2015 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, SEPT. 17–20, 2015, BOSTON, USA

FILTERING OF FREQUENCY-TRANSFORMED IMAGE FOR PRIVACY PRESERVING FACIAL RECOGNITION (DRAFT V1.0)

Author(s) Name(s)

Author Affiliation(s)

ABSTRACT

This paper examines the use of filters on feature vectors for privacy protection in facial recognition. Feature vectors are the results of Fast Fourier Transform and Wavelet Transform on the Yale and Olivetti datasets. Several filters are proposed. Filters based on the signal to noise ratio and t test select feature which prevent privacy compromising reconstruction without sacrificing accuracy. The use of phase removal for FFT and normalization are also shown to protect privacy.

***Index Terms*—** One, two, three, four, five

1. INTRODUCTION

In modern times, the growth of the Internet and digital sensors has lead to the collection of massive amount of personal information. Videos, photos, emails, banking transactions, browsing history, GPS tracks, and other data is collected and store in the cloud This data may be circulated around the Internet and severs without the owner's knowledge. The value and scale of the data creates the risk of privacy leakage. While encryption has been the traditional solution to this problem, recently a novel privacy preserving methodology, compressive privacy, was developed. In the compressive privacy paradigm, the data owner controls the privacy of the data before uploading it to the cloud [need to cite Prof Kung].

As noted in [1], 9/11, Edward Snowden incident and other events have resulted in both a demand for recognition systems and a concern for privacy violation by such systems. A critical competent of such system is biometric identification of people, mainly by facial recognition which can be performed in without consent and at a distnce by suvellance camreas. While a lot of research has gone into improving facial recognition systems - [2], [3], [4], [5], [6] among others - relatively little research has been done on incorporating privacy into such systems; some examples being [7], [8], and [9].

The primary approach to privacy in facial recognition is cryptography. In [7], Eigenfaces recogition system is used on homomorphically encrypted data. In this first cryptographic system for facial recognition [7], the client wants the server to identify a face without revealing the image to the server. The server also does not want to reveal the contents of it's database. In this approach, data is quantized for Pailler encryption and server and client share computations needed for matching faces. [7] Experimentally, 96% accuracy was achieved in [7] on "ORL Database of Faces". [8] improved the algorithm presented in [7] by reducing the time and space complexity with the use of garbled circuits.

Along a different line of research, [9] used Helper Data Systems to provide privacy. The approach generates binary feature vector by determining reliable bits based on statistics from sample data. While the proposed process is similar to compressive privacy, it leaves pirvacy up to the cloud server and [.....]

Our compressive privacy appoach to facial recognition rests on the idea that privacy is compromised when an image can be visually inspected by an unautorized individual. Therefore, as long as the reconstruction of an image from a feature vector is not meaningful to a human, privacy has been maintained. Facial recognition systems often utilize Fast Fourier Transform (FFT) or Wavelet Transform (WT) as part of feature engineering. For example [3], [4], [5], and [6] use FFT or WT. In recognition systems like these, it is possible to alter the output of FFT or WT to reduce the quality of the reconstructed image without sacrificing accuracy of the classification.

2. OUR CLASSIFICATION SYSTEMS

We can break down a facial recognition system into two compnents, feature engineering and classification. Both compenents can be made up of several parts, for example several sequecnail WT transforms. We limit our feature engineering to one transform to keep image reconstruction simple, however this may be a real constraint in time or power sensative applications. As pictured in Figure 1, our classification systems for this investigation, begin with an application of FFT

Thanks to XYZ agency for funding.

978-1-4673-7454-5/15/$31.00

or WT to an image. Then a filter is applied as part of feature selection. Classification is accomplished with an SVM. For all of the following experiments an SVM with a leaner kernel and C = 1 was found to produce the best results.

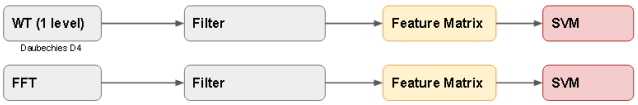


Fig. 1. Block diagrams of classification systems used.

3. FILTERS

Part of our compressive privacy results from the use of filters for feature selection of FFT and WT frequencies. A filter is a binary matrix which selects features based on some selecting function. Let $F_{x,d}$ be a filter $F_{x,d} \in \{0,1\}^{dxm}$ where m is the number of features in each example, x is one of the selecting functions, and d is the number of features selected by the filter.

$F_{x,d,i,j} = 1$ if $W_x(j)$ is with in the d largest W_x, otherwise $F_{d,i,j} = 0$. Thus m dimensional feature vectors are compressed to d dimension feauter vectors, based on the ordering of W_x.

Let X be feature vectors for an example:

$$X_{filtered} = F_{x,d}X \tag{1}$$

Selection functions assign a weight to each feature and divide filters into three categories: data independent, unsupervised, and supervised. Data indepednent filters do not utalize a traning set when determining which features should be selected. In our investigation the Rectangle and Triangle filters are data indepednet as they select a predefined region of features. The variance based filter is unsuperived in the sense that it does not consider labels of the training examples. Feature selection is done purely based on the variance of each feature accross the training set. The idea for this filter originates in [3].

We devised four supervised filters based on the signal to noise ratio, Fisher discriminant ratio, symmetric divergence and t statistical tests. The supervised filters require labels for positive and negative classes. During training, for each individual the training examples are divided into two classes. Positive class contains the pictures of that individual and the negative class contains all other pictures. The signal to noise ratio, Fisher discriminant ratio, symmetric divergence and t tests are computed based on those two classes for each feature, and the final weight for a feature is the mean of weights cross all of the individual in the training set.

The filters are mathematically defined below. For the equations below, let μ_j^+, σ_j^+, and N_j^+ be the mean, standard deviation and number of examples for a target individual and let μ_j^-, σ_j^-, and N_j^- be the mean, standard deviation and number of examples for all other people in the training set. Let $\bar{X}$ be the mean of X with respect to the training set in the case of unsurprised filters, and with respect to all individuals.

3.1. Variance

$$W_{VAR}(j) = \sigma_j^2 \tag{2}$$

3.2. Signal to Noise Ration

$$W_{SNR}(j) = \frac{\overline{|\mu_j^+ - \mu_j^-|}}{\sigma_j^+ + \sigma_j^-} \tag{3}$$

3.3. Fisher Discriminant Ratio

$$W_{FDR}(j) = \overline{\frac{(\mu_j^+ - \mu_j^-)^2}{(\sigma_j^+)^2 + (\sigma_j^-)^2}} \tag{4}$$

3.4. Symmetric Divergence

$$W_{SD}(j) = \overline{\frac{1}{2}\frac{(\mu_j^+)^2}{(\mu_j^-)^2} + \frac{(\mu_j^-)^2}{(\mu_j^+)^2} + \frac{1}{2}\frac{(\mu_j^+ - \mu_j^-)^2}{(\sigma_j^+)^2 + (\sigma_j^-)^2} - 1} \tag{5}$$

3.5. T

$$W_T(j) = \overline{\frac{\mu_j^+ - \mu_j^-}{\sqrt{\frac{(\sigma_j^+)^2}{N_j^+} + \frac{(\sigma_j^-)^2}{N_j^-}}}} \tag{6}$$

3.6. Rectangle

let J be a rectangular region of features

$$W_{Rect}(j) = \begin{cases} 1 & \text{if } j \in J \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

3.7. Triangle

Along with the Rectangle filter, this filter was inspired by results in [3], where features with high varince after FFT are the amplitues of low frequencies and they form a trianglar region.

let J be a triangular region of features

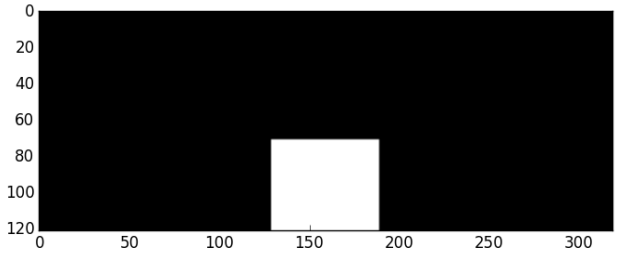


Fig. 2. White pixels represent selected features. The region is indented to select amplitudes of low frequencies in the FFT transform.

$$W_{Tri}(j) = \begin{cases} 1 & \text{if } j \in J \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

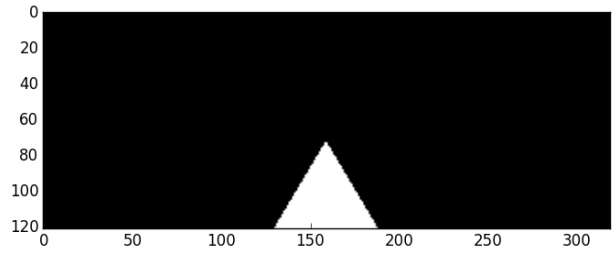


Fig. 3. White pixels represent selected features. The region is indented to select amplitudes of low frequencies in the FFT transform.

4. METHODS

4.1. Classification

Ten fold cross validation was used to test classification accuracies. Grid Search was used to optimize the SVM parameters for each fold. For all of the following experiments an SVM with a leaner kernel and C = 1 was found to produce the best results. 100 cross validations were performed and the training of supervised and unsupervised filters was constrained to the appropriate fold. Since the rectangle and triangle filters are parameterized by a region they were used to determine d for all other filters. For FFT, the filters where only applied to the amplitudes of frequencies. For WT, the filters were applied to all four [.....].

4.2. Reconstruction

To reconstruct original images from filtered feature vectors, we set the values of all filtered features which were filtered out to be zero. The operation can be viewed as muliplication by the transpose of the filter.

$$\tilde{X} = F_{x,d}^{T} X_{filtered} \quad (9)$$

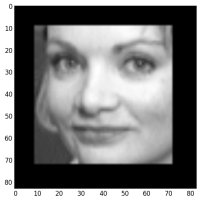
Then the inverse of the FFT or WT is applied on $\tilde{X}$. As mentioned above, only one transform was used at a time to keep this process simple.

5. EXPERIMENTAL RESULTS

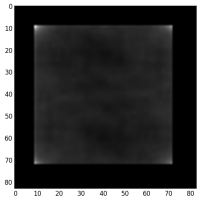
Our baseline accuracy with the FFT transform is 0.741 for the Yale database and 0.975 for the Olivetti database. With the WT transform, our baseline for Yale database is 0.807 and for the Olivetti database is 0.963.

5.1. Phase Removal

FFT produces amplitudes and phases for transformed image. If the phase is removed and the image is reconstructed based on the amplitude the result is a faceless image as seen in Figure 4. This clearly protects privacy. The cost in accuracy is small. The accuracy drops 0.006 to 0.735 for the Yale database and by 0.001 to 0.974 for Olivetti database.



Original Image



Reconstruction

Fig. 4. Result of reconstruction of a FFT transformed image with the phase removed.

5.2. Normalization

The amplitudes and phases of the FFT results for the training set can be normliazed. Normalization is done by computing the mean and standard deviation for each feature and then from each corresponding feature subtracting the mean and dividng the difference by the standard deviation. Performing reconstruction on the amplitudes and phases without undoing the transformation produces privacy as seen in Figure 5. On the Olliveti database the accuracy was 0.975. On the Yale database the accuracy was 0.739.

5.3. Filtering

5.3.1. FFT

Based on accuarcies in Figure 6, the T filter has the best or near best accuarcy across both datasets. Figure 7 shows that it also preserves privacy. Comapored to the other filters, the face

Original Image Reconstructed Image

Fig. 5. Reconstruction from normalized FFT output.

is brealy visible in the reconstrcution from feaures selcted by the T filter.

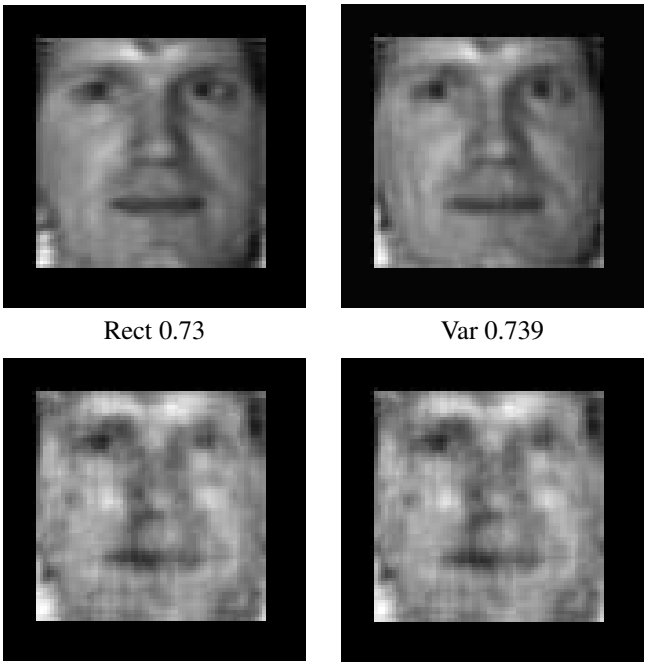
Filter	Yale Acc	Olliveti Acc
Rect	0.737	0.967
Var	0.739	**0.970**
SNR	0.739	0.968
FDR	0.735	0.965
SD	0.686	0.952
T	**0.741**	0.968

Fig. 6. Mean accuracy for the filters based on 100, 10 fold cross validations. Top 399 features are selected.

5.3.2. WT

Using the result of the wavelet transform, the T filter still produces best or second best accuracires across the two datasets.

As seen in Figure 9, the T filter preserves provcary. The reconstructions for Var and T are similar to reconstrcutiosn for the other filters. The privacy protections appears to be a property of the WT tranform rather than any particlar filter.



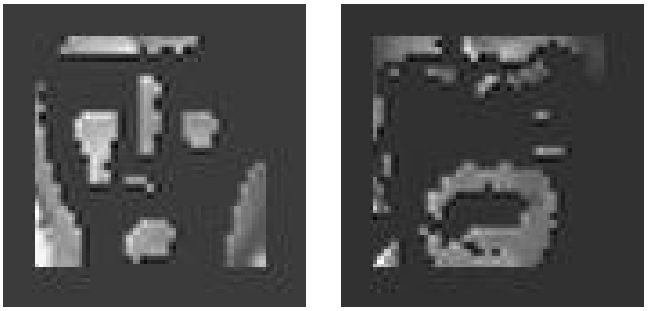
Rect 0.73 Var 0.739

SNR 0.739 T 0.741

Fig. 7. Reconstructed images after a FFT transform and filter $d = 399$. The filter name and accuracy on the Yale dataset is listed under each image.

Filter	Yale Acc	Olliveti Acc
Var	0.737	0.941
SNR	**0.813**	0.968
FDR	0.807	**0.969**
SD	0.672	0.883
T	0.812	**0.969**

Fig. 8. Mean accuracy for the filters based on 100, 10 fold cross validations. Top 399 features are selected.



Var 0.737 T 0.812

Fig. 9. Reconstructed images after a WT transform and filter $d = 399$. The filter name and accuracy on the Yale dataset is listed under each image.

6. DISCUSSION

There is a lot of potential for privacy protection with proper alterations in the feature engineering process. Across all of the methods, the cost, in the form of accuracy loss, is very small. In fact, for the wavelett transform, the accuracy increased when filters where applied. Additionally, the reconstructions from wavelett transform features yieled much better privacy by completely remoeving the eyes and mouth in the image. It appears that the wavelett transform forms a very good basis for privacy preserving feature vectors.

Mathod	From Yale	From Olliveti
Phase Removal	-0.006	-0.001
Normalization	-0.002	0.000
Filtering (FFT)	0.000	-0.006
Filtering (WT)	0.005	0.006

Fig. 10. Changes in mean accuracy relative to the baseline for the four different methods. The T filter is used for filtering accuracies.

While FFT based systems lost accuracy under all of these methods, it is easier to see why the filters preserve accuarcy. The supervised statistical filters select both high and low fequecies as seen in Figure 11. Therefore, the reconstrcution is not just a blurried image as is the case with selecting just the low frequcies. Fewer of those frequcies are selected thus the image is even more blurred, but accuarcy is preserved by the high frequeices. The mix avoid details form being reconstructed fully.

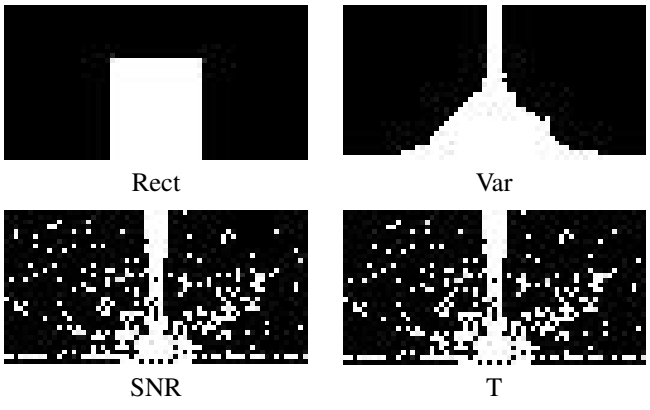


Fig. 11. Amplituedes selected by the filter (marked in white).

Regarding the filters, which should be applied even for performance improvements, SNR and T are the most useful for privacy protection. For both FFT and WT, these filters allow a large dimension reduction with almost no accuracy loss.

In future research consecutive search methods can be utilized to produce filters. Additionally, the reconstruction for WT should be explained in terms of the underlying bands. Lastly, a fully algorithm utilizing the filters and taking into the account the client server model should be developed.

This paper represents my own work in accordance with University regulations.

7. REFERENCES

[1] Kevin W Bowyer, "Face recognition technology: security versus privacy," *Technology and Society Magazine, IEEE*, vol. 23, no. 1, pp. 9–19, 2004.

[2] Anissa Bouzalmat, Jamal Kharroubi, and Arsalane Zarghili, "Comparative study of pca, ica, lda using svm classifier," *Journal of Emerging Technologies in Web Intelligence*, vol. 6, no. 1, pp. 64–68, 2014.

[3] Hagen Spies and Ian Ricketts, "Face recognition in fourier space," in *Vision Interface*, 2000, vol. 2000, pp. 38–44.

[4] Anissa Bouzalmat, Arsalane Zarghili, and Jamal Kharroubi, "Facial face recognition method using fourier transform filters gabor and r_lda," *IJCA Special Issue on Intelligent Systems and Data Processing*, pp. 18–24, 2011.

[5] Zhang Dehai, Ding Da, Li Jin, and Liu Qing, "A pca-based face recognition method by applying fast fourier transform in pre-processing," in *3rd International Conference on Multimedia Technology (ICMT-13)*. Atlantis Press, 2013.

[6] Ahmed Shabaan Samra, Salah Gad El Taweel Gad Allah, and Rehab Mahmoud Ibrahim, "Face recognition using wavelet transform, fast fourier transform and discrete cosine transform," in *Circuits and Systems, 2003 IEEE 46th Midwest Symposium on*. IEEE, 2003, vol. 1, pp. 272–275.

[7] Zekeriya Erkin, Martin Franz, Jorge Guajardo, Stefan Katzenbeisser, Inald Lagendijk, and Tomas Toft, "Privacy-preserving face recognition," in *Privacy Enhancing Technologies*. Springer, 2009, pp. 235–253.

[8] Ahmad-Reza Sadeghi, Thomas Schneider, and Immo Wehrenberg, "Efficient privacy-preserving face recognition," in *Information, Security and Cryptology–ICISC 2009*, pp. 229–244. Springer, 2010.

[9] Tom AM Kevenaar, Geert Jan Schrijen, Michiel van der Veen, Anton HM Akkermans, and Fei Zuo, "Face recognition with renewable and privacy preserving binary templates," in *Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on*. IEEE, 2005, pp. 21–26.