Video Compression Technique

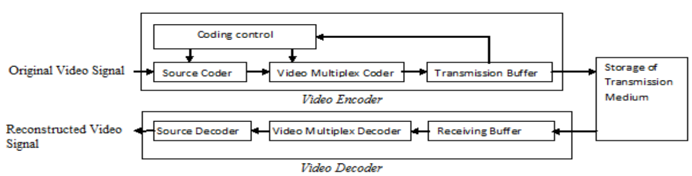
Rajeshwar Dass Member IEEE, Lalit Singh, Sandeep Kaushik

Abstract—Video compression has been the object of intensive research in the last thirty years. Video compression technique is now mature as is proven by the large number of applications that make use of this technology. This paper gives the idea about different techniques available for video compression. H.264/AVC exhibits superior coding performance improvement over its predecessors. The next generation standards are being generated by both VCEG and MPEG. In this paper, we also summarized the progress of those next generation video coding standard projects and existing new video coding techniques.

Index Terms— H.264/AVC, JVT, Video Codec, Issues, KTA, MPEG, VCEG.

1 INTRODUCTION

Video coding techniques provide efficient solutions to represent video data in a more compact and robust way so that the storage and transmission of video can be realized in less cost in terms of size, bandwidth and power consumption. ITU-T and ISO/IEC these are the main two international organizations which decides the standards for video compressions. ISO/IEC MPEG standard includes MPEG-1, MPEG-2, MPEG-4, MPEG-4 Part 10 (AVC), MPEG-7, MPEG-21 and M-JPEG. ITU-I VCEG standard includes H.26x series, H.261, H.263, and H.264. Currently, both VCEG and MPEG are launching their next-generation video coding project. This new generation aims to meet the new requirements future applications may impose on the video coding standard. The entire compression and decompression process requires a codec consisting of a decoder and an encoder. The encoder compresses the data at a target bit rate for transmission or storage while the decoder decompresses the video signals to be viewed by the user. This whole process is shown in fig.1. In general decoding is considerably less complex than encoding. Due to this reason research and implementation efforts are more focused on encoding.



Rest of the paper is organized as: section II describes various basic techniques available for the video compression; section III presents the implementation strategies of video compression techniques; section IV describes various generations &latest trend in video compression; section V describe the different issues related to emerging technologies and discussion of results & conclusion made in section VI.

- Rajeshwar Dass is working as Assistant Professor, Deptartment of Electronics and Communication Engineering ,Deenbandhu Chotu Ram University of Science & Technology Murthal, Sonepat(HR), India.
- Lalit Singh Pursuing Masters in Technology, Deptartment of Electronics and Communication Engineering ,DCRUST, Sonepat(HR), India
- Sandeep Kaushik Pursuing Masters in Technology, Deptartment of Electronics and Communication Engineering ,DCRUST, Sonepat(HR), India

2 BASIC TECHNIQUES

All the video coding standards based on motion prediction and discrete cosine transform produce block artifacts at low data rate. To reduce blocking artifacts a lot of work using post processing techniques has been done [1- 3]. A great deal of work has been done to investigate the use of wavelets in video coding. To reduce blocking artifacts mainly two directions are there:

1. The first one is to code the prediction error of the hybrid scheme using the DWT [4].
2. The second one is to use a full 3-D wavelet decomposition [5-6].

These approaches have reported coding efficiency improvements with respect to the hybrid schemes. These schemes are intended to provide further functionalities such as scalability and progressive transmission. Long-term memory prediction extends motion compensation from the previous frame to several past frames with the result of increased coding efficiency. One of the approaches that reports major improvements using the hybrid approach is the one proposed in [7]. The approach is combined with affine motion compensation. Data rate savings between 20 and 50% are achieved using the test model of H.263+. The corresponding gains in PSNR are between 3 and 0.8 dB. MPEG-4 and H.263+ represent the state of the art in video coding. H.263+ provides a framework for doing frame-based low to moderate data rate compression. It (MPEG-4) combines frame-based and segmentation-based approaches along with the mixing of natural and synthetic content allowing efficient coding as well as content access and manipulation. No doubt, there is that other schemes may improve the coding efficiency established in MPEG-4 and H.263+ but no significant breakthrough has been presented to date. During the last ten years, the hybrid scheme combining motion compensated prediction and DCT has represented the state of the art in video coding. This hybrid scheme approach is used by the ITU H.263 and H.261 standards as well as for the MPEG-2 and MPEG-1 standards. In 1993, the need to add new content-based functionalities and to provide the user the possibility to manipulate the audio-visual content was recognized and a new standard effort known as MPEG-4 was launched. Functionalities in addition to these, MPEG-4 provides also the possibility of combining natural and synthetic content. MPEG-4 phase 1 became an international standard in 1999 [8]. MPEG-4 is having difficulties finding wide-spread use, main reason is due to the protection of intellectual property and the need to develop automatic and efficient segmentation schemes. Mainly the frame-based part of MPEG-4 which incorporates error resilience tools, in finding its way in the mobile communications and Internet streaming. H.263, and several variants of it [9], are also very much used in mobile communication and streaming and it will be interesting to see how these two standards compete in these applications. The natural video part of MPEG-4 is also based in motion compensation prediction followed by the discreet cosine transform; the difference here made is coding of object shapes. The use of the most efficient coding techniques, due to its powerful object-based approach and the large variety of data types that it incorporates, the MPEG-4 represents today the state-of-the-art in terms of visual data coding technology [10]. How MPEG-4 is deployed and what applications will make use of its many functionality is still an open question.

3 IMPLEMENTATION STRATEGIES

There are a number of techniques to implement the video compression and these techniques can be broadly divided into two categories: hardware-based implementation and software-based implementation.

3.1 Hardware-Based Approach

The most common approach is to design a dedicated VLSI circuit for video compression [11]. One can have function specific hardware, such as associated inverse operations, DCT, VLC and block matching. Due to the exploitations of the data flow of the algorithm and special control, the processing capability of these approaches can be increased tenfold compared to those of conventional microprocessors [12]. However, function specific approaches provide limited flexibility and cannot be modified for further developments. In addition, the architecture design usually requires a regular control paradigm and data flow of the algorithms that may not be useful for solving the complexities of circuit designs. Furthermore, the complexity limitations of the circuit design, such as processing speed, throughput, the number of translators, silicon area, also restrict its implementation potential for growing multimedia applications. A more cost-effective alternative is provided by programmable processors, such as programmable DSP or VSP[13]. Such an approach can execute different tasks under software control, it can avoid cost intensive hardware redesign. Programmable processors are flexible in the way it allows the implementation of various video compression algorithms without the need for a hardware redesign. Moreover, multiple algorithms can be executed on the same hardware and their performance can be optimized as well [14]. Consequently, their implementation time and cost increase accordingly. Furthermore, they also incur significant costs in software development and system integration. Usually, programmable processors require silicon area for program storage and control unit, and dissipate more power than dedicated VLSI solutions. Fig.2 shows various implementation approaches and hardware–software relationship.

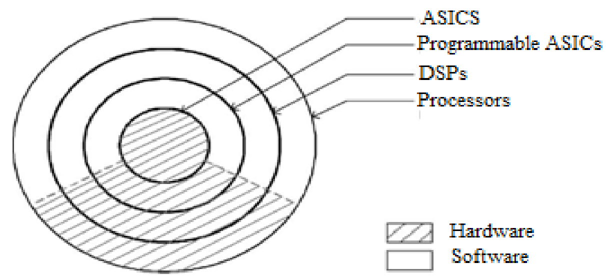


Fig 2. Various Implementation Approaches And Hardware–Software Relationship

3.2 Software-Based Approach

Software-based approaches are becoming more popular because the performance of general-purpose processors has been increasing rapidly. Further, more and more emerging multimedia standards emphasize high-level interactivity, extensibility, and flexibility, posing significant opportunities for software-based solutions Furthermore, the rapid evolution of multimedia techniques has dramatically shortened the required time for market making it very difficult to come up with a new hardware design for each updated technique. The inherent modular nature of various video compression algorithms allows experimenting and hence improving various parts of the encoder independently, including ME, DCT algorithm and rate-controlled coding. The major advantage of using the software-based approach is that it allows incorporating new research ideas and algorithms in the encoding process for achieving a better picture quality at a reduce bit rate for a desired level of picture quality, or on given bit rate. The software-based approach is also flexible in that it allows tuning of various parameters for multiple passes for optimization. Some more benefits of software-based approach are flexibility to adapt to the continuing changes in multimedia applications and portability. Encoding is more challenging due to enormous amount of computation required, whereas decoding can be done easily in software. Real-time performance for high-quality profiles is still quite difficult, but encoding for simple video profiles of various standards can now be done on a single processor. To speed up the compression a natural alternative is to be utilize the accumulated processing capability of parallel processing [15]. However, parallelism can be exploited in different ways, there is no unique philosophy for the best solution, ranging from simultaneous instructions execution within massively parallel processors (MPPs) and a single processor, to distributed networks. It is important to recognize that parallel processing alone may not be enough in software-based implementation, this includes efficient algorithms for DCT, a fast ME and other parts of the encoder [16]. In addition, low-level programming primitives that take advantage of the machine architecture must be harnessed to accelerate the computation [17][18]. Finally, several issues should be addressed in software-based parallel processing such as I/O [19], memory access [20][21], and achieving better rate control [22-25].

4 GENERATION OF VIDEO COMPRESSION TECHNIQUES

4.1 Defining Generations of Video Compression Techniques.

It is of the utmost importance to understand the sequence of image n video coding development expressed on the basses of "generation based" coding approaches. Table 1 shows this classification according to [45]. It can be seen from this classification that the coding community has reached third generation video coding techniques.

Table 1
Generations Of Video Compression Techniques[45]

Coding Generations	Approach	Technique
0^{th} Generation	Direct waveform coding	PCM
1^{st} Generation	Redundancy removal	DPCM,DCT,DWT, VQ
2^{nd} Generation	Coding by structure	Image segmentation
3^{rd} Generation	Analysis and Synthesis	Model-based coding
4^{th} Generation	Recognition and reconstru-ction	Knowledge based coding
5^{th} Generation	Intelligent coding	Semantic coding

4.2 Existing New Technology

The advances of video coding techniques were contributed by various groups and organizations. Hence to provide a software platform to collect and evaluate these new techniques, a Key Technical Area (KTA) [26] platform was developed based on JM11 reference software, where the new coding tools are added very frequently. The major new coding tools added to KTA platform can be summarized as follows:

1 ***Intra Prediction:*** In [27], H.264 intra prediction is enhanced with additional Bi-directional Intra Prediction (BIP) modes, where BIP combines prediction blocks from two prediction modes using a weighting matrix. Furthermore, Mode-Dependent Directional Transform (MDDT) using transforms derived from KLT is applied to capture the remaining energy in the residual block.

2 ***Inter Prediction:*** To further improve inter prediction efficiency, finer fractional motion prediction and better motion vector predictions were proposed. Increasing the resolution of the displacement vector from 1/4-pel to 1/8-pel to obtain higher efficiency of the motion compensated prediction is suggested in [28]. A competing framework for better motion vector coding is proposed which include SKIP mode, in which both spatial and temporal redundancies in motion vector fields are captured [29]. Moreover, [30] suggests extending the macroblock size up to 64x64 so that new partition sizes 64x64, 64x32, 32x64, 32x32, 32x16, and 16x32 can be used. Instead of using the fixed interpolation filter from H.264/AVC, Adaptive Interpolation Filters (AIF) are proposed, such as 2D AIF [31], Separable AIF [32], Directional AIF [33], Enhanced AIF [34], and Enhanced Directional AIF [35].

3 ***Quantization:*** To achieve better quantization, optimized quantization decision at the macroblock level and at different coefficient positions are proposed. Rate Distortion Optimized Quantization (RDOQ), was added to the JM reference software, it performs optimal quantization on a macroblock. It does not require a change of H.264/AVC decoder syntax. More recently, [36] gives an improved, more efficient RDOQ implementation. In [37], Adaptive Quantization Matrix Selection (AQMS), where different quantization steps can be have by different coefficient positions, a method deciding the best quantization matrix index; is proposed to optimize the quantization matrix at a macroblock level.

4 ***Transform:*** For motion partitions bigger than 16x16, a 16x16 transform is suggested in addition to 4x4 and 8x8 transforms [38]. Moreover, transform coding is not always a must. Either spatial domain coding can be adaptively chosen or standardized transform coding to be chosen; this is proposed for each block of the prediction error [16].

5 ***In-loop Filter:*** In KTA, besides the deblocking filter, an additional Adaptive Loop Filter (ALF) is added to improve coding efficiency by applying filters to the deblocked-filtered picture. Adaptive Loop Filter adopted Two different techniques so far: one is Quad-tree based Adaptive Loop Filter (QALF) [39] and second is Block-based Adaptive Loop Filter (BALF) [40].

6 ***Internal bit-depth increase:*** By using 12 bits of internal bit depth for 8-bit sources, so that the internal bit-depth is greater than the external bit-depth of the video codec, the coding efficiency can be further improved [41]. There are many contributions not added to KTA yet. For example, [42-44] proposed three methods, respectively, to use Decoder Side Motion Estimation (DSME) for B-picture motion vector decision, which improves coding efficiency by saving bits on B-picture motion vector coding. Also, some new techniques are under investigation and will be presented in the responses for call for proposals.

4.3 H.264/AVC

The current H.264/AVC compression standard is based on the picture-wise processing and waveform-based coding of video signals. The technology now being considered for the new standardization project on high-efficiency video coding (HEVC) is a generalization of this approach which promises significant gains through innovations such as improved intra-prediction, larger block sizes, more flexible ways of decomposing blocks for inter- and intra-coding and better exploitation of long-term correlations and picture dependencies. It will support a wide range of encoder modes, which are typically optimized using mean-squared-error-based or related distortion measures.

5 ISSUE ON EMERGING TECHNOLOGIES

The predicted growth in demand for bandwidth, driven largely by video applications, is probably greater now than it has ever been. There are four primary drivers for this:

1. Recently introduced formats such as 3-D and multiview, coupled with pressures for increased dynamic range, spatial resolution and framerate, all require increased bit-rate to deliver improved levels of immersion or interactivity.

2. Video-based web traffic continues to grow and dominate the internet through social networking and catch up TV. In recent years, Youtube has accounted for 27% of all video traffic and, by 2015, it is predicted that there will be 700 billion minutes of video downloaded. That represents a full-length movie for every person on the planet.

3. User expectations continue to drive flexibility and quality, with a move from linear to nonlinear delivery. Users are demanding My-Time rather than Prime-Time viewing.

4. Finally new services, in particular mobile delivery through 4G/LTE to smart phones. Some mobile network operators are predicting the demand for bandwidth to double every year for the next 10 years!

6 CONCLUSION

In this paper we have concluded the basic different techniques available for video compression and the latest technique (H.264/AVC) available for video compression is also included. We have seen here that H.264/AVC has been developed by both the ISO/IEC (MPEG) and ITU-T (VCEG) organizations. It has various improvements in terms of coding efficiency, like flexibility, robustness and application domains. No doubt as per the requirements and applications, there will be always new development in video compression technique. From the review of various video compression papers it infers that there are still lots of possibilities for the improvement of video compression technique.

7 REFRENCES

[1] K. K. Pong and T. K. Kan, “Optimum loop filter in hybrid coders,” *IEEE Transactions on Circuits and Systems in Video Technology*, vol. 4, no.2, pp. 158-167, 1997.

[2] T. O’Rourke and R. L. Stevenson, “Improved image decompression for reduced transform coding artifacts,” *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 5, no. 6, pp. 490-499, December 1995.

[3] R. Llados-Bernaus, M. A. Robertson and R. L. Stevenson, “A stochastic technique for the removal of artifacts in compressed images and video,” in *Recovery Techniques for Image and Video Compression and Transmission*, Kluwer, 1998.

[4] K. Shen and E. J. Delp, “Wavelet based rate scalable video compression,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 1, February 1999, pp. 109-122.

[5] C. I. Podilchuk, N. S. Jayant, and N. Farvardin, “Three-dimensional subband coding of video,” *IEEE Transactions on Image Processing*, vol. 4, no. 2, pp. 125-139, February 1995.

[6] D. Taubman and A. Zakhor, “Multirate 3-D subband coding of video,” *IEEE Transactions on Image Processing,* vol. 3, no. 5, pp. 572-588, September 1994.

[7] T. Wiegand, E. Steinbach, and B. Girod, “Long term memory prediction using affine compensation,” *Proceedings of the IEEE International Conference on Image Processing*, Kobe, Japan, October 1999.

[8] ISO/IEC ISO/IEC 14496-2: 1999: Information technology – Coding of audio visual objects – Part 2: Visual, December 1999.

[9] G. Côté, B. Erol, M. Gallant, and F. Kossentini, "H.263+: Video coding at low bit rates," *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 8, no. 7, November 1998.

[10] F. Pereira, “Visual data representation: recent achievements and future developments,” *Proceedings of the European Signal Processing Conference (EUSIPCO)*, Tampere, Finland, September 5-8, 2000.

[11] D.J. Mlynek, J. Kowalczuk, VLSI for digital television, Proceedings of the IEEE 83 (7) (1995) 1055–1070.

[12] P. Pirsch et al., VLSI architectures for video compression––a survey, Proceedings of the IEEE 83 (2) (1995) 1055–1070.

[13] G. Frantz, Digital signal processor trends, IEEE Micro (2000) 52–59.

[14] I. Kuroda, T. Nishitani, Multimedia processors, Proceedings of the IEEE 866 (1998) 1203–1221.

[15] K. Shen, G.W. Cook, L.H. Jamieson, E.J. Delp, An overview of parallel processing approaches to image compression, in: Proceedings of the SPIE Conference on Image and Video Compression, vol. 2186, San Jose, California, February 1994 pp. 197–208.

[16] S.M. Akramullah, I. Ahmad, M.L. Liou, A software based H.263 video encoder using network of workstations, in: Proceedings of SPIE, vol. 3166, 1997.

[17] A. Peleg, U. Weiser, MMX technology extension to the intel architecture, IEEE Micro 16 (4) (1996) 42–50.

[18] M. Tremblay, VIS speeds new media processing, IEEE Micro 164 (1996) 10–20.

[19] I. Ahmad, S.M. Akramullah, M.L. Liou, M. Kafeel, A scalable off-line MPEG-2 encoder using a multiprocessor machine, Parallel Computing, in press.

[20] I. Tamitani, A. Yoshida, Y. Ooi, T. Nishitani, A SIMD DSP for real-time MPEG video encoding and decoding, in: Workshop on VLSI Signal Processing, 1992, pp. 119–128.

[21] J. Tanskanen, J. Niittylahti, Parallel memories in video encoding, in: Data Compression Conference, 1999, p. 552.

[22] S. Bozoki, R.L. Lagendijk, Scene adaptive rate control in a distributed parallel MPEG video encoder, in: International Conference on Image Processing, vol. 2, 1997, pp. 780–783.

[23] K. Nakamura, M. Ikeda, T. Yoshitome, T. Ogura, Global rate control scheme for MPEG-2 HDTV parallel encoding system, in: International Conference on Information Technology: Coding and Computing, 2000, pp. 195–200.

[24] S. Nogaki, A study on rate control method for MP@HL encoder with parallel encoder architecture using picture partitioning, in: 1999 International Conference on Image Processing, vol. 4, 1999, pp. 261–265.

[25] P. Tiwari, E. Viscito, A parallel MPEG-2 video encoder with look-ahead rate control, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, 1996, pp. 1994–1997.

[26] http://iphome.hhi.de/suehring/tml/download/KTA .

[27] Y. ye and M. Karczewicz, “Improved Intra Coding,” *ITU-T Q.6/SG16 VCEG, VCEG-AG11*, Shenzhen, China, 2007.

[28] J. Ostermann and M. Narroschke, “Motion compensated prediction with 1/8-pel displacement vector resolution,” *ITU-T Q.6/SG16 VCEG, VCEG-AD09*, Hangzhou, China, 2006.

[29] G. Laroche, J. Jung, and B. Pesquet-Popescu, “RD Optimized Coding for Motion Vector Predictor Selection,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, 2008, pp. 1681-1691.

[30] P. Chen, Y. Ye, and M. Karczewicz, “Video Coding Using Extended Block Sizes,” *ITU-T SG16/Q6, doc. C-123*, 2009.

[31] Y. Vatis and J. Ostermann, “Prediction of P- and B-Frames Using a Two-dimensional Non-separable Adaptive Wiener Interpolation Filter for H.264/AVC,” *ITU-T Q.6/SG16 VCEG, VCEG-AD08*, Hangzhou, China, 2006.

[32] S. Wittmann and T. Wedi, “Simulation results with separable adaptive interpolation filter,” *ITU-T Q.6/SG16 VCEG, VCEG-AG10*, Shenzhen, China, 2007.

[33] D. Rusanovskyy, K. Ugur, and J. Lainema, “Adaptive Interpolation with Directional Filters,” *ITU-T Q.6/SG16 VCEG,AG21*, Shenzhen, China, 2007.

[34] Y. Ye and M. Karczewicz, “Enhanced Adaptive Interpolation Filter,” *ITU-T SG16/Q.6 Doc. T05-SG16-C-0464*, Geneva, Switzerland, 2008.

[35] T. Arild Fuldseth, D. Rusanovskyy, K. Ugur, and J. Lainema, “Low Complexity Directional Interpolation Filter,” *ITU-T Q.6/SG16 VCEG, VCEG-AI12*, Berlin, Germany, 2008.

[36] M. Karczewicz, Y. Ye, and I. Chong, “Rate Distortion Optimized Quantization,” *ITU-T Q.6/SG16 VCEG, VCEG-AH21*, Antalya, Turkey, 2008.

[37] A. Tanizawa and T. Chujoh, “Adaptive Quantization Matrix Selection,” *ITU-T Q.6/SG16 VCEG, D.266*, 2006.

[38] M. Narroschke and H.G. Musmann, “Adaptive prediction error coding in spatial and frequency domain for H.264/AVC,” *ITU-T Q.6/SG16 VCEG, VCEG-AB06*, Bangkok, Thailand, 2006.

[39] T. Chujoh, N. Wada, and G. Yasuda, “Quadtree-based adaptive loop filter,” *ITU-T Q.6/SG16 VCEG,C.181*, 2009.

[40] G. Yasuda, N. Wada, T. Watanabe, and T. Yamakage, “Block-based Adaptive Loop Filter,” *ITU-T Q.6/SG16 VCEG, VCEG-AI18*, Berlin, Germany, 2008.

[41] T. Chujoh and R. Noda, “Internal bit depth increase for coding efficiency,” *ITU-T Q.6/SG16 VCEG, VCEG-AE13*, Marrakech, Morocco, 2007.

[42] S. Klomp and J. Ostermann, “Response to Call for Evidence in HVC: Decoder-side Motion Estimation for Improved Prediction,” *ISO/IEC JTC 1/SC 29/WG 11 M16570*, London, UK, 2009.

[43] T. Murakami, “Advanced B Skip Mode with Decoder-side Motion Estimation,” *ITU-T Q.6/SG16 VCEG, VCEG-AK12*, Yokohama, Japan, 2009.

[44] S. Kamp, M. Evertz, and M. Wien, “Decoder Side Motion Vector Derivation,” *ITU-T Q.6/SG16 VCEG, VCEG-AG16*, Shenzhen, China, 2007.

[45] H. Harashima, K.Aizawa, and T. Saito, “Modelbased analysis synthesis coding of videotelephone images – conception and basic study of intelligent image coding,” *Transactions IEICE*, vol. E72, no. 5, pp. 452-458, 1989.

Rajeshwar Dass received M.E Degree in Electronics & Communication Engineering from National Institute of Technical Teachers Training and Research Center (NITTTR) Chandigarh in the year 2007 and B.E. (Honors) degree in Electronics & Communication Engineering from Apeejay College of Engineering Sohna Gurgaon (Affiliated to M.D.U Rohtak) in 2004. He is pursuing his Ph.D from DCRUST Murthal. In August 2004 he joined Department of Electronics & Communication Engineering of Apeejay College of Engineering Sohna Gurgaon. He joined Department of Electronics & Communication Engineering of Deenbandhu Chhotu Ram University of Science and Technology (D.C.R.U.S.T) Murthal Sonepat(India) as Assistant professor in October 2008. His interest includes Medical Image Processing, Neural Network, Wireless Communication and Soft Computing Techniques. He has contributed more than twenty technical papers in International Journals and Conferences. He has written a book on wireless communication. He is the member of IEEE and reviewer of IJSTR, IJECT and WASET.

Lalit Singh received the bachelor's degree in Electronics and Communications Engineering with honors from the Kurukshetra University, Kurukshetra, in 2011. He is currently pursuing Masters of Technology in Electronics and Communications from Deenbandhu Chhotu Ram University of Science & Technology, Murthal. His interest includes Image Processing and Signal Processing.

Sandeep Kaushik received the bachelor's degree in Electronics and Communications Engineering with honors from the Kurukshetra University, Kurukshetra, in 2011. He is currently pursuing masters of Technology in Electronics and Communications from Deenbandhu Chhotu Ram University of Science & Technology, Murthal.