

## **4 SUGERENCIAS Y TÉCNICAS**

En los tres primeros capítulos de esta guía se proporciona información sobre la instalación del sistema y sobre las funciones básicas que permiten familiarizarse con el funcionamiento de TextBridge.

En este capítulo se describen los métodos para mejorar el resultado del reconocimiento en TextBridge. Más concretamente se tratan los siguientes temas:

- Cómo obtener un reconocimiento de texto excelente.
- Sugerencias para un procesamiento más eficaz.
- Guardar imágenes de página.
- Ejecutar TextBridge OCR desde otras aplicaciones.

### **CÓMO OBTENER UN RECONOCIMIENTO DE TEXTO EXCELENTE**

El software TextBridge OCR permite obtener el mejor reconocimiento de caracteres en una gran variedad de documentos. Sin embargo, existen algunas acciones que pueden ayudarle a obtener el mejor reconocimiento de caracteres posible para un documento en particular.

En esta sección se ofrecen algunas sugerencias para mejorar el reconocimiento de texto. A continuación se indican los temas que van a tratarse:

- Uso y mantenimiento del escáner.
- Ajustar el brillo del escáner.
- Ajustar los colores.
- Los filtros de fax.
- Procesar varios documentos por separado.
- El verificador de palabras.
- El comando 'Invertir' en 'Muestra'.

## Uso y mantenimiento del escáner

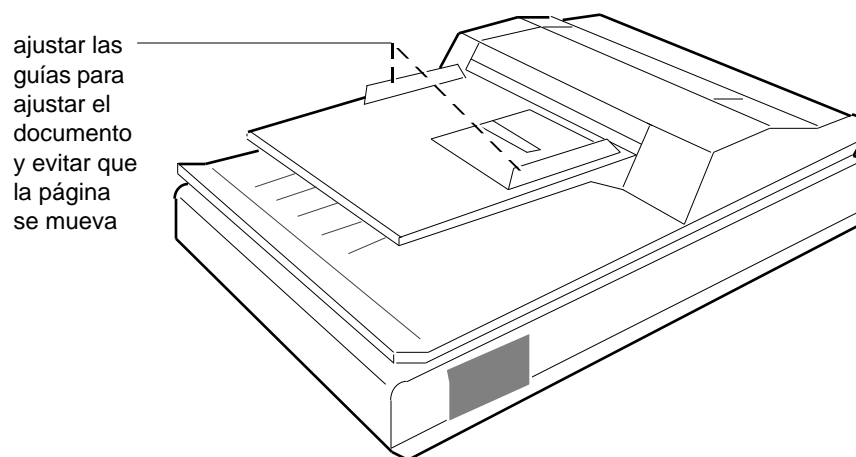
El uso y mantenimiento que se haga del escáner puede suponer una gran diferencia en la calidad de la exploración:

- **Conocer el escáner.** Lea detenidamente y documentación del escáner.
- **Mantenimiento del escáner.** Mantenga el escáner limpio y sin polvo. El lector de vidrio (superficie plana) del escáner debe mantenerse limpio y sin marcas que puedan ser captadas durante la exploración.

- **Cargar el escáner correctamente.** Asegúrese explorar el documento no de lado. Esto dificulta el reconocimiento de caracteres.

Si se usa un alimentador de documentos automático, asegúrese de que las guías estén alineadas correctamente con las páginas que se están explorando (Figura 4-1).

Si se usa la superficie plana, asegúrese de que la imagen de página está bien alineada con la superficie. Algunas veces la imagen está torcida en relación con el resto del texto de la página. Corrija este tipo de problema tanto como sea posible antes de empezar la exploración.

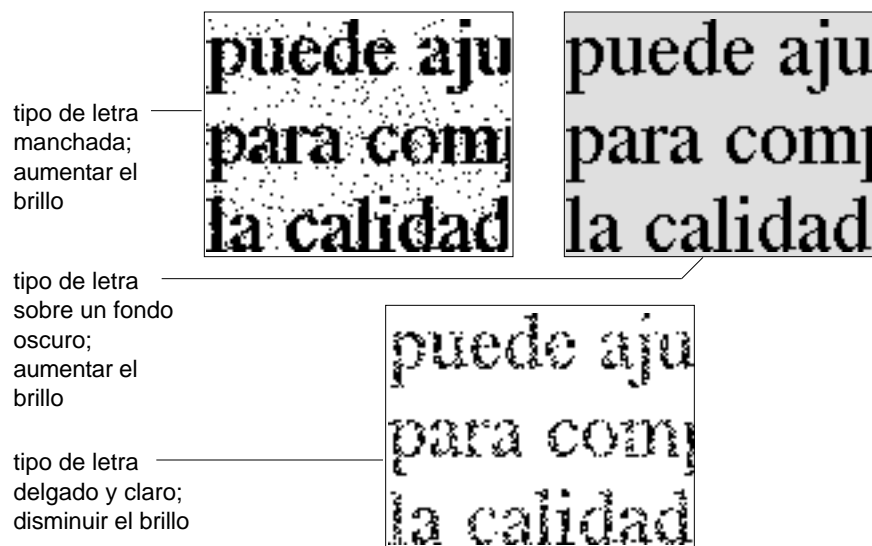


**Figura 4-1.** Colocación de la página en el escáner.

## Ajustar el brillo del escáner

Durante la exploración, una de las asignaciones más importantes y que afecta el resultado del reconocimiento de caracteres es el **brillo** del escáner. Como se ilustra en la Figura 4–2, la calidad de los documentos puede variar considerablemente.

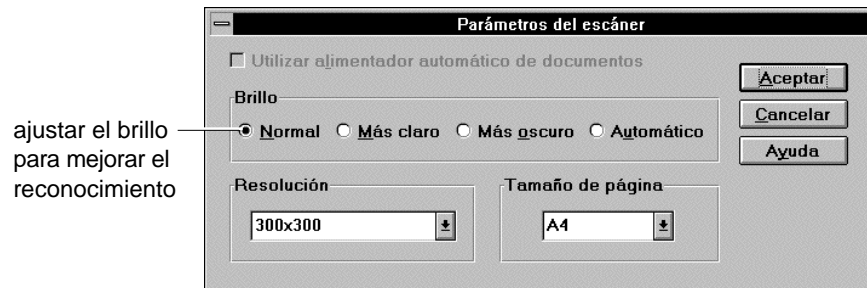
Tanto la oscuridad del texto, la claridad del fondo y la cantidad de **suciedad** en la página (manchas, tinta corrida, huellas, escritura a mano y otras marcas) pueden afectar la precisión del reconocimiento de caracteres.



**Figura 4–2.** Documentos originales y brillo.

Desde el cuadro ‘Parámetros del escáner’ (Figura 4–3), puede ajustarse el brillo para compensar la calidad de la impresión y el fondo del documento. Los ‘Parámetros del escáner’ se acceden a partir del cuadro ‘Preferencias’.

Si los caracteres impresos son demasiado oscuros, están muy juntos, son borrosos o están separados por un espacio muy pequeño (como en algunas revistas) seleccione el nivel de brillo 'Más claro'. El reconocimiento de caracteres en documentos con manchas en el fondo o con el fondo de color, puede mejorar considerablemente aumentando la asignación de brillo.



**Figura 4-3.** Asignaciones de brillo del escáner.

Seleccione un nivel de brillo 'Más oscuro' si los caracteres de la página están poco perfilados, aparecen quebrados o son muy finos.

Si el escáner respalda el brillo automático, seleccione este parámetro para obtener el mejor nivel de brillo en cada página de un documento.

Otra forma de saber si el brillo es adecuado o no para el texto que se está explorado es mostrar primero la página y acercar la imagen a una resolución completa en la ventana 'Muestra'. Esto permite ver la salida del escáner que "ve" el sistema (Figura 4-4). Si la imagen mostrada no parece tener el brillo adecuado, puede ajustar este parámetro y volver a explorar el documento.



**Figura 4-4.** Ampliación de la página en 'Muestra'.

## Ajustar los colores

Todo escáner tiene uno o varios colores que no puede leer. Estos son llamados **colores omitidos**. Remítase a la documentación del escáner para determinar los colores omitidos de su escáner.

- + Si en la documentación del escáner no se especifican los colores omitidos, examine el color de la luz del escáner al moverse por la superficie plana del escáner. El color de la luz determina el color omitido. Muchos escáneres tienen una luz verde pálida. Entonces, el color omitido es el verde pálido.

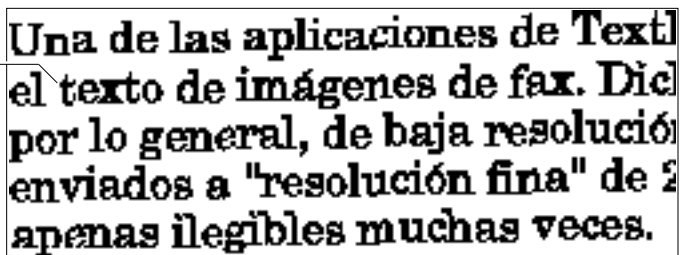
Además de los colores omitidos, puede haber otros colores con los que el escáner tenga problemas. Si el texto (o imagen) que se está explorando es de color o si está impreso sobre un fondo de color, intente ajustar la asignación de brillo.

Si todavía no obtiene el resultado adecuado, intente fotocopiar la página y explorar la copia en blanco y negro.

## Los filtros de fax

Una de las funciones de TextBridge es la de reconocer el texto en imágenes de fax. Las imágenes de fax son comúnmente de baja resolución (100 por 200, 200 por 100 ó 200 por 200 puntos por pulgada). Incluso los fax con una “Alta resolución” de 200 por 200 ppp son por lo general sólo leíbles parcialmente (Figura 4–5).

las imágenes de fax tienen una resolución baja y a menudo están torcidas



**Figura 4–5.** Imagen de fax.

Para reconocer imágenes de fax, TextBridge proporciona la asignación **Fax** en el cuadro ‘Preferencias’. Este filtro de ‘Calidad del documento’ efectúa un paso previo al procesamiento que mejora la imagen de fax antes de efectuar el reconocimiento.

El interruptor de fax funciona en imágenes de fax almacenadas en archivos TIFF y en fax impresos explorados en resoluciones más altas (por ejemplo, 300 ppp).

---

**Nota** No use el filtro 'Fax' en otros documentos explorados o en línea que no sean fax. De hacerlo, la exactitud del reconocimiento puede verse afectado. Si observa que el reconocimiento no es de gran calidad en imágenes de fax sintetizadas (por ejemplo, un documento creado en una aplicación de procesamiento de palabras "impreso" en un módem fax) desactive el filtro 'Fax' e intente verificar parte del texto durante el reconocimiento.

---

## Procesar varios documentos por separado

El software TextBridge OCR usa una gran variedad de técnicas de inteligencia artificial para reconocer texto.

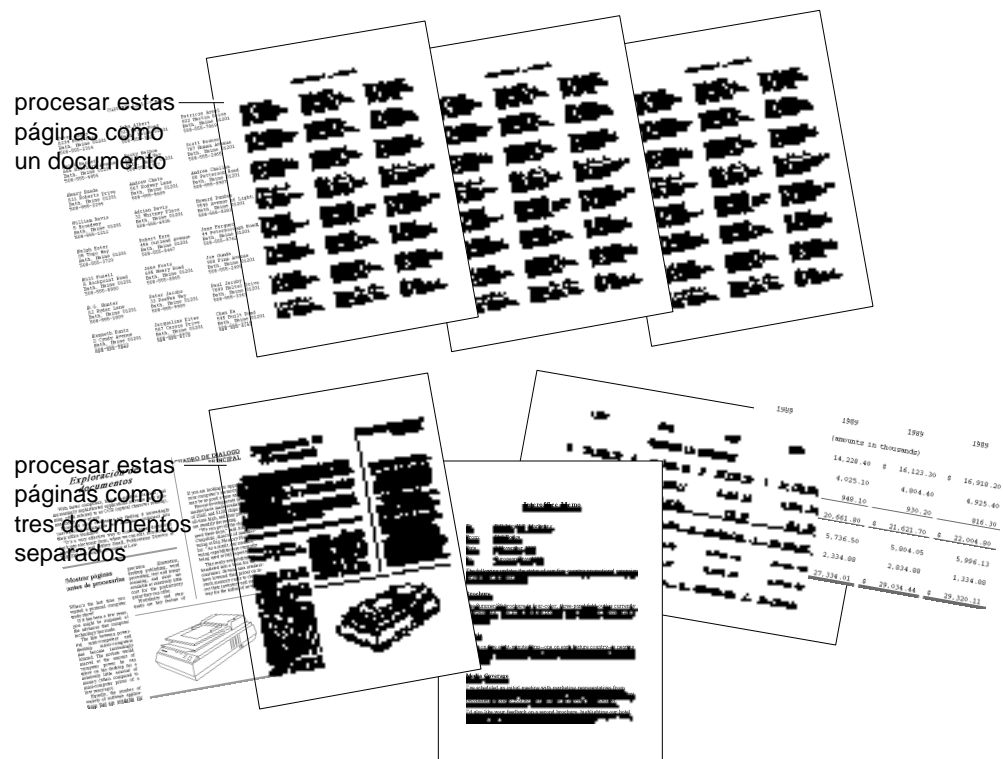
Con estas técnicas, TextBridge se instruye así mismo sobre lo que está reconociendo.

De este modo, TextBridge mejora la precisión del reconocimiento y la velocidad a medida que avanza la exploración y el reconocimiento de las páginas del documento.

Sin embargo, esta capacidad de aprender puede verse afectada si se procesan páginas de distintos documentos al mismo archivo de salida.

TextBridge espera la segunda página y las siguientes de un documento tengan el mismo tipo de letra que la reconocida en la primera página (Figura 4–6).





**Figura 4-6.** Procesamiento de varios documentos.

Si la segunda página es un documento totalmente distinto con distintos tipos de letra y tamaño, el conocimiento adquirido sobre el texto en la primera página ya no es válido.

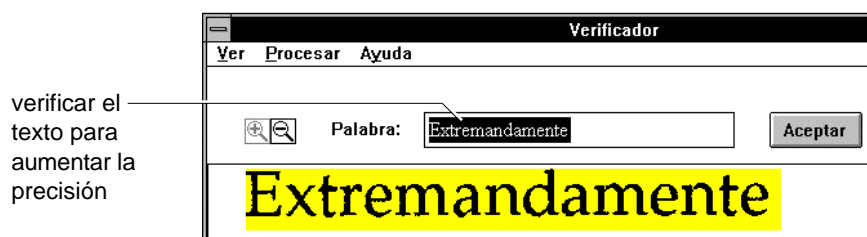
TextBridge debe volver a empezar el proceso de aprendizaje de nuevo para la segunda página (y sucesivas).

Si desea explorar varios documentos y obtener los mejores resultados, explore cada documento como documentos separados.

## El verificador de palabras

Si el resultado obtenido en un documento en particular, no es satisfactorio use el verificador de palabras par mejorar la exactitud del reconocimiento.

Cuando el usuario interactúa con el proceso OCR en la ventana 'Verificador' (Figura 4–7), se está enseñando a TextBridge a reconocer los caracteres y palabras del documento. Esto mejorará la exactitud del reconocimiento.



**Figura 4–7.** Ventana 'Verificador'.

Cada palabra que TextBridge es incapaz de reconocer claramente se muestra en la casilla de edición 'Palabra' en la parte superior de la ventana 'Verificador'. La imagen de la palabra está resaltado en contexto en la parte inferior.

Con el 'Verificador' puede examinar el texto reconocido y aceptar o corregir las decisiones tomadas por TextBridge. Esto ayuda a TextBridge a mejorar el reconocimiento del texto a medida que avanza el documento.

Por lo general en un documento de varias páginas, se verifican una o dos páginas y después se finaliza la verificación. A partir de este momento TextBridge usará las correcciones y verificaciones del usuario para tomar decisiones más acertadas en el resto del documento.

Para obtener el mejor resultado en documentos con pocas páginas, puede verificarse el documento completo.

Para más información sobre cómo usar el verificador de palabras y sus opciones, remítase al capítulo 3.

## El comando Invertir en Muestra

TextBridge es capaz de reconocer archivos TIFF en línea provenientes de módem fax o de otras fuentes.

A veces, los datos de imagen son guardados con los elementos de la foto (**pixels**) invertidos en el archivo resultante: el fondo de la página blanco es negro y la impresión es blanca. Esto suele suceder en archivos Intel FAXability.

TextBridge **no puede** reconocer estos archivos. Para poder reconocer un archivo TIFF en línea es necesario que la imagen contenga el texto en negro sobre un fondo blanco.

Para poder efectuar un reconocimiento de archivos con imágenes invertidas, TextBridge proporciona el comando 'Invertir' en el menú 'Ver' de la ventana 'Muestra' (Figura 4-8).



Figura 4-8. Invertir un documento.

Si no está seguro de si el archivo está invertido o no, visualícelo en la ventana 'Muestra' antes de empezar el reconocimiento. Si se muestra con el texto en blanco sobre un fondo negro, abra el menú 'Ver' y haga clic en el comando 'Invertir'. TextBridge invierte la imagen.

Ahora ya puede empezar el proceso de reconocimiento. Observe que la inversión debe efectuarse manualmente para cada archivo TIFF almacenado de esta forma.

Para más información sobre el uso de 'Muestra', remítase al capítulo 3.

## **SUGERENCIAS PARA UN PROCESAMIENTO MÁS EFICAZ**

Al principio de usar TextBridge, quizás le sea más fácil explorar un documento sin ajustar las preferencias predeterminadas. Las preferencias predeterminadas proporcionan muy buenos resultados en la mayoría de casos.

Sin embargo si desea obtener un rendimiento inmejorable, siga las sugerencias que se indican a continuación:

- use la herramienta de zona en 'Muestra'.
- use la asignación 'Ignorar fotos/medios tonos'.
- use la orientación de página automática cuando sea necesario.
- use la segmentación automática para documentos con varias columnas.

Usando estas prestaciones se asegura que el sistema va a procesar únicamente las partes de una página que son esenciales y las procesa correctamente.

A largo plazo estas prestaciones pueden ahorrarle mucho tiempo.

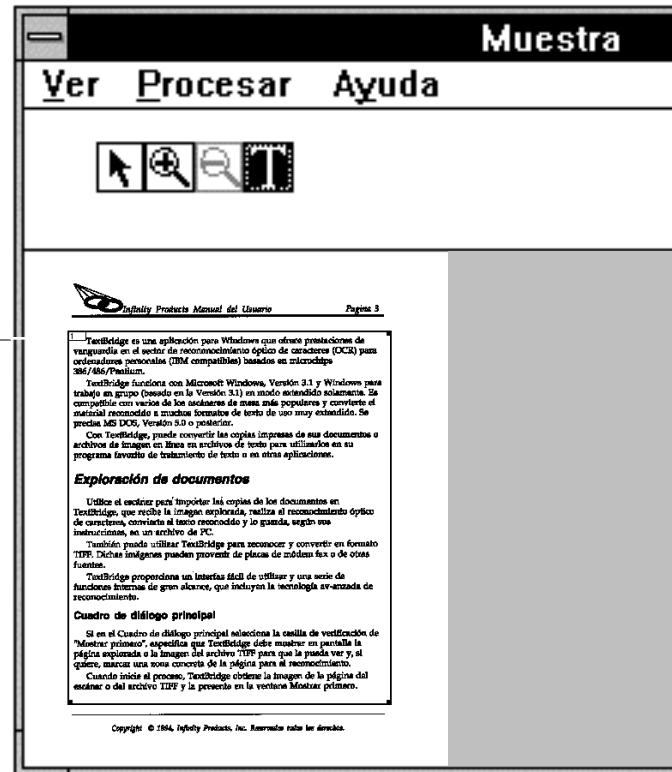
## Definir una zona para capturar sólo los datos deseados

Algunos documentos pueden contener logotipos, gráficos, encabezados y pies de página, así como otros elementos de página que no son necesarios capturar y que hacen que el sistema vaya más despacio.

Con la herramienta para definir zonas de la ventana 'Muestra', puede identificarse la parte de la página que desea capturar (Figura 4–9).

Remítase al capítulo 3 para más información sobre cómo usar las herramientas de muestra.

definir zona  
para procesar  
sólo el texto  
que desee



**Figura 4–9.** Definición de una zona en 'Muestra'.

## La opción Ignorar fotos/medios tonos

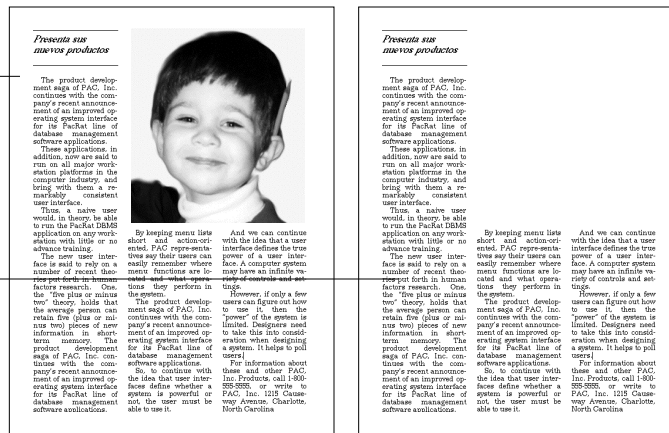
En un documento impreso, una fotografía de medios tonos está compuesta por puntos negros de distintos tamaño. Normalmente, TextBridge tarda bastante tiempo ya que intenta reconocer los puntos de medios tonos como si fueran texto.

Eventualmente, TextBridge reconocer que ha estado intentando reconocer medios tonos y cesa de hacerlo. Para agilizar el proceso de reconocimiento en textos que contienen medios tonos, puede desactivarse la definición **‘Ignorar fotos/medios tonos’** antes de empezar el reconocimiento.

Con la opción ‘Ignorar fotos/medios tonos’ activada, TextBridge explora rápidamente la imagen de página y oculta los medios tonos **antes** de empezar el reconocimiento de caracteres (Figura 4–10). De esta forma se agiliza el reconocimiento de caracteres a la vez que se hace más eficiente.

TextBridge  
intenta  
reconocer los  
medios tonos  
como texto

con el filtro, los  
medios tonos  
son eliminados  
antes de  
empezar el  
reconocimiento



**Figura 4–10.** El filtro ‘Ignorar fotos/medios tonos’.

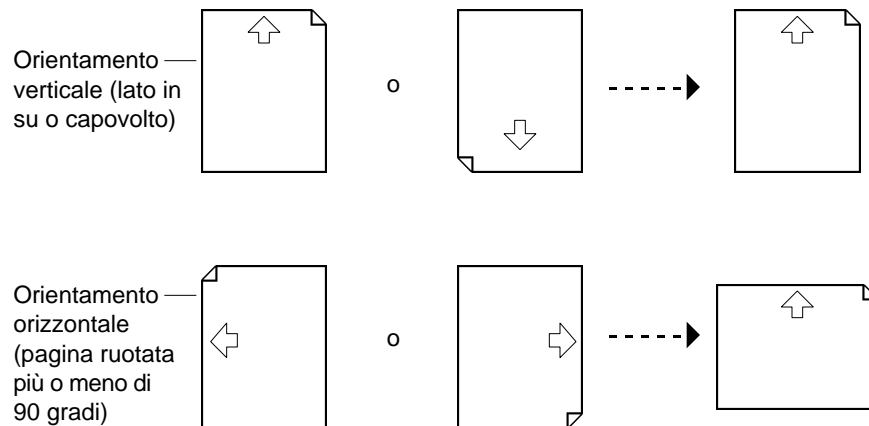
Para usar la opción ‘Ignorar fotos/medios’ tonos, seleccione el botón ‘Preferencias’ del menú principal. En el cuadro ‘Preferencias’, haga clic para marcar la casilla ‘Ignorar fotos/medios tonos’.

- + Aunque el procedimiento de filtrar las fotos y los medios tonos sea muy rápido, no desactive la opción si no contiene medios tonos.

## La orientación automática de página

TextBridge proporciona una herramienta que determina automáticamente la orientación de una página, la gira en memoria de ser necesario y después empieza el reconocimiento (Figura 4–11).

Especifique **‘Orientación automática de página’** en el cuadro ‘Preferencias’ al que puede accederse a partir del cuadro principal. Esta opción puede ser de utilidad en casos como:



**Figura 4–11.** La orientación de página automática.

- al procesarse documentos con páginas con varias orientaciones.
- al procesarse un archivo TIFF y no se sabe con certeza la orientación de la imagen que contiene.

En el primer caso, podría procesarse un documento cuyas páginas fueran en su mayoría verticales y algunas horizontales.

TextBridge explora cada página, determina si es vertical, horizontal (90 grados o 270 grados) o bien si la página está invertida, de ser este el caso la gira (0 grados), antes de empezar el reconocimiento.

En el segundo caso, si la imagen TIFF que se va a reconocer está de lado o invertida, TextBridge la girará correctamente y después efectuará el reconocimiento.

- + La orientación automática es un paso del procesamiento que tiene lugar antes del reconocimiento. Por lo que para obtener el reconocimiento más rápido, use la orientación automática **sólo** cuando sea necesario.

La segmentación automática de página

TextBridge proporciona una herramienta que ubica automáticamente las **regiones** de texto en la página, define el orden y después efectúa el reconocimiento. Esta función de **segmentación de página automática** es crítica para el reconocimiento de páginas que tiene más de una columna y/o disposiciones de página poco comunes (Figura 4–12).

la segmentación automática analiza la distribución de la página

después, identifica y ordena las regiones

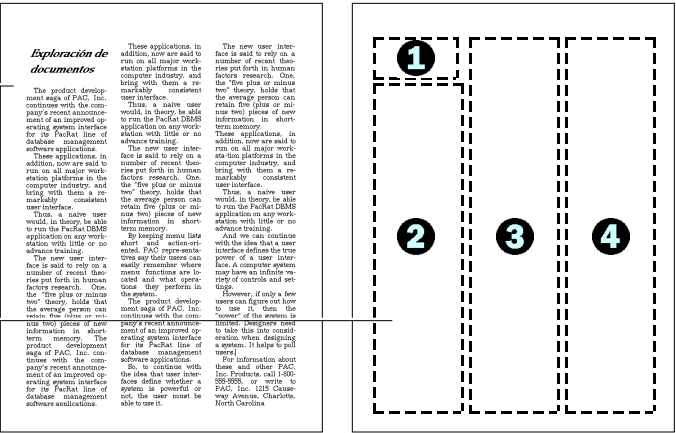


Figura 4–12. Segmentación de página y disposición de las regiones.



**Importante**

---

Es **necesario** activar la segmentación de página si van a procesarse páginas con más de una columna de texto. De lo contrario, TextBridge puede generar regiones de texto reconocido en el orden incorrecto. No utilice la segmentación de página automática en documentos con una sola columna.

---

La segmentación de página es un estado previo al procesamiento y ocurre antes de empezar el reconocimiento.

Observe que puede crearse una zona en 'Muestra' y usar la segmentación de página. La segmentación de página funcionará únicamente dentro de la zona. Por ejemplo, si se dibuja una zona alrededor de dos columnas en una página de tres columnas, la segmentación de página detectará y colocará en orden las dos columnas antes de empezar el reconocimiento.

## GUARDAR IMÁGENES DE PÁGINA

Una de las opciones de las casillas de verificación en el cuadro principal es **‘Guardar imagen de página’**. Esta opción, disponible al definir ‘Entrada’ por ‘Escáner’, permite guardar una imagen binaria (blanco y negro) de cada página explorada durante la sesión de reconocimiento óptico de caracteres de TextBridge.

---

**Nota** TextBridge guarda las imágenes de página como archivos TIFF con compresión CCITT Group 3. Group 3 es un estándar especificado por CCITT (siglas de Comité Consultivo Internacional de Telefonía y Telegrafía), organización de estándares internacionales.

---

Al hacer clic sobre ‘Empezar’ y explorarse la primera página, TextBridge visualiza el cuadro ‘Guardar imagen de página como’ (Figura 4–13).



**Figura 4–13.** Cuadro ‘Guardar imagen de página como’.

El cuadro ‘Guardar imagen de página como’ es muy parecido al cuadro ‘Guardar como’ estándar de Windows en el cual puede especificarse un nombre de archivo, el tipo de archivo, el directorio de salida y la unidad de disco.

El nombre del archivo es el **nombre base** sobre el que se crea los nombres del archivo de imagen de página. También es el **nombre del documento** predeterminado que aparece en el cuadro ‘Guardar como’ de TextBridge cuando se ha completado el reconocimiento. El nombre del archivo es *títular* con la extensión *.tif*.

Por ejemplo, supongamos que se ha especificado el nombre “guía” en el cuadro ‘Guardar imagen de página como’. Los archivos de imagen de página van a guardarse con un nombre en el formato:

*guiannnn.TIF*

donde *nnnn* es el número de página del documento precedido de ceros (por ejemplo, 0001, 0002 etc.).

Los archivos de imagen se les asigna un nombre secuencial dentro del directorio. Si ya existe un archivo con el mismo nombre (por ejemplo, “*guia0001.tif*”), TextBridge asignará un nombre con el número siguiente.

Al finalizar el trabajo, el nombre del documento (por ejemplo “*guia*”) se visualiza automáticamente en la casilla ‘Nombre de archivo’ del cuadro ‘Guardar como’.

En el cuadro ‘Guardar imagen de página’, el directorio de trabajo inicial es el directorio desde el que se lanza la aplicación TextBridge:

*C:\TXBRIDGE\BIN*

Sin embargo puede especificarse cualquier otra unidad de disco y directorio al cual almacenar las imágenes de página. Este será ahora el nuevo **directorio de trabajo** para este trabajo y como el nombre del documento, va a visualizarse en el cuadro ‘Guardar como’ al finalizar el reconocimiento.

Para el formato de archivos de imagen de página, el menú 'Mostrar archivos de tipo' proporciona una sola selección, TIFF CCITT-3 Intel. Las imágenes de página son guardadas exactamente de la misma manera que son exploradas en formato binario (blanco y negro).

Observe que si se selecciona 'Cancelar' en el cuadro 'Guardar imagen de página como', el cuadro se cierra y TextBridge finaliza el trabajo. El cuadro principal permanece listo para empezar de nuevo. Por ejemplo, si no desea guardar las imágenes de página, puede desactivar la opción 'Guardar imagen de página' y volver a empezar el trabajo.

## EJECUTAR TEXTBRIDGE DESDE OTRAS APLICACIONES

TextBridge OCR para Windows es un grupo de aplicaciones que permiten ejecutar OCR desde cualquier aplicación Windows.

Además de la utilidad principal que se ejecuta como un programa independiente y es además el más completo, TextBridge OCR se proporciona en otros dos formatos:

- **Servidor de aplicaciones TextBridge**, programa que actúa con un elemento de menú desde cualquier aplicación de texto Windows (tratamiento de texto, programas de edición electrónica, hoja electrónica, bases de datos, etc.).
- **Impresora TextBridge OCR (TextBridge OCR Printer)**, capacidad que permite enviar una imagen en cualquier formato a una versión de TextBridge OCR que funciona como un controlador de impresión común.

En esta sección se proporciona información sobre el uso de TextBridge OCR en estos formatos.

---

**Nota** TextBridge respalda también la interfaz DDE. Si desea más información póngase en contacto con el servicio técnico de Xerox Imaging Systems.

---

## El Servidor de aplicaciones TextBridge

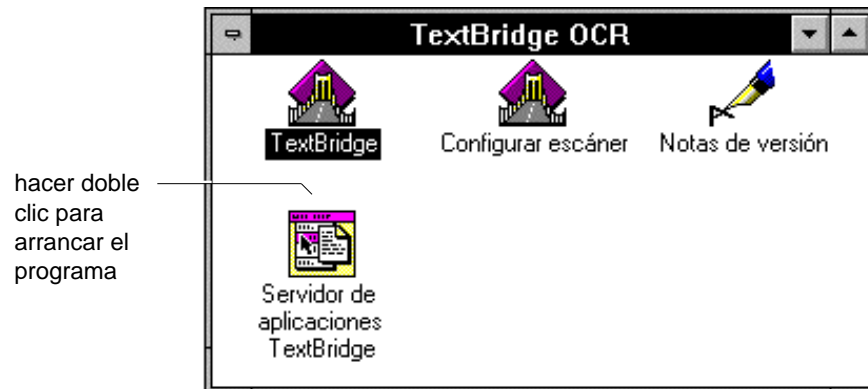
El ‘Servidor de aplicaciones TextBridge’ es un programa Windows que puede “anexarse” a otras aplicaciones de texto Windows por lo que puede ejecutarse desde Windows.

Una vez anexado, aparece el servidor de aplicaciones TextBridge en el menú ‘Archivo’ de la aplicación principal como el comando **‘TextBridge OCR’**. Al seleccionar ‘TextBridge OCR’, aparece el cuadro de dialogo principal de TextBridge como si fuera un cuadro de la aplicación principal. Desde aquí puede instalarse y arrancar OCR como si estuviera en el programa TextBridge.

### Arrancar el Servidor de aplicaciones TextBridge e inscribir las aplicaciones

El ‘Servidor de aplicaciones TextBridge’ se instala durante la instalación de TextBridge como se describe en el capítulo 2 de esta guía.

Al finalizar la instalación, el programa crea un grupo de programas TextBridge OCR que incluye el elemento ‘Servidor de aplicaciones TextBridge’ (Figura 4–14).



**Figura 4–14.** Elemento ‘Servidor de aplicaciones TextBridge’.

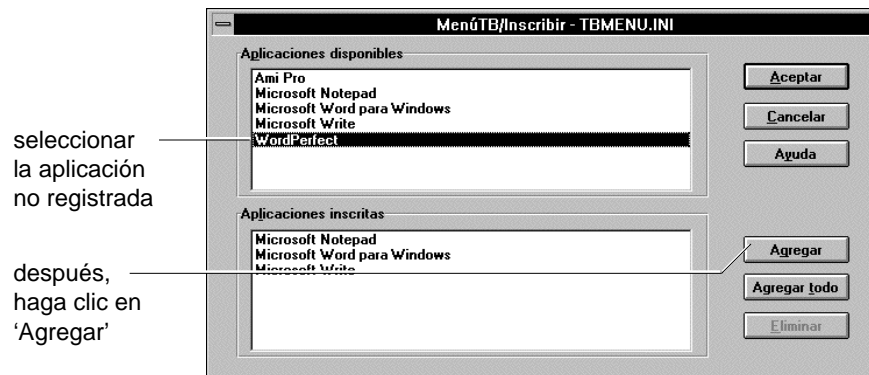
Antes de poder ejecutar el ‘Servidor de aplicaciones TextBridge’ desde una aplicación, es necesario iniciar el programa e inscribir la aplicación.

- 1. Haga doble clic en el elemento ‘Servidor de aplicaciones TextBridge’ del grupo TextBridge OCR.**

El programa se inicializa y aparece como un icono minimizado en el escritorio Windows.

- 2. Haga doble clic en el icono del escritorio Windows.**

Aparece el cuadro para inscribir el ‘Servidor de aplicaciones TextBridge’ (Figura 4–15).



**Figura 4–15.** Cuadro de inscripción del servidor.

### 3. Inscriba la aplicación de ser necesario.

- En la parte superior del cuadro de inscripción, señale la aplicación que desea registrar.
- Haga clic en el botón ‘Agregar’ para agregar la aplicación a la lista de ‘Aplicaciones inscritas’ en la parte inferior de la pantalla.
- Al finalizar la inscripción de las aplicaciones, haga clic en ‘Aceptar’.

Ahora ya puede usar TextBridge OCR desde la aplicación inscrita.

### Ejecutar el Servidor de aplicaciones TextBridge desde la aplicación

En el menú ‘Archivo’ de cualquier aplicación inscrita activa, el comando ‘TextBridge OCR’ aparece como el último comando antes del comando ‘Salir’.

---

**Nota** Para que el ‘Servidor de aplicaciones TextBridge’ funcione, la aplicación principal debe tener un menú ‘Archivo’ y aquí un comando ‘Salir’. La mayoría de aplicaciones Windows usan este estándar.

---

Como ejemplo, en la Figura 4–16 se muestra el comando ‘TextBridge OCR’ en el menú ‘Archivo’ de Microsoft Write®.



**Figura 4–16.** El comando ‘TextBridge OCR’.

Para ejecutar ‘Servidor de aplicaciones TextBridge’ e importar el texto reconocido directamente en el documento abierto de la aplicación principal, siga los pasos descritos a continuación.

- 1. Arrancar el ‘Servidor de aplicaciones TextBridge’.**

Haga doble clic en el elemento de programa en el grupo ‘TextBridge OCR’ (remítase a la Figura 4–14).

- + Para arrancar el ‘Servidor de aplicaciones TextBridge’ automáticamente al arrancar Windows, colóquelo en el grupo ‘Inicio’.

- 2. Asegúrese de que la aplicación principal esté inscrita.**

Remítase al procedimiento de la sección anterior “Arrancar el Servidor de aplicaciones TextBridge e inscribir las aplicaciones”.



### 3. Arrancar la aplicación.

Abra un nuevo documento o uno ya existente desde la aplicación principal a la cual desea importar el texto reconocido.

### 4. Abra el menú ‘Archivo’ de la aplicación principal y seleccione el comando ‘TextBridge OCR’.

Aparecen los mensajes de estado:

Los servicios TextBridge se están conectando...

Los servicios TextBridge se han conectado.

Aparece el cuadro principal TextBridge.

### 5. Configure y arranque OCR desde el cuadro principal.

En el cuadro principal, elija ‘Archivo’ o ‘Escáner’ como fuente de entrada, especifique las ‘Preferencias’ y continúe como si estuviera usando TextBridge como una aplicación independiente y guarda el texto reconocido en una aplicación de tratamiento de textos o cualquier otros formato de texto.

- + Las funciones de ‘Muestra’, ‘Verificador’ y ‘Guardar imágenes de página’ **no** están disponibles en la versión de ‘Servidor de aplicaciones TextBridge’. Si necesita estas funciones, ejecute TextBridge como una aplicación individual y guarde el texto reconocido en la aplicación de tratamiento de texto que desee.

Remítase al capítulo 3 para más información sobre cómo usar TextBridge; remítase a las secciones anteriores de este capítulo para sugerencias y técnicas.

Al completarse el reconocimiento, el ‘Servidor de aplicaciones TextBridge’ se cierra y el texto reconocido aparece en la posición del cursor en el documento abierto de la aplicación seleccionada listo para ser editado.

- + El ‘Servidor de aplicaciones TextBridge’ usa el portapaples Windows para cortar y pegar el texto reconocido a la aplicación en formato **RTF** (siglas de Rich Text Format) en texto **ASCII**. Si la aplicación respalda el formato RTF pegado del portapapeles, entonces se utiliza RTF. De lo contrario, el texto reconocido es pegado como texto y se pierde el formato (negrita, itálica, etc.).

## La impresora OCR

En los otros formatos, TextBridge OCR para Windows puede reconocer los archivos de imagen sólo si están almacenados en formato TIFF. Ciertas aplicaciones y en particular algunos programas de fax, almacenan las imágenes de página de fax sólo en PCX, DCX o en otro formato de propiedad.

Para ejecutar el OCR en imágenes que no son TIFF, puede usarse la **impresora TextBridge OCR**. Ésta está diseñada para que aparezca en las aplicaciones Windows como cualquier otra impresora. Permite “imprimir” una imagen desde una aplicación Windows y producir un archivo de texto reconocido y formateado.

Uno de los usos más comunes de la impresora OCR es reconocer una imagen de página directamente desde una aplicación de fax o de imagen. La impresora OCR es similar al modelo de los programas fax que usan una función similar para enviar fax. Esto es, la imagen de fax es “impresa” al módem fax y enviada a otro módem fax o a otra máquina de fax.

La impresora OCR ofrece la ventaja de poder reconocer virtualmente cualquier formato de imagen (DCX, PCX, Corel, TIFF, etc.). Cualquier programa diseñado para gestionar imágenes puede usar la impresora OCR.

Para poder usar la impresora OCR, remítase a la siguiente sección “Agregar la impresora OCR”. Después remítase a “Usar la impresora OCR en una aplicación de imagen” para más instrucciones sobre cómo enviar una imagen y produce un archivo de texto.

## Agregar la impresora OCR

Los archivos de programa de la impresora OCR son instalados cuando se instala la aplicación principal TextBridge.

Sin embargo, como con los controladores de impresora actuales, es necesario agregar la impresora OCR a la lista de impresoras disponibles a las aplicaciones Windows de la PC.

En este procedimiento se supone que ya se ha instalado TextBridge como se describe en el capítulo 2.

- 1. Desde el ‘Administrador de programas’ Windows, abra el grupo de programas principal y haga doble clic en el icono del ‘Panel de control’:**

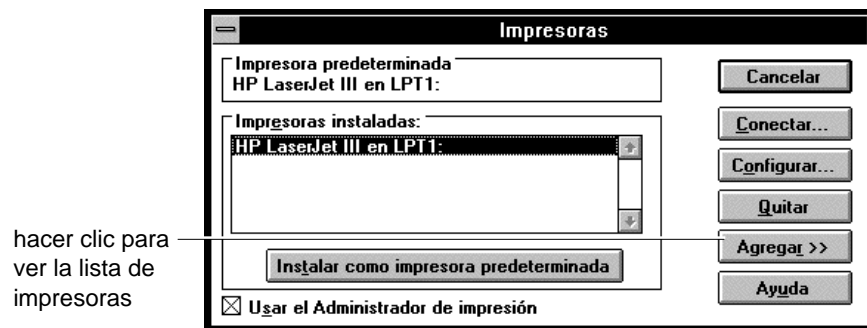


Se abre la ventana ‘Panel de control’ con los iconos de varios elementos del sistema.

2. Haga doble clic en el icono 'Impresoras' en la ventana 'Panel de control'.



Esto abre el cuadro 'Impresoras' de Windows (Figura 4-17).



**Figura 4-17.** Cuadro 'Impresoras'.

3. Haga clic en el botón 'Agregar' del cuadro 'Impresoras'.

El cuadro se expande y muestra la lista de impresoras que pueden agregarse.

4. Señale el elemento siguiente de la lista y después haga clic en el botón 'Instalar'.

Instalar impresora actualizada o no listada

Esta acción muestra el cuadro 'Instalar controladores' en el cual se le indica que especifique la unidad y directorio del controlador de impresión.

5. En el cuadro 'Instalar controladores', introduzca la ruta de acceso al directorio TextBridge BIN:

c:\txbridge\bin

6. Haga clic en 'Aceptar' (o presione Intro).

Esto visualiza el cuadro 'Agregar impresora actualizada o no listada' (Figura 4-18).



**Figura 4-18.** Cuadro 'Agregar impresora actualizada o no listada'.

7. Seleccione la entrada de impresora OCR y haga clic en 'Aceptar'.

El cuadro 'Agregar impresora actualizada o no listada' se cierra y el cuadro 'Impresoras' permanece abierto.

8. Haga clic en 'Cerrar' en el cuadro 'Impresoras' para finalizar el proceso de agregar.

Ahora puede usar la impresora OCR de TextBridge como se describe en la próxima sección.

## Usar la impresora OCR en una aplicación de imagen

La impresora OCR puede utilizarse con cualquier aplicación Windows que pueda abrir y visualizar un archivo de imagen.

Por ejemplo, WinFax Pro (de Delrina Technology Inc.) proporciona un visualizador de imagen que permite ver, manipular e imprimir imágenes de fax. Con el comando 'Imprimir' del menú 'Archivo' de WinFax Image Viewer, puede especificarse y usarse la impresora OCR para que efectúe un reconocimiento en la imagen de fax.

La impresora OCR permite acceder a la función 'Preferencias' antes de empezar el reconocimiento. Después de finalizar el reconocimiento, puede especificar el nombre del archivo de salida, la ubicación y el formato en el cuadro 'Guardar como' de TextBridge.

Para usar la impresora OCR en la aplicación, siga los pasos siguientes:

### **1. Abra la aplicación de imagen y visualice la imagen a ser reconocida.**

- + La imagen debe ser binaria (en blanco y negro) y dentro del rango de resoluciones respaldadas por TextBridge. TextBridge puede reconocer imágenes de 100 por 200, 200 por 100, 200 por 200, 300 por 300 y 400 por 400 puntos por pulgada.

### **2. En el cuadro 'Especificar impresora' de la aplicación especifique la 'Impresora TextBridge OCR' como impresora de destino.**

Es necesario haber agregado la impresora OCR en el programa 'Impresora' del 'Panel de control' de Windows, como se describe en la sección anterior "Agregar la impresora OCR".

### 3. También puede definir las preferencias TextBridge.

- Haga clic en el botón ‘Opciones’ o ‘Configurar’ del cuadro para especificar la impresora de la aplicación de imagen. Se muestra un cuadro de instalación secundario.
- Haga clic en el botón ‘Preferencias’ en el cuadro de instalación para visualizar el cuadro ‘Preferencias’.
- Especifique la calidad del documento estándar o de fax, orientación de página, segmentación de página automática, etc.
- Al acabar, haga clic en ‘Aceptar’ en el cuadro ‘Preferencias’ y salga de los otros cuadros.
  - + En el cuadro para especificar la impresora de la aplicación, si hay una opción para usar la resolución actual de impresora, **actívela**.

### 4. Arranque el proceso OCR en la imagen visualizada.

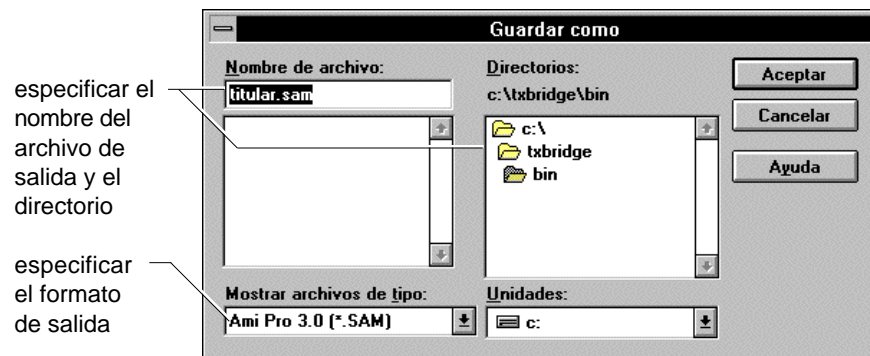
- Abra el menú ‘Archivo’ desde la aplicación y seleccione ‘Imprimir’.
- Haga clic en el cuadro ‘Imprimir’. Aparecen mensajes de procesamiento a medida que se efectúa el reconocimiento de caracteres:

Procesando...

Obteniendo imagen...

Reconociendo texto...

Al finalizar el reconocimiento, aparece el cuadro ‘Guardar como’ de TextBridge (Figura 4–19).



**Figura 4–19.** Cuadro ‘Guardar como’.

- 5. Especifique el nombre del archivo de salida, formato, unidad y directorio de destino, después haga clic en ‘Aceptar’.**

El archivo de texto de reconocido es convertido al formato especificado y escrito en el disco duro.