Compressive sensing using optimized sensing matrix for face verification

Endra Oey, Jeffry, Kelvin Wongso, Tommy
Computer Engineering Department, Faculty of Engineering, Bina Nusantara University, Jakarta, Indonesia 11480
endraoey@binus.edu

Abstract. Biometric appears as one of the solutions which is capable in solving problems that occurred in the usage of password in terms of data access, for example there is possibility in forgetting password and hard to recall various different passwords. With biometrics, physical characteristics of a person can be captured and used in the identification process. In this research, facial biometric is used in the verification process to determine whether the user has the authority to access the data or not. Facial biometric is chosen as its low cost implementation and generate quite accurate result for user identification. Face verification system which is adopted in this research is Compressive Sensing (CS) technique, in which aims to reduce dimension size as well as encrypt data in form of facial test image where the image is represented in sparse signals. Encrypted data can be reconstructed using Sparse Coding algorithm. Two types of Sparse Coding namely Orthogonal Matching Pursuit (OMP) and Iteratively Reweighted Least Squares - ℓp (IRLS - ℓp) will be used for comparison face verification system research. Reconstruction results of sparse signals are then used to find Euclidean norm with the sparse signal of user that has been previously saved in system to determine the validity of the facial test image. Results of system accuracy obtained in this research are 99% in IRLS with time response of face verification for 4.917 seconds and 96.33% in OMP with time response of face verification for 0.4046 seconds with non-optimized sensing matrix, while 99% in IRLS with time response of face verification for 13.4791 seconds and 98.33% for OMP with time response of face verification for 3.1571 seconds with optimized sensing matrix.

Keywords: Face Biometric, Compressive sensing, sensing matrix, face verification system.

1. Background
Secure data storage is an important requirement in these days. To ensure that important data stored in computer is not accessible to anyone other than user himself or herself, the most common technique used is providing a password. But as time flows by, the usage of password has become more and more unsafe. This is because it has a greater risk to be stolen or hacked. In addition, the ease of forgetting and recalling passwords that have been used, especially if the passwords are varying to each electronic device may also become a weakness in using a password. On the other hand, there are many users who tend to use the same password for every device they use, but this will decrease the levels of data security [1].
Biometric is a way to capture the physical characteristics and behavior of an individual that can be used in the identification process. Biometric is able to solve the problems stated above because the biometric information cannot be removed, forgotten and estimated easily. So that biometric bring convenience to users, there is nothing need to be remembered and taken. Biometric system is basically a recognition system that works by extracting a set of biometric features of an individual and comparing the data with the data stored in the database [3]. There are several types of biometric available today, such as fingerprint, voice, face, hand geometry, eye geometry, iris and ear geometry recognition. The biometric system can be either identification or verification. This depends on the application. In verification mode, biometric data that is obtained are compared with user's own biometric stored inside the database to validate a person's identity. Identity verification aims for positive recognition or prevents one identity to be used by more than one person [4].

This research focuses on one of the biometric, the face verification by using Compressive Sensing (CS) to generate facial image data that is compressed and sparse coding in the form of Orthogonal Matching Pursuit (OMP) and iteratively reweighted Least Square - ℓp (IRLS-ℓp) to reconstruct and calculate the coefficients of sparse contained in the user's face image to compare with the coefficient values that already exist in the database to determine whether the input face image is valid or not. Based on research that has been done before [4], the level of accuracy obtained from non-optimized systems reached more than 80% and for optimized is able to reach more than 90%. In this research, by comparing Gradient Descent used previous research [5] with different sensing matrix optimization algorithms based on equiangular tight frame [6] but uses two kinds of sparse coding, namely Orthogonal Matching Pursuit (OMP) [7] which is the same as the previous research and IRLS-ℓp for comparison that is expected to reach a better level of accuracy [8].

2. Compressive Sensing Theory
Compressive Sensing (next will be referred as CS) was first introduced by Donoho [9]. In the conventional signal sampling process, the process of sampling performed on the input signal must satisfy the Shannon-Nyquist theorem, in which the sampling rate must reached at least two times greater than the maximum frequency of the input signal (called the Nyquist-rate). With the CS technique, the sampling process can still be made even if the above condition is not met. Based on CS theory [10], it could be happened because the signal is converted into a domain which has a sparse representation. Then the signals are reconstructed from the samples as results of the sampling process. Furthermore, the CS technique can provide relief in terms of computational load on the hardware because the CS technique only senses the signal that has been compressed, which is different to the traditional method where the input signal is being sampled first before compression is done so that most of the samples are discarded [10].

In CS theory, a compressed signal can be obtained by projecting the original signal \( x \in \mathbb{R}^{N \times 1} \) using a sensing matrix \( \Phi \in \mathbb{R}^{M \times N} \) where the dimensions of the matrix's row is smaller than the dimensions of matrix's column \( M \ll N \) which fulfill the condition where dictionary and sensing matrix do not depend on each other (incoherence). Sensing matrix \( \Phi \in \mathbb{R}^{M \times N} \) is a matrix which is used to project original signal to a compressed signal while the dictionary \( \Psi \in \mathbb{R}^{N \times K} \) with \( K \gg N \) is used to represent signal \( x \) as a linear combination of columns \( \Psi \cdot \theta \):

\[
x = \sum_{i=1}^{K} \theta^i \Psi_i \Delta \Psi \theta
\]

(1)

where \( \theta \in \mathbb{R}^{K \times 1} \) is a vector coefficients. CS requires that the signal \( x \) is sparse which means \( \| \theta \|_0 \ll N \), where \( \| \theta \|_0 \) denotes the number of non-zero elements in \( \theta \). Given the sensing matrix and dictionary, the sparse coefficient can be obtained through a sparse coding and the signal can be recovered by using (1). Random matrix (random matrix) is often used in early development as a
sensing matrix because of its high probability to not depend or incoherence of the dictionary which is widely used in general [11]. However, the sensing matrix can further be optimized by minimizing the averaged mutual coherence of the equivalent dictionary. In this paper, sensing matrix optimization algorithms based on equiangular tight frame [6] is used.

3. Method and Experimental Design

Fig. 1 describes the process of face verification system using compressive sensing with optimized sensing matrix. Discrete cosine transform is used as dictionary and sensing matrix used in CS process in client laptop or reconstruction in server laptop is random sensing matrix with addition of a key and has go through the optimization process using algorithm in [6]. It needs to send request to server laptop to send dictionary as in the client laptop does not have dictionary. After server receives the request, it will send the dictionary to the client. Image is taken by a camera with a size of 152x152 pixels is converted to 8 bit gray scale with double data type to simplify the calculation. The face images are then go through blocking process with each block sized 8x8 pixels so that there will be 361 blocks. Compressive sensing will be performed to those blocks by involving a sensing matrix that is generated by the Client laptop which has been added with a key that is the combination of time such as year, month, day, hour and minute which was taken during image capture to increase the level of data security. Compressive sensing is done to minimize the size of image dimension without removing important information from it. The result of compressive sensing will be sent to the Client to reconstruct along with the key that is obtained before. After the results has been received by the Server, then image reconstruction will be done by involving a dictionary and sensing matrix that is generated previously along with the key that was sent by Client which will be added into the sensing matrix in the server. Sparse coefficient value from reconstruction will be compared with sparse coefficient value from related user face image to determine the minimum value of euclidean norm from the comparison of test image and 12 related user images in database. The minimum euclidean norm will be compared with the threshold to determine the validity of the particular user.

![Figure 1. Flow chart of face verification system using compressive sensing with optimized sensing matrix.](image-url)
3.1. Determining Threshold

Threshold is the boundary value that is used to determine the validity of user face image from user who tries to access data in system compared to face image from card owner that used in the process of card verification. User is valid if the minimum Euclidean norm from the comparison result of user face image with card owner face image in database is less than or equal to the threshold value. The smaller the value Euclidean norm means that the smaller the difference between the coefficients of sparse reconstruction with sparse coefficient value of the respective card owner of the system. Threshold can be determined by taking 10 data from valid user and his outlier (other user beside that valid user himself) compared to valid user face image in database whereas the minimum value of Euclidean norm taken in each data. In this research, every 1 user has 2 outliers so that there are 30 minimum Euclidean norm value. In user part, 10 minimum Euclidean norm which obtained from the 10 data and then find the maximum Euclidean norm value $U_{\text{max}}$ out of it, while in the outlier, 10 minimum Euclidean norm which obtained from the 10 data and then find the minimum value (Omin) out of it, so that there will be 2 $O_{\text{min}}$ and 1 $U_{\text{max}}$. In this stage, the threshold value can be determined

$$U_{\text{max}} < \text{Threshold} < O_{\text{min}}$$

System testing in this research involve 3 user (1 user has 2 outliers) with 3 expressions (normal, smile and smile with teeth exposed) and 3 lux range which are 200-220, 230-250, and 260-280 (1 expression consists of 3 lux range) in determining the threshold value in respect to one of the parameter which is measured, the light sensitivity. If there are 3 user, for example user A, user B and user C, so when user A is the valid user, then outlier for user A are user B and user C. If user B is the valid user, then the outlier for user A are user B and user C. If user B is the valid user, then the outlier for user A are user B and user C.

3.2. Accuracy and System Response Time

In system testing in respect to accuracy and time response, if 30 data are taken for each expression in 1 lux range (10 data for user valid and 20 data for 2 outliers) in previous testing, then the total data need to be taken for each expression in 1 lux range in this stage of testing are 100 data (20 data valid user and 80 data for 4 outliers). In this stage, total user that will be tested increased from 3 to 5 users. This is done with the aims to increase the validity of the data in respect to other users by using the threshold value that is determined in previous testing. Data which are collected in this stage is not much different from data collection in determining threshold, it is just that Euclidean norm value is not shown in data table anymore, only showing rejection or acceptance. In here, user’s card numbers will be used as reference to determine image from which will be compared is going to be used as comparison image for user image that will be tested. In determining the level of accuracy also involve FAR (False Acceptance Rate) and FRR (False Rejection Rate) where FAR is a condition where valid user is rejected to access system and FRR is a condition where invalid user is accepted to access the system.

Besides that, time measurement is also done either in card verification or face verification process. Time measurement in card verification card is started when user swipe the card to magnetic card reader until client laptop receive response from server laptop which determine the succession of card verification, while for face verification is started after webcam capture user face image until client laptop receive response from server laptop to determine the validity of captured face image.
4. Experiment Results and Discussion

From experiment results for accuracy level testing in optimized IRLS-ℓP system, it can be noticed for the result of accuracy level for each expression from the taken data. For normal expression reach 99% with 1% of FRR (1 face image rejected), smile with teeth exposed expression reach 98% with 2% of FRR (2 face image rejected) and for smile expression reach 100% with 0% of FRR. For whichever expression, there is no FAR. Obtained data are exactly the same as data in non-optimized IRLS-ℓP system for the accuracy, but the euclidean norm obtained in this system is much better. In accuracy level graphic, there are results of both non-optimized and optimized IRLS-ℓP method as comparison. For non-optimized can reach 99% accuracy with 1% of FRR (only 3 face images rejected) and for optimized also reach 99% with 1% of FRR too. For both system, there is no FAR.

The table 1 shows the comparison results of face verification system using OMP and IRLS-ℓP sparse coding.

Table 1. The results of face verification system results using OMP and IRLS-ℓP sparse coding.

| Measurement Variable | OMP       | IRLS- ℓP  |
|----------------------|-----------|-----------|
|                      | Non-Opt   | Opt       | Non-Opt   | Opt       |
| Accuracy (%)         | 96.33     | 98.33     | 99        | 99        |
| Time Response (second) | 0.7364   | 3.4828    | 5.26      | 13.82     |

Face verification using optimized sensing matrix is able to achieve better accuracy compared with non-optimized sensing matrix in OMP sparse coding method, with accuracy of 98.33% compared to 96.33% (by a margin of 2%). On the other hand, IRLS-ℓP is able to achieve same accuracy, with accuracy of 99% on both system even though the computational load (processing time) on optimized sensing matrix is higher or longer. Face verification using IRLS-ℓP is able to generate better accuracy (99%) compared with OMP in sparse coding method. This is clearly visible on the non-optimized and optimized face verification system, but the processing time of both IRLS-ℓP systems take much longer than OMP, with 4.917 second in IRLS-ℓP system and 0.4046 in OMP system for non-optimized system and with 13.4791 second in IRLS-ℓP system and 3.1571 second in OMP system for optimized system. The accuracy of IRLS-ℓP method in non-optimized and optimized system reaches the same value of accuracy, which is 99%. However, IRLS-ℓP method in optimized system is better than IRLS-ℓP method in non-optimized system because the overall value of the Euclidean norm in optimized system is lower than non-optimized system (low value of Euclidean norm indicates that image is much more similar to user’s face image in database) even though computational load is longer. Various of facial expression does not significantly influence the result of accuracy in both non-optimized and optimized system which the maximum difference for all expression is 5% in non-optimized OMP method system, 4% in optimized OMP method system, and 2% for both non-optimized and optimized IRLS-ℓP method system. The threshold value can still be used in the illumination range between 220 – 280 lux, which is 23 for threshold value in OMP method system and 22.5 for both optimized and non-optimized IRLS-ℓP method system.

5. Conclusion

From the results, it showed that the optimized sensing matrix based on equiangular tight frame [6] can be improved the performance of face verification system using compressive sensing. Future improvements can be done by considering the noise in verifying a face so that the system can be implemented in practice also communication between client and server can be developed further so that it can communicate via the internet.
References
[1] Naseem, I. (2010). Person Identification Using Face and Speech Biometrics (Doctoral dissertation, University of Western Australia).
[2] Uludag, U. (2006). Secure biometric systems (Doctoral dissertation, Michigan State University. Computer Science & Engineering).
[3] Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology, 14(1), 4-20.
[4] Yang, A., Hadinata, A. N., Salim, F., Oey, E., & Hedwig, R. (2015). Face Verification by Using Sparse Representation Algorithm in Compressive Sensing. INTERNETWORKING INDONESIA, 7(1), 25-29.
[5] Abolghasemi, V., Ferdowsi, S., Makiabadi, B., & Sanei, S. (2010, August). On optimization of the measurement matrix for compressive sensing. In Signal Processing Conference, 2010 18th European (pp. 427-431). IEEE.
[6] Xu, J., Pi, Y., & Cao, Z. (2010). Optimized projection matrix for compressive sensing. EURASIP Journal on Advances in Signal Processing, 2010(1), 560349.
[7] Tropp, J. A., & Gilbert, A. C. (2007). Signal recovery from random measurements via orthogonal matching pursuit. IEEE Transactions on information theory, 53(12), 4655-4666.
[8] Chartrand, R., & Yin, W. (2008, March). Iteratively reweighted algorithms for compressive sensing. In Acoustics, speech and signal processing, 2008. ICASSP 2008. IEEE international conference on (pp. 3869-3872). IEEE.
[9] Donoho, D. L. (2006). Compressed sensing. IEEE Transactions on information theory, 52(4), 1289-1306.
[10] Candès, E. J., & Wakin, M. B. (2008). An introduction to compressive sampling. IEEE signal processing magazine, 25(2), 21-30.
[11] Candes, E. J., & Tao, T. (2006). Near-optimal signal recovery from random projections: Universal encoding strategies. IEEE transactions on information theory, 52(12), 5406-5425.