

Tacking Pergerakan Tangan Menggunakan Skeleton Tracking Kinect

Fachri Yanuar Rudi F

Jurusan Teknologi Informasi dan Komputer Politeknik Negeri Lhokseumawe
Jln. B.Aceh Medan Km.280 Buketrata 24301 INDONESIA

fachri@pnl.ac.id

Abstrak—Gerakan tangan merupakan salah satu cara yang interaktif dalam melakukan interaksi manusia dan komputer. Dengan menggunakan gerakan tangan batasan dalam berinteraksi dengan komputer akan semakin kecil dan akan semakin menarik. Salah satu kendala dalam melakukan tracking gerakan tangan adalah sulitnya memisahkan tangan dengan latar belakang. Cara yang umum dilakukan untuk mengatasi hal tersebut adalah dengan mengatur latar belakang ataupun menggunakan algoritma *deep learning* untuk memisahkan tangan dengan latar belakang. Kinect kamera menyediakan fasilitas penangkapan kedalaman (*depth image*) yang mana hasil dari *depth image* tersebut dapat digunakan untuk melakukan *Skeleton tracking* pada manusia. *Skeleton tracking* tersebut dapat digunakan untuk memudahkan tracking pergerakan tangan.

Kata kunci— Gerakan tangan, kinect, skeleton tracking, *depth image*, background.

Abstract— Hand gesture is the interactive ways to interact with computer. Using hand gesture could reduce limitation in human computer interaction and make it a lot of attractive. One of the problem in tracking hand gesture is the difficulty of separating the hands from the background. A common way to deal with this is to set the background or use a deep learning algorithm to separate the hand from the background. Kinect camera provides a depth image capture facility where the results from the depth image can be used to perform Skeleton tracking on humans. Skeleton tracking can be used to make it easier to track hand movements.

Keywords— hand gesture, kinect, skeleton tracking, *depth image*, background.

I. PENDAHULUAN

Gerakan tangan atau biasa disebut dengan *Hand Gesture* merupakan salah satu cara interaktif dalam melakukan interaksi manusia dengan komputer (*Human Computer Interaction*) [1]. Salah satu cara dalam mengenali gerakan tangan adalah menggunakan kamera kemudian hasil tangkapan dari kamera tersebut akan diterjemahkan oleh komputer [2].

Langkah awal dari pengenalan gestur tangan adalah menentukan dimana letak tangan yang ditangkap oleh kamera atau disebut dengan *hand tracking*. *Hand tracking* telah diteliti sejak beberapa dekade yang lampau [3], [4]. Banyak penelitian yang telah dilakukan untuk mendeteksi lokasi tangan dan hal tersebut masih terbuka untuk dikembangkan lebih lanjut. Banyak juga alat bantu yang telah dikembangkan untuk membantu dalam hal *hand tracking* seperti sarung tangan berwarna [5], kamera khusus leap motion [6] dan sebagainya. Masing-masing metode memiliki kelebihan dan kekurangan tersendiri. Salah satu kekurangan alat bantu tersebut adalah interaksi antara manusia dan komputer dengan pengenalan gestur tangan menjadi kurang alami dan juga memiliki keterbatasan tertentu.

II. METODOLOGI PENELITIAN

Beberapa penelitian terdahulu telah melakukan deteksi posisi tangan dengan menggunakan masukan dari kamera baik kamera webcam ataupun kamera CCTV yang terhubung dengan komputer [7] maupun penggunaan data foto dari beberapa model posisi tangan [8]. Salah satu kesulitan dalam penggunaan kamera berwarna dalam mendeteksi posisi dan Gerakan tangan adalah memisahkan tangan dengan latar belakang terutama latar belakang yang memiliki warna yang sama atau hampir sama dengan kulit tangan manusia sehingga untuk mempermudah hal tersebut dilakukan pengondisian latar belakang dengan menggunakan latar berwarna putih ataupun

warna yang berbeda dengan kulit tangan seperti pada gambar 1. Cara yang lainnya adalah dengan menggunakan algoritma *deep learning* dalam mendeteksi posisi tangan [8]. Kelemahan dari metode pengondisian latar belakang adalah metode tersebut akan sulit untuk diterapkan di dalam kehidupan nyata mengingat akan banyak gangguan dan sulitnya untuk selalu melakukan penyesuaian latar belakang. Sedangkan penggunaan algoritma *deep learning* memiliki kekurangan pada proses komputasi yang besar sehingga akan membutuhkan *resource* computer yang besar dan berharga mahal.



Gambar 1. Pengondisian Latar Belakang Putih

Untuk mengatasi kedua permasalahan tersebut dapat digunakan kamera Kinect. Kamera Kinect merupakan kamera yang dikembangkan oleh Microsoft yang digunakan sebagai *depth sensor* pada konsol permainan Xbox. Cara kerja kamera Kinect adalah dengan memancarkan *array* dari inframerah kemudian akan mengukur jarak dari objek di depannya sehingga didapatkan bentuk tiga dimensi dari objek yang ada didepannya. Bentuk tiga dimensi tersebut berupa *width* (lebar), *height* (tinggi) dan *depth* (kedalaman atau jarak). Bentuk kamera Kinect dapat dilihat pada gambar 2. Hasil tangkapan dari kamera Kinect dapat dilihat pada gambar 3.

Hasil tangkapan *depth image* dari kamera kinect memiliki resolusi 640 x 480 piksel. Walaupun memiliki resolusi yang cukup kecil untuk standar masa kini, hasil tangkapan dari

kamera kinect dapat dipergunakan untuk melakukan tracking tangan.

Penggunaan kamera kinect dapat menjadi solusi untuk permasalahan dari kesulitan dalam mendeteksi posisi tangan. Dengan menggunakan data hasil tangkapan *depth image* kamera kinect akan diperoleh data jarak objek manusia yang berada di depan kamera yang kemudian data tersebut dapat diolah kembali menjadi data *skeleton*. Data *skeleton* dari kinect merupakan hasil pemrosesan *depth image* terhadap objek manusia yang kemudian dijadikan bentuk kerangka dari badan manusia. Bentuk kerangka tersebut dapat mewakili bagian-bagian tubuh manusia. Setiap bagian tubuh tersebut akan memiliki id tersendiri sehingga dengan adanya id tersebut dapat digunakan untuk melakukan tracking anggota tubuh salah satunya tracking tangan.



Gambar 2. Kamera Kinect

Data skeleton yang telah dikenali tersebut dapat digunakan untuk melakukan tracking tangan yang akan digunakan dalam pengenalan gestur tangan.



Gambar 3. Hasil *Depth Image* dari Kamera kinect

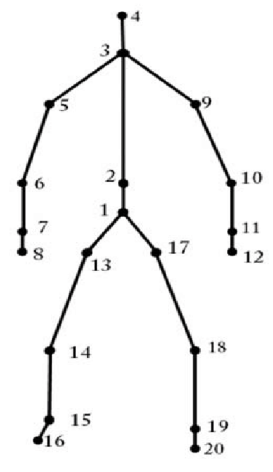
Gambar 4 menunjukkan hasil skeleton dari kinect beserta nomor dari *skeleton joint* pada setiap sendi tubuh. Untuk dapat melakukan tracking dari anggota tubuh hanya diperlukan nomor dari *skeleton joint* sehingga akan memudahkan proses tracking. Sebagai contoh jika ingin melakukan tracking tangan kanan maka dilakukan tracking terhadap *skeleton joint* nomor 12.

Untuk dapat melakukan tracking *skeleton* maka diperlukan minimal data *depth* dari orang yang ditangkap dari kepala sampai dengan pinggang. Jika posisi orang yang akan ditracking tidak dari kepala sampai dengan pinggang maka *tracking skeleton* akan gagal dilakukan.

III. HASIL DAN PEMBAHASAN

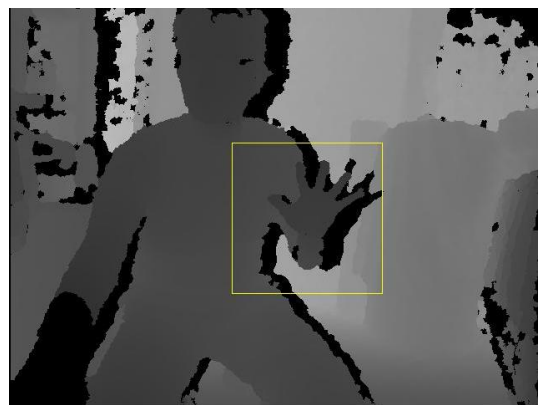
Hasil penelitian dari tracking pergerakan gestur tangan ini menunjukkan keberhasilan dari penggunaan kinect sebagai alat

bantu untuk melakukan *tracking* gesture tangan. Penggunaan kamera kinect untuk *tracking* gestur tangan mempermudah proses tracking dan juga dapat meringankan proses komputasi dan juga kesulitan yang dialami pada saat melakukan tracking gestur tangan.



Gambar 4. Skeleton pada kinect

Gambar 5(a) dan 5(b) menunjukkan hasil dari tracking tangan dengan menggunakan *skeleton joint* dari kamera kinect dan hasilnya diberi tanda kotak berwarna kuning.



(a)



(b)

Gambar 5. Gambar hasil tracking tangan menggunakan kinect skeleton

Agar proses *tracking* dapat berhasil maka posisi orang di depan kinect harus dapat ditangkap mulai dari kepala hingga paha. Jika bagian tubuh yang ditangkap oleh kamera kinect kurang dari kepala hingga paha maka akan gagal dalam proses *tracking skeleton* dari tubuh.

Hasil dari tracking pada gambar 5 kemudian dilakukan *cropping* atau pemotongan yang kemudian hasil pemotongannya dapat dilihat pada gambar 6 (a) sampai dengan 6 (d).



(a)



(b)



(c)



(d)

Gambar 6. Hasil cropping tracking gestur tangan

Setelah berhasil dilakukan *cropping* dari tangan yang telah ditacking maka dapat diproses ke langkah berikutnya berupa pengenalan gestur tangan. Dalam melakukan pengenalan gestur tangan nantinya akan diterapkan metode pengenalan gestur tangan. Baik berupa metode sederhana dengan mengubah citra hasil *cropping* menjadi grafik [9], menggabungkan data hasil tracking dengan gerakan tracking secara tiga dimensi [10], maupun hasil *cropping* dikenali menggunakan algoritma deep learning [11].

IV. KESIMPULAN

Penggunaan kamera kinect dalam melakukan tracking tangan dapat mempermudah dalam proses tracking tangan. Dengan menggunakan kamera kinect hasil tracking menjadi lebih tepat dan dapat mengurangi beban komputasi sehingga beban komputasi dapat dipergunakan untuk proses selanjutnya yang pada akhirnya akan didapat sistem yang dapat mengenali gestur tangan secara *real-time* atau minimal dapat mengenali gestur tangan secara cepat. Penggunaan kamera kinect dalam melakukan tracking gestur tangan juga dapat dikembangkan menjadi tracking gestur tangan secara tiga dimensi dikarenakan kamera kinect menyediakan data hasil tangkapan berupa data tiga dimensi. Pengolahan data tiga dimensi tersebut pada akhirnya dapat dikembangkan menjadi sistem pengenalan tangan yang lebih maju dan lebih intuitif dalam penggunaannya.

Salah satu kekurangan dari penggunaan kamera kinect adalah rendahnya resolusi hasil tangkapannya. Resolusi yang rendah dapat berakibat kepada hasil tangkapan akan banyak derau sehingga akan mengurangi sensitifitas dan keakuratan sistem pengenalan gestur tangan. Akan tetapi kekurangan tersebut dapat saja diakali dengan menggabungkan hasil tangkapan depth sensor dari kamera kinect yang digabungkan dengan kamera lainnya yang beresolusi lebih tinggi dengan pastinya proses penggabungan tersebut harus dikalibrasi terlebih dahulu agar proses penggabungannya dapat berjalan dengan baik atau dapat juga menggunakan kamera kinect V2 yang menyediakan resolusi yang lebih tinggi dan juga keakuratan yang lebih baik tapi dengan harga yang lebih tinggi daripada harga kinect xbox360. Kendala lainnya yang dihadapai oleh peneliti yang ingin mengembangkan tracking gestur tangan dengan menggunakan kamera kinect adalah ketersediaan kamera kinect yang sudah mulai langka dikarenakan Microsoft selaku pengembang kinect sudah tidak lagi mengembangkan dan menjual kinect.

REFERENSI

- [1] A. Haria, A. Subramanian, N. Asokkumar, S. Poddar, and J. S. Nayak, "Hand Gesture Recognition for Human Computer Interaction," *Procedia Comput. Sci.*, vol. 115, pp. 367–374, 2017, doi: 10.1016/j.procs.2017.09.092.
- [2] M. Panwar and P. Singh Mehra, "Hand gesture recognition for human computer interaction," in *2011 International Conference on Image Information Processing*, Nov. 2011, pp. 1–7, doi: 10.1109/ICIIP.2011.6108940.
- [3] J. M. Rehg and T. Kanade, "Visual tracking of high DOF articulated structures: An application to human hand tracking," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 801 LNCS, pp. 35–46, 1994, doi: 10.1007/bfb0028333.
- [4] B. Stenger, A. Thayananthan, P. H. S. Torr, and R. Cipolla, "Model-based hand tracking using a hierarchical Bayesian filter," *IEEE Trans Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1372–1384, 2006, doi: 10.1109/TPAMI.2006.189.
- [5] R. Y. Wang and J. Popović, "Real-time hand-tracking with a color glove," p. 1, 2009, doi: 10.1145/1576246.1531369.
- [6] O. Kainz and F. Jakab, "Approach to Hand Tracking and Gesture Recognition Based on Depth-Sensing Cameras and EMG Monitoring," *Acta Inform. Pragensia*, vol. 3, no. 1, pp. 104–112, 2014, doi: 10.18267/j.aip.38.
- [7] Z. Chen, J. Kim, J. Liang, J. Zhang, and Y. Yuan, "Real-Time Hand Gesture Recognition," *Int. J. Intell. Commun. Comput. Networks*, vol. 02, no. 02, 2021, doi: 10.51735/ijiccn/001/30.
- [8] A. Mujahid *et al.*, "Real-time hand gesture recognition based on deep learning YOLOv3 model," *Appl. Sci.*, vol. 11, no. 9, 2021, doi: 10.3390/app11094164.
- [9] F. Fachri Yanuar Rudi and E. M. Yuniarno, "Contour to Centroid Distance Graph as Feature in Hand Gesture Recognition," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 536, no. 1, 2019, doi: 10.1088/1757-899X/536/1/012150.
- [10] X. Luo, A. Amighetti, and D. Zhang, "A Human-Robot Interaction for a Mecanum Wheeled Mobile Robot with Real-Time 3D Two-Hand Gesture Recognition," *J. Phys. Conf. Ser.*, vol. 1267, no. 1, Jul 2019, doi: 10.1088/1742-6596/1267/1/012056.
- [11] E. I. Nikolaev, P. V. Dvoryaninov, Y. Y. Lensky, and N. S. Drozdovsky, "Using virtual data for training deep model for hand gesture recognition," *J. Phys. Conf. Ser.*, vol. 1015, no. 4, May 2018, doi: 10.1088/1742-6596/1015/4/042045.