

Deteksi Ekspresi Wajah Menggunakan Fitur Gabor dan Haar Wavelet

Claudia Primasiwi, Handayani Tjandrasa, Dini A. Navastara
Departemen Informatika, FTIK, Institut Teknologi Sepuluh Nopember (ITS)
Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

Abstrak—Manusia bersosialisasi secara verbal dan non-verbal. Salah satu bahasa non-verbal yang sering digunakan manusia untuk berinteraksi adalah ekspresi wajah.

Pada studi ini, pengenalan ekspresi wajah manusia dilakukan dalam beberapa tahapan. Sebelum mendeteksi ekspresi, dilakukan pendeteksian wajah menggunakan algoritma Viola-Jones. Perbaikan kontras dilakukan menggunakan *histogram equalization*. Selanjutnya dilakukan perhitungan terhadap fitur Gabor dan fitur Haar. Untuk melakukan klasifikasi, digunakan SVM *One vs. All*. Fitur Gabor digunakan karena invarian terhadap rotasi, penskalaan, dan translasi. Sedangkan fitur Haar digunakan karena komputasinya yang efisien dan efektif dalam merepresentasikan sinyal dalam dimensi rendah dan tetap mempertahankan energinya.

Pada penelitian ini, digunakan kombinasi fitur Gabor dan Haar untuk dibandingkan dan digabungkan dengan fitur Landmark yang telah disediakan di database. Kombinasi fitur Gabor dan Haar memiliki akurasi sebesar 92%, sedangkan kombinasi fitur Gabor dan Landmark memberikan akurasi tertinggi sebesar 94,8%, dengan kelas terbaik berupa ekspresi senang.

Kata Kunci— Ekspresi Wajah, Gabor, Haar Wavelet, Viola-Jones, SVM.

I. PENDAHULUAN

Manusia bersosialisasi secara verbal dan non-verbal. Salah satu bahasa non-verbal yang sering digunakan manusia untuk berinteraksi adalah ekspresi wajah. Pada studi [1] didapatkan bahwa ekspresi wajah dapat menyampaikan hingga 55% informasi dalam interaksi, dibandingkan dengan vokal dan verbal keseluruhan hanya dapat memberi informasi hingga sebesar 45%. Ekspresi wajah dapat digunakan untuk mengembangkan interaksi yang natural antara manusia dan komputer. Beberapa pengenalan ekspresi [2], [3], [4], [5] telah dikembangkan dalam penelitian terdahulu. Penelitian ini memfokuskan pada pengenalan ekspresi wajah dari citra digital yang telah diekstraksi beberapa fiturnya.

Pada penelitian ini digunakan metode Viola-Jones untuk mendeteksi area wajah, Gabor dan Haar Wavelet serta Landmark sebagai fitur yang diekstraksi, dan SVM untuk mengklasifikasi ekspresi wajah. Viola-Jones menggunakan integral image yang dapat memproses citra lebih cepat dibandingkan dengan menggunakan pixel.

Proses yang diterapkan dalam penelitian ini diharapkan dapat membantu pengenalan emosi dengan

pendekatan komputasi. Pada awal tahap, dilakukan *preprocessing* citra dengan *histogram equalization* untuk mendapatkan detail-detail wajah yang lebih jelas pada citra. Setelah dilakukan *pre-processing*, dideteksi area wajah menggunakan metode Viola-Jones. Deteksi area wajah dengan metode yang diusulkan oleh Viola-Jones terdiri dari empat langkah, yaitu ekstraksi fitur Haar-like, integral image, AdaBoost, dan *cascade classifier* [6]. Dari citra wajah yang didapatkan tersebut, sistem mengekstraksi fitur-fitur wajah yang selanjutnya akan dikomputasi untuk klasifikasi ekspresi. Pada penelitian ini, fitur-fitur yang akan digunakan adalah fitur Gabor dan Haar yang akan dibandingkan dan digabungkan dengan fitur Landmark yang telah didapatkan pada database [7][8]. Tahap selanjutnya adalah melakukan klasifikasi ekspresi wajah dengan output berupa prediksi ekspresi.

II. METODE PENELITIAN

A. Ekspresi Wajah

Ekspresi wajah dapat dinotasikan dengan menggunakan *Facial Action Coding System* (FACS) yang terdiri dari 46 *Action Unit* (AU) [8][9]. Pada database Cohn-Kanade Extension (CK+), intensitas AU dari sebuah citra diterjemahkan menjadi bahasa sehari-hari secara manual dengan paduan yang telah disediakan [10].

B. Histogram Equalization

Citra yang akan digunakan dalam penelitian ini tidak diketahui *exposure* yang dimiliki, apakah *overexposure* atau *underexposure* yang disebabkan oleh pencahayaan lingkungan saat pengambilan gambar. *Exposure* yang tidak tepat dapat menyebabkan hilangnya detail-detail pada gambar. Histogram equalization berfungsi untuk memperbaiki kontras citra.

C. Algoritma Viola-Jones

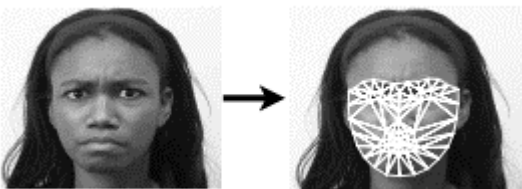
Untuk mendeteksi area wajah, sudah ada penelitian dari *Robust Real-Time Face Detection* [6] dari Viola dan Jones pada tahun 2004. Penelitian ini membahas tentang bagaimana caranya untuk mendeteksi wajah dan senyum dari suatu video secara *real time* menggunakan metode yang dirancang oleh Viola-Jones yang berisikan empat tahap, yaitu fitur Haar-like, *integral image*, AdaBoost, serta *Cascade*

classification.

D. Fitur Landmark

Titik Landmark adalah titik korespondensi pada setiap obyek yang sesuai antara dan di dalam populasi. Pada penelitian ini digunakan titik landmark yang telah diekstraksi pada penelitian sebelumnya menggunakan *Active Appearance Model* (AAM). Dari penelitian [7][8] tersebut, dihasilkan Landmark sebanyak 68 titik. Titik-titik tersebut disimpan ke dalam matriks berukuran 68x2 yang berisi koordinat x dan y dari titik-titik Landmark tersebut.

Titik landmark yang telah diekstraksi dari AAM disertakan dalam bentuk *wire-frame mesh* seperti terlihat pada Gambar . Titik landmark yang diekstraksi dari citra wajah akan digunakan sebagai fitur wajah.



Gambar 1. Visualisasi fitur Landmark dengan AAM [7][8]

E. Fitur Gabor

Fitur Gabor digunakan pada penelitian [2] dan [3]. Fitur Gabor didapatkan pada citra yang diproses menggunakan *Gabor Filter Bank*. Fitur Gabor memiliki keunggulan berupa invarian terhadap rotasi, skalasi, dan translasi [3].

Gabor Filter dapat dinotasikan seperti pada persamaan (1)

$$G(x, y) = \frac{f^2}{\pi\gamma\eta} \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp(j2\pi f x' + \phi) \quad (1)$$

untuk x' dan y' masing-masing dirumuskan pada persamaan (2) dan persamaan (3).

$$x' = x \cos \phi + y \sin \phi \quad (2)$$

$$y' = -x \sin \phi + y \cos \phi \quad (3)$$

dengan

f = frekuensi sinusoida

θ = orientasi dari *Gabor filter* terhadap *parallel stripes*

ϕ = phase offset

σ = standar deviasi dari *Gaussian envelope*

γ = rasio yang menentukan kelonjongan dari Gabor

Pada penelitian ini, fitur yang didapatkan berasal dari perhitungan Gabor dengan frekuensi yang digunakan adalah 5 dan orientasi yang digunakan adalah 8.

F. Fitur Haar

Pada penelitian [4] dan [5], fitur Haar diperoleh dari matriks coefficient. Pada penelitian tersebut fitur Haar digunakan karena kemampuannya melakukan komputasi secara efisien dan efektif pada dimensi rendah dengan tetap mempertahankan energi sinyalnya.

Dalam transformasi *Haar wavelet*, citra didekomposisi

menjadi beberapa komponen koefisien pada berbagai level. Komponen-komponen koefisien dapat direkonstruksi kembali menjadi citra seperti semula. Pada penelitian ini, fitur Haar didapatkan dari *coefficient matrix* level-2 dari proses dekomposisi *Haar wavelet*.

G. Principal Component Analysis (PCA)

PCA merupakan metode untuk membentuk variabel-variabel baru yang terdiri dari kombinasi linear variabel-variabel awal. Variabel-variabel yang diperoleh tidak saling berkorelasi dan jumlahnya dapat direduksi, dengan jumlah maksimum sama dengan jumlah variabel awal.

Langkah-langkah PCA [12] adalah:

1. Standardisasi data.
2. Menghitung *eigenvector* (*principal component*) dan *eigenvalue*. *Eigenvector* digunakan untuk menentukan arah dari fitur baru, sedangkan *eigenvalue* menentukan magnitude atau varian data dari fitur baru.
3. Memilih *principal component* atau komponen utama dengan cara mengurutkan *eigenpairs* dari nilai *eigenvalue* terbesar.
4. Menentukan banyaknya komponen utama dengan mempertahankan 99% nilai varians.
5. Proyeksi matriks ke *space* fitur baru.

H. Klasifikasi Support Vector Machine (SVM)

SVM merupakan salah satu klasifier yang banyak digunakan. SVM dapat digunakan dalam ruang dimensi tinggi dan mudah untuk dilatih. SVM telah terbukti memiliki kerja yang lebih baik saat digunakan dalam ruang yang padat atau pada ruang yang memiliki relevansi tinggi antar fitur-fiturnya [13]. SVM dapat digunakan untuk klasifikasi data dalam bentuk linear dan non-linear.

Multiclass mereferensikan klasifikasi yang terdiri dari lebih dari dua kelas. Pada penelitian ini diperlukan tujuh kelas (netral, sedih, terkejut, bahagia, takut, marah, jijik). Tipe klasifikasi *multiclass* yang akan diimplementasi adalah *One-vs-All* dengan kernel linear. Pada teknik ini, dibuat *N-binary classifier*. Untuk model klasifikasi ke- i , data positif merupakan seluruh data yang berada di kelas i , sedangkan data negatif adalah seluruh data yang tidak berada di kelas i .

Akurasi klasifikasi dihitung dengan menggunakan *k-fold cross validation* dengan k sama dengan 10. Penggunaan 10 *fold* ini dianjurkan karena merupakan jumlah *fold* terbaik untuk uji validitas [14]. Akurasi didefinisikan seperti pada persamaan (4).

$$\text{akurasi} = \frac{\text{jumlah klasifikasi benar}}{\text{jumlah data}} \times 100\% \quad (4)$$

III. HASIL DAN PEMBAHASAN

Secara umum, ekspresi yang paling banyak diprediksi benar adalah ekspresi senang dengan tingkat akurasi pada setiap mayoritas fitur sebanyak 100%. Sedangkan untuk

ekspresi yang banyak diprediksi salah adalah ekspresi merendahkan dan sedih yang paling banyak diprediksi sebagai ekspresi marah. Untuk perbandingan akurasi selengkapnya dapat dilihat pada Tabel 1.

Tabel 1
Tabel performa fitur

Fitur	Akurasi	Kelas Terbaik	Kelas Terburuk	Misklasifikasi Terbanyak
Landmark (L)	91%	Senang	Merendahkan	Marah
Gabor (G)	93,5%	Senang	Sedih	Marah
Haar (H)	84,3%	Senang	Merendahkan	Takut, Terkejut
GL	94,8%	Senang	Merendahkan	Marah
HL	89%	Senang	Merendahkan	Marah
GH	92%	Senang	Sedih	Marah
LGH	93,5%	Senang	Sedih	Marah

Dilihat dari jumlah data, ekspresi senang memiliki data yang cukup banyak, 69 citra atau 21% dari keseluruhan data. Sedangkan untuk ekspresi merendahkan memiliki jumlah data yang relatif sedikit dibandingkan dengan ekspresi lainnya. Data ekspresi merendahkan hanya 5,8% dari total keseluruhan data yaitu 18 citra.

Jika dilihat dari Gambar 2, kedua citra memiliki persamaan pada bagian bibir yang terkatup dan bagian alis dalam yang naik tetapi keduanya memiliki intensitas yang berbeda.

Untuk fitur yang memiliki akurasi terbaik adalah fitur Gabor dan Landmark dengan akurasi sebesar 94,8%. Sedangkan untuk fitur yang memiliki akurasi terendah adalah fitur Haar dengan akurasi sebesar 70%.



Gambar 2. Citra ekspresi merendahkan (a) dan sedih (b) [7][8]

IV. KESIMPULAN

Setelah dilakukan perancangan, pengamatan dan pengujian, dari penelitian ini dapat disimpulkan bahwa sistem dapat mendeteksi area wajah dengan sempurna pada seluruh citra. Selain itu, sistem juga mampu mendeteksi ekspresi wajah. Akurasi yang didapatkan dengan menggunakan fitur yang diusulkan, yaitu Gabor dan Haar wavelet sebesar 92%. Sedangkan akurasi tertinggi didapatkan dengan menggunakan fitur Gabor dan Landmark yang menghasilkan 94,8%. Untuk akurasi terendah, didapatkan dengan menggunakan fitur Haar yaitu sebesar 70%. Dari penelitian ini juga didapatkan kelas yang memiliki jumlah prediksi benar terbanyak merupakan kelas ekspresi senang,

sedangkan kelas dengan misklasifikasi terbanyak merupakan kelas ekspresi merendahkan dengan prediksi kelas ekspresi marah.

Untuk pengembangan pada penelitian selanjutnya, dapat dipertimbangkan untuk memperbanyak jumlah dan variasi ekspresi data *training* yang digunakan untuk menjangkau keberagaman model tiap-tiap ekspresi. Selain itu dapat juga mencoba alternatif metode ekstraksi fitur lain yang lebih optimal dalam memproses citra wajah.

DAFTAR PUSTAKA

- [1] A. Mehrabian, "Communication without words," *Psychology Today*, vol. 2, pp. 52-55, 1968.
- [2] V. Lekshmi and M. Sasikumar, "Analysis of facial expression using Gabor and SVM," *International Journal of Recent Trends in Engineering*, vol. 1, no. 2, pp. 47-50, 2009.
- [3] M. Haghighat, S. Zonouz and M. Abdel-Mottaleb, "CloudID: Trustworthy cloud-based and cross-enterprise biometric identification," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7905-7916, 2015.
- [4] M. Alwakeel and Z. Shaaban, "Face recognition based on Haar wavelet transform and principal component analysis via Levenberg-Marquardt backpropagation neural network," *European Journal of Scientific Research*, vol. 42, no. 1, pp. 25-31, 2010.
- [5] M. Goyani and N. Patel, "Multi-Level Haar Wavelet based Facial Expression Recognition using Logistic Regression," *Indian Journal of Science and Technology*, vol. 10, no. 9, 2017.
- [6] P. Viola and M. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.
- [7] T. Kanade, J. F. Cohn and Y. Tian, "Comprehensive database for facial expression analysis," in *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00)*, Grenoble, France, 2000.
- [8] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *IEEE Computer Society Conference*, San Francisco, CA, USA, 2010.
- [9] P. Ekman and E. (. Rosenberg, *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*, USA: Oxford University Press, 1997.
- [10] Muscle Anatomy, "Muscles Of The Face And Their Functions Facial Expression Recognition (Face Recognition Techniques) Part 1," [Online]. Available: <https://anatomybodysystem.com/muscles-of-the-face-and-their-functions/muscles-of-the-face-and-their-functions-smile-line-muscles-of-the-facial-expression-smile-line/>. [Accessed 30 November 2017].
- [11] M. B. Stegmann, R. Fisker, B. K. Ersbøll, H. H. Thodberg and L. Hyldstrup, "Active Appearance Models: Theory and Cases," 20 September 2000. [Online]. Available: <http://www2.imm.dtu.dk/~aam/main/node34.html>. [Accessed 24 November 2017].
- [12] S. Raschka, "Principal Component Analysis in 3 Simple Steps," 2015. [Online]. Available: http://sebastianraschka.com/Articles/2015_pca_in_3_steps.html. [Accessed 2017 Agustus 2017].
- [13] D. Roth, M. Yang and N. Ahuja, "Learning to recognize three dimensional objects," *Neural Computation*, vol. 14, pp. 1071-1103, 2002.
- [14] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *International Joint Conference on Artificial Intelligence (IJCAI)*, Montreal, 1995.