# FACIAL PARTS RECOGNITION USING LIFTING WAVELET FILTERS LEARNED BY KURTOSIS-MINIMIZATION

Koichi Niijima
*Department of Informatics, Kyushu University*
*6-1, Kasuga-koen, Kasuga, Fukuoka 816-8580, Japan*

Keywords: Facial parts recognition, lifting wavelet filter, learning, kurtosis, ill-posed problem, regularization.

Abstract: We propose a method for recognizing facial parts using the lifting wavelet filters learned by kurtosis-minimization. This method is based on the following three features of kurtosis: If a random variable has a gaussian distribution, its kurtosis is zero. If the kurtosis is positive, the respective distribution is super-gaussian. The value of kurtosis is bounded below. It is known that the histogram of wavelet coefficients for a natural image behaves like a supergaussian distribution. Exploiting these properties, free parameters included in the lifting wavelet filter are learned so that the kurtosis of lifting wavelet coefficients for the target facial part is minimized. Since this minimization problem is an ill-posed problem, it is solved by employing the regularization method. Facial parts recognition is accomplished by extracting facial parts similar to the target facial part. In simulation, a lifting wavelet filter is learned using the narrow eyes of a female, and the learned lifting filter is applied to facial images of 10 females and 10 males, whose expressions are neutral, smile, anger, and scream, to recognize eye part.

## 1 INTRODUCTION

Facial parts recognition is an important problem for face expression recognition. Many face recognition methods have been proposed so far. Principle component analysis is a traditional classification technique for face recognition (Pentland et al., 1994). A framework of hidden Markov models has been used for recognition of eye movement (Jaimes et al., 2001). Support vector machine is a new tool for solving the classification problems (Vapnik, 1998). Recently, an approach using AdaBoost, which is one of the machine learning techniques, has attracted considerable attention as a method for face recognition (Tieu and Viola, 2000).

Unlike such recognition techniques, we have presented a method of person identification, which uses the learned lifting wavelet filters (Takano et al., 2003; Takano et al., 2004; Takano and Niijima, 2005). The learning technique employed therein is to maximize the cosine of an angle between a vector whose components are lifting filters and a vector consisting of pixels in the facial part. In person identification, a slight difference of facial parts such as eyes, nose, and lips must be distinguished. So, we learned several lifting wavelet filters at the center of each of the facial parts so that they can capture the features of the objects. However, since the designed filters are low-pass filters, a recognition method using them is not robust for changing brightness. More recently, we presented a fast objects detecting method using the lifting wavelet filters learned by variance-maximization (Niijima, 2005). Although this method is fast enough for online processing, it extracts unnecessary objects as well as the target one. This suggests that only the use of variance, which is the second order statistics, is not sufficient for the exact detection of objects.

In this paper, we propose a method for recognizing facial parts exploiting the lifting wavelet filters learned by kurtosis-minimization. One of the features of kurtosis is that if a random variable has a gaussian distribution, its kurtosis is zero. If the kurtosis is positive, the respective distribution is supergaussian, which has a sharper peak and longer tails than the gaussian distribution. This implies that the variance of gaussian distribution is bigger than that of supergaussian one. Another very important feature of kurtosis is that the value of kurtosis is bounded below.

It is known from numerical experiments that the histogram of wavelet coefficients for a natural image



Niijima K. (2006).
FACIAL PARTS RECOGNITION USING LIFTING WAVELET FILTERS LEARNED BY KURTOSIS-MINIMIZATION.
In *Proceedings of the First International Conference on Computer Vision Theory and Applications*, pages 41-47
DOI: 10.5220/0001367900410047

behaves like a supergaussian distribution. Therefore, by learning free parameters contained in the lifting wavelet coefficients so as to minimize their kurtosis, we can make the variance of the coefficients large. The large values of the obtained lifting wavelet coefficients have the features of the facial part in question, and the learned filter can be considered as a recognizer of the target facial part. The facial part is called positive data, and the training image except for it negative data. For the lifting wavelet coefficients of the negative data, the free parameters are learned so that their variance becomes small.

Such a minimization problem is a kind of inverse problem and ill-conditioned. So, we apply the regularization method to solve the problem. The solutions of the problem can be found by exploiting various gradient methods such as the steepest descent method and the conjugate gradient method. However, these techniques usually need a lot of time to obtain convergence results. In this paper, we replace the problem by a problem of seeking stationary points of the corresponding functional, and obtain them by Newton's method. The stationary points are local minima of the functional. Different local minima can be found depending on the starting values of Newton's iteration. We seek the solution by starting the iteration from zero-vector.

We extract facial parts from a query image similar to the target facial part by applying the learned lifting wavelet filter to the query image. The extracted facial part is recognized as the target one.

In simulation, the anger face of a female is used as a training image. The target facial part is her narrow eyes. A lifting filter including the learned parameters is applied to a variety of human faces whose expressions are standard, smile, anger, and scream. It is also checked whether the proposed method is robust for changing brightness for some illuminated facial images.

The remainder of this paper is organized as follows. Section 2 describes the relation between a lifting dyadic wavelet filter and an elliptic-type of partial differential operator. Our learning algorithm is presented in Section 3. We describe an extraction method in Section 4, and a recognition method in Section 5. Section 6 is simulation. Finally, we conclude with Section 7.

## 2 LIFTING WAVELET FILTERS

Let $\{h_n^o, g_n^o, \tilde{h}_n^o, \tilde{g}_n^o\}$ be a set of dyadic wavelet filters (Mallat, 1998). The filters $h_n^o$ and $g_n^o$ are called low-pass and high-pass analysis filters, respectively, and the filters $\tilde{h}_n^o$ and $\tilde{g}_n^o$ are low-pass and high-pass synthesis filters, respectively. A lifting scheme for the dyadic wavelet is described as follows:

$$\begin{aligned} h_n &= h_n^o, \\ g_n &= g_n^o - \sum_k \lambda_k h_{n-k}^o, \\ \tilde{h}_n &= \tilde{h}_n^o + \sum_k \lambda_{-k} \tilde{g}_{n-k}^o, \\ \tilde{g}_n &= \tilde{g}_n^o. \end{aligned} \tag{1}$$

This scheme generalizes Sweldens' biorthogonal lifting scheme (Sweldens, 1996). We proved that the lifted filters $\{h_n, g_n, \tilde{h}_n, \tilde{g}_n\}$ also become a set of dyadic wavelet filters (Abdukirim et al., 2005). Here $\lambda_k$'s denote free parameters. In this paper, we only use the lifted filter (1).

We denote an image by $u_{i,j}$. By applying the low-pass analysis filter $h_n^o$ in vertical direction to $u_{i,j}$, we can get

$$C_{m,k}^{col} = \sum_j h_j^o u_{m,k+j}.$$

Next, an application of the lifted filter (1) in horizontal direction to $C_{m,k}^{col}$ yields the following lifting wavelet coefficients

$$D_{m,k} = \sum_i g_i^d C_{m+i,k}^{col}. \tag{2}$$

Here $g_i^d$'s are given by

$$g_i^d = g_i^o - \sum_{l=-L}^{L} \lambda_l^d h_{i-l}^o, \quad i = -L-M, ..., L+M+1,$$

where $\lambda_l^d$'s represent free parameters in horizontal direction and we assumed that the index $i$ of the filter $h_i^o$ moves from $-M$ to $M+1$. Similarly, we obtain lifting wavelet coefficients in vertical direction

$$E_{m,k} = \sum_j g_j^e C_{m,k+j}^{row}. \tag{3}$$

Here $C_{m,k}^{row}$ is given by

$$C_{m,k}^{row} = \sum_i h_i^o u_{m+i,k},$$

and $g_j^e$'s are determined as follows:

$$g_j^e = g_j^o - \sum_{l=-L}^{L} \lambda_l^e h_{j-l}^o, \quad j = -L-M, ..., L+M+1,$$

where $\lambda_l^e$'s represent free parameters in vertical direction.

We choose the initial high-pass filters $g_n^o$ as $g_0^o = g_2^o = -0.25\sqrt{2}$, $g_1^o = 0.5\sqrt{2}$ and $g_i^o = 0$ otherwise. Such dyadic wavelet filters have been provided in (Mallat, 1998). We put

$$w_{m,k} = D_{m,k} + E_{m,k}. \tag{4}$$

From (2) and (3), the sum $w_{m,k}$ can be expressed as

$$w_{m,k} = \sum_i g_i^o C_{m+i,k}^{col} + \sum_j g_j^o C_{m,k+j}^{row} - \left( \sum_{l=-L}^{L} \lambda_l^d C_{m+l,k} + \sum_{l=-L}^{L} \lambda_l^e C_{m,k+l} \right) \quad (5)$$

with $C_{m,k} = \sum_{i,j} h_i^o h_j^o u_{m+i,k+j}$. Therefore, a lifting wavelet filter defined by $w_{m,k}$ approximates a partial differential operator $L(\lambda^d, \lambda^e)$ defined by

$$L(\lambda^d, \lambda^e)u = -\left( \frac{\partial^2}{\partial x^2}(I_y u) + \frac{\partial^2}{\partial y^2}(I_x u) \right) - I(\lambda^d, \lambda^e)u. \quad (6)$$

Here $I_y u$, $I_x u$ and $I(\lambda^d, \lambda^e)u$ represent the integral versions of $C_{m,k}^{col}$, $C_{m,k}^{row}$ and the last term of (5), respectively, and

$$\lambda^d = (\lambda_{-L}^d, ..., \lambda_L^d), \qquad \lambda^e = (\lambda_{-L}^e, ..., \lambda_L^e).$$

# 3 KURTOSIS-MINIMIZATION LEARNING

We start with the definition of kurtosis. Kurtosis is defined in the zero-mean case by the equation

$$\text{kurt}(w) = < w^4 > -3 < w^2 >^2, \quad (7)$$

where $w$ is a random variable and $< w >$ denotes the expectation of $w$. In case of kurt$(w) > 0$, the distribution of $w$ is said to be supergaussian. If $w$ has a gaussian distribution, then the kurtosis is zero, i.e., kurt$(w) = 0$. A typical supergaussian probability density has a shaper peak and longer tails than the gaussian probability density function (pdf). Therefore, the variance of gaussian distribution is bigger than that of supergaussian pdf. It is known that the value of kurtosis is bounded below. Such properties of kurtosis are useful for our analysis.

Let us denote a training image also by $u_{i,j}$, and its domain by $\Omega_d$. We extract a facial part such as eyes, nose, and lips from the training image. The region of the extracted facial part is denoted by $\omega_d$, and the number of pixels in $\omega_d$ by $P$. The facial part is called positive data, and the image in the region $\Omega_d \setminus \omega_d$ negative data. The number of pixels in $\Omega_d \setminus \omega_d$ is denoted by $N$. Using these positive and negative data, we learn free parameters $\lambda_l^d$'s and $\lambda_l^e$'s appeared in (5). Our learning method is to minimize the kurtosis of the lifting wavelet coefficients for the positive data, and to minimize the variance of those for the negative data.

We extend the facial part periodically in horizontal and vertical directions, and compute the wavelet coefficients $w_{m,k}$, defined by (4), of the facial part. Assume that $\lambda_l^d$'s and $\lambda_l^e$'s satisfy

$$\sum_{l=-L}^{L} (\lambda_l^d + \lambda_l^e) = 0. \quad (8)$$

Then, we can prove

$$< w > = \frac{1}{P} \sum_{(m,k)\in\omega_d} w_{m,k} = 0. \quad (9)$$

Therefore, the kurtosis of the lifting wavelet coefficients $w_{m,k}$ is given by

$$J_1 = \frac{1}{P} \sum_{(m,k)\in\omega_d} w_{m,k}^4 - 3 \left( \frac{1}{P} \sum_{(m,k)\in\omega_d} w_{m,k}^2 \right)^2, \quad (10)$$

which is minimized.

For the negative data in the region $\Omega_d \setminus \omega_d$, we compute their lifting wavelet coefficients $w_{m,k}$, and minimize the variance

$$J_2 = \frac{1}{N} \sum_{(m,k)\in\Omega_d\setminus\omega_d} w_{m,k}^2. \quad (11)$$

Thus, our learning algorithm of free parameters $\lambda_l^d$'s and $\lambda_l^e$'s is a process of minimizing the sum of (10) and (11) under the condition (8).

On the other hand, the continuous version of this minimization problem is to minimize

$$\int_\omega \left( L(\lambda^d, \lambda^e)u(x,y) \right)^4 dxdy - 3 \left( \int_\omega \left( L(\lambda^d, \lambda^e)u(x,y) \right)^2 dxdy \right)^2$$

and

$$\int_{\Omega\setminus\omega} \left( L(\lambda^d, \lambda^e)u(x,y) \right)^2 dxdy$$

subject to the constraint (8). Here $\Omega$ is a region corresponding to $\Omega_d$, and $\omega$ a region corresponding to $\omega_d$. The operator $L(\lambda^d, \lambda^e)$ has been given in (6). This continuous problem is an inverse problem which is ill-conditioned. Therefore, the discrete version is also ill-conditioned. To overcome this difficulty, we employ the regularization method. Thus, a functional to be minimized is provided by

$$J = \frac{1}{4} J_1 + \frac{K_0}{4} J_2 + \frac{K_1}{2} \left( \sum_{l=-L}^{L} (\lambda_l^d + \lambda_l^e) \right)^2 + \frac{\delta}{2} \sum_{l=-L}^{L} \left( (\lambda_l^d)^2 + (\lambda_l^e)^2 \right). \quad (12)$$

Here the third term of the right hand side comes from (8), and $K_i, i = 0, 1$ are penalty constants. The last term means regularization and $\delta$ is a sufficiently small positive number.

Although the functional (12) is a polynomial of fourth degree with respect to the free parameters $\lambda_l^d$'s and $\lambda_l^e$'s, it has a possibility of having many local minima. Since it is difficult to obtain a global minimum, we seek local minima. Various gradient methods are often used for computing local minima. However, these methods are slow in convergence. In this paper, we employ Newton's method to solve the problem fast. Newton's method is applied to a system of simultaneous nonlinear equations:

$$\frac{\partial J}{\partial \lambda_l^d} = 0, \qquad l = -L, ..., L, \tag{13}$$

$$\frac{\partial J}{\partial \lambda_l^e} = 0, \qquad l = -L, ..., L. \tag{14}$$

The process of solving (13) and (14) by Newton's method gives our algorithm for learning the free parameters $\lambda_l^d$'s and $\lambda_l^e$'s.

## 4 FACIAL PARTS EXTRACTION

The learned parameters $\lambda_l^d$'s and $\lambda_l^e$'s make the variance of the lifting wavelet coefficients for the positive data large, and that for the negative data small. Therefore, we can extract facial parts of a query image similar to the training facial part by selecting the large wavelet coefficients, which are computed using the learned lifting filter. Our facial extraction algorithm involves the following steps.

1. Compute wavelet coefficients in horizontal and vertical directions by applying the initial dyadic wavelet filters to a query image $u_{i,j}$.
2. By combining them with the parameters $\lambda_l^d$'s and $\lambda_l^e$'s learned for the training image, compute the lifting wavelet coefficients $D_{m,k}$ and $E_{m,k}$ defined by (2) and (3), respectively.
3. Calculate the sum $w_{m,k} = D_{m,k} + E_{m,k}$.
4. Find the locations $(m, k)$ such that $|w_{m,k}| \geq R\sigma$, where $\sigma$ is the standard deviation of lifting wavelet coefficients computed for the training facial part and $R$ denotes some constant.
5. Extract an image region, in which the detected locations are concentrated, as an object similar to the training facial part.

## 5 RECOGNIZER

The learned filter can extract only a facial part similar to the training one. For example, if the training pattern is narrow eyes, it does not extract closed eyes. Therefore, the learned filter is a recognizer of the training facial part. It is important to indicate that a lifting filter has to be learned per facial part. The positive data may be constructed by combining the same type of several facial parts such as large eyes, narrow eyes, and closed eyes.

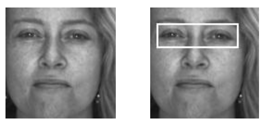
## 6 SIMULATION

We conducted our experiments on the AR face database (Martinez and Benavente, 1998). The initial filters we use are the cubic spline dyadic wavelet filters listed in Table 1 (Mallat, 1998). The number of free parameters in each direction is 15, i.e., $L = 7$. Therefore, 30 free parameters are determined exploiting our learning algorithm. The gray scales of images are converted into the range $[0, 1]$ for stable computation.

Table 1: Cubic spline dyadic wavelet filters (only low-pass and high-pass analysis filters).

| $n$ | $h_n^o/\sqrt{2}$ | $g_n^o/\sqrt{2}$ |
|---|---|---|
| -2 | 0.03125 | |
| -1 | 0.15625 | |
| 0 | 0.31250 | -0.25 |
| 1 | 0.31250 | 0.50 |
| 2 | 0.15625 | -0.25 |
| 3 | 0.03125 | |

The training pattern is the facial image of a female having $120 \times 120$ size, which is shown in Figure 1(a). The positive data are her narrow eyes and the negative data are the image except for the eye region. The positive data are illustrated in Figure 1(b).



(a) (b)

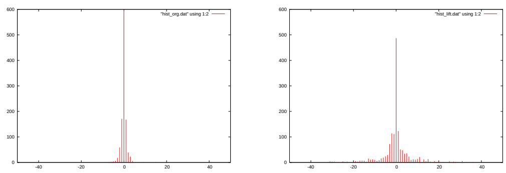
Figure 1: (a) Training facial image, (b) Positive data of narrow eyes.

The penalty constants $K_i, i = 0, 1$ appeared in (12) are chosen as $K_0 = 50$ and $K_1 = 1000$, respectively. The regularization constant $\delta$ is selected as $\delta = 0.0001$. Newton's iteration for solving (13) and (14) was started from zero-vector. We list the learned parameters in Table 2.

Figure 2 shows the histograms of the wavelet coefficients $w_{m,k}, (m, k) \in \omega_d$ obtained exploiting the initial and the learned lifting filters. As seen from Figure 2, the latter histogram is flatter compared with the

Table 2: Learned parameters for the positive and negative data of the facial image illustrated in Figure 1.

| $l$ | $\lambda_l^d$ | $\lambda_l^e$ |
|---|---|---|
| -7 | -0.026973 | -0.013236 |
| -6 | 0.103753 | 0.107106 |
| -5 | -0.166372 | -0.291907 |
| -4 | 0.056374 | 0.200126 |
| -3 | 0.222947 | 0.288220 |
| -2 | -0.427482 | -0.631153 |
| -1 | 0.760302 | 0.906046 |
| 0 | -2.181300 | -2.181300 |
| 1 | 3.417086 | 3.379362 |
| 2 | -1.808546 | -1.953494 |
| 3 | -0.912882 | -0.740206 |
| 4 | 1.503153 | 1.561924 |
| 5 | -0.382691 | -0.580599 |
| 6 | -0.360328 | -0.224570 |
| 7 | 0.204865 | 0.171818 |



(a) (b)

Figure 2: Histograms of the wavelet coefficients $w_{m,k}, (m,k) \in \omega_d$ obtained by using (a) the initial filter, (b) the learned lifting filter.

former one. This means that the variance of the latter histogram is bigger than that of the former one. Actually, the standard derivations of the former and the latter distributions for the positive data were 0.0155 and 0.0770, respectively.

Using the learned filter, we tried to extract the eyes from the faces of 10 females and 10 males, whose expressions involve standard, smile, anger and scream. The constant $R$ in the extraction algorithm was chosen as $R = 0.4$. Figures 3 and 4 show the experimental results for the females, and for the males, respectively. We see from Figures 3 and 4 that many of the narrow eyes have been extracted. In several examples, a part of the large eyes has been detected, because it is similar to that of the training narrow eyes. The learned filter never extracts the closed eyes. However, the eyes of persons wearing the glasses can not be extracted enough using the present learned filter. The learned filter is a recognizer of narrow eyes.

We also tested our algorithm for some illuminated faces, which contain large eyes. The experimental results are shown in Figure 5 for the females, and in Figure 6 for males. Since a part of the training narrow eyes is similar to that of large eyes, almost all large eyes have been extracted, independent of illumination change. The learning time was 1 msec and the detection time was 0.1 msec per face, by using the laptop computer with Pentium M, 1.1GHz.

For comparison, we carried out numerical experiments using the variance-maximization method proposed in (Niijima, 2005) for the same faces shown in Figures 3 through 6. Although the same training and target images were used, eyebrows, lips and teeth as well as the narrow eyes were extracted for many of the faces. For some facial images, the target eyes could not be extracted.

Figure 3: Narrow eyes extracting results for females.

Figure 4: Narrow eyes extracting results for males.

## 7 CONCLUSION

We have proposed a facial parts recognition method. The method is based on the kurtosis-minimization learning of the lifting wavelet filters. Our learning and recognition algorithms are very fast, because only one set of free parameters is learned and only one pair of lifting wavelet filters with the learned parameters is applied to a query image for finding facial parts similar to the target facial part.

In simulation, we succeeded to extract and recognize the narrow eyes for almost all facial images. The learned filter never extracts the closed eyes. Our filter extracts large eyes for illuminated facial images. It is a future work to construct a lifting wavelet filter recognizable the eyes of a person wearing glasses.

## REFERENCES


Abdukirim, T., Niijima, K., and Takano, S. (2005). Design of biorthogonal wavelet filters using dyadic lifting scheme. In *Bulletin of Information and Cybernetics*. Kyushu University.

Jaimes, A., Pelz, J., Grabowski, T., Babcock, J., and Chang, S. (2001). Using human observers' eye movements in automatic image classifiers. In *SPIE Human Vision and Electronic Imaging*.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. Academic Press, London, 2nd edition.

Martinez, A. and Benavente, R. (1998). The ar face database. In *CVC Technical Report 24*. Purdue University.

Niijima, K. (2005). Fast objects detection by variance-maximization learning of lifting wavelet filters. In *SPARS'05, International Workshop on Signal Processing with Adaptive Sparse Structured Representation*.

Pentland, A., Moghaddam, B., and Starner, T. (1994). View-based and modular eigenspaces for face recognition. In *CVPR'94, IEEE Conference on Computer Vision and Pattern Recognition*.

Sweldens, W. (1996). The lifting scheme: A custom-design construction of bi-orthogonal wavelets. In *Applied and Computational Harmonic Analysis*. Elsevier.

Takano, S. and Niijima, K. (2005). Person identification using fast face learning of lifting dyadic wavelet filters. In *CORES'05, 4th International Conference on Computer Recognition Systems*. Springer.

Takano, S., Niijima, K., and Abudukirim, T. (2003). Fast face detection by lifting dyadic wavelet filters. In *ICIP'03, IEEE International Conference on Image Processing*.

Takano, S., Niijima, K., and Kuzume, K. (2004). Personal identification by multiresolution analysis of lifting dyadic wavelets. In *EUSIPCO'04, 12th European Signal Processing Conference*.

Tieu, K. and Viola, P. (2000). Boosting image retrieval. In *CVPR2000, IEEE Conference on Computer Vision and Pattern Recognition*.

Vapnik, V. (1998). *Statistical Learning Theory*. John Wiley & Sons Inc., New York.

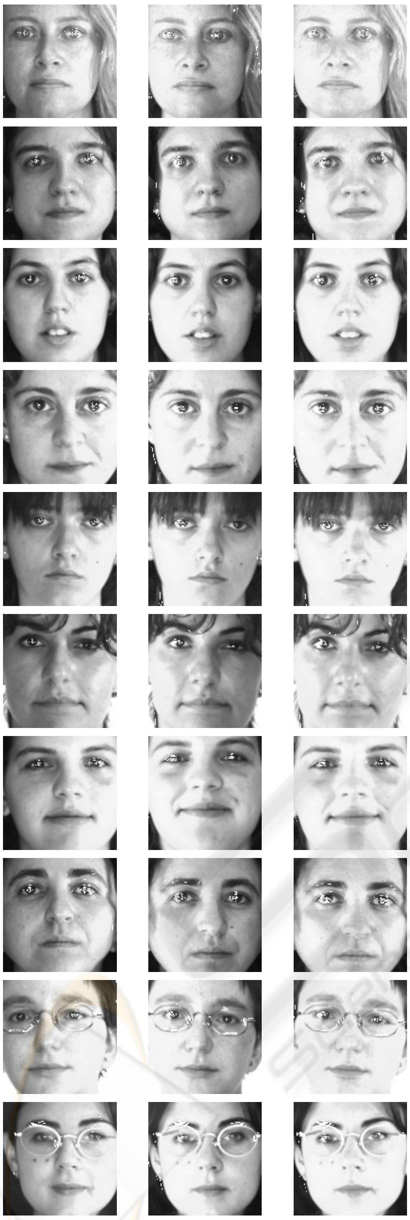


Figure 5: Extraction results for the illuminated faces of females.

Figure 6: Extraction results for the illuminated faces of males.