

Formats de fichiers vidéos

Illustré à partir du gratuiciel PC « Super »



1) Le container ou conteneur

Un fichier vidéo sur ordinateur doit être contenu dans un fichier (un emballage informatique).

Les formats les plus usuels sont : Quicktime (.mov), Windows Media Video (.wmv), AVI (Audio Video Interlaced) (.avi), MP4, Flash (.flv)...

Le conteneur vidéo décrit comment les données sont organisées dans le fichier, pas comment les données sont encodées.

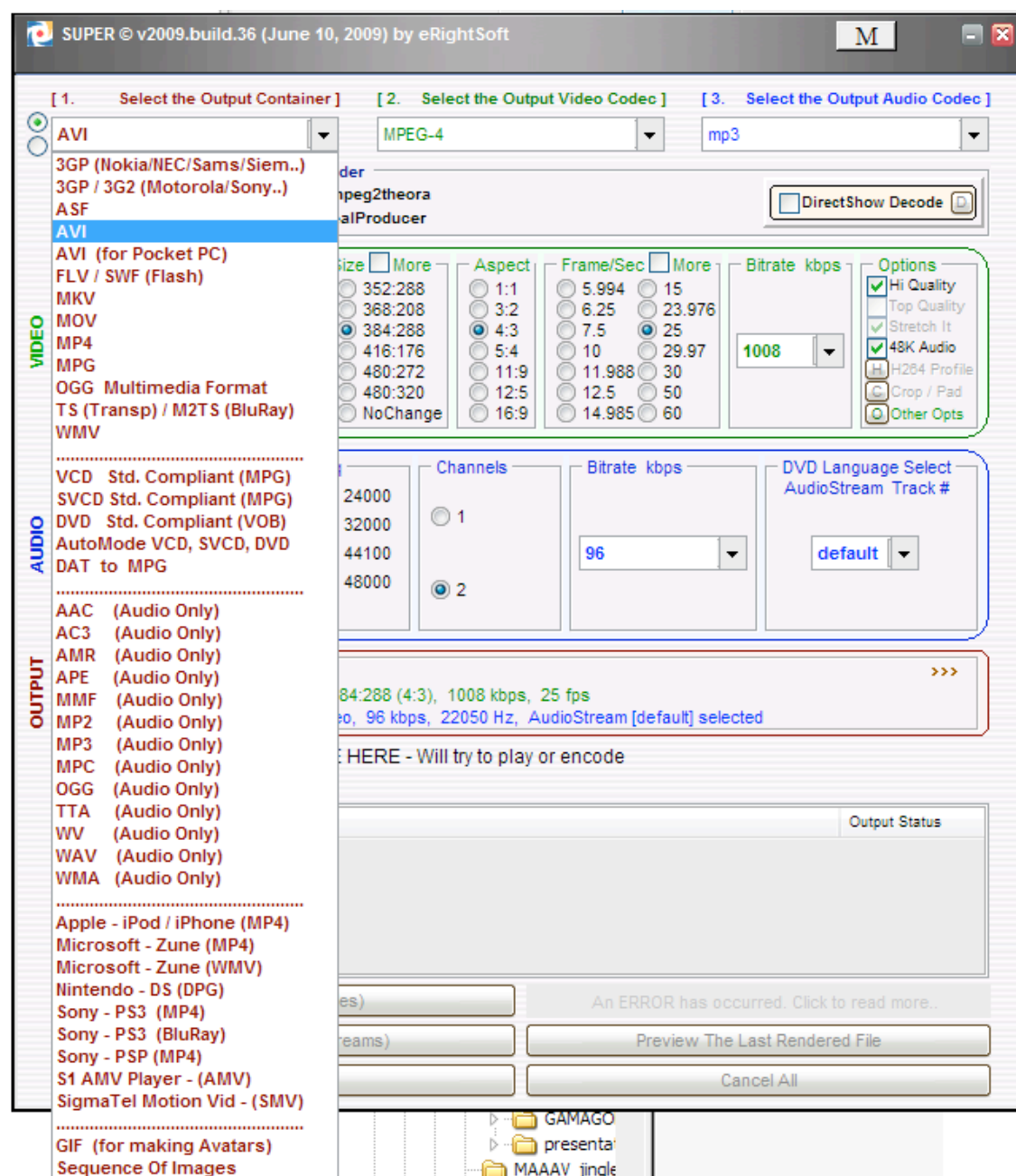


Figure 1 : Le menu 1 de « Super » propose une très grande variété de formats de containers.

2) Le codec ou format de la vidéo

Je parle ici de la vidéo elle-même, pas de son emballage ou conteneur !

Il existe de très nombreux formats vidéo, qui ont leur raison d'être : historique, technologique, guerre de concurrence, grand public ou pro, vidéo en direct, réduction de l'encombrement pour le transport ou au contraire rapidité de lecture...

Le **CODEC vidéo** décrit l'algorithme de COmpression/DÉCompression de l'information vidéo. Une vidéo numérique contient un tel flux de données qu'elle doit forcément être compressée.

Les codecs les plus usuels sont : DivX, H.264, Xvid, x264, Apple Pro Res, MPEG-2, MJPEG, MPEG-4, JPEG 2000, DV, AVCHD, WMV...

Voici la présentation de trois containers courants (AVI, MOV et MP4) susceptibles de contenir des dizaines de codecs ou de formats vidéo différents.

a) AVI

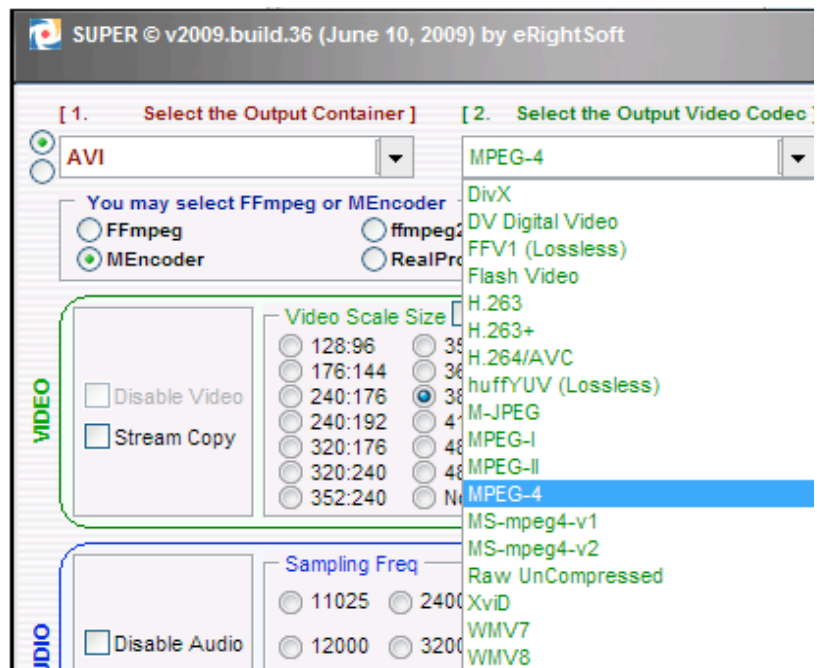
Conteneur traditionnel sur PC.
(= fichier d'emballage)

CODEC à éviter :

- DivX : grand public, trop compressé, impossible à insérer dans un logiciel audionumérique
- DV : caméscopes, gros fichiers, pas assez compressés, trop lourds pour un logiciel audionumérique
- Flash Vidéo : compressé pour le streaming internet
- WMV7 ou 8 : format uniquement PC

CODEC à privilégier :

- H.264, MPEG-4, MPEG2
- M-JPEG



b) MOV

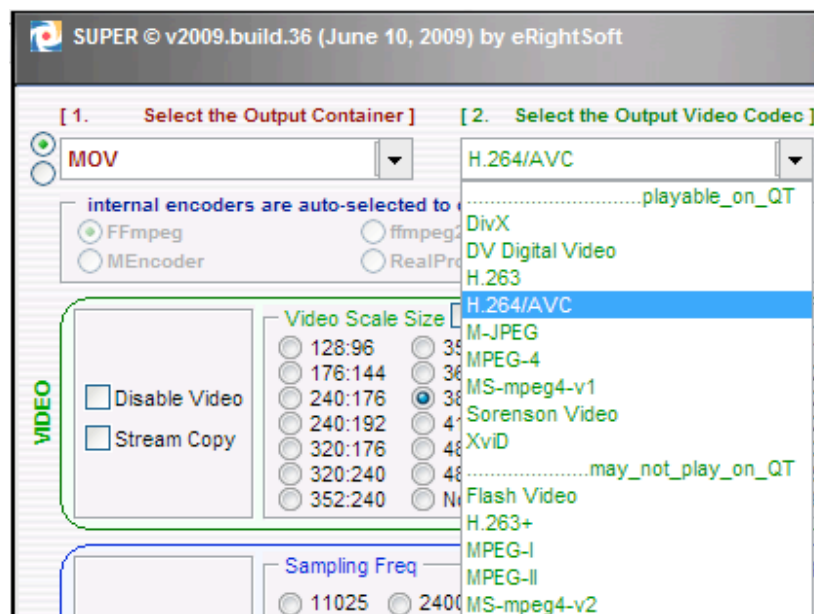
Conteneur standard sur Mac.
(= fichier d'emballage)

CODEC à éviter :

- DivX : grand public, trop compressé, impossible à insérer dans un logiciel audionumérique
- DV : caméscopes, gros fichiers, pas assez compressés

CODEC à privilégier :

- H.264, MPEG-4, MPEG2
- M-JPEG
- Sorenson vidéo (ancien, moyennement performant, mais très compatible)



c) MP4

Container moderne sur Mac, PC, Internet, multimédia.
(= fichier d'emballage)

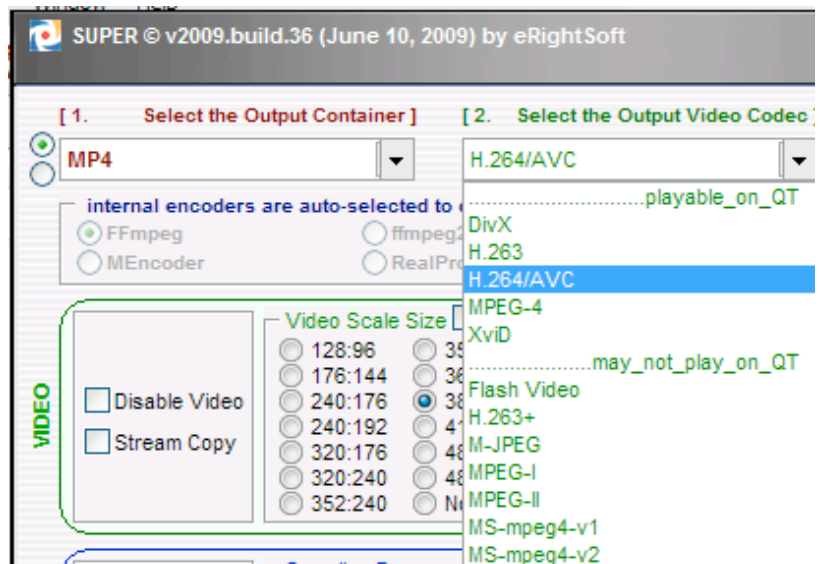
CODEC à privilégier :

(idem ci-dessus)

- H.264, MPEG-4
- M-JPEG

CODEC à éviter :

- DivX, XviD
- Flash Vidéo



3) Relations entre conteneur et codec

Un fichier vidéo est défini par son conteneur ET son codec.

Le choix du conteneur est souvent lié au choix des codec audio et vidéo internes au fichier.

Exemple : Un conteneur MPEG-2 (technologie ancienne) ne peut contenir que du flux vidéo MPEGv2 et des flux audio MP1, MP2, MP3 ou AAC.

Le choix du conteneur et du codec sont liés aux applications et logiciels qui vont être utilisés.

Exemples :

- Windows Media Player peut lire un fichier *Quicktime* qui utilise un codec *h.264*,
- mais ne peut pas lire un fichier *Quicktime* qui utilise un codec *Apple Intermediate codec*.
- Le format DIVX est utilisé pour compresser de façon sévère des films long métrage,
- mais ce format n'est pas lisible par les logiciels audionumériques.

4) Les autres paramètres de la vidéo

« Super » propose de faire des conversions vidéos, en 4 étapes :

1) choix du conteneur

2) choix du codec vidéo

3) choix du codec audio : stéréo, 48kHz, AIF ou WAV

4) Observez ensuite :

les dimensions : 768 x 576, 1280 x 720, 1920 x 1080, etc.

les proportions : 4 :3 ou 16 :9 (dépend de la vidéo originale)

le nombre d'images par seconde (25 ips en général – à ne pas changer)

la compression ou le débit

Compression

- La qualité d'un flux vidéo est déterminée par son débit binaire : c'est-à-dire le nombre de bits par seconde exprimé en **kbps** (kilobit par seconde).

Le débit binaire ou le taux de compression dépendent du contenu vidéo (images complexes et mobiles ou simples et statiques) et de la taille de l'image.

Dimensions

Le DVD-vidéo imposait un format d'image : 768 x 576
 Les proportions 4:3 ou 16:9 étaient souvent obtenues par déformations des pixels.

Aujourd'hui (2017), il existe plusieurs standards de vidéo haute-définition (HD) :

- 720p ou 720i 1280 x 720 pixels
- 1080p ou 1080i 1920 x 1080 pixels
- La lettre (p ou i) correspond au standard progressif ou entrelacé (*interlaced*).

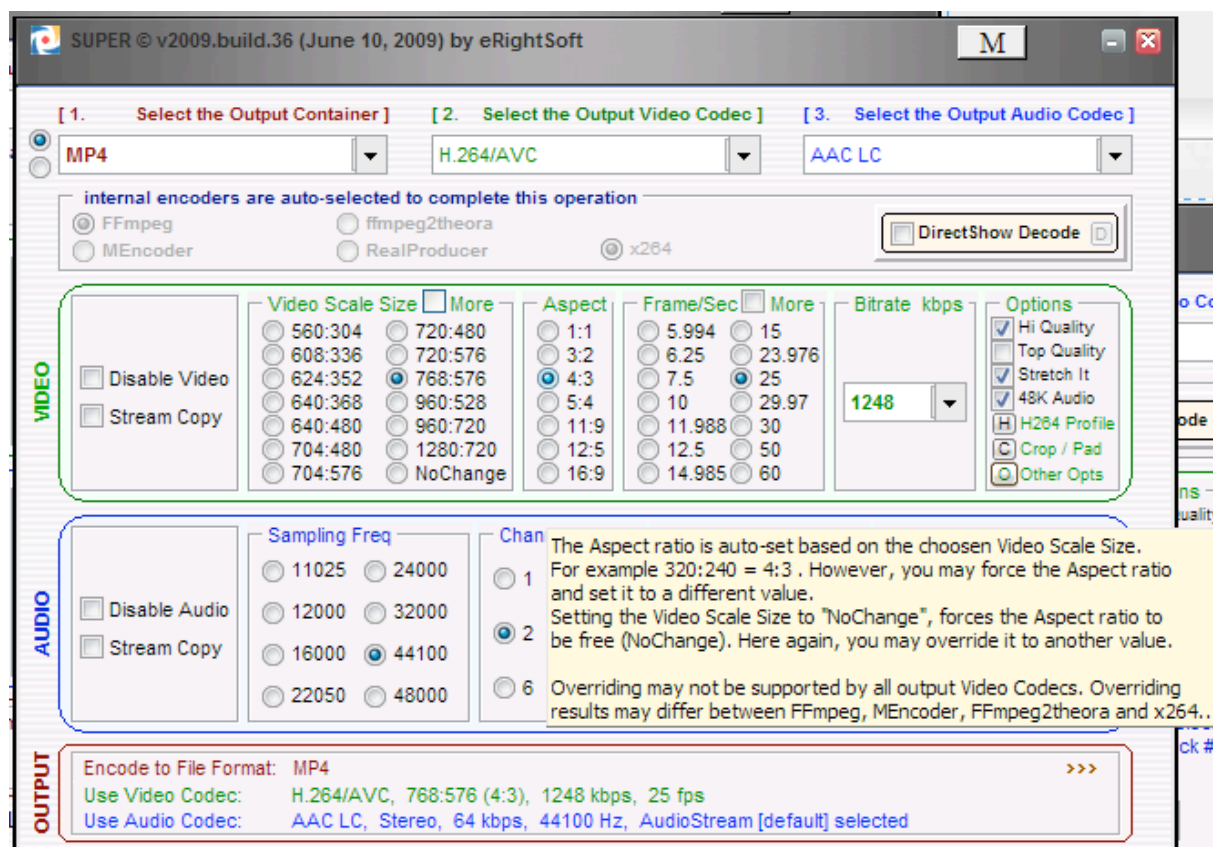


Figure 5 : Vue d'ensemble des multiples réglages vidéo
 container : MP4, codec H.264.

5) Conclusions / Conseils

a) Vous recevez une vidéo : analysez ses caractéristiques !

Y compris la taille du fichier vidéo !

Exemples : un fichier MP4 sans son pèse environ 10 à 15 Mo/minute
 un DVD-vidéo de 1h30 pèse 4 Go en MPEG2
 donc un exercice de 6 minutes ne peut pas peser 1 Go !!!

b) Vous composez de la musique à l'image

Pour des raisons techniques (contraintes du logiciel audionumérique), vous pouvez être amené à :

- changer de codec (qui ne serait pas « digéré » par FruityLoop, Cubase ou Logic...)
 - réduire les dimensions (afin d'optimiser l'occupation de votre écran)
 - compresser plus (afin d'économiser de la RAM et de la puissance du processeur d'affichage)
- Ces modifications sont internes et personnelles.

Le résultat de la composition est un fichier audio qui sera collé sur la vidéo d'origine.

c) Rendu

- demandez au réalisateur le format souhaité du rendu ;
- en l'absence d'information, contentez vous de coller votre fichier audio sur la vidéo originale, sans changer son codec ! ni ses dimensions ! ni surtout ses proportions !

Pour info, la taille du fichier audio + vidéo devrait généralement augmenter de 10-25 % pas plus.

Exemple : un fichier audio stéréo 16 bits 44.1 kHz pèse 10 Mo/minute

d) C'est compliqué !

Il y a de nombreuses causes d'erreurs !

Si vous souhaitez vraiment communiquer,

si vous souhaitez avoir l'air d'un pro,

- **comprenez ce que vous faites avec les vidéos qui vous sont confiées ;**
- **demandez au réalisateur le format souhaité du rendu ;**
- **ne changez pas le format (codec, dimensions, proportions, ips...) (surtout en faveur de formats « touristiques ») ;**
- **faites des tests avant d'envoyer = ouvrez la vidéo dans Quicktime (Mac ou PC) ou VidéoInspector (PC) et analysez de nouveau toutes les caractéristiques. Comparez les avec celles de la vidéo d'origine ;**

Vérifiez bien la taille du fichier vidéo (à comparer avec l'original).

Note : a priori, ce que vous envoyez (vidéo + musique) doit être (à la louche) 5 ou 10% plus « lourd » que l'original.

Si vous envoyez du n'importe quoi :

- cela risque de bloquer la chaîne de production
- et vous allez passer pour un incompetent.

JPEG: DE L'IMAGE FIXE A L'IMAGE ANIMÉE

Le principe de base de JPEG (Joint Photographic Expert Group) consiste à utiliser les caractéristiques limitées de l'œil humain de façon à économiser des données. Ainsi, le fait que nous soyons moins sensibles aux variations de couleur (chrominance) qu'aux variations d'intensité lumineuse (luminance) permet aux algorithmes d'éliminer certaines informations à notre insu et d'obtenir des images 10 à 25 fois moins gourmandes en data que l'original.

En fait, l'image de départ est décomposée en petites sous-images de 8X8 points. Sur chaque sous-image, on opère une Transformée en Cosinus Discrète (TCD ou DCT en Anglais), procédé que l'on peut comparer à une transformée de Fourier (FFT) pour l'analyse spectrale en son. On obtient ainsi une matrice 8x8 de laquelle on extrait le premier point (de coordonnées (1,1)) qui servira de valeur moyenne de référence en fonction de laquelle seront exprimés les 63 autres points de la matrice. Prenons le cas d'une zone rouge uniforme, le premier point en haut à gauche aura la valeur correspondant à la couleur rouge tandis que tous les autres seront exprimés par des zéros. Il sera ensuite facile d'économiser des données en numérisant avec un algorithme de compression classique de type Huffman.

Pour un signal vidéo européen on aura donc 25 traitements de ce type en une seconde !

MPEG VIDEO (Moving Picture Expert Group)

La norme MPEG s'attache à détecter la redondance temporelle (sic!) d'une séquence d'image animée. En clair, dans une scène filmée où le décor est fixe mais où les acteurs bougent, l'algorithme sera à même d'économiser les données relatives à la numérisation du décor puisqu'il reste identique et essaiera de rendre compte des déplacements par des méthodes d'analyse et de déduction. Chaque image est donc décomposée en petites sous-images comme pour JPEG. L'algorithme s'attachera ensuite à déceler les sous-images qui sont relativement identiques parmi les images qui suivent et qui précèdent l'image concernée. De plus, le système ne va pas numériser toutes les images et certaines seront créées par interpolation grâce à l'analyse des images adjacentes. Ainsi, plus la séquence sera prédictible (zoom, travelling) plus les résultats seront satisfaisants. Toutefois, il est plus facile de prédire lorsque l'on connaît le passé et le futur et ce type de procédé reste difficile dans le cadre d'opérations temps-réel car il réclame une forte puissance de calcul et d'analyse prédictive qui peut dépasser en temps la quantité d'image à compresser. Toutefois si ce temps de conversion est pris au transfert, le taux de compression peut être gigantesque et la qualité de l'image sans faille, d'où une très grande économie de données.

Recommandations de débit pour l'encodage vidéo

Taille	Format	Débit minimum recommandé
640 x 480	4:3	Les vidéos qui comportent peu de mouvements, tel que les personnes parlant en gros plan, peuvent offrir de très bons résultats avec seulement 900 kbit/s. Pour une qualité optimale générale, quel que soit le niveau de mouvement, utilisez au moins un débit de 2 672 kbit/s. Ceci s'applique également aux sources 720 x 480, car ce sont des vidéos en 640 x 480 avec un format différent.
640 x 360	16:9	Dans la plupart des cas, le débit vidéo peut descendre jusqu'à 2 132 kbit/s sans problème.
1 024 x 768	4:3	Ce format est principalement utilisé dans les captures d'écran ou les sources HD redimensionnées. Dans le cas de la vidéo, un débit de 4 864 kbit/s est suffisant. Dans le cas de captures d'écran, une telle quantité de données est superflue, car il y a peu de mouvements d'une scène à l'autre. Le débit minimum du 720p (3 136 kbit/s) suffit. C'est en fait légèrement supérieur à ce dont nous avons besoin, mais un débit plus grand autorise une production d'images clés plus fréquente. Comme nous ne voulons pas avoir une rémanence de la souris, nous vous recommandons de disposer d'images clés fréquentes placées à distance égale l'une de l'autre, et ce, deux fois par seconde (soit toutes les 15 images dans le cas d'une capture d'écran très fluide). Si le nombre d'images par seconde de la capture d'écran n'est pas de 30 (ou de 29,97), le débit peut être ajusté en parallèle. Donc, si vous disposez d'une capture d'écran à 15 images par seconde, le débit sera de 1 568 kbit/s. C'est une bande passante assez conséquente pour une capture d'écran, mais néanmoins nécessaire pour éliminer la rémanence et l'effet de découpage de l'image qui se produisent souvent lors du passage au H.264.
1 280 x 720	16:9	Le débit de ce format peut descendre jusqu'à 3 136 kbit/s pour un présentateur en gros plan. Lorsque le nombre de mouvements augmente ne serait-ce qu'un peu, le débit doit être augmenté pour compenser. Dans le cas d'un faible nombre de mouvements (le sujet bouge, mais la caméra reste statique), si l'éclairage est optimal, un débit de 4 512 kbit/s suffit. Pour une vidéo riche en mouvements, ou pour compenser d'autres facteurs, un débit de 5 928 kbit/s suffit (avec un peu de marge).
1 920 x 1 080	source étirée en 16:9	Plusieurs caméras HDV et AVCHD, professionnelles ou non, utilisent des capteurs 4:3 et étirent l'image d'une résolution de 1 440 x 1 080 en 1 920 x 1 200 pour imiter la haute définition. Au final, cela donne souvent une image aussi fine en 960 x 540. C'est pour cette raison que nous vous recommanderons ce format pour 960 x 450 et 1 920 x 1 080. En guise de compromis, vous pouvez charger vos vidéos en 720p ; les paramètres 720p précédents s'appliquent ici. Le format 1 920 x 1 080 exige un débit de 8024 kbit/s pour les vidéos avec beaucoup de mouvements. Une vidéo de présentateur en gros plan peut descendre jusqu'à 6 000 kbit/s, mais un débit de 8 024 kbit/s couvre la plupart des cas de figure. Un tel débit couvre une durée légèrement supérieure à une heure et demie. Avec le format 960 x 540, vous pouvez utiliser le même débit de 3 136 kbit/s qu'une vidéo de présentateur en gros plan en 720p si vous ne prévoyez pas de redimensionnement vers le haut et si le mouvement est assez faible.

Compression – exemple de calcul du débit binaire

Exemple de calcul du débit binaire d'un vidéo non-compressé:

- Nombre de pixels dans une image : $640 \times 480 = 307\,200$ pixels
- Nombre de bits par pixel : $8 \text{ bits/couleur} \times 3 \text{ couleurs (RVB)} = 24 \text{ bits}$
- Nombre de bits par image : $307\,200 \text{ pixels} \times 24 \text{ bits/pixel} = 7\,372\,800 \text{ bits} = 7\,372,8 \text{ kb}$
- Nombre d'image par seconde : $25 \text{ images/seconde}$
- Nombre de bits par seconde : $7\,372,8 \text{ kb} \times 25 \text{ images/secondes} = 184\,320 \text{ kbps}$
le débit binaire

- Taille d'un fichier non-compressé d'une heure : $184\,320 \text{ kbps} \times 3\,600 \text{ s} = 663\,552\,000 \text{ kb}$
 $= 82\,944 \text{ Mo} = 82,9 \text{ Go}$
- Taille d'un fichier compressé au débit de 1000 kbps :
 $1000 \text{ kbps} \times 3\,600 \text{ s} = 3\,600\,000 \text{ kbit} = 450\,000 \text{ Ko} = 450 \text{ Mo}$

On comprend ici la nécessité de la compression.

Et l'efficacité du DIVX (par exemple) qui va ramener le poids du fichier à environ 4 Go.