

Studying The Added Value of Visual Attention in Objective Image Quality Metrics

Wellington Y.L. Akamine and Mylène C.Q. Farias
Department of Electrical Engineering,
University of Brasília (UnB),
Brasília, Brazil

Abstract—In this work, we investigate the benefits of incorporating saliency maps obtained with visual attention *computational* models into three image quality metrics. In particular, we compare the performance of simple quality metrics with quality metrics that incorporate saliency maps obtained using three popular visual attention computational models. Results show that performance of simple quality metrics can be improved by adding visual attention information. Nevertheless, gains in performance depend on the precision of the visual attention model, the type of distortion, and the characteristics of the quality metric.

Keywords—image quality, ssim, degradations, image processing

I. INTRODUCTION

Objective visual quality metrics can be classified as *data metrics*, which measure the fidelity of the signal without considering its content, or *picture metrics*, which estimate quality considering the visual information contained in the data. Customarily, quality measurements in the area of image processing have been largely limited to a few data metrics, such as mean absolute error (MAE), mean square error (MSE), and peak signal-to-noise ratio (PSNR), supplemented by limited subjective evaluation. Although over the years data metrics have been widely criticized for not correlating well with perceived quality measurements, it has been shown that such metrics can predict subjective ratings with reasonable accuracy as long as the comparisons are made with the same content, the same technique, or the same type of distortions.

One of the major reasons why these simple metrics do not generally perform as desired is because they do not incorporate any human visual system (HVS) features in their computation. It has been discovered that, in the primary visual cortex of mammals, an image is not represented in the pixel domain, but in a rather different manner. Unfortunately, the measurements produced by metrics like MSE or PSNR are simply based on a pixel to pixel comparison of the data, without considering what is the content and the relationships among pixels in an image (or frames). In the past few years, a big effort in the scientific community has been devoted to the development of better image and video quality metrics that incorporate HVS features (i.e. picture metrics) and, therefore, correlate better with the human perception of quality [1][2].

A recent development in the area of image quality consists of trying to incorporate aspects of visual attention in the design of visual quality metrics[3], mostly using the assumption that

visual distortions appearing in less salient areas might be less visible and, therefore, less annoying [4], [5]. This research area is still in its infancy and results obtained by different groups are not yet conclusive, as pointed out by Engelke *et al* [6]. Some researchers have reported that the incorporation of saliency maps increases the performance of quality metrics, while others have reported no or very little improvement. Among the works that have reported some improvement, most use *subjective* saliency maps, i.e. saliency maps generated from eye-tracking data obtained experimentally [7]. But, although subjective saliency maps are considered as the ground-truth in visual attention, they cannot be used in real-time applications.

To incorporate visual attention aspects into the design of image quality metrics, we have to use visual attention *computational* models to generate *objective* saliency maps. This raises the question of how the metric performance is affected by the “precision” of the saliency map and the integration model. Another open question is how the distortion type affects the saliency map and, consequently, the metric performance. Very few works tested the incorporation of specific computational attention models into image quality metrics [8]. Up to date, there has been no work that compared the incorporation of visual attention computation models versus subjective saliency maps.

In this work, we investigate the benefit of incorporating objective saliency maps into full-reference and no-reference image quality metrics. We compare the performance of the original quality metrics with the performance of quality metrics that incorporate *objective* saliency maps. Also, we study the effects that different types of degradations have on the computational model and, consequently, on the performance of the final metric.

II. INCORPORATION OF VISUAL ATTENTION MODELS

Visual attention is a feature of the HVS that is responsible for defining which areas of the scene are relevant and should be attended. There are two visual selection mechanisms: *bottom-up* and *top-down*. The bottom-up mechanism is an automated selection that is controlled mostly by the signal. It is fast and short lasting, being performed as a response to low-level features that are perceived as *visually salient*. The top-down mechanism is controlled by higher cognitive factors and external influences, such as semantic information,

viewing task, and personal preferences, context. It is slower and requires a voluntary effort.

In this work, we consider three popular bottom-up visual attention computational models: Itti’s model [9], Achanta *et al.*’s model [10], and GAFFE model (Gaze-Attentive Fixation Finding Engine) [11]. For a given image, these models generate a gray-scale saliency map indicating image regions that are most likely to attract attention. In the saliency maps, higher luminance values correspond to higher saliency pixels, while lower values correspond to lower saliency ones. Figs. 1(a) and (b) depicts the images ‘Rapids’ and ‘Caps’, respectively, while the corresponding saliency maps generated using Itti’s, Achanta’s, and GAFFE models are depicted in Figs. 1.(c)-(h). Notice that the saliency maps are able to capture the most salient areas of the images: the colorful caps and the boat. But, the several saliency maps are very different from each other. We used the subjective saliency maps from the TUD LIVE Eye Tracking database as our visual attention *ground-truth* [12]. These saliency maps were collected in a subjective experiment that used twenty-nine source images from the LIVE database [13]. Figs. 1(i) and (j) depict the subjective saliency maps corresponding to the images ‘Rapids’ and ‘Caps’.

We combine the information from the saliency maps into three different full-reference (FR) image quality metrics: Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity (SSIM) index [14]. The MSE and PSNR error maps are calculated using the following equations:

$$MSE(x, y) = (I_o(x, y) - I_t(x, y))^2, \quad (1)$$

and

$$PSNR(x, y) = 20 \log_{10} \left(\frac{MAX_i}{\sqrt{MSE(x, y)}} \right) \quad (2)$$

where $I_o(x, y)$ is the original image pixel, $I_t(x, y)$ is the test image pixel, MAX_i is the highest intensity level of the pixels, and x and y are the horizontal and vertical coordinates.

The third metric, SSIM, is a more complex and robust full-reference (FR) image quality metric [13]. The general equation for SSIM is:

$$SSIM(I_o, I_t) = \frac{(2\mu_o\mu_t + C_1)(2\sigma_{ot} + C_2)}{(\mu_o^2 + \mu_t^2 + C_1)(\sigma_o^2 + \sigma_t^2 + C_2)}, \quad (3)$$

where μ is the average intensity, σ is the standard deviation, and σ_{ot} is the covariance between the original image (I_o) and the test image (I_t). The variables C_1 , C_2 , C_3 are control constants used to avoid problems when the denominator reaches values close to zero.

The combination or integration process consists of using the gray-scale pixel values of the saliency maps as *weights* for the error maps generated by the three quality metrics. The modified saliency-based quality metrics for the corresponding FR metrics are given as

$$SM-MET = \frac{\sum_{x=1}^M \sum_{y=1}^N MET(x, y) \cdot SAL(x, y)}{\sum_{x=1}^M \sum_{y=1}^N SAL(x, y)}, \quad (4)$$

where $SAL(x, y)$ is the saliency map pixel and $MET(x, y)$ is the error map pixel calculated using one of the FR quality

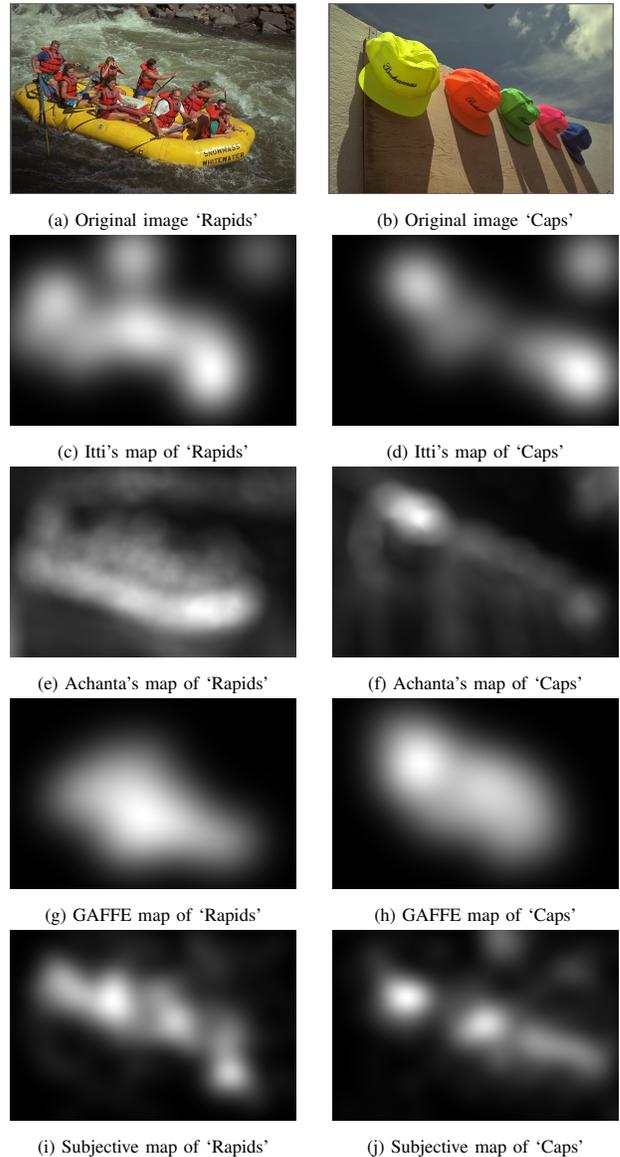


Fig. 1. Saliency maps corresponding to the images ‘Rapids’ and ‘Caps’.

metric (SSIM, PSNR or MSE). This particular combination process was used because it was the simplest solution that allowed that the same model was used for all metrics. This approach makes it easier to compare the performance of the different metrics tested. Other combination models can be found in the work of Redi *et al.* [15].

III. SIMULATION RESULTS

The performance of an image quality metric is measured by how well its output scores (quality estimates) correlate with the Mean Observer Scores (MOS) given by observers in a subjective experiment. To compare the performance of the three original quality metrics with the saliency-based quality metrics (Eq. 4-), we used the LIVE database [13] which contains 779 images with the following degradations: JPEG (175 images), JPEG2k (169 images), Gaussian Blur (GB, 145 images), Fast Fading (FF, 145 images), and White Noise

TABLE I
SPEARMAN CORRELATION COEFFICIENTS FOR MSE METRIC.

Model	JPEG	JPEG2k	GB	FF	WN	All
MSE	0.90117	0.88872	0.78249	0.88549	0.98564	0.87270
Su-MSE	0.91891	0.91338	0.81171	0.90894	0.98526	0.89080
Io-MSE	0.91702	0.90620	0.80246	0.90200	0.98530	0.88650
It-MSE	0.91895	0.90500	0.80360	0.90250	0.98550	0.88620
Ao-MSE	0.91187	0.90757	0.80522	0.90432	0.98548	0.88650
At-MSE	0.91170	0.90470	0.79640	0.90180	0.98560	0.88480
Go-MSE	0.91670	0.91020	0.79140	0.90510	0.98530	0.88790
Gt-MSE	0.91780	0.91150	0.79030	0.90430	0.98520	0.88830
Sw-MSE	0.90091	0.89250	0.74067	0.87540	0.98497	0.86860

TABLE II
SPEARMAN CORRELATION COEFFICIENTS FOR PSNR METRIC.

Model	JPEG	JPEG2k	GB	FF	WN	All
PSNR	0.90120	0.88872	0.78249	0.88549	0.98564	0.87270
Su-PSNR	0.91891	0.91338	0.81159	0.90894	0.98523	0.89080
Io-PSNR	0.91696	0.90620	0.80246	0.90200	0.98530	0.88240
It-PSNR	0.91696	0.90500	0.80370	0.90250	0.98550	0.88620
Ao-PSNR	0.91187	0.90757	0.80522	0.90432	0.98548	0.88650
At-PSNR	0.91170	0.90470	0.79640	0.90180	0.98560	0.88480
Go-PSNR	0.91670	0.91020	0.79140	0.90510	0.98530	0.88790
Gt-PSNR	0.91780	0.91150	0.79030	0.90430	0.98520	0.88830
Sw-PSNR	0.90091	0.89250	0.74070	0.87550	0.98500	0.86860

TABLE III
SPEARMAN CORRELATION COEFFICIENTS FOR SSIM METRIC.

Model	JPEG	JPEG2k	GB	FF	WN	All
SSIM	0.96958	0.95060	0.92506	0.93681	0.96410	0.92210
Su-SSIM	0.97029	0.95188	0.92709	0.94480	0.96889	0.93250
Io-SSIM	0.96874	0.94850	0.92657	0.94287	0.96760	0.93090
It-SSIM	0.95795	0.95060	0.92730	0.94360	0.96960	0.93200
Ao-SSIM	0.96821	0.95000	0.93663	0.94462	0.96473	0.93320
At-SSIM	0.96830	0.94840	0.93580	0.94370	0.96700	0.92990
Go-SSIM	0.96990	0.95560	0.92770	0.94660	0.96860	0.93350
Gt-SSIM	0.96880	0.95410	0.92550	0.94580	0.96940	0.93300
Sw-SSIM	0.96388	0.94593	0.89801	0.93302	0.95786	0.91700

(WN, 145 images). The LIVE database was chosen because it contains different kinds of distortions that reflect a broad range of image impairments commonly found in image processing applications. Also, this database is the same used by the researchers in TuDelft to generate a database of subjective salience maps [7][12]. Therefore, using this database allows us to compare differences in performance of the metrics caused by the incorporation of subjective maps versus the incorporation of ‘computed’ salience maps.

For each image in the database, we evaluate its quality using the original quality metric and the saliency-based quality metrics. The database contains the DMOS (Differential of MOS) for each image, which is the difference between the score given by a subject to the original image and the score given by a subject to the degraded image. To understand if metric performance improves with the saliency information, we calculate the Spearman correlation coefficient between the metric output values and the DMOS [16]. The correlation coefficients are calculated separately for each degradation type and for the set containing all degradation types.

Although the saliency map is generally estimated using the original images, since we want to analyze how the performance of the saliency-based metrics is affected by the use of degraded maps, we estimate saliency maps using both the original and test images. We test the incorporation of subjective (su) saliency maps and the saliency maps obtained with the tested computational models. To make sure that the differences in performance are not by chance, we also test the performance with incorporation of ‘switched’ (sw) saliency maps, which consisted of picking a random saliency map corresponding to another image in the database. To identify the different models, we substitute the initials SM in Eq. 4 by the first letter of the saliency map used (‘I’ for Itti, ‘A’ for Achanta, and ‘G’ for GAFFE) followed by ‘o’ (original) or ‘t’ (test), indicating whether the saliency map is obtained using the original or test image. In Tables I-III, we present the Spearman correlation coefficients for MSE, PSNR and SSIM and their saliency-based versions. Correlation values of saliency-based metrics that represent a gain in comparison to the original metrics are depicted in **bold**.

For saliency-based MSE and PSNR metrics, when we consider each individual distortion (columns 2-6 of Tables I and II), the correlation coefficients improve for almost all distortions, with performance gains varying from 1.2% to 2.1%. The only exception is the degradation White Noise for which the performance decreases with the incorporation of any type of saliency map. Because of the similarity between PSNR and MSE, the correlation values of their corresponding saliency-based metrics are very similar. The best gains are obtained for the subjective maps (1.7% to 2.5%) and GAFFE objective maps (1.6% to 2.2%). Achanta’s model presents the best performance for the degradation Gaussian Blur.

For saliency-based SSIM, considering again only the individual distortions (columns 2-6 of Table III), the performance improves when subjective and GAFFE saliency maps are used. Although GAFFE is the computational model with the best performance, the gains in performance vary with the distortions. The gain for JPEG is only 0.03%, while for other distortions they range from 0.2% to 1%. When Achanta and Itti models are used there is no improvement for JPEG and JPEG2k. For the other degradations, using Itti and Achanta models provides improvement gains from 0.06% to 1.2%. The degradation corresponding to the worst performance is White Noise, with gains from 0.06% to 0.4%. For Gaussian Blur, Achanta’s model incurs in a higher performance gain (1.2%) than Itti’s model (0.16%) or GAFFE (0.28%).

For the set containing all types of distortions (‘All’ – column 7 of Tables I-III), the correlation coefficients of saliency-based metrics show gains ranging from 1.1% to 1.9%. The subjective saliency maps show the highest gain in performance, followed by the GAFFE saliency maps generated from original images. The saliency maps obtained from test images presented an inferior performance for both GAFFE and Itti models, but a better performance for Achanta. For the switched saliency maps, the performance was worse than with any other saliency map. These results seem to point out that the precision of the saliency map has an impact on the performance of the metrics. The correlation values are comparable to the values found by other researchers [7].

Overall, the performance gains for MSE and PSNR were higher than for SSIM. This is expected since SSIM already includes some of the same parameters (e.g. contrast and texture) that are taken into account by attention models. The computational model that presents the best performance is GAFFE. Although the results of GAFFE for SSIM are not as significant as for PSNR and MSE, the gains in performance are close (sometimes higher) than what is obtained with subjective saliency maps.

For Gaussian Blur, the best performance model is Achanta’s – the simplest of the three models tested. Blur removes image details making it easier for simpler models to find salient areas. Most models shows no or very small gain in performance for White Noise. Noise adds more details to the saliency map, making more difficult to find salient areas. These results are in agreement with other studies that show that the importance of saliency maps depends on the type of distortion. To exemplify

this effect, in Fig. 2 the saliency maps corresponding to the images ‘Caps’ with 2 levels of blur are depicted. It can be observed that the saliency maps do not change significantly with the increase in blur, from level 1 to level 2. On the other hand, in Fig. 3 the saliency maps corresponding to the images ‘Caps’ with 2 levels of noise are depicted. It can be observed that the saliency maps change significantly with the increase in noise, from level 1 to level 2. In this particular case, the salient areas have increased.

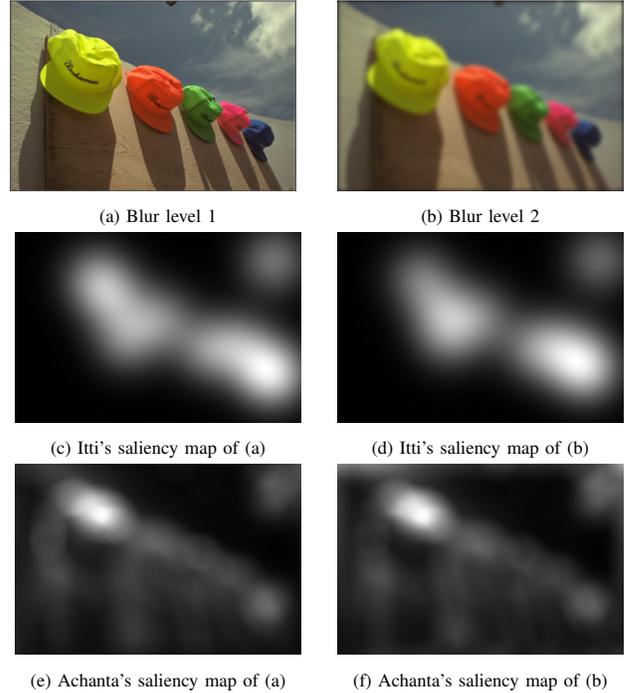


Fig. 2. Saliency maps of the image ‘Caps’ with two levels of blur, obtained using Itti’s and Achanta’s model.

IV. WORK UNDER PROGRESS ON NO-REFERENCE METRICS

Requiring the reference image or even a small portion of it becomes a serious impediment in many real-time transmission applications. For such applications, it is important to develop ways of blindly estimating the quality of an image using *No-Reference* (NR) image quality metrics. It turns out that, although human observers can usually assess the quality of an image without using the reference, designing a NR metric is a difficult task because it is hard to differentiate the natural content from the *defects* that cause *impairments*. One of the possible approaches to this problem consists of taking a multidimensional feature extraction approach by recognizing that the perceived quality of an image can be affected by a variety of artifacts and that the strengths of these artifacts contribute to the overall annoyance. The assumption here is that it is easier to detect artifact signals and estimate their strength because we know their appearance and the type of process which generates them. The next step of our work consists of trying to incorporate saliency maps into NR metrics. In particular,

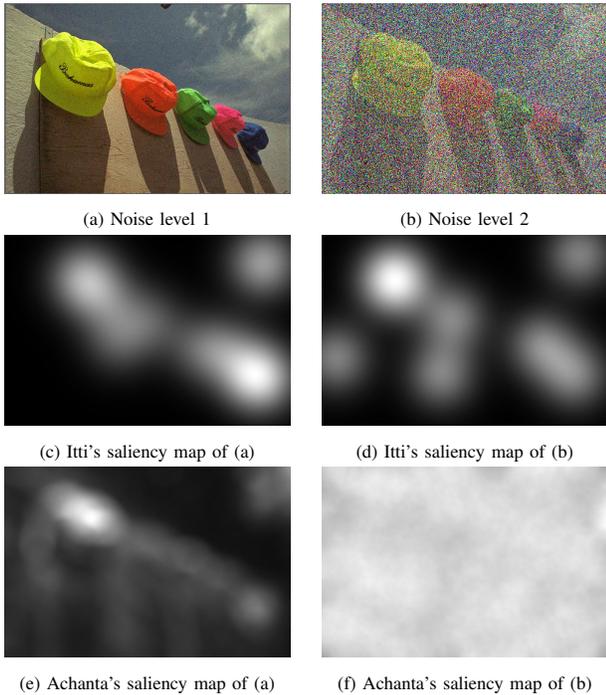


Fig. 3. Saliency maps of the image ‘Caps’ with two levels of noise, obtained using Itti’s and Achanta’s model.

we tested two artifact strength NR metrics: a blurriness and a blockiness metrics.

Most of the existing blurriness metrics are based on the idea that blur makes the edges larger or less sharp [17]. In this work, we have implemented a NR blurriness metric which makes use of this very simple idea. The algorithm estimates the amount of blurriness by measuring the width of the edges in the frame. The first step consists of finding strong edges using the Canny edge detector algorithm [18]. The output of the Canny algorithm gives the magnitude of the edge pixels, $M(i, j)$, and their orientation, $O(i, j)$. We select only the strong edges of the frame and estimate the width of a particular edge by measuring the distance (in pixels) between two local extreme values of the function $M(i, j)$ in the direction given by the orientation of the edge. The blurriness measure for an image is obtained by averaging widths over all strong edges of this frame, as given by:

$$\text{Blur} = \sum_{\substack{i=0, j=0 \\ M(i, j) > 25}}^{N, M} \frac{\text{width}(i, j)}{L}. \quad (5)$$

To integrate the saliency maps with the described blurriness metric, we create a blurriness map where each pixel corresponds to its border width. If the pixel is not a border pixel, its border width is set to zero. Therefore, to integrate the saliency maps with the described blurriness metric, we use the same combination process described in the previous section (see eq. 4).

The blockiness metric used in this work calculates the blockiness strength by estimating luminance differences in

the image. We assume that the blocks used by the coding algorithm have $b_s \times b_s$ pixels, with $b_s = 8$. The first step of the algorithm consists of calculating vertical and horizontal spatial differences in the luminance component of the image. Given the luminance component $Y(i, j)$ of an image of size $M \times N$, the map of horizontal differences is given by:

$$D_h(i, j) = Y(i, j) - Y(i + 1, j); \quad (6)$$

while the map of vertical differences is given by:

$$D_v(i, j) = Y(i, j) - Y(i, j + 1); \quad (7)$$

To calculate the blockiness strength, the algorithm selects the pixels in the map of differences corresponding to the borders of the blocks, i.e. the pixels located at spatial positions that are multiples of b_s , resulting in difference maps $D_v^{b_s}$ and $D_h^{b_s}$. Then, the sum of the new horizontal and vertical difference maps are normalized, using the following equations

$$M_v = \frac{\sum_{i=1}^{M/b_s} \sum_{j=1}^M (D_v^{b_s}(i, j))}{\sum_{i=1}^M \sum_{j=1}^M (D_v(i, j))} \quad (8)$$

and

$$M_h = \frac{\sum_{i=1}^M \sum_{j=1}^{M/b_s} (D_h^{b_s}(i, j))}{\sum_{i=1}^M \sum_{j=1}^M (D_h(i, j))}. \quad (9)$$

Finally, the blockiness metric is calculated taking an average of the above expressions:

$$\text{Block} = \frac{M_v + M_h}{M \cdot N}. \quad (10)$$

Since this blockiness metric does not generate an ‘error’ map as a final step, we use as error maps the individual horizontal and vertical luminance difference maps of the selected pixels on the borders of the image ($D_v^{b_s}$ and $D_h^{b_s}$). To incorporate saliency information, we combine the saliency maps with the horizontal and vertical difference maps individually and, then, calculate the blockiness metric as described above.

TABLE IV
SPEARMAN CORRELATION COEFFICIENTS FOR NO-REFERENCE METRICS.

Model	JPEG	JPEG2k	GB	FF
Block	0.9336	0.5650	0.6103	0.6761
Su-Block	0.9317	0.5798	0.3169	0.6399
Blur	0.4775	0.3527	0.7052	0.4922
Su-Blur	0.7548	0.2459	0.7680	0.6575

Table IV shows the Spearman correlation coefficients for the blockiness and blurriness metrics, with and without the addition of *subjective saliency maps*. The results are separated by distortion type, since for the two artifact metrics used in this work it does not make sense to test the complete set or the white noise dataset. The only image dataset that contains blockiness is the JPEG one. As expected, the blockiness metric ‘Block’ shows a good correlation only for this dataset. Unfortunately, the addition of the subjective saliency maps did not improve its performance. The only case for which the performance improved was for the distortion JPEG2k. But, in

this case the correlation values for the metric without saliency maps was not big.

All datasets contain some type of blurriness. But, the Gaussian Blur dataset is, obviously, the one with greatest amount of blurriness. As expected, this is dataset for which the blurriness metric performed the best. The addition of saliency information did improve the performance. Interestingly, the addition of saliency maps was also able to improve the performance for two other datasets: JPEG and FastFading.

This is still a work under progress and the results obtained may be improved if different approaches are used to combine the saliency information with the NR metrics. Also different results may be obtained for other NR artifact metrics. In particular, the results obtained with the tested blockiness metrics were not good. The reason for this was probably the fact that its performance was already good for images with blockiness. Therefore, adding the saliency information interfered with the calculation of blockiness. On the other hand, metrics like the blurriness metric tested that do not have very good performance can benefit the addition of saliency information.

V. CONCLUSIONS

Our results show that the computational models were able to improve the performance of the image quality metrics tested. The computational model that presented the best performance was GAFFE with gains slightly lower than the subjective saliency maps. Nevertheless, the improvement in performance was higher for the simpler metrics (PSNR and MSE) than for the more complex metric (SSIM). The results also showed that the performance depended on distortion type, with White Noise presenting the lowest gains. We also tested the combination of saliency information with NR metrics. Although the initial results were not as good as for FR metrics, the performance seems to depend on the artifact type, the metric's algorithm, and the combination model.

ACKNOWLEDGMENT

This work was supported in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Brazil and in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) Brazil.

REFERENCES

- [1] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it?" *IEEE Signal Processing Magazine*, vol. 1, no. January, pp. 98–117, 2009.
- [2] S. Chikkerur, V. Sundaram, M. Reisslein, and L. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *Broadcasting, IEEE Transactions on*, vol. 57, no. 2, pp. 165–182, June 2011.
- [3] U. Engelke, H. Kaprykowsky, H.-J. Zepernick, and P. Ndjiki-Nya, "Visual attention in quality assessment," *Signal Processing Magazine, IEEE*, vol. 28, no. 6, pp. 50–59, Nov. 2011.
- [4] C. Oprea, I. Pirnóg, C. Paleologu, and M. Udrea, "Perceptual Video Quality Assessment Based on Salient Region Detection," in *Telecommunications, 2009. AICT '09. Fifth Advanced International Conference on*, May 2009, pp. 232–236.
- [5] J. Redi, H. Liu, P. Gastaldo, R. Zunino, and I. Heynderickx, "How to apply spatial saliency into objective metrics for jpeg compressed images?" in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, Nov. 2009, pp. 961–964.

- [6] U. Engelke, H. Kaprykowsky, H.-J. Zepernick, and P. Ndjiki-Nya, "Visual attention in quality assessment," *Signal Processing Magazine, IEEE*, vol. 28, no. 6, pp. 50–59, Nov. 2011.
- [7] H. Liu and I. Heynderickx, "Studying the added value of visual attention in objective image quality metrics based on eye movement data," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, Nov. 2009, pp. 3097–3100.
- [8] A. Moorthy and A. Bovik, "Visual importance pooling for image quality assessment," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 3, no. 2, pp. 193–201, April 2009.
- [9] L. Itti and C. Koch, "Computational modelling of visual attention," *Nature Reviews Neuroscience*, vol. 2, no. 3, pp. 194–203, 2001. [Online]. Available: http://www.nature.com/nrn/journal/v2/n3/abs/nrn0301_194a.html
- [10] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk, "Salient region detection and segmentation," in *Computer Vision Systems*, ser. Lecture Notes in Computer Science, A. Gasteratos, M. Vincze, and J. Tsotsos, Eds. Springer Berlin / Heidelberg, 2008, vol. 5008, pp. 66–75.
- [11] U. Rajashekar, I. van der Linde, A. Bovik, and L. Cormack, "Gaffe: A gaze-attentive fixation finding engine," *Image Processing, IEEE Transactions on*, vol. 17, no. 4, pp. 564–573, April 2008.
- [12] H. Liu and I. Heynderickx, "Tud image quality database: Eye-tracking release 1," 2009. [Online]. Available: http://mmi.tudelft.nl/iqlab/eye_tracking_1.html
- [13] H. Sheikh, Z. Wang, L. Cormack, and A. Bovik, "Live image quality assessment database release 2," <http://live.ece.utexas.edu/research/quality>.
- [14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [15] J. A. Redi, H. Liu, P. Gastaldo, R. Zunino, and I. Heynderickx, "How to apply spatial saliency into objective metrics for JPEG compressed images?" in *IEEE ICIP2009 International Conference on Image Processing*, Nov. 2009. [Online]. Available: http://mmi.tudelft.nl/pub/hantao/IEEE_ICIP09_JR.pdf
- [16] W. Hays, *Statistics for the social sciences*, 3rd ed. Madison Avenue, New York, N.Y.: LLH Technology Publishing, 1981.
- [17] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Application to JPEG2000," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 163–172, 2004.
- [18] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Jun 1986. [Online]. Available: <http://dx.doi.org/10.1109/TPAMI.1986.4767851>