

Annexe

DESCRIPTION DU PRODUIT

Cette annexe donne une description détaillée de Readiris.

Reconnaissance

technologie Connectionnist
extraction de caractéristiques omnifonte
routines d'auto-apprentissage basées sur un
réseau neuronal propriétaire
analyse contextuelle et linguistique
module d'apprentissage puissant

Support linguistique

55 langues (français, afrikaans, albanais, allemand, anglais américain, anglais britannique, basque, biélorusse, brésilien, breton, bulgare, catalan, corse, croate, danois, écossais (gaélique), espagnol, estonien, finnois, frison, grec, hongrois, indonésien, irlandais (gaélique), islandais, italien, lithuanien, luxembourgeois, macédonien, malais, néerlandais, norvégien, nynorsk, polonais, portuguais, roumain, russe, serbe, slovaque, slovénien, swahili, suédois, tagalog, tchèque, turc et ukrainien)
7 modes de langues mixtes (biélorusse-anglais, bulgare-anglais, grec-anglais, macédonien-



anglais, russe-anglais, serbe-anglais et ukrainien-anglais)

modules linguistiques utilisant des connaissances sur des syllabes et un lexique de 100.000 mots par langue

l'option logicielle "Module d'OCR asiatique" offre la reconnaissance du japonais et du chinois simplifié

afin d'afficher et d'éditer des documents asiatiques, une version asiatique, japonaise ou chinoise du système Windows est installée, ou un logiciel "d'émulation" spécialisé (tel que UnionWay AsianSuite ou TwinBridge AsianBridge) est utilisé sur une version occidentale de Windows

Jeux de caractères

tous les jeux de caractères américains et européens, en ce y compris les alphabets de l'Europe de l'Est, les langues cyrilliques ("russes") et baltiques, le grec, le japonais et le chinois simplifié

utilisation de jeux de caractères mixtes pour la reconnaissance de mots "occidentaux" (noms propres, noms de marques, etc.) dans les documents en cyrillique ("russe") et en grec ainsi que dans les documents asiatiques

mode numérique pour la reconnaissance de tableaux de chiffres

apprentissage possible pour tous les symboles spéciaux: symboles mathématiques ou scientifiques, dingbats, etc.

GUIDE DE L'UTILISATEUR

Qualité d'impression

tous types d'impression: impression, dactylographie, impression laser, impression à jet d'encre, impression matricielle en qualité "courrier" ("25 pin") ou "brouillon" ("9 pin")
tolérance de 7 degrés pour la courbe de ligne

Fontes

toutes les fontes à espacement proportionnel ou fixe
styles normal, gras, italique et souligné
détection et recréation du type de fontes (serif-sans serif, proportionnel-fixe, normal-condensé), du style (gras, italique, souligné) et de la taille des caractères du document original en mode "autoformatage"
tailles des caractères de 6 à 72 points (0,21 à 2,54 cm)
reconnaisances des "letrines"

Interface utilisateur

utilisation basée sur des barres d'icônes
assistant OCR guide l'utilisateur à travers le processus d'OCR
mode automatique qui permet l'OCR "en un clic"
vérification des mots ajustable par l'utilisateur "avec contexte dans l'image"
fonctionnement indépendant ou par appel à partir d'une application Windows grâce à la fonction "Connect"
jeu de commandes DDE disponible pour intégrer la fonction OCR



Prévisualisation

prévisualisation de l'image avec 5 niveaux de zoom et 3 commandes de rotation

analyse de page automatique très précise, tolérante par rapport aux courbes, fait la différence entre zones de texte, zones graphiques et zones de tableau

détection du texte blanc sur fond noir avec inversion automatique au moment de la reconnaissance et inversion de l'image de toutes les pages numérisées

fenêtrage ajustable manuellement des zones intéressantes en utilisant des fenêtres rectangulaires et polygonales

enregistrement de gabarits de fenêtres pour un emploi ultérieur

Sortie texte

tout traitement de texte et format de texte populaire, entre autres les formats Microsoft Word (DOC), RTF et HTML

sortie vers le presse-papiers Windows

exportation directe, automatique des résultats OCR vers le traitement de texte Microsoft Word, le tableur Microsoft Excel et l'éditeur Windows WordPad

détection des paragraphes qui assure la saisie en continu

"l'autoformatage" recrée une copie facsimile du document original grâce au *formatage des mots* - type de fontes (serif-sans serif, proportionnel-fixe, normal-condensé), style (gras, italique, souligné) et taille des caractères -, *formatage des*

GUIDE DE L'UTILISATEUR

paragraphes - alignement (gauche, centré, droite et justifié) et indentation - ainsi que le *formatage des pages* - place des blocs de texte, colonnes, graphiques et tableaux

Sortie de tableaux

l'analyse de page détecte les tableaux encadrés et non encadrés

recréation des tableaux cellule par cellule dans les tableurs et traitements de texte (objets tableaux dans les fichiers Microsoft Word (DOC) et RTF)

exportation directe, automatique de tableaux vers le tableur Microsoft Excel et le traitement de texte Microsoft Word

Sortie graphique

graphiques en noir et blanc en formats TIFF (comprimé) et Paintbrush (PCX)

graphiques inclus dans un fichier RTF quand "l'autoformatage" est activé

Support des numériseurs

supporte une gamme étendue de numériseurs à plat, à alimentation, à main ainsi que les périphériques hybrides ou "tout en un"

compatible avec le standard de numérisation Twain

supporte les plate-formes de numérisation Visioneer PaperPort, Hewlett-Packard Document Assistant, Hewlett-Packard JetSuite, Microtek PageSuite et Primax PaperEase



Entrée graphique

images en noir et blanc en format TIFF (non comprimé, comprimé en packbits, Groupe 3 et Groupe 4), Paintbrush (PCX) et Windows bitmap (BMP)

permet la reconnaissance directe des fichiers télécopies

possibilité de “glisser-déposer” des images dans la zone image Readiris

Exigences du système

un PC 486 Intel avec 16 MO RAM et 30 MO disponible sur le disque dur. Un Pentium PC avec 32 MO RAM est recommandé sous Windows NT 4.0 et pour utiliser la fonction “Connect”

le système d’opération Windows 98, 95 ou Windows NT 4.0

afin d’afficher et d’éditer des documents asiatiques, une version asiatique, japonaise ou chinoise du système Windows est installée, ou un logiciel “d’émulation” spécialisé (tel que UnionWay AsianSuite ou TwinBridge AsianBridge) est utilisé sur une version occidentale de Windows

version MacOS disponible

Versions linguistiques

français, allemand, anglais, espagnol, grec, italien, néerlandais, polonais, russe et tchèque
consultez le site web d’I.R.I.S. pour connaître la disponibilité d’autres version localisées