

**Colloque International Francophone
sur l'Ecrit et le Document
(CIFED'2002)**

Dans le cadre de CFD'2002

BIBLIOTHEQUE DU CERIST

CIFED'2002

Colloque International Francophone sur l'Ecrit et le Document

Publié sous la direction de

A. Belaïd

H. Ben Ghézala

Dans le cadre de la
Conférence Fédérative sur le Document
CFD'2002

Hammamet, Tunisie
21, 22, 23 Octobre 2002

Présentation du CIFED'2002

Depuis sa création en 1992 par le GRCE (Groupement de Recherche en Communication Ecrite), le CNED, devenu le CIFED, s'est donné pour objectif de rassembler toute la communauté francophone : universitaires, industriels et spécialistes, pour exposer et échanger des idées autour des thématiques fortes de l'écrit et du document et qui concernent la lecture et la rétro-conversion de documents papiers tels que les formulaires, l'indexation pour le catalogage numérique, le tri du courrier, la lecture de chèques.

De nouveaux besoins sont apparus récemment autour des sciences et technologies de l'information et de la communication (STIC). La miniaturisation des moyens de communication (mobiles, assistants personnels...) nécessite de développer des interfaces sans claviers. L'analyse du geste écrit et l'interaction avec le stylo électronique sont nécessaires.

L'explosion de l'Internet ces dernières années amène également à une réflexion sur les formats d'échange, l'accessibilité et l'indexation des documents complexes sur le Web, la distribution des traitements et la veille technologique.

Ce colloque permet de faire le point sur la problématique de l'écrit et du document dans ce contexte évolutif. On y présente les derniers développements qui découlent de recherches concernant les thèmes suivants : *représentation et codage de documents (avec compression) : incidence de la qualité sur l'exploitation, reconnaissance de documents : modèles de structures et formats d'échange, segmentation, stratégies d'interprétation, multilinguisme, traitement de formulaires : analyse, classification, lecture et correction, architecture distribuée et gestion du « workflow », combinaison de classifieurs et évaluation de systèmes, reconnaissance de symboles et de documents techniques, reconnaissance de l'écriture manuscrite : prise en compte du contexte par combinaison de niveaux d'analyse et de connaissances : morphologie, syntaxe et sémantique, vérification de signatures : liens avec la sécurisation des documents, interfaces de lecture et de correction, encre électronique, applications : e-book, bureau virtuel, bibliothèque numérique, vente par correspondance, veille technologique sur Internet, etc.*

Pour la première fois, dans le cadre de la Conférence Fédérative sur le Document (CFD), le CIFED sera organisé conjointement à différentes manifestations abordant d'autres aspects du traitement des documents (fouille de textes, usages des documents numériques, édition...). Hammamet (Tunisie) sera ainsi l'occasion de réunir la communauté scientifique impliquée dans le traitement du document dans son acception la plus large.

Nous avons choisi la tenue de cette conférence francophone en Tunisie pour impliquer davantage le bassin méditerranéen dans cette initiative, renforcer les actions francophones autour de l'ingénierie des langues, du document, du multilinguisme de l'information scientifique et culturelle, s'ouvrir sur le monde arabe dont la richesse de la production documentaire est très importante, et permettre de conforter plusieurs relations bilatérales déjà existantes entre plusieurs laboratoires maghrébins et les autres pays participants.

Organisation du CIFED'2002

Président du colloque

A. Belaïd
LORIA, Université Nancy2, France

Coprésident du colloque

H. Ben Ghézala
Centre de Calcul El Khawarizmi, Tunisie

Président du comité de programme

Rolf Ingold
IUFribourg, Suisse

Coprésident de comité de programme

Noureddine Ellouze
ENIT, Tunisie

Membres du comité de programme de CIFED'2002

D. Aboutajdine Univ. Mohamed V, Rabat, Maroc
A. M.Alimi Univ. de Sfax, Tunisie
A. Amin Univ. de New South Wales, Australie
L. Andriamampianina, FADES, Madagascar
J.P. Asselin AUF, Canada
D. Barba IRESTE, Nantes, France
M. Batouche Univ. de Constantine, Algérie
A. Ben Hamadou Univ. de Sfax, Tunisie
R. Benslimane Univ. Sidi M. Ben Abdellah, Fès, Maroc
F. Bortolozzi PUCPR, Brésil
J. Camillerap IRISA, INSA Rennes, France
M. Cheriet ETS, Montréal, Canada
B. Dorizzi INT Evry, France
N. Ellouze ENIT, Tunisie
A. El Yacoubi PUCPR, Brésil
H. Emptoz INSA Lyon, France
C. Faure ENST, France
W. Gdoura ISD, Tunisie
M. Gilloux SRTP, Nantes, France
A. Habacha ENSI, Tunisie
Y. B. Kwon, Chungang Univ., Seoul, Corée du Sud
F. Lebourgeois INSA Lyon, France
E. Lecolinet ENST Paris, France
Y. Lecourtier PSI, Univ. Rouen, France
Y. Le Cun ATT Bell, USA
H. Le Men IGN, Saint Mandé, France
L. Likforman-Sulem ENST, France
G. Lorette IRISA, Univ. Rennes 1, France
D. Mammass Univ. Ibn Zohr, Agadir, Maroc
M. Milgram Univ. Paris VI, France
R. Mullot Univ. Rochelle, L3i, France
L. Najman Océ-Industries SA, France
Ch. Olivier Univ. de Poitiers, France
Th. Paquet PSI, Univ. Rouen, France
R. Plamondon EPM, Montréal, Canada
R. Sabourin ETS, Montréal, Canada
M. Sellami Univ. Badji Mokhtar, Annaba, Algérie
B. Taconet Univ. du Havre, France
K. Tombre Loria-Emin, France
Ch. Viard-Gaudin IRESTE, Nantes, France
N. Vincent Univ. de Tours, France

Table des matières

Session 1 : Ecritures Arabes et Asiatiques

N. E. Ben Amara, Ecole Nationale d'Ingénieurs de Monastir, Tunisie «Problématique et orientations en reconnaissance de l'écriture arabe »	1
S. Snoussi Maddouri, A. Belaïd, Ch. Choisy, H. Amiri, LSTS, ENIT, Tunisie «Modèle perceptif neuronal à vision globale-locale pour la reconnaissance de mots manuscrits arabes ».....	11
S. Kanoun, A. Alimi, A. Ennaji, Y. Lecourtier, LPSI, France «Reconnaissance de mots arabes par approche affixale ».....	21
Y. B. Kwon, Dept of Computer Engineering, Chung-Ang University, Seoul «Tree Classifier for 2D Type Clustering H/V Strokes on Hangul Documents ».....	31

Session 2 : Structures Physiques

I. Leplumey, Ch. Queguiner, INSA-IRISA, France «Un graphe de voisinage base sur l'utilisation des distances discrètes ».....	41
M. Crucianu, C. Faure, F. Tappa, M. Venet, N. Vincent, LI/E3I - Université de Tours, France «Détection et usage des composantes typographiques et graphiques dans les pages de documents ».....	51
S. Ramdane, B. Taconet, A. Zahour, A. Faure, G.R.E.A.H, G.E.D, Université du Havre, France «Apprentissage non supervisé des modèles de Markov cachés: application au tri de formulaires ».....	61
V. Eglin, S. Bres Laboratoire de Reconnaissance de Formes et de Vision, France «Le texte et sa texture : analyse et interprétation de sa prégnance sur la page »...71	

Session 3 : Posters

P.R. Martins, J. Facon «Étude comparatives de techniques de multi-seuillage, Application aux chèques bancaires ».....	81
N. Azizi, T. Sari, L. Souici-Meslati, M. Sellami, Laboratoire LRI, Département, Informatique, Université Badji Mokhtar, Algérie «Reconnaissance de mots arabes manuscrits par combinaison floue de classifieurs neuronaux».....	89

B. Mignot, M. Wack, N. Cottin, LST, Univ. de Technologie de Belfort-Montbéliard « <i>Algorithme de secret partagé, études comparatives et version parallèle</i> ».....	97
A. Gagneux, H. Emptoz, INSA Lyon, France « <i>Le document Web, de la structure à la lisibilité</i> ».....	105
E. Trinh, F. Lebourgeois, H. Emptoz, INSA Lyon, France « <i>Un format de document ouvert et optimisé pour une exploitation collaborative</i> ».....	113
M. Delalandre, P. Héroux, S. Adam, E. Trupin, J-M. Ogier, LPSI-L3I, France « <i>Une approche statistico-structurale pour la reconnaissance de symboles exploitants une représentation XML des données</i> ».....	121
M. Pechwitz, S. Snoussi Maddouri, V. Märgner, N. Ellouze, H. Amiri, IFN/ENIT « <i>IFN/ENIT: Database of handwritten arabic words</i> ».....	129
S. Adam, R. Mullot, J.M. Ogier, J. Gardes, P. Loonis, L3I, LPSI, France Télécom « <i>Système d'interprétation à base de règles</i> ».....	137
P. Loonis, J-M Ogier, S. Adam, R. Mullo, L3I, LPSI, France « <i>Approche multi-sources pour la reconnaissance de formes sur des documents techniques</i> ».....	145
B. Taconet, A. Zahour, S. Ramdane, P. Mercy, Université du Havre, France « <i>Sur la probabilité a priori des classes, Cas des modèles probabilistes représentés</i> ».....	153
L. Souici-Meslati, L. Rahim, M. Zemehri, LRI, Université Badji Mokhtar, Algérie « <i>Système connexionniste à représentation locale pour la reconnaissance de montants littéraux arabes</i> ».....	161
So. Touj, Sa. Touj, N. Ben Amara, H. Amiri, ISET, ENIM, ENIT, Tunisie « <i>Reconnaissance hors ligne de caractères arabes isolés manuscrits</i> ».....	169
B. Taconet, A. Zahour, S. Ramdane, P. Mercy, Université du Havre, France « <i>Distance d'édition et modèle segmenté</i> ».....	177

Session 4 : Images et formats de documents

E. Smigiel, J. Gannard, F. Rousselot, Laboratoire d'informatique et d'intelligence artificielle, ENSAIS Strasbourg, France « <i>Le projet HyperCodex : documents en mode graphique augmentés de liens hypertextes</i> ».....	185
---	-----

F. Lebourgeois, I. Karoui, LRVF, INSA Lyon France «Compression de documents imprimés numérisés ».....	195
--	-----

B. Allier, H. Emptoz, LRVF, INSA Lyon France «Reconstruction d'images de caractères dégradés par contours actifs: étude exploratoire ».....	205
--	-----

C. Wolf, J.M. Jolion, LRVF, INSA Lyon France «Détection et extraction de texte de la vidéo ».....	215
--	-----

Session 5 : Structures Logiques

B. Couasnon, J. Camillerap, IRISA / INSA - Département Informatique, France «DMOS : une méthode générique de reconnaissance de documents : évaluation sur 60000 formulaires du XIXème siècle».....	225
--	-----

K. Hadjar, O. Hitz, L. Radey, R. Ingold, DIUF, Université de Fribourg, Suisse «2(CREM) Une méthode de reconnaissance interactive pour les documents à structure complexe ».....	235
---	-----

G. Bochnia, J. Facon, PUCPR, PPGIA, Brésil «Segmentation du bloc adresse : application aux enveloppes postales complexes ».....	245
---	-----

P. Vaillant, L. Likforman-Sulem, F. Yvon, ENST, LTCl, France «Exploitation d'informations spatiales et textuelles en analyse de documents : le cas des télécopies ».....	255
--	-----

Session 6 : Signatures et Indexation

S. Tabbone, L. Wendling, D. Girardeau-Montaut, LORIA, EMN, Nancy, France «Mesures de rectangularité et d'ellipticité à partir de la transformée de Radon ».....	265
---	-----

M. El Yassa, A. Lahyane, A. Chalifour, D. Mammass, F. Nouboud, LIMI, Unité de Recherche en TII, Faculté des Sciences, Maroc «Vérification statique de signatures manuscrites : revue de la littérature et proposition d'une approche basée sur la prétopologie et les réseaux de neurones».....	275
---	-----

A. Bensefia, L. Heutte, T. Paquet, A. Nosary, LPSI, France «Identification du scripteur par représentation graphèmes ».....	285
--	-----

M. Fuentes, D. Mostefa, J. Kharroubi, S. Garcia-Salicetti, B. Dorizzi, G. Chollet, INT dépt EPH, ENST, Lab CNRS-LTCl, France	
---	--

«Vérification de l'identité par fusion de données biométriques : signatures en ligne et parole ».....	295
---	-----

Session 7 : Reconnaissance en ligne

N. Ragot, E. Anquetil, IRISA, INSA de Rennes, France «Combinaison hiérarchique de systèmes d'inférence floue : application à la reconnaissance en ligne de chiffres manuscrits ».....	305
--	-----

E. Poisson, Ch. Viard-Gaudin, P.M. Lallican, Ecole polytechnique de l'université de Nantes, France «Combinaison de réseaux de neurones à convolution pour la reconnaissance de caractères manuscrits en ligne ».....	315
---	-----

A. L. Koerich, Y. Leydier, R. Sabourin, C. Y. Suen, Lab. d'Imagerie, de Vision et d'Intelligence Artificielle, Ecole de technologie supérieure, Canada «Système hybride de reconnaissance de mots manuscrits sur un grand vocabulaire utilisant des réseaux neuronaux et des modèles de Markov cachés ».....	325
---	-----

L. Oudot, L. Prevost, M. Milgram, Université Pierre & Marie Curie, France «Système de lecture multi contextuel pour la reconnaissance de textes manuscrits dynamiques ».....	335
---	-----

Session 8 : Manuscrit

Ch. Choisy, A. Belaïd, LORIA, France «Couplage d'une vision locale par HMM et globale par RN pour la reconnaissance de mots manuscrits ».....	345
--	-----

A. Nosary, L. Heutte, T. Paquet, LPSI, France «Reconnaissance de mots manuscrits par segmentation-reconnaissance : apports d'une reconnaissance lettres par niveau avec rejet ».....	355
---	-----

A. Nosary, T. Paquet, L. Heutte, LPSI, France «Reconnaissance de textes manuscrits par adaptation au scripteur ».....	365
--	-----

N. Gauthier, T. Artieres, LIP6, Université Paris 6 «Grammaires probabilistes appliquées à la segmentation de documents Manuscrits en ligne ».....	375
--	-----

Index aux auteurs	385
-------------------------	-----