

## REKONTRUKSI MODEL 3D DARI BANYAK GAMBAR MENGGUNAKAN ALGORITMA STRUCTURE FROM MOTION (SFM) DAN MULTI VIEW STEREO (MVS) BERBASIS COMPUTER VISION

Moch. Kholil<sup>1</sup>, Ismanto<sup>2</sup>, M. Nur Fu'ad<sup>2</sup>

<sup>1,2,3</sup>Akademi Komunitas Negeri Putra Sang Fajar Blitar  
Jl. dr. Sutomo No. 29, Kota Blitar, Indonesia

<sup>2</sup>Departemen, Fakultas, Institusi  
Alamat, Kota, Negara

e-mail: [moch.kholil89@gmail.com](mailto:moch.kholil89@gmail.com)<sup>1</sup>, [ism.ismanto@gmail.com](mailto:ism.ismanto@gmail.com)<sup>2</sup>, [nurfuadmail@gmail.com](mailto:nurfuadmail@gmail.com)<sup>2</sup>

Received : August, 2020

Accepted : October, 2020

Published : October, 2020

### Abstract

*With the development of the field of Information and Computer Technology (ICT), three-dimensional technology (3D) is also growing rapidly. Currently, the need to visualize 3D objects is widely used in animation and graphics applications, architecture, education, cultural recognition and virtual reality. 3D modeling of historical buildings has become a concern in recent years. 3D reconstruction is a documentation effort for reconstruction or restoration if the building is destroyed. By using a 3D model reconstruction approach based on multiple images using the Structure From Motion (SFM) and Multi View Stereo (MVS) algorithm, it is hoped that the 3D modeling results can be used as an effort to preserve 3D objects in the cultural heritage area of Penataran Temple. This research was conducted by taking an object in the form of photos as many as 61 pictures in the area of the Blitar Penataran Temple. The resulting photos are reconstructed into a 3D model using the Structure From Motion algorithm in the meshroom. In this study, a test was carried out on the original image with the compressed image for reconstruction to be compared to the 3D reconstruction process from the two input data. From 61 images processed using the Structure Form Motion algorithm, 33 camera pose and 3D point data were obtained, both original and compressed images. For the number of iterations the compressed image is 1.4% less than the original image and takes 43.53% faster than the original image.*

**Keywords:** 3D Reconstruction, SFM, MVS, Meshroom, Penataran Temple

### Abstrak

*Semakin berkembangnya bidang Teknologi Informasi dan Komputer (TIK), teknologi tiga dimensi (3D) juga ikut berkembang pesat. Saat ini, kebutuhan untuk memvisualisasikan obyek 3D banyak digunakan dalam aplikasi animasi dan grafis, arsitektur, pendidikan, pengenalan budaya dan Virtual Reality. Pemodelan 3D dari bangunan bersejarah sudah menjadi perhatian dalam beberapa tahun terakhir. Rekonstruksi 3D merupakan upaya dokumentasi untuk rekonstruksi atau restorasi jika bangunan tersebut hancur. Dengan menggunakan pendekatan rekonstruksi model 3D berdasarkan banyak gambar menggunakan algoritma Structure from Motion (SfM) dan Multi View Stereo (MVS) diharapkan hasil pemodelan 3D ini bisa dimanfaatkan sebagai upaya melestarikan obyek 3D di kawasan cagar budaya Candi Penataran. Penelitian ini dilakukan dengan cara mengambil obyek berupa foto sebanyak 61 gambar pada kawasan Candi Penataran Blitar. Hasil foto yang diperoleh direkonstruksi menjadi model 3D menggunakan algoritma Structure From Motion pada meshroom. Pada penelitian ini dilakukan uji coba terhadap gambar asli dengan gambar yang telah dikompres untuk dilakukan rekonstruksi untuk*

dibandingkan proses rekonstruksi 3D dari kedua data input tersebut. Dari 61 gambar dengan diproses menggunakan algoritma Structure Form Motion diperoleh data pose kamera dan titik 3D disempurnakan sebanyak 33 baik gambar asli maupun gambar kompres. Untuk jumlah iterasi gambar kompres 1.4 % lebih sedikit dari pada gambar asli dan membutuhkan waktu 43.53% lebih cepat dari gambar asli.

**Kata Kunci:** Rekonstruksi 3D, SFM, MVS, Meshroom, Candi Penataran

## 1. PENDAHULUAN

Undang undang Republik Indonesia Nomor 11 Tahun 2010 disebutkan pula bahwa yang disebut dengan pelestarian adalah upaya dinamis untuk mempertahankan keberadaan cagar budaya dan nilainya dengan cara melindungi, mengembangkan, dan memanfaatkannya.

Semakin berkembangnya bidang Teknologi Informasi dan Komputer (TIK), teknologi tiga dimensi (3D) juga ikut berkembang pesat. Saat ini, kebutuhan untuk memvisualisasikan obyek 3D banyak digunakan dalam aplikasi animasi dan grafis, arsitektur, pendidikan, pengenalan budaya dan Virtual Reality[1]. Pemodelan 3D dari bangunan bersejarah sudah menjadi perhatian dalam beberapa tahun terakhir. Rekonstruksi 3D merupakan upaya dokumentasi untuk rekonstruksi atau restorasi jika bangunan tersebut hancur.

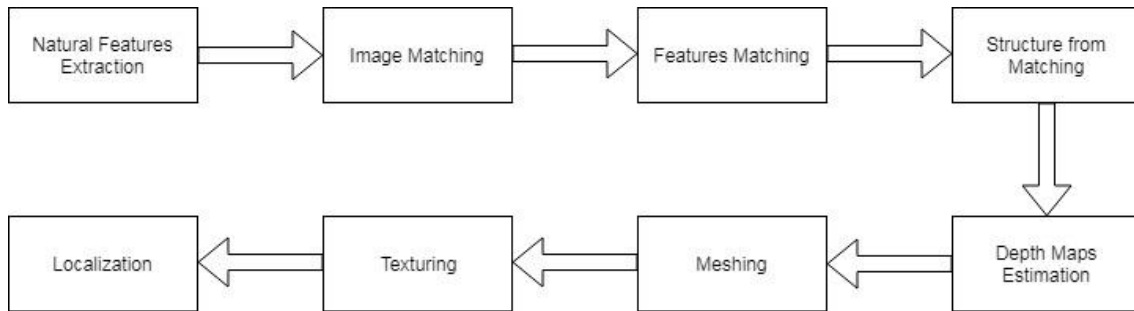
Undang-Undang No. 11 pasal 53 ayat 4 juga mengatur bahwa pelestarian cagar budaya harus didukung oleh kegiatan pendokumentasian sebelum dilakukan kegiatan yang dapat menyebabkan terjadinya perubahan keasliannya. Pendokumentasian tersebut tidak hanya terbatas untuk mengetahui dimensi geometri cagar budaya, namun juga terkait dengan seberapa besar perubahan dimensi geometri yang terjadi dalam rentang waktu tertentu. Salah satu metode pendokumentasian cagar budaya yang saat ini sedang mengalami perkembangan adalah metode pemodelan tiga dimensi. Pemanfaatan metode pendokumentasian dengan pembuatan model tiga dimensi dari benda atau kawasan cagar budaya memberikan banyak keuntungan di antaranya dapat diperoleh data dokumentasi yang memiliki bentuk dan dimensi obyek yang teliti dan mudah untuk disimpan. Oleh karena itu, dewasa ini pembuatan model tiga dimensi untuk kepentingan dokumentasi benda maupun kawasan cagar budaya sangat diperlukan dalam kegiatan pelestarian sehingga

mampu mempertahankan unsur-unsur karya budaya yang ada dalam keadaan cukup lengkap sedemikian rupa sehingga masih mampu memberikan gambaran yang utuh tentang cagar budaya yang ada dan mencerminkan nilai-nilai penting yang dikandungnya [1].

Pemilihan kawasan Candi Penataran sebagai tempat penelitian untuk memodelkan obyek 3D dikawasan tersebut didasarkan Candi Penataran merupakan salah satu Candi termegah dan terluas di Jawa Timur yang terletak di lereng barat Gunung Kelud, di sebelah utara kota Blitar dengan ketinggian 450 meter di atas permukaan air laut. Teknik rekonstruksi obyek 3D terbagi menjadi 2 kategori, yaitu teknik aktif dan teknik pasif. Teknik aktif (object scanning) memerlukan kendali pada cahaya yang terstruktur. Beberapa peneliti menggunakan sebuah proyektor atau viewer untuk menghasilkan cahaya terstruktur[2]. Peneliti lain menggunakan laser beam dan sebuah kamera video. Teknik pasif dilakukan dengan mengambil menggunakan dua atau lebih citra dari sebuah obyek dari berbagai macam posisi dengan kamera [3][4]. Teknik ini sering dikenal dengan adopsi photogrammetry atau structure from motion. Dengan menggunakan pendekatan rekonstruksi model 3D berdasarkan banyak gambar menggunakan algoritma Structure from Motion (SfM) dan Multi View Stereo (MVS) diharapkan hasil pemodelan 3D ini bisa dimanfaatkan sebagai upaya melestarikan obyek 3D di kawasan cagar budaya Candi Penataran.

## 2. METODE PENELITIAN

Tahapan penelitian ini dilakukan secara berurutan dan sistematis (sequence) mulai dari pengumpulan data, proses sampai dengan menghasilkan luaran agar memudahkan dalam proses pengerjaannya. Pada penelitian ini menggunakan model Fotogrammetry Pipeline. Berikut tahapan penelitian yang dilakukan untuk menghasilkan objek 3D dari gambar 2D



Gambar 1. *Photogrammetry Pipeline*

#### A. *Natural Features Extraction*

Tujuan dari langkah ini adalah untuk mengekstraksi kelompok piksel yang berbeda, sampai batas tertentu, tidak berubah pada sudut pandang kamera yang berubah selama akuisisi gambar. Oleh karena itu, fitur dalam adegan harus memiliki deskripsi fitur yang serupa di semua gambar.

#### B. *Image Matching*

Tujuan dari bagian ini adalah untuk menemukan gambar yang mencari ke area yang sama dari pemandangan itu. Untuk itu, digunakan teknik pengambilan gambar untuk menemukan gambar yang membagikan beberapa konten tanpa biaya penyelesaian semua kecocokan fitur secara terperinci. Ambisinya adalah untuk menyederhanakan gambar dalam deskriptor gambar ringkas yang memungkinkan untuk menghitung jarak antara semua deskriptor gambar secara efisien.

#### C. *Feature Matching*

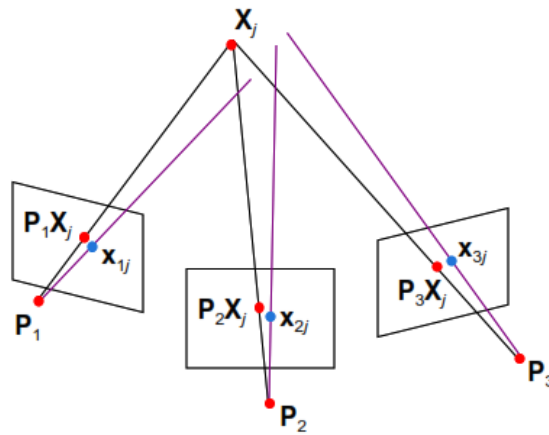
Tujuan dari langkah ini adalah untuk mencocokkan semua fitur antara pasangan gambar kandidat. Pertama, dilakukan pencocokan fotometrik antara set deskriptor ydari 2 gambar masukan. Untuk setiap fitur dalam gambar A, mendapatkan daftar fitur kandidat pada gambar B. Karena ruang deskriptor bukan ruang linier dan terdefinisi dengan baik, maka tidak dapat mengandalkan nilai jarak absolut untuk mengetahui apakah kecocokan tersebut valid atau tidak (hanya

dapat memiliki jarak terikat mutlak lebih tinggi). Untuk menghapus kandidat yang buruk, diasumsikan bahwa hanya ada satu kecocokan yang valid di gambar lain. Jadi untuk setiap deskriptor fitur pada gambar pertama, maka mencari 2 deskriptor terdekat dan menggunakan ambang relatif di antara gambar. Asumsi ini akan mematikan fitur pada struktur berulang tetapi telah terbukti menjadi kriteria yang kuat. Ini memberikan daftar kandidat pencocokan fitur hanya berdasarkan kriteria fotometrik. Temukan 2 deskriptor terdekat dalam gambar kedua untuk setiap fitur yang intensif secara komputasi dengan pendekatan brute force, tetapi ada banyak algoritma yang dioptimalkan. Yang paling umum adalah Perkiraan Tetangga Terdekat, tetapi ada alternatif seperti, *Cascading Hashing*.

#### D. *Structure From Motion*

Tujuan dari langkah ini adalah untuk memahami hubungan geometris di balik semua pengamatan yang disediakan oleh gambar input, dan menyimpulkan struktur adegan yang kaku (titik 3D) dengan pose (posisi dan orientasi) dan kalibrasi internal semua kamera dengan menggunakan *non-linear method bundle adjustment* untuk memperbaiki struktur dan gerak maupun meminimalkan kesalahan proyeksi ulang.

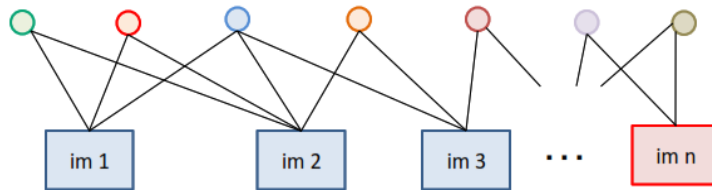
$$E(P, X) = \sum_{i=1}^m \sum_{j=1}^n D(X_{ij}, P_i X_j)^2 \quad (1)$$



Gambar 2. Proses *Bundle Adjustment*

Selanjutnya langkah rekonstruksi akan dihitung secara otomatis dari dua tampilan awal yang diperpanjang secara iterative dengan

menambahkan pandangan baru atau disebut dengan *Incremental SFM*.



Gambar 3. *Incremental SFM*

#### E. Depth Maps Estimation

Untuk semua kamera yang telah diselesaikan oleh SFM, maka ingin mengambil nilai kedalaman setiap piksel menggunakan nilai pendekatan *Semi-Global Matching (SGM)* atau *ADCensus* yang diterapkan di *AliceVision* pada *Meshroom* [13].

#### F. Meshing

Tujuan dari langkah ini adalah untuk membuat representasi permukaan geometris yang padat dari pemandangan.

#### G. Texturing

Tujuan dari langkah ini adalah untuk tekstur mesh yang dihasilkan. Jika mesh tidak memiliki UV terkait, maka akan menghitung peta UV secara otomatis. *AliceVision* mengimplementasikan pendekatan pemetaan UV dasar untuk meminimalkan ruang tekstur.

#### H. Localization

Berdasarkan hasil SFM, dapat dilakukan lokalisasi kamera dan mengambil gerakan kamera animasi di tempat rekonstruksi 3D.

##### 1) Kalibrasi kamera

Parameter kamera internal dapat dikalibrasi dari beberapa tampilan kotak-kotak. Ini

memungkinkan untuk mengambil parameter panjang fokus, titik utama dan distorsi.

##### 2) Pelokalan kamera tunggal

Menggunakan algoritma yang disajikan di bagian pencocokan gambar untuk melokalkan gambar terdekat di hasil SFM. Kemudian melakukan pencocokan fitur dengan gambar-gambar tersebut serta dengan N frame sebelumnya. Kemudian langsung mendapatkan asosiasi 2D-3D, yang digunakan untuk melokalkan kamera.

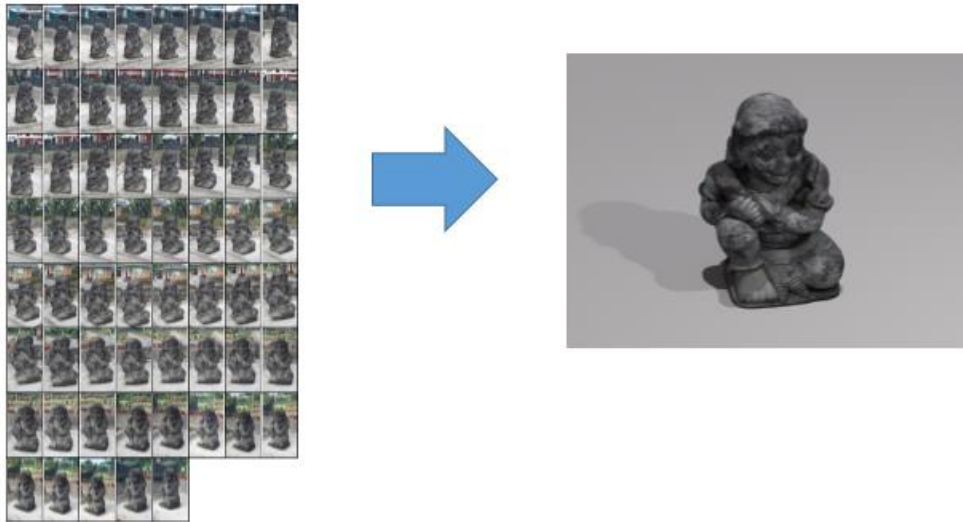
##### 3) Banyak kamera

Jika rig kamera digunakan, dapat melakukan kalibrasi rig. Melokalkan kamera satu per satu di seluruh urutan. Kemudian menggunakan semua pose yang valid untuk menghitung pose relatif antara kamera rig dan memilih nilai yang lebih stabil di seluruh gambar. Kemudian menginisialisasi pose relatif rig dengan nilai ini dan melakukan Penyesuaian Bundel global pada semua kamera rig. Ketika rig dikalibrasi, dapat menggunakannya untuk secara langsung melokalisasi pose rig dari sistem multi-kamera yang disinkronkan dengan pendekatan.

### 3. HASIL DAN PEMBAHASAN

Penelitian rekonstruksi obyek 3D dibangun dari foto sebanyak 61 gambar dari berbagai sudut. Selanjutnya gambar direkonstruksi menggunakan algoritma *Structure From Motion* pada aplikasi *Meshroom*[13] untuk mnghasilkan

model 3D lengkap dengan tekstur. Sedangkan foto sendiri diambil menggunakan kamera Canon EOS-100M. Gambar 4 menampilkan bentuk Rekontruksi Model 3D menggunakan metode *Multi View Stereo* dan *Structurre From Motion*.



Gambar 4. Hasil Rekonstruksi 3D Arca Dwarapala

Tabel 1 merupakan data perbandingan pengolahan data dari pengambilan gambar menggunakan kamera Canon EOS-100M pada gambar asli atau sebelum data dikompres dan

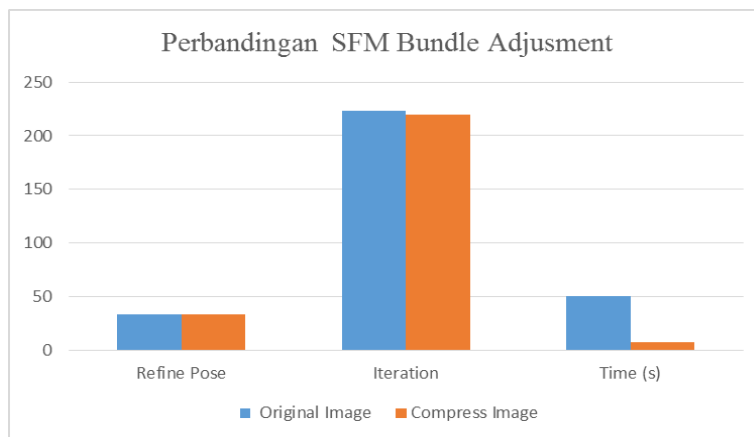
sesudah gambar dikompres guna untuk membandingkan proses dan hasil yang optimal pada tahapan *Structure Form Motion Budle Adjustment*.

Table 1: *Structure Form Motion Budle Adjustment*

No	SFM Bundle Adjustment	Gambar Asli	Gambar Terkompres
1	Refine Pose	33	33
2	Iteration	223	220
3	Time (s)	50.30	6.77

Gambar 5 merupakan grafik hasil perbandingan data dari Tabel 1 yang diperoleh dari pengambilan gambar menggunakan kamera Canon EOS-100M pada gambar asli atau sebelum data dikompres dan sesudah gambar dikompres guna untuk menganalisa hasil dari

perbandingan proses dan hasil yang optimal pada tahapan *Structure Form Motion Budle Adjustment* dengan mengambil nilai dari *Refine Pose*, *Iteration* dan *Time* yang diperlukan dalam proses rekontruksi 3D.



Gambar 5. Grafik Perbandingan SFM *Budle Adjustment*

#### 4. KESIMPULAN

Penelitian ini dilakukan dengan cara mengambil obyek berupa foto sebanyak 61 gambar pada kawasan Candi Penataran Blitar. Hasil foto yang diperoleh direkonstruksi menjadi model 3D menggunakan algoritma *Structure From Motion* pada *Meshroom*[13]. Pada penelitian ini dilakukan uji coba terhadap gambar asli dengan gambar yang telah dikompres untuk dilakukan rekonstruksi untuk dibandingkan proses rekonstruksi 3D dari kedua data input tersebut. Dari 61 gambar dengan diproses menggunakan algoritma *Structure Form Motion* diperoleh data pose kamera dan titik 3D disempurnakan sebanyak 33 baik gambar asli maupun gambar kompres. Untuk jumlah iterasi gambar kompres 1.4 % lebih sedikit dari pada gambar asli dan membutuhkan waktu 43.53% lebih cepat dari gambar asli.

Hasil dari rekonstruksi 3D dapat dikembangkan lebih lanjut menjadi Museum Cagar Budaya dalam bentuk digital atau *Virtual Reality* sehingga keberlanjutan dan catatan sejarah terkait dengan warisan nenek moyang bangsa Indonesia tetap terjaga.

#### DAFTAR PUSTAKA

- [1] Mulyadi, Y., Mengoptimalkan Zonasi Sebagai Upaya Pelestarian Cagar Budaya, [http://www.academia.edu/2923484/Mengoptimalkan\\_Zonasi\\_Sebagai\\_Upaya\\_Pelestarian\\_Cagar\\_Budaya](http://www.academia.edu/2923484/Mengoptimalkan_Zonasi_Sebagai_Upaya_Pelestarian_Cagar_Budaya) (akses tanggal 1 Agustus 2019).
- [2] Dipanda, Albert & Woo, Sungyun. (2005). Towards a real-time 3D shape reconstruction using a structured light system. *Pattern Recognition*. 38. 1632-1650. 10.1016/j.patcog.2005.01.006.
- [3] Prakoonwit, Simant & Benjamin, Ralph. (2007). 3D surface point and wireframe reconstruction from multiview photographic images. *Image and Vision Computing*. 25. 1509-1518. 10.1016/j.imavis.2006.12.019.
- [4] Park, Jong-Seung. (2005). Interactive 3D reconstruction from multiple images: A primitive-based approach. *Pattern Recognition Letters*. 26. 2558-2571. 10.1016/j.patrec.2005.05.009.
- [5] Suharsana, 1997, *Fotogrametri Dasar*, Jurusan teknik Geodesi, Fakultas Teknik Universitas Gadjah Mada, Yogyakarta.
- [6] Atkinson, K. B., 1996, *Close Range Photogrammetry and Machine Vision*. Whittles Publishing. London, UK.
- [7] Hwang, Jin-Tsong & Chu, Ting-Chen. (2016). 3D BUILDING RECONSTRUCTION BY MULTIVIEW IMAGES AND THE INTEGRATED APPLICATION WITH AUGMENTED REALITY. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XLI-B1. 1235-1241. 10.5194/isprs-archives-XLI-B1-1235-2016.
- [8] Scharstein, Daniel & Szeliski, Richard. (2002). A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision*. 47. 7-42. 10.1023/A:1014573219977.
- [9] Okutomi, Masatoshi & Kanade, Takeo. (1993). A Multiple-Baseline Stereo.. *IEEE Trans. Pattern Anal. Mach. Intell.* 15. 353-363. 10.1109/CVPR.1991.139662.
- [10] Tsai, Roger. (1983). Multiframe Image Point Matching and 3-D Surface Reconstruction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on. PAMI-5*. 159 - 174. 10.1109/TPAMI.1983.4767368.
- [11] Baker, Simon & Scharstein, Daniel & Lewis, J.P. & Roth, Stefan & Black, Michael & Szeliski, Richard. (2007). A Database and Evaluation Methodology for Optical Flow. *International Journal of Computer Vision*. 92. 1-31. 10.1007/s11263-010-0390-2.
- [12] Furukawa, Yasutaka & Hernandez, Carlos. (2015). Multi-View Stereo: A Tutorial. *Foundations and Trends® in Computer Graphics and Vision*. 9. 1-148. 10.1561/06000000052.
- [13] <https://alicevision.org/#photogrammetry> (akses tanggal 1 Agustus 2019).