

Offre d'emploi : EXTRACTION D'INDICES SONORES PAR « WORD-SPOTTING », IDENTIFICATION DE CONCEPTS VISUELS ET RECHERCHE DE SEMANTIQUE DANS LES IMAGES

Type de contrat : CDD	Type de poste : Post-doctorant	Formation requise : doctorat
Durée du contrat : 12 mois	Rémunération : Catégorie A en fonction de l'expérience, sur la base de la grille de l'enseignement supérieur et de la recherche, ingénieur d'études	Prise de fonction : 02/01/2018

L'Université de Lyon (UdL), communauté d'Universités et d'établissements, regroupant 12 établissements d'enseignement supérieur de Lyon Saint Etienne recherche un post-doc (H / F)

DESCRIPTION DU POSTE

Le candidat (H / F) exercera ses missions au sein du laboratoire LIRIS (équipe Imagine) sous la responsabilité hiérarchique de Stéphane Bres (LIRIS) et Mylène Pardoën (ISH-CNRS). Le suivi sera également assuré par les spécialiste de la RI multimodale de l'équipe DRIM du LIRIS avec Léa Laporte (LIRIS).

Organisation du travail : 37h

Temps complet

Lieu de travail : Laboratoire LIRIS, INSA de Lyon, Bâtiment Jules Verne, 69621 Villeurbanne Cedex

Le LIRIS, regroupant 320 membres, est une unité mixte de recherche (UMR 5205) dont les tutelles sont le CNRS, l'INSA Lyon, l'Université Claude Bernard Lyon 1, l'Université Lumière Lyon 2 et l'Ecole Centrale de Lyon. Le champ scientifique de l'unité est l'Informatique et plus généralement les Sciences et Technologies de l'Information. Les activités scientifiques du LIRIS sont structurées en 6 pôles de compétences, dont le pôle vision intelligente et reconnaissance visuelle dont l'équipe Imagine est membre (19 permanents, 30 doctorants). Les thématiques principales sont les suivantes: Construction d'algorithmes visant la compréhension de données multimédia (images, vidéo, documents numérique, scènes 3D) et se déclinant en termes d'acquisition/reconstruction, d'indexation, de modélisation, de classification ou de reconnaissance automatique du contenu (objets, actions, concepts)

Encadrants principaux : Stéphane Bres (MCF au LIRIS-Imagine), Mylène Pardoën (chercheuse au CNRS, ISH-EVS, **co-encadrement** : Bruno Tellez (MCF au LIRIS-Imagine), Léa Laporte (MCF au LIRIS-Drim)

Missions principales : L'objet principal de la mission vise la mise en place d'un système de recherche multimodale, dynamique et interactif qui intègre différentes sources de données (données sources issues du patrimoine numérisée des bibliothèques et archives en Rhône-Alpes-Auvergne, partenaires du projet) et autorise l'association de **requêtes à la fois textuelles et de caractéristiques visuelles** pour permettre une recherche d'images par le texte, par le contenu visuel ou par les deux à la fois. Le système sera conçu selon une conception reposant sur un bouclage de pertinence avec l'utilisateur, permettant le raffinement de la recherche et de l'exploration des sources.

PROFIL RECHERCHE

Savoirs : indexation et classification d'images, reconnaissance de formes, techniques d'apprentissage, recherche d'information par le contenu, word-spotting

Savoir-faire : maîtrise des environnements de développement, en particulier de la programmation C++ et usage de la librairie OpenCV, réalisation de prototype fonctionnel, capacité à s'exprimer aisément en français à l'écrit et à l'oral (les corpus de textes sont en français)

Savoir être : aisance dans le cadre d'un travail de groupe pluridisciplinaire, ouverture et communication avec des spécialistes des sciences humaines et sociales, autonomie et capacité à prendre des initiatives

CANDIDATURES

Renseignements sur le poste : veronique.eglin@insa-lyon.fr, stephane.bres@insa-lyon.fr, lea.laporte@insa-lyon.fr

Envoi des candidatures : lettre de motivation + recommandation + CV exclusivement par e-mail avant le 15/10/2017 minuit à veronique.eglin@insa-lyon.fr et isabelle.diraimondo@universite-lyon.fr

DETAIL DU PROFIL

L'objectif du second travail postdoctoral est de proposer une approche de la recherche d'information reposant à partir de requêtes multi-modales texte-image. Habituellement découplées, les recherches d'information en mode « plein texte » et « image » permettent d'extraire des sources documentaires des fragments informationnels pertinents. Chacun des domaines de recherche (texte et image) dispose de son lot de techniques qui ont fait leur preuve dans des domaines où l'accès au contenu ne pouvait se faire que par une seule modalité d'interrogation. La recherche dans les images de textes manuscrits par exemple est habituellement résolue par les techniques alternatives à l'OCR (techniques dites de *word-spotting* dont l'équipe Imagine du LIRIS est spécialiste) lorsque la transcription est trop complexe à obtenir. On procède ainsi pour un corpus d'images à des approches de ce type basées sur des requêtes soit images, soit textuelles soit encore symboliques qui aboutissent à la détection de fragments d'images similaires à la requête. L'idée ici est de pousser plus loin la recherche par le contenu en créant une véritable interaction texte/image dès la formation de la requête qui sera multimodale. L'utilisation de séquences visuelles (images de mots, séquences de descripteurs) répondant aux ontologies sonores décrites pour le projet et modélisant les concepts sonores à différents niveaux : indices directs (ex : « le chien aboie ») à indirects (ex : « le forgeron travaillait dans la rue ») et les ambiances sonores associées (ex : « la rue était bruyante » ou encore « on entendait souffler le vent ») sera couplée à celle d'une recherche textuelle basée sur les métadonnées et les annotations textuelles des documents, ou au contenu plein texte s'il est disponible. Des travaux précurseurs allant dans le sens de requêtes multimodales ont d'ores et déjà ouvert la voie à une recherche d'information par annotation sémantique d'images naturelles, gravures ou photographies à base de descripteurs profonds et de descripteurs dérivés de métadonnées textuelles. C'est dans ce contexte que s'inscrit ce travail de post-doctorat. La partie principale du sujet sera réservée à la mise au point d'un système robuste de



recherche de mots (word-spotting) décrit ci-dessous. Une étude de faisabilité de l'introduction d'une identification de concepts sémantiques sera proposée dans la seconde partie.

Pour le développement de l'outil de fouille de texte en mode image, il sera question d'explorer le dictionnaire terminologique produit dans le projet. Cela permettra d'aboutir à une recherche contextualisée dans les images permettant de ne retenir que les fragments pertinents rassemblant des mots et des notions faisant sens dans un voisinage maîtrisé. Ce travail donnera lieu à la mise au point d'un outil interactif de prévisualisation des données permettant aux utilisateurs (tout d'abord les praticiens et les experts d'EVS) de naviguer et de consulter des pages en profondeur et d'aide au raffinement de la requête. Il sera initialement conçu avec l'aide des experts en archéologie du paysage sonore et les praticiens du projet. Ainsi, nous pouvons décrire un scénario possible du point de vue de l'utilisateur final. Le voici décrit. Une nouvelle série de documents sources est disponible. Après une expertise rapide de son contenu probable, qui permet d'orienter grossièrement le type d'indices sonores potentiellement présents, une analyse automatique est effectuée, par OCR sur de l'imprimé ou par « word spotting » si c'est du manuscrit. Les mots à rechercher sont fournis en amont par l'ontologie de termes sonores. Des indices directs, indirects ou d'ambiances sont recherchés. Selon les résultats de la recherche, on pourra essayer de désambigüiser les termes par des recherches associées ou poursuivre sur de nouveaux termes. La requête elle-même est construite par mosaïque de lettres pour reproduire les formes probables que présenteraient les mots recherchés. Cela nécessite un apprentissage préalable de la forme de l'écriture imprimée ou manuscrite, avec une estimation de sa variabilité. Cette phase est essentielle est faite pleinement partie du travail de recherche que nous envisageons. La façon de réaliser cet apprentissage ne peut s'envisager qu'en collaboration étroite avec la spécialiste en archéologie du paysage sonore, Mylène Pardoën aidée des praticiens experts des données. Les premiers résultats de la recherche automatique donneront des points d'accroches à l'expert pour un examen plus détaillé de document. Une nouvelle phase de recherche sera ensuite envisagée sur les indications et premières conclusions tirées de ces résultats. Des étapes de bouclage successif sont ensuite rendus indispensables pour avoir des résultats pertinents au final. Les compétences techniques sur la question du bouclage de pertinence et de la hiérarchisation (ranking) des réponses du système d'interrogation expertes seront celles de l'équipe Drim du LIRIS. Au cœur de ce dispositif seront étudiées les pratiques en archéologie du paysage sonore décrites par la spécialiste du projet afin de proposer une amélioration des interfaces et des dispositifs d'interaction en phase avec un usage spécialisé du système d'interrogation des sources. Le prototype d'interrogation final qui sera produit dans le cadre du postdoc sera enfin soumis à l'évaluation par les praticiens sur les lieux où sont conservées les sources documentaires.

