A No-Reference Metric for Perceived Ringing Artifacts in Images

Hantao Liu, *Student Member, IEEE,* Nick Klomp, and Ingrid Heynderickx

Abstract—**A novel no-reference metric that can automatically quantify ringing annoyance in compressed images is presented. In the first step a recently proposed ringing region detection method extracts the regions which are likely to be impaired by ringing artifacts. To quantify ringing annoyance in these detected regions, the visibility of ringing artifacts is estimated, and is compared to the activity of the corresponding local background. The local annoyance score calculated for each individual ringing region is averaged over all ringing regions to yield a ringing annoyance score for the whole image. A psychovisual experiment is carried out to measure ringing annoyance subjectively and to validate the proposed metric. The performance of our metric is compared to existing alternatives in literature and shows to be highly consistent with subjective data.**

Index Terms—**Human vision model, image quality assessment, objective metric, ringing artifact annoyance.**

I. INTRODUCTION

OBJECTIVE metrics have the aim to automatically provide a quantitative measure for image quality aspects, and to eventually serve as computational alternative for expensive image quality assessments by human observers. They are of fundamental importance to a broad range of applications, such as the optimization of digital imaging systems, benchmarking of image and video coding, and quality monitoring and control in displays [1]. They are generally classified into full-reference (FR) metrics and no-reference (NR) metrics, depending on the use of the original image or video. FR metrics are based on measuring the similarity or fidelity between the distorted image and its original version, which is considered as a distortion-free reference. The most widely used FR metrics are mean squared error and peak signal-to-noise ratio. These metrics, however, have long been criticized for their poor correlation with perceived image quality [1]. A lot of research effort is devoted to the development of FR metrics that can reflect the way human beings perceive image quality [2]. Improved alternatives of FR metrics include the structural similarity index [3] and the visual information

fidelity index [4]. Since FR metrics require the access to the original, which is however not always available in real-world applications, they are usually employed as tools for in-lab testing of image and video processing algorithms. NR metrics instead are more practical because the quality prediction is based on the distorted image only. However, designing NR metrics is still an academic challenge mainly due to the limited understanding of the human visual system (HVS).

In the last decades, considerable progress on the development of NR metrics is made, and some successful methods are reported in the literature [5]–[19]. In [5], natural scene statistics are used to blindly measure the quality of images compressed by JPEG2000. The approach in [5] relies on the assumption that typical natural images exhibit strong statistical regularities, and therefore, reside in a tiny area of the space containing all possible images. Based on this assumption it quantifies image quality by detecting variations in statistical image features in the wavelet domain. In [6] and [7], NR image quality assessment is formulated as a machine learning problem, in which the HVS is treated as a black box whose input–output relationship, such as the one between image characteristics and the quality rating, is to be learned. After appropriate training with subjective data, these models proved to be able to consistently predict the perceived quality of JPEG compressed or otherwise distorted images.

A large number of NR metrics, proposed, e.g., in [8]–[19], are based on directly measuring a specific type of artifact created by a specific image distortion process, such as blur caused by acquisition systems, sensor noise, and compression artifacts. In such a scenario, the design of the NR metric can make use of the specific characteristics of the artifact, and therefore, generally obtains a higher reliability with perceived quality degradation [1]. Fortunately, in many practical applications, the distortion processes involved are known, and thus, the design of specific NR metrics turns out to be much more realistic and useful. They can, for example, be combined to predict the overall perceived quality. Various examples of this approach are given in literature. A blockiness metric (see, e.g., [8]–[11]) can be combined with a flatness metric (see, e.g., [12] and [13]) to evaluate the quality of images or video after block-based compression. A ringing metric and a blur metric are often combined to assess the image quality of wavelet-based compression (see, e.g., [14]–[16]). In [17] and [18], multiple artifact metrics are adopted to predict the overall quality of still images or video. In addition to assessing the overall image quality, these specific artifact metrics individually are

beneficial for optimizing real-time digital imaging systems [20]–[22]. In the video chain of current television (TV)-sets, various NR metrics, which quantify the quality of the incoming video based on the occurrence of individual artifacts, are used to adapt the parameter settings of the video enhancement algorithms accordingly (see, e.g., [23] and [24]). To optimize the performance of both applications mentioned above, reliably modeling specific types of artifacts has clear added value.

Since the widespread use of compression, research on NR metrics is mainly dedicated to compression artifacts and transmission errors [25]. Especially, the blocking artifact, which is one of the most annoying artifacts introduced by block-based compression algorithms [26], such as JPEG or MPEG/H.263, got a lot of attention. Another compression artifact, especially visible at relatively high-bit rates of block-based compression [21], [26], but also in wavelet compression [27], is ringing. Unlike the blocking artifact, whose spatial location is very regular and thus easily predictable, the location of ringing is edge dependent, and as such also image content dependent. This makes the task of quantifying ringing annoyance much more difficult. In this paper, we present our recent efforts to develop a NR ringing metric, validate its performance using a subjective study of ringing annoyance in JPEG compressed images, and compare its performance against existing ringing metrics. Before discussing our approach (Section III) and its performance (Sections IV and V), a more extended explanation of the occurrence and visibility of ringing, and an overview of existing ringing metrics are given in Section II.

II. BACKGROUND

A. Perceived Ringing Artifacts

1) *Physical Structure:* Current image and video coding techniques are based on lossy data compression, which contains an inherent irreversible information loss. This loss is due to coarse quantization of the image's representation in the frequency domain. The loss within a certain spectral band of the signal in the transform domain reveals itself most prominently at those spatial locations where the contribution from this spectral band to the overall signal power is significant (see [26], [27], and [38]). Since the high-frequency components play a significant role in the representation of an edge, coarse quantization in this frequency range (i.e., truncation of the high-frequency transform coefficients) consequently results in apparent irregularities around edges in the spatial domain, which are usually referred to as ringing artifacts. More specifically, ringing artifacts manifest themselves in the form of ripples or oscillations around high-contrast edges in compressed images. They can range from imperceptible to very annoying, depending on the data source, target bit rate, or underlying compression scheme [38]. As an example, Fig. 1 illustrates ringing artifacts induced by JPEG compression on a natural image.

The occurrence of ringing spreads out to a finite region surrounding the edges, depending on the specific implementation of the coding technique. For example, in discrete cosine transform (DCT) coding ringing appears outwards from the edge up to the encompassing block's boundary [26]. An
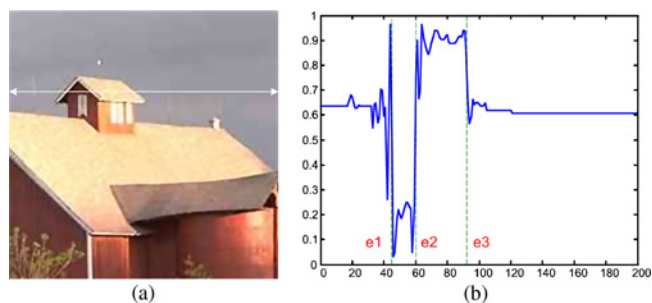

Fig. 1. Illustration of ringing artifacts. (a) Natural image compressed with JPEG (MATLAB's *imwrite* function with "quality" of 30). (b) Gray-scale intensity profile along one row of the compressed image [indicated by the solid double arrowhead line in (a)]. Dashed lines "e1," "e2," and "e3" indicate the position of the sharp intensity transitions (i.e., edges) along that arrow. Ringing can be perceived as fluctuations in the gray-scale values around the edges at "e1," "e2," and "e3," while the image content here should be uniform.

example of how to calculate the extent of the ringing region in a particular codecs is given in [38]. In addition to the edge location dependency, the behavior of ringing also depends on the strength of the edges. It is found in [14], [29], and [38] that, over a wide range of compression ratios, the variance of the ringing artifacts is proportional to the contrast of the associated edge. These important findings have great potential in the design of a reliable ringing metric, and therefore, are explicitly adopted in our algorithm.

2) *Masking of the HVS:* Taking into account the way the HVS perceives artifacts, while removing perceptual redundancies, can be greatly beneficial for matching objective artifact measurement to the human perception of artifacts [39]. Masking designates the reduction in the visibility of one stimulus due to the simultaneous presence of another, and it is strongest when both stimuli have the same or similar frequency, orientation, and location [41]. It is basically due to the limitations in sensitivity of a certain cell or neuron at the retina in relation to the activity of its surrounding cells and neurons. There are two fundamental visual masking effects highly relevant to the perception of ringing artifacts [28]–[31]. The first one is luminance masking, which refers to the effect that the visibility of a distortion (such as ringing) is maximum for medium background intensity, and it is reduced when the distortion occurs against a very low or very high intensity background [40]. This masking phenomenon happens because of the brightness sensitivity of the HVS, where the average brightness of the surrounding background alters the visibility threshold of a distortion [42]. The second masking effect is texture masking, which refers to the observation that a distortion (such as ringing) is more visible in homogenous areas than in textured or detailed areas [40]. In textured image regions, small variations in the texture are masked by the macro properties of genuine high-frequency details, and therefore, are not perceived by the HVS [38]. The effect of luminance and texture masking on ringing artifacts is illustrated in Figs. 2 and 3, respectively.

B. Existing Ringing Metrics

Until recently, only a limited amount of research effort was devoted to the development of a ringing metric. Some of these
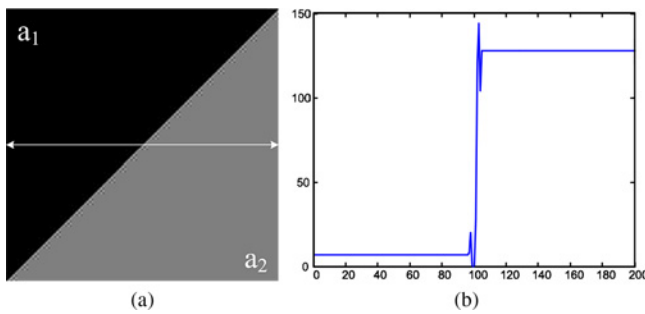
Fig. 2. Example of luminance masking on ringing artifacts. (a) Image patch compressed with JPEG (MATLAB's *imwrite* function with "quality" of 30). (b) Pixel intensity profile along one row of the compressed image patch [indicated by the solid double arrowhead line in (a)]. Original image includes two adjacent parts with different gray-scale levels (i.e., 5 for "a1" and 127 for "a2"). Note that although both sides of a step edge exhibit ringing artifacts, the visibility of ringing differs.
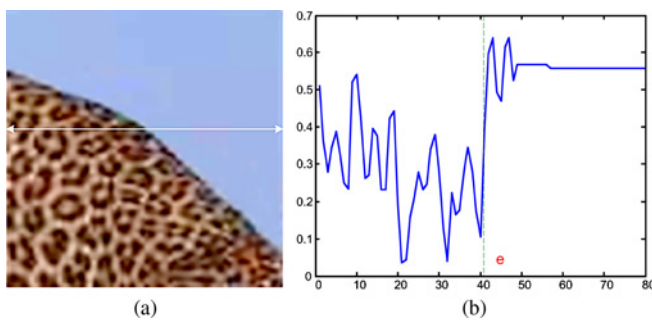

Fig. 3. Example of texture masking on ringing artifacts. (a) Image patch extracted from a JPEG compressed image of bit rate 0.59 bits per pixel (b/p). (b) Pixel intensity profile along one row of the compressed image patch [indicated by the solid double arrowhead line in (a)]. Dashed line "e" indicates the object boundary edge. Note that although both sides of the edge at "e" exhibit ringing artifacts, the visibility of ringing differs.

metrics are FR, others NR. A FR approach presented in [14] starts from finding important edges in the original image (noise and insignificant edges are removed by applying a threshold to the Sobel gradient image), and then measures ringing around each edge by calculating the difference between the processed image and the reference. Since this metric needs the original image, it has its limitations, e.g., for the application in a TV chain. The NR ringing metric, proposed in [17], performs a anisotropic diffusion on the image and measures the noise spectrum filtered out by the anisotropic diffusion process. The basic idea behind this metric is that due to the effectiveness of anisotropic diffusion on deringing, the artifacts would be mostly assimilated into the spectrum of the filtered noise. The NR ringing metric described in [16] indentifies the ringing regions around strong edges in the compressed image, and defines ringing as the ratio of the activity in middle low over middle high frequencies in these ringing regions. An obvious shortcoming of the metrics defined in [14], [16], and [17] is the absence of masking, typically occurring in the HVS, with the consequence that these metrics do not always reflect perceived ringing. Typical masking characteristics, such as luminance and texture masking, are explicitly considered in the metrics defined in [28] and [29], in which ringing regions are no longer simply assumed to surround all strong

edges in an image, but are determined by a model of the HVS. Including a HVS model in an objective metric might improve its accuracy, but often is computationally intensive for real-time applications. For example, the HVS model used in the metric presented in [28] largely depends on a parameter estimation procedure, which requires a number of calculations to achieve an optimal selection. The model described in [29] is based on a computationally heavy clustering scheme, including both color clustering and texture clustering. From a practical point of view, it is highly desirable to reduce the complexity of the HVS-based metric without compromising its overall performance.

The essential idea behind most of the existing metrics mentioned so far (see, e.g., in [14], [16], and [28]) is that they consist of a two-step approach. The first step identifies the spatial location, where perceived ringing occurs, and the second step quantifies the visibility or annoyance of ringing in the detected regions. This approach intrinsically avoids the estimation of ringing in irrelevant regions in an image, thus making the quantification of ringing annoyance more reliable, and the calculation more efficient. Additionally, a local determination of the artifact metric provides a spatially varying quality degradation profile within an image, which is useful in, e.g., video chain optimization as mentioned in Section I. Since ringing occurs near sharp edges, where it is not visually masked by local texture or luminance, the detection of ringing regions largely relies on an edge detection method followed by a HVS model. Existing methods (such as, e.g., [14], [16], [28], and [29]) usually employ an ordinary edge detector, where a threshold is applied to the gradient image to capture strong edges. Depending on the choice of the threshold, this runs the risk of omitting obvious ringing regions near nondetected edges (e.g., in case of a high threshold) or of increasing the computational cost by modeling the rather complex HVS near irrelevant edges (e.g., in case of a low threshold). This implies that to ensure a reliable detection of perceived ringing while maintaining low complexity for real-time applications, an efficient approach for both detecting relevant edges and modeling the HVS is needed. Quantification of the annoyance of ringing in the detected areas can be easily achieved by calculating the signal difference between the ringing regions and their corresponding reference, as used in the FR approach described in [14]. However, for a NR ringing metric, the quantification of ringing becomes more challenging mainly due to the lack of a reference. Metrics in literature (such as in [16] and [28]) estimate the visibility of ringing artifacts from the local variance in intensity around each pixel within the detected ringing regions, and average these local variances over all ringing regions to obtain an overall annoyance score. This approach, however, has limited reliability, since it does not include background texture in the ringing regions, which might affect ringing visibility.

To validate the performance of a ringing metric, its predicted quality degradation should be evaluated against subjectively perceived image quality. To prove whether a ringing metric is robust against different compression levels and different image content, the correlation between its objective predictions and subjective ringing ratings must be calculated. Unfortunately,
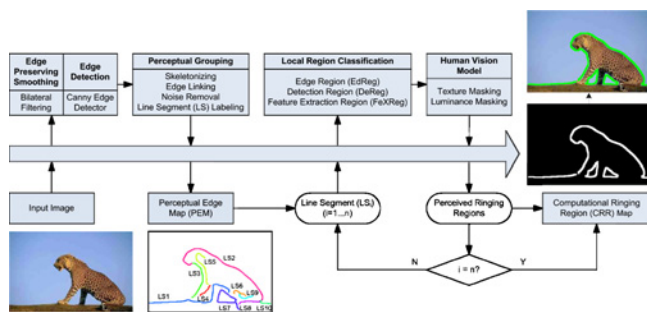
Fig. 4. Schematic overview of the proposed ringing region detection method. In PEM, each perceptually relevant LS is labeled in a different color. In the CRR map, the white areas indicate the detected perceived ringing regions, and the spatial location of these regions is illustrated in a separate image by green areas.

only the performance of the metric reported in [14] is evaluated against subjective data of perceived ringing. For all other metrics (such as the ones in [15]–[17] and [28]) nothing can be concluded with respect to their performance in predicting perceived ringing. Since we had no access to the data used in [14] for our metric evaluation, we performed our own subjective experiment.[1]

In this paper, we propose a NR ringing metric based on the same two-step approach mentioned above. For the first step, we rely on our ringing region detection method (see [30] and [31]), the performance of which in terms of extracting regions with perceived ringing has been shown to be promising [31]. Therefore, we consider this part of the metric readily applicable for the second step, in which the ringing annoyance is quantified. To quantify ringing annoyance, we consider each detected ringing region as a perceptual element, in which the local visibility of ringing artifacts is estimated. The contrast in activity between each ringing region and its corresponding background is calculated as the local annoyance score, which is then averaged over all ringing regions to yield an overall ringing annoyance score. It should be noted that the proposed metric is built upon the luminance component of images only in order to reduce the computational load. The performance of the NR metric is evaluated against subjective ringing annoyance in JPEG compression.

III. PROPOSED NR RINGING METRIC

A. Perceived Ringing Region Detection

For the design of our ringing region detection method (see [30] and [31]), we explicitly exploited the specific physical structure of ringing artifacts and some properties of the HVS. The overall proposed algorithm is schematically shown in Fig. 4, which mainly consists of two processing steps: 1) extraction of edges relevant for ringing, which results in a perceptual edge map (PEM), and 2) detection of perceived ringing regions, which yields a computational ringing region (CRR) map. This method is already described in more detail in [30] and [31], and is only briefly repeated here.

[1]The data collected from this experiment are available to the image quality assessment community on the website http://mmi.tudelft.nl/~ingrid/ringing.html.

To extract the most relevant edges for the purpose of ringing detection, an advanced edge detector is used. It adopts a bilateral filter [32] to largely smooth "irrelevant edges" (i.e., in textured areas), while the position of the "relevant edges" (e.g., contours of objects) is retained. Subsequently, a Canny edge detector [33] is applied on the filtered image to obtain the "relevant edges." The detected edges are combined into line segments [hereafter referred to as line segment (LS)], which are defined as elements of connected edge pixels. These LSs are constructed over the Canny edge map by a simple grouping process, including skeletonizing, edge linking, noise removal, and LS labeling. Fig. 4 shows the extracted PEM, which is formed by a set of these LSs. It clearly illustrates the selection of the edges more relevant for ringing (i.e., the contours of the leopard) in combination with the avoidance of the irrelevant edges (i.e., the texture in the skin of the leopard).

To select the edges around which ringing is actually perceived each LS of the PEM is examined individually on the occurrence of perceived ringing. To this end, the region around a LS is divided into three zones: 1) the edge region (i.e., EdReg); 2) the detection region (i.e., DeReg); and 3) the feature extraction region (i.e., FeXReg). First, the level of texture or detail is estimated from the FeXReg, and those parts of the DeReg, in which the visibility of ringing is masked by texture, are discarded. Subsequently, the average luminance in each remaining part of the DeReg is calculated and those parts with a value above or below a certain threshold are discarded. In this way, only those regions around each LS, in which ringing is visible, are extracted, and then accumulated in the CRR map as illustrated in Fig. 4.

B. Ringing Annoyance Estimation

The CRR map indicates the spatial location of perceived ringing, but it does not give any information yet on how annoying the ringing artifacts in the detected region are. To quantify ringing annoyance, we first split up the detected region in the CRR map into so-called ringing objects (ROs). Fig. 5 illustrates the definition of an RO. It starts from the LSs of the PEM, shown in Fig. 4. Each LS is considered to be split up in a set of connected components (i.e., objects) depending on the local level of texture and averaged luminance in its DeReg (as defined in [30] and [31]). Then, by using the model of the HVS, the visibility of ringing in each object is determined. By removing the objects, in which ringing is invisible due to masking, the remaining objects are defined as ROs. As an example, illustrated in Fig. 5(b) the LS1 of the PEM in Fig. 4 is split up in two ROs, while the LS2 remains as one RO. Some of the LSs, e.g., LS5, LS6, LS8, and LS9, do not result in an RO, since no visible ringing is detected around this LS based on the HVS. So, each RO intrinsically is a single cluster resulting from the application of the human vision model to the LSs of the PEM. Hence, the definition of an RO fully relies on the local image content, and as such, is independent of scaling or cropping the image. Once the ROs are defined [as illustrated in Fig. 5(c)], a ringing annoyance score (RAS) is calculated for each of them, and the overall annoyance score for the image is simply the mean of the RAS over all ROs.
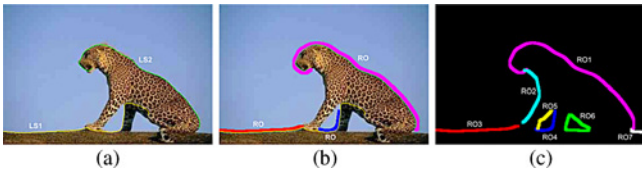
Fig. 5. Illustration of the definition of an RO. (a) Original JPEG image and two (out of ten) of its detected LSs (i.e., LS1 and LS2 of the PEM in Fig. 4). (b) Implementation of the human vision model to LS1 and LS2, resulting in two separate ROs for LS1 and one RO for LS2. (c) All detected ROs as a result of applying the human vision model to the whole PEM (i.e., ten LSs); they are indicated with different colors.
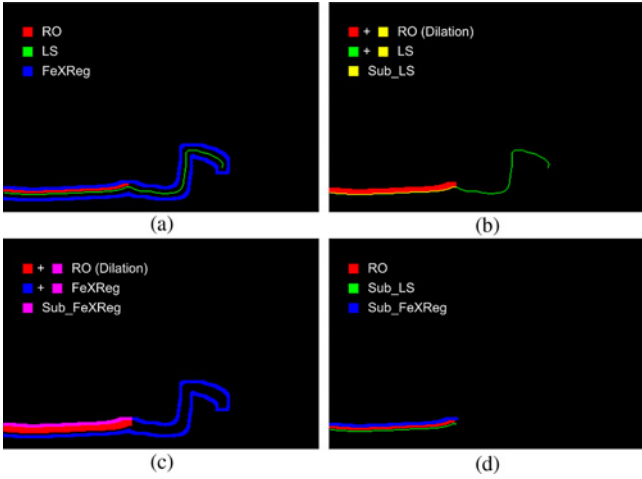

Fig. 6. Illustration of region assignment. (a) RO [see "RO3" in Fig. 5(b)] with its corresponding LS and feature extraction region (FeXReg). (b) Corresponding edge of LS covered by the dilated RO is assigned as the Sub-LS. (c) Corresponding region of FeXReg covered by the dilated RO is assigned as the Sub-FeXReg. (d) Results of region assignment.

The approach taken to quantify perceived ringing is inspired by the basic idea used in the FR metric [14], and is accomplished by the following two steps: 1) calculating the activity of each RO; and 2) comparing that activity to the activity in the neighboring background to which the RO belongs.

1) *Region Assignment:* To implement the two steps mentioned above, we first assign two relevant components to each RO in the CRR map: 1) the edge corresponding to each LS (i.e., referred to as Sub-LS), which is used to determine whether a pixel in the RO is a visible ringing pixel, and 2) the corresponding FeXReg region (i.e., referred to as Sub-FeXReg), which is employed as the reference for the RO. The FeXReg is located far away from the LS, and thus unlikely to be impaired by ringing artifacts. This region assignment is implemented by thickening an RO with a dilation operation. The corresponding LS and FeXReg which are covered by the RO during the dilation process are referred to as the Sub-LS and Sub-FeXReg, respectively. Fig. 6 illustrates this procedure. A specific RO (i.e., "RO3" in the CRR map of Fig. 5) with its corresponding LS and FeXReg are shown in Fig. 6(a). When dilating the RO with a square structuring element of 5 pixels width (e.g., for an image of 256×384 (height $\times$ width) pixels), the region of LS which is covered by the expanded RO is assigned as the Sub-LS (i.e., the yellow region in Fig. 6(b)). The Sub-FeXReg [i.e., the purple region

Fig. 7. Illustration of the list of coordinates as the result of region assignment (the total number of RO in the CRR map is n).

in Fig. 6(c)] is assigned in the same way by dilating the RO with a square structuring element of 9 pixels width. The resulting Sub-LS and Sub-FeXReg are shown in Fig. 6(d). It is noted that the size of the structuring element should be linearly scaled with the image size. The region assignment mentioned above is performed for each RO in the CRR map to eventually obtain a list of coordinates, which indicates the spatial location of each individual RO and its corresponding Sub-LS and Sub-FeXReg. Fig. 7 indicates the format of such a resulting list of coordinates. This way of working intrinsically facilitates the subsequent local analysis and processing of image characteristics.

2) *Local Visibility of Ringing Pixels:* Since ringing manifests itself in the form of artificial oscillations in the spatial domain, its local behavior can be reasonably described as the intensity variance of pixels in the neighborhood [28], [29]. In this paper, determining whether a pixel in an RO is a visible ringing pixel is based on calculating the local variance (LV) in intensity in its 3×3 neighborhood, which is formulated as

$$\text{LV}(i, j) = \frac{1}{9} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} \left[I(k, l) - \frac{1}{9} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} I(k, l) \right]^2$$
$$(i, j) \in Coord\{\text{RO}_n\} \qquad (1)$$

where $\text{LV}(i, j)$ denotes the local variance computed over a 3×3 template, centered at pixel (i, j) having an intensity $I(i, j)$ within the nth ringing object (i.e., RO_n).

The LV only yields an accurate result in case the RO is originally smooth around the edge; indeed, otherwise the LV can be high due to the activity of a textured or edge pixel. One would expect that the issue of considering texture as ringing is efficiently avoided by the application of a texture masking model in the ringing region detection phase (see [30] and [31]). However, we experienced that the dilation operation used in the human vision model may misclassify certain edge or texture components into an RO. In addition, there might be pixels in the RO exhibiting no or a very small intensity variance in their neighborhoods, which means they are not impaired by ringing artifacts (e.g., in higher bit rate compression). This implies that an RO still possibly contains spurious ringing pixels, which manifest themselves either as "noisy pixels" (i.e., misclassified edge or texture pixels) or as "unimpaired pixels" (i.e., pixels with a very low variance in intensity in the neighborhood). Fig. 8 gives an example of the image content underneath a detected RO (i.e., "RO2" as illustrated in Fig. 5), where noisy pixels and unimpaired pixels coexist with real ringing pixels. Calculating the LV
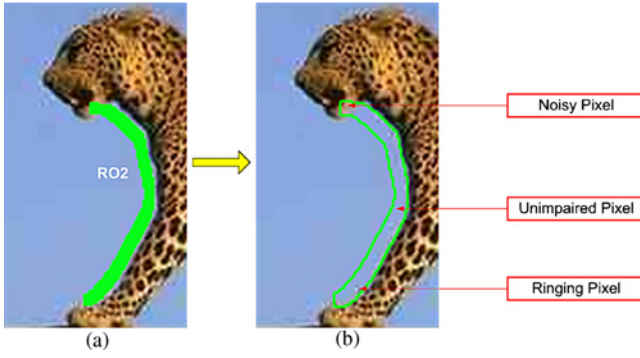
Fig. 8. Illustration of three types of pixels within an RO. (a) Original JPEG image and a detected RO [see "RO2" in Fig. 5(b)]. (b) Illustration of the image content underneath the corresponding RO, in which noisy pixels and unimpaired pixels coexist with real ringing pixels.

over these spurious ringing pixels may degrade the accuracy of measuring the actual ringing activity. The effect of the spurious ringing pixels on the RAS is avoided by applying two thresholds: 1) high threshold (Thr_vc_high), and 2) low threshold (Thr_vc_low). A pixel with its LV value above or equal to Thr_vc_high is considered as a "noisy pixel," and its visibility is set to "0." A pixel with its LV value below or equal to Thr_vc_low is considered as an "unimpaired pixel," and its visibility is also set to "0." Hence

$$
\text{VC}(i, j) = \begin{cases} \text{LV}(i, j), & Thr\_vc\_low < \text{LV}(i, j) \\ & \qquad < Thr\_vc\_high \\ 0, & \text{otherwise} \end{cases} \quad (2)
$$

where $\text{VC}(i, j)$ indicates the visibility coefficient at location (i, j) within the RO_n. After parameter optimization the value of Thr_vc_low is chosen to be zero, and the value of Thr_vc_high is chosen to scale with the strength of the corresponding edge, since we found that the actual LV range corresponding to a visible ringing pixel depends on the strength of its corresponding Sub-LS. Thus, Thr_vc_high is defined as

$$
Thr\_vc\_high = \alpha \cdot MAX\, [\text{LV}(i, j)],
$$
$$
(i, j) \in Coord\{Sub\_LS_n\} \quad (3)
$$

where LV is calculated over the Sub-LS (i.e., Sub_LS_n) assigned to the RO_n, and α (specified in Section V) is used to adjust the value of the high threshold.

All visible ringing pixels are extracted from each individual RO, and their visibility is indicated by a visibility coefficient (VC) according to (2). Fig. 9 illustrates the extraction of visible ringing pixels in an image, in which their visibility is indicated by a different color in a color bar.

3) Ringing Annoyance Estimation: The visibility coefficient for each ringing pixel in itself is yet insufficient to reflect the way human beings perceive ringing. It is the contrast between the visibility of a ringing artifact and its corresponding background that causes the perception of ringing annoyance [29], [38]. More strongly visible ringing pixels against a smoother background are most annoying. Since the Sub-FeXReg is already assigned to each RO to represent its
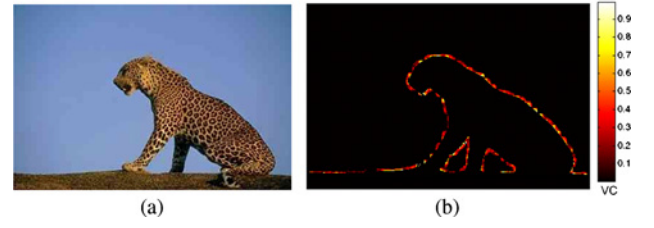

Fig. 9. Illustration of visible ringing pixel extraction. (a) Original JPEG image. (b) Extracted visible ringing pixels with their visibility indicated with a color ranging from black (no visibility) to yellow for the highest visibility.

local background, the activity of the Sub-FeXReg is readily calculated as the mean local variance (MLV)

$$
\text{MLV(Sub-FeXReg}_n) = \frac{1}{N_s} \sum \text{LV}(i, j),
$$
$$
(i, j) \in Coord\{\text{Sub-FeXReg}_n\} \quad (4)
$$

where N_s indicates the total number of pixels within the nth Sub-FeXReg (i.e., Sub-FeXReg$_n$), and $\text{LV}(i, j)$ indicates the local variance calculated at pixel location (i, j) within the Sub-FeXReg$_n$. For the corresponding RO (i.e., RO_n), its activity is defined as

$$
\text{MLV(RO}_n) = \frac{1}{N_r} \sum \text{VC}(i, j),
$$
$$
(i, j) \in Coord\{\text{RO}_n\} \text{ and } \text{VC}(i, j) \neq 0 \quad (5)
$$

where N_r indicates the total number of visible ringing pixels within the RO_n, and $\text{VC}(i, j)$ indicates the visibility coefficient [see (2)] calculated at pixel location (i, j) within the RO_n.

Once the activity of an RO and of its corresponding Sub-FeXReg is calculated, the difference between them is used to quantify the ringing annoyance for this RO. Hence, the ringing annoyance score (RAS) is defined as

$$
\text{RAS(RO}_n) = N_o \times [\text{MLV(RO}_n) - \text{MLV}(Sub\_FeXReg_n)] \quad (6)
$$

where N_o indicates the total number of pixels within the RO_n.

Based on the annoyance score per RO the overall ringing annoyance score for an image is calculated according to the procedure schematically shown in Fig. 10. It contains removal of ROs, for which the amount of visible ringing pixels is below a threshold R. In our algorithm, R is set as a pre-defined percentage (specified in Section V) of the total number of pixels in the RO. This is done with the estimation accuracy and speed in mind, since these ROs contain a too small number of visible ringing pixels to contribute to the overall perception of ringing annoyance. Eventually, the proposed ringing metric is defined as the mean of the ringing annoyance scores (MRAS) over all remaining ROs, which is formulated as

$$
\text{MRAS} = \frac{1}{T} \sum_{n=1}^{N} \text{RAS(RO}_n) \quad (7)
$$

where N indicates the total number of ROs, excluding the discarded ones, and T indicates the total number of pixels within these N ROs.
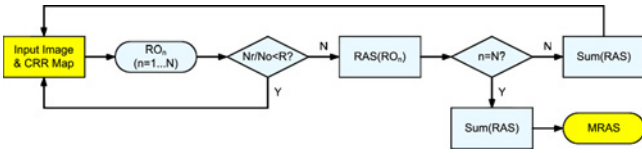
Fig. 10.   Schematic overview of the proposed ringing metric (i.e., MRAS).
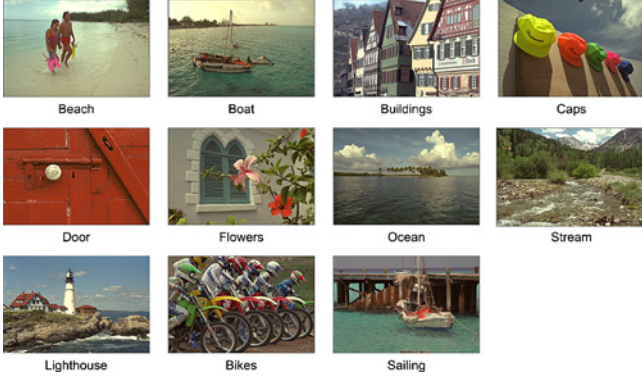

Fig. 11.   Source images used in the subjective quality study.

IV. Psychovisual Experiment

To validate our proposed ringing metric, a subjective experiment was carried out, in which participants scored the annoyance of ringing artifacts in compressed images.

A. Experimental Procedure

1) *Image Database and Test Environment:* A set of 11 source images, reflecting adequate diversity in image content, was taken from the "Kodak Lossless True Color Image Suite" [35]. Fig. 11 shows these source images. They were high-resolution and high-quality color images of size 768×512 (width $\times$ height) pixels. Some images have high activity, while others are mostly smooth. These images were JPEG compressed at four different compression levels (i.e., quality $Q = 25, 40, 55, 70$) using MATLAB's *imwrite* function. This yielded a test database of 55 stimuli (including the originals). The compression level was varied over such a range of quality levels that images with a broad range of ringing annoyance, from imperceptible to high levels of impairment, were generated. The stimuli were displayed on a Philips Cineos 37" liquid-crystal display screen with a native resolution of 1920×1080 pixels and a screen refresh rate of 60 Hz. The experiment was conducted in a standard office environment [34] and the viewing distance was approximately 60 cm.

2) *Test Methodology:* A single-stimulus method was used in our experiment, which means that subjects had to score the ringing annoyance for each stimulus in the absence of a reference. The scoring scale ranged from 0 to 100, where "0" means no ringing annoyance and "100" means highest ringing annoyance. The quality scale included additional semantic labels (i.e., "low," "average," and "high" ringing annoyance) at intermediate points for reference as illustrated in Fig. 12.

The participants of the study were recruited from the MSc program of the Department of Mediamatics at the Delft
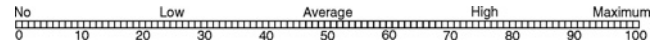

Fig. 12.   Quality scale used in the ringing annoyance assessment.

University of Technology. The 20 students, being 14 males and six females, were inexperienced with image quality assessment and coding artifacts. Before the start of the experiment, an instruction about the goal and procedure (e.g., the type of assessment, the scoring scale and the timing) of the experiment was given to each individual subject. A training session was conducted showing three examples of synthetic ringing, synthetic blocking and synthetic blur, followed by three real-life images in which ringing, blocking and blur were the most annoying artifacts, respectively. When the subject reported to understand ringing and to be able to distinguish it from other types of compression artifacts, a set of ten images covering the same range of ringing annoyance as used in the actual study was presented to the subject in order to familiarize him or her with how to use the range of the scoring scale. Then, three stimuli were shown one by one and the participant exercised how to indicate ringing annoyance on the scoring scale. The images used in the training session were different from those used in the actual experiment. After training, the test images were shown in a random order to each subject in a separate session.

B. Processing of the Raw Data

1) *Outlier Detection and Subject Rejection:* Before the actual data analysis, a simple outlier detection and subject rejection model was implemented on the raw annoyance scores. An individual score for an image was considered to be an outlier if it was outside an interval of two standard deviations around the mean score for that image. All annoyance scores of a subject were rejected if more than five of his/her scores were outliers. Overall, 1 subject out of 20 was rejected, and about 3% of the scores were rejected as outliers.

2) *MOS Scores:* After outlier removal and subject rejection, the scores of the remaining subjects were calibrated using z-scores [36]

$$z_{ij} = \frac{r_{ij} - \mu_i}{\sigma_i} \tag{8}$$

where r_{ij} and z_{ij} indicate the raw score and z-score for the ith subject and jth image, respectively. μ_i is the mean of the raw scores over all images scored by subject i, and σi is the corresponding standard deviation. The z-scores were then averaged across subjects to yield a mean opinion score (MOS) for the jth image

$$\text{MOS}_j = \frac{1}{S} \sum_{i=1}^{S} z_{ij} \tag{9}$$

where S is the total number of subjects (after subject rejection).

V. Performance Evaluation

Our proposed ringing metric is validated with respect to the data resulting from the psychovisual experiment, and its

performance is compared to three alternatives recently published in literature: one FR ringing metric, which is referred to as FRRM [14], and two NR ringing metrics, which are referred to as NRRM [16] and VRM [28], respectively. In literature, these metrics are all proved to be promising in measuring ringing artifacts in compressed images. It should be noted that we implemented these three metrics ourselves based on the information available in the papers and tuned their parameters to yield the highest performance possible for the set of test images used in our experiments. This is done to ensure a fair comparison between the results from different metrics. The parameters used for our proposed metric are specified as follows: 1) for the ringing region detection: $\sigma_d = 3$ and $\sigma_r = 100$ for the bilateral filter, $threshold\_high = 0.85$ and $threshold\_low = 0.4$ for the Canny edge detector, $Thr\_txt = 0.9$ and $Thr\_lum = 0.75$ for the human vision model, and the EdReg, DeReg, and FeXReg are determined with a square structuring element whose width is 3, 9, and 17, respectively (see [30] and [31]); and 2) for the ringing annoyance estimation: $Thr\_vc\_low = 0$, $\alpha = 0.5$, and $R = 0.75$. It should be noted that these parameter settings are empirically determined. The first set of parameters for the ringing region detection was defined based on subjective data for ringing region visibility (see [30] and [31]), and is proved in this paper to be robust for a new set of images. The remaining parameters used for the ringing annoyance estimation are determined in pilot experiments on both synthetic patterns and natural images. The performance of the metric is fairly insensitive to variations in the range of [0, 0.1] and [0.4, 0.6] for the values of Thr_vc_low and α, respectively. The parameter R is mainly used to speed up the algorithm, and thus, hardly affects the prediction accuracy of the metric.

A. Evaluation Criteria

As prescribed by the Video Quality Experts Group [25] the performance of an objective metric can be quantitatively evaluated with respect to its ability to predict subjective quality ratings (the MOS), based on the Pearson linear correlation coefficient to indicate prediction accuracy, the Spearman rank order correlation coefficient to indicate prediction monotonicity, and the outlier ratio to indicate prediction consistency. As suggested in [39], the metric's performance can also be evaluated with nonlinear correlations using a nonlinear mapping function for the objective predictions before computing the correlation. For example, a logistic function may be applied to the objective metric results to account for a possible saturation effect. A nonlinear fitting usually yields higher correlation coefficients in absolute terms, while generally keeping the relative differences between the metrics [39]. On the other hand, without a sophisticated nonlinear fitting (often including various parameters) the correlation coefficients cannot mask a bad performance of the metric itself. To better visualize differences in performance we propose to avoid any nonlinear fitting and to directly use linear correlation between the metric's predictions and the subjective data. However, to demonstrate the effect of a nonlinear mapping, both the linear and nonlinear correlations are given in this paper.
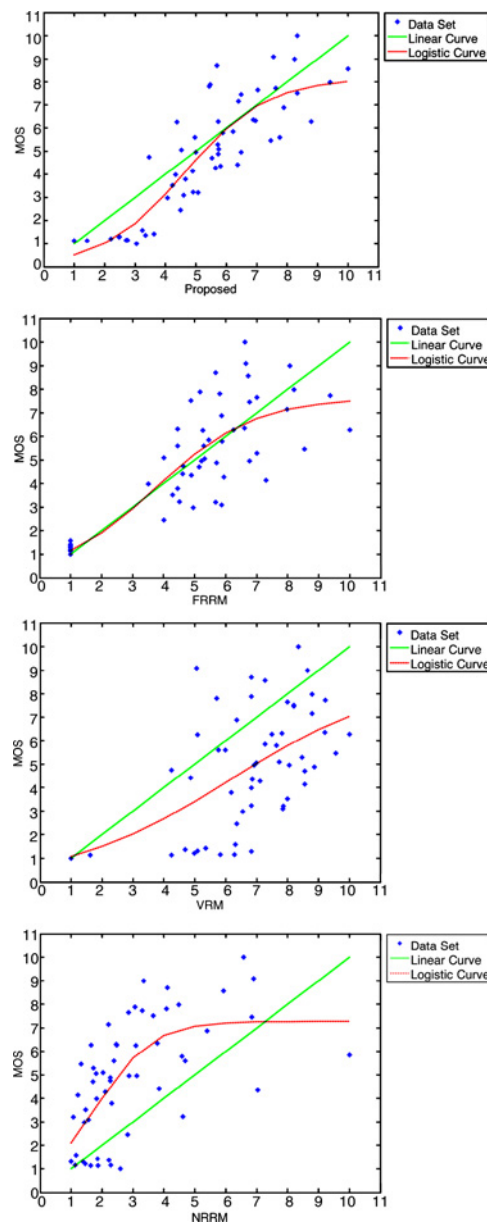

Fig. 13. Scatter plots of MOS versus the ringing metrics FRRM [14], VRM [28], NRRM [16], and our proposed metric. Full-line curves show the linear fit between the MOS and metric results, while the dashed-line curves show the nonlinear logistic fit.

B. Experimental Results

Our proposed ringing metric and the three alternative metrics (i.e., FRRM, NRRM, and VRM) are applied to our database of 55 stimuli. Fig. 13 shows the scatter plots of the MOS versus our proposed metric, FRRM, NRRM, and VRM, respectively. Table I lists the correlation coefficients. To also show the nonlinear correlation, a four-parameter logistic function suggested in [25] was used to fit the metric's predictions to the MOS. The resulting curve fits are included in Fig. 13, and the correlation coefficients are listed in Table II.

Fig. 13 and Table I demonstrate that our proposed NR ringing metric outperforms the existing metrics in the prediction of ringing annoyance. In comparison to the FR ringing metric FRRM our metric shows a higher correlation to the subjective

TABLE I
PERFORMANCE COMPARISON OF FOUR RINGING METRICS (OUR PROPOSED METRIC, FRRM [14], VRM [28], AND NRRM [16]) WITHOUT NONLINEAR FITTING

Metric	Pearson Linear Correlation	Spearman Rank Order Correlation	Outlier Ratio
Proposed	0.851	0.850	0
FRRM	0.793	0.744	0
VRM	0.519	0.498	0.291
NRRM	0.561	0.649	0.218

Threshold to determine the outlier ratio is set to 1.5 standard deviations of the MOS [25].

TABLE II
PERFORMANCE COMPARISON OF FOUR RINGING METRICS (OUR PROPOSED METRIC, FRRM [14], VRM [28], AND NRRM [16]) AFTER A LOGISTIC FIT OF THE METRICS' PREDICTIONS TO THE MOS

Metric	Pearson Linear Correlation	Spearman Rank Order Correlation	Outlier Ratio
Proposed	0.868	0.850	0
FRRM	0.824	0.744	0.127
VRM	0.521	0.498	0.218
NRRM	0.667	0.649	0.21

Threshold to determine the outlier ratio is set to 1.5 standard deviations of the MOS scores [25].

data, i.e., the gain in the Pearson correlation coefficient is $\Delta P = 5\%$, and in the Spearman correlation coefficient is $\Delta S = 11\%$. The lower correlation for the FRRM compared to our metric most probably is due to the absence of a HVS model in the FRRM. It simply assumes that ringing occurs unconditionally in regions surrounding strong edges in an image, neglecting possible luminance and texture masking effects. As a consequence, measuring ringing annoyance in the regions where ringing is invisible to the human eye potentially degrades the prediction performance of this metric. Our metric does contain a model for visual masking, and so, intrinsically avoids the estimation of ringing in irrelevant regions (e.g., texture areas) in an image, thus making the quantification of ringing annoyance more accurate.

Compared to the alternative NR ringing metrics, our metric manifests a much higher prediction performance relative to VRM and NRRM. The measured gain of our metric compared to VRM is $\Delta P = 33\%$ and $\Delta S = 35\%$, and compared to NRRM is $\Delta P = 29\%$ and $\Delta S = 20\%$. A possible reason for the lower performance of the NRRM is that it does not take into account spatial masking by the HVS, thus inevitably measuring ringing in some textured regions. Actually the metric may misclassify texture components into ringing artifacts, which may heavily degrade the prediction accuracy of a ringing metric. Comparing the performance of NRRM to that of FRRM (both without a masking model), it is clear that a NR metric is more sensitive to misclassified textured regions than a FR metric. A FR approach can account for the texture by comparing the region to the same unimpaired, but textured region in the reference. As a result, the error of misclassifying texture as ringing is expected (and confirmed) to be smaller.

It should be noted that exactly the same conclusions can be drawn from Table II as discussed above for Table I. This confirms the statement already reported in [39] that nonlinear mapping of the metric's predictions to the MOS affects the absolute values, but not the relative differences between metrics.

C. Discussion

The experimental results tend to validate our approach in the design a no-reference ringing metric, existing of: 1) a reliable ringing region detection model, and 2) a refined ringing annoyance estimation method. The importance of a reliable ringing region detection method can be seen by comparing the metric VRM to the one reported in [37] (which is a previous version of the one reported here, not including yet the comparison of the variance with the background and the detection of spurious ringing pixels). In both metrics, the annoyance score is simply defined as the intensity variance in the detected ringing regions. The only difference between them lies in the HVS model included in the metric of [37] for detecting *perceived* ringing regions. Therefore, the performance gain of the metric of [37] (with a Pearson correlation coefficient of 0.8) over VRM (with a Pearson correlation coefficient of 0.519) is attributed to the HVS included in the ringing region detection model. The added value of the refined ringing annoyance quantification (including the comparison of the variance with the background and the detection of spurious ringing pixels) can be validated by comparing the performance of the metric reported in this paper to its previous version reported in [37]. The gain in performance of the metric reported here over the one reported in [37] corresponds to an increase in the Pearson correlation coefficient from 0.80 to 0.851. This implies that quantifying ringing annoyance as the absolute intensity variance is effective, but is still too sensitive to remaining texture present in detected ringing regions. The perceived annoyance level is better addressed by comparing the local variance to the activity of its corresponding local surrounding.

It should be noted that the metric proposed in this paper is only validated for ringing perceived in JPEG compressed images, while ringing is also obviously present in JPEG2000 compressed images. There are, however, a couple of reasons, based on which one can expect a similar performance of our metric on JPEG2000 or H.264 compressed images. First
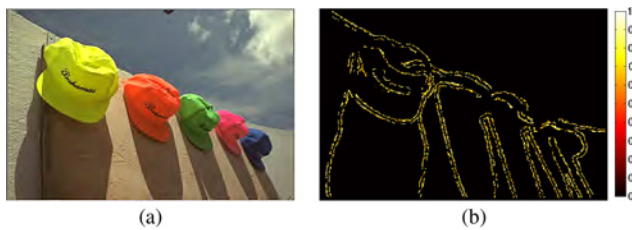
Fig. 14. Illustration of how the proposed NR ringing metric works on JPEG2000 compression. (a) JPEG2000 coded image (bit rate 0.099 b/p) from Laboratory for Image and Video Engineering database [43]. (b) Extracted visible ringing pixels with their visibility indicated with a color ranging from black (no visibility) to yellow for the highest visibility.

of all, most ringing metrics, and also ours, measure ringing in the spatial domain of the decoded image. As such, these metrics only rely on the characteristics of ringing artifacts (e.g., spatial edge information) rather than on the coding parameters (e.g., DCT coefficients or wavelet coefficients). As a consequence, one would not expect that these metrics need to be intrinsically changed for any of the existing image or video coding standards, but rather can be immediately used or at most need to be slightly modified for measuring ringing artifacts in any type of compressed image. This is confirmed by the claim already made for the metric VRM, namely that it is independent of the particular coding method employed [28], [38]. Additionally, it can be shown that the metric FRRM has a comparable performance for predicting perceived ringing in both JPEG2000 and JPEG compressed images. Indeed, its performance was characterized with a Pearson correlation coefficient of 85% for JPEG2000 compressed images in [14], while we found a Pearson correlation coefficient of 80% for JPEG compressed images in this paper. To illustrate the implementation of our proposed NR ringing metric on a JPEG2000 compressed image, an example is given in Fig. 14. It can be seen that the metric successfully identifies and quantifies ringing artifacts in the image. However, to fully evaluate the metric's performance subjective ringing ratings (not the overall quality scores) of JPEG2000 compressed images are needed, which we currently do not have to our availability.

Last, but not least, it should be noted that our performance evaluation with a subjective experiment is limited with respect to the amount of test stimuli, the number of human subjects and the display devices used. Adding more experimental data to the performance evaluation would be highly beneficial, but also is very time-consuming. To facilitate further benchmarking of ringing metrics, apart from developing computational models, future work should also focus on collecting and distributing more reliable subjective data.

VI. CONCLUSION

In this paper, a novel no-reference metric for perceived ringing artifacts in compressed images was presented. This metric relies on the existing perceived ringing region detection method [30], [31], and includes ringing annoyance estimation in the perceptually relevant regions in an image. For each individual ringing region, a ringing annoyance score is calculated by first estimating the local visibility of ringing

artifacts, and then by comparing it to the local background activity. An overall ringing annoyance score was obtained by averaging the local annoyance scores over all ringing regions. A psychovisual experiment was conducted to measure ringing annoyance subjectively and to validate our proposed ringing metric. The performance of our metric was compared to existing alternatives in literature. It demonstrated that our metric outperforms state-of-the-art metrics in predicting perceived ringing annoyance. Combined with its reliability and computational efficiency, our metric can be a good alternative for real-time implementation.

REFERENCES

[1] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment* (Synthesis Lectures on Image, Video and Multimedia Processing). San Rafael, CA: Morgan and Claypool, 2006.

[2] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.

[3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[4] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[5] H. R. Sheikh, A. C. Bovik, and L. K. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1918–1927, Dec. 2005.

[6] R. V. Babu, S. Suresh, and A. Perkis, "No-reference JPEG-image quality assessment using GAP-RBF," *Signal Process.*, vol. 87, no. 6, pp. 1493–1503, 2007.

[7] P. Gastaldo and R. Zunino, "Neural networks for the no-reference assessment of perceived quality," *J Electron. Imaging*, vol. 14, no. 3, p. 033004, 2005.

[8] H. R. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," *IEEE Signal Process. Lett.*, vol. 4, no. 11, pp. 317–320, Nov. 1997.

[9] R. Muijs and I. Kirenko, "A no-reference blocking artifact measure for adaptive video processing," in *Proc. 13th Eur. Signal Process. Conf.*, Sep. 2005.

[10] S. Liu and A. C. Bovik, "Efficient DCT-domain blind measurement and reduction of blocking artifacts," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 12, pp. 1139–1149, Dec. 2002.

[11] H. Liu and I. Heynderickx, "A perceptually relevant no-reference blockiness metric based on local image characteristics," *EURASIP J. Adv. Signal Process.*, vol. 2009, article ID 263 540, 2009.

[12] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. Sep. 2002, pp. 477–480.

[13] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, S. Yao, Z. Lu, and X. Yang, "A locally adaptive algorithm for measuring blocking artifacts in images and videos," *Signal Process.: Image Commun.*, vol. 19, no. 6, pp. 499–506, Jul. 2004.

[14] P. Marziliano, F. Dufax, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Application to JPEG2000," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 163–172, Feb. 2004.

[15] H. Tong, M. Li, H. Zhang, and C. Zhang, "No-reference quality assessment for JPEG2000 compressed images," in *Proc. IEEE Int. Conf. Image Process.*, vol. 5. Sep. 2004, pp. 3539–3542.

[16] R. Barland and A. Saadane, "Reference free quality metric for JPEG-2000 compressed images," in *Proc. Int. Symp. Signal Process. Applicat.*, vol. 1. Aug. 2005, pp. 351–354.

[17] X. Li, "Blind image quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. Sep. 2002, pp. 449–452.

[18] M. C. Q. Farias and S. K. Mitra, "No-reference video quality metric based on artifact measurements," in *Proc. IEEE Int. Conf. Image Process.*, vol. 3. 2005, pp. 141–144.

[19] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, Apr. 2009.

[20] K. Zon and W. Ali, "Automated video chain optimization," *IEEE Trans. Consum. Electron.*, vol. 47, no. 3, pp. 593–603, Aug. 2001.

[21] C. C. Koh, S. K. Mitra, J. M. Foley, and I. Heynderickx, "Annoyance of individual artifacts in MPEG-2 compressed video and their relation to overall annoyance," in *Proc. 10th Soc. Photo-Optic. Instrum. Eng. Human Vision Electron Imaging*, vol. 5666. Mar. 2005, pp. 595–606.

[22] M. Shen and C. Kuo, "Review of postprocessing techniques for compression artifact removal," *J. Visual Commun. Image Process.*, vol. 9, no. 1, pp. 2–14, Mar. 1998.

[23] R. Muijs and J. Tegenbosch, "Quality-Adaptive sharpness enhancement based on a no-reference blockiness metric," in *Proc. 2nd Int. Workshop Video Process. Quality Metrics Consum. Electron.*, Jan. 2006.

[24] I. O. Kirenko, R. Muijs, and L. Shao, "Coding artifact reduction using nonreference block grid visibility measure," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2006, pp. 469–472.

[25] VQEG. (2003, Aug.). Final report from the video quality experts group on the validation of objective models of video quality assessment [Online]. Available: http://www.vqeg.org

[26] M. Yuen and H. R. Wu, "A survey of hybrid MC/DPCM/DCT video coding distortions," *Signal Process.*, vol. 70, no. 3, pp. 247–278, Nov. 1998.

[27] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*. Norwell, MA: Kluwer, 2001.

[28] S. H. Oguz, Y. H. Hu, and T. Q. Nguyen, "Image coding ringing artifact reduction using morphological post-filtering," in *Proc. IEEE 2nd Workshop Multimedia Signal Process.*, 1998, pp. 628–633.

[29] X. Feng and J. P. Allebach, "Measurement of ringing artifacts in JPEG images," in *Proc. Soc. Photo-Optic. Instrum. Eng.*, vol. 6076. Feb. 2006, pp. 74–83.

[30] H. Liu, N. Klomp, and I. Heynderickx, "perceptually relevant ringing region detection method," in *Proc. 16th Eur. Signal Process. Conf.*, Aug. 2008.

[31] H. Liu, N. Klomp, and I. Heynderickx, "A perceptually relevant approach to ringing region detection," *IEEE Trans. Image Process.*, to be published.

[32] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vision*, Jan. 1998, pp. 836–846.

[33] J. Canny, "A Computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[34] *Methodology for the Subjective Assessment of the Quality of Television Pictures*, ITU-R Recommendation BT.500-11, Geneva, Switzerland, 2002.

[35] R. Franzen, *Kodak Lossless True Color Image Suite* [Online]. Available: http://www.r0k.us/graphics/kodak/

[36] A. M. Dijk, J. B. Martens, and A. B. Watson, "Quality assessment of coded images using numerical category scaling," in *Proc. Soc. Photo-Optic. Instrum. Eng.*, vol. 2451. Mar. 1995, pp. 90–101.

[37] H. Liu, N. Klomp, and I. Heynderickx, "A no-reference metric for perceived ringing," in *Proc. 4th Int. Workshop Video Process. Quality Metrics Consum. Electron.*, Jan. 2009.

[38] S. H. Oguz, "Morphological post-filtering of ringing and lost data concealment in generalized lapped orthogonal transform based image and video coding," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Wisconsin, Madison, 1999.

[39] S. Winkler, "Vision models and quality metrics for image processing applications," Ph.D. dissertation, Dept. Elect. Eng., Ecole Polytechnique Fédérale de Lausanne, Lausanne, 2002.

[40] A. B. Watson, *Digital Image and Human Vision*. Cambridge, MA: MIT Press, 1993.

[41] R. J. Safranek, T. N. Pappas, and J. Chen, "Perceptual criteria for image quality evaluation," in *Handbook of Image and Video Processing*, A. Bovik, Ed. San Diego, CA: Academic, 2004.

[42] B. A. Wandell, *Foundations of Vision*. Sunderland, MA: Sinauer Associates, Inc., 1995.

[43] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, *LIVE Image Quality Assessment Database*, release 2 [Online]. Available: http://live.ece.utexas.edu/research/quality
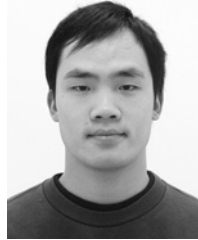
Hantao Liu (S'07) was born in Chengdu, China, in 1981. He received the M.S. degree in signal processing and communications from the University of Edinburgh, Edinburgh, U.K., in 2005. Since 2006, he has been a Ph.D. student at the Department of Mediamatics, Delft University of Technology, Delft, The Netherlands.

He is currently working on a research project supported by Philips Research Laboratories, Eindhoven, The Netherlands, developing no-reference objective metrics for perceived artifacts in compressed images. His current research interests include image analysis, visual perception, and signal processing.

Nick Klomp received the B.S. and M.S. degrees from Delft University of Technology, Delft, The Netherlands, in 2006 and 2008, respectively. From 2007 to 2008, he worked on a Master thesis at the Department of Mediamatics, Delft University of Technology, developing a no-reference ringing artifact metric for real-time video applications.

Ingrid Heynderickx received the Ph.D. degree in physics from the University of Antwerp, Belgium, in December 1986.

In 1987, she joined Philips Research Laboratories, Eindhoven, The Netherlands, and worked in different areas of research: optical design of displays, processing of liquid crystalline polymers, and functionality of personal care devices. Since 1999, she has been the Head of research activities on visual perception of display and lighting systems, and in 2005 she was appointed as a Research Fellow in the group Visual Experiences. She is a Member of the Society for Information Displays (SID), and for the SID, she was the Chairman of the Applied Vision subcommittee from 2002 to 2007. In 2008, she was a Fellow of the SID, and the Chairman of the European Committee of the SID. In 2005, she was appointed as a Guest Research Professor at the Southeast University of Nanjing, Nanjing, China, and a Part-Time Full Professor at the Delft University of Technology, Delft, The Netherlands.