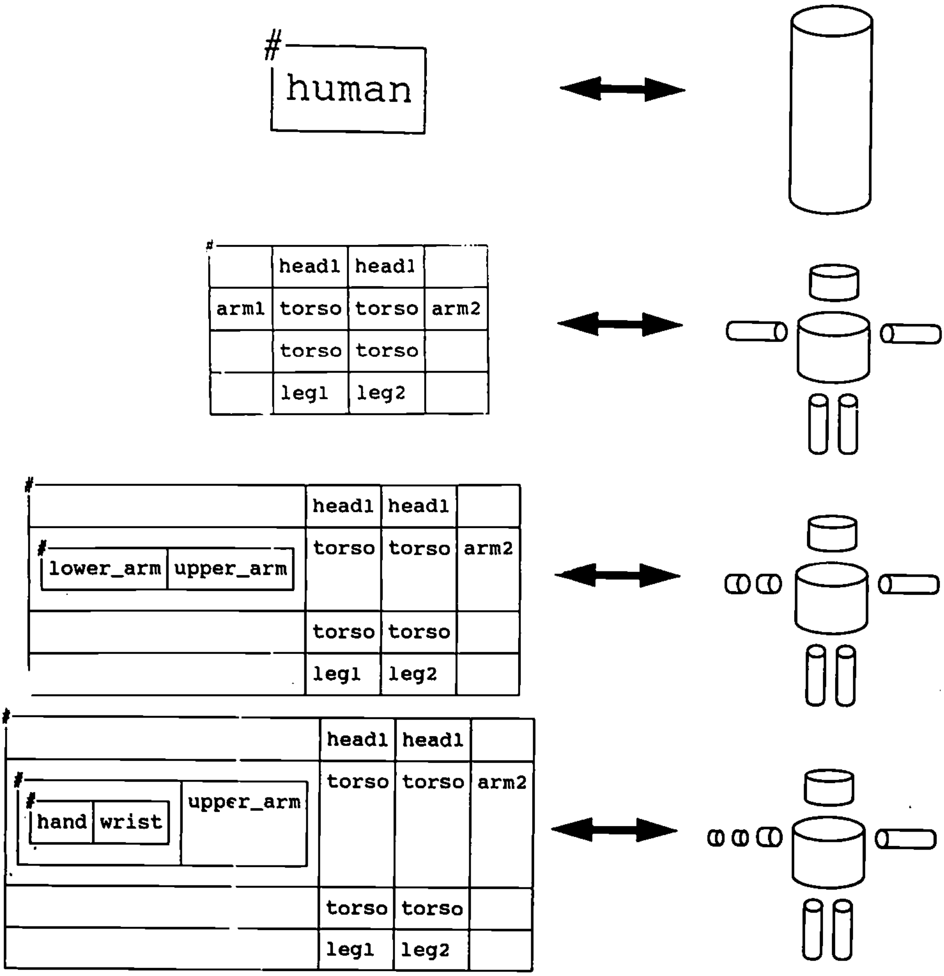


Figure 9. Spatial representation for topological shape description

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:

$$\textit{tallest is operation } A \ \{\textit{first } A\}.$$

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).

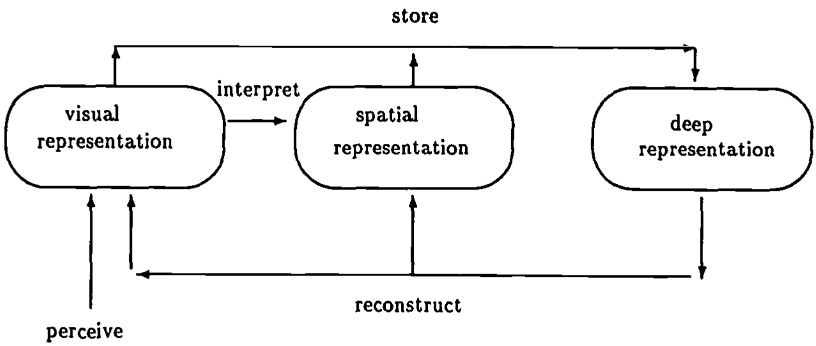


Figure 10. Stages of image representation

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.

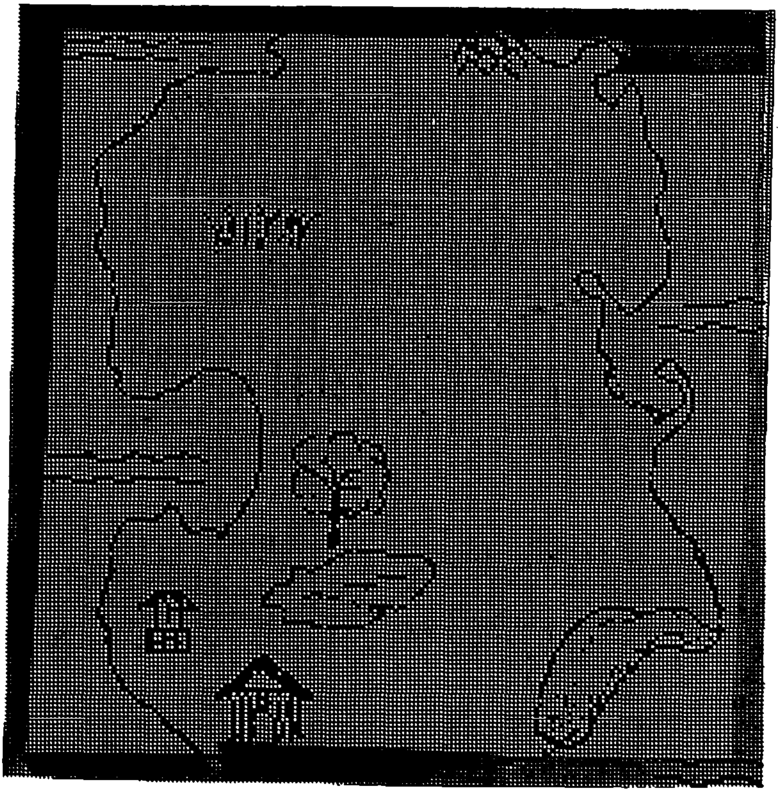
The most relevant previous contribution to imagery is the work of Kosslyn (1980), who proposed a computational theory for mental imagery. In that theory, images have two components: a surface representation (a quasi-pictorial representation that occurs in a visual buffer), and a deep representation for 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.

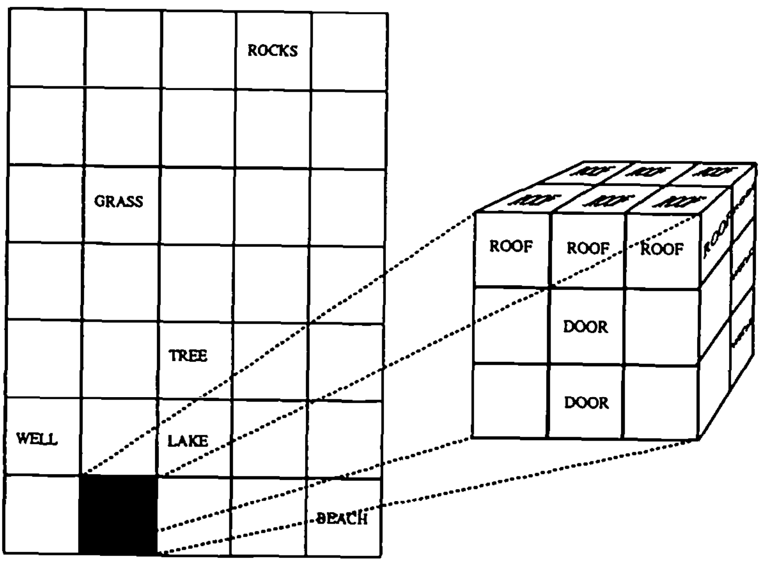
Possibly the most important distinction between our approach to computational imagery and Kosslyn's computational theory is the underlying motivation behind the two pieces of work. Kosslyn's model was initially developed to simulate and test a particular theory for mental imagery. Whereas computational 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.