

MODIFIKASI METODE CAMSHIFT UNTUK PENGENALAN CITRA WAJAH SECARA *REAL TIME* BERDASARKAN WARNA KULIT WAJAH

Sultoni¹⁾

¹⁾Pendidikan Teknologi Informatika dan Komputer, Universitas Muhammadiyah Sidoarjo
Jl. Mojopahit No. 666 B Sidoarjo
Email : sulton@umsida.ac.id

Abstrak. Penelitian pengenalan wajah, saat ini sudah menjadi topik penelitian yang ditekuni oleh banyak peneliti baik dalam negeri maupun luar negeri karena banyak manfaat dan guna dari hasil penelitian ini. Beberapa manfaat dari penelitian pengenalan wajah antara lain untuk system keamanan dan robotika. Peneliti melakukan pengembangan dari penelitian sebelumnya, dimana penelitian sebelumnya dalam pengenalan wajah menggunakan metode camshift untuk deteksi dan tracking wajah, hasil dari deteksi dan tracking kemudian di capture dan digunakan sebagai data pelatihan. Dalam melakukan pengenalan wajah terdiri dari beberapa tahapan yaitu preprocessing, filtering, ekstraksi baru kemudian dilakukan pencocokan atau pengenalan citra wajah. Metode yang digunakan ini masih ada beberapa titik lemah yaitu metode camshaft akan mendeteksi dan men-tracking objek yang memiliki warna apabila objek tersebut mirip atau sama dengan warna citra wajah walaupun itu bukan wajah dan apabila ada kumpulan titik atau pixel yang memiliki warna mirip wajah juga akan dianggap wajah, walaupun sebenarnya itu adalah noise. Untuk itu, penelitian kali ini akan mencoba memodifikasi metode camshift dengan menambahkan metode dilatasi dengan tujuan untuk meminimalisir noise atau gangguan sehingga akan meningkatkan akurasi pengenalan. Hasil ujicoba terhadap 20 mahasiswa PTIK Umsida dimana masing masing mahasiswa di ambil 10 citra wajah sebagai data pelatihan menunjukkan adanya peningkatan pengenalan apabila dibandingkan sebelum adanya penambahan metode dilatasi.

Keyword : Citra wajah, Real Time, Pengenalan Wajah, Camshift, Akurasi, Dilatasi.

1. Pendahuluan

Beberapa penelitian pengenalan wajah yang pernah dilakukan antara lain adalah [1] melakukan penelitian pengenalan wajah secara *real time* menggunakan *eigenfaces*, dalam penelitiannya ini dilakukan melalui beberapa proses yaitu proses filtering menggunakan Gaussian untuk mengambil nilai-nilai eigen dengan menggunakan teknik *Eigenfaces*, berdasarkan hasil ujicoba menghasilkan nilai akurasi yang cukup baik, akan tetapi operator ini akan menyebabkan pelebaran tepi obyek sehingga batas – batas obyek wajah akan terlihat kurang jelas karena terdapat riak pada tepi tepi objek sehingga akan mengurangi keakuratan dalam melakukan pengenalan wajah. [2] melakukan penelitian pengenalan ekspresi wajah secara *real time* dengan menggunakan DCT2D dan *Neural Network*, berdasarkan hasil ujicoba menghasilkan akurasi yang cukup baik yaitu sekitar 89,18%, menurut [3] hasil ini masih bisa ditingkatkan dengan penambahan beberapa proses sebelum ekstraksi, yaitu dengan menambahkan proses *filtering* sebelum dilakukan ekstraksi menggunakan DCT2D, dari segi waktu komputasi juga kurang efisien, karena *Neural Network* akan selalu melakukan pembelajaran dari awal apabila ada data baru, sehingga semakin banyak data yang digunakan waktu komputasinya juga semakin lama. [3] melakukan penelitian pengenalan wajah secara *real time* menggunakan metode *Camshift* dan LoGDCT2D, dimana dalam penelitiannya ini untuk melakukan deteksi dan *tracking* wajah menggunakan metode *camshift*, berdasarkan hasil percobaan menghasilkan deteksi dan *tracking* wajah yang cukup baik, hasil pengenalannya yang menggunakan LoGDCT2D juga meningkat apabila dibandingkan dengan penelitian terdahulu dengan waktu komputasi yang tidak jauh berbeda dengan penelitian sebelumnya. namun penelitian yang dilakukan oleh [3] ini juga masih memiliki beberapa

kekurangan, dimana *camshift* akan melakukan deteksi dan *tracking* wajah berdasarkan warna yang menyerupai wajah, dimana apabila ada *pixel* atau titik yang cukup banyak dan memiliki warna yang mirip wajah walaupun sebetulnya bukan wajah oleh *camshift* akan di deteksi sebagai wajah. Dengan melihat dari kekurangan metode *camshift* ini, maka peneliti mencoba melakukan modifikasi terhadap metode ini dengan harapan akan ada peningkatan akurasi, yaitu dengan menambahkan proses dilatasi dalam proses deteksi dan *tracking* wajah. Dilatasi adalah operasi morfologi yang akan menambahkan *pixel* pada batas antar objek dalam suatu citra digital. Operasi ini menggunakan aturan sebagai berikut: “Untuk gambar *grayscale* maka nilai hasil operasi (*output pixel*) adalah nilai maksimal yang diperoleh dari himpunan *pixel* tetangganya. Dalam *binary image*, jika ada *pixel* tetangga yang bernilai 1 maka *output pixel* akan diset menjadi 1” [4,5].

2. Pembahasan

Dalam makalah ini, pembahasan yang akan diuraikan adalah yang terkait dengan pengujian warna kulit wajah, pengujian hasil pengenalan, gambar hasil pengenalan dan rumus – rumus yang digunakan.

2.1. Tabel

Tabel – tabel yang disajikan dalam artikel ini adalah tabel yang terkait dengan pengujian, dalam penelitian ini terdapat beberapa kali pengujian, yang pertama adalah pengujian terhadap warna kulit citra wajah manusia, untuk menentukan nilai Hue yang merupakan warna yang bisa menyerupai warna aslinya. Berdasarkan hasil ujicoba terhadap 10 orang, didapatkan hasil sebagaimana tabel 1 di bawah.

Tabel. 1 Hasil Pengujian Warna Kulit Wajah Manusia

No	Warna Kulit	Nilai		
		Hue	Saturation	Value
1	Orang 1	0.09	0.39	0.77
2	Orang 2	0.05	0.35	0.77
3	Orang 3	0.04	0.20	0.71
4	Orang 4	0.11	0.29	0.75
5	Orang 5	0.05	0.13	0.44
6	Orang 6	0.09	0.13	0.62
7	Orang 7	0.08	0.30	0.73
8	Orang 8	0.08	0.55	0.52
9	Orang 9	0.09	0.58	0.28
10	Orang 10	0.07	0.43	0.64

Berdasarkan hasil ujicoba sebagaimana terlihat pada tabel 1 di atas, warna kulit manusia untuk nilai Hue berkisar antara 0,04 – 0,11 atau kalau dirubah kedalam bentuk derajat maka nilai Hue berkisar antara 14^0 – 40^0 dan nilai Saturation berkisar antara 0,13 – 0,58, atau dapat dibuat suatu *threshol* untuk warna kulit wajah manusia sebagai berikut:

$0 < H < 50$ dan $0,13 < S < 0,58$ -> Dengan $H = [0,360]$ dan $S = [0,1]$ hal ini senada dengan penelitian yang sudah dilakukan oleh [6] tentang identifikasi kulit manusia.

Tabel 2 di bawah merupakan tabel hasil pengenalan citra wajah secara *real time* setelah dilakukan modifikasi terhadap metode *camshift* yaitu dengan menambahkan proses dilatasi di dalamnya.

Tabel 2. Hasil Pengenalan Citra Wajah

No	Nama Orang	Hasil Pengenalan (%)	Waktu Pengenalan (detik)
1	Orang ke 1	91	1.0076
2	Orang ke 2	90	0.7891
3	Orang ke 3	95	1.0005
4	Orang ke 4	96	1.0309

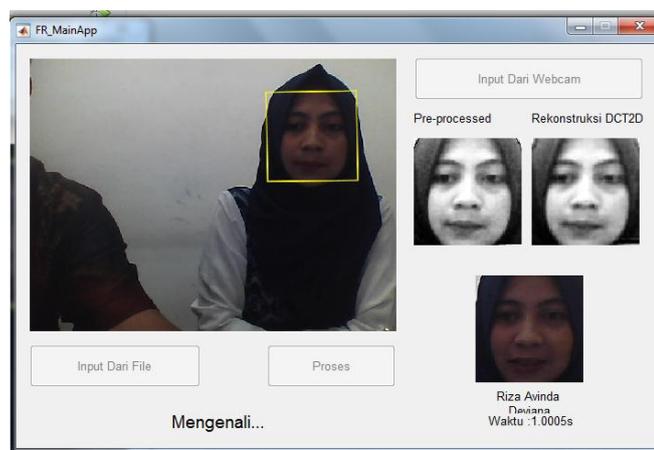
No	Nama Orang	Hasil Pengenalan (%)	Waktu Pengenalan (detik)
5	Orang ke 5	97	1.0657
6	Orang ke 6	97	0.5089
7	Orang ke 7	90	0.9761
8	Orang ke 8	91	1.0903
9	Orang ke 9	95	1.0802
10	Orang ke 10	96	1.0023
11	Orang ke 11	97	0.7681
12	Orang ke 12	96	1.0901
13	Orang ke 13	96	1.0045
14	Orang ke 14	97	0.9831
15	Orang ke 15	92	0.0891
16	Orang ke 16	95	1.0076
17	Orang ke 17	96	0.1008
18	Orang ke 18	95	1.0812
19	Orang ke 19	92	1.0091
20	Orang ke 20	92	1.0902
Rata – rata		94.05	0.89

Berdasarkan data yang ada pada tabel 2 di atas, maka dapat kita simpulkan bahwa akurasi pengenalan rata-rata setelah dilakukan modifikasi terhadap metode *camshift* menunjukkan adanya peningkatan apabila dibandingkan dengan penelitian yang sebelumnya. Dimana penelitian yang dilakukan oleh [3] memiliki akurasi rata-rata sekitar 93,91% sedangkan setelah dilakukan modifikasi meningkat menjadi 94,05% artinya ada peningkatan akurasi pengenalan sekitar 0,14% dari penelitian sebelumnya.

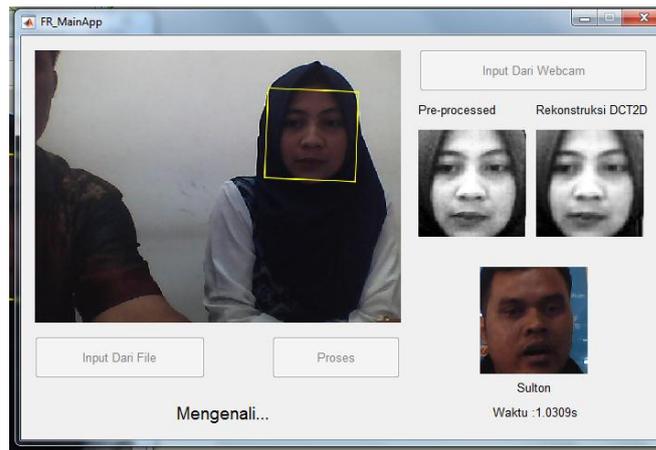
Di tinjau dari waktu komputasi masih lebih cepat setelah dilakukan modifikasi terhadap metode Camshift, apabila sebelumnya waktu komputasi yang diperlukan rata – rata sekitar 2,14 detik. Setelah dilakukan modifikasi waktu komputasi meningkat menjadi 0,89 detik, ini dikarenakan dalam penelitian ini tidak menggunakan proses filtering dengan menggunakan *Laplacian of Gaussian (LoG)* sedangkan penelitian sebelumnya terdapat proses filtering menggunakan LoG

2.2. Gambar Dan Keterangan Gambar

Gambar – gambar yang ditampilkan ini merupakan gambar terkait dengan pengenalan wajah yaitu gambar pengenalan wajah yang benar dan gambar pengenalan wajah yang tidak benar:



Gambar 1. Hasil pengenalan yang benar



Gambar 2. Hasil Pengenalan yang salah

2.3. Persamaan

Persamaan yang digunakan dalam artikel ini adalah yang terkait dengan dasar teori yaitu tentang teori Citra, preprocessing, camshift, Dilatasi dan Discrete Cosine Transform 2 Dimension.

a. Citra

Citra adalah suatu representasi (gambaran), kemiripan, atau imitasi dari suatu obyek. Citra sebagai keluaran suatu sistem perekaman data dapat bersifat optik berupa foto, bersifat analog berupa sinyal-sinyal video seperti gambar pada monitor televisi, atau bersifat digital yang dapat langsung disimpan pada suatu media penyimpanan [7]

Sedangkan citra digital pada umumnya berbentuk empat persegi panjang dengan dimensi ukurannya dinyatakan sebagai lebar x tinggi y [8]. Citra digital memiliki koordinat spasial, dengan tingkat kecerahan atau intensitas cahaya (skala keabu-abuan) yang memiliki numeric yang diskrit dipresentasikan dalam bentuk fungsi matematis $f(x,y)$ yang menyatakan intensitas cahaya pada titik (x,y) itu sendiri

b. Preprocessing

Preprocessing adalah tahap sebelum tahap inti yang bertujuan untuk mempermudah analisis berikutnya, dalam penelitian ini beberapa preprocessing yang dilakukan adalah *grayscale*, *resize* dan *cropping*.

c. Camshift

Deteksi dan penajakan (tracking) wajah dalam penelitian ini menggunakan metode *Camshift* yang merupakan singkatan dari *Continously Adaptive Mean Shift* dan hasil pengembangan dari algoritma Mean Shift yang dilakukan secara terus menerus untuk melakukan adaptasi atau penyesuaian terhadap distribusi probabilitas warna yang selalu berubah tiap pergantian frame dari video sequence .

Langkah – langkah dari algoritma mean shift adalah sebagai berikut [9]:

- Langkah 1 : tentukan ukuran search window
- Langkah 2 : tentukan lokasi awal search window
- Langkah 3 : hitung daerah mean dalam search window
- Langkah 4 : posisikan search window ke tengah daerah mean seperti dihitung pada langkah 3
- Langkah 5 : ulangi langkah 3 dan 4 hingga konvergen (atau hingga pergeseran daerah mean kurang dari treshold / batas yang ditentukan)

Sedangkan langkah-langkah algoritma CamShift adalah sebagai berikut:

- Langkah 1 : tentukan ukuran awal search window
- Langkah 2 : tentukan lokasi awal dari search window
- Langkah 3 : tentukan daerah kalkulasi (calculation region) pada bagian tengah search window dengan ukuran lebih besar dari search window

- Langkah 4 : frame citra video dikonversi ke dalam sistem warna HSV (hue, saturation,value) kemudian langkah selanjutnya adalah membuat histogram citra untuk mengetahui distribusi probabilitas warna.
- Langkah 5 : lakukan algoritma meanshift seperti diatas (satu atau banyak iterasi) dengan input berupa ukuran dan lokasi search window serta citra distribusi probabilitas warna dan simpan zeroth moment
- Langkah 6 : set nilai x,y,z yang diperoleh dari langkah 5
- Langkah 7 : nilai x dan y dipakai untuk menentukan titik tengah search window, sedangkan (2 X area^{1/2}) untuk menentukan ukuran search window
- Langkah 8 : ulangi langkah 3 untuk setiap pergantian frame citra video.

Pada algoritma meanshift, untuk citra berdistribusi probabilitas warna, daerah mean (*centroid*) didalam *search window* bisa dicari dengan persamaan:

Zeroth moment bisa dicari dengan persamaan :

$$M00 = \sum_{xy} I(x,y) \tag{1}$$

First moment untuk x dan y bisa dicari dengan persamaan

$$M01 = \sum_{xy} I(x) \tag{2}$$

$$M10 = \sum_{xy} I(y) \tag{3}$$

Maka lokasi mean dalam *search window (centroid)* adalah

$$Xc = M10/M00 \tag{4}$$

$$Yc = M01/M00 \tag{5}$$

Dimana $I(x,y)$ adalah nilai warna pixel di posisi (x,y) pada citra dan x,y berada dalam search window dan M10 adalah *first moment* untuk x, M01 adalah *first moment* untuk y sedangkan M00 adalah *zeroth moment*.

Sedangkan untuk algoritma Meanshift, untuk citra berdistribusi probabilitas warna, daerah mean (*centroid*) didalam search window bisa dicari dengan persamaan [10]:

$$M20 = \sum_x \sum_y x^2 X I(x,y) \tag{6}$$

$$M02 = \sum_x \sum_y y^2 X I(x,y) \tag{7}$$

$$M11 = \sum_x \sum_y x X y X I(x,y) \tag{8}$$

Orientasi utama dari sumbu x,y dan skala distribusi adalah hal yang menentukan dalam penemuan search windows memiliki hasil pengukuran yang sama dengan pengukuran dari probabilitas distribusi gambar 2D. Orientasi dapat ditentukan dengan menggunakan rumus sebagai berikut:

$$\theta = \frac{\arctan\left(\frac{2 \cdot \frac{M11}{M00} - xc \cdot yc}{\left(\frac{M20}{M00} - xc^2\right) - \left(\frac{M02}{M00} - yc^2\right)}\right)}{2} \tag{9}$$

Yang pertama, nilai *eigenvalue* (panjang dan lebar dari distribusi probabilitas warna) dapat dihitung dengan menggunakan rumus sebagai berikut:

$$a = \frac{M20}{M00} - xc^2 \tag{10}$$

$$b = \frac{M11}{M00} - xc X yc \tag{11}$$

$$c = \frac{M02}{M00} - yc^2 \tag{12}$$

Untuk mencari nilai pusat distribusi warna antara panjang dan lebar dari search windows dapat menggunakan rumus sebagai beriku;

$$l = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}} \tag{13}$$

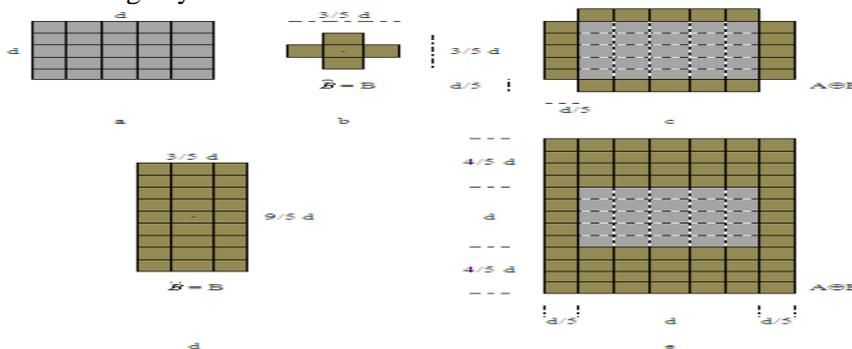
$$w = \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}} \tag{14}$$

d. Dilatasi

- a. Proses dalam dilasi adalah “penumbuhan” atau “penebalan” dalam citra biner.
- b. Jika A dan B adalah anggota Z^2 , dilasi antara A dan B dinyatakan $A \oplus B$ dan didefinisikan dengan:[5]

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \phi\}$$

- o Persamaan ini didasarkan pada perefleksian B terhadap originnya, dan penggeseran refleksi oleh z.
- o Dilasi A oleh B adalah himpunan semua displacement z, sebagaimana B dan A overlap oleh paling sedikit satu elemen.
- c. Dilasi ini sangat berguna ketika diterapkan dalam obyek-obyek yang terputus dikarenakan hasil pengambilan citra yang terganggu oleh noise, kerusakan obyek fisik yang dijadikan citra digital, atau disebabkan resolusi yang jelek,
 - o misalnya teks pada kertas yang sudah agak rusak sehingga bentuk hurufnya terputus, dan sebagainya



(a) Himpunan obyek; (b) Strel +; (c) Dilasi a oleh b; (d) Strel vertikal; (e) Dilasi a oleh d

e. Discrete Cosine Transform 2 Dimensi

DCT-2D merupakan pengembangan dari DCT-1D [11,12] bentuk persamaan dari discrete cosine transform dua dimensi adalah sebagai berikut:

$$c(u, v) = \frac{2}{\sqrt{MN}} a(u)a(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \cos \left[\frac{\pi(2x+1)u}{2N} \right] \cos \left[\frac{\pi(2y+1)v}{2M} \right] \dots\dots\dots(15)$$

Dimana $u = 0, 1, 2, \dots, N-1$ dan $v = 0, 1, 2, \dots, M-1$

Invers DCT2D dapat dituliskan dalam bentuk persamaan:

$$f(x, y) = \frac{2}{\sqrt{MN}} a(u)a(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} c(u, v) \cos \left[\frac{\pi(2x+1)u}{2N} \right] \cos \left[\frac{\pi(2y+1)v}{2M} \right] \dots\dots\dots(16)$$

3. Simpulan

Berdasarkan hasil pengujian yang dilakukan terhadap 20 mahasiswa PTIK Fakultas Keguruan dan Ilmu Pendidikan Universitas Muhammadiyah Sidoarjo terkait modifikasi Camshift untuk pengenalan wajah secara real time, maka dapat disimpulkan sebagai berikut:

1. Hasil pengenalan wajah dengan melakukan penambahan metode Dilatasi dapat meningkatkan akurasi pengenalan dibandingkan dengan penelitian yang sebelumnya, walaupun pengaruhnya tidak begitu besar yaitu hanya sekitar 0.14% saja
2. Dari segi waktu komputasi juga tidak jauh berbeda dengan penelitian sebelumnya, ini mengindikasikan dengan menambahkan metode dilatasi dapat meningkatkan akurasi pengenalan dengan waktu komputasi yang tetap stabil.

UcapanTerimaKasih

Melalui makalah ini, peneliti mengucapkan terimakasih yang sebesar – besarnya kepada mahasiswa PTIK FKIP Umsida yang sudah membantu dalam penelitian ini, dan kepada pihak universitas yang telah mengiizinkan penggunaan laboratorium untuk melakukan penelitian dan juga bantuan pembiayaannya.

DaftarPustaka

- [1]. Georgescu Daniel , “A real-Time Face Recognition System Using Eigenfaces”, Journal of mobile, Embedded and Distributed System. Vol.III.No.4.2011.
- [2]. Wu Shiqian, et.al, “High Speed Face recognition Based on Discrete Cosine Transform and RBF Neural Networks”, IEEE Transaction On Neural Networks, Vol.16 No.13. May 2005.
- [3]. Sultoni, Dachlan H S, Mudjirahardjo M, Dan Rahmadwati. “Pengenalan Wajah Secara Real Time Menggunakan Metode Camshift, LoG dan DCT2D (LoGDCT2D)” , Nero.Vol.2.No.3.2016
- [4]. Wirayuda BAT, ”Pemanfaatan Operasi Morphologi Untuk Proses Pendeteksian Sisi Pada Proses Pengolahan Citra Digital”, Seminar Nasional dan Informatika. Bali 2006
- [5]. Sutoyo T, Mulyanto E, Suhartono VNurhayati O D, dan Wijanarto. “**Teori Pengolahan Citra Digital**”, Andi, Yogyakarta. 2009
- [6]. Yang Ming H, David J. K. “*Detecting Faces in Images: A Survey*“, Senior Member, IEEE, and Narendra Ahuja, Fellow, IEEE
- [7]. Purnomo, M.H., dan Muntasa, A., 2010. *Konsep Pengolahan Citra Digital dan Ekstraksi Fitur*, Graha Ilmu, Yogyakarta
- [8]. Dimas W.W., “Perhitungan Jumlah dan Jenis Kendaraan Menggunakan Metode Fuzzy C-means dan segmentasi Deteksi Tepi Canny”, Malang, 2013
- [9]. Bradski and T. Ogiuchi, and M Higashikubo, “Visual Tracking Algorithm Using Pixel-Pair Feature”, International Confrence on Pattern Recognition, No 4, pp. 1808 – 1811. 2010
- [10]. Jammoussi Y A and Salhi Afef. “ Object Tracking Using Camshift, Meanshift and Kalman ilter”. International Journal of Vol. 6 No. 4, pp. 598 – 603. 2012
- [11]. R.C. Gonzalez and R.E. Woods, *Digital Image Processing, Prentice Hall*, Upper Saddle River, NJ. 2002
- [12].Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, *Digital Image Processing Using Matlab*