

Video Compression Algorithm Based on Directional All Phase Biorthogonal Transform and H.263

Chunxiao Zhang, Chengyou Wang* and Baochen Jiang

School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, China

sdwhzcx@126.com, wangchengyou@sdu.edu.cn, jbc@sdu.edu.cn

Abstract

The vertical or horizontal edges don't dominate in some frames of video sequence, so the conventional discrete cosine transform (DCT) may not be the best choice for those frames. Directional DCT framework behaves better than conventional DCT in coding performance for images where directional edges dominate. In addition, the all phase biorthogonal transform (APBT), which is used in image compression instead of DCT, can also help to improve the performance of compression. In the light of directional DCT and APBT, directional APBT (D-APBT) is proposed and applied to H.263 video coding. Experimental results show that this framework can indeed improve the coding performance remarkably.

Keywords: *Video Coding, Discrete Cosine Transform (DCT), All Phase Biorthogonal Transform (APBT), Directional APBT (D-APBT)*

1. Introduction

As the growth in the living standard, the digital technology plays an important role in people's life today. As we all know, in the raw format, the video data usually needs a large volume of bits to represent. Therefore, some sort of compression has to be done before it is transmitted or stored. Over the past several decades, many video coding methods have been developed for the compression task, such as predictive coding, transform coding and vector quantization. Among various coding techniques, the block-based transform approach has become particularly successful, thanks to its simplicity, excellent energy compaction in the transform domain, super compromise between bit rate and quantization errors. Consequently, many international standards for image and video coding developed so far have adopted some sort of transform, such as JPEG [1], H.263 [2], MPEG-4 [3], H.264 [4], and high efficiency video coding (HEVC) [5]. It's worth mentioning here that HEVC is the next-generation video compression standard, and more than one sort of transform has been applied to the transform stage. The main goals of HEVC design are to improve coding efficiency and increase use of parallel processing architectures [6]. Some recent optimization works have been done for the development of HEVC's parallel processing architectures. For example, Yan et al. have proposed a parallel framework to decouple HEVC motion estimation for different partitions on many-core processors [7] and an efficient parallel framework for HEVC intra-prediction [8]. All these coding standards mentioned above almost cover all applications we can enumerate from our daily life, such as digital cameras, video phones and video-conferencing, storage-based applications, Internet media, digital TV broadcasting and HDTV.

* Corresponding Author

Generally, at low bit rates and poor signal environment, the H.263 video coding standard has been widely used. It is firstly designed for low-bit communication, but not limited to low bit rates. Nowadays, it is used widely in video telecommunication and video conference system. In recent years, there is lots of research related to H.263 conducted. Alparone et al. proposed the theory of vector median filters for motion vector smoothing [9]. To improve the compression efficiency, Ho *et al.*, proposed a modified H.263 encoder which supports real-time content-based scalable video coding [10]. Kholaif et al. presented some modifications for power-saving access point of wireless local area networks in [11]. Then they tried to achieve the transmission of sudden real-time H.263 video over these points. In addition, Goshi *et al.*, applied the unequal loss protection algorithm in motion compensation [12]. Kim et al. proposed an adaptive edge-preserving smoothing and detail enhancement methods for video preprocessing [13].

In H.263 standard, the 2-D discrete cosine transform (DCT), which is applied on individual block of a square size $N \times N$, is adopted. In practice, this conventional 2-D DCT is always implemented separately through two-point DCTs, along the vertical and horizontal directions, respectively. For many occasions, the vertical and horizontal edges dominate in an image. In this case, the conventional DCT seems to be the best choice for these blocks. On the other hand, there may also be other directions in one image block that are perhaps as equally important as the vertical/horizontal direction, *e.g.*, two diagonal directions. Thus, the directional information is not fully used in the conventional DCT. Such information is ever used to perform a postprocessing on coded videos in [14] and [15] to reduce the blocking artifacts. Aiming at the problem proposed above, Zeng et al. proposed an improved transform—directional DCT (D-DCT) [16]. In this improved transform, 7 modes are proposed including the conventional DCT. Its basic idea is to choose the best transform mode from all directional DCTs according to the direction of dominating edges within image. For those images in which vertical or horizontal directions don't dominate, D-DCT performs better. Experimental results show that the coding performance could be improved remarkably and certain coding gain is obtained. Additionally, the conventional DCT has some other unpleasant defects. Once bit rate is changed, more time is wasted in the complex multiplications due to the complex quantization table. Additionally, the image with DCT has serious blocking artifacts at low bit rates. Considering these issues, Hou *et al.*, proposed the all phase digital filtering theory [17]. The key idea of all phase digital filtering is to consider the all possible 2 interception phase of data block in the process of the orthogonal transform. On this basis, the new concept of all phase inverse discrete cosine biorthogonal transform (APIDCBT) was proposed in [18]. Hou et al. deduced the specific forms of all phase biorthogonal transform (APBT) matrices based on inverse DCT transform for subsequent application to still image compression [18]. In APBT-based coding framework, uniform quantization table is adopted instead of complex quantization table, so the amount of computation decreases while the bit rate is changed. What's more, the visual quality is also improved obviously at low bit rates in this framework.

In the light of directional DCT and APBT, directional APBT (D-APBT) was proposed in [19] and applied to grey image compression. Based on the previous work, we try to apply D-APBT to video compression in this paper. The remainder of this paper is organized as follows. In Section 2, we introduce the D-APBT algorithm. Then, the video compression framework based on D-APBT is presented in Section 3. In Section 4, the experimental results and comparisons with D-DCT for video compression framework are given. Finally, conclusions are presented in Section 5.

2. APBT and Directional APBT

2.1. All Phase Biorthogonal Transform (APBT)

On the basis of all phase digital filtering [17], three kinds of all phase biorthogonal transforms based on the Walsh-Hadamard transform (WHT), DCT and IDCT were proposed and the matrices of APBTs were deduced in [18]. For example, the elements of $N \times N$ APIDCBT matrix V are expressed as Eq. (1).

$$V(i, j) = \begin{cases} \frac{1}{N}, & i = 0, j = 0, 1, \dots, N - 1, \\ \frac{N - i + \sqrt{2} - 1}{N^2} \cos \frac{i(2j + 1)\pi}{2N}, & i = 1, 2, \dots, N - 1, j = 0, 1, \dots, N - 1. \end{cases} \quad (1)$$

Similar to the DCT matrix, it can also be used in image compression and transform the image from spatial domain to frequency domain.

2.2. Directional APBT for the Diagonal Down-Left Mode

In H.264, there are in total eight directional prediction modes (the DC mode—Mode 2—is not counted) for the blocks of size 4×4 . Among these modes, one is the vertical prediction (Mode 0), one is the horizontal prediction (Mode 1), and the remaining are named as diagonal down-left (Mode 3), diagonal down-right (Mode 4), vertical-right (Mode 5), horizontal-down (Mode 6), vertical-left (Mode 7), and horizontal-up (Mode 8), respectively. This idea can be readily applied to any block of size $N \times N$ to define the same eight directional modes. For instance, Figure 1 shows six directional modes (Modes 3~8) for $N = 8$.

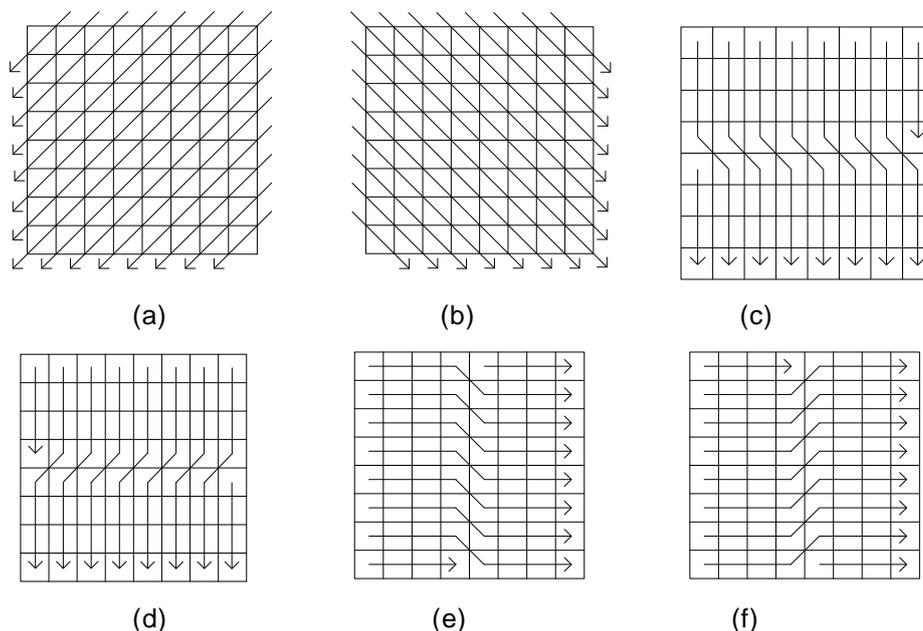


Figure 1. Six Directional Modes Defined in a Similar Way as was Used in Directional DCT for the Block Size 8x8: (a) Mode 3, (b) Mode 4, (c) Mode 5, (d) Mode 6, (e) Mode 7, (f) Mode 8

It is easy to find that Mode 4 can be obtained by flipping Mode 3 either horizontally or vertically; Mode 6 can be obtained by transposing Mode 5, and Mode 7/8 can be obtained by flipping Mode 5/6 either horizontally or vertically. To make our results general

enough, we consider an arbitrary block of size $N \times N$ and will first discuss a truly directional APBT for the diagonal down-left mode, and then discuss the extension to other modes.

As shown in Figure 2, the first 1-D APBT will be performed along the diagonal down-left direction, *i.e.*, for each diagonal line with $i + j = k, k = 0, 1, \dots, 2N - 2$. There are in total $2N - 1$ diagonal down-left APBTs to be done, whose lengths are $[N_k] = [1, 2, \dots, N - 1, N, N - 1, \dots, 2, 1]$. After these APBTs, all the coefficients are expressed into a group of column vectors

$$Y_k = [Y_{0,k}, Y_{1,k}, \dots, Y_{N_k-1,k}]^T, k = 0, 1, \dots, 2N - 2. \quad (2)$$

Notice that each column of Y_k has a different length N_k , with the DC component placed at top, followed by the first AC component and so on; see Figure 3 for $N = 8$.

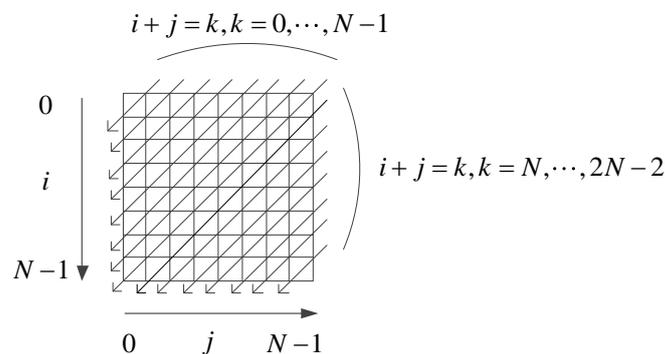


Figure 2. $N \times N$ Image Block in which the First 1-D APBT will be Performed along the Diagonal down-left Ddirection

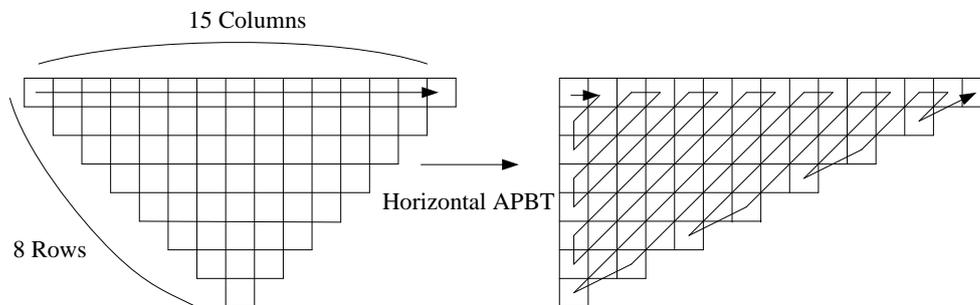


Figure 3. Example of $N = 8$: Arrangement of Coefficients after the First APBT (left) and Arrangement of Coefficients after the Second APBT with the Modified zig-zag scanning (right)

Next, the second 1-D APBT is applied to each row that can be expressed as $[Y_{u,v}]_{v=0:2N-2-2u}$ for $u = 0, 1, \dots, N - 1$. The coefficients after the second APBT are pushed horizontally to the left and denoted as $[Y_{u,v}]_{v=u:2N-2-u}$ for $u = 0, 1, \dots, N - 1$. The right sub-image of Figure 3 shows a modified zig-zag scanning that will be used to convert the 2-D coefficient block into a 1-D sequence so as to facilitate the variable length coding based on run length.

2.3. Extension to Other Directional Modes

Extension to other directional modes is straightforward. For instance, for the diagonal down-right mode (Mode 4), we can simply flip it, and then the flipped image block will fall into the case as discussed above. For Modes 6~8, we can simply flip or transpose the image block first and then the manipulated block will fall into the case of Mode 5. Furthermore, Modes 0 and 1 just are the cases of conventional APBT and thus produce exactly the same result with conventional APBT; we need to consider 7 modes only.

3. Directional APBT-based Video Encoding

The conventional H.263 standard uses DCT in the transform stage, with which the motion-compensated prediction is combined. Moreover, intra-frame and inter-frame coding are defined in this system. For the intra frame, the original image data is directly operated by DCT, quantization, entropy encoding and the inverse procedure to reconstruct the current frame, which is a reference frame to the next frame. However, the inter frame is different from the intra frame, which needs to be operated by the motion-compensated prediction. According to the reference frame, the current frame subtracting the motion-compensated prediction can form the residual image. Then the process to the residual image is the same with intra frame. Finally, combined with the prediction frame by motion compensation, the reconstructed image of the current frame is formed.

In the improved H.263 video standard, D-APBT is used to replace DCT. Then in the quantization process, uniform quantization table is adopted instead of complex quantization table. On the decoder end, the inverse D-APBT and uniform quantization table are also used to replace the conventional process.

In D-APBT, we only consider 7 directional modes as described above. Naturally, the best and most straightforward strategy for coding of an entire image block is to choose the most appropriate directional mode for each block in the image. The simulation adopts the exhaustive traversal method, coding for 7 modes separately, then to choose the best directional mode according to the minimum mean-squared error. In our simulations to be presented in the next section, we use 1 bit to indicate the conventional mode or nonconventional mode, and 3 bits are further used to differentiate Modes 3~8. Because the uniform value is adopted in the quantization process, we can change the bit rate by changing the quantization value. In some sense, our D-APBT is highly compatible with many existing international standards that employ the conventional DCT. Additionally, in the stage of entropy coding, the DC value of each block is coded with variable length instead of the fixed one.

4. Experimental Results

The simulation environment is Visual Studio 2005. To compare the performance of D-APBT and D-DCT algorithm, we apply them to the encoding process of H.263. In the encoding part, luminance and chrominance data are transformed separately. Without loss of generality, we just apply these algorithms to luminance data. The first frames of two video sequences in QCIF format shown in Figure 4 are selected as test images. Figure 5, in which different color represents different modes, shows the directional modes of the test frames using directional DCT.



Figure 4. Video Sequences used in Experimental Simulations: (a) Miss_am, (b) Suzie

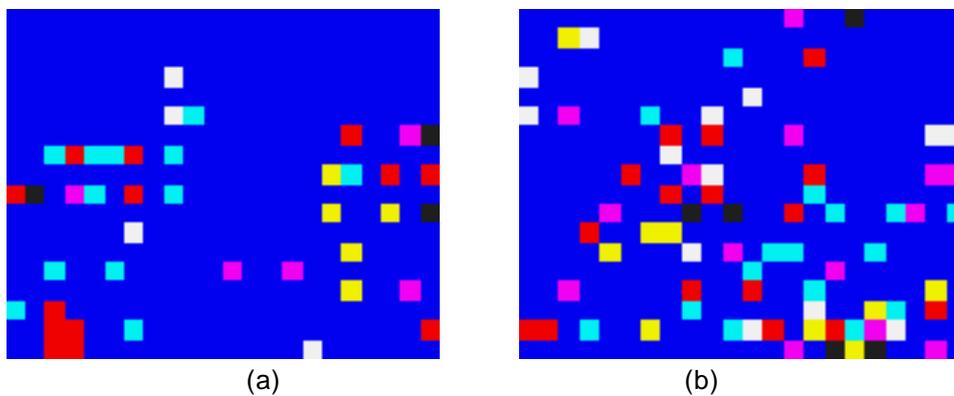


Figure 5. Directional Modes of the Test Frames (QP=25): (a) Miss_am, (b) Suzie

In D-APBT based H.263, we change the bit rate through changing the quantization factor. While in D-DCT based H.263, bit rate varies with different QP value. Table 1 and Table 2 show the experimental results of directional DCT and directional APIDCBT (D-APIDCBT) in H.263. It is clear that the D-APBT based H.263 performs better than that of D-DCT at low bit rates, while at higher bit rates, the D-DCT based H.263 performs better.

Table 1. Comparison of Directional DCT and Directional APIDCBT for H.263 (Miss_am.qcif)

Bit rate/bpp	PSNR/dB	
	D-DCT	D-APIDCBT
0.36	35.93	35.53
0.34	35.30	34.94
0.31	34.56	34.30
0.30	33.90	33.94
0.28	33.23	33.23
0.27	32.78	33.10

As shown in Figure 5, there is a lot of directional information in each frame. Additionally, we know that the APBT performs better than DCT in image coding [18]. Combined with the experimental results shown above, the proposed D-APBT based H.263 video compression framework makes full use of the directional information and the advantage of APBT at low bit rates.

Table 2. Comparison of Directional DCT and Directional APIDCBT for H.263 (Suzie.qcif)

Bit rate/bpp	PSNR/dB	
	D-DCT	D-APIDCBT
0.38	31.75	31.18
0.35	30.94	30.61
0.33	30.58	30.35
0.32	30.33	30.12
0.31	29.88	29.92
0.30	29.74	29.91

5. Conclusion

In this paper, the D-APBT is applied to video coding. Compared with the D-DCT based H.263, the D-APBT based H.263 performs better at low bit rates. Even at higher bit rate, it performs close to D-DCT based H.263. Since uniform quantization is used in the stage of quantization, the computational complexity decreases obviously. Therefore, the D-APBT could be successfully applied in video compression. It not only extends the application range of the APBT theory, but also proves the effectiveness of the directional transform theory.

Although better performance has been achieved in D-APBT based H.263, the procedure of choosing the best mode still needs an amount of calculation. Hence, the further work will be focused on exploring a better method to choose the appropriate mode.

Acknowledgements

This work was supported by the promotive research fund for excellent young and middle-aged scientists of Shandong Province, China (Grant No. BS2013DX022) and the National Natural Science Foundation of China (Grant No. 61201371). The authors would like to thank Qiming Fu, Fanfan Yang, and Liping Wang for their kind help and valuable suggestions. The authors also thank the anonymous reviewers and the editors for their valuable comments to improve the presentation of the paper.

References

- [1] ISO/IEC and ITU-T, "Information Technology -- Digital Compression and Coding of Continuous-tone Still Images -- Part 1: Requirements and Guidelines", ISO/IEC 10918-1: 1994 | ITU-T Rec. T.81, (2011) Sep. 22.
- [2] ITU-T, "H.263: Video Coding for Low Bit Rate Communication", ITU-T Rec. H.263, (2005) Jan. 13.
- [3] ISO/IEC, "Information Technology -- Coding of Audio-Visual Objects -- Part 2: Visual", ISO/IEC 14496-2: 2004, (2014) Dec. 17.
- [4] Joint Video Team of ITU-T and ISO/IEC, "Information Technology -- Coding of Audio-Visual Objects -- Part 10: Advanced Video Coding", ITU-T Rec. H.264 | ISO/IEC 14496-10: 2014, (2014) Aug. 27.
- [5] ISO/IEC, "Information Technology -- High Efficiency Coding and Media Delivery in Heterogeneous Environments -- Part 2: High Efficiency Video Coding", ISO/IEC 23008-2: 2013, (2013) Nov. 25.
- [6] G. J. Sullivan, J.-R. Ohm, W.-J. Han and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, (2012) Dec., pp. 1649-1668.

- [7] C. G. Yan, Y. D. Zhang, J. Z. Xu, F. Dai, J. Zhang, Q. H. Dai and F. Wu, "Efficient parallel framework for HEVC motion estimation on many-core processors", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 12, (2014) Dec., pp. 2077-2089.
- [8] C. G. Yan, Y. D. Zhang, F. Dai, J. Zhang, L. Li and Q. H. Dai, "Efficient parallel HEVC intra-prediction on many-core processor", *Electronics Letters*, vol. 50, no. 11, (2014) May, pp. 805-806.
- [9] L. Alparone, M. Barni, F. Bartolini and L. Santurri, "An improved H.263 video coder relying on weighted median filtering of motion vectors", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 2, (2001) Feb., pp. 235-240.
- [10] W. K.-H. Ho, W.-K. Cheuk and D. P.-K. Lun, "Content-based scalable H.263 video coding for road traffic monitoring", *IEEE Transactions on Multimedia*, vol. 7, no. 4, (2005) Aug., pp. 615-623.
- [11] A. M. Kholiaif, T. D. Todd, P. Koutsakis and A. Lazaris, "Energy efficient H.263 video transmission in power saving wireless LAN infrastructure", *IEEE Transactions on Multimedia*, vol. 12, no. 2, (2010) Feb., pp. 142-153.
- [12] J. Goshi, A. E. Mohr, R. E. Ladner, E. A. Riskin and A. Lippman, "Unequal loss protection for H.263 compressed video", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 3, (2005) Mar, pp. 412-419.
- [13] J.-H. Kim, J. W. Lee, R.-H. Park and M.-H. Park, "Adaptive edge-preserving smoothing and detail enhancement for video preprocessing of H.263", in: *Proceedings of the International Conference on Consumer Electronics, Las Vegas, USA, (2010) Jan. 11-13*, pp. 337-338.
- [14] L. Atzori, F. G. B. De Natale and F. Granelli, "Adaptive anisotropic filtering (AAF) for real-time visual enhancement of MPEG-coded video sequences", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 5, (2002) May, pp. 285-298.
- [15] X. C. Gan, A. W.-C. Liew and H. Yan, "Blocking artifact reduction in compressed images based on edge-adaptive quadrangle meshes", *Journal of Visual Communication and Image Representation*, vol. 14, no. 4, (2003) Dec., pp. 492-507.
- [16] B. Zeng and J. J. Fu, "Directional discrete cosine transforms—A new framework for image coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 3, (2008) Mar., pp. 305-313.
- [17] Z. X. Hou and X. Yang, "The all phase DFT filter", in: *Proceedings of the 10th IEEE Digital Signal Processing Workshop and the 2nd IEEE Signal Processing Education Workshop, Pine Mountain, Georgia, USA, (2002) Oct. 13-16*, pp. 221-226.
- [18] Z. X. Hou, C. Y. Wang and A. P. Yang, "All phase biorthogonal transform and its application in JPEG-like image compression", *Signal Processing: Image Communication*, vol. 24, no. 10, (2009) Nov., pp. 791-802.
- [19] C. Y. Wang, "Directional APBT and its application in image coding", in: *Proceedings of the 11th IEEE International Conference on Signal Processing, Beijing, China, vol. 1, (2012) Oct. 21-25*, pp. 728-731.

Authors



Chunxiao Zhang, He was born in Shandong province, China in 1990. He received his B.S. degree in electronic information science and technology from Shandong University, Weihai, China in 2012 and his M.E. degree in circuits and systems from Shandong University, China in 2015. Now he is with the Huawei Technologies Co. Ltd., Suzhou, China. His research interests include digital image and video processing.



Chengyou Wang, He was born in Shandong province, China in 1979. He received his B.E. degree in electronic information science and technology from Yantai University, China in 2004, and his M.E. and Ph.D. degree in signal and information processing from Tianjin University, China in 2007 and 2010 respectively. Now he is an associate professor in the School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, China. His current research interests include digital image & video processing and analysis, multidimensional signal and information processing.



Baochen Jiang, He was born in Shandong province, China in 1962. He received his B.S. degree in radio electronics from Shandong University, China in 1983, and his M.E. degree in communication and electronic systems from Tsinghua University, China in 1990. Now he is a professor in the School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, China. His current research interests include signal and information processing, image & video processing, and smart grid technology.

