

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Berikut beberapa penelitian atau karya yang berkaitan dengan pemanfaatan *stereo vision* sebagai pengukur jarak :

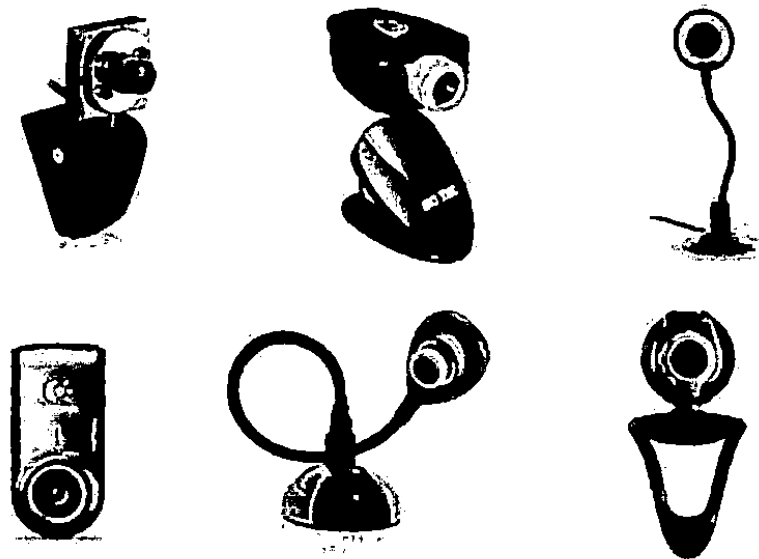
- Metode *Stereo Vision* Untuk Memperkirakan Jarak Objek Dari Kamera. Izzati Muhimmah, Novian Mahardika Putra, Mellita, Deny Rahmalianto, Dhomas Hatta Fudholi. 2012. Merupakan sebuah penelitian tentang *stereo vision* untuk memperkirakan jarak berbasis *Template Matching*.
- Perancangan Sistem Kendali dan Navigasi Berbasis *Stereo Vision* untuk *Obstacle Avoidance* Pada *Three Wheels Omidirectional Robot*. Faikul Umam. 2012. Sebuah karya tentang penerapan *Stereo Vision* pada Robot penghindar rintangan dengan metode *Neuro Fuzzy*.
- *Stereo Vision* untuk Pengukuran Jarak Objek dengan Mendeteksi Tepi *Sub Image*. Dwiretno Istiyadi Swasono. 2009. Pada karya ini metode yang digunakan adalah mendeteksi tiap tepi pada gambar-gambar yang telah dipecah menjadi beberapa bagian dari sebuah gambar utuh yang tertangkap oleh webcam.
- Penentuan Jarak Objek dari Kamera dengan Menggunakan *Stereo Vision* Pada *Mobile Phone. Smart Phone* sebagai pengukur jarak menggunakan

- Sistem Penjejakan Obyek dengan *Stereo Vision*. Tri Arief Sardjono, Djoko Purwanto, Achmad Fiqhl Ibadillah. 2011. Penerapan metode *Stereo Vision* menggunakan teknik pendeteksian warna sebagai objek titik tengah.
- High Accuracy Stero Vision System for Far Distance Obstacle Detection. Sergiu Nedevschi, Radu Danescu, Dan Frentiu, Tiberiu Marita, Florin Oniga, Ciprian Pocol, Rolf Schmidt, Thorsten Graf. 2004. Sebuah karya peelitian penerapan *Stereo Vision* di jalan raya perkotaan dengan hasil yang sangat akurat.

2.2 Dasar Teori

2.2.1 Webcam

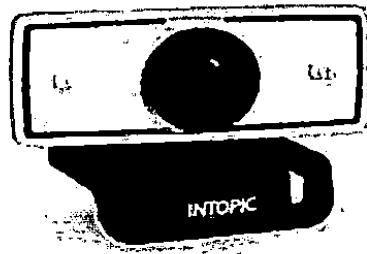
Webcam adalah gabungan dua kata yang berasal dari bahasa inggris yakni *web* dan *camera*. Webcam dapat menampilkan gambar dalam bentuk waktu nyata atau keadaan pada saat itu juga yang gambarnya dapat dilihat pada *world wide web* (www), program pengolahan pesan cepat, aplikasi pemanggilan berbasis video. Penamaan webcam dapat terbentuk sesuai dengan aplikasi penggunaan dimana webcam tersebut digunakan, misal VolcanoCam adalah kamera yang memperlihatkan keadaan gunung berapi, TraffiCam yang dapat digunakan untuk memantau trafik jalan raya, StreetCam yang menampilkan gambar-gambar jalan, LiveCam biasanya untuk aplikasi *chatting*. Biasanya webcam dihubungkan ke komputer melalui konektor USB atau COM. Beberapa vendor mengeluarkan produk dengan berbagai spesifikasi dari tingkat rendah ke tingkat tinggi



Gambar 2.1 Webcam Dengan Berbagai Macam Model Dan Merk

(<http://ahlikompie.com/wp-content/uploads/2012/03/contoh-macam-webcam.jpg> akses 05/07/2014,
08.55.33 WIB)

Sala satu penyedia webcam adalah perusahaan Intopic yang berbasiskan di Taiwan. Webcam intopic livecam 720 yang akan digunakan dalam penelitian ini. Seri tersebut hadir dengan desain *portable* sehingga dapat ditempatkan dimana saja (Clip Anywhere) baik di monitor komputer, meja, dinding, badan robot, dan lain lain. Kelebihan lainnya adalah dapat diarahkan kemana saja secara manual karena terdapat fungsi pivot dan *sturdy clip* di bagian bawah nya. Webcam ini sudah terintegrasi mikrofon internal dan instalasinya sangat mudah karena tidak membutuhkan driver dan juga mendukung banyak program aplikasi seperti Microsoft MSN, Skype, Yahoo Messenger, Facebook, dan lain lain, serta terdapat program bawaan dari webcam yakni CamApp, dengan program tersebut pengguna dapat mengakses webcam baik untuk pengambilan gambar ataupun video. Kelebihan yang utama juga adalah karena webcam dipatok dengan harga yang



INTOPIC

Gambar 2.2 Webcam Intopic Livecam 720

(<http://img.udn.com/image/product/S0001137/APPROVED/U000438347/20140331140311795.jpg>)

akses 05/07/2014, 22.30.20 WIB)

Adapun spesifikasi detail webcam tersebut adalah : desain padat, inovatif dan fasion; desain bagian bawah yang mampu bergeser, 9 macam efek gambar dan fungsi pelacakan wajah, 7 *frame* foto yang dapat di pilih, tombol untuk memilih menu sebanyak 10 mode efek pada gambar, didesain untuk konferensi video atau akting web, dan memonitor rumah dari jarak jauh, cocok untuk komputer desktop ataupun notebook, USB 2.0 kecepatan tinggi, plug dan play+UVC, mikrofon dengan kualitas tinggi.

Antarmuka Ke PC	USB 2.0 High-Speed (Plug & Play)
Sensor Type	1/6 CMOS VGA Sensor, 48.0 Mega pixels (interpolated)
Resolution	VGA (640 x 480) Format, Max. 5.0 Mega Pixels Dynamic Preview Resolution
Lens Features	High Quality 3 Layer Glasses Lens
Frame Rate	Frame Rate up to 30fps

Picture Control	Auto/Manual White Balance
Focus Type	Manual Focusing Ring (5.0 cm ~ infinity)
Built-In Microphone	Available (USB Connection)
Face Tracking	Available
Still Image Capture	2560 x 2048 Pixel (Max), 1-Touch Snapshot, BMP Image Format
LED Lightning Support	Not Available
Pan / Tilt Model	Manual, 360° Horizontal Pan, 30° Vertical Tilt
Operating System Support	PC MS Windows Vista, XP2 System Plug & Play (UVC)
Other Features	RGB 24 (True Color: 24bit) Color, Online Snapshot w/ Situation Model Button
Dimension (WHD)	62(L) x 42.5(W) x 39(H) mm
Weight	62 gram
Others	Made In Taiwan
Manufacturer website	www.intopic.com.tw

Webcam ini menggunakan konektor antar muka USB (*Universal Serial Bus*) 2.0 berkecepatan tinggi dengan fitur "*Plug & Play*", fitur yang tidak memerlukan penginstalan *software driver* terlebih dahulu. Ketika webcam diubungkan dengan komputer maka secara otomatis *driver* yang dibutuhkan akan terinstal dengan sendirinya.

Sensor yang digunakan adalah tipe "1/6 CMOS VGA". VGA (*Video Graphic Array*) adalah standar grafis komputer resolusi 640 x 480 ketika *mode* pengambilan video diaktifkan. CMOS (*Complimentary Metal Oxide Semiconductor*) merupakan tipe sensor yang sering dijumpai di kamera SLR, berbeda dengan CCD (*Charge Coupled Device*) yang membutuhkan daya yang lebih boros dan tidak sensitif terhadap cahaya. Webcam ini juga mampu untuk menginterpolasikan gambar dengan kemampuan sampai 48.0 Megapixel. Interpolasi di webcam digunakan untuk memperbesar gambar dengan jumlah 48 juta titik tiap 1 inci.

Terdiri dari 3 lapis lensa dengan kualitas tinggi sehingga keamanan sensor menjadi tetap terjaga. Mampu menghasilkan *Frame Per Second* tinggi, dalam satu detik mampu menangkap gambar hingga 30 bingkai. Dilengkapi dengan fitur keseimbangan gambar secara otomatis. Bisa difokuskan sampai jarak 5 cm dari posisi semula. Dilengkapi dengan mikrofon dengan konektor USB, dengan demikian webcam dapat dimanfaatkan sebagai alat *chatting*. Webcam berkualitas 5.0 Megapixel dengan masing masing jumlah maksimal pixel di sumbu X sebesar 2560 dan di sumbu Y 2048 titik tiap 1 inci. Format gambar yang dihasilkan adalah

arah vertikal. Dapat digunakan dengan sistem operasi Windows dan Linux.

Dimensi 62(P)x42.5(L)x39(T) mm, berat 62 gram.

2.2.2 Pengolahan Citra

Pengolahan Citra adalah suatu cabang ilmu informatika yang hasil merupakan kombinasi dari ilmu pengetahuan lain seperti komputer, elektronik, matematika, fisika, fotografi yang bertujuan untuk mentransformsikan suatu citra atau gambar menjadi citra lain. Pada umumnya pengolahan citra melibatkan perangkat-perangkat elektronik digital sehingga kemampuan dalam mengolah citra menjadi lebih baik.

Mentransformasikan suatu citra membutuhkan beberapa teknik pengolahan yang dikenal dengan algoritma. Beberapa algoritma yang biasa digunakan dalam dunia pengolahan citra adalah sebagai berikut :

- Algoritma berbasis histogram
- Algoritma berbasis matematika
- Algoritma berbasis konvolusi
- Algoritma berbasis morfologi
- Algoritma berbasis penurunan

Dari semua algoritma yang disebutkan di atas, yang akan diterapkan di penelitian ini adalah algoritma berbasis matematika dengan operasi biner (berbasis operasi boolean untuk memanipulasi data), operasi aritmetika (berbasis

Pengolahan citra adalah bagian dari penglihatan komputer (*computer vision*) dan penglihatan mesin (*machine vision*). Namun dalam penelitian ini teori pengolahan citra yang digunakan diambil dari penglihatan komputer. Beberapa penerapan penglihatan komputer yang menggunakan prinsip pengolahan citra adalah pelacakan video, pengenalan objek, pendeteksi insiden, pendeteksi gerakan, rekonstruksi gambar 2D menjadi 3D.

Seiring dengan meningkatnya kebutuhan suatu sistem yang menggunakan aplikasi pengolahan citra seperti untuk rekonstruksi sebuah lingkungan menjadi gambar 3D yang akan sangat bermanfaat untuk analisis, akuisisi, memahami kondisi lingkungan tersebut maka dibutuhkan sebuah kamera atau webcam yang mampu memindai objek tersebut ataupun untuk estimasi kecepatan laju sebuah benda berbasis gambar 3D, estimasi jarak dengan metode transformasi jarak (*distance transform*), sedangkan kemampuan webcam yang ada saat ini hanya mampu menghasilkan gambar 2D, hanya mampu melihat ketinggian, dan kelebaran objek yang mewakili sumbu X dan Y. Untuk menghasilkan gambar 3D dengan kemampuan melihat kedalaman objek sebagai perwakilan sumbu Z, dibutuhkan 2 webcam atau setidaknya 2 layar LCD atau lensa yang diposisikan sejajar. Disiplin ini dikenal dengan istilah *stereo vision* atau *stereoscopy*. Sejatinya prinsip yang digunakan *stereo vision* saat ini telah diperkenalkan oleh Sir Charles Wheatstone tahun 1838.

- Citra Gray
- Kontur
- Deteksi Tepi
- *Thesh Hold Binary*
- *Stereo Vision*
- *Distance Transform*

a. Citra RGB

Citra RGB adalah suatu citra yang terdiri dari tiga komponen warna *Red* (merah), *Green* (hijau), *Blue* (biru). Citra dalam komputer tidak lebih dari sekumpulan sejumlah warna dimana setiap warna terdiri atas variasi tingkat keterangan (*Brightness*) dari elemen merah, hijau, dan biru. Representasi dalam citra, elemen warna RGB akan terdiri dari 3 angka yang mengatur intensitas merah, hijau, dan biru. Setiap elemen akan merepresentasikan 1 pixel (elemen gambar). Suatu elemen dengan nilai 100, 100, 100 berarti akan mengeset nilai

• 100 100 100 100 100 100. Angka angka RGB ini sering disebut

b. Citra Gray

Dalam dunia komputasi, citra Gray atau *Grayscale* adalah suatu citra dimana nilai dari setiap pixel merupakan contoh tunggal. Citra yang ditampilkan dari citra jenis ini terdiri atas warna abu-abu, bervariasi pada warna hitam pada bagian yang intensitas terlemah dan warna putih pada intensitas terkuat. Citra gray berbeda dengan citra hitam-putih, dimana pada konteks komputer, citra hitam-putih hanya terdiri dari dua warna saja yaitu “hitam” dan “putih” saja.

Citra grayscale seringkali merupakan perhitungan dari intensitas cahaya pada setiap pixel pada spektrum elektromagnetik *single band*. Citra jenis ini disimpan dalam format 8 bit untuk setiap pixel, dimana memungkinkan sebanyak 256 intensitas cahaya. Format ini sangat membantu dalam pemrograman karena manipulasi bit yang tidak terlalu banyak. Pada aplikasi lain seperti pada aplikasi medical imaging dan remote sensing biasa juga digunakan format 10, 12 maupun 16 bit

c. Kontur

Kontur adalah salah satu metode dalam segmentasi pengolahan gambar digital untuk memisahkan objek dengan objek lainnya. Kontur dapat memisahkan tiap objek berdasarkan pendeteksian tepi objek. Kontur dalam pengolahan citra

... dan umumnya dibuat dengan garis bentuk berupa kotak atau lingkaran

d. Deteksi Tepi

Analisis pada citra pada dasarnya terdiri dari tiga tahapan , yaitu ekstraksi ciri, segmentasi, dan klasifikasi. Faktor kunci dalam mengekstraksi ciri adalah kemampuan mendeteksi keberadaan tepi di dalam gambar. Pendeteksian tepi dalam pengolahan citra merupakan tahapan untuk melingkupi informasi di dalam citra. Tepi mencirikan batas objek dan karena itu tepi berguna untuk proses segmentasi dan identifikasi objek di dalam citra. Tepi adalah perubahan nilai intensitas derajat keabuan yang mendadak besar dalam jarak yang singkat. Perbedaan intensitas inilah yang memperlihatkan rincian pada gambar. Tepi dapat diorientasikan dengan satu arah, dan arah ini berbeda-beda, tergantung pada perubahan intensitas.

e. Tresh Hold Binary

Adalah batas ambang berupa biner yang diterapkan di segmentasi gambar. Fungsi ini diperlukan ketika ingin mengkonversikan citra gray ke citra biner, karena citra gray terdiri dari 256 intensitas sedangkan citra biner hanya memiliki 2 intensitas maka dibutuhkan sebuah batas ambang yang dapat membedakan nilai sekian termasuk dari 0 atau 1. Misal batas ambang dari citra gray adalah 128

1. Jika nilai tersebut adalah di atas menjadi hitam (0)

f. *Stereo Vision*

Stereo Vision adalah proses pengembalian struktur 3D menggunakan dua webcam atau lebih terhadap gambar 3D. Sistem ini meniru sistem penglihatan umum yang digunakan oleh hewan dan manusia bermata dua dengan syarat tidak memiliki kelainan seperti kebutaan, mata juling, perbedaan fokus mata, dan lain-lain. Teknik ini sejatinya meniru prinsip kerja mata pada manusia ataupun hewan yang dikenal dengan sebutan *Binocular Vision*¹

Stereo vision diterapkan di banyak ilmu seperti bidang robotika untuk mengenali lingkungan sekitar, kedokteran untuk pemetaan tubuh manusia, *Fashion* untuk *scanning* berbagai model baju sehingga beberapa *sample* tersebut dapat dikombinasikan untuk menghasilkan model-model terbaru, transportasi untuk mengukur laju kendaraan, dan masih banyak lagi.

Terdapat dua kategori *stereo vision* dari segi cara untuk mendapatkan informasi 3D, *stereo vision* aktif dengan cara menembakkan sinar laser atau cahaya, gelombang ultrasonik, infrared, dan lain-lain. Sedangkan *stereo vision* aktif adalah menggunakan kamera atau webcam. Metode pasif menjadi sangat populer karena disamping untuk mendapatkan informasi 3D juga dapat untuk mendeteksi serta melacak benda bergerak, menghitung jumlah objek. Sehingga kosakata *stereo vision* seringkali dianggap sebuah metode yang hanya menggunakan webcam.

1. *Kelebihan dan Kekurangan Sistem Stereo Vision*

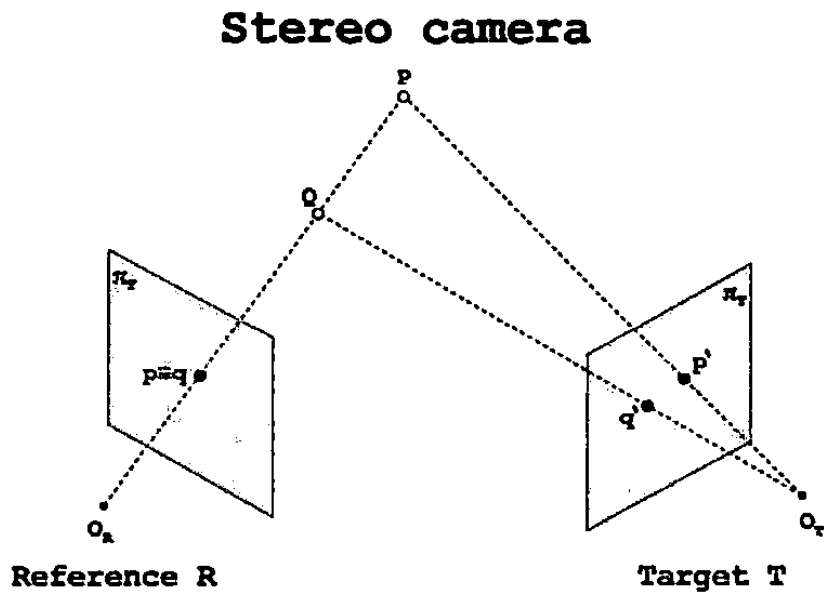
Webcam yang digunakan sebagai inputan *stereo vision* akan melibatkan pengolahan citra digital, dan membutuhkan manipulasi elemen elemen matriks. Sedangkan operasi dasar yang digunakan adalah operasi boolean, aritmatik, dan geometri dengan demikian akan melibatkan pixel-pixel atau titik-titik penyusun gambar menjadi objek yang dapat dilihat oleh manusia.

Algoritma-algoritma *stereo vision* adalah :

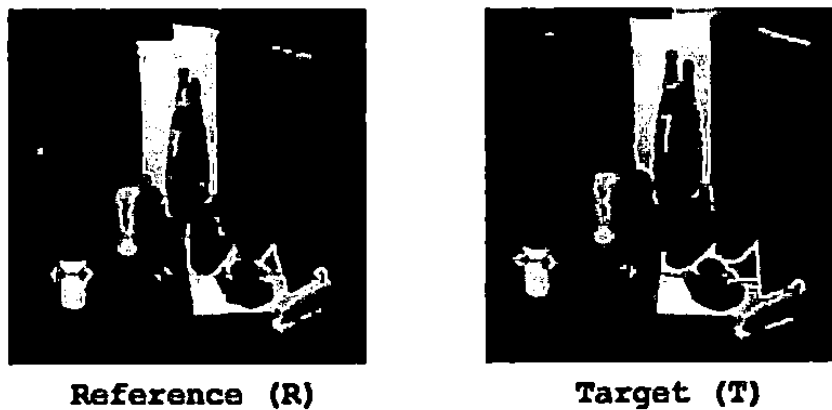
- Epipolar Constraint
- Ordering Constraint
- Figural Continuity
- Dynaming Programming (Dijkstra)

Epipolar Constraint atau Geometry adalah penerapan geometri di Stero Vision, dimana ketika dua webcam atau kamera dibiarkan untuk melihat gambar 3D dari posisi yang berbeda, terdapat nilai-nilai geometri yang berhubungan antara titik-tik 3D dan diproyesikan ke gambar 2D yang menjadi kendala.

Untuk mengetahui bagaimana webcam dapat mendeteksi kedalaman persepsi objek atau target dengan teknik epipolar constraint adalah seperti ditunjukkan pada gambar di bawah ini, dimana webcam sebelah kiri mendeteksi dua objek Q dan P dalam bidang yang sama dan tegak lurus, sedangkan webcam sebelah kanan mendeteksi objek Q dan P juga akan tetapi di garis yang berbeda, dengan demikian dapat diketahui perbedaan posisi Q dan P jika gambar yang dihasilkan oleh kedua webcam dibandingkan dan dikombinasikan sehingga



Gambar 2.3 Cara Stereo Vision Mendeteksi Kedalaman

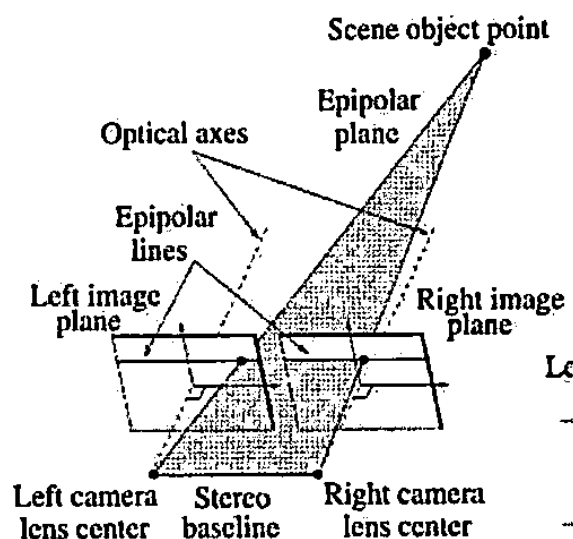


Gambar 2.4 Proses Pencarian Kedalaman Dengan Block Matching

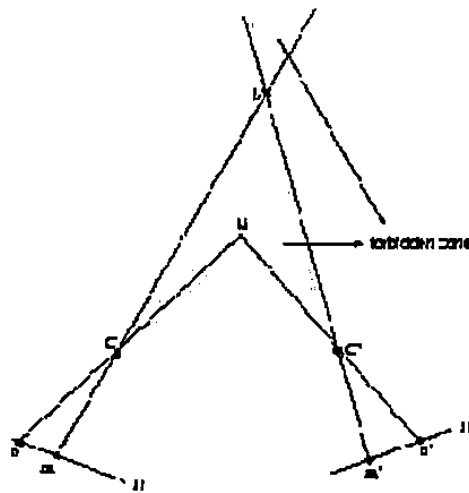
Pada gambar di atas adalah penjelasan berdasarkan data gambar sesungguhnya, dimana Webcam sebelah kiri menjadi sumber referensi penentuan *Disparity* atau *Displacement* yang ditandai dengan garis lurus merah. Webcam sebelah kanan yang menjadi sumber untuk menentukan tingkat kedalaman objek

Berikut beberapa istilah di bidang *Stereo Vision* :

- *Fixation Point* : adalah titik pertemuan sudut optik (*Optical Axis*)
- *Baseline* : adalah jarak antara pusat proyeksi
- *Epipolar Plane* : adalah papan yang dilalui pusat dan titik proyeksi
- *Epipolar Line* : adalah interseksi epipolar plane dengan papan gambar
- *Conjugate Pair* : adalah titik dimana saja yang dapat diliha oleh kedua webcam akan diproyeksikan terhadap pasangan titik gambar ke dua gambar
- *Disparity / Displacement* : adalah jarak antara titik yang sesuai ketika kedua gambar disatukan
- *Disparity / Displacement Map* : adalah peta semua titik-titik perpindahan atau kesenjangan



Langkah-langkah dengan teknik ordering constraint adalah teknik untuk menyepadankan dua gambar menjadi satu. Berbeda dengan epipolar yang cara penyepadannya menggunakan teknik tingkat kemiripan maka ordering dengan cara memberikan pengurutan terhadap objek. Ketika webcam sebelah kiri mendeteksi tiga objek dan webcam sebelah kanan mendeteksi objek hanya dua, maka objek pertama dari webcam sebelah kiri akan disepadankan dengan objek pertama webcam sebelah kanan.



Gambar 2.6 Proses Penyepadanan Dengan Teknik Ordering

(http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OWENS/LECT11/img20.gif akses

09/08/2014, 21.27.00 WIB)

Cara lain adalah figural continuity, dimana menurut Marr dan Poggio, 1979 mengatakan bahwa disparitas dari gambar yang disepadankan harus bervariasi dengan rata pada semua gambar. Namun algoritma ini gagal dan tidak digunakan

Pada intinya algoritma *dynamic programming* tidak hanya digunakan di *stereo vision*, akan tetapi juga diterapkan di banyak bidang. Di *stereo vision* algoritma ini dimanfaatkan untuk koordinat suatu tempat seperti sketsa biografis sebuah jalan. Namun untuk pemanfaatan mengukur jarak, algoritma ini tidak maksimal dalam penerapannya.

g. Distance Transform

Transformasi Jarak juga disebut Peta Jarak dan Bidang Jarak berasal dari representasi gambar digital. Pemilihan istilah tergantung pada sudut pandang objek tersebut apakah gambar awal ditransformasikan menjadi representasi lain atau hanya ditambahkan beberapa beberapa fungsi sehingga seolah-olah gambar menjadi gambar dengan bentuk lain. Dalam bidang ini juga dapat menentukan apakah titik berada di dalam bentuk atau di luar bentuk gambar. Transformasi jarak melibatkan peta-peta pixel terdekat untuk menentukan jarak terdekat pixel terhadap penghalang atau objek. Peta pixel dapat ditemukan pada papan catur seperti berikut dimana warna papan telah di transformasikan ke gambar biner.

Dalam bidang ini juga berkaitan dengan Geometri Citra. Geometri Citra adalah salah satu operasi pengolahan citra yang berfungsi untuk merubah koordinat pixel namun intensitas tetap, geometri citra adalah kebalikan dari aritmetika citra. Pada umumnya transformasi jarak dikategorikan secara garis besar menjadi tiga bagian menurut teori yang berkaitan dengan metrik matematika yang

Salah satunya adalah Jarak Euklidides, Jarak Manhattan, Jarak Papan Catur, Bidang

ini banyak diterapkan di robotika, perencanaan perpindahan (*motion planning*), perintis jalan (*path finding*).

Salah satu aplikasi yang dapat diterapkan dengan bidang ini adalah untuk mengkalkulasikan jarak antara 2 pixel dalam satu gambar tapi untuk sampai ke proses tersebut membutuhkan proses komputasi yang cukup rumit dan berat. Jarak antar pixel dapat ditemukan dengan rumus Euclidean : $D_{euclid} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$. Jika pixel memiliki koordinat di $x_2, x_1, y_2,$ dan y_1 maka untuk mencari jarak antar pixel adalah dengan rumus di atas. Dapat juga dicari dengan persamaan *City Block* : $D_{city} = |x_2 - x_1| + |y_2 - y_1|$. Persamaan tersebut juga dikenal dengan istilah Jarak Manhattan. Disamping menggunakan kedua rumus di atas, dapat juga menggunakan rumus dari Jarak Papan Catur : $D_{chess} = \max(|x_2 - x_1| + |y_2 - y_1|)$. dari ketiga rumus tersebut, teori Euclidean adalah yang paling lambat untuk mengkomputasi persamaan.

2.2.3 Bahasa Pemrograman C++

C++ adalah bahasa pemrograman komputer yang dibuat oleh Bjarne Stroustrup pada awal 1980-an dan merupakan pengembangan dari bahasa pemrograman C. Pada C++ ditambahkan konsep-konsep baru seperti *class* dengan sifat-sifatnya yang *inheritance* dan *overloading*. Salah satu perbedaan mendasar

antara C dan C++ adalah perbedaan terhadap konsep pemrograman berorientasi

Namun demikian C++ menggunakan sintaks yang sama dengan C akan tetapi bahasa C adalah bahasa prosedural, dimana penyelesaian masalah dibagi-bagi kedalam beberapa sub-submasalah yang lebih kecil, berbeda dengan bahasa C++ ketika menyelesaikan masalah dengan cara menjelaskan kelas-kelas yang merupakan anak kelas yang dibuat sebelumnya sebagai abstraksi objek-objek fisik. Kelas tersebut berisi keadaan objek, anggota-anggotanya dan kemampuan dari objeknya. Setelah beberapa kelas dibuat maka kemudian masalah dipecahkan dengan kelas.²

2.2.4 Pustaka OpenCV

OpenCV adalah singkatan dari *Open Source Computer Vision Library* merupakan sebuah pustaka perangkat lunak yang ditujukan untuk pengolahan citra dinamis secara *real time* yang dibuat oleh Intel dan saat ini didukung oleh Willow Garage dan Itseez. Program ini bebas dan berada dalam naungan sumber terbuka dari lisensi BSD. Pustaka ini merupakan pustaka lintas platform. Program ini didedikasikan sebagian besar untuk pengolah citra secara *real time*. OpenCV pertama kali diluncurkan secara resmi pada tahun 1999 oleh Inter Research sebagai lanjutan dari bagian proyek bertajuk aplikasi intensif berbasis CPU, *real time ray tracking*, dan tembok penampil 3D. Para kontributor utama dalam proyek ini termasuk mereka yang berkecimpung dalam bidang optimasi di Intel Rusia dan juga tim pustaka performansi intel. Pada awalnya tujuan utama dari proyek OpenCV ini dideskripsikan sebagai berikut :

² <http://id.wikipedia.org/wiki/C%2B%2B>

- a. Penelitian penginderaan citra lanjutan tidak hanya melalui kode program terbuka, tetapi juga kode yang telah teroptimasi untuk infrastruktur penginderaan citra
- b. Menyebarluaskan ilmu penginderaan citra dengan menyediakan infrastruktur bersama di mana para pengembang dapat menggunakan secara bersama-sama sehingga kode akan tampak lebih mudah dibaca dan ditransfer
- c. Membuat aplikasi komersial berbasis penginderaan citra, di mana kode yang telah teroptimasi tersedia secara bebas dengan lisensi yang tersedia secara bebas yang tidak mensyaratkan program itu harus terbuka atau gratis.³



Gambar 2.7 Logo Pustaka OpenCV

(http://upload.wikimedia.org/wikipedia/commons/thumb/3/32/OpenCV_Logo_with_text_svg_version.svg/487px-OpenCV_Logo_with_text_svg_version.svg.png akses 06/08/2014, 19.55.00 WIB)

³ <http://id.wikipedia.org/wiki/OpenCV>

OpenCV dioptimalkan dengan 2500 lebih pustaka algoritma. OpenCV menyediakan fitur *Integrated Performance Primitives* (IPP) Intel sehingga bisa lebih mengotimalisasikan aplikasi *vision* jika menggunakan *processor* intel.

OpenCV terdiri dari lima pustaka yaitu :

- CV : pustaka untuk algoritma pengolahan citra dan penglihatan
- ML : pustaka untuk pembelajaran mesin
- Highgui : pustaka untuk GUI, gambar, video *input/output*.
- CXCORE : pustaka untuk struktur data, mendukung XML, dan fungsi-fungsi grafis

Berikut beberapa aplikasi yang dapat dijalankan dengan OpenCV :

- Manipulasi Data (*Data Manipulation*)
- Fitur Toolkit untuk 2D dan 3D (*2D and 3D feature toolkits*)
- Estimasi egomosi (*Egomotion Estimation*)
- Sistem pengenalan wajah (*Face Recognition*)
- Pengenalan sikap, gerakan, dan isyarat (*Gesture Recognition*)
- Interaksi Komputer-Manusia (*Human-Computer Interaction*)
- Perangkat robotika (*Mobile Robotic*)
- Identifikasi objek (*Object Identification*)
- Pemahaman gerakan (*Motion Understanding*)
- Segmentasi dan pengenalan (*Segmentation and Recognition*)
- Pelacakan gerak (*Motion Tracking*)
- Penglihatan *Stereo Stereopsis* (*Stereopsis Stereo vision: depth perception from 2 cameras*)

- Analisa Struktur (*Structure Analysis*) : *Distance Transform, Contour Processing, Template Matching*
- Struktur dari gerak (*Structure From Motion* (SFM))
- Ditambah kenyataan (*Augmented reality*)⁴

2.3 Bidang Yang Diterapkan Pada Pembuatan Alat Pengukur Jarak

2.3.1 Pengolahan Citra

2.3.1.1 *Motion Detection*

Mendeteksi pergerakan adalah proses pendeteksian perubahan posisi sebuah objek yang tidak mutlak terhadap sekitar atau perubahan sekitar terhadap objek yang tidak mutlak. Ada dua cara untuk mendeteksi pergerakan, pertama dengan algoritma ini : kalkulasikan nilai rata-rata warna yang telah ditentukan di bingkai pertama, tunggu beberapa detik, kalkulasikan nilai rata-rata warna yang telah ditentukan di bingkai kedua, pergerakan terdeteksi. Algoritma kedua adalah : kalkulasikan nilai tengah massa (*middle mass*) di bingkai pertama, tunggu beberapa detik, kalkulasikan nilai tengah massa (*middle mass*) di bingkai kedua, jika (nilai bingkai pertama – nilai bingkai kedua) > nilai batas ambang, pergerakan terdeteksi. Dari kedua algoritma di atas, algoritma kedua yang diterapkan pada alat ini. kelebihan algoritma kedua adalah dapat diaplikasikan untuk menghitung jarak sebuah objek, kecepatan atau percepatan sebuah objek, nilai *displacement*. Sehingga dapat diterapkan di MOT (*Multiple Object Tracking*).

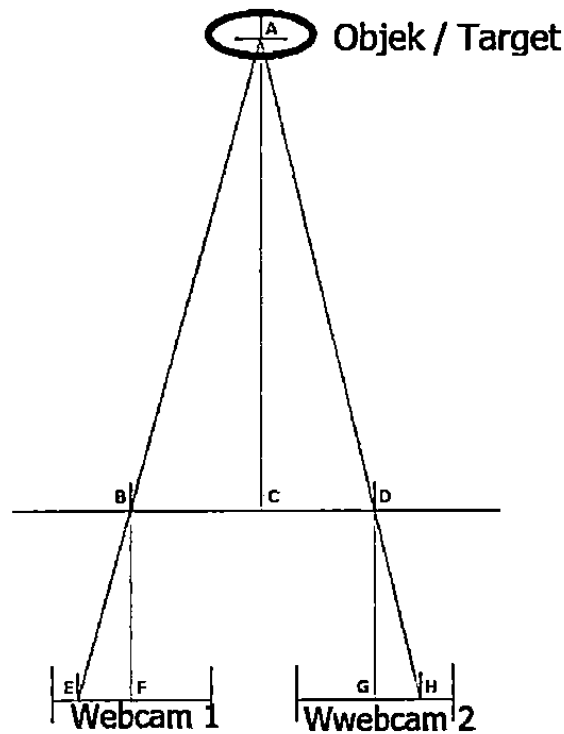
⁴ <http://en.wikipedia.org/wiki/OpenCV>

Tantangan dalam pendeteksian pergerakan adalah seberapa sensitif nilai batas ambang, jika nilai tersebut sensitif maka sistem akan mampu mendeteksi pergerakan yang cepat bahkan sangat cepat sekalipun.

2.3.1.2 Multiple Object Tracking (MOT)

Teknik ini sejatinya menirukan cara kerja mata yang mampu mendeteksi banyak objek dalam satu waktu. Sejatinya teknik ini merupakan pengembangan dari pendeteksian pergerakan. Namun kemampuan algoritma yang ditingkatkan sehingga mampu mendeteksi banyak objek sekaligus. Contoh algoritma yang diterapkan di bidang ini adalah untuk mengukur jarak : $(kecepatan * per_pixel) / (nilai\ bingkai\ pertama - nilai\ bingkai\ kedua)$. Algoritma tersebut hanya untuk pelacakan satu objek, agar dapat banyak objek maka dengan algoritma ini : deteksi perubahan disekitar bingkai 1, hitung jumlah perubahan disekitar objek bergerak, deteksi perubahan disekitar bingkai 2, hitung jumlah perubahan disekitar objek bergerak, banyak objek bergerak terdeteksi dan terlacak. Dengan algoritma tersebut di atas, maka alat ini mampu menghitung jarak terhadap banyak objek

2.3.2 Stereo Vision



Gambar 2.8 Model Segitiga Pada Sistem *Stereo Vision*

Pada gambar di atas terlihat sebuah objek sama-sama ditangkap oleh 2 webcam di waktu yang bersamaan sehingga antara objek dan kedua webcam membentuk bentuk triangulasi. Adapun bentuk segitiga yang dibentuk tidak pasti sama sisi, bisa juga berbentuk segitika lancip, siku-siku, sembarang tergantung letak objek.

Titik A adalah objek yang menjadi target bersama dari dua webcam, titik B adalah titik dimana pertemuan dua garis yang berseberangan dari webcam 1, lebih dikenal dengan istilah *Pinhole Camera*. Titik D adalah *Pinhole Camera* webcam 2 dan E adalah titik paling kiri lensa webcam 1, titik H adalah titik paling kanan lensa webcam 2. F dan G adalah titik tengah lensa kedua webcam. $BD =$

Terdapat empat segitiga yang terbentuk dari gambar tersebut, ACB dan BFE, ACD dan DGH, sehingga *displacement / disparity* $d = EF + GH = BD (BF/AC) = k/z$, dimana $k = BD BF$ dan persamaan $z = AC$. AC adalah jarak dari webcam ke objek. Dengan asumsi papan tempat webcam datar dan berada di satu papan maka *displacement* sumbu Y di antara dua gambar adalah $d = k/z$. K adalah jarak antara dua webcam dikalikan dengan jarak lensa ke gambar. Rumus untuk mencari kedalaman kedua gambar adalah $z_2(x,y) = \min (\{v:v = z_1 (x,y - (k/z_1(x,y))))\}$. Sedangkan $z_1(x,y) = \min (\{v:v = z_2 (x,y - (k/z_2(x,y))))\}$.

Untuk mendeteksi kedalaman atau jarak sebuah objek dengan bantuan komputer, ada beberapa yang harus dilakukan terlebih dahulu. Pertama adalah *Block Matching* : misal terapat dua gambar seperti gambar di bawah ini



Gambar 2.9 Gambar Dari Webcam Kiri Sebagai Target

(http://www.gophoto.it/view.php?i=http://chrisjmccormick.files.wordpress.com/2014/01/left_wtemplate_crop.png#U-oLQKOT6IN akses 12/08/2014, 08.00.00 WIB)

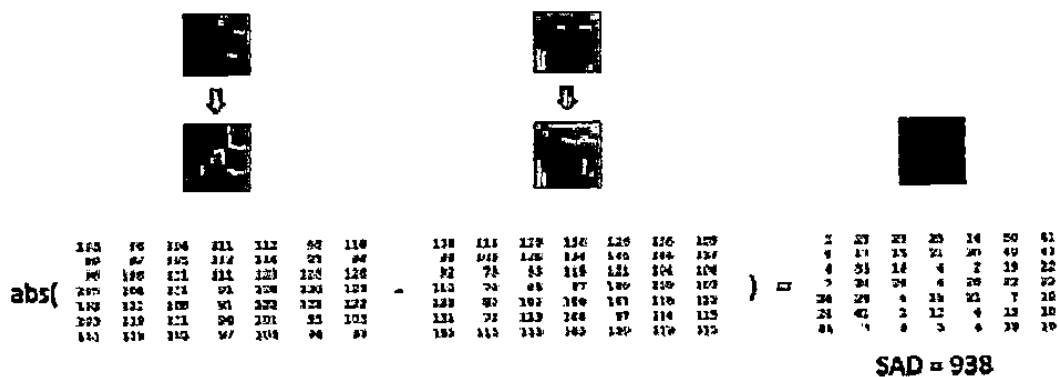


Gambar 2.10 Gambar Dari Webcam Kanan Sebagai Referensi

(http://www.gophoto.it/view.php?i=http://chrisjmccormick.files.wordpress.com/2014/01/right_wse

Gambar pertama adalah hasil tangkapan webcam pertama, dan webcam berikutnya adalah hasil tangkapan dari webcam kedua. Gambar kedua sebagai referensi dan gambar pertama sebagai target, target yang lebih spesifik adalah ujung bentuk berwarna merah, yang pada gambar pertama dikelilingi kotak hitam. Karena webcam kedua webcam sama-sama mendeteksi ujung merah tersebut, berarti pendeteksian objek yang akan dicocokkan sudah pas. Untuk mencocokkan pada gambar yang sama, webcam kedua mencari koordinat gambar ujung merah, dan ditemukan dengan ditandai garis kotak hijau. Setelah gambar cocok maka disparitas atau *displacement* adalah jarak horizontal kotak hijau ke kotak putih.

Setelah nilai disparitas atau *displacement* ditemukan maka jarak nya sudah dapat diketahui, untuk menghitung nilai disparitas dapat menggunakan SAD (*Sum Absolute Difference*). Seperti terlihat pada gambar di bawah ini, nilai SAD adalah penjumlahan nilai yang berbeda dari kedua gambar. Hasil SAD tersebut akan sebagai patokan *disparity* atau *displacement*.



Gambar 2.11 Proses Pencarian Nilai SAD