

# A Saliency Dispersion Measure for Improving Saliency-Based Image Quality Metrics

Wei Zhang, Ralph R. Martin, and Hantao Liu

**Abstract**—Objective image quality metrics (IQMs) potentially benefit from the addition of visual saliency. However, challenges to optimizing the performance of saliency-based IQMs remain. A previous eye-tracking study has shown that gaze is concentrated in fewer places in images with highly salient features than in images lacking salient features. From this, it can be inferred that the former are more likely to benefit from adding a saliency term to an IQM. To understand whether these ideas still hold when using computational saliency instead of eye-tracking data, we first conducted a statistical evaluation using 15 state-of-the-art saliency models and 10 well-known IQMs. We then used the results to devise an algorithm, which adaptively incorporates saliency in IQMs for natural scenes, based on saliency dispersion. Experimental results demonstrate that this can give significant improvements.

**Index Terms**—Dispersion, image quality, saliency.

## I. INTRODUCTION

Image quality metrics (IQMs) lie at the heart of algorithms to automatically predict perceived image quality [1]. Reliably predicting image quality as perceived by humans, however, remains challenging [2]. A significant trend in IQM research [3]–[12] is to investigate the significance of visual attention, an important mechanism in the human visual system that allows effective selection of the most relevant information in a visual scene [13].

Psychophysical studies have been undertaken to understand visual attention in relation to image quality assessment [3]–[7]. Integrating visual attention data obtained from *eye tracking* leads to improved ability of IQMs. Eye tracking is cumbersome and impractical in many circumstances, however. A more realistic way to integrate visual attention into IQMs is to use *computational saliency*. State-of-the-art saliency-based IQMs [8]–[12] generally weight local distortions with local saliency, resulting in a more sophisticated means of quality prediction. However, determining optimal use of computational saliency in IQMs is not straightforward [11], [12].

Our previous research based on eye tracking [14] revealed that the inter-observer agreement (IOA) for human fixations—the degree of agreement between observers freely viewing the same visual stimulus—is strongly image content dependent. Furthermore, this measure predicts the extent to which a certain image may profit from adding saliency information to an IQM. As the observation also revealed from eye-tracking studies in [15] and [16], if an image has highly salient objects, then most viewers will concentrate their fixations around them, whereas if there is no obvious object of interest, viewers' fixations will appear as a more evenly distributed pattern. Thus, images with salient objects tend to have less variation in fixations between viewers (i.e., higher IOA) than images without

Manuscript received May 27, 2016; revised October 15, 2016 and December 14, 2016; accepted December 27, 2016. Date of publication January 10, 2017; date of current version June 4, 2018. This paper was recommended by Associate Editor G. Cheung.

W. Zhang is with the School of Telecommunications Engineering, Xidian University, Xi'an 710071, China (e-mail: wzhang@xidian.edu.cn).

R. R. Martin and H. Liu are with the School of Computer Science and Informatics, Cardiff University, Cardiff CF243AA, U.K. (e-mail: ralph@cs.cardiff.ac.uk; hantao.liu@cs.cardiff.ac.uk).

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/TCSVT.2017.2650910

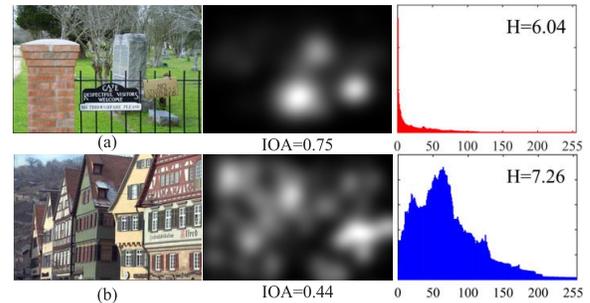


Fig. 1. Natural scenes, their ground truth saliency maps, corresponding IOA scores, and entropy calculated from the histogram of saliency intensity values. (a) Image with a few highly salient objects; IOA is high. (b) Image lacking salient objects; IOA is low. IOA values and saliency maps were determined from human eye fixations of 20 observers [3]. Entropy  $H$  of the saliency map was computed using (1).

salient objects. As illustrated in Fig. 1, images of the former kind [see Fig. 1(a)] have concentrated saliency maps, while the latter [see Fig. 1(b)] have more dispersed maps. When saliency is spread throughout the scene, incorporating saliency in an IQM is less likely to benefit image quality prediction [3], [14], as different observers tend to look at different parts of the image. Incorporating saliency into an IQM may give a low weight to some region with high distortion, and therefore weighting the IQM might unhelpfully downplay the importance of distortion in this region. To make better use of saliency in IQMs, a sophisticated integration strategy is needed, taking into account *image content*, particularly in terms of the dispersion of saliency.

The contributions of this paper are: 1) a statistical evaluation of whether conclusions concerning the content-dependent nature of benefits of adding saliency information to an IQM, determined from eye-tracking data, still hold when computational saliency is used in its place and 2) an algorithm that can provide a reliable proxy for IOA, for use in content adaptive IQM which incorporates computational saliency.

## II. EFFECT OF IMAGE CONTENT DEPENDENCE

Image content dependence of the improvement to IQMs by incorporating saliency has been demonstrated by use of eye-tracking data in [14]. In that study, ground truth saliency and IOA [calculated as the average correlation coefficient (CC) between the mean saliency map and each observer's saliency map] were measured for the LIVE database [17]; based on IOA for scene content, the entire database was divided into three subsets: images with low, intermediate, and high IOA. The measured saliency was integrated into three IQMs to assess quality of images within individual subsets. The result was that the average performance gain of IQMs increases as IOA increases.

A realistic IQM, however, will use a computational model of saliency rather than eye tracking. To determine whether content dependence still remains significant, and potentially useful, we conducted a statistical evaluation using 15 state-of-the-art saliency models and 10 of the best-known IQMs. The methodology established in [3] and [12] was used to assess the added value of computational saliency in IQMs: saliency is incorporated by weighting

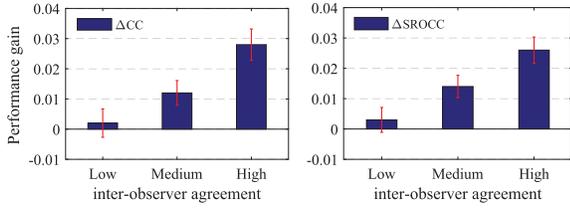


Fig. 2. Performance gain (i.e.,  $\Delta CC$  and  $\Delta SROCC$ ) of saliency-augmented IQMs for three degrees of IOA. Error bars indicate a 95% confidence interval.

the distortion map calculated by an IQM using the saliency map computed from the original scene. For each subset of images, we quantified the performance gain of a saliency-based IQM over its original form without saliency. The IQMs included 6 full-reference metrics, including PSNR [1], UQI [18], SSIM [19], MS-SSIM [20], VIF [21], and FSIM [22]; and four no-reference (NR) metrics, including GBIM [23], NBAM [24], NPBM [25], and JNBM [26]. The IQMs used were implemented in the spatial domain. Note that other well-known IQMs formulated in the transform domain (e.g., [27], [28]) were not included in our study but could be considered in the future work. Also note that specific advanced IQMs (see [9]) already incorporate well-established saliency aspects. Adding saliency information to these IQMs is not very meaningful, as it would duplicate some aspects of how they work already. As suggested in [12], for all NR metrics, saliency was computed from the original scene rather than the distorted scene. Such saliency was either assumed to be practically available [e.g., as a side information, in which case the framework is analogous to the reduced-reference case], or considered to be plausibly approximated from the distorted image (e.g., by filtering out distortion). The 15 saliency models were AIM [29], AWS [30], CA [31], CBS [32], DVA [33], GBVS [34], ITTI [35], PQFT [36], SDCD [37], SDFS [38], SDSR [39], SR [40], SUN [41], SVO [42] and Torralba [43], representing the best performing saliency models in terms of the capability of improving the performance of IQMs [12].

The study thus resulted in 150 possible combinations (10 IQMs  $\times$  15 saliency models). The performance of an IQM was quantified by the Pearson linear CC and Spearman rank order CCs (SROCC) between the IQM's output and the subjective quality ratings [44]. Fig. 2 illustrates the performance gain averaged over all 150 cases for different degrees of IOA. Results of *t*-test (preceded by a Kurtosis test for the assumption of normality [3]) show that the difference in performance gain between each pair of subsets is statistically significant with  $p < 0.05$  at the 95% confidence level. This confirms that the benefits of inclusion of computational saliency in IQMs depend on image content. For images with low IOA, incorporating saliency runs the risk of reducing IQM's performance (i.e., the performance gain can appear negative as shown in Fig. 2).

### III. PROPOSED SALIENCY DISPERSION MEASURE

To optimize the saliency integration by incorporating the above observation, we propose an algorithm to measure the saliency dispersion and use that as a proxy for the variation in human fixation (i.e., IOA) on natural scenes. Reliably quantifying saliency dispersion in agreement with IOA is very challenging, despite research on the topic. Existing methods have either limited sophistication (e.g., the simple saliency coverage measure in [15]) or limited applicability to real-world systems (e.g., the complex approaches in [45] and [46]). We have thus devised our own simple, but reliable, method.

Our method is based on Shannon entropy, which is a measure of the randomness or uncertainty of a variable [47]. We analyze saliency maps as realizations of random variables. Fig. 1(a) shows a

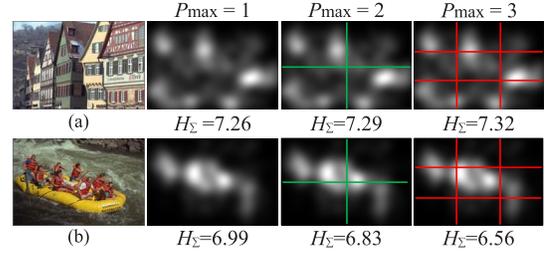


Fig. 3. Calculation of multilevel entropy  $H_{\Sigma}$ . At each level the saliency map is partitioned into blocks of equal size.  $H_{\Sigma}$  is found by adding the entropies computed at each level of partition.  $P_{\max}$  is the level with finest partitioning.

ground truth saliency map (grayscale values represent the intensity of saliency) of a natural scene derived by accumulating human fixations of 20 observers [3]. The normalized histogram of the saliency map represents an estimate of the underlying probabilities of pixel intensities:  $p(i) = h(i)/K$ , where  $h(i)$  is the histogram entry for intensity value  $i$  in the saliency map  $S$ , and  $K$  is the total number of pixels in  $S$ . The entropy of the saliency map is given by

$$H(S) = - \sum_i p(i) \log p(i). \quad (1)$$

For the saliency map in Fig. 1(a), it is 6.04 b. The entropy calculated for the saliency map of a different natural scene shown in Fig. 1(b) is 7.26 b. Saliency in Fig. 1(a) is more concentrated in fewer areas than in Fig. 1(b), which results in a smaller value of entropy.

Note, however, that even a single large salient object may also lead to a spread-out saliency map. For example, the saliency map in Fig. 3(b) is more concentrated than the saliency map in Fig. 3(a), but the entropy values are similar [i.e.,  $H = 7.26$  for Fig. 3(a) and  $H = 6.99$  for Fig. 3(b)]. This is because entropy is a single value summarizing the whole image; it does not consider spatial characteristics and relations of fixation patterns [48]. To perform a more refined saliency dispersion analysis, we use a multilevel approach to entropy calculation. To do so, the saliency map is partitioned at level  $P$  into  $P \times P$  nonoverlapping blocks of equal size (see Fig. 3). At  $P = 2$  the original map is subdivided into four equal quadrants, at  $P = 3$ , into 9 equal partitions, and so on. We then define the multilevel entropy of the saliency map to be

$$H_{\Sigma}(S) = \frac{1}{P_{\max}} \sum_{P=1}^{P_{\max}} \sum_{B=1}^{N_{\max}} H(B) \quad (2)$$

where  $P_{\max}$  is the finest level of division, and  $N_{\max} = P_{\max}^2$ ;  $B$  runs over each block. In the case illustrated in Fig. 3, the disparity in entropy between saliency maps increases as the number of partitions increases, which allows the multilevel entropy to better distinguish the two saliency maps than the whole-image entropy, giving the more compact saliency map a lower entropy.

To determine the number of levels to use, we use an empirical approach, based on quantifying the correlation between the estimated saliency dispersion and its ground truth counterpart (i.e., IOA). Fig. 4 plots the absolute value of the Pearson correlation between  $H_{\Sigma}$  for different choices of  $P_{\max}$ , and ground truth IOA values determined for the same set of images from three independent eye-tracking databases [14]. While correlation increases with  $P_{\max}$ , saturation starts to occur at about  $P_{\max} = 4$ . Hypothesis testing is performed to verify whether there is a significant difference between the use of  $P_{\max} = 4$  and a higher level of  $P_{\max}$ . A Wilcoxon signed rank test (i.e., a nonparametric version of *t*-test in the case of nonnormality) based on the residuals between  $H_{\Sigma}$  and IOA [12] was conducted; and the results (i.e.,  $p < 0.05$  at the 95% confidence level) showed that there was no statistically significant difference between  $P_{\max} = 4$

TABLE I

PERFORMANCE FOR 10 IQMs (IN TERMS OF CC, WITHOUT NONLINEAR REGRESSION) ON ALL IMAGES OF THREE DATABASES, USING VERSIONS WHICH DID NOT USE SALIENCY, ALWAYS USED SALIENCY, OR ADAPTIVELY USED SALIENCY ACCORDING TO SALIENCY DISPERSION. THE 95% CONFIDENCE INTERVALS FOR CC VALUES RANGE FROM 0.001 TO 0.006. NOTE THE LIVE RESULTS ARE BASED ON THE REALIGNED SUBJECTIVE DATA

		PSNR	UQI	SSIM	MS-SSIM	VIF	FSIM	GBIM	NPBM	JNBM	NBAM
CSIQ	original metric	0.751	0.829	0.765	0.815	0.883	0.805	0.731	0.814	0.793	0.625
	saliency-based metric	0.769	0.851	0.834	0.846	0.862	0.794	0.740	0.815	0.796	0.654
	adaptive-saliency-based metric	<b>0.782</b>	<b>0.876</b>	<b>0.852</b>	<b>0.858</b>	<b>0.891</b>	<b>0.809</b>	<b>0.757</b>	<b>0.831</b>	<b>0.811</b>	<b>0.666</b>
TID2013	original metric	0.478	0.615	0.653	0.707	0.608	0.820	0.841	0.787	0.794	0.765
	saliency-based metric	0.485	0.668	0.694	0.740	0.562	0.822	0.859	0.773	0.788	0.799
	adaptive-saliency-based metric	<b>0.497</b>	<b>0.687</b>	<b>0.728</b>	<b>0.759</b>	<b>0.587</b>	<b>0.828</b>	<b>0.875</b>	<b>0.798</b>	<b>0.803</b>	<b>0.811</b>
LIVE	original metric	0.859	0.898	0.825	0.830	0.945	0.859	0.773	0.843	0.833	0.836
	saliency-based metric	0.872	0.915	0.867	0.865	0.935	0.851	0.802	0.872	0.852	0.855
	adaptive-saliency-based metric	<b>0.883</b>	<b>0.929</b>	<b>0.882</b>	<b>0.887</b>	<b>0.952</b>	<b>0.872</b>	<b>0.815</b>	<b>0.886</b>	<b>0.866</b>	<b>0.874</b>

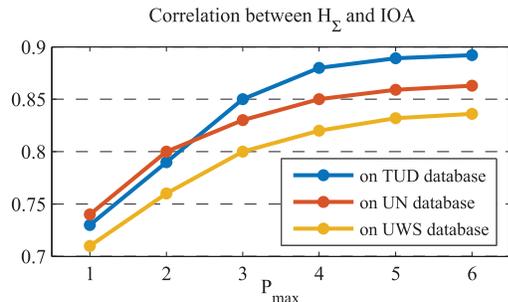


Fig. 4. Absolute value of the Pearson correlation (as shown for each data point) between estimated saliency dispersion,  $H_\Sigma$ , and its ground truth counterpart IOA, for different choices of  $P_{\max}$ . IOA values were determined for the same set of images from three independent eye-tracking databases [14].

and  $P_{\max} = 5$ , and between  $P_{\max} = 4$  and  $P_{\max} = 6$ . We therefore use  $P_{\max} = 4$  in our experiments.

#### IV. PROPOSED OPTIMIZATION METHOD

We now consider how to use the above formula for assessing saliency dispersion to improve saliency-based IQMs.

Suppose we are given a particular saliency model and an IQM. For an input scene of size  $M \times N$ , we can compute a saliency map together with its degree of dispersion  $H_\Sigma$ . The key idea is to *only include saliency in the computation of image quality if the dispersion is not too large*, in line with the observation that using saliency in cases of low IOA may be of no benefit to or even reduce the IQM performance.

In principle, we wish to do the following. If  $H_\Sigma$  is below a threshold  $T$ , saliency is combined with the pre-existing IQM to provide a modified method of quality assessment, as

$$I' = \sum_{x=1}^M \sum_{y=1}^N D(x, y)S(x, y) / \sum_{x=1}^M \sum_{y=1}^N S(x, y) \quad (3)$$

where  $D$  represents the distortion map measured by an IQM, and  $S$  indicates the saliency map generated by the saliency model. If the saliency dispersion is large, the saliency of the scene contains much uncertainty, and so is ignored: the pre-existing IQM is used directly without saliency.

However, using a hard threshold will lead to a discontinuous IQM, and two very similar scenes whose saliency dispersions are just above and below the threshold may end up with significantly different quality scores. To avoid such sudden changes, instead of using a step function to switch between using saliency, or not, a sigmoid function  $\sigma(\cdot)$  is applied to smooth the IQM near the transition region. Our integrated image quality metric  $I''$  is given by

$$I'' = \sigma(H_\Sigma)I + (1 - \sigma(H_\Sigma))I' \quad (4)$$

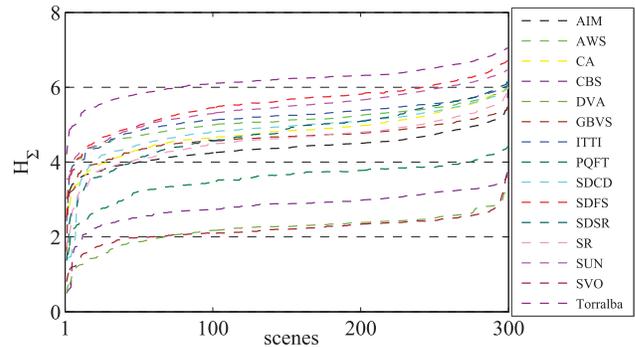


Fig. 5.  $H_\Sigma$  calculated for 300 scenes from the MIT300 database [49], using saliency values generated by 15 state-of-the-art saliency models.  $H_\Sigma$  values are ordered from lowest to highest for each model.

where  $I$  is the original IQM value and  $\sigma(x)$  is defined as

$$\sigma(x) = \frac{1}{1 + e^{-\tau(x-T)}} \quad (5)$$

where  $T$  is the threshold value and  $\tau$  controls the steepness of the sigmoid function.

As different saliency models lead to intrinsically different scales of entropy measurements (i.e., different ranges of  $H_\Sigma$  values),  $T$  should be individually determined for each saliency model. To ensure generality of the technique and to perform a more rigorous procedure to determine reliable parameters,  $\tau$  and  $T$  were empirically determined from a separate larger-scale saliency database to that used in our experiments; we used the MIT300 database [49] containing 300 natural scenes and a wide diversity of content. Fig. 5 gives  $H_\Sigma$  for these scenes, ordered from lowest to highest value, for the 15 saliency models considered in Section II. The median  $H_\Sigma$  value for each saliency model was used as the corresponding threshold  $T$  (e.g.,  $T = 4.38$  for AIM), while the slope of the envelope of the values between the 25th and 75th percentiles was used to determine an appropriate value of the steepness control  $\tau$ ; in practice, these were similar, so we used  $\tau = 20$  for all saliency models. Note that other saliency databases (see [50], [51]) may be used to estimate these parameters, but we do not expect it to change the results significantly.

#### V. EXPERIMENTAL RESULTS

The performance of each IQM was evaluated against three recognized image quality databases: CSIQ [28], TID2013 [52], and LIVE. In each case, we compared its performance between *no* use of saliency, *fixed* use of saliency, and *adaptive* use of saliency according to saliency dispersion. Table I shows the performance (in terms of CC) in each case, averaged over 15 saliency models (SROCC values exhibit similar trends and thus are not presented here). Following

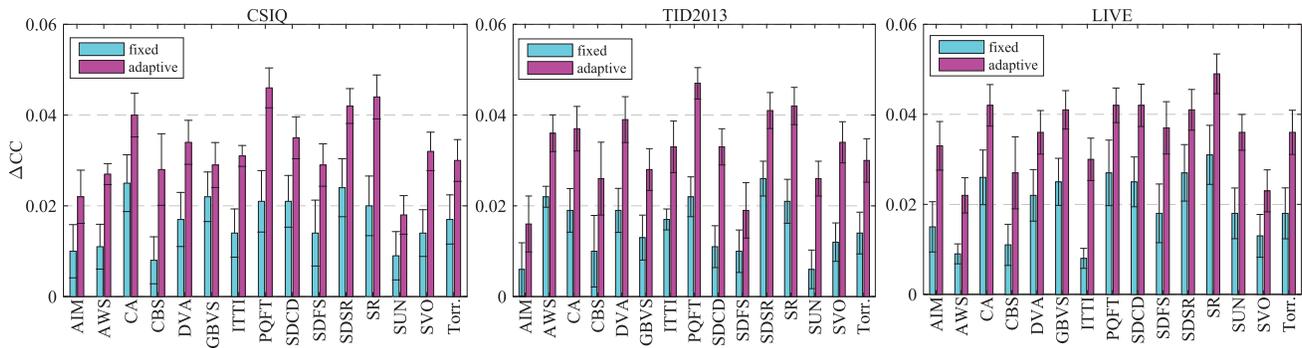


Fig. 6. Comparison of performance gain (i.e.,  $\Delta CC$ ) between saliency-augmented IQMs using fixed and adaptive use of saliency for each saliency models.

the approach taken in [12], CC values are reported without nonlinear fitting in order to better visualize differences in IQM performance. As can be seen, the adaptive approach outperforms fixed use of saliency in all cases. On average (over all databases), VIF and FSIM do not benefit from fixed use of saliency, but are improved by using adaptive saliency. Note that VIF and FSIM obtain relatively small gain by adding saliency. This is probably due to the fact that some well-established saliency aspects are already embedded in VIF and FSIM, which consequently causes a saturation effect in saliency optimization. More detailed results are given in Fig. 6, which shows the performance gain (i.e., increase in correlation when using fixed or adaptive saliency approaches), averaged over all IQMs, for individual saliency models. On average, the gain achieved by adaptive use of saliency is more than double that of always using saliency. As well as the observed *relative* difference in gain, Fig. 6 also gives the *absolute* gain of the adaptive approach for individual saliency models—this can be easily used to decide which of these models are more useful for IQMs. For example, by applying a threshold  $\Delta CC = 0.04$  to all databases picks out the good models to be PQFT, SDSR, SR. However, we again note that the purpose of this paper is not to find the best IQM (or to target specific IQMs), but rather to compare *fixed* use of saliency to *adaptive* use of saliency according to saliency dispersion.

A paired sample *t*-test analysis (preceded by a test for the assumption of normality) was performed, selecting the integration strategy as the independent variable and the performance as the dependent variable. Using the  $150 \times 2 \times 3$  data points contained in Table I demonstrated with  $p < 0.01$  at the 95% confidence level that an adaptive strategy is statistically significantly better than fixed inclusion of saliency.

## VI. CONCLUSION

This paper considered how to reliably measure saliency dispersion in natural scenes and how it can be used to adaptively incorporate computational saliency into image quality metrics. Results show that adaptive use of saliency according to saliency dispersion significantly outperforms the fixed use of saliency in improving IQMs. We also intend to investigate the dependence of gain on saliency model dependence to maximize the IQM's performance as future work.

## REFERENCES

- [1] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment*. San Rafael, CA, USA: Morgan Kaufmann, 2006.
- [2] K. Ma *et al.*, "Group MAD competition—A new methodology to compare objective image quality models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1664–1673.
- [3] H. Liu and I. Heynderickx, "Visual attention in objective image quality assessment: Based on eye-tracking data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 7, pp. 971–982, Jul. 2011.
- [4] E. C. Larson, C. Vu, and D. M. Chandler, "Can visual fixation patterns improve image fidelity assessment?" in *Proc. 15th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2008, pp. 2572–2575.
- [5] A. Ninassi, O. Le Meur, P. Le Callet, and D. Barba, "Does where you gaze on an image affect your perception of quality? Applying visual attention to image quality metric," in *Proc. 14th IEEE Int. Conf. Image Process.*, vol. 2, Sep. 2007, pp. II-169–II-172.
- [6] J. Redi, H. Liu, R. Zunino, and I. Heynderickx, "Interactions of visual attention and quality perception," *Proc. SPIE*, vol. 7865, p. 78650S, Jan. 2011.
- [7] F. Röhrbein, P. Goddard, M. Schneider, G. James, and K. Guo, "How does image noise affect actual and predicted human gaze allocation in assessing image quality?" *Vis. Res.*, vol. 112, pp. 11–25, Jul. 2015.
- [8] A. K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp. 193–201, Apr. 2009.
- [9] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Aug. 2014.
- [10] D. V. Rao, N. Sudhakar, I. R. Babu, and L. P. Reddy, "Image quality assessment complemented with visual regions of interest," in *Proc. Int. Conf. Comput. Theory Appl.*, Mar. 2007, pp. 681–687.
- [11] U. Engelke, H. Kaprykowsky, H. Zepernick, and P. Ndjiki-Nya, "Visual attention in quality assessment," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 50–59, Nov. 2011.
- [12] W. Zhang, A. Borji, Z. Wang, P. Le Callet, and H. Liu, "The application of visual saliency models in objective image quality assessment: A statistical evaluation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1266–1278, Jun. 2016.
- [13] A. Borji, D. N. Sihite, and L. Itti, "Quantitative analysis of human-model agreement in visual saliency modeling: A comparative study," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 55–69, Jan. 2013.
- [14] H. Liu, U. Engelke, J. Wang, P. Le Callet, and I. Heynderickx, "How does image content affect the added value of visual attention in objective image quality assessment?" *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 355–358, Apr. 2013.
- [15] D. S. Wooding, "Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps," *Behavior Res. Methods Instrum. Comput.*, vol. 34, no. 4, pp. 518–528, 2002.
- [16] M. Mancas and O. L. Meur, "Memorability of natural scenes: The role of attention," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 196–200.
- [17] L. C. H. R. Sheikh, Z. Wang, and A. Bovik, *LIVE Image Quality Assessment Database Release 2*, accessed 2006. [Online]. Available: <http://live.ece.utexas.edu/research/quality>
- [18] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.

- [23] H. R. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," *IEEE Signal Process. Lett.*, vol. 4, no. 11, pp. 317–320, Nov. 1997.
- [24] R. Muijs and I. Kirenko, "A no-reference blocking artifact measure for adaptive video processing," in *Proc. 13th Eur. Signal Process. Conf.*, Antalya, Turkey, Sep. 2005, pp. 1–4.
- [25] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," in *Proc. 9th IEEE Int. Conf. Image Process.*, Rochester, NY, USA, Sep. 2002, pp. 57–60.
- [26] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, Apr. 2009.
- [27] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.
- [28] 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. 3267105, 2010.
- [29] N. D. B. Bruce and J. K. Tsotsos, "Saliency, attention, and visual search: An information theoretic approach," *J. Vis.*, vol. 9, no. 3, pp. 1–5, 2009.
- [30] A. Garcia-Diaz, X. R. Fdez-Vidal, X. M. Pardo, and R. Dosiil, "Saliency from hierarchical adaptation through decorrelation and variance normalization," *Image Vis. Comput.*, vol. 30, no. 1, pp. 51–64, Jan. 2012.
- [31] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 1915–1926, Oct. 2012.
- [32] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, and S. Li, "Automatic salient object segmentation based on context and shape prior," in *Proc. 22nd Brit. Mach. Vis. Conf.*, Dundee, U.K., Sep. 2011, pp. 1–7.
- [33] X. Hou and L. Zhang, "Dynamic visual attention: Searching for coding length increments," in *Proc. 22nd Annu. Conf. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, Dec. 2008, pp. 681–688.
- [34] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Proc. 20th Conf. Adv. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, Dec. 2006, pp. 545–552.
- [35] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.
- [36] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 185–198, Jan. 2010.
- [37] Y. Fang, Z. Chen, W. Lin, and C.-W. Lin, "Saliency detection in the compressed domain for adaptive image retargeting," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3888–3901, Sep. 2012.
- [38] J. Li, M. D. Levine, X. An, and H. He, "Saliency detection based on frequency and spatial domain analysis," in *Proc. 22nd Brit. Mach. Vis. Conf.*, Dundee, U.K., Sep. 2011, pp. 1–11.
- [39] H. J. Seo and P. Milanfar, "Static and space-time visual saliency detection by self-resemblance," *J. Vis.*, vol. 9, no. 12, p. 15, 2009.
- [40] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Minneapolis, MN, USA, Jun. 2007, pp. 1–8.
- [41] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *J. Vis.*, vol. 8, no. 7, p. 32, Dec. 2008.
- [42] K.-Y. Chang, T.-L. Liu, H.-T. Chen, and S.-H. Lai, "Fusing generic objectness and visual saliency for salient object detection," in *Proc. 14th IEEE Int. Conf. Comput. Vis.*, Barcelona, Spain, Nov. 2011, pp. 914–921.
- [43] A. Torralba, "Modeling global scene factors in attention," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 20, no. 7, pp. 1407–1418, 2003.
- [44] "Final report from the video quality experts group on the validation of objective models of video quality assessment," VQEG, Tech. Rep., 2000.
- [45] O. Le Meur, T. Baccino, and A. Roumy, "Prediction of the inter-observer visual congruency (IOVC) and application to image ranking," in *Proc. 19th ACM Int. Conf. Multimedia*, 2011, pp. 373–382.
- [46] S. Rahman and N. D. B. Bruce, "Factors underlying inter-observer agreement in gaze patterns: Predictive modelling and analysis," in *Proc. 9th Biennial ACM Symp. Eye Tracking Res. Appl.*, 2016, pp. 155–162.
- [47] R. M. Gray, *Entropy and Information Theory*. New York, NY, USA: Springer, 1990.
- [48] M. R. Sabuncu, "Entropy-based image registration," Ph.D. dissertation, Dept. Elect. Eng., Princeton Univ., Princeton, NJ, USA, 2006.
- [49] Z. Bylinskii *et al.* *MIT Saliency Benchmark*, accessed 2012. [Online]. Available: <http://saliency.mit.edu/>
- [50] G. Kootstra and L. Schomaker, "Prediction of human eye fixations using symmetry," in *Proc. 31st Annu. Conf. Cog. Sci.*, 2009, pp. 56–61.
- [51] T. Judd, K. Ehinger, F. Durand, and A. Torralba, "Learning to predict where humans look," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep. 2009, pp. 2106–2113.
- [52] N. Ponomarenko *et al.*, "Color image database TID2013: Peculiarities and preliminary results," in *Proc. 4th IEEE Eur. Workshop Vis. Inf. Process. (EUVIP)*, Jun. 2013, pp. 106–111.