Independent Feature Detector for Image Quality Assessment

Kshama H G Kanchugar[1], Ullas Pv[2], Rijo John Jacob[3], Sushmita Nair[4], Rekha K S[5]

[1]kshamagk@gmail.com
[2]ullaspv94@gmail.com
[3]rijohn7@gmail.com
[4]sushmita.mys@gmail.com
[5]rekhakowdle@nie.ac.in
Department of CSE,
NIE, Mysuru

---***---

Abstract - *Image quality assessment plays a key role in evaluating and optimizing image processing systems. Such Image Quality Metrics (IQM) are largely based on visual perception characteristics as the human visual system (HVS) is the ultimate receiver of images. Independent Component Analysis (ICA) models are similar to the primary visual cortex which is an important part of the HVS. Inspired by this, we make use of Independent Feature Similarity (IFS) index for the full reference IQA. The proposed technique involves a feature detector which is trained on the natural images by ICA. Independent features are acquired from this feature detector. The computation of IFS involves two components: feature and luminance. Based on these two components the IFS index is computed. The performance score of IFS is also obtained.*

Key Words: **Image quality assessment, Visual perception, Independent component analysis, Visual cortex, Independent Feature Detector**

1. INTRODUCTION

The human visual system is believed to efficiently encode natural visual information. This is done by decreasing the redundancy of the visual information at each stage of processing. Natural visual input contains features such as edges and homogenous color patches, which make the patterns of light falling on the retina highly redundant and which give rise to statistical dependencies between neighboring regions of the visual image. To reduce redundancy, the visual system might use these features as a basis for representing visual input [2,3].

Sophisticated Image Quality Assessment (IQA) models which are based on physiological response properties of the human visual system (HVS) have been developed to achieve perceptual consistency in the quality prediction. Independent Component Analysis (ICA) when applied to achromatic natural images gives independent components that are very much similar to the achromatic spatial receptive fields of the simple cells in the primary visual cortex, even though the produced independent components are only approximately independent. This is the reason why ICA is widely used method for redundancy – reducing encoding of data like natural visual scenes.

In our paper we have proposed an Image Quality Metric (IQM) called Independent Feature Similarity for full reference or reduced reference IQA. The image quality is evaluated based on the similarity between the independent features which are extracted by a feature detector. This feature detector is trained on natural images by ICA. The entire computation is divided into two main stages: training and similarity computation. Finally, the IFS quality index is computed. The performance score of the IFS is also computed.

2. LITERATURE SURVEY

In [6], Anush Krishna Moorthy and Alan Conrad Bovik have discussed, under the hypothesis that certain regions in an image may be visually more important than others, methods used to spatially pool the quality scores from the SSIM maps are an appealing possibility for improving SSIM scores. In, the effect of using different pooling strategies was evaluated, including local quality-based pooling. It was concluded that the best possible gains could be achieved by using an information-theoretic approach deploying "information content-weighted pooling." They have further investigated quality based pooling and also considered pooling based on predicted human gaze behavior.

In [7], J H van Hateren and A van der Schaaf have presented that performing ICA on a large set of calibrated images, and comparing a series of properties of the resulting components with those of receptive fields measured in simple cells of HVS. It was found that there is a good similarity between the distributions for spatial frequency bandwidth, orientation tuning bandwidth, aspect ratio and receptive field length.

3. EXISTING SYSTEM

Subjective IQA is the most accurate and reliable technique as humans are the ultimate users of many of the multimedia applications. However, subjective evaluations are expensive and time consuming, which makes them impractical in real-world applications Therefore, it is necessary to design mathematical models that are capable of predicting the quality evaluation of an average human observer.

The simplest and most widely used IQA scheme is the mean squared error (MSE) or peak signal-to-noise ratio (PSNR). However, it is well known that MSE/PSNR does not always agree with the subjective evaluation results.

In [5], ICA was used to get a feature detector by training on samples of each reference image. Because the samples come from each corresponding reference image, the training process should be done in every evaluation, which takes a lot of time.

Some systems like structural similarity, visual information fidelity and gradient similarity can be applied only to grey scale images.

4. PROPOSED SYSTEM

The IFS works on color images. As IFS involves removal of perceptually unimportant information, it can also be seen as a reduced – reference IQM. The feature detector is trained on color natural images by ICA. The perceived distortion of the color information within a given image can be predicted by IFS because of the ability of the ICA to find statistical regularities of chromatic perception from the training samples.

The entire computation process is divided into training and similarity computation. The similarity computation involves two components: feature and luminance. The feature component is used to evaluate the texture and the structure distortions whereas the luminance component evaluates brightness distortion. The difference between higher-order and lower-order statistical information are also indicated by these two components. The IFS quality score is computed by the combination of the feature and the luminance values.

5. METHODOLOGY

5.1 Feature Detector Training

First part is the Feature Detector. Feature Detector is used to extract the independent features, used by IFS to assess the quality. It is a matrix **W**, got from the training stage, for extracting independent features from reference and distorted image patches. Later IFS QI (Quality Index) is computed by making use of Extracted Features in Similarity Computation stage.

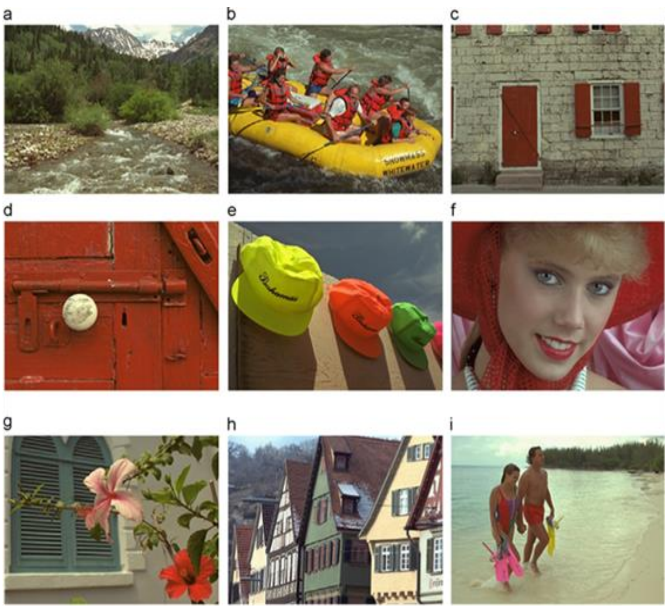

Fig-1: Images for training of feature detector.

Fig-1 shows images selected randomly from TID2008 dataset [8,9]. Initially, random image patches, 9000 in number, each of size 8x8 are taken form the natural images shown in Fig-1. Image patches are treated like random vectors **X** which are whitened into **Z** by PCA. Later **W** will be trained on data, **Z** by **FastICA** [10,11].

This process is considered a model of visual processing. Whitening combined with dimension reduction simulates retina and lateral geniculate nucleus [12-14], and ICA simulates cortical neurons [15-17].

5.2 Preprocessing

Each patch is converted to column vector by scanning numerical values in the patch. As here the patches have three channels, the length of the vector will be 8x8x3=192. It is then centered by subtracting mean pixel value in each patch as it is one independent component and has negative peak or kurtosis. Finally all vectors form a matrix **X**.

5.3 Whitening and Dimension Reduction by PCA

Earlier Sensory Neurons removed statistical redundancy in sensory input [14,18,19]. Therefore PCA is made use of to reduce dimension of vectors **X** without losing too much detail as neighboring elements of central pixel are closely correlated [20]. PCA captures this correlation allowing almost same data but with much smaller vector while another task for PCA is also Whitening.

PCA can be done by eigenvalue decomposition of data covariance matrix. Let $\mathbf{D}=\mathrm{diag}(d_1, \ldots, d_M)$ and $E=(\mathbf{e}_1, \ldots, \mathbf{e}_M)$ respectively show M largest eigenvalues and respective eigenvectors for matrix **X**. Here only 8 principal components of samples are retained for training, so dimension of weighted vector is reduced from 192 to M=8 for each of them. Whitening Matrix **V** is given by

$$\mathbf{V} = \mathbf{D}^{-1/2} \times \mathbf{E}^{T} \quad (1)$$

Finally **Z** is given by

$$\mathbf{Z} = \mathbf{V} \times \mathbf{X} \quad (2)$$

5.4 Training by FastICA

Finally feature detector learns from **Z**, which is a whitened data, by an ICA algorithm, Here, FastICA, a fixed-point algorithm [20], is used. The basic form of iteration involved in FastICA is given by

$$\mathbf{h}_{(p)} = E\{\mathbf{Z}_g((\mathbf{h}_{(p-1)})^T \mathbf{Z})\} - E\{g'((\mathbf{h}_{(p-1)})^T \mathbf{Z})\} (\mathbf{h}_{(p-1)}) \quad (3)$$

h denotes the vector of feature detector in the whitened space, p denotes iteration step, $E(.)$ represents mathematical expectation.

g(.) is the derivative of function $G(.)$.

$$G(u) = \frac{1}{\eta}\log\cosh(\eta u) \quad (4)$$

$$g(u) = \tanh(\eta u) \quad (5)$$

where $1 \le \eta \le 2$, while in this paper, $\eta=1$.

A random orthogonal matrix to start, each iteration consists of updating each row $\mathbf{h}^T$ of **H** by Eq. **(3)**, followed by orthogonalization of matrix through

$$\mathbf{H} \leftarrow (\mathbf{H}\mathbf{H}^T)^{-1/2}\mathbf{H} \quad (6)$$

This procedure will repeat until it reaches convergence. The feature detector should be transformed from whitened space to Original Space, after learning process, by

$$\mathbf{W} = \mathbf{H} \times \mathbf{V} \quad (7)$$

6. Computation of independent component similarity

IFS index consists two components Feature and Luminance. Reference and Distorted images are first divided into 8 x 8 non-overlapping patches that is scanned by a sliding window. Since images are divided the same way, each reference and distorted counterpart form a patch pair which is later considered pair of vectors.

6.1. Feature Component

The mean values are removed and the pixel values of the image patches form a column vector with zero mean. They form reference and distorted image matrices, $\mathbf{X}^{ref}$ and $\mathbf{X}^{dis}$ and fidelity computation is done by following 3 steps:

6.1.1 Selection of patch pairs:

Supra-threshold distortions can be more expressive and is easily caught by human eye, hence HVS is sensitive to poor quality areas than good ones.

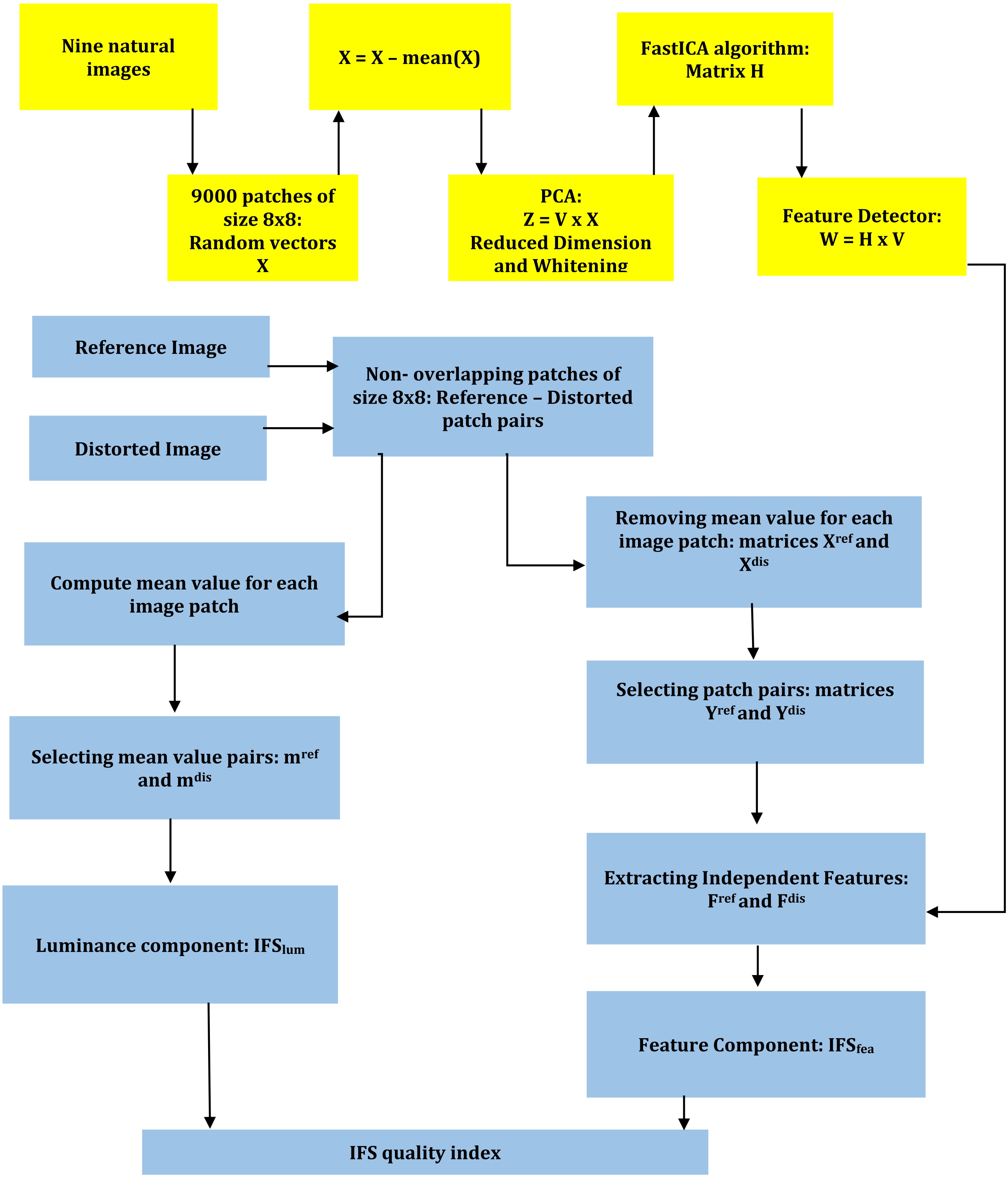

Fig-2: Block diagram for IQA using Independent Feature Detector

Many researches proposed weighting strategies putting more weight on low quality regions [21,22] which increases complexity. So here it is made simple by the proposal of using only reference-distorted patch pairs with large differences, here difference between pair of reference-distorted patch vectors is computed by Mean Absolute Error, Let $\mathbf{x}^{ref}$ and $\mathbf{x}^{dis}$ denote reference and distorted patch vectors respectively. Then, $\mathbf{x}^{ref}$ and x^{dis} is given by.

$$\mathrm{MAE}(\mathbf{x}^{ref}, \mathbf{x}^{dis}) = \frac{1}{N}\sum_i^N |x_i^{ref} - x_i^{dis}| \quad (8)$$

MAE values between reference and distorted patches form a vector, $\mathbf{b} = (b_i \mid i=1,\dots,L)$, where $b_i = MAE(\mathbf{x}^{ref},\mathbf{x}^{dis})$, and L is the number of patches.

A threshold or limit is selected to select set of patches that can be considered, called *TH*, *TH* is chosen on median value **b**. Here, it is shown that value of *TH* depends on strength of distortions, usually a larger value if distortion level is high. And smaller value when distortion level is low. We choose 512 x 512 as standard image size to analyze whether an image is of higher or lower resolution, as *TH* is smaller in high resolution images and vice versa.

The threshold is computed by

$$\mathbf{TH} = \begin{cases} \mathbf{median(b)}, & \mathbf{median(b)} < \mathbf{T_x} \\ \frac{1}{5}(\mathbf{max(b)} + 4\,\mathbf{median(b)}), & \mathbf{otherwise} \end{cases} \quad (9)$$

$$\mathbf{T_x} = \mathbf{T}\frac{\mathbf{S_h} \times \mathbf{S_w}}{512 \times 512} \quad (10)$$

where S_h and S_w depict height and width of input image and T=7(adjust for T_x).

If MAE value of patch is greater than *TH*, then that patch is selected, with vectors $\mathbf{y}^{ref}$ and $\mathbf{y}^{dis}$ denoting retained reference for reference and distorted image vectors, and thus all vectors form $\mathbf{Y}^{ref}$ and $\mathbf{Y}^{dis}$ by

$$(\mathbf{Y}^{ref}, \mathbf{Y}^{dis}) = \{(\mathbf{y}^{ref}, \mathbf{y}^{dis}) \mid \mathrm{MAE}(\mathbf{y}^{ref}, \mathbf{y}^{dis}) \geq TH\} \quad (11)$$

With this, computation is based on part of reference data, hence IFS is called Reduced Reference IQM.

6.1.2 Feature Extraction:

Independent features are extracted by matrix multiplication of

$$\mathbf{F}^{ref} = \mathbf{W} \times \mathbf{Y}^{ref}\,;\ \mathbf{F}^{dis} = \mathbf{W} \times \mathbf{Y}^{dis} \quad (12)$$

where **F** denotes two sets of independent features extracted by **W**.

6.1.3 Similarity measurement:

Feature component IFS and IFS_{fea}, is computed by:

$$\mathbf{IFS_{fea}} = \frac{1}{M \cdot N}\sum_j^M \sum_i^N \frac{2F_{ij}^{ref}\,F_{ij}^{dis} + C}{(F_{ij}^{ref})^2 + (F_{ij}^{dis})^2 + C} \quad (13)$$

M=8, length of feature vector. N shows number of retained vectors obtained by feature detector in training stage, $\mathbf{F}^{ref}_{ij}$ and $\mathbf{F}^{dis}_{ij}$ denote elements in i-th column and j-th row in $\mathbf{F}^{ref}$ and $\mathbf{F}^{ref}$, C stabilizes result.

6.2 Luminance Component

Computed based on mean values of image patches. But not all mean values are required as lower differences won't affect perceptions of regions with higher difference, thus considering large differences, done by selection operation, luminance component can be computed. Experiments show that only 20% of mean values are enough for assessing brightness distortions. Difference is measured by absolute error.

6.2.1 Selection of mean value pairs

This is done by sorting absolute errors between mean values of reference and distorted patches. So after absolute error computation, all mean value pairs are sorted in low-first order and selecting last 20% pairs, two vectors $\mathbf{m}^{ref}$ and $\mathbf{m}^{dis}$, for reference and distorted images.

6.2.2 Correlation Measurement

After selection, luminance component of IFS, $IFS_{lum,}$ is calculated by

$$IFS_{lum} = \frac{\sum_i\left(\left(m_i^{ref} - \mu(m^{ref})\right)\cdot\left(m_i^{dis} - \mu(m^{dis})\right)\right) + C_m}{\sqrt{\sum_i(m_i^{ref} - \mu(m^{ref}))^2 \cdot \sum_i(m_i^{dis} - \mu(m^{dis}))^2} + C_m} \quad (14)$$

where $C_m > 0$, to avoid dividing by zero, m^{ref}_i *and* m^{dis}_i denote $\mathbf{m}^{ref}$ and $\mathbf{m}^{dis}$ and $\mu(\mathbf{m}^{ref})$ $\mu(\mathbf{m}^{dis})$ are means of $\mathbf{m}^{ref}$ and $\mathbf{m}^{dis}$.

6.3 Independent feature similarity

Finally, IFS Index is computed by combining IFS_{lum} and IFS_{fea} into a quality score.

$$\boldsymbol{IFS} = \left(\left(\boldsymbol{IFS_{fea}}\right)^{\alpha} \cdot \left(\boldsymbol{IFS_{lum}}\right)^{\gamma}\right)^{\frac{1}{(\alpha+\gamma)}} \quad \mathbf{(15)}$$

where α > 0 and γ > 0 are parameters to adjust relative importance. For simplicity, α= γ=1 here.

7. RESULTS

We have employed lena image to show the experimental results. Four types of distortions: Additive White Gaussian Noise (AWGN), JPEG compression (JPEG), JPEG 2000 compression (JP2K) and Gaussian Blur (GB) are applied on the image and the IFS quality score is computed.

Eight metrics are used to measure the performance of IFS. The performance metrics being used are: Maximum Difference (MD), Mean Square Error (MSE), Root Mean Square Error (RMSE), Normalized Cross- Correlation (NK), Peak Signal to Noise Ratio (PSNR), Mean Absolute Error (MAE), Spearman rank-order correlation coefficient (SRCC) and Pearson linear correlation coefficient (PLCC).

Some of the metrics are used on red, green and blue colour components of the image. Fig-3 shows the input image and Table-1. shows the different performance metric scores obtained for IFS for different distortions. Fig-4. Shows the image on applying different types of distortion.

Fig-3: Input Image

	AWGN	JPEG	JP2K	GB
IFS Score	0.927555030274699	0.734249810189218	0.980218219100393	0.912458036730783
MD				
RED	0	254.3073	14	120
GREEN	0	219.4511	-1	129
BLUE	0	207.3228	-1	119
MSE				
RED	2.3384e+03	7.0383e+03	20.3272	90.8929
GREEN	2.0783e+03	2.7232e+03	73.9990	88.5600
BLUE	1.6530e+03	1.9229e+03	104.2163	89.6898
RMSE				
RED	48.3567	83.8945	4.5086	9.5338
GREEN	45.5883	52.1847	8.6023	9.4106
BLUE	40.6575	43.8513	10.2086	9.4705
NK				
RED	1.4938	0.0039	0.9649	0.9889
GREEN	1.8414	0.0038	1.1642	0.9768
BLUE	1.9090	0.0038	1.2319	0.9678
PSNR				
RED	14.4417	9.6561	35.0500	28.5455
GREEN	14.9537	13.7799	29.4385	28.6584
BLUE	15.9480	15.2911	27.9514	28.6034
MAE	9.9423e+04	5.5940e+04	6.0808e+04	5.6367e+04
SROCC	1.4183e+05	1.4711e+05	1.4804e+05	1.4448e+05
PLCC				
RED	1.4183e+05	1.4711e+05	1.4804e+05	1.4448e+05
GREEN	1.4183e+05	1.4711e+05	1.4804e+05	1.4448e+05
BLUE	1.4183e+05	1.4711e+05	1.4804e+05	1.4448e+05

a) Additive White Gaussian Noise

b) JPEG Compression

c) JPEG 2000 Compression

d) Gaussian Blur

Fig-4: Input Image on applying different distortions.

8. FUTURE ENHANCEMENT

This paper is implemented based on Reduced-Reference methodology, because of this, the memory utilized is reduced further as compared to Full-Reference image quality assessment. As patches input will be reduced to just 20%.

Though this reduces the complexity and the utilization of memory in a system, it does still use considerable memory.

Further it still requires a clean image for reference, which usually is not the case in reality.

In future we need to implement the algorithm that takes no reference images for assessing the quality and further if the same algorithm takes much lesser memory, it can be a better enhancement too.

9. CONCLUSION

This algorithm outputs the Quality Score and also show the Performance Metrics to easily comprehend the quality of the image. Due to its reduced memory usage and lower complexity, this algorithm is a great improvement over a Full reference Image Quality Assessment algorithms that consider all images and patches which sometimes might not be required, thus wasting a lot of memory processing them. This also enhances the time taken to evaluate the images, and hence it is faster.

REFERENCES

[1] Chang, Hua-wen, Qiu-wen Zhang, Qing-gang Wu, and Yong Gan. "Perceptual image quality assessment by independent feature detector." *Neurocomputing* 151 (2015): 1142-1152.

[2] Olshausen, Bruno A. "Principles of image representation in visual cortex." *The visual neurosciences* 2 (2003): 1603-1615.

[3] Caywood, Matthew S., Benjamin Willmore, and David J. Tolhurst. "Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning." *Journal of Neurophysiology* 91, no. 6 (2004): 2859-2873.

[4] Mohammadi, Pedram, Abbas Ebrahimi-Moghadam, and Shahram Shirani. "Subjective and objective quality assessment of image: A survey." *arXiv preprint arXiv:1406.7799* (2014).

[5] Chang, Hua-wen, and Ming-hui Wang. "Sparse correlation coefficient for objective image quality assessment." Signal processing: Image communication 26, no. 10 (2011): 577-588.

[6] Moorthy, Anush Krishna, and Alan Conrad Bovik. "Visual importance pooling for image quality assessment." IEEE journal of selected topics in signal processing 3, no. 2 (2009): 193-201.

[7] van Hateren, J. Hans, and Arjen van der Schaaf. "Independent component filters of natural images compared with simple cells in primary visual cortex." Proceedings of the Royal Society of London B: Biological Sciences 265, no. 1394 (1998): 359-366.

[8] N. Ponomarenko, K. Egiazarian, Tampere Image Database 2008 TID2008[Online], Available: ⟨http://www.ponomarenko.info/tid2008.htm⟩.

[9] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, et al., TID2008—a database for evaluation of full-reference visual quality assessment metrics, Adv. Mod. Radioelectron. 10 (2009) 30–45.

[10] A. Hyvärinen, E. Oja, A fast fixed-point algorithm for independent component analysis, Neural Comput. 9 (1997) 1483–1492.

[11] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Trans. Neural Netw. 10 (1999) 626–634.

[12] J.J. Atick, Could information theory provide an ecological theory of sensory processing? Netw.: Comput. Neural Syst. 3 (1992) 213–251.

[13] Y. Dan, J.J. Atick, R.C. Reid, Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory, J. Neurosci. 16 (1996) 3351–3362.

[14] H.B. Barlow, Possible principles underlying the transformation of sensory messages, Sens. Commun. (1961) 217–234.

[15] J.H. van Hateren, A. van der Schaaf, Independent component filters of natural images compared with simple cells in primary visual cortex, Proc. R. Soc. Lond. Ser. B—Biol. Sci. 265 (1998) 359–366.

[16] A.J. Bell, T.J. Sejnowski, The independent components of natural scenes are edge filters, Vis. Res. 37 (1997) 3327–3338.

[17] M.S. Caywood, B. Willmore, D.J. Tolhurst, Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning, J. Neurophysiol. 91 (2004) 2859–2873.

[18] D.J. Field, What is the goal of sensory coding? Neural Comput. 6 (1994) 559–601.

[19]] D.J. Field, Relations between the statistics of natural images and the response properties of cortical cells, J. Opt. Soc. Am. A 4 (1987) 2379–2394.

[20] E.P. Simoncelli, B.A. Olshausen, Natural image statistics and neural representation, Annu. Rev. Neurosci. 24 (2001) 1193–1216.

[21] Z. Wang, X. Shang, Spatial pooling strategies for perceptual image quality assessment, in: Proceedings of

the IEEE International Conference on Image Processing, 2006, pp. 2945–2948.

[22] A.K. Moorthy, A.C. Bovik, Visual importance pooling for image quality assessment, IEEE J. Sel. Top. Signal Process. 3 (2009) 193–201.