SONY

SNC-P1/P5/CS10/CS11/DF40/DF70/RZ25 series
Documentation of MPEG4 user data and JPEG application segment

version 1.0

Feb / 20 / 2009

Sony Corporation

1. General description

This document describes the specification of so-called "user data" and "application segment" which are inserted in its MPEG4 bitstream and motion JPEG bitstream of SONY Network Camera SNC-P1 / P5 / CS10 / CS11 / DF40 / DF70 / RZ25 series (hereafter called as SNC cameras).

NOTE:
This information may be changed without any announcement for the reason of upgrading products and so on.

2. Specification of the video bitstream in the SNC cameras

You can acquire the video bitstream from SNC cameras by using CGI commands listed as follows. There are several ways to acquire the video bitstream in accordance with its video mode and its streaming type. Refer to respective CGI command manual for the details.

<MPEG4 mode>

Acquisition of HTTP bitstream

```
http://ip_address/mpeg4
```

Acquisition of RTP(Unicast) bitstream

```
http://ip_address/mpeg4?UdpMode=unicast&UdpPort=<Port number>
```

Acquisition of RTP(Multicast) bitstream

```
http://ip_address/mpeg4?UdpMode=multicast
```

<JPEG mode>

Acquisition of HTTP bitstream

```
http://ip_address/mjpeg?speed=<parameter>
```

In terms of MPEG4 bitstream, there are three ways to acquire the video bitstream. Whatever the streaming type of acquisition is, you will be able to acquire the same video bitstream if you get rid of application layer headers such as HTTP or RTP.

Some camera-specific information e.g. date & time when the image is encoded, is inserted in the video bitstream of SNC cameras. When the video mode is set to MPEG4, the information is embedded in compliance with MPEG4 "user data". When the video mode is set to JPEG, the information is embedded in compliance with JPEG "application segment". The way of embedding

the information is explained in the following chapters.

3. User data in MPEG4 mode

In this chapter, the explanation of the syntax of the MPEG4 bitstream goes on as well as how the user data is inserted.

In terms of SNC cameras, one GOP (Group of pictures) which consists of one I-Frame(I-VOP) and the continuous P-Frame(V-VOP) will form a "Visual Object Sequence". Generally, the bitstream of the "Visual Object Sequence" will be ended by termination code of **visual_object_sequence_end_code**. But this termination code is not added in the case of SNC cameras.

The following shows the syntax of the bitstream of "Visual Object Sequence" and the syntax of each elements in the "Visual Object Sequence".

```
VisualObjectSequence() {
    do {
        visual_object_sequence_start_code   = "00 00 01 B0" (h)
        profile_and_level_indication        = "03" (h) : Simple Profile/ Level 3
        VisualObject()
    } while ( End_of_VisualObjectSequence )

}
```

Fig. 1 : Syntax of Visual Object Sequence

```
VisualObject() {
    visual_object_start_code          = "00 00 01 00"  (h)
    is_visual_object_identifier       = 1 (b)
        visual_object_verid           = 0001 (b)
        visual_object_priority        = 001 (b)
    visual_object_type                = 0001 (b) = "Video ID"
    video_signal_type()
    next_start_code()
    video_object_start_code           = "00 00 01 00" (h)
    VideoObjectLayer()
}
```

Fig. 2: Syntax of Visual Object

```
video_signal_type() {
    video_signal_type                 = 0
}
```

Fig. 3 : Syntax of video_signal_type

```
VideoObjectLayer() {
    video_object_layer_start_code          = "00 00 01 20" (h)
    random_accessible_vol                  1 bit
    video_object_type_indication           = 00000001(b) = Simple Object Type
        is_object_layer_identifier         = 0 (b)
    aspect_ratio_info                      = 0001 (b) = Square pixel
    vol_control_parameters                 = 0 (b)
    video_object_layer_shape               = 00 (b) = rectangular
    marker_bit                             = 1 (b)
    vop_time_increment_resolution          16 bits
    marker_bit                             = 1 (b)
    fixed_vop_rate                         = 0 (b)
        marker_bit                         = 1 (b)
        video_object_layer_width           13bits
        marker_bit                         = 1 (b)
        video_object_layer_height          13bits
        marker_bit                         = 1 (b)
    interlaced                             = 0 (b)
    obmc_disable                           = 1 bit
    sprite_enable                          = 0 (b)

    not_8_bit                              1 bit
    quant_type                             = 0 (b)

    complexity_estimation_disable          = 1 (b)
    resync_marker_disable                  = 1 (b)
    data_partitioned                       = 0 (b)

    scalability                            = 0 (b)

    next_start_code()

    do {
        user_data()
        VideoObjectPlane()                 -> I-VOP or P-VOP
    } while ( next_bits() == user_data_start_code || next_bits() == vop_start_code)

}
```

Fig. 4 : Syntax of Video Object Layer

As you can see the above figure, the user data of the MPEG4 bitstream in the SNC cameras will be emerged ahead of "Video Object Plane (I-VOP, P-VOP)". When you want to take this user data, you can achieve it by searching for the user_data_start_code ("00 00 01 B2" (h)) and then take the followed string.

The data format of MPEG4 user data

The following shows data format of MPEG4 user data.

Size	Description	Example		Remarks
Fixed	Date and time	CamTim: 2004-6-22 Tue 12:00:05\r\n	yyyy-mm-dd day hh:mm:ss	
Variable	Frame rate	FrmRate: 30\r\n		
Fixed	Time stamp	TimStamp: 0000110309\r\n	10 digits	
Fixed	PTZ position	CamPos:111pppptttttzzzzm\r\n	16 digits	Applicable for PTZ models (SNC-P5 / RZ25)
Fixed	Alarm event	AlmEvent: 0011\r\n	4 digits	Inserted when there is alarm

Fig. 5 : The data format of MPEG4 user data

Date and time information (fixed length)
This field shows the date and time when the image is encoded in the camera.

Frame rate information (variable length)
This field shows the current frame rate of the video bitstream.

Time stamp information(10 digits)
This field shows the time stamp when the image is encoded in the camera.

PTZ position
This field will only be inserted for PTZ control capable models such SNC-P5 and SNC-RZ25. First 3 digits are fixed as "111". Following 12 digits indicatates each position of Pan and Tilt and Zoom that is represented by hexadecimal 4 digits. The last digit indicates whether the camera PTZ is moving or not. "M" indicates that it is moving, while "S" indicates it is not moving.

Alarm event information (fixed length. will be inserted when there is some alarm events)
This field will be inserted only when there are some alarms detected. This field shows there is some alarm events such as motion detection or sensor input.

* Alarm event field
In the alarm event field, 4 digits will be displayed, each of which is "0" or "1". Each digit matches the following alarm event. When some digit is set to "1", that shows that matched alarm event occurs.

AlmEvent: 0 0 1 1\r\n

Reserved   Sensor input   Motion detection
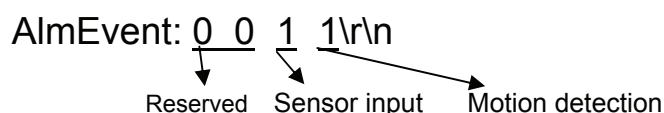
4. Application segment in JPEG mode

In this chapter, the explanation of the motion JPEG bitstream goes on as well as how the application segment is inserted.

When the video mode is set to JPEG, the video bitstream is output in the form of HTTP "server-push". The HTTP "server-push" indicates each of the encoded JPEG image will be separated from the other images by HTTP server-push header. The following figure shows the example of the motion JPEG bitstream.

```
HTTP/1.0 200 OK\r\n
Content-Type: multipart/x-mixed-replace;boundary=--myboundary\r\n
--myboundary\r\n
Content-Type: image/jpeg\r\n
CamTim: 2004-05-18 Tue 10:13:05\r\n
\r\n
<JPEG image data>\r\n
--myboundary\r\n
Content-Type: image/jpeg\r\n
CamTim: 2004-05-18 Tue 10:13:05\r\n
\r\n
<JPEG image data>\r\n
--myboundary\r\n
Content-Type: image/jpeg\r\n
CamTim: 2004-05-18 Tue 10:13:06\r\n
\r\n
<JPEG image data>\r\n
--myboundary\r\n
        .
        .
```

Fig. 6 : Example of motion JPEG bitstream

About an application segment in JPEG

The part of < JPEG image data > stands for actual JPEG image data segment. This < JPEG image data > field has an application segment in which there is camera-specific information as follows.

The application segment is to be inserted between SOI (Start of Image) marker and DQT (Define Quantization Table) marker.

Marker Name		Marker Code(Hex)	Description
SOI	Start of Image	FF D8	Indicates starts of image
APPn	Application Segment	FF E0 - FF EF	For applications
DQT	Define Quantization Table	FF DB	Definition of quantization table
DHT	Define Huffman Table	FF C4	Definition of Huffman table
SOF	Start of Frame	FF C0	Indicates start of frame
SOS	Start of Scan	FF DA	Indicates start of scan data
EOI	End of Image	FF D9	Indicates end of image

Fig. 7 : JPEG marker description

The application segment used in the SNC cameras will start with its application marker as "FF E2" (h). The following table describes the data format of the application segment.

Size	Description	Example		Remarks
2	Marker	FF E2	comment marker	
2	Length of field	00 nn	nn byte (variable)	
Fixed	Jpeg data length	DataLen : 00012345\r\n	8 digits	
Fixed	Date and time	CamTim: 2004-6-22 Tue 12:00:05\r\n	yyyy-mm-dd day hh:mm:ss	
Variable	Frame rate	FrmRate: 30\r\n		
Fixed	Time stamp	TimStamp: 0000110309\r\n	10 digits	
Fixed	PTZ position	CamPos:111ppppttttzzzzm\r\n	16 digits	Applicable for PTZ models (SNC-P5 / RZ25)
Fixed	Alarm event	AlmEvent: 0011\r\n	4 digits	Inserted when there is alarm

Fig. 8 : The data format of the application segment

Marker (2bytes)
　　Application marker which indicates the start of the application segment. This value is specified in the JPEG standard. In the SNC cameras, the value is set to "FF E2"(h).

Marker length (2bytes)
　　Indicates the entire length of the segment except the marker.

Jpeg data length(fixed length)
　　Indicates the length of the < JPEG image data >.

Date and time information(fixed length)
　　This field shows the date and time when the image is encoded in the camera.

Frame rate information (variable length)
　　This field shows the current frame rate of the JPEG bitstream.

Time stamp information(10 digits)
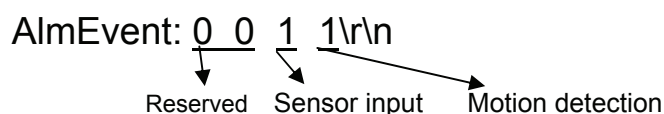　　This field shows the time stamp when the image is encoded in the camera.

PTZ position
　　This field will only be inserted for PTZ control capable models such SNC-P5 and SNC-RZ25. First 3 digits are fixed as "111". Following 12 digits indicatates each position of Pan and Tilt and Zoom that is represented by hexadecimal 4 digits. The last digit indicates whether the camera PTZ is moving or not. "M" indicates that it is moving, while "S" indicates it is not moving.

Alarm event information (fixed length. will be inserted when there is some alarm events)
　　This field shows there is some alarm events such as motion detection or sensor input.

* Alarm event field
　　In the alarm event field, 4 digits will be displayed, each of which is "0" or "1". Each digit matches the following alarm event. When some digit is set to "1", that shows that matched alarm event occurs.

AlmEvent: 0　0　1　1\r\n

Reserved　Sensor input　Motion detection

Revision history

Version	Date	Comment
1.0	Feb/20/2009	First issue