

Projet Libre - Rapport bibliographique

Jean CAILLÉ

Octobre 2013 - Juin 2014

1 Saliency

Afin de mieux comprendre les problématiques de traitement de vidéos, et de construire une base bibliographique, j'ai commencé par étudier la saillance dans les images. Fondamentalement, le problème consiste à trouver dans une image les objets intéressants, la définition des termes "objets" et "intéressant" étant assez flous.

1.1 Définition du problème

Le problème de Saillance dans une image est en soit mal posé. Il n'existe pas en effet de référence globale, et la définition d'un "objet intéressant" peut varier de personnes en personnes. Généralement, la qualité des algorithmes décrits est évaluée par comparaison avec des résultats obtenus par des opérateurs humain sur un jeu de données particulier. La première consiste à suivre le regard d'une personne lorsque l'image lui est présentée grâce à des outils d'eye-tracking. La seconde consiste à demander à des opérateurs humains d'indiquer à la main quels parties de l'image paraissent intéressantes.

1.2 Objectifs de l'analyse bibliographique

Chez Stupeflix : Obtenir des résultats corrects, en un temps raisonnable (jj 1s) avec une implémentation sur iPhone.

1.3 Itti's algorithm

1.4 Saliency detection : A spectral residual approach

1.5 Salient Region Detection and Segmentation

L'idée principal de cet article est similaire à celle d'Itti et al, mais se distingue par plusieurs points clés. La saillance est toujours obtenue par comparaison entre un pixel et une région avoisinante, mais cette comparaison est ici effectuée dans un espace échelle à 4 dimensions (3 dimensions de couleur, 1 dimension d'échelle). Les cartes de saillance [A DEFINIR] obtenus sont à priori de meilleur qualité que ceux résultant de la méthode d'Itti et al., et permettent selon les auteurs une segmentation peu couteuse et sans supervision des objets saillants dans l'image.

1.5.1 Espace Couleur CIELab

La saillance des objets à beaucoup à voir avec la perception par l'oeil humain des couleurs et des formes. En particulier, les distances entre couleurs, plus que les couleurs en elle mêmes joue un rôle important. Or, l'espace de couleur classique utilisé en informatique (i.e. espace RGB) ne retranscrit pas bien les différences de couleurs perçues

[[FIGURE — Deux paires de couleurs à distances euclidiennes égales, mais à distances perçues importantes]]

L'espace CIEL*a*b*, développé en 1976 par la Commission International de l'Éclairage a quand à lui été créé pour que les distances calculées entre deux couleurs (grâce à la distance euclidienne dans R^3) correspondent aux différences perçues par l'oeil humain. Une couleur est représentée par trois composantes L, a et b. La clarté L est définie de 0 (noir) à 100 (blanc), la couleur est encodée par deux valeurs a et b appelées chrominance, qui (en informatique) varient de -128 à 128 et peuvent donc être codées sur un octet chacun. Cet espace étant défini de manière absolue, la formule de conversion entre RGB et CIELab dépend généralement de la calibration de l'écran employé (en particulier la température du blanc pour l'écran employé). A noter que cette conversion n'est pas linéaire. Le standard whitepoint65, employé par plusieurs grands acteurs de l'image informatique (Adobe Photoshop, OpenCV, ...) donne pour formule de conversion :

[[FORMULE]]