Pattern Recognition 61 (2017) 153–168

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

# MDID: A multiply distorted image database for image quality assessment

Wen Sun [a,b], Fei Zhou [a,b,*,1], Qingmin Liao [a,b]

[a] Department of Electronic Engineering, Tsinghua University, China
[b] Shenzhen Graduate School, Tsinghua University, China

ARTICLE INFO

*Article history:*
Received 16 July 2015
Received in revised form
19 July 2016
Accepted 22 July 2016
Available online 26 July 2016

*Keywords:*
Image database
Image quality assessment
Multiply distorted images
Pair comparison sorting
Nonlinear regression

ABSTRACT

In this paper, we present a new database, the multiply distorted image database (MDID), to evaluate image quality assessment (IQA) metrics on multiply distorted images. The database contains 20 reference images and 1600 distorted images. The latter images are obtained by contamination of the former with multiple distortions of random types and levels, so multiple types of distortions appear in each distorted image. Pair comparison sorting (PCS) is used as a new subjective rating method to evaluate image quality. This method allows subjects to make equal decisions on images whose difference in quality cannot be easily evaluated visually. A total of 192 subjects participated in the subjective rating, in which mean opinion scores and standard deviations were obtained. In IQA research, subjective scores and algorithm predictions are generally related by a nonlinear regression. We further propose a method to initialize the parameters of the nonlinear regression. The experiments of IQA metrics conducted on MDID validate that this database is advisable and challenging.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

### 1.1. Background

Image quality assessment aims to assess the perceptual quality of visual signals in various applications, such as image and video compression (e.g., [1–3]), image denoising (e.g., [4,5]), image reconstruction (e.g., [6,7]), and image synthesis (e.g., [8]). Given that the subjective evaluation of image quality is reliable yet laborious, the research of IQA mostly focuses on objective evaluation. Objective evaluation algorithms can be categorized into three types: full-reference (FR), reduced-reference (RR), and no-reference (NR). Among them, the FR and NR problems have attracted the greatest attention. For the FR problem, the pixel-wise information of reference images is available, and many FR algorithms have been presented, including Peak Signal to Noise Ratio (PSNR), Structural Similarity (SSIM [9]), Visual Signal to Noise Ratio (VSNR [10]), Visual Information Fidelity (VIF [11]), Information-Contented Structural Similarity (IW-SSIM [12]), Feature Similarity (FSIM [13]), Scalable Image Quality Assessment (SIQA [47]), Setr [48] and Gradient Magnitude Similarity Deviation (GMSD [14]). By contrast, the NR algorithms take no advantage of reference images (e.g., Blind Image Quality Indice (BIQI [15]), Blind Reference Image Spatial Quality Evaluator (BRISQUE [16]), and Quality Aware Clustering (QAC [17])).

To evaluate the performance of IQA metrics, various IQA databases have been established and made public (see Table 1). Experiments show that many IQA algorithms [9–17,47,48] have been successfully applied in existing databases [18–22]. Nevertheless, the human visual system (HVS) is highly complex, and we cannot build an accurate model to analyze how humans perform visual perceptions on different objects and scenes. Existing IQA metrics are mostly based on modeling some characteristics of HVS. Hence, many challenges remain unsolved [26], and one of the challenges is Multiple Types of Distortions (MTD). Here, MTD means that images simultaneously contain multiple types of distortions. MTD introduces new problems to the design of IQA algorithms because of compound effects of different distortions. In spite of this, MTD needs to be addressed since real images may be multiply distorted [25] in the capturing, transmitting, and displaying processes. In fact, NR algorithms cannot produce satisfactory results with multiply distorted images [25].

### 1.2. Contributions of our work

To provide the benchmark to IQA metrics in the case of MTD, we establish a new MTD database, the multiply distorted image database, which can be accessed at [49]. MDID differentiates itself

* Corresponding author at: Department of Electronic Engineering/Shenzhen Graduate School, Tsinghua University, China.
*E-mail address:* flying.zhou@163.com (F. Zhou).
[1] Present/permanent address: Room 205B, Building H, Tsinghua Campus, XiLi University Town, Shenzhen, Guangdong Province, China.

http://dx.doi.org/10.1016/j.patcog.2016.07.033
0031-3203/© 2016 Elsevier Ltd. All rights reserved.

**Table 1**
Comparison of MDID database with publically available IQA databases.

| Database | IVC | LIVE | MICT | TID2008 | CSIQ | BID | MDIQ | TID2013 | CID2013 | MDID |
|---|---|---|---|---|---|---|---|---|---|---|
| Type of database | STD | STD | STD | STD | STD | MTD | MTD | MTD | MTD | MTD |
| Ref. images available | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | No | Yes |
| Number of ref. images | 10 | 29 | 14 | 25 | 30 | N/A | 15 | 25 | N/A | 20 |
| Number of dis. images | 185 | 779 | 196 | 1700 | 866 | 585 | 405 | 3000 | 480 | 1600 |
| Distortion types | 4 | 5 | 2 | 17 | 6 | N/A | 3 | 24 | 12–14 | 5 |
| Distortion levels | 5 | 5 | 6 | 4 | 4–5 | N/A | 4 | 5 | N/A | 4 |
| Number of distortions in an image | 1 | 1 | 1 | 1 | 1 | N/A | 2 | 1–2 | N/A | 1–4 |
| Detailed list of distortions | Yes | Yes | Yes | Yes | Yes | No | Yes | Yes | No | Yes |
| Subjective evaluation | DSIS | SS | SS | PC | N/A | SS | SS | PC | ACR-DR | PCR |
| Screen | CRT | 21″CRT | 17″CRT | 19″LCD | 24″LCD | 17″CRT | LCD | 19″LCD | 24″LCD | 19″LCD |
| Illumination | N/A | Indoor | Low | Varied | N/A | N/A | Indoor | Varied | Constant | Constant |
| Viewing distance | 6 Hs | 2–2.5 Hs | 4 Hp | Varied | 70 cm | N/A | 4 Hs | Varied | ~80 cm | ~2 Hs |
| Number of subjects | 15 | 20–29 | 16 | 838 | 25 | 180 | 37 | 971 | 188 | 192 |
| Number of ratings for each image | 15 | 23 | 16 | 33 | 5–7 | 11 | 18–19 | ~30 | 31 | 33–35 |
| Data format | DMOS+$\sigma$ | DMOS | RAW | MOS+$\sigma$ | DMOS+$\sigma$ | RAM | DMOS+$\sigma$ | MOS+$\sigma$ | RAW | MOS+$\sigma$ |
| Range of scores | 1–5 | 0–100 | 1–5 | 0–9 | 0–1 | 0–5 | 0–100 | 0–9 | 0–100 | 0–8 |
| Averaged $\sigma$ | 11.9103 | N/A | 14.7294 | 7.2782 | 7.8000 | N/A | 15.7126 | 7.2126 | N/A | 6.4771 |
| Image format | BMP | BMP | BMP | BMP | PNG | JPG | BMP | BMP | JPG | BMP |

N/A means the information is not available or unknown. '~' means about.
Ref. denotes Reference, ref. means reference, and dis. means distorted.
In the row of Subjective evaluation scheme, DSIS is Double-Stimulus Impairment Scale, SS is Single Stimulus, PC is Pair Comparison, ACR-DR is Absolute Category Rating-Dynamic Reference [39], and PCR stands for Pair Comparison Sorting.
In the index of Viewing distance, Hs represents the height of screen and Hp is the height of picture.
MOS is mean opinion score, DMOS is difference mean opinion score, and $\sigma$ represents the standard deviation (normalized by $\sigma$/max(MOS)$\times$100).
MTD is Multiple Types of Distortions, STD is Single Type of Distortion.

from other databases as follows: (1) Each distorted image has multiple types of distortions. In fact, the number (from 1 to 4) of distortion types and level (from 1 to 4) of distortions for each distorted image are random (see Section 4 for more details). (2) Detailed information of distortions in each distorted image is available. (3) MDID is applicable to both the FR and NR image quality assessment problems.

Furthermore, in the process of subjective evaluation, we find that subjects have some difficulties in comparing two quality-indistinctive images. We propose a new subjective scheme, pair comparison sorting, to address such a problem. PCS allows subjects to make an "equal" decision besides a "larger" or "smaller" decision. The second-to-last row of Table 1 compares the normalized standard deviation of MDID with that of other databases, and the results demonstrate the effectiveness of PCS. In addition, when testing IQA metrics on MDID, we find that the nonlinear regression between subjective scores and IQA predictions may not converge or fall into a local optimal state. Accordingly, we propose a method to estimate the initial parameters of nonlinear regression.

The main contributions of this work are as follows: (1) we establish an MTD image database, in which each distorted image has multiple types of distortions; (2) we propose PCS to perform a more reasonable subjective evaluation of image quality; (3) we propose a new method to initialize the parameters of the regression between IQA predictions and mean opinion score (MOS) and thus avoid the non-convergence or the local optimal problems.

The rest of the paper is organized as follows. Section 2 presents the review of existing IQA databases. Section 3 presents the selection and analysis of reference images. Section 4 discusses the generation of distorted images in detail. In Section 5, we describe our PCS scheme and data processing, as well as conducting some discussion about the evaluation results. The experimental results of various IQA metrics conducted on MDID are presented in Section 6. Finally, Section 7 concludes the paper.

## 2. Review of existing IQA databases

In the field of IQA, many databases have been created, including Image and Video Communication (IVC [18]), Laboratory for Image and Video Engineering (LIVE [19]), Multimedia Information and Communication Technology (MICT [20]), Tampere Image Database 2008 (TID2008 [21]), and Categorical Subjective Image Quality (CSIQ [22]). Table 1 presents the detailed information of these IQA databases. IVC has 4 distortion types and 5 distortion levels. However, the number of distorted images is only 185, which is not enough to assess IQA metrics. LIVE is one of the most widely used databases in IQA research, it has 779 distorted images with 5 distortion types and 5 distortion levels. TID2008 is also widely used to evaluate IQA metrics. In TID2008, there are totally 1700 distorted images with 17 distortion types. It is necessary to note that although many types of distortions are introduced in above databases, each distorted image contains only single type of distortion. Recently, several image databases that focus on images with two or more types of distortions have been developed such as Multiply Distorted Image Quality (MDIQ [23]) and Tampere Image Database 2013 (TID2013 [24])). Moreover, based on camera images, Camera Image Database (CID2013 [25]) aims at evaluating real-world camera-captured images.

As shown in Table 1, most of existing IQA databases, such as IVC, MICT, LIVE, TID2008 are Single Type of Distortion (STD) image databases. Blur Image Database (BID [27]), MDIQ, TID2013, and CID2013 start to pay attention to multiply distorted images. However, the subjective evaluation of BID is performed only in terms of the effects of blurriness. Although MDIQ and TID2013 demonstrate some improvements over LIVE and TID2008, only one or two types of distortions are contained in each distorted image, which still cannot sufficiently describe the distortions in real images. CID2013 is a camera-based image database. However, given that the distortion types and levels cannot be listed in details, the effects of distortions cannot be detected.

## 3. Selection of reference images

A reference image is regarded as the source image in IQA databases. To our knowledge, different image contents may impose varying effects on IQA [28]. Therefore, the selection of reference images is highly important for an IQA database. The Video Quality

Download English Version:

https://daneshyari.com/en/article/533080

Download Persian Version:

https://daneshyari.com/article/533080

Daneshyari.com