TEXTURAL 3-DIMENSIONAL MULTISCALE ANALYSIS OF MRI VOLUMES OF THE BRAIN

Harry Hatzakis, Stephen Roberts and Ioannis Matalas
Department of Electrical and Electronic Engineering
Imperial College of Science, Technology and Medicine
London SW7 2BT, U.K.
email: h.hatzakis@ic.ac.uk

Abstract

We describe a method for textural feature extraction of MRI volumes of the brain and, based upon those features, a method for classification and assessment of the anatomical malformations of the brain, due to Alzheimer's Disease (AD). In our research, we make the hypothesis that there is enough detectable textural evidence from a 3D analysis of MR images of the brain to detect and identify the earliest structural changes of AD.

To uniquely characterise structural malformations we construct a database of statistical information for 3D textures at different scales, using wavelet operators. The major goal at this stage of our research is to explore the inherent constraints imposed by the structure of the texture and its symbolic description. Our representation benefits from a unique method of parameter reduction, which gives an unambiguous description of the textures of the brain in 3D. One of the key attributes of this model is that, in the case of conflicting statements, it generates a low confidence estimate, thus allowing a local measure of reliability.

1 Introduction

Although a common sense definition of texture is easily given, it is not easy to construct computationally efficient algorithms whose performance is comparable and compatible with the human visual system. So far, computer vision has failed to unravel the general problem of perception, but it seems reasonable to hope that a simple problem of segmentation would be understood relatively well. Such aims have never been achieved however. A statistician may argue furthermore that perception is further complicated by various types of uncertainty. The examination of 3D MR images of the brain strongly suggests that segmentation, as performed by humans, is not simply a pixel-by-pixel classification, but a procedure which includes a wider range of pixels, making the need for a local operator essential. But this leads to the question: How large should this operator be? The answer is related to the nature of the texture. If the texture in question exhibits small scale characteristics, small sized operators should be used. On the other hand, for macrotextures relatively large windows may be preferable.

The selection of the window size for the filtering function introduces yet another problem: the element of uncertainty in our texture classification, with a trade-off between spatial size and accuracy of textural property estimates in a noisy environment. Such considerations are reminiscent of the discussion of invariants in Marr's approach to the problem of vision [1].

We thus wish to extract ideal properties of the symbolic description of a texture, which are inherently constrained by the structure of the texture itself upon which the transformations operate.

The major goal of our research at this stage is to explore the nature and extent of these constraints and to construct a database of statistical information for 3D textures, which will uniquely characterise and classify MR images of the brain. The method we propose here performs 3D textural analysis using optimised multi-scale wavelet filters. We have applied our method to MRI volumes of the brain, in an attempt to classify asymptomatic and pre-symptomatic Alzheimer's pathologies. Our initial results have proved very encouraging but further investigation is needed to further develop the procedure and to validate it with more data.

2 The transform

The basis of our approach is the fact that the local gradient directions and their consistency in a 3-D image contain important information for textural analysis [2]. Local orientation is important, and for most neighbourhoods, a well defined local feature of multi-dimensional data with respect to the scale of analysis is needed. A set of operators covering four different dyadic scales is implemented in an attempt to extract information from different scales and also to overcome uncertainty limitations. Our idea is to use the outputs at several different scales of multi-directional filters to produce a reliable gradient estimation. Statistical analysis of the result follows.

2.1 The 2-D transform.

To improve multi-directional homogeneity at each scale, we propose wavelet templates oriented at $0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$. Note, however, that for "optimal" homogeneity in N dimensions we need at least $2^{(N-1)}+1$ templates. For two dimensions therefore, we need at least three templates with directions $0, \frac{\pi}{3}, \frac{2\pi}{3}$. This minimal implementation, although computational efficient, is very difficult to implement, especially for small operator sizes. We construct our templates by convolving a Gaussian, $G(x)$, and the first derivative of a Gaussian, $dG(x)$. Thus for each scale we have 4 templates:

$$\begin{aligned} t_1(x,y) &= G(x) * dG(y) \\ t_2(x,y) &= G(\tfrac{x+y}{2}) * dG(\tfrac{x-y}{2}) \\ t_3(x,y) &= G(y) * dG(x) \\ t_4(x,y) &= G(\tfrac{x-y}{2}) * dG(\tfrac{x+y}{2}) \end{aligned} \qquad (1)$$

as shown in Figure 1.

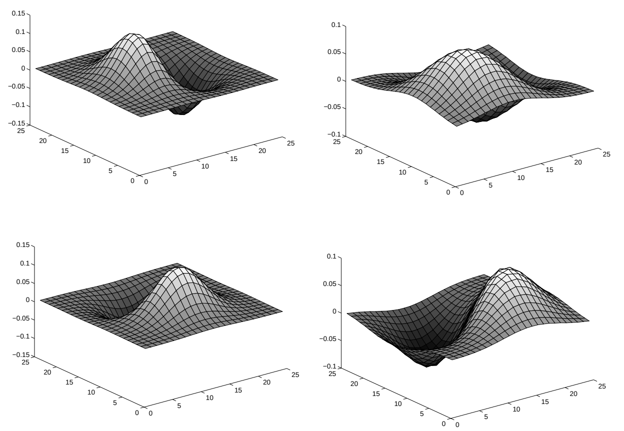

Figure 1: The four templates as given in eq.(1).

We use the standard approach and apply the wavelet transform at dyadic scales as a scale-space hierarchy, i.e. wavelet transforms at scales $\alpha = 2^j$, where $j \in Z$. For our MR brain images (180×180 pixels) scales greater than 2^4 appear to hold no significant information for an analytic descriptions. We use four scales : 3×3, 7×7, 13×13 and 25×25.

Transformation of the image results in a feature space containing several components. A representation of all the convolution outputs would necessarily result in a 16-dimensional feature vector q of the state of a neighbourhood. But, there is a correlation between components. This can be shown to be the case if we have a sufficiently simple model with local dominance. In this case within this neighbourhood the content can be represented by the dominant component only. In consequence, we state that two such features q_A and q_B can be sufficiently well represented by only one variable, as:

$$z_A = q_A - q_B \qquad (2)$$

This follows a similar approach described in [2].

For a vertical edge this model produces a positive value and for an horizontal edge it produces a negative value. For each scale, if q_1, q_2, q_3 and q_4 are the computed features for the orientations $0, \frac{\pi}{4}, \frac{\pi}{2}$ and $\frac{3\pi}{4}$ respectively, we can formulate the output at that scale as a complex number z defined as:

$$z = z_1 + iz_2 = (q_1 - q_3) + i(q_2 - q_4) \qquad (3)$$

thus

$$M = \sqrt{(q_1 - q_3)^2 + (q_2 - q_4)^2} \qquad (4)$$

and

$$A = \tan^{-1}\left(\frac{q_1 - q_3}{q_2 - q_4}\right) \qquad (5)$$

where M and A are the magnitude and argument respectively.

The information about the orientation of z is extremely important to retain, as it turns out to be more robust to noise and consequently more useful than the magnitude of z. In Figure 2 we demonstrate the above representations applied to an MRI slice (size 180×180) transformed at the second scale (7×7).

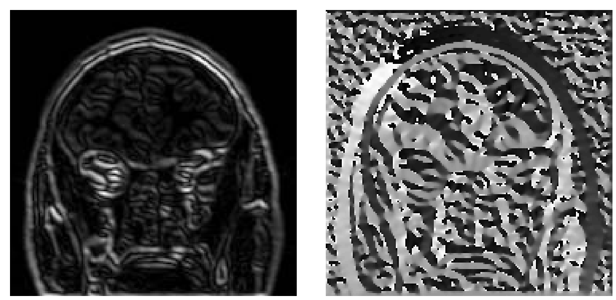

Figure 2: The transformation of an MRI slice.

2.2 The 3-D transform.

We have applied the feature estimator described above to three sets of 2-dimensional slices, one for each plane; the coronal xy-plane, the axial xz-plane and the sagittal yz-plane. Thus for each individual voxel we calculate a gradient in 12 directions.

For each scale of the transform and each voxel, we have a 6D vector with three magnitudes $m_i, i = 1,2,3,$ and three angle estimators $\alpha_i, i = 1,2,3$. To estimate a single 3D vector we compute the projections of each individual vector onto x,y and z axes:

$$\begin{aligned} X &= m_1 \cos(\alpha_1) + m_2 \sin(\alpha_2) \\ Y &= m_1 \cos(\alpha_1) + m_3 \sin(\alpha_3) \\ Z &= m_2 \cos(\alpha_2) + m_3 \cos(\alpha_3) \end{aligned} \qquad (6)$$

We then combine these three to estimate the magnitude M, the elevation angle E and the rotation angle R:

$$\begin{aligned} M &= \sqrt{X^2 + Y^2 + Z^2} \\ E &= \tan^{-1}\left(\frac{Y}{\sqrt{Z^2+X^2}}\right) \\ R &= \tan^{-1}\left(\frac{Z}{X}\right) \end{aligned} \qquad (7)$$

The natural domain for the representation of digital images is the spatial domain. In accordance with traditions in digital signal processing, a volume will be considered to be a three dimensional set of real numbers $u(x, y, z) \in \Re$. In addition to this scalar space, which is the input for the transformation described above, we have produced for each image a 3D "vector image". The result can be denoted as $\underline{u}(x, y, z) \in \Re^k$, where $(\underline{.})$ indicates a vector rather than a scalar in the 12D real space ($k = 12$).

2.3 Statistical treatment.

To describe and quantify texture, one needs to examine the interactions between structure organisations and not the grey levels of voxel neighbourhoods. Thus, it is natural to seek a robust statistical basis to describe our results. One obvious way to improve performance is to treat the problem as a decision problem, in which the statistics of the total population of voxels are considered to be a mixture of those of the component regions into which the texture of the image is to be quantified.

By examining texture by its statistical appearance, we also overcome the problem of isotropy as the central moments of a given structure remain unchanged under orientation changes.

For each image data set we obtain, we compute for each voxel the variance and the skewness of the magnitude M, the elevation angle A, and the rotation angle R, in a 3D ellipsoidal neibourhood of the size comparable to the scale in question.

2.4 Abnormality detection

Based upon the statistical information, we estimate a rich representation of the 'normal' data, against which we may assess a degree of 'abnormality' of any unseen data item. From our scalar image we have extracted a set of features which we hope will be able to discriminate between normal and abnormal anatomy of brain structures with reference to AD. We make two important assumptions [6]:

- An appropriate feature is one that the examples of abnormality appear in the tails of the distributions obtained for all normal data. (i.e. they occur in low density regions of 'normal' data).
- In the absence of any other information, the abnormalities are considered to be uniformly distributed outside the boundaries of normality.

We begin by representing the normal data set and thus determining these boundaries. In a statistical sense, the best description is the unconditional probability density function (pdf) $p(x)$ where x is the 12D feature vector of the normal data. If, a test vector x belongs to a region of input space where $p(x)$ is below a statistically determined threshold, then the vector is said to be 'novel' or abnormal. This however assumes that the *pdf* of the novel data is constant over some large region of input space. To obtain the probability density estimates we will use a semi-parametric Gaussian mixture model mainly because it has the flexibility to model general distributions and has convenient analytic properties. The free parameters of the Gaussian are estimated from a maximum likehood approach.

3 Discussion

3.1 Uncertainty and reliability.

Maybe the two most important aspects of this work concern uncertainty and scale treatment. Uncertainty is a direct consequence of the need for our system to respond in an objective and specific way to certain type of transformation of the image itself (such as spatial shifts or noise) when forced to make a decision. We have defined our goal as the classification of MR volumes based upon regional textural properties of the image. We follow the assumption that both 2D and 3D textures have regions of constant properties, separated by well defined boundaries, which is far from being realistic. Apart from any degradations occurring during the process of image acquisition (such as motion, noise etc.), surface texture generally affects the grey level of the image. Thus, our initial assumption is not always valid and some statistical treatment of the problem must be utilised. Moreover, natural textures generally contain an element of randomness or variability which affects any fixed-size local texture measure used. The use of various scales for textural analysis with different size operators, seems, once more, inevitable.

Furthermore, when a space-scale transformation is performed over a region, one must consider the variations in the neighbourhood which are significantly smaller than the filter kernel and which are perceived as fluctuations rather than as members of separate classes. This strongly suggests that the use of a local smoothing operator, beforehand, may reduce noise fluctuations. Such a smoothing procedure can decrease noise within a region without, hopefully, introducing too much interference between regions. But there is always a degradation of the spatial information, which is due to the removal of the higher spatial frequency components needed to give a precise estimate of the position. Thus, the price we pay for precision in regional properties is the degradation of boundary localisation. To explore this idea further, there is an inverse relationship between the separation of regions and the loss of precision in boundary position which is very similar to the uncertainty principle:

$$\Delta x \Delta g \geq k \tag{8}$$

where k is some positive number and Δx and Δg are the conflicted measures of uncertainty.

It is inevitable that there is uncertainty in the decision of *"what is it"* and of *"where is it"* processing model. Of course, we only repeat Marr's ideas [1], when he simply defined segmentation (and to some extent classification), as the knowledge of "what is where" by observing an image. But Marr had a mammalian visual system in mind when he made such aphorisms. And while the goal of imitating the human vision is desirable, in some applications, the goal of an automatic classifier is not simply to mimic human performance, but, if possible, to exceed it.

Our model tries to reduce the uncertainty of *what is it* and partly of *where is it* by proclaiming zero or low confidence in the case of conflicting statements. This is because two feature measures that have the same magnitude and are incompatible produce a resulting output which is zero or very close to zero. Thus, incompatible pairs of measurements should always act to cancel out each other. This, of course, implies an undecided observation of the state within a neighbourhood.

3.2 Scale-space approach.

In our work, we use the spatial consistency of the outputs from several different operators at different scales to produce a more reliable estimate of the gradient vector that can be obtained by a single filtering operation. It is well known that textural attributes carrying valuable information may emerge at any range of scales in an image. We must expect our operators of being capable to make image textural features explicit over a wide range of sizes as it is obvious that different levels of organisation do not correspond simply to what would be seen through a band-pass filter. Although several types of organisation can be detected using a simple filter-based approach, many cannot. We make, therefore, the important assumption that the textural organisation of a 3D MR image is generated by a number of different processes, each operating at a different scale. Consequently, a vector representation of an image must be capable of capturing changes in attribute values applied to structures that span over a wide range of sizes of the image. In other words, the primitives of our representation must:

- work at a number of different scales, and,
- be scale invariant by using operators of constant shape but different sizes.

3.3 The search for isotropy.

When we examine a certain structure with respect to its 3-D textural appearance, it is essential that we can recognise the structure, irrespective of all the variations of its appearance. This a familiar situation when we think of the accidental, random variations that affect any repetition of a constant shape. For example we can never draw exactly the same triangle twice on a blackboard, yet we still keep talking about the "triangle" because at any certain level of abstraction we know that the attributes of the triangle are satisfied. It is always less trivial and more interesting to consider the systematic variations as a fundamental powerful procedure. This idea becomes more precise when we examine the description of a 3D texture, as different kinds of properties can be defined. By indicating the transformations of the space in which the textels (texture elements) are embedded, we derive some properties which must be left unchanged. This element is very essential for our research: to preserve the invariant nature of our transform to the topological properties of the textures of the brain. But does this desire apply to all of the properties? Do we require a 3D region with homogeneous texture to be classified the same, even after enlargement, rotation or shift? And of what use may the answers to these questions be, when we want to classify an MR image of a normal control subject or of an AD patient? These are some of the questions that remain to be answered. One can hope, by preserving isotropy, to escape the uncertainty inherent in the understanding or the description of the complex structure of nerve fibres, as we do not yet know which geometric properties of the transformation of brain structure are worth identifying, which are due to chance, and which are irrelevant because they are too general. If there is a possibility to track the texture of a certain structure through some natural variation within the brain, one must assume that the aspects of the texture remain invariant throughout all the transformations. A triangle is always a triangle and grey matter structure is always grey matter, regardless how they are oriented or shifted. Our method preserves a good level of isotropy in the 3D space, a characteristic which we believe is essential for a classification system without having too large a computational overhead.

References

[1] D.Marr, "*Vision*," W.H. Freeman and Co., 1982.

[2] G.H.Grandlund, H.Knuttson, "*Signal Processing for Computer Vision*," Knluwer Academic Publishers,1995.

[3] R. M. Haralick and L. G. Shapiro, "*Computer and Robot Vision*," vol. 1, Addison-Wesley, 1992.

[4] T.N.Tan, A.G.Constantinides, "Texture analysis based on a human visual model," *ICASSP-90, New Mexico*, 1986.

[5] R.Wilson, M. Spann, "*Segmentation and Uncertainty*," Research Studies Press Ltd., 1988.

[6] C.M.Bishop, "*Neural Networks for Pattern Recognition*," Oxford Univ. Press, 1995.