Reprint

Compressed Disparity Information Transmission over Constant Bit Rate ATM Channels

D. Papadimatos, T. Antonakopoulos and V. Makios

Letter

Compressed Disparity Information Transmission over Constant Bit Rate ATM Channels [*]

DIONYSIS PAPADIMATOS, THEODORE ANTONAKOPOULOS , VASSILIOS MAKIOS

Department of Electrical Engineering and Computers Technology, University of Patras, 26500 Patras, Greece

{antonako, V.Makios}@ee.upatras.gr

Abstract. This letter presents a real-time lossless compression/decompression unit for disparity map information used in 3D teleconferencing systems. A lossless compression algorithm is used to compress the disparity map data in real-time, resulting in a variable bit-rate data stream that has to be transmitted through a constant bit-rate channel. The system uses a controlled-data-loss method for data rate adaptation and for minimizing the loss of information.

1 INTRODUCTION

The new generation of 3D telepresence videoconferencing systems aims to provide the user with an enhanced illusion of true contact. The goal is to achieve a true telepresence illusion for remote partners, while achieving efficient utilization of the transmission channel. To achieve the telepresence illusion, remote views at arbitrary positions must be synthesized from the views of a stereoscopic camera with a large baseline. The actual viewpoint is adapted, according to the head position of the viewer, in such a way that the impression of motion parallax is produced. This involves the use of disparity information. In order to reduce the complexity of the receiver, the disparity data are produced at the transmitter and transmitted to the receiver. An example of such a system was developed in the framework of PANORAMA [1], where an ATM CBR channel is used as the communication link. The whole system, shown in Figure 1, consists of a disparity estimator, a synchronizer, a stereoscopic MPEG2 encoder, a disparity encoder and a multiplexer on the transmitter side and a demultiplexer, a disparity decoder, a MPEG2 decoder, an interpolator with viewpoint adaptation and an autostereoscopic display at the receiver side. This letter describes the units that support the transmission of the disparity stream.

The disparity stream comprises of disparity maps, produced after processing two synchronous video streams taken by two cameras. The disparity estimator, which generates two unidirectional vector fields, produces the disparity stream: a left-to-right (L→R) vector field and a right-to-left (R→L) vector field. The disparity estimator implements a hierarchical block-matching scheme described in [2]. In order to reduce its complexity, the disparity estimator was optimized for head and shoulder scenes with a uniform background. Both vector fields contain approximately the same information and are highly redundant, thus the disparity vectors are transformed to chain maps [2], [3] and compressed by the disparity encoder.
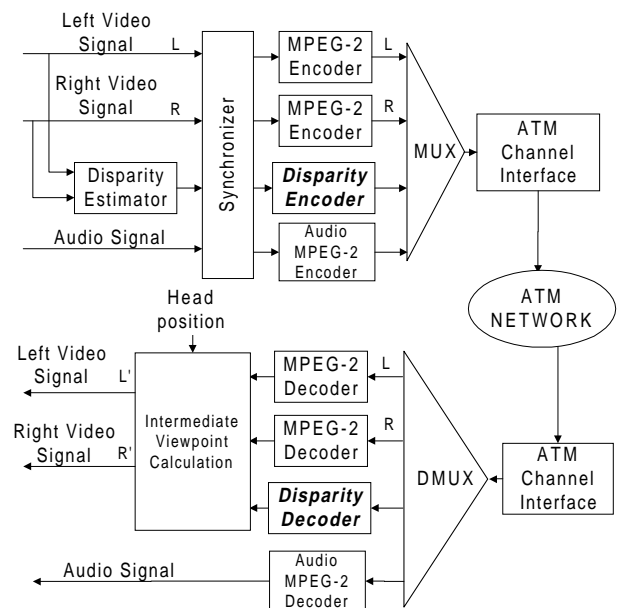

Figure 1: The 3D-transmission chain.

When teleconferencing scenes are involved, vectors from both vector maps are used according to their position. Simulations showed that R→L disparities point more reliably in left occlusions (parts of the scene present only in the left-view image), while R→L disparities produce better image quality in right occlusions (parts of the scene present only in the right-view image) [2]. Therefore vectors from the right-to-left map are used for the left part of the image, and vice versa. For other scenes, one of the two disparity maps is used. On the receiver side, the stereoscopic image is produced by the intermediate viewpoint calculation unit [1], [2] and presented to an "autostereoscopic display". Detailed description of this transformation is given in [3].

The encoder uses lossless encoding techniques, which achieve a smaller compression rate without introducing errors that result to incorrect 3D-impression [4]. Such algorithms have the disadvantage that the bit-rate produced varies according to the statistics of the input data stream. In this letter we describe a simple method of confining this data rate to the allocated channel bit-rate.

2 SYSTEM DESCRIPTION

The function of the encoder unit is to compress the disparity data stream losslessly and transmit it over a constant bit-rate transmission link. After evaluating the performance of various lossless compression algorithms, we decided to use a Lempel-Ziv encoder-decoder [5] and to use a hardware architecture based on re-programmable logic that could also implement other algorithms that could augment the compression efficiency.

In the absence of extensive studies on the effect of disparity map errors on image quality, a method called Controlled Data Loss (CDL) was chosen to bind the output bit-rate to a threshold value during decreased compression. This method is shown graphically in Figure 2. Disparity data are organized into disparity maps that correspond to disparity data generated every 40ms (video frame rate). In CDL, the compression ratio is evaluated on a map-to-map basis. If the size of a compressed map is larger than its original size, then this map is transmitted uncompressed, otherwise the compressed map is transmitted. The ATM CBR channel is divided into fixed-size time slots and each compressed disparity map is allocated to a time slot. In some cases, the time slots cannot carry whole compressed disparity maps, so transmission continues into the next time slot. Meanwhile the next map is not transmitted. At the decoder side, the previously received map replaces a rejected map. In slow moving scenes, like those that occur in teleconferencing applications, spatial characteristics are retained at the expense of some temporal flicker.

The CDL method was chosen since it allows a simple implementation, independent of the compression algo-
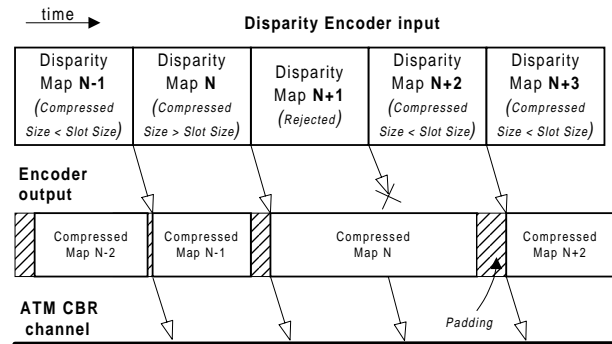

Figure 2: Control of data rate using the Controlled Data Loss (CDL) method.

rithm, and also allows progressive degradation of the depth impression. When map rejections occur, the disparity map prior to the rejected map is reused. All parts of the reconstructed stereoscopic image have valid disparity data, either synchronized with the video data or delayed by 40 msec. In teleconferencing situations, the subjects are not expected to move very suddenly, but even if they do, the errors caused by using a disparity map delayed by 40 ms in relation to the image data are barely noticeable in the interpolated stereoscopic image.

The disparity encoder and decoder units use unidirectional synchronous interfaces for input and output. The input interface of the encoder unit and the output interface of the decoder unit conform to the CCIR (now ITU-R) 601/656 standards [6]. The CCIR 601 standard describes the digital coding of color video signals for systems where the sampling rate for the luminance and the two color-difference signals are related by the ratios 4:4:4 and 4:2:2. The CCIR 656 describes the means of interconnecting digital television equipment complying with 4:2:2 encoding parameters. It supports either an 8-bit parallel interface, which is used in our case, or a bit-serial interface. The standards divide the data streams into logical entities called video data blocks. These data blocks are delineated by two timing reference signals: the start of active video (SAV) and end of active video (EAV) which are used to synchronize the disparity maps with the video data. The interface is clocked at 27 MHz and the bit-rate of the valid disparity map is 5.184 Mbps.

The output interface of the encoder unit and the input interface of the decoder unit use a synchronous interface to transfer the compressed disparity bit-stream to/from an ATM multiplexing/demultiplexing unit. The interface is clocked at 13.5 MHz derived from the CCIR 601/656 interface. The output data are organized into equally spaced, fixed-size blocks, called ATM blocks. Each block occupies a time slot. Each block contains a header with information on the size of its payload and its status. The header is protected using a CRC byte, by also implementing a single bit error correction function [8]. Padding bytes are

used to keep the size of the block constant. The CBR channel allocated for the disparity data has a fixed bit-rate of between 2.7 and 4 Mbps, which is determined during the call set-up. In all cases, the rate of the ATM blocks is set to 25 blocks/sec, thus having 40-msec inter-block time. Constant delay between video and disparity streams is achieved by bounding the delay introduced by the video and disparity encoder/decoder units and multiplexing the data into one ATM CBR transport stream.

3   SYSTEM ANALYSIS

After outlining the requirements of the disparity encoder and decoder units, we proceeded in decomposing the functionality into various modules, as they are shown in Figure 3. The disparity encoder unit is decomposed into the Disparity Data Reception Module, the Compression Module and the Data Framer. The Disparity Data Reception Module (DDRM) filters the disparity data from the CCIR 601/656 standard input bit-stream and transfers the data to the compression module. The Compression Module (CM) undertakes the compression of the disparity data using the lossless algorithm. The Data Framer (DFr) implements the mapping of the variable bit-rate compressor output to the constant bit-rate transmission channel.

The disparity decoder unit is decomposed into the Data Deframer, the Decompression Module and the Disparity Data Transmission Module. The Data Deframer (DDFr) extracts the compressed data and transfers them to the decompression module. The Decompression Module (DM) decompresses and produces the original disparity data, while the Disparity Data Transmission Module (DDTM) encapsulates the data as a CCIR 601/656 standard bit-stream.

In order to outline the logical relationships between the functions that comprise the encoder and decoder units, we used the object-modeling-technique (OMT) [8]. We derived a static description of the system, as shown in Figure 4. Both units share a data-path architecture where the data passes through the various logical objects with a microprocessor having the operational overview of each unit. These object models form the blueprint for the hardware architecture. Some logical objects correspond directly to hardware devices, such as the microprocessor, while others describe the required functionality.

In more detail, the CCIR interface hardware, in conjunction with the FIFO, enables the transfer of data between the CCIR interface and the compression/decompression engine. The microprocessor carries out control functions and allows user access to various parameters and status information.

In the encoder, the transfer of each map to the input interface takes place during the time between the CCIR end-of-active-video (EAV) messages. The CCIR interface uses the EAV message to produce a synchronization signal for the data framer and passes the disparity map
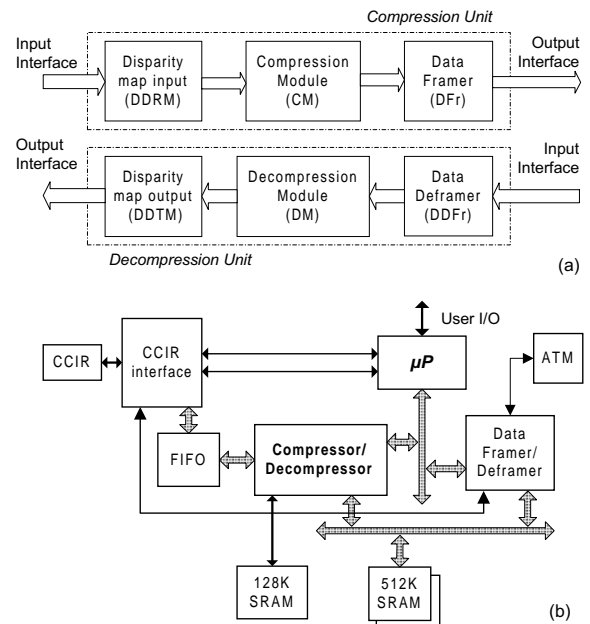

Figure 3: The flow of data through the compression and decompression units and the resulting hardware architecture.

data to the compression engine. The microprocessor measures the compression ratio and transfers this information to the CDL data framer.

The CDL data framer controls the flow of disparity data to the ATM channel according to the compression rate. It encapsulates the disparity data into ATM blocks and controls the output timing of these blocks according to the programmed delay. The CDL logic can be in either of two states, an uncompressed map transmission state and a compressed map transmission state. Disparity maps are compressed and stored temporarily. The compression rate, in conjunction with the state of the CDL logic at the previous ATM block transmission, controls the type of transmission for the current time slot. Compressed disparity maps, which do not fit in a time slot, are fragmented between two ATM blocks in order to bound the disparity data rate and to reduce the artifacts in the reconstructed 3D image.

The aim of the decoder unit is to reverse the compression process and produce the initial disparity map. When the CDL data deframer inputs a block from the ATM demultiplexer, it decodes the header and stores the payload in the buffer.

The header provides information on the status of the payload, i.e., if it is compressed or uncompressed and if it contains a whole disparity map or part of it. If the payload contains compressed data, these are decompressed and stored in a FIFO. After a predefined delay, the output interface retrieves the unformatted disparity map and outputs the original disparity map encapsulated in a CCIR 601/656 video data block. If the map was fragmented between two ATM blocks, it is reassembled. An auto-
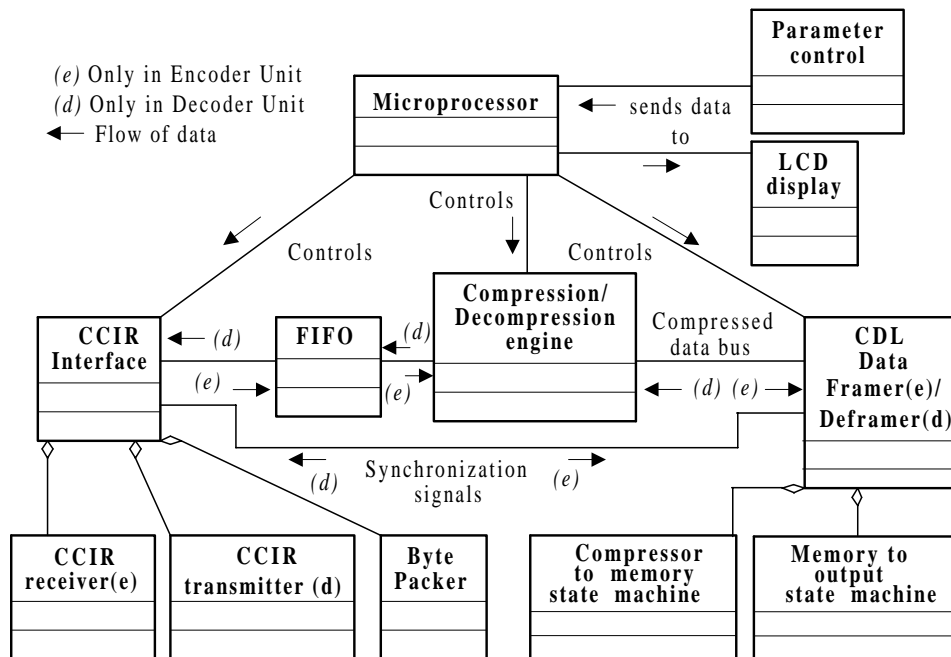
Figure 4. Object model diagrams of encoder and decoder units

matic delay control module is used so that the synchronizer is constantly fed with uncompressed disparity maps. The CDL data deframer synchronizes the CCIR interface at constant intervals. The introduced delay is set to more than 80ms, due to the fact that an uncompressed map may occupy two incoming slots and thus requires 80 ms to reassemble.

The decoder includes two error detection mechanisms: header error detection and payload data error detection. The data deframer checks for errors in the header while the decompression engine checks for erroneous code words. Errors in the received bit-stream do not allow the successful decoding of the compressed information. This situation is handled by resetting the compression process periodically. During the study of the behavior of the compression algorithm on test data, we found that resetting the dictionary of the LZ algorithm on every map boundary does not degrade the compression rate significantly. This allows us to restart the compression process at the beginning of each map and thus decode a map without using the dictionary altered by the previous one.

4 EXPERIMENTAL RESULTS

The final system was designed in order to take advantage of available compression solutions with the flexibility of reprogrammability in both hardware and software. Thus keeping in line with the various simulations on existing disparity maps we decided to use an ASIC compressor implementing the high speed lossless dictionary-based (LZ78) algorithm with a 2.5 Mbytes/sec maximum

throughput, augmented by FPGAs. Compression rates on test sequences ranged from 45% to 75%. This translates to effective bit-rates in the range of 1.5 Mbps to 3 Mbps. The various state machines have been implemented using a combination of FPGAs, FIFOs and SRAMs. The block diagram of the encoder and decoder units is shown in Figure 3. The delay introduced is programmable and ranges between 170 msec and 1 sec, both units inclusive. This high delay is required for delay matching with commercially available MPEG2 codecs.

We have built and tested two fully operational units, an encoder unit and a decoder unit, for the needs of the EU ACTS AC92 PANORAMA project. Preliminary tests were carried out in the beginning of 1998 in Berlin, Germany, on the system chain excluding the ATM connection equipment, the video codecs and the input cameras, while the system chain was successfully demonstrated in October 1998 in Berlin, Germany in a teleconferencing setup. During these tests, the encoder achieved a compression rate of between 45% and 70%, so a 3 Mbps channel was sufficient for the lossless transmission of the disparity data. When less channel capacity was allocated (2.7 Mbps), the CDL mechanism didn't introduce noticeable errors.

5 CONCLUSIONS

In this letter we have presented the design and architecture of lossless compression and decompression units, capable of compressing disparity data for videoconferencing applications. The encoder unit is capable of com-

pressing the disparity data losslessly, where possible, permitting efficient use of an AAL1 ATM transmission channel by using the Controlled Data Loss (CDL) methodology. Data reduction is achieved without sacrificing the stability of the introduced delay, an extremely important parameter when synchronization with video bit-streams is required. The introduced delay is also programmable so that it can match the delay introduced by MPEG2 video codecs. The decoder unit reproduces the original signal while retaining the delay stability. Experimental results showed that by using a lossless compression algorithm along with the CDL method would achieve efficient disparity data transmission over limited bandwidth channels, without introducing noticeable errors.

Manuscript received on 9/7/1999

REFERENCES

[1] EU ACTS Project AC092 PANORAMA. Package for New Operational Autostereoscopic Multiview Systems and Applications, Technical Annex, 1995.

[2] J-R. Ohm, K. Grueneberg, E. Hendriks, E. Izquierdo M., D. Kalivas, M. Karl, D. Papadimatos, A. Redert. A real-time hardware system for stereoscopic videoconferencing with viewpoint adaptation. In *Image Communication, special issue on 3D technology*, Jan. 1998.

[3] Andre Redert, Emile Hendriks. Disparity map coding for 3D teleconferencing applications. In *Proceedings of the SPIE conference on Visual Communications and Image Processing (VCIP)*, Pages 369-379, 1997.

[4] M. Ziegler. Region-Based Analysis and Coding of Stereoscopic Video, Ph. D thesis, Akademischer Verlag Muenchen, Germany, 1997.

[5] D. Papadimatos, T. Antonakopoulos, V. Makios. Real-Time Disparity Information Compression in 3D Teleconferencing Systems. In *EUROMICRO Conference*, August 1998.

[6] CCIR Recommendation 656. Interface for Digital Component Video Signals in 525-Line and 625-Line Television Systems, 1990.

[7] A. Maniatopoulos, T. Antonakopoulos and V. Makios. Implementation Issues of the ATM cell delineation mechanism. In *Electronics Letters*, IEE, Vol. 32, No 1, Pages 963-964, May 1996.

[8] J. Rumbaugh, et al. Object-oriented Modeling and Design, Prentice Hall, 1991.