From Sparse Coding Significance to Perceptual Quality: A New Approach for Image Quality Assessment

Ayyoub Ahar, *Student Member, IEEE*, Adriaan Barri, *Member, IEEE*, and Peter Schelkens, *Member, IEEE*

***Abstract*— An increasing number of image processing applications require an automated quality prediction of the visual content as perceived by humans. Since, sparse coding is suggested to be an underlying strategy of the brain's neural system, it would be logical to assume that specific tasks like quality assessment also attempt to adhere to this strategy. However, existing perceptual quality predictors, often mimicking the different stages of the human visual system and deploying machine learning strategies, such as neural networks, rarely integrate the concept of sparse coding in their design. In this paper, we first investigate the validity of such assumption by performing an empirical analysis on the relation between the structural information of the scene—captured via sparseness significance—and perceptual quality. Subsequently, we propose a new approach to integrate the significance of sparse coding features in the future imagequality measure (IQM) designs. We utilize the Fourier transform as a case study, which leads to a new IQM called sparseness significance ranking measure (SSRM). This measure essentially deploys a Fourier basis for sparse coding, a ranking mechanism based upon the amplitudes of the sparse coefficients and subsequently a complex correlation metric that assesses the correspondence between the ranked coefficient amplitude profiles of the reference and the distorted image. Moreover, we introduce a new methodology, namely separation ratio analysis, to assess the prediction quality of individual features or quality predictors given a target perceptual quality. The quality predictions by the proposed SSRM show excellent compatibility with perceptual quality scores. A set of routine benchmarking experiments utilizing the LIVE and CSIQ, IVC and TID2008 databases indicates a highly competitive performance with state of the art IQMs. Moreover, it delivers this performance at a low computational cost.**

***Index Terms*— Image quality assessment, structural information, perceptual quality predictors, human visual system, sparse coding, Fourier analysis, sparseness significance, SSRM.**

Manuscript received January 27, 2017; revised July 24, 2017 and October 27, 2017; accepted November 2, 2017. Date of publication November 8, 2017; date of current version November 28, 2017. This work was supported in part by the European Research Council through the European Union's Seventh Framework Programme under Grant FP7/2007- 2013 and in part by ERC Grant Agreement under Grant 617779 (INTERFERE). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Kalpana Seshadrinathan. *(Corresponding author: Ayyoub Ahar.)*

The authors are with the Department of Electronics and Informatics, Vrije Universiteit Brussel, B-1050 Brussels, Belgium, and also with IMEC, B-3001 Leuven, Belgium (e-mail: aahar@etrovub.be; abarri@etrovub.be; peter.schelkens@vub.ac.be).

This paper has supplementary downloadable material available at http://www.erc-interfere.eu/downloads, provided by the author. The material includes MATLAB code for SSRM. Contact Ayyoub Ahar (aahar@etrovub.be) for further questions about this work.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2017.2771412

I. Introduction

RATING visual quality of digital images is a crucial step for many image processing applications and more particularly image compression. Two popular approaches exist to tackle the quality assessment problem in the literature. Although both of them are more or less inspired by the Human Visual System (HVS), they refer to specific paradigms regarding how the quality of the input visual content is processed and rated. One approach is mainly related to the error sensitivity paradigm, where the visual quality is assumed to be rated by the error signal. Over the past few decades, numerous objective quality measures have been introduced following this approach where one can consider the Minkowski-based distance metrics family (e.g. MSE, PSNR...) as a simple example. More sophisticated methods of this category either use the available computational models of neurons in the primary visual cortex or use transformation-based spatial frequency decompositions to mimic the neurons' response to certain visual stimuli. However, a typical challenge in this approach would be to differentiate between various types of errors, which may have the same value but a different degree of visibility, and can as such lead to misleading perceptual quality predictions.

The alternative approach that received more attention since the introduction of the Universal image Quality Index (UQI) [1] and the Structural SIMilarity index (SSIM) [2], refers to the widely accepted hypothesis that the HVS concentrates on the essence of the scene by essentially acquiring the structural information to better handle the enormous amount of received visual stimuli. This is compatible with the hierarchical structure of the HVS, where it performs a set of interconnected real-time pre-processes taking place in the retina, the lateral geniculate nucleus and different layers of primary visual cortex in order to enable extracting complex object level features (e.g. in extrastriate areas such as the inferior temporal cortex) [3]. Following this hypothesis, many contemporary IQMs inspired by SSIM have been proposed to measure changes in structural information. Among them Multi Scale-SSIM(MS-SSIM) [4] has been successfully utilized in television and broadcasting industry. Despite the fact that the notion of structural information and the way it should be acquired is still vague and open for discussion, these methods, which mainly consider the remaining information after separating contrast and luminance dependencies as structural information and the spatial cross-correlation as its measure,

proved to be very efficient in the prediction of visual quality of digital images. Another set of IQMs in this category use gradient magnitude comparisons of local image patches as a measure of structural changes. A recently proposed technique calculates for example the Gradient Magnitude Similarity Deviation (GMSD) [5]. Another way to measure changes in structural content is a combination of phase congruency and gradient magnitude, as used by the Feature SIMilarity (FSIM) method [6].

Apart from the mentioned approaches, Larson and Chandler noticeably proposed a combination of dual strategies differentiating between the way low and high quality images should be treated [7]. Another alternative is to measure the quality by finding the mutual information between the reference and distorted image. This idea was first applied by Sheikh et al. to design the Visual Information Fidelity (VIF) index [8]. The VIF measures mutual information between the reference and distorted image using a Gaussian mixture model of wavelet coefficients in different subbands. Some frameworks specifically address the colour perception characteristics of the HVS [9]. Apart from Full-Reference methods, a number of Reduced Reference (RR) quality measures have been recently introduced, among them some prominent examples [10]–[14]. These methods essentially provide quality judgement based on a limited set of features representing the general image characteristics in combination with a sensitivity to a particular — often application dependent — set of distortion types. Nonetheless, RR methods are out of scope of this research and will not be analyzed further.

A. Sparse Coding and Perceptual Quality: Motivation

To smoothly process the continuous stream of visual stimuli projected on the retina, the HVS relies on abstraction processes happening in early stages of the visual system that essentially decrease the redundancy of the input signal. The underlying neural behaviour of these processes have been extensively investigated in literature [3], [15]–[17]. Numerous experimental evidences referring to the activity of different neurons in the brain imply that not only the visual cortex but also other parts of brain corresponding to different sensory modalities follow a specific strategy to reduce mentioned redundancies. This strategy can be modelled by sparse coding. It is suggested that utilizing such sparse patterns of neural activity enables an efficient means of representing data found in the natural world. Moreover, sparse coding provides a means of efficiently forming associations and storing memories, and it achieves all of this with relatively small amounts of energy [18]. So, if we assume sparse coding as one of the general underlying strategies in our brain [19], then any neural activity implemented for specific tasks including quality assessment should be compatible with this main framework.

From a signal processing point of view, such hypothesis is also a very attractive choice especially for signal compression methods where accurate perceptual quality assessment is mainly considered as an extra processing burden, such that in most of the popular compression methods, it is fundamentally downgraded into very simplistic error sensitivity routines based on MSE. In a nutshell, the primary assumption in most compression schemes is to give more weight to features that contribute more significantly to the perceived reconstruction quality. Therefore an efficient perceptual quality assessment method that is consistent with the core assumptions of compression frameworks and that can be integrated in the encoder is highly desirable. Despite the fact that such strong motivations are available, less research has been devoted to pursue the relation between quality/distortion and the importance of features in terms of sparsity. An exception is the Sparse Feature Fidelity (SFF), a machine learning based IQM that uses Independent Component Analysis (ICA) to find the most sparse representation of a received image using small image patches as building blocks [20]. Another exception is [21], introducing a No-Reference quality measure called SNRSS, where a subset of images was selected to extract their natural scene statistics in wavelet domain and to subsequently identify a sparse representation of extracted features by conducting a training session against DMOS scores. However, dimensionality reduction techniques are not the only way to obtain a sparse representation. Our aim is to design a method that can work for any sparse image representation. Moreover, to avoid biasing the design of the IQM and to facilitate reproducibility, we avoid deploying machine learning techniques.

Due to the availability of fast implementations and its vast application for many different types of data modalities, we focus on the well-known Fourier transform to sparsely represent images in our current analysis and IQM design. It is also known that the analysis of image redundancies, specifically the ones related to the second order statistics, can be most easily performed by means of a Fourier spectrum analysis. Explaining correlation between two points in an image as a function of their distance and orientation [22] is directly related to the autocorrelation of this image. Although the complete set of Fourier sinusoids may not yield the most sparse representation in the frequency domain, it is a good choice to demonstrate and validate our approach.

A natural way of comparing two images in the Fourier domain would be to directly measure the difference between the amplitudes and phases of their sinusoid components. Such an approach would of course not consider any of the HVS characteristics and nonlinearities, such as contrast sensitivity and masking effects. Exploiting these effects, in [23] also a Fourier transform based IQM was proposed where a non-uniform binning of frequency components is utilized. After dividing the image into non-overlapping blocks, the 2D DFT coefficients are grouped into different categories. The lower frequency components are not binned at all, hence end up in separate categories. The higher frequency components are categorized in growing bin sizes, where each bin is represented by its arithmetic mean.

Normally no other explicit distinction other than masking effects and contrast sensitivity has been imposed on high and low frequency components. Consequently, here we analyze the possibility of categorizing spatial frequency components by considering another HVS characteristic namely the importance of the frequency components with respect to the sparsity criterion. Nonetheless, it has to be noted that the proposed

approach does not exclude the possibility to integrate these principles as well. For example, CSF can be immediately exploited in Fourier domain by scaling the transform coefficients with their corresponding CSF weights before applying the sorting algorithm. Also, texture and luminance masking can be envisaged to be exploited by deploying a windowed Fourier transform, which will enable the local tuning of the frequency weights based upon spatially local texture complexity and background luminance measurements. Likewise, other transforms that support a better balancing between spatial/frequency domain localisation can be deployed. However, these improvements are subject of future research.

B. Proposed IQA Solution

In this research, we address the quality assessment problem following a psychovisually-inspired 'sparse coding' strategy that does not require training of a machine learning algorithm. To practically verify the compatibility of our assumptions compared to perceptual quality, we propose a new IQM, which is solely designed based on the sparsity hypothesis. The proposed method utilizes a spatial frequency transformation to mimic the neural behaviour, and measures structural information by appropriately weighting the coefficients along the frequency components. To do so, a ranking statistical approach is utilized to differentiate the most and least sparsifying image features in Fourier domain and then a combination of complex correlation coefficients and relative differences will be calculated to quantify the overall quality of the distorted image.

Summarized this paper brings the following novelties: (1) we introduce the exploitation of sparse coding in the context of full reference visual quality assessment, (2) we exploit a complex amplitude representation for categorizing the significance of the perceptual quality predictors, i.e. sparse features, (3) we propose a methodology based on separation ratio analysis that allows for testing the responsiveness of perceptual quality predictors, given a particular level of perceptual quality of the stimuli and (4) we propose a new quality metric based on the above methodologies that proves to be competitive with state of the art solutions at a low computational cost.

The remainder of this paper is structured as follows. In section II, we will discuss ranking the spatial frequency features of digital images according to their significance for sparse coding and demonstrate their relation with the structural content of images. In section III, the proposed IQM is described. In section IV, a separation ratio analysis is performed to investigate the accuracy of partial quality estimators produced by categorizing features based on their sparseness significance. Also we provide a set of comparative results between the proposed method and a large set of the state-of-the-art and well-known IQM algorithms tested on standard annotated databases. In section V, we conclude the paper and summarize the contributions.

II. Sparsity Significance and Ranking Analysis

A. Ranking Analysis of Fourier Space

Utilizing Fourier sinusoids as a representative structured dictionary for sparse coding, our objective is to reorder the Fourier coefficients based on rank of their complex amplitudes, creating a set of signal approximations from subgroups of Fourier components and then investigating the correlation between such approximations and the perceptual quality. In the following paragraphs, we will explain how this ranking organization in subgroups - formalized in (5) - can be obtained starting from the concept of sparse coding.

Sparse coding comes down to finding a linear combination of a set of basis functions ϕ_i, $i = 1, 2, \ldots, M$, using weights w_i of which only a few are far from zero [24]. By sparse coding, we want to find a small subset of weights $\{w_{R(j)} | j = 1, \ldots, k\}$ with $k \ll M$ such that $f(x, y)$ can be approximated by

$$f(x, y) \approx \sum_{j=1}^{k} w_{R(j)} \phi_{R(j)} \tag{1}$$

$R(j)$ is a function that maps the ranking indices j to the corresponding indices i of the non-ranked weights, i.e. $i = R(j)$, whereas $j = 1$ points to the coefficient that has the largest magnitude.This image approximation is restricted to the most sparsifying components, i.e. the components with the largest weights. However, if we do not discard the remaining components, $f(x, y)$ can be reconstructed:

$$f(x, y) = \sum_{j=1}^{k} w_{R(j)} \phi_{R(j)} + \sum_{j=k+1}^{M} w_{R(j)} \phi_{R(j)}. \tag{2}$$

Assuming that the weights $w_{R(j)}$ are decreasing with j, we can repeat the subgroup decomposition in (2) for a set of thresholds $\{k_l | l = 1, \ldots, m\}$, which categorizes the representation into $m + 1$ subgroups ordered by their sparsity significance:

$$f(x, y) = \sum_{j=1}^{k_1} w_{R(j)} \phi_{R(j)} + \sum_{j=k_1+1}^{k_2} w_{R(j)} \phi_{R(j)} + \ldots + \sum_{j=k_m+1}^{M} w_{R(j)} \phi_{R(j)}. \tag{3}$$

Given a maximum allowed error ϵ, the above representation can be used to select the smallest value $k_\epsilon \in \{k_1, k_2 \ldots, k_m\}$ for which $\left\| f(x, y) - \sum_{j=1}^{k_\epsilon} w_{R(j)} \phi_{R(j)} \right\| \leq \epsilon$. When $\epsilon = 0$, this becomes a coarse discretization of the ℓ_0-sparsity minimization problem described in [25]. In the context of this paper, we consider a sparse coding system using a fixed dictionary based on the Fourier sinusoids, but the same procedure can be generalized to other structured or adaptive dictionaries. It is obvious that the most and least sparsifying Fourier coefficients will be the ones with the biggest and smallest amplitudes respectively. Normally, a logical step would be to check the distribution of amplitudes based on their spatial frequency coordinates (u, v) to utilize it as a prior knowledge for their categorization. Nonetheless, that will not be an option, because by looking into the provided models in the literature e.g. in [22], it turns out that the distribution of individual images may significantly deviate from such models both in shape and total power depending on the content of the scene. Among several others, few natural scene image samples shown in Figure1 clearly demonstrate such deviations.

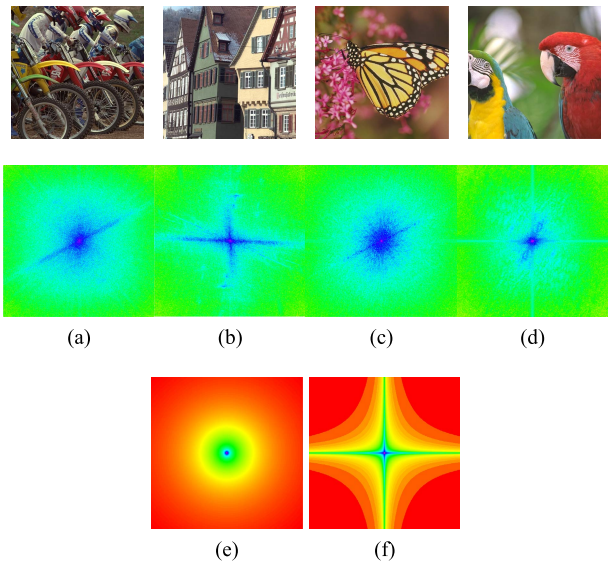


Fig. 1. Figures (a) to (d): sample images selected from the LIVE still image database and their corresponding logarithmic spectral amplitude distribution with values scaled into the range of 0 to 255. (e) shows the circularly averaged logarithmic amplitude distribution model scaled to the same range, extracted from the spectral power distribution model. (f) also depicts the amplitude distribution calculated from the spectral power model with orientation considerations.

Instead, we will need a highly adaptive method for this purpose able to conform to the content of every single image to be analyzed. To model the content-dependency of their distribution, we rank the coefficients' amplitudes and then assign coefficients into separate bins based on their rank (e.g. their amplitude value). This can be considered as non-uniform quantization of the ranks where the step size is not fixed. The variable quantization step sizes can be elegantly described using the quantile function. A quantile function of a probability distribution is the inverse of its cumulative distribution function (CDF). Given a random variable X with a strictly monotonic CDF f, the quantile function Q_f assigns to each probability p the value x for which $P(X \geqslant x) = p$. Assume X is a discrete variable and let $\{x_k | k = 1 \ldots, m\}$ represent the possible outcomes of X ranked from high to low. When $p_k = P(X = x_k)$ then Q_f is a step function with step jumps in $q_k = \sum_{i=1}^{k} p_k$ and

$$Q_f(q_k) = x_k \tag{4}$$

Since the ranked amplitudes of the Fourier coefficients are also discrete and monotonic, the amplitude range can be divided to a finite set of n intervals $(p^{kth}, p^{k+1th}]$, with equal number of coefficients existing in each interval. Given a 2D periodic, discrete signal $f(x, y)$ (e.g. digital image), from (3) and (4) for the N quantiles of Fourier amplitudes the proposed scheme can be formalized into:

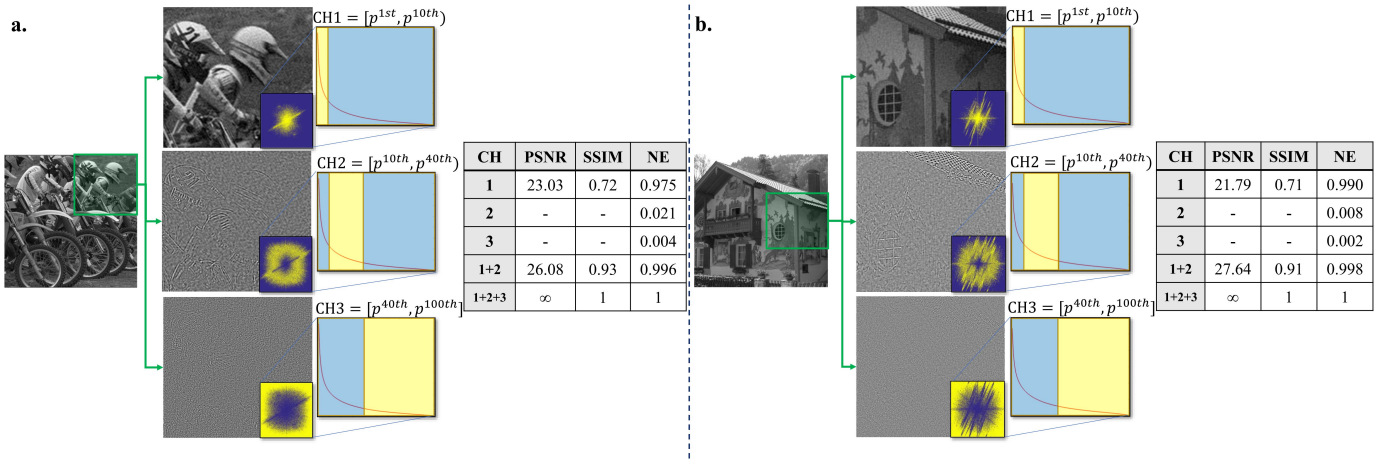
$$f(x,y) = \sum_{j=1}^{p^{1st}} wf_{R(j)}\phi f_{R(j)} + \sum_{j=p^{1st}+1}^{p^{2nd}} wf_{R(j)}\phi f_{R(j)} + \ldots + \sum_{j=p^{(k+1)th}+1}^{N} wf_{R(j)}\phi f_{R(j)} \tag{5}$$

where wf and ϕf, correspond to the complex Fourier amplitudes and Fourier bases respectively. The value for n can vary between 1 up to the ordinal number of distinct amplitude values, which itself will always be smaller or equal to the number of coefficients. This allows for adjusting the resolution of the analysis and choosing a smaller N to decrease the computational cost, or larger value to increase the accuracy of the process. Adjusting the resolution of analysis, i.e. choosing value of N directly depends on the spectral distribution of testing data. The higher the spectral entropy of the data, the higher resolution required to accurately characterize its behaviour. For the further analysis and discussion in this paper, we consider $N = 100$, a rather high resolution seen the typical spectral distribution of the test imagery. Although, our further experiments explained in section IV-A demonstrates that for the test imagery such a resolution is probably overkilling – and $N = 10$ would probably suffice – we have observed experimentally larger N would result in equal metric performance. Algorithm (1) formalizes the overall procedure stemmed from (5), which used in this research for the quality assessment of digital images. It should be noted that the current procedure should not be mixed up with the popular pooling strategy, where typically from a pool of locally calculated quality scores the ones showing the highest impairment (i.e. lowest 5%) are selected [26] to represent the higher sensitivity of HVS for highly distorted regions. In [27] among others, it is argued that HVS will penalize such images more heavily. Instead, we are using the quantile categorization as a strategy to differentiate between Fourier coefficients based on their amplitude values, which themselves represent discrete levels for sparsity significance of the coefficients.

B. Structural Information via Sparseness Significance

So far, we have designed a relatively simple and fast way to discriminate between most and least sparsifying components in Fourier domain. It is also known that sparsity is meant to extract "higher orders of statistical dependencies" which cannot be simply described at the image pixel level [17], [19], [28]. The relation between sparse coding and high-order correlations in visual cortex is also experimentally observed and reported in [29]. Here, we suggest that such correlations and dependencies, which mainly encode information about available oriented lines, edges, curves and textures in the scene, correspond to building blocks of observed object structures, hence "structural information" of the scene. Figure 2 shows two examples, where after calculation of amplitude percentiles, three hard threshold filters were created that only selected a certain number of coefficients based on their sparseness significance plus DC component and the rest were replaced with zero. The selected intervals includes $[p^{1st}, p^{10th})$, $[p^{10th}, p^{40th})$ and $[p^{40th}, p^{100th}]$ respectively. An inverse Fourier transform was applied after filtering to reconstruct the visual information each set of coefficients was carrying.

It is important to realize that the filters corresponding to the first few percentiles – the first 10 in the example – discard some low frequency components (due to their lower



CH	PSNR	SSIM	NE
1	23.03	0.72	0.975
2	-	-	0.021
3	-	-	0.004
1+2	26.08	0.93	0.996
1+2+3	∞	1	1

CH	PSNR	SSIM	NE
1	21.79	0.71	0.990
2	-	-	0.008
3	-	-	0.002
1+2	27.64	0.91	0.998
1+2+3	∞	1	1

Fig. 2. Demonstration of the relation between sparsity significance and structural information of the scene. For illustration purposes, three hard threshold filters, selecting coefficients from a particular percentile range, are applied in Fourier domain. Subsequently, the images are reconstructed with an inverse Fourier transform to visualize the visual information encoded within a particular percentile range. Three close-up views of the reconstructed images based upon the percentile ranges $CH1 : [p^{1st}, p^{10th})$, $CH2 : [p^{10th}, p^{40th})$ and $CH3 : [p^{40th}, p^{100th}]$ are shown. The yellow colour indicates the selected percentiles and blue represents omitted percentiles both in the filtered Fourier spectra and the percentile plots. PSNR, SSIM and Normalized Energy (NE) values are shown per channel as a quantitative representation of signal to noise ratio, structural information and contained normalized energy.

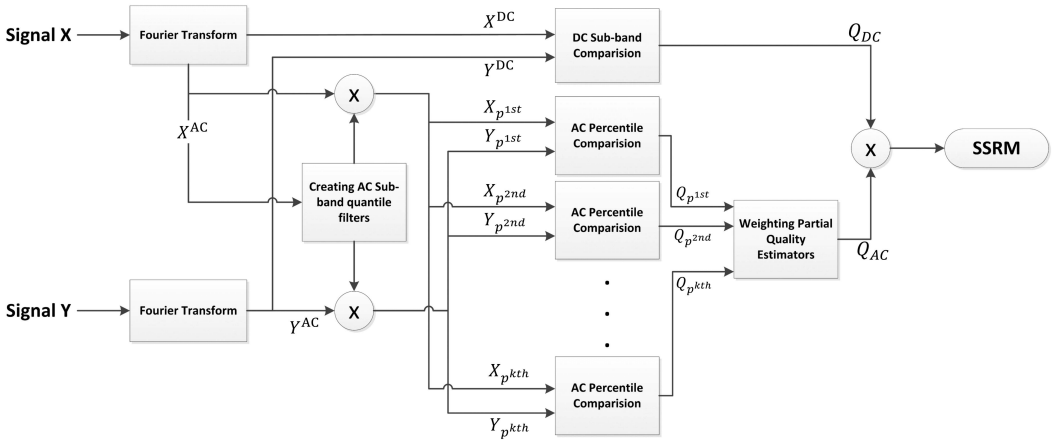


Fig. 3. Diagram of Sparseness Significance Ranking Measure system.

importance in terms of sparsity), while mid and high frequency components which are more significant in terms of sparsity are selected instead. The reconstructions from the second interval in the example, i.e. $[p^{10th}, p^{40th})$, demonstrate that these contain some fine-scale details especially around highly contrasted edges and patterns (e.g. white helmets and the dark background in case Figure 2.a and the roof tiles and window in Figure 2.b). The amount of encoded visual information and their perceptual significance in the second and third intervals are though not comparable with the first one in the sense of realizing the overall structure of the scene. This observation implies that comparison results coming from those should also be significantly under-weighted. Nonetheless, their role especially for the cases of comparing images with very high perceptual quality is yet undeniable. In such cases [7], instead of looking for the overall structural changes, the HVS utilizes an alternative strategy seeking for the small distortions in the presence of high structural similarities between reference and the test image. While the most sparsifying components (i.e. main structural information) will be less affected from relative perspective, it is the comparison of the coefficients in the last intervals that will let us to predict the behaviour of the HVS in rating the quality. We would like to stress that we presented this illustration solely to provide some more insight in the nature of the data contained by the different percentile intervals. In the proposed IQM algorithm, weights based on the magnitude of the contained coefficients, are automatically assigned to the each percentile. In the next section, we will explain the approach utilized to select the weights for different intervals of the coefficients.

C. The Sparseness Significance Ranking Measure (SSRM)

To experimentally verify our primary assumption regarding the close relation between significance of the features for sparse coding and perceptual quality, we introduce a simple objective quality measure. To have a more transparent estimation of the efficiency of the aforementioned approach, we have intentionally avoided – as explained in section II.A – any explicit consideration for other HVS characteristics like contrast sensitivity functions, masking effects..., which most likely should increase the performance of our IQM. Figure 3 shows

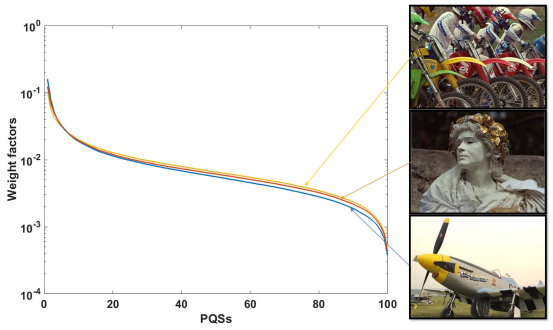


Fig. 4. Example of weights chosen for AC partial quality scores in logarithmic scale. Images I05, I17 and I20 selected from TID2008 database.

the flowchart of the proposed method. After preprocessing the reference and distorted image signals, respectively X and Y following the recommended procedure in [30] for converting both signals into the appropriate scale, we apply a Fast Fourier Transform (FFT) to obtain their complex valued features. In a next step, the features are categorized (ranked) following the procedure explained in section II.A . Since the DC component and the lowest frequency AC components have the biggest amplitudes, we always categorize them into a separate bin, called the DC category. They are important because these coefficients carry information about larger areas in the image and distortion of these coefficients have typically significant impact on the overall image quality. The number of features for the DC category is experimentally chosen such that it only contains a very small portion of total number of features. In our particular instantiation, we have chosen heuristically the lowest 25 frequency coefficients.

Next, we use quantile function in (4) to find the N intervals of the amplitude range of signal X. We use the boundary values of each interval as a threshold to find the features in X that fall within that interval and we also temporarily store the indices of the selected features to find the corresponding coefficient values in Y. Since we choose $N = 100$, for every k_{th} N-tile (percentile in this case) denoted by p^{kth} such that $k \leqslant N$, we will have 2 complex valued vectors p_X^{kth} and p_Y^{kth} of which their values should be compared.

Next step is to calculate the weights for each of the partial quality scores (PQSs) which are the scores coming from each percentile based on (10). We simply consider the median of the values in each percentile as the weight factor for the corresponding PQS. Figure 4 illustrates the chosen percentile weights on a logarithmic scale for a few sample images selected from the TID2008 database. It can be observed that while the weight functions are depending on the scene content, the different weighting curves are very similar and highly weighing the first percentiles. The selection of coefficient amplitude as a weight for that coefficient is due to the fact that the behaviour of such a set of weight values for digital images is extremely sublinear and approximately in an exponential manner. Therefore, it is compatible with the definition of sparsity cost function where it is supposed to more penalize the coefficients near zero (or more reward the ones far from zero).

Similar sparsity functions are also suggested in [31]. From this point of view, it can be seen that while our proposed weight function indeed has a sublinear behaviour, the amount of sublinearity chosen for input images is not fixed or arbitrarily chosen but it is completely scene-dependent. This means it is the content of the input data (thus distribution of coefficients) itself that dictates the amount of sublinearity for the weight function. Nevertheless this does not limit the possibility for using other sparsity functions (e.g. logarithmic) as a weight function to impose more or less sublinearity. However, investigating performance and compatibility of other possible weight functions for the proposed method can be the subject of another more specific research and is not discussed here.

The final SSRM score will be calculated by multiplication of the final score from weighted sum of all PQSs and the Q_{DC} as it is formalized in next section. The overall algorithm is provided in Algorithm (1) and details of the processing steps will be provided in the next section.

III. Sparseness Significance Ranking Measure

A. Distortion Analysis of Fourier Features in Complex Plane

To find an appropriate way to measure the dissimilarity between image pairs, we analyze various distortions in Fourier domain after categorizing the features – in this particular case the Fourier components – based on sparsity considerations. Hence, we imposed different types of distortions, each with different levels of severity, on a sample image to understand the behaviour of the distortion vectors of the Fourier components in complex plane. To make our comparison more tractable, we categorize reference image Fourier features into different bins using (4) with $N = 100$ (percentiles) and then use their indices to find the corresponding distorted image features. Next, we compare the ring shaped distributions of these bins in the complex plane. To avoid cluttering, only a subset of bins including percentiles 3, 10, 25, 75 and 100 are selected for demonstration here. To clearly represent the individual changes of the complex values for each distortion type, we calculated *distortion vectors* representing the exact vectorial distance between the complex amplitude of the coefficient of a sample Fourier basis function of the reference image and the coefficient of the corresponding function of the distorted image. Figure 5 represents such complex plane-distortion vector diagrams of 6 different distortion types including additive white gaussian noise, blur, JPEG, JPEG 2000, contrast and frequency noise for image “1600” selected from CSIQ database. A quick look at the plots in Figure 5 reveals that apart from contrast change, which has almost equal size unidirectional distortion vectors for each percentile without any phase distortion, the other diagrams clearly show random vectors, both in terms of amplitude and phase. From information theory point of view, such randomness can directly be translated into increasing overall and percentile-wise entropy, and statistically speaking, increase of the dispersion for their distributions. Also Figure 6 confirms entropy or dispersion rise for each percentile is directly related to the severity of the imposed distortion. Contrast noise is a special case where

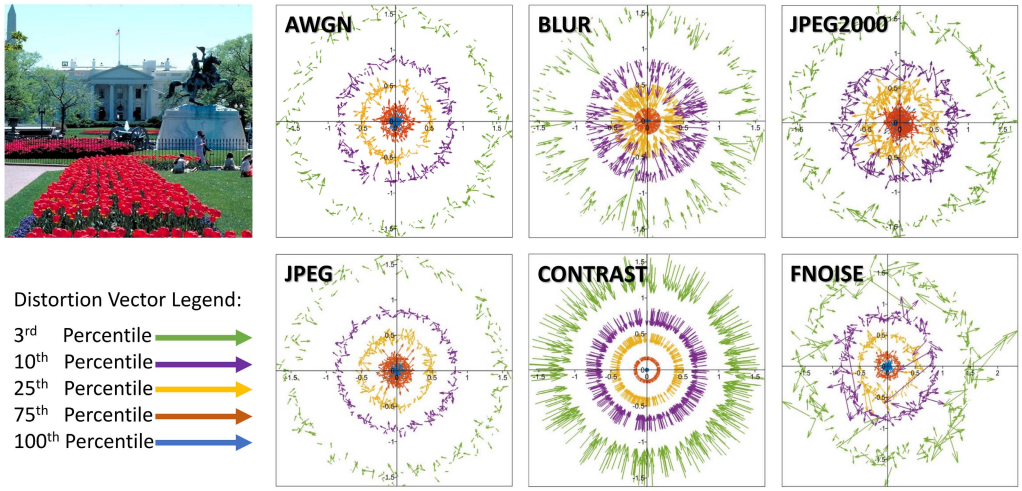


Fig. 5. Distortion Analysis in complex plane of Fourier Coefficients for percentiles 3, 10, 25, 75, 100 and for distortion types: AWGN, Blur, JPEG 2000, JPEG, Contrast and Fnoise. Each distortion vector, starts from a Fourier coefficient value in the reference image and ends in the corresponding distorted value in the distorted image. Distortion vectors of each percentile are represented with the same colour. The reference image is selected from the CSIQ database.

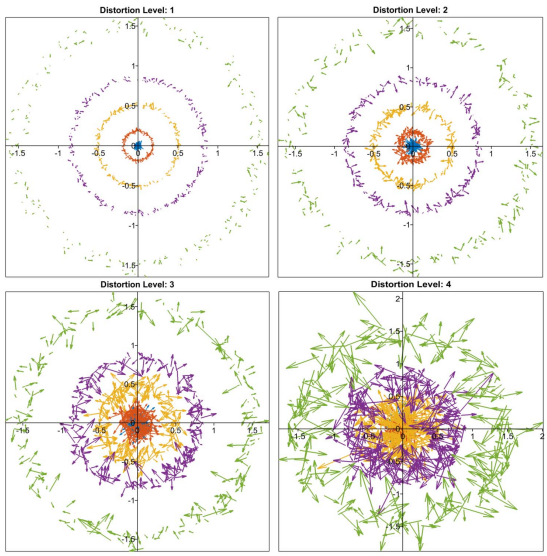


Fig. 6. Distortion analysis of Fourier Coefficients for percentiles 3, 10, 25, 75, 100 in complex plane for 4 levels of JPEG2000 compression distortion from low to high compression level (Higher to lower quality) respectively.

only the overall scale of the distribution changes without any phase dispersion. So, a good comparison method should be able to capture these two main characteristics of the distortion vectors namely dispersion and scale changes to detect the impairment between reference and distorted image. Apart from the contrast case, dispersion of quantile distributions which causes an increase in the distance between inner and outer circle boundaries of the distribution for every quantile bin, can be measured by calculating radial correlation coefficient between the coefficient values in the reference image quantile and those in the corresponding distorted image quantile.

However, such a measurement will be completely invariant to contrast changes and any other type of distortion, which may behave similarly in complex plane, i.e. solely inducing scale changes. Because contrast *distortion vectors* shown in Figure 5 are in the radial direction, which obviously results in a perfect linear correlation between amplitude values in reference and distorted images seen the linear nature of the distortion; i.e. the distortion cannot be captured by the mentioned measurement. This problem cannot be solved by simply decomposing distortion vectors in their real and imaginary basis vectors and separately calculating correlation coefficients for each part; changes in both real and imaginary parts will be linear too. We will address this complication in section III.

Besides the overall distribution behaviour of the distortion vectors, it is also desired to have an estimation of impairment between individual pairs of coefficients. To calculate the amount of such pairwise dissimilarities, one can use the luminance similarity formula proposed in the Structural Similarity Index Metric SSIM method, which can be rewritten based on the normalized l_2 distance as:

$$S(x_i, y_i) = \frac{2x_i y_i + C}{x_i^2 + y_i^2 + C} = 1 - \frac{(x_i - y_i)^2}{x_i^2 + y_i^2} \tag{6}$$

It should be noted that calculating the similarity of complex values in polar form – i.e. directly on the amplitude and phase factor – is not desired considering the fact that the perceptual impact of the distortions on the phase and amplitude is not balanced and combining their similarity scores should be weighted unequally, mostly in favour of phase part [23], [32], [33]. To avoid training sessions to find the mixing parameters for polar components and also to bypass wrapping issues coming from periodic nature of phase factors, it is preferred to calculate the similarity on the real and imaginary parts where amplitude and phase are naturally combined and available in both parts.

B. Formalization of the Partial Quality Scores for Each Percentile

Following the analysis procedure proposed in the previous section, we conduct the comparisons for each percentile by

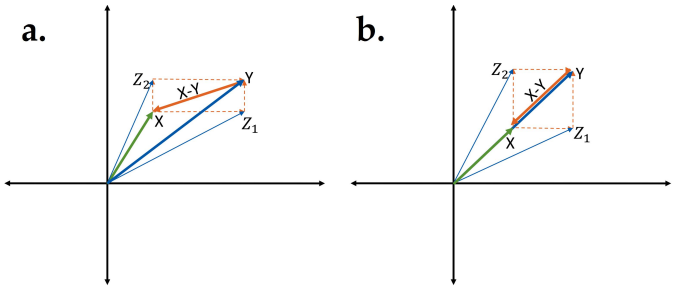


Fig. 7. Representation of the intermediary values z_1 and $z_2$in complex plain: (a) regular distortion affected both amplitude and phase. (b) distortion affected the amplitude only (case of contrast change).

computing both the similarity between their real and imaginary components and the amplitude of complex Pearson correlation coefficient. To circumvent the complication in cases like contrast distortion where the correlation coefficient is invariant to linearly correlated distortion vectors, we create two intermediary complex vectors Z_1 and Z_2 such that for p^{kth} :

$$Z1_{p^{kth}} = \mathrm{Re}(Y_{p^{kth}}) + \mathrm{Im}(X_{p^{kth}})i \tag{7}$$

$$Z2_{p^{kth}} = \mathrm{Re}(X_{p^{kth}}) + \mathrm{Im}(Y_{p^{kth}})i \tag{8}$$

and then calculate their complex valued correlations with the reference signal separately. Figure 7 (a) represents the example of a value from signal X and Y and their corresponding intermediary values z_1 and z_2 in the complex plane. Figure 7 (b) represents the special case where x and y are inline and calculating direct correlation will not react to the distortion. Instead, z_1 and z_2 clearly will not be inline with the reference signal as long as the amplitude of the distortion vector is not zero ($|x - y| \neq 0$). Mathematically, one can write a sample vector $\vec{y}$ based on $\vec{x}$ where both vectors start from the origin as:

$$\vec{y} = \alpha(\vec{x}) + \vec{w} \tag{9}$$

where w can be considered as a non-linear term such that if $|\vec{w}| = 0$ (Figure 7 (b)), independent from the value of constant α, vectors $\vec{x}$ and $\vec{y}$ will be linearly correlated. By introducing intermediary vectors according to (8) and (7), we guarantee their nonlinear relation with $\vec{x}$ by separately adding the real and imaginary bases of distortion vector ($|\vec{x} - \vec{y}|$) as their non-linear terms. It is obvious that this non-linear term will be zero for both intermediary vectors iff $|\vec{x} - \vec{y}| = 0$. Multiplication of real and imaginary similarity values with correlation amplitude of intermediary vectors with reference values and then averaging over all creates the final quality score for that percentile:

$$Q_{p^{kth}} = \frac{1}{T}\sum^{T}([|r^{kth}_{XZ1}| \times S(Re(X_{p^{kth}}), Re(Y_{p^{kth}}))] \bullet [|r^{kth}_{XZ2}| \times S(Im(X_{p^{kth}}), Im(Y_{p^{kth}}))]) \tag{10}$$

where T is equal to the total number of features in percentile p^{kth}, $|r_{XY}|$ is the amplitude of the complex correlation coefficient, S the similarity formula proposed in SSIM method (6) and " $\bullet$ " represents element wise multiplication of values in two vectors.

C. Formalization of Sparseness Significance Ranking Measure

The final quality score for AC percentiles will be calculated by:

$$Q_{AC} = \sum_{k=1}^{N}(Q_{p^{kth}} \times W_{p^{kth}}) \tag{11}$$

where $N = 100$ is the number of quantiles and

$$W_{p^{kth}} = \frac{median(|X_{p^{kth}}|)}{\sum_{k=1}^{N} median(|X_{p^{kth}}|)} \tag{12}$$

To calculate the Q_{DC}, the same procedure will be repeated to measure the intermediary correlations and similarities for real and imaginary parts. To apply the same policy as AC components, the average of the real and imaginary similarities for each pair of features will be weighted based on the amplitude of the reference feature:

$$W_j^{DC} = \frac{|X_j^{DC}|}{\sum_{j=1}^{J}|X_j^{DC}|} \tag{13}$$

$$S_j^{DC} = \frac{S(Re(X_j^{DC}), Re(Y_j^{DC})) + S(Im(X_j^{DC}), Im(Y_j^{DC}))}{2} \tag{14}$$

$$Q_{DC} = |r^{DC}_{XZ1}||r^{DC}_{XZ2}|(\sum_{j=1}^{J}(S_j^{DC} \times W_j^{DC})) \tag{15}$$

where J is the total number of feature pairs in DC category. Finally, the SSRM will be calculated by simply multiplying (11) and (15):

$$SSRM = (Q_{AC})^l \times (Q_{DC})^m \tag{16}$$

The powers m and l are added to balance the effect of each score on the final score, Although in this research we used l, $m = 1$. So far, we succeeded to design a simple method - formalized in Algorithm (1) - which is specifically designed to extract structural information in frequency domain only by considering the importance of the Fourier features for sparse coding. However, it is important to remember that the main goal in this research was not designing an IQM to outperform all the available methods, but to analytically and experimentally examine the hypothesis of relation between perceptual image quality and importance of features for sparse coding.

IV. Experimental Results and Discussion

In this section, we first perform a separation ratio analysis for the partial quality scores to assess their prediction accuracy.. Then we demonstrate the performance of the SSRM method on 4 well-known subjectively annotated test databases. We will also provide results for 11 state-of-the-art and/or well-known full reference IQMs to facilitate comparison of the advocated approach for image quality assessment with the best currently available methods in the field. The other tested methods include: dss–index [34], FSIM [6], GMSD [5], IQM2 [35], IWSSIM [26], MAD [7], MS-SSIM [4], SSIM [2], VIF [8], and PSNR.

Algorithm 1 The Algorithm for Calculating SSRM

procedure MAIN
  $X \leftarrow$ *Reference Image*
  $Y \leftarrow$ *Test Image*
  $N \leftarrow$ *desired number of AC Quantiles*
  $J \leftarrow$ *desired number of DC components*
  X^{DC}, $X^{AC} \leftarrow$ FFTSplit *(X)*
  Y^{DC}, $Y^{AC} \leftarrow$ FFTSplit *(Y)*
  $Q_{AC} \leftarrow$ SimFuncAC(X^{AC}, Y^{AC}, N)
  $Q_{DC} \leftarrow$ SimFuncDC(X^{DC}, Y^{DC}, J)
  $SSRM \leftarrow Q_{AC} \times Q_{DC}$
procedure FFTSPLIT(X)
  $K \leftarrow$ *FFT(X)*
  $DC \leftarrow$ *The J lowest freq. components of K*
  $AC \leftarrow$ *The remaining freq. components of K*
procedure SIMFUNCAC(X^{AC}, Y^{AC}, N)
  sort (X^{AC})
  apply the same sort order for Y^{AC}
  find N *quantiles of* X^{AC}
  for $k \leftarrow 1 : N$ **do**
    calculate the $Q_{p^{kth}}$ and $W_{p^{kth}}$ using **(10)** and **(12)** respectively
  end for
  calculate the Q_{AC} using equation **(11)**
procedure SIMFUNCDC(X^{DC}, Y^{DC}, J)
  for $j \leftarrow 1 : J$ **do**
    calculate W_j^{DC} and S_j^{DC}using respectively **(13)** and **(14)**
  end for
  calculate Q_{DC} from equation **(15)**

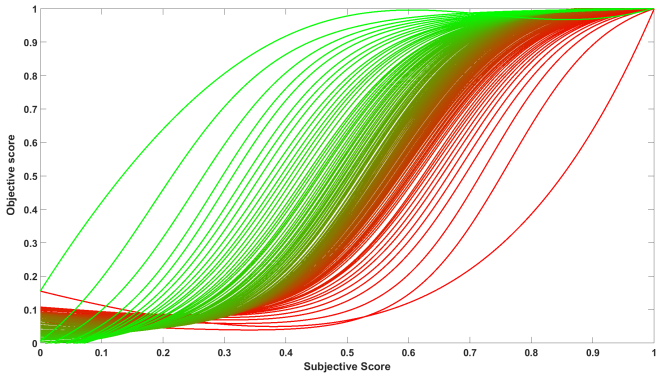


Fig. 8. Logistic regression curves for each percentile predicted quality scores vs subjective scores tested on LIVE database.

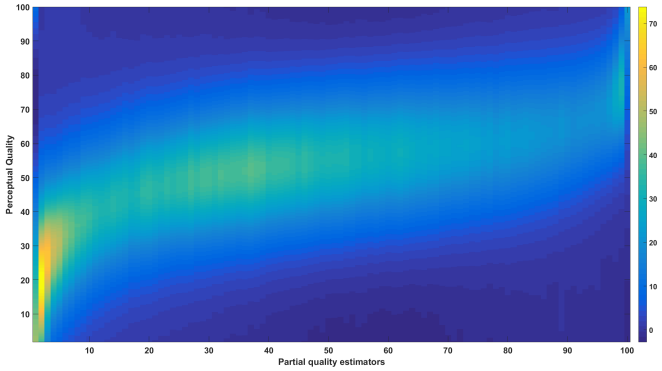


Fig. 9. Separation Ratio (SR) colour map for all partial quality estimators. The colour map encodes the SR values, where a higher SR indicates a better prediction accuracy. Each column represents the colour coded SR values representing prediction accuracy for different parts of the perceptual quality range normalized from 0 to 100 for all images in LIVE database.

A. Separation Ratio Analysis of Partial Quality Scores

Creating partial quality estimators from the quantile wise comparison of data according to the explained procedure for SSRM, provides us the opportunity to test the prediction accuracy of each PQS separately. This can be useful to understand the behaviour of the frequency components which have been categorized, not according to their indices but only based on their amplitude ranking or sparseness significance. We deployed the suggested procedure in [36] for calculating Separation Ratio (SR) to measure the local prediction accuracy of each partial quality estimator through the whole range of perceptual quality. We used all images in the LIVE database and calculated the PQSs for each image (100 scores per image). Hence, the results provided in this analysis are for all the distortion types in this database. Then we re-grouped the results for all images into 100 bins where each set of PQSs was calculated from the corresponding quantile. Next, a logistic regression was performed using (18) to fit the PQSs calculated per quantile to the DMOS scores provided in the database. Figure 8 shows the corresponding regression curves for 100 partial quality estimators of the SSRM. Their indices have been coded by green to red colours such that first percentile is green and the last one is red. The SR in each point is calculated as:

$$\text{SR} = \frac{\text{Steepness of Regression curve}}{\text{Regression fit error}} \tag{17}$$

We simplified the formulation by considering a fixed fit error equal to the median of the fit errors over all regression curve values. Figure 9 shows the calculated SR values. Each column represents SR values of one partial quality estimator for different parts of the perceptual quality range between 0 to 100. Interestingly, a number of important results can be deducted from the results. A very specific pattern is observable, illustrating that the prediction accuracy for each percentile is maximum at a different point in the perceptual quality scale. The first percentiles depict high accuracy when predicting the quality of highly distorted images (irrespective to the distortion type), while the last percentiles are preferred to predict the quality of images only with distortions very near to the JND threshold. This confirms that to have an accurate prediction through the whole perceptual quality range, results from all percentiles should be considered, as we implemented for SSRM. It is very interesting to observe that these results also re-confirm assumptions about the dual-strategy of the HVS when evaluating high and low perceptual quality images [7]. The lower percentiles, characterized as those carrying the larger scale structural information of the

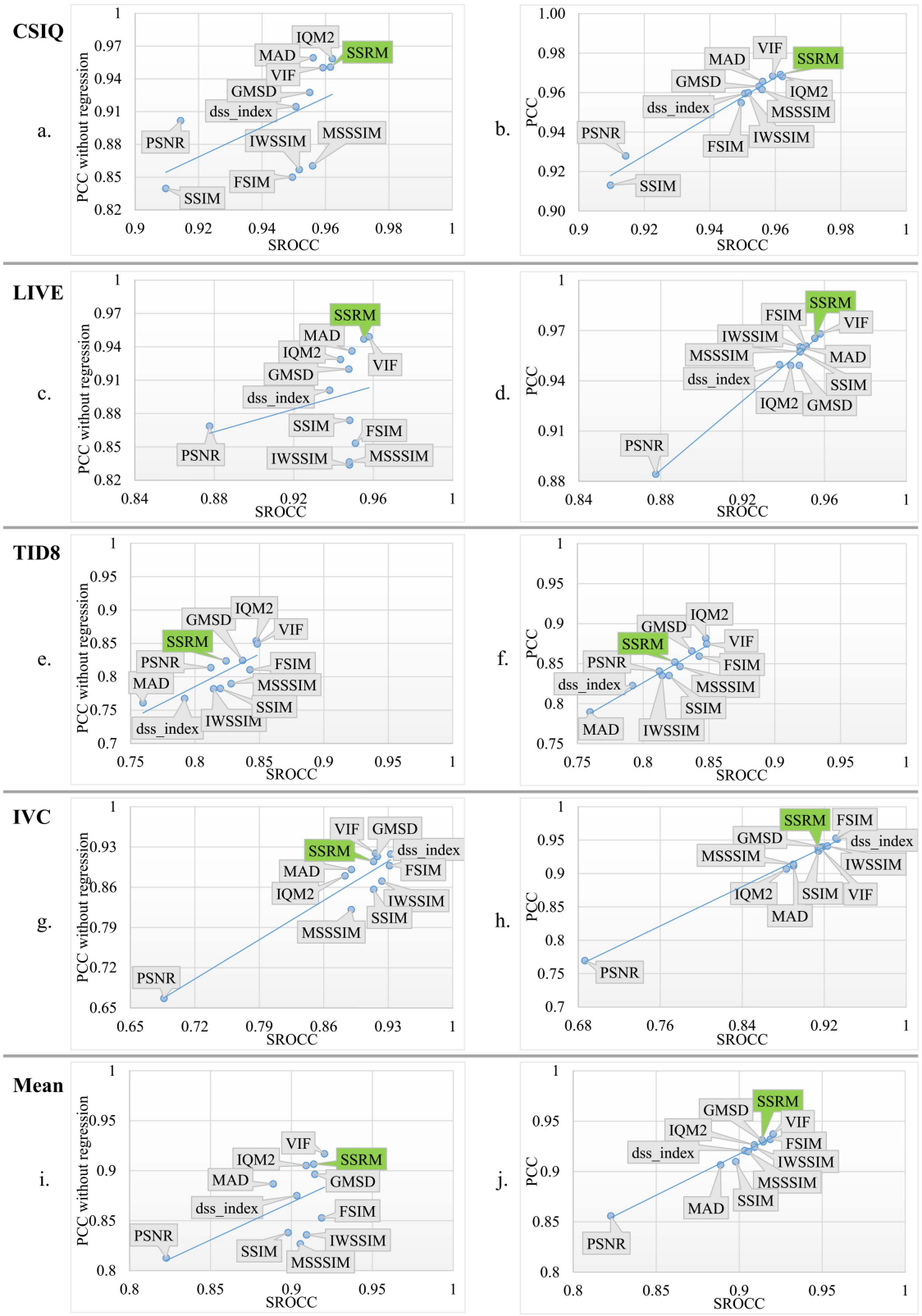


Fig. 10. Diagrams of PCC versus SROCC averaged over all distortion types for different databases. In first column PCC values shown before nonlinear regression and in second column PCC values are calculated after nonlinear regression.

scene, have more success in predicting the quality of highly distorted images, i. e. the first strategy. The last percentiles, which contain mostly very high spatial frequency content and thus fine details of the scene, have better accuracy when the larger scale structural information of the scene remained relatively seen mainly intact and hence the HVS looks for more subtle changes in the data, i.e. the second strategy. Noticeably, the transition between the two so-called strategy

TABLE I

AVERAGE COMPUTATIONAL COST FOR 10 IMAGES SELECTED FROM LIVE DATABASE

IQM:	PSNR	GMSD	SSIM	MSSSIM	dss_index	SSRM	IQM2	FSIM	IWSSIM	VIF	MAD
Average duration (s)	0.0085	0.0108	0.019	0.0732	0.0995	0.1601	0.1866	0.4241	0.56	1.4196	2.2312
Normalized duration c.f. PSNR	1	1.26	2.22	8.57	11.65	18.74	21.84	49.63	65.53	166.13	261.09

TABLE II

PEARSON CORRELATION COEFFICIENT *Before* NONLINEAR REGRESSION CALCULATED FOR ELEVEN METHODS AND FOUR DATABASES

PCC-Before Mapping		dss_index	FSIM	GMSD	IQM2	IWSSIM	MAD	MSSSIM	PSNR	VIF	SSIM	SSRM
CSIQ	AWGN	0.868	0.764	0.922	0.944	0.787	**0.949**	0.825	0.944	**0.959**	0.804	0.933
	BLUR	0.940	0.882	0.946	0.963	0.899	**0.971**	0.867	0.908	0.963	0.869	**0.974**
	JPEG	0.958	0.903	0.944	0.985	0.909	**0.970**	0.906	0.790	0.959	0.916	**0.975**
	Contrast	0.905	0.873	0.892	**0.944**	0.910	0.931	0.900	0.889	0.929	0.767	**0.943**
	fnoise	0.865	0.770	0.898	0.936	0.742	**0.955**	0.779	0.953	**0.955**	0.784	0.897
	jpeg2000	0.952	0.907	0.963	0.978	0.893	**0.981**	0.884	0.927	0.936	0.897	**0.983**
	Average	0.915	0.850	0.928	0.958	0.857	**0.959**	0.860	0.902	0.950	0.840	**0.951**
	Overall	0.928	0.805	**0.929**	0.926	0.795	**0.950**	0.772	0.751	0.922	0.792	0.910
LIVE	fastfading	0.903	0.851	0.928	0.904	0.832	0.940	0.812	0.875	**0.956**	0.862	**0.966**
	gblur	0.920	0.910	0.951	0.946	0.867	0.941	0.846	0.774	**0.957**	0.889	**0.973**
	jp2k	0.885	0.870	0.937	**0.955**	0.834	**0.952**	0.844	0.873	0.936	0.871	**0.942**
	jpeg	0.845	0.728	0.819	**0.913**	0.730	0.861	0.763	0.842	**0.932**	0.790	0.874
	wn	0.952	0.908	0.966	0.925	0.906	**0.988**	0.918	0.979	0.963	0.958	**0.980**
	Average	0.901	0.853	0.920	0.929	0.834	0.936	0.836	0.869	**0.949**	0.874	**0.947**
	Overall	0.861	0.780	0.865	0.903	0.740	0.885	0.679	0.800	**0.941**	0.744	**0.918**
IVC	LAR	0.922	0.909	**0.940**	0.885	0.910	**0.939**	0.862	0.656	0.894	0.880	**0.933**
	blur	0.924	0.888	0.933	0.951	0.860	0.971	0.869	0.862	**0.987**	0.886	**0.974**
	jpeg	**0.932**	0.907	0.914	0.888	0.888	**0.955**	0.824	0.596	0.923	0.833	0.863
	jpeg_lumi	0.904	**0.908**	0.870	0.769	0.860	0.631	0.789	0.504	**0.913**	0.881	0.874
	jpeg_lumi_chromi	**0.946**	0.918	0.911	0.869	0.885	**0.930**	0.832	0.568	0.894	0.872	0.870
	jpeg2000	0.876	0.853	0.913	**0.917**	0.820	**0.919**	0.749	0.814	0.903	0.785	**0.914**
	Average	**0.918**	0.897	0.913	0.880	0.871	0.891	0.821	0.667	**0.919**	0.856	0.905
	Overall	**0.896**	0.860	**0.889**	0.853	0.795	0.839	0.784	0.642	0.882	0.814	0.877
TID2008	Color_noise	0.791	0.819	0.890	0.878	0.754	0.826	0.776	**0.921**	**0.893**	0.775	0.802
	Contrast_change	0.603	0.729	0.551	0.738	**0.778**	0.258	0.769	0.582	**0.881**	0.516	0.702
	Gaussian_blur	0.798	0.908	0.887	**0.956**	0.893	0.923	0.874	0.844	0.939	0.894	**0.949**
	Gaussian_noise	0.812	0.783	**0.883**	0.878	0.715	0.817	0.744	**0.934**	0.866	0.750	0.781
	High_freq_noise	0.836	0.838	0.909	0.941	0.802	0.893	0.818	**0.968**	**0.944**	0.823	0.895
	Image_denoising	0.969	0.932	**0.970**	**0.976**	0.920	0.961	0.916	0.943	0.897	0.921	**0.961**
	Impulse_noise	0.552	0.673	0.604	0.778	0.585	0.041	0.621	**0.856**	**0.814**	0.620	0.706
	JPEG	**0.976**	0.925	**0.980**	0.975	0.906	0.949	0.928	0.860	0.933	0.932	**0.959**
	JPEG_trans_error	0.801	0.842	0.861	**0.862**	0.831	0.856	0.815	0.626	**0.872**	0.837	0.856
	JP2K	0.969	0.955	**0.985**	0.965	0.934	0.973	0.937	0.863	0.917	0.949	**0.978**
	JP2K_trans_error	0.836	0.788	**0.860**	**0.871**	0.773	0.830	0.797	0.853	0.831	0.825	0.846
	Masked_noise	0.229	0.768	0.540	0.842	0.820	0.756	0.785	**0.863**	**0.890**	0.750	0.841
	Mean_shift	0.586	0.671	0.672	0.542	0.653	0.571	0.674	**0.685**	0.590	**0.684**	0.645
	Pattern_noise	**0.813**	0.726	0.755	0.740	0.738	**0.824**	0.664	0.583	0.737	0.669	0.729
	Quant_noise	0.806	0.783	0.863	**0.886**	0.723	0.798	0.761	**0.873**	0.744	0.724	0.821
	Spatial_corr_noise	0.853	0.794	0.903	**0.900**	0.713	0.860	0.758	**0.953**	0.858	0.768	0.809
	blockwise_dists	0.813	0.841	**0.903**	0.787	0.756	0.801	0.786	0.628	0.834	**0.866**	0.723
	Average	0.767	0.810	**0.825**	**0.854**	0.782	0.761	0.790	0.814	**0.849**	0.782	**0.824**
	Overall	0.866	0.830	**0.872**	**0.876**	0.809	0.829	0.790	0.519	0.777	0.740	0.838
Total Average		0.875	0.853	0.896	0.905	0.836	0.887	0.827	0.813	**0.917**	0.838	**0.907**

levels happens smoothly as its depicted in Figure 9. Also, these results are particularly helpful to find the required analysis resolution for perceptual quality prediction of the natural scene imagery (here represented by images in the LIVE database). In particular, the behaviour of majority of middle percentiles is highly similar, while only few first and last percentiles deviate significantly. This means even considerably coarser analysis resolution i.e. lower value of N suffices to effectively characterize the overall characteristics of the depicted trend in Figure 9.

B. Subjectively Tested Databases and Performance Measures

Subjective test data sets perform a significant role in the evaluation of IQMs. Seen the fact that the method of conducting subjective tests, the number of participants, the selected sets of images, the levels of the distortions imposed on the images and many other environmental factors may vary for these databases, IQMs typically show different performance on different databases. In our experiments we involved the CSIQ [37], LIVE2 [38], IVC [39] and TID2008 [40] databases. We calculate the Spearman Rank order correlation(SROCC) to evaluate the monotonicity behaviour of the IQMs under test. The Pearson Correlation Coefficient (PCC) is computed to assess the linear relationship between the subjective scores (perceptual quality) and the prediction scores delivered by the IQMs. Due to the fact that many IQMs have non-linear behaviour, a non-linear regression with a logistic function is normally performed on the predicted quality scores and then PCC will be calculated. However, logistic function selection and its parameters setting is a very sensitive process which can affect the evaluation results. It has to be noted that the parameters are learned after the prediction of the perceptual quality and hence these are completely dependent on the image training sets (i.e. they vary from database to database). Also, we believe that an ideal IQM should have a linear relationship with perceptual quality, thus be able to achieve high PCC scores even without any fitting process. Hence, in our experiments, we have measured PCC both before and after performing a logistic regression using a 5-parameter logistic function:

$$Q_i^{\text{mapped}} = c_1\left(0.5 - \frac{1}{e^{c_2(Q_i - c_3)}}\right) + c_4(Q_i) + c_5 \qquad (18)$$

where c_1, c_2, c_3, c_4, and c_5 are the parameters and Q_i is the prediction score for the image i. We used the MATLAB

TABLE III
SPEARMAN RANK ORDER CORRELATION CALCULATED FOR ELEVEN METHODS AND FOUR DATABASES

SROCC		dss_index	FSIM	GMSD	IQM2	IWSSIM	MAD	MSSSIM	PSNR	VIF	SSIM	SSRM
CSIQ	AWGN	0.949	0.926	**0.968**	**0.964**	0.938	0.954	0.947	0.936	**0.957**	0.897	0.954
	BLUR	0.974	0.972	0.971	**0.977**	**0.978**	0.968	0.971	0.929	0.975	0.961	**0.978**
	JPEG	0.965	**0.966**	0.965	0.965	**0.966**	0.961	0.963	0.888	**0.970**	0.955	**0.966**
	Contrast	0.924	0.942	0.904	**0.957**	**0.954**	0.921	0.953	0.862	0.935	0.792	**0.953**
	fnoise	0.919	0.923	**0.950**	0.934	0.906	**0.957**	0.933	0.934	**0.951**	0.892	0.945
	jpeg2000	0.973	0.968	0.972	**0.976**	0.968	**0.975**	0.968	0.936	0.967	0.961	**0.974**
	Average	0.951	0.950	0.955	**0.962**	0.952	0.956	0.956	0.914	0.959	0.910	**0.962**
	Overall	**0.956**	0.924	**0.957**	0.938	0.921	0.947	0.913	0.806	0.919	0.876	0.937
LIVE	fastfading	0.941	0.950	0.941	0.916	0.944	**0.957**	0.947	0.890	**0.965**	0.955	**0.969**
	gblur	0.923	**0.971**	0.957	0.950	**0.972**	0.947	0.954	0.782	**0.973**	0.952	**0.970**
	jp2k	**0.957**	**0.957**	**0.958**	0.952	0.950	0.953	0.953	0.890	0.953	0.953	0.949
	jpeg	0.910	**0.913**	0.909	0.911	0.907	0.905	**0.913**	0.841	**0.913**	0.912	0.908
	wn	0.959	0.965	0.974	**0.989**	0.967	0.984	0.973	**0.985**	**0.986**	0.969	0.981
	Average	0.938	0.951	0.948	0.944	0.948	0.949	0.948	0.878	**0.958**	0.948	**0.955**
	Overall	0.919	**0.923**	0.910	0.915	0.921	0.919	0.903	0.820	**0.953**	0.900	**0.930**
IVC	LAR	0.911	0.873	0.897	0.895	0.898	**0.928**	0.859	0.699	0.888	0.887	**0.919**
	blur	0.930	**0.967**	0.950	0.942	0.962	0.964	0.946	0.805	**0.973**	0.932	**0.976**
	jpeg	**0.959**	**0.956**	0.946	0.908	**0.947**	**0.947**	0.922	0.674	0.924	0.923	0.906
	jpeg_lumi	0.901	**0.921**	0.871	0.782	0.883	0.618	0.822	0.528	0.900	**0.922**	0.874
	jpeg_lumi_chromi	**0.949**	0.931	0.908	0.844	0.899	**0.943**	0.863	0.562	0.878	0.892	0.868
	jpeg2000	**0.946**	0.941	0.937	0.927	**0.950**	0.940	0.927	0.850	0.936	0.930	**0.942**
	Average	**0.933**	**0.931**	0.918	0.883	**0.923**	0.890	0.890	0.686	0.916	0.914	0.914
	Overall	**0.916**	**0.923**	0.906	0.856	0.904	0.844	0.881	0.651	0.897	0.891	0.895
TID2008	Color_noise	0.805	0.850	**0.897**	0.882	0.791	0.824	0.804	**0.899**	0.875	0.802	0.827
	Contrast_change	0.526	0.647	0.464	0.636	0.629	0.271	0.637	0.585	**0.818**	0.524	0.620
	Gaussian_blur	0.797	0.947	0.897	**0.964**	**0.964**	0.920	0.956	0.870	0.954	0.955	**0.961**
	Gaussian_noise	0.843	0.858	**0.918**	0.896	0.787	0.839	0.809	**0.907**	0.880	0.811	0.814
	High_freq_noise	0.877	0.909	**0.919**	0.909	0.866	0.886	0.869	**0.927**	0.907	0.873	0.894
	Image_denoising	0.965	0.960	**0.975**	**0.971**	0.947	0.943	0.958	0.942	0.916	0.953	**0.955**
	Impulse_noise	0.582	0.744	0.660	0.797	0.644	0.062	0.689	**0.872**	**0.832**	0.671	0.731
	JPEG	0.932	0.929	**0.953**	**0.946**	0.918	0.927	0.932	0.873	0.917	0.925	**0.934**
	JPEG_trans_error	0.828	**0.871**	0.862	0.854	0.859	0.866	**0.868**	0.752	0.858	**0.867**	0.858
	JP2K	0.966	0.978	**0.980**	**0.982**	0.974	0.971	0.970	0.814	0.971	0.963	**0.977**
	JP2K_trans_error	0.847	0.854	**0.882**	**0.899**	0.820	0.839	**0.861**	0.831	0.850	0.858	0.850
	Masked_noise	0.615	0.804	0.710	**0.848**	0.811	0.735	0.813	**0.852**	**0.870**	0.782	0.843
	Mean_shift	0.590	0.671	0.650	0.491	**0.708**	0.517	**0.735**	0.697	0.511	**0.724**	0.590
	Pattern_noise	**0.805**	0.750	0.760	0.748	0.772	**0.829**	0.738	0.582	0.762	0.711	0.742
	Quant_noise	0.837	0.855	**0.887**	**0.895**	0.817	0.816	0.859	0.870	0.797	0.853	0.858
	Spatial_corr_noise	0.851	0.850	**0.914**	**0.906**	0.773	0.869	0.822	**0.917**	0.871	0.816	0.824
	blockwise_dists	0.797	**0.849**	**0.897**	0.784	0.763	0.797	0.755	0.619	0.833	**0.846**	0.730
	Average	0.792	**0.843**	0.837	**0.848**	0.814	0.760	0.828	0.812	**0.848**	0.820	0.824
	Overall	0.870	**0.880**	**0.891**	**0.885**	0.856	0.834	0.854	0.553	0.749	0.775	0.833
Total Average		0.903	**0.919**	**0.914**	0.909	0.909	0.889	0.905	0.823	**0.920**	0.898	**0.914**

optimization toolbox to find the best parameter values to maximize the correlation between the subjective and prediction scores. To aid the optimization algorithms to avoid local minima, we implemented an exhaustive search by systematically repeating the regression process for different parameter initializations to ensure finding the best logistic fit and consequently highest PCC for every regression (about 1000 initializations were tested for each regression and for each initialization 200 iterations were performed by the optimization algorithm). Since, each database consisted of images subjected to different distortion types, we calculated as well the average of the individual correlation scores for different distortion types as an indicator of how well a IQM predicts the quality of specific distortion types.

C. Test Results and Discussion

The database-wise average for different distortion types is presented in Figure 10, where the horizontal axis corresponds to the SROCC and the vertical axis corresponds to the PCC. In the first and second column the PCC is calculated respectively without and after a regression step. From the results shown in the graphs it is pretty obvious that there is no single method that consistently outperforms the others. However, it can be easily observed that SSRM depicts a highly competitive performance by always belonging to the top performing group both in terms of SROCC and PCC. Seen the rather simple structure of SSRM and the fact that we intentionally have not embedded other characteristics of the HVS in its design and no block-wise operations were performed, the achieved results are highly encouraging the validity of the assumptions in our approach, i.e. the validity of the sparsity criterion. For the CSIQ database, the SSRM and IQM2 closely share the top position in terms of SROCC. Interestingly while in Figure 10.b all the IQMs except PSNR and SSIM stay competitive in terms of PCC, In Figure 10.a where no regression performed before calculating PCC, a distinct difference in performance can be observed where generally SSIM based IQMs and FSIM are clearly underperforming. A very similar trend can be observed in Figure 10.c for the LIVE database. Only VIF exchanges the top rank position with IQM2. The rather vertical alignment of IQMs both in cases c and d, represents their highly competitive performance in terms of SROCC. In case of the TID database (Figure 10.e and 10.f), while staying on top half, SSRM is performing a bit lower in terms of SROCC which is expected seen the fact that some of the distortions in TID database are more localized and require gathering spatial information as well to score more accurately. An improvement that is envisageable for SSRM is to utilize block-wise transformation of images with an effective pooling strategy. This is expected to significantly increase the performance of the baseline SSRM and its robustness especially for such cases. In case of the

TABLE IV

PEARSON CORRELATION COEFFICIENT *After* NONLINEAR REGRESSION CALCULATED FOR ELEVEN METHODS AND FOUR DATABASES

PCC-After Mapping		dss_index	FSIM	GMSD	IQM2	IWSSIM	MAD	MSSSIM	PSNR	VIF	SSIM	SSRM
CSIQ	AWGN	0.946	0.928	**0.967**	**0.965**	0.938	0.956	0.947	0.953	**0.961**	0.898	0.956
	BLUR	**0.979**	0.965	0.969	0.977	**0.980**	0.976	0.966	0.925	**0.979**	0.950	**0.979**
	JPEG	0.983	**0.984**	**0.984**	**0.985**	**0.984**	0.983	0.982	0.896	**0.988**	0.979	**0.984**
	Contrast	0.940	0.944	0.921	**0.959**	**0.961**	0.933	0.953	0.898	0.946	0.790	**0.961**
	fnoise	0.926	0.928	**0.958**	0.939	0.915	**0.961**	0.943	**0.953**	**0.958**	0.892	0.949
	jpeg2000	0.983	0.981	0.979	**0.985**	0.981	**0.984**	0.978	0.947	0.978	0.969	**0.985**
	Average	0.960	0.955	0.963	**0.968**	0.960	0.966	0.962	0.929	**0.968**	0.913	**0.969**
	Overall	**0.961**	0.912	**0.954**	0.930	0.914	**0.951**	0.899	0.800	0.928	0.861	0.930
LIVE	fastfading	0.941	0.946	0.928	0.914	0.944	**0.958**	0.947	0.893	**0.970**	0.955	**0.966**
	gblur	0.938	0.974	0.962	0.951	**0.977**	0.949	0.957	0.785	**0.975**	0.948	**0.976**
	jp2k	**0.966**	**0.965**	0.937	0.955	0.958	**0.962**	0.957	0.896	0.962	0.957	0.956
	jpeg	0.938	**0.947**	**0.946**	0.941	0.941	0.939	0.943	0.860	0.943	**0.944**	0.941
	wn	0.965	0.972	0.974	0.984	0.982	**0.990**	0.984	0.988	**0.990**	0.983	**0.989**
	Average	0.950	**0.961**	0.949	0.949	0.960	0.960	0.958	0.884	**0.968**	0.957	**0.965**
	Overall	0.918	0.923	0.911	0.912	**0.925**	0.922	0.911	0.826	**0.950**	0.909	**0.931**
IVC	LAR	**0.960**	0.938	0.949	0.900	0.938	**0.951**	0.927	0.731	0.906	0.942	**0.954**
	blur	0.961	**0.988**	0.972	0.985	0.983	**0.988**	0.957	0.896	**0.990**	0.947	0.977
	jpeg	**0.965**	**0.963**	0.955	0.947	0.957	**0.963**	0.945	0.747	0.942	0.943	0.932
	jpeg_lumi	**0.916**	**0.940**	0.891	0.781	0.888	0.673	0.821	0.721	**0.937**	0.918	0.914
	jpeg_lumi_chromi	**0.959**	**0.947**	0.928	0.896	0.928	**0.954**	0.907	0.669	0.919	0.917	0.914
	jpeg2000	**0.948**	0.942	0.943	0.931	**0.954**	0.940	0.925	0.853	0.936	0.934	**0.944**
	Average	**0.951**	**0.953**	0.940	0.907	**0.941**	0.911	0.914	0.769	0.938	0.934	0.939
	Overall	**0.924**	**0.935**	**0.918**	0.870	0.916	0.859	0.893	0.687	0.905	0.903	0.906
TID2008	Color_noise	0.831	0.867	**0.915**	**0.900**	0.804	0.845	0.820	**0.927**	0.899	0.817	0.839
	Contrast_change	0.720	0.772	0.760	**0.918**	0.780	0.370	0.791	0.696	**0.906**	0.623	0.883
	Gaussian_blur	0.802	0.945	0.898	**0.962**	**0.955**	0.926	0.950	0.876	0.944	0.951	**0.954**
	Gaussian_noise	0.840	0.850	**0.914**	0.890	0.782	0.824	0.801	**0.934**	0.873	0.801	0.804
	High_freq_noise	0.903	0.926	**0.947**	**0.957**	0.883	0.905	0.886	**0.972**	0.946	0.890	0.916
	Image_denoising	**0.975**	0.969	**0.979**	**0.979**	0.958	0.964	0.966	0.947	0.931	0.961	**0.964**
	Impulse_noise	0.579	0.737	0.669	0.791	0.652	0.157	0.691	**0.863**	**0.829**	0.679	0.729
	JPEG	**0.981**	0.974	**0.985**	0.978	0.959	0.962	0.961	0.870	0.955	0.954	**0.965**
	JPEG_trans_error	0.848	**0.886**	0.873	0.869	0.873	0.878	**0.882**	0.773	0.877	**0.883**	0.872
	JP2K	0.975	0.980	**0.985**	**0.984**	0.976	0.979	0.974	0.877	0.973	0.967	**0.981**
	JP2K_trans_error	0.841	0.854	**0.873**	**0.899**	0.823	0.838	0.845	**0.865**	0.862	0.856	0.859
	Masked_noise	0.692	0.827	0.752	0.872	0.844	0.768	0.837	**0.876**	**0.891**	0.817	0.858
	Mean_shift	0.642	0.713	0.685	0.626	**0.749**	0.649	**0.765**	0.724	0.612	0.738	0.700
	Pattern_noise	**0.834**	0.742	0.767	0.758	0.773	**0.833**	0.741	0.594	0.755	0.706	0.730
	Quant_noise	0.843	0.856	0.895	**0.901**	0.828	0.820	0.861	0.880	**0.904**	0.861	0.861
	Spatial_corr_noise	0.858	0.856	**0.920**	0.909	0.778	0.880	0.828	**0.954**	0.870	0.821	0.828
	blockwise_dists	0.822	0.854	**0.903**	0.801	0.779	0.826	0.792	0.663	0.845	**0.872**	0.743
	Average	0.823	0.859	**0.866**	**0.882**	0.835	0.790	0.846	0.841	**0.875**	0.835	0.852
	Overall	**0.877**	0.874	**0.880**	**0.887**	0.858	0.832	0.845	0.586	0.809	0.773	0.845
Total Average		0.921	**0.932**	0.930	0.927	0.924	0.907	0.920	0.856	**0.937**	0.910	**0.931**

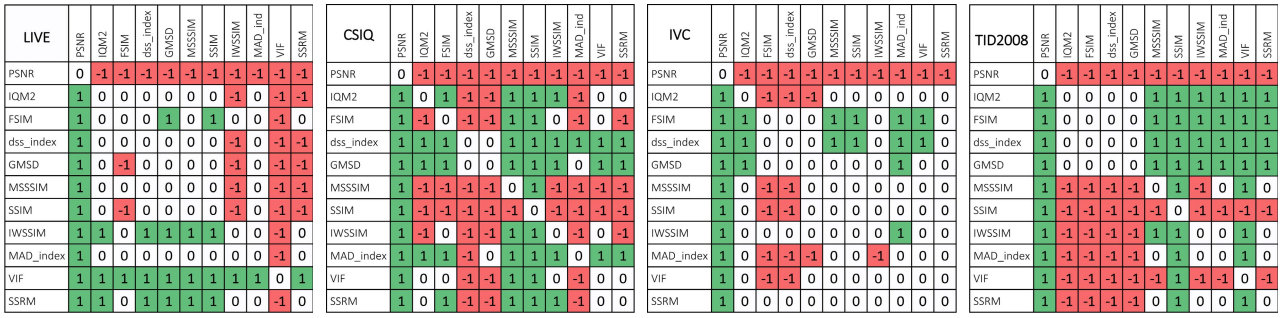


Fig. 11. Results of statistical significance study for tested methods on 4 databases. A value of 1 (Green highlight) represents the significance of method in the row comparing with method in the column, while -1 (Red colour) indicates the vice versa. The value of 0 means corresponding row and column methods are not significantly different from each others.

IVC database IQMs like dss–index and FSIM perform better while SSRM again stays competitive. In Figure 10.i and 10.j, the results for all databases have been averaged. In terms of linear behaviour with respect to the perceptual scores, VIF, SSRM and IQM2 are showing the best overall results respectively while after regression VIF, SSRM, GMSD and FSIM provide the highest scores. It should be noted that the rather better performance of IQMs like VIF or MAD comes with much higher computational cost. The Table I represents the computational costs for the tested IQMs calculated by averaging prediction times for 10 images selected from the LIVE database. The calculations were conducted in MATLAB using a laptop HP Zbook14 with Intel Core i7-4600U CPU and 8GB DDR3 RAM. Their calculation times are also compared relatively with PSNR. It can be noticed that among the top performing IQMs (VIF, SSRM, IQM2 and FSIM), SSRM has the lowest computational cost. Tables II, III and IV provide the complete correlation results for all the distortion types. These results also confirms the tight competition between tested IQMs where top positions for different distortion types are distributed. SSRM though shows a consistently high performance with regards to distortion types like BLUR, JPEG and JPEG2000, while it does not fail or show poor performance for any specific type of distortion type. One exception is Mean

shift error where also all other methods struggle to predict the perceptual score. Also, as an alternative representation of global performance, a statistical significance test is conducted on the fitted scores of the tested IQMs. The results for the 4 databases are presented in Figure 11. Although the achieved results do not show a clear winner through all databases. For example VIF while dominant in LIVE database, consistently underperforms in TID and IVC and stands somewhere in the middle-low range for CSIQ. MAD also shows high statistical significance in CSIQ, but falls behind in IVC and TID. Nonetheless, SSRM shows a rather more consistent statistical significance for all databases along with methods like FSIM and dss-index followed by GMSD and IQM2.

V. Conclusion

Considering the fact that sparse coding is known to be one of the main underlying strategies implemented in our brain, we hypothesized that specific HVS tasks, in particular visual quality assessment, should also be compatible with such a framework. To provide a practical method to investigate our assumption, an alternative categorization of frequency components was proposed where their importance for sparse coding, namely "*sparseness significance*" was considered as the basis for differentiation. Content adaptivity and simplicity are two distinguishable properties of the newly proposed approach. The proposed algorithm does not exclude exploiting visual masking or the contrast sensitivity function, though this considered to be the subject for future research. The main focus was currently on assessing the validity of the sparsity assumption. Considering the Fourier domain as an example for a sparse representation domain, we managed to present a new characterization of *structural information* in the frequency domain based on the same idea. After evaluating the behaviour of Fourier components of the distorted images in complex plane utilizing the proposed notion of *distortion vectors*, we designed a perceptual quality predictor called SSRM, which was only based on the mentioned assumptions. Later on, we performed a novel prediction accuracy analysis using the separation ratio for each partial quality estimator. It is important to note at this stage that the proposed accuracy analysis method is generic and can deploy in the context of the design of any visual quality assessment measure. Finally, through a series of routine benchmarking experiments, a performance assessment of the proposed method was provided. It was demonstrated that without considering any other characteristics of the HVS nor any spatial information of the image and using only the proposed way of extracting structural information in frequency domain, SSRM delivers a competitive performance with respect to state of the art techniques such as VIF, GMSD and FSIM. Important to notice is that all the results are achieved by the proposed method without involvement of machine learning approaches or training sessions for any parameter setting. We believe that current research is just a starting point which motivates a number of opportunities for further investigation. Among them, additional research is needed to propose an efficient way of combining the notion of sparseness significance and visual masking effects. The Fourier domain used in our experiments can be considered as the simplest case – providing however remarkably promising results – while the ideas proposed in this research can be applied to almost any other sparse representation, potentially improving the performance. The provided results with separation ratio analysis may be used to propose even more advanced weight functions, leading to an effective reduced reference IQM. Finally, we would like to remark that the utilized representation model fits fairly well with the internal representation models used in image and video codecs. Hence, holding the promise to perform a more effective rate-distortion control in the transform domain.

References

[1] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[3] N. Kruger *et al.*, "Deep hierarchies in the primate visual cortex: What can we learn for computer vision?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1847–1871, Aug. 2013.

[4] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2. Nov. 2003, pp. 1398–1402.

[5] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.

[6] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[7] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *J. Electron. Imag.*, vol. 19, no. 1, p. 011006, 2010.

[8] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[9] L. He, X. Gao, W. Lu, X. Li, and D. Tao, "Image quality assessment based on S-CIELAB model," *Signal, Image Video Process.*, vol. 5, no. 3, pp. 283–290, 2011.

[10] Y. Zhang, T. D. Phan, and D. M. Chandler, "Reduced-reference image quality assessment based on distortion families of local perceived sharpness," *Signal Process., Image Commun.*, vol. 55, pp. 130–145, Jul. 2017.

[11] D. Tao, X. Li, W. Lu, and X. Gao, "Reduced-reference IQA in contourlet domain," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 6, pp. 1623–1627, Dec. 2009.

[12] X. Gao, W. Lu, D. Tao, and X. Li, "Image quality assessment based on multiscale geometric analysis," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1409–1423, Jul. 2009.

[13] D. Liu, Y. Xu, Y. Quan, and P. Le Callet, "Reduced reference image quality assessment using regularity of phase congruency," *Signal Process., Image Commun.*, vol. 29, no. 8, pp. 844–855, 2014.

[14] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E.-H. Yang, and A. C. Bovik, "Quality-aware images," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1680–1689, Jun. 2006.

[15] H. Barlow, "Single units and sensation: A neuron doctrine for perceptual psychology?" *Perception*, vol. 1, no. 4, pp. 371–394, 1972.

[16] H. B. Barlow, "Single units and sensation: A neuron doctrine for perceptual psychology?" *Perception*, vol. 38, no. 6, pp. 795–798, 2009.

[17] Y. Karklin and M. S. Lewicki, "Is early vision optimized for extracting higher-order dependencies?" in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 18. 2006, pp. 635–642.

[18] B. A. Olshausen and D. J. Field, "Sparse coding of sensory inputs," *Current Opinion Neurobiol.*, vol. 14, no. 4, pp. 481–487, 2004.

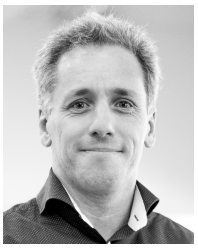
[19] P. J. Garrigues, "Sparse coding models of natural images: Algorithms for efficient inference and learning of higher-order structure," Ph.D. dissertation, EECS Dept., Univ. California, Berkeley, Berkeley, CA, USA, May 2009. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-71.html

[20] H.-W. Chang, H. Yang, Y. Gan, and M.-H. Wang, "Sparse feature fidelity for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 4007–4018, Oct. 2013.

[21] L. He, D. Tao, X. Li, and X. Gao, "Sparse representation for blind image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2012, pp. 1146–1153.
[22] A. van der Schaaf and J. H. van Hateren, "Modelling the power spectra of natural images: statistics and information," *Vis. Res.*, vol. 36, no. 17, pp. 2759–2770, 1996.
[23] M. Narwaria, W. Lin, I. V. McLoughlin, S. Emmanuel, and L. T. Chia, "Fourier transform-based scalable image quality measure," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3364–3377, Aug. 2012.
[24] Z. Wang *et al.*, *Sparse Coding and Its Applications in Computer Vision*. Singapore: World Scientific, 2015.
[25] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 5, pp. 2197–2202, 2003.
[26] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.
[27] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
[28] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by V1?" *Vis. Res.*, vol. 37, no. 23, pp. 3311–3325, 1997.
[29] I. E. Ohiorhenuan, F. Mechler, K. P. Purpura, A. M. Schmid, Q. Hu, and J. D. Victor, "Sparse coding and high-order correlations in fine-scale cortical networks," *Nature*, vol. 466, no. 7306, pp. 617–621, 2010.
[30] Z. Wang. *The SSIM Index for Image Quality Assessment*. Accessed: Jan. 2017. [Online]. Available: https://ece.uwaterloo.ca/~z70wang/research/ssim
[31] P. Foldiak and D. Endres, "Sparse coding," *Scholarpedia*, vol. 3, no. 1, p. 2984, 2008, doi: 10.4249/scholarpedia.2984.
[32] W. H. Hsiao and R. P. Millane, "Effects of Fourier-plane amplitude and phase errors on image reconstruction. I. Small amplitude errors," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 24, no. 10, pp. 3180–3188, 2007.
[33] X. S. Ni and X. Huo, "Statistical interpretation of the importance of phase information in signal and image reconstruction," *Statist. Probab. Lett.*, vol. 77, no. 4, pp. 447–454, 2007.
[34] A. Balanov, A. Schwartz, Y. Moshe, and N. Peleg, "Image quality assessment based on DCT subband similarity," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 2105–2109.
[35] E. Dumic, S. Grgic, and M. Grgic, "IQM2: New image quality measure based on steerable pyramid wavelet transform and structural similarity index," *Signal, Image Video Process.*, vol. 8, no. 6, pp. 1159–1168, 2014.
[36] A. Barri, A. Dooms, B. Jansen, and P. Schelkens, "A locally adaptive system for the fusion of objective quality measures," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2446–2458, Jun. 2014.
[37] E. C. Larson and D. Chandler. (2010). *Categorical Image Quality (CSIQ) Database*. [Online]. Available: http://vision.okstate.edu/csiq
[38] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
[39] P. Le Callet and F. Autrusseau. (2005). *Subjective Quality Assessment IRCCyN/IVC Database*. Accessed: Jul. 2016. [Online]. Available: http://www.irccyn.ec-nantes.fr/ivcdb/
[40] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008—A database for evaluation of full-reference visual quality assessment metrics," *Adv. Mod. Radioelectron.*, vol. 10, no. 4, pp. 30–45, 2009.

Ayyoub Ahar received the B.Sc. degree in applied physics from PNU, Iran, and the M.Sc. degree in IT-multimedia computing from Multimedia University, Malaysia. He is currently pursuing the Ph.D. degree in electrical engineering with Vrije Universiteit Brussel (VUB), Belgium. He is a member of the Interfere Team, a multidisciplinary research group at the Department of Electronics and Informatics, VUB. He holds a full scholarship as a part of EU-ERC Consolidator Grant focusing on digital holography. He is also a Researcher at IMEC, Leuven, Belgium. His current research interests include digital image and video processing, complex data analysis, bio-inspired computing, perceptual quality assessment and standardization with an emphasis on emerging 3D data modalities in particular digital holography and light field imaging.

Adriaan Barri received the M.Sc. degree in mathematics from Vrije Universiteit Brussel (VUB), Belgium, in 2010, and the Ph.D. degree from the Department of Electronics and Informatics (ETRO), VUB, in 2015. In 2016, he was with ETRO as a Post-Doctoral Researcher, where he was involved in machine learning and perceptual image and video quality assessment.



Peter Schelkens (M'98) holds a professorship with the Department of Electronics and Informatics, Vrije Universiteit Brussel, Brussels, Belgium, and is a Research Group Leader with IMEC, Leuven, Belgium. In 2013, he received an EU ERC Consolidator Grant focusing on digital holography. He (co-)authored over 250 journal and conference publications. His current research focuses on the multidimensional signal processing while especially focusing on cross-disciplinary research. He is a member of the ISO/IEC JTC1/SC29/WG1 (JPEG) standardization committee. He is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and *Signal Processing: Image Communications*. He is the Co-Editor of the books *The JPEG 2000 Suite* (Wiley, 2009) and *Optical and Digital Image Processing* (Wiley, 2011).