No-Reference Image Quality Assessment with Orientation Selectivity Mechanism

Jinjian Wu[1], Man Zhang[1], Guangming Shi[1], Xuemei Xie[1] and Weisi Lin[2], Fellow, IEEE

***Abstract*—No-reference (NR) image quality assessment (IQA) technology is greatly required in quality-orientated visual signal processing systems. However, without the guidance of the reference information, it is still a great challenge for NR IQA to perform consistent with the subjective perception. Researches on cognitive neuroscience state that the human visual system (HVS) presents substantially orientation selectivity mechanism, within which the visual structures are extracted in the local receptive fields for scene understanding. Inspired by this mechanism, a set of orientation selectivity based visual patterns are designed. By analyzing the quality degradation on those patterns, a novel visual pattern degradation based NR IQA method is proposed. Experimental results on large databases demonstrate that the proposed method outperforms the existing NR IQA methods.**

***Index Terms*—No-Reference (NR), Image Quality Assessment (IQA), Orientation Selectivity Mechanism, Visual Pattern Degradation**

I. Introduction

Image quality assessment (IQA) technique, which aims to automatically measure the perceptual quality of images, is widely used for optimizing the quality-oriented visual signal processing system [1, 2], e.g., image/video acquisition, transmission, and restoration systems.

During the past decades, a large amount of IQA methods have been proposed [3–5], which can be classified into three categories according to the participant of the reference information: 1) Full reference (FR), which needs the whole reference information; (2) Reduced reference (RR), which requires part of the reference information; (3) No reference (NR), which directly predicts the quality with the distorted image only. With the guidance of the reference information, the FR/RR IQA methods performs highly consistently with the subjective perception. However, the reference image is always not available for most realistic case. In such condition, the NR IQA method is greatly required. In this paper, we focus on the NR IQA modeling.

It is still a large challenge for NR IQA to perform consistently with the subjective perception. At the early stage, the prior knowledge of the distortion type is used to design distortion specific NR IQA models, e.g., blockiness measurement for JPEG compression [6], sharpness calculation for blur distortion [7], and so on. These models can only work for a certainty type of distortion. Recently, many no distortion specific NR IQA methods have been proposed. Following the characteristic of the natural scene statistic (NSS), Moorthy et al. analyzed the distribution of the wavelet coefficients with Generalized Gaussian Gistribution (GGD) and proposed a distortion identification-based image verity and integrity evaluation (DIIVINE) method [8]. Moreover, the NSS is analyzed in other domains, such as DCT domain [9] and spatial domain [10], for NR IQA [11]. Besides, some learning based NR IQA methods, e.g., visual codebook [12] and gradient magnitude [13], are proposed for NR IQA. Though these methods have significantly promoted the performance of the NR IQA, there still exists a large gap between them and the subjective perception.

The human visual system (HVS) can efficiently evaluate the perceptual quality. Thus, we try to explore the inner mechanism of the HVS. Researches on cognitive science indicate that HVS is adapted to extract these repeated visual patterns for visual content perception [14]. Moreover, the local receptive field in the primary visual cortex presents orientation selectivity mechanism for visual pattern extraction during scene perception [15]. Inspired by this, we aim to design a set of novel visual patterns to extract the image content for quality assessment.

By mimicking the interactions among neurons in a local receptive field (i.e., excitatory and inhibitory interactions [16]), the correlations between a pixel and its neighbors are analyzed to design novel visual patterns. Next, the visual content of an image is extracted with these visual patterns. Finally, the content degradation is calculated on these visual patterns for quality prediction, and the novel visual pattern degradation based NR IQA model (NRVPD) is proposed.

The rest of this paper is organized as follows. The visual pattern designing for quality assessment is introduced in Section II. In Section III, the degradation on the visual patterns is analyzed and the novel NRVPD method is proposed. Finally, conclusions are drawn in Section IV.

II. Visual Pattern Degradation Based IQA

In this section, the orientation selectivity based visual pattern is firstly designed. Next, the degradation on these visual patterns are analyzed and a novel NRVPD model is introduced.

A. Visual Pattern Designing

Visual patterns represent the repeated information and play an extremely important role for visual content understanding.

[1]Jinjian Wu (jinjian.wu@mail.xidian.edu.cn), Man Zhang, Guangming Shi, and Xuemei Xie are with School of Electronic Engineering, Xidian University, Xi'an, Shannxi, China

[2]Weisi Lin is with School of Computer Engineering, Nanyang Technological University, Singapore

*This work was supported by the National Natural Science Foundation of China (Nos. 61401325, 61632019).

*This work was supported by the National Natural Science Foundation of China (Nos. 61401325, 61472301, 61632019, 61621005).

Different from other visual processing tasks, IQA demands a type of visual patterns which can not only extract the visual content, but can also characterize the quality degradations from distortion. Hence, we aim to design a novel type of visual patterns for IQA.

Researches on cognitive neuroscience state that the HVS presents substantial orientation selectivity mechanism (OSM) for visual pattern extraction [15]. The OSM arises from the arrangement of interactions among cortical cells in a local receptive field [16]. Moreover, there are two opponent types of interactions, i.e., excitatory and inhibitory. The two types of interactions directly relates to the preferred orientations that the cells presented: 1) excitatory interaction usually appears at neurons with similar preferred orientations, and inhibitory interaction appears for dissimilar preferred orientations [17]. Inspired by the OSM, we calculate the correlations among pixels from a local neighborhood ($\mathcal{R}$), and the preferred orientation is calculated as the gradient direction (θ),

$$\theta(i) = \arctan \frac{G_y(i)}{G_x(i)}, \tag{1}$$

where G_y (G_x) represents the luminance contrast alone the vertical (horizontal) direction.

Next, the correlation between the central pixel (i) and the other pixel from its neighborhood ($\Delta \in \mathcal{R}$) is judged as their orientation similarity,

$$\mathcal{S}(i, i{+}\Delta) = \begin{cases} 1 & \text{if } |\theta(i) - \theta(i{+}\Delta)| < \mathcal{T} \\ 0 & \text{else} \end{cases}, \tag{2}$$

where '1' ('0') means that the two pixels i and $i{+}\Delta$ present similar (dissimilar) orientations, and $\mathcal{T}$ is the judging threshold. With the subjective test about the masking effect, $\mathcal{T}$ is set as 6° [18].

By considering the correlation $\mathcal{S}(i, i{+}\Delta)$ between a central pixel i and pixels in its neighborhood $\mathcal{R}$, an orientation selectivity based visual pattern is designed

$$\mathcal{P}(i) = \{\mathcal{S}(i, j|j{=}i{+}\Delta, \Delta{\in}\mathcal{R})\}. \tag{3}$$

Since the pattern number increases exponentially with the size of the neighborhood (e.g., for $5{\times}5$ neighborhood size, there are 2^{24} different types of patterns), we need to find out there fundamental patterns for visual content representation.

Generally, local patches with a same content but different main orientations should be represented by a same visual pattern. Thus, we realigned the arrangement of the correlations according to the main orientation (Fig. 1) to combine such types of patterns, and the number of patterns can be reduced in to 1/10 of the original ones.

Moreover, these patterns with similar format represent similar contents, which should be combined for pattern reduction. Here, the K-Means clustering procedure is employed for similar pattern combination,

$$\{\hat{\mathcal{P}}_n, n = 1, 2, \cdots, K\} = \arg\min \sum_{n=1}^{K} \sum_{m=1}^{M} w_m \cdot ||(\mathcal{P}_m - \hat{\mathcal{P}}_n)||^2, \tag{4}$$

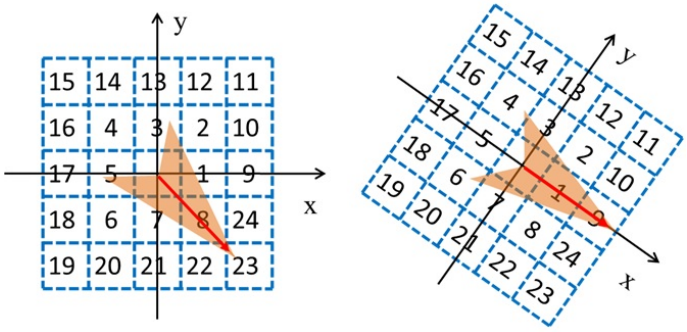


Fig. 1: An example of realigning the arrangement of correlations in a local receptive field with the main orientation, where (a) is the original pattern and (b) is the realigned one.

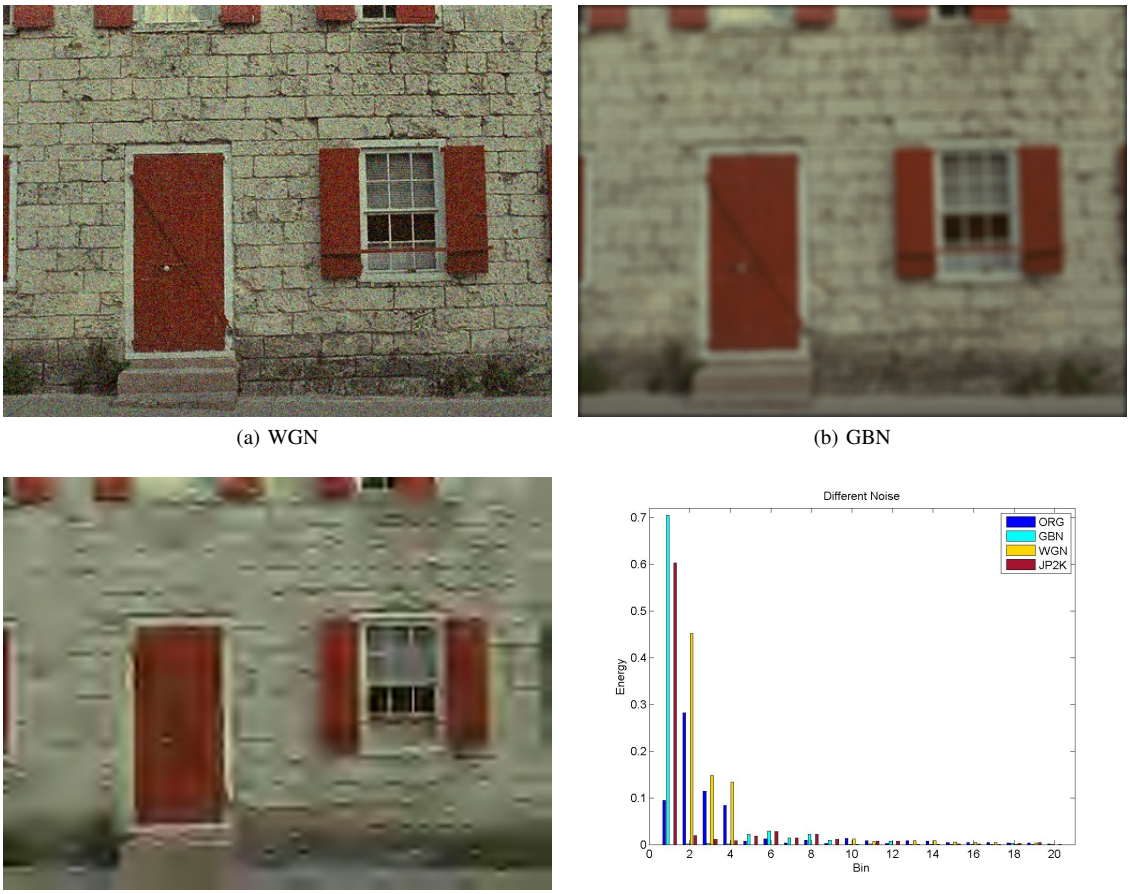
TABLE I: Comparison among these basical visual patterns clustered from different image sets

Databases	Overlap No.	Overlap Proportion
DB-C1 VS. DB-C2	94	98.83%
DB-C1 VS. LIVE	86	96.51%
DB-C1 VS. TID	90	96.83%

where $\hat{\mathcal{P}}_n$ is the nth clustering center, K is the number of centers, and w_m is the weight of the mth pattern which is determined by its probability of occurrence.

To achieve a set of fundamental patterns ($\{\hat{\mathcal{P}}_n, n = 1, 2, \cdots, K\}$) from the K-Means clustering, a large number of natural images are required. Here, we choose a publicly available image database provided by Achanta et al. [19], which contains 1000 different images and is popular in visual processing tasks. Meanwhile, those images are randomly classified into 2 sets (which are called as DB-C1 and DB-C2 here, and each with 500 images) to verified the stability of the clustering result. For DB-C1, the K-Means algorithm is implemented with the following steps: 1) The rotate invariant pattern of each pixel is calculated. 2) All of the pixels from DB-C1 are mapped into a histogram according to their pattern formats, and the weight w_m for each bin is set as the percentage of the current bin. 3) These patterns are ranked according to their weights (with descending order). 4) We set the first K patterns as the initial centers for clustering. Finally, the fundamental patterns ($\{\hat{\mathcal{P}}_n, n = 1, 2, \cdots, K\}$) are acquired.

In order to verify the stability of these fundamental patterns, we do the same process on DB-C2, LIVE (which contains 29 reference images) [20], and TID2013 (which contains 25 reference images) [21]. The comparison between the clustering results are listed in Table I (in which we set $K{=}100$). As can be seen, the fundamental patterns clustered from different datasets are quite similar. Moreover, pixels which belong to these same centroids take up the majority (more than 95%). Therefore, the proposed fundamental patterns can efficiently extract these common features for visual content representation.



(a) WGN (b) GBN

(c) J2K (d) Histogram

Fig. 2: Visual pattern degradation under different distortion types, where (a)-(c) are distorted images with white Gaussian noise (WGN), Gaussian blur noise (GBN), and JPEG2000 (J2K) compression noise, respectively, and (d) is the visual pattern based histograms. For a clear view, we only select the first 20 bins

.

B. Image Quality Assessment

The quality degradation is measured as the structural degradation on these patterns. For an input image, the visual pattern of each pixel is calculated with the procedure introduced in the former subsection. Next, the visual content of the image is mapped into a visual pattern based histogram,

$$\mathcal{H}(k) = \sum_{i=1}^{N} \mathcal{C}_l(i) \cdot \delta(\hat{\mathcal{P}}(i), \hat{\mathcal{P}}_k) \tag{5}$$

$$\delta(\hat{\mathcal{P}}(i), \hat{\mathcal{P}}_k)) = \begin{cases} 1 & \text{if } \hat{\mathcal{P}}(i) = \hat{\mathcal{P}}_k \\ 0 & \text{else} \end{cases}, \tag{6}$$

where $\mathcal{C}_l$ is the contrast value, which is calculated as the gradient magnitude,

$$\mathcal{C}_l(i) = \sqrt{G_y(i)^2 + G_x(i)^2}. \tag{7}$$

Since different types of distortions cause individual degradation on those visual patterns, we measure the quality with the changes on the visual pattern based histogram. Thus, a regression module from feature space to quality score is demanded for quality prediction. As an efficient procedure for high dimensional data processing, the support vector regression (SVR) is popular for feature regression. Here, we employ the LIBSVM package to implement the SVR for quality assessment with a radial basis function kernel,

$$\mathcal{Q}(I) = \text{SVR}\left(\mathcal{H}, \mathcal{MOD}\right), \tag{8}$$

where $\mathcal{Q}$ is the predicted quality, and $\mathcal{MOD}$ is a regression model.

III. Experimental Result Analysis

The proposed visual patterns can efficiently capture the content degradation from different types of distortion. An example is shown in Fig. 2. As can be seen, different types of distortion result in different changes on these visual patterns. As the WGN mainly increases the disturbance, the energies for most image regions are increased, as the yellow bars shown in Fig. 2 (d). On the contrary, the GBN smoothen all of the image regions, the energies for most regions are

TABLE II: IQA Performance Comparison on *TID2013* Database

Dist.	Algo. / Crit.	NR						FR	
		VPD	IL-NIQE	NIQE	BRISQUE	CBIQ	DIIVINE	PSNR	MS-SSIM
J2K	PLCC	**0.9475**	0.9152	0.9250	0.9156	0.9003	0.9136	0.9305	0.8783
	SRCC	**0.9392**	0.9106	0.9054	0.9015	0.8942	0.9073	0.8907	0.8780
	RMSE	**0.4870**	0.6803	0.6393	0.6838	0.7399	0.6855	0.6227	0.3383
JPG	PLCC	**0.9663**	0.9105	0.9389	0.9313	0.8637	0.9438	0.9177	0.9678
	SRCC	**0.9171**	0.8835	0.8826	0.8757	0.8477	0.9034	0.9188	0.9315
	RMSE	**0.3700**	0.6282	0.5173	0.5429	0.7445	0.4954	0.5919	0.3752
WGN	PLCC	0.8450	0.9008	0.8619	**0.9359**	0.7475	0.9241	0.9625	0.9678
	SRCC	0.8502	0.8923	0.8493	**0.9323**	0.8037	0.9208	0.9436	0.9532
	RMSE	0.3675	0.3150	0.3567	**0.2529**	0.4696	0.2705	0.1920	0.4283
GBN	PLCC	0.9022	0.8611	0.8682	0.8893	0.9015	**0.9253**	0.9664	0.9681
	SRCC	0.9011	0.8616	0.8450	0.8939	0.9023	**0.9292**	0.9666	0.9714
	RMSE	0.5207	0.6309	0.6112	0.5648	0.5293	**0.4703**	0.3201	0.3122
Overall	PLCC	**0.9361**	0.8749	0.8346	0.9089	0.8209	0.9225	0.9140	0.9346
	SRCC	**0.9211**	0.8830	0.8450	0.8930	0.7938	0.9079	0.9244	0.9188
	RMSE	**0.4870**	0.6731	0.7687	0.5812	0.8014	0.5376	0.5671	0.4973

TABLE III: IQA Performance Comparison on *LIVE* and *CSIQ* Databases

Dist.	Algo. / Crit.	NR						FR	
		VPD	IL-NIQE	NIQE	BRISQUE	CBIQ	DIIVINE	PSNR	MS-SSIM
LIVE	PLCC	**0.9540**	0.9200	0.9254	0.9486	0.9200	0.9259	0.8813	0.9527
	SRCC	0.9520	0.9187	0.9261	**0.9529**	0.9190	0.9279	0.8841	0.9563
	RMSE	**8.023**	10.616	10.209	7.363	12.361	8.886	12.738	8.205
CSIQ	PLCC	0.9030	0.8837	0.9097	**0.9266**	0.8300	0.8958	0.9073	0.9581
	SRCC	0.8520	0.8867	0.9012	**0.9018**	0.8173	0.8757	0.9219	0.9542
	RMSE	0.1180	0.1305	0.1152	**0.1036**	0.1582	0.1237	0.1188	0.0809

greatly decreased (the light blue bars shown in Fig. 2 (d)). The J2K removes many structure during the compression, thus the energy for most bins are decreased except for the first bin (as the red bars shown in 2 (d)). Meanwhile, the J2K creates some new ringing edges, which increase the energies of several bins (e.g., for the 5-9 bins). Therefore, the proposed visual patterns can efficiently represent the degradation from different distortion types.

Besides, the performance of the proposed NRVPD method is verified on three large-scale benchmark databases, i.e., TID2013 [21], LIVE [20], and CSIQ [22]. Since there are 4 common types of distortions for the three different databases (i.e., J2K, JPG, WGN, and GBN), we mainly consider the four types of distortions in this work. Five latest and classical NR IQA methods (i.e., IL-NIQE [11], NIQE [23], BRISQUE [10], CBIQ [12], and DIIVINE [8] and two classical FR IQA methods, i.e., PSNR and MS-SSIM [24], are also chosen for comparison. Moreover, three metrics are employed for performance comparison. i.e., the Spearman rank order correlation coefficient (SRCC), the Pearson linear correlation coefficient (PLCC) and the root mean squared error (RMSE).

We firstly verify these IQA methods on the TID2013 database. Since we use the SVR for quality prediction in the proposed model, a train-test procedure is demanded. To ensures that no overlap occurs between the training set and the test set, we randomly select 80% reference images and their corresponding distorted images for training, and the rest distorted images for testing. To eliminate performance bias, we repeat the overall train-test procedure for 1000 times, and use the median performance as the final result. The performances of these IQA methods are listed in Tab II (the best results (among these NR IQA methods) are highlighted in bold font). As can be seen, the proposed NRVPD performs the best on J2K and JPG, and performs a slightly worse than the best one on the other two distortion types. Meanwhile, the overall performance shows that the proposed NRVPD performs the best among these NR IQA methods (has a significant improvement against the existing NR IQA methods). Moreover, by comparing with the two classical FR IQA metrics, the proposed NRVPD outperforms the PSNR and MS-SSIM.

In addition, the overall performances on the LIVE and CSIQ databases are listed in Tab. III. As can be seen, the proposed NRVPD outperforms the other NR IQA methods on the LIVE database, and even better than the two FR IQA methods on this database. For the CSIQ database, the proposed VPD method performs worse than the best two (i.e., BRISQUE and NIQE), while performs better than other NR IQA methods. In summary, the proposed NRVPD performs more consistent with the subjective perception than the existing NR IQA methods.

IV. Conclusion

In this paper, we have introduced a novel visual pattern degradation based NR IQA method. Inspired by the orientation selectivity mechanism in the primary visual cortex for visual structural extraction, we have designed a new set of visual patterns for visual content representation. Next, the structural degradation caused by different types of distortion was analyzed within those visual patterns. Finally, the quality was predicted with a regression procedure from these visual patterns with SVR. Experimental results demonstrated that the proposed NRVPD outperforms the state-of-the-art NR IQA methods.

REFERENCES

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

[2] R. A. Manap and L. Shao, "Non-distortion-specific no-reference image quality assessment: A survey," *Information Sciences*, vol. 301, pp. 141–160, 2015.

[3] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 43 –54, 2013.

[4] R. Soundararajan and A. Bovik, "RRED indices: Reduced reference entropic differencing for image quality assessment," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 517 –526, Feb. 2012.

[5] L. He, D. Tao, X. Li, and X. Gao, "Sparse representation for blind image quality assessment," in *IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2012, pp. 1146 –1153.

[6] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, S. Yao, Z. Lu, and X. Yang, "A locally adaptive algorithm for measuring blocking artifacts in images and videos," *Signal Processing: Image Communication*, vol. 19, no. 6, pp. 499–506, Jul. 2004.

[7] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Transactions on Image Processing*, vol. 18, no. 4, pp. 717–728, Apr. 2009.

[8] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.

[9] M. Saad, A. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339 –3352, Aug. 2012.

[10] A. Mittal, A. Moorthy, and A. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.

[11] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.

[12] P. Ye and D. Doermann, "No-reference image quality assessment using visual codebooks," *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3129 –3138, Jul. 2012.

[13] W. Xue, X. Mou, L. Zhang, A. C. Bovik, and X. Feng, "Blind image quality assessment using joint statistics of gradient magnitude and laplacian features," *IEEE Transactions on Image Processing*, vol. 23, no. 11, pp. 4850–4862, Nov. 2014.

[14] M. Bar, "The proactive brain: memory for predictions," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 364, no. 1521, pp. 1235–1243, May 2009.

[15] D. Hansel and C. Vreeswijk, "The mechanism of orientation selectivity in primary visual cortex without a functional map," *The Journal of Neuroscience*, vol. 32, no. 12, pp. 4049–4064, 2012.

[16] J. A. Cardin, L. A. Palmer, and D. Contreras, "Stimulus feature selectivity in excitatory and inhibitory neurons in primary visual cortex," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 27, no. 39, pp. 333–344, 2007.

[17] D. Ferster and K. D. Miller, "Neural mechanisms of orientation selectivity in the visual cortex," *Annual review of neuroscience*, vol. 23, pp. 441–471, 2000.

[18] J. Wu, W. Lin, G. Shi, Y. Zhang, W. Dong, and Z. Chen, "Visual orientation selectivity based structure description," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4602–4613, Nov. 2015.

[19] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection." IEEE CVPR 2009, Jun. 2009, pp. 1597–1604.

[20] H. R. Sheikh, K. Seshadrinathan, A. K. Moorthy, Z. Wang, A. C. Bovik, and L. K. Cormack. (2006) Image and video quality assessment research at live.

[21] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. Kuo, "Color image database TID2013: Peculiarities and preliminary results," in *2013 4th European Workshop on Visual Information Processing (EUVIP)*, Jun. 2013, pp. 106–111.

[22] E. C. Larson and D. M. Chandler. (2004) Categorical image quality (csiq) database.

[23] A. Mittal, R. Soundararajan, and A. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, 2013.

[24] 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.