

## INTRODUCTION

---

Avec ce numéro de la revue *Document Numérique* qui rassemble les contributions marquantes étendues issues de la dernière conférence sur l'écrit et le document (*Conférence Internationale Francophone sur l'Écrit et le Document*, CIFED 2014), preuve est encore faite de la richesse et de la diversité de cette thématique de recherche dans la communauté francophone. Depuis les traiteurs de signaux, transformant les pixels de leurs images pour faire émerger des blocs d'information porteurs de sens, aux spécialistes des méthodes d'apprentissage proposant des algorithmes prenant en compte la spécificité du domaine des documents, tous ont en commun la prise en compte d'un système complexe, où il est indispensable d'imaginer une solution complète. Cela suppose une maîtrise de bout en bout d'un ensemble pluri-technologique où chaque maillon est un défi pour assurer en temps contraint des performances optimales. La clé du succès repose sur une double exigence. Il s'agit d'une part de s'appuyer sur une démarche méthodologique très rigoureuse qui puise et nourrit des avancées fondamentales pour faire progresser les connaissances pour chaque nouvelle étude présentée, mais aussi, et la sélection de travaux présentés ici le montre bien, de proposer systématiquement des mises en situation expérimentales permettant d'évaluer de façon objective les propriétés des méthodes étudiées.

*De l'imprimé au multimodal, analyse et reconnaissance du document numérique* est composé de cinq contributions qui illustrent bien la diversité et les challenges introduits ici. En effet, les données disponibles correspondent soit à des documents imprimés, pour les deux premiers articles, ou bien à des données manuscrites, pour les deux articles suivants, enfin, le cinquième article traite de l'écriture manuscrite en-ligne enrichie de données issues de la parole.

La segmentation de pages de documents à la structure complexe, mixant des zones de textes, des traits séparateurs et des photos, est l'objet du papier *Approche hybride de segmentation de pages à base d'un descripteur de traits*. La méthode est robuste à l'inclinaison du texte qui est d'abord estimée et corrigée par une nouvelle méthode s'appuyant sur une transformée de Radon puis une transformée *Ridgelets*. Ensuite, un nouveau descripteur estimant la variation de la largeur du trait pour chaque composante connexe dans l'image permet d'extraire de l'image les candidats textes et traits.

Traitant également de documents imprimés, l'article *Classification de flux de documents évolutifs avec apprentissage de classes inconnues* propose un algorithme semi-supervisé actif pour la classification de flux continu de documents. Il s'agit de repérer les documents les plus informatifs à l'aide d'une mesure d'incertitude pour demander leur étiquette à un opérateur. Pour cela un modèle sous forme d'un graphe à topologie dynamique dont les nœuds sont des représentants de documents étiquetés est utilisé. La méthode proposée atteint des performances comparables aux méthodes supervisées.

De nombreux défis scientifiques sont associés à la recherche par le contenu dans les images de documents, en particulier lorsque ceux-ci contiennent de l'écriture manuscrite. L'article *Représentation des mots manuscrits par graphe pour la recherche par similarité* propose une nouvelle approche de recherche de mots reposant sur une structure de graphes intégrant des informations sur la topologie, la morphologie locale des mots ainsi que des informations contextuelles du voisinage de chaque point d'intérêt. Là encore, des expérimentations à grande échelle sont réalisées, notamment sur la base de George Washington et la base de registres de mariages de la cathédrale de Barcelone.

Avec l'article *Reconnaissance de mots manuscrits hors-vocabulaire en utilisant des ressources web*, les auteurs cherchent à améliorer la reconnaissance des mots hors-vocabulaires, et cela sans augmenter le dictionnaire statique. À cet effet, des ressources dynamiques provenant du web sont intégrées au système de reconnaissance conçu autour de réseaux récurrents de type *BLSTM*. Les performances, évaluées sur la base RIMES, montrent des améliorations par rapport à la reconnaissance avec dictionnaire statique.

Le dernier article de ce numéro, *Approches multimodales pour la reconnaissance d'expressions mathématiques*, propose une ouverture sur de la fusion de données hétérogènes pour aider à désambigüiser des situations de reconnaissance complexe. En disposant à la fois du signal de la parole et de l'écriture manuscrite correspondant à une même expression mathématique, une architecture permettant de combiner ces deux modalités est étudiée. La base bimodale HAMEX permet de montrer la complémentarité de ces deux sources d'information.

Nous ne doutons pas que tous ces travaux vont contribuer à asseoir la place privilégiée qu'occupe la recherche française au plan international. Rappelons que grâce aux travaux antérieurs de notre communauté de nombreuses sociétés françaises ont essaimé et sont aujourd'hui leaders sur la scène internationale dans le domaine du traitement de l'écrit et du document.

Je remercie très chaleureusement l'ensemble des contributeurs – auteurs et relecteurs – pour leur précieux travail.

CHRISTIAN VIARD-GAUDIN

Institut de Recherche en Communications  
et Cybernétique de Nantes, Université de Nantes

#### COMITÉ DE LECTURE

Eric Anquetil – IRISA, INSA de Rennes  
Sylvie Calabreto – LIRIS, INSA de Lyon  
Christophe Garcia – LIRIS, INSA de Lyon  
Laurence Likforman-Sulem – TSI, Telecom ParisTech  
Harold Mouchère – IRCCyN, Université de Nantes  
Jean-Yves Ramel – LI, Université de Tours  
Nicole Vincent – LIPADE, Université de Paris Descartes  
Laurent Wendling – LIPADE, Université de Paris Descartes