Detecting Changes in Nonisotropic Images

K.J. Worsley,^{1,*} M. Andermann,¹ T. Koulis,¹ D. MacDonald,² and A.C. Evans²

¹Department of Mathematics and Statistics, McGill University, Montreal, Québec, Canada

²Montreal Neurological Institute, McGill University, Montreal, Québec, Canada

Abstract: If the noise component of image data is nonisotropic, i.e., if it has nonconstant smoothness or effective point spread function, then theoretical results for the P value of local maxima and the size of suprathreshold clusters of a statistical parametric map (SPM) based on random field theory are not valid. This assumption is reasonable for PET or smoothed fMRI data, but not if these data are projected onto an unfolded, inflated, or flattened 2D cortical surface. Anatomical data such as structure masks, surface displacements, and deformation vectors are also highly nonisotropic. The solution offered here is to suppose that the image can be warped or flattened (in a statistical sense) into a space where the data are isotropic. The subsequent corrected P values do not depend on finding this warping; it is sufficient only to know that such a warping exists. *Hum. Brain Mapping* 8:98–101, 1999. © 1999 Wiley-Liss, Inc.

Key words: SPM; random fields; PET; fMRI

INTRODUCTION

The theoretical results for P values of local maxima and size of suprathreshold clusters of a statistical parametric map (SPM) are not valid if the noise component of the image data is nonisotropic [Worsley et al., 1996]. One of the conditions for isotropy, or “flatness” (in the statistical sense), is that the FWHM should be constant in all directions and across all voxels in the image. This assumption is reasonable for PET data or smoothed fMRI data, but not for two new types of image data. The first is PET or fMRI data projected onto an unfolded, inflated, or flattened 2D cortical surface [Drury et al., 1998; Fischl et al., 1999], where the different amounts of stretching of the surface alter the original constant FWHM, making it nonisotropic. The second is anatomical data such as 3D

binary masks of a structure [Zijdenbos et al., 1998], 2D surface displacements [MacDonald et al., 1998], and 3D vector deformations required to warp the structure to an atlas standard [Collins et al., 1998; Thompson et al., 1999]. In all these cases, the smoothness of the images varies considerably from region to region, so they too are not isotropic. The purpose of this report is to present a simple method for overcoming these problems so that random field theory can be applied to most nonisotropic images.

METHODS

The first step is to transform the data to a triangular (2D) or tetrahedral (3D) lattice. Pixels on a square lattice can be easily transformed by subdividing the squares (of 4 adjacent pixels) into two triangles. Voxels on a cubical lattice can be transformed by subdividing each cube of eight adjacent voxels into five tetrahedra as shown in Figure 1. Note that no new vertices (voxels) are created; only their connectivity is altered.

*Correspondence to: K.J. Worsley, Department of Mathematics and Statistics, McGill University, 805 Sherbrooke St. West, Montreal, Québec, Canada H3A 2K6. E-mail: worsley@math.mcgill.ca.

Received for publication 12 April 1999; accepted 10 August 1999

Figure 1.

Voxels (balls) on a cubical lattice can be subdivided into five tetrahedra per cube in an alternating checkerboard arrangement.

Some data, particularly on cortical surfaces, are already triangulated, so this step is unnecessary.

The second step is to estimate the *effective* FWHM (*eFWHM*) along each edge of the lattice, defined as the FWHM of a Gaussian kernel that would produce the same local smoothness of the noise component of the observed images. This is based on the normalized residuals from fitting a linear model at each voxel. Label the two voxels at either end of an edge by 1 and 2. We now fit a linear model to data from n images at each voxel by least squares. For image i , let r_{i1} denote the residual (observed—fitted) at voxel 1, and let r_{i2} denote the residual at voxel 2, $i = 1, \dots, n$. The normalized residuals at the ends are

$$u_{ij} = \frac{r_{ij}}{\sqrt{\sum_{i=1}^n r_{ij}^2}}, \quad i = 1, \dots, n, \quad j = 1, 2.$$

We first estimate the roughness of the noise, defined as the standard deviation of the derivative of the noise divided by the standard deviation of the noise itself. First, let

$$\Delta u = \sqrt{\sum_{i=1}^n (u_{i1} - u_{i2})^2}.$$

Then, an unbiased estimator of the roughness is

$$\lambda = \Delta u / \Delta x$$

where Δx is the length of the edge [Worsley, 1999]. The effective FWHM along the edge is then

$$eFWHM = \sqrt{4 \log_e 2} / \lambda.$$

If the image is isotropic, or “flat,” then the *eFWHM* should be constant; departures from this indicate nonisotropy.

A first attempt at correction is to warp the coordinates of the voxels so that the *eFWHM* is approximately constant, or equivalently, $\lambda \approx 1$. If the new edge length is $\Delta \tilde{x}$, then this implies that $\Delta \tilde{x} \approx \Delta u$. This is equivalent to a local multidimensional scaling that makes the new edge length $\Delta \tilde{x}$ proportional to the old edge length Δx divided by the *eFWHM*, which means that edges with low *eFWHM* are stretched and those with high *eFWHM* are shrunk (relative to the average). This was achieved by minimising

$$S = \sum_{\text{edges}} (\Delta \tilde{x}^2 - \Delta u^2)^2$$

for each point separately, holding all others fixed, then iterating till convergence. By approximating $\Delta \tilde{x}$ as a linear function of the warp (ignoring the quadratic term), there is a simple matrix expression for the optimal warp at each iteration.

The usual random field theory for P values of local maxima and cluster sizes can then be applied to the resulting isotropic, or “flattened,” images with constant roughness equal to 1, or $FWHM = \sqrt{4 \log_e 2}$.

For large search regions, the P value of local maxima of a SPM T in D dimensions is then well approximated by

$$P(\max T \geq t) \approx \text{Resels } \rho_D(t),$$

where $\rho_D(t)$ is the D -dimensional EC density function from Worsley et al. [1996], and

$$\text{Resels} = V(4 \log_e 2)^{-D/2},$$

where V is the area ($D = 2$) or volume ($D = 3$) of the search region in the flattened image. The P value of cluster sizes above a threshold also can be evaluated by simply measuring cluster size in the flattened space and applying the usual formulas [Friston et al., 1995; Cao, 1999]. Note that an additional correction is required to account for the randomness of the cluster *Resels* themselves. (This will be the subject of a later report.)

However, the flattening step is not entirely necessary. Inspection of the preceding calculations shows that *Resels* can be derived directly from the normalised residuals without actually carrying out the flattening. To see this, let $u_0, u_1, \dots, u_D$ be the n -vectors of the normalised residuals at each vertex of a triangular ($D = 2$) or tetrahedral ($D = 3$) component of the search region. Define the $n \times D$ matrix of differences by

$$\Delta u = (u_1 - u_0, \dots, u_D - u_0).$$

Then, the *resels* of a region or cluster becomes

$$\text{Resels} = \frac{1}{D!} \sum_{\text{components}} |\Delta u' \Delta u|^{1/2} (4 \log_e 2)^{-D/2}.$$

It can be shown that $|\Delta u' \Delta u|$ does not depend on which vertex is labelled as 0, and that *Resels* is unbiased, with no adjustment for degrees of freedom [Worsley, 1999]. Note also that *Resels* does not depend on the actual Euclidean coordinates of the vertices (voxels), only the information about how the voxels are connected to form components.

How does this compare with current methods for a square or cubical lattice of voxels? At present, most software for the statistical analysis of SPMs uses the following:

$$\text{Resels} = \frac{N}{D!} \left| \sum_{\text{components}} \frac{\Delta u' \Delta u}{N} \right|^{1/2} (4 \log_e 2)^{-D/2}.$$

where N is the number of components that must all be oriented and labelled in the same way. The discrepancy

is then due to summing before or after taking the square root of the determinant, but this discrepancy is slight in practice.

However, for cluster size statistics there can be a very large discrepancy. The reason is simple: by chance alone, large size clusters will occur in regions where the images are very smooth, and small size clusters will occur in regions where the image is very rough. The distribution of cluster sizes will, therefore, be considerably biased toward more extreme cluster sizes, resulting in more false positive clusters in smooth regions. Moreover, true positive clusters in rough regions could be overlooked because their sizes are not large enough to exceed the critical size for the whole region. The proposed method will compensate by replacing cluster size by resel size, which is invariant to differences in smoothness.

RESULTS

As a test, the method was applied to detecting local shape differences between the cortex of normal males ($n = 83$) and females ($n = 68$) using smoothed 3D binary masks [Zijdenbos et al., 1998], 2D surface normal displacements [MacDonald et al., 1998] and 3D deformation vectors [Collins et al., 1998]. In all cases the $eFWHM$ varied considerably from 5 mm to 30 mm depending on location, so the images were highly nonisotropic. Despite this, the P values for local maxima were not greatly affected, but the P values for cluster sizes were very sensitive to nonisotropy. Details are to be reported elsewhere.

CONCLUSIONS

We have derived a theoretical method for calculating P values for local maxima and cluster sizes of nonisotropic image data inside a large search region. The only statistical requirement is that *there exists* a sufficiently high dimensional space in which the image data can be warped to flatness. Knowing the dimension of this space, or the actual warping into it, is not required. Thus the method can be applied simply and efficiently to a very wide range of nonisotropic image data.

For small search regions, the entire flattened image is still necessary to find the remaining boundary correction terms for the unified P value of local maxima, which is accurate for search regions of almost any shape or size [Worsley et al., 1996]. Once again, the flattening can be avoided by the following trick. We note that the unified P value formula does not depend on the dimension of the space used to embed the

warped image; a higher dimensional space could be used to achieve a more successful flattening. Taking this to the limit, it can be shown that *exact* flatness can be achieved by warping the data into a space whose dimensionality equals the number of images: the coordinates are just the normalized residuals. Although this cannot be visualized, the resulting boundary corrections to the P value can be easily calculated from the resels of the component tetrahedra, triangles, and edges alone. This will be pursued in a subsequent report [Worsley, 1999].

REFERENCES

- Cao J. 1999: The size of the connected components of excursion sets of χ^2 , t and F fields. *Advances Applied Probability* (in press).
- Collins DL, Zijdenbos AP, Evans AC. 1998: Improved automatic gross cerebral structure segmentation. *NeuroImage* 7:S707.
- Drury HA, Corbetta M, Shulman G, Van Essen DC. 1998: *NeuroImage*, 7:S728.
- Fischl B, Sereno MI, Dale AM. 1999: Cortical surface-based analysis: II. Inflation, flat-tening, and a surface-based coordinate system. *NeuroImage* 7:S740.
- Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC. 1994: Assessing the significance of focal activations using their spatial extent. *Hum Brain Mapp* 1:214–220.
- MacDonald D, Avis D, Evans AC. 1998: Automatic segmentation of cortical surfaces from MRI with partial-volume correction. *NeuroImage* 7:S703.
- Thompson PM, Woods RP, Mega MS, Toga AW. 1999: Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. *Hum Brain Mapp*.
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. 1996: A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp* 4:58–73.
- Worsley KJ. 1999: Exceedence probabilities of local maxima of non-isotropic random fields, with an application to shape analysis via surface displacements (in prep.).
- Zijdenbos AP, Giedd JN, Blumenthal JD, Paus T, Rapoport JL, Evans AC. 1998: Automatic quantitative analysis of 3D brain data sets: application to a pediatric population. *NeuroImage* 7:S727.