Scanning Issues

Preparing and Converting the Text

The first step in creating your topics is to prepare the text for use by Word for Windows. Original text can be scanned or typed, or it might already exist in electronic format containing embedded attribute codes (such as typesetting or text-editor codes). Prepare your text based on the following guidelines:

- New text should be entered with Microsoft Word for Windows.
- Printed text must be retyped or converted to ASCII using an optical character recognition (OCR) scanner, then formatted and saved using Word for Windows.
- Electronic files must be converted to a format readable by Word for Windows. All text files must be saved in RTF format from Word.
- Text for a Viewer title can include most types of Word for Windows character and paragraph formatting, including tables. Table borders and titles are not supported. To preserve the alignment of table columns, you might want to format tables so they won't wrap if the user resizes the Viewer window.

To create a nonwrapping table:

1. Select the paragraph containing the table.
2. From the Format menu, choose Paragraph.
   The Paragraph dialog box appears.
3. Select the Together check box.
4. Choose OK.

You can also use the same method to format any type of text you don't want to wrap.

Purchasing Images

Digital image libraries are becoming quite popular. The quality, variety, and resolution of computer-generated and digitized or scanned images are going up as fast as the price is going down.

Note :
Think twice before digitizing images from magazines, books or television—most of them are copyrighted. If you find some images that you can't do without, ask permission to use them or you will be violating copyright law. The best places to start are the public relations or marketing departments of the companies claiming copyright ownership. Plan
ahead as it can take several months to obtain permission.

Scanning

Many bitmaps used by a multimedia application will come from photographs. A photograph uses continuous tones and shades—colors that blend smoothly from one to another. By using a scanner or special digitizing equipment, you can transform photographs into bitmap images.

Scanning is the most common way to rapidly create full-color electronic images from photographic prints, slides, or pieces of flat artwork. The main drawback of scanners is that the scanning process is time-consuming. Scanning a large image at high resolution can take up to a minute or more of processing time. This adds up when you have a few thousand images to digitize. If a speedy production cycle matters more than having the highest quality, consider buying a good video camera and a frame grabber.

Image Processing Software

After you've scanned or digitized an image, you may have to do additional processing to prepare it for use in your product. This is where image processing software and paint tools come in.

There is often a fine line of distinction between digitizing software, image enhancement software, and paint software. Some products have all three capabilities built in, while others specialize in one capability or another.

Digitizing Software
At its simplest level, digitizing software controls the scanner or digitizer you are using to capture an image. It may allow you to set the image size, select the portion of the image to digitize, specify the resolution and number of colors, and select the format in which to save the image file. Most of the high-end digitizing

software also includes image enhancement and painting features.

You can probably get by just fine using simple scanning software and the MDK's BitEdit and PalEdit tools if you don't need to do extensive editing to enhance your images.

Image Enhancement Software
Image enhancement software is designed to convert images to different formats, spatial resolutions, and color resolutions; to modify saturation, hue, tint, contrast, and brightness; to sharpen or blur edges; to modify palette assignments; and to flip, rotate, crop, and resize. Think of the image enhancement software as a digital photo-retouching product. You normally use it to make global changes to an entire image, such as blurring the background, changing all blues to greens, and so on.

One function that you will probably use on every image you scan or digitize is color reduction. When you scan or digitize a natural image at high color resolution, you capture millions of colors. Most of these colors, however, are simply subtle shades of a relatively few colors.

If you want to display the image on a computer screen that supports only 16 or 256 colors, then you will have to merge or delete most of the captured colors. Scanning and image enhancement software provide a variety of ways to do this algorithmically and manually.

BitEdit and PalEdit from the Multimedia Development Kit provide simple edit functions for images and image palettes. See the MDK's Data Preparation Tools User's Guide for descriptions of these tools.

Paint Software
Paint software is used to actually edit the contents of an image. Paint software is sometimes used in the scanning process to add special effects to an image or to work on color gradations and hues at a pixel level. Use paint software to make minor changes to images. You can cut, copy, and paste segments of the image, or use tools such as a paintbrush or airbrush to add elements to the image.

File Conversion Software
The images used in your multimedia application can come from a variety of sources and computer platforms. Although your application may be designed to import any graphic format, it will run most efficiently if it imports graphics in Windows DIB format. The MDK's BitEdit and Convert utilities let you convert from a number of the most common graphicformats to DIB and back.

The following table shows the different formats supported by BitEdit and Convert:

Format	Extension	Capability
Apple Macintosh PICT	.PIC	Read, Write
AutoCAD Import	.PLT	Read
CompuServe GIF	.GIF	Read
Computer Graphics Metafile	.CGM	Read
Encapsulated PostScript	.EPS	Read
HP Graphic Language	.HGL	Read
Lotus 1-2-3 Graphics	.PIC	Read
Micrografx Designer/Draw	.DRW	Read
Microsoft RIFF DIB	.RDI	Read, Write
Microsoft RLE DIB	.DIB	Read, Write
Microsoft RLE RIFF DIB	.RDI	Read, Write
Microsoft Windows BMP	.BMP	Read
Microsoft Windows DIB	.DIB	Read, Write
Microsoft Windows Metafile	.WMF	Read
PC Paintbrush	.PCX	Read, Write
Tagged Image File Format (TIFF)	.TIF	Read
Truevision TGA	.TGA	Read, Write

If your images are stored in other formats, you may need to use a two-step approach, first converting them to one of these standard formats and then to the final format.

Scanners

Anytime you want to turn a photographic print, slide or flat artwork into a digital image, use a full-color scanner and scanning software. The scanner builds a digital representation of the photograph and creates a corresponding image file. Scanning software can typically store images in PICT (Apple Macintosh) or PCX (PC Paintbrush) formats. Scanners can produce a far higher image resolution than most cameras. The best scanners digitize at least 300 dots per inch resolution with various color depths per pixel (1 to 24 bits per pixel).

Scanners come in two basic varieties: flat-bed scanners and slide scanners. Flat-bed scanners are used to scan printed materials and photographic prints. Slide scanners are used to scan photographic slides. Use a flat-bed scanner if the bulk of your images are flat art or prints. If most of your images are slides, use a slide scanner for quality reproductions.

Many different types of scanner hardware and software are available. If possible, use a scanner that can capture at 24 bits per pixel and at least 300 dots per inch resolution. You'll probably want to edit and archive your scanned images in the

original 24 bit format, and then reduce the images to either 8-bit or 4-bit for the actual application.
This lets you work with the highest quality image until satisfied it's ready for conversion to the lower resolution.

Capturing and Preparing Images

Although scanning and digitization are both relatively straightforward processes, the quality of your final images and the amount of work required to produce them can vary drastically, depending on how well you plan the project. The next few pages describe the following aspects of the image preparation process:

Choosing images with the proper characteristics
Capturing images with a scanner
Capturing images with a video camera Enhancing captured images

Capturing with a Scanner

The best way to describe the scanning process is to explain the steps involved in converting a photographic print to a Windows DIB (Device Independent Bitmap). The DIB format is the standard format for all Windows bitmaps—it is the recommended target format for all your bitmapped images.

Each image starts out as a separate photographic print with its own unique palette. The process of getting an image ready for the final application goes something like this:

Adjust Your Monitor
Your images may look great on your display, but if your production monitor is improperly adjusted the final image may not look right when displayed on other monitors. Since you can never be sure how well adjusted the delivery system monitor will be, make sure you have at least created the image in the truest, most accurate color.

The best way to adjust a production monitor is to buy a color bar generator that outputs pure RGB and plug it directly into your monitor. You can also look for software that generates a color bar, and then adjust your monitor settings accordingly. Without these tools, adjustments are purely subjective as they rely solely on your ability to visually evaluate color.

Choose the Image Depth
Many scanners offer several image-depth settings. Because Windows with Multimedia supports 1-bit, 4-bit, and 8-bit bitmaps, any palette can contain up to

256 unique colors. The quality of a scanned bitmap depends on how well a system can re-create the effect of a continuous-tone image using these 256 colors. Whenever possible you should always create and enhance bitmaps using large image depths.

Good color scanners can scan with a color resolution of up to 24 bits per pixel, allowing 16 million colors in the palette. Most scanning programs let you reduce the colors in the image from 24 to 8 bits or less. If you want to scan an image once and store the highest quality image possible, scan it at 24 bits and reduce it after you're totally sa
tisfied with the results. If you don't need the original high-quality image, you might as well reduce it as you scan, since it will require far less space to store.

Adjust Image To Proper Size

Set the size or resolution of the screen image by specifying the number of dots per inch (DPI) in the image. Adjust the DPI setting (sometimes called the scanning resolution) to be as close to the desired image size as possible. VGA screens with 640 by 480 resolution display images at about 72 DPI; a scan setting of 72 DPI produces approximately a 1:1 ratio in size.

You control the size of the screen image with the scanning resolution. To shrink the screen image of an illustration, set the scanning resolution to less than 72 DPI. To enlarge the screen image, set the scanning resolution to greater than 72 DPI.

A quick way to determine the desired scanning resolution is to divide the number of pixels you want to cover on the screen in one dimension by the same dimension of the area of the picture you want to use. For example, say your original image is 10-by-8 inches and you want it to fill half the screen (320 by 240). Divide 320 by 10 and you obtain a scanning resolution of 32 DPI.

Hint:

Although paint programs allow you to resize images, building a digitized image of the correct dimensions with the digitizing software gives you a better image. Paint programs make intelligent guesses when reducing or enlarging an image. In contrast, digitizing software doesn't guess; it uses information from the original illustration to build the
digitized image.

Identify Cropping Boundaries

Always pre-scan the image. Pre-scanning takes only a few seconds and provides a quick, low-resolution scan of the entire scanning bed. Pre-scanning your image lets you set cropping boundaries for the digitizing software and saves time during scanning. Pre-scanning is present in all decent capturing programs.

After pre-scanning, eliminate the portion of the picture you don't want. This not only limits the size of the image file, it also reduces the total number of colors

included in its color palette. Since you'll probably have to adjust the image's palette if you want to display it with other images, a smaller palette can simplify this process.

Scan the Image

Now scan in the image. After you've scanned the image, look at it and see what adjustments you might like to make. Scanner software sometimes includes a paint or draw package to clean up any problems introduced during the scan. If you have the time, cut out all unnecessary elements of each image, especially in the background. Again, this makes for a smaller image file and a smaller palette.

Transfer the Image

If you capture digital images on a different computer than your multimedia application development system, you'll need to transfer your images to the development platform. There are several ways to move images from one system to another, but using a network is probably the fastest and most efficient method if you are moving lots of large files, such as images.

Video Frame Grabbing

You can use a video camera hooked to a digitizing board in your computer to capture images. The digitizing board, often called a frame-grabber, converts the analog signal from the camera to a digital format that can be read and enhanced by software the same as a scanned image.

The biggest difference between a scanner and a digitizer is that a scanner can only capture an image from a two-dimensional source, such as a photograph or slide, while a digitizer can capture any video image. A digitizing board usually captures two-dimensional images faster than a scanner, but it doesn't necessarily provide the best quality for the money. So, unless you're willing to get the best equipment, you might be better off buying a scanner to capture two-dimensional images.

Scanning Text

The concept of feeding printed material into a machine that recognizes each letter and feeds it into a text file sounds ideal. Fortunately, such technology exists and is called optical character recognition (OCR). Under the proper circumstances it can be a very efficient way to get text into a computer.

An OCR system consists of a scanner (quite possibly the same one used to scan images), a computer, and some software. The scanner converts a page of text into a bitmapped image, and the software analyzes the letter shapes and converts them into ASCII letters. The number of predefined typefaces is usually limited to less than a dozen, although many systems have a learning facility to include new characters and typefaces.

You scan every page and then you run various utility programs (such as a spell checker) to detect misreads and other scanner errors. For the final step, print and proofread the text. The following table lists some of the benefits and drawbacks associated with using OCR technology.

Benefits	Drawbacks
Scanning requires little upfront labor.	Some scanners can read only a limited set of typefaces. Other scanners read more typefaces, but you first have to train them by running samples through the scanner and then calibrating its interpretation of the text.
An OCR scanner can quickly convert large amounts of printed information into electronic files (less than a minute per page).	Assuming an accuracy rate of 99%, scanned text contains an average of one error for every two lines of text. This error rate can mean hundreds of thousands of errors for long texts. Make sure to schedule time for editing and proofreading.
Scanning usually costs less than re-keying.	You can lose special symbols (such as Greek or other foreign characters) and complex formatting (such as tables, mathematical formulas, or special fonts).

Retyping Text

Although typing is labor-intensive, it is often the most economical method to get large amounts of printed material into a computer. You can either have your staff re-key the text or you can contract with a service bureau. Typists enter the text directly into the system from the printed material.

There are a number of companies offering such a service. These companies generally claim 99.9% or higher accuracy through the use of double or triple key verification (also called double or triple blind typing). This means that they have two or three people type the same material, and then use a computer program to spot differences. The assumption is that several people won't make the same mistake at the same place in the file.

The cost of this service typically varies with the volume and complexity of the original material and the accuracy, turn-around speed, and extra services (such as format tagging) provided by the service company. The following table lists some of the benefits and drawbacks associated with re-keying the text.

Benefits	Drawbacks
Re-keying printed documents is an established type of service, and you can reliably estimate the costs and time associated with such a project.	
	Re-keying usually takes longer than scanning.
The typists can include formatting and structural information into the text as they re-key it. This can reduce the time necessary to prepare the text for the retrieval software.	
	A labor intensive job such as this can become costly. (However, the cost of verifying scanned data is also high.) You still must schedule time for proofreading and correction.

End.