Permutation Entropy Based Full-Reference Image Quality Assessment

Mohtashim Baqar*, Sian Lun Lau* and Mansoor Ebrahim†
Department of Information Systems and Computing, Sunway University, Malaysia *
Iqra University (Main Campus), Karachi, Pakistan†
mohtash.b@imail.sunway.edu.my, *sianlunl@sunway.edu.my*, *mebrahim@iqra.edu.pk*

***Abstract*—Due to the increasing proliferation of multimedia signals, specifically, image, video and their applications in our daily life, it is indispensable to have methods that can efficiently predict the visual quality of images with high measures of accuracy. Image processing procedures often introduce undesirable distortion in images that require fixing; preferably consistent with a human visual system (HVS). Therefore, an image quality assessment(IQA) framework should be highly accurate as well as computationally efficient; making it viable to be used with different image processing applications, especially, with real-time applications. Motivated by the need of appropriate objective models, we propose a novel objective IQA algorithm, namely, Permutation Entropy Deviation Index (PEDI), based on the working principle of permutation entropy (PE). Permutation entropy helps in detecting and visualizing changes related to structures with the correlation between successive samples instead of considering magnitudes of the signal, and since, perception of an image to the HVS changes more because of structural changes in an image rather than that of visible errors i.e. MSE. Therefore, in this work, we have exploited this property to predict image quality efficiently. Moreover, entropy itself is sensitive to variations, whereas the permutation entropy captures pattern variations in an image.**

Furthermore, each local patch in the distorted image undergoes a different level of distortion due to structural differences. This motivates us to use permutation entropy to exploit the global variations in the local quality map for image quality assessment. With standard deviation as the pooling strategy, we observed that permutation entropy between reference and distorted images could predict image quality with high measures of accuracy. Experimental results on a subjective database, CSIQ, have shown that the proposed model outperforms most of the existing STOA image quality assessment models and highly correlates with subjective judgements. [1]**.**

***Index Terms*—Image quality assessment, permutation entropy, visual perception, full reference, standard deviation pooling.**

I. Introduction

Usage of multimedia signals and their applications have increased tremendously in our daily life. These visual signals get contaminated with different types of distortions during the acquisition, transmission and reception process. Most importantly, human vision is the ultimate receiver of these signals. Therefore, it is an essential task of a modern day image processing system to produce high-quality visuals compatible with human perception.

[1]The MATLAB code of the proposed model will be made available online after thorough performance evaluation.

Several calibrated databases have been developed to explore areas related to image quality assessment (IQA) and its applications. The most popular full-reference IQA metric is mean square error (MSE), calculated by averaging the square root difference of intensities of two images. It is rather simple to calculate and have well defined physical meanings, however, its results are not consistent with human visual perception [1]–[9]). For years now, many efforts have been put into place to develop image quality assessment methods consistent with the human visual system (HVS). Generally, quality assessment models are categorized as full-reference, reduced reference and no reference IQAs. In FR-IQA, the reference image is available. In RR-IQA, partial information or features of the reference image are made available whereas, in NR-IQA, there is no reference image available [10]. Many FR-IQA models have been developed using a two-step common framework; calculation of local quality map followed by pooling strategy to predict the final quality score of an image. [11]–[15]. For our proposed model, we have used this two-step framework with permutation entropy (PE) being used as the similarity function for local quality map formation, followed by standard deviation as the pooling mechanism.

In this work, we have utilized the property of entropy in capturing structural changes, we have used it as a similarity measure in our proposed model to build the local quality map. Rest of the paper is organized as follows. In Sect.II, a more detailed description for PE is provided including an illustrative example. Also, complete algorithmic details of the proposed model are also presented in this section. The experimental results with details of database and performance evaluation measures are presented in Sect.III. The paper is concluded in Sect.IV.

II. Methodology

In this section, the algorithmic details of the proposed criterion, PEDI, for IQA are presented. Moreover, the section starts with brief details of permutation entropy, followed by the explanation of the local quality map, pooling strategy and lastly, complete details of the proposed model are discussed.

A. Permutation Entropy

Permutation entropy [16] is a robust complexity measure works on the principle of correlation between neighbouring elements in a sequence. In recent years, entropy and its variants,

i.e. sample entropy, approximate entropy, maximum relevance and permutation entropy, have been used for feature extraction, noise removal and anomaly detection etc., in various areas of biological signal processing, computer science and electrical engineering [17]–[19]. PE, in particular, has been used a lot with EEG signals to capture and detect changes in epileptic seizures, in analyses and classification of heart rate variability data, in discriminating sleeping stages, and in distinguishing brain states etc. [20]–[22]. It takes into consideration the structural variations emanating from successive elements in a sequence. Initially, a permutation order is set and based on that the input sequence is encoded and arranged into symbols [23]. Then, the normalized permutation entropy is computed for the coded sequences as defined in equation 1 [16],

$$PE_{nor} = -\frac{1}{(n-1)} \sum_{c=1}^{d!} p_c log_2(p_c) \tag{1}$$

where, d and p_c reflects the order of permutation and relative frequencies for all permutation possibilities, respectively. Another important parameter used in calculating p_c is τ, which indicates the delay between two successive samples in the sequence.

Moreover, complete details of PE, including, algorithmic explanation and recommendations for selection of input parameters can be found in [16]. The examples in fig. 1 and 2 illustrate how PE works on real-world problems. An utterance of an underwater acoustic sonar sequence [24] of a ship is considered. The signal is of 9-second duration, sampled at 44.1 kHz. It is analyzed at $\tau = 1$, means one sample shift per window with each window having 512 samples.

Moreover, an approximate of 397K windows during 9 seconds were processed in less than a second, which implies that permutation entropy can be useful in time-constrained applications. Since each window contained samples less than 6!, therefore, the permutation order was set to be $d = 5$. Furthermore, additive white Gaussian noise (AWGN) was added to the signal at an SNR of 15 dB to keep the difference between the original sequence and noisy sequence to a bare minimum. The noise was added to the signal after a time lapse of 5 seconds, it can also be seen in the two plots of fig. 1. Normalized permutation entropy was calculated for both the original and noisy sequences. Though the magnitude of noise added to the sequence is very low, but even then the drastic change in entropy values of both the sequences can be observed clearly from fig. 2.

In our proposed method, permutation entropy is calculated for both the reference and distorted images in horizontal and vertical directions. The computed directional entropy of both the images are combined respectively to form quality maps for the two respective images i.e. m_{ref} and $mdis$. Further mathematical details are presented in Algorithm 1.

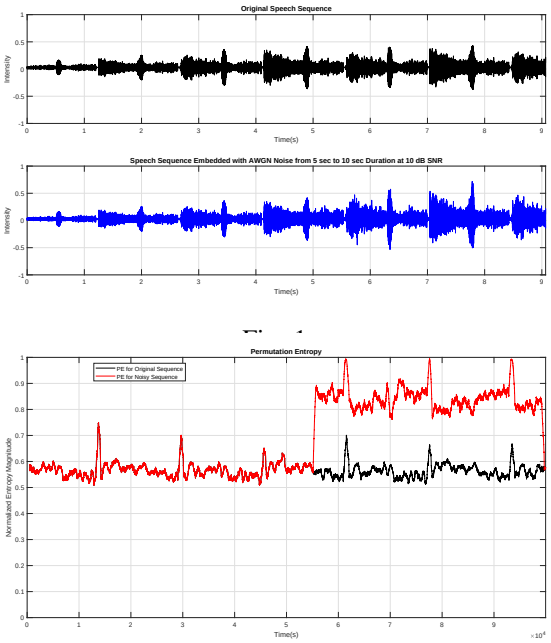


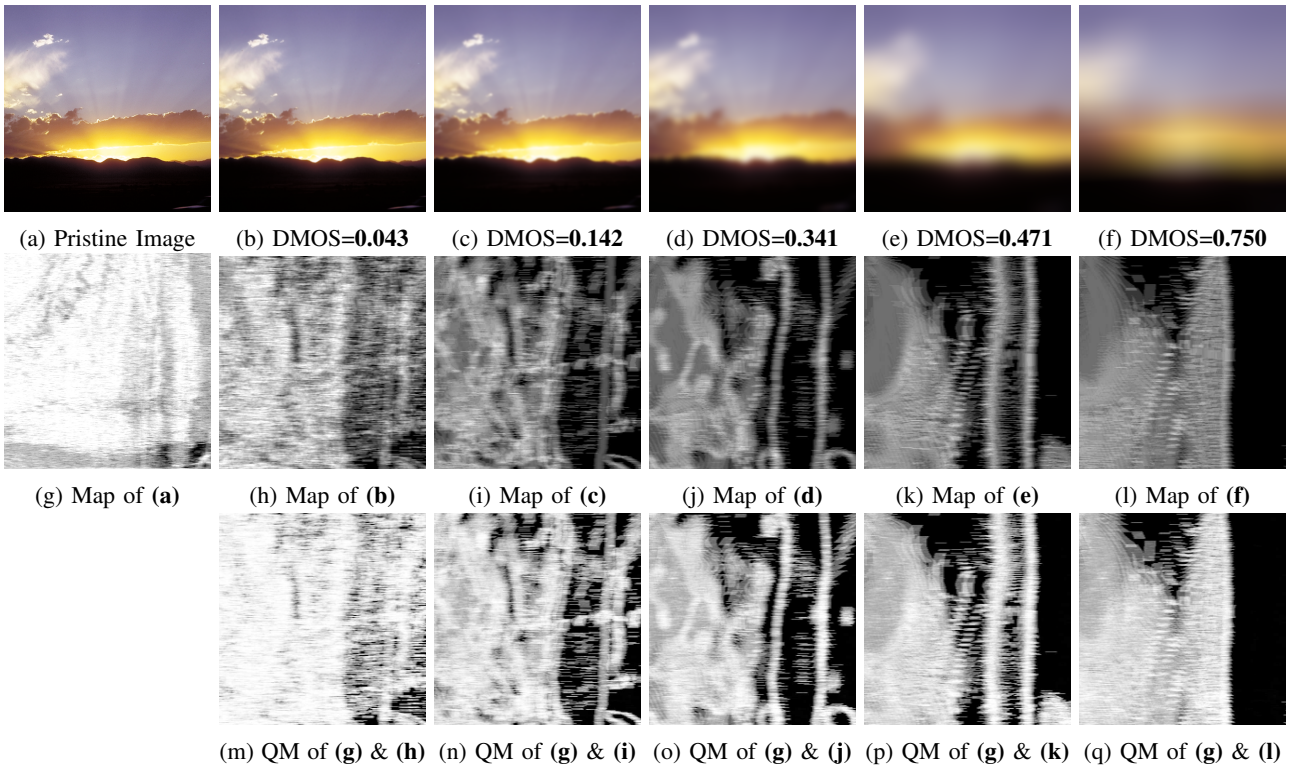
Fig. 2

Algorithm 1 Permutation Entropy Deviation Index (PEDI)

Inputs: $\underbrace{Ref.\ Img.}_{\text{GS},R(.)\in[0,1]} R(x,y)$, $\underbrace{Dis.\ Img.}_{\text{GS},D(.)\in[0,1]} D(x,y)$, Order $d = 3$, Delay $\tau = 1$, Window Size $w_n = 4$
Output: Quality Score $Q_s \in [0,1]$
function $PEDI(I(x,y), d, \tau), w_n$
Initialize: $T \leftarrow size_{row\ or\ col}(R(x,y))$, $\eta = 0.05$, Total Perm. Patterns., $\pi_j, \leftarrow for\ d;\ (j = 1, \ldots, d!)$
for $t = 1\ to\ T$ **do**
  $Seq_{R_r} \leftarrow R(t,:)$, $Seq_{R_c} \leftarrow R(:,t)$, $Seq_{D_r} \leftarrow D(t,:)$ $and\ Seq_{D_c} \leftarrow D(:,t)$
  for $i = 1\ to\ T \Rightarrow Step \rightarrow w_n$ **do**
    Calculate Rank:(of all Seq(s) in 6), which leads to,$\{r_i, \ldots, r_{i+n-1} \leftarrow x_i, \ldots, x_{i+n-1}\}$.Where ranks, r_i, are indices of values, x_i sorted in ascending order.
    if $\underbrace{Compare(\pi_k, r_i)}_{\{k=1,\ldots,d!\},\{i=1,\ldots,i+n-1\}}$ **then**
      $Z_k = Z_k + 1$
    end if
    Compute: $\underbrace{Prob.}_{\forall \pi_j}$, $p'_j \leftarrow z_j / \sum z_k$
    Select $p'_j > 0$, $\forall Seq_{(R_r,R_c,D_r,D_c)}$
    Compute: $(I_{x_{ref}}, I_{y_{ref}}, I_{x_{dis}}, I_{y_{dis}}) = \frac{-1}{(d-1)} \sum_{j=1}^{d!} p'_j log_2(p'_j)$
  end for
end for
Compute: $m_{ref} \leftarrow \sqrt{I^2_{x_{ref}} + I^{2\,\prime}_{y_{ref}}}$
$m_{dis} \leftarrow \sqrt{I^2_{x_{dis}} + I^{2\,\prime}_{y_{dis}}}$
Compute: $LQM \leftarrow \frac{2m_{ref} \times m_{dis} + \eta}{m_{ref}.^2 + m_{dis}.^2 + \eta}$ (See Eq:2)
Compute: $Q_s \leftarrow STD(LQM)$ (See Eq:3)

B. Local Quality Map and Pooling Strategy

Local quality map (LQM) reflects deformity levels of small regions/ patches in a distorted image. After calculation of PE



(a) Pristine Image (b) DMOS=**0.043** (c) DMOS=**0.142** (d) DMOS=**0.341** (e) DMOS=**0.471** (f) DMOS=**0.750**

(g) Map of **(a)** (h) Map of **(b)** (i) Map of **(c)** (j) Map of **(d)** (k) Map of **(e)** (l) Map of **(f)**

(m) QM of **(g)** & **(h)** (n) QM of **(g)** & **(i)** (o) QM of **(g)** & **(j)** (p) QM of **(g)** & **(k)** (q) QM of **(g)** & **(l)**

Fig. 3: **(a)-(f)** Example of reference **(r)** and Gaussian Blur (GB) distorted **(d)** images taken from the Computational and Subjective Image Quality (CISQ) [25] database with respective DMOS scores indicating different levels of distortion. **(g)-(l)** Permutation entropy magnitude images (m_{ref} and m_{dis}) of reference and distorted images illustrated in first row. **(m)-(q)** Illustrates the associated quality maps between permutation entropy map of reference and distorted images, where brighter grey-level means higher similarity. For each presented distorted image from left to right, the calculated objective quality scores with **PEDI** are **0.1877**, **0.3218**, **0.3698**, **0.3818** and **0.3878**, respectively. Further, the values for the three performance evaluation criterion; SRC, PCC and RMSE, for these set of images were calculated to be $\mathbf{1}$, $\mathbf{1}$, and $\mathbf{1.2428 \times 10^{-16}}$, respectively

map for both reference m_{ref} and distorted m_{dis} images, LQM can be calculated as in equation 2,

$$LQM(i,j) = \sum_{\forall i}\sum_{\forall j} \frac{2m_{ref}(i,j)m_{dis}(i,j) + \eta}{m^2_{ref}(i,j) + m^2_{dis}(i,j) + \eta} \tag{2}$$

Pooling strategy is applied to LQM to quantify overall image quality score. Most algorithms use mean or weighted mean as the pooling technique. However, as we know that each local patch in an image undergoes different levels of distortion because of structural differences. So, averaging would neutralize the actual deformity level in an image. Therefore, we have used standard deviation as the pooling mechanism, as it will be able to capture the global variation of distortion levels in local patches of the distorted. The pooling operation on LQM can be mathematically defined as in equation 3,

$$Q_s = \left(\frac{1}{\sum_{\forall i} 1 \sum_{\forall j} 1}\right) \sum_{\forall i}\sum_{\forall j} LQM(i,j) \tag{3}$$

III. Experimental Results

Quality assessment (QA) research depends upon calibrated data as well as good testing mechanisms to make predictions consistent with human observers. Subjective experiments are conducted in controlled environment to develop databases for QA research. In this section, thorough experimental results for the proposed algorithm on a publicly available subjective database i.e. Computational and Subjective Image Quality (CSIQ) Image Quality Database [25], have been presented.

A. Database

In this study, experiments have been conducted using the CSIQ database. It is the second largest subjective database available till date. It has 30 pristine reference and 866 distorted images. The distorted set of images were generated by applying six type of impairments to the reference images at different intensity levels. Namely, the distortion types are; additive white noise (AWN), JPEG2000 compression, global contrast decrements (CTD), additive pink Gaussian noise (APDN), JPEG compression and Gaussian blur (GB). Furthermore, every IQA database is evaluated by human subjects under controlled settings and based on their assessments, a quality score is assigned to each image i.e. difference mean-opinion score (DMOS). These assigned scores act as a reference in evaluation and validation of quality assessment models.

B. Performance Evaluation Criteria

Performance evaluation of the proposed model, PEDI, has been made based on three criterion; Pearson linear correlation coefficient (PCC), Spearman rank-order correlation (SRC) coefficient and root mean square error (RMSE). Moreover, SRC measures the prediction monotonicity, while PCC evaluates the prediction accuracy. Index values for both the criterion are in the range of $[0, 1]$, with higher correlation value indicating greater coherence with subjective assessments. At first, non-linear regression is performed using a five parameter logistic function as in equation 4, to map predicted objective scores to subjective human scores.

$$Q(S) = q_1(\frac{1}{2} - \frac{1}{1 + e^{(q_2 s - q_3)}}) + q_4 s + q_5 \tag{4}$$

where s and $Q(s)$ are the subjective and mapped-objective scores, respectively. Whereas, the coefficients, $q_i (i = 1, 2, 3, 4, 5)$ are calculated during curve fitting.

After regression, the three aforementioned performance evaluation measures are computed; Pearson linear correlation coefficient (PCC) is used to evaluate the prediction accuracy between Q_p and S and is defined as in equation 5,

$$PCC(Q_p, S) = \frac{\overline{Q}_p^T \overline{S}}{\sqrt{\overline{Q}_p^T \overline{Q}_p \overline{S}^T \overline{S}}} \tag{5}$$

where Q, Q_p, S, $\overline{Q}_p$ and $\overline{S}$ are the objective scores, objective scores after regression, subjective scores, zero-mean objective scores and zero-mean objective scores after regression, respectively.

The second index, Spearman rank-order correlation (SRC) coefficient is used to evaluate the prediction accuracy between Q and S and is defined as in equation 6,

$$PCC(Q, S) = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{6}$$

where n indicates the total number of samples and d is the calculated difference between the ranks of each pair of samples in Q and S.

The third coefficient, root mean square error (RMSE) is used to evaluate the prediction consistency of an IQA model. It is calculated between Q_p and S, and is defined as in equation 7,

$$RMSE(Q_p, S) = \sqrt{(Q_p - S)^T (Q_p - S)/n} \tag{7}$$

To demonstrate performance of the proposed model, PEDI, a comparison with 12 STOA IQA algorithms have been made. Including, PSNR, SSIM [15], MS-SSIM [26], FSIM [14], IFC [27], VIF [28], MAD [25], IW-SSIM [29], G-SSIM [13], GSD [12], GS [30] and GMSD [31].

C. Implementation and Performance Comparison

Simulations presented in this paper have been conducted in MATLAB. In our proposed algorithm, we have used a constant, η, to ensure numerical stability. We set $\eta = 0.05$, after thorough simulation trials. In table I, performance of the STOA IQA models are compared with the proposed model in terms of SRC, PCC and RMSE scores. For each criterion, top three performing IQA models are shown in boldface. From table, it can be seen that the top three IQA models are **GMSD**, **MAD** and **PEDI** (proposed model). It can also be observed that the proposed model PEDI outperforms most of the STOA FQ-IQA models based on all three criteria (SRC, PCC and RMSE) including the models dealing with information content of an image.

TABLE I: Comparison of the Proposed PEDI model with other STOA FR-IQA models. The FR-IQA models were compared based on SRC, PCC and RMSE scores. Further, simulations were conducted on CISQ Database. Moroever, top three models in each category are highlighted with boldface.

IQA Methods	CSIQ Database - 866 Images		
	SRC	PCC	RMSE
PSNR	0.806	0.751	0.173
IFC	0.767	0.837	0.144
GSD	0.854	0.854	0.137
G-SSIM	0.872	0.874	0.127
SSIM	0.876	0.861	0.133
VIF	0.919	0.928	0.098
GS	0.911	0.896	0.116
MS-SSIM	0.913	0.899	0.115
MAD	**0.947**	**0.950**	**0.082**
GMSM	**0.929**	0.913	0.107
IW-SSIM	0.921	0.914	0.106
FSIM	0.924	0.912	0.108
GMSD	0.957	**0.954**	**0.079**
PEDI	0.916	**0.930**	**0.096**

IV. Conclusion

In this study, a novel permutation entropy based FR-IQA model has been proposed, named, Permutation Entropy Deviation Index (PEDI). We have demonstrated the usefulness and effectiveness of permutation entropy-based model in capturing local structural variations between corresponding images. As each local patch in an image has diverse structural features, this attribute causes each patch to suffers a different level of degradations. Therefore, keeping the aforementioned in consideration, the standard deviation was used as the pooling

mechanism to quantity global variations in the PE map to predict an image quality score. Simulation results based on SRC, PCC and RMSE coefficients have shown that the algorithm is both efficient & accurate. Moreover, in terms of the three criteria as mentioned above, PEDI is in top three models for PCC and RMSE index values, whereas the SCC index value was also found to be competitive to almost all the STOA IQA models. These outcomes reiterate the fact that PEDI has the potential to be used in a wide range of image processing applications. The future work for this research will move into performance evaluation on more subjective databases.

REFERENCES

[1] B. GIROD, "What's wrong with mean-squared error?" *Digital Images and Human Vision*, pp. 207–220, 1993. [Online]. Available: https://ci.nii.ac.jp/naid/10026807239/en/

[2] P. C. Teo and D. J. Heeger, "Perceptual image distortion," in *Human Vision, Visual Processing, and Digital Display V*, vol. 2179. International Society for Optics and Photonics, 1994, pp. 127–142.

[3] A. M. Eskicioglu, P. S. Fisher, and S.-Y. Chen, "Image quality measures and their performance," 1994.

[4] M. P. Eckert and A. P. Bradley, "Perceptual quality metrics applied to still image compression," *Signal processing*, vol. 70, no. 3, pp. 177–200, 1998.

[5] S. Winkler, "Perceptual distortion metric for digital color video," in *Human Vision and Electronic Imaging IV*, vol. 3644. International Society for Optics and Photonics, 1999, pp. 175–185.

[6] B. L. Evans and W. S. Geisler, "Rate scalable foveated image and video communications," 2001.

[7] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE signal processing letters*, vol. 9, no. 3, pp. 81–84, 2002.

[8] Z. Wang, "Demo images and free software for a universal image quality index," 2001.

[9] Z. Wang, A. C. Bovik, L. Lu *et al.*, "Why is image quality assessment so difficult?" in *ICASSP*, vol. 4, 2002, pp. 3313–3316.

[10] Z. Wang and A. C. Bovik, "Modern image quality assessment (synthesis lectures on image, video, and multimedia processing)," *Morgan & Claypool Publishers*, 2006.

[11] D.-O. Kim, H.-S. Han, and R.-H. Park, "Gradient information-based image quality metric," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 930–936, 2010.

[12] G. Cheng, J. Huang, C. Zhu, Z. Liu, and L. Cheng, "Perceptual image quality assessment using a geometric structural distortion model," in *2010 IEEE International Conference on Image Processing*. IEEE, 2010, pp. 325–328.

[13] G.-H. Chen, C.-L. Yang, and S.-L. Xie, "Gradient-based structural similarity for image quality assessment," in *2006 International Conference on Image Processing*. IEEE, 2006, pp. 2929–2932.

[14] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," *IEEE transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.

[15] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli *et al.*, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[16] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical review letters*, vol. 88, no. 17, p. 174102, 2002.

[17] Aaron Raymond See and Chih-Kuo Liang, "A study on sleep eeg using sample entropy and power spectrum analysis," in *2011 Defense Science Research Conference and Expo (DSR)*, Aug 2011, pp. 1–4.

[18] R. Tiwari, V. K. Gupta, and P. Kankar, "Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier," *Journal of Vibration and Control*, vol. 21, no. 3, pp. 461–467, 2015. [Online]. Available: https://doi.org/10.1177/1077546313490778

[19] K. El-Darymli, E. W. Gill, C. Moloney, P. McGuire, and D. Power, "Permutation entropy for signal analysis: A case study of synthetic aperture radar imagery," in *2015 IEEE 14th Canadian Workshop on Information Theory (CWIT)*, July 2015, pp. 66–70.

[20] "Ordinal pattern based similarity analysis for eeg recordings," *Clinical Neurophysiology*, vol. 121, no. 5, pp. 694 – 703, 2010.

[21] D. Li, X. Li, Z. Liang, L. J. Voss, and J. W. Sleigh, "Multiscale permutation entropy analysis of EEG recordings during sevoflurane anesthesia," *Journal of Neural Engineering*, vol. 7, no. 4, p. 046010, jun 2010. [Online]. Available: https://doi.org/10.1088%2F1741-2560%2F7%2F4%2F046010

[22] N. Nicolaou and J. Georgiou, "The use of permutation entropy to characterize sleep electroencephalograms," *Clinical EEG and Neuroscience*, vol. 42, no. 1, pp. 24–28, 2011, pMID: 21309439.

[23] C. S. Daw, C. E. A. Finney, and E. R. Tracy, "A review of symbolic analysis of experimental data," *Review of Scientific instruments*, vol. 74, no. 2, pp. 915–930, 2003.

[24] K. J. V. RAPOSA, G. SCOWCROFT, C. KNOWLTON, and P. F. WORCESTER, "The discovery of sound in the sea web site: An educational resource," *Bioacoustics*, vol. 17, no. 1-3, pp. 348–350, 2008. [Online]. Available: https://doi.org/10.1080/09524622.2008.9753873

[25] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, p. 011006, 2010.

[26] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2. Ieee, 2003, pp. 1398–1402.

[27] H. R. Sheikh, A. C. Bovik, and G. De Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Transactions on image processing*, vol. 14, no. 12, pp. 2117–2128, 2005.

[28] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3. IEEE, 2004, pp. iii–709.

[29] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp. 1185–1198, 2011.

[30] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1500–1512, April 2012.

[31] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684–695, Feb 2014.