

AVTEXT Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 2015

J. Samuelsson
Ericsson
M. Coban
Qualcomm
S. Wenger
Vidyo
December 15, 2014

**Reference Picture Verification Information in the
RTP Audio-Visual Profile with Feedback (AVPF)
draft-samuelsson-avtext-rpvi-00.txt**

Abstract

This document specifies an extension to the feedback messages defined in the Audio-Visual Profile with Feedback (AVPF). The new Reference Picture Verification Information (RPVI) feedback message conveys information about available reference pictures in the decoded picture buffer of a video decoder in the receiver of an RTP video stream.

By including information related to Decoded Picture Hash (DPH) values, media senders and media receivers can verify that reference pictures used for prediction by the video encoder and the video decoder are aligned. It is also possible to use the RPVI feedback message to indicate that a specific reference picture has incorrect sample values (i.e. a mismatch in the DPH value between encoder and decoder) or that a specific reference picture has been lost.

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on June 15, 2015.

Copyright Notice

Copyright (c) 2014 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.

Table of Contents

1.	Introduction.....	2
1.1.	Applicability.....	3
2.	Terminology.....	4
2.1.	Standards Language.....	4
2.2.	Glossary.....	4
3.	Reference Picture Verification Information.....	4
3.1.	Message Format.....	6
4.	SDP Signaling.....	8
5.	Security Considerations.....	9
6.	IANA Considerations.....	9
7.	References.....	9
7.1.	Normative References.....	9
7.2.	Informative References.....	10
8.	Acknowledgments.....	10

[1. Introduction](#)

This document defines a new RTCP feedback message to augment those defined in [[RFC4585](#)], [[RFC5104](#)] and [[RFC6642](#)], for use together with

video codecs that exploits temporal prediction through the use of one or more reference pictures, e.g. [H.264], VP8 [RFC6386] and [HEVC].

1.1. Applicability

The video codecs [H.264] and [HEVC] both use temporal prediction in order to achieve efficient compression without compromising the visual quality of the compressed video. Video data (frames/pictures) are encoded together with non-video data (such as parameter sets) and an abstraction layer is used to structure the encoded bits in a format suitable for network transportation.

A stream encoded according to H.264 or HEVC, and packetized according to [RFC6184] and [I-D.ietf-payload-rtp-h265], respectively, is typically transmitted from a media sender to a media receiver. The media sender encodes the video and the media receiver decodes the video. During the entire session (or, more specifically, within a coded video sequence, it is crucial that the process performed at the decoder is aligned with the process performed at the encoder. Even the slightest difference in the sample values of a decoded picture can result in severe visual degradation when the picture is used for prediction by following pictures.

There are several factors that can affect the alignment of encoding and decoding processes:

- o Loss of data. In many applications it is possible to detect the loss of RTP packets and perform appropriate actions for repairing the loss without delivering corrupt data to the video decoder. However, in some applications such methods may not be available (for example due to delay constraints) or they may fail.
- o Bit errors. If the receiver does not have means for detecting individual bit errors, such errors may occur in the data that is delivered to the video decoder.
- o Random access. When performing random access into a stream it might be difficult for the decoder to deduce if it is operating with the correct parameters and reference pictures.
- o Hardware failure. The hardware in the decoder could be malfunctioning, for example if it is not able to correctly store decoded pictures used for prediction.

- o Incorrect implementations. Ideally all video encoders and video decoders would be implemented impeccably according to the codec specification. However, in practice there is unfortunately the risk of misinterpretation of the specification as well as the risk of implementation bugs.

The feedback message specified in this memo can be utilized to detect misalignment between encoder and decoder reference pictures. Other mechanisms (such as sending IDR pictures) not specified herein, can be utilized to combat the potential negative effects of an encoder/decoder misalignment.

2. Terminology

2.1. Standards Language

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 [RFC 2119](#) [[RFC2119](#)].

2.2. Glossary

AVPF - Audio-Visual Profile with Feedback

DPH - Decoded Picture Hash

FCI - Feedback Control Information [[RFC4585](#)]

IDR - Instantaneous Decoder Refresh

RPVI - Reference Picture Verification Information

SEI - Supplemental Enhancement Information

3. Reference Picture Verification Information

A Reference Picture Verification Information (RPVI) feedback message can be sent by media receivers to report which reference pictures are available in the decoded picture buffer. Along with identifiers of the available reference pictures it is possible to transmit the result of verifying the Decoded Picture Hash (DPH) values or to transmit the actual DPH values (see [section 3.1](#)). The feedback message can be sent at any time during an RTP session. This memo does

not describe the process for handling incorrect DPH values. However, in order to achieve good media quality and recover from errors in the sample values of decoded pictures it is strongly recommended that a media sender (encoder) takes appropriate actions upon the detection of an incorrect DPH value or negative acknowledgements (NACK). Such actions could for example include:

- o Transmission of data that resets the state of the decoder, e.g. an Instantaneous Decoder Refresh (IDR) picture. By providing a refresh-point, the media sender can ensure that errors that have occurred in decoded reference pictures do not propagate to future pictures.
- o Encoding following pictures using "old" reference pictures that have been received, decoded and preferably verified to have correct sample values. Excluding all references to pictures with incorrect sample values will give the same effect as providing a refresh-point: errors that are present in decoded reference pictures do not propagate to future pictures.
- o Retransmission of parameter sets. If an update of parameter sets is lost, there is a risk that the decoder uses some parameters incorrectly (e.g. too strong deblocking filter) without detectable errors in the decoding process. By retransmitting the parameter sets the encoder can make sure that the correct parameters are used but it is not by its own sufficient for recovering from errors in sample values of decoded reference pictures. This action is recommended to be combined with one of the first two actions in this list.
- o Changing encoder settings or parameters to avoid configurations that cause incorrect decoder state. When errors continuously appear (even after performing one or both of the first two actions in this list) a media sender can try to change the configuration of the encoder in order to find a setting that does not result in errors in the decoded pictures.

3.1. Message Format

The RPVI message is identified by RTCP packet type value PT=PSFB and FMT=TBD. The Feedback Control Information (FCI) for RPVI consists of one or more FCI entries, the content of which is depicted in Figure 1. Each entry applies to a different reference picture, identified by its Reference Picture Identifier.

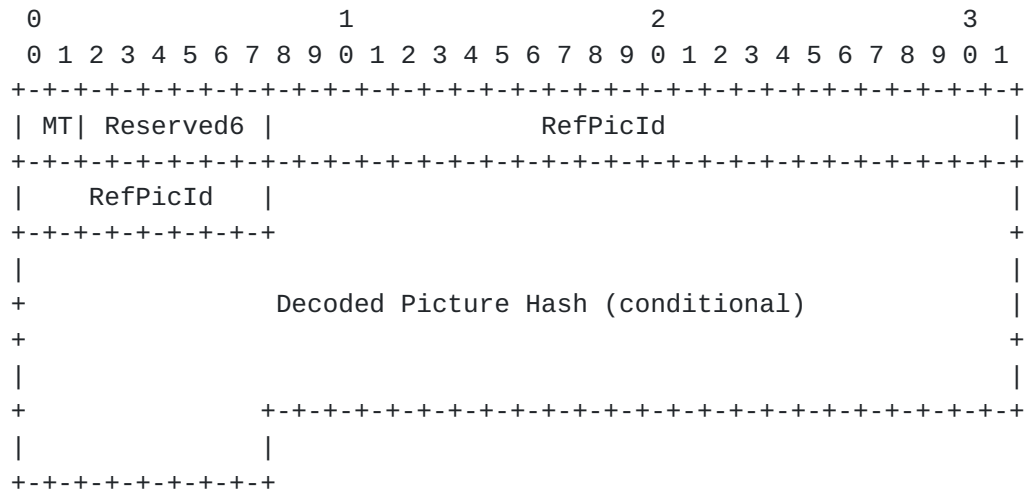


Figure 1 Syntax of an FCI Entry in the RPVI message

The semantics of the fields are as follows:

MT: 2 bits

Indicates the picture status information as follows:

- | | |
|----|---|
| 0: | No hash information regarding the correctness of the reference picture is available. |
| 1: | The Decoded Picture Hash of the reference picture is included in the Reference Picture Description. |
| 2: | The indicated picture is entirely or partially lost, hence not fully decodable. |
| 3: | The Decoded Picture Hash has been used to verify the reference picture to be incorrect. |

When MT equals 0 or 1, the reference picture identified by the current entry is indicated as being available at receiver's decoded picture buffer which may be available at the sender's decoded picture buffer for reference when encoding the next picture to be encoded at the reception of the RPVI feedback message. For MT equals to 1 with the exception that if the encoder finds that the provided hash of the reference picture does not match the encoder's hash value, then it MUST NOT use the reference picture.

Informative note: When a feedback message contains one or more RPVI entry with MT equals to 0 or 1, the encoder may select one or more of the identified pictures and/or inferred reference pictures from the availability of the indicated pictures to be used for reference. The selection of which picture(s) to use for reference is out of scope of this memo but may for example be based on maximizing compression efficiency.

When MT equals 2 or 3 the reference picture identified by the current entry MUST NOT be used for reference for the next picture or any picture that follows the next picture. Other reference pictures that use the reference picture identified by the current entry SHOULD NOT be used for reference, unless their Decoded Picture Hash has been verified to be correct.

Reserved6: 6 bits

This field is reserved for future definition. In the absence of such a definition, the bits in this field MUST be set to zero and ignored by the receiver of the RPVI feedback message.

RefPicId: 32 bits

If the video codec used for the media stream is HEVC, RefPicId represents the value of the PicOrderCntVal (in network byte order) of the reference picture, as defined in [[HEVC](#)].

If the video codec used for the media stream is H.264, RefPicId represents the value of the frame_num (in network byte order) of the reference picture, as defined in [[H.264](#)].

If the video codec used for the media stream is neither HEVC nor H.264, the picture identifier RefPicId SHOULD be defined outside of this specification.

Decoded Picture Hash: Variable number of bytes

Present only if MT equals 1. Represent the Decoded Picture Hash Supplemental Enhancement Information (SEI) data (in network byte order), see D.2.19 of [\[HEVC\]](#), of the decoded picture. The Decoded Picture Hash data starts with a one byte type field, which can be used to calculate the amount of hash data. For video encoded with three color components, such as YCbCr and RGB, the total length of the Decoded Picture Hash will be 49 bytes when the first byte equals 0, 7 bytes when the first byte equals 1 and 13 bytes when the first byte equals 2.

Informative note: At the time of writing this memo, the Decoded Picture Hash SEI message is only specified for HEVC. However, the DPH calculations defined in D.3.19 of [\[HEVC\]](#) operate only on decoded sample values and is therefore codec agnostic. The DPH SEI message defined in D.2.19 of [\[HEVC\]](#) does not contain any HEVC specific information and can therefore easily be replicated in the context of any video codec that decode encoded data into arrays of sample values, such as H.264.

4. SDP Signaling

A new "ack" and "nack" feedback parameter "rpvi" is defined to indicate the usage of the RPVI feedback message.

(In the following ABNF [\[RFC5234\]](#), rtcp-fb-ack-param, rtcp-fb-nack-param is used as defined in [\[RFC4585\]](#).)

```
rtcp-fb-ack-param =/ SP "rpvi"
```

```
rtcp-fb-nack-param =/ SP "rpvi"
```

The following parameter is defined in this document for use with 'ack':

- o 'rpvi' stands for Reference Picture Verification Information and indicates the use of RPVI messages as defined in [Section 3](#).

The following parameter is defined in this document for use with 'nack':

- o 'rpvi' stands for Reference Picture Verification Information and indicates the use of RPVI messages as defined in [Section 3](#).

The offer/answer rules for these SDP feedback parameters are specified in the RTP/AVPF profile [[RFC4585](#)].

Methods and rules for when to send RPVI messages are out of scope of this memo. When the RPVI message is used in "ack" mode it may for example be sent at a regular interval or for all pictures that fulfills certain requirements (such as being coded as Intra pictures). However, it is possible in both "ack" mode and "nack" mode to send the RPVI message in response to a specific event (such as a picture loss). When the "ack" mode is used for MT equal to 2 or 3 it can be said to represent an acknowledgement of having received enough data to derive the PictureID of the indicated picture but that there appears to be some data missing (MT equal to 2) or the sample values seems to be incorrect (MT equal to 3).

5. Security Considerations

The security considerations documented in [[RFC4585](#)] are also applicable for the RPVI message defined in this document.

More specifically, a malicious group member can report incorrect DPH values in RPVI feedback messages to make the sender throttle the data transmission and increase the amount of redundancy information or take other action to deal with the pretended incorrect DPH value (e.g. change encoder configuration). This may result in a degradation of the quality of the reproduced media stream.

A solution to prevent such attack with maliciously sent RPVI feedback messages is to apply an authentication and integrity protection framework for the feedback messages. This can be accomplished using the RTP profile that combines Secure RTP [[RFC3711](#)] and AVPF into SAVPF [[RFC5124](#)].

6. IANA Considerations

A new RPVI Feedback Message Type should be registered with IANA in "FMT Values for PSFB Payload Types".

7. References

7.1. Normative References

[H.264] ITU-T Recommendation H.264, "Advanced video coding for generic audiovisual services", February 2014, <<http://www.itu.int/rec/T-REC-H.264-201402-P>>.

- [HEVC] ITU-T Recommendation H.265, "High Efficiency Video Coding", April 2013, <<http://www.itu.int/rec/T-REC-H.265-201304-I>>.
- [RFC4585] Ott, J., Wenger, S., Sato, N., Burmeister, C., and J. Rey, "Extended RTP Profile for Real-Time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", [RFC 4585](#), July 2006.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", [RFC 3711](#), March 2004.
- [RFC5124] Ott, J. and E. Carrara, "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)", [RFC 5124](#), February 2008.

7.2. Informative References

- [RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", [RFC 5104](#), February 2008.
- [RFC6642] Wu, Q., Xia, F., and R. Even, "RTP Control Protocol (RTCP) Extension for a Third-Party Loss Report", [RFC 6642](#), June 2012.
- [RFC6184] Wang, Y., Even, R., Kristensen, T., and R. Jesup, "RTP Payload Format for H.264 Video", [RFC 6184](#), May 2011.
- [I-D.ietf-payload-rtp-h265]
Wang, Y., Sanchez, Y., Schierl, T., Wenger, S. and M. Hannuksela, "RTP Payload Format for High Efficiency Video Coding", [draft-ietf-payload-rtp-h265](#) (work in progress), August 2014.

8. Acknowledgments

The authors would like to thank Bo Burman, Rickard Sjöberg and Magnus Westerlund for valuable feedback during the development of this memo.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Jonatan Samuelsson
Ericsson
Farogatan 6, 164 80, Stockholm, Sweden
Phone: +46 761 26 35 91
Email: jonatan.samuelsson@ericsson.com

Muhammed Coban
Qualcomm
Email: mcoban@qti.qualcomm.com

Stephan Wenger
Vidyo
Email: stewe@stewe.org