Convolutional auto-encoder based multiple description coding network

Lili Meng[1,2*], Hongfei Li[1], Jia Zhang[1,2], Yanyan Tan[1,2], Yuwei Ren[1,2], Huaxiang Zhang[1,2]
[1]School of Information Science and Engineering, Shandong Normal University
Jinan, Shandong, China
[2]Institute of Data Science and Technology, Shandong Normal University
Jinan, Shandong, China
*Corresponding author: Lili Meng
E-mail:mengll_83@hotmail.com

Abstract

When data is transmitted over an unreliable channel, the error of the data packet may result in serious degradation. The multiple description coding (MDC) can solve this problem and save transmission costs. In this paper, we propose a deep multiple description coding network (MDCN) to realize efficient image compression. Firstly, our network framework is based on convolutional auto-encoder (CAE), which include multiple description encoder network (MDEN) and multiple description decoder network (MDDN). Secondly, in order to obtain high-quality reconstructed images at low bit rates, the encoding network and decoding network are integrated into an end-to-end compression framework. Thirdly, the multiple description decoder network includes side decoder network and central decoder network. When the decoder receives only one of the two multiple description code streams, side decoder network is used to obtain side reconstructed image of acceptable quality. When two descriptions are received, the high quality reconstructed image is obtained. In addition, instead of quantization with additive uniform noise, and SSIM loss and distance loss combine to train multiple description encoder networks to ensure that they can share structural information. Experimental results show that the proposed framework performs better than traditional multiple description coding methods.
.

1. Introduction

$\mathbf{W}$ith the continuous development of information technology, high-quality transmission of image and video signals in the network has become more and more important. The MDC method is mainly used to solve the problem of unreliable transmission caused by packet loss or bit errors in image communication. [1]. The MDC divides the source into M descriptions containing M subsets, these descriptions are transmitted through different channels, which can ensure that even if some descriptions are lost, a reconstructed image with acceptable quality can be obtained. Therefore, the more descriptions the receiver gets, the better the quality of the reconstructed image.

MDC method can achieve higher compression efficiency than conventional single-description compression technology. The MDC method is widely used in the fields of image, video and multimedia signals [2-8]. In [2], the first practical MDC scheme-Multiple Deseription Scalar Quantization (MDSQ) is proposed, which involves the problem of index allocation. In [3], a multiple descriptive scalar quantizer with entropy constraints is proposed. In [4], a universal multiple description scalar quantizer (UMDSQ) is proposed, which can realize a continuous trade-off between center and side distortion at high rates, without extensive training. In [5], a novel algorithm for optimal generalized multiple description vector quantizer (GMDVQ) is proposed, which shows good performance in the case of extensive packet loss. In [6], a lattice-based MDVQ is designed. In [7], a MDVQ with a lattice codebook is designed, which solves the main problem(the label problem) in the design. In [8], a MD lattice vector quantization technique for both descriptions is introduced, in which the fine and coarse codebooks are lattices.

Image transform coding refers to converting an image described in the form of a pixel in a spatial domain to a transform domain, and expressing it in the form of transform coefficients. Appropriate transformation can transform the dispersion distribution of image energy in the spatial domain into a relatively concentrated distribution in the transform domain, to achieve the purpose of removing redundancy. Combined with other coding techniques such as quantization, "z" scanning and entropy coding, effective compression of image information can be obtained. In [9], introduces an MDC framework based on pairwise correlation transforms that introduces correlations between pairs of transform coefficients through pairwise correlation transformations. In [10], introduces a method for generalized multiple description coding (GMDC) using correlation transforms, which is a generalization of [11], extends the technique to more than two descriptions, and is very effective in improving robustness with a small amount of redundancy. In [12], two modified MDC schemes based on random offset quantizer (MDROQ) and uniform offset quantizer (MDUOQ) are proposed. Both schemes use two-rate prediction coding and sequence prediction. [13,14] is an extension of MDROQ. In [15], a new MDC method using transform coding framework is proposed. The transform cardinality is selected so that the coefficients are correlated in pairs, and then each pair of correlation coefficients is split between the two descriptions.

MDC is an effective method of combating bursty packet loss in the Internet and wireless networks. In [16], the application of two generalized multiple description coding(GMDC) [17,18] image coding methods is reported. In [19], based on the characteristics of human visual model, a MDC scheme based on just noticeable difference (JND) is proposed. In [20], error resilient data compression algorithms based on wavelets, MDSQ and erasure-resilient codes are introduced. [21] introduces the principle of using a time-predicted design MD video

encoder, experiments show that combining MDC with multiple path transport (MPT) can bring significant performance gains.[22] introduces a new framework for multiple description video coding (MDVC), and the multiple description motion compensation (MDMC) encoder can operate in a lower redundancy range than the video redundancy coding (VRC) method [23]. [24] introduces a dual-channel distributed MDVC scheme based on MDWZ codec, which solves the problem of MDVC drift on the packet loss channel. Improve the balance between the two descriptions through a bit plane extraction scheme. In [25], a method of domain-based MDC of images and video is proposed.

The convolutional auto-encoder utilizes the unsupervised learning method of the traditional auto-encoder, combines the convolution and pooling operations of the convolutional neural network (CNN) to achieve feature extraction, and finally implements a deep neural network through stack. [26] introduces a lossy image compression method based on CAE to realize a high quality image reconstruction, and principal component analysis (PCA) is used to decorrelate each feature map. In [27], the performance comparison of three compression architectures of convolutional autoencoders (CAEs), generative adversarial networks (GANs) as well as super-resolution (SR) is proposed, CAEs has better performance for lossy compression. In [28], a compression auto-encoder method for lossy image compression is proposed, and the sub-pixel architecture [29] is set to further improve the computational efficiency.

Recently, MDC has been combined with CNN for image processing. In [30], according to the contextual features of the image, a MDC framework based on CNN is proposed. As far as we know, this is the first work that combines CNN and MDC for image compression. In [31], propose a symmetric CAE-based MDC framework.

In this paper,we propose a multiple description coding network (MDCN) based on CAE, the entire network framework is implemented on tensorflow. Our main contributions are lists as follows:

1) We design a MDCN based on CAE, the CAE network architecture used in multiple description encoder networks (MDEN) and multiple description decoder networks (MDDN) to achieve efficient image compression.

2) The MDDN based on CAE is achieved, which includes side decoder network (SDN) and central decoder network (CDN). When the decoder receives only one of the two multiple description code streams, side decoder network is used to obtain an side decoded image of acceptable quality by eliminating compression artifacts. When two descriptions are received, they are taken as the input of the central decoder network at the same time, and the high quality reconstructed image is obtained.

3) Because the rounding function in quantization is not differentiable, the additive uniform noise is added to imitate the quantization noise during the optimization process and we train our whole network in an end-to-end manner.

4) SSIM loss and distance loss combine to train multiple description encoder networks to ensure that they can share structural information even when divided into multiple descriptions.

The rest of this paper is oranzied as follows. First, The Section 2 introduces the related work. And the Section 3 introduces the proposed framework, including the MDEN network and the MDDN network. The experimental results are given in the Section 4. We conclusion this paper in the Section 5.

2. Related Work

2.1 Multiple description coding method

The MDC method can prevent the problem of image and video quality degradation, which is caused by packet loss in image communication of a noisy channel, thereby saving transmission cost. The MDC divides the source into M descriptions containing M subsets, these descriptions are transmitted through different channels, and the probability of simultaneous errors on each channel is very low. By generating a plurality of independently decodable descriptions of the encoding, it is ensured that even if some descriptions are lost, a reconstructed image with acceptable quality can be obtained, and as the received description increases, the quality of the image increases. Reconstruct a quality acceptable image by using part of the information, therefore, the MDC method plays a very important role in the fields of image coding and video coding.
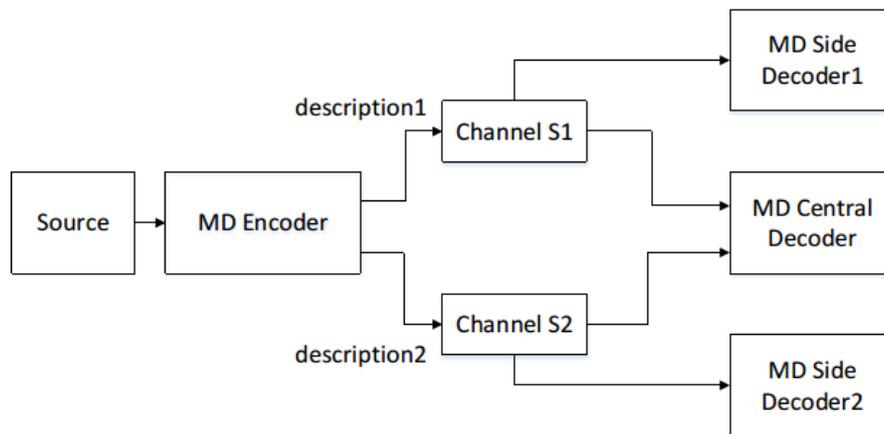

Fig. 1. The basic model of MDC

The MDC was officially presented at the Shannon Theory Research Conference in September 1979, when Gersho, Ozarow, et al raised the following questions [32]: If a source is represented by two separate descriptions, what are the restrictions on the quality of the source when these descriptions are separated or combined? This problem is called a MD problem. **Fig. 1** shows the basic model of MDC. The source generates two descriptions by the MD encoder, which are transmitted to the receiver on two separate channels S1 and S2, respectively. The receiver uses different decoders, if all the descriptions are completely received, the signals pass through the central decoder, and obtain high-quality reconstruction effects according to the important information of each description; If only a part of the description is received, the signals pass through the side decoder, and recovers the lost part of the information from the redundant information carried by the received description, so as to obtain an acceptable reconstruction quality. Therefore, the more descriptions receiver gets, the better the quality of the reconstructed image.

2.2 Additive uniform noise

Noise often appears as an isolated pixel or pixel block that causes strong visual effects on the image. Additive noise generally refers to thermal noise, shot noise, etc, and their relationship to signals is additive. In the general communication, additive randomness is regarded as the background noise of the system. The additive noise of the channel is independent of the useful signal, which always interferes with useful signals, therefore it is inevitable to cause harm to

the channel. The noise $n(t)$ is always present with or without a signal, therefore, it is usually called additive noise or additive interference. The sources of additive noise in a channel are generally divided into three areas: artificial noise, natural noise and internal noise.

The interference effect of noise on the signal in the channel is represented by the additive relationship with the signal, this channel is called an additive channel and this noise is called additive noise.

$$I(t) = O(t) + n(t) \tag{1}$$

where $I(t)$, $O(t)$ and $n(t)$ represent random signals of input, output of the channel and random signals of noise, respectively.

Since the derivative of the quantizer is discontinuous, in order to solve this problem, an additive uniform noise source is used instead of the deterministic quantizer [33,28], and approximated the objective function with one that is continuously differentiable. In [33], an end-to-end optimization framework for nonlinear transform coding is proposed, which uses additive uniform noise instead of quantization to relax the discontinuity problem to the differentiable problem. In [28], a new method of optimizing lossy image compression auto-encoder is proposed to solve the non-differentiability of trained auto-encoder.

3. The proposed framework

In this paper, a MDCN network is introduced to efficiently compress images, which can solve the problem of severe degradation of image and video quality due to packet loss or bit error for unpredictable channels.

3.1 Framework of the proposed scheme

In this paper, we propose a MDCN network based on CAE, and we design a MDEN network and a MDDN network, which consists of SDN network and CDN network, as depicted in the **Fig. 2**. The framework is built upon CAE network [34]. In order to obtain high-quality reconstructed images at low bit rates, the MD-encoder network and MD-decoder network are integrated into an end-to-end compression framework [35,36]. We get the number of bits used by calculating different quality factors (QF).

In the MDEN network, represented by the encoding function $Y_1 = f_\theta(X) + u$ and $Y_2 = f_\theta(X) + u$, and the network is responsible for generating different descriptions Y_1 and Y_2 from the original image X of size $M \times N$. Here, θ is the optimized parameters in the MDEN network. We designed a series of convolution operations, replacing the pooling layer with the convolutional layer to preserve as much image information as possible. In this paper, the additive uniform noise, $Y = Y + u$, which replace the quantization during the optimization process. [28] introduces the effect of rounding function, additive uniform noise and stochastic rounding function (which similar to the binarization proposed in [37]) when used as replacements in JPEG compression.

In the MDDN network, the decoder network mirrors the architecture of the encoder network. The MDDN network consists of SDN1 network, SDN2 network and CDN network, and the networks are respectively represented by the decoding function $\hat{X}_1 = g_\phi(Y_1)$, $\hat{X}_2 = g_\phi(Y_2)$ and $\hat{X} = g_\phi(Y_1 + Y_2)$. Here, X, $\hat{X}_1, \hat{X}_2, \hat{X}$ represents the original image, the side reconstructed image 1, the side reconstructed image 2, and the central reconstructed image, and ϕ is the optimized parameters in the MDDN network.

As far as we know, it is not easy to jointly train the MDEN network and the MDDN network, we train our whole network in an end-to-end manner. And the additive uniform noise is added to imitate the quantization noise during the optimization process
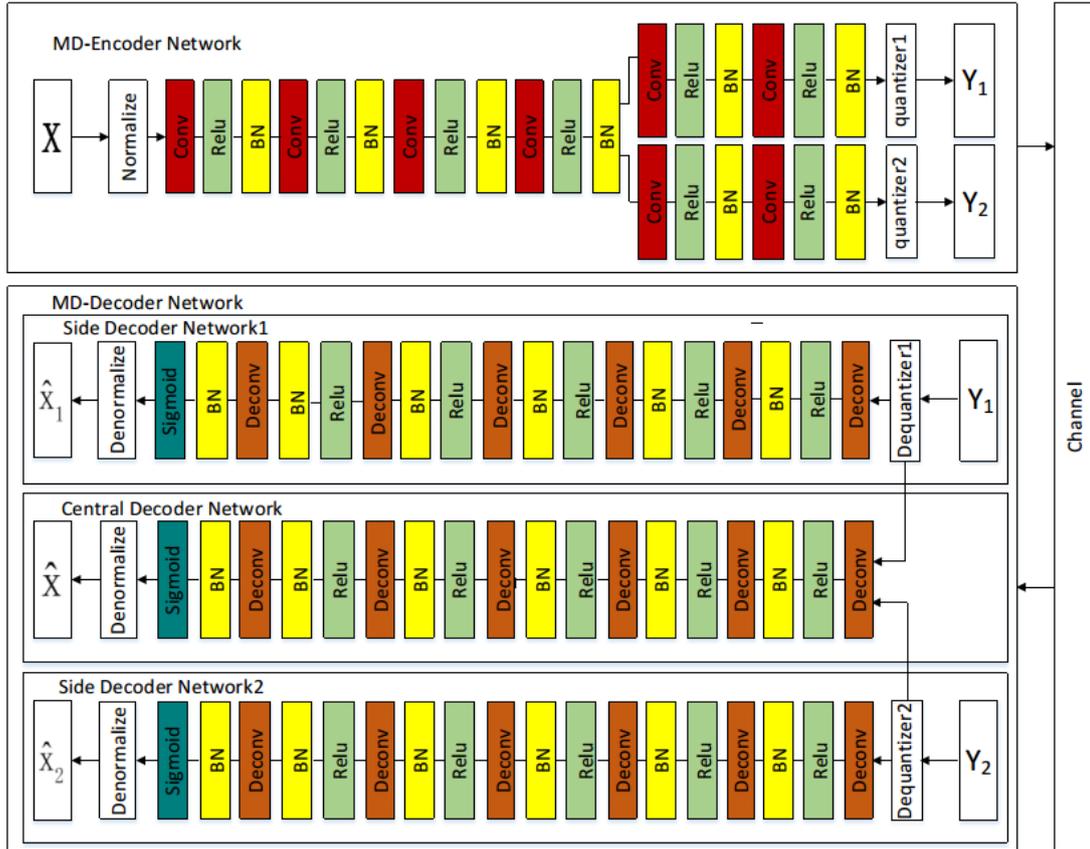

Fig. 2. The framework of MDC network based on convolutional auto-encoder(CAE)

3.2 Objective function

The objective function for our framework is written as follows:

$$L_{MAE}(X, \hat{X}_1, \hat{X}_2, \hat{X}) + L_{SSIM}(X, Y_1, Y_2, \theta) + \beta L_{reg} \tag{2}$$

where, X, $\hat{X}_1, \hat{X}_2, \hat{X}$ represents the original image, the side reconstructed image 1, the side reconstructed image 2, and the central reconstructed image, θ is the optimized parameters in the MDEN network, β is a hyper-parameters.

L2 loss function, Mean Square Error (MSE), is the most commonly used regression loss function in traditional single-description image compression. However, if the difference between the real value and the predicted value is greater than 1, the MSE will further increase the error. Therefore, a model using MSE gives greater weight to the outliers than using Mean Absolute Error (MAE) to calculate losses. Therefore, our framework uses the MAE loss function as the first part of the MD reconstruction loss for side reconstructed images and central reconstructed images, which can be written as follows:

$$L_{MAE}(X, \hat{X}_1, \hat{X}_2, \hat{X}) = \frac{1}{M \cdot N} \sum_i (|X_i - \hat{X}_{1i}|) + \frac{1}{M \cdot N} \sum_i (|X_i - \hat{X}_{2i}|) + \frac{1}{M \cdot N} \sum_i (|X_i - \hat{X}_i|) \tag{3}$$

where, i represents the pixel i.

Combined with the research of neural network, it is found that when humans measure the distance between two images, they pay more attention to the structural similarity of the two images, instead of calculating the difference between the two images pixel by pixel. In order to counteract the defect that L1 loss and L2 loss can not measure image similarity, the SSIM method is proposed to measure the similarity of two images. Our framework uses the SSIM loss [38] and distance loss combine to train multiple description encoder networks, which ensure that they can share structural information even when divided into multiple descriptions. The whole loss function can be written as follows:

$$L_{SSIM}(X,Y_1,Y_2,\theta) = L_{SSIM}(X,Y_1) + L_{SSIM}(X,Y_2) + \alpha L_{dis} \tag{4}$$

where α is a hyper-parameters, SSIM is equivalent to normalizing the data, and then calculating the image block illuminance $l(X,Y_1)$ (the mean of the image block), the contrast $c(X,Y_1)$ (the variance of the image block) and the normalized pixel vector $s(X,Y_1)$, and multiplying them.

$$L_{SSIM}(X,Y_1) = -\frac{1}{M \cdot N}\sum_i L_{SSIM}(X_i - Y_{1i}) \tag{5}$$

$$L_{SSIM}(X_i - Y_{1i}) = l(X_i - Y_{1i}) \cdot c(X_i - Y_{1i}) \cdot s(X_i - Y_{1i})$$
$$= \frac{(2\mu_{X_i}\mu_{Y_{1i}} + c_1)(2\sigma_{X_iY_{1i}} + c_2)}{(\mu_{X_i}^2 + \mu_{Y_{1i}}^2 + c_1)(\sigma_{X_i}^2 + \sigma_{Y_{1i}}^2 + c_2)} \tag{6}$$

where μ_{X_i},σ_{X_i} represents the mean value and variance of the pixel i of the image X, respectively, and similarly, $\mu_{Y_{1i}},\sigma_{Y_{1i}}$ are also expressed in this way. $\sigma_{X_iY_{1i}}$ represents the covariance of the pixels i of the image X and Y_1. $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$ are constants (e.g., $k_1 = 0.01$ and $k_2 = 0.03$).

3.3 Network architecture

The MDEN network proposed in this paper includes six convolutional layers. As the activation function of each convolutional layer, we apply the Recetified Linear Unit (ReLU) function. And we add Batch Normalization (BN) layer after each convolutional layer, which speeds up network convergence. In order to get more image information, there is no pooling layer. Except that the convolution kernel of the first layer and the last layer is set to 5×5, the convolution kernel of the remaining layers is set to 3×3, and the stride of the last layer is set to 2. The network inputs a source, but outputs two descriptions, as shown in the **Fig. 2**.
The MDDN network includes SDN1 network, SDN2 network, and CDN network. The SDN1 network and the SDN2 network can share parameter settings. And the CDN network uses the input of the SDN1 network and the SDN2 network as its own input to obtain the central reconstructed image. They all use six deconvolutional layers [39]. Use ReLU as the activation function for each layer, the last layer does not add the activation function, and each deconvolutional layer adds the BN layer. As shown in the **Fig. 2**, the CDN network can have two descriptions as input at the same time, and the other two networks have only one input. At the end of the network, we add sigmoid to map the real field to the [0,1] space. We set the first layer and the last layer convolution kernels of the network to 5×5, and the remaining layer convolution kernels to 3×3. The stride of the last deconvolutional layer is set to 2, and the remaining layer with stride of 1.

4. Experimental Results

To measure the performance of the proposed framework, experimental results of objective and visual quality are given, which use PSNR and MS-SSIM [40] to measure objective quality, and the rate is measured in terms of bit per pixel (bpp).


（a1）


（a2）

（a3）

（a4）



（b1）

（b2）

（b3）   （b4）
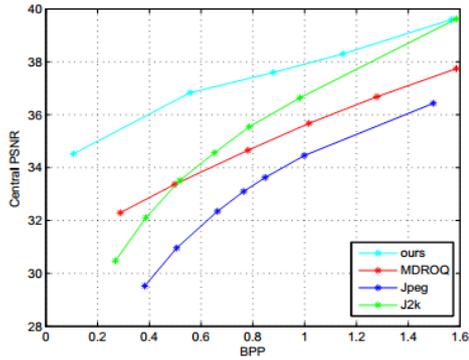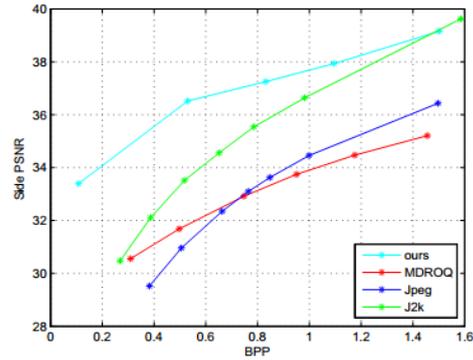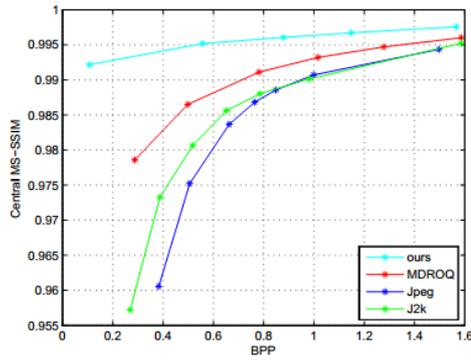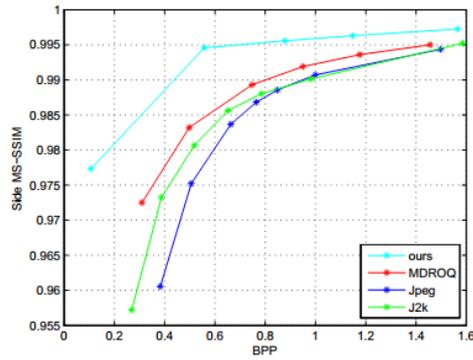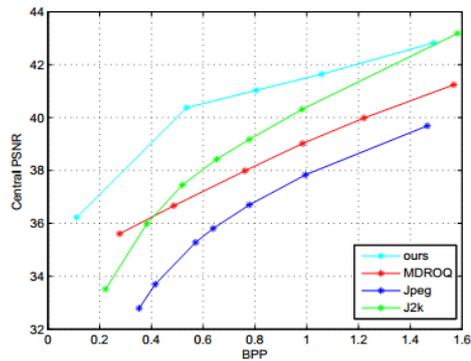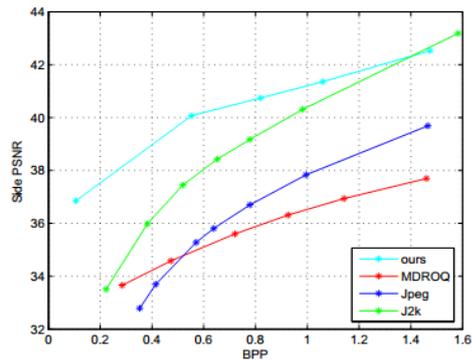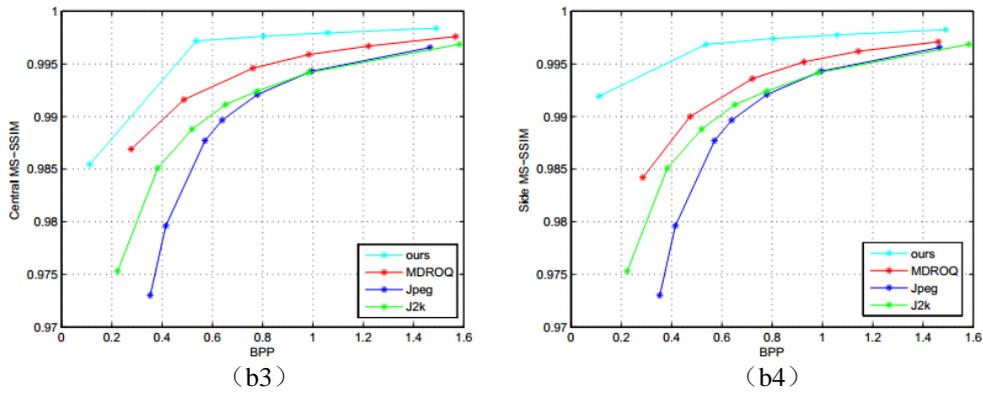
Fig. 3. Objective quality comparison of different methods for image boat and lena

（c1）   （c2）

（c3）   （c4）

（d1）   （d2）

（d3）   （d4）
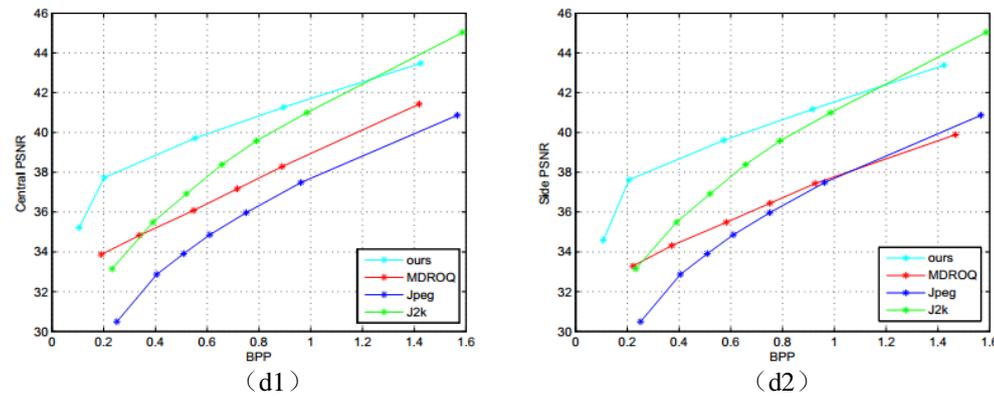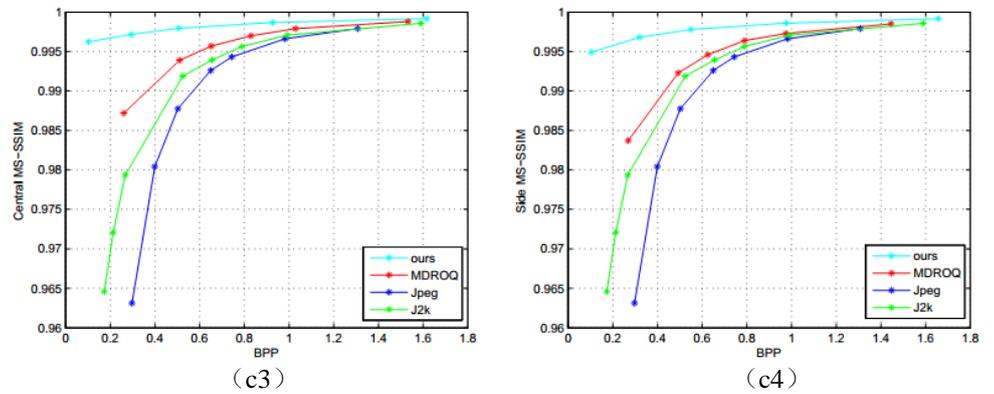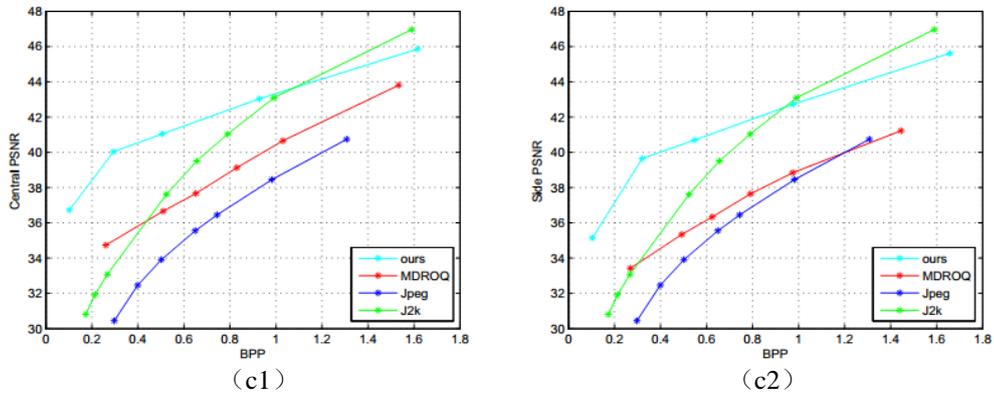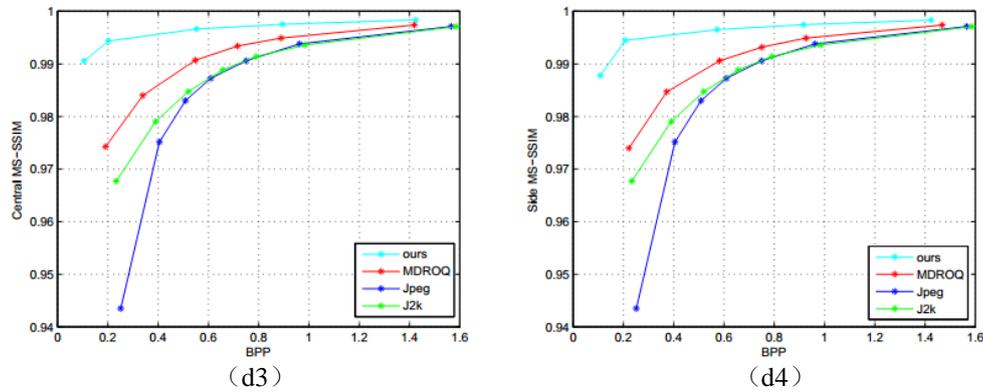
Fig. 4. Objective quality comparison of different methods for image red door and girl
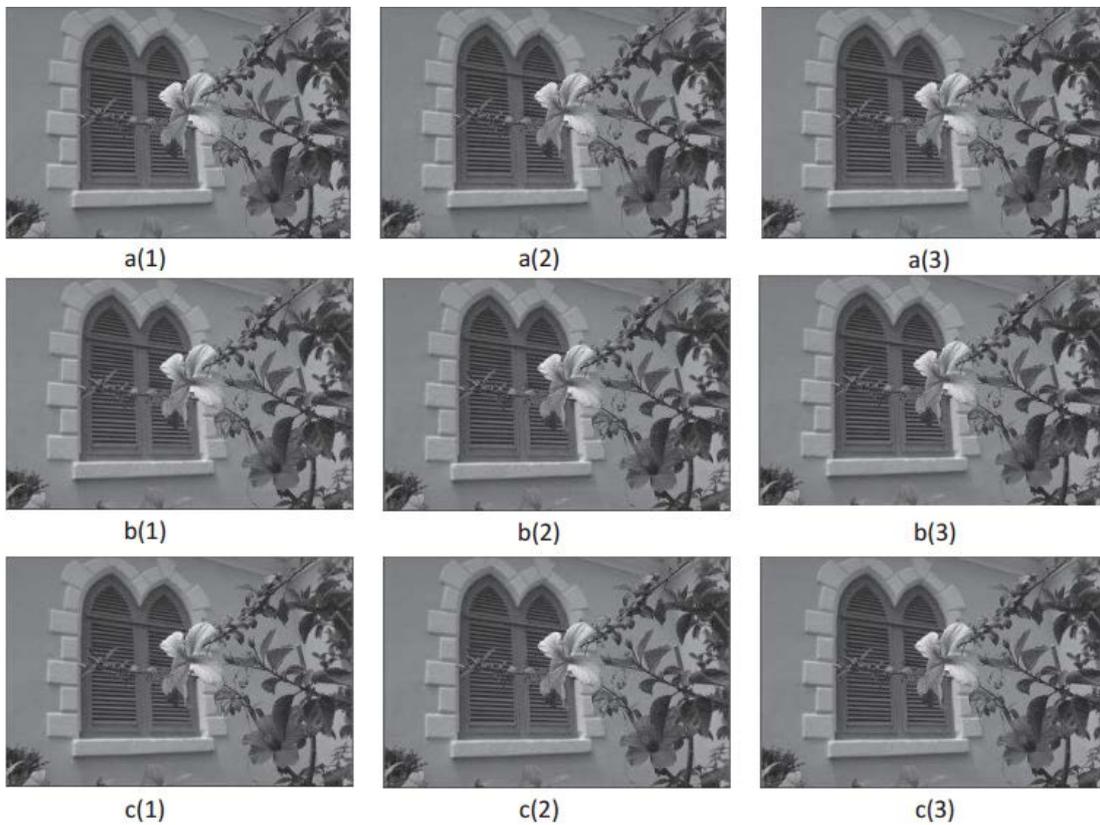
Fig. 5. Comparison of visual quality of different methods of red door.

4.1 Experimental data

Our framework is based on tensorflow [41]. We trained our network framework on 400 images with size 180×180 from [42], and the images were flipped and cropped to get the final 1600 images with size 160×160 for the training set.. Set4 and Kodak PhotoCD dataset with size of 768×512 or 512×768 are our testing set. The network uses Adam optimization algorithm [43].

4.2 Experimental results of objective and visual quality

We use PSNR and MS-SSIM to measure the performance of the objective quality of the proposed framework. We compare our framework with JPEG [44], JPEG2000 and MDROQ [12], which is denoted as "MDROQ". The proposed framework is marked as "ours". This paper gives a comparison of the PSNR and MS-SSIM of lena, boat, red door and girl.

The **Fig. 3** shows comparison of objective quality measures of the side reconstructed image and the central reconstructed image of boat and lena. (a1, b1), (a2, b2) represent the PSNR values of the central reconstructed image and side reconstructed image of the two images, respectively. (a3, b3), (a4, b4) represent the MS-SSIM values of the central reconstruction image and side reconstruction image of the two images, respectively. And the **Fig. 4** shows comparison of objective quality measures of the side reconstructed image and the central reconstructed image of red door and girl. (c1, d1), (c2, d2) represent the PSNR values of the central reconstructed image and side reconstructed image of the two images, respectively. (c3, d3), (c4, d4) represent the MS-SSIM values of the central reconstruction image and side reconstruction image of the two images, respectively. From the **Fig. 3** and the **Fig. 4**, it can be seen that our method has more PSNR and MS-SSIM gain than the other methods, especially at low bit rates. But the PSNR value of our proposed method is lower than JPEG 2000 at high bit rate.

In this paper, we compare the visual quality of the proposed framework with JPEG, JPEG2000 and MDROQ, as shown **Fig. 5**. (a1) represent the original image; (a2-a3) represent the reconstructed images obtained by the JPEG (0.5bpp) and JPEG2000 (0.52bpp), respectively; (b1-b3) represent the visual image of the central reconstruction image and two side reconstruction image of MDROQ (0.51bpp), and (c1-c3) represent the visual image of the central reconstruction image and two side reconstruction image of "ours" (0.5bpp). It can be seen that the visual quality of the reconstructed image of the proposed framework is better than the other methods.

5. Conclusion

In this paper, we propose a multiple description coding network based on CAE, which includes multiple description encoder network, multiple description side decoder network and multiple description central decoder network. Firstly, the CAE network architecture used in both multiple description encoder networks and multiple description decoder networks to achieve efficient image compression. Secondly, the two networks are integrated with an end-to-end compression framework to get high quality reconstructed images at low bit rate. In the proposed scheme, instead of quantization with additive uniform noise in our network, SSIM loss and distance loss combine to train multiple description encoder networks to ensure that they can share structural information even when divided into multiple descriptions. Through the test of two data sets, it is verified that our proposed multiple description coding network method has better performance than the traditional multiple description image compression methods.

References

[1] Goyal,V.K, "Multiple description coding:compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74-93, 2001. Article (CrossRef Link)

[2] Vaishampayan V A, "Design of multiple description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp.821-834, 1993. Article (CrossRef Link)

[3] Vaishampayan V A, Domaszewicz J, "Design of entropy-constrained multiple-description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 245-250, 1994. Article (CrossRef Link)

[4] Tian C, Hemami S S, "Universal multiple description scalar quantization: analysis and design," *IEEE Transactions on Information Theory,* vol. 50, no. 9, pp. 2089-2102, 2004. Article (CrossRef Link)

[5] Fleming M, Effros M, "Generalized multiple description vector quantization," in *Proc. of DCC'99 Data Compression Conference (Cat. No. PR00096). IEEE*, pp. 3-12, 1999. Article (CrossRef Link)

[6] Servetto S D, Vaishampayan V A, Sloane N J A, "Multiple description lattice vector quantization," in *Proc. of DCC'99 Data Compression Conference (Cat. No. PR00096). IEEE*, pp. 13-22, 1999. Article (CrossRef Link)

[7] Vaishampayan V A, Sloane N J A, Servetto S D, "Multiple Description Vector Quantization with Lattice Codebooks: Design and Analysis," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1718-1734, 2001. Article (CrossRef Link)

[8] Goyal V K, Kelner J A, Kovacevic J, "Multiple description vector quantization with a coarse lattice," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 781-788, 2002. Article (CrossRef Link)

[9] Wang Y, Orchard M T, Vaishampayan V, et al, "Multiple description coding using pairwise correlating transforms," *IEEE Transactions on Image Processing,* vol. 10, no. 3, pp. 351-366, 2001. Article (CrossRef Link)

[10] Goyal V K, Kovacevic J, "Generalized multiple description coding with correlating transforms," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2199-2224, 2001. Article (CrossRef Link)

[11] Orchard M T, Wang Y, Vaishampayan V, et al, "Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms," in *Proc. of of International Conference on Image Processing. IEEE*, vol.1, pp. 608-611, 1997. Article (CrossRef Link)

[12] Meng L, Liang J, Samarawickrama U, et al, "Multiple description coding with randomly and uniformly offset quantizers," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 582-595, 2013. Article (CrossRef Link)

[13] Zong J, Meng L, Tan Y, et al, "Adaptive reconstruction based multiple description coding with randomly offset quantizations," *Multimedia Tools and Applications*, vol. 77, no. 20, pp. 26293-26313, 2018. Article (CrossRef Link)

[14] Zong J, Meng L, Tan Y, et al, "Perceptual multiple description coding with randomly offset quantizers," in *Proc. of 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA). IEEE*, pp. 1-5, 2016. Article (CrossRef Link)

[15] Wang Y, Orchard M T, Reibman A R, "Multiple description image coding for noisy channels by pairing transform coefficients," in *Proc. of of First Signal Processing Society Workshop on Multimedia Signal Processing. IEEE*, pp. 419-424, 1997. Article (CrossRef Link)

[16] Goyal V K, Kovacevic J, Arean R, et al, "Multiple description transform coding of images," in *Proc. of International conference on image processing*, vol. 1, pp. 674-678, 1998. Article (CrossRef Link)

[17] Goyal V K, Kovacevic J, "Optimal multiple description transform coding of Gaussian vectors," in *Proc. of DCC'98 Data Compression Conference (Cat. No. 98TB100225). IEEE*, pp. 388-397, 1998. Article (CrossRef Link)

[18] Goyal V K, Vetterli M, Kovacevic J, "Multiple description transform coding: Robustness to erasures using tight frame expansions," in *Proc. of 1998 IEEE International Symposium on Information Theory (Cat. No. 98CH36252).IEEE*, pp. 408, 1998. Article (CrossRef Link)

[19] Zong J, Meng L, Zhang H, et al, "JND-based Multiple Description Image Coding," *KSII Transactions on Internet and Information Systems*, vol. 11, no. 8, pp. 3935-3949, 2017. Article (CrossRef Link)

[20] Servetto S D, Ramchandran K, Vaishampayan V A, et al, "Multiple description wavelet based image coding," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 813-826, 2000. Article (CrossRef Link)

[21] Wang Y, Reibman A R, Lin S, "Multiple description coding for video delivery," *Proceedings. of the IEEE*, vol. 93, no. 1, pp. 57-70, 2005. Article (CrossRef Link)

[22] Wang Y, Lin S, "Error-resilient video coding using multiple description motion compensation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 6, pp.438-452, 2002. Article (CrossRef Link)

[23] Wenger S, "Video redundancy coding in H. 263+," in *Proc. of 1997 International Workshop on Audio-Visual Services over Packet Networks*, 1997. Article (CrossRef Link)

[24] Fan Y, Wang J, Sun J, "Distributed Multiple Description Video Coding on Packet Loss Channels," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 20, no. 6, pp. 1768, 2001. Article (CrossRef Link)

[25] Bajic I V, Woods J W, "Domain-based multiple description coding of images and video," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1211-1225, 2003. Article (CrossRef Link)

[26] Cheng Z, Sun H, Takeuchi M, et al., "Deep convolutional autoencoder-based lossy image compression," in *Proc. of 2018 Picture Coding Symposium (PCS). IEEE*, pp. 253-257, 2018. Article (CrossRef Link)

[27] Cheng Z, Sun H, Takeuchi M, et al., "Performance Comparison of Convolutional AutoEncoders, Generative Adversarial Networks and SuperResolution for Image Compression," in *Proc. of CVPR Workshops*, pp. 2613-2616, 2018. Article (CrossRef Link)

[28] Theis L, Shi W, Cunningham A, et al, "Lossy Image Compression with Compressive Autoencoders," *arXiv preprint arXiv:1703.00395*, 2017. Article (CrossRef Link)

[29] Shi W, Caballero J, Huszr F, et al, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. of of the IEEE conference on computer vision and pattern recognition*, pp. 1874-1883, 2016. Article (CrossRef Link)

[30] Zhao L, Bai H, Wang A, et al, "Multiple Description Convolutional Neural Networks for Image Compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 8, pp. 2494-2508, 2019. Article (CrossRef Link)

[31] Li H, Meng L, Zhang J, et al, "Multiple Description Coding Based on Convolutional Auto-encoder," *IEEE Access*, vol. 7, pp. 26013-26021, 2019. Article (CrossRef Link)

[32] Gamal A A E, Cover T M, "Achievable Rates for Multiple Descriptions," *IEEE Transactions on Information Theory*, vol. 28, no. 6, pp. 851-857, 1982. Article (CrossRef Link)

[33] Ball, Johannes, Laparra V, Simoncelli E P, "End-to-end optimization of nonlinear transform codes for perceptual quality," in *Proc. of 2016 Picture Coding Symposium (PCS). IEEE*, pp. 1-5, 2016. Article (CrossRef Link)

[34] Dumas T, Roumy A, Guillemot C, "Autoencoder based image compression: can the learning be quantization independent?," in *Proc. of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE*, pp. 1188-1192, 2018. Article (CrossRef Link)

[35] Tao W, Jiang F, Zhang S, et al, "An End-to-End Compression Framework Based on Convolutional Neural Networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 3007-3018, 2018. Article (CrossRef Link)

[36] Ball J, Laparra V, Simoncelli E P, "End-to-end optimized image compression," *arXiv preprint arXiv:1611.01704*, 2016. Article (CrossRef Link)

[37] Toderici G, O'Malley S M, Hwang S J, et al, "Variable Rate Image Compression with Recurrent Neural Networks," *arXiv preprint arXiv:1511.06085*, 2015. Article (CrossRef Link)

[38] Wang Z, Bovik A C, Sheikh H R, et al, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, 2004. Article (CrossRef Link)

[39] Zeiler, Matthew D., Graham W. Taylor, and Rob Fergus, "Adaptive deconvolutional networks for mid and high level feature learning," *ICCV*, Vol. 1, No. 2, 2011. Article (CrossRef Link)

[40] Wang Z, Simoncelli E P, Bovik A C, "Multiscale structural similarity for image quality assessment," in *Proc. of The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003. Ieee*, Vol. 2, pp. 1398-1402, 2003. Article (CrossRef Link)

[41] Abadi M, Agarwal A, Barham P, et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *arXiv preprint arXiv:1603.04467*, 2016. Article (CrossRef Link)

[42] Chen Y, Pock T, "Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, no. 6, pp. 1256-1272, 2015. Article (CrossRef Link)

[43] Kingma D P, Ba J, "Adam: A method for stochastic optimization," in *Proc. of conference paper at ICLR 2015*, pp.1-15, 2015. Article (CrossRef Link)

[44] Wallace G K, "The JPEG still picture compression standard," *IEEE transactions on consumer electronics*, vol. 38, no. 1, pp. xviii-xxxiv, 1992. Article (CrossRef Link)

Lili Meng is with School of Information Science and Engineering, Shandong Normal University, Jinan, China. She received the B.E. degree from Shandong University, Jinan, China, in 2005, and received the PhD degree from Beijing Jiaotong University, Beijing, China, in 2013. During 2010.03~2010.06, she visited the National Kaohsiung University of Applied Sciences, Taiwan. During 2010.09~ 2011.09, she was a Visiting Student at Simon Fraser University, Canada. During 2017.08~ 2018.08, she was a Visiting Scholar at Simon Fraser University, Canada. Her research interests include image/video coding and machine learning.

Hongfei Li received her bachelor degree in software outsourcing from Qilu Normal University, China, in 2017. She is currently pursuing the master degree in the School of Information Science and Engineering from Shandong Normal University. Her research interests include deep learning and image comprkession.

Jia Zhang is with School of Information Science and Engineering, Shandong Normal University, Jinan, China. She received the B.E. and M.E. degrees from the School of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, China, in 2006 and 2009, respectively. She received her Ph.D. degree in communication and information systems in Shandong University in 2013. Her research interests include MIMO radio techniques, joint resource allocation and optimization in multicell cellular networks, dynamic programming, interference coordination in heterogeneous cellular networks, etc.

Yanyan Tan received her Ph.D. degree in intelligent information processing from Xidian University, Xi'an, China, in 2013. She is currently a lecturer at the School of Information Science and Engineering, Shandong Normal University, Jinan, China. Her main research interests are computational intelligence, multi-objective optimization, data analysis, and machine learning.

Yuwei Ren received the M.S. degree from Yanshan University, Qinhuangdao, China, and the Ph.D. degree from the Institute of Automation Chinese Academy of Sciences, Beijing, China, in 2007, and 2011, respectively. She is currently with the Department of Information Science and Engineering, Shandong Normal University, China. Her current research interests include SDC, collaborative FTC.

Huaxiang Zhang is currently a professor with the School of Information Science and Engineering, and the Institute of Data Science and Technology, Shandong Normal University, China. He received his Ph.D. from Shanghai Jiaotong University in 2004, and worked as an associated professor with the Department of Computer Science, Shandong Normal University from 2004 to 2005. He has authored over 160 journal and conference papers and has been granted 9 invention patents. His current research interests include machine learning, pattern recognition, evolutionary computation, cross-media retrieval, web information processing, etc.