〔公　開〕

TR－C－0093

What You Say Is What You See

－Interactive Generation, Manipulation and Modification of 3－D Shapes Based on Verbal Descriptions－

ジュリ　ティヘリノ　　安部　伸治　　宮里　勉
Juri　TIJERINO　　Shinji ABE　　Tsutomu MIYASATO

１９９３　１２．２８

ＡＴＲ通信システム研究所

What You Say Is What You See
—Interactive Generation, Manipulation and Modification of 3-D Shapes Based on Verbal Descriptions—

Yuri A. TIJERINO, Shinji ABE, Tsutomu MIYASATO, and Fumio KISHINO

ATR Communication Systems Research Laboratories
2-2 Hikaridai, Seikacho, Sorakugun, Kyoto 619-02 Japan
E-mail: yuri@atr-sw.atr.co.jp

Abstract: The advent of virtual reality (VR) introduced a paradigm for human-to-human communication in which 3-D shapes can be manipulated in real time in a new kind of computer supported cooperative workspace (CSCW)(Takemura and Kishino, 1992). However, mere manipulation—either with 3-D input devices (e.g., the DataGlove™ [1]) or with spoken language (Mochizuki and Kishino, 1991)—does not do justice to this new paradigm, which could prove to be revolutionary for human-to-human and human-to-machine—communication . This paper discusses the possibility to provide the means for VR-based CSCW participants not only to interactively manipulate, but also to generate and modify 3-D shapes with verbal descriptions, aided with simple hand gestures. To this end, the paper also proposes a framework for interactive indexing of knowledge-level descriptions (Newell, 1982; Tijerino and Mizoguchi, 1993) of human intentions to a symbol-level representation based on deformable superquadrics (Pentland, 1986; Horikoshi and Kasahara, 1990; Terzopoulos, 1991). This framework, at least, breaks ground in integration of natural language with interactive computer graphics.

1. Introduction

Computers provide a nice medium for simulating 2-D and 3-D spaces, which people may employ to communicate with one another. Although, 2-D representations have achieved an advanced state-of-the-art—which currently may be multimedia-based representations—so far real-time communication in 3-D representations has not been exploited as one might expect. Recently, virtual reality (VR) has captured the imagination of researchers as well as of the mass media, in part, because it promises to add that one more dimension which we all seem to be so enamored of—i.e., depth—to human-to-human and human-to-machine communication. Nevertheless, research on VR environments have predominantly focused on manipulation of 3-D shapes or interactive playback of scripts, and have essentially ignored more important aspects—i.e., real-time generation and modification of those shapes—which are fundamental for VR-based communication of thoughts and ideas.

At ATR we have been developing a Virtual Teleconferencing System (Kishino, 1990) that brings participants together to a computer-generated virtual conference room, even though they might be at distant locations (See figure 1). Recently, Takemura and Kishino (1992), acknowledged the importance of this type of virtual environments for computer-supported cooperative workspaces (CSCW). However, most research efforts currently focus on issues such as real-time reconstruction of facial expressions, hand gestures of participants, manipulation of virtual objects and so on.

Though Mochizuki and Kishino (1991) recognized the need for integrating natural language commands with hand gestures for manipulation of 3-D shapes, only recently the authors have embarked on efforts to provide interactive generation or modification of such shapes (Tijerino et

1 The DataGlove™ is a trademark of VPL Corp.

al., 1993). This paper further helps to break ground on this direction, by proposing a framework for interactive indexing of knowledge-level descriptions (Newell, 1982; Tijerino and Mizoguchi, 1993) of human intentions to a symbol-level representation based on deformable superquadrics (Pentland, 1986; Horikoshi and Kasahara, 1990; Terzopoulos, 1991).

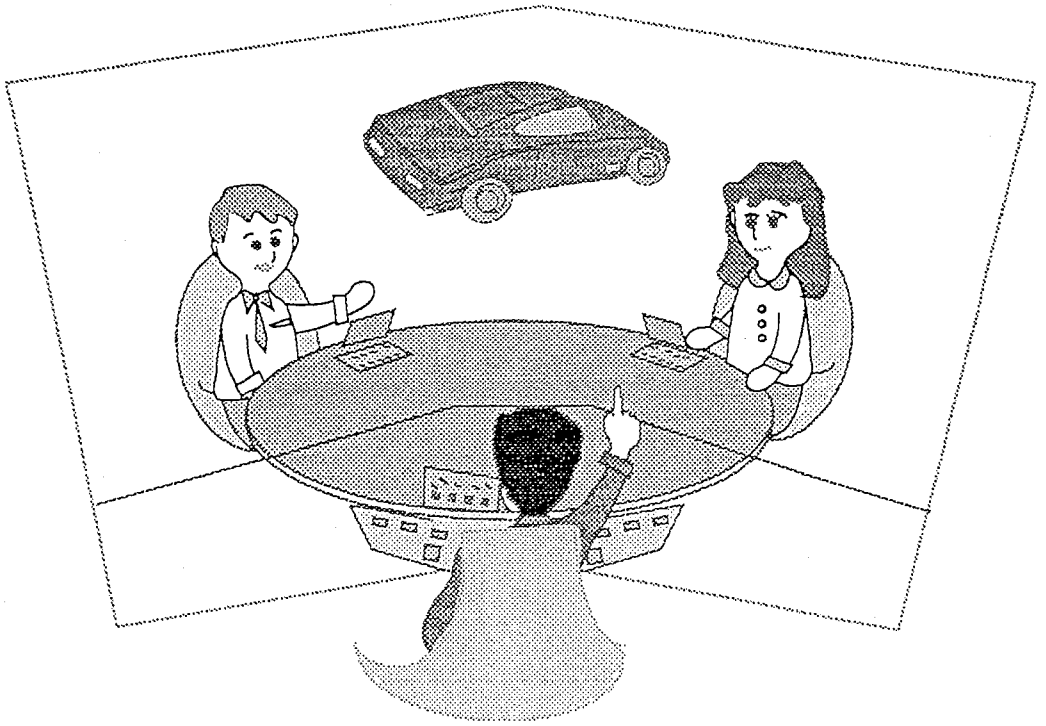

Figure 1. A virtual teleconferencing room. Participants take part in a conference, with very realistic sensations, even when they may be at different distant locations.

This framework promises to create a means in which we can communicate our intentions or thoughts to the computer with natural language and simple hand gestures to generate, manipulate and modify 3-D shapes in real time so that we can "see" what we really want to "say." We coined the term WYSIWYS (What You Say Is What You See) for these type of interactions. Though WYSIWYS has numerous applications in fields such as design, art and computer imagery, it is specially helpful for virtual space teleconferencing and CSCW, because it serves to enhance human-to-human communication.

This paper is organized in the following manner. Section 2 proposes the existence of a two-level ontology—that is, a knowledge level and a symbol level—for 3-D visual knowledge. Section 3 presents some preliminary results in an experiment to acquire knowledge-level 3-D visual concepts. Section 4 introduces a candidate symbol-level representation for symbol-level 3-D shape ontologies. Section 5 discusses the implications and applications of the framework for WYSIWYS interactions and describes an implementation effort taking place at ATR CSRL. Section 6, addresses some problems and provides some concluding remarks.

2. A Two-Level Ontology for 3-D Visual Knowledge

The word ontology means the study of existence in philosophy, but in artificial intelligence it usually means the set of most primitive terms or concepts that canonically describe some particular field of knowledge.

Though the term ontology is widely accepted in the AI community, lexicon or taxonomy may also be appropriate words to describe the concept. In this section, a two-level ontology for 3-D visual knowledge is proposed to help bridge the gap between the knowledge level and the symbol level (Newell, 1982; Tijerino and Mizoguchi, 1993). The knowledge level corresponds to concepts humans apply when describing 3-D visual knowledge. The symbol level corresponds to the symbolic representation or mechanism that operationalizes the descriptions in a visual manner, in this case, 3-D shapes.

The existence of a knowledge level was first proposed by Newell (1982), but was meant to explain how knowledge engineers interpret a particular domain expert's knowledge and translates it to a symbolic level in which a computer might make use of that knowledge to accomplish some task. However, only recently the knowledge-based systems research community has benefited from knowledge-level ontologies for knowledge-based system construction (e.g.; Neches et a., 1991; Gruber, T., 1992; Steels, L., 1992; Tijerino and Mizoguchi, 1993). In addition, Mizoguchi et al. (1992) recently proposed an interactive translation approach in which it is the end-user, and not some intermediary such as a knowledge-engineer, whom interactively translates the knowledge level to the symbol level which the computer understands. In this section, we propose the existence of a knowledge-level 3-D visual ontology, that can be useful for modeling 3-D shapes via verbal descriptions, and a symbol-level 3-D shape ontology, that provides operationalization of shape concepts and modification of those shapes. Figure 2 illustrates the relation of the knowledge-level to the symbol level ontologies.

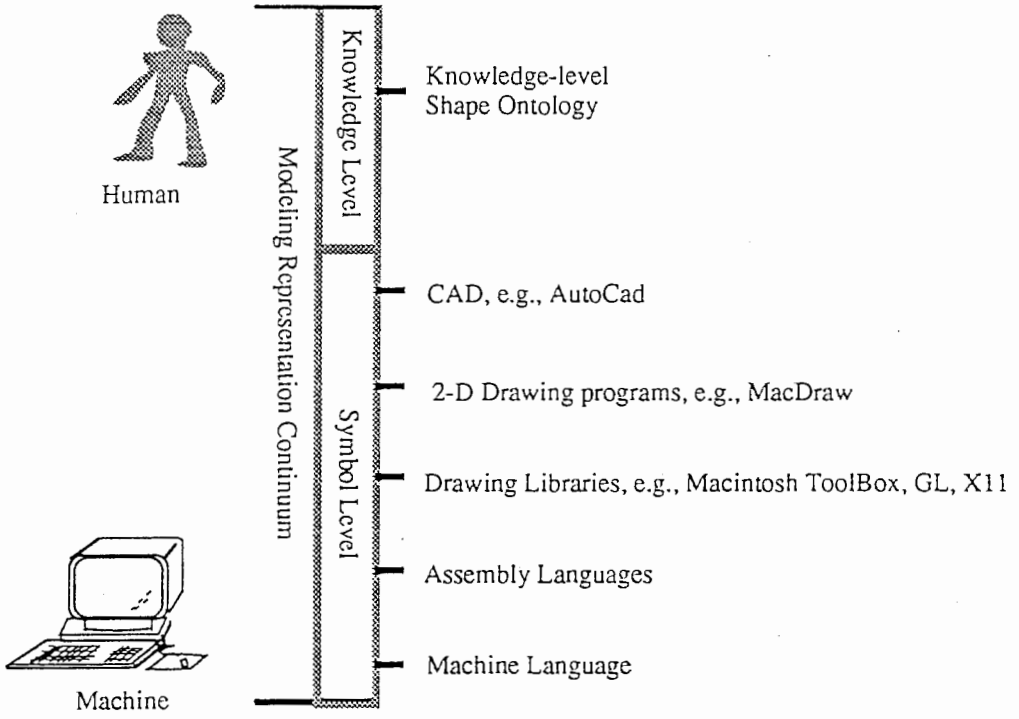

Figure 2. The knowledge-level 3-D visual ontology taxonomizes concepts that are intuitive to humans because of their high level of abstraction, while the symbol-level 3-D shape ontology present a computer representation of those concepts and an operationalization of modifications to shapes.

The knowledge-level 3-D visual ontology which we propose consists of simple 3-D visual concepts and their combinations, which in most cases might be hierarchical, repetitious, or both. Early work on various fields has demonstrated that these visual concepts do exist in nature and

in perceptual forms (Wertheimer, 1923; Thompson, 1942; Johansson, 1950; Rosch, 1973; Stevens, 1974). What we perceive as basic shapes—i.e., cubes, cylinders, spheres, cones, prisms, pyramids and so on—constitute the basic concepts for our knowledge-level 3-D visual ontology. It is not difficult to demonstrate that most shapes in nature, animate or inanimate, can be broken down into these basic shapes. Nishihara (1981) noticed that this type of volumetric primitives seemed to be enough to decompose more complex shapes. Figure 3, illustrates how one might combine these basic shapes to assemble shapes that one perceives as different concepts.

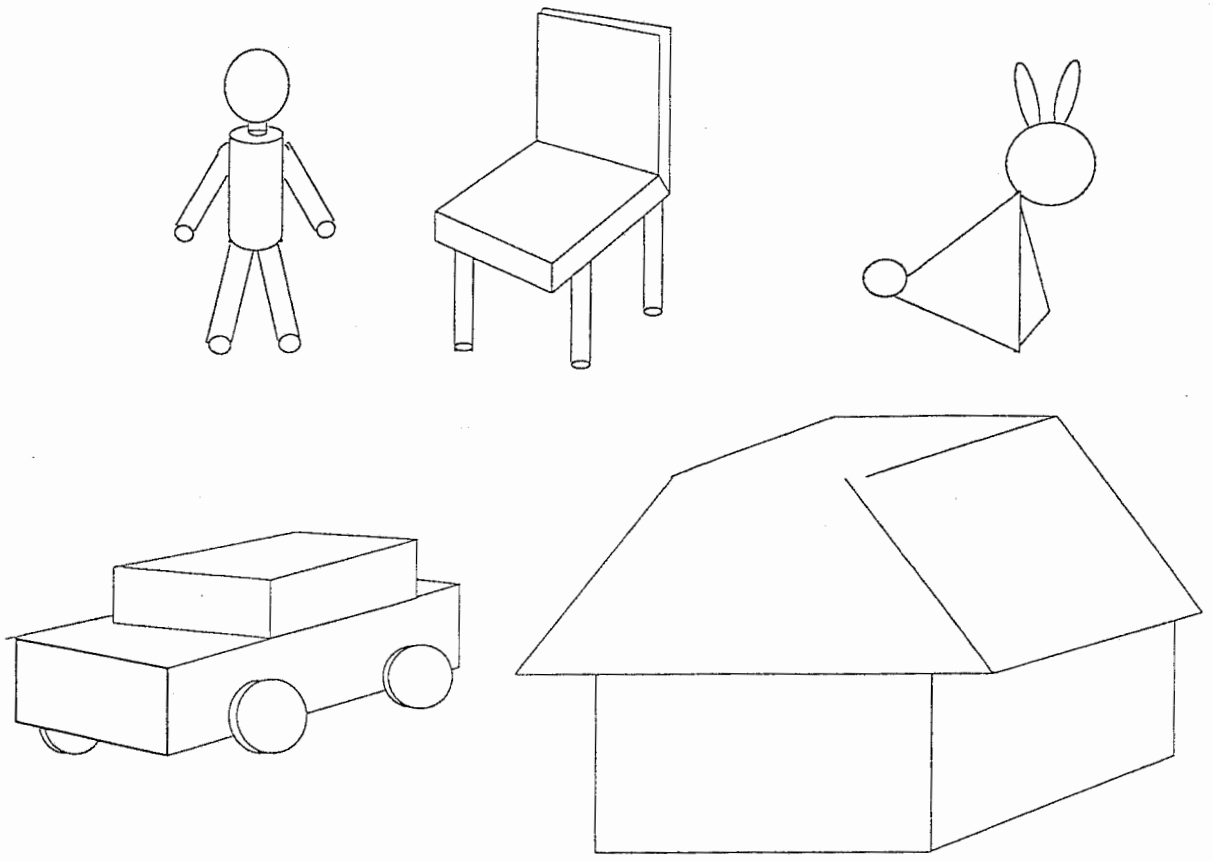

Figure 3. Combination of basic shapes such as cubes, cylinders, spheres, cones, prisms, pyramids can be arranged to produce more complex ones as this figure illustrates.

These basic shapes are not useful enough for our knowledge-level 3-D visual ontology if we cannot perform operations on and with them. That is, our ontology also embraces concepts such as bending, twisting, tapering, rounding, swelling, sharpening, and so on. There are also, descriptors that help to specify states and features of shapes. Where features may be geometrical (e.g., round), visual (e.g., color), functional (e.g., for sitting down) quantitative (e.g., measure of size) or qualitative (e.g., large). Similarly, states may be quantitative positional (e.g., 50 cm to the left) or qualitative positional (e.g., to the left of A). Table 1, summarizes a preliminary tabulation of our knowledge-level 3-D visual ontology.

There seems to be some intricate relation in how we organize these basic shapes to generate more complex ones. Figure 4, illustrates this point. First, figure 4-a presents some primitive shapes in a random pattern. Then, in figure 4-b the same shapes are organized in an specific pattern that we recognize as a car. It was not necessary to give detailed attributes to the car—such as an engine, doors, windows and so on—to recognize that it has the basic shape of car. Yet, we still insist in

recognizing it as a car. Changing the primitive shapes of our car, we can represent more complex concepts—such as that of a racing car—without major effort (see figure 4-c).

Table 1. A preliminary knowledge-level 3-D visual ontology

Primitive shapes	
Concepts	Alternate words
Cube	block, box
Ingot	Rectagular block
Sphere	elipsoid, oval, globe, ball
Cylinder	Trunk
Pyramid	
Cone	
Disc	round sheet
Cylindrical Rod	C. bar, C. stick, C. slab, C. wire
Rectangular Rod	R. bar, R. stick, R. slab, R. wire
Cylindrical tube	hollow C. Rod, hollow C. bar, hollow C. stick, hollow C. slab
Rectangular tube	hollow R. Rod, hollow R. bar, hollow R. stick, hollow R. slab
Prism	
Sheet	
Hoop	doughnut, wheel, ring,
...	
Primitive shape operator concepts	
Shape generation	
generating, combining, creating,...	
Shape modification	
Enlarging, twisting, bending, tapering, rounding, swelling, squaring...	
Primitive descriptors	
Features descriptors	
geometrical	
roundish, cylindric, cubic, flat, sharp,...	
visual	
dark, color, shaded,...	
functional	
for sitting down, for rolling,...	
quantitative	
x centimeters high, y meters wide, z inches thick,...	
qualitative	
large, small, aerodinamic, sporty, ...	
State descriptors	
quantitative positional	
x centimeters left of A, y meters, behind B,...	
qualitative positional	
close, far, left of A, next to B, behind, C,...	
Specification	
This, that, these, those..	
Complex shapes	
Cars, houses, people, rabits, trees, moutains, ...	

These basic shape organization relations seem to be canonical, but are not easily described just in geometric terms. That is, we may be able to recognize figs. 4-b and 4-c as a car and a sports car respectively, because at some point in our lives we have seen what a car and a racing car look like. However, someone that has never seen a car in its life might still recognize the primitive shapes of fig. 4-a. Interestingly enough, Rosch (1973) found that even primitive New Guinea tribesmen also recognized these type of primitive shapes and used them to describe more complex shapes unknown to them. Therefore, there seems to be some kind of background knowledge implicit in the recognition process.

Moreover, there also seems to be some boundary conditions to what we recognize as the shape for something. For instance, consider that the tires of the car represented with the four cylinders in figure 4-b were not placed at the same height in relation to the body or to close together. Would that still be considered a car?

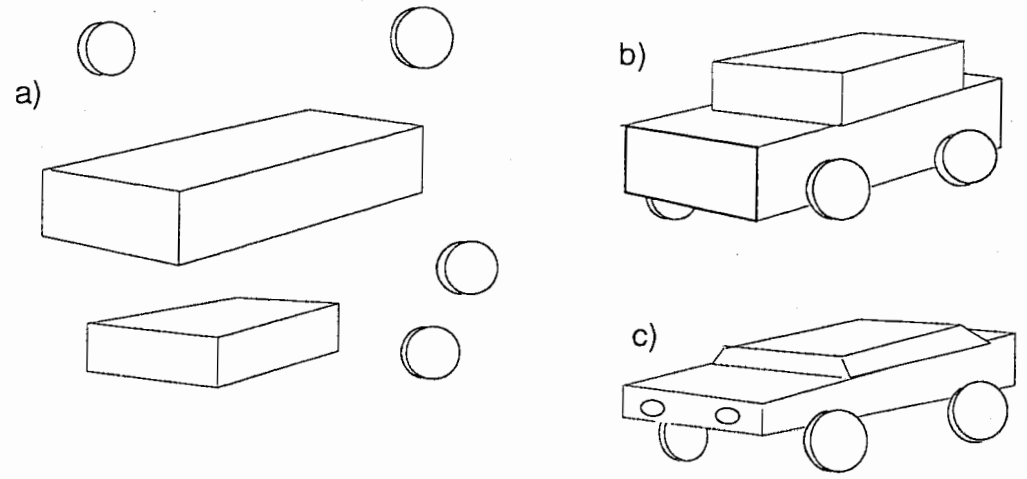

Figure 4. Basic shapes can describe more complex ones: a) presents some basic shapes in random order without any particular meaning, b) organizes the shapes into the simple form of a car, and c) shows how deformations can represent the more complex concept of a sports car.

The knowledge-level 3-D visual ontology that we have presented above lends itself to all sorts of criticisms. For instance, one can argue that it does not provide any kind of insight, since the shapes that we perceive daily are far too complex and, therefore, cannot be captured in any useful way, even with primitive shapes. Howbeit, there are numerous techniques that can be employed to analytically and statistically classify concepts with computer assistance. For instance, machine learning has had great success in learning from examples with techniques such as those for induction (Quinlan, 1986) and explanation-based learning (DeJong, 1986), which could be adapted for common sense visual knowledge. In section 3, we will present results on an experiment which adapted personal construct psychology (Kelly, 1955), a technique that has had great success in knowledge-acquisition for knowledge-based systems (Boose, 1986; Boose and Bradshaw, 1987; Bradshaw et al., 1993; Diederich et al., 1987; Ford et al., 1990; Garg-Janardan and Salvendy, 1987; Shaw and Gaines, 1987), to the acquisition of 3-D visual concepts.

Another criticism that may arise to this knowledge-level 3-D shape ontology is that it is only an approximation, and a relatively poor one on many occasions. Newell (1982) calls it a radically incomplete approximation, but explains that such incompleteness can be tolerated because it leads to unstructured descriptions that can be translated to implementable symbol levels. In this paper, we present approach in our framework for WYSIWYS which takes advantage of knowledge-level unstructured descriptions and interactively translates them to the symbol level.

A third criticism that may arise is that this ontology is in no way at the knowledge level, because traditionally knowledge is represented as being something symbolic in nature. However, Chandrasekaran and Narayanan propose that it is necessary to also provide a theory of common sense visual reasoning (1990). This aroused interest in different research circles and resulted in a field of research known as diagrammatic reasoning (Chandrasekaran et al., 1993). Nonetheless, most work in this area

focuses on 2-D diagrams. To this end, we believe that 3-D shapes should also be included in visual reasoning research and that the type of ontology we propose above makes a significant contribution towards this direction. This type of ontology could make even a larger contribution to knowledge sharing and large common sense knowledge base research, if some adaptations—to accommodate 3-D visual representations—are made to symbolic knowledge sharing languages such as Ontolingua (Gruber, 1992) or CycL (Lenat, 1990).

In short, our knowledge-level 3-D visual ontology consists of verbal concepts of shape encoded in people's memory. Though very recent experiments support the existence of these concepts (Lass et al., 1993), these concepts are not useful for our purposes if no means for machine interpretation is provided. To this end, a symbol level 3-D shape ontology is necessary to which a translation can be performed in order to operationalize the generation, manipulation and modification of computer-simulated 3-D shapes from verbal description.

We define a symbol-level ontology as the set of algorithms or mechanisms that operationalize canonical concepts in a particular task. Consequently, what we need for our symbol-level 3-D shape ontology is: 1) a symbolic representational language that supports feature and state descriptions of the basic and complex shapes, and 2) a set of mechanisms that operationalize transformations of those features and states.

Commercially available CAD applications, such as AutoCAD, WaveFront, Alias and so on, partially fulfill both requirements. CAD applications provide, in most cases, primitives such as cubes and spheres plus a set of menu- or command-driven mechanisms that help modify those primitives. Nonetheless, interaction between the users and CAD applications occurs at a very low level, because CAD applications only provide a set of primitives and a set of tools, which may or may not be the most appropriate. That is, although the basic shape primitives are available (e.g., cubes, spheres, lines, faces, etc.), their combinations into complex shapes have no particular meaning to the application (e.g., the two blocks and four cylinders of fig. 3a will never be a car to the computer, no matter how they are combined), but only to the user. Also, the tools for modification provide only low level modification of the primitives (e.g., extruding, twisting, etc.) and it is left up to the user to combine the tools in some particular sequence to achieve higher level modifications (e.g., a longer car), which may not be intuitive at all.

What is needed to overcome these insufficiencies is a kind of language that can represent higher level concepts (e.g., cars, cups, chairs, etc) and that at the same time provides higher level mechanisms for modification. In section 4 we will introduce a candidate representation, deformable superquadrics, that fulfills both requirements to higher degree than current CAD applications.

3. An Experiment on Visual Concept Acquisition

The basic primitives of an ontology are the concepts from which it is formed. It is extremely difficult to create an ontology that contains all common sense concepts, specially when it comes to visual concepts. Though there is actual research taking place on that direction (Lenat, 1990), it is useful to constraint the context of the concepts to some specific domain or task (Mizoguchi, et al., 1992). This section reports preliminary results of an experiment on visual concept acquisition for description of cars (Umamichi and Tijerino, 1993).

The basic question that we wanted to answer before performing the experiment was: Are there any primitive concepts behind people verbal descriptions of cars? The reason for choosing cars as the object of our

experiment was more arbitrary, than anything else. The fact that more than 200 3-D CG models of cars were already available in a 3-D dead set catalog, courtesy of Viewpoint Datalabs[2], was a determining factor.

We used Personal Construct Psychology (Kelly, 1955) to acquire visual descriptive concepts for cars from a total of 8 subjects, 4 males and 4 females. To insure that the concepts being acquired were of common sense nature, none of the subjects were car designers. To restrain the scope of the concepts, the information shown visually was constrained to 2-D wireframe snapshots of the 3-D models exhibited from random perspectives. That is, with these simplified models, we could guarantee that only descriptive concepts about shape of cars could be acquired.

We begun the experiment by choosing 20 models (listed in table 2) from a stack of 85 randomly. Then for each subject (listed in table 3) the following experiment steps were followed:

1) Randomly choose 3 cards from the stack of 20 and put them in front of the subject.
2) Ask the subject to select two models with a similar feature which the third model does not have.
3) Ask the subject to shortly describe what is similar in the two models and what makes the third one different from the first two. These two opposite descriptions constitute a so-called dichotomy.
4) Repeat steps 1 through 3 until it becomes difficult to describe new dichotomies.
5) For every dichotomy, draw a straight line showing the features at opposite ends of the line.
6) Ask the subject to identify where, on each of the lines, every one of the 20 models can be placed. That is, for every car how would each of the features identified in step 3 best fit.
7) Divide each dichotomy line in 10 equal segments and label each dividing marker from 0 to 10.
8) For each model on each dichotomy line select the closest numeric label as its rating on the dichotomy.

For all dichotomies identified by each subject, we created rating grids showing the name of the subject on the left-most column, each concept of the dichotomy on the next two columns and the names of the models on the top line. Then we rated each model on the dichotomy according to the data gathered on step 8. Table 4, shows the results of the ratings.

Careful analysis of table 4 reveals that some concepts, though differently labeled, have show a very similar pattern of the ratings. This means that the concepts may be the same. The reason for this concept duplication may be attributed to the lack of data (i.e., not enough car models) or bias towards one concept. To overcome this difficulty we chose three car models from the ones in each duplicate dichotomy and performed steps 2 through 8 above for each subject. This time the subjects were asked to think of concepts different from the ones already identified. Only few new concepts were identified with this second run of the experiment (see table 4).

Then, we selected dichotomies with similar labels or that were thought to have the same meaning and normalized the ratings for all subjects, by taking their average. We then entered the results into a grid that contained the normalized dichotomies (See table 5).

2 Viewpoint Datalabs is a company specialized on custom design of 3-D CG data based on Orem, Utah.

Table 2. Labels of 20 models used for the experiment on concept acquisition.

Car Label	Car Type
CarA	91 Lexus 400
CarB	90 Mitsubishi Mirage
CarC	78 AMC Concord
CarD	84 Pontiac 6000
CarE	88 Ford Tempo
CarF	88 Chevy Baretta
CarG	63 Ford Lincoln
CarH	VW Beetle
CarI	83 Cutlass Cierra
CarJ	76 Cadillac Eldorado
CarK	83 Toyota Corolla
CarL	79 Fiat 128
CarM	88 Mazda 929
CarN	90 BMW M5
CarO	90 SAAB 9000
CarP	93 Lancia
CarQ	92 Ford Taurus
CarR	91 Toyota Corolla
CarS	55 Proche Spider
CarT	79 Fiat X/19

Table 3. Subjects for the experiment on concept acquisition.

SubjA	Male
SubjB	Female
SubjC	Male
SubjD	Male
SubjE	Female
SubjF	Female
SubjG	Male
SubjH	Female

From this simple experiment we could learn that it is possible to acquire common sense visual concepts by applying psychological techniques such as PCP. Though the experiment is very simple, there are still many other techniques in PCP that we didn't explore. However, the fact that we could identify, with simple visual analysis, that some of the concepts seem to be canonical among the subjects, makes PCP an appropriate tool for visual concept acquisition. The results of the experiments could also serve to provide numeric rating to concepts for easy database retreival.

Figure 5 shows some of the 2-D snapshots of the 3-D wireframe models of cars used for our experiment. The simplicity of the models makes them ideal for acquiring visual concepts only about shapes; this explains the small amount of concepts acquired with the experiment. On the other hand, the fact that these are just 2-D snapshots takes away the feeling of depth which may be a determining factor to identify other 3-D visual concepts. Also, it may be that attaching textures and color to the models helps the subjects identify other types of concepts (i.e., not about shape).

We are currently implementing a VR-based program for concept acquisition which allows subjects to handle colored 3-D models with 6 degrees of freedom. This way, the subjects are able to see the models from different perspectives as they wish. We hope that this will result in the subjects identifying more concepts that have to do with the 3-dimensional shape of the models. Results on this experiment will be presented in later papers. In the mean time we have constructed a preliminary knowledge-level ontology with concepts shown in table 5.

Table 4. Rating grid for 20 car models (top line). First column to the left shows the subjects. Second and third columns to the left show labels of opposite concepts that compose a dichotomy. Data from the second run is also included.

		CarType →	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
	Concept(0-rat)	Concept(10-rat)																				
SubjA	slow	fast	7	5	5	4	5	6	2	3	5	5	4	1	5	7	8	6	7	3	8	9
SubjA	wide cabin	narrow cabin	3	5	7	3	5	3	7	6	5	2	6	7	5	6	7	6	4	7	9	8
SubjA	economic	luxurious	6	2	6	6	5	5	1	4	4	8	3	1	5	5	8	4	6	2	8	1
SubjA	narrow bag com	wide bag comp	5	5	7	5	9	6	10	2	5	5	5	5	9	10	3	5	8	5	1	8
SubjA	compact	heavy	7	3	8	7	4	4	3	4	6	8	3	1	4	4	6	7	5	3	2	2
SubjA	old design	new design	7	7	4	4	9	5	1	1	5	2	5	2	9	6	7	7	7	6	2	3
SubjA	squarish	roundish	4	4	2	2	3	3	1	9	2	1	2	2	3	2	4	3	4	3	8	1
SubjA	sedan	sports car	4	2	2	2	2	5	3	3	2	5	4	2	2	3	5	2	3	2	9	8
SubjA	expensive	cheap	6	5	5	6	5	4	6	2	5	6	5	3	5	5	7	5	6	6	7	7
SubjA	good mileage	bad mileage	5	2	6	3	4	5	3	6	5	8	7	1	3	4	7	5	5	3	8	7
SubjA	japanese style	western style	1	1	4	4	2	3	5	9	2	5	1	8	2	5	3	2	5	1	7	2
SubjB	cheap	expensive	6	5	3	3	7	5	8	8	3	10	3	7	5	5	9	8	7	5	9	2
SubjB	old	new	6	6	3	4	5	5	2	3	3	1	2	1	5	5	7	6	6	5	10	2
SubjB	squarish	roundish	4	4	3	3	3	3	2	10	3	2	3	2	3	4	5	4	5	4	9	3
SubjB	normal car	sports car	6	6	4	5	5	5	0	0	3	2	7	3	5	5	6	5	8	5	10	7
SubjB	japanese style	western style	2	2	7	7	3	3	9	10	7	10	5	9	3	2	9	3	6	2	9	6
SubjC	huge	tiny	5	1	7	8	3	6	7	5	6	10	6	7	3	6	2	4	5	0	7	9
SubjC	smooth	sharp	5	6	4	3	4	1	4	9	4	2	3	2	7	5	4	8	8	6	10	0
SubjC	kind	agressive	4	6	2	2	5	8	1	1	7	8	1	1	4	5	3	4	5	3	9	10
SubjC	cheap	expensive	3	7	2	5	6	9	1	4	5	10	3	0	4	5	4	7	5	5	8	9
SubjC	old	new	6	8	3	4	7	9	1	2	4	1	3	0	7	7	7	6	8	8	2	10
SubjC	thin	fat	0	4	6	6	6	4	6	8	7	6	6	6	5	7	5	7	6	2	10	#
SubjC	squarish	roundish	7	8	2	3	5	4	2	10	4	0	3	0	9	5	5	6	9	7	10	1
SubjC	sedan	sports car	4	6	3	3	4	8	0	1	3	2	3	1	5	5	6	5	6	4	10	9
SubjC	short	long	3	4	5	8	4	3	10	0	5	9	5	5	4	5	4	5	4	5	1	9
SubjC	sophisticated	simple	5	5	4	2	4	3	7	10	2	0	5	3	4	5	1	4	5	5	8	3
	Concepts added on second run of experiment																					
SubjC	normal	strange shape	0	0	4	2	0	5	10	6	2	7	4	5	0	0	0	0	0	0	8	5
SubjD	slow	fast	7	8	3	7	6	8	2	1	6	7	4	3	4	4	8	6	7	5	8	9
SubjD	smooth	rough	4	6	3	5	5	5	7	8	5	7	4	6	5	4	4	4	3	4	8	7
SubjD	cheap	expensive	5	2	6	2	3	6	1	3	3	8	3	1	4	3	7	2	5	2	3	7
SubjD	old	new	5	5	2	4	5	7	4	0	4	3	2	2	5	4	6	5	5	5	2	8
SubjD	sedan	sports car	2	3	2	2	4	6	0	7	3	1	2	0	3	3	4	3	5	2	9	8
SubjD	straight lined	round lined	4	5	3	4	5	3	1	10	6	1	2	4	6	4	6	4	7	4	9	2
SubjD	unsafe	safe	6	4	7	5	5	4	3	3	6	9	3	2	5	7	8	6	2	5	2	2
	Concepts added on second run of experiment																					
SubjD	short	long	7	4	6	6	5	2	8	0	4	9	1	1	5	5	6	4	4	5	2	3
SubjD	thin	thick	5	7	8	3	5	4	1	9	5	2	4	3	6	7	7	8	7	5	6	2
SubjE	slow	fast	5	4	5	6	4	5	1	1	5	8	6	2	7	5	5	6	5	4	10	5
SubjE	strong	weak	7	0	6	5	8	1	2	9	3	0	3	3	1	2	4	1	6	5	6	1
SubjE	cheap	expensive	4	4	5	5	4	6	5	2	5	1	5	5	6	7	5	9	6	5	1	4
SubjE	narrow-bag com	wide bag comp	2	3	4	5	3	8	10	0	6	10	7	2	7	5	3	7	4	4	0	4
SubjE	squarish	roundish	8	9	7	1	8	0	0	10	1	0	0	0	5	5	6	6	8	6	10	0
SubjE	sedan	leasure car	2	1	2	4	3	0	9	10	5	6	4	5	2	4	6	5	1	2	1	3
SubjE	small	big	4	5	7	5	3	9	8	1	9	10	7	2	7	10	6	10	5	1	8	9
SubjF	cheap	expensive	5	5	6	6	1	5	#	8	4	9	5	6	2	3	5	1	5	2	7	8
SubjF	modern	classic	0	0	6	5	0	0	#	8	5	7	6	7	0	0	0	0	0	0	#	#
SubjF	current	futuristic	2	3	0	1	1	3	6	0	5	2	1	5	5	3	4	0	5	4	9	8
SubjF	squarish	roundish	5	5	8	8	6	5	1	9	3	2	8	8	6	8	5	5	6	5	8	2
SubjF	passive	active	7	0	2	0	2	0	5	0	0	0	0	0	2	1	2	0	5	2	0	0
SubjF	normal	sports car	5	4	2	1	2	1	0	0	2	0	1	0	1	2	3	0	5	4	7	8
SubjF	normal	work car	0	0	0	0	0	0	8	0	0	0	0	2	0	0	0	0	0	#	#	#
SubjF	cold	warm	5	5	3	3	4	4	0	8	5	3	5	6	5	4	4	5	2	5		1
SubjF	japanese style	western style	2	3	9	8	2	4	#	#	8	9	7	7	5	7	2	1	5	2	#	#
SubjG	sloppy	sharp	9	9	6	5	5	5	2	8	4	4	7	3	8	7	9	4	7	8	6	2
SubjG	soft suspension	hard suspensior	4	6	2	5	4	5	0	7	3	0	6	4	4	6	10	4	2	4	10	8
SubjG	family car	couple car	2	2	3	3	3	8	8	0	3	0	8	0	2	3	9	3	0	0	10	10
SubjG	old	new	9	7	4	4	6	5	3	0	5	2	7	0	9	4	5	5	9	9	3	5
SubjG	rear seat center	front seat cente	5	5	3	6	6	1	10	5	6	0	1	5	5	6	1	5	4	5	10	10
SubjG	sedan	sports car	6	5	7	8	6	3	10	10	7	10	5	10	8	6	0	7	7	6	1	2
SubjG	small	big	8	3	5	7	7	7	8	1	7	1	3	1	5	8	7	7	8	4	1	1
SubjG	luxurious	basic	3	9	5	5	8	3	3	10	5	0	3	10	8	5	1	4	4	9	1	1
SubjG	popular	gorgeous	8	2	6	5	5	7	3	0	5	10	3	0	3	6	10	6	6	1	10	7
SubjG	straight lined	round lined	6	7	3	2	3	3	0	10	3	0	0	0	5	3	3	3	8	6	10	0
SubjG	passive	aggressive	5	5	3	3	4	7	0	0	6	0	7	0	1	1	10	5	6	2	8	7
	Concepts added on second run of experiment																					
SubjG	light	heavy	8	4	9	7	4	6	7	4	7	10	3	1	5	7	7	8	8	3	2	2
SubjG	bad mileage	good mileage	3	7	1	3	6	3	1	5	3	0	6	6	6	3	2	2	2	9	3	4
SubjG	cold	worm	7	5	8	5	5	3	8	8	5	2	5	6	5	5	3	2	7	5	6	3
SubjH	nice style	poor style	5	5	3	3	5	4	0	3	2	0	3	2	5	4	5	5	5	5	1	2
SubjH	squarish	roundish	8	8	2	1	5	4	0	10	0	0	0	0	6	4	9	6	6	1	2	2
SubjH	small	big	8	9	6	6	3	6	2	4	9	7	2	1	7	9	7	10	5	5	9	2
SubjH	ugly	cute	5	6	5	3	6	6	3	8	3	1	4	3	5	6	6	5	6	6	2	4
	Concepts added on second run of experiment																					
SubjH	cold	warm	6	5	5	4	5	3	2	8	5	2	5	6	5	4	3	3	5	5	6	2

Table 5. Rating grid for normalized ratings of all 8 subjects.

		CarType =	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
	Concept(0-rat)	Concept(10-rat)																				
1	slow	fast	6	6	4	6	5	6	2	2	5	7	5	2	5	5	7	6	6	4	9	8
2	cheap	expensive	5	4	5	5	4	6	3	5	4	8	4	3	4	5	6	5	6	4	6	5
3	narrow bag comp.	wide bag comp.	4	4	6	5	6	7	10	1	6	8	6	4	8	8	3	6	6	5	1	6
4	light weight	heavy weight	8	4	9	7	4	5	5	4	7	9	3	1	5	6	7	8	7	3	2	2
5	old style	new style	7	7	3	4	6	7	3	1	4	2	4	1	7	5	6	6	7	7	4	6
6	squarish	roundish	6	6	4	3	5	3	1	10	2	1	3	2	5	5	6	5	6	4	8	2
7	sedan	sports car	5	4	3	4	4	5	2	4	3	3	4	3	4	4	4	4	6	4	8	7
8	small	big	7	6	6	6	4	7	6	2	8	6	4	1	6	9	7	9	6	3	6	4
9	short	long	5	4	6	7	5	3	9	0	5	9	3	3	5	5	5	5	4	5	2	6
10	straight lined	round lined	5	6	3	3	4	3	1	10	5	1	1	2	6	4	5	4	8	5	10	1
11	bad mileage	good mileage	4	5	4	3	5	4	2	6	4	4	7	4	5	4	5	4	4	6	6	6
12	cold style	warm style	6	5	5	4	5	3	3	8	5	2	5	6	5	4	3	3	5	5	6	2
13	japanes style	western style	2	2	7	6	2	3	7	10	6	8	4	8	3	5	5	2	5	2	8	4

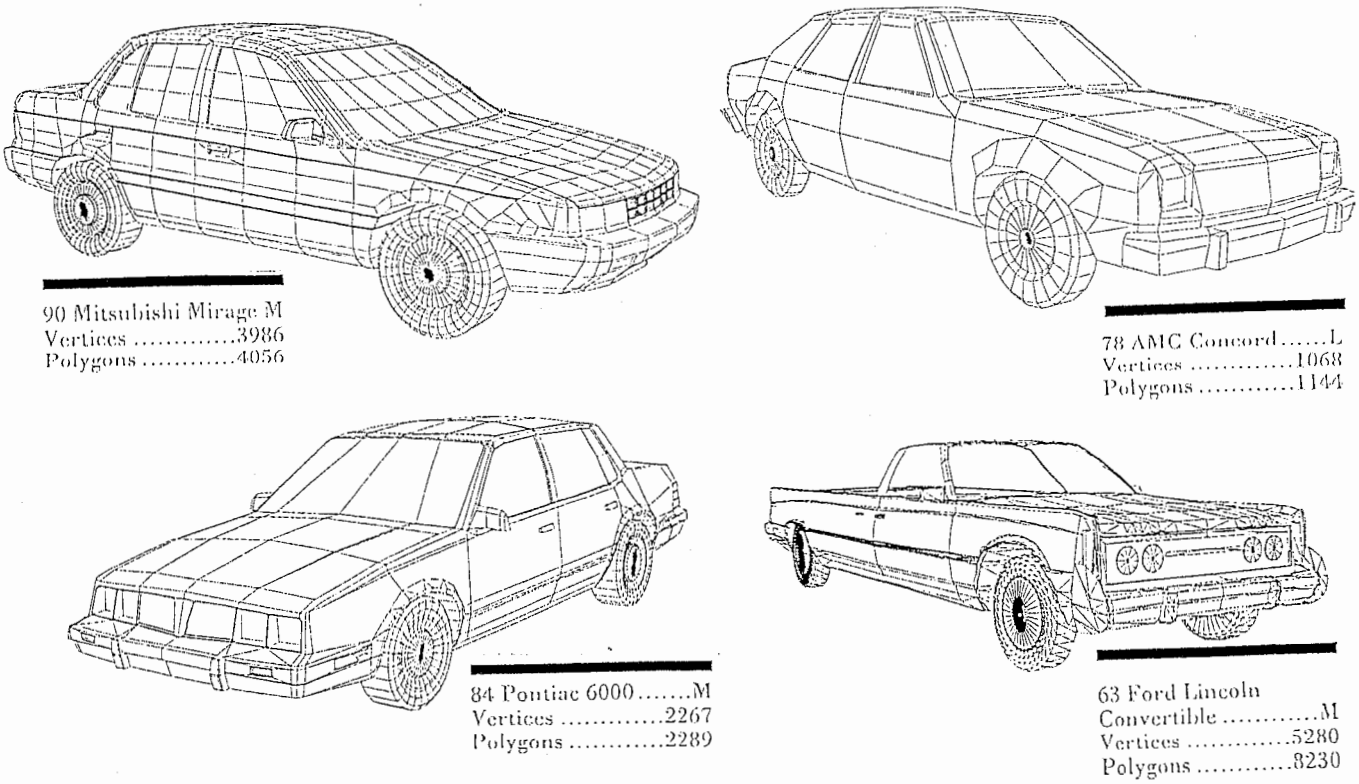

Figure 5. Some examples of 2-D snapshots of car 3-D wireframe models.

4. A Candidate Representation for Symbol-Level 3-D Shape Ontology

It is now appropriate to describe a candidate representation for our symbol-level 3-D shape ontology. The representation we present here is only a candidate because, as explained in previous sections, the only requirements for a representation are that it allows us to symbolically represent feature and state descriptions of basic and complex shapes. Superquadrics satisfactorily complies with this requirement and we have, therefore, chosen it for representation of our symbol-level 3-D shape ontology. This section describes the representation as well as how it fulfills our requirements.

Though superquadrics, which was first discovered by Hein; see (Gardiner, 1965), has recently gathered much attention as a representation useful for shape reconstruction from 2-D and 3-D data (Pentland, 1986; Horikoshi, 1990; and Terzopolous, 1991), has also been recognized as a powerful representation for intuitively building 3-D models.

Superquadrics can be defined by the following vector:

$$\begin{aligned} x &= a1 \cos^{\varepsilon 1}\alpha \cos^{\varepsilon 2}\omega \\ y &= a2 \cos^{\varepsilon 1}\alpha \sin^{\varepsilon 2}\omega \\ z &= a3 \sin^{\varepsilon 2}\omega \end{aligned} \qquad (1)$$

where x, y and z are the coordinates of a surface point on a superquadrics ellipsoid. The parameters a1, a2 and a3 define the scale in the x, y and z directions respectively. The angles α and ω each represent the degrees of latitude and longitude respectively on the the superellipsoid. Finally, $\varepsilon 1$ represents the squareness parameter along the z-y plane and $\varepsilon 2$ the squareness parameter along the x-y plane. Though by just changing the squareness and scale parameters we can represent a few basic shapes such as those illustrated in figure 6, these shapes are only allowed to be symmetrical.

We can substitute the scale parameters a1 and a2 with modifying functions f(z) and g(z) to yield:

$$\begin{aligned} x &= f(z) \cos^{\varepsilon 1}\alpha \cos^{\varepsilon 2}\omega \\ y &= g(z) \cos^{\varepsilon 1}\alpha \sin^{\varepsilon 2}\omega \\ z &= a3 \sin^{\varepsilon 2}\omega \end{aligned} \qquad (2)$$

which will in turn allow us to bend, taper or twist the basic shapes into non-symmetrical ones.

Pentland (1986) demonstrates superquadrics utility in building 3-D models through a system called SuperSketch, a Symbollics-3600-based 3-D modeling application. In this system, users create "lumps," change their squareness/roundness, stretch, bend, taper them, and make Boolean combinations of them in real time by moving the mouse through the relevant parameter space, controlling which parameter is being varied by using the mouse buttons. Pentland noticed that because primitives, operations and combining rules used by the computer are well matched to those of the human operator, the interactions described above were surprisingly effortless. Pentland further states that descriptions couched in this representation are similar to people's (naive) verbal descriptions and appear to match peoples's (naive) perceptual notion of a "a part", and that this correspondence is strong evidence that the descriptions we form will be good spatial primitives for a theory of common-sense reasoning. However, he noticed that domain experts formed descriptions differently than naive observers, reflecting their understanding of the domain-specific

formative processes and their more specific, limited purposes, and that accounting for expert descriptions will require additional, more specialized models. It is this last point in which we believe that the two-level ontology can be the most helpful, because abstract and specialized descriptions at the knowledge level can be indexed to descriptions at the symbol level which provide proper operationalization.

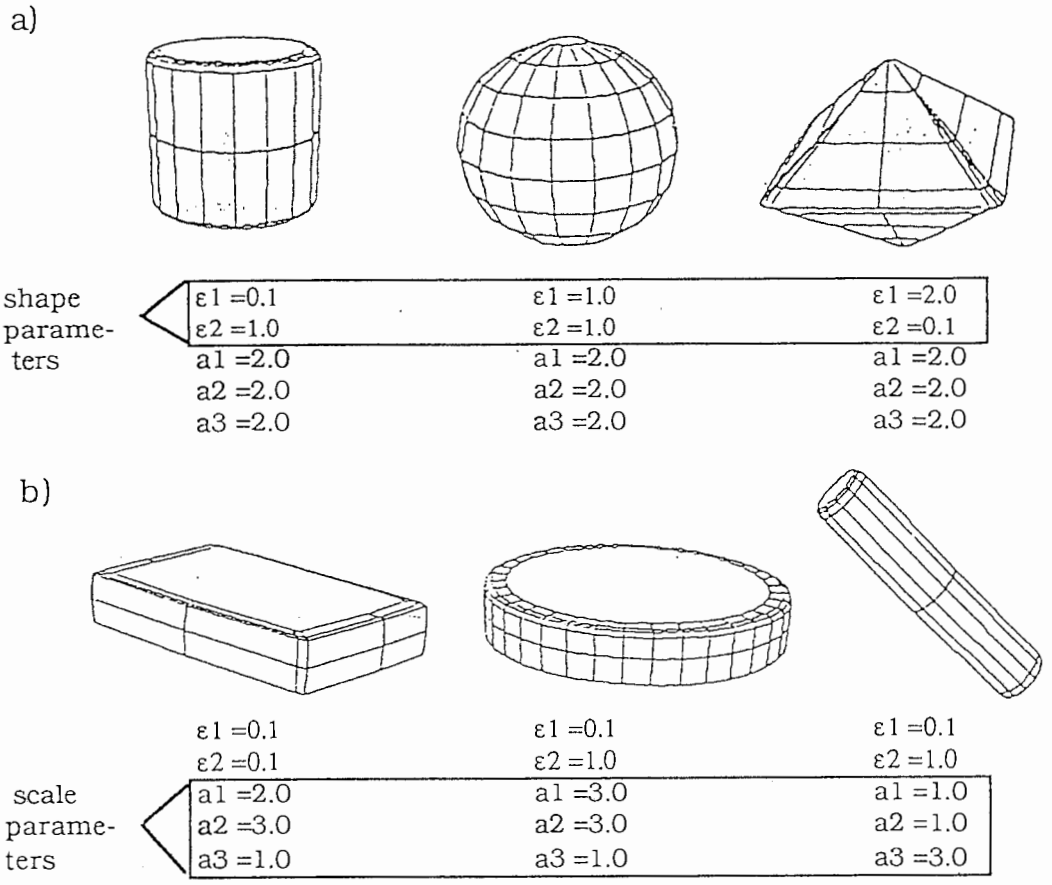

Figure 6. 6 symmetrical shapes that can be represented with simple superquadrics.: a) represents shape modification through squareness parameters and b) represents changes through scale parameters (adapted from [Horikoshi and Kasahara, 1990]).

On a separate research effort, Horikoshi and Kasahara (1990) proposed superquadrics as a multiple purpose indexing language for 3-D models. They demonstrated that superquadrics parameters can be easily mapped to descriptive words, of shape and transformations of shapes. They also employ superquadrics parameters to index and search in a database of 3-D objects with help of two orthogonal free-hand sketches. They further claim that this is an unified 3-D indexing language that can interconnect descriptive words or 2-D images with 3-D information. The power of their approach becomes clear with the fact that simple words such as cube, sphere, cone and pyramid can be indexed to shapes with small variations but that can be considered to fall into those categories. Also, simple words such as pinching, collapsing, denting, sharpening, etc, can be used to represent transformations on the shapes. Figure 7 shows an example of how these type of words can be indexed to this type of transformations.

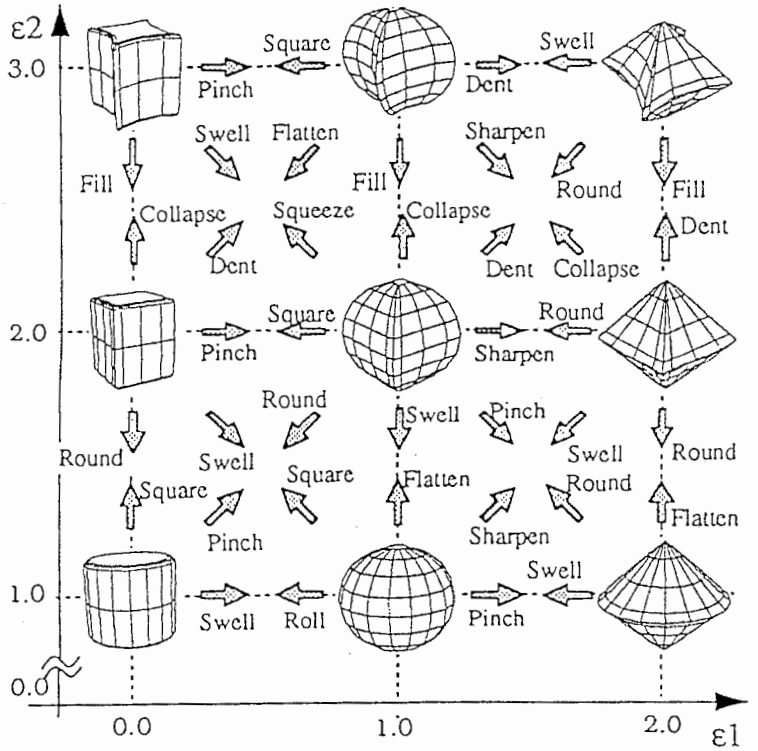

Figure 7. The relation between shape parameters and words (from [Horikoshi and Kasahara, 1990]).

Later on, Terzopolous makes further contribution to our representation by introducing deformable superquadrics (1991) which incorporate the global shape parameters of conventional superquadrics with the local degrees of freedom of a spline. This combination of local and global deformation parameters make superquadrics particularly useful to both provide global salient part descriptors for efficient indexing into a database of stored models and to reconstruct the local details of complex shapes that the global abstraction misses. Terzopolous applies equations of motion to govern the behavior of deformable superquadrics by employing local finite element basis functions because they provide greater shape flexibility for local deformations.

Superquadrics allows us to index numerous 3-D shape classes, with both global and local variations. Thus, superquadrics are appropriate for representing symbol-level 3-D shape ontologies, because the number of parameters involved is very small and also have been proven to map very intuitively to words which describe objects and their transformations. However, a one-to-one mapping from words to arrangement of parameters or transformation is not desirable for our WYSIWYS framework, because it doesn't take into consideration the vagueness in natural language. For this reason, it is necessary to make the mapping between concepts of the knowledge-level 3-D visual ontology described in a previous section. This way, the vagueness in natural language can be managed to some extent by indexing concepts with one or more labels (i.e., words).

5. The WYSIWYS Framework

People tend to picture in their minds what an object might look like after reading or listening to its description (Anderson, 1978). If the object is known, we would automatically retrieve from our memory a prototype of the object (Lass et al., 1993). When it is not known, then we try to imagine what it looks like by relating it to known objects (Rosch, 1973). With the WYSIWYS framework we can actually see what someone is describing as the description progresses. This framework allows people to verbally describe 3-D objects an transformations performed on those

objects as the objects are being displayed by the computer. If the framework is implemented taking advantage of natural language processing and virtual reality technologies, it promises to revolutionize human-to-human and human-to-machine communication.

5.1 Characteristics of the WYSIWYS framework

A system based on the WYSIWYS framework requires at least the following characteristics:

1) flexibility in interpretation of vagueness in verbal descriptions,
2) real-time visual simulation of verbal descriptions,
3) accessibility of simulated shapes through some intuitive input device, and
4) flexibility in selection and assignment of spatial attributes.

Flexibility in interpretation of vagueness in verbal descriptions:

In our two-level ontology approach, words-to-concept mapping permits the flexibility required by verbal descriptive vagueness when describing objects. Take for instance the concept of "cube" in table 1, the labels cube, box and block are all indexed into it; this allows us to refer to the same object in three different manners. Correspondingly, we could define a fuzzy set that takes into account the squareness parameters range to prescribe the meaning of "cube" in a superquadrics representation (Horikoshi and Kasahara, 1990), which are necessary to, among other things, index the concept of "cube" at the symbol level and to display the object.

The indexing of words to concepts and concepts to symbolic representations should be interactive. In the case of words to concepts mapping, interactiveness of mappings allows customizing. Interactive indexing can be accomplished by allowing an acquainting period for the user and the system, in which—for instance—the system gets acquainted to how the user calls its primitive shapes and primitive operators. This is simple to attain because—considering that the number of primitives remains small—all the system has to do is show the basic shapes or animate the primitive operations and request their labeling to the user. The system could then keep an active glossary (Klinker et al., 1992) of the labels for the concepts. In the other case—indexing of concepts to symbolic representation—interactivity provides easy expandability of the scope of the two-level ontology. That is, when a person wants to define a new concept, all is needed for that person to do is to define what the concept means in terms of the symbolic representation. With deformable superquadrics, for instance, if the concept of a new shape is to be defined, the person has only to provide the range of values of the global and local parameters that define the new shape. Because in both cases, instantaneous visual feedback can be obtained, the need for a more complex natural language processing sub-system can then be avoided.

Real-time visual simulation of verbal descriptions:

It is very important to provide real-time visual feedback in the form of perspective monoscopic or stereoscopic display of 3-D shapes and operations being performed on the shapes. This is the essence of the "see" part in the WYSIWYS framework. A system based on this framework should be able to match the intentions of the verbal descriptions, i.e., the "say" part of the WYSIWYS framework, with adequate changes to the visual display environment. This makes possible for a WYSIWYS-based system to avoid dependance on a extrict natural language processing. Because instantaneous visual feedback allows the user to "see" what the system "thinks."

Accessibility of simulated shapes through some intuitive input device:
A WYSIWYS-based system should also support some kind of pointing device, preferably a DataGlove™ or similar device, that at least permits actions such as moving in the 3-D space, pointing or selecting objects, identifying a point in space, etc.

Flexibility in selection and assignment of spatial attributes
When we verbally describe objects, we tend to make gestures with our hands which may or may not aid a listener to better understand what we are describing. We often employ simple hand gestures that describe rough position, direction and orientation. Concepts such as here, there, in this direction and so on, depend on abstract hand gestures, which do not have to be precise as long as they convey the intention of the concept. Superquadrics allows us to adjust scale parameters to hand positions—provided that a DataGlove™-like input device is being used—for concepts such as enlarging or reducing the size of an object. In a similar manner, direction and proportion of movement of our hands can also help to adjust position or orientation vectors.

5.2 OVO-Genesys: a system that implements the WYSIWYS framework
We are currently developing an Ontology-based Virtual Object GENerating SYStem: OVO-GENESYS (Tijerino, et al., 1993), which implements the WYSIWYS framework described above. The architecture is illustrated in figure 8 and is composed of the following.
Concept database: This consists of a knowledge-level 3-D visual ontology about cars. Concepts may or may not be hierarchical. Each concept maybe associated with one or more labels. Concepts are classified in four major groups: state descriptors, feature descriptors, state operators and feature operators. Features may be of geometrical (e.g., round, square), visual (e.g., color), functional (e.g., for sitting down) quantitative (e.g., size) or qualitative (e.g., large) nature. Similarly, states may be quantitative positional (e.g., 50 cm to the left), qualitative positional (e.g., to the left of A).
Object database: A database that contains 3-D models of cars along with indexing to relevant concepts.
Natural language parser: Which uses labels associated to concepts in the concept database as its basic vocabulary.
Gesture understanding module: Which interprets simple hand gestures—such as extended hand and pointing with one finger—, translation and rotation.
Object Cataloguer/Browser: Which permits interactive indexing of labels to concepts and concepts to deformable superquadrics representations.
Object Analogy Engine: This module parses a natural language phrase and searches in the concept database for an adequate state or feature descriptor/operator.
Object Modification Engine: This module parses natural language phrases and maps concepts which express object modification in terms of feature or state operators to specific actions on the objects. The operators are coupled with hand gestures for modification.
2-D Sketch Understanding module: The 2-D sketch understanding module allows the user to input 2-D sketches of objects. The objects are then transformed into 3-D representations with help of verbal descriptions of the object. The 2-D sketch serves as the rough representation of the object, its spatial characteristics and its proportional geometry. The verbal description of the object is mapped to state and feature descriptions and coupled with hand gestures in a similar manner to the Object Modification Engine.

As this paper is being written, Ovo-Genesys is still in an early development stage, therefore it is still difficult to say how accurate conceptual descriptions can be. The system is being implemented on a Onyx RE^2 graphics super computer in both C++ and CLOS. Currently we are working with a database of automobiles but plan to extend it to include more general types of objects. A later paper will report on experiments with Ovo-genesys to test for accuracy of mappings from verbal descriptions to visual display.

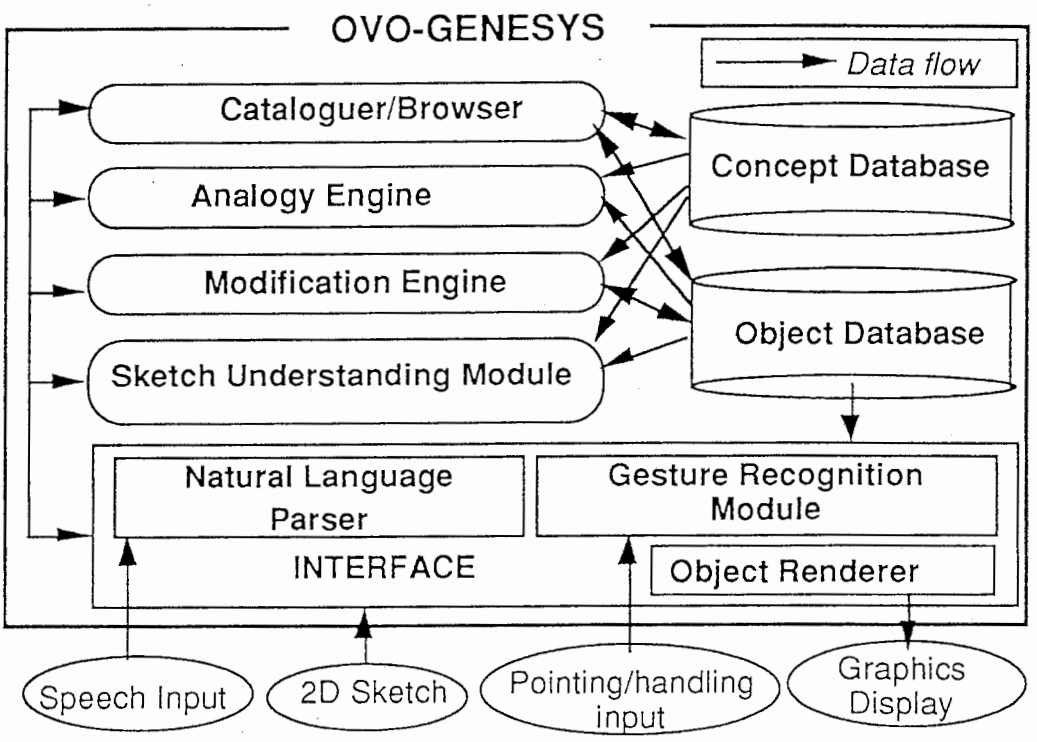

Figure 8. OVO-GENESYS's general architecture.

6. Discussion

The WYSIWYS framework proposed here has numerous implications to human-to-human and human-to-machine communications. It allows for people to "show" to each other what they mean in a computer simulated 3-D medium which can be manipulated in a very intuitive manner through verbal descriptions and simple hand gestures. This in itself, has many applications to such fields as CAI, CG art, CSCW, VR, virtual space teleconferencing, and so on. Another implication in human-to-human communication, is that it provides guidelines for research in common sense 3-D visual knowledge and reasoning that are essential for building very large 3-D visual knowledge bases and also for allowing 3-D visual knowledge sharing.

For human-to-machine, in the other hand, the implications of the WYSIWYS framework are no less important. For instance, CAD applications could incorporate support for knowledge-level 3-D visual ontologies on top of their existing symbol level representations and thus allow intuitive verbal interactions. This is, of course, just one of the many applications that can benefit from the WYSIWYS framework. Others include but are not limited to, intelligent 3-D database search through verbal instructions and 2-D sketches (Horikoshi and Kasahara, 1990), 3-D scientific simulations, machine learning and so on.

A very important contribution of the framework is in the integration of natural language and computer graphics. It provides guidelines on how to avoid too much dependance on natural language processing and take

adavantage of real time computer graphics interpretations. This way, the people can "see" what the computer is "interpreting" from natural language. This instantaneous feedback allows people to interactively guide the interpretations process. The framework also makes an important to VR technology by providing an alternative means for interactive generation, manipulation and modification of 3-D models, an area in which not much research has been focused (Mochizuki and Kishino, 1991; and Butterworth et al., 1992).

There are many problems that have to be solved before the WYSIWYS framework can be widely applied. Probably, the most important problem is the generation an adequate knowledge- and symbol-level 3-D shape ontologies for various domains. The domain chosen in this paper was for cars, but there are numerous domains that can be more useful, not only for common sense 3-D knowledge, but also for 3-D visual knowledge sharing. For instance, human anatomy, architecture, sculpture, and so on. Likewise, natural language understanding is another problem; it also has to be improved. Though simple hand gestures might help towards this end (Mochizuki and Kishino, 1991), there is still much to be expected from speech understanding. Additionally, there remains the problem of resolving segmentation of shapes to support common sense 3-D visual knowledge. That is, we must answer the question of whether everyone segments all shapes in a consistent manner (e.g., can the shape of a house always be segmented in a prism for the roof and a block for the body?).

In this paper, we proposed the existence of a two-level ontology—that is, a knowledge level and a symbol level—for 3-D visual knowledge that supports a WYSIWYS framework. We also presented some preliminary results in an experiment to acquire knowledge-level 3-D visual concepts to prove the feasibility of the framework. We introduced a candidate symbol-level representation, deformable superquadrics, for symbol-level 3-D shape ontologies. In addition, we described the essential characteristics of the framework for WYSIWYS interactions and described an implementation effort taking place at ATR CSRL. This paper at least, makes an initial contribution for the WYSIWYS framework for human-to-human and human-to-machine communication.

Concluding remarks

Virtual reality technology has been revolutionizing fields that benefit from computer visualization. However, interactions have been limited to input devices, such as the DataGlove™, the SpaceBall, Joysticks and so on. Although in a few cases there has also been incorporation of voice input, the WYSIWYS framework outlines guidelines for more intuitive verbal interactions. The two-level 3-D visual ontology conception introduced in this supports the framework so that intentions are better captured from verbal descriptions. In other words, it makes it possible for a word to be worth a million pictures.

References

Anderson, J. R. (1978). Arguments concerning representations for mental images. *Psychological Review*, Vol. 85, 249-277.

Boose, J. H. (1986). *Expertise Transfer for Expert Systems*, Amsterdam: Elsevier.

Boose, J. H. and Bradshaw, J. M. (1987). Expertise transfer and complex problems: using AQUINAS as a knowledge acquisition workbench for knowledge-based systems, *International Journal of Man-Machine Studies*, Vol. 26, pp. 3-28.

Bradshaw, J. M., Ford, K. M. Adams-Webber, J. R. and Boose, J. H. (1993). Beyond the repertory grid: new approaches to contructivist

knowledge acquisition tool development, *Internationals Journal of Intelligent Systems*. Vo. 8, No. 2, pp. 287-333.

Butterworth, J., Davison, A., Hench, S. and Marc Olano, T. (1992). 3DM: a three dimensional modeler using a head-mounted display. ACM 0-89791-471-6/92/0003/0135.

Chandrasekaran, B. and Narayanan, N. H. (1990). Towards a theory of commonsense visual reasoning, in: Nori, K. V, and Veni Madhavan, C. E. (eds.), *Lecture Notes in Computer Science 472*, Springer-Verlag, Berlin, 1990, PP. 388-409.

Chandrasekaran, B., Narayanan, N. H., and Iwasaki, Y. (1993). Reasoning with diagrammatic representations—a report on the spring symposium—, *AI Magazine*, Summer 1993, pp. 49-56.

Dejong, G.F. (1986) Explanation-Based Learning. In R.S. Michalski, J. G. Carbonell, T. M. Mitchell (eds.), *Machine Learning: An artificial intelligence approach Volume II*. Los Altos, CA: Morgan Kaufmann.

Diederich, J., Ruhmann and May, M. (1987). KRITON: A knowledge acquisition tool for expert systems, *International Journal of Man-Machine Studies*, Vol. 26, No. 1, pp. 29-40.

Ford, K. M., Cañas, A., Jones, J., Stahl, H., Novak, J., and Adams-Webber, J. (1990). ICONKAT: an integrated constructivist knowledge acquistion tool, *Knowledge acquisition*, Vol. 3, No. 2, pp. 215-236.

Gard-Jarnadan, C., and Salvendy, G. (1987). A conceptual framwork for knowledge elicitation, *International Journal of Man-Machine Studies*, Vol. 26, No. 4, pp. 521-531.

Gardiner, M. (1965). The superellipse: a curve between the ellipse and the rectangle, *Scientific American*, Vol. 213, pp. 222-234.

Gruber, T. (1992). A translation approach to portable ontology specifications, Stanford University KSL Technical Report KSL 92-72

Horikoshi, T. and Kasahara, H. (1990). 3-D Shape indexing language, in *Proc. of the 1990 International Conference on Computers and Communications*, pp. 493-499.

Johansson, G., (1950). *Configurations in Event Perception*, Almqvist and Wiksell, Stockholm.

Kishino, F. Communication with realistic sensations (1990). *3-D image*, 4, 2. (in Japanese)

Klinker, G., Marques, D., McDermott, J., Marsereau, T., Stintson, L. (1992). The active glossary: taking integration seriously. In Proc. of 7th Knowledge Acquisition for Knowledge-Based Systems Workshop, pp. 14-1 to 14-19.

Lass, U., Lüer, G., Ulrich, M. ans Werner, S. (1993). Access to analog representations in memory for visually preceived forms: the facilitating effect of declarative knowledge, in: Strube, G. and Wender, K. F. (eds), *The Cognitive Psychology of Knowledge*, North-Holland, Elsevier Science Publishers B. V, pp. 75-96.

Lenat, D. B. and Guha, R. V. (1990). Cyc: toward Programs with common sense. *Communications of the ACM*, **33**(8), 30-49.

Mizoguchi, R., Tijerino, Y. A. and Ikeda, M. (1992). Two-level mediating representation for a task analysis interview system, In *Proc. of AAAI-92 Workshop for Knowledge Representation Aspects of Knowledge Acquisition*. San Jose, Ca., pp. 107-114.

Mochizuki, K. and Kishino, F (1991). A 3-D shape access interface considering an individual variations of spatial indication concepts, in *Proc. of 7th Symp. on Human Interface*, pp. 51-54.

Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, T., Snator, T. and Swartout, W. R. (1991). Enabling technololy for knowledge sharing, *AI Magazine*, Vol. 12, No. 3, pp. 36-56.

Newell, A. (1982). The knowledge level, *Artificial Intelligence*, **18**, 1, pp.

87-127
Pentland, A. P. (1986). Perceptual organization and the representation of form, *Artificial Intelligence*, **28**, pp. 293-331.
Quinlan, R. (1986). Induction of decision trees, *Machine Learning*, Vol. 1, No. 1, pp. 81-106.
Rosch, E. (1973). On the internal structure of perceptual and semantic categories, in: Moore, T. E. (ed.), *Cognitive Development and the Acquisition of Language*, Academic Press, New York.
Shaw, M. L. G., and Gaines, B. R. (1987). KITTEN: Knowledge initioation and transfer tools for experts and novices, *International Journal of Man-Machine Studies*, Vol. 27, No. 3, pp. 251-280.
Steels, L. (1992). End-user configuration of applications, *In Proc. of 2nd Japanese Knowledge Acquisition for Knolwedge-Based Systems Workshop*, pp. 47-64.
Stevens, S., (1974). *Patterns in Nature*, Atrantic-Little, Brown Books, Boston, MA.
Takemura, H. and Kishino, F (1992). Cooperative work environment using virtual workspace, in *Proc. of CSCW'92*, pp. 226-232.
Terzopoulos, D. (1991), Dynamic 3D Models with local and global deformations: deformable superquadrics, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, pp. 703-714.
Thompson, D-A., (1942). *On Growth and Form*, University Press, Cambridge, U.K., 2nd ed.
Tijerino, Y. A., Abe, S., Miyasato, T. and Kishino, F. (1993), in *Proc. of 47 National Conference of the Information Processing Society of Japan*, Vol. 2, pp. 385-386.
Tijerino, Y. A. and Mizoguchi, R. (1993). MULTIS II: enabling end-users to design problem-solving engines via two-level task ontologies, in Aussenac, N., Boy, G., Gaines, B., Linster, M., Ganascia, J. G. and Kodratoff, Y. (eds) Lecture Notes in Artificial Intelligence 723—Knowledge Acquisition for Knowledge-Based Systems—, Springer-Verlag, pp. 340-359.
Umamichi, T., and Tijerino, Y. A. (1993). A report on the acquireability of descriptive concepts for cars based on personal construct psychology. ATR Technical Report TR-C-0092. (In Japanese)
Wertheimer, M. (1923). Laws of organization in perceptual forms, in: Ellis, W. D. (ed.), *A source Book of Gestalt Psychology*, Harcourt Brace, New York.