An image based retrieval system for engineering drawings

Chapter · January 2012

2 authors:

Renata Maria Abrantes Baracho
Federal University of Minas Gerais

Beatriz Valadares Cendon
Federal University of Minas Gerais

Renata Maria Abrantes Baracho,
Beatriz Valadares Cendón

Chapter 13. An image based retrieval system for engineering drawings

Abstract: The study presents the conceptual model, the classification scheme and the prototype of an image retrieval system for engineering drawings. The model integrates concepts and techniques of information science and computer science, joining human interpretation and automated processing, and using textual and visual metadata as access points for retrieval. The paper discusses the tests and the case study used to optimize and validate the system proposed. The tests, using a database of 332 drawings, show the feasibility of the system which presented 90.3 % recall using a similarity rate of 70 % on average. Human and automatic processing for indexing and the classification scheme were important factors in reducing processing time.

Keywords: Image retrieval, image based retrieval system, image retrieval system, concept based retrieval, engineering drawing retrieval system, information retrieval system, technical drawings, engineering drawings, technical drawing indexing, visual metadata

Renata Maria Abrantes Baracho (corresponding author), Assistant Professor, School of Information Science, Federal University of Minas Gerais, renatabaracho@eci.ufmg.br
Beatriz Valadares Cendón, Associate Professor, School of Information Science, Federal University of Minas Gerais

Introduction

The motivation for the current study was the difficulty in information retrieval of engineering drawings required for decision-making, as well as the need to achieve greater efficiency in the process of encountering information in large sets of drawings. Current information retrieval systems for engineering drawings use textual data to index and retrieve drawing files, and do not normally consider visual data. Information retrieval in data-bases that may contain thousands of drawings frequently depends on someone who participated in the project, and it is impaired when that person is no longer part of the process.

First, this chapter surveys concepts, techniques and research in image retrieval, from the perspectives of both information science and computer science. Next, the chapter discusses details of the categories and components of engineering and architectural design including explanations about the geometric representation of objects. Further elucidations on the contents of an engineering drawing that allow its interpretation by humans follow.

The information above constituted the basis to develop the conceptual model for indexing and retrieval of engineering drawings which underlies the system proposed. The chapter details the theoretical basis, the conceptual model and the classification scheme, briefly discussed in Baracho and Cendon (2009), and expands this previous publication by presenting the development of the prototype which implemented the model and the final validation of the system through a case study. The prototype brings together knowledge from three different areas resulting in a system which integrates concepts of information and computer science as well as engineering as shown in Figure 1.

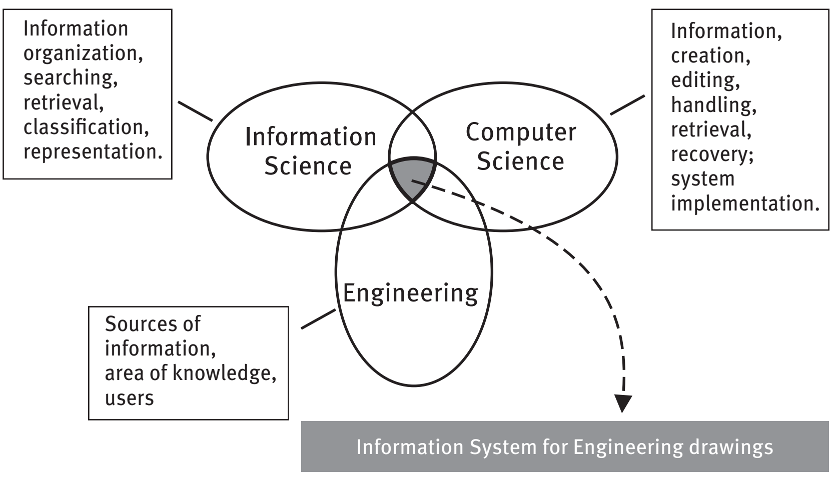


Figure 1. Contributions of three disciplines for the model of a system for image retrieval.

Literature review

The literature review led to two different approaches in image indexing and retrieval research. The first, based on the concepts and foundations of information science, uses descriptive data to index and retrieve information in images. The second, based on concepts and foundations of computer science, achieves content based information retrieval through the graphical properties and shape of the image.

Organization and retrieval of images in information science

In information science, subject analysis, definition of access points, interpretation, categorization, classification, and indexing of documents are used for the organization and retrieval of information. To describe a document, subject analysis, or the determination of the concepts that represent the content of a document, is conducted. The description of a document requires the condensing of its content into the concepts it contains, considering its most perfect conceptual core. The concepts are then translated into indexing terms, which become subject descriptors after being incorporated into an indexing language. Indexing terms, as well as classification schemes and categories are metadata that represent the concepts and subject of a document (Frohmann, 1990). The interpretation, description, and representation of the documents that will be part of the system are part of the construction of online databases. In the process of representation, the document or set of documents in a database can be replaced by a set of metadata to enable its efficient location and retrieval by the user (Baeza-Yates & Ribeiro-Neto, 2011). The metadata used to represent the electronic documents in databases containing text, images and other media constitute access points to the documents. According to Hjørland (1992), the definition of access points is one of the problems in retrieval.

Information retrieval is done through the definition of the documents within a collection which contain the keywords of a user's inquiry. Often, this is not enough to satisfy the need of the user seeking information about a particular subject and not on a particular data or word. In an attempt to satisfy the information needs of the user, the information retrieval system seeks to interpret the content of the information items in a collection. Interpretation involves the matching of syntactic and semantic information in the documents with the information needs of the user (Baeza-Yates & Ribeiro-Neto, 2011). The retrieval process is complete when the user is satisfied with the results of his search.

Most existing image information retrieval systems are based in information science concepts and utilize a textual model in describing this information. Nonetheless, due to the fact that different people comprehend a figure in a number of different ways, depending on the context, the description of an image may vary from one individual to another. In addition to this, all the problems inherent to the different vocabulary used for indexing the images arise. In a system that manipulates images, it would be ideal if the user described the search question using images. In this case, the model would lose less in terms of abstraction and would be able to retrieve more relevant information.

Organization and retrieval of images in computer science

In computer science, content-based image retrieval (CBIR) is achieved by detection of the image features as well as identification and classification of its visual characteristics. CBIR systems usually use algorithms to test all or part of an image in order to identify similar images. The system detects the visual characteristics of the image based on colour, texture and form, and it classifies these characteristics, which are used in feeding the database to retrieve the image desired. Retrieval is attained through comparison and detection of similarity between the visual content of the searched image and the images in the database. In order to begin a search, the user selects the characteristic he is looking for, and defines a similarity measure. The image searched for can be defined by the user or obtained from an example as shown in Figure 2. Zachary, Iyengar, and Barhen (2001) emphasize that the fundamental aspect of image retrieval systems is the determination and representation of a visual feature that effectively distinguishes between pairs of images. Retrieval by picture similarity considers image variation, which enhances the level of information retrieval.

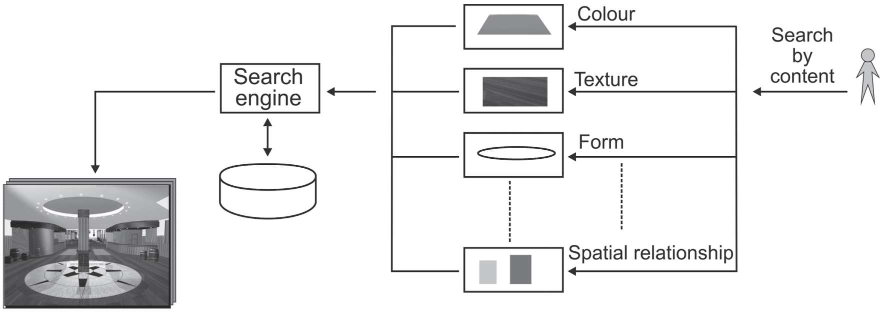


Figure 2. Example of search by colour, texture and form.

The basic principle of CBIR is the use of image visual properties in image search and retrieval instead of a textual description. In addition to this approach, known as the Query by Image Content (QBIC) system (developed by IBM), other systems have been developed with the same goal.

Concepts and techniques of digital image processing are used for detection based on the shape of the object. According to Gonzales and Woods (2002), the term image or monochrome image refers to the two-dimensional function of light intensity, f (x, y), where x and y denote spatial coordinates and the value of f at any point (x, y) is proportional to the brightness (or gray levels) of the image at that point as shown in Figure 3b. It can be considered as an array whose indices of rows and columns identify a point (Figure 3a) in the array. The corresponding value of the array identifies the gray level at that point.

Currently, CBIR systems present search features for organization and representation of images based on syntactic interpretation and recognition of attributes. The systems do not present solutions for defining the concepts carried by an image, which is, at this point, dependent on human interpretation.

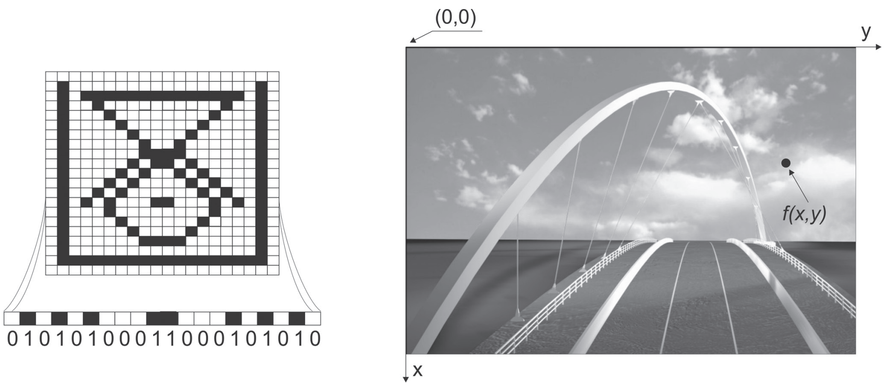


Figure 3. (a) Image matrix. (b) Axes representation digital image.

The gap

Enser (2000), Heidorn (1999), and Smeulders, Worring, Santini, Gupta, and Jain (2000), point out that there is a gap between these two approaches that must be filled in the search for innovative solutions in image retrieval. According to Enser (2000), the collections of images consider the paradigm of retrieval based on the concept in which the image search is verbalized by the user and solved by means of text operations. The text is the verbalization of the image. The author points out other research involving content based visual retrieval, comments on the growth of this line of research since 1990, and underlines the importance of research on image retrieval. This research should investigate the use of hybrid systems in order to get beyond the lack of continuity that exists between high level human processing and low level computational processing.

The study carried out by Smeulders et al. (2000) analyses 200 articles about image retrieval. It concludes that there is an obstacle to overcome in the process of image comprehension, in which computational researchers should be able to identify characteristics required for the iterative understanding of the subject of the image, in addition to the use of automatic techniques for image feature comparisons. The authors point out the growth of research in content based image retrieval (that is: form of the image) due to the greater availability of digital sensors as well as the strengthening of information made available on the Internet and the drop in prices of storage mechanisms. However, as the author notes, it is necessary to surpass the problem of image comprehension.

To the present moment, the gap between high-level semantic concepts of the images and the low-level visual features remains an unsolved challenge. In an effort to better understand the problem, Colombino, Martin, Grasso, and Marchesotti (2010) draw on ethnographic studies of design professionals who routinely engage in image search tasks. Other examples of recent research studies are presented by several authors who refer to the gap issue and proposed various approaches to address it. For instance, Ajorloo and Lakdashti (2011) and Rosa et al. (2008) report on the use of relevance feedback from the user as a way to reduce the semantic gap in a CBIR system. Guan, Antani, Long, and Thoma (2009) and Ion (2009) attempt to deal with the problem through the use of a learning-based scheme for extraction of semantic meaning from image databases. Belkhatir and Thiem (2010) propose the use of three abstract levels of representation to address the problem. Lakdashti, Moin, and Badie (2009) and Ciocca, Cusano, Santini, and Schettini (2011) propose fuzzy rule-based methods to overcome the issues.

Thus, the lack of connection between the two distinct lines of research for the retrieval of images remains. One line considers the semantic understanding of the image and the determination of the concepts that represent the documents. In this line, the description of the image is done through text. The other line of research considers the syntactic processing of the image, seeking to define the contents and properties of existing symbols and icons through digital processing of the image. Both methods present shortcomings when used alone.

In the current study, these two concepts were merged and used as the basis for a model, a classification scheme and a prototype of a retrieval system for engineering drawings. Currently, most engineering drawing retrieval systems do not incorporate the visual content of drawings, and utilize just the textual attributes to represent the documents. Rather than substituting one technology for another, it is relevant to conduct research which permits the enhancement of existing technology, allowing them to consider visual attributes in addition to textual attributes and, therefore, making them more efficient.

The engineering drawing

This section presents an overview of the object of this research: the engineering drawing. An engineering drawing is a type of technical drawing which communicates, through a graphical language, the ideas conceived by the designer of an object and the worker who will build it.

There are specific issues in the organization of information in engineering technical drawings. Engineering has various branches such as agronomic, aeronautical, agricultural, food, environmental, civil, computer, electrical, structural, forest, mechanical, mining, naval, production, chemical, sanitary, safety, software, telecommunications, and transportation, amongst others. Each branch has its specificities and characteristics that determine a series of subdivisions. For instance, civil engineering alone involves the architectural, structural, hydraulic, electrical, fire fighting, fire prevention, and air conditioning projects, amongst others. Each branch of engineering produces a set of drawings needed for the construction of an object.

The engineering project is a set of standardized technical drawings needed for the execution and representation of engineering works. It is usually developed in well-defined steps executed in a linear sequence, which include preliminary design, executive project, detailed design, and project presentation. Each step in the development of an engineering/architectural project is composed of a group of drawings which represent the object to be built in different views.

The set of graphical records is the expression or representation of the form, dimension and location of objects, according to the different needs of the different branches of engineering and architecture. The drawing is a code for a language established between the sender (professional from the field) and the receiver (reader of the drawing), enabling its understanding. Its interpretation requires knowledge both from the sender and from the receiver (Chu, 2001).

The representation and interpretation of the technical drawing requires specific training. Engineering drawings use flat figures in representing spatial forms. The technical drawing consists basically of geometric representations such as lines, surfaces and a set of symbols, signs, dimensions, and texts that complement each order. The set of drawings, through the different views, provides a representation for the constructions of the object. The icons present in the technical drawing allow inferences and conclusions about the project as a whole, defining the type of project (e.g. civil), the process (e.g. executive) and the form (e.g. elevation). Usually, it contains information regarding constructive elements, such as the drawing scale, position and measures of walls, doors, and windows, the name of each ambient, and its respective level.

The development of technical drawings follows policies and standards that vary according to each field of engineering. The standardization of documents is an important step in the creation of a graphical language. This standardization is done by means of technical norms which result from the efforts of those interested in establishing technical codes. According to Maher and Rutherford (1997), engineering drawings are created by the use of a convention of graphical elements or a common syntax of symbolism, as shown in Figure 4. This standardized information improves the understanding of the drawing and the collaboration among users involved in the process. The information contained in the draft can be enhanced by the explicit semantics or meaning of the common syntax of the symbolism.

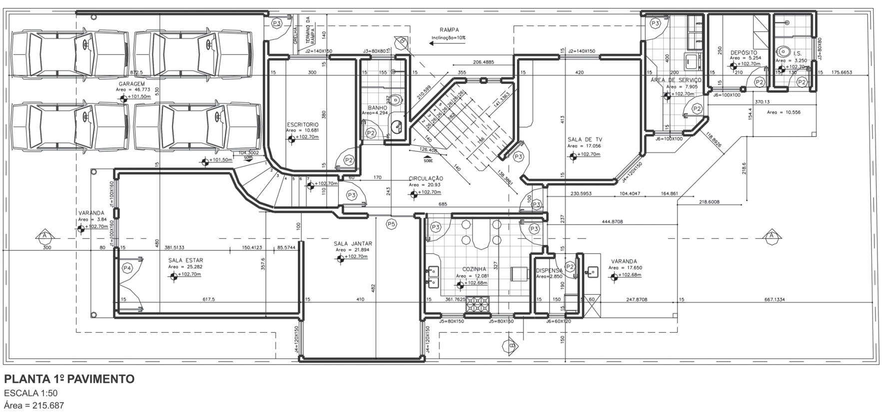


Figure 4. Example of document/technical drawing of an architectural project.

Technological progress and the standardization of drawings made possible the creation of computer aided design (CAD) software, which is now globally used for project development. To increase the performance of CAD software, it became necessary to develop libraries of icons. The libraries, used in project development, are made up of standard sets of icons. They typically represent objects that are repeated in the same or in different drawings and their definition depends on the context. The same library can be used in different projects, institutions, cities and states; however, each institution usually has its own. Figure 5 shows an example of a library of icons.

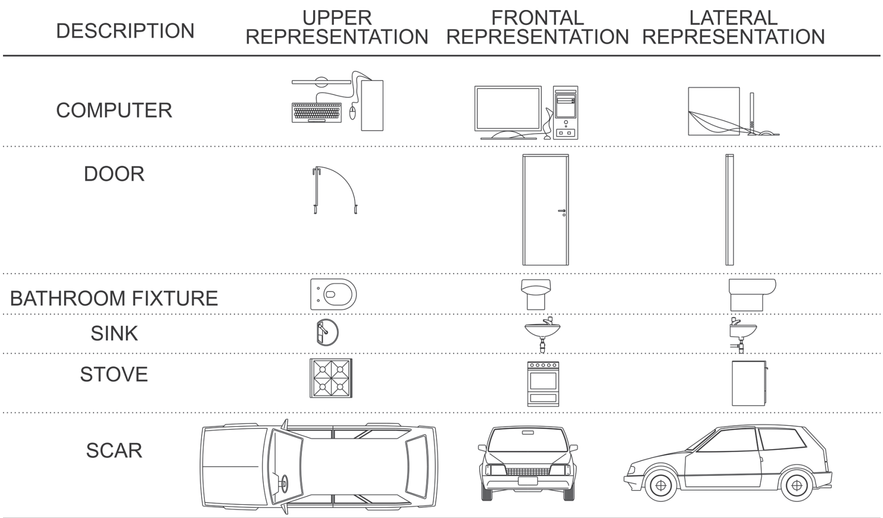


Figure 5. Example of a library of icons.

Conceptual model

The model proposed herein considers human analysis and interpretation of the image as well as automated processing of visual content for the definition of the metadata to be used for retrieval of engineering drawings. Thus, the model considers a hybrid system, consisting of textual and visual data.

To develop the model, the two lines of research described above were considered. The line of research originated in information science considers, for representation and retrieval, the semantics of the image expressed textually through categories or subject descriptors. Computer science considers, on the other side, low level computational processing and interpretation for the representation and retrieval of images. The current research merges these two aspects by proposing the use of textual and visual metadata, as shown in Figure 6.

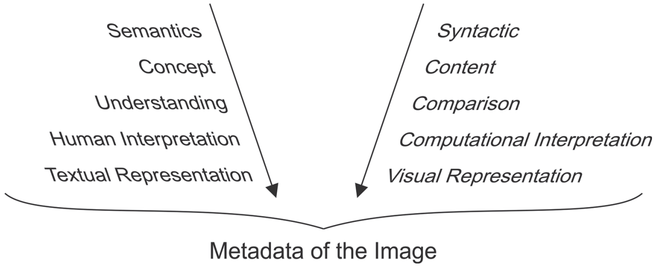


Figure 6. Two types of metadata in image retrieval.

The model proposed presents the interpretation of the drawing in two steps. The first one involves iterative understanding of the drawing, the determination of its subject on the semantic level and its analysis, in order to define what it represents. The second is the syntactic interpretation of metadata, with the definition of administrative, technical, and visual metadata of the drawing, as shown in Figure 7.

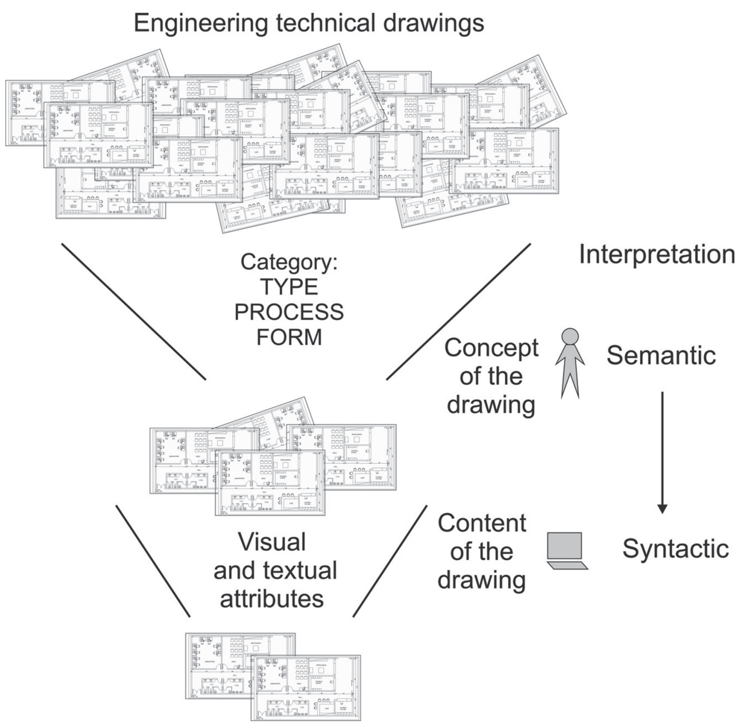


Figure 7. Two steps in drawing interpretation (Baracho, 2007).

Semantic interpretation takes place by means of the inferences made when a human subject observes the drawing and mentally defines the object depicted. It is done through the human definition of three categories (type, process, and form) that define the subject of the document. Syntactic interpretation takes place through the human reading of the textual, administrative, and technical metadata, as well as the automatic processing of the icons present in the document.

The textual metadata is composed of the categories, the administrative metadata, and the technical metadata, and are attributed by a human indexer. Visual metadata is composed of the icons present in the drawing. The identification of icons is done automatically according to digital image processing techniques for extraction of features of form.

Administrative metadata contain data for the drawing that are usually used for administrative control (e.g. number, name, file, company, date, address). Technical metadata refers to the technical characteristics of the project and can be obtained from the stamp or label of the drawing (e.g. scale, title, district, total area, area to build).

The visual metadata are the geometric representation of a symbol and are not contextualized. The attribute of a door, in top view, is simply the representation of a line and an arch. The icon of a door presents different geometric representations according to its position in the drawing. Through syntactic interpretation and the recognition of icons, the definition of the image content is obtained. The icons alone can be seen without considering their context.

The creation of the database starts with the classification of the drawing through the interpretation and analysis of the document by a human indexer who defines the categories, classifies the document, and attributes the administrative and technical metadata. The classification scheme developed is presented in detail in the next section. As Figure 8 demonstrates, the combination of the three categories determines a specific set of icons in a visual metadata table, which might be present in the draft. The drawings are then automatically scanned to locate, index, and store these icons. Drawing files and their textual and visual metadata compose the database.

The first step in the retrieval process is the choice of an image (icon) to be searched for, previously stored in a table of standard icons. To this end, the user defines the three categories (type, process, and form) which lead to the table of icons (see Figure 8). After selecting the icon to be searched for in this table, the user defines the similarity rate between the searched icon and the icons in the documents. Next, the user defines the directory where the search will be conducted. The directory contains a set of drawing files which are then searched for the user's selected icon.

Classification scheme

The following classification scheme was first proposed by Baracho (2007) and is the result of observation and professional experience in engineering companies. The classification scheme encompasses the formal categories of engineering/architectural technical drawings, herein defined as Type, Process, and Form. In classification theory, formal categories have the property of being mutually exclusive, in such a way that each object may belong to only one category at the moment of information organization and retrieval. Thus, the classification scheme funnels and filters technical drawings. According to Hjørland (1992), categories may be defined as access points.

Category: Type

The first category, Type, defines the type of project represented in the technical drawing, among the various branches of engineering. The category Type defines whether the project is architectural, structural, electrical, or hydraulic, amongst others. The project Type defines a set of symbols and representations which may be present in the drawings. By the examination of a technical drawing, a specialist is capable of identifying the type of a given project.

A Type may be:

- architectural: includes the representation of elements to be built (in civil engineering) and the definition of spaces;
- structural: includes information concerning the structure to be executed, pillars, beams and other structural elements;
- electrical: includes information concerning electrical circuits, energy distribution and supply;
- hydraulic: includes information concerning all the hydraulic network to be launched and water and sewage distribution;
- fire prevention: involves information concerning fire prevention and fighting, in compliance with the standards established by the Fire Department;
- mechanical: involves information needed for the manufacturing of mechanical parts;
- others.

The “others” field broadens the classification scheme. In other words, if the project belongs to a category different from those on the list, it will fit in as others. The same others concept can be used in the process and in the form categories.

Category: Process

The second category, called “Process”, contains the development stage. Process is defined by interpreting the level of detail of the project. Most engineering/architecture projects are divided into phases. The cycle of development in a project involves many steps. These steps are usually independent and follow a linear order; when one step ends, one moves on to the next. Project development phases may vary according to the type of project and the type of user. A complete architectural project must contain the following steps:

- pre-project: usually the first study presented in the development of a project. In this phase the parameters and general forms of the project are defined. It is considered a draft and it is subject to changes;
- preliminary project: the project itself, with all definitions ready for execution. Usually, this project is submitted to departments or public agencies responsible for project approval. It is also called a “legal or licensing project”;
- executive project: the project used in the building grounds. It contains a higher level of detail with complexity suited to building procedures;
- detail: the project also used in the execution of building works, contains more information than the previous ones and is usually composed of specific parts of the project on a larger scale;
- presentation project: the kind of project used in presentations to clients and non-expert people. Those are projects containing graphical representations easier to be interpreted, also used for sales and publicity;
- others.

Category: Form

The third category, called Form, defines the graphical representation of the design. The graphical representation of the project is divided into view, ground plan, section, perspective, and others. Each development stage of the project generates a series of technical drawings. These series may contain one or more representations, depending on the position in space of the object to be represented. For example, an architectural project during the development of the executive project may contain some or all technical drawings listed below:

- floor plan: the most representative part of the project. It is the representation seen from above. It defines a horizontal cross section at 1.40m from the ground, and represents information cut and seen on this plane;

– upper view: representation of the object totally viewed from above, as if in a 90º angle in relation to the stand point;
– vertical section: representation of the object vertically cut and containing information concerning the heights of the project;
– elevation: representation of the object seen from outside, in a front or lateral view;
– front view: representation of the object in a front view;
– right lateral view: representation of the object in a right lateral view;
– left lateral view: representation of the object in a left lateral view;
– rear view: representation of the object in a rear view;
– perspective: representation of the object seen from a given angle, defining a vision in perspective;
– others.

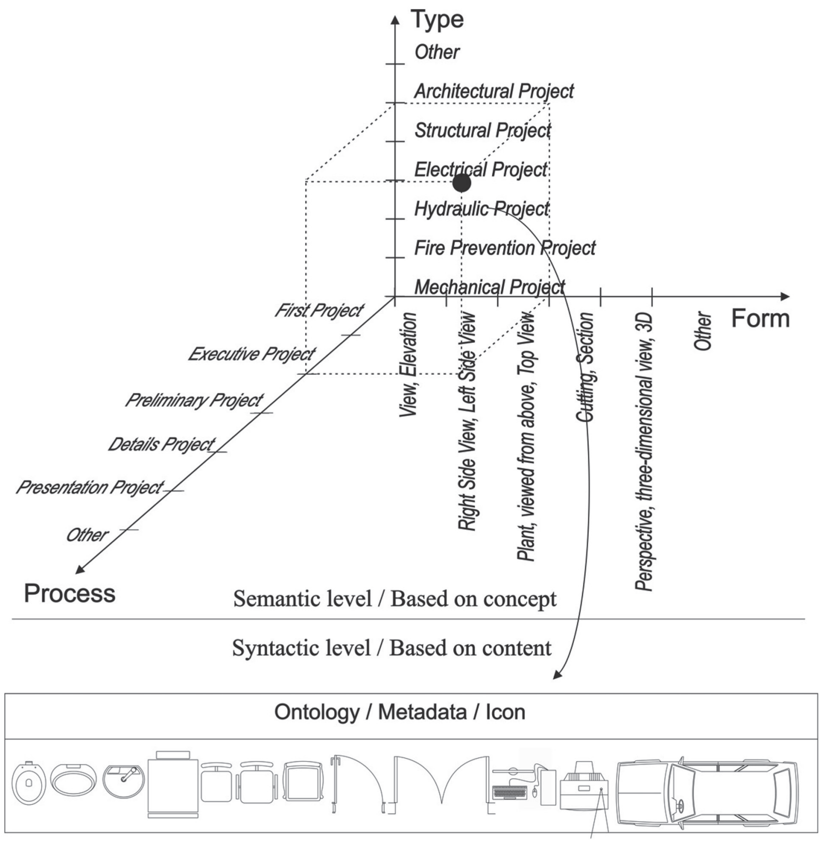


Figure 8. Classification scheme (Baracho & Cendón, 2009).

These categories are used in the model to classify the technical drawing and determine which textual and visual metadata should be used. The combination of the three categories – Type, Process and Form – for a given drawing determines a set of symbols and representations that may be present in the drawing and, consequently, the iconic metadata table that should be used when scanning the drawing. For example, a drawing with Type 'architectural project,' Process 'executive project,' and Form 'floor plan' is classified as the technical drawing of an architectural executive project ground plan. It may contain the representation, belonging specifically to this classification, of an upper view of walls, doors, windows, layout, and impermeable areas. This is different from the attributes of an electrical executive drawing, which contains the symbols for outlets, lamps, circuits, and others. In Figure 8, the three axes of the graphic contain the "others" option, characterizing the openness of the model which can be extended and applied to other categories not discriminated in each axis. This characteristic opens the classification scheme and enables its adaptation to other contexts. Visual and textual metadata are used for indexing and for retrieving documents with consideration of both semantic and syntactic interpretation, both the concept and the content of the drawing.

Prototype

The prototype uses visual and textual metadata for information organization and retrieval. As explained in the previous section, the textual metadata consist of three categories, Type, Process and Form, which a human indexer determines by analysing the drawing and inputting through the prototype interface. These categories are used for conceptual, semantic retrieval. Other textual metadata, consisting of administrative (e.g. number, title, file, company, data, address) and technical (e.g. scale, title, district, total area, etc.) metadata are also implemented in the prototype, and provide additional points of access. The visual metadata consists of icons which may be present in the drawings, according to the iconic metadata table defined for each combination of the categories (Type, Process and Form) for a given drawing. The icons are represented in the drawings as flat, that is, two dimensional monochrome images which consist of arrays of black and white points. The automatic identification of the icons provides content-based retrieval, and it is accomplished through techniques for extraction of shapes or forms of objects. Therefore, the content-based retrieval used by the prototype does not consider colour or texture, since these features are not present in the kinds of engineering drawings used.

The four types of metadata, that is, categories, icons, administrative data, and technical data compose the inverted index used to search the database. Figure 9 shows the complete process described in this section.

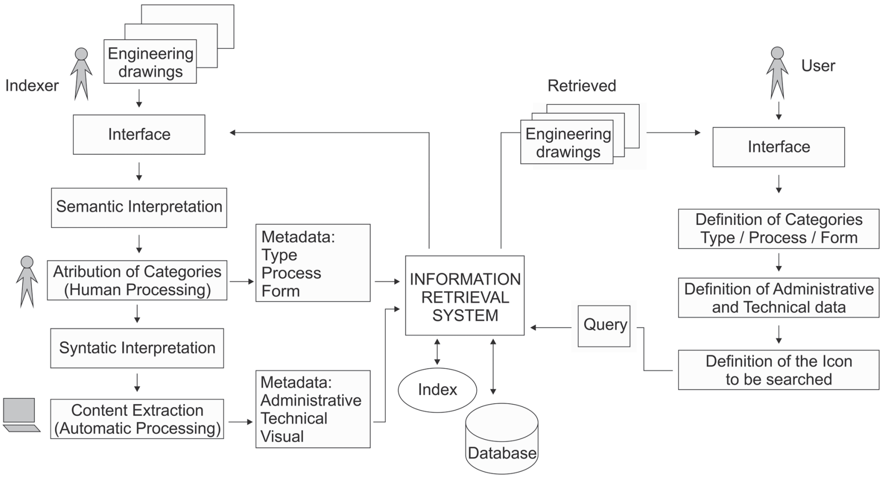


Figure 9. Overview of the proposed system for engineering drawing organization and retrieval (Baracho & Cendón, 2009).

The prototype was developed in the C programming language. For better performance, the graphical user interface was developed in Java. The prototype implemented the model conceptual base in a system which interpreted, classified, and indexed the images in a database. This section presents a general view of the logic and the implementation of the algorithm in the prototype. More technical details on this topic can be found in Baracho (2007).

A sweeping algorithm was developed to locate drawings having icons equal or similar to the searched icon in a database. The general conception of the algorithm is the retrieval of two-dimensional geometric shapes which compose the engineering drawings based on their geometric characteristics. For instance, when looking for the icon of a door, which is represented by a line and an arch (Figure 10), the algorithm searches the drawing for similar arrays of black and white dots, retrieving all the documents where they are present. The central idea of the algorithm is to detect shapes similar to a pre-determined shape, within a larger context.

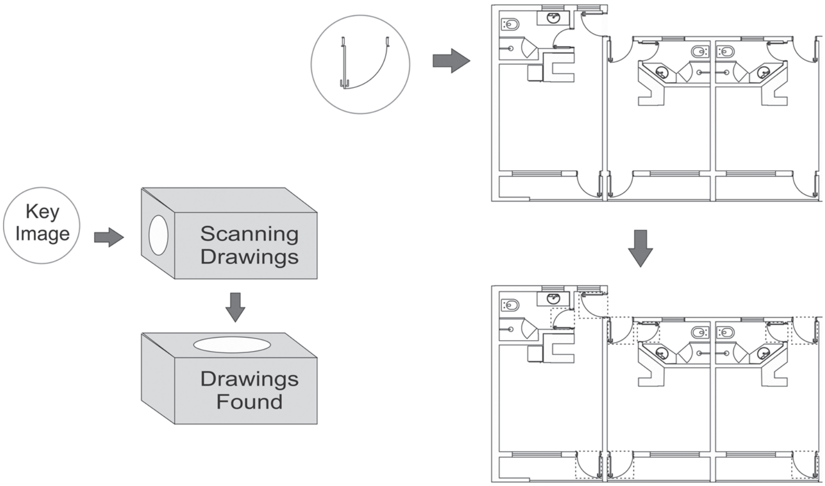


Figure 10. Principle of the algorithm (Baracho & Cendón, 2009).

The algorithm is robust in its search for rapid response time, considering matrix images in vector or matrix format, as well as scales, lines of interference, and rotation of the searched images (Figure 11), independently of the drawing.

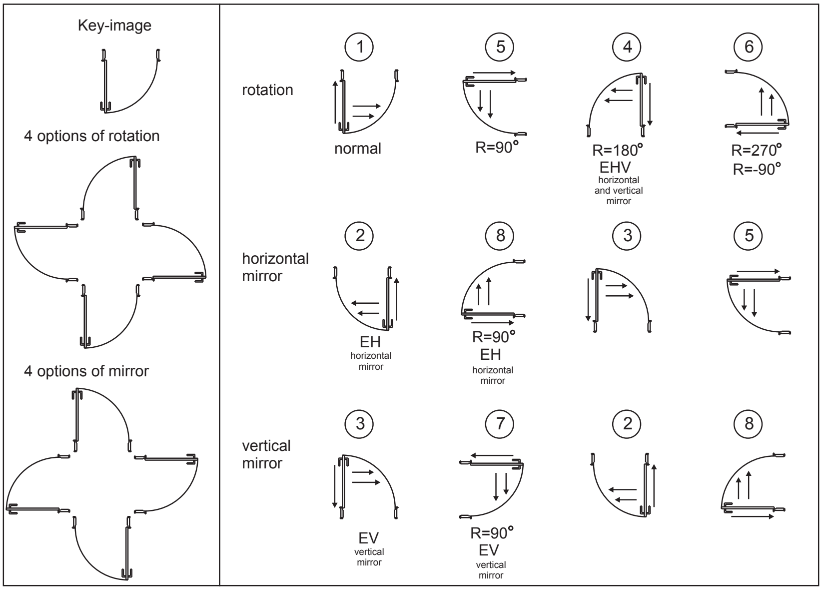


Figure 11. Image with rotation and mirror variation.

It scans the drawings for any image identical or similar to the searched image according to the basic graphic computer transformations, that is, change in translation, scale, rotation, and mirroring. Rotations of 0°, 45°, and 90° were defined. A rotation outside this range in another quadrant is an abstraction of rotation from 0°, 45°, and 90°, and the algorithm is able to automatically recognize this abstraction at 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°, and 360°. The prototype also allows an image to be searched in various scales which are determined by the user. For instance, the user could choose to search for a door in the scales 1.0, 0.75, and 0.50 and rotated in 0°, 45°, and 225°. The basis for the implement of these search features were the classic mathematical principles of geometric transformations for graphic computing (Foley, van Dam, Feiner, Hughes, & Phillips, 1994; Hearn & Baker, 1996), the Fourier Transform (Gonzales & Woods, 2002) and Affine Transform (Chih-Chin & Ying-Chuan, 2011; Mehrotra & Gary, 1995; Mokhtarian & Abbasi, 2002; Yang & Yang, 1999).

The prototype also allows the user to define a similarity rate between the searched icon and the icon found in the drawing as a parameter. The ideal value for the similarity rate would be 100 %, that is, the images should be identical. However, even when this occurs, small differences between the two images may occur due to various reasons such as differences in their resolution or the presence of other lines in the drawing that are not part of the image but overlap with it. The definition of a similarity rate takes this factor into consideration in the search. As shown in Figure 12, the similarity rate is determined by the percentage of pixels that are equal in the two images.

Figure 12. Similarity rate.

Figure 13 shows the interface of the prototype. Through the option “Icones”, in the top menu bar, the searched image can be selected. In Figure 13, the searched image is the icon for a computer in top view. The drawings found are displayed with the hits surrounded by a dotted rectangle. The interface also shows statistical data with the processing time and number of hits for each drawing retrieved.

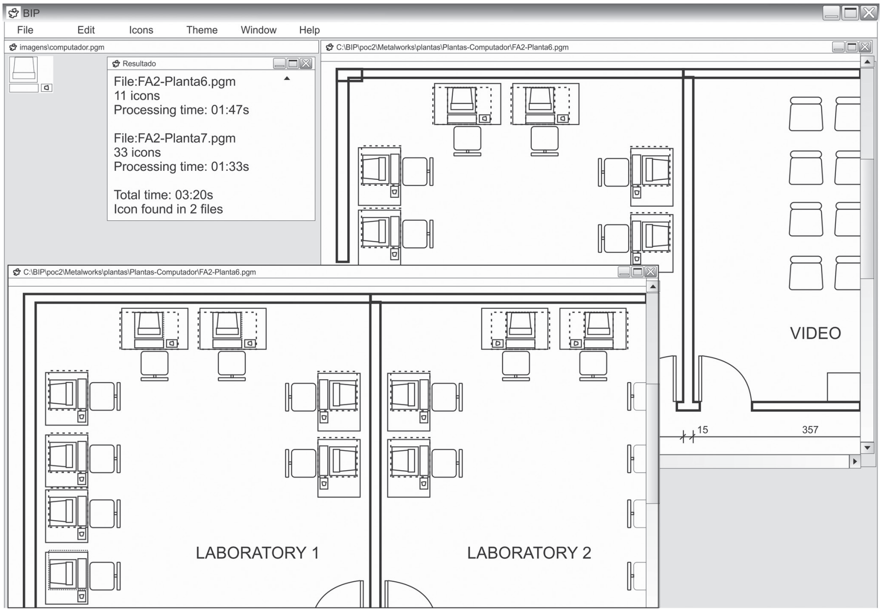
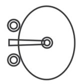
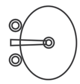
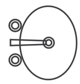
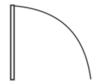
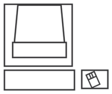


Figure 13. Interface with presentation of the images found (Baracho, 2007).

Tests of the prototype

During the first stage of the research, tests of the prototype were carried out using a corpus composed of 10 engineering technical drawings. Six tests will be discussed here which resulted in optimizations of the system. The results are presented in Figure 14. More details about the tests and optimizations implemented may be found in Baracho (2007).

	Test 1	Test 2	Test 3	Test 4	Test 5
Image					
Pixel	25x29	25x29	25x29	51x51	42x38
Scale	1.0	1.0	1.0	1.0, 0.87, 0.75	1.0
Rotation	0º e 45º	0º e 45º	0º e 45º	0º e 45º	0º e 45º
Rate	70%	80%	50%	70%	70%
	I N R T	I N R T	I N R T	I N R T	I N R T
Drawing 1	2 2 0 43s	1 2 0 43s	2 2 0 43s	8 8 3 6min	0 0 0 2min
Drawing 2	2 2 0 43s	1 2 0 43s	2 2 0 43s	9 10 4 5min	0 0 0 2min
Drawing 3	0 6 0 43s	0 6 0 43s	0 6 0 43s	2 18 3 5min	0 0 0 2min
Drawing 4	0 2 0 44s	0 2 0 44s	2 2 0 44s	8 8 2 6min	1 1 0 2min
Drawing 5	0 1 0 43s	0 2 0 43s	2 2 0 43s	8 10 2 6min	0 0 0 2min
Drawing 6	1 3 0 50s	0 3 0 50s	4 4 0 50s	6 10 4 6min	8 8 0 2min
Drawing 7	2 4 0 45s	0 4 0 45s	4 4 0 45s	5 9 4 6min	25 25 0 2min
Drawing 8	0 0 0 44s	0 0 0 44s	0 0 0 44s	0 0 3 6min	11 20 0 3min
Drawing 9	1 2 0 48s	0 2 0 48s	2 2 0 48s	8 8 2 6min	1 1 0 2min
Drawing 10	1 1 0 48s	0 1 0 48s	0 1 0 48s	2 4 0 6min	2 4 0 3min
Total	9 24 0 451	2 24 0 451	18 25 0 451	56 85 27 58min	48 59 0 22min

I: Images correctly retrieved
N: Total number of images in the drawing
R: Images incorrectly retrieved
T: Provessing time

Figure 14. Results of the tests of the prototype.

Impact of the similarity rate in recall and precision

In tests 1, 2 and 3, the searched icon was a sink. The scale search parameter was set at 1.0, that is, the system should find images that have the same size as the selected icon and should locate images rotated both at 0º and 45º. The similarity rate was set at 70 % in test 1, 80 % in test 2, and 50 % in test 3.

In test 1, 24 icons should have been located. Of these, 9 were found. No incorrect icons were retrieved. Therefore, recall was 37.5 % and precision was 100 %. In test 2 there were 24 icons that should have been located. Two icons were found. No incorrect icons were retrieved. Therefore, recall was 8 % and precision was 100 %. In test 3, 25 icons should have been located. Of these, 18 were found. No incorrect icons were retrieved. Therefore, recall was 75 % and precision was 100 %.

Results of tests 1, 2, and 3 show that a similarity rate of 50 % increases the recall rate and does not affect precision. Test 3 also shows that recall was 0 % for drawings 3 and 10 while for the other drawings, it was 100 %. Examination of these drawings showed that the reason for the low recall was a difference in scale and rotation between the searched icon and the icons present in drawings 3 and 10.

In test 6, shown in Figure 15, the icon searched for is a door. The result shows that a similarity rate lower than 50 % reduces precision, since two icons were incorrectly retrieved.

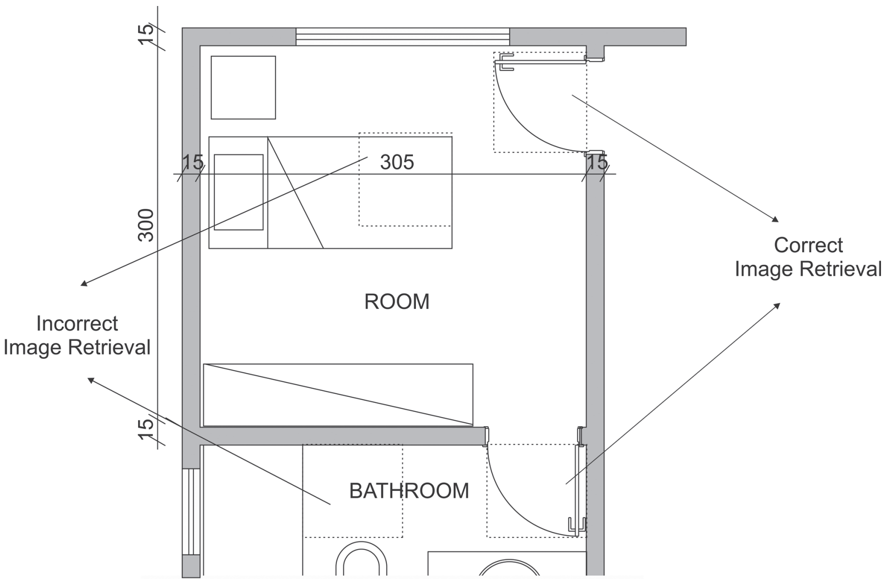


Figure 15. Drawing showing incorrect icons retrieved.

Impact of the parameter scale in retrieval

In test 4, the searched icon was a door. The selected search parameters for the scale were set at 1.0, 0.87, and 0.75. That is, in each drawing the prototype should look for the searched icon in three different sizes, given that in the same building there can be doors of various widths. The prototype should locate images rotated both at 0º and 45º. Because the icon was a simple geometric form, the similarity rate was set at 70 %.

In test 4, 85 icons should have been located. Of these, 56 were correctly found, and 27 false retrievals occurred. Therefore, recall was 67 % and precision was 67.5 %. This test, especially if compared to tests 1, 2 and 3, showed that the search for the same icon in a number of different scales reduces precision. The test also shows increase in processing time if more than one scale is defined for the search. Processing time was 5.8 minutes in average while in tests 1, 2, and 3 it was less than 50 seconds.

Impact of geometric form complexity in processing time

In test 5, the searched icon was a computer, a more complex icon than the one in tests 1, 2, 3, and 4. The scale was set at 1.0 (i.e. the found image should have the same size as the searched image) and the prototype should locate images rotated both at 0º and 45º. The similarity rate was set at 70 %.

In this test, average processing time was 2.2 minutes. This test shows that the size of the searched image in pixels affects processing time. Since resolution and complexity of the searched icon (e.g. number of lines, the presence of curves) influence the size of the image, this finding leads also to the conclusion that resolution and complexity of the icons affect processing time.

The tests above allowed conclusions about factors that affected recall, precision, processing time, and the overall feasibility of the model and prototype. Some of these findings resulted in optimizations which were implemented in the prototype with the objective of improving the performance of the algorithm. The first optimization was implemented to reduce processing time by eliminating the blank spaces. That is, if an area corresponding to the searched image is white, the algorithm jumps a number of columns equal to the length of the image. With the second optimization, the algorithm was improved to ignore the edges of the whole image if a column or row of the image (mask) gets out of the area of the drawing. With the third, the algorithm was modified to ignore the spaces or areas where the image was previously found. That is, if an image has been found in a given position of the technical drawing then this particular position will not be considered in further searches. Other optimizations implemented the transformations for graphic computing, described in the previous section, for better retrieval of rotated images.

Case study: Military Fire Department of the State of Minas Gerais (MFDMG), Brazil

For validation of the model and prototype in a real situation, a case study was developed with the Military Fire Department of the State of Minas Gerais (MFDMG), Brazil, which used projects of fire and panic prevention. The archive of the institution is made up of about 30,000 projects for fire prevention and fire fighting, and each project contains an average of three technical drawings composed of plans, sections, details, specifications and tables.

The corpus used for the validation of the model and prototype was composed of a set of 332 drawings. In the Fire Department's archives, 100 projects were selected and copied with proper authorization to perform this research. The criterion for selection of the projects was the date of approval. All projects selected were approved in 2006 and delivered in DWG format, or developed in AutoCAD.

The icons that comprised the visual metadata for this study, shown in Figure 16, were obtained from the table of graphic symbols in the Legislation for Security against Fire and Panic in Buildings and Risk Areas in the State of Minas Gerais.

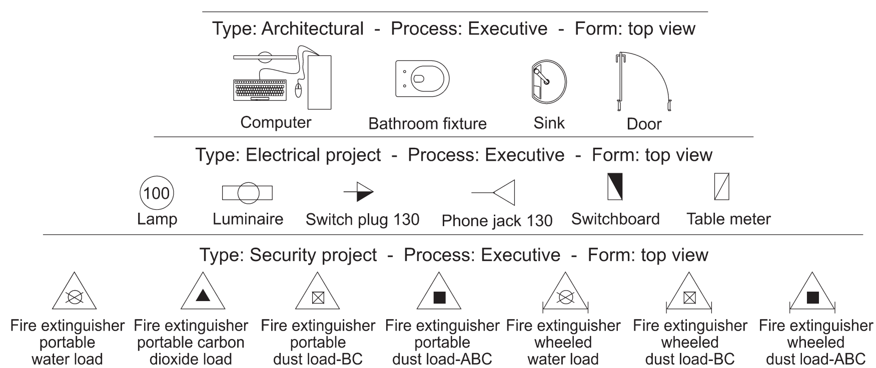


Figure 16. Graphic symbols for safety projects.

After tailoring the prototype to receive the MDFMG data, the graphic symbols table which contained the icons was implemented. Each of the 332 drawings that comprised the corpus was edited in AutoCAD software checking for scales and standard icons, and was placed in a directory. AutoCAD creates the Portable Network Graphics (PNG) image printing file, with parameters of measurement equalling 3.600 x 2.550 pixels. This size was defined after testing which proved its feasibility. Smaller files have lower resolutions which can compromise the performance of the prototype and bigger files occupy far too much memory space and processing time. Afterwards, each technical drawing, in .PNG format, was opened in Irfanview software and stored with a PGM extension which is the format used by the prototype.

Each technical drawing was interpreted to determine the values for the three categories (type, process and shape). For data entry in the prototype, the indexer entered the drawing number, the categories, filled in the textual administrative and technical attributes, and selected the icons to be located.

Test results in the case study database

The database described in the previous section was used to validate the model and the prototype. All 332 drawing files in the database belonged to the same type (fire fighting) and the same process (executive project). In regards to the Form category, they consisted of 213 floor plans, 74 cross-sections, and 45 details. The time to manually categorize all technical drawings was 2100 seconds, that is, it took an average of 6.2 seconds of human processing time to categorize each technical drawing, while it takes 213.30 seconds of computational processing time to automatically accomplish the same task, on average. Thus, the processing time of an expert for the interpretation and classification of the technical drawings is lower than the time a machine would take for the same task, which confirms that the choice for a human indexer was correct.

By defining the category “Form of the Project”, which in this corpus comprises ground plans, cross sections and details, the number of drawings which had to be searched for a given query presented a reduction of 36 % to 86 % (Figure 17). This is representative of the reduction in processing time for image retrieval due to the use of the classification scheme.

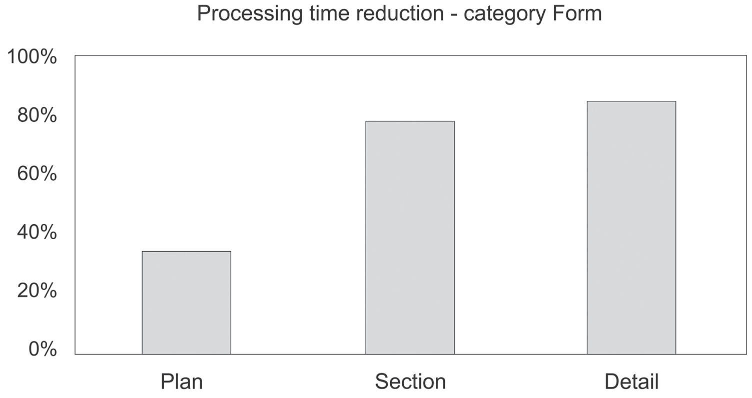


Figure 17. Processing time reduction – category Form.

Two tests of the system will be reported here that demonstrate the success of the optimizations implemented in the prototype based on the results of the first set of experiments. Because the main concern was image retrieval, the administrative and technical metadata were not used in these tests, although they are present in the system. In the first test, the system looked for all seven types of fire extinguishers which are represented as triangles, as shown in Figure 16. Because the shape of the icon was simple, the similarity rate was set as 70. The prototype retrieved 140 of the 155 icons in the drawings that corresponded to the searched image, which represents a recall rate of 90.3 %. No incorrect items were found (100 % precision).

The second test was a repetition of the previous one, with the similarity rate set as 80 %. The prototype retrieved 85 % of the icons in the drawings that corresponded to the searched images and no incorrect icon was retrieved (i.e. the recall rate was 85 %, and precision was 100 %). Examination of the drawings showed that this 15 % error rate occurred mainly because the resolution was low and the prototype could not distinguish the mobile water extinguisher from the mobile powder extinguisher.

Figure 18 shows an example of a technical drawing with the icons found *surrounded by a dotted rectangle.*

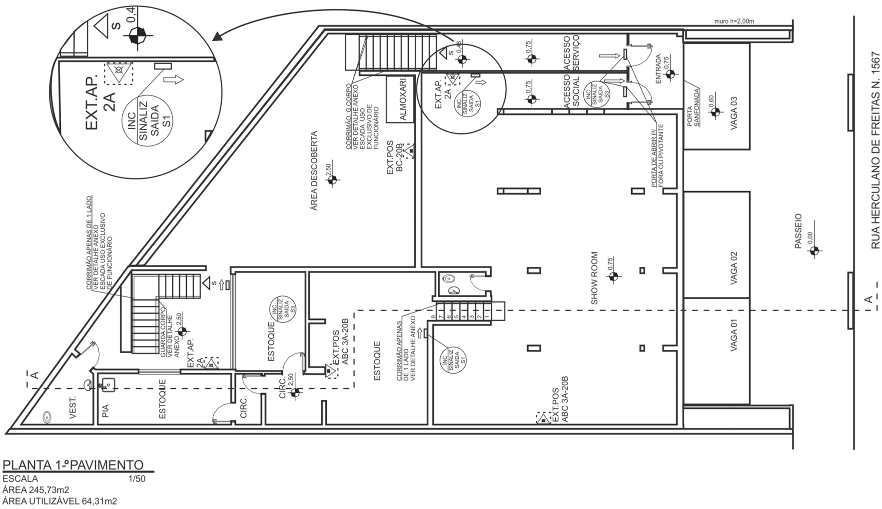


Figure 18. *Example of fire prevention drawings with icons found surrounded by dotted rectangles.*

Conclusions

This study proposed a model for an image retrieval system for technical engineering drawings which was implemented in a prototype and tested in a real database of 332 drawings. The model is innovative in that the indexing of the drawings considers a combination of visual and textual metadata which provides more access points to the user and improves retrieval. In addition to technical and administrative metadata, such as number, author, or area, which are traditionally used for engineering drawing retrieval, the drawings are also indexed by the icons they contain. Both human and automatic processing were used to index the drawings. While the textual metadata (categories, administrative, and technical) are defined and input by a human indexer, the visual metadata are automatically extracted. An inverted index containing the textual and visual metadata was created.

A classification scheme was used at the beginning of the indexing process, which was central to the feasibility of the system. At indexing time, the combination of the three categories for each drawing defines a specific table of icons that may be present in the drawing, therefore reducing the time needed to scan each drawing and extract the visual metadata. At the search stage, the definition of the three categories by the user leads to a table of icons to be selected to be located and also reduces search processing time by limiting the number of drawings that will be searched.

Other important factors for the feasibility of the system were the combined use of a human indexer and the automatic extraction of the visual metadata. This combination offers great potential as it makes the system more efficient using computer processing or human participation where it is most effective. Human cognitive processing was used in the interpretation of the drawing, determination of categories, and indexing of the administrative and technical metadata.

The validation of the model and of the prototype through the case study in the MFDMG shows the time of 6.2 seconds was used by an indexer to interpret, classify, and index the technical drawing. The prototype shows 90.3 % recall using a similarity rate of 70 % on average. Recall drops to 85 % when the similarity rate is set at 80 %. In future developments, in order to achieve a higher level of success rate in the retrieval of icons similar but not identical to the searched image, other optimizations with more accurate functions which consider other variations in scale and rotation of the icons as well as lines of interference should be implemented in the prototype.

Among the theoretical and methodological contributions of this research is the merging of knowledge and research from two areas of knowledge – information science and computer science – enabling the building of a system for retrieving images by images and the creation of a method of development of the prototype involving an iterative process. The resulting system adds human interpretation to automated processing, linking textual and visual metadata and proposing the intersection of techniques and concepts of information science and computer science.

The current study can be applied to the solution of problems concerning the organization and retrieval of information in any field of engineering. The model proposed can be adapted to any institution which deals with engineering drawings requiring only that the parameters, specific metadata tables, categories, and graphical symbols database be defined and adapted for each application. Examples of applications are:

1. How many computers are there in a building or groups of buildings? For this question, the image to be searched would be the icon for a computer. Through this information, the number of laboratories and computers could be defined and factors such as energy consumption and maintenance staff needed could be estimated.
2. How many fire extinguishers are there in a given region and where are they located for direct access in case of an emergency? The system proposed here could provide immediate response for this question, rapidly locating

this data within thousands of drawings which would permit more efficient action by Fire Brigades and public safety services.

Broadly speaking, the system herein proposed can help services and decision making areas such as sanitation, energy, roads, infrastructure, safety, health-care, and education. The model developed can also be adapted to other situations needing retrieval of images.

References

Ajorloo, H., & Lakdashti, A. (2011). A feature relevance estimation method for content-based image retrieval. *International Journal of Information Technology & Decision Making (IJITDM), 10*(5), 933–961.

Baeza-Yates, R. A., & Ribeiro-Neto, B. (2011). *Modern information retrieval*: *The concepts and technology behind search*. New York: Addison-Wesley Longman.

Baracho, R. M. A. (2007). *Sistema de recuperação de informação visual em desenhos técnicos de engenharia e arquitetura: modelo conceitual, esquema de classificação e protótipo.* (Unpublished doctoral dissertation). Federal University of Minas Gerais, Brazil.

Baracho, R. M. A., & Cendon, B. V. (2009, December). *Information retrieval for engineering projects: Using images to search for images.* Paper presented at the 1st International Conference on Information Science and Engineering (ICISE), Nanjing.

Belkhatir, M., & Thiem, H. C. (2010). A multimedia retrieval framework highlighting agents and coordinating their interactions to address the semantic gap. *Expert Systems with Applications, 37*(12), 8903–8909.

Chih-Chin, L., & Ying-Chuan, C. (2011). A user-oriented image retrieval system based on interactive genetic algorithm. *IEEE Transactions on Instrumentation and Measurement, 60*(10), 3318–3325.

Chu, H. (2001). Research in image indexing and retrieval as reflected in the literature. *Journal of the American Society for Information Science and Technology, 52*(12), 1011–1018.

Ciocca, G., Cusano, C., Santini, S., & Schettini, R. (2011). Halfway through the semantic gap: Prosemantic features for image retrieval. *Information Sciences, 181*(22), 4943–4958.

Colombino, T., Martin, D., Grasso, A., & Marchesotti, L. (2010). A reformulation of the semantic gap problem in content-based image retrieval scenarios. *Proceedings of COOP 2010, Computer Supported Cooperative Work* (pp. 45–56).

Enser, P. (2000). Visual image retrieval: Seeking the alliance of concept-based and content-based paradigms. *Journal of Information Science, 26*(4), 199–210.

Foley, J. D., van Dam, A., Feiner, S. K., Hughes, J. F., & Phillips, R. L. (1994). *Introduction to computer graphics*. Reading, MA: Addison-Wesley Professional.

Frohmann, B. (1990). Rules of indexing: A critique of mentalism in information retrieval theory. *Journal of Documentation, 46*(2), 81–101.

Gonzales, R. C., & Woods, R. E. (2002). *Digital image processing*. Upper Saddle River, NJ: Prentice Hall.

Guan, H., Antani, S., Long, L. R., & Thoma, G. R. (2009). *Bridging the semantic gap using ranking SVM for image retrieval*. Paper presented at the Sixth IEEE International Conference on Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA.
Hearn, D., & Baker, M. P. (1996). *Computer graphics*. New Delhi: Prentice-Hall of India Private.
Heidorn, P. B. (1999). Image retrieval as linguistic and nonlinguistic visual model matching. *Library Trends 48*(2), 303–325.
Hjørland, B. (1992). The concept of "subject" in information science. *Journal of Documentation, 48*(2), 172–200.
Ion, A. L. (2009). *Algorithms for reducing the semantic gap in image retrieval systems*. Paper presented at the 2nd Conference on Human System Interactions, Catania, Italy.
Lakdashti, A., Moin, M. S., & Badie, K. (2009). Reducing the semantic gap of the MRI image retrieval systems using a fuzzy rule based technique. *International Journal of Fuzzy Systems, 11*(4), 232–249.
Maher, M. L., & Rutherford, J. H. (1997). A model for synchronous collaborative design using CAD and database management. *Research in Engineering Design, 9*(2), 85–98. Mehrotra, R., & Gary, J. E. (1995). Similar-shape retrieval in shape data management. *Computer, 28*(9), 57–62.
Mokhtarian, F., & Abbasi, S. (2002). Shape similarity retrieval under affine transforms. *Pattern Recognition, 35*(1), 31–41.
Rosa, N. A., Felipe, J. C., Traina, A. J. M., Traina, C., Rangayyan, R. M., & Azevedo-Marques, P. M. (2008). *Using relevance feedback to reduce the semantic gap in content-based image retrieval of mammographic masses.* Paper presented at the 30th Annual International Conference of the IEEE, Vancouver, BC.
Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*, 1349–1380.
Yang, J. D., & Yang, H. J. (1999). A formal framework for image indexing with triples: Toward a concept-based image retrieval. *International Journal of Intelligent Systems, 14*(6), 603–622.
Zachary, J., Iyengar, S. S., & Barhen, J. (2001). Content based image retrieval and information theory: A general approach. *Journal of the American Society for Information Science and Technology, 52*(10), 840–852.