A NEW METHOD FOR BLOCKING ARTIFACTS DETECTION BASED ON HUMAN VISUAL SYSTEM FEATURES

Authors:
Luis Fernando Abanto León
Guillermo Kemper Vásquez
Joel Tellles Castillo

ABSTRACT

- Block-based coding has become the state-of-the-art approach for image and video compression.
- When compression is performed at low bitrates annoying blocking artifacts appear in the block boundaries.
- In this paper we propose :
  - A No-Reference method for spatially locating blocking artifacts
  - An objective quality measure for images and video signals using neural networks.
- The proposed method takes into account Human Visual System features in order to achieve better correlation with subjective opinion.
- In the experiments, we have obtained a correlation coefficient of about 0.90, between subjective scores and the proposed objective quality metric

BACKGROUND

- One of the most annoying visual disturbances is the blocking effect.
- Spatial location of blocking artifacts is of great interest for many applications, e.g. blockiness post-filtering, monitoring of communication systems, and objective video/image quality assessment.
- Traditional approaches such as PSNR and MSE do not effectively correlate with human visual perception.
- Consequently much effort has been devoted for developing objective methods to correlate well with the human visual perception.

BLOCK DIAGRAM OF THE PROPOSED METHOD

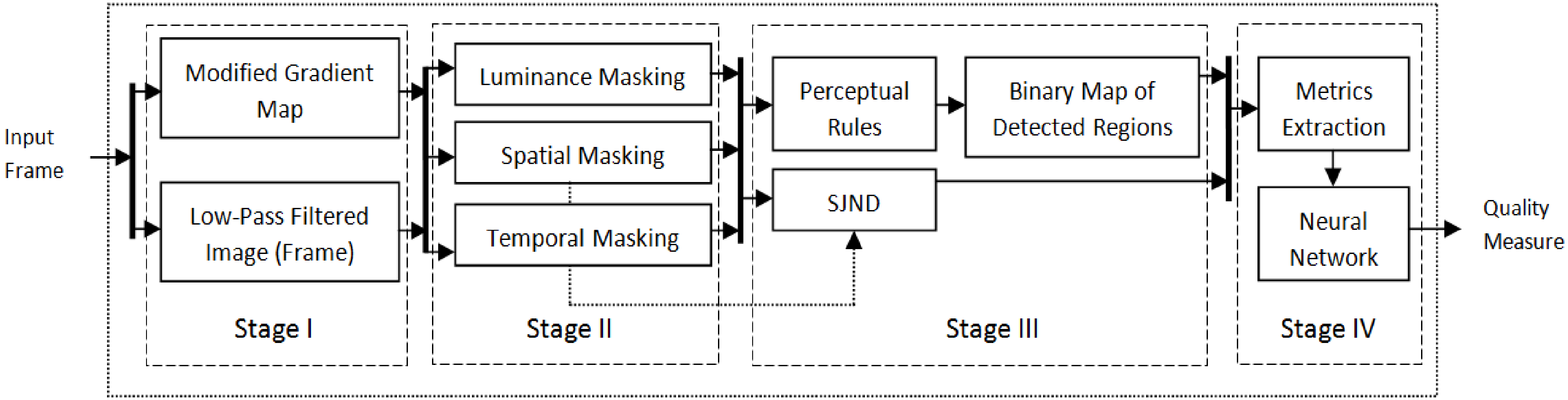


Figure 1. Block Diagram of proposed method

GRADIENT MAP COMPUTATION

- When dealing with images, blocking artifacts manifest itself as an abrupt discontinuity between neighboring blocks.
- When dealing with video, blocking artifacts have no fixed position of appearance due to motion estimation.
- In order to identify potential regions affected by blocking artifacts, we process video frames using a modified version of one-dimensional filters, which are based on the first-order partial derivatives approach.
- BJND values were experimentally obtained according to a method in which an image of constant gray level is gradually added noise of fixed amplitude.

$$C_1 = \begin{cases} 1 & \mathit{if} \quad L(i,j) < L(i,j+1) \\ 0 & \mathit{otherwise} \end{cases}$$

$$C_2 = \begin{cases} 1 & \mathit{if} \quad L(i,j+1) > L(i,j+2) \\ 0 & \mathit{otherwise} \end{cases}$$

$$HM(i,j) = \begin{cases} \max\{L(i,j) - L(i,j+1), 0\} & \mathit{if} \;\; C_1 + C_2 = 2 \\ 0 & \mathit{if} \;\; C_3 + C_4 = 2 \\ |L(i,j) - L(i,j+1)| & \mathit{if} \;\; C_5 = 1 \end{cases}$$

$$C_3 = \begin{cases} 1 & \mathit{if} \quad L(i,j) - L(i,j+1) \geq 0 \\ 0 & \mathit{otherwise} \end{cases}$$

$$C_4 = \begin{cases} 1 & \mathit{if} \quad L(i,j) - L(i,j+1) \leq BJND(L(i,j)) \\ 0 & \mathit{otherwise} \end{cases}$$

$$GM(i,j) = \min\left\{\sqrt{HM(i,j)^2 + VM(i,j)^2},\ 255\right\}$$

$$C_5 = \begin{cases} 1 & \mathit{if} \quad \{C_1 + C_2 < 2\} \ \mathit{and} \ \{C_3 + C_4 < 2\} \\ 0 & \mathit{otherwise} \end{cases}$$

LOW-PASS FILTERING

- A smoothed version of the frame under examination is computed. The outcome of this stage is called the Low-Pass Filtered Map

1	1	1	2	0	2	1	1	1
1	1	2	4	0	4	2	1	1
1	2	4	6	0	6	4	2	1
2	4	6	8	0	8	6	4	2
0	0	0	0	0	0	0	0	0
2	4	6	8	0	8	6	4	2
1	2	4	6	0	6	4	2	1
1	1	2	4	0	4	2	1	1
1	1	1	2	0	2	1	1	1

Figure 2. LPF Mask

$$F(i,j) = \frac{\sum_{x=0}^{2\cdot S+1} \sum_{y=0}^{2\cdot S+1} L(i-S+x, j-S+y)\cdot C(x,y)}{\sum_{x=0}^{2\cdot S+1} \sum_{y=0}^{2\cdot S+1} C(x,y)}$$

LUMINANCE MASKING, TEXTURE MASKING AND TEMPORAL MASKING

- HVS properties have been considered in the proposed method in order to reflect the image/video quality more efficaciously.
- It was found that the human visual system's sensitivity to variations in luminance mainly depends on the local mean luminance. We can summarize that high visibility thresholds are assumed in both very dark or bright regions, and low thresholds in regions of medium gray levels.

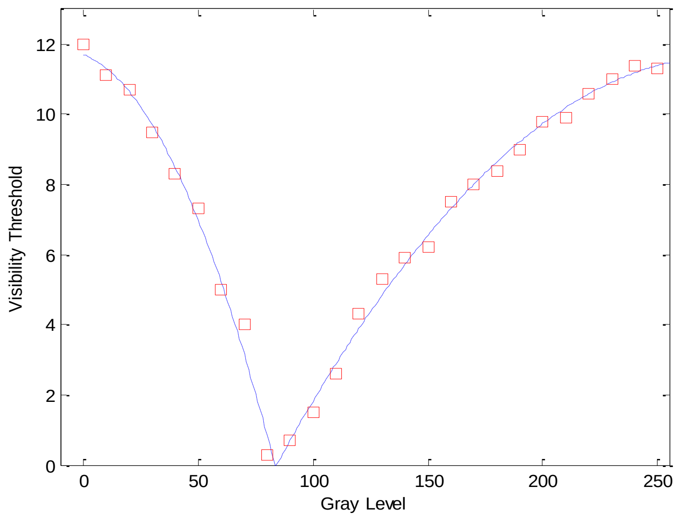


Figure 3. Visibility threshold due to luminance

$$W_l(i,j)=\begin{cases}A_1\cdot I^2(i,j)+B_1\cdot I(i,j)+C_1\\ A_2\cdot I^2(i,j)+B_2\cdot I(i,j)+C_2\end{cases}$$

A_1 = 0.000103895, B_1 = 0.0025918, C_1 = 0.009428690, A_2 = 0.0000220653, B_2 = -0.0125636, C_2 = 1.885972. In this case I represents the average luminance of a group of neighboring pixels.

LUMINANCE MASKING, **TEXTURE MASKING** AND TEMPORAL MASKING

- Another important mechanism of masking is the texture masking. Texture masking suggests that distortions in images may be either hidden or revealed, depending on the region in which they occur.
- The experimental results have shown that high values of texture masking lead to low values of visibility. Distortion visibility decreases near regions with spatial details.

$$T_M(i,j) = -\frac{1}{S_1}\left(\frac{S_2}{1+S_3\cdot I(i,j)} + S_4\cdot I(i,j) + S_5\right) + S_6$$

T_M represents the visibility threshold. Consider the following constants, $S_1 = 100$, $S_2 = 120.73853$, $S_3 = 0.026$, $S_4 = 0.062362$, $S_5 = -20.806316$ and $S_6 = 1.1$. In this paper I has been treated as the variance of luminance level within a group of neighboring pixels.

LUMINANCE MASKING, TEXTURE MASKING AND **TEMPORAL MASKING**

- Temporal masking is mainly based in the interframe difference.
- Larger inter-frame luminance differences result in larger temporal masking effect.
- The visibility thresholds for identifying distortions are determined as a function of the interframe variations and the average background luminance.
- The weighting function yields values between 0 and 1.

$$TJND(i,j,t) = \begin{cases} -\max\left(\tau, \frac{H_1}{2}\cdot\exp\left(\frac{-0.15}{2\pi}(\Delta(i,j,t)+255)\right)+\tau\right)+Z \\ \quad if \qquad \Delta(i,j,t)\le 0 \\ -\max\left(\tau, \frac{H_2}{2}\cdot\exp\left(\frac{-0.15}{2\pi}(255-\Delta(i,j,t))\right)+\tau\right)+Z \\ \quad if \qquad \Delta(i,j,t) > 0 \end{cases}$$

$$\Delta(i,j,t) = \frac{L(i,j,t)-L(i,j,t-1)+F(i,j,t)-F(i,j,t-1)}{2}$$

$\tau = 0.8$, $H_1 = 8$, $H_2 = 3.2$ and $Z = 5.2090$.

Due to the fact that $H_1 > H_2$, high to low luminance transitions induces a more relevant masking than low to high transitions.

RULES FOR BLOCKING ARTIFACTS DETECTION

- Five rules were developed in order to efficiently detect blocking artifacts. These rules are applied to both directions, horizontal and vertical.
- Each rule is composed of nine conditions. If one of the rules meets the conditions, the pixels under examination are considered to be disturbed by blocking artifacts.

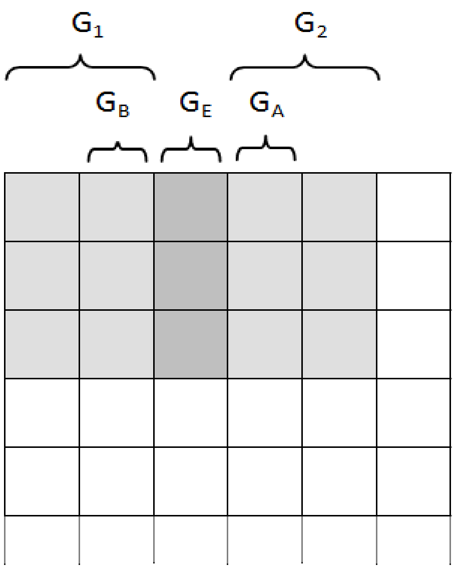


Figure 4. Gradient Map

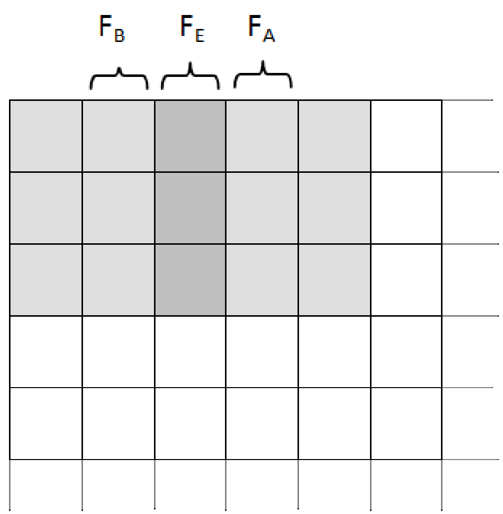


Figure 5. LowPass Filtered Map

Conditions

$$uG_1 = Mean(G_1)$$

$$uG_2 = Mean(G_2)$$

$$uG_E = Mean(G_E)$$

$$sm_{G_E} = \lambda_1 \cdot TextureMasking(\sqrt{Variance(G_E)})$$

$$sm_{G_B} = \lambda_2 \cdot TextureMasking(\sqrt{Variance(G_B)})$$

$$sm_{G_A} = \lambda_3 \cdot TextureMasking(\sqrt{Variance(G_A)})$$

$$lm_{F_E} = \alpha \cdot LumianceMasking(Mean(F_B, F_E, F_A))$$

$$tm_{F_E} = TemporalMasking(Mean(F_B, F_E, F_A))$$

RULES FOR BLOCKING ARTIFACTS DETECTION

- Nine conditions implement each rule

$$rule_n = \begin{cases} true & if \quad \prod_{k=1}^{9} \Theta_{k,n} = 1 \\ false & otherwise \end{cases}$$

$$\Theta_{1,n} = \begin{cases} 1 & if \quad uG_1 \le \gamma_n \\ 0 & otherwise \end{cases}$$

$$\Theta_{2,n} = \begin{cases} 1 & if \quad uG_2 \le \Omega_n \\ 0 & otherwise \end{cases}$$

$$\Theta_{3,n} = \begin{cases} 1 & if \quad uG_E \ge \varphi_n \\ 0 & otherwise \end{cases}$$

$$\Theta_{4,n} = \begin{cases} 1 & if \quad uG_E < \xi_n \\ 0 & otherwise \end{cases}$$

$$\Theta_{5,n} = \begin{cases} 1 & if \quad sm_{G_E} < \kappa_n \\ 0 & otherwise \end{cases}$$

$$\Theta_{6,n} = \begin{cases} 1 & if \quad lm_{F_E} > \psi_n \\ 0 & otherwise \end{cases}$$

$$\Theta_{7,n} = \begin{cases} 1 & if \quad tm_{F_E} > \rho_\pi \\ 0 & otherwise \end{cases}$$

$$\Theta_{8,n} = \begin{cases} 1 & if \quad sm_{G_B} < \sigma_n \\ 0 & otherwise \end{cases}$$

$$\Theta_{9,n} = \begin{cases} 1 & if \quad sm_{G_A} < \varsigma_n \\ 0 & otherwise \end{cases}$$

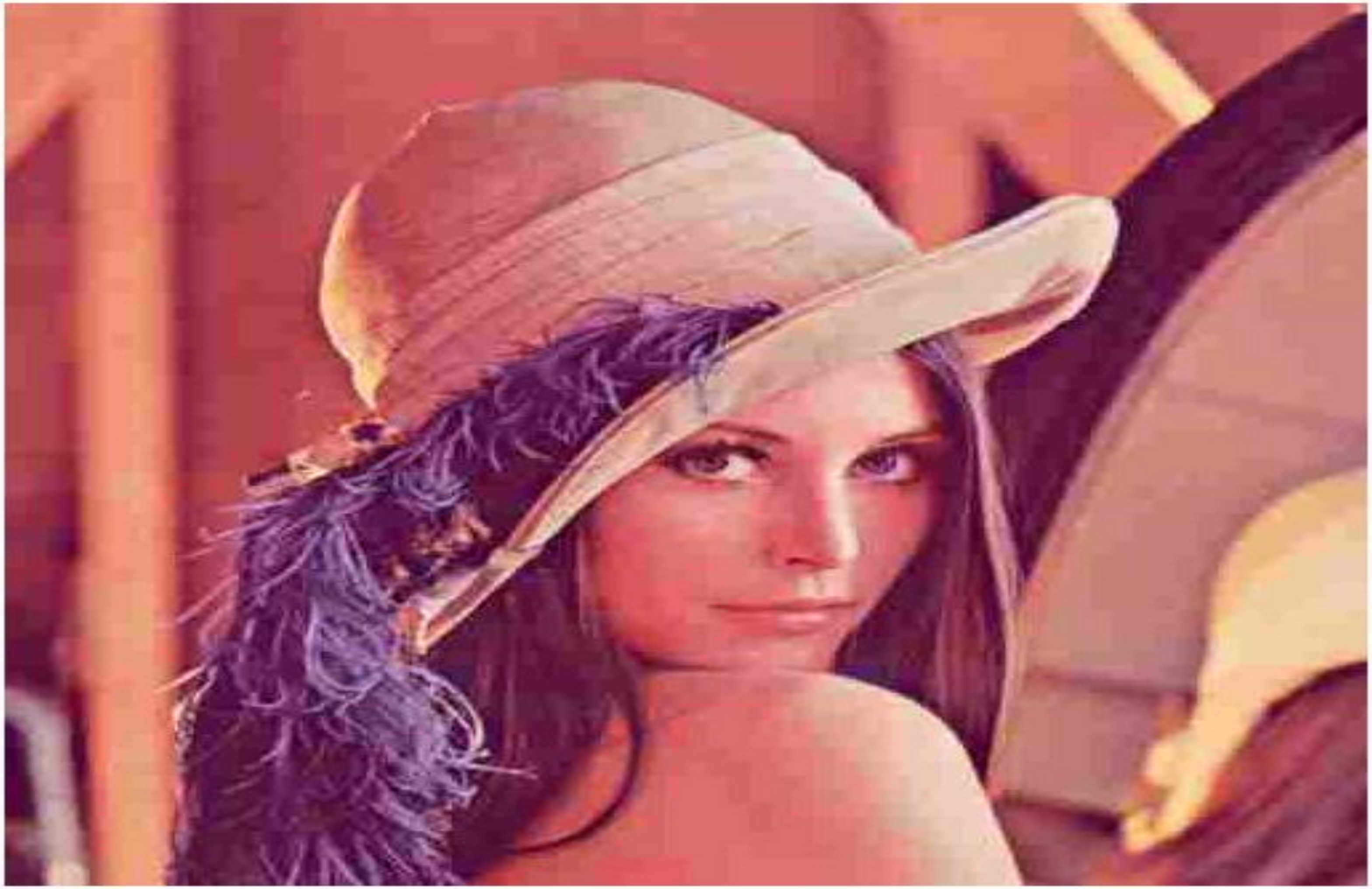
The five rules have to be evaluated at each displacement of the window.
The overlapping between the current position with next position of the sliding window is 80%.

n	γ_n	Ω_n	Φ_n	ξ_n	κ_n	ψ_n	ρ_n	σ_n	ζ_n
1	0.5	0.5	3.0	12.0	0.2174	0.30	0.78	1.0	1.0
2	2.0	2.0	7.2	20.0	0.2174	0.17	0.77	1.0	1.0
3	4.0	4.0	15.4	20.0	0.2853	0.17	0.78	1.0	1.0
4	5.0	5.0	15.0	20.0	0.4076	0.2	0.79	1.0	1.0
5	5.0	5.0	5.0	20.0	0.21	0.4	0.78	0.14	0.14

BINARY DISTURBANCE MAP

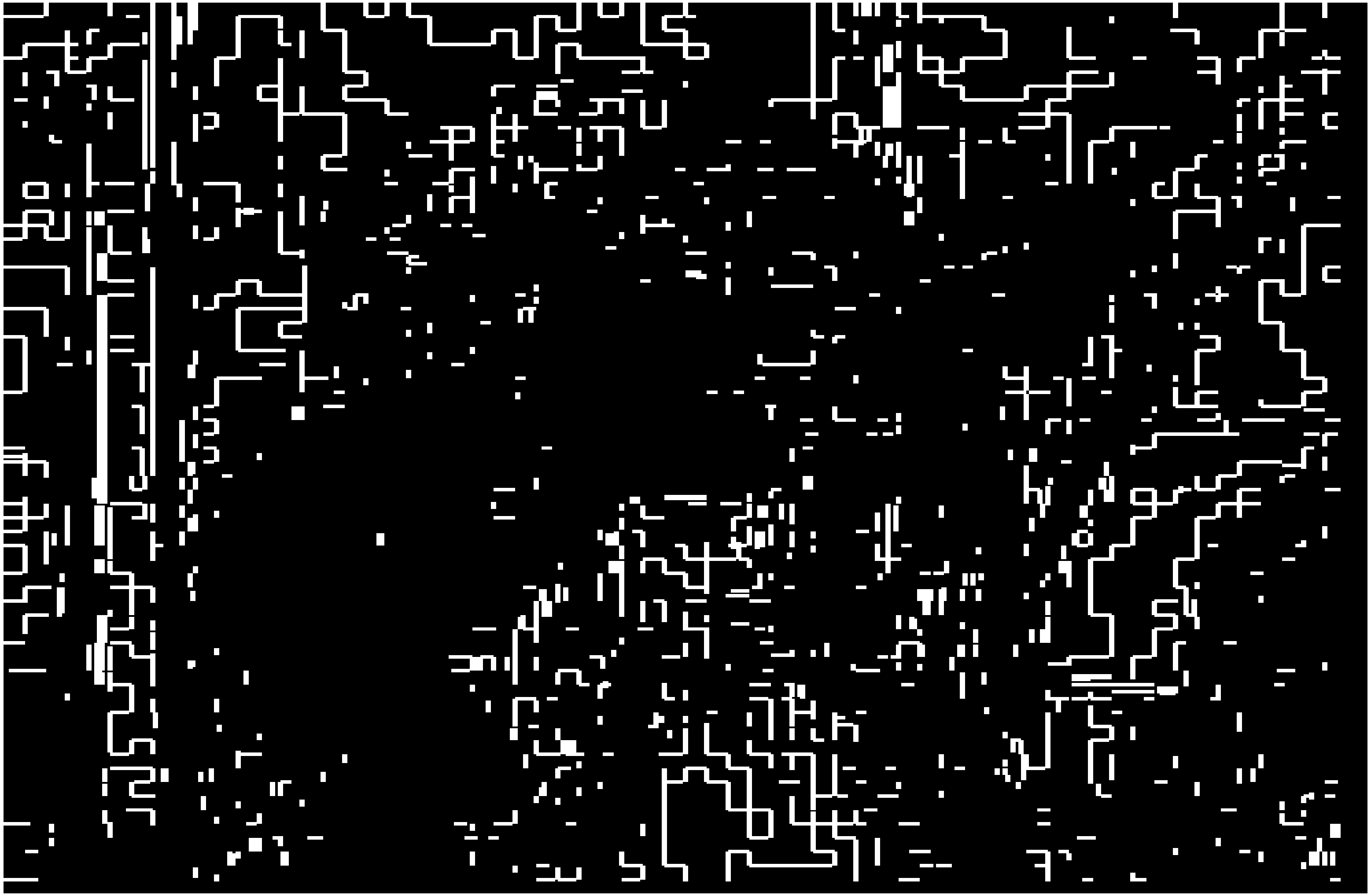
- The Binary Disturbance Map is directly related with the fulfillment of the Perceptual Rules.
- The Binary Disturbance Map (BDM) is a matrix with the same dimensions of the frame or image under evaluation.
- The five rules are applied in every region of the image/frame and if any of the five rules meets all the conditions, then that region is considered to be distorted
- Then the BDM is updated with the value of “1” in the same position that distortions were encountered. BDM will represent the presence or absence of distortions with a binary value : “0” or “1”.

BINARY DISTURBANCE MAP



BINARY DISTURBANCE MAP

BINARY DISTURBANCE MAP



METRICS

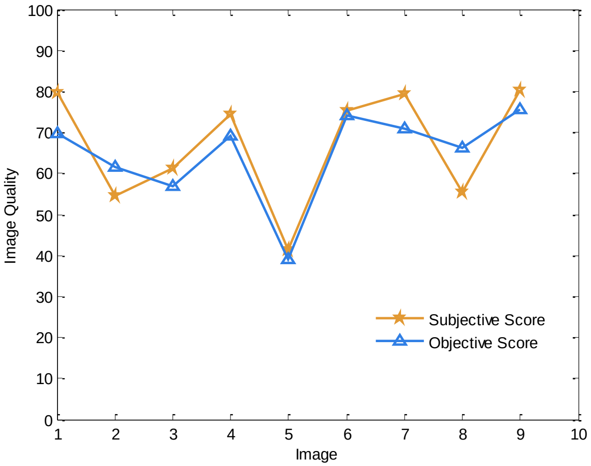
- DTP : Distorted Pixels To Total Pixels Ratio
- This metric is expressed as the ratio between the total pixels in the frame and the total distorted pixels in the frame

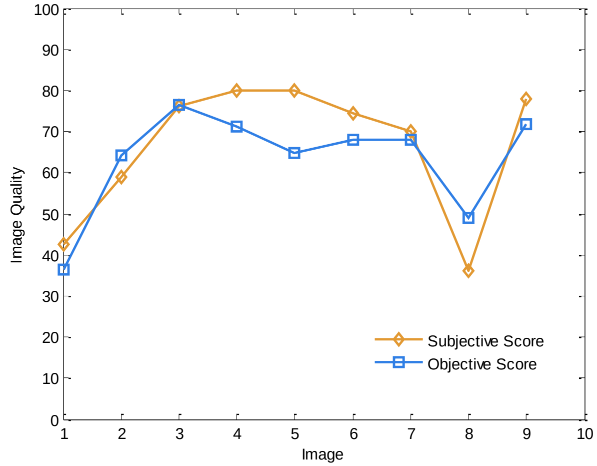
$$DTR = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} BDM(i, j)}{M \cdot N}$$

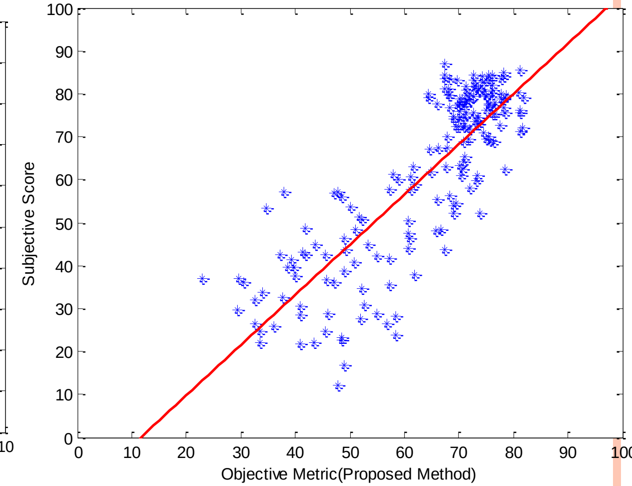
- APD : Average Pixel Distortion
- This metric is based in the computation of the surplus or difference of GM over the SJND Map, i.e. the numerator represents the total distortion of the image. The denominator represents the number of distorted pixels, sot the APD metric represent the average distortion per pixel. Typically values for this metric range between 0 and 15.

$$APD = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \max(\min(GM(i, j), 15) - SJND(i, j), 0) \cdot BDM(i, j)}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} BDM(i, j)}$$

EXPERIMENTAL RESULTS







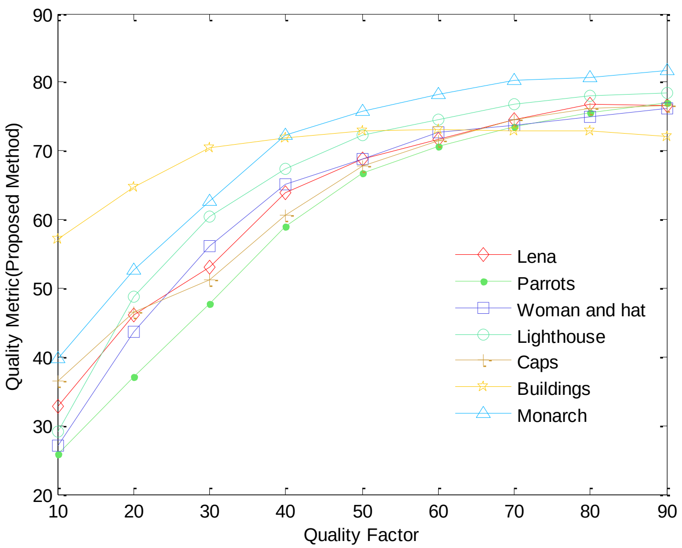


Image Number	Methods/Metrics					
	DF_{IMAGE}	*S*	*NPBM*	*WSBM*	*Proposed Method*	*Subjective Score*
7	0.1236	0.444	0.0929	2.5018	70.09	74.3
16	0.3326	0.4445	0.2743	2.8735	61.93	38.0
37	1.5766	0.9993	0.4323	6.6912	37.74	32.5
49	0.3642	0.1082	0.4724	3.1186	61.27	57.75
54	0.2415	0.5942	0.0942	2.9487	71.36	79.65
69	0.8038	0.3016	0.2235	7.6194	57.3	41.65
81	0.4048	0.6282	0.3392	3.6066	60.84	46.4
101	0.1795	0.0019	0.2006	2.2221	76.06	75.35
113	0.0888	0.4337	0.1152	2.5344	71.33	78.3
125	0.346	0.0507	0.1873	3.0378	81.11	76.07
133	0.3463	0.5302	0.1104	3.0492	75.41	84.38
152	0.8518	0.8336	0.4508	5.2679	50.7	40.92
178	0.7687	0.8108	0.1213	5.7435	64.57	67.07
Corr. Coef	-0.692	-0.405	-0.767	-0.641	0.875	

CONCLUSIONS

- In regard to the objective method for image and video quality assessments that we are proposing, the results correlate well with the Mean Subjective Opinion, a correlation degree of about 90%.

- In respect to the method for blocking artifacts location, we think that the block-based analysis with overlapping is the best approach to detect and estimate blocking artifacts in video sequences. We believe that the introduction of additional HVS features could improve the correlation with the subjective opinion.