

Model Based Segmentation of Radiological Images of the Cranium

M. G. Cawley and K. Natarajan

School of Computer Science
The University of Birmingham
Edgbaston, Birmingham. B15 2TT

This work is part of the Alvey project MMI/134 "Model based processing of radiological images".

Different imaging modalities, for example X-Ray CT and NMR, generate markedly different images from the same anatomy. Investigating the nature of the images produced has allowed us to develop a low level domain independent segmentation algorithm.

A scheme to perform further merging has been developed, based on rules and a current world model derived from the anatomy, the modality and non-visual knowledge⁷.

This final stage produces a segmented world representation of the image which is directly comparable to the original current world model.

The X-Ray CT and NMR imaging techniques provide a non-invasive window for the examination of *in vivo* human morphology and physiology, and generate a two dimensional representation of tissue behaviour under that imaging modality. These two-dimensional views are not directly comparable to the knowledge of morphology and physiology which exists in medicine. The interpretation of radiological images, therefore, requires new skills. This is particularly evident when non-standard scans are taken or when pathology has severely distorted the normal anatomy. Radiological training, therefore, takes many years and relies on experience gained through reporting on many scans. These scans or grey level images are the only visual means of expressing the relative measurement of the physical attributes involved in the imaging technique. Automatic segmentation of X-Ray CT and NMR images is therefore reliant on the correct use of domain constraints to produce plausible groupings of these grey levels. By combining radiological and physical knowledge in an appropriate scheme a robust segmentation can be achieved.

The domain constraints derived from the application of such knowledge provide a requirement for the initial segmentation, that is in the instantiation of a *current world* model which can provide the cues dependent on the anatomical feature to be found. These cues are an expectation of how difficult the feature will be to find and therefore how finely the image should be segmented.

At the lowest level of the segmentation this allows the process to be data-driven and prevents inappropriate tessellations from being imposed on the image⁹[10].

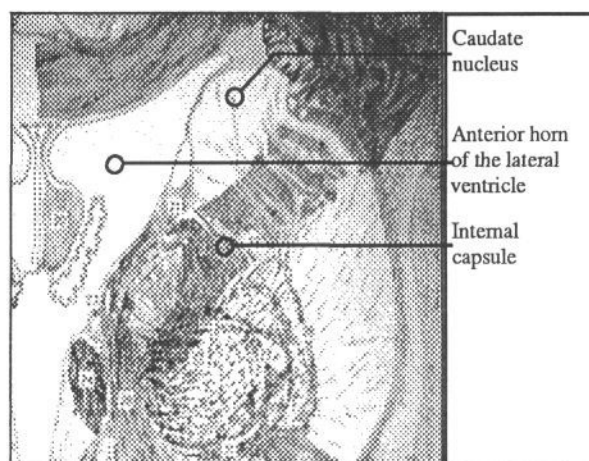


Figure 1:
Section through the interventricular foramen and the hypothalamus.

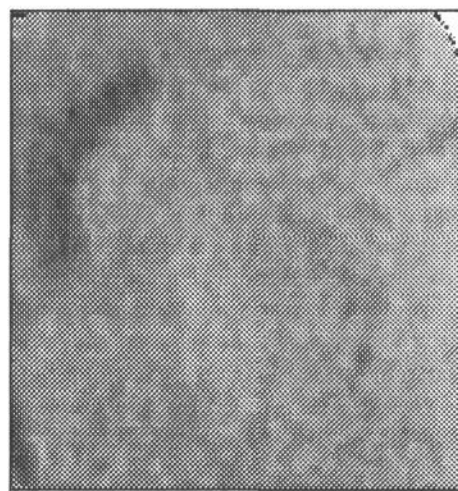


Figure 2:
A 64x64 pixels X-ray CT scan showing the gross structures in figure 1.

Figure 1¹ shows some of the neuroanatomical structures, their positions with respect to landmarks, mutual

¹This diagram is taken from [5].

spatial relationships and relative sizes. The figure 2² shows some of these structures but under the CT imaging modality. The CT scan clearly shows how the imaging modality can drastically change the appearance, and even visibility, of anatomical structures.

Overview

The components of the model based radiological interpretation system described in this paper are shown in figure 3.

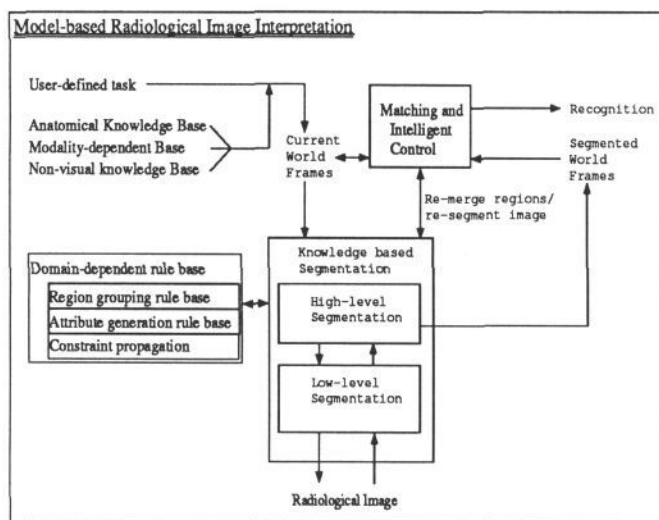


Figure 3:
Schematic view showing the components of the system

The static anatomical model is transformed using modality dependent knowledge into a dynamic *Current World* model^[7] which can be changed or updated depending on the results of the segmentation. Initially the current world describes how the anatomy in a specified slice of a particular patient is expected to appear under one of the imaging modalities; X-Ray CT or NMR. The current world comprises a number of frames³ one for each discernible anatomical feature. References to other features in the static anatomical model that are not expected to appear in the selected slice are removed from the current world. Within each frame there is a set of feature vectors describing that anatomical structure and its relationship to other structures in the current world model by means of a *vocabulary*. It is through this vocabulary that constraints are propagated to the segmentation phases. The main attributes used so far in the segmentation task are: contrast, relative position, tissue type and absolute position with respect to a grid. At the lowest level of the segmentation stage, these

constraints manifest themselves in two ways, firstly suggesting a suitable clustering criterion and secondly suggesting image transformation pre-processing (for example blurring). At the higher level of segmentation, the constraints are used as *expectation* to invoke the appropriate domain-dependent knowledge.

The segmentation stage covers the conversion of parts of the image into a sets of segmented world (SW) frames. Each of these frames is a possible interpretation of the image under the goal derived from the current world. The development of segmentation precepts has been accomplished by using combinations of operators (for example standard deviation and gradient maps).

To summarise, the task of the segmentation is to provide a plausible partitioning of the image given a set of constraints. These constraints are derived from the current world and are therefore tailored to the particular image and patient details. The current world can be subjected to alterations instigated by the matcher but in terms of the segmentation stage the current world acts only in the production of a goal.

Low level segmentation

The first stage in the segmentation of the image is the production of a symbolic representation which encompasses some attribute derived from the image. The representation is in terms of regions of the image, each region conforming to an attribute value.

The experience gained in dealing with radiological images builds up image processing knowledge or model of tissue behaviour within the image. This is distinct from physiology and from the imaging modality modelling. The tissue model is a combination of both the anatomical structure and the physical behaviour of materials when exposed to the imaging process. This type of knowledge allows us to experiment with standard segmentation algorithms. For the low-level segmentation we have chosen to segment the image into a number of homogeneous regions. This was done using a single criterion split and merge^[2] approach in the first instance^[6] and latterly an adaptive split and merge^[9] algorithm. The homogeneous split and merge algorithm exploits the assumption that for a single tissue the texture is uniform. The adaptive split and merge takes as input a "criterion map" which means that any property, for instance statistical measures, can be used to guide the low level segmentation. The importation of parameters into the split and merge algorithm allows for the passing of domain-dependent constraints via the use of rules. This does not control the segmentation but simply changes the tightness of the criterion and thus the data driven aspect is retained.

These rules represent the more fundamental sequences of operations, which under the control of parameters allow reasonable execution times for processing whilst retaining flexibility. The split and merge algorithm now uses a split criterion map rather than a single input value. Rules can still be used, however, to generate this

²This image is not a CT scan of figure 1, but rather, intends to show how the anatomical structures, such as the lateral ventricles, caudate nucleus, globus pallidus and internal capsule, appear under the CT imaging modality.

³In our implementation the dynamic current world are Minsky-like^[4] frames.

split criterion using pre-defined operators. The rules that are used are of a standard form and are invoked from a rule interpreter. The basis for this interpreter is a modified version of that developed at Edinburgh for the Simple Blackboard System^{[1],[3]}.

The output of the low level segmentation stage is a list of regions each of which is a list of blocks. The blocks remain from the initial quadtree mechanism used to split the image into parts. The low level merging stage uses the same criterion as the split stage and so the quadtree should not impose any systematic structure on the data. The blocks, which are represented in symbolic form, are the basic description of the image and are defined according to the split and merge criterion. Variation of this criterion does not change the symbolic vocabulary. So for instance the blocks are always considered to be smooth and yet by increasing the split and merge criterion the actual composition of the blocks can become much more variable.

The low-level segmentation stage, when employing current world constraints such as determining an appropriate focus of attention, would typically yield 900 regions comprising approximately 4000 blocks.

High level segmentation

The use of more abstract knowledge is required to merge together regions presented to the high level segmentation stage. Unlike the low level stage, this knowledge has to bridge the gap between the goal, as derived from the current world, and the symbolic representation derived - with constraints - from the image. The possibility also arises of merging regions using the same attributes as the initial segmentation.

The high level knowledge is split into three parts, each represented using the same form⁴. Firstly, the control knowledge transforms the rule requests into a form suitable for the more domain specific rules. For instance the region lists are, when applicable, converted into segmented world (SW) frames. The SW frames are in of same form as the current world frames but are not compatible with the original region lists. The control knowledge produces a common format from either the SW frames or the region lists. This also applies to the current world frames when necessary. This could be required if, for instance, analysis of the shape of both the SW and the CW were required to match them. This type of knowledge is primarily concerned with limiting the search space. In practice this is achieved by selecting the most appropriate attribute or conjunction of attributes (strategy knowledge), given expectation from the current world, in order to generate a *candidate region list*. For example, a simple rule would form a candidate region list using a threshold taken from the current world for the particular goal feature.

Secondly there are rules which reorganise the data and

⁴The knowledge representation scheme adopted by the integral task is described in [11]

facilitate the writing of merging or grouping rules. These take the formatted information, derived from the control rules and try to trigger on the premises which may have been directed through a selected strategy rule. These intermediate rules generally refer to some attribute, or combination of attributes, which can be derived from the image. At present we use attributes such as contrast and average gradients, but because the regional representation of the image is an accurate representation other attributes may be subsequently calculated. The first of these, the contrast between two regions, is calculated from the block descriptions but now yields a symbolic term such as low, or medium or high. This is shown in figure 4 below. Each of these terms does not have a fixed numerical value although the relationships between the terms are fixed, i.e high is always greater than medium, etc.

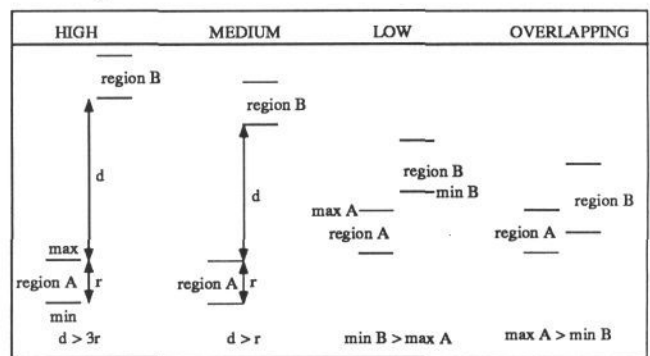


Figure 4: Contrast between neighbouring regions in four cases. Other possibilities are taken into account but are not shown here.

Thirdly attribute rules are the means by which the actual values are found from the data. They form a link between the rules and the predicates and allow the rules to call for such information without duplication of actions. The rules in the end are property based.

The result of the high-level segmentation is a set of SW frames containing regions that have been grouped together according to the expectations dictated by the current world. When this is not possible, regions are grouped together according to the actual regional data attributes. These SW frames, which are in the same format as the current world frames, are subsequently compared to each other by an intelligent matcher. If a sufficient level of match occurs then the recognition task is completed. Failing this, merging rules that had not previously been chosen or fired (possibly due to prioritisation) are activated, or alternatively a new low-level segmentation is requested.

Results

The following shows model based segmentation being applied to the area of interest depicted in figure 2.

The first verticle set of three images are for the non-adaptive split and merge.

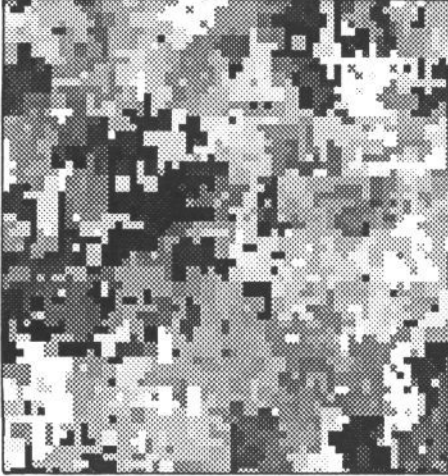


Figure 5:
Split criterion set to 2.5 (smallest σ for expected tissue types).

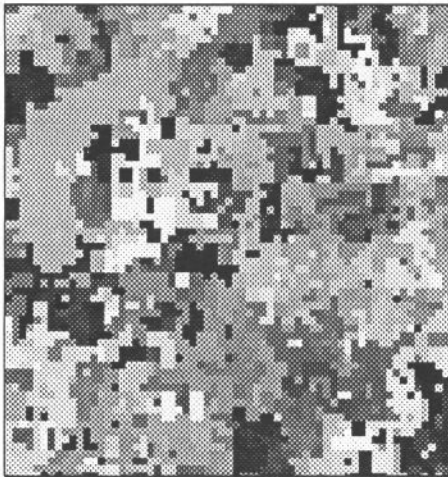


Figure 6:
Results of high level grouping according to spatial and contrast criteria (rule based) of figure 5.

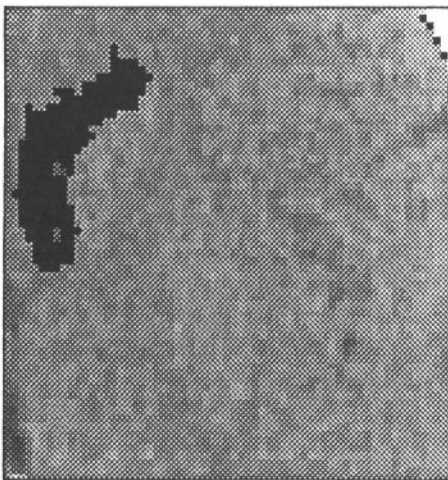


Figure 7:
Results of high-level grouping super-imposed the on original image.

The second showing the results when the adaptive method of split and merge is applied.

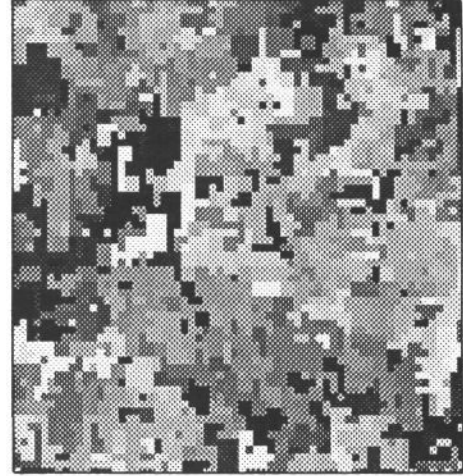


Figure 8:
Adaptive split and merge - using standard deviation map.

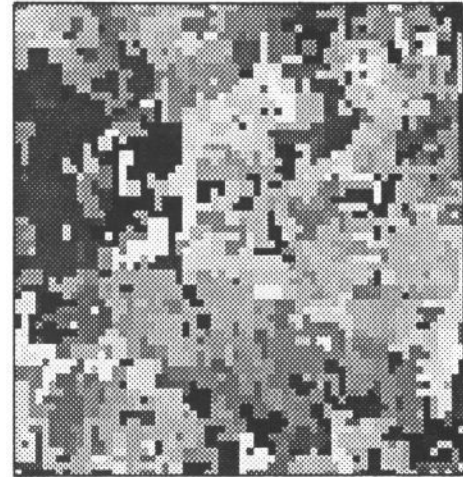


Figure 9:
Results of high level grouping according to spatial and contrast criteria (rule based) of figure 8.

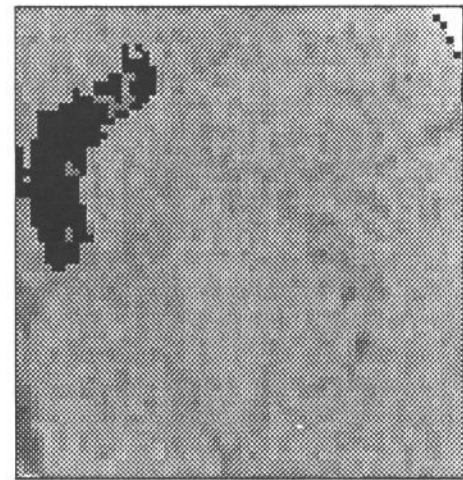


Figure 10:
Results of high-level grouping super-imposed on the original image.

In order to illustrate the segmentation we use as input Figure 2 and attempt to label the anterior horn of the lateral ventricle, (that is the dark curved structure in the top left of Figure 2). This scan is firstly subjected to non-adaptive and adaptive split and merge, figures 5 and 8 respectively.

These low-level segmentations are reasonably accurate representations of the original image⁵ with the data now representing, symbolically, the underlying homogeneity within the image. Figure 5 comprises 3200 blocks grouped together as 1040 regions, whereas figure 8 contains 3150 blocks and 985 regions. These low-level segmentations are now grouped together according to expectations dictated by the left anterior horn's current world frame. At present this takes the form of selecting a list of candidate regions whose intensity is within a range defined by the tissue type- cerebral spinal fluid under CT. This has the effect of limiting the search space to approximately 100 regions. Figure 6 and 9 are the results of applying spatial relationship and contrast constraints. Finally figures 7 and 10 show the high-level segmentation results super-imposed on the original image. We can see from these images that both have extracted the anterior horn and this knowledge can now be used for further recognition and refinement tasks. It is also worth noting that had we used the current world frame for the caudate nucleus (to the right of the anterior horn), this structure would also have been extracted.

Summary

The aim of the work, developing an automated image interpretation system, has been to create expert modules for each step in the interpretation task. Within the segmentation stage itself there are two separable tasks, low-level and high-level segmentation. Low-level segmentation is predominantly data driven employing constraints primarily derived from the imaging modality. High level segmentation propagates these constraints for low level segmentation and also sets up goals derived from knowledge about neuroanatomy and patient details. Further to this the segmentation stage is driven by an intelligent controller whose task is to define a strategy using procedures such as neurological landmarking.

The result to date shows that the use of simple attribute based rules, when invoked through a knowledge based system providing goal and data driven segmentation, applied with domain-dependent knowledge, can go a long way in the segmentation of a fairly complex but closed world scenes. In the future the addition of more attributes to the base rules will be investigated. The flexibility of the system gives easy access to new rules whilst catering for their application through the use of the strategy rules.

References

[1] Baldock R. and Towers S., First Steps Towards A

⁵When compared through the use of discrepancy graphs.

Blackboard Controlled System for Matching Image and Model in the Presence of Noise and Distortion, Proceedings of the 4th International Conference Pattern Recognition ,pages 429-438, Cambridge U.K, March 1988.

[2] Pavlidis T., Structural Pattern Recognition, Springer-Verlag, 1977.

[3] Clark P., "SCS - A Simple Control System for the Control of Software 'Experts' Performing a Collaborative Task", Alvey MMI/134 Internal Report MOBPRIM/TI/REP2/880505.May 1988.

[4] Minsky M., "A Framework for Representing Knowledge",in The Psychology of Computer Vision, P.Winston (Ed.), McGraw-Hill, New York 1975.

[5] Nieuwenhuys R., Voogd J., van Huijzen C., The Human Central Nervous System: A synopsis and atlas, Springer-Verlag, 1988.

[6] Sokolowska, E., Newell, J.A., "Multi-layered image representation: Structure and application in recognition of parts of brain anatomy", Pattern Recognition Letters, Vol.4, 223-230, 1986.

[7] Cawley, M.G. and Natarajan, K., "Development of a model for use in medical image interpretation", Proceedings of the fifth Alvey Vision Conference, AVC89, 305-308, 1989.

[8] Natarajan, K., Cawley, M.G. and Newell, J.A., "A Knowledge-based System Paradigm for Automatic Interpretation of CT Scans", Accepted for the Biological Engineering Society meeting "Computer Aided Interpretation of Medical Images", London, June 1990.

[9] Cawley, M.G., Natarajan, K. and Newell, J.A., "Data representation for subsequent image interpretation" Accepted for the Biological Engineering Society meeting "Computer Aided Interpretation of Medical Images", London, June 1990.

[10] Cawley, M.G., Natarajan, K. and Newell, J.A., "Region based segmentation of CT scans", Alvey MMI/134 Internal Report, April 1990.

[11] Natarajan, K., Cawley, M.G. and Newell, J.A., "An Integrated Frame and Rule based system for Radiological Image Segmentation", Alvey MMI/134 Internal Report, to be released.

