CCITT SGXV, WP XV/1
Experts Group on ATM Video Coding
Paris, May 1991

Source: Germany (Daimler-Benz)
Purpose: Information

Title: H.261 Compatible 2-Layer Video Codec with High Cell Loss Resiliance

Summary:

A 1 layer ATM video codec according to CCITT rec. H.261 has some cell loss concealment capabilities. But the loss of data describing e.g. intra coded macroblocks or macroblocks with great displacements will cause in most cases annoying errors. With two layered coding, using the priority bit of the ATM cells, the problem can be solved even at high cell loss rates as 10E-2. This paper proposes the very simple technique of coding one picture twice, the packets of the first pass beeing transmitted with high priority, the packets of the second pass with low priority. The result is a data stream, without packetizing completely compatible whith H.261 and very resiliant to cell loss rates up to 10E-2. The S/N performance is nearly identical to a 1-layer H.261 codec. Except packetizing hard- and software of the decoder is identical to a 1-layer H.261 decoder; the changes in the encoder are the quantizer control for the enhancement layer and the addressing of the pictures to be encoded. Problems with synchronisation of the two layers are avoided. The expense is a slightly higher processing power compared to a 1-layer codec.

Sensible information in the H.261 data stream

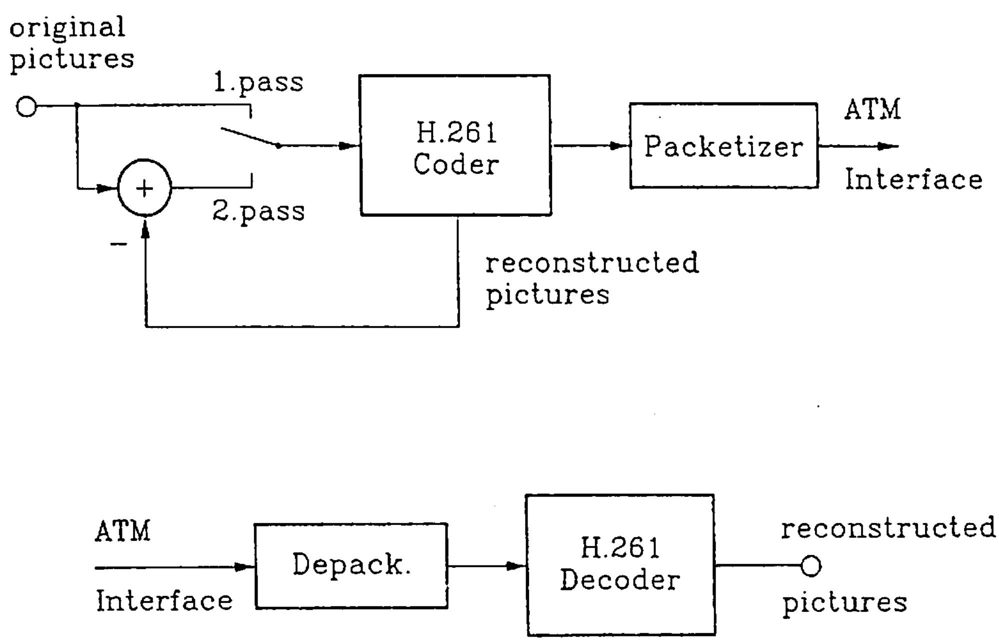
A 1 layer H.261 codec running on an ATM net has inherent cell loss concealment capabilities: normally only one or two out of ten lost cells produces an annoying error on the screen. This can be achieved by interrupting the decoding procedure if a not allowed situation in the decoding tree occurs or just at the point of a lost cell. So the monitoring of obviously senseless picture content is avoided in most cases. The decoding procedure is started again at the next Group of Block Start Code. During the time of interrupted decoding the display of uncorrectly decoded data is substituted by the picture content of the preceeding picture. If the difference is small (this is normally the case in unmoved areas), the visibilty of the degradation is low - the errors are concealed. But in the case of an intra coded macroblock the concealment normally does'nt work. The result can be a very annoying completely wrong picture content (except forced update in unmoved areas). In the case of movement, the errors can be spread over large picture areas.

The techniques described in a lot of other papers and proposed in the ATM Compendium solve this problem by using the priority bit in the cells and splitting off the data stream in a H.261 compatible base layer (with very low allowed celloss rate) and an enhancement layer, where a high (up to 10E-2) cell loss rate can be allowed. In the base layer a relatively good picture is transmitted using all features of H.261. In the enhancement layer the difference between the original picture and the reconstructed base picture is coded again (in time or spectral domain) and transmitted. No use is made of the frame to frame prediction in order to prevent error propagation in the case of lost cells. The enhancement layer raises the resolution of the reconstructed pictures in cases where the base picture, normally produced with fixed rate, is too bad (fast movement, scene cuts). Due to the fact that the frame to frame prediction is switched off for the enhancement layer, the S/N performance of such a 2-layer codec is less than the performance of a 1-layer codec. The data stream in the additional layer is not compatible with H.261 and extra hardware for video mux/demux and other signal processing is needed.

The trick: Coding one picture twice

But there is a simple possibility to make use of the priority bit in order to prevent the loss of sensitive data and simultaneously maintaining fully H.261 compatibility without any extra hardware: Coding one picture twice. In the first pass the actual picture is coded according to H.261 at fixed (or variable) bitrate. Then the same picture is encoded again according to H.261, the picture used for prediction is now the reconstructed picture out of the first pass. The packets out of the first pass are transmitted with high priority, the packets out of the second pass with low priority.

In the second pass no intra coded or moved macroblocks should occur and, in consequence, should be forbidden. The temporal reference transmitted in the picture header is the same for both passes - therefore the decoder knows which picture has to be displayed: the picture after the second pass. Packets containing a picture header or part of it have to be transmitted with high priority.



Structure of the 2-layer 2-pass coder and decoder

The results of the described technique are:

- The data stream (packetizing excluded) is completely compatible to H.261.
- The S/N performance should be nearly identical to a 1-layer H.261 codec. At bitrates in the range of 2 Mbit/s the loss due to the not optimal video multiplex structure for the second pass can be neglected. The simulations even showed a sligthly better performance of the 2 pass coding.
- No soft- or hardware change has to be performed in the decoder or the encoder except the quantizer control for the variable rate enhancement layer and the addressing of the picture to be encoded.
- Problems with cell sequence numbering and layer synchronisation are avoided
- The processing speed is higher: at 30 Hz frame frequency now 60 pictures per second have to be encoded and decoded. The speed factor is less than

two due to the fact that in the second pass no motion estimation and compensation is needed.

- packets with high and low priority are now occuring in bursts; if this is not favourable for the ATM network a rearrangement is possible.

<u>How performs this technique at high cell loss rates?</u>

The techniques mentioned before and studied so far within COST211ter and RACE 1018 HIVITS don't use frame to frame prediction for the enhancement layer in order to avoid possible error accumulation and propagation; if a cell is lost, some fine detail will not be reproduced on the decoder side. The technique proposed here uses frame to frame prediction for the enhancement layer and in principle accumulation of errors and error propagation is now possible. The accumulation is stopped by intra coded macroblocks (forced update). As described in the next chapter, a lot of simulations showed that this accumulation does not reach annoying picture quality degrading levels at expectable cell loss rates:

at CLR=10E-3: no artefacts could be observed, even for experts

at CLR=10E-2: normally no errors are visible; in difficult sequences with movement over a longer period of time (e.g.Tabletennis), experts can see some degradation.

at CLR=10E-1: degradations are clearly visible and start to be annoying

If in the second pass the enhancement signal is not added to the reconstructed picture of the coding resp. decoding loop, the 2 pass coding technique can be considered as a special kind of implementing the method described e.g. in the ATM Compendium, where the enhancement signal is formed out of the difference between original and reconstructed signal in the time domain.

<u>Simulation results</u>

The simulations have been performed with the seqences Salesman, Tabletennis and Swing at cell loss rates of 10E-3, 10E-2 and 10E-1 for the data stream in the enhancement layer. Results for 2 codec types are given here:

a) fixed rate base layer, 640 kbit/s
   fixed rate enhancement layer, 1280 kbit/s (sum rate is 1920 kbit/s)

b) fixed rate base layer, 960 kbit/s

variable rate enhancement layer; 4 Mbit/s peak rate allowed

Forced update by intra coded macroblocks was 3 macroblocks per picture (first pass only). Frame frequency was fixed to 30 Hz. Buffer size and quantizer control for the fixed rate layers has been choosen according to RM8. The variable rate layer had a minimum quantizer step size of 6, maximum of 12. The cell format was 2 octets header, 45 octes (360 bits) video data, followed by 1 octet trailer.

The following table gives some key figures of the simulations; diagrams showing the amount of bits/picture in front of the encoder buffer and the signal to noise ratios are annexed.

Scene	codec type	S/N mean base only	S/N mean base+enh.	mean rate base only	mean rate enhanc.	peak rate base+enh.
Salesman	cbr/cbr	39.01 dB	40.67 dB	0.64 Mbps	1.28 Mbps	1.92 Mbps
Swing	cbr/cbr	37.87 dB	38.89 dB	0.64 Mbps	1.28 Mbps	1.92 Mbps
Tabletennis	cbr/cbr	33.66 dB	35.99 dB	0.64 Mbps	1.28 Mbps	1.92 Mbps
Tabletennis	cbr/cbr	34.41 dB	35.82 dB	0.96 Mbps	0.96 Mbps	1.92 Mbps
Salesman	cbr/vbr	39.49 db	40.68 dB	0.96 Mbps	1.04 Mbps	2.15 Mbps
Tabletennis	cbr/vbr	35.20 dB	37.50 dB	0.96 Mbps	1.53 Mbps	3.96 Mbps
Tabletennis	1-layer cbr	35.48 dB		1.92 Mbps		1.92 Mbps

Conclusion

There are several possibilities to solve the cell loss problem for video coding in an ATM net with two layered coding using the priority bit, even at high cell loss rates as 10E-2. This paper presents simulation results of a very simple technique: coding one picture twice. In the first pass the picture is encoded according to H.261 and the packets are transmitted with high priority; in the second pass the picture is encoded again according to H.261 using the reconstructed picture out of the first pass for prediction, but the packets are transmitted with low priority. The result of this technique is a data stream, without packetizing completely compatible with H.261 and very resiliant to cell loss rates up to 10E-1. Hard- and software of the codec kernels are identical to H.261 except the quantizer control for a variable rate enhancement layer and the addressing of the pictures to be encoded. Problems with layer-synchronisation are avoided. The S/N performance is nearly identical to a 1-layer H.261 codec. The expense compared to a 1-layer H.261 codec is a slightly higher processing power. The technique of 2 pass coding is not limited to H.261 codecs; the technique can be extended to other coding schemes like e.g. MPEG I.

References

Compendium ATM - Coding Procedures
ETSI NA3.1 Doc 90/28
COST211ter, February 11, 1991
CCITT SGXV, WP/1, AVC-13, November 1990

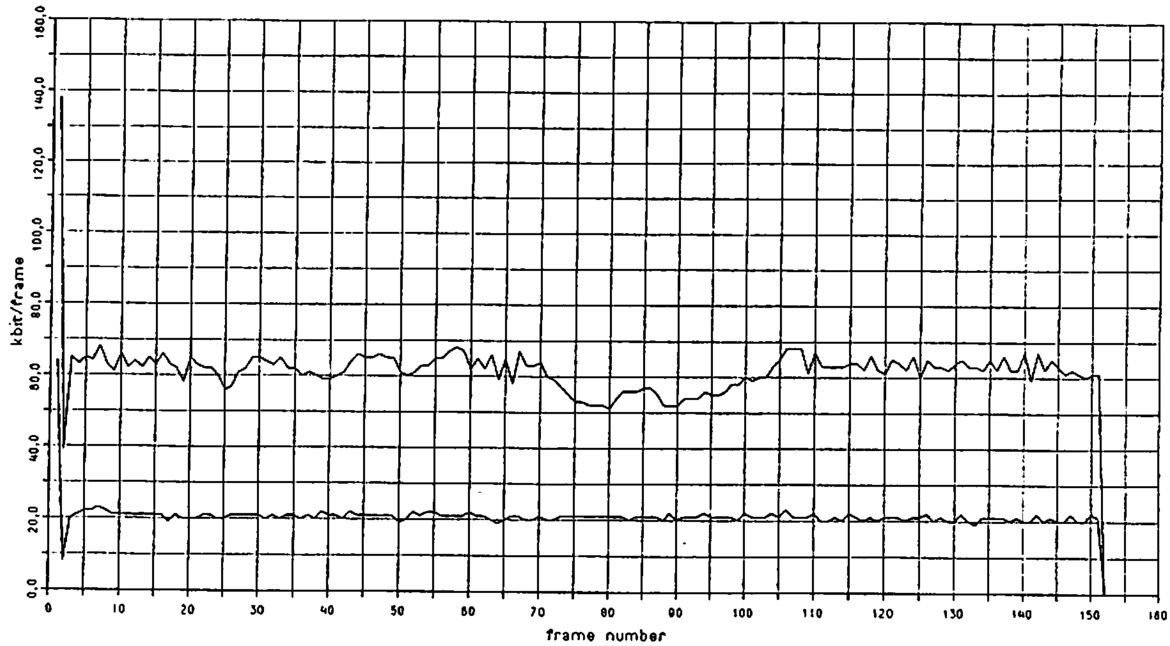
Geoffrey Morrison
Variable Bitrate Video Coding for Asynchronous Transfer Mode Networks
British Telecom Technology Journal, Vol.8, No.3 July 1990

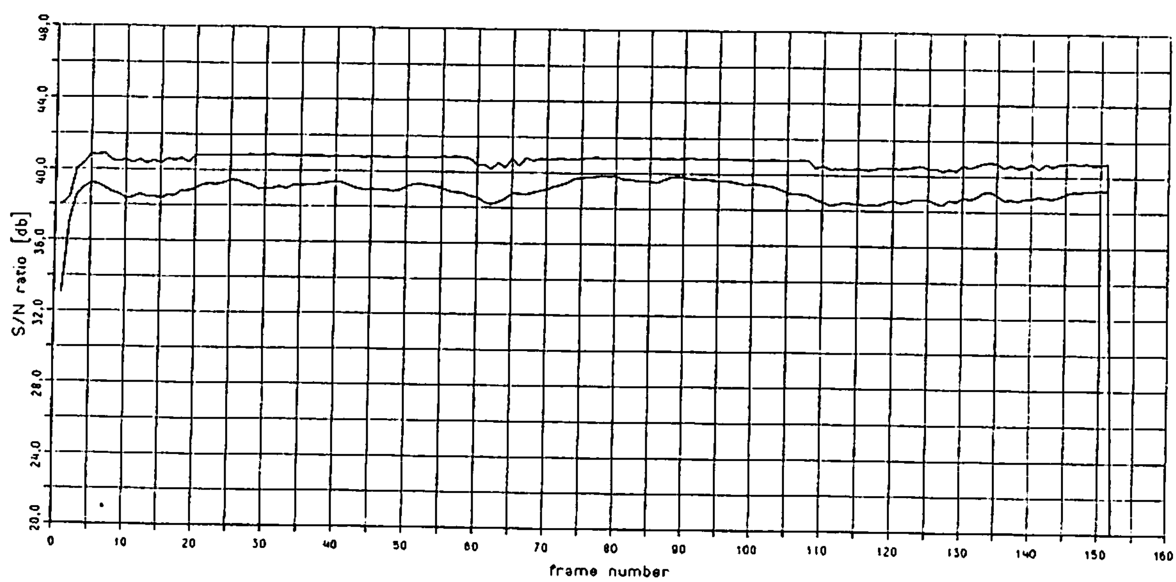
Geoffrey Morrison, David Beaumont
Two-Layer Video Coding for ATM Networks
Image Communication, March 1991

Appendix: Bitrate and S/N diagrams

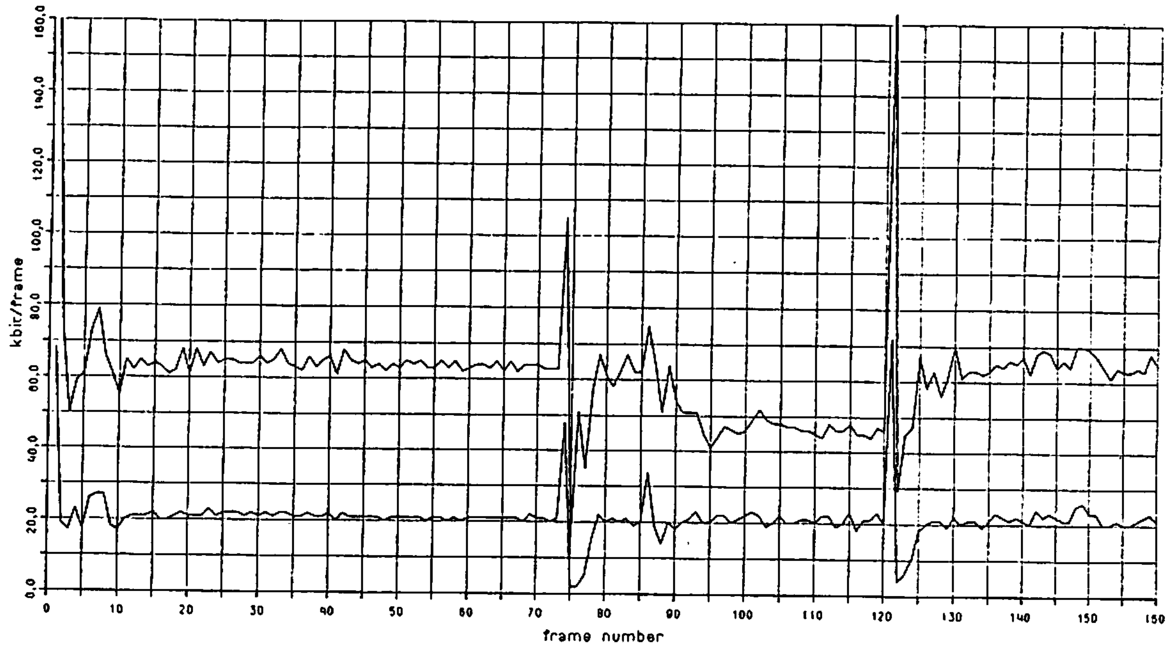
The following diagrams show for the different sequences the amount of bits/picture in front of the buffer (nevertheless at the output of the buffer the rate is fixed for fixed rate layers) and the signal to noise ratios. The lower curves give values for the base layer only, the higher curves for base + enhancement layer. Frame rate is always 30 Hz.

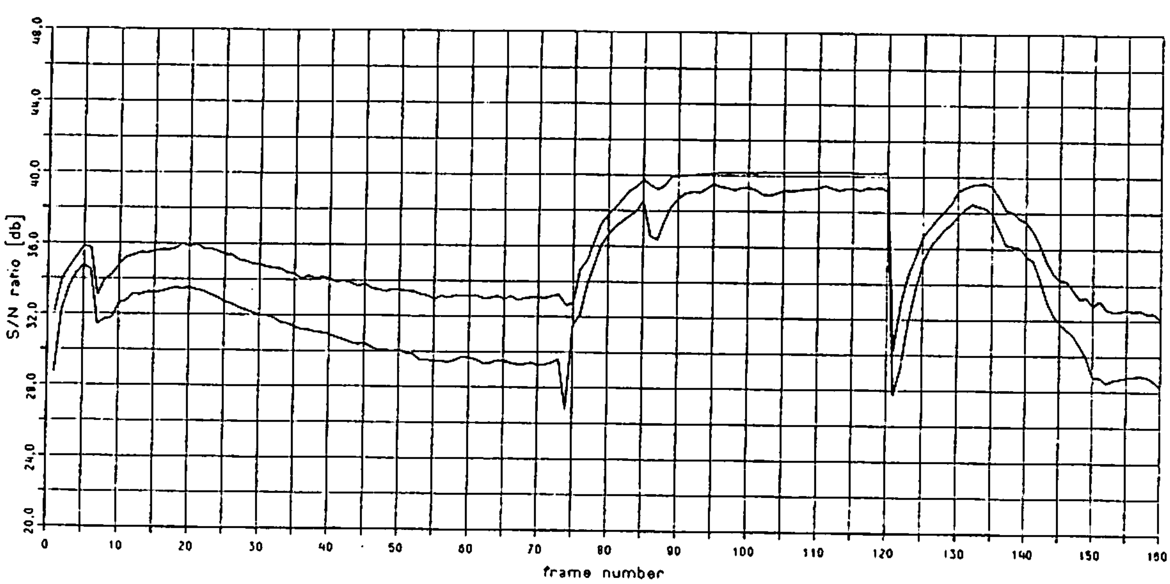
This work has been funded by the European Community under contract RACE 1018 HIVITS.



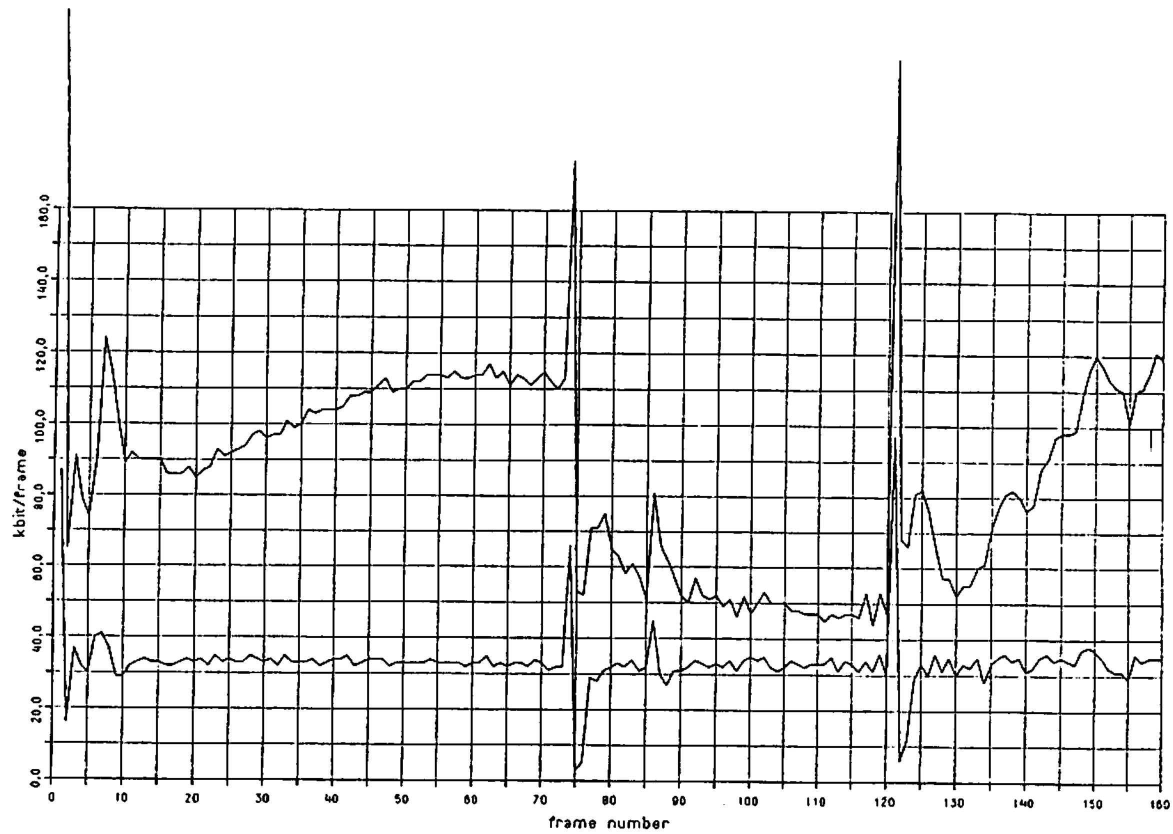


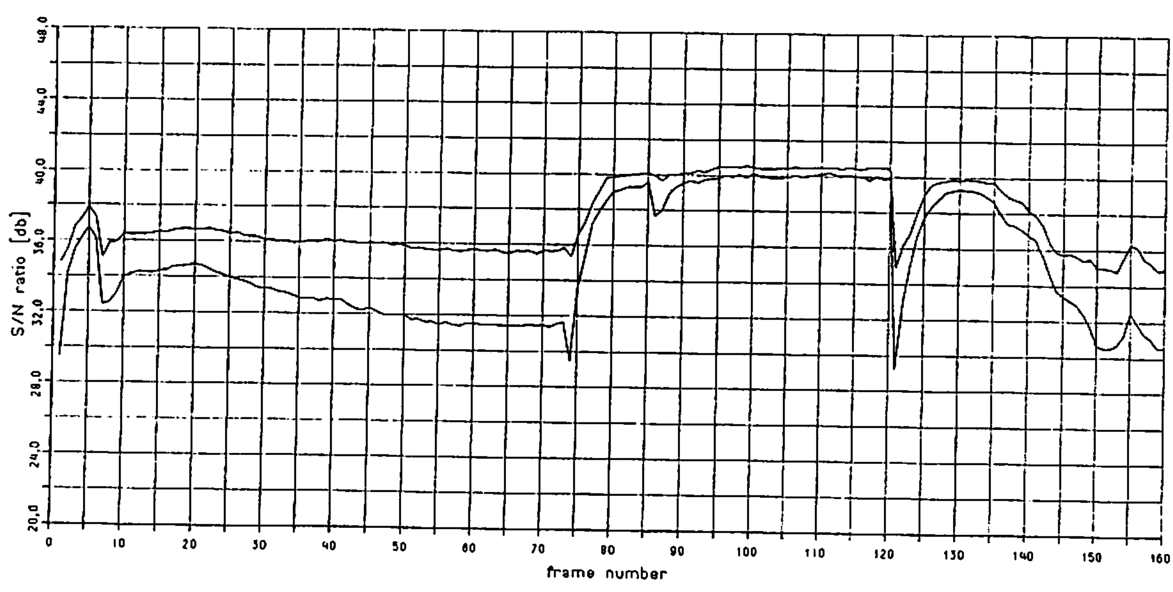
Salesman, 640 kbit/s base layer, 1280 kbit/s (fixed) enhancement layer



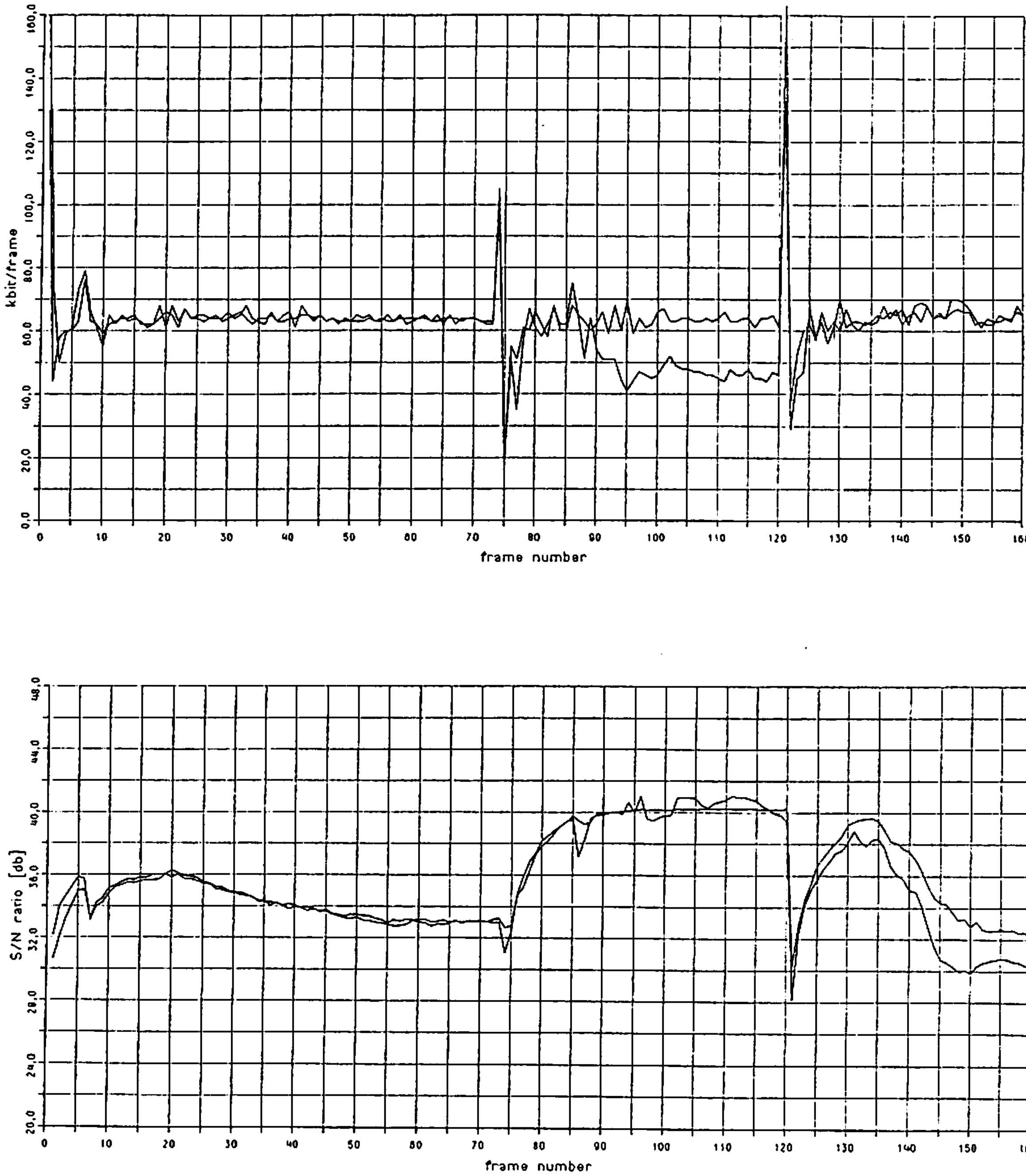


Tabletennis, 640 kbit/s base layer, 1280 kbit/s (fixed) enhancement layer





Tabletennis, 0,96 Mbit/s base layer; variable rate enhancement layer with mean: 1.53 Mbit/s; peak: 2.99 Mbit/s



Tabletennis: Comparison between 1 layer H.261 Codec at 1920 kbit/s and
the described 2 layer codec with 640+1280 = 1920 kbit/s