

mmusic  
Internet-Draft  
Intended status: Standards Track  
Expires: April 25, 2013

B. Greevenbosch  
Y. Hui  
Huawei Technologies  
October 22, 2012

SDP attribute to signal parallax  
draft-greevenbosch-mmusic-sdp-parallax-01

Abstract

This document introduces a "ParallaxInfo" attribute in SDP. This attribute can be used in stereoscopic applications, to signal the depth positioning of 2D media data within the 3D space.

Internet-Draft

SDP attribute to signal parallax

October 2012

#### Note

Discussion and suggestions for improvement are requested, and should be sent to [mmusic@ietf.org](mailto:mmusic@ietf.org).

#### Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 25, 2013.

#### Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Internet-Draft

SDP attribute to signal parallax

October 2012

Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">4</a>
<a href="#">2.</a>	Requirements notation . . . . .	<a href="#">5</a>
<a href="#">3.</a>	Definitions . . . . .	<a href="#">6</a>
<a href="#">4.</a>	The ParallaxInfo attribute . . . . .	<a href="#">7</a>
<a href="#">5.</a>	Example . . . . .	<a href="#">9</a>
<a href="#">6.</a>	Remarks . . . . .	<a href="#">10</a>
<a href="#">7.</a>	Security Considerations . . . . .	<a href="#">12</a>
<a href="#">8.</a>	IANA Considerations . . . . .	<a href="#">13</a>
<a href="#">9.</a>	Normative References . . . . .	<a href="#">14</a>
	Authors' Addresses . . . . .	<a href="#">15</a>

## 1. Introduction

To see a 3D scene, the human brain interprets two different views as perceived by the left and right eyes, and fuses these views into a single 3D perception. The depth of the object is perceived by interpreting the horizontal shift of that object between the views. This shift is called "parallax".

In stereoscopic 3D multimedia applications, there are various ways to transmit media streams in 3D. One way is to transmit two different streams, one for the left eye and one for the right eye. These streams are then projected to the relevant eyes using the appropriate technology.

When sending text streams in 3D, the solution mentioned above would imply sending the same text stream twice. Since the two streams would only differ in horizontal positioning, this introduces a lot of unnecessary overhead.

This document specifies a "ParallaxInfo" attribute in SDP [[RFC4566](#)], which is used to transfer the parallax information. It eliminates the need to send two different streams separately, as they can be calculated from a single stream and the "ParallaxInfo" attribute value.

The attribute transfers this information as two parameters: one indicating which view (left/right/centre) is carried by the stream, and another to signal the parallax between the objects.

The attribute is not restricted for signalling the parallax of text streams, but it can also be used to place other 2D objects in the 3D space. Examples include a channel logo, an electronic programme guide and on-screen display of an audio volume indicator.

## 2. Requirements notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[RFC2119](#)].

### [3.](#) Definitions

#### L-view

A visual entity that is to be projected to the left eye. In the case of video, the L-view is a video frame designated for the left eye. In the case of text, the L-view is the text positioned for viewing by the left eye.

#### L-stream

A sequence of L-views, which is streamed to the device.

#### R-view

A visual entity that is to be projected to the right eye. In the case of video, the R-view is a video frame designated for the right eye. In the case of text, the R-view is the text positioned

for viewing by the right eye.

#### R-stream

A sequence of R-views, which is streamed to the device.

#### C-view

The centre view: a visual entity as seen from a viewpoint between the left and right eyes. The C-view can be used to calculate the L- and R-views.

#### C-stream

A sequence of C-views, which is streamed to the device.

#### stereoscopic device

A device that is able to produce and display different images for the left and right eyes, such that the viewer can experience a 3D view.

#### 2D device

A device that is not able to produce and display different images for the left and right eyes.

#### 2D media stream

A sequence of two dimensional visual entities (such as text or 2D graphics), which is streamed to the device.

#### [4.](#) The ParallaxInfo attribute

The SDP attribute "ParallaxInfo" is used to transmit the depth positioning of 2D media data, such as a text stream, in the 3D space.

The attribute has the following syntax:

```
a=ParallaxInfo:<transmitted position> <parallax>
```

The <transmitted position> indicates whether the L-, C- or R-stream is transmitted, whereas <parallax> indicates the parallax (i.e. shift) between corresponding L- and R-views in pixels, normalised to a screen width of 11520 pixels. To convert the value of <parallax> to match the display video screen width  $W$ , it has to be divided by a factor  $F=11520/W$ .

The <transmitted position> can have the following values:

"L" indicates that the transmitted stream is the L-stream. A stereoscopic device MUST calculate the corresponding R-views by shifting the L-views  $\langle\text{parallax}\rangle/F$  pixels towards the right.

"C" indicates that the transmitted stream is the C-stream. A stereoscopic device MUST calculate the corresponding L-views by shifting the C-views  $\langle\text{parallax}\rangle/(2*F)$  pixels towards the left, and the R-views by shifting the C-views  $\langle\text{parallax}\rangle/(2*F)$  pixels towards the right. <parallax> SHOULD be even. If it is odd, the divided value MUST be rounded off towards zero.

"R" indicates that the transmitted stream is the R-stream. A stereoscopic device MUST calculate the corresponding L-views by shifting the R-views  $\langle\text{parallax}\rangle/F$  pixels towards the left.

<parallax> MAY be negative. In this case, the shift direction is reversed.

The "ParallaxInfo" attribute can be a session-level attribute or a media-level attribute. As a session-level attribute, it specifies the default parallax value which can be applied to all the 2D media streams in the session being described. As a media-level attribute, it specifies the parallax value which can be applied to the associated 2D media stream, overriding any session-level parallax value specified.

The stereoscopic device MAY use the session-level attribute value for on-screen display, for example an audio volume indicaton, channel indication or electronic programme guide.



attribute will ignore it, and therefore display the 2D media data on the position as transmitted.

## 5. Example

The following is an example of an SDP description of a session with an audio stream, a video stream and a 3GPP timed text stream (see [3gpp-tt]), streamed using RTP as per [RFC4396]. If the display resolution is 1280x720, the parallax is scaled down with a factor  $F=11952/1280=9$ . The transmitted text stream is the L-stream, and with the example display resolution each R-view is  $144/9=16$  pixels on the left of the L-view. This corresponds to the virtual positioning of the text in front of the screen.

```
v=0
o=Alice 2890844526 2890842807 IN IP4 131.163.72.4
s=The technology of 3D-TV
c=IN IP4 131.164.74.2
t=0 0
a=ParallaxInfo:L -180
m=video 49170 RTP/AVP 99
a=rtpmap:99 H264/90000
m=video 52888 RTP/AVP 97
a=rtpmap:97 3gpp-tt/1000
a=ParallaxInfo:L -144
m=audio 52890 RTP/AVP 10
a=rtpmap:10 L16/16000/2
```

Notice that the default value "-180" is overridden by the value "-144" for the text stream. However the "-180" value is still signalled for on-screen display of e.g. volume control and other 2D graphics.

In case each R-view is 24 pixels (216 pixels in the reference resolution) on the right of the associated L-view, i.e. the virtual positioning of the text is behind the screen, then the three lines defined for 3gpp-tt can be replaced as follows:

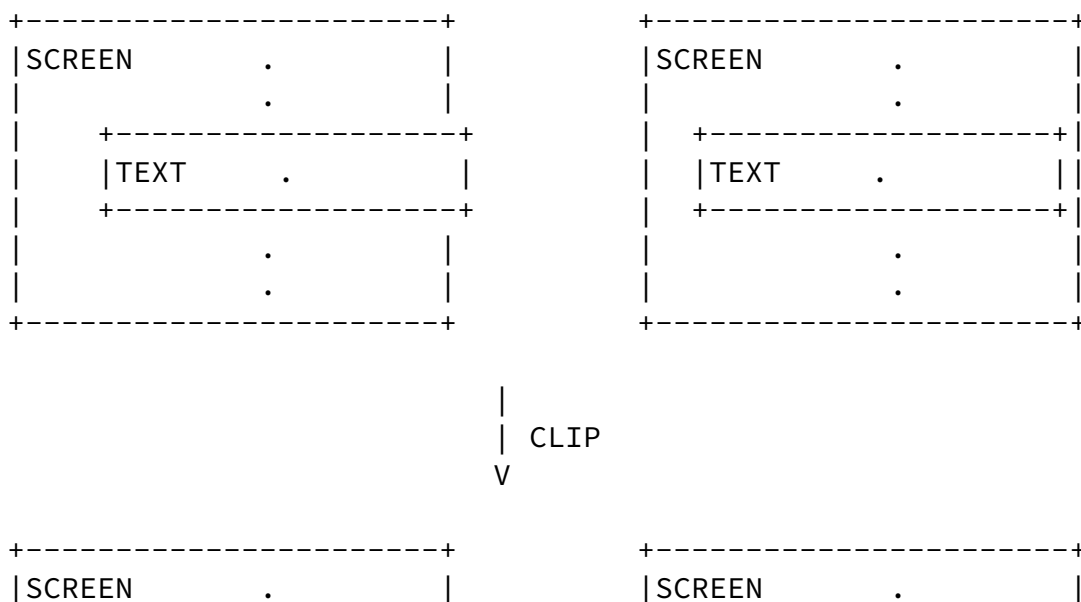
```
m=video 52888 RTP/AVP 97
a=rtpmap:97 3gpp-tt/1000
a=ParallaxInfo:L 216
```

## 6. Remarks

A positive parallax value indicates virtual positioning of the 2D media data behind the screen. This is the case when the objects in the L-view are on the left of the same objects in the R-view. Similarly, a negative parallax value indicates that the objects in the R-view are on the left of the same objects in the L-view, and corresponds to virtual positioning in front of the screen.

Since the "ParallaxInfo" attribute indicates a shift of the transmitted stream, it might happen that the L- or R-view trespasses the boundaries of the display. In this case, clipping is necessary, as illustrated by Figure 1.

Similarly, there are areas which are covered by the L-view but not by the R-view and vice versa. These areas need to be filled in a sensible way. Since the "ParallaxInfo" attribute is designed for objects that overlay other video data (e.g. subtitles), it is trivial to fill in uncovered areas by using the underlying video data. However, if there is no underlying video data, other mechanisms to fill in the uncovered areas need to be defined. Definition of these mechanisms are out of scope of this document.



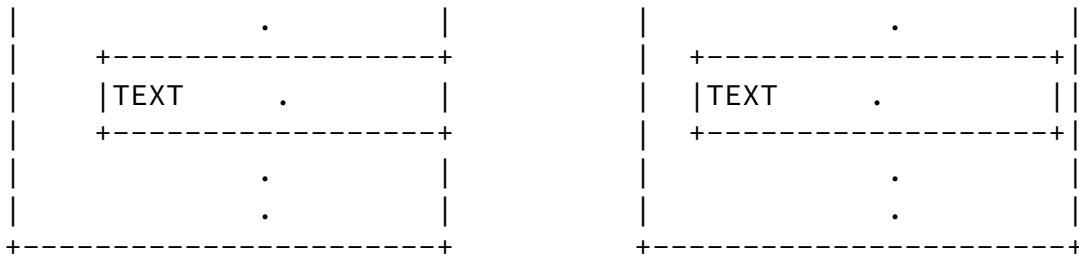


Figure 1

The normalisation of the parallax towards a reference screen width of 11520 pixels was chosen to ensure correct display of the same stream on different video display resolutions. This is especially useful when the underlying video is coded with a scalable video codec, which does not have a fixed video resolution.

## 7. Security Considerations

The authors foresee no security issues in addition to those already listed in [[RFC4566](#)].

## 8. IANA Considerations

Following the guidelines in [[RFC4566](#)], the SDP attribute has to be registered at IANA:

- o Contact name/email: authors of this RFC
- o Attribute name: ParallaxInfo
- o Long-form attribute name: Parallax info for the depth positioning of 2D media data in the 3D space
- o Type of attribute: session level and media level
- o Subject to charset: no

This attribute is used to signal how 2D media data is to be displayed in the 3D space. It indicates the shift of the respective left and

right views.

The attribute has the following ABNF (see [[RFC4234](#)]) description:

```
ParallaxInfo          = "a=ParallaxInfo:" TransmittedPosition Parallax
TransmittedPosition  = "C"/"L"/"R"
Parallax              = num-val
```

The attribute transfers this information as two parameters:  
"TransmittedPosition" indicates which view of the 2D media data (left "L", right "R" or centre "C") is carried by the stream, and  
"Parallax" signals the parallax (in pixels) of objects in the 2D media stream.

## 9. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC4234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", [RFC 4234](#), October 2005.
- [RFC4396] Rey, J. and Y. Matsui, "RTP Payload Format for 3rd Generation Partnership Project (3GPP) Timed Text", [RFC 4396](#), February 2006.

- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", [RFC 4566](#), July 2006.
- [3gpp-tt] 3GPP, "Transparent end-to-end packet switched streaming service (PSS); Protocols and codecs (Release 5)", TS 26.234 v5.3.0, December 2002.

#### Authors' Addresses

Bert Greevenbosch  
Huawei Technologies Co., Ltd.



Huawei Industrial Base  
Bantian, Longgang District  
Shenzhen 518129  
P.R. China

Phone: +86-755-28978088  
Email: bert.greevenbosch@huawei.com

Hui Yu  
Huawei Technologies Co., Ltd.  
Huawei Nanjing R&D Center  
101 Software Avenue  
Yuhuatai District  
Nanjing 210012  
P.R. China

Phone: +86-25-56620323  
Email: huiyu@huawei.com