

**PEMBANGUNGAN SISTEM MULTI-KAMERA UNTUK
MENGEVALUASI REPROJECTION ERROR ESTIMASI 3D
PENANDA SISTEM TANGKAP GERAK SEDERHANA**

**Laporan Tugas Akhir – Capstone Pembangunan Sistem Tangkap
Gerak Multi-Kamera dengan Penanda sebagai Pembanding
Penilaian Kinerja Sistem Tangkap Gerak Monokuler Tanpa
Penanda**

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
MENGEVALUASI REPROJECTION ERROR ESTIMASI 3D
PENANDA SISTEM TANGKAP GERAK SEDERHANA**

**Laporan Tugas Akhir – Capstone Pembangunan Sistem Tangkap
Gerak Multi-Kamera dengan Penanda sebagai Pembanding
Penilaian Kinerja Sistem Tangkap Gerak Monokuler Tanpa
Penanda**

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 23 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 <i>Mocap</i> 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	Sarana pembanding sistem tangkap gerak dengan metode Avatar Digital
4	13519149	Syihabuddin Yahya Muhammad	<i>object detection</i> dan <i>object tracking</i>

Bandung, 23 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, 21 Juli 2023



Moses Ananta

NIM 13519076

ABSTRAK

PEMBANGUNGAN SISTEM MULTI-KAMERA UNTUK MENGEVALUASI REPROJECTION ERROR ESTIMASI 3D PENANDA 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. Kinerja sistem akan dinilai dengan metrik *reprojection error*.

Sistem multi-kamera yang dibangun diuji dengan 4 rentang rotasi antar kamera yang berbeda dan dari hasil percobaan serta analisis terhadap sistem yang dibuat didapati bahwa *reprojection error* dapat digunakan sebagai indikasi hasil estimasi 3D yang salah, rotasi tidak mempengaruhi *reprojection error*, dan triangulasi dengan 3 kamera sekaligus menghasilkan *reprojection error* yang paling rendah serta hasil 3D yang terbaik jika semua kamera sinkron dalam mengambil gambar.

Kata kunci: sistem tangkap gerak, multi-kamera, triangulasi, *reprojection error*.

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 Mengevaluasi *Reprojection Error* Estimasi 3D Penanda 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, 21 Juli 2023

Penulis

DAFTAR ISI

BAB I PENDAHULUAN.....	1
I.1 Latar Belakang.....	1
I.2 Rumusan Masalah.....	2
I.3 Tujuan	3
I.4 Batasan Masalah	3
I.5 Metodologi.....	4
I.6 Sistematika Pembahasan.....	5
BAB II STUDI LITERATUR	6
II.1 Pembentukan Gambar pada Kamera.....	6
II.2 Distorsi Lensa	9
II.3 Kalibrasi Kamera Tunggal.....	14
II.4 Perbaikan Distorsi.....	15
II.5 Triangulasi	17
II.6 Kalibrasi Stereo.....	20
II.7 Multi-Kamera.....	21
II.8 <i>Reprojection Error</i>	23
II.9 Bundle Adjustment	23
II.10 OpenCV	24
BAB III PENGEMBANGAN SISTEM MULTI-KAMERA.....	25
III.1 Analisis Persoalan.....	25
III.2 Analisis Solusi	26
III.3 Rancangan Solusi.....	30

III.3.1	Kalibrasi Kamera Tunggal.....	30
III.3.2	Penghilangan Distorsi Gambar	32
III.3.3	Pemilihan Kamera Utama	32
III.3.4	Kalibrasi Kamera Stereo	33
III.3.5	Estimasi 3D Penanda dengan Triangulasi	34
III.3.6	Optimasi Hasil Estimasi 3D dengan <i>Bundle Adjustment</i>	37
BAB IV EVALUASI SISTEM		39
IV.1	Kalibrasi Kamera Tunggal dan Penghilangan Distorsi Gambar.....	39
IV.2	Kalibrasi Stereo.....	42
IV.3	Estimasi 3D dan Optimasi Hasil Estimasi 3D	44
BAB V KESIMPULAN DAN SARAN		50
V.1	Kesimpulan	50
V.2	Saran	51

DAFTAR GAMBAR

Gambar II.1. Model kamera lubang jarum (Bradski & Kaehler, 2008).....	6
Gambar II.2. Dampak bukaan kamera terhadap kejelasan gambar (Hata & Savarese)	10
Gambar II.3. Penggunaan lensa untuk memperjelas gambar (Hata & Savarese) .	10
Gambar II.4. Jenis-jenis distorsi kamera (Hata & Savarese)	11
Gambar II.5. Lensa yang tidak sejajar dan memiliki sudut θt terhadap titik pusat bidang gambar (Steward, 2021)	12
Gambar II.6. Bidang dengan warna hitam adalah gambar sebelum adanya distorsi dan gambar dengan warna jingga adalah gambar yang terdistorsi tangensial (Steward, 2021).....	13
Gambar II.7. Permasalahan proyeksi balik ke 3D dengan hanya 1 <i>view</i> saja (Kitani, Triangulation, 2007)	17
Gambar II.8. Gambaran Triangulasi dengan 2- <i>view</i> (Kitani, Epipolar Geometry, 2007)	18
Gambar II.9. 3 <i>view</i> melihat titik yang sama pada objek (Snavely & Li, 2018)...	22
Gambar II.10. Visualisasi <i>error</i> estimasi 3D. Gambar (a) estimasi 3D dengan hanya 2 <i>View</i> . Gambar (b) estimasi 3D dengan hanya 4 <i>View</i> (Snavely & Li, 2018)	22
Gambar III.1. Pola kalibrasi papan catur yang digunakan.	31
Gambar III.2. Gambaran urutan kamera yang digunakan.....	33
Gambar III.3. Gambar pola kalibrasi diambil secara bersamaan oleh kamera kiri (a), kamera tengah (b), dan kamera kanan (c).....	34
Gambar III.4. Kondisi ruangan. Bagian merah mengindikasikan daerah ruangan yang tidak boleh terlihat oleh kamera.	35

Gambar III.5. Konfigurasi peletakan kamera untuk setiap rentang rotasi. (a) rentang rotasi dengan kamera tengah 0° - 10° . (b) rentang rotasi dengan kamera tengah 25° - 35° . (c) rentang rotasi dengan kamera tengah 40° - 50° . (d) rentang rotasi dengan kamera tengah 51° - 70°	36
Gambar IV.1. Gambar (a) gambar asli sebelum dilakukan perbaikan distorsi. Gambar (b) gambar setelah dilakukan perbaikan distorsi.....	40
Gambar IV.2. Gambar (a) gambar asli sebelum perbaikan distorsi. Gambar (b) gambar setelah perbaikan distorsi yang masih salah. (c) gambar setelah perbaikan distorsi yang sudah diperbaiki. Daerah hitam, yang menunjukkan perbaikan distorsi pada gambar (c) lebih kecil dibandingkan gambar (b), yang mana gambar (c) lebih mirip dengan gambar (a) dibandingkan gambar (b).....	41
Gambar IV.3. Gambaran aktor dari <i>view</i> dan rotasi yang berbeda-beda	44
Gambar IV.4. Visualisasi gerakan kiri-kanan. Struktur 3D “mengapung” mengapung ke atas. Sumbu x menyatakan arah kanan-kiri. Sumbu y menyatakan arah atas-bawah. Sumbu z menyatakan arah depan-belakang.	47
Gambar IV.5. (a), (b), (c) Gambar kamera kiri-tengah-kanan secara berturut-urut. Penanda yang terdeteksi ditandai dengan titik biru.....	47
Gambar IV.6. Dua gambar di bagian atas adalah gambar yang sama dengan Gambar IV.4, gambar bawah adalah visualisasi <i>reprojection error</i> untuk tiap <i>frame</i> dengan bagian yang diberi warna kuning adalah <i>reprojection error frame</i> saat ini	48

DAFTAR TABEL

Tabel III.1. Koordinat 3D tepi dalam petak-petak papan catur.....	32
Tabel IV.1. Parameter intrinsik dan distorsi lensa hasil kalibrasi kamera tunggal	42
Tabel IV.2. Parameter intrinsik dan distorsi lensa hasil kalibrasi kamera tunggal	42
Tabel IV.3. Tabel rotasi dan translasi kamera.....	43
Tabel IV.4. Tabel rata-rata kesalahan proyeksi balik	45

DAFTAR SIMBOL NOTASI

$\mathbf{X} = (X, Y, Z)$	Titik 3D pada koordinat dunia
$\mathbf{X}_{kamera} = (x, y, z)$	Titik 3D pada koordinat kamera
$\mathbf{x} = (\hat{x}, \hat{y})$	Titik 2D pada koordinat gambar
(u, v)	Titik 2D pada koordinat gambar yang sudah ternormalisasi
k	Faktor skala pada koordinat gambar
f	<i>Focal length</i>
(f_x, f_y)	<i>Focal length</i> untuk sumbu x dan y pada kamera
(c_x, c_y)	<i>Offset</i> titik pusat kamera untuk sumbu x dan y
K	Matriks intrinsik kamera
R	Matriks rotasi 3x3
t	Matriks translasi 3x1
$[R t]$	Matriks ekstrinsik kamera
$P = K[R t]$	Matriks proyeksi kamera 3x4
A	Matriks triangulasi
$d(x, PX)$	Fungsi Jarak antara titik 2D pada gambar dan proyeksi balik titik 3D ke titik 2D-nya.

BAB I

PENDAHULUAN

I.1 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.

I.2 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 gerak 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. Untuk mengukur kinerja dari sistem yang akan dibangun tersebut, digunakan metrik *reprojection error* yang mengukur besarnya perbedaan jarak antara data 2D penanda dengan proyeksi balik estimasi 3D penanda ke koordinat 2D-nya di bidang gambar. Dengan demikian, permasalahan yang akan diselesaikan pada tugas akhir ini adalah “Pembangunan sistem multi-kamera untuk mengestimasi letak 3D penanda titik-titik penting aktor dan mengevaluasi *reprojection error* hasil estimasi 3D-nya pada sistem tangkap gerak sederhana.”

I.3 Tujuan

Berdasarkan rumusan masalah yang telah disampaikan sebelumnya, tujuan dari tugas akhir ini adalah,

1. membangun sebuah sistem multi-kamera yang dapat melakukan estimasi 3D penanda titik-titik penting aktor,
2. mengevaluasi *reprojection error* dari hasil estimasi 3D sistem multi-kamera yang telah dibangun.

I.4 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 mencakup distorsi radial dan distorsi tangensial,
4. 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.

I.5 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.

2. 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

3. 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.

4. 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.

5. Evaluasi sistem

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

I.6 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.

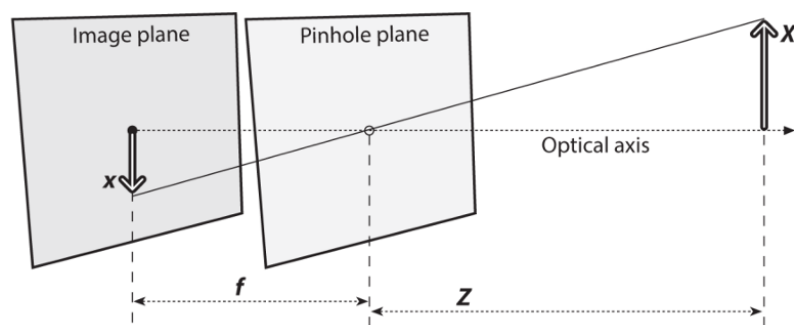
BAB II

STUDI LITERATUR

II.1 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.



Gambar II.1. Model kamera lubang jarum (Bradski & Kaehler, 2008)

Pada model kamera lubang jarum, titik pada koordinat dunia $\mathbf{X} = (X, Y, Z)$ dipetakan ke dalam titik \mathbf{x} pada bidang gambar seperti yang terdapat pada Gambar II.1. Menggunakan perbandingan segitiga serupa, penjabarannya sebagai berikut

$$\begin{aligned}\frac{x}{f} &= \frac{X}{Z} \\ x &= f \frac{X}{Z} \\ \mathbf{x} &= f \frac{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}{Z} \\ \mathbf{x} &= \begin{bmatrix} \frac{fX}{Z} \\ \frac{fY}{Z} \\ f \end{bmatrix}\end{aligned}\tag{II. 1}$$

Dengan membuang koordinat terakhir pada bidang gambar, titik 2D yang terbentuk adalah $\mathbf{x} = \left(\frac{fX}{Z}, \frac{fY}{Z}\right)$. Persamaan berikut menggambarkan transformasi menggunakan sistem koordinat homogen,

$$k \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 \\ f & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}\tag{II. 2}$$

Yang mana (u,v) adalah koordinat pada bidang gambar yang sudah ternormalisasi, 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 *offset* antara titik asal koordinat bidang gambar dengan titik utamanya sehingga *offset* tersebut perlu diperhitungkan ke dalam persamaan sehingga membentuk persamaan berikut,

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{fX}{Z} + c_x \\ \frac{fY}{Z} + c_y \end{bmatrix} \quad (\text{II. 3})$$

Yang mana c_x dan c_y adalah *offset* pada sumbu x dan sumbu y.

Dengan persamaan pada sistem koordinat homogen sebagai berikut,

$$k \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & c_x & 0 \\ f & c_y & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (\text{II. 4})$$

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

$$K = \begin{bmatrix} f & c_x \\ f & c_y \\ 1 & 1 \end{bmatrix} \quad (\text{II. 5})$$

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

$$\mathbf{x} = K[I|0] \mathbf{X}_{kamera} \quad (\text{II. 6})$$

Yang mana I adalah matriks identitas dan $[I|0]$ adalah matriks 3x4 dengan matriks 3x3 pertama adalah matriks identitas dan kolom terakhir yang berisi nilai 0.

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.

$$K = \begin{bmatrix} f_x & c_x \\ f_y & c_y \\ 1 & 1 \end{bmatrix} \quad (\text{II. 7})$$

Di mana f_x adalah *focal length* f dibagi dengan panjang *pixel* dan f_y adalah *focal length* f dibagi dengan lebar *pixel*.

Adapun persamaan II.6 mengasumsikan bahwa titik (X,Y,Z) 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,

$$\mathbf{X}_{kamera} = [R|t] \mathbf{X} \quad (II.8)$$

Dengan \mathbf{X}_{kamera} adalah titik pada kerangka koordinat kamera, R adalah matriks rotasi berukuran 3×3 , t adalah matriks translasi berukuran 3×1 , dan \mathbf{X} yang adalah titik pada kerangka koordinat dunia.

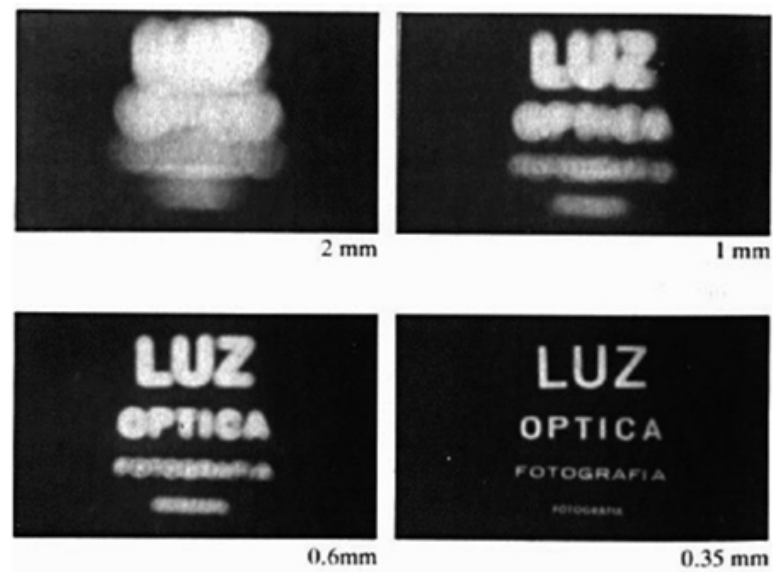
Dengan menggabungkan persamaan II.6 dengan persamaan II.4, terbentuk persamaan berikut,

$$\mathbf{x} = K[R|t] \mathbf{X} \quad (II.9)$$

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

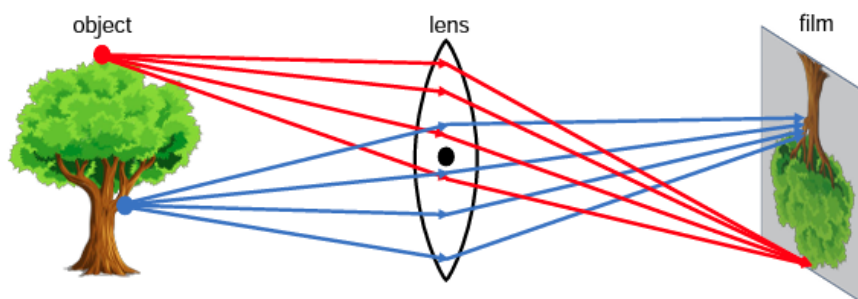
II.2 Distorsi Lensa

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. Seperti pada Gambar II.2 berikut,



Gambar II.2. Dampak bukaan kamera terhadap kejelasan gambar (Hata & Savarese)

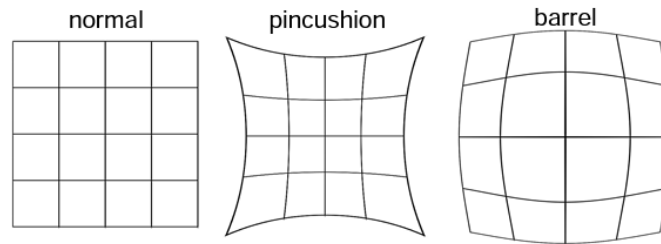
Salah satu upaya untuk mengatasi kendala ini adalah dengan menggunakan lensa yang dapat memfokuskan atau menghamburkan cahaya. Menggunakan lensa ini, sinar-sinar cahaya yang dipancarkan oleh suatu titik pada dunia dapat difokuskan oleh lensa ke suatu titik pada bidang gambar sehingga hasil proyeksi gambar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



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)

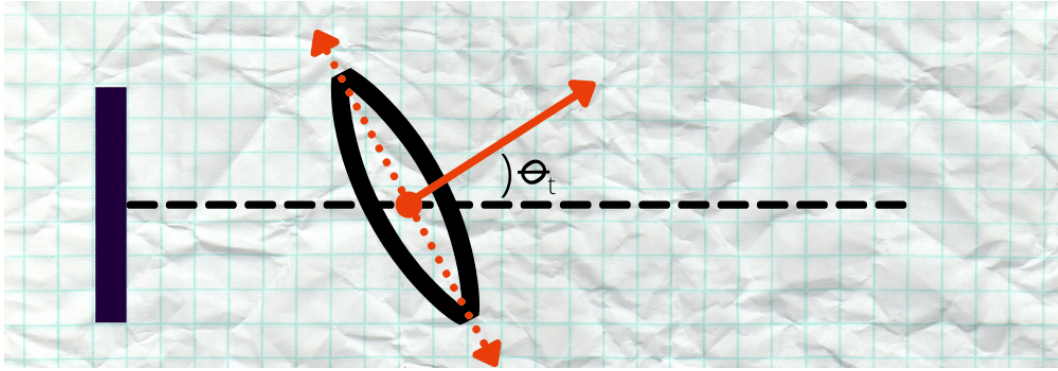
Distorsi radial dapat dimodelkan dengan persamaan matematis berikut,

$$\begin{aligned}x_{terdistorsi} &= x(1 + k_1r^2 + k_2r^4 + k_3r^6) \\y_{terdistorsi} &= y(1 + k_1r^2 + k_2r^4 + k_3r^6) \\r &= \sqrt{x^2 + y^2}\end{aligned}\tag{II. 10}$$

Di mana x dan y adalah titik pada koordinat kamera sebelum terdistorsi, r adalah panjang radius dari tengah gambar, dan k_1, k_2, k_3 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.

$$\begin{aligned}
x_{terdistorsi} &= x \frac{(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)}{(1 + k_4 r^2 + k_5 r^4 + k_6 r^6)} \\
y_{terdistorsi} &= y \frac{(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)}{(1 + k_4 r^2 + k_5 r^4 + k_6 r^6)} \\
r &= \sqrt{x^2 + y^2}
\end{aligned} \tag{II. 11}$$

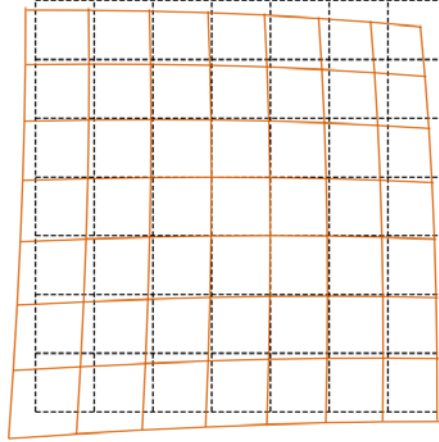
Adapun jenis distorsi lain yang dapat muncul akibat penggunaan lensa yaitu distorsi tangensial atau disebut juga sebagai *de-centering distortion*. Distorsi ini muncul dikarenakan peletakan lensa kamera yang tidak parallel atau sejajar dengan titik pusat (*center*) dari bidang gambar seperti contohnya pada Gambar II.5 berikut,



Gambar II.5. Lensa yang tidak sejajar dan memiliki sudut θ_t terhadap titik pusat bidang gambar (Steward, 2021)

Umumnya kesalahan peletakan lensa ini muncul karena kesalahan pada proses manufaktur kamera. Akibat dari distorsi ini, gambar yang didapati saat pengambilan gambar adalah seperti Gambar II.6. berikut,

Tangential (Decentering) Distortion



Gambar II.6. Bidang dengan warna hitam adalah gambar sebelum adanya distorsi dan gambar dengan warna jingga adalah gambar yang terdistorsi tangensial (Steward, 2021)

Didapati gambar yang terdistorsi terlihat miring dan terotasi terhadap suatu sumbu tertentu, contohnya pada Gambar II.6. di atas, bagian kiri bawah gambar terlihat lebih maju dan bagian kanan atas gambar terlihat lebih maju dibandingkan bagian tengah gambar, yang mana hal tersebut disebabkan oleh efek rotasi yang telah disebutkan.

Brown (1966) memodelkan distorsi tangensial dengan persamaan matematis berikut,

$$\begin{aligned} x_{terdistorsi} &= x + [2p_1xy + p_2(r^2 + 2x^2)] \\ y_{terdistorsi} &= y + [p_1(r^2 + 2y^2) + 2p_2xy] \\ r &= \sqrt{x^2 + y^2} \end{aligned} \quad (II. 12)$$

Yang mana x dan y adalah titik pada koordinat kamera sebelum terdistorsi, r adalah panjang radius dari tengah gambar, dan p_1 dan p_2 adalah koefisien distorsi tangensial.

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

$$parameter_distorsi = [k_1, k_2, p_1, p_2, k_3, k_4, k_5, k_6] \quad (II.13)$$

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

II.3 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 H_0, \dots, H_M 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.

II.4 Perbaikan Distorsi

Seperti yang telah dijelaskan pada subbab II.2, penggunaan lensa pada kamera menimbulkan distorsi pada gambar hasil pengambilan gambar. Oleh karena itu, penting dilakukan perbaikan distorsi.

Cara paling sederhana untuk memperbaiki distorsi adalah menemukan fungsi pemetaan yang memetakan gambar asli yang belum terdistorsi ke gambar yang terdistorsi. Setelah didapatkan fungsi pemetaan tersebut, dimungkinkan untuk melakukan pemetaan balik yang memetakan gambar terdistorsi ke gambar asli yang tidak terdistorsi. Dengan demikian, distorsi dapat diperbaiki.

Jika diberikan suatu gambar kosong (gambar hitam), yang mana gambar ini adalah representasi gambar hasil perbaikan distorsi, dengan ukuran gambar sama dengan ukuran gambar yang akan diperbaiki distorsinya, menggunakan matriks intrinsik yang diperoleh dari metode kalibrasi kamera Zhang, akan dilakukan

proyeksi balik titik-titik pada gambar kosong tersebut ke sistem koordinat kamera sebagai berikut,

$$\begin{aligned}x' &= \frac{u - c_x}{f_x} \\y' &= \frac{v - c_y}{f_y}\end{aligned}\tag{II. 13}$$

Dengan (u,v) adalah titik pada koordinat gambar kosong yang telah ternormalisasi . serupa seperti pada persamaan II.2, (x', y') adalah titik hasil proyeksi balik dari koordinat gambar ke koordinat kamera yang sudah ternormalisasi, c_x, c_y adalah *offset* titik utama pada sumbu x dan y, dan f_x, f_y adalah *focal length* pada sumbu x dan y seperti pada persamaan II.7.

Proyeksi balik dilakukan dikarenakan operasi distorsi diasumsikan dilakukan sebelum pemetaan sistem koordinat kamera 3D ke sistem koordinat gambar 2D.

Selanjutnya akan dilakukan operasi distorsi pada titik hasil proyeksi balik dengan menggabungkan persamaan distorsi II.11 dan II.12 seperti berikut,

$$\begin{aligned}x'' &= x' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + 2p_1 x' y' + p_2 (r^2 + 2x'^2) \\y'' &= y' \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + p_1 (r^2 + 2y'^2) + 2p_2 x' y'\end{aligned}\tag{II. 13}$$

Setelah didapati (x'', y'') yang adalah titik koordinat kamera yang telah terdistorsi, dengan memetakan kembali titik tersebut ke koordinat gambar seperti pada persamaan berikut,

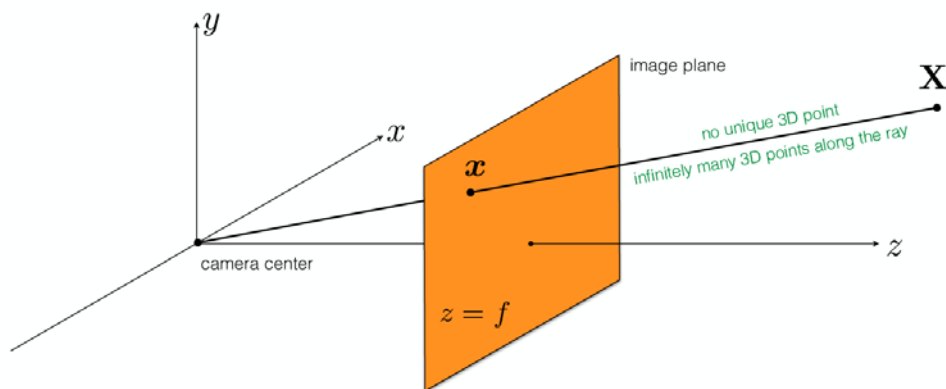
$$\begin{aligned}f_x(u, v) &= x'' f_y + c_x \\f_y(u, v) &= y'' f_y + c_y\end{aligned}\tag{II. 13}$$

Didapati pemetaan gambar yang asli yang tidak terdistorsi ke gambar terdistorsinya. $f_x(u, v)$ dan $f_y(u, v)$ adalah fungsi pemetaan gambar asli ke gambar terdistorsinya untuk sumbu x dan sumbu y. Dengan memanfaatkan gambar

kosong tersebut dan fungsi pemetaan ini, jika diberikan suatu gambar terdistorsi, maka nilai *pixel* gambar tidak terdistorsinya pada titik misalnya (0,0) adalah nilai *pixel* titik $(f_x(0,0), f_y(0,0))$ pada gambar yang terdistorsi, pada titik (0,1) adalah titik $(f_x(0,1), f_y(0,1))$ pada gambar yang terdistorsi, dan seterusnya, hingga didapati gambar asli yang tidak terdistorsi. Untuk pemetaan yang menghasilkan titik di luar gambar terdistorsi, maka nilai *pixel* pada gambar hasilnya akan tetap bernilai kosong (hitam).

II.5 Triangulasi

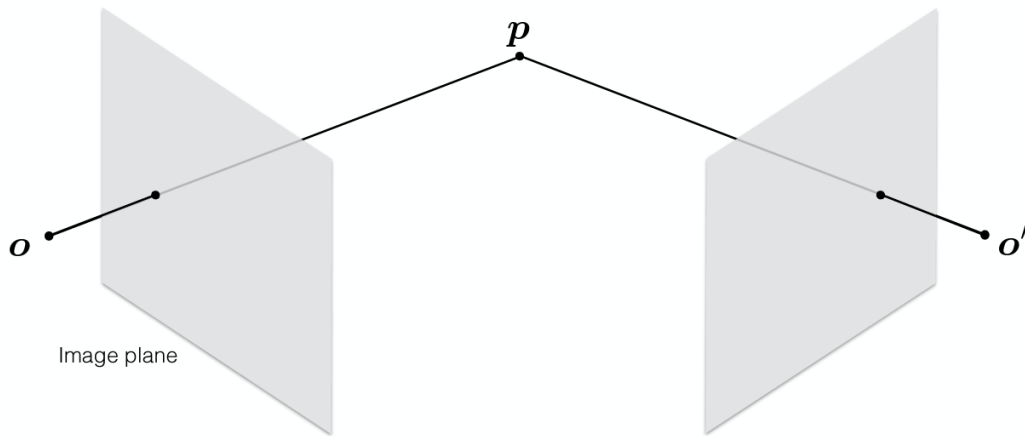
Pada dasarnya matriks intrinsik kamera saja tidak cukup untuk dapat melakukan estimasi 3D dengan baik. Hal ini dikarenakan terdapat tak hingga jumlah nilai 3D yang memenuhi persamaan pemetaan titik 3D ke titik 2D bidang gambar menggunakan matriks intrinsik kamera. Hal ini tergambar seperti pada Gambar II.7 di bawah ini. Oleh karena itu, penggunaan matriks intrinsik saja tidak bisa menghasilkan estimasi 3D yang akurat.



Gambar II.7. 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.8 di bawah ini.



Gambar II.8. Gambaran Triangulasi dengan 2-view (Kitani, Epipolar Geometry, 2007)

Jika titik 3D yang dilihat secara bersamaan pada 2 *view* jatuh pada titik 2D \mathbf{o} pada bidang gambar *view* pertama dan \mathbf{o}' pada bidang gambar *view* kedua, garis proyeksi dari *view* pertama dan *view* kedua akan saling berpotongan pada titik \mathbf{p} di mana titik \mathbf{p} 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 \mathbf{P} adalah sebuah matriks 3x4 hasil perkalian matriks intrinsik dan matriks ekstrinsik.

$$\mathbf{P} = K[R|t] \quad (\text{II. 14})$$

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

$$\mathbf{x} = \mathbf{P}\mathbf{X} \quad (\text{II. 15})$$

$$k \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{P}\mathbf{X}$$

Yang mana titik \mathbf{X} adalah titik pada koordinat dunia, \mathbf{x} titik pada bidang gambar, (u,v) adalah koordinat pada bidang gambar yang telah ternormalisasi, dan k adalah factor skala.

Jika \mathbf{P}^i adalah baris ke i dari matriks \mathbf{P} , maka didapatkan 3 persamaan berikut

$$\begin{aligned} ku &= \mathbf{P}^1\mathbf{X} \\ kv &= \mathbf{P}^2\mathbf{X} \\ k &= \mathbf{P}^3\mathbf{X} \end{aligned} \quad (\text{II. 16})$$

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

$$\begin{aligned} (\mathbf{P}^3\mathbf{X})u &= \mathbf{P}^1\mathbf{X} \\ (\mathbf{P}^3\mathbf{X})v &= \mathbf{P}^2\mathbf{X} \end{aligned} \quad (\text{II. 17})$$

Di mana,

$$\begin{aligned} u\mathbf{P}^3\mathbf{X} - \mathbf{P}^1\mathbf{X} &= 0 \\ v\mathbf{P}^3\mathbf{X} - \mathbf{P}^2\mathbf{X} &= 0 \end{aligned} \quad (\text{II. 18})$$

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

$$\begin{bmatrix} u_1 \mathbf{P}^3 - \mathbf{P}^1 \\ v_1 \mathbf{P}^3 - \mathbf{P}^2 \\ u_2 \mathbf{P}^3 - \mathbf{P}^1 \\ v_2 \mathbf{P}^3 - \mathbf{P}^2 \end{bmatrix} \mathbf{X} = \mathbf{0} \quad (\text{II. 19})$$

Di mana u_1 dan v_1 adalah koordinat gambar pada *view* 1 dan u_2 dan v_2 adalah koordinat gambar pada *view* 2. Anggap matriks \mathbf{A} adalah matriks kiri persamaan II.19.

$$\mathbf{A} = \begin{bmatrix} u_1 \mathbf{P}^3 - \mathbf{P}^1 \\ v_1 \mathbf{P}^3 - \mathbf{P}^2 \\ u_2 \mathbf{P}^3 - \mathbf{P}^1 \\ v_2 \mathbf{P}^3 - \mathbf{P}^2 \end{bmatrix} \quad (\text{II. 20})$$

Untuk triangulasi lebih dari 2 *view*, matriks \mathbf{A} tersebut dapat diperpanjang sebagai berikut,

$$\mathbf{A} = \begin{bmatrix} u_1 \mathbf{P}^3 - \mathbf{P}^1 \\ v_1 \mathbf{P}^3 - \mathbf{P}^2 \\ u_2 \mathbf{P}^3 - \mathbf{P}^1 \\ v_2 \mathbf{P}^3 - \mathbf{P}^2 \\ \vdots \\ u_i \mathbf{P}^3 - \mathbf{P}^1 \\ v_i \mathbf{P}^3 - \mathbf{P}^2 \end{bmatrix} \quad (\text{II. 21})$$

Sehingga didapati persamaan,

$$\mathbf{AX} = 0 \quad (\text{II. 22})$$

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

II.6 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 R_i dan T_i adalah matriks rotasi dari kamera i yang melakukan pemetaan tersebut, bisa didapatkan sebuah matriks rotasi R dan matriks translasi T sehingga

$$R_2 = RR_1 \quad (II.23)$$

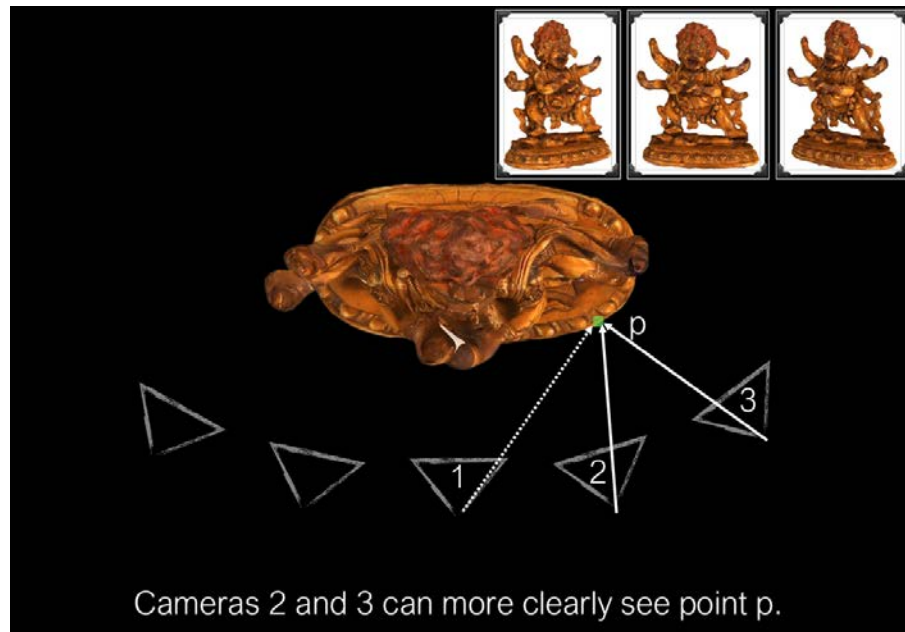
$$T_2 = RT_1 + T$$

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,

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ 1 \end{bmatrix} \quad (II.24)$$

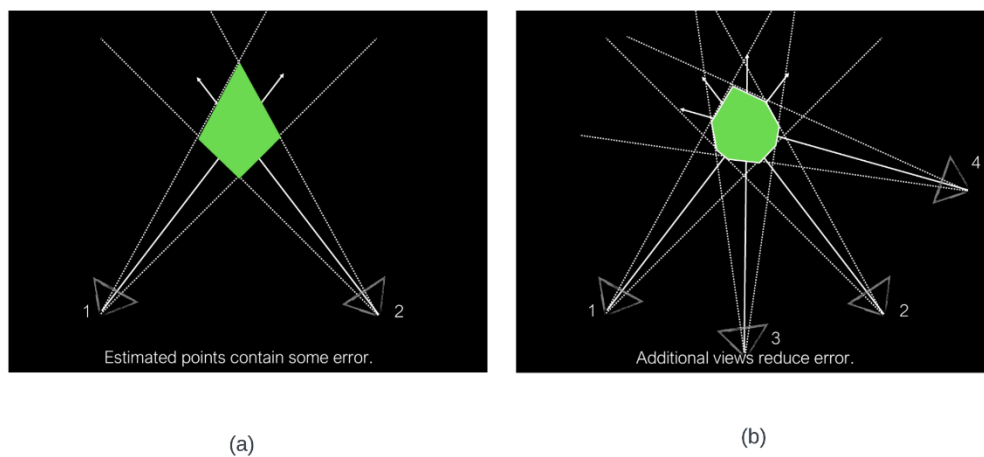
II.7 Multi-Kamera

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.9 berikut, *View* pada kamera 2 dan 3 dapat melihat suatu titik pada objek dengan lebih jelas dibandingkan kamera 1.



Gambar II.9. 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.10 berikut,



Gambar II.10. 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.10 ketika menggunakan banyak *view* akan menyebabkan daerah hijau tersebut menjadi lebih kecil.

II.8 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.

$$reprojection_error = d(\mathbf{x}, \mathbf{P}\mathbf{X}) \quad (\text{II. 25})$$

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

II.9 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,

$$\min_{\mathbf{P}_i, \mathbf{X}_j} \sum_i \sum_j d(\mathbf{P}_i \mathbf{X}_j, \mathbf{x}_{ij})^2 \quad (\text{II. 26})$$

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

II.10 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*, kalibrasi kamera stereo dengan fungsi *cv.stereoCalibrate*, dan penghilangan distorsi dengan fungsi *cv.undistort* yang mana fungsi *cv.calibrateCamera* dan fungsi *cv.stereoCalibrate* melakukan kalibrasi seperti yang dijelaskan pada subbab II.3 kalibrasi kamera tunggal dan subbab II.5 kalibrasi stereo dan fungsi *cv.undistort* melakukan perbaikan distorsi seperti yang dijelaskan pada subbab II.4.

BAB III

PENGEMBANGAN SISTEM MULTI-KAMERA

III.1 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.

III.2 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 distorsi dari distorsi 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. Teknik perbaikan distorsi dibahas lebih lanjut pada subbab II.4.

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.5, 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.5., 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.5 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.6 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.7.

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 optimasi 3D ini dapat dilihat pada subbab II.9. 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*.

III.3 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.

III.3.1 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 dipilih 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 $Z = 0$ 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)

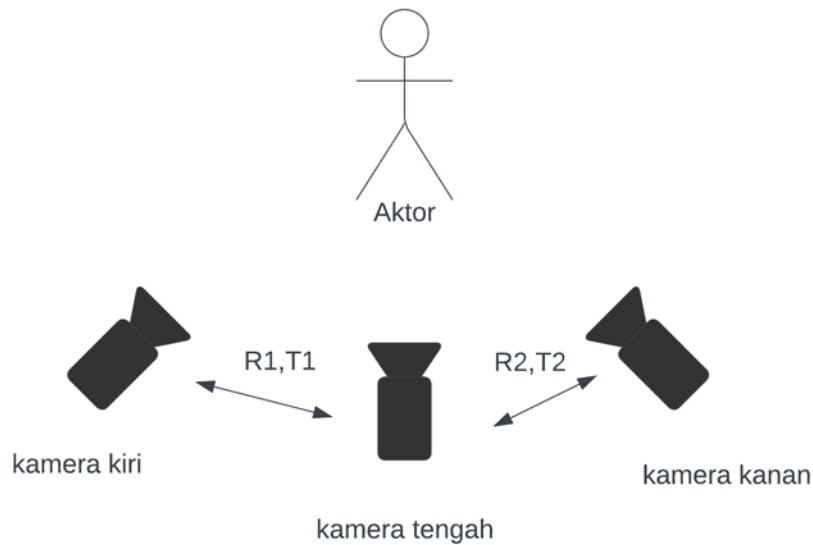
Untuk mengestimasi matriks intrinsik dan parameter distorsi kamera, akan digunakan fungsi *cv.calibrateCamera* dari pustaka OpenCV. Fungsi tersebut menerapkan metode kalibrasi Zhang seperti yang pada dasar teori subbab II.3 yang mana fungsi tersebut menerima data 2D dan data 3D tepi-tepi petak dan menghasilkan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

III.3.2 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. Cara kerja dari fungsi ini serupa seperti yang dijelaskan pada subbab II.4.

III.3.3 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,



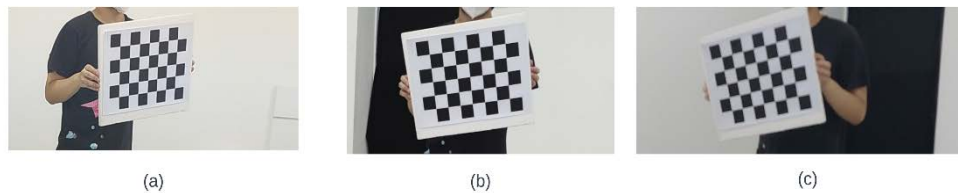
Gambar III.2. Gambaran urutan kamera yang digunakan

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. Selbihnya mengenai nilai rotasi dan translasi kamera yang digunakan untuk evaluasi sistem akan dibahas pada subbab III.3.5.

III.3.4 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.



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.

III.3.5 Estimasi 3D Penanda dengan Triangulasi

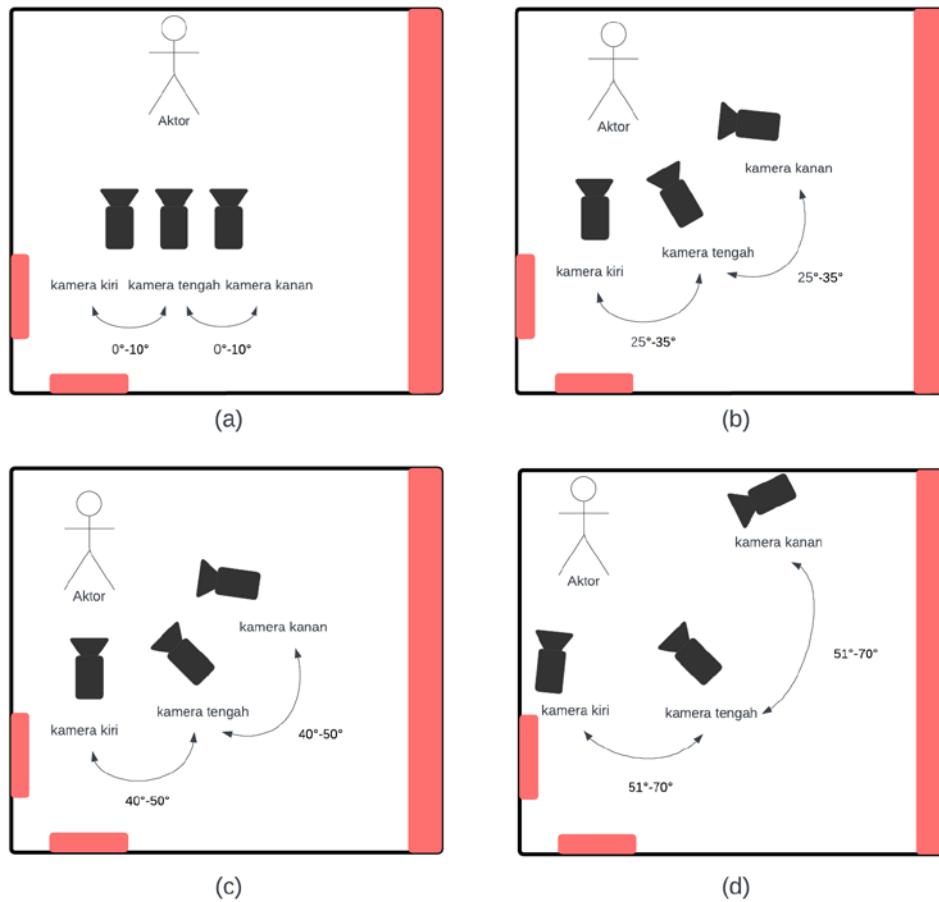
Pada bagian ini, akan ditetapkan beberapa konfigurasi rotasi dan translasi dari tiap pasangan kamera. Hal ini dilakukan untuk mencari konfigurasi rotasi dan translasi dari pasangan kamera yang menghasilkan estimasi 3D yang terbaik. Untuk mempermudah pengujian dan dikarenakan keterbatasan ruangan, ditetapkan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



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

2. 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.
3. 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 Gambar III.5 berikut



Gambar III.5. Konfigurasi peletakan kamera untuk setiap rentang rotasi. (a) rentang rotasi dengan kamera tengah 0° - 10° . (b) rentang rotasi dengan kamera tengah 25° - 35° . (c) rentang rotasi dengan kamera tengah 40° - 50° . (d) rentang rotasi dengan kamera tengah 51° - 70° .

4. 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.8.

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.5. Penerapan proses triangulasi ini dibuat dengan menggunakan pemahaman terhadap proses triangulasi yang ada pada 2 sumber berikut yaitu (Batpurev, 2021) dan (AmyTabb, 2021).

III.3.6 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).

BAB IV

EVALUASI SISTEM

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

IV.1 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.

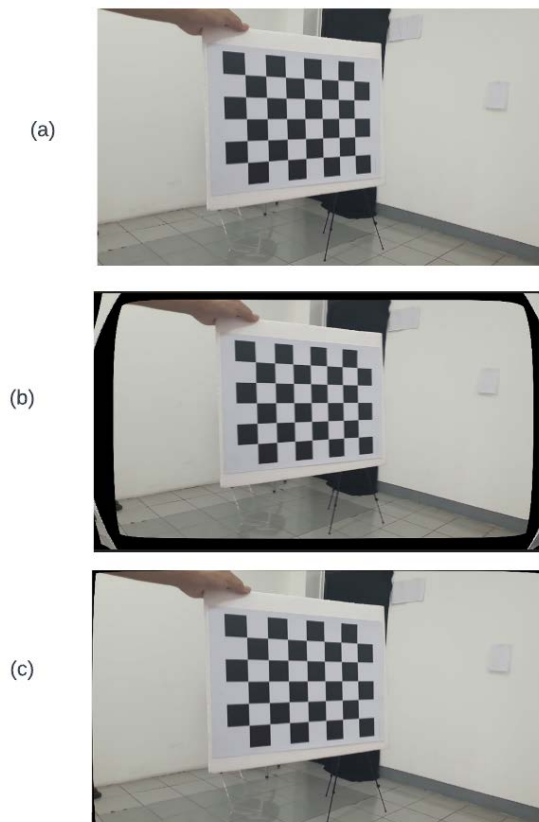


Gambar IV.1. Gambar (a) gambar asli sebelum dilakukan perbaikan distorsi. Gambar (b) gambar setelah dilakukan perbaikan distorsi.

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 perbaikan distorsi. Gambar (b) gambar setelah perbaikan distorsi yang masih salah. (c) gambar setelah perbaikan distorsi yang sudah diperbaiki. Daerah hitam, yang menunjukkan perbaikan distorsi pada gambar (c) lebih kecil dibandingkan gambar (b), yang mana gambar (c) lebih mirip dengan gambar (a) dibandingkan gambar (b).

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	$\begin{bmatrix} 1436.854 & 0 & 951.71 \\ 0 & 1437.605 & 548.9 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 6.37 \\ -46.576 \\ 0.0022 \\ -0.0011 \\ 138.793 \\ 6.389 \\ -47.004 \\ 139.1384 \end{bmatrix}$
Kamera tengah	$\begin{bmatrix} 1473.503 & 0 & 950.21 \\ 0 & 1471.862 & 545.757 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -2.194 \\ 79.834 \\ 0.0003 \\ 0.0013 \\ -115.973 \\ -2.485 \\ 8 \\ -116.852 \end{bmatrix}$
Kamera kanan	$\begin{bmatrix} 1346.381 & 0 & 983.342 \\ 0 & 1346.383 & 530.83 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -17.635 \\ -26.373 \\ 0.0036 \\ 0.0008 \\ 1353.113 \\ -17.941 \\ -18.649 \\ 1301.006 \end{bmatrix}$

Matriks intrinsik kamera disusun seperti yang terdapat pada persamaan IV. 7 untuk matriks intrinsik kamera dan persamaan IV. 13 untuk parameter distorsi.

IV.2 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.

Hasil rotasi dan translasi dari proses kalibrasi stereo untuk setiap rentang pengujian ada untuk proses selanjutnya ada pada Tabel IV.3 berikut.

Tabel IV.3. Tabel rotasi dan translasi kamera

No. Uji (Rentang Uji)	Kamera	Sumbu	Rotasi Kamera	Translasi Kamera
Uji rentang 1 (0°-10°)	Kiri	x	2.350	-8.004
		y	-5.857	71.802
		z	3.035	10.348
	Kanan	x	-2.046	-8.004
		y	3.396	71.802
		z	-0.184	10.348
Uji rentang 2 (25°-35°)	Kiri	x	0.138	142.416
		y	-25.507	2.622
		z	-3.863	24.636
	Kanan	x	-4.4	-8.004
		y	28.068	71.802
		z	2.922	10.348
Uji rentang 3 (40°-50°)	Kiri	x	-5.107	130.589
		y	-43.253	-2.923
		z	-7.118	33.548
	Kanan	x	-5.13	-152.713
		y	46.452	18.2
		z	5.7732	67.064
Uji rentang 4 (51°-70°)	Kiri	x	2.375	234.541
		y	-54.465	4.299
		z	-8.408	108.937
	Kanan	x	-8.004	-170.365
		y	71.802	-16.114
		z	10.348	76.514

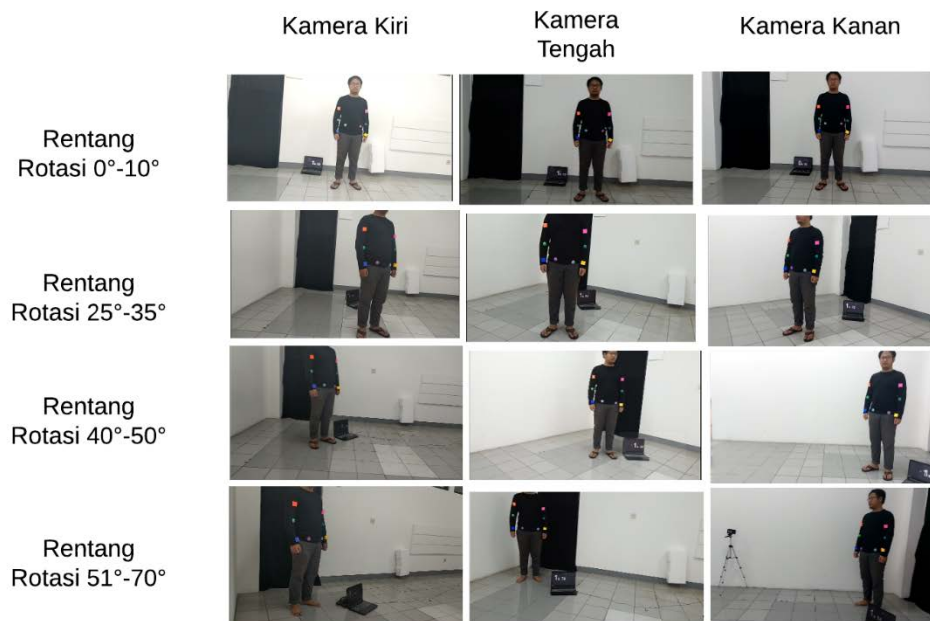
Perlu diperhatikan bahwa rotasi dan translasi kamera-kamera tersebut diukur dari besar perbedaan sudut dan perpindahan dari kamera utama yaitu kamera tengah. Besar perbedaan sudut dan perpindahan kamera tengah sendiri adalah 0 karena kamera tengah tersebut tidak memiliki perbedaan sudut atau perpindahan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

IV.3 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.

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.3. Gambaran aktor dari *view* dan rotasi yang berbeda-beda

Hasil *reprojection error* untuk tiap pengujian terdapat pada Tabel IV.4 di bawah ini

Tabel IV.4. Tabel rata-rata kesalahan proyeksi balik

No. Uji	Pasangan Kamera Triangulasi	<i>Reprojection Error (pixel)</i>	
		Triangulasi	Hasil Optimasi <i>Bundle Adjustment</i>
Uji rentang 1	Kiri-tengah	2.634	0.563
	Tengah-kanan	3.176	0.855
	Kiri-kanan	2.845	0.487
	3 kamera bersamaan	2.516	0.557
Uji rentang 2	Kiri-tengah	2.423	0.831
	Tengah-kanan	3.789	1.757
	Kiri-kanan	2.316	0.811
	3 kamera bersamaan	2.299	0.719
Uji rentang 3	Kiri-tengah	2.84	0.757
	Tengah-kanan	2.8	0.626
	Kiri-kanan	2.124	0.661
	3 kamera bersamaan	2.114	0.69
Uji rentang 4	Kiri-tengah	2.076	0.839
	Tengah-kanan	2.878	0.715
	Kiri-kanan	1.9	1.072
	3 kamera bersamaan	1.866	0.777

Dari Tabel IV.4 tersebut didapati beberapa temuan sebagai berikut,

1. Tidak ada pengaruh perubahan rotasi kamera terhadap *reprojection error*. Hal ini membuktikan kalibrasi matriks intrinsik dan ekstrinsik kamera yang dilakukan dan data 2D yang digunakan adalah benar.
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 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

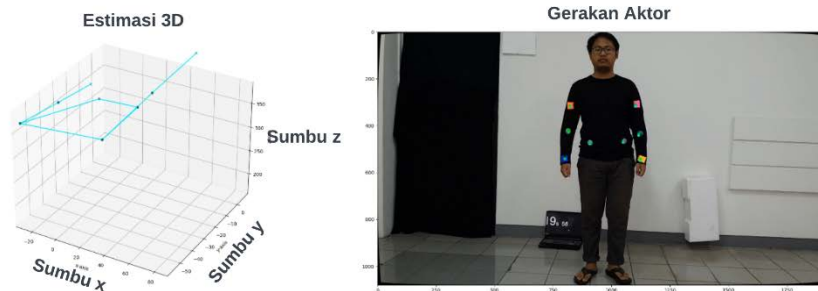
kesalahan proyeksi balik yang paling rendah. Hal ini dapat dikarenakan faktor estimasi awal untuk pasangan kamera triangulasi yang bersangkutan lebih baik dibandingkan estimasi awal triangulasi dengan 3 kamera sekaligus dalam memudahkan algoritma optimasi untuk menemukan titik yang optimal. Mengingat bahwa algoritma optimasi yang digunakan menggunakan prinsip turunan sehingga estimasi awal sangat berpengaruh untuk menemukan lokal minimum terbaik.

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 3 uji gerakan yaitu pose-A yang dinaikan dan diturunkan di samping badan aktor, bergerak ke kiri dan ke kanan, dan gerakan 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 perpindahan horizontal seluruh badan.. Gerakan ketiga 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.

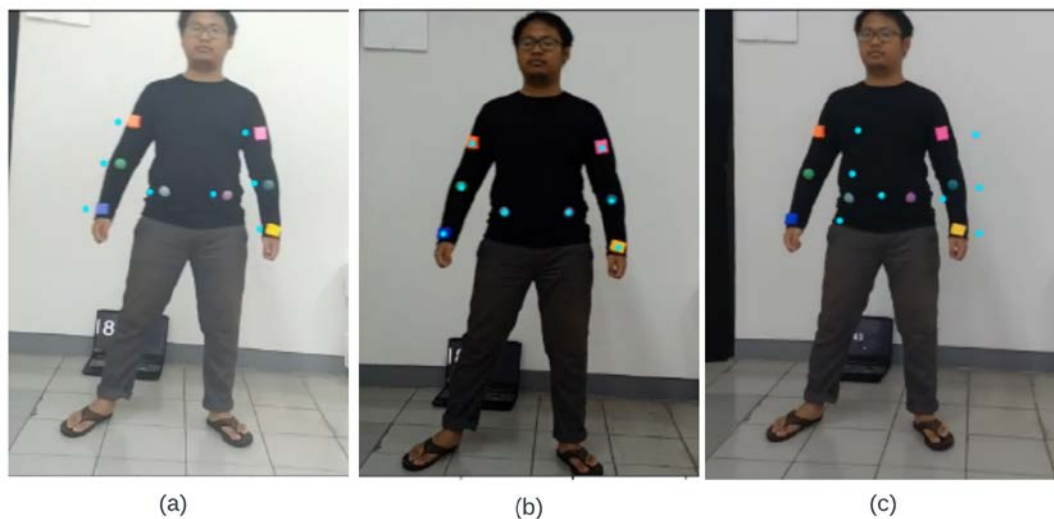
Dari percobaan yang telah dilakukan didapati beberapa temuan sebagai berikut,

1. Umumnya, seluruh gerakan dapat ditangkap dengan benar. Namun, tidak jarang didapati hasil struktur 3D dengan beberapa titik yang tidak sesuai atau salah dalam mengestimasi kedalaman. Sebagai contoh perhatikan Gambar IV.4 berikut.



Gambar IV.4. Visualisasi gerakan kiri-kanan. Struktur 3D “mengapung” mengapung ke atas. Sumbu x menyatakan arah kanan-kiri. Sumbu y menyatakan arah atas-bawah. Sumbu z menyatakan arah depan-belakang.

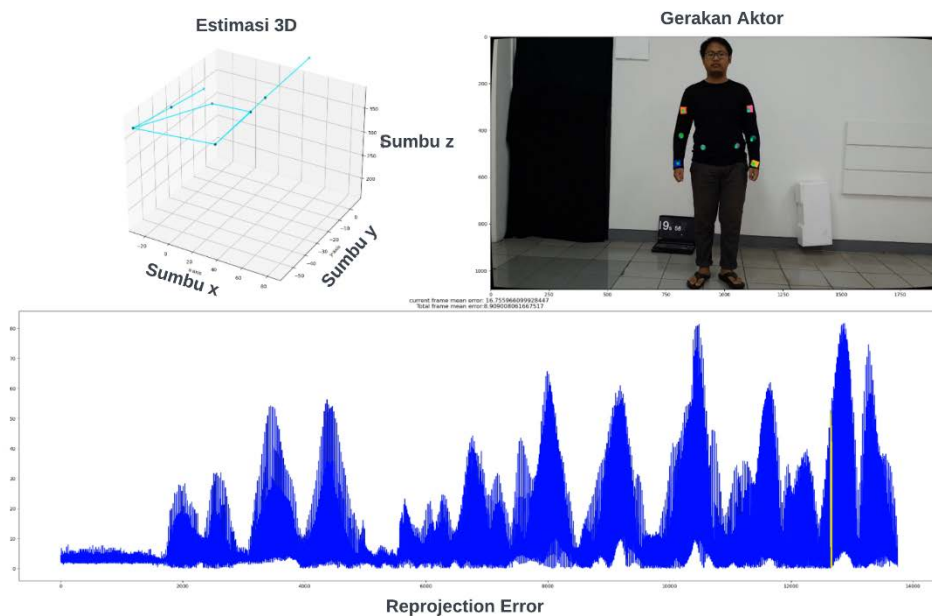
Pada Gambar IV.4. struktur 3D yang dihasilkan “mengapung” yang mana merupakan sebuah kesalahan dalam mengestimasi kedalaman. Setelah dianalisis penyebabnya, didapati kesalahan ini disebabkan karena *frame* antar kamera yang digunakan tidak sinkron seperti pada Gambar IV.5. berikut.



Gambar IV.5. (a), (b), (c) Gambar kamera kiri-tengah-kanan secara berturut-urut. Penanda yang terdeteksi ditandai dengan titik biru.

Akibat *frame* yang tidak sinkron ini, pasangan titik 2D keliru dan menimbulkan kesalahan estimasi kedalaman. Berdasarkan analisis ini pun, didapati bahwa dikarenakan ketiga kamera tersebut sering memiliki perbedaan *frame*, maka ditarik juga kesimpulan bahwa estimasi 3D dengan 3 kamera tidak selalu menghasilkan hasil yang terbaik dikarenakan 3 kamera tersebut mendeteksi titik yang berbeda-beda.

2. *Reprojection error* dapat dijadikan indikasi hasil estimasi 3D yang tidak tepat. Pada gambar IV.6. berikut,



Gambar IV.6. Dua gambar di bagian atas adalah gambar yang sama dengan Gambar IV.4, gambar bawah adalah visualisasi *reprojection error* untuk tiap *frame* dengan bagian yang diberi warna kuning adalah *reprojection error frame* saat ini .

ketika hasil 3D akan mengapung seperti pada gambar IV.4, *reprojection error*-nya pun meningkat seperti pada bagian yang diberi warna kuning pada *panel* bawah. Hal ini menandakan terdapat korelasi antara *reprojection error* dan hasil 3D yang dihasilkan. Hal ini dikarenakan

proyeksi balik dari hasil 3D yang salah tidak akan jatuh pada titik 2D-nya yang ada pada data.

BAB V

KESIMPULAN DAN SARAN

Pada bagian ini akan dibahas kesimpulan dari evaluasi kinerja sistem yang telah dibangun beserta saran terkait pengembangan sistem selanjutnya.

V.1 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 serta mengevaluasi estimasi 3D tersebut dan dengan demikian menjawab permasalahan yang telah didefinisikan.

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

1. *reprojection error* dapat dijadikan sebagai sebuah indikasi hasil estimasi 3D yang salah,
2. rotasi tidak terlalu mempengaruhi *reprojection error* dan dengan demikian tidak terlalu berpengaruh pada hasil estimasi 3D,
3. estimasi 3D dengan 3 kamera sekaligus menghasilkan *reprojection error* yang paling rendah dan hasil 3D yang terbaik jika semua gambar yang diambil oleh masing-masing secara sinkron diambil pada saat yang bersamaan.

V.2 Saran

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

1. melakukan perbandingan lebih lanjut terhadap pengaruh penambahan kamera terhadap hasil estimasi 3D,
2. melakukan percobaan dengan gerakan yang lebih variatif dan melihat pengaruhnya terhadap hasil 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>
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.
- Bradski, G., & Kaehler, A. (2008). *Learning OpenCV*. O'Reilly Media, Inc.
- Brown, D. C. (1966). Decentering Distortion of Lenses.
- 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.
- 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
- Snavely, N., & Li, Z. (2018). Diambil kembali dari Multi-view stereo:
https://www.cs.cornell.edu/courses/cs5670/2018sp/lectures/lec16_mvsv.pdf
- Steward, J. (2021, Agustus 6). *Camera Modeling: Exploring Distortion and Distortion Models, Part I*. Diambil kembali dari Tangram Vision:

<https://www.tangramvision.com/blog/camera-modeling-exploring-distortion-and-distortion-models-part-i#tangential-de-centering-distortions>

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