ANALYSIS ON PERCEPTUAL-BASED QUALITY METRICS FOR IMAGE AND VIDEO SERVICES

Analysis on Perceptual-Based Quality Metrics for Image and Video Services

Ranjeet Yadav

Research Scholar, Sathyabama University, Chennai, Tamilnadu

Abstract – We will focus on the design of objective metrics for visual quality assessment in wireless image and video communication. The aim is to quantify the end-to-end distortions induced during transmission and relate them to quality degradations as perceived by the end-user. These metrics may then replace the conventional link layer metrics to allow for precise perceptual quality monitoring. The application of perceptual image and video quality assessment in a communication context, as we consider it throughout this study, is illustrated.

------------------------◆----------------------------

INTRODUCTION

In the current scenario, the received image or video may from artifacts, and consequently quality degradations, from both the source encoder and the error-prone wire-less channel. The impact of the source coding artifacts is somewhat easier to predict since for different codec's certain artifacts can be expected. On the other hand, the time variant nature of the fading channel makes the range of artifacts in the received signal much more unpredictable. With JPEG source encoding and a Rayleigh °at fading wireless channel with additive white Gaussian noise (AWGN). The symbiosis of this setup resulted in a wide range of different artifacts, hence, substantially complicating the assessment of the artifacts and the related visual quality. The additional shaded boxes in comprise of the necessary components to facilitate perceptual quality assessment, as we propose it in this study. The blocks surrounded by dashed lines indicate optional parts of the quality assessment which are applied when reference features are extracted from the transmitted image to support the quality assessment, hence, facilitating RR quality assessment. On the other hand, if these blocks are omitted then quality assessment is performed solely on the received image, thus, following the NR approach. However, as we aim to quantify quality degradations induced during transmission, we need some reference information from the transmitted image or video frame. Therefore, we incorporate the reference feature extraction into our metric design to establish RR objective quality metrics. In this case, the reference features may be concatenated to the transmitted image or video frame to be available at the receiver for quality assessment.

REVIEW OF LITERATURE:

The number of bits associated with the reference features the overhead for each of the images it is concatenated with and accordingly it is desired to be kept small. In particular the NHIQM metric, as discussed earlier, comprises of only one single value as additional overhead. Extensions and variations to this metric, as proposed in this study, may have slightly larger overhead but allow for tracking of each of the single features included in the metric. This may provide further insights into the cause of induced artifacts during transmission. In order to avoid additional overhead one may alternatively embed the reference features into the image or video frame using data hiding techniques [14]. Due to the limited capacity of these techniques, however, too large reference information may cause visible distortions in the image. Consequently, the aim to keep the number of reference features small remains also when applying these techniques. The metrics developed in this study are designed with respect to two goals. Firstly, the extracted features need to cover the broad range and precisely quantify the appearance of artifacts as induced in the images by both the loss source encoding and the error prone channel. Therefore, feature metrics were selected according to the artifacts that may be observed in images which are distorted due to transmission over a wireless link. Secondly, the objectively measured artifacts need to be related to quality degradations as subjectively perceived by a human observer. This latter goal is followed by incorporating several characteristics of the HVS into the metric design to allow for superior quality prediction performance as compared to metrics that purely measure similarity between images [1]. In order to further support the design of the objective metrics we have conducted subjective image quality experiments at the Western Australian

Telecommunications Research Institute (WATRI) in Perth, Australia, and at the Blekinge Institute of Technology (BTH) in Ronneby, Sweden. The mean opinion scores (MOS) obtained from these experiments allowed us to relate the different measures incorporated in the objective metrics to subjectively perceived visual quality. The MOS further enabled evaluation of the quality prediction performance of the metrics on both a set of training images that were used for the metric design and a set of validation images that were unknown during metric training. Unlike previously proposed HVS based quality metrics [1-3] that incorporate a large number of HVS properties, we focus on a few simple approximations of HVS characteristics that have been shown to be essential for the visual perception of quality. Specifically, the basis for the metric designs is motivated by the phenomenon that the HVS is adapted to extraction of structural information [4]. Thus, a number of structural features are extracted that accurately quantify the artifacts observed in wireless image and video communication. An additional weighting then controls the impact of each feature on the overall metric. The weights are derived in relation to the MOS from the experiments and thus account for the perceptual relevance of each of the artifacts. Additional HVS characteristics, such as, multiple-scale processing and regional attention will be shown to further enhance the metrics quality prediction performance. The latter characteristic has been supported by an additional subjective experiment that we conducted at BTH to identify regions-of-interest in the set of reference images and thus allow for implementation of region-selectivity in the metric design. To account for non-linear quality processing in the HVS, all metrics are in a last step subject to an exponential mapping. The mapping translates the metric values into so-called predicted MOS which aim to measure the quality as it would be rated by a human observer.

Perceptual-based quality metrics for image and video services:

This part consists of a survey of contemporary image and video quality metrics. The work is a result of an intensive literature research which has been carried out to investigate previously conducted image and video quality research and also to identify open issues that need to be addressed. Only few reviews and surveys about image and video quality metrics have been published in the past [5-8]. In contrary to these related works, this survey concentrates on metrics that aim to predict quality as perceived by a human observer and further belong to the class of NR and RR metrics. The latter property enables quality prediction of a distorted image/video without a corresponding reference image/video to be available. Hence, these metrics are readily applicable in wireless and wire line image and video communication, where the original image or video is unavailable for quality assessment at the receiver. The survey provides a detailed classification of the available quality assessment

methods. It further discusses the advantages and drawbacks of a broad range of available NR and RR metrics that have been proposed in the past. Two extensive tables provide direct overviews with the aim for the reader to easily identify the appropriate metric for a given task. The tables provide information about the artifacts (blocking, blur, etc.), the domain (spatial, frequency, etc.), the source codecs (JPEG, MPEG, etc.), and the typical image/frame size which the metrics have been designed for. Finally, some open issues in image and video quality assessment are outlined in the conclusions.

Reduced-reference metric design for objective perceptual quality assessment in wireless imaging:

In this part a RR metric, NHIQM, is proposed for wireless imaging quality assessment. The metric is based on the work conducted earlier in [9-10]. The various extensions to the previous work can be summarized as follows:

- **Extreme value feature normalization:**

The structural feature algorithms included in the objective metric are implemented according to algorithms as outlined in different publications [11-12]. Consequently, the ranges of the different features were strongly varying. In this work, we therefore introduce extreme value normalization [13] in order for the features to fall into a defined interval.

- **Perceptual relevance weighted Lp-norm for feature pooling:**

An alternative feature pooling based on a perceptual relevance weighted Lp-norm [13] is proposed. The resulting metric provides similar quality prediction performance as NHIQM while at the same time allowing tracking the structural degradations independently for each of the features. Thus, insight into the artifacts induced during transmission may be gained using this feature pooling.

- **Statistical analysis of subjective experiments and objective features:**

An in-depth statistical analysis is provided for the subjective experiments that we conducted in two independent laboratories. The analysis reveals the relevance of the subjective scores obtained in the experiments. In addition, a detailed analysis of the objective feature scores from the experiment test images is discussed revealing insight into the artifacts that were objectively quantized by the feature metrics.

- **Metric training and validation:**

The concept of metric training and validation has further been introduced to the work to verify that the metric design does not result in to the training data

but rather allows for good generalization to unknown images.

- **Motivation for a non-linear mapping function:**

Using the training and validation approach we further motivate the use of an exponential prediction function to account for the non-linear processing in the HVS. Other prediction functions can be excluded due to inferior goodness of measures, visual inspection, and overrating on the training set of images.

- **Comparison to state of the art visual quality metrics:**

State of the art visual quality metrics are considered in this work for comparison of quality prediction accuracy, prediction monotonicity, and prediction consistency [13] on both the training and the validation set of images. The evaluation reveals the superior quality prediction performance of NHIQM with respect to all three criteria.

CONCLUSION:

The work as it has developed until today is composed of different methods that were successfully applied to design and improve objective metrics that accurately predict visual quality as it would be perceived by a human observer. The focus was thus far set on spatial feature extraction and the related quantization of artifacts as observed in the spatial domain. This approach shall in future be extended to temporal feature extraction, hence, accounting for temporal artifacts and masking effects that may occur in wireless video sequences.

REFERENCES:

[1]   T. Welch, "A technique for high-performance data compression", Computer Magazine, vol.17 (6), pp. 8-19, 1984

[2]   T. Boutell, "PNG (Portable Network Graphics) specification", ftp://ftp.uu.net/graphics/png/documents/

[3]   P. Deutsch, "DEFLATE compressed data format specification", RFC1951, http://www.faqs.org/rfcs/rfc1951.html, 1991.

[4]   J. Ziv, A. Lempel, "A universal algorithm for sequential data compression", IEEE Trans. on Information Theory, vol. 23(6), pp. 337-343, 1977

[5]   International Telegraph and Telephone Consultative Committee (CCITT), "Facsimile Coding Schemes and Coding Control Functions for Group 4 Facsimile Apparatus," Recommendation T.6, 1984.

[6]   TIFF Revision 6.0, Adobe Developers Association, Adobe System Incorporated. http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf, 1992.

[7]   ITU-T recommendation T.82, "Information technology – coded representation of picture and audio information – progressive bi-level image compression", 1993.

[8]   J. J. Rissanen, Langdon G. G., "Arithmetic coding", IBM Journal of Research, Development 23: 146-162, 1979.

[9]   W. Pennebaker, J. Mitchell, "Probability estimation for the Q-coder", IBM Journal of Research, Development 32(6), pp. 737-759, 1988.

[10]   G. G. Langdon and J. Rissanen, "Compression of black-white images with arithmetic coding," IEEE Trans. on Communications, vol. 29, no. 6, pp. 858-867, June 1981.

[11]   J. Rissanen, "A universal data compression system", IEEE Trans. on Information Theory, vol. 29 (5), pp. 656–664, 1983.

[12]   ITU-T recommendation T.88, "Information technology – coded representation of picture and audio information – lossy/lossless coding of bi-level images", 2000.

[13]   P. Howard, F. Konssentini, B. Martins, S. Forchhammer, W. Rucklidge, "The emerging JBIG2 standard", IEEE Trans. On Circuits and Systems for video technology, vol. 8(7), pp. 838-848, 1998.

Ranjeet Yadav