Invited submission for special session on “Advanced applications of objective video quality metrics and methods”

Statistical Approaches to Quality Assessment for Image Restoration

Hiroyuki Takeda, Hae Jong Seo, and Peyman Milanfar
Department of Electrical Engineering, University of California at Santa Cruz
{htakeda,rokaf,milanfar}@soe.ucsc.edu

***Abstract*— In this paper, we propose two novel (non-parametric) statistical measures of image and video quality. Both of these are based on the fundamental notion of singular value decomposition, applied to local pixel and higher order information derived from the data. These methods enable measurement of similarity between images, and the measurement of image quality without a reference. A possible derivative application of the latter statistical measure is to optimize parameters of the relevant restoration algorithms being applied.**

I. Introduction

Over the past several years, the state of the art in the classical problems of denoising, deblurring, and super-resolution has been significantly advanced. Therefore, the need for better and more consistent measures of quality of the reconstructions is stronger than ever. When a reference image is available, the Mean-Squared Error (MSE) has often been used. However a small MSE does not always correspond to a visually appealing result. Therefore, other metrics that correspond more closely to the perception of high quality are sought after [1]. Indeed, such quantitative metrics must often be computed in the absence of the refrenece image.

In this paper, we propose two novel statistical measures of image and video quality. The first is based on the application of the Singular Value Decomposition (SVD) in windowed fashion to estimated gradient vector fields across the image. By statistical analysis of the resulting singular values, we can not only decide whether this window contained structures of different sorts (e.g. flat region, edge, etc.), but also whether these structures are present in dominant fashion or not (e.g. sharp edge, versus a shallow edge.) The latter, when applied to an image restoration problem, can indicate how well an image has been reconstructed in terms of recovery of edges and other salient features.

The second statistical tool we propose is a natural generalization of the SVD applied to multiple data sets. The concept of Canonical Correlation Analysis (CCA)[2] is designed to measure the similarity between two general data sets. When applied to a pair of images, it enables a geometric comparison of the pixel values, gradient fields, or higher order derivatives of these images against one another. This measure can be thought of as a more general way of measuring image quality than the popular Structural Similarity (SSIM) measure [1]. Specifically, in the context of image restoration, it enables the comparison of the two images in terms of preservation of discontinuities, hence giving an indication of quality.

A possible derivative application of the above statistical measures is to optimize parameters of the relevant restoration algorithms being applied.

II. Statistical Measures of Image Quality

A. A Reference-based Canonical Correlations Measure

We propose a statistical method based on the notion of *canonical correlations*. The key idea behind CCA[2] is to find unit direction vectors along which the data are maximally correlated. The CCA, in addition to maximizing the mutual correlations, has an affine-invariant property which is desirable for any similarity measure. Despite the apparent utility of CCA, we believe this is the first time CCA has been used for image quality assessment. We can define a data model in $\mathcal{N}$-D using k-th order Taylor series as

$$z(\mathbf{x}_i) \approx \alpha_0 + \boldsymbol{\alpha}_1^T(\mathbf{x}_i - \mathbf{x}) + \boldsymbol{\alpha}_2^T \operatorname{vech}\left\{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T\right\} + \cdots ,$$

where $\operatorname{vech}(\cdot)$ is the half-vectorization operator which lexicographically orders the lower triangular portion of a symmetric matrix, and $\alpha_0, \boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}$ are a pixel value, gradient vectors, and hessian vectors respectively. Suppose that $z_1(\mathbf{x})$ and $z_2(\mathbf{x})$ are a pair of images. Namely, a clean image and a degraded (or restored) image, respectively. For each image, employing a local analysis window W_i, we can gather the image pixel, gradient, an possibly higher order information into a data matrix as follows:

$$\boldsymbol{\Delta}_i = [\boldsymbol{\delta}_{\mathbf{0}i}, \boldsymbol{\delta}_{\mathbf{1}i}, \cdots] = \begin{bmatrix} \vdots & \vdots & \\ \alpha_0(\mathbf{x}_j) & \boldsymbol{\alpha}_1^T(\mathbf{x}_j) & \cdots \\ \vdots & \vdots & \end{bmatrix}, \mathbf{x}_j \in W_i. \quad (1)$$

The Canonical Correlations[2] Similarity Measure (CCSIM) between two data sets $\boldsymbol{\Delta}_i^1$ and $\boldsymbol{\Delta}_i^2$ can be computed over the entire image in a patch-based fashion.

$$\mathrm{CCSIM}(\boldsymbol{\Delta}_i^1, \boldsymbol{\Delta}_i^2) = \rho_i, \quad (2)$$

where ρ_i are the canonical correlations and found by solving the following eigenvalue equations:

$$\begin{aligned} \mathbf{C}^{-1}_{\mathbf{\Delta}^1_i\mathbf{\Delta}^1_i}\mathbf{C}_{\mathbf{\Delta}^1_i\mathbf{\Delta}^2_i}\mathbf{C}^{-1}_{\mathbf{\Delta}^2_i\mathbf{\Delta}^2_i}\mathbf{C}_{\mathbf{\Delta}^2_i\mathbf{\Delta}^1_i}\mathbf{U}_{\mathbf{\Delta}^1_i} &= \rho_i^2\mathbf{U}_{\mathbf{\Delta}^1_i}, \\ \mathbf{C}^{-1}_{\mathbf{\Delta}^2_i\mathbf{\Delta}^2_i}\mathbf{C}_{\mathbf{\Delta}^2_i\mathbf{\Delta}^1_i}\mathbf{C}^{-1}_{\mathbf{\Delta}^1_i\mathbf{\Delta}^1_i}\mathbf{C}_{\mathbf{\Delta}^1_i\mathbf{\Delta}^2_i}\mathbf{U}_{\mathbf{\Delta}^2_i} &= \rho_i^2\mathbf{U}_{\mathbf{\Delta}^2_i}, \end{aligned} \quad (3)$$

where $\mathbf{U}_{\mathbf{\Delta}^1_i}$,$\mathbf{U}_{\mathbf{\Delta}^2_i}$ are the normalized canonical correlation basis vectors and $\boldsymbol{C}$ is the total covariance matrix:

$$\boldsymbol{C} = \begin{bmatrix} \mathbf{C}_{\mathbf{\Delta}^1_i\mathbf{\Delta}^1_i} & \mathbf{C}_{\mathbf{\Delta}^1_i\mathbf{\Delta}^2_i} \\ \mathbf{C}_{\mathbf{\Delta}^2_i\mathbf{\Delta}^1_i} & \mathbf{C}_{\mathbf{\Delta}^2_i\mathbf{\Delta}^2_i} \end{bmatrix}. \quad (4)$$

Then the overall quality measure of the entire image is computed by averaging the CCSIM values across all the local windows.

CCSIM is based on CCA and the data hierarchy while SSIM[1] is based on a simple correlation coefficient for structural similarity comparison. Besides, CCSIM can easily be extended to the video quality assessment without explict motion compensation. In this vein, CCSIM can be considered as a more general way of measuring image quality than SSIM.

In Fig.1, the mean CCSIM and SSIM graphs with respect to JPEG quality factors and White Gaussian Noise level (Input and Output) are shown. It is clear that our proposed measure appears to work similarly to SSIM, and we believe that SSIM is a special case of our method.

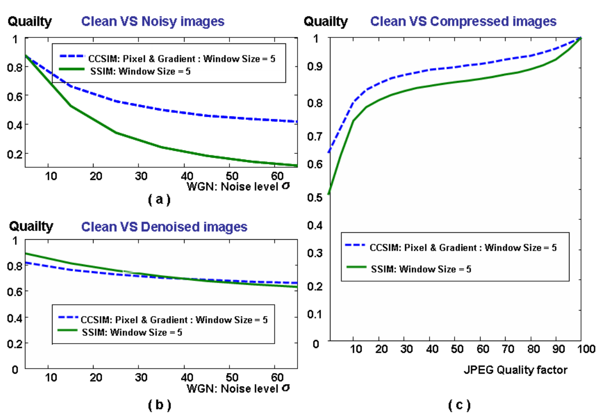

Fig. 1. Quality graphs : (a) with respect to WGNs : Input(Clean VS Noisy images) (b) with respect to WGNs : Output(Clean VS Denoised images by OSA[3]) (c) with respect to Jpeg Quality factor

B. A No-Reference SVD-Based Measure

One way to measure image quality is to analyze the singular values given by a matrix $\boldsymbol{\delta}_i$ constructed from local gradient vectors. Using the singular values s_1 and s_2 of $\boldsymbol{\delta}_i$, the quality measure, defined as

$$R_i = \frac{s_1 - s_2}{s_1 + s_2}, \quad s_1 > s_2, \quad 0 < R_i < 1, \quad (5)$$

indicates the local orientation dominance [4]. If there is one edge at a local region, R_i is close to 1. On the other hand, if the local region is flat or textured (or pure noise), R_i becomes small. We measure the image quality by the mean value of R_i's for all pixels. Fig. 2 shows the mean of R_i's ($\overline{R}$) for a variety of images (Lena, Barbara, Pepper, Cameraman, and pure noise) by adding white Gaussian noise with different PSNR values[1]. As seen in the figure, for the pure noise image, there is no local dominant orientation, and therefore $\overline{R}$ is a the same small constant in for all noise levels. On the other hand, for the ordinary images, since they have many kinds of structures locally (such as flat, edge, corner and texture), $\overline{R}$ has some values. However, when we add strong noise, those structures are buried in noise, and eventually $\overline{R}$ becomes small. In the additive white Gaussian noise case, we have a closed form of the statistics of R_i [4]. Applied to the image itself or the residual image, we can perform a significance test to measure how effectively a filter has removed noise

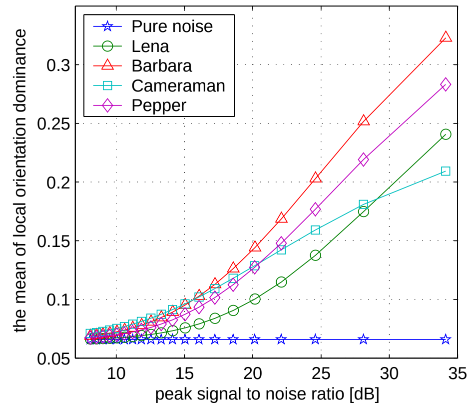

Fig. 2. The mean values of local orientation dominance ($\overline{R}$) for a variety of test images by adding white Gaussian noise with different PSNR values.

III. Conclusions

We briefly introduced two types of quality measures based on the SVD and Canonical Correlations. In the case of CCSIM, we illustrated that patch-based application of this measure yields intuitively correct behavior and can be regarded as a generalization of SSIM[1]. In the SVD-based method, a function of the locally computed singular values of gradients was introduced as a quality measure. Both concepts can be naturally extended to video as well.

References

[1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 600–612, Apr 2004.

[2] G. A. F. Seber, *Multivariate Observations*. New York: John Wiley And Sons, Inc, 1984.

[3] J. Boulanger and C. Kervrann, "Optimal spatial adaptation for patch-based image denoising," *IEEE Transactions on Image Processing*, vol. 15, pp. 2866–2878, 2006.

[4] X. Feng and P. Milanfar, "Multiscale principal components analysis for image local orientation estimation," *Proceedings of the 36th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA*, November 2002.

[1]Peak Signal to Noise Ratio $= 10\log_{10}\left(\frac{255^2}{\text{Mean Square Error}}\right)$ [dB].