

# Penjejakan Wajah Dengan Menggunakan *Improved Mean Shift Tracking*

<sup>1</sup> Sulfan Bagus Setyawan, <sup>2</sup> Hanum Arrosida

<sup>1,2</sup> Teknik Komputer Kontrol, Politeknik Negeri Madiun, Madiun  
<sup>1</sup> sulfan@pnm.ac.id, <sup>2</sup> hanumarrosida@pnm.ac.id

## Article Info

### Article history:

Received December 18<sup>th</sup>, 2019

Revised January 25<sup>th</sup>, 2020

Accepted February 28<sup>th</sup>, 2020

### Keyword:

Meanshift  
Penjejakan Wajah  
Oklusi

## ABSTRACT

Penjejakan wajah secara visual adalah suatu image processing untuk mengikuti pergerakan wajah yang menjadi target. Pada penelitian sebelumnya, telah dilakukan sistem penjejakan objek dengan menggunakan metode haar cascade, tetapi pada sistem penjejakan objek tersebut tidak tahan terhadap oklusi dan gangguan noise background. Maka pada paper ini, dibangun sistem penjejakan wajah yang tahan terhadap oklusi, background noise. Pada penelitian ini, wajah yang dijejakan akan dideteksi dengan menggunakan metode haarcascade viola jones, kemudian akan dibangun Algoritma Improved Mean Shift, yaitu dengan mengabungkan metode Mean Shift dan Corrected Background Weigthed Histogram. Mean Shift digunakan untuk mendeteksi koordinat pada target yang tahan terhadap gangguan oklusi. Metode Corrected Background Weighted Histogram ditambahkan pada metode Mean Shift untuk menghilangkan noise fitur background yang memiliki fitur yang dengan objek yang dijejakan. Hasil dari metode ini, target dapat tahan terhadap oklusi dan noise background. Improved Mean Shift tracking memiliki error rata-rata 3.87 pixel dengan standart deviasi 1.54 pixel

Copyright © 2020 Jurnal FORTECH.  
All rights reserved.

Abstrak—Visual face tracking is methods tracking that face as a target. In the previous research, an object tracking system was carried out using the haar cascade method, but in the object tracking system it was not resistant to occlusion and background noise interference. So in this paper, face tracking system that is resistant to occlusion and background noise is developed. In this paper, the faces traced will be detected using the haarcascade viola jones method, then an Improved Mean Shift Algorithm will be used, which is by combining the Mean Shift method and Corrected Background Weigthed Histogram. Mean Shift is used to detect coordinates on targets that are resistant to occlusion interference. The Corrected Background Weighted Histogram method is added to the Mean Shift method to eliminate noise background that have a same fitur with object target. The results of this methods, the target can be resistant to occlusion and background noise. Improved Mean Shift tracking has average error 3.87 pixel, with standart deviation 1.54 pixel

Keywords : meanshift, face tracking, occlusion

## I. Pendahuluan

Dengan kemajuan teknologi yang semakin pesat pada era modern ini, banyak sekali bermunculan sistem yang dapat mempermudah kehidupan dan pekerjaan manusia. Disiplin ilmu yang ada diantaranya adalah pengolahan citra.

Penjejakan suatu objek benda merupakan salah satu bidang pengolahan citra yang cukup berkembang dewasa ini. Salah satu aplikasinya dalam bidang keamanan (security system) seperti pengawasan lokasi (surveillance) dimana manusia sebagai objek target [1].

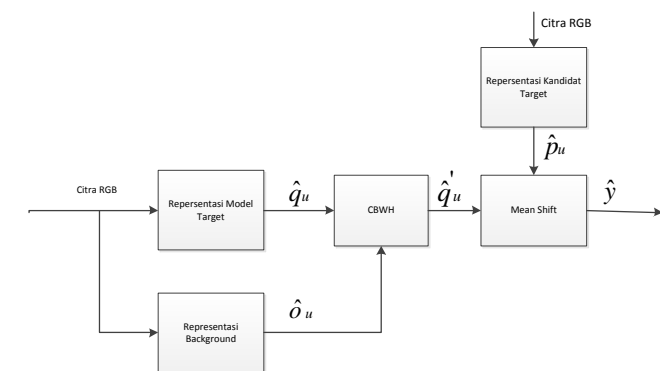
Berbagai metode yang banyak digunakan untuk sistem penjejakan objek. Zhiwen Chen dkk mengembangkan metode baru untuk menjejakan objek dengan menggunakan optical flow. Pada penelitian tersebut, optical flow digunakan untuk mendapatkan kontur dari objek. Langkah pertama yang digunakan untuk menjejakan objek adalah dengan menggunakan algoritma Horn-Schunk untuk mendapatkan kecepatan vector pada video [2]. Dari vektor tersebut maka akan diperoleh posisi dari piksel yang bergerak antara frame. Langkah berikutnya mendapatkan kontur objek dan mengekstaksi objek. Langkah ketiga menghitung posisi dan kecepatan dari objek yang bergerak. Dari hasil penelitian tersebut dapat diketahui bahwa algoritma tersebut dapat menjejakan objek. Tetapi pada algoritma ini rentan terhadap noise lingkungan. Xinghua Li dkk, melakukan pendeteksian dan penjejakan benda yang bergerak dengan menggunakan PTZ kamera. Untuk mendeteksi objek maka digunakan algoritma background subtraction untuk mendapatkan gambar biner setelah segmentasi gambar. Segmentasi gambar tersebut menghasilkan banyak noise sehingga digunakan metode expansion dan corrosion untuk menghilangkannya [3]. Kemudian kamera akan bergerak rotasi untuk mengikuti objek

tersebut dengan menggunakan algoritma timely replacement berdasarkan central region. Metode yang dibangun berhasil mendeteksi pergerakan objek, tetapi metode tersebut akan terjadi masalah ketika ada dua buah objek yang bergerak. Miaohui Zhang dkk melakukan penelitian tentang penjejak benda yang secara efektif terhadap benda 2-D affine motion pada benda [4].

Dari penelitian yang telah ada, maka dapat dikembangkan sistem penjejak wajah yang tahan terhadap noise background. Dibangun sistem penjejak wajah yang tahan terhadap oklusi, background noise. Pada paper ini, wajah yang dijejak akan dideteksi dengan menggunakan metode *haarcascade viola jones*, kemudian akan dibangun Algoritma *Improved Mean Shift*, yaitu dengan menggabungkan metode *Mean Shift* dan *Corrected Background Weighed Histogram*. *Mean Shift* digunakan untuk mendeteksi koordinat pada target yang tahan terhadap gangguan oklusi. Metode *Corrected Background Weighed Histogram* ditambahkan pada metode *Mean Shift* untuk menghilangkan noise. Hasil dari penelitian ini, target dapat tahan terhadap oklusi dan noise background. Pada paper ini bab II akan menjelaskan tentang metodologi penelitian, bab III akan menjelaskan tentang hasil dan pembahasan dan bab terakhir akan menjelaskan tentang kesimpulan

## II. Metodologi penelitian

Penjejukan objek yang baik merupakan penjejukan objek yang dapat tahan terhadap *noise background*, oklusi/ Pada paper ini, dikembangkan metode penjejukan wajah yang dapat tahan terhadap *noise background* dan oklusi. Metode dasar yang akan dikembangkan adalah algoritma *Mean Shift*. Algoritma *Mean Shift* dipilih karena metode ini efisien digunakan dalam penjejukan objek dan tahan terhadap oklusi sebagian (*partial occlusion*). Tetapi metode tersebut tidak tahan terhadap *noise background*. Pada paper ini, dikembangkan algoritma *Mean Shift* agar dapat mengurangi noise background dan tahan terhadap oklusi. Input dari metode *mean shift* ini adalah wajah. Wajah yang dijejak akan dideteksi dengan menggunakan metode *haarcascade viola jones*



Gambar 1. Diagram blok sistem penjejukan wajah

Algoritma Mean Shift adalah pendekatan yang efisien yang digunakan dalam penjejukan objek dimana digunakan histogram probabilitas distribusi warna sebagai fitur yang dijejak. Histogram distribusi probabilitas fitur warna pada target dijadikan sebagai representasi wajah target yang akan dijejak. *Noise background* pada objek tersebut direduksi dengan menggunakan metode CBWH (*Corrected Background Weighed Histogram*). Histogram distribusi probabilitas dari fitur warna target pada frame selanjutnya dijadikan sebagai representasi kandidat target. Lokasi target pada frame selanjutnya diperoleh dari perhitungan bobot dari representasi objek target dan representasi kandidat pada frame tersebut. Pada Gambar 1 merupakan diagram blok sistem penjejukan yang dirancang pada paper ini.

Representasi dari model target objek adalah histogram probabilitas warna dari wajah target yang terdeteksi. Dari wajah yang terdeteksi tersebut maka akan diperoleh nilai tinggi dan lebar dimana probabilitas warna berupa  $u = 1 \dots m$ , dapat didefinisikan pada Persamaan (1).

$$\hat{q}_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta(b(x_i^*) - u) \tag{1}$$

dimana  $m$  adalah jumlah warna yang digunakan. Warna yang digunakan pada penjejukan objek ini warna RGB (*Red, Green, Blue*) dengan masing-masing nilai dinormalisasi menjadi  $16 \times 16 \times 16$ . Sehingga untuk nilai  $u=1$  merupakan nilai dari  $R=1, G=1, B=1$ , untuk nilai  $u=2$  merupakan nilai dari  $R=1, G=1, B=2$ , dan seterusnya hingga  $u = 4096$  merupakan nilai dari  $R=16, G=16, B=16$ . Normalisasi ini bertujuan untuk mempercepat proses perhitungan untuk mendapatkan koordinat hasil penjejukan.  $\{x_i^*\}_{i=1 \dots n}$  adalah posisi koordinat piksel dari objek target. Posisi koordinat piksel ini berada pada persegi panjang yang dipilih oleh user.  $b(x_i^*)$  menunjukkan nilai RGB pada piksel yang terletak pada  $x_i$ .  $k(\|\cdot\|)$  merupakan fungsi kernel. Kernel yang digunakan adalah kernel Epanechnikov.  $\delta(a)$  adalah kroneker delta function. Maka dapat disimpulkan bahwa nilai probabilitas dari fitur warna dapat bernilai pada Persamaan (2).

$$\hat{q}_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \text{ jika } \delta(b(x_i^*)) = u \tag{2}$$

dimana  $C$  merupakan konstanta normalisasi yang didefinisikan pada Persamaan (3).

$$C = \sum_{i=1}^n k(\|x_i^*\|^2)^{-1} \tag{3}$$

Representasi *background* adalah histogram distribusi probabilitas dari warna *background* yang terdapat pada target. Ukuran *background* diperoleh dari 2 kali ukuran target yang

berupa wajah. Representasi *background* diperoleh dari histogram distribusi probabilitas fitur dari area *background*. Probabilitas dari warna  $u = 1 \dots m$ , dapat didefinisikan pada Persamaan (4).

$$\hat{\sigma}_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta(b(x_i^*) - u) \quad (4)$$

dimana  $m$  adalah jumlah fitur warna yang digunakan. Fitur warna yang digunakan pada penjejakan objek ini warna RGB (Red, Green, Blue) dengan masing-masing nilai dinormalisasi menjadi  $16 \times 16 \times 16$ . Sehingga untuk nilai  $u=1$  merupakan nilai dari  $R=1, G=1, B=1$ , untuk nilai  $u=2$  merupakan nilai dari  $R=1, G=1, B=2$ , dan seterusnya hingga  $u = 4096$  merupakan nilai dari  $R=16, G=16, B=16$ . Normalisasi ini bertujuan untuk mempercepat proses perhitungan untuk mendapatkan koordinat hasil penjejakan.  $\{x_i^*\}_{i=1 \dots n}$  adalah posisi koordinat pixel dari background.  $b(x_i^*)$  menunjukkan nilai RGB pada pixel yang terletak pada  $x_i$ .  $k(\|\cdot\|)$  merupakan fungsi kernel. Kernel yang digunakan adalah kernel Epanechnikov. Maka dapat disimpulkan bahwa nilai probabilitas fitur warna pada background dapat bernilai pada Persamaan (5).

$$\hat{\sigma}_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \quad \text{jika} \quad \delta(b(x_i^*)) = u \quad (5)$$

$C$  merupakan konstanta normalisasi yang didefinisikan pada Persamaan (6).

$$C = \sum_{i=1}^n k(\|x_i^*\|^2)^{-1} \quad (6)$$

*Corrected Background Weighted Histogram* adalah metode yang digunakan pada *Mean Shift* yang digunakan untuk meningkatkan akurasi lokasi pada penjejakan wajah. Ukuran *background* yang digunakan adalah 2 kali dari ukuran target objek. Nilai transformasi dibutuhkan mengetahui hubungan antara background dan model target yaitu berupa wajah. koefisien  $v_u$  pada Persamaan (7) merupakan nilai transformasi antara wajah dan background.

$$\left\{ V_u = \min \left( \frac{\hat{\sigma}_u^*}{\hat{\sigma}_u}, 1 \right) \right\}_{u=1 \dots m} \quad (7)$$

dimana nilai koefisien  $v_u$  dibutuhkan mereduksi nilai fitur *background* yang terdapat pada wajah sebagai target.  $\hat{\sigma}_u^*$  adalah nilai minimal non zero dari  $\{\hat{\sigma}_u\}_{u=1 \dots m}$ . Probabilitas dari warna yang baru  $u = 1 \dots m$ , dapat diturunkan pada Persamaan (8).

$$\hat{q}'_u = C' v_u \sum_{i=1}^n k(\|x_i^*\|^2) \delta(b(x_i^*) - u) \quad (8)$$

dimana  $m$  adalah jumlah fitur warna yang digunakan. Fitur warna yang digunakan pada penjejakan wajah ini warna RGB (Red, Green, Blue) dengan masing-masing nilai dinormalisasi menjadi  $16 \times 16 \times 16$ .  $\{x_i^*\}_{i=1 \dots n}$  adalah posisi koordinat pixel dari background.  $b(x_i^*)$  menunjukkan nilai RGB pada pixel yang terletak pada  $x_i$ .  $k(\|\cdot\|)$  merupakan fungsi kernel. Kernel yang digunakan adalah kernel Epanechnikov. Dimana nilai normalisasi nilai probabilitas warna pada Persamaan (9).

$$C' = \sum_{i=1}^n k(\|x_i^*\|^2)^{-1} \quad (9)$$

Representasi dari kandidat target wajah adalah histogram probabilitas objek target berupa wajah yang berada pada frame selanjutnya. Ukuran kandidat target berubah –ubah sesuai dengan ukuran target objek wajah, dimana ukuran awal kandidat target sama dengan ukuran target model. Probabilitas fitur kandidat target objek,  $u = 1 \dots m$ , dapat didefinisikan pada Persamaan (10).

$$\hat{p}_u = C_h \sum_{i=1}^n k(\|x_i^*\|^2) \delta(b(x_i^*) - u) \quad (10)$$

dimana  $m$  adalah jumlah fitur warna yang digunakan. Fitur warna yang digunakan pada penjejakan objek ini warna RGB (Red, Green, Blue) dengan masing-masing nilai dinormalisasi menjadi  $16 \times 16 \times 16$ .  $\{x_i^*\}_{i=1 \dots n}$  adalah posisi koordinat pixel dari kandidat target.  $b(x_i^*)$  menunjukkan nilai RGB pada pixel yang terletak pada  $x_i$ .  $k(\|\cdot\|)$  merupakan fungsi kernel. Lokasi dari koordinat yang baru  $y_1$  dapat ditemukan dengan mengabungkan representasi model target yang berupa wajah dan representasi kandidat target objek. Koordinat yang baru dapat didefinisikan pada Persamaan (11).

$$y_1 = \frac{\sum_{i=1}^n x_i \omega_i g(\|y_0 - x_i\|)^2}{\sum_{i=1}^{n_h} \omega_i g(\|y_0 - x_i\|)^2} \quad (11)$$

dimana  $\{x_i\}_{i=1 \dots n}$  adalah posisi koordinat pixel dari kandidat target.  $g(\|\cdot\|)$  merupakan fungsi kernel. Kernel yang digunakan adalah kernel Epanechnikov.  $\hat{y}_0$  merupakan posisi koordinat lama dari objek. Bobot pixel  $\omega_i$ , dapat didefinisikan dengan Persamaan (12).

$$\omega_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \delta(b(x_i^*) - u) \quad (12)$$

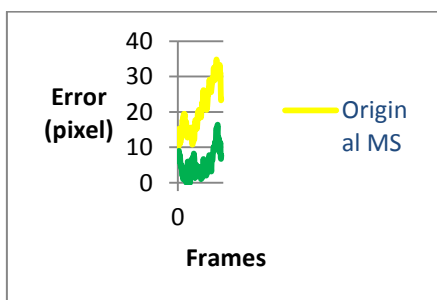
dimana  $m$  adalah jumlah fitur warna yang digunakan.  $b(x_i^*)$  menunjukkan nilai RGB pada pixel yang terletak pada  $x_i$ .  $u$  adalah warna yang digunakan.  $\hat{q}_u$  merupakan distribusi probabilitas model target,  $\hat{p}_u$  merupakan distribusi probabilitas kandidat target.  $\delta(a)$  adalah fungsi kroneker delta.

Algoritma dari penjejakan wajah visual dengan menggunakan *Improved Mean Shift* dapat disimpulkan sebagai berikut,

1. Mendeteksi wajah
2. menghitung probabilitas dari model target wajah
3. menghitung probabilitas dari *background* dengan
4. menghitung probabilitas dari model target yang baru
5.  $k=0$ ,
6. menghitung probabilitas dari kandidat target dengan
7. menghitung bobot piksel dengan menggunakan
8. menghitung posisi baru pada wilayah kandidat target
9. menghitung nilai error dan iterasi

### III. Hasil dan pembahasan

Pada bab III ini akan dibahas tentang hasil pengujian dan pembahasan. Pengujian Penjejakan objek ini bertujuan untuk menguji tingkat akurasi metode pada saat melakukan penjejakan objek. Pada pengujian ini, dibandingkan antara dua metode yaitu original *Mean Shift* dan *Improved Mean Shift*. Pengujian ini dilakukan dengan *me-load* video pengujian. Video pengujian yang digunakan ada dua jenis, yang pertama diuji dengan wajah yang memiliki fitur berbeda dengan background dan yang kedua pengujian dengan fitur wajah target yang memiliki fitur yang sama dengan background. Objek target yang digunakan adalah wajah, dengan background pada pengujian pertama berupa tembok berwarna putih, sedangkan pada pengujian yang kedua berupa tembok berwarna coklat yang memiliki fitur warna yang sama dengan objek yang dipilih (wajah). Pada pengujian ini, wajah dideteksi dengan menggunakan *viola jones* kemudian nilai koordinat per frame disimpan dalam format *\*txt*. Pada pengujian pertama terdapat 160 frame. Untuk mengetahui tingkat akurasi antara dua metode tersebut maka dapat diketahui dari error yang diperoleh dari penjejakan objek yang dilakukan.



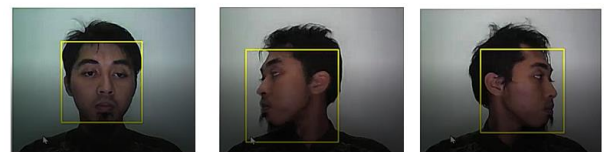
Gambar 2. Pengujian penjejakan

Pada Gambar 2 dapat diketahui nilai error rata – rata dengan menggunakan metode *Mean Shift* adalah 9.01 piksel dengan standard deviasi 5.77 piksel, sedangkan dengan menggunakan metode *Improved Mean Shift*.dapat diketahui bahwa nilai error rata – rata 3.87 piksel dengan standard deviasi 1.54 piksel. Dari error rata-rata tersebut dapat

diketahui bahwa nilai error dengan menggunakan *Improved Mean Shift*. lebih kecil dibandingkan dengan menggunakan metode *Mean Shift*. Penjejakan objek dengan menggunakan metode *Improved Mean Shift* lebih baik dibandingkan dengan menggunakan metode *Mean Shift*, disebabkan karena target model yang digunakan dalam metode *Improved Mean Shift*. tidak terpengaruh oleh *noise background*.



Gambar 3. Pengujian ketahanan ketika wajah dengan metode *viola jones*



Gambar 4. Pengujian ketahanan dengan menggunakan metode *Improved MeanShift* tracking

Pengujian ini bertujuan untuk mengetahui ketahanan objek pada saat penjejakan dari gangguan yang dapat membuat penjejakan objek hilang. Pengujian pertama dilakukan untuk mengetahui ketahanan metode penjejakan dari ketika objek target berotasi. Pengujian dilakukan mendeteksi wajah yang ditunjukkan dalam Gambar 3. Pada pengujian ini dilakukan dengan membandingkan metode tracking wajah secara konvensional yaitu dengan metode *viola jones*. Pada Gambar 3 menunjukkan bahwa hasil tracking yang dihasilkan ketika wajah menghadap ke kanan dan ke kiri sistem penjejakan wajah akan hilang. Kemudian sistem penjejakan wajah diuji dengan menggunakan metode *improved Mean Shift* Tracking. Hasil penjejakan wajah ini dapat dilihat pada Gambar 4. Pada Gambar 4 dapat diketahui bahwa wajah dapat dijejak dengan baik dengan menggunakan metode yang diusulkan walaupun wajah menghadap kekanan dan ke kiri.



Gambar 5. Pengujian ketahanan dengan oklusi dengan metode *viola jones*



Gambar 6. Pengujian oklusi dengan menggunakan metode *Improved MeanShift* tracking

Pengujian ini bertujuan untuk mengetahui ketahanan objek pada saat penjejakan dari gangguan yang dapat membuat penjejakan objek hilang yaitu oklusi. Oklusi sendiri merupakan penghalang terhadap wajah yang dijejak. Pengujian pertama dilakukan menggunakan metode konvensional sedangkan pada pengujian kedua digunakan metode *Improved MeanShift Tracking*. Pengujian dilakukan mendeteksi wajah yang ditunjukkan dalam Gambar 5. Pada pengujian ini dilakukan dengan membandingkan metode tracking wajah secara konvensional yaitu dengan metode *Viola Jones*. Pada Gambar 5 menunjukkan bahwa hasil tracking yang dihasilkan ketika wajah terkena oklusi, sistem penjejakan wajah akan hilang. Kemudian sistem penjejakan wajah diuji dengan menggunakan metode *Improved Mean Shift Tracking*. Hasil penjejakan wajah ini dapat dilihat pada Gambar 6. Pada Gambar 6 dapat diketahui bahwa wajah dapat dijejak dengan baik dengan menggunakan metode yang diusulkan meskipun wajah terkena oklusi.

#### IV. Kesimpulan

Setelah dilakukan pengujian terhadap sistem yang dibangun dengan menggunakan metoda-metoda yang telah ditentukan, dapat disimpulkan bahwa metoda *Improved Mean Shift Original Mean Shift tracking* memiliki error rata-rata 9.01 pixel dan standart deviasi 5.66 pixel. Sedangkan untuk *Improved Mean Shift tracking* memiliki error rata-rata 3.87 pixel dengan standart deviasi 1.54 pixel. Dari hasil pengujian yang dilakukan *Improved Mean Shift* tahan terhadap objek rotasi dan oklusi dibandingkan metode *Original Mean Shift*. Sehingga Metode *Improved Mean Shift* dapat digunakan untuk menjejak wajah dengan baik.

#### V. Daftar pustaka

- [1] Yilmaz, A. dkk (2006). "Object tracking: A survey", *ACM Comput. Surv.*2006. 38.
- [2] Zhiwen, Chen dkk. "Tracking of moving object based on optical flow detection" *International Conference on Computer Science and Network Technology*, Harbin,. hal 1096 - 1099.2011
- [3] Xinghua Li, dkk "Detection and Tracking of Moving Object Based on PTZ camera", 4th ICCCNT, Nanjing, hal 493-397 .2012
- [4] Zheng, Bin dkk. "Object Tracking Algorithm Based on Combination of Dynamic Template Matching and Kalman Filter", 4th International Conference on Intelligent Human-Machine Systems and Cybernetics. Nanchang, Jiangxi,. hal 136 – 139.2012