

A LOG-GABOR FEATURE-BASED QUALITY ASSESSMENT MODEL FOR SCREEN CONTENT IMAGES

Heng Guo¹, Kai-Kuang Ma¹, and Huanqiang Zeng²

¹School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

²College of Information Science and Engineering, Huaqiao University, Xiamen, China 361021

ABSTRACT

In this paper, an *image quality assessment* (IQA) model for conducting objective evaluations of *screen content images* (SCIs) is proposed, called the *log-Gabor feature-based model* (LGFM). From the standpoint of signal representation, the log-Gabor filters outperform the classical Gabor filters since the outputs of the log-Gabor filters are more consistent with the perception of visual cortex in human visual system (HVS). Furthermore, the following two remarkable characteristics of the log-Gabor filters are highly beneficial to develop a more accurate IQA model; i.e., (i) zero response at the DC, and (ii) stronger response at high frequencies. In our proposed LGFM, the log-Gabor filters are used to extract features from the *luminance* of the reference SCIs and that of the distorted SCIs for measuring their degree of similarity. Together with the measurements from the other two chrominance components, the final LGFM score will be arrived at the output of the pooling stage. Extensive simulation results have shown that our proposed LGFM is highly consistent with the human perception, compared to other state-of-the-art IQA models.

Index Terms— Image Quality Assessment, Screen Content Images, Gabor features, log-Gabor.

1. INTRODUCTION

With rapid development of information technology, *screen content images* (SCIs), which contain a mixture of natural images together with texts, charts, symbols, and/or computer-generated images, are often encountered in various multimedia applications. SCIs usually contain a large amount of sharp edges, hence the perception of SCIs are quite different from that of *natural* images. Conventionally, the *mean square error* (MSE) and *peak signal-to-noise ratio* (PSNR) have been exploited to measure the visual quality of natural images. However, they are often inconsistent with the judgments made by the human visual system (HVS) [1]- [2]. Thus, an accurate perceptual-based IQA model is needed.

This work was supported in part by Singapore Ministry of Education (MOE) Academic Research Fund Tier-1 under the Grant 2017-T1-002-110 and Tier-2 under the Grant 2015-T2-2-114.

Among the existing image quality assessment (IQA) algorithms for the SCIs, Ni *et al.* [3] developed the *Gabor feature-based model* (GFM) that is able to deliver state-of-the-art performance. This was motivated by the fact that the Gabor filter's output values are highly consistent with the response of HVS when the image is viewed by our eyes [4]. In [5], Field further suggested that the *log-Gabor* is a better signal representation scheme than Gabor, since the images under the log-Gabor representation yield a Gaussian distribution on the logarithmic frequency scale. Another support on the use of the log-Gabor function for shape recognition application can be found in Massot *et al.* [6]. Inspired by these, a new log-Gabor-based IQA model is proposed in this paper, called the *log-Gabor feature-based model* (LGFM), which exploits two remarkable characteristics of the log-Gabor filter as follows.

First, the log-Gabor function has no DC component. Note that the DC essentially captures the global information, and it does not convey the details of image contents. For example, the shape of a bridge should be perceived the same regardless it is viewed at night or in the daytime. That is, it might be more desirable to carry out perceptual-related works in the absent of the DC response. Second, the log-Gabor function provides a tunable bandwidth. Compared to the Gabor function whose bandwidth is limited to one octave, the log-Gabor can extend its bandwidth to higher frequency range. This is advantageous from the standpoint of signal representation, since the image content of distorted SCIs and their corresponding reference SCIs are quite similar, the subtle difference can only be reflected in the higher frequency range.

In Section II, the proposed LGFM IQA model for evaluating SCIs is presented. In Section III, extensive performance evaluations are conducted and compared with that of state-of-the-art IQA models. Section IV concludes this paper.

2. PROPOSED IQA MODEL BASED ON THE LOG-GABOR FEATURES

2.1. Overview

The framework of the proposed log-Gabor feature-based IQA model for evaluating the quality of a distorted SCI with re-

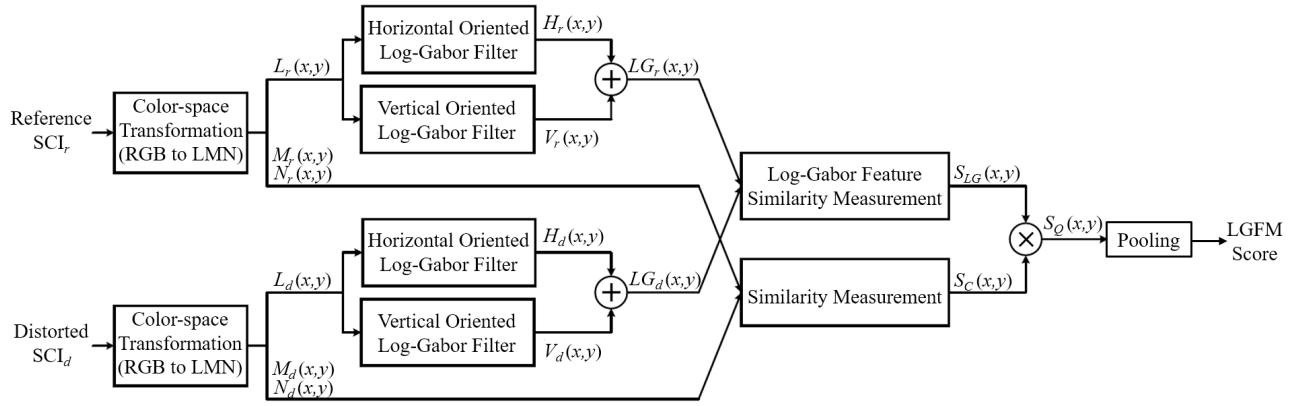


Fig. 1: Framework of the proposed log-Gabor feature-based image quality assessment (IQA) model for the screen content images (SCIs).

spective to its reference SCI is illustrated in Fig. 1. In the first stage, both the reference SCI (denoted by r in the subscript) and the distorted SCI (denoted by d in the subscript) are converted from the RGB color space to the LMN color space based on the following formula:

$$\begin{bmatrix} L_\zeta(x,y) \\ M_\zeta(x,y) \\ N_\zeta(x,y) \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{bmatrix} \begin{bmatrix} R_\zeta(x,y) \\ G_\zeta(x,y) \\ B_\zeta(x,y) \end{bmatrix} \quad (1)$$

where ζ denotes either the reference SCI r or the distorted SCI d . Components $L_\zeta(x,y)$, $M_\zeta(x,y)$ and $N_\zeta(x,y)$ denote the luminance and two chrominance at the pixel location (x,y) , respectively.

In the second stage, the log-Gabor filters are used to convolve with the luminance component of the reference SCI and that of the distorted SCI to generate their log-Gabor feature maps. In the third stage, the degree of similarity between their log-Gabor feature maps (for luminance) will be measured. The same measurement practice will be applied to chrominance components separately. In the last stage, these measurements will be fused to arrive at the final IQA evaluation score. The details of each stage will be further described in the following sub-sections.

2.2. Why the Log-Gabor Filter?

The Gabor filter is one of the most popular choices in many image processing algorithms, since its filtered result is quite consistent with the HVS perception. The Gabor filter $G(\omega)$ is defined as [7]

$$G(\omega) = n_c \cdot \exp \left[-\frac{\sigma^2}{2} (\omega - \omega_0)^2 \right] \quad (2)$$

where ω_0 is the filter's center frequency and n_c is a normalization constant. However, the Gabor filter has a limitation on the selection of its bandwidth: when the bandwidth is set

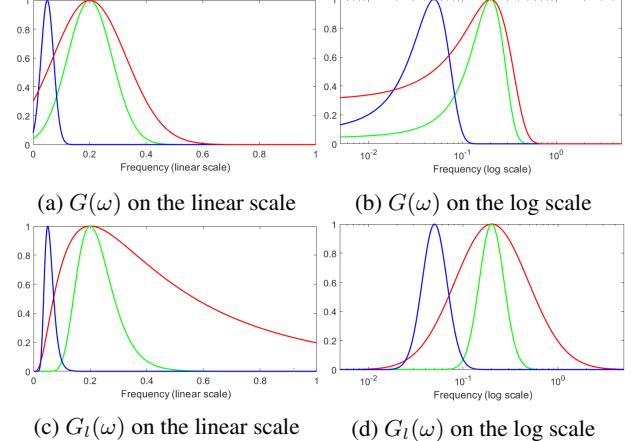


Fig. 2: Illustrations of the Gabor and log-Gabor filters' frequency responses. At the DC, the response of the Gabor filter $G(\omega)$ is non-zero, while that of the log-Gabor $G_l(\omega)$ is always zero, besides its stronger response at high frequencies.

to a value that is larger than 1 octave, its DC response becomes non-negligible [8] as illustrated in Fig. 2 (a) and (b), from which one can see more details in the log scale in (b) than that of the linear scale in (a).

To overcome the bandwidth limitation of the Gabor filter as above-mentioned, Field [5] proposed the log-Gabor filter $G_l(\omega)$, which is defined as

$$G_l(\omega) = n_c \cdot \exp \left[-\frac{\ln^2(\omega/\omega_0)}{2 \ln^2(\kappa_\beta)} \right], \quad (3)$$

where ω_0 is the filter's center frequency, and parameter κ_β (where $0 < \kappa_\beta < 1$) controls the shape of the filter, which is

related to the bandwidth β of the filter via the following:

$$\beta = -\frac{2\sqrt{2}}{\sqrt{\ln 2}} \ln(\kappa_\beta) \quad \text{or} \quad \kappa_\beta = \exp\left(-\frac{\beta}{4}\sqrt{2 \ln 2}\right). \quad (4)$$

Refer to Fig. 2 for some comparisons of the Gabor and the log-Gabor filters in their frequency domains. Note that the *linear* scale is used in (a) and (c), while the *logarithmic* scale is used in (b) and (d). The color legends of these plots denote various settings: (Blue) center frequency at 0.05 and 1-octave bandwidth; (Green) center frequency at 0.2 and 1-octave bandwidth; (Red) center frequency at 0.2 and 3-octave bandwidth. Note that, in the Gabor filter, when its bandwidth is larger than 1 octave, its DC response is non-negligible. On the other hand, the log-Gabor filter has a zero DC response and a large response in higher bandwidth.

In summary, compared with the Gabor filters, the log-Gabor filters produce (i) zero response at the DC, and (ii) stronger response at high frequencies. For (i), as mentioned previously, the DC component captures the global information (e.g., the level of lighting), which is much less relevant to the evaluations of image quality. For (ii), image distortions tend to yield high frequencies. This matches the log-Gabor filters characteristics on signal representation and therefore helps to conduct image quality assessment more accurately.

According to Kovesi's numerical analysis [9], $\kappa_\beta = 0.74$ will result in a filter bandwidth of 1 octave, $\kappa_\beta = 0.55$ with a bandwidth of 2 octaves, and $\kappa_\beta = 0.41$ with a bandwidth of 3 octaves. Psychophysical research indicated that the bandwidth of the transfer functions of the cells in the visual cortex of mammals is between 1 octave and 3 octaves [10]. Consider that the information acquired from the high frequency range is more useful for the IQA application, thus κ_β is set to 0.41 with a bandwidth of 3 octaves in this paper.

The log-Gabor filter as shown in (3) is convolved with the *luminance* component of each SCI for extracting their features. Through an independent convolution along each direction, the horizontal and vertical responses are generated as

$$\begin{aligned} H_\zeta(x, y) &= G_l^h(x, y) \otimes L_\zeta(x, y); \\ V_\zeta(x, y) &= G_l^v(x, y) \otimes L_\zeta(x, y), \end{aligned} \quad (5)$$

where the symbol \otimes denotes the convolution operator. A summation of the obtained horizontal and vertical responses is defined as the *log-Gabor feature* of the luminance component in this paper; that is,

$$G_{l\zeta}(x, y) = H_\zeta(x, y) + V_\zeta(x, y) \quad (6)$$

Fig. 3 (e) and (f) demonstrate an example of the extracted log-Gabor features of a reference SCI_r and its distorted version SCI_d, respectively.

2.3. Similarity Measurements

In compliance with many existing state-of-the-art works such as [2, 3, 11], the degree of similarity of the log-Gabor fea-

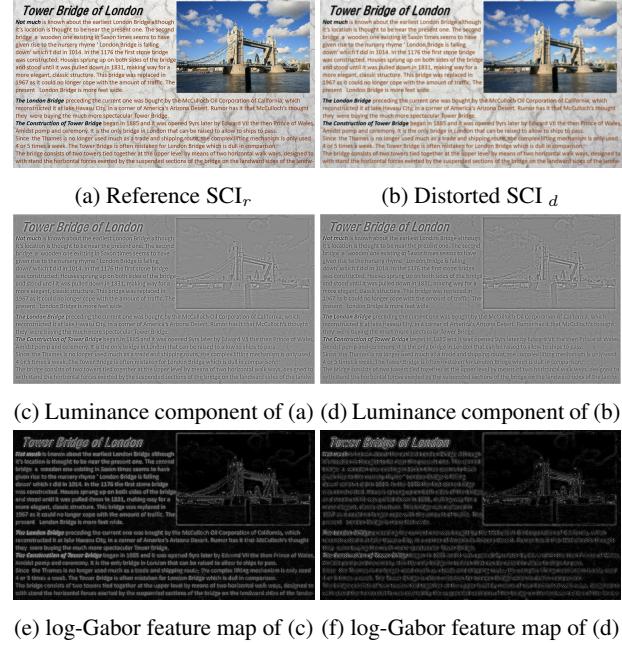


Fig. 3: An illustration of the extracted log-Gabor features of SCIs: (Left column) reference image SCI_r and (Right column) distorted image SCI_d.

tures, between the luminance components of the reference and distorted SCIs, S_{G_l} , can be calculated according to

$$S_{G_l}(x, y) = \frac{2G_{lr}(x, y) \cdot G_{ld}(x, y) + c_1}{G_{lr}^2(x, y) + G_{ld}^2(x, y) + c_1} \quad (7)$$

where $G_{lr}(x, y)$ and $G_{ld}(x, y)$ are the log-Gabor feature maps of the SCI_r and SCI_d, respectively, and c_1 is a constant inserted for avoiding numerical instability when the denominator becomes to zero.

Besides the luminance distortion, chrominance components (denoted as $M_\zeta(x, y)$ and $N_\zeta(x, y)$) also affect the perceptual assessment of the visual quality of color images to a large extent [12]. The chrominance similarity is used to reflect the quality degradation caused by color distortions, which is defined as

$$\begin{aligned} S_C(x, y) &= \frac{2M_r(x, y) \cdot M_d(x, y) + c_2}{M_r^2(x, y) + M_d^2(x, y) + c_2} \\ &\quad \cdot \frac{2N_r(x, y) \cdot N_d(x, y) + c_2}{N_r^2(x, y) + N_d^2(x, y) + c_2} \end{aligned} \quad (8)$$

where $M_r(x, y)$, $N_r(x, y)$ and $M_d(x, y)$, $N_d(x, y)$ are the chrominance components of the SCI_r and SCI_d, respectively. Likewise, constant c_2 plays the same role as c_1 .

After obtaining the similarity measurement on the luminance and two chrominance components, the local quality score $S_Q(x, y)$ is then computed as

$$S_Q(x, y) = [S_{G_l}(x, y)]^{\alpha_1} \cdot [S_C(x, y)]^{\alpha_2} \quad (9)$$

where α_1 and α_2 are two positive constants that are used to adjust the relative importance of $S_{G_l}(x, y)$ and $S_C(x, y)$. In this paper, these two parameter values are empirically determined as $\alpha_1 = 1$ and $\alpha_2 = 0.04$.

2.4. Log-Gabor Feature Pooling Strategy

Intuitively, humans will pay more attention on those pixels with stronger features and result in stronger visual perception from this location. This suggests that pixels which yielded a larger log-Gabor feature value should be given more consideration. The log-Gabor feature pooling strategy is thus designed to accommodate such intuition as follows.

At each pixel location (x, y) , the log-Gabor feature response from the reference SCI (i.e., $G_{lr}(x, y)$) and the distorted SCI (i.e., $G_{ld}(x, y)$) will be considered, and the larger value will be exploited as the weighting factor to reflect the phenomenon as above-mentioned. Thus, a *weight map* $\omega(x, y)$ can be generated according to

$$\omega(x, y) = \max \{ |G_{lr}(x, y)|, |G_{ld}(x, y)| \} \quad (10)$$

where $|\cdot|$ means the absolute value of the argument will be performed. The final image quality assessment score, *LGF Score*, can be obtained by performing the weighted average over all the pixel locations (x, y) on the local quality map $S_Q(x, y)$; that is,

$$LGF Score = \frac{\sum_{(x,y)} \omega(x, y) \cdot S_Q(x, y)}{\sum_{(x,y)} \omega(x, y)} \quad (11)$$

3. EXPERIMENTAL RESULTS

In our experiments, the most recent and largest publicly available database SCID [11] is used for conducting performance evaluation of our proposed LGFM IQA model. In this database, there are 1,840 SCIs, with nine different types of image distortions, including Gaussian noise (GN), Gaussian blur (GB), motion blur (MB), contrast change (CC), color saturation change (CSC), color quantization with dithering (CQD), JPEG compression (JPEG), JPEG2000 compression (J2K), and HEVC-SCC.

Three standard performance evaluation procedures as suggested in [16] are used, namely, the Pearson linear correlation coefficient (PLCC), the Spearman rank order correlations coefficient (SROCC), and the root mean square error (RMSE). PLCC reflects the linear relationship between our model's predicted score and the ground truth mean opinion score (MOS). SROCC measures how well our model ranks the images in the same order as MOS. Higher values of PLCC and SROCC suggest a stronger linear correlation and a better ranking consistency. RMSE quantifies the difference between the model's predicted scores and MOS; a smaller RMSE value means a better prediction.

Table 1: PLCC, SROCC AND RMSE COMPARISON OF VARIOUS IQA MODELS UNDER DIFFERENT DISTORTION TYPES ON SCID DATABASE.

	Distortions	PSNR	SSIM [2]	SCQI [13]	SIQM [14]	SVQI [15]	ESIM [11]	GFM [3]	LGF M
PLCC	GN	0.9530	0.9354	0.9319	0.9269	0.9362	0.9563	0.9497	0.9599
	GB	0.7772	0.8711	0.8244	0.9266	0.9130	0.8700	0.9156	0.9309
	MB	0.7615	0.8794	0.8147	0.9152	0.8997	0.8824	0.9023	0.9002
	CC	0.7435	0.6903	0.8353	0.7821	0.8266	0.7908	0.8787	0.8336
	JPEG	0.8393	0.8581	0.9036	0.9226	0.9356	0.9421	0.9392	0.9318
	J2K	0.9176	0.8586	0.9312	0.9076	0.9513	0.9457	0.9226	0.9453
	CSC	0.0622	0.0890	0.8393	0.0683	0.0919	0.0694	0.8728	0.8759
	HEVC-SCC	0.7991	0.7914	0.8708	0.8316	0.8496	0.9108	0.8740	0.9036
	CQD	0.9210	0.7810	0.8823	0.8385	0.9047	0.9005	0.8928	0.9203
SROCC	Overall	0.7622	0.7343	0.7489	0.8303	0.8604	0.8630	0.8760	0.9023
	GN	0.9424	0.9171	0.9556	0.9133	0.9191	0.9460	0.9370	0.9511
	GB	0.7702	0.8698	0.8638	0.9232	0.9079	0.8699	0.9081	0.9287
	MB	0.7375	0.8588	0.8587	0.9006	0.8842	0.8608	0.8892	0.8933
	CC	0.7265	0.6564	0.7465	0.7435	0.7705	0.6182	0.8225	0.8575
	JPEG	0.8321	0.8490	0.9171	0.9158	0.9287	0.9455	0.9281	0.9324
	J2K	0.9074	0.8439	0.9270	0.8935	0.9367	0.9359	0.9085	0.9342
	CSC	0.0908	0.0963	0.8970	0.0617	0.0790	0.1037	0.8736	0.8901
	HEVC-SCC	0.8074	0.8263	0.8721	0.8517	0.8665	0.9036	0.8712	0.8914
RMSE	COD	0.9080	0.7766	0.9099	0.8301	0.8957	0.8868	0.8907	0.9152
	Overall	0.7512	0.7146	0.7814	0.8086	0.8386	0.8478	0.8759	0.9046
	GN	3.8093	4.4458	4.5600	4.8222	4.4179	3.6760	3.9378	3.5248
	GB	6.6633	5.1998	5.9943	4.0989	4.3194	5.2213	4.2566	3.8679
	MB	7.0843	5.2044	6.3394	4.7388	4.7709	5.1431	4.6121	4.7625
	CC	5.9867	6.4767	4.9217	6.1281	5.0374	5.4790	4.2732	4.9446
	JPEG	8.1718	7.7179	6.4390	6.7341	5.3053	5.0373	5.2011	5.4587
	J2K	6.3222	8.1562	5.8002	7.2951	4.9058	5.1695	6.1385	5.1898
	CSC	9.8203	9.8003	5.3503	9.8394	9.7977	9.8156	4.8031	4.7744
HEVC-SCC	8.4009	8.5037	6.8407	8.1970	7.3381	5.7446	6.7590	5.9596	
	CQD	4.9814	7.9855	6.0188	7.1976	5.4481	5.5607	5.7592	5.0151
	Overall	9.1682	9.6133	9.3846	7.8920	7.2178	7.1552	6.8310	6.1052

We compared the proposed LGFM model with the classical and state-of-the-art IQA models as documented in Table 1. Note that all the methods (except PSNR and SSIM [2]) are designed for the evaluation of the SCIs. In this table, the first-ranked, the second-ranked, and the third-ranked performance evaluation result of each measurement criterion are boldfaced in red, blue, and black, respectively.

The proposed LGFM model achieved the highest overall PLCC and SROCC values and the lowest overall RMSE value. This indicates that the proposed LGFM model is most consistent with subjective judgments made by the HVS, compared with other existing state-of-the-art IQA models.

4. CONCLUSION

In this paper, a novel image quality assessment model for the evaluation of screen content images (SCIs), called the *log-Gabor feature-based model* (LGFM), is proposed. The success of our approach is due to the fact that the log-Gabor filters are exploited to extract salient features from the SCIs under evaluation. As a result, the extracted *log-Gabor features* are highly consistent with the perception of HVS, and therefore better reflect the distortions for quality assessment. Extensive experimental results have demonstrated that the proposed LGFM model outperforms other state-of-the-art IQA models on objectively evaluating the image quality of the SCIs under various types of distortions.

5. REFERENCES

- [1] Z. Wang and A. C. Bovik, “Mean squared error: Love it or leave it? a new look at signal fidelity measures,” *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009.
- [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] Z. Ni, H. Zeng, L. Ma, J. Hou, J. Chen, and K.-K. Ma, “A Gabor feature-based quality assessment model for the screen content images,” *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4516–4528, 2018.
- [4] J. G. Daugman, “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters,” *Journal of the Optical Society of America A, Optics and image science*, vol. 2, pp. 1160–9, 1985.
- [5] D. J. Field, “Relations between the statistics of natural images and the response properties of cortical cells,” *Journal of the Optical Society of America A*, vol. 4, no. 12, pp. 2379–2394, 1987.
- [6] C. Massot and J. Hérault, “Model of frequency analysis in the visual cortex and the shape from texture problem,” *International Journal of Computer Vision*, vol. 76, no. 2, pp. 165–182, 2008.
- [7] D. Boukerroui, J. Noble, and M. Brady, “On the choice of bandpass quadrature filters,” *Journal of Mathematical Imaging and Vision*, vol. 21, pp. 53–80, 2004.
- [8] C. Ronse, “On idempotence and related requirements in edge detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 5, pp. 484–491, 1993.
- [9] P. Kovesi, *Invariant measures of image features from phase information*, Ph.D. thesis, University of Western Australia, 1996.
- [10] S. J. Anderson and D. C. Burr, “Spatial and temporal selectivity of the human motion detection system,” *Vision Research*, vol. 25, no. 8, pp. 1147 – 1154, 1985.
- [11] Z. Ni, L. Ma, H. Zeng, J. Chen, C. Cai, and K.-K. Ma, “ESIM: Edge similarity for screen content image quality assessment,” *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4818–4831, 2017.
- [12] L. Zhang, Y. Shen, and H. Li, “VSI: A visual saliency-induced index for perceptual image quality assessment,” *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4270–4281, 2014.
- [13] S.-H. Bae and M. Kim, “A novel image quality assessment with globally and locally consilient visual quality perception,” *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2392–2406, 2016.
- [14] K. Gu, S. Wang, G. Zhai, S. Ma, and W. Lin, “Screen image quality assessment incorporating structural degradation measurement,” in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2015, pp. 125–128.
- [15] K. Gu, J. Qiao, X. Min, G. Yue, W. Lin, and D. Thalmann, “Evaluating quality of screen content images via structural variation analysis,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 10, pp. 2689–2701, 2018.
- [16] VQEG, “Final report from the video quality experts group on the validation of objective models of video quality assessment,” 2015.