

AVIS DE PRESENTATION DE THESE EN SOUTENANCE POUR L'OBTENTION DU DIPLOME NATIONAL DE DOCTEUR

Monsieur Bao QUOC DANG

Présentera ses travaux intitulés :

« **Localisation d'information dans des très grands corpus de documents numérisés** »

Spécialité : **Informatique et Applications**

Le 6 avril 2018 à 15h00

Lieu :

**Université de La Rochelle
Maison des Sciences de l'Ingénieur
Amphi 100 (rez-de-chaussée)
Av. Becquerel
17000 LA ROCHELLE**

Composition du jury :

**M. CAO DE Tran
Mme EGLIN Véronique
M. FALQUET Gilles
M. MARINAI Simone
M. OGIER Jean-Marc
M. RUSINOL Marcal
Mme VINCENT Nicole**

**Maitre de conférences, HDR, Université de Can Tho
Professeur, INSA de Lyon
Professeur, Université de Genève
Professeur, Université de Florence
Professeur, Université de la Rochelle
Chercheur, Université Autonome de Barcelone
Professeur, Université Paris Descartes**

Résumé :

Ce travail vise à développer un cadre générique qui est capable de produire des applications de localisation d'informations à partir d'une caméra (webcam, smartphone) dans des très grands dépôts d'images de documents numérisés et hétérogènes via des descripteurs locaux. Le système développé dans cette thèse repose sur une requête image (acquise via une caméra) et le système est capable de renvoyer le document qui correspond le mieux à la requête, et également d'indiquer la zone visée par la caméra.

Ainsi, dans cette thèse, nous proposons d'abord un ensemble de descripteurs qui puissent être appliqués sur des contenus aux caractéristiques génériques (composés de textes et d'images) dédié aux systèmes de recherche et de localisation d'images de documents. Nos descripteurs proposés comprennent SRIF, PSRIF, DELTRIF et SSKSRIF qui sont construits à partir de l'organisation spatiale des points d'intérêts les plus proches autour d'un point-clé pivot. Tous ces points sont extraits à partir des centres de gravité des composantes connexes de l'images. A partir de ces points d'intérêts, des caractéristiques géométriques invariantes aux dégradations sont considérées pour construire nos descripteurs. SRIF et PSRIF sont calculés à partir d'un ensemble local des m points d'intérêts les plus proches autour d'un point d'intérêt pivot. Quant aux descripteurs DELTRIF et SSKSRIF, cette organisation spatiale est calculée via une triangulation de Delaunay formée à partir d'un ensemble de points d'intérêts extraits dans les images. Cette seconde version des descripteurs permet d'obtenir une description de forme locale sans paramètres. En outre, nous avons également étendu notre travail afin de le rendre compatible avec les descripteurs classiques de la littérature qui reposent sur l'utilisation de points d'intérêts dédiés, comme par exemple SURF ou SIFT, de sorte qu'ils puissent traiter la recherche et la localisation d'images de documents à contenu hétérogène.

La seconde contribution de cette thèse porte sur un système d'indexation de très grands volumes de données à partir d'un descripteur volumineux. Ces deux contraintes viennent peser lourd sur la mémoire du système d'indexation. En outre, une la très grande dimensionnalité des descripteurs peut amener à une réduction de la précision de l'indexation, réduction liée au problème de dimensionnalité. Nous proposons donc trois techniques d'indexation robustes, qui peuvent toutes être employées sans avoir besoin de stocker les descripteurs locaux dans la mémoire du système. Cela permet, in fine, d'économiser la mémoire et d'accélérer le temps de recherche de l'information, tout en s'abstrayant d'une validation de type distance. Pour cela, nous avons proposé trois méthodes s'appuyant sur des arbres de décisions : « randomized clustering tree indexing » qui hérite des propriétés des kd-tree, « kmean-tree » et les « random forest » afin de sélectionner de manière aléatoire les K dimensions qui permettent de combiner la plus grande variance expliquée pour chaque nœud de l'arbre. Nous avons également proposé une version pondérée de la distance euclidienne entre deux points de données, afin d'orienter celle-ci vers la dimension avec la variance la plus élevée. Enfin, pour améliorer la recherche de l'information, une fonction de hachage a été proposée en second lieu pour indexer et récupérer rapidement les contenus sans stocker les descripteurs dans la base de données. Nous avons également proposé une fonction de hachage étendue pour l'indexation de contenus hétérogènes provenant de plusieurs couches de l'image.

Comme troisième contribution de cette thèse, nous avons proposé une méthode simple et robuste pour calculer l'orientation des régions obtenues par le détecteur MSER, afin que celui-ci puisse être combiné avec des descripteurs dédiés (par exemple SIFT, SURF, ORB, etc.). Comme la plupart de ces descripteurs visent à capturer des informations de voisinage autour d'une région donnée, nous avons proposé un moyen d'étendre les régions MSER en augmentant le rayon de chaque région. Cette stratégie peut également être appliquée à d'autres régions détectées afin de rendre les descripteurs plus distinctifs. Là encore, nous avons utilisé une méthode d'indexation basée sur une fonction de hachage étendue afin d'indexer des contenus hétérogènes aux caractéristiques multiples (textes, graphiques, ...) à partir d'une décomposition des images en couches. Ce système est donc applicable pour les contenus uniformes (un seul type d'information), mais également pour plusieurs types d'entités à partir de plusieurs couches séparées.

Enfin, afin d'évaluer les performances de nos contributions, et en nous fondant sur l'absence d'ensemble de données publiquement disponibles pour la localisation d'information hétérogène dans des images capturées par une caméra, nous avons construit trois jeux de données qui sont disponibles pour la communauté scientifique. Cet ensemble de données contient des parties d'images de documents acquises via une caméra en tant que requête. Il est composé de trois types d'informations: du texte, des contenus graphiques et enfin des contenus hétérogènes.