FACIAL IMAGE DE-IDENTIFICATION USING IDENTIY SUBSPACE DECOMPOSITION

Hehua Chi[1,2], Yu Hen Hu[2]

[1]State Key Laboratory of Software Engineering, Computer School, Wuhan University, China
[2]Electrical and Computer Engineering, University of Wisconsin-Madison, USA

ABSTRACT

How to conceal the identity of a human face without covering the facial image? This is the question investigated in this work. Leveraging the high dimensional feature representation of a human face in an Active Appearance Model (AAM), a novel method called the identity subspace decomposition (ISD) method is proposed. Using ISD, the AAM feature space is deposed into an identity sensitive subspace and an identity insensitive subspace. By replacing the feature values in the identity sensitive subspace with the averaged values of k individuals, one may realize a k-anonymity de-identification process on facial images. We developed a heuristic approach to empirically select the AAM features corresponding to the identity sensitive subspace. We showed that after applying k-anonymity de-identification to AAM features in the identity sensitive subspace, the resulting facial images can no longer be distinguished by either human eyes or facial recognition algorithms.

Index Terms—face recognition, identification of persons, data privacy, active appearance model

1. INTRODUCTION

Recent advances in camera technology have made it significantly easier to deal with large amounts of visual data. This enables a wide range of new usage scenarios involving the sharing of images. However, many of these applications, such as the Google Street-view service, are plagued by privacy problems concerning the people visible in the scene [8]. In July 18, 2012, Youtube launched a new tool that allows user to blur all faces in a video before uploading it to the site. The purpose of this new tool is to protect the identity of protestors of Arab Spring movement so that they are less likely to be arrested by oppressive government agencies based on video clips disseminated on Youtube's human right channel Witness [18]. Clearly, the importance of protecting identities of face images has been widely accepted.

However, the majority of privacy protection methods for facial images focus on identifying the faces in an image or video sequence and apply "pixelation" or "blurring" to conceal the identity [10, 14, 18]. However, blurred or pixelated facial images degrade the aesthetic quality, and hence usability of images and videos in many applications. In rare cases, blurred faces may actually attract the attention of perpetrator, defeating the original purpose of hiding the identity of the subject. Furthermore, in studies of human facial expressions, emotions, mental states, it is essential to observe these facial features from subjects while the subject's identity needs to be concealed by privacy laws such as HIPPA (www.hhs.gov/ocr/privacy/), or FERPA (www.ed.gov/policy/gen/guid/fpco/ferpa/index.html).

In this work, we seek to answer the following question: "How to conceal the identity of a human face without covering the facial image?" We note that in movies such as Mission Impossible, there are identity concealing facial masks that allows a human user to impersonate another person without concealing the facial features. However, our goal here is not to impersonate a specific person, but de-identify the facial image of a particular subject. More specifically, we want to convert the facial image of a particular person into that of an "averaged" person, while preserving identity un-related yet useful facial expressions and movements of facial features.

Previously, based on the notion of k-anonymity introduced by Sweeney [2], algorithms, known as k-same [1], k-same-select [5, 7], and k-same-m [6], have been reported. The basic idea is to blend the facial features of k different subjects such that the blended face is equally similar to that of each of the k subjects (probability of correct classification = 1/k). The basic approach of these k-same family algorithms is to average the k most similar faces to yield an averaged face [3, 4, 5, 9]. While from privacy protection point of view, this approach guarantees the desired k-anonymity, the utilities of the resulting averaged faces are often compromised. Specifically, the averaged face may have degraded viewing quality (unlike a human face) or may lose many useful facial signatures (e.g. expressions) due to averaging. For example, the ability to recognize gender or facial expressions from de-identified images may be compromised.

To address these problems exhibiting in the k-same algorithm, in this work, we propose a novel identity subspace decomposition (ISD) method. ISD is based on the hypothesis that the appearance of a human face is influenced by many factors in additional to identity. These factors include, pose, expression, illumination, gender and others

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

[16]. In a multi-dimensional representation of face appearance, it is possible to empirically determine a subspace that is most influenced by the subject identity and the complementary subspace by the remaining factors. By performing averaging on feature values in the identity sensitive subspace for k-anonymity compliance while maintaining original feature values in the identity insensitive subspace will allow one to achieve k-anonymity de-identification while preserving high utility of the resulting averaged faces.

In this work, we employed the active appearance model (AAM) as the appearance-based representation of facial images [12, 13]. For each AAM facial feature, a stepwise linear discriminant analysis (LDA) is applied to estimate its relevance to identity of subjects empirically [16]. Based on the F-statistic of the estimates, the AAM features are ranked from most sensitive to the identity to the least sensitive to the identity in descending order. The anonymity and data utility trade-off analysis then may be performed by deciding how many identity sensitive features need to be selected to achieve the performance objective. We call this identity subspace decomposition (ISD) approach because the subspace spanned by AAM features that are sensitive to identity is explored. For AAM features that are deemed sensitive to identity, we apply the k-anonymity procedure to average the AAM features of at least k subjects. Then combining with remaining AAM features (identity insensitive), a k-anonymized facial image may be obtained. We evaluate the effectiveness of this method using a JAFFE database [15]. Specifically, we have applied a face recognition algorithm to classify the 10 subjects in the database for different ISD partitions. It is shown that as the identity sensitive dimension increases, the receiver operating curve approaches the diagonal line, showing diminishing discriminating power of those k-anonymized facial images. We also presented the facial images at several key stages for visual inspection by the readers. It is clearly that as identity sensitive subspace dimension increases, it becomes more difficult to distinguish different subjects.

The remaining of the paper is organized as follows: the notion of k-anonymity and the face recognition problem are reviewed in Section 2. The ISD procedure is developed in Section 3. Simulations and discussions are reported in Section 4. Finally, the conclusion is in Section 5.

2. BACKGROUNDS

2.1. k-Anonymity

Sweeny introduced the k-anonymity (KA) model for tabulated query output [2]. In particular, it stipulates that a k-anonymity guaranteed query table must contain at least k instances of each value in each of the so called quasi-identifiers. Quasi-identifiers are a subset of attributes that when combined will reveal unique identity of a data record. If a query table meets the k-anonymity condition, then the probability that the identity of an individual may be revealed based on this table is bounded by 1/k. However, some noted attacks do exist that may compromise the k-anonymity. Related works include l-diversity [20] and t-closeness [19].

A database (universe) U is a collection of data records. Each data record is represented by an n-dimensional tuple (vector) with a unique index i and n attributes. Each data record (or equivalently, its index i) is often associated to a particular individual whose privacy is to be protected.
A query is a mapping f of a subset of the database, denoted by D ($D \subset U$), to a real-valued scalar (or vector) $f(D) \in \Re$. If the outcome of a query f yields a specific data record i^*, then the identity of individual i^* will be breached if query f is served by the database to a perpetrator. On the other hand, if one or more queries' outcomes can only reveal the collective identity of k data records with equal probabilities, then the perpetrator will not be able to uniquely determine the identity of a specific individual, and hence the privacy of individuals in the database is protected in the sense that

$$\Pr.\ \{i^* \text{ is identified} | f_1, f_2, \ldots, f_x\} \le 1/k \qquad (1)$$

2.2. Face Recognition

We define face recognition as function Φ: $\mathrm{IR}^m \times (\mathrm{IR}^m)^u \rightarrow \{1, 2, \ldots, u\}$ which given a probe image $p \in \mathcal{P}$ and a gallery $\mathcal{G}$, with $|\mathcal{G}| = u$ outputs the index of the subject most likely to correspond to the subject seen in the probe image: $\Phi(p, \mathcal{G}) = j$, $1 \le j \le u$. By convention we extend Φ to apply as well to a probe set $\mathcal{P}$ with $|\mathcal{P}| = v$ as $\Phi : (\mathrm{IR}^m)^v \times (\mathrm{IR}^m)^u \rightarrow (\{1, 2, \ldots, \mathrm{u}\})^v$ [6, 11].
We evaluate face recognition performance by computing receiver operating characteristic (ROC) curve for a given face recognition function Φ, gallery set $\mathcal{G}$ and probe set $\mathcal{P}$.

In particular, we use LDA (linear discriminant analysis)-based face algorithm to do the evaluation.

3. IDENTITY SUBSPACE DECOMPOSITION (ISD)

3.1. Background: Active Appearance Models

An Active Appearance Model (AAM) consists of a set of statistical features, including object shape and gray level appearance to describe the appearance of an object (say, human face) [12, 13]. It has been found quite useful for many computer vision related tasks. An active appearance model space (A_1, A_2, … , A_i) can be used to illustrate the appearance-based representation of facial images. In our work, our hypothesis is that the appearance of a human face is influenced by many factors in additional to identity. These factors include, pose, expression, illumination, gender and others [16]. In a multi-dimensional representation of face appearance, it is possible to empirically determine a subspace that is most influenced by the subject identity and the complementary subspace by the remaining factors.

3.2. Stepwise linear discriminant analysis

The model feature vectors $\mathcal{V} = \{F_1, F_2, \ldots, F_n\}$ of *n* facial images are submitted to the stepwise linear discriminate analyses (LDA) which examines the facial factors (identity, pose, expression, illumination, gender and others)[16], In our work, we only consider the identity factor. The analysis operates by evaluating the active appearance model space (A_1, A_2, … , A_i) that shows small within-category variability relative to total variability. For all of the discriminant analyses reported, the criteria are based on an F-statistic.

This is a measure of change in Wilk's lamba for each AAM facial feature A_i to evaluate the sensitivity to the identity. This stepwise LDA operates by ranking the AAM facial features from the most to the least sensitive to the identity in descending order. Therefore, it is possible to empirically determine a subspace that is most influenced by the subject identity and the complementary subspace by the remaining factors.

Therefore, we can get the sensitivity ranking $\mathcal{R}(A_i)$ of the AAM facial features $\{A_i\}$ related to the categorization of the identity. This ranking allows one to select an integer *m* such that subspace spanned by the AAM space (A_1, A_2, … , A_i) with ranks higher than *m* are deemed sensitive to identity, and the subspace spanned by remaining AAM facial features will be deemed identity insensitive. The anonymity and data utility trade-off analysis then may be performed by deciding how many identity sensitive features need to be selected to achieve the performance objective.

The identity subspace decomposition (ISD) algorithm can be summarized below:

Input: Gallery $\mathcal{G} = \{G_1, G_2, \ldots, G_y\}$, An integer *m*, Active Appearance Model $\mathcal{M}$, Stepwise LDA
Output: the identity sensitive subspace $\{A_m\}$
1 Generate AAM facial features (A_1, A_2, … , A_i) with respect to $\mathcal{M}$ and $\mathcal{G}$
2 Label each image G_y as one of identity variables $\mathcal{L}_z$
3 Compute the F-statistic value of each A_i with respect to stepwise LDA, $\mathcal{G}$ and $\mathcal{L}_z$
4 Compare and rank the F-statistic values $\mathcal{R}(F)$
5 Obtain the sensitivity ranking $\mathcal{R}(A_i)$ of $\{A_i\}$ with respect to $\mathcal{R}(F)$
6 Given *m*, divide $\{A_i\}$ into identity sensitive subspace $\{A_m\}$ (with ranks higher than *m*) and an identity insensitive subspace (the remaining $\{A_{i-m}\}$) with respect to $\mathcal{R}(A_i)$

3.3. k-anonymity Compliance

Once the ISD is performed, and the identity sensitive dimension $\{A_m\}$ is determined, one may proceed to impose k-anonymity by averaging the AAM model features within the identity sensitive subspace while maintaining original feature values in the identity insensitive subspace.

The k-anonymity compliance algorithm can be summarized below:

Input: Probe set $\mathcal{P}$, Identity sensitive subspace $\{A_m\}$, An integer *k*.
Output: k-anonymized facial image set $\mathcal{D}_{A_m}^k(\mathcal{P})$
1 Compute AAM feature vectors $\mathcal{V} = \{F_1, F_2, \ldots, F_n\}$ of *n* facial images in $\mathcal{P}$
2 Apply the k-anonymity procedure to average the subspace $\{A_m\}$ of at least *k* subjects while maintaining original feature values in the identity insensitive subspace $\{A_{i-m}\}$
3 Reconstruct the k-anonymized facial images
3 Obtain the k-anonymized set $\mathcal{D}_{A_m}^k(\mathcal{P})$

4. SIMULATIONS AND DISCUSSIONS

4.1. Simulations

To verify the effectiveness of our proposed identity subspace decomposition (ISD) method, we conducted the simulations below.

4.1.1. Data source

Experiments were run on facial images from the JAFFE Face Database which contains 213 images of 7 facial expressions posed by 10 Japanese female models [15]. For the purposes of this paper, it is necessary to manually annotate the facial structure for each facial image using 68 landmarks: eyebrows, eyes, nose, mouth and contour.

4.1.2. Data Analysis

We follow the identity subspace decomposition (ISD) algorithm introduced in Sections 3.1 and 3.2 to analyze the facial images. The active appearance model (AAM) was used to represent the face appearance and the active appearance model space (A_1, A_2, … , A_{30}) was generated. Thus, for facial images, 30-dimension model feature vectors $F_n = (a_1^n, a_2^n, \ldots, a_{30}^n)$ were extracted.

The model vectors $\mathcal{V} = \{F_1, F_2, \ldots, F_{213}\}$ of all the facial images from JAFFE database were submitted to the stepwise linear discriminate analyses (LDA) which examines the facial identity factor. Prior to the analysis, the 'signature' of each facial image is identified as one of the 10 categories of human faces. After the data analysis, the sensitivity ranking $\mathcal{R}(A_i)$ of AAM facial features $\{A_i\}$ related to identity is generated as follows: A_7, A_6, A_2, A_{11}, A_4, A_3, A_8, A_9, A_{10}, A_{13}, A_1, A_5, A_{12}, A_{16}, A_{14}, A_{18}, A_{19}, A_{20}, A_{15}, A_{27}, A_{17}, A_{25}, A_{26}, A_{30}, A_{21}, A_{24}, A_{22}, A_{28}, A_{23}, A_{29}. In our work, we use the Statistical Product and Service Solutions (SPSS) software to run the stepwise LDA [17].

The sensitivity ranking $\mathcal{R}(A_i)$ was used for the facial image de-identification experiments below.

4.1.3. Evaluation of Privacy Protection

We follow the k-anonymity compliance algorithm in Section 3.3 to run the face de-identification experiments. 7 different values {0, 5, 10, 15, 20, 25, 30} were respectively assigned for the integer *m* and let *k* of the k-anonymity be 10. Therefore, we can obtain 7 k-anonymized facial image sets: $\mathcal{D}_{A_0}^{10}$ $\mathcal{D}_{A_5}^{10}$, $\mathcal{D}_{A_{10}}^{10}$, $\mathcal{D}_{A_{15}}^{10}$, $\mathcal{D}_{A_{20}}^{10}$, $\mathcal{D}_{A_{25}}^{10}$, $\mathcal{D}_{A_{30}}^{10}$. Then the ROC curves were computed and compared over all seven configurations which were named as ROC1, ROC2, ROC3, ROC4, ROC5, ROC6 and ROC7 shown in Fig. 1. In particular, we used LDA (linear discriminant analysis)-based face algorithm to do this evaluation.

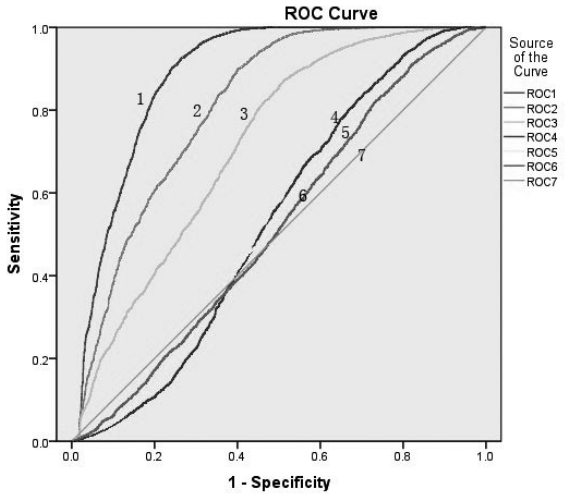

Fig. 1. The results of this experiment: ROC curves for 7 k-anonymized facial image sets based on linear discriminate analysis (LDA)

4.1.4. Evaluation of Data Utility

Table 1. The reconstructed de-identification facial images based on seven different facial expressions of the person KL at seven key stages

	KL.AN	KL.DI	KL.FE	KL.HA	KL.NE	KL.SA	KL.SU
$\mathcal{D}_{A_0}^{10}$							
$\mathcal{D}_{A_5}^{10}$							
$\mathcal{D}_{A_{10}}^{10}$							
$\mathcal{D}_{A_{15}}^{10}$							
$\mathcal{D}_{A_{20}}^{10}$							
$\mathcal{D}_{A_{25}}^{10}$							
$\mathcal{D}_{A_{30}}^{10}$							

In order to visualize and evaluate the data utility after face de-identification experiments in section 4.1.3, we presented the facial images at seven key stages for visual inspection by the readers.

In the Table 1, the seven different expressions of the person KL were selected from the JAFFE database to present the reconstructed de-identification facial images at seven key stages.

4.2. Discussions

We evaluate the effectiveness of the identity subspace decomposition (ISD) method using JAFFE database [15]. Specifically, we have applied the linear discriminate analysis (LDA) face recognition algorithm to classify the 10 subjects in the database for different ISD partitions: $\mathcal{D}_{A_0}^{10}$ $\mathcal{D}_{A_5}^{10}$, $\mathcal{D}_{A_{10}}^{10}$, $\mathcal{D}_{A_{15}}^{10}$, $\mathcal{D}_{A_{20}}^{10}$, $\mathcal{D}_{A_{25}}^{10}$, $\mathcal{D}_{A_{30}}^{10}$. It is shown that as the identity sensitive dimension increases, the receiver operating curve approaches the diagonal line, showing diminishing discriminating power of those k-anonymized facial images. Therefore, we can decide how many identity sensitive features need to be selected to achieve the performance objective. We also presented the facial images at several key stages for visual inspection by the readers. It is clearly that as identity sensitive subspace dimension increases, it becomes more difficult to distinguish different subjects.

5. CONCLUSION

To address these problems exhibiting in the k-same family algorithm, in this work, we proposed a novel identity subspace decomposition (ISD) method. Using ISD, the AAM feature space was deposed into an identity-sensitive subspace and an identity insensitive subspace. By replacing the feature values in the identity sensitive subspace with the averaged values of k individuals, one may realize a k-anonymity de-identification process on facial images. We applied this approach to the JAFFE database. We showed that by performing averaging on feature values in the identity sensitive subspace for k-anonymity compliance while maintaining original feature values in the identity insensitive subspace will allow one to achieve k-anonymity de-identification while preserving high utility of the resulting averaged faces.

6. ACKNOWLEDGMENT

This paper is supported by the Fundamental Research Funds for the Central Universities (No.2012211020202), National Natural Science Fund of China (No. 61173061) and financial support from the program of China Scholarships Council (No.201206270037)

7. REFERENCES

[1] E.Newton, L.Sweeney, and B.Malin, "Preserving Privacy by De-identifying Facial Images," IEEE Trans. Knowledge and Data Engineering, vol.19, no.2, pp. 232-243, Feb.2005.

[2] L.Sweeney, “k-anonymity: a model for protecting privacy,” International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, vol.10, no.5,pp.557-570, 2002.
[3] R.Gross, L.Sweeney, J.Cohn, F.de la Torre, and S.Baker, “Face De-identification. Protecting Privacy in Video Surveillance, ” A. Senior, editor. Springer, 2009.
[4] R.Gross, and L.Sweeney, F.Cohn, F.de la Torre, and S.Baker, “Multi-Factor De-Identification of Facial Images, ” Proceedings of the 2008 American Medical Informatics Association Annual Symposium, 2008.
[5] R.Gross, and L.Sweeney, “Towards Real-World Face De-Identification, ” Proc. IEEE Conference on Biometrics, Washington, DC, Sep. 2007.
[6] R. Gross, L.Sweeney, F.de la Torre, and S.Baker, “Model-based face de-identification, ” Proc. IEEE Workshop on Privacy Research in Vision, 2006.
[7] R.Gross, E.Airoldi, B.Malin, and L.Sweeney, “Integrating Utility into Face De-Identification, ” Proc. Workshop on Privacy-Enhanced Technologies, 2005.
[8] Gross, R., Sweeney, L., Cohn, J., de la Torre, F. and Baker, S. “Face De-identification. Protecting Privacy in Video Surveillance, ” A. Senior, editor. Springer, 2009.
[9] R.Gross, and L.Sweeney, “Semi-Supervised Learning of Multi-Factor Models for Face De-Identification, ” Proc. Computer Vision and Pattern Recognition, Anchorage, AK, Jun, 2008.
[10] R. Agrawal and R. Srikant, “Privacy-preserving data mining,” Proc. SIGMOD, 2000
[11] D. Blackburn, M. Bone, and P. J. Phillips, “Facial recognition, ” 2000.
[12] T. Cootes, G. Edwards, and C. Taylor, “Active appearance models,” IEEE PAMI, 23(6):681–685, 2001.
[13] I. Matthews and S. Baker, “Active appearance models revisited,” IJCV, vol.60, no.2, pp:135–164, 2005.
[14] C. Neustaedter, S. Greenberg, and M. Boyle, “Blur filtration fails to preserve privacy for home-based video conferencing,” ACM TOCHI, 2005.
[15] Michael J. Lyons, Miyuki Kamachi, Jiro Gyoba. Japanese Female Facial Expressions (JAFFE), Database of digital images, 1997.
[16] A.J. Calder, A.M. Burton, P. Miller, et al. “A principal component analysis of facial expressions,” Vision research, vol (41), no.9, pp: 1179-1208, 2001.
[17] Nie N H, Bent D H, Hull C H. SPSS: Statistical package for the social sciences [M]. New York: McGraw-Hill, 1975.
[18] D. Netburn, “Youtube’s new face-blurring tool designed to protect activists,” Activists, Los Angeles Times, LA, USA, July, 2012.
[19] Li N, Li T, Venkatasubramanian S, “t-closeness: Privacy beyond k-anonymity and l-diversity, ” IEEE 23rd International Conference on Data Engineering, pp:106-115, 2007.
[20] Machanavajjhala A, Kifer D, Gehrke J, et al. “l-diversity: Privacy beyond k-anonymity, ” ACM Transactions on Knowledge Discovery from Data (TKDD), 2007.