



PROYEK AKHIR TAHAP 1

SIMULTANEOUS LOCALIZATION AND MAPPING (SLAM) BERBASIS 3D VISUAL-INERTIAL ODOMETRY

Satria Taranika
NRP. 2210171015

Dosen Pembimbing 1:
Prof. Ir. Dadet Pramadihanto, M.Eng., Ph.D.
NIP. 19620211 198811 1 001

Dosen Pembimbing 2:
Dr.Eng. Bima Sena Bayu D, S.ST., M.T.
NIP. 19761215 199903 1 003

PROGRAM STUDI D4 TEKNIK KOMPUTER
DEPARTEMEN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
2021



PROYEK AKHIR TAHAP 1

SIMULTANEOUS LOCALIZATION AND MAPPING (SLAM) BERBASIS 3D VISUAL-INERTIAL ODOMETRY

Satria Taranika
NRP. 2210171015

Dosen Pembimbing 1:
Prof Ir. Dadet Pramadihanto, M.Eng., Ph.D.
NIP. 19620211 198811 1 001

Dosen Pembimbing 2:
Dr.Eng. Bima Sena Bayu D., S.ST., M.T.
NIP. 19761215 199903 1 003

PROGRAM STUDI D4 TEKNIK KOMPUTER
DEPARTEMEN TEKNIK INFORMATIKA DAN
KOMPUTER
POLITEKNIK ELEKTRONIKA NEGERI SURABAYA
2021

***SIMULTANEOUS LOCALIZATION AND MAPPING (SLAM)
BERBASIS 3D VISUAL-INTERTIAL ODOMETRY***

Oleh :

Satria Taranika

NRP. 2210171015

Proyek Akhir ini Digunakan Sebagai Salah Satu Syarat Untuk
Memperoleh Gelar Sarjana Sains Terapan (S.ST)

di

Politeknik Elektronika Negeri Surabaya

2021

Disetujui Oleh :

Tim Penguji Proyek Akhir:

Dosen Pembimbing :

1.

1. Prof. Ir. Dadet Pramadihanto, M.Eng., Ph.D.

NIP. 19620211 198811 1 001

2.

2. Dr.Eng. Bima Sena Bayu D., S.ST., M.T.

NIP. 19761215 199903 1 003

3.

Mengetahui

Ketua Program Studi D4 Teknik Komputer

Departemen Teknik Informatika dan Teknik Komputer

Politeknik Elektronika Negeri Surabaya

Rivanto Sigit, S.T., M.Kom., Ph.D.

NIP. 19700811 199512 1 001

LEMBAR PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan dibawah ini dengan sebenarnya menyatakan bahwa Proyek Akhir ini saya susun tanpa tindakan plagiarisme sesuai dengan peraturan yang berlaku di Politeknik Elektronika Negeri Surabaya (PENS)

Nama : Satria Taranika
NRP : 2210171015
Program Studi : Teknik Komputer
Departemen : Teknik Informatika dan Komputer

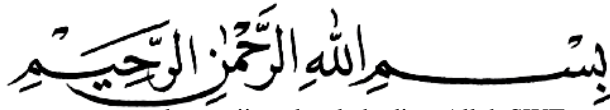
Jika dikemudian hari saya terbukti melakukan tindakan plagiarisme, saya akan bertanggung jawab sepenuhnya dan menerima sanksi yang dijatuhkan oleh PENS kepada saya.

Surabaya, 28 Januari 2020

Satria Taranika
NRP. 2210171015

KATA PENGANTAR

Assalamualaikum, Wr. Wb



Dengan mengucapkan puji syukur kehadiran Allah SWT atas rahmat dan hidayah-Nya, penulis dapat menyelesaikan proyek akhir yang berjudul :

SIMULTANEOUS LOCALIZATION AND MAPPING (SLAM) BERBASIS 3D VISUAL-INERTIAL ODOMETRY

Buku Proyek Akhir ini dimaksudkan untuk memenuhi persyaratan guna menyelesaikan studi Diploma IV di Politeknik Elektronika Negeri Surabaya (PENS).

Terdapat beberapa literatur dan teori baik yang diperoleh dalam perkuliahan maupun dari luar perkuliahan yang digunakan dalam penyelesaian proyek akhir ini, dan juga tidak lepas dari dukungan dosen pembimbing serta pihak-pihak lain yang telah banyak memberikan semangat dan bantuan.

Penulis menyadari bahwa buku proyek akhir ini masih memiliki banyak kekurangan. Untuk itu penulis memohon maaf sebesar-besarnya atas kekurangan yang ada pada buku proyek akhir ini. Selain itu penulis mengharapkan saran dan kritik dari semua pihak demi kesempurnaan buku ini.

Besar harapan penulis agar buku proyek akhir ini dapat bermanfaat khususnya bagi penulis dan pembaca pada umumnya serta untuk menambah wawasan dan pengetahuan.

Surabaya, 28 Januari 2021

Penulis

ABSTRAK

Lokalisasi dan pembuatan peta memainkan peran penting dalam navigasi robot. Mereka bergantung satu sama lain pada saat robot berada lingkungan yang tidak dikenali. Lokalisasi robot membutuhkan peta lingkungan yang akurat, dan pada gilirannya robot perlu mengetahui lokasi tepat mereka untuk membangun peta. *Simultaneous Localization and Mapping (SLAM)* adalah salah satu teknologi yang biasa digunakan untuk membangun kemampuan lokalisasi dan membangun peta pada robot. Akan tetapi sistem *Visual-only SLAM* masih memiliki beberapa kelemahan termasuk ketika menemui medan yang *low texture* ataupun saat robot melakukan gerakan yang berlebih sehingga menyebabkan gambar menjadi *blur*. Kemudian *Visual-Inertial SLAM (VI-SLAM)* muncul sebagai hasil dari peningkatan sistem *SLAM* yang hanya menggunakan kamera sebagai sensornya. *Monocular camera*, *stereo camera*, *omnidirectional camera* dan *RGB-D camera* adalah sensor utama dalam membangun sistem *Visual-Inertial SLAM (VISLAM)*. Pada saat yang sama, sensor *Inertial Measurement Unit (IMU)* menyediakan informasi pergerakan dirinya sendiri. Pada proyek akhir ini akan dibahas tentang proses lokalisasi dan pembuatan peta *3D* pada robot *humanoid*. Sistem yang digunakan untuk proses lokalisasi dan pembuatan peta *3D* dalam proyek akhir ini adalah sistem *Visual-Inertial SLAM* dengan mengkombinasikan kamera *RGB-D* dengan sensor *IMU* yang ada pada *ZED stereo* kamera. Masukan dari sistem ini berupa *streaming* dari *ZED stereo* kamera yang kemudian dilakukan ekstraksi fitur menggunakan metode *FAST feature detector* dan *matching feature* dengan metode *Forward-Backward Lucas Kanade Optical Flow*. Sensor *IMU* digunakan untuk memperbaiki estimasi pose dari kamera. Hasil dari sistem ini berupa peta *3D* dan *3D point cloud* dari peta tersebut dapat disimpan ke dalam *database*. Presentase dari keseluruhan tahapan pengerjaan yang telah selesai dikerjakan pada proyek akhir ini sebesar 30%.

Kata Kunci : *3D-SLAM*, *Visual-Inertial SLAM*, *Visual Inertial Odometry*.

ABSTRACT

Localization and map building play an important role in robot navigation. They depend on each other at a time when robots are in an unknown environment. The robot localization requires accurate map of environment, and in turn the robots need to know their precise locations to build map. Simultaneous Localization and Mapping (SLAM) is one of the technologies commonly used to build localization and map building capabilities on a robot. However, the Visual-only SLAM system still has some disadvantages, including when encountering low texture terrain or when the robot performs excessive movements causing the image to blur. Then Visual-Inertial SLAM (VI-SLAM) appeared as a result of the improvement of the SLAM system that only uses the camera as its sensor. Monocular camera, stereo camera, omnidirectional camera and RGB-D camera are the main sensors in building Visual-Inertial SLAM (VISLAM) systems. At the same time, the Inertial Measurement Unit (IMU) sensor provides its own movement information. In this final project will be discussed about the process of localization and building of 3D map on humanoid robots. Systems used for localization and 3D map building in this final project is a Visual-Inertial SLAM system by combining RGB-D camera with IMU sensor on ZED stereo camera. Input from this system in the form of streaming from ZED stereo camera which is then extracted features using FAST feature detector and matching feature method with forward-backward lucas kanade optical flow method. IMU sensors are used to correct pose estimation from the camera. The result of this system in the form of 3D maps and 3D point cloud of the map can be stored into the database. The percentage of all stages of work completed on this final project is 30%.

Keywords: 3D-SLAM, Visual-Inertial SLAM, Visual Inertial Odometry.

DAFTAR ISI

LEMBAR PERNYATAAN BEBAS PLAGIARISME	iv
KATA PENGANTAR	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan dan Manfaat Penelitian	3
1.4.1 Tujuan.....	3
1.4.2 Manfaat.....	4
1.5 Metodologi Penelitian.....	4
1.5.1 Studi Literatur.....	5
1.5.2 Desain dan Perancangan Sistem	5
1.5.3 Pembuatan Sistem.....	5
1.5.4 Pengujian dan Analisa Sistem.....	5
1.5.5 Pembuatan Laporan	6
1.6 Sistematika Penulisan Laporan.....	6
BAB II KAJIAN PUSTAKA DAN DASAR TEORI	9
2.1 Kajian Pustaka	9
2.1.1 Simultaneous Localization And Mapping (SLAM)	9

2.1.2 Visual-Inertial Simultaneous Localization And Mapping (VI-SLAM).....	9
2.1.3 A Stereo Based Visual-Inertial Odometry for SLAM (2019)10	
2.1.4 Stereo Visual-Inertial SLAM with Point and Lines (2018) ..	10
2.1.5 Robust 3D-SLAM With Tight RGB-D-Inertial Fusion (2019)11	
2.1.6 A Stereo Visual-Inertial SLAM Approach for Indoor Mobile Robots in Unknown Environments Without Occlusions (2019).....	12
2.2 Dasar Teori	12
2.2.1 Computer Vision.....	12
2.2.2 Stereo Camera.....	13
2.2.3 <i>RGB-D</i>	14
2.2.4 Point Cloud	15
2.2.5 Feature Extraction.....	16
2.2.6 <i>Feature Tracking</i>	16
BAB III PERANCANGAN SISTEM.....	18
3.1 Perancangan Sistem.....	18
3.1.1 Akuisisi Data dan Inisialisasi.....	19
3.1.2 Point Cloud Representation	23
3.1.3 Feature Tracking.....	23
3.1.4 Keyframe Decision	24
3.1.5 Loop Closing	25
BAB IV PENGUJIAN DAN ANALISA.....	28
4.1 Perangkat dan Lingkungan Uji Coba.....	28
4.2 Laporan Sistem.....	29
4.3 Tahap Ekstraksi Fitur.....	31
4.4 Tahap Representasi dalam <i>3D Point Cloud</i>	36

BAB V PENUTUP	40
5.1 Kesimpulan.....	40
5.2 Hal yang dilakukan selanjutnya.....	40
DAFTAR PUSTAKA.....	42

DAFTAR GAMBAR

Gambar 1. 1 Metodologi Penelitian.....	4
Gambar 2. 1 Contoh Penerapan Computer Vision.....	13
Gambar 2. 2 ZED Stereo Camera	13
Gambar 2. 3 Contoh penerapan RGB-D.....	14
Gambar 2. 4 Point Cloud Library	15
Gambar 2. 5 Konsep dari Feature Extraction.....	16
Gambar 2. 6 Contoh Feature Extraction dengan metode Kanade Lucas-Tomasi	17
Gambar 3. 1 Desain Sistem	18
Gambar 3. 2 Ilustrasi Sliding Window	25

DAFTAR TABEL

Tabel 4.1 Spesifikasi Perangkat Pengujian.....	27
Tabel 4.2 Tahapan yang Telah Dilakukan	28
Tabel 4.3 Gambar <i>RGB</i>	30
Tabel 4.4 Hasil Ekstraksi Fitur	32
Tabel 4.5 Hasil <i>3D Point Cloud</i>	35

BAB I

PENDAHULUAN

Pada bab 1 ini akan dijelaskan mengenai pendahuluan, rumusan masalah, batasan masalah, tujuan dan manfaat, metodologi, sistematika penulisan penelitian dari proyek akhir ini.

1.1 Latar Belakang

Perkembangan *Artificial Intelligence (AI)* dan *computer vision* telah menyebabkan pertumbuhan yang luar biasa dalam industri robotika. Terdapat banyak jenis robot yang telah diproduksi seperti *mobile robot* dan *humanoid robot*. Robot yang paling populer tentunya adalah robot yang bisa bekerja secara *autonomous*. Namun agar sebuah robot dapat bekerja secara *autonomous*, maka salah satu komponen penting yang harus dimiliki oleh robot tersebut adalah kemampuan untuk melakukan lokalisasi dan pemetaan secara *realtime*. Ketika sebuah robot diaplikasikan di tempat yang belum dikenali, robot tersebut perlu untuk mengerti tubuhnya sedang berada dimana, membuat keputusan berdasarkan posisi tubuhnya dalam peta, serta berjalan secara *autonomous* untuk menyelesaikan tugas yang telah diberikan kepadanya.

Simultaneous Localization And Mapping (SLAM) merupakan salah satu teknologi yang umum digunakan untuk membangun kemampuan lokalisasi dan pemetaan pada sebuah robot. *SLAM* memungkinkan sebuah robot untuk dapat mengerti informasi tentang di lingkungan sekitarnya berdasarkan data yang diperoleh dari sensor seperti kamera, *lidar*, dan *ultrasonic range finder*. Kamera *monocular* merupakan salah satu sensor yang ideal untuk *SLAM* karena ukurannya yang kecil, murah, dan pengoperasiannya yang mudah. Tetapi terdapat banyak permasalahan dalam sistem *SLAM* ketika hanya menggunakan kamera [1].

Dalam beberapa tahun terakhir, *Visual-Inertial SLAM (VI-SLAM)* telah muncul sebagai hasil dari peningkatan kemampuan komputasi dari sistem *SLAM* yang hanya menggunakan kamera

sebagai sensornya. Dibandingkan dengan kombinasi beberapa sensor yang ada, kombinasi dari kamera *stereo* dan sensor *IMU* menghasilkan performa yang luar biasa dan memiliki beberapa keuntungan. Pertama, *VI-SLAM* dapat menghasilkan pose yang presisi dalam situasi dimana sinyal pada sistem satelit navigasi global mengalami kegagalan. Walaupun robot berada dalam situasi tersebut, robot masih bisa untuk membangun peta *3D*. Kedua, kamera *stereo* memiliki lebih banyak informasi yang dapat digunakan untuk pengenalan tempat dan *loop closure*. Kamera *stereo* juga menawarkan informasi yang memungkinkan sebuah robot untuk dapat mengukur skala dan melakukan lokalisasi. Ketiga, sensor *IMU* menyediakan informasi gerakan yang memungkinkan pemulihan dari skala yang terbaca dari kamera. Sensor *IMU* juga dapat mengestimasi arah dari gravitasi dan menentukan dengan tepat *pitch* dan *roll* dari sensor. Keempat, meskipun robot dalam keadaan diam atau bergerak pada kecepatan konstan, kamera *stereo* bisa mendapatkan informasi skala dan menekan arus yang disebabkan oleh sensor *IMU*. Terakhir, kombinasi dari kamera *stereo* dan sensor *IMU* lebih murah dan mengkonsumsi lebih sedikit energi daripada kombinasi dari beberapa sensor lain [1].

Pada proyek akhir ini akan dibangun sebuah sistem dimana nantinya sebuah robot yang bekerja secara *autonomous* dapat memiliki kemampuan untuk melakukan lokalisasi dirinya ketika berada pada lingkungan yang belum dikenali. Pada sistem ini juga nantinya sebuah robot dapat membangun peta *3D* secara *realtime* dan dapat menyimpan serta menggunakan peta tersebut untuk merencanakan rute perjalanan dan menghindari rintangan dalam menyelesaikan tugas yang telah diberikan padanya.

1.2 Rumusan Masalah

Permasalahan yang dihadapi pada proyek akhir ini adalah :

1. Bagaimana cara mengakuisisi data dari kamera dan sensor *IMU*.

2. Bagaimana cara mengkombinasikan data kamera dan sensor *IMU* untuk melakukan lokalisasi.
3. Bagaimana cara membangun peta *3D* berdasarkan data yang diperoleh robot dari lingkungan sekitarnya.
4. Bagaimana cara menyimpan peta yang telah dibangun ke dalam *database* dan dapat menggunakannya sebagai referensi untuk perencanaan rute atau untuk menghindari rintangan.

1.3 Batasan Masalah

Adapun batasan-batasan yang akan diterapkan dalam pembuatan proyek akhir ini sebagai berikut :

1. Sistem yang dibangun menggunakan kamera *stereo* dan sensor *IMU* yang ada di dalam *ZED stereo camera*.
2. Kamera yang digunakan adalah *ZED stereo camera*.
3. Sistem yang dibangun hanya dapat diimplementasikan di dalam ruangan (*indoor*).
4. Sistem *SLAM* ini diimplementasikan pada robot yang dikembangkan oleh RoISC PENS.

1.4 Tujuan dan Manfaat Penelitian

Pada sub bab ini akan menjelaskan mengenai tujuan serta manfaat dari proyek akhir ini.

1.4.1 Tujuan

Tujuan dari proyek akhir ini antara lain :

1. Dapat mengakuisisi data dari kamera *stereo* dan sensor *IMU*.
2. Dapat menggunakan data kamera *stereo* dan sensor *IMU* untuk melakukan lokalisasi.
3. Dapat membangun peta *3D* berdasarkan data dari lingkungan yang ada disekitar robot.
4. Dapat menyimpan peta *3D* ke dalam *database* dan membuat data peta tersebut dapat dipanggil kembali (*recall*).

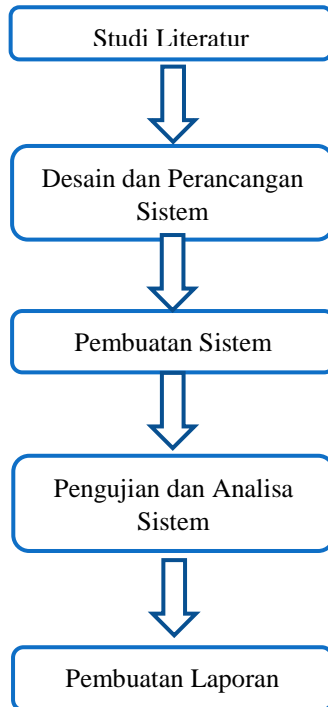
1.4.2 Manfaat

Manfaat dari proyek akhir ini antara lain :

1. Memberikan kemampuan pada sebuah robot *humanoid* untuk melakukan lokalisasi dan pembuatan peta *3D*.
2. Diharapkan mampu memberikan kontribusi pada pengembangan robot di industri robotika.

1.5 Metodologi Penelitian

Penyelesaian sistem akan dikerjakan dalam beberapa tahap. Metodologi yang digunakan dalam pengerjaan proyek akhir adalah seperti pada Gambar 1.1.



Gambar 1.1 Metodologi Penelitian

1.5.1 Studi Literatur

Pada tahap studi literatur dilakukan pengumpulan data referensi mengenai keseluruhan metode yang akan diterapkan pada penelitian sistem, yang terdiri dari studi tentang *Computer Vision*, kamera *stereo*, *RGB-D*, *Point Cloud*, *Feature Extraction*, *Feature Tracking*, serta literatur lainnya yang berhubungan dengan proyek akhir ini. Studi dilakukan dengan membaca buku, *paper*, jurnal ilmiah baik dalam maupun luar negeri dan melalui penelusuran internet.

1.5.2 Desain dan Perancangan Sistem

Desain dan perancangan sistem dilakukan sebelum sistem diimplementasikan agar memperoleh hasil yang maksimal. Perancangan sistem dimulai dari memahami kinerja sistem baik sub sistem maupun keseluruhan sistem, alat dan bahan serta merencanakan metode yang digunakan.

1.5.3 Pembuatan Sistem

Dari rancangan sistem yang telah dibuat, maka dilakukan implementasi sistem. Dalam pembuatan sistem harus menggunakan referensi yang diperoleh pada tahap studi literatur. Program dibuat menggunakan library OpenCV, library Point Cloud, dan aplikasi Visual Studio Code yang berjalan pada sistem operasi UBUNTU 18.04.

1.5.4 Pengujian dan Analisa Sistem

Pengujian dan analisa sistem dilakukan terhadap sistem yang dibuat dalam proyek akhir ini agar dapat berfungsi sesuai dengan yang diharapkan serta menganalisis hasil dari tiap tahapan proses dan hasil keseluruhan. Pengujian sistem dapat dilakukan dengan simulasi atau dengan pengujian secara real. Untuk pengujian berbasis simulasi akan dilakukan pada software Gazebo yang berjalan di Operating System Ubuntu. Sedangkan untuk pengujian

real dengan melakukan streaming menggunakan kamera *stereo*. Proses *streaming* dilakukan di dalam ruangan atau *indoor* dengan menggerakkan kamera dari satu titik ke titik yang lain.

Sistem dikatakan berhasil apabila ketika kamera bergerak dari satu titik ke titik yang lain, sistem dapat menghasilkan beberapa titik dari posisi awal hingga posisi akhir kamera berhenti. Yang menandakan bahwa sistem ini dapat melakukan proses lokalisasi. Kemudian bersamaan dengan itu peta *3D* juga dapat dihasilkan oleh sistem. Yang menandakan bahwa sistem ini dapat melakukan proses *mapping*.

1.5.5 Pembuatan Laporan

Pembuatan laporan ditujukan sebagai dokumentasi dan penjelasan seluruh kegiatan dalam pembuatan proyek akhir. Selain itu juga untuk mengetahui apakah hasil telah sesuai dengan capaian yang diharapkan atau tidak. Pembuatan laporan dilaksanakan pada akhir tahapan pembuatan proyek akhir.

1.6 Sistematika Penulisan Laporan

Dalam menyajikan hasil laporan tugas akhir ini, penulis menggunakan sistematika penulisan laporan yang terdiri dari beberapa Bab sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi tentang pendahuluan yang terdiri dari latar belakang, batasan masalah, tujuan, manfaat, metodologi penelitian serta sistematika penulisan dari proyek akhir ini.

BAB II KAJIAN PUSTAKA DAN DASAR TEORI

Bab ini membahas mengenai teori-teori yang berkaitan dengan penyelesaian tugas akhir, yang didapatkan dari berbagai macam buku serta sumber-sumber terkait lainnya yang berhubungan dengan pembuatan proyek akhir ini.

BAB III PERANCANGAN SISTEM

Bab ini membahas mengenai perancangan sistem, tentang bagaimana proses akuisisi data dari kamera *stereo* dan sensor IMU, proses ekstraksi fitur, proses *tracking* fitur, penentuan keyframe untuk dapat membangun sebuah peta.

BAB IV PENGUJIAN DAN ANALISA

Bab ini menyajikan informasi hasil uji coba yang meliputi hasil uji coba ekstraksi fitur, representasi dalam 3D Point Cloud, tracking fitur, penentuan keyframe, hingga dapat menghasilkan sebuah peta.

BAB V PENUTUP

Bab ini berisi kesimpulan dari uji coba aplikasi, dan saran untuk pengembangan, perbaikan serta penyempurnaan terhadap aplikasi yang telah dibuat.

DAFTAR PUSTAKA

Berisi tentang referensi-referensi yang telah digunakan sebagai landasan selama pembuatan proyek akhir ini

-----Halaman ini sengaja dikosongkan-----

BAB II

KAJIAN PUSTAKA DAN DASAR TEORI

Dalam bab 2 ini membahas mengenai kajian pustaka serta teori-teori yang berkaitan dengan penyelesaian proyek akhir, yang didapatkan dari berbagai macam buku serta sumber-sumber terkait lainnya yang berhubungan dengan pembuatan proyek akhir ini.

2.1 Kajian Pustaka

2.1.1 Simultaneous Localization And Mapping (SLAM)

Teknologi *Simultaneous Localization And Mapping (SLAM)* pertama kali diusulkan oleh Smith [2,3], yang diterapkan dalam robotika dengan tujuan membangun peta *realtime* dari lingkungan berdasarkan data sensor di lingkungan yang tidak diketahui. Karena kemajuan terbaru dalam teknologi *CPU* dan *GPU*, metode *SLAM visual* telah melihat peningkatan minat karena informasi *visual* yang kaya tersedia dari kamera berbiaya rendah dibandingkan dengan sensor lainnya. Ada banyak metode *SLAM visual* yang sangat baik yang telah meningkatkan pengembangan teknologi *SLAM* seperti *MonoSLAM*, *PTAM*, *RatSLAM*, *DTAM*, *KinectFusion*, dan *ORB-SLAM*. Teknologi *SLAM* telah mengalami tiga iterasi besar selama 30 tahun terakhir. Saat ini, teknologi *SLAM* berkembang dan kuat. Teknologi *SLAM* dengan presisi tinggi dan *realtime* sangat dibutuhkan dalam robotika [4].

2.1.2 Visual-Inertial Simultaneous Localization And Mapping (VI-SLAM)

Visual-Inertial Simultaneous Localization And Mapping (VI-SLAM) yang memadukan kamera dan data sensor *IMU* untuk lokalisasi dan persepsi lingkungan telah menjadi semakin populer untuk beberapa alasan. Pertama, teknologi ini digunakan dalam robotika, terutama dalam penelitian dan aplikasi yang luas melibatkan *Micro Aerial Vehicles (MAV)*. Kedua, *Augmented*

Reality (AR) dan *Virtual Reality (VR)* berkembang pesat. Ketiga, teknologi tak berawak dan kecerdasan buatan telah berkembang pesat [4].

VI-SLAM umumnya dibagi menjadi dua pendekatan yaitu berbasis-*filtering* dan berbasis-optimalisasi. *Maplab* dan *VINS-mono* adalah tipikal dari kedua metode ini, dan keduanya adalah *open source*. *Maplab* adalah sistem *VI-SLAM* berbasis *filtering* yang juga menyediakan komunitas penelitian dengan koleksi alat pemetaan multi-sesi termasuk penggabungan peta, *loop closure*, dan optimalisasi *visual-inertial*. *VINS-mono* adalah sistem *VI-SLAM* berbasis optimalisasi *real-time* yang menggunakan *sliding window* untuk memberikan odometri presisi tinggi. Selain itu, fitur *IMU* efisien untuk pra-integrasi dengan koreksi bias, inisialisasi estimasi otomatis, kalibrasi ekstrinsik *online*, deteksi kegagalan, dan deteksi *loop* [4].

2.1.3 A Stereo Based Visual-Inertial Odometry for SLAM (2019)

Penelitian yang dilakukan oleh Yong Li dan ShiBing Lang mengembangkan sistem *visual-inertial SLAM* dengan kamera *stereo RGB-D* dan sensor *IMU*. Metode yang digunakan untuk ekstraksi dan *tracking* fitur adalah metode *KLT-sparse Optical Flow*. Sedangkan untuk *close loop* menggunakan metode *Bags of binary Word (DBoW)*.

Hasil dari penelitian ini adalah estimasi posisi dari sistem *visual-inertial SLAM* memiliki tingkat akurasi dan kecepatan proses yang lebih baik dibandingkan dengan sistem *Visual-only SLAM*.

2.1.4 Stereo Visual-Inertial SLAM with Point and Lines (2018)

Penelitian yang dilakukan Oleh Yanqing Liu, Dongdong Yang, Jiamao Li mengembang sistem *visual-inertial SLAM* dengan mengkombinasikan fitur titik, garis, dan sensor *IMU*. Pada saat

menemui teksur yang kurang baik atau pencahayaan yang berubah-ubah, menyebabkan sulitnya deteksi dan *tracking* fitur titik. Kebanyakan gambar yang memiliki teksur kurang baik mungkin mengandung segmen garis terutama pada benda-benda buatan manusia. Ini sangat memungkinkan untuk mengekstrak fitur garis dari gambar tersebut. Penelitian ini membangun *Stereo Visual-Inertial Point Line-SLAM (SVIPL-SLAM)* dengan dibangun berdasarkan *framework ORB-SLAM2*.

Hasil dari penelitian ini adalah kombinasi dari fitur titik dan garis menghasilkan tingkat akurasi yang terbaik meskipun dalam keadaan pencahayaan yang berubah-ubah dan menemui teksur yang kurang baik.

2.1.5 Robust 3D-SLAM With Tight RGB-D-Inertial Fusion (2019)

Penelitian yang dilakukan oleh Xiangyu Deng, Gumin Jin, Meng Wang, Jianxung Li mengembangkan sistem *visual-inertial SLAM* dengan kamera *stereo RGB-D* dan sensor *IMU*. Karena gambar yang *blur*, gerakan yang berlebih, serta kurangnya fitur dari gambar masih menjadi masalah pada *visual-only SLAM*. Metode yang digunakan untuk ekstraksi fitur gambar adalah *FAST feature detector* dan untuk tracking fitur gambar menggunakan *BRIEF descriptor*. Sedangkan untuk *close loop* menggunakan metode *Bags of binary Word 2 (DBoW2)*.

Hasil dari penelitian ini adalah meskipun dalam kondisi gambar yang *blur*, gerakan yang berlebih, dan kurangnya fitur dari gambar, sistem ini masih memiliki tingkat akurasi yang lebih baik dibandingkan dengan *visual-only SLAM* terutama dalam rekonstruksi peta dan estimasi pose.

2.1.6 A Stereo Visual-Inertial SLAM Approach for Indoor Mobile Robots in Unknown Environments Without Occlusions (2019)

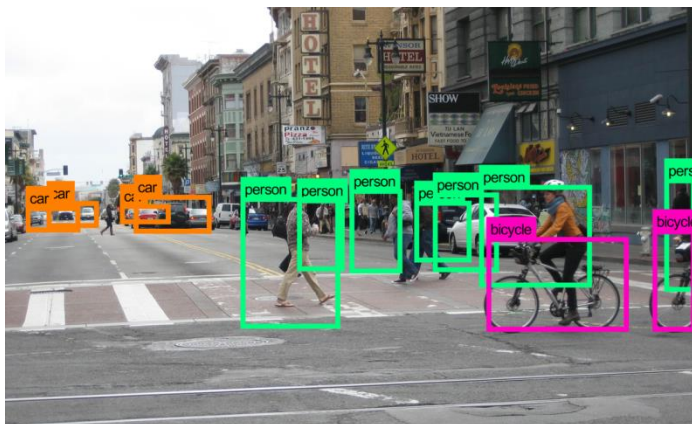
Penelitian yang dilakukan oleh Chang Chen, Huang Zhu, Lei Wang, Yu Liu mengembangkan sistem *visual-inertial SLAM* dengan kamera *stereo RGB-D* dan sensor *IMU* serta dengan metode *matching* satu kali putaran. Metode *matching* satu kali putaran digunakan untuk mengatasi permasalahan *tracking* fitur pada gambar. Paper ini membangun *Sequence of The Circle Matching for the time after space-SLAM (STCM-SLAM)*. Metode yang digunakan untuk ekstraksi fitur gambar adalah *FAST feature detector*. Setelah fitur diekstrak dengan metode *FAST*, selanjutnya fitur tersebut ditentukan menggunakan *circle matching* dan menghilangkan yang tidak sesuai dengan fitur. Untuk *tracking* fitur gambar menggunakan *Forward-Backward bidirectional optical flow Lucas Kanade*.

Hasil dari penelitian ini adalah *STCM-SLAM* memiliki beberapa kelebihan dalam beberapa aspek seperti *scale error*, *running frequency*, *CPU loads*, dan akurasi pada estimasi lintasan.

2.2 Dasar Teori

2.2.1 Computer Vision

Computer vision adalah bidang studi tentang bagaimana komputer dapat melihat dan memahami dari suatu gambar atau video. *Computer vision* mencakup semua tugas yang dilakukan oleh sistem penglihatan secara biologi, termasuk melihat atau merasakan stimulus *visual*, memahami, dan mengekstraksi informasi kompleks ke dalam bentuk yang dapat digunakan oleh proses lain.



Gambar 2. 1 Contoh Penerapan Computer Vision
(Sumber : <https://industrywired.com>)

Secara luas *computer vision* berhubungan dan dapat diterapkan atau dikombinasikan dengan bidang lain seperti *Artificial Intelligence (AI)*, robotika, otomasi industri, pengolahan sinyal, optik fisik, *neurobiology*, dan lain-lain.

2.2.2 Stereo Camera

Stereo camera merupakan tipe dari kamera yang secara umum memiliki 2 sensor kamera, sehingga kamera dapat mensimulasikan penglihatan *binocular* seperti manusia. Dengan adanya sensor kamera yang lebih dari satu, dapat memberikan kemampuan untuk mengolah data dari kedalaman atau *depth*. Salah satu contoh dari *stereo camera* adalah *ZED stereo camera*.



Gambar 2. 2 ZED Stereo Camera
(Sumber : <https://www.stereolabs.com>)

ZED Stereo camera adalah suatu produk yang dikembangkan oleh *stereolabs.com*. Kamera ini memiliki beberapa fitur yang mendukung penelitian terhadap *3D detection* dan *recognition*. Beberapa fitur yang diberikan oleh *ZED Stereo Camera* adalah dapat menangkap atau *capture 3D* video dengan resolusi *2K*, dapat memberikan informasi *depth* atau jarak antara 0.1 meter hingga 20 meter baik pada kondisi *indoor* atau *outdoor*, dapat melakukan deteksi obyek dan *plane* atau suatu bidang untuk melakukan *tracking* pada *movement* dari obyek dalam *3D*, dan dapat melakukan *spatial mapping*.

2.2.3 RGB-D

RGB-D merupakan kombinasi data dari warna *RGB* dengan data *depth (D)* pada setiap *pixel*. Teknologi ini *RGB-D* dikembangkan oleh *PrimeSense* yang sudah dipatenkan dengan nomor paten *US7433024*. *RGB-D* dapat diimplementasikan dengan beberapa *stereo camera* seperti *Microsoft Kinect Camera*, *Asus Xtion Pro*, dan *ZED Stereo Camera*.



Gambar 2. 3 Contoh penerapan RGB-D
(Sumber : <https://vision.in.tum.de>)

Pada *RGB-D*, data *depth* disediakan oleh *sensor infrared* kemudian data tersebut dikolerasikan dengan data *RGB* yang telah dikalibrasi sehingga menghasilkan data *RGB* dengan kedalaman yang terkait pada setiap *pixel*. Dengan adanya sensor dan *RGB-D*, warna menjadi dapat menjadi sebuah fitur. Selain itu, permukaan tiap *point* yang diperkirakan juga akan disimpan dalam *point cloud*.

2.2.4 Point Cloud

Point Cloud merupakan sekumpulan atau dataset dari *point* atau titik yang merepresentasikan data dalam bentuk *3D*. Pada setiap *point* memiliki titik koordinat *X*, *Y*, dan *Z*. *Point cloud* sering digunakan dalam *photogrammetry* atau *remote sensing*. *Photogrammetry* adalah suatu teknik pemetaan melalui foto baik di udara maupun di darat.

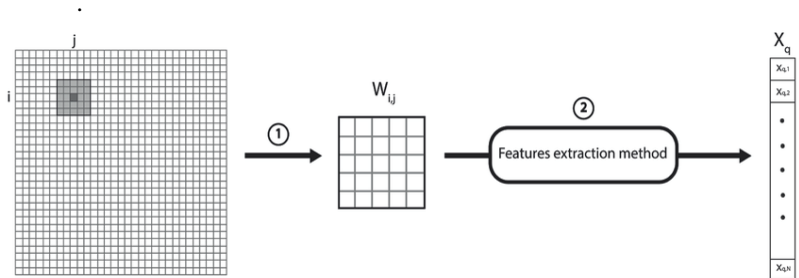


Gambar 2. 4 Point Cloud Library
(Sumber : <https://pointclouds.org>)

Pada *point cloud*, terdapat *library* yang bernama *Point Cloud Library (PCL)*. *Point Cloud Library* adalah *library* yang berdiri sendiri dengan skala yang besar dan terbuka dengan *project 2D* atau *3D* serta *point cloud*. *Point Cloud Library* rilis dibawah ketentuan dari lisensi *BSD* dan gratis digunakan untuk keperluan komersial atau riset.

2.2.5 Feature Extraction

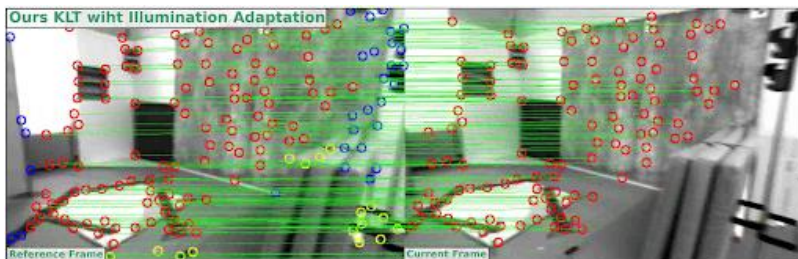
Feature Extraction merupakan suatu pengambilan ciri dari suatu bentuk yang nantinya nilai yang didapatkan akan dianalisis untuk proses selanjutnya. *Feature Extraction* bertujuan untuk mencari daerah fitur yang signifikan pada gambar tergantung pada karakteristik intrinsik dan aplikasinya. Wilayah tersebut dapat didefinisikan dalam lingkungan global atau lokal dan dibedakan oleh bentuk, tekstur, ukuran, intensitas, sifat statistik, dan sebagainya.



Gambar 2. 5 Konsep dari Feature Extraction
(Sumber : Texture Image Segmentation Approach Based on Neural Networks(2018))

2.2.6 Feature Tracking

Feature Tracking merupakan suatu pencocokan ciri antara dua buah gambar atau dua buah *frame*. Ciri dari gambar atau *frame* pertama didapatkan setelah melakukan ekstraksi ciri yang memiliki keunikan dalam suatu titik tertentu dan kemudian mencari kesamaan ciri tersebut di dalam gambar atau *frame* kedua.

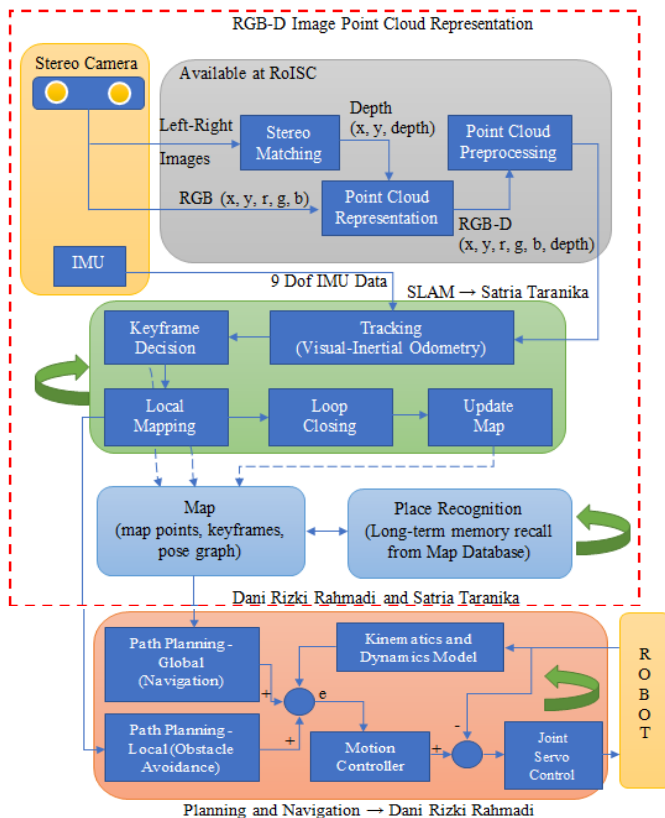


Gambar 2. 6 Contoh Feature Extraction dengan metode Kanade-Lucas-Tomasi
(Sumber : <http://www.luohanjie.com>)

BAB III PERANCANGAN SISTEM

Pada bab 3 ini, dilakukan perancangan sistem keseluruhan dan implementasi sistem yang terdiri dari tahapan-tahapan yang dilakukan dalam penyelesaian proyek akhir ini

3.1 Perancangan Sistem



Gambar 3. 1 Desain Sistem

Gambar 3. 1 Desain Sistem merupakan blok diagram yang menunjukkan bagaimana proses dari sebuah robot dapat melakukan lokalisasi dan pembuatan peta 3D sampai dengan melakukan *path planning* dan *obstacles avoidance* . Namun, pada tugas akhir ini hanya berfokus pada proses lokalisasi dan pembuatan peta 3D. *Input* dari sistem tersebut adalah *streaming* dari kamera *stereo* dan pengukuran sensor *IMU*. Kemudian proses lokalisasi hingga pembuatan peta dilakukan dalam beberapa tahapan proses berikut :

3.1.1 Akuisisi Data dan Inisialisasi

Tahap pertama, pengambilan gambar atau obyek menggunakan *ZED* kamera *stereo*. Sebelum gambar atau obyek yang ditangkap oleh kamera *stereo* diproses, kamera *stereo* akan melakukan *stereo matching* dimana berfungsi sebagai proses untuk mencari *pixel* pada pandangan dari *multiscope* atau lebih dari satu pandangan yang sesuai dengan titik 3D yang sama dalam suatu *scene* atau gambar. Bersamaan dengan itu, kamera *stereo* juga mengambil data *RGB* dan *depth* dari gambar. Data *RGB* dan *depth* yang didapat saat pertama kali kamera dijalankan akan digunakan sebagai *keyframe* pertama. Data *RGB* diperoleh dengan melakukan proses ekstraksi fitur menggunakan metode *Features from Accelerated Segment Test (FAST)*. Persamaan untuk mendapatkan nilai *depth* dari suatu gambar adalah sebagai berikut :

$$\frac{z}{d} = \frac{f}{x_l - x_r} \Rightarrow z = \frac{df}{x_l - x_r} \quad (1)$$

Dimana :

z = nilai *depth* dari gambar.

f = *focal length*.

d = jarak antara lensa kiri dan kanan.

x_l = nilai *disparity* titik (x,y) pada lensa kiri.

x_r = nilai *disparity* titik (x,y) pada lensa kanan.

Tahap kedua, akuisisi data dari sensor *IMU*. Pengukuran *gyroscope* dan *accelerometer* dapat dihitung dengan persamaan berikut :

1. *Gyroscope*

$${}_{\text{mea}}\mathbf{w}(t) = \mathbf{w}^b(t) + \mathbf{b}_g(t) + \boldsymbol{\eta}_g(t) \quad (2)$$

Dimana :

${}_{\text{mea}}\mathbf{w}(t)$ = nilai pengukuran dari *gyroscope*.

$\mathbf{w}^b(t)$ = nilai *real* dari *gyroscope*.

$\mathbf{b}_g(t)$ = *bias* dari *gyroscope*.

$\boldsymbol{\eta}_g(t)$ = *noise* dari *gyroscope*.

2. *Accelerometer*

$${}_{\text{mea}}\mathbf{a}(t) = \mathbf{R}^T(t)(\mathbf{a}(t) - \mathbf{g}) + \mathbf{b}_a(t) + \boldsymbol{\eta}_a(t) \quad (3)$$

Dimana :

${}_{\text{mea}}\mathbf{a}(t)$ = nilai pengukuran dari *accelerometer*.

$\mathbf{R}^T(t)$ = *coordinates transformation*.

$\mathbf{a}(t)$ = nilai *real* dari *accelerometer*.

\mathbf{g} = vektor gravitasi pada *world coordinates*.

$\mathbf{b}_a(t)$ = *bias* dari *accelerometer*.

$\boldsymbol{\eta}_a(t)$ = *noise* dari *accelerometer*.

Setelah data berhasil diakuisisi, proses berikutnya yaitu *pre-integration*. *Pre-integration* digunakan untuk mengintegrasikan rotasi, kecepatan, dan posisi. Proses *pre-integration* dari sensor *IMU* adalah sebagai berikut :

1. Rotasi

$$\mathbf{R}(t + \Delta t) = \mathbf{R}(t) \text{Exp} \left(\left({}_{\text{mea}}\mathbf{w}(t) - \mathbf{b}_g(t) - \boldsymbol{\eta}_{gd}(t) \right) \Delta t \right) \quad (4)$$

2. Kecepatan

$$v(t + \Delta t) = v(t) + g\Delta t + R(t)(_{\text{mea}}a(t) - b_a(t) - \eta_{\text{ad}}(t))\Delta t \quad (5)$$

3. Posisi

$$p(t + \Delta t) = p(t) + v(t)\Delta t + \frac{1}{2}g\Delta t^2 + \frac{1}{2}R(t)(_{\text{mea}}a(t) - b_a(t) - \eta_{\text{ad}}(t))\Delta t^2 \quad (6)$$

Dimana Δt adalah *time interval*, η_{gd} adalah kovariansi waktu diskrit yang diperoleh dari $\frac{1}{\Delta t} \text{Cov}(\eta_g(t))$, dan η_{gd} adalah kovariansi waktu diskrit yang diperoleh dari $\frac{1}{\Delta t} \text{Cov}(\eta_a(t))$. Berdasarkan persamaan (4), (5), dan (6), Maka proses *pre-integration* diantara 2 *keyframe* dapat dituliskan sebagai berikut :

1. Rotasi diantara 2 *keyframe*

$$R_j = R_i \prod_{k=i}^{j-1} \text{Exp}((_{\text{mea}}w_k - b_{gk} - \eta_{\text{gd}k})\Delta t) \quad (7)$$

2. Kecepatan diantara 2 *keyframe*

$$v_j = v_i + g\Delta t_{ij} + \sum_{k=i}^{j-1} R_k(_{\text{mea}}a_k - b_{ak} - \eta_{\text{ad}k})\Delta t \quad (8)$$

3. Posisi diantara 2 *keyframe*

$$p_j = p_i + \sum_{k=i}^{j-1} v_k\Delta t + \frac{1}{2}g\Delta t_{ij}^2 + \frac{1}{2} \sum_{k=i}^{j-1} R_k(_{\text{mea}}a_k - b_{ak} - \eta_{\text{ad}k})\Delta t^2 \quad (9)$$

Dimana :

$$\Delta t_{ij} = \sum_{k=i}^{j-1} \Delta t \text{ .}$$

$$(\cdot)_i = (\cdot)(t_i).$$

Kemudian untuk menghitung nilai *error* dari sensor *IMU* yang dihasilkan setelah melakukan *pre-integration*, dapat didefinisikan dengan persamaan berikut :

$$E_{IMU}(i, j) = p \left([e_R^T \ e_v^T \ e_p^T] \sum_I [e_R^T \ e_v^T \ e_p^T]^T \right) + p \left(e_b^T \sum_R e_b \right) \quad (10)$$

Nilai dari e_R , e_v , e_p , dan e_b masing-masing dapat dihitung dengan persamaan (11), (12), (13), dan (14) berikut :

$$e_R = \text{Log} \left(\left(\Delta R_{ij} \text{Exp}(J_{\Delta R}^g b_{gj}) \right)^T R_i^{bw} R_j^{bw} \right) \quad (11)$$

$$e_v = R_i^{bw} \left(w^{vj} - w^{vi} - g^w \Delta t_{ij} \right) - \left(\Delta v_{ij} + J_g^{\Delta v} b_g^j + J_a^{\Delta v} b_a^j \right) \quad (12)$$

$$e_p = R_{BW}^i \left(w p_B^j - w p_B^i - w v_B^i \Delta t_{ij} - \frac{1}{2} g_w \Delta t_{ij}^2 \right) - \left(\Delta p_{ij} + J_{\Delta p}^g b_g^j + J_{\Delta p}^a b_a^j \right) \quad (13)$$

$$e_b = b^j - b^i \quad (14)$$

Dimana Σ_I adalah bentuk *pre-integration* dari matriks informasi yang diperoleh dari teknik *Mur-Artal and Tardos (2017)*. Σ_R adalah *random walk deviation*. J_a dan J_g adalah *first-order approximations*

dari efek variasi *bias* ketika pra-integrasi tidak secara eksplisit dihitung. Dan P adalah *Huber's weight function*.

3.1.2 Point Cloud Representation

Pada tahap ketiga, data yang telah diperoleh berupa *RGB*, *depth*, dan posisi x dan y dari kamera selanjutnya akan masuk kedalam ruang dataset yang bernama *point cloud representation*. *Point cloud representation* adalah sekumpulan data yang mempresentasikan obyek yang ditangkap dengan koordinat geometri titik x , y , dan z atau *depth* dari suatu permukaan obyek. Hasil dari *point cloud representation* akan diproses dalam *point cloud preprocessing*.

3.1.3 Feature Tracking

Tahap keempat, proses *tracking* fitur. Proses *tracking* fitur ini dilakukan dengan menggabungkan data dari *3D point cloud* dan data dari sensor *IMU*. Proses *tracking* ini menggunakan metode *Forward-Backward Lucas Kanade Optical Flow*. Dasar dari metode ini adalah *Lucas-Kanade Optical Flow* yang dilakukan dengan *Forward-Backward* untuk menghindari kegagalan *tracking* saat robot melakukan gerakan yang sangat cepat.

Metode *Optical Flow* memiliki asumsi yang kuat, yaitu tingkat kecerahan yang diukur dari titik yang sama di setiap perspektif diasumsikan tetap sama. Asumsi ini gagal saat kamera menyesuaikan dengan kecerahan *scene* untuk eksposur otomatis. Maka diperlukan parameter untuk mengatasi perubahan pencahayaan tersebut. *Fotometric residual* diantara gambar I_{k-1} dan I_k adalah sebagai berikut :

$$r_u^I(\xi) = \alpha I_{k-1}(u) + \beta - I_k(u') \quad (15)$$

Nilai α dan β didapatkan dari persamaan (16) dan (17) berikut :

$$\alpha = \frac{\sum_{u \in \Omega_L} I_1^1(u) I_2^1(u')}{\sum_{u \in \Omega_L} I_2^1(u') I_2^1(u)} \quad (16)$$

$$\beta = \frac{1}{|\Omega_L|} \sum_i \left(I_1^1(u') - \alpha I_2^1(u) \right) \quad (17)$$

Dimana :

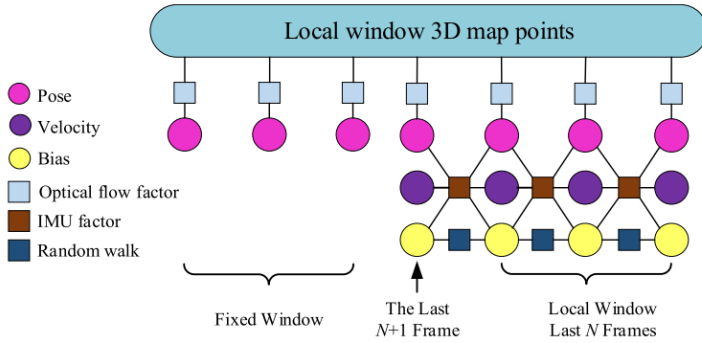
$\Omega_L = \text{set of interior point.}$

$\xi = \text{transisi dari gambar } I_k \text{ ke gambar } I_{k-1}.$

$$u = \pi \left(T_{\xi} \pi^{-1}(u, D_1(u)) \right).$$

3.1.4 Keyframe Decision

Tahap kelima, proses penentuan *keyframe*. Proses penentuan *keyframe* ini bertujuan untuk melakukan pengenalan terhadap lingkungan di sekitar robot dengan mencari kecocokan antara *frame* sebelumnya yang telah ditangkap oleh kamera dan *frame* saat yang sedang dilihat saat ini. Proses penentuan *keyframe* ini menggunakan metode *Sliding Window Filtering*. *Sliding Window* berfungsi untuk membandingkan status dari *IMU* dan kamera yang telah ditangkap sebelumnya dengan keadaan saat ini. *Sliding window* akan diberikan batas tertentu berapa *frame* yang dapat ditampung. Misalkan diberikan batasan lima *frame*. Ketika sudah mencapai lima *frame* dan ada *frame* keenam berupa *frame* saat ini, maka antara *frame* kelima dan keenam akan dicari kecocokannya berdasarkan *keyframe*. Jika *frame* kelima dan keenam memiliki kecocokan, maka *frame* yang paling awal akan dihapus. Jika *frame* kelima dan keenam tidak memiliki kecocokan, maka *frame* kelima akan dihapus dan digantikan dengan *frame* keenam. Ilustrasi dari *sliding window* adalah sebagai berikut :



Gambar 3. 2 Ilustrasi Sliding Window
(Sumber : A Stereo Visual-Inertial SLAM Approach for Indoor Mobile Robots in Unknown Environments Without Occlusions (2019))

Pose frame saat ini dioptimalkan dengan meminimalkan kesalahan proyeksi fitur dari kamera dan kesalahan *IMU*. Dapat didefinisikan dengan persamaan berikut :

$$\mu = \arg \min_{\mu} \left(\sum r_u^l(\xi) + E_{IMU}(i, j) \right) \quad (18)$$

Dimana $\sum r_u^l(\xi)$ adalah persamaan *error* dari kamera yang telah didapatkan pada persamaan (10). $E_{IMU}(i, j)$ adalah persamaan *error* dari sensor *IMU* yang telah didapatkan pada persamaan (15). Setelah *sliding window* selesai , maka akan dihasilkan sebuah *local map* yang nantinya dapat digunakan untuk kebutuhan lain seperti *obstacle avoidance*.

3.1.5 Loop Closing

Tahap keenam, proses *loop closing*. *Loop closing* digunakan untuk mengurangi akumulasi *error* ketika sebuah robot kembali ke

tempat yang sudah dipetakan. Metode yang digunakan pada proses *loop closing* ini adalah dengan model *Bags-of-Word (DBoW)*. *DboW* merupakan sebuah *database image* yang digunakan untuk *image retrieval* sehingga memungkinkan peta baru dapat digabungkan ke peta yang dibangun sebelumnya.

Terakhir, *output* dari sistem tersebut menghasilkan sebuah peta *3D*. Yang dapat disimpan dalam *database* dan dapat digunakan sebagai referensi untuk melakukan *global path planning* dan *obstacles avoidance*.

-----Halaman ini sengaja dikosongkan-----

BAB IV PENGUJIAN DAN ANALISA

Dalam bab ini dijelaskan mengenai pengujian terhadap sistem yang telah dibuat untuk mengetahui apakah sistem dapat bekerja sesuai perancangan.

4.1 Perangkat dan Lingkungan Uji Coba

Sistem dibuat dan diimplementasikan menggunakan *Personal Computer*(PC). Pengujian dilakukan dengan mempertimbangkan komponen-komponen yang digunakan dalam pembuatan sistem. Berikut ini spesifikasi perangkat keras dan perangkat lunak yang ditunjukkan pada tabel 4.1.

Tabel 4. 1 Spesifikasi Perangkat Pengujian

Perangkat Keras dan Perangkat Lunak		
No.	Deskripsi	Spesifikasi
1.	Processor	Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz (4 CPUs), ~2.7GHz
2.	RAM	12.00 GB
3.	Sistem Operasi	Ubuntu 18.04.5 LTS
4.	Graphic Card	NVIDIA Geforce 930MX
5.	Software Build	Microsoft Visual Studio Code
6.	Library	OpenCV 3.1
		PCL-1.8

Pengujian sistem dilakukan untuk mengetahui sejauh mana keberhasilan sistem yang telah dibuat. Terdapat 2 tahap yang telah selesai dilakukan pengujian yaitu :

- 1) Tahap Ekstraksi Fitur
- 2) Tahap Representasi dalam 3D *Point Cloud*

4.2 Laporan Sistem

Tabel 4.2 merupakan tahapan yang telah dilakukan hingga progress akhir ini.

Tabel 4. 2 Tahapan yang Telah Dilakukan

Tahapan Pengerjaan Proyek Akhir					
No.	Tahapan	Status	Bobot	Selesai(%)	Keterangan
1.	Melakukan stereo matching atau kalibrasi pada kamera stereo.	0	0.05	0%	Belum dapat melakukan uji coba dengan kamera <i>ZED</i> .
2.	Melakukan ekstraksi fitur gambar yang dihasilkan oleh kamera <i>stereo</i> dan melakukan akuisi data dari sensor <i>IMU</i> .	70	0.2	14%	Gambar yang di uji berasal dari dataset <i>RGB-D</i>
3.	Pembuatan program untuk memproses hasil ekstraksi fitur dari kamera <i>stereo</i>	80	0.2	16%	Gambar yang di uji berasal dari dataset <i>RGB-D</i>

	dengan <i>point cloud processing</i> hingga dihasilkan sebuah dataset <i>3D point cloud</i> .				
4.	Pembuatan program untuk melakukan proses <i>tracking</i> berdasarkan <i>3D point cloud</i> dan data dari sensor <i>IMU</i> .	0	0.15	0%	Belum dapat melakukan uji coba dengan kamera <i>ZED</i> .
5.	Pembuatan program untuk <i>keyframe selection</i> sehingga dapat menghasilkan sebuah <i>local map</i> .	0	0.2	0%	Belum dapat melakukan uji coba dengan kamera <i>ZED</i> .
6.	Pembuatan program untuk <i>loop closing</i> sehingga menghasilkan <i>map</i> dan menyimpan <i>map</i> tersebut ke dalam	0	0.2	0%	Belum dapat melakukan uji coba dengan kamera <i>ZED</i> .


	<i>database.</i>				
--	------------------	--	--	--	--


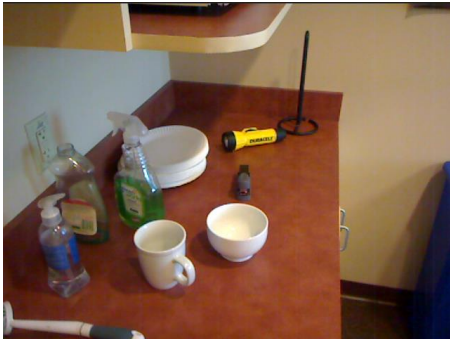
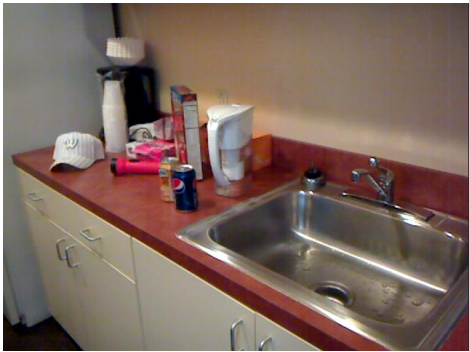
4.3 Tahap Ekstraksi Fitur

Pada tahap ini dilakukan proses ekstraksi fitur gambar dengan metode *Features from Accelerated Segment Test* (FAST). Pengujian dilakukan dengan memproses gambar *RGB* yang diperoleh dari dataset yang di unduh pada halaman website <https://rgb-dataset.cs.washington.edu/dataset/rgb-d-scenes/>. Dataset yang diuji merupakan sebuah gambar yang diambil didalam ruangan tertutup.


Gambar *RGB* dan hasil ekstraksi fitur dari gambar *RGB* disajikan dalam tabel 4.3 dan tabel 4.4 berikut :

Tabel 4.3 Gambar *RGB*


No.	Gambar <i>RGB</i>
1.	
2.	


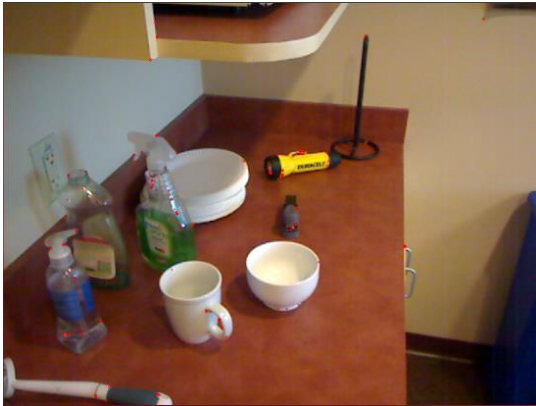
			
3.			
4.			


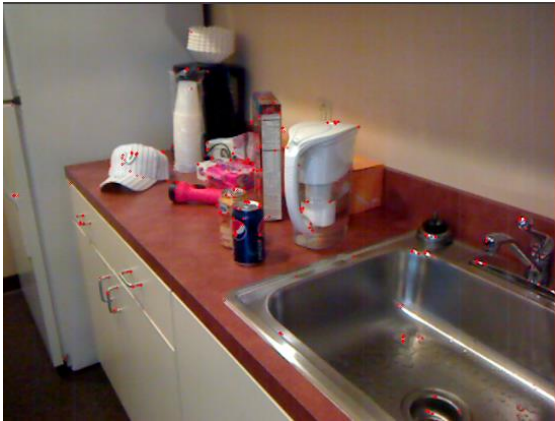
5.

A photograph of a kitchen counter and sink area. The counter is a reddish-brown laminate. On the counter, from left to right, there is a white baseball cap, a black coffee maker with a white pitcher on top, a box of cereal, a white water filter pitcher, a blue can of soda, and a pink object. To the right of the counter is a stainless steel double sink with a chrome faucet. The background is a plain, light-colored wall.

Tabel 4.4 Hasil Ekstraksi Fitur

No.	Hasil Ekstraksi Fitur
1.	
2.	

	
3.	
4.	

	
5.	

Pada pengujian ekstraksi fitur tersebut, pada tabel 4.4 terdapat tanda berupa lingkaran kecil berwarna merah yang merupakan hasil dari proses ekstraksi fitur. Fitur yang diekstraksi berupa sudut atau *corner*. Pada gambar tersebut, masih terdapat beberapa sudut yang


belum terdeteksi sebagai *corner*. Hal ini dapat diakibatkan oleh kurangnya pencahayaan ataupun *blur* pada gambar yang diproses.


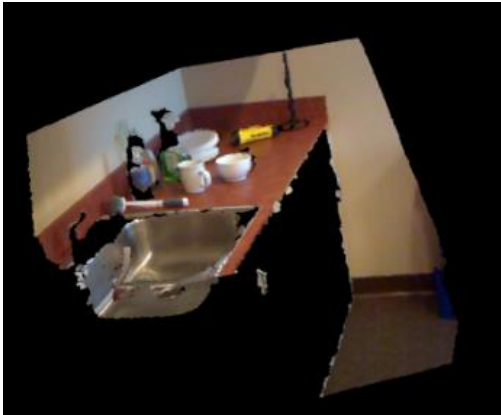
4.4 Tahap Representasi dalam 3D Point Cloud


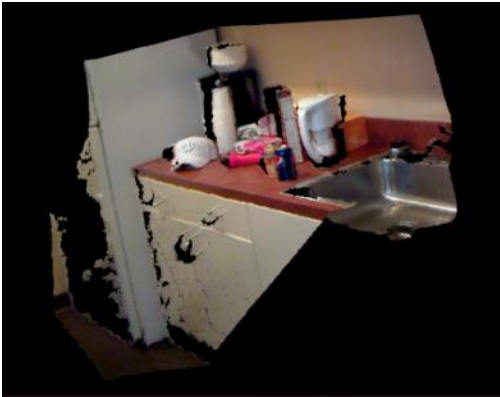
Pada tahap ini dilakukan proses representasi ke dalam bentuk *3D Point Cloud*. Pengujian dilakukan dengan memproses gambar *RGB* dan gambar *Depth* yang diperoleh dari dataset yang di unduh pada website <https://rgb-dataset.cs.washington.edu/dataset/rgb-scenes/>. Dataset yang diuji merupakan sebuah gambar yang diambil didalam ruangan tertutup.

Hasil representasi dalam bentuk *3D Point Cloud* disajikan dalam tabel 4.5 berikut :

Tabel 4.5 Hasil *3D Point Cloud*

No.	Hasil <i>3D Point Cloud</i>
1.	
2.	

	
3.	
4.	

	
5.	

Pada pengujian representasi *point cloud* tersebut, hasil yang didapatkan belum sempurna kemungkinan disebabkan karena dataset tersebut diambil dengan menggunakan kamera stereo selain *ZED* kamera. Sehingga banyak data yang seharusnya dapat direpresentasikan ke dalam *3D Point Cloud* menjadi tidak terlihat.

-----Halaman ini sengaja dikosongkan-----

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisa yang telah dilakukan maka dapat diambil kesimpulan sebagai berikut:

1. Sistem dapat melakukan ekstraksi fitur pada gambar *RGB*, akan tetapi hasil yang didapatkan dipengaruhi oleh tingkat pencahayaan serta tingkat kejelasan dari gambar. Dan perlu dilakukan uji coba dengan kamera *ZED* untuk mengetahui apakah program sudah dapat diterapkan pada kondisi *real*.
2. Sistem dapat merepresentasikan gambar *2D* menjadi bentuk *3D Point Cloud*. Dan perlu dilakukan uji coba dengan kamera *ZED* untuk mengetahui apakah program sudah dapat diterapkan pada kondisi *real*.

5.2 Hal yang dilakukan selanjutnya

Pada bagian ini dijelaskan mengenai tahapan atau pekerjaan selanjutnya yang harus diselesaikan. Adapun tahapan akan dikerjakan antara lain:

1. Menguji program yang telah selesai dibuat dengan menggunakan kamera *ZED*.
2. Pembuatan program untuk melakukan proses *tracking* berdasarkan *3D point cloud* dan data dari sensor *IMU*.
3. Pembuatan program untuk *keyframe selection* sehingga dapat menghasilkan sebuah *local map*.
4. Pembuatan program untuk *loop closing* sehingga menghasilkan *map* dan menyimpan *map* tersebut ke dalam *database*.

-----Halaman ini sengaja dikosongkan-----

DAFTAR PUSTAKA

- [1] C. Chen, H. Zhu, L. Wang and Y. Liu, "A Stereo Visual-Inertial SLAM Approach for Indoor Mobile Robots in Unknown Environments Without Occlusions," in IEEE Access, vol. 7, pp. 185408-185421, 2019.
- [2] Smith, Randall & Cheeseman, Peter. (1987). *On the Representation and Estimation of Spatial Uncertainty. The International Journal of Robotics Research.* 5. 10.1177/027836498600500404.
- [3] Smith, Randall & Self, Matthew & Cheeseman, Peter. (1986). Estimating Uncertain Spatial Relationships in Robotics. Autonomous Robot Vehicles. 1. 435-461. 10.1109/ROBOT.1987.1087846.
- [4] Chang, Chen & Zhu, Hua & Li, Menggang & You, Shaoze. (2018). A Review of Visual-Inertial Simultaneous Localization and Mapping from Filtering-Based and Optimization-Based Perspectives. Robotics. 7. 45. 10.3390/robotics7030045.
- [5] Y. Li and S. Lang, "A Stereo-Based Visual-Inertial Odometry for SLAM," 2019 Chinese Automation Congress (CAC), Hangzhou, China, 2019, pp. 594-598.
- [6] Y. Liu, D. Yang, J. Li, Y. Gu, J. Pi and X. Zhang, "Stereo Visual-Inertial SLAM With Points and Lines," in IEEE Access, vol. 6, pp. 69381-69392, 2018.
- [7] X. Deng, G. Jin, M. Wang and J. Li, "Robust 3D-SLAM With Tight RGB-D-Inertial Fusion," 2019 Chinese Control Conference (CCC), Guangzhou, China, 2019, pp. 4389-4396, doi: 10.23919/ChiCC.2019.8866265.

- [8] Chang, Chen & Zhu, Hua. (2018). Visual-inertial SLAM method based on optical flow in a GPS-denied environment. *Industrial Robot*. 45. 10.1108/IR-01-2018-0002.
- [9] Forster, Christian & Carlone, Luca & Dellaert, Frank & Scaramuzza, Davide. (2015). IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation. 10.15607/RSS.2015.XI.006.