ARCHITECTURAL COMPUTER VISION: AUTOMATED RECOGNITION OF ARCHITECTURAL DRAWINGS

A. KOUTAMANIS and V. MITOSSI
Faculty of Architecture
Delft University of Technology
Berlageweg 1, 2628 CR Delft
The Netherlands

ABSTRACT. Computer vision offers the ability to transform digitized drawings into documents that can be used with computer systems. Recognition of digitized drawings can occur at the levels of (a) geometric elements, (b) building elements, and (c) spatial articulation. The last two levels apply not only to digitized images but also to computer-produced ones. The enormous burden placed on the user for inputting and manipulating CAD drawings suggests that automated recognition can add to the capabilities of CAD by making the computer more flexible with respect to inputting design information and more responsive to the actual concerns of the designer.

1. Digitized Drawings

One of the pressing problems in the computerization of architecture is the transfer of existing architectural drawings to the computer. This is done either to preserve and disseminate historically significant documents or to make the (usually massive) archives of a professional design firm accessible to computer-aided design (CAD) systems. The traditional manual digitization of drawings using tracing devices like graphic tablets to input existing drawings to CAD systems is too labour-intensive and unreliable to offer a viable large-scale solution. Optical digitization of architectural drawings resolves the problem to a degree only. Digitized drawings are reasonably accurate and detailed copies of the original. Their resolution and portability are adequate for any practical purpose such as reproduction with laser printers and plotters. However, digitized images may also demand enormous storage capacity and are often slow to recall and manipulate. In particular, digitized images are more difficult to modify than CAD documents, while they are incompatible with design analysis and evaluation programs.

One way to reduce storage sizes, improve recall/manipulation times, and enhance compatibility/utilization with CAD and other computer programs is the automated recognition of digitized architectural drawings. Recognition of digitized drawings can occur at three levels. The first is recognition of geometric primitives such as lines and shapes in a drawing, which amounts to a direct transformation of a digitized image into a CAD document. The second level is recognition of building elements, such as doors, walls and columns. This represents a transformation of the digitized drawing into a more abstract, comprehensive and knowledgeable description that remains compatible with CAD and other computer programs. The third level is recognition of the more implicit spatial articulation of a design. This level is the most relevant to architectural design thinking, as it facilitates analysis and evaluation from several points of view, as well as the execution of corresponding design manipulations.

2. CAD Drawings

Another problem that architects will sooner or later have to face concerns the arduous nature of

H. Timmermans (ed.), Design and Decision Support Systems in Architecture, 17–25.

current computerized drafting, which offers little more than a digital equivalent of the drafting table. In computerized drafting, the user has to input the description of a design in great detail and with high precision. The computer facilitates certain drafting operations such as copying and pasting or repetition of a standard element far better than the adhesive sheets of symbols and drafting templates used in hand drawing. Moreover, it allows what appears to be extremely effortless and versatile manipulation of drawings. Any individual line and shape can be modified in position or dimensions with great ease and accuracy, while grouping capabilities simplify such operations on more than one graphic element.

All these, together with the unprecedented compactness of storage, paint a very positive and optimistic picture of computerized drafting. This picture certainly conveys all the advantages but fails to acknowledge some of the problems that arise in practice. One such problem is currently obscured by the employment of specialized staff to run the computer facilities instead of the architects who design a project. This problem is that effectiveness, efficiency, reliability, and flexibility in the modification or other manipulation of a computer-produced drawing is primarily dependent upon the structure of the graphic elements which comprise the drawing. Placing a single element on the wrong layer, in the wrong group, or connecting by mistake the dimensions of unrelated building elements does not simply reduce the speed of manipulation. It may also render useless the documents which were automatically produced on the basis of the drawing, such as bills of materials and cost estimates. In order to avoid the frustration and financial loss caused by such events, it is imperative that the structure of a CAD document is always impeccable.

The enormous intellectual burden placed on the user for inputting and manipulating computer-produced drawings can be reduced through the automated recognition of the building elements and spatial characteristics in the configuration of lines and shapes input by the user. This effectively means that the computer interprets the user input and automatically elaborates/structures the representation of a design with respect to both general graphic/geometric aspects (Pavlidis 1988) and the specific domain-dependent issues discussed in the present paper. The computer can recognize the building and spatial elements in the configuration of lines input by the user and allocate each identified element to the correct layer and group, as well as establish the appropriate connections between identified elements. The user can therefore input the drawing more freely and flexibly, more or less as he would present the design to a skilled draftsman. This type of computer support to user interaction represents a logical and fruitful evolution of drawing software toward the established goals of design automation. In fact, recognition of building elements and spatial arrangement in a CAD document is far easier than in a digitized image because of the former's reliance on vector graphics as opposed to the latter's pixel basis.

3. Geometric Primitives

Recognition of geometric primitives such as lines and shapes is essentially similar to what has already been attempted for the automated conversion of engineering drawings to CAD documents (Hofer-Alfeis 1987; Joseph 1989). Chain coding and edge following are employed in order to identify the straight and curved line segments in a drawing and subsequently also the two-dimensional shapes (mostly closed loops) these line segments form. Figure 1 is an example of the application of chain coding and edge following techniques to the recognition of geometric primitives in a floor plan. First, the appropriate chain code is allocated to each black pixel of the image and then uninterrupted sequences with the same (or, for curves, related) coding are grouped together. Further grouping is possible on the basis of rule systems that examine the relationships between connected or adjacent groups of pixels (i.e., line segments) against prototypical or typical line configurations in standard symbolizations of doors and windows, for instance. In architectural drawings, the width of lines varies more than in schematic or mechanical engineering drawings. This variety poses more problems for the already tenacious line following (problems which can be remedied by thinning the image prior to line following), but it also allows for greater precision in the identification of shapes composed by the recognized line segments, because in many cases only intersections of segments with the same width are considered.

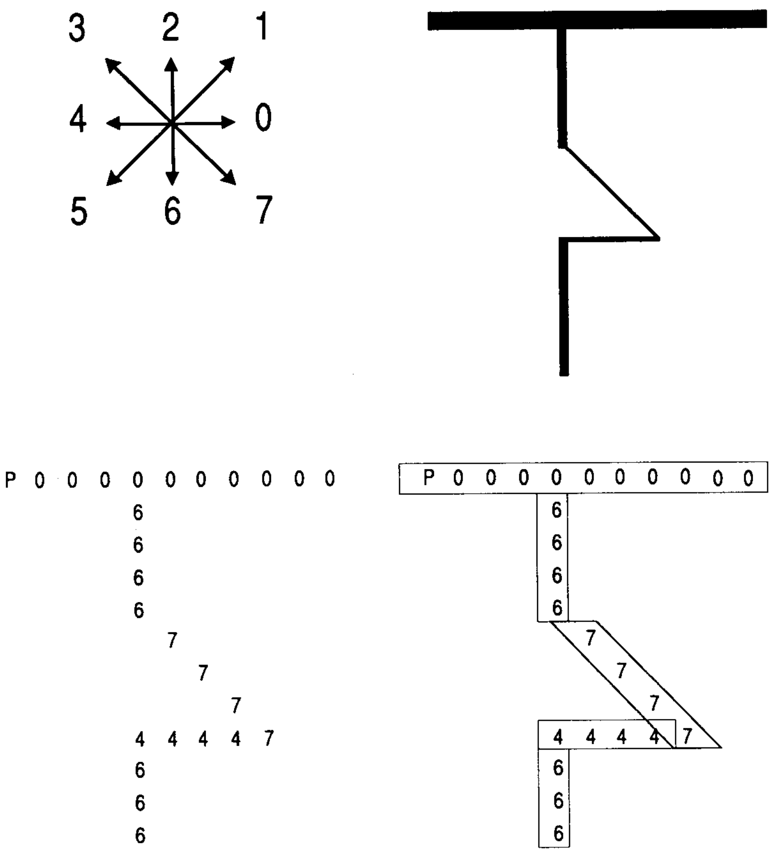


Figure 1. Recognition of geometric primitives using line following: an example
top left: chain coding scheme; top right: floor plan part; bottom left: coding of pixels (starting from an arbitrary point P); bottom right: grouping of pixels

In orthographic projections (floor plans, sections, elevations) recognition of geometric primitives transforms the digitized image into a collection of two-dimensional shapes that is fully compatible with two-dimensional drafting software. In perspective projections (which may even include shadows and shading), three-dimensional shapes formed by the line segments can be recognized by adaptations of the technique used for the recognition of line drawings of scenes with trihedral objects (Waltz 1975). These recover the volumetric composition of the scene and thus render the digitized image compatible to three-dimensional CAD software. Recognized axonometrics and isometrics normally require few modifications and corrections, while other perspective projections may require extensive dimensional corrections.

Recognition of lines and shapes effectively makes the digitized document compatible with

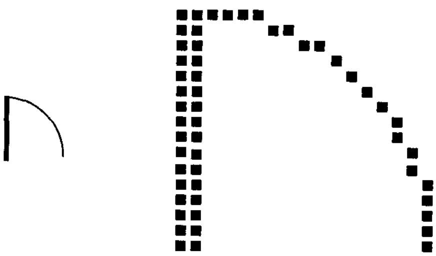
CAD programs, as these also employ geometric primitives to describe a design. Recognition at this level is therefore extremely useful for the computerization of existing drawings and their subsequent manipulation with computer programs. However, as related research into the transfer of other types of engineering drawings has shown, it is also computationally tenacious and expensive (Hofer-Alfeis 1987; Joseph 1989). Recognition techniques and algorithms are in several cases only poorly served by the widely available computer technology, and the current patterns of change do not suggest great optimism for the immediate future. Moreover, the utilization of a CAD document is always dependent upon its structure (subdivision of graphic elements into layers and groups, annotations). This means that in order to use the recognized drawing for anything beyond simple graphic operations, it is necessary to develop its structure interactively. This is a time-consuming and labour-intensive process. In conclusion, it appears that the main benefit from this level of recognition is compacting the size of a digitized image while at the same time improving the usability of the image with the types of software most often used by architects and designers.

4. Building Elements

The second level at which recognition can be applied to architectural drawings is that of building elements such as walls, columns, doors or windows. Most of these elements have a more or less standardized appearance in each type of architectural drawing, as manifested by the intensive use of predefined symbols in architectural drafting (on adhesive sheets or drafting templates in hand drawing and in the form of menus and 'CAD libraries' in CAD). Such standardization invites the use of *recognition* (as opposed to drafting) *templates* (Figure 2). Template-based recognition of building elements in architectural drawings is essentially similar to template-based optical character recognition (OCR). Recognition is achieved by matching each segment of the image to a number of known examples or prototypes which represent the types of symbols possible in an image of a particular type (i.e., a floor plan or a section). As in OCR, a certain degree of training and adaptation is required, and recognition can be efficient only if the number of possible patterns is relatively low.

The number of basic types of building elements in each kind (mostly floor plans and sections) and scale (typically 1:200 to 1:50) of architectural drawings is indeed limited. But considerable variations in the symbolization of building elements can be caused by differences in the materials and form of the real elements, as well as by differences in drawing conventions. The situation is analogous to OCR of multi-font texts where similar variability is confronted by appropriate structuring of the template database and/or abstraction of the patterns (e.g. skeletonization of characters so as to eliminate font attributes). More effective in the case of architectural drawings and especially for floor plans, the most essential drawings for conveying information on building elements, is the use of elastic templates and templates with 'don't care' pixels (Persoon 1988). Such templates can easily account for local perturbations, small variations in orientation, and other imperfections caused either by real differences in the original image or by the process of digitization.

One of the key problems in the use of templates for the recognition of building elements in architectural drawings is that some elements do not occur in standard sizes. In a floor plan at the scale of 1:100, for example, one can expect that practically all door symbols in a residential building will be approximately 1 cm wide. The use of elastic templates or templates with 'don't care' pixels can account for small variations in size or type. The length of wall symbols, on the other hand, may vary from a few millimetres to several centimetres. The use of adaptive hierarchical partitioning techniques like quadtrees can resolve such problems through the subdivision of the image into parts containing a single building element symbol. Partitioning on the basis of a uniform grid determined by the size of key elements, such as doors, is also possible. In some cases (e.g. in drawings of buildings which make use of standardized elements or modular coordination), partitioning is more relevant to the structure of the particular design.



```
00000000000000000000
01111110000000000000
01100001100000000000
01100000011000000000
01100000000100000000
01100000000010000000
01100000000001000000
01100000000000100000
01100000000000010000
01100000000000010000
01100000000000001000
01100000000000001000
01100000000000000100
01100000000000000100
01100000000000000100
01100000000000000100
01100000000000000100
```

Figure 2. A door template
left: a common representation of a door in floor plans; middle: pixel array corresponding to the representation; right: a template for the recognition of the representation: ones match to black and zeros to white pixels

Recognition of building elements, being a knowledge-based process, is more demanding than that of geometric primitives. Still, recognition at this level can be more efficient and reliable, as well as more relevant to the designer user. The resulting description of the image as an array of building element symbols can always be transformed into a metrically precise and detailed drawing by substituting each instance of a symbol with an accurate drawing of the denoted building element. This technique is in fact more effective and efficient than the structuring of the elements produced by recognition at the level of geometric primitives, especially since it can accommodate different conventions or habits in computerized drawing (i.e., different approaches to the use of layers and grouping relations) and hence also ensures compatibility with a greater number of drafting software.

Another advantage of describing a drawing in terms of its building elements, even if the description is rather abstract, is that certain design analyses and evaluations (e.g. structural and thermal analyses) apply directly to the configuration of building elements. For these evaluations, the type of building elements and their physical/spatial relationships are generally more important than their precise dimensions and detailed appearance. In the framework of the above transformation of the abstract symbol array into precise and detailed drawings, these evaluations can even play a role in determining the precise geometric attributes of the building elements.

5. Spatial Articulation

The third and more advanced level at which automated recognition can be applied to visual architectural representations concerns the spatial articulation of a building and, through that, the most essential aspects of architectural design. On the basis of the spaces in a floor plan and the arrangement of these spaces with respect to geometry, adjacency and physical access, for instance, it is possible to evaluate the design with respect to design requirements concerning circulation (Koutamanis and Mitossi 1991) or even formal goodness (Koutamanis 1990).

Recognition of spatial articulation is best performed on the basis of *skeletonized* versions of the digitized image. The purpose of skeletonization is to reduce information existing in the image to the absolutely essential for the recognition of spatial characteristics. This is achieved by abstracting or even removing whole building elements or properties of elements that are only secondary for the overall spatial articulation of a building. These may be the number of stairs and the exact form of a staircase, the thickness of walls, the type and precise position of doors and windows, as well as the furnishing of a space. A skeletonized floor plan, for example, includes only the boundaries of each space in the building, reduced to unity thickness (Figure 3).

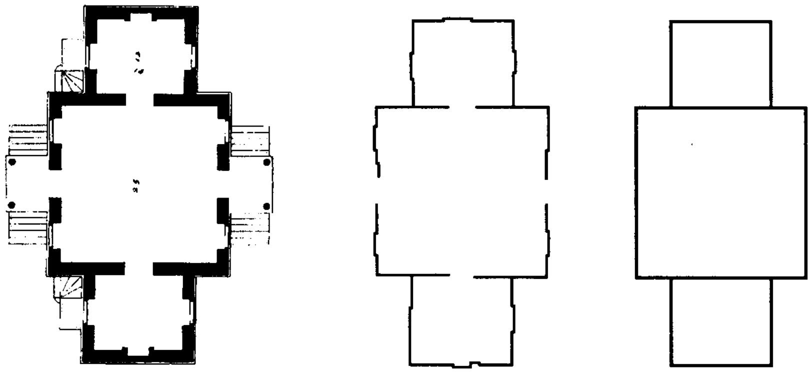


Figure 3. Skeletonization of a floor plan
left: plan; centre: the axes of its components; right: normalized skeleton

Skeletonization is achieved by first thinning the digitized image, whereby lines of unity thickness in the original digitized image (normally symbolizations of stairs, furniture, the swing of a door and similar elements) are eliminated, then filling small gaps that have emerged in the skeleton (normally openings in the boundaries of spaces or intercolumniation spaces), and finally normalizing the skeleton (on the basis of a small set of chain coding transformation rules, as in Bowie and Young (1977) so as to eliminate accidental or insignificant perturbations.

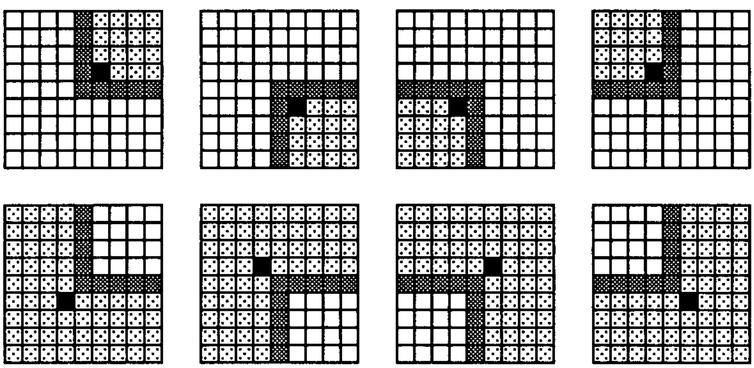


Figure 4. The typology of space corners in rectangular floor plans
top row (from left to right): types I, II, III, IV; bottom row: ¬I, ¬II, ¬III, ¬IV; dark cells represent black pixels (skeleton lines); light cells represent white pixels which belong to the space under investigation; black cells represent the corner under investigation

Line following can be used to identify boundaries of spaces in the skeletonized floor plan in the form of closed loops of line segments (Lawson and Riley 1982; Joseph 1989). An alternative, less tenacious, and conceptually more consistent alternative is to employ a variation of the technique used for the recognition of line drawings of three-dimensional scenes with trihedral objects (Waltz 1975), which also relates to the perception of illusory (subjective) contours and figures. According to this technique, recognition of spaces in a floor plan is based on (i) a typology of edge junctions (i.e., interior space corners; Figure 4), in the form of templates that are matched to the appropriate parts of the image, and (ii) the connectivity of different junction types, which allows the development of expectations concerning the type and position of the neighbours of each junction (i.e., other corners of the same space directly linked to it with edges; Figure 5) (Koutamanis 1990).

vertex	up	right	down	left
I	II,¬III	IV,¬III	*	*
II	*	III,¬IV	I,¬IV	*
III	*	*	IV,¬I	II,¬I
IV	III,¬II	*	*	I,¬II
¬I	III,¬II	III,¬IV	*	*
¬II	*	IV,¬III	IV,¬I	*
¬III	*	*	I,¬IV	I,¬II
¬IV	II,¬III	*	*	II,¬I

Figure 5. Connectivity of space corner types for rectangular floor plans

The conceptual advantage of the above technique is that, while edge following deals with the elements that *bound* a space, this technique concentrates on the space itself, in very much the same way that illusory shape perception operates. The resulting practical advantages include direct recognition of the shape of a space and relative freedom from local perturbations and missing parts in the boundary of a space. The latter can be of particular significance for the recognition of drawings of buildings with semi-unified or flow spaces.

The resulting description of a building in terms of its interior spaces forms the basis of most aspects of architectural design thinking and paves the way for the integration of the corresponding analyses and evaluations on the basis of standard design representations. This integration is traditionally one of the toughest problems for the development of coherent and comprehensive computer-aided design systems (Schmitt 1988). An example of the degree of integration allowed by automated recognition is that the description of a building's spatial articulation can be used to automatically create geometric and topological representations for the analysis and evaluation of dynamic design aspects such as the circulation of people, goods, or even information in a building (Koutamanis and Mitossi 1991). The topological representation is automatically derived from the (geometric) floor plan and correlated with programmatic requirements so as to determine the various circulation patterns (route and flow) at an abstract topological level. These patterns are then transferred back to the floor plan in order to obtain a more accurate representation of the route and its possible conflicts with the form of the building (Figure 6).

A more advanced case of analysis and evaluation is that of the formal goodness of a building, which includes automated recognition of grouping relationships between interior spaces (such as alignment, repetition and adjacency) and direct correlation of the resulting spatial configuration with architectural formal constraints. These make explicit in the drawing which elements and relationships pertain to a particular architectural formal system ('style'), such as classicism, and hence allow for an accurate and precise consideration of the design with respect to this formal system (Koutamanis 1990).

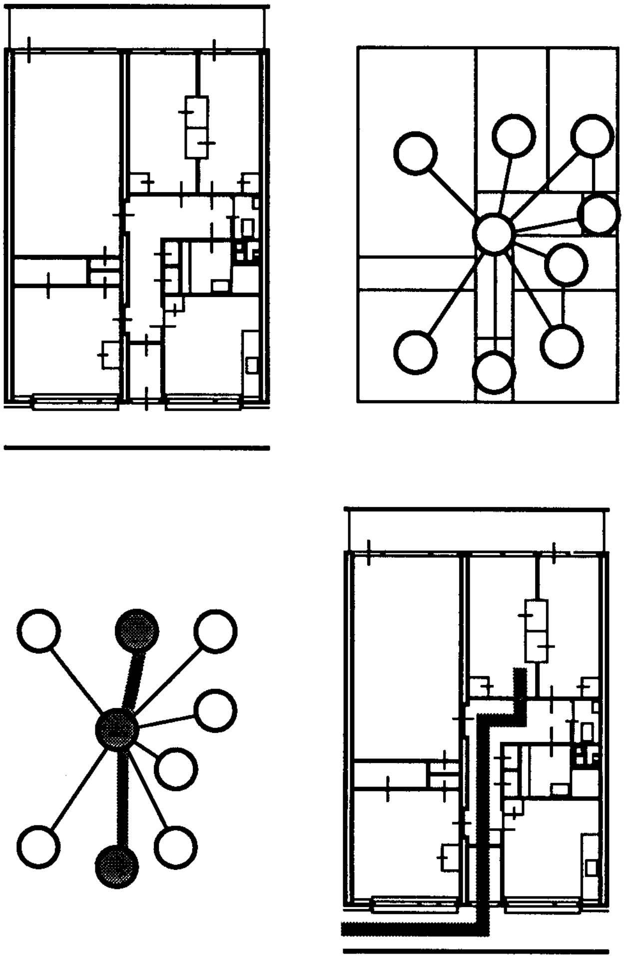


Figure 6. Transformation of a geometric design representation into a topologic one, for the analysis and evaluation of pedestrian circulation in dwellings

6. Extensions

Automated recognition of architectural drawings has obvious significance for the two problem areas we have already outlined: the transfer of existing architectural drawings to the computer, and the automated interpretation of user input to CAD systems. The transformation of the graphic elements comprising a drawing into a dual correlated network of building and spatial elements presents further possibilities which can have far-reaching extensions. One such extension concerns the use of autonomous robots in building construction.

The complexity of building tasks and the inability to use standard spatial set-ups like an assembly line on all sites suggest the potential of autonomous robots in building. Such autonomous robots can operate only if they have clear answers to what should be done and where. Architectural drawings can serve as the basis for the answer to both questions. The configuration of building elements clearly represents the final state of the building process. Moving a building component which arrives on site to its proper position involves the ability to recognize this component as part of a particular structure which has been specified somewhere in the drawings. Recognition of the building elements in a drawing can therefore provide answers concerning where something should be placed and in what relation to other objects.

The movement of a robot on a site, as well as in a finished building, represents another class of problems whose solution is at least facilitated by the recognition of the spatial elements represented in a drawing. The configuration of spaces in a floor plan obviously offers a measure by which the robot can judge the final product of its endeavours. At the same time, it offers a comprehensive spatial framework which facilitates understanding of not only how the robot can move from one space to another but also where exactly a building component should be positioned even if no related components are already in place. This spatial framework and its relationships with the building structure also facilitate the use of autonomous robots after construction, for instance in the diagnosis of structural and other faults, maintenance work and security supervision.

References

Bowie, J.E. and Young, I.T. (1977) "An Analysis Technique for Biological Shape —II", *Acta Cytologica* **21**, 455–464.

Hofer-Alfeis, J. (1987) "Automated Conversion of Existing Mechanical Engineering Drawings to CAD Data Structures: State of the Art", in K. Bø, L. Estensen, P. Falster, and E.A. Warman (eds.), *Computer Applications in Production and Engineering, CAPE'86* , North-Holland, Amsterdam, pp. 259–267.

Joseph, S.H. (1989) "Processing of Engineering Drawings for Automatic Input to CAD", *Pattern Recognition* **22**, 1–11.

Koutamanis, A. (1990) Development of a Computerized Handbook of Architectural Plans, Ph.D. Dissertation, Delft University of Technology.

Koutamanis, A. and Mitossi, V. (1991) "On the Representation and Analysis of Dynamic Aspects of Architectural Design", in F.H. Post and A.J.S. Hin (eds.), *Proceedings Second Eurographics Workshop on Visualization in Scientific Computing*, Delft.

Lawson, B.R. and Riley, J.P. (1982) "ISAAC: A Technique for the Automatic Interpretation of Spaces from Drawn Building Floor Plans", in A. Pipes (ed.), *CAD 82: 5th International Conference and Exhibition on Computers in Design Engineering*, Butterworths, Guildford, pp. 663–666.

Pavlidis, T. (1988) "Image Analysis and Graphics", in R.A. Earnshaw (ed.), *Theoretical Foundations of Computer Graphics and CAD*, Springer Verlag, Berlin, pp. 1177–1195.

Persoon, E. (1988) "A Pipelined Image Analysis System using Custom Integrated Circuits", *IEEE Transactions on PAMI* **10**, 110–116.

Schmitt, G. (1988) *Microcomputer Aided Design for Architects and Designers,* John Wiley and Sons, New York.

Waltz, D. (1975) "Understanding Line Drawings of Scenes with Shadows", in P.H. Winston (ed.), *The Psychology of Computer Vision*, McGraw-Hill, New York, pp. 19–91.