Research Journal of Applied Sciences, Engineering and Technology 8(6): 795-802, 2014
DOI:10.19026/rjaset.8.1036
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: June 11, 2014 Accepted: July 13, 2014 Published: August 15, 2014

Research Article

INFACE and EPOCH Database: A Benchmark for Face Recognition in Uncontrolled Conditions

[1]M. Parisa Beham and [2]S. Md. Mansoor Roomi
[1]Vickam College of Engineering,
[2]Thiagarajar College of Engineering, Madurai, Tamil Nadu, India

Abstract: The main focus of our study is to build a database of labeled face images display with wide variations in pose, lighting, appearance and age. Recognizing human faces amid natural setting is emerging as a critically important and technically challenging computer vision problem. Most of the previous cases of analysis to study the specific variations of the face recognition problem focused on recognition of faces captured under controlled environment in standard laboratory setting. These variations include position, pose, lighting, background, camera quality and gender. But in real environment, there are innumerable applications in which there is little or no control over such variations. In this study, we introduce two novel database viz. "INFACE" and "EPOCH", which can effectively contribute to the face recognition research, in general and suited to the Indian setting (appearance, expressions, etc.) in particular. The former data base ("INFACE"), includes around 7400 face images of 37 individuals collected from approximately 45 Indian movies. The images of later data base ("EPOCH") portray the variations with age. It contains 1000 images of 10 individuals spanning four age groups. The data base specimens of INFACE have high degree of variability (in scale, location, orientation, pose, expression, illumination, appearance and degree of occlusion) which one could ever see in natural world. EPOCH database is providing a large set of unconstrained face images which includes age variance of an individual. In addition to the individuality of the face, these databases provide additional information on pose, gender, expression, location etc. Specific annotations for each image have been stored as XML for easy retrieval. These databases will be made public to catalyze research and development for accurate face recognition.

Keywords: Aging database, expression variation, face annotation, face detection, movie database, pose invariant

INTRODUCTION

Face Recognition (FR) is a fascinating and popular research area within the domain of computer vision and pattern recognition. Many versatile and exhaustive data bases are available for bench-marking of face recognition algorithms (Georghiades *et al*., 2001; Martnez and Benavente, 1998; Albert and Ricanek, 2008; Sim *et al*., 2003; Huang *et al*., 2007). Endurance of these databases has helped extensively to define the expectations on face recognition algorithms. Over the years, there has been steady improvement in the performance of recognition algorithms. Their efficiency has improved typically from about 60% to almost 100%. In this study, we are interested in introducing two novel challenging databases that can critically help the face recognition research in general and in an Indian setting, in particular. Our main motivation to bring in a new database is to provide a large set of unimpeded face images which includes all the issues one can face in a Face recognition system. A sample set of face images of a specific individual showing variety of expressions is shown in Fig. 1. As more face recognition algorithms are being developed, new approaches that use multiple features such as skin color, size, shape and presence of facial features are being developed. The facial database of LFW and PubFigs which are collected from web encompass only limited ranges of color and expression etc. and may not cover the effect of natural aging. These enumerated constraints have inspired the authors to spot light their focus on acquiring exhaustive data base with minimum environmental and aging constraints so that such an algorithm turns out to be a real standard benchmark for face recognition method.

Fig. 1: Scale of variations w.r.t. poses, expression, occlusion and illumination, of an actor Sudeep and scale of variations w.r.t. age of an actor Kamalhassan faces in INFACE and EPOCH databases respectively

Corresponding Author: M. Parisa Beham, Vickam College of Engineering, Madurai, Tamil Nadu, India

Yale (Georghiades *et al*., 2001), AR (Martnez and Benavente, 1998), FERET (NIST, 2003), PIE (Sim *et al*., 2003) and LFW (Huang *et al*., 2007) are the well-liked face databases. Establishing accuracies on databases like Yale, AR and PIE have emerged as the mandatory requirement for reporting progress in face recognition. After several decades of research, steady performance has been achieved on these databases. This has catalyzed the systematic progress in this area, especially for face recognition in laboratory settings (or in cooperative environments). We are interested in introducing a new database that can help advancing face recognition in the next few decades in a natural settings and unconstrained environments. We have kept the important design guidelines in mind while constructing the database. First, we would like to build a database of images that are captured using cameras with different resolutions. Unlike in many of the previous cases (e.g., Yale, AR), we would like to capture faces in a wild. This results in face images with wide variations in Age, pose, expressions and illumination and partial occlusions. Construction of database is a onetime effort. However, we would like to cover faces which were captured over years. This can help in designing recognition schemes that can robustly recognize individuals independent of the age. This demanded the use of images that were captured over a wide span of years and not captured as part of a specific effort. Second, we would like to cover a wide variety of facial expressions that one could ever see in natural world around. This is often hard to obtain in a laboratory setting. Capturing faces with variations in facial expressions may need the faces to be captured from natural videos. In order to make the expressions as natural as possible and therefore capture these expressions from the renowned movie actors and in the best possible setting. In addition to that, there are very few albums of faces for use in research, none of which adequately represent the different ages of faces across the lifespan. This lack of a range of ages has limited the majority of researchers to do age invariant FR.

We find that faces captured from Indian movies could meet the diverse requirements as listed above. In this study, we present two face databases viz, INFACE database (7400 face images of 37 individuals), EPOCH database (1000 face images of 10 individuals) that are captured, cropped and carefully annotated from approximately 45 Indian movies (including short Movies and TV serials). Specific details of the database are discussed in Section below. In addition to the impact on face recognition research, this database has utility for various related vent tasks like:

- Face pose estimation
- Age estimation
- Expression recognition
- De-Identification
- Photometric and geometric calibration of cameras
- Image forensics etc

INFACE and EPOCH are very complex databases. Algorithms that can give more than 90% accuracy on popular databases yield poorer performance on our database. This variation in performance is attributed to the large size of INFACE database with numerous facial expressions, varied illuminations, temporal variations and wide variety of occlusions for each individual also wide and hard variations of EPOCH database. These databases as well as annotations will be made publicly available for research.

MATERIALS AND METHODS

In this section, the evolution of "INFACE" and "EPOCH" databases is described. The images in these databases display wide variations in pose, lighting, background and appearance. The database of ageing images of a same person has been collected starting from 18 years of age to 60 years. The ageing database also has wide variations of pose, illumination, expressions and occlusions. The existing databases do not provide such additional characteristics. Some of the variations in face appearance are due to factors such as motion, occlusions and expressions, which portray real characteristics as they are acquired through uncontrolled environmental conditions. The lavishness of the images included in this collection, however, motivated us to build an index of all of the faces present in a subset of images from this collection. We believe that these databases will greatly help in devising an apt benchmarking of face recognition algorithms with realistic estimates of their expected performance in unconstrained settings.

INFACE database: Most of the popular face databases are captured faces only in the constrained environments with artificial light settings (Georghiades *et al*., 2001; Martnez and Benavente, 1998). Recent years had also seen introduction of novel databases which are captured in the wild. They are often harvested from the web. Prominent among them are Labeled Faces in the Wild (LFW) (Huang *et al*., 2007) and PubFig (Setty *et al*., 2013). The below paragraph, list the most notable face image database taken in controlled settings:

- **FERET database (NIST, 2003):** In this database, each image consists of an individual (head and neck visible only) on a uniform and uncluttered background covered only frontal and 3 pose variations. Note that FERET deals with a very large number of classes, but the number of classes is not the issue in this study. Our main concern here is with the problems caused by insufficient data per class available for learning.
- **CMU/PIE (Sim *et al*., 2003):** It consists of frontal and near frontal view images on a cluttered background. These databases are more useful as training sets rather than test sets for face detection purposes.

Fig. 2: Face images from INFACE database: widely varying appearance of faces based on (a) illumination, (b) make-up, (c) occlusions, (d) pose, (e) expressions

- **Yale and extended yale B (Georghiades *et al*., 2001):** Yale consists of 10 individuals and Extended Yale B consists of 38 individuals with 2412 images. Both Yale databases are varying in high degree of illuminations taken under room lights ON and OFF conditions.

Introduction of these databases have exposed a number of challenging problems in face recognition in recent years. In our study, we have collected Indian movie videos (more from Tamil movies) of various types from different languages. The Viola-Jones face detection algorithm (Jones and Viola, 2003) is applied for each movie individually. Then extracted all the faces and faces having more variance are collected. The corresponding frames are also extracted and saved for the purpose of auxiliary information. The resolution of the images is restricted to 80×80. An actor/actress faces are collected from different movies and scenes. We provide bounding box level annotations of the faces which are not present in most of the face databases. Notably, all our faces are from Indian movies and faces of people of Indian origin. PubFigs often want to retain the identity (appearance, dress patterns, expressions) over sessions, while actors in the movies want to change the appearance and expressions and adapt to the new role present in the movie. Thus INFACE database has totally 37 individuals with 200 samples per each person. This makes INFACE a unique database to work with. We have also done annotation for each and every image in the INFACE database. To obtain annotations we carefully analyzed each and every image across above seven parameters and stored respective values.

In addition to above information name of the character with order, image, frame numbers and name of the movie with year also mentioned. In this way each face is completely annotated and such information is stored in the form of XML files. Figure 2 shows an example of 60 face images from INFACE database. From every 37 individuals one or two sample images have been taken. This gives a quick glimpse of all the Indian actors selected for constructing the INFACE database.

EPOCH database: In recent years, a new dimension has been added to the problem of face recognition. Age as an attribute related to human faces is being increasingly studied and there has been a growing interest in problems such as face recognition across ages, automatic age estimation from face images, appearance prediction across aging etc., (Ramanathan *et al*., 2009). Given the difficulty in compiling face datasets that comprise age-separated face images, there are very few publicly available datasets (BERC, 2010; Albert and Ricanek, 2008) that specifically address facial aging in comparison to the datasets (FGNET, 2009; NIST, 2003) that address other problems in face recognition. There are three publicly available databases that comprise age separated face image samples namely, the MORPH (Albert and Ricanek, 2008), the FG-NET (FGNET, 2009), BERC (BERC, 2010) and PAL (Minear and Park, 2004) aging database. The available database faces from LFW (Huang *et al*., 2007) and PubFigs (Setty *et al*., 2013) are collected from internet sources and in turn they may not cover momentous age variations. Inspired by this we proposed an aging database called EPOCH database by collecting images of the various actors and actresses, acting from their age from 18 to 60 years. To show the age variance in the database, we collected the images from the movies of the actors for age span of 18-25, 26-32, 33-45 and 46- 60. An example of face images from the four age ranges of EPOCH database is shown in Fig. 3. Each row shows aging variations of an individual.

Fig. 3: Face images from EPOCH database: face images of an individual illustrating the variations due to aging

Related aging databases: In this section we provide a brief account on popular, publicly available databases, in relevance to facial aging. The BERC database was collected by Bio-metrics Engineering Research Center (BERC) in order to estimate the ages of Koreans. The database contains 390 people's images in the age range 383 years old. The facial images were obtained using a digital camera (NIKON cool-pix p5000) with a very high resolution of 3648×2736 pixels. All of the images have no variance in light and expression. The PAL aging database contains 430 Caucasians with age range 1893 years old (Minear and Park, 2004). The images in the database were captured using a digital camera with fixed light and position conditions. The database includes various expressions such as smiling, sadness, anger, or neutral faces. The FG-Net aging database (FGNET, 2009) is the most frequently used database for estimating age in the previous works, as it is a publicly available database. The database has 1002 images composed of 82 Europeans in the age range 0-69 years old. Individuals in the database have one or more images included at different ages. Images were obtained by scanning, unlike the other databases. Therefore, there are extreme variations in lighting, expression, background, pose, resolution and noise from scanning. Since the images were retrieved from real-life albums of different subjects, aspects such as illumination, head pose and facial expressions, etc. are uncontrolled in this dataset. Similarly the MORPH database (Albert and Ricanek, 2008) comprises face images of adults taken during different ages. The database has been organized into two albums: MORPH Album 1 and MORPH Album 2. MORPH Album 1 comprises 1690 digitized images of 515 individuals under the age range 1568 years. MORPH Album 2 comprises 15,204 images of nearly 4000 individuals. Apart from the face images, the database also provides meta-information that is critical for the task of studying age progression such as age, sex, ethnicity, height and weight. In all the above standard databases the images are taken under the controlled environment or from the internet. The face images are taken under controlled illumination, occlusions, expression and pose variations. Motivated by this, we have taken an initial attempt to build an aging database where the images are captured in unconstrained environment. In our proposed EPOCH database, we have identified four age groups spanning the age of 18 to 60 years with wide variations in illumination, make-up, occlusions, expressions and pose. To extract the faces from the movies the same experimental setup used for INFACE has also been implemented for EPOCH database. Thus to account for real life and natural conditions, our proposed aging database include age variations as well as changes in the face due to illumination, pose, occlusion and make-ups.

Construction details: In this section, we provide an analysis of the database that gives some insights about its construction. INFACE and EPOCH is designed with a careful observation in order to have a diverse range of variations for each subject. These attributes and annotations are motivated by the large variety of problems related to face recognition (Beham and Roomi, 2013). Some of them are subjective (e.g., expression) while some of them are objective (e.g., age,

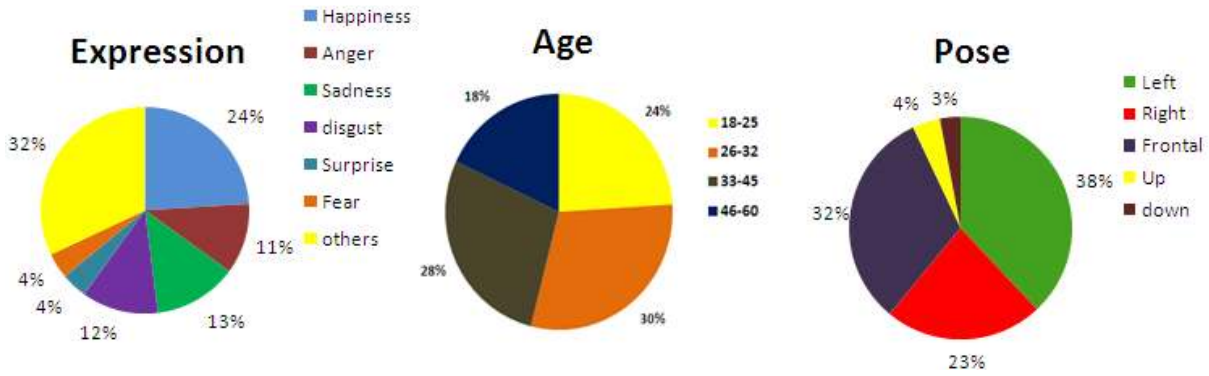

Fig. 4: Statistics of expressions, pose and age variations of INFACE and EPOCH database

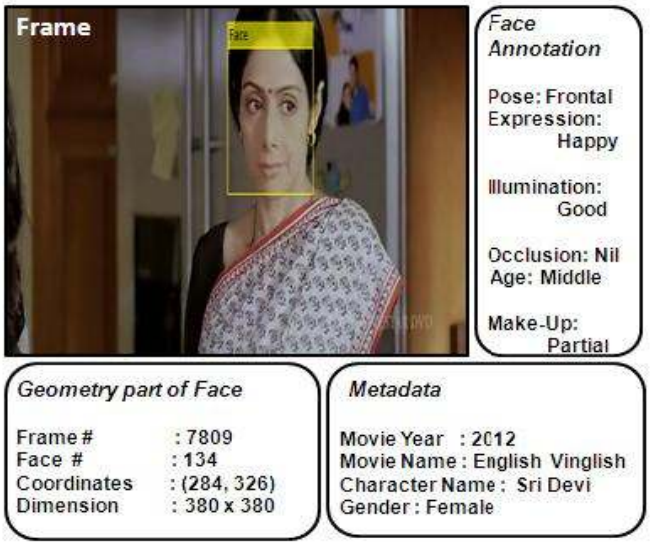

Fig. 5: Face annotation, geometry and metadata of Sridevi Kapoor (character) from English Vinglish movie

which is validated from facts about the actors and the release dates of the movies). Even though all possible set of values for certain attributes such as occlusion, pose and expression could be quite large, we tried to annotate for most commonly occurring attribute values in the movies. The Parameters that are considered for annotation in INFACE database are listed below:

- **Expressions:** Anger, Happiness, Sadness, Surprise, Fear and Disgust
- **Illumination:** Bad and Good
- **Pose:** Frontal, Left, Right, Up and Down
- **Occlusion:** Sun-glasses, Beard, Ornaments, Hair and others
- **Makeup:** Partial and Over
- **Gender:** Male and Female

Annotated information is stored as XML description. An XML specification consists of four main nested elements such as:

- Movie information consisting of movie name and year of its release
- Character information specifying actor name, gender and instance number assigned to it
- Facial variations specifying the type of facial expression, illumination, pose, occlusion, age and makeup
- Bounding box consisting of (x: y) co-ordinates of top left corner in the frame along with height and width of the cropped face

The statistics of variations in facial expressions and pose in INFACE and Age variations in EPOCH database are shown in Fig. 4.

In addition to face annotation attributes (as stated above) for each face, geometric and metadata details such as movie information, bounding box, character information are also mentioned as shown in Fig. 5. Also Fig. 6 shows XML specification of a sample face under consideration. The INFACE consists of both gray scale and RGB face images. Seven kinds of facial expressions, five kinds of pose and four categories of

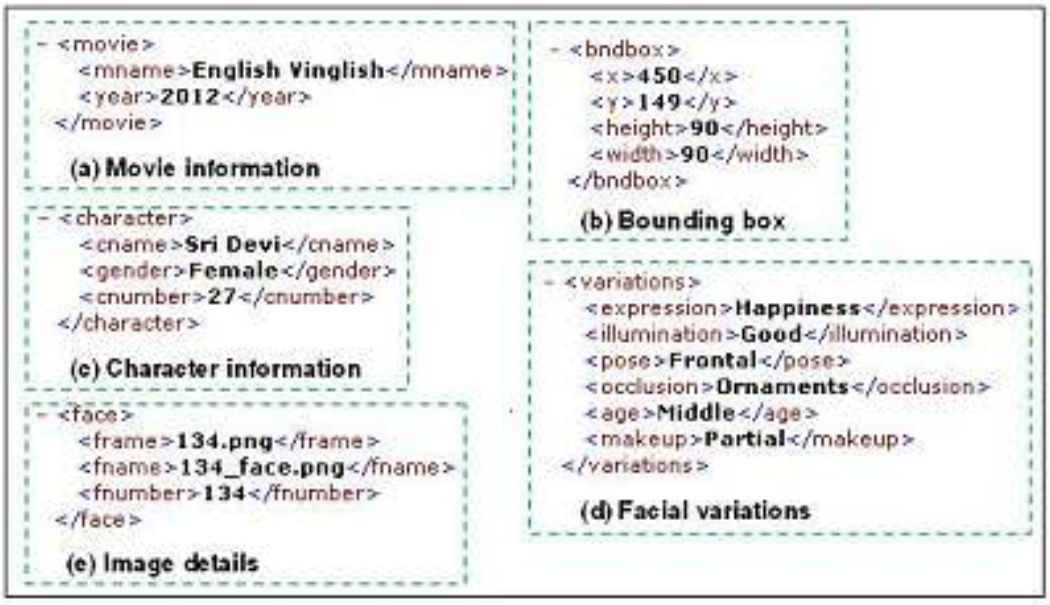

Fig. 6: An example of XML annotation label for each face

age variations are considered. From the extent of INFACE database, we can infer that the range of facial expression is highly varied. The first and the most important potential use of this database are to test the face recognition algorithms robustness in a real world scenario. There is also a potential to test various image preprocessing algorithms (enhancing image quality by filtering techniques), as some of these images are of extremely low quality and resolution. By including different pose images of subjects, we made it possible to use this database in face modeling and 3D face recognition. Other potential uses of this database include but are not restricted to: evaluation of head pose estimation algorithms, evaluation of face recognition algorithms robustness to different poses, evaluation of natural illumination normalization algorithms, indoor face recognition (in uncontrolled environment), low resolution images influence, several occlusions (beard, glasses, hair style), age invariant face recognition etc. The potential applications of this database can be person based image retrieval, character classification, etc.

RESULTS AND DISCUSSION

The primary focus of this study is to present the INFACE and EPOCH database and share it with the research community. However, we also establish some baseline results to show the complexity of our database. Sparse coding has emerged as the natural and effective representation for face images. Wright *et al*. (2009) present a Sparse Representation (SRC) in their highly cited work. Sparse representation based classifier represents each test image as a sparse linear combination of training images (dictionary). The given test image is then assigned the label of a class that gives minimum reconstruction error. SRC classifier is used to report our results on INFACE database. To prove the complexity of the proposed database, we have also compared the results with two popular databases namely Extended Yale B and AR. Features have been obtained using state of the art methods such as Principal Component Analysis (PCA) (Belhumeur *et al*., 1997), Dense Scale Invariant Feature Transform (DSIFT) (Ho and Chellappa, 2012) and Local Binary Pattern (LBP) (Pan *et al*., 2011). MATLAB Vlfeat library (Vedaldi and Fulkerson, 2008) is used to extract D-SIFT, LBP and PCA features. For evaluation, we have selected 3 classes from each database. All the face images in these datasets are already cropped, aligned, centered and resized to 80×80. During the experimentation we randomly divided training and testing images using permutation and combination such that the training set images in each class does not overlap with testing set images of its own. The recognition results are shown in Table 1. From the recognition results it is clear that, the performance of state of the art SRC is poorer due to complexity of our INFACE database compared to other standard databases. INFACE and EPOCH is designed with a careful observation in order to have a diverse range of variations for each subject. A possible way to measure the complexity of the database is by understanding the distribution of faces. Figure 7 shows the corresponding mean faces and Eigen vectors for the same subset of databases. While the average face of INFACE is more spread, it resembles the structure of face for LFW and PubFigs indicating fewer variations. This is attributed to limited pose variations constrained by the detector.

Table 1: Recognition accuracy results on INFACE database: feature extraction using PCA, LBP and D-SIFT using SRC classifier (feature dimension set as 100)

Feature extraction	AR database	Extended yale B database	INFACE database
No of samples/class	14	38	200
PCA	94.46	92.28	73.59
LBP	100	97.60	87.54
D-SIFT	100	100	92.23

Fig. 7: Top 10 eigen vectors of subset of (top row) LFW, (middle row) PubFigs and (bottom row) INFACE of 5 male actors

CONCLUSION

We have introduced two new databases called INFACE and EPOCH. They have been developed with the intention of providing an enduring benchmark for new algorithms to be developed. It contains color images that provide real world challenges to face recognition algorithms by including faces with a large variety in size, shape, orientation, expression and images that have varying lighting conditions, resolution and backgrounds. The images are taken from various Indian movies. It is available to all researchers working in the field and can be downloaded by contacting the authors. We hope this will provide another stimulus to the vibrant research area of face recognition. The applications of future work could be:

- Facial Expression Analysis
- Age and Pose invariant face recognition
- Analysis and annotating actors in movies
- Person specific feature representation

ACKNOWLEDGMENT

We would like to thank Shankar Setty, Moula Husain, Jyothi Gudavalli, Menaka, Radhesyam Vaddi, Vidyagouri Hemadri, J C Karure, Raja Raju, B Rajan, Vijay Kumar and Dr. C V Jawahar for their contribution in compiling the database and its annotations.

REFERENCES

Albert, A.M. and K. Ricanek, 2008. The MORPH database: investigating the effects of adult craniofacial aging on automated face-recognition technology. Forensic Sci. Commun., 10(2).

Beham, M.P. and S.M.M. Roomi, 2013. A review of face recognition methods. Int. J. Pattern Recogn., 27(04).

Belhumeur, P.N., J.P. Hespanha and D.J. Kriegman, 1997. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE T. Pattern Anal., 19(7): 711-720.

BERC, 2010. BERC database. Biometric Engineering Research Center, Yonsei University. Retrieved form: http://berc.yonsei.ac.krS.

FGNET, 2009. The FG-NET Aging Database. Retrieved form: /http:// www.fgnet.rsunit.com.

Georghiades, A.S., P.N. Belhumeur and D.J. Kriegman, 2001. From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE T. Pattern Anal., 23(6): 643-660.

Ho, H.T. and R. Chellappa, 2012. Automatic head pose estimation using randomly projected dense SIFT descriptors. Proceeding of 19th IEEE International Conference on image processing (ICIP, 2012), pp: 153-156.

Huang, G.B., M. Ramesh, T. Berg and E. Learned-Miller, 2007. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, Mass, USA.

Jones, M.J. and P.A. Viola, 2003. Fast Multi-view face detection. Technical Report TR2003-96, Mitsubishi Electric Research Laboratories, Cambridge, MA.

Martnez, A. and R. Benavente, 1998. The AR face database. CVC Technical Report, Vol. 24.

Minear, A.M. and D.C. Park, 2004. A lifespan database of adult facial stimuli, behavior research methods (PAL). Instrum. Comput., 36(4): 630-633.

NIST, 2003. FERET: The Color FERET Database. National Institute of Standards and Technology. Retrieved form: http://www.itl.nist.gov/iad/humanid/colorferet/home.html.

Pan, H., S.Y. Xia, L.Z. Jin and L.Z. Xia, 2011. Illumination invariant face recognition based on improved local binary pattern. Proceeding of 30th Chinese Control Conference (CCC, 2011), pp: 3268-3272.

Ramanathan, N., R. Chellappa and S. Biswas, 2009. Age progression in human faces: A survey. J. Visual Lang. Comput., 15: 3349-3361.

Setty, S., M. Husain, P. Beham, J. Gudavalli, M. Kandasamy *et al*., 2013. Indian movie face database: A benchmark for face recognition under wide variations. Proceeding of the 4th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG, 2013), pp: 1-5.

Sim, T., S. Baker and M. Bsat, 2003. The CMU pose, illumination and expression database. IEEE T. Pattern Anal., 25(12): 1615-1618.

Vedaldi, A. and B. Fulkerson, 2008. VLFeat: An Open and Portable Library of Computer Vision Algorithms. Retrieved form: http://www.vlfeat.org/.

Wright, J., Y.Y. Allen, A. Ganesh, S.S. Sastry and Y. Ma, 2009. Robust face recognition via sparse representation. IEEE T. Pattern Anal., 31(2): 210-227.