

Rekonstruksi 3D Untuk Pemetaan Ruang Menggunakan Kamera Stereo

¹ Septiantar Tebe Nursaputro, ² Dadet Pramadihanto, ³ Bima Sena Bayu Dewantara

^{1,2} Teknik Elektro, Politeknik Elektronika Negeri Surabaya

E-mail : ¹septiantartebenursaputro@gmail.com, ²dadet@pens.ac.id, ³bima@pens.ac.id

Abstrak

Pembuatan lokalisasi peta oleh pemikiran manusia dengan merekam segala situasi di sekitar mereka sangat mudah dilakukan bahkan tanpa harus berpikir ulang. Proses yang dilakukan oleh manusia seakan tidak membutuhkan suatu pembelajaran, hanya dengan melihat sekilas dapat memperoleh peta seluruh sudut ruang, beserta obyeknya. Dengan menggunakan kebiasaan manusia ini, komputer telah merekayasa indera penglihatan manusia dengan mengaplikasikan kombinasi dua kamera tersusun secara stereo. Mengacu pada penelitian sebelumnya, paper ini telah dilakukan melanjutkan bagaimana membentuk peta ruang 3D dengan penglihatan 360°. Dengan menggunakan algoritma *stereo variant* dan RANSAC untuk mendapatkan *descriptor* mempermudah menggabungkan gambar-gambar membentuk sebuah peta.

Kata kunci : Pemetaan 3D, *stereo variant*, RANSAC

Abstract

Making localization maps by human thought by recording all the situations around them is very easy to do even without the urge to rethink. The process carried out by humans as if it does not require a learning, just by glancing can get a map of all corners of space, along with their objects. By using this human habit, the computer has engineered the senses of human vision by applying a combination of two cameras arranged in stereo. Referring to previous research, this paper has been carried out to continue how to form 3D space maps with 360o vision. Using the stereo variant algorithm and RANSAC to get a descriptor makes it easy to combine images to form a map

Keywords : Mapping 3D, *stereo variant*, RANSAC

I. PENDAHULUAN

Era robotika saat ini diinginkan suatu robot yang dapat merasakan kondisi sekitar dengan bantuan sensor secara mandiri. Diantara sensor tersebut ialah kamera, sebagai indera penglihatan. Untuk mencapai teknologi indera penglihatan layaknya mata manusia, dibutuhkan dua kamera yang terintegrasi bekerja bersama, menjadi sebuah kamera stereo. Sistem mata manusia adalah melihat dan menangkap gambaran obyek di sekitar, kemudian disimpan sebagai data, sehingga data tersebut salah satunya difungsikan untuk membangun sebuah peta tiga dimensi.

Pembuatan peta atau obyek tiga dimensi, diperlukan beberapa gambar dua dimensi memuat informasi tentang peta tersebut. Bidang *computer vision* memiliki banyak algoritma membangun beberapa gambar dua dimensi tersebut menjadi gambar bidang tiga dimensi [1,2,3], yaitu dengan memanfaatkan gambar kedalaman (*depth*). Gambar kedalaman merupakan representasi informasi jarak obyek terhadap kamera dalam satuan piksel.

Pemanfaatan kamera stereo untuk membangun gambar kedalaman telah dilakukan oleh banyak studi, diantaranya, [4] dengan menggunakan kamera Minoru 3D webcam. Bima dkk membuat pendeteksi tubuh manusia dengan memanfaatkan siluet yang dihasilkan gambar kedalaman. Bima menggunakan algoritma *stereoSGBM* untuk membuat gambar kedalaman dan *histogram of oriented gradien* (HOG) sebagai pendekteksi ciri tubuh manusia.

Penelitian yang lain dilakukan oleh Waskitho [5] untuk membuat gambar bidang tiga dimensi, menggunakan dua kamera mono yang dikombinasi dengan *structure sensor*. Penelitian dilanjutkan oleh Anjarlistiawan [6] menggunakan dua kamera mono dan *structure sensor* yang sama untuk membangun bidang tiga dimensi, dengan dimensi ruang 120° perputaran kamera. Hasil keluaran gambar kedalaman dari dua kamera mono tersebut dikalibrasi dengan hasil dari *structure sensor*. Waskito dan Anjarlistiawan menempatkan koordinat titik piksel x dan y dari gambar dua kamera mono dan koordinat z dari *structure sensor* untuk membentuk bidang tiga dimensi.

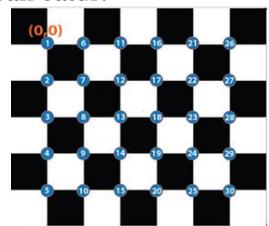
Dalam pelenlitian ini, dibangun peta ruang tertutup bidang tiga dimensi, dengan perputaran kamera 360° (sekeliling ruangan). Penelitian ini menggunakan algoritma *stereoVar* [3] sebagai metode membetuk gambar kedalaman dan algoritma *Random Sample Consensus* (RANSAC)

sebagai metode menggabungkan beberapa gambar dua dimensi menjadi satu kesatuan panorama ruang tiga dimensi.

II. METODE PENELITIAN

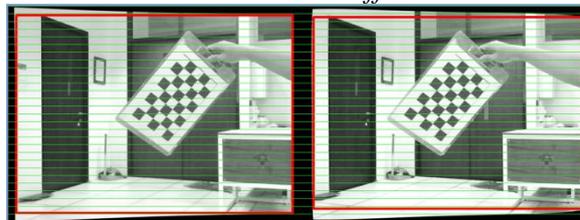
2.1. Kalibrasi Kamera

Kamera stereo perlu dikalibrasi sebelum digunakan mengambil data kedalaman (*depth*). Proses kalibrasi, direkomendasikan menggunakan pose papan catur. Pertama definisikan dulu besar ukuran lebar dan panjang dari papan catur, sehingga bisa didapatkan titik yang berkesesuaian antar sudut papan catur.



Gambar 1. Definiskan ukuran papan catur

Masing-masing kamera akan menangkap gambar papan catur, sehingga akan tampak variabel translasi dan rotasi dari kedua kamera yang berkesesuaian, seperti ditunjukkan oleh Gambar 2. Agar hasil tangkapan kamera jelas dan rotasi tidak berubah, kedua kamera harus diatur besar fokus dan kontras secara manual. Pengaturan auto fokus dan auto kontras, juga *white balance* dalam kondisi *off*.



Gambar 2. Titik sudut pose papan catur yang berkesesuaian

2.2. Pengambilan Gambar Depth

Setelah dilakukan kalibrasi, didapatkan nilai parameter intrinsik (persamaan 1) dan ekstrinsik (persamaan 2) dari kamera yang digunakan. Kedua unsur tersebut akan dijadikan parameter untuk membuat dua gambar berpasangan dari kamera stereo menjadi satu gambar kedalaman.



Gambar 3. Data set gambar stereo dari middlebury: (a) gambar kiri; (b) gambar kanan; (c) gambar kedalaman (*depth*)

Unsur intrinsik merupakan unsur yang memuat informasi fisik kamera, seperti *focal length*, pengaturan sensor kamera, dan jarak antar fokus kamera (*baseline*).

$$A = \begin{bmatrix} \alpha_x & \gamma & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

Keterangan: α_x dan α_y merepresentasikan sebagai besar *focal length* kamera, γ merupakan koefisien antara koordinat x dan y, u_0 dan v_0 adalah titik utama koordinat gambar, atau titik pusat O (0,0).

$$C = -R^{-1}T = -R^T T \quad (2)$$

Unsur ekstrinsik merepresentasikan koordinat transformasi di dunia nyata, yang berisi informasi rotasi dan translasi obyek terhadap titik pusat kamera.

Algoritma untuk membuat gambar kedalaman (*depth*) dalam penelitian ini menggunakan metode *stereoVar*. *Stereo variant* adalah algoritma stereo yang memandang obyek pergeseran titik piksel berkesesuaian dari dua gambar stereo sebagai obyek eksponensial, bukan piksel diskret yang tetap :

$$E(u(x,y)) = \iint dS \|I_1(x,y) - I_2(xu(x,y),y)\|^2 \quad (3)$$

(a) Keterangan: $E(u(x,y))$ adalah pengaturan *gray-scale* untuk kedua gambar kamera stereo I_1 dan I_2 . Kemudian gambar keabuan E diperhalus untuk memperbaiki masalah luasan piksel kedalaman yang ikut terkalkulasi dari piksel tetangga. Metode memperhalus gambar kedalaman secara umum :

$$\Psi(|\nabla u(x,y)|^2) \quad (4)$$

Maka, secara keseluruhan dapat dirumuskan, yang mana φ adalah *weighting factor* parameter kehalusan :

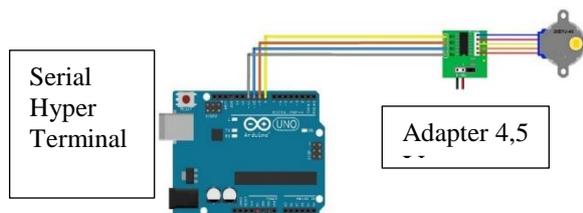
$$(c) E(u(x,y)) = \iint_{\Omega} \iint \|I_1(x,y) - I_2(x + u(x,y),y)\|^2 + \varphi \cdot \Psi(|\nabla u(x,y)|^2) dS \quad (5)$$

2.3. Rancangan perangkat elektronik



Gambar 4. Rangkaian perangkat keras dan kamera

Perangkat penggerak yang dibuat adalah berupa rangkaian motor stepper dan menempatkan kamera di poros (as), sehingga kamera dapat berputar 360°. Kamera yang digunakan dalam penelitian ini adalah kamera stereo Minoru 3D Webcam. Motor stepper dikendalikan oleh mikrokontroler Arduino UNO, melalui driver motor ULN 2003. Motor stepper yang digunakan adalah jenis 28BYJ-48 dengan tegangan input 5V. Input di driver ULN 2003 dihubungkan dengan port digital Arduino., dengan berurutan IN 1, IN 2, IN 3, IN 4 (driver ULN 2003) dihubungkan dengan port 8, 9, 10, 11 (Arduino).



Gambar 5. Desain perangkat elektronik

Tegangan arduino menggunakan serial dari laptop, sedangkan tegangan driver motor menggunakan AC-DC adapter yang diseting tegangan di 4,5 V. Data serial dari Hyper Terminal yang dikirimkan berupa data *string* dan kemudian dibaca oleh Arduino, sehingga Arduino dapat memberi perintah agar kamera berputar ke langkah motor stepper selanjutnya. Setelah motor stepper berputar sekian derajat searah jarum jam. Perputaran dilakukan setiap 22,5 derajat lalu berhenti memberikan waktu *Visual Code* melakukan perintah *snapshot* menangkap beberapa gambar. Setelah menangkap gambar, user memberikan perintah serial kembali untuk perputaran derajat selanjutnya.

2.4. Stitching Gambar

Metode *stitching* merupakan cara menggabungkan dua gambar atau lebih yang memiliki informasi obyek yang saling berkesesuaian (*overlap*). Semakin besar obyek yang berkesesuaian, maka semakin banyak titik-titik piksel yang sama (*descriptor*), dan akan semakin mudah kedua gambar untuk disatukan. Salah satu metode *stitching* yang digunakan untuk penelitian ini adalah *Random Sample Consensus (RANSAC)*. RANSAC adalah sistem yang stabil untuk mencari *descriptor* dari dua atau lebih gambar yang memiliki obyek berkesesuaian, dan memiliki solusi yang tepat untuk menggabungkan titik-titik *overlap* tersebut.

Dalam penelitian ini, cara menggabungkan menjadi gambar tiga dimensi, terlebih dahulu mencari *descriptor* berdasarkan gambar dua dimensi yang telah diambil kamera stereo. Hal ini mirip saat membuat gambar panorama melalui kamera telepon seluler. Dalam kasus panorama [7], bisa diatur jumlah obyek yang memiliki informasi sama (sesuai r , dan perhitungan homografi kedua gambar adalah H . Bila diumpamakan memiliki *descriptor* sebanyak $n = 500$. Bila probabilitas informasi yang cocok (*matching*) adalah π , maka probabilitas informasi gambar yang cocok dari kedua gambar sebagai berikut:

$$p(H_{is\ correct}) = 1 - (1 - (\pi)^r)^n \quad (6)$$



Gambar 6. Dua gambar kamera stereo yang saling *overlap* ditandai sebagai *descriptor*

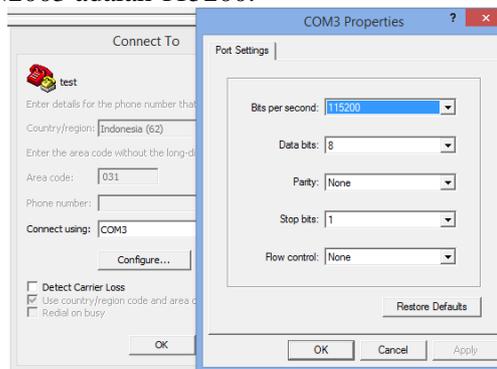
III. HASIL DAN PEMBAHASAN

3.1. Setup eksperimen/percobaan

Penelitian ini menggunakan perangkat lunak *Microsoft Visual C++ 2010 Express*, *OpenCV 2.4.9*, dan *HyperTerminal*. Pergerakan rotasi kamera dikendalikan melalui komunikasi serial. Komunikasi serial ini dihubungkan ke aplikasi *HyperTerminal* dengan memberi sinyal karakter 'c'. Gambar dari kamera stereo ditangkap sebanyak 1 kali, setiap sudut putar 22,5°.

$$\left(\frac{360}{22,5}\right) \text{derajat} = 16 \quad (7)$$

Didapatkan 16 kali penangkapan gambar dari kamera stereo dan langsung diubah ke gambar kedalaman. Pada aplikasi *HyperTerminal* disesuaikan *port* yang digunakan dan *baud rate* dengan Arduino UNO. Baud rate dari driver ULN2003 adalah 115200.



Gambar 7. Pengaturan di aplikasi *HyperTerminal*

3.2. Percobaan pembuatan gambar Depth



Gambar 8. Kalibrasi pose papan catur

Dilakukan berbagai macam pose kalibrasi papan catur, agar untuk mendapatkan perkiraan rotasi dan translasi dari obyek gambar yang ditangkap setelah proses kalibrasi. Didapatkan unsur intrinsik dan ekstrinsik, dan unsur intrinsik merupakan unsur utama dalam kamera stereo.

TABEL 1. TABEL UNSUR INTRINSIK KAMERA

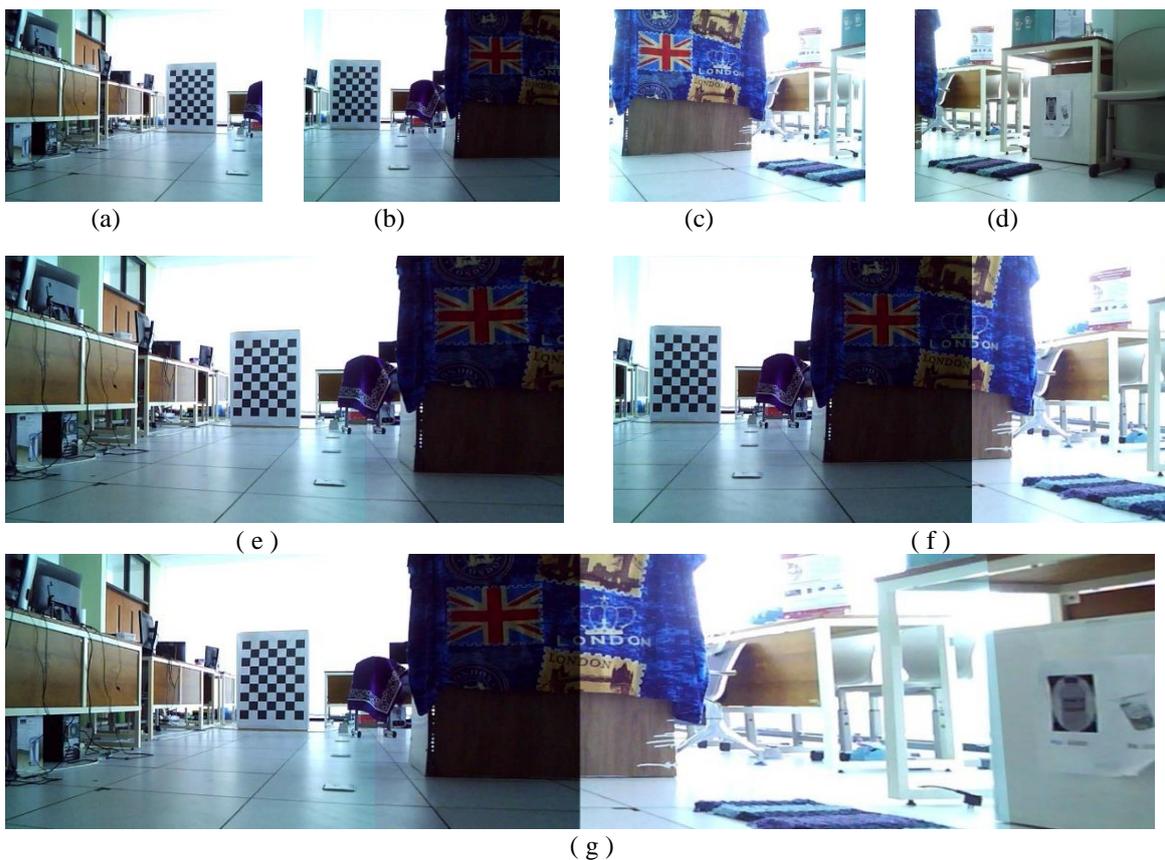
Parameter	Kamera kiri	Kamera kanan
<i>Tangetial Distortion</i>	[0,0]	[0,0]
<i>Focal length</i>	8.4901429475093266	8.4901429475093266
<i>Baseline</i>	60 mm	

Kemudian lanjut menggunakan unsur intrinsik dan ekstrinsik tersebut untuk membuat semua pasang gambar kamera stereo menjadi gambar kedalaman.



Gambar 9. Gambar kedalaman dari gambar kamera kiri dan gambar *depth*

3.3. Percobaan *stitching* gambar

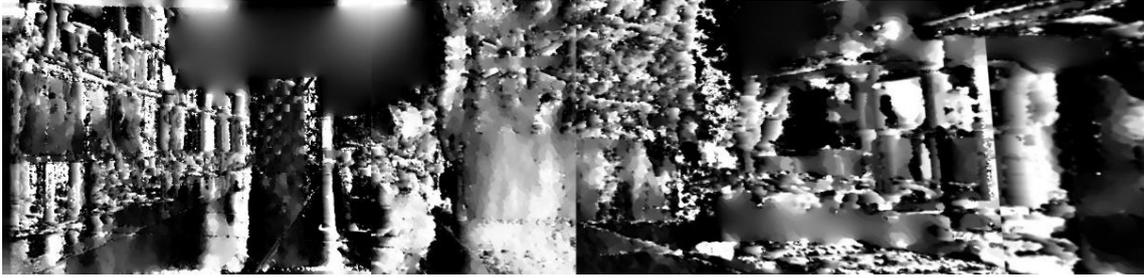


Gambar 10. Gambar proses *stitching*, (a,b,c,d) gambar data referensi kamera stereo (e, f) gambar proses *stitchng* 1; dan (g) Gambar proses *stitching* 2

Mengingat data gambar memiliki resolusi yang besar, proses *stitching* dilakukan tiga kali. Proses pertama dan kedua dengan mencari *descriptor* antar tiap 4 data gambar dua dimensi,

kemudian digabungkan sesuai informasi yang saling *overlap*. Proses ketiga adalah dengan hasil *descriptor* yang telah disimpan, kemudian

dimunculkan kembali sebagai parameter menggabungkan data gambar kedalaman.



Gambar 11. Gambar *stitching depth* dari gambar 10.g

Gambar tiga dimensi dalam penelitian ini dimaksudkan adalah bidang tiga dimensi yang memuat informasi gambar dua dimensi (Gambar 10.g) di koordinat x dan y, kemudian ditambah dengan informasi koordinat z, yaitu jarak obyek dengan kamera, yang terdapat di gambar kedalaman (Gambar 11).

3.4. Waktu komputasi

Waktu komputasi yang telah dilakukan rata-rata adalah 380 ms sampai dengan 410 ms, untuk melakukan *capture* tiap gambar kamera stereo dan langsung dijalankan proses gambar kedalaman.

IV. KESIMPULAN

Membangun peta ruang bidang tiga dimensi dapat dilakukan dengan menggabungkan titik-titik piksel dari gambar referensi dua dimensi dan gambar kedalaman (*depth*). Dari gambar referensi dua dimensi terdapat informasi dalam proyeksi koordinat x dan y, sedang kan gambar kedalaman terdapat informasi jarak obyek terhadap benda, dalam proyeksi koordinat z.

Saran untuk penelitian selanjutnya, dilakukan dalam bentuk visualisasi *point cloud* sehingga dapat nampak jelas jarak obyek dengan kamera dan jarak obyek satu dengan yang lainnya. Selain itu, dipertimbangkan pula jenis obyek yang memiliki sifat homogen dan tanpa tekstur, karena dapat menimbulkan *miss-matching* saat proses *stitching* dilakukan.

DAFTAR PUSTAKA

- [1] Faugeras, Olivier, Bernard Hotz, Hervé Mathieu, Thierry Viéville, Zhengyou Zhang, Pascal Fua, Eric Théron et al. *Real time correlation-based stereo: algorithm, implementations and applications*. Doctoral Dissertation. France: Institut National de Recherche en Informatique et Automatique (INRIA), 1993.
- [2] Hirschmuller, Heiko. *Stereo processing by semiglobal matching and mutual information*. IEEE Transactions on pattern analysis and machine intelligence 30.2 (2007): 328-341.
- [3] Sergey Kosov, Thorsten Thormählen and Hans-Peter Seidel: *Accurate real-time disparity estimation with variational methods*. In: Advances in Visual Computing. Springer Berlin Heidelberg, 2009. 796-807.
- [4] Dewantara, Bima Sena Bayu, Fernando Ardilla, and Ardiansyah At Thoriqy. *Implementation of Depth-HOG based Human Upper Body Detection On A Mini PC Using A Low Cost Stereo Camera*. 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIIT). IEEE, 2019.
- [5] Waskitho, Suryo Aji, et al. *FloW vision: Depth image enhancement by combining stereo RGB-depth sensor*. 2016 International Conference on Knowledge Creation and Intelligent Computing (KCIC). IEEE, 2016.
- [6] Anjarlistiawan, Rimsya, et al. *Multi process matching depth image between stereo camera and structure sensor in realtime*. 2017 International Electronics Symposium on Engineering Technology and Applications (IES-ETA). IEEE, 2017.
- [7] Brown, Matthew, and David G. Lowe. *Automatic panoramic image stitching using invariant features*. International journal of computer vision 74.1 (2007): 59-73.