Matching Coarse Object and Model Descriptions for Generic Recognition

Robert Bergevin Martin D. Levine

McGill Research Center for Intelligent Machines
3480 University Street, Montréal, Québec Canada H3A 2A7

Abstract

We describe an approach by which coarse part-based three-dimensional object descriptions are matched to object models in a similar representation. In the context of a large number of known object models, it is necessary to remove, as early as possible, a large proportion of these models from consideration. Indexing the database of models according to the coarse shape of the object parts permits such an initial filtering. Matching the remaining models to a computed description is realized by means of part and connection assignment trees. A proposed measure for the quality of the match reflects both the similarities and differences between the object and the model. We have adopted this approach for the final matching stage of PARVO, a generic object recognition system. In PARVO, an object is recognized as being a member of a class defined by qualitative geometrical properties and topological structure. Similarity between different 3D views of any member of a given class is a result of building coarse part-based structural descriptions from limited information. Experiments with complete and imperfect line drawings demonstrate that the system has reached the type of competence expected.

Résumé

Nous décrivons une approche pour apparier une description grossière d'un objet basée sur ses parties constituantes à des modèles d'objets similairement représentés. Dans le cas où le nombre de modèles connus est très élevé, il est important de limiter le nombre de ceux qui seront traités plus en détail. Une indexation de la base de données sur la base des parties constituantes des objets permet ce filtrage initial. L'appariement des modèles restants à la description calculée se fait au moyen d'arbres d'assignation pour les parties et leurs connexions. Une mesure est proposée qui reflète à la fois les similitudes et différences entre un objet et un modèle appariés. Nous utilisons cette approche lors de l'étape finale du traitement de PARVO, un système de reconnaissance générique des objets. PARVO reconnait qu'un objet donné fait partie d'une classe définie par des propriétés géométriques et topologiques qualitatives. La similitude entre différents points de vue d'un objet résulte du calcul de descriptions grossières basées sur les parties constituantes des objets à partir d'une information limitée. Des expériences utilisant des dessins au trait souffrant de diverses imperfections ont permis de démontrer que le système a atteint le type de compétence attendu.

Keywords: generic object recognition, coarse descriptions, labelled graphs, partial matching, qualitative similarity

1 Introduction

Specific model-based object recognition [1, 2] relies on accurate prior knowledge for the possible appearances of objects in images. In contrast, generic object recognition systems are based on coarse, qualitative models representing classes of objects. For this reason, instantiation of a model on the basis of only a few low-level features [3] is not possible for the latter. That is, a generic object description must be computed bottom-up, before any matching can take place between the image and model structures.

A computer vision system has been built that computes such type of descriptions from single view 2D *line drawings* of 3D objects (see Figure 1). It is called PARVO [4, 5, 6] and is based on the *Recognition by Components* (RBC) theory of human image understanding [7]. Any object within the scope of PARVO must satisfy an identified set of assumptions related to the generic shape of its parts, the generic modes by which these are spatially connected to each other, and the generic properties of the viewpoints from which objects are observed.

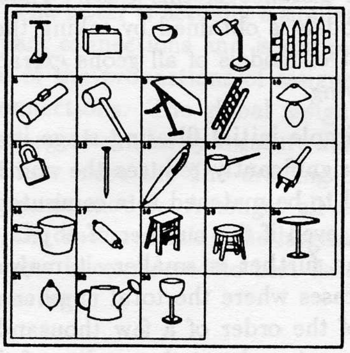

Figure 1: Typical drawings analyzed by PARVO. These drawings have been adapted from both a database of objects available in the psychology literature [8], and a set of drawings used by Biederman in his experiments [7].

In PARVO, both the computed descriptions and the object models are represented as a graph in which each node corresponds to a single 3D part and each link between two nodes refers to a 3D connection between the two associated parts. A geon label [7] is stored at each node. It categorizes the qualitative shape of the part and a symbolic aspect ratio which determines whether the part is approximately *elongated*, *flat*, or *blob-like*. The following information is associated with each link: (i) the connection mode (***end-to-end*** or ***end-to-side*** with the estimated approximate position of the join along the part axes) between the two parts, and (ii) their relative sizes.

The graph corresponding to a particular model contains all the parts of the object and all the connections they form (see Figure 2). As such, our models are actually coarse object-centered descriptions. However, there is no global coordinate system associated with them. That is, the qualitative spatial structure of the object is only implicitly contained in the explicit set of all two-part connections. This is an absolute requirement in order to tolerate partial occlusions and some noise in the line drawings. Nevertheless, this representation is not limited to 2D information since it contains aspects of the 3D structure of objects in terms of the volumetric shape of the parts and their spatial connections. Note that the actual graph corresponding to the description computed from a single view of the object might contain a smaller number of (visible) parts and connections. In fact, only the properly labelled parts and connections are included in the computed description graph. The reader is referred to earlier publications [9, 10, 4, 5, 6, 11] for a more detailed account of RBC and PARVO. There, the requirements of the primal access task addressed by RBC, its relationship to generic object recogntion, and the constraints it imposes on PARVO are described and justified. In particular, the proposed geon-based shape representation and the classes of non-accidental and viewpoint-invariant features from which it is derived are discussed. Also, the specific details about how PARVO actually computes coarse 3D descriptions of objects from single 2D line drawing views are presented in these previous papers.

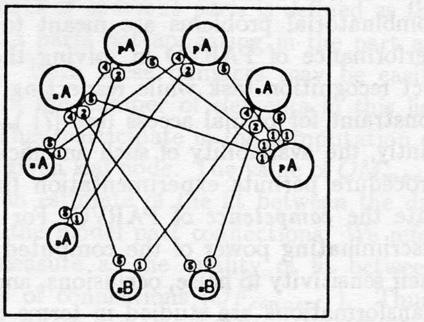

Figure 2: Graphical representation of a STEPLADDER model. At each node is stored the *aspect ratio* (s: stick, p: plate) and the *geon label* of a part (here, A or B). To each link is associated a symbolic *connection mode* (here, numbers in the circles), and the relative sizes of the two connected parts.

PARVO assumes only twenty-two different geon labels [1], three aspect ratio categories, two connection modes, five coarse join positions, and five relative size categories (see [9] for details). Thus, there are sixty-six possible labels for each part node and fifty possible labels for each connection link. The number of possible descriptions, even for an object having a relatively small number of parts and connections, is obviously very high.

The specific problem we address in this paper is the object identification stage of PARVO. This stage follows the computation of a coarse 3D description of an object from its 2D line drawing. It consists of finding, among all known object mod-

---

[1] The original set of thirty-six has been recently reduced to twenty-four by Biederman. In PARVO, we actually consider two of them to be redundant [9].

els, the one which is the most similar to the computed description. This stage involves at least two computationally intensive procedures. First of all, given that there may be a large number of object models known to the system (thousands, say), the task of matching a computed description to each model is very demanding. As a possible solution to this problem, we describe an indexing scheme of the database of models based on the geon categorization of the parts. Secondly, matching any given model to the computed description involves finding the best possible assignment of the parts and connections in this description to those in the model. This is done by building so-called assignment trees of parts and connections. A definition of the compatibility between parts and connections in terms of their coarse labels gives rise to an efficient pruning of these trees. The proposed solutions for these two combinatorial problems are meant to improve the performance of PARVO in solving the generic object recognition task while respecting the efficiency constraint for primal access (see [7]). More importantly, the availability of such an efficient matching procedure permits experimentation meant to evaluate the *competence* of PARVO. For instance, the discriminating power of the computed descriptions, their sensitivity to noise, occlusions, and rigid spatial transformations are studied in terms of the impact on recognition.

In this paper, we describe our proposed solution to the two combinatorial problems arising at the object identification stage. We also present the experimental results obtained by PARVO in matching coarse descriptions of line drawings to models.

2 An approach to determining the best object model

Given an attributed-graph corresponding to a computed 3D object description, the object identification (or matching) task consists of determining, among the graphs of all known object models, the one most similar to the computed one.

2.1 Initial filtering of the database of models

The first problem is the large size of the model database. The purpose of indexing the memory of object models is to permit a fast reduction of the search space associated with the matching task. In PARVO, only the models which have, among their parts, *all* the geons extracted from the line drawing are actually matched to the computed description. In order to perform *initial filtering* of the database, PARVO keeps, for each of the twenty-two geon classes, a list of the models which contain at least one instance of this class. Thus, the filtered list of models is obtained by taking the intersection of the list of models of all geons extracted from the line drawing.

This simple initial filtering stage implemented in PARVO significantly reduces the number of models that need to be matched to a computed description. However, even if the number of object models to be considered further is smaller, it may still be quite large in cases where the total number of known objects is of the order of a few thousand. Therefore, PARVO must evaluate the quality of the match between each object model that has survived the initial filtering stage and the extracted line drawing description. In this way, the best matching model will be accepted as long as its quality of match (QM_{model}) is sufficiently high. This QM_{model} could be determined independently in parallel for each object model. However, in our experiments, matching object descriptions to models is performed sequentially. That is, a measure of similarity (see below) is computed for each model which has passed the initial filtering stage and the given line drawing object description.

The measure of similarity is computed after the best possible assignment (pairing) of object parts to model parts is obtained. In principle, all possible assignments should be generated and compared in order to find the best one. However, in PARVO, an object part can be assigned to (paired with) a model part if and only if they have the same geon label and aspect ratio. In such a case, we say that the two parts are *compatible*. An acceptable *global* assignment consists of a set of object/model pairs of parts where each object (model) part is assigned to at most one compatible model (object) part. There may be more than one solution to this *part assignment* problem for a given model since more than one part may have the same geon label and aspect ratio. However, the overall value of QM_{model} will generally be different for each solution since the connections involving these parts will usually differ. Our approach is to generate all acceptable global assignments for a given model and select the one for which QM_{model} is maximal. That is, for each such global assignment of compatible parts, a measure of the similarity between the two descriptions is computed. This similarity measure depends on (i) what proportion of the object parts are assigned to model parts, (ii) what proportion of model parts are assigned to object parts, and (iii) how similar the connections between each pair of object parts are to the connections between the pair of model parts to which they are assigned. The latter is actually a similarity measure for the part connections and is computed as a weighted sum of the proportions of assigned object and model connections. The global assignment resulting in the largest similarity measure value is our estimation of how close the line drawing of an object is to a generic model. Amongst all models in the database, the one producing the largest similarity measure value is considered to best describe the object line drawing.

2.2 Generating the assignment trees of compatible parts and connections

The generation of all global assignments may be conceptualized as a process in which a *part assignment tree* is built. Each *leaf* node of this tree contains a solution to the part assignment problem. We recall that, for the purpose of building the part assignment tree, a model part and a data part are said to be *compatible* whenever they have the same *geon label* and *aspect ratio*. The labels are of course one of the twenty-two possible geons and the aspect ratio assumes one of the three symbolic values mentioned earlier. By imposing the restriction that a model part may be assigned to a data part if and only if the two are compatible, the part assignment tree is effectively pruned and the number of leaf nodes at which QM_{model} is computed is significantly reduced (see [9]).

The part assignment tree has at most $D+1$ *levels*, where D is the number of data parts. Each level, except the last, is associated with a unique datum part. Starting with a single root node at level one, the tree is built by creating for each node at a given level one child node for each model part which is compatible with the datum part at this level. In the situation where no model part is compatible with the datum part at a given level $n \leq D$, this datum part is discarded and the number of levels of the tree is reduced by one. Otherwise, the *part assignment list* of each node at level $n+1$ is obtained by appending a $(DP\ MP)$ pair to the *part assignment list* of the parent node at level n. In this appended pair, DP is the datum part at level n and MP is one of the model parts compatible with DP. Of course, the part assignment list at the root node is empty. The fact that each model part may be assigned to at most one datum part also gives rise to an additional pruning of the tree. This certainly makes sense if we consider that only labelled parts participate in the object matching stage. After the tree is generated, each leaf node contains a part assignment list of length $l \leq D$ for which QM_{model} is computed.

2.3 Computing the quality of match at leaf nodes

The number of *missing parts* at a given leaf node is defined as the number of model parts not appearing in the *part assignment list* at this node. Similarly, the number of *spurious parts* is defined as the number of data parts not appearing in the part assignment list. Both these numbers may be easily computed from the number of elements in this list (see below). They participate in the computation of QM_{model} at a given leaf node. The value of QM_{model} also includes an estimate of the fit between the datum part and the model part connections. We refer to this last measure as the quality of fit between these two sets of connections ($QF_{connect}$). Thus, $QF_{connect}$ is only one aspect of QM_{model} . The latter is the ultimate measure by which the PARVO system determines the identity of the viewed object.

Computing $QF_{connect}$ involves the construction of a *connection assignment tree* for each $(DP\ MP)$ pair in the part assignment list at a leaf node of the part assignment tree. These trees are built in a similar fashion to the part assignment trees. Each datum part connection DPC that the datum part DP from the part assignment list pair $(DP\ MP)$ makes with another datum part DP' may participate in the generation of the tree. In this context, a model part connection MPC for which the model part MP makes a connection with another model part MP' is said to be *compatible* with the datum part connection DPC whenever the pair $(DP'\ MP')$ also belongs to the part assignment list. In other words, the two datum parts DP, DP' of a datum part connection DPC which is compatible with a model part connection MPC must be, respectively, compatible with the two model parts MP, MP' of the latter.

The *connection assignment tree* is built by adding compatible pairs of datum and model part connections at each level. Once all the connection assignment lists are generated, $QF_{connect}$ is computed at each node of the tree and the highest value is selected as the quality of fit of the connections for the given (*DP MP*) pair of matched parts.

The number of ***spurious connections*** and ***missing connections*** participate in the computation of $QF_{connect}$ (see below). These are defined in a similar fashion to the ***spurious parts*** and the ***missing parts***, respectively. In addition, the similarity of the ***connection modes*** of each pair of connections in the assignment list and the similarity of the ***relative size*** of their parts are taken into account in the computation of $QF_{connect}$.

Let us call *dp* the number of parts in the description, *op* the number of parts in the object model, and *mp* the number of matched parts in the assignment list at a given leaf node. The final values of both $QF_{connect}$ and QM_{model} are obtained as a weighted sum of their constituent measures. More precisely, the quality of match for a given assignment of model parts to data parts is:

$$QM_{model} = \max\ (0.0, (1.0 - f(MP, SP, CM)))$$

where

$$f = (w_1 MP + w_2 SP + w_3 CM),$$

MP is the number of missing parts:

$$MP = \frac{dp - mp}{dp}$$

SP is the number of spurious parts:

$$SP = \frac{op - mp}{op}$$

and CM (defined below) is the ***connection-mismatch*** value. It embodies the minimum discrepancy, for all part-connection assignments between the data and model part-connections.

Let us call *dpc* the number of part connections in the description, *opc* the number of part connections in the object model, and *mpc* the number of matched part connections in the assignment list at a given leaf node. Thus MC, the number of missing connections, is:

$$MC = \frac{dpc - mpc}{dpc}$$

Similarly SC, the number of spurious connections, is:

$$SC = \frac{opc - mpc}{opc}$$

Finally, let us call *mf* the number of matching "from" labels, *mt* the number of matching "to" labels, and *rs* the number of matching "relative-size" labels. Thus MM, which measures how well the modes of the matched connections fit, is:

$$MM = \frac{2mpc - mf - mt}{2mpc}$$

Also MR, which determines how well the relative sizes of the connected parts fit, is:

$$MR = \frac{mpc - mrs}{mpc}$$

Given these definitions, CM, the connection-mismatch becomes:

$$CM = \max\ (0.0, (1.0 - QF_{connect}))$$

where $QF_{connect}$ is computed as follows:

$$QF_{connect} = 1.0 - \min\ g(MC, SC, MM, MR)$$

with

$$g = (w_4 MC + w_5 SC + w_6 MM + w_7 MR)$$

and the minimum is taken over all the part-connection assignments.

Presently, all weights w_i are equal. We have found that this produces a good discrimination between the models. Moreover, the QM_{model} values obtained are intuitively satisfying. This is not true if normalized weights accord more importance to missing parts than to spurious ones.

The quality of match for each model that has survived the initial filtering step is taken as the highest value of QM_{model} among the assignment lists in its part assignment tree. The model having the highest quality of match is then identified as the viewed object. For most objects and views, only one best model should result from this computation. However, it is possible that similar objects or specific views may give rise to ambiguities where no single model clearly matches the viewed object better than any other model in the database. In this situation, a finer discrimination between the models could be obtained by considering some ***global*** relations involving three or more parts of the object [9].

3 Experimental results

In this section, we show object recognition results obtained by analysing both complete and imperfect line drawings. In particular, each line drawing in Figure 1 was analyzed by PARVO. Their successful description and recognition by PARVO shows that many objects of significant variety and complexity are within its scope.

Line drawings representing alternate 3D views of objects were also analyzed. Their successful recognition shows the relative insensitivity of PARVO to 3D rigid transformations between the object and the viewer, as long as the resulting drawing respects our assumptions. Similarly, 2D transformed versions of line drawings were processed by PARVO. The depicted objects were also recognized as efficiently as the originals.

Finally, a series of experiments was performed in which a line drawing was contaminated by various imperfections. Missing and spurious features at the different levels in the hierarchy of representations are shown to have "structurally local" effects. However, the lack of feedback mechanisms between the different modules of PARVO inhibits the introduction of backtracking strategies. As a result, the effect of imperfections is always propagated to the higher levels until they are invalidated.

3.1 Complete line drawings

Three main aspects need to be addressed. First of all, the sensitivity of the initial filtering stage must be evaluated. For instance, should the aspect ratio be added to the geon label in order to prune inappropriate candidate models ? Secondly, the contention of the RBC theory (summarized by the principle of componential recovery [7]) according to which a small subset of its parts is sufficient to recognize an object must be tested. Finally, the competence of PARVO in recognizing line drawings representing 2D or 3D rigid transformations of objects needed to be demonstrated.

3.1.1 Initial filtering

Initial filtering aims at removing from the database the candidate models not satisfying simple criteria (see Section 2.1). Its advantage is that it can be performed very efficiently by indexing the database of models on the basis of their properties. For instance, one could keep various lists of objects possessing one or many properties in common. There is clearly a trade-off between the complexity added to the database and its organization and the greater search efficiency obtained. PARVO uses geon-indexing in its database organization and performs an initial filtering of the object models.

With such a scheme, a single spurious part (wrong geon label) in a line drawing description could filter out from consideration the right object model to which it should be associated. Fortunately, it is more likely that PARVO will *reject* an imperfect part (one which does not satisfy one or another of our assumptions) than produce a wrong geon label. Moreover, if ever a part would end up with a wrong geon label, it is reasonable to expect that no other spurious model would produce a high enough score when matched to the inexact description, due to its inappropriate part connections, for instance. This is exactly what happens for some objects in Figure 1, as will be shown below.

Another question is whether we should take the aspect ratio into account at the initial filtering stage. Since it is quantified into three classes, the number of lists to keep would be three times as much, but this is still a small number since there are few geon classes anyway. We have found that the aspect ratio is not so easy to capture from single projected views because of possible foreshortening. Moreover, it is hard to associate any kind of confidence value with this measure since PARVO has no explicit understanding of the spatial orientation of a part. For these reasons, we have limited the use of the aspect ratio to the evaluation of the compatibility of object and model parts to be paired.

Figure 3 presents the final matching scores from experiments in which a description of each of the twenty-three line drawings of Figure 1 was computed by PARVO and compared to each of the corresponding twenty-three object models which form the database. In the figure, the radius of a black disk represents the score (normalized from 0 to r_{max}) obtained by matching an object description computed from a line drawing (one column entry) to a model (one row entry). The corresponding object models are ordered from top to bottom in the figure. Thus, a perfect description and recognition of all objects would produce only a diagonal of maximal radius disks.

All objects in Figure 1 were segmented by PARVO and the computed descriptions were mostly complete and correct. Out of the twenty-three drawings, fifteen were recognized unambiguously (the ones with no vertical line in Figure 3) [9]. Among the other eight, two obtained perfect scores for two identical models (14,16). Four of the remaining six have part structures violating some of our assumptions (2,11,12,15). As such, they are not totally acceptable to PARVO. Thus, only the *ironing board* and *ladder* drawings (9 and 10, respectively) presented a real problem. For both of them, a wrong geon label

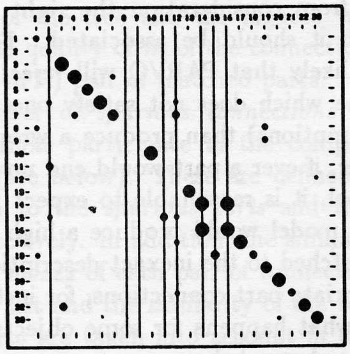

Figure 3: Matching scores for complete line drawings. The larger the radius of a disk, the higher the score obtained by matching the description of a line drawing object (column) to an object model (row).

arose from an inaccurately drawn part. This illustrates the sensitivity of the geon labelling stage of PARVO to small or elongated thin parts.

Other experiments [9] have shown that the scores of the spurious matches (not on the diagonal) are higher when the compatibility of matched parts does not take into account the aspect ratio. The importance of this aspect ratio results from the fact that many part descriptions differ only by these ratios. For instance, the two parts of the seventh drawing (the *hammer*) have the same geon label but different aspect ratios. Nevertheless, the scores on the main diagonal are not adversely affected by the absence of the aspect ratio.

If no initial filtering is performed, all models are matched to each computed object description. In general, the spurious models produce very low or null scores. Moreover, the incorrectly labelled parts still reduce the scores of the correct models, without filtering them out this time. Thus, the scores are very similar to the ones in Figure 3. However, the overall computation time is increased in such a case.

Also, an experiment in which the *geon label* was not taken into account in comparing the structures of objects and models (each line drawing part was considered compatible with every model part as long as it had the same aspect ratio) produced many more spurious matches [9]. Their scores were usually low. Overall, only four additional drawings did not produce the highest score for the correct model. A final experiment, in which even the aspect ratios were not taken into account (only the part connection structures matter), produced very poor discrimination between the object models.

3.1.2 Matching partial descriptions

A major prediction of the RBC theory is that an object may be recognized whenever a small number of its parts is identified together with the connections they make with each other. We tested this prediction by matching only a fixed subset of the parts extracted from the line drawing.

Our original experiments (Figure 3) were done with this number of parts fixed at five. For all but the most complex objects this meant that all parts participated to the matching process. We repeated the experiment with the number of matched object parts successively taking the values four, three, and two [2]. The number of drawings which are *not* associated with the correct model is the same with five and four parts. This number only increases by two when the three largest parts are matched. However, only four models *are* recognized when the matching includes only two parts.

3.1.3 Viewpoint independence

Two additional experiments are reported in [9] to support our contention that PARVO satisfies the constraints of the primal access task. In the first one, a line drawing is used which represents a *different view* of the DESK LAMP model than that shown in the fourth drawing in Figure 1. If the drawing respects our assumptions (see [9]), the parts extracted and the description graph produced will be very similar to the original one. Note that once the generic class to which an object belongs is identified, it might also be possible to use the information present in the drawing self-occlusions (at part connections, for instance) to infer the qualitative pose or attitude of the depicted object. This extension is not yet implemented in PARVO.

The second experiment involves a 2D rigidly transformed (rotated, translated, and scaled) version of the *desk lamp* line drawing. Here, the visible structure is exactly the same as the original, relative to a transformed image coordinate frame. Again, the resulting description graph is identical to the original one. Thus, PARVO does not recognize objects more easily in stable poses. This contrasts with the so-called "familiarity factor" or "priming effect" which is known to exist in human image understanding [12] . An explanation for this "lack of

---

[2] In these experiments, the largest parts are selected first. The computed area of the cross-section times the length of the projected axis provides an approximate measure of the volume occupied by the part. This measure is also used in computing the relative sizes of parts in a connection.

an effect" is that PARVO relies exclusively on an object-centered description. It would be very easy to add a viewer-centered component to the models in order to enforce expectations about the pose of the most common objects (using top/bottom and left/right descriptors, say).

3.2 Imperfect line drawings

Two *principles* summarize the competence of PARVO in analyzing imperfect line drawings. First of all, the adverse effect of either introducing a spurious feature or removing a good one, at any level in the hierarchy of representations, is always limited to the object substructure to which it belongs. In terms of the description graph, the probable result is the removal of only one or a few connected part nodes. Secondly, the strictly bottom-up organization of PARVO, which could eventually be augmented with local feedback paths, propagates a missing or spurious feature upwards. Fortunately, a resulting wrong substructure (e.g. an object face or part) will most probably violate some of our assumptions and be rejected by PARVO at a higher level. This is why some spurious elements could end up having no effect at all on the final description. In the same vein, the higher the level at which an element is missing, the easier it would be to recover from it, if this were attempted. Both these principles are closely related to the robustness of PARVO with respect to imperfect line drawings.

In [9] these two principles are illustrated by examples from the analysis of line drawings having either missing or spurious elements at the different levels in the hierarchy of representations. These two types of imperfections mimic the effect of both noise and occlusions in the line drawings.

4 Conclusions

We have presented an approach by which coarse descriptions computed by PARVO, a generic object recognition system, may be matched to attributed-graph models. Initial filtering on the indexed database of models is performed prior to matching the models to the computed description. A part assignment tree is computed for each pair of computed description and remaining model. A connection assignment subtree is also built for each compatible pair of object and model parts in that main tree. A simple measure for the quality of the match was proposed that reflects both the similarities and differences between an object and a model.

Our experiments with complete and imperfect line drawings have demonstrated that PARVO has reached the type of competence expected. Many objects of significant variety and complexity can be described in a totally bottom-up fashion. Their subsequent recognition is obtained by comparing the computed descriptions to coarse qualitative models. The importance of associating a geon label and an aspect ratio to each part in the description has been illustrated. Moreover, the contention of the RBC theory that only a few parts and connections are sufficient for recognition has been confirmed with our set of representative man-made objects. Finally, the constraints imposed by our assumptions have been shown not to limit the ability of PARVO to perform generic object recognition, while displaying the necessary robustness and flexibility.

In future, a larger database would permit a better evaluation of the discriminating power of our representations and measures. As a result, some refinements to the computed and stored (model) descriptions could be proposed. For instance, we have already realized that the matching process, and especially its initial filtering stage, is very sensitive to a wrongly labelled part. More robust model descriptions, in which a few similar geons would be associated with the most ambiguous parts (e.g. the smallest ones), have been proposed as a possible solution to this problem. In the longer term, another possible extension to PARVO would be to include functional knowledge as a way to group similar objects into common categories. For instance, chairs with different structures (parts and connections) could be assigned to a single category on the basis of such higher level cognitive processes.

Acknowledgements

The authors would like to thank I. Biederman for helpful discussions on his human image understanding theory. This research was partially supported by an NSERC Postgraduate Scholarship to R. Bergevin and the NSERC Grant A1733 to M.D. Levine. M.D. Levine would like to thank the Canadian Institute for Advanced Research for its support.

References

[1] T.O. Binford. Survey of model-based image analysis systems. *The Inter. J. of Robotics Res.*, 1(1):18–64, Spring 1982.

[2] P.J. Besl, R.C. Jain. Three-dimensional object recognition. *Computing Surveys*, 17(1):75–145, March 1985.

[3] D.G. Lowe. Perceptual organization and visual recognition. Technical Report STAN-CS-84-1020, Department of Computer Science, Stanford University, Stanford, CA, September 1984.

[4] R. Bergevin, M.D. Levine. Generic object recognition: Building coarse 3d descriptions from line drawings. In *Proc. of the IEEE Comput. Soc. Workshop on Interpretation of 3D Scenes*, Austin, Texas, 27-29 November, 1989.

[5] R. Bergevin, M.D. Levine. Part-based description and recognition of objects in line drawings. In *Proc. of the Soc. of Photo-Optical Instrumentation Engineers Conf. on Comput. Vision*, Philadelphia, Pennsylvania, 5-10 November 1989.

[6] R. Bergevin, M.D. Levine. Hierarchical decomposition of objects in line drawings. In *Proc. Vision Interface '89*, London, Ontario, Canada, 19-23 June, 1989.

[7] I. Biederman. Human image understanding: Recent research and a theory. *Comput. Vision, Graphics and Image Processing*, 32(1):29–73, 1985.

[8] J.G. Snodgrass, M. Vanderwart. A Standardized Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity. *J. of Experimental Psych.: Human Learning and Memory*, 6(2):174-215, 1980.

[9] R. Bergevin. *Primal Access Recognition of Visual Objects*. PhD thesis, McGill University, Department of Electrical Engineering, October 1989.

[10] R. Bergevin, M.D. Levine. Extraction of line drawing features for object recognition. In *Proceedings of 10th Inter. Conf. on Pattern Recognition*, Atlantic City, New Jersey, USA, June 1990.

[11] R. Bergevin, M.D. Levine. Recognition of 3-d objects in 2-d line drawings: An approach based on geons. Technical Report TR-CIM-88-24, McGill Research Center for Intelligent Machines, Montréal, Québec, Canada, November 1988.

[12] I. Biederman. Recognition-by-components: A theory of human image understanding. *Psychological Rev.*, 94(2):115–147, 1987.