Quality Assessment Metrics vs. PSNR under Packet Loss Scenarios in MANET Wireless Networks

M.Martinez-Rach, O.López, P.Piñol, M.P.Malumbres
Departament of Physics and Computer Engineering
Miguel Hernández University (UMH)
{mmrach,otoniel,pablop,mels}@umh.es

J. Oliver, Carlos T. Calafate
Departament of Computer Engineering
Polytechnic University of Valencia (UPV)
{joliver,calafate}@disca.upv.es

ABSTRACT
It is well known that PSNR does not always rank quality of an image or video sequence in the same way that a human being. There are many other factors considered by the human visual system and the brain. So, a lot of efforts were required to find an objective video quality metric that is able to measure the quality distortion similarly to the one perceived by the destination user. We analyze the behaviour of some of the most relevant objective quality metrics when they are applied to video compressed by a H264/AVC codec at different bit-rates and with error resilience options enabled. Video data is transmitted in a wireless MANET environment and packet losses are modelled for different scenarios including variable congestion and mobility states. We take as reference the PSNR metric and try to find out if there is a more accurate metric in terms of human quality perception that could substitute PSNR in the performance evaluation of different coding proposals under packet loss scenarios.

Categories and Subject Descriptors
C.2.1 [**Network Architecture and Design**]: *Wireless communication;* G.3 [**Probability and statistics**]: Markov processes; I.4 [**Image processing and computer vision**]: *Compression (coding), Feature Measurements.*

General Terms
Algorithms, Measurement, Performance,

Keywords
Quality Assessment Metrics, Wireless Ad-hoc Networks, Markov Models, Video Compression, Error Resilience

1. INTRODUCTION
The most reliable way of assessing the quality of a video is subjective evaluation. The Mean Opinion Score (MOS), is a subjective quality metric obtained from a panel of human observers. It has been regarded for many years as the most reliable form of quality measurement technique. However, this method is too cumbersome, slow and expensive for most applications. Objective quality metrics provide video designers with means for making meaningful quality evaluations without convening viewer panels.

MV'07, September 28, 2007, Augsburg, Bavaria, Germany.

There is a consensus on a primer classification of objective quality metrics [16] attending to the availability of original non-distorted info (video reference) to measure the quality degradation of an available distorted version. *Full Reference* (FR) metrics perform the distortion measure having full access to the original image/video. *No Reference* (NR) metrics have no access to reference image/video estimating distortion only from the distorted version. And finally, *Reduced Reference* (RR) metrics work with some information about the original video to perform quality measurements. The most widely used FR objective video quality metrics are Mean Square Error (MSE) and PSNR. They are simple and quick to calculate, providing a good way to evaluate the video quality [3]. However, these metrics do not always capture the distortion perceived by the Human Visual System (HVS). In the last years, new objective image and video quality metrics have been proposed, mostly for FR/RR Quality Assessment (QA). They emulate human perception of video quality since they produce results which are very similar to those obtained from subjective methods.

We are going to evaluate different available objective quality metrics. The main goal is to find candidates that replace the classical PSNR metric to better assess the reconstructed video quality of encoded streams that were delivered through error prone networks, like MANETs (Mobile Ad-Hoc NETworks). For that purpose, we have used the H.264/AVC video coder to produce bitstreams at different bitrates and a HMM model that accurately reproduce packet loss patterns in MANET networks. Then, we perform a bitstream erasure process based on the lost patterns suggested by the HMM model at different MANET scenarios. The resulting packet streams are delivered to de H.264/AVC decoder in order to get the resulting Hypothetical Reference Circuits (HRC) that will be used to compare the QA results from the different objective quality metrics under study.

The organization of the paper is as follows: In the next section we will describe the main frameworks defined around objective QA metrics. In section 3, we describe the HMM Model for packet loss patterns in MANETS. In section 4, the models and methods used for metric comparison are explained. Section 5 analyzes the behaviour of the different metrics under the selected scenarios. And finally in section 6 some conclusions are given.

2. OBJECTIVE QUALITY METRICS

Different frameworks have been proposed in the literature that group QA metrics depending on the way they are designed. We briefly describe ideas behind the proposed frameworks.

2.1 Error Sensitivity

The Error Sensitivity framework (ESF) cover all metrics that were designed taking into account different models based on the current knowledge of the HVS. Generally, the emulation of HVS is a bottom-up approach that begins with the first retina processing steps followed with different models about the visual cortex behaviour. Also, some metrics deal with cognitive issues about the human visual processing system. Usually the HVS models first decompose the input signal into spatio-temporal subbands in both the reference and distorted signal. Then, an error normalization and weighting process is carried out to obtain the estimated degradation measure [5,6,8,15,24].

2.2 Structural Distortion/Similarity

The Structural Distortion Framework (SDF) is focused on a top-down approach, analyzing the HVS to emulate it at a higher abstraction level. Authors supporting this framework argument that the main function of the human eyes is to extract structural information from the viewing field, being the HVS highly adapted for this purpose. Therefore, a measurement of structural distortion should be a good approximation to perceived distortion [16-20].

2.3 Statistics of natural images

This framework is related with the statistical behaviour of natural images and we will refer to it as Statistics of Natural Images Framework (SNI). Here, a natural image/video is defined as those captured with high quality devices working in the visual spectrum (natural scenes). Authors supporting this framework argument that the HVS has evolved with the statistical patterns (spatial and temporal) found in the signals captured form the visual field. Also, they state that these statistical patterns of natural scenes have modulated the biological system, adapting the different processing layers to these statistics. So, the metrics defined under this framework will extract the information from a visual input signal in the form of statistical information [10,11,13,21].

3. HMM MODEL FOR PACKET LOSSES PATTERNS IN MANETS NETWORKS

The IEEE 802.11 standard for wireless LANs defines several mechanisms for reliable packet transmission in noisy wireless channels. Since all data is protected with a CRC field, it is unlikely that a corrupted packet gets to the destination, even if using an unacknowledged service, like with broadcast or multicast traffic. We can therefore assume that either the whole packet is received or nothing at all is received. In a previous work [1] we found that routing related losses can provoke quite large packet loss bursts in MANET environments.

We assume that stations belonging to the MANET are found in different routing states (e.g. route available, route discovery, re-routing, etc.). Anyway, independently of the routing state, packet losses can occur for a variety of other reasons (collisions, channel noise, queue dropping, etc.). Therefore, an outside observer cannot relate a packet loss with a certain routing state. We deal with a situation where the observation is a probabilistic function of the state, that is, only the output of the system and not the state transitions are visible to an observer. We will therefore try to solve the classification problem using a *hidden Markov model* [9].

HMMs are well known for their effectiveness in modeling bursty behavior, relatively easy configuration, quick execution times achieved and general application. So, we consider that they fit our purpose of accelerating the evaluation of QA metrics for video delivery applications on MANET scenarios, while offering similar results as with simulation or real-life testbeds.

3.1 General methodology

Relatively to the methodology proposed in our previous work [2] we start selecting a single data stream for analysis, as well as the criteria for considering a packet good or unusable by the application. We then have to map each packet sequence number with values 1 - considering the packet good - or 0 if the packet does not arrive to destination or does not meet any of the chosen criteria. This output mapping is stored in a trace file named *ST*, that will be parsed to obtain the distributions of consecutive packets arriving (CPA) and consecutive packets lost (CPL) stored respectively in trace files *C1* and *C0*. We then use the latter two to tune the proposed HMMs.

In a HMM, the number of states is not defined by the possible output events. So, in [2], the study was focused in defining a very simple 2-state model: one of the states models a currently broken path and the other models path availability where the probability for a packet to reach destination is given by function *h(s)*, being *s* is the packet size. This function models packet losses mostly due to collisions, but also due to channel noise, packet fragmentation, buffers overflow, and the type of MAC used.

However, as it was proven in the aforementioned work, the accuracy of the two-state HMM model was not satisfactory enough for certain MANET routing protocols, like DSR and other reactive routing protocols, and so a three-state model was proposed where the routing behavior was better represented, increasing the accuracy of the model. In our experiments we did not have to use more than three states, showing that the model complexity can be kept low and still provide the desired behavior.

3.2 Three-state HMM packet loss burst model

In this section we present the HMM that is able to model large packet lost bursts in mobile ad hoc networking scenarios based on reactive routing protocols (like DSR). As shown in Figure 1, states L and R represent those situations where a path towards the destination is lost and no packet can arrive successfully.

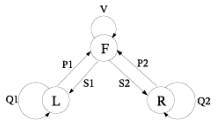

Figure 1: Three-state Markov chain for packet loss burst model in MANETs

In those situations, state L models short path breakages, mainly due to transient congestion on the path to destination and interference errors, and state R models the route discovery process which is triggered to find out a new path to the destination just after declaring the current path lost.

The state F represents the situation when packets correctly arrive to destination, with a probability defined by function $h(s)$, where s

is the packet size. In states L and R all packets are lost. Mapping state L as 0, state F as 1, and state R as 2, we obtain the following transition probability matrix:

$$A_3 = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} Q_1 & P_1 & 0 \\ S_1 & V & S_2 \\ 0 & P_2 & Q_2 \end{bmatrix} \quad (1)$$

Then, using the ns-2 network simulation tool, we define the network scenario we are going to model (number of nodes, working area, mobility and traffic patterns, network architecture model parameters, etc.) and obtain the trace file *ST* as mentioned in previous subsection. With this simulation trace we will be able to tune our model and obtain and estimation of the model parameters $v_e = (S_{1_e}, S_{2_e}, P_{1_e}, P_{2_e}, V_e)$ that will represent the desired MANET scenario. For more details about this process and the validation of the model, please refer to [2].

4. EVALUATION OF QUALITY METRICS

4.1 HMM model description and methods

We will use the three-state model presented in previous section. The MANET network scenario is composed of 50 nodes moving in an 870x870 square meters area. Node mobility is based on the random way-point model, and speed is fixed constant at 4 m/s. The routing protocol used is DSR. Every node is equipped with an IEEE 802.11g/e enabled interface, transmitting at the maximum rate of 54 Mbit/s up to a range of 250 meters. Notice that a QoS differentiated service is provided by IEEE 802.11e [4]. Concerning traffic, we have six sources of background traffic transmitting FTP/TCP traffic in the Best Effort MAC Access Category throughout the entire simulation time (runs of 300 seconds long). The foreground traffic will be composed with real traces of an H.264 video encoded (using the Foreman CIF video test sequence) at a target data rate of 1 Mbit/s. The video source is mapped to the Video MAC Access Category.

We describe two scenario classes: (a) congestion related scenarios, and (b) mobility related scenarios. The first class is composed of 6 scenarios (from M1 to M6) with 1 to 6 video sources (increasing levels of congestion are represented with this class). The second class is composed of 3 scenarios (S1 to S3) with only one video source but with increasing degrees of mobility (from 1 to 4 m/s).

We obtain losses of less than 7 consecutives packets (isolated burst), losses of several small packets (consecutive small burst) and large packet bursts (large bursts) with even more than 1000 losses. We have used the Foreman CIF seq. (300 frames at 30 fps) to build an extended sequence by repeating the original up to the desired frame number.

We have use the H264/AVC codec, properly setting the error resilience and concealment options, so that the decoder is able to reconstruct sequences even with such amount of packets lost. Codec produces one I frame every 30 P frames, with no B frames. Encoder is set to get 7 slices per frame, forcing the decoder to put each slice into a separate packet and to generate its output in RTP packet format. We force 1/3 of the macroblocks of each frame to be encode in intra mode. These error resilience parameters where proposed for MANET network scenarios in [1] as an appropriate error resilience configuration.

We get a RTP bitstream for each extended sequence. We applied them the packet loss patterns obtained through the model eliminating the lost packets from, getting a *'filtered bitstream'*. This process simulates packet losses in the MANET scenarios. The *'filtered bitstream'* will be the input to the decoder.

The decoder behaviour varies with the packet loss burst type. For isolated bursts the decoder applies error concealment to repair the frames affected by the burst. With consecutive bursts error resilience is used to repair the affected frames. Quality decreases when losses occur, increasing slowly with time by the use of the intra coded macroblocks until the next I frame completely cancels error. When the decoder faces large bursts it will stop decoding and wait until new packets arrives. This produces a shorter bitstream in the decoder than the produced by the encoder. Therefore both bitstreams are not directly comparable. A more realistic behaviour would be to freeze the last completely decoded frame until the loss bursts ends, and with the new incoming packets try to reconstruct the frame progressively So, the observer will see a frozen frame and not a jump in the sequence.

We post processed the decoded bitstream in order to achieve the same length than the encoded one by repeating the last decoded frame as many times as lost frames are produced by the burst.

4.2 Metrics comparison

Once we have the original and the decoded sequences we are able to run the Quality Assessment Metrics against them. This run will produce two set of quality: the objective quality values of the metric and the rescaled values in DMOS. We will enumerate the evaluated metrics:

Mean Structural SIMilarity index (MSSIM) [22] a FR-Image metric from the SDF. *Visual Information Fidelity (VIF)* measure [12] located in the SNI framework, a FR-Image metric. *No-Reference JPEG Quality Score* (NRJPEGQS) [24] a NR-Image metric designed specifically for JPEG compressed images. *No-Reference JPEG2000 Quality Assessment* (NRJPEG2000) [19] a NR-metric that use Natural Scene Statistics models. *Reduced-Reference Image Quality Assessment* (RRIQA) [21] RR-metric based on a Natural Image Statistic model. *PSNR-DMOSp*, the traditional PSNR in the predicted DMOS space (see [7]), that we call PSNR-DMOSp.

In [7] authors concluded that the PSNR-DMOSp metric can be taken as the *'subjective'* counterpart of the traditional PSNR in a DMOS scale. In this scale PSNR-DMOSp models correctly the effect of the saturation of quality in both ends. When the quality is quite good, higher values of PSNR are not visually perceptible; on the other end, when quality is too bad, it is difficult to decide whether a frame with lower PSNR value is worse than the former. This effect is correctly scored by all the evaluated metrics.

As explained in [7] each metric runs its code producing an objective quality value for each frame. This value is measured in its own scale. So we can not compare them directly. In order to compare the metrics we have to rescale their objective quality values to the common DMOS scale. For scale conversion we use the equation 1 of [7] with the corresponding parameters as explained in the referred paper. The goodness of each metric depends on how well the metric follows the quality values in subjective experiments. The metric which offered better adjustments to human subjective quality scores, according to the experiments in [7] was VIF and the worst was PSNR-DMOSp. In the DMOSp scale (linear scale) quality increases as the value of DMOSp decreases. So, higher quality scores have lower DMOSp

values. As pointed out in [7] noticeable differences correspond to at least 4.91 DMOSp points.

The FR metrics take as input both the original sequence and the reconstructed one for scoring quality. NR metrics score quality only with the information available on the reconstructed sequence. RR metrics do the job in two steps: (a) they first extract some relevant information from the original frame, and then (b) this information will be compared with the same type of information extracted from the reconstructed frame.

In Figure 3 we can observe that the behaviour of the metrics is quite different between them. This behaviour is what we have to analyze in order to find, if any, if there is a metric that could substitute the traditional PSNR metric for evaluate performance of codecs under packet losses scenarios.

5. Analyzing the metrics behaviour

After running the metrics over the testing sequence we obtain the objective quality values and the DMOSp values. In Figure 2 for model M2, we can see the objective quality value of the traditional PSNR in dB at three different compression levels (Low compression Qp=26, Medium compression Qp=36 and High compression Qp=44) for large packet loss bursts. We observe the evolution of quality during the burst duration.

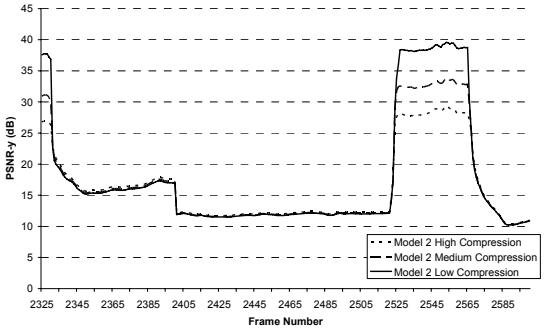

Figure 2. PSNR comparison at different compression levels for a wide burst of lost packets

Quality drops drastically with the first frame affected by the burst, and decrease even more as the differences of the frozen frame towards the original frame (where movement continues) increases. Differences with the original frame vary depending on the similitude with the reconstructed frame. Nearly at the middle of the burst an additional drop of quality can be observed. It corresponds with a scene change, that is, with the beginning of a new cycle of the foreman sequence, producing a drastically scene change that makes differences even higher. When the burst ends, quality rapidly increases with the arriving of packets belonging to a new frame that can be compared with its corresponding original one. As shown in Figure 2, at the scene change, the traditional PSNR scores the additional differences with even worse PSNR values arriving even up to 10-15 dBs where perception of quality changes at these levels is quite difficult.

What the observer really sees during this large burst is a frozen frame, with more or less quality depending of the compression level. If some quality metric takes only into account the quality of this frozen frame and not the differences with the original changing one, the effect of the burst in the scene would go unnoticed for this metric, and then the average quality of the sequence would be the same as with no packet lost bursts. For evaluating the quality of packet loss scenarios in MANET networks the chosen metric must have a drop of quality in these cases. This can only be achieved by comparing with the original changing frame during the duration of these large bursts.

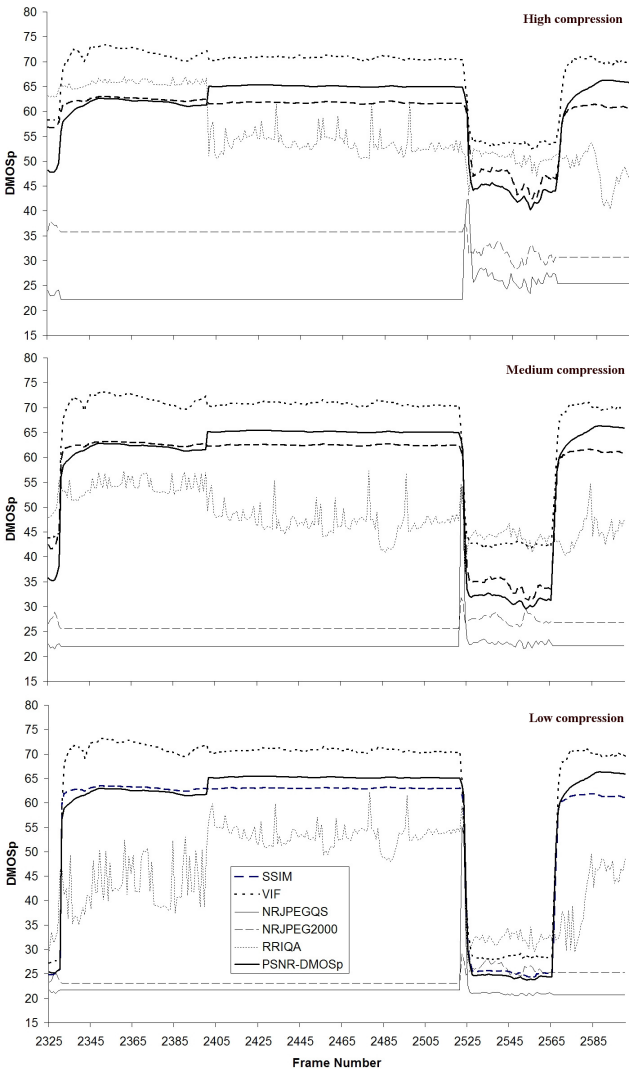

Figure 3. Metric comparison for a wide burst in DMOSp

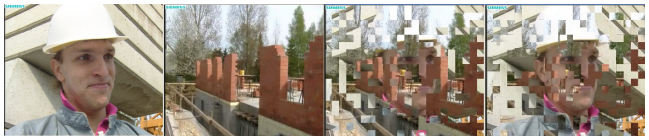

Figure 4. Packets arriving after a wide burst

Figure 3 shows the evolution of the metrics during a large burst, the same burst as in Figure 2 but in the DMOSp space. There is a panel for each compression level. We observe some interesting effects.

The NR metrics, do not detect the presence of a frozen frame as expected. Variations in quality for these metrics can be scored only when the packet loss burst affect partially to the frozen frame producing impairments on it, as in the two right pictures of Figure 4. When the large burst ends, packets belonging to perhaps P type frames begin to arrive. Then the decoder represents, only

the intra coded macroblocks belonging to these packets, over the last correctly decoded frame (the same that we have frozen). The whole frame will be progressively drawn and quality increases in the same manner. In Figure 3 this happens around frame 2521 where the NR metrics reacts to and score down quality, while the FR metrics begin to increase its quality score. When the frame is fully reconstructed then the NR metrics follows again the scored quality for the compression level like the FR metrics. The NRJPEGQS metric reacts better (higher quality differences) than the NRJPEG2000 because it is designed for detecting blockiness introduced by the discrete cosine transform.

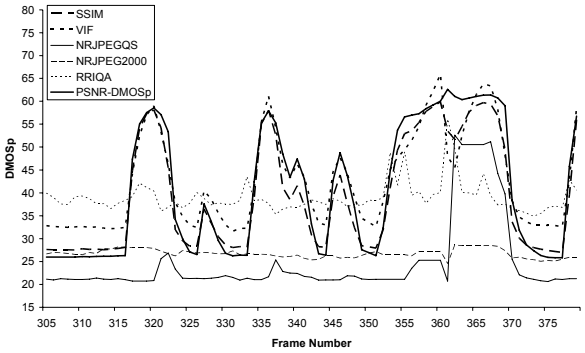

Figure 5. Metric comparison for consecutive short bursts

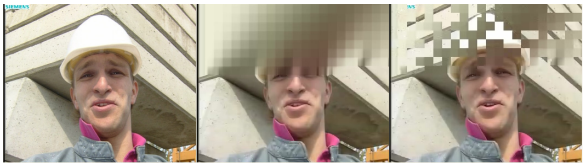

Figure 6. Decoded frames between two consecutive bursts

In Figure 4 (from left to right) the first picture corresponds to the original frame and the rest of the pictures correspond to reconstructed frames. The second picture corresponds to the last frozen frame of the burst. The third and fourth pictures correspond to the firs frames after the large burst where can be observe how the decoder represents progressively the intra macroblocks as they arrive in the incoming packets. In these frames occurs the mentioned reaction of the NR metrics.

Looking again to Figure 3 we see that the DMOSp value before the burst and after it, increases as the compression does (different panels) for all metrics (more compression lower quality), except for metric NRJPEGQS that can not detect as good as the rest of the metrics (included the other NR metric) the blur effect introduced by the encoder. The RRIQA metric shows high variability in its scores between consecutive frames. These variations are higher inside the large burst and as the compression decreases. Outside the burst the variations are closer to the noticeable differences in the DMOSp scale, but in any case its behaviour differs too much from the homogeneous scoring of the FR metrics. Besides, the RRIQA do not recover the quality scored before the burst, except for the lower compression scenario.

Another interesting behaviour of metrics occurs in periods between large bursts, that is, where no frames are lost and where no packets are lost. There, differences in DMOSp values between metrics increases as compression do.

MSSIM, VIF and PSNR-DMOSp show a similar behaviour being the quality scores almost parallel between them, except in some punctual situations. MSSIM and PSNR-DMOSp have closer quality scores between them than VIF, which has a better adjustment to subjective scores, and therefore could be taken as reference between the FR metrics. The drawback as mentioned in [7] is its high computational cost in comparison with the other two FR metrics, and of course with the PSNR-DMOSp one.

Another interesting effect present in Figure 3 occurs at the point where the scene changes and the foreman sequence begin again. There, a change in the quality is only scored clearly by RRIQA (not in all compression levels in the same direction) and to a lesser extend by PSNR-DMOSp. The rest of the FR metrics seems to have achieved its saturation 'bad quality' level at this point.

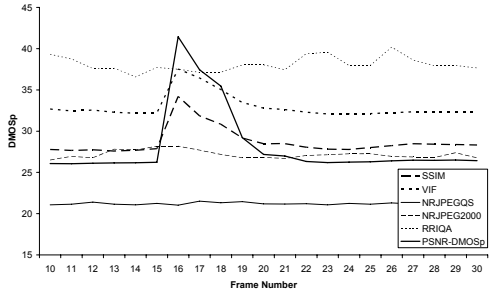

Figure 7. Metric comparison for consecutive short bursts

Figure 8. Packet lost affecting only one frame

Figure 5 shows the effect of the consecutive short bursts that causes the lost of only few frames or part of them. Here we have use model M6 and a quantization parameter of Qp=29. We find the same effects mentioned for Figure 3. Additionally we notice a new effect produced around frame numbers 362 and 363. Frames beside these two frames are frozen. The first picture of Figure 6 (from left to right) represents the original frame, the other two frames corresponds to frames 362 and 363. These two frames follows in the lower half of the frame correctly the original frame and in the upper half only intra macroblocks are rendered. So for these two frames we have partially (for the lower half) perfect correspondence with the original one, and therefore quality must increase at least in some extent. This is only reflected by the VIF and MSSIM metrics, even PSNR-DMOSp is not able to catch this effect because is computed with the information of the whole frame. After frame 363 quality decreases again because the next frame is frozen. So, VIF and MSSIM counts two burst where PSNR-DMOSp counts only a larger one.

Figure 7 shows an isolated burst, (model M3, Qp=29). The blur impairment introduced by the encoder is perceived by all metrics except by the NRJPEGQS. The error concealment mechanisms need up to 6 frames to achieve the quality score before the burst. Figure 8 shows (from left to right) the original frame and two subsequent pictures where the effect of the lost packets is quite good concealed.

Figure 9 presents only the quality variation produced by the packet loss without the variation introduced by compression. Only during the burst duration we get variation in quality. RR and NR metrics scores an increment of quality in some regions inside the

duration of the burst. In general it is expected that quality will be scored worst in some extent when packets are lost. We also see how MSSIM and PSNR-DMOSp react to the scene change with a higher extent than in Figure 3 because the DMOSp value has not reached the saturation value jet. However the VIF metric scores uniformly during the whole burst.

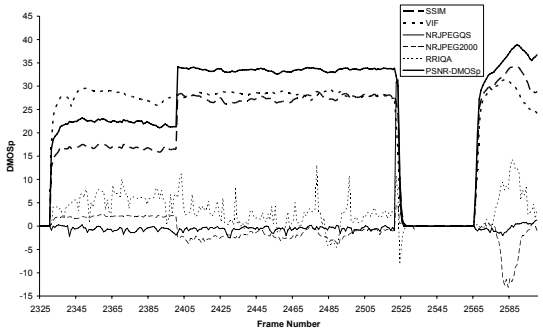

Figure 9. Metric comparison for a wide burst in DMOSp with no compression effect

6. CONCLUSIONS

We have analyzed the behaviour of different QA metrics when measuring reconstructed video quality sequences encoded and delivered through error prone wireless networks, like MANETs. The analysis results are the following ones: (1) NR metrics are not able to proper detect and measure the sharp quality loss due to the lost of several consecutive frames.(2) The RR metric has a non deterministic behaviour under the presence of packet losses, having difficulties to identify and measure this effect when the video is encoded with moderate to high compression rates. (3) The other metrics, SSIM, PSNR-DMOSp and VIF show a similar behaviour in all cases, as expected. Although they exhibit slight differences, we propose the MSSIM metric as a trade-off between a high quality measurement process (near human visual perception) and its computational cost.

7. REFERENCES

[1] Carlos T. Calafate, Manuel P. Malumbres, and Pietro Manzoni. Performance of H.264 compressed video streams over 802.11b based MANETs. In International Conference on Distributed Computing Systems Workshops, March 2004

[2] Carlos T. Calafate, Pietro Manzoni, Manuel P. Malumbres,, " Speeding up the evaluation of multimedia streaming applications in MANETs using HMMs ", Int. Symp. on Modeling, Analysis and Simulation of Wireless and Mobile Systems, 1-58113-953-5,pp. 315-322, 2004, Venice, Italy

[3] B.Girod, What's wrong with mean-squared error. Digital Images and Human Vision, 1993

[4] IEEE 802.11 WG. 802.11e Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements, 2005.

[5] C.Lambrecht and O.Verscheure, Perceptual Quality Measure using a Spatio-Temporal Model of the Human Visual System. In Proc. of the SPIE, 1996

[6] J.Malo, A.M. Pons, J.M. Artigas, Subjective image fidelity metric based on bit allocation of the human visual system in the DCT domain. Image and Vision Computing, 1997

[7] M. Martinez-Rach, O. Lopez, P. Pinol, J. Oliver, M.P. Malumbres, "A Study of Objective Quality Assessment Metrics for Video Codec Design and Evaluation", ism, pp. 517-524, Eighth IEEE International Symposium on Multimedia (ISM'06), 2006

[8] M.Masry, S.S.Hemami, Y.Sermadevi, A Scalable Wavelet-Based Video Distortion Metric and Applications. IEEE Trans. On Circuits and Systems for Video Techn. 2006

[9] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2):257–286, 1989.

[10] H.R.Sheikh, A.C. Bovik, L.Cormack, No-reference quality assessment using natural scene statistics: jpeg2000. IEEE Trans. Image Process. 2005

[11] H.R.Sheikh, A.C.Bovik, G.Veciana, An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Trans. Image processing, 2005

[12] H.R.Sheikh, A.C. Bovik, Image information and visual quality. IEEE Transactions on Image Processing, 2006

[13] E.P.Simoncelli, Modeling the joint statistics of images in the wavelet domain. Proc. SPIE 44th Annual Meeting, 1999

[14] Z.Wang, H.R.Sheikh, A.C.Bovik Objective Video Quality Assessment. Chap.41 The Handbook of Video Databases: Design and Applications, 2003

[15] Z.Wang, A.C.Bovik, L.Lu and J.Kouloheris, Foveated wavelet image quality index. SPIE's 46th Annual Meeting, Proc. SPIE, Application of digital image proc. XXIV, 2001

[16] Z.Wang, A.C.Bovik, H.R. Sheikh, E.P. Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Trans. Image Proc. vol. 13 no. 4, 2004

[17] Z.Wang, E.P. Simoncelli, An adaptative linear system framework for image distortion analysis. Proc. IEEE Inter. Conf. on Image Processing 2005

[18] Z.Wang, E.P.Simoncelli, Traslation insensitive image similarity in complex wavelet domain. Proc. IEEE Inter. Conf. Acoustic, Speech & Signal Processing, 2005

[19] Z.Wang, L.Lu, A.C.Bovik. Video quality assessment using structural distortion measurement. IEEE Inter. Conf. on Image Processing, 2002

[20] Z.Wang, A.C.Bovik, B.L.Evans, Blind Measurement of Blocking Artifacts in Images. Int. Conf. on Image Proc. 2000

[21] Z.Wang, E.P. Simoncelli, Reduced-reference image quality assessment using a wavelet-domain natural image statistics model. Human Vision and Elect. Imag. X, Proc. SPIE, 2005

[22] Z.Wang, A.C.Bovik, H.R.Sheikh, E.P. Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Trans. Image Processing, 2004

[23] A.B.Watson, J.Hu, and J.F.McGowan, Digital video quality metric based on human vision. J. Electronic Imaging, 2001

[24] Z.Wang, H.R.Sheikh, A.C.Bovik, No-reference perceptual quality assessment of jpeg compressed images. Int. Conf. Image Processing, 2002