2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

A Statistical Evaluation of Sparsity-based Distance Measure (SDM) as an Image Quality Assessment Algorithm

Manasa Priya K. V. S. N. L., Manasa K., Sumohana S. Channappayya
Dept. of Electrical Engineering, IIT Hyderabad, Yeddumailaram, AP - 502205
{EE12M1020, EE12P1002, sumohana}@iith.ac.in

Abstract— **Sparsity-based Distance Measure (SDM), a sparse reconstruction-based image similarity measure was recently proposed and shown to have promising applications in image classification, clustering and retrieval. In this paper, we present a statistical evaluation of SDM's performance as an image quality assessment (IQA) algorithm. This evaluation is carried out on the LIVE image database. We show that the SDM performs fairly in comparison with the state-of-the-art while possessing several attractive properties. Specifically, we demonstrate its robustness to rotation (90^o, 180^o), scaling, and combinations of distortions – properties that are highly desirable of any IQA algorithm.**

I. INTRODUCTION

The vital role that digital multimedia plays in our lives (ranging from medical diagnosis to security to entertainment to online society) is no longer a question for debate but rather a well-accepted norm. This change to our lifestyle has led to a rapid proliferation of multimedia content that needs to be managed (compressed, stored, and communicated) efficiently and effectively. The role of automated or objective multimedia quality assessment to manage multimedia cannot be overemphasized - especially given the cost of subjective evaluation and the massive scale of multimedia data.

Automated or objective image and video quality assessment algorithms have made giant strides in the past decade. The invention of the Structural SIMilarity (SSIM) index [1] heralded a wave of significant improvements in the automatic assessment of image quality and in turn video quality as well. Several excellent full-reference (FR) [2], [3], reduced-reference (RR) [4], and no-reference (NR) [5] image quality assessment (IQA) algorithms have since been proposed. Each of these algorithms take us a step closer to the ultimate goal of being able to mimic the human visual system's assessment of image quality. Given the context of the proposed work, we will restrict our focus to full-reference IQA algorithms.

The underlying principles of the state-of-the-art FR IQA algorithms have ranged from attempting to model the physiology of the human visual system [6] to using abstract notions from information theory [3]. An excellent exposition of these principles can be found in [7]. The success of these varied principles leads one to believe that there could either be several different approaches to solving the FR IQA problem or that these approaches are yet to converge to the true solution. Recent works by Guha et. al. [8], [9], [10] provide yet another approach to measuring image similarity that is based on sparse representations of natural images. This is a promising approach given its close analogy with sparse representations in the human visual system [11]. Their works provide several flavors of sparsity-based image similarity measurement that are tailored to various applications including FR IQA.

In this paper, we consider one such flavor – the SDM, that we feel is intuitively well-suited for FR IQA, and attempt to determine its efficacy. A preliminary evaluation of the SDM as an FR IQA has been carried out by Guha et. al. [8]. The main contributions of this work are: (i) a comprehensive statistical performance evaluation of the SDM on the LIVE image database [12], [13], and (ii) a demonstration of several useful properties of the SDM that make it an attractive FR IQA algorithm.

The paper is organized as follows. Section II defines the SDM and its motivation, Section III describes the statistical evaluation of the SDM, Section IV demonstrates the useful properties of the SDM and Section V concludes the paper and discusses directions for future work.

II. SPARSITY-BASED DISTANCE MEASURE (SDM)

An interesting trend seen in image similarity measurement is the use of Kolmogorov complexity-inspired [14] formulations. An information distance between the two strings x and y can be defined as $max\{K(x|y), K(y|x)\}$ where $K(x|y)$ is the Kolomogorov complexity of x relative to y and vice-versa for $K(y|x)$. To convert it to a normalized symmetric metric, a novel normalized information distance (NID) measure was defined by Li et. al. [15] as follows:

$$NID(x,y) = \frac{max\{K(x|y), K(y|x)\}}{max\{K(x), K(y)\}} \quad (1)$$

where $K(x|y)$ is the conditional Kolmogorov complexity of x relative to y. While NID has nice analytical properties, it is not practical since computing the Kolmogorov complexity is an NP-hard problem. Recent methods attempt to approximate Kolmogorov complexity using quantities that can be computed using fast algorithms. To the best of our knowledge, the first such approach to measure image similarity was introduced by Nikvand et. al. [16] where the size of the encoded bitstream from a lossless image coder was used to approximate Kolmogorov complexity.

Guha et.al. [8] related sparsity and Kolomogorov complexity based on the inference that the number of components required to represent a signal increases with signal complexity. The SDM was then defined to measure image similarity as follows.

$$SDM(X,Y) = \frac{N(X|Y) + N(Y|X)}{N(X) + N(Y)}. \qquad (2)$$

where X is the reference image; Y is the test image; $N(X)$ and $N(Y)$ represents the number of components required to represent the image from the dictionary learnt from the patches of X and Y respectively. $N(X|Y)$ and $N(Y|X)$ represent the number of components required to represent the image from the dictionary learnt from the patches of Y and X respectively. $N(X) < N(X|Y)$; since number of components required to represent the current image X from the dictionary learnt from the patches of X is always less than the dictionary learnt from the patches of Y. Hence lower values of SDM indicate better similarity between images under consideration and is always greater than or equal to one.

In our work, the SDM has been implemented using the K-SVD algorithm [17] to find $N(X), N(Y)$ and the cross term $N(X|Y), N(Y|X)$. A randomly chosen set of 3000 $8{\times}8$ images patches were used for learning a dictionary containing 128 atoms [8].

III. STATISTICAL EVALUATION

One of the main contributions of this work is to perform a statistical evaluation of the SDM as an FR IQA algorithm. The results of the statistical evaluation and an intuitive explanation of the performance are presented in the following subsections.

A. Evaluation

The SDM was evaluated over the LIVE database [12] that consists of 779 images covering a range of 5 types of distortions. There are 29 reference images and distortion types include fast fading, white noise, JPEG, JPEG 2000 and gaussian blur. SDM is compared with the state of art full reference algorithms such as SSIM [1], MSSSIM [2] and VIF [3]. The SDM scores were fit to the subjective scores (DMOS) using the four parameter exponential logistic function specified in [18].

The results of the statistical evaluation are presented in Fig. 1 and Table 2. Fig. 1 shows the scatter plots for each of the distortion types in the database along with an overall scatter plot. It is clear that the SDM performs best when the distortion type is either blur or additive white noise and performance drops for JPEG and fast fading distortions. We present an intuitive explanation for this performance in the following subsection. From Table 2, we see that the SDM performs fairly when compared to the state-of-the-art using Spearman Rank Ordered Correlation Coefficient (SROCC). However, we show in Section IV that the SDM has several useful properties that the state-of-the-art IQA algorithms lack thereby making the SDM a very promising IQA algorithm.

B. Intuition

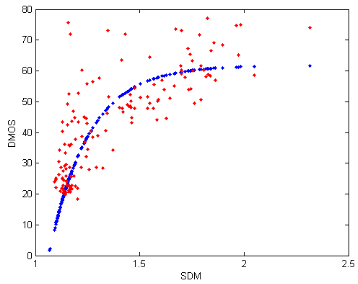
We present an intuitive explanation for the performance of the SDM using images distorted with white noise. Fig. 3 and Table 4 corroborate the inference made in [8] about requiring a large number of dictionary elements to represent complex signals (for e.g., images corrupted with noise). The loss in sparsity is clearly seen in Table 4. As the noise variance increases, $N(Y)$, $N(Y|X)$, and $N(X|Y)$ increase suggesting (expectedly) that noise cannot be sparsely represented. We also observed the opposite effect for blurred images i.e., a decrease in the aforementioned quantities. The other distortion types (fast fading, JPEG and JPEG2000) do not bring about changes to the images that significantly affect their sparsity, thereby explaining the SDM's average performance for these distortions. These qualitative observations combined with the statistical evaluation in the previous subsection suggest that the SDM is able to quantify departure of images from "naturalness" that correlates fairly well with subjective evaluation.

IV. SALIENT PROPERTIES OF THE SDM

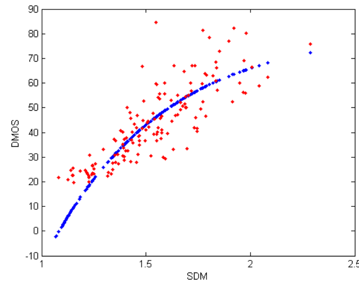
In this section, we demonstrate salient properties of the SDM that make it a very promising IQA algorithm and distinguish it from the state-of-the-art IQAs. Specifically, we demonstrate SDM's robustness to rotation, scaling, and combinations of distortions. The top row of Fig. 5 shows various distortions and corresponding SDM, VIF, and MSSSIM scores. It is clear from Figs. 5b, 5c, and 5d that the SDM outperforms both VIF and MSSSIM for the mentioned distortion types. Fig. 5 is an illustrative example of the robustness that has been consistently observed over a much larger dataset. From Fig. 5c, it is worth highlighting that unlike MSSSIM and VIF, the SDM does not require the reference and distorted image sizes to match. It is to be noted that a score close to 1 means low distortion for all the algorithms considered.

We present an empirical explanation of the robustness of SDM to rotation, scaling, and combinations of distortions. The bottom row of Fig. 5 shows the histogram of the maximum pair-wise correlation between the atoms of the reference and distorted image dictionaries. Let D_R and D_D be the reference and distorted dictionaries respectively. Let $D_R = [a_1^R, a_2^R, \ldots, a_{128}^R]$, $D_D = [a_1^D, a_2^D, \ldots, a_{128}^D]$ where a_i^R is the column vector of size 64 representing the i^{th} atom of the reference dictionary D_R, and a_j^D is the column vector of size 64 representing the j^{th} column of D_D. We construct the correlation matrix R where R_{ij} is the correlation between the a_i^R and a_j^D. The maximum value of row i in R represents the best matching atom in D_D to a_i^R. The histograms in the bottom row of Fig. 5 correspond to the row-wise maximum correlation for each distortion type. Note that the histograms in Fig. 5 correspond to the images directly above them.

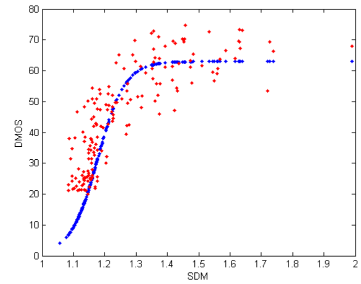
From these histograms, we see that there are a large number of atom-pairs with high correlation (> 0.9) for the distortions where SDM is robust. This can be interpreted as the dictionaries D_R and D_D being composed of atoms that are "similar". This in turn implies that the cross terms in



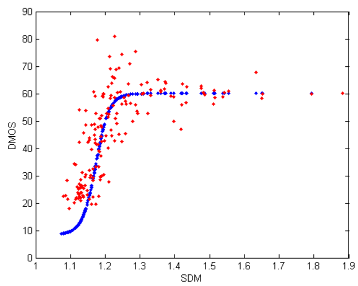
(a) Fast fading.



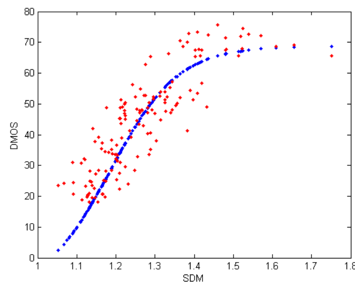
(b) Gaussian blur.



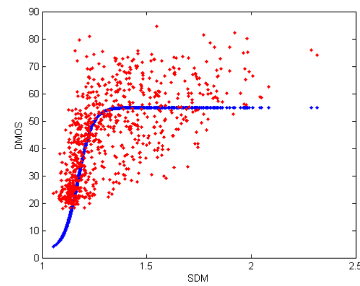
(c) JPEG2000.



(d) JPEG.



(e) White noise.



(f) All distortions.

Fig. 1: Scatter plots of SDM vs DMOS for various distortions. The blue line represents the best fit function.

the SDM index $(N(X|Y), N(Y|X))$ would be similar and therefore the robustness of the SDM.

V. CONCLUSIONS AND FUTURE WORK

We have presented a statistical evaluation of the SDM and shown that it performs fairly when compared to the state-of-the-art. However, we have shown that the SDM possesses several useful properties such as robustness to rotation, scaling and distortion combinations that make it appealing in a wider variety of applications than most popular full-reference image quality assessment algorithms. The strength of the SDM as an objective function has already been demonstrated in image classification, clustering and retrieval applications [8].

We believe that the SDM opens up interesting avenues for further investigation in the measurement of image similarity with potential extensions to video similarity as well. As future work, we plan to explore these avenues with a particular emphasis on video similarity measurement.

REFERENCES

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

[2] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Signals, Systems and Computers, 2003. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 2, pp. 1398–1402, IEEE, 2003.

[3] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *Image Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 430–444, 2006.

[4] R. Soundararajan and A. C. Bovik, "Rred indices: reduced reference entropic differencing for image quality assessment," *Image Processing, IEEE Transactions on*, vol. 21, no. 2, pp. 517–526, 2012.

[5] A. Mittal, A. Moorthy, and A. Bovik, "No-reference image quality assessment in the spatial domain," 2012.

[6] J. Lubin, "A human vision system model for objective picture quality measurements," 1997.

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

[8] T. Guha and R. K. Ward, "Image similarity using sparse representation and compression distance," *arXiv preprint arXiv:1206.2627*, 2012.

[9] T. Guha and R. K. Ward, "Image similarity measurement from sparse reconstruction errors," in *Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on*, 2013.

[10] T. Guha, E. Nezhadarya, and R. K. Ward, "Sparse representation-based image quality assessment," *arXiv preprint arXiv:1306.2727*, 2013.

[11] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by v1?," *Vision research*, vol. 37, no. 23, pp. 3311–3325, 1997.

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

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

[14] M. Li, *An introduction to Kolmogorov complexity and its applications*. Springer, 1997.

[15] M. Li, X. Chen, X. Li, B. Ma, and P. M. Vitányi, "The similarity metric," *Information Theory, IEEE Transactions on*, vol. 50, no. 12, pp. 3250–3264, 2004.

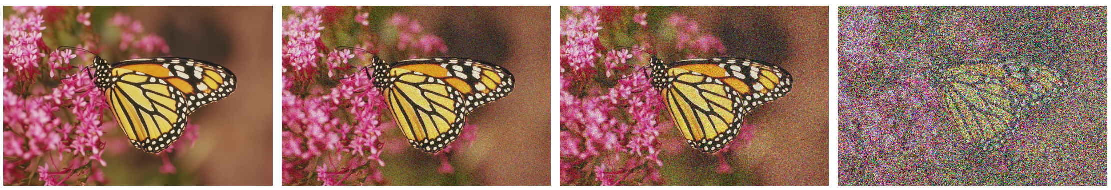
[16] N. Nikvand and Z. Wang, "Generic image similarity based on kolmogorov complexity," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pp. 309–312, IEEE, 2010.

[17] M. Aharon, M. Elad, and A. Bruckstein, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.

[18] A. M. Rohaly, P. J. Corriveau, J. M. Libert, A. A. Webster, V. Baroncini, J. Beerends, J.-L. Blin, L. Contin, T. Hamada, D. Harrison, *et al.*, "Video quality experts group: Current results and future directions," in *Visual Communications and Image Processing 2000*, pp. 742–753, International Society for Optics and Photonics, 2000.

	FF	Blur	JPEG	JP2K	AWGN	All
SSIM	0.9629	0.9481	0.9266	0.8711	0.9903	0.9298
MSSSIM	0.8499	0.9274	0.9445	0.962	0.9865	0.924
VIF	0.9587	0.976	0.9025	0.9355	0.8852	0.8677
SDM	0.8277	0.9102	0.8188	0.843	0.8913	0.7885

Fig. 2: Performance of the SDM on the LIVE image database measured using SROCC. Also shown are state-of-the-art IQA algorithms.



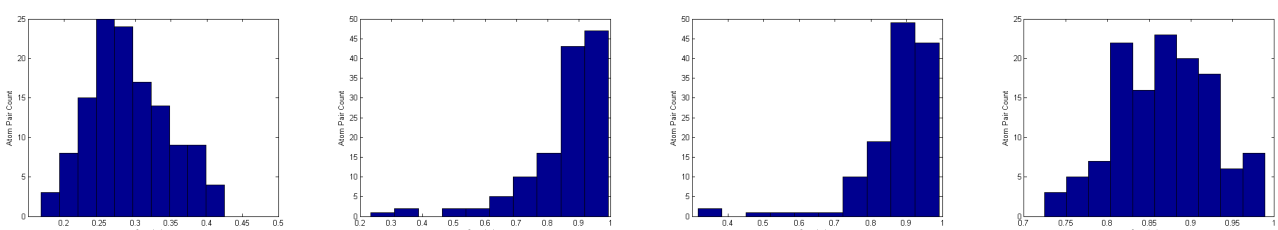
(a) Original. (b) σ = 0.117. SDM = 1.84. (c) σ = 0.187. SDM = 1.95. (d) σ = 1.0. SDM = 2.39.

Fig. 3: Intuition behind SDM's performance illustrated using noisy images.

Noise σ	$N(X)$	$N(Y)$	$N(X\|Y)$	$N(Y\|X)$	SDM
0.117	8.441	23.3747	15.2877	43.1803	1.8377
0.187	8.4407	23.9843	16.6767	46.654	1.9531
1.0	8.2673	24.0303	26.8163	50.3037	2.3878

Fig. 4: An intuitive explanation of the SDM's ability to measure image similarity. These values correspond to the images in Fig. 3.

(a) AWGN. SDM = 2.0765, VIF = 0.0303, MSSSIM = 0.0024. (b) Rotation. SDM = 1.1026, VIF = 0.0158, MSSSIM = 0.0473. (c) Scaling down. SDM = 1.1311, VIF, SSIM require size match. (d) Combo. SDM = 2.2032, VIF = 0.0075, MSSSIM = 0.1725.



(e) White noise. (f) Rotation by 180^o. (g) Scaled down by 0.8. (h) FF and rotation by 180^o.

Fig. 5: Robustness of the SDM to rotation, scaling, and a combination of distortions. Top row showing various distortion types. Bottom row showing histogram of maximum correlation between atom pairs formed from reference and distorted image dictionaries.