

A Fast Computation Method for IQA Metrics Based on their Typical Set

Vittoria Bruni^{1,2} and Domenico Vitulano²

¹Dept. of SBAI, Faculty of Engineering, Univ. of Rome "Sapienza", Via A. Scarpa 16, 00161 Rome, Italy

²Istituto per le Applicazioni del Calcolo "M. Picone" - C.N.R., Via dei Taurini 19, 00185 Rome, Italy

Keywords: Information Theory, SSIM, Asymptotic Equipartition Property, Image Quality Assessment.

Abstract: This paper deals with the *typical set* of an image quality assessment (IQA) measure. In particular, it focuses on the well known and widely used Structural SIMilarity index (SSIM). In agreement with Information Theory, the *visual distortion typical set* is composed of the least amount of information necessary to estimate the quality of the distorted image. General criteria for an effective and fruitful computation of the set will be given. As it will be shown, the typical set allows to increase IQA efficiency by considerably speeding up its computation, thanks to the reduced number of image blocks used for the evaluation of the considered IQA metric.

1 INTRODUCTION

Several neurological studies proved that a few points attract human attention in the early vision (Monte et al., 2005; Frazor and Geisler, 2006). Those points are able to code the significant content of the scene and are known as *fixation points* (Monte et al., 2005; Frazor and Geisler, 2006; Winkler, 2005). They allow to synthesize (and understand (Grunwald, 2004)) image information in a very small lapse of time — 200-300 msec per fixation point. A wide literature focused on methods and algorithms able to find these characteristic points and most of them rely on the concept of *saliency maps* (Wang et al., 2010; Rivera et al., 2007; Benabdelkader and Boulemden, 2005; Bruni et al., 2011) i.e., those maps that label image content in a hierarchical way, according to its visual appearance. As a matter of fact, bearing in mind some well-known concepts of Information Theory (IT), the aforementioned set of points can be seen as the *typical set* of the "source image" (Cover and Thomas, 1991) i.e., the one that contains all the information concerning the visual content of the image; that is why we will refer to it as the *visual typical set*. On the other hand, recently the growing need of measures that correlate with the Human Visual System (HVS) better than the classical Signal-to-Noise Ratio (SNR) (Winkler, 2005; Gonzalez and Woods, 2002) led to the definition of new Full Reference (FR) quality measures, that compare the original image I with a distorted version J (Sheikh et al., 2005; Sheikh and Bovik, 2006;

Zhang and Jernigan, 2006; Wang and E.P.Simoncelli, 2005; Wang and Li, 2011; Bruni et al., 2013a; Bruni et al., 2013b). Despite their high correlation with HVS, most of the proposed FR measures are computationally more demanding than SNR or PSNR (Peak Signal to Noise Ratio) and then less attractive for real time applications, especially video processing based applications. The objective of this paper is to ask whether there exists a *Visual Distortion Typical Set* A_M^ϵ for a given FR quality measure M . In other words we are wondering if a given FR measure can be successfully evaluated from a reduced number of image pixels. In fact, as it happens in the observation process of a single image, an observer usually 'looks at' just some salient regions in the observed image, rather than checking all its pixels, before assigning a quality score to it. It turns out that it is reasonable to assume that there exists an 'absolute' *Visual Distortion Typical Set*: the one employed by HVS in the observation process. For image quality assessment, A_M^ϵ will be then composed of a subset of corresponding pixels in the original image I and in the distorted J and it will also depend on the FR quality measure M and on ϵ . The latter represents the distance between M estimated on A_M^ϵ (i.e. \hat{M}) and M estimated on the whole images I and J (i.e. \bar{M}). Since the search of A_M^ϵ seems to be not straightforward at all, this paper will focus on the A_M^ϵ for a specific Full Reference measure, namely the Structural SIMilarity index (SSIM) (Wang et al., 2004a; Wang et al., 2004b), and it will be studied from both a practical and theoretical point of

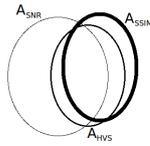


Figure 1: Venn diagram for the typical sets of SSIM, SNR and HVS (unknown).

view. We expect that any FR measure has a different A_M^ε , with a partial (hopefully wide) intersection with the absolute one, as depicted in Fig. 1, according to what extent the FR measure correlates with HVS. The estimation of the *visual distortion typical set* A_M^ε is twofold advantageous. From a practical point of view, it would allow to estimate \bar{M} just from a subset of the available information (I and J) within a small error ε , with a considerable computational saving. From a more theoretical point of view, it would permit to better understand the HVS and *i*) to design new and more precise FR measures, able to simulate the complexity of the human vision *ii*) to embed FR quality measures in the minimization of 'HVS based functionals', *iii*) to make a basis for a formal explanation of Visual Information Theory (Bruni et al., 2013b), with effects on new visible image coding schemes etc..

In this paper we will focus on the reduction of the computational complexity of FR measures, with particular reference to SSIM. The aim is to show that A_M^ε allows to speed up SSIM evaluation with a very small estimation error i.e., SSIM can be estimated from a selected subset of blocks with high precision. Experimental results on LIVE database show the robustness of the proposed method to different kinds of degradation.

The outline of the paper is the following. Next section addresses the problem of how to define A_M^ε and how to look for a sequence belonging to it. Some theoretical findings that guide a correct SSIM computation and its complexity reduction will be presented. Section 3 presents some experimental results that confirm the theoretical findings of Section 2. Concluding remarks and guidelines for future research are given in the last Section.

2 VISUAL DISTORTION TYPICAL SET

The estimation of the *visual distortion typical set* A_M^ε yields as side effect some interesting theoretical results that will be a good ground basis for practical purposes. In order to better understand them, we will consider one of the most effective and widely used HVS-based FR IQA measures: SSIM (Wang et al.,

2004a; Wang et al., 2004b). Starting from an $W_1 \times W_2$ (original) image I and a distorted version J , SSIM can be computed via the following simple algorithm:

1. Split I into a set of N_0 overlapping blocks $\{b_i\}$ of size $l \times l$ and centered at each pixel of I (note that $\frac{W_1 \times W_2}{l \times l} \leq N_0 \leq W_1 \times W_2$). Make the same for the distorted version J , achieving blocks $\{d_i\}$.
2. For each couple of blocks (b_i, d_i) , estimate SSIM: $M_i = \frac{2\mu_{b_i}\mu_{d_i}+C_1}{\mu_{b_i}^2+\mu_{d_i}^2+C_1} \frac{2\sigma_{b_i d_i}+C_3}{\sigma_{b_i}^2+\sigma_{d_i}^2+C_2}$, where C_1, C_2 and C_3 are numerical stabilizing constants (see (Wang et al., 2004a) for details). The array M (or the matrix, as to each pixel of I or J can be assigned the corresponding SSIM value) is then produced.
3. Compute the mean¹ of M : $\bar{M} = \frac{1}{N_0} \sum_{i=1}^{N_0} M_i$.

SSIM is computed in correspondence to each pixel of the image, it involves block-based operations and it adopts a pooling strategy by assigning equal weights to each pixel. However, one may wonder if these are the best implementation choices. For instance, one may ask for:

1. **Reduction of the Information.** Is the whole I and J 's information really important?
2. **Selection of the Best Reduction Domain.** Is it more convenient to reduce I and J 's information or to reduce the M 's information?
3. **Locality of the Selected Information.** Is it more convenient to select I (J) samples from local regions (for instance, blocks) or to select them randomly (not locally) from I (J)?
4. **Overlapping Blocks.** In the case of blocks based measures, have blocks to be overlapped?
5. **Procedure for Finding A_M^ε .** Is there a formal (and possibly fast) procedure to find this reduced information?

A formal answer to these questions is given below.

2.1 Reduction of Information

From a qualitative point of view, the *visual distortion typical set* A_M^ε can be defined as a subset of all sequences composed of samples of I (and the corresponding ones of J) such that they give an approximated value (\hat{M}) of the expected value \bar{M} of M within an error ε , i.e.: $|\hat{M} - \bar{M}| < \varepsilon$. More formally, A_M^ε can be thought in terms of Information Theory quantities (Cover and Thomas, 1991). Shannon typical set is

¹If the mean is computed in the whole image for $N_0 = W_1 \times W_2$, it corresponds to the expected value of M , i.e. $E[M]$. However, in order to make the notation less heavy, the symbol $\bar{\cdot}$ will be used in the paper.

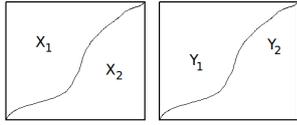


Figure 2: The original image I is composed of two homogeneous regions X_1, X_2 . Y_1 and Y_2 are the corresponding regions in the distorted copy J .

defined as the set of sequences of fixed size whose entropy is close to the entropy of the source. Similarly, we can think about the original image I as the first source associated to the variable X , its distorted version J associated to the variable Y while the variable $Z = M(X, Y)$ characterizes the source M which, in turn, depends on X and Y . A_M^ϵ also depends, even though not explicitly, on the kind of degradation D_J that produces J starting from the original image I : $J = D_J(I)$. However, D_J can be considered 'embedded' in J and it will not explicitly mentioned in the sequel. A_M^ϵ will be then composed of the subset of sequences $\{X_1, \dots, X_{N_r}, Y_1, \dots, Y_{N_r}\}$ of size $2N_r < 2N_0$ such that for a fixed $\epsilon > 0$ it holds

$$|\overline{M}(X, Y) - \overline{M}(X_1, \dots, X_{N_r}, Y_1, \dots, Y_{N_r})| < \epsilon. \quad (1)$$

The existence of A_M^ϵ is guaranteed by existing information theoretic results (Cover and Thomas, 1991), that is why we will focus on how to select and characterize the sequence $(X_1, \dots, X_{N_r}, Y_1, \dots, Y_{N_r})$ such that eq. (1) is satisfied. This problem can be seen as an application of the weak law of large numbers, which states that for i.i.d. r.v.s X_i it holds $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\mathbf{P}} E[X]$ $n \rightarrow \infty$, where \mathbf{P} indicates the convergence in probability and $E[X]$ is the expected value of X . However, it is more convenient to use the equivalent concept, known as the Asymptotic Equipartition Property (AEP) (Cover and Thomas, 1991), i.e.: $\frac{1}{n} \log \frac{1}{p(X_1, X_2, \dots, X_n)} \xrightarrow{\mathbf{P}} H(X)$ $n \rightarrow \infty$, where $H(X)$ is the entropy of X and X_i are i.i.d. r.v.s. That is why in the sequel just the entropy will be considered. Entropy is more mathematically tractable as it gradually increases as the number of samples grows (Cover and Thomas, 1991), while it is not so for the mean value, as proved by the following Proposition whose proof is in the Appendix.

Prop. 1 Let $X \sim Q$ with finite alphabet χ and $\{X_1\} \sim p_1, \{X_1, X_2\} \sim p_2, \dots, \{X_1, X_2, \dots, X_n\} \sim p_n, \dots$ while μ_n be the mean value of the pdf p_n , μ be the mean value of the pdf Q and $M_n = \max_{x \in \chi} |x|$. Then

1. in general, the sequence $\{\mu_n\}$ is not monotonic for increasing n ;
2. $|\mu_n - \mu|^2 \leq 2M_n^2 D_{KL}(p_n || Q), \quad \forall n$



Figure 3: Original images of LIVE database used in this paper: Ocean, Stream, Lighthouse, Flowersonih35, House, Sailing4.

2.2 Best reduction Domain

In order to get a typical subsequence $\{X_1, \dots, X_{N_r}, Y_1, \dots, Y_{N_r}\}$, one may ask whether it is more convenient to reduce information of the sources X and Y and then to estimate \overline{M} from them (and then Z) or to leave X and Y unchanged, whereas to reduce Z 's information. This is the topic of the following Proposition:

Prop. 2 $H(Z) \equiv H(M(X, Y)) \leq H(X, Y)$.

The proof is omitted since it straightforwardly derives from the well-known result: $H(f(X)) \leq H(X)$, for any function f and random variable X (Cover and Thomas, 1991). In practice, since part of the information of X and Y is lost in the computation of $M(X, Y)$, it is more convenient to leave X and Y unchanged and to reduce Z .

2.3 Locality of Information

Though the results above would lead to select information directly from M , it is necessary to find a strategy to only take part of the information directly from X and Y . In fact, with regard to SSIM, it is useless to firstly build the whole vector M to take just a subset of its samples. It corresponds to the use of two weights (1 or 0) for M_i in the pooling step (step 3 of SSIM algorithm). More formally, we can think of the subsequence $\{X_1, \dots, X_{N_r}, Y_1, \dots, Y_{N_r}\}$ to be built in a progressive manner, i.e. $\{X_1, Y_1\}, \dots, \{X_1, \dots, X_{N_r}, Y_1, \dots, Y_{N_r}\}$, till the constraint in eq. (1) is verified with ϵ fixed 'a priori' — i.e., the precision required to the estimation is fixed. Since the original image I can be supposed to be composed of a finite number of 'homogeneous' regions (for instance 'grass', 'sky', 'sea', 'buildings', etc.), without lack of generality, we can consider only two regions, as in Fig. 2, and prove the following Proposition:

Prop. 3 Be $X = \begin{cases} X_1 & \text{with prob. } \alpha \\ X_2 & \text{with prob. } 1 - \alpha \end{cases}$ and $Y = \begin{cases} Y_1 & \text{with prob. } \alpha \\ Y_2 & \text{with prob. } 1 - \alpha, \end{cases}$ with $\alpha \in \mathbf{R}, 0 \leq \alpha \leq 1$, X_1 and X_2 disjoint variables (the same for Y_1 and Y_2), and let $Z = M(X, Y)$. By denoting with p_* is the pdf of the variable $*$, then

$$H(p_Z) \leq H(\alpha) + H(p_{M(X_1, Y_1)}) + H(p_{M(X_2, Y_2)}).$$

Proof is in Appendix. For a suitable cardinality (> 2) of the alphabet of M , $H(\alpha)$ (whose maximum is equal to 1) can be neglected and then the mixture leads to a lower entropy. Hence, in order to build the subsequence $\{X_1, \dots, X_{N_r}, Y_1, \dots, Y_{N_r}\}$ belonging to the typical set of M , it is more convenient to select them in local regions of the images I and J , rather than pointwise randomly in the whole image domain. In this way, we maximize $\{X_1, \dots, X_{N_r}, Y_1, \dots, Y_{N_r}\}$ entropy by minimizing, at the same time, its length $2N_r$. This theoretical result shows that the practical choice of a blockwise implementation of SSIM is really the most convenient: local isolated blocks better capture image information. It is not fortuitous that also HVS follows the same procedure (see for instance (Monte et al., 2005; Frazor and Geisler, 2006)): fixation points are foveated, since they depend on the local content of the fovea region, and they are the ones that maximize the entropy of the visual contrast in the fovea region.

2.4 Overlapping Blocks

According to previous results, next proposition, whose proof is in the Appendix, proves that non overlapping blocks maximize the entropy of the measure M (i.e. the variable Z).

Prop. 4 If Z_1, Z_2, \dots, Z_T are defined considering repetitions of X and Y 's samples while $\bar{Z}_1, \bar{Z}_2, \dots, \bar{Z}_R$ are the ones achieved by considering X and Y 's samples just one time, (i.e. they are independent and such that $R < T$), then

$$\frac{H(Z_1, Z_2, \dots, Z_T)}{T} \leq \frac{H(\bar{Z}_1, \bar{Z}_2, \dots, \bar{Z}_R)}{R} \quad T > R.$$

2.5 How to Find an A_M^ε Sequence

The objective of this paper is to go beyond the theoretical existence of A_M^ε . We want to identify at least ONE subsequence $\in A_M^\varepsilon$ with the least size (i.e. the least N_r) — and we want to do that with a low computational effort, if possible. Mathematically, if $L = 2N_r$, we ask for the existence of a subset of indices

$$\{i_1, \dots, i_L\} : \operatorname{argmin}_L |\Delta M| < \varepsilon, \quad (2)$$

with $\Delta M = \overline{M}(X, Y) - \overline{M}(x_{i_1}, \dots, x_{i_L}, y_{i_1}, \dots, y_{i_L})$. Though the search of $\{i_1, \dots, i_L\}$ may be performed with a more sophisticated and efficient strategy, in this paper it will be done by randomly selecting non overlapping blocks within the original image I and its distorted version J , in agreement with the theoretical observations of previous subsections. As it will be shown in the experimental results, this suboptimal criterion still allows to get satisfactory results.

3 EXPERIMENTAL RESULTS

The theoretical findings above have been validated on several images contained in different databases. In this section, only a representative subset of images will be considered. They are shown in Fig. 3 and belong to LIVE database (Sheikh et al.,). The latter is composed of 779 images having different amount of (five kinds of) distortion: Fast Fading, Gaussian Blur, JPEG2K, JPEG and Additive Gaussian Noise. In the sequel, we will first test the theoretical findings in Section 2. Then, based on these criteria, for each degraded image a subset of information is extracted for approximating the corresponding SSIM value.

Reduction of Information. To verify that not all I and J samples are really necessary to get a quite precise estimate of SSIM, Fig. 4 refers to Ocean image and its copy distorted by fastfading. Even though this work focuses on SSIM, Fig. 4 also contains results for SNR. This allows us to show that theoretical findings can be properly extended to other FR measures. 16×16 non overlapping blocks have been considered in all tests. They have been randomly selected in the original image and the corresponding ones have been extracted from the degraded image. In addition, the curves depicted in Fig. 4 have been normalized (divided by the corresponding maximum value) in order to design them on the same plot. Finally, a uniform quantization step with bins of width (Δ) equal to 10^{-4} has been used for storing IQA measure — this reduces the quantization distortion that is proportional to $\log(\Delta)$ (Cover and Thomas, 1991). In their first part, SSIM and SNR curves oscillate till they approach values close to the true ones. On the contrary, the corresponding entropy curves have an increasing trend with a critical curvature after which they tend to the entropy of the whole available sample (i.e. the true one). Both the fast ascending trend of entropies and the oscillating trend of SSIM and SNR stop in correspondence to quite the same point. It is worth outlining that the behavior plotted in Fig. 4 is common to all the analysed images, for each kind and level of distortion. These preliminary results give a clear evidence of the fact that it is possible to drastically reduce the information sent by the two sources I and J in order to assess quite precisely the visual quality of J , independently of the involved quality measure. In other words, the visual distortion typical set is composed of few samples of image pixels, in agreement with the Shannon's typical set. Fig. 4 also suggests that this reduced information can be easily found by measuring the entropy of the samples of the metric under study rather than the metric itself, because of the more regular entropy behaviour.

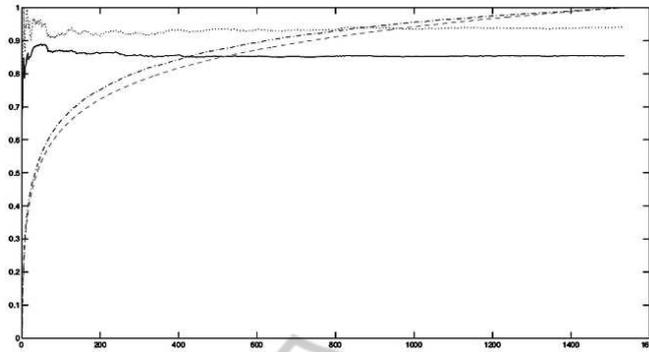


Figure 4: Ocean image and its fastfaded copy (img7_ocean in Table 1): SSIM (solid line), SNR (dotted), SSIM entropy (increasing dot-line) and SNR entropy (increasing dashed) curves versus the number of blocks.

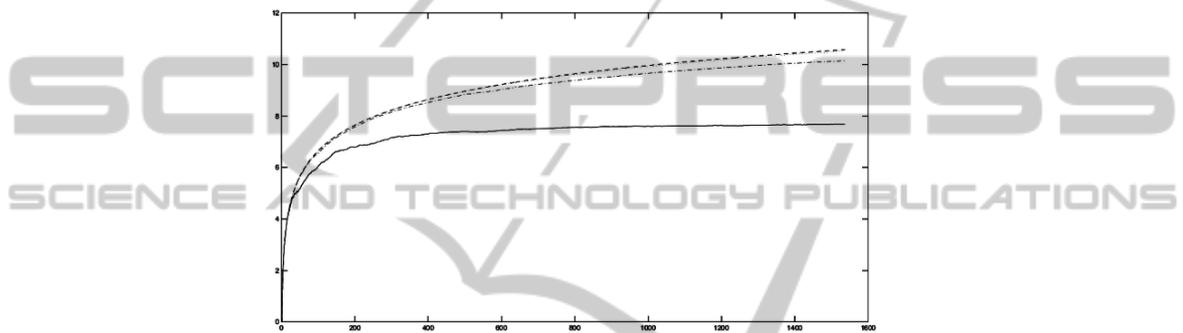


Figure 5: Ocean image and its Gaussian blurred copy (img57_ocean in Table 1). Entropy versus number of blocks (or an equivalent number of random pixels). (Bottom-up): SSIM entropy via random pixels (solid) and non overlapping blocks (dashdot), SNR entropy via random pixels (dotted) and non overlapping blocks (dashed).

Table 1: Images in Fig. 3: Entropy of: the original image ($H(X)$), the distorted image given the original one ($H(Y|X)$), SSIM with overlapping blocks ($H(Z)$) and SSIM for non overlapping blocks ($H(\bar{Z})$). For each kind of distortion the parameters used in LIVE database have been given: the standard deviation of the gaussian kernel for the gaussian blurring, the SNR of the distortion strength for fastfading, the quality score for jpeg and the standard deviation of the noise distribution for the white noise. More details in (Sheikh et al.,).

Original Image	Distorted Image	Distortion kind	$H(X)$	$H(Y X)$	$H(Z)$	$H(\bar{Z})$
ocean	img57_ocean	Gaussian blur (1.48)	7.1785	4.8630	5.0798	5.3230
stream	img58_stream	Gaussian blur (3.08)	7.4230	6.6357	5.9525	5.9997
lighthouse	img97_lighthouse	Gaussian blur (1.48)	7.3799	5.3012	5.1560	5.4239
sailing4	img127_sailing4	Gaussian blur (1.51)	6.8476	5.2000	5.0161	5.1369
ocean	img7_ocean	fastfading (18.9)	7.1785	4.7998	4.6848	4.8755
house	img73_house	fastfading (20.3)	7.1803	4.6494	4.0312	4.3385
stream	img100_stream	jpeg (0.29)	7.4230	6.4137	5.2691	5.5687
flowersonih35	img27_flowersonih35	jpeg (0.93)	7.7161	5.6465	3.4555	3.6996
ocean	img118_ocean.bmp	white noise (0.035)	7.1785	4.6308	4.2645	4.6762
house	img109_house.bmp	white noise (0.125)	7.1803	6.4434	6.1604	6.1460
flowersonih35	img72_flowersonih35	white noise (0.070)	7.7161	5.5789	3.6160	3.8698

Selection of the Best Reduction Domain. To test the results in Section 2.2, the entropy $H(X)$ has been computed for each image in Fig. 3. Bins width, necessary to build the empirical p.d.f., has been set equal to 1. With regard to the distorted image J , for each example the conditional entropy $H(Y|X)$ has been considered. The latter has been estimated from $H(X - Y)$, i.e. by looking at the distortion as an additive term — even though the process may be much more complicated. However, $X - Y$ really gives the difference of

information between the original image I and the distorted one J . The size of the bin width of $H(Y|X)$ has been set equal to 1, while 32×32 blocks have been used. The entropy $H(Z)$ of the SSIM vector M has been computed by quantizing M with a bin width equal to .01. Using these settings, Table 1 shows that $H(Z) < H(X) + H(Y|X)$. This behavior does not change for a different setting of parameters. Hence, SSIM naturally reduces entropy of the original images.

Locality of the Selected Information. Section 2.3 proves that it is more convenient to build FR quality measure samples using blocks rather than random pixels in I and J . Fig. 5 shows that SSIM entropy curve, that has been built using non overlapping and randomly selected blocks, always assumes values larger than those of SSIM entropy curve that has been built using randomly selected pixels. The same happens for SNR but the effect is strongly less visible: curves are very close to each other. In all tests, quantization bins have been set to 10^{-4} , even though different settings confirm the same trend.

Overlapping Blocks. Theoretical results in Section 2.4 state that it is more convenient to select non overlapping blocks from I and J rather than taking overlapping ones. A simple practical proof has been made by taking all possible (non overlapping) blocks from images in Fig. 3 and computing SSIM using them. Considering a bin width of .01 for this new vector of measures, the corresponding entropy $H(\bar{Z})$ has been computed. Tests have been performed using 32×32 blocks. The last column of Table 1 shows that non overlapping blocks lead to a higher entropy and then they convey a greater amount of information.

How to find a A_M^e Sequence. In order to manually estimate a sequence belonging to A_M^e , the sequence size has been fixed and the corresponding error has been measured. Table 2 contains the results achieved on Ocean image and its copies distorted by Gaussian blur, fastfading and white noise. Fig. 6 shows 100 randomly selected 16×16 blocks that have been used for the evaluation of SSIM. Both SSIM and SNR have been considered and the number of samples has been set equal to 100. On the contrary, the size of the blocks on I and J has been changed: 8×8 , 16×16 and 32×32 . The blocks on I (and the corresponding in J) have been randomly selected. That's why the results in Table 2 have been achieved as average on 30 trials, i.e. 30 different choices of 100 non overlapping blocks, in order to get a more fair evaluation of the estimation error for the considered IQA metric. Specifically, the mean and the standard deviation of the the relative error

$$\frac{|\bar{M} - \hat{M}|}{\bar{M}} \quad (3)$$

has been computed. For each image and each distortion kind, the estimation errors always are smaller than 5% of the true IQA measure. Table 3 contains the results obtained using different sizes for the selected sequence of FR values, respectively 50, 100 and 200 blocks on both the original I and the distorted image J . Again, apart from just one case (in bold) errors are always under 5%, confirming that few blocks (a little part of the available information) are required to give



Figure 6: 100 Randomly selected 16×16 blocks used for the estimation of SSIM of Ocean image.

a good estimate of the involved FR measure.

The speed up obtained in the computation of the considered image quality assessment metric on the typical set depends on the number of blocks that are used for the evaluation of SSIM. The computational gain is $G = \frac{N_0}{N_r}$, where N_0 is the number of blocks that are used for the computation of SSIM in the whole image (using the standard algorithm), while N_r is the number of blocks belonging to the *visual distortion typical set* i.e. the reduced set of blocks from which it is possible to get a quite precise estimation of SSIM. For example, if an image can be partitioned into 1536 non overlapping blocks, the gain using just 50, 100 or 200 non overlapping blocks respectively is 30.72, 15.36 and 7.68. This gain increases if the image is partitioned into non overlapping blocks, while it decreases for images having small dimension. For example, for images composed of 1280 non overlapping blocks the gain becomes 27, 13.5 and 6.75 respectively for 50, 100 or 200 non overlapping blocks in the visual distortion typical set.

4 CONCLUSIONS

The paper has presented a study concerning the definition of a *visual distortion typical set* in agreement with the general concept of asymptotic equipartition property and the neurological studies on those points that attract human attention in the early vision. General criteria for the characterization and the practical definition of this typical set have been given. The typical set has been used for reducing the amount of information necessary to assess the quality of an image using a standard full reference image quality assessment measure. This typical set makes the FR IQA metric less computational demanding. Achieved results are very encouraging since they are robust to changes of image subject, parameters settings and distortion kinds. Future research will focus on designing an optimized procedure to find the minimum budget of information for a fixed error.

Table 2: Ocean image and its copy distorted by respectively Gaussian blurring (*top*), fastfading (*middle*) and white noise (*bottom*). Mean and standard deviation (in the brackets) of the relative error (%) for the estimation of the true value of SSIM and SNR using just 100 blocks with size equal to: 8×8 , 16×16 and 32×32 .

Block size = 8 /No. Blocks = 100		Block size = 16 /No. Blocks = 100		Block size = 32 /No. Blocks = 100	
Total blocks: 6144		Total blocks: 1536		Total blocks: 384	
SSIM error	SNR error	SSIM error	SNR error	SSIM error	SNR error
1.84% (1.41%)	2.46%(2.14%)	1.25%(1.01%)	2.43%(1.63)%	1.40%(1.20%)	2.2% (1.72%)
1.33%(1.42%)	2.75%(2.00%)	1.39%(0.80%)	2.50%(1.90%)	0.79%(0.55%)	2.25%(1.71%)
1.14%(0.87%)	1.31%(1.03%)	1.09%(0.20%)	1.07%(0.70%)	0.71%(0.54%)	0.81%(0.57%)

Table 3: Images in Fig. 3 with different kinds of distortions. Mean value and standard deviation (in brackets) of the relative error (%) as in eq. (3) for the estimation of the true value of SSIM and SNR considering just 50, 100 and 200 non overlapping blocks with size equal to 16×16 .

Image	Dist	Blks	Blk size:16 / Sel. Blks:50		Blk size:16 / Sel. Blks:100		Blk size:16 / Sel. Blks:200	
			SSIM error	SNR error	SSIM error	SNR error	SSIM error	SNR error
Ocean	Gb	1536	2.53% (2.49)	4.15% (2.70)	1.25% (1.01)	2.43% (1.63)	1.16%(1.03)	1.82% (1.18)
Stream	Gb	1536	3.50%(3.00)	2.65%(2.25)	3.45%(3.01)	2.29% (1.80)	2.43%(2.01)	1.66%(1.02)
Lighth.	Gb	1350	2.45%(2.20)	3.72%(2.01)	2.30%(1.80)	2.39%(2.22)	1.56%(1.21)	1.75%(1.15)
Sail4	Gb	1536	2.13%(2.00)	3.53%(2.99)	1.47%(0.80)	3.26%(2.22)	1.02%(0.45)	1.47%(0.89)
Ocean	Ff	1536	1.53% (1.20)	3.18%(2.05)	1.39%(0.80)	2.50%(1.90)	0.83%(0.52)	1.86%(1.01)
House	Ff	1536	1.08%(0.99)	2.68%(1.90)	0.75%(0.57)	1.73%(1.60)	0.55%(0.21)	1.36%(0.94)
Stream	Jp	1536	2.93%(2.90)	2.78%(2.09)	1.95%(1.50)	1.84%(1.30)	1.56%(0.98)	1.11%(0.80)
Flower	Jp	1280	0.58%(0.50)	4.01%(3.03)	0.30%(0.15)	2.62%(2.15)	0.23%(0.09)	1.89%(1.01)
Ocean	Wn	1536	1.19%(0.70)	1.29%(0.90)	1.09%(0.80)	1.07%(0.70)	0.61%(0.30)	0.72%(0.31)
House	Wn	1536	5.46% (2.09)	2.04%(1.77)	4.40%(2.87%)	1.93%(1.35)	2.72%(2.51)	0.93%(0.80)
Flower	Wn	1280	2.82%(2.77)	1.45% (1.35)	1.86%(1.52%)	0.90%(0.50)	1.09%(1.08)	0.56% (0.53)

REFERENCES

Benabdelkader, S. and Boulemden, M. (2005). Recursive algorithm based on fuzzy 2-partition entropy for 2-level image thresholding. In *Pattern Recognition*, 38. Elsevier.

Bruni, V., Rossi, E., and Vitulano, D. (2013a). Jensen-shannon divergence for visual quality assessment. In *Signal Image and Video Processing* 7(3). Springer.

Bruni, V., Vitulano, D., and Ramponi, G. (2011). Image quality assessment through a subset of the image data. In *Proc. of ISPA 2011*. IEEE.

Bruni, V., Vitulano, D., and Wang, Z. (2013b). Special issue on human vision and information theory. In *Signal Image and Video Processing* 7(3). Springer.

Cover, T. M. and Thomas, J. A. (1991). *Elements of Information Theory*. John Wiley & sons.

Frazor, R. and Geisler, W. (2006). Local luminance and contrast in natural in natural images, 46. In *Vision Research*.

Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing*. Prentice Hall, 2nd edition.

Grunwald, P. D. (2004). A tutorial introduction to the minimum description length principle. In *Advances in Minimum Description Length: Theory and Applications*. Myung Grunwald, Pitt.,

Monte, V., Frazor, R., Bonin, V., Geisler, W., and Corandin, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system 8(12). In *Nature Neuroscience*.

Rivera, M., Ocegueda, O., and Marroquin, J. L. (2007). Entropy-controlled quadratic markov measure field models for efficient image segmentation. In *IEEE Trans. on Image Proc.* 16(12).

Sheikh, H. R. and Bovik, A. C. (2006). Image information and visual quality. In *IEEE Trans. on Image Proc.*, 15(2).

Sheikh, H. R., Bovik, A. C., and Veciana, G. D. (2005). An information fidelity criterion for image quality assessment using natural scene statistics. In *IEEE Trans. on Image Proc.*, 14(12).

Sheikh, H. R., Wang, Z., Cormack, L., and Bovik, A. C. Live image quality assessment database release 2. [Online]. Available: <http://live.ece.utexas.edu/research/quality>.

Wang, W., Wang, Y., Huang, Q., and Gao, W. (2010). Measuring visual saliency by site entropy rate. In *Proc. of CVPR10*. IEEE.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004a). Image quality assessment: From error visibility to structural similarity. In *IEEE Trans. on Image Proc.*, 13.

Wang, Z. and E.P.Simoncelli (2005). Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. In *Proc. of Human Vision and Electronic Imaging X*, vol. 5666. SPIE.

Wang, Z. and Li, Q. (2011). Information content weighting for perceptual image quality assessment. In *IEEE Trans. on Image Proc.* 20(5).

Wang, Z., Lu, L., and Bovik, A. (2004b). Video quality as-

assessment based on structural distortion measurement. In *Signal Processing: Image Communications*, 19(2).
 Winkler, S. (2005). *Digital Video Quality, Vision Models and Metrics*. Wiley.
 Zhang, D. and Jernigan, E. (2006). An information theoretic criterion for image quality assessment based on natural scene statistics. In *Proc. of ICIP 2006*. IEEE.

APPENDIX

Proof 1 1. $\mu_{n+1} - \mu_n = \sum_{x \in \mathcal{X}} x(p_{n+1}(x) - p_n(x)) = \sum_{x \in \mathcal{X}} x \left(\frac{x}{n+1} p_n(x) + \frac{x_{n+1}}{n+1} - \frac{x}{n+1} \right) = \frac{1}{n+1} (x_{n+1} - \mu_n)$. The sign of the difference between two successive mean values depends on x_{n+1} and the convergence is not monotonic. 2. $|\mu_n - \mu|^2 = \left| \sum_{x \in \mathcal{X}} x(p_n(x) - Q(x)) \right|^2 \leq M_n^2 V^2(p_n, Q)$ where $V(p_n, Q)$ is the variational distance between p_n and Q i.e., $V(p_n, Q) = \sum_{x \in \mathcal{X}} |p_n(x) - Q(x)|$. Since the Kullback-Leibler divergence $D_{KL}(p_n || Q) = \sum_x p_n(x) \log \left(\frac{p_n(x)}{Q(x)} \right)$ is such that $D_{KL}(p_n || Q) \geq \frac{1}{2} V^2(p_n, Q)$, then $|\mu_n - \mu|^2 \leq 2M_n^2 D_{KL}(p_n || Q)$. Hence, for n : $D_{KL}(p_n || Q) \leq \frac{\epsilon}{2M_n^2}$, $\epsilon > 0$, then $|\mu_n - \mu|^2 \leq \epsilon$.

Proof 3 $H(X) = H(\alpha) + \alpha H(X_1) + (1 - \alpha)H(X_2)$ (Cover and Thomas, 1991) and the same holds for $H(Y)$. Since $Z = M(X, Y)$, then $Z = \begin{cases} M(X_1, Y_1) & \text{with prob. } \alpha \\ M(X_2, Y_2) & \text{with prob. } 1 - \alpha \end{cases}$ where $M(X_1, Y_1)$ and $M(X_2, Y_2)$ are not disjoint. $Z \sim p_z$ where $p_z = \alpha p_{M(X_1, Y_1)} + (1 - \alpha) p_{M(X_2, Y_2)}$. Hence, $H(p_z) \leq H(\alpha) + \alpha H(p_{M(X_1, Y_1)}) + (1 - \alpha)H(p_{M(X_2, Y_2)})$. In fact, let's suppose that

$$H(p_z) > H(\alpha) + \alpha H(p_{M(X_1, Y_1)}) + (1 - \alpha)H(p_{M(X_2, Y_2)}) \quad (4)$$

and let us consider the Jensen-Shannon divergence $D_{JS}^\alpha(p_{M(X_1, Y_1)} || p_{M(X_2, Y_2)}) = H(p_z) - \alpha H(p_{M(X_1, Y_1)}) - (1 - \alpha)H(p_{M(X_2, Y_2)})$, then $D_{JS}^\alpha(p_{M(X_1, Y_1)} || p_{M(X_2, Y_2)}) > H(\alpha)$, that is absurd since $0 \leq D_{JS}^\alpha(p_{M(X_1, Y_1)} || p_{M(X_2, Y_2)}) \leq H(\alpha)$. Since $H(p_{M(X_1, Y_1)}) \leq H(p_{(X_1, Y_1)})$ and $H(p_{M(X_2, Y_2)}) \leq H(p_{(X_2, Y_2)})$ we have $H(p_z) \leq H(\alpha) + \alpha H(p_{(X_1, Y_1)}) + (1 - \alpha)H(p_{(X_2, Y_2)})$.

Proof 4 Let $\hat{Z}_j = \{Z_1, Z_2, \dots, Z_{N_j}\}$ be a collection of N_j variables Z_i selected in $\{Z_1, Z_2, \dots, Z_T\}$ and let be K the number of possible N_j -ples such that: $\bigcup_{j=1}^K \hat{Z}_j = \{Z_1, Z_2, \dots, Z_T\}$. Since for generic variables S_1, \dots, S_n , it holds $H(S_1, S_2, \dots, S_n) \leq \sum_{i=1}^n H(S_i)$, then $H(Z_1, Z_2, \dots, Z_T) \leq \sum_{j=1}^K H(\hat{Z}_j) \leq \sum_{j=1}^K H(\bar{Z}_1, \bar{Z}_2, \dots, \bar{Z}_R) = KH(\bar{Z}_1, \bar{Z}_2, \dots, \bar{Z}_R) = KR \frac{H(\bar{Z}_1, \bar{Z}_2, \dots, \bar{Z}_R)}{R}$.