Stevens Institute of Technology

**MPEG Video Compression**

Course: CS550 Computer Organization and Programming

Junjie Chen

Professor: Edward Banduk

December 2021

1. **MPEG And Video Compression introduction**

MPEG is known as the Moving Picture Experts Group which is a working group of ISO (International Organization for Standardization) and IEC(International Electrotechnical Commission) in charge of the development of international standards for compression, decompression, processing, and coded representation of moving pictures, audio and their combination.

In multimedia processing, the most basic requirement is to be able to process sound, animation, and video signals dynamically. However, the data volume of images is very large, and without compression processing of video data, real-time cannot be achieved. For example, a color digital video image with medium resolution (640 × 480) the amount of data is about 7.37Mbits per frame, if the frame rate of 30 frames per second, the transmission rate of the video signal is about 221.1Mbits per second, a minute of video requires more than 13266Mbits of hard disk space. Such a large amount of data, it is difficult to seek huge storage devices to store these data, and it is also difficult for the computer to transfer these data from the memory to CPU in real time, therefore, video data compression technology has become a key technology for developing video processing in multimedia systems.

The analysis of the video metadata reveals that the redundancy existing in the original video data provides the possibility of data compression implementation. First, each image has a large spatial redundancy in the data due to the fact that neighboring pixels within a frame are spatially correlated with each other. Second, there is also a strong correlation between successive frames of the video sequence signal, indicating that there is a large temporal redundancy in the data. For example, for a speaker image sequence in television, there may be only minor differences between adjacent frames caused by small changes in the head, eyes, and mouth. Furthermore, in the application area of multimedia system, human is the main receiver and eye is the receiving end of image information, so that the high compression rate can be achieved by taking advantage of the fact that human vision is insensitive to sharp edge changes and eye is sensitive to image brightness information and weak to color resolution, so that the image signal recovered from compressed data still has satisfactory subjective quality.

The availability of excellent video compression technology makes it possible to process high quality motion pictures dynamically and in real time at low cost, with low rate requirements and limited bandwidth. Among them, MPEG-1, MPEG-2, MPEG-4 and MPEG-H Part2 standards which have been established by MPEG play an important role in modern multimedia video compression.

1. **MPEG Series Standards of Video Compression History**
   1. MPEG-1

MPEG-1 was established in 1993 and is designed as an industrial-grade standard for devices with different bandwidths, such as CD-ROM, Video-CD, CD-I. It can compress images at SIF standard resolution (352×240 for NTSC; 352×288 for PAL) at a transfer rate of 1.5 Mbps, 30 fps, with CD sound quality and a quality level basically equivalent to VHS. MPEG-1 can be encoded at rates up to 4 to 5 Mbps, but the decoded image quality decreases as the rate increases. MPEG-1 has a wide range of applications, the most successful of which are VCD products. asymmetric digital subscriber lines (ADSL), video-on-demand (VOD), and educational networks.

* 1. MPEG-2

MPEG-2 was created in 1995 and is designed for advanced industry standard image quality and higher transfer rates. MPEG-2 offers transfer rates between 3 and 10 Mbits/sec and resolutions up to 720 x 480 in NTSC. MPEG-2 also features a wide range of compression ratios to accommodate different picture quality, storage capacity, and bandwidth requirements.

MPEG-2 technology is the standard technology for DVD implementation and can also be used to provide broadcast-quality digital video for broadcast, cable, and satellite broadcasts.

Since the excellent performance of MPEG-2 was already suitable for HDTV, MPEG-3, which was originally intended for HDTV, was abandoned before it was even introduced.

* 1. MPEG-4

MPEG-4 incorporates many of the features of MPEG-1, MPEG-2, and other related standards, as well as new ones such as extended VRML support for 3D rendering, object-oriented composite files with audio, video, and VRML objects, support for externally specified Digital Rights Management, and a variety of interactivity options.

MPEG-4 is a standard that is still in development and is separated into several components. MPEG-4 Part 2 including Advanced Simple Profile, which is used by codecs like DivX, Xvid, Nero Digital, and 3ivx, as well as QuickTime 6, and MPEG-4 Part 10 including Advanced Video Coding, which is used by the x264 encoder, Nero Digital AVC, QuickTime 7, and high-definition video media like Blu-ray Disc.

* 1. MPEG-H

MPEG-H is one of the newest standards that is still being promoted and not widely used. It was published in 2011. In contrast to previous standards, larger block structures with flexible mechanisms of sub-partitioning are introduced in HEVC as MPEG-H part2. The HEVC draft specification defines variable-block-sized coding units (CUs) for this, which describe a rectangular sub-partitioning of a picture. The coding unit structure replaces the macroblock structure of prior video coding standards with a variable-block-sized similar structure, and each CU contains one or more variable-block-sized prediction unit(s) (PUs) and transform unit(s) (TUs). Prior to transferring the frame into a reference decoded picture buffer, an adaptive loop filter (ALF) is applied within the prediction loop, improving objective and subjective quality. In addition to a deblocking filter with similar characteristics to the deblocking filter found in AVC, the ALF has been incorporated as a feature in the HEVC architecture. In a similar way as AVC, context-adaptive entropy coding techniques are used.

1. **Main technologies of MPEG video compression**
   1. Motion compensation

Motion compensation is a method of describing the difference between adjacent frames (adjacent in this case means adjacent in coding relationship, two frames may not be adjacent in playback order), specifically how each small block of the previous frame (adjacent in this case means preceding in coding relationship, not necessarily preceding the current frame in playback order) moves to a certain position in the current frame. This approach is often used by video compression/video codecs to reduce time-domain redundancy in video serial. It can also be used for deinterlacing and motion interpolation operations.

A video serial contains a certain number of pictures - usually called frames. Adjacent pictures are usually very similar, i.e., they contain a lot of redundancy. The purpose of using motion compensation is to improve the compression ratio by eliminating this redundancy.

Usually, image frames are processed in groups. The first frame of each group (usually the first frame) is encoded without using motion estimation, and this frame is called an intra frame or I frame. The other frames in the group are encoded using an inter frame, usually a P frame. This coding method is often called IPPPP, which means that the first frame is an I frame and the other frames are P frames when coding.

When predicting, not only the current frame can be predicted from the past frame, but also the future frame can be used to predict the current frame. Of course, at the time of encoding, the future frames must be encoded earlier than the current frames, i.e., the order of encoding and playback is different. Usually such a current frame is predicted using both past and future I-frames or P-frames, and is called a bidirectional prediction frame, or B-frame. An example of the encoding order for this type of encoding is IBBBPBBPBBPBB.

According to the actual needs, there are also improved various motion compensation algorithms, for instance, BMC for block motion compensation, VBSMC for Variable Block Size Motion Compensation, OBMC for Overlapped block motion compensation.

* 1. DCT(Discrete Cosine Transform)

A discrete cosine transform (DCT) expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. The DCT, first proposed by Nasir Ahmed in 1972.

The DCT-II, also known as simply the DCT, is the most important image compression technique. There, the two-dimensional DCT-II of N \* N blocks are computed and the results are quantized and entropy coded. In this case, N is typically 8 and the DCT-II formula is applied to each row and column of the block. The result is an 8 × 8 transform coefficient array in which the (0,0) element (top-left) is the DC (zero-frequency) component and entries with increasing vertical and horizontal index values represent higher vertical and horizontal spatial frequencies.

* 1. Flexible interlaced-scan

Each frame of an interlaced-scan image consists of two fields, so there are three coding methods for interlaced-scan images: 1) combining two fields into one frame for coding; 2) coding two fields separately; 3) combining two fields into one frame, but at the macroblock level, dividing one frame macroblock into two field macroblocks for coding.

The first two coding methods above are called Picture-adaptive frame-field (PicAFF or PAFF) and the third one is called Macroblock-adaptive frame-field (MBAFF).

PAFF: For a motion image, the spatial correlation between two adjacent lines in a frame is smaller compared to a progressive scan image because of the large scan interval between the fields, so it is more stream-efficient to encode the two fields separately. For a non-motion image, there is a large spatial correlation between two adjacent rows, and it is more efficient to combine the two fields into one frame for encoding. So the PAFF coding method can be adaptively adjusted for the whole image, choosing frame coding or field coding.

MBAFF: When there are both motion and non-motion areas in the image, the disadvantage of PAFF becomes obvious, and the adaptive granularity of PAFF is too coarse to achieve to meet the finer coding requirements, so MBAFF comes into being. MBAFF's choice of frame coding or field coding is based on macroblocks, MBAFF combines two fields into one frame for coding, but divides each frame macroblock (16x16) is divided into field macroblock pairs (8x16), and for each frame macroblock, the code stream sizes generated by frame coding and field coding are compared and the most stream-saving method is used.

**4.** **Conclusion**

Pursuing more efficient compression algorithms and higher compression ratios has become a major challenge in computer science and technology. MPEG has made great contributions in this area, especially in audio and video images, and the technical details of which are worthy of in-depth study.

**References**

[1] MPEG | The Moving Picture Experts Group website. (n.d.). Retrieved December 3, 2021, from <https://mpeg.chiariglione.org>.

[2] 吴明芳. MPEG-4视频编码算法研究及软件实现. (Doctoral dissertation, 重庆大学).

[3] Moving Picture Experts Group - Wikipedia. (n.d.). Retrieved December 4, 2021, from <https://en.wikipedia.org/wiki/Moving_Picture_Experts_Group>.

[4] MPEG-4 - Wikipedia. (n.d.). Retrieved December 4, 2021, from <https://en.wikipedia.org/wiki/MPEG-4>.

[5] 运动补偿- 维基百科，自由的百科全书. (n.d.). Retrieved December 5, 2021, from <https://zh.wikipedia.org/wiki/%E8%BF%90%E5%8A%A8%E8%A1%A5%E5%81%BF>.

[6] Advanced Video Coding - Wikipedia. (n.d.). Retrieved December 5, 2021, from <https://en.wikipedia.org/wiki/Advanced_Video_Coding>.