RRED INDICES: REDUCED REFERENCE ENTROPIC DIFFERENCING FRAMEWORK FOR IMAGE QUALITY ASSESSMENT

Rajiv Soundararajan and Alan C. Bovik

Dept. of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712, USA
Email: rajiv.s@mail.utexas.edu, bovik@ece.utexas.edu

ABSTRACT

We study the problem of automatic "reduced reference" image quality assessment algorithms from the point of view of image information change. Algorithms that measure differences between the entropies of wavelet coefficients of reference and distorted images are designed. A family of algorithms are presented, each differing in the amount of data on which information change is predicted and ranging from almost full reference to almost no reference. A special case of this are algorithms that require just a single number from the reference for quality assessment. The algorithms are shown to correlate very well with subjective quality scores as demonstrated on the LIVE Image Quality Assessment Database.

Index Terms— Image quality, reduced reference quality assessment, natural scene statistics, image information

1. INTRODUCTION

Image and video quality assessment (QA) algorithms can be broadly classified into full reference (reference available or FR) and no reference (reference not available or NR) algorithms. The mean squared error (MSE) has been used as a quality metric for a very long time, owing to its simplicity, despite having a very poor correlation with human perception [1]. The last decade has seen significant progress in the field of objective full reference image/video QA algorithms. The structural similarity index (SSIM) [2] and visual information fidelity (VIF) [3] are examples of successful full reference algorithms which have been shown to perform very well in predicting the quality scores of human subjects.

Most of the progress on NR QA that has been possible, has been accomplished by relaxing the no reference assumption in various ways. One approach is to devise NR algorithms by exploiting the prior knowledge about the distortion process afflicting the image. Alternately, partial information about the reference can be made available, which can be used along with the distorted image to predict quality. This paradigm is known as reduced reference (RR) QA, which may or may not require knowledge of the distortion type.

RR QA algorithms involve sending or supplying some amount of information about the reference along with the distorted image that is useful in quality computation. For example, the concept of *quality aware* images was proposed in [4], where partial reference image information is embedded within the image and can be extracted reliably despite distortions. The information embedded could for example, be the statistical parameters of the distribution of wavelet coefficients obtained by a multi scale-space-orientation decomposition of the reference image. The quality is based on computing the Kullback-Leibler (KL) divergence between the wavelet coefficients of the reference and the distorted images. This idea is further extended in [5], where an additional divisive normalization transform step is introduced before computing the KL divergence to improve performance. However, this algorithm depends on a number of parameters that need to be trained on databases.

In this paper, we develop a new framework of reduced reference QA algorithms that are information theoretic. The algorithms compute the average difference between scaled entropies of wavelet coefficients of reference and distorted images, obtained at the output of a neural noise channel. A family of algorithms are proposed depending on the subband in which the quality computation is carried out and the amount of information required from the reference image. The algorithms allow for bidirectional computation of quality, by which we mean that the quality of the distorted image can also be computed at the reference, if relevant information from the distorted image is made available. This feature has potential applications in image/video quality monitoring in networks, which requires feeding back the quality at different nodes in the network to the sender. Another interesting feature of these algorithms is that they are not dependent on any parameters that need to be trained on databases. Depending on the bandwidth available for information, one algorithm from this class may be picked for desired applications. A special case of these algorithms are algorithms that require merely a single number from the reference for quality computation. The framework also allows users to choose an algorithm from this class for general purpose or distortion specific quality assessment.

2. RR QA ALGORITHMS

The source and distortion models considered here closely follow the assumptions in [6]. The resulting RR algorithms utilize the wavelet coefficients obtained by a steerable pyramid decomposition of the reference and distorted image into subbands at different orientations and scales [7]. Let K be the total number of subbands obtained as a result of this decomposition. The wavelet coefficients in subband k, $k \in \{1, 2, \ldots, K\}$, are partitioned into M_k non-overlapping blocks, each block containing N coefficients of size $\sqrt{N} \times \sqrt{N}$. Non-overlapping blocks are assumed independent and identically distributed (i.i.d.).

2.1. Source Model

Let $\bar{C}_{mk} = (C_{1mk}, C_{2mk}, \ldots, C_{Nmk})$ denote the vector of coefficients in block m, $m \in \{1, 2, \ldots, M_k\}$ of subband k of the reference image. We assume a Gaussian scale mixture model for natural images. Therefore

$$\bar{C}_{mk} = S_{mk}\bar{U}_{mk},$$

where $\bar{U}_{mk} \sim \mathcal{N}(0, \mathbf{K}_{U_k})$ and S_{mk} is a scalar random variable that modulates the covariance matrix of the block $\bar{C}_{mk}$. Also, S_{mk} and $\bar{U}_{mk}$ are independent. Subband k is associated with the covariance matrix $\mathbf{K}_{U_k}$ and $\bar{U}_{mk}$ for each block m, is distributed identically. Thus, the wavelet coefficient block $\bar{C}_{mk}$, when conditioned on the realization $S_{mk} = s_{mk}$, is distributed according to a Gaussian model with a covariance matrix $s_{mk}^2 \mathbf{K}_{U_k}$. Further, S_{mk} and $\bar{U}_{mk}$ are each independent over m and k.

2.2. Distortion Model

Distortions introduced in a natural image may take it outside the space of natural images. As a result, it is possible that the wavelet coefficients of these images do not follow a Gaussian scale mixture distribution. We measure quality as a distance between the reference and a natural image approximation of the distorted image by modelling the wavelet coefficients of the distorted image as well as a Gaussian scale mixture distribution. We show later that the approach of approximating the wavelet coefficients of a distorted image by a GSM model results in RR QA algorithms that perform very well in predicting the quality scores of images. Denote $\bar{D}_{mk} = (D_{1mk}, D_{2mk}, \ldots, D_{Nmk})$ as the vector of coefficients in block $m \in \{1, 2, \ldots, M_k\}$ of subband k of the distorted image. We have,

$$\bar{D}_{mk} = T_{mk}\bar{V}_{mk},$$

where $\bar{V}_{mk} \sim \mathcal{N}(0, \mathbf{K}_{V_k})$ and T_{mk} is the scalar premultiplier random variable as in the reference image. The independence assumptions are similar to the reference image. We now describe the RR index.
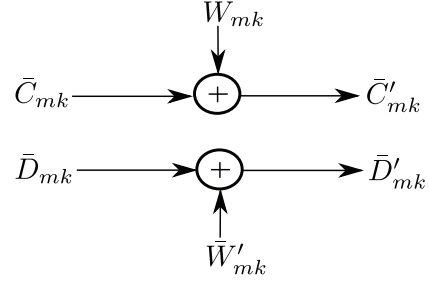

Fig. 1. System Model

2.3. RR Quality Index

Since quality assessment is a perceptual problem, we additionally model the perceived reference and distorted images as having passed through an additive neural noise channel, where the noise is assumed to be a zero mean Gaussian random vector for each block of coefficients. The resulting system model is shown in Fig. 1. We have

$$\bar{C}'_{mk} = \bar{C}_{mk} + W_{mk} \quad \bar{D}'_{mk} = \bar{D}_{mk} + W'_{mk},$$

where $W_{mk} \sim \mathcal{N}(0, \sigma_W^2 \mathbf{I})$ and $W'_{mk} \sim \mathcal{N}(0, \sigma_W^2 \mathbf{I})$, W_{mk} and W'_{mk} are independent of each other, independent of $\bar{C}_{mk}$ and $\bar{D}_{mk}$ and independent across the indices m and k. The reduced reference quality indices that we introduce, and which we term Reduced Reference Entropic Difference (RRED) indices, are defined as the average of the absolute value of the difference between the scaled entropies of the neural noisy reference and distorted images, conditioned on the realizations of the respective premultiplier random variables in a subband.

Let the eigen values of $\mathbf{K}_{U_k}$ be $\alpha_{1k}, \alpha_{2k}, \ldots, \alpha_{Nk}$ and the eigen values of $\mathbf{K}_{V_k}$ be $\beta_{1k}, \beta_{2k}, \ldots, \beta_{Nk}$. In the following, assume that $\mathbf{K}_{U_k}$ and $\mathbf{K}_{V_k}$ are full rank matrices. If this is not true, then the index is calculated by using the positive eigen values alone. The entropies of wavelet coefficient blocks indexed by m in subband k belonging to the reference and distorted images, conditioned on $S_{mk} = s_{mk}$ and $T_{mk} = t_{mk}$ respectively, are given by

$$h(\bar{C}'_{mk}|S_{mk} = s_{mk}) = \sum_{n=1}^{N} \frac{1}{2} \log \left[(2\pi e)(s_{mk}^2 \alpha_{nk} + \sigma_W^2) \right]$$

$$h(\bar{D}'_{mk}|\bar{T}_{mk} = t_{m,k}) = \sum_{n=1}^{N} \frac{1}{2} \log \left[(2\pi e)(t_{mk}^2 \beta_{nk} + \sigma_W^2) \right].$$

Define scaling factors,

$$\gamma_{mk}^r = \log(1 + s_{mk}^2) \text{ and } \gamma_{mk}^d = \log(1 + t_{mk}^2).$$

The entropies conditioned on the realizations of the premultiplier random variables are multiplied by the above scalars before computing the difference. These scalars lend a local character to the algorithm by imposing additional local effects on

the entropy terms. Moreover, the scalars may be interpreted as facilitating contrast masking while preventing over masking. They are also useful in the context of extremely small neural noise variance, where they help saturate the entropy terms at locations having extremely small premultiplier random variable realizations. This helps avoid numerical instabilities in the computation of the index.

We present a family of algorithms, by varying the subband in which quality is evaluated and the amount of information that is required from the reference for quality computation. First, we discuss algorithms obtained by varying the subband in which quality computation is carried out, alone. In these algorithms, the scaled entropies at each block in one particular subband k, $\{\gamma_{mk}^r h(\bar{C}_{mk}'|S_{mk} = s_{mk})\}_{m=1}^{M_k}$ of the reference image are required to evaluate quality. Since different subbands have different sizes, the number of blocks M_k, reduces from the subbands at the finest to the subbands at the coarsest scales of the wavelet decomposition.

The reduced reference QA index corresponding to subband k, when M_k scalars are available from the reference is given by

$$RRED_k^{M_k} = \frac{1}{L_k} \sum_{m=1}^{M_k} |\gamma_{mk}^r h(\bar{C}_{mk}'|S_{mk} = s_{mk}) - \gamma_{mk}^d h(\bar{D}_{mk}'|T_{mk} = t_{mk})|,$$

where L_k is the size (number of coefficients) of the subband k. The above index is a reduced reference index since M_k is less than the size of the image. Note that either image can compute the index using the entropy information from the other image. The absolute value of the difference is calculated, since the nature of the distortion process could lead to either an increase or a decrease in entropy.

The amount of information required from a subband can also be reduced by summing scaled entropy terms over local patches and sending the sum of these scaled entropies instead of all the entropy terms. This is equivalent to filtering the image of scaled entropies in a subband using rectangular windows of sizes $b \times b$ and subsampling by b in each dimension, where b is a natural number that represents the size of the patches. Let Λ_k denote the number of subsampled blocks and let $\lambda \in \{1, 2, \ldots, \Lambda_k\}$ index the block. Every $m \in \{1, 2, \ldots, M_k\}$ belongs to one of the subsampled blocks $B_{\lambda k}$ in subband k. Define

$$g_{\lambda k}^r = \sum_{m \in B_{\lambda k}} \gamma_{mk}^r h(\bar{C}_{mk}'|S_{mk} = s_{mk})$$

$$g_{\lambda k}^d = \sum_{m \in B_{\lambda k}} \gamma_{mk}^d h(\bar{D}_{mk}'|T_{mk} = t_{mk}).$$

Then the RR quality index in subband k when Λ_k, scalars

are available from the reference, is given by

$$RRED_k^{\Lambda_k} = \frac{1}{L_k} \sum_{\lambda=1}^{\Lambda_k} |g_{\lambda k}^r - g_{\lambda k}^d|.$$

The superscript denotes the number of scalars required from the respective subband. For example, for $b = 2$, the subband is filtered by windows of size 2×2 and subsampled by a factor of 2 in each dimension, and the number of entropy terms required reduces from M_k to $\Lambda_k = M_k/4$. $RRED_k^1$ denotes the algorithm in which all the scaled entropy terms in the subband are added and only the sum, which is a single scalar, is required for quality computation. Since only a single number is needed, this may be considered as an almost NR algorithm.

2.4. Estimation of Parameters

In order to compute the QA index, it is necessary to estimate s_{mk}, t_{mk}, $\mathbf{K}_{U_k}$ and $\mathbf{K}_{V_k}$ for $m \in \{1, 2, \ldots, M_k\}$ and $k \in \{1, 2, \ldots, K\}$. The procedure outlined here is similar to the estimation of reference image parameters in [6]. We obtain maximum likelihood estimates for the above parameters. Without loss of generality, assume $\sum_{m=1}^{M_k} s_{mk}^2 = 1$ and $\sum_{m=1}^{M_k} t_{mk}^2 = 1$. Therefore, the ML estimates of $\mathbf{K}_{U_k}$ and $\mathbf{K}_{V_k}$ are given by

$$\hat{\mathbf{K}}_{U_k} = \sum_{m=1}^{M_k} \frac{\bar{C}_{mk}\bar{C}_{mk}^T}{M_k} \quad \hat{\mathbf{K}}_{V_k} = \sum_{m=1}^{M_k} \frac{\bar{D}_{mk}\bar{D}_{mk}}{M_k}.$$

Since the wavelet coefficients are conditionally Gaussian distributed, the ML estimates of s_{mk}^2 and t_{mk}^2 are given by

$$\hat{s}_{mk}^2 = \frac{\bar{C}_{mk}^T \mathbf{K}_{U_k}^{-1} \bar{C}_{mk}}{N} \quad \hat{t}_{mk}^2 = \frac{\bar{D}_{mk}^T \mathbf{K}_{V_k}^{-1} \bar{D}_{mk}}{N}.$$

3. RESULTS AND DISCUSSION

We conducted experiments on the LIVE Image Quality Database [8] of distorted images and perceptual scores. Both the reference and distorted images are decomposed into different subbands using a steerable pyramid wavelet decomposition using 6 orientations at 4 scales [7]. Thus there are a total of 26 subbands in the wavelet decomposition. The algorithm was implemented using blocks of size 3×3 in each subband, implying a value of $N = 9$.

We evaluated the performance of algorithms obtained by computing $RRED_k^{M_k}$ for all the horizontal and vertical oriented subbands at different scales. The analysis in [6] as well as in the vision literature suggests that human subjects are more sensitive to horizontal and vertical orientations than others. Further, we observe that the performance obtained by choosing the vertical subbands is marginally better than the horizontal subbands. The vertically oriented subbands are

indexed by $k = 4, 10, 16, 22$ from the coarsest to the finest scale, i.e. at levels 1 through 4. Note that the $RRED$ indices evaluated in each subband can also be extended to the case where information from different subbands can be combined suitably. Due to space limitations, we do not present a detailed analysis of the performance of many such algorithms belonging to this framework. Here, we only present the results of one algorithm ($RRED_{16}^{M_{16}}$) that performs almost as well as the best FR algorithms and one algorithm ($RRED_{22}^1$) that just uses a single number. We choose $\sigma_W^2 = 0.1$ [3]. Recall that L is the size (total number of pixels) of the image.

Table 1. SROCC between objective algorithms and LIVE Image Database scores

Distortion Type	$RRED_{16}^{M_{16}}$	$RRED_{22}^1$	$PSNR$
JPEG2000	0.9580	0.9514	0.8951
JPEG	0.9759	0.9152	0.8812
AWGN	0.9780	0.9447	0.9853
Gaussian Blur	0.9678	0.9038	0.7812
Fast fading errors	0.9427	0.9183	0.8904
Overall	0.9429	0.7978	0.8754
No.of scalars	L/36	1	L

In Table 1, the Spearman rank order correlation coefficient (SROCC) between the scores of two particular RRED indices and subjective (DMOS) scores from the LIVE image database are shown. We also included the performance of PSNR, which is an FR QA algorithm for comparison. One important observation is that $RRED_{16}^{M_{16}}$ *performs nearly as well as the best performing FR QA algorithms* such as VIF [3]. Further, $RRED_{16}^{M_{16}}$ significantly outperforms PSNR, which is an FR algorithm.

Observe that the single number algorithm also possesses very good performance within individual distortion categories. The reduced performance on the overall database can be attributed to the different ranges of quality scores for the different distortion types. Certain distortions such as JPEG and AWGN noise lead to different locations in the same subband yielding an increase and decrease in entropy simultaneously. For example, in JPEG, increases in entropy occur due to the introduction of discontinuous blocking artifacts, while decreases in entropy occur in smoother regions that are heavily quantized. As a result, when scaled entropies are summed up, the gain and loss of entropies tend to cancel each other leading to lower quality ranges. However, for a particular distortion process, the single number algorithms could be very useful in predicting qualities for that process by using hardly any information from the reference.

4. CONCLUSION

We studied the problem of reduced reference image quality assessment by measuring the changes in suitably scaled

entropies between the reference and distorted images in the wavelet domain. The algorithms differ in the subbands in which the quality is evaluated and the amount of information required from the reference. When the number of scalars required is around 1/40 of the image size, the algorithm achieves a performance which is nearly as good as the best performing full reference QA algorithms. Moreover, the algorithms perform significantly better than the mean squared error, which is an FR algorithm. Even when only a single scalar is obtained from the reference image, the algorithm achieves state of the art performance within each distortion category without knowing anything about the type of distortions that the image might have been subjected to. The overall performance of the algorithm in such cases may be improved further by better aligning the scores obtained for different distortion categories. This is a subject of future research.

5. REFERENCES

[1] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? - A new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

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

[3] 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.

[4] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. H. Yang, and A. C. Bovik, "Quality-aware images," *IEEE Trans. Image Process.*, vol. 15, no. 5, pp. 1680–1689, 2006.

[5] Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," *IEEE Journal of Selected Topics in Signal Processing: Special issue on Visual Media Quality Assessment*, vol. 3, pp. 202–211, 2009.

[6] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.

[7] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *Proc. IEEE Int Conf on Image Processing*, 1995.

[8] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "Live image quality assessment database release 2," available at http://live.ece.utexas.edu/research/quality.