

Spatial-frequency Image Denoising for Face Recognition

Jianlin Chen, Gaoyong Luo*, Fasheng Zhou, Haitao Cao

School of Physics and Electronic Engineering, Guangzhou University, Guangdong 510006, China

*gyl@gzhu.edu.cn

Keywords: Face Recognition, Wiener Filter, Kalman Filter, Wavelet Filter, Neural Network

Abstract: Image denoising has played an important role in face recognition. Face features are more associated with edge detection. However, noise reduction methods such as adaptive Wiener or Kalman filtering based on minimizing mean square error achieve suboptimal results but at the same time blur image by filtering out high frequency contents related to edge. This paper presents a spatial-frequency image filtering method by wavelet decomposition to achieve better edge preservation while reducing noise significantly. Experimental results show that the proposed method performs better than other adaptive filtering methods for feature extraction of face recognition by neural network with multilayer perceptron.

1. Introduction

Image face recognition (FR) has been widely used for the Internet of Things (IoT) applications such as smart home and smart buildings. Home automation by IoT technology has now been developed and implemented, where face recognition delivers a high level of security and privacy protection [1, 2, 3]. Generally speaking, face recognition algorithm includes five aspects: face images collection, images pre-process, face detection and location, face feature extraction and recognition [4, 5]. When face images are collected [6], noise will inevitably be added into the images. The added noise will influence the subsequent steps and degrade the performance of FR model dramatically [7, 8, 9]. Hence image denoising, which acts as an important pre-processing step, is required in order to obtain a good accuracy when the FR tasks are conducted with face feature extraction, where image features are more associated with edge detection [10].

Two classic and simple spatial image denoising methods are commonly used to process the noisy images: median filter and mean filter. Their performance of smoothing noise usually can not meet the requirements of high FR accuracy, due to the fact that images will be totally blurred after filtering. The key to denoising polluted face images is to reduce the noise while preserve the details related to high frequency contents and edges [11]. In order to obtain a better performance of noise filtering, we investigate two Minimum Mean Square Error (MMSE) based denoising methods: adaptive Wiener filter and Kalman filter. They both process signal in either the time domain or frequency domain, and their performance of denoising are generally good but can not achieve high FR accuracy required for security protection. In this paper, we propose a spatial-frequency image denoising algorithm based on wavelet decomposition to better preserve face feature with a popular neural network (NN) model with multilayer perceptron (MLP) trained to recognize the denoised face images. Some powerful strategies [12, 13, 14] have been used to train the MLP model.

2. Algorithms

Wiener and Kalman filtering approaches are all based on minimizing mean square error (MMSE) [15, 16]. The MMSE method is generally good at suppressing high error but not sensitive at low error values.

2.1. Wiener Filter

The principle of designing a Wiener filter is to find a system function $h(n)$ which has the minimum mean square error between the output signal $y(n)$ and the desired signal $d(n)$ [17]. The

output signal by filtering input signal $x(n)$ is

$$y(n) = x(n) * h(n) = \sum_{m=0}^{+\infty} h(m)x(n-m) \quad (1)$$

and the error is

$$e(n) = d(n) - y(n) \quad (2)$$

The mean square error can be written by mathematical expectation as

$$E[|e(n)|^2] = E[|d(n) - y(n)|^2] \quad (3)$$

To minimize $E[|e(n)|^2]$, we have

$$\frac{\partial E[|e(n)|^2]}{\partial h_j} = 0 \quad (4)$$

Thus we obtain

$$\mathbf{h} = \mathbf{R}_{xx}^{-1} \mathbf{R}_{xd} \quad (5)$$

Where \mathbf{R} is correlation function. If the cross-correlation function between the desired signal $d(n)$ and the input signal $x(n)$, and the auto-correlation function of $x(n)$ are known, then the optimal solution of the Wiener filtering can be achieved. So to design a Wiener filter we need to estimate the desired signal $d(n)$. Estimating the desired signal thus will be the first step. The more accurate the estimated desired signal is, the better denoising performance the Wiener filter can achieve.

2.2. Kalman Filter

Kalman filter is based on state space, where the concept of state variable is introduced [18]. If x_k is the state variable of a system in moment k , then the state equation and measurement equation can be written as

$$x_{k+1} = A_k x_k + \omega_k \quad (6)$$

$$y_k = C_k x_k + v_k \quad (7)$$

Where y_k is observation sample, ω_k is system noise and define its variance matrix as Q_{k-1} , v_k is observation noise and define its variance matrix as R_k , and state-transition matrix as A_k , observation matrix as C_k . The process of Kalman filtering can be expressed as follows:

1) Calculate P'_k

$$P'_k = A_k P_{k-1} A_k^T + Q_{k-1} \quad (8)$$

2) Calculate gain matrix H_k

$$H_k = P'_k C_k^T (C_k P'_k C_k^T + R_k)^{-1} \quad (9)$$

3) Compute state estimate \hat{x}_k

$$\hat{x}_k = A_k \hat{x}_{k-1} + H_k (y_k - C_k A_k \hat{x}_{k-1}) \quad (10)$$

4) Update P_k

$$P_k = (I - H_k C_k) P'_k \quad (11)$$

5) Repeat step 1) - 5).

Kalman filtering method requires that the initial state first be set, and that would determine the time of the system when it gets stable. The performance of filtering depends on the choice of the parameters like state-transition matrix A_k , system noise ω_k and observation noise v_k . To make the filtering effective, it is important to estimate the initial parameters correctly.

2.3. Wavelet Filter

Wavelet filter is based on time-frequency analysis by frequency decomposition. Wavelet

decomposition is a kind of series expansion method which acts on signal $f(t)$, with scaling functions $\varphi_{j,k}(t)$ and wavelet functions $\Psi_{j,k}(t)$. The signal expansion equation can be written as

$$f(t) = \sum_k c_{j0}(k) \varphi_{j0,k}(k) + \sum_k \sum_{j=j_0}^{\infty} d_j(k) \Psi_{j,k}(k) \quad (12)$$

where,

$$c_j(k) = \langle f(t), \varphi_{j,k}(t) \rangle = \int f(t) \varphi_{j,k}(t) dt \quad (13)$$

$$(k) = \langle f(t), \Psi_{j,k}(t) \rangle = \int f(t) \Psi_{j,k}(t) dt \quad (14)$$

This is the form of discrete wavelet transform (DWT), where $c_j(k)$ and $d_j(k)$ represent the low frequency components and high frequency components of the signal $f(t)$, respectively. Instead of removing all of the high frequency parts by generally low pass filtering such as Wiener and Kalman filtering discussed, most of the high frequency details can be preserved while noise can be suppressed by processing the coefficients $d_j(k)$ in the time-frequency wavelet domain, where the denoising method is advantageous over other time or frequency domain filtering methods [19, 20].

The basic principle of wavelet transform based denoising algorithm proposed in this paper is divided into two steps thresholding: at first step we obtain the wavelet coefficients of different scales after wavelet decomposition. The wavelet coefficients whose values are lower than the thresholds are considered as noise to filter out while the others are replaced by the difference between the coefficients and the thresholds for general denoising. We rewrite $x_j(n)$ as the sum of the original signal $s_j(n)$ and noise $v_j(n)$ at scale j

$$x_j(n) = s_j(n) + v_j(n) \quad (15)$$

At this step, the thresholds are obtained from

$$\lambda_j = \frac{\sigma_{x_j}}{\sigma_{s_j}} \quad (16)$$

Where σ_{s_j} , σ_{x_j} are the standard deviation of original signal $s_j(n)$ and noise added signal $x_j(n)$, respectively. The next step is to derive an optimal denoising process. Define $y_j(n)$ as the the denoised signal, then for the thresholding filter $\hat{\lambda}_j$ we have

$$\hat{\lambda}_j x_j(n) = y_j(n) \quad (17)$$

The error function is $e_j(n) = s_j(n) - y_j(n)$ and its mean square value can be expressed as

$$\begin{aligned} E[|e_j(n)|^2] &= E[|s_j(n) - y_j(n)|^2] = E[|s_j(n) - \hat{\lambda}_j x_j(n)|^2] \\ &= E[s_j^2(n) - 2\hat{\lambda}_j s_j(n) x_j(n) + \hat{\lambda}_j^2 x_j^2(n)] \end{aligned} \quad (18)$$

To minimize the mean square error, we have

$$\begin{aligned} \frac{\partial E[|e_j(n)|^2]}{\partial \lambda_j} &= \frac{\partial E[s_j^2(n) - 2\hat{\lambda}_j s_j(n) x_j(n) + \hat{\lambda}_j^2 x_j^2(n)]}{\partial \lambda_j} \\ &= E[-2s_j(n) x_j(n) + 2\hat{\lambda}_j x_j^2(n)] = 0 \end{aligned} \quad (19)$$

The solution of equation (19) is

$$\hat{\lambda}_j = \frac{E[s_j(n)]}{E[x_j(n)]} \quad (20)$$

It follows

$$\hat{\lambda}_j = \sqrt{\frac{E[|s_j(n)|^2]}{E[|x_j(n)|^2]}} \quad (21)$$

From equation (15), we have

$$\hat{\lambda}_j = \sqrt{\frac{E[|s_j(n)|^2]}{E[|s_j(n)+v_j(n)|^2]}} \quad (22)$$

Since $s(n)$ is not related to $v(n)$, we obtain the optimal wavelet filter

$$\hat{\lambda}_j = \sqrt{\frac{E[|s_j(n)|^2]}{E[|s_j(n)|^2 + |v_j(n)|^2]}} = \sqrt{\frac{E[|s_j(n)|^2]}{E[|s_j(n)|^2] + E[|v_j(n)|^2]}} = \sqrt{\frac{\sigma_{s_j}^2}{\sigma_{s_j}^2 + \sigma_{v_j}^2}} \quad (23)$$

Where the variance $\sigma_{s_j}^2$ is estimated by calculating every 5 wavelet coefficients centered on a sliding window at scale j and $\sigma_{v_j}^2 = \sigma_{x_j}^2 - \sigma_{s_j}^2$, $\sigma_{x_j}^2$ is estimated by calculating every 5 wavelet coefficients centered on a sliding window at scale j after general denoising.

3. Experiments

In order to evaluate the performance of different denoising algorithms, experiments were conducted based on Matlab platform. At first, we used Wiener filter, Kalman filter and wavelet filter to process the noisy 1D sinusoidal signal. Then we implemented the three algorithms on polluted face images as well, which is a 2D spatial-frequency signal. At last, we built a simple face dataset and recognized the denoised face images with different signal to noise ratios (SNR). The face recognition experiments were implemented on Tensor Flow framework [21], with MLP model.

3.1. Sinusoidal Signal Denoising

We generated a sinusoidal signal whose frequency is 1Hz, and amplitude is a unit, as shown in figure 1(a). Then the Gaussian noise (shown in figure 1(b)) whose mean is 0 and variance is 0.05 was added to the original sinusoidal signal. The noisy signal is shown in figure 1(c).

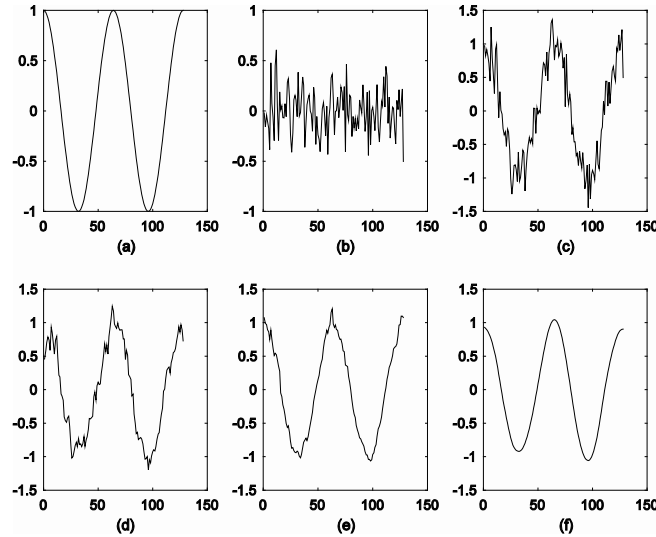


Fig. 1 The denoising experiments on the 1D sinusoidal signal.

Table 1 The SNR improvement results of three denoising methods implemented on 1D sinusoidal signal.

Methods	SNR	Improvement
Noisy signal	10.7058	-
Wiener	14.6038	36.41%
Kalman	21.7428	103.09%
Wavelet	23.5383	119.86%

The simulation result of Wiener filtering, Kalman filtering and wavelet denoising algorithm is shown in figure 1(d), (e), (f), respectively. Obviously, the time-frequency wavelet analysis method achieves the best performance and its denoised signal is very close to the original signal. While Kalman filter roughly recovers the shape of sinusoidal signal but still contains a little noise. However, Wiener filter achieves the worst performance, where noise is reduced but residual is still significant. The results of SNR improvement are shown in table 1, which are in accordance with visual assessment.

3.2. Face Image Denoising

Similarly, we added the Gaussian noise image (shown in figure 2(b)) whose mean is 0, standard deviation is 25 to the original face grayscale image (shown in figure 2(a)) whose resolution is 512×512 . The denoising results are shown in figure 2(d), (e), (f) and the SNR improvements are listed in table 2. From the figures we can see that the denoising performance of Kalman filter seems to be better than wavelet filter. However, the Kalman method cannot deal with the noise on the edge which is the advantage of wavelet denoising algorithm. Wiener method has the worst performance. It is noted that the wavelet decomposition method achieves the best SNR improvement.

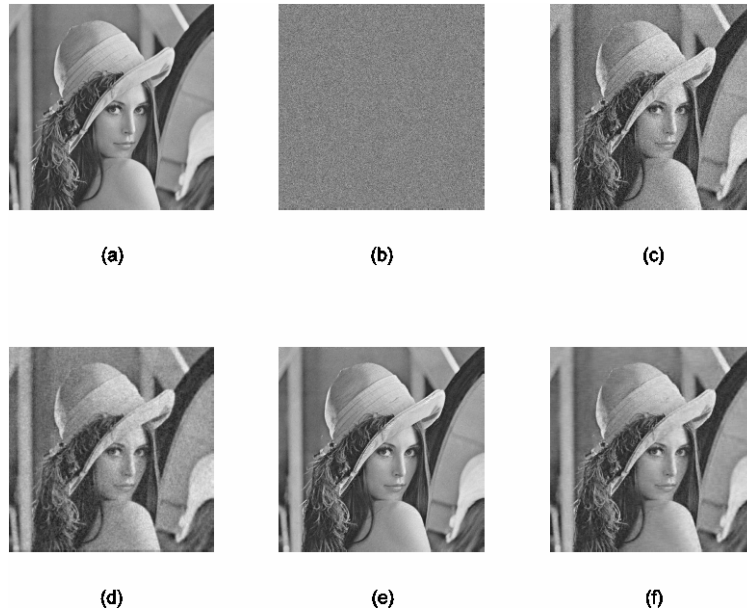


Fig. 2 The denoising experiments on the 2D face image

Table 2 The SNR improvement results of three denoising methods implemented on 2D face image.

Methods	SNR	Improvement
Polluted image	14.8712	-
Wiener	19.6933	32.43%
Kalman	21.5939	45.21%
Wavelet	24.0921	62.01%

Table 3 Hyper-parameters of the MLP model.

Model	Hyper-parameters
Multilayer Perceptron	architecture: [4096 50 30 20 4], batch size: 128, epochs: 500, optimizer: Adam, learning rate: 0.00001, activation function: ReLU batch normalization



Fig. 3 Some face samples of different postures and expressions in the dataset.

3.3. Face Recognition

We built a small face image dataset which contains 2000 samples for 4 people. These samples (such as shown in Figure 3) have different postures, expressions and illumination. Then we took 1400 samples for training and 600 samples for testing to feed into a simple 3-layer MLP model. Some key hyper-parameters of the model are listed in table 3. The back propagation algorithm [22] was used to train the model. Since the dataset is not complicated and the resolution of the face images is merely 64×64 , the MLP model can be easily trained resulting in recognition accuracy up to 98.33%.

Table 4 Recognition results of different denoising algorithms.

Metrics	Methods	Noisy SNR			
		9.1067	13.249	17.656	20.383
Denoised SNR Improvement	Wiener	67.95%	30.73%	10.99%	4.05%
	Kalman	108.66%	50.34%	18.93%	6.66%
	Wavelet	94.79%	66.17%	44.42%	34.48%
Denoised Accuracy	Noisy	78.83%	92.67%	97.17%	96.67%
	Wiener	97.33%	97.33%	97.67%	97.83%
	Kalman	97.33%	97.83%	97.67%	98.00%
	Wavelet	97.83%	98.50%	98.17%	98.33%

As discussed, we added Gaussian noise on the test dataset and performed three denoising algorithms on the polluted images. Then the trained MLP model made prediction on the denoised images. The accuracy results are listed in table 4. From the results we can see that the three denoising algorithms can all recognize the face using the desnoised images, while the wavelet method has the highest recognition accuracy due to spatial-frequency analysis for face feature extraction, and better preservation of high frequency edges. Moreover, the wavelet denoising algorithm can improve the SNR dramatically in the case of high SNR.

4. Conclusion

Spatial-frequency wavelet decomposition method can effectively reduce the noise on the face images and better preserve high frequency contents related to edge and face features. The SNR of polluted images can also be improved significantly by wavelet filtering. Compared with Wiener filter and Kalman filter, wavelet decomposition can be implemented by a faster algorithm developed and proposed in this paper. Experimental results demonstrate that the denoising performance of spatial-frequency wavelet filter achieves higher recognition accuracy than the two filters in either the time or frequency domain. With a neural network (NN) model with multilayer

perceptron (MLP) trained, image face recognition by wavelet denoising can provide a high level of security and privacy protection.

References

- [1] A. Patel, A. Verma: International Journal of Computer Applications Vol.172 (7) (2017), p. 11-17.
- [2] R.K. Kodali, V. Jain, S. Bose, L. Boppana: International Conference on Computing, Communication and Automation IEEE (2017), p. 1286-1289.
- [3] M.H. Asghar, A. Negi, N. Mohammadzadeh: International Conference on Computing, Communication & Automation IEEE (2015), p. 427-431.
- [4] R. Rathi, M. Choudhary, B. Chandra: International Journal of Computer Technology & Applications Vol. 03(01) (2012).
- [5] E. Bagherian, R.W.O.K. Rahmat: International Symposium on Information Technology Vol. 2 (2008), p. 1-9.
- [6] A. Thakur, A. Kausar, A. Iqbal: Cloud System and Big Data Engineering (2016), p. 508-512.
- [7] A. Krizhevsky, I. Sutskever, and G. E. Hinton: Advances in neural information processing systems (2012), p. 1097-1105.
- [8] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner: Proceedings of the IEEE Vol. 86(11) (1998), p. 2278-2324.
- [9] X. Glorot and Y. Bengio: Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (2010), p. 249-256.
- [10] S. Gupta and Meenakshi: International Conference on Computing for Sustainable Global Development (2014), p. 972-976.
- [11] S. Madabusi and S. V. Gangashetty: International Conference on Pattern Recognition (2012), p. 2689-2693.
- [12] S. Ioffe and C. Szegedy: International Conference on Machine Learning (2015), p. 448-456.
- [13] V. Nair and G. E. Hinton: Proceedings of the 27th international conference on machine learning (2010), p. 807-814.
- [14] I. Sutskever, J. Martens, G. Dahl, and G. Hinton: International conference on machine learning (2013), p. 1139-1147.
- [15] M.J. Grimble: Decision and Control, 1982 IEEE Conference on IEEE (1982), p. 221-227.
- [16] R. Singer and P. Frost: IEEE Transactions on Automatic Control Vol. 14(4) (2003), p. 390-394.
- [17] N. Wiener: Mit Press Vol. 113(21) (1949), p. 1043-1054.
- [18] R.E. Kalman: Journal of Basic Engineering Transactions Vol. 82(1960), p. 35-45.
- [19] D.L. Donoho, I.M. Johnstone, G. Kerkyacharian and D. Picard: Journal of the Royal Statistical Society Vol. 57(2) (1995), p. 301-369.
- [20] C.S. Burrus, R.A. Gopinath and H.T. Guo: Introduction to wavelets and wavelet transform: a primer (Pearson Education, Schweiz Ag 1998).
- [21] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard et al: 12th USENIX Symposium on Operating Systems Design and Implementation Vol. 16(2016), p. 265-283.
- [22] D.E. Rumelhart, G.E. Hinton, R.J. Williams et al: Cognitive modeling Vol. 5(3) (1988), p. 1.