IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 23, Issue 1, Ser. I (Jan. – Feb. 2021), PP 36-43
www.iosrjournals.org

Using ANN to Enhance Error Concealment in Digital Image Decoding

Dr. Brice EKOBO AKOA* ,
Electronic, Mechatronic and Signal Processing Laboratory, NASEY, University of Yaounde 1, Yaounde, Cameroon

Abstract
This paper proposes an approach to perform error concealment based on a Video Quality Measurement Tool (VQMT) correlated with the Human Visual System (HVS). The VQMT is expected to replace human judgment when rating video quality in subjective experiments. To the best of our knowledge, this is the first time that results are reported on image quality improvement algorithms, which have been dynamically controlled by a neural network based machine learning process, after having been carefully trained on video and image databases with associated subjective quality ratings (MOS). The underlying study consists of selecting fundamental quality metrics based on Human Visual System models and using artificial intelligence solutions as well as statistical analysis. This new combination enables suitable video quality ratings while taking as input multiple quality metrics. The designed tool is based on a neural network based machine learning process and will also be relevant for image quality evaluation. The efficiency of the proposed method is demonstrated by comparing their results with those of existing work done on synthetic video artefacts. The VQMT is used to optimize an error concealment algorithm applied on image (that means only with the spatial aspects of the video) with a wiener filter in dynamic mode, where coefficients are adjusted based on the optimization criteria of least mean squares. The performances of the resultant error concealment algorithm are measured and results are discussed.
Keywords: *Error Concealment; Video Quality; Artificial Neural Network; Peak Signal to Noise Ratio; Blur Metric; Visual Artifacts.*

Date of Submission: 18-01-2021 Date of Acceptance: 03-02-2021

Introduction

With the evolution of multimedia technologies and the users requirements for quality pictures and videos, digital decoders must provide satisfactory quality to decoded videos. Therefore, a new challenge arises in the field of image processing which is to ensure a good perceptual quality to the images in the output of monitors and other multimedia devices and, with respect to human visual perception. It is in this sense that this paper establishes a loop of visual artefacts concealment in the output of a digital decoder, using a VQMT (video/image quality measuring tool) to enhance error concealment performances and to reach a satisfactory level of visual quality in regard with human visual system.

This manuscript presents the results of an additional research work which is a continuation of an idea developed in [1] describing the VQMT. In the present manuscript we try to prove that the VQMT can help in optimizing any error concealment algorithm in the sense of the visual aspect of HVS, based on a practical example using the Weiner filter, while the article [1] just presents the realization of the VQMT.

We are interested here in blur and noise artefacts, which are the most influential and well known in digital video and image decoding. To reduce the complexity of this project, we have just considered spatial aspects of a video sequence, so the work described here is done in an image (a video can then be seen as a group of images).

The usage of artificial intelligence helped in various do-mains to overcome optimization issues in Microelectronics [1-3]. This idea has also been carried in image and video processing to help detecting and concealing visual artefacts. In [1,4], the image quality evaluation by the VQMT is compared to subjective evaluation through mean observers scores (MOS) given by humans during a subjective experiment. By so doing, the expectations of human eye are better attained. The rest of the paper is organized as follows:

Section 2 describes existing works, and presents the database used and selected metrics; section 3 compares the VQMT with few existing tools; section 4 presents the error concealment algorithm, and section 5 draws a conclusion.

---

*Permanent address: Department of Elect. & Telecom. Engineering, NASEY (National Advanced School of Engineering of Yaounde), P.O Box 8390, Yaounde, Cameroon; Email address: *ekobobrice@gmail.com*, *brice.ekobo@univ-yaounde1.cm*

Previous Works

The most reliable method [5] in assessing the quality of a video is subjective assessment. The video streams are viewed by a group of observers who will rate them according to a given scale. Subsequently, a Mean Opinion Score (MOS) is computed as the average of the scores given by observers. This MOS represents an index on the visual quality level of the video. However, this method cannot be applied to real time scenarios directly, due to some time-consuming repetitive tasks. So it's not practical to use such method of Image or Video Quality Evaluation (VQA) for image error concealment algorithm in digital decoding process.

Some researchers have thus proposed ways of extrapolating subjective VQA towards objective VQA [6-9]; as well as [10] where a reference metric is created using neural network approach. The achieved results demonstrate strong correlation with MOS. Among these objective VQAs, there are solutions that need to refer to the original version of the video sequence to rate the quality, called full-reference VQA; those in which just little information about the original video sequence is required, are called reduced-reference VQA; and those in which only the video to assess is needed together with information about its distortion, are called no-reference VQA.

2.1 VQMT Database and Metrics Description

The EPFL database described in [11] provides the video on which the experiences have been carried. The 4 quality metrics selected are : PSNR, PLR, SI and blurMetric. The database and the 3 first quality metrics PSNR, PLR and SI are described in [1, 4].

The blurMetric:

The metric used in this project to measure the amount of blur was established by F. Crete and Al. [12], observing that the augmentation of blur effect in an image varies depending on the amount of blur already present in this image. The algorithm for computing the blur metric "*blurMetric*" consists in blurring that image again with a low pass filter and comparing the differences between adjacent pixels before and after passing the low-pass filter.

Let U be the luminance component of an image or a video frame of size of $m \times n$ pixels. To summarize, we obtain a no-reference perceptual blur estimation ranging from 0 to 1 which are respectively the best and the worst quality in term of blur perception. The **Error! Reference source not found.** shows the simplified flow-chart for the computation of the blur metric. The convenience and simplicity of design of this metric are the reasons for the choice of its use in this project.

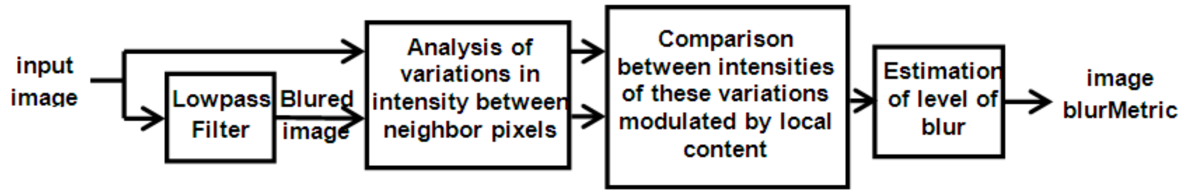

Fig. 1 : Simplified flow-chart of the blur estimation principle

2.2 ANN-based VQMT

The article [4]describes the work done on the elaboration of the ANN-based VQMT. In the scope of the current paper, a supervised learning neural network system has been preferred. The input features correspond to video streams and the target features correspond to their respective MOS values, both taken from the database available in [5].

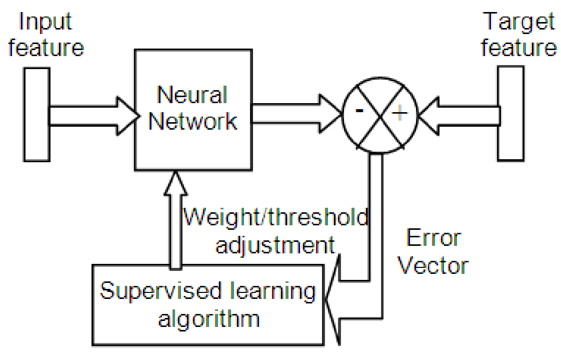

Fig. 2 : Supervised learning scheme

An ANN-based VQMT have been developed with an architecture of 4 inputs corresponding to the 4 metrics (PLR, PSNR, SI & blurMetric), 1 output corresponding to the suitable score for the video to rate, 10% of the initial 78 samples have been taken to perform tests and 10% for validation. The ANN implemented has 2 layers with 4 neurons for the first (hidden) layer and 1 neuron for the second layer (Fig. 2). The MSE obtained at epoch 19 is about 5.10^{-2} that's very close to value zero. The regression (R=0.98) obtained as results and the regression proves the good match between experience MOS and the VQMT estimated output scores. Further details on the performances of this ANN designed for the VQMT are described in [3].

VQMT-assisted error concealment algorithm

Given an input image *g*, deteriorated or not, the system would be able to conceal all the visual artefacts appearing on that image and recover the closest as possible the original image, as seen by human eyes. In the rest of the paper, we will only consider MPEG family type of artefacts. To simplify the algorithm, we will consider only two visual artefacts, blur and noise whose influences on image quality are illustrated in **Error! Reference source not found.** and **Error! Reference source not found.** respectively.

3.1 Analysis and Prioritization of Visual Artefacts

The analysis here consists in reproducing the considered MPEG visual artefacts and elaborating metric to measure the two associated type of annoyance. To do this, a synthetic study is done which allows the recognition of the different types of artefacts used here, create distortion parameters and create quality metrics closely related to these parameters.

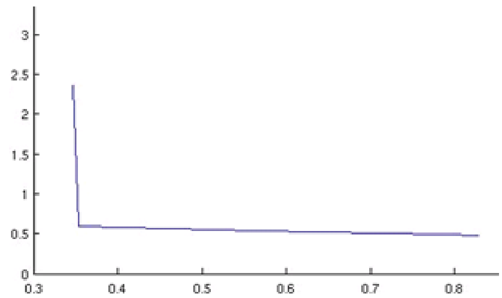

Fig. 1 : Image quality dependence with its blurMetric

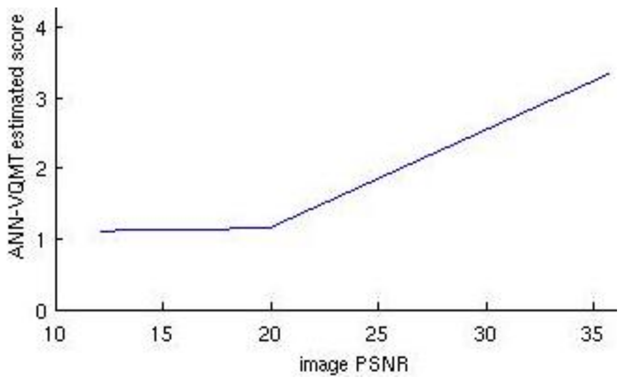

Fig. 2: Image quality dependence with its PSNR

The Created metrics should be normalized. That is to say, the distortions produced on synthetic videos must be indistinguishable from the original artefacts found in the decoding of the MPEG family. The choice of metrics and parameters was done so that each of the two artefacts synthesis is strongly correlated to a given metric. This match will then have some sort discrimination between types of artefacts in prioritizing the use of Spearman coefficients between metrics and quality measurement.

Measurement of "Noise" annoyance

Depending on the type of noise (Gaussian noise, white noise, additive noise, etc.) there are many desired parameters that appear in the synthesis of noise in a video. The easiest way to create an easy to configure as noise is additive Gaussian noise. This method involves adding the input image to a randomly generated signal. For a uniform distribution on the image, a random signal of the Gaussian distribution is selected. Indeed, the Gaussian noise is obtained by adding this signal to each pixel of an image according to a random variable Gaussian probability law (Fig. 3). A Gaussian distribution law of variance σ and mean μ is in the form:

$$G_{\sigma,\mu}(s) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(s-\mu)^2}{2\sigma^2}} \quad \textbf{(1)}$$

By increasing the parameter σ, benefits the image by adding noise is degraded.

Measurement of “Blur” annoyance

In practice, we generated blur annoyance on images using Matlab by *gaussfir* and *imfilter* functions that operate in the following way: create a low-pass Gaussian filter finite-impulse response with the following characteristics:

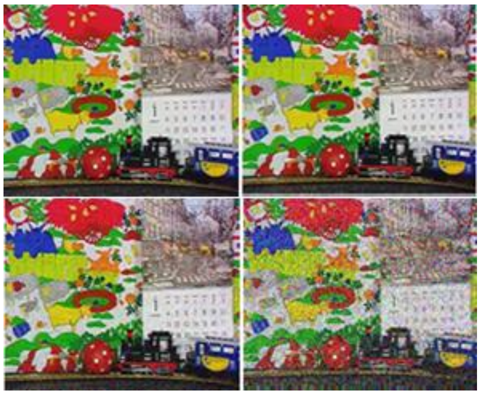

Fig. 3 : A frame of mobile & calendar with additif gaussien white noise : de letf to right : original (from EPFL DB [11]); blurred with σ = 0,001 ; 0 ,01 et 0,1 respectively

BT: the product of the period and the bandwidth at -3 dB. B is the unilateral bandwidth in Hertz; T the period, expressed in seconds.
NT: the number of symbol periods between the beginning of the impulse response of the filter and the maximum point of the signal.
OF: the oversampling factor, the number of samples per symbol. If property is not specified, DE = 2 is used.
This function performs a convolution between the PSF filter and the image *u* which lead to the blurred picture *g* (Fig. 6).

Fig. 4 : A foreman video frame gaussian blurred: from left to right : original (from EPFL DB [11]); blurred with BT = 0,1 ; 0 ,06 et 0,08 respectively

Visual Artefact Prioritization:

However, there are cross-influences between the distortions related with blurriness and those related to noise. That is why a study was conducted to know the degree of interdependence between these two artefacts and to get an idea on the comparison of the degree of their impact on the overall quality of the image. Then the algorithm will start first by removing the artefact given as the more influential on the quality of the video.

We consider as "priority", the artefact with the highest Spearman with the image score computed by the VQMT. This idea implies two assumptions: One, monotony between each metric representing 2 levels of annoyance along the 2 artefacts and VQMT computed score; and the other involves very low cross between metric artefacts correlations, given the interdependence existing between the sources of these artefacts.

Given these cross influences, it would be better to use an error concealment algorithm which involve both annoyance blurriness and noise so that when correction one of these artefacts, we could keep holding control on the other. The adaptive wiener filter is an alternative to this issue. In fact, the Wiener filter does not characterize the signal and the annoyance in their analytical form but by their statistical properties. This is a unique finite impulse response filter adapted to both the attenuation of noise and reduction of blur in an image.

3.2 Artefacts Concealment Algorithm using Wiener filter

Let's consider a degraded image $g = h * u + b$, where u is the original image to restore, h a positive symmetric kernel convolution (impulse answer of "blur" filter) and b is a noise probability distribution (identical for each pixel) μ. Here we take a Gaussian white noise of standard deviationσ_b.

Let's model the problem with a formulation in least squares sense considering the squared error: Minimization of the MSE in (2):

$$\begin{aligned} \|e\|^2 &= \|g - h * u + b\|_{L^2}^2 \\ &= \hat{g}^2 + (\hat{h}\hat{u} + \hat{b})^2 - 2\hat{g}(\hat{h}\hat{u} + \hat{b}) \end{aligned} \quad (2)$$

Research of the optimal solution$\hat{u}$:

In practice, the reference image u is unknown, but its power spectrum Pu is needed in the error correction algorithm and can be estimated. An estimate of the power spectrum Pu of u is obtained from the original image g correct, assuming that both images spectrum are from the same form. So the problem in real can be illustrated as shown in Fig. 7.

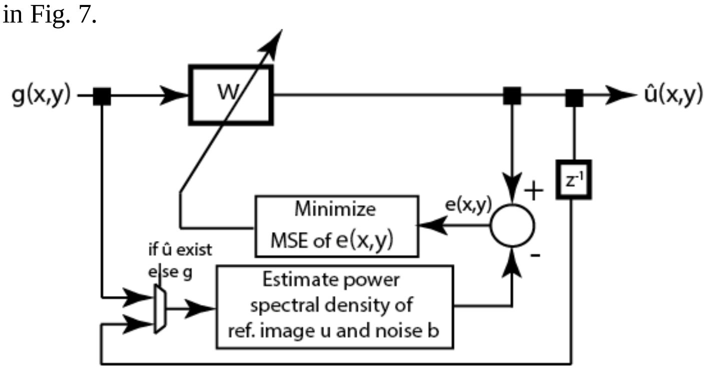

Fig. 5 : VQMT-assisted Error Concealment Algorithm Deseign

Impose optimization constraints in the research of $\hat{u}$ allow to control the parameters of the wiener filter, σ_b(the standard deviation of Gaussian noise) and N (the dimension of h) and thus allow increasing the quality of the optimal image $\hat{u}$

$$g = h * u + b \quad (3)$$

Solution with Minimization of MSE and optimization using gradient function:

Let's solve the equation (3). The goal is to estimate reference image u by its estimation $\hat{u}$ as a function of input image g. Especially, we need to find the LMMSE to estimate $\hat{u}$ based on (all or part of) g. The FIR filter version of this problem is expressed in (4).

$$\hat{u}(x,y) = \sum_{i=n-N}^{n} \sum_{j=m-N}^{m} h_{x+i,y+j} * g(x,y) \quad (4)$$

Where x and y are the pixel coordinates in the image.

We set: $\hat{u}$(x,y) = $\alpha g(x,y) + \beta$

Image and noise are assumed to be independent. We consider an additive Gaussian white noise of varianceσ_b^2.

α and $\beta$are choosen so as to minimize the MSE of e(x,y).The gradient of this new problem can then be calculated as:

$$J(\alpha,\beta) = E\left[\left(\hat{u}(x,y) - u(x,y)\right)^2\right] \quad (5)$$

After derivation and cancelling this derivative to find the extreme values using the fact that $\overline{g(x,y)} = \overline{u(x,y)}$ and σ_g^2=$\sigma_u^2 + \sigma_b^2$, we obtain:

$$\begin{cases} \alpha = \dfrac{\sigma_u^2}{\sigma_u^2 + \sigma_b^2} \\ \beta = (1 - \alpha)\bar{g}(x,y) \end{cases} \quad (6)$$

The estimated image is then:

$$\hat{u}(x,y) = \frac{\sigma_u^2}{\sigma_g^2} g(x,y) + \frac{\sigma_b^2}{\sigma_g^2} \bar{g}(x,y) \quad (7)$$

The choice of N and $\sigma_b^2$is very important. N must be at least equal to 3. These two values are dependent by the expression of LLMMSE filter, which can be interpreted as a local response or mask whose weights w_n $(0 \le n < N)$are functions of local statistics in the input. One can also determinate a local frequency response whose bandwidth is a function of local signal-to-noise variance ratio.

Now let's move to frequency domain: (x, y) are replaced by (u, v) and any lowercase letter referring to a signal or a function in spatial domain is now written in uppercase
The restored image is given by:

$$\widehat{U}(u,v) = W(u,v).G(u,v) \quad (8)$$

The wiener filter is chosen so as to minimize the MSE$E\left[\left|\widehat{U}(u,v) - W(u,v).G(u,v)\right|^2\right]$.
Let's suppose that image is independent of noise. Then the wiener filter is given by:

$$W(u,v) = \frac{H(u,v)}{|H(u,v)|^2 + \dfrac{S_b(u,v)}{S_u(u,v)}} \quad (9)$$

Where $S_u(u,v)$ is the power spectrum of the reference image U and $S_b(u,v)$ is the power spectrum of noise B.

In our case, for an additive Gaussian white noise of variance $\sigma_b^2$for a pixel, and a mxn image resolution, we have:

$$S_b(u,v) = M \times N \cdot \sigma_b^2 \quad (10)$$

Where $\sigma_b^2$is estimated through local variances of low regions (smooth) of low frequency in the image.
$S_u(u,v)$ can be estimated using the reference image G(u,v),since wiener filter is not sensible to low variations of power spectrum and transformations on an image do not (or do very small) change its power spectrum.

3.3 VQMT-assisted Error Concealment: Simulation and results

First method: The cost functions to measure image optimization are artefact metric

For a fixed value of PSF matrix size N (initial value where taken N=3 to the maximum N = 11), we started by concealing blur in the image at the first iteration for the fact that its influence on VQMT as seen in the beginning of this section, is higher than noise influence.

We have proven in the first paragraph of this section that blurMetric and PSNR have monotonous dependencies with image quality measured through ANN-based VQMT. Additionally, these 2 metrics have dependencies on the wiener filter parameters N. But they also have monotonous dependencies on the standard deviation; this is proven on the equations (6) and (7). So we assume on the base of these hypotheses that we can use blur and noise metric as cost functions to our image visual quality optimization algorithm. The simulation results of this method are shown on Fig. 8.

Second method: ANN-based VQMT estimated scores are used as cost function to measure image optimization

An estimation $\hat{u}_i$ of reference image u is computed in each iteration i and the corresponding score $VQMT\_Score_i$ is evaluated by the VQMT in order to measure the sense of evolution of the error concealment process. Then the monotonicity seen between PSF size N and *blurMetric* in the one hand and image *blurMetric* and its VQMT estimated score *on* the other hand allowed us to set an optimal value of N, N_{opt} (corresponding to the iteration i_0 where $\hat{u}$ has the higher score $VQMT\_Score_{i0}$) from which the other wiener filter parameter σ_b have been varied from the initial value 0.01 to 100.

In the same way, the monotonicity seen between the image noise standard deviation σ_b and the image *PSNR* in the one hand and image *PSNR* and its VQMT estimated score on the other hand allowed us to obtained the higher score $VQMT\_Score_{opt}$ as the highest image score possible, and the corresponding $\hat{u}_{opt}$ has been taken as the optimal estimation of the reference image u. The simulation results of this method are shown on Fig. 9.

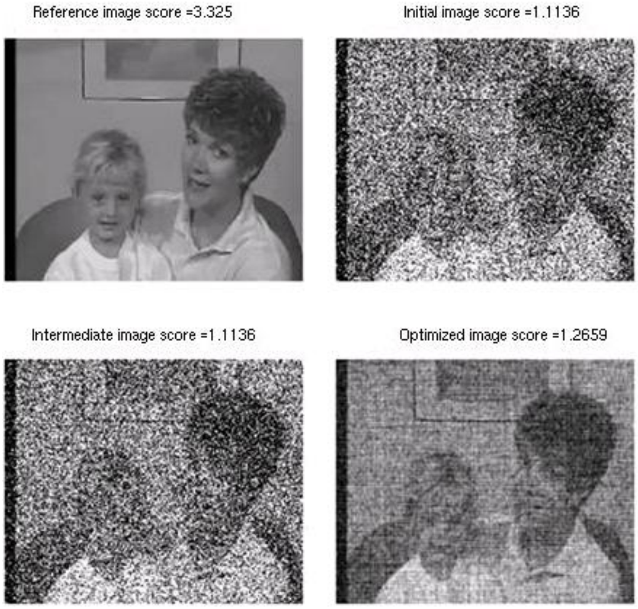

Fig. 6: VQMT-assisted image concealment Results with artefacts metrics taken as cost functions

Comparison with previous works

4.1 Coherence with human judgment

The company Tektronix [13] has developed a next-generation tool for the analysis of the quality images called PQA. This tool is based on the visual system concepts human. His release PQA600 provides a suite of repeatable quality objective measures that closely match the human visual subjective evaluation. These measures provide valuable information to engineers working to optimize compression and decoding digital video or images. On the technique used for the evaluation of the quality of the videos or images, the CAA is also based on the HVS, as the VQMT. Both evaluations of the image quality use the results of network statistics such as loss rate packets. However, the relevance of the results of the VQMT is shown in the development of performance by direct comparison with the subjective assessment of the quality of videos made by humans.

Both experiments are working on the basis of digital decoding of the MPEG family. On the correction of errors, no experimental approach has been made (at least to the public) to show how the PQA using compression and digital decoding videos. However, this paper has shown that the error correction algorithm designed uses the video quality measuring tool (VQMT) in real-time to ensure that the quality desired in image optimization is well appreciated by the human being. This is the main advantage of the correction algorithm assisted by the VQMT.

4.2 The concealment approach improvement

The paper [14] proposes a new method of EC improvement. The main idea is that spatial and temporal correlations are jointly utilized for EC of both intra and inter frames. The proposed EC consists in three main steps: Scene change detection, Motion activity detection and MV retrieval. In a first time, the proposed EC use its scene change detection algorithm to decide whether the scene change occurs or not, then it decide what type of temporal information will be use with a suitable motion activity detection algorithm. Lastly, the MVs are derived with the proposed MV recovery algorithm.

The experimental results have shown that this algorithm reduces artefacts but there is no way to ensure that the quality is being improved in the sense of human eye visual perception. In the opposite, we have proven in this paper how the VQMT ensure a convergence of the reduction of artefacts toward the better visual quality of images.

Conclusion

In this paper, we have described the conception of a new method for video quality measurement tool using an artificial neural network. The developed tool enables objective evaluation of a given video in close correlation with human visual system perception. The VQMT offers a simple and generic architecture that could either be used to monitor video quality in multi-media networks or as a video quality assessment in digital video decoders.

This paper contributes in video quality assessment in creating an objective video quality assessment correlated with human visual system. In this sense, a demonstration have been shown to show how we ensure a visually ameliorated quality to decoded image been corrected by the VQMT-assisted Error Concealment algorithm. We emphasize that the method elaborated here is applicable to every type of video format and competitive in terms of performances with the latest video quality assessment methods.

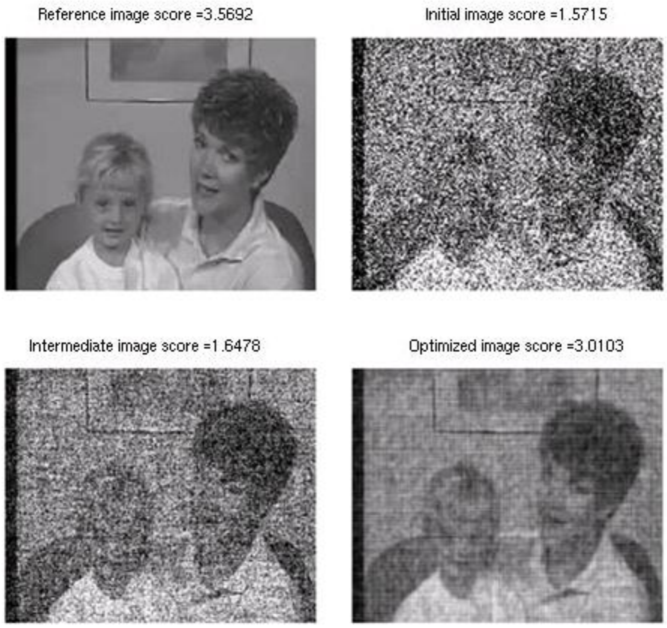

Fig. 7: VQMT-assisted image concealment Results VQMT estimated score taken as cost function

References

[1] B. Ekobo Akoa, E. Simeu et F. Lebowsky, «Using statistical analysis and artificial intelligence tools for automatic assessment of video sequences,» chez Proc. SPIE 9015, Color Imaging XIX: Displaying, Processing, Hardcopy, and Applications, San Francisco, 2014.

[2] L. Lupka, E. Simeu, H. Stratigopoulos, L. Ruffer, S. Mir et O. Tumova, «Signature analysis for MEMS pseudorandom testing using neural networks,» chez 12th IMEKO Joint Symposium on Man Science and Measurement, Annecy, 2008.

[3] H. Stratigopoulos et Y. Makris, «Error moderation in low-cost machine-learning-based analog/RF testing,» IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems, pp. 339-351, 2008.

[4] B. Ekobo Akoa, E. Simeu et F. Lebowsky, «Using Artificial Neural Network for Automatic Assessment of Video Sequences,» chez 27th IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA'13), Barcelona, 2013.

[5] F. De Simone, M. Naccari, M. Tagliasacchi, F. Dufaux, S. Tubaro et T. Ebrahimi, «Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel,» chez Proc. Int. Conf. QoMEX 2009, 2009.

[6] A. Eden, "No-Reference Image Quality Analysis for Compressed Video Sequences," IEEE transaction on broadcasting, vol. 54, no. 3, September 2008.

[7] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin et F. Battisti, «Two new full-reference quality metrics based on HVS,» chez Proceedings of the Second International Workshop on Video Processing and Quality Metrics, Scottsdale, 2006

[8] D. M. Chandler et S. S. Hemani, online supplement to 'vsnr: A visual signal-to-noise ratio for natural images based on near-threshold and suprathreshold vision, 2007.

[9] .C. Charrier, O. Lézoray et G. Lebrun, «Mesure de la qualité d'images couleur par combinaison de classifieurs,» chez Colloques sur le traitement du Signal et des Images (GRETSI), 2009.

[10] A. Chetouani, A. Beghadi, S. Chen et G. Mostafaoui, «A novem free reference image quality metric using neural network approach,» Proc. Int. Workshop Video Process. Qual. Metrics Cons. Electrn., pp. 1-4, 2010.

[11] F. De Simone, M. Tagliasacchi, S. Tubaro et T. Ebrahimi, «A H.264/AVC video database for the evaluation of quality metrics,» chez Proceedings of IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP), Dallas, 2010.

[12] F. Crete, T. Dolmiere, P. Ladret et M. Nicolas, «The Blur Effect: Perception and Estimation with a New No-Reference,» chez SPIE Electronic Imaging Symposium Conf Human Vision and Electronic Imaging, San Jose, 2007.

[13] Tektronix. [En ligne]. Available: http://www.tek.com/video-quality-monitors.Accessed 20 June 2014.

[14] L. Su, L. Zhang, W. Gao, Q. Huang et Y. Lu, «Improved Error Concealment Algorithms Based on H.264/AVC Non-nrmative Decoder,» chez Proc. Int. Conf. Multimedia Expo (ICME), 2004.