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

DOCUMENT #191
March 1987

Source : NTT, KDD, NEC and FUJITSU

Title : EFFECT OF VARIABLE BLOCK SIZE PROCESSING

---

1. Introduction

Variable block size (denoted hereafter as VB) processing has already been proposed by France in Docs. # 93 and #94, and Japan has suggested in Doc. #156 tat this be an item for future inclusion. In the present paper we reconfirm the effectiveness of VB, and offer some additional comments to the proposals of Doc. #156. We shall also refer to simulation results obtained with flexible hardware (denoted hereafter as FH), of which is now being implemented in Japan.

We shall first examine the statistical nature of motion-compensated interframe prediction error signals based on RM3, and describe a comparative study of the effectiveness of KLT and DCT of MC prediction error signals. Next, we outline specifications for variable block size processing, and show that it can be incorporated fairly easily in the FH which are now under consideration. We will then make a comparison between fixed block size processing (denoted hereafter as FB) and the VB method based on FH through the results of the simulation.

2. Comparative effectiveness of KLT, DCT and VB-DCT

A mean error energy has been calculated for following four methods for the purpose of comparing the VB method with the FB method. The calculation of mean error energy is based on MC error signals of the significant blocks of RM3.

a) FB with KLT
   Fixed block size 8*8 is used.
b) FB with DCT
   Fixed block size 8*8 is used.
c) VB with KLT
   One of block size 8*8, 8*4 and 4*4 is selectively used.
d) VB with DCT
   One of block size 8*8, 8*4 and 4*4 is selectively used.

The result is shown in Table 1. From table 1, the mean error energy in the VB method is less than one in the FB method. In other word, the VB method give higher energy compaction than the FB method.

As a result, it is concluded that the VB with DCT method is effective to improve the energy packing capability.

3. Specifications of VB method, and implementation of hardware

The basic idea of the VB method is to further divided an 8*8 block into sub blocks, and to carry out DCT only on sub-blocks which are significant. Basically, therefore, it is only necessary to define the procedure used to inform the receiving side where such sub-blocks are located in the 8*8 block. However, as some of these correspond to 8*8 DCT circuits which are already used in FH, it is necessary to define also methods for performing 4*4, 4*8 and 8*4 DCT. The required specifications are listed in Appendix 1 to this paper.

As the appendix lists many items, it would appear that the realization of VB requires complicated hardware, but in FH which are presently under consideration, the essential part has already been implemented.

Item 1 of the Appendix defines a specification for color signal components wherein all significant blocks are forced into the 8*8 mode (the 8*8 mode in the VB method is the same as in the FB method).

Items 2 - 4 can be easily implemented if, in designing the Inter/Intra judgement part (and the Significant/Insignificant Judgement part which shares the same part of the circuit), information is also obtained for 4*4 sub-block units.

Item 5 can be realized, based on Significant/Insignificant sub-block information, by applying a 2 digit shift to input signals or gating operation with zero output.

Item 6 is a similar kind of specifications to item 5, in addition to these X - Y address of the block memory are transposed, when the significant sub-blocks are horizontally aligned.

Items 7,8 and 10 have already been implemented in current FH.
In order to realize the scanning method in the VB method, extra two scanning classes, which is shown in Fig 2, are added to the scanning class set in the FB method. Since the number of the scanning classes are four in FB, the number of scanning classes in VB are six in total. Therefore, the scanning method in VB is easily realized within FH.

For Item 9, the signals obtained by adding the significant sub-block addresses to the bits indicated by the scanning class, may be defined as a VLC which respecifies the scanning class.

Items 11 and 12 can be realized by using part of the same types of circuit as for Items 5 and 6.

4. Simulation Results

A simulation of the VB method was carried out according to the specifications of Appendix 1.

Since Significant/Insignificant block judgement isn't employed in RM3, it is more careful to apply the VB method to RM3. We have now been carrying out the simulation based on RM3. In next meeting, the simulation result will be shown.

If the coding control or the processing of the part not covered by the recommendation is different, the method could give a different result. we therefore carried out another simulation on FH for which prototypes are now being implemented in Japan. VB processing was added to this, and a comparison was made of FB and VB.

The results of the processing will be shown on the VTR. From this, it is seen that the VB method is effective in reducing mosquito noise and block distortion. Reduction of mosquito noise is especially evident from the checked jacket head and shoulders, and from the area around the head on the split image. Decrease of block distortion is evident from Miss America's face and throat, or from Mr. Trevor's face.

Table 2 gives resultant coding parameters for the two methods, and Fig. 3 shows the change in S/N ratio. In the table, the S/N ratio also shows that the VB method is slightly superior. in particular, by comparing the ratio on non-zero coefficients to zero coefficients in significant blocks (obtained by dividing the value of the 4th term in the table with the value of the 5th), it can be seen that the proportion of non-zero coefficients in the VB method is ordinarily high, except transient duration of chart signals. This may be ascribed to the fact that significant DCT coefficients are scanned effectively, and to the fact that energy is more highly concentrated with the method as we mentioned in Chapter 2.

The iecrease of S/N ratio despite the increase of Mean step size with VB in Table 2 is because, in the 4*4 and 4*8 modes, DCT is carried out after the input signal has been amplified 4 or 2 times respectively. In these blocks, therefore, the quantizer step is equivalently smaller than the value shown in the table.

5. Conclusion

It is evident that, from the viewpoint of energy packing capability, the VB method is superior to the FB method. From simulation experiments, moreover, the method reduces block distortion and mosquito noise. Concerning S/N ratio, there was an improvement of about 0.1 - 0.6 dB, and it would appear that, from the viewpoint of hardware implementation, VB processing can be incorporated fairly easily in present FH.

We therefore propose that the VB method be included as an additional core element of FH.

(Ref.)

(1) CCITT SG XV,SGCV Document #93 "The use of a N-DCT device to perform a N/2-DCT", France

(2) CCITT SG XV, SGCV Document #94 "Variable block size hybrid coding scheme", France

(3) CCITT SG XV, SGCV Document #156 "Block type identification method, Variable block size and Overall performance", NTT, KDD, NEC and Fujitsu

Appendix to Document #191

Specifications of Variable Block Size Processing

1. Processing is carried out only on the luminance signal component.
2. Each block is divided into sub-blocks of 4 picture elements * 4 lines.
3. THe 4*4 mode is selected when the block has only one significant sub-block and the 4*8 mode is selected when two significant sub-block in the vertical or horizontal direction.
4. Significant sub-blocks of 4*4 mode or 4*8 mode are coded in inter-frame mode or intra-frame mode.
5. In the 4*4 mode, the input values for the significant sub-blocks are quadrupled and insignificant sub-blocks are forcibly set to zero.
6. In the 4*8 mode, the input values for a significant sub-block are doubled and insignificant sub-blocks are forcibly set to zero. Also, when the significant sub-blocks are horizontally aligned, the X - Y coordinates of the blocks are transposed and the significant sub-blocks are aligned vertically.
7. 4*4 mode block and 4*8 mode block signals are transformed as 8*8 block.
8. The transformed coefficient values of the 4*4 mode and 4*8 mode blocks are transmitted by the scanning, for example, indicated in Fig. 2 of this document.
9. Significant sub-block address information is appended to the scanning class indication signal.
10. Decoding is performed by inverse transformation with 8*8 block size under the condition that the non-transmitted coefficient values are set to zero.
11. In the 4*8 mode, when the significant sub-block address indicates that the significant sub-blocks are horizontally aligned, the X - Y coordinates inside the block are transposed again and the significant sub-blocks aligned horizontally.
12. After inverse transformation the insignificant sub-blocks are forcibly set to zero.

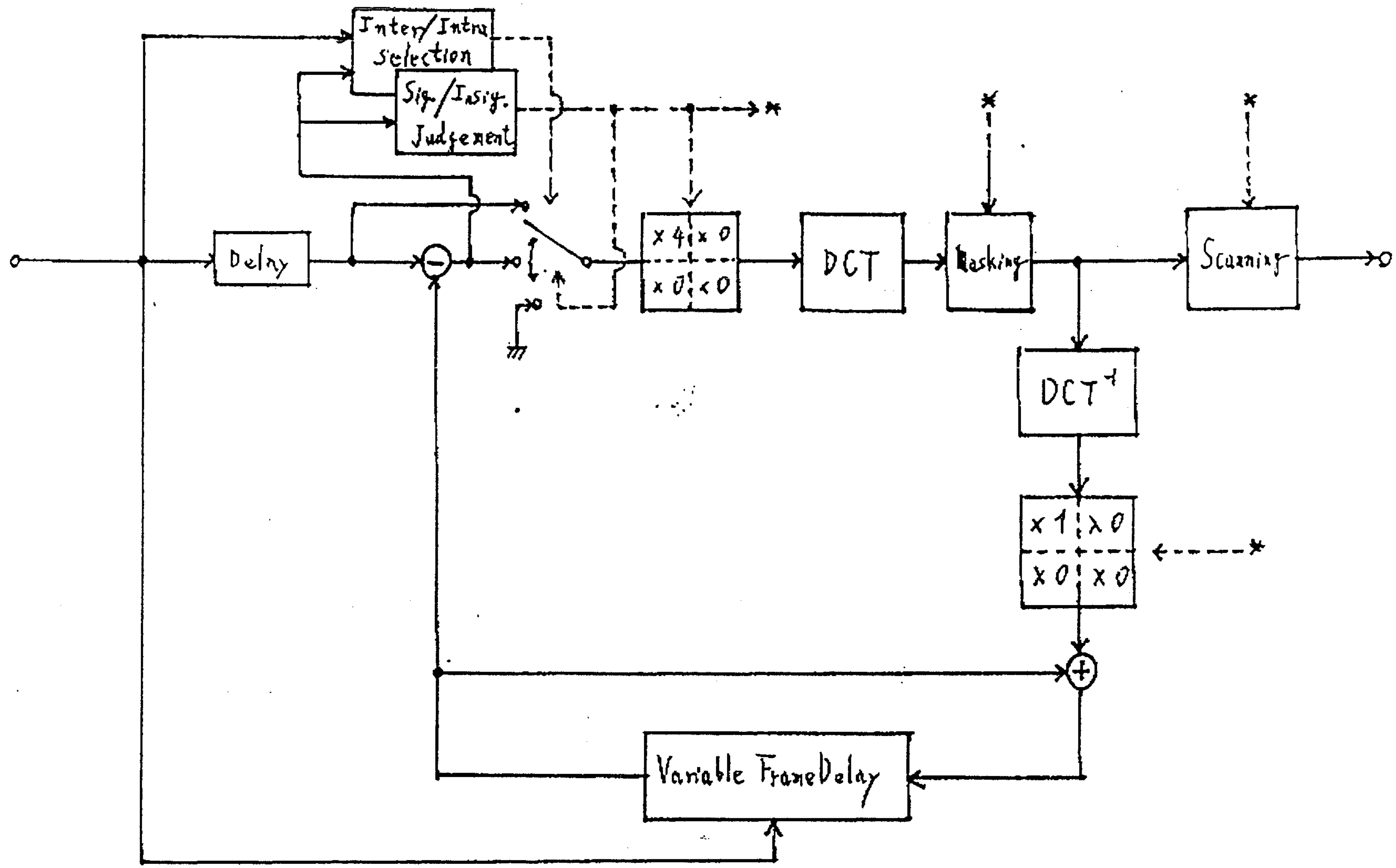

Figure 1. Significant/insignificant, Inter-frame/intra-frame and Variable Block Size Processing

Table 1. mean error energy of the FB method and the VB method.
sequence : split 46'th frame

Th/pel*			one**	three	six
3.5	fixed	DCT	100.53%	102.14%	103.38%
		KLT	100.00%	100.00%	100.00%
	variable	DCT	99.21%	99.08%	99.66%
		KLT	98.40%	96.52%	95.76%
4.0	fixed	DCT	100.55%	102.18%	103.55%
		KLT	100.00%	100.00%	100.00%
	variable	DCT	98.82%	98.72%	99.12%
		KLT	98.11%	96.01%	94.71%
4.5	fixed	DCT	100.55%	102.20%	103.66%
		KLT	100.00%	100.00%	100.00%
	variable	DCT	98.25%	97.52%	97.67%
		KLT	96.89%	94.46%	92.34%

* threshold of significant/insignificant judgement.
** number of lowest sequency coefficients

1 –	2 –	3 –	4 –
– –	– –	– –	– –
5 –	6 –	7 –	8 –
– –	– –	– –	– –
9 –	10 –	11 –	12 –
– –	– –	– –	– –
13 –	14 –	15 –	16 –
– –	– –	– –	– –

1 –	2 –	3 –	4 –
5 –	6 –	7 –	8 –
9 –	10 –	11 –	12 –
13 –	14 –	15 –	16 –
17 –	18 –	19 –	20 –
21 –	22 –	23 –	24 –
25 –	26 –	27 –	28 –
29 –	30 –	31 –	32 –

(– : don't care)

(a) 4 * 4 block mode (b) 4 * 8 block mode

Figure 2. Scanning for Variable Block Size Processing

Table 2-1 Comparison between FB and VB *1 (Miss America, Checked Jacket)

Item			MA FB	MA VB	CJ FB	CJ VB	FB	VB
0)Frame dropping rate			1.96	2.04	1.92	1.92		
1)R.M.S for luminance *2			2.20	2.18	2.49	2.45		
2)SNR for luminance *2			41.32	41.40	40.23	40.40		
1')R.M.S for luminance *3			2.65	2.63	2.70	2.66		
2')SNR for luminance *3			39.69	39.75	39.54	39.68		
3)Mean value of the step size	DC part		5.7	6.4	5.5	6.0		
	AC parts		8.2	8.9	8.0	8.5		
4)Mean value of the number of non-zero coefficients			2.97	3.15	4.46	4.77		
5)Mean value of the number of zeros before the last non-zero coefficients			4.42	4.36	8.60	8.19		
5') 4) / 5)			0.67	0.72	0.52	0.58		
6) Block type of Y	Intra		7	16	11	15		
	Fixed(Intra/No MC/No coded)		1071	1086	1138	1161		
	Inter(Inter/No MC/Coded)		315	282	360	333		
	Fixed MC(Inter/MC/No coded)		43	42	4	3		
	Inter MC(Inter/MC/Coded)		145	156	68	70		
	Filtered		189	198	72	74		
7) Block type of C	Intra		10	12	0	0		
	Fixed (Inter/No coded)		516	639	679	689		
	Inter (Inter/Coded)		265	240	112	102		
	Filtered		0	0	0	0		
8) Number of Bits	Attributes	Y	2116	2083	1682	1598		
		Cr	436	421	225	212		
		Cb	460	440	199	185		
		Total	3012	2946	2108	1996		
	Classification indexes		938	1640	881	1376		
	EOB words		2235	2124	1659	1568		
	Motion Vectors		1194	1262	491	500		
	Coefficients	Y	8475	9340	13107	13212		
		Cr	1706	1454	499	402		
		Cb	2001	1566	261	230		
		Total	12183	12361	13868	13845		
	Total		19562	20333	19007	19285		

*1 Average value (exclude first two coded frames data)

*2 Difference between temp.filter output to decoder output.

*3 " input "

Table 2-2 Comparison between FB and VB *1 (Split and Trevor)

Item			SP FB	SP VB	TR FB	TR VB	S+T FB	S+T VB
0)Frame dropping rate			3.00	3.08	2.08	2.30	2.34	2.52
1)R.M.S for luminance *2			4.04	4.02	2.85	2.77	3.18	3.12
2)SNR for luminance *2			36.02	36.07	39.05	39.29	38.20	39.38
1')R.M.S for luminance *3			4.17	4.16	2.99	2.91	3.32	3.26
2')SNR for luminance *3			35.74	35.76	38.63	38.86	37.82	37.98
3)Mean value of the step size	DC part		12.7	14.9	8.3	8.4	9.5	10.2
	AC parts		13.6	15.0	10.5	10.6	11.4	11.9
4)Mean value of the number of non-zero coefficients			3.68	3.54	3.78	3.92	3.75	3.81
5)Mean value of the number of zeros before the last non-zero coefficients			3.20	2.89	3.68	3.83	3.54	3.56
5') 4) / 5)			1.15	1.22	1.02	1.02	1.06	1.07
6) Block type of Y	Intra		115	173	85	114	93	131
	Fixed(Intra/No MC/No coded)		762	734	1019	1011	947	933
	Inter(Inter/No MC/Coded)		253	250	176	161	198	186
	Fixed MC(Inter/MC/No coded)		122	92	48	33	69	49
	Inter MC(Inter/MC/Coded)		329	333	253	263	275	283
	Filtered		452	425	302	296	344	333
7) Block type of C	Intra		64	70	11	12	26	29
	Fixed (Inter/No coded)		575	598	695	692	662	6[illegible]6
	Inter (Inter/Coded)		151	123	84	86	103	96
	Filtered		0	0	0	0	0	0
8) Number of Bits	Attributes	Y	3045	3093	2180	2192	2422	2447
		Cr	390	373	286	281	315	307
		Cb	378	356	294	297	318	313
		Total	3815	3823	2761	2771	3056	3068
	Classification indexes		1398	2538	1031	1826	1133	2027
	EOB words		2746	2854	1835	1915	2090	2180
	Motion Vectors		3508	3244	2275	2211	2620	2503
	Coefficients	Y	15806	16249	12202	13587	13212	14340
		Cr	1371	1214	390	393	664	625
		Cb	1073	952	360	370	560	535
		Total	18251	18416	12953	14351	14437	15500
	Total		29718	30875	20855	23074	23336	25278

*1 Average value (exclude first two coded frames data)

*2 Difference between Temp. filter output to decoder output.

*3 " input " .

Table 2-3 Comparison between FB and VB *1 (Chart 1 and 2)

Item			CH1 FB	CH1 VB	CH2 FB	CH2 VB	1+2 FB	1+2 VB
0)Frame dropping rate			1.16	1.09	1.53	1.56	1.35	1.32
1)R.M.S for luminance *2			1.97	1.84	2.85	2.80	2.43	2.31
2)SNR for luminance *2			42.22	42.79	39 40	39.61	40.75	41.22
1')R.M.S for luminance *3			1.97	1.84	2.85	2.80	2.43	2.31
2')SNR for luminance *3			42.22	42.79	39.40	39.61	40.75	41.22
3)Mean value of the step size	DC part		3.0	3.0	4.5	4.6	3.8	3.8
	AC parts		4.9	4.9	6.3	6.4	5.6	5.7
4)Mean value of the number of non-zero coefficients			0.79	1.09	1.00	1.01	0.90	1.05
5)Mean value of the number of zeros before the last non-zero coefficients			1.03	1.69	1.77	2.11	1.42	1.89
5') 4) / 5)			0.77	0.64	0.56	0.48	0.63	0.56
6) Block type of Y	Intra		0	2	1	1	1	2
	Fixed(Intra/No MC/No coded)		873	1043	866	903	870	974
	Inter(Inter/No MC/Coded)		709	537	713	676	711	606
	Fixed MC(Inter/MC/No coded)		0	0	0	0	0	0
	Inter MC(Inter/MC/Coded)		0	0	1	1	0	0
	Filtered		0	0	2	2	1	1
7) Block type of C	Intra		0	0	0	0	0	C
	Fixed (Inter/No coded)		537	551	443	445	488	499
	Inter (Inter/Coded)		254	240	348	346	303	292
	Filtered		0	0	0	0	0	0
8) Number of Bits	Attributes	Y	2412	2062	2347	2316	2378	2187
		Cr	383	372	412	411	398	391
		Cb	394	383	395	388	395	385
		Total	3190	2818	3156	3116	3172	2965
	Classification indexes		1420	1506	1433	1487	1427	1497
	EOB words		2895	2342	3195	3078	3052	2704
	Motion Vectors		0	0	20	21	10	10
	Coefficients	Y	3187	3639	5914	6517	4613	5056
		Cr	597	567	657	664	629	615
		Cb	515	477	589	595	554	535
		Total	4300	4684	7162	7778	5797	6207
	Total		11805	11350	14966	15480	13458	13383

*1 Average value (exclude first two coded frames data)

*2 Difference between temp. filter output and decoder output.

*3 " input

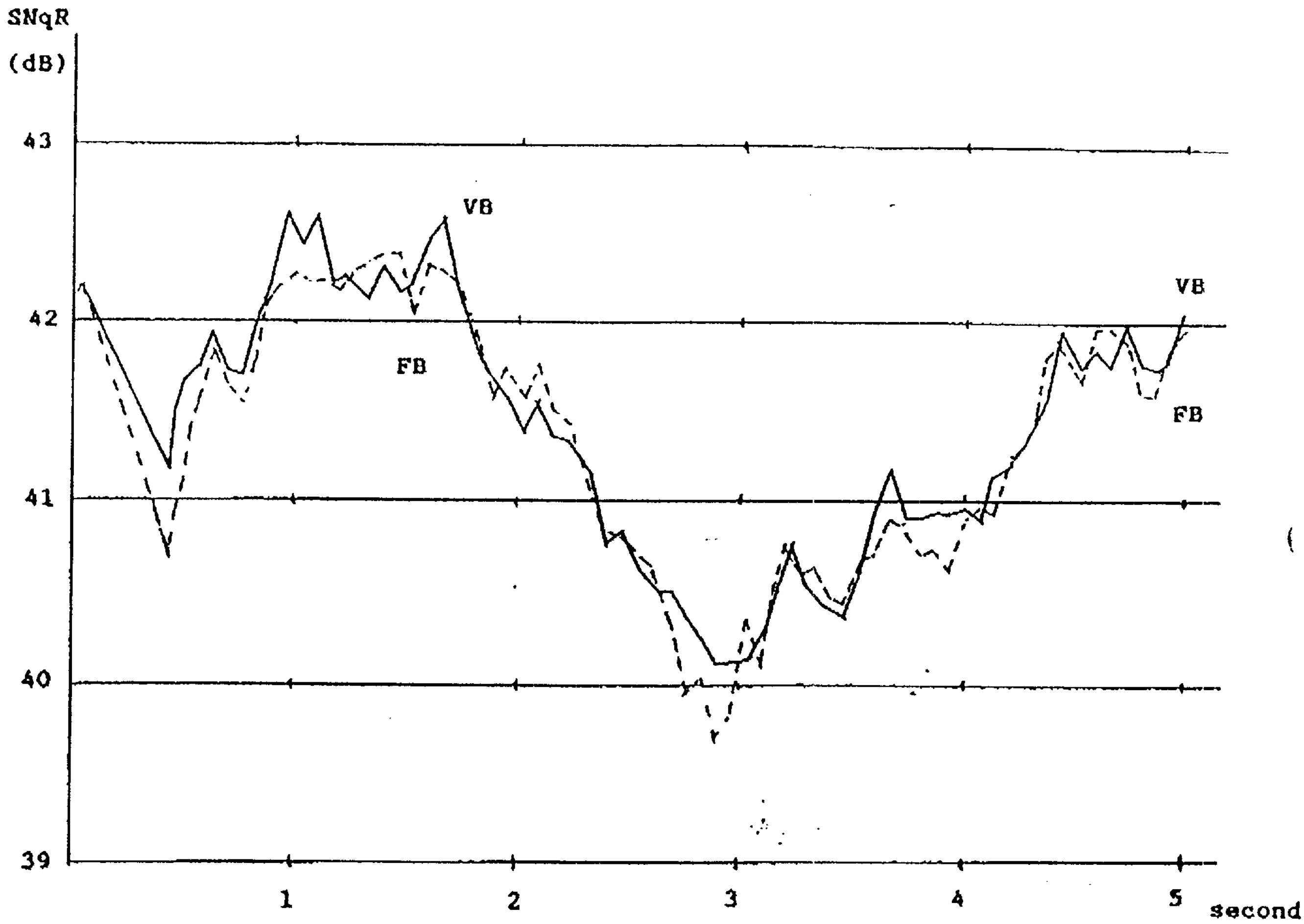

Fig. 3-1 SNqR (Miss America)

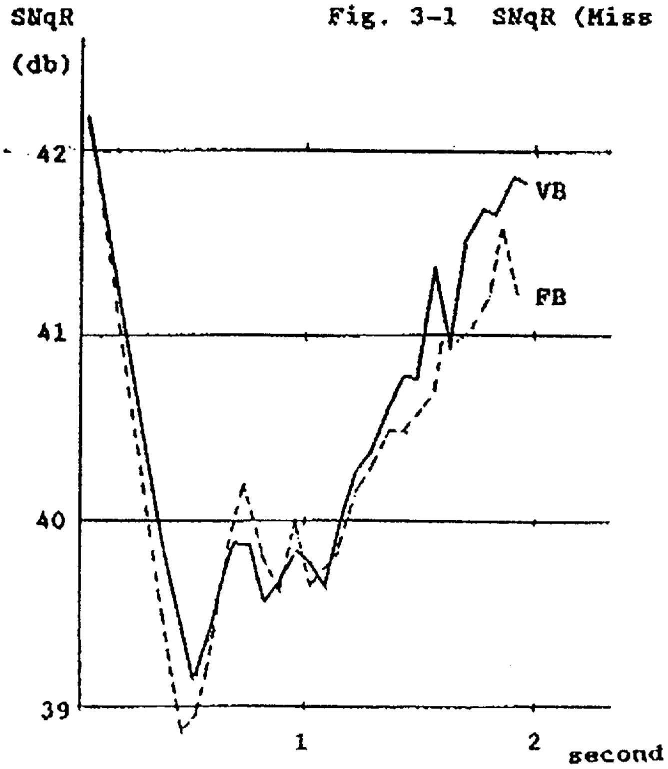

Fig. 3-2 SNqR (Checked Jacket)

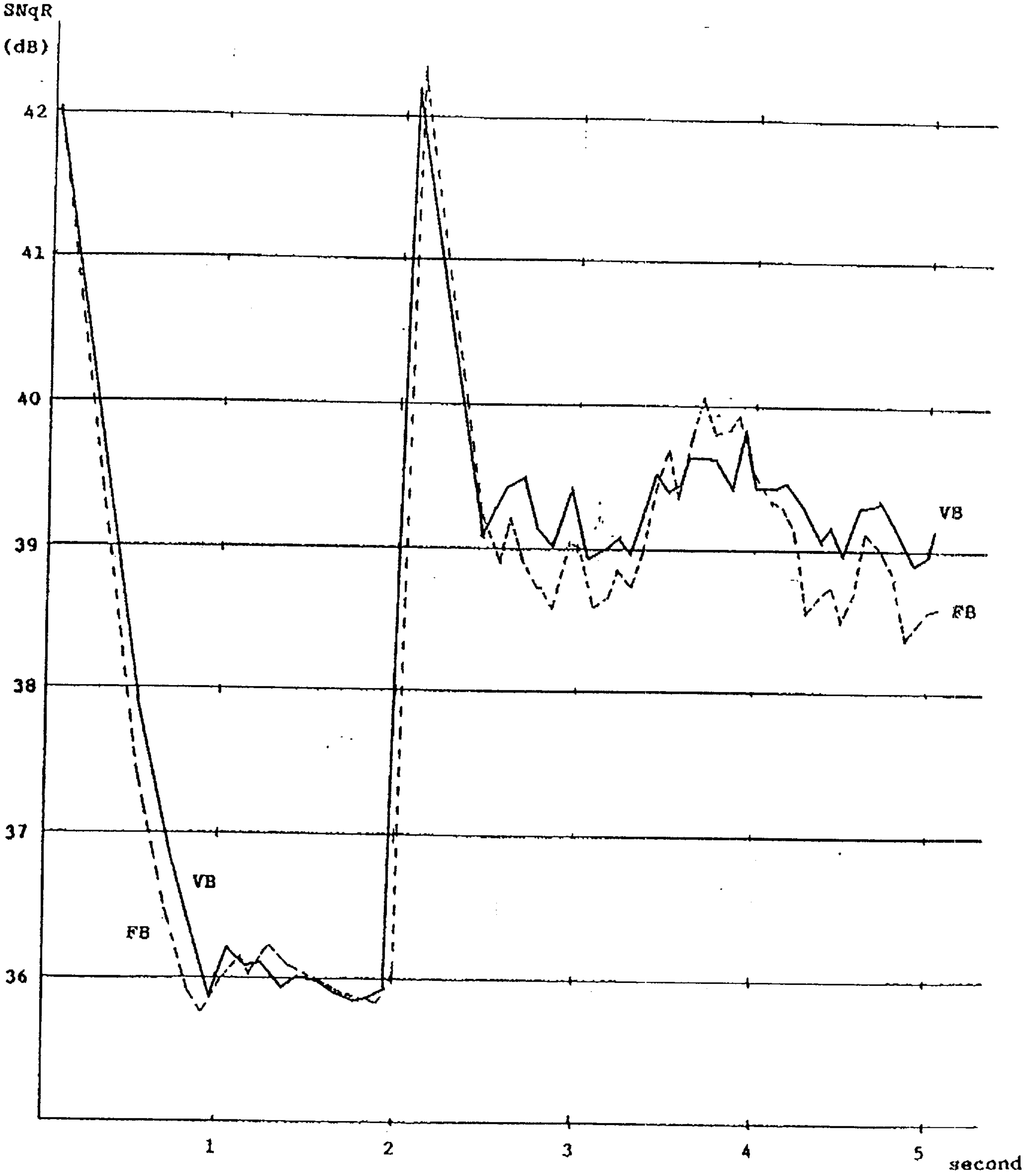

Fig. 3-3 SNqR (Split and Trevor)

I₁ original の ... scene change?

1 sec で stable になる
2—3

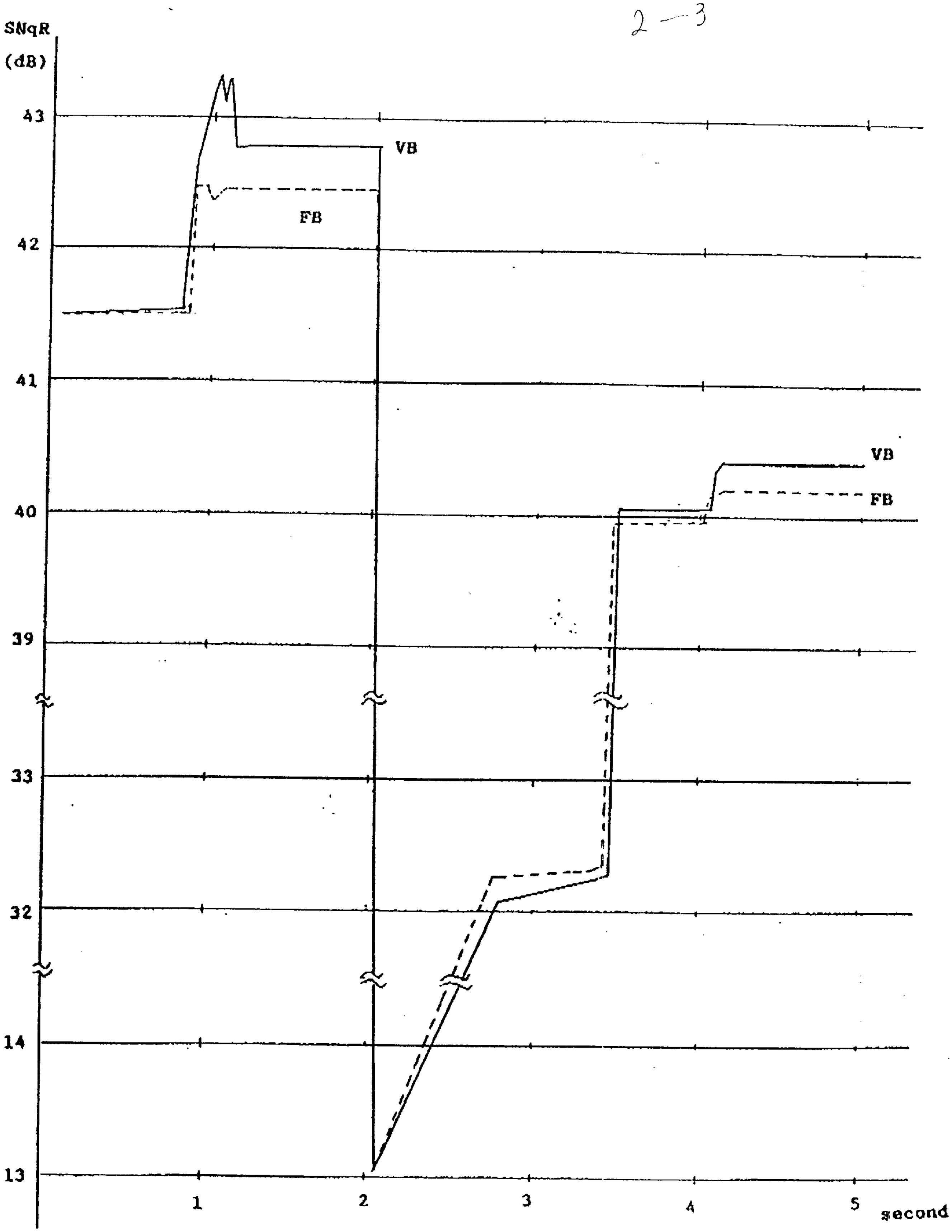

Fig. 3-4 SNqR (Chart 1 and 2)