Face Image Quality Assessment

A MINOR PROJECT REPORT

Submitted by

Susnata Das[RA2011030010043]
Rohan Kumar[RA2011030010050]
Aniruddha Ghosh[RA2011030010038]

Under the guidance of

VIJAYALAKSHMI M

(Assistant Professor, Department of Computing Technologies)

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING of

FACULTY OF ENGINEERING AND TECHNOLOGY

S.R.M.Nagar, Kattankulathur, Chengalpattu District **NOVEMBER 2022**

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY (Under Section3 of UGC Act,1956)

BONAFIDE CERTIFICATE

Certified that 18CSP107L minor project report [18CSP108L internship report] titled "**Face Image Quality Assessment**" is the bonafide work of **Susnata das[RA2011030010043], Rohan Kumar[RA2011030010050] ,Aniruddha Ghosh [RA2011030010038]** who carried out the minor project work under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE:

VIJAYALAKSHMI M
ASSISTANT PROFESSOR
18CSE357T-Biometrics
Course Faculty
Department of Computing Technologies

ABSTRACT

Face image quality is an important factor to enable high-performance face recognition systems. Face quality assessment aims at estimating the suitability of a face image for recognition. Previous works proposed supervised solutions that require artificially or human labeled quality values. However, both labeling mechanisms are error-prone as they do not rely on a clear definition of quality and may not know the best characteristics for the utilized face recognition system. Avoiding the use of inaccurate quality labels, we proposed a novel concept to measure face quality based on an arbitrary face recognition model. By determining the embedding variations generated from random sub networks of a face model, the robustness of a sample representation and thus, its quality is estimated. The experiments are conducted in a cross-database evaluation setting on three publicly available databases. We compare our proposed solution on two face embedding against six state-of-the-art approaches from academia and industry. The results show that our unsupervised solution outperforms all other approaches in the majority of the investigated scenarios. In contrast to previous works, the proposed solution shows a stable performance over all scenarios. Utilizing the deployed face recognition model for our face quality assessment methodology avoids the training phase completely and further outperforms all baseline approaches by a large margin. Our solution can be easily integrated into current face recognition systems and can be modified to other tasks beyond face recognition.

TABLE OF CONTENTS

INTRODUCTION

Face Image Quality Assessment (FIQA) refers to the process of taking a face image as input to produce some form of “quality” estimate as output. A **FIQA algorithm (FIQAA)** is an automated FIQA approach. While FIQA and general **Image Quality Assessment (IQA)** are overlapping research areas, there are important distinctions. Most of the published FIQA literature focuses on single face image input in the visible spectrum. Therefore, unless otherwise specified in this survey, FIQA(A) refers to single-image Face Image Quality Assessment (Algorithms) in the visible spectrum, with a **Quality Score (QS)** output that can be represented by: (A) a single scalar value or (B) a vector of quality values measuring different quality-related features. For a discussion of (F)IQA that instead compares two image variants, i.e., full/reduced-reference method.

The term “quality” is an intrinsically subjective concept that can be defined in different ways, with ISO/IEC 29794-1 differentiating between three aspects referred to as character, fidelity, and utility. In the context of facial biometrics these can be described as follows

- **Character:** Attributes inherent to the source biometric characteristic being acquired (e.g., the face topography or skin texture) that cannot be controlled during the biometric acquisition process (e.g., scars).
- **Fidelity:** For a biometric sample, e.g., a face image, fidelity reflects the degree of similarity to its source biometric characteristic . For instance, a blurred image of a face omits detail and has low fidelity .
- **Utility:** The fitness of a sample to accomplish or fulfill the biometric function (e.g., face recognition comparison), which is influenced, i.e, by the character and fidelity. Thus, the term utility is used to indicate the value of an image to a receiving algorithm.

LITERATURE SURVEY

This survey considers “utility” as the primary definition of what a quality score should convey, which is in accordance to the quality score definition of ISO/IEC 2382-37 and the definition in the ongoing **Face Recognition Vendor Test (FRVT)** for face image quality assessment . Thus, a QS should be indicative of the **Face Recognition (FR)** performance. Note that this entails that the output of a specific FIQAA may be more accurate for a specific FR system, so the FIQA utility prediction effectivity ultimately depends on the combination of both, the FIQAA and the FR system. To facilitate interoperability, it is, however, desirable that the FIQAA is predictive of recognition performance in general for a range of relevant systems, instead of being dependent on a single FR technology.

In short, under this survey’s definitions, a FIQAA is typically meant to output a scalar quality score to predict the FR performance from a single face input image. Being able to predict FR performance without necessarily running an FR algorithm makes FIQA useful for a variety of scenarios, which are described further. FIQA as a predictor for FR performance has attracted the predominant interest of researchers so far and is thus the main focus in the present survey. FIQA for other tasks in the field of face biometrics, such as emotion analysis, attention-level estimation, gender or other soft biometrics recognition , and so on, may open interesting research lines in the future and can take advantage of current developments that employ FIQA for FR performance prediction.

The contributions of this survey are:

- An introduction to FIQA , i.e., including the distinction against general IQA the conceptual problem with single-image utility assessment , and an overview of both common and uncommon FIQA application areas.
- A categorization of the surveyed FIQA approaches with a taxonomy that differentiates between factor-specific and monolithic approaches, in addition to various other aspects.

- A survey of more than 60 FIQAA publications from 2004 to 2021 including condensed overview tables for the publications and their used datasets This part is meant for literature overview purposes and does not have to be read in sequence.

  Prior work listed varying publication numbers, with Hernandez-Ortega et al. being a recent example that contained a summary for some prior publications ranging from 2006 to 2020. A fingerprint/iris/face quality assessment survey by Bharadwaj et al. considered less than 10 FIQAA publications from 2005 to 2011. The European JRC-34751 report also listed some FIQAAs from 2007 to 2018. To our knowledge, this FIQA survey is the most comprehensive one to date.

- An introduction for the **Error-versus-Reject-Characteristic (ERC)** evaluation methodology , which is a standardization candidate in addition to being commonly used in recent FIQA literature, and a subsequent concrete evaluation that includes a variety FIQA approaches. The ERC introduction mentions details not considered in recent FIQA literature, and the evaluation discusses its weaknesses to note opportunities and challenges for future work.

- A detailed discussion of various FIQA issues and challenges, including avenues for future work.

This survey considers “utility” as the primary definition of what a quality score should convey, which is in accordance to the quality score definition of ISO/IEC 2382-37 and the definition in the ongoing **Face Recognition Vendor Test (FRVT)** for face image quality assessment. Thus, a QS should be indicative of the **Face Recognition (FR)** performance. Note that this entails that the output of a specific FIQAA may be more accurate for a specific FR system, so the FIQA utility prediction effectivity ultimately depends on the combination of both, the FIQAA and the FR system. To facilitate interoperability, it is, however, desirable that the FIQAA is predictive of recognition performance in general for a range of relevant systems, instead of being dependent on a single FR technology.

In short, under this survey's definitions, a FIQAA is typically meant to output a scalar quality score to predict the FR performance from a single face input image. Being able to predict FR performance without necessarily running an FR algorithm makes FIQA useful for a variety of scenarios, which are described further in Section . FIQA as a predictor for FR performance has attracted the predominant interest of researchers so far and is thus the main focus in the present survey. FIQA for other tasks in the field of face biometrics, such as emotion analysis attention-level estimation gender or other soft biometrics recognition, and so on, may open interesting research lines in the future and can take advantage of current developments that employ FIQA for FR performance prediction.

The contributions of this survey are:

- An introduction to FIQA ,i.a., including the distinction against general IQA (Section), the conceptual problem with single-image utility assessment (Section), and an overview of both common and uncommon FIQA application areas (Section).
- A categorization of the surveyed FIQA approaches with a taxonomy that differentiates between factor-specific and monolithic approaches, in addition to various other aspects.
- A survey of more than 60 FIQAA publications from 2004 to 2021 including condensed overview tables for the publications and their used datasets This part is meant for literature overview purposes and does not have to be read in sequence.

  Prior work listed varying publication numbers, with Hernandez-Ortega et al. being a recent example that contained a summary for some prior publications ranging from 2006 to 2020. A fingerprint/iris/face quality assessment survey by Bharadwaj et al. considered less than 10 FIQAA publications from 2005 to 2011. The European JRC-34751 report also listed some FIQAAs from 2007 to 2018. To our knowledge, this FIQA survey is the most comprehensive one to date.
- An introduction for the **Error-versus-Reject-Characteristic (ERC)** evaluation methodology (Section), which is a

standardization candidate in addition to being commonly used in recent FIQA literature, and a subsequent concrete evaluation that includes a variety FIQA approaches (Section). The ERC introduction mentions details not considered in recent FIQA literature, and the evaluation discusses its weaknesses to note opportunities and challenges for future work.

- A detailed discussion of various FIQA issues and challenges including avenues for future work.

2.2 FIQA versus IQA

FIQA can be seen as a specific application within the wider field of **Image Quality Assessment (IQA)**, which is a very active research area of image processing. Even though related to IQA, FIQA has been mainly developed within the biometric context and focuses on distinctive face features. Consequentially, general **IQA algorithms (IQAA)** have shown poor performance when directly applied to FIQA, and, conversely, the very specific FIQA algorithms usually do not generalize to the broader application field of IQA.

General non-biometric IQA typically aims to assess images in terms of subjective (human) perceptual quality, meaning that technically objective quality scores generated by such IQAAs usually intent to predict or model subjective perceptual quality.

Biometric FIQA, however, is usually concerned with the assessment of the biometric utility for facial biometrics, which can be objectively defined in the context of specific FR systems. FIQA works may also test or train FIQAAs using ground truth data stemming from human quality assessments, but for biometric purposes the intent still differs from general perceptual quality assessment, insofar that the question is how well the images can be used for facial biometrics, versus how good/undistorted the images look overall for a human.

It can be expected that perceptual quality and biometric utility coincide to some degree, thus general IQA can be utilized for FIQA as well. The reverse is less likely, since FIQA algorithms may be specifically developed for face images, so results for non-face images are not expected to be useful. This also means that FIQA can perform better for the purpose of biometric utility prediction than a general IQA that has not been developed with facial biometrics in mind. Some of the surveyed FIQA literature tested known IQA algorithms together with specialized FIQA algorithms.

2.3 Full/Reduced/No-reference Quality Assessment

IQA literature draws a distinction between approaches that require a "reference" version of the input and those that do not (not to be confused with biometric references e.g., in an FR database):

- **Full-reference:** IQA that compares the input image against a known reference version thereof, i.e., a version that is known to be of higher or equal quality. Conversely, the input image can be seen as a potentially degraded (e.g., blurred) version of the reference image.
- **Reduced-reference/Partial-reference:** Similar to full-reference IQA, a reference version of the input image has to exist first, but only incomplete information of the reference is known and used for the IQA, e.g., some statistics of the image. The distinction between full-reference and reduced-reference approaches is not necessarily clear, since full-reference approaches may also "reduce" their input to a different representation, with information loss, before the comparison step.
- **No-reference:** No reference version of the input image is required for the IQA. Note that such an IQAA can still use other forms of internal data: An IQAA could, e.g., utilize some fixed set of images unrelated to the input image and still be categorized as no-reference IQA. Likewise, machine learning IQA models are not automatically classified as reduced-reference IQA just because they incorporate information from training images.

3. Implementation

When processing a face, the features like variations in light, image quality, persons' pose, facial expressions and more should be taken into account. In order to identify the individuals correctly these variations must be minimized. In order to make the image more suitable for recognition purposes, the images need to be preprocessed. Image pre-processing and normalization is important part of face recognition systems as variations in lighting conditions dramatically decrease recognition performance.

3.2 FEATURE EXTRACTION BASED ON MPCA AND LPP

Generally, an image of size nxm pixels is exercised to represent a face image by means of a vector in a nxm dimensional space. Feature extraction or dimensionality reduction is a methodology to transform a high dimensional data set into a low-dimensional equivalent representation that assumes to retain most of the information regarding the underlying structure or the original physical phenomenon [10]. The main tendency of using feature extraction is its representation of data in a lower dimensional space that computes through a linear or non-linear transformation satisfying certain properties.

3.2.1DIMENSIONALITY REDUCTION USING MPCA

MPCA is a multilinear subspace learning method that extracts features directly from multi-dimensional objects. MPCA receives the set of face image samples of the same dimensions as input for feature extraction. The resultant output of the MPCA is the dimensionally reduced feature projection matrix of face images. MPCA algorithm for dimensionality reduction can be referred in

3.2.2 FEATURE MATRIX EXTRACTION USING LPP

Locality Preserving Projection (LPP) is one of the linear approximation obtained from the nonlinear Laplacian Eigenmap [11].

The dimension reduced feature projection matrices of face image samples obtained using MPCA is then fed as an input to the LPP algorithm. The LPP algorithm is available in.

3.3 FACE RECOGNITION USING L2 DISTANCE MEASURE

The dimensional reduced feature matrices of the training sample images obtained using the MPCA and LPP techniques are stored in a database. While we are testing the face images, the aforesaid techniques are applied to generate the feature matrix and thereby a similarity measure is carried out on the sample face images. Various face recognition systems may use different distance measures while matching query images with the nearest database images. Our Face recognition approach used here is performed using L2 distance measure. The L2 distance is computed between the face images present in the database and the query image for matching process. The similarity distance measure for a pair of face images is computed in which a threshold determines whether the face pair is classified as same or different.

EXPERIMENTAL RESULTS AND COMPARATIVE ANALYSIS

The proposed methodology is tested using the FERET database and AT&T database of faces. Performance is measured by the procedures of FERET and AT&T facial images. In particular, all the images were preprocessed using a simple geometric normalization, followed by resizing of the images. The images are divided into two mutual exclusive sets, the training set and the test set. The training set is used to initialize and prepare the system to recognize arbitrary images. The test set is the set of images used to evaluate the performance of the system once the training phase is completed. Here, in FERET database, we use nearly 80 images for training and 160 images for testing. In AT&T database, we take 100 images for training and 200 images for testing process. The size of the image is 32x32, and the experiments were conducted using the L2 distance measure. The equation to measure the biometric recognition accuracy is Accuracy=100-(FAR/FRR)/2 where FAR is the false acceptance rate and FRR is the false rejection rate. The percentage of the recognition accuracy for both the approaches using the FERET and AT&T databases are given in the following table 1 and 2.

Table 1: Recognition rates and Accuracy values on the FERET database

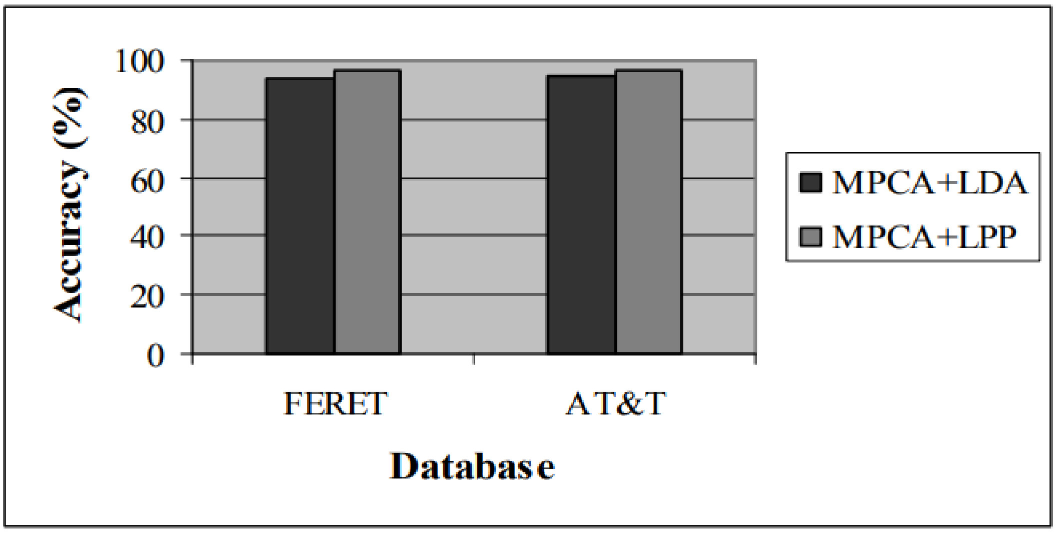
Performance Comparison		
Database	**Approach**	**Accuracy(%)**
FERET	MPCA+LDA	93.75
	MPCA+LPP	96.5

Table 2: Recognition rates and Accuracy values on the AT&T database

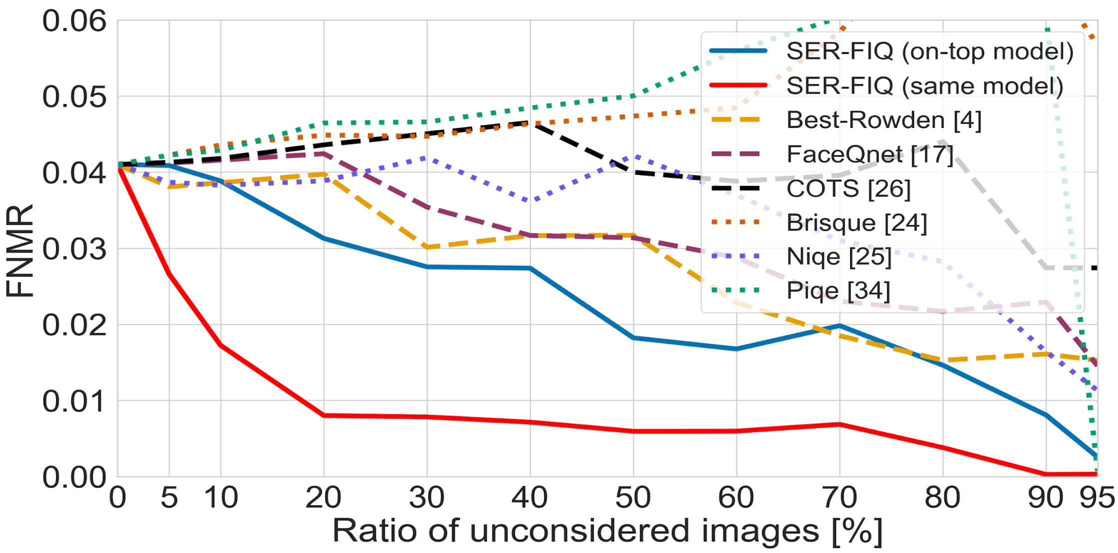
Performance Comparison		
Database	**Approach**	**Accuracy(%)**
AT&T	MPCA+LDA	95
	MPCA+LPP	96.5

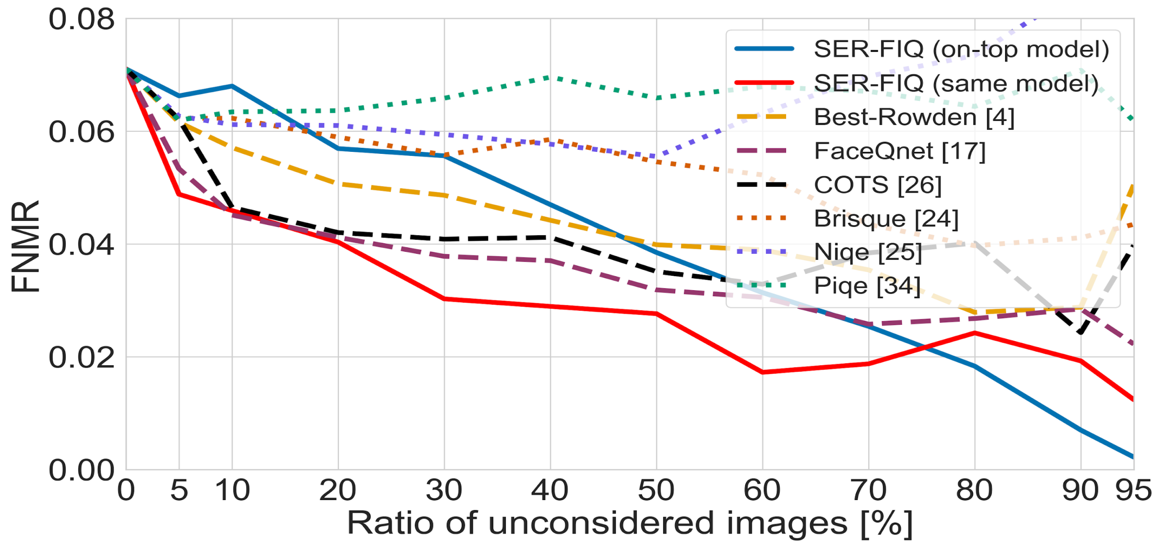
The comparative results are analyzed with the recognition rates of the feature extraction methods such as MPCA plus LDA and MPCA plus LPP by means of charts. The charts clearly shows that the recognition performance of the proposed approach MPCA plus LPP is more efficient when compared to MPCA plus LDA. The face recognition results of our approach is compared with MPCA+LDA based on the evaluation measures of recognition accuracy. The comparative result of each recognition rate is shown in the following chart.

Figure 1: Comparative results of recognition accuracy on MPCA+LDA and MPCA+LPP



Face image quality assessment results are shown below on LFW (left) and Adience (right). SER-FIQ (same model) is based on ArcFace and shown in red. The plots show the FNMR at FMR as recommended by the best practice guidelines of the European Border Guard Agency Frontex. For more details and results, please take a look at the paper.





4. CONCLUSION

Over two decades of research have resulted in successful techniques for recognizing the 2D facial images. In this 285 article, the face recognition method by combining the two popular appearance based techniques such as MPCA and LPP is presented. It also includes the comparison of face recognition approaches MPCA plus LDA and MPCA plus LPP. The combined appearance based technique such as MPCA and LPP yield to produce a high face recognition rate compared to the existing MPCA and LDA technique. Experimental results on FERET and AT&T database demonstrated the effectiveness of the proposed approach with improved recognition accuracy in comparison with the existing approach. In future, combination of various face recognition approaches could be experimented to identify the efficient approach in face recognition.

5. References

[1] https://github.com/pterhoer/FaceImageQuality

[2] https://ieeexplore.ieee.org/abstract/document/5752512

[3] https://sci-hub.se/10.1109/IEEEGCC.2011.5752512

[4] https://ieeexplore.ieee.org/document/8496511

[5] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=34173