New Approach: Dominant and Additional Features Selection Based on Two Dimensional-Discrete Cosine Transform for Face Sketch Recognition

Arif Muntasa arifmuntasa@trunojoyo.ac.id
Informatics Department Trunojoyo University
Kamal, Bangkalan, East Java,
Indonesia

Abstract

Modality reduction by using the Eigentransform method can not efficiently work, when number of training sets larger than image dimension. While modality reduction by using the first derivative negative followed by feature extraction using Two Dimensional Discrete Cosine Transform has limitation, which is feature extraction achieved of face sketch feature is included non-dominant features. We propose to select the image region that contains the dominant features. For each region that contains dominant features will be extracted one frequency by using Two Dimensional-Discrete Cosine Transform. To reduce modality between photographs as training set and face sketches as testing set, we propose to bring the training and testing set toward new dimension by using the first derivative followed by negative process. In order to improve final result on the new dimension, it is necessary to add the testing set pixels by using the difference of photograph average values as training sets and the corresponding face sketches average as testing sets. We employed 100 face sketches as testing and 100 photographs as training set. Experimental results show that maximum recognition is 93%.

Keywords: Face sketch, one frequency, new dimension, dominant and additional features selection.

1. INTRODUCTION

During the last several years, feature detection was conducted by many researches [1], [2], [3], [4]. Besides feature detection, extraction feature has been also conducted for face recognition, such as Principal Component Analysis [10], Linear Descriminant Analysis [11], [12], Radial Basis Function [6], Linear Locality Preserving Projection [7], Elastic Bunch Graph Matching [8], Kernel Principal Component Analysis [5], Kernel Linear Descriminant Analysis [9] and Independent Component Analysis [13]. However, face sketch recognition by using photograph as training set has not been conducted by researchers [1], [14], [15]. Besides high dimension, problem in face sketch recognition is difference modality, where face photograph images as training set and face sketch images as testing set have big difference modality. Popular method used to reduce the difference in modality is the Eigentransform [15], but this method can not efficiently work, when number of training sets larger than image dimension, because number of covariance value dimension larger than image dimension. To overcome this problem, we propose to improve modality reduction and feature extraction process. Our dimension reduction proposed method does not depend on number of training set. To improve modality reduction process, it is necessary to bring photograph as training set and sketch as testing set toward new dimension by using the first derivative followed by negative process. Lastly, to enhance result of modality reduction on new dimension, the difference between training and testing set average value is added to the corresponding testing set. To achieve dominant feature on new dimension of the training and testing set, we propose to select regions that contain face dominant features and additional features to eliminate of feature's location errors.

2. FACE SKETCH RECOGNITION [15]

Face recognition has become interesting issue in biometrics research field. Unfortunately, there are many researchers who have not thought about face sketch interpretation [1] and recognition [15]. In fact, when there was a criminal case, while the camera is not installed around the scene, the witnesses and sketcher will be main key to solve this case. Therefore, it is necessary to conduct research about 'face sketch recognition'. In face sketch recognition, face photograph images as training set and face sketch images as testing set have different modality. Distance between two face photograph images of different person is smaller than distance between face photograph images and face sketch images of same person [1], [15]. Therefore, to recognize face sketch without reducing modality is not supported.

Xiaoou Tang transformed face photograph images to face sketch images by using the Eigentransform [15]. It is method for reducing difference of modality based on the Principal Component Analysis [15]. Consider face photograph images *MxN*, *N* represents number of dimension and *M* represents number of training set [P_1, P_2, ... , P_M] and P_i is the Eigenface of training set, covariant of training set can be computed by using the following equation

$$C=\sum_{i=1}^{M}(P_i-m_p)(P_i-m_p)^T = \mathrm{AA}^{\mathrm{T}} \tag{1}$$

m_p represents face photograph image average, A_p represents training set vector, which is composed in the matrix form

$$A_p=[P_1-m_p, P_2-m_p, \ldots\ldots\ldots\ldots, P_m-m_p] \tag{2}$$

Based on the Singular Value Decomposition (SVD) method, E_p can be expressed by using the following equation

$$E_p = A_p V_p \lambda_p^{-1/2} \tag{3}$$

V and Λ are the Eigenvector and Eigenvalue respectively. New image *P* can be reconstructed by using the following the equation

$$P_r = E_p W_p + m_p \tag{4}$$

W_p represents weight vector that can be computed from image projection to *Eigen faces*

$$W_p = E_p^T (P-m_p) \tag{5}$$

Based on equation (4) and (5), reconstruction of training set can be re-represented by using the following the equation

$$\begin{aligned} P_r &= A_p V_p \lambda_p^{-\frac{1}{2}} w_p + m_p \\ &= \mathrm{A_p} C + m_p \end{aligned} \tag{6}$$

However, the Eigentransform has a weakness, if number of image dimensions is smaller than number of training set, then this method can not efficiently work.

3 PROPOSED METHOD

Photograph and sketch image are two images with different modality. Two photograph images of different person have smaller distance than photograph and sketch of same person. It shows that, besides feature extraction process, modality is main key to recognize face sketch before feature extraction is conducted. To maximize feature extraction on face photograph and sketch image, it is necessary to remove non-face and non-dominant face feature regions. Examples of non-face image are the image background, neck and clothes. Non dominant face feature, included hair and ear. Dominant face features, included left eyebrow, right eyebrow, left eye, right eyebrow, nose and lips. To reduce error in feature extraction, it is necessary to add additional features region closed to dominant face features region. We proposed framework as seen in Figure 1.

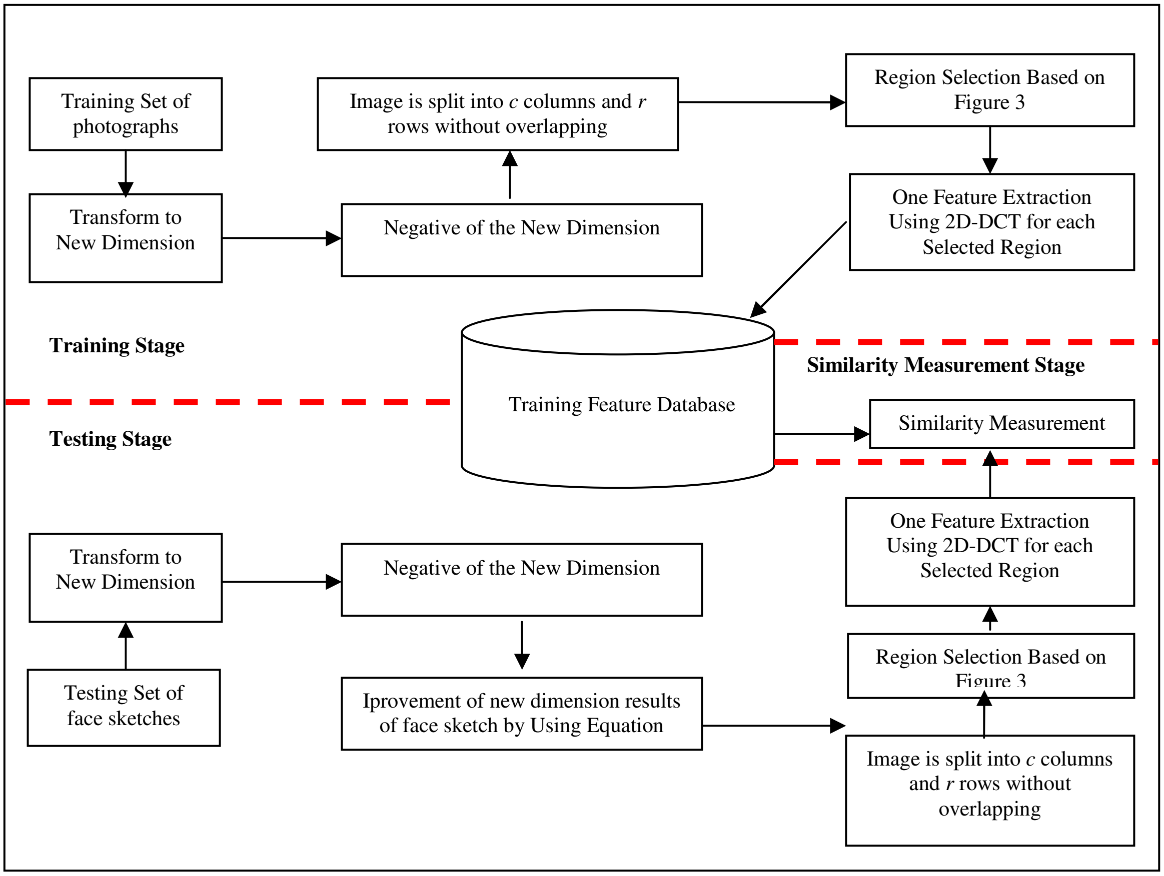

FIGURE 1 Framework of Proposed Method

3.1. Transformation toward New Dimension

Consider $f(x, y)$ and $g(x, y)$ are training and testing set respectively. To reduce the differences in the modality, for both training and testing set should be brought toward new dimension. First derivative is used to apply it. To improve the result of the first derivative, it is necessary to bring negative form. The first derivative of the training set can be expressed by using the following equation

$$f' = \begin{bmatrix} f'_x \\ f'_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} \\ \frac{\partial f(x,y)}{\partial y} \end{bmatrix} \tag{7}$$

where

$$\frac{\partial f(x,y)}{\partial x}=\frac{f(x+\Delta x,y)-f(x,y)}{\Delta x} \tag{8}$$

$$\frac{\partial f(x,y)}{\partial y}=\frac{f(x,y+\Delta y)-f(x,y)}{\Delta y} \tag{9}$$

The Sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively inexpensive in terms of computations. Vertical and horizontal derivative can be approximated by using the following equation

$$\begin{aligned} f'_x = {} & [f(x+1,y-1)+2f(x+1,y)+f(x+1,y+1)]- \\ & [f(x-1,y-1)+2f(x-1,y)+f(x-1,y+1)] \end{aligned} \tag{10}$$

$$\begin{aligned} f'_y = {} & [f(x-1,y+1)+2f(x,y+1)+f(x+1,y+1)]- \\ & [f(x-1,y-1)+2f(x,y-1)+f(x+1,y-1)] \end{aligned} \tag{11}$$

Equation (10) and (11) can be written in the matrix form as follows

$$f'_x=\begin{bmatrix}-1 & -2 & -1\\ 0 & 0 & 0\\ 1 & 2 & 1\end{bmatrix} \tag{12}$$

$$f'_y=\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix} \tag{13}$$

The magnitude of these equations can be calculated by using the following equation

$$f'[f(x,y)]=\sqrt{(f'_x)^2+(f'_y)^2} \tag{14}$$

To simplify the convolution result using the Sobel Operator on face photograph images as training set, it can be written in the following equation

$$f(x,y)\xrightarrow[\text{1'}st\text{ Derivative}]{}f'(x,y) \tag{15}$$

And negative of Equation (15) can be written in the following equation

$$\sim f'(x,y)=255-f'(x,y) \tag{16}$$

Similarly, it can be applied into face sketch as testing set. The result of the sobel operator on face sketch images as testing set can be expressed as follows

$$g(x,y)\xrightarrow[\text{1'}st\text{ Derivative}]{}g'(x,y) \tag{17}$$

And negative of Equation (17) can be written in the following equation

$$\sim g'(x,y)=255-g'(x,y) \tag{18}$$

To improve the result of new dimension on face sketch, it is necessary to add the testing set by using the difference of average photograph values as training sets and face sketches average on the corresponding testing sets. It can be modeled by using the following equation

$$\sim g'(x,y)\approx\sim g'(x,y)+|\,\bar{f}_{training}-\bar{g}_{testing}\,| \tag{19}$$

Where $\bar{f}_{training}$ represents training sets average value and $\bar{g}_{training}$ is the corresponding testing set average value, therefore $\bar{f}_{training}$ is constant, while $\bar{g}_{training}$ based on testing set.

3.2. Localizing and Feature Selection

If photograph image is represented by using *f(x, y)* as training set and face sketch is represented by using *g(x, y)* as testing set, then images will be split into *c* columns and *r* rows. To achieve feature dominant, it is important to select regions that contain dominant features, such as left eyes, right eyes, left eyebrow, right eyebrow, nose and lips. In this research, image will be split into 7 columns and 18 rows, so there are 126 regions for each image as seen in Figure 1. Based on Figure 1, we have selected 13 regions that contain dominant and 12 additional features regions to reduce feature's location errors as seen in Table 1. In order to reduce feature's location error, we added 2 regions above right eyebrow location (region feature labels are 31 and 38), 2 regions above left eyebrow location (region feature labels are 33 and 40), 2 regions above nose feature (region feature labels are 46 and 53) and 6 regions below lips feature (region feature labels are 87, 88, 89, 94, 95 and 96). The result of feature selection for Figure 2 can be seen in Figure 3.

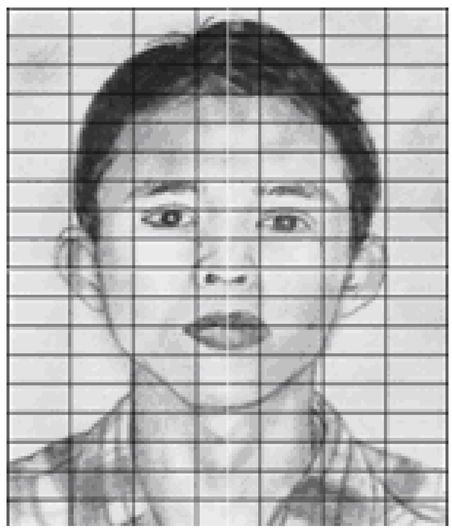

FIGURE 2 Image was split into 126 regions (7 columns and 18 rows)

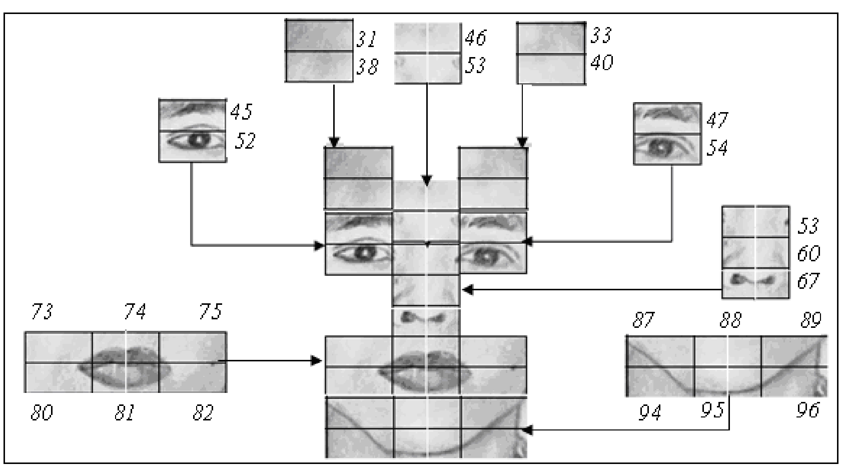

FIGURE 3 Selection of Feature's Regions to achieve Feature's Dominant and Their Additional to Reduce Feature's Error Location

While feature's region labels can be seen in Table 1

No	Features	Region Feature Labels	Additional Region Feature Labels
1	Right Eyebrow	45	31 and 38
2	Right Eye	52	-
3	Left Eyebrow	47	33 and 40
4	Left Eye	54	-
5	Nose	53 , 60 and 67	46 and 53
6	Lips	73, 74, 75, 80, 81 and 82	87, 88, 89, 94, 95 and 96.

TABLE 1 Label of Feature's Region and Their Additional

3.3. Feature extraction by Using Two Dimensional Discrete Transform (2D-DCT)

In order to achieve image features, we proposed to apply 2D-DCT for each region selected. To preserve region selected, for both feature and additional feature regions, we extracted features based on original feature location. Feature extraction resulted is based on location for each region. If number of regions selected are 25 region, then number of frequencies resulted are 25. 2D-DCT can be written in the following equation

$$C(u,v) = \frac{2}{\sqrt{M.N}}\alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y).\cos\theta.\cos\vartheta \quad (20)$$

Where

$$\theta = \left[\frac{\pi(2x+1)u}{2N}\right] \quad (21)$$

$$\vartheta = \left[\frac{\pi(2y+1)v}{2M}\right] \quad (22)$$

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & Untuk \quad \mathrm{u} = 0 \\ \sqrt{\frac{2}{N}} & Untuk \quad \mathrm{u} <> 0 \end{cases} \quad (23)$$

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & Untuk \quad \mathrm{v} = 0 \\ \sqrt{\frac{2}{N}} & Untuk \quad \mathrm{v} <> 0 \end{cases} \quad (24)$$

Consider an image *f(x, y)*, *w* and *h* are width and height of image respectively. If it is divided into *c* columns and *r* rows, then the area for each region (L_i) and *i* expresses region index, for each region can be written in the following equation

$$L_i = \frac{w.h}{c.r} \quad (25)$$

If number of selected regions and additional regions are *R* and *A* respectively, then number of processes required to extract an image with one frequency on each region using equation (20), (21), (22), (23) and (24) can be written by using the following equation

$$nL = \frac{w.h}{c.r} * (R + A) \quad (26)$$

If the features extraction used is selected regions only, then number of processes required to extract an image can be expressed by using the following equation

$$nL = \frac{w.h}{c.r} * R \tag{27}$$

Based on equation (26) and (27), the greater number of selected regions and additional regions, then the longer time the process required.

3.4. Similarity Measurement

To classify the result of feature extraction, four similarity measurement methods are used, which are the Euclidian Distance (d_1), Angular Separation (d_2) and the Canberra (d_3), as written in the following equation

$$d_1(f(x,y), g(x,y)) = \sqrt{\sum_{k=1}^{nF} (f(x,y)_k - g(x,y)_k)^2} \tag{28}$$

$$d_2\big(f(x,y), g(x,y)\big) = \sum_{k=1}^{nF} \frac{f(x,y)_k . g(x,y)_k}{\|f(x,y)_k\| . \|g(x,y)_k\|} \tag{29}$$

$$d_3\big(f(x,y), g(x,y)\big) = \sum_{k=1}^{nF} \frac{| f(x,y)_k - g(x,y)_k |}{| f(x,y)_k | + | g(x,y)_k |} \tag{30}$$

4 EXPERIMENTAL RESULTS AND DISCUSSION

In this research, we employed 100 face photograph images as training set (image samples can be seen in Figure 4) and 100 face sketch images as testing set (image samples can be seen in Figure 5). We perform two experiments. The first experiment was conducted by using 13 regions that contain feature dominant and secondly experiment was conducted by using 25 regions, which are 13 regions that contain feature dominant and 12 additional regions as seen in Table 1. Experimental results of our proposed method can be seen In Table 2.

On the first experiment, errors occurred are caused by deviation on feature selection of face sketch, for both the training and the testing set. Error of Features selection has caused the frequency difference between features of the corresponding training and the testing set. This difference will cause errors in classification. Based on Table 2 can be shown that the first experiment accuracies for the Euclidian Distance, the Angular Separation and the Canberra are 91%, 90% and 91%.

The weakness of the first experiments can be improved by additional features for each the corresponding features. On the second experiment, additional features of the area around of the corresponding feature causes number of generated frequencies more than the first experiment. However, the frequency difference between the corresponding training and testing set produces smaller frequencies than the first experiment. It can reduce the error rate that occurred in the first experiment. The Accuracies of the second experimental results are 91%, 92% and 93%. It shows that additional region selection can improve face sketch recognition of the first experimental.

Our proposed method has higher recognition than our previous research. On our previous research obtained maximum recognition accuracy reached only 89% [3].

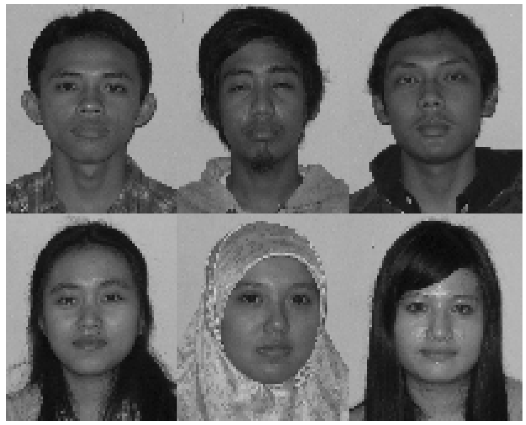

FIGURE 4 samples of 6 person face photograph images

FIGURE 5 samples of 6 person face sketch images

Experimental Results	The number of Regions Used	Face Sketch Recognition (%)		
		Euclidian Distance	Angular Separation	Canberra
1^{st}	13	91	90	91
2^{nd}	25	91	92	93
Average		91	91	92
Maximum		91	92	93

Table 2 Experimental Results of Our Proposed Method

5 CONCLUSION AND FUTURE WORK

In this paper, we can show that

1. Combining of feature extraction between dominant features and additional features can produce higher face sketch recognition than feature extraction using dominant features only.
2. Improvement of new dimension transformation results on face sketch by adding the difference of average photograph values face sketches average and dominant features selection followed by dominant features selection, additional features selection and

feature extraction can obtain higher face sketch recognition than without dominant feature and additional features selection
3. Errors that occurred in face sketch recognition are caused error in dominant features and additional features selection.
4. Generally, Errors occurred on the first and the second experiment are caused by errors in the dominant features selection of face sketch, it can produce high frequency difference value of corresponding data training and data testing. This difference will cause errors in classification.

For future work, we will detect feature location first to obtain more accurate position of features before feature extraction is conducted. To improve similarity measurement, the Support Vector Machine method will be used.

6 REFERENCES

[1]. A. Muntasa, M. Hariadi., M. H. Purnomo. "*Maximum Feature Value Selection Of Nonlinear Function Based On Kernel Pca For Face Recognition"*. In Proceeding of The 4th Conference On Information & Communication Technology and Systems, Surabaya, Indonesia, 2008

[2]. A. Muntasa, M. Hariadi., M. H. Purnomo. *"A New Formulation of Face Sketch Multiple Features Detection Using Pyramid Parameter Model and Simultaneously Landmark Movement"*. IJCSNS International Journal of Computer Science and Network Security, 9(9): 2009

[3.] A. Muntasa. "*A Novel Approach for Face Sketch Recognition Based on the First Derivative Negative and 2D-DCT with Overlapping Model*". International Journal of Computer Science, (Accepted), 2010

[4]. D. Cristinacce and T.F. Cootes. *"Facial Feature Detection using ADABOOST with Shape Constraints"*. In Proceeding .BMVC, Vol.1, 2000

[5]. Lu J., P. K.N. , V. A.N.. *"Face Recognition Using Kernel Direct Discriminant Analysis Algorithms"*. IEEE Trans. Neural Networks, 14(1):117-126, 2003

[6]. Su, H., Feng D., and Zhao R.-C. *"Face Recognition Using Multi-feature and Radial Basis Function Network"*. In Proceeding of the Pan-Sydney Area Workshop on Visual Information Processing, Sydney, Australia, 2002

[7]. X. He, S. Yan, Y. Hu, P. Niyogi, Hong-Jiang Zhang. *"Face Recognition Using Laplacianfaces"*. IEEE Transactions On Pattern Analysis And Machine Intelligence,27(3):328-340, 2005

[8]. L. Wiskott, J.M. Fellous, N. Kruger, C. von der Malsburg. *"Face Recognition by Elastic Bunch Graph Matching"*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 19(7): 775- 779, 1997

[9]. A. Muntasa, M. Hariadi., M. H. Purnomo. *"Maximum Feature Value Selection Of Nonlinear Function Based On Kernel Pca For Face Recognition"*. In Proceeding of The 4th Conferrence On Information & Communication Technology and Systems, Surabaya, Indonesia, 2008

[10] M. Turk, A. Pentland. *"Eigenfaces for recognition"*. *Journal of Cognitive Science, 71–86, 1991*

[11] J.H.P.N. Belhumeur, D. Kriegman. *"Eigenfaces vs. fisherfaces: Recognition using class specific linear projection"* IEEE Trans. on PAMI, 19(7):711–720, 1997

[12]. Yambor, W.S. *"Analysis of PCA-Based and Fisher Discriminant-Based Image Recognition Algorithms"*. Tesis of Master, Colorado State University, 2000

[13]. Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. *"Face recognition by independent component analysis"*. IEEE Trans. on Neural Networks, 13(6): 1450-1464, 2002

[14]. R.G. Uhl and N.d.V. Lobo. "A Framework for Recognizing a Facial Image from A Police Sketch". In *Proceedings of CVPR*, 1996

[15]. X. Tang, X. Wang. *"Face Sketch Recognition".* IEEE Transactions on Circuits and Systems for Video Technology, 14(1):50-57, 2000