



# Sparse linear regression for reconstructing muscle activity from human cortical fMRI

G. Ganesh<sup>a,b,\*</sup>, E. Burdet<sup>b</sup>, M. Haruno<sup>a</sup>, M. Kawato<sup>a</sup>

<sup>a</sup> Department of Computational Neurobiology, ATR International, Computational Neuroscience Laboratories, 2-2-2 Hikaridai, “Keihanna Science City”, Seika-cho, Soraku-gun, Kyoto, 619-0288, Japan

<sup>b</sup> Department of Bioengineering, Imperial College of Science, Technology and Medicine, South Kensington Campus, London, SW7 2AZ, UK

## ARTICLE INFO

### Article history:

Received 28 February 2008

Revised 23 May 2008

Accepted 13 June 2008

Available online 25 June 2008

## ABSTRACT

In humans, it is generally not possible to use invasive techniques in order to identify brain activity corresponding to activity of individual muscles. Further, it is believed that the spatial resolution of non-invasive brain imaging modalities is not sufficient to isolate neural activity related to individual muscles. However, this study shows that it is possible to reconstruct muscle activity from functional magnetic resonance imaging (fMRI). We simultaneously recorded surface electromyography (EMG) from two antagonist muscles and motor cortices activity using fMRI, during an isometric task requiring both reciprocal activation and co-activation of the wrist muscles. Bayesian sparse regression was used to identify the parameters of a linear mapping from the fMRI activity in areas 4 (M1) and 6 (pre-motor, SMA) to EMG, and to reconstruct muscle activity in an independent test data set. The mapping obtained by the sparse regression algorithm showed significantly better generalization than those obtained from algorithms commonly used in decoding, i.e., support vector machine and least square regression. The two voxel sets corresponding to the activity of the antagonist muscles were intermingled but disjoint. They were distributed over a wide area of pre-motor cortex and M1 and not limited to regions generally associated with wrist control. These results show that brain activity measured by fMRI in humans can be used to predict individual muscle activity through Bayesian linear models, and that our algorithm provides a novel and non-invasive tool to investigate the brain mechanisms involved in motor control and learning in humans.

© 2008 Elsevier Inc. All rights reserved.

## Introduction

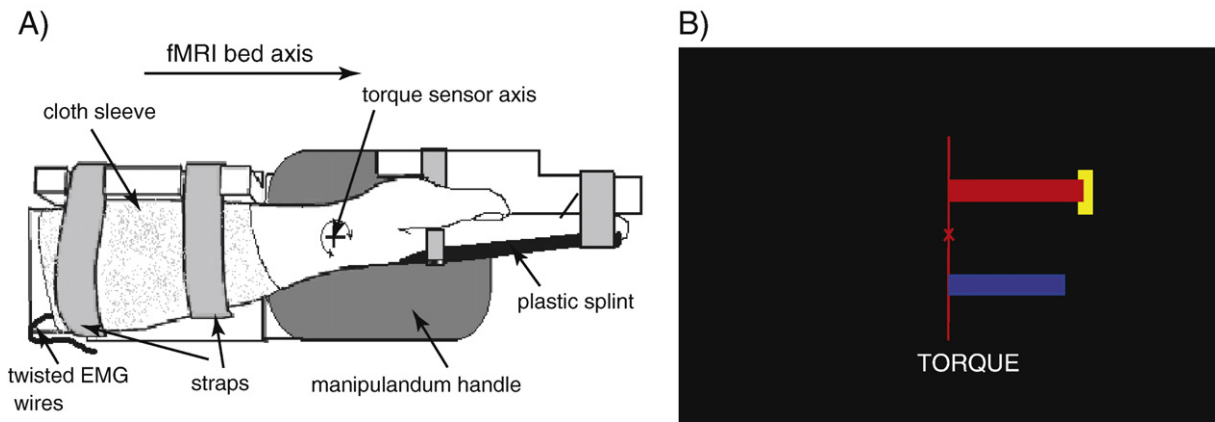
The Central Nervous System (CNS) controls the dynamic interaction with the environment using muscles, via their force and impedance properties (Hogan, 1984; Burdet et al., 2001). Observing the evolution of muscle activity is critical in understanding fundamental brain processes such as learning (Franklin et al., 2003) and motor synergies (d'Avella et al., 2006). Therefore, studying brain activity and regions related to individual muscles may provide important clues in understanding the motor system. Hence electrophysiological studies with monkeys investigating interactions between muscle activity, kinematics and brain activity have revealed important information about the organization of the motor system, such as the coordinate frames used by the brain (Kakei et al., 1997) and its functional topography (Cisek et al., 2003), and

has enabled decoding of brain activity corresponding to various task variables such as trajectories (Koike et al., 2006), relative positions of hand and target (Pesaran et al., 2006) and arm orientation (Scott et al., 1997a,b).

How can one infer the areas of the CNS related to motor control in humans, where electrophysiology is restricted to exceptional cases? Functional magnetic resonance imaging (fMRI) enables us to determine specific brain areas involved in the control of motor tasks in humans with good spatial accuracy. Human motor imaging experiments have studied the neural correlates of force control (Pope et al., 2005; Dai et al., 2001; Ehrsson et al., 2000) and various high level motor processes, such as internal model loading (Bursztyn et al., 2006), model switching (Imamizu et al., 2004) and differences in control strategies (Schaal et al., 2004; Diedrichsen et al., 2005; Milner et al., 2007). However, to our knowledge, no study using fMRI to seek for neural correlates of individual muscle level processes has been reported in the literature. This may stem from the belief that the fMRI resolution is not sufficient to isolate neural activity related to individual muscles. Also, only recently have methods been developed to monitor muscle activity during fMRI (Van Duinen et al., 2005; Ganesh et al., 2007).

\* Corresponding author. Department of Computational Neurobiology, ATR International, Computational Neuroscience Laboratories, 2-2-2 Hikaridai, “Keihanna Science City”, Seika-cho, Soraku-gun, Kyoto, 619-0288, Japan. Fax: +81 774 95 1236.

E-mail address: [gganesh@atr.jp](mailto:gganesh@atr.jp) (G. Ganesh).



**Fig. 1.** Experiment setup and visual display. (A) The subject hand was fixed in an isometric position onto an fMRI compatible manipulandum. Straps and a plastic splint helped restrict motion to a wrist rotation. (B) Visual feedback of the applied torque (red bar) and total EMG activity (blue bar) is provided. During the torque condition shown here, a target (yellow block) appears in front of the torque bar and the subject had to match it for the period it was displayed. During the co-contraction condition the target block appears in front of the co-contraction bar.

However, the results of this paper will demonstrate that it is possible to use fMRI to reconstruct activity corresponding to individual muscles. The paper first introduces a novel method to map fMRI to muscle activity. Data was obtained from surface electromyography activity (EMG) recorded from two muscles simultaneously to fMRI in conditions requiring various levels of reciprocal activation and co-activation. We adapt a Bayesian sparse regression method to map voxels from areas 4 (M1) and 6 (pre-motor, SMA) of the brain to the quantitative muscle activity (EMG) of the two muscles.

The prediction power of the learned mapping is then tested on a data set distinct from the training set, and compared with mapping identified using support vector machine regression (SVR) and ordinary least square regression (OLS), which are commonly used methods in decoding. The distribution of selected activity voxels is analyzed and compared to previous findings from electrophysiological studies in monkeys. The results demonstrate, for the first time, that the fMRI resolution is sufficient to reconstruct individual muscle EMG from brain activity in humans.

## Experiment

### Protocol

Six healthy male subjects aged between 23 and 40 years participated in the study. The institutional ethics committee approved the experiments and subjects gave informed consent prior to participation. An fMRI compatible interface (Gassert et al., 2006) was used to restrain the subject's wrist to an isometric posture with the help of straps and a plastic splint (Fig. 1A) while the subject contracted. This device has a custom optical torque sensor which was used to collect wrist joint torque during the experiment.

The task consisted of isometric wrist contractions in both flexion and extension directions performed during scanning. Subjects were presented with visual feedback of the torque applied to the manipulandum and their total muscle activity (summation of the rectified, smoothed EMG activity from antagonist wrist muscles) at all times (Fig. 2B) during alternating *torque*, *co-contraction* and *rest* conditions.

During the *torque condition* a flexion/extension target torque level in the set  $\{-3.6, -2.4, -1.2, 0, 1.2, 2.4, 3.6\}$  Nm

was displayed in a pre-decided sequence every 4 s. The subjects were instructed to slowly increase the torque to the target level and hold it throughout the target display period, which lasted 3 s, after which they could relax for 1 s before a new target appeared and they contracted again to reach the new torque target.

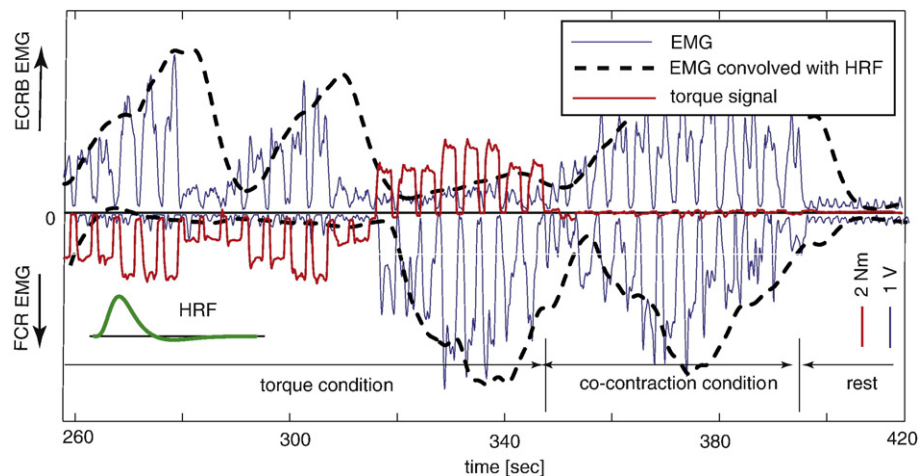
In the *co-contraction condition* muscle activation targets were presented with similar time sequence and the subjects were required to increase muscle activation to the target levels by co-contracting the wrist muscles without applying any torque. The co-contraction target was evaluated by the mean of the EMG recorded during the immediately preceding torque condition. The co-contraction target was presented randomly either to the left or right of the screen so as to make the visual display similar to that in the torque condition. All subjects were trained before the fMRI experiment so as to familiarize with the paradigm and visual feedback.

The three conditions covered the possible combinations of muscle activities, which are when mainly either the flexor or extensor is active (*torque*), when they are both active equally (*co-contraction*) and when they are both inactive (*rest*). The torque and muscle activation target sequences were chosen so as to get slow varying but distinct antagonist muscle EMG profiles after convolution with a haemodynamic response function (HRF).

An example of this is shown in Fig. 2B where during the *torque condition* the presented target sequence  $\{-2.4, -2.4, -2.4, -3.6, -3.6, -3.6, -1.2, -1.2, 2.4, 2.4, 2.4, 3.6, 3.6, 3.6, 2.4, 2.4\}$  Nm. The selected HRF function corresponds to the canonical form used in SPM<sup>1</sup> (Fig. 2). The slow varying EMG profile was necessary to 1) ensure detection by fMRI and 2) reduce effects of errors in the HRF profile. The alternating activation and rest periods limited subject fatigue and movement artifacts while still providing a changing brain activity varying sufficiently slowly to be detected during fMRI.

Two 13 minutes long sessions were conducted for three of the subjects. In each session the regressor pattern repeated four times with a rest period of 30 s between repetitions. For the other three subjects, the two sessions were combined into one long experiment.

<sup>1</sup> <http://www.fil.ion.ucl.ac.uk/spm/>.



**Fig. 2.** Creation of muscle activity regressands. The figure shows the (rectified and integrated) EMG activity recorded on the ECRB (blue plot above zero), FCR (represented by negative blue plot below zero) and torque data over a short period of 160 s during the experiment. These values were collected online during scanning and used to give visual feedback to the subject as shown in Fig. 1B. The expected brain activation (dash black) is constructed by convolving the muscle activity with the HRF (green). Both the (blue) muscle activity and (red) torque pulsate with a time period of 4 s corresponding to the target presentation. Note that in the torque condition there is a non zero torque while either one of the muscle activities is low depending on the direction of the torque. In the co-contraction condition both FCR and ECRB activities are high while the torque signal is close to zero, showing the subject contracted his wrist muscles without applying any directional torque.

The data collected was divided into three sets: a *regression set* and *selection set*, which together form a *training set*, and a distinct *test set*. The first half of the experiment (the first session, for the subjects who did two sessions; the first half of the session, for the subjects who did a single long session) was used as regression set. The first 25% of the second half of the data formed the selection set while the remaining data was used as test set. The regression was performed on the regression set while tuning the parameters using the selection set. The trained function was then tested on the test set to check for the prediction power of the identified mapping.

### Electromyography

EMG was recorded from two muscles acting at the wrist (flexor carpi radialis (FCR) and extensor carpi radialis brevis (ECRB), which are the major contributors to wrist flexion and extension in a sideways position (Haruno and Wolpert, 2005).

After electrode placements for each muscle were determined using functional movements, the area was cleansed with alcohol and abrasive gel (Nuprep, D.O. Weaver & Co, USA). EMG electrodes designed for use in the MR environment (NE-706A, Nihon Kohden, Japan) were filled with EEG electrode paste (Biotech, GE Marquette Medical Systems, Japan) and firmly fixed to the subjects skin with tape. Two electrodes were positioned on the belly of each muscle separated by approximately 1 cm. An elastic cloth sleeve was placed over the electrodes and wires, fixing them against the subjects forearm to avoid any accidental electrode removal and to minimize movements of the electrode wires during the scanning. Once the subject was positioned in the scanner the long braided electrode wires were firmly fixed to prevent movement in the magnetic field. To avoid external noise being carried into the shielded MR room, the electrode wires were passed through multiple ferrite filters before passing through wave guides in the penetration panel of the MR room. EMG channels were cleaned online during scanning and used to provide EMG feedback using the methodology of Ganesh et al. (2007).

### fMRI

A 1.5 T MR scanner (Shimadzu-Marconi ECLIPSE 1.5T Power Drive 250) was used to obtain blood oxygen level-dependent (BOLD) contrast functional images. Images weighted with the apparent transverse relaxation time were obtained with a gradient-echo echoplanar imaging (EPI) sequence. Data was collected from the upper cortex located in the top of the brain to about the center of the ventricles, so as to include areas 4 and 6 completely. Scanning was performed at a *repetition time* of  $TR = 1.5$  s with echo time, 47 ms; flip angle,  $60^\circ$ , 15 slices (thickness 3 mm, gap 0 mm) of  $64 \times 64$  in-plane voxels (in-plane field of view of  $224 \text{ mm}^2$ ).

### Reduction of movement artifacts

Movement artifacts are a major problem in motor experiments as the artifacts are correlated with the movements and thus difficult to separate from actual activity. In our experiment the forearm is fixed, the task is isometric and performed with a distal link, namely the wrist, the movement artifacts are thus expected to be small, if any. However to further reduce any artifacts, the following steps were taken:

While movement artifacts occur immediately on start of the movement and last as long as the movement (Birn et al., 2004), the brain activity is more delayed and long lasting. The experiment was designed to exploit this property in order to reduce potential artifacts in the data. By employing a pulsed activation paradigm (3 s of target presentation followed by 1 s of rest) the movement artifacts are highly correlated with the pulsating torque regressor (Fig. 2) but the correlation of the artifacts and brain activity is weakened. Further, due to the pulsating paradigm, the frequency content of the movement artifacts is much higher than those of the brain activity, so a low-pass filter can be used to filter the artifacts before the regression analysis. In our analysis the cut-off frequency for the low-pass filter was set to 0.2 Hz (5 second period) so as to attenuate movement artifacts which repeat every 4 s.

A custom made bite bar was used in all experiments to reduce head movement. In this way, head motion amplitude was below 1 mm and 1° for all the subjects. This was first detected by the SPM movement parameters and confirmed in a separate experiment using the OPTOTRACK 3020 from Northern Digital Inc. with optical markers fixed on the head providing a measurement at 100 Hz.

The images obtained from the experiment were realigned in SPM before performing the regression.

## Algorithm

### Preprocessing

The brain images from the two sessions were realigned to the first scan of session 1 to ensure there was no displacement among scans from the two sessions. For each subject the region of interest (ROI), in our case area 4 (M1) and area 6 (pre-motor, SMA), were manually mapped on the individual T2 structure scans according to literature (Talairach and Tournoux, 1988; Picard and Strick, 2001). The voxel activity from the ROI was extracted and band-pass filtered between 0.003 Hz (300 s) and 0.2 Hz (5 s) before the regression analysis. The low frequency cut-off value was chosen so as to remove any drift in the fMRI signal while allowing signals corresponding to the muscle activity which repeated every ~196 s. The high cut-off of 0.2 Hz was chosen to remove any possible activity corresponding to movement artifacts. This preprocessed data was used as the regressor in the regression analysis.

The integrated EMG profile obtained after the use of the MR artifact cleaning algorithm (Ganesh et al., 2007) was convolved with a haemodynamic response function (HRF) and used as the regressand for the regression analysis.

### Regression analysis

Our aim was to reconstruct the muscle activity represented by EMG using brain voxels and thus get a functional mapping between a set of voxels in the brain and the activity of a particular muscle. We used the Bayesian linear sparse regression algorithm proposed by Figueiredo (2003) for the mapping and developed a recursive procedure to optimize the function complexity as described below. Linear regression has been previously shown to efficiently reconstruct muscle activity from brain in (monkeys) electrophysiological studies (Townsend et al., 2006; Ting et al., 2005; Koike et al., 2006; Morrow and Miller, 2003). We decided to use a sparse regression method for the following two reasons:

- Sparseness leads to simple functions and helps avoiding over-fitting, as explained in the later section.
- Sparse regression methods set the coefficients of irrelevant voxels to zero thus removing any effects of their activity in the final result. Therefore given muscles will be connected to specific brain activity centers, as is generally agreed in the literature.

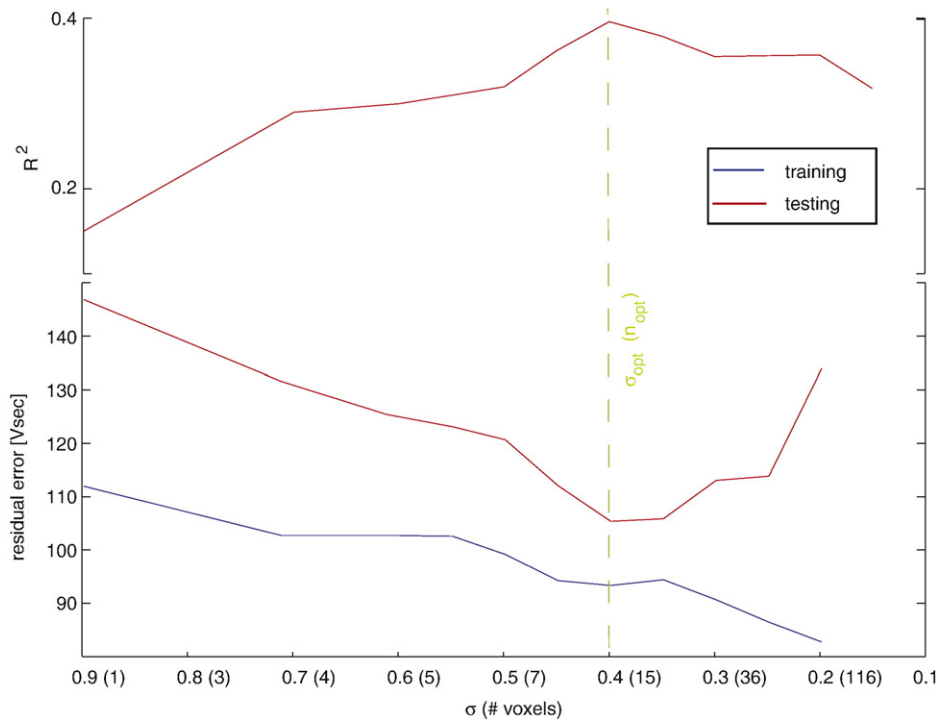
The linear regression function is in the form

$$y = H\beta + w \quad (1)$$

where  $y = [y_1, \dots, y_n]^T$  is the muscle activity (regressand) to be reconstructed,

$$H = \begin{bmatrix} x_1^1 & \dots & x_1^p \\ \vdots & \ddots & \vdots \\ x_n^1 & \dots & x_n^p \end{bmatrix}$$

is the design matrix with the time series of a set of  $p$  voxels from the selected ROI in the 'regression set', and  $w = [w_1, \dots, w_n]^T$



**Fig. 3.** Evolution of error with number of voxels selected during training. The plot shows the evolution of training and selection set testing error and COD values with change in  $\sigma$  for the FCR muscle of subject 2. With the decrease in the value of  $\sigma$  (and increase in the number of selected voxels) the (blue) training error decreases. However the testing error typically shows a (red) U-shaped behavior with respect to  $\sigma$ . The binary search algorithm is used to find the value of  $\sigma_{opt}$  at which the testing error is minimum and COD is maximum. The voxel number corresponding to  $\sigma_{opt}$  is taken as the optimal ( $n_{opt}$ ) for the reconstruction of the EMG.



**Table 1**

The coefficient of determination ( $R^2$ ) values achieved by the mapping during the training and testing phases with a sparse selection of voxels (#)

Sub	FCR reconstruction ( $R^2$ )		# voxels	ECRB reconstruction ( $R^2$ )		# voxels
	Training	Testing		Training	Testing	
1	0.82	0.38	10	0.85	0.62	31
2	0.78	0.38	15	0.46	0.57	52
3	0.58	0.51	3	0.64	0.42	18
4	0.70	0.39	13	0.89	0.63	6
5	0.74	0.42	8	0.77	0.44	29
6	0.76	0.41	5	0.65	0.47	9

is the noise for  $n$  brain scans recorded in an experiment. The parameters  $\beta = [\beta_1, \dots, \beta_p]$  will be computed as a solution of

$$\beta' = \arg \min_{\beta} \left\{ \|H\beta - y\|_2^2 + 2\sigma^2 \alpha \|\beta\|_1 \right\}. \quad (2)$$

The first term with the Euclidean norm  $\|\cdot\|_2$  expresses the least square approximation of the regressand, while the second term with  $\|\beta\|_1 = \sum |\beta_i|$  is to produce a sparse representation. The larger  $\sigma$  will be, the larger weight will be put on minimization of  $\|\beta\|_1$  thus on sparseness. The  $\|\cdot\|_1$  norm, corresponding to the absolute value, has a larger gradient than the  $\|\cdot\|_2$  norm in the vicinity of the minimum and thus should converge fast to zero, i.e. to a sparse representation. This criterion was named as *least absolute shrinkage and selection operator* or LASSO (Tibshirani, 1996).

The  $\alpha$  parameter, coming from a Laplacian prior (Figueiredo, 2003)

$$p(\beta|\alpha) = \prod_i \frac{\alpha}{2} \exp \{-\alpha|\beta_i|\} = \left(\frac{\alpha}{2}\right)^k \exp \{-\alpha\|\beta\|_1\},$$

can be eliminated using a Jeffrey's non-informative hyper-prior as explained in Figueiredo (2003). This allows us to use the expectation-maximization (EM) algorithm to implement the LASSO criterion in two steps (Osborne et al., 2000; Tibshirani, 1996):

E-step

$$U_{(t)} = \text{diag}\{|\beta'_{1,(t)}|, \dots, |\beta'_{k,(t)}|\} \quad (3)$$

M-step

$$\beta_{(t+1)} = U_{(t)}(\sigma^2 I + U_{(t)} H^T H U_{(t)})^{-1} U_{(t)} H^T y \quad (4)$$

$\beta$  is computed iteratively through these steps and determined by the value of  $\sigma$ . The optimal value of  $\sigma$  can be determined by utilizing *over-fitting* as a cost.

### Determination of $\sigma$ to avoid over-fitting

To achieve a good generalization of mapping, it is necessary to control the complexity of the learnt function. A too complex function will reproduce unimportant details of the *training set* thus over-fitting the training data. Such a function will fit the training data well but will be unable to predict a different test data well. Conversely, an overly simple function will not be able to capture the true mapping between the regressor and the regressand, thus under-fitting the training data. In our case the number of voxels chosen during training decides the ability of the model to generalize. The typical behavior of our data set is represented in Fig. 3A. The regression algorithm has the parameter  $\sigma$  to adjust the number of voxels used during training. We utilize over-fitting, detected by the performances on a representative *selection set*, as a criterion to select an optimal value of  $\sigma$ . The EM algorithm loop starts with  $\sigma=1$ . The mapping vector  $\beta(\sigma)$  is computed as explained above and the corresponding coefficient of determination ( $R^2$ ) is evaluated in the *selection set*. A binary search is used to determine the  $\sigma$  maximizing  $R^2$ .

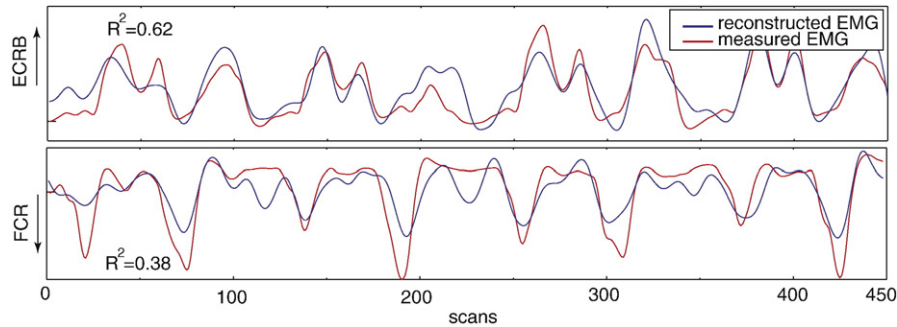
## Results

### Reconstruction performance

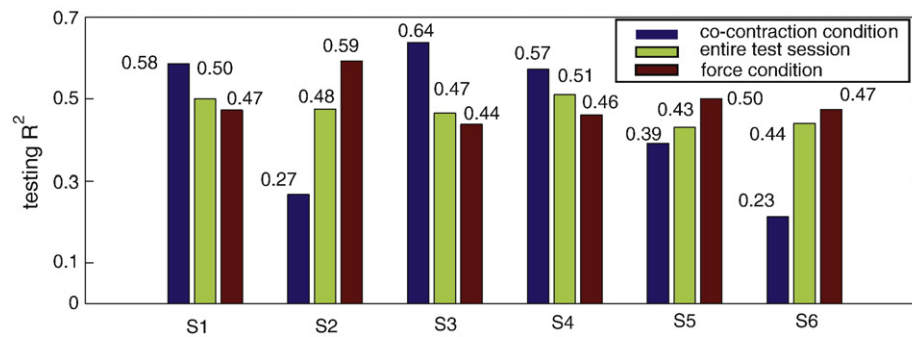
Table 1 shows the performance of the method in reproducing FCR and ECRB EMG activity for 6 subjects. The performance has been quantified using the coefficient of determination  $R^2 = 1 - \frac{\sum (y - y_r)^2}{\sum (y - \bar{y})^2}$ , where  $y$  represents the muscle regressand value,  $\bar{y}$  its average and  $y_r$  the reconstructed value. For the first three subjects, the training and test sets were chosen from the single long session, while for the last three subjects the training regression and test sets were taken from separate sessions.

The  $R^2$  values obtained on the test for the different subjects vary between 0.38 and 0.64, corresponding to a good reconstruction. The  $R^2$  values can be seen to be higher when the training and testing are done in the same session. Fig. 4 shows a representative time series of the predicted and actual EMG for subject 1.

To test if the performance of the algorithm is good enough indicator of actual brain activity and is above statistical chance, an additional analysis was also carried out for each of the subjects, in which a similar regression analysis was performed with voxels from the right prefrontal cortex. This area is generally not expected to have activity related to individual muscle activations.



**Fig. 4.** Predicted waveform. The figure shows the actual (red) and reconstructed (blue) EMG waveforms for the two muscles, ECRB (represented by plot above zero) and FCR (represented by negative plot below zero) for subject 2.



**Fig. 5.** Comparison of performance during different conditions. The plot shows the COD values for each subject, averaged for the two muscles when the testing is done only in the co-contraction condition (blue), force condition (red) and on the entire test data (green). Overall the performance is relatively similar in the two conditions with half the subjects showing a better performance in one condition and half in the other. The performance with the entire test data is roughly the average of the performance in the two conditions.

For five of the six subjects, the trained voxels could not give  $R^2$  values more than 0.18 in testing while for one subject the value reached 0.21 for one muscle. Overall the values were well below the values obtained with the voxels from areas 4 and 6.

To compare the reconstruction performance in the *force* and *co-contraction* conditions (Fig. 5), reconstruction was carried out in the two conditions separately. It was found that the performance of the algorithm was comparable in the two conditions with three subjects showing a better performance during *co-contraction* and three showing better in the *force* condition. The performance of the entire test set was roughly the average of the performance in the individual conditions.

#### Comparison with other regression methods

The results from our algorithm were compared with two other methods commonly used in decoding: *ordinary least square regression* (OLS) and *support vector machine regression* (SVR). OLS was implemented in Matlab using the *pinv* function while SVR was implemented according to the method of Gao et al. (2003). Fig. 6 compares the performance of the sparse linear regression (SLR) algorithm with ordinary least square regression (OLS) and support vector regression (SVR). The OLS and SVR methods performed extremely well in training giving  $R^2$  value of 1 in each case. The weights  $\beta$  (Eq. (1)) obtained from the two methods were different, demonstrating the possibility of distinct mappings fitting the data. In comparison, the SLR gave lower values of  $R^2$  in training, but much higher values in the test phase. This shows that the SLR is able

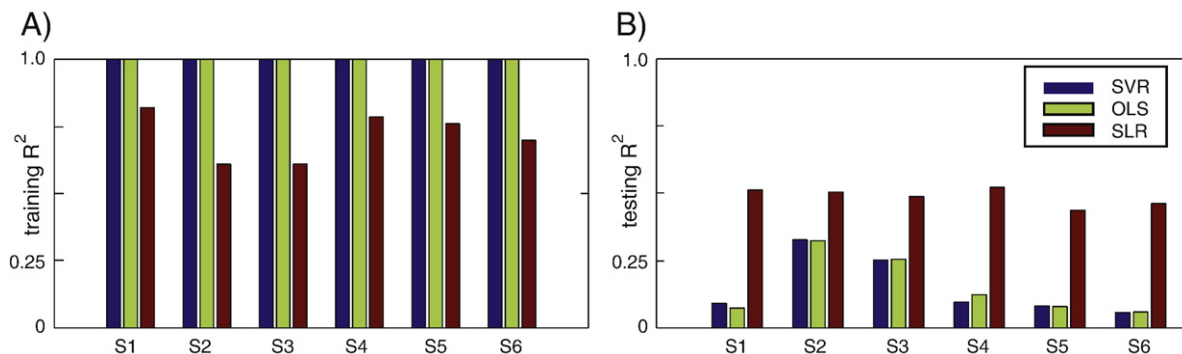
to generalize beyond the training set, in contrast to the other two methods tested here.

#### Distribution of brain activity to distinct muscles

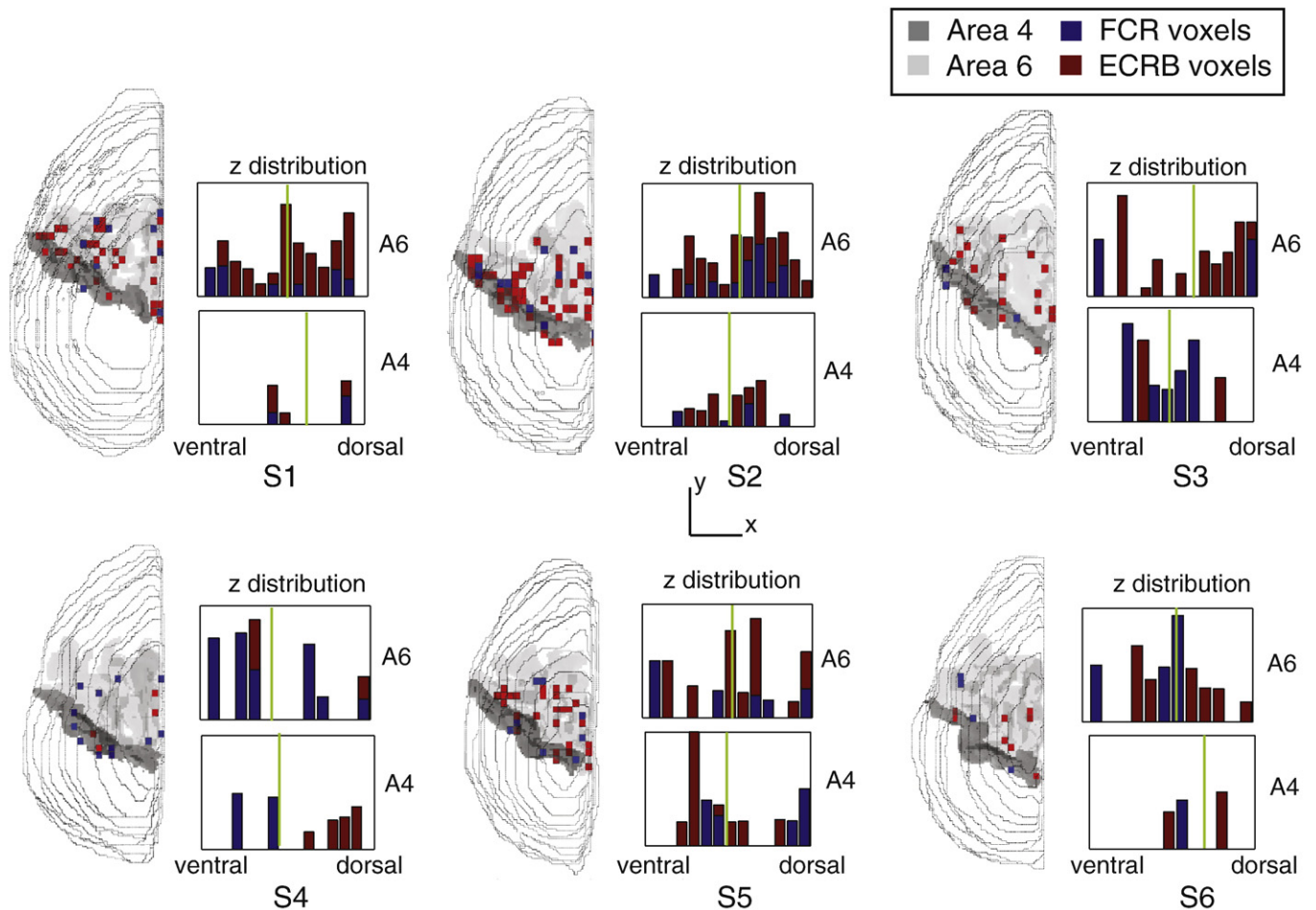
An interesting aspect of the results is the distribution of voxels selected for the FCR and ECRB muscles shown in Fig. 7. The bar charts in the figure give the weighted histogram of the voxels along  $z$ . Each bar represents the sum of the absolute weights as obtained from the sparse regression of voxels in one particular slice. The weights have been normalized by the maximum weight. The sum of the FCR weights (blue bar) and ECRB weights (red bar) indicates the importance of the slice in the reconstruction of that particular muscle activity. Green lines indicate the center of mass of the  $z$ -distribution.

To check that the reconstruction selected the functionally important voxels, we adopted a procedure where the voxels selected by the algorithm for a muscle were removed from the data, after which the regression procedure was re-run. The results showed a drop of about 50% in the coefficients of determination for all subjects when the few selected voxels were not used.

The FCR and ECRB voxel distributions (Fig. 7) were found to spread over cortical areas 4 and 6. In general there were less selected voxels in area 4 than 6. However, except for the FCR voxels of subject 6, for all the subjects and for both the muscles, at least one of the two top voxels in regression was located in area 4. The spread of voxels was roughly over the medial and medial-dorsal region in areas 4 and 6.



**Fig. 6.** Comparison with algorithms generally used for decoding brain activity. The average  $R^2$  values obtained for the two muscles from support vector regression (SVR), ordinary least square (OLS) and sparse linear regression (SLR) are shown for (A) training set and (B) test set. While SVR and OLS give an  $R^2$  value of 1 during training illustrating a perfect fit, they perform poorly on the test set. The SLR is able to generalize much better than SVR and OLS, as exemplified by the similar correlation coefficients in the training and test sets.



**Fig. 7.** Voxel distribution shown as glass brain representation of the left cortex for the six subjects. The grey lines represent the z contours which increase from  $z=0$  mm to  $z=45$  mm in steps of 3 mm, corresponding to the 15 fMRI slices. Areas 4 (dark grey) and 6 (light grey) corresponding to each subject have been marked on the plots. Voxels representing FCR (red) and ECRB (blue) were found to be distributed throughout areas 4 and 6. Note that the voxel size has been magnified in the figure to improve visibility. The bar graphs show the weighted z-distribution. Each bar represents the sum of absolute normalized weights assigned to voxels corresponding to FCR (blue bar) and ECRB (red bar) in one particular slice with the green line representing the center of mass of the distribution.

For all subjects, there was almost no overlap between voxels chosen to reconstruct the FCR and ECRB muscles, even though the two muscles were evaluated in separate regression analyses. The number of voxels chosen to reconstruct the extensor (ECRB) activity (Table 1) was found to be larger than for the flexor reconstruction.

## Discussion

This study showed that fMRI activity related to individual muscle activity can be reliably identified, and provided an algorithm to map fMRI activity to muscle activity. A technique based on Bayesian sparse regression was utilized to identify a linear mapping from fMRI data in areas 4 (M1) and 6 (pre-motor, SMA) of the brain, which have been shown to be directly associated with muscle activation (M1: Morrow and Miller, 2003; Rathelot and Strick, 2006; Koike et al., 2006; Jackson et al., 2007; Townsend et al., 2006 and pre-motor: Cisek et al., 2003; Crammond and Kalaska, 2000; Kurata and Wise, 1988; Ojakangas et al., 2006) to the EMGs of antagonist wrist muscles. The proposed method does not require any manual parameter tuning or thresholding. This method was implemented on six subjects, who had performed an isometric wrist task requiring control of both force and co-

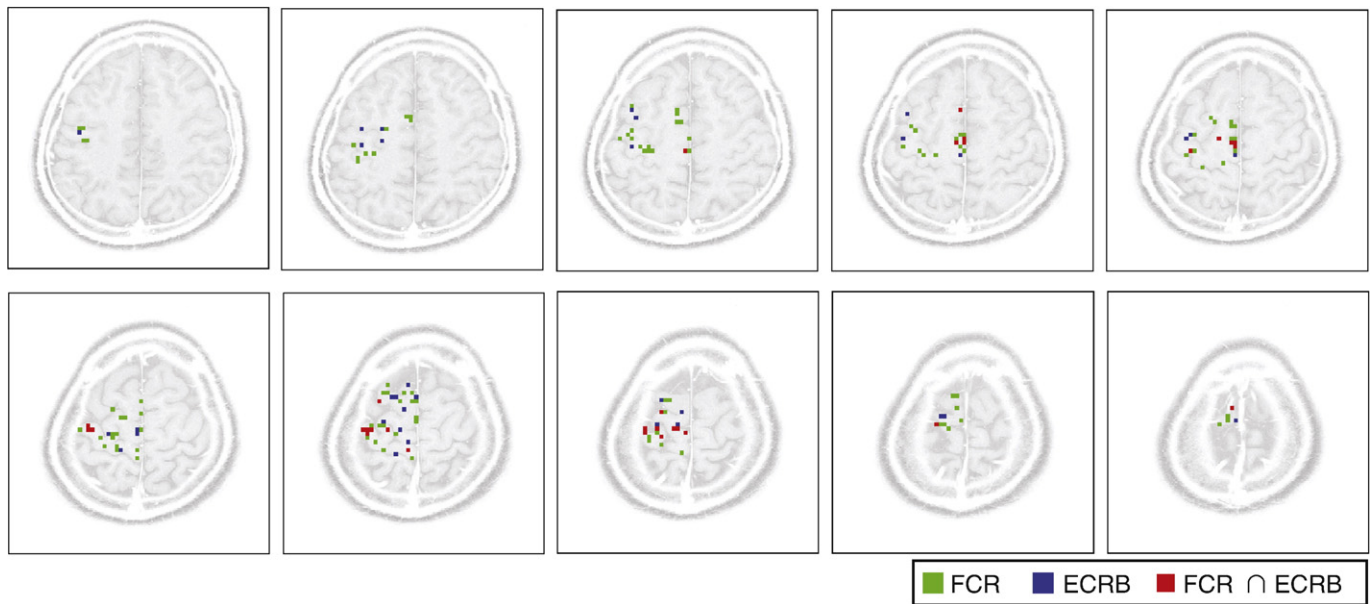
contraction. The results demonstrated that fMRI can be reliably mapped to EMG by testing the mapping on data distinct from the training data.

## Role of spinal reflexes

In this study muscle activity was predicted from the brain activity patterns alone and reflex contributions through direct spinal pathways were neglected. As in most non-human primate studies, we considered voluntary and smooth movements for the reconstruction. Errors due to neglect of reflexes may become prominent in the case of real movements with perturbations (Jackson et al., 2007). While it may be impossible to estimate the reflex contributions from observation of the brain activity alone, comparison of the predicted and actual muscle activities may be used to isolate and study reflex components from EMG signals recorded during experiments.

## SPM and ordinary linear regression

In a system like ours, where single muscle activity is regressed with a brain activity matrix of large dimension, ordinary linear regression provides a perfect fit (Fig. 6A).



**Fig. 8.** Regression of muscle activity using SPM. The images from the training session of a subject were realigned to the first image of the session and analyzed, without smoothing, using a standard SPM analysis with the two muscle activity, and six re-alignment parameters (provided by SPM) as regressors. The figure shows the voxels selected by the SPM regression analysis  $p \leq 0.001$ , uncorrected) for only the FCR (green square), ECRB (blue square) and both FCR and ECRB (red square) in the motor cortex. The voxels selected for the two muscles overlap significantly with around 50% of the ECRB voxels being common for both muscles at this threshold. Three of the top five and 7 voxels of the top 10 peak ECRB voxels are also selected for FCR.

However, the many possible mappings and over-fitting due to the high dimensionality of the regressors are major concerns. These concerns cannot be addressed by the ordinary least square algorithm and the mapping identified is usually data specific and cannot be generalized for test data other than the data used for training (Fig. 6B).

Statistical toolboxes such as Statistical Parametric Mapping (SPM)<sup>1</sup>, used commonly for brain image processing, often rely on simple linear regression to isolate brain activity related to physical processes. Regression in SPM is set up with the brain voxels as the regressand  $y$ , and the physiological parameters, muscle activity in our case, constituting the regressor  $H$  (Eq. (1)). As the physiological parameters are usually small in number, the dimension of the regressor matrix in SPM is small leading to fewer possible mappings. Thresholding of the results in SPM helps in isolating motor cortex voxels from the other brain regions and thus reducing over-fitting. However, within the motor cortex, the voxels corresponding to individual muscles often significantly overlap in space (Fig. 8).

We chose to take motor cortex activity as the regressor as we were interested specifically in the reconstruction of muscle activity without having to account for other neural processes that may be prevalent during the experiment. Sparse regression helps us to generalize well beyond the training set, despite the relative large dimension of our regressor matrix.

The mapping in our sparse regression algorithm and in SPM are essentially inverse of each other (Eq. (1)). Use of both kinds of mapping during analysis can complement and improve insights that can be obtained from the same pair of data sets. For example, the definition of a functional region of interest using SPM may provide additional information in interpreting the sparse regression analysis, while, during analysis of high level motor processes using SPM, sparse regression may be used to equalize or remove activity corresponding to low level muscle processes.

#### Local versus global mapping

The technique presented in this paper, based on fMRI, has lower temporal and spatial resolution compared to electrophysiology. However, electrophysiology usually involves recording a set of neurons locally in the brain during completion of a task, thus it is not possible to know the entire population of the neurons related to, for example, a muscle. To get a complete map of the neurons associated with particular muscles in monkeys, Rathelot and Strick (2006) introduced the rabies virus into a muscle and analyzed the brain area retrogradely labeled by this virus via two stages of synapses. This method, however, cannot be applied to healthy humans, and gives no information about the strength of the connections. In contrast, the technique introduced in this paper can provide a rough but quantitative mapping of the brain areas associated with muscles. It is non-invasive and can be routinely used on humans.

Transcranial magnetic stimulation (TMS) has been used for isolating neural correlates of muscles, however getting focused stimulation with TMS has been difficult (van Elswijk et al., 2008) and the spatial resolution of stimulation is limited to few centimeters (Wassermann et al., 1992, Hallet, 2007). TMS being a top-down process, mapping of all the brain regions associated with a muscle would require stimulation and checking of each individual brain site which is tedious and still the quantitative contribution of brain sites may be difficult to determine. However TMS can help isolating the specific brain regions involved in the activation of a muscle. As we use fMRI to detect brain activity, it is difficult for our algorithm to determine if the selected voxels activate the muscle or receive feedback from the muscle.

While we recorded surface EMG from the FCR and ECRB muscles, other agonist and antagonist wrist muscles may have similar activation patterns and thus similar regressands as FCR and ECRB respectively (Hoffman and Strick 1999). This means



that the brain voxels isolated in this study may correspond to other functionally similar muscles. However, our purpose here was to show that the method works for a given set of regressands. To map individual muscles more precisely, a task could be chosen which generates rich data able to differentiate the activity pattern of each of the involved muscles.

### Activity distribution

For each muscle, the distribution of the voxels selected by our algorithm spread over areas 4 and 6 of the brain, with a bias towards the medial–dorsal region. In area 4, the voxel distribution was not limited to the wrist areas, but was also spread in areas generally associated with shoulder, arm and finger muscle activations. This is in agreement with the distributed brain activity reported for finger muscles in monkeys (Rathelot and Strick 2006), and confirms a previous study in humans (Sanes et al., 1995). While neurons in the dorsal pre-motor region are known to be associated with muscle activity (Cisek et al., 2003; Crammond and Kalaska 2000; Kurata and Wise 1988), dorsal pre-motor is also known to be involved in coding of hand, target position (Pesaran et al., 2006) and gaze shifts (Sylvestre and Cullen 2006). In our experiment, gaze shift was minimized by keeping the targets close to the center and providing the subject with a fixation point. Furthermore, gaze and visual feedback, which may be correlated to the force produced, are not correlated to the activations of individual muscles, especially during the co-contraction condition while the reconstruction of our algorithm is comparable in both conditions (Fig. 5).

While the voxel distribution for both FCR and ECRB muscles was found to be distributed over the motor cortex, the two distributions were found to be orthogonal with an overlap of less than 5%. While it is generally agreed that muscles are controlled by distributed neural sites, the sites have usually been reported to be overlapping (Townsend et al., 2006). The voxels isolated by our algorithm may not represent the complete set of brain areas associated with muscle activation. Minimally contributing voxels and noise may be pruned by the regression process to achieve better generalization, and activity may be missed due to the linear nature of our regression model. Thus, while we cannot conclude that the neural correlates of individual muscle activations are orthogonal, our results show that the voxels most prominently correlated to individual muscle activations have distinct distributions.

The orthogonal mapping of these prominent voxels may give important clues regarding the coordination of muscles in complex actions. The similarity of the brain maps obtained in this study and that of Rathelot and Strick (2006) regarding patchy, intermingled but disjoint representations of individual muscles is in sharp contrast to usual fMRI analysis results; continuous map and large overlap between antagonist muscles (Fig. 8). This might indicate that reconstruction of EMG from fMRI could provide a more causal and anatomical representation of muscle control than its inverse mapping.

### Possible applications

Mapping brain to muscle activity can be used to examine how brain processes involved in motor control and learning in humans are coded in the brain. In particular it may be used to infer muscle combinations such as muscle synergies (d'Avella et al., 2006), or reciprocal and co-activation required to control force and impedance (Haruno and Ganesh 2007, Burdet et al.,

IEEE EMBC, 2004). Mapping between brain and muscles may be used in rehabilitation studies and diagnostics. Similar to electrophysiological studies of post stroke monkeys (Nudo et al., 1996), it will enable precise characterization of the migration of brain activity related to muscles in humans, and may thus explain how brain functions are affected by rehabilitation procedures. Finally, in studies of high level motor processes using fMRI, detected muscle related voxels may be used to help equalize or remove activity corresponding to low level muscle processes.

### Acknowledgments

We thank Satoshi Tada, Ichiro Fujimoto and Yasuhiro Shimada for their valuable help in the preparation of experiments and collection of imaging data. This study was funded by the Human Frontier Science Program (HFSP).

### References

- Birn, R.M., Cox, R.X., Bandettini, P.A., 2004. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. *NeuroImage* 23, 1046–1058.
- Burdet, E., Osu, R., Franklin, D.W., Milner, T.E., Kawato, M., 2001. The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* 414, 446–449.
- Burdet, E., Franklin, D.W., Osu, R., Tee, K.P., Kawato, M., Milner, T.E., 2004. How are internal models of unstable tasks formed? *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 6, 4491–4494.
- Bursztyn, L.L.C.D., Ganesh, G., Imamizu, H., Kawato, M., Flanagan, R.J., 2006. Neural correlates of internal-model loading. *Curr. Biol.* 16, 2440–2445.
- Cisek, P., Crammond, D.J., Kalaska, J.F., 2003. Neural activity in primary motor and dorsal premotor cortex in reaching tasks with the contralateral versus ipsilateral arm. *J. Neurophysiol.* 89, 922–942.
- Crammond, D.J., Kalaska, J.F., 2000. Prior information in motor and premotor cortex: activity during the delay period and effect on pre-movement activity. *J. Neurophysiol.* 84, 986–1005.
- Dai, T.H., Liu, J.Z., Sahgal, V., Brown, R.W., Yue, G.H., 2001. Relationship between muscle output and functional MRI-measured brain activation. *Exp. Brain Res.* 140(3), 290–300.
- Diedrichsen, J., Hashambhoy, Y., Rane, T., Shadmehr, R., 2005. Neural correlates of reach errors. *J. Neurosci.* 25(43), 9919–9931.
- d'Avella, A., Portone, A., Fernandez, L., Lacquaniti, F., 2006. Control of fast-reaching movements by muscle synergy combinations. *J. Neurosci.* 26(30), 7791–7810.
- Ehrsson, H.H., Fagergren, A., Jonsson, T., Westing, G., Johansson, R.S., Forssberg, H., 2000. Cortical activity in precision- versus power-grip tasks: an fMRI study. *J. Neurophysiol.* 83(1), 528–536.
- Figueiredo, M.A., 2003. Adaptive sparseness for supervised learning. *IEEE Trans. Pattern Anal. Mach. Intell.* 25(9), 1150–1159.
- Franklin, D.W., Osu, R., Burdet, E., Kawato, M., Milner, T.E., 2003. Adaptation to stable environments achieved by combined impedance control and inverse dynamics model. *J. Neurophysiol.* 9, 3270–3282.
- Ganesh, G., Franklin, D.W., Gassert, R., Imamizu, H., Kawato, M., 2007. Accurate real-time feedback of surface EMG during fMRI. *J. Neurophysiol.* 97, 912–920.
- Gao, J.B., Gunn, S.R., Harris, C.J., 2003. Mean field method for the support vector machine regression. *Neurocomputing* 50, 391–405.
- Gassert, R., Moser, R., Burdet, E., Bleuler, H., 2006. MRI/fMRI-compatible robotic system with force feedback for interaction with human motion. *IEEE/ASME Trans. Mechatron.* 11, 216–224.
- Hallett, M., 2007. Transcranial magnetic stimulation: a primer. *Neuron* 55(2), 187–199.
- Haruno, M., Ganesh, G., 2007. Dissociable neural correlates of force control and muscle co-activation control revealed by fMRI and on-line EMG feedback. Poster at Neural Control of Movement (NCM) 2007, Seville, Spain.
- Haruno, M., Wolpert, D., 2005. Optimal control of redundant muscles in step-tracking wrist movements. *J. Neurophysiol.* 94(6), 4244–4255.
- Hoffman, D.S., Strick, P.L., 1999. Step-tracking movements of the wrist. IV. Muscle activity associated with movements in different directions. *J. Neurophysiol.* 81(1), 319–333.
- Hogan, N., 1984. Neural, mechanical and geometric factors subserving arm posture in humans. *J. Neurosci.* 4, 2745–2754.
- Imamizu, H., Kuroda, T., Yoshioka, T., Kawato, M., 2004. Functional magnetic resonance imaging examination of two modular architectures for switching multiple internal models. *J. Neurosci.* 24(5), 1173–1181.
- Jackson, A., Mavoori, J., Fetz, E.E., 2007. Correlations between the same motor cortex cells and arm muscles during a trained task, free behavior, and natural sleep in the macaque monkey. *J. Neurophysiol.* 97, 360–374.
- Kakei, S., Hoffman, D., Strick, P.L., 1997. Muscle and movement representations in the primary motor cortex. *Science* 24, 285(5436), 2136–2139.
- Koike, Y., Hirose, H., Sakurai, Y., Iijima, T., 2006. Prediction of arm trajectory from a small number of neuron activities in the primary motor cortex. *Neurosci. Res.* 55, 146–153.

- Kurata, K., Wise, S.P., 1988. Premotor cortex of rhesus monkeys: set-related activity during two conditional motor tasks. *Exp. Brain Res.* 69, 327–343.
- Morrow, M.M., Miller, L.E., 2003. Prediction of muscle activity by populations of sequentially recorded primary motor cortex neurons. *J. Neurophysiol.* 89, 2279–2288.
- Milner, T., Franklin, D.W., Imamizu, H., Kawato, M., 2007. Central control of grasp: manipulation of objects with complex and simple dynamics. *NeuroImage* 36 (2), 388–395.
- Nudo, R.J., Wise, B.M., SiFuentes, F., Milliken, G.W., 1996. Neural substrates for the effects of rehabilitative training on motor recovery after ischemic infarct. *Science* 272 (5269), 1791–1794.
- Ojakangas, C.L., Shaikhouni, A., Friebs, G.M., Caplan, A.H., Serruya, M.D., Saleh, M., Morris, D.S., Donoghue, J.P., 2006. Decoding movement intent from human premotor cortex neurons for neural prosthetic applications. *J. Clin. Neurophysiol.* 23 (6), 577–584.
- Osborne, M., Presnell, B., Turlach, B., 2000. A new approach to variable selection in least squares problems. *IMA J. Numer. Anal.* 20, 389–404.
- Pesaran, B., Nelson, M.J., Andersen, R.A., 2006. Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reach planning. *Neuron* 51, 125–134.
- Picard, N., Strick, P.L., 2001. Imaging the premotor areas. *Curr. Opin. Neurobiol.* 11 (6), 663–672.
- Pope, P., Wing, A.M., Praamstra, P., Miall, R.C., 2005. Force related activations in rhythmic sequence production. *NeuroImage* 27 (4), 909–918.
- Rathelot, J.A., Strick, P.L., 2006. Muscle representation in the macaque motor cortex: an anatomical perspective. *PNAS* 103, 8257–8262.
- Sanes, J.N., Donoghue, J.P., Thangaraj, V., Edelmann, R.R., Warach, S., 1995. Shared neural substrates controlling hand movements in human motor cortex. *Science* 268 (5218), 1696–1698.
- Schaal, S., Sternad, D., Osu, R., Kawato, M., 2004. Rhythmic arm movement is not discrete. *Nat. Neurosci.* 7 (10), 1136–1143.
- Scott, S.H., Kalaska, J.F., 1997. Reaching movements with similar hand paths but different arm orientations: I. Activity of individual cells in motor cortex. *J. Neurophysiol.* 77 (2), 826–852.
- Scott, S.H., Sergio, L.E., Kalaska, J.F., 1997. Reaching movements with similar hand paths but different arm orientations: II. Activity of individual cells in dorsal premotor cortex and parietal area 5. *J. Neurophysiol.* 78 (5), 2413–2426.
- Sylvestre, P.A., Cullen, K.E., 2006. Premotor correlates of integrated feedback control for eye–head gaze shifts. *J. Neurosci.* 26 (18), 4922–4929.
- Talairach, J., Tournoux, P., 1988. *Co-planar Stereotaxic Atlas of the Human Brain, 3-Dimensional Proportional System: An Approach to Cerebral Imaging*. Thieme Medical Publishers, Inc, New York.
- Tibshirani, R., 1996. Regression shrinkage and selection via the LASSO. *J. Royal Statistical Soc. (B)* 58, 267–288.
- Ting, J., D'Souza, A., Yamamoto, K., Yoshioka, T., Hoffman, D., Kakei, S., Sergio, L., Kalaska, J., Kawato, M., Strick, P., Schaal, S., 2005. Predicting EMG Data from M1 Neurons with Variational Bayesian Least Squares. NIPS conference.
- Townsend, B.R., Paninski, L., Lemon, R.N., 2006. Linear encoding of muscle activity in primary motor cortex and cerebellum. *J. Neurophysiol.* 96 (5), 2578–2592.
- Van Duinen, H., Zijdwind, I., Hoogduin, H., Maurits, N., 2005. Surface EMG measurements during fMRI at 3t: accurate EMG recordings after artifact correction. *NeuroImage* 27, 240–246.
- Van Elswijk, G., Kleine, B.U., Overeem, S., Eshuis, B., Hekkert, K.D., Stegeman, D.F., 2008. Muscle imaging: mapping responses to transcranial magnetic stimulation with high-density surface electromyography. *Cortex* 44 (5), 609–616.
- Wassermann, E.M., McShane, L.M., Hallett, M., Cohen, L.G., 1992. Noninvasive mapping of muscle representation in the human motor cortex. *Electroencephalogr. Clin. Neurophysiol.* 85 (1), 1–8.