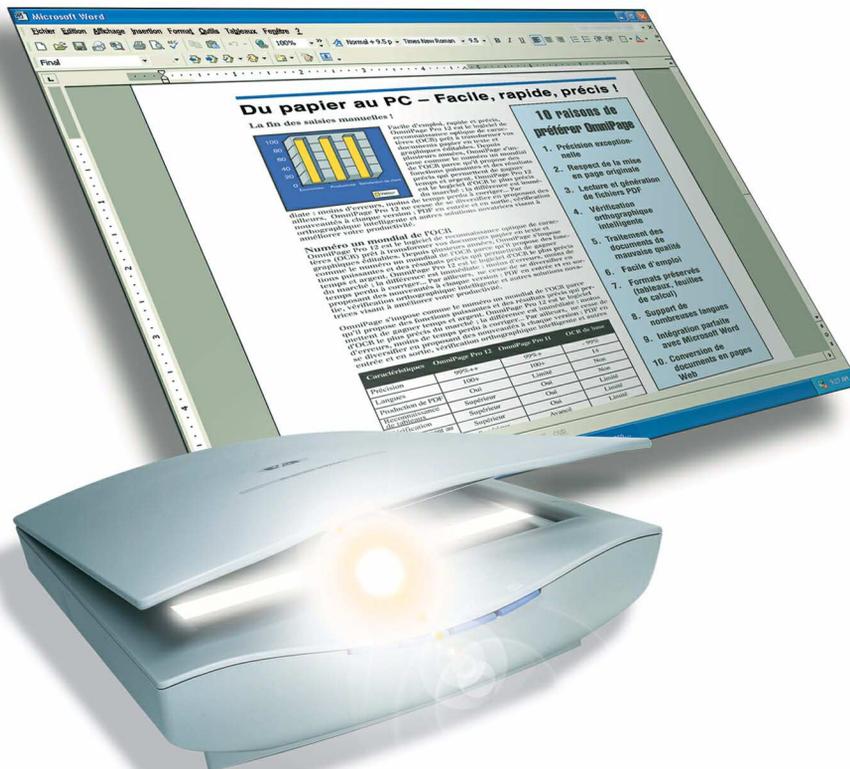


ScanSoft®



OmniPage® SE

Manuel de l'utilisateur

AVIS JURIDIQUES

Copyright © 2002 ScanSoft, Inc. Tous droits réservés.

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 garantie d'aucune sorte, expresse ou implicite, y compris, mais 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 SE, OmniPage Pro, PaperPort, Pagis, True Page et Direct OCR 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 société ou de produit cités peuvent être des marques de leurs propriétaires respectifs.

ScanSoft, Inc.

9 Centennial Drive
Peabody, MA 01960
États-Unis

ScanSoft Belgium BVBA

Guldensporenpark 32
BE-9820 Merelbeke
Belgique

Numéro de référence : 58-281201-01A



T A B L E D E S M A T I È R E S

	PRÉSENTATION GÉNÉRALE	7
	Utilisation de ce manuel	8
	Consultation de l'aide en ligne	9
	Aide HTML en ligne	9
	Aide contextuelle	9
	Notes techniques	10
	Glossaire	10
	OmniPage SE	10
1	INSTALLATION ET CONFIGURATION	11
	Environnement requis	12
	Installation de OmniPage SE	13
	Configuration d'un scanner pour utilisation avec OmniPage SE	14
	Lancement de OmniPage SE	16
	Enregistrement de OmniPage SE	17
	Nouvelles fonctions d'OmniPage Pro 12 OmniPage SE et OmniPage Pro 12	17 19
2	INTRODUCTION	21
	Reconnaissance optique de caractères (OCR)	22
	Technologie OCR d'OmniPage SE	22
	Documents dans OmniPage SE	23
	Principales étapes de l'OCR	23
	Bureau OmniPage	24
	Barre des menus	25

	Barres d'outils	25
	Fenêtre d'image	26
	Éditeur de texte	26
	Boîte à outils OmniPage	27
	Gestion de documents	28
	Vignettes	28
	Gestionnaire de documents	29
	Personnaliser des colonnes dans le gestionnaire de documents	30
	Supprimer des pages d'un document	30
	Imprimer un document	31
	Fermer un document	31
	Documents OmniPage	31
	Avantages du format OPD	32
	Enregistrer au format OPD	32
	Réglages	33
3	TRAITEMENT DES DOCUMENTS	35
	Guide de démarrage rapide	36
	Charger et reconnaître les exemples de fichiers image	36
	Numériser et reconnaître un document d'une page	36
	Présentation du traitement	38
	Traitement automatique	40
	Arrêter et relancer le traitement automatique	41
	Traitement manuel	42
	Traitement combiné	43
	Traitement à l'aide de l'Assistant OCR	45
	Traitement à partir d'autres applications	46
	Configuration de Direct OCR	47

Utilisation de Direct OCR	47
Utilisation de OmniPage SE avec PaperPort	48
Traitement de documents avec Planifier OCR	49
Définition de la source des images de page	50
Importer des fichiers image	50
Importer des documents à partir d'un scanner	51
Numériser avec un chargeur ADF	53
Numériser sans chargeur ADF	53
Définition de la mise en page du document	54
Zones et fonds	55
Découpage en zones automatique	56
Traçage de zones manuel	56
Types et propriétés de zone	58
Gestion des zones	60
Gestion des grilles du tableau dans l'image	62
Utilisation des gabarits de zone	64
4 VÉRIFICATION ET ÉDITION	67
Modes d'affichage de l'éditeur de texte	68
Vérification des résultats de l'OCR	69
Vérification du texte reconnu	70
Dictionnaires utilisateur	72
Apprentissage	73
Apprentissage manuel	74
IntelliTrain	74
Fichiers d'apprentissage	75
Édition du texte et des images	77
Édition en direct	79
Lecture du texte à voix haute	80

5	ENREGISTREMENT ET EXPORTATION	83
	Enregistrement des images originales	84
	Enregistrement des résultats de l'OCR	85
	Enregistrement d'un document lors d'une session de travail	86
	Sélection d'un niveau de formatage	87
	Sélection des options d'enregistrement évoluées	88
	Enregistrement d'un document au format PDF	90
	Copie des pages dans le Presse-papiers	90
	Ajout de pages à un e-mail comme pièce jointe	91
6	INFORMATIONS TECHNIQUES	93
	Dépannage	94
	Solutions de premier recours	94
	Test de OmniPage SE	95
	Augmentation de la quantité de mémoire disponible	96
	Augmentation de l'espace disque	96
	Reconnaissance de texte incorrecte	97
	Reconnaissance de fax incorrecte	98
	Problèmes de système ou de performances pendant l'OCR	98
	Support ODMA	99
	Fonctions évoluées du module Planifier OCR	99
	Types de fichier supportés	100
	Types de fichier pour l'ouverture et l'enregistrement des images	100
	Types de fichier pour l'enregistrement des résultats de l'OCR	102
	Désinstallation du logiciel	103



Présentation générale

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

Manuel de référence

Le présent manuel fournit des informations relatives aux principes d'utilisation d'OmniPage SE (Special Edition) : 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 manuel est proposé au format PDF ; vous pouvez ainsi utiliser les outils de navigation de votre afficheur PDF et les liens hypertexte intégrés pour passer aux différentes sections du document.

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. Voir « Consultation de l'aide en ligne », page 9.

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 renseignements divers

Le site Web de ScanSoft (www.scansoft.com) fournit des informations actualisées au sujet du programme. La section « Scanner Guide » contient des données régulièrement mises à jour et relatives aux scanners supportés et à d'autres questions y afférentes ; ScanSoft effectue des tests sur les 25 modèles les plus répandus. 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 de ce 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. Reportez-vous, si nécessaire, 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>	Indique le nom des rubriques de l'aide en ligne. Indique le texte développé des options dans les boîtes de dialogue.
Police sans empattement	Indique les noms de fichier, tels que exemple.tif
	Présente une remarque qui introduit des informations supplémentaires.
	Présente un conseil qui introduit des suggestions sur l'accomplissement de tâches spécifiques à l'aide des fonctions du logiciel.
	L'icône SE, placée en marge de certaines sections, vous permet d'identifier rapidement les différences entre OmniPage SE et OmniPage Pro 12. Voir « OmniPage SE », page 10.

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 ou 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. Cette combinaison affiche l'aide relative aux commandes des menus contextuels.



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.com) propose des notes techniques (Tech Notes) concernant les problèmes les plus courants concernant l'utilisation d'OmniPage®. Ces pages Web peuvent également inclure des conseils sur l'installation et le dépannage du logiciel.

Glossaire

Ce manuel ne comporte pas de glossaire. En revanche, l'aide en ligne inclut un glossaire exhaustif, assorti d'un index alphabétique et d'une table des matières. Consultez-le pour trouver la définition d'un terme utilisé dans ce manuel ou dans 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 12. 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 12. Voir « OmniPage SE et OmniPage Pro 12 », page 19.

Chapitre 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
- ◆ Installation de OmniPage SE
- ◆ Configuration d'un scanner pour utilisation avec OmniPage SE
- ◆ Lancement de OmniPage SE
- ◆ Enregistrement de OmniPage SE
- ◆ Nouvelles fonctions d'OmniPage Pro 12
- ◆ OmniPage SE et OmniPage Pro 12

Environnement requis

Pour installer et utiliser OmniPage SE 2.0, vous devez prévoir l'environnement minimum suivant :

- ◆ Processeur Pentium ou supérieur
- ◆ Microsoft Windows 98 (à partir de la Deuxième Édition), Windows Me, Windows NT 4.0 (avec au moins Service Pack 6), Windows 2000 ou Windows XP
- ◆ 64 Mo de mémoire RAM (128 Mo recommandés)
- ◆ 90 Mo d'espace disque disponible destiné aux fichiers d'application, plus 5 Mo d'espace disponible lors de l'installation
- ◆ 5 Mo pour Microsoft Installer (MSI) si absent de votre système (ce module est inclus dans le système d'exploitation Windows Me, Windows 2000 et Windows XP)
- ◆ Affichage SVGA avec 256 couleurs. Couleur 16 bits et résolution de 800 x 600 pixels recommandées [65536 couleurs dans Windows 2000 et Moyenne (16 bits) dans XP].
- ◆ Souris (ou équivalent) compatible Windows
- ◆ Lecteur de CD-ROM (pour l'installation)
- ◆ Scanner compatible avec le pilote approprié (si vous prévoyez de numériser des documents). Pour consulter la liste des scanners supportés, reportez-vous au guide des scanners (Scanner Guide) sur le site Web de ScanSoft (www.scansoft.com).



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

Installation de 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 :

- ◆ Fermez les applications actives, en particulier les logiciels antivirus.
- ◆ Si vous installez OmniPage SE sur un système Windows NT, Windows 2000 ou Windows XP, ouvrez une session avec privilèges d'administrateur.
- ◆ Si une version antérieure d'OCR de ScanSoft 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 version du module Texte-parole et la langue de l'interface du programme. L'interface du programme affiche dans la langue choisie des éléments tels que les commandes de menu, les boîtes de dialogue, les fenêtres de message, etc. Vous pouvez modifier la langue d'interface ultérieurement, via les options du logiciel, mais la sélection effectuée lors de l'installation détermine la version du module Texte-parole installée. 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.



Certains problèmes peuvent être résolus en désinstallant puis en réinstallant OmniPage SE. Voir « Désinstallation du logiciel », page 103.



Dans OmniPage Pro 12, 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. Voir « Lecture du texte à voix haute », page 80.

Configuration d'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, avec le pilote approprié, et tester son fonctionnement. Les pilotes du scanner ne sont pas fournis avec OmniPage SE.

L'installation et la configuration du scanner s'effectuent via l'Assistant du scanner, que vous pouvez lancer vous-même, en appliquant les procédures décrites. Sinon, l'Assistant du scanner apparaît la première fois que vous tentez une numérisation.

Pour configurer votre scanner avec OmniPage SE à l'aide de l'Assistant du scanner, suivez les étapes suivantes :

- ◆ Choisissez Démarrer ▶ Programmes ▶ ScanSoft OmniPage SE 2.0 ▶ Assistant du scanner.

ou cliquez sur le bouton Configuration dans l'onglet Scanner de la boîte de dialogue Options.

ou choisissez un paramètre de numérisation dans la zone de liste déroulante Importer une page de la boîte à outils et cliquez sur le bouton Importer une page.

L'Assistant du scanner démarre. La première fenêtre s'affiche uniquement lors de la configuration initiale, à partir d'OmniPage SE.

- ◆ Choisissez Sélectionner un scanner ou un appareil photo numérique, puis cliquez sur Suivant. La liste de tous les pilotes de scanner TWAIN détectés s'affiche ; le scanner par défaut du système est sélectionné.
- ◆ Cliquez une fois pour sélectionner le pilote de scanner voulu. Si vous devez rechercher un pilote, cliquez sur Autres pilotes. Si vous voulez masquer l'interface de votre scanner au cours de la numérisation ou modifier la méthode de transfert des images, sélectionnez Configurer les paramètres évolués pour ouvrir une fenêtre supplémentaire. Cliquez sur Suivant.
- ◆ Choisissez Oui pour tester la configuration de votre scanner, puis cliquez sur Suivant. L'Assistant vérifie la connexion de l'ordinateur au scanner. À l'issue de l'opération, cliquez sur Suivant.

- ◆ Insérez une page de test dans le scanner. L'Assistant est prêt à effectuer une numérisation de base au moyen du logiciel fourni par le fabricant du scanner. Cliquez sur Suivant. L'interface utilisateur native de votre scanner s'affiche.
- ◆ Cliquez sur Numériser pour lancer la numérisation de la page de test.
- ◆ Si nécessaire, cliquez sur Image inversée ou Image manquante et effectuez les sélections appropriées.
- ◆ Si l'image affichée dans la fenêtre est correcte, cliquez sur Suivant.
- ◆ Sélectionnez l'élément qui décrit le mieux votre scanner, puis cliquez sur Suivant.
- ◆ Cliquez sur Suivant pour passer à la sélection de la taille de page.
- ◆ Les tailles de page que l'Assistant du scanner attribue à votre scanner sont indiquées dans la fenêtre. Pour modifier les tailles de page, cliquez sur Évolué, effectuez les changements voulus et cliquez sur Suivant.
- ◆ Insérez dans le scanner une page contenant du texte, mais pas d'image. Cliquez sur Suivant pour lancer une numérisation en noir et blanc.
- ◆ Si nécessaire, cliquez sur Image inversée ou Image manquante et effectuez les sélections appropriées.
- ◆ Si l'image affichée dans la fenêtre est correcte, cliquez sur Suivant.
- ◆ Si vous disposez d'un scanner couleur, insérez une photographie couleur ou une page comportant une image couleur. Cliquez sur Suivant pour lancer une numérisation en couleur. Si nécessaire, cliquez sur Image inversée ou Image manquante et effectuez les sélections appropriées. Si l'image affichée dans la fenêtre est correcte, cliquez sur Suivant. Si votre scanner ne gère pas la couleur, ignorez cette étape.
- ◆ Insérez dans le scanner une photographie ou une page contenant une image. Cliquez sur Suivant pour lancer une numérisation en niveaux de gris. Si nécessaire, cliquez sur Image inversée ou Image manquante et effectuez les sélections appropriées. Si l'image affichée dans la fenêtre est correcte, cliquez sur Suivant.
- ◆ Vous venez de configurer votre scanner pour utilisation avec OmniPage SE. Cliquez sur Terminer.

Pour modifier les réglages du scanner par la suite ou pour configurer un autre scanner, ouvrez de nouveau l'Assistant de configuration du scanner depuis le menu Démarrer de Windows ou depuis l'onglet Scanner de la boîte de dialogue Options. Pour tester et dépanner un scanner ne fonctionnant pas correctement, ouvrez l'Assistant du scanner depuis le menu Démarrer de Windows et

sélectionnez Tester scanner/appareil photo numérique dans la première fenêtre, puis appliquez la procédure décrite ci-dessus.

Lancement de 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 2.0 ▶ OmniPage SE 2.0.
- ◆ 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. Voir « Documents OmniPage », page 31.

Lorsque vous ouvrez l'application, l'écran d'accueil puis le bureau OmniPage SE s'affichent. Voir « Bureau OmniPage », page 24. Cette section présente les principales zones de travail du programme.

Vous pouvez également exécuter le programme avec une interface limitée :



- ◆ Ouvrez le programme Planifier OCR. Dans la barre des tâches Windows, cliquez sur Démarrer et choisissez Programmes ▶ ScanSoft OmniPage Pro 12.0 ▶ Planifier OCR. Voir « Traitement de documents avec Planifier OCR », page 49. Cette fonction n'est pas disponible dans OmniPage SE.
- ◆ Dans une application déclarée dans Direct OCR™, choisissez Acquérir texte dans le menu Fichier. Voir « Configuration de Direct OCR », page 47.
- ◆ 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.
- ◆ Pour ajouter des services d'OCR, utilisez OmniPage SE avec les produits de gestion de documents de ScanSoft, PaperPort® ou Pagis®. Voir « Utilisation de OmniPage SE avec PaperPort », page 48.

Enregistrement de OmniPage SE

L'Assistant d'enregistrement ScanSoft est lancé à l'issue de l'installation du programme. 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 accomplir l'enregistrement.

Si vous n'enregistrez pas votre copie du logiciel pendant son installation, le programme vous invitera à intervalles réguliers à accomplir cette formalité. Vous pouvez visiter le site Web (www.scansoft.com) pour vous enregistrer en ligne. Cliquez sur Support, puis dans l'écran de support principal, choisissez Enregistrer dans la colonne de gauche.

Pour plus de détails sur l'utilisation de vos données d'enregistrement, reportez-vous à la clause relative à la confidentialité de la politique de ScanSoft.

Nouvelles fonctions d'OmniPage Pro 12

La famille de produits OmniPage® compte désormais OmniPage Pro 12 et OmniPage SE. Cette section présente les améliorations apportées à OmniPage Pro 12. Comme l'indique la section suivante, certaines de ces fonctionnalités sont également comprises dans OmniPage SE. Nouvelles fonctionnalités de OmniPage Pro 12 par rapport à OmniPage Pro 11:

- ◆ **Importante augmentation de la précision**
De nouveaux niveaux de précision avec l'amélioration de la synergie entre les moteurs de reconnaissance, le support des dictionnaires professionnels et l'apprentissage possible des caractères choisis par l'utilisateur.
- ◆ **Rationalisation de l'interface**
Vous pouvez effectuer le traitement automatique et manuel directement à partir de la boîte à outils OmniPage, sans avoir recours à des barres d'outils séparées (voir la page 27). Les vignettes apparaissent dorénavant dans la fenêtre d'image ; vous pouvez choisir d'afficher la page courante, la page des vignettes ou les deux (voir page 28). La vue Détails de la version précédente a été remplacée par le gestionnaire de documents, qui inclut une colonne de remarques pour les commentaires et les mots-clés explorables.
- ◆ **Nouveaux concepts de zones**
La fonction de découpage en zones « en direct » permet de modifier des zones « à la volée », sans devoir relancer la reconnaissance de la page

entière (voir page 79). Le fond des pages est affecté du type Traiter (découpage en zones automatique) ou Ignorer ; par conséquent l'intégralité des instructions de découpage en zones apparaît sur la page et vous pouvez enregistrer celles-ci dans des gabarits de zone (voir page 55). Les procédures à suivre pour tracer des zones irrégulières, ainsi que diviser et joindre des zones, ont été simplifiées et elles n'exigent aucun outil séparé (voir page 60).

◆ **Amélioration de la vérification**

La boîte de dialogue Vérification de l'OCR affiche maintenant les mots douteux dans un contexte étendu. Un vérificateur dynamique peut rester ouvert pendant la vérification du texte, tandis que l'affichage et la fenêtre d'image suivent le contexte d'édition (voir page 69).

◆ **Niveaux de formatage pour l'affichage et l'enregistrement**

L'affichage dans l'éditeur de texte accepte trois niveaux de formatage (voir page 68). Le niveau de formatage de sortie est désormais sélectionné au moment de l'exportation ; les sélections possibles dépendent du type de fichier spécifié. L'option d'exportation Pages en continu est une version améliorée de la vue Conserver l'enchaînement des colonnes de la version précédente. Elle conserve la mise en page, si possible sans cadres et blocs de texte, pour la disposition du texte entre les colonnes (voir page 87).

◆ **Meilleure analyse des pages**

Le transfert du formatage des tableaux a été amélioré, en particulier la détection des tableaux dépourvus de traits de grille dans les pages originales. Les adresses Web et e-mail peuvent être détectées et transférées vers l'éditeur de texte ; vous pouvez également insérer des hyperliens. L'ordre de lecture peut maintenant être affiché et modifié après reconnaissance dans la vue True Page® de l'éditeur de texte. Voir à partir de page 77.

◆ **Meilleur traitement PDF**

OmniPage Pro 12 recherche le texte de fond dans les PDF qu'il ouvre, afin de garantir une précision de reconnaissance supérieure. Un nouveau type de fichier, PDF édité, assure une bonne conservation du format sur les pages modifiées dans l'éditeur de texte à l'issue de la reconnaissance.

◆ **Options d'enregistrement évoluées**

La gamme des options de sauvegarde a été enrichie pour chaque type de fichier de sortie. Les types de fichier de sortie définis par l'utilisateur peuvent être créés sur la base de réglages personnalisés (voir page 88). Si votre version d'OmniPage Pro 12 inclut les nouveaux formats d'enregistrement XML et eBook, voir page 102.

OmniPage SE et OmniPage Pro 12



La liste suivante répertorie les fonctionnalités de la version professionnelle d'OmniPage Pro 12 qui n'ont pas été ajoutées à OmniPage SE :

- ◆ Meilleur niveau de reconnaissance.
- ◆ Accès au module d'apprentissage, IntelliTrain et aux fichiers d'apprentissage 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.
- ◆ Numérisation en vis-à-vis pour numériser les livres en un clin d'œil.
- ◆ Des fonctionnalités de formatage des pages en continu permettant une conservation parfaite de la mise en page.
- ◆ Planifiez l'OCR pour que vos travaux soient traités de façon automatique et selon un horaire défini.
- ◆ Gestion des formats d'image TIFF LZW et GIF en entrée et sortie.
- ◆ Exportation aux formats eBook et XML.
- ◆ Prise en charge de la sortie WYSIWYG HTML 4.0.
- ◆ Nombre de langues reconnues passant de 50 à plus de 100. Accès à des dictionnaires juridiques et médicaux dans certaines langues.
- ◆ Accès au module Texte-parole *RealSpeak*, permettant la lecture à voix haute des textes reconnus.

Pour obtenir une mise à niveau, consultez le site www.scansoft.com.

Chapitre 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 ordinateur, vous devez trouver un moyen de le faire, en évitant la saisie manuelle, source d'erreur et synonyme de perte de temps.

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.

Ce chapitre propose les sections suivantes :

- ◆ Reconnaissance optique de caractères (OCR)
 - Documents dans OmniPage SE
 - Principales étapes de l'OCR
- ◆ Bureau OmniPage
- ◆ Gestion de documents
- ◆ Documents OmniPage
- ◆ 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.



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.

Documents 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. Voir « Bureau OmniPage », page 24.

Principales étapes de l'OCR

OmniPage SE propose deux modes de traitement des documents : automatique et manuel. Voir « Traitement automatique », page 40 et « Traitement manuel », page 42. 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 sous la forme de vignettes dans la fenêtre d'image ; l'image de la première page est affichée. Les informations relatives aux pages concernées s'affichent dans le gestionnaire de documents. Voir « Définition de la source des images de page », page 50.

2. Exécuter l'OCR pour obtenir du texte éditable

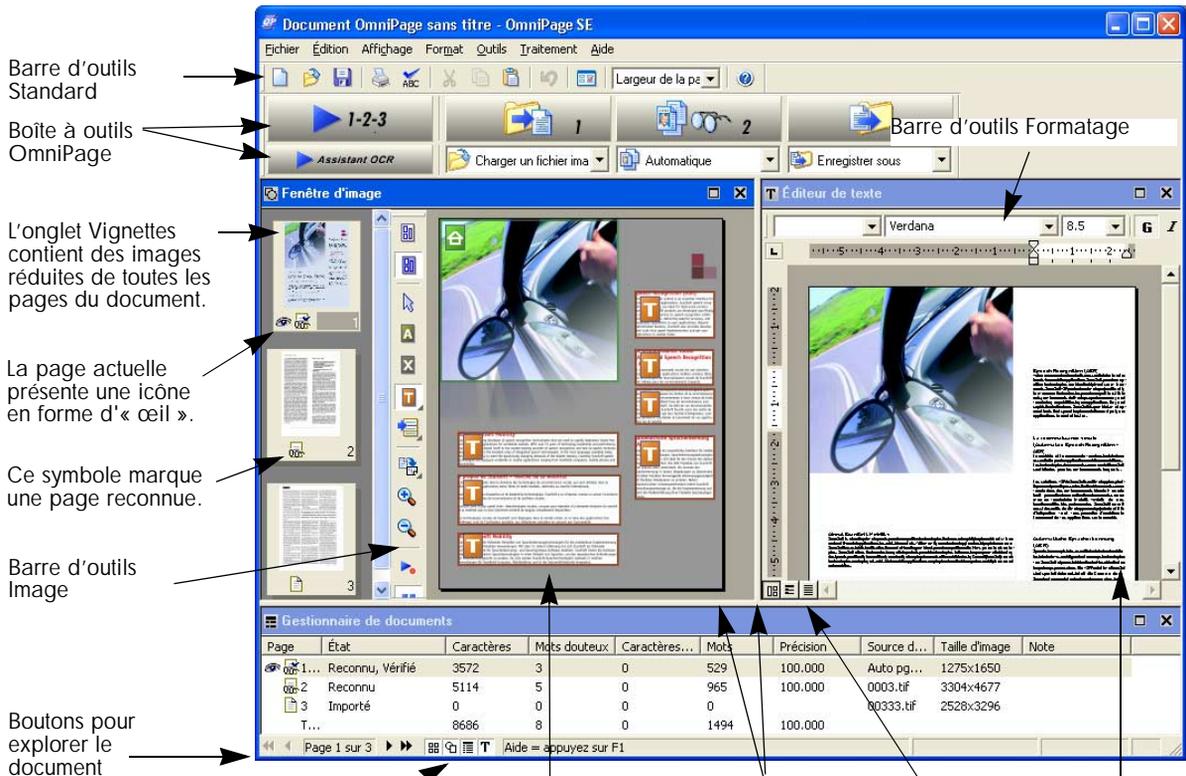
Pendant la reconnaissance, OmniPage SE trace des **zones** autour des parties de l'image à traiter, puis interprète le contenu textuel et graphique de chaque zone. Vous pouvez également créer des zones manuellement ou en appliquant un gabarit de zones prédéfini. À l'issue de la reconnaissance, vous pouvez 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 de document OmniPage (OPD) selon la procédure décrite ci-après. Un même document peut être enregistré plusieurs fois avec différents réglages d'emplacement, de type de fichier, de formatage, etc., selon vos besoins. Voir « Enregistrement et exportation », page 83.

Bureau OmniPage

Le bureau OmniPage 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 documents, fenêtre d'image et éditeur de texte. Dans leur partie supérieure, ces trois espaces de travail comprennent les boutons de fermeture, d'agrandissement et de réduction de la fenêtre. La Fenêtre d'image est munie de la barre d'outils Image et la fenêtre Éditeur de texte de la barre d'outils Formatage.



Boutons permettant d'afficher ou de masquer le gestionnaire de documents, l'éditeur de texte et l'affichage des vignettes et de la page actuelle pour la fenêtre d'image. Vous pouvez également effectuer ces opérations depuis le menu Affichage.

Fenêtre d'image :
 Cette fenêtre affiche l'image de la page actuelle et indique les zones tracées. La fenêtre d'image peut afficher la page actuelle, la page des vignettes ou les deux.

Faites glisser ces barres de fractionnement pour redimensionner les zones de travail.

Boutons de l'éditeur de texte offrant trois niveaux de formatage.

Éditeur de texte :
 Affiche le résultat de la reconnaissance de la page actuelle en mode True Page.

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 la partie inférieure de chaque vignette indique l'état de la page correspondante.

Les boutons de la barre d'état vous permettent d'afficher et de masquer 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 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.

Barres d'outils

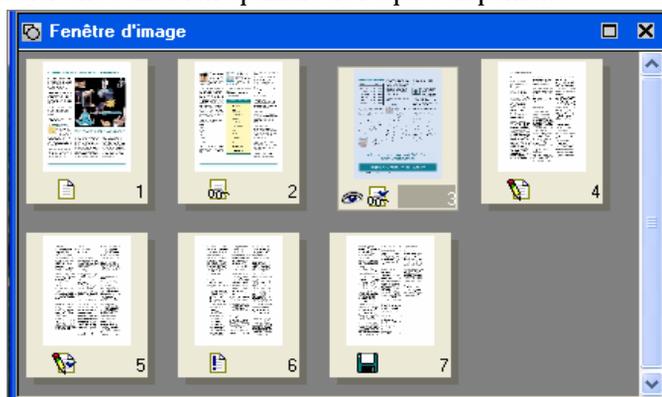
Le programme propose trois barres d'outils principales, qui peuvent toutes être flottantes, si besoin. Le menu Affichage permet de les afficher, de les masquer et de les personnaliser. L'aide contextuelle explique l'action de chaque outil. Deux barres d'outils supplémentaires permettent de piloter des tâches spécifiques.

Barre d'outils	Emplacement par défaut	Autres emplacements d'ancrage	Action
Standard	Horizontalement sous la barre de menus	Tout bord du bureau OmniPage	Exécution des fonctions de base du programme. Voir page 31 et page 69.
Image	Verticalement à gauche de la page d'image actuelle	Verticalement à droite de la page d'image actuelle	Opérations sur les images, les zones et les tableaux. Voir page 55 et page 62.
Formatage	Horizontalement en haut de l'éditeur de texte	Aucun	Formatage du texte reconnu dans l'éditeur de texte. Voir page 77.
Vérificateur	Placez le curseur sur la fenêtre du vérificateur pour afficher cette barre d'outils flottante.		Contrôle de l'emplacement et de l'aspect du vérificateur. Voir page 70.
Réagencement	Cliquez sur l'outil <i>Changer l'ordre de lecture</i> . Cette barre d'outils remplace la barre d'outils <i>Formatage</i> .		Modification de l'ordre des éléments dans les pages reconnues. Voir page 77.

Fenêtre d'image

Lorsque cette fenêtre affiche l'image de page actuelle, la barre d'outils Image est disponible. Les images de page possèdent toutes une valeur de fond : Traiter ou Ignorer. Les zones peuvent être tracées manuellement sur les images de page ou placées automatiquement à l'issue de la reconnaissance. Il existe cinq types de zones : Traiter, Ignorer, Texte, Tableau et Graphique. Les zones sont automatiquement tracées pour les éléments placés à l'intérieur des zones Traiter et sur un fond Traiter en dehors d'autres zones, et leurs types de zones sont déterminés pendant le traitement. Voir « Zones et fonds », page 55.

Si l'image de page actuelle est masquée, les vignettes apparaissent en rangées pour assurer une utilisation optimale de l'espace disponible.



Éditeur de texte

L'éditeur de texte affiche les résultats de la reconnaissance selon trois niveaux de formatage :

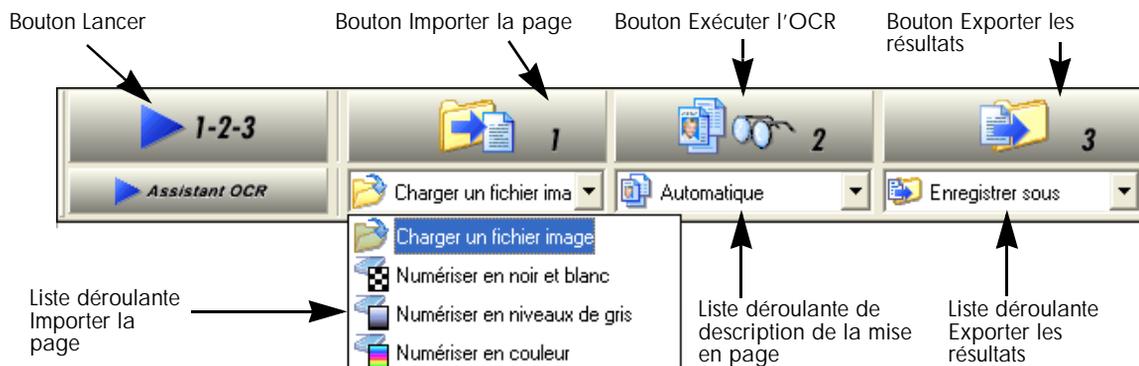
- ◆ Aucun formatage
- ◆ Conserver le formatage polices/paragraphes
- ◆ True Page

Le mode True Page conserve la mise en page sur la base des cadres et des blocs de texte, de tableaux et d'images. True Page peut afficher des zones multicolonnées, afin de présenter des blocs de texte pouvant être assimilés à un enchaînement de colonnes au moment de l'exportation. **True Page** correspond également à un niveau de formatage d'exportation, de même que **Pages en continu**, qui conserve la mise en page sans blocs ni cadres. Voir « Modes d'affichage de l'éditeur de texte », page 68. OmniPage SE ne prend pas en charge la fonction de formatage des Pages en continu.



Boîte à outils OmniPage

La boîte à outils d'OmniPage SE vous permet de gérer toutes les étapes du traitement. Elle est placée par défaut dans la partie supérieure du bureau OmniPage, 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.

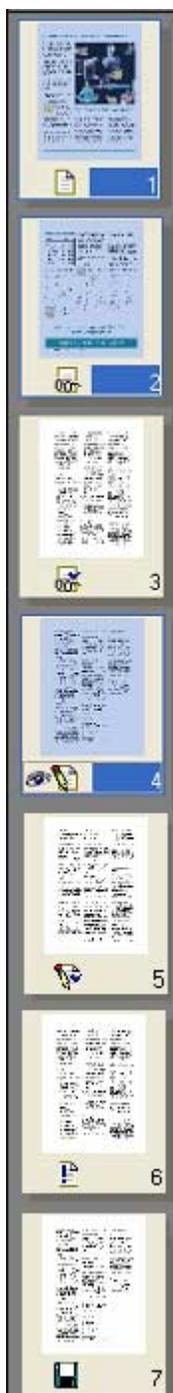


Le **traitement automatique** est lancé, arrêté et relancé à l'aide du bouton Lancer (1-2-3). Voir « Traitement automatique », page 40.

Le **traitement manuel** permet de traiter les documents page par page et étape par étape. Vous pouvez démarrer chacune des étapes avec le bouton correspondant : Importer la page (1), Exécuter l'OCR (2) et Exporter les résultats (3). Voir « Traitement manuel », page 42.

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. Voir « Traitement combiné », page 43.

L'**Assistant OCR** est conçu pour les nouveaux utilisateurs. Voir « Traitement à l'aide de l'Assistant OCR », page 45. Si un document est ouvert lorsque vous lancez l'Assistant OCR, le programme vous invite à l'enregistrer, puis le ferme. Si vous utilisez l'Assistant OCR 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.



Gestion de documents

La gestion de documents s'effectue à partir des vignettes de la fenêtre d'image ou du gestionnaire de documents (partie inférieure du bureau OmniPage). Les deux fenêtres comportent des informations succinctes sur les pages du document traité et sont synchronisées. Les illustrations utilisées dans ce manuel présentent ces deux fenêtres avec le même document de sept 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 d'image). L'état de la page est indiqué comme suit :

Page	État	Icône	Opérations effectuées sur la page
1	Importée		Page importée, mais pas encore reconnue.
2	Reconnue		Page reconnue, non vérifiée ou vérification interrompue.
3	Reconnue, vérifiée		Page reconnue et entièrement vérifiée.
4	Modifiée		Page reconnue, avec exécution dans l'éditeur de texte d'au moins une modification ou un changement de formatage.
5	Modifiée, vérifiée		Page reconnue, modifiée dans l'éditeur de texte et entièrement vérifiée.
6	En instance		Page importée et peut-être reconnue, avec certains changements de type de zone stockés, mais pas traités.
7	Enregistrée		Page reconnue et enregistrée au moins une fois.

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 présente une icône en forme d'« œil ». Vous pouvez sélectionner plusieurs pages dans un document ; toute page sélectionnée présente un aspect « distinctif ». Utilisez les vignettes pour les opérations de page, comme suit :

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. Vous pouvez aussi exporter les pages sélectionnées.



Pour afficher des informations sur une image importée, placez le curseur sur la vignette correspondante (les bulles d'aide doivent être activées). 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.

Gestionnaire de documents

Cette fonction fournit 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 définies par l'utilisateur.

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

Page	État	Caractères	Mots douteux	Caractères...	Mots	Précision	Note
1	Importé	0	0	0	0		Notes
2	Reconnu	4047	33	6	719	99.852	Intro
3	Reconnu, Vérifié	1185	0	0	187	100.000	diagramme
4	Modifié	7102	3	0	1168	100.000	menu
5	Modifié, Vérifié	812	0	0	141	100.000	tableau
6...	En instance	7102	3	0	1168	100.000	Europe
7	Enregistré	4822	9	0	619	100.000	Asie
		25070	48	6	4202	99.976	

Entrez des commentaires ou des mots-clés consultables.

La page actuelle présente une icône en forme d'« œil ». Dans le gestionnaire de documents, vous pouvez effectuer les opérations suivantes :

Pour passer à une page, cliquez sur la partie gauche de la rangée de pages ou double-cliquez dans sa rangée.

Pour modifier l'ordre d'apparition d'une page, cliquez sur sa rangée et faites-la glisser à l'emplacement requis. Un indicateur 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.

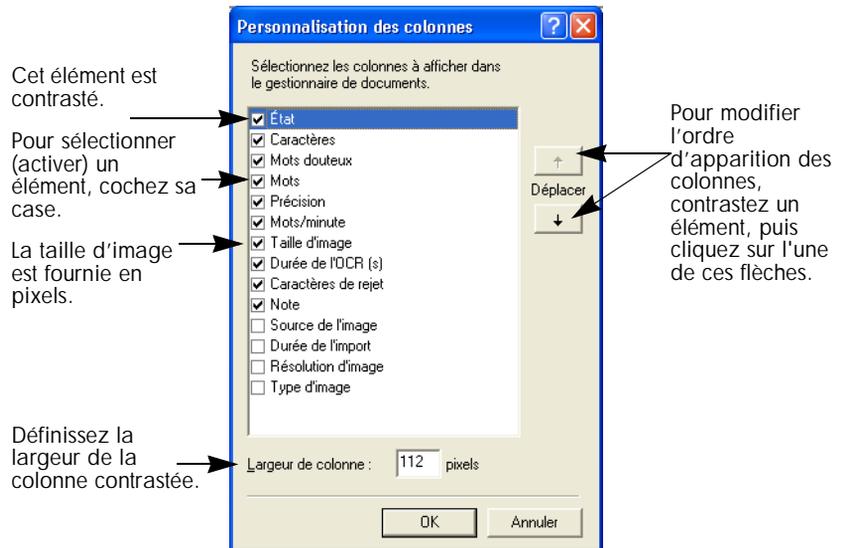
Pour sélectionner plusieurs pages adjacentes, maintenez enfoncée la touche Maj. et cliquez sur la première, puis sur la dernière rangée du groupe requis. Pour sélectionner plusieurs rangées non adjacentes, maintenez enfoncée la

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. Vous pouvez aussi exporter les pages sélectionnées.

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

Personnaliser des colonnes dans le gestionnaire de documents

Vous pouvez spécifier les colonnes que vous souhaitez inclure dans le gestionnaire de documents. Dans le menu Affichage, cliquez sur Personnaliser les colonnes 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 *Personnaliser les colonnes du gestionnaire de documents* 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 documents ou dans la page des vignettes, appuyez sur la touche Suppr ou utilisez la commande Effacer du 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. L'outil Imprimer de la barre d'outils Standard envoie à l'imprimante des images ou du texte, en fonction de la fenêtre active (fenêtre d'image 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 de document OmniPage (*.OPD), reportez-vous à la section qui suit. Un message vous demande également d'enregistrer les nouvelles données d'apprentissage, si vous avez sélectionné l'option Demander l'enregistrement des données d'apprentissage à la fermeture des documents 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 (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 (corrigée et prétraitée) 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 de document OmniPage, le logiciel sauvegarde également les réglages en cours (et les nouvelles données d'apprentissage). Lors de l'ouverture d'un document OmniPage, ses réglages sont appliqués, remplaçant les paramètres du programme.



Un document OmniPage créé et enregistré dans OmniPage SE n'inclut pas de données d'apprentissage. Si vous ouvrez un fichier OPD contenant des données d'apprentissage dans OmniPage SE, ces données seront ignorées.

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 au cours de la session en cours.
- ◆ Vous souhaitez transmettre le document à d'autres utilisateurs disposant d'OmniPage SE ou Pro (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 accessible. (le texte reconnu permet d'effectuer des recherches par mot-clé et accepte d'autres techniques de récupération de documents).



Avant d'installer une mise à jour d'OmniPage, vous devez enregistrer les résultats de la reconnaissance stockés dans les fichiers OPD. En effet, ces fichiers ne seront peut-être pas compatibles avec les nouveaux formats de fichier OPD ; il se peut aussi que seules les images soient conservées. Lorsque vous ouvrez un fichier OPD créé dans OmniPage Pro 10, seules les images sont chargées. Lorsque vous ouvrez un fichier OPD créé dans OmniPage Pro 11 et son édition spéciale, les images et les pages reconnues sont chargées, mais les zones ne sont pas conservées.

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 sauvegardez 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 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.

La barre de titre indique le nom du fichier dans lequel le document entier a été enregistré pour la dernière fois.

Réglages



Les réglages OmniPage SE sont regroupés dans la boîte de dialogue Options. Vous pouvez y accéder depuis la barre d'outils Standard ou le menu Outils. 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 ou professionnel, des caractères de rejet et le mappage des polices. Pour sélectionner ou désélectionner une langue, cochez ou décochez la case correspondante. Vous pouvez effectuer plusieurs sélections, mais limitez-vous aux seules langues apparaissant dans le document à reconnaître. Les éléments indiqués en premier correspondent aux dernières langues sélectionnées. Pour afficher directement une langue, entrez ses premières lettres. OmniPage SE ne prend pas en charge les dictionnaires professionnels.



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. Voir « Importer des documents à partir d'un scanner », page 51. L'onglet Scanner ne s'affiche pas si vous aviez spécifié l'affichage de l'interface TWAIN native de votre scanner lors de la configuration de celui-ci. Voir « Configuration d'un scanner pour utilisation avec OmniPage SE », page 14.

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. Voir « Configuration de Direct OCR », page 47.

Traitement

Permet de définir l'emplacement d'insertion de nouvelles images, d'activer le message demandant la suite des pages lors de la numérisation, de spécifier la numérisation sur deux pages pour le traitement des livres et de définir d'autres

paramètres. Cet onglet vous donne aussi la possibilité de modifier la langue d'interface. OmniPage SE ne prend pas en charge la numérisation en vis-à-vis.



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. Voir « Vérification des résultats de l'OCR », page 69. 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). Voir « Définition de la mise en page du document », page 54.

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. Voir « Édition du texte et des images », page 77.

Dans OmniPage Pro 12, si vous pouvez accéder à un système DMS à partir de votre ordinateur et que votre édition d'OmniPage Pro 12 inclut le support ODMA, une fenêtre ODMA peut également d'afficher. Voir « Support ODMA », page 99.



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 activés immédiatement dans les pages existantes (par exemple, retour à la ligne automatique ou unité de mesure dans l'éditeur de texte).

Chapitre 3



Traitement des documents

Ce chapitre 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
- ◆ Présentation du traitement
- ◆ Traitement automatique
- ◆ Traitement manuel
- ◆ Traitement combiné
- ◆ Traitement à l'aide de l'Assistant OCR
- ◆ Traitement à partir d'autres applications (Direct OCR, PaperPort)
- ◆ Traitement de documents avec Planifier OCR

Les instructions détaillées portent sur les opérations suivantes :

- ◆ Définition de la source des images de page
- ◆ Définition de la mise en page du document
- ◆ Zones et fonds
 - Découpage en zones automatique
 - Traçage de zones manuel
 - Types et propriétés de zone
 - Gestion des zones
- ◆ Gestion des grilles du tableau dans l'image
- ◆ Utilisation des gabarits de zone

Guide de démarrage rapide

Cette section vous présente toutes les étapes de la procédure d'OCR de base.



Charger et reconnaître les exemples de fichiers image

Les exemples de fichiers image (comportant une seule page et multipages) sont mis à votre disposition dans le dossier Programmes. Traitez ces fichiers en appliquant la procédure décrite ci-dessous, mais dans un premier temps ignorez toute référence au scanner. Voir « Importer des fichiers image », page 50. Les résultats vous donneront 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.



Numériser et reconnaître un document d'une page

Mettez votre scanner sous tension et assurez-vous qu'il fonctionne correctement. Pour ce test, choisissez une page au texte net.

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 le menu Outils) et choisissez Valeurs par défaut.

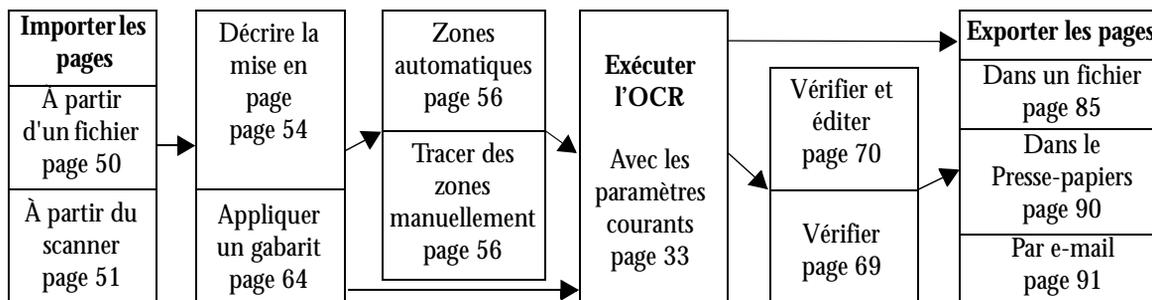
Le document sera traité automatiquement et les résultats de l'OCR seront enregistrés dans un fichier. Vous allez effectuer la vérification, mais non l'édition, du document dans l'éditeur de texte.

	Action	Résultat
1.	Le cas échéant, configurez le scanner à l'aide de l'Assistant du scanner.	OmniPage SE est configuré pour l'utilisation avec votre scanner.
2.	Sélectionnez Démarrer ► Programmes ► ScanSoft OmniPage SE 2.0 ► OmniPage SE 2.0.	Le programme OmniPage SE est ouvert.
3.	Placez le document soigneusement dans le scanner.	
4.	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.
5.	Dans le menu déroulant de description de la mise en page, 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.
6.	Dans le menu déroulant Exporter les résultats, vérifiez que l'option Enregistrer sous est sélectionnée.	Vous pouvez attribuer un nom au fichier d'exportation après la vérification du document.
7.	Cliquez sur Lancer.	OmniPage SE lance la numérisation de votre document. Une vignette apparaît, assortie d'une jauge de progression. La fenêtre Vérification de l'OCR apparaît.
8.	Utilisez la fenêtre Vérification de l'OCR pour modifier les mots que le programme considère comme reconnus incorrectement.	Le module 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 la reconnaissance. Les marques de vérification sont supprimées une fois que vous avez terminé l'opération.
9.	Cliquez dans la fenêtre de l'éditeur de texte. Sélectionnez les différents modes de l'éditeur de texte pour comparer les modifications d'affichage de la page.	Chaque mode de l'éditeur de texte définit un niveau de formatage différent. Vous pouvez ainsi choisir le niveau voulu lors de l'enregistrement.
10.	Cliquez sur Reprendre pour relancer la vérification. Lorsque le message <i>La vérification de l'OCR est terminée</i> , s'affiche, cliquez sur OK.	La vérification de la reconnaissance est terminée. La boîte de dialogue Enregistrer sous apparaît.
11.	Choisissez le nom et le type de fichier, ainsi que le chemin et le niveau de formatage requis pour l'enregistrement du document reconnu. Cliquez sur OK.	Par défaut, la fonction Enregistrer et lancer est activée et votre document est ouvert automatiquement dans le programme de traitement de texte associé au type de fichier sélectionné.
12.	Vérifiez 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.

Si vous obtenez de bons résultats lors de la reconnaissance des exemples de fichiers image, mais non de la page numérisée, vérifiez l'installation et la configuration de votre scanner, notamment la luminosité et la résolution d'image. Voir « Importer des documents à partir d'un scanner », page 51. Vous obtiendrez un modèle de réglage de luminosité optimal. Consultez également les rubriques *Configurer le scanner* et *Dépanner le scanner* de l'aide en ligne.

Présentation du traitement

Le diagramme suivant récapitule les principales étapes :



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.

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 boîte à outils OmniPage, puis cliquez sur **Lancer**. 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.



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, cliquez sur l'un des trois boutons correspondants dans la boîte à outils OmniPage.

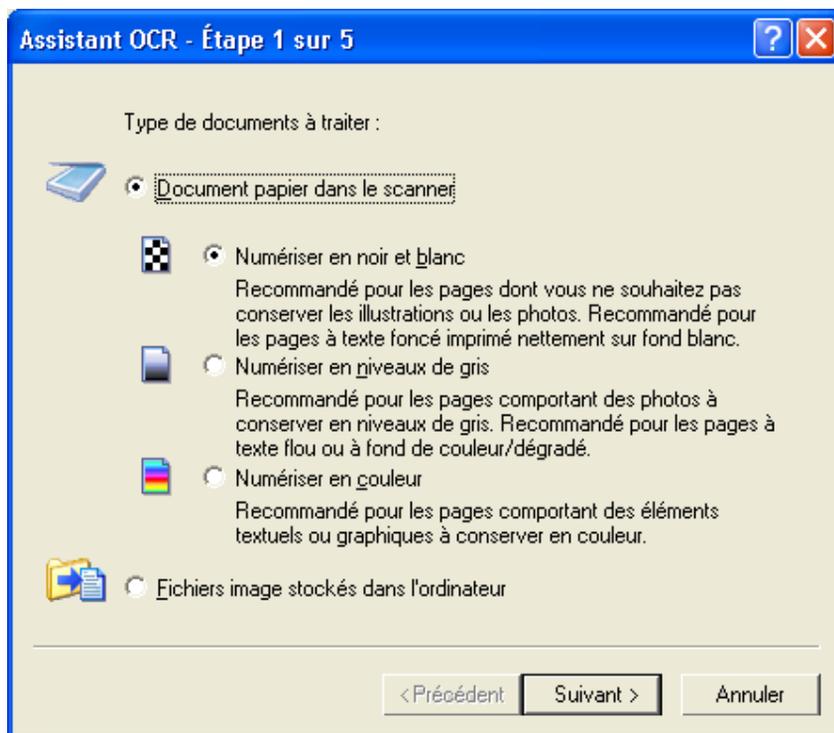
Combiné

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. Vous pouvez aussi importer les images manuellement, tracer des zones sur tout ou partie d'entre elles, puis appliquer à toutes les pages le traitement automatique.



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 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.

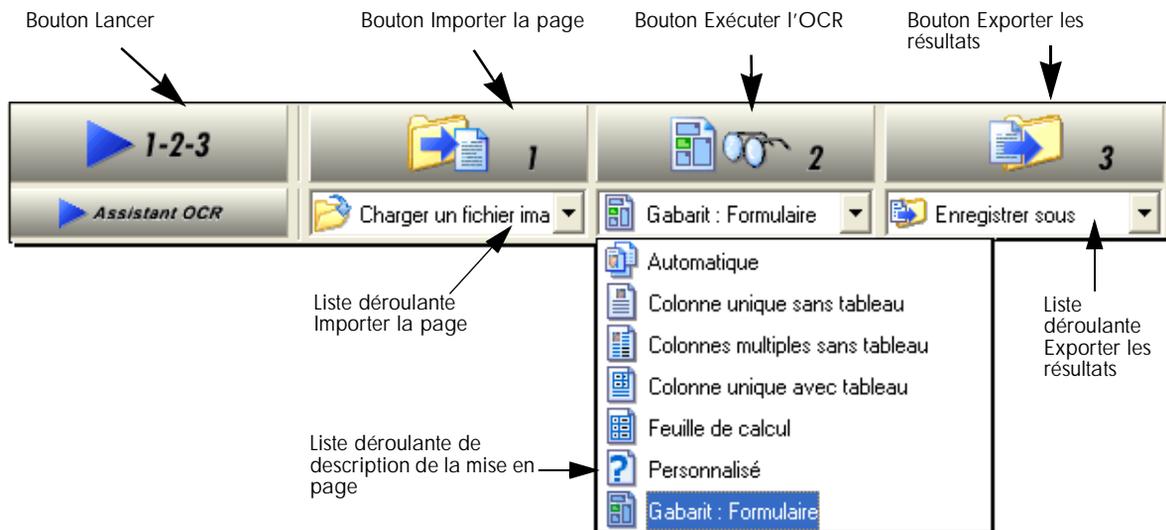
Traitement différé (Planifier OCR)

Le module Planifier OCR permet de différer la reconnaissance automatique des documents (par exemple, à une heure où vous n'utilisez pas l'ordinateur). Pour spécifier les réglages et l'heure de lancement du traitement, utilisez l'Assistant de nouveau travail de ce module. OmniPage SE ne prend pas en charge la fonction Planifier OCR.



Traitement automatique

Le mode automatique représente une méthode efficace pour le traitement de documents, notamment de documents volumineux. Vous définissez les réglages pertinents, puis utilisez le bouton Lancer 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. Dans la liste déroulante du bouton Importer la page, sélectionnez la commande requise. Vous définissez la source du document (fichier image ou scanner). Voir « Définition de la source des images de page », page 50.
2. Dans la liste déroulante de description de la mise en page, 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. Voir « Définition de la mise en page du document », page 54.
3. Dans la liste déroulante du bouton Exporter les résultats, sélectionnez la commande requise. Vous pouvez enregistrer le document reconnu dans un fichier au format document OmniPage. Vous pouvez enregistrer les pages voulues (actuelle, sélectionnées, toutes) dans un fichier, les copier dans le Presse-papiers ou les envoyer comme pièce jointe d'un courrier électronique. Voir « Enregistrement et exportation », page 83.

4. Choisissez la commande  dans la barre d'outils Standard ou 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. Voir « Réglages », page 33.
5. Cliquez sur Lancer ou choisissez Lancer le traitement automatique dans le menu Traitement. Les pages du document sont traitées les unes après les autres ; 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.

Arrêter et relancer le traitement automatique



Arrêter : Lorsque le traitement automatique est en cours, le bouton Arrêter remplace le bouton Lancer. Cliquez sur ce bouton pour arrêter l'opération. Cette commande est utile lorsque vous souhaitez modifier certains réglages alors que le traitement est déjà entamé.

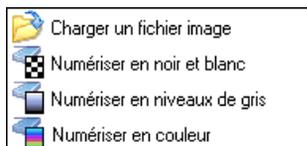


Relancer : Lorsque le traitement automatique est terminé, le bouton Lancer réapparaît. Cliquez sur ce bouton pour relancer le traitement. La boîte de dialogue Traitement automatique vous permet d'effectuer les actions suivantes :

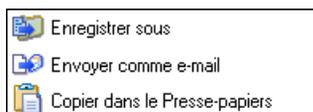
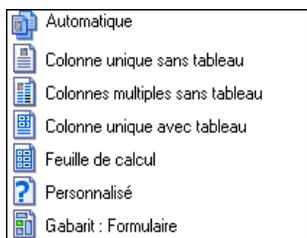
- ◆ **Terminer** le traitement des pages non reconnues et non vérifiées, puis exporter les résultats.
- ◆ **Exporter** un document déjà enregistré, avec ou sans modification, sous un autre nom, emplacement ou type de fichier, ou avec un autre niveau de formatage.
- ◆ **Ajouter des pages** de la même source ou d'une autre source, en conservant ou modifiant les réglages.
- ◆ **Recommencer le traitement de toutes les pages** pour éliminer tous les résultats de l'OCR et relancer la reconnaissance de toutes les pages du document avec des réglages différents. Vous pouvez activer le découpage en zones automatique ou appliquer un fichier de gabarit. Le relancement du traitement est utile 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, le chargement ou le redécoupage en zones de toutes les images.

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 le fond de chaque page. Pour lancer une étape, vous cliquez sur l'un des trois boutons correspondants dans la boîte à outils OmniPage.



1. Cliquez sur la commande  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. Voir « Réglages », page 33.
2. Dans la liste déroulante du bouton Importer la page, sélectionnez l'option requise. Vous définissez la source du document (fichier image ou scanner). Si vous effectuez une numérisation, sélectionnez le mode voulu et effectuez les réglages pertinents dans les fenêtres Scanner et Traitement de la boîte de dialogue Options. Voir « Définition de la source des images de page », page 50.
3. Cliquez sur le bouton Importer la page. Cette action affiche la boîte de dialogue Charger un fichier image (permettant d'indiquer les noms des fichiers image voulus) ou lance la numérisation. Des vignettes de chaque page apparaissent dans la fenêtre d'image, avec l'image de page actuelle. Les boutons de la barre d'état vous permettent d'afficher ou de masquer ces images. Les informations relatives aux pages concernées s'affichent dans le gestionnaire de documents.
4. Les images de page importées dans le programme sont toujours dotées de la valeur de fond Traiter. Si vous ne tracez aucune zone sur ces pages, elles seront soumises à un découpage en zones automatique lors du lancement de la reconnaissance.
5. Vous pouvez tracer et modifier manuellement des zones sur une ou plusieurs images, et leur attribuer des propriétés. Les boutons de la barre d'état vous permettent de passer d'une page à l'autre. Dès que vous tracez une zone sur une page, son fond reçoit la valeur Ignorer. Tracez des zones à valeur Traiter pour définir le découpage en zones automatique sur des parties d'une page. Voir « Zones et fonds », page 55.



6. Dans la liste déroulante du bouton Exécuter l'OCR, sélectionnez l'option requise. Choisissez la mise en page du document original. La valeur spécifiée est prise en compte dans le processus de traçage de zones automatique. Voir « Définition de la mise en page du document », page 54. Vous pouvez également sélectionner un gabarit pour appliquer ses zones à la page actuelle. Voir « Utilisation des gabarits de zone », page 64.
7. Cliquez sur le bouton Exécuter l'OCR pour reconnaître la page actuelle. Pour reconnaître plusieurs pages, sélectionnez le groupe de pages requis à l'aide des vignettes ou dans le gestionnaire de documents (Voir « Gestion de documents », page 28.) et cliquez sur le bouton Exécuter l'OCR. Les pages reconnues s'affichent dans l'éditeur de texte.
8. 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 « Vérification des résultats de l'OCR », page 69.
9. Continuez à charger les pages, à exécuter l'OCR et à procéder aux opérations d'édition et de vérification selon vos besoins. Vous pouvez changer l'ordre de lecture des éléments de page dans l'éditeur de texte. Voir « Édition du texte et des images », page 77.
10. Dans la liste déroulante du bouton Exporter le texte reconnu, sélectionnez l'option requise. Vous pouvez enregistrer le document reconnu au format document OmniPage. Vous pouvez enregistrer les pages voulues (actuelle, sélectionnées, toutes) dans un fichier, les copier dans le Presse-papiers ou les envoyer comme pièce jointe d'un courrier électronique. Cliquez sur le bouton Exporter les résultats. Voir « Enregistrement et exportation », page 83.

Traitement combiné

Le traitement automatique garantit rapidité et efficacité. S'il exige une plus grande attention, le traitement manuel assure un meilleur contrôle sur les résultats. Vous pouvez combiner ces deux avantages lorsque vous traitez un seul document.

Pour lancer une reconnaissance automatique et terminer l'opération en mode 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. Ajustez certains paramètres et/ou modifiez des zones, puis passez au mode manuel pour relancer 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 et supprimez ou modifiez les zones existantes dans la fenêtre d'image. Vous pouvez également charger un gabarit pour appliquer ses zones à la page actuelle. Tracez les nouvelles zones selon vos besoins. Voir « Zones et fonds », page 55.
4. Modifiez les autres réglages en fonction des particularités de la page actuelle. Voir « Réglages », page 33.
5. 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. Vous pouvez aussi choisir d'appliquer le traitement en direct pour appliquer les modifications apportées aux zones sans devoir relancer la reconnaissance de la page entière. Voir « Édition en direct », page 79.
6. Pour relancer la reconnaissance de plusieurs pages, sélectionnez le groupe de pages voulu dans la page des vignettes ou dans le gestionnaire de documents avant de cliquer sur le bouton Exécuter l'OCR.
7. Lorsque toutes les pages soumises au nouveau traitement fournissent un résultat satisfaisant, réenregistrez le document.

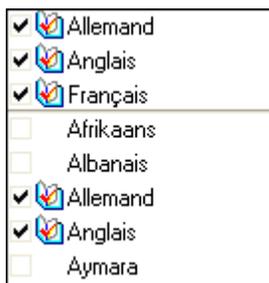
Pour lancer une reconnaissance manuelle et terminer l'opération en mode automatique :

1. Cliquez sur le bouton Importer la page afin de définir les réglages et d'importer les images pour le document.
2. Examinez les pages pour en déterminer la luminosité, l'orientation et le contenu. Renumériser ou faites pivoter les images insatisfaisantes. Réordonnez les pages selon vos besoins.

3. Tracez manuellement les zones dans les pages dont vous ne voulez traiter qu'une partie ou pour lesquelles vous voulez fournir des instructions de zones précises. Appliquez la valeur Ignorer aux fonds ou aux zones que vous voulez exclure du traitement. Appliquez la valeur Traiter aux fonds ou aux zones pour lesquels vous voulez définir le découpage en zones automatique.
4. Cliquez sur le bouton Lancer, puis choisissez *Terminer le traitement des pages existantes* dans la boîte de dialogue Traitement automatique.
5. À l'issue de la vérification (le cas échéant), vous pouvez enregistrer ou exporter le document.

Traitement à l'aide de l'Assistant OCR

Vous pouvez utiliser l'Assistant OCR pour lancer le traitement d'un nouveau document. Si un document est ouvert lorsque vous sélectionnez l'Assistant OCR, ce dernier se ferme. En vous guidant à travers cinq 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. Pour afficher le premier écran de l'Assistant, cliquez sur le bouton Assistant OCR dans la boîte à outils OmniPage :



1. La première fenêtre vous permet de définir la source du document : scanner ou fichier image. Voir « Définition de la source des images de page », page 50. 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. Voir « Définition de la mise en page du document », page 54.
3. La troisième fenêtre vous permet de définir les langues de reconnaissance. Les langues associées à un dictionnaire sont accompagnées de l'icône de livre ouvert. Les sélections les plus récentes apparaissent au début de la liste.
4. La quatrième 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. Voir « IntelliTrain », page 74. La référence au module IntelliTrain ne s'applique pas à OmniPage SE.

5. 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.
6. 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. Une fois la vérification terminée ou arrêtée, la boîte de dialogue Copier dans le Presse-papiers ou Enregistrer sous apparaît pour vous permettre de spécifier les réglages d'exportation pour le fichier, notamment un intervalle de pages et un niveau de formatage.
7. 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 d'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.



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 cliquer sur *Lancer* dans la boîte à outils OmniPage.



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.

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.

Configuration de 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. Sélectionnez les options de traitement pour la vérification et le découpage en zones. Les réglages effectués s'appliquent à l'ensemble des opérations futures de Direct OCR jusqu'à leur éventuelle modification ; ils ne sont pas pris en compte si vous utilisez OmniPage SE sans Direct OCR.
3. 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.

Utilisation de 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 Configuration d'acquisition du texte dans le menu Fichier de l'application cible 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 Configuration d'acquisition du 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 poursuivre 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.



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.



Pour plus de détails, reportez-vous aux rubriques relatives à Direct OCR dans l'aide en ligne, notamment à *Direct OCR – Questions et réponses*. Le fichier ReadMe/ LisezMoi et le site Web de ScanSoft peuvent proposer des informations plus récentes sur certaines applications cibles spécifiques.

Utilisation de 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 doté d'OmniPage SE, les services d'OCR de ce dernier étendent les fonctionnalités de PaperPort. Pour choisir un programme d'OCR, cliquez avec le bouton droit sur le lien PaperPort d'application de texte, sélectionnez Préférences, puis spécifiez OmniPage SE 2.0 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 cet exemple, 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 le lien Word de PaperPort. 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 un nouveau document sans titre de l'application cible.

Traitement de documents avec Planifier OCR



OmniPage SE ne prend pas en charge la fonction Planifier OCR. Le texte suivant ne s'applique qu'à OmniPage Pro.

Le module Planifier OCR permet de différer le traitement automatique de vos travaux d'OCR (dans la limite de huit jours 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 Pro 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 Pro 12.0 ▶ Planifier OCR.
2. La boîte de dialogue Planifier OCR apparaît. Cliquez sur Nouveau pour accéder à l'Assistant de nouveau travail. 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 : scanner avec chargeur ADF ou fichier.
4. Les deux fenêtres suivantes sont similaires aux écrans de l'Assistant OCR, mais vous pouvez en outre spécifier un dictionnaire utilisateur ou spécialisé et un fichier d'apprentissage. L'activation de la fonction IntelliTrain dépend des réglages actifs dans OmniPage Pro au moment du lancement du travail.
5. Les fenêtres suivantes vous permettent de spécifier le nom, le type et l'emplacement du fichier d'exportation, ainsi qu'une option de séparation de fichier et un niveau de formatage.
6. La dernière fenêtre permet de définir l'heure du lancement du travail et (le cas échéant) l'heure d'arrêt, ainsi que la conservation ou la suppression des fichiers d'origine après traitement. Pour fermer l'Assistant, cliquez sur Terminer.



La boîte de dialogue Planifier OCR affiche tous les travaux planifiés et indique leur état : En attente, En cours, En pause, Erreur ou Terminé. Pour modifier les réglages d'un travail en attente, sélectionnez Modifier le travail. Vous pouvez afficher, 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.

Définition de 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 fichier image compatibles avec OmniPage SE est très étendue. Voir « Types de fichier pour l'ouverture et l'enregistrement des images », page 100. Sélectionnez *Charger un fichier image* dans la liste déroulante Importer la page. Les fichiers figurent dans la boîte de dialogue Charger un fichier image, qui apparaît au moment du lancement d'un traitement automatique. Lors d'un traitement manuel, cliquez sur le bouton Importer la page 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.

Dossier actuel.

Pour placer plusieurs fichiers dans la zone Nom de fichier, utilisez les séquences Maj.+clac ou Ctrl+clac.

Spécifiez le ou les types de fichiers à afficher.

Cette option est destinée aux fichiers TIFF, DCX et MAX multipages.

Blank.tif est un fichier vide destiné à l'option d'enregistrement Un fichier après chaque page blanche.

Sélectionnez cette option pour afficher une vignette du fichier sélectionné. Elle n'est pas disponible lorsque plusieurs fichiers sont sélectionnés.

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

Permet d'ajouter des fichiers à partir de plusieurs dossiers et de contrôler avec précision leur ordre d'intégration.

Utilisez ces flèches pour modifier l'ordre des fichiers.

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. Voir « Configuration d'un scanner pour utilisation avec OmniPage SE », page 14. Le programme offre plusieurs modes de numérisation que vous choisissez en fonction des deux facteurs suivants :

- ◆ 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. Elle s'applique uniquement aux scanners couleur. 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 n'offre pas de 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 ou dans l'interface de votre scanner. 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 Créer un fichier après chaque page blanche, insérez également des pages blanches dans la pile. Voir « Enregistrement des résultats de l'OCR », page 85.

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 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.



OmniPage Pro 12 vous permet de numériser simultanément deux pages d'un livre. Le programme divise les images entrantes en deux pages distinctes qu'il redresse séparément. Cette fonction n'est pas proposée dans OmniPage SE.

Définition de 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 ne contenant aucune zone ou fond à valeur Traiter. Lors d'un traitement manuel, le traçage de zones automatique est parfois activé. Reportez-vous à la rubrique *Quand exécuter le découpage en zones automatique* de l'aide en ligne. Pour décrire le document original, vous disposez des options suivantes :



Automatique

Choisissez cette option pour laissez à OmniPage SE le soin d'effectuer le découpage en zones automatique. Le programme détermine si l'élément est du texte en colonnes, un graphique, du texte à reconnaître ou un tableau à conserver. Choisissez l'option Automatique si le document contient des pages dont la mise en page est hétérogène ou de type inconnu. Sélectionnez cette option pour une page à colonnes multiples et comportant un tableau, et pour toutes pages comportant plus d'un tableau.



Colonne unique sans tableau

Choisissez cette option si les pages du document contiennent une seule colonne de texte et aucun tableau, comme c'est le cas des circulaires et pages de livre. Sélectionnez cette option pour une page dont les mots et nombres sont agencés en colonnes et si vous ne souhaitez pas les placer dans un tableau ni les séparer en plusieurs colonnes distinctes. Cette option permet la détection des graphiques.



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). Pour conserver les colonnes au moyen de cadres dans le document de sortie, sélectionnez le format True Page lors de l'exportation. La fonction de formatage des pages en continu d'OmniPage Pro 12 permet également de conserver les colonnes sans cadres. En cas de données tabulaires, celles-ci seront le plus probablement traitées comme une colonne unique. Cette option permet la détection des graphiques.



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. Cette option permet la détection des graphiques.



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 une valeur de fond, avec ses zones et propriétés, à toutes les pages importées. Les zones de gabarit sont également appliquées à la page actuelle, en remplacement des zones existantes, ainsi qu'aux pages préexistantes sans zones au moment de leur (nouvelle) reconnaissance. Voir « Utilisation des gabarits de zone », page 64.

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.

Zones et fonds

Les zones définissent les sections de la page destinées à être traitées ou ignorées. Elles peuvent prendre une forme rectangulaire ou irrégulière (avec des côtés verticaux et horizontaux). Les images de page d'un document possèdent une valeur de fond : Traiter ou Ignorer (cette dernière étant la plus répandue). Vous pouvez utiliser les outils indiqués pour modifier les valeurs de fond. Vous pouvez également utiliser les outils indiqués pour tracer des zones sur les fonds de page :

Fonds	Traiter	Ignorer			
Zones	Traiter	Ignorer	Texte	Tableau	Graphique

Traiter – le document est découpé automatiquement en zones (dans les zones ou les fonds de type Traiter) lors de son envoi pour reconnaissance.

Ignorer – les zones (dans les zones et fonds de type Ignorer) sont exclues du traitement. Le texte n'est pas reconnu et aucune image n'est transférée.

Découpage en zones automatique

Le découpage en zones automatique permet au programme de détecter les éléments de page (blocs de texte, en-têtes, images et autres) et de les inclure dans une zone. OmniPage attribue des types et des propriétés à ces zones. Le découpage en zones automatique s'applique aux pages entières dans le cas d'un traitement automatique, sauf si vous avez chargé un gabarit, et lors de l'utilisation de l'Assistant OCR. Vous pouvez également spécifier le découpage en zones automatique dans le cadre d'un traitement manuel, en procédant comme suit :

Découper automatiquement une page entière en zones

Importez une page ; celle-ci s'affiche avec un fond de type Traiter. Ne tracez aucune zone et vérifiez dans la liste déroulante de description de la mise en page qu'aucun gabarit de zone n'est chargé. Cliquez sur le bouton Exécuter l'OCR. Vous pouvez sélectionner plusieurs pages sans zones afin de les découper automatiquement et de les reconnaître.

Découper automatiquement en zones une partie de page

Importez une page ; celle-ci s'affiche avec un fond de type Traiter. Tracez une zone. La valeur de fond passe à Ignorer. Tracez les zones Texte, Tableau ou Graphique afin de délimiter manuellement les zones de type voulu. Tracez les zones à valeur Traiter afin de délimiter les zones à affecter du traitement automatique. À l'issue de la reconnaissance, les zones à valeur Traiter seront remplacées par une ou plusieurs zones Texte, Tableau ou Graphique.

Découper automatiquement un fond de page en zones

Importez une page ; celle-ci s'affiche avec un fond de type Traiter. Tracez une zone. La valeur de fond passe à Ignorer. Tracez les zones Texte, Tableau ou Graphique afin de délimiter manuellement les zones de type voulu. Cliquez sur l'outil Fond à valeur Traiter (voir illustration) afin de définir un fond de ce type. Tracez des zones à valeur Ignorer sur les parties inutiles de la page. À l'issue de la reconnaissance, la page réapparaît : elle comprend un fond à valeur Ignorer et de nouvelles zones encadrant tous les éléments du fond.

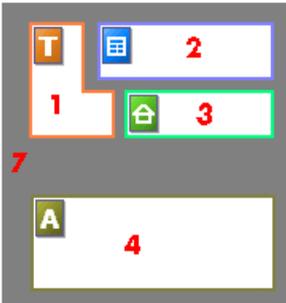


Traçage de zones manuel

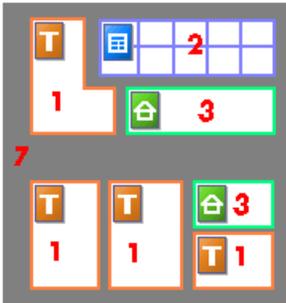
Cette section présente d'abord deux exemples relatifs aux zones et aux fonds, décrit ensuite les types de zone, et enfin, indique les méthodes à appliquer pour tracer et utiliser les zones. Pour ces exemples, les nombres renvoient au tableau de la page suivante.

Tracer des zones sur un fond à valeur Ignorer :

Avant la reconnaissance :



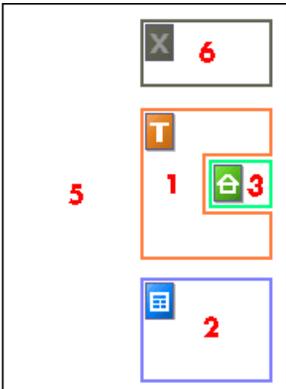
Après la reconnaissance :
Le fond conserve la valeur Ignorer.



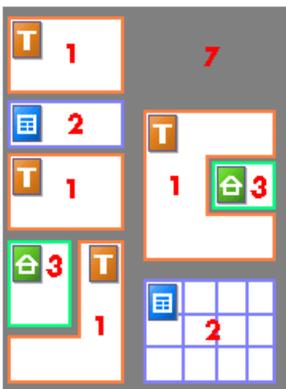
La zone 4 correspond à un groupe de zones, permettant en l'occurrence de gérer trois colonnes de texte et une photo.

Tracer des zones sur un fond à valeur Traiter :

Avant la reconnaissance :



Après la reconnaissance :
La valeur de fond passe à Ignorer.



La zone 6 est incluse dans le fond. Les zones figurant dans la partie gauche de la page ont été créées automatiquement.

Numéro	Type	Résultat
1	Zone Texte	L'OCR est exécuté et génère le texte.
2	Zone Tableau	L'OCR est exécuté et le texte est placé dans une grille de tableau.
3	Zone Graphique	L'image est incorporée dans la page reconnue.
4	Zone à valeur Traiter	Le découpage en zones automatique crée une ou plusieurs zones, leur attribue un type et en traite le contenu.
5	Fond à valeur Traiter	
6	Zone à valeur Ignorer	Aucun résultat
7	Fond à valeur Ignorer	

Les zones tracées automatiquement et les zones de gabarit ont des bordures pleines :



Les zones tracées manuellement et les zones modifiées ont des bordures en pointillés :



Aucun ordre de lecture n'est attribué aux zones. Vous pouvez réordonner les éléments reconnus dans l'éditeur de texte. Voir « Édition du texte et des images », page 77. Le chapitre 4 présente le traçage de zones « en direct ». Voir « Édition en direct », page 79.

Types et 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 effectuez les sélections souhaitées. Pour changer les propriétés de plusieurs zones simultanément, sélectionnez le groupe de zones requis (combinaison Maj.+clic) avant de choisir les nouvelles valeurs.

La barre d'outils Image propose cinq outils de traçage de zones (un pour chaque type). Le type d'une zone est indiqué par l'icône placée dans son angle supérieur gauche, ainsi que par la couleur de cette icône et celle de la bordure de la zone. La liste ci-après indique les outils et les couleurs disponibles :

Zone à valeur Traiter (vert olive)



Cet outil permet de tracer une zone à valeur Traiter, afin de définir une section qui sera soumise au traçage de zones automatique. À l'issue de la reconnaissance, cette zone est remplacée par une ou plusieurs zones dont les types sont déterminés automatiquement. Vous tracez normalement les zones à valeur Traiter sur un fond à valeur Ignorer. Tracez une zone à valeur Traiter pour entourer les colonnes de texte à traiter automatiquement. Le texte sera affiché sous une seule colonne dans le mode Aucun formatage et Conserver le formatage polices/paragraphes de l'éditeur de texte, mais les colonnes seront conservées en mode True Page.



Zone à valeur Ignorer (gris)

Cet outil permet de tracer une zone à valeur Ignorer, afin de définir une section à ne pas transférer vers l'éditeur de texte. La fonction de découpage en zones automatique omet ces zones pendant le traitement. Pour exclure une section donnée d'un grand nombre de pages (par exemple, un en-tête ou les numéros de page), placez une zone à valeur Ignorer dans un gabarit. Vous tracez normalement les zones à valeur Ignorer sur un fond à valeur Traiter.



Zone Texte (marron)

Cet outil permet de tracer une zone Texte sur un bloc de texte unique. Le contenu des zones est assimilé à un enchaînement de texte, sans aucune colonne. Pour assurer le traitement automatique des colonnes de texte, entourez-les d'une zone à valeur Traiter.



Zone Tableau (bleu)

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 éditable. Pour la plupart des types de fichier de sortie, vous pouvez exporter le tableau sous la forme de grille ou de colonnes de texte séparées par tabulation.



Zone Graphique (vert)

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 images incorporées peuvent être exportées avec le document vers les applications cibles supportant les graphiques.



Les zones Texte et Tableau sont affectées d'un réglage de contenu. 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 leur ponctuation. Toute lettre est ignorée. Le menu contextuel de la zone vous permet de modifier ce réglage.



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.

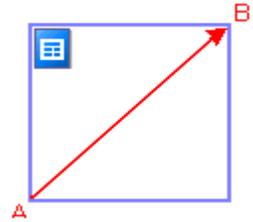
Gestion des zones



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. Certains de ces outils sont groupés. Dans ce cas, seul le dernier outil sélectionné dans le groupe est visible. Pour sélectionner un outil visible, cliquez dessus. Pour sélectionner un outil masqué, positionnez la souris sur le triangle placé dans l'angle inférieur droit de l'outil visible, appuyez sur le bouton de la souris et maintenez-le enfoncé jusqu'à ce que d'autres outils apparaissent, puis cliquez sur l'outil voulu.

Tracer une zone unique

Sélectionnez l'outil de traçage de zone du type voulu, puis cliquez et faites glisser le curseur. Dans ces exemples, cette procédure est illustrée par le déplacement de la flèche de A vers B. Vous pouvez aussi effectuer ce mouvement de l'angle supérieur gauche vers l'angle inférieur droit.



Vous pouvez uniquement tracer des zones rectangulaires ; les zones tracées (à l'exception des zones Tableau) peuvent être converties en zones irrégulières.

Pour **redimensionner** une zone, sélectionnez-la (en cliquant dessus), amenez le curseur sur un côté ou un angle, activez une poignée et amenez-la à l'emplacement voulu.



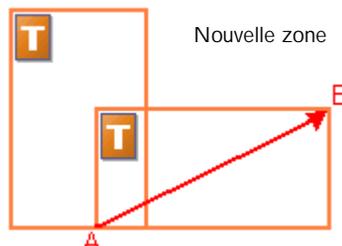
Pour **déplacer** une zone, sélectionnez-la à l'aide d'un outil de sélection de zone et déplacez-la selon vos besoins. Vous ne pouvez pas déplacer une zone pour la superposer sur une autre.

Créer une zone irrégulière par addition

Tracez une zone recouvrant partiellement une autre zone du même type :

Zone existante

Zone résultante



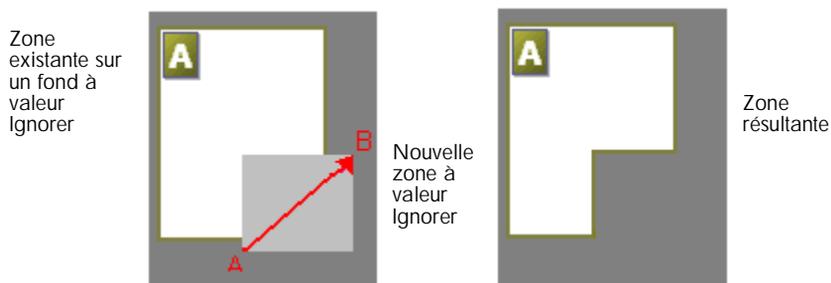
Joindre deux zones du même type

Tracez une zone recouvrant des deux zones du même type.



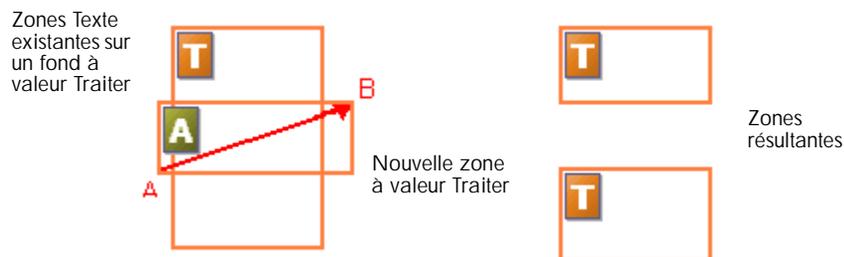
Créer une zone irrégulière par soustraction

Tracez une zone recouvrant une autre zone du même type que le fond. (Dans cet exemple, il y a un fond à valeur Ignorer).

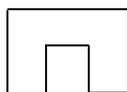


Diviser une zone

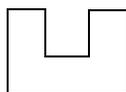
Tracez une zone du même type que le fond divisant une autre zone. (Dans cet exemple, il y a un fond à valeur Traiter).



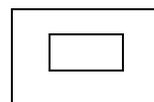
Les formes de zone suivantes sont interdites :



Retrait le long de la partie inférieure



Retrait le long de la partie supérieure



Avec trou

Pour agrandir une zone plus rapidement qu'au moyen des poignées de redimensionnement, entourez cette zone d'une zone du même type. La plus petite des zones est remplacée par la plus grande. Pour remplacer un groupe de zones (quel que soit le type) par une zone unique, entourez ces zones d'une zone plus grande du type voulu. Les zones plus petites sont toutes remplacées par la plus grande.

Lorsque la nouvelle zone tracée recouvre partiellement une zone existante d'un type différent, une superposition complète ne se produit pas ; la nouvelle zone remplace seulement la partie recouverte de la zone existante.

Pour plus de détails, reportez-vous à la rubrique *Tracer des zones manuellement* de l'aide en ligne.

Gestion des 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 gauche. Vous pouvez modifier le type affecté à une zone rectangulaire via son menu contextuel. Vous pouvez également tracer des zones de type Tableau, mais elles doivent toujours être rectangulaires.

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 tracer ou redimensionner une zone Tableau (à condition de conserver sa forme rectangulaire) pour éliminer les colonnes ou les rangées inutiles des bords extérieurs d'un tableau.

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. Si l'outil voulu n'est pas visible, cliquez sur le triangle placé dans l'angle inférieur droit de l'outil visible pour afficher tous les outils, puis cliquez sur l'outil voulu.

Utilisez les outils de tableau et leurs curseurs comme suit :



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 divisent un texte.



Insérer séparation de colonnes

Cliquez sur cet outil puis sur l'emplacement dans la zone Tableau où vous souhaitez placer une séparation de colonnes.



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 divisent un texte.



Supprimer séparation

Cliquez sur cet outil puis sur la séparation individuelle de rangée ou de colonne 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.



Insérer/Supprimer les séparations

Cliquez sur cet outil, puis sur son icône de curseur à l'intérieur d'une zone Tableau sans séparations. Les séparations sont détectées et placées automatiquement. Pour les éliminer, placez le curseur dans un tableau avec séparations et cliquez sur cet outil.

Pour placer, déplacer ou supprimer une séparation dans la cellule actuelle uniquement, utilisez la combinaison Ctrl+clic.

Vous pouvez spécifier le formatage de lignes pour les bordures et les grilles de tableau à partir d'un menu contextuel. Vous disposerez de davantage d'options d'édition des bordures et de fond dans l'éditeur de texte à l'issue de la reconnaissance.

Utilisation des gabarits de zone

Un fichier de gabarit contient une valeur de fond de page, un groupe de zones et leurs propriétés. Pour appliquer les zones d'un gabarit lors de la reconnaissance, chargez un fichier de gabarit de zone via la liste déroulante du bouton de description de la mise en page ou à l'aide du menu Outils.

Lorsque vous chargez un gabarit, son fond et ses zones sont insérées :

- ◆ Sur la page actuelle (les zones déjà insérées sont remplacées)
- ◆ Sur toutes les pages importées par la suite
- ◆ Sur les pages préexistantes envoyées pour (nouvelle) reconnaissance sans aucune zone

En mode de traitement manuel, vous pouvez afficher et modifier les zones de gabarit dans les deux premiers cas avant la reconnaissance.

En mode de traitement automatique, vous ne pouvez afficher et modifier les zones de gabarit qu'après la reconnaissance.

Ce comportement demeure tant que le gabarit est chargé.

Les gabarits acceptent les zones et les fonds à valeur Ignorer et Traiter. Par conséquent, ils peuvent être très utiles pour définir les parties de la page destinées au traitement à l'aide du découpage en zones automatique et les parties à ignorer. Pendant la reconnaissance, vous pouvez remplacer les zones ou les fonds à valeur Traiter par un groupe de zones plus petites, auxquelles seront attribués des types de zones spécifiques.



Enregistrer un gabarit de zones

Sélectionnez une valeur de fond et définissez les zones d'une page, puis vérifiez leur emplacement et leurs propriétés. Cliquez sur Gabarit de zones dans le menu Outils. Dans la boîte de dialogue affichée, sélectionnez [zones sur la page] et cliquez sur Enregistrer ; entrez le nom voulu et cliquez sur OK.

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 dans la liste déroulante de description de la mise en page du document. 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 de description de la mise en page, ou ouvrez la boîte de dialogue Fichiers de gabarit de zones, sélectionnez le gabarit souhaité et cliquez sur le bouton Définir comme actuel. Les zones du nouveau gabarit sont également appliquées à la page actuelle, en remplacement des zones existantes ; elles s'appliquent également comme indiqué ci-dessus.

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.



Les gabarits sont disponibles dans Direct OCR, mais non dans l'Assistant OCR.

Chapitre 4



Vérification et édition

Les résultats de la reconnaissance sont placés dans l'éditeur de texte. Il peut s'agir d'un texte reconnu, de tableaux et de graphiques incorporés. L'éditeur aux capacités d'affichage fidèle (« WYSIWYG ») offre les fonctions énumérées ci-dessous et décrites en détail dans le présent chapitre :

- ◆ Modes d'affichage de l'éditeur de texte
- ◆ Vérification des résultats de l'OCR
- ◆ Vérification du texte reconnu
- ◆ Dictionnaires utilisateur
- ◆ Apprentissage
- ◆ Édition du texte et des images
- ◆ Édition en direct
- ◆ Lecture du texte à voix haute

Modes d'affichage de l'éditeur de texte

L'éditeur de texte affiche le texte reconnu et, en fonction de vos réglages, signale les mots qui ont été considérés comme douteux pendant la reconnaissance, en les soulignant d'une ligne ondulée.

- ◆ **Vert – Mots absents du dictionnaire** : La reconnaissance est sûre, mais ces mots ne se trouvent dans aucun dictionnaire actif (standard, utilisateur ou spécialisé).
- ◆ **Bleu – Mots comportant un caractère douteux** : Ces mots contiennent des caractères non reconnus ; ils peuvent également figurer dans un dictionnaire, mais contiennent des caractères pour lesquels la reconnaissance n'est pas aussi sûre.
- ◆ **Rouge – Mots douteux** : Ces mots, vraisemblablement absents du dictionnaire, contiennent au moins un caractère douteux. Le problème peut cependant être imputable à une autre raison.

Vous pouvez activer ou désactiver les marques des mots absents du dictionnaire dans l'onglet Vérification de la boîte de dialogue Options. Les marques peuvent être affichées ou masquées selon les options sélectionnées dans l'onglet Éditeur de texte de cette même boîte. 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.

OmniPage SE peut afficher les pages selon trois niveaux de formatage. Pour passer d'un niveau à l'autre, sélectionnez l'un des trois boutons présents dans l'angle inférieur gauche de l'éditeur de texte ou utilisez le menu Affichage. Les graphiques et les tableaux sont compatibles avec tous les modes disponibles. Les modes de l'éditeur disponibles présentent les caractéristiques suivantes :



Aucun 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

Le mode True Page® permet de 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 blocs et des cadres. L'ordre de lecture peut être indiqué par des flèches (voir à partir de la page 77).

Le niveau de formatage pour l'exportation est choisi séparément au moment de l'exportation.

Vérification des 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 au moment de votre choix. Procédez comme suit :



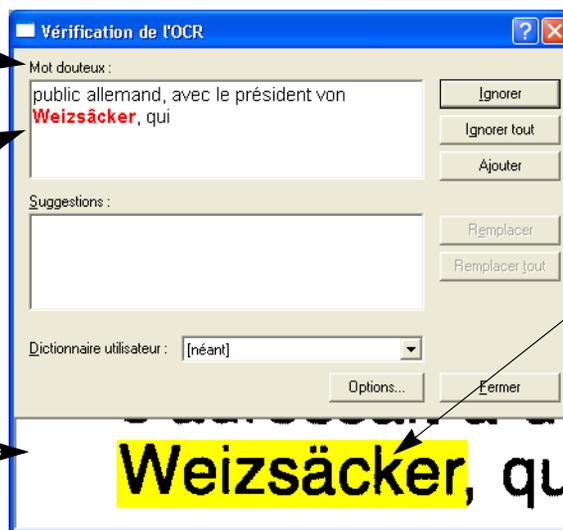
1. Cliquez sur l'outil Vérifier l'OCR dans la barre d'outils Standard ou choisissez Vérifier l'OCR dans le menu Outils.
2. La vérification démarre au début du document, mais omet le texte déjà vérifié. Si un mot douteux est détecté, la boîte de dialogue Vérification de l'OCR affiche ce mot dans son contexte,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

Raison pour laquelle le mot est marqué.

Fenêtre d'édition :

Le mot marqué est affiché en couleur (rouge, bleu ou vert).

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



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 fenêtre d'édition ou sélectionnez une suggestion du dictionnaire. Cliquez sur Remplacer ou Remplacer tout pour valider votre correction et passer au mot douteux suivant. Cliquez sur Ajouter pour ajouter le mot corrigé au dictionnaire utilisateur actuel et 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.



La vignette de toute page entièrement vérifiée porte l'icône de vérification  dans le gestionnaire de documents.



Si vous avez masqué les marques dans l'éditeur de texte au lancement de la vérification ou activé l'option *Rechercher prochain mot douteux*, les marques apparaissent et restent affichées après la vérification.



Si l'option *Marquer les mots absents des dictionnaires* est désactivée dans l'onglet Vérification de la boîte de dialogue Options, la vérification signale uniquement les mots marqués en rouge ou en bleu, mais non les mots absents des dictionnaires. Cette fonction est utile pour les pages qui contiennent un fort pourcentage de mots absents des dictionnaires, par exemple, des catalogues de produits avec codes et des bibliographies avec noms propres.



L'option *Revérifier la page actuelle* du menu Outils permet de lancer une nouvelle vérification orthographique sur une page déjà vérifiée. Vous pouvez ainsi contrôler les mots qui ont été tapés ou collés dans l'éditeur de texte après la vérification. Cette fonction s'applique même si l'option *Marquer les mots absents des dictionnaires* est désactivée dans l'onglet Vérification.

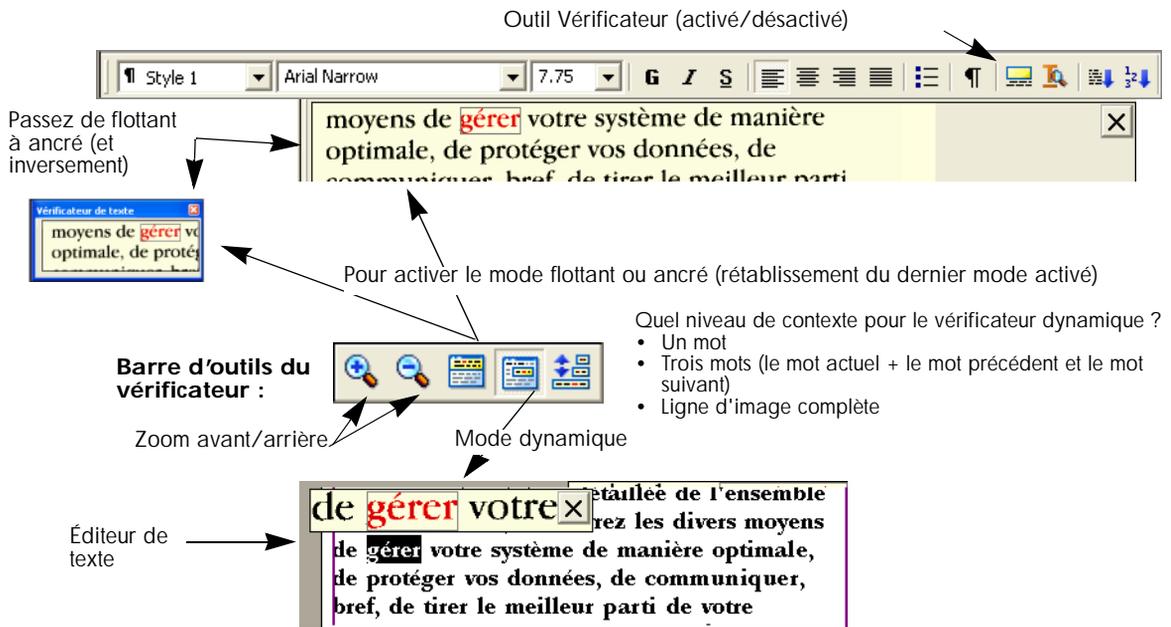
Vérification du texte reconnu

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 :

Action requise	Séquence utilisée
Activer le vérificateur	F9 ou outil Vérificateur
Désactiver le vérificateur	Échap, F9 ou outil Vérificateur
Activer/désactiver temporairement le vérificateur	F8 : appuyer et maintenir la touche enfoncée
Afficher le vérificateur jusqu'au prochain appui sur une touche	Double-cliquer sur le mot
Zoom avant	Alt + Num + ou cliquer dans le vérificateur
Zoom arrière	Alt + Num - ou cliquer dans le vérificateur
Définir le mode du vérificateur : dynamique ou ancré/flottant	Alt + Num /
Contexte dynamique (faire défiler 3 valeurs)	Alt + Num *



L'outil Vérificateur se trouve dans la barre d'outils de formatage. Vous pouvez également l'activer/le désactiver à partir du menu Outils. Placez le curseur sur la fenêtre du vérificateur pour afficher sa barre d'outils. Utilisez cet outil pour effectuer les opérations suivantes :





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.



Le module Texte-parole d'OmniPage Pro 12 vous permet de vérifier le texte reconnu à l'aide d'une lecture à voix haute. La lecture peut porter sur les lettres, les mots, les lignes ou les phrases ou s'effectuer sur l'ensemble de la page. Reportez-vous à la section « Lecture du texte à voix haute », page 80. 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 télécharger 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 téléchargement, sélectionnez [néant].

Éditer ou supprimer 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 en cliquant sur Édition dans la boîte de dialogue Fichiers de dictionnaire utilisateur. Le bouton Supprimer permet de supprimer le dictionnaire utilisateur sélectionné.



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 standard. Chaque mot doit être placé sur une ligne séparée et il ne doit être ni précédé ni suivi d'aucun signe de ponctuation.



Des dictionnaires spécialisés sont disponibles dans OmniPage Pro 12 pour certaines langues et certaines professions (juridiques et médicales, pour le moment). Ces dictionnaires ne sont pas proposés dans OmniPage SE.



Le programme identifie la langue utilisée pour les textes reconnus et l'affiche dans la barre d'état. Ces sélections de langue sont exportées avec le document. Pour modifier la langue du texte sélectionné, utilisez la commande Sélectionner langue du menu Outils. Cette option n'a aucune incidence sur les langues de reconnaissance spécifiées.

Apprentissage



L'apprentissage, le module IntelliTrain et les fichiers d'apprentissage ne sont pas supportés par OmniPage SE. Ces fonctionnalités sont uniquement disponibles dans OmniPage Pro 12. Toutes les données d'apprentissage pouvant accompagner un fichier OPD seront ignorées lors de son ouverture dans OmniPage SE.

L'apprentissage correspond à la procédure permettant de modifier les solutions d'OCR attribuées aux caractères de l'image. 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 » :

Suggérer Suggérer Suggérer Suggérer

Les deux premiers exemples n'exigent pas l'apprentissage, car les formes de la lettre « g » utilisées sont normales et le programme peut les traiter sans problème. Le troisième exemple pourrait bénéficier de l'apprentissage, 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 de l'apprentissage, 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.

Vous pouvez utiliser l'apprentissage pour améliorer la reconnaissance des symboles spéciaux, tels que @, ® et ©, ou pour accroître la fiabilité de la reconnaissance des caractères accentués supportés. L'apprentissage n'a pas pour but d'enseigner au programme la lecture des langues ou alphabets non supportés.

OmniPage Pro 12 propose deux types d'apprentissage : manuel et automatique (via IntelliTrain). Les données provenant des deux types d'apprentissage sont combinées et vous pouvez les enregistrer dans un fichier d'apprentissage.

Lorsque vous fermez une page pour laquelle des données d'apprentissage ont été générées, le programme vous demande de spécifier la manière de les appliquer aux autres pages existantes du document.

Apprentissage manuel

Pour lancer un apprentissage manuel, placez le point d'insertion devant le caractère à ajouter ; vous pouvez aussi sélectionner un groupe de caractères (un mot maximum) et choisir Nouveau caractère dans le menu Outils ou le menu contextuel. Une vue agrandie des caractères à ajouter s'affiche, assortie de la solution d'OCR actuelle. Remplacez-la par la solution voulue et cliquez sur OK. Le programme accepte cet apprentissage et examine le reste de la page. S'il trouve des mots susceptibles d'être modifiés, il les affiche dans la boîte de dialogue Vérification de l'apprentissage. Vous devrez recommencer l'apprentissage des mots incorrects avant de valider la liste.

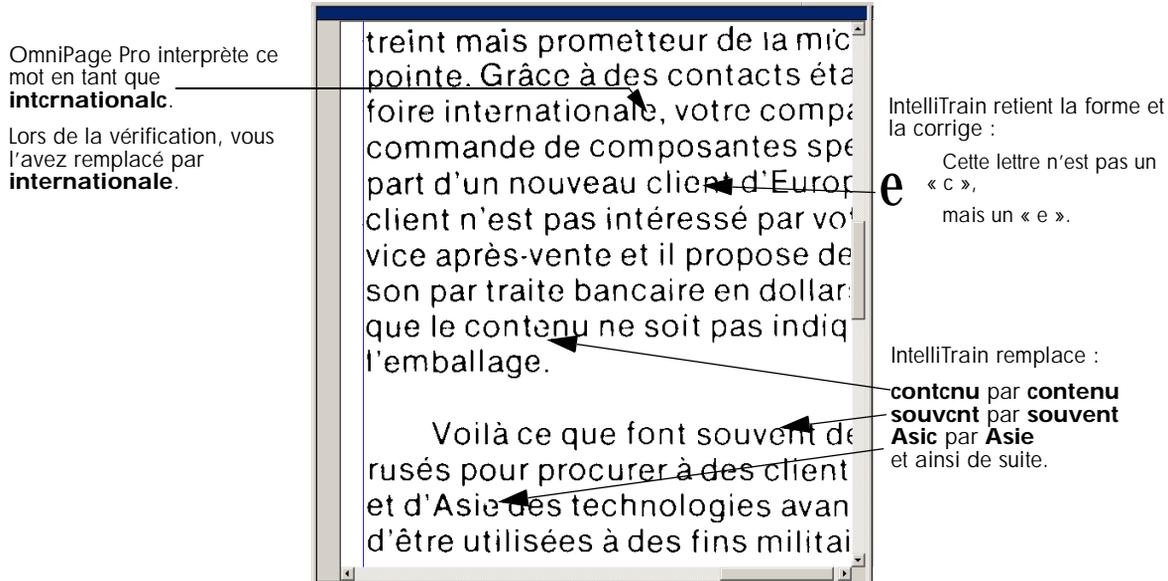
Pour plus de détails sur les boîtes de dialogue Nouveau caractère et Vérification de l'apprentissage, reportez-vous à l'aide contextuelle ou à la rubrique *Apprentissage manuel* de l'aide en ligne et à ses rubriques connexes.

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 peut être activé et désactivé via l'onglet OCR de la boîte de dialogue Options.

L'illustration ci-dessous explique le fonctionnement d'IntelliTrain. Cet 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 et les ajoute, le cas échéant, à l'apprentissage manuel que vous avez effectué. Vous pouvez les enregistrer dans un fichier d'apprentissage pour réutilisation avec des documents similaires.

Fichiers d'apprentissage

Si vous souhaitez que le programme vous demande d'enregistrer les données d'apprentissage non sauvegardées à la fermeture d'un document, sélectionnez l'option correspondante dans l'onglet Vérification de la boîte de dialogue Options. Les données d'apprentissage non enregistrées sont stockées dans un document OmniPage. Si vous n'enregistrez pas le document dans un fichier OPD, tout apprentissage non enregistré est éliminé lorsque vous fermez le document.

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 décharger un fichier d'apprentissage.



Cliquez sur ce bouton pour éditer le fichier d'apprentissage sélectionné dans la boîte de dialogue Modification du fichier d'apprentissage.

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

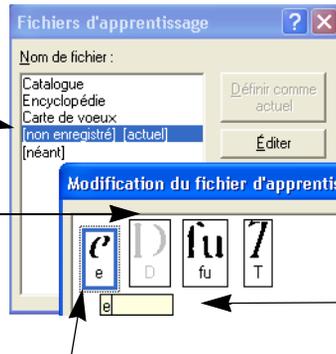
Des données d'apprentissage non enregistrées peuvent être éditées dans la boîte de dialogue Modification du fichier d'apprentissage ; un astérisque s'affiche dans la barre de titre à la place du nom de fichier. Les données ne sont pas sauvegardées à la fermeture de la boîte de dialogue Modification du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

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 Modification du fichier d'apprentissage 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 proposée. Pour passer au cadre précédant ou suivant, utilisez les touches fléchées.

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 proposée.

Édition du texte et des images

L'éditeur de texte d'OmniPage SE est doté de capacités d'affichage fidèles (« WYSIWYG »), et ses fonctions d'édition sont très proches des possibilités offertes par les applications de traitement de texte les plus courantes.

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 Aucun 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 utilisé pour l'exportation, ni transmis aux autres modes ; ceux-ci gardent leurs anciens paramètres.

Ouvrez la boîte de dialogue Mappage des polices (via l'onglet OCR de la boîte de dialogue Options) avant d'effectuer la reconnaissance 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 options de tabulation évoluées peuvent être spécifiées dans 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'attribuer 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 pouvez exporter également ses styles de paragraphe. Cette option est disponible uniquement si l'application cible supporte les styles de paragraphe.

Graphiques

Pour éditer des zones 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. Modifiez le graphique, puis fermez l'éditeur d'image pour réintégrer le graphique dans l'éditeur de texte. 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 certains types de fichiers spécifiques, vous pouvez exporter les tableaux en tant que grilles ou colonnes séparées par des tabulations ou par des espaces.

Hyperliens

Les adresses de courrier électronique et celles des pages Web peuvent être détectées et placées en tant que liens dans le texte reconnu. Pour éditer un lien existant ou pour en créer un nouveau, choisissez Hyperlien dans le menu Format. Ce nouveau lien peut renvoyer à une page Web ou à un fichier. Pour supprimer un lien, utilisez le menu contextuel.

Édition dans True Page

Les éléments de page sont placés dans des blocs de texte, de tableau et d'image, correspondant généralement aux zones Texte, Tableau et Graphique dans l'image. Cliquez dans un élément pour afficher la bordure du bloc ; la couleur est identique dans les zones correspondantes. La rubrique *True Page* de l'aide en ligne propose des détails sur les opérations décrites ci-après.

Les **cadres** ont des bordures grises et contiennent au moins un bloc. Ils sont placés lorsqu'une bordure visible est détectée dans une image. Vous pouvez définir les bordures et les dégradés d'un cadre et d'un tableau au moyen du menu contextuel ou de la commande Tableau du menu Format. Pour spécifier le dégradé d'un bloc de texte, utilisez le menu contextuel. Pour afficher un menu contextuel, cliquez avec le bouton droit de la souris dans un élément autre qu'un mot marqué.

Les **zones multicolennes** présentent des bordures roses et contiennent au moins un bloc. Détectées automatiquement, elles indiquent le texte qui sera traité comme un enchaînement de colonnes au moment de l'exportation. Les menus contextuels permettent de dégroupier les cadres et les zones multicolennes afin de modifier leurs éléments. Vous pouvez aussi grouper les éléments en cadres ou en zones multicolennes.



L'**ordre de lecture** peut être affiché et modifié. Pour visualiser les flèches indiquant cet ordre, cliquez sur l'outil Afficher ordre de lecture dans la barre d'outils de formatage. Cliquez de nouveau pour supprimer les flèches. Pour remplacer la barre d'outils de formatage par un groupe de boutons de

réagencement, cliquez sur l'outil *Changer l'ordre de lecture*. Leur utilisation est décrite dans l'aide contextuelle et dans la rubrique *Ordre de lecture* de l'aide en ligne. L'ordre est modifié dans les fenêtres *Aucun formatage* et *Conserver le formatage polices/paragraphes*. Ce changement modifie le mode de déplacement du curseur dans la page, lorsque le document est exporté au format *True Page*.

Édition en direct

Cette fonction vous permet de modifier une page reconnue par un nouveau découpage de zones, sans devoir traiter une nouvelle fois la page entière. Lorsque l'édition en direct est activée, les changements de zone (suppression, définition, redimensionnement, changement de type) s'appliquent immédiatement à la page reconnue. De même, lorsque vous modifiez des éléments dans la fenêtre *True Page* de l'éditeur de texte, les changements effectués sont répercutés sur les zones de la page. Vous pouvez aussi utiliser le découpage en zones « en direct » sur les pages non reconnues.

Deux outils liés de la barre d'outils *Image* contrôlent le découpage en zones en direct. L'un de ces outils est toujours actif lorsque aucune reconnaissance n'est en cours.



Cliquez sur ce bouton pour activer l'édition en direct. La couleur rouge indique qu'aucune modification de découpage en zones n'est stockée.



Cliquez sur ce bouton pour désactiver l'édition en direct. Les modifications apportées au découpage en zones sont stockées et l'outil affiche la couleur verte. Pour appliquer ces modifications, effectuez l'une des opérations suivantes :



Cliquez sur l'outil d'édition en direct doté de la couleur verte. Les modifications apportées au découpage en zones sont répercutées dans l'éditeur de texte.



Cliquez sur le bouton *Exécuter l'OCR* pour reconnaître la page entière, avec les changements de zone.

Pour plus de détails sur le traitement des changements dans le cadre du découpage de zones en direct, ainsi que sur leur effet dans les fenêtres de l'éditeur de texte, reportez-vous à la rubrique *Traitement en direct* dans l'aide en ligne.

Lecture du texte à voix haute



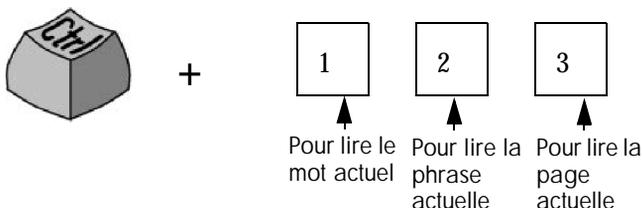
Le module Texte-parole n'est pas disponible dans OmniPage SE. Ce module n'est disponible que dans OmniPage Pro 12.

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	Ctrl + PgPrec, PgSuiv ou boutons de navigation
Caractères saisis	Chaque caractère saisi est prononcé individuellement, y compris les signes de ponctuation.

La fonction de synthèse vocale est activée ou désactivée via la commande Mode synthèse vocale du menu Outils ou via la touche F5. Le sous-menu Réglages de synthèse vocale permet de sélectionner le type de voix (masculine ou féminine dans une langue donnée) et de configurer le débit et le volume pour la lecture.

Les trois touches principales utilisées pour la lecture à voix haute sont regroupées sur le bloc numérique.



Vous disposez également des séquences clavier suivantes :

Action requise	Séquence utilisée
Interrompre/Reprendre	Ctrl + Cinq du bloc numérique (5)
Augmenter le débit	Ctrl + Signe plus du bloc numérique (+)
Réduire le débit	Ctrl + Signe moins du bloc numérique (-)
Rétablir le débit	Ctrl + Astérisque du bloc numérique (*)

Le programme de synthèse vocale devrait être proposé dans les langues 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 de synthèse vocale, la version anglaise est installée.

Si votre ordinateur est muni de systèmes de synthèse vocale compatibles avec la norme SAPI pour d'autres langues, ceux-ci sont détectés et mis à votre disposition. Leurs voix sont proposées dans la boîte de dialogue Réglages de synthèse vocale. Dès qu'une voix a été attribuée 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.

Chapitre 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 (une seule page, les pages sélectionnées ou le document entier) vers l'application cible de votre choix à l'aide d'une des méthodes suivantes : enregistrement du document dans un fichier, copie du document dans le Presse-papiers ou ajout du document à un e-mail en tant que pièce jointe. Vous pouvez également enregistrer le fichier en tant que document OmniPage.

Ce chapitre traite les sujets suivants :

- ◆ Enregistrement des images originales
- ◆ Enregistrement des résultats de l'OCR
 - Enregistrement d'un document lors d'une session de travail
 - Sélection d'un niveau de formatage
 - Sélection des options d'enregistrement évoluées
 - Enregistrement d'un document au format PDF
- ◆ Copie des pages dans le Presse-papiers
- ◆ Ajout de pages à un e-mail comme pièce jointe

Le document étant conservé dans OmniPage SE après l'exportation, vous pouvez enregistrer, copier ou envoyer ses pages 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. À l'issue de la reconnaissance (ou de la vérification, selon vos réglages) de la dernière page disponible, la boîte de dialogue d'exportation s'affiche.

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.

Enregistrement des images originales

Les images originales d'un document peuvent être enregistrées sur disque sous différents types de fichiers. Voir « Types de fichier pour l'ouverture et l'enregistrement des images », page 100.

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. Indiquez le nom du fichier.
2. Sélectionnez l'enregistrement de l'image de zone actuelle uniquement, de l'image actuelle, des images de pages sélectionnées ou de toutes les images du document. Dans ces deux derniers cas, vous pouvez placer l'ensemble des images dans un seul fichier image multipage, à condition de spécifier le type de fichier TIFF, MAX 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.



Pour afficher la taille et la résolution originale d'une image, placez le curseur sur sa vignette dans la fenêtre d'image.



Vous pouvez enregistrer vos documents dans cinq versions du format PDF dans OmniPage Pro. Deux de ces versions permettent d'enregistrer les images originales, tandis que les autres enregistrent les résultats de l'OCR. Pour plus d'informations, reportez-vous aux sections suivantes. L'enregistrement au format PDF n'est pas pris en charge par OmniPage SE.

Enregistrement des résultats de l'OCR

Les images reconnues d'un document peuvent être enregistrées sur disque sous différents types de fichiers. Voir « Types de fichier pour l'enregistrement des résultats de l'OCR », page 102.

1. Choisissez Enregistrer sous dans le menu Fichier, ou cliquez sur le bouton Exporter les résultats dans la boîte à outils OmniPage 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.

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

Cliquez sur ce bouton pour afficher et modifier les options de sortie pour le type de fichier actuel.

Sélections possibles :
Toutes les pages
Page actuelle
Pages sélectionnées
Sélectionnez les pages au moyen des vignettes ou dans le gestionnaire de documents.

Sélections possibles :
Regrouper les pages en un seul fichier
Créer un fichier par page
Créer un fichier après chaque page blanche
Créer un fichier par image

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. Sélectionnez ensuite un niveau de formatage pour le document. Voir « Sélection d'un niveau de formatage », page 87.
4. Indiquez le nom du fichier. Cliquez sur le bouton Évolué pour spécifier un intervalle de pages, une option de séparation de fichiers ou d'autres options d'enregistrement et sélectionnez les options voulues. Voir « Sélection des options d'enregistrement évoluées », page 88.

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 à ce type de fichier dans Windows ou dans les options d'enregistrement évoluées pour le filtre du type de fichier sélectionné.



Les graphiques, les grilles de tableaux et les autres propriétés sont enregistrés dans le document uniquement si le type de fichier sélectionné les supporte, et s'ils sont configurés pour conservation dans les paramètres évolués (Options de filtre) du type de fichier actuel.



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 *Créer un fichier après chaque page blanche*, vous pouvez placer des pages blanches dans le document. Voir « Importer des fichiers image », page 50.



Si vous activez l'option *Créer un fichier par image*, aucun nom de fichier n'est requis. Le nom du fichier de sortie est obtenu à partir du fichier d'entrée qui l'a généré ; seule l'extension change.

Enregistrement d'un document lors d'une session de travail



Cliquez sur l'outil 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 type de fichier du dernier enregistrement, tel qu'indiqué dans la barre de titre. Ces données s'affichent uniquement si le document entier a été enregistré.

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 dans un autre format, 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. Voir « Documents OmniPage », page 31. Les références aux données d'apprentissage ne s'appliquent pas à OmniPage SE.

La boîte de dialogue Enregistrer sous affiche les types de fichier disponibles dans la liste déroulante Enregistrer sous. 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.

Sélection d'un niveau de formatage

Le niveau de formatage pour l'exportation est défini au moment de l'opération, dans la boîte de dialogue Enregistrer sous, Copier dans le Presse-papiers ou Envoyer comme e-mail. Trois de ces niveaux correspondent aux fenêtres de formatage du même nom dans l'éditeur de texte. Cependant, le niveau à appliquer pour l'enregistrement ne dépend pas de la fenêtre de formatage affichée dans l'éditeur de texte. Lorsque vous exportez vers un fichier ou un e-mail, commencez par spécifier le type de fichier : ce choix détermine les niveaux de formatage disponibles. Vous trouverez un récapitulatif dans un tableau du chapitre 6. Voir « Types de fichier pour l'enregistrement des résultats de l'OCR », page 102.

Les niveaux de formatage sont les suivants :



Aucun formatage

Pour afficher le texte en une seule colonne avec alignement à gauche, dans un seul style et corps de police. Les éléments graphiques et les tableaux ne sont pas supportés pour l'exportation vers des fichiers de type Texte ou Unicode. Vous pouvez exporter le texte brut vers pratiquement tout type de fichier et application cible ; dans ce cas, vous pouvez conserver les éléments graphiques, les tableaux et les puces.



Conserver le formatage polices/paragraphes

Pour afficher le texte en une seule colonne et conserver les styles de police et de paragraphe, avec des éléments graphiques et des tableaux. Cette option est disponible pour la plupart des types de fichier.



Pages en continu

Pour conserver la mise en page du document original (y compris la disposition en colonnes). Si possible, les colonnes et les retraits, et non les cadres ou les blocs de texte, sont utilisés pour assurer l'enchaînement du texte. La fonction de

formatage des pages en continu n'est pas proposée dans OmniPage SE. Elle n'est disponible que dans OmniPage Pro.



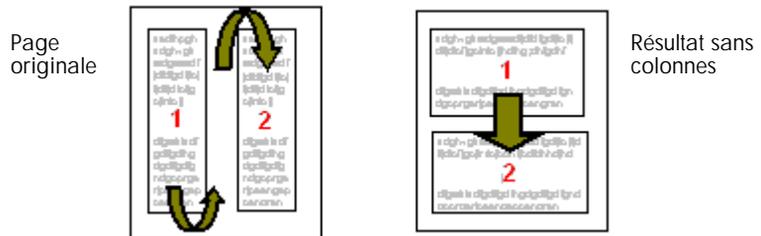
True Page

Pour conserver la mise en page du document original (y compris la disposition en colonnes). Les cadres et les blocs de texte, d'image et de tableau sont utilisés. Cette option concerne uniquement les applications cibles supportant ce type d'opération.

Feuille de calcul

Cette option exporte les résultats de l'OCR dans un tableau, adapté pour les applications de feuille de calcul.

La conversion du texte en une colonne unique pour les exportations en mode Aucun formatage et Conserver le formatage polices/paragraphes s'effectue de gauche à droite et de haut en bas :



Avant l'exportation, assurez-vous que l'ordre sans colonnes des éléments est correct dans la fenêtre Aucun formatage et Conserver le formatage polices/paragraphes. Si tel n'est pas le cas, passez à True Page et cliquez sur l'outil Afficher ordre de lecture pour indiquer l'ordre voulu par des flèches. Pour spécifier un ordre différent, utilisez l'outil Changer l'ordre de lecture. Les zones multicolonnées indiquent les colonnes liées. Si ces liens ne sont pas adaptés, dégrouperez la zone et modifiez l'ordre des éléments qu'elle inclut.

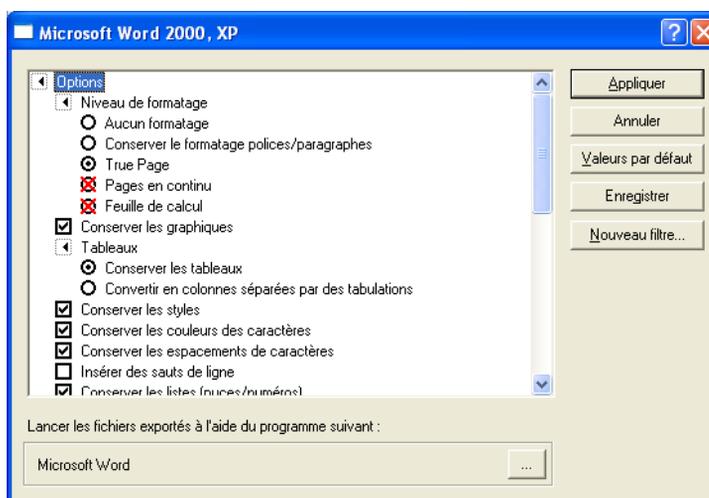
Sélection des options d'enregistrement évoluées

Pour disposer d'un contrôle précis sur l'exportation, cliquez sur le bouton Options de filtre dans la section Évoluée de la boîte de dialogue Enregistrer sous. Une boîte de dialogue portant le nom du type de fichier actuel s'affiche. Elle propose une série d'options adaptées à ce type de fichier. En premier lieu, confirmez ou modifiez le niveau de formatage ; ces données déterminent les autres options affichées. Sélectionnez ensuite les options voulues. Vous trouverez la description de ces opérations dans l'aide en ligne.

Pour appliquer les paramètres modifiés à l'enregistrement actuel uniquement, cliquez sur Appliquer.

Pour rétablir la valeur par défaut de tous les paramètres du type de fichier actuel, cliquez sur Valeurs par défaut.

Pour appliquer les paramètres modifiés à l'enregistrement en cours et les stocker en tant que paramètres à utiliser ultérieurement, si ce type de fichier est encore une fois sélectionné pour enregistrement, cliquez sur Enregistrer.



Le programme actuellement associé au type de fichier choisi pour la fonction Enregistrer et lancer s'affiche au bas de la boîte de dialogue. Pour spécifier un programme différent, cliquez sur le bouton portant les points de suspension.

Pour créer votre propre filtre personnalisé, définissez vos paramètres, cliquez sur Nouveau filtre, entrez le nom voulu, puis cliquez sur OK. Vous pouvez aussi commencer par nommer le filtre, puis modifier les paramètres et cliquer sur Enregistrer. Les filtres personnalisés sont utiles pour les tâches répétitives, par exemple, la publication d'un hebdomadaire. Vous pouvez ainsi exporter toutes les pages reconnues, avec un formatage adapté à leur utilisation prévue. Pour un type de fichier donné, vous pouvez aussi créer un groupe de filtres personnalisés définissant les options d'enregistrement appropriées pour chaque niveau de formatage, par exemple : RTF Aucun formatage, RTF Conserver le formatage polices/paragraphes et RTF True Page.

Il est possible de modifier les options de filtre sans effectuer aucun enregistrement dans un fichier. Ouvrez la boîte de dialogue Filtres d'exportation à partir du menu Outils. Sélectionnez le filtre voulu et cliquez sur le bouton Options. Le bouton Appliquer n'est pas disponible dans ce cas.

Enregistrement d'un document au format PDF



Cette section ne s'applique pas à OmniPage SE.

Dans OmniPage Pro 12, lors de l'enregistrement au format PDF (Portable Document Format), vous avez le choix entre cinq options :

PDF (format 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.

PDF édité

Utilisez cette option si vous avez effectué un grand nombre de modifications d'édition dans les résultats de l'OCR. Vous disposez de trois options pour le niveau de formatage, dont True Page. Le fichier PDF peut être affiché, exploré et édité.

PDF avec image sur texte

Le fichier PDF peut être affiché, mais il ne peut être modifié dans un éditeur PDF. Les images originales sont exportées, mais les images sont toutes assorties d'un fichier texte lié pour permettre d'effectuer une recherche dans le texte. Tout mot trouvé est contrasté dans l'image.

PDF avec substitution d'image

Ce format est similaire au format PDF standard, mais l'image originale se substitue aux mots contenant des caractères de rejet et douteux, ce qui assure un affichage identique au document original. Le fichier PDF peut être affiché, exploré et édité.

PDF avec image uniquement

Les images originales sont exportées. Le fichier PDF peut être affiché, mais le texte ne peut être ni modifié ni exploré dans un éditeur PDF.

Copie des pages dans le Presse-papiers

Vous pouvez copier les résultats de l'OCR dans le Presse-papiers, à partir de la page actuelle, des pages sélectionnées ou de toutes les pages du document. 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 6.0/95. Sinon, seules les données textuelles ou Unicode sont copiées. Les graphiques sont conservés dans les applications gérant l'insertion d'images.

▼ Pour copier des pages dans le Presse-papiers :

- Dans le cas d'un traitement automatique, sélectionnez la commande Copier dans le Presse-papiers dans la liste déroulante Exporter les résultats de la boîte à outils OmniPage ou dans la fenêtre de l'Assistant OCR. La boîte de dialogue Copier dans le Presse-papiers s'affiche dès que la dernière page disponible a été reconnue ou vérifiée.
- Dans le cas d'un traitement manuel, sélectionnez la commande Copier dans le Presse-papiers dans la liste déroulante Exporter les résultats, puis cliquez sur le bouton correspondant. La boîte de dialogue Copier dans le Presse-papiers s'affiche immédiatement.
- Spécifiez l'intervalle de pages et le niveau de formatage voulu, puis cliquez sur OK pour lancer la copie.



Vous pouvez utiliser la fonction Glisser-déplacer pour effectuer un copier-coller pour la page actuelle. Appuyez sur les touches Ctrl+Tab, et tout en les maintenant enfoncées, cliquez sur la page actuelle dans la fenêtre d'image, puis faites glisser le curseur vers une application cible avec un document ouvert. Le contenu de la page est collé à la position du curseur. La reconnaissance est exécutée, si nécessaire.

Ajout de pages à un e-mail comme pièce jointe

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

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



1. Cette boîte de dialogue permet de spécifier le type de fichier, un intervalle de page, un niveau de formatage et les options d'envoi : une pièce jointe pour toutes les pages, une pièce jointe par page, une pièce jointe après chaque page vierge ou une pièce jointe pour chaque fichier d'entrée. 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.



Le programme peut détecter des adresses électroniques à mesure qu'il reconnaît les pages et qu'il les transmet à l'éditeur de texte. Si vous cliquez sur une adresse, votre application de messagerie s'ouvre ; elle affiche un nouveau message vide indiquant uniquement cette adresse.

Chapitre 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 section du site consacrée aux scanners (Scanner Guide) fournit régulièrement des informations détaillées et mises à jour 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
 - Test de OmniPage SE
 - Augmentation de la quantité de mémoire disponible
 - Augmentation de l'espace disque
 - Reconnaissance de texte incorrecte
 - Reconnaissance de fax incorrecte
 - Problèmes de système ou de performances pendant l'OCR
- ◆ Support ODMA
- ◆ Fonctions évoluées du module Planifier OCR
- ◆ Types de fichier supportés
 - Types de fichier pour l'ouverture et l'enregistrement des images
 - Types de fichier pour l'enregistrement des résultats de l'OCR
- ◆ Désinstallation du 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. Voir « Environnement requis », page 12.
- ◆ Vérifiez que le scanner est sous tension et que ses câbles sont correctement connectés.
- ◆ Pensez à visiter le site Web de ScanSoft (www.scansoft.com), qui propose des notes techniques concernant les problèmes les plus courants afférents à l'utilisation d'OmniPage. Ces pages Web peuvent également inclure des conseils sur l'installation et le dépannage du logiciel.
- ◆ 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 le site Web de ScanSoft (accessible via le menu Aide), notamment la page consacrée aux scanners (Scanner Guide).
- ◆ Exécutez l'utilitaire ScanDisk (Windows 98 et Me) ou Check Disk (Windows NT, 2000 et XP) pour vérifier l'intégrité du disque dur. Pour plus de détails, reportez-vous à l'aide en ligne de Windows.
- ◆ Défragmentez le disque dur. Pour plus de détails, reportez-vous à l'aide en ligne de Windows.
- ◆ Désinstallez et réinstallez OmniPage SE (voir la dernière section de ce chapitre, « Désinstallation du logiciel », page 103).

Test de OmniPage SE

Le redémarrage de Windows 98, Me, 2000 et XP 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.



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 98, 2000, Me et XP) :

1. Redémarrez le système : 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 l'un des exemples de fichiers image.
 - 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 fin de ce chapitre), 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. Relancez 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, par exemple l'un des exemples de fichiers fournis.



Vous pouvez également exécuter OmniPage SE depuis une ligne de commande en mode sans échec. Choisissez Démarrer ▶ Exécuter, recherchez le fichier OmniPage.exe et ajoutez l'option de ligne de commande /safe. Cette opération lance le programme, mais sans prendre en compte les réglages enregistrés et sans essayer de récupérer des documents fermés incorrectement.

Augmentation de la quantité de mémoire disponible

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 superflues.
- ◆ Défragmentez le disque dur (pour plus de détails, consultez 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. Voir « Environnement requis », page 12.

Augmentation de l'espace disque

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 ou Check Disk.
- ◆ 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.
- ◆ Videz le cache de votre navigateur Web et diminuez sa taille.

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, consultez la section « Gestion des zones », page 60.
- ◆ 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. Voir « Types et propriétés de zone », page 58.
- ◆ 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 sur la vignette de page voulue pour afficher une mini-fenêtre. 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. Les références aux données d'apprentissage ne s'appliquent pas à OmniPage SE.
- ◆ Exécutez un apprentissage manuel ou éditez les données d'apprentissage existantes pour supprimer toute valeur incorrecte.
- ◆ 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 ou des blocs de texte. Si l'un de ces blocs de texte est trop petit, une partie du texte peut être masquée. Pour afficher le texte caché, placez le curseur dans ce bloc de texte 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.



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. Voir « Importer des fichiers image », page 50.

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 d'insuffisance de mémoire vive et d'espace disque. Voir « Test de OmniPage SE », page 95.
- ◆ 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. 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. Voir « Gestion des zones », page 60.

- ◆ Relancez Windows 98, 2000, XP ou 2000 en mode sans échec (ou Windows NT en mode VGA), et testez OmniPage SE en effectuant la reconnaissance des exemples de fichiers image.

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

Support ODMA



Ceci ne s'applique pas à OmniPage SE. Si votre réseau local inclut un système DMS capable de supporter les clients ODMA, OmniPage Pro peut vous permettre de l'utiliser. Une fenêtre ODMA s'affiche alors dans la boîte de dialogue Options, vous permettant de spécifier les types de fichier autorisés et d'autres paramètres. Une interface ODMA remplace les boîtes de dialogue Charger un fichier image et Ouvrir un document OmniPage (*.OPD). Vous pouvez ainsi charger les fichiers image et OPD un par un, à partir du système de fichiers de réseau ou de votre ordinateur local. La boîte de dialogue Enregistrer sous comporte le bouton Enregistrer dans DMS, qui permet d'enregistrer dans ce système les documents reconnus. Pour plus de détails, reportez-vous à l'aide en ligne.

Fonctions évoluées du module Planifier OCR



Ceci ne s'applique pas à OmniPage SE. Le module Planifier OCR vous permet de désigner les fichiers d'entrée pour un travail. Dans certaines éditions d'OmniPage Pro, vous pouvez spécifier le traitement de tous les fichiers d'un type donné dans un ou plusieurs dossiers. Ces versions proposent également la fonction de la surveillance des dossiers. Le premier écran de l'Assistant de nouveau travail comprend deux boutons : Fichiers et Dossiers. Il inclut également la case à cocher *Surveiller l'arrivée de fichiers dans les dossiers*. Sélectionnez cette option et spécifiez au moins un dossier. Tous les fichiers du ou des types spécifiés entrant dans le ou les dossiers sélectionnés seront traités à l'arrivée. Dans la cinquième fenêtre de l'Assistant, vous pouvez définir une heure de départ et d'arrêt pour la procédure de surveillance des travaux. Vous pouvez également décider de ne pas spécifier une heure d'arrêt lors de la configuration d'un travail. Dans ce cas, utilisez le bouton Modifier du module Planifier OCR pour entrer l'heure d'arrêt par la suite.

Vous pouvez mettre les travaux en pause et les reprendre. Le bouton Affichage vous permet de visualiser, fichier par fichier, le journal des traitements terminés d'un travail sélectionné. Pendant l'exécution d'un travail, vous pouvez consulter le journal des travaux afin de suivre (fichier par fichier) la progression de l'opération et les problèmes signalés.

La quatrième fenêtre de l'Assistant de nouveau travail vous permet de spécifier un type de fichier et de choisir les options de séparation des fichiers. Si vous choisissez l'option *Un fichier de sortie pour chaque fichier d'entrée*, vous spécifiez uniquement un dossier ; les fichiers de sortie conservent leurs noms d'entrée, mais avec d'autres extensions. Si vous choisissez une option exigeant plusieurs fichiers de sortie, vous indiquez un nom de fichier et le programme y ajoute automatiquement des suffixes numériques afin de générer des noms uniques. Si l'entrée spécifiée provient d'un groupe de fichiers, vous pouvez définir un dossier de sortie différent pour chaque dossier d'entrée.

Types de fichier supportés

OmniPage gère de nombreux types de fichier pour les images et le texte.

Types de fichier 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
MAX	max	Oui	Ouvert. et enreg.	Tous
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	Ouvert. et enreg.	Noir et blanc
TIFF compressé G4	tif	Oui	Ouvert. et enreg.	Noir et blanc
TIFF compressé LZW	tif	ND	ND	ND
TIFF FX	xif	Oui	Ouverture	Tous
TIFF PackBits	tif	Oui	Ouvert. et enreg.	Tous
TIFF non compressé	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. Cela signifie que ces images ne sont pas appropriées pour un traitement d'OCR ultérieur.

Placez le curseur sur une vignette de page pour afficher une mini-fenêtre indiquant la taille et la résolution de l'image originale.



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 suggère une conversion en noir et blanc.



Seul OmniPage Pro 12 gère l'enregistrement au format PDF et propose à l'utilisateur cinq versions de ce format, dont deux (*Avec image uniquement* et *Avec image sur texte*) permettent l'exportation des images originales. Vous pouvez sélectionner le format requis dans la boîte de dialogue Enregistrer sous. Voir « Enregistrement d'un document au format PDF », page 90. Veuillez noter que OmniPage SE ne gère pas les fonctionnalités PDF. OmniPage SE ne gère pas non plus les fichiers GIF et TIFF LZW.

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

Le tableau qui suit indique les niveaux de formatage disponibles pour chaque type de fichier.

Type de fichier	Extension	Aucun formatage	Conserver le formatage polices/paragraphes	Pages en continu (1)	True Page	Feuille de calcul	Éléments graphiques	Tableaux	
eBook (voir remarque 1)	opf	●	●				●●	●	
Excel 97, 2000	xls	●	●			●	●●	●	
Excel 3.0 à 7.0	xls	●	●			●		●	
FrameMaker 5.5.3	mif	●	●				●●	●●	
Freelance Graphics	txt	●	●					○●	
Harvard Graphics	txt	●	●					○●	
HTML 4.0 (voir remarques 1 et 2)	htm	●	●		●		●●	●	
HTML 3.2 (voir remarque 2)	htm	●	●			●	●●	●	
Microsoft PowerPoint 97	rtf	●	●				●●	●●	
Microsoft Publisher 98	rtf	●	●				●●	●●	
Microsoft Word 6.0, 97, 2000, XP	doc	●	●	●	●		●●	●●	
PageMaker 6.5.2	doc	●	●				●●	●●	
Quattro Pro pour Windows 4.0, 8	xls	●	●			●		●	
PDF (format standard) (1)	pdf				●		●●	●	
PDF édité (voir remarque 1)	pdf	●	●		●		●●	●●	
PDF avec image sur texte (1)	pdf				(●)		●	●	
PDF avec substitution d'image (1)	pdf				●		●	●	
PDF avec image uniquement (1)	pdf				(●)		●	●	
RTF Word 2000, 97, 95/6.0 (3)	rtf	●	●	●	●		●●	●●	
Ventura Publisher	doc	●	●				●	●●	
WordPad	rtf	●	●				●●	●●	
WordPerfect 8, 9, 10	wpd	●	●	●	●		●●	●●	
WordPerfect 6.0, 6.1	wpd	●	●		●		●●	●●	
WordPerfect 5.1, 5.2	wp5	●	●		●		●●	●●	
XML (voir remarque 1)	xml				●		●	(●)	
Texte et Texte avec sauts de ligne (voir remarque 4)	txt	●	●					○●	
Texte avec données séparées par des virgules (voir remarque 4)	csv	●	●					●	
Texte formaté (voir remarque 4)	txt				●			(●)	
Document OmniPage (5)	opd	Conserve les propriétés d'affichage						●	●

Éléments graphiques

- Le type de fichier supporte les graphiques
- Le type de fichier supporte les graphiques, avec la possibilité, au moment de l'exportation, de conserver ou d'omettre les graphiques.

Tableaux

- Le type de fichier supporte les tableaux avec grilles, sans options de tableau au moment de l'exportation
 - Le type de fichier supporte les tableaux et vous pouvez choisir d'utiliser des grilles ou des colonnes séparées par des tabulations
 - Le type de fichier ne supporte pas les grilles des tableaux ; vous pouvez choisir une conversion en colonnes séparées par des tabulations ou des espaces
- 1 Ces formats et Pages en continu ne sont pas pris en charge par OmniPage SE.
 - 2 Lors de l'enregistrement au format HTML, tous les graphiques sont enregistrés en tant que fichiers image séparés, sous le format JPEG.
 - 3 Les résultats de l'OCR sont placés dans le Presse-papiers au format RTF 95/6.0, puis sont collés dans ce format si possible, et sinon, au format Unicode ou ASCII.
 - 4 Les formats de texte sont tous disponibles en tant que Texte ou Unicode. Le format Unicode est compatible avec le plus grand nombre de caractères accentués.
 - 5 Vous pouvez rouvrir les Documents OmniPage créés à l'aide de l'édition standard ou spéciale d'OmniPage Pro 12. Vous pouvez également ouvrir dans OmniPage SE des fichiers OPD créés à l'aide d'OmniPage Pro 10 ou de l'édition standard ou spéciale d'OmniPage Pro 11. Ces fichiers sont importés en tant que documents sans nom. Pour conserver un fichier OPD sous son format d'origine et l'enregistrer dans un nouveau fichier OPD, choisissez un nom différent pour ne pas remplacer le fichier d'origine.

Désinstallation du logiciel

Certains problèmes peuvent être résolus en désinstallant puis en réinstallant OmniPage SE. Veillez également à désinstaller OmniPage SE avant d'installer OmniPage Pro 12 ou toute autre version d'évaluation du programme. Le programme de désinstallation d'OmniPage Pro 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, Windows 2000 ou Windows XP, 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, 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 jusqu'à la fin de la procédure.



I N D E X

- A**
acquérir texte, réglages, 47
ADF, 33, 50, 53
afficher/masquer les marques, 68
aide
 contextuelle, 9, 25, 33
 en ligne, 9
ajouter
 à une zone, 61
 apprentissage aux fichiers
 d'apprentissage, 76
 des mots à un dictionnaire utilisateur,
 70
 pages à un document, 41
alphanumérique, zone, 58
ancrée/flottante
 barre d'outils, 25
ancrée/flottante, barre d'outils, 70
apprentissage, 73
 charger des fichiers d'apprentissage, 76
 créer des données d'apprentissage, 76
 décharger des fichiers d'apprentissage,
 76
 demander l'enregistrement des données,
 31
 données d'apprentissage non
 enregistrées, 31
 éditer des fichiers d'apprentissage, 76
 enregistrer des fichiers d'apprentissage,
 76
 fichiers d'apprentissage, 76
 manuel, 74
Assistant
 configurer un scanner, 14
 de nouveau travail, 49, 99
 Planifier OCR, 49, 99
 pour le traitement, 45
attributs, caractère, 77
Aucun formatage, mode, 68, 87
augmentation
 de l'espace disque, 96
 de la mémoire disponible, 96
automatique
 mise en page, 54
 traitement, 27, 40
avantages
 de l'apprentissage, 73
 du format OPD, 32
- B**
barre d'outils
 ancrée, 25
 flottante, 25
- Formatage, 24, 25
 Image, 24, 25
barre des menus, 25
blocs, 26, 78
 texte reconnu, 97
boîte à outils
 OmniPage, 24, 27, 40
bureau, 24
- C**
cadres, 26, 78, 88, 97
caractères
 douteux, 68
 non imprimables, 68
 non reconnus, 68
changer de mode de traitement, 27, 43
charger
 dictionnaire utilisateur, 72
 fichiers d'apprentissage, 76
 fichiers OPD, 31
 gabarits de zones, 55, 64
Charger un fichier image, boîte de dialogue,
 50
chargeur de feuilles automatique (ADF), 33,
 50, 53
colonne unique
 avec tableaux, mise en page, 54
 pages, 54, 58
colonnes
 dans le gestionnaire de document, 30
 dans les tableaux, 63
 de remarques dans le gestionnaire de
 documents, 30
colonnes multiples, mise en page, 54
combiné, traitement, 43
Comparaison des fonctionnalités
 d'OmniPage SE et de celles d'OmniPage
 Pro, 8, 10, 19
comparer le texte reconnu à l'image originale,
 70
configurer
 Direct OCR, 47
 un scanner, 14
conseils pour la résolution des problèmes, 94
conservation de la mise en page, 69
Conserver le formatage polices/paragraphes,
 mode, 68, 87
conserver les styles de paragraphe, 85
consultation de l'aide en ligne, 9
contextuelle, aide, 33
contraste, 33, 52, 97
contrôle du traitement, 42
conversion d'images, 101
copier des pages dans le Presse-papiers, 90
- couleur
 images, 84
 marques, 70
 numériser, 52
créer des données d'apprentissage, 76
- D**
décharger un dictionnaire utilisateur, 72
découper en zones
 en direct, 79
 fond, 55
dégrouper les éléments, 78
demander l'enregistrement des données
 d'apprentissage, 31
démarrer le programme, 14
dépannage, 93, 94
déplacer séparations de tableau, 63
désinstaller OmniPage SE, 103
dictionnaire, 45, 70
 ajouter des mots, 70
 charger, 72
 décharger, 72
 démarrer, 72
 éditer, 72
 légal, 70
 médical, 70
 spécialisé, 70
 utilisateur, 70
différer le traitement, 31
Direct OCR, 46
 Options, boîte de dialogue, 33
 tracer des zones, 47
diviser des zones, 58
documents
 avec mise en page variable, 54
 copier dans le Presse-papiers, 45, 90
 dans OmniPage SE, 23
 emplacement des nouvelles pages, 34
 enregistrer, 32, 83
 enregistrer lors d'une session de travail,
 86
 exporter, 23, 40, 43, 83
 fermer, 31
 gérer, 28
 mise en page, 54
 mise en page hétérogène, 54
 non terminés, 31
 recto verso, 53, 54
 terminer, 41
documents OmniPage, 86
 avantages du format OPD, 32
 contenu, 86
 définition, 31
 enregistrer sous, 32, 86

données d'apprentissage, 31
dossiers surveillés, 99

E

éditer

- attributs de caractère, 77
- attributs de paragraphe, 77
- dans True Page, 78
- dictionnaire utilisateur, 72
- éléments graphiques, 78
- en direct, 79
- et découper en zones en direct, 79
- fichiers d'apprentissage, 76
- PDF, format d'exportation, 90
- tableaux, 62, 78
- texte reconnu, 77

éditeur de texte, 24, 26, 34, 68

éléments graphiques

- dans l'exportation, 85
- dans les fichiers HTML, 103

éditer, 78

exportation, 102

emplacement des nouvelles pages, 34

en instance, pages, 28

en ligne

- aide HTML, 9
- enregistrer, 17

enregistrés, pages, 28

enregistrer

- applications, Direct OCR, 47
- au format OPD, 32, 85
- dans un fichier, 46, 84
- dictionnaire utilisateur, 72
- documents, 83
- documents lors d'une session de travail, 86
- en tant que document OmniPage, 32, 86

- Enregistrer et lancer, option, 86
- fichiers d'apprentissage, 76
- gabariets de zones, 64
- images originales, 84, 100
- OmniPage SE, 17
- résultats de l'OCR, 85
- texte, 85

entendre le texte lu à voix haute, 81

entrée pour Planifier OCR, à partir d'un dossier, 99

environnement requis, 12

Envoyer comme e-mail, boîte de dialogue, 91

Exécuter l'OCR, bouton, 40, 43

exemples

- de l'apprentissage, 73
- fichiers image, 95

exporter

- dans le Presse-papiers, 90
- dans un fichier, 85, 102
- éléments graphiques, 85, 103
- Pages en continu, mode, 88
- plusieurs fois, 83, 86
- True Page, 88
- types de fichiers et niveaux de formatage, 102

- vers des applications cibles, 23, 42, 84
- vers PDF, 90, 102
- vers une messagerie, 91

Exporter le texte reconnu, bouton, 41

Exporter les résultats, bouton, 43, 85

F

fax, reconnaissance incorrecte, 98

fenêtre d'image, 24, 26

fermer un document, 31

feuille de calcul, mise en page, 55

fichier PDF

- exportation, 101
- importation, 50, 101

fichiers

- conservés après désinstallation, 103
- destination d'exportation, 84
- en tant que source de l'image, 50
- options de séparation, 85, 92
- types, 85
- types pour l'exportation, 87, 102
- types supportés, 100

fichiers d'apprentissage, télécharger, 76

fichiers image

- exemples, 36, 95
- importer, 22, 50
- ordre de lecture, 50
- ouvrir, 50, 100
- types, 100

fichiers image multipages, 50, 84

fichiers OPD

- avantages, 32
- définition, 31
- enregistrer, 32

fonctions, nouvelles, 17

fond, 26

- à valeur Ignorer, 55
- à valeur Traiter, 55

G

gabariets de zones, 55, 64, 97

- télécharger, 64
- modifier, 64
- remplacer, 64

générer les séparations d'un tableau, 63

gérer

- des documents, 28
- les zones, 60

gestionnaire de documents, 24, 28, 29

personnaliser les colonnes, 30

Gestionnaire des tâches, 95

Graphique, type de zone, 59

gras, texte, 77

grouper les éléments, 78

guide de démarrage rapide, 36

H

hyperliens, 78

I

Ignorer, valeur des zones, 59

image originale, enregistrer, 84, 101

images

- conversion, 101
- couleur, 84
- éditer, 78
- enregistrer, 84, 100
- fonds, 55
- importer, 23, 42
- niveaux de gris, 84
- noir et blanc, 84
- qualité, 52
- résolution, 29, 84, 97, 101
- substitution dans les fichiers PDF, 90
- taille, 29

importées, pages, 28

importer

- à partir du scanner, 51
- depuis des fichiers PDF, 50, 100
- depuis un fichier image, 50
- des images, 23, 42

Importer la page, bouton, 40, 42

Importer la page, liste déroulante, 42

imprimer

- documents, 31
- pages reconnues, 31
- indicateurs d'en-tête et de bas de page, 68
- informations techniques, 93
- insérer des séparations dans un tableau, 63
- insertion de nouvelles pages dans le document, 34

installer

- OmniPage SE, 13
- scanners, 14

IntelliTrain, 34, 49, 97

- apprentissage automatique, 74
- apprentissage manuel, 74

interface, langue, 34

interrompre le traitement automatique, 41

irrégulières, zones, 60

italique, texte, 77

J

joindre

- des pages à un e-mail, 91
- des zones, 61

L

lancer

- l'application cible, 86
- un dictionnaire utilisateur, 72

langues

- d'interface utilisateur, 34
- de reconnaissance, 33, 45, 97

lien vers une page Web, 78

lire

- le texte à voix haute, 80
- ordre des fichiers image, 50

liste déroulante

- Exporter les résultats, 43
- Importer la page, 42

livres, numérisation, 34

luminosité, 33, 52, 97

M

- manuel
 - traçage des zones, 42, 55
 - traitement, 27
- marques, 68, 70
- menus contextuels, 59
- messagerie, 41, 91
- messages électroniques, pièces jointes, 91
- Microsoft Outlook, 91
- minimum
 - environnement requis, 12
- mise en page
 - automatique, 54
 - description, 40, 45, 54
- modes
 - Aucun formatage, 68
 - Conservé le formatage polices/
paragraphe, 68
 - de l'éditeur de texte, 26, 68
 - sans échec, 95
 - True Page, 69
 - VGA, 95
- modifiées, pages, 28
- modifier
 - ordre de lecture, 79
 - partie d'une page, 79
 - types de zones, 59
- mots
 - absents du dictionnaire, 68
 - douteux, 68
 - marqués dans l'éditeur de texte, 68
- multicolonne, zones, 26, 78
- multipages, fichiers image, 100

N

- niveaux de formatage, 49, 68, 102
 - types de fichiers, 102
- niveaux de gris
 - images, 84
 - numériser, 52
- noir et blanc
 - images, 84
 - numériser, 52
- Nouveau fichier après chaque page blanche,
option, 50
- nouvel apprentissage, 74
- nouvelles fonctions, 17
- numérique, zone, 58
- numérisation
 - contraste, 33
 - livres, 34
 - luminosité, 33
 - sur deux pages, 34
- numériser
 - Assistant, 14
 - couleur, 52
 - images, 52
 - importer, 52
 - luminosité, 52
 - niveau de gris, 52
 - noir et blanc, 52

O

- OCR
 - Assistant, 39, 45, 46
 - définition, 22
 - Direct OCR, 33, 46
 - étapes, présentation, 23
 - exécuter l'OCR, 23
 - Planifier OCR, 49
 - principales étapes, 23
 - problèmes de performances, 98
 - réglages, 33
 - réglages, Direct OCR, 46
 - traitement automatique, 27, 40
 - traitement manuel, 27, 42
 - travaux, Planifier OCR, 49
 - vérifier les résultats, 69, 70
- omettre des graphiques lors de l'exportation,
85
- OmniPage bureau, 24
- OmniPage Pro
 - nouvelles fonctions, 17
- OmniPage SE
 - comparaison de ses fonctionnalités avec
celles d'OmniPage Pro, 8, 10, 19
 - démarrer, 14
 - désinstaller, 103
 - documents dans, 23
 - enregistrer, 17
 - installer, 13
 - réinstaller, 103
 - tester, 95
 - versions antérieures, 13
- optimiser la luminosité, 33, 52
- options
 - de traitement, 34
 - de vérification, 34, 69
- Options, boîte de dialogue, 33
- ordre
 - de lecture, 79
 - des éléments de page, 79
- ouvrir des fichiers image, 50, 100

P

- page Web, liens, 78
- pages
 - colonne unique, 54, 58
 - colonne unique avec tableaux, 54
 - colonnes multiples, 54
 - copier dans le Presse-papiers, 45, 90
 - diviser en zones, 28
 - en instance, 28, 79
 - enregistrées, 28
 - état, 28
 - feuille de calcul, 54
 - fichiers image multipages, 50, 84, 101
 - importées, 28
 - Importer la page, bouton, 40, 42
 - insertion dans le document, 34
 - joindre à un e-mail, 91
 - mise en page, 54
 - modifiées, 28
 - navigation, 24, 81
 - Nouveau fichier après chaque page
blanche, option, 50

- parcourir, 28
- plusieurs colonnes, 54
- recommencer la reconnaissance, 43
- reconnues, 28
- relancer la reconnaissance, 41
- réordonner, 28
- sélectionner plusieurs, 28, 29
- supprimer, 28, 30
- vérifiées, 28
- Pages en continu, mode, 88
- PaperPort, 48
- paragraphe
 - conserver les styles de paragraphe, 86
 - éditer les attributs, 77
 - styles, 77, 85
- parcourir pages, 28
- PDF, explorer, 90
- performances, problèmes pendant l'OCR,
98
- personnalisée, mise en page, 34, 55
- personnaliser
 - barres d'outils, 25
 - colonnes dans le gestionnaire de
documents, 30
- pilotes TWIN pour scanners, 14
- Planifier OCR, 49
 - dossiers surveillés, 99
 - entrée à partir de dossiers, 99
- plusieurs pages, sélectionner, 29
- précision
 - améliorer, 51, 73, 74
 - incidence de la formation, 73
 - incidence du mode de numérisation, 51
 - influence de la luminosité, 52
- présentation
 - document, 28
 - étapes de l'OCR, 23
 - traitement, 38
- Presse-papiers, 41, 90
- principales étapes de l'OCR, 23
- problèmes
 - espace disque insuffisant, 12, 96
 - mémoire insuffisante, 12, 96
 - reconnaissance de fax, 98
 - système ou de performances pendant
l'OCR, 98
- propriétés de zone, 58

Q

- qualité de l'image, 52

R

- rangées d'un tableau, 63
- recherche
 - mots absents du dictionnaire, 69
 - mots douteux, 69
- reconnaissance
 - accélérer, 98
 - enregistrer les résultats, 85
 - excessivement lente, 98
 - exécuter, 42
 - langues, 33, 45, 97
 - optique de caractères, 22

- précision, 52, 73
- reconnaissance de fax incorrecte, 98
- reconnues, pages, 28
- rectangulaires, zones, 60
- redimensionner les zones, 60
- réglages, 34
 - Acquérir texte, commande, 47
 - dans l'Assistant OCR, 46
 - dans la boîte de dialogue Options, 33
 - Direct OCR, 47
 - types de zones, 62
- réinstaller OmniPage SE, 103
- relancer traitement automatique, 41
- réordonner les pages, 28
- résolution, 29, 84, 97, 101
- retour à la ligne automatique, 34

S

- scanners, 51, 98
 - configurer, 14
 - pilotes, 14
 - recto verso, 53
- sélectionner plusieurs pages, 29
- solutions aux problèmes de performances, 94
- souligné, texte, 77
- Soustraire de la zone, outil, 58
- spécifier la mise en page du document, 40, 54
- Standard
 - barre d'outils, 24, 25
- stocker les changements du découpage en zones, 79
- suggestions pendant la vérification, 70
- support
 - DMS, 99
 - ODMA, 99
- supprimer
 - dictionnaire utilisateur, 72
 - fichiers d'apprentissage, 76
 - gabarits de zones, 64
 - les séparations d'un tableau, 63
 - pages, 28, 30
- synthèse vocale
 - fonction, 13, 81

T

- tableau
 - colonnes, 63
 - dans les pages à colonne unique, 54

- déplacer des séparations, 63
- éditer, 78
- générer les séparations, 63
- insérer des séparations, 63
- modifier les séparations, 62
- rangées, 62
- supprimer les séparations, 52
- traiter un tableau dans l'éditeur de texte, 78
 - zones, 59, 62
- terminer le traitement d'un document, 41
- tester OmniPage SE, 95
- texte
 - enregistrer, 85
 - impression matricielle, 98
- texte reconnu, comparer à l'image originale, 70

- Texte, zone, 59

- TIFF, fichiers image, 100

- touches pour la lecture à voix haute du texte reconnu, 81

- traitement

- à partir d'autres applications, 46
- automatique, 27, 40
- automatique, incomplet, 41
- automatique, interrompre, 41
- automatique, relancer, 41
- avec l'Assistant OCR, 45
- changer de mode de traitement, 27, 43
- combiné, 27, 43
- étapes, présentation, 38
- manuel, 27, 42
- par étapes, 23, 42
- présentation, 36
- ultérieure des documents, 31

- traitement automatique

- incomplet, 41
- interrompre, 41
- relancer, 41

- Traiter, valeur des zones, 59

- travaux. Planifier OCR, 49

- True Page, 26

- édition, 78
- exporter, 88
- mode, 69

- types

- de fichier supportés, 100
- de zone, 58

U

- unité de mesure, 34
- URL, 78
- utiliser Direct OCR, 47

V

- vérificateur dynamique, 70
- vérification

- différée, 31
- options, 34, 69
- ultérieure, 31

- Vérification de l'OCR, boîte de dialogue, 69

- vérifiées, pages, 28

- vérifier les résultats de l'OCR, 69, 70
- vignettes, 24, 26, 28

Z

- zones, 26

- à valeur Ignorer, 59
- ajouter à, 61
- alphanumériques, 58
- automatiques, 26, 34, 40, 54, 59
- décharger un gabarit, 65
- définir les types, 63
- diviser, 59
- enregistrer un gabarit, 64
- gabarit, 55, 64, 97
- gérer, 60
- Graphique, 59
- irrégulières, 60
- joindre, 61
- manuelles, 55, 97, 99
- modifier le type, 59
- modifier un gabarit, 64
- numériques, 58
- propriétés, 58
- rectangulaires, 60
- redimensionner, 60
- remplacer un gabarit, 64
- soustraire de, 59
- supprimer un gabarit, 64
- sur la page, 28
- Tableau, 59, 62
- Texte, 59
- tracer dans Direct OCR, 47
- Traiter, 59
- types, 26, 58, 97
- zoom, valeurs, 70