The Content-Based Image Retrieval as an Assistance Tool to the Architectural Design

Sabrina Kacher, Jean Claude Bignon, Gilles Halin and Gerald Duffing
CRAI (Research Center in Architecture and Engineering) UMR MAP, N°694 CNRS
Nancy School of Architecture
2 rue Bastien Lepage
Nancy, France.

ABSTRACT

The architectural design requires a research of ideas and a documentation to help the designer in its creation work. It is a domain where the use of pictures (drawing, photography,…) is essential, because the information transmitted by photographic images are often more easily to understand than the one transmitted by texts. The goal of this work is to show the help that can bring the research of pictures indexed by visuals criteria, as colour, texture and shape, in the architectural design domain. If we accept the principle that " an image is better than 10000 words ", we can make the hypothesis that an image research indexed by visual criteria can bring a supplementary help to the designer when he tries to resolve design problems. We tested a research tool resting on image indexation with graphic attributes. Two types of corpora have been used. The first one contains images illustrating building products and the second one shows buildings or parts of buildings, which illustrate the wood architecture domain. The objective of this experiment is to evaluate the relevance of this type of image indexing according to identified users needs. We try to determine which type of visual criteria is the most appropriate to help the designer in the various phases of the design process.

1 INTRODUCTION

The starting point of these research works is the assistance of Architecture design. More precisely, we work on the assistance to the information retrieval relative to the construction systems and building products. The target users belongs to the architectural domain, such as architects, designers… We identify three types of user's needs, which can be answered by developing a system of information retrieval by images:

1. The watching brief need: The user searches for new concepts without any preconceived ideas of the precise nature of the construction system or building products. He only wants to be informed of new solutions available by visualising images.
2. The vague idea need: The user searches for ideas, but he doesn’t know precisely what he is looking for. He visualises images and then during his navigation he starts to progress in solving its design problem.
3 The precise idea but no formulating need: The user knows precisely what he searches for, but he doesn't know precisely how to formulate it. For example the

user searches for an element of roof, but he doesn't know how to name it.

Image is widely used as an information support in the design field because it can convey a great number of information that will help the designer to solve his design problem in a minimum of time. Several researches shows that image bring two kind of information features, the first one concerns its graphic content such as colour, shape, texture (Del Bimbo 1999) and the other one concerns its semantic content relative to the interpretation of the representing elements appearing on the images.

2 SEMANTIC INDEXING FOR IMAGES

To help the designer to find solutions to problems in the different phases of the design process, we have developed an information retrieval system by image. This system allows the user to perform an interactive and progressive image retrieval (Nakapan, Halin, Bignon, Wagner, Humbert, 2000), based on a textual indexing of the images. To index images more efficiently we have developped a thesaurus aimed to describe graphicaly and sementicaly the element illustrated by images. To improve the image indexing, we defined a specific thesaurus in 3 hierarchical levels

2.2. **Thesaurus levels**

The terms are grouped on a functional approach inside 3 levels (Figure 1) (Rosch 1973):

1 The current level: is defined as containing the general terms that express every general architectural concept. This level is also called the basic level; it represents the most important level, because the terms belonging to this level are the most differentiated. Indeed, these terms are more important in the language and are therefore those that we learn in the first place, example (column, beam…)

2 The superior level: represents a more general level. It contains the architectural concepts that are more abstract than the current level. For example "structure vertical system " and "roofing cover ".

3 The lower level: contains more specific architectural concepts than those belonging to the current level. for example " joist " and "bracket"

The transition from a hierarchical level to another is carried out according to some logical rules. The categorisation is performed while leaving the current level or the basic level then operating an " inference " (a column is a vertical structure system). To reach the lower category (specification) is performed by searching for the distinctive features (ex: a "bracket" possesses the distinctive features than a "joist") (kacher, Bignon and Halin 2002). The thesaurus based on this approach is applied to indexing the image representing constructive functions, which we have mentioned previously. This indexing allows only:

1 To identify the elements presented in the images.

2 To complete the images describing with the type of view, the author's name…

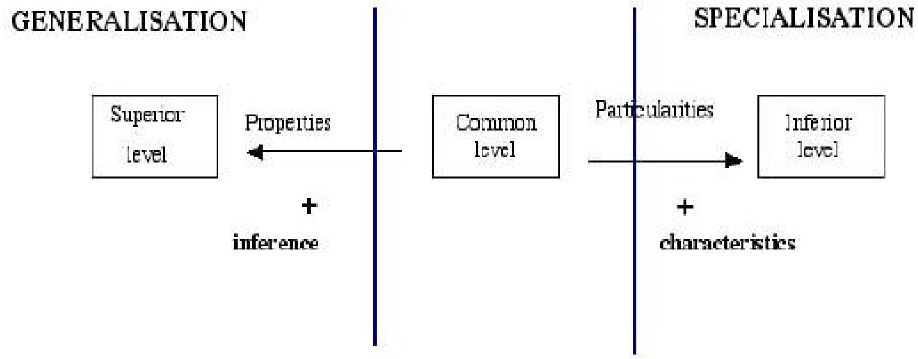

Figure (1): **The hierarchical organisation of the thesaurus**

2.3. **Thesaurus for the semantic indexing:**

The definition of this thesaurus allows to describe the content of an image from a semantically point of view. From this point of view, however, there seems to be some limits of the textual indexing of the image. We should consider for example:

1 The thesaurus terms must be in the common interest of all the persons belonging to the same domain of application.
2 There are some graphics features, which are difficult to be described by words.
3 This approach needs an expert of the domain in order to optimise the description of those represented elements on the image.
4 There is a problem of a language barrier.
5 Thesaurus needs to be structured in a hierarchical organisation in order to represents all the selected concept.
6 Manual indexing is time consuming.

Actually, the image possesses graphics qualities that encode some information more efficiently than the text. To complete this approach and to use in a suitable way the evocative character of the image, we would aboard an other kind of indexing which is: the indexing by visual criteria. The description of the image by its own graphics attributes could be partially automated and would bring:

1 An economy of the indexing time required.
2 The recognition of qualities which are difficult to describe by the text, such as the atmosphere and the architectural concepts in a global way.

To evaluate the implication of the graphics attributes of images we test a content based image retrieval system.

3 THE EXPERIMENTATION

We have experimented an image retrieval system rested on the images content indexing. However, it is important to differentiate two types of content-based images retrieval. The first type refers to the research systems applied to specific image databases. Systems applied to specific data bases belonging to very specific domain such as faces recognition or fingerprint recognition (Boujemaa and Vertan 2001) or more over to object recognition (Ullman 2000) these systems proceed by setting a signature of the specific object to look for on the data base. And the second one refers to the research systems applied to generic image (Vertan and Boujemaa 2000) Generic systems rested on analysing the whole image, not only one specific element as the first type systems. The tested system belongs to the second type, because the images of the corpora are generics and several object are illustrated on one image. The objective of this experimentation is to evaluate the contribution of this kind of indexing for each type of users' needs identified before. This experimentation has been led through a set of tests respecting the following protocol:

Step 1:
This step consists in the selection of a representative image on the database as an **image query**. Images judged similar to these images query, from a visual point of view, are selected manually by the users. The images selected manually will be compared with those retrieved by the system.

Step 2:
Queries to this system are formulated. The system finds the similar images to the image query by the application of the visual criteria

Step 3*:*
This step consists in evaluating the relevance of the system while comparing the images retrieved by the queries with those selected before manually. This evaluation is performed by calculating:

1 Noise: the fraction of the non-relevant images which has been retrieved by the system from the total number of the images retrieved.
2 Silence: the fraction of relevant images which has not been retrieved by the system from all the whole relevant images on the database.
3 Recall : the fraction of the relevant images which has been retrieved from the total number of the relevant images on the database.
4 Precision: the fraction of the relevant images retrieved by the system from the total number of the images retrieved by the system.

3.1 **The image databases**

The images used in this experimentation belong to two different corpora. The first one contains the images illustrating building products (example : brick, roof tile, taps, …). These products are illustrated individually but are often showed in their immediate environment. These kind of images gives more information concerning the use of the product, especially their definitive appearance once integrated into the project. The second corpora contain images representing buildings or parts of buildings belonging to the wood architecture domain (example : post, beam, cladding…). These images represent projects in their environment, or more exactly architectural elements that constituent a constructive function.

3.2 **The tested system**

Before starting the tests, the visual retrieval system needs to index these images with their visual criteria. The system (Duffing 2001) indexes the images in each corpus separately; then calculations on colour and texture for each image are done. Indexing images with their visual criteria rests on two phases. The first one concerns the analysis of all the image on the databases. The system proceeds to characterise images with colours. First, it establishes a list of the colours used by all the images in the corpora. Then, it reduces all the identified colour in a colour palette regrouping the 128 dominant colours in the databases. This palette is reused for the transformation of each image. The colour of each pixel in the image is compared to the colours on the palette and hence replaced by its nearest colour. This process constitutes a simplified "image" as the raw material of the system. These simplified images are used in two distinct techniques of the image retrieval system. When the system lists the dominant colour of the whole image, the resulting technique is called " global"; when the system cuts up images into a parcel of 32*32 pixels, where the calculation is based on. When the system lists the dominant colours and calculate by lot of 32*32 pixels part of the image, the technique is called "local".
To calculate texture the system extract indications of texture for a grey level image derived from a colour image in RGB. It consists of determining the difference of luminance intensity between two successive pixels by taking in account their orientation.

3.3. **Indexing criteria of the tested system**

The visual retrieval system for our experimentation is based on three indexing criteria:

Global colour:
This technique lists and characterises the dominant colours of all the image in the corpora. This characterisation is represented, for each image, by a colour histogram describing the frequency of the 128 colours of the reduced palette. The histogram is created by simply counting the number of pixels representing each colour. In this technique a single histogram is calculated for each image. Two images are said to be similar if they contain the same proportion of near colour in their totality.

Local colour:

This technique uses the same strategy presented above, but at the local level. For every parcel of 32*32 pixels of the image. Two images are said to be similar if they contain the same proportion of near dominant colours locally.

Texture:

This technique called "local" because every part of 32*32 pixels is characterised by indications of textures. Because the rate will be very insignificant and it should be irrelevant to calculate only one global rate because there are several texture on an image. Two images are said to be similar if texture indications met in each image are near.

With these techniques mentioned above six tests have been carried out for each corpus. Three tests are directly related to the indexing criteria and the three others represent a combination of the three previous techniques:

1). Global colour
2). Local colour
3.) Texture
4). Global colour and texture (mixed, with percentage of colour 60% and texture 40%)
5). Local colour and texture (mixed, with percentage of colour 75% and texture 25%)
6). Global colour + local colour (mixed with global colour 60% + local colour 40%).

The main objective of these tests are firstly, to evaluate the relevance of the visual retrieval system, and secondly, to identify the most suitable indexing criteria to answer the users needs identified previously by calculation of relevance weights.

4 THE RESULTS

The results obtained by the tests are presented as a set of 9 images mosaic. These images are sorted in decreasing order of degree of similarity.

4.1 **The constructive function image corpora**

The average of recall and precision varies between 0 and 0,37. We know that these rates for classical research tools vary between 0,6 and 0,8 for recall, and between 0,2 and 0,8 for precision. The best rates obtained for recall and precision have been performed by querying for the picture (Figure 2) illustrating a footbridge in the middle of a lake which is characterised by a big part of landscape. The tested tool was developed to retrieve "generic" pictures. The images selected manually in the first step of the experimentation are showed on the (figure : 3). We can note that the best rates (Table 1) of recall and precision was obtained with the test of global colour (Figure 3) and the mix of global colour + local colour. These results have been obtained because images constituted this bases are complex and information concerning "global" colour of the image is easier to be detected on this kind of images than the features concerning the local colour and texture. For this reason, with this technique, the users' choices correspond better to the systems answers sets.

Figure 2: **The results obtained with the global colour test**

Table 1: **rate of the global colour test**

	Noise	Silence	Recall	Precision
C G	**0,77**	**0,75**	**0,25**	**0,33**

4.2 **Building products image corpora**

When we compared the selected images manually (Figure 5) with those retrieved by the system we obtained rates of recall between 0 and 0,7. Rates of precision going from 0 to 0,7. The average of rates of recall and precision varies between 0,2 and 0,46. These results are more satisfying than those of the constructive functions databased. The best rates (Table 2) have been obtained with the mix of the global colour and the local colour, and with the test unique of the global colour (figure 4). The best results can be explained by the difference of the images used for the two tests.

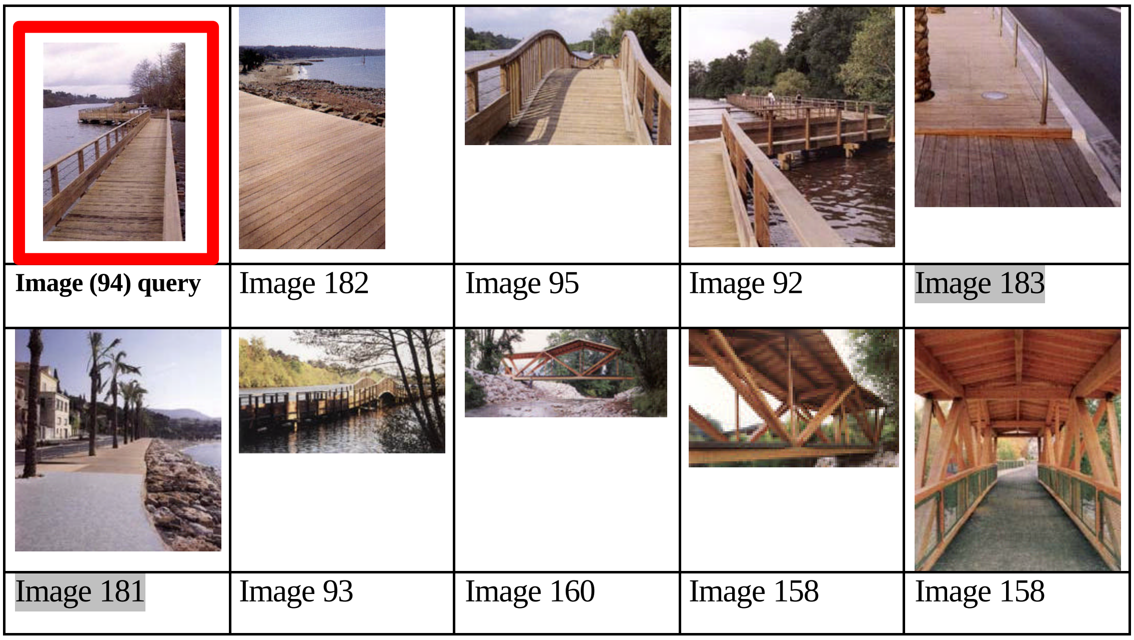

Figure 3: **The image selected manually in the step of the experimentation**

Indeed, the images used for tests of building products are a little particular. They illustrate a single product (roof tile, brick, door…) centralised and contrasting strongly with their backgrounds.

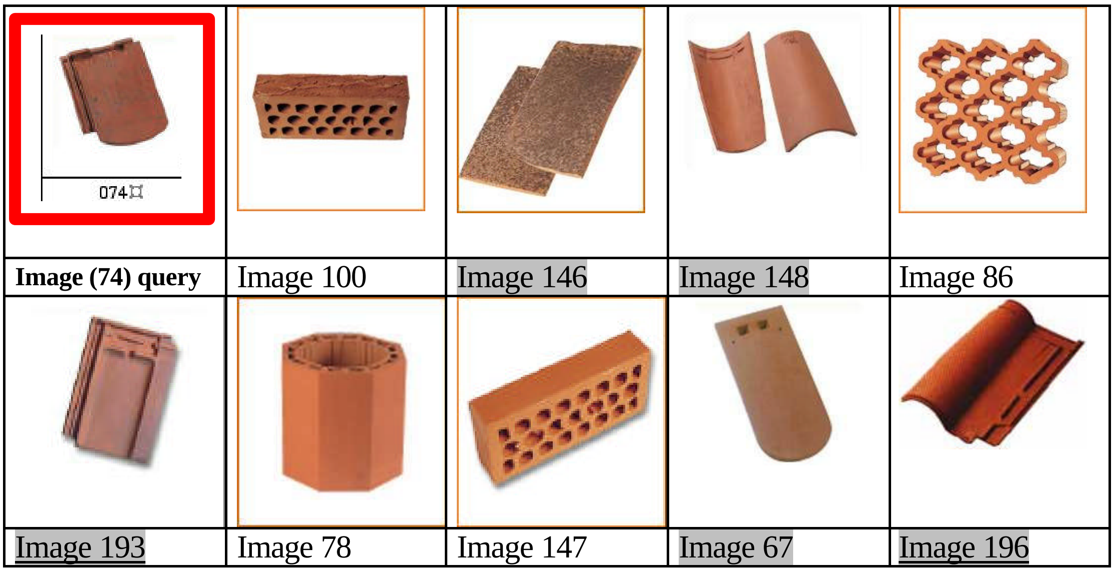

Figure 4**: Test results of global colour**

Table 2: **Rates of the global colour test**

	Noise	Silence	Recall	Precision
C L	**0,3**	**0,3**	**0,7**	**0,7**

				
Image (74) query	Image 66	Image 67	Image 83	Image 211
				
Image 111	Image 146	Image 148	Image 193	Image 196

Figure 5**: The image selected manually in the step of the experimentation**

5 CONCLUSION

In this article we have presented our approach to help the designer in its creation work. The image research using textual indexing can answer to a certain type of needs of information, especially when they are noted down in terms and the structure of the thesaurus. To enlarge this solution, a content based image indexing can bring help, particularly when the user's need is not very precise. Our experimentation shows the limits of those approaches. Only two kinds of needs are concerned; the watching brief and the vague idea need. A research of very specifics building products and a research of ambiences by the global colour could be guided by the graphical characteristics of the image. The solution that we are studying at the moment aimed to integrate some graphics criteria to the currently textual indexing in order to increase the relevance of our system.

6 REFERENCES

N. Boujemaa and C. Vertan, (Boujemaa and Vertan 2001), Integrated Colour Texture Signature for Image Retrieval, ICISP 2001- Agadir - Morocco

Alberto Del BIMBO, (Del Bimbo 1999), VISUAL Information Retrieval, University Press, Great Britain,1999

G. Duffing, (Duffing 2001), Thematico-Visual Image Retrieval : How to Deal With Partially Indexed Corpora, Internet Imaging II, San Jose, California, 2001

Gerald DUFFING and M.SMAIL, (Duffing and Smail 2000), A Novel Approach for accessing Partially Indexed Image Corpo, In 4th Int. Conf. Visual Information Systems-- VISUAL'2000, 244--256, Lyon, France, 2000

S Kacher and J C Bignon and G Halin, (Kacher and Bignon and Halin 2002), research by image applied to the image based reasoning, The Ninth International Conference on Computing in Civil and Building Engineering, Taipei, Taiwan, April 3-5, 2002.

W.Nakapan, G. Halin, J-C Bignon, M Wagner and P Humbert, (Nakapan, Halin, Bignon, Wagner, Humbert, 2000), Building Product Image Extraction from the Web, International Conference on Intelligent Technologies, Intech'2000, Bangkok, Decenber 13-19 2000.

Eleanor Rosch, (Rosch 1973), Natural Categories, Cognitive Psychology, 1973.

Shimoun Ullman, (Ullman 2000), High-Level Vision ; object recognition and visual cognition, the MIT Press, London England, 2000.

C. Vertan and N. Boujemaa, (Vertan and Boujemaa 2000) Spatially constained colour distributions for image indexing, International Conference on Colour Graphics and Image Processing CGIP2000, St. Etienne, 1-4 Oct. 2000.