COGNITIVE SCIENCE 16, 355–394 (1992)

Computational Imagery

JANICE GLASGOW AND DIMITRI PAPADIAS
Queen's University

After many years of neglect, the topic of mental imagery has recently emerged as an active area of research and debate in the cognitive science community. This article proposes a concept of computational imagery, which has potential applications to problems whose solutions by humans involve the use of mental imagery. Computational imagery can be defined as the ability to represent, retrieve, and reason about spatial and visual information not explicitly stored in long-term memory.

The article proposes a knowledge representation scheme for computational imagery that incorporates three representations: a long-term memory, descriptive representation and two working-memory representations, corresponding to the distinct visual and spatial components of mental imagery. The three representations, and a set of primitive functions, are specified using a formal theory of arrays and implemented in the array-based language Nial. Although results of studies in mental imagery provide initial motivation for the representations and functionality of the scheme, our ultimate concerns are expressive power, inferential adequacy, and efficiency.

Numerous psychological studies have been carried out and several, often conflicting, models of mental imagery have been proposed. This article does not present another computational model for mental imagery, but instead treats imagery as a problem-solving paradigm in artificial intelligence (AI). We propose a concept of computational imagery, which has potential applications to problems whose solutions by humans involve the use of mental imagery. As a basis for computational imagery, we define a knowledge representation scheme that brings to the foreground the most important visual and spatial properties of an image. Although psychological theories are used as a guide to these properties, we do not adhere to a strict cognitive model: Whenever possible, we attempt to overcome the limitations of the

The research described in this article was supported by grants from the Intelligent Robotics and Information Systems Federal Network Center of Excellence, the Information Technology Research Center of Ontario, and the Natural Science and Engineering Research Council of Canada. We would like to acknowledge crystallographers Suzanne Fortier and Frank Allen for their collaboration and contributions in the area of molecular scene analysis.

Correspondence and requests for reprints should be sent to Janice Glasgow, Department of Computing and Information Science, Queens University, Kingston, Ontario K7L 3N6, Ontario.

human information-processing system. Thus, our primary concerns are efficiency, expressive power, and inferential adequacy.

Computational imagery involves tools and techniques for visual-spatial reasoning, where images are generated or recalled from long-term memory and then manipulated, transformed, scanned, associated with similar forms (constructing spatial analogies), pattern matched, increased or reduced in size, distorted, and so on. In particular, we are concerned with the reconstruction of image representations to facilitate the retrieval of visual and spatial information that was not explicitly stored in long-term memory. The images generated to retrieve this information may correspond to representations of real physical scenes or to abstract concepts that are manipulated in ways similar to visual forms.

The knowledge representation scheme for computational imagery separates visual from spatial reasoning and defines independent representations for the two modes. Whereas visual thinking is concerned with *what* an image looks like, spatial reasoning depends more on *where* an object is located relative to other objects in a scene (complex image). Each of these representations is constructed, as needed, from a descriptive representation stored in long-term memory. Thus, our scheme includes three representations, each appropriate for a different kind of processing:

- An image is stored in long-term memory as a hierarchically organized, descriptive, *deep representation* that contains all the relevant information about the image.
- The *spatial representation* of an image denotes the image components symbolically and preserves relevant spatial properties.
- The *visual representation* depicts the space occupied by an image as an occupancy array. It can be used to retrieve information such as shape, relative distance, and relative size.

While the deep representation is used as a permanent store for information, the spatial and visual representations act as working (short-term) memory stores for images.

A formal theory of arrays provides a meta-language for specifying the representations for computational imagery. Array theory is the mathematics of nested, rectangularly arranged data objects (More, 1981). Several primitive functions, which are used to retrieve, construct, and transform representations of images, have been specified in the theory and mapped into the functional programming language, Nial (Jenkins, Glasgow, & McCrosky, 1986).

The knowledge representation scheme for computational imagery provides a basis for implementing programs that involve reconstructing and reasoning with image representations. One such system, currently under investigation, is a knowledge-based system for molecular scene analysis.

Some of the concepts presented in this article will be illustrated with examples from that application area.

Research in computational imagery has three primary goals: a cognitive science goal, an AI goal, and an applications goal. The *cognitive science goal* addresses the need for computational models for theories of cognition. We describe a precise, explicit language for specifying, implementing, and testing alternative, and possibly conflicting, theories of cognition. The *AI goal* involves the development of a knowledge representation scheme for visual and spatial reasoning with images. Finally, the *applications goal* involves incorporating the knowledge representation scheme for computational imagery into the development of programs for solving real-world problems.

The article begins with an overview of previous research in mental imagery, which serves as a motivation for the representations and processes for computational imagery. This is followed by a detailed description of the deep, visual, and spatial representations for imagery, and the primitive functions that can be applied to them. It concludes with a summary of the major contributions of computational imagery to the fields of cognitive science, AI, and knowledge-based systems development, and a discussion of the relationship between our scheme and previous research in the area.

MENTAL IMAGERY

In vision research, an image is typically described as a projection of a visual scene of the back of the retina. However, in theories of mental imagery, the term "image" refers to an internal representation used by the human information-processing system to retrieve information from memory.

Although no one seems to deny the existence of the phenomenon called "imagery," there has been a continuing debate about the structure and the function of imagery in human cognition. The imagery debate is concerned with whether images are represented as *descriptions* or *depictions*. It has been suggested that descriptive representations contain symbolic, interpreted information, whereas depictive representations contain geometric, uninterpreted information (Finke, Pinker, & Farah, 1989). Others debate whether or not images play any causal role in the brain's information processing (Block, 1981). According to Farah (1988a), in depictive theories the recall of visual objects consists of the top-down activation of perceptual representation, but in descriptive theories visual recall is carried out using representations that are distinct from those in vision, even when it is accompanied by the phenomenology of "seeing with the mind's eye." Further discussions on the imagery debate can be found in various sources (e.g., Anderson, 1978; Block, 1981; Kosslyn & Pomerantz, 1977).

This article does not attempt to debate the issues involved in mental imagery, but to describe effective computational techniques for storing and manipulating image representations. To accomplish this, however, requires an understanding of the broad properties of representations and processes involved in mental imagery.

Research Findings in Mental Imagery

Many psychological and physiological studies have been carried out in an attempt to demystify the nature of mental imagery. Of particular interest to our research are studies that support the existence of multiple image representations and describe the functionality of mental imagery processes. In this section we overview relevant results from such studies, and based on these results, propose some important properties of mental imagery, which we use to motivate our representation scheme for computational imagery.

Several experiments provide support for the existence of a visual memory, distinct from verbal memory, in which recognition of verbal material is inferior. Paivio's (1975) dual-code theory suggests that there is a distinction between verbal and imagery processing. This theory leaves the exact nature of mental images unspecified, but postulates two interconnected memory systems—verbal and imaginal—operating in parallel. The two systems can be independently accessed by relevant stimuli but they are interconnected in the sense that nonverbal information can be transformed into verbal and vice versa. Furthermore, it has been indicated that visual memory may be superior in recall (Standing, 1973).

The issue of visual memory is an important one for computational imagery. What it implies to us is the need for separate descriptive and depictive representations. This is reinforced by the experiments carried out by Kosslyn (1980) and his colleagues, who concluded that images preserve the spatial relationships, relative sizes, and relative distances of real physical objects. Pinker (1988) suggested that image scanning can be performed in two- and three-dimensional space, providing support for Kosslyn's proposal that mental images capture the spatial characteristics of an actual display. Pinker also indicated that images can be accessed using either an object-centered or a world-centered coordinate system.

A series of experiments suggest that mental images are not only visual and spatial in nature, but also structurally organized in patterns, that is, they have a hierarchical organization in which subimages can occur as elements in more complex images (Reed, 1974). Some researchers propose that under certain conditions images can be reinterpreted: They can be reconstructed in ways that were not initially anticipated (Finke et al., 1989). Experiments also support the claim that creative synthesis is performed by composing mental images to make creative discoveries (Finke & Slayton, 1988 (Finke & Slayton, 1988).

The relationship between imagery and perception was considered by Brooks (1968), who demonstrated that spatial visualization can interfere with perception. Farah (1988a) also suggested that mental images are visual representations in the sense that they share similar representations to those used in vision, but noticed that this conclusion does not imply that image representations are depictive because both imagery and perception might be descriptive. Farah argued, from different evidence, however, that they are in fact spatial.

Findings, provided by the study of patients with visual impairments, point toward distinct visual and spatial components of mental imagery. Mishkin, Ungerleider, and Macko (1983) showed that there are two distinct cortical visual systems. Their research indicated that the temporal cortex is involved in recognizing *what* objects are, whereas the parietal cortex determines *where* they are located. Further studies have verified that there exists a class of patients who often have trouble localizing an object in the visual field, although their ability to recognize the object is unimpaired (De Renzi, 1982). Other patients show the opposite patterns of visual abilities: They cannot recognize visually presented objects, although they can localize them in space (Bauer & Rubens, 1985). Such patients are able to recognize objects by touch or by characteristic sounds. It has also been suggested that the preserved and impaired aspects of vision in these patients are similarly preserved or impaired in imagery (D. Levine, Warach, & Farah, 1985). In experimental studies, subjects with object identification problems were unable to draw or describe familiar objects despite being able to draw and describe in detail the locations of cities on a map, furniture in a house, and landmarks in a city. Patients with localization problems were unable to describe relative locations, such as cities on a map, although they could describe from memory the appearance of a variety of objects. Such findings have been interpreted by some researchers (e.g., Kosslyn, 1987) as suggesting two distinct components of mental imagery, the spatial and the visual, where the spatial component preserves information about the relative positions of the meaningful parts of a scene and the visual component preserves information about how (e.g., shape, size) a meaningful part of a scene looks.

Although there are varying strategies for retrieving spatial information and solving problems concerning spatial relations, research has suggested that humans typically use mental imagery for spatial reasoning (Farah, 1988b). Experimental results also support an isomorphism between physical and imaged transformations (Shepard & Cooper, 1982). A premise of Kritchevsky (1988) is that behavior can be divided into spatial and nonspatial components. For example, determining the color of an object is a nonspatial behavior, whereas determining relative positions of objects is a spatial behavior. Kritchevsky assumed that the spatial component of behavior is understood in terms of elementary spatial functions. Furthermore, these functions are independent of any particular sensory modality (Ratcliff, 1982).

Although individually the results described previously do not imply a particular approach to computational imagery, collectively they infer several properties that we wish to capture in our approach. Most importantly, an image may be depicted and reasoned with visually or spatially, where a visual representation encodes what the image looks like and the spatial representation encodes relative location of objects within an image. As well, images are inherently three-dimensional and hierarchically organized. This implies that computational routines must be developed that can decompose, reconstruct, and reinterpret image representations. Results from studies comparing imagery and vision imply that the representations and processes of imagery may be related to those of high-level vision. Thus, we should also consider the representations and functionality of object recognition when defining computational imagery. Finally, we must be able to consider an image from either an object-centered or a viewer-centered perspective.

The numerous experiments that have been carried out in mental imagery not only suggest properties for the representation scheme, but also support the premise that mental imagery is used extensively to reason about real-world problems. Thus, computational imagery is an important topic to investigate in relation to AI problem solving.

The subjective nature of mental imagery has made it a difficult topic to study experimentally. Qualities like clarity, blurring, and vividness of images are not directly observable and may differ from one person to another. Furthermore, it has been argued by some researchers that it is impossible to resolve the imagery debate experimentally because depictive and descriptive representations do not have distinct properties from which behavioral consequences can be predicted (Anderson, 1978). As a result, several alternative accounts have been proposed to explain the findings mentioned previously. The most important of these are tacit knowledge, experimenter bias, eye movements, and task-induced characteristics (Intons-Peterson, 1983). These difficulties involved in experimental studies emphasize the need for computer models for mental imagery. Although the knowledge representation scheme for computational imagery is not meant to model a particular theory of imagery, it does provide the tools for specifying, testing, and formally analyzing a variety of theories, and thus can contribute to resolving the imagery debate.

Theories and Principles of Mental Imagery

Pylyshyn (1981), a forceful proponent of the descriptive view, argued that mental imagery simply consists of the use of general thought processes to simulate perceptual events, based on tacit knowledge of how these events happened. Pylyshyn disputed the idea that mental images are stored in a raw uninterpreted form resembling mental photographs, and argued for an abstract format of representation called propositional code. Kosslyn's (1980) model of mental imagery is based on a depictive theory, which claims that images are quasi-pictorial, that is, they resemble pictures in several ways but lack some of their properties. According to Kosslyn's model, mental images are working memory, visual representations generated from long-term memory, deep representations. A set of procedures, which is referred to as the "mind's eye," serves as an interface between the visual representations and the underlying data structures, which may be decidedly nonpictorial in form. Hinton (1979) disputed the picture metaphor for imagery and claimed that images are more like generated constructions. In this approach, as in Marr and Nishihara's (1978) 3D model, complex images can be represented as a hierarchy of parts.

Finke (1989) took a different approach to the imagery debate. Instead of proposing a model, Finke defined five "unifying principles" of mental imagery:

- The principle of *implicit encoding* states that imagery is particularly useful for retrieving information about physical properties of objects and relations among objects whenever this information was not previously, explicitly encoded.
- The principle of *perceptual equivalence* states that similar mechanisms in the visual system are activated when objects or events are imagined, as when the same objects or events are actually perceived.
- The principle of *spatial equivalence* states that the spatial relations between objects are preserved, although sometimes distorted, in mental images.
- The principle of *structural equivalence* states that the structure of images corresponds to that of perceived objects, in the sense that the structure is coherent, well organized, and can be reinterpreted.
- The principle of *transformational equivalence* states that imagined and physical transformations exhibit similar dynamic characteristics and follow the same laws of motion.

These principles provide a basis for evaluating the representations and functions for computational imagery; in the development of our scheme we have attempted to address each of the underlying principles for mental imagery.

Stages of Image Representations

The hypothesis of multiple representations for mental imagery can explain several experimental results that cannot be explained independently by either a propositional, a spatial, or a visual representation. For instance, after a series of experiments, Atwood (1971) concluded that memory for high-image phrases is disrupted if followed by a task requiring the subject to process a visually presented digit in contrast to abstract phrases. Although other researchers found difficluty in replicating Atwood's experiments, Jannsen (1976) succeeded consistently over several experiments and claimed that other failures stemmed from using an interfering task that is spatial rather than visual. Baddeley and Lieberman (1980) interpreted these results as pointing towards distinct visual and spatial components of mental imagery.

When images are retrieved, it is possible to recall information about which objects constitute a scene and their spatial relationships with other objects without remembering what the object looks like. Furthermore, we are able to recognize objects independent of any context. Distinct spatial and visual components for imagery can explain such phenomena, where the spatial component can be considered as an index that connects visual images to create a scene.

Intuitively, we can distinguish between visual and spatial representations by considering the type of information we wish to retrieve. Consider, for example, answering the following questions: *How many windows are there in your home? What city is farther north, Seattle or Montreal? What objects are sitting on top of your desk? Who was sitting beside Mary in class?* These questions can typically be answered without constructing an explicit visual image, that is, you could possibly recall that John was sitting beside Mary without knowing what John looked like or what clothes he was wearing. Each of these questions does rely on knowing the relative locations of objects within a recalled image, information that is embodied in a spatial representation. Now consider questions such as: *What is the shape of your dog's ears? What does a particular image look like if you rotate it ninety degrees? What is larger, a rabbit or a racoon? Is Montreal or Toronto closer to Ottawa?* To answer these questions you may need to reconstruct a representation that preserves information such as size, shape, or relative distance, information that is embodied in a visual representation.

From the computational point of view, a single representational system cannot always effectively express all the knowledge about a given domain; different representational formalisms are useful for different computational tasks (Sloman, 1985). In perceptual systems, for instance, multiple representations have been proposed to derive cognitively useful representations from a visual scene. For computational imagery, we propose three stages of image representation, each appropriate for a different type of information processing (Papadias & Glasgow, 1991). The deep representation stores structured, descriptive information in terms of a semantic network, long-term memory model. The working-memory representations (spatial and visual) are consciously experienced and generated as symbolic and occupancy arrays, as needed, using information stored in the deep representation. Details about the computational advantages of each of the image representations involved in the scheme will be presented in the following section.

KNOWLEDGE REPRESENTATION SCHEME

Research in AI has long been concerned with the problem of knowledge representation. AI programs rely on the ability to store descriptions of a particular domain and formally manipulate these descriptions to derive new knowledge. Traditional approaches to knowledge representation include logic representations, which denote the objects and relations in the world in terms of axioms, and structural knowledge representation schemes, which denote concepts and relations in terms of structural hierarchies.

In addition to general schemes, there exist specialized schemes concerned with the representation of the visual representation of images. In discrimination trees, objects are sorted by discriminating on their coordinates, as well as other quantitative and qualitative discriminators (McDermott & Davis, 1984). A simple way of describing volume or shape is with occupancy arrays, where cells of the array denote objects filling space. For computer vision applications, an occupancy array is often called a gray-level description, because the value of the cells encode the intensity of light on a gray scale from white to black. For our molecular scene analysis application, we use three-dimensional occupany arrays that correspond to electron density maps resulting from X-ray diffraction experiments. The values of the cells in such maps correspond to the electron density in a unit cell of a crystal.

According to Biederman (1987), the visual representation for objects can be constructed as a spatial organization of simple primitive volumes, called geons. Other researchers have proposed alternative primitive volumes, like generalized cones, spheres, and so forth. A major contribution in representational formalisms for images is the progression of primal sketch, 2½D sketch, and 3D sketch (Marr & Nishihara, 1978). The primal sketch represents intensity changes in a 2D image. The 2½D sketch represents orientation and depth of surface from a particular viewer perspective. Finally, the 3D sketch represents object-centered spatial organization.

The representation schemes discussed before are not suggested as structures for representing human knowledge and do not necessarily commit to addressing questions about mental processes. Whereas many AI researchers believe that the best way to make true thinking machines is by getting computers to imitate the way the human brain works (Israel, 1987), research in knowledge representation often is more concerned with expressiveness and efficiency, rather than explanatory and predictive power. Thus, although our knowledge representation scheme attempts to preserve the most relevant properties of imagery, whenever possible we try to overcome the limitations of the human information-processing system. For example, theories of divided attention argue that attention can be concentrated on, at most, a few mental processes at a time. Our proposed scheme has the capability of relatively unrestrictive parallel processing of spatial images. Furthermore, although the resolution of mental images is limited by the capabilities of the human mind, in the knowledge representation scheme the resolution restrictions are imposed by the implementation architecture.

A theory of arrays provides a formalism for the representations and functions involved in computational imagery. Array theory (More, 1981) is the mathematics of nested, rectangularly arranged collections of data objects. Similar to set theory, array theory is concerned with the concepts of nesting, aggregation, and membership. Array theory is also concerned with the concept of data objects having a spatial position relative to other objects in a collection. Thus, it provides for a multidimensional, hierarchical representation of images, in which spatial relations are made explicit.

We consider computational imagery as the ability to represent, retrieve, and reason about information not explicitly stored in long-term memory. In particular, we are concerned with visual and spatial information. Recall that the visual component of imagery specifies how an image looks and is used to retrieve information such as shape, size, and volume, whereas the spatial component of imagery denotes where components of an image are situated relative to one another and is used to retrieve information such as neighborhoods, adjacencies, symmetry, and relative locations. As illustrated in Figure 1, the long-term memory representation is implemented as a description of the image, and the working-memory representations correspond to representations that make explicit the visual and spatial properties of an image. In the remainder of this section, we describe each of the representations in detail and discuss the primitive functions that operate on them. First, though, we overview the theory or arrays that provide the basis for describing and implementing the representations and functions for computational imagery.

Array Theory

Results of empirical studies suggest that images may be organized using both a hierarchical and a spatial structure. Components of an image may be grouped into features and stored based on their topological relations, such as adjacency or containment, or their spatial relations, such as *above, beside, north-of,* and so on. Because of the relevance of storing and reasoning about such properties of an image, we base the development of the knowledge representation scheme for computational imagery on a theory of arrays. This mathematical theory allows for a multidimensional, hierarchical representation of images in which spatial relations are made explicit. Furthermore, functions can be defined in array theory for constructing, manipulating, and retrieving information from images represented as arrays. For example, functions that compose, translate, juxtapose, and compare images have been defined within the theory.

The development of array theory was motivated by efforts to extend the data structures of APL and has been influenced by the search for total operations that satisfy universal equations (More, 1981). In this theory, an array is a collection of zero or more items held at positions in a rectangular arrangement along multiple axes. Rectangular arrangement is the concept of data objects having a position relative to other objects in the collection.

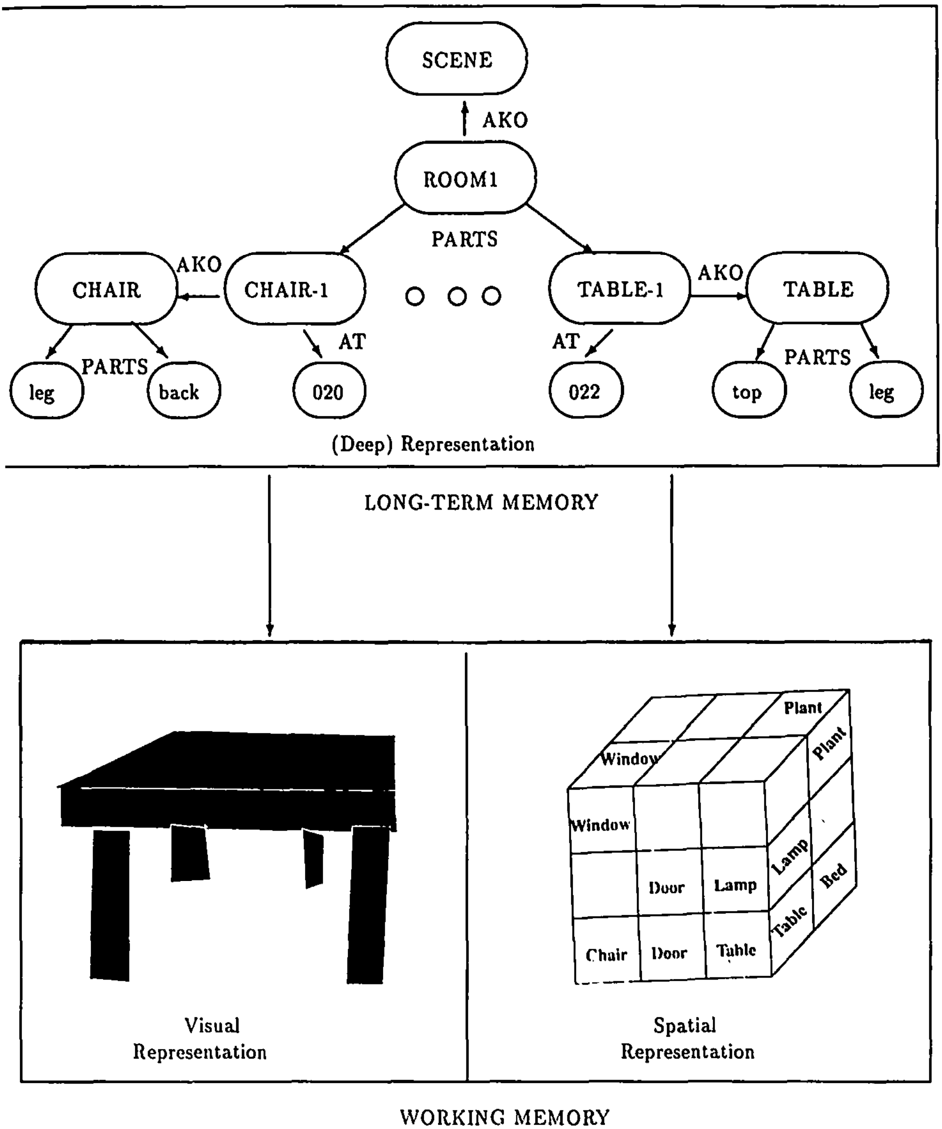


Figure 1. Representations for computational imagery

The interpretation of this structure can be illustrated using nested, box diagrams. Consider the array diagram in Figure 2. In this array the pair formed from 7 and 9 is an array nested within the larger array. Nesting is the concept of having the objects of a collection be collections themselves. This is an important concept in array theory because it is the power of aggregating arbitrary elements in an array that gives the theory much of its expensive

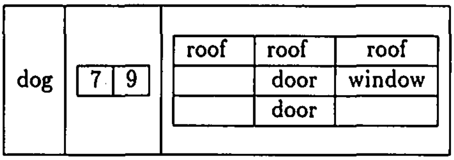


Figure 2. Example of nested array diagram

power. The third element of the array is a symbolic array, which denotes an image of a house containing three parts. The indexing of the array allows us to make explicit such properties as *above(roof,door)* and *left-of(door, window)* in a notation that is both compact and accessible.

Array theory has provided a formal basis for the development of the *N*ested *I*nteractive *A*rray *L*anguage, *Nial*. This multiparadigm programming language combines concepts from APL, Lisp, and FP with conventional control mechanisms (Jenkins et al., 1986). The primitive functions of array theory have all been implemented in Q'Nial (Jenkins & Jenkins, 1985), a commercially available, portable interpreter of Nial developed at Queen's University.

Operations in array theory are functions that map arrays to arrays. A large collection of total, primitive operations are described for the theory. They are chosen to express fundamental properties of arrays. Nial extends array theory by providing several syntactic forms that describe operations, including composition, partial evaluation of a left argument, and a lambda form. Array theory also contains second-order functions called transformers that map operations to operations.

It has previously been shown that the syntactic constructs of array theory facilitate both sequential and parallel computations (Glasgow, Jenkins, McCrosky, & Meijer, 1989). This is an important feature when considering computational imagery as a basis for specifying cognitive processes, which themselves may be sequential or parallel. The potential parallelism in array theory comes from three sources: inherent parallelism in the primitive operations, parallelism expressed by syntactic constructs, and parallelism in operation application controlled by primitive transformers. The potential parallelism of the primitive operations results from treating an entire array as a single value; each array takes an array as a single argument and returns an array as its result. Array theory includes transformers that allow expression of the parallel application of an operation to subparts of an array.

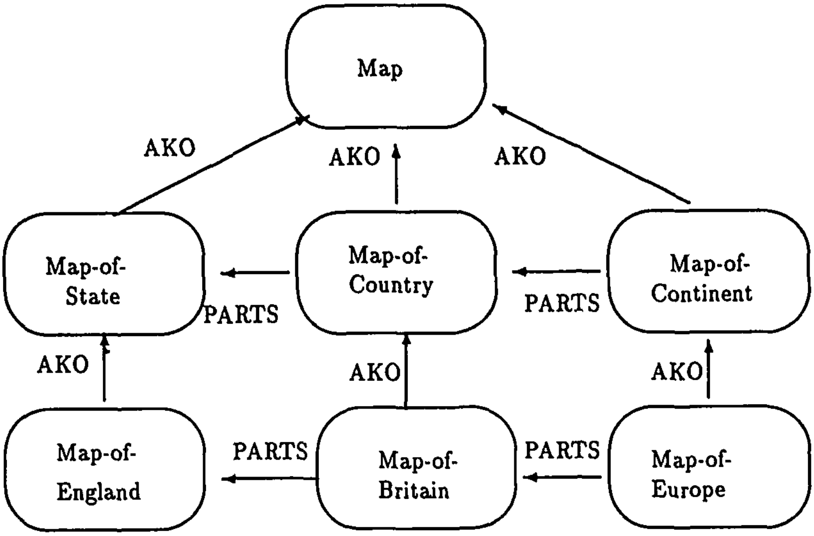
The software development associated with AI problem solving in general, and with computational imagery in particular, differs from traditional computer applications. AI problems are solved at the conceptual level, rather than a detailed implementation level. Thus, much of the programming effort is spent on understanding how to represent and manipulate the knowledge associated with a particular problem, or class of problems. This imposes certain features on a programming language, including interactive program development, operations for symbolic computation, dynamically created data structures, and easy encoding of search algorithms. Although Lisp and Prolog both address capabilities such as these, they provide very different and complementary approaches to problem solving. The language Nial is an attempt to find an approach to programming that combines the logic and functional paradigms of Prolog and Lisp (Glasgow & Browse, 1985, Jenkins et al., 1986). It has been demonstrated that array theory and Nial can provide a foundation for logic programming (Glasgow, Jenkins, Blevis, & Feret, in press), as well as other descriptive knowledge representation techniques (Jenkins et al., 1988). These techniques have been implemented and tested on a variety of knowledge-based applications.

Deep (Long-Term Memory) Representation

The deep representation for computational imagery is used for the long-term storage of images. Earlier work has suggested that there exists a separate long-term memory model that encodes visual information descriptively (Kosslyn, 1980; Pinker, 1984). This encoding can then be used to generate depictive representations in working memory. As pointed out in Marschark, Richman, Yuille, and Hunt (1987), most of the research in vision and imagery has focused on the format of the on-line conscious representations, excluding long-term storage considerations. Our point of view is that the deep representation falls more in the limits of research in long-term memory than imagery, and we base its implementation on the hierarchical network model of semantic memory (Collins & Quillian, 1969). This model is suitable for storing images because they have a structured organization in which subimages can occur as elements in more complex images.

The deep representation in our scheme is implemented using a frame language (Minsky, 1975), in which each frame contains salient information about an image or class of images. This information includes propositional and procedural knowledge. There are two kinds of image hierarchies in the scheme: the AKO (a kind of) and the PARTS. The AKO hierarchy provides property inheritance: Images can inherit properties from more generic image frames. The PARTS hierarchy is used to denote the structural decomposition of complex images. The deep representation for imagery can be characterized as nonmonotonic because default information (stored in specific slots, or inherited from more generic frames) is superseded as new information is added to a frame.

A frame corresponding to the image of a map of Europe and part of the semantic network for a map domain is illustrated in Figure 3. Each node in the network corresponds to an individual frame and the links describe the relationships among frames. The AKO slot in the frame of the map of



a) Semantic network representation

FRAME	Map-of-Europe
AKO	Map-of-Continent
PARTS	Sweden (0 4) Britain (1 0) ...
POPULATION	'find-population'
...	...

b) Frame representation

Figure 3. Example of deep representation

Europe denotes that the frame is an instance of the concept "Map-of-Continent." The PARTS slot contains the meaningful parts that compose the map, along with an index value that specifies their relative locations. The POPULATION slot contains a call to a procedure that calculates the population of Europe, given the populations of the countries. As well, the frame could incorporate several other slots, including ones used for the generation of the spatial and visual representations.

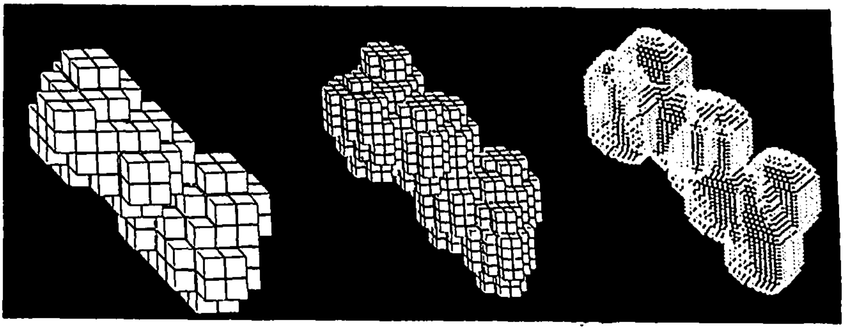
For the molecular scene analysis application, the frame hierarchy is more complex than the simple map example. The structure of a protein is described in terms of a crystal, which consists of a regular three-dimensional arrangement of identical building blocks. The structural motif for a protein crystal can be described in terms of aggregate (complex or quaternary), three-dimensional structures. Similarly, tertiary structures can be decomposed into secondary structures, and so on. Each level in this decomposition hierarchy corresponds to a conceptual frame denoting a molecular fragment at a meaningful level of abstraction. If we consider a fully determined crystal as a molecular scene, there exist databases containing over 90,000 images of small molecules and over 600 images of protein structures (Allen, Bergerhoff, & Sievers, 1987). These databases include the three-dimensional geometry of the molecular scenes that forms a basis for our long-term memory model for molecular images.

Semantic networks and frames have previously been suggested as representations for images in vision research. One example of this deals with the interpretation of natural scenes (M. Levine, 1978). In Levine's system, the spatial relations are represented as arcs such as *left-of, above,* or *behind.* A classic example of the use of semantic networks is the work of Winston (1975) on structural descriptions. In that study on scene understanding, common structures, such as arches and pedestals, are represented in terms of their decomposition into parts and a description of the spatial relations among the parts. Although this approach may be useful for some applications, we argue later that explicitly representing spatial relations in terms of an indexed array provides increased computational efficiency for spatial reasoning.

Our implementation of the deep representation has several attractive properties. First, it provides a natural way to represent knowledge because all the information about an image (or a class of images) can be stored in a single frame, and the structure of images is captured by the PARTS hierarchy. It is assumed that a property is stored at the most general level possible (highest level in the conceptual hierarchy) and is shared by more specific levels, thus providing a large saving in space over propositional or database formulations of property relations. The deep representation also incorporates the psychological concept of semantic networks in an implementation that provides features such as procedural attachment. The nonmonotonic feature of the frame allows for reasoning with incomplete information; default information can be stored in conceptual frames and inherited and used for depicting or reasoning about subconcepts or instances of images. Despite its attractive properties, however, the deep representation is not the most suitable representation for all of the information processing involved in imagery. Thus, we require alternative representations to facilitate the efficiency of the scheme.

Working-Memory Representations

Mental images are not constantly experienced. When an image is needed, it is generated on the basis of stored information. Thus, unlike the deep representation, which is used for the permanent storage of information, the



a) 1.0A b) 0.5A c) 0.15A

Figure 4. Example of occupancy arrays for visual representations

working-memory representations of an image exist only during the time that the image is active, that is, when visual or spatial information processing is taking place.

The distinct working-memory representations were initially motivated by results of cognitive studies that suggest distinct components in mental imagery (Kosslyn, 1987). More importantly, separate visual and spatial representations provide increased efficiency in information retrieval. The visual representation is stored in a format that allows for analysis and retrieval of such information as shape and relative distance. Because the spatial representation makes explicit the important features and structural relationships in an image while discarding irrelevant features such as shape and size, it provides a more compact and efficient depiction for accessing spatial and topological properties.

Visual Representation. The visual representation corresponds to the visual component of imagery, and it can either be reconstructed from the underlying deep representation or generated from low-level perceptual processes. Similar to Kosslyn's (1980) skeletal image, this representation is depictive and incorporates geometric information. Unlike Kosslyn's approach, we assume that the visual representation can be three-dimensional and viewer-independent.

For the current implementation of the visual representation we use *occupancy arrays.* An occupancy array consists of cells, each mapping onto a local region of space and representing information such as volume, lightness, texture, and surface orientation about this region. Objects are depicted in the arrays by patterns of filled cells isomorphic in surface area to the objects. Figure 4 illustrates depictions of three-dimensional occupany arrays corresponding to a molecular fragment at varying levels of resolution. These arrays were constructed using geometric coordinates and radii of the atomic components of the molecule.

Representing occupancy arrays explicitly in long-term memory can be a costly approach. As a result, other approaches to storing or generating this information (like generalized shapes) have been developed. Such approaches can be incorporated into an application of the scheme for computational imagery.

Spatial Representation. A primary characteristic of a good formalism for knowledge representation is that it makes relevant properties explicit. Although an occupancy array provides a representation for the visual component of imagery, it is basically uninterpreted. For the spatial component of imagery we are best served by a representation that explicitly denotes the spatial relations between meaningful parts of an image, corresponding to the mental maps created by humans. Thus, we use a multidimensional symbolic array to depict the spatial structure of an image, where the symbolic elements of the array denote its meaningful parts (Glasgow, 1990). The symbolic array preserves the spatial and topological relationships of the image features, but not necessarily relative sizes or distances. The arrays can be interpreted in different ways depending on the application. If, for example, we use the scheme to reason about geographic maps, interpretations could include predicates such as *north, east, south,* and *west;* if the array is used to represent the image of a room, then the interpretation would involve predicates such as *above, behind, left-of, beside,* and so on. For molecular scene analysis we are more concerned with properties such as *symmetry* and *adjacency* (bonding), which are made explicit by a symbolic array. The spatial representation can also denote nonspatial dimensions. For example, the symbolic array could be used to index features such as height or speed.

The symbolic array representation for the spatial component of imagery is generated, as needed, from information stored explicitly in the frame representation of an image. For example, in Figure 3 the PARTS slot contains the indices needed to reconstruct the spatial representation for a simplified map of Europe. Figure 5 illustrates this symbolic array. Note that some parts occupy more than one element in an array (e.g., Italy, France). This is necessary to capture all the spatial relationships of the parts of an image. We may also wish to denote more complex relations, such as one object being "inside" another. This is illustrated in Figure 6, which displays a spatial image of a glass containing water.

According to Pylyshyn (1973), images are not raw, uninterpreted, mental pictures, but are organized into meaningful parts that are remembered in terms of their spatial relations. Furthermore, we can access the meaningful parts, that is, we are able to focus attention on a specific feature of an

				Sweden	
Britain			Denmark		
		Holland	Germany	Germany	
		Belgium			
	France	France	Italy	Yugoslavia	Yugoslavia
Portugal	Spain		Italy		Greece

Figure 5. Example of symbolic array for spatial representation

glass	water	glass
glass	glass	glass

Figure 6. Symbolic array depiction of inside relation

image. Nested symbolic arrays capture these properties by representing images at various levels of abstraction as prescribed by the PART hierarchy of the deep representation; each level of embedding in an array corresponds to a level of structural decomposition in the frame hierarchy. For instance, focusing attention on Britain in the array of Figure 5 would result in a new array in which the symbol for Britain is replaced by its spatial representation (see Figure 7). This subimage is generated using the PARTS slot for the frame of Britain in the deep representation.

It has been suggested that people can reconstruct and reinterpret mental images (Finke, 1989). The proposed scheme also provides the capability to combine and reconstruct images, using special functions that operate on the symbolic array representations. For instance, we can combine a portion of the array of Figure 5 with a portion of the array that corresponds to the map of Africa and create a new array containing Mediterranean countries.

Recall that Pinker (1988) pointed out that images are represented and manipulated in three dimensions. Similar to the visual representation, a symbolic array can be two- or three-dimensional, depending on the application. In the domain of molecular scenes, fragments of molecules are represented as three-dimensional symbolic arrays at varying levels of abstraction, corresponding to the level of decomposition in the frame hierarchy. For example, a protein can be represented as a three-dimensional array of symbols denoting high-level structures, which can be decomposed into nested

<table>
<tr><td></td><td></td><td></td><td></td><td>Sweden</td><td></td></tr>
<tr><td><table><tr><td></td><td>Scotland</td></tr><tr><td>Wales</td><td>England</td></tr></table></td><td></td><td></td><td>Denmark</td><td></td><td></td></tr>
<tr><td></td><td></td><td>Holland</td><td>Germany</td><td>Germany</td><td></td></tr>
<tr><td></td><td></td><td>Belgium</td><td></td><td></td><td></td></tr>
<tr><td></td><td>France</td><td>France</td><td>Italy</td><td>Yugoslavia</td><td>Yugoslavia</td></tr>
<tr><td>Portugal</td><td>Spain</td><td></td><td>Italy</td><td></td><td>Greece</td></tr>
</table>

Figure 7. Embedded symbolic array representation

arrays of symbols denoting progressively more detailed substructures. Because of the size and complexity of molecular structures, it is essential to be able to reason at multiple levels of abstraction when analyzing a particular molecular scene. Figure 8 depicts a three-dimensional image of a fragment of a protein secondary structure, and an embedded amino acid residue substructure containing symbols denoting atoms. Bonding at the residue and atomic level is made explicit through structural adjacency in the representation.

For image recognition and classification, it is necessary to pick out characteristic properties and ignore irrelevant variations. One approach to image classification is on the basis of shape. Although the visual representation provides one approach to shape determination, the spatial representation allows for a hierarchical, topological representation for shape. This approach is particularly useful in applications where images are subject to a large number of transformations. For example, a human body can be configured many ways depending on the positions of the arms, legs, and so forth. Although it is impossible to store a separate representation for every possible configuration, it is possible to represent a body using a symbolic array that makes explicit the parts of the body and the relations among parts that remain constant under allowable transformations. Figure 9 illustrates such a spatial representation. Combined with a primitive shape descriptor (such as generalized cylinder), the spatial representation provides for multidimensional shape descriptors as proposed by Marr (1982).

The spatial representation can be thought of as descriptive because it can be expressed as a propositional representation, where the predicates are spatial relationships and the arguments are concrete, imaginable objects. Although information in the spatial representation can be expressed as

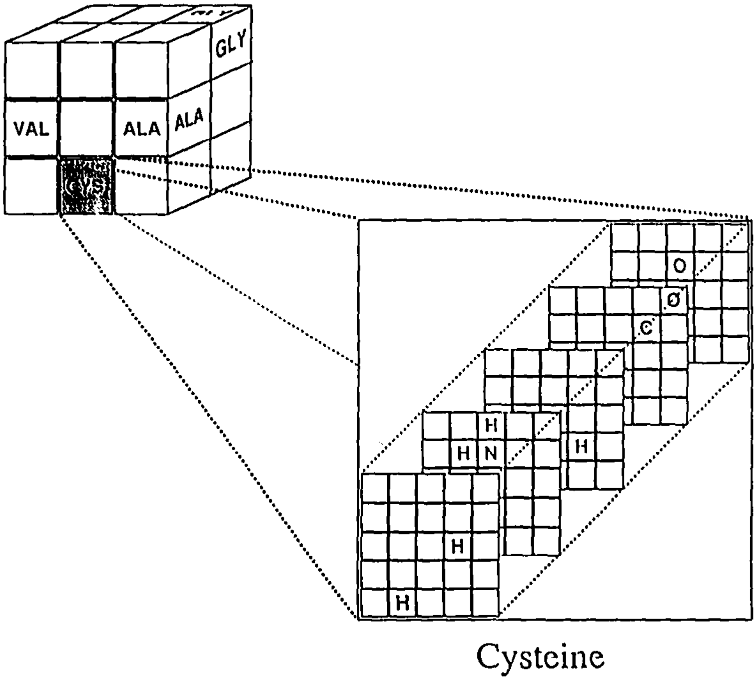


Figure 8. Symbolic array of molecular fragment

propositions, the representations are not computationally equivalent, that is, the efficiency of the inference mechanisms is not the same. The spatial structure of images has properties not possessed by deductive propositional representations. As pointed out by Lindsay (1988, p. 231), these properties help avoid the "combinatorial explosion by correct but trivial inferences that must be explicitly represented in a propositional system." Lindsay also argued that the spatial image representations (symbolic representations in our case) support nondeductive inference using built-in constraints on the processes that construct and access them. Consider, for example, the spatial representation of the map to Europe. To retrieve the information about what countries are north of Germany, we need only search a small portion of the symbolic array. Alternatively, in a propositional approach, the spatial relations would be stored as axioms such as

north-of(Britain,Portugal), north-of(France,Spain), north-of(Holland,Belgium)...,

and general rules such as

$$north\text{-}of(X,Y) \wedge north\text{-}of(Y,Z) \rightarrow north\text{-}of(X,Z).$$

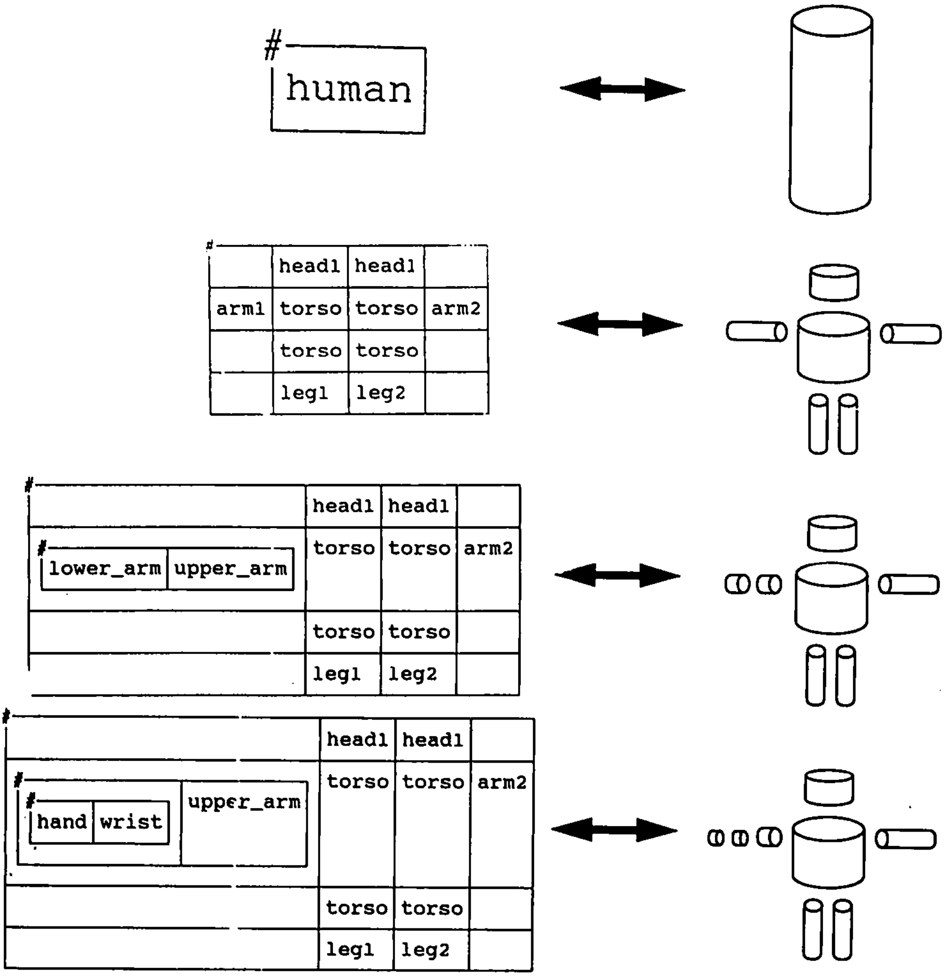


Figure 9. Spatial representation for topological shape description

To determine what countries are north of Germany using this representation involves considering all axioms plus recursive calls to the general rule. Thus, although the information embodied in the spatial representation is derivable from propositional knowledge, the indexing of this information using an array data structure can make spatial reasoning more efficient.

Another advantage of symbolic arrays, with respect to propositional representations, concerns temporal reasoning. Any cognitive system, natural or artificial, should be able to deal with a dynamic environment in which a change in a single item of knowledge might have widespread effects. The problem of updating a system's representation of the state of the world to reflect the effects of actions is known as the *frame problem* (Raphael, 1971). Representing an image as a symbolic array has advantages when considering this problem. Consider, for example, changing the position of a country in our map of Europe. In a propositional representation we would have to consider all of the effects that this would have on the current state. Using the symbolic array to store the map, we need only delete the country from its previous position and insert it in the new one. Because spatial relationships are interpreted, not logically inferred, from image representations, we eliminate some of the problems associated with nonmonotonicity in domains involving spatial and/or temporal reasoning. There still remains, however, the problem of dealing with truth maintenance if we desire to preserve relations as changes are made.

The representation scheme provides the ability to extract propositional information from symbolic arrays and to create or manipulate symbolic arrays with respect to propositional information. It should be noted, though, that the spatial representation does not provide the full expressive power of first-order logic: We cannot express quantification or disjunction. For example, it is not possible to represent an image of Europe that denotes the fact that Britain is either north of *or* south of Portugal. But mental images cannot express such information either. The representation scheme can be integrated with a logic representation through Nlog, a logic programming environment based on the theory of nested arrays (Glasgow et al., 1991). In this environment, the spatial information extracted through imagery processes can be used as propositions in logical deductions.

Primitive Functions for Computational Imagery

Approaches to knowledge representation are distinguished by the operations performed on the representations. Thus, the effectiveness of our scheme can be partially measured by how well it facilitates the implementation of imagery-related processes. In this section we review some of the primitive imagery functions that have been defined for the scheme. We also discuss how these functions provide the building blocks for more complex processes.

In his computational model for imagery, Kosslyn (1980) considered three basic categories of image processes: procedures for image generation (mapping deep representations into visual representations), procedures for evaluating a visual image, and procedures for transforming an image. Although we attempt to capture much of the functionality of the procedures described by Kosslyn, and in fact can categorize our operations similarly, the nature of our representations imply great difference in the implementations. For example, we define operations for both visual and spatial reasoning of three-dimensional images. Also, because our images can be organized hierarchically, we have defined functions that allow us to depict parts of an image at varying levels of abstraction using embedded arrays. When considering spatial functions, we were also influenced by the work of Kritchevsky (1988), who defined (but did not implement), a classification scheme for elementary spatial functions that include operations for spatial perception, spatial memory, spatial attention, spatial mental operations, and spatial construction. As well as attempting to capture much of the functionality derived from cognitive studies of behavior, we have been influenced by our desire to incorporate our tools in reasoning systems for knowledge-based system development. Thus, we have been concerned with issues such as efficiency and reusability of our primitive functions.

The implementation of the imagery functions assumes global variables corresponding to the current states of long-term and working memory. The primitive functions modify these states by retrieving images from memory, transforming the contents of working memory or storing new (or modified) images in long-term memory.

We consider the primitive functions for imagery in three classes corresponding to the three representations: deep, visual, and spatial. Functions for deep and visual memory have been considered previously in research areas such as semantic memory, vision, computational geometry, and graphics. Thus, we provide a brief overview of these classes and concentrate on the more novel aspect of our research, the functions for spatial reasoning. We also discuss the processes involved in transforming one representation into another, a powerful feature of our knowledge representation scheme. Note that the proposed functions have been specified using array theory and implemented in the programming language Nial.

Long-Term Memory Functions. The frame concept was initially proposed as a model for analogy-driven reasoning (Minsky, 1975). In the context of imagery, this type of reasoning involves the understanding of an image in a new context based on previously stored images. The functions for the deep representation of imagery are exactly those of the Nial Frame Language (Hache, 1986). In this language, imagery frames contain information describing images or classes of images, where knowledge is organized into slots that represent the attributes of an image.

Like most frame languages, the Nial frame language uses a semantic network approach to create configurations of frame taxonomies. The hierarchical network approach supports AKO links for implementing an inheritance mechanism within the frame structure. Frames in the language are implemented and manipulated as nested association lists of slots and values. Creating a generic or instance frame for an image requires assigning values to its slots, which is achieved using the function *fdefine*. Information is modified, added to, or deleted from an existing frame using the *fchange*, *fput,* and *fdelete* operators. Knowledge is retrieved (directly or through inheritance) from frames using the *fget* function. These and many other frame functions are implemented as part of the Nial AI Toolkit (Jenkins et al., 1988).

The decomposition of images into their components is an important concept of computational imagery. This is achieved through a PARTS slot that contains the meaningful parts of an image and their relative location. Because the spatial representation of an image is stored relative to a particular axis, an instance frame may also contain an ORIENTATION slot. As described later, the PARTS and ORIENTATION slots allow for reconstruction of the spatial representation of an image.

Functions for Visual Reasoning. Functions for visual reasoning have been studied extensively in areas such as machine vision and graphics. Similar to previous work, we consider visual images as surface or occupancy representations that can be constructed, transformed, and analyzed.

The occupancy array representation for the visual component of imagery can be constructed in a number of ways, depending on the domain of application. For example, the visual representation can be stored as generalized shape descriptions and regenerated at varying levels of resolution. They may also be reconstructed from geometric information stored in the deep representation.

Imagery functions for manipulating occupancy arrays include *rotate, translate,* and *zoom,* which change the orientation, location, or size of a visual image. Functions for retrieving *volume* and *shape* are also being implemented. Whereas many of these functions are generic, domain-specific functions can also be implemented for a particular application. For example, when considering molecular scenes we are concerned with a class of shape descriptors that correspond to the shape of molecular fragments at varying levels of abstraction (e.g., residues, secondary structure, molecule, etc.)

Functions for Spatial Reasoning. Whereas functions for visual and memory-based reasoning have been studied previously, the primitive functions for spatial imagery are more unique to our representation. The importance of spatial reasoning is supported by research in a number of areas, including computer vision, task planning, navigation for mobile robots, spatial databases, symbolic reasoning, and so on (Chen, 1990). Within the imagery context we consider spatial reasoning in terms of a knowledge representation framework that is general enough to apply to various problem domains. We also consider the relationship of spatial image representations to visual and deep representations.

As mentioned earlier, the functions for computational imagery are implemented assuming a global environment consisting of a frame knowledge base and the current working-memory representation. Generally, the

TABLE 1
Primitive Functions for Spatial Reasoning

Name	Mapping	Description
retrieve	$DM \times N \rightarrow WM$	Reconstruct spatial image
put	$WM \times N \times N \times L \rightarrow WM$	Place one image component relative to another
find	$WM \times N \rightarrow L$	Find location of component
delete	$WM \times N \rightarrow WM$	Delete image component
move	$WM \times N \times L \rightarrow WM$	Move image component to new location
turn	$WM \times Direction \rightarrow WM$	Rotate image 90°in specified direction
focus	$WM \times N \rightarrow WM$	Replace specified subimage with its spatial representation
unfocus	$WM \rightarrow WM$	Return to original image
store	$WM \times DM \times N \rightarrow DM$	Stores current image in long-term memory
adjacent	$WM \times N \rightarrow N^*$	Determine adjacent image components

working-memory representation consists of a single symbolic array (for spatial reasoning) or an occupancy array (for visual reasoning). One exception to this case is when we are using the spatial array to browse an image by focusing and unfocusing attention on particular subimages. In this case we need to represent working memory as a stack, where we push images onto the stack as we focus and pop images from the stack as we unfocus. Table 1 presents a summary of some of the functions for spatial imagery. We specify these functions as mappings with parameters corresponding to deep memory (DM), working memory (WM), image name (N) and relative or absolute location (L).

In order to reason with images, it is necessary to provide functions that allow us to interpret the spatial representations in terms of propositions within a given domain. For example, consider the three-term series problem: *John is taller than Mary, Sam is shorter than Mary, who is tallest?* It has been suggested that people represent and solve such a problem using an array where the spatial relationships correspond to the relative heights (Huttenlocker, 1968):

John	Mary	Sam

As discussed earlier, describing and solving such a problem using a propositional approach involves an exhaustive search of all the axioms describing the relation. The symbolic array representation allows direct access to such information using a domain-specific array theory function *tallest,* which returns the first element of the array:

tallest is operation A {first A}.

If our array is representing a map domain, we could similarly define domain-specific domain-specific functions for *north-of, east-of, bordering-on,* and so forth.

Cognitive theories for pattern recognition support the need for *attention* in imagery, where attention is defined as the ability to concentrate tasks on a component (or components) of an image. The concept of attention is achieved using the spatial representation by defining a global variable that corresponds to a region of attention (and possibly an orientation) in a spatial representation of an image and implementing functions that implicitly refer to this region. For example, we have defined functions that initialize a region of attention (*attend*), shift attention to a new region (*shift*), retrieve the components in the region of attention (*at-attend*), focus on region of attention to retrieve detail (*focus-attend*), and so on. These functions are particularly useful for applications where we wish to describe and reason about a scene from an internal, rather than external, perspective. Consider, for example, a motion-planning application where the spatial representation reflects the orientation and current location of the moving body.

Complex Functions for Imagery. Using the primitive functions for computational imagery we can design processes corresponding to more complex imagery tasks. For example, a function for visual pattern matching can be defined using the *rotation* and *translation* functions to align two visual representations of images, and a primitive *compare* function to measure the similarity between these occupancy arrays.

To retrieve properties of an image, it may be necessary to focus on details of subimages. For example, we may wish to determine all the regions of countries on the border of an arbitrary country *X*. This can easily be determined by applying the *focus* function to the countries *adjacent* to country *X* and then determining the *content* of these subimages. This can be expressed as the array theory function definition *border,* where the body of the definition is enclosed by the curly brackets:

border is operation X {content (EACH focus) adjacent X}.

A key feature of our approach to knowledge representation for imagery is the underlying array theory semantics, which allows us to consider all representations as array data structures and implement functions that transform one representation of an image to another. Figure 10 illustrates the transformations supported by the scheme. Although the implementation of functions used for storage, retrieval, and interpretation may be complex and domain specific, the primitive functions for imagery provide a basis for their implementation. For further details of the use of imagery for image interpretation in the domain of molecular scene analysis see Glasgow, Fortier, and Allen (1991).

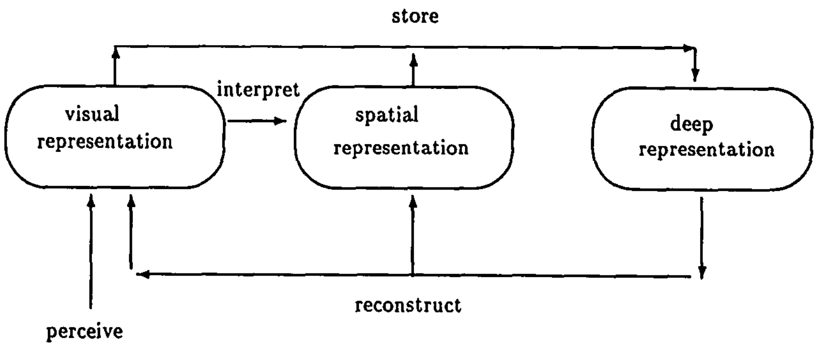


Figure 10. Stages of image representation

CONTRIBUTIONS OF COMPUTATIONAL IMAGERY

In the introduction we proposed three goals for our research in computational imagery: the cognitive science goal, the AI goal, and the applications goal. Combined, these goals attempt to address the fundamental question: *What are the underlying processes involved in mental imagery, and how can corresponding computational processes be efficiently implemented and used to solve real-world problems?* We do not believe that the three goals can be approached independently. The representations and functionality of computational imagery are motivated by empirical results from cognitive science, as well as the pragmatic needs of applications in AI. Also, the tools that have been developed for computational imagery can be used to implement and test cognitive theories and thus increase our understanding of mental imagery. In this section we discuss the major contributions of computational imagery to each of the prescribed goals.

Cognitive Science Goal

A primary objective of research in cognitive science is to study and explain how the mind works. One aspect of work in this area is the theory of computability. If a model is computable, then it is usually comprehensible, complete, and available for analysis; theories that are implemented can be checked for sufficiency and used to simulate new predictive results. In a discussion of the issues of computability of cognitive theories for imagery, Kosslyn (1980) expressed frustration with existing implementation tools:

> There is a major problem with this approach however; the program will not run without numerous "kluges," numerous ad hoc manipulations required by the realities of working with a digital computer and a programming language like ALGOL or LISP. (p. 137)

Kosslyn went on to state that:

> The ideal would be a precise, explicit language in which to specify the theory and how it maps into the program. (p. 138)

Array theory, combined with the primitive functions and representations for computational imagery, provides such a meta-language. Moreover, it allows us to represent an image either visually or spatially, and provides for the implementation and testing of alternative, and possibly conflicting, models for mental imagery.

Consider the problem of mental rotation. Although empirical observations conclude that rotation involves an object representation being moved through intermediate orientations (Shepard & Cooper, 1982), a still unresolved issue is the actual content of the representation used. One obvious representation is a visual depiction of the object that preserves detailed three-dimensional shape information. An alternative approach is one in which the object is represented as vectors corresponding to the major axes of the object (Just & Carpenter, 1985). This type of representation can be considered as spatial in nature: It preserves connectivity of parts but discards surface information about the image. Furthermore, whereas some researchers argue that images encode size (e.g., Kosslyn, 1980), others claim that mental images preserve information about relative positions but not size (e.g., Kubovy & Podgorny, 1981). This conflict, as possibly others, could be attributed to the different representations used by subjects in the different experimental tasks. Using the primitives of computational imagery and array theory, such theories could be simulated and analyzed. Although we are not interested in entering into the imagery debate, we suggest that such simulations could contribute to discussions in this area. As another example, consider that Pylyshyn's (1981) main criticism of depictive theories of imagery is that they confuse physical distance in the world with the representation of distance in the head. The visual representation for computational imagery does, in fact, attach a real distance to the representation, in terms of the number of cells in the array depicting the image. The spatial representation, on the other hand, does not preserve distance information. Thus, the distinct representations could be used to model conflicting theories of image scanning.

The use of abstract representations for storing and manipulating three-dimensional images has been supported by research in cognition. Attneave (1974) suggested that humans represent three-dimensional objects using an internal model that at some abstract level is structurally isomorphic to the object. This isomorphism provides a "what–where" connection between the visual perception of an object and its location in space. A similar connection exists between the visual and spatial representations for imagery.

The human brain is often compared to an information-processing system where computations can either be serial or parallel. Ullman (1984) suggested that there may be several forms of parallelism involved in mental imagery.

One form is spatial parallelism, which corresponds to the same operations being applied concurrently to different spatial locations in an image. Functional parallelism occurs when different operations are applied simultaneously to the same location. Funt (1983) also argued that many spatial problems are amenable to parallel processing. In developing a parallel computational model for the rotation problem, Funt was able to simulate the linear-time behavior corresponding to the human solution of the problem.

As well as allowing for multiple representations for testing cognitive theories, the array theory underlying computational imagery also provides both sequential and parallel constructs for specifying the processes involved in imagery. For example, the *EACH* transformer of array theory is a primitive second-order function that applies an operation to all of the arguments of an array, that is, $EACH\, f[A_1,\ldots,A_n] = [f(A_1),\ldots, f(A_n)]$. Thus, we could specify a spatial parallel operation such as *EACH focus,* which would simultaneously reconstruct all of the subimages in a given image. Functional parallelism can be captured using the *atlas* notation of array theory. An atlas is a list of functions that may be applied in parallel to an array. For example, the expression $[f_1, f_2, \ldots f_n]\ A$ specifies simultaneous application to the functions $f_1, \ldots, f_n$ to array A. Using the atlas construct and the functions of computational imagery we can specify such spatial parallelism as [*turn, move*], which expresses the simultaneous updating of working and deep memory to reflect the translation and rotation of an image.

A full study of the relationship between parallel processing in mental imagery and computational parallelism is a topic for future research. It has previously been demonstrated that the constructs of array theory are powerful enough to express a wide gambit of concurrent processing (Glasgow et al., 1989). It may then be possible to analyze the limitations of parallel processing in cognitive tasks by analyzing the limitations when specifying these in array theory; if we cannot express a parallel algorithm for a task, then perhaps it is inherently sequential, cognitively as well as computationally.

A detailed discussion of the relationship between mind and computer was presented by Jackendoff (1989), who addressed the issue of studying the mind in terms of computation. More specifically, Jackendoff suggested that to do so involves a strategy that divides cognitive science into studies of structure and processing. Our functional approach to computational imagery is complimentary to this philosophy; image representations are array data structures, which can be considered distinctly from the array functions that operate on them. Jackendoff also supported the possibility of different levels of visual representation with varying expressive powers.

In summary, the underlying mathematics for computational imagery satisfies Kosslyn's ideal by providing a precise and explicit language for specifying theories of mental imagery. Visual and spatial representations are implemented as arrays and manipulated using the primitive functions of computational imagery, which themselves are expressed as array theory operations. Finally, the primitives of array theory and computational imagery have been directly mapped into Nial programs, which run without any "kluges" or "ad hoc manipulations." Note that the theory can also provide the basis for other implementations of computational imagery, as illustrated by the Lisp implementation of Thagard and Tanner (1991).

AI Goal

AI research is concerned with the discovery of computational tools for solving hard problems that rely on the extensive use of knowledge. Whereas traditional approaches to knowledge representation have been effective for linguistic reasoning, they do not always embody the salient visual and spatial features of an image. Also, they do not allow for an efficient implementation of the operations performed on this information, such as comparing shapes and accessing relevant spatial properties.

Whereas representations and operations for visual reasoning have previously been studied in imagery, as well as other areas such as computer vision and graphics, there has been little attention given to knowledge representations for spatial reasoning. We suggest that the proposed scheme for representing and manipulating spatial images has several advantages over visual or propositional representations. First, the spatial structure imposed by symbolic arrays supports efficient, nondeductive inferencing. Furthermore, the symbolic array representation for images can deal more easily with dynamic environments.

The symbolic array representation for computational imagery has also provided the basis for analogical reasoning in spatial problems (Conklin & Glasgow, 1992; Glasgow, 1991). A thesis of this work is that the structural aspects of images, in particular the spatial relations among their parts, can be used to guide analogical access for spatial reasoning. Preliminary results in the conceptual clustering of chess game motifs has illustrated that computational imagery can be applied to the area of image classification. Currently, we are extending this work to include classification of molecular structures based on spatial analogies (Conklin, Fortier, Glasgow, & Allen, 1992).

Applications Goal

Since the time of Aristotle, imagery has been considered by many as a major medium of thought. Einstein stated that his abilities did not lie in mathematical calculations but in his visualization abilities (Holton, 1971). Similarly, the German chemist Kekulé stated that it was spontaneous imagery that led him to the discovery of the structure of benzene (MacKenzie, 1965). Mental simulations provide insights that contribute to effective problem-solving techniques. Thus, it is only natural to use the representations and functions of computational imagery to develop knowledge-based systems that incorporate the imagery problem-solving paradigm. One such system is an application to the problem of molecular scene analysis (Glasgow et al., 1991), which combines tools from the areas of protein crystallography and molecular database analysis, through a framework of computational imagery.

In determining structures, crystallographers relate the use of visualization or imagery in their interpretation of electron density maps of a molecular scene. These maps contain features that are analyzed in terms of the expected chemical constitution of the crystal. Thus, it is natural for crystallographers to use their own mental recall of known molecular structures, or of fragments thereof, to compare with, interpret, and evaluate the electron density features. Because molecular scenes can be represented as three-dimensional visual or spatial images, this mental pattern recognition process can be implemented using the primitive functions of computational imagery.

In molecular scene analysis, we attempt to locate and identify the recognizable molecular fragments within a scene. As in Marr's (1982, p. 3) definition of computational vision, it is the "process of discovering what is present in the world, and where it is." The integrated methodology for molecular scene analysis is being implemented as a knowledge-based system, through the development of five independent, communicating processes: (1) retrieval and reconstruction of visual representation of anticipated motifs from the long-term memory (deep representation) of molecular images; (2) enhancement and segmentation of the visual representation of the three-dimensional electron density map molecular scene; (3) visual pattern matching of the segmented image features with the retrieved visual motifs; (4) analysis and evaluation of the hypothesized, partially interpreted spatial representation of the perceived image; and (5) resolution and reconstruction of the molecular image. These processes are applied iteratively, resulting in progressively higher resolution images, until ultimately, a fully interpreted molecular scene is reconstructed.

The organization of the comprehensive information of crystal and molecular structures into a deep representation is crucial to the overall strategy for molecular scene analysis. This representation stores concepts and instances of molecular scene in terms of their structural and conceptual hierarchies. A serious problem in this domain, and in general, is to find appropriate visual and spatial depictions. This involves determining what features (visual or spatial) we wish to preserve in each of the representations. Initial algorithms have been developed to construct visual representations that depict the surface structure of an image and spatial representations that preserve bonding and symmetry information. Whether these are the most appropriate structures for all our reasoning in the domain is still an open question.

A full implementation of the knowledge-based system for molecular scene analysis is an ambitious and on-going research project. To date, we have been encouraged by preliminary results in the development of a long-term memory model (deep representation) for molecular scenes and the implementation of some of the essential tasks of molecular imagery. These tasks include transforming geometric information into spatial and visual representations, evaluation of partially interpreted images, classification and retrieval of images, and visual and spatial comparison of molecular scenes.

Although molecular scene analysis shares many features with visual scene analysis, it also differs in many ways. Both tasks involve segmentation of perceived images, retrieval and reconstruction of image templates, and pattern matching for object classification. The problem of molecular scene analysis is more tractable, however. Molecular images are perceived in three dimensions, thus eliminating the bottleneck of early vision routines. As well, the molecular domain is highly constrained: Molecular interactions and symmetry constraints impose hard restrictions on the image representations. Finally, there exists a wealth of knowledge about molecular scenes and molecular interactions in existing crystallographic databases. Using machine-learning techniques, we hope, ultimately, to generalize, correlate, and classify this information.

Although molecular scene analysis is only one of many potential applications for computational imagery, we feel that it is important to apply our reasoning paradigm to a complex problem that involves extensive imagery abilities when carried out by humans. Because of the experience embodied in existing crystallographic databases and algorithms, the availability of experts in the field and the natural constraints that exist in the domain, we believe that the important and real problem of molecular image reconstruction is an ideal test case for the concepts and implementations of computational imagery. It also suggests that the multiple representations of the scheme provide the framework for a complete computational model for the complex reasoning tasks involved in scene analysis.

Other potential applications for imagery-based systems include haptic perception and medical imaging. Literature in haptic perception provides evidence for an interdependence between haptic perception and visual imagery (Katz, 1989). Of special interest, are applications such as motion planning and game playing, which combine spatial and temporal reasoning. As suggested earlier, the spatial representation for computational imagery facilitates nondeductive reasoning, thus precluding many of the nonmonotonicity problems involved in deductive approaches in these areas. Preliminary work in imagery and machine learning has demonstrated that the spatial representation for imagery can be used to depict and reason about structural motifs in a chess game (Conklin & Glasgow, 1992). As well, the representations for computational imagery have been used to describe the role of visual thinking in such complex domains as atomic theory development (Thagard & Hardy, 1992).

DISCUSSION

This article introduces the concept of computational imagery, which treats imagery as a problem-solving paradigm in AI. By proposing a knowledge representation scheme that attempts to capture the fundamental principles of mental imagery, we provide a foundation for implementing systems relying on imagery-based reasoning.

Aside from related research in perception and early work in frame representations, the AI community has given little attention to the topic of imagery. Thus, we rely on relevant theories of cognition to provide initial guidance for our research. We are also driven by the need to apply the scheme to real-world applications. The representation scheme is not intended to be a model of mental imagery; we do not claim that in human working memory two "mind's eyes" exist that process visual and spatial representations identical to the ones we have implemented. What we do suggest is that the internal image representations are informationally equivalent to representations involved in our scheme, that is, information in one representation is inferable from the other (Larkin & Simon, 1987).

The knowledge representation scheme for computational imagery includes three image representations, each appropriate for a different kind of information processing. A set of primitive functions, corresponding to the fundamental processes involved in mental imagery, has been designed using the mathematics of array theory and implemented in the functional array language Nial. These functions provide the building blocks for more complex imagery-related processes.

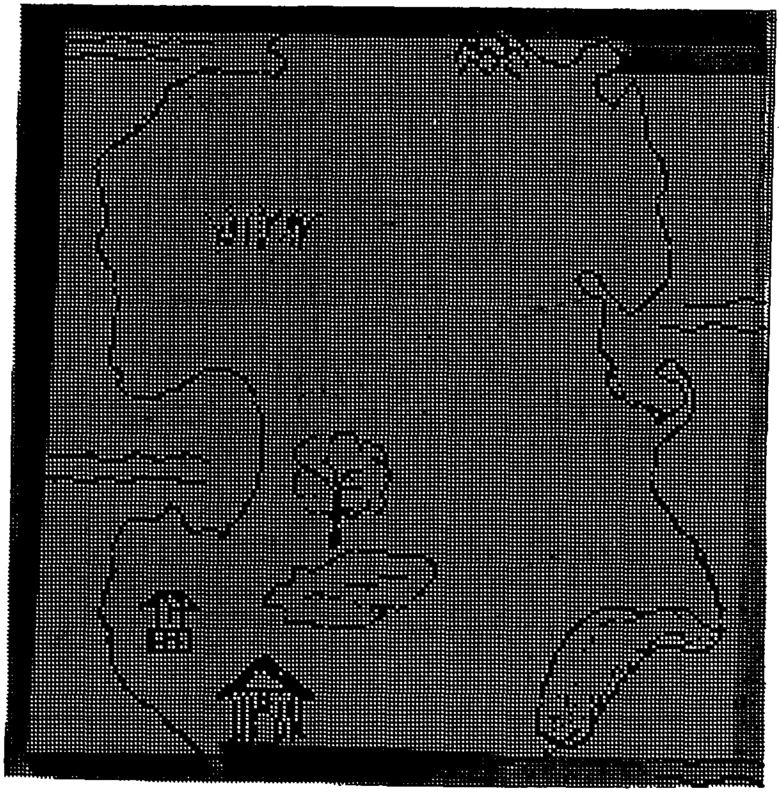
The most relevant previous contribution to imagery is the work of Kosslyn (1980), who proposed a computational theory for mental imagery. In that theory, images have two components: a surface representation (a quasi-pictorial representation that occurs in a visual buffer), and a deep representation for information stored in long-term memory. Like Kosslyn, we consider a separate long-term memory model for imagery, that encodes visual information descriptively. Unlike Kosslyn, we consider the long-term memory to be structured according to the decomposition and conceptual hierarchies of an image domain. Thus, we use a semantic network model, implemented using frames, to describe the properties of images. The long-term memory model in Kosslyn's theory is structured as sets of lists of propositions, stored in files.

The surface representation in Kosslyn's theory has been likened to spatial displays on a cathode ray tube screen; an image is displayed by selectively filling in cells of a two-dimenstional array. Our scheme for representing images in working memory is richer in two important ways. First, we treat images as inherently three dimensional, although two-dimensional images can be handled as special cases. As pointed out by Pinker (1988), images must be represented and manipulated as patterns in three dimensions, which can be accessed using either an object-centered or a world-centered coordinate system. Second, we consider two working-memory representations, corresponding to the visual and spatial components of mental imagery. Just as the long-term memory stores images hierarchically, the visual and spatial representations use nested arrays to depict varying levels of resolution or abstraction of an image. Although the functionality of many of the primitive operations for computational imagery were initially motivated by the processes defined by Kosslyn's theory, their implementation varies greatly because of the nature of the image representations.

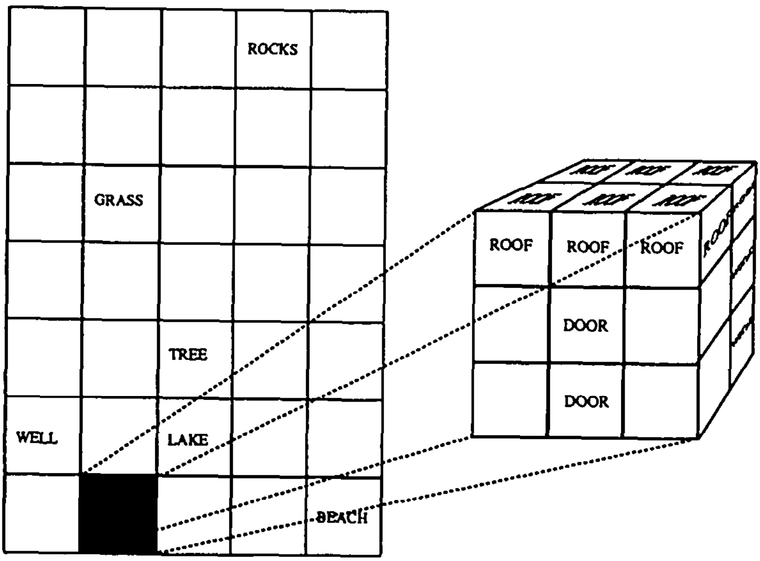
Possibly the most important distinction between our approach to computational imagery and Kosslyn's computational theory is the underlying motivation behind the two pieces of work. Kosslyn's model was initially developed to simulate and test a particular theory for mental imagery. Whereas computational imagery can be used to specify and implement cognitive theories, its development was based on the desire to construct computer programs to solve hard problems that require visual and spatial reasoning. Thus, efficiency and expressive power, not predictive and explanatory power, are our main concerns.

As a final illustration of the knowledge representation scheme, consider the island map used by Kosslyn (1980) to investigate the processes involved in mental image scanning. Figure 11 presents a visual depiction of such a map, as well as a spatial representation that preserves the properties of closeness (expressed as adjacency) and relative location of the important features of the island. It does not attempt to preserve relative distance. Consider answering such questions as: *What is the shape of the island? Is the beach or the tree closer to the hut?* These properties can be retrieved using the visual representation of the map. For example, we could analyze the surface of the island and compare this with known descriptions in the deep representation to retrieve shape information. Now consider the queries: *What is north of the tree? What is the three-dimensional structure of the hut?* Although it may be possible to derive this information from the visual representation, it would be a costly process. Using the symbolic array representation, however, we can easily access and retrieve spatial information using an efficient constrained search procedure. Although it may be argued that it is also initially costly to construct the spatial representation, the process of determining the structure of this representation can be carried out once, and then the results stored in the deep representation for later use.

More detailed information can be accessed from the spatial representation using the *focus* function to construct and inspect spatial images at lower levels of the structural hierarchy. For this particular example, there is not sufficient information to determine all of the three-dimensional features of the hut from the two-dimensional visual depiction. Using the computational imagery paradigm, which incorporates inheritance in the deep representation, we can construct the three-dimensional symbolic array using



Visual Representation



Spatial Representation

Figure 11. Visual and spatial representation of Kosslyn's (1980) island map

information stored in the generic frame for the concept "hut" to fill in missing details.

It is worth noting here that the spatial representation is not just a low-resolution version, or approximation, of the visual representation of an image. As well as capturing the structural hierarchy of an image, the symbolic array may discard, not just approximate, irrelevant visual information. For example, in particular molecular applications we are primarily concerned with bonding information, which is made explicit using adjacency in a three-dimensional symbolic array. Visual and spatial properties such as size, distance, relative location (i.e., *above, behind, left-of,* etc.) may not be important for such applications and thus are not preserved.

Another approach to visual reasoning was presented by Funt (1980), who represented the state of the world as a diagram, and actions in the world as corresponding actions in the diagram. Similar to Kosslyn, Funt used two-dimensional arrays to denote visual images. A more recent model describes how visual information can be represented within the computational framework of discrete symbolic representations in such a way that both mental images and symbolic thought processes can be explained (Chandrasekaran & Narayanan, 1990). Although this model allows a hierarchy of descriptions, it is not spatially organized.

One way of evaluating our approach to computational imagery is in terms of the fundamental principles of mental imagery, as described in Finke (1989). In particular, the scheme was designed around the principle of *implicit encoding,* which states that imagery is used to extract information that was not explicitly stored in long-term memory. We retrieve information such as shape and size using the visual representation and information pertaining to the relative locations of objects in an image using the spatial representation for working memory. The principle of *perceptual equivalence* is captured by our assumption that perception and imagery share common representations. In fact, the processes involved in transforming a visual representation to a spatial representation are just those of scene analysis: taking a raw, uninterpreted image (visual representation) and identifying the subcomponents and their relative positions (spatial representation). The spatial representation captures the principle of *spatial equivalence,* because there is a correspondence between the arrangement of the parts of a symbolic array of an image, and the arrangement of the actual objects in the space. Note, though, that Finke argued for a continuous space of mental images, whereas the spatial representation assumes a discrete space. The principle of *structural equivalence* is preserved by the deep and the spatial representations, which capture the hierarchical organization of images. Furthermore, images in our representation scheme can be reorganized and reinterpreted. The scheme captures the functionality required of the principle of *transformational equivalence* by providing primitive array functions that can be used to manipulate both the visual and spatial representations of images.

When questioned on the most urgent unresolved difficulties in AI research, Sloman (1985) replied:

> I believe that when we know how to represent shapes, spatial structures and spatial relationships, many other areas of AI will benefit, since spatial analogies and spatial modes of reasoning are so pervasive. (pp. 386-387)

Experimental results suggest that people use mental imagery for spatial reasoning. Thus, by facilitating an efficient implementation of the processes involved in mental imagery, computational imagery provides a basis for addressing the difficulties suggested by Sloman and developing AI systems that rely on representing, retrieving, and reasoning about visual and spatial properties of images.

REFERENCES

Allen, F.H., Bergerhoff, G., & Sievers, R. (Eds.). (1987). *Crystallographic databases.* Chester, England: IUCr.

Anderson, J.R. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85,* 249-277.

Attneave, F. (1974). Apparent movement and the what-where connection. *Psychologica, 17,* 108-120.

Atwood, G.E. (1971). An experimental study of visual imagination and memory. *Cognitive Psychology, 2.*

Baddeley, A.D., & Lieberman, K. (1980). Spatial working memory. In R. Nickerson (Ed.), *Attention and performance* (Vol. 8, pp. 521-617). Hillsdale, NJ: Erlbaum.

Bauer, R.M., & Rubens, A.B. (1985). Agnosia. In K.M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (pp. 187-241). New York: Oxford University Press.

Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review, 94,* 115-147.

Block, N. (Ed.). (1981). *Imagery.* Cambridge, MA: MIT Press.

Brooks, L.R. (1968). Spatial and verbal components in the act of recall. *Canadian Journal of Psychology, 22,* 349-368.

Chandrasekaran, B., & Narayanan, N.H. (1990). Integrating imagery and visual representations. *Proceedings of the 12th Annual Conference of the Cognitive Science Society* (pp. 670-677). Hillsdale, NJ: Erlbaum.

Chen, S. (Ed.). (1990). *Advances in spatial reasoning* (Vols. 1 & 2). Norwood, NJ: Ablex.

Collins, A.M., & Quillian, M.P. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240-247.

Conklin, D., Fortier, S., Glasgow, J.I., & Allen, F.A. (1992). Discovery of spatial concepts in crystallographic databases. *Proceedings of the Machine Discovery Workshop* (pp. 111-116).

Conklin, D., & Glasgow, J.I. (1992). Spatial analogy and subsumption. In D. Sleeman & P. Edwards (Eds.), *Machine Learning: Proceedings of the Ninth International Conference ML* (92) (pp. 111-116). San Mateo: CA: Morgan Kaufmann.

De Renzi, E. (1982). *Disorders of space exploration and cognition.* New York: Wiley.

Farah, M.J. (1988a). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review, 95,* 307-317.

Farah, M.J. (1988b). The neuropsychology of mental imagery: Converging evidence from brain-damaged and normal subjects. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi (Eds.), *Spatial cognition—Brain bases and development* (pp. 33-56). Hillsdale, NJ: Erlbaum.

Finke, R.A. (1989). *Principles of mental imagery*. Cambridge, MA: MIT Press.
Finke, R.A., Pinker, S., & Farah, M.J. (1989). Reinterpreting visual patterns in mental imagery. *Cognitive Science, 13*, 51–78.
Finke, R.A., & Slayton, K. (1988). Explorations of creative visual synthesis in mental imagery. *Memory & Cognition, 16*, 252–257.
Funt, B.V. (1980). Problem solving with diagrammatic representations. *Artificial Intelligence, 13*, 201–230.
Funt, B.V. (1983). A parallel-process model of mental rotation. *Cognitive Science, 7*, 67–73.
Glasgow, J.I. (1990). Artificial intelligence and imagery. In S.M. Humphrey & B.H. Kwasnick (Eds.), *Advances in classification research: Proceedings of the 1st ASIS SIG/CRF Classification Workshop*. Learned Information, Inc. for the American Society for Information Science.
Glasgow, J.I. (1991). Imagery and classification. In S.M. Humphrey (Ed.), *Advances in classification research and application*. Medford, NJ: Learned Information.
Glasgow, J.I., & Browse, R. (1985). Programming languages for artificial intellgence. *Computers and Mathematics with Applications, 11*, 431–448.
Glasgow, J.I., Fortier, S., & Allen, F.H. (1991). Crystal and molecular structure determination through imagery. *Proceedings of the Seventh Conference on Artificial Intelligence* (pp. 98–105).
Glasgow, J.I., Jenkins, M.A., Blevis, E., & Feret, M. (September, 1991). Logic programming with arrays. *IEEE Transactions on Knowledge and Data Engineering, 3*(3), 307–319.
Glasgow, J.I., Jenkins, M.A., McCrosky, C., & Meijer, H. (1989). Expressing parallel algorithms in Nial. *Parallel Computing, 11*, 331–347.
Hache, L. (1986). *The Nial frame language*. Unpublished master's thesis, Queen's University, Kingston, Canada.
Hinton, G. (1979). Some demonstrations of the effects of structural descriptions in mental imagery. *Cognitive Science, 3*, 231–250.
Holton, G. (1971). On trying to understand scientific genius. *American Scholar, 41*, 95–119.
Huttenlocker, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review, 4*, 277–299.
Intons-Peterson, M.J. (1983). Imagery paradigms: How vulnerable are they to experimenters' expectations? *Journal of Experimental Psychology: Human Perception and Performance, 9*, 394–412.
Israel, D.J. (1987). Some remarks on the place of logic in knowledge representation. In N. Cercone & G. McCalla (Eds.), *The knowledge frontier* (pp. 80–91). New York: Springer Verlag.
Jackendoff, R. (1989). *Consciousness and the computational mind*. Cambridge, MA: MIT Press.
Jannsen, W.H. (1976). Selective interference during the retrieval of visual images. *Quarterly Journal of Experimental Psychology, 28*, 535–539.
Jenkins, M.A., Glasgow, J.I., Blevis, E., Chau, R., Hache, E., & Lawson, E. (1988). The Nial AI toolkit. *Annual Proceedings of Avignon '88 Eighth International Workshop on Expert Systems and their Applications, 2*, 37–52.
Jenkins, M.A., Glasgow, J.I., & McCrosky, C. (1986). Programming styles in Nial. *IEEE Software, 86*, 46–55.
Jenkins, M.A., & Jenkins, W.H. (1985). *The Q'Nial Reference Manual*. Kingston, Canada: Nial Systems Ltd.
Just, M.A., & Carpenter, P.A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review, 92*, 137–172.
Katz, D. (1989). *The world of touch*. Hillsdale, NJ: Erlbaum.
Kosslyn, S.M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
Kosslyn, S.M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review, 94*, 148–175.

Kosslyn, S.M., & Pomerantz, J.P. (1977). Imagery, propositions, and the form of internal representations. *Cognitive Science, 9,* 52–76.

Kritchevsky, M. (1988). The elementary spatial functions of the brain. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi (Eds.), *Spatial cognition—brain bases and development.* Hillsdale, NJ: Erlbaum.

Kubovy, M., & Podgorny, P. (1981). Does pattern matching require the normalization of size and orientation? *Perception and Psychophysics, 30,* 24–28.

Larkin, J., & Simon, H.A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 10,* 65–99.

Levine, D.N., Warach, J., & Farrah, M.J. (1985). Two visual systems in mental imagery: Dissociation of "what" and "where" in imagery disorders due to bilateral posterior cerebral lesions. *Neurology, 35,* 1010–1018.

Levine, M., (1978). A knowledge-based computer vision system. In A. Hanson & E. Riseman (Eds.), *Computer vision systems* (pp. 335–352). New York: Academic.

Lindsay, R.K. (1988). Images and inference. *Cognition, 29,* 229–250.

MacKenzie, N. (1965). *Dreams and dreaming.* London: Aldus Books.

Marr, D. (1982). *Vision: A computational investigation in the human representation of visual information.* San Francisco: Freeman.

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three dimensional shapes. *Proceedings of the Royal Society, B200,* 269–294.

Marschark, M., Richman, C.L., Yuille, J.C., & Hunt, R. (1987). The role of imagery in memory: On shared and distinctive information. *Psychological Bulletin, 102,* 28–41.

McDermott, D.V., & Davis, E. (1984). Planning routes through uncertain territory. *Artificial Intelligence, 22,* 107–156.

Minsky, M. (1975). A framework for representing knowledge. In P. Winston (Ed.), *The Psychology of computer vision* (pp. 211–277). New York: McGraw Hill.

Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neuroscience, 6,* 414–417.

More, T. (1981). Notes on the diagrams, logic and operations of array theory. In P. Bjorke & O. Franksen (Eds.), *Structures and operations in engineering and management systems.* Trondheim, Norway: Tapir Publishers.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition, 3,* 635–647.

Papadias, D., & Glasgow, J.I. (1991). A knowledge representation scheme for computational imagery. *Proceedings of the 13th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition, 18,* 1–63.

Pinker, S. (1988). A computational theory of the mental imagery medium. In M. Denis, J. Engelkamp, & J.T.E. Richardson (Eds.), *Cognitive and Neuropsychological Approaches to Mental Imagery* (pp. 17–36). Dordecht/Boston/Lancaster: Martinus Nijhorff.

Pylyshyn, Z.W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80,* 1–24.

Pylyshyn, Z.W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 88,* 16–45.

Raphael, B. (1971). The frame problem in problem-solving systems. In X. Findler & X. Meltzer (Eds.), *Artificial intelligence and heuristic programming* (pp. 159–169). Edinburgh: Edinburgh University Press.

Ratcliff, G. (1982). Disturbances of spatial orientation associated with cerebral lesions. In X. Potegal (Ed.), *Spatial abilities: Development and physiological foundations* (pp. 301–331). New York: Academic.

Reed, S.K. (1974). Structural descriptions and the limitations of visual images. *Memory & Cognition, 2,* 329–336.

Shepard, R.N., & Cooper, L.A. (1982). *Mental images and their transformations.* Cambridge, MA: MIT Press.
Sloman, A. (1985). Artificial intelligence-Where are we. In Bobrow & Hayes (Eds.), *Artificial Intelligence, 25*(1), 386.
Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25,* 207–222.
Thagard, P., & Hardy, S. (1992). Visual thinking in the development of Dalton's atomic theory. In J. Glasgow & R. Hadley (Eds.), *Proceedings of AI '92* (pp. 30–37).
Thagard, P., & Tanner, C. (1991). *Lisp implementation for computational imagery* [Computer program]. Princeton, NJ: Princeton University.
Ullman, S. (1984). Visual routines. In S. Pinker (Ed.), *Visual Cognition* (pp. 97–159). Cambridge, MA: MIT Press.
Winston, P.H. (1975). Learning structural descriptions from examples. In P.H. Winston (Ed.), *Psychology of computer vision.* New York: McGraw-Hill.