CCITT SGXV
Working Party XV/1
Specialist Group on Coding for Visual Telephony

Document # 61
January, 1986

SOURCE : NTT, KDD, NEC and FUJITSU

TITLE : DCT-BASED CODING ALGORITHM

---

1. Introduction

This document describes the DCT-based coding algorithm for motion-compensated interframe prediction errors.

Outline of the overall structure of the codec is described in Document # 60. This document describes the coding algorithms for C and F components mentioned in Document # 60. An example of simulation results is also presented in Annex.

2. Orthogonal Transform

Motion-compensated interframe prediction errors are segmented into blocks. After significant/insignificant detection of each block, significant blocks are two-dimensional discrete cosine transformed with N x M pels. N and M are both eight at present.

3. C-component and F-component in DCT-based Coding Algorithm

Motion-compensated interframe prediction errors contain components which have relatively small variations and components which have sharp variations in magnitude. These two types of components usually exist in the same significant block according to the performance of motion-compensated prediction. Thus, significant coefficients are produced in higher orders as well as in lower orders of the DCT domain. This makes the coding problem of motion-compensated interframe prediction errors difficult.

In order to perform efficient coding, we consider motion-compensated interframe prediction errors as a combination of the following two types of components :

- C-components ... components for which DCT coding works well,
- F-components ... components for which DCT coding does not work well.

Figure 1 shows the configuration of a DCT-based codec.

A separation filter is first applied to the motion-compensated interframe prediction errors in order to remove components having sharp magnitude variations. This causes reduction of the number of significant DCT coefficients in the higher order domain. One example of the separation filter detects an element which has a sharp magnitude variation such as a pulsive or edge-like magnitude variation. The motion-compensated interframe prediction error of such an element is replaced by the median value of prediction errors in neighboring area. The performance and effectiveness of the separation filter depends on to what extent components having sharp magnitude changes exist in motion-compensated prediction errors, and on how DCT coefficients are quantized. In a variation algorithm, little attention is payed to the usage of the separation filter.

The C-components, that is, the outputs of the separation filter are two-dimensional discrete cosine transformed. Three kinds of adaptive coding methods for DCT coefficients have been studied to improve the quality of decoded pictures and coding efficiency. They are described in section 4.

Difference signals between motion-compensated interframe prediction errors and decoded outputs of DCT coding are treated as F-components. When the separation filter is applied to motion-compensated interframe prediction errors so as to remove components having sharp magnitude changes, the F-components correspond to those removed. Furthermore, F-components contain the error signals caused by the quantization process for DCT coefficients.

The ratio of C-components and F-components in a significant block depends on factors such as the characteristics of input images, thus characteristics of motion-compensated interframe prediction errors (e.g. amount of components having sharp magnitude changes in significant blocks), how well the coded results of C-components represent motion-compensated interframe prediction errors, the characteristics of the separation filter and so on.

The appropriate balance of bit allocation for C-components and F-components requires further study.

4. Coding of C-component

The following adaptive coding algorithms have been studied to quantize DCT coefficients. These can be assembled to improve the quality of decoded pictures and coding efficiency.

(1) Adaptive level coding

Quantization characteristics such as dead zone and step size are adaptively controlled depending on movement in the picture and the power of prediction errors. The number of quantizing levels for each order is also adaptively assigned. The quantizing indices of significant pels are Huffman coded. The position of significant pels are run-length coded using the Huffman code. Considering the characteristics of DCT, that is, the concentration of powers into lower order components, the zigzag scan from lower order to higher order is employed to represent positions of significant pels.

(2) Adaptive zone coding

Input blocks are classified into several classes depending on movement in the picture and the power of prediction errors. Higher order coefficients in the block are selectively discarded. The zone size and zone shape can vary on a block-by-block basis. The quantizing indices of significant pels are Huffman coded. The position of significant pels are run-length coded using the Huffman code.

(3) Adaptive coding using bit-allocation table

Input blocks are classified into several classes by a matching method e.g. using a vector quantization technique. Each classified coefficient block is quantized according to the quantizing table which is defined for each class.

5. Coding of F-component

Since the correlation of each element in the F-components to neighboring elements is normally very low, the scalar quantization (SQ) in the pel region is employed.

The magnitude of significant pels are scalar quantized, and their positions are run-length coded using the Huffman code. In this case, the scanning direction is not fixed because it probably does not affect coding efficiency.

6. Control

The coding and quantizing parameters listed below are controlled to keep the generated quantity of information within an appropriate range.

- quantization characteristics for C-components
- quantization characteristics for F-components
- threshold value for significant/insignificant blocks
- number of omitted frames

(how to treat the omitted frames is for further study).

Both feed-forward and feedback controls are employed. These controls are performed using the buffer memory occupancy, together with other factors such as the ratio of significant blocks in an input frame and so on.

7. Information to be Transmitted

In the above coding algorithm :

- significant/insignificant block identification flag*
- identification flag for combination of DCT and SQ*
- motion vector
- indices for C-components
    - vector quantizing set index*
    - scalar quantizing index
        (magnitude of non-zero elements and run-length of zero elements)
- indices for F-components
    - scalar quantizing index
        (magnitude of non-zero elements and run-length of zero elements)

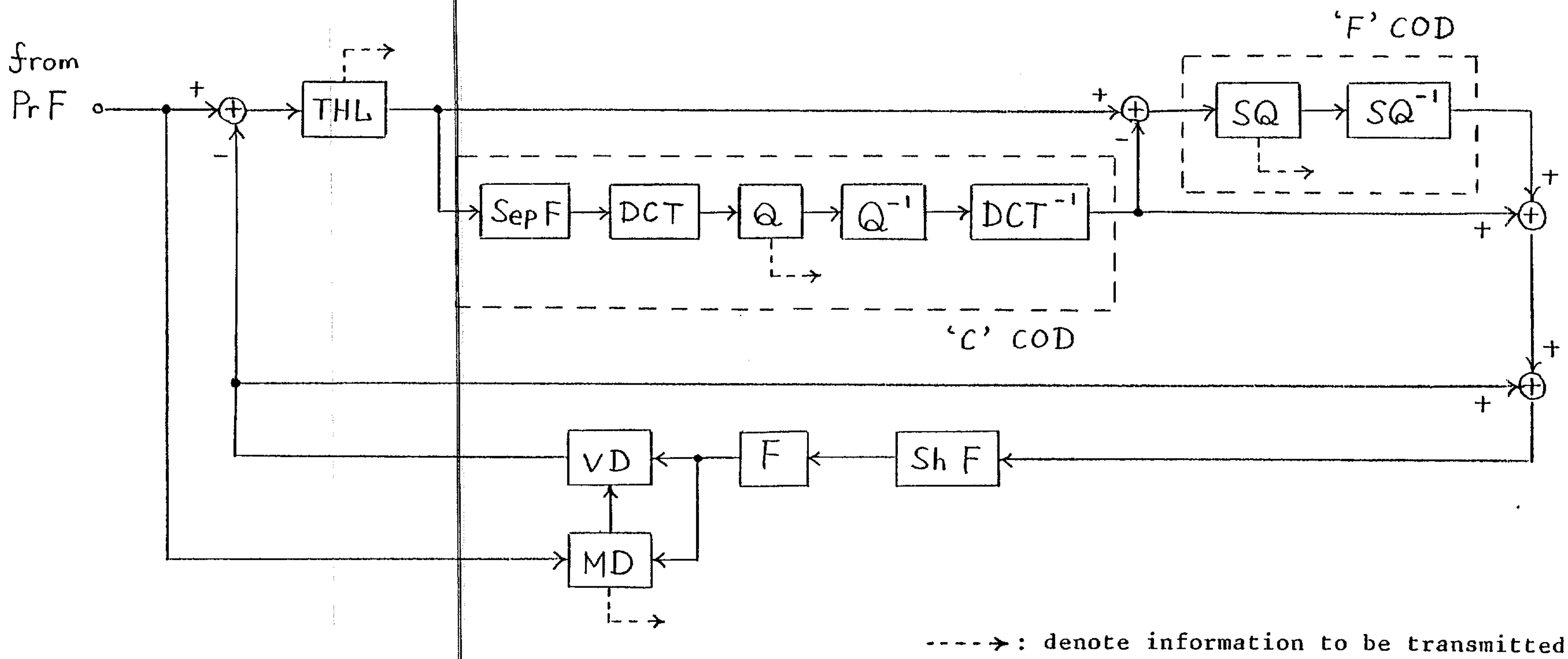
should be transmitted. Mark * represents information transmitted in variation algorithms.

8. Simulation Results

An example of the simulation results is given in Annex. An outline of the experimental coding method, coding parameters and some characteristics such as bit allocation for each transmitted information are also shown. Examples of decoded pictures will be demonstrated in the meeting.

9. Conclusion

Current results in the study of the DCT-based coding algorithm are described in this document. The most important problem is how to code C-components efficiently. Further studies are required to improve the quantization methods for DCT coefficients.



PrF	: Pre-filter
MD	: Motion Detector
F	: Frame Memory
THL	: Significant/insignificant Block Detector
'C' COD	: C-component Encoder and Quantizer
'F' COD	: F-component Encoder and Quantizer
ShF	: Shaping Filter
VD	: Variable Delay

Sep F	: Separation Filter
DCT	: Discrete Cosine Transform
DCT^{-1}	: Inverse Discrete Cosine Transform
Q	: Quantizer
Q^{-1}	: Quantization Decoder
SQ	: Scalar Quantizer
SQ^{-1}	: Scalar Quantization Decoder

Fig.1 Block diagram of DCT-based coding algorithm.

ANNEX / Doc.#61

AN EXAMPLE OF SIMULATION RESULT OF DCT-BASED CODING

1. INTRODUCTION

The simulation results demonstrated in the Ipswich meeting is obtained under the generic codec structure described in the Document #60. This annex describes a specific coding algorithm and parameters used in the simulation.

2. Coding algorithm

The total block diagram of the DCT coding algorithm was shown in Figure 1/Doc.#61 which is agreed as a generic structure in Japan. Therefore, the remaining coarse coding algorithm is presented in this section.

Figure A.1 shows the coarse coding part of the simulated coding algorithm. In this algorithm, motion compensated interframe difference signals are first segmented into blocks and the blocks are discrete cosine transformed. Then, each transformed block is classified into plural number of classes using vector quantization technique. Namely the most approximate vector for each input block is chosen by matching to pre-defined vector set based on a minimum power criterion. The vector index is coded and the residua of input coefficients and represented vector are sent to the following scalar quantizer. Finally, the coefficients are quantized according to the pre-defined bit-allocation-table set which is assigned to each vector index.

The following two factors improve coding efficiency and picture quality.

First, VQ technique improve the coding efficiency compared with scalar quantization alone.

Secondly, it provides one of the most appropriate classification methods which adapt to the input signals. Since input signals vary from frame to frame depending on contained movement, adaptive quantization is indispensable in which input signal is classified into plural number of classes based on some classification methods. Among several classification methods vector quantization technique is one of the most appropriate methods, since it classifies the blocks into optimum sub-classes in which the multi-dimension space is divided so as to minimize the overall variance of probability function.

3. Simulation parameters

3.1 Pre-processing

Temporal filter which is described in Figure 3/Doc.#60 is introduced to reduce noise in the signal.

3.2 Threshold to identify significant/insignificant block

In the closed loop operation, the threshold is controlled by the rate of significant block in the previous frame.

3.3 Threshold to identify intraframe/interframe mode

When the power of prediction error in a block exceeds a certain threshold (400.0), input signals are directly discrete cosine transformed (intraframe mode). When otherwise, interframe prediction errors are processed (interframe mode).

3.4 Coarse/fine separation filter
Not used in the simulation.

3.5 Coarse Quantization
The number of VQ index for interframe mode : 1024
The number of VQ index for intraframe mode : 1024
The number of VQ dimensions : 20 (lower 20 coefficients except DC component in a 8 by 8 pel block are vector quantized as shown in Figure A.2.)
The number of Bit-allocation table for interframe mode : 1024
The number of Bit-allocation table for intraframe mode : 1024
Each bit-allocation table is modified to regulate the information rate using scaling factor in closed loop operation.

3.6 Fine Quantization
The difference between original prediction error and output of coarse coder is quantized. The difference which exceeds a threshold is scalar quantized and the addresses are run length coded.
The dead zone and step size are modified depending on pre-estimated information rate (see 3.10).

3.7 Shaping filter
Median filter with cross-shaped window (5 elements) is introduced as a spatial filter.

3.8 Frame dropping
Nominal frame dropping rate 2:1. When a scene-change occurs, 11 frames can be dropped at maxmum. Frame dropping rate is controlled depending on significant block rate in the frame and buffer memory occupancy.

3.9 Buffer memory
The necessary buffer size is at least 10 kbit in this simulation. The total transmission delay is estimated as 140 msec. It is estimated as 50 msec for television standard conversion (1.5 field for each conversion), 30 msec for information estimation which is for closed loop operation, 30 msec for 2:1 frame dropping and 20 msec for buffering.

3.10 Closed loop operation
Pre-estimated information generation rate and buffer memory occupancy activate parameter control.
In advance of scalar quantization, information generation rate in a frame is estimated taking advantage of classified vector index histogram and bit-allocation tables which are defined for vector indexes. Scaling factor of MAX quantizer in the coarse coder mainly control the information generation rate. This factor is derived from estimated information rate and modified by the difference between nominal buffer occupancy and present occupancy.

3.11 Bit Assignment
Following information are coded.
- Significant/insignificant block identification : Fixed word length coded.
- Intraframe/interframe mode identification : Fixed word length coded.

- Moving vector : Variable word length coded in which the code table is modified from Part 3 codec.
- Vector index for coarse quantization : Fixed word length coded.
- Scalar index for coarse quantization : Fixed word length coded for DC coefficient of intraframe mode blocks. Variable word length coded for DC coefficient of interframe mode blocks and AC coefficients for both mode blocks.
- Scalar index for fine quantization : Levels and addresses are variable word length coded.

4. Results of the simulation

Table A.1 shows an example of information rate assigned to each code. Figure A.3 and Figure A.4 show signal to noise ratio and fluctuation of buffer occupancy as the simulation results of test sequence of "Miss America", "Checked Jacket" and "Split-Trevor". Signal to noise ratio in Figure A.3 is derived from original intermediate format signals and decoded signals.

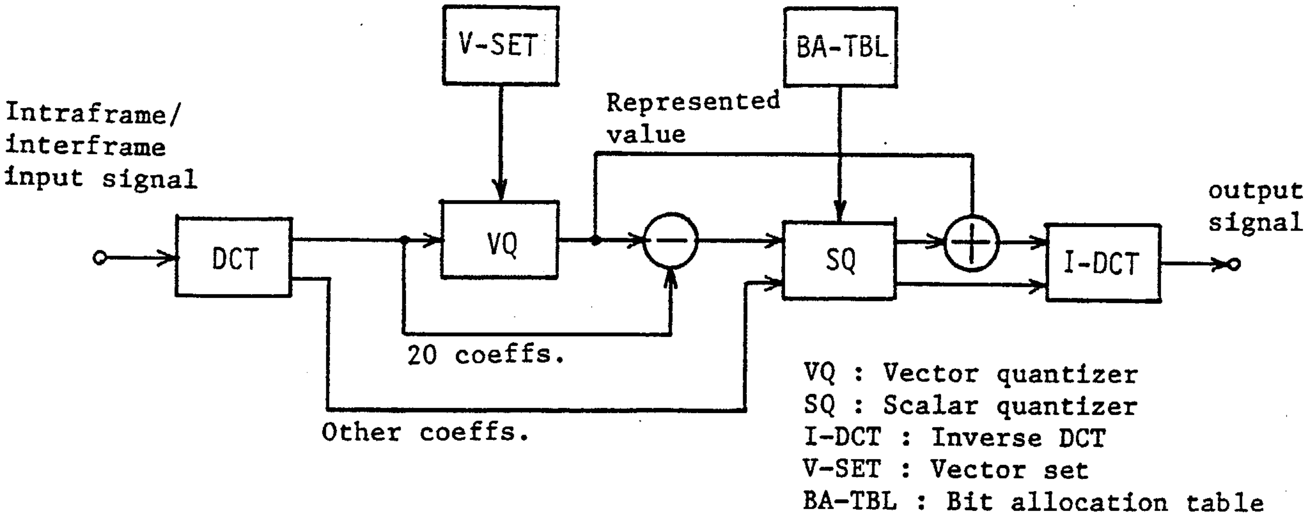
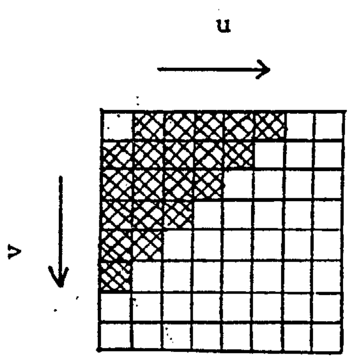


Figure A.1 The coarse coding part block diagram of DCT-based coding algorithm



Meshed zone is vector quantized prior to scalar quantization
White zone is only scalar quantized.

Figure A.2 The area of vector quantization in a coefficient domain

Table A. 1 Information rate assigned to each code

Test Sequence	Miss America	Checked Jacket	Split	Trevor
Frame No.	60	40	30	100
identification of sig/insig and intra/inter	10.6 %	11.0 %	6.7 %	11.4 %
mc vector	20.4 %	10.7 %	24.0 %	27.2 %
vq index	11.4 %	15.1 %	19.5 %	18.5 %
coarse sq	55.5 %	58.7 %	47.1 %	41.0 %
fine sq	2.1 %	4.5 %	2.7 %	1.9 %
Total Bit (kbit)	20.2	20.0	40.0	20.0

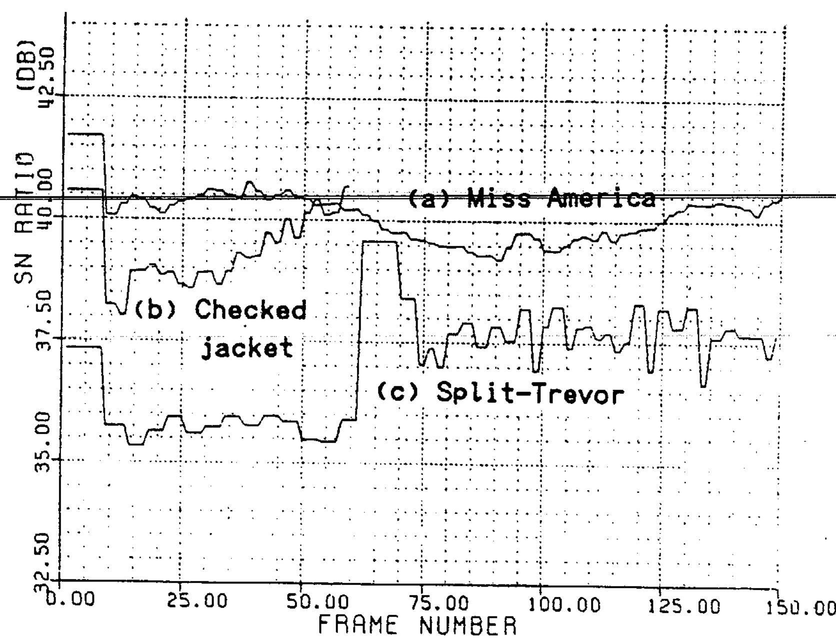


Figure A. 3 Signal to noise ratio of decoded test sequence

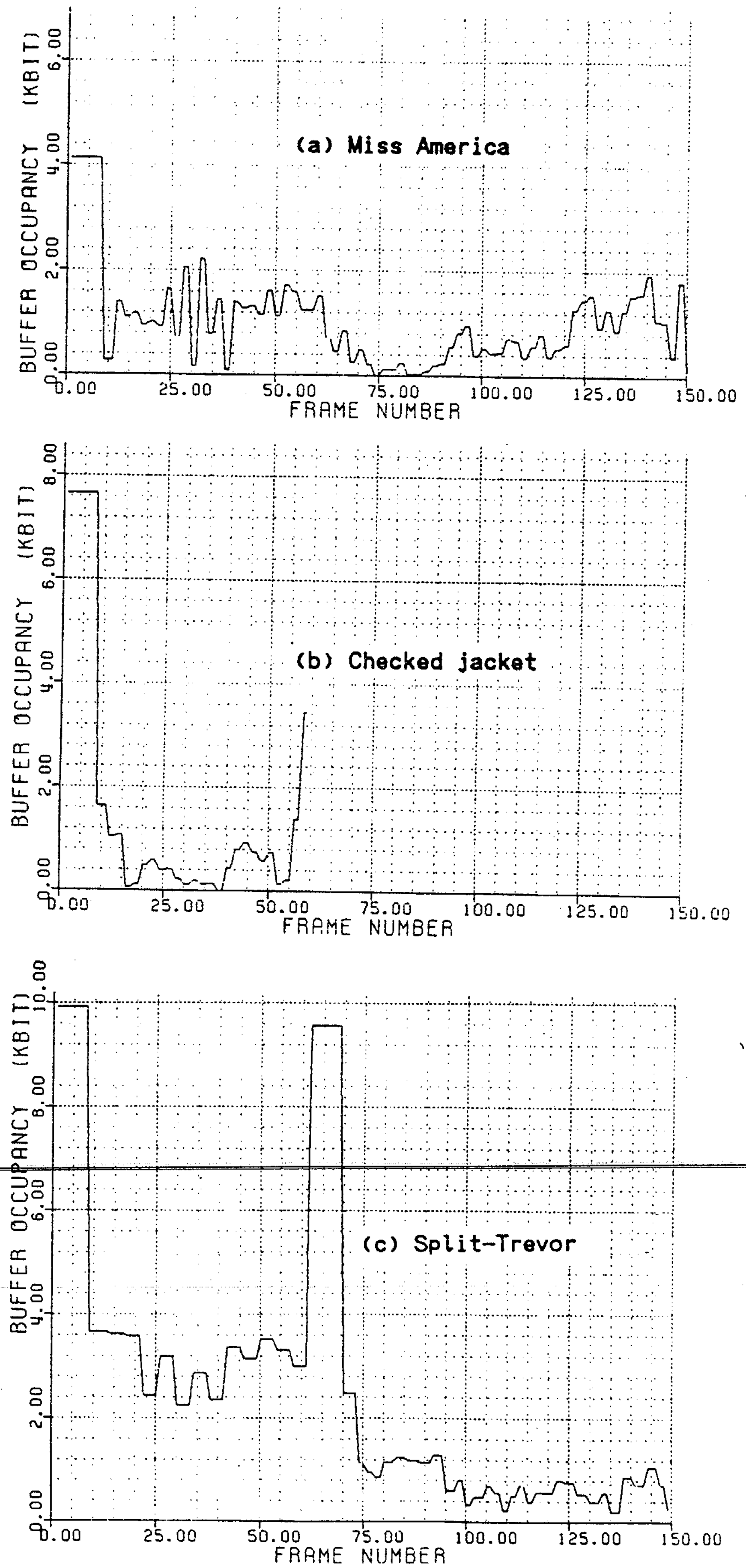


Figure A. 4 Fluctuation of buffer occupancy