

IMAGE QUALITY IN THE ERA OF PERSONAL AND SOCIAL MEDIA

Stefan Winkler

Advanced Digital Sciences Center (ADSC), Singapore

ABSTRACT

With the wide-spread use of digital cameras, imaging software, photo-sharing sites, social networks, and other related technologies, media production and consumption patterns have become much more multi-faceted and complex than they used to be. User-generated content in particular has grown tremendously. As a result, the popular concept of “Quality of Experience” (QoE) and quality assessment (QA) must also be looked at from a different angle. This paper contrasts some of the traditional quality assessment approaches with new perspectives on quality for personal and social media.

1. BACKGROUND

Image quality assessment (QA) dates back to the 1970’s, with the beginnings of studies of the visual cortex, psychophysics, and vision modeling. Algorithms based on models of the human visual system now compete with more pragmatic, image- or feature-based methods – Lin and Kuo provide a recent survey [1]. Video quality assessment has seen a similar, albeit shorter development [2]. Most of the work has focused on fidelity (i.e. how closely the processed image/video resembles some reference).

Ubiquitous and affordable digital cameras (and other image capture devices) now enable users to take pictures and videos everywhere and anytime. This has led to an explosion of the amount of picture material both amateurs and professionals have to work with. As an example, Flickr, a popular online photo sharing site, reported the 5th billion picture upload in September 2010; about 15 million pictures are currently uploaded to Flickr every day.

Methods for automatic QA have focused mainly on traditional media and processing steps, i.e. the compression, transmission, and enhancement of images and video. In many cases, there exists an explicit reference (e.g. the source image/video), which passes through the system under test and undergoes certain changes (e.g. loss of fidelity, compression artifacts, packet losses, noise removal, etc.)

It is very much a linear process, in that there is generally a single (high-quality) source which undergoes various processing steps such as those mentioned above,

which deteriorate (or sometimes enhance) the quality of the material.

For such high-value, professionally produced content, quality assessment in the production process is typically done manually. Afterwards, the content is prepared for and distributed to many paying consumers via different ways (cinema, broadcast TV, etc.), which makes automatic QA interesting at that stage.

For user-generated content, the criteria for quality assessment and enhancement are not only image- or content-specific (e.g. impairments, scene composition), but also user-centric (i.e. what is most relevant to the user in this collection), contrary to the more traditional QA approaches.

Automated quality assessment for user-generated content becomes interesting in the production process already, for a number of reasons:

- Even if QA could be done manually by the user, it would be too time-consuming for most. Besides, the average user is not a professional and can use guidance for producing high-quality content.
- Quality becomes a much more personal concept. Traditional media is designed for a wide audience, therefore the “average user” and “mean opinion score” (MOS) are the benchmark. For user-generated content, it is mainly the user himself who matters, and (perhaps to a lesser extent) the circle of people he/she shares the content with.

Using the example of picture collections, I will discuss these aspects in more detail.

2. PICTURE COLLECTIONS

The most common type of user-generated content today is still pictures. These collections typically comprise all pictures taken during an event of holiday/trip, possibly from multiple users and/or devices, by images shared in social networks and online sites, and third-party repositories. Devices could include single-lens reflex (SLR) cameras, point-and-shoot cameras, camera phones, etc.

The role of QA here is primarily the selection of the “best” pictures from the collection, because users will typically want to reduce the number of photos before creating an album for personal use or for sharing.

Let us now look at the criteria users apply to select a subset. We can roughly distinguish between two tasks:

1. The selection of a photo from a group of similar ones (typically multiple shots of the same scene).
 2. The selection of scenes for the album, often with the purpose of telling a story, or sharing an experience.
- Users would generally apply different criteria of “quality” in these two steps; however, the aspect of personalization is important in both.

2.1. Selection

Since memory is cheap, it has become standard practice to take multiple pictures of the same scene. Selecting the best picture from such a group typically includes evaluating the following aspects [3]:

- Lighting conditions (e.g. flash vs. no flash), camera exposure, white balancing.
- Framing and perspective of the scene (e.g. an object may be cut off).
- Postures, actions and faces of the important people in the scene (e.g. how many are looking at the camera).
- Image quality (e.g. some pictures may be more blurred, noisy, compressed than others, in particular if they come from different devices).

Traditionally, the quality of a picture is assessed by comparing it to a reference with the same content, but without impairments (full-reference comparison), or on its own (no-reference). In this case, there is no reference image, so traditional full-reference methods do not apply. No-reference methods could be used in principle, but generally work best along a single impairment dimension (e.g. quantization or blur) of the same image.

The problem here is more general and revolves around comparing pictures that have similar, related (but not identical) content, different quality/ impairment dimensions and levels (one image may be blurred, while another may be underexposed), and choosing the best among them. This also requires a good understanding of the effects different impairment dimensions have on perception.

2.2. Summarization

Selecting the most representative pictures from a set is similar to storytelling or summarization. The key component of this task is identifying which scenes the user considers important for the “story”. There may not be a unique set of representative pictures for a collection, because of the large number of possible subsets and different possible themes.

Parameters and criteria that people use in their choice of pictures from a collection have been studied before; examples include specific people, variety of places, or general image quality [3]. These parameters can be used to guide the selection process.

In automating the process, it can be helpful that some events such as weddings or birthdays (within certain cultural boundaries) have a specific sequence of events

and a number of important “milestones” that need to be included in the picture selection process.

Finally, it can be desirable to find pictures that may not be part of the initial set, but that are relevant to the story and can be sourced from external collections. Examples of this are a map of the places visited, or a better picture of a popular sight if the ones present in the collection are not satisfactory.

2.3. Personalization

For user-generated content, the personal and social aspects are much more important than for traditional professional content, because user-generated content is often only meaningful to a small group of people – typically the user, his/her family and friends, etc.

Indeed, the user aspect of the picture selection, enhancement, and organization process has not received much attention so far. However, due to the subjectivity involved in these steps, taking a user-centric approach is essential.

Personalization also implies that it is not the criteria of the “average” user (as exemplified by the traditional MOS) that matter, but the criteria of the specific user, which is quite different to the way the topic is usually approached.

For example, photos in a family album are meaningful to family members, but not to outsiders. Therefore, the selection of certain shots from such a collection by a random person is unlikely to be meaningful or relevant. Likewise, a user may have certain preferences in terms of perspective, lighting, color, enhancements, subjects, expressions, poses, and so on. Any QA system for such content should thus be able to offer personalized suggestions of any image selection or enhancement steps according to the user’s individual taste and preferences.

ACKNOWLEDGEMENTS

This research is supported by the Advanced Digital Sciences Center (ADSC) under a grant from the Agency for Science, Technology and Research of Singapore (A*STAR).

REFERENCES

- [1] W. Lin, C.-C. J. Kuo: “Perceptual visual quality metrics: A survey.” *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297-312, May 2011.
- [2] S. Winkler, P. Mohandas: “The evolution of video quality measurement: From PSNR to hybrid metrics.” *IEEE Trans. Broadcasting*, vol. 54, no. 3, pp. 660-668, Sept. 2008.
- [3] A. E. Savakis, S. C. Etz, A. Loui: “Evaluation of image appeal in consumer photography.” *Proc. SPIE Human Vision and Electronic Imaging*, vol. 3959, San Jose, CA, Jan. 2000.