

Queste de savoir

1 heure pour encoder 5 secondes de vidéo

6 janvier 2023

Table des matières

Introduction	1
Contenu masqué	8

Introduction

Le [codec AV1](#) est un codec de dernière génération, libre et ouvert, qui est le fruit de la collaboration de plusieurs grands noms, cela prend la forme d'un consortium: l'[Alliance for Open Media](#) (AOM).

AV1 promet des performances de compression excellentes en maintenant une belle qualité d'image quand on le compare à H.264 qui reste, encore aujourd'hui, le codec de choix dans énormément d'applications à destination du Web.

Cependant, AV1 se plie mal—vraiment très mal—à l'encodage logiciel. J'ai testé l'encodeur de référence (de l'AOM) via `ffmpeg`. Voyons ce que ça donne.

Version source

Mon fichier source est une vidéo **1920×1080 en 60 FPS** prise depuis un iPhone 8. Il s'agit de la source directe, telle qu'enregistrée par l'iPhone au moment de la capture. Voici les caractéristiques du flux vidéo de ~5 secondes de la source:

- HEVC, profil **Main L4.1**
- 1920×1080, 60 FPS (constant)
- Bitrate de **12.0 Mb/s**

Version H.264

Une version H.264 a été produite, laissant à `ffmpeg` le soin de sélectionner les bons presets H.264. L'encodeur est ici [libx264](#), un encodeur logiciel rapide et réputé. L'encodeur matériel via [VA-API](#) de mon CPU Intel est plus performant encore, mais il ne produit pas un rendu aussi bon à bitrate égal.

Nous obtenons ainsi un fichier vidéo affichant un bitrate moyen de **9.2 Mb/s**, l'encodage se fait en **quelques secondes** sur mon Intel Core i7-10710U.

Version AV1

Pour la version AV1, c'est un résultat bien différent: c'est bien **1 heure et 17 minutes** d'encodage qui ont été nécessaires! Le flux final a un bitrate moyen de **8.3 Mb/s**, ce qui est certes meilleur que celui de la version H.264, mais ça ne semble pas être un gain mirobolant, surtout quand on passe de quelques secondes d'encodage à plus d'une heure.

Différence de qualité

AV1 se targue d'avoir une compression très performante. Ce n'est pas probant ici. Cependant, nous pouvons approcher le même problème sous un autre angle. En comparant le flux H.264 (9.2 Mb/s) et le flux AV1 (8.3 Mb/s), quel est celui qui a le meilleur rendu visuel à la fin?

J'ai pris quelques captures grâce à `ffmpeg` pour comparer certains éléments de la vidéo, tantôt sur la version H.264, tantôt sur la version AV1. Ces captures ont été prises en PNG sans compression. Les captures sont assez lourdes (2.5 MB pour chaque fichier PNG), donc je ne mets ici que des éléments rognés, eux non plus sans compression.

Mes observations sont dans les blocs secrets. Ainsi, vous pouvez vous amuser à distinguer des éléments entre les deux versions sans être influencés.

Numéro d'immeuble



FIGURE 0.1. – AV1 - numéro 16



FIGURE 0.2. – H264 - Numéro 16

Mon avis subjectif:

© Contenu masqué n°1

Visage moyennement proche de la caméra



FIGURE 0.3. – AV1 - Visage



FIGURE 0.4. – H264 - Visage

Mon avis subjectif:

© Contenu masqué n°2

Poteau



FIGURE 0.5. – AV1 - Poteau

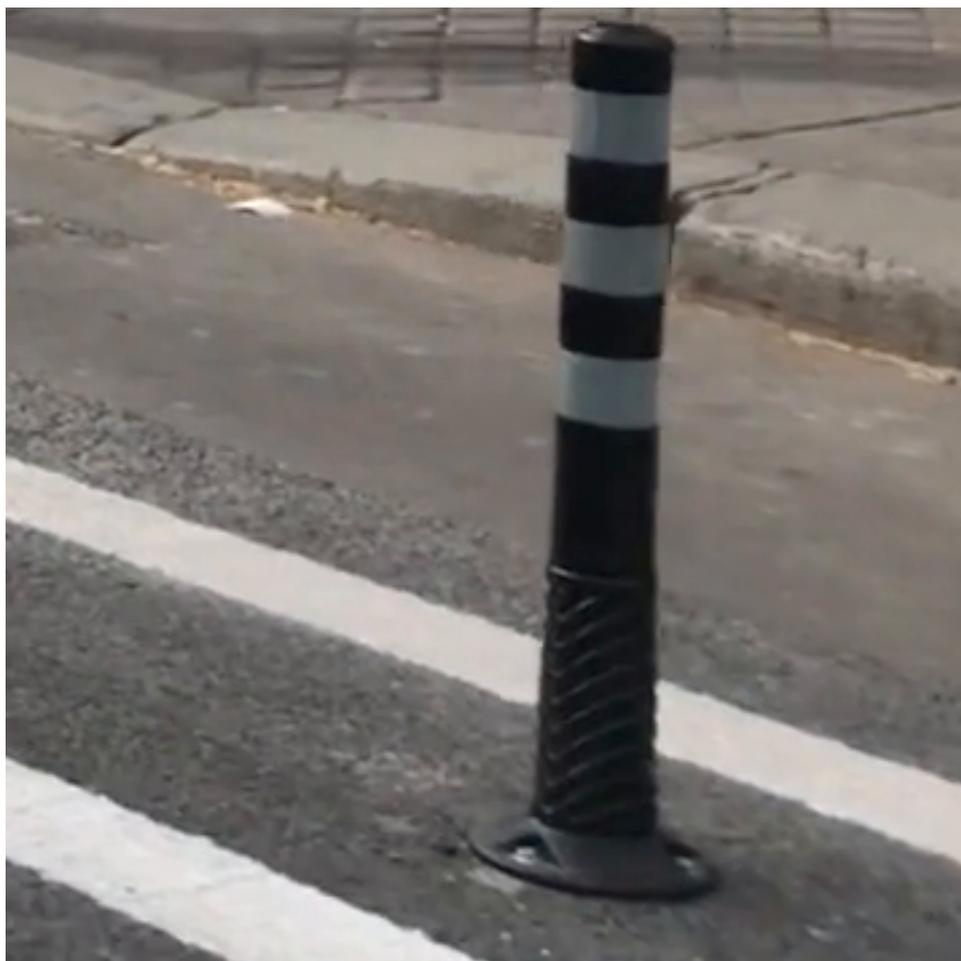


FIGURE 0.6. – H264 - Poteau

Mon avis subjectif:

👁️ Contenu masqué n°3

Conclusion

L'encodeur *software* de référence de l'AOM me semble produire une meilleure qualité à bitrate plus réduit (de 1 Mb/s) que H.264. Cependant, le test n'est pas probant et mériterait certainement d'être suivi d'une analyse plus approfondie. Notamment en testant sur des fichiers 4K (qui ont vraiment besoin d'une grosse compression pour être praticables) et en s'assurant de comparer ces deux codecs avec des *flags* et des *presets* rendant la comparaison plus pertinente. Il faudrait aussi viser absolument les bitrates qui sont la norme dans la distribution Web (pour du 1080p 60 FPS, nos bitrates sont plutôt raisonnables en l'occurrence).

À noter qu'entre H.264 et AV1 nous sautons une génération, celle de HEVC (H.265) et de VP9. Les tests comparent en général H.264 directement avec AV1 car à part sur YouTube où VP9 s'est généralisé pour la distribution finale, H.264 reste quand même le codec dominant du Web. La comparaison est donc peu équitable techniquement, mais elle l'est du point de vue de

Contenu masqué

l'expérience de l'internaute, puisque l'ambition d'AV1 est de devenir le nouveau codec standard et omniprésent du Web.

Enfin, même si l'encodage logiciel semble peu viable pour AV1, rassurez-vous: il est tout à fait possible de décoder décentement des flux AV1 logiciellement. Si vous consultez YouTube, il est même possible que vous ayez déjà eu des vidéos en AV1 décodées de façon logicielle, sans le moindre souci. Ma machine, peinant tant à encoder en AV1, décode un flux 4K encodé en AV1 sans *drop* et sans faire tourner les ventilateurs à fond, avec le décodeur logiciel [dav1d](#) [↗](#) .

Contenu masqué

Contenu masqué n°1

Le numéro me semble plus distinct sur la version AV1. Cependant, cela est plus probant avec l'image de 1920×1080 complète. [Retourner au texte.](#)

Contenu masqué n°2

Le visage semble présenter moins d'artefacts (blocs carrés) sur la version AV1. La délimitation entre le verre des lunettes et la peau bave moins et est plus propre sur la version AV1. La zone ombrée par les lunettes est plus propre sur la version AV1. [Retourner au texte.](#)

Contenu masqué n°3

Sur les bandes blanches réfléchissantes des poteaux, la version AV1 semble conserver plus de détails. [Retourner au texte.](#)