

Scanned Drawing Cleanup with Autodesk Raster Design

Extend the value of scanned engineering drawings and maps, aerial photos, and satellite imagery, with Autodesk® Raster Design software.

Effective use of such information translates into dollars saved. With tools such as Autodesk® Raster Design, CAD (computer-aided design) and GIS (geographic information system) departments can tap into this wealth of information to get the most from existing raster assets and enhance the impact of visualizations, presentations, and maps. Raster Design complements AutoCAD® software-based CAD and GIS solutions with powerful raster data preparation, translation, editing, and analysis.

This paper explains the planning considerations involved in the use of scanned drawings. It covers various revision strategies, emphasizing the pros and cons of each. It details the Autodesk Raster Design editing and conversion features. Finally, it compares Raster Design revision and conversion methods.

Contents

The Hard-Copy Dilemma	3
Advantage of Scanning	3
Scanning Considerations	3
Drawing Cleanup with Autodesk Raster Design	6
Drawing Management Needs	6
Raster and Vector Data	7
Making Changes to Documents	7
Raster Changes to Documents Using Autodesk Raster Design	8
Hybrid Edits	8
Raster-to-Vector Conversion	9
Conversion Methods	9
Interactive Conversion using Autodesk Raster Design	10
Color and Grayscale Hybrids	11
Conclusion	12

The Hard-Copy Dilemma

Existing paper drawings preserve a lot of engineering and design expertise. They may contain the only record of historical or current conditions. Ready access to these drawings may be required for ISO 9000 or OSHA compliance. Even companies that make extensive use of CAD may continue to rely on paper drawings to meet these requirements.

However, managing years of accumulated paper drawings presents a huge operational challenge. Additional problems include the following:

- Paper drawings are easy to lose or misfile.
- Paper drawings are difficult to keep current, and an out-of-date drawing may be used accidentally.
- Paper drawings are expensive to correct or update.
- Paper drawings deteriorate with time.
- Paper drawings are costly to distribute.

Advantage of Scanning

Scanning an archive of paper drawings reduces the time it takes to locate drawings. A quick electronic search replaces a costly manual one. Scanned drawings can be safely stored in an electronic archive. Since the “original” is no longer pulled to make a copy, misfiling becomes a thing of the past. Improved cataloging and an electronic archive greatly reduce the chances of accidentally using an outdated print. And most important, scanning prevents the further deterioration of hard-copy drawings as they are handled and filed. Every copy and print of a drawing from an electronic archive is the same as the first.

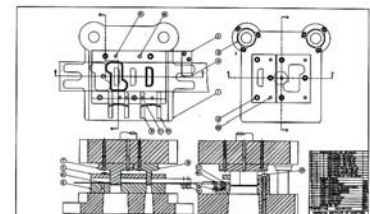
Scanning Considerations

Before scanning an archive of drawings, organizations need to consider several questions about the types of drawings in the collection and the amount of detail needed in the scanned output.

- *What types of data do I have?* The three basic data types, shown at right, are bitonal, grayscale, and color. Each data type has its own considerations for scanning density and format.
- *What scanning density should I use?* Bitonal drawings are typically scanned from 200 to 400 dots per inch (DPI). The best resolution to use depends on the amount of detail in the drawing versus its size. The narrowest lines in the drawing should be represented by raster a couple of pixels wide. For consistency, the organization should consider establishing a uniform DPI to use for all bitonal drawings. The resolution for color and grayscale images should be sufficient to capture desired features in the photo or to optimize DPI for output. Color and grayscale images can typically be scanned from 75 to 600 DPI.
- *What format should I use?* For bitonal scans, use a popular and well-supported format such as TIFF or CALS/MIL. These images respond well to compression so a compressed format is recommended. RLC and CCITT Group 4 compression



Grayscale Data



Bitonal Data



Color Data

are good alternatives, with Group 4 achieving the highest degree of compression. Grayscale images are normally scanned into 8-bit TIFF format. Color images may be scanned into TIFF or JPEG format, depending on how they will be used. JPEG is a highly compressed format that alters data as it compresses. It is best for high-color images (24 bit) where some alteration of the original data can be tolerated.

How Much Storage Space Do Images Need?

The amount of space required for a bitonal image depends on its original size and the type of compression used. A raw bitmap file is uncompressed, while RLC is a line-by-line compression. Group 4 compresses 2D areas. The following table shows typical file sizes for a 300 DPI scan with different combinations of drawing size and compression technique.

***Table 1: Bitonal Data and Compression Results**

Drawing Size	Raw Bitmap	RLC	CCITT G4
A (A4)	1 MB	200 K	40 K
B (A3)	2 MB	400 K	75 K
C (A2)	4 MB	820 K	150 K
D (A3)	8 MB	1.6 MB	300 K
E (A0)	16 MB	3.2 MB	580 K

Clearly, bitonal scans respond well to compression. Note the large differences between RLC and Group 4 compression results. These algorithms are very efficient with bitonal data because a typical line drawing contains large areas of blank space.

Given the efficiency of scanning into compressed formats and the low cost of storage media, creating a digital archive is desirable. The following table shows the approximate number of drawings (300 DPI, Group 4 compressed) that can fit on a single CD (650 MB), compared to the original drawing size.

***Table 2: Approximate Drawing Storage per CD**

Drawing Size	A (A4)	B (A3)	C (A2)	D (A3)	E (A0)
Drawings per CD	16,000	8,000	4,000	2,000	1,000

Cost of Scanning

Many factors affect the cost of scanning an archive besides the actual scanning step itself. How much processing or cleanup is performed as the drawings are scanned? Are images indexed as they are scanned? Is a database being created in the process? What are the costs of transporting drawings? What are the operational costs to an organization if frequently used drawings are sent away for scanning?

SCANNED DRAWING CLEANUP WITH AUTODESK RASTER DESIGN

In general, these added costs represent a small fraction of the basic scanning costs. Typically, a scanning bureau charges a few dollars per document. The following table shows typical prices quoted by scanning bureaus.

***Table 3: Approximate Cost of Scanning Drawings at 200 DPI in Quantities of 100 (2006)**

Drawing Size	A (A4)	B (A3)	C (A2)	D (A3)	E (A0)
Cost Per Scanned Drawing	\$2	\$3	\$4	\$5	\$7

The DPI setting for scanned documents can significantly affect the cost. The dollar amounts showing in the preceding table, for example, could double for 300 DPI scans.

How to Scan

The first step is to determine how many drawings are in the archive. An archive of several hundred to a few thousand drawings has different requirements than an archive of hundreds of thousands of documents.

Size of the archive often dictates whether scanning is performed in house or at a scanning bureau. The need to scan a large archive may justify the cost of a large-format scanner and personnel to operate it. Other considerations may also come into play:

- Are the drawings of such crucial nature that they should not be allowed outside the facility?
- Are the drawings heavily used, requiring access during the conversion process?

These considerations may require the purchase of an in-house scanner even when it may not be justified on volume alone.

Regardless of where scanning is performed, the process should start with a pilot project.

- The first phase of the pilot is to create sample scans of five to ten drawings. Choose drawings of varying quality and size. Use this sample phase to experiment with file formats, scanning resolution, and file compression. Do all the applications that use the drawings work with your chosen file formats and compression? Are the drawings readable at the chosen scanning resolution? Can your archiving system handle files of this size?
- The second phase is to execute a small production run of 5 percent or less of your archive. Here is where organizations can work out the logistics of transporting drawings, staff time, throughput, cataloging, and populating the archive. Also you get an idea of process exceptions such as the percentage of improper scans, missing scans, and drawings in poor condition that can't be scanned.

Armed with the knowledge gained from the pilot project, organizations can begin scanning the entire archive. You want to scan your archive only once. Quality is everything!

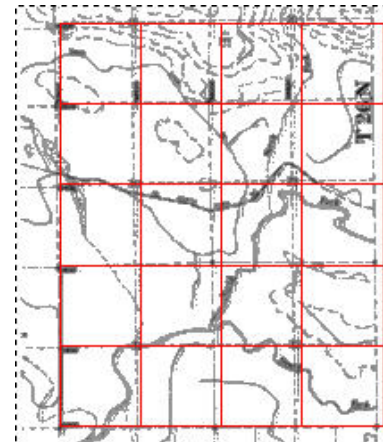
Drawing Cleanup with Autodesk Raster Design

Drawings do not typically emerge from the scanning process ready to use. Depending on the sophistication of its hardware/software and services, the scanning bureau may perform some cleanup before delivery of the scanned files. In other cases the user must do the cleanup. Autodesk® Raster Design software provides several powerful drawing cleanup tools.

Deskew enables you to align points in the drawing (usually a border) to a specific angle, making it exactly horizontal or vertical, for example. Use this capability to fix drawings that have a slight tilt or skew as a result of being fed less than perfectly through the scanner.

Invert reverses the foreground and background tones in a drawing, so you can see black lines on a white background or white lines on a black background, whichever best suits your requirements. You can also use the Invert command with the image color property to apply colors other than black and white to the foreground, background, or both.

Rubbersheet is used to warp or stretch drawings in a precisely controlled manner to correct distortions that can occur in the scanning process. For example, circles may look more like ellipses or the drawing may no longer be orthogonal. The drawing on the right shows grid rubbersheeting, which uses a regular grid of control points to correct the shape of a drawing.



Grid Rubbersheeting

Despeckle can clean up dirty or faded drawings that show spots or speckles after scanning. This command removes all spots smaller than a specified size from the drawing.

Touchup tools enable you to make precise corrections to lines, arcs, or even individual pixels. The drawing on the right shows part of the Touchup tool palette and an example of a diagonal tool that you can use to repair lines in a drawing.



Touchup Tool

Filters can quickly clean up blurred bitonal linework by making lines thinner, thicker, or smoother in a controlled, iterative process. You can also use filters with grayscale images to permanently improve the resolution of details and image contrast.

Raster Entity Manipulation (REM) enables you to select individual raster entities (lines, arcs, or circles) and smooth their linework.

Rub commands can selectively remove stains, tear marks, wrinkles, and tape shadows from scanned drawings.

Crop is useful when you need to cut out unwanted areas of a drawing. This can happen if your scanner has captured a large sheet of paper, and the scan extends beyond the border of the drawing.

Drawing Management Needs

Once drawings have been scanned into raster format, the manual process for storage, retrieval, and reproduction of paper drawings no longer applies. A drawing management system should be reliable and easy to use, while also preserving data integrity. Other, less obvious requirements are as follows:

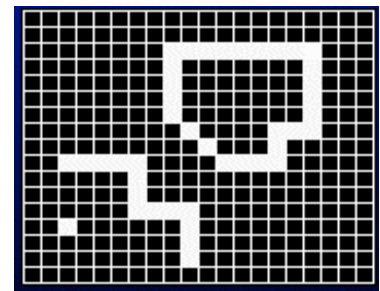
- *Compatibility with raster editor, database, and CAD system.* Without compatibility with other systems that need to assess the drawing archive, the document management system cannot serve its purpose.
- *Customizable by network manager.* Every company and every department within that company have differing needs for the document management system. For control, efficiency, and cost reasons, customizing the management system to meet those needs is best done by in-house personnel.
- *File locking.* Document management system should not allow changes while drawings are checked out.
- *Device independence.* Network servers are not necessarily using the same operating system as workstations attached to it. The document management system must be able to handle different operating systems and hardware seamlessly.
- *Ability to view in native format.* The user can be assured of accuracy only if documents can be viewed in their native format. Translations can result in data loss or misinterpretation.

Raster and Vector Data

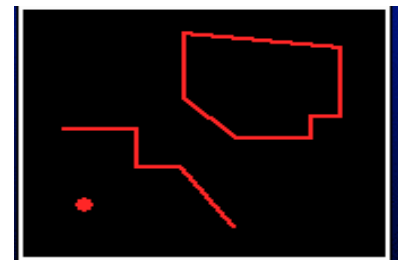
It's important to differentiate between the two basic ways drawings can be represented electronically. Raster format is the result of the scanning process, while vector format is the form constructed in a CAD drawing.

In a raster image, individual pixels create a mosaic or bitmap of the original image. Pixels can be either black and white in the case of a bitonal scan, or multiple shades of gray or color in the case of a grayscale or color scan. With hundreds of pixels to the inch, a very fine mosaic is created, quite faithful to the original. This mosaic is created on a grid. Pixels can fit only within the grid and cannot be moved by fractional amounts.

Vector data, on the other hand, is a mathematical description of geometry. For instance, the coordinates of its two endpoints define a line. Since the entire line was created from a single description, any point on the line can be used to "trace back" the entire geometry. Since vector entities are not constrained to a grid, they normally appear much smoother and thinner than the equivalent raster geometry.



Bitonal Data



Vector Data

Making Changes to Documents

Invariably, the scanned documents now stored as raster files require changes and updates. Three choices exist for making these changes:

- Make changes in raster to revise the raster document
- Make changes in vector, preserving unchanged portions in raster, resulting in a hybrid drawing
- Convert the entire document to vector, incorporating changes in the process

Each method has implications for efficiency, accuracy, and cost.

From a cost standpoint, scanning a document is inexpensive, costing only a few dollars per sheet. In contrast, converting or redrawing that document is very expensive, usually costing a couple of days of drafting time, or several hundred dollars. Hybrid or raster changes take a relatively short time, as only relevant parts need to be touched. All other factors being constant, hybrid and raster changes are the most cost effective.

A rough guideline to determine which revision method to use is based on the amount of change involved:

- Changes involving less than 5 percent of the drawing content are performed in raster.
- Changes involving 5 to 40 percent of the drawing are done as hybrid edits.
- Changes involving more than 40 percent of the drawing are done in vector format.

Raster Changes to Documents Using Autodesk Raster Design

Autodesk Raster Design has several tools for performing raster edits on scanned drawings within AutoCAD® or AutoCAD-based software. Three basic approaches are available:

- Edit the existing raster
- Create new raster
- Draw in vector format and merge the resulting geometry into the raster image

Raster editing is particularly useful for tasks such as replacing a title block, translating the text in a drawing, cut-and-paste drawing composition, inserting as-built revisions, and doing cut-and-paste operations to other Microsoft® Windows® applications. Touchup tools are used both for erasing unwanted geometry, say making way for updates, and for drawing new raster as part of the updates themselves.

Raster Entity Manipulation (REM) tools enable you to capture raster geometry, and then manipulate it by moving, copying, grip editing, deleting, or transferring to the Windows clipboard. The Vector Merge command creates new raster from vector entities. For instance, you can produce text in AutoCAD software and then merge it as raster into the drawing.

Hybrid Edits

Hybrid edits are often the most efficient means to update drawings. In this process, new portions of the drawing are done in vector and unchanged portions are left in raster. Since only part of the drawing is touched, revisions are fast. An advantage of this technique is that as a drawing is updated over time, it is converted to vector incrementally.

Autodesk Raster Design Rub, REM, Raster Snap, and Mask commands are used to clear the way for vector updates. Upon completion of the edits, saving the DWG™ file saves correlation between the raster and vector components. If an external correlation file is needed, the Autodesk Raster Design Export command allows one to be written.

Raster-to-Vector Conversion

Raster-to-vector conversion of a drawing archive is a resource-intensive undertaking, often requiring large amounts of time and money. Organizations must have a clear understanding of the purpose for converting a large number of raster drawings.

To remain useful, some drawings must be converted to CAD. Examples include drawings to be used for N/C (numerical control) machine tooling, parametric design, process modeling, simulations, and 3D modeling.

In other cases, a paper drawing may have deteriorated too much to be useful in its current state. Years of erasures and modifications can make a drawing almost impossible to scan into bitonal format because of inconsistent background and linework. A drawing like this can still be scanned as a grayscale image to pick up details that would otherwise be lost. Conversion can then proceed on the grayscale scan, using image enhancement and image processing tools as necessary to clarify raster details.

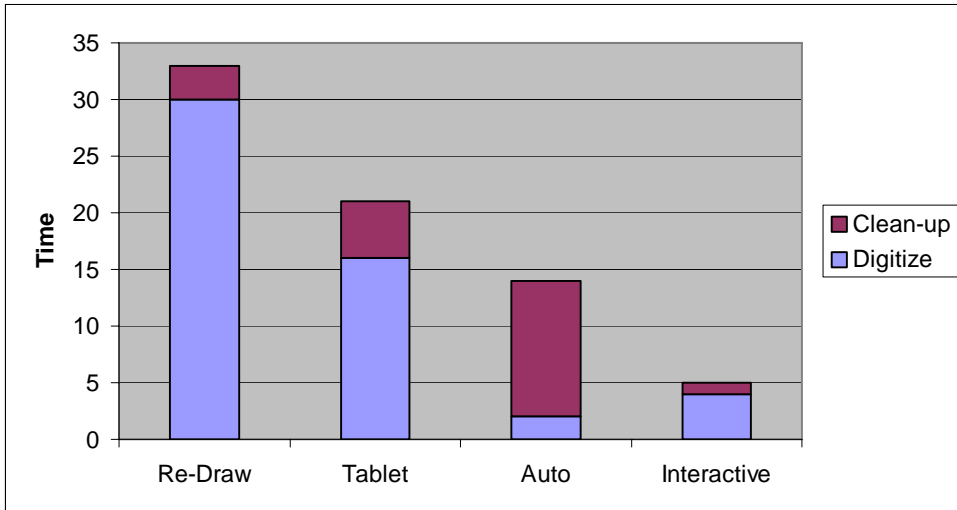
Conversion may also be dictated to satisfy contractual requirements, such as delivering DWG format drawings. Conversion may also be justified because the drawing would be more useful in vector form. Perhaps drafting standards have changed sufficiently that a new CAD drawing would be much more readable, or color vectors can be used to communicate additional information.

Conversion Methods

Several methods are available to perform conversion. Some are basic while others use sophisticated tools.

- *Manual redraw* is the simplest method. Here the user simply redraws the information using CAD tools, referring to the original drawing as necessary. This is perhaps the slowest method of conversion because of the poor ergonomics of continuously switching attention from the paper drawing to the computer screen.
- *Digitizing* by taping the drawing to a digitizing tablet is an improvement over manual methods, as geometry is more efficiently transferred from the drawing to the screen. This method still involves much attention switching from tablet to screen.
- *Automatic conversion* is a tempting alternative but suffers from several shortcomings. First, the result is only as accurate as the original drawing. For mechanical parts, this level of accuracy is usually not sufficient. Second, post-processing cleanup can be time consuming, fixing what the vectorizer failed to recognize. And third, the quality of the resultant drawing may not be sufficient to meet the needs of mechanical applications. For example, several line segments may be used to represent what should be one long line. Circles and arcs may be represented by polylines instead of primitives.
- *Interactive conversion* methods are often the most effective. Here the computer does what it's best at, and the user intervenes when decisions are necessary. Because of this interaction, accuracy and geometry are controlled from the beginning, requiring little cleanup.

*The following graph compares total conversion time (digitizing time plus cleanup time) for each method.



Time Required for Different Conversion Methods

Interactive Conversion using Autodesk Raster Design

Autodesk Raster Design provides several tools to assist with interactive conversion. These tools are effective in achieving quick, accurate, and geometrically correct conversions.

Raster Snap allows accurate selection of points, because the AutoCAD cursor “jumps” to the appropriate point on the raster geometry. Raster Snap also lets the user work at a higher zoom level. As a result, the user gains a better perspective on the drawing and is assured of accurate picks without having to zoom in and out. Productivity is thus enhanced.

Accurate picks notwithstanding, extraction of geometry from raster is only as accurate as the original drawing. The original drawing, in many cases, was never intended to be used in this manner, and dimension information was supplied through annotation instead. Geometry verification allows exact dimensions to be specified based on the user’s interpretation of the drawing as the vector entity is being created. Autodesk Raster Design **VTools** commands include this verification option as part of the command. For example, a raster circle being converted using Raster Snap and the **VCircle** command may extract a radius of .48. The verification step allows a new value, .5 in this example, to override the extracted value.

Text recognition is a semiautomatic means of replacing raster text or tables with vector versions. Use the text recognition tools to obtain vector-based text that you can then standardize, edit, or extract from the drawing to use in word processing, spreadsheet, or database applications.

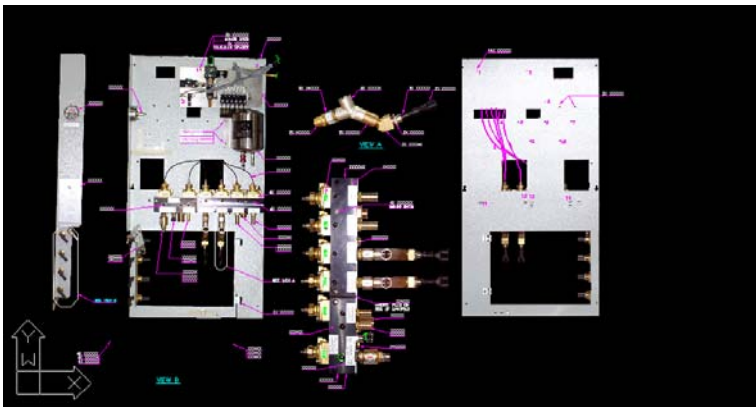
SCANNED DRAWING CLEANUP WITH AUTODESK RASTER DESIGN

As conversion progresses, the old raster geometry becomes less important. A convenient feature of Raster Design is the incorporation of **AutoRub** into **VTools** commands. Here the underlying raster is automatically erased as vector entities are created.

Interactive conversion has proven to be the best method for converting drawings of mechanical parts. Interactive conversion recognizes the inherent inaccuracies of raster geometry and the need to specify geometry precisely in the finished DWG file. **Raster Snap** and **AutoRub** help to make the conversion process more efficient.

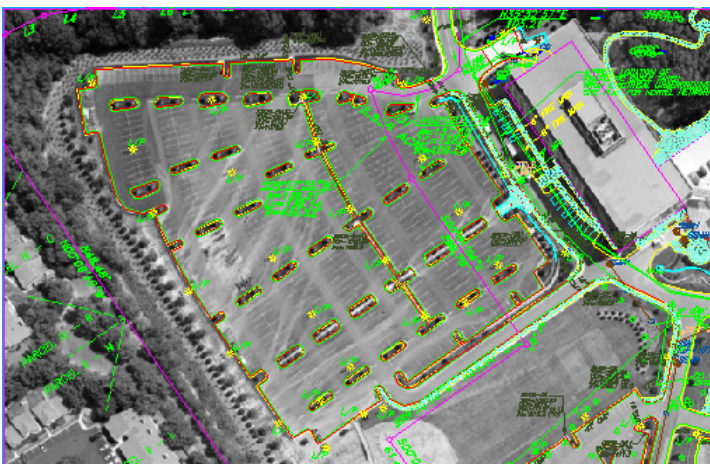
Color and Grayscale Hybrids

Images are often useful as backdrops to CAD drawings, containing a wealth of information that is often unproductive to digitize. In mechanical CAD applications, digital photos of parts and assemblies can be combined with AutoCAD linework to illustrate repair, assembly, or maintenance procedures. The following figure illustrates assembly annotations for a control panel.



Digital Photo Used as Backdrop for Assembly Drawing

In geographic CAD applications, aerial photos provide a visual context for roads, land parcels, and other civil engineering data. The following figure illustrates an excerpt from a facility design with an underlying photo of the existing parking lot, buildings, and trees.



Aerial Photo Showing Context for a Facility Design

Conclusion

Every company that maintains an archive of paper drawings needs to take advantage of scanning. The benefits of a scanned archive far outweigh the costs of creating it.

Your organization will be more productive with a scanned archive. Drawing revisions using the methods described here are much more efficient than manual drafting.

ISO 9000 compliance may depend on an electronic archive to quickly locate and distribute documentation. Without improved access to legacy data, compliance may be difficult to achieve.

With the information presented in this paper your company should be able to confidently move forward with the creation of a scanned archive for your paper drawings. With AutoCAD or AutoCAD-based software and Autodesk Raster Design, you can work efficiently with drawings from that archive as requirements for updates, changes, conversion, and illustration arise.

* The performance results and statistical information reported in this paper were derived from tests and research intended to determine average values. Actual results may vary based on machine, procedures, and source material.

Autodesk, AutoCAD, and DWG are registered trademarks or trademarks of Autodesk, Inc., in the USA and/or other countries. All other brand names, product names, or trademarks belong to their respective holders. Autodesk reserves the right to alter product offerings and specifications at any time without notice, and is not responsible for typographical or graphical errors that may appear in this document.

© 2006 Autodesk, Inc. All rights reserved.