J. Vis. Commun. Image R. 30 (2015) 1–9

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

MCL-V: A streaming video quality assessment database ☆

Joe Yuchieh Lin [a], Rui Song [b,*], Chi-Hao Wu [a], TsungJung Liu [a], Haiqiang Wang [a], C.-C. Jay Kuo [a]

[a] University of Southern California, Ming Hsieh Department of Electrical Engineering, 3740 McClintock Avenue, Los Angeles, CA, United States
[b] Xidian University, 710071 P.O. Box 103, 2nd Taibai South Road, Xi'an, Shaanxi, China

ARTICLE INFO

Article history:
Received 9 June 2014
Accepted 18 February 2015
Available online 5 March 2015

Keywords:
Video quality
Quality assessment
Quality database
Video streaming
Compression distortion
Video re-sizing
Video coding
Perceptual quality

ABSTRACT

A high-definition video quality assessment (VQA) database that captures two typical video distortion types in video services (namely, "compression" and "compression followed by scaling") is presented in this work. The VQA database, called MCL-V, contains 12 source video clips and 96 distorted video clips with subjective assessment scores. The source video clips are selected from a large pool of public-domain high-definition (HD) video sequences with representative and diversified contents. Both distortion types are perceptually adjusted to yield distinguishable distortion levels. An improved pairwise comparison method is adopted for subjective evaluation to save evaluation time. Several existing image and video quality assessment (IQA and VQA) algorithms are evaluated against the MCL-V database. The MCL-V database is publicly accessible in the link – http://mcl.usc.edu/mcl-v-database/ to facilitate future video quality assessment research of the community.

1. Introduction

The high-definition video broadcasting and streaming services are blooming nowadays. Consumers can enjoy on-demand video services from Netflix, Hulu or Amazon, and watching high-definition (HD) programs becomes the mainstream for video content consumption. According to the report in [1], more than half of US population watches on-line movies or dramas. Specifically, the viewers have increased from 37% in 2010 to 51% in 2013. The watched video programs vary in bit rates and resolutions due to the available bandwidth of their networks. Different sizes of video are transmitted at lower bit rates and up-scaled for display on HDTV (e.g., playing a 720p movie on the 1080p screen). This is common in people's daily life [2], but users' video quality of experience on HD video has not yet been extensively studied in the past.

There are quite a few video quality assessment databases available to the public [3–28]. They were however limited in the following areas [29,30]. First, the source video set is not representative or diversified enough. For example, they do not contain dark scenes, sports scenes, traditional cartoon, and computer animation. The lack of these contents will not provide an extensive evaluation of viewers' experience. Second, the video resolution is low. The resolution of sequences in all VQA databases except [3,8,20,26,28] are lower than 1920×1080. Third, the distortion is not complete for the target application. For example, all above-mentioned VQA databases except [20,10,11] do not cover video up-scaling, which is encountered frequently in our daily life. Although the work in [20] includes practical distortion types, it has only three video sources. Being motivated by these observations, we build a new VQA database called MCL-V to address the shortcomings of existing VQA databases. The MCL-V database provides 12 source video clips, 96 distorted video clips and their associated mean opinion scores (MOS). In this paper, we will elaborate on the methodology of building MCL-V including collecting suitable video sources, generating distortions and conducting subjective evaluation.

One key issue in our design is to choose an appropriate subjective test procedure to collect opinion scores. Several subjective test methodologies have been recommended in VQEG [25,31] and ITU [32,33] as shown in Table 1. Since the precision of the final MOS is not improved by adopting the continuous scale [34,35], the discrete scale is adopted in this work for user friendliness. Furthermore, we use an improved pairwise comparison method to make the final MOS more stable and meaningful.

The rest of this paper is organized as follows. Section 2 describes ways to choose representative and diversified reference sequences, to generate practical distortion types and to determine the reasonable distortion levels. Section 3 presents an improved pairwise comparison method for subjective evaluation and elaborates on the process of collecting and normalizing opinion scores in the subjective test. We study the MOS values and analyze the performances of several existing IQA and VQA metrics against the MCL-V database in Section 4. Finally, concluding remarks are

☆ This paper has been recommended for acceptance by Prof. M.T. Sun.

* Corresponding author.
E-mail address: rsong@xidian.edu.cn (R. Song).

http://dx.doi.org/10.1016/j.jvcir.2015.02.012
1047-3203/

given in Section 5. The whole database is publicly available on the USC Media Communication Lab website http://mcl.usc.edu/mcl-v-database/.

2. Construction of MCL-V database

2.1. Source video selection

We selected 12 uncompressed HD video clips as the source sequences. Some sequences are originally in YUV444p or YUV422p, and we converted them into YUV420p using FFMpeg [36] to make all videos included in the MCL-V database be YUV420p at a fixed resolution of progressive 1920×1080. The frame rates of the sequences range from 24 fps to 30 fps, and the length of each video is 6 s. Fig. 1 shows all reference videos with a single frame.

The selected sequences are freely available from several sources, including HEVC test sequences [37], TUM dataset [38], CDVL [39], and others [40–42]. They were professionally acquired and recorded in digital form. We select some of them to construct the MCL-V database based on the following two criteria.

First, some prior databases [13,23] contain scenes that are not representative in video applications. For example, there are video clips with a close view on the water surface or the blue sky in the LIVE database [23]. These sequences were used for video coding performance test since they contain specific contents which are difficult to encode. However, they are not common scenes in movies or dramas. We prefer more representative scenes since they can better reveal human visual experience.

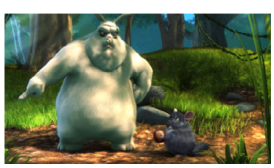
Second, the database should have sufficient diversity in terms of several characteristics. We list various characteristics for diversity consideration in Table 2. They are categorized into three groups: (1) high-level video genres, (2) mid-level video semantics and (3) low-level video features. We aim to make the database cover a wide range of characteristics given in the table.

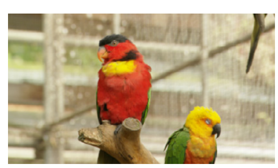
For video genres, we take several new genre types such as animation and sports into account. These video genres have different characteristics from others. For instance, cartoons scenes contain clear edges and simple color components while sports scenes contain fast moving objects with simple background. These videos are commonly seen in applications and should be included in the MCL-V database.

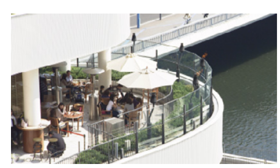
For video semantics, we consider factors that will have a great impact on human visual perception. For example, while other databases usually do not include video scenes with a close-up face, we take this feature into consideration since it is typical in many dramas. In addition, the human face is typically a region of visual salience which attracts human attention.

Table 1
Classification of subjective testing methods.

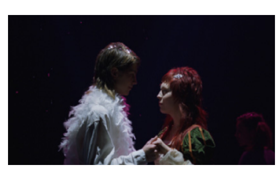
	Discrete scale	Continuous scale
Single Stimulus	Absolute Category Rating (ACR) [32]	Single Stimulus Continuous Quality Evaluation (SSCQE) [33]
Double Stimulus	Degradation Category Rating (DCR) [32]	Double Stimulus Continuous Quality Scale (DSCQS) [33]
	Comparison Category Rating (CCR) [33]	


(a) Big Buck Bunny (BB)

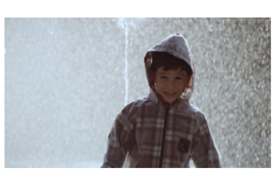

(b) Birds in Cage (BC)

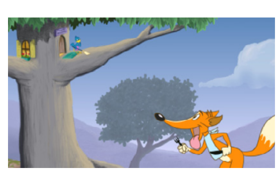

(c) BQ Terrace (BQ)


(d) Crowd Run (CR)

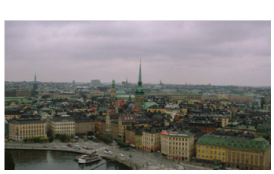

(e) Dance Kiss (DK)

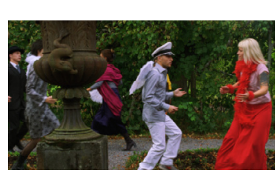

(f) El Fuente A (EA)


(g) El Fuente B (EB)


(h) Fox Bird (FB)


(i) Kimono (KM)


(j) Old Town Cross (OT)


(k) Seeking (SK)

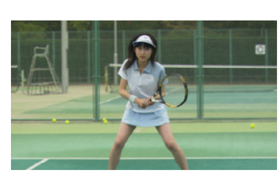

(l) Tennis (TN)

Fig. 1. Selected source video sequences.