Modeling and Optimization of PET-Redundancy Assignment for MPEG Sequences

by Rainer Storn[1)]

TR-95-018

May 1995

Abstract

Priority Encoding Transmission (PET) is an encoding scheme which provides multiple levels of redundancy in order to protect the different contents of a data set according to their importance. The task of optimally assigning redundancies for the PET encoding scheme is investigated for the special case of MPEG-1 encoded video sequences. The prerequisites for this optimization problem and the way of proceeding for its solution are outlined and several suggestions for further improvements are given.

[1)]International Computer Science Institute, 1947 Center Street, Berkeley, CA 94704-1198, Suite 600, Fax: 510-643-7684. E-mail: storn@icsi.berkeley.edu. On leave from Siemens AG, ZFE T SN 2, Otto-Hahn-Ring 6, D-81739 Muenchen, Germany. Fax: 01149-636-44577,
E-mail: rainer.storn@zfe.siemens.de.

1. Introduction

The number of computer applications which make use of MPEG-1 encoded video sequences is continuously on the rise. At the same time many of these applications are intended for usage in a heterogeneous networking environment which exhibits links with different bandwidths and workstations with varying computational power. A major goal of priority encoding transmission (PET) is to provide graceful degradation for MPEG-1 encoded video streams in order to allow users with different HW-equipment to share the same applications [1]. To this end PET protects the different contents of an MPEG-1 video according to their importance via a multilevel redundancy scheme and information spreading. The intention is to increase the likelihood that the more important parts of the information can still be recovered if the video stream is corrupted by packet losses. These packet losses can be caused by different events like network congestion, fading in satellite transmission, insufficient computational power of the receiving workstations and the like. Coding the video streams in such a way that a potential degradation is perceived as graceful is a challenging task. First it must be clear how the different quality reductions in a video sequence are perceived by the human visual system. These reductions affect basically colour, spatial and temporal behaviour of the video as well as overall precision of the pixels. A thorough understanding of the MPEG-1 as well as the PET encoding process is then necessary to analyze what kind of changes lead to which effect, assuming that a certain packet loss behaviour is given. An optimized encoding scheme should yield the best video quality for a certain pattern of disturbance.

2. MPEG-1 Compression

MPEG (Moving Picture Experts Group) is a standardization group which has defined a widely accepted compression standard for video sequences in the 1.5Mbps range. This standard, the development of which was primarily driven by the rapidly expanding CD-ROM-Video market, is often called MPEG but should be called more accurately MPEG-1 as there is a second proposal, MPEG-2 [2], which addresses digital television and is still under scrutiny.

An MPEG-1-video is represented by a layered bistream as being described in [3]. The top layer is called a sequence which is encapsulated by the sequence delimiters header and trailer. The sequence header contains information such as picture size, picture rate and bit rate. The trailer consists of a 32bit end code.

The header is followed by any number of GOPs (Group of Pictures). A GOP provides random access, since it is the smallest coding unit which, under certain circumstances, can be decoded independently.

A GOP consists of a GOP header being followed by different types of frames. MPEG-1 employs three frame types: intraframe frames (I-frames), predictive frames (P-frames) and bi-directionally interpolated frames (B-frames) as indicated in Fig.1. I-frames, which are self contained and can be decoded independently, are basically encoded like JPEG pictures as illustrated in Fig. 2. JPEG stands for Joint Photographic Experts Group and also defines an international digital image compression standard for continuous-tone (multilevel) still images [4]. In every GOP there is one I-frame which immediately follows the GOP header if the GOP is viewed in bitstream order (further details concerning bitstream and display order are described in [2]). Due to the perceptive properties of the human visual system the luminance component is sampled with more spatial accuracy than the chrominance components. Each chrominance

component, chrominance red and chrominance blue, has one-half the vertical and horizontal resolution of the luminance component. A so-called macroblock in MPEG-1 consists of 4 luminance blocks and just 2 chrominance blocks as shown in Fig. 3.

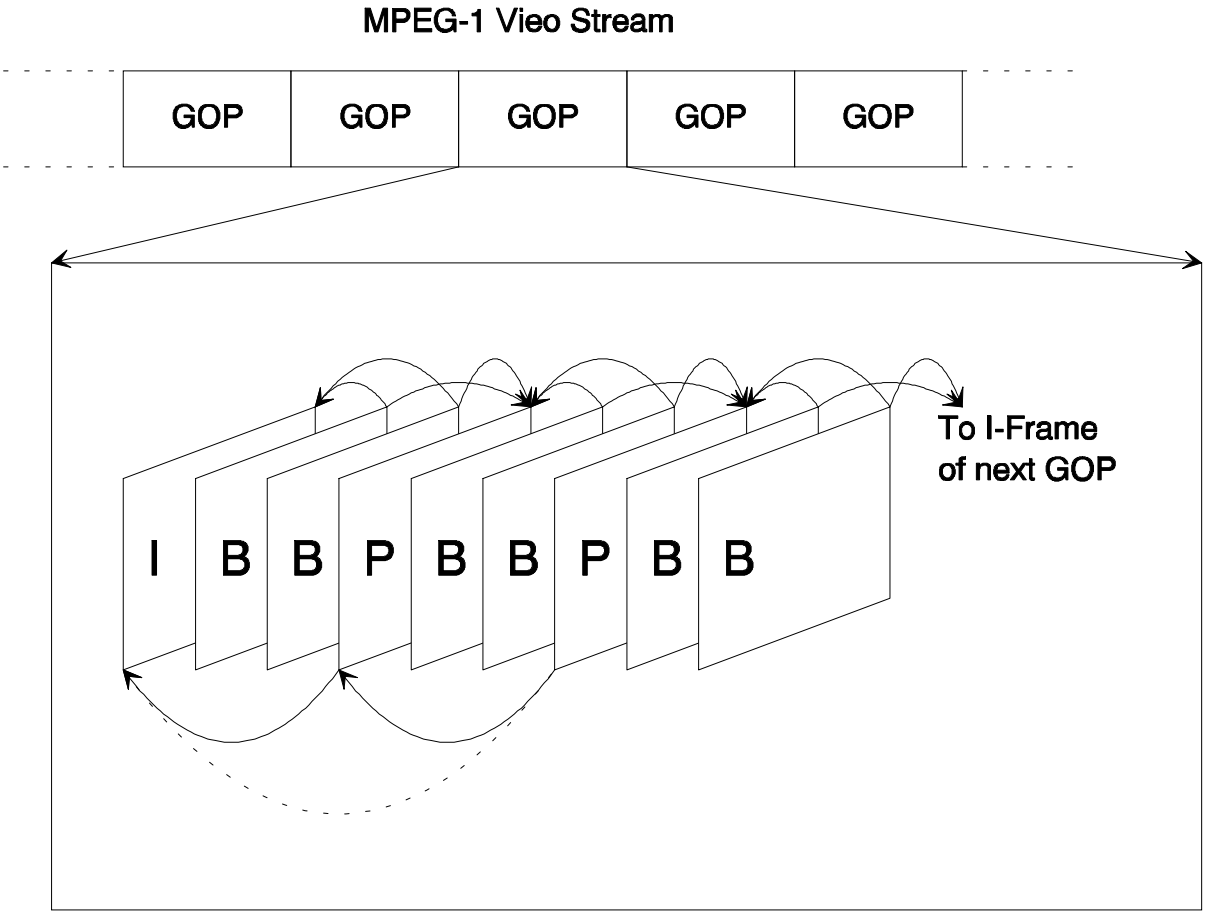

Fig.1: Example of Interframe coding of an MPEG-1 video stream without the GOP-flag being set.

As indicated in Fig. 1 P-frames need a past I- or P-frame for decoding making use of motion-compensated prediction. Motion-compensated prediction assumes that "locally" the current picture can be modeled as a translation of the picture at some previous time. Hence, in order to code a certain block of a P-frame, the previous I- or P-frame is used as a reference and searched for a block which exhibits the least difference to the block to be encoded (Fig. 4). Once a good match has been found, the displacement and an error term are used to encode the block under scrutiny. The comparison search is fairly computation expensive which is why the search is restricted to the luminance component of the block. After the search, however, the actual coding is done on luminance as well as chrominance.

B-Frames are encoded with motion vectors from a past (I- or P-) and a future (I- or P-) frame via interpolative techniques. B-frames are not used for prediction of other frames so the errors in them do not propagate, as opposed to the errors in I- or P-frames. This makes the B-frames the least critical of the three frame types. Errors in I- and P-frames propagate until the end of the GOP if the GOP has its "closed GOP"-flag set [2]. If the flag is not set the error in a P-frame can even propagate into the next GOP, and an error in the I-frame might affect B-frames in the previous GOP.

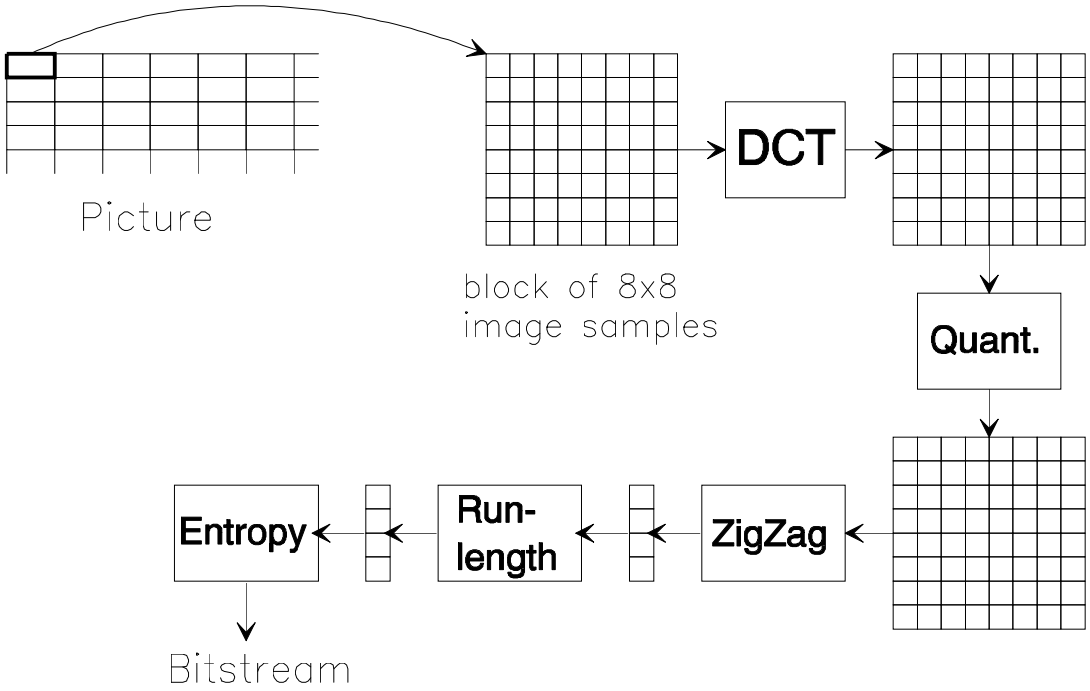
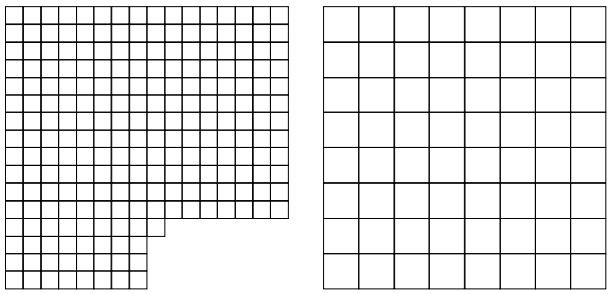
Fig.2: Block Transfer Encoding, the basis of JPEG and MPEG encoding.