

PROPOSITION DE STAGE

Campagne 2012

(*) Supprimer la mention inutile



Laboratoire L3i

Sujet Stage :

Analyse de la trajectoire oculaire pour la classification d'images en fonction de retours utilisateurs implicites

Résumé du travail proposé :

Dans ce stage, on se propose d'utiliser un dispositif de eye-tracking [Facelab] pour construire un signal de retour de pertinence [Wood,1998 ;Huang,2001] implicite dans un contexte de recherche d'image par le contenu [Veltkamp, 2002]. Nous proposons d'utiliser un paradigme de « préférence visuelle » (introduit by Fantz in 1958 [F58]) qui consiste à présenter 2 images côte à côte et déterminer laquelle est préférée par l'utilisateur via l'analyse de ses trajectoires visuelles. On propose par ailleurs de coupler ce système à un apprentissage actif déjà développé au laboratoire [Picard,2010]. On s'attend à ce que le retour visuel soit beaucoup plus rapide et permette donc de faire converger l'algorithme d'apprentissage beaucoup plus rapidement.

Mots clés :

Recherche d'images par le contenu, CBIR, eye-tracking, retour de pertinence

Informations complémentaires :

Encadrant(s) : Arnaud REVEL

Axe thématique : « Données complexes, Images et Documents », « Systèmes Interactifs et Image »

Axe stratégique : « Pertinence Contenu-Interaction »

Cadre de coopération :

Date de début du stage :

Durée du contrat :

Contexte de l'étude:

Le bouclage de pertinence est une technique utilisée en apprentissage automatique afin de prendre en considération le retour utilisateur dans la construction du modèle d'apprentissage. Dans les systèmes de recherche d'images par le bouclage de pertinence a permis d'obtenir de très bons résultats. En pratique, au court de plusieurs itérations, l'utilisateur est invité à classifier les images qui lui sont proposées comme « bonnes » ou « mauvaises » jusqu'à atteindre un

certain niveau de satisfaction dans les performances du classifieur (on considère comme « acceptable » la sollicitation de 100 clicks auprès de l'utilisateur). Il est toutefois illusoire de considérer qu'un tel système soit acceptable par un utilisateur dans la vie de tous les jours (par exemple dans une utilisation mobile).

Description du sujet :

L'idée de ce stage est d'explorer la possibilité d'utiliser un système d'acquisition de la trajectoire oculaire (système Facelab) pour obtenir un retour de pertinence de l'utilisateur sans retour explicite. En particulier, on s'attend à ce que le retour visuel soit beaucoup plus rapide et permette donc de faire converger l'algorithme d'apprentissage beaucoup plus rapidement.

Prérequis et contraintes particulières :

Connaissances du C++ ou bon niveau en programmation

Références bibliographiques :

[Facelab] <http://www.seeingmachines.com/product/facelab/>

[Veltkamp, 2002] R. Veltkamp, "Content-based image retrieval system: A survey," University of Utrecht, Tech. Rep., 2002.

[F58] Fantz R. (1958). Pattern vision in young infants. The Psychological Record, vol. 8, 43–47, 1958.

[Wood,1998] M. Wood, N. Campbell, and B. Thomas, "Iterative refinement by relevance feedback in content-based digital image retrieval," in ACM Multimedia 98, Bristol, UK, September 1998, pp. 13–20.

[Huang,2001] T. Huang and X. Zhou, "Image retrieval with relevance feedback: From heuristic weight adjustment to optimal learning methods," in International Conference in Image Processing (ICIP'01), vol. 3, Thessaloniki, Greece, October 2001, pp. 2–5.

[Picard,2010] D. Picard, A. Revel, M. Cord, An application of swarm intelligence to distributed image retrieval, Information Sciences, 2010

Contacts – liens :

Email : arevel@univ-lr.fr

Lien vers le fichier de description : (PDF)

Présentation libre :

Cette partie ne sera visible QUE sur le fichier PDF dont le lien est présenté ci-dessus. Cette partie permet de présenter certains aspects non développés dans la partie WEB qui reste nécessairement globale.

Le bouclage de pertinence est une technique utilisée en apprentissage automatique afin de prendre en considération le retour utilisateur dans la construction du modèle d'apprentissage. Dans les systèmes de recherche d'images par le contenu (« Content-based Image Retrieval »,

CBIR), le bouclage de pertinence a permis d'obtenir les meilleurs résultats [Wood, 1998; Huang,2001]. En pratique, au court de plusieurs itérations, l'utilisateur est invité à classifier les images qui lui sont proposées comme « bonnes » ou « mauvaises » jusqu'à atteindre un certain niveau de satisfaction dans les performances du classifieur. Un raffinement de ces techniques de bouclage est l'apprentissage « actif » qui se propose de présenter à l'utilisateur les exemples les plus « pertinents » pour accélérer la convergence du système d'apprentissage. Dans les compétitions internationales permettant de comparer les performances des différents algorithmes dans un même cadre (VOC par exemple) on considère comme « acceptable » la sollicitation de 100 clicks auprès de l'utilisateur.

Il est toutefois illusoire de considérer qu'un tel système soit acceptable par un utilisateur dans la vie de tous les jours. L'idée de ce stage est d'explorer la possibilité d'utiliser un système d'acquisition de la trajectoire oculaire (système Facelab) pour obtenir un retour de pertinence de l'utilisateur sans retour explicite. En particulier, on s'attend à ce que le retour visuel soit beaucoup plus rapide et permette donc de faire converger l'algorithme d'apprentissage beaucoup plus rapidement.

Cependant, on doit valider l'idée que les annotations obtenues grâce au système de eye-tracking sont équivalents à ceux obtenus avec de « vrais » clicks. Certains travaux [HZ08] vont dans cette direction puisqu'ils montrent que le temps passé sur une image est une information fiable par rapport à son intérêt pour l'utilisateur. Pour pallier toute difficulté d'interprétation, nous proposons d'utiliser le paradigme de « préférence visuelle » (introduit by Fantz in 1958 [F58]). Celui-ci consiste à présenter 2 images côte à côte et déterminer laquelle est préférée par l'utilisateur. Il a été montré qu'une telle tâche mène à des saccades très rapides vers l'image préférée [KT06]. On propose par ailleurs de coupler ce système à un apprentissage actif déjà développé dans le cadre de la thèse de D.Picard [Picard, 2010] en sélectionnant les images à annoter telles qu'elles soient placées de chaque coté de la marge d'un classifieur SVM.

[HJZ+08] Hoi, S. C., Jin, R., Zhu, J., Lyu, M. R. (2008). Semi-supervised svm batch mode active learning for image retrieval, IEEE Conference on Computer Vision and Pattern Recognition 1–7, 2008.

[KT06] Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. Vision Research, 46, 1762–1776.

Argumentaire interne (non visibles vis-à-vis de l'extérieur) :

Enjeu du stage par rapport à la politique scientifique du laboratoire :

Le sujet proposé est à l'intersection des axes thématiques IDDC et S2I puisqu'il se propose d'utiliser des données d'interaction pour alimenter un système de recherche d'images par le contenu.

Positionnement du travail dans le cadre de la politique de capitalisation du laboratoire (1) :

Le stage s'appuie sur la réutilisation du matériel « facelab » peu ou plus utilisé au sein du laboratoire. L'idée est de pouvoir réutiliser ce matériel afin de mettre au point une démo mettant en valeur les compétences du laboratoire à la fois en indexation d'images et en interaction.

Mise à disposition de livrables :

Démo de recherche d'images par le contenu via eye-tracking.

(1) Rappel : les ressources communes du laboratoire ne seront désormais attribuées que si elles rentrent dans le cadre de la politique de capitalisation du laboratoire.