WebRTC Video Processing and Codec Requirements
draft-ietf-rtcweb-video-01

Abstract

   This specification provides the requirements and considerations for
   WebRTC applications to send and receive video across a network. It
   specifies the video processing that is required, as well as video
   codecs and their parameters.

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 30, 2015.

Copyright Notice

Table of Contents

1. Introduction

   One of the major functions of WebRTC endpoints is the ability to send
   and receive interactive video. The video might come from a camera, a
   screen recording, a stored file, or some other source. This
   specification defines how the video is used and discusses special
   considerations for processing the video. It also covers the video-
   related algorithms WebRTC devices need to support.

   Note that this document only discusses those issues dealing with
   video codec handling. Issues that are related to transport of media
   streams across the network are specified in
   [I-D.ietf-rtcweb-rtp-usage].

2. Terminology

   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. Pre and Post Processing

   This section provides guidance on pre- or post-processing of video
   streams.

   Unless specified otherwise by the SDP or codec, the color space
   SHOULD be sRGB [SRGB].

TODO: I'm just throwing this out there to see if a specific proposal, even if wrong, might draw more comment than "TBD". If you don't like sRGB for this purpose, comment on the rtcweb@ietf.org mailing list. It has been suggested that the MPEG "Coding independent media description code points" specification [IEC23001-8] may have applicability here.

3.1. Camera Source Video

This document imposes no normative requirements on camera capture; however, implementors are encouraged to take advantage of the following features, if feasible for their platform:

o Automatic focus, if applicable for the camera in use

o Automatic white balance

o Automatic light level control

3.2. Screen Source Video

If the video source is some portion of a computer screen (e.g., desktop or application sharing), then the considerations in this section also apply.

Because screen-sourced video can change resolution (due to, e.g., window resizing and similar operations), WebRTC video recipients MUST be prepared to handle mid-stream resolution changes in a way that preserves their utility. Precise handling (e.g., resizing the element a video is rendered in versus scaling down the received stream; decisions around letter/pillarboxing) is left to the discretion of the application.

Additionally, attention is drawn to the requirements in [I-D.ietf-rtcweb-security-arch] section 5.2 and the considerations in [I-D.ietf-rtcweb-security] section 4.1.1.

TODO: Do we want to define additional metadata to indicate whether a stream is sourced from a camera versus a screen capture? This would allow the receiving party to tune, e.g., output filters. It would appear that H.263 has this kind of indicator built into its bitstream, but I found no analog in H.264 or VP8.

4. **Stream Orientation**

   In some circumstances - and notably those involving mobile devices -
   the orientation of the camera may not match the orientation used by
   the encoder. Of more importance, the orientation may change over the
   course of a call, requiring the receiver to change the orientation in
   which it renders the stream.

   While the sender may elect to simply change the pre-encoding
   orientation of frames, this may not be practical or efficient (in
   particular, in cases where the interface to the camera returns pre-
   compressed video frames). Note that the potential for this behavior
   adds another set of circumstances under which the resolution of a
   screen might change in the middle of a video stream, in addition to
   those mentioned under "Screen Sourced Video," above.

   To accommodate these circumstances, RTCWEB implementations SHOULD
   support generating and receiving the R0 and R1 bits of the
   Coordination of Video Orientation (CVO) mechanism described in
   section 7.4.5 of [TS26.114]. (TODO: Is "SHOULD support" the right
   level here?) They MAY support the other bits in the CVO extension,
   including the higher-resolution rotation bits.

   Further, some codecs support in-band signaling of orientation (for
   example, the SEI "Display Orientation" messages in H.264 and H.265).
   If CVO has been negotiated, then the sender MUST NOT make use of such
   codec-specific mechanisms. However, when support for CVO is not
   signaled in the SDP, then such implementations MAY make use of the
   codec-specific mechanisms instead.

5. **Codec-Specific Considerations**

   WebRTC endpoints are not required to support the codecs mentioned in
   this section.

   However, to foster interoperability between endpoints that have
   codecs in common, if they do support one of the listed codecs, then
   they need to meet the requirements specified in the subsection for
   that codec.

   SDP allows for codec-independent indication of preferred video
   resolutions using the mechanism described in [RFC6236]. If a
   recipient of video indicates a receiving resolution, the sender
   SHOULD accommodate this resolution, as the receiver may not be
   capable of handling higher resolutions.

Additionally, codecs may include codec-specific means of signaling
maximum receiver abilities with regards to resolution, frame rate,
and bitrate.

Unless otherwise signaled in SDP, recipients of video streams are
MUST be able to decode video at a rate of at least 20 fps at a
resolution of at least 320x240. These values are selected based on
the recommendations in [HSUP1].

Encoders are encouraged to support encoding media with at least the
same resolution and frame rates cited above.

5.1. VP8

If VP8, defined in [RFC6386], is supported, then the endpoint MUST
support the payload formats defined in [I-D.ietf-payload-vp8]. In
addition it MUST support the 'bilinear' and 'none' reconstruction
filters.

In addition to the [RFC6236] mechanism, H.264 encoders MUST limit the
streams they send to conform to the values indicated by receivers in
the corresponding max-fr and max-fs SDP attributes.

TODO: There have been claims that VP8 already requires supporting
both filters; if true, these do not need to be reiterated here.

5.2. H.264

If [H264] is supported, then the device MUST support the payload
formats defined in [RFC6184]. In addition, they MUST support
Constrained Baseline Profile Level 1.2, and they SHOULD support H.264
Constrained High Profile Level 1.3.

Implementations of the H.264 codec have utilized a wide variety of
optional parameters. To improve interoperability the following
parameter settings are specified:

packetization-mode: Packetization-mode 1 MUST be supported. Other
   modes MAY be negotiated and used.

profile-level-id: Implementations MUST include this parameter within
   SDP and SHOULD interpret it when receiving it.

max-mbps, max-smbps, max-fs, max-cpb, max-dpb, and max-br: These par
   ameters allow the implementation to specify that they can support
   certain features of H.264 at higher rates and values than those
   signalled by their level (set with profile-level-id).
   Implementations MAY include these parameters in their SDP, but

SHOULD interpret them when receiving them, allowing them to send
the highest quality of video possible.

sprop-parameter-sets: H.264 allows sequence and picture information
to be sent both in-band, and out-of-band. WebRTC implementations
MUST signal this information in-band; as a result, this parameter
will not be present in SDP.

TODO: Do we need to require the handling of specific SEI messages?
One example that has been raised is freeze-frame messages.

6. Mandatory to Implement Video Codec

Note: This section is here purely as a placeholder, as there is not
yet WG Consensus on Mandatory to Implement video codecs. The issue
is more complicated than may be immediately apparent to newcomers,
who are strongly encouraged to familiarize themselves with the
previous discussions on the topic before engaging on this issue.

The currently recorded working group consensus is that all
implementations MUST support a single, specified mandatory-to-
implement codec. The remaining decision point is a selection of this
single codec.

6.1. Temperature of Working Group

To capture the conversation so far, this section summarizes the
result of a straw poll that the working group undertook in December
2013 and January 2014. Respondents were asked to answer "Yes,"
"Acceptable," or "No" for each option. The options were collected
from the working group at large prior to the initiation of the straw
poll.

```
                                            Yes  Acc  No
                                            ---  ---  ---
 1. All entities MUST support H.264         48%  11%  41%
 2. All entities MUST support VP8           41%  17%  42%
 3. All entities MUST support both H.264 and VP8   9%  38%  53%
 4. Browsers MUST support both H.264 and VP8, other
    entities MUST support at least one of H.264
    and VP8                                 11%  34%  55%
 5. All entities MUST support at least one of
    H.264 and VP8                           10%  16%  74%
 6. All entities MUST support H.261          5%  23%  72%
 7. There is no MTI video codec             12%  30%  58%
 8. All entities MUST support H.261 and all entities
    MUST support at least one of H.264 and VP8   4%  28%  68%
 9. All entities MUST support Theora         7%  26%  67%
```

```
10. All entities MUST implement at least two of
    {VP8, H.264, H.261}                          5%  30%  65%
11. All entities MUST implement at least two of
    {VP8, H.264, H.263}                          5%  25%  70%
12. All entities MUST support decoding using both
    H.264 and VP8, and MUST support encoding using
    at least one of H.264 or VP8                 7%  20%  73%
13. All entities MUST support H.263              6%  19%  75%
14. All entities MUST implement at least two of
    {VP8, H.264, Theora}                         6%  27%  67%
15. All entities MUST support decoding using Theora 1%  15%  84%
16. All entities MUST support Motion JPEG        1%  25%  74%
```

7. Security Considerations

This specification does not introduce any new mechanisms or security
concerns beyond what the other documents it references. In WebRTC,
video is protected using DTLS/SRTP. A complete discussion of the
security can be found in [I-D.ietf-rtcweb-security] and
[I-D.ietf-rtcweb-security-arch]. Implementers should consider
whether the use of variable bit rate video codecs are appropriate for
their application based on [RFC6562].

8. IANA Considerations

This document requires no actions from IANA.

9. Acknowledgements

The authors would like to thank Gaelle Martin-Cocher, Stephan Wenger,
and Bernard Aboba for their detailed feedback and assistance with
this document. Thanks to Cullen Jennings for providing text and
review. This draft includes text from draft-cbran-rtcweb-codec.

10. References

10.1. Normative References

[H264]   ITU-T Recommendation H.264, "Advanced video coding for
          generic audiovisual services", April 2013.

[HSUP1]   ITU-T Recommendation H.Sup1, "Application profile - Sign
          language and lip-reading real-time conversation using low
          bit rate video communication", May 1999.

[I-D.ietf-payload-vp8]

          Westin, P., Lundin, H., Glover, M., Uberti, J., and F.
          Galligan, "RTP Payload Format for VP8 Video", draft-ietf-
          payload-vp8-11 (work in progress), February 2014.

   [IEC23001-8]
          ISO/IEC 23001-8:2013/DCOR1, "Coding independent media
          description code points", 2013.

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
          Requirement Levels", BCP 14, RFC 2119, March 1997.

   [RFC4175] Gharai, L. and C. Perkins, "RTP Payload Format for
          Uncompressed Video", RFC 4175, September 2005.

   [RFC4421] Perkins, C., "RTP Payload Format for Uncompressed Video:
          Additional Colour Sampling Modes", RFC 4421, February
          2006.

   [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.

   [RFC6184] Wang, Y.-K., Even, R., Kristensen, T., and R. Jesup, "RTP
          Payload Format for H.264 Video", RFC 6184, May 2011.

   [RFC6236] Johansson, I. and K. Jung, "Negotiation of Generic Image
          Attributes in the Session Description Protocol (SDP)", RFC
          6236, May 2011.

   [RFC6386] Bankoski, J., Koleszar, J., Quillio, L., Salonen, J.,
          Wilkins, P., and Y. Xu, "VP8 Data Format and Decoding
          Guide", RFC 6386, November 2011.

   [RFC6562] Perkins, C. and JM. Valin, "Guidelines for the Use of
          Variable Bit Rate Audio with Secure RTP", RFC 6562, March
          2012.

   [SRGB]   IEC 61966-2-1, "Multimedia systems and equipment - Colour
          measurement and management - Part 2-1: Colour management -
          Default RGB colour space - sRGB.", October 1999.

   [TS26.114]
          3GPP TS 26.114 V12.7.0, "3rd Generation Partnership
          Project; Technical Specification Group Services and System
          Aspects; IP Multimedia Subsystem (IMS); Multimedia
          Telephony; Media handling and interaction (Release 12)",
          September 2014.

[10.2](). Informative References

[I-D.ietf-rtcweb-rtp-usage]
          Perkins, C., Westerlund, M., and J. Ott, "Web Real-Time
          Communication (WebRTC): Media Transport and Use of RTP",
          draft-ietf-rtcweb-rtp-usage-06 (work in progress),
          February 2013.

[I-D.ietf-rtcweb-security-arch]
          Rescorla, E., "WebRTC Security Architecture", draft-ietf-
          rtcweb-security-arch-09 (work in progress), February 2014.

[I-D.ietf-rtcweb-security]
          Rescorla, E., "Security Considerations for WebRTC", draft-
          ietf-rtcweb-security-06 (work in progress), January 2014.

Author's Address

    Adam Roach
    Mozilla
    \
    Dallas
    US

    Phone: +1 650 903 0800 x863
    Email: adam@nostrum.com