Pattern recognition approach to the detection of single-trial event-related functional magnetic resonance images

D. Burke[1,4] **K. Murphy**[2] **H. Garavan**[2,3] **R. Reilly**[4]

[1]Rehabilitation Engineering, National Rehabilitation Hospital, Dún Laoghaire, Ireland
[2]Department of Psychology and Institute of Neuroscience, Trinity College, Dublin, Ireland
[3]Department of Psychiatry & Behavioral Medicine, Medical College of Wisconsin, Milwaukee, USA
[4]Department of Electronic & Electrical Engineering, National University of Ireland, Dublin, Ireland

Abstract—*Functional magnetic resonance imaging (FMRI) is an imaging technique for determining which regions of the brain are activated in response to a stimulus or event. Early FMRI experiment paradigms were based upon those used in positron emission tomography (PET), i.e. employing a block design consisting of extended periods of 'on' against 'off' activations. More recent experiments were based on event-related FMRI, harnessing the fact that very short stimuli trains or single events can generate robust responses. FMRI data suffer from low signal-to-noise ratios, and typical event-related experiment paradigms employ selective averaging over many trials before using statistical methods for determining active brain regions. The paper reports a pattern recognition approach to the detection of single-trial FMRI responses without recourse to averaging and at modest field strengths (1.5 T). Linear discriminant analysis (LDA) was applied in conjunction with different feature extraction techniques. Use of the unprocessed data samples as features resulted in single-trial events being classified with an accuracy of 61.0 ± 9.5% over five subjects. To improve classification accuracy, knowledge of the ideal template haemodynamic response was used in the feature extraction stage. A novel application of parametric modelling yielded an accuracy of 69.8 ± 6.3%, and a matched filtering approach yielded an accuracy of 71.9 ± 5.4%. Single-trial detection of event-related FMRI may yield new ways of examining the brain by facilitating new adaptive experiment designs and enabling tight integration with other single-trial electrophysiological methods.*

Keywords—*Functional magnetic resonance imaging, Event-related, Single-trial, Parametric modelling, Linear discriminant analysis, Pattern recognition*

Med. Biol. Eng. Comput., 2004, **42**, 604–609

1 Introduction

FUNCTIONAL MAGNETIC resonance imaging (FMRI) is a relatively new imaging technique based on the acquisition of magnetic susceptibility–weighted images of the brain taken in quick succession (KWONG *et al.*, 1992).

In contrast to other methods, such as the electro-encephalogram (EEG), FMRI does not measure electrical activity of the brain but rather changes in local blood flow, volume and oxygenation. These *haemodynamic* changes typically occur within a few seconds of neural activity and result in a tomographic imaging technique with millimetre spatial accuracy and a temporal resolution of seconds (COHEN and BOOKHEIMER, 1994).

Correspondence should be addressed to Dr David Burke; email: david.burke@ucd.ie

Paper received 5 March 2004 and in final form 27 May 2004

MBEC online number: 20043930

Detecting haemodynamic responses in the background data is analogous to evoked or event-related potential extraction in the EEG. At 1.5 T, the haemodynamic changes that occur in the FMRI signal during activation typically constitute only about 2–5% of the total amplitude (COHEN and BOOKHEIMER, 1994), thus making it difficult to detect the change. In addition, periodic motion, vibration, scanner instabilities and timing errors in echo-planar imaging sequences will cause FMRI images to have 'ghosts': faint duplicates of portions of an image displaced to incorrect locations. Placing an object, such as a research participant, inside the magnet's bore will further influence the field owing to the material's magnetic susceptibility.

Gross motion of the subject being scanned also affects the resultant signal, with small motions affecting the signal greatly. Motion outside the scanned region, e.g. swallowing or speaking, will affect the signal owing to magnetic susceptibility altering the magnetic field within the brain. Other physiological motions, e.g. breathing, can have similar effects.

Early FMRI experiment paradigms were based upon those used in PET, i.e. employing a block design consisting of

extended periods of 'on' against 'off' activations (ROSEN *et al.*, 1998). Whereas radio-pharmaceutical-based imaging such as PET requires extended periods of activation to achieve a quasi-equilibrium physiological state for about 1 min, it has been shown that stimuli as short as 34 ms can elicit FMRI responses (ROSEN *et al.*, 1998), which has led to the event-related FMRI experiment paradigm. Event-related FMRI allows investigators to study new tasks not possible with the conventional block design, e.g. discrete, unpredictable responses such as response inhibition (GARAVAN *et al.*, 1999), or the effect of novelty, as in the oddball paradigm (MCCARTHY *et al.*, 1997).

The most widely used technique for FMRI activation analysis is the correlation method. A reference time series that has the shape of the expected FMRI response is generated by convolving an ideal haemodynamic response with a vector indicating the position of relevant events (i.e. 1 for an event, 0 for a non-event, for each time point). The correlation of each voxel time series to this reference time series is computed. Voxels with a large correlation coefficient are deemed to be active.

Another approach is to calculate an impulse response function (the average shape, over the whole time series, following a particular type of event) for each voxel using a multiple regression technique. Measurements based on this impulse response function (i.e. peak height, the integral) yield an activation measure for that voxel. The approach used by GARAVAN *et al.* (1999) fitted a gamma-variate model to the impulse response function and then calculated the model's area under the curve, which was expressed as a percentage of that voxel's baseline activity.

In this paper, we report on a pattern recognition approach to the detection of single-trial FMRI responses without averaging at modest field strengths (1.5 T). Linear discriminant analysis (LDA) is employed in the pattern recognition task for this paper, in conjunction with three strategies for feature extraction. The first approach simply involves using the unprocessed data samples as features. To improve accuracy, information regarding the underlying haemodynamic response is included by way of matched filtering and a novel application of an autoregressive with exogenous input (ARX) model for feature extraction.

The ARX approach is inspired by an application for filtering event-related potentials in the EEG by CERUTTI *et al.* (1988). The template haemodynamic response constitutes the exogenous input of the model, and the noise contributions in the FMRI signal are approximated by white noise filtered by an autoregressive (AR) model. For the matched filtering approach, an optimum filter is created using a template haemodynamic response.

The focus of this paper is on the feature extraction stage, with the choice of LDA being motivated by its relative simplicity and wide application in the pattern recognition literature (BISHOP, 1995).

2 Methods

2.1 *FMRI principles*

MRI is based on the physics of nuclear magnetic resonance (NMR). Rather than pepper the text with references in what follows, we suggest the interested reader consult ROSEN *et al.* (1998), COHEN and BOOKHEIMER (1994) and KWONG *et al.* (1992) for an insightful introduction to the field.

Protons by themselves, as in the hydrogen nucleus, are slightly magnetic. Under the influence of a large magnetic field $\vec{B}$, there is a tendency for the protons to align themselves with the field. As a result, water (H_2O) in a magnetic field becomes slightly magnetised, with the relaxed magnetisation vector $\vec{M}$ aligned with the externally applied field. This alignment can be disturbed by the application of microwave radiofrequency (RF) radiation. When the applied RF radiation is turned off, the magnetisation relaxes back to alignment with the external field by precessing (rotating) around $\vec{B}$, during which time the proton re-emits RF radiation (see Fig. 1).

The frequency of the RF radiation emitted depends on the strength of the external magnetic field. By varying the strength of the magnetic field in space and by detecting the emitted signal over a range of radiofrequencies, the strength of the RF signal originating from different locations can be reconstructed. The result is an image that is sensitive to the number of H_2O molecules at each location in space. By varying the properties of the applied external field and the RF radiation, different types of tissue can be highlighted in the image.

In the majority of FMRI experiments, the signal change of interest is due to the haemodynamic oxygenation phenomenon and is referred to as the blood oxygen level dependent (BOLD) effect. The BOLD effect is a result of over-supply of oxyhaemoglobin as a consequence of neural activation. Haemodynamic changes are considerably slower than electrophysiological measures: for example, changes related to oxygenation take 4–5 s to reach a peak and another 5 s to return to baseline and are often followed by an undershoot for a subsequent 10 s (BOYNTON *et al.*, 1996). The FMRI BOLD response can be conveniently modelled by the γ-variate model (COHEN, 1997; GARAVAN *et al.*, 1999)

$$y = kt^r e^{-t/b} \tag{1}$$

and can be considered as low pass-filtered neural activity. Fig. 2 illustrates a modelled response with realistic parameter values (GARAVAN *et al.*, 1999), as given in the Figure caption.

2.2 *Feature extraction*

A central goal of feature extraction in pattern recognition problems is to process the data in such a way as to improve the overall classification accuracy. To establish a 'benchmark' accuracy, we initially considered a feature set of the unprocessed data samples (with the exception of removing the mean). To improve accuracy, we applied two techniques, each incorpor-

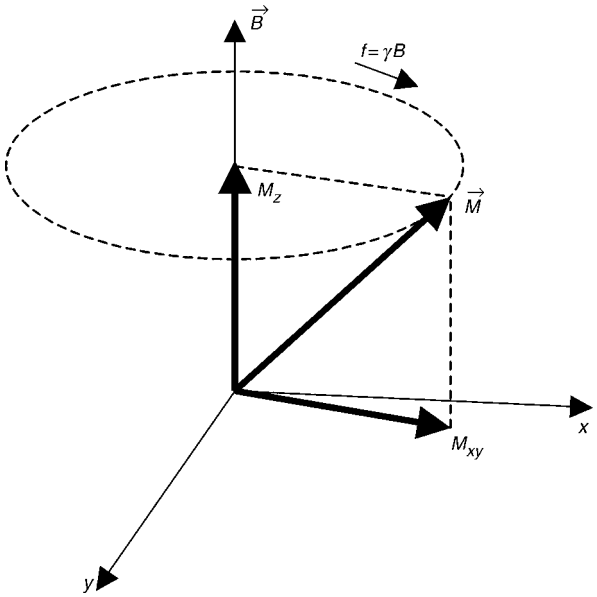

Fig. 1 *When disturbed from the relaxed state, $\vec{M}$ precesses about $\vec{B}$ in clockwise direction at Larmor frequency $f = \gamma B$, while emitting RF radiation. Precession rate is much larger than rate at which $\vec{M}$ relaxes back to $\vec{B}$ direction*

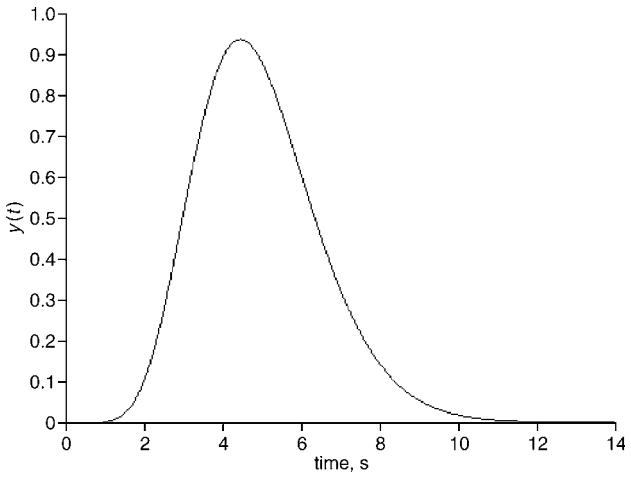

Fig. 2 *γ-variate function model of typical haemodynamic response $y = kt^r e^{-t/b}$. Parameter values: $k = 0.01$, $r = 9$, $b = 0.5$*

ating information related to the underlying haemodynamic response in different ways. To simplify notation, we assume a sampling interval of unity in what follows.

2.2.1 *ARX-based feature extraction:* For this approach, we attempted to model both the signal and the noise. The haemodynamic response was modelled as an autoregressive moving average (ARMA)-filtered ideal response to account for trial-by-trial variability. The noise contribution was modelled as an AR-filtered white noise source. Employing a common denominator resulted in the ARX model structure illustrated in Fig. 3.

The equation for the ARX model is

$$y(t) = -a_1 y(t-1) - \ldots - a_{n_a} y(t-n_a) + b_1 s(t-k) + \ldots + b_{n_b} s(t-k-n_b+1) + e(t) \quad (2)$$

where $s(t)$ is the exogenous input, a_i and b_i are the coefficients of the n_a, n_b order model, k is the delay (which we took as 0), and $e(t)$ is white noise. Denoting θ as the vector of model parameters, the *forward prediction* of the ARX model is obtained by dropping the zero mean noise term in (2)

$$\hat{y}(t) = -a_1 y(t-1) - \ldots - a_{n_a} y(t-n_a) + b_1 s(t-k) + \ldots + b_{n_b} s(t-k-n_b+1) \quad (3)$$

This results in an expression where the predicted output is expressed in terms of the old values of $y(t)$ and $s(t)$. The forward prediction error is given by

$$\epsilon(t, \theta) = y(t) - \hat{y}(t, \theta) \quad (4)$$

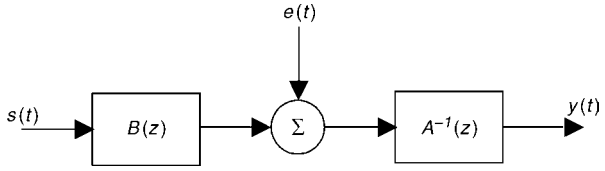

Fig. 3 *ARX model structure. s(t) is template haemodynamic response, e(t) is white noise, and y(t) is output of model, which is fitted to real data to obtain parameter values for A(z) and B(z)*

To fit the model to the data $y(t)$ collected over a period $t = 1, \ldots, N$, we chose $\hat{\theta}$ such that it minimises

$$E(\theta) = \frac{1}{N} \sum_{t=1}^{N} \epsilon^2(t, \theta) \quad (5)$$

An expression for calculating the coefficients of the ARX model directly from the data can be obtained by restating (3) as a linear regression and using (4) and (5) to solve explicitly (LJUNG, 1987). In addition, inspired by approaches in speech recognition to improve classification results (DELLER *et al.*, 2000), we also included an energy feature given by

$$E_y = \sum_{t=0}^{t=T} |s(t)|^2 \quad (6)$$

2.2.2 *Matched filtering:* To apply a match-filtered approach to processing FMRI data, we assumed that the underlying haemodynamic response was deterministic but corrupted by additive random noise. We can write

$$y(t) = s(t) + e(t) \quad (7)$$

where $s(t)$ is the ideal haemodynamic response, and $e(t)$ is noise. Applying a filter with impulse response $h(t)$ results in (* denotes convolution)

$$z(t) = y(t) * h(t) = z_s(t) + z_e(t) \quad (8)$$

where

$$z_s(t) = s(t) * h(t)$$
$$z_e(t) = e(t) * h(t)$$

The signal-to-noise ratio at the output of the filter at any time τ is given by

$$(SNR)^2 = \frac{|z_s(\tau)|^2}{E(|z_e(\tau)|^2)} \quad (9)$$

The goal of the matched-filter technique is to find the optimum filter $h_{opt}(t)$ such that (9) is maximised. It can be shown (BAHER, 2001) that (9) is maximised for a filter with a transfer function given by

$$H_{opt}(j\omega) = c \frac{S^*(j\omega)}{P_{ee}(j\omega)} \exp(-j\omega\tau) \quad (10)$$

where c is a constant, and $P_{ee}(j\omega)$ is the power spectrum of the noise $e(t)$. Assuming contamination of white noise for simplicity, $P_{ee}(j\omega) = A$, we can write (10) as

$$H_{opt}(j\omega) = \frac{c}{A} S^*(j\omega) \exp(-j\omega\tau)$$
$$h_{opt}(t) = \frac{c}{A} s(\tau - t) \quad (11)$$

i.e. $h_{opt}(t)$ is a scaled, time-reversed, shifted version of the ideal haemodynamic response $s(t)$. We used the data samples filtered by $h_{opt}(t)$ as features for classification. As with the ARX approach, we included an energy feature (6).

2.3 *Classification*

Given the set of features for a single trial, the classification task at hand was to determine whether a particular voxel exhibited a response or a non-response. Consider a set of c classes C_k. Given an input vector x of features, the classification problem can be formulated in terms of a set of discriminant functions $y_1(x)$, $y_2(x)$, $\ldots$, $y_c(x)$, where an input vector x is assigned to class C_k if

$$y_k(x) > y_j(x) \quad \text{for all} \quad j \neq k \quad (12)$$

Table 1 Results (% mean classification accuracy)

Subject	Session 1			Session 2			Session 3			Session 4		
	U	A	M	U	A	M	U	A	M	U	A	M
1	59.2	70.0	72.3	49.1	66.0	70.7	47.8	61.8	66.3	67.2	68.8	63.3
2	56.5	64.6	70.5	62.1	67.9	72.1	50.8	68.2	70.2	53.1	71.3	69.3
3	58.4	73.9	72.3	55.4	62.7	64.9	63.3	67.6	77.6	56.8	60.4	63.7
4	71.1	74.8	80.3	52.9	77.0	74.1	81.8	87.7	81.4	73.1	77.3	77.5
5	70.1	73.6	73.8	53.3	64.9	64.8	75.3	66.8	76.1	63.6	70.7	76.3

U-unprocessed; A-ARX; M-matched filter

i.e. choose the class for which the corresponding discriminant function is largest. We can derive a discriminant function in terms of Bayes' theorem given by

$$P(C_k|x) = \frac{p(x|C_k)P(C_k)}{p(x)} \tag{13}$$

Bayes' theorem relates the *posterior* probability to the product of the class conditional probability and *prior* probability. The $p(x)$ term serves as a normalisation term so that posterior probabilities sum to unity. The usefulness of Bayes' theorem (13) stems from the fact that it is easier to calculate the right-handside than the left hand side. One of the simplest discriminant functions is $y_k = P(C_k\,|\,x)$, which minimises the probability of misclassification (BISHOP, 1995). This discriminant function can be modified further by omitting the class independent probability from (13), taking the logarithm (valid for any monotonic function) and assuming a normal distribution to yield

$$y_k(x) = \mu_k^T \Sigma^{-1} x - \frac{1}{2}\mu_k^T \Sigma^{-1} \mu_k + \ln(P(C_k)) \tag{14}$$

where μ_k is the mean vector for class k, and Σ is the covariance matrix (where we have assumed equal covariance matrices across classes).

2.4 *Experiment protocol*

The experiment paradigm was that of GARAVAN *et al.* (1999) and employed the same data. Subjects were presented with a continuous serial stream of letters, with each letter appearing for a duration of 500 ms. A response was required for every occurrence of the alternating target letters X and Y, unless the alternation order was broken, in which case a response withholding was required. Stimuli were back-projected onto a screen at the subject's feet and were viewed with the aid of prism glasses attached to the inside of the radio-frequency head coil. Foam padding was used to limit head movements within the coil.

Contiguous 7 mm sagittal slices covering the entire brain were collected on a 1.5T scanner* using a blipped gradient-echo, echo-planar pulse sequence (TE = 40 ms, TR = 2000 ms; FOV = 24 cm; 64 × 64 matrix; 3.75 mm × 3.75 mm in-plane resolution). GARAVAN *et al.* (1999) identified regions (strongly lateralised to the right hemisphere) that were activated when subjects correctly withheld the motor response.

In our experiment, 12 voxels from each subject were identified as showing a strong response (6 voxels) and no response (6 voxels), respectively, by a goodness of fit to the ideal haemodynamic shape (1) being performed. A haemodynamic template was obtained by ensemble averaging over the events from the active voxels per session using a time window of 16 s. This template formed the exogenous signal to the ARX model and was also used to construct $h_{opt}(t)$ for the matched filter approach (11). To avoid problems with ill-conditioned matrices when calculating ARX model parameters (as a result of the coarse sampling period of 2 s), the FMRI data were resampled at five times the original rate using low-pass interpolation.

To obtain an estimate of accuracy, a cross-validation procedure was employed (BISHOP, 1995). The trials for each session were first randomly shuffled and subsequently divided into N distinct segments. $N-1$ segments were used to train an LDA classifier, and the remaining segment was used as the test set. This process was repeated for each N possible test sets, and the mean test set accuracy was computed. Finally, the complete process with a new random shuffle was repeated M times to yield an overall mean accuracy and standard deviation.

3 Results

Classification results were based on data obtained from five subjects over four sessions, each lasting 136 s. This translates to a total of 45 min of data for analysis. FMRI events were separated by an average of 20 s to reduce the possible effects of overlapping responses. A total of 15 segments ($N = 15$) and 20 shuffles ($M = 20$) were used for the cross-validation procedure. Table 1 illustrates the mean classification accuracy for each feature extraction technique (unprocessed, ARX and matched filter). ARX model orders were selected on the basis of the highest classification accuracy obtained: parameter values $n_a = 8$, $n_b = 4$ were applied throughout.

Fig. 4 illustrates the mean classification accuracy and standard deviation measured over the four sessions for each subject. In all

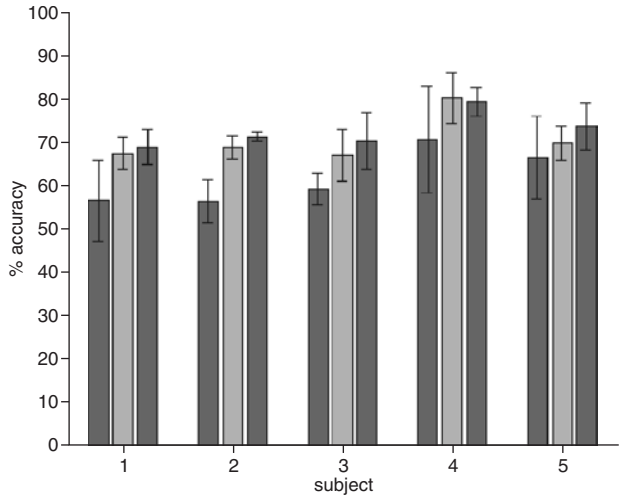

Fig. 4 *Classification accuracies averaged over 4 sessions for each subject. Bars are grouped according to feature extraction method (from left to right): unprocessed, ARX, matched filter. Standard deviation is illustrated by vertical lines*

*GE Signa

but one case (subject 4), the unprocessed data gave the lowest classification accuracy, with the matched filter approach yielding the highest accuracy. Averaging over all subjects yielded a classification performance of $61.0 \pm 9.5\%$, $69.8 \pm 6.3\%$ and $71.9 \pm 5.4\%$, for the unprocessed, ARX and matched filter approaches, respectively.

4 Discussion and conclusions

The LDA classifier yielded mediocre results on a feature set consisting of unprocessed data. By employing more sophisticated feature extraction strategies, improved mean classification accuracies were obtained, with the matched filter approach yielding the best results when calculated over the entire dataset. The matched filter and ARX approaches shared a similar concept, whereby information relating to the underlying haemodynamic shape was incorporated into the feature extraction algorithm itself. Both techniques yielded higher mean classification accuracies compared with the use of unprocessed data as features, which is an intuitively satisfying result.

Furthermore, the application of the matched filtering or ARX approach increased the robustness of the classification accuracy estimate, which was evident in the reduction of the standard deviation calculated over all subjects. Although the matched filter and ARX methods yielded consistent improvements over the unprocessed data method for each subject, the results were not significantly different (at the 95% confidence level) for the three methods.

At 1.5 T (the field strength most commonly employed in brain imaging experiments), the haemodynamic changes were typically very small in comparison with the background noise, which was not well correlated with the stimuli (WEISSKOFF *et al.*, 1993). Detection of these small haemodynamic signals is a key challenge in FMRI and is often achieved by selective averaging coupled with statistical tests. Single-trial detection of haemodynamic responses poses a more difficult challenge and is analogous to event-related or evoked potential detection in the EEG. Robust access to single-trial information in FMRI opens up new avenues for investigation, for example by

(i) facilitating new adaptive experiment designs
(ii) providing a real-time measure of mental performance
(iii) allowing for biofeedback and direct brain interfacing applications.

The temporal resolution of BOLD FMRI is fundamentally limited by the haemodynamic mechanisms that produce it. The EEG and magneto-encephalogram (MEG) are alternative, non-invasive techniques capable of resolving brain activity to short time scales. Combining the spatial advantages of FMRI with the temporal advantages of simultaneous EEG or MEG measurement may yield a new and intriguing source of data for studying the brain (see, for example, HUANG-HELLINGER *et al.* (1995) and DALE and SERENO (1993)). The ability to perform single-trial detection and estimation of event-related FMRI should prove useful to enable tight integration with established evoked and event-related potential techniques.

The ARX model performs well in improving classification results and is based on the assumption that the stimulus-elicited haemodynamic signal is linearly summated with the background noise fluctuations. As a first approximation, the background noise appears as white noise, with certain preferred peaks corresponding to breathing and heart beat dynamics (WEISSKOFF *et al.*, 1993). The AR structure (within the ARX structure) is an all-pole model and is well suited to modelling this type of data. The ARMA structure serves to model the trial-by-trial variability in the haemodynamic response.

Although our main concern in this paper was the detection of single-trial event-related responses, the ARX model can also be applied as a filter, as presented in the original publication applied to EEG event-related potentials by CERUTTI *et al.* (1988). Whereas averaging yields a generalised haemodynamic response, examination of individual haemodynamic responses by ARX filtering may provide access to unique information about aspects of an isolated event.

We have presented preliminary results for the detection of single-trial event-related FMRI. Modest classification performances have been obtained at average magnetic field strengths (1.5 T). As the signal-to-noise ratio scales with magnetic strength (COHEN and BOOKHEIMER, 1994), improved results may be possible at higher magnetic field strengths. However, the classification performance determined by our techniques may be ultimately limited by the reproducibility of activation patterns. Indeed TEGELER *et al.* (1999) reported that the reproducibility of FMRI activation patterns was no better at 4 T than at 1.5 T and suggested that physiology rather than physics limits the usable BOLD signal on a single-trial basis.

The focus of this paper has been on the feature extraction stage of the pattern recognition process. The LDA, although elegant in its simplicity and popular in the pattern recognition literature, is nonetheless limited to linearly separable problems. In the future, we plan to study single-trial analysis of FMRI data with more sophisticated classifiers, such as quadratic discriminant analysis and artificial neural networks.

References

BAHER, H. (2001): 'Analog and digital signal processing', 2nd edn (John Wiley and Sons, 2001)

BISHOP, C. M. (1995): 'Neural networks for pattern recognition' (Clarendon Press, Oxford, 1995)

BOYNTON, G. M., ENGEL, S. A., GLOVER, G. H., and HEEGER, D. J. (1996): 'Linear systems analysis of functional magnetic resonance imaging in human VI', *J. Neurosci.*, **16**, pp. 4207–4221

CERUTTI, S., CHIARENZA, G., LIBERATI, D., MASCELLANI, P., and PAVESI, G. (1988): 'A parametric method of identification of single-trial event-related potentials in the brain', *IEEE Trans. Biomed. Eng.*, **35**, pp. 701–711

COHEN, M. S. (1997): 'Parametric analysis of fMRI data using linear systems methods', *NeuroImage*, **6**, pp. 93–103

COHEN, M. S., and BOOKHEIMER, S. Y. (1994): 'Localization of brain function using magnetic resonance imaging', *Trends Neurosci.*, **17**, pp. 268–277

DALE, A. M. and SERENO, M. I. (1993): 'Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach', *J. Cognit. Neurosci.*, **5**, pp. 162–176

DELLER, J. R., HANSEN, J. H. L., and PROAKIS, J. G. (2000): 'Discrete-time processing of speech signals', (IEEE Press, 2000)

GARAVAN, H., ROSS, T. J., and STEIN, E. A. (1999): 'Right hemispheric dominance of inhibitory control: An event-related functional MRI study', *Proc. Natl. Acad. Sci.*, **86**, pp. 8301–8306

HUANG-HELLINGER, F. R., BREITER, H. C., MCCORMACK, G., COHEN, M. S., KWONG, K. K., SUTTON, J. P., SAVOY, R. L., WEISSKOFF, R. M., DAVIS, T. L., BAKER, J. R., BELLIVEAU, J. W., and ROSEN, B. R. (1995): 'Simultaneous functional magnetic resonance imaging and electrophysiological recording', *Human Brain Mapping*, **3**, pp. 13–25

JOSEPHS, O., TURNER, R., and FRISTON, K. (1997): 'Event-related fMRI', *Human Brain Mapping*, **5**, pp. 243–248

KWONG, K. K., BELLIVEAU, J. W., CHESLER, D. A., GOLDBERG, I. E., WEISSKOFF, R. M., PONCELET, B. P., KENNEDY, D. N., HOPPEL, B. E., COHEN, M. S., TURNER, R., *et al.* (1992): 'Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation', *Proc. Natl. Acad. Sci.*, **89**, pp. 5675–5679

LJUNG, L. (1987): 'System identification: theory for the user' (Prentice Hall, 1987)

MCCARTHY, G., LUBY, M., GORE, J., and GOLDMAN-RAKIC, P. (1997): 'Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI', *J. Neurophysiol.*, **77**, pp. 1630–1634

ROSEN, B. R., BUCKNER, R. L., and DALE, A. M. (1998): 'Event-related functional MRI: past, present, and future', *Proc. Natl. Acad. Sci.*, **95**, pp. 773–780

TEGELER, C., STROTHER, S. C., ANDERSON, J. R., and KIM, S.-G. (1999): 'Reproducibility of BOLD-based functional MRI obtained at 4 T', *Human Brain Mapping*, **7**, pp. 267–283

WEISSKOFF, R. M., BAKER, J. R., BELLIVEAU, J. W., DAVIS, T. K., KWONG, K. K., COHEN, M. S., and ROSEN, B. R. (1993): 'Power spectrum analysis of functionally-weighted MR data: What's in the noise?', *Soc. Magnet. Resonance Med.*, p. 7

Authors' biographies

DAVID BURKE received his BE, MEngSc and PhD from the National University of Ireland, Dublin, in 1997, 1998 and 2003, respectively. His research interests include dynamic systems modelling of the EEG and biomedical signal processing. Currently employed at Voxpilot Ltd, he works as an architect for next-generation internet systems and serves as editor and contributor to W3C and IETF specifications.

KEVIN MURPHY received a BA in theoretical physics and an MSc in high performance computing from Trinity College, Dublin, in 2000 and 2001, respectively. He is currently studying for a PhD in neuroscience in the Department of Psychology. His research topic is analytic methods in event-related FMRI, specifically questions relating to sample size, optimum experimental design and variability in haemodynamic shapes.

HUGH GARAVAN is a cognitive neuroscientist who received his PhD in 1995, from Bowling Green State University in Ohio. His research interests include the brain basis for human behavioural control, specifically looking at processes such as attention switching, inhibitory control, error detection and correction and how these functions might be affected in drug abusers. He is currently a lecturer in psychology at Trinity College, Dublin.

RICHARD REILLY received his BE, MEngSc and PhD in 1987, 1989 and 1992 respectively, from the National University of Ireland. He is currently senior lecturer in the Department of Electronic & Electrical Engineering at University College, Dublin and researches in biomedical signal processing, focusing on neuroscience. He was the 1999/2001 Silvanus P. Thompson International Lecturer for the IEE. In 2004 he was awarded a US Fulbright Award and a UCD President's Research Fellowship for sabbatical leave at the Nathan Kline Institute for Psychiatric Research, New York.