

# Resolving some Ambiguities in Real-time Design Drawing Recognition by means of a Decision Tree for Agents

ACHTEN Henri

*Department of Architecture, Building and Planning, Eindhoven University of  
Technology, The Netherlands*

**Keywords:** multi-agent system, decision tree, pattern recognition, sketch

**Abstract:** In this paper, we present a theoretical study on automated understanding of the design drawing. This can lead to design support through the natural interface of sketching. In earlier work, 24 plan-based conventions of depiction have been identified, such as grid, zone, axial system, contour, and element vocabulary. These are termed graphic units. Graphic units form a good basis for recognition of drawings as they combine shape with meaning. We present some of the theoretical questions that have to be resolved before an implementation can be made. The contribution of this paper is: (i) identification of domain knowledge which is necessary for recognition; (ii) outlining combined strategy of multi-agent systems and online recognition; (iii) functional structure for agents and their organisation to converge on sketch recognition.

## 1 GRAPHIC REPRESENTATIONS

During the design process, roughly three classes of graphic representations are utilised: the diagram, the design drawing, and the sketch. The diagram is a clear, well-structured schematic representation of some state of affairs. It is typically used in the early phase of design, very often as an analytical tool. The design drawing is a well-drafted comprehensive drawing applying the techniques of plan, section, and façade. It is typically utilized after concept design, to communicate ideas with the patron or other participants in the design process. The sketch is a quickly produced rough outline of a design idea, typically created in the early phase of design. Within each class there are many differences due to production technique, time constraints, conventions of depiction and encoding, and personal style.

The diagram and sketch are very apt for early design. They enable fast production of numerous ideas, thereby allowing the architect to engage in an iterative process using internal and external memory, reflection, exploration, and test (Goel 1995; Verstijnen 1997). In this work we are concerned with computational interpretation of such drawings. We are motivated from a *design support* and *design research*

## **Resolving some ambiguities in real-time design drawing recognition**

perspective. Much of the output in early design is produced by means of diagrams and sketches. Understanding these drawings supports early design assessment, can provide relevant knowledge, aid in information exchange, and facilitate transfer of information between design phases or applications.

In order to achieve understanding of graphic representations, we need a theory for computational interpretation of drawings. Such a theory is lacking in the design research field (Lawson and Loke 1997, 174-175; Verstijnen et al. 1998, 520; Dahl et al. 2001, 6) and in general recognition research (Tombre 1995, Mundy 1999, 23). An additional complication is the general oversight of context (Song et al. 2002) and annotations (McGown et al. 1998, 432-433), the lack of which leads to unrealistic assumptions about image recognition. As a consequence, we have to draw our foundations, evidence, and intuitions from many disciplines including design research, image and pattern recognition, handwriting recognition, artificial intelligence, and multi-agent systems. The contribution of this paper is: (i) identification of domain knowledge which is necessary for recognition; (ii) outlining combined strategy of multi-agent systems and online recognition; (iii) functional structure for agents and their organisation to converge on sketch recognition.

Design sketches have been subject of much study. Although there is now general agreement that the sketching activity is structured (Kavakli et al. 1998; Scrivener et al. 2000), there is no consensus in terminology and framework how to categorise this structuring. This seems to be mainly caused by the great variety of research questions, design domains, and agendas pursued (compare for example Goel 1995; Purcell and Gero 1998; McGown et al. 1998; Rodgers et al. 2000; Lim et al. 2003; Akin and Moustapha 2004; Tovey et al. 2003). Most authors employ their own categorisation to differentiate between kinds of drawing (e.g. Goel 1995; 128-136, Verstijnen et al. 1998; and McGown et al. 1998).

Applied research on sketch recognition in the engineering areas has focused almost exclusively of the construction or identification of three-dimensional objects based on sketches or line drawings (e.g. Jansen and Krause 1984; Cooper 1997; Tovey 1997; Parodi et al. 1998; Bimber et al. 2000). With regard to formalised content of architectural sketch drawings, most work has been done on the plan representation. (Koutamanis 1990) outlines a computational analysis of space formation in architectural plans. (Cha and Gero 1998) formalise a number of shape configurations that may occur without consideration of design content. (Do et al. 2000) and (Gross 1996) deal with graphic shorthands of for example 'table,' 'chair,' and 'house.' (Koutamanis and Mitossi 2001) focus on local and global co-ordinating devices such as 'door in wall' and 'proportion system' of Palladio. (Leclercq 2001) has implemented a system that takes sketch input and recognises grids, spaces, and functions. In earlier work, we have classified 24 conventions that are used in drawings such as 'grid,' 'zone,' 'contour,' and 'axial system' (Achten 1997).

### **1.1 Basic Assumptions**

Domain knowledge is important for drawing recognition as it informs what should be looked for. Based on the authors above, we can state that the kind of knowledge

we need combines meaning with shape. Such meaning necessarily is based on agreement or convention within a domain: it is not intrinsic to a graphic representation. In our work, we have the following basic assumptions:

- There is nothing inherently ambiguous in graphic representations. Ambiguity through multiple interpretation is what the architect does with the graphic representation, not the graphic representation itself.
- We primarily consider sketches that are made with some care for clarity (and which can thus be verified by an outside observer). Sketches that are purposefully unclear fall outside our scope. Within all possible conventions of depiction, we only look at plan representations.
- Although designers habitually reinterpret sketches, they do so in an orderly fashion and in a limited way. Reinterpretations do not wildly diverge and are relatively close to each other. The question therefore, is not to investigate why sketches are ambiguous, but rather which clues architects employ and which interpretations they allow.
- We limit our definition of computational interpretation of a sketch to instrumental meaning (what do the graphic entities map to the design task), not architectural theoretical meaning or other domains of discourse. Instrumental meaning is domain-specific and context dependent. We take the previously identified 24 graphic units as instrumental meaning in graphic representations.

## 1.2 Graphic Units

The definition of a graphic unit is: “a specified set of graphic entities and their appearance that has a generally accepted meaning within the design community” (Achten 1997, 22). The appeal to ‘generally accepted meaning,’ even though it introduces methodological problems, is necessary because meaning is in many cases the only distinguishing factor between otherwise similar shapes (e.g. “is this set of straight lines a grid or a close packing of squares,” and “is this rectangle a table or a column?”) Graphic units can be divided in structuring and descriptive graphic units.

*Structuring* graphic units build up and organise the design. Their elements only indirectly map on elements in the built environment, and they are typically left out in final documentation drawings. They are *measurement device, zone, schematic subdivision, modular field, grid, refinement grid, tartan grid, structural tartan grid, schematic axial system, axial system, proportion system, and circulation system.*

*Descriptive* graphic units map to objects in the built environment: *simple contour, contour, specified form, elaborated structural contour, complementary contours, function symbols, element vocabulary, structural element vocabulary, combinatorial element vocabulary, functional space, partitioning system, and circulation.*

Sketch recognition therefore, translates in our view to recognition of graphic units in a plan-based diagrammatic drawing or sketch which has been created with some care for clarity.

## 2 GRAPHIC UNIT RECOGNITION

Image recognition research mainly focuses on photo- or video-like sources, or aims to mimic some functionality of the human visual system. Line art or line drawings are quite distinct from these kind of sources. In particular handwriting recognition has received a lot of attention, but in the area of drawings there has not been much work. Comparatively much effort has been given to automated document management, in particular in the transfer from paper-based maps to the electronic format of GIS and CAD (Ogier et al. 1998). Other work includes the reconstruction of telephone system manhole drawings (Arias et al. 1995), conversion of machine engineering line drawings (Wenyin and Dori 1998), and symbol detection in architectural drawings (Ah-Soon and Tombre 2001). In all cases the processed sources are the precise class of design documentation drawings.

Most approaches in image recognition are built on multiple levels of (similar or mixed) recognisers that incrementally process or reason about information delivered from lower levels. A recent trend is to incorporate statistical techniques (Forsyth 1999), although there is concern to which extent this can link to domain knowledge (Mundy 1999). Questions in this area are about the often manual tuning of the system, incorporation of domain knowledge, and the problem of low performance when variations in the drawings occur. It is to be expected that this latter problem is much larger when the input material is a sketch drawing. This calls for a strategy which can deal with much uncertainty in the recognition process.

We propose to combine two techniques to tackle the question of drawing recognition: (1) multi-agent approach; and (2) online recognition. We discuss these two strategies in the context of a hypothetical drawing system that recognises graphic units. The hypothetical system has the following components: (a) drawing area; (b) drawing pen (similar to the technique reported in Cheng 2004); (c) module for tracking and segmenting strokes made by the pen; (d) multi-agent module for determining which graphic units are present in the drawing; (e) visual display for feedback of d-module.

### 2.1 Multi-agent Approach

It has become increasingly productive to recast the multiple classifier approach in terms of multi-agent systems. Although the classifiers remain the same (Ng and Singh 1998; Giacinto and Roli 2001 point out the utility of aggregating various classifiers), the approach adds to the conceptual level a more autonomous role to each classifier; it acknowledges explicitly the limited capabilities per classifier; and it realises that classifiers (in the guise of agents) should communicate with other classifiers to settle ambiguities (Vuurpijl and Schomaker 1998; van Erp et al. 2002). The parallelism inherent in multi-agent systems is another major motivation. In particular when multiple interpretations are possible, resolution critically depends on a weighed and balanced exchange of viewpoints. In sequential processing this may lead to long waiting times for a decision-module to gather all relevant evidence.

Earlier, we have established a multi-agent framework that forms the basis for building a drawing recognition system (Achten and Jessurun 2003). The functional behaviour of an agent  $A_i$  is described by its properties  $\{P_i, G_i, F_i, C_i, S_i\}$ , that are informally defined as follows.

The purpose  $P_i$  of an agent  $A_i$  is to recognise one particular graphic unit. Thus, there is a *Grid-agent*, *Zone-agent*, *Circulation System-agent*, *Simple Contour-agent*, etc.

The goal  $G_i$  of an agent  $A_i$  is to determine whether the graphic unit  $P_i$  is present in the current state of the drawing. Thus, the *Grid-agent* continuously checks for grids.

Each agent  $A_i$  employs a set of features  $F_x$  to decide if graphic unit  $P_i$  occurs. Thus, the *Grid-agent* looks for occurrences of parallel ( $F_1$ ) aligned ( $F_2$ ) straight lines ( $F_3$ ) drawn consecutively ( $F_4$ ) in the same direction ( $F_5$ ) of roughly the same length ( $F_6$ ) in two directions ( $F_7$ ) in which the parallel lines keep the same distance from each other ( $F_8$ ), and the lines in two directions overlap each other in an implied grid area ( $F_9$ ). Obviously, multiple agents can use the same features.

Each agent  $A_i$  has a set of criteria  $C_y$  which establish a relative measure  $RM_i$  of the degree to which graphic unit  $P_i$  is detected. Thus, the *Grid-agent* wants to detect at least three consecutive lines that fulfil  $F_1 - F_9$  before it considers the possibility that a grid occurs. If the number of consecutive lines increases that fulfil  $F_1 - F_9$ , the relative measure also increases according to an S-shaped curve.

Each agent  $A_i$  has a set of segmentation measures  $S_z$  to determine if it should continue tracking the current sequence of strokes graphic unit  $P_i$ . This basically defines the activation window during which the agent considers the current sequence of strokes. Thus, the *Grid-agent* stops tracking a series of consecutive lines when curved lines occur, or when new lines no longer fall within the implied grid area without extending it, and so forth.

## 2.2 Online Recognition

Online recognition means that computer interpretation takes place while the designer is drawing. In particular in the handwriting recognition area, numerous researchers opt for online recognition of text (Williams 2002 contains many examples), motivated especially by the high efficiency of the stroke direction feature (Liu et al. 2003, 2271). To the best of our knowledge, there is no online computer *interpretation* of sketches applied in architectural design, and only sporadically in engineering design (Dijk and Mayer 1997 and Qin et al. 2001 are notable exceptions). This is an omission since the creation order of strokes and their direction information can only be derived from an online process. Especially in the case of diagrams and sketches, where the appearance of elements shows high degree of variety, these are important clues to derive what is being drawn (e.g. features  $F_4$  and  $F_5$  of the Grid-agent example).

The c-module in the hypothetical sketch system creates a data stream of strokes which forms the input for the multi-agent system in the d-module. The agents continuously parse the stream in the manner described above; this is in fact their

## Resolving some ambiguities in real-time design drawing recognition

way of 'seeing' the drawing. An agent  $A_i$  annotates the drawing by placing four kinds of markers in the stream:  $M_s$ ,  $M_h$ ,  $M_e$ , and  $M_c$ . The start marker  $M_s$  designates the first stroke of a sequence that an agent is tracking. A new start marker can only be placed after the agent has placed an end marker or claim marker. The hypothesis marker  $M_h$  designates the end of a sequence while the agent is anticipating that in the near future it will be completed (in this manner, interfering non-fitting strokes can be ignored). The end marker  $M_e$  designates the last stroke of a sequence that an agent is tracking. A new end marker can only be placed after the agent has placed a start marker. The claim marker  $M_c$  is placed once the agent reaches the threshold value and a graphic unit has been identified.

At any point in the stream, any agent can place its own marker; each marker is associated with the agent that has put it there. After parsing and annotating the data stream, the stream is stored for later reference.

### 2.3 Resolving the Decision Process

Whenever an agent reaches its threshold value and places a claim, all other agents respond by polling their current relative measures  $RM_i$ . In order to converge to a decision, we adopt a decision tree developed earlier for architects to decide which graphic unit is present (Achten 2000); see Figure 1.

In the decision tree, the path ABDGK17 is the series of questions that must be answered to determine the graphic unit. These are currently stated in natural language: *A*. Is it a graphic or symbol element; *B*. Is it a closed shape or a set of one or more lines; *D*. Is it a coordinating system or not; *G*. Is it a zone, grid, or proportion system; and *K*. Is it a modular field, grid, refinement grid, tartan grid, or structural tartan grid. The leaves 1-27 are the specific graphic units.

The decision tree essentially divides graphic units in groups: text (A7); multiple shapes on building level (ABCEI); multiple shapes on element level (ABCEJ); single shapes on building level (ABCF); area structuring devices (ABDG and ABDGK); and building structuring devices (ABDH, ABDHL, ABDHM, ABDHN). The groups are distinguished from each other through the nodes *A-N*. The decision in each node to decide between one or the other further branch is made on a specific set of features for that node. Therefore, each graphic unit is characterised by a unique decision cluster of features (the aggregation of the nodes leading to that graphic unit).

At any given point during the online recognition process, it is likely that multiple agents have a relative measure  $RM_i$  which is higher than 0. In other words, features are likely not to activate one single agent (restricted to a single path in the decision tree), but multiple agents (distributed over the nodes of the whole tree). So we can view the distribution of the features over the 24 distinct paths as additional evidence for a ranking which can modify a preliminary ranking based on  $RM_i$  only. Qualitatively speaking, the path which has the greatest collection of features presents the most likely candidate. We still have to determine a quantitative means of aggregating the features along a path and balancing this measure against  $RM_i$ .

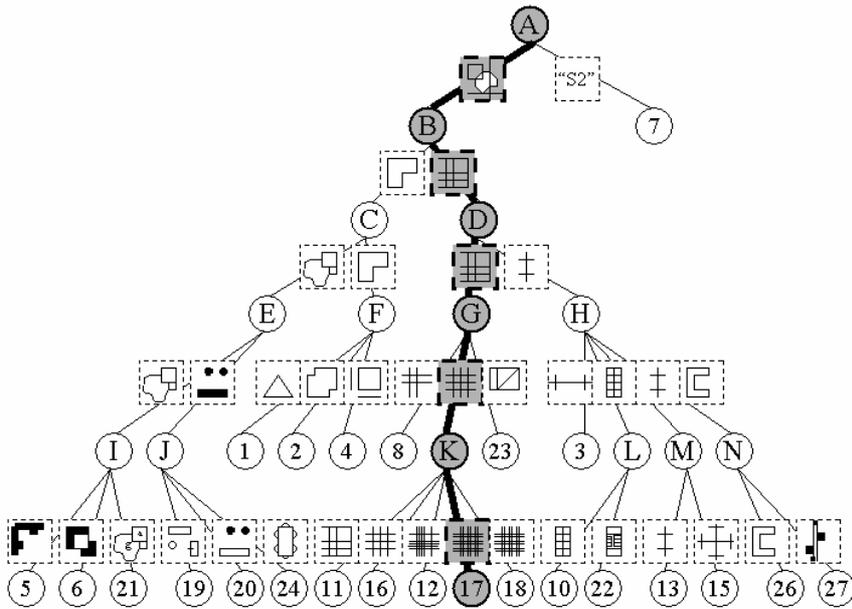


Figure 1 Decision Cluster ABDGK For Graphic Unit *Tartan Grid*

### 3 FUTURE WORK

We have outlined a multi-agent system for online recognition of graphic units in diagrammatic and sketch drawings. Empirical work is still needed to calibrate the stroke segmentation of the c-module. The utility of the decision tree is to separate the major groups, and show near the leaves how ambiguity occurs through alternative interpretations. The different depths of the decision tree give a relative indication which groups are easier to recognise than others. A major step we still have to take is the formalisation into features of the natural language questions in the nodes A-N. This also requires an additional revision of the decision tree into truly binary branches per node. Based on the resulting pool of features, we can determine which aggregation policy has the most potential.

A running implementation will provide feedback about the viability of the theoretical work. Supposing that we find some virtue in the current work, further theoretical questions then concern whether an agent should keep a record of its own performance when recognition has to take place; if it is desirable to have each agent segment the input by itself (and how to aggregate over feature activations that are built on different segmentations); and whether an agent should follow multiple tracks in the data stream.

## REFERENCES

- Achten, H.H. 1997. *Generic Representations*. PhD-diss., Eindhoven University of Technology.
- Achten, H.H. 2000. Design case retrieval by generic representations. In *Artificial intelligence in design '00*, ed. J.S. Gero: 373-392. Dordrecht: Kluwer.
- Achten, H.H., and J. Jessurun. 2003. Learning from mah jong. In *Digital design: research and practice*, ed. M.L. Chiu et al.: 115-124. Dordrecht: Kluwer.
- Ah-Soon, C., and K. Tombre. 2001. Architectural symbol recognition using a network of constraints. *Pattern Recognition Letters* 22: 231-248.
- Akin, O., and H. Moustapha 2004. Strategic use of representation in architectural massing. *Design Studies* 25(1): 31-50.
- Arias, J.F., Lai, C.P., Surya, S., Kasturi, R., and A. Chhabra. 1995. Interpretation of telephone system manhole drawings. *Pattern Recognition Letters* 16: 355-369.
- Bimber, O., Encarnação, L.M., and A. Stork. 2000. A multi-layered architecture for sketch-based interaction within virtual environments. *Computers & Graphics* 24(6): 851-867.
- Cha, M.Y., and Gero, J.S. 1998. Shape pattern recognition. In *Artificial intelligence in design '98*, ed. Gero, J.S., and F. Sudweeks: 169-187, Dordrecht: Kluwer.
- Cheng, N.Y.-W. 2004. Stroke sequence in digital sketching. In *Architecture in the Network Society [eCAADe 2004 proceedings]*, eds. B. Rüdiger, B. Tournay, and H. Orbak: 387-393. Copenhagen: eCAADe.
- Cooper, M.C. 1997. Interpreting line drawings of curved objects with tangential edges and surfaces. *Image and Vision Computing* 15: 263-276.
- Dahl, D.W., Chattopadhyay, A., and G.J. Gorn. (2001). The importance of visualization in concept design. *Design Studies* 22(1): 5-26.
- Dijk, C.G.C. van, and A.A.C. Mayer. 1997. Sketch input for conceptual surface design. *Computers in Industry* 34(1): 125-137.
- Do, E., Gross, M.D., Neiman, B., and G. Zimring (2000). Intentions in and relations among design drawings. *Design Studies* 21(5): 483-503.
- Erp, M. van, Vuurpijl, L., and L. Schomaker. 2002. An overview and comparison of voting methods for pattern recognition. In *Proceedings of IWFHR'02*, ed. Williams, A.D.: 195-200, Los Alamitos: IEEE Computer Society.
- Forsyth, D. 1999. An empirical-statistical agenda for recognition. In *Shape, contour and grouping in computer vision*, ed. Forsyth, D.A., Mundy, J.L., di Gesù, V., and R. Cipolla: 9-21, Berlin: Springer Verlag.
- Giacinto, G., and F. Roli. 2001. An approach to the automatic design of multiple classifier systems, *Pattern Recognition Letters* 22: 25-33.

- Goel, V. 1995. *Sketches of thought*, Cambridge: The MIT Press.
- Gross, M. (1996). The Electronic Cocktail Napkin - a computational environment for working with design diagrams. *Design Studies* 17(1): 53-69.
- Jansen, H., and F.-L. Krause. 1984. Interpretation of freehand drawings for mechanical design processes. *Computers & Graphics* 8(4): 351-369.
- Kavakli, M., Scrivener, S.A.R. and L.J. Ball 1998. Structure in idea sketching behaviour. *Design Studies* 19(4): 485-517.
- Koutamanis, A. (1990). *Development of a computerized handbook of architectural plans*. PhD. diss., Delft University of Technology.
- Koutamanis, A., and V. Mitossi (2001). On representation. *Design Research in the Netherlands 2000*, ed. H. Achten, B. de Vries, and J. Hennessey: 105-118, Eindhoven: Eindhoven University of Technology.
- Lawson, B. and S.H. Loke 1997. Computers, words and pictures. *Design Studies* 18(2): 171-183.
- Leclercq, P.P. 2001. Programming and assisted sketching. In *Computer aided architectural design futures 2001*, ed. B. de Vries and J.P. van Leeuwen, and H.H. Achten:15-31. Dordrecht: Kluwer.
- Lim, S., Qin, S.F., Prieto, P., Wright, D., and J. Shackleton. 2003. A study of sketching behaviour to support free-form surface modelling from on-line sketching. *Design Studies* 25(4): 393-413.
- Liu, C.L., Nakashima, K., Sako, H., and H. Fujisawa 2003. Handwritten digit recognition: benchmarking of state-of-the-art techniques. *Pattern Recognition* 36(10): 2271-2285.
- McGown, A., Green, G., and P.A. Rodgers 1998. Visible ideas: information patterns of conceptual sketch activity. *Design Studies* 19(4): 431-453.
- Mundy, J. 1999. A formal-physical agenda for recognition. *Shape, contour and grouping in computer Vision*, ed. Forsyth, D.A., Mundy, J.L., di Gesù, V., and R. Cipolla: 22-27, Berlin: Springer Verlag.
- Ng, G.S., and H. Singh. 1998. Democracy in pattern classification: combinations of votes in various pattern classifiers, *Artificial Intelligence in Engineering* 12: 189-204.
- Ogier, J.M., Mullot, R., Labiche, J., and Y. Lecourtier. 1998. Multilevel approach and distributed consistency for technical map interpretation: Application to cadastral maps. *Computer Vision and Image Understanding* 70(3): 438-451.
- Parodi, P., Lancewicki, R., Vjih, A., and J.K. Tsotsos. 1998. Empirically-derived estimates of the complexity of labelling line drawings of polyhedral scenes. *Artificial Intelligence* 105: 47-75.
- Purcell, A.T., and J.S. Gero 1998. Drawings and the design process. *Design Studies* 19(4): 389-430.

## Resolving some ambiguities in real-time design drawing recognition

- Qin, S.-F., Wright, D.K., and I.N. Jordanov. 2001. On-line segmentation of freehand sketches by knowledge-based nonlinear thresholding operations, *Pattern Recognition* 34(10): 1885-1893.
- Rodgers, P.A., G. Green, and A. McGown. 2000. Using concept sketches to track design progress. *Design Studies* 21(5): 451-464.
- Scrivener, S.A.R., Ball, L.J. and W. Tseng. 2000. Uncertainty and sketching behaviour. *Design Studies* 21(5): 465-481.
- Song, X.B., Abu-Mostafa, Y., Sill, J., Kasdan, H., and M. Pavel. 2002. Robust image recognition by fusion of contextual information. *Information Fusion* 3(4): 277-287.
- Tombre, K. 1995. Graphics recognition – general context and challenges. *Pattern Recognition Letters* 16(9): 883-891.
- Tovey, M. 1997. Styling and design: intuition and analysis in industrial design. *Design Studies* 18(1): 5-31.
- Tovey, M., Porter, S., and R. Newman. 2003. Sketching, concept development and automotive design. *Design Studies* 24(2): 135-153.
- Verstijnen, I.M. 1997. *Sketches of creative discovery*. PhD. diss, Delft University of Technology.
- Verstijnen, I.M., J.M. Hennessey, C. van Leeuwen, R. Hamel, and G. Goldschmidt. 1998. Sketching and creative discovery. *Design Studies* 19(4): 519-546.
- Vuurpijl, L., and L. Schomaker, L. 1998. Multiple-agent architectures for the classification of handwritten text. In *Proceedings of IWFHR6*, 335-346. Taejon, Korea.
- Wenyin, L., and D. Dori. 1998. A generic integrated line detection algorithm and its object-process specification. *Computer Vision and Image Understanding* 70(3): 420-437.
- Williams, A.D. (ed.). 2002. *Proceedings Eighth International Workshop on Frontiers in Handwriting Recognition*, Los Alamitos: IEEE Computer Society.