

Groupe de travail Réseau
Request for Comments : 4175
 Catégorie : Sur la voie de la normalisation
 Traduction Claude Brière de L'Isle

L. Gharai, USC/ISI
 C. Perkins, University of Glasgow
 septembre 2005

Format de charge utile RTP pour vidéo non compressée

Statut de ce mémoire

Le présent document spécifie un protocole de l'Internet en cours de normalisation pour la communauté de l'Internet, et appelle à des discussions et suggestions pour son amélioration. Prière de se référer à l'édition en cours des "Normes officielles des protocoles de l'Internet" (STD 1) pour connaître l'état de la normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

Notice de copyright

Copyright (C) The Internet Society (2005).

Résumé

Le présent mémoire spécifie un schéma de mise en paquet pour encapsuler de la vidéo non compressée dans un format de charge utile pour le protocole de transport en temps réel, RTP. Il prend en charge une gamme de formats de vidéo standard et de haute définition, incluant les formats courants de télévision tels que l'UIT BT.601, et des normes de la Société des ingénieurs de télévision et d'images animées (SMPTE), comme SMPTE 274M et SMPTE 296M. Le format est conçu pour être applicable et extensible aux nouveaux formats de vidéo qui seront développés.

Table des matières

1. Introduction.....	1
2. Conventions du document.....	2
3. Dessin de la charge utile.....	2
4. Mise en paquet RTP.....	3
4.1 En-tête RTP.....	3
4.2 En-tête de charge utile.....	4
4.3 Données de charge utile.....	4
5. Considérations relatives au RTCP.....	6
6. Considérations relatives à l'IANA.....	7
6.1 Enregistrement de type MIME.....	7
6.2 Enregistrement des paramètres.....	8
7. Transposition des paramètres MIME en SDP.....	8
8. Considérations sur la sécurité.....	9
9. Relation à la RFC 2431.....	9
10. Relation à la RFC 3497.....	9
11. Remerciements.....	9
Références normatives.....	10
Références pour information.....	10
Adresse des auteurs.....	10
Déclaration complète de droits de reproduction.....	11

1. Introduction

Le présent mémoire définit un schéma de mise en paquet de flux vidéo non compressés de qualité studio pour un transport utilisant RTP [RFC3550]. Il prend en charge une gamme de formats de vidéo standard et de haute définition, incluant UIT-R BT.601 [BT.601], SMPTE 274M [SMPTE.274] et SMPTE 296M [SMPTE.296].

Les formats pour la télévision en définition standard non comprimée sont définis par la Recommandation UIT-R BT.601 [BT.601] ainsi que pour les interfaces en série binaire et parallèles dans la Recommandation BT.656 [BT.656]. Ces formats permettent des fonctionnements aussi bien en 625 lignes qu'en 525 lignes, avec 720 échantillons par ligne numérique active, un sous échantillonnage de couleur 4:2:2, et une représentation numérique sur 8 ou 10 bits.

La représentation de télévision haute définition non compressée est spécifiée dans les normes SMPTE 274M [SMPTE.274] et 296M [SMPTE.296]. SMPTE 274M définit une famille de systèmes d'examen avec un format d'image de 1920 x 1080 pixels avec examen progressif et entrelacé, tandis que SMPTE 296M définit des systèmes avec une taille d'image de 1280 x 720 pixels et examen progressif. Dans l'examen progressif, les lignes d'examen sont affichées en séquence du haut en bas d'une trame entière. Dans l'examen entrelacé, une trame est divisée en lignes d'examen impaires et paires (appelées des champs) et les deux champs sont affichés à la suite. SMPTE 274M et 296M définissent des images avec un rapport d'aspect de 16:9, et définissent la représentation numérique pour les composants RGB et YCbCr. Dans le cas des composants YCbCr, les composants Cb et Cr sont sous échantillonnés horizontalement par un facteur deux (codage de couleur 4:2:2).

Bien que ces formats diffèrent dans les détails, ils sont structurellement très similaires. Le présent mémoire spécifie un format de charge utile pour encapsuler ces formats et d'autres formats de vidéo similaires pour le transport en RTP.

2. Conventions du document

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans le BCP 14, [RFC2119].

3. Dessin de la charge utile

Chaque ligne d'examen de vidéo numérique est mise en paquet dans un ou plusieurs paquets RTP. Si les données pour une ligne d'examen complète excèdent la MTU du réseau, la ligne d'examen DEVRAIT être fragmentée en plusieurs paquets RTP, chacun inférieur à la MTU. Un seul paquet RTP PEUT contenir des données de plus d'une ligne d'examen. Seuls les échantillons actifs sont inclus dans la charge utile RTP : les échantillons inactifs et le contenu des blancs horizontaux et verticaux NE DEVRAIENT PAS être transportés. Dans des instances où les données auxiliaires sont transmises, l'expéditeur et le receveur peuvent faire la distinction entre les données auxiliaires et les données vidéo par les numéros de ligne d'examen. C'est-à-dire que les données auxiliaires vont utiliser des numéros de ligne d'examen qui ne sont pas dans la portée de la trame vidéo.

Les numéros de ligne d'examen sont inclus dans l'en-tête de charge utile RTP, ainsi qu'un identifiant de champ pour la vidéo entrelacée.

Pour la vidéo en format SMPTE 296M, les numéros de ligne d'examen valides sont de 26 à 745, inclus. Pour la vidéo en format SMPTE 274M à examen progressif, les lignes d'examen valides sont des lignes d'examen 42 à 1121, inclus. Pour la vidéo de format SMPTE 274M à examen entrelacé, les numéros de ligne d'examen valides pour le champ un (F=0) sont de 21 à 560 et les numéros de ligne d'examen valides pour le second champ (F=1) sont de 584 à 1123. Pour la vidéo de format UIT-R BT.601, les intervalles de blanc définis dans BT.656 sont utilisés pour la vidéo à 625 lignes, les lignes 24 à 310 du champ un (F=0) et 337 à 623 du second champ (F=1) sont valides ; pour la vidéo à 525 lignes, les lignes 21 à 263 du premier champ, et 284 à 525 du second champ sont valides. D'autres formats (par exemple, [SMPTE.372]) peuvent définir des gammes différentes de lignes actives.

L'en-tête de charge utile contient une extension de 16 bits au numéro de séquence standard RTP de 16 bits, étendant par là le numéro de séquence à 32 bits et permettant que le format de charge utile s'accommode de débits de données élevés sans ambiguïté. C'est nécessaire car le numéro de séquence de 16 bits de RTP va revenir à zéro très rapidement pour les hauts débits de données. Par exemple, pour un flux vidéo de 1 Gbit/s avec des tailles de paquet d'au moins 1000 octets, le paquet RTP standard va revenir à zéro en 0,5 seconde, ce qui peut être un problème pour détecter les paquets perdus et déclassés, particulièrement dans des instances où le délai d'aller-retour est supérieur à une demie seconde. Le numéro de séquence étendu de 32 bits permet un délai d'épuisement de l'espace de numéros de séquence d'approximativement neuf heures.

Chaque ligne d'examen comporte un nombre entier de pixels. Chaque pixel est représenté par un nombre d'échantillons. Les échantillons peuvent être codés comme des valeurs de 8, 10, 12, ou 16 bits. Un échantillon peut représenter un composant de couleur ou un composant de luminance de la vidéo. Les échantillons de couleur peuvent être partagés entre des pixels adjacents. Le partage d'échantillons de couleur entre des pixels adjacents est appelé sous échantillonnage de couleur. C'est normalement fait dans l'espace de couleur YCbCr afin de réduire la taille des données d'image.

Les pixels qui partagent des valeurs d'échantillon DOIVENT être transportés ensemble comme un "groupe de pixels". Si on utilise des échantillons de 10 bits ou de 12 bits, chaque pixel peut aussi comporter un nombre non entier d'octets. Dans ce cas, plusieurs pixels DOIVENT être combinés en un groupe de pixels alignés sur l'octet pour la transmission. Ces

restrictions simplifient le fonctionnement des receveurs en assurant que la charge utile complète est alignée sur l'octet, et que les échantillons se rapportant à un seul pixel ne sont pas fragmentés sur plusieurs paquets [ALF].

Par exemple, dans la vidéo YCbCr avec un sous échantillonnage de couleur 4:1:1, chaque groupe de 4 pixels adjacents comporte 6 échantillons, Y1 Y2 Y3 Y4 Cr Cb, les valeurs Cr et Cb étant partagées par tous les 4 pixels. Si les échantillons sont des valeurs de 8 bits, le résultat est un groupe de 4 pixels comportant 6 octets. Si, cependant, les échantillons sont des valeurs de 10 bits, le groupe de 60 bits résultant n'est pas aligné sur l'octet. Pour être à la fois aligné sur l'octet et tramé de façon appropriée, deux groupes de 4 pixels adjacents doivent être collectés, devenant ainsi alignés sur l'octet sur une limite de 15 octets. Cette longueur est appelée la taille de groupe de pixels ("pgroup").

Formellement, le paramètre "pgroup" est la taille en octets du plus petit groupement de pixels tels que 1) le groupement comporte un nombre entier d'octets, et 2) si le sous échantillonnage de couleur est utilisé, les échantillons ne sont partagés qu'au sein du groupement. Quand on met en paquets un contenu de ligne active numérique, les données de vidéo NE DOIVENT PAS être fragmentées au sein d'un pgroup.

Le contenu vidéo est presque toujours associé à des informations supplémentaires telles que des pistes audio un code horaire, etc. Dans les applications de vidéo numérique professionnelles, ces données sont couramment incorporées dans des portions non actives du flux vidéo (périodes de blanc horizontal et vertical) afin qu'une synchronisation précise et robuste soit conservée. Ce format de charge utile exige que les applications qui utilisent de telles données auxiliaires synchronisées DEVRAIENT les livrer dans des sessions RTP séparés qui opèrent concurremment avec la session vidéo. Le mécanisme RTP normal DEVRAIT être utilisé pour synchroniser le support.

4. Mise en paquet RTP

L'en-tête RTP standard est suivi par un en-tête de charge utile de 2 octets qui étend le numéro de séquence RTP, et par un en-tête de charge utile de 6 octets pour chaque ligne (ou ligne partielle) de vidéo incluse. Une ou plusieurs lignes, ou lignes partielles, de données vidéo suivent. Ce format fait que l'en-tête de charge utile de 32 bits est aligné dans le cas courant, où une ligne d'examen (ou un fragment) de vidéo est incluse dans chaque paquet RTP.

Par exemple, si deux lignes de vidéo sont encapsulées, le format de charge utile sera celui montré à la Figure 1.

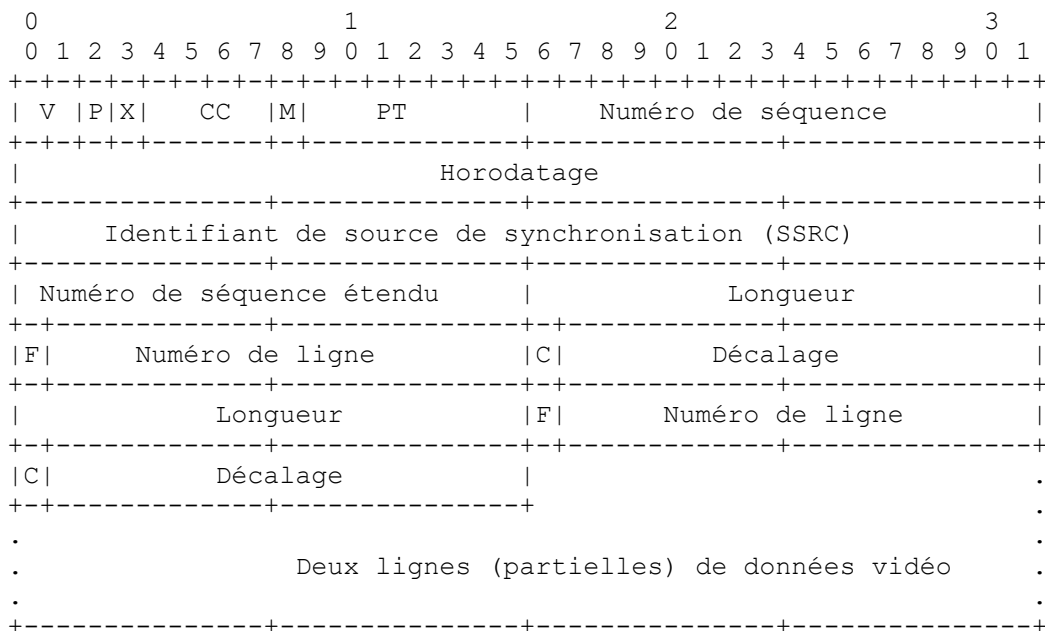


Figure 1 : Format de charge utile RTP montrant deux lignes (partielles) de vidéo

4.1 En-tête RTP

Les champs de l'en-tête fixe RTP ont leur signification usuelle, avec les notes supplémentaires suivantes :

Type de charge utile (PT) : 7 bits

Un champ alloué de façon dynamique Type de charge utile qui désigne la charge utile comme vidéo non compressée.

Horodatage : 32 bits

Pour la vidéo à examen progressif, l'horodatage note l'instant d'échantillonnage de la trame à laquelle appartient le paquet RTP. Les paquets NE DOIVENT PAS inclure de données provenant de plusieurs trames, et tous les paquets appartenant à la même trame DOIVENT avoir le même horodatage.

Pour la vidéo entrelacée, l'horodatage note l'instant d'échantillonnage du champ auquel appartient le paquet RTP. Les paquets NE DOIVENT PAS inclure de données provenant de plusieurs champs, et tous les paquets appartenant au même champ DOIVENT avoir le même horodatage. L'utilisation du champ Horodatage, plutôt qu'un horodatage de trame et d'un bit d'indicateur de champ, est nécessaire pour prendre en charge la décimation 3-2 inverse.

Un horodatage à 90 kHz DEVRAIT être utilisé dans les deux cas. Si l'instant d'échantillonnage ne correspond pas à une valeur entière de l'horloge (comme ce peut être le cas en entrelacement) la valeur DEVRA être tronquée à l'entier inférieur le plus proche, sans ambiguïté.

Bit marqueur (M) : 1 bit

Si une vidéo à examen progressif est transmise, le bit marqueur note la fin d'une trame vidéo. Si c'est de la vidéo entrelacée qui est transmise, il note la fin du champ. Le bit marqueur DOIT être réglé à 1 pour le dernier paquet de la trame/champ vidéo. Il DOIT être à 0 pour les autres paquets.

Numéro de séquence : 16 bits

Ce sont les bits de moindre poids du numéro de séquence RTP. Le numéro de séquence standard de 16 bits est augmenté de 16 autres bits dans l'en-tête de charge utile afin d'éviter des problèmes dus au retour à zéro lors du fonctionnement à haut débit.

4.2 En-tête de charge utile

Numéro de séquence étendu : 16 bits

Les bits de poids fort du numéro de séquence étendu de 32 bits, dans l'ordre des octets du réseau.

Longueur : 16 bits

Nombre d'octets de données inclus provenant de cette ligne d'examen, dans l'ordre des octets du réseau. Ce DOIT être un multiple de la valeur de pgroup.

Numéro de ligne : 15 bits

Numéro de la ligne d'examen des données encapsulées, dans l'ordre des octets du réseau. Les paquets RTP successifs PEUVENT contenir des parties de la même ligne d'examen (avec un numéro de séquence RTP incrémenté, mais le même horodatage) si il est nécessaire de fragmenter une ligne.

Décalage : 15 bits

C'est le décalage du premier pixel des données de charge utile au sein de la ligne d'examen. Si ce sont des données de format YCbCr qui sont transportées, c'est le décalage du pixel de l'échantillon de luminance ; si c'est le format RGB qui est transporté, c'est le décalage du pixel de l'échantillon rouge ; si c'est le format BGR qui est transporté, c'est le décalage de pixel de l'échantillon bleu. La valeur est dans l'ordre des octets du réseau. Le décalage a une valeur de zéro si le premier échantillon de la charge utile correspond au début de la ligne, et s'incrémente de un pour chaque pixel.

Identification de champ (F) : 1 bit

Identifie à quel champ appartient la ligne d'examen, pour les données entrelacées. F=0 identifie le premier champ et F=1 le second champ. Pour les données à examen progressif (par exemple, de format vidéo SMPTE 296M) F DOIT toujours être réglé à zéro.

Continuation (C) : 1 bit

Détermine si un en-tête de ligne d'examen supplémentaire suit l'en-tête de la ligne d'examen en cours dans la paquet RTP. Réglé à 1 si un en-tête supplémentaire suit, ce qui implique que le paquet RTP porte des données pour plus d'une ligne d'examen. Réglé à 0 autrement. Plusieurs lignes d'examen PEUVENT être incluses dans un seul paquet, jusqu'à la limite de la MTU de chemin. La seule façon de déterminer le nombre de lignes d'examen incluses par paquet est d'analyser les en-têtes de charge utile.

4.3 Données de charge utile

Selon le format de vidéo, chaque paquet RTP peut inclure une seule ligne d'examen complète, un seul fragment d'une ligne d'examen, ou une (ou plusieurs) lignes d'examen complètes et des fragments de ligne d'examen. La longueur de chaque ligne d'examen ou fragment de ligne d'examen DOIT être un multiple entier de la taille de pgroup en octets. Les lignes d'examen DEVRAIENT être fragmentées afin que le paquet RTP résultant soit plus petit que la MTU de chemin.

Il est possible que la longueur de la ligne d'examen ne soit pas divisible par le nombre de pixels d'un pgroup, de sorte que le pixel final des données d'une ligne d'examen ne s'aligne pas sur une limite d'octet ou de pgroup. Néanmoins, la charge utile DOIT contenir un nombre entier de pgroup ; l'expéditeur DOIT remplir les bits restants du pgroup final avec des zéros et le receveur DOIT ignorer les données de remplissage. (En pratique, le bord de queue de l'image est noirci jusqu'à une limite de pgroup.)

Pour la vidéo en format RGB, les échantillons sont mis en paquet dans l'ordre Rouge-Vert-Bleu. Pour la vidéo en format BGR, les échantillons sont mis en paquet dans l'ordre Bleu-Vert-Rouge. Pour les deux formats, si des échantillons de 8 bits sont utilisés, le pgroup fait 3 octets. Si des échantillons de dix bits sont utilisés, les échantillons provenant de quatre pixels adjacents forment un pgroup de 15 octets. Si des échantillons de douze bits sont utilisés, les échantillons provenant de deux pixels adjacents forment un pgroup de neuf octets. Si des échantillons de seize bits sont utilisés, chaque pixel forme un pgroup séparé de 6 octets.

Pour la vidéo en format RGBA, les échantillons sont mis en paquets dans l'ordre Rouge-Vert-Bleu-Alpha. Pour le format de vidéo BGRA, les échantillons sont mis en paquet dans l'ordre Bleu-Vert-Rouge-Alpha. Pour les échantillons de 8, 10, 12, ou 16 bits, chaque pixel forme son propre pgroup, avec des tailles d'octet respectives de 4, 5, 6, et 8.

Si la vidéo est en format YCbCr, la mise en paquet des échantillons dans la charge utile dépend du sous échantillonnage de couleur utilisé.

Pour la vidéo en format YCbCr 4:4:4, les échantillons sont mis en paquet dans l'ordre Cb-Y-Cr pour les trames aussi bien entrelacées que progressives. Si des échantillons de huit bits sont utilisés, le pgroup fait 3 octets. Si des échantillons de dix bits sont utilisés, les échantillons provenant de quatre pixels adjacents forment des pgroup de 15 octets. Si des échantillons de douze bits sont utilisés, les échantillons provenant de deux pixels adjacents forment des pgroup de neuf octets. Si des échantillons de seize bits sont utilisés, chaque pixel forme un pgroup séparé de 6 octets.

Pour la vidéo en format YCbCr 4:2:2, les composants Cb et Cr sont sous échantillonnés horizontalement d'un facteur deux (chaque échantillon Cb et Cr correspond à deux composants Y). Les échantillons sont mis en paquet dans l'ordre Cb0-Y0-Cr0-Y1 pour les lignes d'examen aussi bien entrelacées que progressives. Pour les échantillons de 8, 10, 12, ou 16 bits, le pgroup est formé de deux pixels adjacents (respectivement 4, 5, 6, ou 8 octets).

Pour la vidéo en format YCbCr 4:1:1, les composants Cb et Cr sont sous échantillonnés horizontalement d'un facteur quatre (chaque échantillon Cb et Cr correspond à quatre composants Y). Les échantillons sont mis en paquet dans l'ordre Cb0-Y0-Y1-Cr0-Y2-Y3 pour les lignes d'examen aussi bien entrelacées que progressives. Pour les échantillons de 8, 10, 12, ou 16 bits, le pgroup est formé de quatre pixels adjacents (respectivement 4, 5, 6, ou 8 octets).

Pour la vidéo YCbCr 4:2:0, les composants Cb et Cr sont sous échantillonnés d'un facteur deux aussi bien horizontalement que verticalement. Donc, les échantillons de chrominance sont partagés entre certaines lignes adjacentes. La Figure 2 montre la composition des échantillons de luminance et de chrominance pour une grille 6x6 de pixels de vidéo 4:2:0 YCbCr. Le groupe de pixel est un groupe de quatre pixels arrangés dans une matrice 2x2. La taille en octets du pgroup pour la vidéo à examen progressif 4:2:0 avec des tailles d'échantillon de 8, 10, 12, et 16 bits est respectivement de 6, 15, 9, et 12 octets. Pour la vidéo entrelacée 4:2:0, les pgroup correspondants sont de 4, 5, 6, et 8 octets.

```

ligne 0 : Y00 Y01 Y02 Y03 Y04 Y05 Cb00 Cr00 Cb01 Cr01 Cb02 Cr02
ligne 1 : Y10 Y11 Y12 Y13 Y14 Y15

ligne 2 : Y20 Y21 Y22 Y23 Y24 Y25 Cb10 Cr10 Cb11 Cr11 Cb12 Cr12
ligne 3 : Y30 Y31 Y32 Y33 Y34 Y35

ligne 4 : Y40 Y41 Y42 Y43 Y44 Y45 Cb20 Cr20 Cb21 Cr21 Cb22 Cr22
ligne 5 : Y50 Y51 Y52 Y53 Y54 Y55

```

Figure 2 : Composition chrominance/luminance en vidéo 4:2:0 YCbCr

Quand on met en paquets de la vidéo à examen progressif 4:2:0 YCbCr, les échantillons provenant de deux lignes d'examen consécutives sont incluses dans chaque paquet. Le numéro de ligne d'examen dans l'en-tête de charge utile est réglé à celui de la première ligne d'examen de la paire :

ligne 0/1 : Y00-Y01-Y10-Y11-Cb00-Cr00 Y02-Y03-Y12-Y13-Cb01-Cr01 Y04-Y05-Y14-Y15-Cb02-Cr02

ligne 2/3 : Y20-Y21-Y30-Y31-Cb10-Cr10 Y22-Y23-Y32-Y33-Cb11-Cr11 Y24-Y25-Y34-Y35-Cb12-Cr12

ligne 4/5 : Y40-Y41-Y50-Y51-Cb20-Cr20 Y42-Y43-Y52-Y53-Cb21-Cr21 Y44-Y45-Y54-Y55-Cb22-Cr22

Figure 3 : Mise en paquet de vidéo YCbCr 4:2:0 progressive

Pour le transport entrelacé, les échantillons de chrominance sont transportés avec toute autre ligne. Le premier ensemble d'échantillons de chrominance peut être transporté avec la première ligne de champ 0, ou avec la première ligne de champ 1. La Figure 4 illustre le transport des échantillons de chrominance en commençant par la première ligne de champ 0 (signalé par le paramètre MIME "Champ supérieur en premier").

Champ 0 :

ligne 0 : Y00-Y01-Cb00-Cr00 Y02-Y03-Cb01-Cr01 Y04-Y05-Cb02-Cr02

ligne 2 : Y20-Y21 Y22-Y23 Y24-Y25

ligne 4 : Y40-Y41-Cb20-Cr20 Y42-Y43-Cb21-Cr21 Y44-Y45-Cb22-Cr22

Champ 1 :

ligne 1 : Y10-Y11 Y12-Y13 Y14-Y15

ligne 3 : Y30-Y31-Cb10-Cr10 Y32-Y33-Cb11 Cr11 Y34-Y35-Cb12-Cr12

ligne 5 : Y50-Y51 Y52-Y53 Y54-Y55

Figure 4 : Mise en paquet de vidéo entrelacée 4:2:0 YCbCr avec champ supérieur en premier

Les valeurs de chrominance peuvent être échantillonnées avec des décalages différents par rapport aux valeurs de luminance. Par exemple, dans la Figure 2, les valeurs de chrominance sont échantillonnées à la même distance des échantillons de luminance du voisinage. Il est aussi possible qu'un échantillon de chrominance soit co-situé avec un échantillon de luminance, comme dans la Figure 5 :

ligne 0 : Y00-C Y01 Y02-C Y03 Y04-C Y05

ligne 1 : Y10 Y11 Y12 Y13 Y14 Y15

ligne 2 : Y20-C Y21 Y22-C Y23 Y24-C Y25

ligne 3 : Y30 Y31 Y32 Y33 Y34 Y35

ligne 4 : Y40-C Y41 Y42-C Y43 Y44-C Y45

ligne 5 : Y50 Y51 Y52 Y53 Y54 Y55

Figure 5 : Échantillonnage co-situé de vidéo en 4:2:0 YCbCr où C désigne une paire CbCr

En général, les valeurs de chrominance peuvent être placées entre les échantillons de luminance ou co-situés. Les positions peuvent être désignées par un système de numérotation d'entiers commençant de gauche à droite et de haut en bas. Les matrices de position des Figures 6, 7, et 8 s'appliquent respectivement pour la vidéo 4:2:0, 4:2:2, et 4:1:1 :

ligne N : Y[0] [1] Y[2] Y[0] [1] Y[2] [3] [4] Y[5] [3] [4] [5]

ligne N+1 : Y[6] [7] Y[8] Y[6] [7] Y[8]

Figure 6 : Matrice de position de chrominance pour la vidéo 4:2:0 YCbCr

ligne N : Y[0] [1] Y[2] [3] Y[0] [1] Y[2] [3]

ligne N+1 : Y[0] [1] Y[2] [3] Y[0] [1] Y[2] [3]

Figure 7 : Matrice de position de chrominance pour la vidéo 4:2:2 YCbCr

ligne N : Y[0] [1] Y[2] [3] Y[4] [5] Y[6]

ligne N+1 : Y[0] [1] Y[2] [3] Y[4] [5] Y[6]

Figure 8 : Matrice de position de chrominance pour la vidéo 4:1:1 YCbCr

Bien que ces positions n'affectent pas l'ordre de mise en paquet des échantillons de chrominance et de luminance, les informations sont nécessaires pour l'interpolation avant l'affichage et devraient donc être signalées au receveur.

5. Considérations relatives à RTCP

RTCP DEVRAIT être utilisé comme spécifié dans la [RFC3550]. Il faut noter que le compte d'octets de l'envoyeur dans les paquets SR et le nombre cumulé de paquets perdus va revenir rapidement à zéro pour les flux de données à haut débit. Cela signifie que ces deux champs peuvent ne pas représenter de façon précise le compte d'octets et le nombre de paquets perdus depuis le début de la transmission, comme défini dans la RFC 3550. Donc, pour les besoins de la surveillance du réseau, d'autres moyens de garder trace de ces variables DEVRAIENT être utilisés.

6. Considérations relatives à l'IANA

L'IANA a enregistré un nouveau sous type MIME avec un format de charge utile RTP associé, et a créé deux registres de sous paramètres, comme décrit dans ce qui suit.

6.1 Enregistrement de type MIME

Nom de type de support MIME : video

Nom de sous type MIME : raw

Paramètres exigés :

Débit : débit de l'horloge d'horodatage RTP. Les applications qui utilisent ce format de charge utile DEVRAIENT utiliser une valeur de 90000.

Échantillonnage : détermine le mode d'échantillonnage ou sous échantillonnage de couleur du flux vidéo. Les valeurs actuellement définies sont RGB, RGBA, BGR, BGRA, YCbCr-4:4:4, YCbCr-4:2:2, YCbCr-4:2:0, et YCbCr-4:1:1. De nouvelles valeurs peuvent être enregistrées comme décrit au paragraphe 6.2 de la RFC 4175.

Largeur : détermine le nombre de pixels par ligne. C'est un entier entre 1 et 32767.

Hauteur : détermine le nombre de lignes par trame. C'est un entier entre 1 et 32767.

Profondeur : détermine le nombre de bits par échantillon. C'est un entier dont les valeurs normales sont 8, 10, 12, et 16.

Colorimétrie : ce paramètre définit l'ensemble des spécifications colorimétriques et autres caractéristiques de transfert pour la source vidéo, par référence à une spécification externe. Les valeurs valides et leurs spécifications sont :

BT601-5 : Recommandation UIT-R BT.601-5 [BT.601]

BT709-2 : Recommandation UIT-R BT.709-2 [BT.709]

SMPTE240M : Norme SMPTE 240M [SMPTE.240]

De nouvelles valeurs peuvent être enregistrées comme décrit au paragraphe 6.2 de la RFC 4175.

Paramètres facultatifs :

Entrelacé : Si ce paramètre FACULTATIF est présent, il indique que le flux vidéo est entrelacé. Si il est absent, l'examen progressif est implicite.

Champ supérieur en premier : Si ce paramètre FACULTATIF est présent, il indique que les échantillons de chrominance sont mis en paquets en commençant par la première ligne du champ 0. Son absence implique que les échantillons de chrominance sont mis en paquet en commençant par la première ligne du champ 1.

chroma-position : Ce paramètre FACULTATIF définit la position des échantillons de chrominance par rapport aux échantillons de luminance. C'est soit un seul entier, soit une paire d'entiers séparés par une virgule. Les valeurs d'entier vont de 0 à 8, comme spécifié dans les Figures 6 à 8 de la RFC 4175. Un seul entier implique que Cb et Cr sont co-situés. Une paire d'entiers séparés par une virgule désigne la situation des échantillons, respectivement Cb et Cr. En son absence, une seule valeur de zéro est supposée pour la vidéo à sous échantillonnage de couleur (chroma-position=0).

gamma : Valeur FACULTATIVE de correction gamma à virgule flottante.

Considérations de codage :

La vidéo non compressée est transmise sur RTP seulement comme spécifié dans la RFC 4175. Aucun type de support de format de fichier n'a été défini pour l'instant pour aller avec ce type de support de transmission.

Considérations sur la sécurité : voir la section 9 de la RFC 4175.

Considérations d'interopérabilité : aucune

Spécification publiée : RFC 4175.

Applications qui utilisent ce type de support : Vidéocommunication.

Informations supplémentaires : aucune

Adresse personnelle & de messagerie à contacter pour plus d'informations :

Ladan Gharai <ladan@isi.edu>

Groupe de travail Transport audio/vidéo de l'IETF.

Utilisation prévue : Commune

Auteur : Ladan Gharai <ladan@isi.edu>

Contrôleur des changements : groupe de travail AVT de l'IETF sur délégation de l'IESG.

6.2 Enregistrement des paramètres

De nouvelles valeurs du paramètre "échantillonnage" PEUVENT être enregistrées auprès de l'IANA pourvu qu'elles fassent référence à une RFC ou autre spécification permanente et directement disponible (politique "Spécification requise" de la [RFC2434]). Un nouvel enregistrement DOIT définir l'ordre de mise en paquet des échantillons et les combinaisons valides de couleur et de modes de sous échantillonnage.

De nouvelles valeurs du paramètre "colorimétrie" PEUVENT être enregistrées auprès de l'IANA pourvu qu'elles fassent référence à une RFC ou autre spécification permanente et directement disponible des paramètres de colorimétrie et autres caractéristiques de transfert applicables (politique "Spécification requise" de la [RFC2434]).

7. Transposition des paramètres MIME en SDP

Les informations portées dans la spécification de type de support MIME ont une transposition spécifique en champs dans le protocole de description de session (SDP, *Session Description Protocol*) [RFC2327], qui est couramment utilisée pour décrire les sessions RTP. Quand SDP est utilisé pour spécifier des sessions qui transportent de la vidéo non compressée, la transposition est la suivante :

- le type MIME ("video") devient dans SDP "m=" comme nom de support ;
- le sous type MIME (nom de format de charge utile) devient en SDP "a=rtptime" comme nom de codage ;
- les paramètres restants deviennent en SDP l'attribut "a=fmtp" en les copiant directement de la chaîne de type de support MIME comme liste séparée par un point-virgule de paires paramètre=valeur.

Voici un exemple de transposition en SDP pour la vidéo non compressée :

```
m=video 30000 RTP/AVP 112
a=rtptime:112 raw/90000
a=fmtp:112 sampling=YCbCr-4:2:2; width=1280; height=720; depth=10; colorimetry=BT.709-2; chroma-position=1
```

Dans cet exemple, un type de charge utile dynamique de 112 est utilisé pour la vidéo non compressée. L'horloge d'échantillonnage RTP est à 90 kHz. Noter que la ligne "a=fmtp:" a été renvoyée à la ligne pour tenir (*dans la version anglaise*) dans cette page, et sera une seule longue ligne dans le fichier SDP.

8. Considérations sur la sécurité

Les paquets RTP qui utilisent le format de charge utile défini dans la présente spécification sont soumis aux considérations sur la sécurité discutées dans la spécification RTP [RFC3550] et dans tout profil RTP approprié. Cela implique que la confidentialité des flux de supports est réalisée par le chiffrement.

Ce type de charge utile ne présente du côté receveur aucune non uniformité significative de la complexité de calcul pour le traitement de paquet qui causerait une menace potentielle de déni de service.

Il est important de noter que la vidéo non compressée peut avoir d'immenses exigences de bande passante (jusqu'à 270 Mbit/s pour la vidéo de définition standard, et approximativement de 1 Gbit/s pour la vidéo haute définition). C'est suffisant pour causer un déni de service potentiel si elle est transmise sur les chemins les plus couramment disponibles de l'Internet.

En conséquence, si le service "au mieux" est utilisé, les utilisateurs de ce format de charge utile DOIVENT surveiller la perte de paquet pour s'assurer que le taux de perte de paquet reste dans les gammes de paramètres acceptables. La perte de paquet est considérée comme acceptable si un flux TCP sur le même chemin réseau, et rencontrant les mêmes conditions de

réseau, réaliserait un débit moyen, mesuré sur une échelle de temps raisonnable, qui ne serait pas moins que ce que le flux RTP réalise. Cette condition peut être satisfaite en mettant en œuvre des mécanismes de contrôle de l'encombrement pour adapter de taux de transmission (ou le nombre de couches souscrites pour une session de diffusion groupée en couches) ou en s'arrangeant pour qu'un receveur quitte la session si le taux de perte s'élève à un niveau inacceptable.

Ce format de charge utile peut aussi être utilisé dans des réseaux qui fournissent des garanties de qualité de service. Si un service amélioré est utilisé, les receveurs DEVRAIENT surveiller la perte de paquet pour s'assurer que le service demandé est bien celui qui est fourni. Sinon, ils DEVRAIENT supposer qu'il reçoivent un service "au mieux" et en tirer les conclusions qui s'imposent.

9. Relation à la RFC 2431

Par rapport à la RFC 2431, le présent mémoire spécifie la prise en charge d'une plus grande variété de vidéos non compressées, en termes de taille de trame, de sous échantillonnage de couleur et de tailles d'échantillons. Bien que la [RFC2431] permette le transport jusqu'à 4096 lignes d'examen et 2048 pixels par ligne, notre type de charge utile peut prendre en charge jusqu'à 32 768 lignes d'examen et pixels par ligne. Aussi, la RFC 2431 ne s'adresse qu'aux données 4:2:2 YCbCr, tandis que le présent mémoire couvre YCbCr, RGB, RGBA, BGR, BGRA, et la plupart des schémas les plus courants de sous échantillonnage de couleurs. Étant données la variété des types de vidéo couverts, le présent mémoire suppose la signalisation hors bande de la taille d'échantillon et des types de données (la RFC 2431 utilise la signalisation dans la bande).

10. Relation à la RFC 3497

La [RFC3497] spécifie un format de charge utile RTP pour l'encapsulation de la vidéo SMPTE 292M. La norme SMPTE 292M définit une interface numérique en série binaire pour le transport de la télévision de haute définition (HDTV, *High-Definition Television*) en zone locale. Comme moyen de transport, SMPTE 292M utilise des mots de 10 bits et un débit de données fixe de 1,485 Gbit/s (et 1,485/1,001 Gbit/s). SMPTE 292M est normalement utilisé dans l'industrie de la diffusion pour le transport d'autres formats vidéo tels que SMPTE 260M, SMPTE 295M, SMPTE 274M, et SMPTE 296M.

La RFC 3497 définit une émulation de circuit pour le transport de SMPTE 292M sur RTP. Elle est très spécifique de SMPTE 292 et a été conçue pour être interopérable avec les équipements existants de diffusion avec un débit constant de 1,485 Gbit/s.

Le présent mémoire définit un schéma souple de mise en paquet native qui peut mettre en paquet toute vidéo non compressée, à des débits de données variables. De plus, à la différence de la RFC 3497, le présent mémoire ne transporte que des pixels de vidéo actifs (c'est-à-dire que les blancs horizontaux et verticaux ne sont pas transportés).

11. Remerciements

Les auteurs remercient Philippe Gentric, Chuck Harrison, Stephan Wenger, et Dave Singer de leurs commentaires.

Le présent mémoire se fonde sur des travaux soutenus par la U.S. National Science Foundation (NSF) sous le mandat n° 0230738. Toutes les opinions, découvertes et conclusions ou recommandations exprimées dans ce document sont celles des auteurs et ne reflètent pas nécessairement les vues de la NSF.

Références normatives

- [BT.601] Union Internationale des Télécommunications, "Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios", Recommandation BT.601, octobre 1995.
- [BT.709] Union Internationale des Télécommunications, "Parameter Values for HDTV Standards for Production and International Programme Exchange", Recommandation BT.709-2
- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. (MàJ par [RFC8174](#))

- [RFC2434] T. Narten et H. Alvestrand, "Lignes directrices pour la rédaction d'une section Considérations relatives à l'IANA dans les RFC", BCP 26, octobre 1998. (*Rendue obsolète par la RFC5226*)
- [RFC3550] H. Schulzrinne, S. Casner, R. Frederick et V. Jacobson, "[RTP : un protocole de transport pour les applications en temps réel](#)", STD 64, juillet 2003. (*MàJ par RFC7164, RFC7160, RFC8083, RFC8108*)
- [SMPTE.240] Society of Motion Picture and Television Engineers, "Television - Signal Parameters - 1125-Line High-Definition Production", SMPTE 240M-1999.

Références pour information

- [ALF] Clark, D. D., and Tennenhouse, D. L., "Architectural Considerations for a New Generation of Protocols", Proceedings of SIGCOMM '90 (Philadelphia, PA, septembre 1990), ACM.
- [BT.656] Union Internationale des Télécommunications, "Interfaces for Digital Component Video Signals in 525-line and 625-line Television Systems Operating at the 4:2:2 Level of Recommendation ITU-R BT.601 (Partie A)", Recommendation BT.656, avril 1998.
- [RFC2327] M. Handley et V. Jacobson, "SDP : [Protocole de description de session](#)", avril 1998. (*Obsolète; voir RFC4566*)
- [RFC2431] D. Tynan, "Format de charge utile RTP pour codage vidéo BT.656", octobre 1998. (*P.S.*)
- [RFC3497] L. Gharai et autres, "Format de charge utile RTP pour vidéo 292M de la société des ingénieurs d'images animées et de télévision (SMPTE)", mars 2003. (*P.S.*)
- [SMPTE.274] Society of Motion Picture and Television Engineers, "1920x1080 Scanning and Analog and Parallel Digital Interfaces for Multiple Picture Rates", SMPTE 274M-1998.
- [SMPTE.296] Society of Motion Picture and Television Engineers, "1280x720 Scanning, Analog and Digital Representation and Analog Interfaces", SMPTE 296M-1998.
- [SMPTE.372] Society of Motion Picture and Television Engineers, "Dual Link 292M Interface for 1920 x 1080 Picture Raster", SMPTE 372M-2002.

Adresse des auteurs

Ladan Gharai
USC Information Sciences Institute
3811 N. Fairfax Drive, #200
Arlington, VA 22203
USA
mél : ladan@isi.edu

Colin Perkins
University of Glasgow
Department of Computing Science
17 Lilybank Gardens
Glasgow G12 8QQ
United Kingdom
mél : csp@cspkins.org

Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (2005).

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et à www.rfc-editor.org, et sauf pour ce qui est mentionné ci-après, les auteurs conservent tous leurs droits.

Le présent document et les informations contenues sont fournies sur une base "EN L'ÉTAT" et le contributeur, l'organisation qu'il ou elle représente ou qui le/la finance (s'il en est), la INTERNET SOCIETY et la INTERNET ENGINEERING TASK FORCE déclinent toutes garanties, exprimées ou implicites, y compris mais non limitées à toute garantie que l'utilisation des informations ci encloses ne violent aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

Propriété intellectuelle

L'IETF ne prend pas position sur la validité et la portée de tout droit de propriété intellectuelle ou autres droits qui pourraient être revendiqués au titre de la mise en œuvre ou l'utilisation de la technologie décrite dans le présent document ou sur la mesure dans laquelle toute licence sur de tels droits pourrait être ou n'être pas disponible ; pas plus qu'elle ne prétend avoir accompli aucun effort pour identifier de tels droits. Les informations sur les procédures de l'ISOC au sujet des droits dans les documents de l'ISOC figurent dans les BCP 78 et BCP 79.

Des copies des dépôts d'IPR faites au secrétariat de l'IETF et toutes assurances de disponibilité de licences, ou le résultat de tentatives faites pour obtenir une licence ou permission générale d'utilisation de tels droits de propriété par ceux qui mettent en œuvre ou utilisent la présente spécification peuvent être obtenues sur répertoire en ligne des IPR de l'IETF à <http://www.ietf.org/ipr> .

L'IETF invite toute partie intéressée à porter son attention sur tous copyrights, licences ou applications de licence, ou autres droits de propriété qui pourraient couvrir les technologies qui peuvent être nécessaires pour mettre en œuvre la présente norme. Prière d'adresser les informations à l'IETF à ietf-ipr@ietf.org .

Remerciement

Le financement de la fonction d'édition des RFC est actuellement fourni par la Internet Society.