

# **TP3 - Indexation de fichiers multimédia: Décomposition de prises de vues**

INF4710 - Introduction aux technologies multimédia

**Nom de l'étudiant : Félix La Rocque Carrier**  
**Matricule : 1621348**

**Nom de l'étudiant : Mathieu Gamache**  
**Matricule : 1626377**

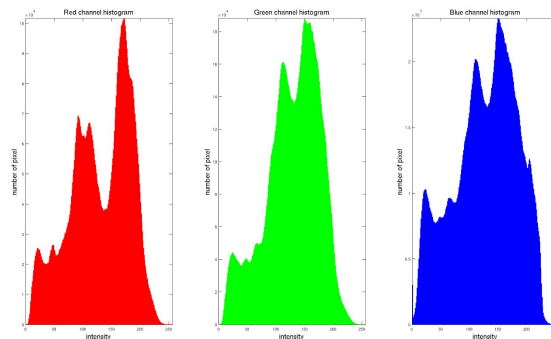
<b>Définition des concepts</b>	<b>1</b>
Histogramme :	1
Convolution:	2
<b>Présentation des fonctions de base</b>	<b>2</b>
Fonction Histo (Histogramme) :	2
Fonction Convo (Convolution) :	3
<b>Présentation logique de détection de transitions</b>	<b>3</b>
<b>Présentation des paramètres</b>	<b>4</b>
<b>Discussion avantages/inconvénients pour stratégie proposée</b>	<b>5</b>
<b>Présentation et discussion des résultats de détection</b>	<b>5</b>
<b>Discussion des améliorations possibles</b>	<b>6</b>

# Définition des concepts

## Histogramme :

L'histogramme d'une image sert à représenter sous forme graphique (ou encore dans une forme facilitant les comparaisons) les couleurs composant une image.

Dans notre cas, nous allons séparer l'image en ses couleurs de bases (Rouge, Vert, Bleu) pour mesurer l'intensité de chacun et générer un graphique sous la forme suivante:



Cet information nous permettra de comparer image par image le vidéo pour détecter des potentiels changement de couleurs important, indiquant un changement de scène.

Cet comparaison par les couleurs n'est pas assez précise pour détecter un changement de scène alors nous devons aussi faire une seconde comparaison.

## Convolution:

La convolution est une méthode de modification de l'image pour en faire ressortir certaines propriétés. On peut en faire ressortir des contours, ou encore enlever le bruit dans des images de basse qualité.

Avec cet opération, nous pouvons avoir une autre base pour la comparaison. Si les contours des objets dans la scène changent brusquement, nous pourrions interpréter cela comme un changement de scène.

Du même principe, si nous avons des images de faible qualité, nous pourrions les traiter pour éliminer une partie du bruit et diminuer le taux de faux dans la détection d'une nouvelle scène.

# Présentation des fonctions de base

## Fonction Histo (Histogramme) :

Dans cet fonction, nous faisons le décompte des intensité pour chaque couleurs. Comme mentionné plus haut, le but est d'obtenir une base similaire de comparaison entre image d'une scène.

Nous parcourons chaque image en regardant les intensités de chaque couleurs par pixel. Pour chacune de ses intensités, nous déterminons dans quel intervalle d'intensité (histogramme) cet intensité se situe en incrémentant le décompte de cet intervalle.

Pour déterminer l'intervalle, nous utilisons la valeur maximale connue (255) et le nombre d'intervalle. En sachant la quantité d'intervalle (N) nous pouvons alors déterminer l'intervalle de n'importe quel valeur entre 0-255. Ceci permet d'éviter d'être trop sensible aux petites variations de couleurs: provenant de la caméra et tout simplement que dans la nature ce n'est pas tout à fait la même couleur. Pour faire simple, on regroupe ensemble les couleurs similaires, le nombre de groupe est égale à N.

## Fonction Convo (Convolution) :

Pour appliquer la convolution (modification de l'image), nous parcourons chaque pixel de l'image en regardant le pixel visé et ces voisins. Selon l'objectif de la convolution (la matrice de convolution spécifié) nous allons obtenir des résultats différents.

Lorsque le pixel visé est un pixel de frontière, nous avons pris le choix de mettre à la valeur la plus proche, la valeur des pixels à l'extérieur de l'image.

La matrice spécifier peut faire ressortir plusieurs caractéristiques de l'image. Une matrice rempli de valeurs identiques appliquera un effet de flou tandis qu'une matrice ayant deux lignes pleines fera ressortir les frontières verticales ou horizontales (matrices de sobel).

-1	0	+1
-2	0	+2
-1	0	+1

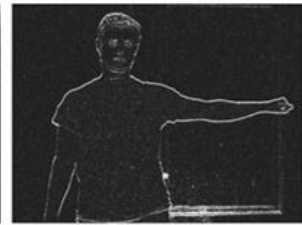
Gx

+1	+2	+1
0	0	0
-1	-2	-1

Gy



Vertical sobel



Horizontal Sobel

## Présentation logique de détection de transitions

Notre première idée pour détecter un changement de scène de façon facile est la comparaison d'histogrammes. Nous avons premièrement fait la différence entre l'histogramme de chaque frame du vidéo et avons déterminé un changement de scène lorsque la différence entre les histogrammes étaient plus élevés qu'un certain point.

Cet approche naïve comprend plusieurs problème, surtout lors d'un changement de scène fondue. Dans ce cas, on détermine un grand nombre de changement de scène consécutifs.

Pour combler ce manque, nous avons ajouté un deuxième seuil pour commencer à détecter là où la scène est en transition, par la même méthode de comparaison d'histogrammes.

Pour pouvoir détecter une vrais transition fondue et non pas qu'une scène en mouvement, nous devons utiliser une autre technique de détection.

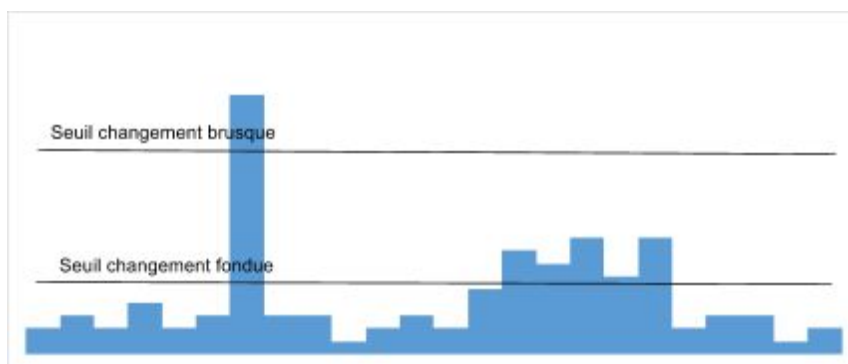


Figure: Histogramme exemple de différence d'histogrammes

Il nous manque un dernier élément, comme dans la dernière scène, on détecte une transition de scène lorsque le camion blanc avance.

# Présentation des paramètres

Pour notre système de détection, nous avons 4 paramètres :

**partition\_count:** Le nombre de partitions à faire dans l'histogramme de comparaisons. Plus il est élevé plus l'algorithme est sensible aux variations de couleurs.

**low\_trigger:** Le seuil minimale à laquel on peut trouver un changement de scène fondue. Plus il est bas, plus il pourra détecter les changements de scènes subtile (par exemple un long fondue), par contre il y a des risques de détecter des faux positifs.

**high\_trigger:** Le seuil maximale à laquel on ne détecte pas des changement de scène fondue mais tous les changement de scène brusque. Plus il est élevé, moins on a de chance de détecter des faux positifs. Par contre, si trop élevé il se pourrait qu'il ne détecte pas un changement de scène dont les couleurs sont dans le même tons.

**frame\_check\_interval:** Le nombre de frame entre chaque comparaison d'histogramme pour détection d'une transition fondue. En modifiant cette valeur on réduit le coût de l'algorithme de détection ainsi qu'on peut diminuer le nombre de faux positif.

C'est ces paramètres qui ont le plus grand impact sur la détection des deux types de transition. En variant ces paramètres, nous sommes arrivé au point où ceux-ci ont été modifiés spécifiquement pour la séquence vidéo donner.

## Discussion avantages/inconvénients pour stratégie proposée

Notre objectif était d'avoir un algorithme le plus léger possible, pour être en mesure de l'appliquer à du flux temps réel ou avec un haut taux de frame rate. Après quelque test on a trouvé que calculer la convolution est une opération coûteuse qu'on veut éviter.

Notre algorithme est idéal dans le cas où on veut détecter toutes les vrais transitions mais qui détecte aussi les grand changement intra-scène. Supposons qu'on voudrait découper la vidéo en plusieurs scènes, il peut être intéressant de découper une même scène mais où se produit un grand changement: par exemple une explosion ou un grand camion qui prend une bonne partie de l'image.

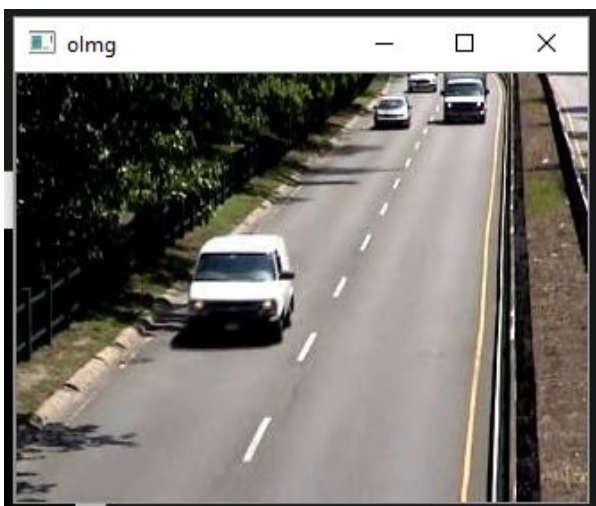
Par contre, cette simplicité est en même temps un inconvénient dans le cas où on veut détecter seulement les vraies transitions.

De plus utilisant, si plusieurs scènes sont dans les mêmes tons de couleurs, notre technique va être incapable de le détecter. Par exemple, s'il y a plusieurs vidéos qui regardent un parc sous différents angles, on ne pourrait détecter lorsqu'on change de flux vidéo, ce qui pourrait être un comportement désiré, mais probablement pas.

## Présentation et discussion des résultats de détection

Nos résultats sont assez précis pour cette séquence vidéo. Toutes les transitions brusques sont bien détectées et ils le sont vite.

Cela est bien sûr expliqué par le fait que ces changements brusques causent une grande différence entre deux histogrammes d'images consécutives. Ayant une stratégie de détection de transition basée sur les différences d'histogrammes, ce résultat n'est point surprenant.



Pour ce qui est de la détection des transitions fondues, nous pouvons les détecter mais avec un petit délai. Puisque leur détection passe par une sélection de 1 frame par plusieurs (**frame\_check\_interval**), il y a un certain délai entre le moment de début du fondu et celui de détection.

Nous pouvons tout de même éliminer certaines section du vidéo ayant les caractéristique au niveau de l'histogramme d'une transition fondue (partie finale avec le camion blanc) avec cet technique.

## Discussion des améliorations possibles

Plusieurs améliorations sont possible. Dans notre cas, nous n'avons pas utilisé l'outil de la convolution pour détecter les contours dans la scène. Si on veut garder notre faible impact il aurait été possible d'utiliser une combinaison des 2: à chaque intervalle de frame utiliser la différence de l'histogramme (qui n'est pas couteux) et si on détecte un changement de scène on pourrait utiliser la convolution, ce qui aurait éliminé la majorité des faux positifs. De plus, cela aurait aussi pu nous donner plus d'information sur le type de transition dans la scène (fade-in, fade-out,...).

Nous aurions aussi pu calculer plusieurs types de différences de l'histogramme, attribuant des poids à chacun pour obtenir une meilleure précision.

Avec une séquence vidéo plus longue et des annotation de début de transition tant brusque que fondue, nous aurions pu générer un modèle et obtenir des valeurs de paramètres beaucoup plus précis dans un cas général.

Nous avons finalement des amélioration possible dans le temps requis pour détecter une transition fondue. Comme nous ne considérons pas plusieurs frames dans le calcul de détection des transitions fondues (**frame\_check\_interval**) nous ne détectons la transition que quelques ms après.