**PEMBANGUNGAN** **SISTEM MULTI-KAMERA UNTUK ESTIMASI LETAK 3D PENANDA PADA SISTEM TANGKAP GERAK SEDERHANA**

**Laporan Tugas Akhir - Capstone**

**Disusun sebagai syarat kelulusan tingkat sarjana**

**Oleh**

**Moses Ananta**

**NIM : 13519076**



**PROGRAM STUDI TEKNIK INFORMATIKA**

**SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA**

**INSTITUT TEKNOLOGI BANDUNG**

**Juli 2023**

**PEMBANGUNGAN SISTEM MULTI-KAMERA UNTUK ESTIMASI LETAK 3D PENANDA PADA SISTEM TANGKAP GERAK SEDERHANA**

**Laporan Tugas Akhir**

**Oleh**

**Moses Ananta**

**NIM : 13519076**

**Program Studi Teknik Informatika**

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung

Telah disetujui dan disahkan sebagai Laporan Tugas Akhir

di Bandung, pada tanggal 11 Juli 2023

Pembimbing,

Nugraha Priya Utama, S.T, M.A., Ph.D.

NIP 118110074

**LEMBAR IDENTITAS**

**TUGAS AKHIR CAPSTONE**

Judul Proyek TA : Pembangunan Sistem Tangkap Gerak Multi-Kamera dengan Penanda sebagai Pembanding Penilaian Akurasi Sistem Tangkap Gerak Monokuler Tanpa Penanda

Anggota Tim dan Pembagian Peran:

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **NIM** | **Nama** | **Peran** |
| 1 | 13519022 | Jose Galbraith Hasintongan | Membuat Mocap berbasis AI untuk gerak sederhana |
| 2 | 13519076 | Moses Ananta | Pembangunan sistem multi-kamera, kalibrasi dan perbaikan distorsi kamera, dan konfigurasi serta estimasi 3D penanda |
| 3 | 13519093 | Wisnu Aditya Samiadji | Integrasi Sistem Tangkap Gerak dengan Avatar Digital |
| 4 | 13519149 | Syihabuddin Yahya Muhammad | object detection dan object tracking |

Bandung,  11 Juli 2023

Mengetahui,

|  |
| --- |
| Pembimbing I,    Nugraha Priya Utama, S.T, M.A., Ph.D. |

NIP. 118110074

**LEMBAR PERNYATAAN**

Dengan ini saya menyatakan bahwa:

1. Pengerjaan dan penulisan Laporan Tugas Akhir ini dilakukan tanpa menggunakan bantuan yang tidak dibenarkan.
2. Segala bentuk kutipan dan acuan terhadap tulisan orang lain yang digunakan di dalam penyusunan laporan tugas akhir ini telah dituliskan dengan baik dan benar.
3. Laporan Tugas Akhir ini belum pernah diajukan pada program pendidikan di perguruan tinggi mana pun.

Jika terbukti melanggar hal-hal di atas, saya bersedia dikenakan sanksi sesuai dengan Peraturan Akademik dan Kemahasiswaan Institut Teknologi Bandung bagian Penegakan Norma Akademik dan Kemahasiswaan khususnya Pasal 2.1 dan Pasal 2.2.

Bandung, 11 Juli 2023



Moses Ananta

NIM 13519076

ABSTRAK

**PEMBANGUNAN SISTEM MULTI-KAMERA UNTUK ESTIMASI LETAK 3D PENANDA PADA SISTEM TANGKAP GERAK SEDERHANA**

Oleh

MOSES ANANTA

NIM : 13519076

Dengan semakin berkembangnya teknologi berbasis kecerdasan buatan (AI), biaya produksi dan kompleksitas dalam membuat sebuah sistem tangkap gerak menjadi lebih murah, sederhana, dan terjangkau bagi industri-industri pembuat konten kecil. Namun, dikarenakan proses estimasi gerakan dari sistem tangkap gerak tersebut tidak berdasarkan penanda, seperti pada sistem tangkap gerak konvensional, dan hanya memperkirakan titik-titik penting tubuh aktor berdasarkan bagian tubuh yang terlihat, akurasi dari sistem tangkap gerak berbasis AI dapat menjadi sebuah permasalahan untuk industri yang menginginkan sistem tangkap geraknya menghasilkan hasil yang akurat tanpa harus mengeluarkan biaya yang besar seperti yang terdapat pada sistem tangkap gerak konvensional.

Berdasarkan permasalahan tersebut, akan dibuat sebuah versi murah dan sederhana dari sistem tangkap gerak konvensional yang memanfaatkan penanda berbasis warna dan beberapa kamera untuk menangkap gerakan aktor. Fokus dari tugas akhir ini adalah membangun sistem multi-kamera yang akan digunakan untuk mengestimasi koordinat 3 dimensi dari penanda-penanda yang terdeteksi menggunakan teknik triangulasi yang kemudian dilakukan optimasi dengan metode *bundle adjustment*.

Sistem multi-kamera yang dibangun diuji dengan 4 rentang rotasi antar kamera yang berbeda. Dari hasil percobaan dan analisis terhadap sistem yang dibuat didapati bahwa penggunaan rentang rotasi yang rendah, triangulasi dengan 3 kamera sekaligus, dan penggunaan *bundle adjustment* berpengaruh besar dalam mempertahankan struktur 3D yang terbentuk dan mengurangi *reprojection error.*

Kata kunci: sistem tangkap gerak, kalibrasi multi-kamera, triangulasi, *bundle adjustment*.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas berkat dan anugerah-Nya yang senantiasa menyertai, menolong, dan memberkati penulis dalam segala proses pembuatan hingga penyelesaian Laporan tugas akhir dengan judul “Pembangunan Sistem Multi-Kamera Untuk Estimasi Letak 3D Penanda Pada Sistem Tangkap Gerak Sederhana”.

Penulis menyadari sepenuhnya bahwa penulisan laporan tugas akhir ini tidak karena kemampuan dan pengetahuan penulis sendiri, tetapi berkat dukungan, bimbingan, doa, dan bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Bapak Nugraha Priya Utama, S.T, M.A., Ph.D. selaku dosen pembimbing tugas akhir yang telah membimbing penulis selama seluruh proses pembuatan tugas akhir ini,
2. Orang tua dan saudara, yang dengan penuh kasih dan sayang telah memberikan dorongan, doa, dan bantuan tiada putusnya,
3. Teman-teman sekelompok tugas akhir ini yang telah bersama berjuang untuk menyelesaikan tugas akhir ini bersama-sama,
4. Para dosen di Jurusan Teknik Informatika yang telah memberikan ilmu dan dasar teoritis dalam peranannya selaku pengajar
5. Saudara – saudara seiman yang telah banyak mendukung dalam doa,
6. Teman-teman lain yang tidak dapat penulis sebutkan satu persatu.

Bandung, 11 Juli 2023

Penulis

DAFTAR ISI

[BAB I PENDAHULUAN 1](#_Toc139993334)

[I.1 Latar Belakang 1](#_Toc139993335)

[I.2 Rumusan Masalah 2](#_Toc139993336)

[I.3 Tujuan 3](#_Toc139993337)

[I.4 Batasan Masalah 3](#_Toc139993338)

[I.5 Metodologi 4](#_Toc139993339)

[I.6 Sistematika Pembahasan 5](#_Toc139993340)

[BAB II STUDI LITERATUR 6](#_Toc139993341)

[II.1 Pembentukan Gambar Pada Kamera 6](#_Toc139993342)

[II.2 Distorsi Lensa 9](#_Toc139993343)

[II.3 Kalibrasi Kamera Tunggal 12](#_Toc139993344)

[II.4 Triangulasi 13](#_Toc139993345)

[II.5 Kalibrasi Stereo 17](#_Toc139993346)

[II.6 Multi-Kamera 18](#_Toc139993347)

[*II.7* *Reprojection Error* 19](#_Toc139993348)

[II.8 Bundle Adjustment 20](#_Toc139993349)

[II.9 OpenCV 20](#_Toc139993350)

[BAB III PENGEMBANGAN SISTEM MULTI-KAMERA 21](#_Toc139993351)

[III.1 Analisis Persoalan 21](#_Toc139993352)

[III.2 Analisis Solusi 22](#_Toc139993353)

[III.3 Rancangan Solusi 26](#_Toc139993354)

[III.3.1 Kalibrasi Kamera Tunggal 26](#_Toc139993355)

[III.3.2 Penghilangan Distorsi Gambar 28](#_Toc139993356)

[III.3.3 Pemilihan Kamera Utama 28](#_Toc139993357)

[III.3.4 Kalibrasi Kamera Stereo 29](#_Toc139993358)

[III.3.5 Estimasi 3D Penanda dengan Triangulasi 30](#_Toc139993359)

[III.3.6 Optimasi Hasil Estimasi 3D dengan *Bundle Adjustment* 33](#_Toc139993360)

[BAB IV EVALUASI SISTEM 34](#_Toc139993361)

[IV.1 Kalibrasi Kamera Tunggal dan Penghilangan Distorsi Gambar 34](#_Toc139993362)

[IV.2 Kalibrasi Stereo 37](#_Toc139993363)

[IV.3 Estimasi 3D Dan Optimasi Hasil Estimasi 3D 39](#_Toc139993364)

[BAB V KESIMPULAN DAN SARAN 48](#_Toc139993365)

[V.1 Kesimpulan 48](#_Toc139993366)

[V.2 Saran 49](#_Toc139993367)

DAFTAR GAMBAR

[Gambar II.1. Model Kamera Lubang Jarum (Bradski & Kaehler, 2008) 6](#_Toc139993685)

[Gambar II.2. Dampak Bukaan Kamera Terhadap Kejelasan Gambar (Hata & Savarese) 9](#_Toc139993686)

[Gambar II.3. Penggunaan Lensa Untuk Memperjelas Gambar (Hata & Savarese) 10](#_Toc139993687)

[Gambar II.4. Jenis-Jenis Distorsi Kamera (Hata & Savarese) 11](#_Toc139993688)

[Gambar II.5. Distorsi Tangensial (Kaehler & Bradski, 2017) 12](#_Toc139993689)

[Gambar II.6. Permasalahan proyeksi balik ke 3D dengan hanya 1 *view* saja (Kitani, Triangulation, 2007) 14](#_Toc139993690)

[Gambar II.7. Gambaran Triangulasi dengan 2-*view* (Kitani, Epipolar Geometry, 2007) 15](#_Toc139993691)

[Gambar II.8. 3 *View* Melihat Titik Yang Sama Pada Objek (Snavely & Li, 2018) 18](#_Toc139993692)

[Gambar II.9. Visualisasi *Error* Estimasi 3D. Gambar (a) Estimasi 3D Dengan Hanya 2 *View*. Gambar (b) Estimasi 3D Dengan Hanya 4 *View* (Snavely & Li, 2018) 19](#_Toc139993693)

[Gambar III.1. Pola Kalibrasi Papan Catur yang Digunakan. 27](#_Toc139993694)

[Gambar III.2. Gambaran Urutan Kamera Yang Digunakan 29](#_Toc139993695)

[Gambar III.3. Gambar Pola Kalibrasi Diambil Secara Bersamaan Oleh Kamera Kiri (a), Kamera Tengah (b), dan Kamera Kanan (c). 30](#_Toc139993696)

[Gambar III.4. Kondisi Ruangan. Bagian Merah mengindikasikan daerah ruangan yang tidak boleh terlihat oleh kamera. 31](#_Toc139993697)

[Gambar III.5. Konfigurasi Peletakan Kamera Untuk Setiap Rentang Rotasi 32](#_Toc139993698)

[Gambar IV.1. Gambar (a) gambar asli sebelum dilakukan perbaikan distorsi. Gambar (b) gambar setelah dilakukan perbaikan distorsi. 35](#_Toc139993699)

[Gambar IV.2. Gambar (a) gambar asli sebelum dilakukan perbaikan distorsi. Gambar (b) gambar setelah dilakukan perbaikan distorsi. 36](#_Toc139993700)

[Gambar IV.3. *Error* Proyeksi Dari Proses Kalibrasi Kamera Tunggal. 38](#_Toc139993701)

[Gambar IV.4. Gambar (a) urutan deteksi dari kiri ke kanan dimulai dengan warna ungu ke merah. Gambar (a) urutan deteksi dari kiri ke kanan dimulai dengan warna merah ke ungu. 39](#_Toc139993702)

[Gambar IV.5. gambaranaktor dari *view* dan rotasi yang berbeda-beda 41](#_Toc139993703)

[Gambar IV.6. Gambar (a),(b),(c),dan (d) secara berturut-turut adalah hasil struktur 3D dari proses triangulasi untuk pasangan kamera kiri-tengah, kanan-tengah, kiri-kanan, dan 3 kamera sekaligus. 44](#_Toc139993704)

[Gambar IV.7. Gambar (a) struktur 3D kanan-tengah sebelum optimasi. Gambar (b) struktur 3D kanan-tengah setelah optimasi. 45](#_Toc139993705)

[Gambar IV.8. Salah satu titik yang dihasilkan tidak sesuai dengan struktur 3D aslinya. 47](#_Toc139993706)

DAFTAR TABEL

[Tabel III.1. Koordinat 3D Tepi Dalam Petak-Petak Papan Catur 28](#_Toc139993390)

[Tabel IV.1. Parameter Intrinsik dan Distorsi Lensa Hasil Kalibrasi Kamera Tunggal 36](#_Toc139993391)

[Tabel IV.2. Parameter Intrinsik dan Distorsi Lensa Hasil Kalibrasi Kamera Tunggal 36](#_Toc139993392)

[Tabel IV.2. Tabel Rotasi dan Translasi Kamera 40](#_Toc139993393)

[Tabel IV.3. Tabel Rata-rata Kesalahan Proyeksi Balik 41](#_Toc139993394)

# PENDAHULUAN

## Latar Belakang

Sistem tangkap gerak merupakan sistem yang menangkap gerakan dari aktor gerakan dan menerjemahkan gerakan yang ditangkap tersebut ke dalam suatu karakter pada dunia digital untuk dapat dilihat visualisasinya. Sistem ini sangat berguna untuk industri animasi baik untuk bidang hiburan atau film maupun untuk bidang gim 3 dimensi karena dapat meringankan kerja dari para pembuat animasi sehingga mereka tidak perlu menganimasikan gerakan setiap poin penting yang terdapat pada karakter digital secara manual tetapi teknisi dapat merekam gerakan aktor dan memproyeksikan gerakan tersebut kepada karakter digital kemudian pembuat animasi dapat memperbaiki beberapa poin penting dalam animasi hasil proyeksi yang menurutnya janggal.

Untuk menghemat biaya produksi dan implementasi sistem tangkap gerak, khususnya untuk perusahaan kecil, sistem tangkap gerak yang digunakan adalah yang berbasis AI (*Artificial Intelligence*) untuk mengestimasi keberadaan poin-poin penting pada tubuh aktor kemudian merekam pergerakan berdasarkan hasil estimasi tersebut. Dengan sistem ini, biaya produksi dapat dikurangi secara signifikan karena dihilangkan komponen penanda dan peralatan lainnya yang umumnya digunakan oleh sistem tangkap gerak profesional digantikan dengan kamera yang dapat berupa kamera amatir atau bahkan kamera biasa yang terdapat pada ponsel pintar.

Kendala utama yang muncul dari penggunaan sistem tangkap gerak berbasis AI ini adalah proses penangkapan poin-poin penting pada tubuh aktor hanyalah merupakan estimasi dikarenakan sistem tidak mengetahui letak pasti dari poin-poin penting tersebut dan hanya mengira-ngira di mana poin-poin tersebut mungkin berada. Masalah ini adalah masalah yang penting karena sistem ini tidak dapat dipastikan apakah benar-benar menangkap gerakan nyata secara akurat.

Untuk mengetahui akurasi dari sistem tangkap gerak berbasis AI tersebut, maka diperlukan sistem tangkap gerak pembanding yang lebih akurat untuk membandingkan perbedaan gerakan yang dihasilkan kedua sistem dan menilai akurasi estimasi gerakan dari sistem tangkap gerak berbasis AI terhadap gerakan dari sistem tangkap gerak ini. Salah satu sistem yang menghasilkan tangkap gerak yang akurat adalah sistem tangkap gerak berbasis penanda yang merupakan standar bagi industri perfilman. Namun sistem tangkap gerak tersebut memerlukan peralatan-peralatan canggih seperti baju penanda, peralatan multi-kamera yang harganya umumnya tidak terjangkau oleh industri-industri kecil. Dibutuhkan suatu sistem tangkap gerak sederhana yang biaya produksinya lebih terjangkau serta memiliki akurasi yang tidak jauh berbeda dari sistem profesional sehingga dapat dijadikan sebagai pembanding untuk menilai akurasi sistem tangkap gerak berbasis AI.

## Rumusan Masalah

Saat ini, sistem tangkap gerak berbasis AI hanya dapat mengestimasi posisi titik- titik penting dari tubuh aktor berdasarkan masukan gambar atau video yang tidak bisa dijamin akurasinya. Untuk mengetahui seberapa akurat sistem tangkap berbasis AI tersebut, diperlukan sistem tangkap gerak pembanding yang lebih akurat daripada sistem tangkap gerak AI seperti sistem tangkap gerak berbasis penanda yang banyak digunakan di industri perfilman. Namun meskipun lebih akurat, sistem tersebut memiliki biaya produksi yang mahal dan peralatan multi-kamera yang canggih yang hanya dapat dijangkau oleh industri-industri profesional. Dari permasalahan tersebut, rumusan masalah keseluruhan tugas akhir *capstone* ini adalah:

“Pembangunan sistem multi-kamera untuk sistem tangkap gerak sederhana berbasis penanda warna sebagai pembanding untuk menilai akurasi dari hasil estimasi gerakan sistem tangkap gerak berbasis AI.”

Adapun sistem yang dihasilkan dapat memproyeksikan gerakan yang ditangkapnya ke dalam suatu karakter digital dan dapat menghitung perbedaan gerakan dari karakter digital yang dibuat oleh sistem ini dengan gerakan dari karakter digital yang diperoleh dari sistem tangkap gerak berbasis AI.

Untuk dapat membangun sistem keseluruhan seperti yang tertera pada permasalahan di atas, salah satu komponen sistem yang penting dan yang menjadi pokok permasalahan pada tugas akhir ini adalah sistem multi-kamera yang berfungsi untuk merekam gerakan aktor dan melakukan estimasi letak 3D penanda titik-titik penting. Dengan demikian, permasalahan yang akan diselesaikan pada tugas akhir ini adalah:

“Pembangunan sistem multi-kamera untuk mengestimasi letak 3D penanda titik-titik penting aktor pada sistem tangkap gerak sederhana.”

## Tujuan

Berdasarkan rumusan masalah yang telah disampaikan sebelumnya, tujuan dari tugas akhir ini adalah membangun sebuah sistem multi-kamera yang dapat melakukan estimasi letak 3D penanda titik-titik penting aktor.

## Batasan Masalah

Agar membatasi lingkup dari permasalahan yang telah disampaikan, didefinisikan beberapa batasan masalah sebagai berikut,

1. Jumlah kamera yang digunakan pada pembangunan sistem berjumlah 3,
2. Kamera-kamera yang digunakan tidak mengalami perpindahan posisi atau rotasi selama penangkapan gerakan,
3. Perbaikan distorsi hanya mencangkup distorsi radial dan distorsi tangensial,
4. gerakan yang dilakukan tidak menimbulkan oklusi ke semua kamera yang digunakan.
5. Fokus sistem hanyalah mengakuisisi dan memproses data kalibrasi serta memproses data penanda titik-titik penting 2D menjadi 3D menggunakan triangulasi dan dioptimasi menggunakan *bundle adjustment*. Akuisisi data penanda, pencarian algoritma triangulasi terbaik, dan optimasi lainnya di luar lingkup tugas akhir ini.

## Metodologi

Metodologi adalah sebagai berikut,

1. Identifikasi Masalah

Pengerjaan tugas akhir dimulai dengan mengidentifikasi permasalahan serta kebutuhan dari permasalahan tersebut. Bagian ini meliputi perumusan masalah, tujuan penyelesaian masalah, dan analisis masalah.

1. Perumusan Solusi

Kemudian akan dikaji teori-teori yang dibutuhkan untuk menyelesaikan masalah yang telah dianalisis sebagai modal dasar untuk merancang sistem utuh yang akan digunakan untuk menjawab masalah. Bagian ini meliputi studi literatur

1. Perancangan Sistem

Sistem sebagai bentuk solusi dari permasalahan yang telah dirumuskan sebelumnya, didefinisikan dan dirangkai berdasarkan analisis solusi yang telah dikaji sebelumnya dan Batasan masalah dan sumber daya yang dimiliki.

1. Implementasi sistem

Sistem kemudian dibangun berdasarkan rancangan yang sudah didefinisikan lalu dicoba di lapangan. Percobaan-percobaan yang dilakukan di lapangan berguna untuk mengumpulkan data yang kemudian akan digunakan untuk menguji dan mengevaluasi bagian-bagian dari sistem.

1. Evaluasi sistem

Setelah data-data yang diperlukan didapatkan, akan dilakukan evaluasi kinerja sistem secara kuantitatif dan secara kualitatif. Evaluasi mencangkup hasil keluaran tiap proses pada sistem, temuan-temuan yang didapati pada hasil proses sistem, dan analisis dari setiap temuan.

## Sistematika Pembahasan

Laporan tugas akhir ini dibagi ke dalam lima bab. Bab pertama akan menjelaskan mengenai dasar dan tujuan dilakukannya tugas akhir ini. Penjelasan tersebut akan dibagi menjadi beberapa bagian yaitu latar belakang, rumusan masalah, tujuan tugas akhir, batasan masalah, dan metodologi yang digunakan.

Bab kedua akan membahas mengenai teori-teori dasar dan studi literatur terkait proses-proses yang digunakan untuk menjawab permasalahan dan membangun sistem utuh.

Bab ketiga akan menjelaskan lebih detail mengenai permasalahan yang telah disampaikan pada bagian latar belakang serta analisis solusi untuk menjawab permasalahan tersebut dan rancangan sistem utuh sebagai implementasi dari solusi yang telah dikemukakan.

Bab keempat akan menjelaskan mengenai hasil implementasi rancangan solusi yang telah dibuat dan evaluasi serta analisis terhadap tiap hasil implementasi tersebut.

Bab kelima akan membahas mengenai kesimpulan dari hasil kinerja sistem secara umum serta memberikan saran-saran terkait proses-proses yang sebaiknya dilakukan atau tidak dilakukan ketika membangun sistem yang serupa.

# STUDI LITERATUR

## Pembentukan Gambar Pada Kamera

Pada dasarnya, pembentukan gambar pada kamera adalah sebuah proses pemetaan titik-titik pada ruang dunia 3 dimensi ke dalam bidang gambar 2 dimensi. Untuk melakukan transformasi pada proses pemetaan tersebut, perlu didefinisikan terlebih dahulu model kamera yang mendasari transformasi yang terjadi pada proses pemetaan. Salah satu jenis model kamera yang umum digunakan sebagai acuan dan titik awal untuk mendefinisikan model matematis yang digunakan pada proses transformasi adalah kamera lubang jarum.

Kamera lubang jarum adalah sebuah model kamera yang paling sederhana dengan memanfaatkan sebuah lubang kecil sebagai titik masuknya cahaya. Ukuran lubang pada kamera lubang jarum dibuat kecil supaya hanya satu sinar cahaya pada suatu titik pada dunia dapat melewati lubang tersebut dan jatuh pada bidang gambar sehingga gambar yang terbentuk pada bidang gambar dapat terlihat lebih jelas.

A diagram of a pinhole plane

Description automatically generated

Gambar II.1. Model Kamera Lubang Jarum (Bradski & Kaehler, 2008)

Pada model kamera lubang jarum, titik pada koordinat dunia dipetakan ke dalam titik ***x***pada bidang gambar seperti yang terdapat pada Gambar II.1. Menggunakan perbandingan segitiga serupa, penjabarannya sebagai berikut

Dengan membuang koordinat terakhir pada bidang gambar, titik 2D yang terbentuk adalah . Persamaan berikut menggambarkan transformasi menggunakan sistem koordinat homogen,

Yang mana (u,v) adalah koordinat pada bidang gambar, dan k adalah faktor skala yang diakibatkan oleh proyeksi sumbu Z dunia ke gambar.

Pada persamaan II.2, diasumsikan bahwa titik asal koordinat bidang gambar terletak pada titik utamanya. Namun pada umumnya, terdapat deviasi antara titik asal koordinat bidang gambar dengan titik utamanya sehingga deviasi tersebut perlu diperhitungkan ke dalam persamaan sehingga membentuk persamaan berikut,

Dengan persamaan pada sistem koordinat homogen sebagai berikut,

Adapun pada umumnya ditulis sebuah matriks *K* yang biasa disebut matriks kalibrasi dengan persamaan seperti berikut,

Sehingga didapati bentuk rangkap dari persamaan II.4 menjadi seperti berikut,

Dikarenakan dimungkinkan bahwa *pixel* pada gambar terdapat *pixel* tidak murni berbentuk persegi namun memiliki besar Panjang dan lebarnya sendiri maka matriks *K* pada persamaan II.5 dapat diubah menjadi persamaan berikut.

Di mana adalah Panjang fokal *f* dibagi dengan Panjang *pixel* dan adalah Panjang fokal *f* dibagi dengan lebar *pixel*.

Adapun persamaan II.6 mengasumsikan bahwa titik terletak sistem koordinat kamera yang disebut kerangka koordinat kamera (*camera coordinate frame*). Namun, kenyataannya adalah titik tersebut terletak pada sistem koordinat yang berbeda yang dinamakan kerangka koordinat dunia (*world coordinate frame*). Kedua kerangka koordinat ini dapat dihubungkan dengan proses transformasi rotasi dan translasi sebagai berikut,

Dengan adalah titik pada kerangka koordinat kamera, *R* adalah matriks rotasi berukuran 3 x 3, *t* adalah matriks translasi berukuran 3 x 1, dan yang adalah titik pada kerangka koordinat dunia.

Dengan menggabungkan persamaan II.6 dengan persamaan II.4, terbentuk persamaan berikut,

*K* biasa disebut unsur intrinsik kamera dan R dan t disebut unsur ekstrinsik kamera.

## Distorsi Lensa

Salah satu kelemahan dari kamera lubang jarum adalah untuk mendapatkan gambar yang lebih jelas, lubang kamera (*aperture*) perlu diperkecil sehingga hanya satu sinar cahaya dari satu titik pada dunia dapat terproyeksi pada bidang gambar sehingga gambar terlihat semakin jelas dan fokus. Namun kelemahan dari lubang kamera yang semakin mengecil adalah gambar yang terproyeksi terlihat semakin gelap dikarenakan kurangnya pencahayaan.

Timeline

Description automatically generated

Gambar II.2. Dampak Bukaan Kamera Terhadap Kejelasan Gambar (Hata & Savarese)

Salah satu upaya untuk mengatasi kendala ini adalah dengan menggunakan lensa yang dapat memfokuskan atau menghamburkan cahaya. Menggunakan lensa ini, sinar-sinar cahaya yang dipancarkan oleh suatu titik pada dunia dapat difokuskan oleh lensa ke suatu titik pada bidang gambar sehingga hasil proyeksi gambar tidak hanya terlihat jelas namun juga lebih cerah dibandingkan jika tidak menggunakan lensa.

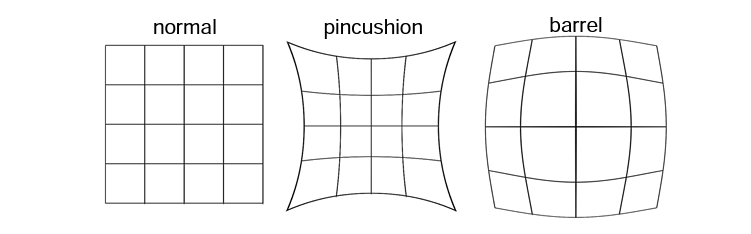
Chart

Description automatically generated

Gambar II.3. Penggunaan Lensa Untuk Memperjelas Gambar (Hata & Savarese)

Namun, penggunaan lensa pada kamera ini menimbulkan beberapa masalah baru yaitu salah satunya yang menjadi pokok bahasan pada subbab ini adalah timbulnya distorsi geometrik pada gambar yang dihasilkan. Pada umumnya, lensa-lensa khususnya lensa-lensa sudut lebar (*wide-angle lenses*) menimbulkan distorsi geometrik yang disebut distorsi radial (*radial distortion*) yang cukup signifikan pada gambar hasil yang menyebabkan efek lengkungan pada daerah gambar dekat sudut-sudut gambar.

Terdapat 2 jenis distorsi radial yaitu distorsi *barrel* dan distorsi *pincushion.* Distorsi *barrel* membuat lengkungan-lengkungan di sekitar sudut-sudut gambar mengarah ke luar menjauhi titik tengah gambar dan membuat efek ‘menggembung’ pada gambar. Distorsi *pincushion* membuat lengkungan-lengkungan di sekitar sudut-sudut gambar mengarah ke dalam mendekati titik tengah gambar dan membuat efek ‘menciut’ pada gambar.



Gambar II.4. Jenis-Jenis Distorsi Kamera (Hata & Savarese)

Distorsi radial dapat dimodelkan dengan persamaan matematis berikut,

Di mana *x* dan *y* adalah koordinat gambar sebelum terdistorsi, *r* adalah panjang radius dari tengah gambar, dan adalah koefisien distorsi radial. Adapun OpenCV, sebuah pustaka terkenal untuk melakukan operasi di bidang *computer vision*, menggunakan koefisien distorsi sampai dengan derajat 6 untuk alasan normalisasi dengan model distorsi sebagai berikut.

Selain daripada distorsi radial, terdapat pula jenis distorsi lainnya dapat terjadi saat pengambilan gambar yaitu distorsi tangensial yang membuat gambar hasil terlihat miring dikarenakan sumbu x dan sumbu y pada sistem koordinat gambar tidak tegak lurus. Umumnya, distorsi ini disebabkan oleh kecacatan pada saat proses manufaktur kamera yang membuat lensa kamera tidak sejajar dengan bidang gambar sehingga distorsi ini muncul.

Diagram

Description automatically generated

Gambar II.5. Distorsi Tangensial (Kaehler & Bradski, 2017)

Distorsi tangensial dapat dimodelkan dengan persamaan matematis berikut,

Dengan demikian, didefinisikan parameter-parameter distorsi yang perlu didapatkan sebagai berikut,

Urutan dari parameter distorsi tersebut mengikuti urutan keluaran parameter distorsi dari proses kalibrasi menggunakan pustaka OpenCV yang akan dibahas pada bagian selanjutnya.

## Kalibrasi Kamera Tunggal

Pada dasarnya kalibrasi kamera adalah pencarian matriks proyeksi yang mentransformasikan titik-titik pada koordinat dunia ke dalam koordinat gambar. Tujuan utama dari dilakukannya kalibrasi kamera ini adalah jika terdapat sejumlah *n* titik pada koordinat dunia dan titik proyeksinya pada koordinat gambar, akan dicari unsur intrinsik dan ekstrinsik kamera yang melakukan transformasi proyeksi tersebut.

Salah satu algoritma kalibrasi kamera yang umum digunakan adalah algoritma kalibrasi Zhang (Zhang, 1998) yang menggunakan sebuah pola kalibrasi datar yang susunan dan dimensinya diketahui. Proses kalibrasinya adalah sebagai berikut:

1. Sejumlah *M* gambar pola kalibrasi diambil dari berbagai sudut yang berbeda dengan menggerakkan pola kalibrasi atau menggerakkan kamera.
2. Untuk setiap gambar yang telah diambil, sejumlah *N* titik pada gambar pola kalibrasi diambil.
3. Definisikan titik-titik 3D pola kalibrasi sejumlah *N* di sistem koordinat sembarang dengan yang mana titik-titik yang diambil tersebut diasumsikan memiliki korespondensi 1:1 dan menggambarkan titik-titik 2D dari gambar pola kalibrasi .
4. Untuk setiap titik pada masing-masing gambar, akan dicari homografi H0,…,HM yang memetakan titik 3D pada bidang pola kalibrasi ke dalam titik 2D yang telah diambil pada gambar.
5. Untuk setiap homografi yang telah didapati, akan diestimasi matriks intrinsik kamera dengan mengabaikan terlebih dahulu parameter distorsi pada kamera.
6. Setelah matriks intrinsik didapati, akan diestimasi matriks ekstrinsik kamera untuk setiap gambar yang diambil.
7. Estimasi distorsi kamera menggunakan *linear least-square*.
8. Melakukan optimasi non-linier untuk setiap parameter-parameter yang telah diestimasi (matriks intrinsik kamera, matriks ekstrinsik kamera, parameter distorsi kamera).

Adapun salah satu kelemahan dari metode Zhang ini adalah metode ini tidak bisa mendapatkan matriks ekstrinsik kamera yang memetakan sistem koordinat dunia ke sistem koordinat kamera. Hal ini wajar dikarenakan metode Zhang hanya menggunakan titik 3D dari sistem koordinat sembarang dan bukan dari sistem koordinat dunia yang umumnya sangat susah untuk didapati. Hasil keluaran dari metode ini adalah matriks intrinsik dan parameter distorsi kamera.

## Triangulasi

Pada dasarnya matriks intrinsik kamera saja tidak cukup untuk dapat melakukan estimasi 3D dengan baik. Hal ini dikarenakan terdapat tak hingga jumlah nilai 3D yang memenuhi persamaan pemetaan titik 3D ke titik 2D bidang gambar menggunakan matriks intrinsik kamera. Hal ini tergambar seperti pada gambar II.6 di bawah ini. Oleh karena itu, penggunaan matriks intrinsik saja tidak bisa menghasilkan estimasi 3D yang akurat.

A diagram of a triangle with text

Description automatically generated

Gambar II.6. Permasalahan proyeksi balik ke 3D dengan hanya 1 *view* saja (Kitani, Triangulation, 2007)

Salah satu pendekatan yang paling umum digunakan untuk mengestimasi 3D adalah menggunakan triangulasi. Prinsip dasar dari triangulasi adalah jika terdapat 2 atau lebih kamera (sering disebut juga sebagai *view*)yang melihat suatu objek, akan dicari suatu titik 3D di ruang dari perpotongan garis-garis proyeksi titik dari masing-masing kamera. Sebagai contoh, perhatikan pada Gambar II.7 di bawah ini.

A diagram of a structure

Description automatically generated

Gambar II.7. Gambaran Triangulasi dengan 2­­-*view* (Kitani, Epipolar Geometry, 2007)

Jika titik 3D yang dilihat secara bersamaan pada 2 *view* jatuh pada titik 2D pada bidang gambar *view* pertama dan pada bidang gambar *view* kedua, garis proyeksi dari view pertama dan view kedua akan saling berpotongan pada titik di mana titik tersebutlah yang merupakan estimasi 3D yang akan dicari.

Untuk memudahkan penjabaran proses matematis dari triangulasi, proses triangulasi akan dimulai dengan menggunakan jumlah view minimalnya yaitu 2. Dengan asumsi bahwa setiap *view* yang digunakan mengikuti model kamera lubang jarum, maka tiap *view* tersebut memiliki matriks intrinsik dan matriks ekstrinsiknya masing-masing. Anggap bahwa **P** adalah sebuah matriks 3x4 hasil perkalian matriks intrinsik dan matriks ekstrinsik.

Matriks **P** ini biasa disebut sebagai matriks proyeksi. Dengan matriks proyeksi ini pemetaan titik 3D objek ke titik 2D gambar menjadi seperti berikut

Jika adalah baris ke *i* dari matriks **P**, maka didapati 3 persamaan berikut

Dengan aljabar linier, 3 persamaan tersebut dapat disusun sebagai berikut,

Di mana,

Persamaan II.18 tersebut dapat diterapkan kepada kedua kamera sehingga jika persamaan-persamaan yang terbentuk digabungkan dan dibentuk sebagai sebuah matriks, terbentuk persamaan berikut,

Di mana dan adalah koordinat gambar pada *view* 1 dan dan adalah koordinat gambar pada *view* 2. Anggap matriks **A** adalah matriks kiri persamaan II.19.

Untuk triangulasi lebih dari 2 *view*, matriks A tersebut dapat diperpanjang sebagai berikut,

Sehingga didapati persamaan,

Namun, dikarenakan pada keadaan nyatanya terdapat *noise* baik dari data 2D ataupun dari matriks proyeksi, garis-garis proyeksi *view* tidak berpotongan sehingga nilai dari sehinga solusinya adalah pencarian nilai dari yang paling yang menghasilkan nilai yang mendekati 0. Permasalahan ini adalah sebuah permasalahan *least-square* dan umumnya dapat diselesaikan dengan metode SVD.

## Kalibrasi Stereo

Stereo kalibrasi adalah proses untuk menentukan relasi geometri antara 2 kamera. Relasi geometri ini umumnya direpresentasikan dengan matriks rotasi dan translasi yang mentransformasikan sistem koordinat suatu kamera terhadap kamera lainnya. Salah satu cara penentuan matriks rotasi dan translasi ini, yang mana cara ini diterapkan oleh proses kalibrasi stereo dari pustaka OpenCV (Bradski & Kaehler, 2008) adalah dengan memanfaatkan matriks rotasi dan translasi dari proses kalibrasi kamera tunggal Zhang yang memetakan titik 3D pola kalibrasi di ruang sembarang ke titik 2D-nya di bidang gambar. Jika diasumsikan bahwa dan adalah matriks rotasi dari kamera *i* yang melakukan pemetaan tersebut, bisa didapatkan sebuah matriks rotasi *R* dan matriks translasi *T* sehingga

Di mana R dan T ini adalah matriks rotasi dan translasi yang mentransformasikan titik-titik pada sistem koordinat kamera 1 ke sistem koordinat kamera 2 sehingga berlaku persamaan berikut,

## Multi-Kamera

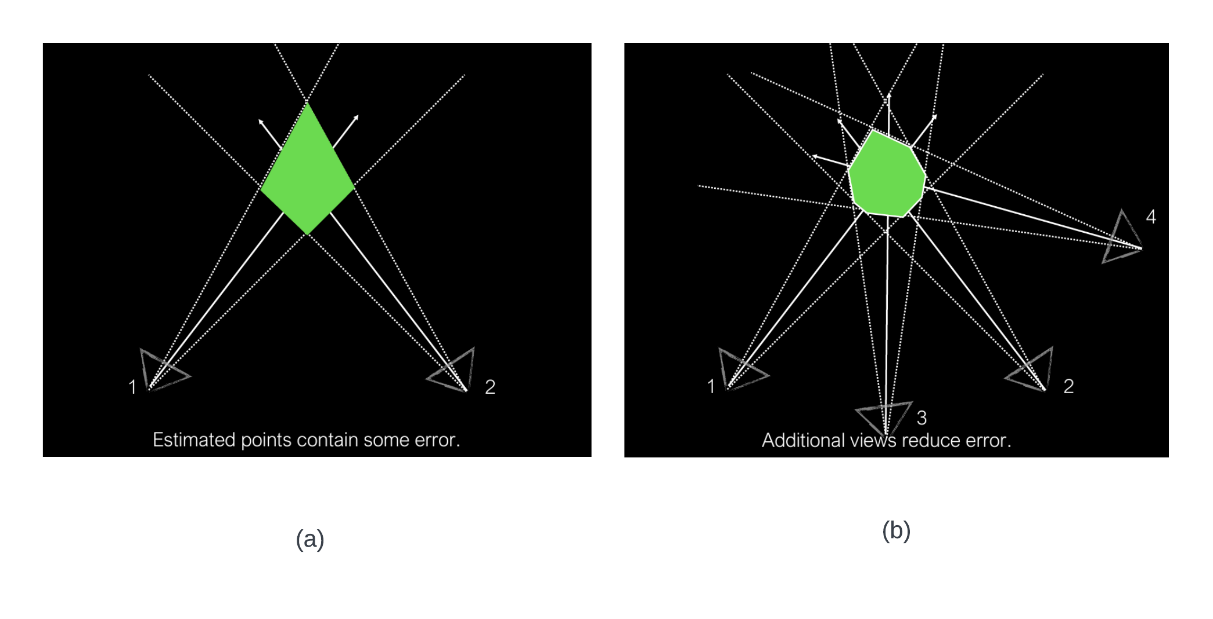
Praktis umum dari estimasi 3D dari gambar adalah untuk menggunakan banyak *view* (sudut pandang atau kamera) yang melihat objek secara bersamaan. Alasan utama mengapa digunakan banyak *view* dalam melakukan estimasi adalah untuk memitigasi pengaruh oklusi pada objek yang mana terdapat *view* tertentu yang dapat melihat bagian objek dengan lebih baik dibandingkan *view* lain. Contohnya seperti Gambar II.8 berikut, *View* pada kamera 2 dan 3 dapat melihat suatu titik pada objek dengan lebih jelas dibandingkan kamera 1.

A close-up of a sculpture

Description automatically generated

Gambar II.8. 3 *View* Melihat Titik Yang Sama Pada Objek (Snavely & Li, 2018)

Alasan lainnya yang juga penting adalah semakin banyak *view* yang melihat objek, semakin banyak pula pengukuran yang dilakukan untuk menentukan struktur 3D objek sehingga *error* dari *noise* dapat terkecilkan. Contohnya seperti pada Gambar II.9 berikut,



Gambar II.9. Visualisasi *Error* Estimasi 3D. Gambar (a) Estimasi 3D Dengan Hanya 2 *View*. Gambar (b) Estimasi 3D Dengan Hanya 4 *View* (Snavely & Li, 2018)

Garis putih adalah garis proyeksi ideal yang tidak mengandung *noise* sementara garis titik-titik adalah garis proyeksi yang disebabkan oleh *noise*. Akibat *noise* tersebut, hasil estimasi 3D menjadi tidak akurat dan ketidakyakinan solusi 3D yang memenuhi pengukuran berada pada daerah yang diberi warna hijau. Seperti pada Gambar II.9 ketika menggunakan banyak *view* akan menyebabkan daerah hijau tersebut menjadi lebih kecil.

## *Reprojection Error*

Istilah *reprojection error* sering digunakan dalam mengukur kualitas estimasi 3D yang dihasilkan. Hal ini dilakukan dengan cara mengukur jarak antara titik 2D dari data dengan titik 2D hasil proyeksi dari estimasi 3D yang telah dilakukan ke gambar. Persamaan matematisnya adalah sebagai berikut.

Di mana *d* adalah fungsi jarak antar 2 titik 2D, *x* adalah titik 2D dari data, *P* adalah matriks proyeksi 3D ke 2D, dan *X* adalah hasil estimasi 3D. *reprojection error* sering digunakan sebagai alternatif pengukuran kualitas 3D dikarenakan pada umumnya data 3D asli tidak bisa didapati.

## Bundle Adjustment

Dikarenakan pada umumnya estimasi 3D tidak sempurna dikarenakan *noise* sehingga hasil proyeksi balik estimasi 3D ke titik 2D-nya tidak sama dengan titik 2D aslinya sehingga muncul *reprojection error* seperti yang dijelaskan sebelumnya. Oleh karena itu, pada umumnya dilakukan optimasi lebih lanjut untuk mengurangi *error*. Optimasi yang umumnya diterapkan bernama *bundle adjustment*. *Bundle adjustment* melakukan optimasi non-linier dengan cara memperbaharui matriks proyeksi tiap kamera dan estimasi 3D sehingga didapatkan *reprojection error* terkecil. Berikut adalah persamaan matematisnya,

Persamaan menggambarkan *bundle adjustment* yaitu akan dicari matriks proyeksi P untuk setiap kamera ke-*i* dan estimasi 3D *X* untuk setiap hasil estimasi 3D dari triangulasi ke-*j* yang meminimalkan jarak dari hasil proyeksi balik kamera *i* untuk titik ke-*j* dengan titik 2D dari data pada kamera *i* pada titik ke *j*

## OpenCV

OpenCV (Bradski G. , 2000) adalah sebuah pustaka pemrograman *open-source* yang dibuat secara khusus untuk menangani persoalan di bidang pemrosesan citra dan penglihatan komputer. Penggunaan pustaka ini sangat membantu dalam melakukan operasi dalam bidang citra, khususnya pada tugas akhir ini dalam hal untuk melakukan kalibrasi kamera tunggal dengan metode Zhang menggunakan fungsi *cv.calibrateCamera* dan kalibrasi kamera stereo dengan fungsi *cv.stereoCalibrate* yang mana kedua fungsi ini melakukan kalibrasi seperti yang dijelaskan pada subbab II.3 kalibrasi kamera tunggal dan subbab II.5 kalibrasi stereo.

# PENGEMBANGAN SISTEM MULTI-KAMERA

## Analisis Persoalan

Seperti yang telah dijelaskan pada bagian latar belakang, kemajuan dalam teknologi AI membuat sistem tangkap gerak menjadi lebih praktis dan murah namun memiliki akurasi hasil gerakan yang dapat dipertanyakan. Sedangkan, sistem konvensional memerlukan bahan, alat, dan teknisi yang menghabiskan biaya yang cukup tinggi yang umumnya tidak terjangkau oleh industri-industri kecil, tetapi dapat menghasilkan akurasi hasil tangkap gerak yang jelas serta lebih akurat dibandingkan sistem tangkap gerak berbasis AI. Namun, dikarenakan pembuatan sistem tangkap gerak konvensional tersebut membutuhkan biaya yang mahal, maka replikasi dari sistem dan kinerjanya tidak dapat dilakukan pada tugas akhir ini. Oleh karena itu, untuk melakukan perbandingan sistem tangkap gerak berbasis AI diperlukan sebuah sistem tangkap gerak yang tidak menggunakan teknologi berbasis AI dan mampu mendekati kinerja dari sistem tangkap gerak konvensional tanpa menggunakan alat-alat dan bahan-bahan mahal untuk melakukan estimasi gerakan aktor. Khususnya pada tugas akhir ini, akan dibahas mengenai cara melakukan estimasi letak 3D penanda-penanda yang ada pada tubuh aktor dengan menggunakan sistem multi-kamera yang akan dibangun.

Estimasi letak 3D dari gambar bukanlah sebuah permasalahan sederhana. Hal ini dikarenakan saat dilakukan pengambilan gambar, informasi 3D objek hilang dipetakan ke sebuah gambar 2D. Bagian selanjutnya akan membahas mengenai proses dan kebutuhan setiap proses untuk merekonstruksi data 2D menjadi 3D menggunakan dasar-dasar teori yang telah dikemukakan sebelumnya.

## Analisis Solusi

Berdasarkan subbab II.1 mengenai cara pembentukan gambar pada kamera, suatu titik 3D pada dunia akan dipetakan ke dalam suatu titik 2D pada kamera menggunakan transformasi seperti yang terdapat pada rumus II.6 yang melibatkan matriks ekstrinsik yang memetakan titik-titik 3D pada sistem koordinat dunia ke dalam titik-titik 3D pada sistem koordinat kamera dan matriks intrinsik yang memetakan titik-titik 3D pada sistem koordinat kamera ke dalam titik-titik 2D pada sistem koordinat gambar. Jika diketahui nilai titik-titik 2D pada sistem koordinat gambar serta nilai dari masing-masing matriks intrinsik dan matriks ekstrinsik kamera, dimungkinkan untuk melakukan proyeksi balik sehingga titik-titik 2D pada sistem koordinat gambar dapat kembali menjadi titik-titik 3D di sistem koordinat dunia. Namun, terdapat sebuah permasalahan yaitu informasi mengenai matriks intrinsik dan matriks ekstrinsik kamera umumnya tidak diketahui atau tidak langsung tersedia pada kamera yang digunakan.

Proses pencarian matriks intrinsik dan matriks ekstrinsik dinamakan kalibrasi kamera atau pada tugas akhir ini sering disebut sebagai kalibrasi kamera tunggal. Menggunakan metode kalibrasi yang telah dibahas pada subbab II.3 yaitu Metode Kalibrasi Zhang, dimungkinkan untuk didapatkan matriks intrinsik kamera serta parameter distorsi kamera menggunakan beberapa gambar pola kalibrasi seperti papan catur yang diketahui jumlah baris dan kolomnya serta berapa sisi setiap persegi pada pola papan catur tersebut. Parameter distorsi kamera yang dimaksud adalah parameter distrosi dari distrosi gambar yang disebabkan oleh penggunaan lensa pada kamera. Informasi mengenai parameter ini penting didapatkan dikarenakan parameter ini dapat digunakan untuk melakukan penghilangan distorsi. Penghilangan distorsi ini penting dilakukan karena akan membantu meningkatkan akurasi dari proses-proses lain yang akan dijelaskan selanjutnya.

Adapun, Metode Kalibrasi Zhang memiliki sebuah kelemahan yaitu metode ini tidak mampu mendapatkan matriks ekstrinsik yang memetakan titik-titik pada sistem koordinat dunia ke sistem koordinat kamera, sehingga tidak dapat dilakukan pemetaan balik dari sistem koordinat 2D gambar ke sistem koordinat 3D dunia. Akan tetapi, dikarenakan telah didapatkan matriks intrinsik dari hasil proses Kalibrasi Zhang, bisa dilakukan pemetaan balik dari sistem koordinat 2D gambar ke sistem koordinat 3D kamera sehingga informasi 3D yang hilang saat pengambilan gambar dapat ditemukan kembali.

Akan tetapi, terdapat sebuah permasalahan baru. Seperti yang telah dibahas pada bagian awal subbab II.4, yaitu meskipun dimungkinkan untuk melakukan pemetaan balik dari sistem koordinat 2D gambar menjadi sistem koordinat 3D kamera dengan hanya satu sudut pandang atau satu *view*, hasil proyeksi 3D yang dihasilkan tidak jelas, khususnya informasi mengenai ke dalam atau nilai sumbu Z suatu titik. Seperti yang juga dijelaskan pada subbab II.4., hal ini disebabkan oleh terdapat jumlah titik 3D yang tidak terhingga yang memenuhi persamaan pemetaan suatu titik 3D ke dalam titik 2D-nya seperti pada gambar II.6. Namun, jika digunakan lebih dari 1 *view* yang melihat suatu titik 3D yang sama, akan didapati setidaknya 1 titik 3D yang memenuhi persamaan pemetaan gambar dari masing-masing *view* seperti yang terdapat pada gambar II.7.

Upaya untuk mencari titik 3D menggunakan lebih dari 1 *view* dinamakan sebagai triangulasi. Seperti pada yang telah dibahas pada subbab II.4 triangulasi pada dasarnya menggunakan ilmu dasar aljabar linear yaitu mencari titik perpotongan dari 2 atau banyak garis, di mana dalam hal ini titik perpotongan yang dimaksud adalah titik 3D yang akan dicari sementara garis yang dimaksud adalah sebuah matriks proyeksi yang memetakan titik 3D ke titik 2D. Fungsi proyeksi yang dimaksud adalah hasil perkalian dari matriks intrinsik dan matriks ekstrinsik seperti pada persamaan II.14. Dikarenakan untuk mendapati titik perpotongan tersebut dibutuhkan setidaknya 2 garis yang tidak paralel, maka untuk mendapati titik 3D dibutuhkan setidaknya 2 *view* yang melihat suatu titik 3D yang sama. Setelah, matriks proyeksi masing-masing kamera dan titik 2D dari masing-masing kamera yang melihat titik 3D yang sama digabungkan menjadi sebuah matriks *A* seperti yang terdapat pada persamaan II.20. penyelesaian persamaan II.22 akan mendapatkan titik 3D perpotongan dari masing-masing garis proyeksi kamera.

Pada kasus nyata, titik 3D perpotongan masing-masing pemetaan balik kamera tidak bisa didapatkan karena umumnya garis pemetaan balik masing-masing kamera tidak saling berpotongan yang umumnya disebabkan oleh *noise* baik dari titik 2D yang dilihat masing-masing kamera atau matriks proyeksi kamera. Oleh karena itu, permasalahan pencarian titik 3D perpotongan dari garis-garis proyeksi kamera diubah menjadi permasalahan pencarian titik 3D terbaik yang dapat merepresentasikan perpotongan dari garis-garis proyeksi. Untuk menyelesaikan permasalahan ini, umumnya diterapkan metode SVD yang melakukan dekomposisi terhadap matriks *A* sehingga didapati titik 3D terbaik yang paling mendekati pemenuhan persamaan perpotongan garis-garis proyeksi kamera.

Adapun satu permasalahan yang masih terdapat pada proses sampai saat ini. Mengingat bahwa matriks proyeksi yang digunakan pada proses triangulasi adalah hasil perkalian dari matriks intrinsik dan matriks ekstrinsik dari masing-masing kamera. Permasalahan yang masih terdapat pada proses triangulasi ini adalah proses triangulasi tersebut mengasumsikan bahwa matriks ekstrinsik yang digunakan adalah matriks ekstrinsik yang mentransformasikan sistem koordinat kamera ke sistem koordinat dunia di mana sampai tahap ini matriks ekstrinsik ini belum bisa didapatkan. Namun, solusi lain dapat diterapkan yaitu selain daripada melakukan transformasi sistem koordinat masing-masing kamera ke sistem koordinat dunia sehingga semua operasi triangulasi berjalan dalam konteks sistem koordinat dunia, dapat dilakukan alternatif lain yaitu melakukan operasi triangulasi hanya dalam konteks sistem koordinat salah satu kamera saja sehingga yang perlu dicari adalah matriks transformasi yang mentransformasikan sistem koordinat kamera ke sistem koordinat kamera tujuan. Upaya pencarian transformasi ini jauh lebih mudah dibandingkan mencari matriks ekstrinsik masing-masing kamera dan proses pencarian ini disebut sebagai kalibrasi stereo.

Serupa dengan matriks ekstrinsik yang melakukan rotasi dan translasi terhadap sistem koordinat dunia ke dalam sistem koordinat kamera, kalibrasi stereo mencari matriks yang melakukan rotasi dan translasi yang mentransformasikan sistem koordinat suatu kamera ke sistem koordinat kamera lain. Menggunakan teknik kalibrasi stereo subbab II.5 yang memanfaatkan penggunaan pola kalibrasi papan catur, seperti pada teknik kalibrasi tunggal, yang dilihat oleh kedua kamera secara bersamaan, akan menghasilkan matriks rotasi dan translasi yang diperlukan.

Sampai pada tahap ini, telah didapatkan matriks intrinsik serta matriks rotasi dan translasi dari satu kamera ke kamera lain. Informasi-informasi yang telah diperoleh sampai saat ini sudah cukup untuk mendapatkan estimasi titik 3D menggunakan teknik triangulasi. Adapun ketika berbicara mengenai struktur hasil estimasi 3D menggunakan triangulasi, penggunaan jumlah kamera lebih dari 2 *view* dapat meningkatkan kualitas dari hasil estimasi titik 3D. Hal ini dikarenakan estimasi 3D menggunakan lebih dari membuat proses estimasi titik 3D tersebut lebih *robust* terhadap *noise* dari parameter-parameter yang digunakan sehingga menghasilkan *error* yang lebih kecil. Selebihnya mengenai estimasi 3D dengan lebih dari 2 *view* dapat dilihat pada subbab II.6.

Tahap terakhir, yang mana tahap ini adalah tahap opsional tetapi merupakan aspek penting yang sudah menjadi standar saat melakukan estimasi 3D adalah melakukan optimasi terhadap hasil estimasi 3D dari proses triangulasi. Pada tahap ini umumnya digunakan metode yang dinamakan *bundle adjustment* yang melakukan optimasi terhadap matriks proyeksi yang digunakan dan titik 3D hasil proses triangulasi sehingga kesalahan dari proyeksi balik titik 3D ke titik 2D-nya dari setiap kamera yang digunakan dapat diminimalkan. Selebihnya mengenai estimasi 3D dengan lebih dari 2 *view* dapat dilihat pada subbab II.7. Setelah dilakukan tahap optimasi ini, diharapkan bahwa hasil estimasi 3D yang sudah teroptimasi dapat semakin akurat.

Dengan demikian, jika disimpulkan proses-proses yang dilakukan untuk menghasilkan estimasi titik-titik 3D penanda dari titik-titik 2D-nya dapat dijabarkan sebagai berikut:

1. Mendapatkan matriks intrinsik dan parameter distorsi tiap kamera dengan kalibrasi kamera tunggal,
2. Menghilangkan distorsi pada pengambilan gambar kamera,
3. Memilih satu dari antara beberapa kamera yang digunakan untuk dijadikan kamera utama yang mana proses triangulasi kamera-kamera lainnya menggunakan sistem koordinat kamera tersebut,
4. Mendapatkan matriks rotasi dan translasi tiap kamera yang mentransformasikan sistem koordinat kamera yang bersangkutan ke sistem koordinat kamera utama dengan kalibrasi stereo,
5. Melakukan triangulasi dari informasi data 2D penanda (yang didapatkan dari sistem lain) dan informasi matriks intrinsik serta matriks rotasi dan translasi tiap kamera untuk mendapatkan estimasi 3D penanda.
6. Melakukan optimasi terhadap hasil estimasi 3D penanda dari proses triangulasi dengan metode *bundle adjustment.*

## Rancangan Solusi

Bagian ini akan membahas implementasi dari tiap proses-proses yang dilakukan untuk mendapatkan data letak 3D penanda dari data 2D-nya seperti yang telah dibahas pada bagian sebelumnya.

### Kalibrasi Kamera Tunggal

Metode Kalibrasi Zhang, melibatkan pengambilan gambar beberapa gambar pola kalibrasi untuk kemudian diproses untuk menghasilkan estimasi matriks intrinsik kamera dan parameter distorsi kamera. Pola kalibrasi yang dipilah adalah sebuah papan catur dengan 6 jumlah baris dan 8 jumlah kolom dengan besar sisi untuk setiap petak pada catur adalah 4.5 cm, seperti pada gambar berikut.



Gambar III.1. Pola Kalibrasi Papan Catur yang Digunakan.

Kemudian akan dilakukan pencarian titik-titik 2D dari tepi dalam petak-petak yang ada pada papan catur menggunakan fungsi *cv.findChessboardCorners* dari pustaka OpenCV. Titik-titik 2D inilah yang akan dijadikan masukan untuk proses kalibrasi kamera tunggal.

Untuk masukan letak 3D dari tepi-tepi yang bersangkutan, akan diasumsi bahwa titik-titik tersebut memiliki nilai dan akan dibuat seolah-olah papan catur tersebut datar dan tegak lurus secara sempurna terhadap kamera sehingga misalkan pada kasus ini digunakan pola kalibrasi papan catur dengan 6 jumlah baris dan 8 jumlah kolom dengan besar sisi untuk setiap petak pada catur adalah 4.5 cm, maka akan didapati 5x7 tepi dalam petak-petak dengan koordinat 3D tiap tepi petak digambarkan sebagai tabel matriks berikut.

Tabel III.1. Koordinat 3D Tepi Dalam Petak-Petak Papan Catur

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| (0,0,0) | (4.5,0,0) | (9,0,0) | (13.5,0,0) | (18,0,0) |
| (0,4.5,0) | (4.5,4.5,0) | (9,4.5,0) | (13.5,4.5,0) | (18,4.5,0) |
| (0,9,0) | (4.5,9,0) | (9,9,0) | (13.5,9,0) | (18,9,0) |
| (0,13.5,0) | (4.5,13.5,0) | (9,13.5,0) | (13.5,13.5,0) | (18,13.5,0) |
| (0,18,0) | (4.5,18,0) | (9,18,0) | (13.5,18,0) | (18,18,0) |
| (0,22.5,0) | (4.5,22.5,0) | (9,22.5,0) | (13.5,22.5,0) | (18,22.5,0) |
| (0,27,0) | (4.5,27,0) | (9,27,0) | (13.5,27,0) | (18,27,0) |

Untuk mengestimasi matriks intrinsik dan parameter distorsi kamera, akan digunakan fungsi *cv.calibrateCamera* dari pustaka OpenCV. Fungsi tersebut menerapkan metode kalibrasi Zhang seperti yang pada dasar teori subbab II.3 yang mana fungsi tersebut menerima data 2D dan data 3D tepi-tepi petak dan menghasilkan matriks intrinsik dan parameter distorsi kamera yang dicari.

### Penghilangan Distorsi Gambar

Untuk melakukan penghilangan distorsi pada gambar yang diambil kamera, akan digunakan fungsi *cv.undistort* dari pustaka OpenCV. Fungsi ini menerima masukan gambar yang ini diperbaiki distorsinya, matriks intrinsik kamera, dan parameter distorsi kamera. Matriks intrinsik dan parameter distorsi kamera diperoleh dari proses sebelumnya.

### Pemilihan Kamera Utama

Dikarenakan keterbatasan sumber daya yang dimiliki, jumlah kamera yang digunakan pada tugas akhir ini adalah 3 kamera. Ketiga kamera tersebut akan disusun seperti pada gambar III.2 berikut,

A diagram of a person's diagram

Description automatically generated

Gambar III.2. Gambaran Urutan Kamera Yang Digunakan

Di mana terdapat suatu kamera yang berada di antara kedua kamera lainnya yang disebut kamera tengah dan terdapat kamera yang berada di sebelah kiri kamera tengah yang dinamakan kamera kiri dan terdapat kamera yang berada di sebelah kanan kamera tengah yang dinamakan kamera kanan. Kamera tengah akan dijadikan sebagai kamera utama karena kamera tengah akan diposisikan saling berhadapan dengan aktor sehingga hasil estimasi 3D dapat terlihat jelas dan tidak miring. Pengukuran rotasi dan translasi kamera lain akan diukur dari besar sudut dan perpindahan kamera tersebut dari kamera tengah. Selebihnya mengenai nilai rotasi dan translasi kamera yang digunakan untuk evaluasi sistem akan dibahas pada subbab III.3.5.

### Kalibrasi Kamera Stereo

Untuk melakukan kalibrasi kamera stereo, untuk mendapatkan nilai rotasi dan translasi yang mentransformasikan sistem koordinat kamera kiri dan kamera kanan ke sistem koordinat kamera tengah, dilakukan pengambilan gambar pola kalibrasi papan catur yang sama seperti yang digunakan pada tahap kalibrasi kamera tunggal di mana pola kalibrasi tersebut dapat terlihat oleh pasangan kamera kiri-tengah dan pasangan kamera kanan-tengah secara bersamaan. Pengambilan gambar pola kalibrasi di mana pola tersebut dapat terlihat dengan jelas oleh ketiga kamera secara sekaligus dapat mempersingkat pengambilan karena pengambilan gambar untuk masing-masing pasangan kamera dapat diambil secara bersamaan. Seperti pada contoh gambar III.3 berikut.

A person holding a checkerboard

Description automatically generated

Gambar III.3. Gambar Pola Kalibrasi Diambil Secara Bersamaan Oleh Kamera Kiri (a), Kamera Tengah (b), dan Kamera Kanan (c).

Serupa seperti yang dilakukan pada kalibrasi kamera tunggal, akan dicari titik-titik 2D tepi dalam dari setiap petak pada pola kalibrasi papan catur untuk masing-masing gambar pola kalibrasi yang diambil dengan fungsi *cv.findChessboardCorners* dari pustaka OpenCV. Titik-titik tepi yang telah didapatkan dari setiap pasangan kamera bersamaan dengan nilai 3D dari tiap titik-titik tepi (nilai 3D ini menggunakan konfigurasi yang sama seperti yang digunakan pada kalibrasi kamera tunggal) akan digunakan untuk mengestimasi matriks rotasi dan translasi yang dibutuhkan dengan bantuan fungsi *cv.stereoCalibrate* dari pustaka OpenCV.

### Estimasi 3D Penanda dengan Triangulasi

Pada bagian ini, akan ditetapkan beberapa konfigurasi rotasi dan translasi dari tiap pasangan kamera. Hal ini dilakukan untuk mencari konfigurasi rotasi dan translasi dari pasangan kamera yang menghasilkan estimasi 3D yang terbaik. Untuk mempermudah pengujian dan dikarenakan keterbatasan ruangan, ditetapkan beberapa Batasan dalam melakukan pencarian rotasi dan translasi yang menghasil estimasi 3D terbaik sebagai berikut,

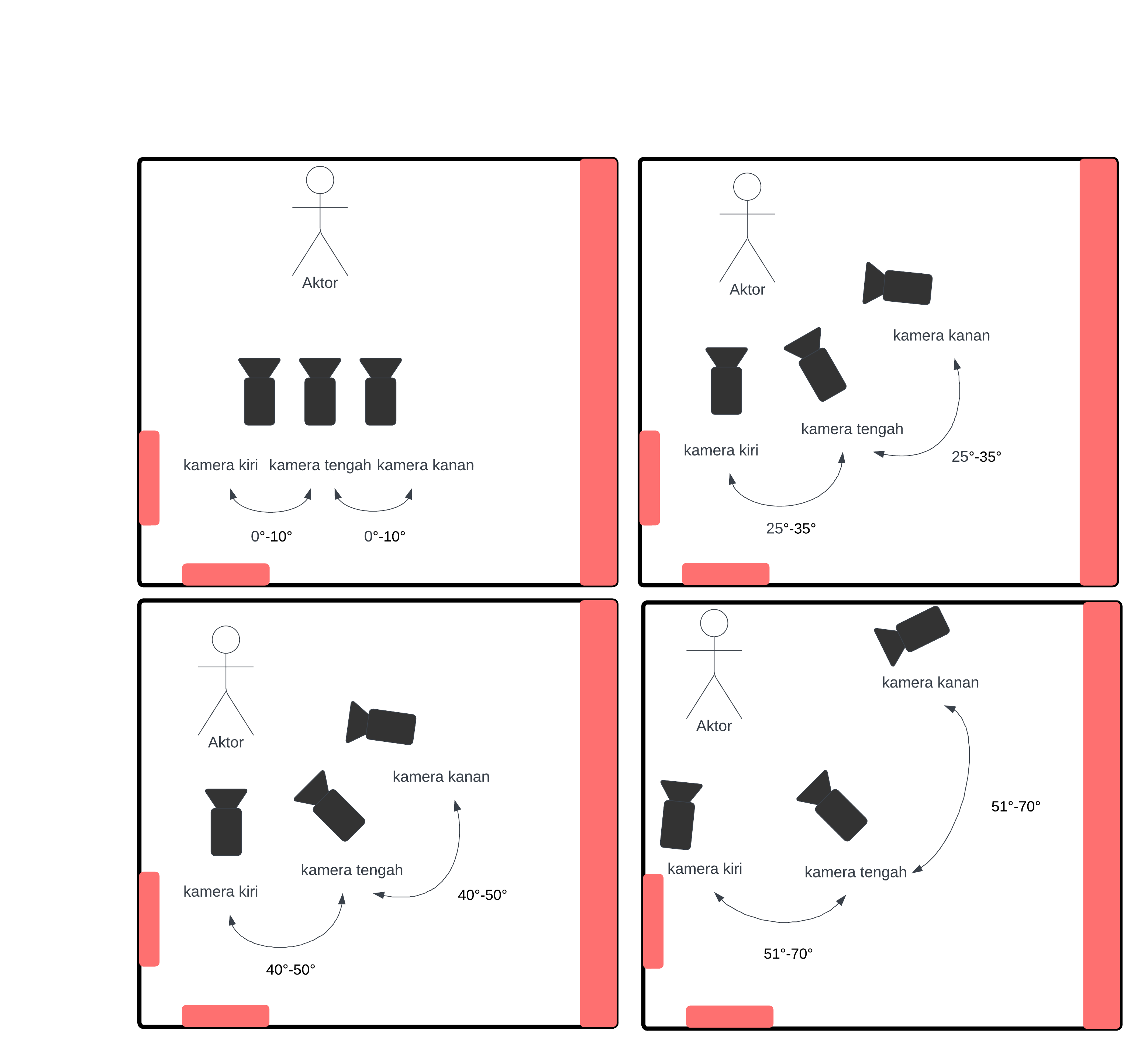
1. Pengujian lebih menekankan perubahan rotasi dari pasangan kamera di mana aspek translasi disesuaikan sedemikian rupa dengan kondisi ruangan yang terlihat pada kamera sehingga kamera tidak melihat bagian ruangan yang berpotensi merusak pengambilan data 2D penanda. Selebihnya mengenai bagian ruangan yang dapat digunakan dapat dilihat pada gambar III.4 berikut

A white rectangle with black lines

Description automatically generated

Gambar III.4. Kondisi Ruangan. Bagian Merah mengindikasikan daerah ruangan yang tidak boleh terlihat oleh kamera.

1. Rentang perubahan rotasi kamera yang digunakan berkisar antara 0 hingga 70 derajat dari kamera yang bersangkutan ke kamera tengah. Rentang ini diestimasi dari perkiraan pengukuran yang dilakukan di lapangan.
2. Rentang perubahan rotasi akan dibagi ke dalam 4 rentang yang menggambarkan perubahan sudut dari kecil ke besar untuk menentukan pengaruh perubahan rotasi yang terhadap hasil estimasi 3D. 4 rentang tersebut yaitu 0°-10°, 25°-35°, 40°-50°, 51°-70°. Selebihnya mengenai gambaran posisi dan rotasi kamera yang dilakukan dapat dilihat pada III.5 berikut



Gambar III.5. Konfigurasi Peletakan Kamera Untuk Setiap Rentang Rotasi

1. Percobaan penentuan rotasi terbaik dilakukan dengan kondisi di mana rotasi setiap pasangan kamera diuji pada rentang yang sama. Hal ini berarti jika pasangan kamera yang satu diuji pada rentang 0°-10° berarti pasangan kamera yang lain pun diuji pada rentang tersebut. Hal ini dilakukan untuk mempermudah pengujian dengan tidak melakukan pengujian yang mencoba seluruh kombinasi rentang yang mungkin.

Dengan ditetapkan Batasan seperti yang telah ditetapkan tersebut, maka perlu didapati 1 matriks rotasi dan translasi untuk tiap rentang pengujian untuk tiap pasangan kamera sehingga perlu dilakukan 4 kalibrasi stereo untuk tiap pasangan kamera.

Adapun selain dari nilai rotasi dan translasi, terdapat aspek lain yang dapat mempengaruhi kualitas estimasi 3D. Seperti yang telah dibahas sebelumnya, penggunaan multi-kamera dipercaya untuk dapat meningkatkan kualitas estimasi 3D sehingga perlu dilakukan percobaan untuk membuktikan pernyataan tersebut. Percobaan yang akan dilakukan adalah membandingkan penggunaan triangulasi dengan 3 kamera dengan semua pasangan 2 kamera yang mungkin yaitu pasangan kamera kiri-tengah, kamera kanan-tengah, kamera kiri-kanan. Dengan percobaan ini, total percobaan yang akan dilakukan adalah 4 percobaan untuk tiap membandingkan kualitas dari pasangan kamera yang mungkin. 4 percobaan tersebut dilakukan untuk tiap rentang rotasi yang telah ditetapkan sehingga terdapat setidaknya total 16 percobaan. Pengukuran terhadap kualitas estimasi 3D akan ditentukan dari besar *reprojection error* atau error proyeksi balik hasil estimasi 3D ke titik 2Dnya dan dilihat perbedaan titik 2D hasil proyeksi balik dan titik 2D dari data yang digunakan untuk estimasi 3D menggunakan triangulasi. Penjelasan lebih lanjut mengenai proyeksi balik dibahas pada subbab II.4.3.

Untuk melakukan proses triangulasi sendiri, setelah mendapatkan parameter-parameter yang dibutuhkan yaitu data 2D penanda, matriks intrinsik kamera, serta matriks rotasi dan translasi kamera terhadap kamera tengah, parameter-parameter tersebut digunakan masukan untuk proses triangulasi seperti pada proses pada subbab II.4. Penerapan proses triangulasi ini dibuat dengan menggunakan pemahaman terhadap proses triangulasi yang ada pada 2 sumber berikut yaitu (Batpurev, 2021) dan (AmyTabb, 2021).

### Optimasi Hasil Estimasi 3D dengan *Bundle Adjustment*

Pada tahap ini, akan diterapkan optimasi *bundle adjustment* untuk setiap percobaan yang dilakukan, yang berarti total jumlah optimasi yang perlu dilakukan adalah 16 kali yang juga berarti jumlah total perbandingan sekarang adalah 32 perbandingan. Implementasi *bundle adjustment* dibuat menggunakan pemahaman dan adaptasi proses *bundle adjustment* yang terdapat pada sumber berikut (Large-scale bundle adjustment in scipy, 2016).

# EVALUASI SISTEM

Pada bagian ini akan dipaparkan proses dasar, temuan, analisis, dan hasil akhir dari tiap implementasi rancangan solusi.

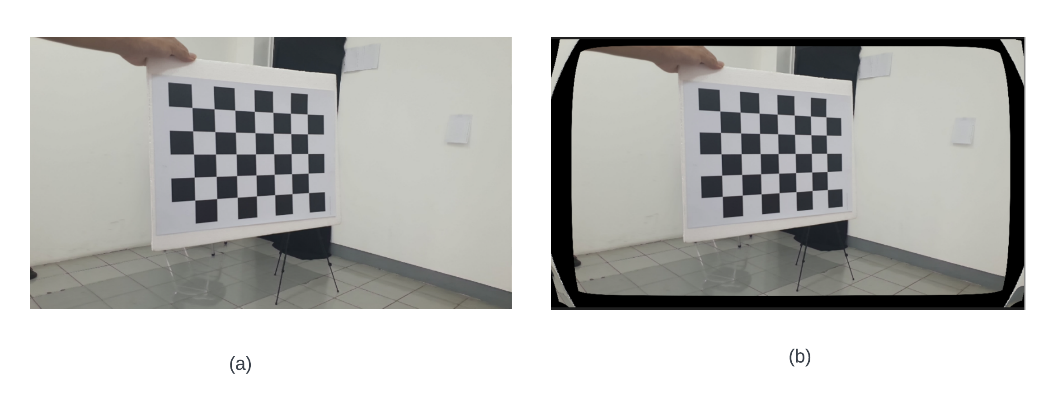
## Kalibrasi Kamera Tunggal dan Penghilangan Distorsi Gambar

Dikarenakan keterhubungan proses kalibrasi kamera dan proses penghilangan distorsi gambar, evaluasi terhadap kedua proses tersebut dilakukan akan disatukan dalam subbab ini.

Proses kalibrasi diawali dengan melakukan pengambilan gambar pola kalibrasi untuk tiap kamera. Jumlah gambar pola kalibrasi papan catur yang diambil oleh tiap kamera ada pada rentang jumlah 10 hingga 20 gambar sesuai dengan arahan yang ada pada (Zhang, 1998). Setelah dilakukan pendeteksian tepi dalam petak-petak pada pola kalibrasi, titik-titik 2D hasil deteksi dan letak 3D-nya dengan nilai seperti yang ada pada Tabel.III.1 dimasukan ke dalam fungsi *cv.calibrateCamera* dan dihasilkan matriks intrinsik kamera dan parameter distorsi kamera.

Salah satu temuan yang konsisten didapati pada saat pengambilan gambar adalah pada setiap kamera, gambar awal yang diambil sebelum dilakukan distorsi tidak memiliki indikasi distorsi yang kasat mata. Hal ini membuahkan asumsi bahwa distorsi yang terdapat pada kamera saat kecil hingga tidak terlihat oleh mata. Hal tersebut, dapat terjadi dikarenakan lensa yang digunakan pada kamera memiliki lengkungan yang tidak terlalu besar sehingga distorsi akibat lensa tidak terlalu besar. Dari temuan ini dapat ditarik sebuah asumsi baru bahwa seharusnya gambar hasil perbaikan distorsi tidak terlalu mengubah gambar awal dikarenakan distorsi gambar awal kecil.

Namun, pada masa awal percobaan. Asumsi tersebut tidak dapat terpenuhi. Gambar hasil penghilangan distorsi bukan hanya tidak menghilangkan distorsi yang kecil namun malah menimbulkan distorsi yang sangat signifikan seperti pada Gambar IV.1. di bawah ini.

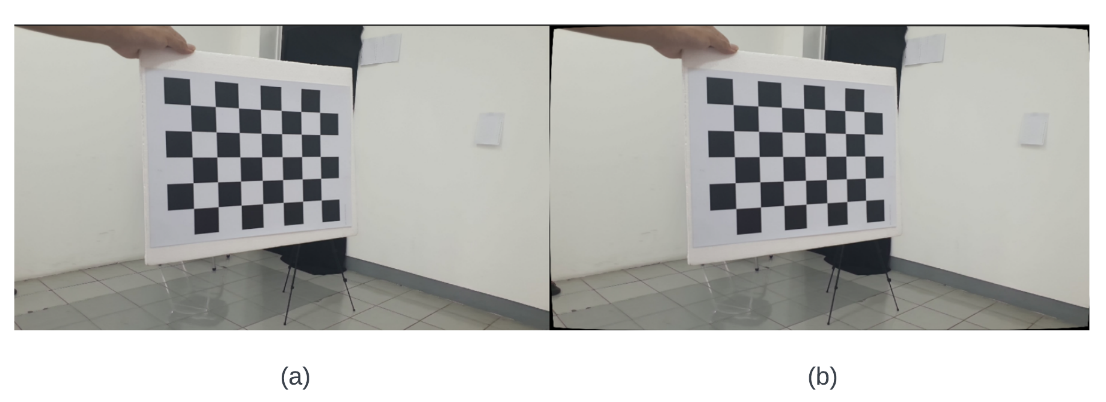


Gambar IV.1. Gambar (a) gambar asli sebelum dilakukan perbaikan distorsi. Gambar (b) gambar setelah dilakukan perbaikan distorsi.

Setelah dilakukan beberapa iterasi percobaan ditemukan 2 faktor yang menyebabkan masalah ini distorsi gambar yaitu,

1. Gambar pola kalibrasi tidak cukup tersebar rata didaerah tepi *view* kamera dan hanya terfokuskan di daerah tengah *view* kamera. Hal ini menyebabkan algoritma kalibrasi untuk memodelkan parameter distorsi hanya di daerah tengah kamera saja dan tidak pada daerah lainnya.
2. Algoritma kalibrasi menggunakan derajat koefisien distorsi yang rendah. Berdasarkan dokumentasi OpenCV , derajat *default* dari koefisien distorsi ditetapkan pada derajat 3 dari total 6 derajat yang bisa digunakan. Selama melakukan percobaan, jumlah derajat 3 dinilai masih terlalu kecil dan tidak dapat memodelkan distorsi dengan baik.

Setelah memperbaiki kedua kesalahan tersebut, dengan mengambil gambar pola kalibrasi yang merata ke seluruh bagian kamera serta meningkatkan derajat distorsi dari 3 ke 6. Didapati hasil perbaikan distorsi seperti gambar IV.2. di bawah ini yang tidak jauh berbeda dengan gambar awalnya.



Gambar IV.2. Gambar (a) gambar asli sebelum dilakukan perbaikan distorsi. Gambar (b) gambar setelah dilakukan perbaikan distorsi.

Pada percobaan ini, tidak dilakukan uji pada derajat-derajat lain yang tersedia pada OpenCV dikarenakan hasil dari perbaikan distorsi dinilai sudah cukup optimal untuk melanjutkan ke proses lain.

Distribusi pengambilan gambar tiap kamera dan hasil matriks kalibrasi serta parameter distorsi tiap kamera ada pada tabel IV.1. dan IV.2. di bawah ini.

Tabel IV.1. Parameter Intrinsik dan Distorsi Lensa Hasil Kalibrasi Kamera Tunggal

| Kamera | Jumlah Pengambilan Gambar Pola Kalibrasi | Jumlah Pola Kalibrasi Yang Berhasil Terdeteksi |
| --- | --- | --- |
| Kamera kiri | 12 | 12 |
| Kamera tengah | 20 | 18 |
| Kamera kanan | 10 | 9 |

Tabel IV.2. Parameter Intrinsik dan Distorsi Lensa Hasil Kalibrasi Kamera Tunggal

| Kamera | Matriks Intrinsik Kamera | Parameter Distorsi |
| --- | --- | --- |
| Kamera kiri |  |  |
| Kamera tengah |  |  |
| Kamera kanan |  |  |

Matriks intrinsik kamera disusun seperti yang terdapat pada persamaan untuk matriks intrinsik kamera dan persamaan untuk parameter distorsi.

## Kalibrasi Stereo

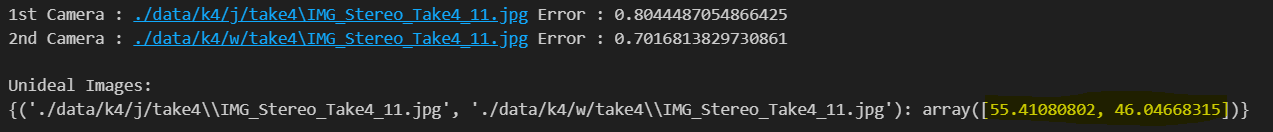
Proses pada tahap ini cukup sederhana yaitu mengambil gambar pola kalibrasi untuk 3 kamera secara sekaligus jika dimungkinkan lalu memisahkan gambar-gambar yang diambil untuk tiap pasangan kamera atau mengambil gambar untuk satu pasang kamera terlebih dahulu lalu mengambil gambar untuk pasangan lain jika tidak dimungkinkan. Hal ini terjadi dikarenakan bergantung dari posisi dan rotasi kamera, ketiga kamera mungkin tidak dapat mengambil pola kalibrasi secara bersamaan.

Namun dari percobaan yang dilakukan, ketiga kamera dapat mengambil gambar pola kalibrasi secara bersamaan untuk tiap rentang rotasi seperti yang telah ditetapkan pada subbab III.3.5. Hal ini mempermudah proses pengambilan gambar yang diperlukan.

Adapun berbeda dari proses sebelumnya, dari percobaan yang telah dilakukan didapati bahwa proses kalibrasi stereo ini tidak memerlukan jumlah gambar yang besar. Proses ini bisa berjalan dengan hanya 1-3 gambar dan tidak memerlukan gambar terdistribusi secara merata ke seluruh *view* kamera. Hal ini dapat disebabkan karena proses ini tidak lagi perlu mengestimasi matriks intrinsik dan parameter distorsi namun hanya memproses matriks rotasi dan translasi dari proyeksi titik 3D ke titik 2D tiap kamera untuk tiap gambar pola kalibrasi. Matriks rotasi dan translasi dari kamera 1 dan kamera 2 tersebut akan kemudian diproses sehingga memenuhi persamaan sehingga didapati matriks rotasi dan translasi dari kamera pertama ke kamera yang lainnya.

Meskipun didapati juga dari percobaan bahwa perbedaan jumlah gambar yang diproses mempengaruhi nilai dari matriks rotasi dan translasi yang dihasilkan serta juga dipercaya bahwa semakin besar jumlah gambar yang diproses akan menghasilkan estimasi yang lebih akurat, namun dikarenakan rotasi dan translasi yang sebenarnya tidak bisa didapatkan, maka jumlah pola gambar yang akan diambil disesuaikan dengan jumlah gambar saat melakukan proses kalibrasi tunggal yaitu kurang lebih 10 gambar.

Adapun satu permasalahan yang terdapat pada proses ini adalah terdapat pasangan gambar tertentu yang menyebabkan algoritma kalibrasi stereo menghasilkan *error* yang besar seperti yang terdapat pada gambar IV.3 pada bagian yang disorot berikut.



Gambar IV.3. *Error* Proyeksi Dari Proses Kalibrasi Kamera Tunggal.

Setelah diselidiki didapati bahwa *error* tersebut disebabkan oleh kesalahan urutan pendeteksian titik pada bagian proses pendeteksian tepi dalam pola kalibrasi kedua kamera seperti pada gambar IV.4 berikut.



Gambar IV.4. Gambar (a) urutan deteksi dari kiri ke kanan dimulai dengan warna ungu ke merah. Gambar (a) urutan deteksi dari kiri ke kanan dimulai dengan warna merah ke ungu.

Hal ini terbukti fatal untuk proses kalibrasi stereo karena jika urutan dari titik yang terdeteksi dari salah, maka korespondensi antar titik kedua kamera pun akan salah. Setelah diselidiki, penyebab dari peristiwa ini adalah fungsi *cv.findChessBoardCorners* mendeteksi tepi dalam dengan urutan yang *random* sehingga fungsi ini tidak bisa diterapkan pada proses ini. Adapun fungsi lain yang melakukan fungsi yang sama namun dapat mendeteksi tepi dalam dengan urutan yang lebih konsisten yaitu *cv.findChessboardCornersSB*. Meskipun pengetahuan mengenai proses dalam dari kedua fungsi tersebut tidak diketahui, namun melalui beberapa iterasi percobaan didapati bahwa fungsi *cv.findChessboardCornersSB* tidak melakukan pendeteksian tepi dalam secara *random* namun dimulai dari tepi dalam petak ujung yang berwarna putih. Permasalahan pada pola kalibrasi yang digunakan adalah terdapat 2 ujung petak yang berwarna putih hingga terdapat 2 kemungkinan urutan awal pendeteksian petak. Namun, untuk mempercepat jalannya percobaan, diputuskan bahwa jika terdapat kasus di mana urutan petak berbeda, pasangan gambar tersebut akan diabaikan sehingga tidak merusak proses kalibrasi stereo.

## Estimasi 3D Dan Optimasi Hasil Estimasi 3D

Untuk memudahkan perbandingan kinerja dari hasil estimasi 3D dengan triangulasi dan hasil optimasi estimasi 3D dengan *bundle adjustment*, pembahasan mengenai hasil dan analisis dari kedua proses tersebut akan dibahas pada subbab ini.

Untuk mendapatkan hasil triangulasi terbaik, akan dilakukan pengujian terhadap 4 rentang rotasi antar kamera yaitu 0°-10°, 25°-35°, 40°-50°, 51°-70° seperti yang telah dibahas pada subbab III.3.5. Hasil rotasi dan translasi dari proses kalibrasi stereo untuk setiap rentang pengujian ada pada Tabel IV.2 berikut.

Tabel IV.2. Tabel Rotasi dan Translasi Kamera

| No. Uji (Rentang Uji) | Kamera | Sumbu | Rotasi Kamera | Translasi Kamera |
| --- | --- | --- | --- | --- |
| Uji rentang 1  (0°-10°) | Kiri | x |  |  |
| y |  |  |
| z |  |  |
| Kanan | x |  |  |
| y |  |  |
| z |  |  |
| Uji rentang 2  (25°-35°) | Kiri | x |  |  |
| y |  |  |
| z |  |  |
| Kanan | x |  |  |
| y |  |  |
| z |  |  |
| Uji rentang 3  (40°-50°) | Kiri | x |  |  |
| y |  |  |
| z |  |  |
| Kanan | x |  |  |
| y |  |  |
| z |  |  |
| Uji rentang 4  (51°-70°) | Kiri | x |  |  |
| y |  |  |
| z |  |  |
| Kanan | x |  |  |
| y |  |  |
| z |  |  |

Perlu diperhatikan bahwa rotasi dan translasi kamera-kamera tersebut diukur dari besar perbedaan sudut dan perpindahan dari kamera utama yaitu kamera tengah. Besar perbedaan sudut dan perpindahan kamera tengah sendiri adalah 0 karena kamera tengah tersebut tidak memiliki perbedaan sudut atau perpindahan dengan dirinya sendiri.

Pada tiap rentang pengujian akan dicari pengujian ke berapa dan pasangan kamera triangulasi yang mana yang menghasilkan *reprojection error* yang terendah. Untuk setiap pengujian akan disediakan data letak 2D penanda pada saat aktor berada pada pose berdiri tegak. Selebihnya mengenai gambar pose aktor untuk setiap uji ada pada gambar IV.5. berikut.



Gambar IV.5. gambaranaktor dari *view* dan rotasi yang berbeda-beda

Hasil *reprojection error* untuk tiap pengujian terdapat pada tabel IV.3 di bawah ini

Tabel IV.3. Tabel Rata-rata Kesalahan Proyeksi Balik

| No. Uji | Pasangan Kamera Triangulasi | *Reprojection Error* | |
| --- | --- | --- | --- |
| Triangulasi | Hasil Optimasi *Bundle Adjustment* |
| Uji rentang 1 | Kiri-tengah | 10.308 | 5.03 |
| Tengah-kanan | 9.282 | 2.451 |
| Kiri-kanan | 8.756 | 4.607 |
| 3 kamera bersamaan | 8.313 | 4.031 |
| Uji rentang 2 | Kiri-tengah | 58.96 | 5.802 |
| Tengah-kanan | 88.511 | 8.622 |
| Kiri-kanan | 51.097 | 7.981 |
| 3 kamera bersamaan | 50.071 | 6.82 |
| Uji rentang 3 | Kiri-tengah | 131.797 | 10.074 |
| Tengah-kanan | 128.11 | 10.251 |
| Kiri-kanan | 138.264 | 17.431 |
| 3 kamera bersamaan | 112.313 | 13.691 |
| Uji rentang 4 | Kiri-tengah | 99.191 | 29.079 |
| Tengah-kanan | 148.457 | 45.992 |
| Kiri-kanan | 615.009 | 116.595 |
| 3 kamera bersamaan | 120.01 | 27.949 |

Dari tabel IV.3 tersebut didapati beberapa temuan sebagai berikut,

1. Semakin besar sudut antar kamera, semakin buruk hasil *reprojection error*-nya. Hal ini dapat dikarenakan semakin besar perbedaan rotasi dengan kamera tengah, semakin besar pula jarak minimal antara garis yang satu dengan garis yang lain di mana jarak minimal tersebut adalah lokasi titik 3D yang diestimasi. Semakin kecil jarak minimal tersebut menandakan garis-garis semakin berpotongan dan semakin besar jarak minimal tersebut menandakan ketidakyakinan perpotongan garis-garis dan lokasi titik 3D semakin membesar sehingga dihasilkan *reprojection error*  yang semakin besar.
2. Penggunaan optimasi *bundle adjustment* terhadap hasil triangulasi sangat mempengaruhi penurunan rata-rata kesalahan proyeksi balik. Hal ini membuktikan bahwa teori penggunaan *bundle adjustment* untuk optimasi estimasi 3D terbukti berhasil dan berguna.
3. Sebelum dilakukan optimasi *bundle adjustment*, triangulasi menggunakan ketiga kamera secara bersamaan relatif menghasilkan nilai rata-rata kesalahan proyeksi balik yang paling rendah. Hal ini membuktikan bahwa teori estimasi 3D menggunakan lebih dari 2 kamera atau multi-kamera mampu menurunkan *reprojection error* .
4. Setelah dilakukan optimasi *bundle adjustment*, triangulasi menggunakan ketiga kamera secara bersamaan tidak selalu menghasilkan nilai rata-rata kesalahan proyeksi balik yang paling rendah. Hal ini dapat dikarenakan faktor estimasi awal untuk pasangan kamera triangulasi yang bersangkutan lebih baik dibandingkan estimasi awal triangulasi dengan 3 kamera sekaligus dalam memudahkan algoritma optimasi untuk menemukan titik yang optimal. Mengingat bahwa algoritma optimasi yang digunakan menggunakan prinsip turunan sehingga estimasi awal sangat berpengaruh untuk menemukan lokal minimum terbaik.

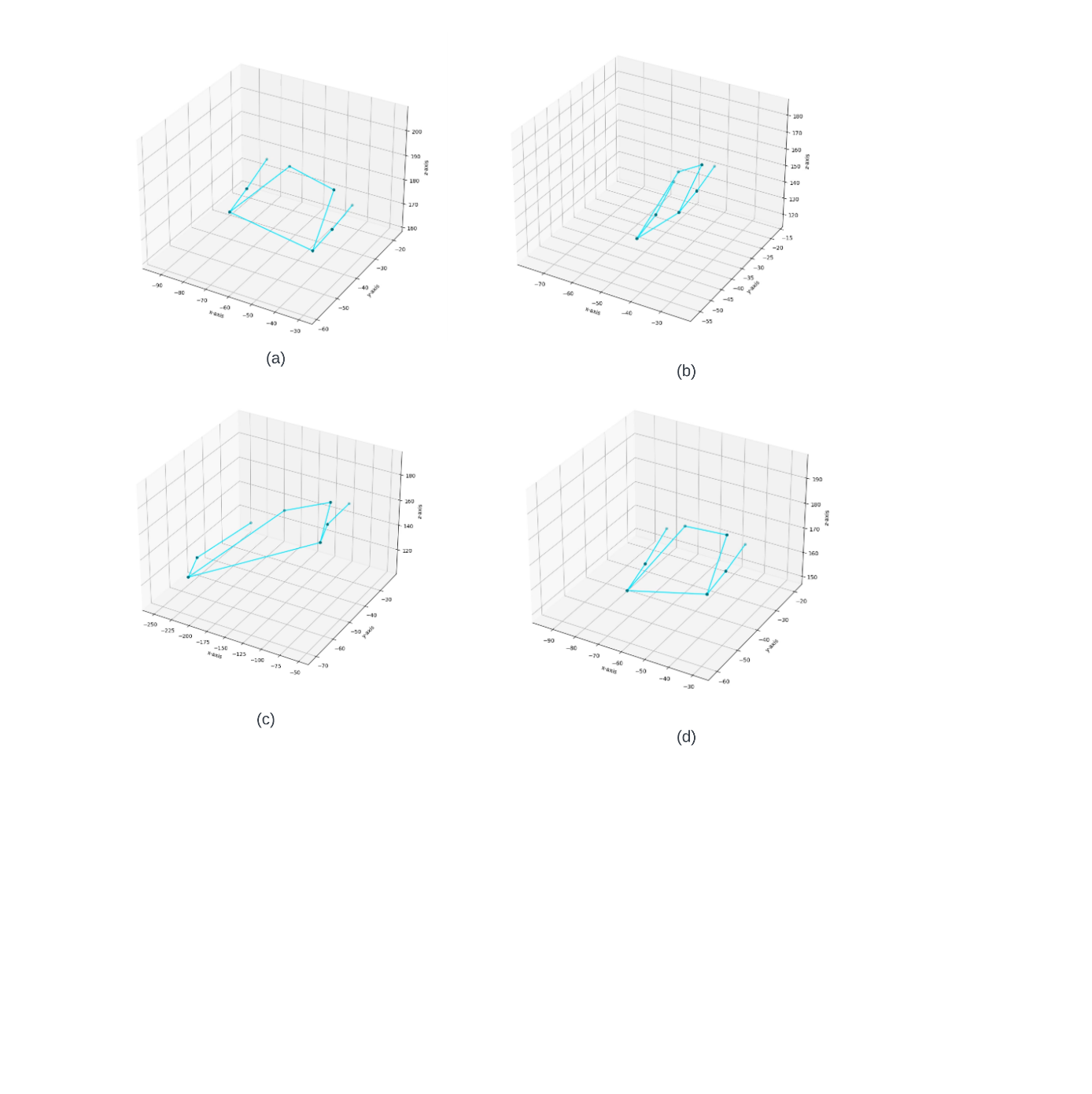
Berdasarkan percobaan yang dilakukan di atas didapati bahwa rentang rotasi terkecil yaitu 0°-10° adalah rentang terbaik untuk mendapatkan *reprojection error* terkecil dengan harapan menghasilkan kualitas estimasi 3D yang terbaik.

Adapun akan dilakukan penilaian lain terhadap kualitas 3D yang dihasilkan dari proses triangulasi dan *bundle adjustment* dengan tidak melihat *reprojection error*-nya namun melihat struktur keseluruhan 3D yang dihasilkan. Dikarenakan tidak bisa didapatkan data 3D asli sehingga penilaian kualitas dilakukan secara kualitatif dan bukan kuantitatif seperti pada penilaian kinerja estimasi 3D terhadap *reprojection*-nya.

Untuk melakukan penilaian terhadap kualitas struktur 3D yang dihasilkan, akan dilakukan 2 uji gerakan yaitu gerakan pertama yaitu pose-A yang dinaikan dan diturunkan di samping badan aktor dan gerakan kedua yaitu maju dan mundur. Uji gerakan pertama digunakan sebagai pengujian dasar sistem untuk menilai apakah sistem dapat menghasilkan kualitas struktur yang baik dari gerakan yang paling sederhana. gerakan kedua digunakan untuk menguji apakah sistem dapat dengan benar mengestimasi kedalaman di mana aktor melakukan gerakan maju dan mundur. Uji gerakan ini juga dilakukan pada 4 rentang uji yang telah ditentukan.

Dari percobaan yang telah dilakukan didapati beberapa temuan sebagai berikut,

1. Penggunaan pasangan kamera berpengaruh terhadap struktur 3D yang dihasilkan. Hal ini semakin terlihat pada saat digunakan nilai rotasi kamera yang besar sebelum dilakukan optimasi *bundle adjustment*. Sebagai contoh perhatikan Gambar IV.6 berikut.



Gambar IV.6. Gambar (a),(b),(c),dan (d) secara berturut-turut adalah hasil struktur 3D dari proses triangulasi untuk pasangan kamera kiri-tengah, kanan-tengah, kiri-kanan, dan 3 kamera sekaligus.

Pada gambar IV.6. tersebut pasangan kamera kiri-tengah menghasilkan struktur 3D yang sedikit miring ke kiri (jika gambar dilihat dari tampak depan), sedangkan pasangan kamera kanan-tengah menghasilkan struktur 3D yang miring ke kanan dengan kemiringan yang cukup tajam, pasangan kamera kiri-kanan menghasilkan gambar yang tidak terlalu miring ke kanan namun struktur 3D yang dihasilkan melebar, dan penggunaan 3 kamera sekaligus menghasilkan gambar dengan kualitas yang baik dengan tidak ada kemiringan yang terlalu terlihat dan bentuk tubuh yang dihasilkan pun terlihat baik. Fenomena kemiringan ini dikarenakan, pada saat sudut kedua kamera cukup lebar, garis-garis dari kedua kamera tersebut akan semakin tegak lurus dan struktur 2D pada kamera yang sejajar terhadap aktor akan semakin menyempit sementara struktur 2D kamera yang berhadapan dengan aktor tetap sama. Penyempitan salah satu struktur 2D inilah yang menyebabkan munculnya efek “penarikan” pada struktur 3D yang dihasilkan dikarenakan perbedaan jarak tiap titik 2D pada satu kamera berbeda dengan kamera yang lain. Adapun jika sudut antara kedua pasang kamera lebih besar lagi sehingga kamera hampir saling berhadapan, penyempitan struktur 2D akan terjadi pada kedua kamera dan dikarenakan garis-garis dari kedua kamera tersebut semakin paralel maka titik perpotongan kedua garis tersebut akan semakin jauh sehingga muncul efek penarikan dari dua sisi pada hasil struktur 3D seperti pada gambar (c) pada gambar IV.5.

Adapun triangulasi menggunakan 3 kamera tidak menyebabkan efek penarikan pada struktur 3D yang dihasilkan. Hal ini dikarenakan garis proyeksi masing-masing kamera saling memperbaiki kemiringan dari pasangan-pasangan kamera triangulasi sehingga tidak terjadi fenomena kemiringan atau penarikan. Dengan kata lain, penambahan kamera membuat efek penarikan struktur 3D menjadi seimbang untuk tiap pasangan kamera sehingga tidak ada efek penarikan yang terlihat.

1. Penggunaan *bundle adjustment* dapat memperbaiki kemiringan struktur 3D hasil triangulasi. Berikut pada Gambar IV.7. adalah salah satu contohnya.

A graph of a graph

Description automatically generated

Gambar IV.7. Gambar (a) struktur 3D kanan-tengah sebelum optimasi. Gambar (b) struktur 3D kanan-tengah setelah optimasi.

Gambar (a) pada gambar IV.7. adalah gambar (b) pada (gambar) sebelum dilakukan optimasi *bundle adjustment* sementara gambar (b) pada (gambar) setelah dilakukan optimasi *bundle adjustment*. Dapat dilihat bahwa gambar (b) tersebut hampir mirip dengan gambar (d) pada (gambar) di mana struktur 3D yang dihasilkan tidak mengalami kemiringan. Hal ini dikarenakan optimasi *bundle adjustment* mempertimbangkan *reprojection error* dari semua kamera dan mengoptimasi hasil estimasi 3D awal supaya *reprojection error* untuk semua *view* minimum. Oleh karena itu, struktur 3D hasil optimasi mampu menghilangkan kemiringan dan membuat hasilnya serupa dengan struktur 3D dengan menggunakan 3 kamera bersamaan.

1. Sistem mampu menangkap struktur dan perubahan kedalaman pada gerakan maju-mundur dengan baik, namun dari segi *reprojection error* semakin dekat aktor ke kamera, semakin besar *reprojection error*-nya. Hal ini dapat disebabkan karena semakin dekat aktor menuju kamera, sudut antara garis-garis proyeksi kamera yang berpotongan semakin menjadi lebih lebar yang mana serupa dengan penjelasan no.1 pada percobaan sebelumnya mengenai pencarian rentang rotasi terbaik, semakin besar sudut perpotongan garis, semakin besar ketidakyakinan estimasi 3D dan semakin besar pula *reprojection error*-nya.
2. Terdapat titik-titik yang terkadang gagal diestimasi 3Dnya, menjadi sebuah outlier, atau tidak sesuai dengan struktur aslinya. Contohnya seperti pada gambar IV.8 berikut,

A graph of a graph with lines and dots

Description automatically generated

Gambar IV.8. Salah satu titik yang dihasilkan tidak sesuai dengan struktur 3D aslinya.

Hal ini bisa jadi disebabkan oleh berbagai hal seperti *noise* dari data 2D, *noise* dari hasil estimasi matriks intrinsik dan ekstrinsik kamera, dll. yang menimbulkan *error* pada proses estimasi 3D.

# KESIMPULAN DAN SARAN

Pada bagian ini akan dibahas kesimpulan dari evaluasi kinerja sistem yang telah dibangun beserta saran terkait proses-proses yang sebaiknya dilakukan atau tidak dilakukan ketika membangun sistem yang serupa dan yang dapat dikembangkan kemudian.

## Kesimpulan

Berdasarkan tujuan dari tugas akhir ini dan evaluasi yang telah dilakukan, bisa diputuskan bahwa tugas akhir ini berhasil membangun sistem multi-kamera untuk melakukan estimasi 3D penanda-penanda pada tubuh aktor dan dengan demikian menjawab permasalahan yang telah didefinisikan.

Adapun berdasarkan temuan dan analisis yang telah dilakukan, dapat disimpulkan bahwa,

1. rentang rotasi terbaik untuk melakukan rotasi 3D adalah rentang rotasi 0°-10° karena pada rentang tersebut diperoleh *reprojection error* yang terendah,
2. estimasi 3D dengan 3 kamera sekaligus menghasilkan estimasi 3D dengan struktur 3D yang terbaik khususnya pada saat sudut dan jarak antar kamera besar serta menghasilkan *reprojection error* yang umumnya paling rendah,
3. penggunaan *bundle adjustment* sangat membantu dalam mempertahankan struktur 3D terbaik khususnya pada saat sudut dan jarak antar kamera besar serta menurunkan *reprojection error* secara signifikan,
4. *noise* dapat merusak struktur hasil estimasi 3D.

## Saran

Adapun berdasarkan evaluasi yang telah dilakukan, untuk pengembangan selanjutnya disarankan,

1. melakukan percobaan dan perbandingan lebih lanjut mengenai pengaruh rotasi dan translasi antar kamera terhadap hasil estimasi 3D,
2. melakukan perbandingan lebih lanjut terhadap pengaruh penambahan kamera terhadap hasil estimasi 3D,
3. melakukan percobaan dengan gerakan yang lebih variatif dan melihat pengaruhnya terhadap hasil estimasi 3D,
4. melakukan percobaan gerakan yang memberikan oklusi kepada Sebagian kamera dan melihat pengaruhinya terhadap hasil estimasi 3D.
5. Melakukan upaya lebih lanjut untuk mengurangi *noise* pada setiap proses estimasi 3D.

DAFTAR REFERENSI

(2016, Oktober 22). Diambil kembali dari Large-scale bundle adjustment in scipy: https://scipy-cookbook.readthedocs.io/items/bundle\_adjustment.html

AmyTabb. (2021, Oktober 31). *N-view triangulation: DLT methods 2 and 3 (2)*. Diambil kembali dari https://amytabb.com/tips/tutorials/2021/10/31/triangulation-DLT-2-3/

Batpurev, T. (2021, Februari 2). *Stereo Camera Calibration and Triangulation with OpenCV and Python*. Diambil kembali dari https://temugeb.github.io/opencv/python/2021/02/02/stereo-camera-calibration-and-triangulation.html

Hartley, R., & Zisserman, A. (2004). *Multiple View Geometry in Computer Vision.* Cambridge University Press.

Hata, K., & Savarese, S. (t.thn.). *CS231A Course Notes 1: Camera Models.* Diambil kembali dari https://web.stanford.edu/class/cs231a/course\_notes/01-camera-models.pdf.

Kaehler, A., & Bradski, G. R. (2017). *Learning OpenCV 3 : computer vision in C++ with the OpenCV library.* O’reilly Media.

Kitani, K. (2007, April 5). *Epipolar Geometry.* Diambil kembali dari http://www.cs.cmu.edu/~16385/s15/lectures/Lecture18.pdf

Kitani, K. (2007, April 5). *Triangulation.* Diambil kembali dari http://www.cs.cmu.edu/~16385/s17/Slides/11.4\_Triangulation.pdf

Zhang, Z. (1998). *A Flexible New Technique for Camera Calibration.*