INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'07
28 - 31 AUGUST 2007, CITE DES SCIENCES ET DE L'INDUSTRIE, PARIS, FRANCE

A VISUAL LEXICON TO HANDLE SEMANTIC SIMILARITY IN DESIGN PRECEDENTS

John Restrepo[1]
[1]Mechanical Engineering Department, Technical University of Denmark

ABSTRACT
Designers make extensive use of design precedents as a tool to formalize vague ideas and as a means of communicating ideas to others. Currently, there are no systems to support them in this process other than experimental lab efforts. This paper presents two approaches that have been used to create such a supporting system. The first one uses an image recognition engine to query a database and another one that uses a semantic lexicon (WordNet) to do searches by similarity. The paper presents illustrations of both and describes the future work that is needed to integrate the two approaches into one system.

Keywords: Design Semantics, Design Precedents, Content Based Retrieval, Design Support

1 INTRODUCTION
Imagining, seeing and drawing are inseparable activities, all of which form an essential part of design, especially during the early stages. The process of translating vague ideas in the designer's mind into sketches and models to communicate requires great skill and effort. Designers make use of tools such as collages, mood boards and other visual means to communicate their ideas for which images of existing products (precedents) are a welcomed aid. Searching for these precedents, however, is difficult, because it requires either exhaustive browsing or verbalization of the ideas to be able to use the existing search engines.

Previous research using image recognition techniques and a query by example approach has shown that having a quick access to large collections of design precedents that can be browsed or searched in a systematic way is a significant step into helping designers satisfy their visual information needs [18].

The same research also showed that using geometrical similarity (image recognition) as the only search criteria is not enough. To humans, similarity is more than appearance. Semantic similarity is important and in cases more easy to identify. This paper presents some ideas for the construction of a visual lexicon that can be queried using the same query by example approach reported on [18], but using a language lexicon instead of image recognition to provide semantic querying of a visual database of design precedents.

The paper is structured as follows. First, it discusses the use of precedents in design. Then it briefly presents previous research done by the author to support the use of design precedents by the use of image recognition techniques and a query by example approach. It discusses the limitations of such approach and proposes a solution. A system describing the use WordNet for the creation of a visual semantic lexicon and its implementation is presented. Finally, the paper discusses the future work needed to combine geometrical similarity with semantic similarity.

2 PRECEDENTS IN DESIGN

Creating collages, sketches and other types of (external) visual representations are used by designers to help in shaping and establishing the vague ideas that appear in early stages of the design process. One image is quite often the only thing a designer needs to represent the essence of a concept, but it is often difficult for the designer himself to find out what image it will be. In our work with designers, we have observed that they often search for visual information by flicking through magazines or image databases. One of our interviewees described such process as "searching for something you cannot describe", trying to find the precedent that carries the kind of knowledge needed to formalize the idea. This visual thinking and visualization of ideas is a process that is inherent to conceptualizing solutions in form and material [11], [7], [21]. At this stage of the design process, visual information is preferred over linguistic information for it fits better the designers' way of thinking. The designer knows, thinks and works in a visual way [5]. For a more complete discussion on the role of precedents in design see [16] and [18].

Supporting designers in their search for images to aid this visual thinking has been researched by many scholars, whom have tried out alternative solutions (for a complete review see [18]). However, most of the systems developed as a product of this research require the addition of metadata to the images in the form of 'design stories' [14], [2], affective associations [14], requirements [6], icons [23], [3] or sketches [8], [24]; or need to be classified within a predetermined structure to browse through [12], [4], [20]. These approaches have some important drawbacks. The most important drawback is that to be useful, a database with precedents has to have a significant amount of images. Describing and indexing each image is time consuming, labor intensive and expensive. The second drawback has to do with the way images are understood and interpreted. Attribution of meaning is very personal, subjective and situational and can hardly be determined beforehand by the editor of a collection.

To date, there is no system in place that allows the indexing, retrieving and organizing of visual information without a significant amount of effort being put into indexing it, specially when this indexing can be outdated even though the visual information is still relevant, e.g. if a curator of a collection describes images as "fashionable". Although the image might still be useful some time later as a design precedent, the label fashionable might not be relevant anymore. The high costs associated with the indexing of the data and the amount of maintenance required to keep it up to date, are major obstacles for the assimilation of these technologies in industry.

The approach used by the abovementioned systems, filtering the information with the keywords given by the designer, can be sufficient when and if the designer has a good idea of what to search for, and if the search criteria are well defined. However, as we have found out in our studies, information needs cannot always be expressed in such a straightforward manner (using language).

Design precedents have many purposes other than being used as a means to explore possible solutions, as a source of knowledge and as a source of inspiration. Precedents can also be used as referents in the codification of messages that have to be transmitted by a product. For instance, if a designer intends to design a product that looks sturdy, sportive or elegant to a certain social group, salient characteristics of other products that are considered sturdy, sportive or elegant by that social group serve as a starting point for the exploration of new forms. This exercise requires not only a good understanding of the codes to be used, but also access to a large number of referents (e.g. products containing those codes). In this case, the filtering approach is no longer adequate. The reason is that images (or what they represent) do not have an intrinsic, fixed meaning that can be completely captured in words by the editor of an image collection. Instead, the meaning of an image is characterized by three factors: It is contextual, it is differential and it is situational [19].

The meaning attribution characteristics and uses of design precedents are not reflected in the design of current systems to support their handling and usage, as most of them are based on the addition of metadata such as descriptions, icons or sketches. In all these approaches, the data must be described beforehand by either the designers themselves or by editors of a collection of precedents, and with few exceptions, the collections must be queried using keywords. This means that designers need to have a

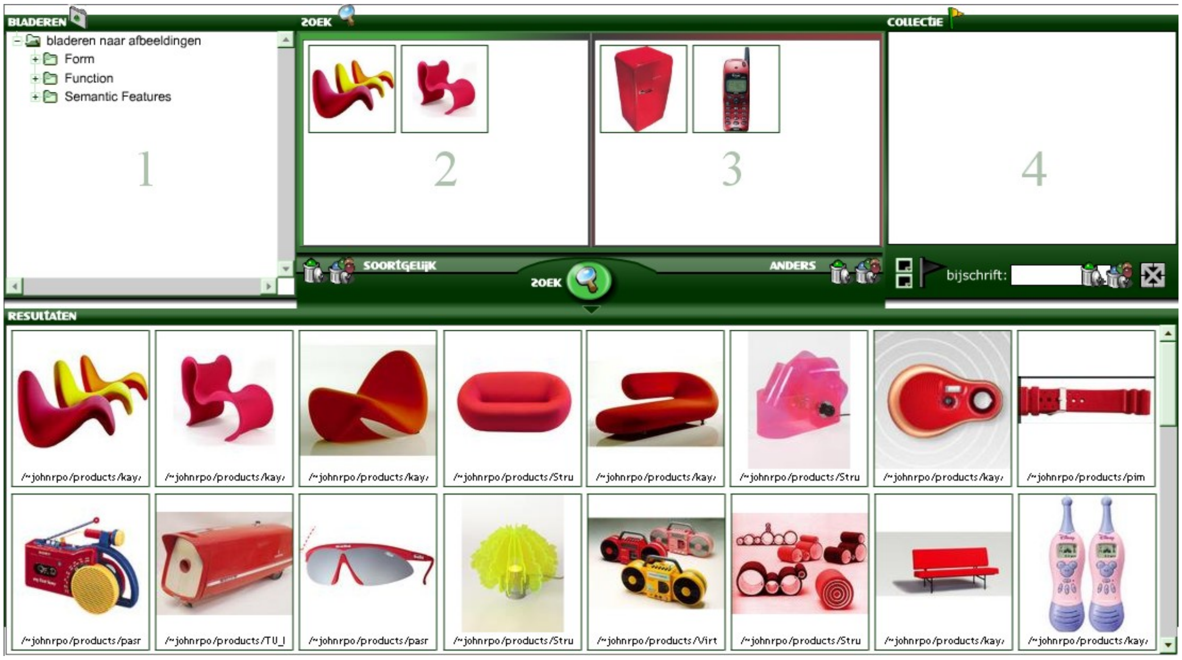

Figure 1 Screenshot of a Query by Example system. In area 2, the user drags the images to be used as examples (should look like this) and area 3 is used for counter examples (should not look like this). The results are ordered by similarity. Area 1 is used to allow the user to browse through a set of examples of products that will serve as seeds for the query. Pressing the search button with an empty query returns a random set of images. Area 4 is used to store images or to compose collages.

good idea of what is being sought, as well as a good idea of how the information has been represented in the collection. Due to the vagueness of the ideas early in the design process, it is difficult to express information needs in the form of keywords.

3 PREVIOUS RESEARCH

Efforts made by researchers in developing systems to support the use of design precedents have made evident the difficulties inherent to the process of describing and representing these precedents adequately [18]. Subjectivity, inflexibility, high maintenance costs and the dependence of expert knowledge are amongst the problems mentioned.

Any system to be develop which does not require a description of the data it contains (metadata) would need to be able to recognize the information it contains. Many systems use natural language processing techniques to make inferences about the content of information contained (as text) and it is still quite a demanding computational task. However, when the information is embedded in an image, then we are referring to what is still considered one of the most formidable challenges of modern computer science: Content Based Image Retrieval Systems (CBIRS) [13]

CBIR systems create automatically a low level representation of the images it contains. This representation is based exclusively on properties of the image, such as colour, texture, shape, spatial relationships, etc. However, the system cannot map this low level representation to the full semantics of the image. Even recognizing that the image contains a TV, a car or a phone, i.e., recognizing a small part of the semantics is still, to a great extent, an unresolved problem.

Even though CBIRS are still quite immature, they contribute to the issue of indexing and retrieving precedents in design in two ways. First, CBIRS allow the automatic creation of a low level representation of the image that is used for indexing purposes. Second, this representation, being based exclusively on properties of the image, is culturally independent and does not represent any particular point of view. In a CBIRS, the search is done using a seed that, in contrast with other systems, is not based on language. If a user needs an image, all what is required is to feed the system with an image that resembles what the user is looking for. This approach is called query by example (QBE) and has been implemented and tested successfully with designers [18]. Figure 1 presents a screenshot of the system used in the studies.

4 LIMITATIONS OF QBE APPROACHES

One of the motivations to use Content Based Image Retrieval System (CBIRS) was the promise of results without having to describe, organize and index each image manually, as is necessary in previous systems to handle design precedents. This allows escaping the subjectivity of the interpretations, escaping the imprecisions of language and avoiding differences in opinion between users. However promising, this approach falls short of fulfilling all the designer's needs for visual information.

The reason is that the algorithms available cannot recognize what the image contains (in semantic terms) but humans can, and with great facility. This ability was reflected in the searching process of the designers in our studies [18]. It is very natural for them to expect living room furniture if using a sofa and a lamp as seeds for a query, because a user can understand that these two are related, and that the common aspect is that they are both elements of a living room. To the QBE system, they are geometrically so different that the results are completely incoherent. Image recognition is useful, but not enough! Similarity to humans is more than geometrical similarity. Semantic similarity is, in many cases, more relevant and easier to identify. Therefore it has been decided to implement a system that searches the database using a QBE approach, but based on semantic similarity.

5 DESIGNERS' EXPRESSION OF INFORMATION NEEDS

A well known fact is that people prefer to ask other people for information rather than using information systems. In Engineering Design, this has been researched demonstrated in various occasions [1]. Expressing information needs is easier when the interlocutor is human. It is possible to use natural language and gestures, the other person helps in re-prahasing the questions and in suggesting who the information in the answer can be used. Moreover, the queries do not have to be precise.

When using information systems, this flexibility to express information needs is supported only to a very limited extent or, in most cases, not supported at all. This is due to restrictions in existing query languages, which are mostly based on keywords. Searching by keywords requires having a clear idea of what is being sought and of how the information in the database is represented and indexed. Due to the ill-defined character of design tasks, and the vagueness of initial ideas, this clarity is rare during the early stages of the design process.

Conclusions from previous studies suggest that it is necessary to keep supporting the queries for precedents using examples as seeds for the queries, but allow the search to be based both on the geometrical features present (search by formal similarity) and by semantic meaning. This does not necessarily mean these queries have to be expressed using keywords, they can still use images as seeds as in the QBE system described above.

Humans have a great ability to make semantic links. That is, to make sense of sets of things and to determine what is there in common among them. For instance, if we consider the objects camera, phone, mp3 player and PDA, it is easy to recognize that they all are portable electronics. It is also easy to know that a PDA is semantically closer to an MP3 player that, say, to a toaster, which is on its turn closer to an iron and a vacuum cleaner. This is the basis for the system described in the next sections.

6 SEMANTIC LINKING

A lexicon organized using semantic relations can provide the engine to drive the search engine sought. In an effort to do so, scientists at Princeton University created WordNet. WordNet is a database mapping the English language (though other projects have mapped other languages as well) into a number of relations. These relations are:

- Lexical relations: Synonymy and Antonymy
- Semantic Relations: Hyponymy and Hypernymy (*cat* is a hyponym of *mammal* and *mammal* is a hyponym of *animal*).

These relations can be used to create a hierarchical semantic structure as usually there is a common super ordinate under which a term exists. Therefore these relations are called familiar relations. People do not need any training to use them in everyday life and that is the true power of representing language this way. Another type of relations that can be used is the part/whole relations, also called meronymy and holonymy. For instance, wheel is a meronym of bicycle because wheel it is a part of a bicycle and bicycle is therefore a holonym of wheel. These relations can be used to construct part structures. Since products have very complex relations of parts and components, these meronymy/holonymy relations are very limited in WordNet database. It is possible though to use existing ontologies to extend the amount of knowledge the database has about parts and components.

Lastly, words can be written using different morphological inflexions (child, children). These two words refer to the same concept (child) which is in the database (though children is not). The system is able to recognize these morphological inflexions and put the elements in the same synset. To illustrate, consider the following example of the noun car:

Car (Noun)

- *Car is a kind of … (hypernyms)*
  - [1] motor vehicle with 4 wheels
    - automobile, machine, motorcar
  - [2] wheeled vehicle adapted to the rails of a railroad
    - rail car, railway car
  - [3] conveyance for passengers on a cable railway
    - cable car
- *… is kind of a car (hyponyms)*
  - auto, automobile, machine, motorcar
  - ambulance
  - cab, hack, taxi
  - convertible, coupe, station wagon, hatchback, hardtop
  - jeep, Landover, limousine, limo
  - […]
- *car is a member of … (member holonyms)*
  - railcar, railway, train, railroad train
  - gondola, cable car
  - elevator car
- *car is a part of … (holonyms)*
  - airship
  - elevator, lift
  - cable car
- *… is part of a car (meronyms)*
  - accelerator, air bag, engine, bumper
  - car door, roof, suspension […]

Besides nouns, WordNet can also handle adjectives, verbs and adverbs, which have their own relations. However, for the implementation of the software described in this paper, only nouns were considered. The reasons for this are twofold. First, using a QBE approach using image recognition was decided as a measure to eliminate the need to add metadata to the images. Since the image recognition technology used hardly allows determining unambiguously which object is depicted in an image, it was decided to add this metadata to the image manually for testing purposes. In the future, this tagging should be done automatically or semi-automatically. Since there is only one noun being attached to each image, the method is more efficient than other systems and as there are no adjectives used, personal interpretations among different users is also avoided. Nouns are less prone to ambiguity when determined by a human.

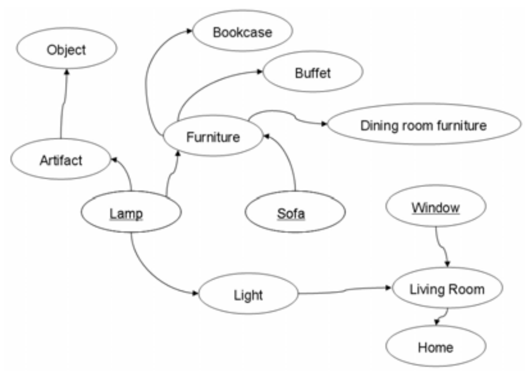

Figure 2 Example of a synset construction using WordNet

In figure 2 it can be seen how using WordNet it is possible to find the common supersets (hypernyms) of the two nouns in the case described before {lamp, sofa}. Lamp is a kind of artefacts and a kind of furniture. Sofa is also a kind of furniture. Therefore all nouns in the furniture synset will appear as related to both lamp and sofa {bookcase, buffet, dinning room furniture, etc.}). The system can search recursively, that is, all images containing either one of the tags in the synset furniture or a related term (a synonym or an inflected form, e.g.) would also appear, as all belong to the same synset.

Other synsets can be pursued depending on how the user gives feedback. For instance, if the user drags to the negative feedback area one or several of the elements belonging to the furniture synset and gives positive feedback to say, a window, the new synset used is living room, then all images related to living room would appear.

7 PRELIMINARY RESULTS

To test the approach, a software implementation was done using the same flash interface and drag-and-drop functionality used in the previous version of the software that uses image recognition. A PHP script was written to query the WordNet database. From the 40000 images in our test database, 1000 were randomly selected and tagged with only one noun per image. Figure 3 shows the results of a query using semantic relations using the same two images used by one of the participants in a previous study (a lamp and a sofa). Note that much of the results are chairs, this is due to the fact that many of the 1000 images tagged were chairs (80 out of a total 150 that were considered furniture). The preliminary tests show that the system is able to successfully construct sets of images by identifying the synsets they belong to. This is a significant improvement over the previous system that uses only image recognition to resolve the queries, as it adds semantic recognition capabilities to the system.

8 FUTURE WORK

The tests with the first system, searching by geometrical similarity have shown that it speeds up the searches and, as results are not necessarily in the same category as the products searched (a hand drill can appear as similar to a hair dryer, for instance), it favours lateral thinking patterns. However, users sometimes also need to search by semantic similarity, or need to restrict the results within one area. Therefore the search engine using semantic similarity was deemed necessary.

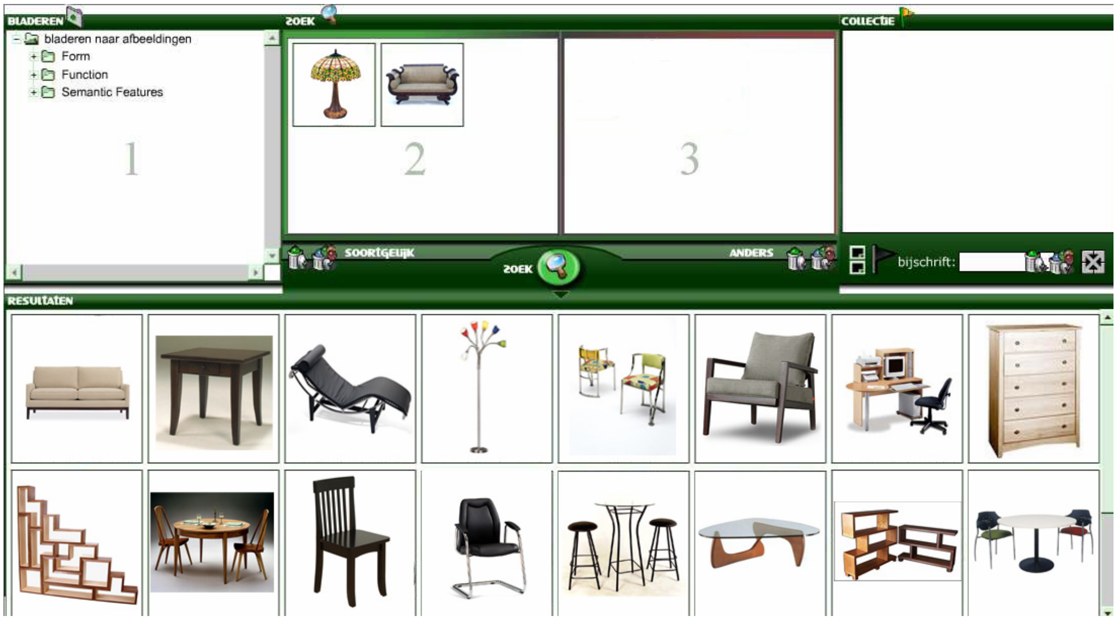

Figure 3 Screenshot of the system using the semantic relations. Note the results do not look similar, but they share the same synset (furniture).

Either system alone would have its own restrictions and limitations. Future work should be done to make the results of a semantic search be combined with geometrical similarity and vice versa. The system should be able to infer users' intentions using the feedback actions (drag and drop to the positive/negative feedback areas). This approach has not been tested with designers yet, but tests will be conducted to evaluate the usefulness of the search by meaning as an added value to searching by appearance.

The use of an image recognition system was motivated by the prospect of not having to add metadata to the images. We found that it was difficult to avoid having a minimum of information (one noun per picture). As image recognition gets more accurate and computers more powerful, this process could be automated, but this has been deemed out of the scope of this research.

It is clear that to be able to create search and get relevant synsets starting with any noun a full map of not only language but products should be created. It has been discussed earlier on this paper that WordNet is limited in the information it contains about hypernyms, hyponyms and holonyms of everyday and technical objects. There have been efforts on creating descriptive ontologies for many technical systems. Such efforts could serve as a basis for enriching or creating complementary WordNets. The purpose of this research has been to make indexing and retrieving images easier, cheaper and neutral (not dependant on the curator of the image collection). The first version of this system indexed images by appearance without the intervention of humans, the version presented here requires one noun per image to be found. Future versions of the system should be able to combine both engines and retrieve images that are both similar in shape and semantically related.

REFERENCES

[1] Ahmed, S. and Wallace, K.M. (2003) Indexing design knowledge based upon descriptions of design processes. in ICED03, 14th International Conference on Engineering Design, Stockholm, Sweden, 691-692

[2] Ashby, M.F. & Johnson, K.W. (2001) Classification and choice in product design., Technical Report CUED/C-EDC/TR108.

[3] Aslandogan, Y.A.; Their, C.; Yu, C.T.; Chengwen, Liu & Nair, K.R. (1995) Design, Implementation and Evaluation of Score (A System for content Based Retrieval of Pictures). Proceedings of the 11th ICED. Pp.280-297.
[4] Clark, R.H. & Pause, M. (1996) Precedents in Architecture. 2nd Ed. Van Nostrand-Reinhold, New York.
[5] Cross, N. (1982) Designerly Ways of Knowing. Design Studies, 3(4), pp.221-227
[6] Flemming, U & Aygen, Z. (2001) A Hybrid Representation of Architectural Precedents. Automation in Construction 10, pp.687-699.
[7] Goldschmidt, G. (1991). The dialectics of sketching. Creativity Research Journal 4(2) p.123-143
[8] Gross, M.D. (1996) The Electronic Cocktail Napking – A Computational Environment for Working with Design Diagrams. Design Studies 17, pp.53-69.
[9] Kuhn, T.S. (1977) Second Thoughts on Paradigms. In Thomas S. Kuhn. The Essential Tension. The University of Chicago Press. Chicago, Il.
[10] Miller, George A., Christiane Fellbaum, Katherine J. Miller. August, (1993) Five papers on WordNet
[11] Muller, W (2001) Order and Meaning in Design. Lemma. Utrecht.
[12] Muller, W., & Pasman, G. (1996). Typology and the Organization of Design Knowledge. Design Studies, 17, 111-130
[13] Müller, W.T.E (2001) Design and Implementation of a Flexible Content Based Image Retrieval Framework. Doctoral Dissertation. Université de Genève.
[14] Nakakoji, K.; Yamamoto, Y. & Ohira, M. (1999) A Framework that Supports Collective Creativity in Design Using Visual Images. Creativity and Cognition 99 pp.166-173.
[15] Oxman, R. (1997). Design by re-presentation: A model of visual reasoning in design. Design Studies 18, pp.329-347
[16] Pasman, G. (2003) Precedents in Design. Doctoral Dissertation. Delft University of Technology. DUP, Delft.
[17] Restrepo, J, Green, W.S. and Christiaans, H. (2004). Structuring Design Problems: Strategies, Hindrances and Consequences for Design Education. Journal of Education in Design, Engineering and Technology. (Forthcoming)
[18] Restrepo, J. (2004) Information Processing in Conceptual Design. Delft University Press
[19] Santini, S., Gupta, A., & Jain, R. (2001). Emergent Semantics Through Interaction in Image Databases. IEEE Transactions on Knowledge and Data Engineering, 13(3), 337-351
[20] Schenider, F. (1994). Floor Plan Atlas. Housing. Birkhäuser, Boston.
[21] Schön, D.A. (1992) Designing as a Reflective Conversation with the Materials of a Design Situation. Research in Engineering Design. 3. p. 131-147
[22] Spiro, R.J.; Vispoel, W.P.; Schmmitz, J.G.; Samarapunga van, A & Boerger, A.E. (1987) Knowledge Acquisition for Application: Cognitive Flexibility and Transfer in Complex Content Domains. In B.K. Britton and S.M. Glynn (Eds.). Executive Control Processes in Reading. Lawrence Earlbaum Associates, London. (1987). Pp177-199
[23] Tsuda, K.; Hirakawa, M.; Tanaka, M.& Ichikawa, T. (1989) Iconic Browser. An Iconic Retrieval System for Object Oriented Databases. IEEE Workshop on Visual Languages. Pp. 130-137
[24] Yi-Luen Do, E. (1998) The Right Tool at the Right Time Investigation of Freehand Drawing as an Interface to Knowledge Based Design Tools

Contact: John Restrepo
Section for Engineering Design and product Development
Mechanical Engineering Department
DTU Byg 404, DK-2800 Kgs. Lyngby
Denmark
+45 45255658
jdrg@mek.dtu.dk
www.kp.mek.dtu.dk/staff/jdrg