

# Multimedia data management course 1

Johnny Nguyen

**Abstract—**Ce document test une synthèse du premier cours de gestion de données multimedia.

## I. INTRODUCTION

Le multimedia n'est pas du texte. Hadoop, le précurseur du Big Data, il utilise un cluster pour manipuler d'énormes données, le système est distribué et robuste. HDFS, lecture sur le disque en parallèle de manière synchrone. HIPI, programmé en JAVA pour traiter du HDFS. Python est le langage le plus connu et beaucoup de bibliothèques ont été importés pour faire du traitement scientifique. Spark, est un langage magique pour le traitement de données massif, mais c'est données doivent être serializable. Le but sera de chercher des visages dans un flux (le Machine Learning est envisageable).

## II. AUDIO SIGNAL

C'est un signal qui doit être entendu. Notre oreille écoute une certaine gamme de fréquence. Une bonne oreille, vers 20 ans, écoute entre 20hz et 20khz. L'objectif est de donner un son entre 20 hz et 600 khz.

### A. The voice

Toute la bouche nous permet de construire le signal. Le poumon permet de générer un signal, modulé par la gorge, la bouge et le nez permettent de sortir ce signal transformé. Ce sont des vibrations sinusoïdes. Nous pouvons créer un signal aléatoire en faisant des soupirs. La comparaison entre un soupir et une élocution nous montre une différence lair sur les signaux obtenus. Dans la voix, il faut préserver le signal élevé et faible. Un instrument de musique subit une vibration, le son est amplifié et la vibration s'atténue.

### B. Masking a sound

Reconnaître rapidement un son nous permet de survivre dans la jungle. Retirer une ondulation dans un signal s'appelle le masquage. Notre cerveau corrige les erreurs dans un signal.

## III. COMPARAISON AVEC LE SIGNAL D'UNE IMAGE

Le capteur des yeux sont élaborés. Le niveau de gris est très sensible au niveau de la périphérie de la rétine. Plus, il y a de battonets, plus la précision est élevée. Si nous avons une image, la luminosité est importante.

### A. Edges

Notre cerveau sait reconnaître les objets très rapidement. La luminosité est importante. C'est très important pour segmenter une image. Notre cerveau sait aussi corriger les erreurs sur les contours. (le masquage) Ce contour sera interprété de manière différente.

### B. Textures

C'est un pattern que nous retrouvons fréquemment dans une image. Cette petite image est définie par sa fréquence, sa couleur et sa distribution. Nous pouvons apercevoir des formes même si sur l'image, cette forme n'est pas censée exister. C'est une interprétation faite par notre cerveau.

## IV. COMPARAISON AVEC LE SIGNAL VIDÉO

Une vidéo est une succession d'images. Chaque scène est interprété par notre cerveau. Ce signal peut être interprété seulement si l'image ne subit pas un grand mouvement. Les statistiques données sur le taux de compression d'image ne sont pas forcément vrais. La détection de contour n'est pas possible dans ce cas. En augmentant la taille de l'image, la qualité de l'image est réduite car le masquage n'est pas appliqué correctement.

### A. Sampling and subsampling

Chaque signal peut avoir une variation de dimension selon la qualité de celle-ci.

- Pour le son, un haut parleur ne doit pas faire le même son qu'un autre sinon le signal est faux.
- Pour les images, il ne faut pas sous-echantillonner les images.

En modifiant la fréquence d'une ondulation, de la taille d'une image ou le nombre d'image successif nous pouvons affecté le repliement de spectre (spectral aliasing)

### B. Compression

Une compression permet de comprimer un signal pour le transmettre. Une bonne compression doit pouvoir réduire la taille des données et faire transiter rapidement cette information. "Qu'est-ce qu'un compandor ?", il faut adapter l'erreur en fonction de la taille du signal. La quantification vectoriel permet de regarder plusieurs dimensions à la fois. C'est long et abandonné. La quantification prédictive quantifie le signal et prédit le prochain signal pour en déduire une erreur. Le schéma MPEG-2 nous permet de prédire en fonction de l'image précédente et suivante.

### C. Transformations

Transformation de Fourier : seulement sur des signaux échantillonnés. Sur une discontinuité, cette transformation n'est pas viable.

Transformation de cosinus ; une bonne transformation sans les inconvénients de la transformée de Fourier.

## REFERENCES

- [1] <http://www.i3s.unice.fr/mathieu/>