Metadata of the chapter that will be visualized in SpringerLink

Book Title	Selfie Biometrics	
Series Title		
Chapter Title	Selfies for Mobile Biometrics: Sample Quality in Unconstrained Environments	
Copyright Year	2019	
Copyright HolderName	Springer Nature Switzerland AG	
Corresponding Author	Family Name	**Lunerti**
	Particle	
	Given Name	**Chiara**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Kent
	Address	Canterbury, CT2 7NT, Kent, UK
	Email	c.lunerti@kent.ac.uk
Author	Family Name	**Guest**
	Particle	
	Given Name	**Richard**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Kent
	Address	Canterbury, CT2 7NT, Kent, UK
	Email	r.m.guest@kent.ac.uk
Author	Family Name	**Blanco-Gonzalo**
	Particle	
	Given Name	**Ramon**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Carlos III of Madrid
	Address	Leganés, Spain
	Email	rbgonzal@ing.uc3m.es
Author	Family Name	**Sanchez-Reillo**
	Particle	
	Given Name	**Raul**
	Prefix	
	Suffix	

Role	
Division	
Organization	University of Carlos III of Madrid
Address	Leganés, Spain
Email	rsreillo@ing.uc3m.es

Abstract

Taking a “selfie” using a mobile device has become a natural gesture in everyday life. This simple action has many similarities to face authentication on a smartphone: positioning the camera, adjusting the pose, choosing the right background and looking for the best lighting conditions. In the context of face authentication, most of the standardised processes and best practice for image quality is mainly focused on passport images and only recently has the attention of research moved to mobile devices. There is a lack of an agile methodology that adapts the characteristics of facial images taken on smartphone cameras in an unconstrained environment. The main objective of our study is to improve the performances of facial verification systems when implemented on smartphones. We asked 53 participants to take a minimum of 150 “selfies” suitable for biometric verification on an Android smartphone. Images were considered from constrained and unconstrained environments, where users took images both in indoor and outdoor locations, simulating real-life scenarios. We subsequently calculated the quality metrics for each image. To understand how each quality metric affected the authentication outcome, we obtained biometric scores from the comparison of each image to a range of images. Our results describe how each quality metric is affected by the environment variations and user pose using the biometric scores obtained. Our study is a contribution to improve the performance and the adaptability of face verification systems to any environmental conditions, applications and devices.

Chapter 7
Selfies for Mobile Biometrics: Sample Quality in Unconstrained Environments

Chiara Lunerti, Richard Guest, Ramon Blanco-Gonzalo and Raul Sanchez-Reillo

7.1 Introduction

Mobile devices have brought a significant change in everyday life. They are ubiquitous both for business and personal tasks including storing sensitive data and information; from saving images to a photo gallery to interacting with financial information. As such, and given the mobile nature of the devices, data has the risk of being accessed by unauthorised users. It is therefore of critical importance to secure mobile devices through appropriate and effective authorisation processes.

Personal Identification Numbers (PIN) and passwords are two techniques that have been traditionally used to protect access to a mobile device across a range of mobile device manufacturers and operating systems (OSs). In 2008, the Android OS also introduced a personalised graphical pattern system that allows the unlocking of the device by the connection of at least four dots on a 3×3 grid. However, all these security methods are vulnerable to attacks such as shoulder-surfing and latent finger traces or are easy to replicate or guess [1, 2].

Biometrics has quickly become a viable alternative to traditional methods of authentication. The use of biometric verification technologies provides many advantages as the authentication is achieved using a personal aspect that users do not need to remember and that is impossible to lose. Adoption of authentication using

C. Lunerti (✉) · R. Guest
University of Kent, Canterbury CT2 7NT, Kent, UK
e-mail: c.lunerti@kent.ac.uk

R. Guest
e-mail: r.m.guest@kent.ac.uk

R. Blanco-Gonzalo · R. Sanchez-Reillo
University of Carlos III of Madrid, Leganés, Spain
e-mail: rbgonzal@ing.uc3m.es

R. Sanchez-Reillo
e-mail: rsreillo@ing.uc3m.es

A. Rattani et al. (eds.), *Selfie Biometrics*, Advances in Computer Vision and Pattern Recognition, https://doi.org/10.1007/978-3-030-26972-2_7

face images as a security mode began in 2011 when Google introduced in Android 4.0 "Ice Cream Sandwich" a face verification system called Face Unlock. In recent years, the system has updated and improved. Now called Trusted Face, starting with Android 5.0 "Lollipop", it has been included as part of the Smart Lock system [3]. In November 2017, Apple Inc. released the iPhone X with FaceID, a verification system that works with a TrueDepth camera system. This technology comprises an infrared camera, a dot projector and a flood illuminator, with a claim to allow high face verification performances even in hostile light condition and robust against facial changes like growing hair and beard [4].

To authenticate on a mobile facial verification system, users need to take a self-portrait using the front-mounted camera of the device. Since this action corresponds to the definition of "taking a selfie", it is possible to identify the relationship between the process of selfie generation and smartphone authentication. However, we can identify substantial differences between these processes depending on the use-context. For instance, to ensure a successful authentication, the selfie should not be taken with other people, as this would add additional processing to the system for selecting the appropriate face to authenticate among the others. Also, the facial expression should be neutral, to avoid variability on the image.

Despite these differences, it is possible to surmise that the massive popularity of posting selfies on social media has helped with the acceptability of mobile face verification. The growth of the use of facial systems on mobile devices has not been without issues. According to a survey of 383 subjects conducted by De Luca et al. in 2015, a shift was observed as to the motivations to cause people to abandon the use of Face Unlock, primarily from overriding privacy concerns to social compatibility. Across the subjects, 29% declared that they stopped using Face Unlock for usability concerns (such as variable performance caused by environmental problems) and for the feeling of awkwardness in taking a selfie in front of other people for authentication [5].

The recent acceptance in the social context of taking selfies in public is playing an essential role in the acceptability of face verification on a smartphone, leading to the socially acceptable possibility of selfie authentication or selfie banking. In work presented by Cook [6] in 2017, the authors underline that an increasing number of users are checking their bank accounts using their mobile devices, and they are willing to use face verification as a biometric over other modalities such as fingerprint, as they considered it more reliable and, through liveness detection, more secure.

It is, however, necessary to understand how taking authentication images in an unconstrained environment influences the quality (and consequently the performances) of a verification system. In face verification, most implementation standards and best practices are focused on the use of facial images in specific scenarios, such as electronic IDs or passports. Best practice needs to be adapted to the additional unconstrained environment parameters that the device mobility introduces. As the user moves the device in an unconstrained manner, both posture and the background may be subject to significant change. Also, the resolution of a device camera is typically lower than those used for taking passport images, so the same quality metrics may not have the same effect in this scenario. In the context of mobile devices, it is crucial to asses a realistic scenario including the variability of unconstrained environments.

Our research aims to contribute to the improvement of the performance of facial verification systems when applied in smartphones. We have analysed how image quality changes in respect to unconstrained environments and what influence this has on the biometric match scores. We also have studied how the user and the smartphone camera introduce variability in the system.

7.2 Biometric Selfies, the Challenges

The ISO/IEC 19794-5:2011 Biometric data interchange formats—Part 5: Face image data standard [7] provides a series of measures and recommendations to consider when collecting images for facial verification. The standard includes the acquisition process, where subjects should be in a frontal position, at a fixed distance from the camera. Images taken in unconstrained environments are mainly influenced by the different postures that users present towards a camera that is considerably smaller in size compared to the Single Lens Reflex (SLR) system generally used for capturing passport images. Mobile devices can also be moved, varying the distance between the subject and the capturing device, resulting in a variation of light and posture. Some existing studies [8, 9] have aimed to improve performance across different lighting conditions and poses of subjects, although the majority focus on video surveillance recognition or passport image application. In the first case, high-quality equipment is usually adopted, and in the second scenario, there is controlled variability in pose and lighting that limits the application in real life scenarios.

One approach to enhance sample quality of a biometric system is to provide real-time feedback to subjects so that they can adjust the device or posture, or they can provide another sample. In work presented by Abaza et al. [10], the authors analysed common metrics used to assess the quality and presented an alternative face image quality measure to predict the matching performance, requesting another sample in the case where a donated image did not conform to quality requirements. The method presented by the authors was to filter low-quality images using a proposed face quality index, resulting in an improvement of the system performance from 60.67 to 69.00% when using a distribution-based algorithm (Local Binary Patterns) and from 92.33 to 94.67% when using commercial software (PittPatt).

Another approach when dealing with low-quality images is presented by Kocjan et al. [11]. Their methodology consists of determining fiducial face points that are robust to different light and posture conditions by using Toeplitz matrices. Their algorithm achieved a 90% success rate when verifying images in unconstrained environments although this only occurred for a database with less than 30 users. Future research is focusing on maintaining the success rate while increasing the database size. AQ2

There are few studies explicitly focused on mobile devices. A study on smartphone and image quality [12] collected 101 subjects' images of which 22 samples from each

462399_1_En_7_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:18/7/2019 Pages: 23 Layout: T1-Standard

person was captured from two different devices: a Samsung Galaxy S7 and an Apple iPhone 6 Plus during two sessions. The variation of the light position and pose of the user were fixed as participants were asked to take two images with a different yaw posture (head turn to the right or the left) and six more variating their posture with roll and pitch (head tilt to the right or the left and the back or the front respectively). The quality was assessed over the collected database using different schemes, and the method proposed by the authors resulted in nearly equal or better performances to the other quality assessment methodologies.

Several databases have been released to assess face verification/identification covering a series of problems and challenges that this modality needs to overcome (for example, the "Labeled Faces in the Wild" [13] database of unconstrained facial images, formed of 13,233 images from 5749 subjects taken in different light conditions and environments). However, there is a lack of a suitable unconstrained environment facial image database with samples taken from a smartphone. Available databases usually focus on a specific environment such as an office or a laboratory and with controlled movements and posture for the user.

The main contribution of our study is the analysis of selfie biometrics considered in real life scenarios where the unconstrained environment introduces variations in quality, interaction and performances. This work builds on our previous study [14] where we described the quality variations in constrained and unconstrained environments considering quality metrics conformant to the standard requirements for passport images.

7.3 Data Collection

With the aim of assessing the impact that different types of environments have on selfies for mobile verification, we carried out an analysis by undertaking our data collection. We designed a collection process lasting about 30 min repeated across three time-separated sessions where participants took selfies suitable for verification on a provided mobile device (a Google Nexus 5). Full local ethics approval was granted prior to the commencement of our data collection.

During the first session, participants were informed as to the nature of the study and demographics were recorded. Information was also recorded regarding participants' previous experience with biometric systems and biometric authentication on mobile devices. Following this process, they received an explanation on smartphone enrolment. Each participant was asked to sit on a chair at a fixed distance from the camera (2 m) in a room with only artificial light and a white background. Six pictures were taken by a supervisor using a Canon EOS 30D SLR following the specification for passport images as described in the standard ISO/IEC 19794-5. Under the same conditions, they were given the smartphone and were asked to take another five images by themselves using the front-mounted camera of the Nexus 5 and this provided data to compare the *ideal conditions* of enrolment across two different cameras.

462399_1_En_7_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:18/7/2019 Pages: 23 Layout: T1-Standard

For the remainder of Session 1, and for the following two sessions, a standard procedure was followed. Participants were required to follow a map of locations where they were to capture a minimum of 5 verification images. The map differed across each capture session. Each map contained a total of 10 locations resulting in a minimum of 150 selfies for each participant. The locations varied: indoors and outdoors, crowded and less crowded, and were representative of locations where smartphones are used in everyday life (cafés, car parks, corridors of a building, etc.).

To collect all the images, we used an Android app that was developed for this study which also helped the participants to keep the count of the number of selfies taken during the session. The only instruction that participants received was to take the selfies for verification: ideally, they were advised to present a neutral expression and a frontal pose to the camera, but they were free to move as required, assessing lighting conditions and background that, in their opinion was ideal to provide their biometrics for verification. We collected a total of 9728 images from 53 participants of which only one participant did not complete all three sessions. Gender of participants was balanced (50.5% F/49.5% M).

7.4 Data Analysis

Based on the research questions that we wished to address, we considered our analysis according to the diagram shown in Fig. 7.1. The figure shows the contributory variables that we wanted to investigate, and their relationships indicated by the arrows. These relationships can be explored across different types of environment.

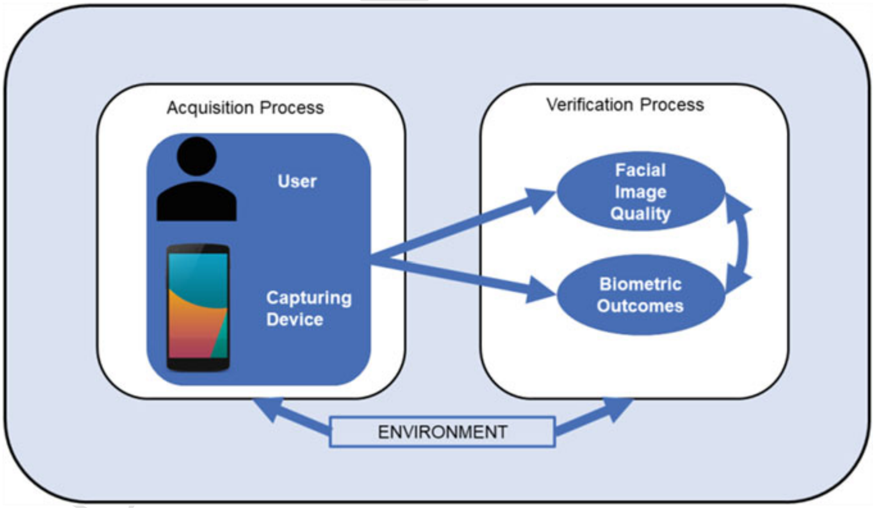

Fig. 7.1 Diagram of relationships considered in a mobile face verification system

The acquisition process in mobile scenarios is not a fixed system. Both the user and the smartphone can move freely. In the verification process, Facial Image Quality and biometric outcome scores receive influence from the user interaction and the capturing sensor. All variables are under the influence of different environments.

7.4.1 Biometric Verification

We first used two different algorithms to assess facial detection, Viola-Jones [15] as an open-source algorithm that is commonly used for this task, and the detection system with a state-of-the-art commercial verification system [16]. The commercial biometric system (CBS) was also used to assess biometric verification performance.

We considered four enrolment scenarios. The first enrolment (E1) included five images taken using the SLR camera under static conditions as previously explained. Under the same static condition, the second type of enrolment (E2) used images taken with the smartphone camera. These first two types of enrolment enabled a comparison of different types of cameras under the same ideal enrolment conditions.

The other two types of enrolment replicate real-life situations where the user is using the face authentication for the first time and need to enrol on the smartphone. We selected five random images taken indoors for the third enrolment (E3) and five random images from the images taken outdoors (E4). We decided to exclude a random combination between images taken indoors and outdoors because we assumed that it would be unlikely that someone will change his or her location from indoors to outdoors (or vice versa) in this situation.

Once all the images had been selected for the enrolment, we then considered all remaining images from that participant for verification. We used the CBS to perform the biometric verification, recording a failure to detect when the CBS could not recognise a face within an image. We calculated a biometric score (BS) as the mean of the comparisons of one verification image against all five enrolment images and a biometric outcome (BO) as either "succeeded" or "failed" depending on the majority between the five comparisons.

7.4.2 The User

The user can introduce two types of influencing factors. Some characteristics are intrinsic to the participant (such as demographic characteristics) and others that can be temporary (such as glasses, type of clothing and facial expression). From the demographics, we considered age, gender and previous experience (both with biometrics in general and in biometrics used on a mobile device) that the users declared before taking part in the experiment. We wanted to verify that there were not any differences in terms of quality and performance assessment within any demographic groups.

462399_1_En_7_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:18/7/2019 Pages: 23 Layout: T1-Standard

We used the CBS to estimate the facial expression that the user made during the image acquisition concerning the level of anger, disgust, fear, happiness, neutral, sadness and surprise. Each expression is recorded as a percentage of confidence that the user exhibits a particular expression in a captured image.

7.4.3 The Capture Device

The capture devices used during the data collection were a Canon EOS 30D SLR and a Nexus 5 smartphone camera. We provided the same model of mobile device to all the participants, to ensure that there were no differences regarding camera resolutions between the images. This decision had been made to obtain results that are device-independent and that the observations made in this study are generally valid in any case of scenarios.

We hypothesised that the images taken with the SLR would be higher quality images and that it would be easier to use for verification over a lower quality image taken from a smartphone camera. The camera specifics for both types of devices are summarised in Table 7.1.

The Exif (Exchangeable image file format) file, providing information related to the image format, was examined from each image to establish the variation capture equipment. Recent phones allow the owner to access, personalise and modify specific characteristics of the frontal camera but with the Nexus 5 that was not possible, and the focus was set to automatic.

The main camera settings that give control over quality are the aperture, ISO and shutter speed [19]. Aperture is the size of the hole within the lens that controls the lights that enters the camera body and consequentially the focus of the subject. In our experiment, it had a fixed value of 2.9 throughout all the images taken with both the smartphone camera and the SLR. Shutter speed is the length of time the camera shutter opens when taking the image. The SLR camera was fixed in position with a tripod, and the shutter speed was set at 1/60 recording images of ideally not moving subjects. When taking selfies with the smartphone, not only the subjects are moving but also the camera can take a different position, depending on how the user is holding the device. It becomes hard to differentiate these types of movements, and for this reason, the settings that we decided to consider in our analysis is the variation in ISO

Table 7.1 Camera specifics for the SLR Canon EOS 30D and the Google Nexus 5 cameras used during the data collection [17, 18]

Camera specifics	Canon EOS 30D	Google Nexus 5
Type	Digital AF/AE SLR	Selfie camera
Pixels	8.5 MP	1.3 MP
Focal length (35 mm)	35 mm	33 mm
Sensor pixel size	22.5 × 15.0 mm	1.95 μm
Autofocus features	Autofocus 9 point	Fixed focus

that measures the sensitivity of the camera sensor. The SLR had a fixed value set to 400, while the smartphone camera ISO variates between 100 and 2000.

7.4.4 Environment

We considered two types of environmental conditions. The experiment room, where there was only a fixed artificial light and participants were sitting on a chair with a white background, presented an indoor environment with ideal conditions. Images taken in this scenario were collected using both the SLR and the smartphone camera (SmrC).

All the selfies taken with the smartphone outside the experiment room have been collected in unconstrained environmental conditions. We analysed separately the images taken in the unconstrained environment when outdoors and when indoors.

7.4.5 Facial Image Quality Metrics

To assess the facial quality of the selfies acquired during the data collection, we followed the recommendations of ISO/IEC TR 29794-5 Technical Report (TR) [20]. Out of the several Facial Image Quality (FIQ) metrics considered in the TR, we selected five metrics as the ones that are commonly used in the state-of-the-art to describe quality features. Image Brightness refers to the overall lightness or darkness of the image. The Image Contrast helps to understand the difference in brightness between the user and the background of the image. The Global Contrast Factor (GCF) determines the richness of contrast in details perceived in an image. The higher the GCF, the more detailed the image. Image Blur quantifies the sharpness of an image. Finally, the Exposure quantifies the distribution of the light in an image.

Below there is a description on how to calculate each FIQ metric:

Image Brightness (B)
Image Brightness is a measure of pixels intensities of an image. As defined in the TR, the image brightness can be represented by the mean of the intensity values h_i, where $i \in \{0, \ldots, N\}$.

The mean of the histogram $\bar{h}$ can be represented by the formula:

$$\bar{h} = \frac{1}{N+1} \sum_{i=0}^{N} h_i$$

where h is the intensity value of each pixel, and N is the maximum possible intensity value.

Image Contrast (C)

Image Contrast is the difference in luminance of the object in the image. There are different ways to define Image Contrast—we chose to calculate it from the histogram of the whole image using the following formula:

$$C = \sqrt{\frac{\sum_{x=1}^{N}\sum_{y=1}^{N}(I(x,y)-\mu)^2}{MN}}$$

where $I(x, y)$ is the image face of size $M \times N$, and μ represents the mean intensity value of the image.

Global Contrast Factor (GCF)

The Global Contrast Factor (GCF) is described in the TR as the sum of the average local contrasts for different resolutions multiplied by a weighting factor. We calculated the GCF following the methodology presented by Matkovic et al. [21]. The local contrast is calculated at the finest resolution that is the original image as the average difference between neighbouring pixels. Then the local contrast is calculated for various resolutions that are obtained combining four original pixels into one super pixels, reducing the image width and height to half of the original ones. This process has been done for a number of R iterations. The global contrast is then calculated as a weighted average of local contrasts:

$$GCF = \sum_{k=1}^{R} w_k C_k$$

where C_k is the local contrast for R a number of resolutions considered, and w_k is the weighting factor. The authors defined the optimum approximation for the weighting factor over R resolution levels as:

$$w_k = \left(-0.406385\frac{k}{R} + 0.334573\right)\frac{k}{R} + 0.0877526$$

where w_k ranges from 1 to the number of resolutions (R) of the image considered.

Image Blur (Blur)

To calculate the blur effect, we studied the work presented by Crete et al. [22]. Their methodology allows the determination of a no-reference perceptual blurriness of an image by selecting the maximum blur among the vertical direction $blur_{ver}$, and the one among the horizontal one $blur_{hor}$.

$$Blur = Max(blur_{ver}, blur_{hor})$$

The metric range is between 0 and 1, where 0 is the best and 1 is the worst quality.

Exposure (E)

Exposure can be characterised by the degree of distribution of the image pixels over the grayscale or over the range of values in each colour channel. As defined in the TR, exposure can be calculated as a statistical measure of the pixel intensity distribution, such as entropy [23].

$$E = -\sum_{i=1}^{N} p_i \log_2 p_i$$

where p_i is the histogram of the intensity level for the N possible intensity levels.

7.5 Results

As a pre-processing stage, we removed the images that were taken by mistake (for example, that did not include a facial image, or contained other people), obtaining a final database of 9420 selfie images. In this paragraph, we illustrate the results obtained according to the different elements considered for image quality, biometric outcomes and user expressions.

7.5.1 *Image Quality*

Our initial investigation was to understand the variations regarding the quality of facial images. We wanted to assess how each metric varies depending on the many factors that affect the system, including different types of environments.

 From Table 7.2 we can observe that the original means have around the same values as the median, so we can assume that extreme scores do not influence the mean. A further analysis assessing the 5% trimmed means confirmed that there were no substantial outliers in the distribution that affect the mean values. From the skewness and kurtosis analysis we can ascertain that all the variables are normally distributed, as their values are between −1.96 and 1.96, except Exposure (E).

We studied the quality metrics under different conditions. Since each FIQ metric has a different range of values, we analysed them separately to understand their relationship with the user and the type of environmental conditions. In Fig. 7.2 we can see the variations of Image Brightness (B) across the 53 participants. This feature could be used to distinguish the images that have been taken in ideal conditions from the ones taken in the unconstrained environment. The threshold that is presented in the graph, as well as in the following figures that describe each quality metrics, represents an example of an empirically selected threshold (120) that can be used to distinguish between images taken in a constrained or unconstrained environment. A

Table 7.2 descriptive statistics of each FIQ metrics for the whole database (9420 selfies)

FIQ metrics	Min	Max	Mean	Median	Std. dev.	Skewness		Kurtosis	
						Stat.	Std. err	Stat.	Std. err.
B	21.19	210.28	108.81	107.98	15.56	0.052	0.025	1.822	0.05
C	6.52	13.79	10.39	10.29	0.822	0.552	0.025	0.624	0.05
GCF	1.05	9.60	5.19	5.19	1.32	−0.256	0.025	0.361	0.05
Blur	0.18	0.49	0.30	0.29	0.042	0.496	0.025	−0.071	0.05
E	5.01	7.98	7.53	7.6	0.29	−1.65	0.025	4.21	0.05

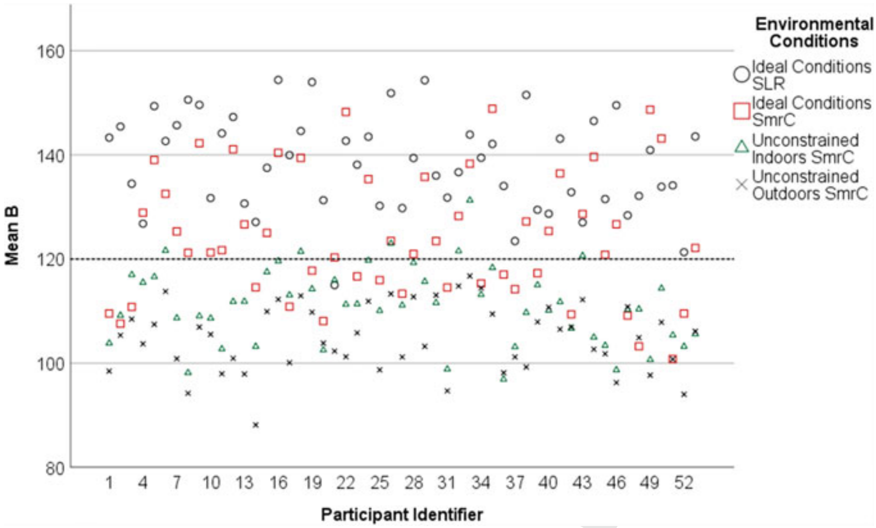

Fig. 7.2 Mean values of Image Brightness across 53 participants

further study needs to be carried out to determine the optimal thresholds that could be generally valid for any type of camera sensors.

The images that have been taken with the SLR in static condition have quality values different from those taken with a smartphone camera in unconstrained environments, indicated separately for indoors and outdoors location and the distinction between static conditions when using the smartphone is less evident. For SLR images, B ranges between 120 and 160 while for images taken indoors and outdoors in the unconstrained environments the range is from 90 to 120. When investigating brightness considering additional influencing factors, we observed that the values appear to be stable across all the three sessions and there are no significant differences between gender and age. Similarly, people that had previous experience with (mobile) biometrics did not result in different images concerning brightness compared to those who had not experienced biometric systems.

From Fig. 7.3 we can see the variation in Image Contrast (C) across all the participants. In this case, SLR images taken in ideal conditions vary across the users with values from around 11–13, while in unconstrained scenarios the images presented values with variation from 9.5 to 11. C provides a clearer division compared to B between ideal conditions and unconstrained environment. No significant differences were identified across demographics.

Contrary to the previous two FIQ metrics, GCF calculated on SLR images, as shown in Fig. 7.4, appear centred between a small range (from 1 to 3) compared to the values of all the images taken by the smartphone.

All the images captured using the smartphone range from 3 to 6.5, including those under ideal conditions, making impossible to distinguish them from the unconstrained environment. GCF resulted in the only quality metric considered that is influ-

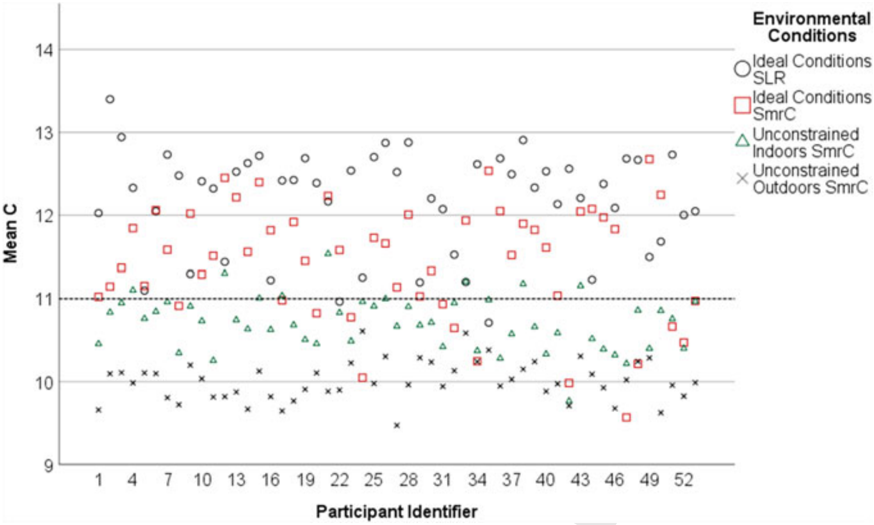

Fig. 7.3 Mean values of Image Contrast across 53 participants

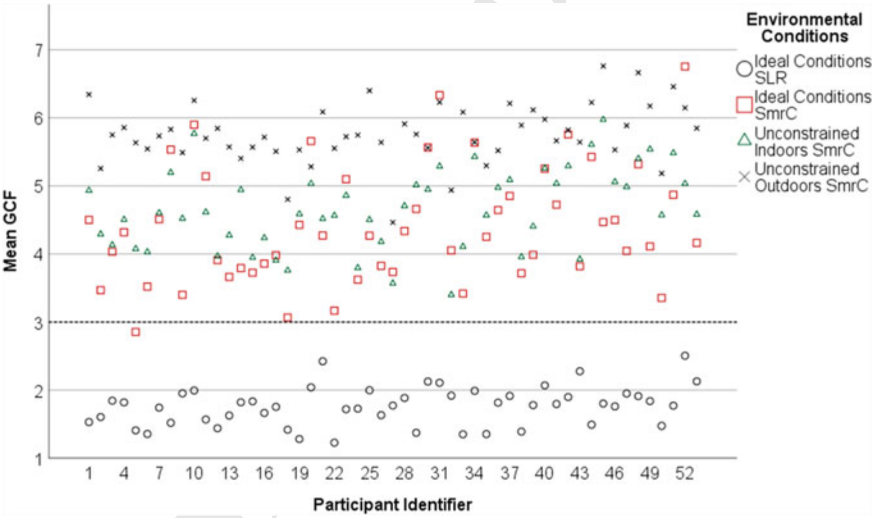

Fig. 7.4 Mean values of GCF across 53 participants

enced by the demographic. There is a negative correlation with age (r = −0.123, n = 9728, $p < 0.001$). It looks like younger participants tented to take images with higher GCF, hence more high defined images. This could be of interest for future analysis.

Like GCF, Image Blur (Fig. 7.5) also presented a distinct range of values for images taken with the SLR compared to when using the smartphone camera under the same ideal conditions. Across the collected facial images there were not many cases of an extreme blur—all the participant reported blurriness less than 0.36. Ideal conditions with the SLR can be detected from having a range of values less than 0.26, while all the images taken with the mobile device range between 0.26 to 0.36. Even though it could be unclear to form a distinction between images taken in ideal conditions with a smartphone and those taken in the unconstrained environments, we can still notice a distinction between images taken when indoors (from 0.31 to 0.36) and outdoors (0.26–0.31). There are no differences regarding sessions, demographics and previous experience.

Exposure values (Fig. 7.6) for SLR images are between the ranges of 6.65–7.35, whereas we can put a threshold to differentiate them from smartphone images taken indoors and outdoors that range from 7.35 to 7.80, we cannot make a distinction with the images taken in ideal conditions with the smartphone. There are no significant differences between sessions, gender and age.

We also inspected the variation of ISO when the images were taken in different environmental conditions in an attempt to analyse the correlation between the camera specifics and the levels of FIQ metrics. ISO distribution does not appear normally distributed, but from the analysis of the scatter plots, we observed a linear correlation that we investigated through a non-parametric Spearman correlation. There were significant results for each of the FIQ metrics, but there was a particularly strong

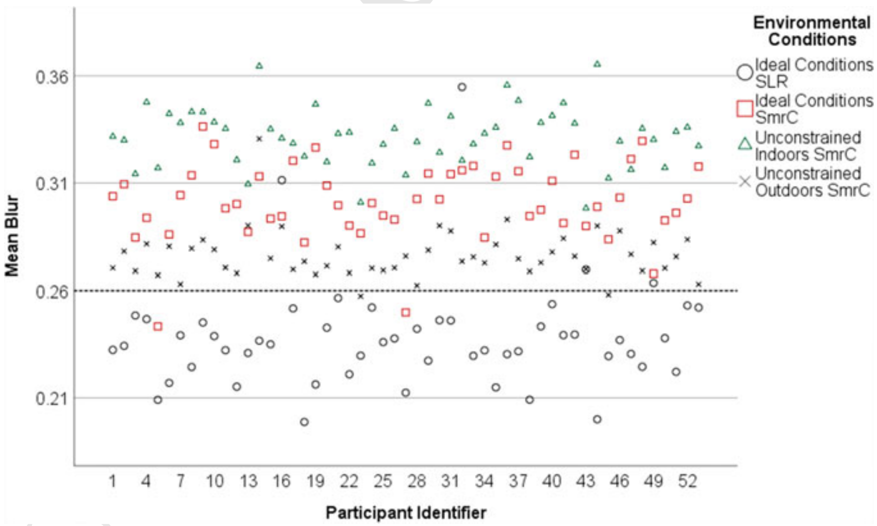

Fig. 7.5 Mean values of Image Blur across 53 participants

462399_1_En_7_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:18/7/2019 Pages: 23 Layout: T1-Standard

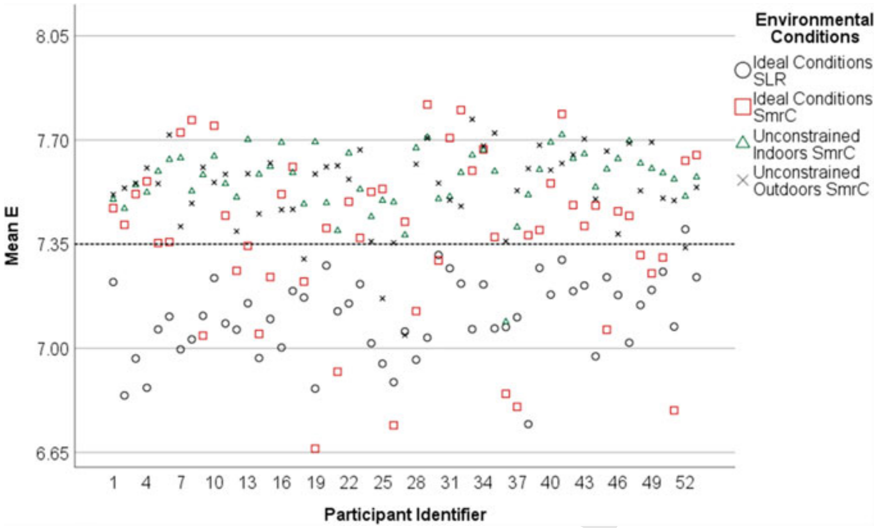

Fig. 7.6 Mean values of Exposure across 53 participants

positive correlation for Blur (r = 0.528, n = 9420, $p < 0.001$) and C (r = 0.451, n = 9420, $p < 0.001$). ISO values have a negative correlation with GCF for r = −0.438, n = 9420, $p < 0.001$. The correlation for B and E is less strong, with correspondently positive values for r = 2.28 and negative for r = −0.072 (n = 9420, $p < 0.001$).

Acknowledging the correlation between each quality metric and ISO specification, we can determine the required FIQ levels that we want to achieve and fix the ISO value on the capturing sensor. Alternatively, it may be possible to predict outcome in quality from the ISO value and be able to provide feedback in real-time or request a new image from the user to ensure that the selfie will appear with the required quality for verification.

7.5.2 Biometric Results

To perform biometric verification, we first detect the facial area of each image in our dataset. A facial area was detected within all the images taken in ideal conditions when using the SLR. Table 7.3 shows the failure to detect (FTD) using the Viola-Jones algorithm and the CBS. Overall, the number of faces detected across the entire database is above 90%. In a controlled environment, CBS was not able to detect three faces, using Viola-Jones, only one facial image was not detected. A higher percentage of FTD is recorded when images were taken outdoors (7.5% for CBS and 5.7% for Viola-Jones).

We analysed the outcomes of the biometric system depending on the type of environment. We aimed to understand how different type of environmental conditions

Table 7.3 Frequency and percentage of FTD recorded by the two algorithms

Environmental conditions			Viola-Jones		CBS	
			Frequency	Percent	Frequency	Percent
Ideal conditions	Valid	FTD	1	0.4	3	1.1
		Detected	264	99.6	262	98.9
		Total	265	100.0	265	100.0
Unconstrained indoors	Valid	FTD	135	3.9	194	5.5
		Detected	3364	96.1	3305	94.5
		Total	3499	100.0	3499	100.0
Unconstrained outdoors	Valid	FTD	306	5.7	400	7.5
		Detected	5032	94.3	4938	92.5
		Total	5338	100.0	5338	100.0

influence the biometric outcome and if there is a relationship between quality and biometric scores. A relationship can be used to regulate a biometric threshold to adapt it to the different conditions and to ensure high performances in any unconstrained environments.

Table 7.4 shows the different percentages of verification success and failure for the different environments.

A higher percentage of users that have been mistakenly rejected by the system is recorded when the enrolment has been performed using the SLR images in ideal conditions (E1), particularly when the verification takes place in an unconstrained environment, where returned results of 8.2% indoors and 11.3% outdoors. Despite having a better resolution, verification comparisons between images taken from an SLR and a smartphone yield poorer results, as already observed in our previous study [14]. This outcome could result from the application of the chosen matching algorithm to two different types of camera sensors, and it highlight the importance of using an accurate cross-sensor matching in the particular scenario between static SLR images and mobile camera images. Future research should focus on addressing

Table 7.4 Percentages of succeeded and failed verification across different environmental conditions when using a smartphone

Environmental conditions	Verification Dataset	Outcome	E1	E2	E3	E4
Ideal conditions	N = 210	Succeeded	96.7	100	99.5	99
		Failed	3.3	0	0.5	1
Unconstrained Indoors	N = 3040	Succeeded	91.8	97.4	98.9	98.1
		Failed	8.2	2.6	1.1	1.9
Unconstrained Outdoors	N = 4683	Succeeded	88.7	96.1	97.7	99.2
		Failed	11.3	3.9	2.3	0.8

this issue analysing images collected using different camera sensors to study the effects that this can have on biometric performances.

Enrolment performed with a smartphone in ideal conditions (E2) obtained the perfect acceptance rate for images taken under the same conditions, as expected, but it also recorded a favourable success rate for both the type of unconstrained environments, with 97.4% for verification performed when indoors and 96.1% when outdoors.

When the enrolment has occurred within an unconstrained environment (E3 and E4), it can be seen that a system is more resilient to the different types of verification environments, meaning that it would be better to enrol ideally under conditions that are adverse in terms of light and background so that we can ensure higher performances across a broad range of environments.

To perform a correlation between biometric scores and quality metrics, we need to check whether the scores are also normally distributed. Table 7.5 shows the descriptive statistics for the biometric scores recorded during the verification of images against the four types of enrolments. Checking the skewness and kurtosis values, we can say that not all the biometric scores form a normal distribution with only a few exceptions. In the table are also reported the minimum and maximum biometric scores recorded in the different environments (and their means and standard deviations).

We performed a non-parametric (Spearman) correlation shown in Table 7.6. The correlation has been performed for all the verification images ($n = 7923$) taken with the smartphone in both constrained and unconstrained environment. We investigated the correlation between the quality metrics recorded for those images and their biometric scores recorded when comparing them against the four types of enrolment.

From Table 7.6 we can observe some significant correlations, but not particularly strong overall (all values of the correlation coefficient, r, are smaller than 0.29). Image Blur has a strong negative correlation with the fourth type of enrolment E4 ($r = -0.288$, $n = 7923$, $p < 0.001$). In a scenario where the enrolment is performed in an unconstrained outdoor environment, the verification images appear to be more sensitive to the blurriness of the image. The correlation indicates that a reduction of blurriness of the image corresponds to a higher biometric score during the verification. Exposure presented a weak correlation that is negative for all the type of enrolments. The other quality metrics tend to have overall a positive correlation with the first three types of enrolment (captured indoors), and a negative correlation for the fourth type of enrolment (captured outdoors).

GCF has the opposite behaviour, having negative correlations with the first three types of enrolment, and a positive correlation with the E4. This can mean that despite having higher values of GCF, hence an image richer in details, in the first three types of enrolment the performances are lower. An explanation for this could be the influence that the GCF receives from local contrast in different areas of the image. For instance, a facial image can have a lower contrast in one side of the image compared to the other one, and this cannot be recorded using the Image Contrast. This difference in contrast on the same image can influence the performances in the first three types of

Table 7.5 Descriptive statistics for the biometric scores recorded in different environments

Environmental conditions		Min	Max	Mean	Std. dev.	Skewness		Kurtosis	
						Stat.	Std. err.	Stat.	Std. err.
Ideal conditions	E1	14.2	219.2	128.29	44.94	−0.188	0.168	−0.614	0.334
	E2	593.8	1193.2	916.1	147.83	0.073	0.168	−1.106	0.334
	E3	47.4	281.8	95.88	31.46	2.544	0.168	11.936	0.334
	E4	43.6	175.2	92.81	29.54	0.798	0.168	−0.003	0.334
Unconstrained indoors	E1	8.4	205.8	80.3	26.51	0.644	0.044	1.023	0.089
	E2	26.0	738.0	95.40	41.10	3.721	0.044	33.865	0.089
	E3	23.8	420.6	115.09	42.29	1.859	.044	5.714	0.089
	E4	30.2	267.0	100.53	31.52	0.890	0.044	1.523	0.089
Unconstrained outdoors	E1	4.2	239.2	76.86	27.73	1.017	0.036	1.870	0.072
	E2	4.4	262.4	92.27	37.02	1.246	0.036	1.857	0.072
	E3	7.4	198.0	99.26	30.57	0.447	0.036	−0.232	0.072
	E4	7.6	687.6	140.63	64.75	1.917	0.036	7.675	0.072

Table 7.6 Correlation between biometric scores and FIQ metrics for n = 7923

		BS_E1	BS_E2	BS_E3	BS_E4
B	Spearman's rho	0.028*	0.076**	0.041**	−0.130**
	Sig. (2-tailed)	0.014	0.000	0.000	0.000
C	Spearman's rho	0.053**	0.057**	0.047**	−0.222**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000
GCF	Spearman's rho	−0.096**	−0.095**	−0.117**	0.202**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000
Blur	Spearman's rho	0.049**	0.042**	0.105**	−0.288**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000
E	Spearman's rho	−0.059**	−0.064**	−0.001	−0.027*
	Sig. (2-tailed)	0.000	0.000	0.896	0.016

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

enrolment as it has been recorded to occur more frequently when the images were taken in indoor locations.

7.5.3 User's Facial Expressions

For most of the images taken with the SLR and the smartphone camera where it has been possible to detect a face (n = 7888), the CBS provided a level of confidence that the user was displaying a series of facial expression. In our study, we wanted to inspect if there is a correlation between the user's facial expressions and the quality level recorded, as well as the outcome from the biometric system, considering the variation that the different type of environmental conditions adds. In Fig. 7.7, we can see the mean of a facial expression's confidence for each environmental condition, indicating the frequency with which each specific expression occurred in different scenarios.

Users were only instructed to take selfies during the data collection that could be used for biometric authentication. The ideal posture would be frontal and with a neutral expression. So as expected, the facial expression that occurs the most is the neutral expression with a mean value above 40% across all scenarios. For images taken with the SLR under ideal conditions, a neutral expression has a confidence level of more than 60%. Another expression with a mean value of more than the 40% is '*surprise*' which notably occurred when using the smartphone camera. It was reported by the participants that in situations of inclement weather when outdoors, particularly with rain and strong wind, it had been harder for them take the selfies for face authentication that conformed to the requirements asked from them and

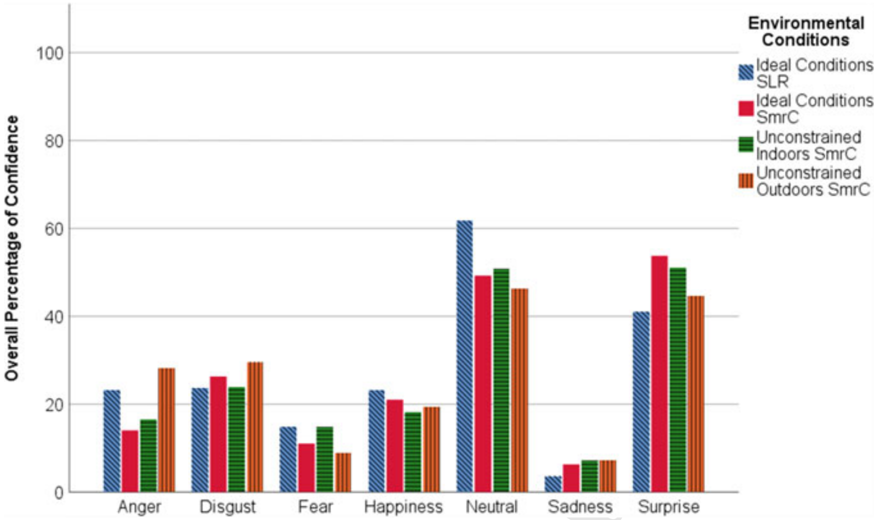

Fig. 7.7 Mean of confidence values for facial expressions

this may explain why the level of disgust and anger is higher for images taken in unconstrained outdoor environment.

Facial expressions do not conform to the normality assumption for a parametric correlation, so a Spearman correlation has been used to assess the relation that different facial expressions have on both quality and biometric performances. We did not find any particularly strong correlations between quality metrics and facial expressions (the correlation coefficient was smaller than 0.18), but we did however observe a correlation with the biometric outcomes. We considered the correlation with all the verification images where it could be possible to estimate facial expressions (n = 7678) and their biometric scores for each of the enrolment type. We noticed a strong positive correlation for neutral expression in each enrolment scenario: under ideal conditions for images taken with the SLR (r = 0.324, n = 7678, $p < 0.001$) and the smartphone (r = 0.318, n = 7678, $p < 0.001$) and for enrolment that was performed in unconstrained environments indoors (r = 0.382, n = 7678, p < 0.001) and outdoors (r = 0.295, n = 7678, $p < 0.001$). Among the other facial expressions estimated, we also observed that an expression of disgust has a strong negative correlation with ideal conditions of enrolment performed with SLR (r = −0.314, n = 7678, $p < 0.001$) and the smartphone camera (r = −0.211, n = 7678, p < 0.001). The correlation was also negative for confidence estimation of disgust presented in the images that recorded biometric scores when compared with unconstrained enrolment scenarios for smartphone images taken indoors (r = −0.232, n = 7678, $p < 0.001$) and outdoors (r = −0.141, n = 7678, $p < 0.001$).

462399_1_En_7_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:18/7/2019 Pages: 23 Layout: T1-Standard

7.6 Conclusions and Future Work

Our study aims to contribute to improving the adaptability and the performance of mobile facial verification systems by analysing how an unconstrained environment affects quality and biometric verification score. Our experimental results describe the variations of FIQ metrics and biometric outcomes recorded under different conditions and provide recommendations for the application of selfies biometrics in real life scenarios.

From the analysis of five different image quality metrics selected from the ISO/IEC Technical Report for Image Quality applied for face verification, we found that Image Brightness and Contrast could be employed to select whether an image has been taken in a constrained or unconstrained environment. Global Contrast Factor, Image Blur and Exposure were not showing different values for ideal and unconstrained conditions as clearly as the other metrics. However, by observing the local contrast and the level of blurriness, it could be possible to observe a difference between images taken in the unconstrained environments when indoors from when outdoors. These interesting results are encouraging and lead to further investigation to assess if there are significant differences between the FIQ metrics values across each type of environments. To have an overall and realistic perspective, future research will focus on analysing results collecting images using a range of different model of devices to ensure that these overall observations can be applied in context with any possible camera model. A further experiment will also be performed to explore deblurring techniques that can improve the biometric performances on those images that presented lower quality characteristics.

Our results also suggest that it is possible to consider camera specification to regulate the quality requirement for facial images when taken on a smartphone. From our study, our recommendations will be considering fixing a value for the ISO that can result in the FIQ desired, or to inspect the variation of ISO values to regulate the thresholds of acceptance of images before verification and request an additional presentation in case of noncompliance of the requirements for quality.

Studying the biometric scores, we can confirm that enrolment under unconstrained conditions ensures the system to be more robust against the variations of the environment regarding verification performances. We reported a linear correlation between quality and biometric scores, although not particularly strong.

The type of the environment is one of the factors that influence users' facial expressions. While there was not a significantly strong correlation between different facial expressions and the quality metrics, we reported positive and negative correlations depending on the type of expressions that affect the biometric outcomes. Future research can use this information to adapt biometric systems depending on the estimation of facial expressions detected in both the enrolment and verification scenarios considering the environment in which the interaction is taking place. The biometric system could send adapted feedbacks when the estimation of the location is possible to remind the user to maintain a neutral expression during the verification process.

References

1. Bonneau J (2012) The science of guessing: analyzing an anonymized corpus of 70 million passwords. In: 2012 IEEE symposium on security and privacy (SP), pp 538–552
2. Uellenbeck S, Dürmuth M, Wolf C, Holz T (2013) Quantifying the security of graphical passwords: the case of android unlock patterns. In: Proceedings of the 2013 ACM SIGSAC conference on computer & communications security, pp 161–172
3. Android (2018) Smart lock [Online]. Available: https://get.google.com/smartlock/
4. Apple Inc. (2018) About face ID advanced technology [Online]. Available: https://support.apple.com/en-gb/HT208108
5. De Luca, A, Hang A, Von Zezschwitz E, Hussmann H (2015) I feel like I'm taking selfies all day!: towards understanding biometric authentication on smartphones. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp 1411–1414
6. Cook S (2017) Selfie banking: is it a reality? Biomet Technol Today 3:9–11
7. ISO/IEC 19794-5:2011—Information technology—biometric data interchange formats—-part 5: face image data [Online]. Available: http://www.iso.org/iso/catalogue_detail.htm?csnumber=50867
8. Sang J, Lei Z, Li SZ (2009) Face image quality evaluation for ISO/IEC standards 19794-5 and 29794-5. In: International conference on biometrics, no. Springer, Berlin, pp 229–238
9. Haghighat M, Abdel-Mottaleb M, Alhalabi W (2016) Fully automatic face normalization and single sample face recognition in unconstrained environments. Expert Syst Appl 47:23–34
10. Abaza A, Harrison MA, Bourlai T, Ross A (2014) Design and evaluation of photometric image quality measures for effective face recognition. IET Biomet 3(4):314–324
11. Kocjan P, Saeed K (2012) Face recognition in unconstrained environment. In: Biometrics and Kansei engineering, no. Springer, New York, pp 21–42
12. Wasnik P, Raja KB, Ramachandra R, Busch C (2017) Assessing face image quality for smartphone based face recognition system. In: 2017 5th international workshop on biometrics and forensics (IWBF), no. IEEE, pp 1–6
13. Huang GB, Mattar M, Berg T, Learned-Miller E (2008) Labeled faces in the wild: a database for studying face recognition in unconstrained environments. In: Workshop on faces in 'Real-Life' images: detection, alignment, and recognition
14. Lunerti C, Guest RM, Blanco-Gonzalo R, Sanchez-Reillo R, Baker J (2017) Environmental effects on face recognition in smartphones. In: 2017 international Carnahan conference on security technology (ICCST), no. IEEE, pp 1–6
15. Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput 57:137–154
16. NEUROtechnology VeriLook 10.0 Standard SDK [Online]. Available: http://www.neurotechnology.com/verilook.html
17. Canon, EOS 30D specifications [Online]. Available: http://web.canon.jp/imaging/eos30d/spec/index.html
18. Ubergizmo, Google Nexus 5 specifications [Online]. Available: https://www.ubergizmo.com/products/lang/en_us/devices/nexus-5/
19. Hawkins M (2017) The exposure triangle: aperture, shutter speed and ISO explained, 12 July 2017 [Online]. Available: https://www.techradar.com/uk/how-to/photography-video-capture/cameras/the-exposure-triangle-aperture-shutter-speed-and-iso-explained-1320830
20. ISO/IEC TR 29794-5:2010—Information technology—biometric sample quality—part 5: face image data [Online]. Available: https://www.iso.org/standard/50912.html
21. Matkovic K, Neumann L, Neumann A, Psik T, Purgathofer W (2005) Global contrast factor-a new approach to image contrast. Comput Aesthetic pp 159–168
22. Crete F, Dolmiere T, Ladret P, Nicolas M (2007) The blur effect: perception and estimation with a new no-reference perceptual blur metric. In: Human vision and electronic imaging XII, vol 6492, no. International Society for Optics and Photonics, p 64920I
23. Gonzalez RC, Eddins SL, Woods RE (2004) Representation and description. In: Digital image processing using MATLAB. Prentice Hall, Upper Saddle River, pp 465–466 (Chapter 11)

Author Queries

Chapter 7

Query Refs.	Details Required	Author's response
AQ1	As chapter-wise Abstract is mandatory, please provide the abstract.	
AQ2	The reference citation author name 'Kocjan et al. [11]' does not match the list. Please check and change accordingly.	
AQ3	The comma has been changed to a decimal point in 'Tables 2–3, 5 and 6'. Please check and correct if necessary.	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	• • • under matter to remain	
Insert in text the matter indicated in the margin		New matter followed by or
Delete	through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	,
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	