

METODE EKSTRAKSI CIRI 2DPCA PADA PENGENALAN CITRA WAJAH DENGAN MATLAB

Pratiwi

Program Studi Teknik Informatika
Fakultas Teknologi Informasi, IKPIA Perbanas Jakarta
wiek.pratiwi@gmail.com dan pratiwi@perbanasinstitute.ac.id

ABSTRACT

Research on face recognition have been carried out and the various methods. However, this study is still interesting to do with the development of technologies and methods used in image processing . Feature extraction method at introduction face recognition much is to identify specific facial image using measurable characteristics . In this study, researchers used a feature extraction 2 Dimensional Principal Component Analysis (2DPCA and using Matlab application. In this process faces read as a matrix and made the same size 112x96 pixel, and then do the conversion to grayscale.. After the feature extraction of face images 2DPCA testing was conducted using Euclidean distance. Method is the method by comparing the Euclidean distance of the test image by the training image data base that has minimal distance. Clustering of training data and test data using a 5 fold crossvalidation . Accuracy of the results obtained during experiments conducted at 10 percent face with eigenvalues 80% , 85% , 90% and 95%, the result is an average of over 96.5%.

Key words : Face Recognition, Fiture Extraction, 2DPCA, Euclidean Distance.

INTISARI

Penelitian mengenai pengenalan wajah telah banyak dilakukan dan dengan berbagai metode. Namun penelitian ini masih menarik dilakukan seiring perkembangan teknologi dan metode yang digunakan dalam pengolahan citra. Metode ekstraksi ciri pada pengenalan wajahpun cukup banyak yaitu dengan mengidentifikasi citra wajah secara spesifik dengan menggunakan karakteristik terukur. Pada penelitian ini peneliti menggunakan ekstraksi ciri *2 Dimensional Principal Component Analysis* (2DPCA) dan menggunakan aplikasi Matlab. Data diambil terlebih dahulu dari digital image kemudian dilakukan *cropping* untuk mendapatnya citra wajah. Proses selanjutnya citra wajah dibaca sebagai matrik dan dibuat berukuran sama 112x96 pixel, kemudian dilakukan konversi ke bentuk *grayscale*. Setelah ekstraksi ciri 2DPCA dilakukan pengujian citra wajah dilakukan dengan metode jarak *euclid*. Metode *euclid* adalah metode dengan membandingkan jarak citra pengujian dengan basis data citra pelatihan yang memiliki jarak minimal. Pengelompokan data latih dan data uji menggunakan *5 fold cross validation*. Hasil akurasi yang diperoleh saat dilakukan percobaan pada 10 wajah dengan prosentase nilai eigen 80%, 85%, 90% dan 95% , hasilnya mencapai rata-rata diatas 96,5%.

Kata kunci : Pengenalan Wajah, ekstraksi ciri, 2DPCA, jarak euclid

PENDAHULUAN

Penelitian mengenai pengenalan wajah telah banyak dilakukan dan dengan berbagai metode seperti yang dilakukan oleh Sudarmilah (2009), Yang and Zhang (2004), Soelaiman dan Kusumoputro (2002). Di dalam proses dalam pengenalan wajah cukup banyak metode yang dapat digunakan, salah satunya dengan menggunakan jaringan syaraf tiruan dengan berbasis ekstraksi ciri. Ekstraksi ciri mengidentifikasi citra wajah secara spesifik dengan menggunakan karakteristik terukur. Karakteristik ini digunakan pada pengenalan wajah untuk mendapatkan ciri-ciri yang penting dari basis data citra pelatihan yang akan menjadi referensi.

Metode ekstraksi ciri yang sudah dikenal yaitu ekstraksi ciri *Principal Componen Analysis* (Lindsay S,2002); *Linier Diskriminant Analysis* (LDA), juga *2 Dimensional Principal Component Analysis* (Yang and Zhang, 2004). Pada ekstraksi ciri PCA matrik citra diubah menjadi vektor kolom yang berukuran cukup besar kemudian dilakukan proses pencarian matrik kovarian dan nilai eigen. Sementara pada 2DPCA matrik citra wajah dibaca secara langsung untuk mendapatkan matrik kovarian. Tentunya ukuran matrik kovarian jauh lebih kecil dan memudahkan untuk mengevaluasi keakuratannya. Ciri atau karakteristik utama dalam 2DPCA berupa vektor yang kemudian digunakan dalam

membuat matrik template bagi data latih dan data uji.

Pengenalan wajah itu sendiri adalah proses identifikasi yang berdasarkan citra wajah yang tersimpan dalam basis data (Sudarmilah, 2009), Pengenalan citra wajah setelah proses ekstraksi ciri dilakukan dengan pencocokan citra uji terhadap basis data dari data latih dengan menggunakan jarak minimal diantara keduanya. Perhitungan jarak diantara 2 citra dapat menggunakan jarak *euclid*, jarak *city-block*, jarak *chebychef*, jarak kotak catur, jarak Sorensen (Khadir dan Susanto, 2013). Pada penelitian pengenalan wajah dengan 2DPCA sebagai ekstraksi ciri yang mereduksi langsung citra wajah diperoleh hasil berupa matrik template, kemudian dilakukan pengenalan dengan metode jarak *euclid*. Metode *euclid* adalah metode dengan membandingkan jarak minimum citra pengujian dengan basis data pada citra pelatihan (Pratiwi, 2010).

Penelitian ini bertujuan untuk melakukan pengenalan citra wajah dengan metode jarak dengan ekstraksi ciri 2DPCA dengan ruang lingkup penelitian meliputi:

1. Data citra masukan untuk pelatihan dan pengujian yang diperoleh dengan mengambil citra digital dengan menggunakan kamera.
2. File yang digunakan berekstensi jpg.
3. Citra wajah yang digunakan adalah posisi menghadap ke depan.
4. Ekstraksi ciri yang digunakan adalah metode 2DPCA.
5. Metode yang digunakan untuk pengenalan citra wajah dalam penelitian ini adalah metode jarak *euclid*.
6. Aplikasi pengolahan citra yang digunakan adalah Matlab 2013.

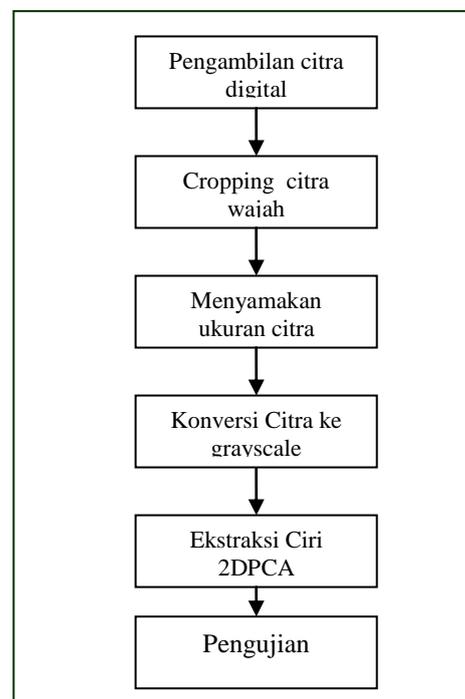
Penelitian ini diharapkan dapat bermanfaat dalam memperkaya kegiatan penelitian pengenalan citra wajah dengan basis ekstraksi ciri terutama metode 2DPCA dan dapat bermanfaat dalam jika diaplikasikan dalam pengembangan teknologi pengolahan citra digital.

METODE PENELITIAN

Pada penelitian ini langkah pertama yang dilakukan tentu saja dengan melakukan studi pustaka yaitu mencari referensi penelitian yang sudah ada sebelumnya dan pengetahuan mengenai ekstraksi ciri 2DPCA.

Langkah-langkah berikutnya yaitu pengambilan citra digital yang kemudian dipotong (*cropping*) hanya pada bagian wajah saja. Setelah itu citra wajah harus memiliki ukuran yang sama dan dilakukan konversi ke *grayscale* untuk dapat dilakukan ekstraksi ciri. *Cropping*, penyamaan ukuran dan konversi ke grayscale merupakan bagian praproses untuk pengenalan citra wajah yang akan dilakukan.

Setelah praproses, dilakukan ekstraksi ciri 2DPCA yang nanti menghasilkan vektor karakteristik yang kemudian dari vektor ini dibuat matrik template untuk setiap data citra wajah. Alur penelitian ini digambarkan pada gambar 1.



Gambar 1. Alur diagram pada penelitian

Pengambilan citra digital dilakukan dengan kamera digital dan diambil data wajah untuk 15 orang laki-laki dan perempuan, serta masing-masing diambil sebanyak 10 citra wajah.



Gambar 2. Contoh citra wajah digital penelitian ini.

Hasil pengambilan citra digital tidak selalu tepat pada wajah maka dilakukan cropping(pemotongan) pada citra tersebut dan hanya diambil citra wajah saja. Citra wajah yang dimaksud adalah bagian dari kepala dari rambut, alis, telinga, hidung, mulut hingga dagu dan sebagian leher. Ukuran citra wajah yang akan diproses untuk ekstraksi ciri 2DPCA harus sama, untuk itu dilakukan penyamaan ke ukuran 112x96 pixel (seperti pada penelitian Yang 2004).

Dalam proses penyamaan ukuran ini pada Matlab digunakan fungsi *imsize*. Pada fungsi *imsize* ini ada 3 jenis yaitu *nearest*, *bilinier*, *bicubic*. Dicontohkan fungsinya adalah:

```
re_foto1=imresize(foto1,[112 96],'nearest');
```

Kemudian dilakukan konversi citra kedalam bentuk *grayscale*. Konversi citra kedalam bentuk *grayscale* pada Matlab menggunakan fungsi *rgb2gray* seperti yang digunakan pada penelitian ini yaitu:

```
fotogray= rgb2gray(re_foto1.jpg).
```



Gambar 3. Contoh citra wajah digital yang sudah mengalami praproses

Dari proses *cropping*, menyamakan ukuran dan konversi dari RGB ke *grayscale* termasuk pada praproses (*preprocessing*) pada pengolahan citra.

Ekstraksi 2DPCA

Citra wajah digambarkan dalam matrik $A_{m \times n}$ kemudian ditransformasi ke bentuk $Y = AX$ lalu diperoleh vektor dengan dimensi m yang merupakan proyeksi vektor ciri dari citra wajah A . Kemudian didefinisikan matrik G_t yaitu

$$G_t = E[(A - EA)^T (A - EA)] \dots \dots \dots (1)$$

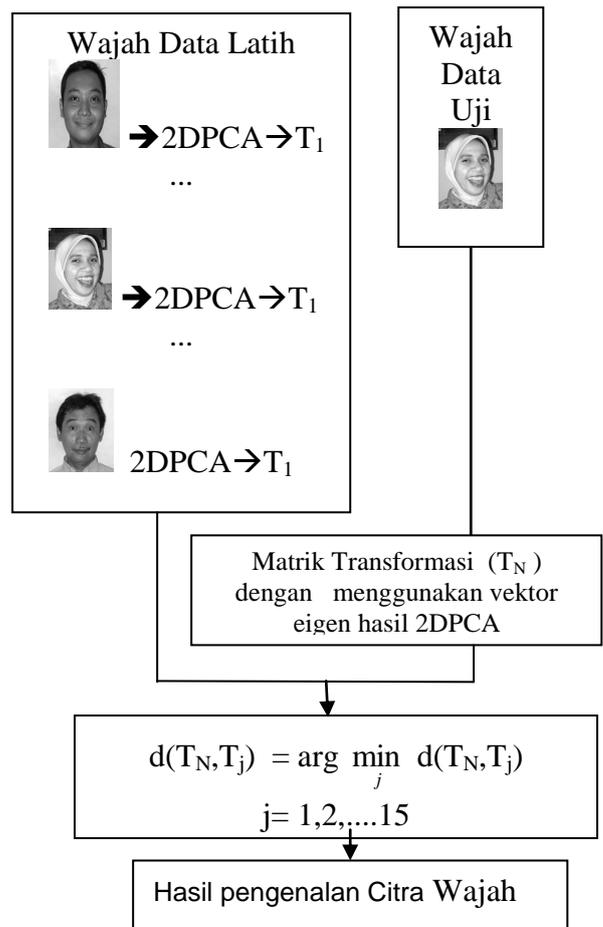
Matrik G_t disebut matrik kovarian citra wajah dan matrik G_t ini dievaluasi dengan citra wajah dari data pelatihan keseluruhan. Pada citra wajah ke- j dinotasikan sebagai matrik $A_{j(m \times n)}$ dimana $j = 1, 2, \dots, M$, serta rata-rata dari seluruh data pelatihan dinotasikan dengan \bar{A} . Untuk itu G_t dievaluasi dengan rumus:

$$G_t = \frac{1}{M} \sum_{j=1}^M (A_j - \bar{A})^T (A_j - \bar{A}) \dots (2)$$

Diperoleh sejumlah proyeksi vektor X_1, \dots, X_d yang disebut komponen utama dalam bentuk vektor dari citra wajah A . Setiap komponen utama 2DPCA adalah vektor dan vektor komponen utama yang digunakan untuk membuat matrik $T_{m \times d} = [Y_1, \dots, Y_d]$ yang disebut matrik ciri dari citra wajah A .

Metode Pengenalan Citra Wajah

Matrik ciri yang diperoleh dari transformasi dengan menggunakan 2DPCA, digunakan untuk membuat matrik template setiap data latih. Pengenalan citra wajah dilakukan dengan mencari jarak minimum antara matrik transformasi pada data uji dengan matrik template dari data latih yang ada. Metode pencarian jarak dengan jarak Euclid, yaitu pengukuran jarak garis lurus antara satu titik dengan titik lain, dalam hal ini jarak antara setiap vektor kolom dalam matriks.



Gambar 4. Metode jarak euclid untuk pengujian.

Didefinisikan jarak antar matrik $T_i = [Y_1^i, Y_2^i, \dots, Y_d^i]$ dan $T_j = [Y_1^j, Y_2^j, \dots, Y_d^j]$ sebagai berikut,

$$d(T_i, T_j) = \sum_{k=1}^d \|Y_k^{(i)} - Y_k^{(j)}\|_2 \dots (3)$$

dimana $\|y_k^i - y_k^j\| =$

$$\sqrt{(y^{1i} - y^{1j})^2 + \dots + (y^{ni} - y^{nj})^2} \dots (4)$$

Berdasarkan data minimum jarak eulid pada data pengujian terhadap pelatihan T_1, T_2, \dots, T_M ($M=$ jumlah data pelatihan) kemudian tiap data diberi identitas klasifikasi ω_k untuk data pelatihan. Apabila suatu data uji dengan matrik transformasi T_N , maka $d(T_N, T_j) = \arg \min_j d(T_N, T_j)$ dengan $j = 1, 2, \dots, 15$, sehingga hasilnya $T_N \in \omega_k$.

K - Fold Cross Validation

Pada pengenalan citra wajah dilakukan pengelompokan data menjadi data latih dan data uji. Pengelompokan data ini terdapat permasalahan dalam menentukan model seleksinya dan hasil estimasi. Salah satu teknik pengelompokan ini adalah menggunakan *cross validation* yang terbagi mejadi 3 yaitu mengelompokan data secara acak (*Random Subsampling*), *k-fold cross validation*, *leave-one-out cross validation*.

K-fold cross validation digunakan dalam metode pembelajaran pada data latih untuk menentukan keakuratan suatu algoritma sehingga dapat menguji data yang tidak terlatih (data uji). Dengan metode ini data dikelompokkan menjadi k kelompok.

Langkah-langkah pengelompokan sebagai berikut :

1. Membagi data yang ada menjadi k kelompok.
2. Untuk setiap k, buat T himpunan data yang memuat semua data latih kecuali yang berada di kelompok ke-k.
3. Kerjakan algoritma yang dimiliki dengan sejumlah T data latih.
4. Uji algoritma ini dengan menggunakan data pada kelompok "k" sebagai data uji.
5. Lakukan pencatatan hasil algoritma.

Data	1	2	3	n-1	n
fold 1						
fold 2						
...						
fold k						

Gambar 5. Ilustrasi k fold cross validation

K-fold cross validation sangat tepat dan berguna ketika menentukan nilai yang tepat untuk k. Metode ini tidak menghabiskan waktu cukup banyak untuk membuat variasi data uji yang ada dan lebih sederhana dan mudah dibandingkan metode *leave-one-out cross validation*.

HASIL DAN PEMBAHASAN

Penelitian dilakukan dengan membagi data yang terdiri dari 150 citra wajah dari 15 individu. Dari 10 citra wajah untuk setiap individu dibagi menjadi 2 bagian utama yaitu data latih dan data uji.

Hasil pengujian diperoleh hasil akurasi mencapai nilai diatas 92,5% ketika menggunakan akumulasi 8 nilai eigen yang memiliki prosentase kontribusi diatas 80 % pada fold 5. Terdapat 30 pengujian untuk kelima pola ini dengan masing masing 2 citra data uji pada setiap fold.

Tabel 1. Hasil pengenalan citra wajah.

Nilai eigen	Akurasi Pengenalan Rata-Rata					Total	
	fold 1	fold 2	fold 3	fold 4	fold 5	Rata-rata	Variansi
80%	98.75%	97.50%	98.75%	97.50%	92.50%	97.00%	0.0005
85%	98.75%	98.75%	98.75%	97.50%	92.50%	97.25%	0.0007
90%	98.75%	97.50%	98.75%	97.50%	92.50%	97.00%	0.0006
95%	98.75%	97.50%	98.75%	95%	92.50%	96.50%	0.00065

Pembagian kelompok data ini mengikuti *5 fold cross validation* .

- fold 1 : Citra 1 dan 2 sebagai data uji dan citra 3,4,5,6,7,8,9,10 sebagai data latih.
- fold 2 : Citra 3 dan 4 sebagai data uji dan citra 1,2,5,6,7,8,9,10 sebagai data latih.
- fold 3 : Citra 5 dan 6 sebagai data uji dan citra 1,2,3,4,7,8,9,10 sebagai data latih.
- fold 4 : Citra 7 dan 8 sebagai data uji dan citra 1,2,3,4,5,6,,9,10 sebagai data latih.
- fold 5 : Citra 9 dan 10 sebagai data uji dan citra 1,2,3,4,5,6,7,8 sebagai data latih

Pada penelitian ini setiap pola memiliki eigen value yang berbeda dan masing-masing dilakukan percobaan untuk 5 prosentase yang berbeda dari akumulasi nilai eigen. Akumulasi nilai eigen dengan prosentase 80%, 85%, 90%, 95 %.

KESIMPULAN DAN SARAN

Proses pengenalan wajah dengan teknik reduksi 2DPCA pada citra wajah seseorang memiliki tingkat akurasi baik dengan rata-rata diatas 92,5%, terlihat pada percobaan dengan *5-fold cross validation* yang mengambil 8 citra wajah sebagai data latih dan 2 citra wajah sebagai data uji.

Secara keseluruhan penelitian tentang pengenalan citra wajah dengan menggunakan 2DPCA memiliki tingkat akurasi lebih dari 96,5%.

Peneliti berharap bahwa penelitian ini masih dapat dikembangkan lebih jauh yang nantinya diharapkan dapat terbentuk suatu sistem pengenalan yang lebih baik, seperti membuat pemodelan aplikasi pengenalan citra wajah dengan 2DPCA ini . Metode pengenalan untuk pengujian dapat menggunakan rumus jarak yang lainnya, tidak hanya dengan jarak *euclid* .

DAFTAR PUSTAKA

Hui Kong and Lei Wang. (2005). *2D Principal Component Analysis for Face Image*

Representation and Recognition, School of Electrical and Electronic Engineering, Nanyang Technological University, Neural Networks 18, 585–594, Nanyang 639798, Singapore.

Kadir,A dan Susanto,A. (2013). *Teori dan Aplikasi Pengolahan Citra*, Yogyakarta, Penerbit Andi

Licesio J ,et all.(2005). *Face Verification Advances Using Spatial Dimension Reduction Methods: 2DPCA & SVM*, Universidad Rey Juan Carlos, F. Roli and S. Vitulano (Eds.): ICIAP 2005, LNCS 3617, pp. 978–985.,Springer-Verlag Berlin Heidelberg 2005.

Lindsay S. (2002), *A Tutorial on Principal Component Analysis*, Chapter 2-3.

McAndrew A. (2004) *An Introduction to Digital Image Processing with Matlab ,Notes for SCM2511 Image Processing 1*, School of Computer Science and Mathematics Victoria University of Technology.

Pratiwi. (2010), *Face Recognition Model Development with Euclidean Distance On Eigen Space with 2DPCA*, Tesis S2 Ilmu Komputer IPB Bogor.

Sudarmilah, Endah. (2009). *Pengenalan Wajah dengan Perbandingan Histogram*, Seminar Nasional Aplikasi Teknologi Informasi 2009.

Soelaiman R, Kusumoputro B. (2002). *Sistem Pengenalan Wajah dengan Penerapan Algoritma Genetika pada Optimasi Basis Proyeksi Metoda Eigenface dan Fisherface*, *Prosiding Ilmu Komputer dan Teknologi Informasi III*, 391, 266-270.

Turk M. and Pentland A. (1991). *Eigenfaces for Recognition*, *J Cognitive Neuroscience*, vol. 3 no 1 , pp 71-86,

Woodward JD., et all. (2003). *BiometricsA Look at Facial Recognition*. Virginia State Crime Commission.

Yang and Zhang. (2004). *Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition*. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol 26 no 1.