Scheme of Color and Shape Invariant Image Retrieval from Databases *

J. M. Fuertes†, M. Lucena†, N. Pérez de la Blanca‡, J. Fdez-Valdivia‡
†Departamento de Informática. Escuela Politécnica Superior de Jaén
Universidad de Jaén - SPAIN
email: {jmf;mlucena}@ujaen.es
‡Departamento de Ciencias de la Computación e I.A. ETSI Informática
Universidad de Granada - SPAIN
email: {nicolas;J.Fdez-Valdivia}@decsai.ugr.es

Abstract

This paper presents a scheme of image retrieval from a database using queries prompted by the color and the shape of the objects present in the different scenes. Of the whole scheme of image retrieval, we focus attention on the modules that allow feature extraction of the component objects from the scenes and the matching of the objects among the different images. The defined scheme enables the indexing of images by measuring the similarity between the integral objects.

Key words: Color image segmentation, color spaces, image retrieval, mathematical morphology, deformable models.

1 Introduction

One of the paradigms, that in the last decade has experienced a wide ranging development throughout computer vision, has been the study of the access techniques of large data bases of graphic and color based documents using key images [7][9].

In this paper, we show the development of a full scheme of data storage and retrieval from an image based database. The attributes used in the interaction with the database will be color and shape. We concentrate on the modules of feature extraction and image matching.

In the selection process of the images that fulfil the conditions expressed in the query, it is necessary to use some kind of similarity ranking, due to the matching that is required between the extracted values of the original image and those given in the query image. In our case, the image matching is carried out by using an index based on the similarity between the component objects of the scenes, on the basis of color and shape attributes in terms of which the query is formulated.

Recent works reveal increasing interest in the use color as an attribute for image segmentation [8][17] as well as for image retrieval from large databases [11][9][7].

The description of the shape has been developed with the theory of non-rigid deformation. This approach considers two shapes to be *similar* according to the amount of deformation that would have to be applied to the object in order to transform it into a different one. There are many references related to the idea of describing objects with a model of non-rigid deformation [20][1][15]. We use here the main ideas from the Pentland deformable model [15], but instead we apply the mode calculation method of the model proposed by Nastar [13].

One of the main features to include in a content-based image retrieval scheme is the recovering of scenes containing some specific object defined in the query. In this sense, we propose a preliminar segmentation of the images for obtaining homogeneous regions supported in a hierarchical structure of clustering and then, we use the color and shape attributes for establishing the similarity betwwen scenes. Our scheme, allows us, as contrasted with other works [7][9][11], to recover images from complete scenes or from objects contained in them.

The layout of this paper is as follows: In section 2, we show the feature extraction of the component objects in the scene. In section 3, the similarity between the different objects are shown. In section 4, we detail our

*This research was partially sponsored by the Dirección General de Enseñanza Superior CICYT, under grant PB98-1374

experiments and finally in section 5, the conclusions of this paper are presented.

2 Feature Extraction

Feature extraction plays an important role in the proposed scheme of image retrieval. We segment the scenes into their component and then we determine the features which allows us to identify those homogeneous regions.

2.1 Edge detection in color images

In order to determine the local variations of color in the image, we need to define a space of representation, a distance and an order inside the chosen space color.

In this paper we use (HSI) as the representation space and it is upon this, that the mechanisms of segmentation have been developed, although the general scheme of image retrieval presented would remain valid with any other form of representation.

In [6] we can find a more detailed development of the above space.

The color space used is inscribed inside a cylinder for which the distance between the two points (colors) $P_1 = (H_1, S_1, I_1)$, $P_2 = (H_2, S_2, I_2)$ is given by:

$$d = ((d_I)^2 + (d_C)^2)^{1/2} \tag{1}$$

where

$$d_I = |I_1 - I_2| \tag{2}$$

and

$$d_C = \begin{cases} ((S_1)^2 + (S_2)^2 - 2S_1S_2\cos\phi)^{1/2} & \text{, if } P_1 \text{ and } P_2 \text{ are chromatic} \\ 0 & \text{, if } P_1 \text{ and } P_2 \text{ are achromatic} \\ S_1 - \Omega_{I_2} & \text{, if } P_1 \text{ is chromatic and } P_2 \text{ is achromatic} \end{cases} \tag{3}$$

gives

$$\Omega_{I_2} = \text{Chromaticity threshold for } I_2 \tag{4}$$

$$\phi = \begin{cases} |H_1 - H_2| & \text{if } |H_1 - H_2| < \pi \\ 2\pi - |H_1 - H_2| & \text{if } |H_1 - H_2| > \pi \end{cases} \tag{5}$$

The value d_I is the difference of intensity between the two points, d_c is the difference between the projection of the points in a chromatic plane, and ϕ is the smallest angle that separates H_1 and H_2. The chromaticity threshold Ω establishes the separation of the colour representation space into two regions: the *chromatic region* and the *achromatic*. This separation responds to the perceptual necessity of using intensity, saturation or tone when determining borders using the colour points analysed. In [6] we can find an empirical approach to its calculation.

As there is not a universally accepted natural way of ordering colors, we determine the order in a simple practical way: the lexicographic order. In our case, given the two pixels (H_1, S_1, I_1) and (H_2, S_2, I_2) we first compare the values of I, then those of S and finally those of H. To compare the latter component, the criteria used is as follows: if the smallest angle to pass from H_1 to H_2 is positive, we choose (H_2, S_2, I_2) as the larger element and in other case, we choose (H_1, S_1, I_1).

One of the classic methods of the detection of discontinuities is to use the gradient vector (magnitude and direction). According to the color representation space discussed in this paper, we have used first order masks based on Gaussian, extending the Canny's filter [2], from monochrome images to color images.

In the discrete case, the convolution of a grey level image with a Gaussian first partial derivative mask can be carried out through the successive convolution of the image with two perpendicular unidimensional masks in the following way:

$$\sum_k m_k x_{i+k} - m_k x_{i-k} \tag{6}$$

where the values m_k represent the mask coefficients and the (x_{i+k}, x_{i-k}) represent the image coefficients.

Grouping the terms and pairing the corresponding pixels equidistant from where the mask is centred, we obtain the following equation:

$$\sum_k m_k (x_{i+k} - x_{i-k}) = \sum_k m_k c(x_{i+k}, x_{i-k}) d(x_{i+k}, x_{i-k}) \tag{7}$$

$d(x_i, x_j)$ being the distance between the colors of the pixels x_i y x_j and

$$c(x_i, x_j) = \begin{cases} 1 & \text{if } x_i > x_j \\ -1 & \text{if } x_i \leq x_j \end{cases} \tag{8}$$

the ordering function above described.

2.2 Image segmentation

The segmentation algorithm proposed uses the information obtained from the gradients, associated to their corresponding boundaries, using a growth region. The main disadvantage of the use of any growth technique is due to the necessity of correctly placing the seeds around which the regions will grow. The detected boundaries are used as *containment walls* for the flooding carried out previously.

The proposed algorithm, detailed in [6][5], calculates two basic steps:

- Sewing the seeds along the detected boundaries.
- Use of the gradient data to carry out the propagation of the seeds.

In the Figure 1 is possible to observe the process of growth of regions starting from the map of edges of the original image, on which the seeds have been located.

The process of simplification of the frontier segments assisting to the location of the seeds can cause that different recovered regions appear on oneself object. To avoid that redundancy of regions is inserted, in a last step, a fused of the same ones. The approach of merge is simple: if the frontier of union of two adjacent regions does not appear in the map of edges they unite both to form a bigger one.

The fusion process also allows recovered regions considered as small can be absorbed by their neighboring regions. The Figure 1.f show (a label of black color has been chosen for its best visualization) the final grouping of regions assisting to an approach of minimum distance among the colors means of the considered regions.

The determined homogeneous regions completely cover the image and hence the proposed method allows us to make a *data organisation*. The simplification of the data allows for the design of structures, capable of collecting the topological relations of the different regions. Joining said regions, according to a similarity criteria, allows us to represent the elements of the scenes. If this approach allows for the definition of hierarchical structure of the homogeneously covered regions of the scenes.

3 Image Matching Module

The matching module between the scheme input consultation and the ser of homogeneously covered regions of the different scenes has been designed with three optional filters. The first analyses the shape similarity. In order to do this, both the morphology of the region occupied by the shape and the deformation of the external contour of

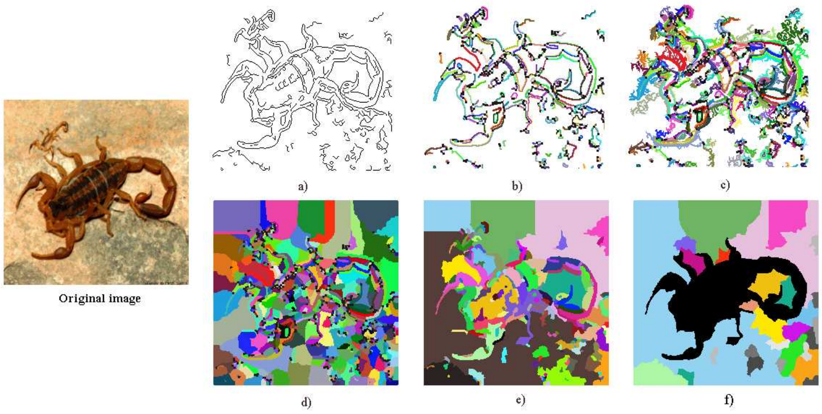

Figure 1: a) Edges map obtained after the application of the discontinuities detection algorithm proposed (mask of size 9x9, $\sigma = 1.0$, upper threshold = 0.9 and lower threshold = 0.5), b) Results of growth regions process after the first iteration, c) after the seventh iteration, d) after the tenth iteration, e) Final filler of not labeled areas, f) Final grouping of regions assisting to of minimum distance approach among the middle colors on considered regions.

the same are characterized. We have also defined a 2D shapes recognition model based on two sequential stages, the second of which analyses the color similarity. The optional character allows a greater flexibility in the type of consultation carried out, with the ability to treat a color, a shape or a combination of both.

3.1 Mathematical morphology

Initially, an invariant characterization of finite subsets, based on the mathematical morphology of image subset retrieval, that possesses a similar shape to that given in the query, has been used. This technique requires the previous retrieval of characteristics with an invariant property with respect to certain geometrical transformations of the image [3]. The two-dimensional representation of the ratios of area that exist between the region occupied by the shape and the background of the image, for the morphological operators of erosion, dilation, aperture and closure [18], on a structural element of increasing size, allows for the characterization of a shape by the region that it occupies [16] (Figure 2). This initial filtering of the database allows for a rapid structuring of images that may contain a morphologically similar object (size and overall shape), to that given in the query, leaving aside the final details of its external contour.

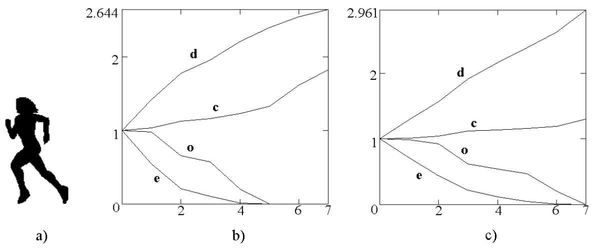

Figure 2: a) Original binary 2D shape. b) The simulation carried out for a circular structural element of increasing radius r ($r \in [1, 40]$) over the original image given a). c) Simulation carried out for a circular structural element of increasing radius r ($r \in [1, 32]$ over the original image a). Curves marked with **d** correspond to dilation, **c** to closing, **a** to opening and **e** to erosion.

3.2 Deformable model

With the subset retrieved images we use a deformable model to measure their similarity with the query. We have chosen a model of shape deformation, initially used by Pentland in Photobook [14], for more specific database. We will use an approach given by Nastar. To establish the deformation we would have to apply to an object in order to transform it into another one, and in this way to be able to measure the energy used as a measure of object similarity. This model gives us an analytical expression of the vibration modes of the nodes situated on a closed curve (we are interested in contours of closed forms). We will define a mechanism to represent the value of the modal range as a function of this mode, as follows: $\tilde{u}_i(t) = f(i)$

We call to this representation *modal spectrum* [13]. In summary, it describes which modes and in which quantity they are sufficiently excited in order to deform a shape enough to make it appear like another. In this way, we can obtain the following indicator of the amount of *deformation energy* required in the process:

$$E_{deformation} = \frac{1}{2}\mathbf{U}^T\mathbf{K}\mathbf{U} = \frac{1}{2}\sum_{i=1}^{N}\omega_i^2\tilde{u}_i^2 \tag{9}$$

where $\mathbf{K}$ is the stiffness matrix of the shape, $\mathbf{U}$ is the displacement matrix, ω_i is the frequency of vibration and $\tilde{u}$ is the range of the $i-$th mode in consideration.

In this context, rigid movements have zero deformation energy and two deformations are deemed to be similar when their corresponding fields of displacement $\mathbf{U}_1$ and $\mathbf{U}_2$, regardless of rigid transformations, are similar. Expressed in another way, *two deformations are similar when their corresponding deformation spectra are similar.*

Recently new approaches to the invariant characterization of flat shapes have appeared [10][12]. However, the calculation of said invariants can be very sensitive to the random noise present int he image or to small non-rigid deformation of the object. The approach used by Pentland allows us to identify the space of the comparison with the selection of adequate eigenvectors which introduces robustness to the method.

3.3 Color

The same scene viewed under two different illuminants induces two different color images. If the two illuminants are the same color but are placed at different positions then corresponding *rgb* pixels are related by simple scale factors. In contrast if the lighting geometry is held fixed but the color of the light changes then it is the individual color channels that are scaling apart.

In order to make the proposed retrieval scheme more general regarding the type of scenes shown in the database, we have used a technique called **comprehensive image normalization** which removes image dependency on lighting geometry and illumination color [4].

Similarity measure based on the normalization of the color between the object given in the query and the contained in the images is established from the histogram intersection of the said [19].

4 Experimental results

The experiments on the general scheme of image retrieval proposed, have been defined with a database composed of 160 images containing multiple objects of different shapes and colors. We varied the positions of the objects and the camera and the lighting conditions of the composed scenes.

The defined scheme supports three different methods of consulting with the designed user interface:

- *Color based query.* This query will extract or retrieve the images that contain objects whose color is the closest to the target object (Figure 3).
- *Shape based query.* Through this type of query, those images that most resemble the shape of the target object can be extracted (Figure 4).
- *Color and shape combination based query.* Here, the images that most resemble the color and shape of the target object can be retrieved (Figure 5).

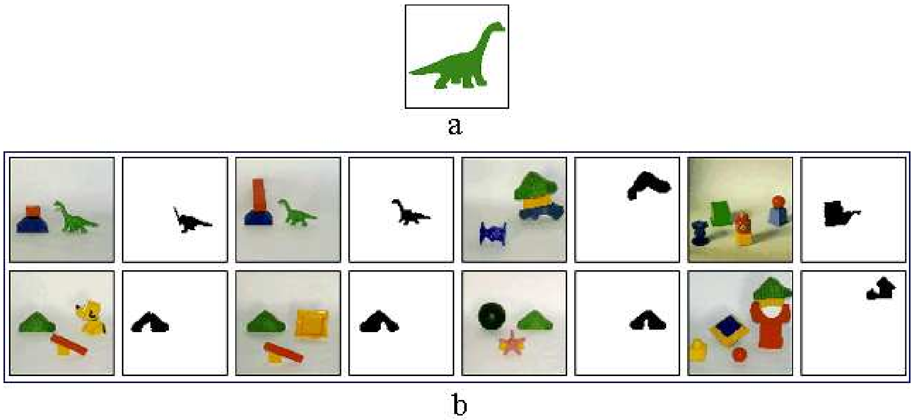

Figure 3: Example of colour based query

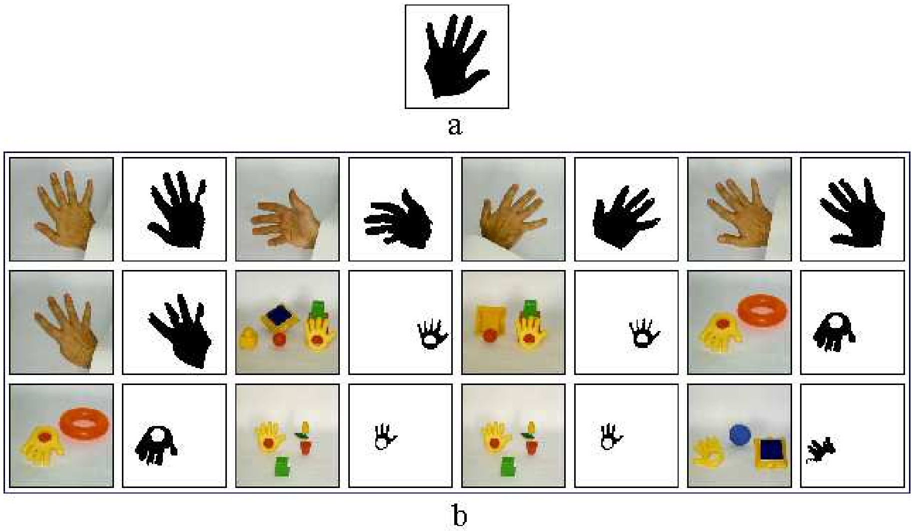

Figure 4: Example of shape based query

We establish 10 queries designed by different users: 7 simples, 1 complex and 2 with a non-linear variation of the rgb channels. In figure 6 we show the different query images used in the retrieval scheme.

For the color and shape combination based query, we compute the recall rates for measuring the accuracy of the retrieval scheme. The recall rate is defined as the percent of retrieved images which are similar to the query among the total number of images similar to the query in the database.

In table 1, we show the recall rate for the carried out queries. Rows show the queries and columns show the number of recovered images.

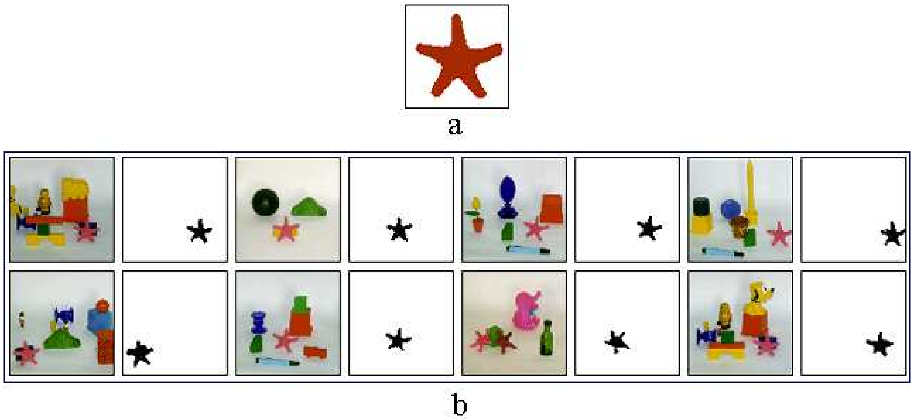

Figure 5: Example of colour and shape combination based query

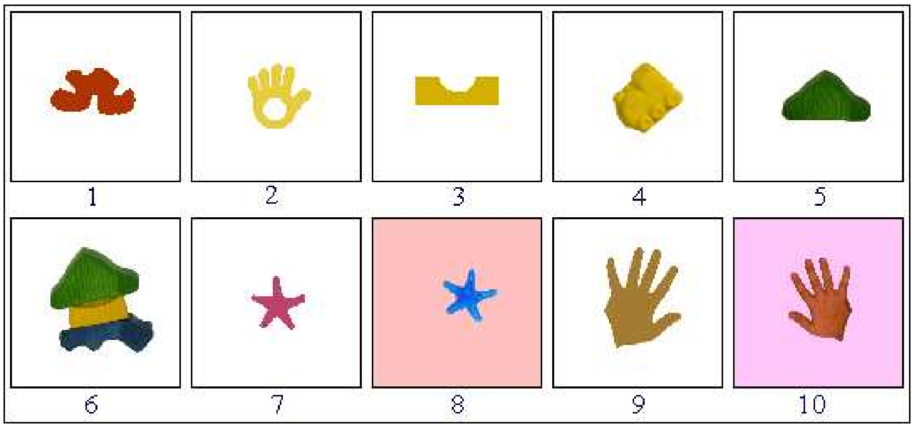

Figure 6: Query images

5 Conclusions

We have presented a complete scheme of image retrieval from a database through the use of queries about the *component* elements of the different scenes.

The defined scheme allows for the complete segmentation of the objects contained in the images from a given color space (in our case, an HSI color space) and a gradient based edge detector (with the use of Canny's algorithm).

The formulation of the query can be established on the basis of color, shape or color and shape combined as considered by the user. Independently of the kind of query chosen by the user, the obtained results are an image indexed relation to the measurement of the attributes considered in each case.

The proposed scheme has been proved on a database of 160 complex images. In our experiments, when we consider the color and shape combination based query, the recall rate is higher the 85% when we recover the 15 more similar images and reach the 93.8% when we consider the 20 more similar images.

Recall Rate	5	10	15	20
Query 1	66,7	100,0	100,0	100,0
Query 2	25,0	75,0	100,0	100,0
Query 3	31,3	56,3	87,5	93,8
Query 4	42,9	57,1	85,7	100,0
Query 5	35,7	71,4	92,9	100,0
Query 6	100,0	100,0	100,0	100,0
Query 7	31,3	62,5	93,8	100,0
Query 8	31,3	62,5	93,8	100,0
Query 9	100,0	100,0	100,0	100,0
Query 10	100,0	100,0	100,0	100,0

Table 1: Recall rates obtained for the different queries when we recover the 5, 10, 15 and 20 *most similar* images

References

[1] F. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[2] J. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8:679–698, 1986.

[3] E. R. Dougherty. *Mathematical morphology and image analysis.* Marcel Dekker, 1993.

[4] G. D. Finlayson, B. Schiele, and J. L. Crowley. Comprehensive colour image normalization. In H. Burkhardt and B. Neumann, editors, *Computer vision, ECCV'98*, number 1406 in Lecture notes in computer vision. Springer, 1998.

[5] J. M. Fuertes, M. Lucena, N. Perez de la Blanca, and J. Chamorro-Martinez. Color-based object recognition. *Pattern Recognition Letters. To apperar*, 2000.

[6] J. M. Fuertes, M. Lucena, N. Perez de la Blanca, and J. Fdez-Valdivia. New methods to segmenting color images. In *VII Symposium Nacional de Reconocimiento de Formas y Analisis de Imagenes*, pages 233–238, Barcelona, 1997.

[7] T. Gevers and A. W. N. Smeulders. Color-based object recognition. *Pattern Recognition*, 32:453–464, 1999.

[8] G. Healey. Segmenting images using normalize color. *IEEE Trans. Syst. Man. Cybern.*, 22:64–73, 1992.

[9] M. S. Kankanhalli, B. M. Mehtre, and H. Y. Huang. Color and spatial feature for content-based image retrieval. *Pattern Recognition Letters*, 20:109–118, 1999.

[10] M. Kliot and E. Rivlin. Invariant-based shape retrieval in pictorial database. In H. Burkhardt and B. Neumann, editors, *Computer vision, ECCV'98*, number 1406 in Lecture notes in computer science. Springer, 1998.

[11] B. M. Mehtre, A. D. Kankanhalli, A. D. Narasimhalu, and G. C. Man. Color matching for image retrieval. *Pattern Recognition Letters*, 16:325–331, 1995.

[12] J. L. Mundy and A. Zisserman. *Geometric invariance in computer vision.* MIT Press, 1992.

[13] C. Nastar and N. Ayache. Frequency-based nonrigid motion analysis: application to four dimensional medical images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(11):1067–1079, November 1996.

[14] A. Pentland, R. Picard, and S. Sclaroff. Photobook:tools for content-based manipulation of images databases. *International Journal of Computer Vision*, 18(3), 1996.

[15] A. Pentland and S. Sclaroff. Closed-form solutions for physically-based shape modeling and recognition. *IEEE Trans. on Patterns Analysis and Machine Intelligence*, 13(7):715–729, July 1991.

[16] B. Ripley. *Statistical inference for spatial processes.* Cambridge University Press, 1988.

[17] R. Schettini. A segmentation algorithm for color images. *Pattern Recognition Letters*, 14:499–506, 1993.

[18] J. Serra. *Image analysis and mathematical morphology.* Academic Press, 1988.

[19] M. J. Swain. Color indexing. *International Journal of Computer Vision*, 7:11–32, 1991.

[20] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: Recovering 3d shape and nonrigid motion. *Artificial Intelligence*, 36:91–123, 1988.