Usability of Skin Texture Biometrics for Mixed-Resolution Images

Hamed Alsufyani, Sanaul Hoque, and Farzin Deravi
School of Engineering and Digital Arts
University of Kent
Canterbury, Kent, United Kingdom
{ha324, s.hoque, f.deravi}@kent.ac.uk

***Abstract*—There is a growing demand for alternative biometric modalities that can handle various real world challenges such as recognising partially occluded individuals. Skin texture has been proposed as a potential alternative; however, such skin texture analysis can become difficult when captured images are at varying resolutions (due to different distances or devices). This paper explores the prospect of using mixed-resolution facial skin images as a source of biometric information. The four facial skin regions investigated here are the forehead, right cheek, left cheek, and chin which were selected because at least one of these are expected to be captured in real-world scenarios. The proposed framework first localises and assesses the usability of the extracted region of interest (ROI) for subsequent analysis. Local Binary Pattern (LBP) descriptors are then used for feature matching because of their reported effectiveness in extracting skin texture information. Experiments conducted using the XM2VTS database suggest that mixed resolution skin texture images can provide adequate information for biometric applications.**

Keywords— skin texture, forensic biometrics, occlusion

I. Introduction

In forensic applications, there is a growing demand for alternative biometric modalities that can handle various real world challenges such as recognising people from their partially occluded faces. Skin texture has been proposed as a potential alternative. As a standalone source of skin biometric information, a high resolution (HR) face image can be acquired at close range. However, several factors may degrade skin image resolution; for example, if the face is not close enough to the camera or the sensor is of low quality, the images will be in a lower resolution (LR). Use of such low resolution images means that system performance may be significantly degraded due to lack of details. Additionally, matching two skin images of the same individual captured from different distances/devices may also be challenging. For example, enrolled template images at high-resolution while the query image in low resolution or vice versa is likely to make the matching outcome less reliable. It is natural that with growing numbers of smartphone users and surveillance cameras, partial views of the faces are most likely be captured at different resolutions. Besides, only a small area of skin may be visible. It is therefore important to how the size or location of the skin patch may impact performance. Therefore, it is very important to be able to exploit all available information efficiently and adequately for successful person identification.

This work focuses mainly on investigating the use of skin texture features at varied resolutions to identify individuals and more importantly exploring how the proposed framework may perform best in difficult conditions (e.g. when using mixed-resolution images). The study contributes to the enhancement of skin-based biometric systems by exploring the impact of resolution in the use of facial skin images for people recognition.

II. Background and Motivations

There are clear evidences that image resolution is crucial for the performance of biometric recognition systems [1]. For example, in identifying criminals by means of surveillance camera or cell phone face recognition, challenges that may affect system performance include expression, age, illumination, pose, occlusion, and distance. Particularly, matching two images of the same subject at different resolutions (e.g. due to different distances/sensors) is of interest in forensic applications. Some work has been reported on recognising faces in high-resolution and low-resolution images. LR images captured by surveillance cameras [2, 3] are often indistinct, making individuals difficult to recognise. Using traditional methods [4, 5], HR images usually provide more detail than LR. To exploit LR images, computer vision and machine learning researchers have employed a technique known as super-resolution (SR) to generate HR images from LR [6]. There are three categories of SR: learning-based, functional-interpolation, and reconstruction-based. Earlier works [8-10] achieved face recognition by applying SR to low resolution images. Studies, therefore, addressed this problem by introducing different SR algorithms [11, 12], but to date, most such studies have focused on identification of whole face images which are sometimes unavailable. All of the above-mentioned methods explored only full-face recognition systems [7]. Also, low-resolution partial face images (e.g. captured from a distance) have not been considered.

Today, high quality facial images are easy to capture from different distances by using imaging devices such as high-resolution cameras, phones, webcams, and other inexpensive computing resources. For example, pan-tilt-zoom (PTZ) cameras can capture a high resolution facial image from up to 12 meters, with automatic tracking and close-ups [13]. Face recognition systems are also developing rapidly with increasing access to high resolution equipment [9]. All of these factors facilitated this exploration of the use of a small part of the facial image (such as facial skin) to identify individuals at a non-fixed location/distance.

HR digital cameras have also facilitated sophisticated analysis of skin texture features, and research has confirmed that skin information extracted from a small part of the facial image can provide meaningful biometric information [14, 15]. It is opportune, then, to investigate the possibility of cross matching the skin information embedded in high- and low-resolution images. The present work investigated the use of skin texture features at different resolutions to identify an individual, assessing how the proposed framework performs in difficult conditions (e.g. using low-resolution images) and

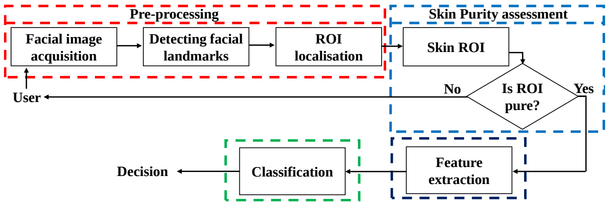


Fig. 1. The block diagram of the proposed skin-based recognition system

how image resolution affects recognition. The study contributes to the enhancement of skin-based biometric systems. This work reveals the relationship between resolution as well as the potential of facial skin images for human recognition.

III. General Approach of the Proposed System

The work presented in this paper investigates the impact of varying resolution images on skin biometric systems, as well as how such images can be successfully used for biometric recognition/applications. As outlined in Fig. 1, the proposed system is composed of the following key steps: (i) pre-processing, including facial image acquisition, facial landmark detection and ROI localisation; (ii) skin purity assessment; (iii) feature extraction, (iv) classification. The following sections provide more details of each step.

A. Facial Landmarks Detection and ROI Localisation

All facial images were rotated so that eye-centres were horizontal. Facial landmarks were then extracted using the Chehra Face Tracker [16, 17]. This software identifies 59 points (49 facial landmarks and 10 eye landmarks) on frontal face images. In this study, facial landmarks such as eye corners, nose, and mouth corners have been used to automatically determine the four skin ROIs (forehead, right and left cheek, and chin).

Geometric measures were used for the localisation of the ROI. All relevant parameters were empirically determined for the XM2VTS database [18]. As the Chehra software does not provide a facial outline, a very conservative static ROI dimension for each skin region (i.e. forehead, right cheek, left cheek, and chin) was adopted. Details of the ROI localisation can be found in [14, 15].

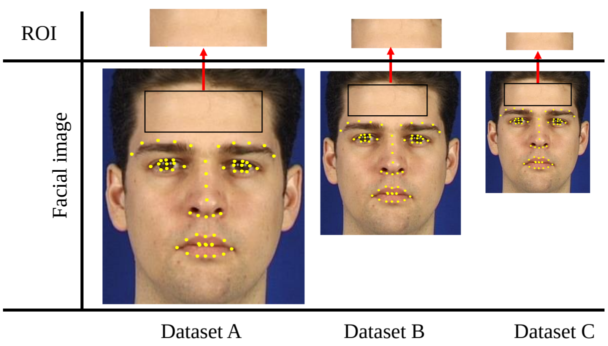


Fig. 2. Example of facial skin images and skin ROI at different scales

B. Generating Facial Images at Different Scale

To simulate the effect of the distance between subject and camera, all facial images in the database were scaled to different resolutions. Commonly used methods of interpolation include nearest neighbour, bilinear interpolation, spline interpolation [19], and bicubic interpolation [20, 21]. Although nearest neighbour and bilinear interpolation seem computationally simple, serious blurring remains a problem for these methods. It is not obvious which interpolation method best is in terms of skin scale, however bicubic interpolation method is generally considered to achieve better results [22].

To investigate the impact of low-resolution imaging on skin texture information, all facial skin regions in this work were scaled down using bicubic interpolation technique with 4x4 neighbourhoods.

For all individuals in XM2VTS database, images were resized to different scales to simulate the varying resolution images. Specifically, for the results presented here, two datasets namely Dataset B and Dataset C were created by scaling the original images (Dataset A) by factors of 0.75 and 0.50 respectively. The 0.50 was chosen as the lowest due to the restriction imposed by the LBP parameters (P, R) when the skin features extracted. It also ensures that the ROI would contain sufficient pixels for the LBP operator in both neighbourhoods (P) and radius (R). Fig. 2 shows one example of scaled skin images extracted from the forehead region.

These three datasets were then utilised for biometric processing of each ROI (forehead, right cheek, left cheek, and chin). Table I shows the size range of the original ROIs and their reduced scale versions from the three datasets. The forehead region is obviously larger than the other skin regions. Both cheek region sizes are smaller, causing feature extraction methods to limited LBP parameters during processing. For example, the cheek region becomes too small, when resized using scaling factor 0.5, for effective representation by LBP.

C. Skin Purity Analysis

The skin purity assessment technique applied here was originally reported in [14]. The scheme implemented skin colour models to ascertain the level of purity of the extracted ROI. The reason for using colour based models is because skin colours are largely invariant to partial occlusion, scaling, pose and rotation and, thus, helpful in differentiating between skin pixels from other artefacts within a facial image. The images that exceeded the skin purity threshold were divided into N non-overlapping sub-regions to extract skin features

TABLE I. ROI SIZE RANGES AFTER BICUBIC INTERPOLATION USING DIFFERENT SCALING FACTORS

ROI	Dataset A	Dataset B	Dataset C
		Scale factor	
	Original size (from–to)	*0.75-scale (from – to)*	*0.50-scale (from – to)*
	Range of ROI size (in pixels)		
Forehead	49 x 122–79 x 97	37 x 92–59 x 148	25 x 61–40 x 99
Right Cheek	24 x 24–56 x 57	18 x 18–42 x 43	12 x 12–28 x 29
Left Cheek	24 x 24–56 x 57	18 x 18–42 x 43	12 x 12–28 x 29
Chin	23 x 41–67 x 120	17 x 31–50 x 90	12 x 21–34 x 60

separately from each of those sub-regions. Finally, all sub-regional features were concatenated to form a single feature vector.

D. *Feature Extraction and Classification*

LBP histogram has been used for feature extraction as it provided more reliable skin features. In order to match LBP feature vectors provided by multiple operators with different (P, R) values and different skin image resolutions, all feature sets were normalised using the following formula:

$$f_i' = \frac{f_i}{\sum_{i=1}^{n} f_i} \quad (1)$$

where f_i is the feature component before normalising; f_i' is the normalised feature vector; and n is the dimension of the feature vector.

A k-Nearest-Neighbour (k-NN) classifier has been used for feature classification. The k value was chosen based on the performances achieved during trials.

IV. Experimental Results and Discussions

This section presents the experimental results and their interpretations. The face images were taken from the XM2VTS database for these experiments. Desired ROIs were automatically localised and segmented using the facial landmarks detected using the Chehra software. In the XM2VTS database, the software could detect landmarks in only 1,128 of the available images, marking all four images for 274 of the 295 available individuals.

Preliminary analysis found that the size of the LBP neighbourhood parameter (P) has less influence on biometric performance, but the changes in the LBP radius (R) clearly affected performance. So, the experiment reported in this paper only used 8 neighbourhoods ($P = 8$) while the radius (R) was varied to determine the best values for each image scale. All LBP feature vectors were normalised using (1).

Experiment 1: The Effect of Resolution on Recognition

The first experiment had two objectives: (i) to investigate the usability of low-resolution skin images in biometric systems and (ii) to explore the impact of low-resolution images on overall skin-based biometric performance. Both the training and the test sets images here were of the same resolution.

The LBP descriptor extracts skin features from ROI with parameters $P = 8$ and $R = n$, where n varied between 1 and 5. LBP features were extracted for all different resolutions, and feature vectors were normalised. The experimental protocol used three skin images from each person as a training set; the remaining images were used for testing, with the same R value for both training and test sets.

Figs. 3–6 show the recognition rates for the various skin regions of interests at different resolutions, for different R values. Different R values were found optimum for different ROI as well as for datasets with differing resolutions.

For the forehead skin region, 5 was the optimal value of R for the Dataset A and the Dataset B. The optimal value of R was 3 for the Dataset C indicating that for low-resolution skin images, it is perhaps preferable to extract LBP features using a small R value. Owing to the relative size of the cheek regions, LBP with R=2 delivered the best performance for

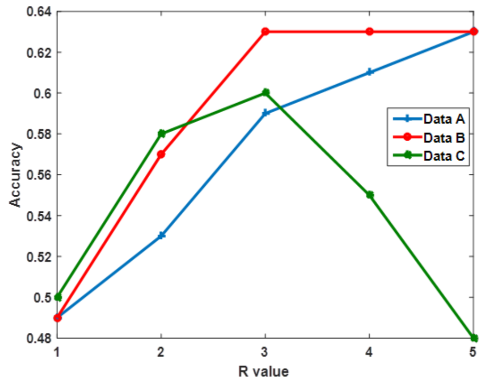


Fig. 3. Recognition rates for forehead region at different resolutions

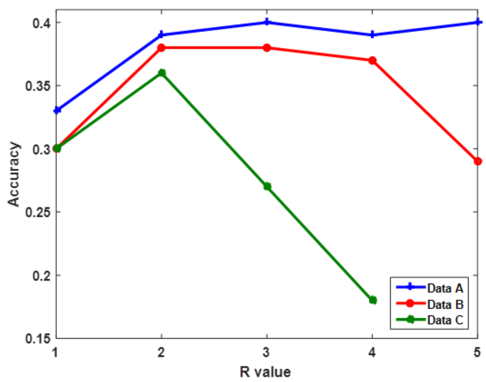


Fig. 4. Recognition rates for right cheek region at different resolutions

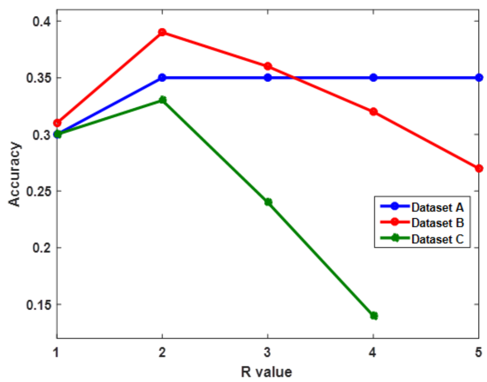


Fig. 5. Recognition rates for left cheek region at different resolutions

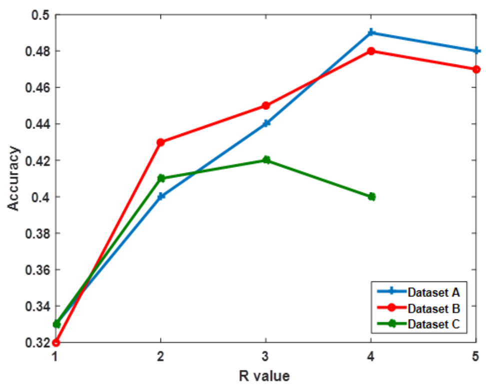


Fig. 6. Recognition rates for chin region at different resolutions

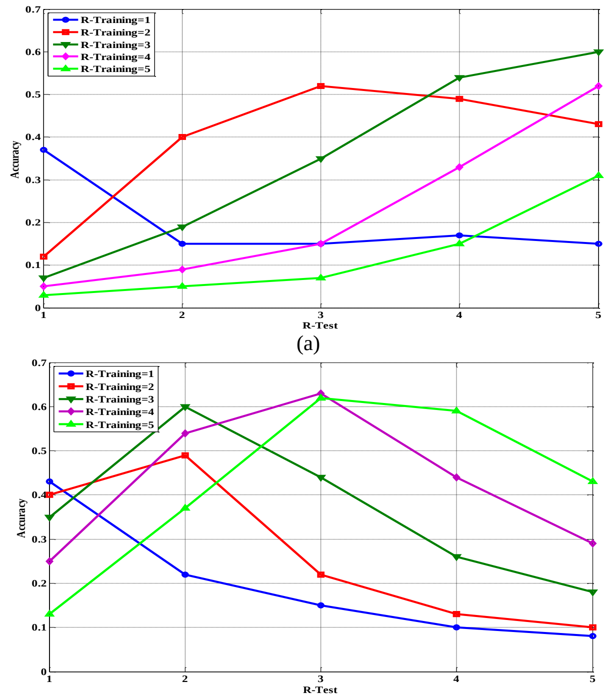
both. For the chin region, *R*=4 yielded the best performance for features extracted from Dataset A and Dataset B.

These results confirm that LBP's radius parameter *R* plays a vital role in extracting discriminating skin information for people recognition. Looking more closely, these results indicate that biometric skin information can be obtained using an appropriate feature extraction method even at mixed resolution scenarios. Planar skin surfaces such as the forehead and chin regions achieved significantly better recognition performance compared to the less-planar cheek regions. While recognition rates for the cheek regions were below 0.40, the forehead region achieved the highest recognition rate of above 0.60. This difference of almost 0.20 in recognition rates suggests the usability of the forehead skin region even at low resolution for biometric purposes.

Experiment 2: The Effect of Mismatched Resolution between the Enrolment and Verification

The objective of Experiment 2 was to investigate recognition of individuals from skin images captured at different resolutions from the enrolment set. In real-world scenarios, resolutions of the face images captured from different distances or by different devices may often differ from template samples. In this experiment, the training and test sets are at different resolutions. To optimise the value of R, different values were analysed.

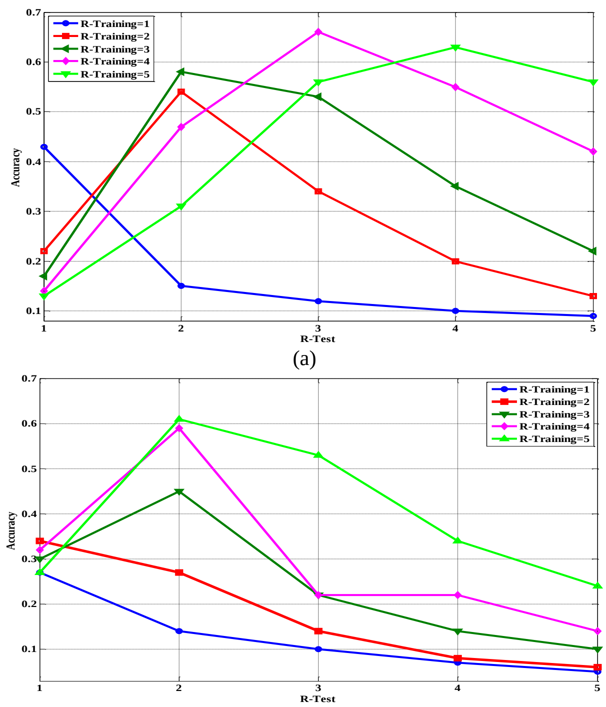
As the forehead region provided the best results for identification in Experiment 1, only this region is investigated in Experiment 2. Figs. 7 to 9 show the recognition accuracy rates achieved using different LBP parameter settings for different combinations of train-test resolutions for the forehead region. These results show that the system naturally performed better when high-resolution images (in this experiment, Dataset A or Dataset B) were used in both training and testing. The system also performed quite well when low-resolution images (Dataset C) were used for testing. However, when low-resolution images were used in enrolment, the performances deteriorated significantly. In addition, images



(a)

(b)

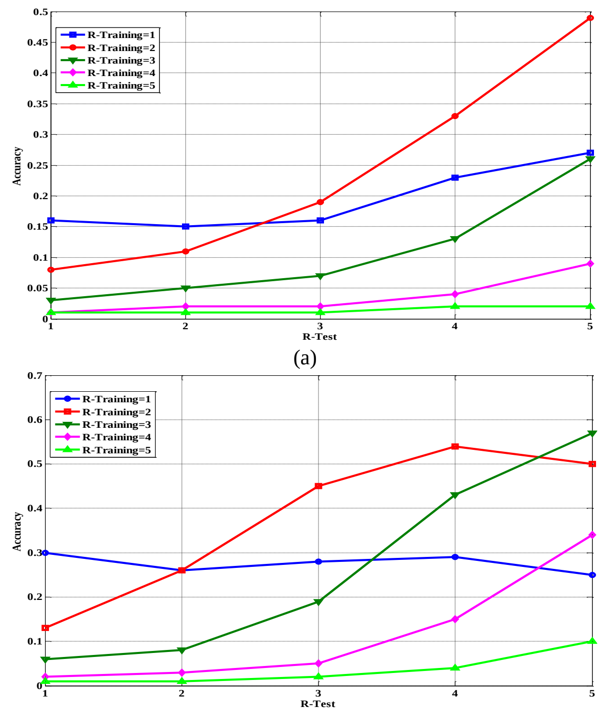
Fig.8 Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train with Dataset B and test using (a) Dataset A, and (b) Dataset C



(a)

(b)

Fig.7. Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train with Dataset A and test using (a) Dataset B and (b) Dataset C



(a)

(b)

Fig.9 Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train with Dataset C and test using (a) Dataset A and (b) Dataset B

TABLE II. BEST RECOGNITION RATES ACHIEVED WITH DIFFERENT *R* VALUES FOR TRAINING AND TEST SETS WHEN MATCHING IMAGES ARE AT DIFFERENT SCALES

ROI	**Experimental protocol**		***R* values**	**Recognition rates**	
	Training images	***Test images***	**($R_{training}$, R_{test})**		***overall***
Forehead	Dataset A	Dataset B	(4,3)	0.66	0.64
		Dataset C	(5,2)	0.61	
	Dataset B	Dataset A	(3,5)	0.60	0.62
		Dataset C	(4,3)	0.63	
	Dataset C	Dataset A	(2,5)	0.49	0.53
		Dataset B	(3,5)	0.57	

that were close in resolution (e.g. Dataset A vs Dataset B or Dataset B vs Dataset C) yielded better recognition rates.

It should also be noted that optimisation of LBP parameters was very important in securing high recognition rates. For example, high-resolution skin images favour larger values of *R* (e.g., $R = 4$ or $R = 5$) while low-resolution images favour smaller values of *R* (e.g., $R = 2$). These results also suggest that high-resolution images are preferable for enrolment for individual recognition.

Form these observations, it can be seen that extracting skin features using a single value of *R* may not capture sufficient biometric information when the training and test images differ in resolution. As the resolution of the skin regions may be highly variable, it may be advisable to incorporate adaptive sets of LBP parameters during the feature extraction to represent the skin ROI. Table II summarises the sets of *R* values that generated high recognition rates for different mixed resolution scenarios for the forehead images. It can be seen that, for forehead region, 4 and 5 are the best *R* values for Dataset A; 3 and 4 are the best for Dataset B; and 2 and 3 are the best for Dataset C, irrespective of whether the images were used for training or testing. It may therefore be advisable to extract multiple features using a subset of optimised LBP parameters from the enrolled images (appropriate for its resolution). This will ensure that optimal matching is likely even if the test resolution is uncertain. Similarly, multiple test features can also be extracted to address if the enrolment dataset resolution is variable. The rightmost column of Table II shows the overall accuracy if such measure is incorporated.

V. CONCLUSIONS

This paper explores the effects of mixed image resolution on skin biometrics. The results suggest that skin images remain usable as a source of biometric information even when images of different resolution are used during training and testing. One notable finding was that the LBP, based on a fixed value of R, is not always the best option for analysing multiresolution skin images and scale changes. It is also found that LBP features using multiple parameters can deliver better recognition performance in such mixed-resolution scenarios. Future work will explore scenarios where the capture conditions are even less constrained such as those covered in the LFW database [23] by dynamically adapting feature capture parameters.

REFERENCES

[1] J. Neves, F. Narducci, S. Barra, and H. Proença, "Biometric recognition in surveillance scenarios: a survey," *Artificial Intelligence Review*, Springer, vol. 46(4), pp. 515-541, December 2016.

[2] L. An and B. Bhanu, "Face image super-resolution using 2D CCA," *Signal Processing*, Elsevier, vol. 103, pp. 184-194, October 2014.

[3] Z. Wang, Z. Miao, Q. M. J. Wu, Y. Wan, and Z. Tang, "Low-resolution face recognition: a review," *The Visual Computer*, Springer, vol. 30(4), pp. 359-386, April 2014.

[4] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Transactions on Neural Networks*, vol. 13(6), pp. 1450-1464, November 2002.

[5] J. Ruiz-del-Solar and P. Navarrete, "Eigenspace-based face recognition: a comparative study of different approaches," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 35(3), pp. 315-325, August 2005.

[6] C.-Y. Yang, C. Ma, and M.-H. Yang, "Single-image super-resolution: A benchmark," in *Proc. of 13th European Conference on Computer Vision (ECCV 2014)*, Zurich, Switzerland, vol. LNCS 8692, Part IV, pp. 372-386, September 2014.

[7] J. Jiang, C. Chen, J. Ma, Z. Wang, Z. Wang, and R. Hu, "SRLSP: A face image super-resolution algorithm using smooth regression with local structure prior," *IEEE Transactions on Multimedia*, vol. 19(1), pp. 27-40, January 2017.

[8] G. Dedeoglu, T. Kanade, and J. August, "High-zoom video hallucination by exploiting spatio-temporal regularities," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004)*, 27 June-2 July 2004, Washington, DC, USA. Vol.2, pp. II-151-II-158.

[9] Y. M. Lui, D. Bolme, B. A. Draper, J. R. Beveridge, G. Givens, and P. J. Phillips, "A meta-analysis of face recognition covariates," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS'09)*. Washington, DC, USA, pp. 1-8, September 2009.

[10] W. W. W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Transactions on Image Processing*, vol. 21(1), pp. 327-340, January 2012.

[11] J. D. van Ouwerkerk, "Image super-resolution survey," *Image and Vision Computing*, Elsevier, vol. 24(10), pp. 1039-1052, October 2006.

[12] K. Jia and S. Gong, "Generalized face super-resolution," *IEEE Transactions on Image Processing*, vol. 17(6), pp. 873-886, June 2008.

[13] U. Park, H.-C. Choi, A. K. Jain, and S.-W. Lee, "Face tracking and recognition at a distance: A coaxial and concentric PTZ camera system," *IEEE Transactions on Information Forensics and Security*, vol. 8(10), pp. 1665-1677, October 2013.

[14] H. Alsufyani, S. Hoque, and F. Deravi, "Exploring the Potential of Facial Skin Regions for the Provision of Identity Information," in *Procceedings of the 7th International Conference on Imaging for Crime Detection and Prevention (ICDP-16)*, Madrid, Spain, pp. 1-6, November 2016.

[15] H. Alsufyani, S. Hoque, and F. Deravi, "Automated skin region quality assessment for texture-based biometrics," in *Seventh International Conference on Emerging Security Technologies (EST)*, Canterbury, UK, pp. 169-174, September 2017.

[16] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Incremental face alignment in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, Ohio, USA, pp. 1859-1866, June 2014.

[17] Intelligent Behaviour Understanding Group (iBUG), "Chehra Face Tracker (CVPR 2014)." [Online]. Available: https://ibug.doc.ic.ac.uk/resources/chehra-tracker-cvpr-2014/ (Last Accessed: 2015).

[18] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "XM2VTSDB: The extended M2VTS database," in *Proc. 2nd International Conference on Audio and Video-based Biometric Person Authentication (AVBPA '99)*, Washington, D.C., USA, pp. 72-77, March 1999.

[19] M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, vol. 13(3), pp. 277-285, March 1991.

[20] H. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26(6), pp. 508-517, December 1978.

[21] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29(6), pp. 1153-1160, December 1981.

[22] S. Battiato, G. Gallo, and F. Stanco, "A locally adaptive zooming algorithm for digital images," *Image and Vision Computing*, Elsevier, vol. 20(11), pp. 805-812, September 2002.

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