An Adaptive Cross-layer Mapping Algorithm for Multiview Video Coding over IEEE 802.11e WLANs

Bhavsingh Maloth, Ashok Sarabu ,B.Vinod Kumar

Dept of Information Technology, G.Pullaiah College of Engineering & Technology
Kurnool, Andhra Pradesh,India
bhavsinghm@gmail.com
sarabu.ashok@gmail.com
vinodkumar.bangaru674@gmail.com

***Abstract*-This paper proposes an Adaptive Cross-layer Mapping algorithm for Multi view Video Coding (ACM-MVC) over IEEE 802.11e Wireless Local Area Networks (WLANs). To support the varying Quality-of-Service (QoS) requirements of emerging applications, four access categories (ACs) have been defined in the IEEE 802.11e standard. Each AC has a different transmission priority, which implies the higher the transmission priority, the better the opportunity to transmit. In order to improve the transmission of video data over IP networks, the Multiview Video Coding (MVC) extension of the H.264/MPEG 4 AVC standard has been introduced. The MVC adopts the layered video scheme. The proposed mapping algorithm utilizes a cross-layer approach and dynamically assigns packets of different layers to one of the four ACs. The experiment results demonstrate that the proposed mapping algorithm achieves better QoS than the conventional methods such as the static mapping algorithm.**

***Keywords*: Multiview Video Coding, cross-layer mapping, wireless network, QoS, Adaptive mapping, static mapping.**

1. INTRODUCTION:

The IEEE 802.11e standard for Wireless Local Area Networks (WLANs) and the **Multiview Video Coding** (**MVC**) is an amendment to H.264/MPEG-4 AVC video compression standard developed with joint efforts by MPEG/VCEG that enables efficient encoding of sequences captured simultaneously from multiple cameras using a single video stream have been proposed to enhance multimedia services, such as Video-on-Demand (VoD), IPTV and video conferencing in wireless network environments. In the IEEE 802.11e standard, four access categories (ACs) have been introduced to support the varying Quality of Service (QoS) requirements. Each AC has different transmission priority; the higher the transmission priority, the better the opportunity to transmit.

Multiview Video Coding (MVC, ISO/IEC 14496-10:2008 Amendment 1) is an extension of the Advanced Video Coding (AVC) standard that provides efficient coding of such multiview video. The overall structure of MVC defining the interfaces is illustrated in the figure 1. The encoder receives *N* temporally synchronized video streams and generates one bitstream. The decoder receives the bitstream, decodes and outputs the *N* video signals.

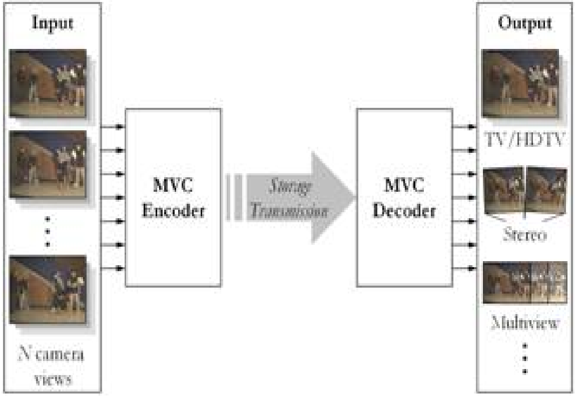

Fig 1.Multiview Video Coding (MVC)

Multiview video contains a large amount of interview statistical dependencies, since all cameras capture the same scene from different viewpoints. Therefore, combined temporal and inter-view prediction is the key for efficient MVC. As illustrated in the figure 2 a picture of a certain camera can not only be predicted from temporally related pictures of the same camera. Also pictures of neighboring cameras can be used for efficient prediction.

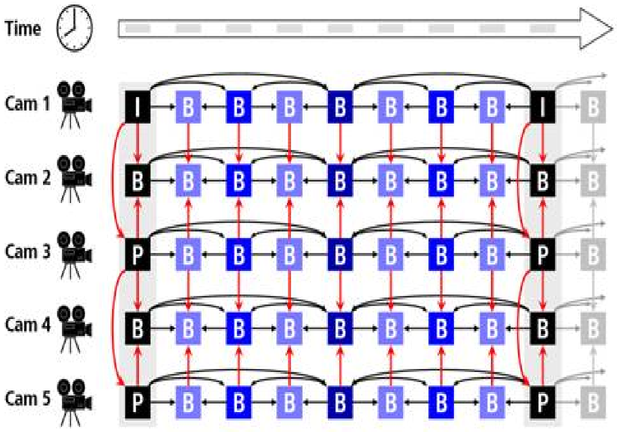

Fig 2. Temporal/inter-view prediction structure for MVC.

2. BACKGROUND:

The scalabilities of SVC have special benefits in wireless networks, where video streaming services are highly required to adapt the varying network conditions and diverse client devices with different capabilities, such as processing power, memory capacity and access bandwidth. The encoded SVC video consists of a base layer and one or more enhancement layers. Different frame rates, spatial resolutions and fidelities of video streams can be supported by selective extraction of enhancement layers. Recently, several studies have been conducted to improve the video streaming over the IEEE 802.11e network. According to the priority of the video data, the video packets are classified. Each class of video packets will be assigned to a predefined AC. However, as the network condition changes constantly, the static mapping may not work well in the dynamic environment. Proposed a MVC adaptive mapping algorithm, which delivers each packet to different ACs depending on network conditions. Even though the algorithm in deals with the adaptive mapping algorithm, it focuses on the MPEG-4 streaming, but it does not consider the layered video format such as MVC.

3.ADAPTIVE CROSS-LAYER MAPPING ALGORITHM FOR MULTIVIEW VIDEO CODING

The multi-view video coding (MVC) is currently being developed as an extension of the ITU-T Recommendation H.264 | ISO/IEC International Standard ISO/IEC 14496-10 advanced video. In addition to implementing the normative decoding process specified in the text of the Joint Draft (JD) for MVC, reference software is provided to demonstrate effectiveness of non-normative encoding techniques to be used with the standard .Figure 1: Inter-view-temporal prediction structure using hierarchical B pictures

As a matter of procedure, new tools, syntax or processes are firstly documented in the JMVM. Further review and software verification are conducted before the technology is adapted to the MVC standard. In observance of that procedure, decoder specifications are considered for the future version of the joint draft (JD).

Figure 3 shows the current reference structure for MVC. This scheme uses a prediction structure with hierarchical B pictures for each view. Additionally, inter-view prediction is applied to every second view: S1, S3 and S5 in Fig. 3. When the total number of views is even, the prediction structure of the last view (S7 in Fig. 3) is similar to those of even views. While B pictures in the even views do not use any inter-view references, B pictures in the last view use one inter-view reference. To allow random access, we start each GOP (S0/T0, S0/T8) with the I-frame.

Figure 4 also depicts that if the total length of the sequence does not fit an integer multiple of the GOP-length, a shortened tail GOP can be realized at the end of the sequence. In Fig. 1, the GOP-length is 8; however, we use GOP-lengths of 12 or 15 for the current test conditions.

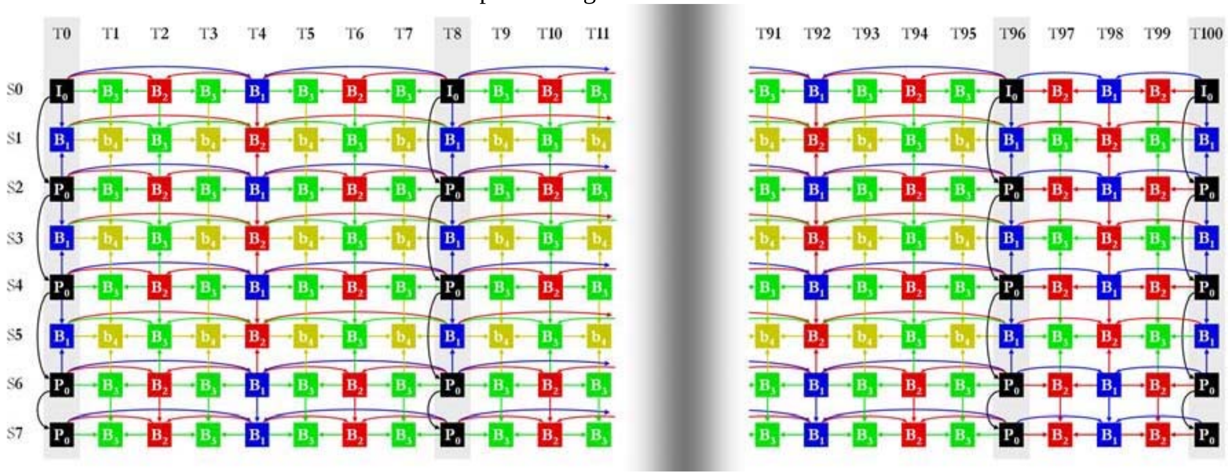

Fig 3 Inter-view-temporal prediction structure using hierarchical B pictures

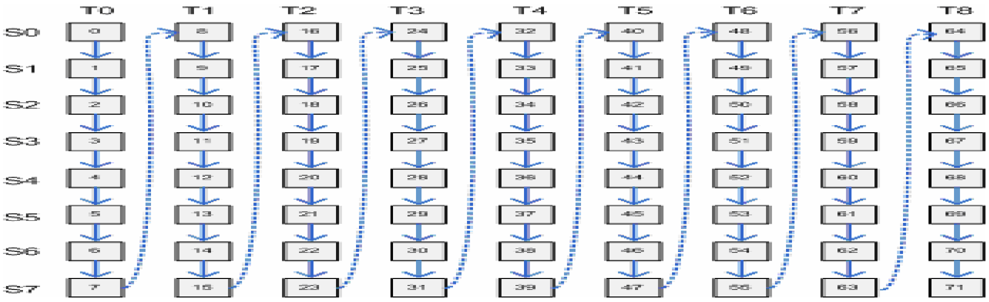

Figure 4: Coding order of multi-view videos

To guarantee the quality of delivered video the proposed mapping algorithm dynamically allocates the video to the most appropriated AC at the MAC layer according to both the significance of video type and the network traffic load. For MPEG-4 video stream, the loss of more important video frames would deteriorate the delivered video quality. For example, one I frame loss will cause all frames in the same GOP to be undecodable; at the same time, one B frame loss just affects itself [1]. Based on the significance of video frame, the channel access priorities used to prioritize the transmission opportunity at the MAC layer are set with the I frame as the highest; the P frame below I but above B’s priority, and the B frame set at the lowest priority. To allocate important video data into higher priority AC queue in 802.11e MAC layer as far as possible, we give different mapping probabilities, defined as ***Prob_TYPE***, to different video frame types according to its coding significance. If allocating a frame into a lower priority queue is inevitable, the transmission allocating probability of lower significant frames is higher than that of important video frames. Less important video frame types will be assigned larger ***Prob_TYPE***. As a result, for the MPEG-4 codec the downward mapping probability relationship of these three video frame types is ***Prob_B > Prob_P > Prob_I***, and all of the probabilities are between 0 and 1.

Furthermore, to support dynamic adaptation to changes in network traffic loads, we use the MAC queue length as an indication of the current network traffic load. According to the IEEE 802.11e specification, when transmitted over an IEEE 802.11e wireless network, MPEG-4 video packets are placed in AC2 category which has better opportunity to access the channel than lower priority ACs. The tradeoff is, when the video stream increases, this queue rapidly jams and drops occur. For this reason, the proposed mapping algorithm re-arrange most recently received video packets into other available lower priority queues, while the AC2 queue is getting filled. We adopted two parameters, ***threshold low*** and ***threshold high***, to predicatively avoid the upcoming congestion by performing queue management in advance. This congestion control idea is derived from the principal of Ran- dom Early Detection (RED) mechanism [2] (used to predict the upcoming congestion) which actively drops packets if needed, and achieves efficient congestion avoidance and an optimal queue management. The integrated function to introduce these two parameters in the algorithm is in the following expression:

$$Prob\_New = Prob\_TYPE * \frac{qlen(AC[2]) - threshold\_low}{threshold\_high - threshold\_low} \qquad (1)$$

In this function, the original predefined downward mapping probability of each type of video frame, ***Prob_TYPE***, will be adjusted according to the current queue length and threshold values, and about the result is a new downward mapping probability, ***Prob_New***. The higher ***Prob_New***, the greater the opportunity for the packet to be mapped into a lower priority queue. Table 2 lists the notations used in the proposed adaptive cross-layer mapping algorithm.

Table 2: Parameter notations in our proposed adaptive mapping algorithm

Prob_TYPE	Downward mapping probability of each type video packet e.g. Prob_I, Prob_P, Prob_B
Prob_New	New calculated downward mapping probability
threshold_low	The lower threshold of queue length
threshold_high	The upper threshold of queue length
qlen(AC[2])	The queue length of access category 2

3.1 The proposed adaptive cross-layer mapping algorithm

```
When a video data frame arrives:
if(qlen(AC[2]) < threshold_low)
    video packet → AC[2];
else if(qlen(AC[2]) < threshold_high) {
    $Prob\_New = Prob\_TYPE * \frac{qlen(AC[2]) - threshold\_low}{threshold\_high - threshold\_low}$
    RN = a random number generated from Uniform function (0.0, 1.0);
    if(RN > Prob_New)
        video frame → AC[2];
    else
        video frame → AC[1];
}
else if(qlen(AC[2]) > threshold_high){
    if(RN > Prob_TYPE){
        video frame → AC[1];
    else
        video frame → AC[0];
}
```

In the mapping algorithm shown, when a video packet arrives, first the queue length of AC2 (***qlen(AC[2])***) is checked and compared against a set of threshold values, ***threshold_high*** and ***threshold_low***. If the queue length is lower than the lower threshold value, ***threshold_low*** (light load), the video data is mapped to AC[2] (no matter what type of video data is being transferred). But if the queue length is greater than the upper threshold value, ***threshold_high*** (heavy video traffic load) the video data is directly mapped to lower priority queues, AC[1] or AC[0]. However, while the queue length of AC[2] falls within ***threshold_high*** and ***threshold_low,*** the mapping decision is determined based on both the mapping probability (***prob_TYPE***) and the current buffering size condition of the queue as given by formula (1). Hence, the video data packet will be mapped to AC[2], AC[1] or AC[0] according to the calculated downward mapping probability. By exploiting such a priority scheme and queue length management strategy, the transmissions prioritized and the drop rate of video minimized, along with efficient utilization of network resources.

4. RESULTS AND DISCUSSIONS

4.1Experiment environment and setting

To evaluate the performance of our proposed cross-layer mapping algorithm, we have conducted simulations using a widely adopted network simulator NS-2 [3], and integrated with Evalvid [4]. The results of the proposed mapping al- gorithm are compared with the results derived from IEEE 802.11e EDCA and the static mapping algorithm in [5].

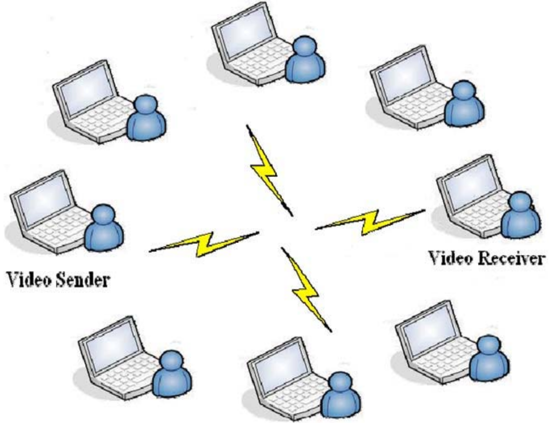

Fig. 5 Network topology used in our simulation tests

The video sources used in the simulation are YUV CIF (352 × 288) [6], Football, and YUV QCIF (176 × 144) [7], Foreman. Football is a shorter (90 frames) and in motion type video, while Foreman is a longer (400 frames) and static type video. Each video frame was fragmented into packets before transmission, and the maximum transmission packet size over the simulated network is 1000 bytes. Table 3 shows the number of video frames and packets of the video sources.

Table 3 The amounts of video frames and packets of the video sources

Video	Format	Frame number			Total	Packet number			Total
		I	*P*	*B*		*I*	*P*	*B*	
Football	CIF	10	20	60	90	135	125	220	480
Foreman	QCIF	45	89	266	400	237	149	273	659

Table 4 Number of traffic streams for each case in scenario 2

	Audio (AC[3])	TCP (AC[1])	UDP (AC[0])
Case 1	1	1	1
Case 2	2	2	2
Case 3	3	3	3
Case 4	5	5	5
Case 5	10	10	10

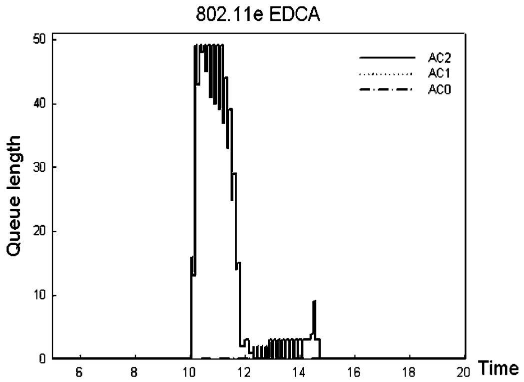

Fig.6(a) Queue space utilization of 802.11e EDCA(Football).

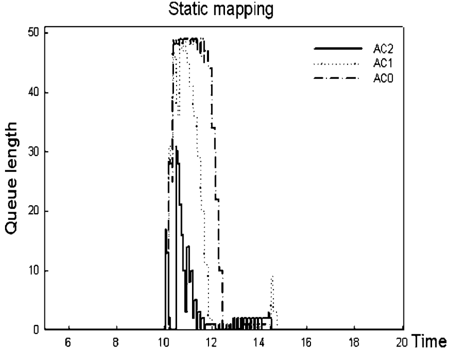

Fig. 6 (b) Queue space utilization of static mapping (Football)

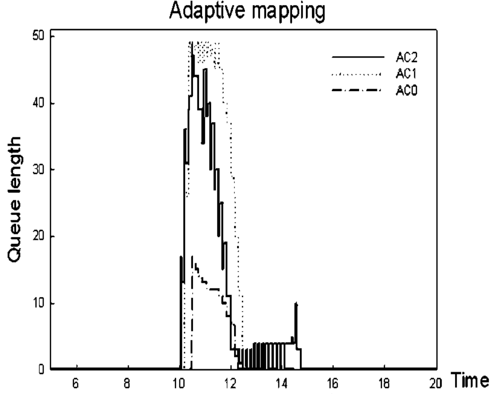

Fig.6(c) Queue space utilization of adaptive mapping (Football)

In the experiments. There are eight ad-hoc wireless nodes where one is video server and another is the video receiver. The video stream is transmitted in unicast mode which the sender does not retransmit to recover loss packets. The data rate of the wireless link is 11 Mbps.

There are two kinds of scenario in the simulations for evaluating the video transmission performance:

- ***Scenario 1***: In this case only one observed video stream is transmitted from the video sender node to the video receiver node. In this scenario, the performance evaluation focused on the frame loss of each video type and the queue space utilization by witnessing the queue length variation of each AC.
- ***Scenario 2***: in this case we used five different loading cases, including different loads of voice traffic (64 k, in

AC[3]), TCP (in AC[1]), and UDP (in AC[0]). Traffic flows were randomly generated and transmitted over the entire simulation environment, and the flow numbers of three kinds of traffic in each case are listed in Table 4. In this scenario, we analyzed the received video quality to evaluate the efficacy of our proposed scheme under various network loading conditions.

The parameter settings of the proposed mapping algorithm are specified as follows: the *threshold_low* value of queue length is twenty packets; the *threshold_high* value of queue length is forty-five packets; the downward mapping probability, *prob_TYPE*, for the 1 frame is 0, for *P* frame is

0.6, and for *B* frame is 0.8. The queue sizes of all ACs are limited to a maximum of fifty packets

4.2 Queue space utilization of mapping algorithms

Table 5 shows the average PSNR and frame loss result of video Football in scenario 1. The packet loss for different frame types is recorded with the video frame significance.

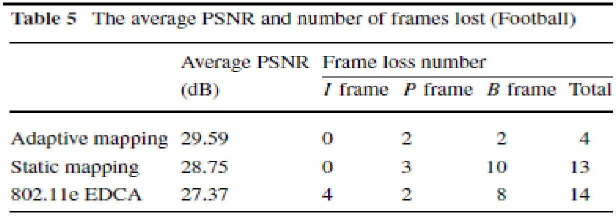

Table 5 The average PSNR and number of frames lost (Football)

	Average PSNR (dB)	Frame loss number			
		I frame	*P* frame	*B* frame	Total
Adaptive mapping	29.59	0	2	2	4
Static mapping	28.75	0	3	10	13
802.11e EDCA	27.37	4	2	8	14

Since the *I* video frames are transmitted with the highest pri- ority in the proposed adaptive mapping algorithm and static mapping algorithm, the number of lost *I* frames are both zero. But the *B* frame is assigned the lowest priority, so the number of lost *B* frames is greater than the other two kinds of frames.

From the view point of the queue space utilization illustrated in Fig. 4, we found that for 802.11e EDCA (as shown in Fig. 6(a)) for which there is no queue management mechanism, all video packets are placed in the AC[2] queue. Consequently, when the queue buffer is full, sub- sequent packets are dropped instantly. In contrast, the static mapping algorithm made use of the two alternate queue spaces AC[1] and AC[0], as shown in Fig. 6(b), yielding a lower packet loss. Furthermore, from Fig. 6(b) of the static approach cannot utilize the higher priority queue space to the maximum, but our proposed mapping algorithm can. By comparing the queue lengths of different ACs, our proposed mapping algorithm attempts to optimize the utilization of the higher priority queue (as shown in Fig.6(c)) to achieve higher channel access probability for video transmissions. As a result, the proposed scheme delivers better video quality (in PSNR) than the other two schemes. Table 5 presents a visual comparison of the video transmission quality of the three schemes considered in the simulations. The quality of the video delivered using the proposed scheme is significantly better than that of the other two schemes.

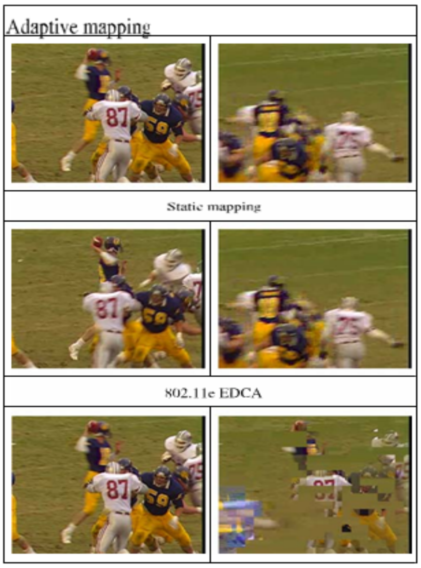

Fig. 7. Visual comparison of the reconstructed video (Football)

Table 5 shows the average PSNR and the number of frames lost for the Foreman video used in scenario 1. The performance results of video transmission are similar to video Football, even though Foreman is a longer video stream. Consequently, as Fig. 8(a), presents the queue space utilization during the video stream transmission period. Our proposed mapping algorithm consistently yields higher utilization of the queue space compared to the other two schemes. Fig.9 illustrates the reconstructed Foreman video, demonstrating that the perceived video quality of our proposed cross-layer mapping algorithm is better than that obtained with the other two alternative schemes.

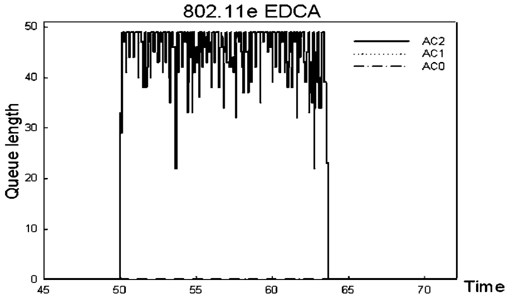

Fig 8(a) Queue space utilization of 802.11e EDCA (Foreman)

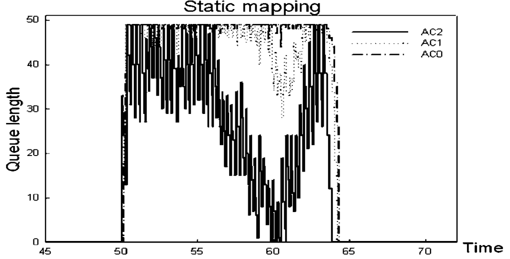

Fig 8(b). Queue space utilization of static mapping (Foreman)

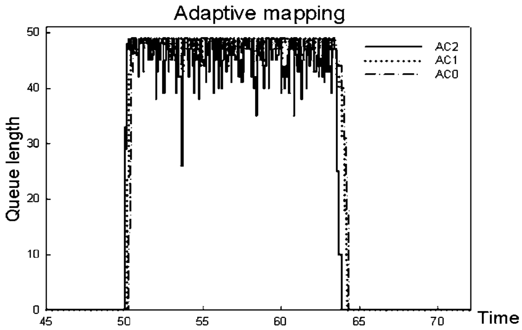

Fig 8(c). Queue space utilization of adaptive mapping (Foreman)

Table 6. The average PSNR and number of frames lost (Foreman)

(dB)		*I* frame	*P* frame	*B* frame	Total
Adaptive mapping	29.73	0	3	8	11
Static mapping	27.51	0	7	44	51
802.11e EDCA	26.59	4	2	59	65

Fig. 9 Visual comparison of the reconstructed video (Foreman)

4.3 Performance comparison of video quality delivered under network loading conditions

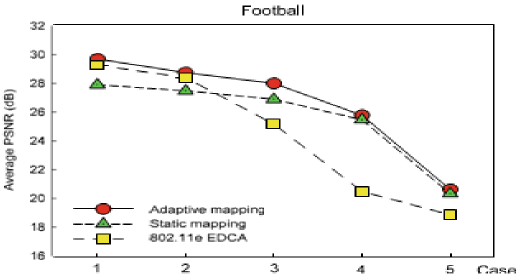

Fig 10. Average PSNR under various loading cases (Football)

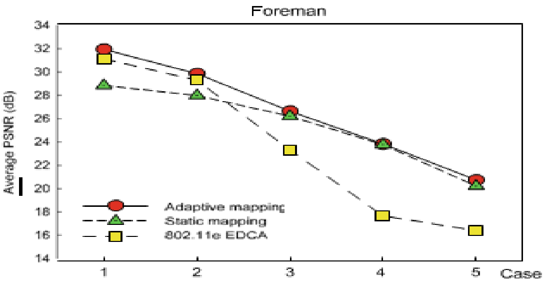

Fig 11.Average PSNR under various loading cases (Foreman)

Figure 10 presents the PSNR variations of video Football for the five different loading cases as mentioned in Table 4. When the network load is light, such as case 1 and case 2, the proposed method gives a similar performance to 802.11e EDCA. The static mapping method results in a lot of unnec- essary packet loss because it places the packets into lower priority queues. However, because of dynamic mapping is based on both the significance of the video data and the net- work traffic load, under heavy network loads (such as case 3, case 4, and case 5) our proposed method yields a better av- erage PSNR than either the static mapping approach or the 802.11e EDCA approach. Fig.11 is the corresponding PSNR trend of video Foreman under five different loading cases. The results obtained are similar to those obtained with the Football video.

CONCLUSION

This paper project an Adaptive Cross-layer Mapping algorithm for Multi view Video Coding (ACM-MVC) over IEEE 802.11e Wireless Local Area Networks (WLANs). To support the varying Quality-of-Service (QoS) requirements of emerging applications, four access categories (ACs) have been defined in the IEEE 802.11e standard. Each AC has a different transmission priority, which implies the higher the transmission priority, the better the opportunity to transmit. In order to improve the transmission of video data over IP networks, the Multiview Video Coding (MVC) extension of the H.264/MPEG 4 AVC standard has been introduced. The MVC adopts the layered video scheme. The proposed mapping algorithm utilizes a cross-layer approach and dynamically assigns packets of different layers to one of the four ACs. The experiment results demonstrate that the proposed mapping algorithm achieves better QoS than the conventional methods such as the static mapping algorithm.

REFERENCES:

1. Ke, C.-H., Lin, C.-H., Shieh, C.-K., Hwang, W.-S., & Ziviani, A. (2007). Evaluation of streaming mpeg video over wireless channels. *Journal of Mobile Multimedia*, *3*(1), 47–64.
2. Floyd, S., & Jacobson, V. (1993). Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, *4*, 397–413
3. NS simulator, http://nsnam.isi.edu/nsnam/index.php/Main_Page.
4. Evalvid, http://www.tkn.tu-berlin.de/research/evalvid/.
5. Ksentini, A., Naimi, M., & Gueroui, A. (2006). Toward an

improvement of H.264 video transmission over IEEE 802.11e through a cross-layer architecture. *IEEE Communications Magazine*, *44*(1), 107–114

6. YUV video sequences (CIF), http://www.tkn.tu-berlin.de/research/evalvid/cif.html.
7. YUV video sequences (QCIF), http://www.tkn.tu-berlin.de/research/evalvid/qcif.html..
8. IEEE Std 802.11e-2005 (2005).*Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 8:medium access control (MAC) quality of service enhancements*.
9. Takeuchi, S., Sezaki, K., & Yasuda, Y. (2005). Dynamic adaptation of contention window sizes in IEEE 802.11e wireless LAN.In *Fifth international conference on information, communicationsand signal processing (ICICSP 2005)* (pp. 659–663).
10. Kim, J. O., Tode, H., & Murakami, K. (2005). Service- based rate adaptation architecture for IEEE 802.11e QoS networks. In *IEEE global telecommunications conference*,
11. Siris, V. A., & Courcoubetis, C. (2006). Resource control for the EDCA mechanism in multi-rate IEEE 802.11e networks. In *International symposium on a world of wireless, mobile and multimedia networks*, WoWMoM 2006 (pp. 419–428), 26–29 June 2006.
12. Bosveld, F., Lagendijk, R. L., & Biemond, J. (1992). Hierarchical video coding using a spatio-temporal subband decomposition.In *IEEE international conference on acoustics, speech, and signalprocessing*, ICASSP-92 (Vol. 3, pp. 221–224), 23–26 March 1992.
13. Wang, Y., & Huang, B. (2004). Performance improvement of IEEE 802.11e. In *The 13th IEEE workshop on local and metropolitan area networks*, LANMAN 2004 (pp. 51–54), 25–28 April 2004.
14. Majkowski, J., & Palacio, F. C. (2006). Enhanced TXOP scheme for efficiency improvement of WLAN IEEE 802.11e. In *Vehicular technology conference*, VTC-2006 Fall. 2006 IEEE 64th (pp. 1–5),Sept. 2006.
15. Majkowski, J., & Palacio, F. C. (2006). Dynamic TXOP configuration for Qos enhancement in IEEE 802.11e wireless LAN. In *International conference on software in telecommunications and computer networks*, SoftCOM 2006 (pp. 66–70), Sept. 29 2006– Oct. 1 2006.
16. Kim, J. O., Tode, H., & Murakami, K. (2005). Service-based rate adaptation architecture for IEEE 802.11e QoS networks. In *IEEE global telecommunications conference*, GLOBECOM'05 (Vol. 6, pp. 3341–3345), 28 Nov.–2 Dec. 2005.

ABOUT AUTHORS:

Bhavsingh Maloth is currently a Assistant Professor teaching in the Department of Information Technology, G.Pullaiah College of Engineering & Technology, JNTU Anantapur University, India. He holds M.Tech degree in Information Technology from University of JNTU Hyderabad (gold medallist), His Current Research areas include Cloud Computing, wireless QoS, He is a Member of Computer Society of India (CSI), and He is a Life Member of International Association of Engineers (IAENG).

Ashok Sarabu is a pursuing B.Tech. in Information Technology at G.Pullaiah College of Engineering &Technology, JNTU Anantapur, India. He is a Member of IEI; He did R&D in DeskStream at GPCET. His current research areas include wireless QoS.MAC Protocols, Cross Layer Techniques

Vinod Kumar is a student in G.Pullaiah College of Engineering & Technology, JNTU Anantapur, India.

His research areas include wireless QoS, Wireless LAN protocols.