

REGISTRASI 3D MENGGUNAKAN SCALE INVARIANT FEATURE TRANSFORM (SIFT)

Farrady A.F. *, Djoko P., Rudy D.

Electronics Department

Institut Teknologi Sepuluh Nopember, Surabaya
Kampus ITS Keputih, Sukolilo, Surabaya 60111, Jawa Timur.
farrady@gmail.com*

Abstrak

Penyelarasan gambar 2D yang telah dipelajari pada bagian lain computer vision, dengan mendeteksi ciri hingga diperoleh matrik homography yang menghubungkan kedua gambar tersebut. SIFT diusulkan dalam penelitian ini karena memiliki kestabilan dalam berbagai situasi. KNN dan RANSAC digunakan untuk menemukan pasangan yang tepat. Meskipun matrik yang dihasilkan dari tersebut tidak mempengaruhi komponen kedalaman (Z) pada gambar 3D, matrik tersebut dapat dimanfaatkan untuk mendapatkan matrik transformasi 3D berdasarkan titik-titik yang berkorespondensi pada kedua gambar dengan melibatkan komponen kedalaman. Sehingga dapat diperoleh sebuah gambar 3D yang merupakan hasil registrasi dari dua buah gambar 3D berdasarkan matrik transformasi 3D tersebut.

Kata kunci : Registrasi 3D, SIFT, Matrik transformasi 3D.

1. Pendahuluan

Dalam proses menyelaraskan dua buah gambar dengan scene yang berbeda didasarkan pada kesesuaian ciri yang ada pada tiap-tiap gambar. SIFT (Scale Invariant Feature Transform) menyajikan stabilitas dalam berbagai situasi kecuali teriluminasi dibandingkan dengan SURF (Speed Up Robust Features). Selain itu SIFT jauh lebih lambat dibandingkan dengan SURF [1]. Dalam penelitian ini kami menggunakan SIFT sebagai pendeteksi ciri karena kami lebih mengutamakan ketepatan dari pada kecepatan. Selain itu kami menggunakan KNN (K-Nearest Neighbour) digunakan untuk menemukan kesesuaian ciri, dan RANSAC (Random Sample Consensus) digunakan untuk menepis ketidak-konsistenan pasangan agar diperoleh pasangan titik yang tepat.

Pada penelitian yang lain SIFT dimanfaatkan untuk aplikasi pengenalan rambu-rambu lalu-lintas [2] dan mosaic [3]. Pada penelitian D.Lowe tidak hanya membahas SIFT tetapi juga membahas tentang kesesuaian titik yang juga didasarkan pada kesesuaian tetangga terdekat. Dalam penelitian tersebut diberikan pengukuran yang efektif dalam pemilihan tetangga dengan membandingkan tetangga terdekat dengan tetangga terdekat kedua.

Gambar 3D diperoleh dari rekonstruksi stereo vision terkalibrasi. Berdasarkan survey yang dilakukan oleh Q.Wei dan rekan bahwa metode multi-ocular lebih presisi

dibandingkan dengan metode mono-ocular. Kalibrasi stereo camera sangat diperlukan untuk memperoleh keserasian pasangan gambar yang dihasilkan. Berdasarkan titik-titik yang berkorespondensi dapat dihitung secara geometri kedalaman pixel tersebut sehingga diperoleh gambar 3D. Dalam proses penyelarasan kami mencoba memanfaatkan ketiga metode tersebut untuk deteksi ciri hingga diperoleh matrik transformasi 2D. Kemudian dengan melibatkan kedalaman pixel pada gambar 3D yang berkorespondensi untuk mendapatkan matrix transformasi 3D.

SIFT terdiri dari 4 tahap utama yaitu : scale-space extrem detection, keypoint localization, orientation assignment dan keypoint descriptor. Pada tahap pertama menggunakan fungsi DOG (Difference of Gaussian) untuk mengidentifikasi titik-titik yang berpotensi [4].

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \dots (1)$$

Pada keypoint localization, terjadi proses penepisan titik dengan kontras rendah dan respon tepi. Matrix Hessian digunakan untuk komputasi kelengkungan dan mengeliminasi keypoint yang memiliki kelengkungan melebihi rasio. Orientation histogram diperoleh dari gradien orientasi dari beberapa titik sampel di sekitar keypoint untuk mendapatkan orientation assignment [4].

2. Pembahasan

Dalam penelitian ini, Kami menggunakan komputer standart yang mampu mengoperasikan EmguCV dan sepasang kamera standar dengan resolusi 640x480 pixel. Sebelum pengambilan gambar, sepasang kamera tersebut dikalibrasi menggunakan 15 pasang gambar papan catur, sehingga diperoleh parameter intrinsik dan ekstrinsik kamera.

Pada bagan 1 adalah beberpa sepasang gambar hasil pengambilan gambar dari beberapa sudut pandang yang telah dikalibrasi. Sedangkan pada bagan 2 adalah gambar 3D hasil dari penentuan kedalaman pixel pada titik-titik yang berkorespondensi dalam bentuk jaring segitiga primitif. Koordinat pada bidang XY disesuaikan dengan gambar 2D sebelah kanan pada tiap pasang gambar.

Penggunaan ketiga metode tersebut pada proses penyelarasan 2D adalah untuk memperoleh matrik transformasi 2D yang menghubungkan antara model dan observer. Karena matrik tersebut tidak mengakomodir

komponen kedalaman *pixel*, maka matrik tersebut harus diubah menjadi matrik transformasi 3D. Ketika *observer* ditransformasikan secara 2D pada menuju *model*, sehingga akan ada bidang yang berhimpitan pada bidang XY. *Vertex* pada bidang tersebut pada sisi *observer* sebagai matrik keluaran (P_O). Sedangkan pada sisi *observer* sebelum ditransformasi secara 2D sebagai matrik masukan (P_M). Berdasarkan persamaan 2 dapat ditentukan matrik transformasi 3D (M_{3D}) yang menghubungkan matrik P_M dengan matrik P_O .

Tidak semua *vertex* pada *model* mendapatkan pasangan tepat berupa *vertex* pada *observer*. Maka pada *model* dibuat sebuah *vertex* maya yang memiliki informasi posisi pada bidang XY sesuai dengan *vertex* yang telah ditransformasi secara 2D, sedangkan informasi kedalamannya diperoleh dari kedalaman relatif dari tiga buah *vertex* di sekitarnya yang membentuk suatu segitiga primitif. Prosedur penentuan *vertex* maya pada *model* adalah seperti pada bagan 3.

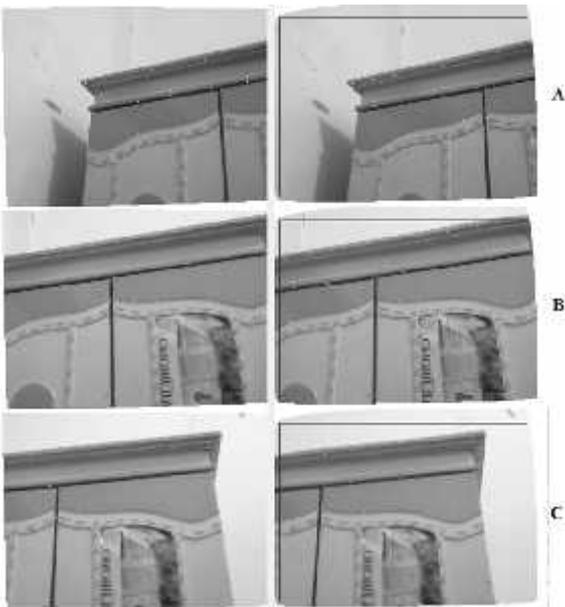
$$P_M = \begin{bmatrix} X_1 & X_2 & \dots & X_n \\ Y_1 & Y_2 & \dots & Y_n \\ Z_1 & Z_2 & \dots & Z_n \\ 1 & 1 & \dots & 1 \end{bmatrix} \quad P_O = \begin{bmatrix} X_1 & X_2 & \dots & X_n \\ Y_1 & Y_2 & \dots & Y_n \\ Z_1 & Z_2 & \dots & Z_n \end{bmatrix}$$

$$P_O = M_{3D} \cdot P_M$$

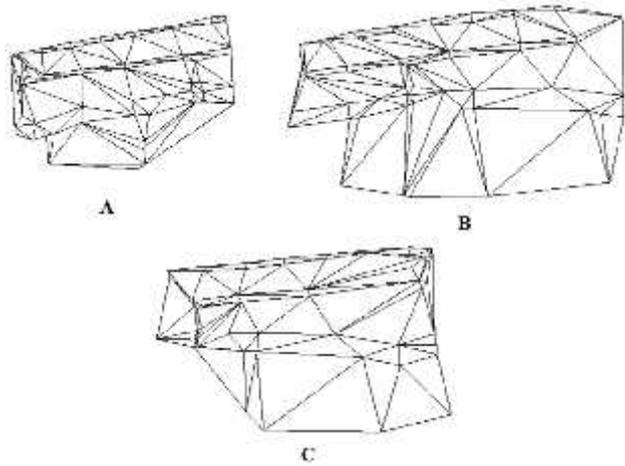
$$M_{3D} = P_O \cdot P_M^+ \dots \dots \dots (2)$$

Dimana

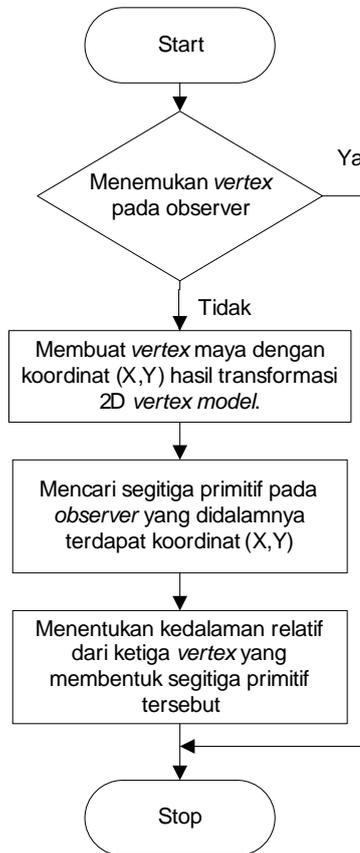
$$P_M^+ = P_M^T (P_M P_M^T)^{-1}$$



Bagan 1 : Gambar objek berupa almari. gambar terkalibrasi dengan sudut pandang yang berbeda.



Bagan 2 : Gambar 3D dalam bentuk wireframe dari tiap-tiap pasang gambar pada bagan 1 dengan titik korespondensi yang dapat terpantau.

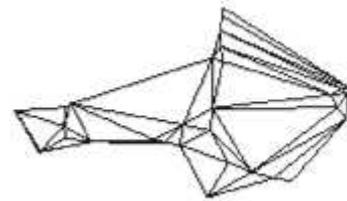


Bagan 3 : Diagram alir untuk menentukan matrik P_O

II. Pengujian

SIFT diterapkan pada gambar yang akan diselarskan pada gambar sebelah kanan dari setiap pasang gambar. Maka akan diperoleh matrik transformasi 2D seperti tampak pada bagan 4. Setelah dihitung matrik transformasi 3D, maka akan tampak seperti pada bagan 5.

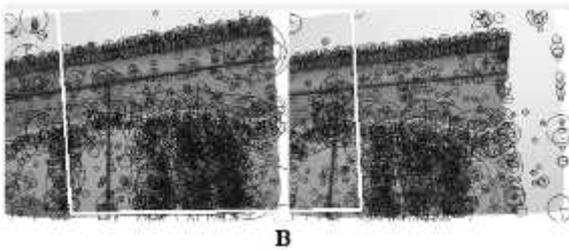
Pada kasus lain, dengan menerapkan cara yang sama, pada bagan 6 adalah gambar objek yang sama tetapi dengan sudut pandang yang lebih ekstrim. Bagan 7 adalah bentuk 3D dari bagan 6 yang akan diregistrasikan dengan bagan 5, hasilnya akan ditampilkan pada bagan 9 dari 3 arah pandang yang berbeda. Pada bagan 8 adalah penerapan SIFT untuk menemukan matrik transformasinya. Pada bagan 10 adalah bentuk *solid* dari bagan 9. Warna *vertex* yang cerah menandakan bahwa *vertex* terletak lebih jauh dari kamera dari pada *vertex* yang berwarna gelap.



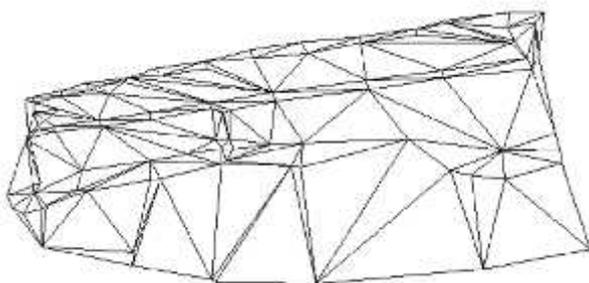
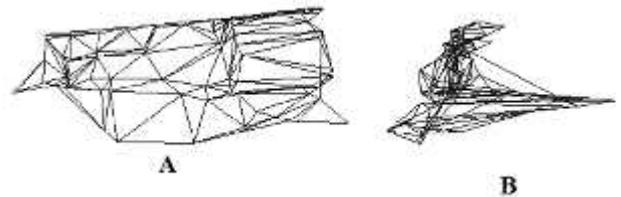
Bagan 7 : Gambar 3D dalam bentuk *wireframe* dari bagan 5 dengan titik korespondensi yang dapat terpantau.



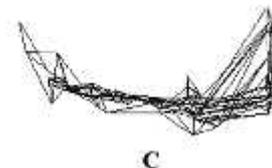
Bagan 8 : Gambar sebelah kiri adalah *model*, dan gambar sebelah kanan adalah *observer*. Garis kotak putih adalah *region of interest model* pada *model*. bagan 6 sebagai *observer* dan bagan 1C sebagai *model*.



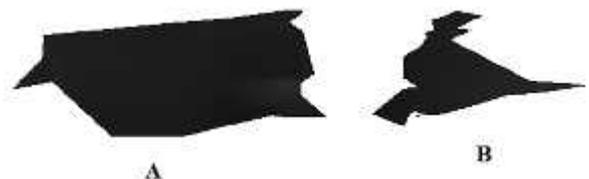
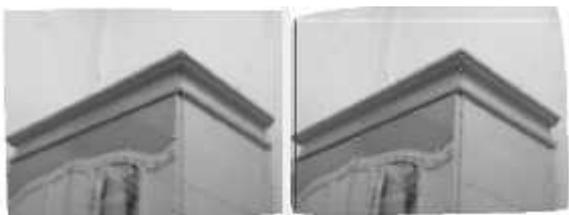
Bagan 4 : Gambar sebelah kiri adalah *observer*, dan gambar sebelah kanan adalah *model*. Garis kotak putih adalah *region of interest model* pada *model*. Pada bagan 4A, bagan 1B sebagai *model* dan bagan 1A sebagai *observer*. Pada bagan 4B, bagan 1B sebagai *model* dan bagan 1C sebagai *observer*.



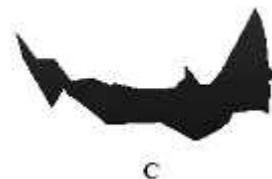
Bagan 5 : hasil registrasi 3D menggunakan matrix transformasi 3D dalam bentuk *wireframe*.



Bagan 9 : hasil registrasi 3D menggunakan matrix transformasi 3D dalam bentuk *wireframe*. Bagan 9A tampak depan, bagan 9B tampak samping kiri, bagan 9C tampak atas.



Bagan 6 : Gambar objek berupa almari. gambar terkalibrasi dengan sudut pandang yang berbeda.



Bagan 10 : Bentuk solid dari bagan 9.

3. Kesimpulan dan penelitian mendatang

SIFT tidak hanya dapat digunakan untuk menyelaraskan gambar 2D, tapi juga dapat dimanfaatkan untuk menyelaraskan gambar 3D. Tidak menutup kemungkinan penggunaan jenis pendeteksi ciri yang lain untuk memperoleh hasil registrasi lebih baik lagi. Selain itu perlu penelitian yang lebih lanjut algoritma khusus untuk pendeteksi ciri dalam 3D.

Daftar Pustaka

- [1] J. Luo and G. Oubong. "A Comparison of SIFT, PCA-SIFT, and SURF". *International Journal of Image Processing (IJIP)* Volume:3, Issue:4. pp.143—152. 2009
- [2] Kus, M.C.; Gokmen, M.; Etaner-Uyar, S. "Traffic sign recognition using Scale Invariant Feature Transform and color classification". *ISCIS '08*. pp: 1-6, Oct. 2008.
- [3] Yang zhan-long and Guo bao-long. "Image Mosaic Based On SIFT", *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp:1422-1425,2008.
- [4] D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints", *IJCV*, 60(2):91–110, 2004.
- [5] Q.Wei, et al. "Converting 2D to 3D: A Survey". *Research Assignment. ITC-University of Technology, Netherlands*. 2005.
- [6] V.Kolmogorov and R.Zabih, "Multi-camera scene reconstruction via graph cuts," in *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part III*. London, UK: Springer-Verlag, pp. 82–96. 2002
- [7] N. Campbell, G. Vogiatzis, C. Hernandez, and R. Cipolla, "Automatic 3d object segmentation in multiple views using volumetric graph-cuts," *Image and Vision Computing*, vol. 28, no. 1, pp. 14–25, 2010.

Biodata Penulis

Farrady Alif Fiolana, memperoleh gelar Sarjana Teknik (S.T), Teknik Elektro ITS, lulus tahun 2006. Saat ini sedang pendidikan Magister Teknik Elektro ITS. Saat ini menjadi Dosen Teknik Elektro UNP-Kediri.

Djoko Purwanto, Saat ini menjadi dosen Teknik Elektro ITS.

Rudy Dikairono, Saat ini menjadi dosen Teknik Elektro ITS.