IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 13, 2018, accepted March 14, 2018, date of publication April 18, 2018, date of current version June 29, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2825368

Visual Information Evaluation With Entropy of Primitive

SONGCHAO TAN[1], SHURUN WANG[2], XIANG ZHANG[2], SHANSHE WANG[2], SHIQI WANG[3], SIWEI MA[2], AND WEN GAO[2], (Fellow, IEEE)
[1]School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China
[2]School of Electronics Engineering and Computer Science, Institute of Digital Media, Peking University, Beijing 100871, China
[3]Department of Computer Science, City University of Hong Kong, Hong Kong

Corresponding author: Songchao Tan (sctan@mail.dlut.edu.cn)

This work was supported in part by the General Financial Grant from the China Postdoctoral Science Foundation under Grant 2016M600866, in part by the National Natural Science Foundation of China under Grant 61571017, and in part by the Top-Notch Young Talents Program of China and Shenzhen Peacock Plan.

ABSTRACT In this paper, we overview the recent work on entropy of primitive (EoP), including its concept, design, extension, and mathematical analysis in evaluating the visual information of natural images. The design philosophy of EoP is establishing an entropy model that quantifies the visual information based on patch-level sparse representation, due to the close relationship between sparse representation and the hierarchical cognitive process of human perception. Furthermore, based on the concept and definition of EoP, we also demonstrate several applications, including just noticeable difference estimation and visual quality assessment. The future research directions of visual information evaluation are also envisioned, where we can perceive both promises and challenges.

INDEX TERMS Entropy of primitive, sparse representation, visual information, quality assessment, just noticeable difference.

I. INTRODUCTION

The evaluation of visual information perceived by the human visual system (HVS) is a fundamental issue that plays an important role in understanding the visual world. In the field of information theory [1], the entropy is an effective measure that quantifies the amount of information missing before reception. For natural images, a traditional way of the visual information quantification is to calculate the entropy at pixel level. For example, the histogram that characterizes the occurrence of each pixel in the image can be built for entropy calculation. However, this may not reveal the visual information perceived by HVS. It has been widely believed that natural image signals are highly structured [2]: their pixels are not independently distributed and exhibit strong dependencies that carry important information about the structure of the objects in the visual scene. As such, the hypothesis that the visual information is perceived in terms of pixels is quite questionable.

Sparse representation has been repeatedly proven to be powerful in characterizing the visual signals based on the sparsity and redundancy of their representations for many visual processing tasks [3]. In [4], the properties of spatially localized, oriented and bandpass properties of the primitives in sparse representation are exhibited to be closely relevant with the human visual system, especially the receptive fields of simple cells. As such, constantly increasing applications have been powered by sparse representation, and promising performance in image quality assessment [5]–[15], image denoising [16], image restoration [17]–[22] and image/video coding [23]–[28] has been achieved. Despite the great success of sparse representation, the visual information evaluation based on patch-level sparse representation is an emerging area.

To perform sparse representation, the typical K-SVD [29] algorithm is a popular method in obtaining the over-completed dictionary based on dictionary training. A series of matching pursuit family algorithms have also been presented to achieve sparse representation [30]. Among them, the orthogonal matching pursuit (OMP) [31] is one that works in a greedy fashion. These powerful tools enable efficient and effective visual signal representation for visual information evaluation. Benefiting from the advantages of sparse representation over the traditional pixel level representation, the Entropy of Primitive (EoP) was proposed to characterize the visual information in a more precise way. In this paper, we investigate and summarize the design and concept of EoP, explore its characteristics and possible extensions, and demonstrate its further applications. The future extensions of visual information evaluation, especially based on deep learning, which has been demonstrated to be more powerful in dealing with the rich, varied and directional information, are discussed. In summary, the paper presents the following contributions:

- We systematically analyze the concept and design philosophy of EoP in an effort to provide an accessible and intuitive overview of these approaches. In particular, the sparse representation and dictionary learning, which serve as the foundations of EoP, as well as the properties of the EoP, are reviewed. The extension of EoP by considering the coefficient energy and its convergency analysis, are also discussed.
- The applications of EoP are introduced, including the typical visual perception relevant tasks such as just noticeable difference (JND) estimation and visual quality assessment. Possible applications of EoP that worth further investigation are also discussed.
- We outline the main challenges and potential further research directions of visual information evaluation powered by the deep neural network, due to wide spectrum visual computing applications of deep learning that can automatically extract meaningful features in a data driven manner.

The rest of the paper is organized as follow. In Section 2, we provide the review of EoP as well as its properties and extensions. Section 3 introduces the applications of EoP. Current challenges and future directions of visual information evaluation are discussed in Section 4, and the paper is concluded in Section 5.

II. ENTROPY OF PRIMITIVE

A. SPARSE REPRESENTATION

The Sparseland model serves as the foundation of the sparse representation [3], and it assumes that natural visual signals $x (x \in \mathbb{R}^n)$ can be well represented by a linear combination over an over-complete dictionary, which can be written as $\forall x, x \approx \Psi\alpha$ and $\|\alpha\|_0 \ll n$. Here, $\Psi (\Psi \in \mathbb{R}^{n\times k})$ is the over-complete dictionary and the primitive is denoted as ψ_i. $\alpha (\alpha \in \mathbb{R}^k)$ is the representation vector corresponding to the coefficients in sparse representation. The notation $\|\bullet\|_0$ represents the ℓ_0 norm. Typically, we assume that $k > n$, implying the dictionary Ψ is redundant to x. To obtain the over-complete dictionary, the K-SVD algorithm [29] is typically employed. In particular, two iterative calculations are performed, including sparse coding and dictionary updating. Given the training samples which are generated by partitioning the input image into patches $x_1, x_2, \ldots, x_i, \ldots$, we can obtain the dictionary that leads to the best representation of the image under the sparsity constraint, which is formulated as follows,

$$(\Psi, \{\alpha_i\}) = \underset{\Psi,\{\alpha_i\}}{\arg\min} \sum\nolimits_i \|x_i - \Psi\alpha_i\|_2^2, \; s.t. \|\alpha_i\|_0 < \mathbb{L}. \quad (1)$$

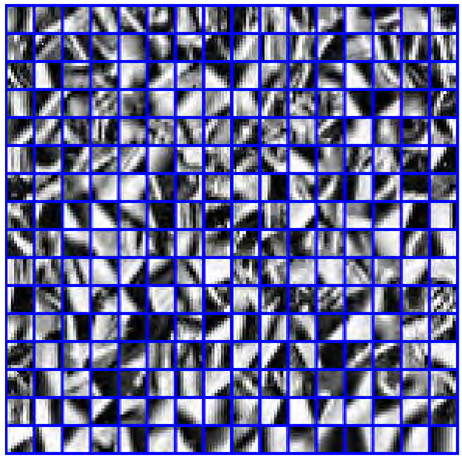


FIGURE 1. The 256 primitives learned in terms of 8 × 8 patches (*Lena*).

Here, $\mathbb{L}$ controls the sparse level. A typical dictionary learned from the *Lena* image is shown in Fig. 1.

Based on the trained dictionary Ψ and the constrains on accuracy and sparsity, sparse representation targets at obtaining the coefficients α_i that represents the visual signal, which is given by,

$$\alpha_i = \underset{\alpha_i}{\arg\min} \|x_i - \Psi\alpha_i\|_2^2, \; s.t. \|\alpha_i\|_0 < \mathbb{L}. \quad (2)$$

This can be solved by a number of approximation algorithms, among which the OMP [31] is a popular one working in a greedy way.

The sparse representation in terms of different values of L is shown in Fig. 2, where L is used to specify a specific iteration in OMP. It is clearly observed that the reconstruction quality increases with the value of L. In particular, the image reconstructed by the first layer ($L = 1$) represents the primary information. With the increasing number of the primitives used in the reconstruction, the artifacts such as blocking and blur are removed and most of the structural information can be preserved after $L = 6$. In essence, such observation is in accordance with the mechanism of HVS. In particular, based on the Internal Generative Mechanism (IGM) theory [32], [33], visual signal can be regarded by the composition of primary visual information and uncertainty. The primary visual information can be accounted by the low-level layers, and the uncertainty corresponds to the details described by the high level layer.

In [34], it is observed that the matching pursuit schemes such as OMP can decompose the image signal into multiple layers, including primary, sketch and non-structural layers. Interestingly, these layers are naturally ordered by perceptual importance, as demonstrated in Fig. 3. As such, the most significant structural information can be reconstructed by the

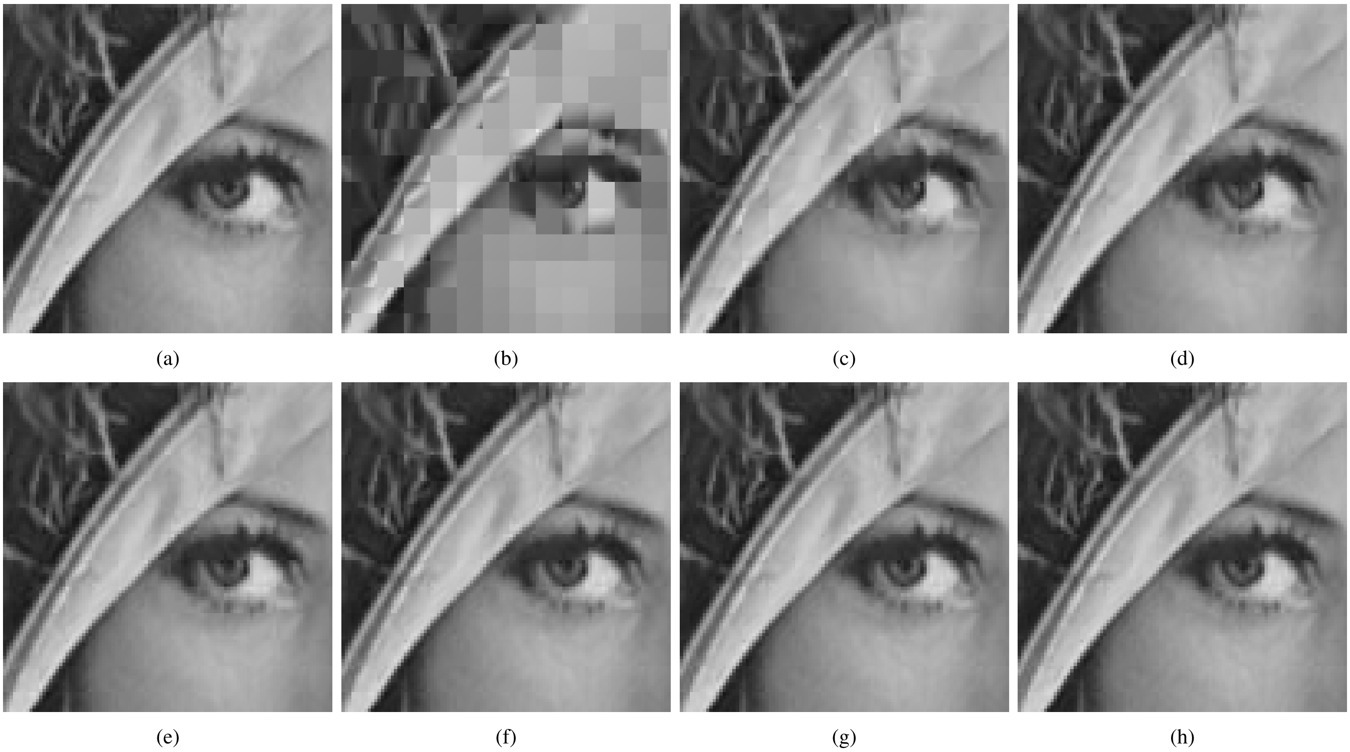


FIGURE 2. **Reconstructed *Lena* images with different number of primitives. (a) Original image. (b) $L = 1$. (c) $L = 3$. (d) $L = 5$. (e) $L = 7$. (f) $L = 9$. (g) $L = 11$. (h) $L = 13$.**

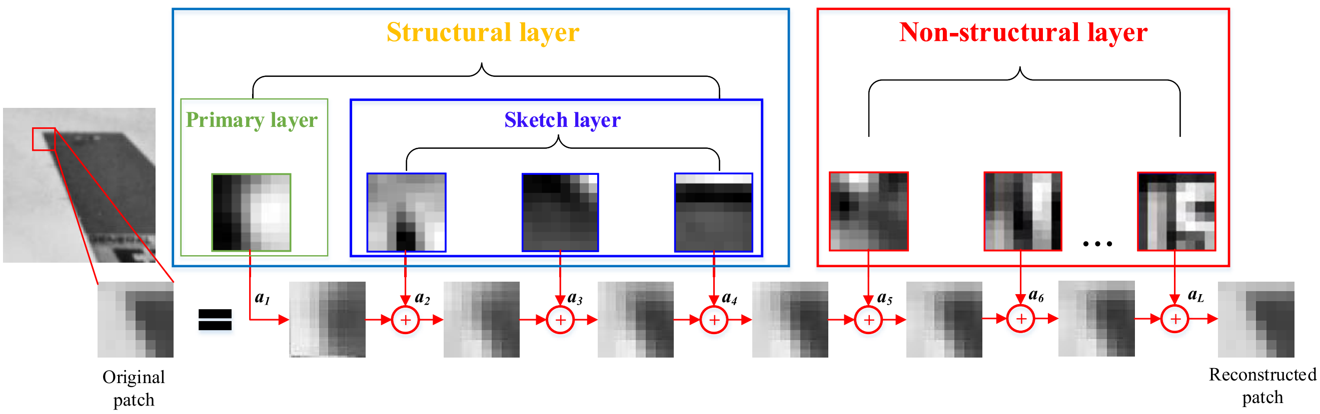


FIGURE 3. **Hierarchical image representation based on different number of primitives [34].**

first layer ($L = 1$), and the following layers reconstruct the detailed information. With the primary and sketch layers, almost all the perceptual information that is sensitive to HVS is adequately represented.

B. ENTROPY OF PRIMITIVE

The design philosophy of EoP [34] is that the visual information perceived in the natural scene shall be evaluated in terms of the patch level representation instead of the pixel level histogram. This originates from the fact that the natural images obey the natural scene statistics [35] and appear to be highly structured [2]. Moreover, due to the close relationship between the patch level sparse representation and human perception, the primitive is adopted as the basis in entropy calculation. In particular, in the sparse representation process, we assume that the total number of the i^{th} primitive used from the first iteration to the L^{th} iteration is defined as $\bar{N}_{L,i}$. The corresponding probability density functions (PDF) for the i^{th} primitive selected is given by,

$$\bar{p}_{L,i} = \frac{\bar{N}_{L,i}}{\sum_{i=1}^{k} \bar{N}_{L,i}}. \tag{3}$$

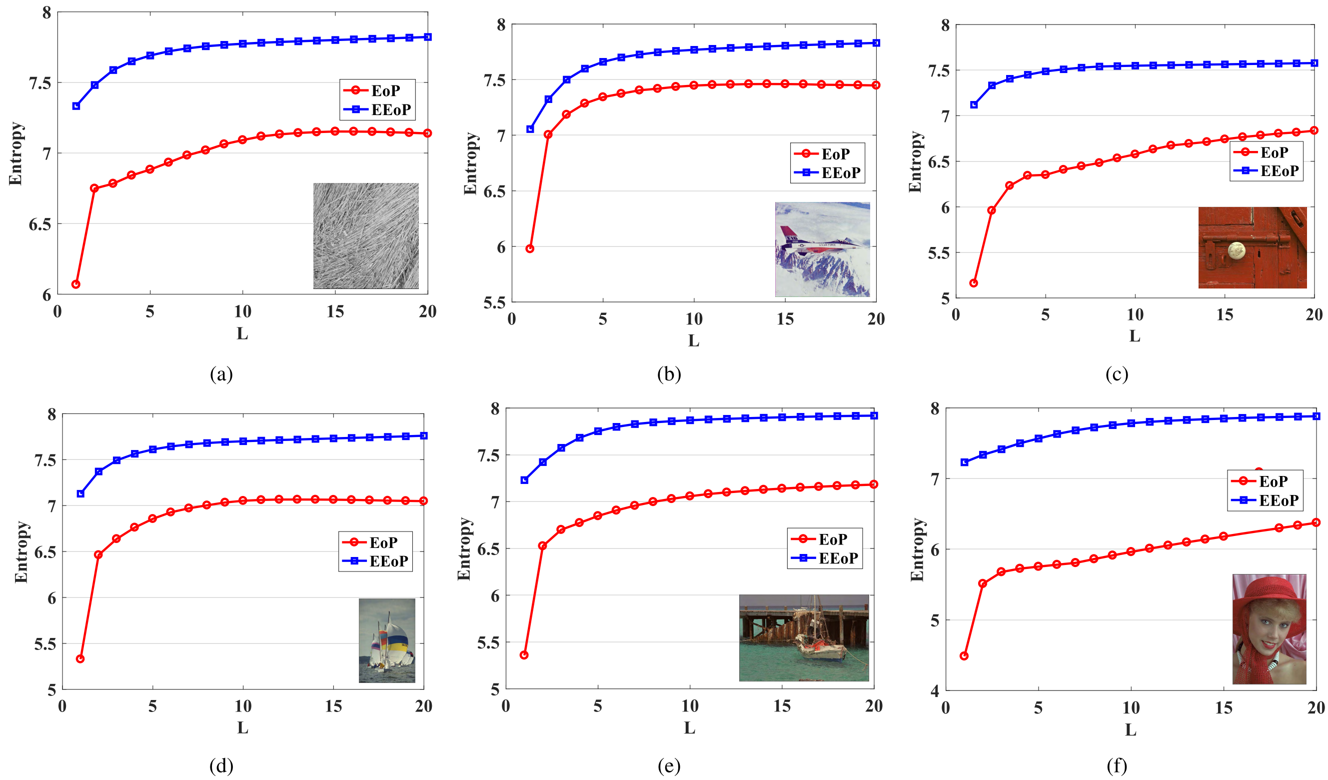


FIGURE 4. **EoP and EEoP curves in terms of the number of iterations *L*. The original image is shown in the corner of each figure.**

Based on the *Shannon* theory [1], the EoP at the L^{th} iteration can be defined as follows,

$$EoP_L = -\sum\nolimits_{i=1}^{k} \bar{p}_{L,i} \log \bar{p}_{L,i}. \tag{4}$$

As indicated in (4), the EoP is essentially defined based on the iteration index L, indicating that at each iteration of reconstruction, the visual information characterized by EoP is variant. In particular, as the visual quality becomes better along with the increase of the iteration index, it is also interesting to see that the EoP increases monotonously with L, as shown in Fig. 4. Another observation is that it converges to a saturation point, which also corresponds to the scenario that the visual quality is perceptually equivalent to the original image as the visual information becomes constant. As such, it is natural to apply the concept of EoP in JND estimation and quality evaluation, as discussed in Section III.

C. EXTENSION OF EOP

However, a major drawback of the EoP is that the coefficients in sparse representation are not taken into consideration. In particular, the PDF in EoP definition is only based on the selected primitive, regardless of the coefficient energy. To tackle this, an extension of EoP (EEoP) based on the ℓ_2 norm of the coefficients is proposed in [36]. First, a PDF considering the coefficient energy can be defined as follows,

$$N_L = \sum_{i=1}^{k} n_{L,i}, \tag{5}$$

$$p_{L,i} = \frac{n_{L,i}}{N_L}, \tag{6}$$

where $n_{L,i}$ denotes the ℓ_2 norm of the coefficients for the i^{th} primitive and the L^{th} iteration. Then, the EEoP is defined as follows,

$$EEoP_L = -\sum\nolimits_{i=1}^{k} p_{L,i} \log p_{L,i}. \tag{7}$$

In [36], the convergence of the EEoP is further verified. In particular, it is proved that the PDFs of two neighboring iterations become similar to each other as the number of iterations increases, i.e., $\lim_{L\to\infty} |p_{L+1,i} - p_{L,i}| = 0$, which can be split to the following two parts,

$$\lim_{L\to\infty} |p_{L+1,i_{\min}} - p_{L,i}| = 0, \tag{8}$$

$$\lim_{L\to\infty} |p_{L+1,i_{\max}} - p_{L,i}| = 0, \tag{9}$$

where $p_{L+1,i_{min}}$ and $p_{L+1,i_{max}}$ are the lower and upper bounds of $p_{L+1,i}$, respectively.

Some examples of the EEoP curves are shown in Fig. 4. We can observe that the curves of EEoP are more stable and robust compared to the original EoP curves. EoP is essentially based on the statistical distribution of the number of nonzero coefficients, i.e., the ℓ_0 norm, while the EEoP is essentially

FIGURE 5. Images noised by different JND models including Yang *et al.*'s [37], Liu *et al.*'s [38], Wu *et al.*'s [39] and EoP based JNDs from left to right, respectively.

based on the coefficient energy, i.e., the ℓ_2 norm. Therefore, EoP could be rather sensitive to the values close to zero and it simply ignores the coefficient amplitude. As such, a significant primitive with large coefficient value and an insignificant primitive with negligible value have the same impact on the entropy calculation in EoP. EEoP addresses this issue by taking the coefficient energy into account, thus achieving a more stable and robust representation compared to EoP. As such, EEoP is a meaningful extension of EoP, which comprehensively reflects the sparse representation process and the relationships between sparse representation and HVS.

III. APPLICATIONS OF EOP

While the fields of visual information evaluation and sparse coding are still quickly evolving, it is interesting to discuss how we could make use of EoP in real-world applications. As a powerful tool in evaluating the visual information, EoP has been successfully applied in the applications of JND estimation and visual quality evaluation. In this section, the applications of EoP will be detailed to show how EoP can play important roles in an even more extended field of scenarios. It is also envisioned that in the future more applications of the EoP or visual information evaluation methods may emerge, especially for the perceptual visual compression which relies on robust visual quality assessment algorithms.

A. JUST NOTICEABLE DIFFERENCE ESTIMATION

As the quality of the reconstructed image improves gradually with the value of L, the visual quality will reach to the saturation point such that further signal level fidelity improvement cannot further improve the visual quality. As such, it is natural to exploit the characteristics of EoP in JND estimation. In particular, given an original image X, the corresponding reconstructed image $\tilde{X}$ can be obtained by sparse representation. As such, the JND profile can be estimated by the difference between X and $\tilde{X}$, which is given by,

$$JND(X) \triangleq \left|\tilde{X} - X\right|, \tag{10}$$

where the notation $|\bullet|$ indicates the absolute operator.

To obtain the reconstruction image $\tilde{X}$, we follow the computation of EoP. In particular, the image X is used to train the dictionary using the K-SVD algorithm. Then each patch is decomposed into a linear combination of a few primitives by the OMP approach. The threshold $\tilde{L}$ can be calculated based on the definition of EoP,

$$\tilde{L} = \arg\min_{i} i, s.t. \frac{EoP_i - EoP_{i-1}}{\max_j \left(EoP_j\right) - \min_j \left(EoP_j\right)} < \varepsilon, \tag{11}$$

where ε determines the convergence condition. As such, the reconstructed image $\tilde{X}$ can be reconstructing the image with the first $\tilde{L}$ primitives, and the following primitives can be discarded. Due to the concept of EoP, we regard the reconstructed image $\tilde{X}$ as having equal quality with the original image or perceptual lossless. In this manner, the JND map can be obtained via (10).

In [34], the JND noised images are compared, as shown in Fig. 5. Though they share approximately identical PSNR value, it is interesting to see that the noise injected images with the JND derived based on EoP and sparse representation have better quality. Here, we also provide the JND maps to better show their differences, as illustrated in Fig. 6. It is also worth noting that all these JND maps contain exactly the same noise energy, as identical quantity of errors are injected into the original images. One can see that the EoP based method can concentrate the noises on the regions that have

FIGURE 6. Illustration of the JND maps including Yang *et al.*'s [37], Liu *et al.*'s [38], Wu *et al.*'s [39] and EoP based JND maps from left to right, respectively.

rich textures and details, and meanwhile keep the smooth areas with minimal distortions, yielding better visual quality. This further provides useful evidence that the JND profile guided with EoP is more effective in shaping the noise.

B. VISUAL QUALITY EVALUATION

Recently, image quality assessment (IQA) has received great interest due to its widely applications in monitoring and optimizing the multimedia systems [40]. Advanced full-reference (FR) [2], [32], [41], reduced-reference (RR) [42]–[47] and no-reference (NR) [48]–[50] IQA models have been developed to access the visual quality of natural images. Moreover, the IQA tasks have also been extended in various ways, such as screen content [51], [52], contrast [53]–[55], 3D and synthesized view [56]–[58], video streaming [59], [60] etc. Recently, various databases [61], [62] and evaluation methods [63] have also been proposed for the validation of the IQA methods.

In view of the significant importance of IQA, EoP has been adopted in IQA models in various ways. In particular, EoP serves as the measure that globally quantifies the visual information, such that the RR-IQA model can be built based on it. More specifically, inspired by the IGM theory, the RR-IQA model in [12] is derived based on the primary visual information and uncertainty. The primary visual information can be regarded as the information that can be understood in the natural scene, and this is naturally consistent with the concept of EoP. In [64], the uncertainty is represented as the discrepancy between the input signal and the best interpretation with the auto-regression model. In the context of sparse representation, by regarding sparse coding as the approximation of visual cognition process, the residuals between the input signal and the reconstructed signal with sparse representation can be treated as the uncertainty. As such, assuming the reconstructed image after the sparse representation is $\tilde{X}$, the uncertainty is defined as the entropy of the difference signal between the original and reconstructed one,

$$F_L(X) = E(X - \tilde{X}_L). \tag{12}$$

Here, E denotes the entropy calculation and again L denotes the iteration in the sparse representation process.

The RR-IQA model is finally defined as the combination of primary information and uncertainty, which are both characterized with sparse representation. In particular, with two images X and Y, which correspond to the original image and the distorted version, the quality measure is defined as

$$Q = EoP_L(X) \cdot F_L(X) - EoP_L(Y) \cdot F_L(Y). \tag{13}$$

As such, the features extracted from the original image X are transmitted and compared with those of Y, such that one value is required to be transmitted to the receiver side, which greatly facilitates the application of the RR-IQA model. The experimental results of the RR-IQA model are listed in Table 1, where the RR-IQA methods such as RRVIF [42], OSVP [43], ROCB [44] and WNISM [45], and the FR-IQA metrics such as PSNR and SSIM [2] are compared. The performance is evaluated based on Pearson linear correlation coefficient (PLCC), Spearman's rank correlation coefficient (SRCC), Kendall's rank correlation coefficient (KRCC), Root mean-squared error (RMSE) and mean absolute error (MAE). It is observed that the RR-IQA model achieves promising performance in terms of both prediction accuracy and monotonicity.

TABLE 1. Performance Comparisons of Visual Quality Prediction Based on LIVE Image Dataset.

	PLCC	SRCC	KRCC	RMSE	MAE
SSIM [2]	0.9042	0.9104	0.7311	11.669	9.228
PSNR	0.8723	0.8756	0.6865	13.360	10.509
RRVIF [42]	0.7543	0.7246	0.5438	17.937	13.675
OSVP [43]	0.8201	0.8218	0.6275	15.633	12.261
ROCB [44]	0.8866	0.8822	0.6966	12.636	9.853
WNISM [45]	0.7512	0.7599	0.5697	18.035	14.020
EoP Based [12]	**0.9146**	**0.9157**	**0.7418**	**11.050**	**8.731**

In addition, the concept of EoP has also been extended in various ways, such that they can be successfully applied in the IQA model. In [13], the ℓ_1 norm instead of ℓ_0 norm is used in the calculation of EoP, and it is interesting to observe that the EoP curve with such modification is more consistent with the perceptual quality measure SSIM [2]. As such, the perceptual stereoscopic image quality assessment method is derived. In [14], the visual primitives are classified into DCprimary, sketch and texture, and the corresponding entropy of classified primitives (EoCP) are calculated. In this manner, the differences of EoCP are used as the feature distance to characterize the perceptual loss. In [15], the EoCP is also applied in the stereoscopic IQA scheme, where the EoCP as well as the mutual information of classified primitives (MIoCP) are used in the quality prediction. In particular, the MIoCP is obtained by the two-view images to indicate the binocular cue.

As a natural extension of EoP, in the future EEoP may also play important roles in visual perception tasks, especially for the near-threshold JND profile estimation and supra-threshold IQA algorithm development. Moreover, it is also envisioned that more ways will be found with EoP and EEoP to benefit various IQA tasks, such as 3D synthesized view, contrast, as well as other relevant IQA topics.

IV. OUTLOOK

It is apparent that the sparse representation based visual information evaluation possesses many favorable properties. However, as one of the first attempt in this research topic, although it has demonstrated promising performance in various visual computing tasks, accurate and efficient visual information quantification is still in its infancy stage. While artificial intelligence is still quickly evolving, deep neural network has been shown to be advanced methods in learning discriminative prior models for natural images [65]. As such, it is natural to investigate the visual information evaluation in the context of deep neural network, which has also been verified to be highly correlated with the cognitive process of the human visual system [66], [67].

In particular, powered by the deep neural network, the current approach can be improved from the following perspectives. Firstly, the deep features instead of the primitives are extracted for visual information evaluation, which provide a more intuitive way in combining the psychological process in the HVS and visual information processing. Secondly, with the flexible representation of the deep neural networks, instead of the primitives, the features which correspond to the coefficients in sparse representation, are used in visual information evaluation. As such, the extracted information are more meaningful, leading to more robust evaluation of visual entropy. Finally, the deep neural network enables adaptive perceptual scale in computing the visual entropy, such that the perceived information can be adaptively computed with the dynamic viewing conditions and image content. However, one important issue of deep learning is that there does not exist a generic deep model that is perfectly designed and trained. As such, the accuracy of visual information evaluation may get improved gradually with the evolution of deep learning methods. This poses new challenges to visual research and opens up new space for future exploration.

Another meaningful research topic is to systematically develop the evaluation framework for validating the visual information prediction methods. In contrast with other visual computing tasks, it is difficult to obtain the ground-truth of the visual information. Therefore, a widely accepted evaluation protocol is necessary in this scenario. Moreover, regarding the application of EoP, most of the efforts focus on the perceptual visual quality assessment, and much less work has been dedicated to the visual analysis tasks. In the future, how the visual information evaluation could benefit the high level visual analysis should also be further investigated.

V. CONCLUSION

In this paper, we review recent findings on visual information evaluation based on sparse representation, and bridge the relationship between visual information and primitive representation with the concept of EoP. More specifically, the design philosophy of EoP is discussed, and the distinct properties of EoP are analyzed. It has also been demonstrated that the EoP can be successfully applied in a wide spectrum of applications such as visual quality evaluation and JND estimation. In the future, it is expected that the intelligence-oriented visual information evaluation can play more important roles in the visual processing and communication, and impact the new development of future visual-related technologies.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 5, no. 1, pp. 3–55, 2001.

[2] 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.

[3] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer, 2010, pp. 79–109.

[4] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.

[5] H.-W. Chang, H. Yang, Y. Gan, and M.-H. Wang, "Sparse feature fidelity for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 4007–4018, Oct. 2013.

[6] L. He, D. Tao, X. Li, and X. Gao, "Sparse representation for blind image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1146–1153.

[7] X. Zhang, S. Wang, S. Ma, S. Liu, and W. Gao, "Entropy of primitive: A top-down methodology for evaluating the perceptual visual information," in *Proc. Vis. Commun. Image Process. (VCIP)*, 2013, pp. 1–6.

[8] S. Wang, X. Zhang, S. Ma, and W. Gao, "Reduced reference image quality assessment using entropy of primitives," in *Proc. Picture Coding Symp. (PCS)*, 2013, pp. 193–196.

[9] J. Zhang, S. Ma, R. Xiong, D. Zhao, and W. Gao, "Image primitive coding and visual quality assessment," in *Advances in Multimedia Information Processing—PCM*, Berlin, Germany: Springer, 2012, pp. 674–685.

[10] X. Zhang, S. Wang, S. Ma, R. Xiong, and W. Gao, "Towards accurate visual information estimation with entropy of primitive," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2015, pp. 1046–1049.

[11] X. Zhang, S. Wang, K. Gu, T. Jiang, S. Ma, and W. Gao, "Sparse structural similarity for objective image quality assessment," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2015, pp. 1561–1566.

[12] S. Wang, S. Wang, K. Gu, S. Ma, and W. Gao, "Internal generative mechanism inspired reduced reference image quality assessment with entropy of primitive," in *Proc. IEEE Vis. Commun. Image Process.*, Dec. 2017, pp. 1–4.

[13] W. Shi, F. Jiang, and D. Zhao, "Image entropy of primitive and visual quality assessment," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 2087–2091.

[14] Z. Wan, Y. Liu, F. Qi, and D. Zhao, "Reduced reference image quality assessment based on entropy of classified primitives," in *Proc. Data Compress. Conf. (DCC)*, 2017, pp. 231–240.

[15] Z. Wan, F. Qi, Y. Liu, and D. Zhao, "Reduced reference stereoscopic image quality assessment based on entropy of classified primitives," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2017, pp. 73–78.

[16] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[17] J. Zhang, C. Zhao, R. Xiong, S. Ma, and D. Zhao, "Image super-resolution via dual-dictionary learning and sparse representation," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2012, pp. 1688–1691.

[18] J. Zhang, D. Zhao, R. Xiong, S. Ma, and W. Gao, "Image restoration using joint statistical modeling in a space-transform domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 6, pp. 915–928, Jun. 2014.

[19] X. Liu, D. Zhai, J. Zhou, S. Wang, D. Zhao, and H. Gao, "Sparsity-based image error concealment via adaptive dual dictionary learning and regularization," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 782–796, Feb. 2017.

[20] J. Jiang, J. Ma, C. Chen, X. Jiang, and Z. Wang, "Noise robust face image super-resolution through smooth sparse representation," *IEEE Trans. Cybern.*, vol. 47, no. 11, pp. 3991–4002, Nov. 2017.

[21] J. Jiang, R. Hu, Z. Wang, and Z. Han, "Face super-resolution via multilayer locality-constrained iterative neighbor embedding and intermediate dictionary learning," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4220–4231, Oct. 2014.

[22] J. Jiang, R. Hu, Z. Han, and Z. Wang, "Low-resolution and low-quality face super-resolution in monitoring scene via support-driven sparse coding," *J. Signal Process. Syst.*, vol. 75, no. 3, pp. 245–256, 2014.

[23] X. Zhang *et al.*, "Rate-distortion optimized sparse coding with ordered dictionary for image set compression," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.

[24] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double sparsity: Learning sparse dictionaries for sparse signal approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1553–1564, Mar. 2010.

[25] H. Xiong, Z. Pan, X. Ye, and C. W. Chen, "Sparse spatio-temporal representation with adaptive regularized dictionary learning for low bit-rate video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 4, pp. 710–728, Apr. 2013.

[26] J.-W. Kang, M. Gabbouj, and C.-C. J. Kuo, "Sparse/DCT (S/DCT) two-layered representation of prediction residuals for video coding," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2711–2722, Jul. 2013.

[27] X. Zhang, W. Lin, R. Xiong, X. Liu, S. Ma, and W. Gao, "Low-rank decomposition-based restoration of compressed images via adaptive noise estimation," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4158–4171, Sep. 2016.

[28] X. Zhang *et al.*, "Low-rank-based nonlocal adaptive loop filter for high-efficiency video compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 10, pp. 2177–2188, Oct. 2017.

[29] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[30] J. Z. Salvatierra, "New sparse representation methods; Application to image compression and indexing," Ph.D. dissertation, Human-Comput. Interaction, Univ. Rennes 1, Rennes, France, 2010.

[31] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

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

[33] D. C. Knill and A. Pouget, "The Bayesian brain: The role of uncertainty in neural coding and computation," *Trends Neurosci.*, vol. 27, no. 12, pp. 712–719, 2004.

[34] S. Ma, X. Zhang, S. Wang, J. Zhang, H. Sun, and W. Gao, "Entropy of primitive: From sparse representation to visual information evaluation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 2, pp. 249–260, Feb. 2017.

[35] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 1193–1216, 2001.

[36] S. Wang *et al.*, "Improved entropy of primitive for visual information estimation," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Nov. 2016, pp. 1–4.

[37] X. Yang, W. Lin, Z. Lu, E. Ong, and S. Yao, "Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 6, pp. 742–752, Jun. 2005.

[38] A. Liu, W. Lin, M. Paul, C. Deng, and F. Zhang, "Just noticeable difference for images with decomposition model for separating edge and textured regions," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1648–1652, Nov. 2010.

[39] J. Wu, G. Shi, W. Lin, A. Liu, and F. Qi, "Just noticeable difference estimation for images with free-energy principle," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1705–1710, Nov. 2013.

[40] Z. Wang, "Applications of objective image quality assessment methods [applications corner]," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 137–142, Nov. 2011.

[41] 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.

[42] J. Wu, W. Lin, G. Shi, and A. Liu, "Reduced-reference image quality assessment with visual information fidelity," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1700–1705, Nov. 2013.

[43] J. Wu, W. Lin, G. Shi, L. Li, and Y. Fang, "Orientation selectivity based visual pattern for reduced-reference image quality assessment," *Inf. Sci.*, vol. 351, pp. 18–29, Jul. 2016.

[44] L. Ma, S. Li, F. Zhang, and K. N. Ngan, "Reduced-reference image quality assessment using reorganized DCT-based image representation," *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 824–829, Aug. 2011.

[45] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," *Proc. SPIE*, vol. 5666, Mar. 2005, pp. 149–159.

[46] Y. Liu, G. Zhai, K. Gu, X. Liu, D. Zhao, and W. Gao, "Reduced-reference image quality assessment in free-energy principle and sparse representation," *IEEE Trans. Multimedia*, vol. 20, no. 2, pp. 379–391, Feb. 2018.

[47] Y. Zhang, T. D. Phan, and D. M. Chandler, "Reduced-reference image quality assessment based on distortion families of local perceived sharpness," *Signal Process., Image Commun.*, vol. 55, pp. 130–145, Jul. 2017.

[48] Z. Wang and A. C. Bovik, "Reduced-and no-reference image quality assessment," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 29–40, Jun. 2011.

[49] Q. Wu *et al.*, "Blind image quality assessment based on rank-order regularized regression," *IEEE Trans. Multimedia*, vol. 19, no. 11, pp. 2490–2504, Nov. 2017.

[50] Q. Wu, Z. Wang, and H. Li, "A highly efficient method for blind image quality assessment," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 339–343.

[51] X. Min, K. Ma, K. Gu, G. Zhai, Z. Wang, and W. Lin, "Unified blind quality assessment of compressed natural, graphic, and screen content images," *IEEE Trans. Image Process.*, vol. 26, no. 11, pp. 5462–5474, Nov. 2017.

[52] K. Gu *et al.*, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1098–1110, Jun. 2016.

[53] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "No-reference quality metric of contrast-distorted images based on information maximization," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4559–4565, Dec. 2017.

[54] S. Wang, K. Ma, H. Yeganeh, Z. Wang, and W. Lin, "A patch-structure representation method for quality assessment of contrast changed images," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2387–2390, Dec. 2015.

[55] Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai, "No-reference quality assessment of contrast-distorted images based on natural scene statistics," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 838–842, Jul. 2015.

[56] K. Gu, V. Jakhetiya, J.-F. Qiao, X. Li, W. Lin, and D. Thalmann, "Model-based referenceless quality metric of 3D synthesized images using local image description," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 394–405, Jan. 2017.

[57] J. Wang, A. Rehman, K. Zeng, S. Wang, and Z. Wang, "Quality prediction of asymmetrically distorted stereoscopic 3D images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3400–3414, Nov. 2015.

[58] Q. Jiang, F. Shao, W. Lin, and G. Jiang, "On predicting visual comfort of stereoscopic images: A learning to rank based approach," *IEEE Signal Process. Lett.*, vol. 23, no. 2, pp. 302–306, Feb. 2016.

[59] Z. Wang, K. Zeng, A. Rehman, H. Yeganeh, and S. Wang, "Objective video presentation QoE predictor for smart adaptive video streaming," *Proc. SPIE*, vol. 9599, Sep. 2015, Art. no. 95990Y.

[60] T. Zhao, Q. Liu, and C. W. Chen, "QoE in video transmission: A user experience-driven strategy," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 285–302, 1st Quart., 2017.

[61] K. Ma *et al.*, "Waterloo exploration database: New challenges for image quality assessment models," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 1004–1016, Feb. 2017.

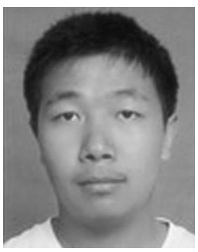
[62] V. Hosu *et al.*, "The Konstanz natural video database (KoNViD-1k)," in *Proc. 9th Int. Conf. Quality Multimedia Exper. (QoMEX)*, May 2017, pp. 1–6.

[63] Q. Wu, H. Li, F. Meng, and K. N. Ngan, "A perceptually weighted rank correlation indicator for objective image quality assessment," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2499–2513, May 2018.

[64] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41–52, Jan. 2012.
[65] L. Zhang and W. Zuo, "Image restoration: From sparse and low-rank priors to deep priors [lecture notes]," *IEEE Signal Process. Mag.*, vol. 34, no. 5, pp. 172–179, Sep. 2017.
[66] A. Berardino, V. Laparra, J. Ballé, and E. Simoncelli, "Eigen-distortions of hierarchical representations," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3533–3542.
[67] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 694–711.

SONGCHAO TAN received the B.S. degree in computer science and technology from the Dalian University of Technology, Dalian, China, in 2009, where he is currently pursuing the Ph.D. degree in computer science and technology. His research interests include video coding and quality assessment.

SHURUN WANG received the bachelor's degree from the School of Mathematical Sciences, Peking University. He is currently pursuing the master's degree with Peking University, Beijing, China. His main research interests include deep learning and machine learning, especially feature-based image compression and understanding.



XIANG ZHANG received the B.S. degree in computer science from the Harbin Institute of Technology, Harbin, China, in 2013. He is currently pursuing the Ph.D. degree with Peking University. His research interests include image/video quality assessment, video compression, and visual retrieval.

SHANSHE WANG received the B.S. degree from the Department of Mathematics, Heilongjiang University, Harbin, China, in 2004, the M.S. degree in computer software and theory from Northeast Petroleum University, Daqing, China, in 2010, and the Ph.D. degree in computer science from the Harbin Institute of Technology, Harbin, in 2014. He currently holds a post-doctoral position with Peking University. His current research interests include video compression, image and video quality assessment.

SHIQI WANG received the B.S. degree in computer science from the Harbin Institute of Technology in 2008 and the Ph.D. degree in computer application technology from the Peking University in 2014. From 2014 to 2016, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. From 2016 to 2017, he was a Research Fellow with the Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore. He is currently an Assistant Professor with the Department of Computer Science, City University of Hong Kong. He has authored over 30 technical proposals to ISO/MPEG, ITU-T, and AVS standards. His research interests include image/video compression, analysis and quality assessment.

SIWEI MA received the B.S. degree from Shandong Normal University, Jinan, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005. He held a post-doctoral position with the University of Southern California, Los Angeles, CA, USA, from 2005 to 2007. He joined the School of Electronics Engineering and Computer Science, Institute of Digital Media, Peking University, Beijing, where he is currently a Professor. He has authored over 100 technical articles in refereed journals and proceedings in the areas of image and video coding, video processing, video streaming, and transmission.

WEN GAO (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from The University of Tokyo, Tokyo, Japan, in 1991. He was a Professor of computer science with the Harbin Institute of Technology, Harbin, China, from 1991 to 1995, and a Professor with the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. He is currently a Professor of computer science with Peking University, Beijing. He has authored extensively, including five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics. He served or serves on the Editorial Board for several journals, such as the IEEE Transactions on Circuits and Systems for Video Technology, the IEEE Transactions on Multimedia, the IEEE Transactions on Image Processing, the IEEE Transactions on Autonomous Mental Development, the *EURASIP Journal of Image Communications*, and the *Journal of Visual Communication, and Image Representation*. He was the Chair of a number of prestigious international conferences on multimedia and video signal processing, such as the IEEE International Conference on Multimedia and Expo and Association for Computing Machinery Multimedia, and served on the advisory and technical committees of numerous professional organizations.