MS-UNIQUE: Multi-model and Sharpness-weighted Unsupervised Image Quality Estimation

Mohit Prabhushankar*[1]*, Dogancan Temel*[1]*, Ghassan AlRegib*[1]*;
[1]*Center for Signal and Information Processing, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA*

Abstract

In this paper, we train independent linear decoder models to estimate the perceived quality of images. More specifically, we calculate the responses of individual non-overlapping image patches to each of the decoders and scale these responses based on the sharpness characteristics of filter set. We use multiple linear decoders to capture different abstraction levels of the image patches. Training each model is carried out on 100,000 image patches from the ImageNet database in an unsupervised fashion. Color space selection and ZCA Whitening are performed over these patches to enhance the descriptiveness of the data. The proposed quality estimator is tested on the LIVE and the TID 2013 image quality assessment databases. Performance of the proposed method is compared against eleven other state of the art methods in terms of accuracy, consistency, linearity, and monotonic behavior. Based on experimental results, the proposed method is generally among the top performing quality estimators in all categories.

Introduction

With the advent of social media and faster wireless networks, high quality digital images are one of the most popular forms of multimedia being shared online. Infact, on an average day, billions of photos are shared through dedicated platforms. It is essential for these platforms to maintain high standards in acquiring, compressing, transmitting, and displaying these images without compromising it's visual quality to the end user. Such a task cannot be manually performed due to it's mechanical and time consuming nature and the sheer volume of data involved. The goal of image quality assessment (IQA) is to automate this process by developing objective quality estimators that can predict subjective scores. In other words, the *perceived* quality of images is measured objectively. Based on the availability of original distortion free images, image quality assessment algorithms are classified into three categories. Full-Reference (FR) metrics require the original image for predicting the quality of distorted image [4, 10, 11, 12, 13, 14, 15, 16]. No-Reference (NR) metrics estimate the quality of a distorted image without requiring access to the corresponding original image [18, 19, 20]. Reduced-Reference metrics require a few feature sets extracted from the original image for quality prediction of the distorted image. In the proposed work, we focus on extending a FR model that we proposed in [1] which was based on a data driven approach.

Data driven approaches are not uncommon in IQA literature. The authors in [17] propose MLIQM, a metric that benefits from the already present IQA theory to construct features, and apply SVM classification to understand the quality class. Then a SVM regression is used to estimate the quality of a distorted image within that quality class. The authors in [18] apply a pre-training step in which they distort high quality images and feed them into their deep network to train a model that predicts the subjective score. The authors in [19] propose an image quality assessment approach based on learning a set of filters through Support Vector Regression. The weights of SVR are learnt through a Stochastic gradient descent algorithm and their responses are used to estimate quality. In [20], the authors propose an unsupervised learning approach to obtain quality-aware filters using distorted images. These filters are used to extract features that are then regressed using a random forest to obtain quality estimates. However, the common thread in all these algorithms, is the requirement of distorted images and subjective scores during training.

In this paper, we explore the combination of unsupervised learning and hand-crafting to extend learning networks to assess quality of images. In [1] we had proposed UNIQUE, a shallow learning architecture to estimate quality. It had one hidden layer which was trained using a sparsity criterion where the weights and bias were considered a domain transformation on non-overlapping patches of images. This technique outperforms majority of the existing methods in LIVE [8] and TID 2013 [9] databases. It is an unsupervised architecture since it does not require any target labels during the neural network training. Also, there is no need for either subjective scores or distorted images during training. Keeping all these advantages intact, we extend UNIQUE and improve it's performance by analyzing the weights which we learnt, utilizing existing IQA literature that stresses the importance of sharpness in measuring quality [7]. We also learn multiple self-contained and reversible representations of undistorted data and use these representations to estimate quality of images. We propose MS-UNIQUE which is a full reference image quality assessment algorithm based on an unsupervised learning approach through distortion-free images.

Methodology

We propose learning a set of weights and bias from a linear decoder. Before using the learning framework, we preprocess the data to make it more descriptive. The learnt weights from a linear decoder are considered as a filter set which are used to estimate the quality of an image. And if linear decoder models with different number of neurons in the hidden layer are trained, we obtain a number of filter sets all learning to model the same input using multiple representations. The filters are also made structure aware by differentiating the ones that capture edges from the ones that capture color.

Color Space Selection

We use luminance (Y channel) as part of our input data. The human visual system is more sensitive towards changes in intensity domain rather than chroma [2]. The authors in [4] claim that structural information can be gleaned from the normalized luma domain. In addition to the Y channel, we use the green channel from RGB color space. Green channel is selected since it contains a large part of the information from R and B color channels. This is verified by measuring the cross correlation between channels of RGB representations - the cross correlation r_{RG} between R and G color channels is 0.98 and r_{GB}, between G and B color channels is 0.94 [3]. We augment the Y and the G channels with the Cr channel after a transformation into YCbCr colorspace. This is done to include chroma information as part of our data. The specific plane Cr is chosen over Cb based on experimental results. The three planes are combined to obtain a descriptive YGCr image.

Data Matrix Preparation

From the ImageNet 2013 test database, $1,000$ images are randomly selected during training. We do not use any annotated metadata associated with the images. Each image is first transformed into YGCr colorspace. From each image, we extract 100 patches of size 8x8x3 randomly. Each patch is then reshaped into a 192x1 column vector. The patch vectors from all images are stacked together to get a 192x100000 input patch matrix. The data matrix is then passed through a Zero Component Analysis (ZCA) Whitening algorithm. Whitening is performed to decorrelate adjacent pixels in raw data so as to lessen redundancy. The authors in [6] show that the HVS performs whitening. Essentially, this converts the input data with a zero mean covariance matrix into whitened data with an identity covariance matrix. The adjacent features in the input matrix are decorrelated and the variance of each is one. ZCA also satisfies the property that the whitening matrix is orthogonal. Note that whitening is not performed on the $100,000$ patches but on the 192 input features in each patch feature vector. Hence, individual pixels inside a patch are decorrelated from other pixels in the same patch. This happens over all $100,000$ patches hence lowering the redundancy fed into the learning architecture from each feature vector [5]. We summarize the data matrix preparation in Figure 1.

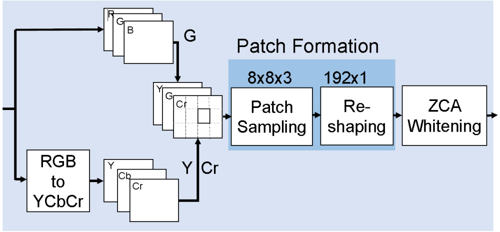


Figure 1: Data matrix preparation.

Linear Decoder

A linear decoder is an unsupervised neural network framework used to represent data in different dimensions. In this work, we use a framework with only one hidden layer. It can be used to sparsify data or learn a compact representation by changing the number of neurons in the hidden layer. The framework functions by transforming the input into hidden layer activations or responses and then reconstructing back the input using these responses. Transformation occurs through a set of weights and bias that are randomly initialized and then adjusted iteratively based on the reconstruction error using backpropagation. The hidden layer responses are obtained as

$$s = sigmoid(W_1^T P + b_1), \tag{1}$$

where s is the response, W_1 and b_1 are the forward weights and bias. Each column in W_1 is a 192x1 vector that filters each patch from the data. Sigmoid is the non-linear layer used in our framework. These hidden layer responses are filtered using another set of backward weights W_2 and bias b_2 to obtain back a reconstructed version of the input $\tilde{P}$ as

$$\tilde{P} = W_2^T s + b_2, \tag{2}$$

Note that there is no sigmoid layer after reconstruction. The objective function for backpropagation $J(W_1, W_2, b_1, b_2)$ is given by

$$J(W,b) = \|(W_2^T s + b_2) - P\|_2^2 + \beta \sum_{j=1}^{N} \mathrm{KL}(\rho || \hat{\rho}_j) + \lambda \|W\|_2^2, \tag{3}$$

where the first term is the reconstructed $L2$ norm error, the second term is the sparsity penalty term and the third is the weight decay or regularization term. N is the total number of patches, which amounts to $100,000$. Sparsity penalty is included to constrain the average activation of neurons to be close to zero and this penalty is obtained using KL-Divergence over all training patches. ρ is the desired average activation which is set to 0.035. The sparse penalty term goes to 0 when the actual average activation $\hat{\rho}$ comes close to ρ. It is weighted by β which is set to 5. The weight decay term λ, which is set to $3e^{-3}$ acts as a regularization term by decreasing the magnitude of weights thereby preventing overfitting. We show the architecture of a linear decoder in Figure 2 in which h corresponds to the number of neurons in the hidden layers. We change the number of neurons h to obtain models that can either sparsely or compactly represent the input data. We visualize the weight sets corresponding to different values of h in Figure 3.

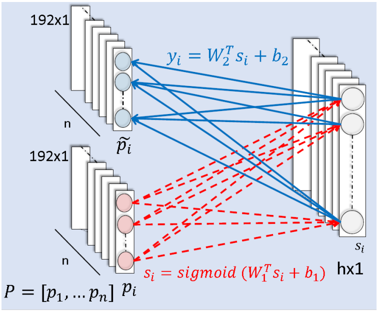
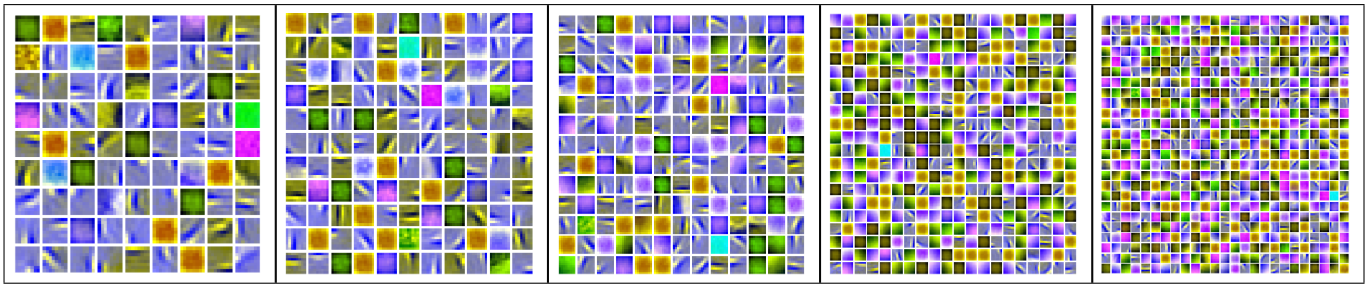
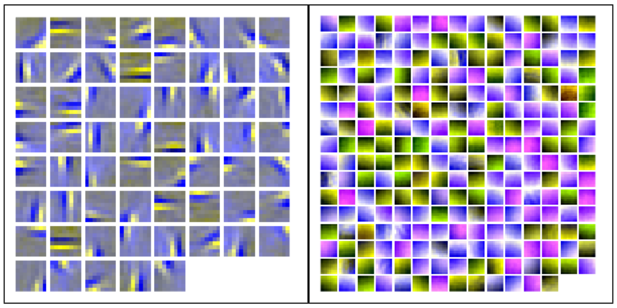


Figure 2: Linear decoder architecture



(a) 81 neurons (b) 121 neurons (c) 169 neurons (d) 400 neurons (e) 625 neurons

Figure 3: Weight Visualizations. In each set, each square can be used to infer input patches that maximally activate it. Each individual square in all sets is of size 8x8x3 and is scaled here for visualization purposes.



(a) Edge Filters (b) Color Filters

Figure 4: Result Visualization of differentiating a 625 filter model into edge and color aware filters.

Multi-model training

The data matrix is fed into a linear decoder model with $h = 81$ and trained for 400 epochs. The trained forward weights and bias are stored. This step is repeated to obtain weights and bias for $h = 121, 169, 400, 625$ separately. The sparsity parameter during training ensures that none of the filters from any model get activated abnormally over the others. This multi-model approach ensures that we represent an image patch both sparsely and compactly and learn multiple filter sets that combine non-linearly to reconstruct it. Also, a sparse filter set learns more localized features while a compact set learns global features.

Sharpness aware filters

Sharpness is an important determining factor in the perceptual quality assessment of images [7]. The HVS is adept at detecting blur and evaluating quality based on sharpness. However, our learning framework does not use any handcrafted features like incorporating edges. Hence we add this feature to the already constructed filter set. We make our filters sharpness aware by analyzing their descriptiveness and then weighing their responses accordingly. We give higher importance to filters that capture edges rather than color. Distinguishing filters based on edge characteristics is performed using the bias corrected implementation of kurtosis. Kurtosis is defined as,

$$k = \frac{E(x-\mu)^4}{\sigma^4} \tag{4}$$

Hence, further away a data point is from the mean of the distribution, larger is it's influence on kurtosis. We theorize that filters that capture edge components consist of more data points that are away from the mean of the overall data making them outliers. The presence of these outliers gives a higher kurtosis score to edge filters. The kurtosis of each vectorized, zero centered, and normalized filter is measured against a threshold to capture it's edge characteristics. Any filter with a kurtosis greater than 5 is labeled as an edge filter while filters with kurtosis less than 2 are labeled as color filters. The results of thresholding on a 625 filter model set is shown in Figure 4.

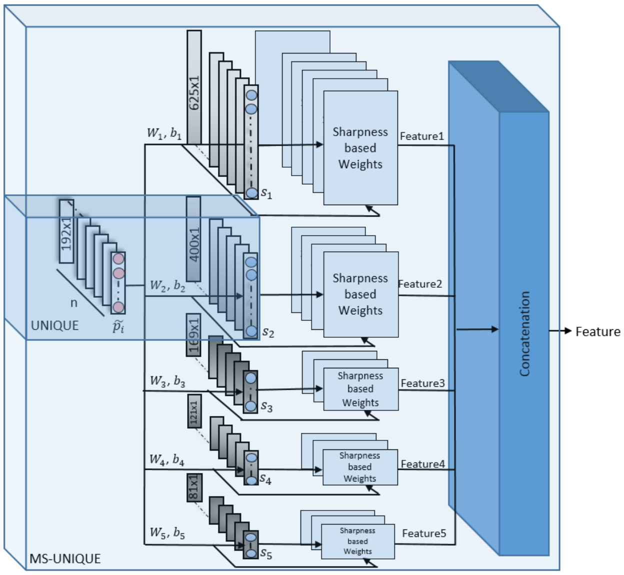


Figure 5: Feature generation

Image Quality Assessment

We preprocess images as described previously and utilize the formulation in Eq.1 to obtain filter responses. These responses are weighted based on the sharpness characteristics of corresponding filters. The edge filter responses are given a higher weightage of 2 while the color responses are lowered by a weight of 0.5. This is performed for all models to obtain one feature vector per image. The feature generation process is summarized in Figure 5. The responses in feature vector that are significantly less than the average activation value set during training are assigned a zero to mimic the suppression mechanisms in the HVS. We generate

Table 1: Performance of image quality estimators.

Methods	**PSNR**	**PSNR HA**	**PSNR HMA**	**SSIM**	**MS SSIM**	**CW SSIM**	**IW SSIM**	**SR SIM**	**FSIMc**	**PerSIM**	**UNIQUE**	**MS-UNIQUE**
		[10]	[10]	[4]	[11]	[12]	[13]	[14]	[15]	[16]	[1]	
	Outlier Ratio											
TID13	0.725	0.615	0.670	0.732	0.697	0.855	0.700	0.632	0.727	0.655	0.640	**0.611**
	Root Mean Square Error											
LIVE	8.61	6.93	**6.58**	7.52	7.43	11.2	7.11	7.54	7.20	6.80	6.76	6.61
TID13	0.87	0.65	0.69	0.76	0.68	1.20	0.68	0.61	0.68	0.64	0.60	**0.57**
	Pearson Correlation Coefficient											
LIVE	0.928	0.953	**0.958**	0.945	0.946	0.872	0.951	0.945	0.950	0.955	0.956	**0.958**
TID13	0.705	0.850	0.827	0.789	0.832	0.227	0.831	0.866	0.832	0.854	0.870	**0.884**
	Spearman Correlation Coefficient											
LIVE	0.909	0.937	0.944	0.949	0.951	0.902	**0.960**	0.955	0.959	0.950	0.952	0.949
TID13	0.700	0.847	0.817	0.741	0.785	0.562	0.777	0.807	0.851	0.853	0.860	**0.870**

Table 2: Distributional difference between subjective scores and objective quality estimates.

Metric	**Difference-LIVE**					**Difference-TID13**				
	EMD	**KL**	**JS**	**HI**	**L2**	**EMD**	**KL**	**JS**	**HI**	**L2**
PSNR-HMA	0.226	0.205	0.053	0.226	0.066	0.360	0.927	0.117	0.360	0.124
IW-SSIM	0.297	0.325	0.072	0.297	0.076	0.500	1.678	0.196	0.500	0.180
UNIQUE	0.236	0.258	0.055	0.236	0.069	0.386	0.855	0.120	0.386	0.109
MS-UNIQUE	**0.209**	**0.176**	**0.038**	**0.209**	**0.057**	**0.357**	**0.734**	**0.108**	**0.357**	**0.103**

feature vectors for both reference and distorted images. The feature vectors corresponding to the original and distorted images are compared using 10^{th} power of Spearman correlation coefficient to fully utilize quality estimation range.

The proposed method is an extension of the quality estimator UNIQUE [1] as shown in Figure 5. It builds on UNIQUE by weighing filter responses. We also propose using multiple decoders with different number of neurons in the hidden layer to abstract local and global characteristics in image patches.

Validation

Database

The proposed quality estimator is validated on the LIVE image quality [8] and TID 2013 [9] databases. The databases have more than 3500 distorted images between them. These images can be classified into 7 categories based on their distortion types - compression artifacts, image noise, communication errors, blur artifacts, color degradations, global, and local distortions. The compression artifacts category consists of the JPEG and the JPEG2000 compressions, and lossy compressions of noisy images. The noise category includes Gaussian noise and additive noise added in color components, spatially correlated noise, masked noise, high frequency noise, impulse noise, quantization noise, image denoising, multiplicative Gaussian noise, and comfort noise. The communication errors category includes the JPEG and the JPEG2000 transmission errors of noisy images. The blur artifacts category consists of Gaussian blur, and sparse sampling and reconstruction. The color degradations category contain changes in color saturation, image color quantization with dither, and chromatic aberrations. The global category includes intensity shifts, and contrast changes while the local category contains non-eccentricity pattern noise, and local blockwise distortions of different intensities.

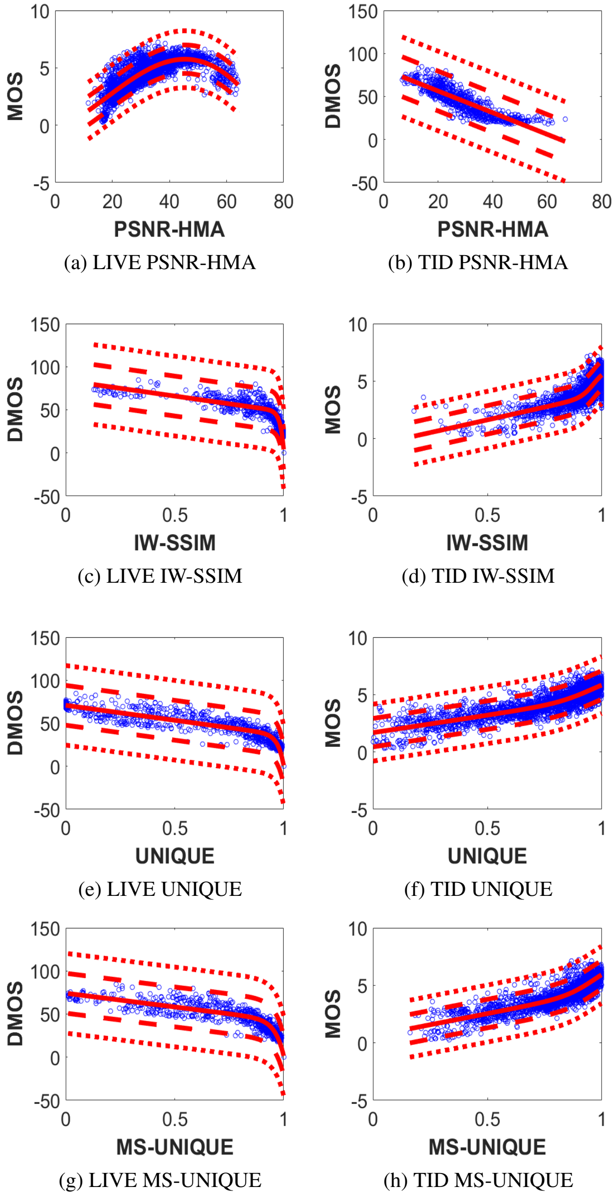
Performance Metrics

Validation of MS-UNIQUE and compared algorithms are carried out in terms of root mean square error, outlier ratio, Pearson and Spearman correlation coefficients. In the outlier ratio calculations, we use those data points that lie two standard deviations away from the average subjective scores. Also, outlier ratio is only reported for TID 2013 database since the standard deviations of subjective scores are not reported in LIVE database. The regression formulation from [8] is used to calculate regress estimates of all methods before comparing. We report the difference between the normalized histograms of subjective scores and the regressed quality estimates through common histogram difference metrics including Earth Movers Distance (EMD), Kullback-Leibler (KL) divergence, Jensen-Shannon (JS) divergence, histogram intersection (HI), and L2 norm.

Results

The proposed quality estimator is compared against eleven other commonly used or state of the art full reference quality assessment methods based on fidelity, perceptually-extended fidelity, structural similarity, feature similarity, and perceptual similarity. The performances of all these metrics are summarized in Table 1 with the highest performing metric in each category displayed in bold. PSNR-HMA, IW-SSIM, UNIQUE, and MS-UNIQUE are among the top performing metrics. MS-UNIQUE outperforms all these estimators in TID13 database among all performance metrics. In the LIVE database it consistently performs well in all but two of the metrics. IW-SSIM performs better in terms of SROCC in this database. However, MS-UNIQUE outperforms IW-SSIM among all the other categories. Both MS-UNIQUE and PSNR-HMA provide similar results in terms of PCC. MS-UNIQUE's results for RMSE are slightly lesser than PSNR-HMA. MS-UNIQUE builds on UNIQUE among all categories except in SROCC in LIVE database.

To better analyze the distribution of subjective scores against



(a) LIVE PSNR-HMA (b) TID PSNR-HMA

(c) LIVE IW-SSIM (d) TID IW-SSIM

(e) LIVE UNIQUE (f) TID UNIQUE

(g) LIVE MS-UNIQUE (h) TID MS-UNIQUE

Figure 6: Scatter plots of top performing quality estimators

the estimated scores, scatter plots of the best performing metrics are shown in Figure 6. The X-axis corresponds to the estimated scores while the Y-axis is the ground truth subjective mean opinion scores (MOS) or differential Mean Opinion Scores (DMOS). For an ideal quality estimator, the scatter plot data should follow a linear curve with low deviation. This is not observed in PSNR-HMA which shows a parabolic curve in LIVE database. There is a much sharper drop off in IW-SSIM with most of the points concentrated on the far end of the curve in LIVE database. UNIQUE and MS-UNIQUE have a far more linear curve with scores extending throughout the range. To numerically differentiate between MS-UNIQUE and other metrics in terms of regressed quality estimates, we present the difference between the normalized histograms of ground truths and regressed results, in Table 2. The best results are highlighted in bold and MS-UNIQUE consistently performs well in both the databases among all compared metrics. Overall, MS-UNIQUE is the best performing metric in 15 out of 17 compared metrics over both databases.

Conclusion

We proposed an extension to the quality estimator UNIQUE, by analyzing the learning network used and handcrafting a weighing scheme that captures sharpness. This is done in the preprocessing and postprocessing blocks by enhancing information acquired from the data, analyzing the edge characteristics of learnt filters so that their responses are weighed based on quality assessment theory. Multiple models of linear decoders, where the number of hidden layer neurons represent the local or global characteristics captured, are used to obtain different levels of abstraction. The performance of MS-UNIQUE shows that performance of metrics that use a data driven approach can be enhanced by handcrafting features.

References

[1] D. Temel, M. Prabhushankar, and G. AlRegib, ”UNIQUE: Unsupervised Image Quality Estimation,” in IEEE Signal Processing Letters , vol.23, no.10, pp.1414-1418

[2] Van den Branden Lambrecht, Christian J. ”Vision models and applications to image and video processing” Springer Science and Business Media, 2013.

[3] M. Tkalcic and J. F. Tasic, Colour spaces: perceptual, historical and applicational background, in *EUROCON 2003. Computer as a Tool. The IEEE Region 8*, Sept 2003, vol. 1, pp. 304308 vol.1.

[4] 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, no. 4, pp. 600612, April 2004.

[5] A.Krizhevsky, and Geoffrey Hinton. ”Learning multiple layers of features from tiny images.” (2009), Appendix A, The ZCA Whitening Algorithm.

[6] Y. Dan, J. J. Atick, and R. C. Reid, ”Efficient coding of natural scenes in the lateral ganiculate nucleus: Experimental test of a computational theory”, The Journal of Neuroscience, vol. 16(10), pp. 33513362, 1996.

[7] Rania Hassen, Zhou Wang, and Magdy Salama. ”No-reference image sharpness assessment based on local phase coherence measurement.” 2010 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2010.

[8] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, IEEE Transactions on Image Processing, vol. 15, no. 11, pp. 34403451, Nov 2006.

[9] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. Jay Kuo, Image database TID2013: Peculiarities, results and perspectives , Signal Processing: Image Communication, vol. 30, pp. 57 77, 2015.

[10] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, and M. Carli, Modified image visual quality metrics for contrast change and mean shift accounting, the proceedings of CADSM, 2011.

[11] Z. Wang, E. P. Simoncelli, and A. C. Bovik, Multiscale structural similarity for image quality assessment, in Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on, Nov 2003, vol. 2, pp. 13981402 Vol.2.

[12] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey, Complex wavelet structural similarity: A new image similarity index, IEEE Transactions on Image Processing, vol. 18, no. 11, pp. 23852401, Nov 2009.

[13] Z. Wang and Q. Li, IW-SSIM: Information content weighted structural similarity index for image quality assessment, IEEE Transactions on Image Processing, vol. 20, no. 5, pp. 11851198, May 2011.

[14] L. Zhang and H. Li, SR-SIM: A Fast and high performance IQA index based on spectral residual, in Image Processing (ICIP), 2012 19th IEEE International Conference on, Sept 2012, pp. 14731476.

[15] 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. 23782386, Aug 2011.

[16] D. Temel and G. AlRegib, PerSIM: Multi-resolution image quality assessment in the perceptually uniform color domain, in Image Processing (ICIP), 2015 IEEE International Conference on, Sept 2015, pp. 16821686.

[17] C. Charrier,O. Lzoray, and G. Lebrun, Machine learning to design full-reference image quality assessment algorithm,Signal Processing:Image Communication, vol. 27, no. 3, pp. 209 219, 2012.

[18] H. Tang, N. Joshi, and A. Kapoor, Blind image quality assessment using semi-supervised rectifier networks, in Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Washington DC, USA, 2014, CVPR 14, pp. 28772884, IEEE Computer Society.

[19] P. Ye, J. Kumar, L. Kang, and D. Doermann, Real-time no-referenceimage quality assessment based on filter learning, in Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 2013, CVPR 13, pp. 987994, IEEE Computer Society.

[20] Z. Gu, L. Zhang, X. Liu, H. Li, and J. Lu, Learning quality-aware filters for no-reference image quality assessment, in Multimedia and Expo (ICME), 2014 IEEE International Conference on, July 2014, pp.16.

Author Biography

Mohit Prabhushankar received the M.S. degree in Electrical and Computer Engineering with a minor in Computer Science from Georgia Institute of Technology, Atlanta, in 2015. Since then, he has been pursuing Ph.D. degree in the Center for Signal and Information Processing (CSIP), Georgia Institute of Technology, USA, as a Research Assistant. His research interests include image quality assessment, image denoising and enhancement feature design through data driven approaches.

Dogancan Temel received an M.S. degree with a minor in Management in 2013, and a PhD degree with a minor in Computer Science in 2016 from the school of Electrical and Computer Engineering in Georgia Institute of Technology, Atlanta. While his studies at Georgia Tech, Dr. Temel worked in the Multimedia and Sensors Lab at the Center for Signal and Information Processing as a Graduate Research Assistant and in Texas Instruments as a Systems Engineering intern. Dr. Temel worked on various projects including perceived image quality assessment, deep learning-based image processing and computer vision, high color range imaging, vital sign monitoring, computational aesthetics, seis- mic interpretation, 3D reconstruction, streaming, and quality assessment.

Ghassan AlRegib is currently Associate Professor at the School of Electrical and Computer Engineering at the Georgia Institute of Technology in Atlanta, GA, USA. His research group is working on projects related to image and video processing and communications, immersive communications, collaborative systems, quality of images and videos, and 3D video processing. Prof. AlRegib is a Senior Member of IEEE. Prof. AlRegib served as the chair of the Special Sessions Program at the IEEE International Conference on Image Processing (ICIP) He was the Track Chair in the IEEE International Conference on Multimedia and Expo (ICME) in 2011 and the co-chair of the IEEE MMTC Interest Group on 3D Rendering, Processing, and Communications, 2010-present. Prof. AlRegib is a member of the Editorial Board of the Wireless Networks Journal (WiNET), 2009-present. Prof. AlRegib co-founded the ICST International Conference on Immersive Communications (IMMERSCOM) and served as the Chair of the first event in 2007. Prof. AlRegib is the founding Editor-in-Chief (EiC) of the ICST Transactions on Immersive Communications to be inaugurated in late 2012. He is also the Chair of the Speech and Video Processing Track at Asilomar 2012.