Image Recognition: from Syntax to Semantics

Benedetto Colajanni and Giuseppe Pellitteri

Keywords: image analysis; semantics.

Abstract

In a previous paper the authors presented an analyser of simple architectural images. It works at syntactical level inasmuch as it is able to detect the elementary components of the images and to perform on them some analyses regarding their reciprocal position and their combinations.

Here we present a second step of development of the analyser: the implementation of some semantic capabilities. The most elementary level of semantics is the simple recognition of each object present in the architectural image. Which, in turn means attributing to each object the name of the class of similar objects to which the single object is supposed to pertain. While at the syntactical level the pertinence to a class implies the identity of an object to the class prototype, at the semantic level this is not compulsory. Pertaining to the same class, that is having the same architectural meaning, can be objects having approximately the same shape. Consequently in order to detect the pertinence of an object to a class, that is giving it an architectural meaning, two things are necessary: a database containing the class prototypes to which the recognized objects are to be assigned and a tool able to "measure" the difference of two shapes.

Introduction

Image recognition is a front topics in many fields of the computer science. It has not been of particular interest in the field of architecture, but it can have some applications, e.g. as a search tool in a case base. Furthermore, it poses a certain number of theoretical questions about the way images represent objects and concepts. It seems then worthy pain to do some exploration in this field. In precedent papers [2] [4] the authors showed a tool able to perform a first analysis of the structure of an image, namely a building facade represented vectorially. The tool was able to detect all the figures (named Areal Objects, AOs) present inside the outline of the facade. In this kind of representation all the figures are drawn as simple perimeters made of primitives (lines and arcs). Once all the AOs have been recognized and classified an analysis of the reciprocal positions is operated. The tool recognizes: classes of identical and symmetrical object, disconnects simple objects comprised in more complex objects, detects the nature of the sequences of identical or symmetrical AOs, singling out equal or alternate distances, arithmetical or geometrical progression, symmetries and symmetry hierarchies. This analyser acted in the field of pure syntax as it recognized only topological relationships. This level leaves the following step. Still in progress is the transition to the semantic level giving the analyser the capability of recognizing the kind and the nature of the singled out objects.

The semantic level is very complex specially referring to the problem of the ambiguity of the meaning. Things are easier if in a first moment we limit ourselves to assign only a precise meaning, a sort of label to each recognizable object.

The main instruments to attain this goal are two: a Dictionary of Architectural Objects and a Measures of Shape Differences.

The Dictionary is structured in Chapters, each of which comprises the representation of the objects, pertaining to a certain architectural environment, that is a period, a style and/or a geographic and temporal limitation. Each object is described by its perimeter (with the sequence of the primitives forming it together with the angles that each one forms with its subsequent) and some other information, namely a name, that in the first instance represents its meaning, and a dimensional interval. This second information is necessary specially when object having a simple shape are dealt with, as it is possible that several of them may have very similar shapes (e.g. a quadrangle) but different dimensions.

The measurer of shape differences is based on the representation of a perimeter with the method of the Rotation Function. The comparison of the two rotation functions allows "measuring the distance" between the two
represented shapes.

This procedure allows the recognition of elementary AOs. But besides the meaning of simple objects there is the meaning of complex objects, i.e. objects comprised of more than a simple component (e.g. the capital of a column, made of several mouldings).

The analyser comprises a procedure that builds complex objects starting from adjacent simple objects. The check of the signification of the compound object is done constructing a general perimeter enclosing all the component objects and comparing it with the perimeter of a higher-level architectural object stored in an higher level of the Dictionary.

Next step will be the storage and checking of the topological structures (reciprocal positions of architectural objects) in order to test their possible correspondence to the composition rules of the architectural environment to which they pertain.

Tests have been done of the efficiency of the shape distance measurer and of the semantic analysis applied to Doric temples. The results seem encouraging. Of Them it will referred in the paper.

The measurer of shape difference

The idea of measure of the difference between two shapes is obviously conventional inasmuch as it pretends to transform in quantities an absolutely qualitative thing as shape. In reality what it is about is finding a Metrics in the set of formal representations of shapes. The algorithm employed [1] [5] to establish such a metrics seems to be to some extent reliable as it gives results enough akin to the ones a human observer would give. The AOs are simple closed outlines composed of lines and arcs, that is, rectilinear or mixtilinear objects. The algorithm is based on a particular kind of representation of polygons (regular or irregular and of any number of sides) the rotation function.

The rotation function of a closed outline is obtained as it is hereafter explained:

- A point ad libitum is chosen on the perimeter of the shape as origin of linear co-ordinates along the perimeter itself.
- A direction is chosen ad libitum, as reference direction.
- The perimeter is normalized i.e. is taken as having length 1. Each point of the perimeter is represented in the Cartesian plane with a point having as abscissa its linear co-ordinate, i.e. its distance from the origin point measured along the perimeter, and as ordinate the angle the tangent to the perimeter in the point does with the reference direction.
- The comparison between two shapes is done by superposing the diagrams of the two rotation functions and computing a quadratic function of the absolute values of the areas comprised between the two diagrams.

To further normalize this measure the obtained value is divided by the area of the maximal diagram, thus limiting the difference measure in the interval 0 - 1. As there are three arbitrary choices (a direction and two points) also the computed "distance" can vary in function of them.

The algorithm searches a minimum of the function representing this distance. The actual problem with the algorithm is its being able to cope only with polygons while many architectural objects, also in their simplified planar representation present curved sides. The provisional solution is approximating the curve with a broken line, but we are trying to solve the problem at least for circular arcs. The set of differences computed on a set of shapes satisfies the conditions of a metrics. A not-trivial problem is the evaluation of the extent of difference which still allows to consider the two shapes pertaining to the same class.

Complex objects

In the language of the architectural elements, more than a level of complexity can be individuated. There are some objects that acquire an architectural meaning only if comprised in a complex object. This happens to objects having simple and regular shapes as, for instance, a rectangle. A rectangle having two sides much longer than the others two, can be a step of a stylobate, a fillet, an anulus, a piece of a cornice, and so on. The diversity of dimensions or the peculiar position can give some help which, however, can not be enough. Transferring the analysis to the level of compositions of objects can turn necessary. Once to all the AOs a first level “meaning”,
that is a name has been assigned, the search for possible second level significant objects begins.

The procedure for detecting this kind of object has been shown in [4]. Here it is shortly summarized. Beginning with an AO whatsoever a search is done for possible other AOs having some part of the perimeter in common. If such a couple is detected the check is done if other couples of the same objects exist. If the answer is positive this may signal the repeated presence of a Complex Areal Object (named CAO) having a meaning of its own besides that of the component objects.

Besides the series of composed object also unique composed objects exist. An example: the stylobate of a Greek temple. Also this type of objects is to be recognized.

In a first phase only symmetric object compositions are taken into consideration. In this case the procedure is the following:

- an intrinsically symmetric AO is chosen;
- a check is done of possible the existence of adjacent objects;
- if such an object exists three checks are done:
  - is the dimension ratio between the two objects comprised between the two pre-fixed values?
  - is this second object intrinsically symmetric?
  - has it the same symmetry axis than its predecessor?
  - if yes the answer is yes does it form with the first object a composition the perimeter of which is still symmetric?
- if the answers to the three questions are positive the discovered composition of objects ca be considered a real CAO; then, if one is running a phase of prototype data base construction, a new prototypes is added; if the analysis of an image is performed the usual procedure of recognition of a CAO is run.

The procedure of detecting adjacencies of AOs and taking into consideration more and more CAOs can be repeated as many times as is wanted. The way to check the signification of the CAO is always the same: the comparison with a set of prototypes. There are two possible ways to encode those prototypes. The first is recording as an AO the outline of the object comprised of the adjacent single AO. The second is recording the sequence of the names of the prototypes composing a significant CAO and comparing it with the corresponding sequence of the found CAOs.

Here some caution has to be taken. The algorithm comprises a filter that in the search of adjacent objects discards those whose dimensional difference exceeds a pre-determined quantity. Managing this quantity interactively it is possible to guide the search obtaining a further control of the signification of the found complex objects. An example. The capital of a Doric column is a composed AO. Its lowest component is a fillet. It is adjacent on its lowest side to the shaft of the column: on the higher one to the cushion. If the capital is to be recognized as a significant architectural object, a filter must exist able to privilege the adjacency between fillet and cushion over the one between fillet and shaft. After the recognition of the capital as an object having a distinct architectural meaning has been accomplished, also the adjacency between fillet and shaft, now between capital and shaft, can be taken into consideration leading to the recognition of the complex object column.

The possible existence of more than one level of complexity of objects implies the hierarchical structuring of the data base. The search of an CAO of the nth order goes on at the corresponding nth level of the hierarchical tree structuring the data base.

Once the recognition and the consequent assignment of the architectural meaning of all the simple and complex object as been completed, the syntactic analysis can begin as explained in [2] and [4]. The set on which the analysis is to be performed is composed by all the CAOs at their highest level and by all the simple AOs never entered in an AO composition. The analysis is performed at as many levels as the levels of complexity of the objects are. Of course not all of them are significant. And there is no unique criterion to which resort. This is another and more advanced field of research because it is just the combination of two plans (rather than levels) of syntax and semantics that gives true signification to the structure of the architectural image.

The interactive procedure

As already said, the functioning of the tool requires the existence of a data base in which the prototypes are stored. This can be made through an interactive procedure analyzing a specimen of the architectural environment on which
further analyses are to be performed. The procedure is described thereafter. An empty list which will contain the prototypes is open. Two important choices are done: the maximum distance between two points allowing to consider them as coincident (as errors cannot be avoided in drawing or acquiring drawings with the scanner), and the maximum difference between the rotation functions of two shapes allowing consider them similar and them having the same architectural meaning. A third choice is the initial ratio between the maximum dimensions of two adjacent objects that can provisionally be considered candidates to form a CAO. Then the analysis begins. As the prototypes list is empty the AO first detected cannot be compared with any prototype. Thus the operator is required to name the object and record in the list of prototypes. All the AOs which in turn are detected are compared with the prototypes already present in that list. If recognized as similar to one of those prototypes they are assigned their name. Same objects are put in an unique list. It is worthy noting that more than one list can exist having the name of the same prototype. In fact belonging to one list has a double signification: the sameness of its members and the resemblance to the prototype. Both information is in the same time quantitative and qualitative but only the second has semantic content. After all the elementary objects are detected, the check of the adjacencies begins. To each level of complexity a new ratio is fixed between the maximum dimensions of adjacent objects to be considered possible components of CAO; a new empty list of possible prototypes is created. When a CAO satisfying the imposed dimensional limitations is met, its outline is computed and compared with the existing outline prototypes. The procedure may be repeated as many times as are the complexity levels interactively considered significant.

The automatic procedure

The automatic procedure hypothesizes the existence of a previously created data base. If an AO is encountered which is not recognized as similar to a prototype of the data base, it is taken as a new prototype and put in a provisional list. At the end of the phase the list is analyzed. Each AO can be discarded or transferred, after having been assigned a name, to the database of the prototypes.

An example

As an example a first analysis of the front of the Teseion of Athens is presented. For brevity sake only some steps are shown. The first phases perform the detection of all kinds of AOs.

Fig.1 shows the phase in which some simple AOs are recognized and listed (syntactic).

Fig.2 shows the phase in which nested AOs are recognised and listed (syntactic).

Fig 3 shows the phase in which composed AOs are. Note that at this level the ratio of maximum dimensions of the objects allowed composing a complex AO is rather low (3).

Fig 4 shows a further phase of CAO recognition in which the ratio has been raised (7). This allows the recognition of the column as CAO.

The second phase performs the assignment of a name to each simple or CAO. In the same fig.5 the object is recognized and assigned the name capital (semantic). In fig. 6 the column is recognized and named.

Then the syntactic phase begins again performing the search of series of equal distance (fig.7), progressions and partial symmetries (fig. 8).

Fig 9 shows the synthesis of the symmetries.
Fig. 1. Recognition of simple Areal Objects (AOs).

Fig. 2. Recognition of nested AOs.
Fig. 3. Complex Areal Object (CAO) of 3th level.
Fig. 4. Complex Areal Object of 4th level.
Fig. 5. Semantic recognition of the capital.
Fig. 6. Semantic recognition of the column.

Fig. 7. Recognition of an equidistant sequence.
Fig. 8. Recognition of local symmetries.
Conclusions

This research may seem a mechanical play rather than a construction of a really useful tool. The limitations of the premises are more than evident. No architecture is perceived as a plane image; also if the façade is plane, the architectural objects inserted in it can have a certain relief. The tool has no pretension of performing a complete analysis of the figurative structure of a façade. It is rather aimed at exploring how far it is possible to extend the automatic capability of interpretation of a figure using for that purpose, both the syntactic and the semantic content of it. The 2D representation of an architectural façade is now only a field in which this research can be more easily carried on. As we already stated in [3] the final aim is to realise a tool which can retrieve in a data base an object whose similarity of structure with the structure of a design problem depends not only on separate graphical or alphanumerical information but on a combination of the two.

References


1 Dipartimento di Progetto e Costruzione Edilizia, Università di Palermo, Viale delle Scienze, 90128 Palermo, Italy, tel + 39 91 234100, fax + 39 91 488562, E-mail: bcolajan@mbox.unipa.it, pellitt@mbox.unipa.it.

2 Here there are some technical difficulties depending on the approximation with which the original image has been captured.

Many cases are possible and each case poses a different problem. In the case a), left, the side B - F is to be stored as three segments B-C, C -G, G-F, otherwise the algorithm will not be able to recognize that the two AOs have a part of the perimeter in common. In the case b), right, it is very unlikely that the algorithm recognizes that the two objects have a part of the perimeter in common however small the distance between lines B-F and C-G are.