FAST OBJECTS DETECTION BY VARIANCE-MAXIMIZATION LEARNING OF LIFTING WAVELET FILTERS

Koichi Niijima

Department of Informatics, Kyushu University
6-1, Kasuga-koen, Kasuga, Fukuoka 816-8580, JAPAN
niijima@i.kyushu-u.ac.jp

ABSTRACT

A fast objects detecting method is proposed, which is based on the variance-maximization learning of lifting dyadic wavelet filters. First, it is shown that the sum of lifting wavelet coefficients in horizontal and vertical directions defines an elliptic-type of discrete operator containing free parameters. The free parameters are learned so as to maximize the variance of lifting wavelet coefficients for a target object. Since this problem is an ill-posed problem, a regularization method is employed to solve it. Objects in a query image similar to the target object are detected by the use of the learned filter. Simulation concerns the detection of narrow eyes from faces.

1. INTRODUCTION

Fast and exact detection of objects in an image is an important problem in the research areas such as computer vision and robot vision. Many objects detection methods have been developed so far [8]. Unlike such detection techniques, we have recently proposed a new detection method using lifting dyadic wavelet filters for person identification [5, 6, 7]. These papers employ a cosine-based learning algorithm of lifting dyadic wavelet filters. Since the goal of these papers lies in person identification, the details of facial parts must be learned. So, we learned several lifting filters at only one location of each facial part so as to detect it exactly. However, the constructed filters are low-pass filters and, therefore, our cosine-based method is not robust for changing brightness.

In this paper, we propose a fast and robust method for detecting objects in an image using the variance-maximization learning of lifting dyadic wavelet filters. It is known in numerical experiments that the histogram of wavelet coefficients for a natural image behaves like supergaussian distribution. Thus, the variance of wavelet coefficients characterizes the image. We utilize this fact to learn free parameters contained in a lifting filter.

First, the sum of lifting dyadic wavelet filters in horizontal and vertical directions is shown to be a discretization scheme for an elliptic-type of partial differential operator containing free parameters. We learn the free parameters so as to maximize the variance of lifting wavelet coefficients for a target object in a training image. In the learning process, we put the condition that a vector of lifting filters has the norm 1 and its each component becomes a high-pass filter. We also put the condition that lifting wavelet coefficients in an image region except for the target object become small. Under these conditions, a functional to be minimized is derived for learning the free parameters. However, this minimization problem is an ill-posed problem. So, we derive a new functional to be minimized by adding a regularization term to the original functional. We can use the steepest descent method to minimize the functional. However, this method is time-consuming to obtain the convergence results. To realize fast learning, we derive a system of simultaneous nonlinear equations by differentiating the functional, and solve the system exploiting Newton's method.

Objects in a query image similar to the target object is detected by applying a lifting filter having the learned parameters to the query image. Actually, we extract the location where the absolute value of a lifting wavelet coefficient exceeds the standard deviation of lifting wavelet coefficients computed for the target object.

In simulation, the anger face of a female is used as a training image. The target object is her right eye. A lifting filter including the learned parameters is applied to various 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. In Section 4, we describe a detection method. Section 5 is simulation. Finally, we conclude with Section 6.

2. CHARACTERIZATION OF LIFTING DYADIC WAVELET FILTERS

Let $\{h^o_n, g^o_n, \tilde{h}^o_n, \tilde{g}^o_n\}$ be a set of dyadic wavelet filters [2]. The filters h^o_n and g^o_n are called low-pass and high-pass analysis filters, respectively, and the filters $\tilde{h}^o_n$ and $\tilde{g}^o_n$ are low-pass and high-pass synthesis filters, respectively. A lifting scheme for the dyadic wavelet is described as follows:

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

This scheme generalizes Sweldens' biorthogonal lifting scheme [4]. We proved that the lifted filters $\{h_n, g_n, \tilde{h}_n, \tilde{g}_n\}$ also become a set of dyadic wavelet filters [1]. Here λ_k's denote free parameters. In this paper, we only use the lifted filter (1).

Let us denote an image by $u_{i,j}$. Applying the low-pass analysis filter h_n^o in vertical direction to $u_{i,j}$, we 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

$$D_{m,k} = \sum_j g_i^d C_{m+i,k}^{col}. \quad (2)$$

Here g_i^d's are given by

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

where λ_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

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

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:

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

where λ_l^e's represent free parameters in vertical direction.

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

$$H_{m,k} = D_{m,k} + E_{m,k}.$$

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

$$H_{m,k} = \sum_i g_i^o C_{m+i,k}^{col} + \sum_j g_j^o C_{m,k+j}^{row} - \left(\sum_{l=-L}^{L} \lambda_l^d C_{m+l,k} + \sum_{l=-L}^{L} \lambda_l^e C_{m,k+l} \right) \quad (4)$$

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

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

Here $I_y u$, $I_x u$ and $I(\lambda^d, \lambda^e)u$ indicate the integral versions of $C_{m,k}^{col}$, $C_{m,k}^{row}$ and the last term of (4), respectively, and

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

3. LEARNING ALGORITHM

We describe a method for learning free parameters λ_l^d and λ_l^e for objects detection. First, we express $H_{m,k}$ in (4) in an inner product form. Let us define a vector $\mathbf{g}$ whose components are the lifting filters g_i^d and g_j^e as

$$\mathbf{g} = (g_{-L-M}^d, ..., g_{L+M+1}^d, g_{-L-M}^e, ..., g_{L+M+1}^e)$$

and a vector $\mathbf{C}_{m,k}$ whose components consist of the low-pass components $C_{m+i,k}^{col}$ and $C_{m,k+j}^{row}$ as

$$\mathbf{C}_{m,k} = (C_{m-L-M,k}^{col}, ..., C_{L+M+1,k}^{col}, C_{m,k-L-M}^{row}, ..., C_{m,k+L+M+1}^{row}).$$

Using these vectors, $H_{m,k}$ can be written as

$$H_{m,k} = \mathbf{g} \cdot \mathbf{C}_{m,k},$$

where $\cdot$ denotes inner product.

Let us denote a training image also by $u_{i,j}$, and its domain by Ω_d. By ω_d, we denote a target region which is a subimage of $u_{i,j}, (i,j) \in \Omega_d$. The number of pixels in ω_d is denoted by P. We call $u_{i,j}$ for $(i,j) \in \omega_d$ positive data, and $u_{i,j}$ for $(i,j) \in \Omega_d \setminus \omega_d$ negative data. Using these positive and negative data, we learn free parameters λ_l^d and λ_l^e appeared in (4).

The variance of $H_{m,k}$ for the positive data is given by

$$\sigma^2 = \frac{1}{P} \sum_{(m,k)\in\omega_d} (\mathbf{g} \cdot (\mathbf{C}_{m,k} - \bar{\mathbf{C}}))^2. \quad (5)$$

Here $\bar{\mathbf{C}}$ denotes the average of $\mathbf{C}_{m,k}$. We extend the target object periodically in horizontal and vertical directions, and impose the condition

$$\sum_{l=-L}^{L} (\lambda_l^d + \lambda_l^e) = 0. \quad (6)$$

Then, we can prove

$$\mathbf{g} \cdot \bar{\mathbf{C}} = 0. \quad (7)$$

This means that $\mathbf{g}$ is a vector of high-pass filters, and that (5) can be written as

$$\sigma^2 = \frac{1}{P} \sum_{(m,k)\in\omega_d} (\mathbf{g} \cdot \mathbf{C}_{m,k})^2. \quad (8)$$

For the negative data, it is desirable to minimize

$$\frac{1}{N} \sum_{(m,k)\in\Omega_d\setminus\omega_d} (\mathbf{g} \cdot \mathbf{C}_{m,k})^2, \quad (9)$$

where N is the number of pixels in $\Omega_d \setminus \omega_d$.

Our learning algorithm for free parameters λ_l^d and λ_l^e is to maximize (8) and to minimize (9) subject to the condition (6) and normalization of the vector $\mathbf{g}$,

$$|\mathbf{g}| = 1. \quad (10)$$

The continuous version of this problem is to maximize

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

and to minimize

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

under the constraint (6) and an integral version of (10). Here Ω is a domain corresponding to Ω_d, and ω a domain corresponding to ω_d. This problem is an ill-posed problem and, therefore, the discrete version is also an ill-posed problem, which yields unstable solutions. To overcome this difficulty, we use a regularization method. Thus, a functional to be minimized is

$$J = -\frac{1}{2P}\sum_{(m,k)\in\omega_d} (\mathbf{g}\cdot\mathbf{C}_{m,k})^2 + \frac{K_0}{2N}\sum_{(m,k)\in\Omega_d\setminus\omega_d} (\mathbf{g}\cdot\mathbf{C}_{m,k})^2 + \frac{K_1}{4}(|\mathbf{g}|^2-1)^2 + \frac{K_2}{2}\left(\sum_{l=-L}^{L}(\lambda_l^d+\lambda_l^e)\right)^2 + \delta\sum_{l=-L}^{L}\left((\lambda_l^d)^2+(\lambda_l^e)^2\right). \qquad (11)$$

Here the third and fourth terms of the right hand side come from (10) and (6), respectively, and K_0, K_1 and K_2 are penalty constants. The last term means regularization and δ is a sufficiently small positive number.

The functional (11) has a possibility of having many local minima. Since it is difficult to gain a global minimum, we seek local minima. The most popular technique for obtaining local minima is the steepest descent method. However, this method needs a lot of time for finding local minima. 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:

$$\frac{\partial J}{\partial \lambda_l^d} = 0, \qquad l = -L, ..., L, \qquad (12)$$

$$\frac{\partial J}{\partial \lambda_l^e} = 0, \qquad l = -L, ..., L. \qquad (13)$$

The process of solving (12) and (13) by Newton's method gives our learning algorithm of free parameters contained in the lifting dyadic wavelet filter.

4. DETECTION ALGORITHM

Our detection 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.
2. By combining them with the parameters λ_l^d and λ_l^e learned for the target object, compute the lifting wavelet coefficients $D_{m,k}$ and $E_{m,k}$ defined by (2) and (3).
3. Calculate the sum $H_{m,k} = D_{m,k} + E_{m,k}$.
4. Find the locations (m,k) such that $|H_{m,k}| \geq R\sigma$, where σ is the standard deviation of lifting wavelet coefficients computed for the target object and R denotes some constant.
5. Detect an image region, in which the extracted locations are concentrated, as an object similar to the target object.

5. SIMULATION

Numerical experiments are carried out using the AR face database, which was provided by Martines [3].

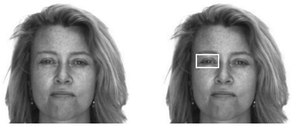
The initial filters we use are the cubic spline dyadic wavelet filters listed in Table 1 [2]. The number of free parameters in each

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	

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 $[0, 10]$ in our experiments. The training pattern is the image of a female having 210×200 size, shown in Figure 1(a). It is an anger face whose eyes are narrow. As a target object, we extract the right eye with 32×17 size from the face, which corresponds to positive data. It is shown in Figure 1(b). Negative data are large wavelet coefficients in an image region except for the target object.



(a) (b)

Fig. 1. (a) Training image, (b) Target object

The penalty constants K_0, K_1 and K_2 appeared in (11) are chosen as $K_0 = 2.0$ and $K_1 = K_2 = 1000$, respectively. The regularization coefficient δ is selected as $\delta = 0.01$. Newton's iteration for solving (12) and (13) was started from 0 vector. We list the learned parameters in Table 2. The standard derivation σ for the target is $\sigma = 2.375139$.

We tested the detection algorithm for the faces of 3 females and 3 males, whose expressions involve standard, smile, anger and scream. The constant R in the detection algorithm was chosen as $R = 1.2$. Figure 2 shows the detection results. We see from Figure 2 that almost all narrow eyes have been detected. However, some eyebrows and teeth have also been extracted by mistake. We also tried to detect objects similar to the target object for some illuminated faces. The experimental results are shown in Figure 3. Some eyes have been detected independent of illumination change.

The learning time was 1 msec and the detection time was 0.1 msec per face, by using laptop computer with Pentium M, 1.1GHz.

6. CONCLUSIONS AND FUTURE WORKS

We have proposed a fast and robust method capable of detecting objects in an image. The method is based on the variance-maximization learning of free parameters in a lifting dyadic wavelet

Table 2. Parameters learned for the training image illustrated in Figure 1(a)

l	λ_l^d	λ_l^e
-7	0.080625	1.479114
-6	-0.836936	-1.974234
-5	1.407438	2.260381
-4	-1.193815	-1.712079
-3	-0.601544	0.229184
-2	2.752888	1.262573
-1	-3.132490	-2.721202
0	-0.637861	0.619456
1	7.083106	2.487422
2	-8.072976	-3.435716
3	3.840926	0.816744
4	0.946699	0.562087
5	-3.284445	-0.287910
6	2.515891	-0.785364
7	-0.954539	1.301003

filter. Our learning and detecting algorithms are very fast, because only one set of free parameters is learned and only one lifting filter with the learned parameters is applied to a query image for finding objects similar to the target object.

We succeeded to detect narrow eyes from a variety of faces independent of illumination change in simulation. However, some other image regions have also been extracted by mistake. This situation changes depending on the learned parameters.

In our approach, we only use the variance of a target object, which is the second order statistics. We need to utilize higher order statistics for achieving exact detection. This is a future work.

Fig. 2. Detection results for the faces of standard, smile, anger and scream

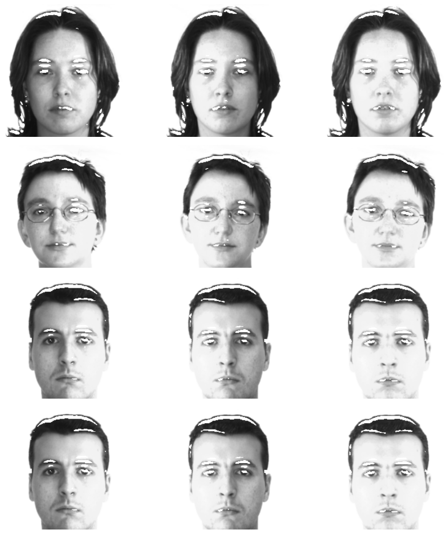


Fig. 3. Detection results for some illuminated faces

7. REFERENCES

[1] T. Abdukirim, K. Niijima, and S. Takano, Design of biorthogonal wavelet filters using dyadic lifting scheme, Bull. of Information and Cybernetics, in press.

[2] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[3] A.M. Martinez and R. Benavente, The AR Face Database, CVC Technical Report 24, Purdue University, 1998.

[4] W. Sweldens, The lifting scheme: A custom-design construction of bi-orthogonal wavelets, Applied and Computational Harmonic Analysis, Vol. 3, pp.186-200, 1996.

[5] S. Takano, K. Niijima, and T. Abudukirim, Fast face detection by lifting dyadic wavelet filters, Proc. of the IEEE International Conference on Image Processing, pp.893-896, 2003.

[6] S. Takano, K. Niijima, and K. Kuzume, Personal identification by multiresolution analysis of lifting dyadic wavelets, Proc. of the 12th European Signal Processing Conference, pp.2283-2286, 2004.

[7] S. Takano and K. Niijima, Person identification using fast face learning of lifting dyadic wavelet filters, Proc. of the 4th International Conference on Computer Recognition Systems, Soft Computing Series, Springer, pp.815-823, 2005.

[8] M.H. Yang, D.J. Kriegman, and N. Ahuja, Detecting faces in images: A survey, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.24, No.1, pp.34-58, 2002.