

ScanSoft®



OmniPage[®] SE

Manuel de l'utilisateur

D R O I T S D ' A U T E U R

Copyright © 2002 ScanSoft, Inc. Tous droits réservés. Aucune partie de cette publication ne peut être transmise, transcrite, reproduite ou stockée dans un système de récupération de données, ni traduite dans une langue ou un langage informatique sous quelque forme ou par quelque moyen que ce soit (mécanique, électronique, magnétique, optique, chimique, manuel, etc.) sans l'accord préalable par écrit du département juridique de ScanSoft, Inc. (Legal Department at ScanSoft, Inc., 9 Centennial Drive, Peabody, Massachusetts 01960, U.S.A.).

Le logiciel présenté dans ce manuel est fourni sous licence et son utilisation et sa copie sont soumises aux termes de la licence.

REMARQUE IMPORTANTE

ScanSoft, Inc. fournit cette publication « en l'état », sans aucune garantie d'aucune sorte, expresse ou implicite, y compris de manière non limitative, toute garantie implicite de qualité commerciale ou d'adéquation à un usage particulier. Certaines juridictions ne reconnaissant pas les limitations de garantie expresse ou implicite dans certaines transactions, cette clause peut ne pas s'appliquer à votre cas. ScanSoft se réserve le droit de réviser et/ou modifier cette publication périodiquement sans obligation de notification de telles révisions ou modifications.

MARQUES COMMERCIALES

ScanSoft, OmniPage, OmniPage Pro, OmniPage SE, PaperPort, Papis, True Page, Direct OCR, AutoOCR et *OCR Proofreader* sont des marques déposées ou des marques commerciales de ScanSoft, Inc. enregistrées aux États-Unis et/ou dans d'autres pays.

Tous les autres noms de produit cités peuvent être des marques de leurs propriétaires respectifs et sont reconnus comme tels.

ScanSoft Inc.
9 Centennial Drive
Peabody, MA 01960
U.S.A.

Numéro de référence : 58-28001-01A



T A B L E D E S M A T I È R E S

PRÉSENTATION GÉNÉRALE	VII
Utilisation du manuel	viii
Consultation de l'aide en ligne	ix
Aide HTML en ligne	ix
Aide contextuelle	ix
Notes techniques	x
Glossaire	x
OmniPage SE	x
1 INSTALLATION ET CONFIGURATION	11
Environnement requis	12
Installer OmniPage SE	13
Configurer un scanner pour utilisation avec OmniPage SE	14
Lancer OmniPage SE	16
Enregistrer OmniPage SE	17
Nouvelles fonctions d'OmniPage Pro 11	18
OmniPage SE et OmniPage Pro 11	19
2 INTRODUCTION	21
Reconnaissance optique de caractères (OCR)	22
Technologie OCR d'OmniPage SE	22
Définition d'un document dans OmniPage SE	23
Principales étapes du traitement d'un document	23
Bureau OmniPage SE	24

	Barre d'outils Standard	25
	Barre des menus	26
	Barre d'outils Image	26
	Barre d'outils Formatage	27
	Boîte à outils OmniPage	27
	Gérer des documents	28
	Fenêtre Vignettes	29
	Fenêtre Détails	29
	Personnaliser des colonnes dans la fenêtre Détails	31
	Supprimer des pages d'un document	31
	Imprimer un document	32
	Fermer un document	32
	Documents OmniPage	32
	Avantages du format OPD	33
	Enregistrer au format OPD	33
	Réglages	34
3	DIDACTICIEL : TRAITEMENT DES DOCUMENTS	37
	Guide de Démarrage Rapide	38
	Charger et reconnaître les fichiers images exemples	38
	Numérisation et reconnaissance d'un document d'une page	38
	Traiter des documents à l'aide de l'Assistant OCR	41
	Traiter des documents automatiquement	44
	Boutons de commande	45
	Traiter des documents manuellement	46
	Traitement combiné automatique/manuel	48
	Traitement à partir d'autres applications	49
	Configurer Direct OCR	49

Utiliser Direct OCR	49
Utiliser OmniPage SE avec PaperPort	51
Traiter des documents avec Planifier OCR	52
Définir la source des images de page	53
Importer des fichiers image	53
Importer des documents à partir d'un scanner	54
Numériser avec un chargeur ADF	56
Numériser des documents volumineux sans chargeur ADF	56
Spécifier la mise en page du document	57
Tracer des zones manuellement	58
Gérer les zones	59
Propriétés de zone	60
Gérer les grilles du tableau dans l'image	61
Utiliser des gabarits de zone	63
4 VÉRIFICATION ET ÉDITION	65
Vérifier les résultats de l'OCR	66
Comparer le texte reconnu à l'original	67
Dictionnaires utilisateur	68
IntelliTrain	69
Modes d'affichage de l'éditeur de texte	72
Éditer du texte et des images	73
Lire le texte à voix haute	74
Ébauche de page	76

5	ENREGISTREMENT ET EXPORTATION	77
	Préparer les résultats de l'OCR pour l'exportation	78
	Enregistrer le document dans un fichier	79
	Enregistrer des images originales	79
	Enregistrer les résultats de l'OCR	80
	Enregistrer un document lors d'une session de travail	81
	Copier un document dans le Presse-papiers	82
	Joindre un document à un e-mail	83
6	INFORMATIONS TECHNIQUES	85
	Dépannage	85
	Solutions de premier recours	86
	Tester OmniPage SE	87
	Mémoire insuffisante	88
	Espace disque insuffisant	88
	Types de fichiers supportés	90
	Types de fichiers pour l'ouverture et l'enregistrement des images	90
	Types de fichiers pour l'enregistrement des résultats de l'OCR	91
	Enregistrer un document au format PDF	92
	Problèmes de reconnaissance	93
	Reconnaissance de texte incorrecte	93
	Reconnaissance de fax incorrecte	94
	Problèmes de système ou de performances pendant l'OCR	95
	Désinstaller le logiciel	96
	INDEX	97



Présentation générale

Nous vous remercions de la confiance que vous portez à OmniPage SE™ (Special Edition). La documentation citée ci-dessous est destinée à vous présenter le programme et à faciliter son utilisation.

Manuel d'utilisation

Le présent manuel fournit des informations relatives aux principes d'utilisation d'OmniPage SE : procédures d'installation et de configuration, description des commandes et zones de travail disponibles, instructions de réalisation de différentes tâches, méthodes de personnalisation et de gestion de traitement et diverses informations techniques. Ce guide est disponible au format PDF. Ce format vous permet d'utiliser les liens hypertexte des références croisées ainsi que les autres outils de navigation offerts par votre programme de consultation de fichiers PDF.

Aide en ligne

L'aide en ligne d'OmniPage SE contient des informations sur les fonctions, les réglages et les procédures relatives au logiciel. Au format HTML, cette aide a été conçue pour garantir une consultation aisée et rapide. Son module d'aide contextuelle très complet assure une assistance suffisante pour vous permettre d'accomplir sans délai les opérations requises. Pour plus de détails, reportez-vous à la section *Consultation de l'aide en ligne*.

Fichier ReadMe/LisezMoi

Ce fichier au format HTML comporte les informations de dernière minute concernant le logiciel. Il est vivement recommandé de le consulter avant d'utiliser OmniPage SE. Pour ouvrir le fichier ReadMe/LisezMoi, vous pouvez, lors de l'installation d'OmniPage SE, sélectionner l'option ReadMe/LisezMoi, ou, à tout moment, choisir la commande correspondante dans le menu « Aide ».

Informations de numérisation et diverses

Le site Web de ScanSoft (www.scansoft.fr) fournit des informations actualisées au sujet du programme. Le « Guide du scanner » (actuellement en anglais) contient des données mises à jour régulièrement et relatives aux scanners supportés et à d'autres questions y afférentes. Pour accéder au site Web de ScanSoft, vous pouvez, lors de l'installation d'OmniPage SE, sélectionner l'option d'accès au site Web, ou, à tout moment, choisir la commande correspondante dans le menu « Aide ».

UTILISATION DU MANUEL

Ce manuel présuppose que vous maîtrisez les principes de travail dans l'environnement Microsoft Windows. Pour tous détails sur cet environnement (utilisation des boîtes de dialogue, des menus déroulants et contextuels, des barres de défilement, de la fonction Glisser-déplacer, etc.), reportez-vous à la documentation Windows appropriée.

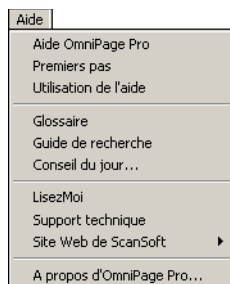
Vous devriez également connaître le fonctionnement de votre scanner et de son logiciel. Avant de procéder à la configuration du scanner avec OmniPage SE, assurez-vous qu'il est installé et qu'il fonctionne correctement. En cas de besoin, reportez-vous à la documentation du scanner.

Ce manuel utilise les conventions suivantes :

Texte en gras	Introduit les nouveaux termes et signale les titres des différentes sections de chaque chapitre.
<i>Texte en italiques</i>	Met en évidence les références aux différentes sections de ce manuel (sauf mention contraire, toute section citée se trouve dans le même chapitre que la référence). Indique les boutons principaux utilisés lors du traitement automatique : <i>Démarrer</i> , <i>Arrêter</i> , <i>Terminer</i> , <i>Autres pages</i> .
Police sans empattement	Indique les noms de fichier, tels que exemple.tif
Remarque:	Introduit des informations supplémentaires.
Astuce:	Introduit des conseils sur l'accomplissement de tâches spécifiques à l'aide des fonctions du logiciel.

CONSULTATION DE L'AIDE EN LIGNE

L'aide en ligne d'OmniPage SE contient des informations sur les fonctions, les réglages et les procédures relatives au programme. Le système d'aide devient disponible dès l'installation du logiciel.



Aide HTML en ligne

Pour accéder à l'aide en ligne d'OmniPage SE, choisissez la commande « Aide OmniPage SE » dans le menu d'aide. Vous pourrez obtenir les informations requises en consultant des rubriques de la table des matières, en sélectionnant des mots-clés dans l'index alphabétique ou en effectuant des recherches sur l'intégralité du texte d'aide. Les autres commandes du menu « Aide » vous permettront d'accéder à de nombreux sujets et pages Web utiles.



Lors de l'utilisation d'OmniPage SE, vous pouvez consulter la rubrique d'aide relative à la boîte de dialogue, la zone d'écran ou le message affiché en appuyant sur la touche F1.

Aide contextuelle

Pour afficher des informations succinctes sur un élément d'interface d'OmniPage SE (commande de menu, bouton de barre d'outils, groupe d'options, boîte de dialogue, etc.), vous pouvez procéder des manières suivantes :

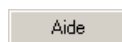


Pour afficher la rubrique d'aide relative à un élément du bureau en dehors d'une boîte de dialogue ou d'une fenêtre de message, cliquez sur le bouton d'aide dans la barre d'outils standard afin d'afficher l'icône d'aide (flèche accompagnée d'un point d'interrogation), puis cliquez sur cette icône et sur l'élément requis.

Vous pouvez également afficher cette icône via la séquence Maj.+ F1.



Pour afficher une mini-fenêtre décrivant un élément de boîte de dialogue, cliquez sur le point d'interrogation situé dans l'angle supérieur droit de la boîte en question puis sur l'élément requis.



Certaines boîtes de dialogue ou fenêtres de message comportent un bouton Aide (icône ou libellé). Pour accéder aux informations sur leurs éléments, cliquez sur ce bouton.

Pour supprimer une fenêtre d'aide contextuelle, cliquez n'importe où sur le bureau.

Notes techniques

Le site Web de ScanSoft (www.scansoft.fr) fournit des informations techniques (Tech Notes) au sujet d'OmniPage SE et propose de l'assistance pour l'installation et le dépannage du logiciel.

Glossaire

Un glossaire complet d'OmniPage SE (disposant d'un index alphabétique et d'une table des matières) est inclus dans l'aide en ligne du programme. Consultez ce glossaire pour trouver la signification d'un terme technique utilisé dans le manuel ou le programme.

OMNIPAGE SE



La version d'OmniPage dont vous disposez est une version spéciale de l'application OmniPage Pro™. Cette version a spécialement été conçue pour certains fabricants de scanners et ne comprend pas l'ensemble des fonctionnalités d'OmniPage Pro 11. Ce guide ainsi que l'aide en ligne offrent une description de l'ensemble des fonctionnalités du produit. L'icône SE, placée en marge de certaines sections, vous permet d'identifier rapidement les différences entre chacune des versions.

Si vous êtes intéressé par les fonctionnalités supplémentaires offertes par la version professionnelle du produit, consultez notre site Web pour effectuer une mise à niveau vers OmniPage Pro 11.



1 Installation et configuration

Ce chapitre contient les informations nécessaires à l'installation et au lancement d'OmniPage SE, soit les sections suivantes :

- ◆ Environnement requis
- ◆ Installer OmniPage SE
- ◆ Configurer un scanner pour utilisation avec OmniPage SE
- ◆ Lancer OmniPage SE
- ◆ Enregistrer OmniPage SE
- ◆ Nouvelles fonctions d'OmniPage Pro 11
- ◆ OmniPage SE et OmniPage Pro 11

ENVIRONNEMENT REQUIS

Pour installer et utiliser OmniPage SE, vous devez prévoir l'environnement minimum suivant :

- ◆ Processeur Pentium ou supérieur
- ◆ Microsoft Windows 95, Windows 98, Windows Me, Windows 2000 ou Windows NT 4.0
- ◆ 32 Mo de mémoire RAM (64 Mo recommandés)
- ◆ 75 Mo d'espace disque disponible destiné aux fichiers d'application, plus 10 Mo d'espace disponible lors de l'installation
- ◆ 9 Mo pour Microsoft Installer (MSI) et 44 Mo pour Internet Explorer, le cas échéant. (Ces logiciels font partie de l'installation de Windows 98, Windows Me et Windows 2000.)
- ◆ Affichage SVGA avec 256 couleurs et résolution de 800 x 600 pixels
- ◆ Souris (ou équivalent) compatible Windows
- ◆ Lecteur de CD-ROM (pour l'installation)
- ◆ Scanner compatible (si vous prévoyez de numériser des documents). Pour consulter la liste des scanners supportés, consultez le « Guide du scanner » (actuellement en anglais) disponible sur le site Web de ScanSoft.

Remarque : Vous obtiendrez de meilleures performances avec un processeur plus puissant, et des valeurs supérieures de RAM et d'espace disque.

INSTALLER OMNIPAGE SE

Pour vous guider dans l'installation, OmniPage SE affiche des instructions sur chaque écran.

Avant de lancer l'installation d'OmniPage SE, respectez les points suivants :

- ◆ Vérifiez que le scanner est compatible avec votre système, qu'il est connecté au PC et sous tension.
- ◆ Fermez les applications actives, en particulier les logiciels antivirus.
- ◆ Si vous installez OmniPage SE sur un système Windows 2000 ou Windows NT, ouvrez une session avec privilèges d'administrateur.
- ◆ Si une version antérieure d'OmniPage est déjà installée sur votre ordinateur, le programme d'installation vous demandera l'autorisation de désinstaller le logiciel existant.

▼ Pour installer OmniPage SE :

1. Insérez le CD-ROM OmniPage SE dans le lecteur approprié. En principe, le programme d'installation démarre automatiquement. Si ce n'est pas le cas, ouvrez l'Explorateur Windows, cliquez sur l'icône du lecteur de CD-ROM, puis cliquez deux fois sur le fichier de programme AUTORUN.EXE.
2. Choisissez la langue d'installation souhaitée. Ce choix détermine également la langue de l'interface du programme (commandes de menu, boîtes de dialogue, fenêtres de message, etc.). Vous pouvez modifier la langue d'interface ultérieurement, mais la sélection effectuée lors de l'installation détermine la version du module Texte-parole installée avec le programme. Veuillez noter que le module Texte-parole n'est pas disponible dans OmniPage SE
3. Pour installer le logiciel, suivez les instructions affichées sur les écrans successifs. Tous les fichiers nécessaires à la numérisation sont copiés automatiquement lors de l'installation.



Remarque : Certains problèmes affectant OmniPage SE peuvent être résolus par l'installation puis la désinstallation du logiciel. Reportez-vous à la section *Désinstaller le logiciel* du chapitre 6.



Remarque : Dans OmniPage Pro 11, le module Texte-parole est disponible en français, anglais (britannique et américain), allemand, italien, portugais et espagnol. Ce module n'est pas disponible dans OmniPage SE. Reportez-vous également à la section *Lire le texte à voix haute* du chapitre 4.

CONFIGURER UN SCANNER POUR UTILISATION AVEC OMNIPAGE SE

Tous les fichiers nécessaires à la configuration et au support du scanner sont copiés automatiquement lors de l'installation du programme. Avant de numériser des documents à l'aide d'OmniPage SE, vous devrez vérifier que votre scanner est installé correctement et tester son fonctionnement.

L'installation et la configuration du scanner s'effectuent via l'Assistant de numérisation. Vous pouvez accéder à ce dernier via la procédure décrite ci-dessous. La fenêtre de l'Assistant de numérisation apparaît également lors de la première numérisation d'un document à partir d'OmniPage SE.

Pour configurer votre scanner pour l'utilisation avec OmniPage SE à l'aide de l'Assistant de numérisation, procédez comme suit :

- ◆ Choisissez Démarrer ► Programmes ► ScanSoft OmniPage SE ► Assistant de numérisation, *ou*
Cliquez sur le bouton Configuration de l'onglet Scanner dans la boîte de dialogue Options, *ou*
Choisissez la commande de numérisation dans la liste déroulante Importer la page dans la boîte à outils OmniPage.
- ◆ Choisissez Sélectionner la source de numérisation, puis cliquez sur Suivant.
- ◆ Cliquez sur le pilote TWAIN de votre scanner pour le sélectionner, puis cliquez sur Suivant.
- ◆ Choisissez Oui pour tester la configuration de votre scanner, puis cliquez sur Suivant.
- ◆ L'Assistant teste la connexion entre l'ordinateur et le scanner. Cliquez sur Suivant.
- ◆ Placez une page de test dans le scanner.

- ◆ L'assistant se prépare à réaliser une numérisation standard à l'aide du logiciel fourni avec le scanner. Cliquez sur Suivant.
- ◆ L'interface native de votre scanner apparaît. Pour lancer un test de numérisation, cliquez sur Numériser.
- ◆ Si nécessaire, cliquez sur Image inversée... ou Image manquante... et effectuez les sélections appropriées.
- ◆ Une fois l'image affichée correctement dans la fenêtre, cliquez sur Suivant.
- ◆ Sélectionnez l'option la plus appropriée à votre scanner et cliquez sur Suivant.
- ◆ Cliquez sur Suivant pour passer à la sélection de la taille de page.
- ◆ Les tailles de page supportées par le scanner et détectées par l'Assistant de numérisation apparaissent. Pour apporter des modifications à la liste affichée, cliquez sur Évolué, effectuez les changements requis et cliquez sur Suivant.
- ◆ Placez dans le scanner une page comportant uniquement du texte (sans images). Cliquez sur Suivant pour lancer la numérisation en noir et blanc.
- ◆ Si nécessaire, cliquez sur Image inversée... ou Image manquante... et effectuez les sélections appropriées.
- ◆ Une fois l'image affichée correctement dans la fenêtre, cliquez sur Suivant.
- ◆ Si vous disposez d'un scanner couleur, placez une photographie ou une page couleur dans le scanner. Cliquez sur Suivant pour lancer la numérisation en couleur. Si nécessaire, cliquez sur Image inversée... ou Image manquante... et effectuez les sélections appropriées. Une fois l'image affichée correctement dans la fenêtre, cliquez sur Suivant. Si votre scanner ne gère pas la couleur, omettez cette étape.
- ◆ Placez une photographie ou une page comportant une image dans le scanner. Cliquez sur Suivant pour lancer la numérisation en niveaux de gris. Si nécessaire, cliquez sur Image inversée... ou Image manquante... et effectuez les sélections appropriées. Une fois l'image affichée correctement dans la fenêtre, cliquez sur Suivant.
- ◆ Votre scanner est désormais configuré pour l'utilisation avec OmniPage SE. Cliquez sur Terminer.

Pour modifier ultérieurement les paramètres de votre scanner, pour configurer un scanner différent ou pour tester et dépanner un scanner installé, rouvrez l'Assistant de numérisation selon une des méthodes suivantes :

- ◆ Démarrer > Programmes ► ScanSoft OmniPage SE ► Assistant de numérisation, ou
- ◆ Démarrer > Programmes ► ScanSoft OmniPage SE ► OmniPage SE ► menu Outils > Options ► Scanner... ► bouton Configuration.

Remarque : Pour tester et dépanner un scanner fonctionnant incorrectement, suivez la procédure ci-dessus en sélectionnant *Tester et configurer la source de numérisation actuelle* au début du processus.

LANCER OMNIPAGE SE

Pour lancer OmniPage SE, effectuez l'une des opérations suivantes :

- ◆ Dans la barre des tâches Windows, cliquez sur Démarrer et choisissez Programmes ► ScanSoft OmniPage SE ► OmniPage SE.
- ◆ Cliquez deux fois sur l'icône OmniPage SE située dans le dossier d'installation d'OmniPage SE ou sur le bureau Windows (si vous l'y avez placée).
- ◆ Cliquez deux fois sur l'icône ou le nom de fichier d'un document OmniPage (OPD) pour le charger. Pour plus de détails, reportez-vous à la section *Documents OmniPage* du chapitre 2.

Lorsque vous ouvrez l'application, l'écran d'accueil puis le bureau OmniPage SE s'affichent. Pour plus d'informations sur le bureau OmniPage SE, consultez le chapitre 2.

Vous pouvez également exécuter le programme avec une interface limitée :

- ◆ Ouvrez le programme de planification de reconnaissance. Dans la barre des tâches Windows, cliquez sur Démarrer et choisissez Programmes ► ScanSoft OmniPage SE ► Planifier OCR. Pour plus de détails, reportez-vous à la section *Traiter des documents avec Planifier OCR* du chapitre 3.
- ◆ Dans une application déclarée dans Direct OCR, cliquez sur Acquérir texte du menu Fichier. Pour plus de détails, reportez-vous à la section *Configurer Direct OCR* du chapitre 3.

- ◆ Cliquez avec le bouton droit sur une icône ou un nom de fichier image pour afficher son menu contextuel. Choisissez la sous-option requise de la commande « Convertir en » pour définir une cible.
- ◆ Utilisez OmniPage SE pour bénéficier des services OCR dans les produits de gestion de documents PaperPort® ou Pagis® de ScanSoft. Voir *Utiliser OmniPage SE avec PaperPort* du chapitre 3.

ENREGISTRER OMNIPAGE SE

À l'issue de l'installation, le programme lance l'Assistant d'enregistrement de ScanSoft. L'enregistrement s'effectue via un formulaire électronique facile à remplir (moins de cinq minutes).

Lorsque vous avez terminé de remplir le formulaire et que vous cliquez sur Envoyer, le programme tente de se connecter à Internet pour effectuer l'enregistrement en ligne sur-le-champ.

Si vous n'enregistrez pas votre copie du logiciel pendant son installation, le programme vous invitera à intervalles réguliers à accomplir cette formalité. Pour procéder à un enregistrement en ligne, rendez-vous sur le site Web de ScanSoft : www.scansoft.fr. Cliquez sur Support, puis choisissez Enregistrer dans la colonne de gauche.

Pour des informations sur l'utilisation des données d'enregistrement, consultez la clause de confidentialité de ScanSoft (Privacy Policy).

NOUVELLES FONCTIONS D'OMNIPAGE PRO 11

La famille de produits OmniPage® compte désormais OmniPage Pro 11 et OmniPage SE. Cette section présente les améliorations apportées à OmniPage Pro 11. Comme l'indique la section suivante, certaines de ces fonctionnalités sont également comprises dans OmniPage SE.

Nouvelles fonctionnalités de OmniPage Pro 11 par rapport à OmniPage Pro 10 :

- ◆ Meilleure précision – Son moteur de reconnaissance révisé confère à la version 11 un niveau de précision supérieur à tout autre produit OmniPage.
- ◆ Mise en page améliorée – OmniPage Pro 11 permet désormais de conserver le formatage de l'original, même si celui-ci contient des tableaux sans grille, des en-têtes, des bas de pages ou des lettrines.
- ◆ Vérification d'OCR plus efficace – La nouvelle fonction IntelliTrain exploite automatiquement les corrections apportées aux documents précédents pour assurer une meilleure reconnaissance du document traité.
- ◆ Compatibilité avec le format PDF – Vous pouvez désormais importer des fichiers PDF (y compris en lecture seule) et les convertir au format de vos applications favorites (Word, Excel, etc.). Vous pouvez également créer des fichiers PDF à partir de documents imprimés ou de fichiers image.
- ◆ Génération optimisée des pages HTML – Grâce à la nouvelle option de génération de documents HTML affichés en mode réel (correspondant à l'impression/la publication), vos pages HTML conservent le formatage du texte, les graphiques et les fonds du document original.
- ◆ Support linguistique étendu – OmniPage Pro 11 supporte maintenant plus de 100 langues (y compris les alphabets grec et cyrillique).
- ◆ Vue Détails – Cette fenêtre offre des possibilités renforcées de personnalisation d'informations sur des pages individuelles, facilitant ainsi la gestion des pages d'un document.

- ◆ Éditeur de texte – Cet éditeur innovant avec affichage en mode réel (correspondant à l'impression/la publication) permet de remanier les résultats de la reconnaissance à l'aide d'une large gamme d'outils d'édition, un puissant support de couleurs et quatre niveaux de formatage pour l'affichage et l'exportation.
- ◆ Meilleur traitement des documents de mauvaise qualité – Un nouveau module d'élimination des imperfections réduit sensiblement le nombre d'erreurs de reconnaissance susceptibles de se produire dans le cas des documents comportant des tâches ou des fonds en dégradé ou de couleur.

OMNIPAGE SE ET OMNIPAGE PRO 11



La liste suivante répertorie les fonctionnalités de la version professionnelle d'OmniPage Pro qui n'ont pas été ajoutées à OmniPage SE :

- ◆ Meilleur niveau de reconnaissance.
- ◆ Accès au module d'apprentissage IntelliTrain pour la reconnaissance de caractères.
- ◆ Capacité d'ouverture et de lecture du contenu de fichiers PDF.
- ◆ Capacité d'enregistrement des fichiers reconnus au format PDF
- ◆ Capacité d'ouverture de fichiers image TIFF FX.
- ◆ Gestion des formats d'image LZW TIFF et GIF en entrée et sortie.
- ◆ Prise en charge de la sortie WYSIWYG HTML 4.0.
- ◆ Nombre de langues reconnues passant de 50 à plus de 100.
- ◆ Accès au module Texte-parole, permettant la lecture à voix haute des textes reconnus.

Pour plus d'informations sur OmniPage Pro 11, ou pour obtenir une mise à niveau, consultez le site www.scansoft.com/opse.



2 Introduction

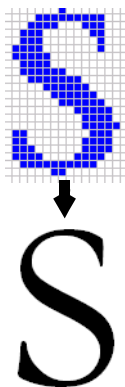
De nos jours, la plupart des activités professionnelles (correspondance, création de rapports, gestion des données, etc.) font appel à un ordinateur et pourtant, certaines sources d'information circulent toujours sous la forme imprimée et ne sont pas directement exploitables par la machine.

Ainsi, si vous voulez insérer un article de magazine au format papier dans un rapport en cours de création sur l'ordinateur, vous devez trouver un moyen de le faire, en évitant la saisie manuelle tout à fait inefficace.

OmniPage SE vous propose une solution sûre : la reconnaissance optique de caractères (OCR). Ce chapitre décrit l'emploi de la technologie OCR par OmniPage SE dans le but de transformer les informations contenues dans vos documents numérisés ou fichiers image en données électroniques éditables par vos applications habituelles.

Il traite les sujets suivants :

- ◆ Reconnaissance optique de caractères (OCR)
 - ◆ Définition d'un document dans OmniPage SE
 - ◆ Principales étapes du traitement d'un document
- ◆ Bureau OmniPage SE
- ◆ Gérer des documents
- ◆ Documents OmniPage (OPD)
- ◆ Réglages



RECONNAISSANCE OPTIQUE DE CARACTÈRES (OCR)

La reconnaissance optique de caractères (OCR) est une opération qui consiste à extraire du texte d'une image obtenue par numérisation d'un document papier ou par ouverture d'un fichier image électronique. Le texte d'une image n'est pas directement utilisable car ses caractères se composent de groupes de points (pixels).

Pendant la reconnaissance, OmniPage SE analyse l'image proposée, identifie les caractères qu'elle contient et définit les solutions permettant de les transformer en texte éditable. À l'issue de la reconnaissance, le texte peut être enregistré dans la plupart des applications (logiciels de traitement de texte ou de PAO, tableurs, etc.).

Technologie OCR d'OmniPage SE

Outre le texte, OmniPage SE peut reconnaître les éléments et caractéristiques suivants du document.

Éléments graphiques

Photos, logos, illustrations, etc.

Formatage du texte

Format des caractères, tel que type (Arial, Times New Roman...), corps (8, 12...) et style (**gras**, *italique*, souligné) d'une police. Format des paragraphes, tel que retraits, tabulations, marges et interlignes.

Mise en page

Agencement des colonnes, agencement des tableaux, emplacement des éléments graphiques et des en-têtes.

Les éléments et caractéristiques conservés par OmniPage SE dépendent des réglages sélectionnés pour chaque document. Pour plus de détails sur le choix des réglages, consultez la rubrique *Conseils de réglages* de l'aide en ligne OmniPage SE.

Remarque : OmniPage SE reconnaît uniquement les caractères imprimés, tels que les sorties imprimante ou le texte tapé à la machine. Vous pouvez toutefois lui demander de conserver un bloc de texte manuscrit (une signature, par exemple) en affectant celui-ci du type de zone Graphique.

Définition d'un document dans OmniPage SE

OmniPage SE gère les documents de façon individuelle. Lorsque vous importez la première image (à partir d'un scanner ou d'un fichier), le programme crée un document. Toutes les images importées par la suite sont ajoutées à ce document, jusqu'à son enregistrement et sa fermeture.

Dans OmniPage SE, un document comporte une ou plusieurs pages, chacune d'elles constituée par une image importée. Après la reconnaissance, le document comportera également le texte reconnu correspondant à ses images (affichable via l'Éditeur de texte) et des graphiques ou des tableaux éventuels. Pour plus de détails sur les différents éléments de l'écran OmniPage SE, reportez-vous à la section *Bureau OmniPage SE*.

Principales étapes du traitement d'un document

OmniPage SE propose deux modes de traitement des documents : automatique et manuel. (Consultez le chapitre 3 *Traiter des documents automatiquement* et *Traiter des documents manuellement*.) Les deux modes impliquent globalement les mêmes étapes principales :

1. Importer des images de document dans OmniPage SE

Vous pouvez numériser un document papier (avec ou sans le chargeur de feuilles automatique ADF) ou charger un ou plusieurs fichiers image. Les images importées apparaissent dans la fenêtre Vignettes et les informations relatives aux pages concernées s'affichent dans la fenêtre Détails (fenêtres faisant partie du gestionnaire de document). L'image de la page en cours figure dans la fenêtre Image originale.

2. Exécuter l'OCR pour obtenir du texte éditable

Pendant la reconnaissance, OmniPage SE trace des **zones** autour des parties de l'image à traiter, puis interprète le contenu textuel et graphique de chaque zone. Vous pouvez également créer des zones manuellement ou en appliquant un gabarit de zones prédéfini. À l'issue de la reconnaissance, vous pouvez détecter et corriger les erreurs éventuelles à l'aide du module de vérification de l'OCR et éditer le document dans l'Éditeur de texte.

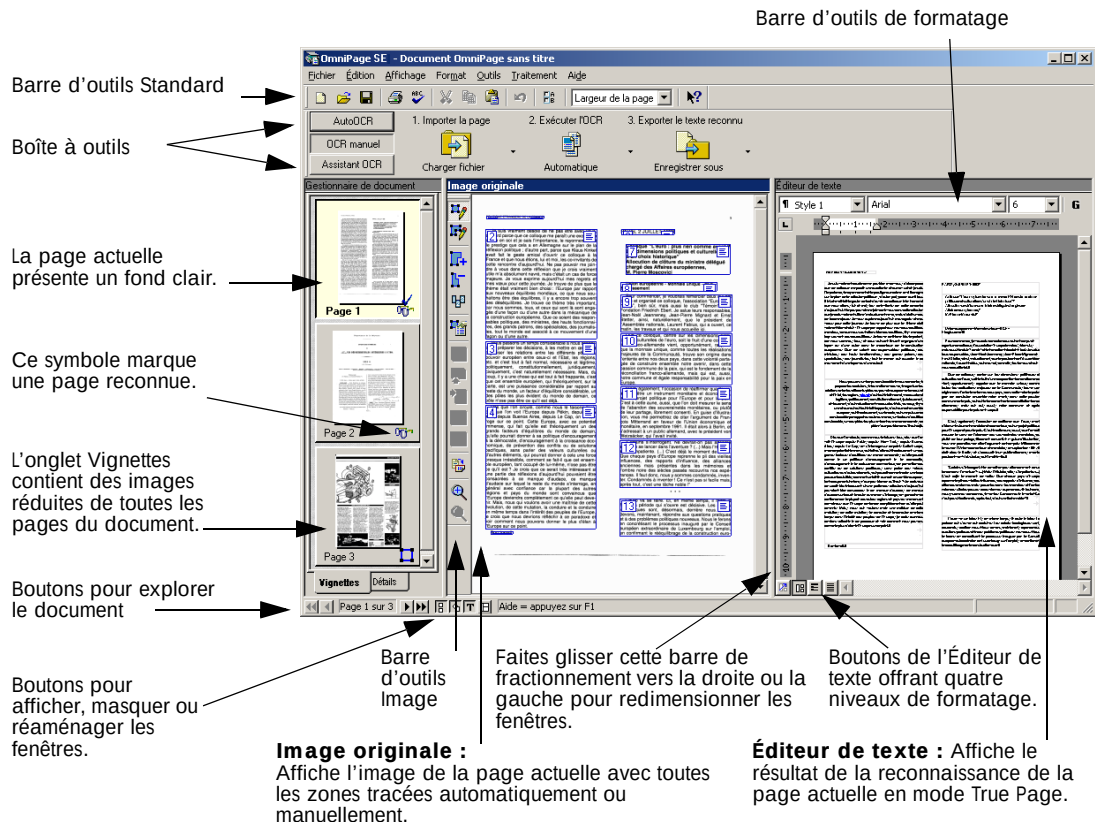
3. Exporter le document vers la destination de votre choix

Une fois le document reconnu, vous pouvez l'enregistrer dans un fichier (en spécifiant le nom et le type voulus), le copier dans le Presse-papiers ou l'envoyer comme pièce jointe d'un e-mail. Vous pouvez également l'enregistrer au format Document OmniPage (OPD) selon la procédure décrite ci-après. Un même document peut être enregistré plusieurs fois

avec différents paramètres d'emplacement, de type de fichier, de formatage, etc., selon vos besoins. Voir le chapitre 5.

BUREAU OMNIPage SE

Le bureau OmniPage SE dispose d'une barre de titre et de menus dans sa partie supérieure et d'une barre d'état dans sa partie inférieure. Il est composé de trois espaces de travail ou fenêtres : Gestionnaire de document, Image originale et Éditeur de texte. La fenêtre Gestionnaire de document présente deux onglets : Vignettes et Détails. La fenêtre Image originale est munie de la barre d'outils Image et la fenêtre Éditeur de texte de la barre d'outils de formatage.



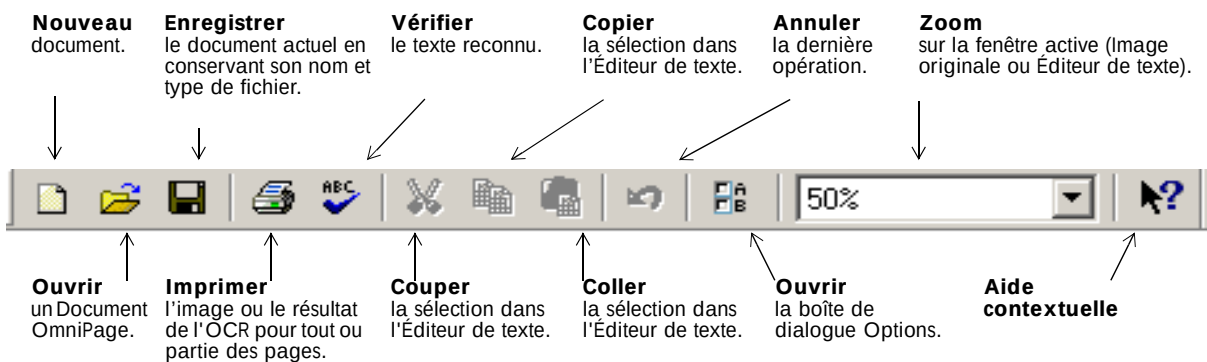
Remarque : Pour configurer l'affichage des fenêtres du bureau OmniPage (Gestionnaire de document, Image originale ou Éditeur de texte), cochez/décochez la ou les cases correspondantes dans le menu Affichage ou utilisez les boutons appropriés de la barre d'état.

La boîte à outils OmniPage permet à l'utilisateur de gérer le traitement. Cette boîte peut prendre trois aspects différents, en fonction du mode de traitement activé (boutons de gauche). L'illustration précédente correspond à l'OCR manuel. L'exemple utilisé s'appuie sur un document de trois pages. La page 1 représente la page actuelle, reconnue et vérifiée ; la page 2 a été reconnue mais non vérifiée ; la page 3 a été importée et divisée manuellement en zones, mais n'a pas été reconnue. L'icône figurant dans l'angle inférieur droit de chaque vignette indique l'état de la page.

Les boutons de la barre d'état vous permettent d'afficher, de masquer et de réaménager les fenêtres du bureau OmniPage SE, et de passer à d'autres pages du document. Un clic avec le bouton droit dans une fenêtre affiche son menu contextuel avec les commandes les plus utiles pour cette fenêtre.

Barre d'outils Standard

La barre d'outils Standard contient des boutons et une liste déroulante permettant d'effectuer les tâches courantes. Elle peut prendre la forme de barre flottante ou être ancrée à tout bord du bureau OmniPage SE. Toutes les fonctions proposées sur cette barre sont également accessibles par l'intermédiaire de menus.

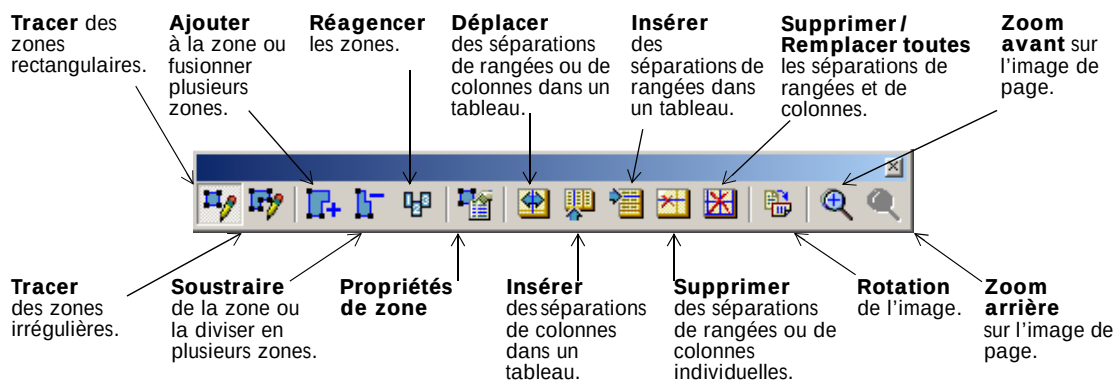


Barre des menus

Pour obtenir des informations succinctes sur une commande de menu, cliquez sur le bouton d'aide contextuelle puis sur la commande concernée. La fenêtre qui apparaît explique la fonction de la commande. Pour fermer cette fenêtre, cliquez n'importe où sur l'écran.

Barre d'outils Image

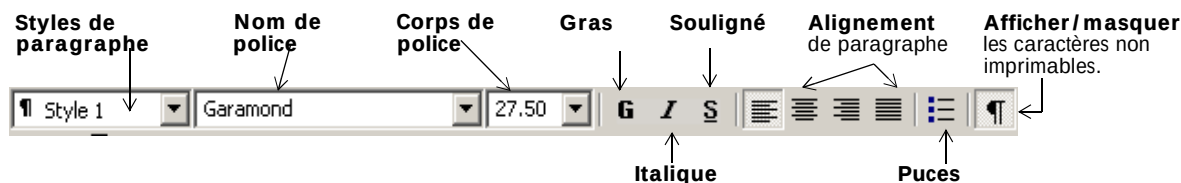
La barre d'outils Image contient des boutons permettant d'agrandir/de réduire et de faire pivoter l'image actuelle, ainsi que de manipuler les zones et les séparations de tableaux présentes sur la page. Ces fonctions sont décrites en détail dans le chapitre 3, dans *Tracer des zones manuellement* et *Gérer les grilles du tableau dans l'image*. L'illustration ci-dessous résume les rôles des différents boutons de la barre. La barre d'outils Image peut prendre la forme de barre flottante (librement déplaçable sur le bureau) ou être ancrée à tout bord de la fenêtre Image originale.



Astuce : Vous pouvez également redimensionner et faire pivoter l'image originale par le biais d'un menu contextuel. Cliquez avec le bouton droit dans la fenêtre Image originale (en dehors d'une zone) et sélectionnez une valeur de zoom ou de rotation.

Barre d'outils Formatage

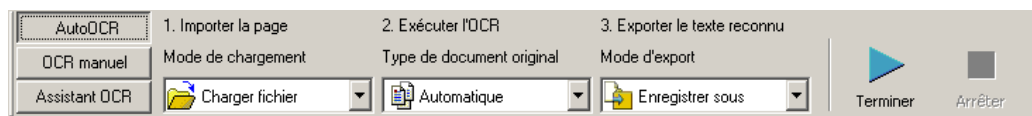
La barre d'outils Formatage contient des boutons permettant de modifier le texte reconnu dans l'Éditeur de texte. Ces fonctions sont décrites en détail au chapitre 4, dans *Éditer du texte et des images*. L'illustration ci-dessous résume les rôles des différents boutons de la barre, qui est toujours visible dans la partie supérieure de l'Éditeur de texte.



Boîte à outils OmniPage

La boîte à outils d'OmniPage vous permet de gérer toutes les étapes du traitement. Elle est placée par défaut dans la partie supérieure du bureau OmniPage SE, au-dessus des fenêtres de travail, mais peut également prendre la forme de fenêtre flottante ou être ancrée le long du bord inférieur du bureau.

La boîte à outils dispose de trois boutons à gauche déterminant le mode de traitement activé : AutoOCR, OCR manuel et Assistant OCR. Lorsque vous cliquez sur un bouton, la boîte à outils affiche les options correspondantes. L'illustration présentée au début de ce chapitre reflète le bureau OmniPage SE en mode OCR manuel. L'illustration ci-dessous présente les réglages disponibles pour l'option AutoOCR.



Le traitement automatique est lancé, arrêté et relancé à l'aide des boutons dans la partie droite de la barre d'outils. L'utilisation de ces boutons et tous les réglages y afférant sont décrits au chapitre 3, *Traiter des documents automatiquement*.

Vous pouvez passer du mode automatique au mode manuel à tout moment, à la seule condition que le programme ne soit pas en train d'effectuer une opération de traitement. Cette propriété signifie que vous pouvez modifier le

mode sélectionné entre les différentes étapes du traitement d'un document. Par exemple, il est possible de traiter automatiquement certaines pages, puis insérer dans le document d'autres pages traitées manuellement. Selon le même principe, vous pouvez appliquer le mode automatique à une pile de pages, puis vérifier le résultat obtenu et, si nécessaire, retraiter certaines pages manuellement. Cette procédure est décrite au chapitre 3 *Traitement combiné automatique/manuel*.

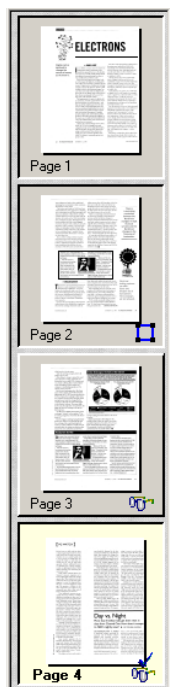
Pour lancer l'Assistant OCR, le bureau OmniPage SE ne doit comporter aucun document. Voir *Traiter des documents à l'aide de l'Assistant OCR*. Si vous utilisez l'Assistant pour traiter et enregistrer un document, celui-ci est par la suite disponible pour des étapes de traitement supplémentaires (ajout de nouvelles pages, retraitement des pages, etc.) en mode automatique ou manuel.

GÉRER DES DOCUMENTS

Le gestionnaire de document est situé dans la partie gauche du bureau OmniPage SE et présente deux onglets : Vignettes et Détails. Pour afficher la fenêtre correspondant à un onglet, cliquez sur celui-ci. Les deux fenêtres comportent des informations succinctes sur les pages du document traité et sont synchronisées (lors du passage d'une fenêtre à l'autre, les pages actuelle et sélectionnées sont conservées). Les illustrations présentent ces deux fenêtres avec le même document de quatre pages. Les pages 1 et 2 sont sélectionnées et la page 4 est la page actuelle (c.-à-d. la page affichée dans la fenêtre Image originale). Le gestionnaire de document indique l'état du traitement d'une page à l'aide des icônes suivantes :

Page	État	Icône (Vignettes)	Icône (Détails)	Opérations effectuées sur la page
1	Importée	—		Page importée, sans zones définies manuellement ou via un gabarit, et non reconnue.
2	Divisée en zones		—	Page importée, avec zones définies manuellement ou via un gabarit, et non reconnue.
3	Reconnue			Page reconnue, non vérifiée ou vérification interrompue.
4	Vérifiée			Page reconnue et entièrement vérifiée.

Fenêtre Vignettes



L'onglet Vignettes contient des images réduites de toutes les pages du document. Vu leur présentation verticale, vous devrez peut-être faire défiler les vignettes pour atteindre l'élément requis. La page actuelle se caractérise par un fond clair et son numéro de page apparaît en gras. Vous pouvez sélectionner plusieurs pages dans un document ; toute page sélectionnée présente un aspect « en relief ». Chaque page dispose d'une icône d'état dans sa partie inférieure droite (voir la section précédente).

Pour passer à une page, cliquez sur sa vignette.

Pour modifier l'ordre d'apparition d'une page, cliquez sur sa vignette et faites-la glisser sur le numéro de page voulu. Les pages sont renumérotées automatiquement.

Pour supprimer une page, sélectionnez sa vignette et appuyez sur la touche Suppr.

Pour sélectionner plusieurs pages adjacentes, maintenez enfoncée la touche Maj. et cliquez sur la première, puis sur la dernière vignette du groupe requis. Pour sélectionner plusieurs vignettes non adjacentes, maintenez enfoncée la touche Ctrl et cliquez sur les pages individuelles pertinentes. Cela permet ensuite de déplacer, supprimer ou soumettre à l'OCR le groupe de pages sélectionnées.

Fenêtre Détails

Cette nouvelle fonction a été introduite dans OmniPage SE pour vous fournir un résumé d'informations sur un document sous la forme d'un tableau. Les rangées représentent les pages ; les colonnes offrent les données statistiques ou d'état des pages, et (le cas échéant) spécifient les statistiques globales du document. L'illustration ci-dessous affiche les colonnes par défaut (à gauche) et quatre colonnes définies par l'utilisateur.

Page	Etat	Mots d...	Car...	Mots	Préci...	Mots/mi...	Duré...	Zones	Tabl...	Gra...	Taille de l'ima...
1	Importé	0	0	0	0.00	0	0.000	0	0	0	2560x3328
2	Importé	0	0	0	0.00	0	0.000	0	0	0	2560x3328
3	Reconnu	3	3659	726	100...	4661	9.344	32	0	4	2560x3328
4	Vérifié	16	4817	979	99.96	5052	11....	14	0	0	2560x3328
		19	8476	1705	99.98	4878	20....	46	0	4	

Pour afficher la vignette d'une page, placez le curseur sur son icône d'état.

Nombre de zones de différents types présentes sur la page.

La page actuelle est contrastée. Dans la fenêtre Détails, vous pouvez effectuer les opérations suivantes :

Pour passer à une page, cliquez sur sa rangée.

Pour modifier l'ordre d'apparition d'une page, cliquez sur sa rangée et faites-la glisser à l'emplacement requis. Une flèche s'affiche pour vous signaler l'endroit d'insertion de la page. Les pages sont renumérotées automatiquement.

Pour supprimer une page, sélectionnez sa rangée et appuyez sur la touche Suppr.

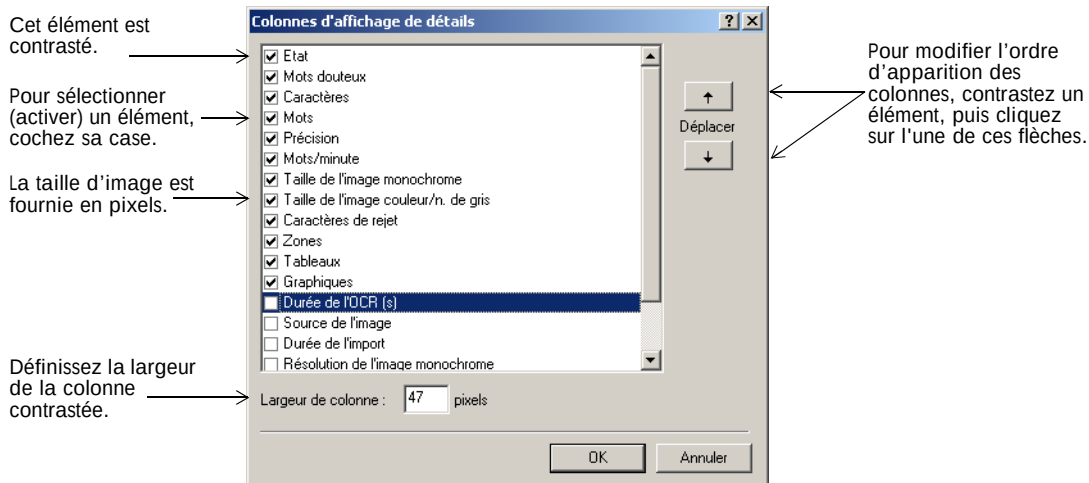
Pour sélectionner plusieurs pages adjacentes, maintenez enfoncée la touche Maj. et cliquez sur la première, puis sur la dernière rangée du groupe requis. Pour sélectionner plusieurs rangées non adjacentes, maintenez enfoncée la touche Ctrl et cliquez sur les éléments individuels pertinents. Cela permet ensuite de déplacer, supprimer ou soumettre à l'OCR le groupe de pages sélectionnées.

La sélection de plusieurs pages n'affecte pas l'état de la page actuelle. Toutes les pages sélectionnées sont contrastées.

Astuce : Pour consulter les informations sur la taille d'une image, placez le curseur sur sa vignette ou en dehors d'une zone dans la fenêtre Image originale. La mini-fenêtre qui s'affiche spécifie la taille de l'image en pixels et en unités de mesure du programme. Vous êtes également informé de la résolution de l'image.

Personnaliser des colonnes dans la fenêtre Détails

Vous pouvez spécifier les colonnes que vous souhaitez inclure dans la fenêtre Détails. Dans le menu Affichage, cliquez sur Personnaliser les détails pour afficher la boîte de dialogue illustrée ci-dessous.



Définissez les colonnes à afficher, leur largeur et leur ordre d'apparition. Le contenu de chaque colonne est décrit en détail dans la rubrique *Colonnes de la fenêtre Détails* de l'aide en ligne. Vous pouvez modifier la largeur des colonnes tout simplement en faisant glisser les séparations de colonne voulues dans la barre de titre.

Supprimer des pages d'un document

Toute suppression de page doit être confirmée et ne peut pas être annulée. Pour supprimer la page actuelle, vous devez recourir à la commande Supprimer la page actuelle du menu Édition. Pour supprimer toutes les pages sélectionnées dans le gestionnaire de document (fenêtre Vignettes ou Détails), appuyez sur la touche Suppr ou utilisez la commande Effacer de son menu contextuel.

Imprimer un document

Pour imprimer un document, utilisez la commande Imprimer du menu Fichier. Vous pouvez décider d'imprimer une image de page ou le texte reconnu (c.-à-d., le résultat de l'OCR affiché dans l'Éditeur de texte). Vous pouvez imprimer tout ou partie des pages d'un document. Le bouton Imprimer de la barre d'outils Standard envoie à l'imprimante des images ou du texte, en fonction de la fenêtre active (Image originale ou Éditeur de texte).

Fermer un document

Pour fermer un document, utilisez la commande Fermer du menu Fichier. Si vous n'avez pas effectué de sauvegarde du document ou si vous y avez apporté des modifications depuis le dernier enregistrement, vous êtes invité à l'enregistrer. Pour plus de détails sur l'enregistrement d'un document au format Document OmniPage (*.opd), reportez-vous à la section qui suit. Un message vous demandera d'enregistrer les nouvelles données d'apprentissage, si vous avez sélectionné l'option « Demander l'enregistrement des données IntelliTrain à la fermeture de document » dans la section Vérification de la boîte de dialogue Options.

La dernière phrase ne s'applique pas à OmniPage SE.



DOCUMENTS OMNIPAGE

Le format Document OmniPage correspond au type de fichier propre à OmniPage SE et Pro (extension .opd). C'est un des types de fichiers disponibles pour l'enregistrement d'un document. Le format OPD est destiné à l'enregistrement des documents sur lesquels vous avez l'intention de retravailler plus tard dans OmniPage SE. Ainsi, vous pourrez terminer le traitement du document, y ajouter des pages ou modifier le texte reconnu. Un Document OmniPage contient l'image de page originale et toute zone tracée sur celle-ci. À l'issue de la reconnaissance, un fichier OPD comporte également le résultat de l'OCR. Les caractères reconnus sont enregistrés avec leurs coordonnées et détails techniques, ce qui permet de conserver le lien entre l'image et le texte correspondant, et rend possible la vérification lors de la réouverture du fichier dans OmniPage SE.

Au moment de l'enregistrement au format Document OmniPage, le logiciel sauvegarde également les réglages en cours (et les nouvelles données d'apprentissage). Lors de l'ouverture d'un Document OmniPage, ces réglages sont appliqués, remplaçant temporairement les paramètres du programme.

Avantages du format OPD

Le format OPD, proposé parmi d'autres types de fichiers disponibles dans OmniPage SE, est particulièrement adapté aux circonstances suivantes :

- Vous ne pouvez pas terminer le traitement d'un document pendant la session en cours.
- Vous souhaitez transmettre le document à d'autres utilisateurs disposant d'OmniPage SE ou Pro 11 (par exemple, l'envoyer à un spécialiste pour vérification). Votre bureau est équipé d'un scanner générant des images destinées à la reconnaissance et la vérification sur plusieurs postes de travail.
- Vous souhaitez mettre en place une archive des documents reconnus dont les images originales restent accessibles (le texte reconnu permet d'effectuer des recherches par mot-clé et accepte d'autres techniques de récupération de documents).

Remarque : Avant de procéder à l'installation de toute mise à jour d'OmniPage, veillez à enregistrer vos résultats de reconnaissance dans un format autre que le format OPD. Ces fichiers peuvent être incompatibles avec les formats OPD plus récents ou conserver uniquement des images.

Enregistrer au format OPD

Si vous avez l'intention de créer un fichier OPD, il est recommandé d'enregistrer votre document sous ce format dès le début de son traitement (à des fins de protection). Pendant une session de travail, utilisez régulièrement le bouton Enregistrer, et resauvegardez le document à la fin de la session.

Le bouton Enregistrer sauvegarde le document actuel sous le même nom et type de fichier, mais vous pouvez enregistrer votre document sous plusieurs formats. Si lors du premier enregistrement du document, vous l'avez affecté d'un format autre que OPD (par exemple, .DOC), utilisez la commande Enregistrer sous du menu Fichier pour le sauvegarder en tant que type de fichier OPD. Si vous enregistrez un document OPD sous un autre format, le programme ne le réenregistre pas automatiquement au format OPD. Au moment de la fermeture du document ou de l'application, vous serez invité à enregistrer le document en tant que fichier OPD.

RÉGLAGES

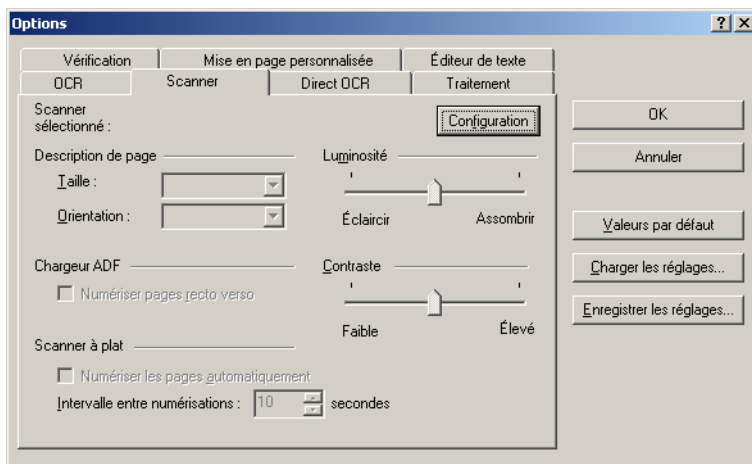
Les réglages OmniPage SE sont regroupés dans les sept onglets de la boîte de dialogue Options. Chaque onglet est présenté sommairement ci-dessous (reportez-vous à l'aide contextuelle en ligne pour obtenir des informations détaillées sur chaque fenêtre de réglages).

OCR

Permet de spécifier la ou les langues de reconnaissance, un dictionnaire utilisateur, des caractères de rejet, la méthode d'OCR requise (optimisée pour vitesse ou précision de traitement) et le mappage des polices.

Scanner

Permet de définir la taille et l'orientation de la page pour la numérisation. Vous pouvez également spécifier les réglages de luminosité et de contraste, et définir les options de numérisation des documents multipages, avec ou sans chargeur ADF. Cet onglet permet aussi de modifier la configuration du scanner, d'installer un nouveau scanner ou de changer de scanner par défaut.



Direct OCR™

Permet de définir les réglages de Direct OCR, fonction assurant les services de reconnaissance directement à partir des applications extérieures (traitement de texte, etc.). Utilisez cet onglet pour enregistrer/désenregistrer des applications pour Direct OCR et activer/désactiver ce service. Vous pouvez également sélectionner le découpage en zones automatique ou manuel, et activer ou désactiver la vérification.

Traitement

Permet de définir l'emplacement d'insertion de nouvelles images et de spécifier d'autres préférences de traitement. Cet onglet vous donne aussi la possibilité de modifier la langue d'interface.

Vérification

Permet d'activer ou de désactiver la vérification automatique à la suite de la reconnaissance. Vous pouvez aussi utiliser cet onglet pour configurer IntelliTrain dans le but de charger ou de travailler avec un fichier d'apprentissage. Pour plus de détails, reportez-vous au chapitre 4, *Vérifier les résultats de l'OCR*.



Les références au module IntelliTrain et aux fichiers d'apprentissage ne s'appliquent pas à OmniPage SE.

Mise en page personnalisée

Permet de décrire en détail la mise en page du document d'entrée. Grâce à cet onglet, vous pouvez configurer le découpage en zones automatique (recherche ou omission de colonnes, de graphique et de tableaux).

Éditeur de texte

Permet d'afficher ou de masquer certaines fonctions de l'Éditeur de texte, de définir l'unité de mesure à appliquer et d'activer ou de désactiver le retour à la ligne automatique.

Remarque : Certains réglages s'appliquent uniquement aux sessions de reconnaissance suivantes (c'est le cas, par exemple, de la langue de reconnaissance, du fichier d'apprentissage et de la luminosité du scanner). De ce fait, il est important de définir correctement ces réglages avant le lancement d'un traitement. Pour appliquer les nouveaux réglages aux pages reconnues, vous devez recommencer l'OCR de celles-ci. Par contre, les autres réglages sont implémentés immédiatement dans les pages existantes (par exemple, retour à la ligne automatique ou unité de mesure dans l'Éditeur de texte).



3 Didacticiel : Traitement des documents

Ce chapitre vous propose un Guide de démarrage rapide, décrit les modes de traitement des documents disponibles sous OmniPage SE et fournit des détails sur les principales étapes des procédures impliquées :

- ◆ Guide de Démarrage Rapide
- ◆ Traiter des documents à l'aide de l'Assistant OCR
- ◆ Traiter des documents automatiquement
- ◆ Traiter des documents manuellement
- ◆ Traitement combiné automatique/manuel
- ◆ Traitement à partir d'autres applications (Direct OCR, PaperPort)
- ◆ Traiter des documents avec Planifier OCR

Les instructions détaillées portent sur les opérations suivantes :

- ◆ Définir la source des images de page
- ◆ Spécifier la mise en page du document
- ◆ Tracer des zones manuellement
- ◆ Gérer les grilles du tableau dans l'image
- ◆ Utiliser des gabarits de zone

GUIDE DE DÉMARRAGE RAPIDE

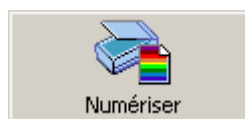
Les paragraphes ci-dessous vous guident à travers les principales étapes de la reconnaissance des caractères.



Charger et reconnaître les fichiers images exemples

Les fichiers images exemples (comportant une seule page et multipages) sont mis à votre disposition dans le dossier Programmes. Traitez ces fichiers en appliquant la procédure ci-dessous, mais dans un premier temps ignorez toute référence au scanner. Pour plus d'informations sur *l'importation des images*, reportez-vous à la section Importer des fichiers image. Vous obtiendrez ainsi une référence indiquant la qualité d'OCR que vous devriez obtenir avec vos propres fichiers aux caractéristiques comparables.

Par la suite, passez à la numérisation d'une page placée dans votre scanner.



Numérisation et reconnaissance d'un document d'une page

Mettez votre scanner sous tension et assurez-vous qu'il fonctionne correctement. Préparez une page au texte net et de bonne qualité.

Ce test présuppose qu'OmniPage SE présente les réglages par défaut et que votre document est rédigé dans la langue qui a été spécifiée comme langue d'interface pendant l'installation. S'il ne s'agit pas de la première utilisation du programme, ouvrez la boîte de dialogue Options (via menu Outils) et choisissez Valeurs par défaut.

Le document sera traité automatiquement et les résultats de l'OCR seront enregistrés dans un fichier. Vous allez effectuer la vérification, mais non l'édition, du document dans l'éditeur de texte d'OmniPage SE.

	Action	Résultat
1.	Le cas échéant, configurez le scanner à l'aide de l'Assistant de numérisation.	OmniPage SE est configuré pour l'utilisation avec votre scanner.
2.	Choisissez Démarrer ➤ Programmes ➤ Scan-Soft OmniPage SE ➤ OmniPage SE	Le programme OmniPage SE est ouvert.
3.	Placez le document soigneusement dans le scanner.	
4.	Cliquez sur les trois onglets à gauche de la boîte d'outils OmniPage. Le bouton AutoOCR devrait être sélectionné. Si ce n'est pas le cas, cliquez dessus.	Le traitement automatique du document par OmniPage SE selon les réglages en cours est activé.
5.	Dans le menu déroulant Importer la page, sélectionnez l'option de numérisation adaptée à votre document : noir et blanc, niveaux de gris ou couleur.	Vous pouvez spécifier l'aspect des images ou des textes et fonds de couleur dans le document exporté. Notez que la numérisation en couleur exige un scanner couleur.
6.	Dans le menu déroulant Type de document original, vérifiez que la commande Automatique est sélectionnée. Ce réglage est le plus approprié à un grand nombre de documents.	OmniPage SE est configuré pour le traçage des zones et spécification des propriétés de zones automatiques.
7.	Dans le menu déroulant Exporter le texte reconnu, vérifiez que la commande Enregistrer sous est sélectionnée.	Vous pouvez attribuer un nom au fichier d'exportation après la vérification du document.
8.	Cliquez sur Démarrer.	OmniPage SE lance la numérisation de votre document.
9.	La fenêtre Vérification de l'OCR apparaît et vous invite à modifier les mots que le programme considère comme reconnus incorrectement.	Le module de vérification de l'OCR est similaire à la fonction de vérification orthographique d'un programme de traitement de texte, mais offre des fonctionnalités supplémentaires spécifiques à l'OCR.
10.	Cliquez dans la fenêtre de l'éditeur de texte. Sélectionnez les différents modes d'éditeur pour comparer les modifications d'affichage de la page. Choisissez le mode d'exportation requis.	Chaque mode de l'éditeur de texte définit un niveau de formatage différent. Le mode spécifié lors de l'enregistrement est appliqué au texte du fichier.
11.	Cliquez sur Reprendre pour relancer la vérification. Lorsque le message La vérification de l'OCR est terminée s'affiche. Cliquez sur OK.	La vérification de l'OCR est terminée. La boîte de dialogue Enregistrer sous apparaît.
12.	Choisissez l'emplacement et le type de fichier requis pour l'enregistrement du document reconnu. Cliquez sur OK.	Par défaut, la fonction Enregistrer et lancer est activée et votre document est automatiquement ouvert dans le programme de traitement de texte associé au type de fichier sélectionné.
13.	Contrôlez le document dans votre programme de traitement de texte.	Vous avez utilisé OmniPage SE avec succès pour reconnaître le document et l'ouvrir dans votre application cible.

Astuce : Si vous obtenez de bons résultats lors de la reconnaissance des fichiers images exemples, mais non de la page numérisée, vérifiez l'installation et la configuration de votre scanner, notamment la luminosité et la résolution d'image. Pour des informations sur le réglage de luminosité optimal, reportez-vous à la section Importer des documents à partir d'un scanner. Consultez également les rubriques Configurer le scanner et Dépanner le scanner de l'aide en ligne.

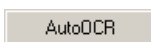
Les méthodes de traitement disponibles dans OmniPage SE sont présentées succinctement ci-dessous. Une description approfondie de chaque mode est fournie dans les pages suivantes du manuel.

Assistant OCR



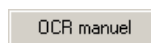
L'Assistant OCR vous guide dans la sélection des réglages et des commandes de reconnaissance en vous posant quelques questions sur le document. Cela fait, l'Assistant lance le traitement en mode automatique. Si vous n'avez pas l'habitude d'utiliser OmniPage SE, n'hésitez pas à faire appel à l'Assistant OCR.

Traitement automatique



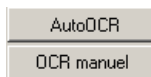
La solution la plus simple et la plus efficace de traiter un document consiste à laisser OmniPage SE s'en charger pour vous. Sélectionnez les réglages dans la boîte de dialogue Options et les commandes dans la barre d'outils AutoOCR, puis cliquez sur *Démarrer*. OmniPage SE soumet chaque page à toutes les étapes du traitement, si possible effectuant plusieurs opérations en parallèle. Sous ce mode, le logiciel se charge également de tracer les zones.

Traitement manuel



Le traitement manuel vous assure un contrôle plus précis du traitement de votre document. Vous pouvez traiter chaque page individuellement avec différents réglages. Le programme marque une pause entre chaque étape du traitement : importation d'images, reconnaissance et exportation. Vous pouvez ainsi tracer des zones manuellement ou modifier les langues de reconnaissance. Pour lancer une étape, vous cliquez sur le bouton correspondant dans la barre d'outils OCR manuel.

Traitement combiné automatique/manuel



Vous pouvez traiter un document automatiquement, puis afficher le résultat obtenu dans l'éditeur de texte. Si la qualité de l'ensemble du document est bonne et que seules certaines pages sont insatisfaisantes, vous pouvez passer au traitement manuel pour modifier les réglages et relancer la reconnaissance des pages incorrectes.

Traitement à partir d'autres applications

La fonction Direct OCR permet de faire appel aux services OCR d'OmniPage SE à partir de votre application de traitement de texte usuelle ou d'une autre application. De plus, OmniPage SE établit un lien automatique avec les programmes de gestion de documents de ScanSoft – PaperPort et Pagis.



Traitement différé (Planifier OCR)

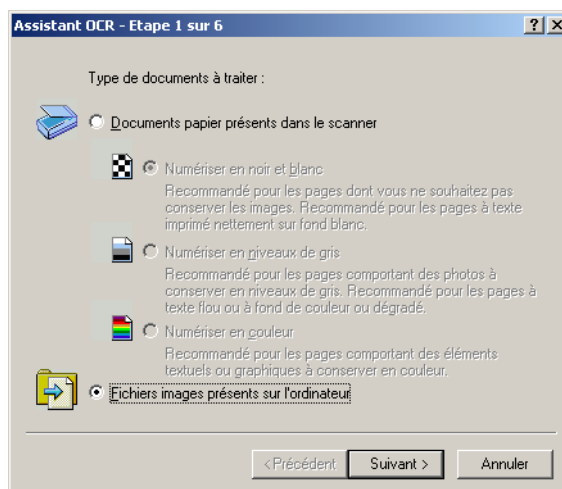
Le module Planifier OCR permet de différer la reconnaissance automatique des documents (par exemple, à une heure où vous n'utilisez pas l'ordinateur). Utilisez l'Assistant de planification OCR pour spécifier les réglages et l'heure de lancement.

TRAITER DES DOCUMENTS À L'AIDE DE L'ASSISTANT OCR

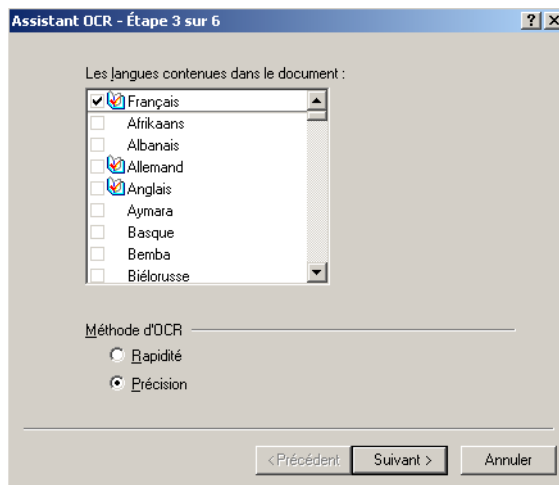
En vous guidant à travers six fenêtres de traitement, l'Assistant OCR vous aide à sélectionner les réglages appropriés à votre document, puis lance le traitement automatique. Toutes les fenêtres de l'Assistant sont munies d'aide contextuelle. L'Assistant OCR ne peut s'exécuter qu'à condition qu'aucun document ne soit ouvert dans OmniPage SE.



Pour afficher le premier écran de l'Assistant, cliquez sur l'onglet Assistant OCR dans la boîte à outils OmniPage, puis cliquez sur le bouton Assistant :



1. La première fenêtre vous permet de définir la source du document : scanner ou fichier image. Pour plus de détails, reportez-vous à la section *Définir la source des images de page*. Effectuez les sélections appropriées et cliquez sur *Suivant*.
2. La deuxième fenêtre vous demande de décrire la mise en page du document original dans le but de faciliter le traçage de zones automatique. Pour plus de détails, reportez-vous à la section *Spécifier la mise en page du document*.
3. La troisième fenêtre (illustrée ci-dessous) permet de définir la ou les langues de reconnaissance et de choisir la méthode d'OCR. Les langues associées à un dictionnaire sont accompagnées de l'icône 



4. La quatrième fenêtre permet de définir le niveau de formatage à appliquer au document lors de l'affichage et de l'exportation. Pour plus d'informations, consultez le chapitre 4, *Modes d'affichage de l'éditeur de texte*.
5. La cinquième fenêtre vous donne la possibilité de vérifier le texte avant de l'exporter. Si vous choisissez l'option *Oui*, vous pouvez aussi éditer le texte avant l'enregistrement. Vous décidez également si vous voulez créer et utiliser des données IntelliTrain pendant la vérification. Pour plus d'informations, consultez le chapitre 4. Les références au module IntelliTrain ne s'appliquent pas à OmniPage SE.



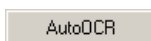
6. La dernière fenêtre vous demande de définir les modalités d'exportation : enregistrement dans un fichier ou copie dans le Presse-papiers. Après avoir effectué la sélection requise, cliquez sur *Terminer* pour fermer l'Assistant et démarrer le traitement automatique.
7. Si vous avez activé l'option de vérification et que le texte contient des mots douteux, la boîte de dialogue Vérification de l'OCR apparaît. À l'issue de la vérification, soit le résultat de la reconnaissance est placé directement dans le Presse-papiers, soit la boîte de dialogue Enregistrer sous apparaît pour vous permettre de spécifier les réglages d'exportation pour le fichier.
8. Le document est conservé dans OmniPage SE. Vous pouvez éditer le résultat de la reconnaissance et l'enregistrer sous différents formats. OmniPage SE vous permet de redéfinir des zones manuellement ou de modifier les réglages, puis utiliser le traitement manuel pour relancer la reconnaissance de certaines pages du document. Le document peut être enrichi de nouvelles pages via le traitement automatique ou manuel.

Remarque : Les fenêtres de l'Assistant contiennent toujours les derniers réglages spécifiés dans le programme. De plus, OmniPage SE applique les réglages sélectionnés dans l'Assistant OCR à toute opération de traitement automatique ou manuel (jusqu'à leur éventuelle modification). Par conséquent, si ces réglages conviennent à plusieurs documents, il suffit de passer à la barre d'outils AutoOCR et de cliquer sur *Démarrer*.

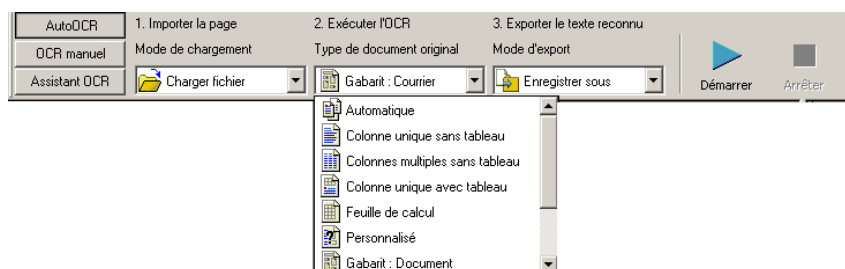
Remarque : Les réglages applicables qui ne sont pas proposés par l'Assistant OCR adoptent les valeurs les plus récentes spécifiées dans le programme (il s'agit principalement des réglages du scanner, du dictionnaire utilisateur et du fichier d'apprentissage). Les gabarits de zones ne peuvent pas être utilisés avec l'Assistant OCR. Si un fichier de gabarit est actif lors du lancement de l'Assistant, il est déchargé et l'option Automatique est appliquée à la description du document original. Vous ne pouvez pas exporter un document reconnu en tant que pièce jointe à un message électronique ; vous devez, pour cela, employer le traitement automatique ou manuel.

TRAITER DES DOCUMENTS AUTOMATIQUEMENT

Le mode automatique représente une méthode efficace pour le traitement de documents, notamment de documents volumineux. Vous définissez les réglages pertinents, puis utilisez la barre d'outils AutoOCR dans la boîte à outils OmniPage pour appliquer toutes les étapes du traitement à un nouveau document, ou pour relancer et terminer le traitement d'un document ouvert.



1. Cliquez sur l'onglet AutoOCR de la boîte à outils OmniPage pour afficher la barre d'outils AutoOCR.



2. Dans la liste déroulante du bouton Importer la page, sélectionnez la commande requise. Vous définissez la source du document (fichier image ou scanner). Pour plus de détails, reportez-vous à la section *Définir la source des images de page*.
3. Dans la liste déroulante du bouton Type de document original, sélectionnez l'option pertinente (voir l'illustration plus haut). La description du type de document importé ou le choix d'un gabarit de zones aide le programme dans le processus de traçage de zones automatique. Pour plus de détails, reportez-vous à la section *Spécifier la mise en page du document*.
4. Dans la liste déroulante du bouton Exporter le texte reconnu, sélectionnez la commande requise. Vous pouvez enregistrer le document reconnu dans un fichier, le copier dans le Presse-papiers ou l'envoyer comme pièce jointe d'un courrier électronique. Pour plus d'informations, consultez le chapitre 5.
5. Choisissez Options dans le menu Outils et vérifiez que les réglages en cours conviennent au document. Vous pouvez, par exemple, spécifier la ou les langues de reconnaissance et activer/désactiver la vérification du texte reconnu. Pour plus de détails, reportez-vous à la section *Réglages* du chapitre 2.
6. Cliquez sur *Démarrer* ou choisissez Démarrer dans le menu Traitement. Les pages du document sont traitées les unes après les autres, mais

OmniPage SE peut effectuer plusieurs tâches simultanément. Ainsi, le programme peut lancer le chargement et la reconnaissance d'une nouvelle page pendant la vérification de la page précédente.

Boutons de commande



Démarrer : Ce bouton permet de lancer le traitement automatique d'un nouveau document.



Arrêter : Ce bouton permet d'interrompre le traitement automatique. Cette commande est utile lorsque vous souhaitez modifier certains réglages alors que le traitement est déjà entamé. En cas d'interruption de traitement, le libellé du bouton *Démarrer* devient *Terminer*.

Le bouton Démarrer adopte un libellé différent en fonction de l'état du traitement (arrêté ou terminé).



Terminer : Ce libellé apparaît si toutes les étapes du traitement n'ont pas été accomplies. Vous pouvez alors effectuer les opérations suivantes :

- ◆ **Terminer** le traitement des pages partiellement traitées.
- ◆ **Exporter** le document sans les pages non reconnues.



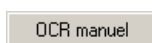
Autres pages : Ce libellé apparaît si toutes les pages du document ont été traitées et exportées. Vous pouvez alors effectuer les opérations suivantes :

- ◆ **Exporter** le document à nouveau, avec ou sans modification, sous un autre nom, emplacement ou type de fichier, ou avec un autre niveau de formatage.
- ◆ **Ajouter des pages** de la même source ou d'une autre source, en conservant ou modifiant les réglages.
- ◆ **Recommencer le traitement de toutes les pages** pour supprimer l'ensemble des résultats de reconnaissance et relancer l'OCR sur toutes les pages du document avec des réglages différents. Vous pouvez activer le découpage en zones automatique ou appliquer un fichier de gabarit.

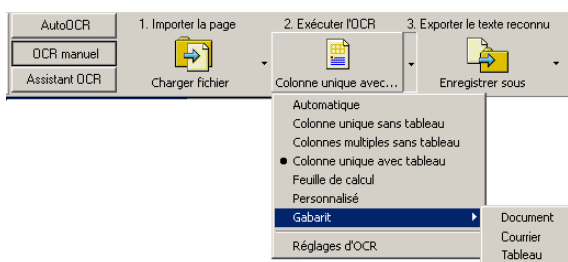
Astuce : Vous pouvez relancer la reconnaissance de toutes les pages si un réglage inadéquat a produit des résultats insatisfaisants dans l'intégralité du document. Une telle intervention peut s'imposer, par exemple, lorsque la spécification d'une langue incorrecte lors du premier traitement a entraîné le marquage de la plupart des mots du document comme douteux pendant la vérification. Cette option permet de réexécuter l'OCR sans avoir à recommencer la numérisation ou le chargement de toutes les images.

TRAITER DES DOCUMENTS MANUELLEMENT

Le traitement manuel vous assure un contrôle plus précis du traitement de votre document. Vous pouvez traiter chaque page individuellement avec différents réglages. Le programme marque une pause entre chaque étape du traitement : importation d'images, reconnaissance et exportation. Vous pouvez ainsi tracer des zones manuellement ou modifier les langues de reconnaissance. Pour lancer une étape, vous cliquez sur le bouton correspondant de la barre d'outils OCR manuel.

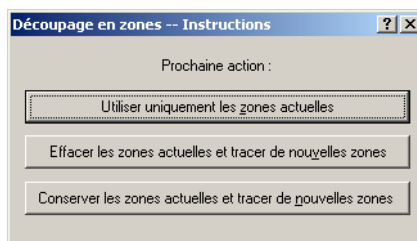


1. Cliquez sur l'onglet OCR manuel de la boîte à outils OmniPage pour afficher la barre d'outils OCR manuel.



2. Cliquez sur  dans la barre d'outils Standard ou choisissez Options dans le menu Outils pour vérifier ou modifier les réglages de la boîte de dialogue Options. Pour plus de détails, reportez-vous à la section *Réglages* du chapitre 2.
3. Dans la liste déroulante du bouton Importer la page, sélectionnez la commande requise. Vous définissez la source du document (fichier image ou scanner). Ouvrez la boîte de dialogue des réglages de scanner et choisissez les valeurs pertinentes. Pour plus de détails, reportez-vous à la section *Définir la source des images de page*.
4. Cliquez sur le bouton Importer la page. Cette action affiche la boîte de dialogue Charger fichier (permettant d'indiquer les noms de fichiers image) ou lance la numérisation. Une ou plusieurs images apparaissent dans le gestionnaire de document et une image s'affiche dans la fenêtre Image originale.
5. À ce stade, vous pouvez tracer et modifier manuellement des zones sur une ou plusieurs images, et leur assigner des propriétés. Les boutons de la barre d'état vous permettent de passer d'une page à une autre. Toute image dépourvue de zones sera soumise à un découpage en zones automatique lors du lancement de la reconnaissance. Pour plus de détails, reportez-vous à la section *Tracer des zones manuellement*.

6. Dans la liste déroulante du bouton Exécuter l'OCR, sélectionnez la commande requise. Vous décrivez la mise en page du document original. La valeur spécifiée est prise en compte dans le processus de traçage de zones automatique. Vous pouvez également sélectionner un gabarit pour appliquer ses zones à la page actuelle. Pour plus de détails, reportez-vous aux sections *Spécifier la mise en page du document* et *Utiliser des gabarits de zone*.
7. Cliquez sur le bouton Exécuter l'OCR pour reconnaître la page actuelle. Pour reconnaître plusieurs pages, sélectionnez le groupe de pages requis dans le gestionnaire de document (voir *Gérer des documents* du chapitre 2) et cliquez sur le bouton Exécuter l'OCR.
8. La boîte de dialogue Découpage en zones -- Instructions apparaît (à moins qu'elle ne soit désactivée).



Lorsque vous choisissez l'une des options proposées, la reconnaissance démarre.

9. Si vous avez activé l'option de vérification, la boîte de dialogue Vérification de l'OCR affiche la liste des mots douteux provenant des pages reconnues. Vous pouvez vérifier et éditer le texte traité (voir *Vérifier les résultats de l'OCR* du chapitre 4).
10. Continuez à charger les pages, à exécuter l'OCR et à procéder aux opérations d'édition et de vérification selon vos besoins.
11. Dans la liste déroulante du bouton Exporter le texte reconnu, sélectionnez la commande requise. Vous pouvez enregistrer le document reconnu dans un fichier (y compris sous le format de Document OmniPage), le copier dans le Presse-papiers ou l'envoyer comme pièce jointe d'un courrier électronique. Vous pouvez enregistrer le document à plusieurs reprises pour obtenir différents formats (pour plus d'informations, consultez le chapitre 5).

Remarque : Si vous désactivez l'option « Tracer des zones en plus de celles du gabarit » (onglet Traitement de la boîte de dialogue Options), la boîte de dialogue Découpage en zones -- Instructions n'apparaît pas et la reconnaissance prend en compte uniquement les zones actuelles.

TRAITEMENT COMBINÉ AUTOMATIQUE / MANUEL

Dans le cas d'un document volumineux dont seules quelques pages exigent une attention particulière de votre part, vous n'avez pas besoin de traiter manuellement l'intégralité du document. Vous pouvez le traiter automatiquement dans un premier temps, puis afficher le résultat obtenu dans l'éditeur de texte, où vous identifiez les pages reconnues correctement, et celles qui nécessitent des réglages différents ou un traçage de zones manuel. Cela fait, vous passez au mode manuel pour modifier certains paramètres et relancer le découpage en zones et la reconnaissance des pages inadéquates.

1. Préparez le document et exécutez un traitement en mode automatique, selon la procédure décrite dans les sections précédentes.
2. Si vous interrompez ou terminez la vérification, vous serez invité à enregistrer le document. Il est recommandé d'effectuer une sauvegarde, même si le document se trouve dans un état intermédiaire.
3. Sélectionnez une page exigeant la redéfinition des zones ou des réglages différents, et cliquez sur l'onglet OCR manuel dans la partie gauche de la boîte à outils OmniPage.
4. Supprimez ou modifiez les zones existantes dans la fenêtre Image originale. Vous pouvez également charger un gabarit pour appliquer ses zones à la page actuelle. Tracez les nouvelles zones selon vos besoins. Pour plus de détails, reportez-vous à la section *Tracer des zones manuellement*.
5. Modifiez les autres réglages en fonction des particularités de la page actuelle. Pour plus de détails, reportez-vous à la section *Réglages* du chapitre 2.
6. Cliquez sur le bouton Exécuter l'OCR pour relancer la reconnaissance de la page actuelle. Dans la fenêtre affichée, confirmez l'annulation des résultats de l'OCR précédents. La boîte de dialogue Découpage en zones -- Instructions apparaît, à moins qu'elle ne soit désactivée.

7. Pour relancer la reconnaissance de plusieurs pages, sélectionnez le groupe de pages voulu dans le gestionnaire de document avant de cliquer sur le bouton Exécuter l'OCR.
8. Lorsque toutes les pages soumises au nouveau traitement fournissent un résultat satisfaisant, réenregistrez le document.

TRAITEMENT À PARTIR D'AUTRES APPLICATIONS

La fonction Direct OCR permet de faire appel aux services OCR d'OmniPage SE à partir de votre application de traitement de texte usuelle ou d'une autre application. Le recours à cette fonction suppose l'établissement préalable d'un lien direct entre OmniPage SE et l'application en question. À l'issue de cette opération, deux commandes ajoutées au menu Fichier de l'application relient celle-ci aux services d'OCR.

Configurer Direct OCR

1. Démarrez l'application à lier à OmniPage SE. Lancez OmniPage SE, ouvrez la boîte de dialogue Options, puis la fenêtre Direct OCR et sélectionnez « Activer Direct OCR ».
2. La fenêtre Non enregistrées affiche la liste des applications actives ou enregistrées dans le passé. Sélectionnez la ou les applications voulues et cliquez sur Ajouter ou recherchez une application absente de la liste au moyen du bouton Parcourir. Sélectionnez les options de traitement souhaitées, qui tiendront lieu de préférences.

Utiliser Direct OCR

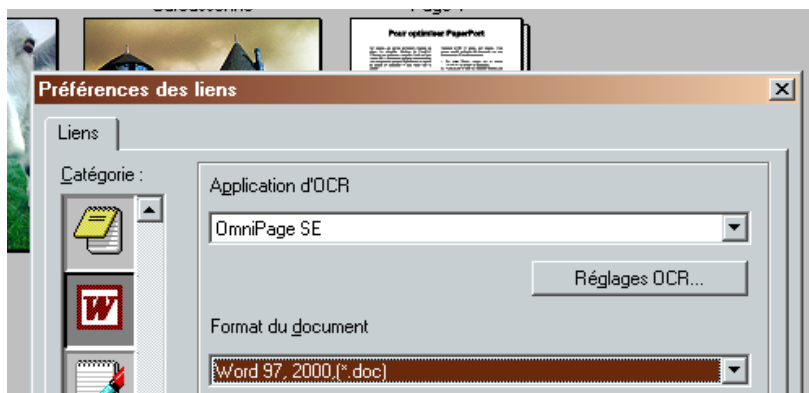
1. Démarrez l'application liée à OmniPage SE et ouvrez un document. Pour importer dans le document les résultats des pages numérisées, placez celles-ci correctement dans le scanner.
2. Sélectionnez la commande Acquérir les réglages de texte du menu Fichier pour spécifier les réglages à appliquer lors de la reconnaissance. Tout réglage non défini conserve la valeur utilisée lors de la dernière session OmniPage SE, et tout réglage modifié dans Direct OCR est reflété dans OmniPage SE.

3. Sélectionnez la commande Acquérir texte du menu Fichier pour importer des images à partir d'un scanner ou d'un fichier.
4. Si vous avez sélectionné l'option « Définir les zones automatiquement » dans la fenêtre Direct OCR de la boîte de dialogue Options ou dans la fenêtre Acquérir les réglages de texte, la reconnaissance est lancée immédiatement.
5. Si vous n'avez pas sélectionné l'option « Définir les zones automatiquement », le programme vous propose de tracer des zones manuellement sur chaque image de page. Cliquez sur le bouton Exécuter l'OCR pour lancer la reconnaissance.
6. Si la fonction de vérification a été activée, cette opération suit l'OCR. Le texte reconnu est ensuite inséré dans votre application à l'emplacement du curseur, avec le niveau de formatage spécifié avec la commande Acquérir les réglages de texte.

Remarque : Si OmniPage SE est ouvert au moment de l'appel de la fonction Direct OCR à partir d'une application cible, une nouvelle session d'OmniPage SE démarre.

Utiliser OmniPage SE avec PaperPort

PaperPort® est un produit de gestion de documents papier de ScanSoft. Ce programme permet de lier vos pages aux applications appropriées. Ces pages peuvent comporter des éléments graphiques et/ou du texte. Si PaperPort est présent sur un ordinateur lors de l'installation d'OmniPage SE, les services d'OCR de ce dernier étendent les fonctionnalités de PaperPort. Pour choisir un programme d'OCR dans PaperPort, cliquez avec le bouton droit sur le lien d'applications texte, sélectionnez Préférences, puis spécifiez OmniPage SE en tant que programme d'OCR à utiliser. À l'instar de Direct OCR, vous pouvez spécifier les réglages d'OCR de votre choix.



Dans l'illustration ci-dessus, OmniPage SE a été sélectionné comme programme d'OCR pour Word 2000. Par la suite, vous pouvez faire glisser des images de page depuis le bureau PaperPort sur l'icône de Word dans la barre des liens. Lors de la reconnaissance du texte, seule une jauge de progression apparaît ; toutefois, l'utilisateur peut demander l'affichage de la fenêtre de découpage en zones manuel ou de vérification d'OmniPage SE. Les résultats de l'OCR sont placés dans l'application cible, dans un nouveau document sans titre.

TRAITER DES DOCUMENTS AVEC PLANIFIER OCR



Le module de planification OCR permet de différer le traitement automatique de vos travaux d'OCR (dans la limite de 24 heures suivant l'activation de cette fonction). Chaque « travail planifié » porte sur un document. Les pages de celui-ci peuvent provenir d'un scanner muni d'un chargeur ADF ou d'un fichier image. Lors du démarrage d'un travail planifié, vous n'avez pas besoin d'être présent et OmniPage SE ne doit pas obligatoirement être actif. L'ordinateur doit être allumé au moment du lancement d'un travail différé, mais vous pouvez l'éteindre entre la configuration de Planifier OCR et l'heure du début du traitement planifié. Si le travail en question implique la numérisation, votre scanner doit être actif au moment du démarrage du travail et les pages correspondantes doivent être placées dans le chargeur ADF. Pour configurer un travail planifié, procédez comme suit :

1. Cliquez sur Planifier OCR dans le menu Traitement ou dans le menu Démarrer de Windows : sélectionnez Programmes ► ScanSoft ► OmniPage SE ► Planifier OCR.
2. La boîte de dialogue Planifier OCR apparaît. Cliquez sur Ajouter un travail pour accéder à l'Assistant du même nom. Celui-ci vous guide à travers six fenêtres de traitement (similaires à celles de l'Assistant OCR).
3. La première fenêtre vous permet de définir la source de l'image. Grâce à une fonction supplémentaire, vous pouvez effectuer le traitement de tous les fichiers image supportés dans le dossier de votre choix.
4. Les trois fenêtres suivantes sont similaires aux écrans de l'Assistant OCR, mais vous pouvez en outre spécifier un dictionnaire utilisateur. OmniPage Pro 11 vous permet d'utiliser un fichier d'apprentissage et/ou d'exécuter le module IntelliTrain. Ces fonctionnalités ne sont pas disponibles dans OmniPage SE.
5. La cinquième fenêtre vous permet de spécifier le nom, le type et l'emplacement du fichier d'exportation, ainsi qu'une option de séparation de fichier.
6. La dernière fenêtre permet de définir le moment du lancement du travail, la conservation ou la suppression des fichiers d'origine après traitement, et l'utilisation ou non d'un fichier journal pour la consignation de l'accomplissement du travail et de tout problème éventuel. Pour fermer l'Assistant, cliquez sur Terminer.



Remarque : La boîte de dialogue Planifier OCR affiche tous les travaux planifiés et indique leur état : En attente, En cours, Erreur ou Terminé. Pour modifier les réglages d'un travail en attente, sélectionnez **Modifier le travail**. Vous pouvez modifier et réutiliser les travaux accomplis pour traiter de nouveaux documents exigeant des réglages similaires ; vous pouvez aussi supprimer les travaux terminés lorsque vous n'en avez plus besoin.

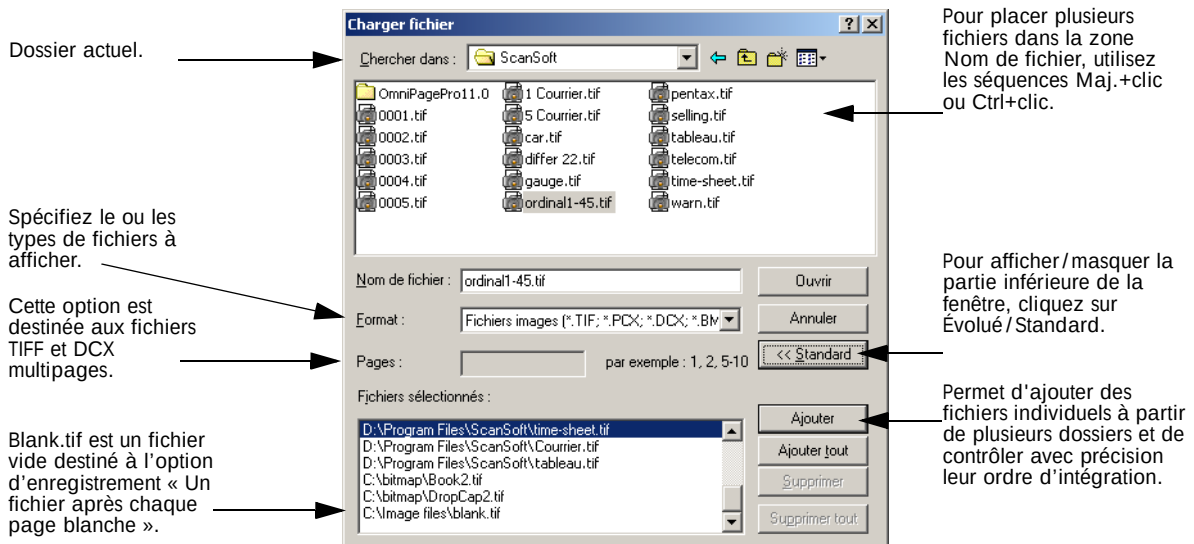
Pour plus de détails, reportez-vous à la rubrique *Planifier l'OCR* dans l'aide en ligne.

DÉFINIR LA SOURCE DES IMAGES DE PAGE

Les images de page peuvent provenir des sources suivantes : fichier image ou scanner. Il existe essentiellement deux types de scanner : à plat et à alimentation feuille à feuille. Pour faciliter la numérisation des documents multipages, un scanner peut être équipé d'un chargeur de feuilles automatique (ADF). Les images des documents numérisés peuvent être soit exportées directement dans OmniPage SE, soit enregistrées avec le logiciel du scanner au format de fichier image pour ouverture dans OmniPage SE.

Importer des fichiers image

Vous pouvez créer un fichier image à partir de votre scanner, ou le recevoir par e-mail ou par fax. La liste des types de fichiers image compatibles avec OmniPage SE est très étendue (voir *Types de fichiers supportés* dans Chapitre 6). Les fichiers image disponibles pour l'importation figurent dans la boîte de dialogue Charger fichier, qui apparaît au moment du lancement d'un traitement automatique. Lors d'un traitement manuel, cliquez sur le bouton Charger fichier ou utilisez le menu Traitement. Les réglages évolués proposés dans la partie inférieure de la boîte de dialogue illustrée ci-dessous peuvent être affichés ou masqués selon vos préférences.



Lorsque vous intégrez un fichier avec le bouton Ajouter, il est placé en fin de liste. Pour insérer un fichier à un emplacement différent, contrastez un fichier dans la liste : le nouveau fichier sera ajouté immédiatement en dessous du fichier contrasté le plus bas.

Importer des documents à partir d'un scanner

Pour bénéficier de cette fonction, vous devez disposer d'un scanner pris en charge, opérationnel et correctement configuré dans OmniPage SE. Pour plus d'informations, consultez le chapitre 1. Le programme offre plusieurs modes de numérisation parmi lesquels vous choisissez en prenant en considération deux facteurs :

- ◆ le type de graphique souhaité dans le document de sortie ;
- ◆ le mode qui produira la reconnaissance la plus précise.

Numériser en noir et blanc

Sélectionnez cette option pour numériser vos documents en noir et blanc. Cette option ne convient pas aux documents dont vous voulez conserver la couleur en sortie ou dont vous souhaitez obtenir des photographies en noir et blanc (celles-ci nécessitent la numérisation en niveaux de gris). Pour de meilleurs résultats de reconnaissance, utilisez cette option pour un texte noir net sur fond blanc ou clair. La rapidité de numérisation et de traitement,



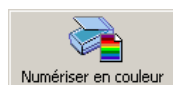
ainsi que le moindre espace disque occupé représentent les deux principaux avantages de ce mode de numérisation par rapport aux deux autres.

Numériser en niveaux de gris



Sélectionnez cette option pour numériser vos documents en niveaux de gris. Ce réglage convient notamment aux photographies en noir et blanc. Pour de meilleurs résultats de reconnaissance, utilisez cette option pour les pages présentant un contraste variable ou réduit (peu de différence entre les zones claires et sombres) et un fond en couleur ou en dégradé.

Numériser en couleur



Sélectionnez cette option pour numériser vos documents en couleur. Ce réglage est réservé aux scanners couleurs. Cette option est particulièrement adaptée si vous souhaitez conserver les graphiques, le texte et les fonds en couleur dans le document de sortie. La numérisation en couleur offre une meilleure précision d'OCR que le mode en niveaux de gris (pour une résolution donnée), mais exige des délais plus importants et une plus grande quantité de mémoire et d'espace disque.

Luminosité et contraste

Les réglages de luminosité et de contraste influent considérablement sur la précision de l'OCR. Vous définissez ces paramètres dans l'onglet Scanner de la boîte de dialogue Options. Le diagramme ci-dessous illustre le réglage de luminosité optimal. Dès le chargement d'une image, vérifiez son aspect. Si les caractères sont épais et collés, augmentez la luminosité ; si les caractères apparaissent fins et partiels, réduisez-la. Une fois les modifications requises effectuées, renumériser la page.



Numériser avec un chargeur ADF

La meilleure méthode pour numériser des documents multipages est de faire appel à un chargeur de feuilles automatique (ADF). Placez les pages dans l'ordre requis dans le chargeur. Si vous souhaitez enregistrer votre document dans plusieurs fichiers de sortie à l'aide de l'option « Un fichier après chaque page blanche », insérez également des pages blanches dans la pile. Voir *Enregistrer le document dans un fichier* dans le chapitre 5.

Si le document à traiter comporte un nombre de pages excédant la capacité du chargeur ADF, sélectionnez l'option « Demander automatiquement la suite des pages » dans l'onglet Traitement de la boîte de dialogue Options. Une boîte de dialogue vous invite à recharger l'ADF et signale la fin de la numérisation.

Un chargeur ADF accepte également la numérisation de pages recto verso. Si votre scanner ne gère pas ce genre d'opération automatiquement, sélectionnez l'option « Numériser pages recto verso » (onglet Scanner de la boîte de dialogue Options). Cette option permet de numériser le document en quelques passes en séparant les pages paires des pages impaires. OmniPage SE se charge de la fusion de l'ensemble des pages.

Numériser des documents volumineux sans chargeur ADF

Même si votre scanner à plat ne dispose pas de chargeur ADF, vous pouvez numériser efficacement des documents multipages. Sélectionnez l'option « Numériser les pages automatiquement » (onglet Scanner de la boîte de dialogue Options) et définissez une valeur de pause en secondes. Ainsi le scanner effectue les passes de numérisation automatiquement, en marquant une pause de durée spécifiée entre chaque numérisation pour vous permettre de charger la page suivante. Une boîte de dialogue vous donne la possibilité de raccourcir ou de prolonger cette pause et de signaler la numérisation de la dernière page.

SPÉCIFIER LA MISE EN PAGE DU DOCUMENT

Avant de commencer la reconnaissance d'un document, OmniPage SE vous demande de décrire la mise en page de celui-ci dans le but de faciliter le traçage de zones automatique. Lorsque vous faites appel à l'assistant OCR, la fonction de découpage en zones automatique est toujours activée. Cette fonction s'applique également en cas de traitement automatique, à moins que vous n'ayez spécifié un gabarit de zones. Lors d'un traitement manuel, le traçage de zones automatique est parfois activé. Pour plus d'informations, consultez l'aide en ligne.

Pour décrire le document original, vous disposez des options suivantes :

Automatique



Choisissez cette option pour instruire le programme de prendre toutes les décisions concernant le découpage en zones. Le programme détermine si le texte est aménagé en colonnes, identifie les graphiques et les éléments textuels, et décide du placement des tableaux. Sélectionnez Automatique si votre document contient des pages présentant des mises en page différentes ou inconnues. Cette option convient également aux pages munies de plusieurs colonnes et d'un tableau, ainsi qu'à toute page comportant plus d'un tableau.

Colonne unique sans tableau



Choisissez cette option si vos pages contiennent une seule colonne de texte et ne comportent aucun tableau. Ce type de mise en page est employé généralement dans les livres et les lettres commerciales. Ce réglage convient également aux pages qui contiennent des mots ou des nombres disposés en colonnes et que vous ne souhaitez pas placer dans un tableau, dans des colonnes séparées ou dans une mise en page sans colonnes. Les graphiques peuvent être détectés.

Colonnes multiples sans tableau



Choisissez cette option si certaines de vos pages contiennent du texte en colonnes que vous souhaitez disposer sans colonnes ou dans des colonnes distinctes (correspondant à la mise en page originale). Les colonnes peuvent être conservées dans le document de sortie avec cadres (option True Page) ou sans cadres (option Conserver l'enchaînement des colonnes). En cas de données tabulaires, celles-ci seront le plus probablement traitées comme une colonne unique. Les graphiques peuvent être détectés.

Colonne unique avec tableau



Choisissez cette option si vos pages contiennent une seule colonne de texte et un tableau. La fonction de traçage de zones automatique n'essaie pas d'identifier des colonnes mais de détecter un tableau et de le placer dans une grille dans l'éditeur de texte. Vous pouvez spécifier par la suite si vous souhaitez

exporter le document sous la forme de grille ou de colonnes de texte séparées par tabulation. Les graphiques peuvent être détectés.

Feuille de calcul



Choisissez cette option si l'intégralité de la page se présente sous la forme d'un tableau que vous souhaitez exporter dans une feuille de calcul ou en tant que tableau unique. Les colonnes uniques et les graphiques ne sont pas détectés.

Personnalisé



Choisissez cette option pour contrôler au maximum le déroulement du traçage de zones automatique. Vous pouvez configurer la détection de colonnes, de graphiques et/ou de tableaux. Effectuez les réglages appropriés dans l'onglet Mise en page personnalisée de la boîte de dialogue Options.

Gabarit



Choisissez un fichier de gabarit de zones si vous souhaitez appliquer ses zones et propriétés à toutes les pages importées. Lors du traitement manuel, les zones de gabarit sont également appliquées à la page actuelle, en remplacement des zones existantes. En plus des zones de gabarit, vous pouvez aussi définir d'autres zones. Pour plus de détails, reportez-vous à la section *Utiliser des gabarits de zone*.

Si le découpage en zones automatique produit des résultats inattendus, utilisez le traitement manuel pour redéfinir les zones sur les pages individuelles et relancer leur reconnaissance.

TRACER DES ZONES MANUELLEMENT



Les zones définissent les sections de la page destinées au traitement. Elles peuvent prendre une forme rectangulaire ou irrégulière (avec les côtés formés par des lignes verticales et horizontales). OmniPage SE n'accepte pas le chevauchement des zones. Chaque zone affiche son numéro dans l'angle supérieur gauche et son icône de type de zone dans l'angle supérieur droit. Pour sélectionner une zone, cliquez dessus ; pour sélectionner plusieurs zones, utilisez la séquence Maj.+clac. Les zones actuelles et sélectionnées sont contrastées. Pour désélectionner une zone, cliquez en dehors de celle-ci. Les zones apparaissent dans la fenêtre Image originale dans les cas suivants :

- ◆ La page a été reconnue.
- ◆ Un fichier de gabarit de zones a été spécifié pour le traitement manuel lorsque la page était actuelle.
- ◆ Vous avez tracé des zones manuelles sur l'image.

Gérer les zones

La barre d'outils Image propose des outils d'édition des zones. Un des ces outils est toujours sélectionné ; pour modifier la sélection en cours, cliquez sur un autre outil. L'outil activé par défaut est généralement l'outil Tracer des zones rectangulaires.

Tracer des zones rectangulaires



Cliquez sur cet outil et faites glisser le curseur dans la direction souhaitée pour définir des zones rectangulaires. Les nouvelles zones adoptent les propriétés de la dernière zone définie ou sélectionnée. Lorsque cet outil est actif, vous pouvez également déplacer ou redimensionner les zones existantes.

Tracer des zones irrégulières



Utilisez cet outil pour définir des zones irrégulières. Cliquez et tracez une ligne, puis répétez cette opération jusqu'à ce que tous les côtés de la zone soient tracés, sauf un. Pour fermer la zone, cliquez deux fois dessus. Les zones irrégulières adoptent une forme rectangulaire si vous leur affectez le type Tableau. Lorsque cet outil est actif, vous pouvez également déplacer ou redimensionner les zones existantes.

Ajouter à la zone



Cliquez sur cet outil pour agrandir une zone irrégulière ou fusionner plusieurs zones. Lorsque cet outil est actif, vous ne pouvez ni déplacer ni redimensionner les zones existantes. Vous ne pouvez pas non plus appliquer cet outil à une zone de type Tableau.

Soustraire de la zone



Cliquez sur cet outil pour supprimer des sections irrégulières d'une zone ou subdiviser une zone. Lorsque cet outil est actif, vous ne pouvez ni déplacer ni redimensionner les zones existantes. Vous ne pouvez pas non plus appliquer cet outil à une zone de type Tableau.

Réagencer les zones



Utilisez cet outil pour modifier l'ordre de reconnaissance des zones. Cliquez sur les zones dans l'ordre de reconnaissance souhaité. Pour que l'ordre défini soit respecté, évitez de tracer des zones de type Colonnes multiples et Automatique et sélectionnez l'option « Utiliser uniquement les zones actuelles ».

Propriétés de zone



Cliquez sur cet outil pour accéder à la boîte de dialogue Propriétés de zone, qui permet de définir le type et le contenu des zones sélectionnées sur la page. Cette opération est également possible à partir du menu contextuel de la zone. Pour plus d'informations, reportez-vous à la section suivante.

Propriétés de zone

Chaque zone est affectée d'un type, et les zones comportant du texte peuvent également présenter un paramètre de contenu de zone : alphanumérique ou numérique. Le type et le contenu d'une zone représentent les « propriétés de la zone ». Pour modifier les propriétés d'une zone, cliquez dessus avec le bouton droit pour afficher son menu contextuel et effectuer les sélections souhaitées. Pour changer les propriétés de plusieurs zones simultanément, sélectionnez le groupe de zones requis avant de choisir les nouvelles valeurs. Cette opération est également disponible via le bouton de propriétés de zone de la barre d'outils Image. Les zones peuvent adopter les types suivants :



Colonne unique

Utilisez ce type pour traiter l'ensemble du contenu de la zone en tant que colonne unique.



Colonnes multiples

Utilisez ce type pour traiter le contenu de la zone en tant que colonnes multiples. Le programme essaie de détecter des colonnes à l'intérieur de la zone. Le texte est extrait des colonnes ou conservé dans celles-ci, en fonction des propriétés de l'éditeur de texte. Lors de la reconnaissance, une zone de type Colonnes multiples peut être remplacée par des zones distinctes pour chaque colonne. Pour effectuer cette opération, vous devez activer le découpage en zones automatique, qui peut toutefois modifier l'ordre de reconnaissance des zones.



Tableau

Utilisez ce type pour traiter le contenu de la zone en tant que tableau. Les grilles du tableau peuvent être détectées automatiquement ou placées manuellement (voir la section suivante). Les zones de type Tableau doivent être rectangulaires. L'éditeur de texte affiche le tableau dans une grille éditables. Vous pouvez spécifier si vous souhaitez exporter le tableau sous la forme de grille ou de colonnes de texte séparées par tabulation.



Automatique

Utilisez ce type pour laisser le soin au programme de définir le type de zone. Pour effectuer cette opération, vous devez activer le découpage en zones automatique, qui peut toutefois modifier l'ordre de reconnaissance des zones. À l'issue de la reconnaissance, vous pouvez vérifier le type appliqué. Si vous utilisez la fonction de découpage en zones automatique pour traiter une page comportant divers contenus, le programme peut remplacer la zone automatique par un certain nombre de zones plus réduites.



Graphique

Utilisez ce type pour entourer une illustration, un diagramme, un dessin, une signature et tout autre élément que vous souhaitez transmettre à l'éditeur de

texte en tant qu'image incorporée plutôt que texte reconnu. Les zones graphiques se distinguent par une bordure verte. Les images incorporées peuvent être exportées avec le document vers les applications cibles supportant les graphiques.



Ignorer la zone

Utilisez ce type pour délimiter une section que vous ne souhaitez pas traiter dans l'éditeur de texte. La fonction de découpage en zones automatique omet ces zones pendant le traitement. Pour exclure une section donnée sur plusieurs pages (l'en-tête ou les numéros de page, par exemple), placez le type Ignorer la zone dans un gabarit et sélectionnez « Tracer des zones en plus de celles actuelles ou du gabarit » dans l'onglet Traitement de la boîte de dialogue Options.

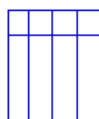


Contenu de zone

Cette propriété est disponible pour les types de zone contenant du texte. Le contenu alphanumérique valide tous les caractères nécessaires pour votre sélection linguistique. Le résultat de la reconnaissance d'une zone numérique comporte uniquement des nombres et la ponctuation y afférent. Toute lettre est ignorée.

Remarque : Pour afficher le menu contextuel correspondant à l'ensemble de l'image, cliquez avec le bouton droit en dehors de toute zone. Les commandes de ce menu permettent d'effectuer un zoom avant et arrière, et de faire pivoter l'image. La rotation d'une image supprime toutes ses zones tracées.

GÉRER LES GRILLES DU TABLEAU DANS L'IMAGE



À l'issue d'un traitement en mode automatique, OmniPage SE peut placer des zones de type Tableau sur la page traitée. Ces zones comportent une icône de zone Tableau dans leur angle supérieur droit. Vous pouvez modifier le type affecté à une zone via son menu contextuel. Vous pouvez également tracer une zone de type Tableau. Si une zone de type Tableau est présente sur la page, sélectionnez-la et tracez la nouvelle zone rectangulaire, qui héritera du type Tableau. Vous pouvez également tracer une zone rectangulaire puis utiliser son menu contextuel pour lui affecter le type Tableau.

La création ou le déplacement des lignes de séparation verticales et horizontales d'un tableau permet de déterminer l'emplacement de la grille lors du placement du tableau dans l'éditeur de texte. Vous pouvez utiliser les outils

Ajouter à la zone ou Soustraire de la zone pour agrandir ou réduire une zone Tableau, mais celle-ci doit conserver sa forme rectangulaire. Les outils mentionnés permettent de supprimer les colonnes ou les rangées superflues.

Chaque fois que la page actuelle comporte une zone de type Tableau, les cinq outils de gestion de tableau de la barre d'outils Image deviennent actifs. Utilisez-les pour effectuer les opérations suivantes :

Déplacer séparation



Cliquez sur cet outil et placez le curseur sur la séparation de rangées ou de colonnes requise. Le curseur prend la forme d'une flèche double. Faites glisser la séparation dans la direction souhaitée (sans dépasser les limites des séparations voisines). Évitez de placer des séparations qui se chevauchent ou qui divisent un texte. Pour appliquer l'opération de déplacement uniquement à la rangée actuelle, maintenez enfoncée la touche Ctrl lors de la manipulation de la séparation de colonnes concernée.

Insérer séparation de colonnes



Cliquez sur cet outil puis sur l'emplacement dans la zone Tableau où vous souhaitez placer une séparation de colonnes. Pour appliquer cette opération uniquement à la rangée actuelle, maintenez enfoncée la touche Ctrl lors du placement de la séparation.

Insérer séparation de rangées



Cliquez sur cet outil puis sur l'emplacement dans la zone Tableau où vous souhaitez placer une séparation de rangées. Évitez de placer des séparations qui se chevauchent ou qui divisent un texte.

Supprimer séparation



Cliquez sur cet outil puis sur la séparation individuelle à supprimer. Cette opération est utile pour éliminer une séparation mal placée ou pour modifier l'apparence du tableau dans le document final. Ainsi, vous pouvez regrouper deux colonnes de données en une seule en supprimant leur trait de séparation.

Supprimer/Remplacer toutes séparations



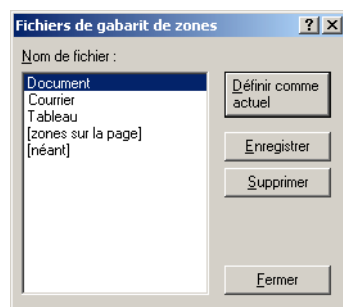
Cliquez sur cet outil, puis à l'intérieur d'une zone Tableau pour masquer toutes les séparations du tableau. Cliquez à nouveau dans le tableau pour (re)détecter les séparations automatiquement. Le placement de séparations a normalement lieu lors de la reconnaissance. Pour afficher et éditer les séparations avant l'OCR, cliquez deux fois sur cet outil.

UTILISER DES GABARITS DE ZONE

Un gabarit regroupe dans un fichier un ensemble de zones, leurs propriétés et les données définissant leur ordre de reconnaissance. Pour appliquer les zones d'un gabarit lors de la reconnaissance, vous chargez un fichier de gabarit de zone via la liste déroulante du bouton Exécuter l'OCR ou à l'aide du menu Outils.

Lorsque vous chargez un gabarit par l'intermédiaire de la **barre d'outils OCR manuel**, ses zones sont immédiatement reflétées sur la page actuelle, en remplacement des zones tracées (les pages existantes ne sont pas affectées). Le gabarit de zones est appliqué à toutes les pages importées jusqu'à son déchargement. Avant de procéder à l'OCR, vous pouvez modifier le gabarit de zones et ajouter de nouvelles zones.

Lorsque vous chargez un gabarit par l'intermédiaire de la **barre d'outils AutoOCR**, celui-ci n'affecte ni la page actuelle ni les pages existantes. Le gabarit de zones est appliqué à toutes les pages suivantes importées jusqu'au déchargement du gabarit. L'onglet Traitement de la boîte de dialogue Options propose l'option « Tracer des zones en plus de celles actuelles ou du gabarit ». Si celle-ci est activée lors d'un traitement automatique, le découpage en zones automatique porte sur les sections de la page non couvertes par le gabarit appliqué.



Enregistrer un gabarit de zones

Tracez des zones sur une page. Vérifiez leurs emplacement, propriétés et ordre de reconnaissance. Cliquez sur Fichiers de gabarit de zones dans le menu Outils. Dans la boîte de dialogue affichée, sélectionnez [zones sur la page] et cliquez sur Enregistrer.

Modifier un gabarit de zone

Chargez un gabarit et importez une image au moyen d'un traitement manuel. Les zones de gabarit apparaissent sur l'image. Modifiez les zones et/ou les propriétés selon vos besoins. Ouvrez la boîte de dialogue Fichiers de gabarit de zones. Le gabarit actuel est sélectionné. Cliquez sur Enregistrer puis sur Fermer.

Décharger un gabarit de zones

Sélectionnez une option non associée à un gabarit pour décrire la mise en page du document dans la liste déroulante du bouton Exécuter l'OCR. Les zones du gabarit ne sont pas supprimées de la page actuelle ou des pages existantes, mais le gabarit n'est plus configuré pour utilisation lors des futurs

traitements. Vous pouvez également ouvrir la boîte de dialogue Fichiers de gabarit de zones, sélectionner [néant] et cliquer sur le bouton Définir comme actuel. Dans ce cas, l'option de description de la mise en page reprend la valeur Automatique.

Remplacer un gabarit de zones

Sélectionnez un autre gabarit dans la liste déroulante du bouton Exécuter l'OCR, ou ouvrez la boîte de dialogue Fichiers de gabarit de zones, sélectionnez le gabarit souhaité et cliquez sur le bouton Définir comme actuel. Lorsque la barre d'outils AutoOCR est active, les zones existantes ne sont pas modifiées et le nouveau gabarit s'applique uniquement aux futurs traitements. Lorsque la barre d'outils OCR manuel est active, les zones du nouveau gabarit sont appliquées à la page actuelle, en remplacement des zones existantes.

Supprimer un gabarit de zones

Ouvrez la boîte de dialogue Fichiers de gabarit de zones. Sélectionnez un gabarit et cliquez sur le bouton Supprimer. Notez que les zones déjà créées à l'aide de ce gabarit ne sont pas supprimées.

Astuce : Les gabarits de zones acceptent les zones de type Ignorer la zone et Automatique. Par conséquent, un gabarit peut être très utile pour définir les parties de la page destinées au traitement et les parties à ignorer.

Remarque : Lors de la reconnaissance, le programme peut remplacer les zones automatiques provenant d'un gabarit par un certain nombre de zones plus réduites ; des types de zones spécifiques sont assignés à ces zones. Les zones à colonnes multiples peuvent également être divisées en zones plus réduites à colonne unique, leur nombre correspondant au nombre de colonnes détecté.

Remarque : Les gabarits et la fonction supplémentaire de découpage en zones automatique sont disponibles dans Planifier OCR et Direct OCR, mais pas dans l'Assistant OCR.



4 Vérification et édition

Les résultats de la reconnaissance sont placés dans l'éditeur de texte. Ce nouveau module aux capacités d'affichage « tel quel » offre les fonctions énumérées ci-dessous et décrites en détail dans le présent chapitre :

- ◆ Vérifier les résultats de l'OCR
- ◆ Comparer le texte reconnu à l'original
- ◆ Dictionnaires utilisateur
- ◆ IntelliTrain
- ◆ Modes d'affichage de l'éditeur de texte
- ◆ Éditer du texte et des images
- ◆ Lire le texte à voix haute
- ◆ Ébauche de page

L'éditeur de texte offre quatre modes d'affichage de pages. Vous pouvez à tout moment passer d'un mode à un autre pour sélectionner le niveau de formatage requis :



Sans Formatage

Pour afficher le texte en une seule colonne dans un seul style et corps de police.



Conserver le formatage polices/paragraphes

Pour afficher le texte en une seule colonne dans plusieurs styles de police et de paragraphe.



True Page

Pour reproduire au maximum le formatage du document original. Les styles de caractère et de paragraphe sont conservés et tous les éléments de page, y compris les colonnes, sont placés dans des cadres.



Conserver l'enchaînement des colonnes

Mode quasiment identique au mode True Page, à cette exception près que l'ordre de la reconnaissance des zones est indiqué par des flèches. Cette particularité influe surtout sur l'exportation des documents (pour plus d'informations, reportez-vous à la section *Préparer les résultats de l'OCR pour l'exportation* du chapitre 5).

VÉRIFIER LES RÉSULTATS DE L'OCR

À l'issue de la reconnaissance d'une page, les résultats de l'opération apparaissent dans l'éditeur de texte. La vérification démarre automatiquement si cette option a été activée dans l'onglet Vérification de la boîte de dialogue Options ou par l'intermédiaire de l'Assistant OCR. Vous pouvez également lancer l'OCR manuellement dès que le programme n'est pas occupé. Procédez comme suit :



1. Cliquez sur le bouton Vérifier l'OCR dans la barre d'outils Standard, ou choisissez Vérifier l'OCR dans le menu Outils.
2. La vérification démarre au début du document, mais omet le texte déjà vérifié. Si un mot douteux est détecté, la boîte de dialogue Vérification de l'OCR affiche le texte erroné et l'image originale correspondante.

Raison pour laquelle le mot est considéré douteux.

Fenêtre présentant le fragment correspondant de l'image originale. Cliquez dedans pour agrandir ou réduire l'affichage.

Mot qu'OmniPage SE n'arrive pas à reconnaître.

L'image du mot douteux est contrastée.

Pour redimensionner la boîte de dialogue, faites glisser un angle ou le bord inférieur.

3. Si le mot reconnu est correct, cliquez sur Ignorer ou Ignorer tout pour passer au mot douteux suivant. Cliquez sur Ajouter pour l'ajouter au dictionnaire utilisateur actuel et passer au mot douteux suivant.
4. Si le mot reconnu est erroné, tapez la version correcte dans la zone Remplacer par ou sélectionnez une suggestion de dictionnaire dans la zone Suggestions. Cliquez sur Remplacer ou Remplacer tout pour implémenter votre correction et passer au mot douteux suivant. Cliquez sur Ajouter pour intégrer au dictionnaire utilisateur actuel le mot affiché dans la zone Remplacer par et pour passer au mot douteux suivant.

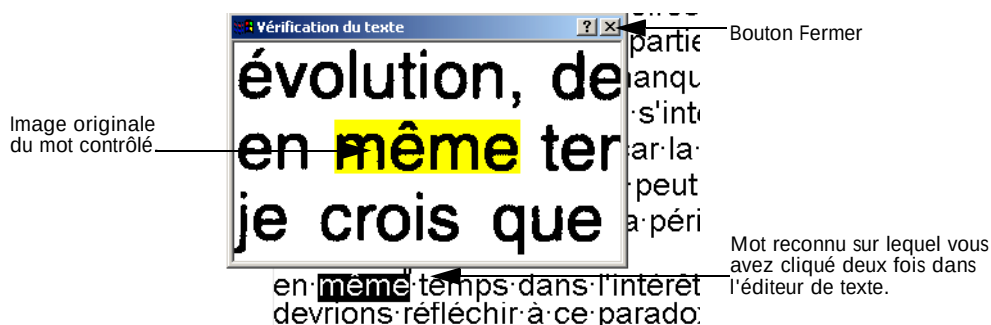
5. Au fur et à mesure que les mots douteux affichés dans l'éditeur de texte sont vérifiés, les marques de couleur les accompagnant disparaissent. Lors de la vérification, vous pouvez passer à la fenêtre de l'éditeur de texte pour y effectuer les corrections requises. Pour relancer la vérification, utilisez le bouton Reprendre. Pour arrêter la vérification avant d'atteindre la fin du document, cliquez sur Fermer.

Remarque : Toute page entièrement vérifiée porte les icônes  et  dans le gestionnaire de document.

COMPARER LE TEXTE RECONNU À L'ORIGINAL

Une fois la reconnaissance d'une page effectuée, vous pouvez comparer tout élément du texte reconnu avec la partie correspondante de l'image originale afin de contrôler la précision de l'opération. Procédez comme suit :

1. Cliquez deux fois sur un mot dans l'éditeur de texte ou sélectionnez un mot et choisissez Vérifier le texte dans le menu Outils. La fenêtre Vérification du texte apparaît et affiche l'image comportant le mot original et le texte adjacent. Modifiez le mot dans l'éditeur selon vos besoins.
2. Pour agrandir ou réduire l'affichage, cliquez à l'intérieur de la fenêtre. L'image est agrandie sur les deux premiers clics et réduite sur les deux suivants.



3. Continuez à cliquer deux fois sur les mots à contrôler et corrigez-les selon vos besoins. L'affichage est modifié en fonction des nouveaux mots sélectionnés.
4. Cliquez sur le bouton Fermer pour quitter la fenêtre de comparaison.

Astuce : Il est conseillé de vérifier l'OCR et contrôler le texte reconnu avant de procéder à une édition importante. Si vous coupez et collez de grands fragments de texte, les liens entre les mots et l'image originale peuvent être perdus.



Astuce : Le module Texte-parole d'OmniPage Pro 11 vous permet de vérifier le texte reconnu à l'aide d'une lecture à voix haute. La lecture peut s'effectuer sur les lettres, les mots, les lignes, les phrases ou sur l'ensemble de la page. Reportez-vous à la section *Lire le texte à voix haute*. Veuillez noter que ce module n'est pas disponible dans OmniPage SE.

DICTIONNAIRES UTILISATEUR

OmniPage SE intègre des dictionnaires en plusieurs langues, qui facilitent la reconnaissance et peuvent proposer des suggestions pendant la vérification de l'OCR. Les dictionnaires intégrés peuvent être remplacés par les dictionnaires créés par l'utilisateur. Vous pouvez enregistrer plusieurs dictionnaires utilisateur, mais ne pouvez charger qu'un seul dictionnaire à la fois. Le programme propose aussi vos dictionnaires utilisateur de Microsoft Word : un dictionnaire personnalisé et le dictionnaire utilisateur Word par défaut.

Ouvrir un dictionnaire utilisateur

Cliquez sur Ajouter dans la boîte de dialogue Vérification de l'OCR (aucun dictionnaire utilisateur n'étant chargé), ou ouvrez la boîte de dialogue Fichiers de dictionnaire utilisateur à partir du menu Outils et cliquez sur Nouveau. Un message vous demande de nommer le dictionnaire.

Charger ou décharger un dictionnaire utilisateur

Réalisez ces opérations depuis l'onglet OCR de la boîte de dialogue Options ou à partir de la boîte de dialogue Fichiers de dictionnaire utilisateur. Pour le chargement, sélectionnez un fichier dictionnaire ; pour le déchargement, sélectionnez [néant].

Éditer un dictionnaire utilisateur

Pour ajouter des mots à un dictionnaire utilisateur, chargez-le et cliquez sur Ajouter dans la boîte de dialogue Vérification de l'OCR. Vous pouvez ajouter et supprimer des mots via la boîte de dialogue Fichiers de dictionnaire utilisateur.

Astuce : Lors de l'édition d'un dictionnaire utilisateur, vous pouvez optimiser l'opération d'ajout de nouveaux mots en important la liste des mots requis à partir d'un fichier texte.

INTELLITRAIN

IntelliTrain représente une nouvelle forme d'apprentissage automatisé, capable de mémoriser les corrections effectuées par l'utilisateur lors de la vérification de l'OCR. Lorsque vous effectuez une modification, IntelliTrain retient la forme du caractère concernée et la correction apportée. Par la suite le module parcourt le document à la recherche des formes de caractère similaires, notamment parmi les mots douteux, et évalue la pertinence de la correction de l'utilisateur.



Le module IntelliTrain et les fichiers d'apprentissage ne sont pas supportés par OmniPage SE. Ces fonctionnalités sont uniquement disponibles dans OmniPage Pro 11. Toutes les données d'apprentissage pouvant accompagner un fichier OPD seront ignorées lors de son ouverture dans OmniPage SE.

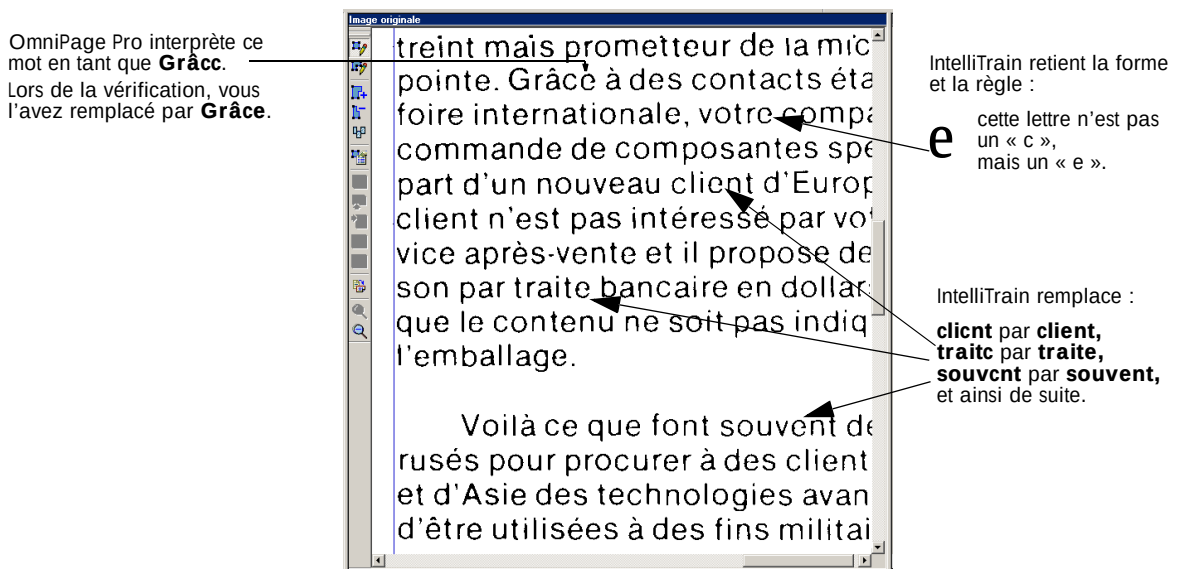
Le module IntelliTrain peut être activé et désactivé via l'onglet OCR de la boîte de dialogue Options. Cette fonction est utile pour les documents présentant une dégradation uniforme ou un type de police inhabituel dans la totalité de leur contenu. Elle ne produit pas d'aussi bons résultats dans les documents présentant des déformations ponctuelles et variées. L'exemple ci-dessous s'appuie sur les différentes impressions de la lettre « g » :

Rugged...Rugged...**Rugged**..Rugged

Les deux premiers exemples n'exigent pas IntelliTrain, car les formes de la lettre « g » utilisées sont normales et le programme peut les traiter sans problème.

Le troisième exemple pourrait bénéficier d'IntelliTrain, parce que la forme de la lettre « g » utilisée est inhabituelle et que toutes les instances de celle-ci dans le texte auront probablement le même aspect. Le quatrième exemple ne convient pas à l'utilisation d'IntelliTrain étant donné que la première instance de la lettre « g » est partiellement imprimée et qu'il est peu probable que cette forme se répète dans le document.

L'illustration ci-dessous explique le fonctionnement d'IntelliTrain. Cette exemple porte sur les lettres **c** et **e**. Avec certaines polices et dans certaines conditions de numérisation, la ligne horizontale de la lettre e risque de devenir très fine et d'entraîner des erreurs de reconnaissance, qu'IntelliTrain se chargera de corriger.



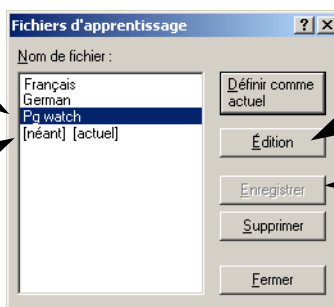
IntelliTrain conserve les données d'apprentissage générées par l'opération ; vous pouvez les enregistrer dans un fichier d'apprentissage pour réutilisation avec des documents similaires. Si vous souhaitez que le programme vous demande d'enregistrer les données d'apprentissage non sauvegardées à la fermeture d'un document, sélectionnez cette option dans l'onglet Vérification de la boîte de dialogue Options. Les données d'apprentissage, qu'elles soient enregistrées ou non, sont stockées dans un document OmniPage.

Vous pouvez enregistrer des données d'apprentissage dans un fichier, ainsi que charger, éditer et télécharger des fichiers d'apprentissage via la boîte de

dialogue Fichiers d'apprentissage. Ouvrez cette fenêtre à partir de l'onglet Vérification de la boîte de dialogue Options ou depuis le menu Outils.

Sélectionnez cette entrée, cliquez sur Enregistrer et tapez un nom pour créer un fichier d'apprentissage.

Sélectionnez cette entrée pour télécharger un fichier d'apprentissage.



Cliquez sur ce bouton pour éditer le fichier d'apprentissage sélectionné (voir ci-dessous).

Ce bouton permet aussi d'enregistrer de nouvelles données d'apprentissage dans le fichier chargé. Le fichier adopte le nom : Nom de fichier [modifié].

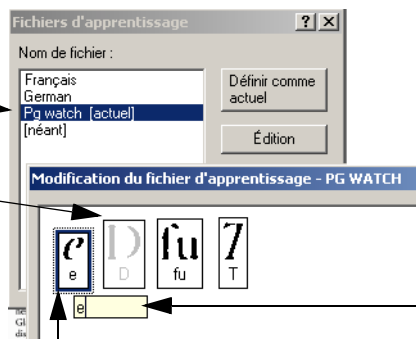
Des données d'apprentissage non enregistrées peuvent être éditées dans la boîte de dialogue Modification du fichier d'apprentissage ; un astérisque s'affiche dans la barre de titre à la place du nom de fichier. Les données ne sont pas sauvegardées à la fermeture de la boîte de dialogue.

Vous pouvez également éditer un fichier d'apprentissage ; dans ce cas, son nom apparaît dans la barre de titre. Si des données d'apprentissage non enregistrées ont été ajoutées au fichier, son nom est accompagné d'un astérisque. Les données d'apprentissage non enregistrées et modifiées sont sauvegardées à la fermeture de la boîte de dialogue.

La boîte de dialogue affiche des cadres comportant une forme de caractère et la solution d'OCR qui lui est assignée. Cliquez sur un cadre pour le sélectionner. Vous pouvez ensuite le supprimer à l'aide de la touche Suppr ou modifier la solution affectée. Pour passer au cadre précédent ou suivant, utilisez les touches flèches.

Vous éditez des données d'apprentissage non enregistrées.

Ce cadre est grisé car il a été supprimé. Pour annuler sa suppression, resélectionnez-le et appuyez sur la touche Suppr. Les caractères marqués en tant que supprimés ne sont effacés qu'à la fermeture de la boîte de dialogue.



Cliquez deux fois sur un cadre ou appuyez sur Entrée pour modifier la solution d'OCR associée. Tapez la nouvelle solution dans la zone de texte qui s'affiche et appuyez sur Entrée. Les assignations modifiées apparaissent en rouge.

Ce cadre est sélectionné. La partie supérieure affiche la forme provenant de l'image. La partie inférieure affiche la solution d'OCR assignée.

MODES D’AFFICHAGE DE L’ÉDITEUR DE TEXTE

L’éditeur affiche le texte reconnu et, en fonction de vos réglages, signale les mots qui ont été considérés comme douteux pendant la reconnaissance. Les mots douteux sont soulignés par une ligne ondulée : la couleur rouge dénote un mot absent du dictionnaire (uniquement pour les langues disposant d’un dictionnaire), alors que la couleur bleue indique un mot comportant un caractère douteux ou de rejet. Ces marques peuvent être affichées ou masquées selon les options sélectionnées dans l’onglet Éditeur de texte de la boîte de dialogue Options. Vous pouvez également afficher ou masquer les caractères non imprimables et les indicateurs d’en-tête et de bas de page. L’onglet Éditeur de texte permet aussi de définir l’unité de mesure du programme et de configurer le retour à la ligne automatique dans tous les modes de l’éditeur de texte, à l’exception du mode Sans Formatage.

Les modes de l’éditeur disponibles présentent les caractéristiques suivantes :

Sans Formatage



Pour afficher le texte en une seule colonne alignée à gauche et appliquer un seul style et corps de police, tout en conservant les sauts de ligne de l’original. La plupart des boutons et des boîtes de dialogue de formatage sont désactivés. Les règles ne sont pas affichées. Ce mode s’avère généralement utile pour la vérification et l’édition du texte.

Conserver le formatage polices/paragraphes



Pour afficher le texte en une seule colonne et appliquer des styles de police et de paragraphe. La règle horizontale est affichée. Ce mode s’avère généralement utile pour la vérification et l’édition du texte et des styles utilisés.

True Page



Pour conserver au maximum le formatage du document original. Les styles de caractère et de paragraphe sont conservés et tous les éléments de page, y compris les colonnes, sont placés dans des cadres. La vérification et l’édition du texte peuvent être moins aisées dans ce mode, en effet il peut être nécessaire de faire défiler le contenu d’un cadre pour accéder à tous ses éléments. Un contenu dépassant les limites du cadre est signalé par une série de flèches.



Conserver l’enchaînement des colonnes



Mode quasiment identique au mode True Page, à cette exception près que l’ordre de la reconnaissance des zones est indiqué par des flèches. Cette particularité influe surtout sur l’exportation des documents (pour plus d’informations, reportez-vous à la section *Préparer les résultats de l’OCR pour l’exportation* du chapitre 5).

Vous sélectionnez le mode requis à l'aide des quatre boutons présents dans l'angle inférieur gauche de l'éditeur de texte ou dans le menu Affichage. Les graphiques et les tableaux sont compatibles avec les quatre modes disponibles.

ÉDITER DU TEXTE ET DES IMAGES

L'éditeur de texte est doté de capacités d'affichage « tel quel », et ses fonctions d'édition sont très proches des possibilités offertes par les applications de traitement de texte les plus populaires.

Édition de caractères

Les modes de l'éditeur autres que le mode Sans Formatage permettent de modifier le type, le corps et les attributs (gras, italique, souligné) de la police appliqués au texte sélectionné. Utilisez la barre d'outils de formatage ou la boîte de dialogue Police accessible par le biais du menu Format. (Cette boîte de dialogue offre également les options d'indice, d'exposant et de texte/fond de couleur.)

En mode Sans Formatage vous pouvez utiliser la barre d'outils de formatage pour spécifier un seul type et une seule taille de police à appliquer à l'ensemble du document. Ce réglage n'est pas transmis aux autres modes ; ceux-ci gardent leur anciens paramètres.

Ouvrez la boîte de dialogue Mappage des polices (via l'onglet OCR de la boîte de dialogue Options) pour spécifier les polices à utiliser dans les textes affichés dans l'éditeur de texte.

Édition de paragraphes

Les modes de l'éditeur autres que le mode Sans Formatage permettent de modifier l'alignement des paragraphes sélectionnés et de les formater avec des puces. Utilisez la barre d'outils de formatage ou la boîte de dialogue Paragraphe accessible depuis le menu Format. Cette boîte de dialogue permet de modifier le retrait, l'interligne et l'espacement des paragraphes. Grâce à la règle horizontale de l'éditeur de texte, vous pouvez définir les emplacements de retrait et de tabulation en toute facilité. Les positions de tabulation évoluées peuvent être spécifiées par le biais de la boîte de dialogue Tabulations (accessible depuis le menu Format).

Styles de paragraphe

Les styles de paragraphe présents dans un document sont détectés automatiquement pendant la reconnaissance. Une liste des styles est créée et présentée dans une boîte de sélection dans la partie gauche de la barre

d'outils de formatage. Cette liste permet d'assigner un style aux paragraphes sélectionnés. Pour renommer/modifier un style ou pour définir un nouveau style, utilisez la boîte de dialogue Style (accessible depuis le menu Format). Lors de l'enregistrement d'un document dans un fichier, vous décidez si vous souhaitez exporter également ses styles de paragraphe. Cette option est disponible uniquement si l'application cible supporte les styles de paragraphe.

Graphiques

Pour éditer des graphiques dans OmniPage SE, votre ordinateur doit être équipé d'un éditeur d'image. Choisissez la commande Éditer l'image dans le menu Outils. Cette sélection active l'éditeur d'image associé au format BMP sur votre système Windows et charge le graphique. Éditez le graphique, puis fermez ce programme pour le réintégrer dans l'éditeur de texte d'OmniPage SE. Ne modifiez pas la taille, la résolution ou le type de graphique, car ces actions empêchent la réintégration.

Tableaux

Les tableaux s'affichent dans l'éditeur de texte sous forme de grilles. Pour modifier un tableau, placez le curseur dans une zone de type Tableau ; il change d'aspect et vous pouvez déplacer les traits de la grille. Vous pouvez également utiliser les règles de l'éditeur de texte : modifiez l'emplacement du texte dans les cellules d'un tableau à l'aide des boutons d'alignement de la barre d'outils de formatage et des commandes de tabulation de la règle. Lors de l'enregistrement du document dans un fichier, vous pouvez décider si vous souhaitez exporter les tableaux en tant que grilles ou colonnes séparées par des tabulations.

LIRE LE TEXTE À VOIX HAUTE



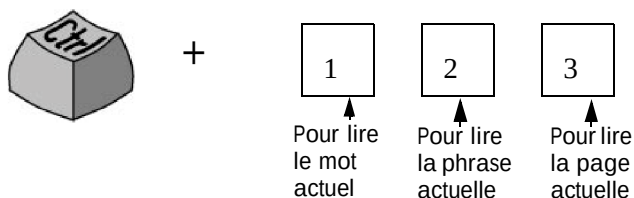
La fonction Texte-parole est activée ou désactivée via la commande Mode texte-parole du menu Outils ou via la touche F5. Le sous-menu Réglages texte-parole permet de sélectionner le type de voix (pour exemple, masculine ou féminine dans une langue donnée) et de configurer le débit et le volume pour la lecture. Le module Texte-parole n'est pas disponible dans OmniPage SE. Ce module n'est disponible que dans OmniPage Pro 11

La fonction de lecture du texte reconnu est avant tout destinée aux personnes malvoyantes ; toutefois, elle peut également faciliter les processus de vérification de l'OCR et de contrôle du texte reconnu à tout utilisateur d'OmniPage Pro. La lecture à voix haute est contrôlée par les mouvements du point d'insertion dans l'éditeur de texte (via la souris ou le clavier).

Modalités de lecture	Touches utilisées
Un seul caractère à la fois, lecture avant ou arrière	Flèche Droite ou Gauche. Les noms de lettre, de nombre et de signe de ponctuation sont prononcés.
Mot actuel	Ctrl + Un du bloc numérique (1)
Mot suivant vers la droite	Ctrl + flèche Droite *
Mot suivant vers la gauche	Ctrl + flèche Gauche *
Une ligne	Placez le point d'insertion dans la ligne
Ligne suivante	Flèche Bas
Ligne précédente	Flèche Haut
Phrase actuelle	Ctrl + Deux du bloc numérique (2)
Depuis le point d'insertion jusqu'à la fin de la phrase	Ctrl + Six du bloc numérique (6)
Depuis le début de la phrase jusqu'au point d'insertion	Ctrl + Quatre du bloc numérique (4)
Page actuelle	Ctrl + Trois du bloc numérique (3)
Depuis le début de la page actuelle jusqu'au point d'insertion	Ctrl + Début
Depuis le point d'insertion jusqu'à la fin de la page actuelle	Ctrl + Fin
Page précédente, suivante ou la page de votre choix	PgPrec, PgSuiv ou boutons de navigation
Caractères saisis	Chaque caractère saisi est prononcé individuellement, y compris les signes de ponctuation.

* Si le curseur est placé au milieu d'un mot, le programme prononce d'abord un fragment du mot, puis sur un deuxième appui sur la touche, le mot dans son intégralité.

Les trois touches de base de réglage texte-parole sont regroupées sur le block numérique.



Vous disposez également des séquences clavier suivantes :

Action requise	Séquence utilisée
<i>Interrompre/Reprendre</i>	Ctrl + Cinq du bloc numérique 5
<i>Augmenter le débit</i>	Ctrl + Signe plus du bloc numérique +
<i>Décroître le débit</i>	Ctrl + Signe moins du bloc numérique -
<i>Rétablir le débit</i>	Ctrl + astérisque du bloc numérique *

ScanSoft a l'intention d'offrir la fonction Texte-parole dans les versions linguistiques suivantes : anglais, français, allemand, italien, portugais et espagnol. Pour obtenir les dernières informations à ce sujet, consultez le fichier ReadMe/LisezMoi. Une seule version linguistique est installée avec OmniPage Pro, en fonction de la sélection effectuée au lancement de l'installation. Si vous spécifiez une langue ne disposant pas de la fonction Texte-parole, la version anglaise est installée.

Si votre ordinateur est muni d'autres systèmes de conversion texte-parole compatibles avec la norme SAPI, ceux-ci sont détectés et mis à votre disposition. Leurs voix sont proposées dans la boîte de dialogue Réglages texte-parole. Dès qu'une voix a été affectée à une langue donnée, OmniPage Pro retient cette sélection et passe à la voix indiquée en fonction de la langue de reconnaissance spécifiée pour votre document.

ÉBAUCHE DE PAGE

La fenêtre Ébauche de page permet de modifier l'ordre de reconnaissance des zones sur une page ou des paragraphes à l'intérieur de celles-ci. Vous pouvez également y définir la disposition du texte pour l'exportation en mode Conserver l'enchaînement des colonnes. Ouvrez la fenêtre Ébauche de page à partir du menu Affichage. Les sections affichées correspondent aux zones appliquées pendant la reconnaissance et aux cadres utilisés dans l'éditeur de texte. Cliquez sur un élément et faites-le glisser à l'emplacement voulu. Les paragraphes réordonnés s'affichent immédiatement dans l'éditeur de texte et sont exportés. Les zones réordonnées apparaissent et sont exportées en modes Sans Formatage et Conserver le formatage polices/paragraphes. Elles n'ont aucun impact pratique en mode True Page. En mode Conserver l'enchaînement des colonnes, une série de flèches indique l'ordre de la disposition du texte. Déplacez les zones pour modifier cet ordre. Les positions des zones ne changent pas, mais les flèches reflètent la nouvelle configuration.



5 Enregistrement et exportation

Après avoir importé au moins une image dans un document, vous pouvez procéder à l'exportation des images acquises vers un fichier. À l'issue de la reconnaissance d'au moins une page, vous pouvez exporter les résultats de l'OCR vers l'application cible de votre choix à l'aide d'une des méthodes suivantes :

1. Enregistrer le document dans un fichier
2. Copier un document dans le Presse-papiers
3. Joindre un document à un e-mail

Le document étant conservé dans OmniPage SE après l'exportation, vous pouvez l'enregistrer, le copier ou l'envoyer à plusieurs reprises en appliquant, le cas échéant, différents niveaux de formatage, types de fichiers, noms ou emplacements. Vous pouvez également y ajouter de nouvelles pages, relancer l'OCR des pages traitées ou modifier le texte reconnu.

Dans le cas du traitement à l'aide de l'Assistant OCR et en mode automatique, vous spécifiez la première destination d'enregistrement avant d'entamer le processus. La reconnaissance (ou la vérification, selon vos réglages) de la dernière page disponible est suivie de l'exportation.

Vous pouvez configurer l'exportation pour toute période pendant laquelle le programme n'est pas occupé. Si vous essayez d'exporter un document comportant des pages non reconnues, un message vous demande si vous souhaitez procéder d'abord à leur reconnaissance. Si vous sélectionnez l'option *Non*, seuls les résultats de l'OCR des pages reconnues sont exportés. Si vous avez modifié des zones sur des pages reconnues, le programme vous demande de relancer l'OCR des pages concernées avant l'exportation.

PRÉPARER LES RÉSULTATS DE L'OCR POUR L'EXPORTATION

Le texte est exporté dans un fichier, dans le Presse-papiers ou dans un e-mail avec le niveau de formatage défini par le mode sélectionné dans l'éditeur de texte au moment de l'exportation (si possible). Toutefois, certains types de fichiers et certaines applications cibles ne supportent pas tous les attributs de formatage. En cas d'incompatibilité, le programme peut afficher un avertissement et vous proposer le meilleur mode possible. Vous pouvez accepter cette option, ou annuler l'exportation, définir un autre mode et relancer l'opération.

Le tableau de la section *Types de fichiers pour l'enregistrement des résultats de l'OCR* dans le chapitre 6 vous indique les types de fichiers et les niveaux de formatage compatibles.

Les différents modes de l'éditeur présentent les modalités d'exportation suivantes :

Sans formatage



Ce mode est requis pour l'exportation aux formats ASCII, Unicode et autres formats portant l'extension .TXT. Ces types de fichiers n'acceptent ni graphiques ni tableaux. Le texte seul peut, bien sûr, être exporté dans tout type de fichier et d'application cible.

Conserver le formatage polices/paragraphes



Ce mode convient à l'exportation de tous les formats, à l'exception de ceux présentant l'extension TXT ou PDF. Ces formats peuvent gérer les graphiques et les tableaux.

True Page



Ce mode est destiné uniquement aux types de fichiers et d'applications cibles capables de gérer les cadres ou les encadrés de texte. Lorsque vous exportez au format PDF, le mode True Page est utilisé en tant que source, quel que soit le mode spécifié dans l'éditeur (format PDF non disponible dans OmniPage SE). L'ordre de lecture des zones et les modifications d'ordre effectuées dans la fenêtre Ébauche de page ne sont pas pris en compte lorsque True Page est sélectionné pour l'exportation.

Conserver l'enchaînement des colonnes



Ce mode permet de conserver la mise en page du document original (y compris la disposition en colonnes) lors de l'exportation. Si possible, les colonnes, et non les cadres, sont utilisées pour assurer l'enchaînement du texte. Le sens de l'enchaînement est indiqué par des flèches. Vous pouvez le

modifier dans la fenêtre Ébauche de page (voir *Ébauche de page* au fin de chapitre 4).

ENREGISTRER LE DOCUMENT DANS UN FICHIER

Le texte reconnu et les images originales d'un document peuvent être enregistrés sur disque sous différents types de fichiers. Pour consulter une liste des types de fichiers compatibles, reportez-vous au chapitre 6: *Types de fichiers supportés*

Enregistrer des images originales

1. Choisissez Enregistrer image dans le menu Fichier. Dans la boîte de dialogue affichée, sélectionnez un emplacement de dossier et un type de fichier pour vos images. Tapez un nom de fichier.
2. Sélectionnez l'enregistrement de l'image actuelle uniquement ou de toutes les images du document. Dans ce dernier cas, vous pouvez placer l'ensemble des images dans un seul fichier image multipage, à condition de spécifier le type de fichier TIFF ou DCX. Sinon, chaque image est placée dans un fichier séparé. OmniPage SE ajoute des suffixes numériques aux noms de fichiers fournis afin de générer des noms uniques.
3. Cliquez sur OK pour enregistrer la ou les images selon les réglages spécifiés. Les zones et le texte reconnu ne sont pas enregistrés. Si possible, le fichier conserve les propriétés d'affichage : noir et blanc, niveaux de gris ou couleur. Les images en noir et blanc sont enregistrées avec la résolution originale. Dans le cas des images en niveaux de gris et en couleur, la résolution est réduite à environ 150 ppp.

Astuce : Pour voir la taille et la résolution originale d'une image, placez le curseur dessus dans la fenêtre Image originale ou sur sa vignette dans le gestionnaire de document.



Remarque : OmniPage Pro vous permet d'enregistrer vos documents dans quatre versions du format PDF, dont « Image uniquement ». Cette option se rapporte au format d'image, non à l'image originale. L'enregistrement au format PDF n'est pas pris en charge par OmniPage SE.

Enregistrer les résultats de l'OCR

1. Choisissez Enregistrer sous dans le menu Fichier, ou cliquez sur le bouton Exporter le texte reconnu dans la barre d'outils OCR manuel avec l'option Enregistrer sous sélectionnée dans sa liste déroulante.
2. La boîte de dialogue Enregistrer sous apparaît (voir l'illustration ci-dessous).

Cochez cette case pour ouvrir automatiquement le fichier enregistré dans l'application cible.

Cochez cette case pour conserver les styles de paragraphe définis dans l'éditeur de texte lors de l'exportation du texte reconnu.

Cliquez sur Évolué pour afficher la partie inférieure de la fenêtre, et sur Standard pour la masquer.

Vous disposez des options suivantes :
Pour regrouper les pages en un seul fichier
Pour créer un fichier par page
Pour créer un fichier après chaque page blanche

3. Sélectionnez un emplacement de dossier et un type de fichier pour votre document. Le type de fichier OPD (document OmniPage) apparaît en dernière position dans la liste des types disponibles.

4. Indiquez le nom du fichier. Cliquez sur le bouton Évolué pour afficher l'ensemble des options d'enregistrement, et sélectionnez les options voulues.
5. Cliquez sur OK. Le document est enregistré sur disque selon les réglages spécifiés. Si l'option Enregistrer et lancer est activée, le fichier exporté apparaît dans son application cible, c.-à-d. l'application associée au type de fichier dans Windows.

Remarque : Les graphiques et les réglages de formatage sont enregistrés dans le document uniquement si le type de fichier sélectionné les prend en charge. Le niveau de formatage pour l'exportation correspond au mode de l'éditeur de texte défini lors de l'enregistrement. En cas d'incompatibilité entre le niveau de formatage et le type de fichier spécifiés, le programme affiche un avertissement.

Remarque : Si plusieurs fichiers d'exportation sont créés, OmniPage SE ajoute des suffixes numériques aux noms de fichiers fournis afin de générer des noms uniques. Si, lors de l'importation des fichiers image, vous activez l'option Pour créer un fichier après chaque page blanche, consultez la section *Importer des fichiers image* du chapitre 3 pour plus d'informations sur le placement des pages blanches.

ENREGISTRER UN DOCUMENT LORS D'UNE SESSION DE TRAVAIL



Cliquez sur le bouton Enregistrer dans la barre d'outils Standard ou choisissez Enregistrer dans le menu Fichier pour stocker les modifications apportées au document actuel lors d'une session de travail. Si vous traitez un document sans titre, la boîte de dialogue Enregistrer sous apparaît.

Dans le cas d'un document nommé, la commande Enregistrer effectue une sauvegarde sous le nom et le format du dernier enregistrement, tel qu'indiqué dans la barre de titre. Si le document est au format Document OmniPage, les images nouvelles ou modifiées, les zones modifiées, les résultats de l'OCR et les données d'apprentissage sont enregistrées. Si le document est au

format texte, seules les modifications apportées aux résultats de l'OCR sont enregistrées.

Si vous souhaitez retraiter un document dans OmniPage SE lors d'une session ultérieure, enregistrez-le en tant que Document OmniPage. Ce format représente un type de fichier de sortie permettant de sauvegarder l'image originale, ainsi que les résultats de l'OCR, les réglages et les données d'apprentissage (consultez la section *Documents OmniPage* du chapitre 2). La boîte de dialogue Enregistrer sous affiche les types de fichiers disponibles dans la liste déroulante Format. Le type de fichier Document OmniPage apparaît en dernière position dans la liste.

Si vous enregistrez d'abord le document au format Document OmniPage (par exemple, memo.opd), puis le modifiez et l'enregistrez au format texte (par exemple, memo.txt), puis le remodifiez et cliquez sur Enregistrer, les dernières modifications sont enregistrées dans le fichier memo.txt, et non dans le fichier OPD. Lorsque vous fermez le document ou quittez le programme, vous êtes invité à enregistrer le fichier s'il n'a pas été enregistré en tant que Document OmniPage ou si des modifications ont été effectuées depuis la dernière sauvegarde au format OPD.



Il est possible d'échanger des documents OmniPage entre les versions SE et Pro 11 de l'application. Dans OmniPage SE, toutes les données d'apprentissage du fichier OPD sont ignorées et l'apprentissage est impossible.

COPIER UN DOCUMENT DANS LE PRESSE-PAPIERS

Vous pouvez copier les résultats de la reconnaissance de toute page d'un document dans le Presse-papiers. Le processus de copie est suivi par une jauge de progression. Vous pouvez ensuite coller le contenu du Presse-papiers dans une autre application.

Le formatage du texte (gras, italique, etc.) est conservé si la copie est réalisée vers une application compatible avec le format RTF. Sinon, seules les données textuelles sont copiées. Les graphiques sont conservés dans les applications gérant l'insertion d'images.

▼ Pour copier un document dans le Presse-papiers :

- Dans le cas d'un traitement automatique, sélectionnez la commande Copier dans le Presse-papiers dans la liste déroulante Exporter le texte reconnu de la barre d'outils OCR automatique ou dans la fenêtre de l'Assistant OCR. Le texte est placé dans le Presse-papiers dès que la dernière page disponible est reconnue ou vérifiée.
- Dans le cas d'un traitement manuel, sélectionnez la commande Copier dans le Presse-papiers dans la liste déroulante Exporter le texte reconnu, puis cliquez sur le bouton correspondant dans la barre d'outils OCR manuel. L'opération de copie démarre sur-le-champ.

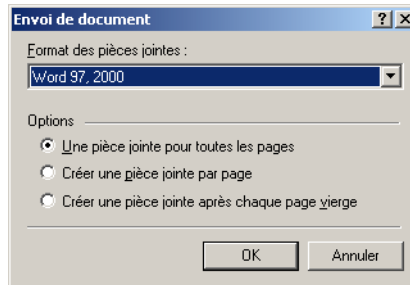
JOINDRE UN DOCUMENT À UN E-MAIL

Si vous disposez d'une application compatible MAPI, telle que Microsoft Outlook, vous pouvez envoyer les résultats de la reconnaissance sous la forme d'un ou plusieurs fichiers joints à un message électronique.

▼ Pour envoyer un document comme pièce jointe à un e-mail :

- Dans le cas d'un traitement automatique, sélectionnez la commande Envoyer comme e-mail dans la liste déroulante Exporter le texte reconnu de la barre d'outils OCR automatique. Dès que la reconnaissance ou la vérification de la dernière page du document disponible est terminée, la boîte de dialogue Envoi de document apparaît.
- Dans le cas d'un traitement manuel, sélectionnez la commande Envoyer comme e-mail dans la liste déroulante Exporter le texte reconnu, puis cliquez sur le bouton correspondant dans la barre d'outils OCR manuel. La boîte de dialogue Envoi de document s'affiche sur-le-champ.

Pendant toute période d'inactivité du programme, vous pouvez afficher la boîte de dialogue Envoi de document en choisissant la commande Envoyer comme e-mail dans le menu Fichier.



1. La boîte de dialogue Envoi de document permet de spécifier le type de fichier et les options d'envoi : une pièce jointe pour toutes les pages, une pièce jointe par page ou une pièce jointe après chaque page vierge. Sélectionnez les options voulues et cliquez sur OK.
2. Au besoin, connectez-vous à votre application de messagerie.
3. La fenêtre qui s'ouvre génère un nouveau message et y joint la ou les pièces spécifiées. Les pièces jointes adoptent le nom utilisé lors du dernier enregistrement du document dans OmniPage SE ou emploient le nom « OmniPage – Sans titre ». L'extension de fichier appropriée est ajoutée, ainsi que des suffixes numériques dans le cas de plusieurs pièces jointes.
4. Spécifiez une ou des adresses de destinataire, tapez le texte du message et cliquez sur le bouton Envoyer.



6 Informations techniques

Ce chapitre regroupe les informations sur le dépannage des problèmes rencontrés lors de l'utilisation d'OmniPage SE et d'autres informations techniques relatives au logiciel.

Consultez également le fichier ReadMe/LisezMoi et l'aide d'OmniPage SE, disponibles en ligne, ou visitez le site Web de ScanSoft. La page du site consacrée aux informations sur les scanners fournit des informations détaillées et mises à jour régulièrement sur la configuration et le support des différentes marques de scanner. Le fichier ReadMe/LisezMoi contient des informations de dernière minute concernant OmniPage SE. Pour accéder au fichier ReadMe/LisezMoi et au site Web de ScanSoft, utilisez le menu Aide.

Ce chapitre traite les sujets suivants :

- ◆ Dépannage
 - Solutions de premier recours
 - Tester OmniPage SE
 - Mémoire insuffisante
 - Espace disque insuffisant
- ◆ Types de fichiers supportés
 - Types de fichiers pour l'ouverture et l'enregistrement des images
 - Types de fichiers pour l'enregistrement des résultats de l'OCR
 - Enregistrer un document au format PDF
- ◆ Problèmes de reconnaissance
 - Reconnaissance de texte incorrecte
 - Reconnaissance de fax incorrecte
 - Problèmes de système ou de performances pendant l'OCR
- ◆ Désinstaller le logiciel

DÉPANNAGE

OmniPage SE est une application très simple d'emploi, mais la diversité des configurations système peut être une source de problème. Lisez toujours attentivement les messages d'erreur ; leur énoncé vous suggère souvent les mesures à prendre – vérifier les connexions, fermer toutes les autres applications pour libérer de la mémoire, etc. Dans bien des cas, vous n'aurez pas besoin d'autres renseignements de dépannage.

Pour plus de détails sur l'optimisation de votre environnement et de vos applications, consultez la documentation Windows.

Solutions de premier recours

Si vous rencontrez un problème lors du lancement ou de l'utilisation d'OmniPage SE, commencez par vérifier les points suivants :

- ◆ Vérifiez que la configuration de votre système respecte les spécifications présentées dans la section *Environnement requis* du chapitre 1.
- ◆ Vérifiez que le scanner est sous tension et que ses câbles sont correctement connectés.
- ◆ Éteignez le scanner et l'ordinateur, puis rallumez-les (le scanner en premier). Vérifiez que vos applications habituelles fonctionnent correctement.
- ◆ Avant d'utiliser votre scanner avec OmniPage SE, exécutez le logiciel livré avec celui-ci pour vérifier qu'il fonctionne correctement.
- ◆ Vérifiez que votre système dispose des versions correctes des pilotes de scanner, d'imprimante et d'affichage vidéo. Pour plus de détails, visitez la page consacrée aux scanners du site Web de ScanSoft (accessible via le menu Aide).
- ◆ Exécutez l'utilitaire ScanDisk (Windows 95, 98 et Me) ou Check Disk (Windows NT et 2000) pour vérifier l'intégrité du disque dur. Pour plus de détails, reportez-vous à l'aide en ligne de Windows.
- ◆ Défragmentez le disque dur (voir l'aide en ligne de Windows).
- ◆ Désinstallez et réinstallez OmniPage SE (voir la dernière section de ce chapitre: *Désinstaller le logiciel*).

- ◆ Visitez la section Support du site Web de ScanSoft, à l'adresse www.scansoft.com, pour consulter les notes techniques (Tech Notes) consacrées aux problèmes les plus courants affectant OmniPage SE. Ces pages Web offrent également de l'assistance pour l'installation et le dépannage du logiciel.

Tester OmniPage SE

Le redémarrage de Windows 95, 98, 2000 et Me en mode sans échec ou de Windows NT en mode VGA permet de tester OmniPage SE sur un système à configuration simplifiée. Cette solution est conseillée en cas de blocage persistant du système ou d'OmniPage SE. Pour plus de détails, reportez-vous à l'aide en ligne de Windows.

Remarque : Les scanners ne fonctionnent pas avec OmniPage SE en mode sans échec et en mode VGA ; ne testez pas les problèmes de scanner sous ces modes.

▼ Pour tester OmniPage SE en mode sans échec (Windows 95, 98, 2000 et Me) :

1. Redémarrez Windows 95 : lorsque le message « Démarrage de Windows » apparaît, appuyez sur <F8>.
2. Lancez OmniPage SE et exécutez l'OCR d'une image. Utilisez un fichier image existant, par exemple un des fichiers exemples fournis.
 - Si OmniPage SE ne démarre pas ou ne fonctionne pas correctement en mode sans échec, cela signifie généralement que le problème se situe au niveau de l'installation. Désinstallez et réinstallez OmniPage SE (voir la dernière section de ce chapitre: *Désinstaller le logiciel*), puis exécutez le programme en mode sans échec.
 - Si OmniPage SE fonctionne correctement en mode sans échec, le problème provient sans doute d'un conflit avec l'un des pilotes de périphérique. Relancez Windows en mode de confirmation étape par étape (pour plus de détails, reportez-vous à l'Aide en ligne de Windows).

▼ Pour tester OmniPage SE en mode VGA (Windows NT) :

1. Redémarrez l'ordinateur.
2. Sélectionnez la commande correspondant au mode VGA et appuyez sur la touche Entrée.
3. Appuyez sur Ctrl+Alt+Suppr et sélectionnez Gestionnaire des tâches.
4. Dans la fenêtre Gestionnaire des tâches, sélectionnez toutes les applications d'arrière-plan et cliquez sur Fin de tâche. Pour plus de détails, reportez-vous à l'aide en ligne de Windows.
5. Lancez OmniPage SE et exécutez l'OCR d'une image. Utilisez un fichier image existant (un fichier exemple au format .tif).

Remarque : Vous pouvez également exécuter le logiciel OmniPage SE en mode sans échec. Choisissez Démarrer ➤ Exécuter, recherchez le fichier OmniPage.exe et tapez /safe dans la ligne de commande. Cette opération lance le programme, mais sans prendre en compte les réglages enregistrés et sans essayer de récupérer des documents fermés incorrectement.

Mémoire insuffisante

Lorsque la mémoire vive disponible (RAM) est insuffisante, les performances d'OmniPage SE se dégradent : cela se manifeste par l'affichage de messages d'erreur, un fonctionnement ralenti, de nombreux accès au disque dur, etc. Dans ce cas, appliquez les solutions suivantes :

- ◆ Relancez l'ordinateur.
- ◆ Fermez les autres applications pour libérer de la mémoire.
- ◆ Fermez les fenêtres OmniPage SE superflues.
- ◆ Défragmentez le disque dur (pour plus de détails, reportez-vous à l'aide en ligne de Windows).
- ◆ Libérez de l'espace sur le disque dur.
- ◆ Augmentez la mémoire vive (RAM).

- ◆ La quantité de mémoire RAM conditionne la vitesse de la reconnaissance. Pour plus de détails, consultez la section *Environnement requis* du chapitre 1.

Espace disque insuffisant

Si l'espace disponible sur le disque dur est insuffisant, les performances d'OmniPage SE risquent de se dégrader. Dans ce cas, appliquez les solutions suivantes :

- ◆ Videz la Corbeille de Windows.
- ◆ Quittez toutes les applications ouvertes et supprimez les fichiers *.tmp présents dans le dossier Temp, (résidant généralement dans le dossier Windows).
- ◆ Exécutez l'utilitaire ScanDisk (Windows 95, 98 et Me) ou Check Disk (Windows NT et 2000) afin d'identifier les erreurs responsables de l'occupation de l'espace disque. Pour plus de détails, reportez-vous à l'aide en ligne de Windows.
- ◆ Créez une copie de sauvegarde des fichiers superflus, puis supprimez ces fichiers du disque dur.
- ◆ Supprimez les applications Windows que vous n'utilisez pas.
- ◆ Défragmentez le disque dur (pour plus de détails, consultez l'aide en ligne de Windows).
- ◆ Videz le cache de votre navigateur Web et diminuez sa taille.

TYPES DE FICHIERS SUPPORTÉS

Le logiciel gère de nombreux types de fichiers, et cette gamme a été enrichie d'éléments importants dans OmniPage SE.

Types de fichiers pour l'ouverture et l'enregistrement des images

Type de fichier	Extension	Multi-page	Ouverture/ Enreg.	Noir et blanc, niv. de gris, couleur
BMP, bitmap	*.bmp	Non	Ouvert. et enreg.	Tous
DCX	*.dcx	Oui	Ouvert. et enreg.	Tous
GIF	*.gif	ND	ND	ND
JPEG	*.jpg	Non	Ouvert. et enreg.	Niv. de gris, couleur
PCX	*.pcx	Non	Ouvert. et enreg.	Tous
PDF	*.pdf	ND	ND (voir remarque)	ND
PNG	*.png	Non	Ouvert. et enreg.	Tous
TIFF compressé G3	*.tif	Oui	Ouverture	Noir et blanc
TIFF compressé G4	*.tif	Oui	Ouvert. et enreg.	Noir et blanc
TIFF compressé LZW	*.tif	ND	ND	ND
TIFF FX	*.xif	ND	ND	ND
TIFF non compressé	*.tif	Oui	Ouvert. et enreg.	Tous
TIFF PackBits	*.tif	Oui	Ouvert. et enreg.	Tous

Les fichiers image importés peuvent présenter une résolution maximale de 600 ppp, mais la valeur de 300 ppp (dans le sens horizontal et vertical) est recommandée pour une précision de reconnaissance optimale. Le programme stocke les images en noir et blanc avec leur résolution originale ; toutefois, la résolution d'enregistrement des images en niveaux de gris et en couleur ne dépasse généralement pas 150 ppp. Placez le curseur sur une image pour afficher une mini-fenêtre indiquant la taille et la résolution de l'image originale.

Remarque : Si vous tentez d'enregistrer une image en noir et blanc au format JPEG, le programme vous propose d'effectuer une conversion en niveaux de gris. Dans le cas des formats TIFF G3 et G4, OmniPage SE suggère une conversion en noir et blanc.



Remarque : Seul OmniPage Pro 11 gère l'enregistrement au format PDF et propose à l'utilisateur quatre versions de ce format. Une d'entre ces options, « avec image uniquement », permet d'exporter les résultats de l'OCR au format d'image, et non l'image originale. Vous pouvez choisir le format requis dans la boîte de dialogue Enregistrer sous. Veuillez noter que OmniPage SE ne gère pas les fonctionnalités PDF. OmniPage SE ne gère pas non plus les fichiers GIF, LZW TIFF et TIFF FX.

Types de fichiers pour l'enregistrement des résultats de l'OCR

Type de fichier	Extension	Niveaux de formatage (modes de l'éditeur de texte)	Support des graphi- ques
Adobe PDF, standard	*.pdf	ND	ND
Adobe PDF avec image uniquement	*.pdf	ND	ND
Adobe PDF avec l'image couvrant le texte	*.pdf	ND	ND
Adobe PDF avec substitution d'image	*.pdf	ND	ND
ASCII, texte ¹	*.txt/.csv	Sans formatage (SF)	Non
Excel (3.0 à 7.0, 97, 2000)	*.xls	SF, CFPP (Feuille de calcul)	Oui
FrameMaker (5.5.3)	*.mif	Tous	Oui
Freelance Graphics	*.txt	Sans formatage (SF)	Non
Harvard Graphics	*.txt	Sans formatage (SF)	Non
HTML (3.2 ou 4.0) ²	*.htm	Tous ²	Oui ²
Microsoft Publisher 98	*.rtf	Tous	Oui
PageMaker 6.5.2	*.doc	Tous	Oui
PowerPoint 97	*.rtf	Tous	Oui
Quattro Pro pour Windows 4.0, 8	*.xls	SF, CFPP (Feuille de calcul)	Non
RTF 6.0/95 ³	*.rtf	Tous	Oui
Unicode, texte ^{1 4}	*.txt/.csv	Sans formatage (SF)	Non
Ventura Publisher	*.doc	Tous	Oui
Word pour Windows (6.0, 97, 2000)	*.doc	Tous	Oui
WordPad	*.rtf	SF, CFPP ⁵	Oui
WordPerfect (5.1, 5.2, 6.0, 6.1, 8, 9, 10)	*.wpd	Tous	Oui
Document OmniPage ⁶	*.opd	Tous	Oui

- 1 Les formats ASCII et Unicode offrent les options d'enchaînement du texte, de sauts de ligne et de virgules de séparation. Cette dernière option utilise l'extension .csv et est utilisée pour le transfert des données textuelles d'un tableau dans une application de type Tableur.
- 2 Lors de l'enregistrement au format HTML, tous les graphiques sont enregistrés en tant que fichiers image séparés, sous le format JPEG. Le format HTML 4.0 n'est supporté que par OmniPage Pro 11. OmniPage SE est limité au format HTML 3.2.
- 3 Les résultats de l'OCR sont placés dans le Presse-papiers sous ce format, puis sont collés au format RTF si possible, et sinon, au format Unicode ou ASCII.
- 4 Le format Unicode est compatible avec le plus grand nombre de caractères accentués.
- 5 Les modes True Page et Conserver l'enchaînement des colonnes (CEC) ne sont pas refusés, mais le texte apparaît en mode Conserver le formatage polices/paragraphes (CFPP), c.-à-d. sans colonnes.
- 6 Les documents OmniPage créés dans OmniPage SE ou OmniPage Pro 11 peuvent être rouverts par OmniPage SE. Le programme peut également ouvrir les fichiers OPD créés avec OmniPage Pro 10 et les fichiers MET provenant d'OmniPage Pro 9. Ces types de fichiers sont conservés dans leur format d'origine et une copie est convertie au format OmniPage SE.



Enregistrer un document au format PDF

Cette section ne s'applique pas à OmniPage SE. Dans OmniPage Pro 11, lors de l'enregistrement des résultats de l'OCR au format PDF (Portable Document Format), vous avez le choix entre quatre options :

PDF standard

Les pages exportées conservent l'apparence du mode True Page de l'éditeur de texte. Le fichier PDF peut être affiché et recherché dans une visionneuse PDF et édité dans un éditeur PDF.

Avec substitution d'image

Ce format est similaire au format décrit ci-dessus, mais l'image originale se substitue aux caractères de rejet et douteux, ce qui assure un affichage identique au document original. Le fichier PDF peut être affiché, recherché et édité.

Avec l'image uniquement

Le fichier PDF peut être affiché, mais le texte ne peut pas être recherché, et non plus modifié dans un éditeur PDF.

Avec l'image couvrant le texte

Le fichier PDF peut être affiché, mais il ne peut être modifié dans un éditeur PDF. Par contre, chaque image dispose d'un fichier texte associé, de sorte que le fichier accepte les opérations de recherche. Tout mot trouvé est contrasté dans l'image.

PROBLÈMES DE RECONNAISSANCE

Cette section fournit des informations sur les problèmes de reconnaissance qui peuvent survenir lors de l'utilisation d'OmniPage SE et suggère des solutions possibles. Les sujets abordés sont les suivants : amélioration de la précision de l'OCR, optimisation de l'importation des fax et problèmes de performances lors de la reconnaissance.

Reconnaissance de texte incorrecte

Si une partie du document original n'est pas convertie en texte correctement pendant l'OCR, vous pouvez avoir recours aux solutions suivantes :

- ◆ Examinez l'image de page originale et vérifiez que tous les blocs de texte sont entourés d'une zone (tout bloc de texte non entouré d'une zone est ignoré lors de la reconnaissance). Pour plus de détails sur la création et la modification de zones, reportez-vous au *Tracer des zones manuellement* dans chapitre 3.
- ◆ Vérifiez que les zones de texte sont identifiées correctement. Si nécessaire, sélectionnez d'autres valeurs de type et/ou de contenu pour les zones existantes et relancez la reconnaissance du document. Pour plus de détails, reportez-vous au *Propriétés de zone* dans chapitre 3.
- ◆ Assurez-vous que vous n'avez pas chargé un gabarit incorrect. Si les bords des zones empiètent sur le texte, la qualité de la reconnaissance en est affectée.
- ◆ Ajustez les curseurs de luminosité et de contraste (onglet Scanner de la boîte de dialogue Options) : vous devrez sans doute essayer plusieurs combinaisons de réglages pour obtenir le résultat escompté.



- ◆ Vérifiez la résolution de l'image originale. Placez le curseur dans la fenêtre Image originale pour afficher une mini-fenêtre contenant les informations de résolution et de taille de l'image. Une résolution bien supérieure ou inférieure à 300 ppp risque d'affecter la reconnaissance.
- ◆ Vérifiez que les langues du document ont été correctement sélectionnées dans les réglages OCR ; seules les langues effectivement présentes dans le document doivent être sélectionnées.
- ◆ Activez IntelliTrain et effectuez les corrections nécessaires. Cette fonction convient tout particulièrement aux polices stylisées et aux documents présentant une dégradation uniforme. Si IntelliTrain était actif lors de l'OCR, désactivez-le car il n'est pas destiné à certains types de documents très dégradés. Ceci ne s'applique pas à OmniPage SE.
- ◆ Si vous activez le mode True Page dans l'éditeur de texte pour l'affichage ou l'exportation, le texte reconnu est placé dans des cadres de formatage. Si l'un de ces cadres est trop petit, une partie du texte peut être masquée. Pour afficher le texte caché, placez le curseur dans ce cadre et utilisez les touches flèches du clavier pour faire défiler le texte dans la direction requise.
- ◆ Vérifiez la vitre, les miroirs et les objectifs du scanner. Éliminez les traces de doigt et de poussière.

Remarque : OmniPage SE reconnaît uniquement les caractères imprimés (machine à écrire, imprimante laser, etc.). Le programme peut traiter les caractères matriciels, mais la précision de l'OCR risque d'être insatisfaisante pour les textes de qualité moyenne. Le programme n'accepte pas les caractères manuscrits ; toutefois, vous pouvez lui demander de conserver un bloc de texte manuscrit (une signature, par exemple) en entourant celui-ci d'une zone de type Graphique.

Reconnaissance de fax incorrecte

Pour améliorer la précision de reconnaissance des fax, appliquez les solutions suivantes :

- ◆ Demandez à votre correspondant d'utiliser un document original propre, à impression nette.

- ◆ Demandez à votre correspondant de sélectionner le mode Fin (ou équivalent) lors de l'envoi de fax. En général, la résolution obtenue est de 200x200 ppp.
- ◆ Si votre PC et celui de votre correspondant sont équipés d'un fax-modem, demandez-lui de transmettre son fichier fax directement à votre ordinateur. Vous pouvez ensuite enregistrer ce fax sous forme de fichier image, puis le charger dans OmniPage SE. Pour plus de détails, reportez-vous à la section *Importer des fichiers image* du chapitre 3.

Problèmes de système ou de performances pendant l'OCR

Si le système se bloque pendant la reconnaissance, ou si le traitement est particulièrement lent, essayez les solutions suivantes :

- ◆ Éliminez les problèmes de mémoire vive (voir section *Tester OmniPage SE*).
- ◆ Éliminez les problèmes d'espace disque insuffisant (voir section *Tester OmniPage SE*).
- ◆ Réduisez en icônes toutes les applications (ou appuyez plusieurs fois sur Alt+Tab) pour visualiser les éventuels messages d'erreur Windows.
- ◆ Vérifiez la qualité de l'image destinée à la reconnaissance.
- ◆ Pour améliorer la qualité des images numérisées, consultez la documentation de votre scanner.
- ◆ Simplifiez le traitement des images de page chargées (texte dense, nombreux graphiques et/ou formatage complexe) en exécutant la reconnaissance sur une partie d'image à la fois. Tracez les zones manuellement ou modifiez les zones créées en mode automatique, puis lancez la reconnaissance sur une partie de page à la fois. Pour plus d'informations sur la création et la modification des zones, reportez-vous au *Tracer des zones manuellement* dans chapitre 3.
- ◆ Relancez Windows 95, 98, Me ou 2000 en mode sans échec (ou Windows NT en mode VGA), et testez OmniPage SE en effectuant la reconnaissance des fichiers image copiés dans votre système lors de l'installation (fichiers exemples au format .tif). Pour plus de détails, consultez la section *Tester OmniPage SE*.

Le traitement est plus lent si vous exécutez plusieurs tâches simultanément (par exemple, OCR et impression).

DÉSINSTALLER LE LOGICIEL

Certains problèmes peuvent être résolus en désinstallant puis en réinstallant OmniPage SE. Veillez également à désinstaller OmniPage SE avant d'installer OmniPage Pro 11 ou toute autre version d'évaluation du programme. Le programme de désinstallation d'OmniPage SE ne supprime pas les fichiers suivants :

- Fichiers de gabarit de zones (*.zon)
- Fichiers d'apprentissage (*.otd) (Ne concerne pas OmniPage SE)
- Dictionnaires utilisateur (*.ud)
- Documents OmniPage (*.opd)

Pour désinstaller OmniPage SE sous Windows NT ou Windows 2000, vous devez ouvrir une session avec privilèges d'administrateur.

▼ Pour désinstaller ou réinstaller OmniPage SE :

- ◆ Fermez OmniPage SE.
- ◆ Dans la barre des tâches Windows, cliquez sur Démarrer et choisissez Paramètres ► Panneau de configuration et cliquez deux fois sur ► Ajout/Suppression de programmes.
- ◆ Sélectionnez OmniPage SE et cliquez sur Ajouter/Supprimer.
- ◆ Dans la boîte de dialogue qui apparaît, cliquez sur Suivant.
- ◆ Sélectionnez l'option de suppression ou d'annulation et cliquez sur Suivant.
- ◆ Suivez les instructions affichées à l'écran.



I N D E X

- A**
- ADF, 34, 53, 56
 - Afficher/masquer les marques, 72
 - Aide
 - contextuelle, ix, 25, 34
 - en ligne, ix
 - HTML en ligne, ix
 - Ajouter
 - des pages à un document, 45
 - Ajouter un travail, Assistant, 52
 - Apprentissage, 69
 - charger un fichier, 71
 - créer des données, 71
 - décharger un fichier, 71
 - éditer un fichier, 71
 - fichiers, 71
 - nouvelles données, 32
 - ASCII, texte, 92
 - Assistant OCR, 40
 - Attributs, caractère, 73
 - Automatique, traitement, 28, 44
 - AutoOCR, 27, 44
- B**
- Barre d'outils
 - Formatage, 24, 27
 - Image, 24, 26
 - Standard, 24, 25
 - Barre des menus, 26
 - Boîte à outils OmniPage, 24, 27, 44
 - Boîte de dialogue Options, 34
 - Boutons de commande pour le traitement automatique, 45
 - Bureau OmniPage SE, 24
- C**
- Cadres, 94
 - Cadres dans l'exportation de documents, 78
 - Caractères non imprimables, 27
- Changer**
 - grilles de tableau, 61
 - mode de traitement, 47
 - modifier, type de zone, 61
- Charger fichier, boîte de dialogue, 53**
- Chargeur de feuilles automatique (ADF), 53, 56**
- Colonne unique**
 - avec tableau, mise en page, 58
 - mise en page, 57
 - type de zone, 60
- Colonnes**
 - modifier la disposition du texte, 76
 - multiples, 57
- Comparer le texte reconnu à l'image originale, 67**
- Configurer Direct OCR, 49**
- Configurer un scanner, 14**
- Conserver**
 - enchaînement des colonnes, mode, 65, 72, 79
 - formatage polices/paragraphes, mode, 65, 72, 78
 - styles de paragraphe, 81
- Conversion d'images, 90**
- Copier**
 - document dans le Presse-papiers, 42, 82
 - et coller du texte, 26
- Correspondance entre le mode de l'éditeur et le type de fichier, 91**
- Correspondance entre niveau de formatage et type de fichier, 78**
- Couleur, numériser, 55**
- Couper et coller du texte, 26**
- D**
- Démarrer le programme, 14
 - Dépannage, 85, 86
 - Désinstaller OmniPage SE, 96
- Dictionnaire utilisateur, 66**
 - ajouter des mots, 66
 - charger, 68
 - créer, 68
 - décharger, 68
 - éditer, 69
 - lancer, 68
- Différer le traitement, 32**
- Direct OCR, 49**
- Document OmniPage, 14**
 - avantages, 33
 - contenu, 82
 - enregistrer, 33
 - fichiers OPD, 32
- Documents**
 - avec mise en page variable, 57
 - décrire la mise en page, 57
 - recto verso, 56
 - spécifier la mise en page, 57
- Documents dans OmniPage SE, 23**
- Documents non terminés, 32**
- E**
- Ébauche de page, 76
 - Éditer des graphiques, 74
 - Éditeur de texte, 35, 65
 - changer de mode, 73
 - modes, 72
 - traiter un tableau, 74
 - Éditeur de texte, fenêtre, 24
 - Édition
 - attributs, 73
 - caractères, 73
 - texte, 72
 - E-mail, destination d'exportation, 83
 - Enregistrer
 - applications pour Direct OCR, 49
 - au format Document OmniPage, 33, 82
 - dans un fichier, 43, 79
- MANUEL OMNIPage SE** **97**

- document au format PDF, 92
- document lors d'une session de travail, 33, 81
- documents, 77
- fichier d'apprentissage, 71
- gabarit de zone, 63
- images originales, 79, 90
- résultats de l'OCR, 80
- Enregistrer et lancer, option, 81
- Entendre le texte lu à voix haute, 74
- Environnement requis, 12
- Envoi de document, boîte de dialogue, 83
- Espace disque, 12
- Espace disque insuffisant, 89
- Exécuter l'OCR, bouton, 47
- Exemples de fichiers image, 87
- Exportation
 - préparatifs, 78
 - types de fichiers, 78
- Exporter, 77, 78
 - dans un fichier, 80
 - document dans le Presse-papiers, 82
 - documents, 23
 - graphiques, 81
 - plusieurs fois, 77, 82
- Exporter le texte reconnu, bouton, 44, 47

F

- Faire pivoter une image, 26
- Fenêtre Détails
 - colonnes, 31
- Fermer un document, 32
- Feuille de calcul, mise en page, 58
- Fichier
 - destination d'exportation, 79
 - options, 80
 - séparation, 80
 - types, 80
- Fichier image
 - exemples, 87
 - importer, 23, 53
 - multipages, 53, 79, 90
 - ordre, 53
 - ouvrir, 22, 90
 - pivoter, 26
 - source, 53
 - TIFF, 90
 - types, 90

Fichiers

- conservés après désinstallation, 96
- exportation, 91
- options, 84
- séparation, 84
- types, 93
- types pour l'exportation, 91
- types supportés, 90, 91
- Formatage, niveaux, 42

G

- Gabarit de zones, 58, 93
- Gérer
 - documents, 28
 - zones, 59
- Gestionnaire
 - de document, 28
 - document, 24
 - fenêtre, 24
 - tâches, 88
- Graphiques
 - éditer, 74
 - exportation, 81, 91
 - fichiers JPEG, 92

I

- Image
 - conversion, 90
 - couleur, 79
 - importer, 46
 - niveaux de gris, 79
 - noir et blanc, 79
 - qualité, 55
 - résolution, 30, 79, 90, 94
 - taille, 30
- Image originale
 - enregistrer, 79
 - fenêtre, 24
 - sauvegarder, 79
- Importer des fichiers image, 53
- Impression matricielle, 94
- Imprimer
 - document, 32
 - images, 26
- Informations techniques, 85
- Installation
 - OmniPage SE, 13
 - scanner, 14

- IntelliTrain, 32, 35, 52, 69, 94
- Interface, langue, 35

J

- Joindre un document à un e-mail, 83

L

- Lancer l'application cible, 81
- Lancer un dictionnaire utilisateur, 68
- Langues
 - installation, 13
 - interface utilisateur, 13, 35
 - module Texte-parole, 13
 - reconnaissance, 42, 94
- Lire le texte à voix haute, 74

M

- Manuel, traitement, 28
- Marques de couleur, 67
- Mémoire insuffisante, 88
- Mémoire requise, 12
- Messages électroniques, pièces jointes, 83
- Microsoft Outlook, 83
- Minimum, environnement requis, 12
- Mise à niveau vers OmniPage Pro, 19
- Mise en page
 - description, 42, 44, 57
 - personnalisée, 35
- Mode sans échec, 87
- Mode VGA, 88
- Modifier le texte reconnu, 27, 73
- Mots marqués dans l'éditeur de texte, 72

N

- Niveaux de formatage, 52, 65, 72, 91
- Niveaux de formatage pour l'exportation, 78
- Niveaux de gris, numériser, 55
- Noir et blanc, numériser, 55

Nouvelles fonctions d'OmniPage
Pro 11, 18

Numérisation
assistant, 14

Numériser
contraste, 55
en couleur, 55
en niveaux de gris, 55
en noir et blanc, 55
images, 54, 55
luminosité, 55

O

OCR

Assistant, 40, 41, 43
AutoOCR, 27, 44
Direct OCR, 34
exécuter, 23
exécuter, bouton, 44
méthode, 34, 42
optimiser, 34
planifier OCR, 52
vitesse, 34

Omettre des graphiques lors de
l'exportation, 81

OmniPage

versions antérieures, 13

OmniPage Pro

mise à niveau vers, 19
nouvelles fonctions, 18

OmniPage SE, x, 19

bureau, 24
démarrer, 14
désinstaller, 96
documents, 23
fonctionnalités, 19
installation, 13
réinstaller, 96
tester, 87

Onglets

Détails, 24, 29
Vignettes, 24, 29

OPD

avantages, 33
enregistrer au format, 33

Optimiser la qualité d'image, 55

P

Page

divisée en zones, 28
état, 28
exploration d'un document, 24
importée, 28
Importer la page, bouton, 44,
46
parcourir les pages d'un docu-
ment, 29
reconnue, 28
réordonner, 29
sélectionner plusieurs, 29
supprimer, 29
vérifiée, 28

PaperPort, 51

Paragraphe

alignement, 27
attributs, 74
édition, attributs, 73
modifier l'ordre, 76
réordonner, 76
styles, 27, 81

PDF

afficher, 92
éditer, 92
explorer, 92
exportation au format, 92
importation, 53

Performances, problèmes pendant
l'OCR, 95

Personnalisée, mise en page, 58

Personnaliser des colonnes dans la
fenêtre Détails, 31

Pièces jointes, 83

Placer des séparations dans un
tableau, 26

Planifier OCR, 52

Pour, 14

Précision

améliorer, 34, 54, 69

Préparer les résultats de l'OCR pour
l'exportation, 78

Presse-papiers, 82

Problèmes

espace disque insuffisant, 89
mémoire insuffisante, 88
performances pendant OCR, 86
reconnaissance de fax, 94
système ou de performances
pendant l'OCR, 95

R

Rechercher

mots absents du dictionnaire,
66
mots douteux, 66

Reconnaissance

exécuter, 46
fax incorrecte, 94
langues, 34
lente, 95
problèmes, 93

Reconnaissance optique de
caractères, OCR, 22

Réglages

acquérir les réglages de texte, 49
Direct OCR, 49
divers, 35

Réinstaller OmniPage SE, 96

Réordonner des paragraphes et des
zones, 76

Résolution

images, 30, 90, 94
images enregistrées, 79

Retour à la ligne automatique, 35

S

Sans Formatage, mode, 72, 78

Scanner, 94

chargeur de feuilles automatique
(ADF), 34
configurer pour OmniPage SE,
14

contraste, 34, 93
importer des documents, 54
luminosité, 34, 93
pilotes, 14

recto verso, 56

Spécifier la mise en page du
document, 44

Styles de paragraphe, 74

Substituts d'image dans des fichiers
PDF, 92

Suggestion de dictionnaire pour la
vérification de l'OCR, 66

Supprimer

pages, 31

T

Tableau

- colonnes, 62
- déplacer les séparations, 62
- éditer, 74
- générer les séparations, 62
- insérer des séparations, 62
- modifier les séparations, 62
- modifier les séparations d'un tableau, 26
- rangées, 62
- séparations, placer dans un tableau, 26
- supprimer les séparations, 62

Tester OmniPage SE, 87

Texte

- enregistrer, 80
- exportation, 92
- gras, 27
- italique, 27
- modifier la disposition, 76
- reconnu, comparer à l'image originale, 67
- sans formatage, 65
- souligné, 27

Texte-parole fonction, 13, 74

Touches utilisées pour la lecture du texte reconnu, 75

Traçage manuel des zones, 26, 46

Tracer des zones, 50, 58

Traitement

- à partir d'autres applications, 49
- automatique, 44, 45
- automatique et manuel, 28
- automatique, incomplet, 45
- automatique, relancer, 45
- changer de mode de traitement, 28
- contrôle, 46

exporter, 46

icônes, 28

interrompre le traitement automatique, 45

manuel, 46

par étapes, 46

relancer la reconnaissance de toutes les pages, 45

terminer, 45

ultérieurement, 32

True Page, mode, 65, 72, 78

TWAIN, 14

U

Un fichier après chaque page

blanche, option, 53

Unicode, exportation de texte, 92

Unité de mesure, 35

V

Vérification

différée, 32

OCR, boîte de dialogue, 66

options, 35, 42, 66

Vérifier

document ultérieurement, 32

mots absents du dictionnaire, 66

mots douteux, 66

résultats de l'OCR, 66, 67

Versions antérieures d'OmniPage, 13

Z

Zones

ajouter à, 59

ajouter des zones aux zones de

gabarit, 63

alphanumérique, 60

automatique, 60

automatiques, 35, 44, 57

colonne unique, 59

colonnes multiples, 60

contenu, 60

décharger un gabarit, 63

découpage en zones, 47

définir, 61

diviser, 59

gabarit, 63

gabarits, 58

graphique, 61

ignorer, 61

instructions, boîte de dialogue, 47

irrégulières, 26, 59

joindre, 59

manuelles, 26, 46, 58, 93, 95

modifier l'ordre, 76

modifier un gabarit, 63

numériques, 60

propriétés, 26, 60

raccourci, 61

réagencer, 26, 59

rectangulaires, 59

redimensionner, 26, 59

remplacer un gabarit, 63

réordonner, 76

soustraire de, outil, 59

supprimer un gabarit, 63

tableau, 60, 61

types, 60

Zoom, 26