CHAPTER 1

Introduction

1.1 Introduction

Since the first IEEE 802.11 [1] wireless network standard was announced in 1997, IEEE 802.11a, 802.11b, 802.11g, and the latest draft of 802.11n come to the public subsequently. On the other hand, many handheld devices, such as PDA, wireless TV (IPTV), wireless phone, are equipped with wireless network access and multimedia support. People come to realize the inevitable trend of the integration of wireless network and multimedia service. IEEE 802.11e [2] was thereby proposed to prioritize packets into different categories for priority scheduling. JVT (Joint Video Team) proposed H.264/AVC (Advanced Video Coding) [3] video compression standard in 2003, and amendments of the second and the third editions in 2005. As Figure 1.1 shows, H.264/AVC can offer not only high-quality service in high-bandwidth network but also acceptable quality service in low-bandwidth network. H.264/AVC is able to adapt to wireless network and is very suitable for video streaming services. This paper will first discuss some essential problems of IEEE 802.11 DCF and IEEE 802.11e EDCA in managing video streaming, and then propose a hybrid design framework to improve the quality of video streaming. This framework consists of a MAC-centric cross-layer architecture, a timebase retransmission mechanism and a single-video multi-level queue. The paper organization is as follows. Chapter 1 introduces fundamental background. Chapter 2 introduces related works. Chapter 3 proposes our hybrid design framework. Chapter 4 presents simulations and results. Last chapter offers brief concluding remarks.

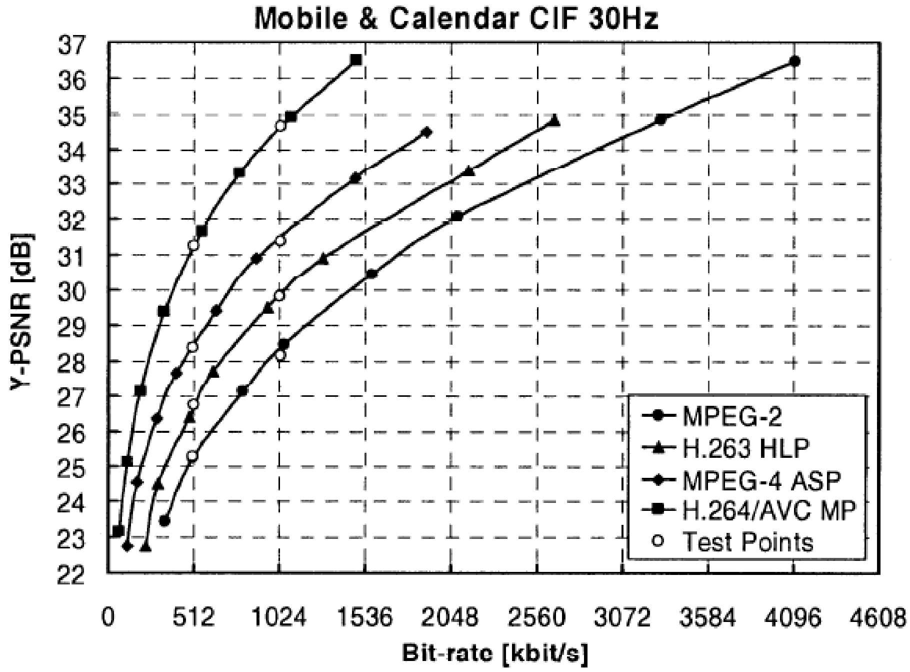

Figure 1.1: PSNR against bit-rate for different video codec [5][6][8].

1.2 Background

1.2.1 IEEE 802.11 DCF (Distributed Coordination Function)

DCF (Distributed Coordination Function), a mechanism of IEEE 802.11 MAC layer, is used to share medium between multiple stations. Before transmitting, each station should check whether the channel is idle. If the channel is busy then the other stations will set backoff times to avoid collisions. Backoff time is determined by CW (Contention Window). As a collision occurs, the station has to double the contention window. As formula 1.1 shows, CW_{min} is minimum of initial CW and the initial value i is 0. As transmission fails, i increases until CW equals to maximum of CW. Formula 1.2 shows backoff time.

$$CW = (CW_{min} \times 2^{i}) - 1, \qquad (1.1)$$

$$\text{backoff time} = \text{random}(0, CW) \times \textit{aSlotTime} \qquad (1.2)$$

In Formula 1.2, *aSlotTime* is 9 μs in IEEE 802.11a and 20 μs in IEEE 802.11b/g. In DCF, once a station finishes a transmission, it waits a DIFS (Distributed Interframe Space) time. If the channel is busy at the time the station starts next transmission, the station moves into a backoff interval a CW time for next transmission, as Figure 1.2 and Figure 1.3 show. A collision may occur as a station transmits a packet. As Figure 1.4 shows, when A and D stations generate the same backoff time in backoff interval, a collision occurs. Then the two stations generate new backoff time to avoid collosions.

Figure 1.2: The channel is idle for a DIFS interval before the transmission [4].

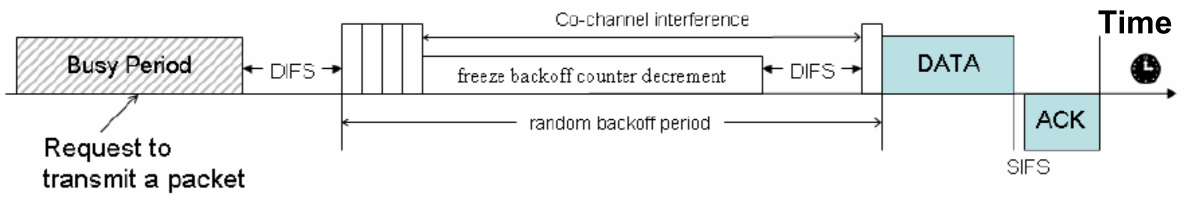

Figure 1.3: The channel is busy during a DIFS interval before the transmission [4].

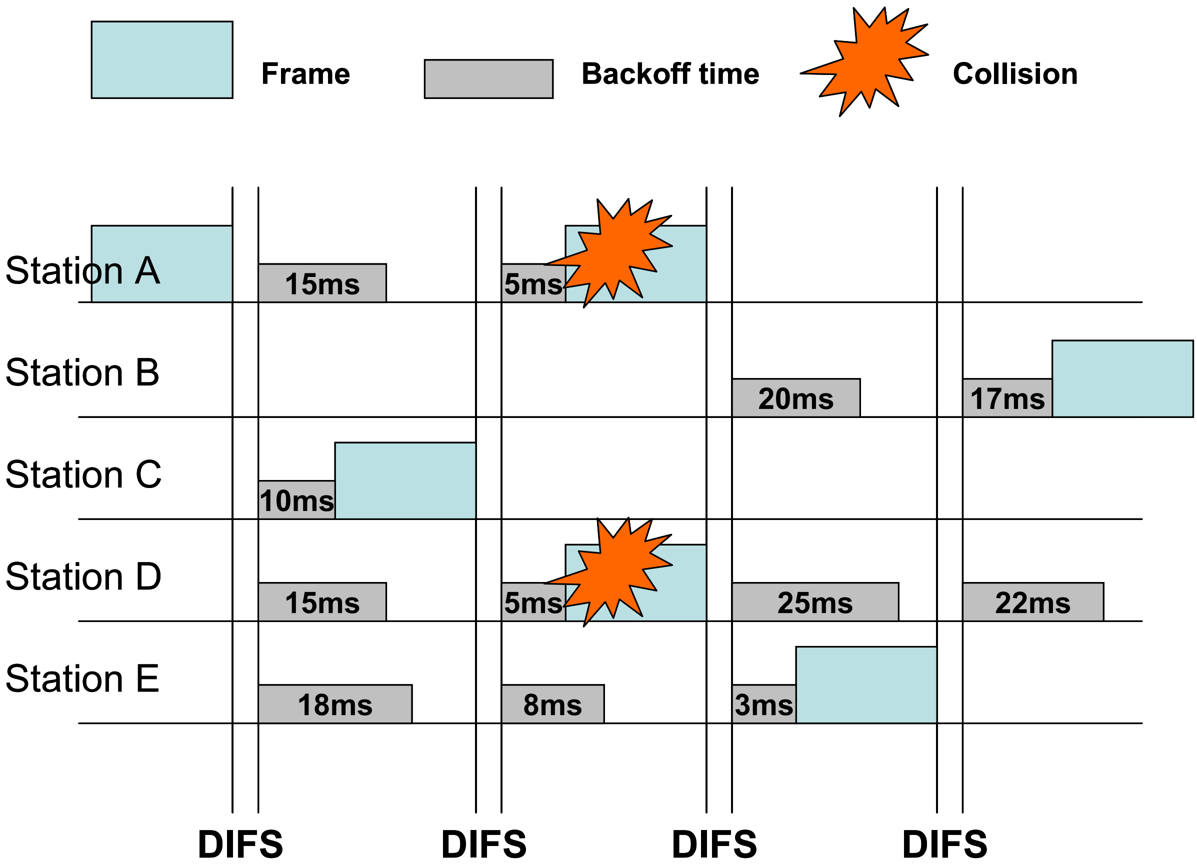

Figure 1.4: An example of collision.

1.2.2 IEEE 802.11e EDCA (Enhanced Distributed Channel Access)

IEEE 802.11e [2], a supplementary protocol to IEEE 802.11, defines a set of QoS (Quility of Service) enhancements for IEEE 802.11. IEEE 802.11e EDCA (Enhanced Distributed Channel Access) enhances DCF single-queue by introducing priority multi-level queue. EDCA makes it compatible with DCF DIFS using AIFS (Arbitration InterFrame Space). IEEE 802.11e defines AIFS and CW in different lengths of time. As Figure 1.5 shows, it defines four ACs (Access Category) of priority queues which are AC_VO (Access Category of Voice), AC_VI (Access Category of VIdeo), AC_BE (Access Category of Best Effort) and AC_BK (Access Category of Background). AC is a queue and each queue has different AIFS time, CW time and retry limit. As Figure 1.6 shows, a higher priority queue will be assigned shorter AIFS time and CW time. Compared with IEEE 802.11 single queue, IEEE 802.11e priority multi-level queue is able to enhance bandwidth utilization.

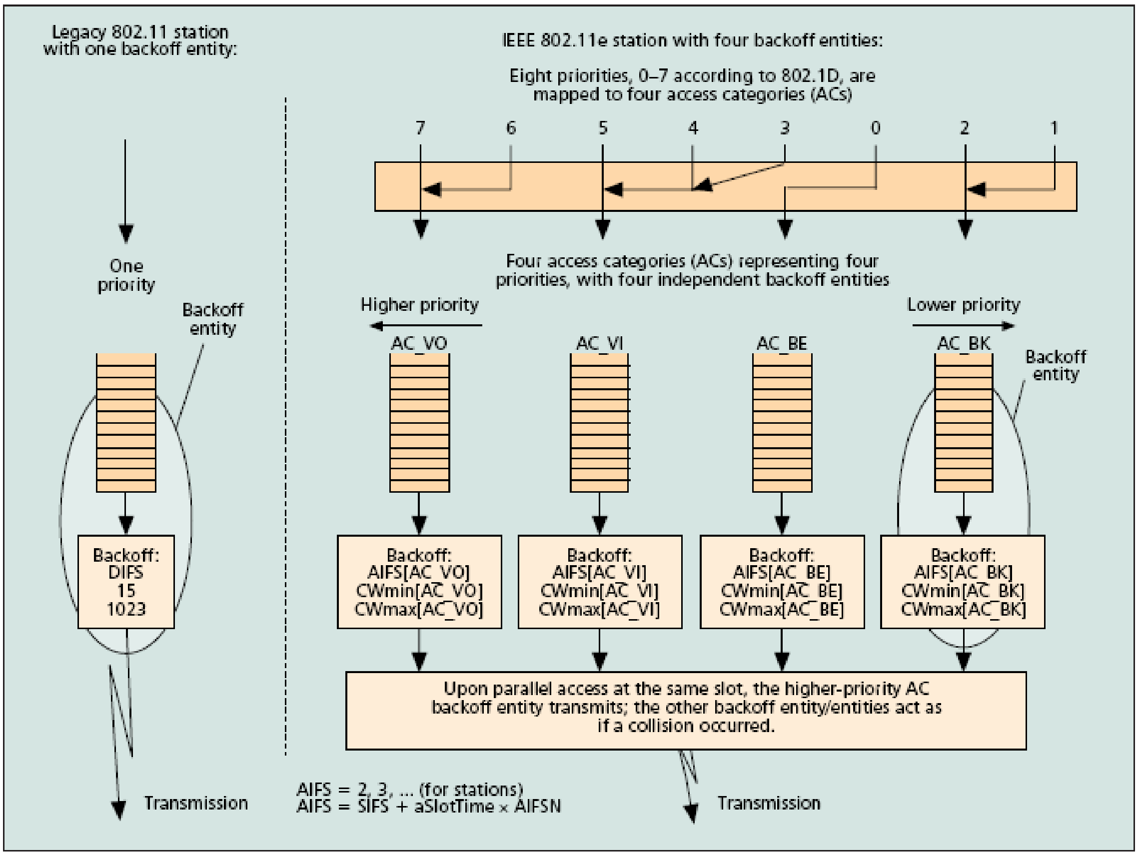

Figure 1.5: Legacy 802.11 station and 802.11e station with four ACs within one station [14].

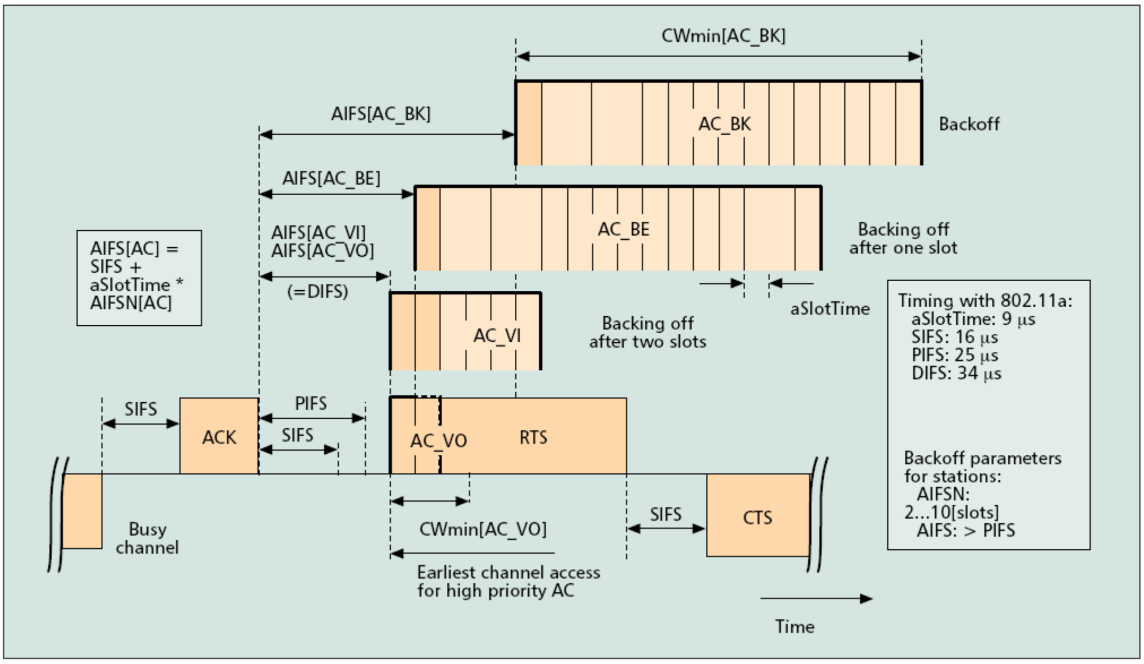

Figure 1.6: In EDCA, multiple backoff entities contend for medium access with different priorities in parallel. The earliest possible medium access time after a busy medium is DIFS [14].

1.2.3 H.264/AVC (Advanced Video Codec)

H.264/AVC [3] is the next generation video compression standard. H.264 is known as AVC or MPEG 4 Part10 proposed by JVT. H.264/AVC has better video quality and compression ratio than that of MPEG 2/4 [5]. H.264/AVC consists of VCL (Video Coding Layer) and NAL (Network Abstraction Layer), and is very suitable for video streaming services. VCL is a encoding / decoding unit for traditional image compression. NAL is used to adapt to the network. Videos are compressed into bit-streams by VCL. NAL will then encapsulate bit-streams into NAL-units and mark the video type in the header.

Architecture of H.264/AVC

H.264/AVC has a strict hierarchy structure. As Figure 1.7 shows, the hierarchy structure is followed by macroblock (MB), slice, slice group, frame/field-picture and sequence. And A

MB is segmented to blocks and sub-blocks according the setting of compression. In fact, a video frame consists of MBs and a sequence consists of many video frames. The first video frame of each sequence must be IDR frame (Instantaneous Decoding Refresh frame) encoded by intra-prediction. Intra-prediction means the frame can be decoded by itself without making any references to other frames. A slice consists of MBs and a video frame consists of many slices (Figure 1.8). Slice is the smallest self-decodable unit for H.264/AVC. The self-decodable unit means a slice can be decoded by itself without making any references to other slices. The advantage is that the slice can be decoded at once while receiving a slice without waiting for all slices of a video frame. In general, we set one slice in one packet while encoding H.264/AVC video. If the slice is lost or wrong while transmitting, it doesn't affect other slices. It is different from MPEG-2. H.264/AVC allows a frame consisting of a slice. As Figure 1.8 shows, the slice of MPEG-2 consists of a row of MBs, but the slice of MPEG-4 or H.264/AVC can consist of more than a row of MBs. The slice for H.264/AVC has a special characteristic called FMO (Flexible Macroblock Ordering). It means that the MBs don't have to limit to the arrangement of raster scan. For example, the arrangement of Figure 1.8(c) is suitable to more slice groups in the foreground and a single slice group in the background. The advantage is that it can use compressional parameters of different qualities for different slice groups. For example, the objects in the foreground are usually interesting to our eyes, so we can use smaller compression ratio to keep better quality. In short, each video frame is divided into many slices. In this paper we assume one slice one packet.

Slice for H.264/AVC

There are five types of H.264/AVC slices which are I-slice (Intra slice), P-slice (Predicted slice), B-slice (Bi-predicted slice), SI-slice (Switch Intra slice) and SP-slice (Switch Predicted slice). I-slice has the richest reference information for P/B slices and is the most critical slice. I-slice is coded by intra-prediction, P-slice by intra-prediction and inter-prediction, B-slice by inter-prediction. SI/SP-slice is used to adapt I/P-slice to different bitrates.

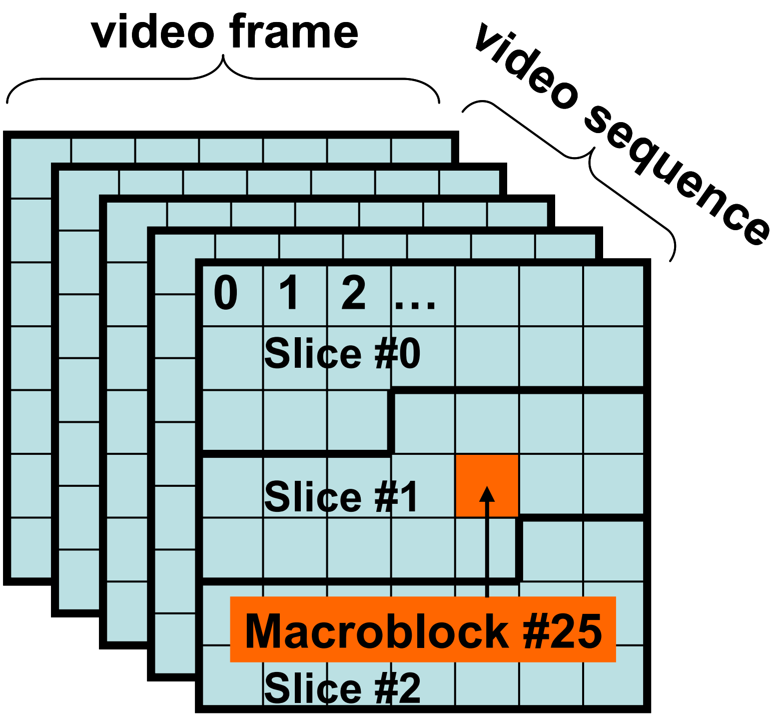

Figure 1.7: Architecture of H.264/AVC.

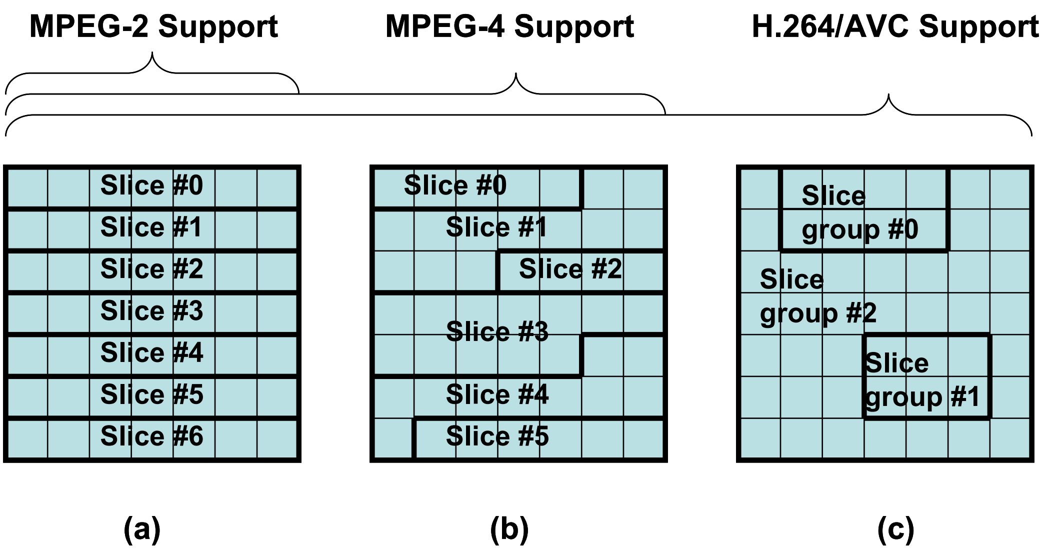

Figure 1.8: (a)(b)Slice without FMO and (c)slice group with FMO.

1.3 Motivation

Video streaming occurs a long delay time or high packet loss rate while transmitting in IEEE 802.11 wireless network because IEEE 802.11 uses CAMA/CA. As the number of nodes increases, the probability of collisions increases. Thus, it leads to packet loss rate increasing [7][9]. Video streaming has several characteristics as follows:

1 Video streaming is tolerant of packet loss.

Video streaming is tolerant of packet loss. For example, MPEG consists of video frames. And each video frame consists I-frames (Intra frames), P-frames (predicted frames) and B-frames (Bidirectional frames). The smallest decoded unit for video streaming is slice, and a video frame consists of many slices [10]. While the video frame doesn't receive enough slices, the video frame cannot be decoded. If fewer video frames aren't decoded, details of the video will be ignored by our eyes. If more video frames are decoded, the video will be interrupted.

2 Video frames (slices) tansmitted in the limited time are valid.

Each video frame (slice) has a deadline to be transmitted because receiver in the client plays the video in real-time, but receiver still needs buffer time to transmit, decode and save the video file. If the video frames (slices) are still transmitting while deadlines were due, video frames (slices) cannot be decoded. Then the video will be interrupted or cannot be played smoothly.

3 Each video frame (slice) has different significances.

As Figure 1.9 shows, it's a playing order. I-frame is intra-coded, but P/B-frame is inter-coded. I-frame is from the compression of original video frame, P-frame is reference to front P-frame or I-frame, and B-frame is reference to front and back I/P-frame. By above, I-frame is the most critical frame for P/B-frame to refer to. For example, we know that the playing order is I_0, B_1, P_2, B_3, P_4, B_5, P_6, B_7, P_8, B_9 and I_{10} as Figure 1.9 shows. So we can calculate the sending order; that is I_0, P_2, B_1, P_4, B_3, P_6, B_5, P_8, B_7, I_{10} and B_9. If all the packets for I-frame are transmitted successfully, it can improve the quality of video.

Unfortunately, the lower layer cannot get the information of packets from higher layer in IEEE 802.11 wireless network. So we propose MAC-centric cross-layer design for video streaming. Thus, using cross-layer framework improves quality of video.

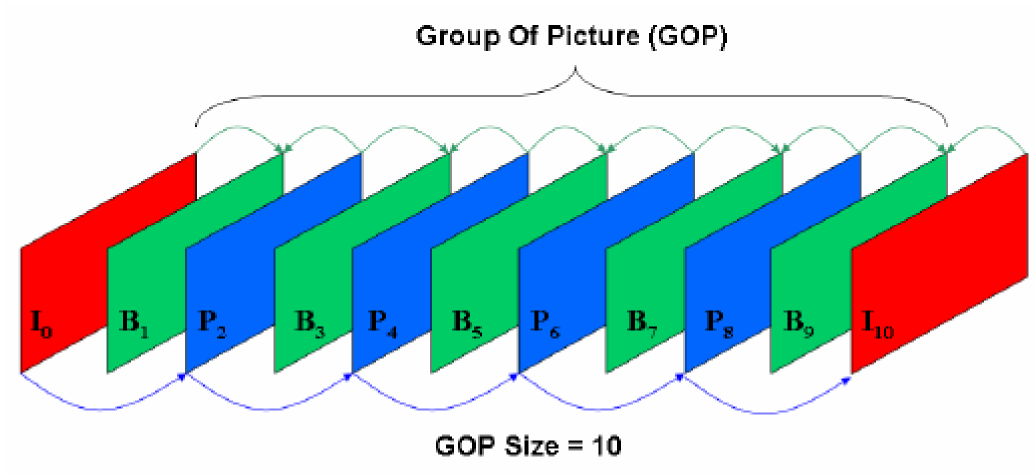

Figure 1.9: Typical inter-coded video frame structure [4].