Telecommunication Standardization Sector
Study Group 15
Experts Group for ATM Video Coding
(Rapporteur's Group on Part of Q.2/15)

Document AVC-541
July 1993

INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO-IEC/JTC1/SC29/WG11
CODED REPRESENTATION OF PICTURE AND AUDIO INFORMATION

ISO-IEC/JTC1/SC29/WG11/MPEG93/633

Source: I. Parke, M.E.Nilsson, BT Labs
Title: Syntax and semantic extensions for spatial scalability
Purpose: Proposal

This document provides additional syntax elements for inclusion into the video working draft, version 4.2 28th April 1993 for spatial scalability, including snr scalability.

Picture spatial scalable extension

picture_spatial_scalable_extension() {	No. of bits	Mnemonic
extension_start_code	32	bslbf
extension_start_code_identifier	4	uimsbf
lower_layer_temporal_reference	10	uimsbf
spatial_scalable_mode	2	uimsbf
horizontal_subsampling_ratio	8	uimsbf
vertical_subsampling_ratio	8	uimsbf
low_resolution_horizontal_offset	15	simsbf
marker_bit	1	"1"
low_resolution_vertical_offset	15	simsbf
next_start_code()		
}		

lower_layer_temporal_reference - An unsigned integer value which indicates the temporal reference of the lower layer picture to be used to provide the spatial prediction. If the lower layer indicates the temporal reference with more than 10 bits, the least significant bits are indicated here. If the lower layer indicates the temporal reference with less than 10 bits, all bits are indicated here, with the extra most significant bits being set to zero.

spatial_scalable_mode - The spatial_scalable_mode indicates the type of spatial scalability used for the picture. If no picture spatial scalable extension data is present for a picture, no spatial scalability is used for that picture. It also indictates the macroblock_type tables to be used. If no picture spatial scalable extension data is present for a picture, tables B.2-1, B.2-2, B.2-3, B.2-8, B.2-9 and B.2-10 should be used.

horizontal_subsampling_ratio - The horizontal_subsampling_ratio is an 8 bit integer that defines the ratio with which to up-sample the lower layer coded picture to generate the spatial prediction. It is composed of 2 four bit integers that define the ratio, values of zero are not allowed. For example, the value 00010010, would represent an up-sample ratio of 1:2. The meaning is illustrated in the figure below. When spatial_scalable_mode is snr_scalability or chrominance_scalability this should have the value "00010001".

Table defining spatial_scalable_mode

codeword	spatial_scalable_mode	macroblock_type tables
00	reserved	
01	spatial scalability	B.2-4, B.2-5, B.2-6, B.2-11, B.2-12, B.2-13
10	snr scalability	B.2-7, B.2-14
11	chrominance scalability	To be defined

vertical_subsampling_ratio - The vertical_subsampling_ratio is an 8 bit integer that defines the ratio with which to up-sample the lower layer coded picture to generate the spatial prediction. It is composed of 2 four bit integers that define the ratio, values of zero are not allowed. For example, the value 00110101, would represent an up-sample ratio of 3:5, which may be used for CIF to NTSC up-sampling. The meaning is illustrated in the figure below. When spatial_scalable_mode is snr_scalability or chrominance_scalability this should have the value "00010001".

low_resolution_horizontal_offset - A signed integer that defines the offset of the upsampled lower layer coded picture relative to the current picture, for spatial prediction. The meaning is illustrated in the figure below. When spatial_scalable_mode is snr_scalability or chrominance_scalability this should have the value 0.

low_resolution_vertical_offset - A signed integer that defines the offset of the upsampled lower layer coded picture relative to the current picture, for spatial prediction. The meaning is illustrated in the figure below. When spatial_scalable_mode is snr_scalability or chrominance_scalability this should have the value 0.

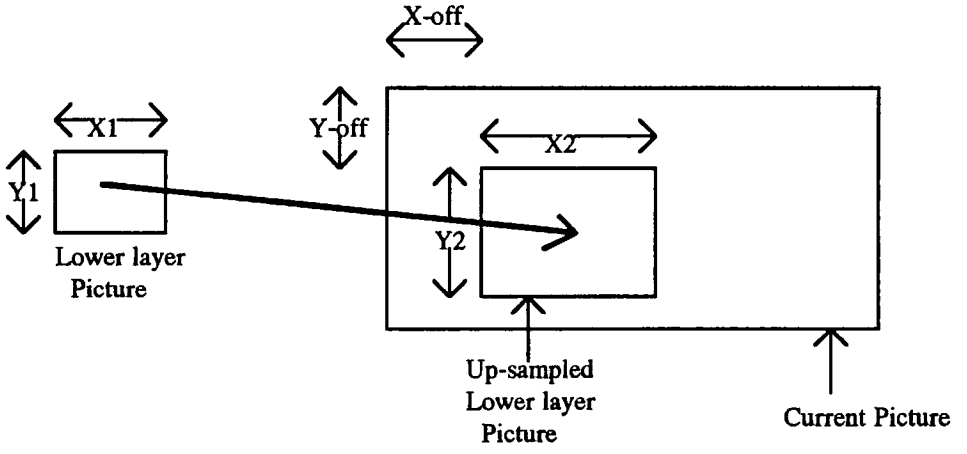

Diagram showing specification of up-sampling.
X-off and Y-off are the horizontal and vertical offsets,
X1:X2, Y1:Y2 are the subsampling ratios.

Specification of spatial scalable up-sampling

This section describes how the parameters in the picture spatial scalable extension are used to generate the spatial prediction from the coded lower layer picture.

When the lower layer picture is progressive the following procedure is used.

The low_resolution_field1_horizontal_offset and low_resolution_field1_vertical_offset are used to offset the top left luminance pel of the lower layer picture onto a luminance pel of the current picture, relative to the top left of the current picture. Positive offsets are measured to the right and downwards.

Bilinear interpolation in the horizontal and vertical direction is performed on the pels of the coded lower layer picture using the horizontal_subsampling_ratio and the vertical_subsampling_ratio. Rounding is

performed to the nearest integer. Pels that can not be calculated from interpolation of the lower layer picture because it does not cover the entire area of the current picture are set to a value of 128.

When the lower layer picture is interlaced the following procedure is used.

The lower layer frame is firstly padded with zeros to form a progressive grid at a picture rate equal to the field rate of the lower layer, and with the same number of lines and pels per line as the lower layer frame. It is then filtered using the relevant two field aperture filter from the table below.

Temporal	Vertical	Filter for first field	Filter for second field
-1	-2	0	-1/16
-1	0	0	1/8
-1	2	0	-1/16
0	-1	1/2	1/2
0	0	1	1
0	1	1/2	1/2
1	-2	-1/16	0
1	0	1/8	0
1	+2	-1/16	0

The resulting progressive fields are then up-sampled as described above for progressive lower layer pictures using bi-linear interpolation using the upsampling_ratio's, and offset relative to the current picture using the horizontal and vertical offsets.

If the current picture is interlaced, the appropriate lines should be selected from these progressive fields. If the current picture is progressive, all the lines are retained. If the current picture is a field picture, the appropriate lines should be selected from the appropriate progressive field.

Notes. I think that two sets of offsets are only required when the lower layer is interlaced. But if these offsets are different, I do not know how to perform the upsampling, particularly when the vertical offsets are different. Also, in the WD, the picture pan scan extension only defines one set of offsets per picture.

Macroblock layer

snr_block(i) {	No. of bits	Mnemonic
while (nextbits != end_of_block) {		
dct_coef_next	3-28	vlcbf
}		
end_of_block	2	"10"
}		

macroblock() {	No. of bits	Mnemonic
if (<sequence extension was not present>)		
while (nextbits() == '0000 0001 111')		
macroblock_stuffing	11	vlclbf
while (nextbits() == '0000 0001 000')		
macroblock_escape	11	vlclbf
macroblock_address_increment	1-11	vlclbf
macroblock_type	1-13	vlclbf
if ((picture_structure == 'frame') &&		
(frame_pred_frame_dct == 0) &&		
(macroblock_intra || macroblock_pattern))		
dct_type	1	uimsbf
if (macroblock_quant)		
quantizer_scale_code	5	uimsbf
if (macroblock_motion_forward ||		
(macroblock_intra && concealment_motion_vectors))		
forward_motion_vectors()	...	...
if (macroblock_motion_backward)		
backward_motion_vectors()	...	...
if (macroblock_intra && concealment_motion_vectors)		
marker_bit	1	
if (macroblock_pattern)		
coded_block_pattern()	...	...
for (i=0; i<block_count; i++) {		
if(spatial_scalable_mode==snr_scalable&&!macroblock_pattern){		
snr_block(i)		
} else {		
block(i)		
}		
}		
if (picture_coding_type == 4)		
end_of_macroblock	1	"1"
}		

macroblock_type -- Variable length coded indicator of the method of coding and content of the macroblock according to the tables indicated by the spatial_scalable_mode.

End of document