Retaining Expressions on De-identified Faces

Lily Meng[*], Zongji Sun, Aladdin Ariyaeeinia
School of Engineering and Technology
University of Hertfordshire
Hatfield AL10 9AB, UK
E-mail:1.1.meng@herts.ac.uk

Ken L. Bennett
Satellite Simulation and Functional Verification
Infrastructure
Astrium Ltd., Stevenage, UK
E-mail: KenBennett@theIET.org

Abstract—**The extensive use of video surveillance along with advances in face recognition has ignited concerns about the privacy of the people identifiable in recorded documents. Prior research into face de-identification algorithms has successfully proposed *k*-anonymity methods that guarantee to thwart face recognition software. However, there has been little investigation into the preservation of the data utility such as gender and expression in the original images. To address this challenge, a new algorithm based on the Active Appearance Model is proposed here. The main attraction of the approach is that of the preservation of the data utility in terms of facial expression, whilst maintaining privacy protection. The former includes not only the preservation of the expression category (e.g. happy or sad), but also the details of the original expression (e.g. the intensity of a smile and movements of the lips). This is considered to be of significant value in real applications of face de-identification, where the given video contains facial images of the same expression with various degrees of intensity.**

Keywords—privacy protection, face de-identification, facial expression transfer, active appearance model, linear discriminant analysis, k-anonymity.

I. INTRODUCTION

Due to the extensive applications of surveillance systems and consequently the rising concern about the privacy of the individuals identifiable in the captured videos, various methods have been proposed for the de-identification of faces in still and moving images. These methods can be divided into two categories: the ad hoc methods (such as masking, pixelation and blurring [1-3] and the *k*-anonymity based methods (such as *k*-Same [4]). The ad hoc methods are usually simple to implement. However, they fail to serve their purpose as they are unable to thwart the existing face recognition software [4, 5]. To achieve privacy protection, the concept of *k*-anonymity was introduced by Sweeney in 2002 [6]. All *k*-anonymity based methods de-identify a face image by replacing it with the average of *k* face images from a given gallery and hence achieve privacy protection by guaranteeing a recognition rate lower than 1/*k*. Within the *k*-anonymity family the most widely used algorithms for face de-identification is the *k*-Same family. However, the original *k*-Same solutions (i.e. *k*-Same-Pixel and *k*-Same-Eigen [4]) were not designed for preserving data utility. As a result, the de-identified version of a male face might look feminine and a happy face may lose its smile. In addition, *k*-Same is appearance based, operating entirely in the image space. As a result, 'ghosting' artefacts tend to appear due to the misalignments of the *k* images involved, even when images are aligned based on a small number of facial

landmarks (e.g. the corners of the eyes and the tip of the nose). Finally, *k*-Same-Pixel/-Eigen and all their extensions achieve decent privacy protection at the cost of large *k*'s, which in turn lead to the demand for a large image gallery or otherwise lack of discrimination among the de-identified faces (number of distinctive faces in the de-identified face set is equal to or less than size of the gallery divided by *k*).

To address the limitations of *k*-Same-Pixel/-Eigen, various approaches have been proposed. To remove ghosting artefacts, the *k*-Same-M algorithm [7] models each of the *k* closest images as an Active Appearance Model (AAM) [8] and then performs averaging on the AAM model parameter vectors. To truly protect privacy, a new member named *k*-Same-furthest has been introduced to the *k*-Same family [9]. Like *k*-Same-M, *k*-Same-furthest is model based and represents face images as AAMs. In contrast to the *k* closest images used in the other *k*-Same algorithms, it averages the *k* most different images to the original and hence the name *k*-Same-furthest.

The most cited attempt on integrating utility preservation into face de-identification was *k*-Same-Select [10]. It partitions its image gallery into mutually exclusive subsets using data utility functions such as an expression or a gender classifier and then applies the standard *k*-Same algorithm merely to the subset which the probe image belongs to. Obviously, *k*-Same-Select requires a large image gallery as there must be an image for each individual in every utility subset. Furthermore, the algorithm is extremely inflexible. Each time a new subset (e.g., another expression) is added to the application, the utility classifier would have to be re-trained to identify the new class. Nearly all cultures recognise facial expressions of six basic emotional categories, namely: joy, sadness, anger, disgust, fear and surprise [11]. To accommodate all six categories of facial expression, *k*-Same-Select would require a large image gallery and a rather complicated expression classifier. Even worse, *k*-Same-Select demands exclusivity among its utility subsets while categories of facial expression are not mutually exclusive. This means that *k*-Same-Select would fail its purpose of preserving data utility when facing, for example, a happily surprised face. Finally, *k*-Same-Select preserves only one aspect of each data utility (i.e. the category) and relies on accurate classification of the probe image. The level of data utility preservation achieved by *k*-Same-Select is equal to the accuracy of its utility classifier.

The algorithm proposed here takes a completely different approach to preserving data utility. It recovers the data utility by transferring/cloning the facial expression from the original

to the de-identified face. The proposed algorithm does not require complicated classifiers or high-level semantic information describing facial expressions and therefore avoids the above-mentioned limitations of *k*-Same-Select.

The remainder of the paper is structured as follows. Section 2 describes the two key techniques of AAM and *k*-Same-furthest, which are adopted in the proposed algorithm. Section 3 reviews facial expression transfer using AAMs and Section 4 presents the proposed algorithm together with example results. Section 5 evaluates the proposed algorithm's ability to protect privacy as well as preserve data utility. Finally, the findings and overall conclusions of this work are given in Section 6.

II. SUPPORTING METHODS

A. Active Appearance Models

In an AAM, the shape and texture data of an object are separated and can be manipulated independently. The shape of an object is defined by the coordinates of a pre-determined set of landmarks. The landmarks of an object are allocated using an iterative fitting process, where the texture of the object is warped to the current fitted shape and the result appearance is compared with the original appearance of the object. If the difference is acceptable, the current fitted shape returns the locations of the landmarks. Otherwise, the fitted shape is updated through changing its Principal Component Analysis (PCA [12]) parameters and the process iterates. To isolate texture of an object from its shape, the face appearance in the original image (bound by the object's shape) is reshaped into the AAM mean shape using piece-wise affine warping. Then the shape and the texture data is projected independently to their own PCA feature spaces for a more efficient representation. The shape and texture PCA parameters can be kept separate as in independent AAMs. Alternatively, they can be concatenated to produce a combined set of object parameters, with the addition of shape weights to allow the concatenated parameters to have a common set of units. PCA can then be applied to produce a new set of appearance parameters that are a combined representative model of both the shape and the texture. This is the case of combined AAMs.

The shape of a 2D AAM is defined by a triangulated mesh and in particular the vertex locations of the mesh, i.e. the positions of the landmarks. The proposed face de-identification algorithm uses 64 landmarks to define the shape of a human face (Fig. 1). The piece-wise affine warping achieves its goal by transferring image pixels from the mean-shape based texture image to the corresponding mesh triangles in the target shape mesh. The ghosting artefacts in *k*-Same are due to the misalignments between the *k* closest face images; while the shapes and the texture images represented in the AAM PCA space are all aligned to the mean shape and therefore with each other. In AAM, alignment between two shapes is achieved through Procrustes analysis, where the shape landmarks are scaled, rotated and translated, to minimise the total sum of distances between landmark pairings of the two shapes.

A more recent reassessment of the performance of standard AAMs by Mathews *et al*. [13] resulted in the development of the more efficient Inverse Compositional AAM (ICAAM). The ICAAM approach has become popular in the area of assessing
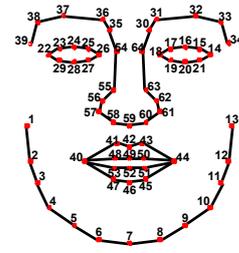
Fig. 1. Facial landmarks used in this work.

the performance of AAM on fitting and analyzing faces [14, 15]. In addition, an ICAAM is an independent AAM that models shape and texture separately. The separate and hence simpler handling of shape and texture was ideal for the expression transfer presented in this paper, allowing different strategies for the shape and the texture transfer. Furthermore, combined AAM does no longer assume the orthogonality between the shape and texture parameters and therefore restricts the fitting algorithm.

B. The k-Same-furthest Algorithm

The newly introduced member of the *k*-Same family, *k*-Same-furthest [9], is a model-based approach, representing faces as AAM features to avoid ghosting artefacts. It identifies the *k* faces that are furthest away from a given probe face image, calculates the average of these faces, and returns this average as the de-identified face of the probe. The *k*-Same-furthest algorithm guarantees that face de-identified by it can never be recognised as the original face as long as the identity distance measure used to recognise the de-identified faces is the same as that used by *k*-Same-furthest [9]. If the de-identified faces are recognised by Eigenface technique [16] which represents faces in their PCA space rather than AAM, *k*-Same-furthest offers a recognition rate very close to zero [9].

III. REVIEW OF EXPRESSION TRANSFER USING AAMS

There have been various attempts to use AAM for Facial Expression Transfer (FET). This can be straightforward when source and destination face shapes and appearances are similar, such as in the case of the animation of human faces [17]. However, when geometric dissimilarity exists between the source and the destination, a more complicated process is needed to ensure the transferred expression merges naturally with the destination face.

The study in [18] presents an effective AAM-based approach to FET between two known subjects, where one AAM is trained for each subject. Each source AAM base vector is expressed as a weighted sum of all the base vectors of the destination AAM space. The mapping between a source vector and each of the destination vectors is computed by their inner products. This mapping is then used to transfer appearance of the source subject to the destination subject via the AAM feature spaces. The mapping proposed in [18] is simple, allowing FET in real-time. The relevant work of Hunty et al. [19] built on other research into FET using AAMs. The proposal assessed how a fitted expression from one AAM could be used to synthesize the same expression realistically onto another person or animated character in a separate AAM. The 'combined' shape and appearance parameters were

transferred by linearly representing the expressions as the combination of the mean appearance and associated expression deviations. These expression variations were transferred from one AAM to the other, either directly or by converting those variations to the possible variations in the destination AAM. Although the work presented here has some similarities with the studies in [18] and [19], it involves a fundamentally different approach. This is due to the fact that the proposed FET process must focus on using a single generic AAM to transfer expressions between subjects, rather than using separate person-specific AAMs as in [18] and [19]. Person specific AAMs would provide more accurate AAM fitting, but this requires new AAMs to be trained for each new subject. In k-Same framework, the de-identified face image is obtained by averaging k images. Expression can be retained on the average face image by first transferring the expression from the source to each of the k subjects in concern and then taking the average. This means, to support k-same de-identification per subject, the processes in [18, 19] would require the training of k additional AAMs and perform FET between the source AAM and each of the k additional AAMs. Furthermore, preparing training images for AAM is usually a lengthy manual process, as they must be annotated as accurately and consistently as possible to prevent shape and appearance noise between the facial training images being included in the model. In [18] it is stated that the subject-specific AAMs require 15-20 training images per subject and images across subjects must display similar expressions; while the training of a generic AAM accepts training images from various subjects displaying unlimited expressions. For the above reasons, the FET process proposed here makes use of a single generic AAM across subjects and can be applied to any source subject without the need of additional AAM training.

IV. The Proposed Algorithm

A. Overall Process Flow

As the proposed algorithm employs k-Same-furthest for privacy protection and model-based FET for data utility preservation, it is named k-Same-furthest-FET. Fig. 2 summarizes its process flow. The input to the algorithm is an expression face image to be de-identified. The k-Same-furthest process generates a new identity for the input face, which is also the destination subject of the FET process. Since no subject-specific AAM is employed, the proposed FET process requires a common initial expression between the original and the new destination identities. Neutral expressions have been chosen to serve this purpose. As shown in Fig. 2, the proposed process achieves expression preservation by transferring the differences between the original identity's neutral and expression faces to the neutral face of the new identity. The adjustments before and after the tranfer of expression difference are explained in the next subsection. The FET process begins with fitting three face images by the trained generic AAM. The same AAM is shared by the k-Same-furthest and the FET processes. To facilitate independent and hence possible parallel processing, shape and texture data of each face are kept separated after the AAM fitting. When the expression transfer is completed with the shape and the texture data independantly, the output expression appearance of the new identity is generated by warping the transferred texture to the transferred shape.

In practice, the neutral face of the original identity may or may not be available explicitly. If not, it can be retrieved from a gallery using standard Eigenface technique [16] or model-based Eigenface technique. This work assumes the neutral face of the original identity is not available explicitly. To avoid additional training, the proposed k-Same-furthest-FET uses AAM-based Eigenface to identify the neutral face of the original identity. When the Eigenface technique returns a face of a wrong identity, this identity difference between the expression and the neutral faces will be transferred from the original subject to the generated faces of the new subject. However, the visual quality of the generated faces of the new subject do not degrade (see Fig. 3 for an example). The study in [20] reports a rank-1 recognition rate of 73% by basic Eigenface technique for 2D expression images and an expression-invariant recognition technique is able to produce a recognition rate higher than 90% (e.g. [21]), meaning faces of the same subject will display an identical identity in most cases.
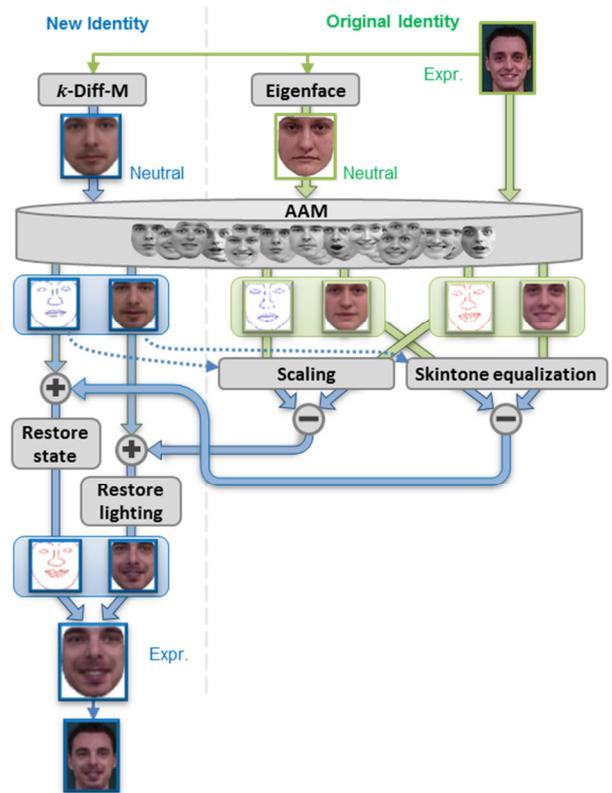
Fig. 2. Process flow of the proposed algorithm, k-Same-furthest-FET.

Original identity			New identity		
Input expression face	Neutral face		Neutral by k-Same-furthest	Output expression with	
	Wrong by Eigenface	Correct		wrong neutral	correct neutral

Fig. 3. Impact of a wrong neutral face of the original identity on the output expression face of the new identity. In this and following figures, face have been cropped for display purpose.

B. AAM Training

The proposed FET demands no control over the identity or the expression in the source image and therefore requires its AAM to be capable of fitting and synthesizing any identity for any facial movements that the FET process would be expected to handle. To meet this requirement, the AAM has to be trained with enough examples of each type of expected facial expression and from a variety of unrelated, but representative set of subjects. The proposed independent manipulation of the shape and texture data means the AAM implementation has to be an 'independent' AAM, such as an ICAAM. The PCA thresholds must be set to a level so that any shape or appearance inaccuracies in the training set do not cause the AAM fitting process to become unstable. An unstable fitting process leads to poor synthesized representations of the FET faces, resulting in poor generated expressions. The PCA thresholds of 75% for shape and 90% for appearance were found to provide stable ICAAM fitting and face synthesis for the dataset used in this work, even though the synthesized faces contain smoother facial shapes and skin textures than the original images.

C. Expression Transfer between Face Shapes

1) Notation

For a convenient notation distinguishing different faces and persons, the shape for a face expression f of identity p is denoted as $\mathbf{s}_p^f$. Therefore, the goal of expression transfer across face shapes can be expressed as that for a given shape $\mathbf{s}_{org}^{expr}$ of an orignal identity org generate a plausible $\mathbf{s}_{new}^{expr}$ of a new identity new such that it presents the same expression features as $\mathbf{s}_{org}^{expr}$. This paper uses the same notation for face textures $\mathbf{t}_p^f$.

2) Scaling of input face shapes

Considering the geometric differences of face shapes among individuals, the face shapes of the original identity are normalized to the shape of the new identity. The normalization is applied independently in the horizontal and vertical directions. The horizontal scaling factor is based on the widest facial landmarks (l_1 to l_2 and l_{12} to l_{13}) and determined by aligning $\mathbf{s}_{org}^{neutral}$ to $\mathbf{s}_{new}^{neutral}$ using Procrustes analysis; while the vertical scaling factor is based on the highest and lowest landmarks (l_{31} to l_{33}, l_{36} to l_{38} and l_6 to l_8).

3) Transfer of expression in the PCA space

Intuitively, the output $\mathbf{s}_{new}^{expr}$ can be calculated using (1).

$$\mathbf{s}_{new}^{expr} = \mathbf{s}_{new}^{neutral} + \mathbf{s}_{org}^{expr} - \mathbf{s}_{org}^{neutral} \tag{1}$$
$$= \sum_{i=1}^{l} \left(p_{i_{new}}^{neutral} + p_{i_{org}}^{expr} - p_{i_{org}}^{neutral} \right) \mathbf{b}_i ,$$

where $\Delta \mathbf{s}_{org} = \mathbf{s}_{org}^{expr} - \mathbf{s}_{org}^{neutral}$. As expressed in (1), each face shape can be represented in terms of its Cartesian coordinates or its projection in the AAM feature space. Calculation of (1) in the two spaces would derive identical result. This is due to the fact that (1) only involves an addition and a subtraction while PCAs are linear combinations of the original Cartesian coordinates. Despite the same results, the PCA space is much more efficient than the Cartesian space since firstly, it deals with much fewer data (e.g. 6 vs.64x2

shape features and 35 vs. 197x178x3 texture features for the dataset and AAM PCA thresholds used in this work); and secondly, no shape alignment is necessary. For these reasons, the proposed algorithm calculates (1) in the AAM PCA space for both shape and texture data. Fig. 4 shows the identical results of $\mathbf{s}_{new}^{expr}$ calculated in the Cartesian and the AAM feature spaces for an example set of input shapes.

An important consideration when calculating (1) is that output $\mathbf{s}_{new}^{expr}$ might go too far from the mean shape and cause implausible appearances (e.g. the upper and lower lip boundaries intersect). To avoid this, a constraint is imposed on the PCA parameters $\{p_{i_{new}}^{expr}\}$ of $\mathbf{s}_{new}^{expr}$ to ensure each $p_{i_{new}}^{expr} \leq \pm 3\sqrt{\lambda_i}$ where λ_i is the ith PCA eigenvalues and measures the variance of parameter p_i across the AAM training images.

4) Extra expression preservation of the eyes and the mouth

The degrees of mouth and eye opening and the overall shape of the mouth plays an important role in the communication of emotion [22]. However, a direct transfer of $\Delta \mathbf{s}_{org}$ using (1) may cause $\mathbf{s}_{new}^{expr}$ to display a different state than $\mathbf{s}_{org}^{expr}$. Fig. 5 demonstrates an example, where $\mathbf{s}_{new}^{expr}$ calculated using (1) (Fig. 5 (d)) has a small mouth opening and different mouth shape than that of $\mathbf{s}_{org}^{expr}$. This is simply because $\mathbf{s}_{org}^{neutral}$ has a slightly opened mouth while $\mathbf{s}_{new}^{neutral}$ is tightly closed. To address this issue, this work prioritizes the preservation of the mouth and eye opening and the overall mouth shape.

For a given facial component (the mouth or an eye), first translate the component to have the midpoint of its corners over the origin of the Cartesian reference frame and then rotate the translated facial component with respect to the origin to have its corners in horizontal. This step will enable a pose-and-position invariant expression transfer. Second, determine the x coordinate of each landmark on the component in $\mathbf{s}_{new}^{expr}$ using

$$x_{new\ i}^{expr} = x_{org\ i}^{expr} \cdot w_{new}^{neutral} / w_{org}^{neutral} \tag{2}$$

where $x_{p\ i}^f$ denotes the x coordinate of the ith landmark l_i of shape $\mathbf{s}_p^f$ and w_p^f the width of the component. Width of a component is defined as the Euclidean distance between its corners. Third, determine the y coordinates of the landmarks that define the opening of the component (i.e. landmarks on the eye lids: 14 to 21 and 22 to 29 in Fig. 1, or landmarks on the inner lip boundaries: 48 to 53 in Fig. 1) using (2) but with ratio between the component height instead of width. Forth, if the component is the mouth use (2) to calculate the desired lip thickness of $\mathbf{s}_{new}^{expr}$ at each landmark on the outer lip boundaries (landmarks 40 to 47 in Fig. 1) and determine the y coordinates

(a) $\mathbf{s}_{org}^{neutral}$ by Eigenface	(b) Input $\mathbf{s}_{org}^{expr}$	(c) output $\mathbf{s}_{new}^{expr}$ based on Cartesian coordinates and AAM parameters	(d) $\mathbf{s}_{new}^{neutral}$ by k-Same-furthest
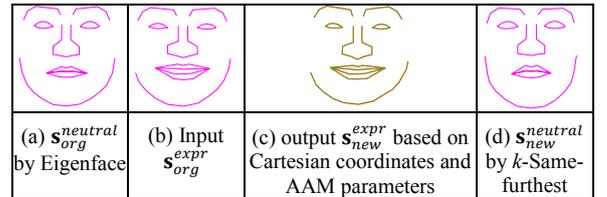
Fig. 4. Results of expression transfer derived in the Cartesian and the AAM spaces, for given inputs (a), (b) and (d).

of these landmarks to achieve the desired lip thickness values. Here, lip thickness at a landmark is the distance between itself and its corresponding landmark on the inner lip. For example, lip thickness at landmark 41 in Fig. 1 is the distance between itself and landmark 48. Finally, restore the position and rotation of the component as those generated using (1). The third step ensures that the mouth and both eyes of $\mathbf{s}_{new}^{expr}$ present the same degrees of opening as those of $\mathbf{s}_{org}^{expr}$. The forth step with the landmarks on the outer lip boundaries together with the third step ensures that $\mathbf{s}_{new}^{expr}$ has identical mouth shape as $\mathbf{s}_{org}^{expr}$. Fig. 5 (c) shows the result of these steps, where the degrees of mouth and eye opening and the overall mouth shape of $\mathbf{s}_{org}^{expr}$ is preserved in $\mathbf{s}_{new}^{expr}$.

D. Expression Transfer between Face Textures

Like shape, $\mathbf{t}_{new}^{expr}$ is calculated using (1) in the PCA space with the limit of $\pm 3\sqrt{\lambda_i}$ applied to each PCA parameter. However, $\mathbf{t}_{new}^*$ often has a different skin tone than $\mathbf{t}_{org}^*$, making the skin tone of $\mathbf{t}_{org}^*$ to be transferred to output $\mathbf{t}_{new}^{expr}$ through calculation of (1). As a result, the output $\mathbf{t}_{new}^{expr}$ appears to have a different skin tone than $\mathbf{t}_{new}^{neutral}$ even when they have identical illumination. Here $\mathbf{t}_p^*$ denotes any texture of person p. To avoid this, the skin tone of $\mathbf{t}_{org}^*$ are align to that of $\mathbf{t}_{new}^{neutral}$ through histogram equalization. To allow $\mathbf{t}_{new}^{expr}$ to blend in naturally with the original background of $\mathbf{t}_{org}^{expr}$, illumination of $\mathbf{t}_{org}^{expr}$ is cloned to the generated $\mathbf{t}_{new}^{expr}$. In this work, all neutral face images display a similar evenly lit face area, allowing the following steps to produce a good cloning result of the illumination. First, align $\mathbf{t}_{org}^{expr}$ to $\mathbf{t}_{org}^{neutral}$ using histgram equalization. Then calculate the difference between the equalized and the original $\mathbf{t}_{org}^{expr}$. Finally, add this difference to $\mathbf{t}_{new}^{expr}$. Fig. 6 displays visual results of these steps.
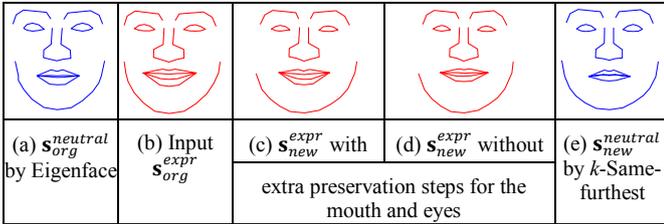
Fig. 5. Results of expression transfer between shapes with and without extra preservation steps for the mouth and eyes.

(a) $\mathbf{s}_{org}^{neutral}$ by Eigenface	(b) Input $\mathbf{s}_{org}^{expr}$	(c) $\mathbf{s}_{new}^{expr}$ with	(d) $\mathbf{s}_{new}^{expr}$ without	(e) $\mathbf{s}_{new}^{neutral}$ by k-Same-furthest
		extra preservation steps for the mouth and eyes		

$\mathbf{s}_{org}^{neutral}$	$\mathbf{s}_{org}^{expr}$	$\mathbf{s}_{new}^{expr}$ (skintone equalized)	$\mathbf{s}_{new}^{expr}$ after restore of illumination	$\mathbf{s}_{new}^{neutral}$
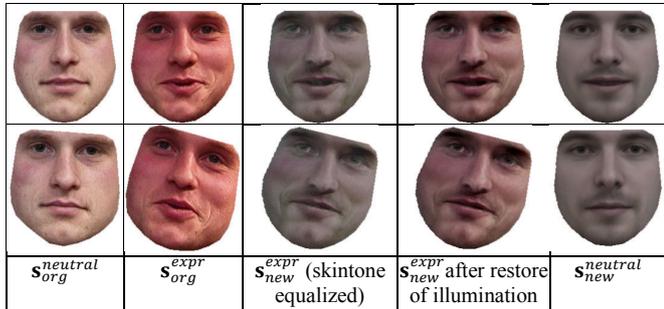
Fig. 6. Results of expression transfer before and after various adjustmets to original textures. First row: the mean-shape face textures. Second row: the face appearances with textures being warped to their corresponding shapes.

V. EXPERIMENTS

A. Dataset

Experiments in this work were conducted with the IMM dataset [23], which contains images of 40 subjects. Only images with a near-frontal pose were used. These include a neutral, a happy and an arbitrary expression face images per subject. There is variation in head pose among the neutral as well as the happy faces. There is variation in both pose and lighting among the arbitrary expression faces.

B. Evaluation of Privacy Protection

Privacy protection ability of the proposed algorithm is measured through recognition experiments using Eigenface technique in the AAM space [16]. Cropped face images showing only the region inside the outline of the AAM-fitted shape are used in the experiments. 70% of the subjects were randomly selected and used for training the Eigenface space with cropped original images. In testing, all cropped original images with various expressions were used as the gallery and the de-identified expression images as the probes. All results reported are based on randomly selecting ten different training and gallery/probe sets and computing the average recognition rate over all configurations.

Fig. 7 shows the rank-1 recognition rates for the cropped original faces and the faces de-identified using AAM-based k-Same-Select and k-Same-furthest with and without FET. The recognition rates of k-Same-M-Select faces stay slightly below the theoretical maximum of $1/k$. As proved and tested in [9], k-Same-furthest without the FET process produces a recognition rate of zero regardless of the value of k. The recognition rates of the k-Same-furthest-FET faces is nearly zero regardless of k, indicating that the expression transfer process after k-Same-furthest de-identification has hardly any impact on the privacy protection performance.

C. Evaluation of Data Utility

Following the work on k-Same-Select [10] which also aims to preserve data utility, data utility preservation ability of the proposed k-Same-furthest-FET algorithm is measured in terms of the accuracy of a facial expression classifier. Here Linear Discriminant Analysis (LDA) is employed as the classifier. The LDA classifier is trained with the original cropped neutral and happy faces and tested with happy faces de-identified using AAM-based k-Same-Select and the proposed k-Same-furthest-FET. Fig. 8 presents the classification accuracy obtained. Some happy faces in the IMM dataset display an ambiguous smile. When any of these faces is one of the k closest happy faces to an original happy face, it will make the de-identified face by k-Same-Select (the average of the k closest) display an even more ambiguous smile that an expression classifier cannot detect. In contrast, the proposed algorithms transfer the original expression after the de-identification process, aiming to generate a face with an identical expression to the original. Hence, when the output faces are classified with a classifier trained with the original expressions, it is more likely for the classifier to detect the expression cloned by k-Same-furthest-FET. Fig. 8 confirms that k-Same-furthest-FET preserves expressions more than k-Same-Select.
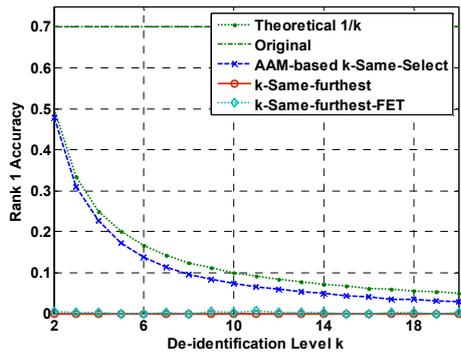
Fig. 7. Recognition rates for cropped original and de-identified faces.
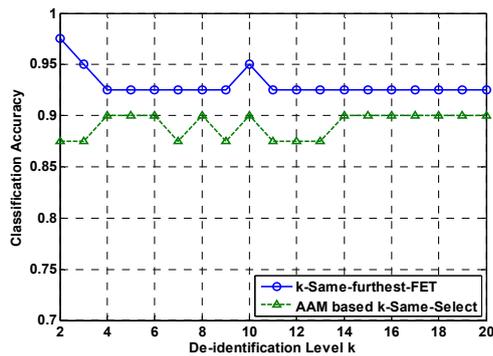
Fig. 8. Data uility preservation ability as measured by facial expression classification accuracy.

VI. CONCLUSION

This paper presents a novel algorithm named *k*-Same-furthest-FET, performing facial expression transfer after the *k*-Same-furthest face de-identification. Despite the fact that the FET process of the proposed *k*-Same-furthest-FET does not satisfy *k*-anonymity, experimental results show that it has hardly any negative impact on the recognition rate, achieving nearly perfect privacy protection for all faces with any value of *k*. In addition, the FET process enables the expression of the original face to be effectively cloned on the de-identified face. As confirmed by the experimental results from data utility evaluation, *k*-Same-furthest-FET preserves expression better than *k*-Same-Select. Furthermore, visual results of the output faces demonstrate that *k*-Same-furthest-FET is able to preserve not only the category but also dynamic details of various facial expressions. This is of significant value in face de-identification in videos, which inevitably contains various dynamic expressions with various degrees of intensity.

REFERENCES

[1]   J. Crowley, J. Coutaz, and F. Berard, "Things that see," Communications of the ACM, vol. 43, pp. 54–64, 2000.

[2]   C. Neustaedter, and S. Greenberg, "Balancing privacy and awareness in home media spaces," Workshop on Ubicomp Communities: Privacy as Boundary Negotiation, in conjunction with the 5th Int'l Conf. Ubiquitous Computing (UBICOMP), Seattle, WA , 2003.

[3]   M. Boyle, C. Edwards, and S. Greenberg, "The effects of filtered video on awareness and privacy," Proc. ACM Conf. Computer Supported Cooperative Work, 2000.

[4]   E. M. Newton, L. Sweeney, and B. Malin, "Preserving privacy by de-identifying face images," IEEE Trans. Knowledge and Data Eng., vol. 12, no. 2, pp. 232 – 243, Feb. 2005.

[5]   C. Neustaedter, S. Greenberg, and M. Boyle, "Blur filtration fails to preserve privacy for home-based video conferencing," ACM Trans. Computer Human Interactions (TOCHI), vol. 13, issue 1, Mar. 2006.

[6]   L. Sweeney, "k-Anonymity: a model for protecting privacy," Int'l J. Uncertainty, Fuzziness, and Knowledge-Based Systems, vol. 10, no. 5, pp. 557–570, 2002.

[7]   R. Gross, L. Sweeney, F. de la Torre, and S. Baker, "Model-based face de-identification," IEEE Workshop on Privacy Research in Vision, 2006.

[8]   G. Edwards, C. Taylor, and T. Cootes, "Interpreting Face Images Using Active Appearance Models," Proc. FG'98, pp. 300–305, Apr. 1998.

[9]   L. Meng, and Z.J. Sun, "Face De-identification with Perfect Privacy Protection," submitted to the 37th Intl. Convention MIPRO, Special Session on BiForD, Opatija, Croatia, May 2014.

[10] R. Gross, E. Airoldi, B. Malin, and L. Sweeney, "Integrating utility into face de-identification," Workshop on Privacy-Enhanced Technologies, 2005.

[11] P. Ekman, "Facial expressions," Chapter 16 of Handbook of Cognition and Emotion, John Wiley & Sons Ltd., 1999.

[12] I. T. Jolliffe, Principal Component Analysis, 2nd ed., New York: Springer-Verlag New York, Inc, 2002.

[13] I. Matthews, and S. Baker, "Active appearance models revisited," Int'l J. Computer Vision, vol. 60, no. 2, pp. 135-164, Nov. 2004.

[14] A. Asthana, J. Saragih, M. Wagner and R. Goecke, "Evaluating AAM fitting methods for facial expression recognition," ACII 2009, pp. 1-8.

[15] L. van der Maaten, and E. Hendriks, "Capturing appearance variation in active appearance models," Computer Vision and Pattern Recognition Workshops (CVPRW) 2010, pp. 34-41.

[16] M. Turk, and A. P. Pentland, "Eigenfaces for recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[17] E.B. Roesch, L. Tamarit, L. Reveret, D. Grandjean, D. Sander, and K.R. Scherer, "FACSGen: a tool to synthesize emotional facial expressions through systematic manipulation of facial action units," J Nonverbal Behav., vol. 35, issue. 1, pp. 1-16, Nov. 2010.

[18] B. Theobald, I. Matthews, J. Cohn, and S. Boker, "Real-Time Expression Cloning Using Appearance Models," Proc. Ninth Int'l Conf. Multimodal Interfaces (ICMI '07), pp. 134-139, 2007.

[19] M. de la Hunty, A. Asthana, and R. Goecke, "Linear facial expression transfer with active appearance models," 20th Int'l Conf. Pattern Recognition (ICPR), pp. 3789-3792, 2010.

[20] K. Bowyer, K. Chang, P. Flynn, "A survey of 3D and multi-modal 3D+2D face recognition," Notre Dame Computer Science and Engineering Technical Report, 2004.

[21] A. M. Bronstein, M. M. Bronstein, R. Kimmel, "Expression-invariant 3D face recognition," Proc. Audio- and Video-based Biometric Person Authentication (AVBPA), Lecture Notes in Comp. Science No. 2688, Springer, pp. 62-69, 2003.

[22] Y.I. Tian, T. Kanade, and J.F. Cohn, "Recognizing action units for facial expression analysis," IEEE Trans Pattern Anal Mach Intell, vol. 23, no. 2, pp. 97-115, 2001.

[23] M.M. Nordstrøm, M. Larsen, J. Sierakowski, and M.B. Stegmann, "The IMM face database - an annotated dataset of 240 face images," Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, May 2004.