

Chapter 5 Spatial Reasoning

Figure 5.1(a) shows a picture with a house, a car and a tree. This picture is called a *symbolic picture*, as opposed to an actual image, because it contains objects that have symbolic names: house, tree, car, etc.

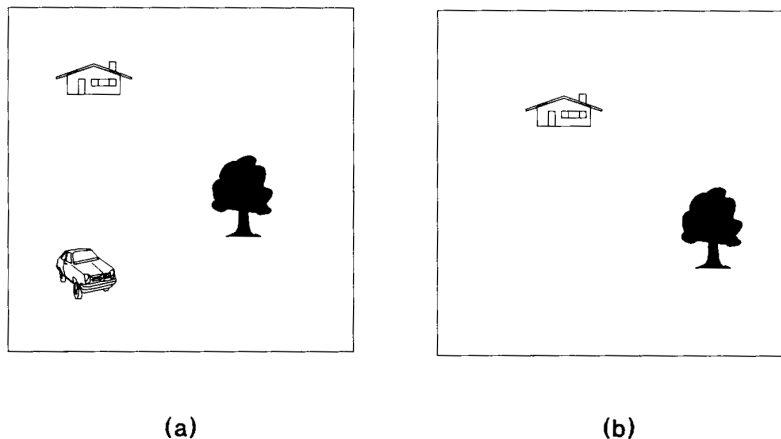


Figure 5.1 A symbolic picture (a) and its subpicture (b).

Fig. 5.1. A symbolic picture (a) and its subpicture (b).

Suppose the objective is to find out whether there is a tree to the southeast of the house. The x-projection of the above symbolic picture can be constructed as follows:

The names of objects in each column of the symbolic picture are projected onto the x-axis. The '<' symbol is inserted to distinguish the objects belonging to different columns. Thus the x-projection is:

x-projection: house car < tree

Similarly, the y-projection is:

y-projection: car < tree < house

Unlike the projections of a mathematical function, the projections of a symbolic picture are strings. A pair of two symbolic projections is called a *2D string*.

The statement "there is a tree to the southeast of a house" corresponds to the symbolic picture shown in Figure 5.1(b). This picture has the following symbolic projections:

x-projection: house < tree

y-projection: tree < house

We immediately notice "house < tree" is a subsequence of "house car < tree" and "tree < house" is a subsequence of "car < tree < house". In this case, the two symbolic pictures can be perfectly reconstructed from the two corresponding pairs of symbolic projections. Therefore, the above statement can be verified to be true, just by checking the subsequence property of the 2D strings involved.

The Theory of Symbolic Projection was first developed by Chang and co-workers based upon the above described intuitive concept. It forms the basis of a wide range of image information retrieval algorithms. It also supports pictorial-query-by-picture, so that the user of an image information system can simply draw a picture and use the picture as a query. Many researchers have since extended this original concept, so that there is now a rich body of theory as well as empirical results. The extended Theory of Symbolic Projection can deal with not only point-like objects, but also objects of any shape and size. Moreover, the Theory can deal with not only one symbolic picture, but also multiple symbolic pictures, three-dimensional pictures, a time sequence of pictures, etc.

A word about terminology. A symbolic picture can have at least two symbolic projections, and in general more than two symbolic projections can be defined for a symbolic picture. We will use the plural form - symbolic projections - to refer to them. However, the capitalized singular form - Symbolic Projection - is reserved to refer to this theory of spatial relations.

In this chapter, Section 5.1 describes what is an image information system. In Section 5.2, we present a conceptual framework for image information system design. Examples are described in Section 1.3 and Section 1.4 to motivate the theoretical considerations.

1. Image Information Systems

Before we introduce our conceptual framework for image representation, image structuring and spatial reasoning, let us first describe what is an image information system.

An image information system is an information system for the input, storage, processing, output and communication of a large quantity of images. An image information system typically has the following five components: (a) image input subsystem, (b) image processing subsystem, (c) image output subsystem, (d) image database subsystem and (e) image communications subsystem. A schematic is illustrated in Figure 5.2.

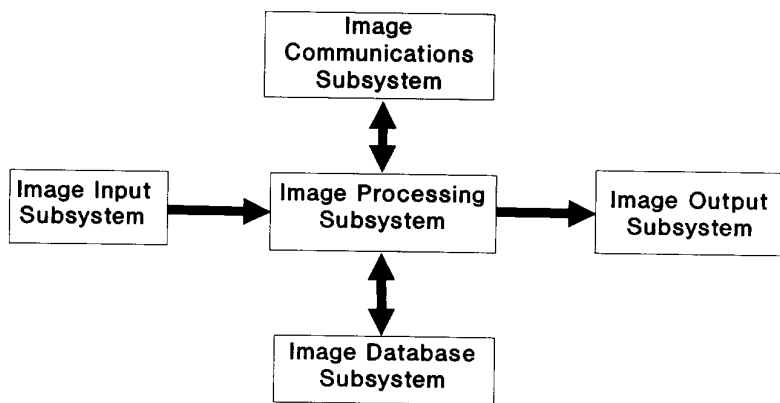


Figure 1.2 Component subsystems of an image information system.

Fig. 5.2. Component subsystems of an image information system.

Recently, advances in image storage technologies have made the creation of very large image databases feasible. Multimedia communications also greatly facilitate the distribution of images across communication networks. Parallel computers lead to faster image processing systems. High resolution graphics and dedicated co-processors enable the design of image output subsystems with superior image quality. The state-of-art image information systems have found their way into many application areas, including geographical information systems (GIS), office automation (OA), medical pictorial archiving and communications systems (PACS), computer-aided design (CAD), computer-aided manufacturing (CAM), computer-aided engineering (CAE), robotics, and scientific databases (SD) applications.

Wider applications also lead to more sophisticated end users. Image information systems, like other types of information systems, have increasingly become knowledge-based systems, with capabilities to perform many sophisticated tasks by accessing and manipulating domain knowledge.

The emphasis on the user has also led to changes in terminology. Increasingly people are using "visual information systems" instead of "image information systems" to emphasize the visual presentation and visualization of stored data, which often include numerical and textual data. For example, the scientific data from a chemistry experiment are not images, but can be visualized by the user as such. The distinction between image information systems and visual information systems tends to be blurred, because image information systems can also handle artificial, man-made images resulting from non-image data. We will stick with the former, although many techniques described in this book often can be applied to both.

So far, image information systems are designed on an ad-hoc basis. The above mentioned technological advances dictate a better methodology to design knowledge-based, user-oriented image information systems. The design methodology, to take into consideration the diversified application requirements and users' needs, should provide a unified framework for image representation, structuring, as well as spatial reasoning. Such a framework will be presented below.

2. A Conceptual Framework

As discussed in Section 1, an image information system typically consists of image input subsystem, image output subsystem, image processing subsystem, image database subsystem, and image communications subsystem. We will now concentrate on the image processing subsystem and image database subsystem, which constitute the heart of the image information system.

A traditional image processing system performs the tasks of image analysis, image enhancement and pattern recognition. Within the context of an image information system, the image processing subsystem and image database subsystem must perform the following three functions, which can also be regarded as three stages in knowledge-based image processing. The three stages are illustrated in Figure 5.3.

Introduction

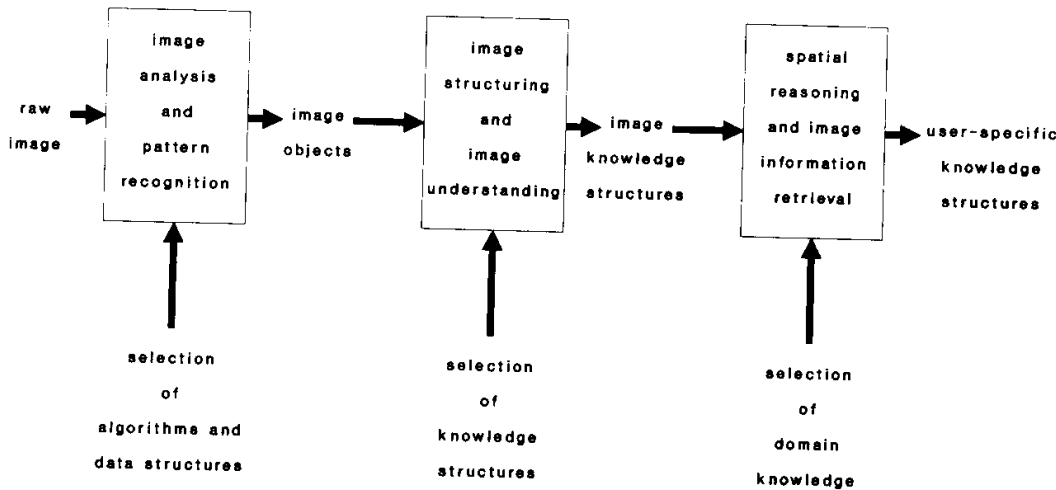


Figure 1.3 Stages in knowledge-based image processing.

Fig. 5.3. Stages in knowledge-based image processing.

(1) Image Analysis and Pattern Recognition: The raw image is analyzed, and the image objects recognized. This stage is almost always present in any image processing system. Extensive techniques are available for image enhancement, normalization, segmentation, and pattern recognition. The end result is a collection of recognized image objects. These image objects are usually encoded in some data structures for ease of access and further manipulation. For example, the image objects may be encoded in run-length codes, polygonal contour codes, quad-trees, oct-trees, etc. The set of image objects constitutes a symbolic picture.

In general, the image objects are objects with attributes including coordinates, geometric properties, etc. The unanalyzed or unprocessed parts of the image can be regarded as image entities which will be analyzed later if needed. Therefore, the system input to this stage includes the selection of various image processing algorithms, and the selection of data structures.

(2) Image Structuring and Understanding: For some applications, it may be sufficient to access and manipulate the image objects, and there is no need for further structuring. However, for many applications, the image objects must be converted into image knowledge structures, so that spatial reasoning and image information retrieval can be supported. The particular image knowledge structure emphasized in this book is based upon 2D strings of symbolic projections. For some applications, 2D strings will be sufficient. For other applications, we must use generalized versions of 2D strings. The 2D strings can also be embedded into hierarchical image knowledge structure called the *S-tree*. Therefore, in this book, the 2D string-based image knowledge structure will be the basic knowledge representation for symbolic pictures. However, it may be desirable to use other image knowledge structures, such as directed graphs of spatial relations, semantic networks, etc. Therefore, the system input to this stage includes the selection of knowledge structures. When 2D strings are used as the image knowledge structure, the system input to this stage includes the selection of cutting mechanism and operator set. **(3) Spatial Reasoning and Image Information Retrieval:** An image information system supports the information gathering needs and problem solving activities of the end users. Some applications require spatial reasoning. Other applications are primarily concerned with image information retrieval. In medical information systems, for instance, the clinician may want to retrieve CAT images of all patients having a tumor similar to the shape present in a specific CAT image. Generally speaking, both spatial reasoning and image information retrieval are needed and may complement each other. To solve specific problems in a domain, a domain knowledge-base is needed. It is also necessary to perform various transformations upon the image knowledge structure, so that the desired image knowledge can be easily accessed and/or manipulated. The transformations include: rotation, translation, change of point-of-view, projection from 3D to 2D views, addition/deletion of image objects, etc. The final result of this stage is a user-specific knowledge structure, such as a navigation plan, a path, a set of retrieved images, an image index or indices, etc.

In the conceptual framework presented above, the three stages can be regarded as three generalized transformations, to transform a raw image first into an image data structure, then into an image knowledge structure, and finally into a user-specific knowledge structure. We first present two examples of image information retrieval and spatial reasoning. Both examples involve path finding in an image. However, the first example is primarily a path finding problem, whereas the second example involves more complex image information processing and spatial reasoning.

3. The Navigation Example of Image Information Retrieval and Spatial Reasoning

Figure 5.4 illustrates a raw image of a coastal region. The specific problem for this application, is to find a path for navigating a ship from point A to point B.



Figure 1.4 An image illustrating the coastal region of the Swedish Archipelago.

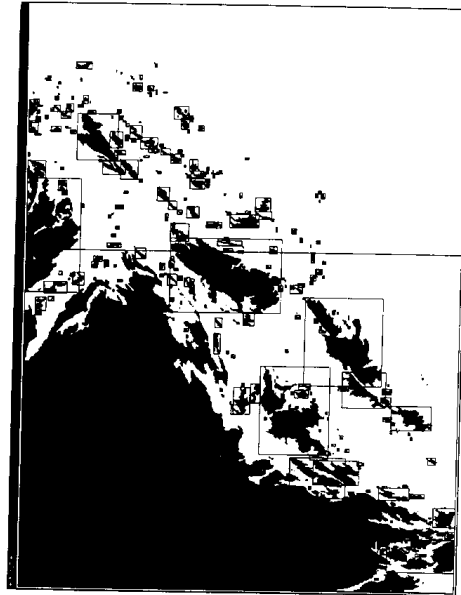


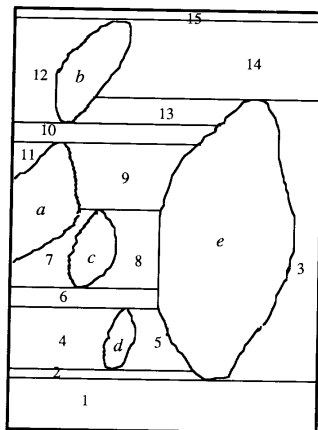
Figure 1.5 Identification of island objects in Figure 1.4, where each object is enclosed in a minimal enclosing rectangle.

Fig. 5.4. An image illustrating the coastal region of the Swedish Archipelago.

Fig. 5.5. Identification of island objects, where each object is enclosed in a minimal enclosing rectangle.

At the Image Analysis & Pattern Recognition (IAPR) Stage, minimal enclosing rectangles for each island object are constructed, as shown in Figure 5.5. The image objects are run-length encoded and stored in the image database. This is the basic image data structure created by the IAPR Stage.

At the Image Structuring & Image Understanding (ISIU) Stage, the run-length encoded data structure is processed, so that a connectivity graph for the "tiles" (the "runs" of white pixels) can be constructed. This is part of the resultant image knowledge structure created by the ISIU Stage. A small symbolic picture extracted from the image of Figure 5.4 is shown in Figure 5.6, where the "tiles" are the "runs" of white pixels. The corresponding tile graph is illustrated in Figure 5.7.



$(a < b < c < d < e, d < c e < a < b)$

Figure 1.6 A small symbolic picture containing five island objects. The 2D string is given below the figure.

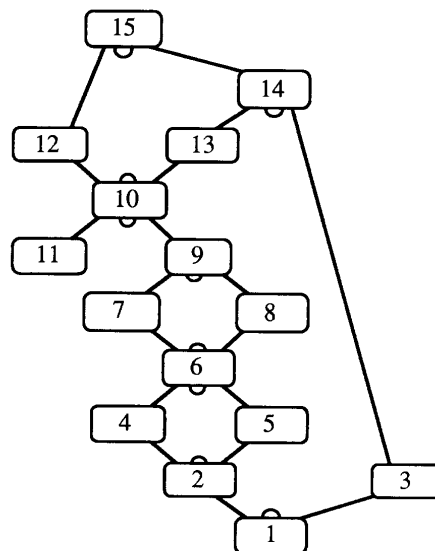


Figure 1.7 The tile graph constructed from Figure 1.6.

Fig. 5.6. A small symbolic picture containing five island objects.

Fig. 5.7. The tile graph constructed from Fig. 5.6.

The spatial knowledge structure can be expressed by the 2D strings as follows:

$(a < b < c < d < e, d < c e < a < b)$

Intuitively, the first string says that "a is to the left of b, which is to the left of c, which is to the left of d, which is to the left of e", and the second string says that "d is below c and e, which are below a, which is below b". Therefore, the 2D strings express the approximate spatial relations among image objects.

The spatial relations can be expressed even more precisely, if the image objects are segmented using horizontal and vertical cutting lines, as illustrated in Figure 5.8.

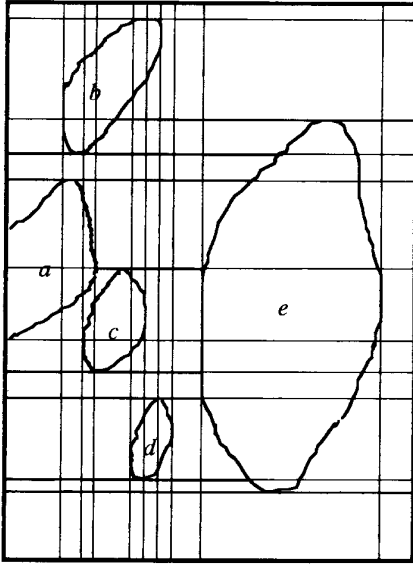


Figure 1.8 Segmentation of the symbolic picture of Figure 1.6 using cutting lines.

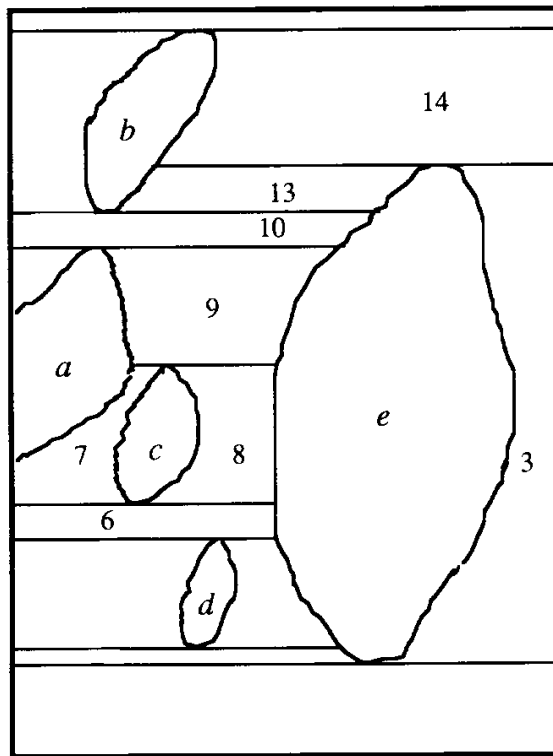
Fig. 5.8. Segmentation of the symbolic picture of Figure 5.6 using cutting lines.

After segmentation, the generalized 2D strings are as follows:

(a | a b | c a b | c b | d c b | d b | d < e, e | d e | e | a c e | a e | e | b e | b)

The generalized 2D strings give a concise description of the picture. In fact, the tile graph and the generalized 2D strings are equivalent image knowledge structures, and there are transformations to convert from one representation into another and vice versa.

Suppose the user wants to find a navigation path from the region of tile 7, to the region of tile 3. At the Spatial Reasoning and Image Information Retrieval (SRIIR) Stage, the tile graph is examined by a path finding algorithm so that different navigation plans can be generated. The resultant user-specific knowledge structure is a plan, or a number of alternate plans, as illustrated in Figure 5.9. Figure 5.9(a) shows the tiles which lie in the navigation path. Figure 5.9(b) is the navigation plan. Details of the spatial reasoning techniques will be explained later.



(a)

tile	direction/context
7	
6	south-north / left-turn
8	north / double-sp-left
9	north / double-up-sp-left
10	north / low-left-up-left-sp
13	north / X-inflection
14	north-south / right-turn
3	

(b)

Figure 1.9 The resultant navigation plan from Figure 1.7.

Fig. 5.9. The resultant navigation plan from Figure 5.7.

We can summarize the knowledge-based image processing activities for this path-finding example as follows:

ACTIVITIES	RESULTS
Image analysis and Pattern Recognition	Identify image objects (islands) and produce run-length encoded image data structure
Image structuring and Image Understanding	Produce tile graph as the image knowledge structure
Spatial Reasoning and Image Information Retrieval	Produce navigation plans as the final output

4. The Forest Fire Example of Image Information Retrieval and Spatial Reasoning

The above example illustrate the transformations from raw images to image data structures, image knowledge structures, and finally navigation plans (user-specific knowledge structures). For more complicated planing problem, such as the Forest Fire Crisis Management system, the transformations are more complex, as will be explained below.

Figure 5.10 illustrates a map which can be displayed on the screen of the Forest Fire Crisis Management system. The raw images which are the maps are either manually digitized or automatically digitized. The map objects such as towns, cities, forests, rivers, roads, etc. are stored in the image database. The image data structure could again be the run-length code.

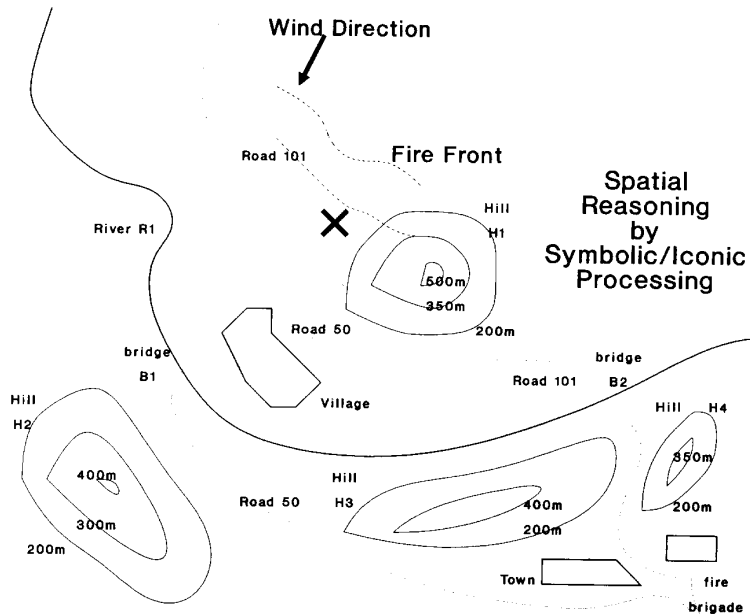


Figure 1.10 A map showing the forest fire.

Fig. 5.10. A map showing the forest fire.

This image database for maps, with its run-length encoded data structure, can be used to answer many map related queries, such as "find the roads within the city boundary of city A". However, to support the planning and problem solving activities for a system such as the Forest Fire Crisis Management system, it will be too time-consuming to process the image database to answer some queries. Therefore, a symbolic spatial data structure called the *S*-tree, which is a hierarchical structure with embedded 2D strings, can be created by the ISIU Stage. From this spatial knowledge structure, an efficient 3D model can be generated and displayed. An example of such an approximate 3D model is illustrated in Figure 5.11. (When there are multiple approximate 3D models, a pyramid is formed and each slice of this pyramid corresponds to an image with a different level of details.)

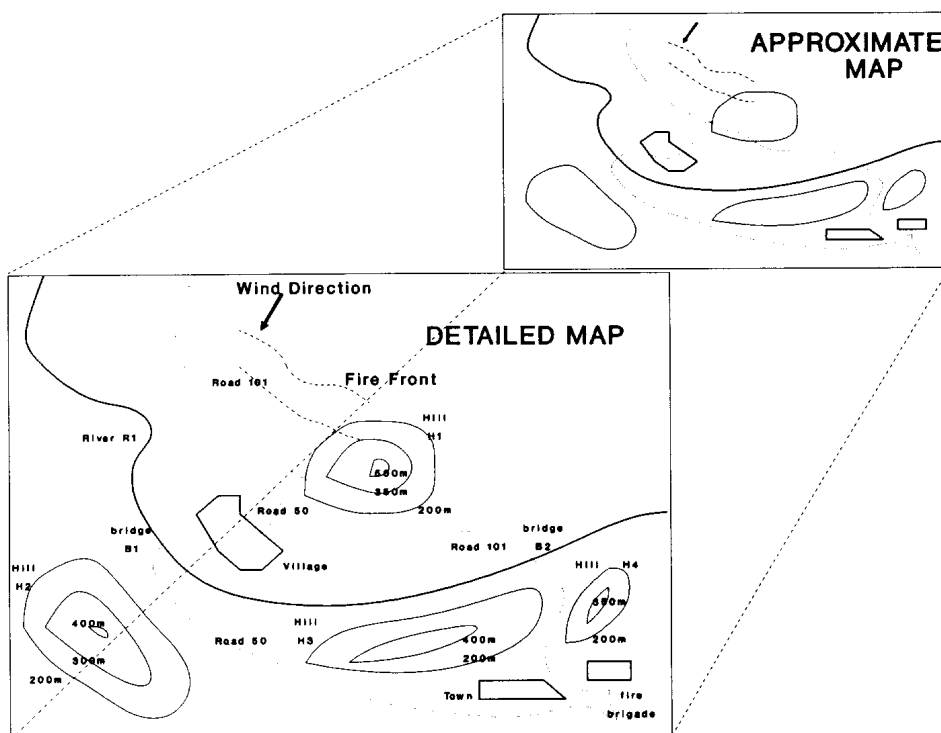


Figure 1.11 The symbolic picture constructed from the spatial data structure, the σ -tree.

Fig. 5.11. The symbolic picture constructed from the spatial data structure.

Therefore, the user of the system can observe both the original map, and the approximate 3D model, on separate screens or separate windows of the same screen. We now have two image structures: a run-length encoded image data structure for the maps, which will be regarded as the low-level image structure; and a spatial knowledge structure called the *S*-tree, which will be regarded as the high-level image structure. In what follows, we illustrate a scenario where the user's problem solving activities are supported by the image information system, which performs both high-level and low-level image processing activities.

Suppose the user poses the following query: "From the current fire front, compute fire fronts 1 hour, 2 hours, 3 hours, etc. from the present time". This query can be answered first by high-level processing of the symbolic spatial data structure as follows. From the

symbolic structure, compute the fire fronts in areas other than the hill H1. Since all objects and their locations are stored in the symbolic structure, this computation can be done quickly.

However, insufficient information is carried by the symbolic structure as far as H1 is concerned. Therefore, low-level processing of the image data structure is necessary, to find out the areas in H1 covered by forest. After such processing, we can also compute fire front in H1 area.

It is worth noting that the high-level and the low-level processing tasks can be carried out simultaneously by parallel processes, so that the user will first receive a quick but incomplete (and less accurate) reply, while more detailed information will be provided later.

In an active image information system, the detection of the fire front may lead to automatic recalculation of the predicted fire front positions. The detected object is a *hot spot* in an active index system that enables an image to automatically react to certain events.

The second query is as follows. The user would place a cross on the symbolic picture, indicating the approximate location of the fire corridor to be constructed by the firemen, and ask "what are the estimated times needed to deploy firemen."

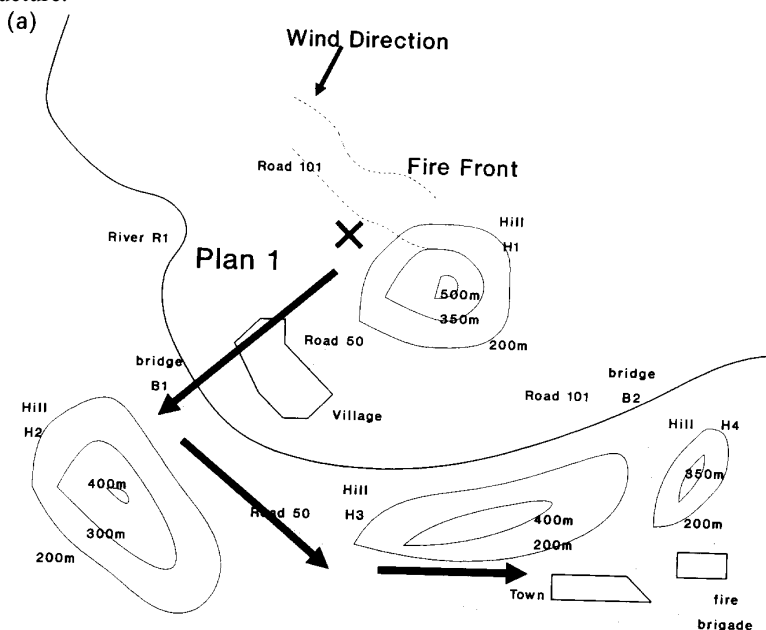
Again, both high-level and low-level processing are necessary. At the high-level processing, we can estimate the approximate times needed to travel from town T to bridge B1 and bridge B2. These are the two alternate paths that the firemen may take. Again, since object locations are known, the travel times can be estimated from the symbolic structure.

At the low-level processing, the approximate times of travel between T and B1 or B2 are replaced by more accurate results, by taking into consideration road types, road conditions, terrain conditions, etc.

Similarly, we can first do high-level processing to estimate the travel times from B1 or B2 to the cross X, and at the same time do low-level processing to recompute the travel times. The refined results are supplied to the user when available. Such information can also be displayed on the symbolic picture which the user can manipulate at will.

The third and final query from the user, is to "determine how to construct a feasible fire corridor."

As shown in Figure 5.12, two plans are available. The first plan is displayed on the symbolic picture. The user may decide bridge B1 could be blocked by evacuating traffic. Certain additional costs (time delays) are added by the user to the symbolic picture, i.e., the symbolic knowledge structure, and the deployment time is recalculated by high-level processing of the symbolic picture. Low-level processing is then initiated, to compute the cost of constructing the fire corridor. The user may move the fire corridor in the symbolic picture, causing recomputation of the cost by the low-level processing. Additional queries may be posed by the user to determine the best location for the fire corridor. For example, he may ask "what is the nearest road to the north of the town", because it will be must less costly, if the fire corridor can be constructed alongside the road. Such queries can be answered by high-level processing of the symbolic structure.



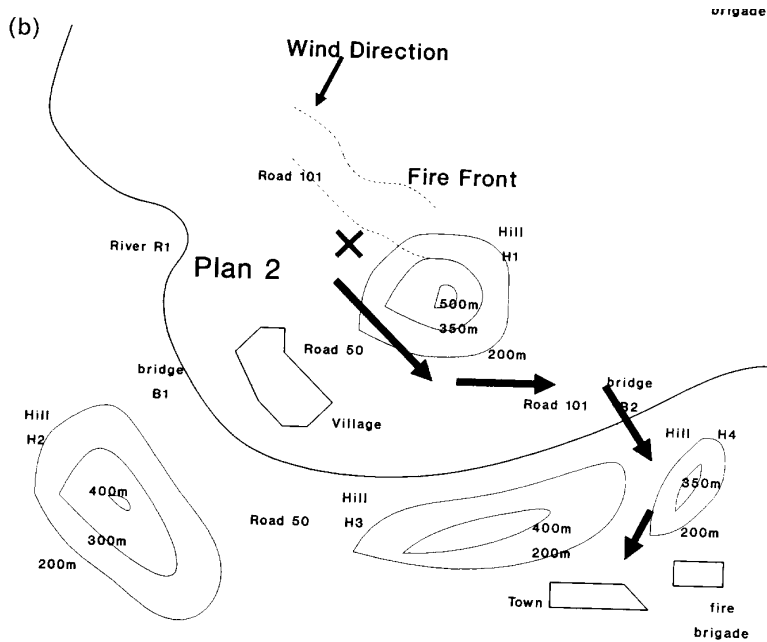


Figure 1.12 Multiple plans on the symbolic picture.

Fig. 5.12. Multiple plans on the symbolic picture.

When the second plan is displayed, the user can again decide whether the terrain in hill H1 will cause a problem or not. Since a crude terrain model is part of the 3D model of the symbolic structure, high-level processing will give the user initial estimates of the terrain passability. Low-level processing will lead to fine tuning of the travel times through H1. Finally, cost in constructing fire corridor can again be estimated.

This second example illustrates the desirability of having two levels of knowledge structures, the high-level image knowledge structure for quick computation and spatial reasoning, and the low-level image data structure for more elaborate accurate computation. However, we must be able to efficiently switch between the two levels of processing, by quickly relating image knowledge structure and image data structure.

As illustrated by these two examples, image information retrieval and spatial reasoning are complex tasks. They include the sub-tasks of image retrieval (to find images satisfying certain query conditions), image measurement, reasoning and interactive decision making. The conceptual framework presented in Section 5.2 can be used to analyze the image information retrieval and spatial reasoning tasks in a logical manner.