TECHNICAL REPORT

ISO/IEC TR 29794-5

First edition
2010-04-01

Information technology — Biometric sample quality —

Part 5: Face image data

Technologies de l'information — Qualité d'échantillon biométrique —

Partie 5: Données d'image de face

Reference number
ISO/IEC TR 29794-5:2010(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword iv

Introduction v

1 Scope 1

2 Normative references 1

3 Terms and definitions 1

4 Abbreviated terms 1

5 Approaches to Face Image Quality 2

6 Categorization of Facial Quality 2

7 Facial Image Quality Analysis 4
7.1 Dynamic Subject Characteristics 5
7.1.1 Subject's Behaviour 5
7.1.2 Analysis Based on Statistical Differences of the Left and Right Half of the Face 5
7.2 Static Characteristics of the Acquisition Process 7
7.2.1 Image Resolution and Size 8
7.2.2 Noise 8
7.3 Characteristics of Image Acquisition 8
7.3.1 Image Properties 8
7.3.2 Image Appearance 9
7.3.3 Illumination Intensity 9
7.3.4 Image Brightness 9
7.3.5 Image Contrast 10
7.3.6 Exposure 11
7.3.7 Focus, Blur and Sharpness 11
7.3.8 Colour 12
7.3.9 Subject-Camera distance 12

Bibliography 13

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard (“state of the art”, for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC TR 29794-5, which is a Technical Report of type 2, was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 37, *Biometrics*.

ISO/IEC 29794 consists of the following parts, under the general title *Information technology — Biometric sample quality*:

— *Part 1: Framework*

— *Part 4: Finger image data* [Technical Report]

— *Part 5: Face image data* [Technical Report]

Introduction

The purpose of this part of ISO/IEC 29794 is to define and specify methodologies for computation of objective, quantitative quality scores for facial images. Furthermore, the purpose, intent, and interpretation of face quality scores are defined.

ISO/IEC 19794-5, *Information technology — Biometric data interchange formats — Part 5: Face image data*, already gives some specifications that are related to

— scene constraints of the facial images,

— photographic properties of the facial images, and

— digital image attributes of the facial images.

Within this part of ISO/IEC 29794, a sample of a classification scheme of facial quality is exemplified and approaches for the determination of certain aspects of quality are introduced.

Information technology — Biometric sample quality —

Part 5:
Face image data

1 Scope

For aspects of quality specific to facial images, this part of ISO/IEC 29794:

— specifies terms and definitions that are useful in the specification, use and testing of face image quality metrics;

— defines the purpose, intent, and interpretation of face image quality scores.

Performance assessment of quality algorithms and standardization of quality algorithms are outside the scope of this part of ISO/IEC 29794.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 29794-1, *Information technology — Biometric sample quality — Part 1: Framework*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 29794-1 and the following apply.

3.1
comparison score
numerical value (or set of values) resulting from a comparison

3.2
face quality assessment algorithm
algorithm that computes a quality score for a given face image sample

3.3
facial image
electronic image-based representation of the portrait of a person

4 Abbreviated terms

CCD Charge-coupled device

DCT Discrete Cosine Transform

GCF Global Contrast Factor

FQAA Face Quality Assessment Algorithm

QS Quality Score

FQS Face Quality Score

QSN Quality Score Normalization

5 Approaches to Face Image Quality

Face Image Quality can be defined in many ways, depending on the application. For the purpose of this part of ISO/IEC 29794 standard Face Image Quality is defined in relation to the use of facial images with automated face recognition systems. The performance of an automated face recognition system is affected by the amount of defect or the degree of imperfection present in the face image. The knowledge of quality can, and is currently being used to, process face images differently, by either invoking some image enhancement or normalization methods prior to feature extraction, invoking different matchers based on quality, or simply changing the threshold. The use of face image quality assessment to enhance the overall performance of the system is increasing [3, 4, 5].

A very important application of real-time quality analysis of faces is Face Recognition in Video, also referred to as Face in a Crowd, Recognition on the move, or Face at a Distance, e.g [21].

This part of ISO/IEC 29794 shows some approaches for estimating Face Image Quality. The aim is to give the reader examples of assessment algorithms. Note, that these algorithms have pros and cons and no one algorithm is likely to be suitable for all facial images. Standardization of these algorithms is out of scope of this part of ISO/IEC 29794.

The following related work is being done in ISO/IEC JTC1 SC37 [1, 2]:

- ISO/IEC 29794-1 suggests the use of Quality Algorithm Identification (QAID), or Quality Score Percentile Rank upon standardization of a Quality Score Normalization Dataset (QSND).

This part of ISO/IEC 29794 adopts the following approach for face sample quality description:

- Specifying characterization of the facial quality and possible defects of face biometric samples in categorized aspects.
- Showing how FQAAs can be used to derive face quality scores (FQSs) related to specific characteristics and associated possible defects. An FQAA typically analyzes a face sample locally at the pixel or feature level and fuses the local analysis results over a global region. An FQS evaluates one or more characteristics and associated potential defects, and provides an indicator of the quality.

A typical approach of a system for generation of quality scores for facial images then takes the atomic FQSs generated by the FQAAs and combines them to a final quality score. The final quality score must predict performance metrics such as either false match or false non-match of an automatic facial image recognition.

6 Categorization of Facial Quality

Different factors affect the quality of the facial image with respect to biometric systems' performance. A successful recognition will be based on the biometric characteristics of the subject and a number of factors that influence these characteristics such as variations (e.g. due to ageing) and the environmental conditions in the acquisition process:

- Influence of subjects characteristics on biometric performance,
- Influence of the acquisition process (including the capturing device) on biometric performance.

This classification is not sufficient, as it does not distinguish between static and dynamic characteristics and properties:

- static subjects characteristics are related to anatomical characteristics of the subject,
- dynamic subject characteristics are related to subjects behaviour during the acquisition process,
- static properties of the acquisition process are related to physical properties of the capturing device and effects caused by the sample processing chain,
- dynamic properties of the acquisition process are related to environmental conditions during the capturing process.

Table 1 shows a classification scheme that differentiates between the dynamic versus static properties as well as the subject versus the acquisition process characteristics affecting facial quality.

Table 1 — Characterization of Facial Quality

	Subject characteristics	**Acquisition process**

Static	Biological characteristics, like
	- anatomical characteristics (e.g. head dimensions, eye positions)
	- injuries and scars
	- ethnic group
	- impairment
	
	Other static characteristics
	- Heavy facial wears, such as thick or dark glasses
	- Makeup
	- Permanent jewellery
		Acquisition process and capture device properties, like
		- image enhancement and data reduction process
		- physical properties (e.g. image resolution and contrast)
		- optical distortions
		- static properties of the background, e.g. wallpaper
		- camera characteristics
		    o sensor resolution
		- scene characteristics
		    o geometric distortion

Dynamic	Subject characteristics and behaviour, like
	- closed eyes
	- (exaggerated) expression
	- hair across the eye
	- head pose
	- subject posing (frontal / non frontal to camera)
		Scenery, like
		- dynamic characteristics of the background like moving objects
		- variation in lightning and related potential defects as
		    o deviation from the symmetric lighting
		    o uneven lighting on the face area
		    o Extreme strong or weak illumination
		- subject posing , e.g.
		    o too far (face too small), or too near (face too big)
		    o out of focus (low sharpness)
		    o partial occlusion of the face
		- Acquisition process and capture device properties, such as
		    o camera characteristics
		    o dynamic range (response to weak and strong lighting)

The classification scheme (as all other content of this part of ISO/IEC 29794) is given for informative purposes only. The proposed classification scheme is certainly not the only possible scheme, but it is very useful since it separates design from character, i.e. it can be used to guide quality by design and hence performance improvement.

This characterization and the related categories of defects, degradations and interferences affect the performance of an automated facial recognition system. What is not considered in this part of ISO/IEC 29794 are the effects of printing on the given facial images (e.g. in a passport production process), which likely introduces further distortions especially with respect to image appearance and noise.

7 Facial Image Quality Analysis

Different aspects have to be considered in a facial image quality analysis. Some of them are already defined in related standardization documents. Different categories can be identified:

1. image properties like the size of the image or its resolution,

2. image appearance characteristics like the exposure or noise,

3. scenery characteristics like lighting or background,

4. characteristics like the consistency between the skin colour on the image and the skin colour of the subject,

5. the behaviour of the subject.

For some of these properties and characteristics metrics already exist. Some properties and characteristics, however, are much harder to be assessed and evaluated like the consistency of the skin colour on the image and the skin colour of the subject.

Furthermore, for some properties and characteristics, like the eye distance (in pixels), requirements are defined in ISO/IEC 19794-5 [37]. Their evaluation requires more complex algorithms and technologies from computer vision and image understanding. Therefore, a simple metric can not be given without considering the implementation that is needed to extract the corresponding features. In addition to this, different core concepts might be possible, e.g. different principles exist to automatically determine the eye positions in facial images. It may be possible to derive normalized quality scores as described in ISO/IEC 29794-1 (QSND). For some metrics, the variation between the enrolled images and that of the query images plays a bigger role in predicting performance than does the absolute metric applied to a single image. For instance [38] shows that performance is more affected by the relationship between the resolution of enrolled images and the query images than by absolute measure of resolution applied to each of them.

An FQAA can examine the image without a segmentation of the facial area (e.g. to assess static characteristics of the acquisition process like the compression rate and resulting compression artefacts, sensor resolution when measuring the size of an image) or perform an analysis on the facial area only (e.g. when estimating the pose of a subject). Local structures of a face may be defined by pixel values (raw or processed) within local regions; they may be fused globally to give a single quality score. Various FQAA can be developed, for different quality aspects related to environment, of camera, and/or subject showing different performance on different data sets. It is out of scope of this part of ISO/IEC 29794 to rate or rank the different approaches.

For some of the quality measures it is assumed that the face has been detected, and the facial area is normalized properly in geometry according to some landmarks such as the eye positions. Only the cropped face region is used for the analysis in this case.

7.1 Dynamic Subject Characteristics

7.1.1 Subject's Behaviour

Typical characteristics that are related to the subject's behaviour include:

- closed or open eyes,
- closed or open mouth,
- any kind of expression, e.g. smiling or neutral,
- head pose, e.g. frontal or rotated in any direction.

Similar to the scenery properties or the characteristics, the quantification of these parameters requires the recognition of background, faces and facial characteristics.

Again, different core algorithms can be implemented and their performance values can be used. A reduction in complexity can be achieved by selecting algorithms or concepts that are most commonly applied if this information is available.

7.1.2 Analysis Based on Statistical Differences of the Left and Right Half of the Face

7.1.2.1 Lighting Symmetry

The following approaches are based on the assumption that the images being analyzed are 2D portrait style images such as those specified in ISO/IEC 19794-5:2005/Amd.1, *Information technology — Biometric data interchange formats — Face image data — Amendment 1: Conditions for taking photographs for face image data*. This relates to facial and environment semantics. Left-right symmetry can be used to evaluate quality of lighting and pose [36]. The face region is divided into left and right halves at the mid-line of the eyes (Figure 1).The symmetry analysis below examines differences between the corresponding left-right locations. The difference value indicates the degree of asymmetry in some local image properties, e.g., raw pixel value, or locally-filtered pixels value. The local image filter can be Gabor filter [9,10], Local Binary Pattern (LBP) filter [11-13], Ordinal filter [14-17], or any other suitable local filter. The left-right difference value provides a quality score for the lighting (i.e. how symmetric the lighting is), or the pose quality (i.e. how frontal the pose is). Although the majority of faces seem to be left –right symmetric some faces could have significant deviation, e.g. caused by marks, discoloration etc. that would affect symmetry based quality analysis metrics Images are taken from Yale face image database [18].

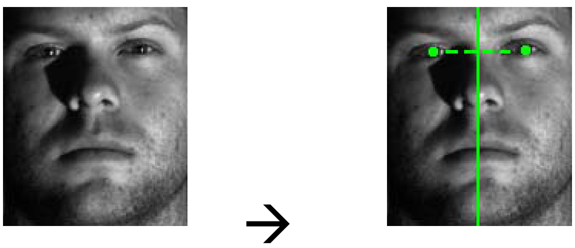

Figure 1 — Division of a face into left and right half regions at the mid-line of the eyes

The difference can be based on histograms H^L_{m*n} and H^R_{m*n} of some local features in the left and right half regions where *m* is the dimensionality of the feature vector, and *n* is the number of bins in the histogram. A histogram difference can be calculated as follows:

$$D_i = \left| H^L_{m*n} - H^R_{m*n} \right|$$

where the metric |.| is some suitable form of histogram distance, e.g., histogram intersection, cross-entropy, Kullback–Leibler divergence. The larger the difference value is, the less left-right symmetric the face image is, and the lower the image quality is in some aspect.

One possibility is using image normalized pixel values. The following presents an example of an FQAA for lighting symmetry:

(1) Normalize the range of pixel values in the cropped face region using a suitable normalization or equalization algorithm.

(2) Calculate the difference between normalized values for each pixel pair of sub windows at left-right mirror 2 locations.

(3) Calculate a suitable sum of the abstract values of the differences.

The sum is a metric of lighting asymmetry. The larger the sum value is, the less the left-right symmetric the face image is, and the lower the image quality is in terms of the lighting symmetry.

Figure 2 shows the lighting asymmetry for two face images. 2300 pairs of pixels are randomly selected across an image. The horizontal axis indexes the pair number. The vertical axis corresponds to the lighting asymmetry. The frontal lighted face image (the one on the left) has lower lighting asymmetry (the darker curve), whereas the sided lighted one has higher lighting asymmetry (lighter curve). Figure 3 shows another example, where the lighting is even more asymmetric. As can be seen, the symmetrically lighted face has much lower difference values.

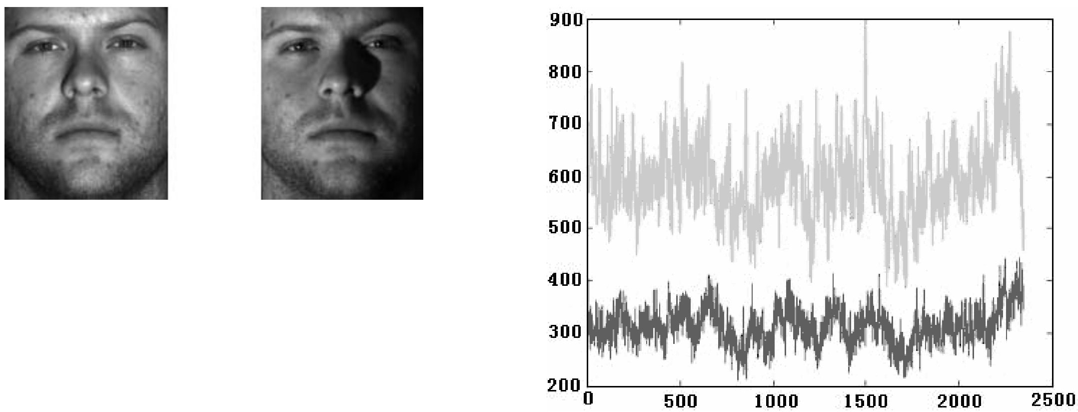

Figure 2 — A result of asymmetric lighting, and the distributions of asymmetry

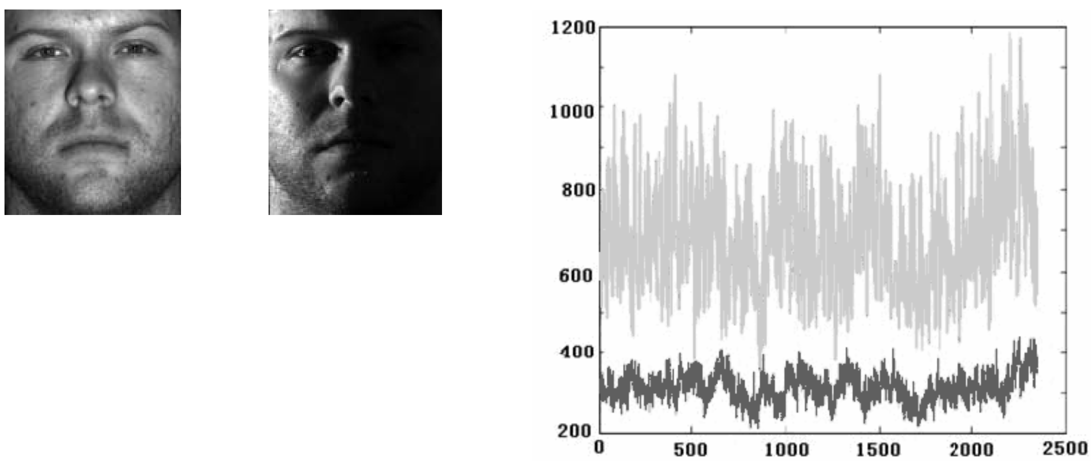

Figure 3 — Another result of asymmetric lighting, and the distributions of asymmetry

Note, that this method only works for frontal or almost frontal images. On the other hand, if we assume that asymmetry of the object is small in comparison to the other factors we can use this to evaluate quality of lighting, to the extent that the image symmetry mismatch is due to these conditions

7.1.2.2 Pose Symmetry

This should be done based on pose-sensitive image properties. One possibility is using Local Binary Pattern (LBP) filtered pixel values. The following presents an example of an FQAA for pose symmetry:

(1) Perform filtering using LBP filters.

(2) Calculate the difference between filtered values for each pixel pair of sub windows at left-right mirror locations.

(3) Calculate a suitable sum of the absolute values of the differences.

The sum is a metric of pose asymmetry. The larger the sum value is the more the face is rotated left or right and the lower the image quality is in terms of pose symmetry.

Figure 4 gives an example of the pose asymmetric values for 4 pose categories (different curves) and 10 people (horizontal axis) of the local differences for the following two face images (one with symmetric lighting and the other not). The differences are calculated between pairs of pixels at 2300 random locations. The means of the differences are plotted. From bottom to top, the curves correspond to the four pose categories from left to right. The curves (from bottom to top) correspond to the face asymmetry values for the four pose categories from left to right.

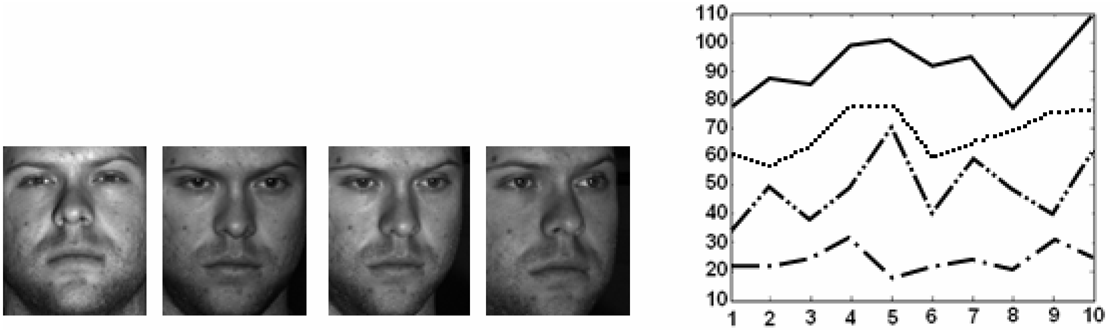

Figure 4 — Result of pose asymmetry

Note, that from the outcome of the described method one cannot distinguish between asymmetry of the face and/or asymmetry of the illumination and/or pose. On the other hand, if we assume that asymmetry of the object is small in comparison to the other factors we can use this to evaluate quality of lighting and pose, to the extent that the image symmetry mismatch is due to these conditions.

7.2 Static Characteristics of the Acquisition Process

Typical scenery characteristics describing the environmental influence are

- image enhancement and data reduction process, i.e. image resolution and size,
- static camera characteristics like resolution,
- static properties of the background like wallpapers.

Depending on the property or the characteristic, the quantification of these parameters requires the recognition of background, faces and facial characteristics.

Here, different core algorithms can be implemented and their performance values can be used. A reduction of the complexity can be achieved by selection of the algorithms or concepts that are applied in the most significant recognition system if this information is available.

7.2.1 Image Resolution and Size

An estimation of the size or the resolution is difficult as important parameters during the image acquisition process cannot be estimated. As an alternative, the numbers of row and column pixels can be used as an indication of nominal resolution for a standard subject-to-camera geometry. More meaningful measures of image resolvability may be made once the head and face are segmented from the image background and particularly once eye coordinates are located. Interpupillary distance in pixels provides a measure of pixel coverage relative to facial features. Moreover, applying a statistical measure of average interpupillary distance, e.g., 63mm [19], pixel density may be converted to spatial sample rate in, e.g., samples/millimetre.

This does not, however, take into account image processing operations that remove detail information and therefore reduce the resolution of the image. Among these operations are

- low-pass filtering and high-frequency noise removal,
- a down-sampling process that is followed by an up-sampling process.

7.2.2 Noise

The noise in facial images depends on the different processes that are required to result in a digital image. The introduced noise is specific according to the device or process involved. Relevant noise sources include

- digital image acquisition devices, e.g. the image sensor of a digital camera,
- analogue image acquisition devices,
- image scanning devices,
- image compression algorithms, e.g. JPEG or Wavelet compression.

7.2.2.1 Image Acquisition Noise

Liu et al. [22] present a method to estimate an upper bound for the CCD acquisition noise. The authors estimate this upper bound based on a piecewise smooth image prior model and measured CCD camera response functions.

7.2.2.2 DCT Compression Noise

As an example, Yang et al. [27] propose a method to estimate the compression artefacts introduced by the DCT compression. For this, blocky compression artefacts in each pixel vector is modelled as a shape vector weighted by the boundary discontinuity. The authors suggest estimating the boundary discontinuity from the difference between the pixel gradient across the block boundary and that of the internal pixels.

Another approach for measuring block artefacts is suggested by Zhou et al. [28]. Furthermore note, that image noise actually reduces other artefacts at a specified quality specification as the additional high frequency content suppresses the compression level at a specified quality level, see Shahnaz et. al. [29].

7.3 Characteristics of Image Acquisition

7.3.1 Image Properties

Different standards exist that allow storing meta data in images directly. For example, today's digital cameras store meta data information in the EXIF format. EXIF “is a specification for the image file format used by digital cameras” [30]. Note, that EXIF is not supported by JPEG 2000.

EXIF information contains information about the state of a camera at the time of capture. Some of this information (e.g. exposure time) is certainly useful for quality assessment. Within the EXIF standard different meta data tags are defined, e.g.

- date and time information,
- camera settings like exposure,
- location information,
- descriptions and copyright information.

Although this information is valuable neither its availability nor its authenticity can be guaranteed. In many passport application processes the applicant will present an analogue image. Therefore, image properties have to be verified without relying on meta data that is included in the image source.

7.3.2 Image Appearance

Image appearance itself depends on the colour distributions in the images. While these effects are significant for the perceived quality for humans, it is not clear how strongly these factors generally affect an automatic recognition system.

7.3.3 Illumination Intensity

This analysis calculates quality score for illumination strength, to evaluate whether the illumination is too strong or too weak. It is performed on the histogram of un-normalized image pixel values of the entire image. The histogram with normal illumination generally spans a wider range. In the case of very dark or very bright illumination, the distribution of the greyscale values is concentrated towards the lower or higher end of the histogram.

Let H_0 be the histogram of standard illumination and H is the histogram of the image being assessed. A quality score could be defined based on the distribution differences between the two histograms H_0 and H. Note that it might be quite difficult to determine the histogram of a standard illumination H_0 universally because H_0 strongly depends on the physical scene related to the image.

7.3.4 Image Brightness

The brightness of an image can be investigated by considering the histogram of the intensity values, e.g. the luminance in the YC_BC_R colour system. For a given histogram h_i of intensity values i with $i \in \{0,...,N\}$, where N is the maximum possible intensity value, of the entire image the histogram can be considered as a probability density functions and can be characterized by its statistical moments:

- Mean

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

- Variance

$$v(h_0,...,h_N) = \frac{1}{N}\sum_{i=0}^{N}(h_i - \bar{h})^2 \text{ respectively } \sigma = \sqrt{Var(h_0,...,h_N)}$$

- Skewness

$$s(h_0,...,h_N) = \frac{1}{N+1}\sum_{i=0}^{N}\frac{\left|h_i - \bar{h}\right|^3}{\sigma^3}$$

- Kurtosis

$$k(h_0,...,h_N) = \left\{\frac{1}{N+1}\sum_{i=0}^{N}\frac{\left|h_i - \bar{h}\right|^4}{\sigma^4}\right\} - 3$$

Kurtosis [31, 32] is a measure of the “peakedness” of the probability distribution function. Data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. A uniform distribution would be the extreme case.

7.3.5 Image Contrast

7.3.5.1 General

While contrast for simple images is well defined, contrast for complex images is not [20].

Real world images and therefore facial images can be considered as complex images while simple images rather contain test patterns like sinusoidal gratings.

As contrast is defined with respect to human perception, the effects of contrast on biometric systems might differ. Thus, it is suggested to consider different contrast definitions.

This analysis calculates quality scores for image contrast. It is performed on the histogram of un-normalized image pixel values of the entire image. It could be calculated via

$$C_I = L_I \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[I(x,y) - \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y) \right]^2} \;,$$

where C_I is the contrast quality score, $I(x,y)$ is the intensity value of the image, M, N are the image width and height, respectively, and L_I is a scaling factor. The following method provides an alternative: In the following, a given facial image is considered as a two-dimensional function f that maps a position in the facial image to an intensity value $I : f : Z^2 \to Z$, respectively $I = f(x,y)$. For a given image $f(x,y)$, $F(k_x,k_y)$ respectively $F(k_\rho,k_\tau)$ represents the corresponding Fourier transformation in the Cartesian respectively in the polar coordinate system.

Michelson defined for I_{max} and I_{min} (maximum respectively the minimum values in the given simple pattern) the contrast according to the formula [26]

$$C_{Michelson} = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}}.$$

Similarly, the Weber contrast is defined as

$$C_{Weber} = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{min}}}.$$

As another approach in [20] the contrast of a complex image in the local frequency space is proposed as

$$C(k_x,k_y) = 2\frac{F(k_x,k_y)}{F(k_x = 0, k_y = 0)}.$$

Alternatively, the energy or information content is used for contrast definition, i.e. the contrast is calculated using the power spectrum of the image

$$C_{pow}(k_x,k_y) = 2\left|\frac{F(k_x,k_y)}{F(k_x = 0, k_y = 0)}\right|^2.$$

A contrast number for the picture can be defined by the sum of all frequency

$$C_{Hess} = \frac{\sum_{k_x}\sum_{k_y} C_{pow}(k_x,k_y)}{w * h}$$

where h and w are the height and the width of the given image.

7.3.5.2 Perceived Contrast Considering Spatial Frequencies

In [26], this idea is extended by consideration of the dependency of human perceived contrast on the spatial frequency. Recently, a Global Contrast Factor (*GCF*) was suggested [23].

This approach is based on the local contrast $l_c(x,y)$ at a given position $f(x,y)$ in a 4 pixel neighbourhood

$$l_c(x,y) = \frac{|f(x,y)-f(x-1,y)| + |f(x,y)-f(x+1,y)| + |f(x,y)-f(x,y-1)| + |f(x,y)-f(x,y+1)|}{4}.$$

Based on this local contrast $l_c(x,y)$, an overall local contrast - for the given image at resolution *res*- is calculated by

$$C_{res} = \frac{1}{w \cdot h} \sum_x \sum_y l_c(x,y).$$

where w and h are the width and height of the image respectively.

The Global Contrast Factor (*GCF*) is defined as the weighted sums of the overall local contrast for different resolutions

$$GCF = \sum_{res=1}^{N} w_{res} C_{res}.$$

Different resolution levels are created by a down sampling with the factors 2, 4, 8, 16, 25, 50, 100 and 200.

Experiments showed that for the given images, the optimum approximation of the weight factors was

$$w_i = (-0.406385 \cdot \frac{i}{9} + 0.334573) \cdot \frac{i}{9} + 0.0877526 \text{ where } i \in \{1,\ldots,9\}.$$

7.3.6 Exposure

Exposure can be partially identified based on brightness and contrast. Exposure might be characterized by the degree to which the image pixel values are distributed over the grey scale or over the range of values available in each colour channel. Accordingly, statistical measures of pixel intensity distribution such as entropy [33] may serve as potential exposure metrics or as the basis for such. Possible metrics are discussed by Shirvaikar [34].

In a well exposed image very few pixels will take extreme values (e.g. 0 or 256 in the case of an 8 bit image). A large number of pixels at the extremes indicate an under/over exposed image [35].

7.3.7 Focus, Blur and Sharpness

Marziliano et al. [25] presented an algorithm to identify the blur in the horizontal direction. It identifies vertical edges in a given image. Each row of the edge image is scanned. Start and end positions of edges are identified and the corresponding edge width is calculated. This is considered as the local blur value for an edge. The global blur measure is derived by averaging the local blur values over all edge locations.

$$b_{global} = \frac{1}{N} \sum_{i=1}^{N} width_i \,'$$

The sharpness of a face image refers to the degree of clarity in both coarse and fine details in the face region. The quality value for sharpness can be calculated via image gradient

$$G = \sum_{x=1}^{M-2} \sum_{y=1}^{N-2} G(x,y)$$

where $G(x,y)$ is the gradient value at (x,y) [35].

7.3.8 Colour

Colour naturalness depends on the subject and that background in the facial image. Deriving an objective metric is difficult. However, if the colour of the background is known it can be used to calibrate the image – i.e. 18% grey.

Colour information is known not to affect current automated face recognition performance of major face recognition systems. That is, recognition from black-and-white photos is the same as from a colour photo. Colour however can be used in face detection and tracking as well as in eye localization. This information could be taken into account when designing a colour-based metric.

Furthermore, colour depth, i.e the number of bits used to represent the colour of a single pixel in a bitmapped image or video frame buffer, can influence face and/or eye-finding. Therefore, typically 8 bits per pixel for a grey-scale image and 24-bits per pixel for a colour image should be used to prevent such issues.

7.3.9 Subject-Camera distance

Subject-camera distance is recommended to be 1,2 m to 2,5 m in a typical photo studio and 0,7 m to 1,0 m in a typical photo booth (ISO/IEC 19794-5:2005/Amd.1); and inter-eye distance to be 120 pixels (ISO/IEC 19794-5:2005, Annex A). As the subject-to-camera distance is inversely related to the size of the face, the inter-eye distance, may be taken as an indication that the subject is not at the distance from the camera, that was anticipated at the time of camera setup Therefore, the inter-eye distance (Dist_{eye}) can be used to estimate the quality score whether the subject is at a proper distance from the camera:

$$QS_{uc\text{-}dist} = \mathrm{D}(\mathrm{Dist}_{eye}, \mathrm{Dist}_0)$$

Where available EXIF data often includes the focal length of the camera. This information may be used to estimate the distance of the subject to the camera.

The number of pixels between the centres of the eyes is one parameter for characterizing performance. This factor interacts with other camera parameters as discussed in sections 7.2.1.

Bibliography

[1] ISO/IEC JTC 1/SC 37 Standing Document 2, Version 9, *Harmonized Biometric Vocabulary*

[2] ISO/IEC JTC 1/SC 37 N 1128 Quality Rapporteur Group Report, 2005

[3] Fronthaler, H.; Kollreider, K.; Bigun, J., “Automatic Image Quality Assessment with Application in Biometrics” , Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW’06) , pp.30-37, June 2006

[4] Frank Weber, Some quality measures for face images and their relationship to recognition performance Cognitec Systems GmbH. NIST Biometric Quality Workshop 2006

[5] Martin Werner and Michael Brauckmann, Quality Values for Face Recognition Viisage Inc. NIST Biometric Quality Workshop 2006

[6] E. Tabassi, C.L. Wilson, “A novel Approach to Fingerprint Image Quality”, Proc. IEEE International Conference on Image Processing, 2005

[7] E. Tabassi, C.L. Wilson and C.I. Waston. Fingerprint Image Quality, NFIQ, National Institute of Standards and Technology, NISTIR 7151, 2004

[8] N.K. Ratha and R.M. Bolle. “Finger image quality estimation”, IBM computer science research report RC21622,1999

[9] Peng Yang, Shiguang Shan, Wen Gao, Stan Z. Li, Dong Zhang. "Face Recognition Using Ada-Boosted Gabor Features". The 6th International Conference on Automatic Face and Gesture Recognition. Seoul, Korea, May, 2004

[10] C. Liu and H. Wechsler: "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition", IEEE Trans. Image Processing, vol. 11, no. 4, pp. 467-476, 2002

[11] T. Ojala, M. Pietikainen and D. Harwood, “A comparative study of texture measures with classification based on feature distributions”, Pattern Recognition, vol. 29 no. 1, pp. 51-59, January 1996

[12] T. Ojala, M. Pietikainen, and M. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification width local binary patterns”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, pp. 971–987, 2002

[13] Guangcheng Zhang, Xiangsheng Huang, Stan Z. Li, Yangsheng Wang & Xihong Wu, “Boosting Local Binary Pattern (LBP)-Based Face Recognition”, Proc. Advances in Biometric Person Authentication: 5th Chinese Conference on Biometric Recognition, SINOBIOMETRICS 2004, Guangzhou, China, pp. 179-186, Dec 2004

[14] J. Sadr, S. Mukherjee, K. Thoresz and P. Sinha, “The fidelity of local ordinal encoding”, In Proceedings of the Fifteenth Annual Conference on Neural Information Processing Systems (NIPS 2001), pp. 1279-1286 Vancouver, British Columbia, Canada, December 3-8 2001

[15] P. Sinha, “Qualitative representations for recognition”, In Proceedings of the Second International Workshop on Biologically Motivated Computer Vision, pp. 249–262, Tubingen, Germany, November 22-24 2002

[16] K. J. Thoresz, “On qualitative representations for recognition”, Master’s thesis, MIT, July 2002

[17] ShengCai Liao, Zhen Lei, XiangXin Zhu, ZheNan Sun, Stan Z. Li, Tieniu Tan, "Face Recognition Using Ordinal Features, In Proceedings of IAPR International Conference on Biometrics”, (ICB-2006), pp. 40-46, Hong Kong, Jan 2006

[18] Athinodoros S, Georghiades and Peter N.Belhumeur, “From Few to many: Illumination cone models for face recognition under variable lighting and pose”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 6, pp 643-660, 2001

[19] N. A. Dodgson. “Variation and extreme of human interpupillary distance,” Proc. SPIE Vol. 5291, Stereoscopic Displays and Virtual Reality Systems XI, A. J. Woods, J. O. Merritt, S. A. Benton and M. T. Bolas (eds.), 19–22 January 2004, San Jose, California, USA, pages 36–46

[20] R. F. Hess, A. B. u. L. P. Contrast-coding in amblyopia. “In Differences in the neural basis of human amblyopia”, Proc. R. Soc. London, pp. 419–422,1983

[21] A. Fourney, R. Laganière, Constructing Face Image Logs that are Both Complete and Concise”, Intern. Workshop on Video Processing and Recognition (VideoRec'07), in Proceedings of Fourth Canadian Conference on Computer and Robot Vision (CRV '07) http://computer-vision.org/VideoRec07/program.html

[22] Ce Liu William, T. Freeman, “Noise Estimation from a Single Image,” Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1, 2006, pp. 901 – 908

[23] Kresimir Matkovic, Laszlo Neumann, Attila Neumann, Thomas Psik, and Werner Purgathofer, “Global Contrast Factor - a New Approach to Image Contrast, in: Computational Aesthetics in Graphics”, L. Neumann, M. Sbert, B. Gooch, W. Purgathofer (Editors), “Visualization and Imaging”, 2005

[24] Pina Marziliano, Frederic Dufaux, “A no-reference perceptual blur metric. In Image Processing Proceedings 2002 International Conference”, vol. 3, pp. III–57–III–60, 2002

[25] Pina Marziliano, Frederic Dufaux, Stefan Winkler and Touradj Ebrahimi, “A No-Reference perceptual Blur Metric, in Proc. International Conference on Image Processing, vol. 3, pp. 57-60, Rochester, NY, September 22-25, 2002

[26] E. Peli, “Contrast in complex images”. Journal of Optical Society of America 7,10, 2032–2040, 1990

[27] Jeonghun Yang, Hyuk Choi, Taejeong Kim, "Noise estimation for blocking artifacts reduction in DCT coded images", IEEE Transactions on Circuits and Systems for Video Technology, vol. 10 no. 7 October pp. 1116-1120, Oct. 2000

[28] W. Zhou and A. C. Bovik, "A universal image quality index," Signal Processing Letters, IEEE, vol. 9, no. 3, pp. 81-84, 2002

[29] R. Shahnaz, J. F. Walkup, and T. F. Krile, "Image Compression in Signal- Dependent Noise," Appl. Opt. 38, 5560-5567 (1999)

[30] Japan Electronics and Information Technology Industries Association, JEITA CP-3451 Exchangeable image file format for digital still cameras: Exif Version 2.2

[31] Kenney, J. F. and Keeping, E. S. "Kurtosis." §7.12 in Mathematics of Statistics, Pt. 1, 3rd ed. Princeton, NJ: Van Nostrand, pp. 102-103, 1962

[32] Joanes, D. N. & Gill, C. A. (1998), Comparing measures of sample skewness and kurtosis. Journal of the Royal Statistical Society (Series D): The Statistician 47 (1), 183–189

[33] R.C. Gonzalez, R.E. Woods, S.L. Eddins, Digital Image Processing Using MATLAB, New Jersey, Prentice Hall, 2003

[34] Mikul V. Shirvaikar, An optimal measure for camera focus and exposure, System Theory, 2004. Proceedings of the Thirty-Sixth Southeastern Symposium on 2004 Page(s):472 – 475

[35] A.K. Jain, Fundamentals of Digital Image Processing, pp 348-350, Prentice Hall, 1989

[36] Xiufeng Gao, Stan Z. Li, Rong Liu, Peiren Zhang. "Standardization of Face Image Sample Quality". In Proceedings of IAPR/IEEE International Conference on Biometrics (ICB-2007), pp 242-251. Seoul, Korea, August 2007

[37] ISO/IEC 19794-5, *Information technology — Biometric data interchange formats — Part 5: Face image data*

[38] Beveridge R. J., Givens G. H., Philips P. J., Draper B. A., Lui Y. M., “Focus on Quality, Predicting FRVT 2006 Performance”, 8th IEEE International Conference on Automatic Face and Gesture Recognition, 2008

ICS 35.040

Price based on 15 pages