MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-15, 1994, Kawasaki

Vector-based Editing Method of Drawings for Facility Maintenance

Hideaki Maehara, Jun'ichi Shibayama and Jun Sawamoto
Mitubishi Electric Corporation

5-1-1 Ofuna, Kamakura, Kanagawa 247, Japan

Abstract

Upon considerations on manual and recognition-based input methods, we propose a vector-based editing method to input facility drawings. The following features are implemented.

(1) Vector modification; The input drawing data are automatically re-edited and modified by predefined rules. This is for giving uniform precision, adding display attributes and keeping disposition rules of elements.

(2) Generation of management information; The facility management system needs also identification numbers, link pointers to other media, etc. not drawn in the drawings. They are generated automatically.

(3) Error detection in input; The input data are checked formally, by comparison, or by conformity conditions, and errors are detected automatically.

The drawings are edited by the upper functions for use in a facility maintenance system. We evaluated the method and proved its effectiveness.

0. Introduction

Recently, facility maintenance systems are becoming popular to improve management, for example, in a railway's track maintenance sections. As such systems have to handle technical drawings, a serious problem arises; how to input efficiently a big amount of drawings and documents, used and updated so far. Then aspired is a technology to solve the problem. In this paper, we propose a new method to input drawings, which we call a 'Vector-based editing method'. The method might be used instead of the already proposed ones, including recognition methods. We implemented its prototype system, evaluated it and proved its efficiency. We report its features, problems, and evaluation results.

1. The aims of the development

(1) Definition of facility drawings

In a facility drawing, lines, symbols and other drawn elements keep one-to-one correspondence to the real world objects, such as pipes, valves, electric wires, electric poles, etc. Electronic or logical charts, architectural drawings are the examples. By the way, topographic maps or machine assembly drawings are not the facility drawings, with many auxiliary lines like contours and dimensions, which are not the real world objects. A facility drawing is not used alone, but used with documents, known as 'facility ledgers'. They contain information of the drawing itself, and of the attributes for the drawn facilities, such as construction date, constructors, and other characteristics. An electric facility drawing is a typical one, used in power suppliers and transportation firms.

(2) Outline of facility maintenance

Facility maintenance has long been operated on paper media drawings. Generally, the drawings are used in the maintenance sections to get status of the facilities, for regular maintenance, emergency recovering from accidents or hazards, or for planning new facilities. A big company, an electric power supplier for example, keeps a large amount of drawings. In such sections, even searching of an adequate drawing took time, so that quick handling of drawings became impossible. Thus introduced are computerized facility maintenance systems, from the later 1980's.

The systems are also used for automatic maintenance and checking operations. Such systems are generally equipped, not only with drawing information, but also with information of the facility elements in characters and numerals. And both information types are unified in database, supporting various retrieval functions. Figure 1 shows the basic structure of a facility maintenance system.

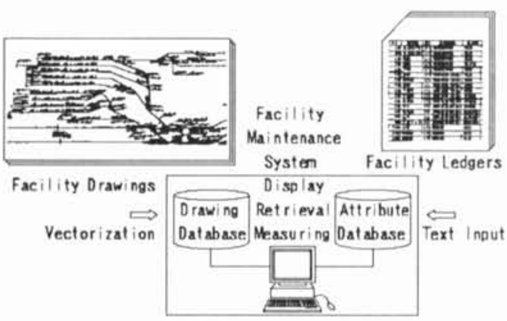


Fig.1 Basic Structure of Facility Maintenance System

(3) Purpose of technology development

A serious problem came up in an introductory phase in a facility maintenance system; how to and by whom to input initially the large amount of drawings, containing many figures and characters. There are image filing systems, which input binary images. But the image data are not adequate to retrieve and to edit the already drawn contents. Thus handling of vector records, or 'vectorized drawings', is necessary for facility maintenance. A primitive way to input the vectorized drawings is manual CAD operations, by a keyboard and a mouse, with an original paper drawing at the operator's side. This is also inadequate, demanding much cost, time, operator's training and job planning. Checking and correcting phase, aside the input itself, are also costly. Therefore, we need a technology, with less cost, less time and fewer errors. It became the purpose of our development, based on various computer technologies.

In the following Chapter 2, a manual input process is reviewed, in the Chapter 3, conventional drawing systems are reviewed, and upon them, in the Chapter 4, our vector-based editing method is introduced. In the Chapter 5 we evaluated our method.

2. Manual input of drawings

We begin with the wholly manual input process of drawings, to illustrate the problems in the conventional input technologies.

Phase 0: - Arrangement of sources -
The paper-based drawings themselves are often not suitable as sources for computer input, because their formats may differ in every local office, and the contents may have lacks, overlaps or other faults. So the materials to be used by an input operator, which we call 'the source documents', must be re-edited from the original drawings.
The facility management system also needs text documents for ledger database, and they might be re-edited and input.

Phase 1: - Input of facility drawings -
In this phase, an operator refers to the source documents and makes up facility drawings in vector format. Namely, the operator reads the paper-based drawingws, and draws by a CAD system.

Phase 2: - Input of management data -
A facility maintenance system needs information to achieve graphical retrievals, for example, identification numbers, retrieval keys, and link pointers linking plural entities. In this phase, such information, called as 'management data' are added.
Some management data already exists in the original drawings, or added onto them, to be input in the phase 1. Some other data may be generated by a software tool. But there are still other data, and they are to be inut by the operator.

Phase 3: - Detection of input errors -
In this phase, the operator should inspect and find errors in the vectorized drawings and the management data.
The drawings are inspected in two manners. In a format inspection, the operator checks the conformance to the rules, formats, layouts etc., only with the printed out drawings. In a content inspection, the operator compares the printed out drawings with the source documents.
On the other hand, it is difficult to find the management data errors by sight, except for the data already drawn on the drawings. The only way to detect such errors are to execute the retrievals; on error, the right data would not come out.

Phase 4: - Correction of errors in input data -
The operator corrects the detected errors, using the CAD system again, or using the facility maintenance system functions.

3. Conventional drawing input methods and their problems

To overcome the problems of manual input of drawings by CAD systems, automatic drawing input methods have been developed, with the academic insight of binary image pattern recognition. But still there remains problems.
One of the methods, appeared in the early time[1], is a batch-type input of drawings, as in figure 2(a). In it, a system automatically tries to recognize a binary image. In spite of development, such a method could recognize only limited kinds of elements, not densely drawn, on a noiseless drawings. Afterwards, interactive-type drawing input method appeared[2], as in figure 2(b), to recognize more noisy and complex drawings. This method tries to improve the recognition rate by adding human judgment, as giving hints for recognition or instructions for error correction. These two methods gave good performance under certain conditions, but common problems are left. We describe the problems in the following sections, according to the phases in Chapter 2.

(1) Problems on phase 1
- limits of recognizable drawings and their accuracy -

The majority of drawings have various disadvantages for the conventional drawing input methods; not precisely drawn originals, deterioration during paper storage, distortion by multiple copy generations, etc. As the pattern recognition ability has not yet reached the human capability, a computer can hardly cope with the common drawings read and used by men. This suggests, that drawings with good quality, as in figure 3(a), are suitable for recognition, but most drawings are not suited. The figure 3(b) is one of such actual drawings, with many dimm lines, cut outs, and overlapped elements. Our vector-based editing method just aims those drawings.
There is another problem, concerning accuracy in drawings. The conventional methods extract the exact positions of the figures as they are on the paper. So are the derived attributes information. This is just the aim of geometrical map recognition. In facility drawings, inclinations and distortions in the original drawings should disappear. Often the end user wants to get a fair copy, even if the paper-based drawings are not 'fair'. New rules, not applied to paper drawings, are sometimes desired; for example, figures should be on discrete grids, or should have certain relative dispositions each other. These are not the primary tasks of recognition, but of our method.

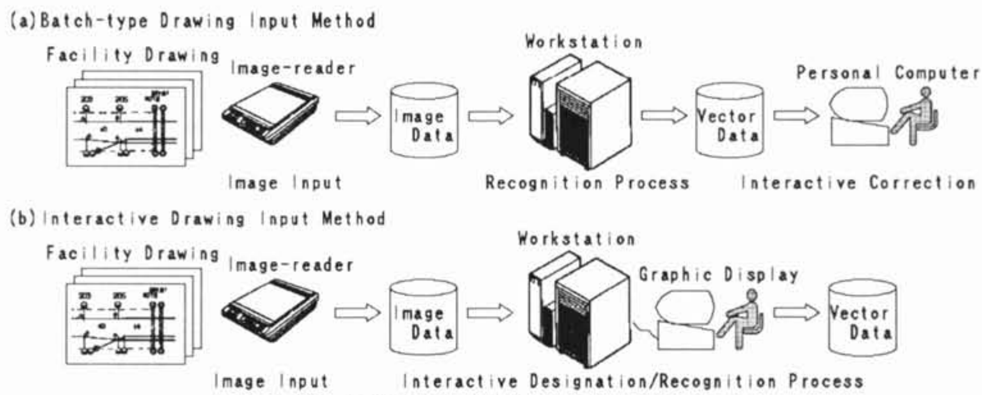
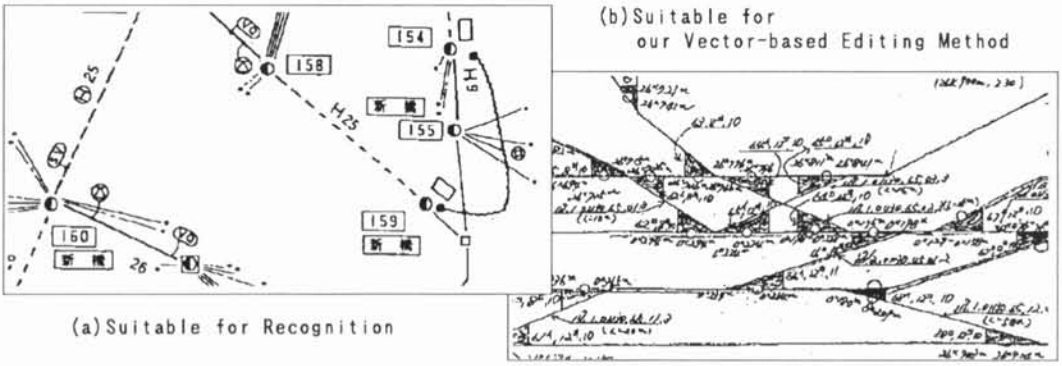


Fig.2 Conventional Drawing Input Technology



(b)Suitable for
our Vector-based Editing Method

(a)Suitable for Recognition

Fig.3 Examples of Drawings

(2) Problems on the phase 2 - lack of management data -

While faithful reproduction of the paper drawings is the purpose of recognition methods, facility maintenance drawings need other conditions, as follows, for management requirements.

- ○ An entity in the real world is often symbolized by plural basic figures. These figures should be grouped, to keep one-to-one correspondence.
- ○ The figures should have correct display attributes, layer informations, retrieval keys etc., for management functions.
- ○ The figures should have link information to other media, as characters, numeric data, or photographs, that show the same real world entity.
- ○ Newly needed information, supplied at the vectorization phase should be added.

These are the requirements for our vector-based editing method.

(3) Problems on phase 3 and 4 - reliability of data -

There are two types of errors after input; those caused by an operator during input, and those inherent in the original source data. Whatever are the causes, the final data should not contain errors, especially in public sectors like transportation and electric power firms. So it is an important task to assure the reliability of the data.

The best known method to prevent errors, is to compare the source data with the computerized data. But this work requires labor, almost equal to input the data again. And there are no guarantees of error elimination. There are few research works found in this field. We tried also in this field.

4. Approach to facility drawing input by Vector-based editing method

As above discussed, neither the drawings input by manual CAD systems, nor by conventional recognition methods, meet the requirements in a facility management system, since they don't have enough precision or informations needed, or don't satisy the drawing rules newly required. Our vector-based editing method is proposed to meet these requirements. In the method, after manual input of vectorized data, the system checkes the data according to the predefined rules or precision limits, and automatically edits and regenerates the data. Thus an operator needs not to draw precisely, but to draw roughly, with time saving, to let the final precise drawings come out from the system.

Moreover, error detection functions and other advanced functions are to be included.

In this chapter, we describe the approaches taken to each phase, and the implementations of our system.

(1) phase 1 - vector modification -

The vector-based editing method processes vectors, drawn by an operator. Thus there are practically no limitations of the source documents, since they are read by human eyes. As the method accepts roughly or partially drawn data and converts them to precise drawings, the precision problems are also solved, and operator's load to keep precision during drawing is reduced.

The functions are predifined and set into the system. There are many types of vector modification functions, and are classified as follows.

- ○ To arrange figures in relatively or absolutely correct coordinate positions.
- ○ To modify the shape of figures according to the drawing rules.
- ○ To give the display attributes etc. according to the drawing rules.
- ○ To re-arrange figures, inadequately overlapped or contacted.

There are two difficulties in the vector modification. It is rather easy to edit and modify a selected figure, since it is a numerical procedure. But the selection from a lot of figures is not easy. For instance, we wish to dispose a figure and its attributive character strings nearly, so the first task is to link them. When the elements are densely disposed, a simple algorithm as to select the nearest element cannot do it. This is a mathematical modeling problem, and we have to use pattern recognition techniques partially. Another difficulty is the need to treat plural figures at the same time. To relocate an overlapped figure, the space to be relocated must be free, for example. This is also a modeling problem of two-dimensional space.

We implemented trial-and-error algorithms for both problems, and the mathematical sides remain as future tasks. Figure 4 shows examples of the implemented vector modification functions in our prototype. In this case, an operator inputs a triangle figure roughly on a position, as in (b), to let the system automatically dispose the right-angled triangle just at the corner of the already drawn strokes, as in (c). Then the operator inputs attributive character strings, as in (d), to let the system dispose them, as in (e) to keep relative location to the triangle. And the made up figure (f) is supplied with necessary display attributes (hatching in a polygon, size of characters, kind/thickness/color of the lines, etc.) and the dispay layers. The operator's task is decreased, to draw (a) at a rough position. Yet the accuracy is guaranteed in the final data for facility management.

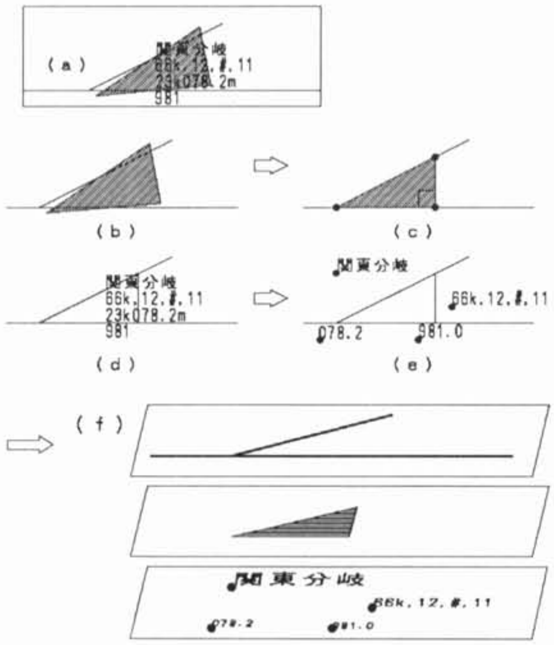


Fig.4 Examples of Vector Modifying

Here let us make supplemental comparison to the CAD operations. The CAD has object-snap function etc. to facilitate accurate drawings. How to input the figure 4(c) is as follows. First an operator selects "Intersection neighborhood mode" and inputs the top position, corresponding to the turning point. Next, he selects "Neighborhood line mode" and inputs the right-angled top. Then he selects "Orthogonal mode" and generates a perpendicular from the last input top. Finally he selects "Neighborhood line mode" and input the third top. But in our vector-based editing method, the operator only has to move the mouse cursor and roughly input the three points. This shows, that our system, although using manual operations, is better and time saving.

(2) phase 2 - automatic generation of management data -

A facility maintenance system is to offer various retrievals among different media. To make them able, the equipments graphically represented must have the identification numbers, retrieval keys, the link pointers, etc., so called management data. When only paper-media were used, operator's experiences and knowedges were used. But for computerization, the information must be in the system, whether by input or by generation.

The management data are often automatically derivable from the input data. This is because of the reduncancy and similarity in the input data. Since there are plural sources for one physical entity, the redundancy is inevitable. Generation of the data is often a simple numeric procedure. But it is not easy to formulate, which data are to be used to get what information. So it is difficult to make general routines to cope this problem.

Among the management data are also graphic entities, as later shown in figure 5(a2) and (b2). They are not facilities, but their attributes. Thus they can be generated from the ledger information. It is hard to draw them manually, since there are no source documents showing their shapes and positions.

We implemented generation methods in a trial-and-error manner. The generalization is one of the future tasks. Figure 5 shows examples of generating the management data in our prototype. The line figure (a2) is generated from character/numeric data (a1), the symbol of (b2) is from the character/numeric data (b1). In this case, not all information for graphic entity, such as display coordinates, is supplied from the ledger data. These lacking data must be supplemented by analogy from the graphic data near the coordinate position of the figures to be drawn. In the examples, only the horizontal coordinates are recorded in the ledeger data, and the vertical coordinates are calculated from the dashed line, the base figure drawn by an operator. The shape of the figure are obtained from the ledgere data.

(3) phase 3 and 4
- automatic extraction of errors in the input data -

In the chapter 3, we described the detection and correction of data errors. Let us consider making data inspection process, to detect contradictory and unmatched data, and correct them if possible, in a certain predifined range, to reduce the operator's load.

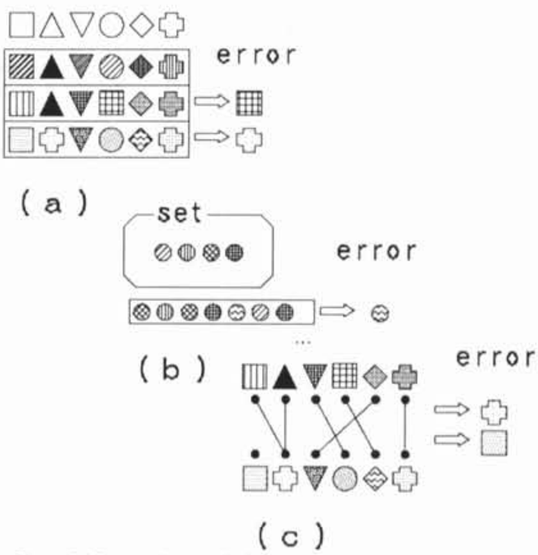


Fig.6 Examples of Data Error Correction

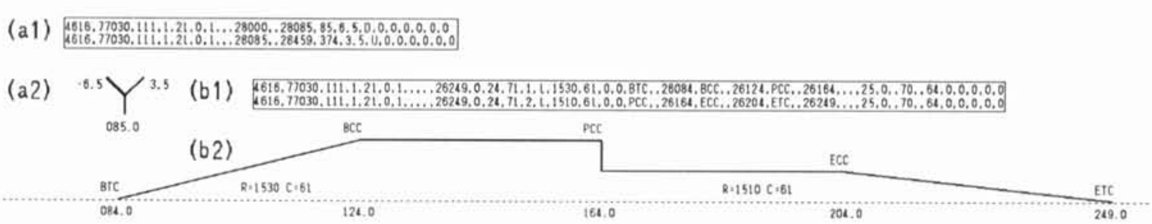


Fig.5 Examples of Automatic Management data Genaration

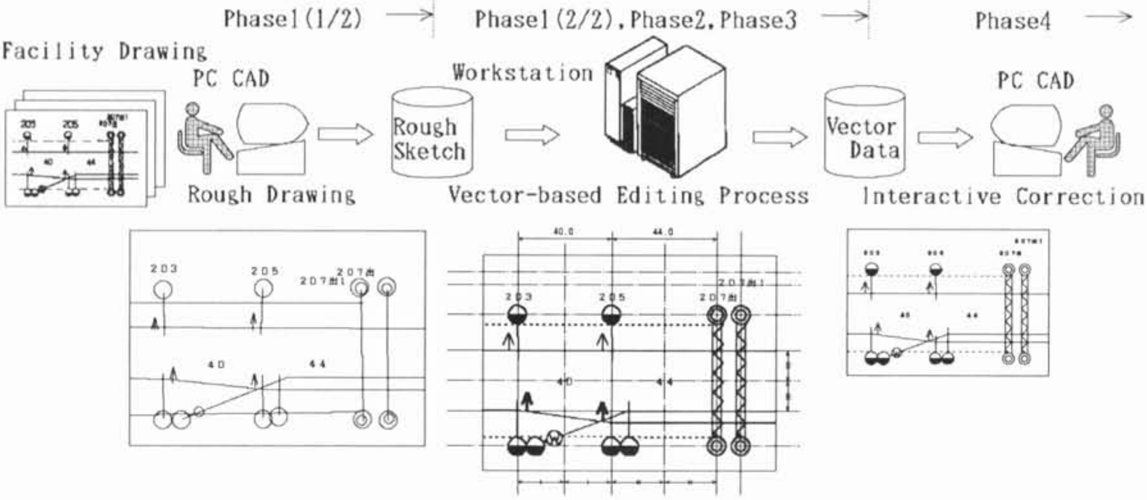


Fig.7 Outline of Vector-based Editing Method

The redundancy and similarity among two or more data are to be used, as in the generation of the management data. This is because a database keeps data with mutual relations, with dependency. The process might also correct the detected errors, when similarity relations are established, but this case is limited. We therefore consider only the detection process, as follows. Here, 1) and 2) are formal checks for structural and format errors. 3), 4) and 5) are content checks, according to the meaning of each data element.

1) To inspect the data structure and the data type, as integers, characters, etc.
2) To inspect the data, using relation conformity, as one-to-one correspondence, order relations, etc.
3) To compare the data with other data set, which is guaranteed to be correct.
4) To collate the two data, which have same information.
5) To collate the data to a knowledge base.

Thus, the patterns of the errors are easy to formulate. When automatic detection methods can be generally formulated, they will greatly contribute in making database.

We implemented some of the error detection methods in a trial-and-error manner. Evaluation and generalization of the methods are future tasks. In our prototype, following error detections are applied; inspection of data type, as in figure 6(a), comparison with the master data of guaranteed quality, as in figure 6(b), and conformity check to one-to-one relation, as in figure 6(c).

We have described the approach of the drawing input by the vector-based editing method above. The outline of the drawing input by this method is shown in figure 7.

5. Evaluation of the vector-based editing method

(1) Qualitative evaluation

We evaluate the time reduction for input of a drawing. Our vector-based editing method aims the phase 1 to input facility drawings, the phase 2 to input management data, and the phase 3 to detect input errors. The conventional methods cope with the phase 1 only, but can utilize our phases 2 and 3. So first we compare the phase 1. The conventional recognition methods give full reduction without operator's interaction and seem better, while our method seems to be only a half way, since it requires manual input operation. Meanwhile, the reduction rate of 100 % would be attained only if no errors occur in recognition. In real cases, an operator has to seek the lacks and errors, to add the unrecognized elements and to correct the errors. This leads to a supposition, that 80 % of recognition rate is not relevant for time reduction. Errors may be caused during recognition in many stages; the initial drawing conditions, multiple generations of copying, deterioration while paper conservation and binarization conditions, etc. On the contrary, the vectorization method is influenced only by the initial drawing conditions. And the CAD operations guarantees better drawing conditions. These considerations sustain the better reduction rate in our method, with lower error rate for paper-media drawings of any quality.

(2) Quantative evaluation

As described, there are 4 phases to input drawings. Let F_i be the operation time for the phase i, and C_i be the manual input ratio of the phase i. The time reduction ratio R is defined as a ratio of input time, to an input time for wholly manual input, namely;

Time reduction ratio $R = \Sigma(F_i * C_i) / \Sigma F_i$.

The smaller R, the better performance is derived. The manual input ratio C_i may differ greatly according to the species of the drawings and the applications of the systems. But prior investigation and analysis of the total amount of data gives some predictions. Also a test input in a small scale gives a predictive value.

(3) Evaluation examples for facility drawings

We took for example a real facility drawing for 6.5 km of railway's track maintenance. We experimentally compared the time for manual input and our vector-based editing method by the example, to obtain a quantative evaluation.

A: manual input time (in minutes)

phase 1: input of a facility drawing by CAD precisely
- input of vector figures (track lines, etc.)
  268

phase 2: input of management data
- additional input of figures (curve information, etc.) not drawn on the original drawing
  713
- input of linking informations, to link with character/numeric based documents
  34

phase 3: detection of errors
- cross check with the original drawings
  97
- confirmation of the linking informations
  17
- cross check with the character/numeric documents
  0 (not done)

phase 4: correction of errors
- correction of the drawings
  58
- correction of the management data
  64

B: input by our vector-based editing method

phase 1: input of a facility drawing by CAD roughly
- input of vector figures (track lines, etc.)
  112

phase 2: input of management data
- additional input of figures (curve information, etc.) not drawn on the original drawing
  0 (automatically generated)
- input of linking informations, to link with character/numeric based documents
  0 (included in the phase 1)

phase 3: detection of errors
- cross check with the original drawings
  65
- confirmation of the linking informations
  0 (automatically done)
- cross check with the character/numeric documents
  0 (automatically done)

phase 4: correction of errors
- correction of the drawings
  58 (the value for A used)
- correction of the management data
  64 (the value for A used)

(assumed that error rates may be the same as A)

According to the experiments, we obtained the following results.

* manual input ratio for phase 1: C_1 = 41.8 %
  phase 2: C_2 = 0.0 %
  phase 3: C_3 = 57.0 %
  phase 4: C_4 = 100.0 %
* time reduction ratio: R = 299/1251 = 23.9 %

In this case, as the time to input management data is dominant, we obtained a good time reduction ratio of less than 25 %. But the vector input of a drawing showed only 40 %. As stated, this value varies, according to the type of the drawings and their applications, and we may expect 30 % of mean time reduction ratio.

We did not carry out cross checking with the character/numeric documents for the manual input, since it was hard to do manually. Such a work with mass data is suitable for computerization. This automatic checking surely helps in finding errors.

(4) conclusion remarks

We see, that the conventional method by recognition should be best, if the recognition rate would be considerably high. But as such is difficult to guarantee, our vector-based editing method is practically the most suitable. As our method treats vector data, it is more compact in system size and speed, than the binary-image based recognition method.

6. Conclusions

We proposed a vector-based editing method as a new method to input facility drawings and related informations, and proved its effectiveness. As there are various applications of drawings, our method may become one of the mainstreams in the input, instead of the recognition technologies. The future of our method depends on the generalization of the application systems.

Our main target in this paper has been input of drawings. The vector-based editing method can be extended farther to input of characters, numerals, documents, images and other media, since the main ideas to detect errors or to generate management data are all common. Input of a drawing has to be integrated in more general data input method.

Another open side is a human-interface oriented approach, since computer operations are closely related to operator's actions. Without such considerations, we would not get the best solutions for maximal efficiency or minimal error rates. We are now confronting such fields.

References

1) T. Futatsumata, G. Shichino, J. Shibayama, and A. Maeda : "Development of an Automatic Recognition System for Plant Diagrams", Proc. IAPR Workshop on Machine Vision Applications, pp. 207-210, 1990.

2) Y. Nagano, H. Kanechika, S. Tanaka, H. Maehara and A. Maeda: "Experimental system using an interactive drawing input method", Proc. SPIE Workshop on Visual Communications and Image Processing '91 : Visual Communication, SPIE Vol.1605, pp. 614-623, 1991