International Conference on Latest Advancements & Future Trends In Engineering, Science & Management
Sambhram Institute of Technology, Bangalore (India)
03rd - 04th May 2019 www.conferenceworld.in ISBN : 978-93-87793-86-6

Streaming Video Quality Assessment and compression in Digital TV under No-Reference Conditions

Dr.S.AnanthaPadmanabhan[1]Dr.M.S.Suresh[2] K.Ezhilarasan[3]

Deepa Gowda N[4]

[1]Professor Electronics Dept.Gopalan Collegeof Engineering, VTU Bangalore india
[2]ProfessorElectrical Dept Sir MVIT Bangalore india VTU
[3]ProfessorElectronics Dept Sambram institute of Technology VTU Bangalore india
[4] Prof.Deepa Gowda N Electronics Dept. Sea College Of Engineering And Technology VTU Bangalore

ABSTRACT:

Assessment of Video (Image) quality has always been a subject of great concern for people engaged in Digital Television, Medical Imaging etc. establishing reliable. While quality assessment in the presence of reference images is an established discipline in the domain of Image Processing, such an assessment in the absence of reference is an extremely difficult task and remains an esoteric subject even today. In this paper, we research and formulate a fairly reliable mechanism of assessing perceptual video quality under no-reference conditions.

Compression of Video Frames is a heavily researched domain in Digital TV Technology and hundreds of researchers have experimented with various techniques of transform coding and quantization mechanisms with varied results. Despite this fact, there is still a tremendous amount of scope for better compression schemes that seek to reduce the utilized band of the spectrum. In real time applications such as Digital TV and streaming Multimedia applications, time plays a crucial role and a successful implementation usually makes a compromise between factors such as (1) quality of reconstructed frames, (2) amount of utilized band owing to compression and (3) ability of the system to cope with increasing frame rates and advanced video profiles that are creeping into Digital TV Standards.Our objective in this research paperis to focus on ways and means of increasing the redundancies by employing Linear Neural Network techniques with feed-forward mechanisms. Our research program targets to achieve the objectives: (1) better compression ratios with even fairly noisy frames (2) better quality of reconstructed frames in terms of very large PSNR and RMS Error tending towards values extremely small and nearer to 0 and (3) honoring the time constraints imposed by frame rates.

Keywords: Burrows Wheeler Transform (BWT), Discrete Cosine Transform (DCT),Neural Networks, Peak Signal to Noise Ratio (PSNR), Root Mean Square Error (RMS) , Rosseta Vector.

ISBN : 978-93-87793-86-6

1. Introduction and Existing Framework:

Traditionally, image quality has been evaluated by human subjects. This method, though reliable, is expensive and too slow for real-world applications. So this presentation is about *objective* image quality Assessment ,where the goal is to have computational models that can automatically predict perceptual image quality. The idea behind this research work is many researchers over the world have experimented with various quality metrics in spatial, Fourier and Wavelet Domains. Structural properties have been mainly the theme in all these activities. Notable amongst the researches are those by *Charles Poynton*, *Hu & Wang* and *Crosier*.

The existing implementations of Compressors and Decompressors (**CODEC** for short) follow (and expected to follow) and adhere to ISO Standard 13818. The standards have evolved from MPEG-1 to MPEG-2, MPEG-4, MPEG-7 and MPEG-21. MPEG2 implements the CODEC based on DCT[1].There are implementations of CODEC based on H264[2] Standard which uses Wavelets instead of DCT. The shortcomings of existing CODECs are: they are not adaptive to the patterns of the pixel population in the video frames. MPEG2 uses a flat and uniform quantization while other implementation uses Vector Quantization techniques. It is important to observe here that Vector Quantization, while being marginally superior in terms of compression ratio, has substantial computing overheads which offset the gain resulting from better compression ratio for a given visual quality metric. There are however various esoteric mechanisms proposed by researchers but none of them are sufficiently suitable for a straightforward and efficient implementation. When we consider the process of decoding pictures at the receiver end, every millisecond of time gained in the decoding process leads to better bandwidth utilization and overall efficiency.

The biggest drawback with these approaches are

1.Performs well across a limited class of images

2.The quality index generated from these approaches not reliable enough

3.The detected quality factor does not seem to agree well with the Human Visual System always

There is therefore a genuine need for better mechanisms of compression which will seek to achieve (1) *better compression ratio* (2) *faster decoding of pictures at the receiver end* and (3) *better quality of reconstructed frames for a given quality factor in terms of Mean Square Error and Peak Signal to Noise Ratio*.

2. Proposed Framework:

Our approach is based on various well established quality metrics which are not applied in isolation but taken together as a whole in establishing the quality. These metrics employed are:

2,1 Our Design Philosophy of No-Reference System.1.Edge Strength that has visual appeal **2**. Contrast and Brightness that are well balanced **3**. Invariant Moments that provide useful pixel spread and Density **4**. Kurtosis that provides vital information about skewness and geometry .

ISBN : 978-93-87793-86-6

Edge strength can be conveniently expressed in terms of the 2-dimensional Laplacian Masked Image.

$$\nabla^2 g(x,y) = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} \qquad (1)$$

Where g(x,y) represents a 2D image. The discrete equivalent of this equation can be employed to ascertain the edge strength is representative of the level of pixel strength in edges.

2.2 Contrast and Brightness:

Contrast and brightness can be established from a 2-dimensional DCT whose equations are given below:

The DC coefficient computed from the Forward DCT Equation is proportional to sum of all the pixels in the MxN subimage. The following values for weighted averaging were used.

$$Y = 0.299R+0.587G+0.144B$$
$$C_b = -0.168736R\pm0.331264G+0.5B$$
$$C_r = -0.5R-0.4186G-0.081312B \qquad (2)$$

2.3 InvariantMoments and Kurtosis :

Image moments are useful to describe objects after segmentation. Simple properties of the image are found *via* image moments include area (or total intensity), its centroid, and orientation information.

The fourth standardized moment is defined as kurtosis gives information of skewness and geometry

The proposed framework is based on adopting an adaptive quantization technique and deploying a linear Neural Network in order to perform the compression. It is a well known fact that quantization controls the amount of information that is being discarded as being non-contributive to the human visual system. A flat quantization as adopted by MPEG2 CODEC does not take into account the significance of the transform coefficients. Owing to vast structural differences in varied types of images, a flat quantizer cannot discriminate between contributing and minimal-contributing coefficients. This being the case, sections of images with excessive details (with significant coefficients spread across the frequency matrix) may lose useful information in the quantization stage and compression may not be the maximum. Therefore, an adaptive quantizer that suits itself to the local characteristics of the blocks performs best.

In this research work, we deploy a linear Artificial Neural Network to generate an adaptive quantizer which runs through limited and predictable iteration cycles to evolve a suitable matrix. The proposed transform coding is the conventional 8x8 DCT, since DCT has very high energy-concentrating properties. Other transforms such as Wavelets, Walsh-Hadamard are also possible. The proposed methodology consists of the following steps:

1. Divide the image into blocks of 8x8
2. For each block the following set of actions are taken in the specified order:
3. The input to the input layer is an 8x8 block (64) pixels
4. DCT is performed on the block, generating an 8x8 Frequency Matrix
5. The default quantization matrix is applied
6. Divide the image into blocks of 8x8
7. For each block the following set of actions are taken in the specified order:

ISBN : 978-93-87793-86-6

8. The input to the input layer is an 8x8 block (64) pixels
9. DCT is performed on the block, generating an 8x8 Frequency Matrix
10. The default quantization matrix is applied
11. The quantized coefficients are reordered using Burrows-Wheeler Transform
12. Huffman Coding is employed to compress the block and the compressed block is stored
13. Inverse Huffman Coding is used to decompress the compressed output of the previous stage
14. Inverse Burrows-Wheeler Transform is applied to the output of the above step
15. The inverse quantization is applied to regenerate the transform coefficients
16. Inverse DCT is performed to reconstruct the 8x8 sub-image and stored
17. The PSNR Value is calculated on the sub-images of steps 7 and 11
18. Adjust the values in the quantization matrix (Weight Adjustment)
19. Go to step 3 if PSNR does not meet the prescribed threshold

At first sight, it appears that the computational steps involved are too high and time consuming but in reality, the situation is different. Most of the modern day processors support threading concept so that several activities (that are mutually exclusive) can run in parallel. Practical implementations of CODEC intended for commercial use employ Digital Signal Processors with pipelines that run in parallel. Even the most modest DSP supports pipelines and parallel processing. Floating point DSPs can execute floating point instructions in a couple of processor clocks. This being the case, we will be justified in assuming that the apparently long sequences of processing steps mentioned above are of no consequence.The following pages describe the researched work in detail under the following various sections:

- Brief Exposition of lossy compression based on transform-coding (specifically DCT)
- Burrows-Wheeler Transform and Adaptive Quantization
- Structure of the proposed Neural Network
- Implementation of the Neural Network
- Results obtained and comparison charts

3. Lossy Compression of Video Frames and Transform Coding.

With lossy compression, we achieve a reduction in data volume, i.e. the compressed data requires much less space for storage. This in turn implies that the amount of information transmitted per second is less. Digital TV frames are streaming in nature and the frames are displayed on the display either at 30 or sometimes 60 frames/ second depending on the MPEG Profile adopted. This means that a tremendous amount of data has to be transmitted per second and reducing this amount will definitely ease the efforts required by the transmitting equipment. We will elaborate the above situation by introducing the concept of bandwidth. In Digital TV terms, a bandwidth can be loosely defined as the number of bits required to be transmitted per second. The following table illustrates the bandwidth requirements for various types of Digital TV display formats:

Table1 : Data comparison

1	2	3	4 = [2] x [3]	5 = [4] x 24	6	7 = [5] x [6]
No	Display Width Pixels	Display Height Pixels	Total Pixels	Total Number of bits at 24 bits/ pixel	Frame Rate	Number of Bits to be transmitted per second
1	352	288	101376	2433024	30	72990720
2	720	576	414720	9953280	30	298598400
3	720	608	437760	10506240	30	315187200
4	960	576	552960	13271040	30	398131200
5	1440	1152	1658880	39813120	60	2388787200
6	1920	1152	2211840	53084160	60	3185049600

From the above table, it is obvious that as the display size increases or the frame rate increases the total number of bits to be transmitted per second increases dramatically. From 1 (low profile display) to 6 (high definition profile), the values in the last column vary between 72 million to 3 billion bits per second. Even for a low profile display, 72 million bits per second is no small a figure. It is therefore important to compress the data before transmission. Compression by even about 50% (can be easily achieved) can exhibit dramatic improvement in performance. The following pages elaborate on the theme of lossy compression.

$$D_{original} \not\equiv D_{decompressed}$$

These axioms are simple enough and don't need any explanation. Lossless compression is required for compression of data files containing numeric values, executable/ DLL files containing code and text files containing readable information. Even a slight loss of information will render data useless for consumption. Popular compression and archival techniques. Thus, transform coding and quantization play a major role in making image compression attractive.The following paragraphs seek to explain very briefly these two concepts.

2-Dimensional Discrete Cosine Transform

The Discrete Cosine Transform and its inverse for 2D Images s defined by the equations:

$$F(u,v) = \frac{2}{\sqrt{WH}} C_u C_v \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} p(x,y) * \cos\frac{2x+1}{2W}\pi u * \cos\frac{2y+1}{2H}\pi v \quad \textbf{Forward DCT}$$

(3)

$$p(x,y) = \frac{2}{\sqrt{WH}} \sum_{u=0}^{W-1} \sum_{v=0}^{H-1} C_u C_v * F(u,v) * \cos\frac{2x+1}{2W}\pi u * \cos\frac{2y+1}{2H}\pi v \quad \textbf{Inverse DCT}$$

and H its height

(4)

The invention of this transform is attributed to**Ahmed** and **Natarajan** [13].This transform is remarkable in several ways – specifically for its energy concentrating properties in the lower order coefficients, generating a large amount of sparseness through quantization. DCT is the corner stone of innumerous compression algorithms and commercial compression programs for video data. It is interesting to note that while DCT is used extensively in lossy compression mechanisms, DCT by itself is perfectly reversible with absolutely no loss except for floating point imprecision and rounding errors which are usually negligible[3][4][5]. The real loss in lossy compression comes from the quantization stage where integer division truncates the fractional part of the coefficients.

ISBN : 978-93-87793-86-6

4. Burrows Wheeler Transform (BWT)

In transform coding, one usually rearranges the transform coefficients in the increasing order of frequency coordinates, the premise being that as the frequency coordinate increases, the transform coefficients tend to decrease, losing significance. With such rearrangement, one usually finds the coefficient values monotonically decreasing in the increasing coordinate direction. This implies an increased level of correlation, leading to better compression. But in certain images with unusual structural features, this is far from truth.

4.1 The Burrows-Wheeler Transform Basics

Michael Burrows and David Wheeler released a research report in 1994 discussing work they had been doing at the Digital Systems Research Center in Palo Alto, California. Their paper, "A Block-sorting Lossless Data Compression Algorithm" presented a data compression algorithm based on a previously unpublished transformation discovered by Wheeler in 1983.

The input to BWT algorithm is a block of data that is rearranged using a sorting algorithm. The output of BWT isa block that contains exactly the same original data that was input, differing only in their ordering. The transformation is reversible in the sense that the original ordering of the data elements can be restored without loss of information integrity.The BWT [6]is performed on an entire block of data at once. Commonly known lossless compression algorithms operate in the streaming mode, reading a single byte or a few bytes at a time. With BWT one can operate on large blocks of data.

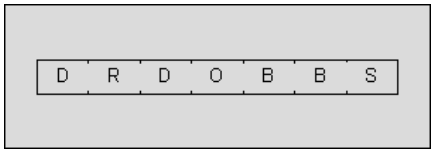


Figure 1 (An sample data set)

To illustrate the workings of BWT, consider the string shown in figure 1 which has 7 bytes of data. We start with the original string and build more strings by rotating one character to the right at a time. So, we end up with 7 strings (including the original). This is depicted in figure 2 below.

String 0	D	R	D	O	B	B	S
S 1	R	D	O	B	B	S	D
S 2	D	O	B	B	S	D	R
S 3	O	B	B	S	D	R	D
S 4	B	B	S	D	R	D	O
S 5	B	S	D	R	D	O	B
S 6	S	D	R	D	O	B	B

Figure 2 (The set of strings associated with the buffer)

The next step is to perform a lexicographical sorting of inputs (usually a *Quick Sort implementation*). There are two important points to note in this picture. First, the strings have been sorted, but we've kept track of which string occupied which position in the original set. So, we know that the String 0, the original unsorted string, has now moved down to row 4 in the array as shown in figure 3.

ISBN : 978-93-87793-86-6

	F						L
S 4	B	B	S	D	R	D	O
S 5	B	S	D	R	D	O	B
S 2	D	O	B	B	S	D	R
S 0	D	R	D	O	B	B	S
S 3	O	B	B	S	D	R	D
S 1	R	D	O	B	B	S	D
S 6	S	D	R	D	O	B	B

Figure 3 (The set of strings after sorting)

We tag the first and last columns with special designations F and L for the first and last columns of the array block. Column F contains all the characters in the original string in sorted order. So our original string "DRDOBBS" is represented in F as "BBDDORS".The characters in column L don't seem to follow any ordering, but interestingly they have a notable property. Each of the characters in L is the *prefix character* to the string that starts in the same row in column F. The actual output of the BWT, oddly enough, consists of two things: a copy of column L, and the *primary index*, an integer indicating which row contains the original first character of the buffer B. So performing the BWT on our original string generates the output string L which contains "OBRSDDB", and a primary index 5. The primary index 5 can be ascertained easily since the original first character of the buffer will always be found in column L in the row that contains S1.Since S1 is simply S0 rotated left by a single character position, the very first character of the buffer is rotated into the last column of the matrix. Therefore, locating S1 is equivalent to locating the buffer's first character position in L. At first sight the above transform doesn't seem reversible. Generally sorting routines are not reversible, so a reversal of Quick Sort doesn't exist – you will never be able to get the original ordering by any means. But in the case of BWT, a reversal is indeed possible through a transformation vector. The transformation vector is an array that defines the order in which the rotated strings are scattered throughout the rows of the matrix of Figure 3. The transformation vector, *T*, is an array with one index for each row in column F. For a given row *i*, T[i] is defined as the row where S[i + 1] is found. In Figure 3, row 3 contains S0, the original input string, and row 5 contains S1, the string rotated one character to the left[8]. Thus, T[3] contains the value 5. S2 is found in row 2, so T[5] contains a 2. For this particular matrix, the transformation vector can be calculated to be { 1, 6, 4, 5, 0, 2, 3 } [7].

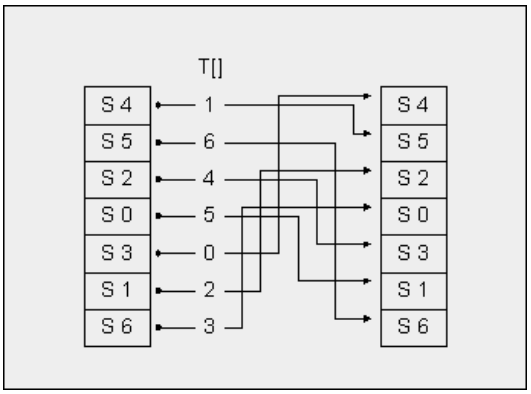


Figure 4 The transformation vector routes S[i)to S[i + 1])

Figure 4 shows how the transformation vector is used to walk through the various rows. For any row that contains S[i], the vector provides the value of the row where S[i + 1] is found. This transformation vector is called '*Rosetta Vector*'.Therefore, given a copy of L, we can calculate the transformation vector for the original

input matrix. And consequently, given the primary index, you can recreate S0, or the input string. The possibility of calculating the transformation vector requires only that we know the contents of the first and last columns of the matrix. This is simply having a copy of L.

5.Adaptive Quantization Mechanism

The contents of the white cell (with 1) will remain unchanged at all quantization levels. All the values in the gray colored cell with increase by 1 (stepping size) at each increase in quantization level. The columns and rows are reckoned with starting value 0, so that rows (**R**) have the range 0 → 7 and cols (**C**) have range 0 → 7. So, given **R** and **C,** we can calculate the cell value using the following equation shown in Figure 5.

1	1	1	2	3	4	5	6
1	1	2	3	4	5	6	7
1	2	3	4	5	6	7	8
2	3	4	5	6	7	8	9
3	4	5	6	7	8	9	10
4	5	6	7	8	9	10	11
5	6	7	8	9	10	11	12
6	7	8	9	10	11	12	13

Figure 5 Quantization Matrix (Level 0)

The next step is to perform a lexicographical sorting of inputs (usually a *Quick Sort implementation*). There are two important points to note in this picture. First, the strings have been sorted, but we've kept track of which string occupied which position in the original set. So, we know that the String 0, the original unsorted string, has now moved down to row 4 in the array[8]. BWT on our original string generates the output string L which contains "OBRSDDB", and a primary index of 5.

$$(R<3)*((C<(3-R))*1+(C>(2-R))*(C+R-1))+(R>2)*(R+C-1) \quad (5)$$

The logic of keeping the values in the white cells unchanged is fairly straightforward. The coefficients at the lower frequency coordiante end are contributing terms and the 6 coefficients in these positions are therefore significant in almost all the cases.

In implementing this quantizer, we employ a increment step size of 1, though larger values are possible. Larger step sizes lead to coarser and more approximated images with smaller values of PSNR and this is not preferable, as our objective is to achieve two distinct but conflicting goals – *better decompressed quality* and *larger compression ratios*.

6. NEURALNETWORK WORK

Figures 6-8 shows the methodology involved

The decoding process is exactly the inverse process of encoding

ISBN : 978-93-87793-86-6

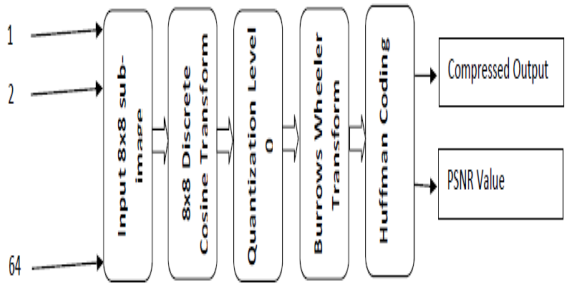


Figure 6.Layout of the Linear Neural Network

1. Determine the quantization level from Block Header and generate the quantization matrix
2. Use Inverse Huffman Coding to decompress the compressed output
3. Use Inverse Burrows-Wheeler Transform to the output of the above step
4. Apply inverse quantization to regenerate the transform coefficients
5. Apply Inverse DCT to reconstruct the image and store it

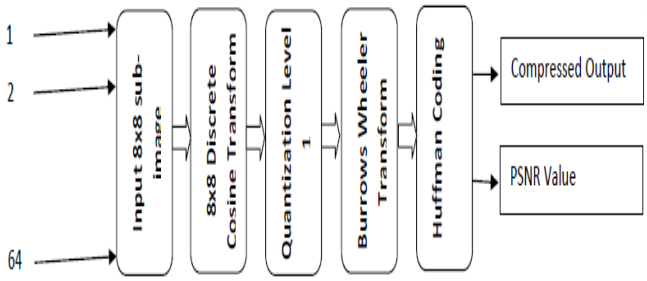


Figure 7. ANN Compression Network

The decoding process is exactly the inverse

A few important observations are in place. Unlike the traditional ANN, [9] training sets and training are not necessary. The number of iterations for converging to a solution is finite and limited to 10 only. [10]

7 Results of the Network with Comparisons (Ours vs. MPEG2)PSNR threshold at 4500 dB

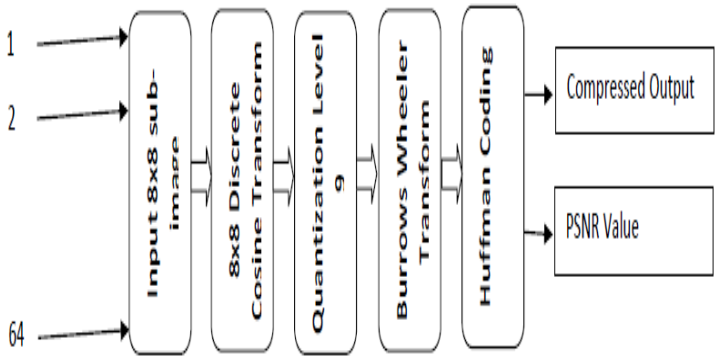
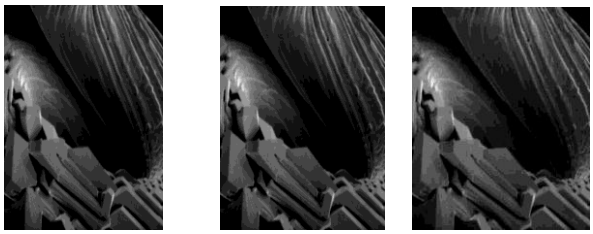


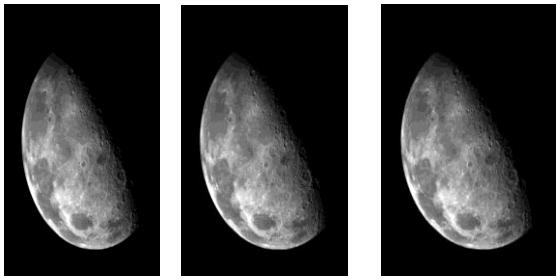
Figure 8. Intermediate Levels 2 to 8

ISBN : 978-93-87793-86-6

Original Image 610 KB | ANN Compression 61KB | MPEG2 Compression 124KB



Original ImageANN 492 KB | Compression 41KB | MPEG2 Compression 102KB



Original Image 398 KB | MPEG2 21KB | ANN 97KB

Figure 9. Comparisons of Various Methods

The results in figure 9. show that the ANN compression method has a better compression as compared MPEG-2 standards[11] and the consistency has been seen for various kinds of frames with slow changes , medium changes and fast changes. The compression ratio is almost 1:50

Methodology of Analysis:

1. EdgeStrength=0.45(E)2.ContrastandBrightness =0.38(C)3.HueandSaturation=0.55(H)4.Invariant Moments =0.44 (I) 5. Kurtosis=0.49 (K) . The parameters E, C, H, I & K are represented on a normalized scale of 0 to 1.0. So, the quality index is as follows.

$$Quality\ Index = 0.45 * E + 0.38 * C + 0.55 * H + 0.44 * I + 0.49 * K$$

Edge Strength	0.53
Contrast and Brightness	0.49
Hue and Saturation	0.46
Invariant Moments	0.46
Kurtosis	0.69
Quality Index	6

Figure 10 quality index existing method

Edge Strength	0.41
Contrast and Brightness	0.52
Hue and Saturation	0.34
Invariant Moments	0.57
Kurtosis	0.71
Quality Index	7

Figure 11 Quality index proposed method

The results for two images are as shown in figure 10 and 11 where the quality index is good and all experiments were implemented using C++.The proposed method gives better quality index and also the neural net method the soft computing technologies gives better PSNR ratio.[9]

8 Conclusion:

Our research experiments lead us to conclude that stable No-Reference quality measure has to take into account all the metrics of quality, giving weighted importance depending upon the level of contribution to quality. The results obtained seem very encouraging to enable us to proceed further in refining the assessment process.

On all the image files without exception, ANN compression produced better compression ratios for a given quality factor. For evaluation of performance in compression time, we performed the compression of several images both by ANN and MPEG2 Codec 500 times in a loop to determine the average time taken. While time for compression understandably varies across different images, the performance of ANN was much better than that of MPEG2 video codec. The algorithm presented here lends itself very easily for implementation both in hardware and software. Future enhancements and extensions to this research work could include deployment of an additional layer clustering based on Markov's Chains. The structure of the implementation is such that any type of transform can be employed in the Neural Network, potential candidates being *wavelets* (in all its variations), *Hadamard* and *Walsh Transforms*.

9 References:

[1] Creusere,C.D., Mitra,S.K., Image-CodingUsing Wavelets Based on Perfect Reconstruction IIR FilterBanks,CirSysVideo(6),No.5,October 1996, pp. 447-458.

[2] Sahiner,B., Yagle,A.E., ImageReconstruction fromProjectionsunderWaveletConstraints,TSP(41), No. 12,1993,pp.3579-3580.

[3] Creusere,C.D., Mitra,S.K., Image- Coding Using Wavelets Based on Perfect Reconstruction IIR FilterBanks,CirSysVideo(6), No. 5, October 1996, pp. 447-458

[4] Dewitte,S., Cornelis,J.P.H., LosslessIntegerWaveletTransform ,SP Letters(4), No. 6, June 1997, pp. 158-160.

[5] Rabadi,W.A., Myler,H.R., IterativeImageReconstruction:AWaveletApproach,SPLetters(5),No.1,January1998,pp.1-3.

[6] Liew,A.W.C., Law,N.F., Nguyen,D.T., DirectReconstructionMethodforWaveletTransformExtremaRepresentation,VISP(144), No. 4, August 1997, pp. 193-198. 9806

[7] Kim,H.J., Li,C.C., LosslessandLossyCompressionUsingBiorthogonalWaveletTransformswithMultiplierlessOperations,CirSysSignal(45),No.8,August1998, pp. 1113-1118. 9809BibRef

[8] Kim,H.J., Compression Using Biorthogonal WaveletTransformswithMultiplierless2DFilterM

ISBN : 978-93-87793-86-6

askOperation,ICIP97(I:648651).

[9] Charles poyntonarles, Theory of Perceptual Quality Assessment, Academic Press 1998

[10] Hu & Wang ,Princapal Component Analysis in Image Quality Morgan ,Kaufmann 2004

[11] Crosier, Estaban & GallandStructural approach to Image Quality Assessment,Prentice Hall 2008

[12] Fellgett, P.B., Linfoot, E.H., On the Assessment of Optical Images,Royal(A-247), No. 931, 1955, pp. 9-407.

[13] Ahmed and Natarajan and K. R. Rao, "Discrete cosinetransform",IEEETransactionsonComputers, Vol. C-23, 1974, pp. 90–93.