

KAJIAN TENTANG REKONSTRUKSI OBJEK 3D MENGUNAKAN STEREO VISION DENGAN METODE HARRIS INTEREST POINT DAN RANSAC

Firbert Oktariko ¹⁾ Lina ²⁾

¹⁾Mahasiswa Program Studi Teknik Informatika
Fakultas Teknologi Informasi Universitas Tarumanagara, Jakarta 11440 Indonesia
email : firbert96@gmail.com

²⁾Dosen Program Studi Teknik Informatika
Fakultas Teknologi Informasi Universitas Tarumanagara, Jakarta 11440 Indonesia
email : lina@untar.ac.id

ABSTRAK

Rekonstruksi 3D merupakan proses untuk memperoleh kembali informasi objek 3D yang ada di dunia nyata dan menyusunnya kembali ke dalam titik-titik pada komputer sehingga komputer dapat mengolah serta menampilkan informasi yang mirip bahkan sama terhadap objek 3D di dunia nyata.

Proses ini diawali dengan mengambil 2 input dari citra kamera yang berbeda tetapi memiliki spesifikasi sama. Proses selanjutnya adalah proses segmentasi otomatis menggunakan Harris Interest Point, Proses Normalized Cross Correlation, dan Random Sample Consensus. Adapun perbandingannya dengan metode segmentasi semi otomatis dengan model warna RGB dan HSV dengan nilai parameter yang telah ditentukan. Dari hasil pasangan titik yang terpilih nantinya akan diambil 8 pasangan titik untuk direkonstruksi 3D.

Keberhasilan rata-rata hasil rekonstruksi objek 3D yang didapatkan pengujian dengan segmentasi otomatis sebesar 0% dari 40 kali pengujian sedangkan untuk pengujian dengan segmentasi semi-otomatis sebesar 37,5% dari 8 kali pengujian.

Kata Kunci :

Harris Interest Point, Normalized Cross Correlation, Random Sample Consensus, Rekonstruksi 3D, Segmentasi

1. Pendahuluan

Rekonstruksi 3D berbasis geometri merupakan rekonstruksi 3D bertujuan untuk mendapatkan nilai jarak, keliling, luas, maupun volume dari hasil olahan nilai vektor setiap masing-masing titik pada piksel citra yang ditangkap.

Rekonstruksi 3 dimensi berbasis geometri sendiri memiliki metode secara garis dibagi menjadi dua, yaitu metode aktif melakukan proses rekonstruksi objek 3D dengan laser scanner atau cahaya terstruktur dan metode pasif yang

menggunakan sebuah atau beberapa citra yang ditangkap.

2. Dasar Teori

Pada sistem yang dirancang terdapat dua kamera memiliki spesifikasi yang sama dan pengaturan titik fokus yang sama. Dua kamera memiliki jarak yang telah diketahui dan dalam posisi sejajar. Kemudian kamera pertama dan kamera kedua mengambil gambar sehingga terdapat dua citra berbeda yang ditangkap dengan posisi objek yang tidak berubah. Dua citra tersebut akan dicari titik-titik serupa dari citra kiri (citra yang ditangkap oleh kamera kiri) dengan citra kanan (citra kamera yang ditangkap oleh kamera kanan) melalui proses segmentasi. Proses segmentasi otomatis dimulai dengan metode Harris Interest Point, Normalized Cross-Correlation (NCC), Singular Value Decomposition (SVD), Direct Linier Transform (DLT), Normalized Direct Linier Transform, dan RANSAC. Adapun pembandingannya dengan metode segmentasi semi otomatis dengan model warna RGB dan HSV dengan nilai parameter yang telah ditentukan. Dari titik-titik tersebut akan dicari nilai koordinat *pixel*-nya (x,y) kemudian akan dicari sebuah kedalaman (*depth*) atau koordinat Z sehingga membentuk koordinat proyeksi (X,Y,Z) menggunakan Stereo Vision. Titik-titik proyeksi tersebut akan dihubungkan dan membentuk objek 3D yang diamati.

Adapun metode yang dipakai sebagai berikut :

1. Harris Interest Point

Perhitungan pada metode ini bertujuan untuk mencari ciri khusus dari sebuah gambar berupa titik bertemunya sudut pada suatu gambar. Persamaan yang digunakan pada perhitungan Harris Interest Point diawali dengan mengkonvolusikan derivative

of Gaussian dengan citra yang digunakan. Adapun masking yang digunakan sebagai berikut sebagai berikut:

$$(Ix) = \frac{\partial G(x, y, \sigma)}{\partial x} = -\left(\frac{x}{2\pi\sigma^4}\right) e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (1)$$

$$(Iy) = \frac{\partial G(x, y, \sigma)}{\partial y} = -\left(\frac{y}{2\pi\sigma^4}\right) e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (1)$$

$$(IxIy) = -\left(\frac{x * y}{2\pi\sigma^4}\right) e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (1)$$

Setelah mencari gradien magnitudenya, barulah mencari *keypoint* menggunakan persamaan sebagai berikut:

$$harris =$$

$$g(Ix)^2 g(Iy)^2 - \{g(IxIy)^2 - a[g(Ix)^2 + g(Iy)^2]^2\} \quad (2)$$

Dilakukannya perhitungan non-maxima suppression untuk 2 buah matriks skeleton dari masing-masing gambar yang bernilai 0 dan 255.

2. Normalized Cross Correlation (NCC)

Normalized cross-correlation adalah suatu metode *template matching* yang lebih spesifik, metode ini dapat digunakan untuk citra *grayscale* atau citra RGB maupun *binary pixel arrays*. Tujuan perhitungan pada metode ini adalah mencari untuk mencari kesamaan korespondensi paling baik antara citra target terhadap citra pembandingnya dengan menghitung nilai pembanding tiap piksel yang telah menjadi targetnya. Adapun persamaan yang digunakan sebagai berikut:

$$NCC =$$

$$\frac{\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (w(x, y) - \bar{w})(f(x + i, y + i) - \bar{f}(i, j))}{\sqrt{\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (w(x, y) - \bar{w})^2} \sqrt{\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (f(x + i, y + i) - \bar{f}(i, j))^2}} \quad (2)$$

Pada metode ini akan dicari nilai terbesar NCC pada masing-masing pasangan sehingga didapatlah pasangan yang ideal untuk dijadikan koresponden titik.

3. Direct Linier Transform (DLT)

Direct Linear Transform (DLT) adalah sebuah algoritma yang digunakan untuk mencari *homography* matriks (matriks H) yang didapat dari koresponden antar titik pada ruang lingkup 2 dimensi. Setiap titik yang berhubungan memiliki 2 persamaan, maka hubungan antara 4 titik akan mencukupi untuk menyelesaikan permasalahan 8 *degree of freedom* dari transformasi proyeksi matriks H.

$$A_i =$$

$$\begin{bmatrix} -x_1 & -y_1 & -1 & 0 & 0 & 0 & u_1x_1 & u_1y_1 & u_1 \\ 0 & 0 & 0 & -x_1 & -y_1 & -1 & v_1x_1 & v_1y_1 & v_1 \\ -x_2 & -y_2 & -1 & 0 & 0 & 0 & u_2x_2 & u_2y_2 & u_2 \\ 0 & 0 & 0 & -x_2 & -y_2 & -1 & v_2x_2 & v_2y_2 & v_2 \\ -x_3 & -y_3 & -1 & 0 & 0 & 0 & u_3x_3 & u_3y_3 & u_3 \\ 0 & 0 & 0 & -x_3 & -y_3 & -1 & v_3x_3 & v_3y_3 & v_3 \\ -x_4 & -y_4 & -1 & 0 & 0 & 0 & u_4x_4 & u_4y_4 & u_4 \\ 0 & 0 & 0 & -x_4 & -y_4 & -1 & v_4x_4 & v_4y_4 & v_4 \end{bmatrix} \quad (2)$$

Setelah itu sebuah matriks H dapat dihitung menggunakan metode SVD, dimana matriks A dapat diasumsikan sebagai berikut:

$$A = USV^T \quad (3)$$

Matriks H merupakan hasil dari *reshaping* nilai pada kolom terakhir matrix V berukuran 1 x 9 menjadi ukuran 3 x 3.

4. Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) merupakan metode yang biasa digunakan untuk melakukan reduksi pada sebuah matriks. SVD merupakan suatu metode analisis matriks yang mempresentasikan dimensi kecil dari sebuah matriks yang awalnya memiliki dimensi yang besar.

Dari persamaan yang dimiliki metode SVD, diasumsikan sebuah matriks X memiliki ukuran m x n sehingga dapat dilihat sebagai berikut:

1. Matriks U, Merupakan matriks berukuran m x n yang memiliki kolom yang ortonormal. Setiap kolom merupakan sebuah vektor dimana nilai perkalian titik (*dot product*) antara dua kolom akan selalu menghasilkan nilai 0, disebut sebagai *left singular vectors*.
2. Matriks V merupakan matriks berukuran m x n yang memiliki kolom yang ortonormal. Matriks V digunakan dalam bentuk yang sudah *di-transpose* untuk dapat menghasilkan matriks yang ortonormal. Matriks ini disebut sebagai *right singular vectors*.
3. Matriks S merupakan matriks diagonal yang berisi *singular value* dari matriks X yang berukuran n x n, dimana semua elemen selain diagonal utamanya bernilai 0.

5. Homography

Matriks transformasi dapat tergabung atas beberapa operasi geometri seperti translasi dengan rotasi, rotasi dengan skala dan lain lain. Saat melakukan proyeksi *homography*, terjadinya beberapa operasi geometri meliputi 2 parameter translasi, 2 parameter rotasi, 2 parameter skala dan 2 parameter *line at infinity* yang menjadikan proyeksi *homography* memiliki 8 *degree of freedom*.

Persamaan geometri pada proyektif *homography* adalah sebagai berikut :

$$\begin{bmatrix} \bar{u} \\ \bar{v} \\ \bar{w} \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (4)$$

6. *Normalized Cross Correlation (NDLT)*

Menurut Hartley dan Zisserman, hasil dari DLT tergantung dari keaslian dan skala dari koordinat. Dengan kata lain, algoritma ini sangat sensitif dengan hubungan antar koordinat dikarenakan DLT tidak pernah dipakai dengan koresponden antar titik lebih dari 4 namun bekerja baik ketika hubungan antar titik memiliki skala yang sama. Alasan tidak konsistennya hasil tersebut ada hubungannya dengan bagaimana cara metode DLT menggunakan SVD pada matriks A untuk mendapatkan solusi dari persamaan $A_h = 0$. Hartley dan Zisserman mengajukan sebuah normalisasi data untuk memastikan hasil konvergen yang benar. Adapun T_{norm} yang dipakai adalah sebagai berikut:

$$T_{norm} = \begin{bmatrix} s & 0 & t_x \\ 0 & s & t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

dengan

$$s = \frac{\sqrt{2}}{\frac{1}{n} \sum_i \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2}} \quad (5)$$

$$t_x = -s * \bar{x} \quad (5)$$

$$t_y = -s * \bar{y} \quad (5)$$

Setelah dilakukan perhitungan matriks T_{norm} dan T'_{norm} , maka algoritma normalized DLT sebagai berikut:

1. Normalisasi setiap pasangan titik titik dengan persamaan :

$$\tilde{x} = T_{norm} * X; \tilde{x}' = T'_{norm} * X' \quad (6)$$

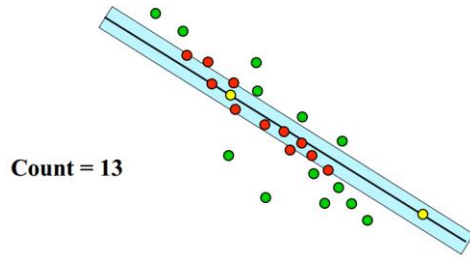
2. Lakukan algoritma DLT kepada $\tilde{x} \rightarrow \tilde{x}'$ untuk mencari matriks H.

3. Lakukan denormalisasi pada matriks H yang didapat dari DLT dengan persamaan:

$$H = T'^{-1}_{norm} * \tilde{H} * T_{norm} \quad (6)$$

7. *Random Sample Consensus (RANSAC)*

Random Sample Consensus (RANSAC) merupakan sebuah metode iterasi untuk memperkirakan parameter dari model matematika pada sebuah *dataset* yang mengandung *outliers* dimana *outliers* tersebut tidak terpengaruh dengan nilai estimasi. Oleh karena itu, RANSAC juga dapat disebut sebagai *outliers detection method*. Pada metode ini, koresponden antar titik yang didapat sebelumnya akan diproses pada metode ini.



Gambar 1 Contoh perhitungan inliers berdasarkan nilai threshold

Pada umumnya nilai P diasumsikan dengan nilai 0.99; Sehingga untuk mencari N atau jumlah iterasi yang diperlukan, didapatlah persamaan berikut¹.

$$N = \frac{\log(1-p)}{\log(1-(1-e)^s)} \quad (7)$$

$$p = \frac{\text{outliers}}{\text{jumlah titik pada dataset}} \quad (7)$$

Keterangan :

e = Kemungkinan *outliers*

S = Banyaknya pemilihan acak s inliers

P = Kemungkinan untuk mendapatkan inliers terbaik

N = Banyaknya iterasi yang harus dilakukan

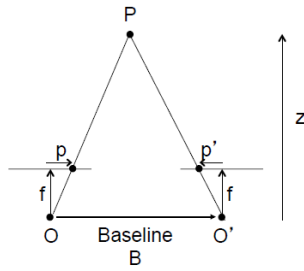
8. *RANSAC untuk estimasi Homography*

Saat mencari matriks H untuk projektif *homography*, diperlukan paling sedikit 4 koresponden titik. Metode RANSAC dapat memecahkan masalah koresponden titik yang termasuk *inliers* dan *outliers*.

serta memilih 4 pasangan titik terbaik untuk dijadikan matriks H pada transformasi projektif. Dengan menerapkan algoritma *singular value decomposition* dan *normalized direct linear transform* serta mencari iterasi optimal menggunakan RANSAC pada sebuah dataset.

9. *Stereo Vision*

Stereo vision adalah proses rekonstruksi model 3D melalui pengolahan citra 2D. Melihat objek atau citra dari dua lebih posisi yang berbeda secara bersamaan memungkinkan untuk membuat kesimpulan tentang struktur 3D, yaitu dengan membandingkan titik-titik yang terkait pada citra. Sistem visual manusia dan beberapa hewan lain menggunakan ini, dan sangat penting dalam upaya untuk mengembangkan sistem visual komputer praktis. *Stereo Vision* menggunakan dua kamera untuk mengamati lingkungan, menemukan titik yang sama di setiap citra, dan mengukur kedalaman pada titik itu dengan triangulasi, yaitu dengan melihat garis – garis yang berpotongan dari setiap kamera terhadap objek.



Gambar 2 Stereo Vision

Dari Gambar 2. dapat dirumuskan bahwa :

$$Z = \frac{B \cdot f}{d} \tag{8}$$

Keterangan :

- Z = kedalaman pada objek yang diamati
- B = *baseline* (jarak antar kamera)
- F = titik fokus
- d = *disparity*

Disparity akan dirumuskan sebagai berikut:

$$d = x_l - x_r \tag{8}$$

Keterangan :

- d = disparity
- x_l = koordinat x pada titik yang bersesuaian pada citra kiri atau citra yang ditangkap oleh kamera kiri
- x_r = koordinat x pada titik yang bersesuaian pada citra kanan atau citra yang ditangkap oleh kamera kanan

Perhitungan titik koordinat objek 3D menggunakan titik kedalaman (Z) sehingga dapat dirumuskan sebagai berikut:

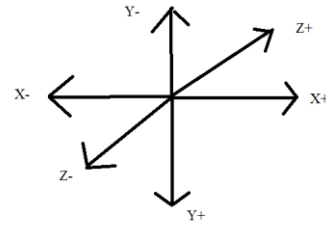
$$X = \frac{x_l \cdot Z}{f} \text{ dan } Y = \frac{y_l \cdot Z}{f} \tag{8}$$

Keterangan :

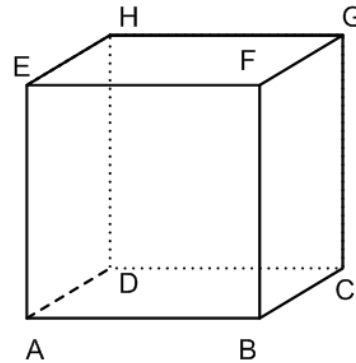
- X = Koordinat X pada titik-titik pada objek 3D
- Y = Koordinat y pada titik-titik pada objek 3D
- x_l = Koordinat X pada titik-titik yang bersesuaian pada citra kiri yang ditangkap kamera kiri
- y_l = Koordinat Y pada titik-titik yang bersesuaian pada citra kiri yang ditangkap kamera kiri

10. Teknik Rekonstruksi Objek 3D

Pada rancangan ini dibatasi objek yang diamati hanya objek yang berbentuk balok maupun kubus sehingga *rule* rekonstruksi objek ini hanya memerlukan 8 titik sudut sehingga dihubungkan menjadi 12 garis yang berbeda. Rekonstruksi ini juga menskalakan ukuran asli objek yang diamati (dengan satuan mm) sama dengan 1 *pixel* gambar. Adapun gambar yang menjadi referensi penamaan pada masing-masing titik sudut sebagai berikut:



Gambar 3 Kaidah tangan kanan Degree of Freedom



Gambar 4 Kubus dengan masing-masing titik sudut yang diberikan nama

Dari referensi Gambar 4, rekonstruksi objek 3D akan dibuat dengan persamaan-persamaan berikut seperti berikut:

1. Buatlah 10 bin koordinat x dengan jarak 32 pada masing-masing citra. Seperti bin 0 merupakan bin berjarak 0-31, bin 1 merupakan bin berjarak 32-63, dst.
2. Masukkan pasangan titik yang terkorresponden tersebut ke dalam masing-masing bin citra.
3. Cek 2 bin yang terbanyak pada masing-masing citra
4. Buatlah bin x_1 dan bin x_2 pada masing-masing citra dengan kondisi jarak bin $x_1 <$ jarak bin x_2 pada masing-masing citra. Maksud bin x_1 dan bin x_2 disini adalah kumpulan titik yang berkoresponden yang menyerupai batang vertikal yang nantinya akan dicari titik sudutnya. Bin x_1 akan mewakili titik A, titik D, titik E, dan titik H. Bin x_2 akan mewakili titik B, titik C, titik G, dan titik H.
5. Masukkan list pasangan titik yang termasuk pada masing-masing bin untuk dicari pasangan titik yang bentuk . Ada list x_1 dan list x_2 .
6. Akan dipilih 2 hasil maksimal dari perkalian koordinat y kiri dengan y kanan dari list x_1 untuk menjadi nilai A dan nilai D. Kemudian dipilih 2 hasil minimum dari list x_1 untuk menjadi nilai E dan nilai H. Kemudian dipilih 2 hasil maximum dari list x_2 untuk menjadi nilai B dan nilai C. Kemudian dipilih 2 hasil minimum dari list x_2 untuk menjadi nilai F dan nilai G. Untuk ilustrasi nya dapat dilihat pada
7. Bentuk garis AB yang berasal dari titik A dan titik B. Kemudian Bentuk garis BC yang berasal

dari titik B dan titik C. Kemudian Bentuk garis CD yang berasal dari titik C dan titik D. Kemudian Bentuk garis AD yang berasal dari titik A dan titik D. Kemudian Bentuk garis EF yang berasal dari titik E dan titik F. Kemudian Bentuk garis FG yang berasal dari titik F dan titik G. Kemudian Bentuk garis GH yang berasal dari titik G dan titik H. Kemudian Bentuk garis EH yang berasal dari titik E dan titik H. Kemudian Bentuk garis AE yang berasal dari titik A dan titik E. Kemudian Bentuk garis BF yang berasal dari titik B dan titik F. Kemudian Bentuk garis CG yang berasal dari titik C dan titik G. Kemudian Bentuk garis DH yang berasal dari titik D dan titik H.

8. Jadilah rekonstruksi 3D berbentuk kubus atau balok.

3. Hasil Percobaan

Hasil pengujian rekonstruksi 3D ini sangatlah dipengaruhi pada hasil segmentasi, perhitungan koordinat (x,y,z), dan penghubungan garis dengan masing-masing pasangan titik. Oleh sebab itulah pengujian ini seringkali bermasalah dengan segmentasi dikarenakan ketidak sesuai parameter diakibatkan nilai *threshold* yang tidak sesuai sehingga tidak didapatkannya pasangan titik yang sesuai. Tidak hanya segmentasi saja, penghubungan garis pada masing-masing titik juga bermasalah karena ketidak-cocokan titik fokus yang sedang digunakan oleh kamera, keakuratan nilai jarak antar kamera, dan peletakan posisi kamera sehingga menyebabkan kesalahan pengambilan citra dua untuk diolah. Hal ini dapat mengganggu proses pembentukan sebuah objek 3D yang baik yang menyerupai kubus atau balok. Perhitungan keberhasilan dihitung dari banyaknya titik sudut dari di koordinat x,y,x yang terbentuk.

4. Kesimpulan

Kesimpulan pembuatan aplikasi rekonstruksi 3D ini adalah sebagai berikut:

1. Hasil titik sudut yang didapat sangat bergantung pada hasil segmentasi otomatis dan pengaturan nilai masking beserta thresholdnya. Sehingga sulit sekali menyesuaikan nilai threshold dan nilai masking yang ada untuk mendapatkan rekonstruksi yang diinginkan
2. Keberhasilan rata-rata titik sudut yang didapat dari hasil percobaan sebesar 57,1875% dari 40 kali pengujian untuk metode otomatis dan 100% dari 8 kali pengujian untuk metode semi-otomatis
3. Keberhasilan rata-rata hasil rekonstruksi objek 3D yang didapatkan pengujian dengan segmentasi otomatis sebesar 0% dari 40 kali pengujian sedangkan untuk pengujian dengan segmentasi semi-otomatis sebesar 37,5% dari 8 kali pengujian.

4. Metode semi-otomatis lebih mudah diterapkan tetapi penentuan titiknya masih manual (dengan melihat warna pada titik sudut yang dimaksud). Berbeda dengan dengan penentuan metode otomatis yang benar-benar mengandalkan nilai parameter yang ada sehingga sulit 8 titik sudut dengan baik.
5. Walaupun beberapa hasil percobaan didapat juga 8 titik sudut atau keberhasilan sebesar 100%, bangun ruang yang terbentuk banyak yang tidak menyerupai kubus atau balok. Hal ini disebabkan tidak tepatnya titik sudut yang terpilih. Contohnya tertukarnya titik a dengan titik d sehingga menyebabkan membuat garis AE dan DH menyilang.

REFERENSI

- [1] Lindeberg , T. "Scale selection properties of generalized scale space interest point detector". Jurnal of mathematical imaging and vision, Vol. 46, Nomor 2, Stockholm: School of Computer Science and Communication KTH Royal Institute of Technology, September 2013.
- [2] Haryono, Amarta Dimas. AUTOMATIC PANORAMIC IMAGE PADA CITRA PERSPEKTIF MENGGUNAKAN METODE HARRIS INTEREST POINT DAN RANSAC. Jakarta: Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara (Skripsi tidak dipublikasikan), 2018
- [3] Nagaria, Geraldi Lukito. Klasifikasi Lahan Mangrove Menggunakan Metode Singular Value Decomposition. Jakarta: Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara (Skripsi tidak dipublikasikan), 2016.
- [4] Csail. Advance in Computer Vision : Homography. <http://6.869.csail.mit.edu/fa12/lectures/lecture13ransac/lecture13ransac.pdf>, 2017.
- [5] Zhang, Rong. Automatic Computational of a homography by RANSAC Algorithm. https://engineering.purdue.edu/kak/computervision/ECE661.08/solution/hw4_s1.pdf, 2017.
- [6] Opsahl, Thomas. Estimating homographies from feature correspondences. http://www.uio.no/~lecture_4_3-estimating-homographies.pdf, 2017.
- [7] Collins, Robert. Robust Estimation RANSAC. <http://www.cse.psu.edu/~rtc12/CSE486, 1 Mei 2017>.
- [8] Li, Fei-Fei. Lecture 9 & 10 : Stereo Vision. http://vision.stanford.edu/teaching/cs131_fall1415/lectures/lecture9_10_stereo_cs131.pdf, 2017.

Firbert Oktariko, mahasiswa Universitas Tarumanagara, Jakarta tahun 2014 sampai 2018. Program studi Teknik Informatika.

Lina, memperoleh gelar S.T dari Universitas Tarumangara, Indonesia pada tahun 2001, dan mendapatkan gelar M.Kom dari Universitas Indonesia, Indonesia pada tahun 2004. Pada tahun 2009 mendapatkan

gelar Ph.D. dari Nagoya University, Jepang. Saat ini sebagai pengajar program studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara, Indones