A STUDY ON USING SPECTRAL SALIENCY DETECTION APPROACHES FOR IMAGE QUALITY ASSESSMENT

Ashirbani Saha and Q. M. Jonathan Wu

Department of Electrical and Computer Engineering, University of Windsor, Canada

ABSTRACT

Recent developments in the field of full reference image quality assessment (FR-IQA) have witnessed the use of spectral residual (SR) based index as a fast measure with high accuracy. Following SR, several variants of spectral measures for visual saliency have come up. These new measures differ in their computational times as well as in performances and have established themselves better than or competitive with SR as measures of visual saliency. The effectiveness of these measures in FR-IQA is still an open question. In this paper, a study to evaluate the performance of the recent spectral approaches for visual saliency (hence spectral saliency) for FR-IQA is presented. We have fixed a framework for FR-IQA to maintain uniformity in the evaluation process. Also, the parameters required by the framework are chosen to bring out the best potential of each measure. Our experiments on six benchmark databases reveal some insightful details about the usage of these measures to form an FR-IQA measure.

***Index Terms*—** Image quality assessment, visual saliency, spectral residual

1. INTRODUCTION

The field of image quality assessment (IQA) aims to predict the quality of any image such that the prediction is consistent with subjective evaluation of the image by the human visual system (HVS). There has been substantial proliferation in this research area due to its variegated applicability in different areas of image processing. There exists clear distinction among the different types of IQA methods, based on the availability of the reference image. Full reference image quality assessment (FR-IQA) is a special type of IQA where the original or reference image is required to evaluate the quality of a query/test image. Since the earlier decade, a lot of research work has been dedicated towards finding more accurate and faster FR-IQA techniques [1]. Also, several public IQA databses have been made available with the research community. However, no FR-IQA index/measure has been found to be consistently performing well across all the databases or across all the distortions. The state-of-the-art FR-IQA methods, therefore, include different types of indices, each with their own characteristic advantages and limitations. The most popular and widely used FR-IQA index is SSIM [2]. Though, computationally fast, it suffers from low accuracy while evaluating the perceptual quality of blurred images. Though, GSSIM [3] improves this limitation, it has overall low accuracy compared to SSIM. Other variants of SSIM include MSSSIM [4] and IWSSIM [5] which are multiscale techniques, hence computationally slower than SSIM but perform with higher accuracy in in all databases. Apart from these structural similarity based approaches, there exists information theory based index called VIF [6] which is computationally slower but highly accurate. Another notable method is MAD [7] which is also computationally expensive but demonstrates better consistency.

Among the recent FR-IQA indices, FSIM/FSIM$_c$ [8] has the highest overall accuracy. The same framework used in FSIM has been deployed for the development of SR-SIM [9] which performs very well in three largest public IQA databases available and has high computational speed. SR-SIM is based on SR [10] which is one of the state-of-the-art techniques in visual saliency detection. Spectral residual has been used earlier for FR-IQA but embedded in a different framework [11, 12]. Other saliency based FR-IQA techniques use visual saliency for the purpose of pooling in the later stages of the algorithm [5]. Since visual attention is one of the integral properties of HVS which the IQA methods aim to imitate, the justification of using image saliency for IQA is intuitive. Inspired by the potential and performance of SR for FR-IQA, this paper attempts to explore and evaluate the performance of several recently developed spectral saliency detection methods in FR-IQA. The spectral methods selected are those which have gradually evolved being inspired from SR and showed better or competitive performance with the same. The current work is an effort to study how these newly emerged spectral saliency methods support FR-IQA and evaluate their performance by comparison with the state-of-the-arts and SR. For the evaluation purpose we have selected an existing framework of FR-IQA. The selected framework and the spectral saliency based methods selected for this study are described in sections 2 and 3. The experiments and results are discussed in section 4. We draw conclusions and present future research directions in section 5.

The work is supported in part by the Canada Research Chair program, and the Natural Sciences and Engineering Research Council of Canada.

2. BASIC FRAMEWORK

In this section, we discuss the framework selected for the purpose of evaluation. The block diagram of the framework is shown in Fig. 1. This is the same framework deployed in FSIM and SR-SIM. The inherent simplicity and the success of FSIM and SR-SIM are the main reasons for choosing this framework. The FSIM and SR-SIM methods work on luminance component (Y of YIQ color space) of images. The variant of FSIM called FSIM$_c$ separately processes the chrominance channels (I and Q) and combines chrominance comparison maps with FSIM map. In the framework used here, the preprocessing steps involve a conversion of colorspace (if required), followed by the image scale selection procedure mentioned in [13]. Some of the spectral saliency methods selected here work on color images only and for them YIQ space is used. For the spectral methods that work on grayscale images, the luminance component Y is used for color images to find the saliency maps similar to SR-SIM. After the preprocessing steps, the saliency maps (SM_R and SM_T) and the gradient maps (G_R and G_T) of the reference (I_R) and test images (I_T) are extracted. The saliency and gradient similarity maps (SM_{sim} and G_{sim}) are computed as

$$SM_{sim}(i) = \frac{2SM_R(i)SM_T(i) + K_1}{SM_R^2(i) + SM_T^2(i) + K_1} \tag{1}$$

and

$$G_{sim}(i) = \frac{2G_R(i)G_T(i) + K_2}{G_R^2(i) + G_T^2(i) + K_2}, i \in N \tag{2}$$

Here, K_1 and K_2 are two parameters which are positive constants used to improve the stability of the equations and N is the total number of pixel positions in the saliency maps. These maps are combined to form the similarity maps (SIM_{map}) between the two images as

$$SIM_{map} = SM_{sim}G_{sim}^{\beta}w \tag{3}$$

where $w = \max(SM_R, SM_T)$ and β is the weight of the gradient saliency maps. Finally, the quality score (Q_S) is calculated as

$$Q_s = \frac{\sum_{i \in N} SIM_{map}(i)}{\sum_{i \in N} w(i)}. \tag{4}$$

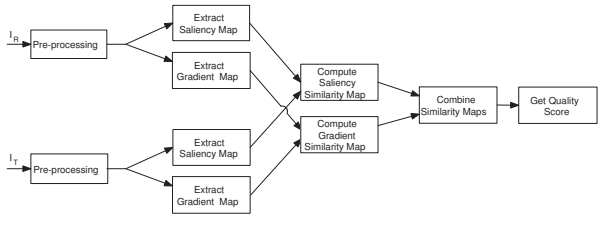

Fig. 1. Block Diagram of the framework

3. SELECTED SPECTRAL APPROACHES

During the evaluation process, we change the approach for saliency extraction method and generate the quality scores. Here, we discuss the spectral saliency methods used during the evaluation. The first approach is SR-SIM itself. SR uses the residual log amplitude (computed as the difference of log amplitude and low-pass filtered log amplitude) and phase of the Fourier transform to generate the saliency maps. The second approach is Phase Fourier Transform saliency (PFT) [14] which uses the only the phase of Fourier transform to compute the saliency maps. As shown in [14], a faster and better computation of saliency map is obtained using PFT. Further innovation led towards the use of phase of the Quaternion Fourier transform (PQFT) to form saliency maps. For the application of Quaternion Fourier transform, the color channels are required and hence this method works on color images only. Another recent technique uses Eigen SR and Eigen PQFT [15] for spectral saliency determination. Eigen PQFT and Eigen SR differs from PQFT and SR respectively in using eigen axes and eigen angles of the quaternion. Another saliency detection approach using Hypercomplex Fourier transform (HFT) based saliency detection has been recently introduced in [16]. The scale at which the saliency map will be perfect is decided by the saliency entropy criteria.

Apart from Fourier Transform, an approach has used discrete cosine transform (DCT) to generate features for image saliency called image signature [17]. This method can extract saliency from both color and grayscale images. This work was further extended to Quaternion DCT (QDCT) [18] for image saliency detection. All of the aforementioned spectral saliency detection methods have been chosen for this study. Therefore, altogether seven saliency determination algorithms have been used to obtain the quality scores from the images.

4. EXPERIMENTAL EVALUATION

In this section, we discuss on two topics. Since, the basic framework consists of the parameters K_1, K_2 and β, the values of these parameters are required for the evaluation. The evaluation is carried out using six benchmark databases : LIVE database [19], CSIQ database [7], A57 database [20], TID2008 database [21], IVC database [22] and Toyama database [23]. The subjective assessment scores in all of these databases have been given in terms of Mean Opinion Score (MOS) or Differential MOS (DMOS) [19]. The five evaluation measures used for the quantitative analysis : Spearman's Rank-Order Correlation Coefficient (SROCC), Kendall's Rank-Order Correlation Coefficient (KROCC), Pearson's Linear Correlation Coefficient (PLCC), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). SROCC and KROCC are useful for judging the $prediction\ monotonicity$ of any IQA method. Higher values of SROCC and KROCC indicate that objective assessment scores are more consis-

tent with MOS/DMOS values and therefore the assessment method is better. On the other hand, PLCC is measure of $prediction\ accuracy$ whereas MAE and RMSE are the measures of $prediction\ error$. Therefore, higher values of PLCC and lower values of MAE and RMSE are preferable. For PLCC, MAE and RMSE, a 5-parameter logistic function [24] is used for mapping between the subjective and objective scores.

4.1. Parameter Selection

First, we fix the value of β as it decides the contribution of gradient similarity maps towards forming the SIM_{map}. The value of β is taken to be as 0.5 similar to that mentioned in [9]. Next, we followed a procedure for the selection of the remaining parameters K_1 and K_2 used to find the corresponding similarity maps. The value of K_1 is varied from 0.05 to 1 with an interval of 0.05 and hence 19 different values are obtained. K_2 is used for gradient images and its value is varied from 10 to 250 with an interval of 10. Therefore, for each value of K_1, we have 25 values of K_2. Together we have 19×25 = 475 different pairs for the values of (K_1, K_2) each of which is denoted by an index value (j) varying from 1 to 475. Starting from the beginning, if a non-overlapping set of 25 indices are formed, each of those sets correspond to a fixed value of K_1 and 25 different values of K_2. The maximum value SROCC can take is 1. For every parameter index, we calculate the mean of squares of deviation of corresponding SROCC value from 1 across all databases. We call this quantity 'correlation error' CE_j which is calculated for the j^{th} parameter index as

$$CE_j = \left(\frac{1}{N_D} \sum_{d=1}^{N_D} (1 - SROCC_{d,j})^2 \right)^{0.5}, \qquad (5)$$

where d represents each of the N_D databases and $SROCC_{d,j}$ is the SROCC value in d^{th} database for j^{th} parameter set. The plot of correlation error for all the parameter indices using different types of spectral saliency based FR-IQA methods is shown in Fig. 2. The K_1 and K_2 values are chosen corresponding to index (j) for which CE_j is minimum. This experiment reveals different sets of parameter values for different saliency based approaches as shown in Table 1. In this method of selecting parameters, equal importance has been given to all of the databases which vary a lot in size, subjective score range and in types of distortion present in images. Since, the FR-IQA measures need to be consistent across several types of distortions, the parameters chosen in this way can have better applicability with many distortions. The range of the correlation error plots is also significant. A small range assures that the variation in parameter values has less effect on the measure whereas a bigger range shows that Q_s is more dependent on the parameters. Again, the lesser the higher limits of CE, the closer is the performance of the measure to perfection. The least higher limit in CE is exhibited by PQFT. Though the selected framework is embedded in the FSIM and SR-SIM, it is quite dependent on the parameters.

Table 1. Parameter values used with different measures

	PFT	PQFT	Eigen SR	Eigen PQFT	HFT	DCT	QDCT
K_1	0.35	0.05	0.05	0.05	0.95	0.95	0.95
K_2	70	50	20	30	220	10	10

4.2. Comparison Results

We have compared the seven aforesaid saliency based FR-IQA measures with each other and also with state-of-the-art measures MSSSIM, IWSSIM, VIF, MAD and FSIM$_c$. Since the quaternion based approaches process color images only, there are no results for these approaches in A57 database which contains grayscale images only. The results for FSIM$_c$ is equal to those obtained using FSIM for grayscale images in A57 database. For SR based IQA, we have used the parameters used in the implementation of SR-SIM [9]. The comparison between different IQA measures is summarized in Table 2. SR-SIM has the best results for TID database. PFT performs better in TID and IVC database. PQFT gives a consistent performance, though it has never been among the top two performers except for one database. It performs consistently well in all databases. Eigen SR performs better than SR-SIM in LIVE, IVC and TOYAMA datasets. In TID, IVC and TOYAMA database, Eigen PQFT performs better than PQFT. HFT and DCT based FR-IQA measures vary a lot from the best results of each database. However, the performance of QDCT based measure is much better in all of the databases. In Table 3, we show the direct and weighted average SROCC values across all databases obtained using different FR-IQA methods. We find that saliency based methods, SR-SIM, PFT, PQFT and Eigen PQFT based methods have higher average SROCC. In [9], SR-SIM was shown to be faster than several FR-IQA techniques. PFT calculation is faster than SR [14]. PQFT, Eigen SR and Eigen PQFT involve quaternion fourier transform and hence are computationally expensive compared to PFT. Therefore, among the methods that have competitive performance with the state-of-the-arts, PFT is the fastest.

5. CONCLUSION

We have presented a comparative study of spectral saliency based approaches for the purpose of FR-IQA. This study enables us to see that several aspects of saliency based FR-IQA. Firstly, the raw values present in the saliency maps formed are very important and though they may form better saliency maps or help in better detection of salient areas, they may not help to design a better FR-IQA measure. Secondly, we show the dependence of the framework (used in FSIM$_c$ and SR-SIM) on the values of the parameters. Thirdly, this study enabled us to evaluate the relative capabilities of the spectral saliency based methods to form an FR-IQA measure. The future work involves using suitable saliency measures for FR-IQA via a framework which has lesser dependence on parameters.

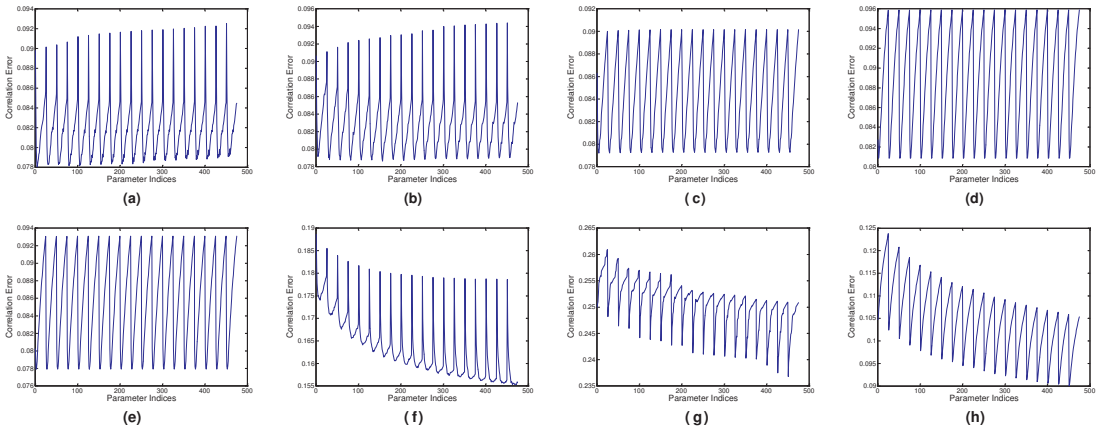

Fig. 2. Correlation Error across the databases for (a) SR, (b) PFT, (c) PQFT, (d) Eigen SR, (e) Eigen PQFT, (f) HFT, (g) DCT and (h) QDCT

Table 2. Performance of different spectral saliency approaches used in the IQA measure

		SR-SIM	PFT	PQFT	Eigen SR	Eigen PQFT	HFT	DCT	QDCT	MSSSIM	IWSSIM	VIF	MAD	$FSIM_c$
	SROCC	0.9618	0.9647	**0.9659**	0.9629	0.9651	0.9403	0.9365	0.9467	0.9513	0.9566	0.9636	**0.9669**	0.9645
	KROCC	0.8299	0.8357	**0.8374**	0.8299	0.8354	0.7796	0.7697	0.8001	0.8049	0.8178	0.8282	**0.8421**	0.8365
LIVE	PLCC	0.9552	0.9615	0.9635	0.9619	**0.9638**	0.9389	0.9336	0.9429	0.9489	0.9519	0.9604	**0.9675**	0.9613
	MAE	6.3263	5.8746	5.6067	5.7533	**5.5954**	7.5383	8.0128	7.3918	6.6701	6.3805	6.0952	**5.2071**	5.8236
	RMSE	8.0792	7.5093	7.3135	7.4662	**7.2821**	9.4050	9.7837	9.0955	8.6143	8.3757	7.6089	**6.9068**	7.5236
	SROCC	0.9319	0.9392	0.9344	0.9279	0.9319	0.9179	0.8570	**0.9454**	0.9132	0.9212	0.9194	**0.9466**	0.9309
	KROCC	0.7719	0.7827	0.7728	0.7631	0.7698	0.7401	0.6634	**0.7933**	0.7386	0.7522	0.7532	**0.7963**	0.7684
CSIQ	PLCC	0.9241	0.9324	0.9227	0.9165	0.9202	0.9088	0.8854	**0.9526**	0.8986	0.9142	0.9277	**0.9505**	0.9186
	MAE	0.0734	0.0694	0.0726	0.0756	0.0735	0.0806	0.0974	**0.0607**	0.0857	0.0789	0.0742	**0.0631**	0.0745
	RMSE	0.1021	0.0960	0.1032	0.1075	0.1052	0.1098	0.1221	**0.0801**	0.1168	0.1082	0.098	**0.0817**	0.1060
	SROCC	**0.9295**	0.9170	-	-	-	-	0.5205	-	0.8435	0.7750	0.6224	0.9023	**0.9181**
	KROCC	**0.7779**	0.7583	-	-	-	-	0.3610	-	0.6529	0.5880	0.4592	0.7233	**0.7639**
A57	PLCC	**0.9247**	0.9177	-	-	-	-	0.5344	-	0.8394	0.7652	0.6137	0.9043	**0.9252**
	MAE	**0.0778**	0.0818	-	-	-	-	0.1592	-	0.1119	0.1182	0.1417	0.0856	**0.0794**
	RMSE	**0.0936**	0.0978	-	-	-	-	0.2079	-	0.1337	0.1587	0.1957	0.1051	**0.0933**
	SROCC	**0.8913**	**0.8870**	0.8826	0.8797	0.8855	0.8279	0.7491	0.8544	0.8528	0.8559	0.7496	0.8340	0.8840
	KROCC	**0.7149**	**0.7030**	0.6952	0.6948	0.7013	0.6287	0.5580	0.6636	0.6543	0.6636	0.5863	0.6445	0.6991
TID	PLCC	**0.8854**	**0.8827**	0.8742	0.8777	0.8813	0.8154	0.7946	0.8660	0.8419	0.8572	0.8075	0.8311	0.8758
	MAE	**0.4543**	**0.4782**	0.4979	0.4872	0.4814	0.5905	0.6531	0.5153	0.5616	0.5245	0.5837	0.5543	0.4868
	RMSE	**0.6246**	**0.6307**	0.6522	0.6433	0.6344	0.7771	0.8149	0.6712	0.7247	0.6915	0.8007	0.7491	0.6482
	SROCC	0.9265	**0.9402**	0.9322	**0.9361**	0.9350	0.7515	0.8953	0.9212	0.8980	0.9125	0.8964	0.9146	0.9293
	KROCC	0.7560	**0.7809**	0.7671	**0.7739**	0.7717	0.5601	0.7083	0.7561	0.7203	0.7339	0.7158	0.7406	0.7636
IVC	PLCC	0.9357	**0.9495**	0.9432	**0.9488**	0.9470	0.7511	0.9073	0.9301	0.9106	0.9228	0.9026	0.9210	0.9390
	MAE	0.3404	**0.3050**	0.3188	**0.3130**	0.3150	0.5877	0.4096	0.3267	0.3701	0.3684	0.4050	0.3673	0.3277
	RMSE	0.4306	**0.3826**	0.4052	**0.3850**	0.3918	0.8137	0.5126	0.4509	0.5103	0.4704	0.5276	0.4753	0.4203
	SROCC	0.8825	0.9054	0.9133	0.9131	0.9154	0.8624	0.9095	0.9145	0.8874	**0.9202**	0.9077	**0.9362**	0.9067
	KROCC	0.6975	0.7303	0.7421	0.7426	0.7462	0.6758	0.7381	0.7434	0.7029	**0.7537**	0.7315	**0.7823**	0.7303
TOYAMA	PLCC	0.8870	0.9056	0.9153	0.9164	0.9182	0.8599	0.9144	0.9198	0.8924	**0.9246**	0.9136	**0.9367**	0.9072
	MAE	0.4444	0.3950	0.3792	0.3755	0.3705	0.4835	0.3804	0.3671	0.4287	**0.3653**	0.4012	**0.3493**	0.4015
	RMSE	0.5782	0.5317	0.5053	0.5017	0.4968	0.6398	0.5071	0.4924	0.565	**0.4775**	0.5096	**0.4384**	0.5275

Table 3. Average performance of different spectral saliency approaches used in the IQA measure

	SR-SIM	PFT	PQFT	Eigen SR	Eigen PQFT	HFT	DCT	QDCT	MSSSIM	IWSSIM	VIF	MAD	$FSIM_c$
Direct Avg.	0.9206	0.9256	**0.9257**	0.9239	**0.9266**	0.8600	0.8113	0.9164	0.8910	0.8902	0.8432	0.9168	0.9223
Weight Avg.	**0.9170**	**0.9189**	0.9161	0.9127	0.9169	0.8709	0.8241	0.9012	0.8908	0.8964	0.8458	0.8972	0.9152

6. REFERENCES

[1] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[2] 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.

[3] G.-H. Chen, C.-L. Yang, and S.-L. Xie, "Gradient-based structural similarity for image quality assessment," in *Int. Conf. on Image Processing*, 2006, pp. 2929–2932.

[4] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *IEEE Asilomar Conf. on Signals, Systems, and Computers*, 2003, pp. 1398–1402.

[5] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp. 1185–1198, 2011.

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

[7] E. C. Larson and D.M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 011006, 2010.

[8] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. on Image Processing*, vol. 20, no. 8, pp. 1–26, 2011.

[9] L. Zhang and H. Li, "SR-SIM: A fast and high performance IQA index based on spectral residual," in *International Conference on Image Processing*, 2012, pp. 1473–1476.

[10] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Computer Vision and Pattern Recognition (CVPR07). IEEE Computer Society*, 2007, pp. 1–8.

[11] Q. Ma and L. Zhang, "Saliency-based image quality assessment criterion," in *International Conference on Intelligent Computing,*, 2008, pp. 1124–1133.

[12] Q. Ma and L. Zhang, "Image quality assessment with visual attention," in *International Conference on Pattern Recognition*, 2008, pp. 1–4.

[13] Z. Wang, "SSIM index for image quality assessment [Online].," 2003, Available: http://www.ece.uwaterloo.ca/ z70wang/research/ssim/.

[14] C. Guo, Q. Ma, and L. Zhang, "Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform," in *Computer Vision and Pattern Recognition*, 2008, pp. 1–8.

[15] B. Schauerte and R. Stiefelhagen, "Quarternion-based spectral saliency detection for eye fixation prediction," in *European Conference on Computer Vision*, 2012.

[16] J. Li, M. D. Levine, X. An, X. Xu, and H. He, "Visual saliency based on scale-space analysis in the frequency domain," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012, http://www.cim.mcgill.ca/ lijian/saliency3.htm.

[17] X. Hou, J. Harel, and C. Koch, "Image signature: Highlighting sparse salient regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 194–201, 2012.

[18] B. Schauerte and R. Stiefelhagen, "Predicting human gaze using quaternion DCT image signature saliency and face detection," in *Workshop on the Applications of Computer Vision*, 2012.

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

[20] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise-ratio for natural images," *IEEE Trans. on Image Processing*, vol. 16, no. 9, pp. 2284–2298, 2007.

[21] N. Ponomarenko, V. Lukin, A. Zelensky, K. Eziazarian, M. Carli, and F. Battisti, "TID2008- A database for evaluation of full reference image quality metrics," *Advances of Modern Radioelectronics*, vol. 10, pp. 30–45, 2009.

[22] Patrick Le Callet and Florent Autrusseau, "Subjective quality assessment-IVC database," 2005, http://www.irccyn.ec-nantes.fr/ivcdb/.

[23] Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. P. Sazzad, "MICT image quality evaluation database [Online]," Available: http://mict.eng.u-toyama.ac.jp/ mict/index2.html.

[24] H.R. Sheikh, M.F. Sabir, and A.C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. on Image Processing*, vol. 15, no. 11, pp. 3440–3451, 2006.