CCITT SGXV

Working Party XV/1

Specialists Group on Coding for Visual Telephony

Document #389

September 19, 1988

Source: FRG

Title: Improvements of RM5

At low transmission rates the Reference Model 5 for p*64 kbit/s coding of moving video produces acceptable picture quality for common video telephone scenes like MissAmerica or Claire. With scenes like Trevor, Salesman or Swing the picture quality could be better. The introduction of the 4 techniques described in this document will result in better picture quality especially for difficult sequences. The techniques are:

1. Two dimensional prediction and VLC of the displacement vectors
2. Inter/Intra switch per macroblock
3. Prevention of buffer overflow
4. Simplified and effective coding of the first picture.

1. Two dimensional prediction and VLC of the displacement vectors

The coding of the displacement vectors with fixed codeword length as in RM5 wastes a lot of bits when the moving area in the scene is large. Also the displacements are restricted to +- 7 pixels. Two dimensional prediction of the displacements and variable length coding reduces the number of bits spent for describing the displacement significantly (see the simulation results chapter 5). In addition bigger displacements than +-7 can be coded in a bit saving manner.

Let k,l be the column and row number of a specific macroblock, let dk(k,l) and dl(k,l) be its displacements, then the prediction errors to be coded are:

$$ddk(k,l) = dk(k,l) - (dk(k-1,l)+dk(k,l-1))/2$$
$$ddl(k,l) = dl(k,l) - (dl(k-1,l)+dl(k,l-1))/2$$

Integer calculation with truncation. In the first row the values for dk(k,l-1) and dl(k,l-1) are set to zero, in the first column the values for dk(k-1,l) and dl(k-1,l).

The one dimensional VLC used has a range of +- 64 since this code is also used for encoding DC values of the cosine transform at the beginning of the sequence or when buffer is nearly full (see chapter 2 and 4). The gain of a special code with +-15 maximum range is neglectible. The code used is:

C	AMPLITUDE		CODE WORD	CW-LENGTH	
C			11111111111111	not allowed	
C	-64		11111111111110	14	
C	-63		11111111111101		
	.		.		
	.		.		
C	-44		11111111101010		
C	-43		11111111101001		
C	-42		11111111101000	23* 14	
C	-41		1111111110011	13	
C	-40		1111111110010		
C	-39		1111111110001		
C	-38		1111111110000		
C	-37		1111111101111		
C	-36		1111111101110		
C	-35		1111111101101		
C	-34		1111111101100		
C	-33		1111111101011		
C	-32		1111111101010		
C	-31		1111111101001		
C	-30		1111111101000	12* 13	
C	-29		111111110011	12	
C	-28		111111110010		
C	-27		111111110001		
C	-26		111111110000		
C	-25		111111101111		
C	-24		111111101110		
C	-23		111111101101		
C	-22		111111101100		
C	-21		111111101011		
C	-20		111111101010		
C	-19		111111101001		
C	-18		111111101000	12* 12	
C	-17		11111110011	11	
C	-16		11111110010		
C	-15		11111110001		
C	-14		11111110000		
C	-13		11111101111		
C	-12		11111101110	6* 11	
C	-11	1014	1111110110	10	
C	-10	1013	1111110101	10	
C	-9	1012	1111110100	10	
C	-8	505	111111001	9	
C	-7	504	111111000	9	
C	-6	251	11111011	8	
C	-5	250	11111010	8	
C	-4	124	1111100	7	
C	-3	30	11110	5	
C	-2	14	1110	4	
C	-1	6	110	3	
C	0	0	0	1	
C	1	4	100	3	
C	2	10	1010	4	
C	3	22	10110	5	
C	4	92	1011100	7	
C	5	186	10111010	8	
C	.	.	.	.	etc

2. Inter/Intra switch per macroblock

The most important reasons for re-introducing the inter/intra switch are:
a) much better performance at scene cuts, too fast movement and in areas of decovered background (e.g. when Trevor raises his hands)
b) better error resilience and very simple implementation of (necessary) forced update

The amount of additional bits is very low, because the 1 extra bit per macroblock is only spent for coded macroblocks (no mc, coded); possible displacement vectors are forced to zero. The formulas written as FORTRAN statements are:

```
DO L=1,16
  OR=O(K,L)
  DIF= OR-S(K,L)
   VAR=VAR+DIF*DIF
   VAROR=VAROR+OR*OR
   MWOR=MWOR+OR
END DO
END DO

VAR=VAR/256
VAROR=VAROR/256 - (MW/256)*(MW/256)

IF (VAROR.LT.VAR .AND.VAR.GT.64) THEN ! CODE ORIGINAL BLOCK
```

with O(K,L) beeing the pixels in the original macroblock
S(K,L) the pixels of the (motion compensated) estimated macroblock.
The lower threshold of 64 for this decision is empirically optimized.
For blocks to be coded in the intra mode, the DC-value is 2d predicted as described in chapter 4.

3. Prevention of buffer overflow

An annoying effect of RM5 is the agreed procedure in case of buffer overflow: the complete freezing of the past in a macroblock. In critical situations this causes e.g. clearly visible rectangular holes in the moving object. A simple strategy helps to overcame the problem:
If the quantizer step size is 64 (that means the buffer tends to overflow) , only the DC component in the transform domain is quantized and coded. The VLC shown on page 2 is used. This code is now much more effective than the 2d-VLC of RM5 and no EOB-codeword has to be transmitted.

4. Simplified and effective coding of the first picture

The procedure in RM5 for coding the first picture is uneffective and wastes a lot of bits. Also the special quantizer for DC-components of macroblocks, the doubling of the bits/picture and the skipping of the second picture to be coded is not favorable for hardware implementations. I nstead of spending 25 bits per macroblock when coding the mean values by an 8 bit FLC, prediction of the mean values of 8*8 blocks from preceeding blocks and encoding them with the VLC given on page 2 is much more effective.

Let k,l be the column and row number of a specific 8*8 block, let m(k,l) be its mean value (with a scaling factor the (1,1) DCT component),then the prediction error to be coded is

$$dm(k,l) = m(k,l) - (m(k-1,l)+m(k,l-1))/2$$

In the first row the value for m(k,l-1) is set to 128, in the first column the value for m(k-1,l).

This prediction error is multiplied by 64 (the scaling factor for same amplitude range as in the DCT transform domain) quantized as described in chapter 3 and VLC coded according to the table on page 2. Because of the fact that the buffer is empty at the beginning, no values of the quantizer step size below 16 is allowed for the first picture.

The performance is much better than in RM5 and up to 3000 bits are saved for the first picture. Results of S/N of the first reconstructed pictures are (in RM5 the first picture is not shown):

	Claire		Swing		Salesman	
	RM5	RM5NS	RM5	RM5NS	RM5	RM5NS
pict.no.						
1	(24.2)	27.0	(18.4)	19.9	(23.3)	25.0
4	34.5	35.5	19.5	20.8	25.1	26.8
7	35.1	36.2	20.5	22.1	26.7	28.3
10	36.2	37.1	21.4	23.1	27.5	29.2
13	36.9	37.5	22.6	24.0	27.8	29.4
16	37.7	38.0	23.7	24.8	28.2	29.6

(RM5NS means RM5 New Start procedure)

5. Results (MissaAmerica, Claire, Trevor, Salesman, Swing, SceneCut)

On the following pages the results with the different improuvements are shown according to the agreed presentation form. RM5-MIQ7 indicates coding according to RM5 with the addition of the four proposed improvements (predictive coding of the displacements, inter/intra switch, new start procedure and prevention of buffer overflow).

For Miss America and Trevor the results with the 3 step search according to RM5 and only predictive coding of the displacements are also given (RM5_BVP); Compared to the FLC of RM5 the number of bits for motion vectors are more than halved and S/N figu res sligthly higher.

As expected the improvements are only small for Miss America and Claire; S/N figueres of Swing are comparable, but the pictures with RM5_MIQ7 are looking better. Significant improvements are visible if difficult sequences like Trevor or Scene Cuts are coded.

STATISTICS RM5	SOURCE: AEG (FRG) DATE: 3-SEP-88 SEQUENCE: TREV,RM5 CODED PICTURES: 23 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: TREV,RM5,BVP CODED PICTURES: 23 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: TREV,RM5,MIQ7 CODED PICTURES: 23 BIT RATE: 60 kbit/s FRAME RATE: 10.00
1. RMS for luminance	7.25	6.77	6.01
for chrominance (u)	3.24	3.23	3.14
for chrominance (v)	2.83	2.81	2.68
2. SNR for luminance	31.12	31.63	32.56
for chrominance (u)	37.94	37.98	38.20
for chrominance (v)	39.16	39.27	39.60
3. Mean value of step size	46.35	44.17	42.78
4. Mean value of the number of non-zero coefficients	1.78	1.87	1.91
5. Mean value of the number of zeros before the last NZ-coefficient	3.74	4.00	4.26
6. Block type of MACRO | not MC, not COD	239.96	236.61	237.91
| MC, COD	88.17	91.65	83.83
| MC, not COD	38.30	35.52	33.00
| not MC, COD	29.57	32.22	41.26
| INTRA	0.00	0.00	14.78
7. Block type of Y | not MC, not COD	1167.61	1159.52	1162.83
| MC, COD	200.22	214.04	189.61
| MC, not COD	153.22	142.09	132.00
| not MC, COD	62.96	68.35	99.57
8. Block type of UV | not COD	779.52	778.43	779.09
| COD	12.48	13.57	12.91
9. Number of bits | Macro attributes	444.74	453.78	501.78
| End of block	1412.87	1486.43	1501.04
| Motion vectors	1011.83	538.74	469.96
| Coefficients | Y	2946.78	3313.30	3317.04
| | U	47.96	45.87	47.70
| | V	54.04	53.22	51.13
| | Total	3048.78	3412.39	3415.87
| Total	5918.22	5891.35	5888.65

STATISTICS RM5			SOURCE: AEG (FRG) DATE: 3-SEP-88 SEQUENCE: MISSA,RM5 CODED PICTURES: 49 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 3-SEP-88 SEQUENCE: MISSA,RM5_BVP CODED PICTURES: 49 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: MISSA,RM5_MIQ7 CODED PICTURES: 49 BIT RATE: 60 kbit/s FRAME RATE: 10.00
1. RMS for luminance			3.30	3.25	3.23
for chrominance (u)			3.08	3.04	3.06
for chrominance (v)			2.89	2.80	2.82
2. SNR for luminance			37.77	37.91	37.95
for chrominance (u)			38.36	38.47	38.41
for chrominance (v)			38.92	39.19	39.13
3. Mean value of step size			21.41	20.39	21.39
4. Mean value of the number of non-zero coefficients			1.43	1.53	1.55
5. Mean value of the number of zeros before the last NZ-coefficient			4.49	4.69	4.67
6. Block type of MACRO	not MC, not COD		226.67	220.00	221.73
	MC, COD		69.88	71.69	70.37
	MC, not COD		25.51	22.51	20.96
	not MC, COD		73.94	81.80	82.94
	INTRA		0.00	0.00	0.22
7. Block type of Y	not MC, not COD		1292.84	1291.53	1297.04
	MC, COD		123.61	128.65	126.45
	MC, not COD		102.04	90.04	83.84
	not MC, COD		65.51	73.78	76.67
8. Block type of UV	not COD		700.49	689.22	689.31
	COD		91.51	102.78	102.69
9. Number of bits	Macro attributes		505.73	526.22	608.12
	End of block		1725.80	1841.88	1839.67
	Motion vectors		763.10	283.14	269.20
	Coefficients	Y	2313.88	2550.31	2517.31
		U	248.57	306.47	295.84
		V	311.59	354.24	349.90
		Total	2874.04	3211.02	3163.04
	Total		5868.67	5862.27	5880.04

STATISTICS RM5

	SOURCE: AEG (FRG) DATE: 3-SEP-88 SEQUENCE: CLAIRE_RM5 CODED PICTURES: 155 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: CLAIRE_RM5,MIQ7 CODED PICTURES: 155 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 3-SEP-88 SEQUENCE: SALESMAN,RM5 CODED PICTURES: 148 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: SALESMAN,RM5_MIQ7 CODED PICTURES: 148 BIT RATE: 60 kbit/s FRAME RATE: 10.00
1. RMS for luminance	3.29	3.27	7.81	7.73
for chrominance (u)	2.76	2.92	3.07	3.22
for chrominance (v)	1.90	2.10	2.85	2.92
2. SNR for luminance	37.79	37.86	30.34	30.41
for chrominance (u)	39.33	39.18	38.39	37.99
for chrominance (v)	42.56	42.58	39.05	38.82
3. Mean value of step size	22.43	22.81	32.95	33.45
4. Mean value of the number of non-zero coefficients	2.39	2.42	1.72	1.71
5. Mean value of the number of zeros before the last NZ-coefficient	8.23	8.08	10.41	10.57
6. Block type of MACRO | not MC, not COD	283.56	279.83	260.06	258.87
| MC, COD	53.17	50.15	35.61	32.24
| MC, not COD	15.87	10.16	8.00	5.85
| not MC, COD	43.40	55.85	92.33	99.03
| INTRA	0.00	0.26	0.00	2.41
7. Block type of Y | not MC, not COD	1311.92	1323.16	1314.54	1314.64
| MC, COD	135.45	132.54	88.74	80.51
| MC, not COD	63.48	40.65	32.00	23.41
| not MC, COD	73.15	87.65	148.72	165.45
8. Block type of UV | not COD	761.82	757.90	782.96	782.29
| COD	30.18	34.10	9.04	9.71
9. Number of bits | Macro attributes	367.45	443.75	460.69	567.48
| End of block	1158.81	1272.08	1535.27	1575.32
| Motion vectors	552.31	198.27	348.86	120.19
| Coefficients | Y	3634.30	3789.81	3531.80	3622.42
| | U	141.04	152.97	24.99	26.67
| | V	57.88	97.72	13.49	13.68
| | Total	3833.21	4040.51	3570.28	3662.77
| Total	5911.78	5954.61	5915.11	5925.77

STATISTICS RM5	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: SWING_RM5 CODED PICTURES: 99 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: SWING_RM5,MIQ7 CODED PICTURES: 99 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: SZW_RM5 CODED PICTURES: 25 BIT RATE: 60 kbit/s FRAME RATE: 10.00	SOURCE: AEG (FRG) DATE: 5-SEP-88 SEQUENCE: SZW_RM5_MIQ7 CODED PICTURES: 25 BIT RATE: 60 kbit/s FRAME RATE: 10.00
1. RMS for luminance	7.53	7.37	13.22	9.74
for chrominance (u)	5.09	5.10	12.78	9.55
for chrominance (v)	4.41	4.52	12.78	9.65
2. SNR for luminance	31.26	31.25	30.65	32.75
for chrominance (u)	34.30	34.17	31.80	34.60
for chrominance (v)	35.43	35.18	31.67	34.27
3. Mean value of step size	27.55	27.97	38.64	33.16
4. Mean value of the number of non-zero coefficients	1.56	1.55	1.28	1.24
5. Mean value of the number of zeros before the last NZ-coefficient	17.97	18.69	4.36	4.28
6. Block type of MACRO | not MC, not COD	275.51	276.00	239.76	235.72
| MC, COD	7.83	5.60	45.12	24.56
| MC, not COD	1.74	1.06	22.64	7.20
| not MC, COD	110.93	113.34	88.48	128.52
| INTRA	0.00	1.47	0.00	48.28
7. Block type of Y | not MC, not COD	1384.94	1389.30	1282.32	1306.76
| MC, COD	20.31	13.62	92.88	50.88
| MC, not COD	6.95	4.24	90.56	28.80
| not MC, COD	171.80	176.84	118.24	197.56
8. Block type of UV | not COD	753.29	754.53	695.64	701.76
| COD	38.71	37.47	96.36	90.24
9. Number of bits | Macro attributes	455.14	571.39	494.88	675.44
| End of block	1425.09	1427.27	1603.20	1836.96
| Motion vectors	76.53	28.65	542.08	94.80
| Coefficients | Y	3457.49	3446.58	2347.12	2601.04
| | U	312.20	296.23	451.88	355.40
| | V	173.75	166.16	541.92	437.72
| | Total	3943.44	3908.97	3340.92	3394.16
| Total	5900.20	5936.28	5981.08	6001.36

6. Video tape presentation

First RM5 with the proposed improvements is compared with RM5
On the left side RM5, on the right side RM5-MIQ7 with the four proposed improvements. Every simulation with 15 bits/macroblock, 60 kbits/s transmission rate, 6400 bit buffer size.

1.	Missa Amerika	50 coded pictures
2.	Scene Cut (from MissA. to Alexis)	24
3.	Trevor	24
4.	SALESMAN	100
5.	SWING	100

Then the following sequences coded with the improved RM5-MIQ7

1.	TREVOR	24
2.	SALESMAN	100
3.	SWING	100
4.	MISSA	50
5.	CLAIRE	100