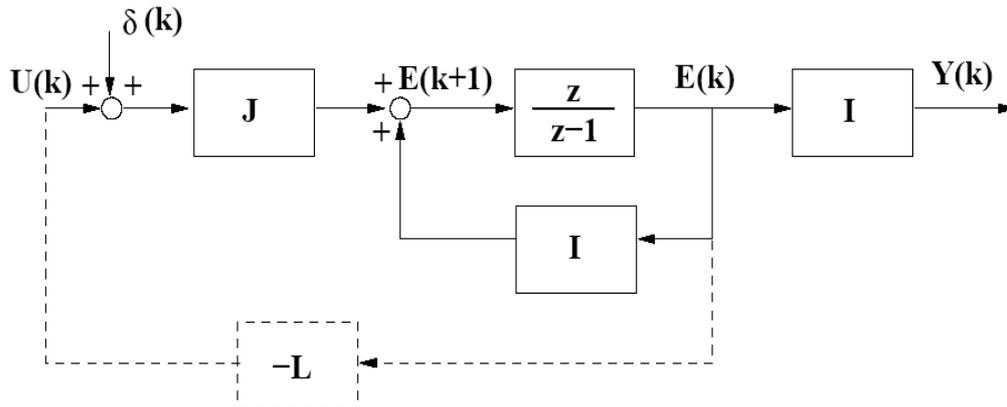
$$f_i(X) = \left( \begin{array}{c} \frac{f_x(\psi - \xi)X}{Z + \xi\sqrt{X^2 + Y^2 + Z^2}} - c_x, \\ \frac{f_y(\psi - \xi)Y}{Z + \xi\sqrt{X^2 + Y^2 + Z^2}} - c_y \end{array} \right)^t$$

Pada kasus, di mana kalibrasi sensor diketahui, fungsi  $f_i$  dapat disederhanakan dengan memisalkan  $f_x(\psi - \xi) = f_y(\psi - \xi) = 1$  dan  $c_x = c_y = 0$ . Maka dapat diasumsikan tanpa kehilangan keumuman bahwa fungsi pemetaan dapat diberikan sebagai :

$$f_i(X) = \left( \begin{array}{c} \frac{X}{Z + \xi\sqrt{X^2 + Y^2 + Z^2}}_x, \\ \frac{Y}{Z + \xi\sqrt{X^2 + Y^2 + Z^2}}_y \end{array} \right)^t$$

9.3 Tracking dan kontrol

Gambar 9.1 menunjukkan perpindahan sebuah objek yang rigid yang diamati oleh sebuah sensor *catadioptric* sentral. Frame referensi  $R_b$  ditambahkan pada tubuh objek yang bergerak,  $R$  merupakan matriks rotasi di antara  $R_b$  dan  $R$  dengan  $T$  merupakan posisi  $O_b$  di dalam koordinat sensor, maka tujuan kita di sini adalah untuk mengestimasi pose dari objek yang bergerak tersebut, dengan koordinat-koordinat  $\{X_b^1, X_b^2, \dots, X_b^n\}$  dari sekumpulan  $n$  titik dari objek tersebut.



Gambar 9.3 Estimasi pose iteratif,  $I$  merupakan matriks identitas  $2n \times 2n$  dan garis yang putus-putus berkorespondensi dengan perulangan timbal-balik

Andaikan  $X_b$  mewakili titik generik model objek, jika pose  $(R, T)$  diketahui maka posisi titik 3D di dalam koordinat sensor adalah  $X = RX_b + T$ . Dari persamaan

$$f_i(X) = \begin{pmatrix} \frac{X}{Z + \xi \sqrt{X^2 + Y^2 + Z^2}} \\ \frac{Y}{Z + \xi \sqrt{X^2 + Y^2 + Z^2}} \end{pmatrix}_i$$

maka titik  $X$  diproyeksikan di dalam  $x_i = f_i(X)$  pada bidang citra *catadioptric*. Perpindahan sebuah objek mengindikasikan perubahan pose dari objek tersebut yang dapat dideskripsikan sebagai pasangan kinematis  $\delta = (\omega, v)^t$ . Pertimbangkan matriks  $3 \times 6$   $J_m = [\bar{X} \mid I]$  di mana  $\bar{X}$  merupakan matriks yang tidak simetris dari  $X$  dan  $I$  merupakan matriks identitas  $3 \times 3$ . Kecepatan 3D dari titik  $X$  berkaitan dengan

perpindahan objek yang setara dengan  $X = J_m \delta$ . Lebih-lebih jika  $J_i$  merupakan matriks Jacobian dari fungsi  $f_i$  dan  $J_g = J_i J_m$  maka kecepatan terkait dalam bidang citra *catadioptric* adalah  $x_i = J_g \delta$ .

$$E = \begin{bmatrix} x_i^1 - \hat{x}_i^1 \\ x_i^2 - \hat{x}_i^2 \\ \vdots \\ x_i^n - \hat{x}_i^n \end{bmatrix} \approx \begin{bmatrix} J_g^1 \\ J_g^2 \\ \vdots \\ J_g^n \end{bmatrix} \hat{\delta} = J \hat{\delta}$$

Andaikan  $\hat{s} = \{\hat{x}_i^1, \hat{x}_i^2, \dots, \hat{x}_i^n\}$  menjadi sekumpulan titik model yang diproyeksikan di dalam citra berdasarkan estimasi pose  $(\hat{R}, \hat{T})$  saat sekarang, dan  $s = \{x_i^1, x_i^2, \dots, x_i^n\}$  merupakan posisi sebenarnya. Vektor  $E$  didefinisikan sebagai  $E = s - \hat{s}$  dan bergantung pada error estimasi pose yang dideskripsikan oleh padangan kinematis  $\delta$ . Dari diskusi di atas dapat dikatakan bahwa  $(x_i^j - \hat{x}_i^j) \approx J_g^j \delta$  dengan  $j = 1, 2, \dots, n$  dan  $J_g^j$  merupakan matriks Jacobian dari  $J_g$  yang dievaluasikan pada titik model ke-  $j$ . Persamaan

$$E = \begin{bmatrix} x_i^1 - \hat{x}_i^1 \\ x_i^2 - \hat{x}_i^2 \\ \vdots \\ x_i^n - \hat{x}_i^n \end{bmatrix} \approx \begin{bmatrix} J_g^1 \\ J_g^2 \\ \vdots \\ J_g^n \end{bmatrix} \hat{\delta} = J \hat{\delta}$$

membentuk hubungan antara error citra yang diukur  $E$  dan error  $\delta$  pada estimasi pose tubuh objek yang rigid.

#### 9.4 Matriks Jacobian *catadioptric*

Pertimbangkan fungsi pemetaan *catadioptric* sentral  $f_i$  yang memetakan titik koordinat 3D,  $X$ , pada koordinat citra  $x_i$ . Matriks Jacobian terkait dihasilkan dengan mendiferensialkan fungsi

$$f_i(X) = \begin{pmatrix} \frac{X}{Z + \xi \sqrt{X^2 + Y^2 + Z^2}} \\ \frac{Y}{Z + \xi \sqrt{X^2 + Y^2 + Z^2}} \end{pmatrix}_x^t$$

Hasil yang dicapai adalah

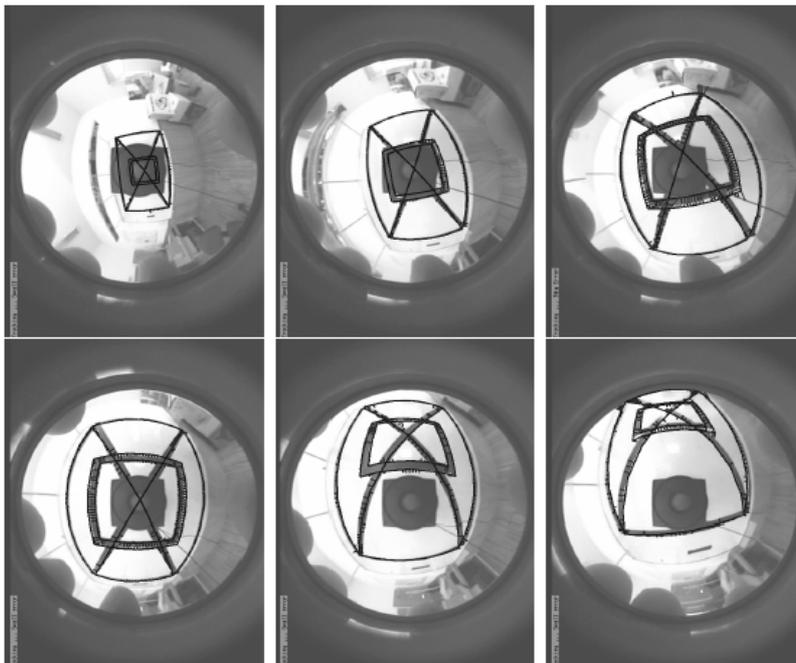
$$\mathbf{J}_i = \frac{1}{\rho(Z + \xi\rho)^2} \begin{bmatrix} \rho Z + \xi(Y^2 + Z^2) & -\xi XY & -X(\rho + \xi Z) \\ \xi XY & -(\rho Z + \xi(X^2 + Z^2)) & Y(\rho + \xi Z) \end{bmatrix}$$

Di mana  $\rho = \sqrt{X^2 + Y^2 + Z^2}$ .

$$\mathbf{J}_g = \begin{bmatrix} x_i y_i & \frac{(1+x_i^2)\Upsilon - y_i^2 \xi}{\Upsilon + \xi} & y_i & \frac{1+x_i^2(1-\xi(\Upsilon+\xi))+y_i^2}{\rho(\Upsilon+\xi)} \\ \frac{(1+y_i^2)\Upsilon - x_i^2 \xi}{\Upsilon + \xi} & x_i y_i & -x_i & -\frac{x_i y_i \xi}{\rho} \\ & & & \frac{x_i y_i \xi}{\rho} & -\frac{x_i \Upsilon}{\rho} \\ & & & \frac{1+x_i^2+y_i^2(1-\xi(\Upsilon+\xi))}{\rho(\Upsilon+\xi)} & -\frac{y_i \Upsilon}{\rho} \end{bmatrix}$$

$$\mathbf{J}_i = \underbrace{\begin{bmatrix} \frac{Z(\rho Z + \xi(Y^2 + Z^2))}{\rho(Z + \xi\rho)^2} & \frac{\xi XY Z}{\rho(Z + \xi\rho)^2} \\ \frac{\xi XY Z}{\rho(Z + \xi\rho)^2} & \frac{Z(\rho Z + \xi(X^2 + Z^2))}{\rho(Z + \xi\rho)^2} \end{bmatrix}}_{\mathbf{J}_c} \begin{bmatrix} \frac{1}{Z} & 0 & -\frac{X}{Z^2} \\ 0 & -\frac{1}{Z} & \frac{Y}{Z^2} \end{bmatrix}$$

$$\mathbf{J} = \underbrace{\begin{bmatrix} \mathbf{J}_c^1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_c^2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}_c^n \end{bmatrix}}_{\mathbf{C}} \underbrace{\begin{bmatrix} \mathbf{J}_p^1 \\ \mathbf{J}_p^2 \\ \vdots \\ \mathbf{J}_p^n \end{bmatrix}}_{\mathbf{P}}$$



Gambar 9.4 Dua baris hasil tracking. Baris pertama, objek berpindah menuju ke kamera. Pada baris kedua, objek yang sama berpindah sejajar dengan bidang citra

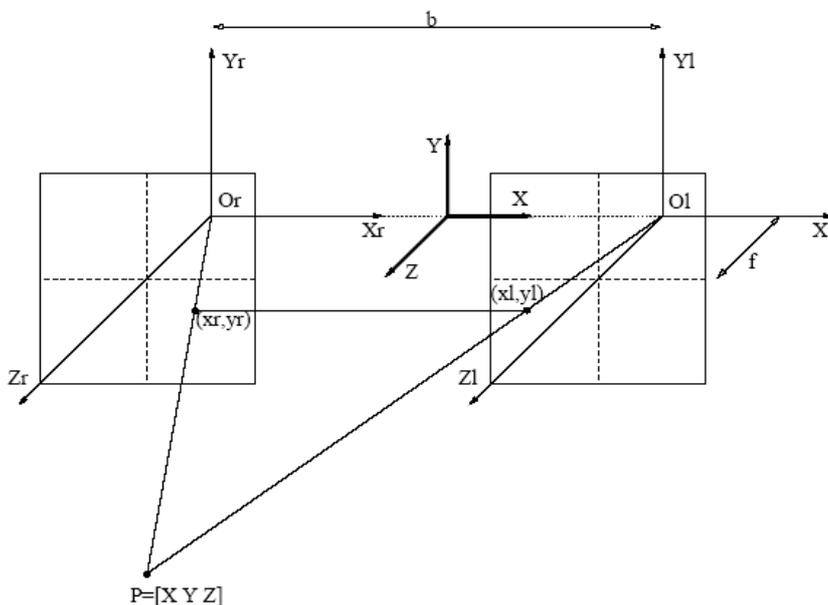
# Bab 10

## ESTIMASI PERGERAKAN 3D DARI CITRA STEREO - FORMULASI DISKRIT DAN DIFERENSIAL

### 10.1 Pendahuluan

Metode untuk penyelesaian masalah gerak (pergerakan) dapat diklasifikasikan menjadi metode kontinu (diferensial) dan metode diskrit. Kedua metode menggunakan sebarisan citra yang bersifat temporal. Metode-metode yang terdahulu sering dikatakan diskrit dikarenakan aplikasi metode tersebut dalam memanfaatkan fitur-fitur objek dan kaitan fitur-fitur tersebut terhadap waktu yang diasumsikan telah diketahui, sebaliknya, metode diferensial menggunakan aliran optis yang kontinu.

Pada bab ini kita akan membahas analisis performance dari kedua metode tersebut.



Gambar 10.1 Sistem koordinat stereo

## 10.2 Estimasi pergerakan

Titik-titik di dalam ruang 3D akan direpresentasikan oleh koordinat vektor  $P = [X \ Y \ Z]^T$ . Posisi kamera dengan kamera lainnya adalah paralel dan dipisahkan oleh *baseline*  $b$ . Vektor aliran dalam bidang citra direpresentasikan oleh  $v^l = (v_x^l, v_y^l)$  untuk citra kiri dan  $v^r = (v_x^r, v_y^r)$  untuk citra kanan.

Kita pertimbangkan suatu pergerakan yang rigid. Andaikan  $V$  merupakan kecepatan total 3D dari sebuah titik  $P$ . Sebarang pergerakan yang rigid dapat diekspresikan oleh suatu komponen translasi yang diberikan sebagai  $t = [t_x \ t_y \ t_z]^T$  dan komponen rotasi  $\Omega = [\Omega_x \ \Omega_y \ \Omega_z]^T$  dengan kecepatan 3D yang diberikan sebagai  $V = \Omega \times P + t$ . Jika kita lebih perluas persamaan kecepatan tersebut, maka kita akan memperoleh komponen ketiga sebagai kecepatan suatu titik sepanjang arah  $Z$ ,  $V_Z$ , di mana  $V_Z = t_z + \Omega_x Y - \Omega_y X$ .

Di dalam pendekatan diskrit, dua buah titik yang berada di dalam ruang sangat berkaitan dengan suatu transformasi linier yang dikomposisikan oleh sebuah matriks rotasi dan vektor translasi, sehingga,  $P' = R \bullet P + T$ , di mana  $P$  merupakan sebuah titik 3D pada waktu  $t$  dan  $P'$  merupakan titik yang sama pada waktu  $t'$ .

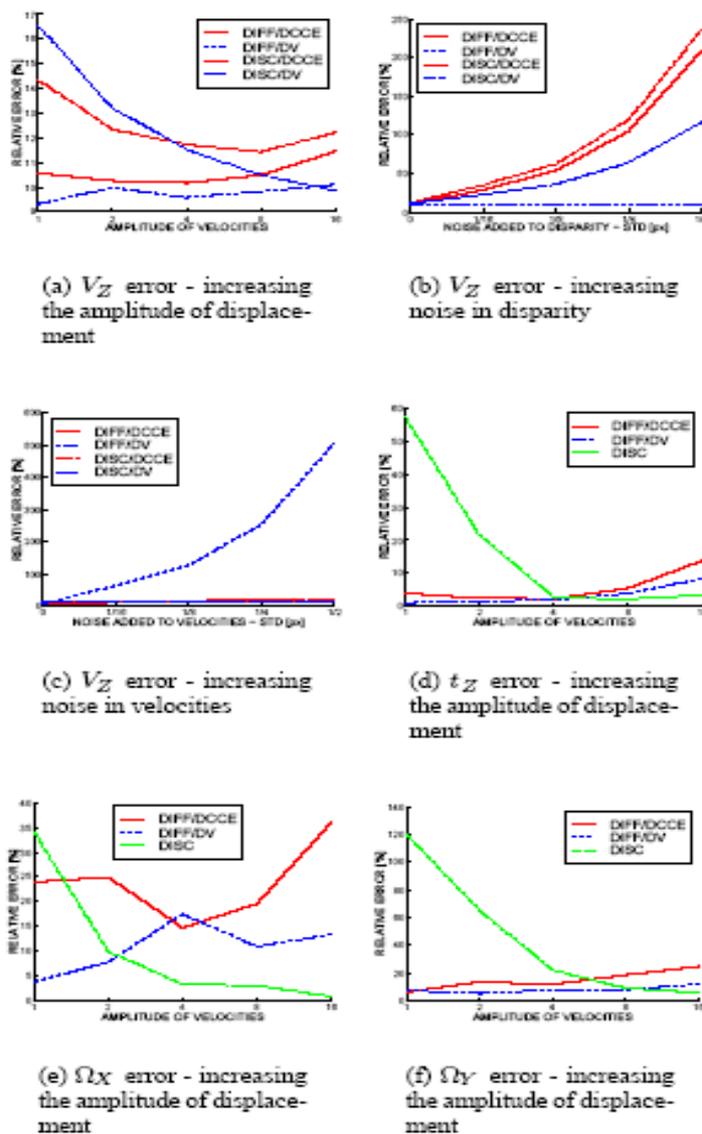
## 10.3 Pendekatan diferensial

$V_Z$  - *Depth constraint*, perubahan kedalaman sebuah titik pada suatu waktu secara langsung dikaitkan pada kecepatannya di dalam ruang 3D.

kedalaman sebuah titik pada waktu  $t'$  merupakan kedalaman titik pada waktu  $t$  ditambah suatu *displacement* pada arah  $Z$  -  $V_Z$ , atau dapat dieskpresikan sebagai :

$$V_Z = Z_t + Z_x \cdot v_x + Z_y \cdot v_y$$

dimana  $Z(x, y, t)$  merupakan suatu kedalaman pada waktu yang diberikan  $t$ ,  $Z_x, Z_y$  dan  $Z_t$  merupakan turunan spatial-temporal dari  $Z(x, y, t)$ , dan  $v_x$  dan  $v_y$  merupakan komponen-komponen aliran optis. Formulasi tersebut sering juga diistilahkan sebagai DCCE linier, *Depth Change Constraint Equation*.



Gambar 10.2 Analisis error dalam citra sintesis

$V_Z$  - *Binocular Flow Constraint*, koordinat  $x$  pada  $P$  adalah  $x_l$  pada citra kiri dan  $x_r$  pada citra kanan. Proyeksi titik  $P$  dalam masing-masing bidang citra  $(x_l, y_l, f)$  dan  $(x_r, y_r, f)$  dan pusat optis  $O_l$  dan  $O_r$  mendefinisikan dua buah segitiga yang sama, sehingga hubungan kesemuanya dapat kita formulasikan sebagai :

$$\frac{Z}{b} = \frac{Z - f}{b - (x^r - x^l)}$$

Bila kita hitung persamaan diferensial temporal persamaan di atas, maka kita akan memperoleh :

$$V_Z = -\frac{bf}{(x^r - x^l)^2} \cdot (v_x^r - v_x^l) = -\frac{Z^2}{bf} \Delta v_x$$

Parameter pergerakan, *Depth Constraint*, bila kita lakukan penukaran pada persamaan

$$V_Z = Z_t + Z_x \cdot v_x + Z_y \cdot v_y$$

kecepatan pada citra dapat ditentukan oleh suatu relasi yang diketahui dengan parameter pergerakan, sehingga kita peroleh :

$$-Z_t = \begin{bmatrix} f \frac{Z_x}{Z} \\ f \frac{Z_y}{Z} \\ -\frac{Z+xZ_x+yZ_y}{Z} \\ -fZ_y - \frac{y}{f}(Z+xZ_x+yZ_y) \\ fZ_x + \frac{x}{f}(Z+xZ_x+yZ_y) \\ xZ_y - yZ_x \end{bmatrix}^T \cdot \vec{\phi}$$

di mana  $\vec{\phi}$  merupakan vektor dengan enam parameter pergerakan yang akan diestimasi.

Parameter pergerakan, *Binocular Flow Constraint*, kecepatan 3D dalam arah  $Z$  dapat diekspresikan sebagai suatu persamaan linier pada tiga dari enam parameter pergerakan yang dapat disubstitusikan di dalam aliran binokular, sehingga dapat kita formulasi sebagai :

$$t_Z + \Omega_X Y - \Omega_Y X = -\frac{Z^2}{bf} \Delta v_x \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} \Delta v_x \\ bf \end{bmatrix} = \begin{bmatrix} -1 & -y & x \\ Z^2 & fZ & Z \end{bmatrix} \begin{bmatrix} t_Z \\ \Omega_X \\ \Omega_Y \end{bmatrix}$$

di mana kita dapat gantikan koordinat titik-titik 3D  $X$  dan  $Y$  dengan persamaan inversi proyeksi perspektifnya.

#### 10.4 Formulasi diskrit

$V_z$  - DCCE diskrit, persamaan DCCE di dalam formulasi diskrit adalah :

$$V_z = Z_t + Z_x \cdot \Delta_x + Z_y \cdot \Delta_y$$

di mana  $Z_t = Z(x, y, t') - Z(x, y, t)$ .

Di dalam formulasi diskrit, persamaan DCCE, kecepatan citra digantikan oleh selisih terhingga dari koordinat titik-titik citra.

$V_z$  - *Binocular Flow*, persamaan aliran binokular diskrit diberikan sebagai :

$$V_z \approx \frac{ZZ'}{bf} (d' - d)$$

Parameter pergerakan, formulasi diskrit, berdasarkan ekspresi  $P' = R \cdot P + T$  dan perluasannya, maka dapat kita bentuk suatu relasi antara parameter pergerakan diskrit ( $R$  dan  $T$ ) dan titik-titik 3D  $P$  dan  $P'$ . Untuk mentransformasikan persamaan menjadi suatu sistem yang dapat ditentukan, maka kita harus menentukan kembali isi dari matriks  $R$  dan vektor  $T$ . Untuk memperbaiki parameter pergerakan secara rotasi, maka terlebih dahulu kita pertimbangkan parameter pergerakan yang konstan setiap waktu, sehingga aproksimasi instan yang dihasilkan oleh :

$$R \approx \begin{bmatrix} 1 & -\Omega_Z \Delta t & \Omega_Y \Delta t \\ \Omega_Z \Delta t & 1 & -\Omega_X \Delta t \\ -\Omega_Y \Delta t & \Omega_X \Delta t & 1 \end{bmatrix}$$

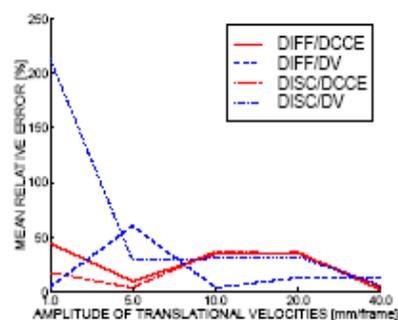
dan parameter pergerakan secara translasi adalah :

$$\vec{t} = V^{-1} \cdot T$$

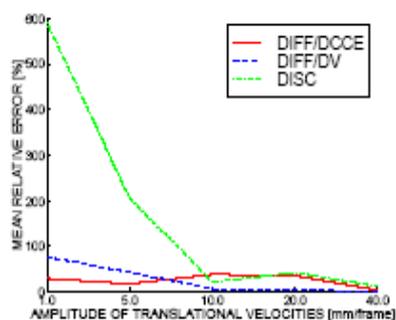
di mana  $V$  merupakan matriks yang diperoleh dari suatu bentuk tertutup dari *trajectory* titik di dalam pergerakan yang rigid.



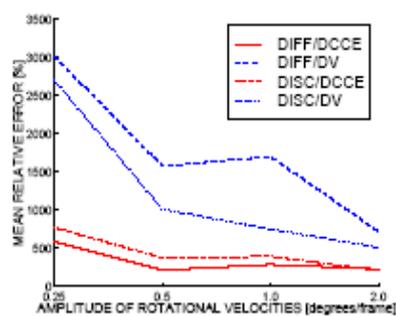
Gambar 10.3 Citra riil dan peta disparitas terkait



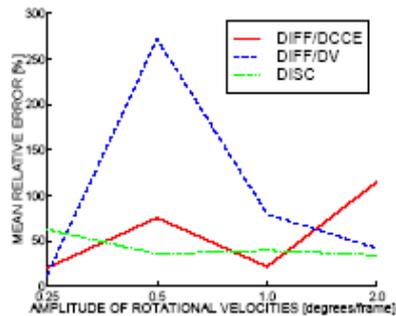
(a) Relative error of  $V_Z$  with increasing displacement



(b) Relative error of  $t_Z$  with increasing displacement



(c) Relative error of  $V_Z$  with increasing rotation



(d) Relative error of  $\Omega_\gamma$  with increasing rotation

Gambar 10.4 Analisis error di dalam citra riil

# Bab 11

## ESTIMASI POSE UNTUK SISTEM CATADIOPTIC SENTRAL : PENDEKATAN ANALITIS

### 11.1 Pendahuluan

Estimasi pose merupakan salah satu isu yang sangat penting dalam bidang komputer visi, dari aplikasi untuk navigasi robot hingga rekonstruksi citra.

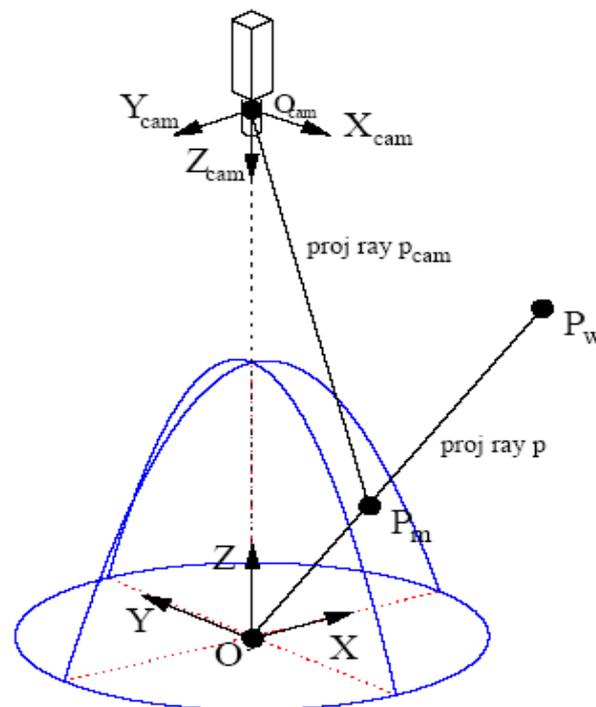
Pada umumnya sistem estimasi pose didasari pada kamera perspektif standar, yang telah diakui memiliki keunggulan di dalam menjangkau hampir keseluruhan sifat-sifat geometris. Di balik keunggulan tersebut, sistem tersebut memiliki juga kelemahan terutama pada jangkauan sudut pandang yang sangat terbatas, yang dapat menjadi fasilitas yang tidak begitu memadai dalam penggunaannya pada berbagai sistem mobil robot. Untuk mengatasi kelemahan tersebut, diimplementasikanlah suatu sistem panoramik, yang memanfaatkan sifat-sifat cermin untuk menjangkau segala sudut pandang yang tidak dapat dijangkau oleh kamera konvensional standar. Bukan berarti tidak ada celah, sistem panoramik juga memiliki ketidaknyamanan di dalam, (1) sifat-sifat proyektif geometris yang tidak begitu mudah diaplikasikan, dan (2) citra yang tidak begitu mudah dipahami oleh mata manusia biasa.

### 11.2 Formasi citra *catadioptric*

Sebagaimana telah diungkapkan sebelumnya dalam berbagai penelitian yang telah dilakukan, ada empat kombinasi kamera-cermin yang diaplikasikan untuk memperoleh sifat-sifat titik pandang tunggal. Untuk menghindari beraneka ragam pengkomputasian pose untuk masing-masing komputasi, maka diimplementasikanlah generalisasi pendekatan untuk tujuan keseragaman komputasi dan interpretasi. Beberapa pendekatan generalisasi untuk formasi citra *catadioptric* adalah seperti apa yang dilakukan oleh, (1) Y. Yagi and S. Kawato, "Panoramic Scene Analysis with Conic Projection", In Proceedings of the International Conference on Robots and Systems, 1990, dan (2) C. Geyer and K. Daniilidis, "A Unifying Theory for Central Panoramic Systems and Practical Implications", In

Proceedings of the European Conference on Computer Vision, pp. 445-461, Dublin 2000.

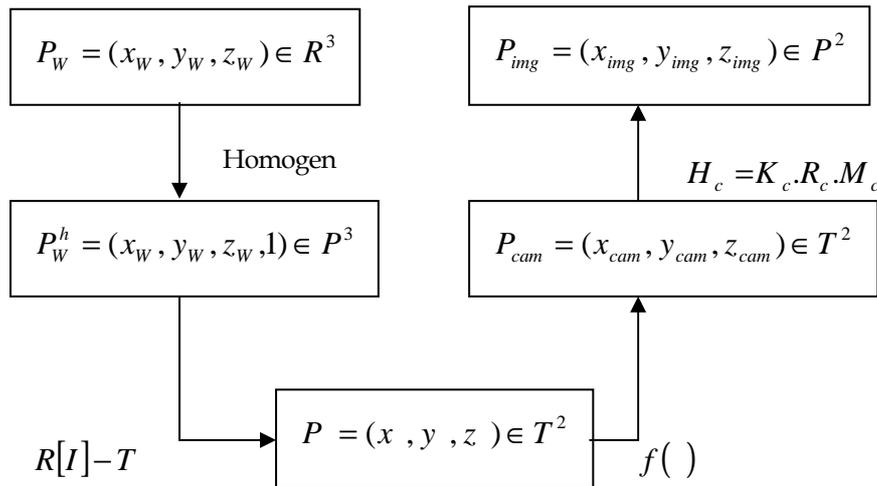
Asumsikan sistem *catadioptric* untuk titik pandang tunggal diilustrasikan seperti gambar 11.1. Titik  $p_w$  di dalam koordinat alam ditransformasikan menjadi koordinat cermin menjadi sinar proyeksi  $p = (x, y, z)$  melalui suatu rotasi  $R$  dan translasi  $-R \cdot T$ , di mana  $T$  merupakan titik foci di dalam koordinat alam. Sinar proyeksi  $p = (x, y, z)$  akan mengiris permukaan cermin pada  $P_m^h = (\lambda \cdot x, \lambda \cdot y, \lambda \cdot z, \lambda)$  untuk beberapa  $\lambda > 0$  dan tergantung dengan jenis dari cermin yang digunakan.



Gambar 11.1 Representasi sistem *catadioptric*

Titik  $P_m^h$  berkorespondensi dengan sinar proyeksi yang berada di dalam koordinat kamera, sehingga  $p_{cam} = P_c \cdot A_c \cdot P_m^h$ , di mana  $P_c$  mendefinisikan jenis dari kamera, proyeksi atau orthografik, dan  $A_c$  mentransformasikan koordinat cermin menjadi koordinat kamera.

Kita dapat memperoleh citra dari suatu titik pandang sebagaimana terlihat dari sebuah cermin panoramik,  $P_{img} = K_c \cdot R_c \cdot x_{cam}$ .



Gambar 11.2 Urutan model pemetaan general

Transformasi yang non-linier dari sinar proyeksi  $p$  menjadi  $p_{cam}$  didefinisikan sebagai fungsi  $f( )$ ,

$$P_{cam} = M_c \cdot f(p)$$

$$M_c = \begin{bmatrix} \psi - \xi & 0 & 0 \\ 0 & \xi - \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$f(p) = \left( \frac{x}{\sqrt{x^2 + y^2 + z^2}}, \frac{y}{\sqrt{x^2 + y^2 + z^2}}, \frac{z}{\sqrt{x^2 + y^2 + z^2}} + \xi \right)^T$$

Ekspresi untuk  $\psi$  dan  $\xi$  sangat tergantung pada jenis cermin dan parameter yang digunakan. Transformasi lengkap dari gambar 11.2 selanjutnya adalah :

$$P_{img} = H_c \cdot f(R.[I | -T].P_w^h)$$

$$H_c = K_c \cdot R_c \cdot M_c$$

Keunggulan dari model ini adalah kelinieran pemetaan dari persamaan di atas, yang memudahkan sistem untuk melakukan inversi. Dikarenakan masing-masing sinar proyeksi  $p$  berkorespondensi terhadap satu dan hanya satu sinar proyeksi  $p_{cam}$ , maka fungsi  $f( )$  adalah injektif dan mempunyai sebuah invers  $f^{-1}( )$ ,

$$f^{-1}(P_c) = (\lambda_{cam} \cdot x_{cam}, \lambda_{cam} \cdot y_{cam}, \lambda_{cam} \cdot z_{cam} - \xi)^T$$

$$\lambda_{cam} = \frac{z_{cam} \cdot \xi + \sqrt{z_c^2 + (1 - \xi^2) \cdot (x_c^2 + y_c^2)}}{x_c^2 + y_c^2 + z_c^2}$$

Sinar proyeksi  $p$  dapat dikomputasikan sebagai :

$$p = f^{-1}(H_c^{-1} \cdot P_{img})$$

### 11.3 Estimasi pose

Pose dikomposisikan oleh suatu operasi rotasi dan operasi translasi. Dengan menggunakan relasi yang telah dibahas sebelumnya, maka untuk beberapa titik pada koordinat alam,  $p_w$  :

$$P = R \cdot P_w - R \cdot T$$

dan matriks rotasi dapat dibentuk dari tiga sudut :

$$R = R_z(\theta) \cdot R_y(\psi) \cdot R_x(\phi)$$

Posisi empat titik dalam koordinat alam harus dipilih untuk memudahkan pengkomputasian.

$$P_{1w} = (x_0, y_0, z_0)^T$$

$$P_{2w} = (x_0, y_0 + L, z_0)^T$$

$$P_{3w} = (x_0, y_0 + z_0 + W)^T$$

$$P_{4w} = (x_0, y_0 + L, z_0 + W)^T$$

#### 11.3.1 Estimasi rotasi

Definisikan  $p_i = (x_i, y_i, z_i)$  sebagai sinar proyeksi untuk koordinat titik alam  $P_{wi}$ , dengan  $i = 1, 2, 3, 4$ , maka relasi berikut akan dipenuhi :

$$\lambda_i \cdot p_i = R \cdot P_{wi} \quad \Leftrightarrow$$

$$\lambda_i \cdot R^{-1} \cdot p_i = \lambda_i \cdot R_x(-\phi) \cdot R_y(-\psi) \cdot R_z(-\theta) \cdot p_i = P_{wi}$$

Sinar-sinar tersebut direpresentasikan pada gambar 11.3. Untuk memudahkan manipulasi aljabar, titik selanjutnya yang berbeda yang berasal dari titik terakhir

yang diperoleh dari hasil perotasian pada sekeliling sumbu  $Z$ , akan digunakan, sehingga :

$$p_i' = (x_i', y_i', z_i')^T = R_Z(-\theta).p_i$$

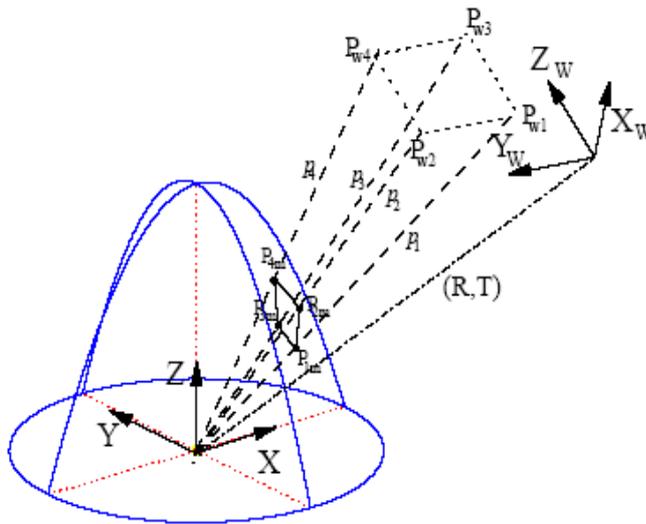
sehingga persamaan,

$$\lambda_i.p_i = R.P_{wi} \Leftrightarrow$$

$$\lambda_i.R^{-1}.p_i = \lambda_i.R_X(-\phi).R_Y(-\psi).R_Z(-\theta).p_i = P_{wi}$$

menjadi

$$\lambda_i.R^{-1}.p_i = \lambda_i.R_X(-\phi).R_Y(-\psi).p_i' = P_{wi}$$



Gambar 11.3 Sinar-sinar proyeksi yang digunakan untuk komputasi pose

Catatan, persamaan terakhir merepresentasikan 3 persamaan, dan, dengan empat titik, kita akan memiliki 12 persamaan yang lain, dan 8 variabel yang tidak diketahui :  $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \phi, \psi, L, W$ . Persamaan dalam  $\psi$  dan  $\phi$  dapat ditulis sebagai berikut :

$$\tan \phi = \frac{x_1'.z_2' - x_2'.z_1'}{(x_1'.y_2' - x_2'.y_1') \cdot \cos \psi + (y_1'.z_2' - y_2'.z_1') \cdot \sin \psi}$$

$$\tan \phi = \frac{x_3'.z_4' - x_4'.z_3'}{(x_3'.y_4' - x_4'.y_3') \cdot \cos \psi + (y_3'.z_4' - y_4'.z_3') \cdot \sin \psi}$$

$$\tan \phi = \frac{x_1'.z_3' - x_3'.z_1'}{(x_1'.y_3' - y_1'.x_3') \cdot \cos \psi + (y_1'.z_3' - z_1'.y_3') \cdot \sin \psi}$$

$$\tan \phi = \frac{x_2'.z_4' - x_4'.z_2'}{(x_2'.y_4' - y_2'.x_4') \cdot \cos \psi + (y_2'.z_4' - z_2'.y_4') \cdot \sin \psi}$$

$$\tan \psi = \frac{(x_3' \cdot z_4' - x_4' \cdot z_3') \cdot (x_1' \cdot y_2' - x_2' \cdot y_1') - (x_1' \cdot z_2' - x_2' \cdot z_1') \cdot (x_3' \cdot y_4' - x_4' \cdot y_3')}{(x_1' \cdot z_2' - x_2' \cdot z_1') \cdot (y_3' \cdot z_4' - y_4' \cdot z_3') - (x_3' \cdot z_4' - x_4' \cdot z_3') \cdot (y_1' \cdot z_2' - y_2' \cdot z_1')}$$

$$\tan \psi = \frac{(x_2' \cdot y_4' - y_2' \cdot x_4') \cdot (x_3' \cdot z_1' - x_1' \cdot z_3') - (x_1' \cdot y_3' - y_1' \cdot x_3') \cdot (x_4' \cdot z_2' - x_2' \cdot z_4')}{(y_1' \cdot z_3' - z_1' \cdot y_3') \cdot (x_4' \cdot z_2' - x_2' \cdot z_4') - (y_2' \cdot z_4' - z_2' \cdot y_4') \cdot (x_3' \cdot z_1' - x_1' \cdot z_3')}$$

Untuk menghitung  $\psi$  dan  $\phi$ ,  $\theta$  dibutuhkan, yang diperoleh dengan mensubstitusikan persamaan :

$$p_i' = (x_i', y_i', z_i')^T = R_Z(-\theta) \cdot p_i$$

ke dalam persamaan :

$$\tan \psi = \frac{(x_3' \cdot z_4' - x_4' \cdot z_3') \cdot (x_1' \cdot y_2' - x_2' \cdot y_1') - (x_1' \cdot z_2' - x_2' \cdot z_1') \cdot (x_3' \cdot y_4' - x_4' \cdot y_3')}{(x_1' \cdot z_2' - x_2' \cdot z_1') \cdot (y_3' \cdot z_4' - y_4' \cdot z_3') - (x_3' \cdot z_4' - x_4' \cdot z_3') \cdot (y_1' \cdot z_2' - y_2' \cdot z_1')}$$

$$\tan \psi = \frac{(x_2' \cdot y_4' - y_2' \cdot x_4') \cdot (x_3' \cdot z_1' - x_1' \cdot z_3') - (x_1' \cdot y_3' - y_1' \cdot x_3') \cdot (x_4' \cdot z_2' - x_2' \cdot z_4')}{(y_1' \cdot z_3' - z_1' \cdot y_3') \cdot (x_4' \cdot z_2' - x_2' \cdot z_4') - (y_2' \cdot z_4' - z_2' \cdot y_4') \cdot (x_3' \cdot z_1' - x_1' \cdot z_3')}$$

sehingga diperoleh :

$$\begin{cases} \tan \psi = \frac{A \cdot \cos \theta + B \cdot \sin \theta}{C} \\ \tan \psi = \frac{D \cdot \cos \theta + E \cdot \sin \theta}{F} \end{cases} \Leftrightarrow \tan \theta = \frac{D \cdot C - A \cdot F}{B \cdot F - E \cdot C}$$

$$\begin{aligned} A &= (x_1 \cdot y_2 - x_2 \cdot y_1) \cdot (x_3 \cdot z_4 - x_4 \cdot z_3) - (x_1 \cdot z_2 - x_2 \cdot z_1) \cdot (x_3 \cdot y_4 - x_4 \cdot y_3) \\ B &= (x_1 \cdot y_2 - x_2 \cdot y_1) \cdot (y_4 \cdot z_3 - y_3 \cdot z_4) - (y_2 \cdot z_1 - y_1 \cdot z_2) \cdot (x_3 \cdot y_4 - x_4 \cdot y_3) \\ C &= (x_2 \cdot z_1 - x_1 \cdot z_2) \cdot (y_4 \cdot z_3 - y_3 \cdot z_4) + (y_1 \cdot z_2 - y_2 \cdot z_1) \cdot (x_4 \cdot z_3 - x_3 \cdot z_4) \\ D &= (x_3 \cdot z_1 - x_1 \cdot z_3) \cdot (x_2 \cdot y_4 - y_2 \cdot x_4) - (x_1 \cdot y_3 - y_1 \cdot x_3) \cdot (x_4 \cdot z_2 - x_3 \cdot z_4) \\ E &= (x_2 \cdot y_4 - y_2 \cdot x_4) \cdot (y_1 \cdot z_3 - z_1 \cdot y_3) - (x_1 \cdot y_3 - y_1 \cdot x_3) \cdot (y_2 \cdot z_4 - z_2 \cdot y_4) \\ F &= (x_3 \cdot z_1 - x_1 \cdot z_3) \cdot (z_2 \cdot y_4 - y_2 \cdot z_4) + (z_1 \cdot y_3 - y_1 \cdot z_3) \cdot (x_2 \cdot z_4 - x_4 \cdot z_2) \end{aligned}$$

### 11.3.2 Estimasi translasi

Ambil sebagai titik pertama,  $P_{w1} = (x_{1w}, y_{1w}, z_{1w})^T$  dan sinar proyeksi terkait  $p_1$ , kita dapat ekspresikan di dalam bentuk formulasi, sinar proyeksi tersebut di dalam koordinat alam :

$$P_{w1} = (x_{1w}, y_{1w}, z_{1w})^T = R^{-1} \cdot p + T$$

Persamaan garis yang melalui titik tersebut adalah :

$$\frac{x - x_{1w}}{x_{1rw} - T_x} = \frac{y - y_{1w}}{y_{1rw} - T_y} = \frac{z - z_{1w}}{z_{1rw} - T_z}$$

dengan mengkombinasikan kedua persamaan di atas, maka diperoleh persamaan berikut :

$$\frac{x - x_{1w}}{r'_{11} \cdot x_1 + r'_{12} \cdot y_1 + r'_{13} \cdot z_1} = \frac{y - y_{1w}}{r'_{21} \cdot x_1 + r'_{22} \cdot y_1 + r'_{23} \cdot z_1} = \frac{z - z_{1w}}{r'_{31} \cdot x_1 + r'_{32} \cdot y_1 + r'_{33} \cdot z_1}$$

yang beririsan dengan bidang  $z = z_{1w} + d$ , maka kita peroleh pusat cermin  $L = (x_c, y_c, z_c)$  di dalam koordinat alam adalah sebagai :

$$\begin{cases} x_c = \frac{d \cdot (r'_{11} \cdot x_1 + r'_{12} \cdot y_1 + r'_{13} \cdot z_1)}{r'_{31} \cdot x_1 + r'_{32} \cdot y_1 + r'_{33} \cdot z_1} + x_{1w} \\ y_c = \frac{d \cdot (r'_{21} \cdot x_1 + r'_{22} \cdot y_1 + r'_{23} \cdot z_1)}{r'_{31} \cdot x_1 + r'_{32} \cdot y_1 + r'_{33} \cdot z_1} + y_{1w} \\ z_c = z_{1w} + d \end{cases}$$

Titik kedua,  $p_{w2} = (x_{2w}, y_{2w}, z_{2w})^T$  selanjutnya diketahui. Definisikan operasi produk skalar  $\overrightarrow{LP_{w1}} \cdot \overrightarrow{LP_{w2}} = |\overrightarrow{LP_{w1}}| \cdot |\overrightarrow{LP_{w2}}| \cdot \cos \alpha$ , cosinus  $\alpha$  dapat kita tulis di dalam koordinat alam dan koordinat cermin sebagai :

$$\cos \alpha = \frac{x_1 \cdot x_2 + y_1 \cdot y_2 + z_1 \cdot z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} \cdot \sqrt{x_2^2 + y_2^2 + z_2^2}}$$

$$\cos \alpha = \frac{(x_{1w} - x_c) \cdot (x_{2w} - x_c) + (y_{1w} - y_c) \cdot (y_{2w} - y_c) + (z_{1w} - z_c) \cdot (z_{2w} - z_c)}{L_{1w} \cdot L_{2w}}$$

$$L_{iw} = \sqrt{(x_{iw} - x_c)^2 + (y_{iw} - y_c)^2 + (z_{iw} - z_c)^2}$$

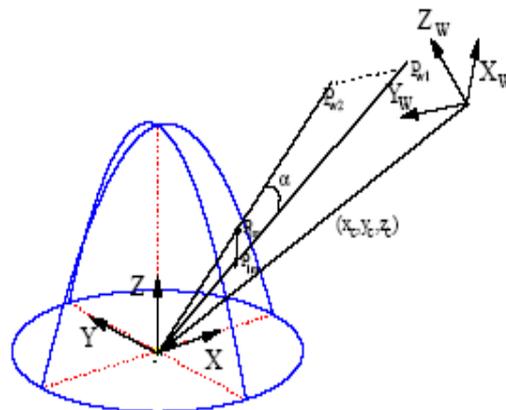
Untuk menyelesaikan persamaan terakhir, substitusikan persamaan :

$$\begin{cases} x_c = \frac{d \cdot (r'_{11} \cdot x_1 + r'_{12} \cdot y_1 + r'_{13} \cdot z_1)}{r'_{31} \cdot x_1 + r'_{32} \cdot y_1 + r'_{33} \cdot z_1} + x_{1w} \\ y_c = \frac{d \cdot (r'_{21} \cdot x_1 + r'_{22} \cdot y_1 + r'_{23} \cdot z_1)}{r'_{31} \cdot x_1 + r'_{32} \cdot y_1 + r'_{33} \cdot z_1} + y_{1w} \\ z_c = z_{1w} + d \end{cases}$$

sehingga persamaan kuadratik pada  $d$  dapat diekspresikan sebagai :

$$\begin{aligned}
 & p \cdot d^2 + q \cdot d + r = 0 \\
 & p = (1 - \cos^2 \alpha) \cdot (a^2 + b^2 + 1) \\
 & \begin{cases} q = -2 \cdot (1 - \cos^2 \alpha) \cdot (a^2 + b^2 + 1) \cdot [a \cdot (x_{2w} - x_{1w}) + b \cdot (y_{2w} - y_{1w}) + (z_{2w} - z_{1w})] \\ r = [a \cdot (x_{2w} - x_{1w}) + b \cdot (y_{2w} - y_{1w}) + (z_{2w} - z_{1w})]^2 - \cos^2 \alpha \cdot (a^2 + b^2 + 1) \cdot [(x_{2w} - x_{1w})^2 + (y_{2w} - y_{1w})^2 + (z_{2w} - z_{1w})^2] \end{cases} \\
 & a = \frac{r'_{11} \cdot x_1 + r'_{12} \cdot y_1 + r'_{13} \cdot z_1}{r'_{31} \cdot x_1 + r'_{32} \cdot y_1 + r'_{33} \cdot z_1}, \quad b = \frac{r'_{21} \cdot x_1 + r'_{22} \cdot y_1 + r'_{23} \cdot z_1}{r'_{31} \cdot x_1 + r'_{32} \cdot y_1 + r'_{33} \cdot z_1}
 \end{aligned}$$

selanjutnya titik  $L = (x_c, y_c, z_c)$  dapat dihitung.



Gambar 11.4 Vektor dengan titik kedua

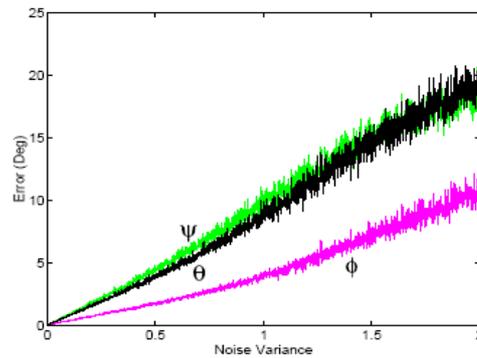
#### 11.4 Eksperimen

Nilai-nilai baku yang ditentukan sebagai definisi awal di dalam sistem adalah :

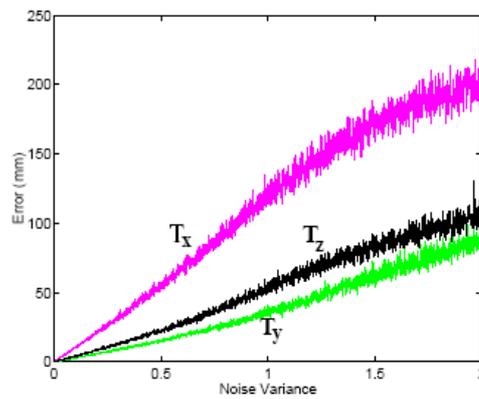
$$\begin{aligned}
 p &= 40, P_1 = [0, 10, 10]^T, \theta = 10^\circ \\
 \phi &= 20^\circ, \psi = 10^\circ, T = [300, 100, 200]^T
 \end{aligned}$$

Gambar 11.5 merepresentasikan berbagai error di dalam sudut rotasi, dan gambar 11.6 merepresentasikan error di dalam translasi. Pada kedua kasus, koordinat-

koordinat citra mengalami pengkorupsian yang disebabkan oleh *Gaussian white noise*, dengan berbagai varians.



Gambar 11.5 Error yang terjadi dalam pengestimasi rotasi



Gambar 11.6 Error yang terjadi dalam pengestimasi translasi

### 11.5 Kesimpulan

Metode yang telah dipaparkan di atas adalah sangat berguna untuk berbagai aplikasi yang membutuhkan perhitungan transformasi antara sistem koordinat alam dan sistem koordinat *catadioptric*.

Menjadi suatu solusi analitis yang memungkinkan jangkauan penggunaannya pada sistem waktu nyata.

## DAFTAR PUSTAKA

- Fu, K.S., Gonzalez, R.C., dan Lee, C.S.G., 1987. *Robotics : Control, Sensing, Vision, and Intelligence*, McGraw-Hill International Editions
- Gonzalez, Rafael C., dan Wintz, Paul. 1987. *Digital Image Processing*, Addison Wesley
- Hearn, D. dan Baker, MP. 1994. *Computer Graphics*. Englewood Cliffs, New Jersey : Prentice-Hall
- Craig, John J., 1986. *Introduction to Robotics : Mechanics & Control*, Addison Wesley
- Low, Adrian. 1991. *Computer Vision & Image Processing: Introductory*, McGraw-Hill International Editions
- Rogers, DF dan Adams, JA.1989. *Mathematical Elements For Computer Graphic* : McGraw-Hill

### Daftar Journal & Proceeding

- António Paulino, Helder Araújo. 2001. *Control Aspects of Maintaining Non-Holonomic Robots in Geometric Formation*, SIRS'2001--Proceedings of the 9th International Symposium on Intelligent Robotic Systems, pp. 18-20, Toulouse, France
- Antonio Paulino, Helder Araújo. 2002. *Pose Estimation for Central Catadioptric Systems: an Analytical Approach*, ICPR'2002—16th International Conference on Pattern Recognition , Quebec City, Canada
- António Paulino, Jorge Batista, Helder Araújo. 2000. *Controlling Multiple Robots in Geometric Formation Using Vision*, in *Mechatronics and Machine Vision*, John Billingsley (Ed.), Research Studies Press, Hertfordshire, UK, pp. 57-64
- C. Thorpe, O. Clatz, D. Duggins, J. Gowdy, R. MacLachlan, J.R. Miller, C. Mertz, M. Siegel, C. Wang, and T. Yata., 2001. *Dependable Perception for Robots*, Proceedings of International Advanced Robotics Programme IEEE, Robotics and Automation Society, Seoul, Korea
- Colbry, Dick., Stockman, George., dan Jain, Anil., 2006. *Detection of Anchor Points for 3D Face Verification*, Journal of Computer Sciences

- 
- Curran, Kevin., Li, Xuelong., dan Mc Caughley, Neil., 2005. *The Use of Neural Networks in Real-time Face Detection*, Journal of Computer Sciences 1 (1), 47 – 62, ISSN 1549 – 3636
- Fadlisyah dkk., 2006. *Edge-Detection Menggunakan Operator Prewitt Termodifikasi*, Journal of Science and Technology (SAINTEK), Vol. 4 No. 2 Nopember 2006, 23 – 31, ISSN 1693-4105
- J. Dias, C. Paredes, I. Fonseca, J. Batista, H. Araújo, A. T. Almeida. 1998. *Simulating Pursuit with Machines--Experiments with Robots and Artificial Vision*, IEEE Transactions on Robotics and Automation, Vol. 14, No. 1
- J. Gao and J. Shi. 2003. *Inferring Human Upper Body Motion Using Belief Propagation*, tech. report CMU-RI-TR-03-06, Robotics Institute, Carnegie Mellon University
- J. P. Barreto, F. Martin, Radu Horaud. 2002. *Visual Servoing/Tracking Using Central Catadioptric Images*, ISER2002 - 8th International Symposium on Experimental Robotics, Ischia, Italy
- Guo, Guodong., dan Dyer, Charles R., 2006. *Face Cyclographs for Recognition*, Computer Sciences Department University of Wisconsin – Madison, Madison, WI 53706
- Guo, Guodong., dan Dyer, Charles R., 2005. *Learning from Examples in the Small Sample Case : Face Expression Recognition*, IEEE Transaction on System, Man, and Cybernetics – Part B: Cybernetics Vol 35, No. 3, 477 – 488
- K. Schmidt and J. Cohn., 2001. *Dynamics of facial expression: Normative characteristics and individual differences*, IEEE International Conference on Multimedia and Expo ( ICME2001), pp. 728 - 731
- K, Stubbs, P. Hinds, dan D. Wettergreen., 2006. *Challenges to Grounding in Human-Robot Collaboration: Errors and Miscommunications in Remote Exploration Robotics*, tech. report CMU-RI-TR-06-32, Robotics Institute, Carnegie Mellon University, 1 - 17
- Lu, Xiaoguang., Jain, Anil K., dan Colbry, Dirk., 2006. *Matching 2.5D Face Scans to 3D Models*, IEEE Transaction on Pattern Analysis and Machine Intelligence Vol 28, No. 1, 31 – 43
- Lu, Xiaoguang., dan Jain, Anil K., 2006. *Deformation Modelling for Robust 3D Face Matching*, Journal of Computer Science
- Lu, Xiaoguang., Chen, Hong., dan Jain, Anil K., 2006. *Multimodal Facial Gender and Ethnicity Identification*, Journal of Computer Science

## DAFTAR PUSTAKA

---

- Lu, Xiaoguang., dan Jain, Anil K., 2006. *Ethnicity Identification from Face Images*, Journal of Computer Science
- N. Duta and A.K. Jain., 1998. *Learning the Human Face Concept in Black and White Images*, Proc. 14th Int'l. Conf. Pattern Recognition, Brisbane, pp. 1365 – 1367
- N. Gonçalves, H. Araujo. 2003. *Mirror Shape Recovery From Image Curves and Intrinsic Parameters: Rotationally Symmetric and Conic Mirrors*, accepted in OMNIVIS'03—Omnidirectional Vision and Cameras Networks, workshop at CVPR'03.
- Nuno Gonçalves, Helder Araújo. 2002. *Estimation of 3D Motion from Stereo Images - Differential and Discrete Formulations*, IROS'2002—IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland
- Park, Unsang., Chen, Hong., dan Jain, Anil K., 2005. *3D Model – Assisted Face Recognition in Video*, Journal of Computer Science
- Park, Unsang., dan Jain, Anil K., 2006. *3D Face Reconstruction from Stereo Video*, Journal of Computer Science
- Peng, Kung., Liming, Chen., Su, Ruan., dan Kukharev, Georgy., 2005. *A Robust Algorithm for Eye Detection on Gray Intensity Face without Spectacles*, JCS&T Vol. 5, No. 3, 127 – 132
- Qian Yu, Helder Araujo, Hong Wang. 2003. *Stereo-Vision Based Real Time Obstacle Detection for Urban Environments*, ICAR'2003—11th Int. Conf. on Advanced Robotics, Coimbra, Portugal.
- Sugimoto, Akihiro., Kimura, Mitsuhiro., dan Matsuyama, Takashi., 2005. *Detecting Human Heads with Their Orientation*, Electronics Letters on Computer Vision and Image Analysis 5(3)
- T.W. Fong, I. Nourbakhsh, dan K. Dautenhahn., 2002. *A Survey of Socially Interactive Robots: Concepts, Design, and Applications*, tech. report CMU-RI-TR-02-29, Robotics Institute, Carnegie Mellon University
- T.W. Fong, J. Scholtz, J. Shah, L. Flueckiger, C. Kunz, D. Lees, J. Schreiner, M. Siegel, L. Hiatt, I. Nourbakhsh, R. Simmons, R. Ambrose, R. Burridge, B. Antonishek, M. Bugajska, A. Schultz, dan J.G. Trafton., 2006. *A Preliminary Study of Peer-to-Peer Human-Robot Interaction*, International Conference on Systems, Man, and Cybernetics, IEEE

Utz, Hans., Kaufmann, Ulrich., Mayer, Gerd., dan Kraetzschmar, Gerhard K., -.  
*VIP - A Framework-Based Approach to Robot Vision*, Advanced Robotic  
System

V. Brajovic., 2001. *An Object Tracking Computational Sensor*, SPIE Proceedings  
Vol. 4019: Critical Technologies for the Future of Computing, SPIE

#### Daftar Dissertasi

C. Dima., 2006. *Active Learning for Outdoor Perception*, doctoral dissertation,  
tech. report CMU-RI-TR-06-28, Robotics Institute, Carnegie Mellon  
University

C. Zitnick., 2003. *Computing Conditional Probabilities in Large Domains by  
Maximizing Renyi's Quadratic Entropy*, doctoral dissertation, tech. report  
CMU-RI-TR-03-20, Robotics Institute, Carnegie Mellon University

D. Nikovski., 2002. *State-Aggregation Algorithms for Learning Probabilistic  
Models for Robot Control*, doctoral dissertation, tech. report CMU-RI-TR-  
02-04, Robotics Institute, Carnegie Mellon University

Guo, Guodong., 2006. *Face, Expression, and Iris Recognition Using Learned-  
Based Approaches*, Ph.D's Dissertation, Computer Science University of  
Wisconsin – Madison, USA

Hsu, Rein-Lein., 2002. *Face Detection and Modelling for Recognition*, Ph.D's  
Dissertation, Department of Computer Science & Engineering Michigan  
State University, USA

P. Chang., 2003. *Robust Tracking and Structure from Motion with Sampling  
Method*, doctoral dissertation, tech. report CMU-RI-TR-02-33, Robotics  
Institute, Carnegie Mellon University

S. Iba., 2004. *Interactive Multi-Modal Robot Programming*,  
doctoral dissertation, tech. report CMU-RI-TR-04-50, Robotics Institute,  
Carnegie Mellon University

Y. Tsin., 2003. *Kernel Correlation as an Affinity Measure in Point-Sampled Vision  
Problems*, doctoral dissertation, tech. report CMU-RI-TR-03-36,  
Robotics Institute, Carnegie Mellon University