

## **4 CONSEILS ET TECHNIQUES**

Les trois premiers chapitres de ce manuel offrent les informations nécessaires pour comprendre, installer et utiliser les fonctionnalités de base de TextBridge.

Le présent chapitre décrit différentes méthodes permettant d'obtenir les meilleurs résultats. Les sujets suivants sont traités :

- Optimisation de la reconnaissance de texte
- Amélioration du traitement des documents
- Enregistrement des images de pages
- Exécution de TextBridge à partir d'autres applications

### **OPTIMISATION DE LA RECONNAISSANCE DE TEXTE**

Le logiciel TextBridge réalise des reconnaissances optiques de caractères très précises sur une grande variété de documents. Cependant, vous pouvez encore améliorer la performance de TextBridge pour un document spécifique.

Cette section propose des suggestions pour optimiser la reconnaissance de caractères. Les sujets suivants sont traités :

- Utiliser et entretenir votre scanner correctement
- Ajuster la luminosité du scanner
- Ajuster les couleurs
- Utiliser le filtre Fax
- Traiter plusieurs documents individuellement
- Utiliser le vérificateur de mots
- Utiliser la commande Inverser de la fenêtre Visualiser

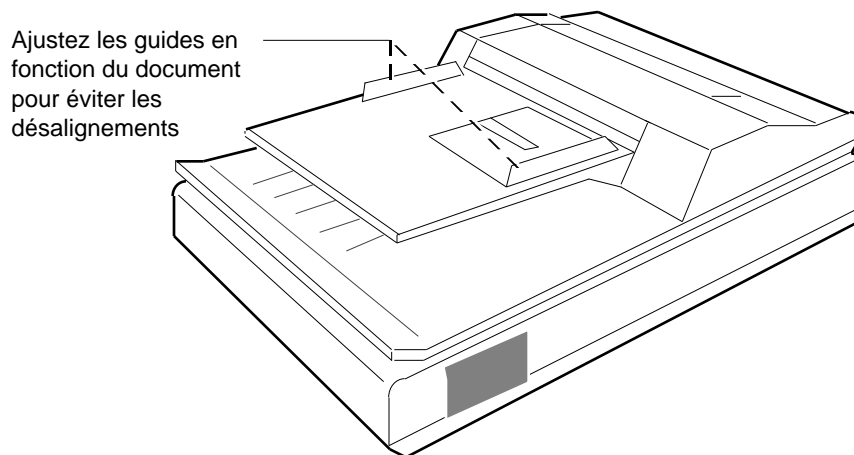
## Utiliser et entretenir votre scanner correctement

La façon dont vous utilisez et entretenez votre scanner peut déterminer l'échec ou la réussite de la numérisation. Suivez ces conseils :

- **Apprenez à connaître votre scanner.** Lisez attentivement toute la documentation fournie avec votre scanner.
- **Entretenez votre scanner.** Maintenez votre scanner propre. Nettoyez la vitre pour éliminer la poussière et les taches qui risqueraient d'être saisies lors de la numérisation.
- **Chargez le scanner correctement.** Assurez-vous que votre document n'est pas positionné de travers. Cela risquerait de compliquer la reconnaissance.

Lorsque vous utilisez le chargeur de document, veillez à positionner les guides correctement en fonction des pages à numériser (figure 4-1).

Si vous placez une page sur la vitre, assurez-vous qu'elle est positionnée contre le bord de celle-ci et correctement alignée. Si l'image est imprimée de travers sur la page, essayez de corriger ce défaut autant que possible avant de procéder à la numérisation.

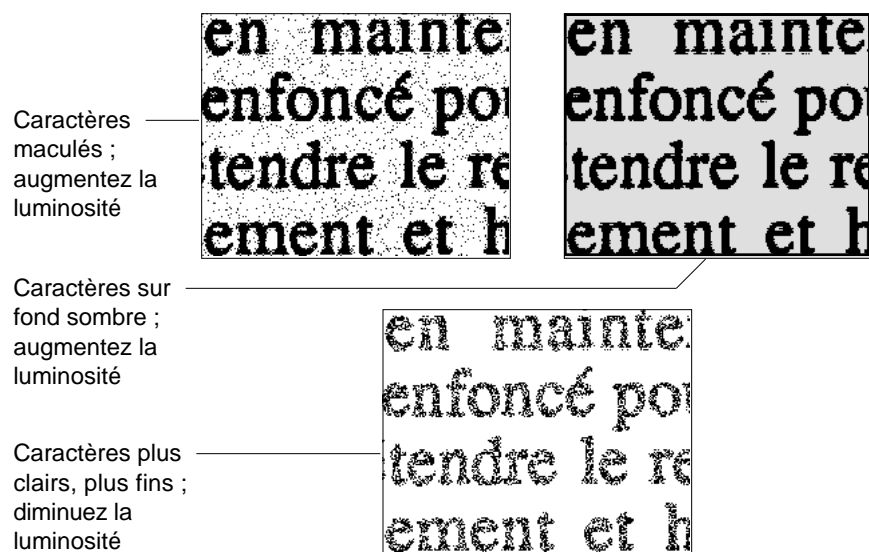


**Figure 4-1.** Positionnement des pages sur le scanner

## Ajuster la luminosité du scanner

Au cours de la numérisation, l'un des facteurs les plus importants pouvant affecter la reconnaissance de caractères est la **luminosité** du scanner. Comme l'illustre la figure 4-2, la qualité des documents originaux peut varier considérablement.

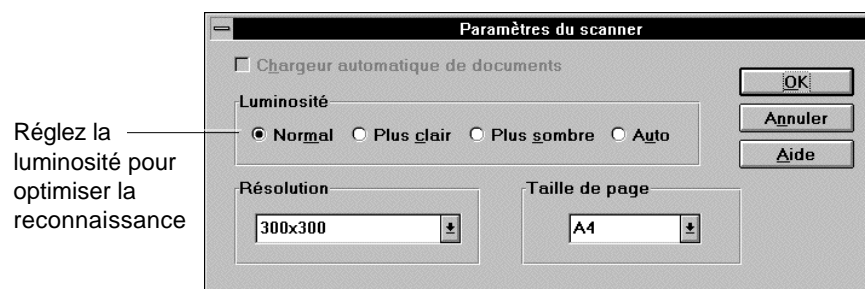
Un texte trop foncé, un fond trop clair ou des **parasites** (saleté, taches d'encre, traces de doigts, notes manuscrites et autres marques) sur la page peuvent affecter la précision de la reconnaissance.



**Figure 4-2.** Documents originaux et luminosité

Ajustez la luminosité en fonction de la qualité d'impression et du fond du document original, dans la boîte de dialogue Paramètres du scanner (figure 4–3). Vous pouvez accéder à celle-ci à partir de la boîte de dialogue Préférences.

Choisissez l'option Plus clair si les caractères imprimés sur votre page sont trop foncés, un peu flous ou se touchent, ou encore si les mots ne sont séparés que par des espaces très petits (comme dans certains magazines). La reconnaissance des documents qui contiennent des parasites, ou dont le fond est coloré ou quadrillé, peut être améliorée considérablement en augmentant la luminosité.

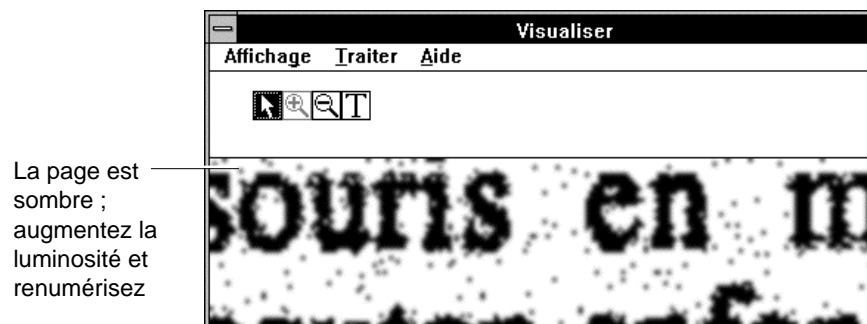


**Figure 4–3.** Paramètres de luminosité du scanner

Choisissez l'option Plus sombre si les caractères imprimés sur votre page sont pâles, découpés, ou très fins.

Si votre scanner supporte le paramètre Auto, sélectionnez-le pour obtenir la meilleure luminosité sur chaque page du document.

L'autre moyen de tester la luminosité consiste à examiner une page et à l'agrandir au maximum dans la fenêtre Visualiser. En procédant ainsi, vous visualisez l'image numérisée telle que le système la "voit" (figure 4–4). Si la luminosité de l'image ne vous semble pas adéquate, changez le paramètre puis renumérisez le document.



**Figure 4-4.** Visualisation d'une page agrandie

## Ajuster les couleurs

Il existe toujours une ou plusieurs couleurs que les scanners ne peuvent pas lire. (On parle de **couleurs invisibles**). Reportez-vous à la documentation de votre scanner pour déterminer la ou les couleurs invisibles.

- + Si la documentation de votre scanner ne mentionne aucune couleur invisible, examinez la couleur du faisceau lumineux de votre scanner. Cette couleur détermine la couleur invisible. Sur la plupart des scanners, ce faisceau est vert clair (dans ce cas, la couleur invisible est donc le vert clair).

Votre scanner a peut-être des difficultés à lire d'autres couleurs. Si le texte (ou l'image) à numériser est en couleur ou imprimé sur un fond coloré, essayez d'ajuster le paramètre de luminosité.

Si vous n'obtenez pas de meilleurs résultats, photocopiez la page et numérisez la copie en noir et blanc.

## Utiliser le filtre Fax

L'une des applications de TextBridge consiste à reconnaître le texte des images de fax. La résolution de ces dernières est généralement faible (100x200, 200x100 ou 200x200 ppp). Même les fax de “haute résolution” (200x200 ppp) sont souvent difficiles à lire (figure 4–5).

Les images des fax sont de faible résolution et souvent imprimées de travers

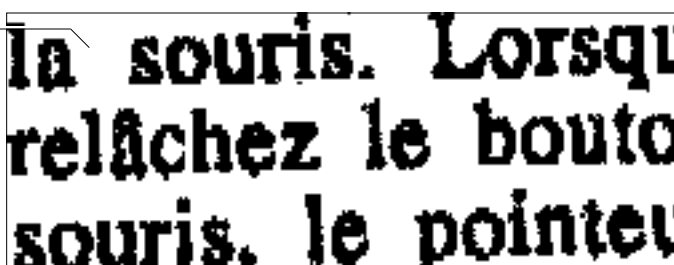


Figure 4–5. Image de fax

TextBridge offre un paramètre **Fax** (dans la boîte de dialogue Préférence) pour la reconnaissance des images de fax. Ce filtre permet d'effectuer un prétraitement pour l'améliorer la qualité de l'image avant sa reconnaissance.

Le filtre Fax traite les images de fax stockées dans des fichiers TIFF et les fax sur support papier numérisés à des résolutions supérieures (par exemple, 300 ppp).

---

**Remarque** N'utilisez **pas** le filtre Fax sur des documents autres que des fax, qu'ils soient numérisés ou en ligne, car la précision de reconnaissance risquerait d'en souffrir. De plus, si vous remarquez que la reconnaissance des images de fax synthétisées (par exemple, un document de traitement de texte "imprimé" sur un fax modem) est médiocre, désactivez le filtre et essayez de vérifier une partie du texte au cours de la reconnaissance.

---

## Traiter plusieurs documents individuellement

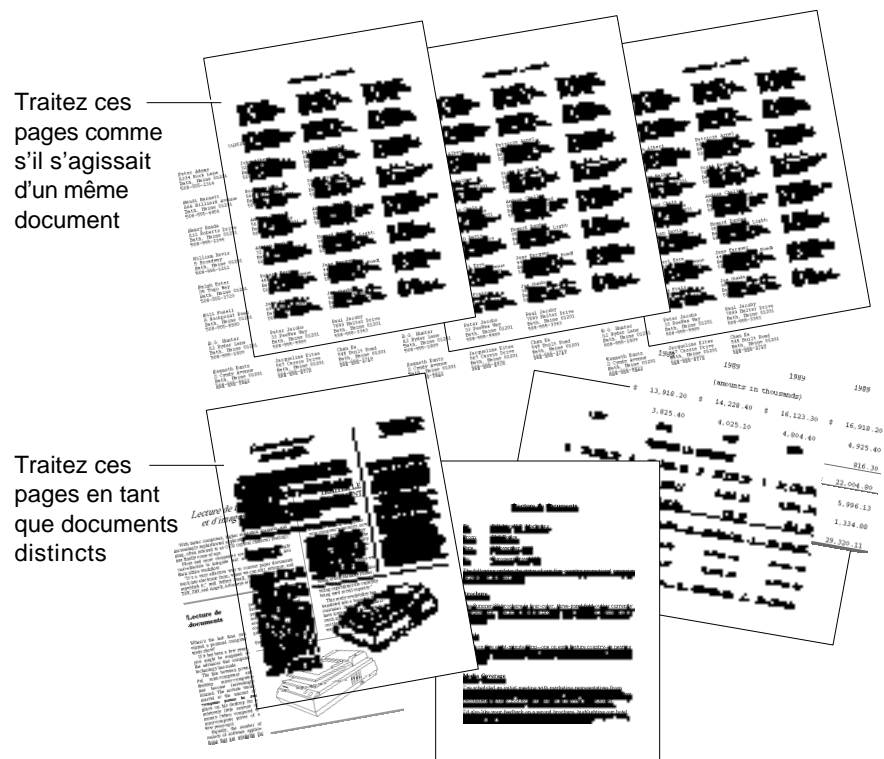
Le logiciel TextBridge utilise diverses techniques relevant de l'intelligence artificielle pour reconnaître les caractères.

Grâce à ces techniques, TextBridge effectue son propre apprentissage au cours de la reconnaissance.

Ainsi, TextBridge peut améliorer la précision et la vitesse de la reconnaissance au fur et à mesure qu'il numérise et reconnaît les pages d'un document.

Toutefois, vous pouvez compromettre cette capacité d'apprentissage en traitant les pages de différents documents dans le même fichier de sortie.

TextBridge s'attend à ce que toutes les pages consécutives d'un document utilisent les polices reconnues sur la première page (figure 4-6).



**Figure 4-6.** Traitement de plusieurs documents

Si la seconde page provient d'un document complètement différent (avec des polices et des tailles de caractères différentes), TextBridge ne peut tirer aucun parti de ce qu'il a "appris" en reconnaissant la première page.

TextBridge doit recommencer son processus d'apprentissage pour la seconde page (et les suivantes).

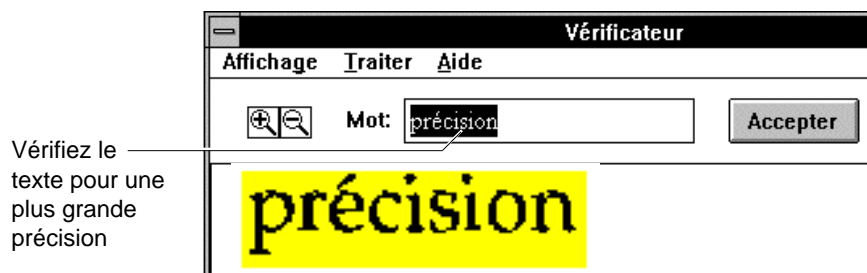
Si vous souhaitez numériser plusieurs documents et obtenir les meilleurs résultats, traitez chaque document séparément.



## Utiliser le vérificateur de mots

Si vous n'êtes pas satisfait des résultats de TextBridge sur un document particulier, utilisez le vérificateur de mots pour améliorer la précision de la reconnaissance.

En intervenant dans le processus de reconnaissance via la fenêtre Vérificateur (figure 4-7), vous initiez TextBridge aux caractères et mots du document, et améliorez ainsi la précision de la reconnaissance de manière significative.



**Figure 4-7.** Fenêtre Vérificateur

Chaque mot que TextBridge ne peut reconnaître ou dont il n'est pas sûr s'affiche dans la boîte d'édition Mot située en haut de la fenêtre Vérificateur. L'image du mot est également mise en surbrillance dans son contexte.

Avec le Vérificateur, vous pouvez vous déplacer dans le texte reconnu et accepter ou corriger les décisions de TextBridge. Vous aidez ainsi TextBridge à améliorer la reconnaissance au fur et à mesure du traitement du document.

Dans le cas d'un document comportant plusieurs pages, nous vous conseillons de vérifier une ou deux pages puis de mettre un terme à la vérification. TextBridge utilisera vos données pour prendre de meilleures décisions au cours de la reconnaissance du reste du document.

Pour obtenir une reconnaissance de la plus haute précision sur les petits documents (une ou deux pages), vérifiez ceux-ci en entier.

Pour des informations supplémentaires sur l'utilisation du vérificateur de mots et ses options, reportez-vous au chapitre 3.

## Utiliser la commande Inverser de Visualiser

TextBridge est capable de reconnaître des fichiers TIFF en ligne, provenant de fax modems ou d'autres sources.

Parfois, les données des images sont enregistrées de telle sorte que les éléments graphiques (**pixels**) du fichier résultant sont inversés : le fond blanc apparaît en noir et les caractères noirs sont affichés en blanc. Ce phénomène se produit souvent avec les fichiers Intel FAXability par exemple.

TextBridge **ne peut pas** reconnaître de tels fichiers. Pour une reconnaissance de fichiers TIFF en ligne, il est important que l'image comporte des caractères en noir sur un fond blanc.

Pour la reconnaissance des fichiers dont l'image est inversée, TextBridge dispose de la commande Inverser, qui figure dans le menu Affichage de la fenêtre Visualiser (figure 4–8).



**Figure 4-8.** Inversion d'un document

Si vous n'êtes pas sûr que le fichier en ligne est inversé, affichez-le dans la fenêtre Visualiser avant de commencer la reconnaissance. S'il apparaît en caractères blancs sur un fond noir, cliquez sur la commande Inverser du menu Affichage. TextBridge inversera l'image.

Vous pouvez ensuite lancer le processus de reconnaissance. Notez que l'inversion doit être corrigée manuellement pour chaque fichier TIFF enregistré ainsi.

Pour des détails sur l'utilisation de Visualiser, reportez-vous au chapitre 3.

## AMELIORATION DU TRAITEMENT DES DOCUMENTS

La première fois que vous utiliserez TextBridge, il vous semblera probablement plus simple de numériser un document sans modifier les préférences par défaut. Vous obtiendrez de très bons résultats avec ces préférences dans la plupart des cas.

Toutefois, si vous désirez tirer le meilleur parti des possibilités de TextBridge, suivez ces quelques conseils avant de commencer la reconnaissance :

- utilisez les outils de la fenêtre Visualiser ;
- utilisez l'option Ignorer photos/demi-tons ;
- utilisez l'orientation automatique autant que possible ;
- utilisez la segmentation automatique pour les documents comportant plusieurs colonnes.

Ces fonctionnalités imposent au système de traiter uniquement (et correctement) les parties essentielles de la page.

De plus, en utilisant ces fonctionnalités, vous économiserez un temps considérable en cas de traitement d'un long document ou de plusieurs documents.

### Encadrer des parties pour ne saisir que les données utiles

Certains documents contiennent des logos, des graphiques, des en-têtes, des pieds de page et d'autres éléments que vous n'avez pas besoin de saisir, et qui pourraient ralentir le processus de reconnaissance.

Avec les outils disponibles dans la fenêtre Visualiser, vous pouvez identifier la portion de la page à saisir (figure 4-9).

Reportez-vous au chapitre 3 pour des informations sur l'utilisation de ces outils.

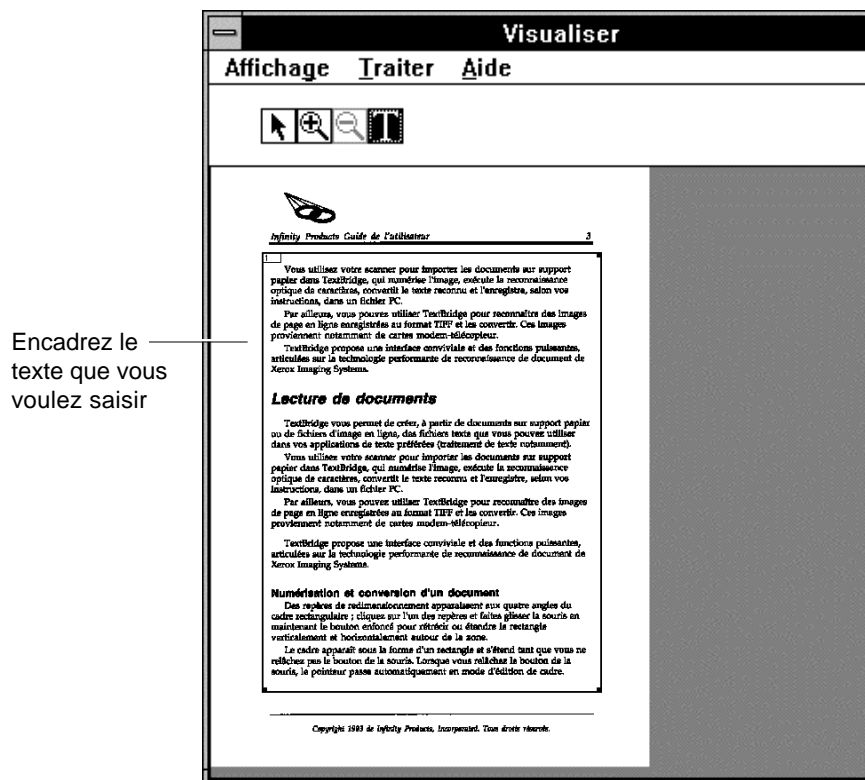


Figure 4-9. Cadre dans la fenêtre Visualiser

## Utiliser l'option Ignorer photos/demi-tons

Sur un document imprimé, une photographie en demi-tons est composée de points noirs de différentes tailles. TextBridge risque de perdre du temps à essayer de reconnaître chacun de ces points en tant que texte.

TextBridge finira par conclure qu'il essayait de reconnaître un demi-ton et il l'ignorerait. Pour accélérer la reconnaissance du texte d'un document qui contient aussi des demi-tons, activez l'option **Ignorer photos/demi-tons** avant que ne commence la reconnaissance.

Lorsque l'option Ignorer photos/demi-tons est activée, TextBridge numérise rapidement l'image de page et cache les demi-tons **avant** de commencer la reconnaissance des caractères (figure 4-10). Ainsi, cette opération est plus rapide et plus efficace.

TextBridge tente de reconnaître les demi-tons en tant que texte

Avec le filtre, les demi-tons sont retirés avant la reconnaissance

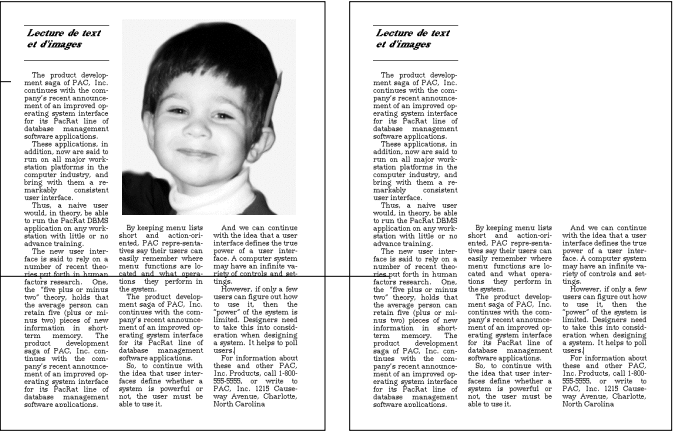


Figure 4-10. Filtre Ignorer photos/demi-tons

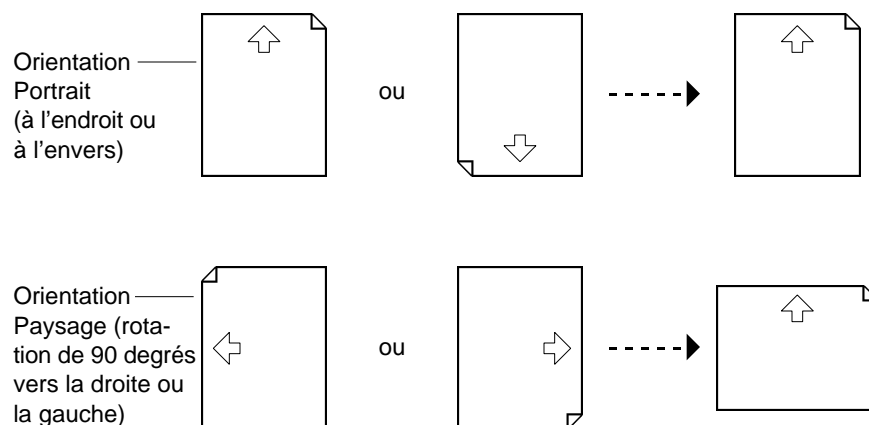
Pour utiliser l'option Ignorer photos/demi-tons, sélectionnez le bouton Préférences dans la boîte de dialogue principale. Cliquez dans la case à cocher Ignorer photos/demi-tons de la boîte de dialogue Préférences.

+ Bien que la procédure de filtrage soit relativement rapide, n'activez pas cette option si votre document ne contient pas de demi-tons.

## Utiliser l'orientation de page automatique

TextBridge dispose d'un outils qui détermine automatiquement l'orientation d'une page, effectue une rotation en mémoire si nécessaire, puis commence la reconnaissance des caractères (figure 4–11).

Choisissez **l'orientation Auto** dans la boîte de dialogue Préférences, accessible via la boîte de dialogue principale. Cette fonctionnalité est utile dans certaines circonstances, par exemple :



**Figure 4–11.** Fonctionnement de l'orientation automatique des pages

- lorsque vous traitez des documents dont les pages sont orientées différemment ;
- lorsque vous traitez un fichier TIFF et que vous ne connaissez pas l'orientation de l'image qu'il contient.

Dans le premier cas, vous traiterez un document composé principalement de pages portrait mais aussi de plusieurs pages paysage.

TextBridge numérise chaque page, déterminera si elle est en orientation portrait ou paysage (90 degrés ou 270 degrés), ou à l'envers, et la placera en orientation paysage (0 degré) avant de commencer la reconnaissance.

Dans le second cas, si l'image TIFF que vous vous apprêtez à reconnaître est de travers ou à l'envers ; TextBridge choisira l'orientation adéquate puis effectuera la reconnaissance.

- + La procédure d'orientation automatique prend place avant la reconnaissance. Par conséquent, pour parvenir à une reconnaissance rapide, utilisez cette fonctionnalité **uniquement** lorsque cela est nécessaire.

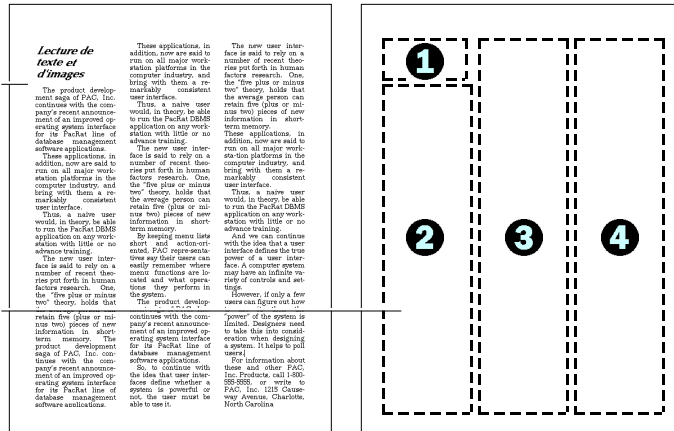
## Utiliser la segmentation de page automatique

TextBridge dispose d'un outil qui repère automatiquement les **régions** de texte sur la page, définit leur ordre, puis commence la reconnaissance. Cette fonctionnalité de **segmentation de page automatique** est particulièrement utile lors de la reconnaissance de pages comprenant plus d'une colonne et/ou dont la présentation est inhabituelle (figure 4-12).



La fonctionnalité de segmentation automatique analyse la présentation de la page

Puis identifie les régions et détermine leur ordre



**Figure 4-12.** Segmentation de page et identification des régions

## Important

**Vous devez activer** la fonctionnalité de segmentation automatique si vous traitez des pages comprenant plusieurs colonnes de texte. Sinon, TextBridge risque de ne pas générer les régions de texte reconnues dans le bon ordre. N'utilisez pas cette fonctionnalité sur les documents à une seule colonne.

La segmentation automatique est effectuée avant que la reconnaissance ne commence.

Notez que vous pouvez créer un cadre dans la fenêtre Visualiser et utiliser la segmentation automatique à l'intérieur de celui-ci uniquement. Ainsi, si vous encadrez deux colonnes sur une page qui en comporte trois, ces deux colonnes seront identifiées puis ordonnées pour la reconnaissance.

## ENREGISTRER DES IMAGES DE PAGES

L'option **Enregistrer images de pages** se trouve dans la boîte de dialogue principale. Elle est disponible lorsque vous définissez Entrée à partir de sur Scanner. Elle vous permet d'enregistrer une image binaire (en noir et blanc) de chaque page numérisée au cours d'une session de reconnaissance TextBridge.

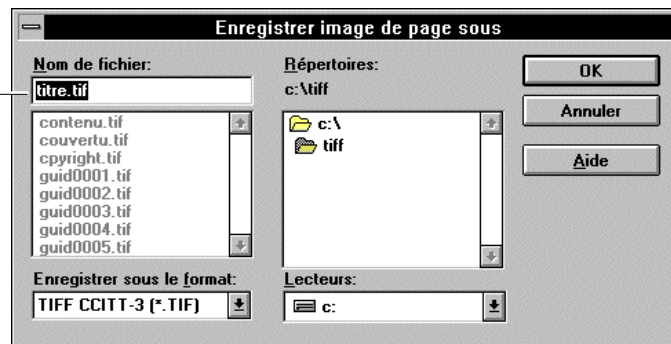
---

**Note** TextBridge enregistre les images de pages en tant que fichiers TIFF comprimés selon la norme CCITT Groupe 3. Le CCITT (Consultative Committee of International Telephone and Telegraph) est une organisation chargée d'établir des normes internationales.

---

Après que vous avez cliqué sur GO! et que la première page a été numérisée, TextBridge affiche la boîte de dialogue Enregistrer image de page sous (figure 4-13).

Tapez un nom de base pour les images de pages



**Figure 4-13.** Boîte de dialogue Enregistrer image de page sous

La boîte de dialogue Enregistrer image de page sous est très similaire à une boîte de dialogue Enregistrer sous de Windows. Elle vous permet en effet de spécifier un nom et un type de fichier, ainsi qu'un répertoire et un lecteur de destination.

Le nom de fichier est le **nom de base** à partir duquel les noms des fichiers d'images sont définis. Il s'agit également du **nom de document** qui s'affiche par défaut dans la boîte de dialogue Enregistrer sous de TextBridge lorsque la reconnaissance optique est terminée. Le nom de fichier par défaut est `titre` avec l'extension `.tif`.

Supposons par exemple que vous avez spécifié le nom "guide" dans la boîte de dialogue Enregistrer image de page sous. Les fichiers d'images seront stockés sous un nom respectant le format suivant :

`guidnnnn.TIF`

où *nnnn* indique le numéro de la page du document, précédé de zéros (par exemple, 0001, 0002, etc.).

Les fichiers des images de pages sont nommés de façon séquentielle dans le répertoire. S'il existe déjà un fichier du même nom (par exemple, "guid0001.tif"), TextBridge commencera par le numéro 0002.

De plus, à la fin de la session, le nom du document (par exemple, "guide") s'affiche automatiquement dans la case Nom de fichier de la boîte de dialogue Enregistrer sous habituelle.

Dans la boîte de dialogue Enregistrer image de page sous, le répertoire de travail initial est celui à partir duquel vous avez lancé TextBridge :

`C:\TXBRIDGE\BIN`

Cependant, vous pouvez spécifier un autre lecteur et un autre répertoire pour le stockage des images de pages. Vous définirez ainsi un nouveau **répertoire de travail** pour cette tâche qui, comme le nom du document, s'affichera dans la boîte de dialogue Enregistrer sous, une fois la reconnaissance optique terminée.

Le menu Enregistrer sous le format ne propose qu'un seul type de format, TIFF CCITT-3 Intel. Les images de pages sont enregistrées exactement telles qu'elles ont été numérisées, en format binaire (noir et blanc).

Notez que si vous cliquez sur Annuler dans la boîte de dialogue Enregistrer image de page sous, celle-ci se referme et TextBridge met fin à la session. La boîte de dialogue principale reste affichée pour la tâche suivante. (Par exemple, si vous ne souhaitez pas enregistrer les images des pages, vous pouvez désactiver l'option Enregistrer images de pages avant de reprendre votre travail.)

## EXECUTION DE TEXTBRIDGE A PARTIR D'AUTRES APPLICATIONS

TextBridge pour Windows est en fait un ensemble d'applications qui vous permet d'effectuer la reconnaissance optique de caractères à partir de pratiquement n'importe quelle autre application Windows.

En plus de l'utilitaire principal, qui tourne en tant que programme autonome et possède le jeu de fonctionnalités le plus complet, TextBridge vous offre :

- **Serveur d'applications TextBridge, ou TAS,** un programme qui fonctionne en tant qu'article de menu à partir de pratiquement toutes les applications de texte Windows (traitement de texte, PAO, tableur, base de données, etc.).

- **Imprimante OCR TextBridge (TextBridge OCR Scanner)**, une fonctionnalité qui permet d'envoyer une image de n'importe quel format à une version de TextBridge OCR qui fonctionne comme un gestionnaire d'impression conventionnel.

Cette section vous apporte des informations sur l'utilisation de TextBridge sous ces deux formes.

---

**Remarque** TextBridge supporte également une interface DDE. Les développeurs et intégrateurs de systèmes intéressés peuvent contacter le Support technique de Xerox Imaging Systems pour plus de détails.

---

## Utiliser le Serveur d'applications TextBridge

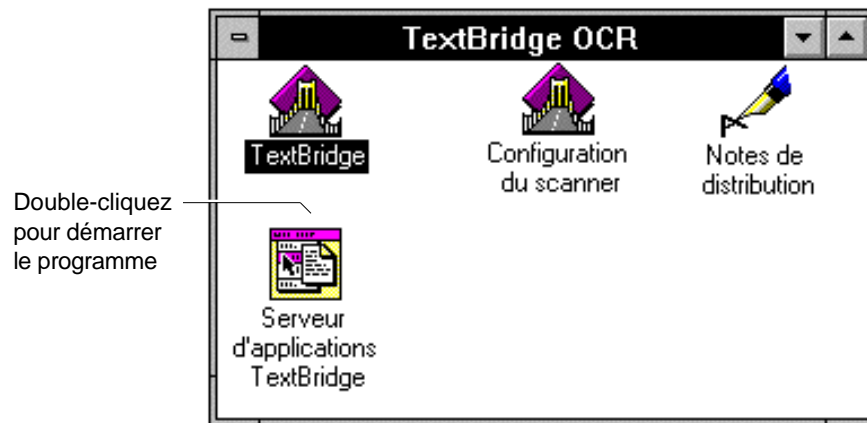
Le Serveur d'applications TextBridge (ou TAS, de l'anglais TextBridge Application Server) est un programme Windows qui peut être "rattaché" à d'autres applications de texte Windows et exécuté à partir de celles-ci.

Une fois relié, TAS s'affiche dans le menu Fichier de l'application hôte en tant que commande **TextBridge OCR**. Lorsque vous sélectionnez cette dernière, la boîte de dialogue principale de TextBridge s'affiche comme s'il s'agissait d'une boîte de dialogue de l'application hôte. Vous pouvez ensuite configurer et lancer la reconnaissance optique exactement comme vous le feriez avec le programme standard TextBridge.

## Démarrer TAS et enregistrer des applications

TAS est implémenté au cours du processus d'installation de TextBridge décrit dans le chapitre 2 de ce manuel.

A la fin de ce processus, le programme d'installation (setup) crée un groupe de programmes TextBridge OCR qui inclut le Serveur d'applications TextBridge (figure 4-14).



**Figure 4-14.** Icône du Serveur d'applications TextBridge (TAS)

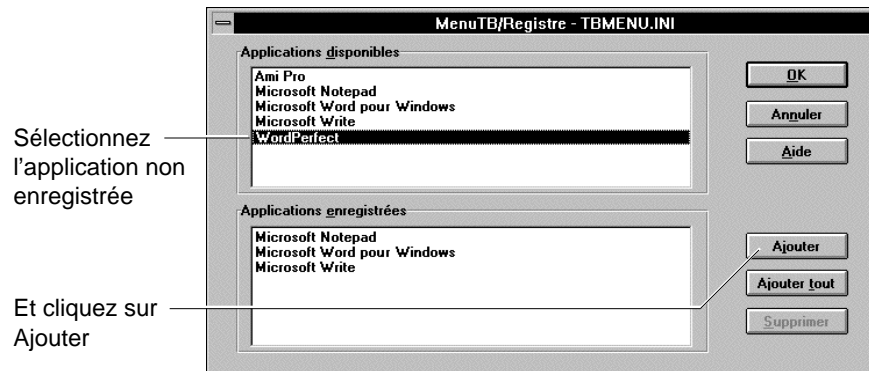
Avant d'exécuter TAS à partir d'une application, vous devez démarrer le programme et, si nécessaire, enregistrer cette application :

- 1. Double-cliquez sur l'icône TAS dans le groupe de programmes TextBridge OCR.**

Le programme démarre et s'affiche en tant qu'icône sur votre bureau Windows.

## 2. Double-cliquez sur cette icône sur le bureau Windows.

La boîte de dialogue d'enregistrement du serveur d'applications TextBridge s'affiche (figure 4–15).



**Figure 4–15.** Boîte de dialogue Enregistrement du Serveur d'applications TextBridge (TAS)

## 3. Enregistrez l'application si nécessaire.

- Mettez en surbrillance l'application que vous souhaitez enregistrer (en haut de la boîte de dialogue Enregistrement).
- Cliquez sur le bouton Ajouter pour ajouter l'application à la liste des applications enregistrées, affichée dans la partie inférieure de la boîte de dialogue.
- cliquez sur OK une fois l'enregistrement de vos applications terminée.

Vous pouvez maintenant utiliser TextBridge à partir d'une application enregistrée.

## Exécuter TAS à partir de votre application

La commande TextBridge OCR apparaît au bas du menu Fichier de l'application enregistrée active (juste au-dessus de la commande Quitter).

---

**Remarque** Pour que TAS puisse fonctionner, l'application hôte doit comporter un menu Fichier incluant une commande Quitter. La plupart des applications Windows disposent d'un tel menu.

---

La figure 4-16 illustre la commande TextBridge OCR dans le menu Fichier de Microsoft Write®.



**Figure 4-16.** Commande TextBridge OCR

Procédez comme suit pour exécuter TAS et importer un texte reconnu directement dans le document (ouvert) de l'application hôte.



**1. Démarrez le Serveur d'applications TextBridge.**

Double-cliquez sur son icône dans le groupe de programme TextBridge OCR (voir figure 4–14).

- + Pour que TAS démarre automatiquement chaque fois que vous lancez Windows, placez-le dans le groupe de programmes Démarrage.

**2. Assurez-vous que l'application hôte est enregistrée.**

Reportez-vous à la procédure énoncée dans la section précédente, “Démarrer TAS et enregistrer des applications”.

**3. Démarrez l'application hôte.**

Ouvrez dans celle-ci le document (nouveau ou existant) dans lequel vous désirez importer le texte reconnu.

**4. Sélectionnez la commande TextBridge OCR dans le menu Fichier de l'application hôte.**

Les messages d'état suivant apparaissent :

Connexion aux services TextBridge...

Connexion aux services TextBridge  
établie.

La boîte de dialogue principale de TextBridge s'affiche après quelques instants.

## 5. Configurez et lancez la reconnaissance optique à partir de la boîte de dialogue principale.

Une fois la boîte de dialogue principale affichée, choisissez la source Fichier ou Scanner, spécifiez vos préférences et procédez exactement comme si vous utilisiez TextBridge en tant qu'application autonome.

- + Les fonctionnalités Visualiser, Vérifier et Enregistrer images de pages ne sont **pas** disponibles dans la version TAS de TextBridge. Si elles vous sont nécessaires, exécutez TextBridge en tant qu'application autonome et enregistrez le texte reconnu dans un format texte.

Reportez-vous au chapitre 3 pour des informations détaillées sur l'utilisation de TextBridge ; consultez également les sections précédentes de ce chapitre pour des conseils et techniques d'utilisation.

Une fois la reconnaissance optique achevée, TAS se ferme et le texte reconnu s'affiche à l'emplacement du curseur, dans le document ouvert à l'intérieur de l'application.

- + TAS utilise le Presse-papiers Windows pour couper et coller le texte reconnu dans votre application, en format **RTF** (Rich Text Format) ou **ASCII**. Si votre application permet de coller un texte en format RTF à partir du presse-papiers, ce format sera utilisé. Sinon, le texte reconnu sera collé en format ASCII et les données de formatage (gras, italique, etc.) ne seront pas conservées.

## Utiliser l’Imprimante OCR TextBridge

Sous ces autres formes, TextBridge pour Windows ne peut reconnaître des fichiers d’images que s’ils sont stockés en format TIFF. Certaines applications, en particulier les programmes de télécopie, ne stockent les images des fax qu’en format PCX, DCX ou autre format propriétaire.

Pour effectuer la reconnaissance optique des images non TIFF, utilisez **l’Imprimante OCR TextBridge** (OCR Printer). Elle apparaît dans les applications Windows en tant qu’imprimante de destination. Elle vous permet d’“imprimer” une image à partir d’une application Windows et d’obtenir un fichier texte reconnu et formaté.

Dans la plupart des cas, vous utiliserez l’Imprimante OCR pour effectuer la reconnaissance d’une image de page directement à partir d’une application de télécopie ou d’imagerie. L’Imprimante OCR est similaire à la fonctionnalité utilisée par de nombreux programmes de télécopie pour envoyer des fax. L’image du fax est “imprimée” sur le fax modem puis envoyée à un autre fax modem ou télécopieur.

L’Imprimante OCR possède l’avantage de pouvoir reconnaître pratiquement tous les formats d’image (DCX, PCX, Corel, TIFF, etc.). Presque tous les programmes Windows conçus pour traiter des images peuvent l’utiliser.

Pour préparer l’Imprimante OCR, reportez-vous aux sections “Ajouter l’Imprimante OCR”. Consultez ensuite “Utiliser l’Imprimante OCR dans votre application d’imagerie” pour obtenir des instructions sur la manière de télécharger une image et de produire un fichier texte.

## Ajouter l'Imprimante OCR

Les fichiers du programme Imprimante OCR sont installés en même temps que l'application principale de TextBridge.

Toutefois, vous devez ajouter l'Imprimante OCR à la liste des imprimantes à la disposition des applications Windows de votre PC (comme s'il s'agissait d'un gestionnaire ou "driver" d'imprimante).

Cette procédure suppose que vous avez déjà installé TextBridge en respectant les instructions du chapitre 2.

1. **Ouvrez le Groupe principal à partir du Gestionnaire de programmes Windows, et double-cliquez sur l'icône Panneau de configuration :**

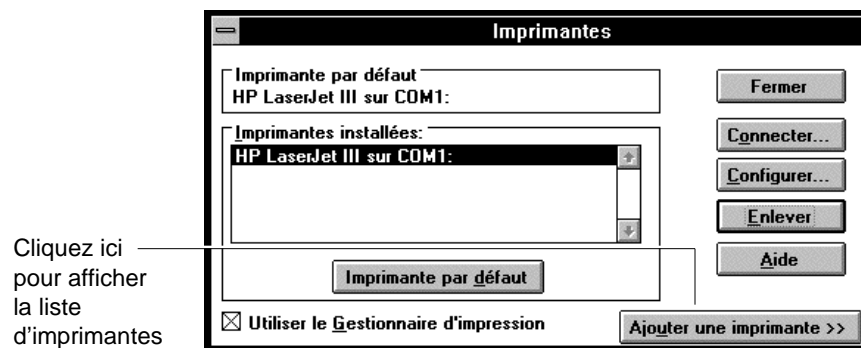


La fenêtre Panneau de configuration s'affiche. Elle contient des icônes pour divers éléments du système.

2. **Double-cliquez sur l'icône Imprimantes dans la fenêtre Panneau de configuration.**



La boîte de dialogue Imprimantes de Windows s'affiche (figure 4-17).



**Figure 4-17.** Boîte de dialogue Imprimantes

- 3. Cliquez sur le bouton Ajouter une imprimante dans la boîte de dialogue Imprimantes.**

La boîte de dialogue s'étend pour afficher la liste des imprimantes pouvant être ajoutées.

- 4. Mettez le choix suivant en surbrillance puis cliquez sur le bouton Installer.**

Imprimante non listée ou mise à jour

La boîte de dialogue Gestionnaires installés s'affiche, vous demandant de spécifier un lecteur et un répertoire pour le gestionnaire d'imprimante.

- 5. Entrez le chemin d'accès au répertoire BIN de TextBridge dans la boîte de dialogue Gestionnaires installés :**

c:\txbridge\bin

**6. Cliquez sur OK (ou appuyez sur Entrée).**

La boîte de dialogue Imprimante non listée ou mise à jour s'affiche (figure 4–18).



**Figure 4–18.** Boîte de dialogue Imprimante non listée ou mise à jour

**7. Sélectionnez Imprimante OCR et cliquez sur OK.**

La boîte de dialogue Imprimante non listée ou mise à jour se ferme ; la boîte de dialogue Imprimantes reste affichée.

**8. Cliquez sur Fermer dans la boîte de dialogue Imprimantes pour terminer la procédure d'ajout.**

Vous pouvez maintenant utiliser l'Imprimante OCR TextBridge comme le décrit la section suivante.

## Utiliser l'Imprimante OCR dans votre application d'imagerie

Vous pouvez utiliser l'Imprimante OCR avec n'importe quelle application Windows capable d'ouvrir et de visualiser un fichier d'image.

Par exemple, WinFax Pro (de Delrina Technology Inc.) propose un afficheur (Image Viewer) pour visualiser, manipuler et imprimer les images des fax. À l'aide la commande Imprimer du menu Fichier de l'Image Viewer WinFax, vous pouvez spécifier et utiliser l'Imprimante OCR pour effectuer la reconnaissance de caractères sur une image de fax.

L'imprimante OCR vous permet d'accéder aux préférences TextBridge avant de commencer la reconnaissance. Une fois la reconnaissance terminée, vous pouvez spécifier le nom du fichier texte de sortie, son emplacement et son format dans la boîte de dialogue Enregistrer sous de TextBridge.

Procédez comme suit pour utiliser l'Imprimante OCR dans votre application.

### **1. Ouvrez l'application d'imagerie et affichez l'image à reconnaître.**

- + L'image doit être binaire (en noir et blanc) et d'une résolution supportée par TextBridge. TextBridge peut reconnaître des images en 100x200, 200x100, 200x200, 300x300 et 400x400 points par pouce.

**2. Spécifiez l’Imprimante OCR TextBridge en tant qu’imprimante de destination dans la boîte de dialogue Configuration d’impression (Print Setup) de votre application.**

A ce stade, vous devriez avoir chargé l’Imprimante OCR dans le Panneau de configuration Windows, comme le décrit la section précédente, “Ajouter l’Imprimante OCR”.

**3. Définissez, si vous le désirez, vos préférences TextBridge.**

- Cliquez sur le bouton Options ou Configurer (Setup) dans la boîte de dialogue Configuration d’impression de votre application. Une autre boîte de dialogue de configuration s’affichera.
  - Cliquez sur le bouton Préférences de la boîte de dialogue de configuration pour afficher la boîte de dialogue Préférences de TextBridge.
  - Spécifiez la qualité de document standard ou fax, l’orientation de page, la segmentation de page automatique, etc.
  - Lorsque vous avez terminé, cliquez sur OK dans la boîte de dialogue Préférences et sortez des autres boîtes de dialogue.
- + Si la boîte de dialogue Configuration d’impression de votre application comporte une option permettant d’utiliser la résolution de l’imprimante, **activez-la**.



**4. Commencez le processus de reconnaissance optique de l'image affichée.**

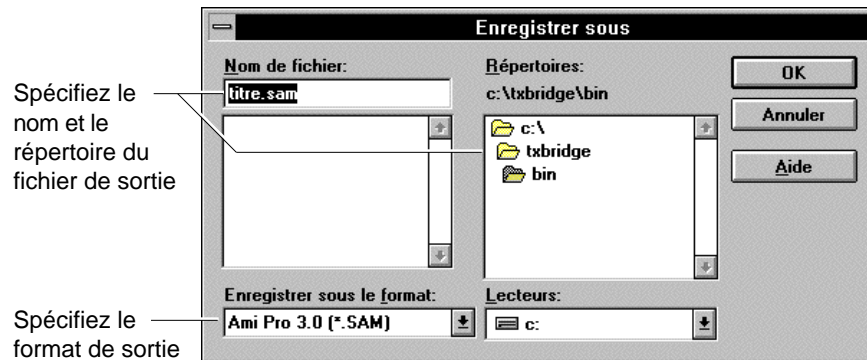
- Sélectionnez Imprimer dans le menu Fichier de votre application.
- Cliquez sur OK dans la boîte de dialogue Imprimer. Les messages suivants s'affichent pendant la reconnaissance :

Traitement...

Acquisition de l'image...

Reconnaissance du texte...

Une fois la reconnaissance terminée, la boîte de dialogue Enregistrer sous de TextBridge s'affiche (figure 4-19).



**Figure 4-19.** Boîte de dialogue Enregistrer sous

**5. Spécifiez le nom du fichier de sortie, son format, le lecteur et le répertoire de destination, puis cliquez sur OK.**

Le fichier texte reconnu est converti dans le format spécifié et écrit sur votre disque dur.