

Pelacakan Gerak Tangan dengan Metode Metode Pelacakan Objek Berbasis Korelasi

Naser Jawas, Ni Kadek Sumiari

Abstrak— Pelacakan tangan (hand tracking) saat ini tengah mendapat perhatian dari para peneliti di bidang visi komputer (computer vision). Tujuan utama dari pelacakan tangan adalah untuk mengetahui lokasi tangan di setiap deretan frame video atau deretan citra. Ada tiga buah fitur yang dapat digunakan untuk melacak objek yakni bentuk, warna, dan gerak. Fitur bentuk susah untuk digunakan melacak tangan, sedangkan fitur warna dan gerak lebih mudah dan lebih reliabel untuk melacak gerak tangan. Fitur warna dan gerak memiliki kelemahan-kelemahan masing-masing, sehingga pada penelitian sebelumnya diusulkanlah penggabungan kedua buah fitur tersebut untuk digunakan dalam melacak tangan secara bersama-sama agar dapat saling menutupi kelemahan masing-masing fitur. Namun melihat dari proses yang dilakukan pada metode yang diusulkan pada penelitian sebelumnya, peran yang lebih memiliki andil adalah peran dari fitur gerak. Pergerakan tangan yang cepat juga belum dapat ditangani dengan baik padahal dalam bahasa isyarat, gestur, dan aktivitas keseharian manusia, tidak jarang tangan bergerak dengan cepat. Oleh karena itu, pada penelitian ini diusulkan sebuah metode pelacakan tangan dengan menggunakan fitur gerak dan warna menggunakan metode pelacakan berbasis korelasi. Secara sederhana, metode berbasis korelasi ini akan melacak tangan di frame selanjutnya secara melingkar di sekitar lokasi sebelumnya dan setelah ditemukan dilakukan pelacakan menjauh dari lokasi awal ke arah yang ditemukan di pelacakan pertama. Fitur warna dan gerak akan digunakan secara seimbang dalam melacak arah pergerakan tangan. Diharapkan dengan menggunakan metode pelacakan yang diusulkan ini, baik gerak tangan yang cepat atau lambat dapat dilacak dengan baik.

Index Terms— Pelacakan gerak tangan, filter korelasi, object tracking.

I. PENDAHULUAN

Pelacakan tangan (hand tracking) saat ini tengah mendapat perhatian dari para peneliti di bidang visi komputer (computer vision). Hal ini dikarenakan pelacakan tangan menjadi salah satu bagian penting yang menjadi pondasi dasar dari beberapa aplikasi potensial seperti pengenalan bahasa isyarat (sign language recognition) (Koller, O. dkk., 2015), pengenalan gestur (geture recognition) (Maqueda, A.I. dkk., 2011; Premaratne, P., dkk., 2016), dan pengenalan

aktivitas manusia (human activity recognition) (Weinland, D., dkk., 2011). Bahkan, keberhasilan dari aplikasi-aplikasi potensial tersebut sangat tergantung dari keberhasilan dari pelacakan tangan (Weinland, D., dkk., 2011; Zhang, X., dkk., 2015). Tujuan utama dari pelacakan tangan adalah untuk mengetahui lokasi tangan di setiap deretan frame video atau deretan citra. Dengan didapatkannya lokasi-lokasi di setiap deret frame video, maka pergerakan tangan dapat diketahui.

Namun pelacakan tangan memiliki tantangan yang cukup banyak sebelum dapat dibawa ke aplikasi nyata. Permasalahan-permasalahan tersebut di antaranya adalah bentuk tangan yang sering berubah-ubah seiring dengan gerak tangan terutama pada bagian-bagian sendi, warna kulit tangan yang serupa dengan bagian tubuh lainnya seperti wajah, serta pergerakan tangan yang bebas mengakibatkan tangan dapat tersembunyi di balik objek lainnya (Yang, R., dkk., 2010). Ketiga permasalahan di atas masing-masing mengacu pada tiga fitur utama dalam pelacakan objek yakni, bentuk (shape), warna (color), dan gerak (motion). Menurut (Zhang, X., dkk., 2015), fitur bentuk tidak cocok digunakan untuk melacak tangan karena karakteristik tangan yang memiliki banyak persendian, fitur bentuk tangan menjadi fitur yang paling sulit dijadikan acuan dalam melakukan pelacakan tangan. Dua fitur lainnya, warna dan gerak, menjadi pilihan yang lebih reliabel dan mudah digunakan sebagai acuan dalam melakukan pelacakan tangan.

Fitur warna sederhananya dapat diambil dengan menggunakan histogram seperti pada penelitian (Comanicu, D., dkk., 2003). Fitur warna ini sangat baik digunakan dalam melakukan pelacakan objek ketika warna objek yang dilacak kontras dengan warna latar belakang atau warna objek-objek di sekeliling objek target. Namun apabila objek target bergerak melewati area latar belakang atau objek lain yang memiliki warna serupa, maka tingkat kesalahan pelacakannya meningkat (Zhang, X., dkk., 2015).

Fitur gerak paling umum digunakan adalah optical flow (Beauchemin, S.S. and Barron, J.L., 1995). Umumnya titik piksel yang digunakan untuk pelacakan optical flow dipilih menggunakan algoritma good features to track (Shi, J. and Tomasi, C., 1994) untuk memudahkan melakukan pemilihan titik piksel yang baik untuk dilacak. Algoritma good features to track mencari titik-titik yang berada di pinggiran objek atau bagian bertekstur dari objek yang dilacak. Karena tangan adalah objek yang tidak terlalu bertekstur, maka

Naser Jawas is with the Information System Departement of STMIK STIKOM Bali, Denpasar, Indonesia (corresponding author provide phone 0341-665144; email naser.jawas@gmail.com)

Ni Kadek Sumiari with the Information System Departement of STMIK STIKOM Bali, Denpasar, Indonesia.

algoritma *good features to track* tidak dapat menghasilkan cukup titik fitur yang dapat digunakan untuk proses pelacakan *optical flow*. Selain itu asumsi tingkat kecerahan yang konstan pada titik-titik yang dilacak oleh *optical flow* sering mengakibatkan kesalahan pelacakan tangan (Zhang, X., dkk., 2015).

Melihat dari kelemahan dari masing-masing fitur di atas, pada penelitian yang dilakukan oleh Zhang (2015) diusulkanlah penggabungan kedua buah fitur tersebut untuk digunakan dalam melacak tangan secara bersama-sama agar dapat saling menutupi kelemahan masing-masing fitur. Zhang (2015) sebuah metode pembangkitan fitur dan seleksi fitur dengan menggunakan 3 buah kriteria yakni perbedaan, kepadatan, perwakilan. Hasil dari penelitian ini menunjukkan bahwa penggabungan kedua buah fitur tersebut berhasil mengungguli metode pelacak yang menggunakan salah satu fitur saja. Namun pada metode yang diusulkan oleh Zhang tersebut, peran yang lebih memiliki andil adalah peran dari fitur gerak. Karena titik-titik fitur dari *optical flow* saja yang digunakan sebagai pelacakan. Fitur warna digunakan saat menentukan kotak (*bounding box*) yang akan digunakan untuk membangkitkan kembali fitur-fitur yang akan dilacak menggunakan *optical flow*. Kotak *bounding box* tersebut juga dihasilkan dari titik-titik pelacakan sebelumnya dari hasil *optical flow* yang telah mengalami proses reduksi. Penggunaan *optical flow* sebagai fitur yang lebih berbobot tersebut dapat mengakibatkan kesalahan dalam hasil pelacakan secara menyeluruh ketika titik-titik fitur hasil *optical flow* mengalami penyebaran yang sangat luas akibat dari pergerakan tangan yang cepat atau pergerakan latar belakang dengan warna serupa dengan tangan. Pergerakan tangan yang cepat ini juga tidak dapat ditangani dengan baik karena titik-titik fitur hasil *optical flow* akan menyebar dan akan terhapus dengan seleksi fitur yang diterapkan menggunakan kriteria kedekatan. Padahal dalam bahasa isyarat, gestur, dan aktivitas keseharian manusia, tidak jarang tangan bergerak dengan cepat.

Oleh karena itu, pada penelitian ini diusulkan sebuah metode pelacakan tangan dengan menggunakan fitur gerak dan warna menggunakan metode pelacakan berbasis korelasi. Metode berbasis korelasi ini akan melacak tangan di frame selanjutnya secara melingkar di sekitar lokasi sebelumnya dan setelah ditemukan dilakukan pelacakan menjauh dari lokasi awal ke arah yang ditemukan di pelacakan pertama. Fitur warna dan gerak akan digunakan secara seimbang dalam melacak arah pergerakan tangan. Diharapkan dengan menggunakan metode pelacakan yang diusulkan ini, baik gerak tangan yang cepat atau lambat dapat dilacak dengan baik.

II. METODE PENELITIAN

Pada penelitian ini metode pelacakan objek yang digunakan adalah adalah algoritma *Kernelized Correlation Filter* (Henriques, J.F., et al., 2015) dan *Discriminative Correlation Filter with Channel and*

Spatial Reliability (Lukežič, A., et al., 2018). Keduanya merupakan metode pelacakan gerak berbasis korelasi.

Tahapan pemrosesan adalah sebagai berikut setiap *sequence* citra diambil satu persatu. Kemudian, di layer pertama ditentukan *initial point* dari tangan yang akan di ikuti gerakannya. Kemudian metode korelasi bekerja untuk mencari bagian gerak di sekitar koordinat di frame sebelumnya pada frame selanjutnya.

Data uji yang digunakan dalam penelitian ini adalah data video berupa rekaman manusia yang menggerakkan tangannya secara bebas. Manusia yang direkam berada di depan kamera dengan latar belakang yang kompleks (tidak 1 warna). Contoh frame dari Data Video dapat dilihat pada Gambar 1.





Gambar 1. Contoh Data Frame Video (sumber: <http://www.votchallenge.net/>)

III. HASIL DAN PEMBAHASAN

A. Implementasi Algoritma Pelacak Gerak

Hasil yang dicapai pada tahapan ini telah mencakup ke pengembangan dan implementasi dasar-dasar algoritma pelacakan gerak kemudian mengujinya pada dataset gerak tangan yang umum digunakan. Implementasi algoritma yang telah diterapkan pada tahapan ini adalah algoritma Kernelized Correlation Filter (Henriques, J.F., et al., 2015) dan Discriminative Correlation Filter with Channel and Spatial Reliability (Lukežič, A., et al., 2018). Kedua algoritma ini adalah algoritma yang memanfaatkan filter korelasi.

1) Kernelized Correlation Filter

Algoritma Kernelized Correlation Filter (Henriques, J.F., et al., 2015) memiliki keunggulan proses yang cepat. Bahkan pada percobaan yang kami lakukan pemrosesan dapat mencapai lebih dari 100 frame per detik.

Namun kelemahan yang kami temui dari hasil implementasi algoritma ini adalah ketika tracker ini tidak berhasil mendapatkan area perpindahan di frame selanjutnya, tracker ini kehilangan petunjuk untuk mengikuti objek dan gagal menunjuk objek yang sedang diikuti. Gambar 2 menunjukkan hasil yang implementasi algoritma ini yang diuji menggunakan dataset citra standar.

2) Discriminative Correlation Filter with Channel and Spatial Reliability

Algoritma Discriminative Correlation Filter with Channel and Spatial Reliability (Lukežič, A., et al., 2018) menunjukkan hasil yang lebih baik dibanding algoritma Kernelized Correlation Filter. Pada pengujian menggunakan citra standar, algoritma ini tetap berhasil memberikan hasil keluaran pelacakan. Namun setelah frame ke-20 algoritma ini gagal melacak lokasi tangan. Hal ini dikarenakan tangan pada dataset tersebut bergerak melewati area wajah yang memiliki warna yang sama dengan area tangan. Hal tersebut membuat algoritma ini gagal dalam melacak gerak di frame selanjutnya.

Gambar 3 menunjukkan hasil yang dicapai oleh algoritma ini. Dapat dilihat pada Gambar 7(c) tangan melwati area wajah dan pada frame selanjutnya tracker tidak dapat menemukan lokasi tangan yang

sesungguhnya dan terjebak di area wajah. Namun, kecepatan proses dari algoritma ini masih kalah cepat dibandingkan algoritma Kernelized Correlation Filter. Algoritma ini memiliki rata-rata proses 35 frame per detik.



Gambar 2. Hasil Percobaan Algoritma Kernelized Correlation Filter:

(a) Frame 0



Gambar 3. Hasil Percobaan Algoritma Kernelized Correlation Filter:

(b) Frame 10



Gambar 4. Hasil Percobaan Algoritma Kernelized Correlation Filter:

(c) Frame 20



Gambar 5. Hasil Percobaan Algoritma Kernelized Correlation Filter: (d) Frame 30.



Gambar 9. Hasil Percobaan Algoritma Discriminative Correlation Filter with Channel and Spatial Reliability: (d) Frame 30.



Gambar 6. Hasil Percobaan Algoritma Discriminative Correlation Filter with Channel and Spatial Reliability: (a) Frame 0



Gambar 7. Hasil Percobaan Algoritma Discriminative Correlation Filter with Channel and Spatial Reliability: (b) Frame 10



Gambar 8. Hasil Percobaan Algoritma Discriminative Correlation Filter with Channel and Spatial Reliability: (a) Frame 0, (b) Frame 10, (c) Frame 20,

IV. KESIMPULAN

Pada tahapan ini didapatkan kesimpulan berupa algoritma terbaru yang menerapkan filter korelasi adalah Kernelized Correlation Filter (Henriques, J.F., et al., 2015) dan Discriminative Correlation Filter with Channel and Spatial Reliability (Lukežič, A., et al., 2018). Keduanya memiliki kelebihan dan kekurangan masing-masing. Dari segi akurasi pelacakan pemrosesan algoritma Discriminative Correlation Filter with Channel and Spatial Reliability unggul, namun dari segi kecepatan algoritma Kernelized Correlation Filter lebih unggul.

REFERENCES

- [1] Beauchemin, S.S. dan Barron, J.L. 1995. The computation of optical flow. *ACM Computing Surveys*. Vol. 27. No. 3. Hal. 433-466.
- [2] Chen, K., Tao, W., Han, S. 2017. Visual object tracking via enhanced structural correlation filter. *Information Sciences*. Vol. 394-395. No. 1. Hal. 232-245.
- [3] Comaniciu, D., Ramesh, V., Meer, P. 2003. Kernel-based object tracking. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. Vol. 25. No. 5. Hal. 564-577.
- [4] Henriques, J.F., Caseiro, R., Martins, P., Batista, J. 2015. High-Speed Tracking with Kernelized Correlation Filters. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. Vol. 37. No. 3. Hal. 583-596.
- [5] Hu, M., Liu, Z., Zhang, J., Zhang, G. 2017. Robust object tracking via multi-cue fusion. *Signal Processing*. Vol. 139. Hal. 86-95.
- [6] Koller, O., Forster, J., Ney, H. 2015. Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers. *Computer Vision and Image Understanding*. Vol. 141. No. 1. Hal. 108-125.
- [7] Lukežič, A., Vojří, T., Zajc, L.Č., Matas, J., Kristan, M. 2018. Discriminative Correlation Filter Tracker with Channel and Spatial Reliability. *International Journal of Computer Vision*. Vol. 126. No. 7. Hal. 671-688.
- [8] Maqueda, A.I., del-Blanco, C.R., Jaureguizar, F., Garcia, N. 2015. Human-computer interaction based on visual hand-gesture recognition using volumetric spatiograms of local binary patterns. *Computer Vision and Image Understanding*. Vol. 141. No. 1. Hal. 126-137.
- [9] Premaratne, P., Yang, S., Vial, P., Iftthikar, Z. 2016. Centroid tracking based dynamic hand gesture recognition using discrete hidden markov models. *Neurocomputing*. Vol. 228. No. 1. Hal.79-83.
- [10] Shi, J. and Tomasi, C. 1994. Good features to track. *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*. 21-23 Juni 1994, Seattle, USA. Hal. 593-600.

- [11] Weinland, D., Ronfard, R., Boyer, E. 2011. A survey of vision-based methods for action representation, segmentation and recognition. *Computer Vision and Image Understanding*. Vol.115. No. 1. Hal. 224-241.
- [12] Yang, R., Sarkar, S. 2010. Handling movement epenthesis and hand segmentation ambiguities in continuous sign language recognition using nested dynamic programming. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. Vol. 32. No. 2. Hal. 462-466.
- [13] Zhang, X., Li, W., Ye, X., Maybank, S. 2015. Robust hand tracking via novel multi-cue integration. *Neurocomputing*. Vol. 157. No. 1. Hal. 296-305.