MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

TWO SYSTEMS FOR CONVERTING RASTER DATA TO NUMERICAL CONTROL DATA

Robert H. Thibadeau
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Dale M. McNulty
Department of Information and Computer Science
University of California, Irvine
P.O. Box 2926
Newport Beach, CA 92663

ABSTRACT

This paper describes two systems constructed for the purpose of converting paper and film documents to their numerical control or digital equivalent. One retools printing wiring board phototools, and the other recreates pen-recorder signals contained in oil well logs. Instead of focusing on particular algorithms, we compare and contrast the two system architectures, their functions, and the way they perform in their respective domains. The aim is to understand fundamental similarities that might suggest general pattern recognition problems and solutions.

INTRODUCTION

Problems of pattern recognition can be studied in many ways. One method is the method of complete systems. Complete systems theory can describe problems well. And, while this paradigm may not be a principal motivator of theory, the problems that it describes can lead to important insights in the study of pattern recognition.

Input	Output	Purpose
VVS:		
Printed wiring artwork, hand-taped, plotted or design drawing	Gerber ASCII Data	Drive Photoplotter
	Excellon Drill	Drive Driller
	AutoCAD (etc)	Edit in CAD
		Reverse Engineer from Artwork
	PostScript, HPGL Pen Plot/Postscript	
		Checkplotting
LIS:		
oil-well log	SEGY ASCII Data	Geophysical Analysis Programs
	Pen Plot	Checkplotting

Table 1. Comparison of the VVS and LIS systems by input, output, and purpose.

This paper describes two complete systems, one created as the Visus Vectorizing System (VVS) and Gerber Model 900 Scanner, and the other created as the LogTrak International System (LIS). Both systems run on the IBM-AT type platform enhanced with custom electronics where required. The algorithms, and the reasoning behind the algorithms, provide a number of insights into the problems of document scanning for the purpose of numerical conversion[1]. Space does not allow a detailed description of the respective input domains, therefore we assume the reader has some familiarity with the two tasks. Table 1 does, however, provide a brief overview of the input, output, and purpose of the VVS and LIS. Figures 1 and 2 provide visual examples of the two domains.

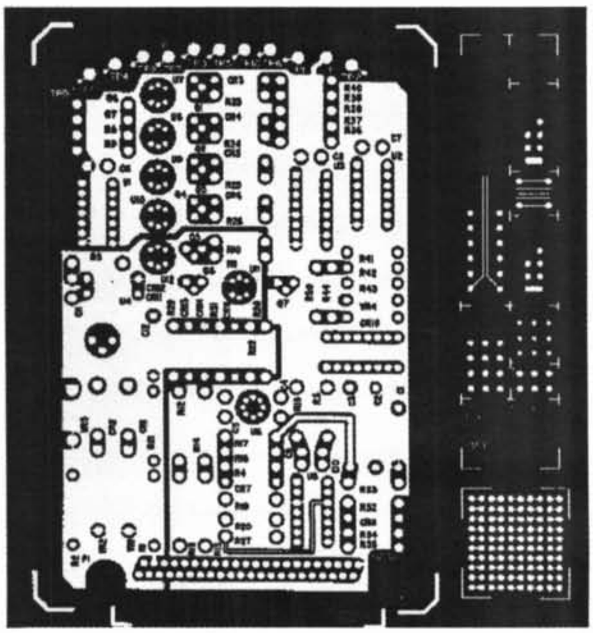

Figure 1. Example VVS input, a PWB drawing. All data must be converted

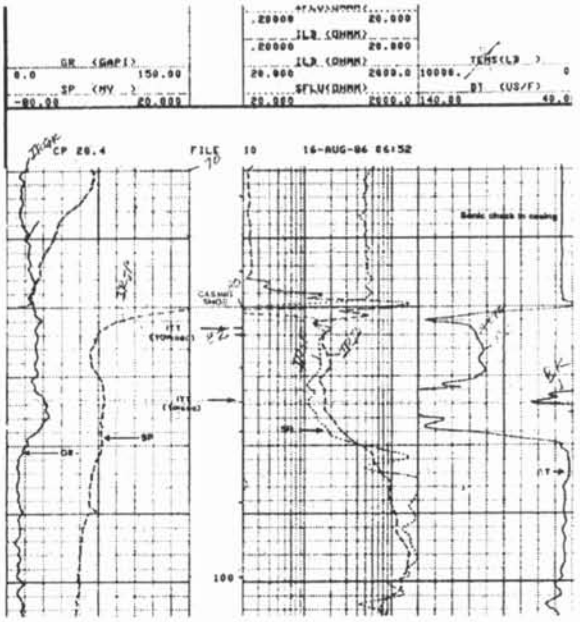

Figure 2. Example LIS input. The textual data is ignored by the LIS.

System	Scanning Resolution and Accuracy
VVS: Minimum 400 dots per inch Accurate to +/-.015 inch in absolute placement over 17X22 Inch area[2]	
LIS: Approximately 200 dots per inch 14 inch X 50 foot scan Scanner accuracy is not important.	

Table 2. The scanning requirements of the VVS and LIS.

	Entity Types	Tolerancing
VVS:	Draws (or "Lines", or "conductor") with an Aperture (or "shape")	Position +/- .001 inch
	On or Off Grid Flashes (or "Pads", "Lands", or "reliefs") with an Aperture (or "shape")	
	ASCII Characters (formed with Draws)	Clearances for Manufacturing
	Routing Contours	
	Registration Marks (formed with flashes)	
	Fill Areas (arbitrary shapes, "ground") (formed by painting draws)	
	Company/Trade Symbols (e.g., special apertures)	
	Negative/Positive Plotting (shape characterized by logical operations on binary images).	
	Layer	Layer to Layer Registration
LIS:	Solid, Dashed, Dotted Pen Recorder Lines	Line Placement (+/- .0875") and identification
	Which Track, Which Curve with respect to Grid	
	Text	Interpretation with respect to physical recording parameters like well depth and ohms/meter.
	Grid Lines	
	Rules of Usage and Jotted Notes and Jotted Exceptions	

Table 3. Comparison by Pattern Recognition Requirement

FUNCTIONAL COMPARISON

The pattern recognition in these systems is reconstructive. Table 2 summarizes the scanning specifications of the two systems. Table 3 summarizes the pattern recognition requirements. Both systems provide three classes of pattern recognition in sequence: recognition of raster skewing (an artifact of physical processes), recognition of grid (a means of confirming or interpreting entity locations), and entity recognition (different types of observable tokens). Taken together, these pattern recognition processes define an act of meaningful reconstruction.

The VVS and LIS utilize a 3X3 neighborhood processor to extract centerline features and outline features which are expressed as Cartesian coordinate pairs. The specific centerlining method (thinning and line-fitting) is arbitrary to these problem domains.

Skew Recognition: Both applications assume that long and nearly vertical or horizontal lines should be, in fact, vertical or horizontal. By inspecting the angular distributions of these lines for evidence of a skew, it is possible to nearly flawlessly deskew. All deskewing is of the affine (or "parallelogram") type.

Grid Recognition: Recognition of the invisible grid in the VVS application proceeds by (1) identifying object centers which are likely to reside on grid points; (2) comparison with a priori knowledge (e.g., .025 inch grids are common in the U.S.); and (3) calculation of grid offset and distance (see [3]). The latter begins with a computation of the Fourier transform of the "signal" provided by the horizontal and vertical projections of "grid" objects. This provides an initial estimate of offset and periodicity, however, the Fast Fourier precision is usually inadequate for the accuracy requirements of this domain. Therefore a least squared fit of the signed deviation between observed and predicted key points is derived from the initial estimate. Accuracy is further enhanced by incorporating knowledge of specific scanner's accuracy from repeatability calibration against a known grid.

Grid recovery from well logs is simpler because the grid lines are explicit. The LIS computes a horizontal and vertical projective histogram of deskewed horizontal and vertical lines and a peak analysis to extract grid. The difficulties are: (a) a trace near a grid line can be ambiguous with a grid line, (b) grid lines are often of poor

[1] Variously called: document conversion, raster to vector conversion (RTV), vectorization, and sometimes simply, but mistakenly, scanning. These systems are two examples of converting documents to digital or vector equivalent. Other domains where numerical conversion has potential include: optical character recognition (OCR), computer aided design (CAD), and numerical control (CAM and other names).

[2] Lower accuracy places greater dependency on accurate grid finding (a typical grid value of .025 inch implies system accuracy must exceed +/-.0125 inch after grid correction). Note that a second version, the VVS2, was created to provide an accuracy exceeding +/-.001 inch over an 18"X24" area.

quality or the grid count is input incorrectly. (c) a logarithmic grid exist. These problems are resolvable. In (a), trace tracking adjusts for traces along grid lines, in (b), the operator can check the grid line assignments, and, in (c), a logarithmic grid is recognized and the grid lines *inferred*.

Entity Recognition: The entities of interest to the LIS are the individual trace records. The LIS attempts to reconstruct the meandering trace lines from the centerline data (Figure 3). To do this it must determine when a trace starts, stops, and where it goes between those points. Traces can be composed of: solid, dashed, or dotted lines, and can crisscross each other and "change scale" or go off scale on the right to return with a larger value on the left of the same grid area.

Trace following is driven by two different constraint systems depending on whether the thinned, line segments can be associated with grid lines or not. When the centerline segments can be associated with grid lines, the interpretation is framed as a constraint system that is analogous to the recording process that created the input data. For example, a trace moving along a grid line can either exit the grid line or continue to move within it. An exit must either be from the vertical grid line, from a horizontal line, or from an adjacent vertical grid line (e.g., not from a vertical grid line two verticals over). To lower computational complexity all line segments associated with grid locations are initially removed from the lists of explicit candidate path lines.

Line segments that can not be associated with grid lines dictate a different constraint system. A few examples illustrate the type of constraints utilized: (1) The strongest constraint is the absence of alternative descriptions. (2) The strongest evidence for a successful path: (a) the end of one line and the beginning of another at the same location; (b) current trace is solid; (c) the "bend" is not near any grid lines; and (d) there are no other candidate line segment end points nearby. (3) If two lines are seen near a join, then a split with a line crossing is preferred.

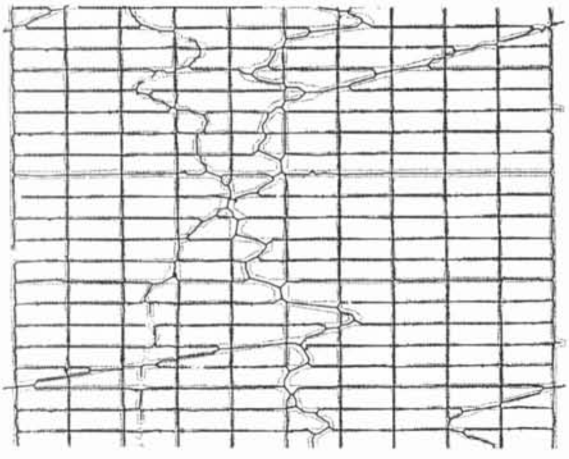

Figure 3. Centerline data for well log with an outline of the original raster image .

Line segments are assigned to the traces and constraints are propagated from the top to the bottom of the recording. A non-sequential method of constraint propagation could be employed whereby constraints would be passed up and down over some local distance. However, given the poor quality of the input data there was no appreciable advantage of this technique over the simpler, top to bottom, approach that reconstructs each trace separately.

The VVS must employ a different approach from the LIS because the entities have shape. The centerline data serves as a basis for an efficient search mechanism for identifying shape. VVS entity recognition techniques include: (a) profile projections to distinguish shapes such as round from square from cut pad, etc. and to characterize lengths and widths (see Figure 4); (b) polygons that replicate fill areas; (c) cycle detection (thinned lines that circle back on themselves) that help detect large pads and pads with holes.

The VVS recognition is highly reconstructive dependent on the constraints in the output language. For example, a hand-taped "pad" with a hole in the center would be interpreted as a solid "pad" in the output directed to certain photoplotters (see [2] for details).

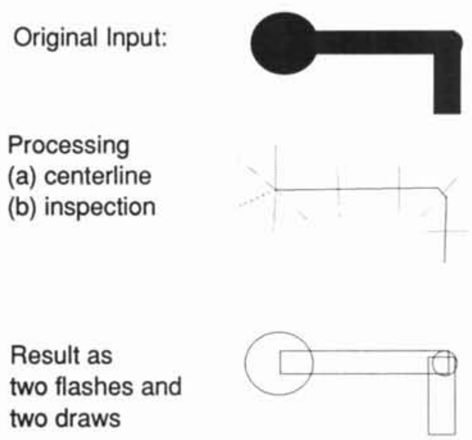

Figure 4. Projections and widths detecting entities.

Manual Editing: Manual editing, or cleanup, is necessary in both systems because: (1) each system must meet accuracy requirements that cannot be assured in the original material; (2) the scan objects are inherently ambiguous with respect to the output language.

In the VVS recognition precedes the edit process while in the LIS entity recognition, grid line recognition, and edit occur simultaneously. This difference is mandated by local, problem solving constraints. In the PWB domain errors do not propagate. The failure to recognize one pad correctly, in the worst case, only propagates to other, similarly shaped pads. In the LIS, on the other hand, the failure to recognize the correct subsequent line segment can propagate to every following assignment over the ten feet or so of log.

One must assume that automatic conversion can fail catastrophically in which case manual conversion would be necessary. Therefore, the editors must provide a complete manual reconstruction capability for their respective domains. Both editors provide for "raster-vector" overlay to enable hand-tracing (see [2] for other editing techniques). The LIS editor which integrates recognition, reconstruction, and editing also provides for recognition and reconstruction retry. In a further improvement the LIS editor (and the later VVS2 editor) can plot on screen from the final numerical control data.

PERFORMANCE ANALYSIS

Both systems have been exercised extensively with many documents, by operators possessing many skill levels. Evaluating system performance is difficult because of varying and subjective performance metrics. One approach is to consider a system's ability to perform recognition (see Table 4).

	Skew Correction	GridFinding	Entity Recognition
VVS:	~98%-100%[3]	~100%-60%[4]	~98%-80%
LIS:	100%	~94%-100%	~95%-75%[5]

Table 4. Quantitative performance expressed as a ratio of hit rate (a correct recognition of one item) as a ratio of the percentage of applicable instances. The values shown are representative results.

While the 98% recognition rate is commendable, these data are misleading because they measure local success not the global success of the system. A more realistic performance measure would be a comparison between the costs of total system throughput between automated and manual conversion methodologies.

One could compare the times to complete an average job (from input to numerical control data) as a function of total machine and operator time. This is problematic, however, because a great many variables influence the economic equation, including: (a) operator expertise and facility (these vary between people and within people daily); (b) the ability to judge the applicability of technology to specific conversion problems; (c) variation in drawing complexity.

[3]VVS skew correction is either all or none, however the results include approximately 2% of PWB artwork that does not have a skew which the system can identify (typically artwork representing analog electronics).

[4]Additionally, only about 60% of PWB artwork is consistent with respect to a standard grid.

[5]In the LIS, a hit is scored if the correct next segment is selected. Since probabilities cascade along a trace, performance can degrade quickly.

Another problem is that this type of data is generally proprietary. However, there is an evaluation method more general than the preceding but encompassing the specific issue: are service bureaus able to use and make money from automated conversion systems in general and the VVS and LIS in particular?

Generally speaking, the systems are promising but the answer is "not always". There are cases where automated conversion vastly outperforms manual digitizing, but often the most efficient manual methods are more economic than the automation achieved in these systems.

This conclusion seems puzzling since the VVS and LIS permit manual digitization at rates comparable to manual digitizing rates. The reason is that the set up, vectorization, quality control, and ancillary processing times associated with the automated techniques are superfluous if comprehensive manual digitization is ultimately required anyway. For example,scanning a typical, complex image with over 100,000 entities and converting to Gerber may take one hour or more. Simply loading the resultant Gerber file into an editor can take 15 to 30 minutes. If, after this time, the operator determines that comprehensive, manual digitization is required and much of the conversion is unsalvageable, then the entire conversion and load time is wasted. Reducing the economic cost of catastrophic reconstruction failure is important and difficult to accomplish.

CONCLUSIONS

This treatise has taken the complete systems point of view to compare and contrast two examples of numerical conversion systems. Some conclusions concerning system design that can be garnered from the comparison include:

(a) The low quality scan objects require robust data reconstruction methods.
(b) The system should provide immediate, interactive, feedback concerning reconstruction and the option to instantly retry the recognition and reconstruction in an updated context.
(c) The automated processes must not add cost to the operation of the system used manually.

In addition to these issues is a general principle best introduced by a VVS deficiency not discussed earlier. The VVS does not recognize a type of entity (symbol) known in the PWB domain as the "groundplane pad-relief." This shape, which is composed of two parts and resembles consecutive, matching parentheses (i.e., "()"), is not recognized by the VVS as a single object. We see three solutions to this problem: (a) manually augment the system for specific symbols; (b) incorporate more powerful, template-based, pattern recognition and provide for system training, or (c) imbue the system with adaptive learning procedures .

While the first is easiest for a limited domain such as well logs, it is economically problematic across large or ill-defined domains such as PWB. The second has been incorporated in the later VVS2 with success. However the last solution is clearly preferable in the long run.

Adaptive learning would examine the output language and its geometric generative capacity to infer input reconstruction. This purely linguistic approach, however, puts great pressure on the system's memory mechanisms and search strategies. The solution to this "bottle-neck" problem is to enable the system to learn so that only occasional operator/trainer intervention is necessary. This, however, presents another problem, a contention between the feedback and non-feedback learning elements (see [1]). Despite these problems, this is the goal for future work.

REFERENCES

1. McNulty, D. Learning, Attention, Feedback, and Non-Feedback in a Visual Recognition Domain. Ph.D. Dissertation (in progress). University of California, Irvine, CA.

2. Thibadeau, R. Problems of Automatic Vectorization of Artwork. Paper presented to the Fall Meeting of the Institute for Interconnecting and Packaging Electronic Components, San Diego, CA., September, 1986. Also, Report CMU-RI-TR-87-2, Robotics Institute, Carnegie-Mellon University, 1987.

3. Thibadeau, R.. Method and apparatus for scanning, storing, and reproducing artwork, U.S. Patent Serial No. 078,231, (awarded, September 1990).

ACKNOWLEDGEMENTS

This work was supported by Visual Understanding Systems, Inc., Pittsburgh, Pennsylvania. Special thanks to Joe Knight, Ph.D., and Ron Nichols of Logtrak Systems Inc., Houston Texas, who provided the support in every aspect of the Logtrak System.