



Images animées et indexation vidéo

PREDICTION OF VISUAL SALIENCY IN VIDEO

SEQUENCES BY DEEP CNN

WOLSKI Axel

DEDIEU Thomas

Chargé de TD : CHAABOUNI Souad

Master 2 Informatique Spécialité Image et Son 2017-2018

Octobre 2017

1 Organisation du projet

TP réalisé en binome: **Thomas DEDIEU & Axel WOLSKI**

Pour compiler et exécuter le programme Lab4 il faut se placer dans le dossier `Lab4/build` et faire la commande suivante:

```
cmake .. && make  
./Lab4 ../Data/model/deploy.prototxt ../Data/model/train/train_iter_11718.caffemodel  
../Data/model/PARKINSON_flou_mean4k.binaryproto ../Data/model/label.txt  
../Data/video/test_rgbframe.avi ../Data/video/test_tmpframe.avi 50  
../Data/ChaboNet_Saliency/
```

Après exécution de `Lab4`, vous trouverez dans le dossier `Lab4/Data/ChaboNet_Saliency` tous les Saliency générés pour chaque frames de la vidéo.

Dans le dossier `Lab4/Data/ChaboNet_Saliency/listFrames` vous trouverez chaque frames en RGB de la vidéo.

Vous trouverez également, après exécution du code, dans le dossier `Lab4/build` deux fichiers texte, `file-list-frames.txt` et `file-saliency-frames.txt` qui contiennent respectivement les chemins d'accès aux frames RGB et saliency.

Pour compiler et exécuter le programme `Lab4_bis` il faut se placer dans le dossier `Lab4_bis/build` et faire la commande suivante:

```
cmake .. && make  
./create_frames_salMix ../../Lab4/build/file-list-frames.txt  
../../Lab4/build/file-saliency-frames.txt ../../data/
```

Après exécution de `Lab4_bis`, vous trouverez dans le dossier `Lab4_bis/data` toutes les frames mélangeant Saliency et RGB de la vidéo.

2 Déroulement du projet

2.1 Principe

Dans un premier temps nous avons eu à comprendre ce qu'est la **Prediction of Visual Saliency** dans une vidéo. Cette prédiction d'objet saillant dans une vidéo est tout simplement l'objet ou l'événement qui va capter notre regard et que l'on va regarder. Par exemple dans un plan immobile nous allons d'abord regarder les objets avec des couleurs vives et si un objet se met en mouvement ou rentre dans la scène nous allons le regarder.

Nous avons donc commencé par tester des images RGB et leurs images TMP correspondantes pour tester leur saillance.

3 PREDICTION OF CHABONET SALIENCY MAP

Dans cette partie nous avons eu à coder la partie manquante dans Lab4.cpp afin de créer les images correspondantes à la saillance d'une image. Le principe est de tester patch par patch l'image (d'une taille que l'on choisie, 25 dans notre cas) pour savoir si il est saillant ou non. Si ce patch est saillant on applique un filtre sur celui-ci et on construit la nouvelle image à l'aide de la modification apportée sur chaque patch.

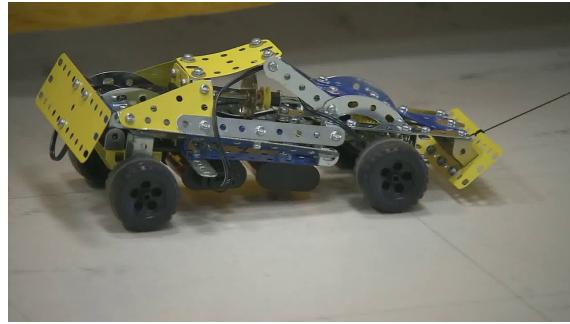


Figure 1: Frame original en RGB

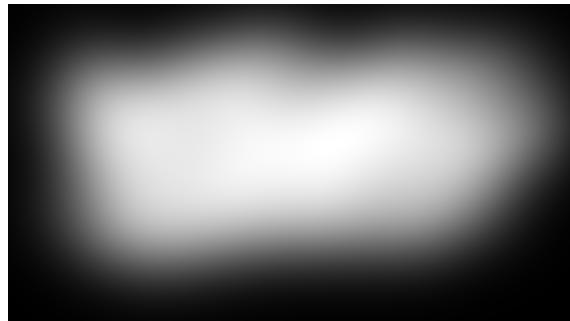


Figure 2: Frame saliency générée

Les images saillantes obtenues sont toutes en noir et blanc, mais à l'aide d'une fonction on peut lui appliquer les couleurs de la colormap.

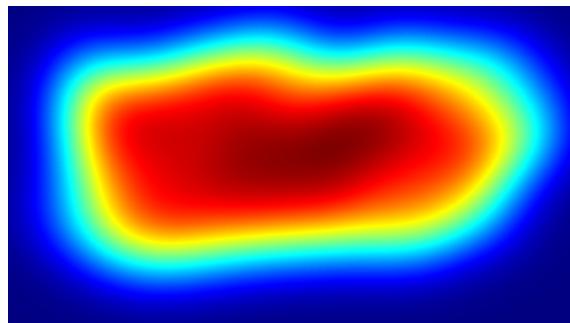


Figure 3: Frame saliency - ColorMap

4 HEAT MAP

L'obtention des images saillantes ne nous permet pas réellement de nous rendre compte de ce que l'on obtient par rapport à l'image d'origine. C'est pourquoi à l'aide du code dans `Lab4_bis` nous pouvons mixer l'image `rgb` et l'image `saillante` afin de nous rendre compte du résultat.

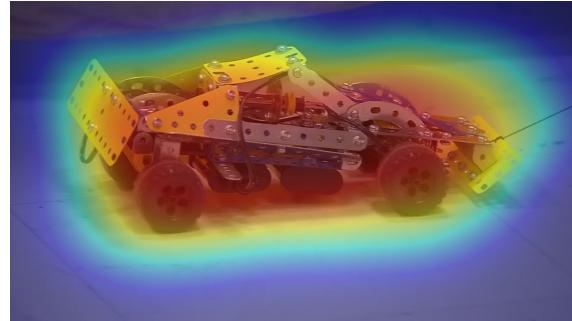


Figure 4: Exécution du code Lab4_bis avec la voiture

Nous pouvons constater que notre saillance suis bien le mouvement de la voiture tout au long de la vidéo. Mais aussi, que lorsque d'autres objets rentrent dans la scène, la saillance les prend également en compte. De plus la caméra et la voiture sont en mouvement et la caméra suit la voiture, notre regard va donc surtout suivre la voiture.

En revanche en ce qui concerne les résultats obtenus avec le train, ils sont un peu plus intrigants. Nous aurions pus penser que le regard allait suivre le mouvement du train, or d'après le résultats, la quasi totalité de la vidéo est saillante.



Figure 5: Frame rgb du train

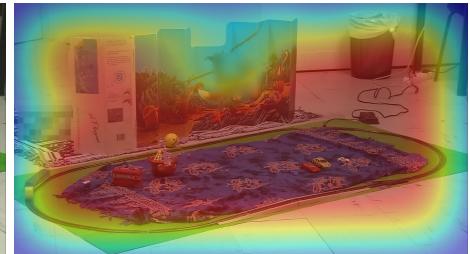


Figure 6: Resultat lab4_bis sur la frame du train

Nous en concluons donc que notre regard est aussi capté par le décors de la scène et pas seulement par le mouvement de l'objet qui y est présent.