The Correlation Dimension: A Video Quality Measure

Bogdan Budescu, Alexandru Căliman, and Mihai Ivanovici

MIV Imaging Venture Laboratory, Department of Electronics and Computers, Faculty of Electrical Engineering and Computer Science, Transilvania University, Str. Politehnicii nr. 1, 500019 Brasov, Romania
{bbudescu,alexandru.caliman,mihai.ivanovici}@gmail.com
http://www.miv.ro

Abstract. Correlation dimension is a measure of the multidimensional complexity of an object. Stemming from the area of chaos theory and having several applications involving the study of the convergence and the recurring patterns of random signals, it has been proven to be a possible way to assess video quality. Based on its meaning in the multidimensional space of color fractals, it can be used, in the context of a fractal's intrinsic similarity to natural shapes and colours, to quantify the aesthetic and harmonic properties of an image. Our approach in the assessment of the perceived quality of a video stream is based on the analysis of the fractal dimension of video signals expressed in the CIE L*a*b* color space. This colour space has a strong resemblance to the human visual perception system, thus making its ΔE_{2000} norm relevant for the measurement of the perceptual difference between colours, and hence useful for image quality assessment. The fractal dimension is computed through the correlation dimension definition. In this paper we expose the experimental results obtained in a simulation of a real-life scenario: the streaming of a video of a football game over a busy network.

Keywords: Correlation dimension, colour fractal dimension, video quality.

1 Introduction

Human perceived quality of visual data is a complex metric influenced by subjective parameters aside from many objective aspects, like spatial frequency, topology and colour. In order to self-adjust the transmission parameters of digital video streams, an automatic assessment mechanism of the perceptual quality of images is required. In order to quantify this metric, there have been several approaches involving statistical observation of the perceived quality, but for the definition of a valid scale it was necessary to minimize the influence of the subjective factors in the study. The techniques that emerged were burdened, for instance, with the need to provide the same environment for every sampling session (see the ITU-T recommendations [1, 2, 3, 4] on video quality measurements regarding viewing distance and room lighting).

L. Atzori, J. Delgado, and D. Giusto (Eds.): MOBIMEDIA 2011, LNICST 79, pp. 55–64, 2012.

However, objective metrics have also been developed. They may be grouped into three categories, depending on the set of data available at the time of assessment. The full reference results of traditional signal processing methods, like RMSE [5] and PSNR, along with the classical Quality of Service metrics have the disadvantage that they are not highly correlated with the magnitude of distortion perceived by the human visual system. Full reference techniques that also take the specific characteristic of the human perception into consideration have also been defined [6] [7] leading to Quality of Experience.

The second category from the reference data availability point of view is represented by reduced reference metrics. These are based on the comparison of a set of synthesized characteristics or features of both the reference and distorted images, thus reducing the information necessary for degradation quantification. Carnec [8] employs a model of the human visual system (HVS) in order to detect points of interest from which features are extracted. A correspondence coefficient based on structural and color information (expressed in the Krauskopf perceptual color space) is then computed for the feature from both the distorted and the reference images. Cheng [9] proposes a reduced-reference metric based on natural image statistical prior computed on the gradient.

There are significantly fewer representatives of the third class of objective metrics, the ones that do not use any reference signal in the assessment process. Lu et al. [10] introduce a no-reference technique based on image structural information gathered after multiscale and multidirectional spectral decomposition.

We shall focus in the following on reduced-reference metrics and propose the usage of fractal measures for the assessment of video quality. The study of fractals have proven that many of the shapes and colours encountered in nature respect fractal-like patterns, providing this way a formal model for the perception of natural, aesthetic and harmonic visual data. We propose to take advantage of this model and to use its formalized metrics to quantify the perceptual quality of images.

The usage of the fractal dimension as a metric for video quality was proved in [11], the argumentation relies on the relationship between the complexity and the power spectrum of a signal. The degradation that affects the video signal is a mixture of several impairments, like blockiness and the sudden occurrence of new colours. The modifications of the image content reflect both in the colour histograms—a larger spread of the histogram due to the presence of new colours—and the spectral representation of the luminance and chrominance (new higher frequencies due to blockiness). The complexity of a continuous random signal can be defined based on its power density spectrum: for a random fractal signal $v(t)$, the power density function varies upon a power law in $\frac{1}{f^\beta}$, thus the Fourier transform $V(f,T)$ computed on a time interval T of $v(t)$ allows to express the spectral density function $S_V(f)$ as:

$$S_V(f) \propto T|V(f,T)|^2 \text{ as } T \to \infty \tag{1}$$

Therefore the intimate relationship between the power law and the fractal dimension [12] allows us to use the fractal dimension as a measure of signal complexity and, eventually of signal quality - based on the hypothesis that the degraded signal exhibits an increase of complexity.

2 Correlation Dimension Estimation

There are two important fractal metrics that are relevant in image quality assessment: fractal dimension and fractal lacunarity [11], and this article focuses on a new way to estimate the fractal dimension of color images. The fractal dimension indicates the complexity of a fractal set, by computing the fraction of its bounding area that it covers. Given the computational complexity of the reference theoretical Hausdorff dimension computation algorithm, implementations usually use more simple algorithms, like the box-counting algorithm, that provide the expected result (with notable exceptions, which are to be carefully taken into consideration) [13] [14] [15] [16].

First formalized in [17] and then applied in fractal theory in [18], correlation dimension is a measure of how much space is occupied by a set of random points. Although this measure has been defined in the context of chaos theory, its intimate connection to fractal dimension has provided an effective computation approach. The correlation integral $C(r)$ for a set of points $X_1, X_2, ..., X_N$ is defined in [18] as:

$$C(r) = \lim_{N \to \infty} \frac{2q}{N(N-1)} \tag{2}$$

where q is the number of pairs (i, j) whose distance $d(X_i, X_j)$ is less than r. The correlation integral is related to the standard correlation function, being the definite integral of it. In [18], the relationship between the Hausdorff dimension and the correlation dimension (ν) is proven, which for a continuous random vector process X, in certain conditions, is smaller than the Hausdorff dimension $dim_H X : 0 \leq \nu \leq dim_H X \leq d$. The theoretical Hausdorff dimension is not used in practice and it is approximated by various equivalent definitions in the discrete domain. For the definition of the Hausdorff dimension the reader is invited to consult [13].

In order to use the correlation dimension for the assessment of the complexity of a colour image, the definition 2 has to be extended to a 5-dimensional space and a colour representation should be chosen. In our approach, the extension to the colour domain is based on the ΔE_{2000} colour distance between the pixels of a colour video frame, which is the latest CIE standard: the CIE recommendations for colour distance evolved from the initial ΔE, to the ΔE_{94} and finally ΔE_{2000}. The equation of colour distance ΔE_{2000} is the following:

$$\Delta E_{2000} = \sqrt{\left(\frac{\Delta L'}{K_L S_L}\right)^2 + \left(\frac{\Delta C'}{K_C S_C}\right)^2 + \left(\frac{\Delta H'}{K_H S_H}\right)^2 + R_T \left(\frac{\Delta C'}{K_C S_C}\right)\left(\frac{\Delta H'}{K_H S_H}\right)} \tag{3}$$

The parameters K_L K_C and K_H weight the formula depending on the conditions of observation. The following terms were added in ΔE_{2000} in order to bring several corrections, and ultimately the ΔE_{2000} has a better behavior that suits the human vision than ΔE and ΔE_{94} for small colour differences:

- S_L: Compensation for lightness, corrects for the fact that Delta gives the predictions larger than the visual sensation for light or dark colours;
- S_C: Compensation for chroma, mitigates for the significant elongation of the ellipses with the chromaticity;
- S_H: Compensation for hue, that corrects for the magnification of ellipses with the chromaticity and hue;
- R_T: To take into account the rotation of the ellipses in the blue.

The steps of the algorithm for the estimation of colour correlation dimension involve the computation of colour distances histogram and the calculus of their cumulative density function, which is then represented in a log-log space. The resulted regression line's slope represents an estimate of the correlation dimension. The regression line is computed through Matlab `robustfit()` function which generates 9 regression lines (based on the flavor of the least square classical technique) from which we chose the mean slope. In [19] the authors concluded that the Andrews approach for the computation of the regression line led to the largest fractal dimension and implicitly to the largest range of values. However, the fractal dimension in that case was estimated using a box-counting definition.

3 Experiments

In order to study the relevance of fractal dimension computed by means of correlation dimension in the perceptual quality of images, we compared the measurements obtained from a video sequence transported over a network with zero loss to the measurements obtained from the received video sequence that has been transported over the same network, but in heavy load conditions in which a certain amount of packet loss also occurred.

3.1 Experimental Setup

We chose an MPEG-4 video streaming application - very sensitive to packet loss, due to the fact that neither UDP itself nor the video streaming application implement a retransmission mechanism: any lost packet in the network will cause missing bits of information in the MPEG video stream. In addition, the generated traffic is inelastic, because the transmission rate is not adapted to network conditions in any way. Therefore, the packet loss is the major issue for an MPEG-4 video streaming application.

In our experiments the induced packet loss percentage varied from 0% to 1.3% - the maximum amount of loss for which the application still functions and tests cannot be performed. We used the Helix streaming server from Real Networks (`http://www.realnetworks.com`) as MPEG-4 streaming server and the MPEG-4 client was mpeg4ip (`http://mpeg4ip.sourceforge.net`). The source code of the client was modified in order to record the received video sequence as individual frames in bitmap format. The test sequence we used was the widely-used 10-second long video test sequence football (250 frames, of 320×240 pixel each). The average transmission rate was approximately 1 Mb/s, which was a constraint imposed by the use of a trial version of the MPEG-4 video streaming server; however it represents a realistic scenario. In the next section the results for a video stream with an induced 1% normally distributed packet loss are presented.

3.2 Experimental Results

In order to quantify the perceived colour diversity of the frames in the original and altered video streams we computed the histograms of ΔE_{2000} distance in the CIELAB colour space between all pairs of different pixels within each frame. We have observed that due to the high amount of irregularities and colours generated artificially by erroneous estimations determined by the lack of data (lost in transmission), the histograms of affected frames expose a much wider distribution than the ones of unaltered frames. Fig. 2, for instance, illustrates the histograms of one original video frame of the unaltered stream (Fig. 1(a)) and the corresponding degraded video frame of the altered stream (Fig. 1(b)).

(a) Original video frame (b) Degraded video frame

Fig. 1. Corresponding Original and Degraded Frames in Unaltered/Altered Video Streams

The cumulative distribution functions of the ΔE_{2000} distance in L*a*b* space encountered within the selected frames expressed in a logarithmic reference system are shown in Figure 3. The robust multilinear regressions of the CDFs are computed through the implementation provided by Mathworks' MATLAB. By default, the algorithm uses iteratively reweighted least squares with a bisquare weighting function, but for precision purposes we experimented with the 9 available weighting functions: Andrews, bisquare(default), Cauchy, fair, Huber, logistic, ordinary least squares (no weighting function), Talwar and Welsch. We also computed the average slopes of the regression lines obtained by using these weighting functions, which we used in further statistical analysis. Table 1 shows the regression slopes for the above mentioned frames. One can notice that for the degraded video frame, clearly exhibiting an increase in complexity, the average estimated correlation dimension is larger than the one of the corresponding original video frame.

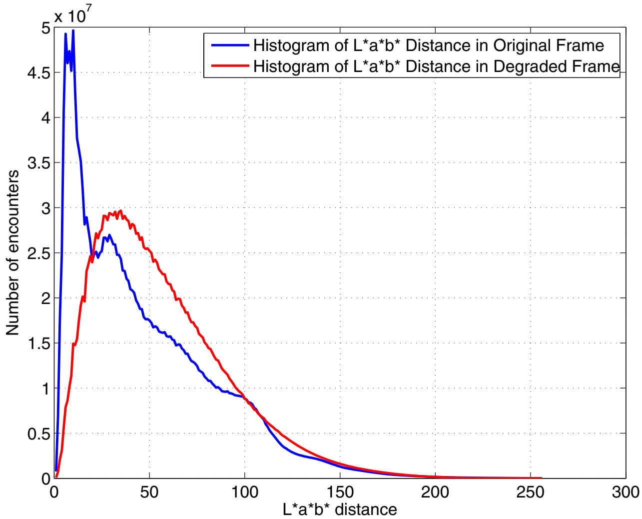
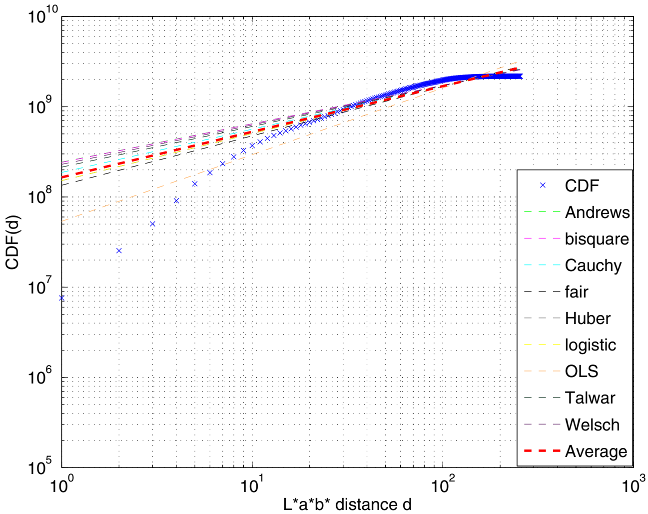


Fig. 2. Histograms of inter-pixel ΔE_{2000} Distance in L*a*b* Space of Corresponding Frames in Unaltered/Altered Video Streams

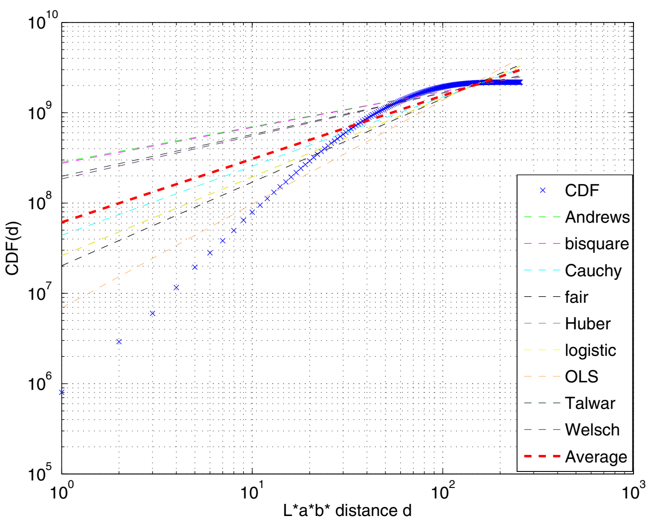
Table 1. Regression slopes of CDF of ΔE_{2000} distance in L*a*b* space computed in logarithmic reference system (i.e. Correlation Dimension) using several weighting functions

	Original Video Frame	**Degraded Video Frame**
Andrews	0.42511	0.39097
bisquare	0.42615	0.39672
Cauchy	0.47903	0.76571
Fair	0.54682	0.9248
Huber	0.52227	0.87212
Logistic	0.52217	0.87241
OLS	0.73887	1.1498
Talwar	0.45088	0.46192
Welsch	0.43661	0.47543
Average	**0.50532**	**0.7011**
σ	**0.098689705**	**0.2763793**

The average slope of the regression line (i.e. the correlation dimension), along with its variation by computation method is illustrated in Figure 4, for both the original (Fig. 4(a)) and the altered (Fig. 4(b)) video streams. The average is represented with a red line and the variation of the correlation dimension with green. One can notice the intervals for which both the larger average correlation dimension and its larger variance indicate the degradation of the video stream, e.g. around frames #50, #130 and #150.

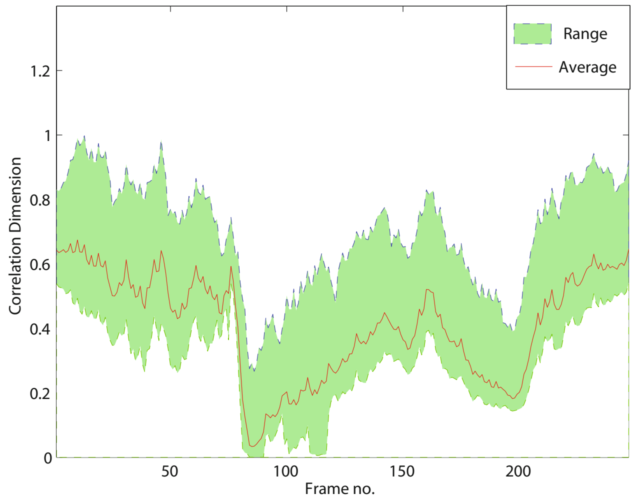


(a) Original video frame

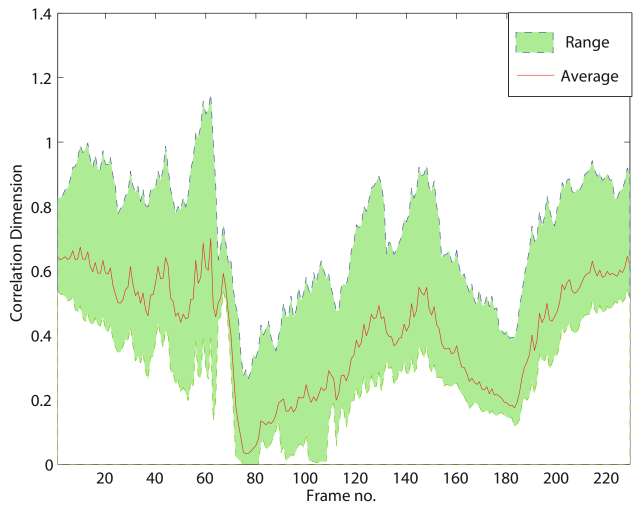


(b) Degraded video frame

Fig. 3. CDFs of inter-pixel ΔE_{2000} Distance in L*a*b* Space of Corresponding Frames in Unaltered/Altered Video Streams



(a) Original Stream



(b) Altered Stream

Fig. 4. Regression Slopes of CDFs of inter-pixel ΔE_{2000} Distance in L*a*b* Space between pixels within correspondent frames of Original/Degraded Video Streams

4 Conclusions

We presented in this article an extension of the correlation dimension to the colour domain, thus we are able to estimate the fractal complexity of a colour image. As a possible application, in the assessment of the video quality, we present the results we obtained for a video sequence, in the case of an MPEG-4 video streaming application. We conclude that the colour correlation dimension can be used for the objective video quality assessment in a reduced-reference scenario. However, the real-time assessment of video quality through fractal metrics remains an open question, given the not-neglectable algorithm complexity. A possible solution we envisage is the use of GPU for the acceleration of the approach.

Acknowledgement. This work and the paper were partially supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), ID76945 financed from the European Social Fund and by the Romanian Government.

References

[1] ITU-R Recommendation BT.500. Subjective quality assessment methods of television pictures. International Telecommunications Union (1998)
[2] ITU-T Recommendation P.910. Subjective video quality assessment methods for multimedia applications. International Telecommunications Union (1996)
[3] ITU-R Recommendation J.140. Subjective assessment of picture quality in digital cable television systems. International Telecommunications Union (1998)
[4] ITU-T Recommendation J.143. User requirements for objective perceptual video quality measurements in digital cable television. International Telecommunications Union (2000)
[5] Wang, Z., Bovik, A.C.: Mean squared errorl: Love it or leave it? IEEE Signal Processing Magazine, 98–117 (January 2009)
[6] Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004)
[7] Sampat, M.H., et al.: Complex wavelet structural similariy: A new image similarity index. IEEE Transactions on Image Processing 18(11), 2385–2401 (2009)
[8] Carnec, M., Le Callet, P., Barba, D.: Objective quality assessment of color images based on a generic perceptual reduced reference. Image Commun. 23, 239–256 (2008) ISSN: 0923-5965, `http://portal.acm.org/citation.cfm?id=1371260.1371348`
[9] Cheng, G., et al.: Image quality assessment using natural image statistics in gradient domain. AEU - International Journal of Electronics and Communications 65(5), 392 (2011) ISSN: 1434-8411, `http://www.sciencedirect.com/science/article/pii/S143484111000172X`
[10] Lu, W., et al.: No-reference image quality assessment in contourlet domain. Neurocomput. 73, 784–794 (2010) ISSN: 0925-2312, `http://dx.doi.org/10.1016/j.neucom.2009.10.012`

[11] Ivanovici, M., Richard, N., Fernandez-Maloigne, C.: Towards video quality metrics based on colour fractal geometry. EURASIP Journal on Image and Video Processing (2010) (2010)
[12] Voss, R.: Random fractals: characterization and measurement. Scaling phenomena in disordered systems 10(1), 51–61 (1986)
[13] Falconer, K.: Fractal Geometry, mathematical foundations and applications. John Wiley and Sons (1990); ISBN: 0-471-9228-70
[14] Allain, C., Cloitre, M.: Characterizing the lacunarity of random and deterministic fractal sets. Physical Review A 44(6), 3552–3558 (1991)
[15] Plotnick, R.E., Garnder, R.H., Hargrove, W.H., Prestgaard, K., Perlmutter, M.: Lacunarity analysis: a general technique for the analysis of spatial patterns. Physical Review E 53(3), 5461–5468 (1996)
[16] Maragos, P., Sun, F.K.: Measuring the fractal dimension of signals: morphological covers and iterative optimization. IEEE Transactions on Signal Processing 41(1), 108–121 (1993)
[17] Grassberger, P., Procaccia, I.: Measuring the strangeness of strange attractors. Physica D: Nonlinear Phenomena 9(1-2), 189–209 (1983)
[18] Bardet, J.-M.: Dimension de corrélation locale et dimension de haussdorf des processus vectoriels continus / local correlation dimension and hausdorff dimension of continuous random fields. Comptes Rendus de l Acadmie des Sciences - Series I - Mathematics 326(5), 589–594 (1998)
[19] Ivanovici, M., Richard, N.: Fractal dimension of colour fractal images. IEEE Transactions on Image Processing 20(1), 227–235 (2011)