Face De-Identification

Aruna Shenoy

National Physical Laboratory (UK)

and

Lily Meng

University of Hertfordshire (UK)

Need for de-identification

- In order to protect the individual.
- If individuals are de-identified then it is easier for data sharing between organizations for research, business, academic, security and many other purposes.
- Ethical and legal responsibilities - restricted data sharing and utilization.

- How can de-identification be achieved –

  - AD-HOC means

    MASKING/PIXELATION/BLURRING

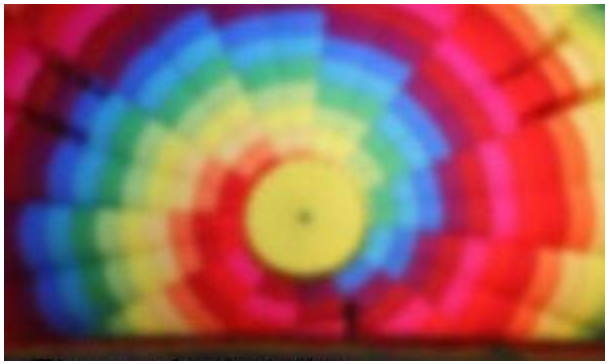
  - *K*- Anonymity based methods such as *k*-same

AD-HOC Methods -Problems

- Masking/Blurring/Pixelation

Original Image

Blurred Image



http://en.wikipedia.org/wiki/Gaussian_blur

- Due to the advances in robust face recognition software, it fails to serve the purpose of thwarting.
- It also significantly distorts the integrity of the entire image.

k-Anonymity based methods

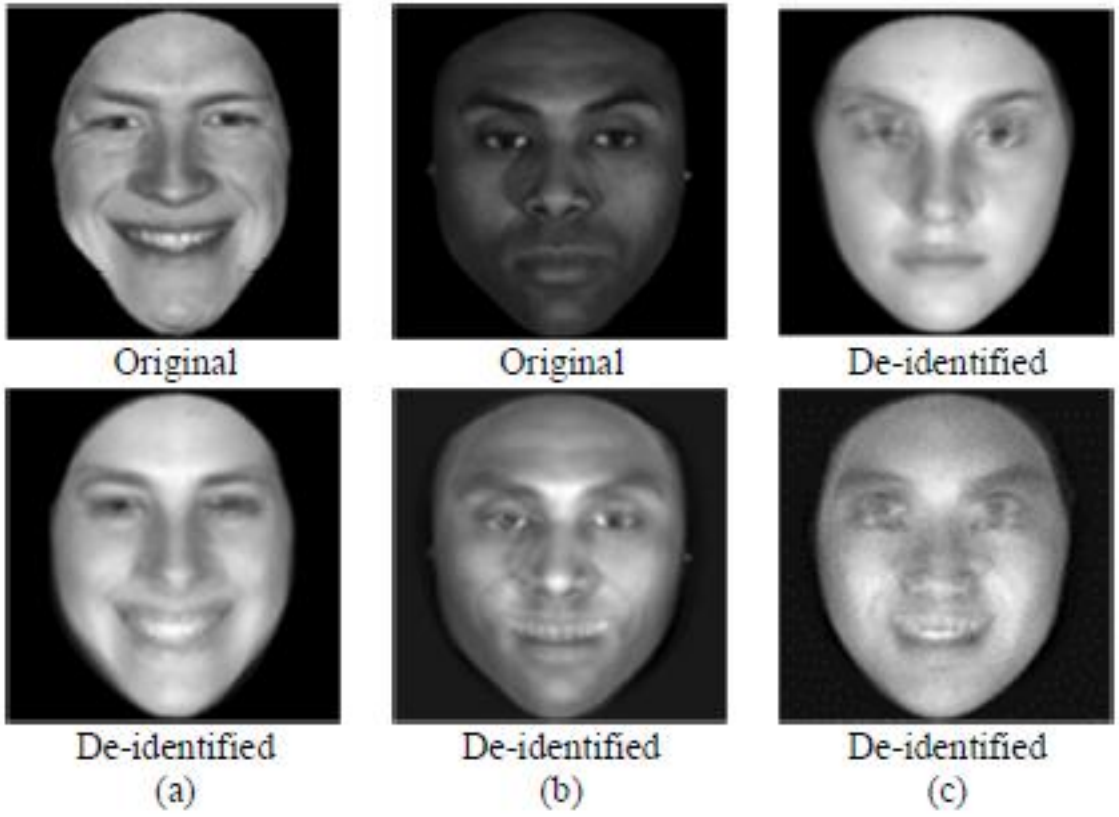
- Overcomes the problems with Ad-hoc methods – The de-identified image can successfully thwart the available face recognition system.

- Principle -– All *k*-anonymity based methods de-identify based on average k faces from the gallery and achieve privacy protection.

- The recognition rate of the de-identified images < 1/*k*

Problems with k- Anonymity

- No data utility
- It loses information : gender, expression, age, race etc.



Problems with the *k*-same algorithm

a) Loss of Gender
b) Loss of expression
c) Ghost artifacts

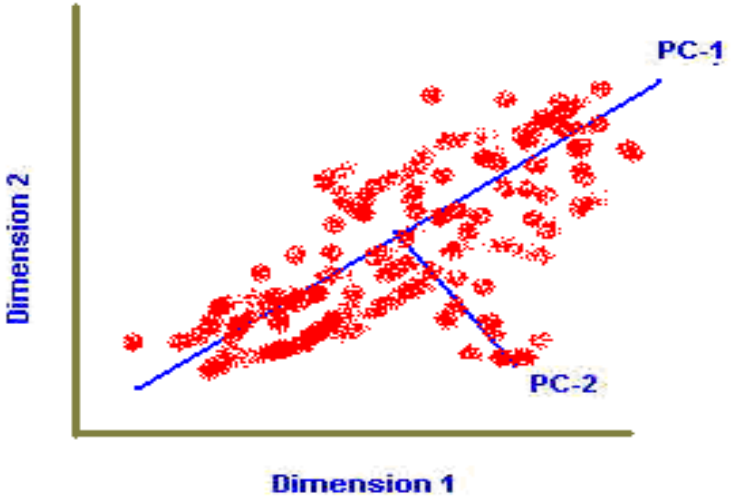
k-Same Class-Eigen – OUR METHOD

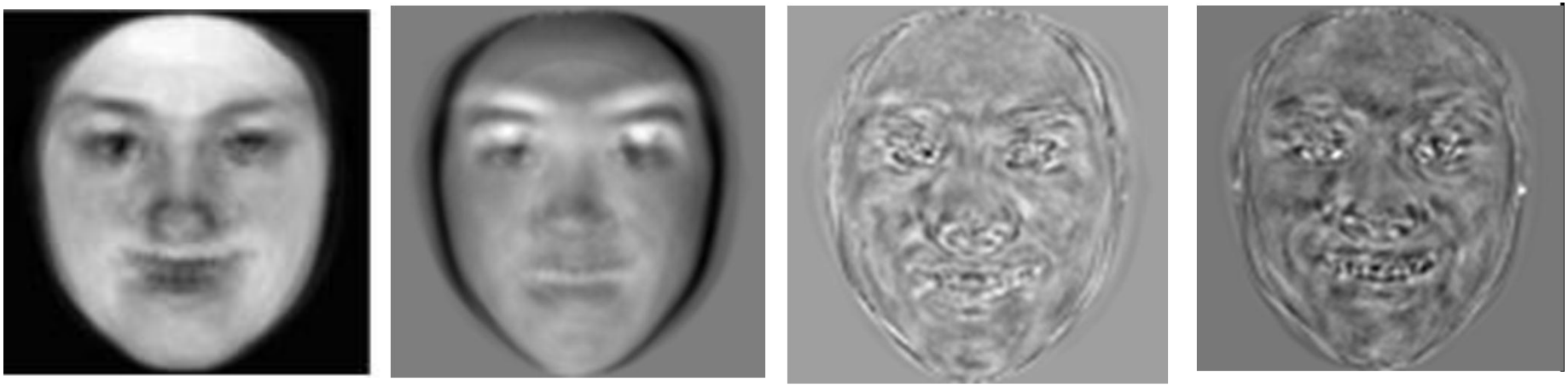
- *k*-same Pixel
- *k*-same Eigen
- Both versions identify the *k* closest faces to the probe in the PCA face space
  - *k*-same Pixel - returns the pixel-wise average of the *k* closest faces.
  - *k*-same Eigen - returns the average of the PCA face space.
- Our method : PCA, LDA, Encoding power.

Principal Component Analysis (PCA)

- It transforms higher dimensional datasets into lower dimensional data while trying to retain as much information as possible.
- The greatest variance by any projection of the data set comes to lie on the first axis, the second greatest variance on the second axis, and so on.
- The first principal component or eigenvector for face is called the eigenface.

Figure shows the first two consecutive principal components.





The first two and last two Eigenfaces in this work.

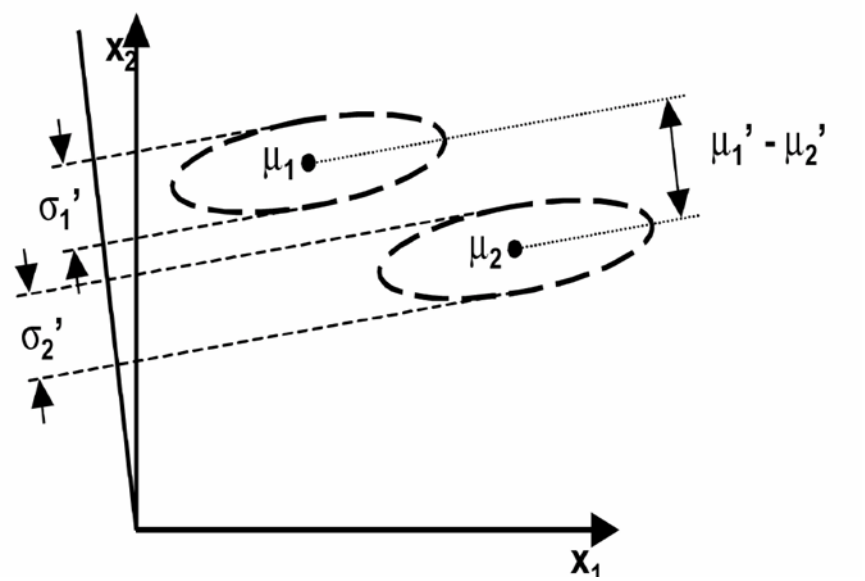
- All eigenvectors with a nonzero eigenvalue are kept to avoid losing information on data utility.

PCA…

- Although PCA is effective in terms of maximizing the scatter among individual face images, it ignores the underlying class structure and the projection axes chosen by PCA might not provide good discrimination power for classification purposes.
- The first few components obtained by PCA may have maximum variance but may not be of interest .
- if the property of interest of the data is encoded by the last few components then this method would be disadvantageous.
- Hence, the selection of the components should be such that they are based on the importance of the property rather than the total variance.

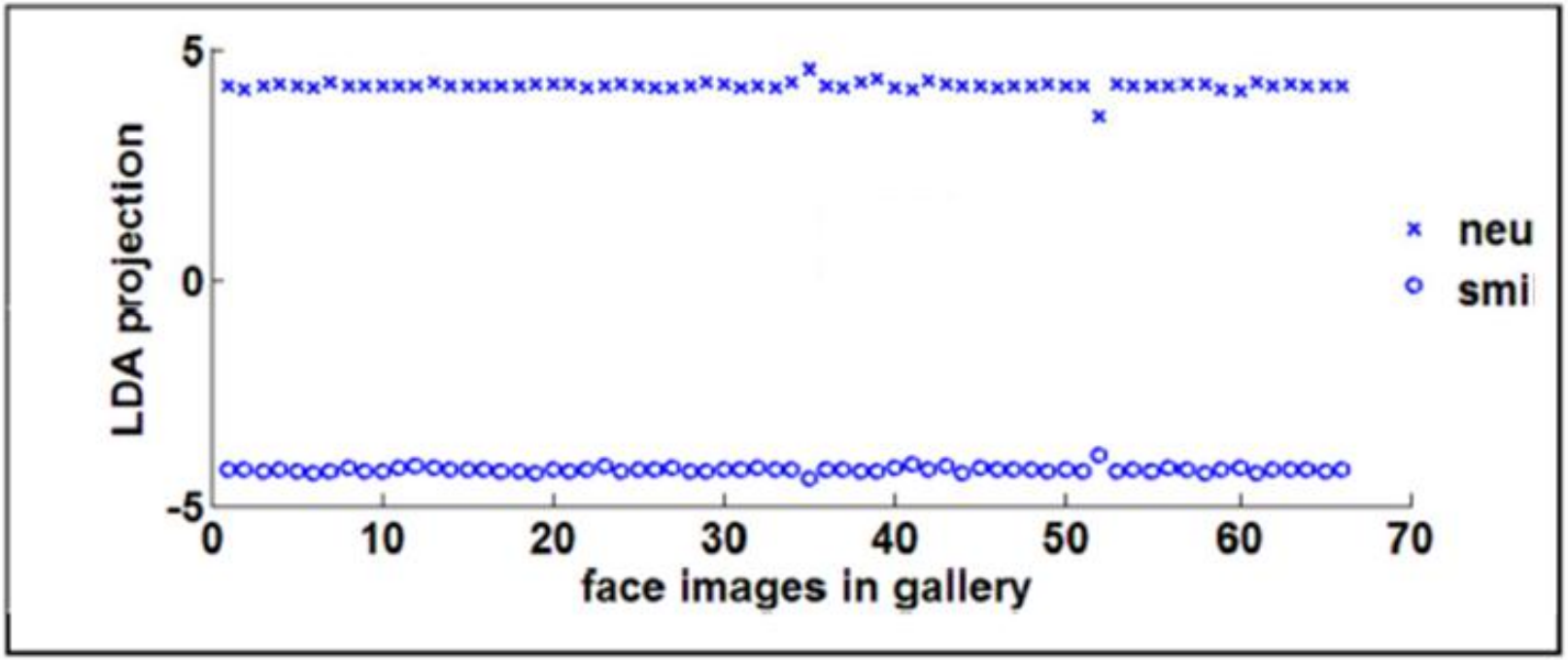
Linear Discriminant Analysis (LDA)

- maximizing the between-class scatter
- minimizing the within-class scatter



The classes can be based one various face properties such as expression, gender, age, identity and race.

LDA projection of gallery images



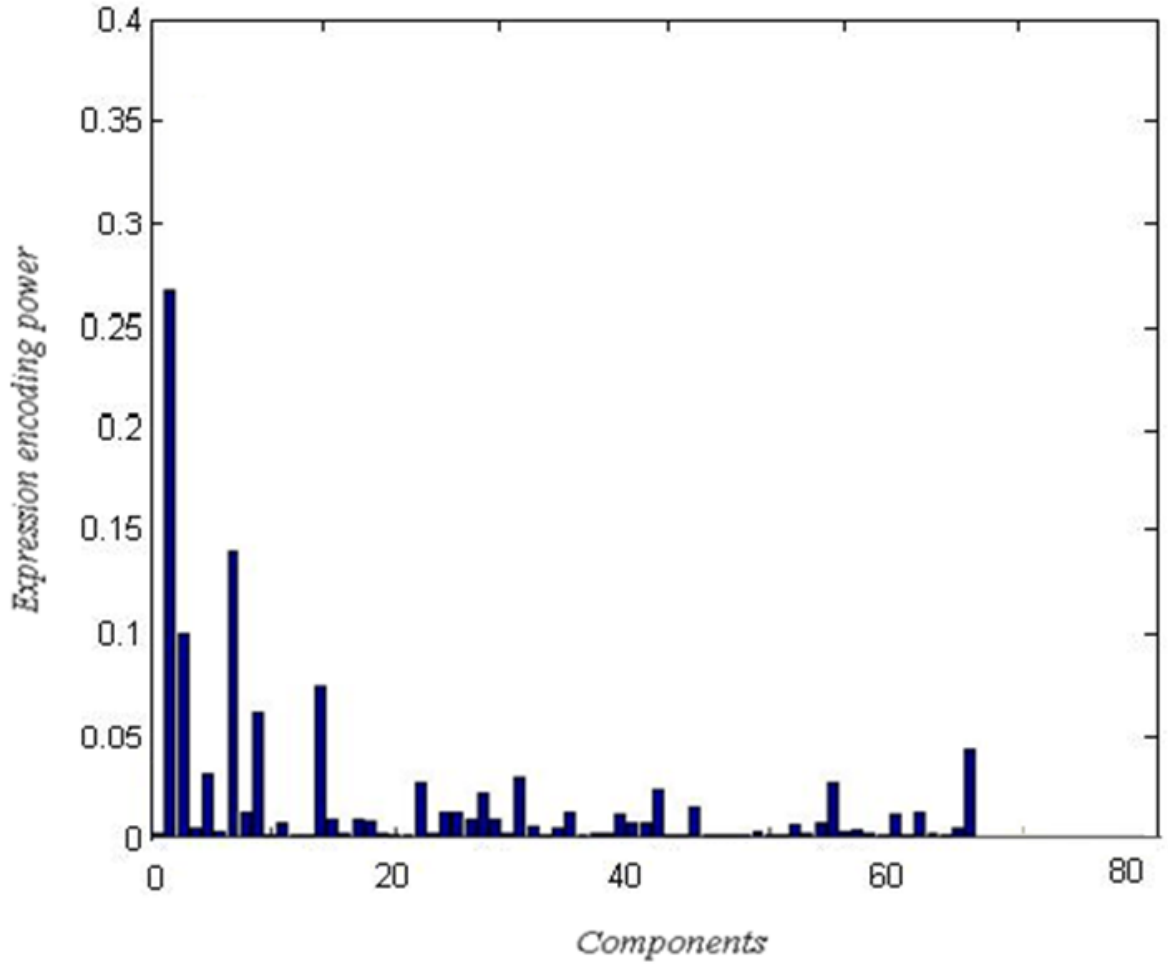
The LDA projection of all the gallery neutral faces used in this work has an average projection value of 4.2 with a small variance of 0.004; while the average LDA projection of the smiley faces is -4.2 with a small variance of 0.011.

1.We first perform PCA to reduce the dimensionality.
2. Perform LDA based on two classes.
3. Find the Encoding power component.

Encoding power

- The discriminating power is defined as the ratio of projection of the between-class variance to the projection of the within-class variance.

- With the PCA of two classes, the first components encode information common to both classes of faces, whilst the latter components encode information not so common between the two classes.



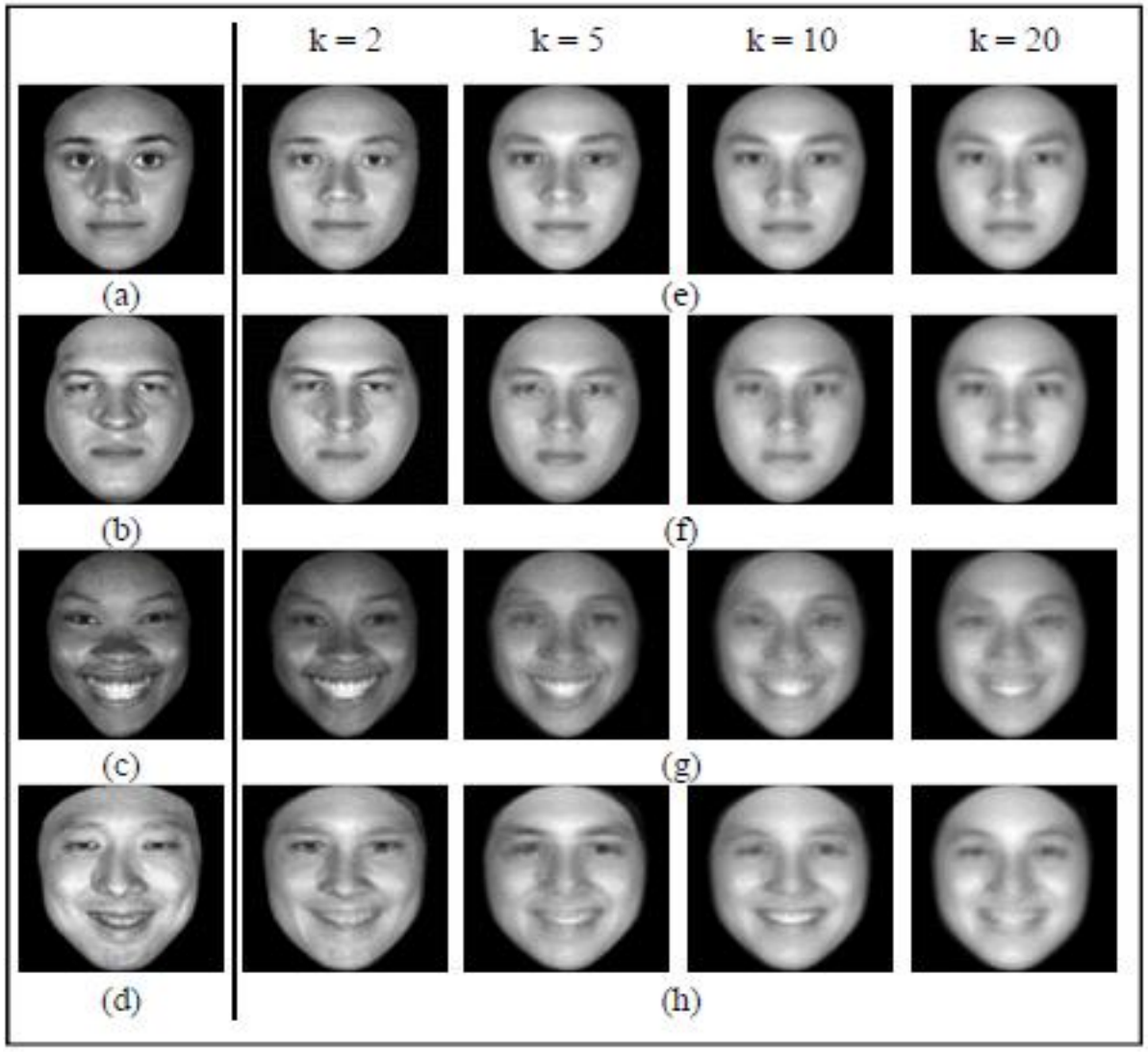
Encdoing power between Neutral and Smiling faces

Dataset used :

Examples of face images from the BINGHAMTON BU-3DFE dataset.

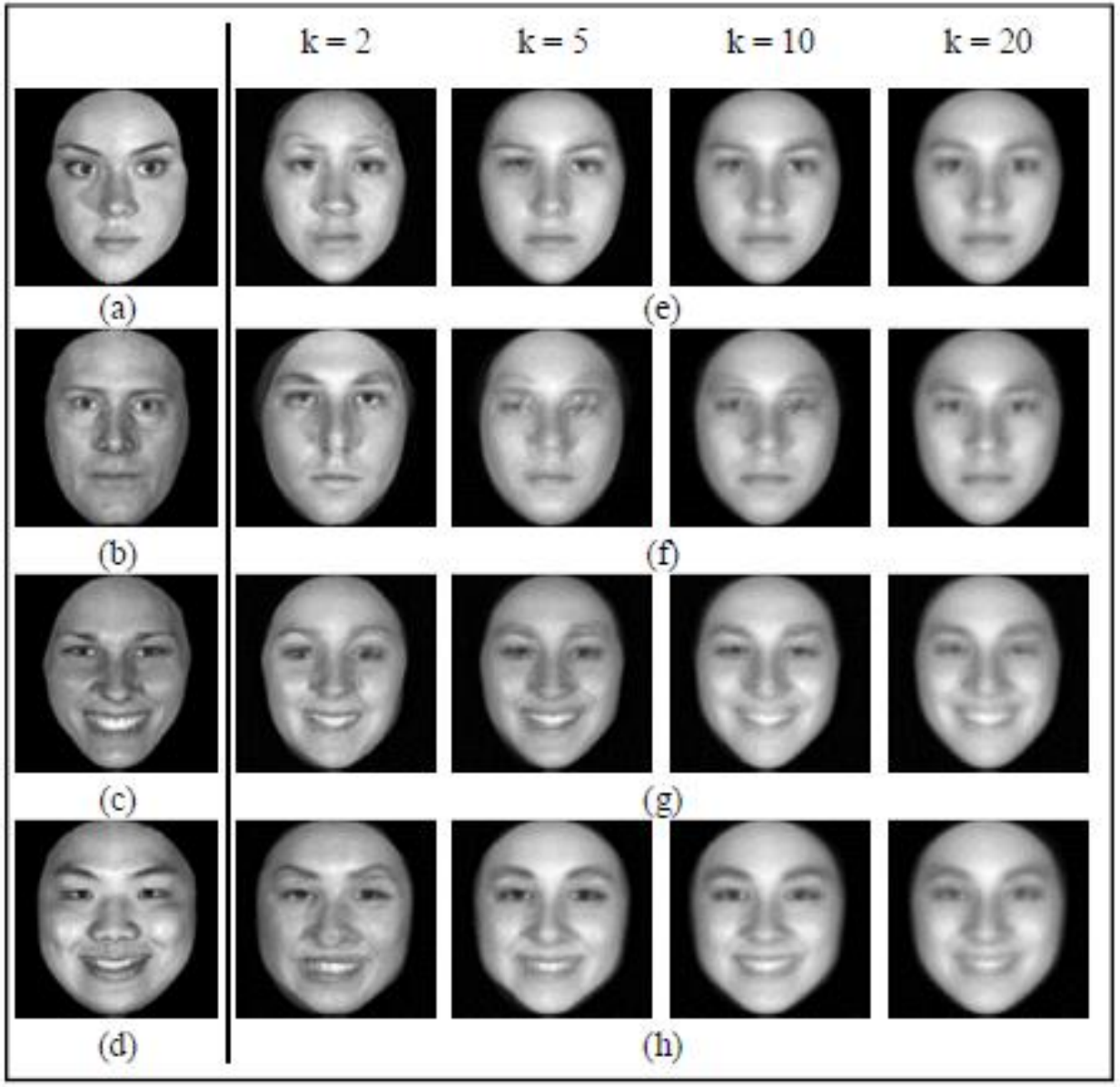
Each row is a subject showing various expression (left to right) neutral (NE), happy (HA), angry (AN), fear (FE), sad (SA), surprise (SU) and disgust (DI).

Results



Examples of original seen face images and their de-identified faces from k-SameClass-Eigen for k = 2, 5, 10, and 20 from left to right:

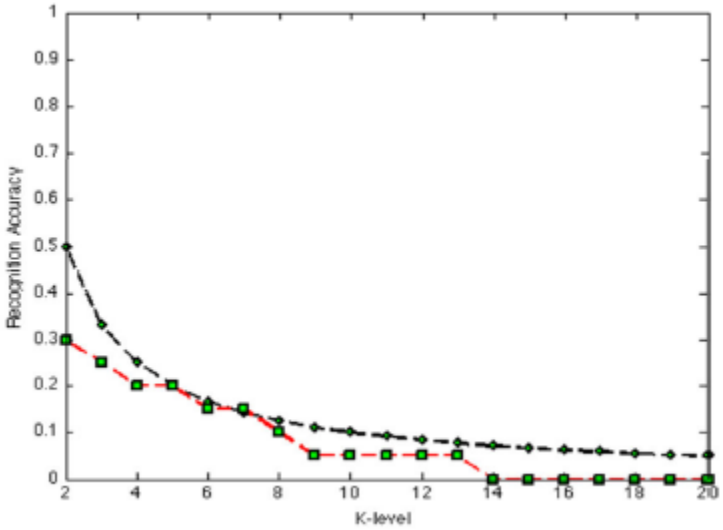
(a) original female neutral
(b) original male neutral
(c) original female smiling
(d) original male smiling
(e) - (h) de-identified results for (a) - (d)



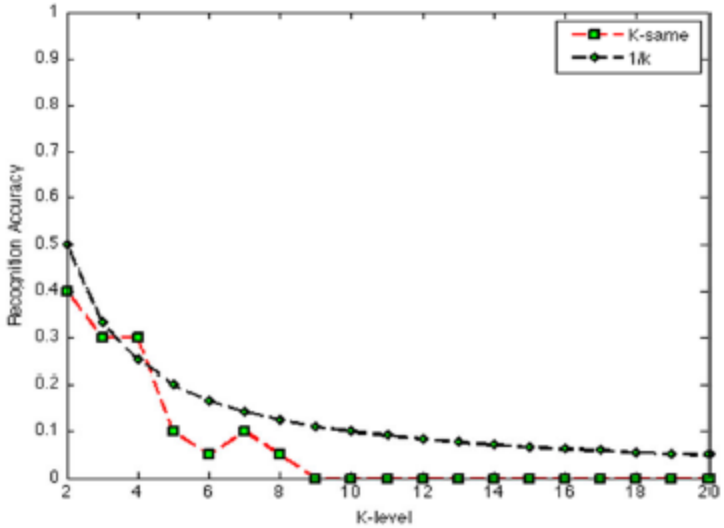
Examples of original unseen face images and their de-identified faces from k-SameClass-Eigen for k = 2, 5, 10, and 20 from left to right:

(a) original female neutral
(b) original male neutral
(c) original female smiling
(d) original male smiling
(e) - (h) de-identified results for (a) - (d)

Evaluation of Data Utility - Identity

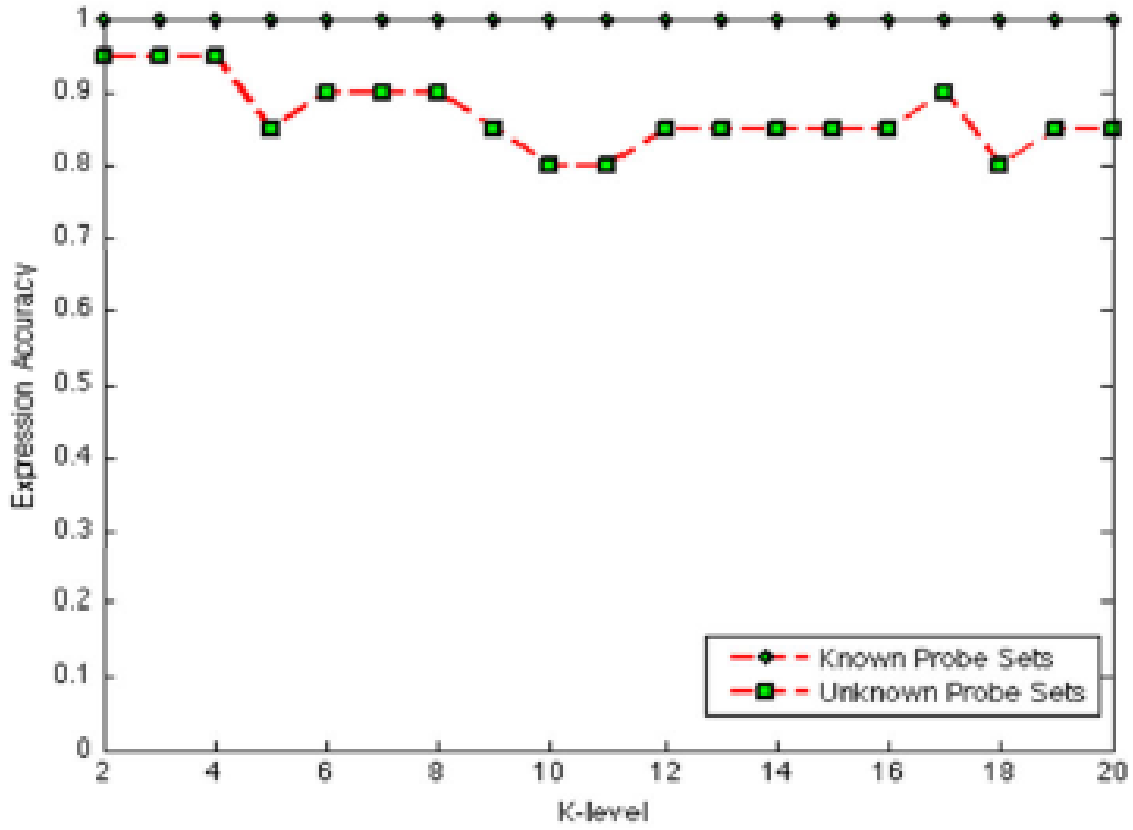


PCA recognition rates for unaltered gallery images and de-identified probe images. Here, probe images are taken from gallery images

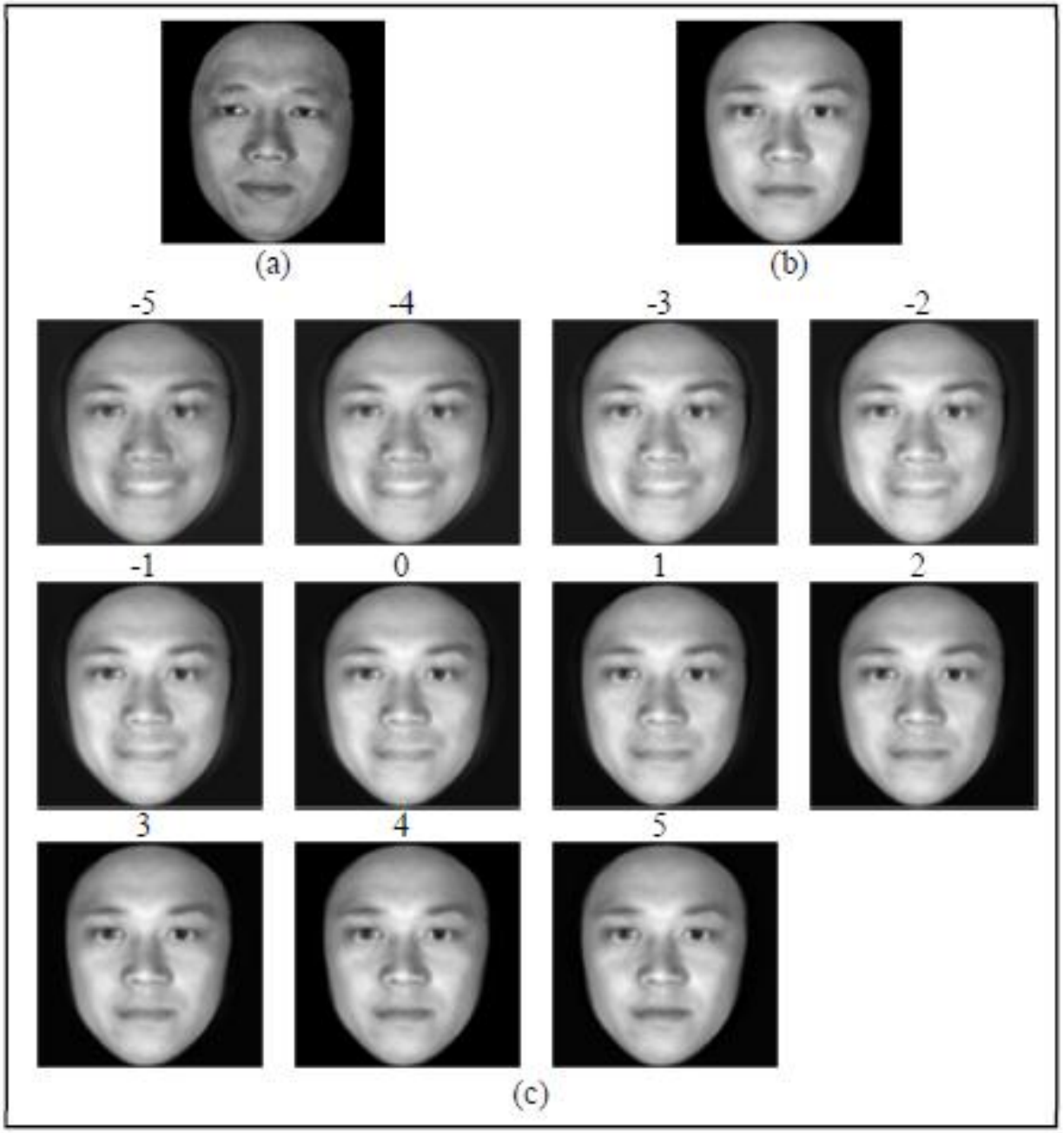


PCA recognition rates for unaltered gallery images and de-identified probe images. Here, probe images are unknown to gallery images

Evaluation of Data Utility - Expression



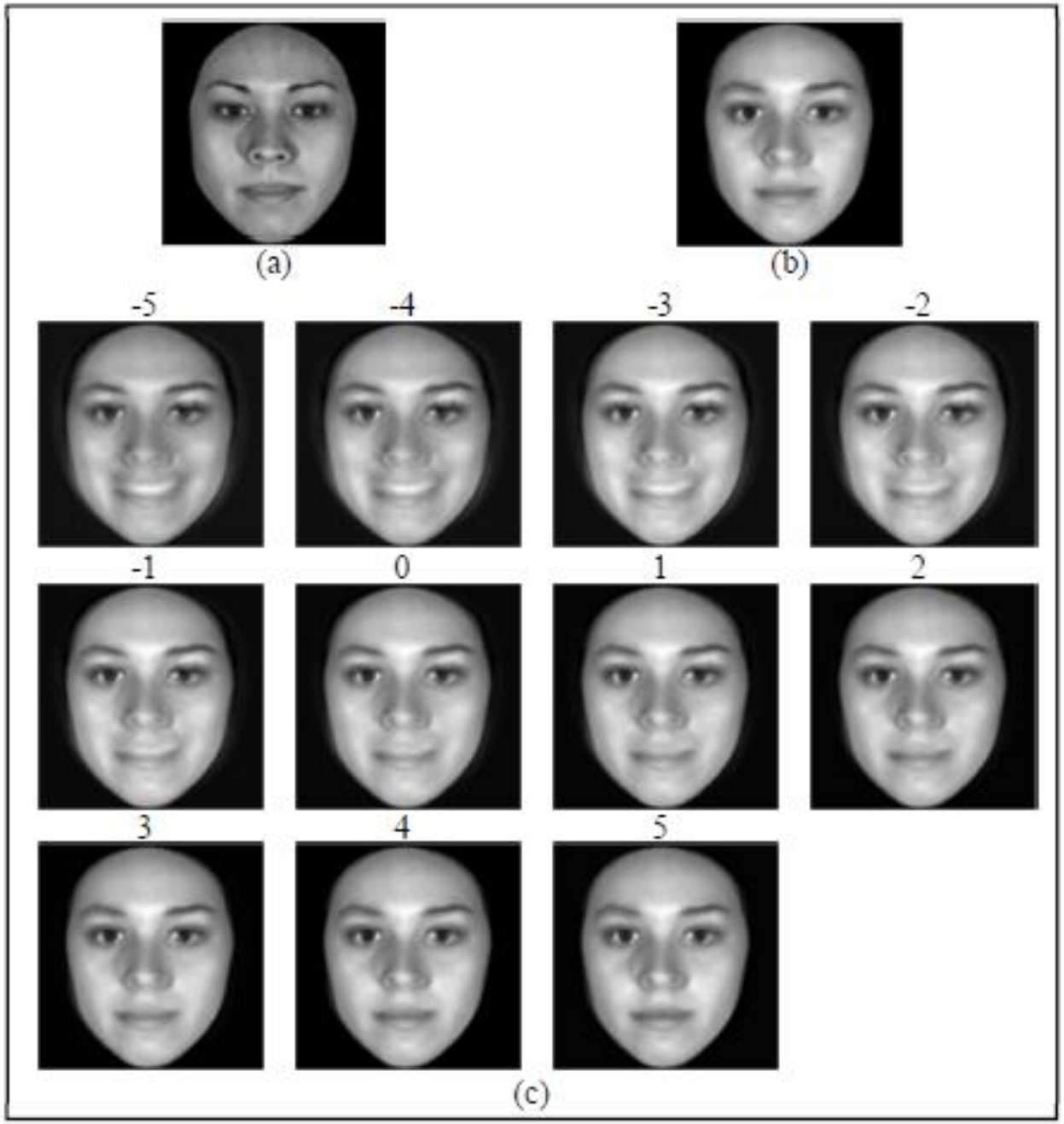
Expression recognition accuracy for both sets of probe images (known and unknown to gallery images)



NPL
National Physical Laboratory

An example male face:

(a) original
(b) de-identified
(c) variations of (b) for various expression intensities (intensity values are above the face images)



NPL
National Physical Laboratory

An example female face:
(a) original
(b) de-identified
(c) variations of (b) for various expression intensities (intensity values are above the face images)

Conclusions

- Our *k*-SameClass-Eigen has been proposed in this work with a goal to preserve privacy as well as retain both the expression class and the expression intensity.
- Experimental results show that it is able to limit recognition rate to below 1/*k*.
- Although *k*-SameClass-Eigen always retains the expression on known face images, it failed with a few unknown faces - a much larger training set may address this problem.
- It is intended to be used for Pain expression dataset.
- Will be extended to other utility such as gender.