

THE POLAR EDGE COHERENCE: A QUASI BLIND METRIC FOR VIDEO QUALITY ASSESSMENT

V. Baroncini(*), L. Capodiferro(*), E. D. Di Claudio(**), G. Jacovitti(**),

(*)Fondazione "Ugo Bordon"i

Via Baldassarre Castiglione 59, I-00142, Rome, Italy

web: <http://www.fub.it>

(**)INFOCOM Dpt., University of Rome "La Sapienza"

Via Eudossiana 18, I-00184, Rome, Italy

web: <http://infocom.uniroma1.it>

ABSTRACT

An extremely reduced reference video quality assessment method for video sequences is presented. It is based on the comparison of indices independently calculated for the examined video sequence and for the reference one. The employed metric is based on a cumulative measure of both frame by frame and interframe local edge coherence measures made with pairs of circular harmonic functions.

1. INTRODUCTION

As video content communications are becoming more pervasive and networks diversify, image processing for video quality assessment is becoming increasingly important. In fact, considerable gain of resource allocation should be achieved in a variety of applications, from HDTV to cell phone videos, by the use of objective measurements well correlated with perceptual judgments.

In the video production area, quality can be easily verified by direct comparison of image sequences under examination to reference full quality images. The degradation is measured by metrics of image differences considered as undesired errors. They range from crude MSE, SNR, PSNR metrics to more sophisticated metrics tuned to known features of the Human Visual System (HVS), or correlated with it, such as SSIM or VIF.

These techniques, usually referred in the specialized literature to as FR (full reference) assessment methods, are unfeasible for video distribution, where reference images are not available. For this reason, RR (reduced reference) methods are employed. They operate with partial information about original image structures or embedded markers for estimating the channel effects. Some other methods, are based on the prior knowledge of the distortion affecting the judged image, such as blocking, blur or ringing. Some examples are the ones presented in [1], [2], [3].

The cited methods can be still employed for video quality assessment by averaging the quality of single frames. Recently in the context of FR methods, specific metrics for measuring the quality impairment generated by imperfect motion reproduction have been proposed in [4]. These measurements compare the differences among patterns observed in the 3D space-time domain of the reference

sequence and of the observed one, using the well known SSIM approach.

In this contribution, an extreme RR quality assessment method recently introduced for limiting the side transmission to the order of one number per image [4] is extended to video sequences, encompassing both interframe and intraframe quality measurements into the unique concept of *polar edge coherence*.

The method can be defined as "quasi blind" for two reasons. From one hand, it implies simple comparison of two indices independently calculated from the observed and the reference image, with a negligible transmission overhead. On the other hand, it is not well specialized for particular distortions, but is sensitive to different image impairments.

The method is based on the local analysis with circular harmonic functions extracted from the so-called Laguerre-Gauss (GL) orthonormal family, applied to single frames and to space time slices. The polar edge coherence is calculated from a pair of GL expansion coefficients. It is recognized that the polar coherence is a rather general metric which tends to be disrupted by the factors causing loss of perceived quality, and in particular by image compression techniques [4]. Before entering into technical issues about video, the notion of polar edge coherence and related metrics is briefly introduced.

2. THE POLAR EDGE COHERENCE

Making reference to the polar coordinates

$r = \sqrt{x^2 + y^2}$, $\varphi = \arctg\left(\frac{y}{x}\right)$ centred around the generic

analysis point, the image $I(r, \gamma)$ can be expanded as follows:

$$I(r, \gamma) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \gamma_{n,k} \mathcal{L}_{n,k}(r, \gamma; \sigma)$$

where $\mathcal{L}_{n,k}(r, \gamma; \sigma)$ are complex orthonormal and polar separable functions defined as :

$$\mathcal{L}_{n,k}(r, \gamma; \sigma) = g_{n,k}(r) \cdot e^{jn\gamma}$$

whose radial profile is:

$$g(r) = \frac{1}{\sqrt{k!(|n|+k)!}} (-1)^k \left(\frac{r}{\sigma}\right)^{|n|} \times L_k^{(|n|)} \left[\left(\frac{r}{\sigma}\right)^2\right] \frac{1}{\sigma\sqrt{\pi}} e^{-\frac{1}{2}\left(\frac{r}{\sigma}\right)^2}$$

and $L_k^{(n)}(\zeta)$ is the generalized Laguerre polynomials defined by the Rodriguez formula

$$L_k^{(n)}(\zeta) = (-1)^{n+k} k! \left[\sum_{p=0}^k (-1)^p \binom{n+k}{k-p} \frac{\zeta^p}{p!} \right]$$

These functions, called Laguerre Gauss Circular Harmonic (LG-CH) functions, are indexed by the integers n (referred to as *angular order*) and k (referred to as *radial order*) and possess many remarkable mathematical properties. Among others, they are self-steerable, i.e., they rotate by the angle α after multiplication by the factor $e^{jn\alpha}$, and keep their shape in invariant under Fourier transformation. They are suitable for multiscale and multicomponent image analysis [5] as an alternative to other bases such as the steerable pyramid, separable orthogonal wavelets, the Gabor family, etc. Because of the orthonormality of the LG-CH functions, the coefficients $y_{n,k}$ of the expansion are:

$$y_{n,k}(\sigma) = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} I(x, y) \mathcal{L}_{n,k}(x, y; \sigma) dx dy$$

Writing them in polar coordinates

$$y_{n,k}(\sigma) = \int_{\gamma=-\pi}^{\pi} e^{jn\gamma} \left\{ \int_{r=0}^{\infty} g_{n,k}(r) I(r, \gamma) r dr \right\} d\gamma$$

reveals that they are Fourier coefficients of a radial tomographic section of the image (which is periodic versus the azimuth γ).

These coefficients can be calculated for any image point of coordinates $x = n_1$ and $y = n_2$ by the convolution

$$y_{n,k}(n_1, n_2, \sigma) = I(x, y) * \mathcal{L}_{n,k}(-x, -y; \sigma)$$

Finite sets of LG functions (each referred to in the sequel as LG *plexus*, LGP) have been employed for different purposes. In this contribution the (1,3) LG plexus is considered, which contains the following LG-CH functions

$$\mathcal{L}_{1,0}(r, \gamma; \sigma) = -\frac{1}{\sigma\sqrt{\pi}} e^{-\frac{r^2}{2\sigma^2}} \sqrt{\frac{r^2}{\sigma^2}} \cdot e^{j\gamma}$$

$$\mathcal{L}_{3,0}(r, \gamma; \sigma) = -\frac{1}{\sqrt{6}} \frac{1}{\sigma\sqrt{\pi}} e^{-\frac{r^2}{2\sigma^2}} \left[\sqrt{\frac{r^2}{\sigma^2}} \right]^3 \cdot e^{j3\gamma}$$

whose shapes are shown in Fig. 1.

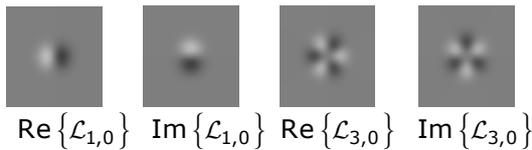


Figure 1. The (1,3) LGP (values in gray levels)

Pairs of real functions have been already employed for estimating the entity of image blur [8], but the role of the (complex) LGP is more involved. In particular, it is shown that for *ideal* edge patterns, i.e. abrupt unitary steps passing through (n_1, n_2)

$$\frac{|y_{3,0}(n_1, n_2, \sigma)|}{|y_{1,0}(n_1, n_2, \sigma)|} = \frac{1}{2}$$

$$\arg[y_{1,0}(n_1, n_2, \sigma)] = 3 \arg[y_{3,0}(n_1, n_2, \sigma)] + \pi$$

This relationship is analogous to the one existing among the first and third harmonic of the Fourier expansion of a periodic square waveform (see Fig. 2). For non centered edges, and for other patterns the magnitude ratio is generally smaller on the average, and the phase relationship does not longer hold.

Based on the above relationship, the *polar edge coherence* (PEC) map is defined as :

$$PEC(n_1, n_2, \sigma) = -\frac{|y_{3,0}(n_1, n_2, \sigma)|}{|y_{1,0}(n_1, n_2, \sigma)|}$$

$$\cos\left\{\arg[y_{3,0}(n_1, n_2, \sigma)] - 3 \arg[y_{1,0}(n_1, n_2, \sigma)]\right\}$$

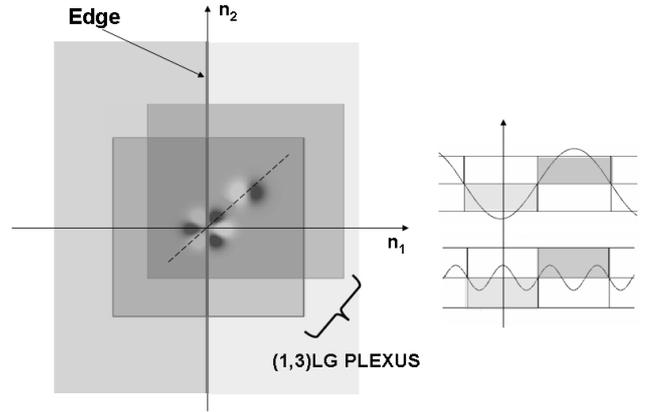


Figure 2. Analogy between the local expansion of an edge into circular harmonic components and the Fourier expansion of a square, periodic wave.

The present approach to quality assessment is inspired to the fact that on the retinal surface natural images are characterized by perfect sharp edges under ideal focusing and in absence of noise, and any edge impairment is perceived as an annoying disturbance, apart notable exceptions. Natural factors impairing edge patterns are optical distortion and neuronal noise, while reproduction technologies originate different kinds of edge distortion. From the above Fourier analogy, it is clear that any edge degradation is revealed by misalignments or attenuation of higher order harmonics with respect to the fundamental harmonic, just as happens for transitions of periodic square waves. It is concluded that good natural images are characterized by high polar edge

coherence, whereas edge disruption cause coherence losses. These arguments support *the conjecture that the human judgment about image quality is consistently correlated with polar coherence losses*

Limiting the analysis to the third harmonic, this conjecture leads to the following definition of a cumulative edge polar coherence over a $N_1 \times N_2$ image observed image, indicated for short as Edge Coherence (ECO) :

$$ECO(\sigma) = - \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} w(n_1, n_2) \frac{|y_{3,0}(n_1, n_2, \sigma)|}{|y_{1,0}(n_1, n_2, \sigma)|} \times \cos\{\arg[y_{3,0}(n_1, n_2, \sigma)] - 3 \arg[y_{1,0}(n_1, n_2, \sigma)]\}$$

where the $w(n_1, n_2)$ is a feature weighting factor. The choice

$$w(n_1, n_2) = |y_{1,0}(n_1, n_2, \sigma)|^2$$

which was made in [4], corresponds to the so-called RTAEC version of RECO.

The ECO metric is an absolute measure, so that it might be used in principle for No Reference quality estimation, provided that a quality unit is defined. However, ECO suffers from the fact that it depends on the image content. For compensating it, the Relative Edge COherence (RECO) index between the images $I(n_1, n_2)$ and a reference image $\tilde{I}(n_1, n_2)$ of the same scene is defined as the ratio:

$$RECO(\sigma) = \frac{ECO(\sigma) + C}{\tilde{ECO}(\sigma) + C}$$

where $\tilde{ECO}(\sigma)$ is the cumulative ECO metric calculated for $\tilde{I}(n_1, n_2)$, and C is a regularizing (small) constant.

The most relevant formal properties of the RECO index are:

2.1 Reciprocity:

$$RECO(\tilde{I} / I, \sigma) = \frac{1}{RECO(I / \tilde{I}, \sigma)}$$

In other words, the RECO does not measure quality impairments, as other methods do. It actually measures the relative quality level, and by consequence also image improvements.

2.2 Multiplicative chaining:

$$RECO(I / \hat{I}, \sigma) = RECO(I / \tilde{I}, \sigma) \cdot RECO(\tilde{I} / \hat{I}, \sigma)$$

This means that the total RECO for a cascade of processing stages is just the *product of partial RECOs*.

2.3 Rotational invariance:

$$RECO(I_R / \tilde{I}, \sigma) = RECO(I / \tilde{I}, \sigma)$$

where I_R is a rotated version of I .

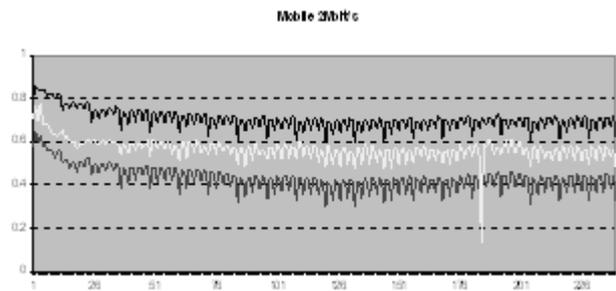
This property comes from the rotational invariance of the PEC, but is valid only if I_R and I share exactly the same content.

3. VIDEO QUALITY TRACKING WITH RECO

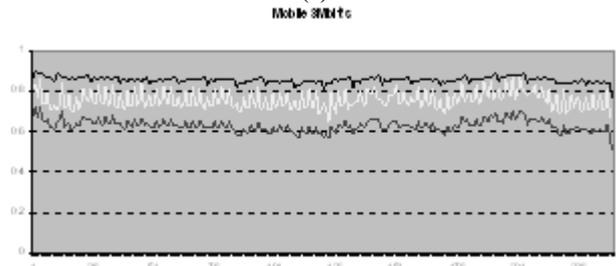
The RECO index does not require full or partial comparison of the observed and of the reference image, other than the cumulative ECO, i.e., a single number.

Therefore, the RECO can be employed for quality assessment of video sequences, inserting the ECO number in the header of each frame or of groups of frames, thus allowing for continuous tracking of the video quality in presence of channel fading, constant bitrate transmission, dynamic multiplexing, etc.

In Fig. 3 the frame by frame track of the RECO index is compared to the tracks of the *full reference* indices SSIM and SCOR on the sequence "Mobile" (see [9] for further examples). The RECO index seems fairly well correlated with the full reference measurements (see also[4]).

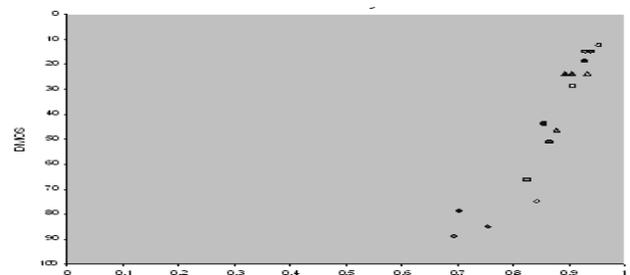


(a)



(b)

Figure 3 (a)(b). Quality tracking of a MPEG video sequence with the quasi-blind metric compared to the tracking made with full reference metrics at two different bitrates. The upper track is the SSIM index, the middle track (bright) is the RECO index, and the lower track is the SCOR index.



(a)

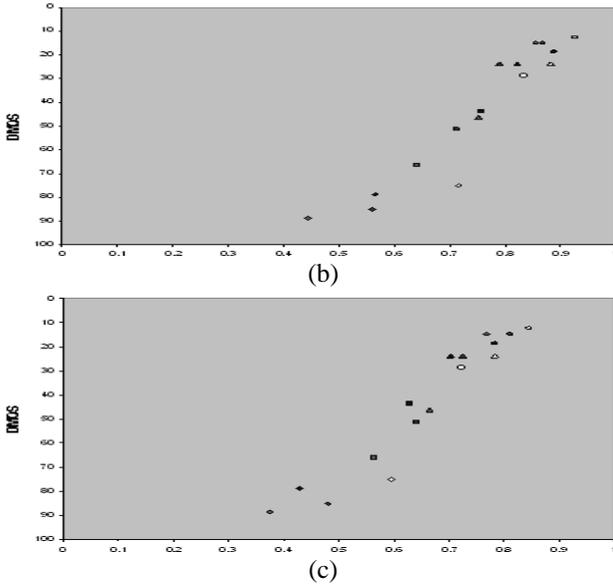


Figure 4(a)(b)(c). Scatterplots of the objective measurements versus the DMOS: (a) SSIM, (b) RECO, (c) SCOR.

In the following table the Spearman Rank Correlation Coefficient (SROCC) between subjective and objective scores for this video dataset is reported. It shows that the quality of the sequence ranking offered by the single-term RECO index is comparable to that of multi-term FR indices, such as SSIM and SCOR.

	RECO	SSIM	SCOR
MOS	0.9461	0.9557	0.9564

Table I. SROCC of quality estimators.

4. THE MRECO METRIC

Still, the frame by frame quality measurements do not fully account for the quality of motion. To compensate for this lack, the concept of edge coherence is extended from the single frames to space-time slices of a video sequence.

To this purpose, let us consider an edge passing through the generic point (n_1, n_2) of the n_3 -th frame of a video sequence, individuated by the space time coordinate (n_1, n_2, n_3) and called *pivot* frame. In a static scene, the edge in the space time is a cylindrical surface aligned with the time axis. In a dynamic scene the edge describes a complicated surface that can be locally approximated by its tangent plane. High quality video sequences are characterized by accurate reproduction of this tangent plane, corresponding to smooth and pleasant motion perception. On the other hand, high compression ratios, sub-sampling, format conversions etc. may cause inaccurate motion compensation, artifacts, jerkiness, etc., resulting in low quality motion perception. The edge coherence concept can be still applied to measure these impairments.

To this purpose, looking at the generic edge point passing through (n_1, n_2, n_3) , the angle orientation α orthogonal to

the edge in the spatial coordinates is first estimated. For the sake of simplicity, but without loss of generality, in Fig. 4 a moving edge is oriented along the n_2 axis, so that $\alpha = 0$.

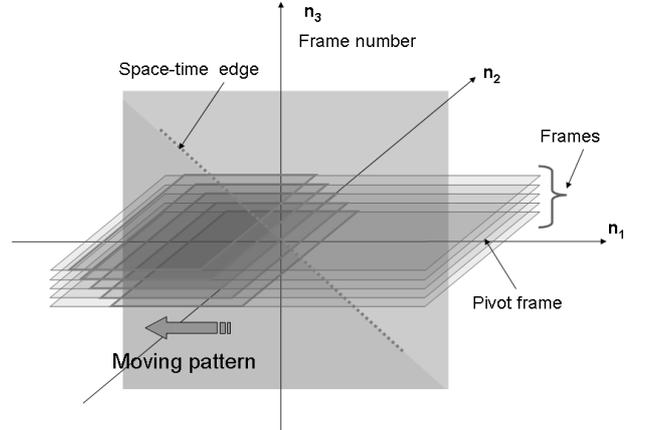


Figure 5. Space-time edge generated by the local motion of a pattern.

The estimate of the angle α is directly obtained by the phase of $\gamma_{1,0}(n_1, n_2, \sigma)$. Let us now consider the set of the frames adjacent to the pivot frame:

$$I(n_1, n_2, m - n_3)$$

and the space-time slice orthogonal to the edge orientation:

$$I_\alpha(s, t),$$

characterized by the transverse coordinate s and the temporal coordinate t (in Fig. 4 s runs along the n_1 axis and t along the n_3 axis).

An ideal moving edge yields an ideal edge in the (s, t) slice, whereas a perturbed motion does correspond to a corrupted edge. This corruption can be again measured through the polar edge coherence based on the LGP defined by the functions:

$$\mathcal{L}_{1,0}(\rho, \psi; \sigma) \quad \mathcal{L}_{3,0}(\rho, \psi; \sigma)$$

where $\rho = \sqrt{s^2 + t^2}$ and $\varphi = \arctg\left(\frac{t}{s}\right)$ are the local polar coordinates. The PEC is now obtained as

$$\gamma_{1,0}^{(\alpha)}(n_1, n_2, n_3, \sigma) = \int_{\psi=0}^{2\pi} \int_{\rho=0}^{\infty} I_\alpha(\rho, \psi) \mathcal{L}_{1,0}(\rho, \psi; \sigma) \rho d\rho d\varphi$$

$$\gamma_{3,0}^{(\alpha)}(n_1, n_2, n_3, \sigma) = \int_{\psi=0}^{2\pi} \int_{\rho=0}^{\infty} I_\alpha(\rho, \psi) \mathcal{L}_{3,0}(\rho, \psi; \sigma) \rho d\rho d\varphi$$

In practice, these quantities are approximated with sums applied to space-time-samples. The scale factor σ must be larger than the one employed for the intraframe edge coherence, in order to involve several frames contiguous to the pivot frame.

Based on these quantities, the motion edge coherence (MRECO) index is calculated over the considered space-time volume as:

$$MRECO(\sigma) = - \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \sum_{n_3=1}^{N_3} w(n_1, n_2) \frac{|y_{3,0}^{(\alpha)}(n_1, n_2, n_3, \sigma)|}{|y_{1,0}^{(\alpha)}(n_1, n_2, n_3, \sigma)|} \times \cos \left\{ \arg \left[y_{3,0}^{(\alpha)}(n_1, n_2, n_3, \sigma) \right] - 3 \arg \left[y_{1,0}^{(\alpha)}(n_1, n_2, n_3, \sigma) \right] \right\}$$

Preliminary experiments were conducted both synthetic and real video sequences to test its effectiveness. To give an idea of how the motion quality index MRECO contributes to the overall objective quality rating, in Fig. 6 the RECO and MRECO indices for a suite of MPEG sequences at different compression ratio, from near perfect quality to critical quality, is shown. It refers to a group of nine frames around a common pivot frame.

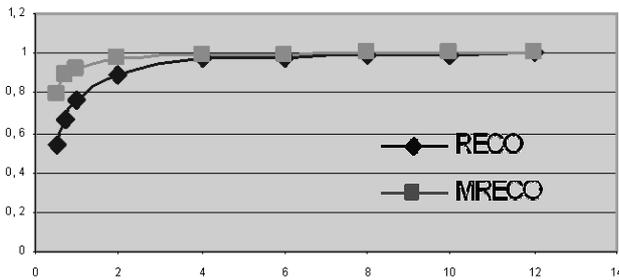


Figure 6. RECO and MRECO scores for the MPEG-2 coded "Coastguard" sequence vs. bitrate in Mb/s.

MRECO indices do have their individual significance, along with color quality indices. Merging these indices into one video quality will be the matter of further statistical analysis on the basis of an extensive number of subjective and objective measurements.

5. CONCLUSION

The polar edge coherence metric grades the quality of images on the basis of self-referenced edge integrity measures not tuned to specific distortions. As such, it is in principle a blind method, even if it requires overall normalization to compensate for the bias due to different image contents. Experiments indicated that it provides quality tracking of compressed sequences well correlated with the estimates extracted from full reference methods, still using only one reference number per frame. This outstanding feature makes the RECO and MRECO metrics very interesting candidates for "quasi blind" continuous monitoring of video quality in multimedia networks.

6. ACKNOWLEDGEMENTS

This work was partially supported by the Italian Ministry for University and Research (MIUR).

REFERENCES

- [1] Z. Wang, E. Simoncelli, "Reduced-Reference Image Quality Assessment Using A Wavelet-Domain natural Image Statistic Model", *Human Vision and Electronic Imaging X, Proc. SPIE 2005*, vol. 5666.
- [2] M. Carenc, P. Le Callet, D. Barba, "Visual Features for Image Quality Assessment with Reduced Reference," *Proc. of the IEEE ICIP 2005*, vol. 1, pp. 421-424.
- [3] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, "Perceptual blur and ringing metrics: application to JPEG2000". *Signal Processing: Image Communication* vol. 19, 2004, pp. 163-172.
- [4] L. Capodiferro, E. Di Claudio, G. Jacovitti, "Short Reference Image Quality Rating Based on Angular Edge Coherence", *Proceedings of 14th European Signal Processing Conference, EUSIPCO 2006, Florence, Italy*.
- [5] G. Jacovitti, A. Neri, "Multiresolution circular harmonic decomposition," *IEEE Trans. on Signal Processing*, vol. 48, no. 11, 2000, pp. 3242-3247.
- [6] Z. Wang, E. Simoncelli, "Local Phase Coherence and the perception of Blur," *Advances in Neural Information processing Systems*, S. Thrun, L. Saul, and B. Scholkopf eds., vol. 16, May 2004.
- [7] Z. Wang, A. C. Bovik, H. R. Sheikh & E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", *IEEE Trans. Image Processing*, vol. 13, no. 4, 2004, pp. 600-612.
- [8] V. Baroncini, L. Capodiferro, E. D. Di Claudio, G. Jacovitti, P. Sità, M. Visca, "Diagnostic quality monitoring of video sequences based on multiple structural analysis", *Proc. of the Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics, WPQM2009, Scottsdale, AZ, USA, Jan. 14-16, 2009*
- [9] V. Baroncini, L. Capodiferro, E. D. Di Claudio, G. Jacovitti, F. Mangiatordi, G. Ridolfi, "Quasi-blind on line video quality tracking based on polar edge coherence", *Proc. of the Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics, WPQM2009, Scottsdale, AZ, USA, Jan. 14-16, 2009*.