

Rekonstruksi Lintasan Objek Bergerak Pada Citra Digital Video Dengan Metode Belief Propagation

Efron Manik¹

¹ Prodi Pendidikan Matematika, Universitas HKBP Nommensen
efmanikl@gmail.com

Abstract

The control device in an auto-navigation car is a computer assisted by two cameras on the front of the car to determine the three-dimensional coordinates of all pixels of the digital image captured by the camera from frame to frame. So that the computer will be able to determine where the streets are flat or potholes. The purpose of this research is to create a program code that is able to determine the disparity from one frame to the next. With the help of recordings from two cameras, the disparity of any object captured on the two cameras will be calculated for each pixel. This data is used to compare the accuracy of the methods developed in this study. We use a program code developed to calculate the disparity of each frame in the video. The model developed in this study still needs to be studied in order to find a better method.

Keyword: belief propagation, disparity, stereo configuration.

Abstrak

Perangkat pengendali pada mobil autonavigasi adalah komputer yang dibantu oleh dua kamera di bagian depan mobil untuk menentukan koordinat tiga dimensi dari semua piksel dari gambar digital yang diambil oleh kamera dari frame ke frame. Sehingga komputer akan dapat menentukan di mana jalan-jalan yang datar atau berlubang. Tujuan dari penelitian ini adalah membuat kode program yang mampu menentukan disparity dari satu frame ke frame berikutnya. Dengan bantuan rekaman dari dua handycam, disparity dari objek apa pun yang ditangkap di dua kamera akan dihitung untuk setiap piksel. Data ini digunakan untuk membandingkan akurasi dari metode yang dikembangkan dalam penelitian ini. Kami menggunakan kode program yang dikembangkan untuk menghitung disparity setiap frame dalam video. Model yang dikembangkan dalam penelitian ini masih perlu dipelajari untuk mendapatkan metode yang lebih baik.

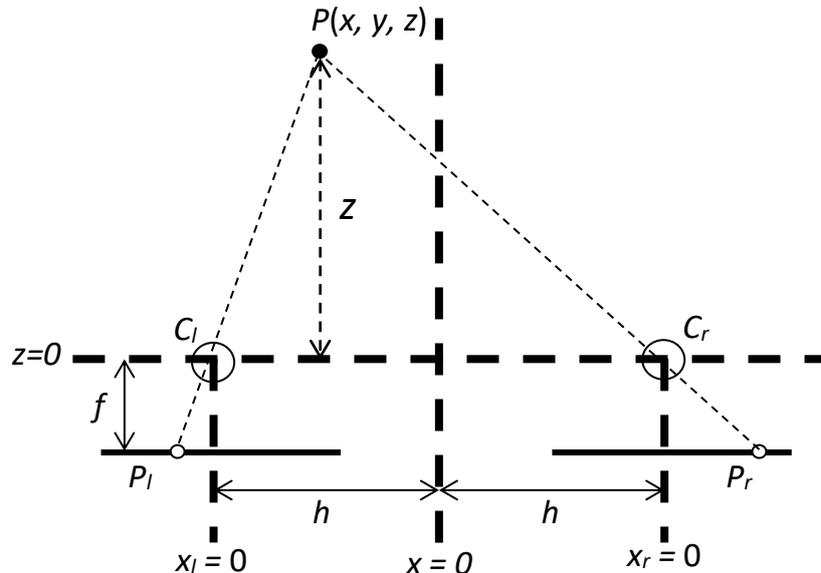
Kata Kunci: belief propagation, disparity, konfigurasi stereo.

1. Pendahuluan

Salah satu masalah besar pada Komputer Vision adalah bagaimana merekonstruksi koordinat dimensi tiga dari setiap pixel objek bergerak yang tampak pada suatu citra digital (invers problem: mengembalikan setiap pixel ke koordinat alamnya). Untuk merekonstruksi pixel pada objek dapat dilakukan dengan cara terlebih dahulu mengubah citra ke dalam segmen-segmen (Manik, 2018, 2020a, 2020b) dan ada juga yang langsung menghitung disparity setiap pixel pada citra stereo. Konfigurasi Stereo Sederhana merupakan penelitian yang banyak dilakukan untuk menyelesaikan masalah tersebut. Dengan menggunakan dua kamera yang diletakkan dengan jarak tertentu dan sumbu fokus kamera sejajar, kita akan mendapatkan dua citra digital (citra digital kiri dan citra digital kanan). Jarak antara suatu pixel yang berkorespondensi antara citra digital kiri dan kanan disebut disparity. Pada penelitian sebelumnya kami telah dapat menentukan disparity setiap pixel dengan metode Belief Propagation, sehingga sistem koordinat dimensi tiga citra digital untuk satu frame telah dapat dihitung. Tetapi untuk mengenali objek bergerak dari frame ke frame berikutnya masih merupakan masalah pada Komputer Vision. Jadi masalah yang akan diselesaikan dalam penelitian

ini adalah bagaimana menemukan metode merekonstruksi lintasan objek bergerak pada citra digital video dengan konfigurasi stereo sederhana berdasarkan koordinat dimensi tiga dari setiap pixel-pixelnya dari suatu frame ke frame berikutnya.

Tujuan penelitian ini adalah membuat kode program yang mampu menentukan disparity video dari satu frame ke frame berikutnya. Sedangkan manfaat dari penelitian ini adalah sebagai masukan untuk pembuatan mobil autonavigasi di kemudian hari.



Gambar 1. Konfigurasi Stereo Sederhana

Konfigurasi Stereo Sederhana adalah konfigurasi dua kamera yang meniru cara penglihatan manusia dengan dua matanya, seperti diperlihatkan pada Gambar 1 (Florczyk, 2005). Misalkan C_l , C_r adalah kamera kiri dan kamera kanan secara berturut-turut yang diletakkan dengan jarak $2h$ dan fokus f . Jika P adalah sebuah titik pada koordinat alam yang proyeksinya pada citra digital kiri adalah pixel P_l dan proyeksinya pada citra digital kanan adalah pixel P_r , maka dengan operasi matematika sederhana kita dapat menghitung bahwa jarak titik P dari garis yang menghubungkan dua lensa kamera adalah

$$z = \frac{2hf}{P_r - P_l} \tag{1}$$

dimana disebut disparity. Jadi jika kita dapat menentukan proyeksi titik P pada citra digital kiri (P_l) dan pada citra digital kanan (P_r) maka kita dapat merekonstruksi koordinat dimensi tiga dari titik P , yaitu $P(x, y, z)$ dimana y bersesuaian dengan tinggi pixel pada citra digital. Jika dua kamera tersebut terhubung dengan komputer maka kita dapat membuat kode program yang dapat menentukan semua koordinat dimensi tiga dari objek di depan dua kamera berdasarkan disparity. Sehingga suatu saat nanti mobil autonavigasi akan dapat diciptakan, yaitu: mobil tanpa sopir atau mobil yang hanya dikendalikan komputer yang terhubung dengan dua kamera di depan.

Tabel 1. Metode Blok Kecocokan secara Metrik

Metrik Matematika	Definisi
Normalized Cross-Correlation (NCC)	$\frac{\sum_{u,v} (I_1(u,v) - \bar{I}_1) \cdot (I_2(u+d,v) - \bar{I}_2)}{\sqrt{\sum_{u,v} (I_1(u,v) - \bar{I}_1)^2 \cdot (I_2(u,v) - \bar{I}_2)^2}}$
Sum of Squared Differences (SSD)	$\sum_{u,v} (I_1(u,v) - I_2(u+d,v))^2$

Normalized SSD	$\sum_{u,v} \left(\frac{I_1(u,v) - \bar{I}_1}{\sqrt{\sum_{u,v} (I_1(u,v) - \bar{I}_1)^2}} - \frac{I_2(u+d,v) - \bar{I}_2}{\sqrt{\sum_{u,v} (I_2(u,v) - \bar{I}_2)^2}} \right)^2$
Sum of Absolute Differences (SAD)	$\sum_{u,v} I_1(u,v) - I_2(u+d,v) $

Rekonstruksi Koordinat Dimensi Tiga dari Setiap Pixel Citra Digital pada Konfigurasi Stereo Sederhana dengan Metode Belief Propagation sangat baik karena sistem koordinat dimensi tiga citra digital dengan Konfigurasi Stereo telah dapat dihitung (Manik, 2013). Tetapi untuk mengenali objek bergerak dari frame ke frame berikutnya masih merupakan masalah pada Komputer Vision (Bhatti 2008).

Tabel 1 (Brown at al., 2003) merupakan metrik matematika yang sering digunakan untuk menentukan korespondensi pada stereo vision. Metode ini sering disebut metode kecocokan korespondensi secara lokal. NCC merupakan metode standar statistika yang digunakan untuk menentukan keserupaan dua objek dan SSD merupakan variasi dari NCC yang sering juga digunakan. Penormalan tersebut bertujuan untuk mengurangi kesensitifan metrik terhadap gangguan dan bias nilai citra digital. SSD disenangi karena perhitungannya mudah tetapi SAD adalah metrik yang paling populer untuk perhitungan yang paling mudah.

Hasil dari metode yang hanya mengandalkan metrik untuk menghitung disparity dari setiap pixel dari frame ke frame berikutnya masih kurang memuaskan. Penyebab pertama adalah bahwa tidak semua pixel yang tampak pada citra digital kiri akan tampak juga pada citra digital kanan. Pixel seperti itu disebut occluded pixel. Kita akan menaksir disparity pixel seperti ini melalui disparity pixel yang ada disekitarnya. Kedua, pixel pada citra digital di kiri mungkin mempunyai pasangan lebih dari satu pixel pada citra digital kanan secara lokal. Pada hal satu pixel di kiri seharusnya hanya berpasangan dengan satu pixel di kanan. Jika hal seperti ini terjadi maka kita akan memperbesar jari-jari lingkungannya sedemikian sehingga diperoleh hanya satu pixel di kanan yang menjadi pasangan paling cocok untuk pixel kiri tersebut. Ketiga, walaupun kedua citra digital diambil dalam waktu yang bersamaan, nilai-nilai pixel dari proyeksi titik yang sama belum tentu sama karena terjadi distorsi dan gangguan pada citra digital kiri atau kanan. Distorsi ini akan lebih besar lagi dari satu frame ke frame berikutnya karena pergerakan objek atau kamera.

Untuk mengatasi ketiga masalah di atas, Bhatti (2008) menggunakan metode: "Real Time Stereo Image Registration for Planar Structure and 3D Sensor Pose Estimation" dengan langkah-langkah: Segmentation, Image Registration, dan Tracking Schemes secara berturut-turut. Tetapi metode ini masih memiliki banyak kelemahan-kelemahan.

Dalam metode tersebut, Region of Interest (ROI) digunakan untuk mensegmentasi citra digital pada langkah Segmentation. ROI merupakan gabungan dari beberapa metode segmentasi yang sudah ada (Lombardi *et al.*, 2005; Jansen *et al.*, 2005; Guzman & Parra, 2007; Alvarez *et al.*, 2008). ROI mempunyai kelemahan karena hasil segmentasi bergantung pemberian nilai awal. Hal ini akan diperbaiki dalam usul penelitian ini dengan menggunakan metode Mean Shift untuk mensegmentasi yang tidak bergantung pada pemberian nilai awal (Manik, 2010).

Wang and Suter (2004) memperkenalkan metode pencocokan bidang yang robust dan menyebutnya metode Adaptive-Scale Residual Consensus (ASRC). Mereka telah menguji metode tersebut dan menyimpulkan bahwa metode ASRC lebih baik dari LMedS, RANSAC, MSAC, RESC, ALKS, dan ASSC. Metode ASRC tidak membutuhkan inliers threshold awal sehingga metode ini cocok digunakan pada komputer vision. Metode "Real Time Stereo Image Registration for Planar Structure and 3D Sensor Pose Estimation" menggunakan metode RANSAC. Sehingga kami yakin bahwa kami mampu mendapatkan metode yang lebih baik karena kami akan menggunakan metode pencocokan bidang yang robust ASRC yang lebih baik.

Dalam penelitian ini kami akan terlebih dahulu membuat kode program untuk menentukan koordinat dimensi tiga dari setiap frame dengan metode yang telah ditemukan dalam penelitian kami sebelumnya. Selanjutnya kode program akan mencari dimana kumpulan titik-titik yang berbeda dari frame yang satu ke frame berikutnya dan berapa jauh pergeseran dan arahnya. Pergeseran dan arah tersebut akan digunakan untuk memprediksi letak objek yang bergerak pada frame- frame berikutnya.

Dalam penelitian ini, kami juga akan membuat data yang memuat titik-titik koordinat sebenarnya dari setiap pixel yang terdapat dalam video dari frame ke frame dengan cara perhitungan langsung berdasarkan pengamatan dari tiga citra digital dari tiga handycam terpasang (Gambar 2). Data ini digunakan untuk menguji bahwa metode yang kami kembangkan lebih baik dari metode "Real Time Stereo Image Registration for Planar Structure and 3D Sensor Pose Estimation" oleh Bhatti (2008).



(a) Frame Pertama



(b) Frame Terakhir

Gambar 2. Video Penelitian

2. Metode Penelitian

Penelitian ini bertujuan untuk membuat algoritma dan kode program yang mampu merekonstruksi lintasan objek bergerak pada citra digital video dengan konfigurasi stereo sederhana berdasarkan koordinat dimensi tiga dari setiap pixel-pixelnya dari frame ke frame. Pertama, video citra digital objek bergerak yang diambil dengan Handycam Stereo Vision tidak bergerak telah diteliti. Sedangkan video citra digital yang diambil oleh Handycam Stereo Vision bergerak akan diteliti untuk bagian kedua.

Penelitian ini bertujuan untuk membuat algoritma dan kode program yang mampu merekonstruksi lintasan objek bergerak pada citra digital video dengan konfigurasi stereo sederhana berdasarkan koordinat dimensi tiga dari setiap pixel-pixelnya dari frame ke frame. Pertama, video citra digital objek bergerak yang diambil dengan Handycam Stereo Vision tidak bergerak telah diteliti. Sedangkan video citra digital yang diambil oleh Handycam Stereo Vision bergerak akan diteliti untuk bagian kedua.

Langkah pertama yang dilakukan dalam penelitian ini adalah membuat/ mempersiapkan video dengan konfigurasi stereo sederhana. Video perjalanan yang diambil dari dalam mobil dengan konfigurasi stereo sederhana. Selanjutnya, menghitung koordinat/ disparity sebenarnya dari semua frame video yang dibuat pada langkah sebelumnya. Cara melakukannya adalah dengan mengukur langsung semua disparity setiap pixel dengan bantuan Matlab.

Menghitung disparity/ koordinat dimensi tiga dari video dengan menggunakan kode program yang sudah dibuat. Akhirnya, menguji ketelitian kode program dengan cara membandingkan disparity yang diperoleh dengan menggunakan metode yang diperoleh pada penelitian ini dan membandingkannya dengan disparity sebenarnya.

3. Hasil Penelitian dan Pembahasan

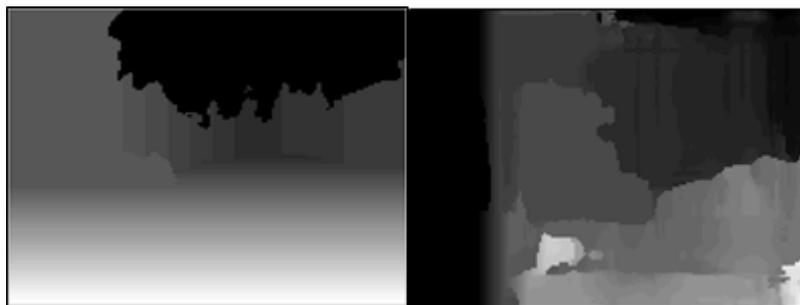
Video yang digunakan dalam penelitian ini adalah video tentang keadaan jalan raya di depan mobil. Ada sebanyak 40 frame dari video tersebut yang menjadi pusat perhatian. Video kiri dan kanan digabungkan menjadi satu yang pixelnya berukuran 120 baris kali 320 kolom. Gambar frame pertama dan frame terakhir dapat dilihat pada Gambar 2.

Dengan menggunakan bantuan rekaman dari dua handycam, disparity dari setiap permukaan objek yang ditangkap handycam dihitung pixel demi pixel. Data ini digunakan untuk membandingkan keakuratan metode yang dikembangkan dalam penelitian ini. Disparity sebenarnya dari frame terakhir dapat dilihat pada Gambar 3.



Gambar 3. Disparity Frame Terakhir

Kami menggunakan kode program yang dikembangkan untuk menghitung disparity setiap frame video. Disparity yang dihasilkan dari frame terakhir dapat dilihat pada Gambar 4.b. Sedangkan Gambar 4.a merupakan disparity sebenarnya dari frame terakhir yang dihitung langsung.



a. Disparity Sebenarnya

b. Disparity Penelitian

Gambar 4. Disparity Frame Terakhir Penelitian

Dengan membandingkan kesamaan warna Gambar 4a dan Gambar 4b, daerah sebelah kiri dari Gambar 4a tidak dapat dihitung disparity-nya karena pasangan pixel pada daerah ini tidak lengkap pada citra kiri atau kanan. Sedangkan disparity gambar kanan atas dapat dihitung oleh kode program dengan baik. Tetapi daerah lainnya mempunyai perbedaan warna yang berbeda-beda.

Kode program yang dikembangkan dalam penelitian ini masih membutuhkan penyempurnaan. Karena disparity antara gambar disparity sebenarnya dengan gambar disparity penelitian masih banyak perbedaan. Langkah pembuatan segmen-segmen pada citra kiri dan kanan sebelum menggunakan kode program penghitung disparity dapat dipertimbangkan untuk memperbaiki penghitungan disparity dari setiap pixel.

4. Kesimpulan

Disparity sebenarnya dan disparity yang dikembangkan dalam penelitian ini sangat banyak perbedaan. Sehingga model yang dikembangkan dalam penelitian ini masih perlu diteliti untuk mendapatkan metode yang lebih baik lagi.

5. Ucapan Terima Kasih

Penelitian ini dibiayai dengan Proyek Pengkajian Dan Penelitian Ilmu Pengetahuan Terapan. Untuk itu, ucapan terima kasih disampaikan kepada Direktorat Pembinaan Penelitian Dan Pengabdian Kepada Masyarakat Direktorat Jenderal Pendidikan Tinggi Kementerian Pendidikan Dan Kebudayaan.

6. Referensi

- Alvarez, J. M., A. López, and R. Baldrich. (2008). Illuminant-Invariant Model-Based Road Segmentation. In *IEEE Intelligent Vehicules Symposium*.
- Bhatti, A. (2008). *Computer Vision*. In Tech. Swedia.
- Brown, M.Z, D. Burschka, and G.D. Hager. (2003). Advances in Computational Stereo. *IEEE Patern Analysis and Machine Intelligence* 25(8): 993-1008.
- Florczyk, S. (2005). *Robot Vision: Video-based Indoor Exploration with Autonomous and Mobile Robots*. Wiley-VCH. Germany.
- Guzmán, A. and C. Parra. (2007). Extraction of Roads from Outdoor Images. In *Vision Systems: Applications*.
- Jansen, P., W. van der Mark, J.C. van der Heuvel, and F.C.A. Groen. (2005). Colour Based off-Road Environment and Terrain Type Classification. In *IEEE Intelligent Transportation Systems*.
- Lombardi, P., M. Zanin, and S. (2005). Messelodi. Switching Models for Vision-Based on-Board Road Detection. In *IEEE Intelligent Transportation Systems*.
- Manik, E. (2020a). Relationship between Segment Edges and Thresholds on Segmentation Generated by Minimum Spanning Trees. *Engineering Letters*, 28(3), pp.796-802
- Manik, E. (2020b). The outlier pixel and sub trees that form segments of image. IOP Conf. Series: *Materials Science and Engineering* 725 (012125) doi:10.1088/1757-899X/725/1/012125
- Manik, E. (2018). The minimum spanning tree in stereo vision. IOP Conf. Series: *Materials Science and Engineering* 420 (012115) doi:10.1088/1757-899X/420/1/012115
- Manik, E. (2013). Rekonstruksi Koordinat Dimensi Tiga dari Setiap Pixel Citra Digital pada Konfigurasi Stereo Sederhana dengan Metode Belief Propagation. *VISI*, 21(1), pp: 1272-1278.
- Manik, E. (2010). Pengaruh Bandwidth terhadap Segmentasi Citra Digital dengan Menggunakan Mean Shift. *Visi* 18(1): 43-49.
- Wang, W., and D. Suter. (2004). Robust Fitting by Adaptive-Scale Residual Consensus. *Proceedings of the 8th European Conference on Computer Vision, Prague*.