Performance Analysis of a Selective Rejection Algorithm on Compressed Disparity Data

D. Papadimatos, T. Antonakopoulos and V. Makios

The 7th International Conference on Advances in Communications and Control – COMCON'7

ATHENS, JUNE 1999

Performance Analysis of a Selective Rejection Algorithm on Compressed Disparity Data

D.Papadimatos, T. Antonakopoulos, V.Makios

Department of Electrical Engineering and Computers Technology,

University of Patras,

26500 Patras, GREECE

Abstract : *This paper describes a novel approach towards achieving efficient transfer of video disparity data, i.e. data conveying depth information of a video scene, over a constant bit-rate ATM channel. Disparity data are produced from the processing of video data from two or more video cameras and are used in conjunction with the video bit-streams to generate intermediate views on an autostereoscopic monitor. Errors in the coded disparity bit-stream cause large synthesis errors in the recreated image even when the lossy algorithms are configured for high quality mode. Thus a lossless compression algorithm has to be used for compressing the disparity information. In our approach, we use a simple but robust method of controlling the output bit-rate of the lossless compression algorithm while reducing the degradation of the recreated autostereoscopic scene. This method minimizes the number of rejected frames and spreads the introduced 'noise' so that it remains visually undetected.*

I. Introduction

Disparity information represents the optical information which must be presented to both eyes in order to take advantage of the stereoscopic nature of human vision. The human visual system uses various cues to estimate the depth of a view. One of them is the ability of both eyes to present a slightly different view to the brain. The brain processes the two images, taking into account other visual cues, like motion parallax etc. and generates the impression of depth. On 2-D displays this impression is missing, so a new generation of systems aims to reproduce this natural impression. The goal of a disparity-based system is to transfer this perception of depth remotely, by using the information generated by two cameras and by exploiting specially designed displays which can present a suitable view to each eye.

In this work, we will consider systems aimed at teleconferencing applications, where disparity information is generated by processing the images from two cameras viewing the same scene. In the ACTS AC092 PANORAMA project [1] such a system was successfully developed and demonstrated. This system has the ability to modify the 3D impression according to the position of the viewer by synthesizing 3D views from the two video bit-streams, the disparity bit-stream and the position of the viewer. Disparity information is generated on the transmitter side in order to reduce the complexity of

---

This work has been performed in the framework of the European Commission ACTS Project AC092 PANORAMA -“Package for New Operational Autostereoscopic Multiview Systems and Applications”

the receiver but at the expense of transmitting two video bit-streams, a disparity bit-stream and accompanying audio. The capacity required for the transmission of two video bit-streams and the disparity bit-stream is such that compression is required for reducing the required bandwidth. Lossy image compression algorithms such as MPEG2 can be used for the video but due to the synthesis of intermediate views the disparity information must be compressed with redundancy reduction methods. Lossy coding can be applied only as a last resort to reduce peak compressed bit-rate.

This paper describes a novel approach towards achieving efficient transfer of disparity data. In section II we give a brief description of the hardware system and the constraints involved. In section III we briefly describe the lossless compression algorithm we used and the ramifications to the transmission of the disparity data over a CBR channel and we present our approach to achieve efficient transmission. In section IV we present the proposed Controlled Data Loss method and describe its functionality. In Section V we present the model of CDL and some variations which improve throughput while trading off delay stability. The last section presents the results achieved in the experimental prototype.

II. System Overview

The lossless compression unit, called disparity encoder on the transmitting side and disparity decoder on the receiving side, is part of a hardware chain that transmits 3DTV to an autostereoscopic display through an ATM network [2]. This compression unit is used for efficient transmission of the system disparity fields. The disparity encoder derives its input from the disparity estimator through a synchronizer and outputs the compressed command bit stream to an ATM multiplexer. The data rate at the input of the command converter is 5.184 Mbps with the data being encapsulated in a CCIR 601/656 format. The disparity decoder accepts the compressed disparity stream from an ATM demultiplexer and recreates the original disparity stream by encapsulating the data in CCIR 601/656 format. The decompression unit outputs the data to the interpolator, synchronized frame per frame with the MPEG decoders. The main functional parts of the disparity encoder/decoder are the following:

1. The *Disparity Data Reception Module* (DDRM). This module filters the disparity data from the CCIR 601/656 standard input bitstream and transfers the data to the compression module.
2. The *Compression Module* (CM). This module undertakes the compression of the disparity data using the lossless algorithm.
3. The *Data Framer* (DFr). This module implements the mapping of the variable bit-rate compressor output to the constant bitrate transmission channel.
4. The *Data Deframer* (DDFr). This module extracts the compressed data and transfers them to the decompression module.
5. The *Decompression Module* (DM).This module decompresses and produces the original disparity data.
6. The *Disparity Data Transmission Module* (DDTM). This module accepts the uncompressed data, adds the appropriate format bytes so that the original CCIR 601/656 format is generated and outputs the original CCIR 601/656 frame.

The disparity encoder unit and decoder units use unidirectional synchronous interfaces for I/O. 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. The CCIR601/656 standard divides the data streams into logical entities called video data blocks, which correspond to 40 msec of video data. These data blocks are delineated by two timing reference signals: the start of active video (SAV) and the end of active video (EAV). Each video data blocks is further subdivided into two fields. Each video data block consists of 625 scanlines. The interface is clocked at 27 MHz but the disparity information data rate equals 5.184 Mbps, so we use every fourth scanline to send the disparity information. For the rest of the paper we will refer to uncompressed disparity data corresponding to 40 ms of video data and encapsulated in a CCIR 601/656 video data block as the *disparity map*.

The video data blocks represent a large amount of data, which in turn requires a large amount of buffering hardware. To reduce complexity at the receiver side the bit-streams must arrive virtually synchronized at the interpolator input. By ensuing synchronization at the encoding hardware, the ATM interfaces and the decoding hardware, the synchronization at the interpolator is achieved with less effort. Thus the disparity encoder-decoder modules had the task to introduce constant delay between interfaces. Synchronization was achieved on video data block boundaries i.e. 40ms, so the synchronization mechanisms didn’t have to deal with the unpredictable introduced delay of the encoder-decoder hardware. The delay between the start of the incoming disparity frame and the output of the corresponding output frame was kept constant.

In order to build the proper command encoder/decoder, we initially faced the problem of selecting the best compression algorithm for that specific application. The algorithm selection criteria were the achievement of good compression rate and the simplicity of its implementation.

III. Lossless Compression Algorithms and the Need for Rate Control

The general aim of compression algorithms is the generation of an alternative representation of a digital source, using a smaller amount of data. Information and coding theory provide a lower bound on the reduction of the necessary bits needed by the representation of the original source. This lower bound can be calculated by using the entropy of the digital source. Lossless algorithms achieve optimum compression by trying to produce representations of the digital source which require a number of bits analogous to the entropy or information content of the source. Arbitrary digital sources can not be defined exactly but various models can be used to predict their behavior. These models are based on information derived from these sources by various means. Appropriate representations are then produced through the utilization of these models. The corresponding representations do not contain information already included in the model. These models change in response to the change of the source behavior thus making the algorithms adaptive.

The above comments have important ramifications in the implementation of a lossless algorithm:

- the compressed representation of a source can vary in size according to the information content of that source, and

- there is no guarantee on the size of the compressed representation if the source exhibits a behavior which the model cannot describe. The bit-rate of a compressed bit-stream may vary in time and thus is characterized as variable-bit-rate (VBR).

On the other hand, the ATM multiplexer allocates constant bit-rate (CBR) for the transmission of the compressed disparity data, so during the degradation of the compression ratio the bit-rate must be reduced, while affecting the 3D impression as little as possible.

In [3], we found that Lempel-Ziv algorithms achieve a good tradeoff between compression efficiency and implementation simplicity. We also found that by resetting the compression process every 40 ms we actually increased compression efficiency. This also increased the robustness of the transmission, since the retrieval of the original data from compressed data, which have been corrupted, is very difficult.

In studies concerning the synthesis of 3D views, it has been found that the human visual system can adapt to slow moving changes in depth information [4]. Since our application concerns two-way teleconferencing applications, the general movement of the objects is limited. A method called Controlled Data Loss (CDL) was developed to limit the output bit-rate to a threshold value during decreased compression efficiency.

In CDL, the compression ratio is evaluated on a map-to-map basis. The ATM CBR output channel is divided into fixed-sized slots. Each compressed disparity map is allocated a slot. If a compressed map size is larger than its original size, then this map is transmitted uncompressed. Since the output slots cannot carry whole uncompressed disparity maps, transmission continues into the next slot. In that case, the next map is not transmitted. Both cases are shown in Figures 1 and 2. At the decoder side, the previously received frame replaces a rejected map. In slow moving scenes, like those, which occur in teleconferencing applications, spatial characteristics are retained at the expense of some temporal flicker. 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 image data or delayed by 40 ms. In teleconferencing situations, the subjects are not expected to move 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 above scheme was chosen since it achieved the most robust synchronization with the other data streams, allowing a simple implementation independent of compression algorithm and a progressive degradation of the whole image. The CDL scheme is an integral part of the data framer. In the following section we present the functionality of CDL and outline its implementation with the use of the Specification and Description Language (SDL) and Message Sequence Charts (MSCs).

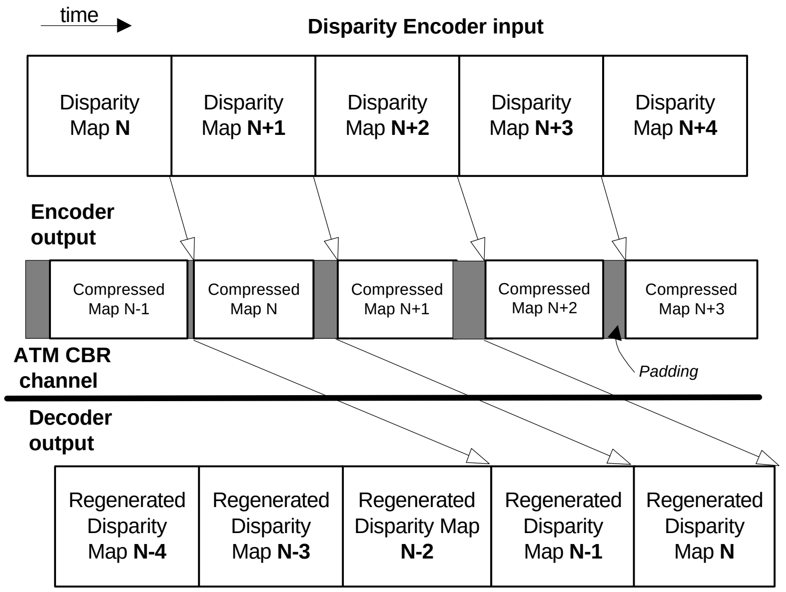

Figure 1. CDL operation while the encoder output data rate is smaller than the available ATM data rate

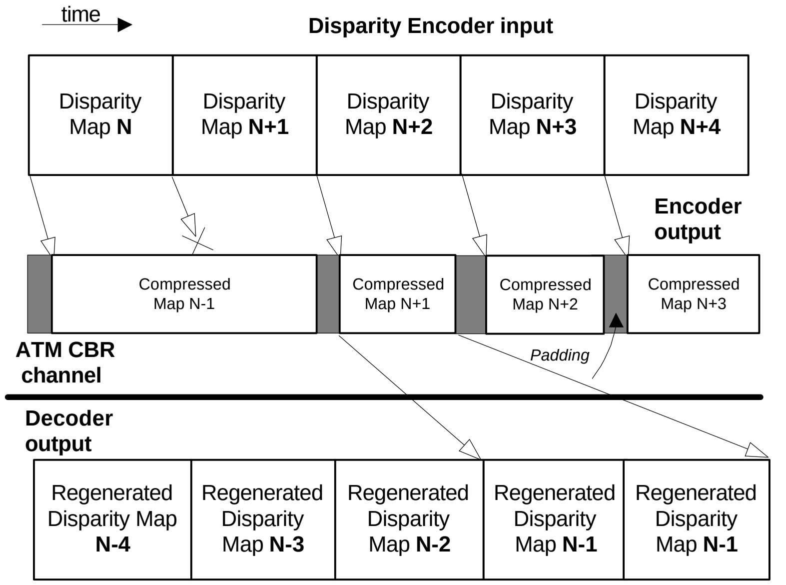

Figure 2. CDL operation during compression rate degradation

IV. The CDL method and the data framer

The data framer encompasses the functionality for transferring the compressor output disparity maps to the ATM multiplexer by implementing the Controlled Data Loss (CDL) scheme. The task of the data framer is to output each disparity map after a predefined delay. In this context we consider the delay introduced to a disparity map to be the time between the decoding of the end of the CCIR input video data block holding the disparity map and the output of the first byte of the ATM block that contains the corresponding disparity map.

A synchronization signal from the input interface controls the output of data to the ATM multiplexer. Both interfaces are synchronized and the data framer decides which compressed map should be inserted in the output stream. These maps have been stored temporarily in local memory, due to the unpredictable processing time of the compression engine. As noted previously, the capacity of each ATM output slot is much smaller than the size of an uncompressed disparity map. This is where the CDL logic comes into play to reduce the output bit-rate when necessary. The data framer consists of three state machines:

- the compressor-to-memory state machine (C2M SM)
- the memory-to-output state machine (M2O SM)
- the CDL logic (CDL)

As shown in Figure 3, the compressor-to-memory state machine carries out the actual transfer of the compressed data from the compression engine output port to the delay buffer memory. It implements the actual handshaking between the compression engine and the buffer memory and allocates the necessary memory area. The memory-to-output state machine creates the output block, which consists of a header and the compressed data. The CDL logic decides when to transmit or reject a disparity map based on its compressed size.The internal operation of the CDL logic can be described efficiently using SDL (Specification and Description Language) [6], as shown in Figure 4. We consider that the CDL state machine is composed of two states, the Compressed map (CM) and the Uncompressed map (UM).

During start-up, the state machine is in the CM state. When the end of the CCIR input video data block is decoded at the input interface, CM starts a timer. Meanwhile, the compression engine compresses the disparity map and signals the data framer upon completion of the compression. The CDL then initiates the C2M state machine to transfer the compressed map to the buffer. It also stores the compression ratio of each stored map. Upon expiration of the timer for the specific map, the CDL accesses the compression ratio and computes the slot size required to transmit the data. If a slot size is sufficient to carry the compressed map, the CDL signals the M2O state machine to transmit it. An appropriate header is inserted at the start of the map. The compressed disparity map follows the header and pad bytes are inserted at the end in order to keep the size of the slot constant. Upon completion of the transmission the CDL returns to the CM state. On the other hand, if more than two ATM slots are required, the CDL instructs the M2O state machine to transmit the uncompressed disparity map data instead.

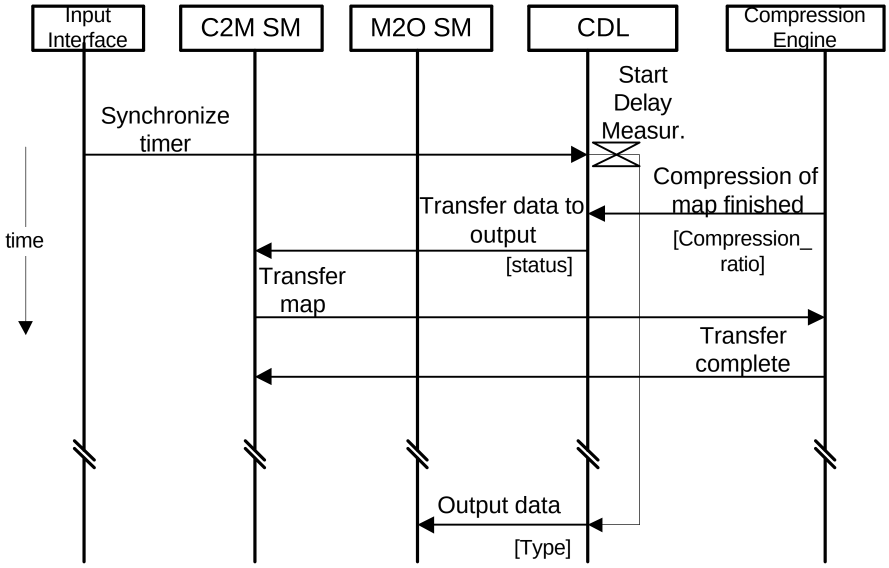

Figure.3. MSC showing the interaction between the processes of the data framer

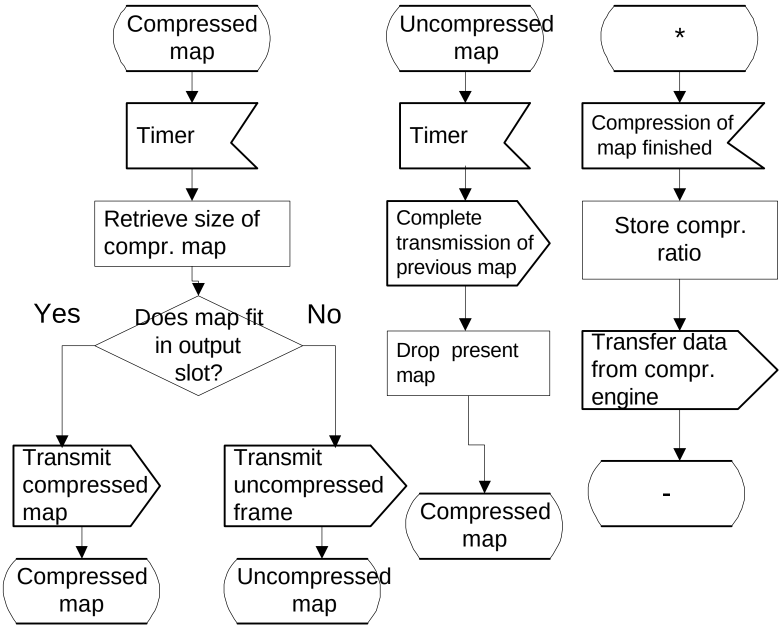

Figure 4. The SDL diagram of the CDL state machine

The ATM slot is filled with the first part of the map while the CDL transitions into the UM state. Upon expiration of the timer for the following map, the remainder of the previous uncompressed map is transmitted in the following ATM slot and the CDL returns to the CM state while the next disparity map is rejected.

V. The model of the CDL method

The metrics of CDL efficiency include the ratio of dropped or rejected disparity maps to the total number of compressed maps and the delay jitter. Both metrics depend on the compression ratio achieved and the allocated channel bit-rate. As noted above, a dropped map is caused when the previously compressed disparity map exceeds its allocated slot size. In pure CDL there is no delay jitter since compressed maps are always transmitted after a constant delay has been added in the disparity encoder unit. The image and disparity bit-streams need to be virtually synchronized i.e. the difference between the start of each data block of the two image bit-streams and the disparity bit-stream must not exceed 40 ms.

The encoder-decoder hardware and the transmission network may introduce delay jitter. We have minimized the delay jitter of the disparity encoder-decoder by synchronizing the output interfaces of both units to the input interfaces and by stabilizing the introduced delay of each module to a user defined value. Thus the transmission network may introduce up to 40 ms of jitter before the system loses synchronization. If the network introduces less jitter, we can tradeoff the delay stability requirement for a reduced rejection rate by allowing the encoder to output the disparity maps using two different delays which do not differ by more than 20 ms. Thus we studied two derivatives of the CDL method where we relax the constant delay requirement for improved throughput.

In essence, we divide the output slot into two mini-slots which have the same capacity as the original slot but are transmitted 20 ms apart. Each disparity map is allocated two mini-slots to be transmitted but can use the following two mini-slots if required. The following disparity map is rejected only if the present disparity map uses both mini-slots allocated to it. In the case where only the first additional mini-slot is used, the disparity map can start transmission in the second mini-slot and continue transmission in the following two mini-slots. Of course if the previous map uses both mini-slots allocated to a disparity map then the next disparity map is rejected. The disparity map that starts transmission in the second mini-slot allocated to it has different delay than the disparity map transmitted in the first mini-slot. In this case, the delay jitter of the network must not exceed 20 ms. We can further divide the original ATM slot into four mini-slots and allow four different delay values that do not differ by more than 10 ms.

We simulated the CDL method and its two derivatives described above considering that the compression ratio follows an exponential distribution. As shown in Figure 5, when the allocated ATM bit-rate is adequate, fewer rejections occur. By allowing more delay jitter, the map rejection rate improves by up to 25 % when the mean compressed map size is not much larger than the capacity of the ATM channel. As the ATM channel gets significantly overloaded, the ratio of the dropped frames converges in all three cases.

Unfortunately the delay jitter increases the likelihood that the disparity data may lose synchronization with the other bit-streams. The difference between the arrival times of the disparity maps and the corresponding image data blocks at the interpolator must not exceed 40 ms. By allowing different delays in the disparity encoder there is an increased chance that this difference is exceeded and synchronization is lost, as it is shown in Figure 6.

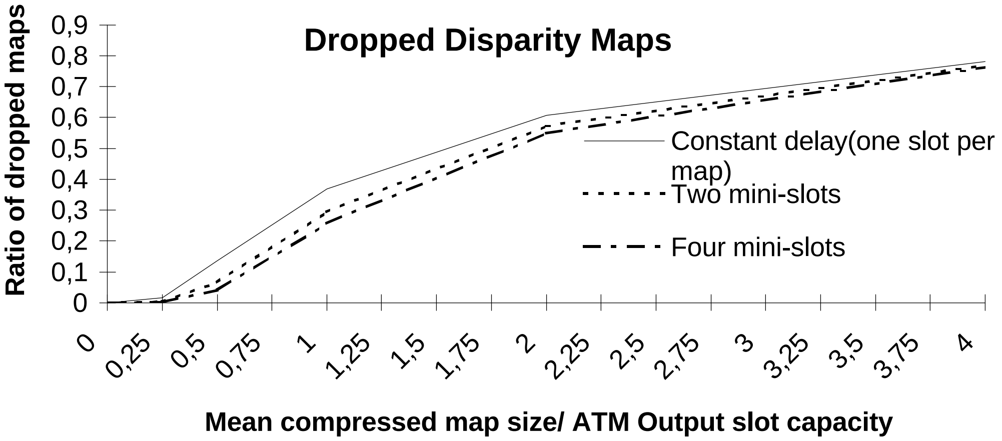

Figure 5. Ratio of dropped maps to mean compressed disparity map size

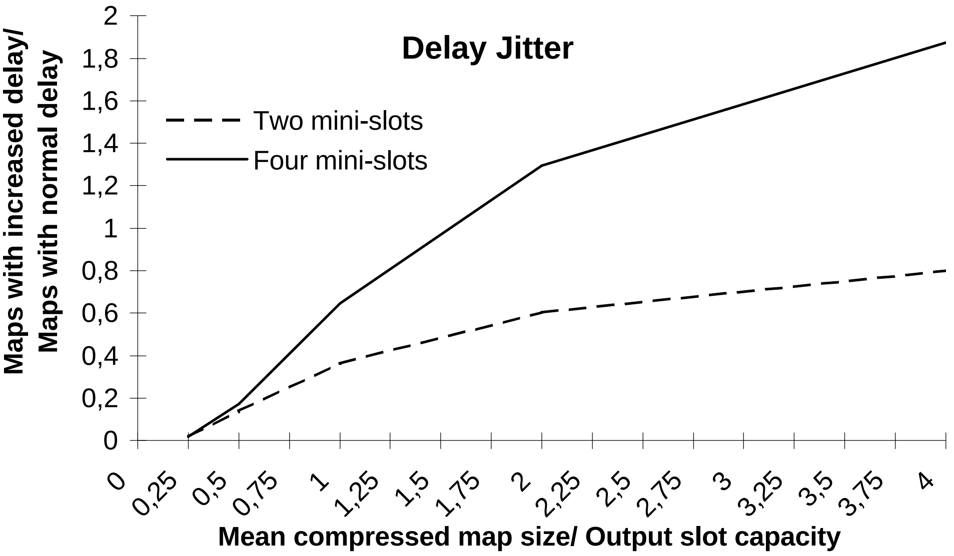

Figure 6. Delay jitter for CDL variants using two and four mini-slots for disparity map transmission

In pure CDL, the encoder introduces constant delay so no delay jitter exists. The two variants allow the encoder to use two or four different delay values. When there is ample ATM channel capacity, most disparity maps are not delayed more than the programmed delay. As the mean compressed map size reaches and exceeds the ATM channel capacity an ever-increasing number of disparity maps has increased delay.

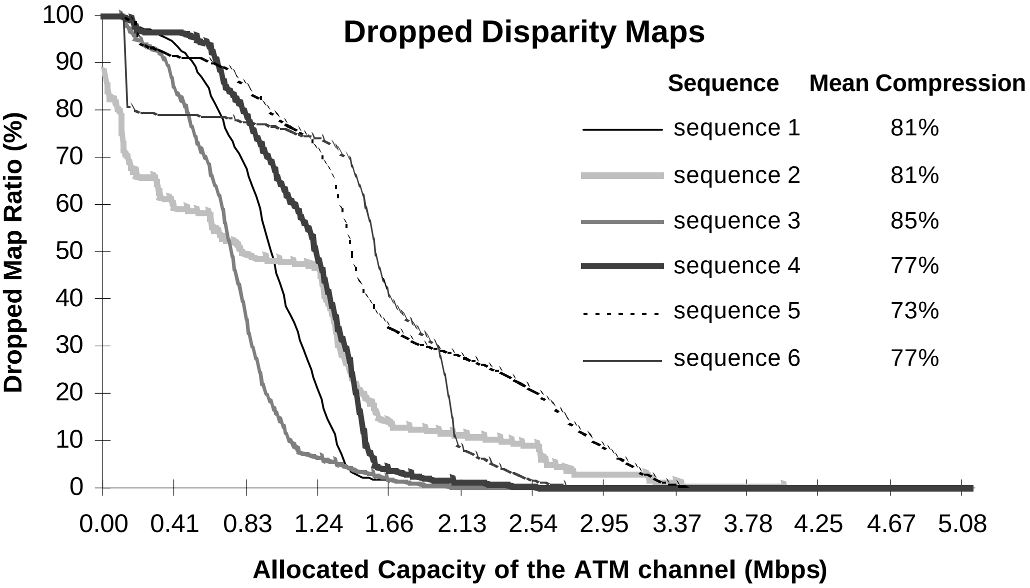

Figure 7. Performance of CDL on real-life disparity data

VI. Experimental results

During the PANORAMA project we gathered test data from typical teleconferencing scenes. We show results from six such sequences in Figure 7. Each disparity map sequence lasts 20 minutes. We tested various ATM channel bit-rates and found that the image quality was very good when data rate between 2 Mbps to 3 Mbps was allocated.

In Figure 7 we have plotted the dropped map rate of these sequences using the pure CDL method with respect to various ATM channel data rates. If the ATM channel bit-rate is half the required uncompressed disparity map bit-rate (circa 2.7 Mbps), in all sequences except the sixth over 90 % of the disparity maps are able to reach the receiver.

VII. Conclusions

In this paper we have presented the CDL method of transmitting compressed disparity information over ATM CBR channels. This method achieves a good balance between the quality of the interpolated stereoscopic image, the resilience to loss of synchronization between the transmitted disparity and video bit-streams and system complexity. Three variations of the method have been studied, which tradeoff the delay jitter requirement and achieved throughput and allow the method to be adapted according to the characteristics of the ATM CBR connection. This method was implemented in the disparity encoder and decoder units for the AC092 PANORAMA project, which was successfully presented in October 1998 and offered satisfactory 3D image quality.

References

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

[2] J.R. Ohm, K. Grueneberg, E. Izquirdo, M. Karl, E.A. Hendriks, P.A. Redert, D. Kalivas, and D. Papadimatos. “A real-time hardware system for stereoscopic videoconferencing with viewpoint adaptation”, *Proceedings of the International Workshop on Synthetic-Natural Hybrid Coding and Three Dimensional Imaging, (IWSNHC3DI'97)*, pages 147-150, Rhodes, Greece, 1997.

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

[4] Colin Ware: “Dynamic Stereo Displays” *CHI 95 Conference Proceedings*, pages 310-316, Denver, Colorado, 1995.

[5] ITU Recommendation Z.120. Message Sequence Charts (MSC), 1996.

[6] ITU Recommendation Z.100. Specification and Description Language (SDL), 1994.