Department of Physiology
Institute of Neuroscience and Physiology
the Sahlgrenska Academy
University of Gothenburg
2009

MACHINE LEARNING FOR IDENTIFICATION OF BRAIN ACTIVITY PATTERNS

WITH APPLICATIONS IN GENTLE TOUCH PROCESSING

MALIN BJÖRNSDOTTER ÅBERG

UNIVERSITY OF GOTHENBURG

Department of Physiology
Institute of Neuroscience and Physiology
the Sahlgrenska Academy
University of Gothenburg
2009

MACHINE LEARNING FOR IDENTIFICATION OF BRAIN ACTIVITY PATTERNS

WITH APPLICATIONS IN GENTLE TOUCH PROCESSING

MALIN BJÖRNSDOTTER ÅBERG

UNIVERSITY OF GOTHENBURG

Cover illustration: Insular cortex activation pattern in response to a gentle caress. Adaptation of a photograph by Andreas Elmquist.

ISBN: 978-91-628-7854-2
Printed by Geson Hyltetryck, Göteborg, 2009.

MACHINE LEARNING FOR IDENTIFICATION OF BRAIN ACTIVITY PATTERNS
WITH APPLICATIONS IN GENTLE TOUCH PROCESSING

Malin Björnsdotter Åberg, Department of Physiology, Institute of Neuroscience and Physiology, University of Gothenburg, Göteborg, Sweden.

Abstract

Since the first mention of artificial intelligence in the 1950s, the field of machine learning has provided increasingly appealing tools for recognition of otherwise unintelligible pattern representations in complex data structures. Human brain activity, acquired using functional magnetic resonance imaging (fMRI), is a prime example of such complex data where the utility of pattern recognition has been demonstrated in a wide range of studies recently (Haynes et al., Nature Reviews Neuroscience, 2006, 7(7), pp. 523-34).

In contrast to conventional methods, pattern recognition approaches exploit the distributed nature of fMRI activity to achieve superior sensitivities in detecting subtle differences in brain responses. The first objective of this thesis was to implement and empirically evaluate such novel machine learning algorithms for detection and, specifically, spatial localization of regional brain response patterns. Two complementary methods are proposed, namely a Monte Carlo approximation designed for coarse whole-brain mapping, and an evolutionary optimization scheme for refined identification of specific brain regions. As demonstrated on real and simulated data, both methods were more sensitive than conventional approaches in localizing differential brain activity patterns.

The second objective was to utilize these methods to study brain processing of gentle touch mediated by a system of thin, unmyelinated mechanoreceptive C tactile (CT) afferents (Vallbo et al., Brain Research, 1993, 628(310), pp. 301-4). These afferents are thought to modulate affective aspects of tactile sensations, and to act in parallel with thick, myelinated $A\beta$ fibers which signal discriminative information (Löken et al., Nature Neuroscience, 2009, 12(5), pp. 547-8). First, the Monte Carlo algorithm identified differential response patterns due to C tactile and $A\beta$ activation in the posterior insular cortex. Second, the evolutionary scheme revealed a C tactile induced somatotopic insular activation pattern similar to that previously described in relation to other thin-fiber mediated sensations such as pain (Björnsdotter et al., Journal of Neuroscience, 2009, 29(29), pp. 9314-20).

In addition to demonstrating the utility of brain response pattern analysis, the results have a number of implications. The findings support the hypothesis that parallel networks of C tactile and Aβ fibers project affective and discriminative aspects of touch, respectively, and that C tactile afferents follow the projection path of other thin fibers. This further solidifies the hypothesized sensory-affective role of the C tactile system in the maintenance of physical well-being as part of a thin-afferent homeostatic network.

Keywords: somatosensory, machine learning, pattern recognition, fMRI, support vector machines, neuroscience, brain, BOLD, signal processing, artificial intelligence, touch, human, unmyelinated, sensory, affective

ISBN: 978-91-628-7854-2

ETT NERVSYSTEM FÖR VÄLBEFINNANDE KARTLAGT MED ARTIFICIELL INTELLIGENS

Malin Björnsdotter Åberg, Sektionen för fysiologi, Institutionen för neurovetenskap och fysiologi, Göteborgs Universitet, Göteborg, Sverige.

Populärvetenskaplig sammanfattning
Artificiell intelligens (AI) har sedan femtiotalet använts inom vitt skilda områden för att detektera och identifiera subtila mönster i komplexa datamängder. I den här avhandlingen har jag utvecklat två nya metoder baserade på AI med vars hjälp jag har undersökt hur människans hjärna bearbetar emotionella aspekter av hudsmekning. Resultaten visar att en speciell typ av nerver i huden - C-taktila nervfibrer - utgör ett separat nätverk som signalerar emotionella beröringsegenskaper, och att de är organiserade på samma sätt som de tunna fibrer som bland annat signalerar smärta. Detta tyder i sin tur på att det C-taktila nervsystemet är del av ett större tunnfiber-nätverk som vidmakthåller kroppens välbefinnande.

AI-modeller som tränats att känna igen små skillnader i regionala blodflödesförändringar, mätt med funktionell magnetresonansavbildning (fMRI), har tillämpats med stor framgång inom hjärnforskning. Sådan analys kan dels avslöja vilka mentala processer som ett visst hjärnaktivitetsmönster motsvarar, dels identifiera skillnader i blodflödesförändringar med högre känslighet än konventionell analys.

Den här avhandlingen presenterar två nya sådana metoder för att kartlägga mönster av blodflödesförändringar. Metoderna användes för att analysera de specifika mönster som uppstår i hjärnan vid den typ av behaglig smekning som aktiverar de tunna C-taktila nervfibrerna. Resultaten visar att blodflödesmönstret vid behaglig och neutral beröring skiljer sig åt i en del av hjärnan som kallas insulära kortex. Dessutom påverkades behaglighetsmönstret i insulära kortex beroende på vilken kroppsdel som stimulerades, på ett sätt som tidigare beskrivits för smärta.

Studierna stärker teorin om att C-taktila nervfibrer utgör ett separat nätverk som signalerar emotionella aspekter av beröring, och att de är organiserade på samma sätt som de tunna fibrer som bland annat signalerar smärta. Det C-taktila nervsystemet kan utgöra en viktig del i ett homeostatiskt tunnfibernätverk för upprätthållande av kroppens fysiologiska balans.

Resultaten visar också att de nyutvecklade AI-metoderna är väl anpassade för att användas direkt i fysiologisk grundforskning för känslig kartläggning av hjärnaktivitetsmönster.

LIST OF PUBLICATIONS

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Malin Björnsdotter, Karin Rylander and Johan Wessberg, A Monte Carlo method for locally-multivariate brain mapping (submitted manuscript).

II Malin Björnsdotter Åberg and Johan Wessberg, An evolutionary approach to the identification of informative voxel clusters for brain state discrimination, in *IEEE Journal of Selected Topics in Signal Processing*, 2008, 2(6), pp. 919-28.

III Malin Björnsdotter, Karin Rylander, Johan Wessberg and Håkan Olausson, Separate neural systems underpin discriminative and affective touch in humans (manuscript).

IV Malin Björnsdotter, Line Löken, Håkan Olausson, Åke Vallbo and Johan Wessberg, Somatotopic organization of gentle touch processing in the posterior insular cortex, in *Journal of Neuroscience*, 2009, 29(29), pp. 9314-20.

In addition to the papers listed above, parts of the following book chapter, freely available at www.intechweb.org, are also included in the thesis.

Malin Björnsdotter, Machine learning for functional brain mapping, in *Machine Learning*, 2009, InTech Education and Publishing, Vienna, ISBN: 978-953-7619-X-X (in press).

TABLE OF CONTENTS

Introduction and objectives

Since McCarthey coined the term artificial intelligence in 1955, machine learning techniques have won significant ground in virtually every niche imaginable – from DNA sequencing (Baldi and Brunak, 2001) to active galaxy detection (Ball and Brunner, 2009), computerized music composition (Miranda and Biles, 2007), and identification of honeybees (Lavine and Vora, 2005). As evidenced by the wide range of applications, machine learning provides highly appealing tools for sophisticated recognition of otherwise unintelligible pattern representations embedded in complex data structures (Duda, Hart, and Stork, 2000).

Human brain activity measured by functional magnetic resonance imaging (fMRI) is a prime example of such complex data, being excessively noisy, high-dimensional and spatially distributed. FMRI signals are traditionally approached by descriptive statistical methods where average signal changes in single locations are related to experimental conditions. Although tremendously productive in the mapping of brain areas which are activated by various conditions, such univariate average measures are poorly suited for capturing temporally variable local and global interactions across neural networks in the cortex.

Pattern recognition approaches, in contrast, provide tools for detection and identification of transient patterns of brain activity, integrated across multiple measuring points (Mitchell et al., 2004; Haynes and Rees, 2006; Norman et al., 2006). These techniques allow computer models to learn desired behaviors from examples, in virtually the same sense that humans learn. A model can be trained to recognize and decode subtle intrinsic signal patterns correlated to given brain states – such as the indistinct fMRI pattern, consisting of tens of thousands of voxels, produced by a single touch stimulus (Beauchamp, LaConte, and Yasar, 2009). Trained models can be applied in mental state tracking (Polyn et al., 2005), lie detection (Davatzikos et al., 2005), the decoding of single visual stimuli – visible (Haxby et al., 2001; Cox and Savoy, 2003; Kamitani and Tong, 2005), as well as invisible (Haynes and Rees, 2005a) – biofeedback (Yoo et al., 2006), and various types of real-time fMRI analyses (LaConte, Peltier, and Hu, 2007; deCharms, 2008). In addition, pattern recognition methods capture and utilize the spatially distributed nature of fMRI activity, and are therefore more sensitive to subtle differences between brain responses than traditional univariate approaches (Kriegeskorte, Goebel, and Bandettini, 2006; De Martino et al., 2008; Björnsdotter Åberg and Wessberg, 2008; Björnsdotter, Rylander, and Wessberg, 2009).

Despite their promising potential, these techniques are only recently beginning to see more than limited use in neuroimaging. This is in part due to the fact that appropriate application of machine learning concepts requires not only an understanding of physiology, but also a solid technical and mathematical background. The first objective of this thesis, consequently, was to implement and empirically evaluate novel machine learning algorithms for ef-

fectively applicable fMRI pattern analysis. Although the identification of single events lies at the core of pattern recognition, information concerning the constitution of the representational pattern is also required for physiologically interpretable results. Accordingly, the focus of this thesis was detection and, primarily, spatial localization of brain response patterns. Two such complementary multivariate brain mapping methods were developed. The first, based on Monte Carlo approximations (see section 5.1 and paper I), was designed for fast, non-specific whole-brain mapping, whereas the second, an evolutionary algorithm optimization scheme (see section 5.2 and paper II), was implemented for refined identification of brain regions in a highly tailored fashion.

A particular aspiration of the thesis research was to implement generic pattern recognition methods directly applicable in a variety of neuroimaging studies. As a second objective, therefore, these novel machine learning algorithms were utilized to explore brain activation patterns in response to gentle tactile stimulation. Specifically, touch mediated through a recently discovered type of cutaneous sensory nerve fibers, termed C tactile (or CT) afferents was studied. These afferents innervate human hairy skin and react vigorously to soft mechanical stimulation, such as a gentle caress (Vallbo et al., 1993; Vallbo, Olausson, and Wessberg, 1999). As opposed to other nerve fibers, C tactile afferent firing rates correlate with the perceived pleasantness of tactile stimulation (Löken et al., 2009), and are thought to project affective, emotional aspects of the tactile experience (Vallbo, Olausson, and Wessberg, 1999; Olausson et al., 2002; Wessberg et al., 2003; McGlone et al., 2007; Olausson et al., 2008a).

Discriminative properties of touch, in contrast, are relayed through thick, myelinated (Aβ) fibers. These are activated by all types of mechanical stimuli (including gentle touch) and project to the primary and secondary somatosensory cortices (Kaas, 2007). In two rare patients who lack Aβ afferents, it was recently revealed that C tactile afferents activate a brain region called the insular cortex but not the somatosensory areas (Olausson et al., 2002; Olausson et al., 2008b). In healthy individuals, however, such projection differences remain to be demonstrated. Hence, one specific aim of this thesis was to investigate differential brain patterns in response to Aβ stimulation and combined C tactile and Aβ stimulation in healthy subjects (section 10.1 and paper III).

C tactile afferents belong to a class of thin fibers that also transmits pain and temperature sensations (see Craig, 2002 for a review of this system). Whereas the specific projections of C tactile fibers are poorly understood, extensive studies have detailed the pain and temperature fiber pathways from the skin to the cortex (Craig et al., 1994; Craig, Zhang, and Blomqvist, 1999; Blomqvist and Craig, 2000; Craig et al., 2000; Craig, 2003b; Craig and Zhang, 2006). In particular, functional imaging of cooling as well as painful stimuli has shown that the posterior portion of the insular cortex is somatotopically organized with upper body afferents activating regions anterior to those of the lower body (Hua et al., 2005; Brooks et al., 2005; Henderson, Gandevia, and Macefield, 2007). A similar organization of C tactile afferent brain projections

would substantiate the link to the pain and temperature fiber pathways. The final specific aim of this thesis, therefore, was to investigate whether C tactile afferent activation patterns in the posterior insular cortex follow a somatotopic organization similar to that shown for pain and temperature projections (section 10.2 and paper IV).

This thesis is divided into two parts. Part I presents a technical background to fMRI and conventional signal processing, as well as considerations regarding machine learning in general and brain activity pattern recognition of in particular. Specifically, the two novel methods developed during the course of this thesis research are presented and evaluated. Part II, in contrast, outlines the physiology underpinning tactile sensation and details how the pattern recognition methods were applied to differentiate brain response patterns produced by C fiber mediated gentle touch. Finally, the two parts are jointly discussed in an attempt to unify methodology and physiology.

Part I:
Brain activity acquisition & processing

1. Functional magnetic resonance imaging (fMRI)

Magnetic resonance imaging (MRI) involves detection and analysis of signals derived from intrinsic atomic properties of matter, and enables noninvasive detailed exploration of biological tissues (Hashemi, Bradley, and Lisanti, 2004). In particular, local blood oxygenation changes in response to neural processing measured with functional MRI (fMRI) provides an effective, albeit indirect, indication of relative levels of brain activity (Norris, 2006).

1.1 Magnetic resonance imaging

MRI utilizes the nuclear spin and magnetism of atoms to obtain information about their environment (Hashemi, Bradley, and Lisanti, 2004). The spin refers to the inherent angular momentum possessed by all atomic nuclei with an odd number of protons and/or neutrons. One such substance is prevalent in organic tissue – the hydrogen atom. As a result of the spin, the hydrogen atom also has a magnetic dipole moment and, therefore, behaves like a small magnet with a north and south pole as illustrated in figure 1.1.

Figure 1.1: A hydrogen atom.

Such magnetic dipole moments in tissue are normally randomly oriented producing a net magnetization of approximately zero (figure 1.2A). Application of an external magnetic field, however, induces gradual movements in the magnetic moments, and a small portion align with the magnetic field (longitudinal relaxation) with a time constant T_1 (typically around 1 second, depending on field strength and type of tissue). The alignment of the magnetic moments results in a net magnetization in the direction of the field (figure 1.2B).

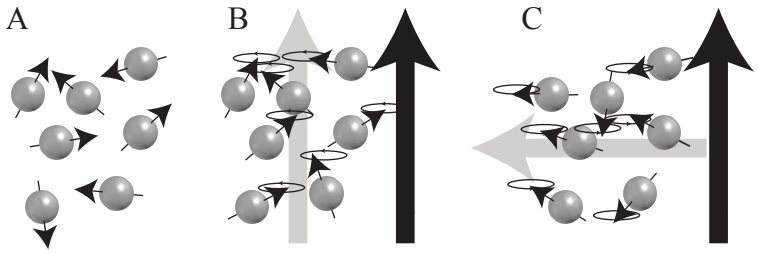

Figure 1.2: Hydrogen magnetic dipole moments in a sample (such as brain tissue) are normally oriented in a random fashion (**A**). When an external magnetic field (black arrow) is applied, a small portion of the dipole moments align with the field which produces a net magnetic moment (**B**; grey arrow). Also, the net magnetization begins to precess around the field. After application of an RF-pulse, the net magnetization is flipped 90° (**C**).

Also, the net magnetization will change the direction of the axis of rotation around the field axis in a process called precession (figure 1.3). The frequency of precession, v_0, is called the Larmor or nuclear magnetic resonance frequency. The Larmor frequency is proportional to the strength of the external magnetic field: $v_0 = \gamma B_0$, where B_0 (measured in Tesla, T) is the external magnetic field strength, and γ (units: MHz/T) is the gyromagnetic ratio. This element-specific ratio expresses the relationship between the angular momentum and the magnetic moment of the nuclei, and the value for hydrogen is 42.58 MHz/T. Magnetic field strengths used for human fMRI range from 1.5 (approximately 30 000 times the earth's magnetic field) to 9 T.

Figure 1.3: A precessing hydrogen atom.

In response to application of energy at the Larmor frequency, the net magnetization begins to resonate and move out of alignment with the external magnetic field. For hydrogen, the Larmor frequency corresponds to the radio frequency (RF) band, and application of rapidly oscillating electromagnetic RF pulses will change the alignment of the hydrogen magnetic moments (while elements with other precessional frequencies are unaffected). The RF pulse can be applied to flip the net magnetization 90°, and thus change the net magnetization from being aligned with the external magnetic field to pointing perpendicular to the field (while still precessing around the field; figure 1.2C). As a result of the flip, the magnetic moments are in phase and produce a precessing net magnetization which can induce an alternating current in a coil placed nearby.

Importantly, the current induced in the coil decays over time (relaxation). The decay is in part due to thermal motion on the molecular level, realigning the net magnetic moments with the external magnetic field (T_1 relaxation). Also, random interactions of nuclei result in a loss of coherence of the precession which reduces the net magnetization (T_2 relaxation). Moreover, inhomogeneities in the magnetic field cause dephasing since the precession frequency of the nuclei is proportional to the strength of B_0. This effect in combination with the random nuclei interactions is referred to as T_2^* relaxation.

Fundamentally important for functional and structural biological imaging, the relaxation times differ between various tissues such as muscle and bone, grey and white brain matter, etc. Structural images can therefore be reconstructed from the acquired relaxation signals. An example of such an anatomical image is shown in figure 1.4A.

1.2 BOLD functional imaging

In addition to structural brain images, magnetic resonance techniques provide a possibility to acquire signals related to active functions of the brain (Norris, 2006). Such functional MRI detects effects of hemodynamics, including blood flow, blood volume and oxygen consumption, on the basis of hemoglobin (the molecule in red blood cells which contains oxygen). Hemoglobin is diamagnetic when oxygenated and paramagnetic when deoxygenated, i.e. possesses

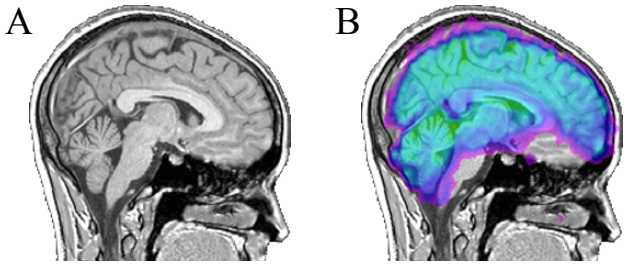

Figure 1.4: **A**. MRI structural image. **B**. fMRI functional image superimposed on a structural image.

different magnetic characteristics depending on oxygenation state. This phenomenon, in combination with the measured T_2^* relaxation (see section 1.1), is used in fMRI to detect magnetic differences between oxygenated and deoxygenated blood in the brain. Specifically, blood-oxygen-level dependent (BOLD) fMRI is used to identify temporal and spatial variations in the proportion of oxygenated to deoxygenated blood, which, in turn, is an indication of blood flow changes (Ogawa et al., 1990). A relative increase in blood flow results in a positive signal, and vice versa.

The fMRI BOLD signal is acquired one brain volume at a time, and each measuring point in the three-dimensional volume grid is referred to as a *voxel* (see figure 1.4B). Depending on the MRI scanner properties, the time required to acquire a single whole brain volume (repetition time, TR) typically ranges from 1-4 seconds with voxel dimensions of 2-4 mm per side. Thus, fMRI has relatively poor temporal resolution and excellent spatial resolution compared to noninvasive electrophysiological measuring techniques such as electroencephalography (EEG). Moreover, fMRI is entirely noninvasive, as opposed to other brain imaging techniques including positron emission tomography (PET).

1.3 Neural correlates of BOLD

A positive correlation between local blood flow and brain cell activation level was first observed in the 1890s (Roy and Sherrington, 1890). The temporal pattern of blood flow changes in response to activated nerve cells is called the hemodynamic response function (HRF), and, although differing between brain regions (Leoni et al., 2008), a double gamma function is thought to be a good estimate of the response in humans (figure 1.5; Büchel et al., 1998). A generic blood flow response to a briefly presented stimulus lasts up to 20 seconds and peaks at approximately 6 seconds. The temporal resolution of fMRI is thus inherently limited due to the delay in the hemodynamic response.

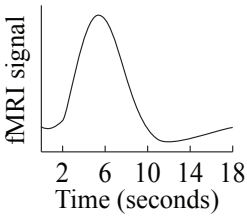

Figure 1.5: A double gamma function estimate of the hemodynamic response function (HRF).

Although it is generally assumed that changes in blood flow are prompted by metabolic effects such as increased oxygen consumption by activated nerve cells, the exact relationship between neural activity and the BOLD signal is not fully understood. The peak of the HRF appears to be a substantial overcompensation (supplying more blood than is required by metabolic demands), and the mechanisms for this are unknown (Norris, 2006). Moreover, the BOLD signal is an indirect measure of neural activity, and is therefore susceptible to influence by a number of physical parameters of non-neural nature. The BOLD signal can, in fact, reflect increased blood flow into an area despite no local neural activity (Sirotin and Das, 2008).

Nonetheless, concurrent intracortical recordings of neural signals and fMRI responses in the visual cortex in monkeys have shown that regional brain activity (and in particular local field potentials) are significantly correlated to the hemodynamic response (Logothetis et al., 2001; and see e.g. Heeger and Ress, 2002, or Goense and Logothetis, 2008, for further details on the relationship between neural activity and the BOLD signal).

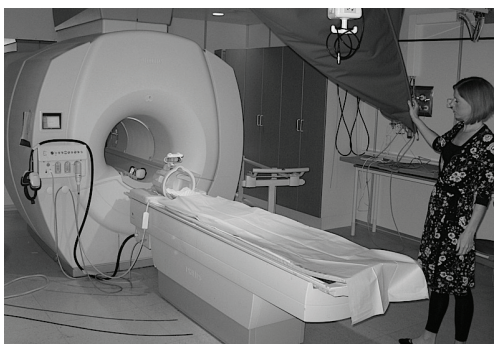

Figure 1.6: The 1.5 Tesla MRI scanner at the Sahlgrenska University hospital in Göteborg, used to acquire data for the studies in this thesis.

2. Data acquisition and preprocessing

A variety of considerations are necessary to effectively acquire and analyze fMRI data for research purposes. Prior to acquisition, the experimental paradigm must be carefully designed in order to ensure that the actual effect of interest is studied, and, subsequently, a number of signal preprocessing steps are required for noise reduction.

2.1 Experimental paradigm

Careful attention needs to be paid to the type and organization of conditions presented during the experiment in order to isolate the effect of interest (as opposed to noise or unwanted cognitive processes). Typically, paradigms involve a number of stimulus conditions which are contrasted in subsequent analysis to remove confounding variables. During the scanning session, the conditions are presented in a predetermined fashion. Influenced by positron emission tomography (PET) imaging where extended stimulation periods are required in order to produce stable activations (Muehllehner and Karp, 2006), fMRI studies often utilize experimental paradigms which alternate extended periods of stimuli being 'on' or 'off' (see figure 2.1A; Turner, Howseman, and Friston, 1998).

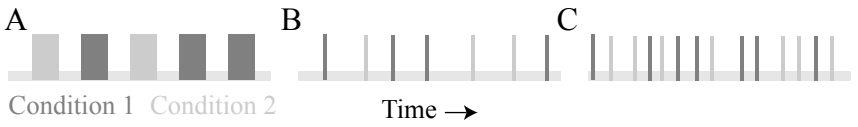

Figure 2.1: Schematic of **A**) block, **B**) slow event-related, and **C**) fast event-related experimental paradigm designs with two stimulus conditions.

These block designs are appealing due to ease of presentation and analysis, as well as to the relatively high signal-to-noise ratios achieved. Brief stimuli can, however, produce a measurable BOLD response (e.g. 34 ms; Rosen, Buckner, and Dale, 1998), which is utilized in event-related designs (figure 2.1B and C; Buckner, 1998). More dynamic responses can be obtained, and, given similar scanning times, more stimulus repetitions can be applied (see e.g. Kriegeskorte et al., 2008). A disadvantage of event-related paradigms is the lower functional signal-to-noise ratio compared to block design paradigms (Bandettini and Cox, 2000). It should be noted, however, that the strict division of paradigms into these categories is idealized and experimental designs follow a range of variants.

2.2 Preprocessing

A variety of software exists for both preprocessing and subsequent statistical analysis of fMRI data, including the freely available Neurolens (neurolens.org), SPM (fil.ion.ucl.ac.uk/spm; Friston, 2007) and AFNI (afni.nimh.nih.gov), as well as commercial software such as BrainVoyager (brainvoyager.com). The following preprocessing steps are typically applied, although all are not necessarily required and further steps can be included to improve the analysis (see e.g. Friston, 2007 or Henson, 2003 for more details).

- *Slice-time correction*: The acquisition of an entire brain volume generally takes of the order of 2-4 seconds (depending on MRI scanner parameters), during which slices of brain tissue are scanned consecutively. The resulting shift in acquisition time between slices can be corrected by resampling the time courses with linear interpolation such that all voxels in a given volume represent the signal at the same point in time.
- *Motion correction*: Even slight head movements have a severe effect on the quality of the data as a result of the high spatial resolution of fMRI, and must therefore be corrected. A variety of algorithms are available in any of the software packages listed above. It should be noted, however, that these algorithms usually only correct temporal changes in spatial alignment, and motion induced effects such as differential signal distortion are often impossible to correct post hoc.
- *Temporal filtering and detrending*: Temporal drifts which can significantly affect the results are usually reduced using temporal high-pass filtering, although more sophisticated methods have shown promising results (Friman et al., 2004).
- *Spatial smoothing*: In order to reflect a degree of spatial integration, spatial smoothing is often applied to the volume time series using a Gaussian FWHM kernel in the range of 3-12 millimeters. For some types of statistical analysis (e.g. the general linear model described in section 3.1), smoothing to whiten the spatial distribution of the signal is required for statistical inference (Worsley et al., 1992; Worsley et al., 1996).
- *Spatial normalization*: Individual brains are highly anatomically variable, and for group analysis and comparison with brain atlases the acquired signals must be projected into a standard brain format such as Talairach (Talairach and Tournoux, 1988) or MNI (Montreal Neurological Institute; Evans et al., 1993) space. Various algorithms have been proposed to this end (Collins et al., 1994).

3. Conventional brain mapping

Conventional fMRI analyses aim to identify brain regions where a certain stimulation or condition of interest produces a significant change in BOLD signal. Such analysis assumes that brain function is highly modular, and has been criticized as a contemporary version of phrenology. Opponents argue that a holistic, distributed model is a better representation of cognitive function (see e.g. Fodor, 1983, and Uttal, 2003 for a discussion favoring the modular and holistic view, respectively). Well-designed and appropriately interpreted brain mapping with deep roots in related research fields (including electrophysiological recordings) have, nevertheless, been a tremendously successful tool for exploring brain function.

As opposed to multivariate pattern recognition methods (see section 4-5), conventional analysis is *univariate*. Univariate analyses treat each voxel as an independent measurement with no interaction with neighboring elements. Numerous variations of univariate fMRI analysis techniques are widely used, and the field is under active research. The following chapter describes two of the most commonly used approaches (also utilized in paper I-IV) namely the general linear model (Friston et al., 1994) and event-related averaging (Rosen, Buckner, and Dale, 1998).

3.1 General linear modeling

A highly lucrative univariate approach is parametric statistical analysis to produce images (statistical parametric maps) which identify brain regions that show significant signal changes in response to the experimental conditions (see e.g. Henson, 2003). A spatially invariant model of the expected BOLD response is fitted independently at each voxel's time course and the differences between estimated activation levels during two or more experimental conditions are tested (Friston et al., 1994).

Most such parametric modeling techniques are versions of the general linear model (GLM). The GLM aims to explain the variation of the time course $\mathbf{y}$, in terms of a linear combination of explanatory variables $\mathbf{x}$ and an error ε term:

$$\mathbf{y} = \mathbf{x}\beta + \varepsilon \tag{3.1}$$

The matrix $\mathbf{x}$ (the design matrix) contains one row per time point and one column per explanatory variable in the model (e.g. representing the presence or absence of a specific condition). In order to detect activations, the magnitude of the parameter β is estimated by solving the following equation:

$$\mathbf{x}^T\mathbf{y} = (\mathbf{x}^T\mathbf{x})\hat{\beta} \tag{3.2}$$

where $\hat{\beta}$ corresponds to the best linear estimate of β. Given that $X^T X$ is invertible, $\hat{\beta}$ can be estimated as:

$$\hat{\beta} = (\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T \mathbf{y} \tag{3.3}$$

A number of additional parameters (regressors) can be included in the GLM analysis, such as cardiac responses, respiration, drift, motion correction parameters or other confounds.

Comparisons between conditions are expressed as contrasts, representing linear combinations of $\hat{\beta}$ values. If the respective contrasts are formulated in a vector c, a t-statistic testing whether the condition combinations specified in c differ significantly from the null hypothesis ($c^T \hat{\beta} = 0$) can be computed in each voxel as follows:

$$t = \frac{c^T \hat{\beta}}{\sqrt{var(\varepsilon) c^T (X^T X)^{-1} c}} \tag{3.4}$$

The obtained t-maps highlight brain locations where the conditions of interest differ, and are usually color-coded and overlaid on structural MRI images in a visually appealing fashion (see figure 3.1).

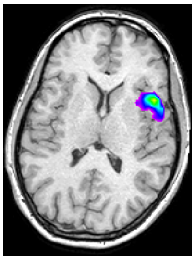

Figure 3.1: GLM t-map overlaid on anatomical image.

Importantly, this massively univariate testing results in one statistic per voxel, and thus produces a classical problem of multiple comparisons (Friston et al., 1994). With an error probability of $p<0.05$, the same test repeated 100 times under the assumption that there is no effect (null hypothesis) will yield five cases of false positives on average. Computing the t-test in equation 3.4 is statistically equivalent to repeating the same test for each voxel, thus, for 100 000 voxels, approximately 5 000 would be assumed (falsely) to be significantly activated by chance. Numerous methods have been proposed for correction of the problem of multiple comparisons, including simple (but overly conservative) Bonferroni correction (Nichols and Hayasaka, 2003) and false discovery rate (FDR) approaches (Genovese, Lazar, and Nichols, 2002; Chumbley and Friston, 2009; Schwartzman et al., 2009).

3.2 Event-related averaging

Temporal brain response dynamics can be explored by computing an event-related average BOLD signal change. The event-related response is usually expressed as percent signal change compared to a baseline, x_b, which can, for example, be estimated as the average across a specified number of time

points preceding each trial. The percent signal change is then calculated as $(x_t - x_b)/x_b$, where x_t is the BOLD value at time t.

Event-related averages are generally computed in regions of interest identified through other methods, and can be particularly useful in pattern recognition studies to examine BOLD response directions (increases or decreases; see e.g. paper III).

4. Brain activity pattern identification

In contrast to conventional univariate brain mapping which estimates voxel-by-voxel signal increases or decreases, pattern recognition approaches identify patterns of activity changes integrated across multiple voxels in a *multivariate* fashion. Specifically, classifier-based machine learning algorithms teach a computer model to recognize complex, spatially distributed brain signal changes related to specific experimental conditions. Since these models can also be used to predict (decode) brain states, the approach is often, somewhat equivocally, referred to as "mind reading" (Norman et al., 2006; deCharms, 2008).

The term multivariate is used interchangeably with multivoxel, and the general approach is often called multivoxel pattern analysis (MVPA; Norman et al., 2006). Akin to conventional univariate techniques, MVPA may be used in brain mapping to identify regions containing multivariately differential BOLD responses patterns. A direct link between instant fMRI activity and brain states is, in addition, provided.

In this chapter, machine learning is introduced along with general technical implementation considerations, and specific aspects for application of multivariate analysis in functional imaging are described.

4.1 Machine learning

Machine learning is concerned with algorithms allowing computer models to "learn" from examples and generalize learned behaviors to make intelligent decisions given new data. The structure of the data and desired behavior of the algorithm determines teaching and learning operations. Although various exploratory algorithms, where no category labels or hypotheses pertaining to the experimental conditions are supplied (i.e. unsupervised learning) have been successfully applied in brain activity analysis (e.g. independent component analysis; ICA; De Martino et al., 2007), this thesis investigates hypothesis-driven brain mapping with clear links to the experimental conditions. This category of machine learning is referred to as supervised learning or classification.

4.1.1 Supervised learning and classification

Supervised learning algorithms train a computer model to recognize characteristic consistencies in signal patterns with the specific aim of relating each pattern to one of the supplied categories. A properly trained model (classifier) can subsequently be used to classify data instances where the category is unknown.

The data points representative of the signal with which the classifier is trained are called features (also known as attributes or variables). In fMRI,

the features correspond intimately with voxels and these terms are used interchangeably. The data categories are called labels, or, in fMRI, (experimental) conditions or brain states. The instances of the data are termed samples, patterns, examples, observations, or, in fMRI, trials or volumes.

A classifier can be described as the mapping from a feature space (e.g. voxels) to a defined set of labels (e.g. experimental conditions), or, mathematically: consider a set of signal samples $\{(\mathbf{x_1}, y_\mathbf{1}), \ldots, (\mathbf{x_n}, y_\mathbf{n})\}$. A classifier is then a process $h : \mathscr{X} \rightarrow \mathscr{Y}$, mapping the object $\mathbf{x} \in \mathscr{X}$ to the corresponding label $\mathbf{y} \in \mathscr{Y}$. Although multi-category maps have been demonstrated (see e.g. Björnsdotter Åberg and Wessberg, 2008), multivariate brain mapping generally involves binary data where samples belong to either of two categories. Given such binary data in the form:

$$\mathscr{D} = \{(\mathbf{x}_i, y_i) | \mathbf{x}_i \in \mathbb{R}^p, y_i \in \{-1, 1\}\}_{i=1}^n, \tag{4.1}$$

where y_i is the category label with a value of either 1 or -1, indicating the category to which the point x_i belongs, and each x_i is a p-dimensional real vector of features, the aim of the classifier is to construct a classification rule in the form of a decision boundary hyperplane that separates the data points x_i where $c_i = 1$ from those where $c_i = -1$. An example of a simple classifier is shown in figure 4.1A.

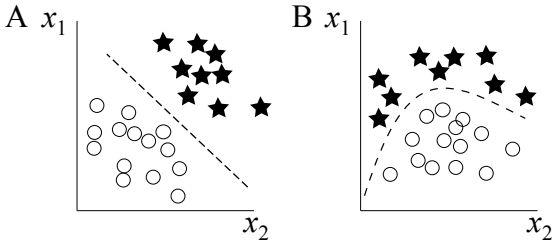

Figure 4.1: Illustration of **A**) linearly and **B**) nonlinearly separable data for a two-feature (x_1 and x_2), binary dataset. For the linearly separable categories the classifier hyperplane (dotted line) corresponds to a straight line of the standard form: $w_1x_1 + w_2x_2 = b$. The classification rule of this classifier is to assign the sample to category ★ if $w_1x_1 + w_2x_2 > b$ and to category ◯ if $w_1x_1 + w_2x_2 \leq b$.

The data categories might not be linearly separable (figure 4.1B), and, if so, a classifier which can capture nonlinear effects is required. Linear classifier rules are based on linear combinations of features, whereas nonlinear classifiers represent more complicated, nonlinear relationships with the features (Theodoridis and Koutroumbas, 2006). Nonlinear classifiers can also capture linear data structures, but generally require more computational resources, reduce model interpretability and are more sensitive to overfitting (Mørch et al., 1997).

4.1.2 Classification algorithms

A wide range of algorithms to construct classifiers have been suggested with varying performance depending on the structure and quality of the data and desired generalization behavior. Some common classifiers include linear discriminant analysis and Fisher's linear discriminant (Fisher, 1936), naïve Bayes classifier (Buntine, 1989), k-nearest neighbor (Cover and Hart, 1967) and artificial neural networks (Haykin, 1999). As noted in section 4.2.5, a substantial amount of MVPA studies (including papers I-IV) utilized linear or nonlinear support vector machines (SVMs; Vapnik and Lerner, 1963) and these are therefore described in more detail.

Support vector machines

SVMs construct a separating hyperplane such that the distance from the hyperplane to the nearest data point is maximized (Vapnik, 1995; Suykens et al., 2002). Given the notation in the previous sections, the SVM algorithm attempts to find the maximum-margin hyperplane which separates the x_i points where $c_i = 1$ from those having $c_i = -1$. Any hyperplane can be written as the set of points x satisfying

$$\mathbf{w} \cdot x - b = 0, \tag{4.2}$$

where $\cdot$ denotes the dot product. $\mathbf{w}$ is a normal vector, perpendicular to the separating hyperplane. As illustrated in figure 4.2, the hyperplane offset from the origin, along $\mathbf{w}$, is determined by the parameter $\frac{b}{||\mathbf{w}||}$. Thus, the SVM selects $\mathbf{w}$ and b so that the margin (the distance between the hyperplanes) is maximally large.

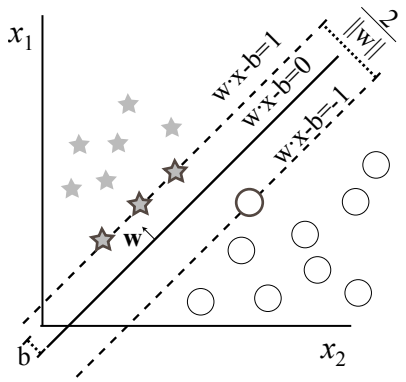

Figure 4.2: Illustration of a support vector machine (SVM) separating hyperplane.

The distance between the hyperplanes is $\frac{2}{||\mathbf{w}||}$, and thus maximizing the distance requires minimization of $||\mathbf{w}||$. In order to prevent data points from falling into the margin, the following constraints are also added: $\mathbf{w} \cdot x_i - b \geq 1$

for x_i of class 1, and $\mathbf{w} \cdot x_i - b \leq -1$ for x_i of class -1. This can be rewritten as:

$$c_i(\mathbf{w} \cdot x_i - b) \geq 1, \text{for all } 1 \leq i \leq n. \tag{4.3}$$

The parameters of the maximum-margin hyperplane, $\mathbf{w}$ and b are solved for in the following optimization problem:

$$\text{minimize } ||\mathbf{w}||, \text{ subject to } c_i(\mathbf{w} \cdot x_i - b) \geq 1, \text{ for all } 1 \leq i \leq n \tag{4.4}$$

This problem depends on $||\mathbf{w}||$ which involves a square root. Instead, the problem can be simplified to:

$$\text{minimize } \frac{1}{2}||\mathbf{w}||^2, \text{ subject to } c_i(\mathbf{w} \cdot x_i - b) \geq 1, \text{ for all } 1 \leq i \leq n \tag{4.5}$$

with retained $\mathbf{w}$ and b. The factor of $\frac{1}{2}$ is added for mathematical convenience.

Nonlinear separability is constructed using the kernel trick. The kernel trick involves replacing every dot-product in the SVM equations with a nonlinear kernel function to transform the (originally linearly non-separable) feature space into a higher-dimesional space where the categories are linearly separable by the maximum-margin hyperplane (Boser, Guyon, and Vapnik, 1992).

A variety of specialized algorithms have been developed for solving the SVM optimization problem (see e.g. Schölkopf and Smola, 2001). Throughout this thesis, the least-squares SVM approach proposed by Suykens et al., 2002, is used, as implemented in the Matlab SVM package LS-SVMlab developed by the group SCD/sista in the department ESAT at the KULeuven, Belgium (Suykens et al., 2002).

4.1.3 Classifier training and evaluation

The process of adapting the classifier parameters (e.g. solving the SVM optimization problem described above) to the given data is called *training*. During training, the classifier is supplied with data samples and corresponding categories. Caution is required during training in order to avoid classifier adaptation to noise (Duda, Hart, and Stork, 2000). Such overfitting yields outstanding results on the training dataset – at the cost of poor generalization performance (figure 4.3). Overfitting may be controlled for by applying the trained model to an independent validation data set to empirically evaluate the generalization performance of the classifier. Depending on the problem at hand, more data partitions may be required to minimize bias (such as during feature selection, described in section 4.1.4).

Generalization performance estimates are often obtained using cross-validation, particularly when the number of available samples is limited, according to one of the following schemes:

N-fold cross-validation: The data samples are partitioned into N parts. The classifier is trained on N-1 parts of the data, and the performance of the trained classifier is evaluated on the Nth part. The process is repeated for each of the N parts, after which a result is formed from the average across all N iterations. Leave-one-out cross-validation is the special case of $N = n$, where n is the number of available samples.

Hold-out validation: The available samples are randomly divided into two sets where each set is used as training and validation data respectively. By averaging over a number of partitions of the same size, a reliable estimate of the generalization performance can be obtained.

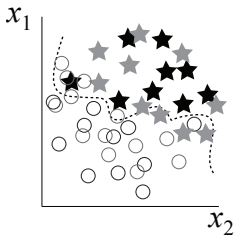

Figure 4.3: An example of an overtrained classifier (dotted line) which correctly classifies all instances of the training data (black) but makes mistakes on the validation data (grey).

4.1.4 Feature selection

Feature selection involves the identification of a subset of variables relevant for the given classification task (Guyon and Elisseeff, 2003). Given the typical number of original variables in contemporary measuring systems (such as fMRI scanners), acquired signals are likely to contain a drastic number of features which are unrelated to the categorization task. Feature selection is fundamentally important for effective classification, particularly for the following two reasons:

1. Mitigate the effect of the curse of dimensionality to improve classifier performance (Bellman, 1961).

   The curse of dimensionality refers to the exponential increase in volume associated with additional dimensions (figure 4.4). For example, 10 samples suffice to fully cover a 10 point one-dimensional interval, whereas only 1% of the space is represented in three dimensions. Thus, the larger the dimensionality, the more samples are required to sufficiently cover the probability space to construct a good classifier. Given the exceedingly high dimensionality of fMRI data (ranging from tens to hundreds of thousands of voxels) compared to the number of available samples (in the range of tens to hundreds), the curse of dimensionality is a highly acute problem and feature selection is fundamental for optimal classifier performance.

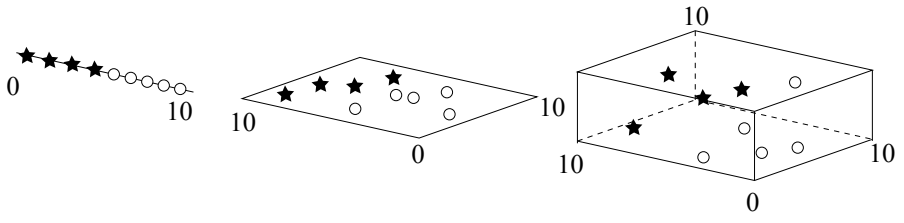

Figure 4.4: The curse of dimensionality: in one dimension (one feature), the ten observations (samples) cover the probability space to 100%. As the number of dimensions increases the probability space grows exponentially resulting in an increasingly poor coverage of the space (Bellman, 1961).

2. Improving model interpretability.

   The explicit identification of informative features provides additional information about the representation of the relevant patterns and thus facilitates data visualization and model interpretation. Feature selection can either focus on identifying variables which are useful to build a good predictor, or, in contrast, with the problem of identifying all potentially relevant variables, and each of the approaches provide different types of information (see the review articles by Kohavi and John, 1997 or Blum and Langley, 1997 for a discussion of usefulness vs. relevance). As detailed in section 4.3, improved model interpretability through voxel selection is of fundamental significance in brain mapping.

In addition, feature selection reduces classifier training times as well as measurement and storage requirements.

Approaches to feature selection can be roughly divided into two categories, namely *filter* and *wrapper* methods (Kohavi and John, 1997; Blum and Langley, 1997).

Filter feature selection utilizes an external measure, independent of the classifier, to estimate the relevance of each feature. Common methods include univariate variable ranking approaches where individual features are scored and selected according to some measuring criteria independent of the classifier. While filter methods are computationally fast, the obtained feature subsets are generally not optimal for the given classifier. Importantly, variables which are not informative individually may provide improved performance when jointly analyzed with other variables, and two variables which are not informative individually may be so when analyzed together (see also figure 4.5; Guyon and Elisseeff, 2003). Filter feature selection, especially of the univariate kind, is therefore a poor choice for multivariate analyses.

Wrapper methods, on the other hand, utilize the intended classifier directly to assess the relative usefulness of feature subsets (e.g. by classification scores for given subsets or using classifier weights). Elements of re-training the clas-

sifier with different feature subsets are typically involved. Notably, the training data set may require further partitioning to obtain reliable classifier generalization performance estimates during such classifier re-training. Wrapper methods produce feature subsets which are specifically tailored to high performance for the given classifier, albeit at a computational cost.

Various feature selection methods in the context of brain mapping are detailed further in section 4.3.

4.2 BOLD response classification

Along with the notation in the previous section, fMRI activity patterns (data samples) can be represented as points in a multidimensional space where the number of dimensions equals that of voxels (features) and the experimental conditions are the labels.

In the simplified situation of a two-voxel brain, each trial can be considered as a point in a plane corresponding to the BOLD magnitude measured in each voxel. The aim of MVPA is to categorize the samples by separating the points belonging to each of the condition (i.e. brain state) classes. As illustrated in figure 4.5, the method of doing so depends on the structure of the data; if the conditions are sufficiently different this can be done on a single voxel level (with conventional univariate statistics; figure 4.5A), but if the voxel distributions overlap, multiple voxels must be jointly analyzed in order to distinguish the conditions (figure 4.5B and C). By virtue of taking the information encoded over multiple signal elements into account, machine learning algorithms have the potential to detect significant signal changes where traditional univariate, voxel-by-voxel methods fail.

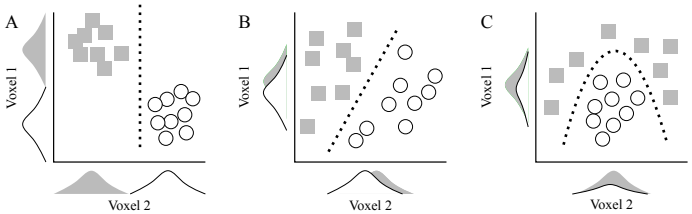

Figure 4.5: Two-voxel illustration of the multivoxel analysis approach, where dark squares and white circles represent two different experimental conditions. In **A**, the response distributions to the conditions (the Gaussian curves) are separable in each single voxel and a univariate statistical approach is capable of differentiating the conditions. In **B**, the two conditions can not be separated by each individual voxel due to the overlap of the distributions, and a univariate measure would fail in distinguishing the conditions. A linear decision boundary (dotted line) can, however, separate the conditions. Similarly, in **C**, the conditions can be separated but a nonlinear decision boundary is required.

Numerous studies have utilized the improved sensitivity of pattern recognition methods, including the decoding of speech (Formisano et al., 2008),

hidden intentions (Haynes et al., 2007), odor quality (Howard et al., 2009), visual stimuli – visible (Haxby et al., 2001; Cox and Savoy, 2003; Kamitani and Tong, 2005) as well as invisible (Haynes and Rees, 2005a) – to name a few.

Not only are multivariate methods more sensitive than univariate approaches, but the predictive power of classifier-based machine learning algorithms can be utilized to identify and distinguish the specific spatial activity patterns produced by single experimental conditions (see e.g. paper II). Numerous studies have shown the utility of such brain state classification, and applications include tracking of mental states over time (Polyn et al., 2005), lie detection (Davatzikos et al., 2005), biofeedback (Yoo et al., 2006; LaConte, Peltier, and Hu, 2007) and brain-computer interfacing (Sitaram, Caria, and Birbaumer, 2009).

MVPA typically follows the procedure shown in figure 4.6. After acquisition, the fMRI signals are preprocessed, partitioned into training and validation data sets and subject to voxel selection and classifier training. Voxel selection and classifier training are often iterated to assess the success of any voxel subset (particularly when a wrapper feature selection approach is utilized), and a map which indicates what brain regions containing differential brain responses due to the different conditions is obtained. The success of the classifier, in combination with the selected voxels, is evaluated using the validation data set. Finally, the obtained classifier can be applied to predict brain states in new fMRI data.

The specific considerations required for appropriate application of classifier-based MVPA analysis are outlined in the following sections.

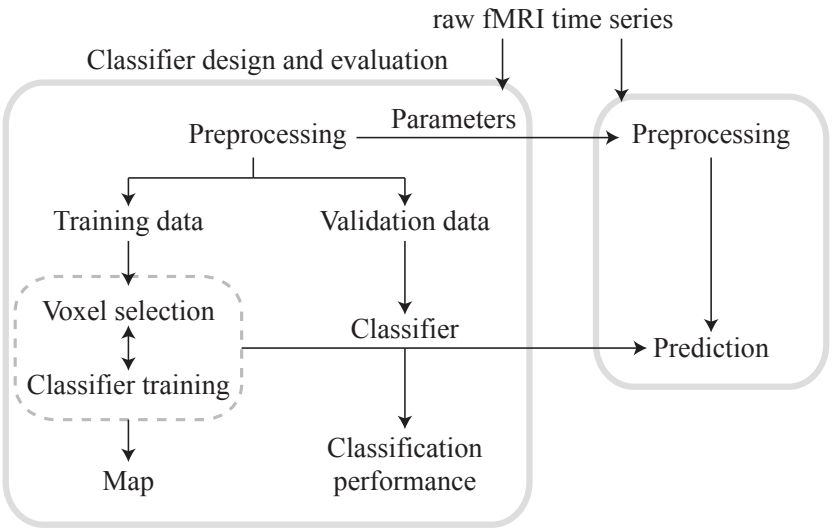

Figure 4.6: Generic multivoxel pattern analysis (MVPA) workflow.

4.2.1 Experimental paradigm

MVPA experimental paradigm considerations are similar to those in conventional approaches, and both block and event-related designs have been used with good results (see e.g. Burke et al., 2004; Beauchamp, LaConte, and Yasar, 2009 and papers I-IV). In particular, event-related designs have the benefit of yielding more independent trials, which, in turn, results in less contaminated estimations of the spatial pattern related to each condition. Although rapid event-related designs risk temporal overlap of hemodynamic responses, various techniques can be applied to reduce this effect (Beauchamp, LaConte, and Yasar, 2009).

4.2.2 Preprocessing

The same preprocessing steps as in conventional analysis are required, with the notable exception of spatial smoothing. If the conditions differ in terms of their fine-grained spatial activation patterns, spatial smoothing will reduce classifier performance (Kriegeskorte, Goebel, and Bandettini, 2006). Also, without smoothing, the spatial resolution provided by the fMRI scanner is preserved and small differences in location can be maximally resolved. Smoothing may, nonetheless, have a beneficial impact on classification performance (LaConte et al., 2003).

4.2.3 Condition response estimation

The continuous fMRI signal, consisting of a series of BOLD values for each voxel across the scanning time course, must be re-represented as single trial responses for subsequent analysis. In particular, the hemodynamic response delay must be accounted for (see section 1.3) such that each single trial label corresponds to the peak BOLD response due to the appropriate condition. A number of condition response representations have been proposed:

1. *Single-volume intensities*: The BOLD values in a single acquisition volume are taken to represent the conditions (Haynes and Rees, 2005a; Mourão-Miranda et al., 2005). A simple approach to compensating for hemodynamic delay is to shift the data labels an appropriate amount of acquisition time points (typically around 6 s).
2. *Volume-average intensities*: The average BOLD signal across a number of consecutive volumes (e.g. in block design studies; Kamitani and Tong, 2005; Mourão-Miranda et al., 2006, and papers II and IV). Typically the first few volumes are discarded. This approach has the added benefit of increased signal-to-noise ratio due to the averaging procedure.
3. *Single-trial GLM fitting*: Estimation of the single condition response based on the hemodynamic response function (e.g. De Martino et al., 2008 and paper I and III). Estimates of the condition responses are obtained by fitting

a GLM to each trial (Friston et al., 1998). At every voxel, the corresponding regressor coefficient (β) is taken to represent the trial response.

4.2.4 Data partitioning

In order to avoid classifier overfitting and biased prediction accuracies (see e.g. Kriegeskorte et al., 2009, for a review of this problem in functional brain imaging), care is required when partitioning the samples into training and validation data.

Potential dependencies between datasets must be carefully avoided; the inherent temporal sluggishness of the hemodynamic response producing temporal dependencies is of particular concern. Thus, any randomization of training and validation samples must be preceded by a single condition response estimate (see section 4.2.3) ensuring no temporal dependencies between samples. Another possibility is to select a temporally independent validation data set from samples collected towards the end of the scanning session.

4.2.5 Choice of classifier

Despite the theoretical superiority of nonlinear classifiers (i.e. nonlinear classifiers can differentiate linear data but not vice versa), linear classifiers dominate the MVPA literature (but see Davatzikos et al., 2005; Hanson, Matsuka, and Haxby, 2004; Polyn et al., 2005), partially since the improvement over linear classifiers is not conclusive (see e.g. Cox and Savoy, 2003). Moreover, a highly appealing advantage of linear classifiers is the direct relation between classifier weights and voxels, providing a means to understand which regions of the brain are multivariately informative (Mourão-Miranda et al., 2005; De Martino et al., 2008).

Classifiers employed for multivoxel pattern analysis of fMRI data range from various versions of linear discriminant analysis (Carlson, Schrater, and He, 2003; O'Toole et al., 2005; Haynes and Rees, 2005a; Haynes and Rees, 2005b; Kriegeskorte, Goebel, and Bandettini, 2006), correlation-based classification (Haxby et al., 2001; Spiridon and Kanwisher, 2002), artificial neural networks (ANNs; Hanson, Matsuka, and Haxby, 2004; Polyn et al., 2005) and Gaussian naïve Bayes (GNB) classifiers (Mitchell et al., 2004). Although there may be little practical difference between linear classifiers (Ku et al., 2008), SVMs dominate in recent MVPA studies (Cox and Savoy, 2003; Mitchell et al., 2004; Kamitani and Tong, 2005; Mourão-Miranda et al., 2005; Mourão-Miranda et al., 2006; LaConte et al., 2005; De Martino et al., 2008; Formisano et al., 2008; Staeren et al., 2009; Mourão-Miranda et al., 2009).

4.3 Voxel selection and brain mapping

Voxel selection is not only of critical importance in order to obtain classifiers with good generalization performance, but it also provides a means to spatially localize brain response patterns. A number of approaches, with and without explicit voxel selection, have been proposed for multivariate brain mapping as described in the following section.

4.3.1 Univariate filter selection and region-of-interest analysis

The feature selection and brain mapping problem may be resolved by region-of-interest (ROI) based methods where classifiers are applied to voxels in anatomically or functionally predefined areas (Cox and Savoy, 2003; Haynes and Rees, 2005a; Kamitani and Tong, 2005). ROI selection methods based on univariate functional ranking include estimates of activation magnitude due to any condition (activation-based voxel selection) or the ability to differentiate the conditions, as quantified by e.g. the standard GLM (discrimination-based voxel selection; Mitchell et al., 2004; Haynes and Rees, 2005a; Mourão-Miranda et al., 2006). Such univariate measures can also be used as an initial ranking scheme for improved speed or accuracy in subsequent multivariate voxel selection (see e.g. De Martino et al., 2008; Niiniskorpi, Björnsdotter Åberg, and Wessberg, 2009).

Coarse brain maps may be obtained by assessing the classification performance in a number of ROIs. However, univariate feature selection disregards any distributed brain activity effects and provides little additional information regarding the localization of brain response patterns compared to the univariate measure on its own.

4.3.2 Locally multivariate mapping

Locally multivariate mapping approaches integrate distributed brain responses, but are restricted to a limited neighborhood of adjacent voxels. Approaches targeting regions of fixed size and shape provide appealing flexibility and simplicity of implementation. In particular, the attractively simple and intuitive "searchlight" algorithm introduced by Kriegeskorte et. al (2006) has proved useful in numerous studies (Kriegeskorte et al., 2007; Haynes et al., 2007; Bode and Haynes, 2009; Clithero, Carter, and Huettel, 2008; Stokes et al., 2009; Pereira, Mitchell, and Botvinick, 2009). Whole-volume maps are produced by computing multivariate brain response measures across fixed-size (typically spherical) search volumes sequentially centered on each voxel in the brain volume. As opposed to region-of-interest approaches, the searchlight requires no a priori spatial hypotheses – with the drawback that an excessive number of information computations (one per voxel, that is, in the order of tens to hundreds of thousands or higher) is required. The Monte Carlo approximation proposed in paper I is a derivative of the searchlight which reduces the required number of information computations dras-

tically (Björnsdotter Åberg and Wessberg, 2009). Other locally multivariate, fixed-size search approaches have followed suit (e.g. particle-swarm mapping; Björnsdotter Åberg and Wessberg, 2009). Such fixed-size, fixed-shape methods rely on the assumption that response patterns are contained within the locality of the search volume, and may fail to detect discriminative patterns encoded across regions of different shape and size.

The evolutionary brain mapping approach presented in paper II, on the other hand, optimizes voxel cluster size, shape and location (Björnsdotter Åberg and Wessberg, 2008), and yields specific information regarding the spatial extent of differential fMRI response patterns. This property is highly useful in studies where the exact extent and location of a specific activity pattern is acute (such as in paper IV).

A collective benefit of these locally multivariate methods is that any arbitrary information measure (including nonlinear classifiers) can be used to detect brain response patterns.

4.3.3 Globally multivariate mapping

Globally multivariate methods jointly analyze voxels in spatially remote regions or across entire brain volumes, and are an appropriate choice when brain responses are expected to be widely distributed or include a number of separate brain regions.

Such methods include massively multivariate methods, where a classifier is applied to all voxels in an entire brain volume simultaneously and individual voxel contributions are estimated from classifier weights (Mourão-Miranda et al., 2005; LaConte et al., 2005; LaConte, Peltier, and Hu, 2007; Beauchamp, LaConte, and Yasar, 2009; Sato et al., 2009). In addition to being limited to linear classifiers providing a direct relationship between individual voxels and classifiers weights, massively multivariate approaches do not alleviate the curse of dimensionality and hence produces suboptimal brain state discrimination sensitivities (see section 4.1.4).

In contrast, recursive feature elimination (RFE; Hanson and Halchenko, 2008; De Martino et al., 2008; Formisano et al., 2008; Staeren et al., 2009) is a voxel selection approach which explicitly identifies maximally discriminative regions of arbitrary size, shape and location. Similar to massively multivariate methods, however, RFE requires linear classifier weight rankings for iterative voxel elimination (or, less appealingly, utilization of a univariate voxel ranking scheme).

4.4 Performance metrics

Two performance metrics are used in MVPA: condition prediction ability (i.e. pattern discrimination performance), and sensitivity of the classification scheme in detecting relevant voxels (i.e. pattern localization performance).

The pattern discrimination performance is usually expressed in terms of classification measures, such as proportion correctly labelled instances of the validation data. More refined receiver operating characteristic curve (ROC) analysis may also be used. The ROC is a plot of the classification sensitivity versus (1-specificity) for varying thresholds. The area under the curve (AUC), where a value of 1 corresponds to perfect classification (no true negatives or false positives) can then be used as a metric (see e.g. paper II and IV). Due to the limited number of samples available in fMRI studies, use of cross-validation is standard.

Pattern localization sensitivity can also be estimated using ROC analysis on simulated data where the true discriminative voxels are known (see e.g. De Martino et al., 2008; Björnsdotter, Rylander, and Wessberg, 2009 and paper I). In studies on real data, sensitivities are typically compared to alternative methods such as the GLM (as in, for example, paper II).

A number of variants for the problem of statistical significance testing for both data classification and the various types of multivoxel brain maps have been proposed (see e.g. Pereira, Mitchell, and Botvinick, 2009, for a summary). Although the statistical method of choice is highly dependent on the application and properties of the data, variants of nonparametric permutation testing is an appealing choice to test whether a group of voxels can classify the experimental conditions to a significant degree (Good, 2004; Golland et al., 2005). The reasoning is as follows: if the classification score is not related to information regarding the categories of the data (i.e. the score was obtained by chance), permuting the data labels should not affect the classification score (i.e. the labels are exchangeable under the null hypothesis). Repeated permutations are used to estimate the empirical cumulative distribution of the classifier error under the null hypothesis, from which a p-value for the true label classification score can be computed.

Group analysis maps reflecting relative regional information content (e.g. measured in proportion correctly classified samples) can be formed by averaging across results obtained on the individual level (see e.g paper I and III). Other formal group-level statistical testing methods include fixed and random effect models (Mourão-Miranda et al., 2005; LaConte et al., 2005; Mourão-Miranda et al., 2006; Wang et al., 2007).

4.5 Available software

A number of the machine learning algorithms described above are implemented in available software. Notably, a Python-based, cross-platform, and open source software toolbox called PyMVPA was recently released by the Department of Experimental Psychology, University of Magdeburg (www.pymvpa.org; Hanke et al., 2009). A similar toolbox for Matlab called the Multi-Voxel Pattern Analysis (MVPA) Toolbox is provided by the Center for the Study of Brain, Mind and Behavior, Princeton (www.csbmb.princeton.edu/mvpa/; Detre et al., 2005). Commercially avail-

able, BrainVoyager QX (brainvoyager.com) has just released a new version where ROI-based classification as well as the searchlight and RFE are implemented.

5. Proposed brain mapping techniques

As described in the previous chapters, pattern recognition approaches constitute an appealing complement to univariate methods and particularly for sensitive brain mapping.

More specifically, the searchlight algorithm (described in section 4.3.2) is an appealing multivariate mapping approach due to simplicity of implementation, as well as interpretation, and flexibility of information measure. The algorithm involves the sequential centering of a fixed-size sphere on each voxel to compute a locally multivariate information measure for that voxel (figure 5.2; Kriegeskorte, Goebel, and Bandettini, 2006). After scanning the entire volume, a map containing one such value per voxel is obtained. However, the excessive number of information computations (one per voxel, that is, in the order of tens to hundreds of thousands or higher) and corresponding time requirements restrict the practicality of the searchlight for whole-brain, multi-subject data – especially in high-field fMRI with improved spatial resolution.

The first proposed brain mapping method, therefore, is a Monte Carlo approximation of the searchlight designed for simple, fast whole brain mapping while still retaining the benefits of multivariate sensitivity over the univariate GLM.

Fixed-shape, fixed-size locally-multivariate methods are, however, restricted to voxel-wise approximative estimates of brain responses in the exact region defined by the search sphere, and may miss patterns distributed across regions of different sizes and shapes. Given a typical brain volume of $64 \times 64 \times 25$ voxels and the combinatorial explosion of possible voxel subsets, an exhaustive search to identify the exact size and shape of a maximally informative region is not feasible. Optimization methods, such as recursive feature elimination (De Martino et al., 2008), have generally resorted to using voxel ranking schemes, either by univariate means or linear classifier weights in order to iteratively discard voxels with low scores. A wrapper-based, classifier-independent voxel optimization approach would therefore be useful.

Evolutionary algorithms (EA, also known as genetic algorithms, GA) have recently received much attention for robustly producing promising results in complex optimization tasks in a wide variety of medical and biological fields (e.g. cancer prediction; Li et al., 2004; Wahde and Szallasi, 2006, EEG single trial analysis; Åberg and Wessberg, 2007, pap smear diagnosis; Marinakis, Dounias, and Jantzen, 2009, nutrition; Kemsley et al., 2007, identification of honeybees; Lavine and Vora, 2005 and numerous bioinformatics studies; e.g. Li et al., 2008; Wu, 2008; Taneda, 2008; Thachuk et al., 2009). The second proposed machine learning brain mapping method is, therefore, a wrapper feature optimization scheme based on evolutionary algorithms (Holland, 1975), that identifies specific regions containing representational brain response patterns.

5.1 Monte Carlo brain mapping (paper I)

Monte Carlo methods approach problems by observing properties obeyed by randomly generated instances (Sobol, 1994), and have been successfully utilized in various fields where the system under investigation is complex and exact computations are not possible or feasible (including e.g. medicine, finance and engineering; Zaidi and Sgouros, 2002; McLeish, 2005; Amar, 2006). Generally, the Monte Carlo method can be described by the following procedure:

1. Specify a domain of variables.
2. Sample variables randomly from the domain.
3. Compute a measure using the variables.
4. Combine the individual measures into a final result.

The algorithm is iterative in nature, and more iterations result in an improved approximation of the true (exhaustive) result. See Sobol, 1994, for an in-depth description of the mathematical theory behind Monte Carlo sampling.

5.1.1 Implementation

Pseudocode describing the proposed Monte Carlo brain mapping method is presented in figure 5.1. Similar to the searchlight algorithm (figure 5.2; Kriegeskorte, Goebel, and Bandettini, 2006), the brain volume is partitioned into voxel clusters of fixed size and shape which are evaluated in terms of information content (e.g using a classifier to differentiate brain response patterns). As illustrated in figure 5.3, however, one iteration of the algorithm consists of the brain volume being randomly divided into a number of clusters (typically, but not necessarily, in the approximate shape of a sphere) such that each voxel is included in *one* (and one only) cluster and the information measure is computed for and assigned to each such cluster. A robust multivariate information map reflecting the *mean* contribution per voxel is subsequently computed by forming an average across the information computed in all constellations in which the voxels took part (as opposed to the searchlight were each voxel is assigned the one value computed when the sphere was centered on that voxel).

Thus, where the searchlight requires as many information computations as there are voxels in the volume in an exhaustive fashion, the sparse Monte Carlo method reduces the required number to the total number of voxels in the volume divided by the number of voxels in the search sphere multiplied by the number of iterations. Notably, larger searchlight spheres yield increased computational requirements (since the number of computations are unchanged but each computation uses more voxels), whereas the contrary holds true for the Monte Carlo approach – with increased size, fewer search spheres are required to cover the brain volume and thus the number of information computations is reduced. As per the nature of Monte Carlo schemes, iterating the algorithm

```
BEGIN
For (number of iterations);
    While (not all voxels are selected);
        Select a random hitherto unselected voxel;
        Select all other hitherto unselected voxels within the specified radius;
        Compute an information measure on these voxels;
    End While
End For
Compute the mean information per voxel across iterations;
END
```

Figure 5.1: Pseudocode for Monte Carlo fMRI brain mapping.

to perform as many information computations as the searchlight yields an exhaustive measure with the difference that the information measure reflects an average contribution.

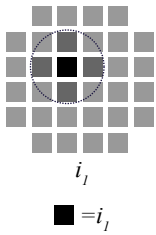

Figure 5.2: Visualization of the searchlight algorithm (Kriegeskorte et al., 2006). A search sphere (circle) is centered on each single voxel and the corresponding measure is assigned to that voxel.

5.1.2 Method evaluation

The searchlight, the Monte Carlo method, and an exhaustive search using the Monte Carlo averaging procedure (denoted Monte Carlo* or MC*) were evaluated and compared on simulated data with realistically modeled discriminative regions varying in size and degree of condition-discriminative information content (i.e. contrast-to-noise ratio, CNR). The data were simulated according to a block design, and the discriminative voxels were assigned to either of two populations ($\text{condition}_1 > \text{condition}_2$; $\text{condition}_2 > \text{condition}_1$) with random spatial distribution within the discriminative regions. The single-trial condition responses were estimated by single-trial GLM fitting (see section 4.2.3).

The brain mapping sensitivities were measured in area under the ROC-curve (AUC), and a standard GLM was computed on the data for comparison.

Both linear and nonlinear (radial basis function kernel, RBF) SVM classifiers were evaluated on various search sphere volumes.

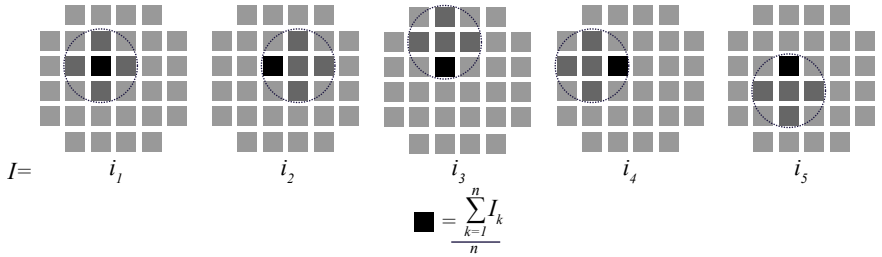

Figure 5.3: Visualization of the Monte Carlo fMRI brain mapping computation for one voxel. For each iteration k, the search volume (circle) is centered on a random voxel (black) and the voxels within the search volume (black and dark grey) are used to compute an information measure i_k (e.g. condition classification accuracy) for that voxel. In subsequent iterations, the voxel is included in different constellations with other neighboring voxels, and a final information measure is computed as the average across all n iterations. Here, $n = 5$ iterations corresponds to an exhaustive search.

The following list summarizes the main results.

1. **More sensitive than the searchlight and the GLM**

   Both the searchlight and Monte Carlo approach outperformed the univariate GLM which achieved a discriminative voxel detection sensitivity of 0.503 (see table 5.1). Interestingly, the Monte Carlo method was significantly more sensitive across search sphere volumes and classifier kernels at a mean AUC of 0.873 (range 0.840-0.899) than the searchlight (0.826; range 0.767-0.857; paired t-test, $p<0.05$; table 5.1). Also, the exhaustive Monte Carlo* method consistently achieved higher mapping sensitivities than the Monte Carlo approximation and the searchlight at an AUC of 0.904 (range 0.877-0.918; table 5.1). Thus, assigning each voxel the average information content across all constellations in which it has been included improves the mapping sensitivity and explains the improved performance of the Monte Carlo approximation compared to the searchlight.

   Figure 5.4 shows the corresponding maps obtained with a search sphere volume of 0.5% and the RBF kernel. Despite a 66% reduction in computer resources the Monte Carlo map is strikingly similar to both the searchlight and exhaustive Monte Carlo* maps. Moreover, as is clearly exemplified in the large discriminative cluster with $CNR = 0.8$ (shown in blue) and in stark contrast to the searchlight map, voxels with similar discriminative information content obtain homogenous values as a result of the information-averaging smoothing effect. Map values obtained with the searchlight algorithm, on the other hand, are deceptively dependent on the regional context such that any voxel value is highly sensitive to the number of discriminative voxels within the search sphere – resulting in substantially higher

Table 5.1: Table of brain mapping sensitivities for the searchlight (Kriegeskorte et al., 2006), Monte Carlo (MC) method and exhaustive Monte Carlo search (MC*) on the simulated data, measured in area under the receiver operating curve (AUC). The best performance for each approach is denoted by ⋆. The number of information computations for the exhaustive searches (searchlight and MC*) was 28 502. The search sphere volume is expressed in percentage of total brain volume. Lin-SVM: support vector machine with a linear kernel, RBF-SVM: support vector machine with a radial basis function kernel; MC nr. computations, Number of information computations required using the Monte Carlo approximation approach.

	Lin-SVM			RBF-SVM		
Search sphere volume (%)	0.1	0.5	1.0	0.1	0.5	1.0
Searchlight	0.767	0.835	0.842	0.809	0.857⋆	0.848
MC	0.840	0.883	0.880	0.850	0.899⋆	0.886
MC*	0.877	0.915	0.911	0.896	0.918⋆	0.909
MC nr. computations	27 778	11 643	6 745	27 778	9 705	6 745
Reduction (%)	2.54	59.15	76.33	2.54	65.95	76.33

values for voxels in the center of the discriminative cluster than in the surrounding voxels (despite the same information content).

2. **Substantially faster than the searchlight**

   For the kernel (RBF) and search sphere volume (0.5% of the total volume) with which all method achieved the highest mapping sensitivities, the number of required information computations (number of times the classifier was trained and tested) was dramatically reduced from 28 502 for the searchlight to 9 705 for the Monte Carlo method – corresponding to a reduction of 66% for no loss in sensitivity (see table 5.1). The more iterations allowed for the Monte Carlo method, the higher the mapping sensitivity, and, consequently, the computational load. As exemplified in figure 5.5A (for a search sphere volume of 0.5%), the Monte Carlo method mapping sensitivity improved dramatically during the initial information computations (up to in the order of 5 000) and approximates the exhaustive Monte Carlo* search performance at relatively few computations. For the RBF kernel and search sphere volume of 0.5% where the exhaustive Monte Carlo* search obtained an AUC of 0.918, for example, the Monte Carlo method reduced the number of required information computations by 75% for a 2.8% reduction in mapping sensitivity (Figure 5.5B).

3. **Impact of search sphere volume**

   All methods achieved highest sensitivities with small to medium sized search spheres, and the best results were obtained at a volume of 0.5% of the total brain volume (see table 5.1 and figure 5.6A; but note that search

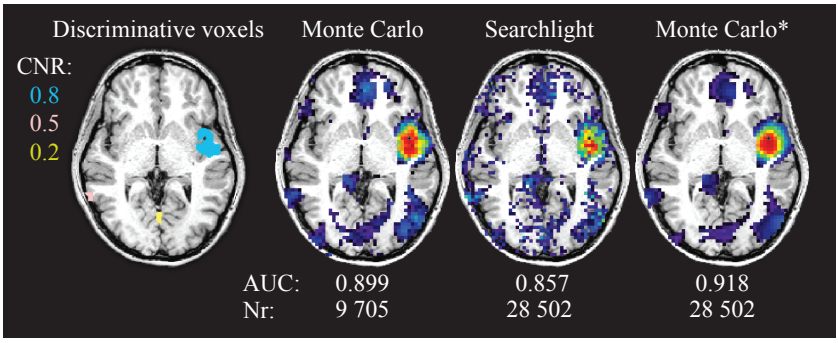

Figure 5.4: Comparison of the maps produced using the Monte Carlo, searchlight and exhaustive Monte Carlo* algorithms on the simulated data with a search sphere volume of 0.5% and the RBF SVM kernel. The position and CNR of the simulated discriminative voxels are shown in the leftmost panel. The Monte Carlo map is strikingly similar to the exhaustive maps despite a reduction in number of required information computations from 28 502 to 9 705. The maps are thresholded to show voxels with values above 0.5. CNR: contrast-to-noise ratio; Nr: Number of information computations (training and testing of the classifier); AUC: area under the receiver operating characteristic curve.

volumes larger than 1% were not used in conjunction with the searchlight and Monte Carlo* search due to excessive time requirements). There was no substantial difference in sensitivity between the small to medium search spheres (0.5 to 2%), despite a dramatic decrease in number of required performance computations up to a volume of 2% (figure 5.6B). All methods, however, obtained better results on the large clusters using larger search spheres (volume of 1% vs. 0.1% and 0.5% vs. 0.1%, $p<0.05$, paired t-test). Across CNRs there was no significant trend.

4. **Nonlinear classifiers are more sensitive**

   Across all discriminative regions and search sphere volumes, the mapping sensitivities obtained with all methods in conjunction with the RBF kernel were significantly higher than those of the linear kernel maps ($p<0.05$, paired t-test). The difference was not substantial, however – the RBF kernel improved the Monte Carlo mapping sensitivity from 0.883 to 0.899 on the combined cluster analysis with a search sphere volume of 0.5%, and the corresponding figures for the exhaustive Monte Carlo* search was 0.915 and 0.918 (table 5.1).

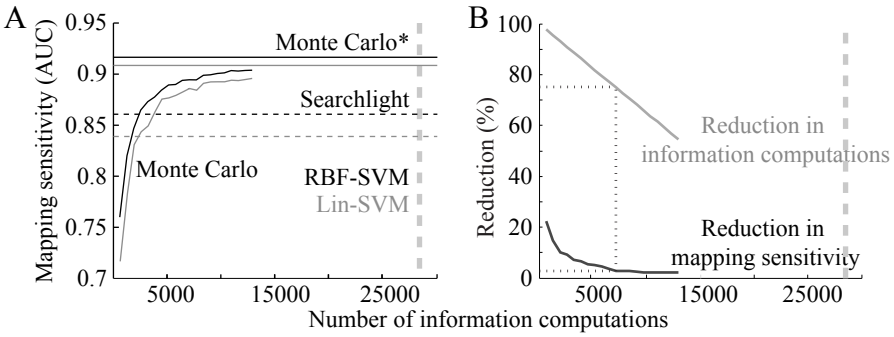

Figure 5.5: **A**) Monte Carlo algorithm brain mapping sensitivity (measured in area under the receiver operating characteristic curve; AUC) as a function of the number of information computations (number of times the classifier is trained and tested), for the RBF (black) and linear (grey) SVM kernel. The searchlight and exhaustive Monte Carlo* search sensitivities are plotted as dotted and straight lines, respectively. **B**) The reduction in number of information computations (grey) and mapping sensitivity (black) of the Monte Carlo method compared to the exhaustive Monte Carlo* search, for the RBF kernel. A reduction in number of information computations by 75% corresponds to a decrease in sensitivity of 2.8%, for example, as indicated by the dotted lines. A search sphere volume of 0.5% of the total brain volume was used in this example, and the vertical thick grey lines represent the number of information computations required for the searchlight and exhaustive Monte Carlo* search.

5.2 Evolutionary brain mapping (paper II)

An evolutionary algorithm is a machine learning optimization method inspired by biological evolution which utilizes operators such as recombination, selection, reproduction and mutation (Holland, 1975). Candidate solutions to the optimization problem are represented by individuals in the population, and an objective function is used to evaluate the fitness of each individual. The individuals continuously evolve to produce a solution which approaches the optimum throughout repeated application of the operators (as illustrated in figure 5.7; and see Reeves and Rowe, 2002, for an in-depth discussion on the theoretical framework of EAs).

5.2.1 Implementation

Feature selection in the traditional sense attempts to minimize the number of redundant features (Guyon and Elisseeff, 2003). A standard feature selection approach would, therefore, extract voxels which are sparse and distributed, yielding maps that are difficult to interpret from a physiological point of view (see e.g. Åberg, Löken, and Wessberg, 2009). The task of the evolutionary algorithm was, instead, to identify spatially coherent voxel clusters of unrestricted size or shape, where maximal brain response differentiation can be obtained using a classifier.

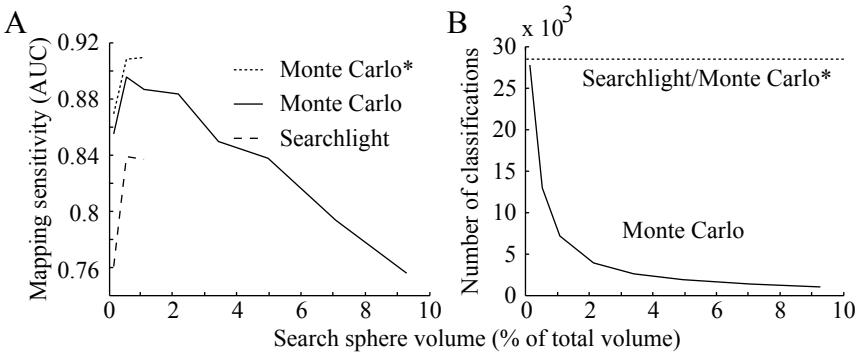

Figure 5.6: The effect of search sphere volume (expressed in percentage of total brain volume) on **A**) the mapping sensitivity (measured in area under the receiver operating characteristic curve; AUC) and **B**) number of required information computations for all methods and the linear support vector machine kernel. There was a dramatic decrease in number of required performance computations up to a volume of 2%. Note that for the searchlight and Monte Carlo* search only the three smallest search sphere volumes were investigated due to excessive time requirements for larger volumes. Similar results were obtained for the support vector machine with a radial basis function kernel.

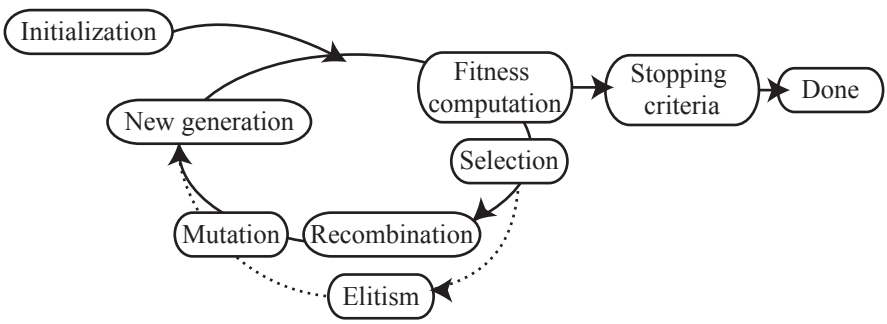

Figure 5.7: Schematic of a general evolutionary algorithm.

Attempts were initially made to identify more than one cluster at any given time, but the increase in search space complexity yielded unstable results with the standard evolutionary algorithm (but see the discussion in section 6 on suggestions for multiple cluster identification). Instead, the algorithm was designed to optimize one single cluster, and, where required, it was iterated to identify more clusters with varying locations. Below follows details on the implementation of the algorithm, and pseudocode is presented in figure 5.8.

Representation: Each individual in the population corresponds to one voxel cluster. Each voxel is represented by its corresponding index in a list the size of the total number of voxels in the brain volume.

Initialization: The population of individuals is initialized in a stochastic fashion, where, for each individual, one seed voxel is randomly selected. The

```
BEGIN
Initialize population;
While (termination criteria not met);
      For (each individual);
            Apply mutation operations;
                  1. Add n_a random voxels;
                  2. Remove n_r random voxels;
                  3. Substitute n_s random voxels;
      End For
      Select parents;
      Reproduce;
      generation = generation + 1;
End While
END
```

Figure 5.8: Pseudocode for evolutionary brain mapping.

voxel cluster is then constructed by the addition of random voxels which neighbor the seed voxel, or, subsequently, any voxel already in the cluster.

Mutation operations: The following mutation operations are implemented in the algorithm: the addition of a number of voxels, the deletion of a number of voxels, and the substitution of a voxel with another voxel. All voxel additions and substitutions are performed on neighboring voxels, that is, voxels within the 26 voxel cube surrounding any voxel already contained in the cluster. Also, deletions or substitutions resulting in voxels disconnecting from the clusters are disallowed. The frequency of mutation is regulated by a constant mutation rate parameter for each mutation operation. In addition, a voxel cluster in the population is occasionally substituted for a new, randomly generated cluster to add fresh genetic material and aid in escaping potential local maxima.

Selection and reproduction: A standard tournament scheme is used for parent selection. In order to retain a variety of the genetic material and maintain searches in widespread regions of the brain, the proportion of parents to discarded individuals is set high. Since all individuals in the population represent different locations and crossover thus would destroy the spatial integrity of the voxel clusters, reproduction is asexual and the new generation is formed by cloning the parents.

Objective function: The objective is to maximize experimental condition classification success computed using a classifier. Any classifier can be applied (including nonlinear schemes; see the discussion in section 4.1.2 on classifiers). To ensure high generalization capability, the algorithm is supplied with three datasets. The first is used in classifier training (training data, 35% of the total volumes) while the second is used for fitness estimation (testing data, 45%). The third dataset is exclusively used with the already trained and optimized classifier and voxel cluster (validation data, 20%). Any fitness measure indicative of classification performance can be used, and since all

relevant voxels are of interest, no penalty for the number of voxels (common in standard feature selection implementations) is involved.

Termination: The algorithm is run for either a predetermined maximum number of generations or until a cluster yielding testing data classification rates above a given threshold (the fitness threshold) is obtained. Since the algorithm is prone to overfitting when allowed to run the full course, the cluster with the best result on the mean of the training and testing data performance is identified and subject to validation classification.

5.2.2 Method evaluation

The performance of the evolutionary brain mapping algorithm was evaluated on data from two authentic fMRI studies using a linear SVM. First, nine subjects tapped their fingers to their thumbs, and, second, six subjects were brushed on their forearm or thigh such that two datasets (arm brushing/thigh brushing/rest) were obtained. All data were preprocessed in a standard fashion, with the exception of 6 mm spatial smoothing of the brushing data, and the single-trial condition responses were estimated as volume-average intensities. The evolutionary algorithm was repeatedly applied to obtain ten voxel clusters per dataset. All reported classification results refer to the classification performance obtained using trained and optimized classifiers on the validation data, measured in the area under the ROC curve (AUC). A standard GLM was also applied for comparison.

The following list summarizes the main results.

1. **Highly accurate prediction of brain states in new subjects**

   A leave-one-out cross-validation was performed on the finger-tapping dataset, where the algorithm was applied to training and testing data sets containing eight subjects. Combinations of the identified clusters were then used on the validation dataset consisting only of the ninth subject. The performance as a function of the (unique) accumulated voxels from the highest through lowest ranked cluster is presented in figure 5.10A. The pattern recognition approach was highly accurate in classifying the brain state of the unknown subject, with a subject-average AUC of over 0.9. The latent (unthresholded) SVM classifier output estimate for subject one is shown in figure 5.9.

2. **EA outperforms GLM feature selection**

   The same SVM leave-one-out cross-validation scheme was applied to the corresponding number of voxels ranked according to the GLM t-map (obtained on the training subjects). As seen in figure 5.10A, the cluster algorithm outperformed the GLM ranking method which obtained classification scores of less than an AUC of 0.85. A closer inspection of the selected vox-

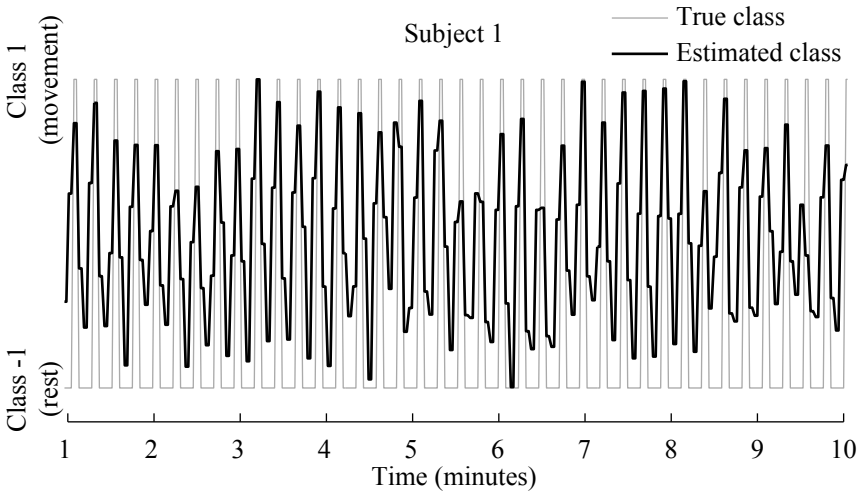

Figure 5.9: SVM classifier predictions (black) for on a subject performing an alternating finger movement task (grey). The SVM was trained on eight other subjects.

els with respective methods shows that, although the general areas are similar and the overlap is large, the cluster algorithm generated voxel subsets slightly more medial and posterior than the GLM t-map ranking method (see figure 5.10B). Also, the latter included voxels in the supplementary motor area (SMA) at an early stage, whereas all of the 10 evolutionary clusters remained in the primary motor and somatosensory area.

3. **More sensitive pattern discrimination than the GLM**

On the brushing dataset, the insular cortex, known to be activated by gentle touch (Olausson et al., 2002; Björnsdotter et al., 2009) was first extracted in each individual subject as a region of interest (ROI). Subsequently, the algorithm was applied to the forearm/rest and thigh/rest datasets to identify clusters within the ROI. All voxels contained in clusters with an AUC larger than 0.5 were combined and a hold-out validation (using a random 80% of the data for training and 20% for validation, 10 repetitions) was performed on each data set. The same validation approach was applied to all voxels within the ROI. Table 5.2 shows the resulting classification AUCs as well as the maximum ROI GLM t-value and the corresponding number of voxels included for classification. In six data sets (subjects 2, 5 and 6 for arm/rest and subjects 1, 4 and 5 for thigh/rest) no significantly activated voxels (false discovery rate; FDR$<$0.05) were found in the ROI according to the GLM. Using the whole ROI for classification, only three data sets achieved significant classification scores, whereas the cluster-based classification achieved significant results in all cases (p<0.05; permutation test with 1000 iterations).

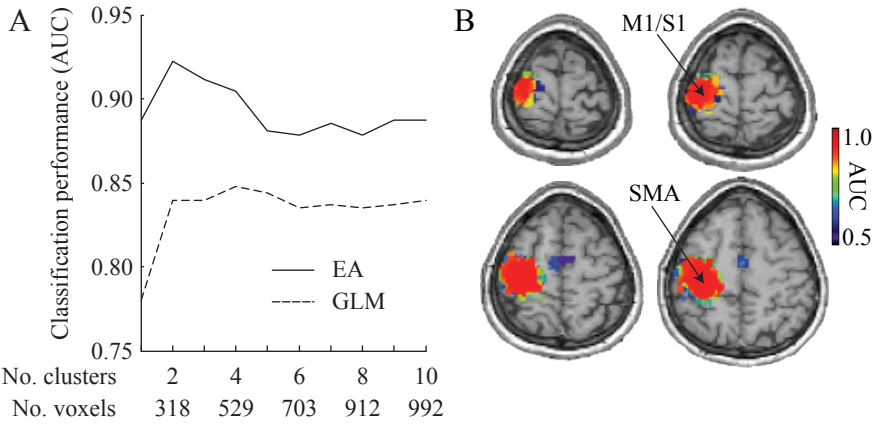

Figure 5.10: **A**) Comparison of classification performances using evolutionary algorithm (EA) and univariate GLM ranking feature selection as a function of the number of clusters/voxels included for classification. **B**) The detected clusters with corresponding classification performance. The contralateral primary motor and somatosensory cortices (MI/SI) produce higher classification results than any other area. The supplementary motor area (SMA) and ipsilateral MI/SI are also detected, but yield lower classification scores.

Table 5.2: Summary of ROI classification results for arm/rest and thigh/rest brushing in six subjects. The ROI classification parameters, including the maximum GLM t-value, classification results for whole ROI classification (measured in area under the ROC-curve, AUC), the number of voxels contained in the entire ROI, the AUC for the cluster analysis and the number of voxels used by the cluster algorithm (voxel subset size). Stars denote significant scores (GLM T: FDR<0.05, AUC: permutation test, p<0.05). S, subject; ROI, region of interest.

Arm	Max -log(p)	ROI AUC	ROI size	Cluster AUC	Cluster size
1	4.438*	0.6533*	753	0.6587*	351
2	2.455	0.5991	904	0.6916*	313
3	4.812*	0.592	787	0.604*	252
4	3.055*	0.5996*	811	0.6178*	304
5	1.355	0.5511	703	0.6649*	202
6	2.113	0.5733	1004	0.6502*	356
Thigh	Max -log(p)	ROI AUC	ROI size	Cluster AUC	Cluster size
1	3.098	0.5173	753	0.6716*	219
2	4.087*	0.6467*	904	0.6773*	195
3	7.660*	0.5698	787	0.632*	312
4	2.190	0.5818	811	0.6178*	405
5	2.391	0.5516	703	0.6036*	128
6	4.326*	0.5578	1004	0.6244*	345

6. Discussion

Machine learning based multivariate localization of representational patterns provides a highly appealing complement to univariate brain mapping, particularly in terms of increased sensitivity to experimental condition differences. The first objective of the thesis was to develop and empirically evaluate such pattern recognition methods which could be effectively used for neuroimaging purposes. Two complementary multivariate brain mapping methods are proposed. While both were demonstrated to be more sensitive than univariate schemes in detecting differential fMRI activity patterns, the Monte Carlo approach produces whole-brain maps whereas the evolutionary algorithm yields tailored regions showing the spatial extent of differential fMRI responses. Additional strengths and weakness of the algorithms, as well as implications, applications and future directions are discussed below.

6.1 Improved sensitivity

The proposed method evaluation demonstrated one appealing benefit of multivariate compared to univariate (GLM) analysis, namely that of increased sensitivity in detecting differential brain responses.

The simulated data modeled a situation with two spatially distributed and intermixed populations of voxels belonging to different conditions (see paper I for details). Here, the superiority of the multivariate approach over the GLM can be explained by the integration of weak univariate condition differences heterogenous with respect to the direction (increase or decrease) of the BOLD response. Notably, spatial smoothing has a destructive effect on such fine-grained response patterns and would reduce the sensitivity of any differentiation attempts (Kriegeskorte, Goebel, and Bandettini, 2006).

Surprisingly, the Monte Carlo approach achieved higher discriminative voxel detection sensitivities also than the searchlight algorithm. The improved sensitivity can be attributed to two effects. First, the searchlight assigned relatively higher map values to voxels with more discriminative neighbors (despite equal information content), while the corresponding Monte Carlo maps produced more homogenous values for voxels of equal information content (as a result of averaging across multiple voxel clusters). Second, the typical variation in performance between classification attempts, a problem which is of particular concern when few fMRI volumes are available, is reduced due to the averaging across numerous performance evaluations.

On the real data, a main effect (e.g. tactile brush stimulation vs. rest in the insular cortex) was investigated, and, presumably, the patterns detected by the evolutionary algorithm reflected a homogenous BOLD increase during stimulus application (particularly since the data was spatially smoothed). The higher sensitivity can therefore be attributed to the integration across multiple weak univariate differences with the same sign, and not necessarily to fine-

grained spatial patterns. At improved signal-to-noise ratios (SNR), such as in higher field scanners, the GLM is therefore likely to be equally efficient at detecting the corresponding differential activity. Indeed, the GLM did identify significant univariate BOLD increases during brushing in some individuals.

Based on the above observations, multivariate mapping is likely to be more sensitive than univariate analyses in two situations: 1. when the differential brain responses are represented by fine-grained spatial patterns, and 2. in low SNR situations when individual univariate differences are not detectable.

Improved sensitivity in lower SNR situations is particularly useful when the number of samples is restricted. These often occur in physiological studies as well as clinical situations, where considerations for subject or patient discomfort limits scanning times.

Advances in high-field imaging warrants improved SNRs as well as increasingly high spatial resolution and corresponding finer-scale neuronal activity representations (down to the submillimeter range; Uğurbil, Toth, and Kim, 2003; Uğurbil et al., 2003; Harel et al., 2006). Whereas univariate and multivariate methods benefit from improved SNR alike, a higher number of dimensions yields an increasingly severe problem of multiple comparisons for the GLM which may be detrimental (Kriegeskorte and Bandettini, 2007).

6.2 Locally vs. globally multivariate mapping

Whereas the Monte Carlo and searchlight approaches are restricted to activity patterns which can be captured in the local neighborhood of a fixed-size search sphere, the evolutionary algorithm tailors the neighborhood to any size and shape (and location) that optimally represents the brain response pattern. Nonetheless, both methods are limited to jointly considering adjacently located voxels and may fail to detect widely distributed or spatially separate activity patterns. In situations where subtle, spatially distant coactivations are expected (e.g. auditory tasks; Formisano et al., 2008), globally multivariate methods, such as recursive feature elimination (De Martino et al., 2008) are more appropriate. Or, as was demonstrated in the single-trial classification task in paper II, a number of individually defined regions can be combined to improve classification rates.

Nevertheless, both the Monte Carlo and evolutionary algorithm can potentially be extended to allow for simultaneous analysis of distant regions. The Monte Carlo method could, for example, easily incorporate more than one stochastically selected search sphere at any given time. Also, the search volume size and shape can be randomized to capture voxel activity patterns whose spatial extend is not known. In addition, a number of niching techniques allowing for simultaneous optimization of multiple voxel clusters have been proposed for evolutionary algorithms (Dunwei, Fengping, and Shifan, 2002; Zhang et al., 2009). Specifically, successful multi-cluster maps have been obtained using an extension of the evolutionary brain mapping method

with elements from memetic algorithms (unpublished data; Zhu, Ong, and Dash, 2007b; Zhu, Ong, and Dash, 2007a).

6.3 Voxel selection vs. scanning

The two proposed methods approach the problem of brain mapping from two entirely different perspectives: the Monte Carlo method scans the entire brain such that each voxel obtains a value, whereas the evolutionary algorithm specifically selects only those voxels which maximize the activity pattern classification. For similar computational costs the evolutionary algorithm tends to produce poor overall detection results (unpublished results on simulated data). Instead, it specifically singles out a majority, if not all, of the voxels in the most discriminative region and yields higher classification results.

The evolutionary algorithm hence lends itself towards studies where either maximal brain state classification is required, or the size, shape and location of a specific brain response pattern is desired. The algorithm is, for example, highly useful in paper IV, where the research question is well-defined and limited to localizing a single, maximal brain response. Providing maps where each voxel is evaluated, the Monte Carlo algorithm, in contrary, is suited for situations where whole-brain explorations are required. An example of this is found in paper III, where differential projection patterns due to different types of tactile stimulation are examined.

Despite various tricks to avoid local minima, the evolutionary algorithm is notoriously sensitive to initialization factors and, unless any of the cluster in the initial population by chance is located near the global maxima it may not be detected. The Monte Carlo algorithm, on the other hand, easily provides a map of the entire brain. Preliminary results on combining these corresponding advantages by initializing the evolutionary (as well as the recently developed memetic) algorithm on clusters obtained from a rough Monte Carlo map are promising. As an alternative, initialization clusters could be derived from massively multivariate linear classifier weights (Mourão-Miranda et al., 2005; LaConte et al., 2005) in the same fashion as recursive feature elimination (De Martino et al., 2008).

6.4 Flexibility of performance measure

A particular aspiration of the thesis was to develop generic pattern recognition methods directly applicable in a variety of neuroimaging studies. In this light, a clear benefit of the proposed methods is that neither is dependent on a specific type of performance measure. Whereas recursive feature elimination (De Martino et al., 2008) and massively multivariate maps (Mourão-Miranda et al., 2005; LaConte et al., 2005) require the use of linear classifiers, nonlinear classifiers can be directly incorporated in the Monte Carlo as well as the evolutionary approach. Nonlinear classifiers have not been extensively studied in

the context of fMRI classification (with the exception of Mørch et al., 1997; Davatzikos et al., 2005; Hanson, Matsuka, and Haxby, 2004; Polyn et al., 2005), partially since the improvement over linear classifiers is not conclusive (see e.g. Cox and Savoy, 2003). In paper I, it was, however found that nonlinear SVM mapping was significantly more sensitive in detecting discriminative voxels than a linear SVM. Discouraging results with nonlinear classifiers are likely to be due to limited number of training samples (and high number of voxels) which restricts the construction of complicated relationships with the voxels (Mørch et al., 1997). FMRI studies explicitly designed for pattern recognition analysis with fast, event-related paradigms yielding more samples in combination with effective voxel selection methods may increase the utility of nonlinear classifiers and possibly reveal patterns not distinguishable with linear classifiers.

6.5 Computational requirements

Whereas the Monte Carlo algorithm is attractively simple and only two parameters require specification, the evolutionary approach is complex to implement and requires a substantial number of parameters to be empirically specified. Such implementation issues aside, the computational requirements for both algorithms are well within a practical range. On a standard PC (3.20GHz processor, 3GB RAM), the Monte Carlo algorithm took in the order of 14 minutes to complete a 34 519 voxel whole brain map in Matlab (The Mathworks, Massachusetts, USA; see paper I in the appendix), and the corresponding time requirements for the evolutionary algorithm was 20 minutes.

While more computer intensive than univariate measures (which take seconds or less to compute), the methods are substantially faster than the searchlight; the Monte Carlo takes 75% of the time required for the searchlight (for no loss in sensitivity). In addition, parallelization of neither the Monte Carlo nor the evolutionary algorithm (Stender, 1993) should not pose technical difficulties if required.

6.6 Applications

Machine learning based classifiers are appealing general purpose tools and, in addition to revealing differential brain activity patterns for basic physiology research purposes as was demonstrated in this thesis, can be highly advantageous in virtually any signal processing setting. Although the proposed methods were primarily developed for brain mapping, they (and particularly the evolutionary algorithm) have great potential for use in tracking of brain states, both offline (as was demonstrated in paper II; see also e.g. Polyn et al., 2005) and for real-time fMRI analysis (LaConte, Peltier, and Hu, 2007; deCharms, 2008). Real-time analysis may provide novel approaches to dynamic exploration of brain activity in adaptive experiment designs, which

could, for example, be advantageous when patients or subjects have trouble adhering to the specified experimental paradigm. Such analysis can also be utilized in brain-machine interfaces that allow patients to communicate non-verbally with a computer or to control robotic devices (Birbaumer and Cohen, 2007), and for biofeedback, where patients can learn to control brain activity (for pain relief or other purposes; Weiskopf et al., 2003; Weiskopf et al., 2004; Yoo et al., 2006). In addition, although controversial and in a heavily restrictive setting, real-time classifier based analysis has shown positive results in lie detection (Davatzikos et al., 2005).

Real-time classifier based fMRI analysis should not, however, be confused with "mind reading". Although single brain states can be decoded using these methods, there are severe limitations; both technical, in terms of the heavily restricted and categorized states which a classifier can learn to detect, and due to inherent properties of the fMRI signal. In addition to being related to metabolic demands rather than direct neural activity, BOLD responses are inherently sluggish (with peak response delays of approximately six seconds) and fMRI is therefore poorly suited for instantaneous detection of transient thoughts. Similarly, suggestions to use fMRI as an objective measure of brain states in legal settings (Thompson, 2005) or to reveal terrorists (Wild, 2005) appear fallacious.

Classifiers that can be trained on a number of individuals for subsequent use in predicting which category a new subject's brain patterns belongs to, as was demonstrated in paper II, could be useful in disease diagnosis. Machine learning algorithms have indeed shown promising results in clinical diagnosis, both with structural (Ecker et al., 2009; Koutsouleris et al., 2009) and functional brain imaging (Marquand et al., 2008; Fu et al., 2008); or on other types of data entirely (Björnsdotter Åberg and Wessberg, 2008).

The proposed machine learning algorithms are not limited to fMRI data. In particular, similar methods have been used to predict various brain states from electroencephalography (EEG; Åberg and Wessberg, 2007), electrocorticograms (Yanagisawa et al., 2009), and direct neural recordings in monkeys (Wessberg et al., 2000), rats (Laubach, Wessberg, and Nicolelis, 2000) and humans (Salarian et al., 2007). The methods for volumetric pattern discrimination and localization developed in this thesis are particularly suited for multivariate analysis of differential patterns in cortical source images of EEG (Grave de Peralta Menendez, Murray, and Gonzalez Andino, 2004).

Finally, machine learning algorithms are not limited to predicting brain states from neuroimaging data. In fact, machine learning algorithms have recently been successfully applied to predict muscle responses from fMRI data, thus providing a direct link between a single muscle contraction to the BOLD response (Ganesh et al., 2008). Similarly, response pattern representations have been used to correlate monkey neural recordings to human fMRI (Kriegeskorte et al., 2008). Methods for such intermodality correlations appear particularly promising for analysis of recently developed concurrent fMRI and EEG recording techniques (Ritter and Villringer, 2006).

Part II:
Central processing of CT mediated gentle touch

7. Background

The vast majority of neuroscience and neurology textbooks describe touch as signaled exclusively by large myelinated (Aβ) fibers. Over sixty years ago, however, it was established that mammal skin is also equipped with an additional afferent system which responds to mechanical stimulation – that of thin, unmyelinated C tactile (or CT) fibers (Zotterman, 1939; Bessou et al., 1971; Kumazawa and Perl, 1977b; Shea and Perl, 1985). C tactile afferents were more recently identified also in humans (Nordin, 1990). They are highly sensitive to soft, slow skin deformations such as a caress (Vallbo et al., 1993), and their firing rates correlate intimately with subjective ratings of perceived pleasantness (Löken et al., 2009). Whereas the Aβ system provides highly acute information regarding discriminative properties of touch, the C tactile system appears to signal affective aspects of the tactile experience.

Contrary to Aβ mediated touch, C tactile brain physiology and function are not well understood. This part of the thesis, therefore, is devoted to exploring the multivariate brain response patterns to C tactile mediated touch in an attempt to further elucidate the central organization of these afferents.

7.1 Cutaneous sensory neurons

The human skin is innervated by a variety of sensory neurons which respond to mechanical stimulation and transmit tactile information through the spinal cord to the brain (Gardner and Martin, 2000). Such mechanoreceptive afferents can be classified into three broad categories based on conduction velocity. Aβ fibers are rapidly conducting, large diameter neurons with a thick layer of myelin, Aδ afferents have a smaller diameter and thinner myelin, and C fibers are thinnest and unmyelinated.

Four types of Aβ afferents innervate the glabrous skin (Johnson, 2001). Slowly adapting fibers with irregular discharge rates (type I, SAI) are implied in high-resolution discrimination (Knibestöl, 1973; Knibestöl, 1975; Johansson and Vallbo, 1980; Vallbo et al., 1984; Phillips, Johansson, and Johnson, 1992; Edin et al., 1995), whereas SA neurons with regular discharge rates (type II; SAII) respond to skin stretch and indentation (Vallbo et al., 1995). Rapidly adapting (RA) type I fibers react to low-frequency flutter vibration, such as moving objects (Mountcastle et al., 1967; Connor et al., 1990; Friedman et al., 2002; Bensmaia et al., 2005), whereas RAII units respond to high-frequency vibration (Bolanowski and Zwislocki, 1984a; Bolanowski and Zwislocki, 1984b; Bolanowski, 1984; Vallbo and Johansson, 1984).

In hairy skin, SAI, SAII and RAII units have been identified, and, additionally, RA hair and field units (Vallbo et al., 1995).

Aβ afferents transmit acute information regarding discriminative properties of the stimulus, such as location, shapes and textures (Vallbo et al., 1995; Johnson, 2001; Hsiao and Bensmaia, 2007). Body parts important for tactile

discrimination, such as the finger tips, are densely innervated by $A\beta$ afferents (Johansson and Vallbo, 1979a).

Aδ and C-fibers, in contrast, signal thermoreception, nociception and chemoreception in addition to C tactile low-threshold mechanoreception (Burgess and Perl, 1967; Bessou and Perl, 1969; Bessou et al., 1971).

7.2 Properties of C tactile afferents

Low-threshold mechanoreceptive C afferents have been described in a number of species, including rats (Lynn and Carpenter, 1982; Leem, Willis, and Chung, 1993), guinea pigs (Sugiura, Terui, and Hosoya, 1989), mice (Liu et al., 2007), cats (Douglas and Ritchie, 1957; Iggo, 1959; Iggo, 1960; Bessou et al., 1971; Iggo and Kornhuber, 1977) and primates (Kumazawa and Perl, 1977b; Kumazawa and Perl, 1977a; Kumazawa and Perl, 1978). In humans, C tactile afferents have been identified in the hairy skin of different body parts, including the face and hairy skin of the extremities, but not in glabrous skin (i.e. the palms or the soles of the feet; Nordin, 1990; Vallbo, Olausson, and Wessberg, 1999; Edin, 2001; Löken, Wessberg, and Olausson, 2007).

C tactile afferents respond vigorously to slow, light mechanical stimulation, such as a caress, and poorly to rapid, vibratory skin deformations (Zotterman, 1939; Douglas and Ritchie, 1957; Iggo and Muir, 1969; Bessou et al., 1971; Kumazawa and Perl, 1977b; Nordin, 1990; Vallbo et al., 1993; Vallbo, Olausson, and Wessberg, 1999; Edin, 2001). In addition, they have been reported to respond to skin stretch (Nordin, 1990; Vallbo, Olausson, and Wessberg, 1999; Edin, 2001).

C tactile fibers are susceptible to fatigue (decrease in response due to repeated stimulation; Iggo, 1960; Bessou et al., 1971; Iggo and Kornhuber, 1977; Lynn and Carpenter, 1982; Wiklund Fernström, 2004); recovery times are highly variable can range from 30 seconds to 30 minutes (Iggo, 1960; Wiklund Fernström, 2004).

The receptive fields associated with C tactile afferents in humans are described to have a number (on the order of one to nine) of small, distributed, non-uniformly responsive spots (Wessberg et al., 2003).

7.3 Spinal cord organization

Classically, mechanoreceptive afferents are assumed to follow the posterior column-medial lemniscus pathway from the dorsal roots through the spinal cord to the brain (Kaas, 2007). The posterior columns primarily contain $A\beta$ fibers which project rostrally towards the brain and terminate in the posterior column nuclei of the medulla. The second-order fibers, in turn, cross over to the opposite side of the spinal cord and continue through the medial lemniscus, after which they ascend to a lateral region of the ventral posterolateral

nucleus (VPL) of the thalamus. Ultimately, third-order neurons terminate in the somatosensory cortices (figure 7.1).

Small-diameter ($A\delta$ and C) fibers, on the other hand, connect through lamina I in the dorsal horn of the spinal cord (figure 7.1). Modality-specific classes of neurons receiving input from functionally distinct groups of small-diameter afferents have been identified in lamina I (Han, Zhang, and Craig, 1998; Craig, Krout, and Andrew, 2001), including neurons responding to gentle mechanical stimuli (Kumazawa and Perl, 1977b; Light and Willcockson, 1999; Sugiura, Lee, and Perl, 1986). Specifically, it has been shown that lamina I neurons responding to noxious and temperature stimuli project somatotopically through the spinothalamic pathway in the ventral horn of the spinal cord through the ventral posterior nucleus of the thalamus (possibly through a distinct nucleus in humans termed the posterior ventromedial thalamus; VMpo; Craig et al., 1994; Dostrovsky and Craig, 1996; for an alternative view see Willis et al., 2002).

It has been speculated that C tactile afferents are organized in a fashion similar to that of the pain and temperature mediating thin-fiber system (Olausson et al., 2002).

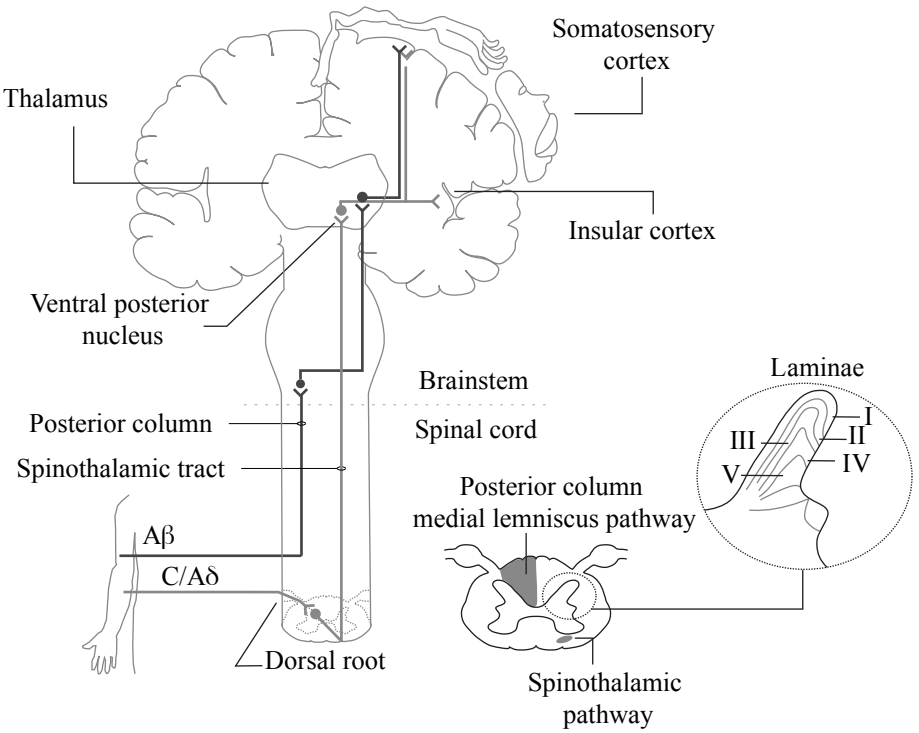

Figure 7.1: Simplified schematic of mechanoreceptive sensory neuron projection paths. The laminae of the dorsal horn of the spinal cord are also shown.

7.4 Central projections

Thick, myelinated $A\beta$ afferents which project along the posterior column-medial lemniscus pathway terminate in the primary somatosensory cortex (SI), which is found in the postcentral gyrus in the anterior parietal lobe of the human brain (figure 7.2A; Krubitzer and Kaas, 1987). All three architectonic brain areas contained in SI, namely 3 (further divided into *a* and *b*), 1 and 2 (as defined by Brodmann in 1909; see e.g. Kandel, Schwartz, and Jessell, 2000), represent differential projection patterns from the contralateral side of the body (see e.g. Hsiao and Bensmaia, 2007 for details on the functional differences between these areas). All regions are somatotopically organized. SI neurons respond to a wide range of tactile stimuli, and are particularly implied in acute discriminative aspects of touch such as texture recognition and high-resolution localization of tactile input (Hsiao and Bensmaia, 2007).

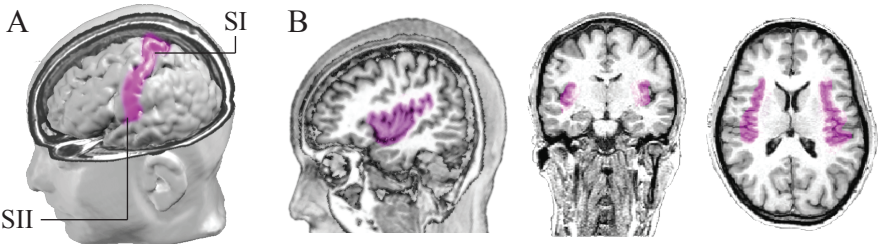

Figure 7.2: Location of the **A**) primary (SI) and secondary (SII) somatosensory cortices and **B**) insular cortex.

The secondary somatosensory cortex (SII), found in the parietal cortex (figure 7.2A), is implicated in object form perception (Murray and Mishkin, 1984; Disbrow et al., 2001; Haggard, 2006). SII receives inputs from the primary somatosensory cortex as well as directly from the thalamus (Krubitzer and Kaas, 1987; Qi, Preuss, and Kaas, 2007), and also appears to be somatotopically organized (Ruben et al., 2001).

Thin-fiber mediated sensations projecting through the spinothalamic tract have been shown to activate the insular cortex (Coghill et al., 1999; Hofbauer et al., 2001; Craig et al., 2000). Specifically, functional imaging of innocuous cooling and painful stimuli has revealed a somatotopic organization of the posterior portion of the insular cortex (Hua et al., 2005; Brooks et al., 2005; Henderson, Gandevia, and Macefield, 2007). The insular cortex (often abbreviated as simply 'the insula') is found bilaterally within the lateral sulcus between the temporal lobe and inferior parietal cortex (figure 7.2B). Parts of the temporal, parietal, and frontal lobes form lids (opercula) which cover the insular cortex.

The insular cortex is described as a site of emotional processing and interoceptive awareness (Craig, 2003a; Craig, 2009). The posterior region of the insular cortex in primates has been proposed to contain a sensory representation of thin-fiber activity, integral in the maintenance of well-being as

an afferent homeostatic network (Craig, 2002; Craig, 2008; Olausson et al., 2008a).

An fMRI study of two unique subjects (GL and IW) with the rare acute neuronopathy syndrome (Sterman, Schaumburg, and Asbury, 1980), who lack Aβ fibers but have intact C fibers, revealed that C tactile stimulation activates the contralateral posterior insular cortex (Olausson et al., 2002; Olausson et al., 2008b), consistent with other thin-fiber mediated sensations such as pain and temperature. Notably, no activity in the somatosensory cortices was observed, indicating that the C tactile - insular system projects differently than he Aβ - somatosensory network. However, the question remains whether the corresponding projection patterns are present also in healthy subjects.

The consistency between C tactile and other small-diameter fibers, particularly those signaling temperature and pain, suggest that C tactile fibers are an integral part of the same thin-fiber afferent network. Hence, posterior insular cortex processing of C tactile stimuli is hypothesized to be somatotopically organized.

7.5 Functional role of CT afferents

Due to the dynamic properties of C tactile afferents (such as slow conduction velocity, susceptibility to fatigue, poor response to vibration, and insensitivity to rapid stimulus change; Bessou et al., 1971), it is considered unlikely that C tactile fibers are involved in acute discriminative aspects of touch (Vallbo, Olausson, and Wessberg, 1999). Instead, the C tactile fiber system was initially proposed to signal tickle (Zotterman, 1939; Bessou et al., 1971; Nordin, 1990). A recent study has, however, revealed that C tactile afferent firing rates correspond highly with subjective perceptions of the pleasantness of the tactile sensation (Löken et al., 2009), and the prevalent view today is that C tactile afferents signal affective touch involved in affiliative behavior (Vallbo, Olausson, and Wessberg, 1999; Olausson et al., 2002; Wessberg et al., 2003; McGlone et al., 2007; Olausson et al., 2008a).

C tactile afferents have also been implied as an integral part of the proposed thin-fiber homeostatic network for maintenance of physical well-being (Craig, 2003a; Craig, 2008; Craig, 2009).

8. Aims of the study

As described in the previous chapters, a number of observations support the hypothesis that C tactile afferents signal light touch in a network which is differentiated from, but parallel to, the $A\beta$ system in humans. Specifically, positive C tactile BOLD responses have been identified in the insular, but not somatosensory, cortical areas of neuronopathy patients (Olausson et al., 2002; Olausson et al., 2008b), indicating that C tactile and $A\beta$ afferents project along separate pathways. Whereas $A\beta$ fibers follow the posterior column medial lemniscus pathway, C tactile afferents are hypothesized to project along the pain and temperature fiber spinothalamic tract, which terminates in a somatotopic fashion in the posterior insular cortex.

The differential response patterns in the insular cortex due to C tactile and $A\beta$ activation remain to be demonstrated in healthy humans, as does the hypothesized somatotopic organization of C tactile projections in the insular cortex. The aims of the study were therefore to:

1. Explore differential brain patterns in response to C tactile and Aβ stimulation in healthy subjects.
2. Investigate whether C tactile afferent activation patterns in the posterior insular cortex are somatotopically organized.

9. Summary of methods

9.1 Subjects and ethics

The study was performed according to the Declaration of Helsinki with approval of the Ethics Committee of the Göteborg University, and informed consent was obtained from all subjects. Handedness was checked with a modified inventory (Varney and Benton, 1975).

In paper III, two healthy participants were included. In paper IV, six neurologically intact volunteers were studied in addition to one subject (GL, age 56, right-handed, female) with sensory neuronopathy syndrome (Sterman, Schaumburg, and Asbury, 1980). At the age of 31, GL suffered permanent specific loss of large-diameter myelinated afferents, leaving unmyelinated and small-diameter myelinated afferents intact (Forget and Lamarre, 1995). Motor nerve conduction and electromyography findings are within the healthy range, and thresholds for temperature and pain detection are largely normal (Olausson et al., 2002; Olausson et al., 2008b). GL routinely denies any ability to identify or localize touch below the level of the nose (Forget and Lamarre, 1995). In a forced choice task she did, however, demonstrate the perception of gentle touch in the hairy but not in the glabrous skin (lacking C tactile afferents), and she failed to detect vibratory stimuli (which poorly excite C tactile afferents) in both types of skin (Olausson et al., 2008b). In a four-alternative forced choice procedure, she identified 72% of soft brush stimuli to the correct extremity (at chance level of 25%). Healthy subjects, in contrast, detect gentle touch as well as vibration without fail in both glabrous and hairy skin, and can localize point indentation on the hairy skin to an accuracy in the range of two centimeters (Norrsell and Olausson, 1994).

9.2 Stimuli

Light mechanical stimulation, known to vigorously activate C-tactile afferents (see section 7.2), was manually delivered using a soft artist's goat hair brush (width: 2.3 - 7 cm, indentation force: 0.5 - 0.7 N, distance 3 - 3.5 cm, velocity: 0.9 to 7.5 cm s^{-1}).

In study IV, vibration (50 Hz), which preferentially activates Aβ afferents (both rapidly and slowly adapting), was also used. The stimulus was applied with a hand-held device consisting of a rectangular piece (40 x 12 x 7 mm) of balsa wood connected to a piezoelectric element (Piezo Systems, Inc., Cambridge, Massachusetts, USA). It should be noted that, although the preferred stimulus for C tactile afferents is slow stroking, responses can be elicited also by other types of stimuli such as vibration (unpublished data). While $A\beta$ fibers are highly activated by all types of touch, including brushing and vibration, C tactile afferents are poorly activated by vibration. Differential brain activa-

tion patterns between brushing and vibration are, therefore, likely to reflect C tactile projections.

In paper III, all stimuli were applied to the right side of the body whereas in paper IV the stimulation was applied to the left side.

9.3 Experimental paradigm

A block-designed paradigm was used throughout the study, and the stimuli were applied manually according to timing cues from the scanner. All subjects were instructed to focus on the stimulus throughout the fMRI scanning session.

In paper III, six-volume blocks of vibratory stimulation, brushing stimulation or rest, each lasting for 21 s were acquired. Two functional scans were obtained during vibration and two during brushing in a balanced order between subjects.

In paper IV, three-volume blocks of forearm brushing, thigh brushing or rest, each lasting 10.5 s were alternated in a pseudo-random order with equal numbers of each of the three conditions. The condition order remained fixed throughout each scan and across participants, and six functional scans were obtained per subject.

9.3.1 fMRI data acquisition

A 1.5 T fMRI scanner (healthy subjects: Philips Intera; GL: Siemens Sonata) was used to collect whole brain anatomical scans using a high-resolution T1-weighted anatomical protocol. Functional scans were acquired using a BOLD (blood oxygenation level dependent) protocol and a T2*-weighted gradient-echo, echo-planar imaging (EPI) sequence (healthy subjects: thickness 6 mm, in-plane resolution 3.6 $\times$ 3.6 mm; GL: thickness 4 mm, in-plane resolution 4 $\times$ 4 mm). The scanning planes were oriented parallel to the line between the anterior and posterior commissure and covered the brain from the top of the cortex to the base of the cerebellum.

9.4 Preprocessing

Standard preprocessing steps were applied to the data, and single-trial condition responses were estimated by either GLM fitting (paper III) or volume-average intensities (paper IV; see section 4.2.3).

For paper IV, the posterior contralateral (left) insula was isolated (Naidich et al., 2004). An initial evolutionary mapping was performed in each subject to identify the region of the insular cortex which responded maximally to brushing of the forearm and thigh combined (compared to rest). In all subjects, this region was located to the posterior portion of the insula, consistent with the activations found in previous studies of the insular cortex in relation to

C tactile activation (Olausson et al., 2002; Olausson et al., 2008b). This area was identified as a region of interest (ROI) to which all further analysis was restricted.

9.5 Analysis

All multivariate maps were computed on individual subjects, and a linear support vector machine was used for classification.

Paper III: Whole brain main effect GLM analysis was performed using BrainVoyager QX in order to detect regions with a significant univariate BOLD response to any of the tactile stimuli. In addition, the Monte Carlo algorithm (see section 5.1; search sphere radius of 6.6 mm) was applied to the whole brain to identify locally-multivariate differential activation patterns between brushing and vibration. Due to the reported susceptibility of C tactile afferents to fatigue, only the first two volumes of each stimulus block were analyzed. Maps were produced in each individual, and a group map was subsequently formed by averaging across the subject-level maps. Finally, in order to resolve whether any identified multivariate differences in brain responses between brushing and vibration were due to variations in fine-grained spatial patterns or differential BOLD response magnitudes, an event-related average of the BOLD temporal response for both conditions was computed using BrainVoyager QX.

Paper IV: The evolutionary clustering scheme (see section 5.2) was applied to the forearm/rest and thigh/rest datasets separately within the posterior insula ROI to explore potential somatotopic response patterns in the insular cortex. A ten-fold cross-validation of both forearm/rest and thigh/rest brushing data was performed in all the identified regions, and significance levels of the resulting classification scores were established using permutation testing.

10. Summary of results

10.1 Differential C tactile and Aβ activation patterns in the insula (paper III)

The GLM analysis demonstrated significant univariate brain activations in response to the tactile stimuli compared to rest in the expected somatosensory cortices (SI and SII), bilateral insular cortex and a number of additional areas (table 10.1).

The Monte Carlo method found differential activation patterns due to soft brush stroking and vibration in a number of regions (table 10.1). Specifically, a region in the left insular cortex with approximate location near regions previously identified as activated by soft brushing (Olausson et al., 2002; Olausson et al., 2008b), was identified and analyzed further (figure 10.1).

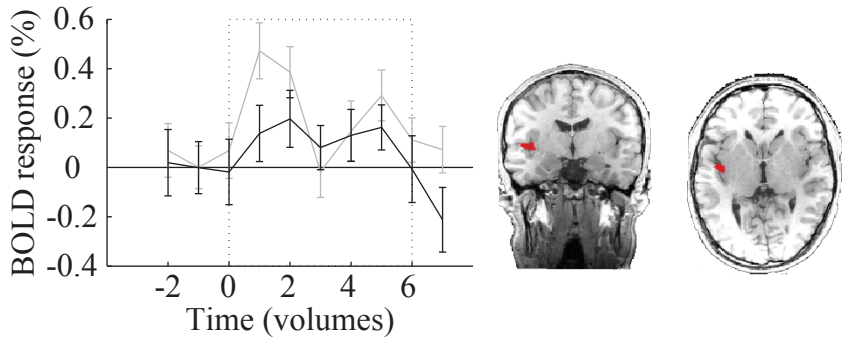

Figure 10.1: Event-related average percentage BOLD change in response to the brushing (grey) and vibration (black) tactile stimuli in the left insular cortex region (contralateral to the stimuli) where differential multivariate response patterns between brushing and vibration was found (shown in red). The dotted line indicates when the stimulus was applied.

Voxels in this region were also significantly univariately activated by brushing and vibration combined according to the GLM (peak t value of 5.22, $p<0.05$), as well as by the GLM contrast brushing vs. rest ($t=4.25$, $p<0.05$) and, although less so, vibration vs. rest ($t=3.12$, $p<0.05$). No significant differentiation was, however, observed between brushing and vibration (max t-value of 2.56, $p>0.05$).

As illustrated by the event-related averaging in figure 10.1, the differential spatial patterns between C tactile and Aβ stimulation in this region of the insular cortex (contralateral to the stimuli) was the result of a relatively larger BOLD increase during the brushing condition than during the vibration condition. The event-related averaging revealed a maximal BOLD response to brushing during the first two to three stimulation volumes, agreeing with the reported fatigue in C tactile afferents (Iggo, 1960;

Bessou et al., 1971; Iggo and Kornhuber, 1977; Lynn and Carpenter, 1982; Wiklund Fernström, 2004).

Also, the vibration condition shows some BOLD response increase during the stimulation, although with a peak magnitude of a third of that of brushing. Although C tactile afferents are clearly preferentially activated by gentle types of stimulation and poorly by vibration, the insular BOLD response due to vibration could nonetheless be the result of a partial C tactile activation.

Table 10.1: Peak group results (t values for the GLM analysis and group average classification scores for the multivoxel analysis) and cluster centroid Talairach coordinates for whole brain activations due to stimulation on the left thigh. The insular region of interest which is analyzed further is marked with a *. BA, Brodmann area; GLM, General linear model; Vib, vibration; contra, contralateral to the stimulated thigh; ipsi, ipsilateral to the stimulated thigh.

Region	Main tactile effect (GLM)	Vib vs. brush (Monte Carlo)
S1 contra	8.49 (-16.0, -43.0, 69.0)	
S2 contra	13.12 (-55.0, -28.0, 27.0)	
	6.60 (-58.0, 2.1, 6.5)	
IC contra	6.92 (-43.0, -7.8, 15.0)	0.61 (-25.0, -2.4, 23.0)
		0.60 (-39.0, -10.0, 1.0)*
S1 ipsi	-7.01 (2.6, -38.0, 63.0)	
S2 ipsi	6.22 (56.0, -19.0, 21.0)	
	6.78 (52.0, -28.0, 30.0)	
	5.95 (53.0, -17.0, 32.0)	
IC ipsi	5.46 (46.0, -0.058, 4.0)	
Other	BA 3, 6, 7, 19	BA 5, 6, 9, 10, 40
	21, 39	

10.2 Somatotopic organization of C tactile response patterns in the insula (paper IV)

The evolutionary mapping approach identified brain regions where the forearm and thigh response patterns were highly significantly differentiated in GL as well as for the healthy volunteers. Forearm and thigh tactile stimulation were found to project to distinctly separate locations in GL, with a substantial euclidean distance between cluster centroids of 8.9 mm (figure 10.2). The forearm cluster centroid was located at MNI (X, Y, Z) coordinates (-34, -10, 4), and the thigh cluster was found at (-34, -18, 0). The distance between clusters was thus maximal in the anterior-posterior (Y) plane at 8 mm, whereas the location differences in the remaining planes were either small or non-existent (X: 0 mm, Z: 4 mm). Validating the pattern observed in GL, the insular responses in the healthy subjects were also arranged in a clear somatotopic fashion. The difference in location was significant only in the Y-plane (anterior-posterior; two-tailed paired t-test, $p<0.05$). The subject mean euclidean distance between the cluster centroids equaled that of GL at 9.3 mm (range 6.6-12 mm).

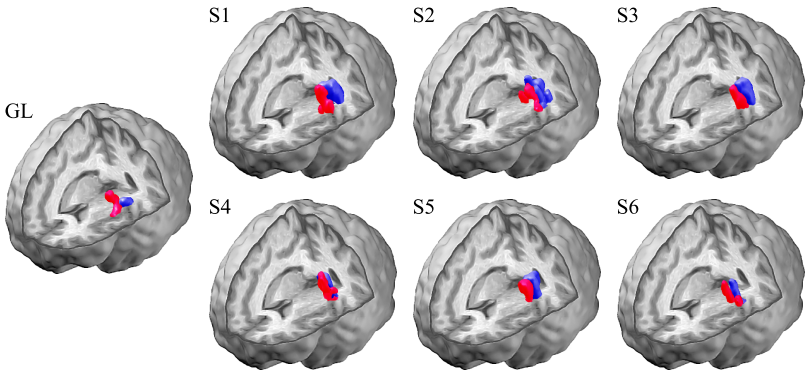

Figure 10.2: The voxel clusters reflecting forearm (red) and thigh (blue) BOLD activation patterns in the neuronopathy syndrome patient GL and six healthy subjects, reflecting the projection of C tactile afferents. There was a significant difference between forearm and thigh cluster centroid location in the Y-plane only (two-tailed paired t-test, $p<0.05$).

The ten-fold cross-validation confirmed that the anterior region contained a localized multivariate BOLD response to forearm brushing and the posterior region to thigh brushing (see figure 10.3).

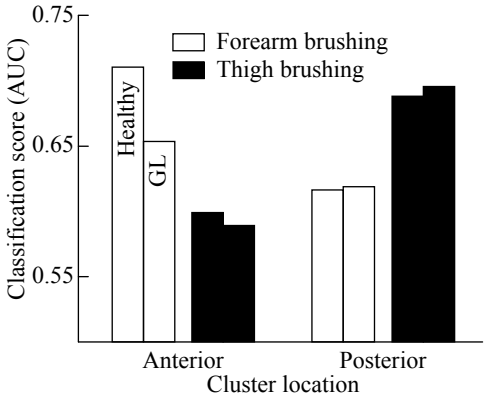

Figure 10.3: The classification scores (measured by area under the receiver operating characteristic curve, AUC) of the forearm (white) and thigh (black) brushing volumes on the posterior insular voxels obtained using the evolutionary mapping algorithm (as seen in figure 10.2). The healthy subject (left bar) forearm brushing volumes were significantly more separable in the forearm cluster than in the thigh cluster, and vice versa for the thigh brushing volumes (two-tailed paired t-test, $p<0.05$). Also, the forearm brushing scores were significantly higher than those of thigh brushing in the anterior cluster, and vice versa (two-tailed paired t-test, $p<0.05$). The neuronopathy patient data (GL; right bar) follows the healthy subject trend. These results verify the forearm - anterior and thigh - posterior pattern demonstrated by the evolutionary algorithm.

11. Discussion

Using multivariate machine learning methods, these studies have demonstrated two distinctive brain response patterns due to gentle tactile stimulation activating C tactile afferents. First, a whole-brain analysis revealed that brushing (to which both C tactile and $A\beta$ afferents respond) and vibration (which primarily activates Aβ fibers) response patterns differ in the posterior insular cortex. Second, C tactile responses to brushing in this portion of the cortex were found to be organized in a somatotopic fashion. These results have a number of implications, as discussed below.

11.1 Afferent activation

Any tactile stimulation activates a range of mechanoreceptors which transmit a complex pattern of inputs to the brain. Deducing how fMRI BOLD responses correlate to a single type of afferent is far from trivial. Moreover, cutaneous afferents rarely respond exclusively to a particular type of stimulation, and although C tactile afferents preferentially respond to soft brush stimuli some degree of activation to other types of stimuli can also be expected (Nordin, 1990; Vallbo, Olausson, and Wessberg, 1999). Also, myelinated ($A\beta$) low-threshold mechanoreceptors respond vigorously to gentle tactile stimulation in healthy subjects (Vallbo et al., 1995). Since she lacks $A\beta$ afferents, the neuronopathy patient GL has therefore played a fundamental role in determining the central projections of C tactile afferents (Forget and Lamarre, 1995; Sterman, Schaumburg, and Asbury, 1980).

The pattern recognition analysis detected differences in the posterior insular cortex between vibration and brushing, and a closer investigation of the event-related BOLD response demonstrated differential temporal responses to the stimuli. Since brushing activates this region in the patient (Olausson et al., 2002; Olausson et al., 2008b; Björnsdotter et al., 2009), it appears likely that a larger increase in BOLD response due to brushing than to vibration reflects the central projections of C tactile afferents. Similarly, the consistency in activation patterns between the group of healthy individuals and the patient suggest that the observed somatotopic organization indeed reflects brain activations due to C tactile fiber projections rather than of myelinated afferents.

11.2 Parallel tactile systems

Soft brushing, which vigorously activates C tactile as well as Aβ afferents, was found to produce multivariately, but not univariately, differential insular BOLD responses compared to vibration, which preferentially activates Aβ fibers. This spatially encoded difference implies that either the univariate effects are small and enforced through the integration across multiple voxels,

or that the spatial BOLD response patterns differ in a fashion such that no univariate differentiation is at all possible. As indicated by the event-related averaging (figure 10.1), however, it is clear that the BOLD responses to the brushing conditions are of a higher magnitude than those of vibration. Confirming the projection patterns observed in the neuronopathy patients (Olausson et al., 2002; Olausson et al., 2008b), these results provide further support for the hypothesis that human touch is processed in parallel by two distinct systems that comprise Aβ afferents with projections to somatosensory cortices, and C tactile afferents with projections to the insular cortex.

Aβ afferents, with the highest innervation density in the glabrous skin of the hand (Johansson and Vallbo, 1979b), provide the central nervous system with rapid and detailed information regarding tactile stimuli (Sinclair and Hinshaw, 1950; Johnson, 2001). Aβ afferents are essential for precision finger movements such as the ability to manipulate tools or discriminate textures (Johansson and Birznieks, 2004; Hsiao and Bensmaia, 2007). Also, $A\beta$ fibers are fundamentally important for proprioception of larger joints (Edin, 2001) and postural control (Backlund et al., 2005), and project directly to the primary and secondary somatosensory cortices (Hsiao and Bensmaia, 2007). Vibration and brushing should therefore produce localized differential activation patterns also in the somatosensory cortices. None such were found, however, likely due to limitations of fMRI spatial and temporal resolution. It should also be noted that only two subjects were used in paper III, and further research including more individuals is required to substantiate these observations.

As opposed to Aβ afferents, the C tactile system provides poor discriminatory information (e.g. regarding physical location of a stimulus; Olausson et al., 2002; Olausson et al., 2008b; Björnsdotter et al., 2009) and C tactile firing rates correlate well with subjective rating of the pleasantness of a tactile stimulation (Löken et al., 2009). It has therefore been proposed that while the $A\beta$ system provides highly acute information regarding discriminative properties of touch, the C tactile network signals emotional and social aspects of tactile stimulation (Vallbo et al., 1993; Vallbo, Olausson, and Wessberg, 1999; Essick, James, and McGlone, 1999; Olausson et al., 2002).

11.3 Discriminative functions of the CT system

Although the Aβ system is clearly dominant for accurate localization of tactile stimuli, the findings of paper IV suggests that there is a sensory-discriminatory functionality to the C tactile system. Consistent with the somatotopic organization in the insular cortex, the neuronopathy subject, despite lacking thick myelinated afferents and denying any ability of sensing touch below the level of the nose in daily life, could localize the soft brush stimulation to the thigh or arm at an accuracy of 97% in the forced-choice situation (and 72% in a previous quadrant study; Olausson et al., 2008b). The contrast compared to neurologically intact individuals is, however, striking – healthy subjects can localize point indentation on hairy skin with an accuracy of about two cen-

timeters (Norrsell and Olausson, 1994). Also, the patient's C tactile system potentially serves an amplified discriminative function due to central sensory representation adaptations, but, given the similarity in fMRI activation pattern with the healthy controls, it is highly unlikely that the insular somatotopy reflects such neuroplastic changes.

Nonetheless, it appears improbable that the C tactile system plays a significant role in acute spatial localization. Yet, it can be presumed that the general stimulus location significantly modulates affective sensations, which, as opposed to $A\beta$ mediated percepts, are intimately related to C tactile activity (Löken et al., 2009). Propagation of such affective information is fundamentally important in the design or preparation of appropriate actions in response to emotionally relevant stimuli. For example, it has been shown in rats, cats, and humans that painful stimuli applied to various body parts result in correspondingly different autonomic responses (Lewis, 1942; Bandler, Price, and Keay, 2000). It can thus be hypothesized that the crude localization capacity of the C tactile system serves a similar function, where, for example, a gentle stroke on the cheek evokes a different emotional and motivational response than that on the arm, thus signaling various affective aspects with corresponding social implications.

11.4 Central organization of CT-afferents and relation to pain and temperature networks

It has been speculated that C tactile afferents are organized in a fashion similar to that of the pain and temperature mediating thin-fiber system projecting through the lamina I spinothalamic pathway to the insular cortex (Olausson et al., 2002; see Craig, 2002 for a review of the pathway). The insular activation pattern due to C tactile excitation observed in paper IV is highly similar to that of painful and cooling stimuli as shown in figure 11.1. These results support the notion that C tactile afferents indeed project along the lamina I spinothalamic pathway to the posterior insular cortex but not to somatosensory regions.

11.5 Role in homeostasis

It has been suggested that thin afferents constitute an anatomically and functionally distinct system (Craig, 2002). According to this view, thinly myelinated and unmyelinated fibers project from the posterior insular cortex to other insular regions and anterior cingulate cortices (Craig, 2002). Insular and anterior cingulate cortices in turn provide descending control of the autonomic nervous system. Hence, these fibers may act as the afferent limb in a system with the autonomic nervous system as the efferent limb, i.e. an interoceptive system for well-being (Craig, 2002; Craig, 2003b; Craig, 2003c). Furthermore, there is evidence suggesting that C tactile afferents inhibit nociceptive

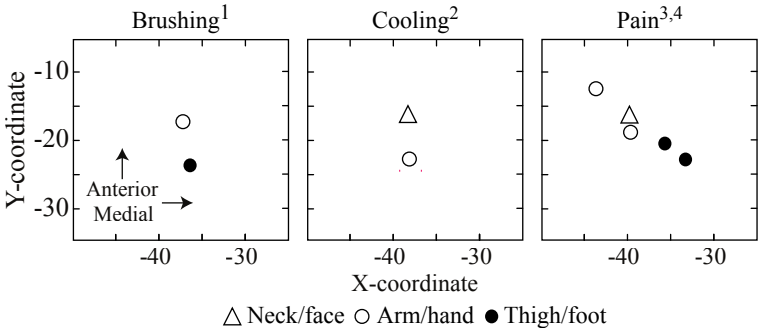

Figure 11.1: Thin-fiber activation center comparison. Posterior insular cortex activation centers (in MNI coordinates) due to application of thin-fiber exciting stimuli (soft brushing, pain and innocuous cooling) on various body parts reported in the present[1] and previous studies (Hua et al., 2005[2], Henderson et al., 2007[3] and Brooks et al., 2005[4]). White markers indicate an upper body stimulus location and black a lower body location. Gentle tactile stimulation (brushing) fit the somatotopic pattern in the posterior contralateral insular cortex well, with upper body stimulations projecting anterior (and slightly lateral) to lower body stimulations.

signals even at the level of the dorsal horn (Lu and Perl, 2003), consistent with a role for regulating well-being. In this light, the C tactile system contributes to the maintenance of physical integrity and well-being (that is, homeostasis) by relaying information regarding the affective tactile status of the body. The tactile well-being of the body is of fundamental importance, especially in primates as illustrated by Harlow's classic study of infant monkeys displaying affection for a surrogate mother in response to tactile comfort (Harlow, 1958). In fact, it has been argued that the C tactile system provides an important sensory underpinning of social behavior and, moreover, plays an integral role in the foundation of self-awareness (McGlone et al., 2007; Olausson et al., 2008; Craig, 2009).

The thin-fiber system is, in addition, proposed not only as an afferent homeostatic pathway, but also to contribute significantly to the construction of the subjective experience of the self and awareness (Craig, 2009). The neuronopathy patient, although denying a general sense of touch, could perceive as well as localize a soft brush stimulation in an experimental situation. While the somatotopically organized posterior insular cortex serves as a primary interoceptive region essential for localization, the reported subsequent mid-to-anterior progression of integration of various physiological representations is likely to play a substantial role in generating an awareness of the tactile stimulation (Burton, Videen, and Raichle, 1993; Spinazzola et al., 2008; Hölzel et al., 2008; Craig, 2009; Lovero et al., 2009). It has previously been observed that soft brushing induces robust activity in the anterior insular cortex (Olausson et al., 2002), suggested as integral in the subjective evaluation of the body's condition ('how you feel'; Craig, 2009).

Concluding remarks

In contrast to conventional univariate statistical analysis where average signal changes in single voxels are detected, machine learning algorithms utilize the inherent multivariate nature of brain activity to identify and localize brain response patterns. As demonstrated in this thesis, such pattern recognition methods are more sensitive than conventional techniques, and also provide a direct link between the BOLD response pattern and corresponding brain state.

A particular aspiration of the thesis research was to implement multivariate brain mapping methods directly applicable in a variety of neuroimaging studies. Two such algorithms for highly sensitive pattern discrimination and localization were proposed, and these were successfully applied to map brain responses to C tactile afferent mediated gentle touch. First, the Monte Carlo approach (paper I) was designed for generic whole brain exploration, and was utilized to identify regions where responses to gentle brushing (combined $A\beta$ and C tactile afferent stimulation) and vibration (mainly activating $A\beta$ fibers) differed (paper III). The evolutionary algorithm (paper II), in contrast, was implemented to customize specific regions of any size and shape to optimally capture regional activity patterns. Such a tailored analysis proved highly useful in the somatotopy study (paper IV), where the specific regions coding for forearm and thigh brushing were desired.

The following list summarizes the main findings of the work in this thesis:

- Multivariate pattern recognition methods are highly useful as a sensitive complement to univariate brain mapping.
- Patterns of brain activity responses to C tactile afferent stimulation differ from those of $A\beta$ fibers in the posterior insular cortex.
- C tactile afferent responses are somatotopically organized in the posterior insular cortex.

Continued interdisciplinary research aiming to refine fMRI pattern recognition methodology promises further insights in both experimental and clinical neuroimaging settings.

Acknowledgements

This work was made possible by the support of the many people in the Department of Physiology and the Department of Clinical Neuroscience and Rehabilitation at the University of Gothenburg. In recognition of all help and encouragement, I would like to express my sincere gratitude to all members of the staff.

A special expression of my deepest appreciation goes to my supervisor Johan Wessberg for his invaluable encouragement and critique. I am particularly grateful for his endless confidence in my research capabilities despite my unorthodox interpretation of space and time.

I am highly thankful to Håkan Olausson for inspiration and support with the physiology part of my thesis, and to Åke Vallbo, my co-supervisor Mikael Elam, Sergei Perfiliev and Ulf Norrsell for the continuous stream of valuable advice and intriguing discussion. I also thank all my senior research colleagues, Anders Hedström, Elena Orekhova, Rick Johnson, Göran Starck, Yrsa Sverrisdottir, Eric Hanse, Ingela Hammar, Bengt Gustafsson, Lars-Gunnar Pettersson, Gunnar Wallin, Fredrik Asztely and Elzbieta Jankowska for expert advice.

I especially would like to thank my wonderful fellow PhD-students, junior research colleagues and master's thesis students, Linda Lundblad, Line Löken, Karin Rylander, Jonny Daborg, My Andersson, Joakim Strandberg, Ilse Riebe, India Morrison, Helena Backlund Wasling, Pontus Wasling, Elin Nilsson, Therese Abrahamsson, Catta Lindquist, Fredrik Hessulf, Åsa Widerstrand, Andreas Björefeldt, Anna Ermund, Malin Stridh, Lotta Olofsson, Irene Perini, Siri Leknes, Nils Löfgren, Johan Löfhede, Josefin Nilsson, Jaquette Liljencrantz, Hector Montes, Sima Shasavari, Simon Bergstrand (SI), Timo Niiniskorpi, Simon Beckmann (SII) and Erik Ziegler, for good times, great laughs and reassuring solidarity.

My sincere gratitude goes to Karin Göthner, Inger Olofsson, Tore Holmström, Kristina Palmgren, Dan Magnusson, Staffan Berg, Johan Kling and Rauni Larsson for assistance with all things practical and technical.

I am particularly grateful to Federico De Martino for his friendship, guidance and invaluable help in the preparation of my thesis. I am thankful to Mattias Wahde who introduced me to the wondrous world of machine learning and to Helge Malmgren and Kajsa Nalin for interesting collaborations on automatic diagnosis systems. I am also highly indebted to Martin Weygandt for his expert advice on the manuscript.

All my appreciation goes to Gabriella Körling, Johanna Lindahl, Sillen Breuer, Benny Alm, Henrik Elofsson Linda, Line, Karin R, Karin G, SI, Erik Z, SII, Lance-Conny Bjerkeli, Calle Jacobson and Mats Hulander who helped me debug my thesis and life. I am abundantly thankful to Erik Z and SII for bulldozing bibtex, and to Daniel Johansson, Simon Casey, Erik Z, Niklas Hansson, Andreas Elmquist and Fabien Buttejat for that marvelous team editing night.

I wish to express my deep love and gratitude to my mother Britt Inger for her ceaseless encouragement, to my father Björn, my siblings Mia, Mattias and Minna, their partners and little ones, and, finally, to all my dear friends. ♡

The work in this thesis was funded by grants from NanoBioTact, the Swedish research council and the Swedish federal government under the LUA/ALF agreement.

Index

References

Åberg, M., L. S. Löken, and J. Wessberg (2009). An evolutionary approach to multivariate feature selection for fMRI pattern analysis. In Pedro Encarnacao, Antonio Veloso, editor, *Proceedings of the First International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS)*, Vol. 2, pp. 302–7, Funchal, Madeira, Portugal. Institute for Systems and Technologies of Information, Control and Communication.

Åberg, M. and J. Wessberg (2007). Evolutionary optimization of classifiers and features for single trial EEG discrimination. *BioMedical Engineering Online* 6(32).

Amar, J. G. (2006). The monte carlo method in science and engineering. *Computing in Science and Engineering.* 8(2): 9–19.

Backlund, W. H., U. Norrsell, K. Göthner, and H. Olausson (2005). Tactile directional sensitivity and postural control. *Experimental Brain Research* 166(2): 147–156.

Baldi, P. and S. Brunak (2001). *Bioinformatics: the machine learning approach*. MIT Press, Cambridge, MA, USA.

Ball, Nicholas M. and Robert J. Brunner (2009). Data mining and machine learning in astronomy .

Bandettini, P. A. and R. W. Cox (2000). Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. *Magnetic Resonance in Medicine* 43(4): 540–548.

Bandler, R., J. L. Price, and K. A. Keay (2000). Brain mediation of active and passive emotional coping. *Progress in Brain Research* 122: 333–349.

Beauchamp, M. S., S. LaConte, and N. Yasar (2009). Distributed representation of single touches in somatosensory and visual cortex. *Human Brain Mapping* (in press).

Bellman, R. E. (1961). *Adaptive Control Processes*. Princeton University Press, Princeton, NJ.

Bensmaia, S. J., Y. Y. Leung, S. S. Hsiao, and K. O. Johnson (2005). Vibratory adaptation of cutaneous mechanoreceptive afferents. *Journal of Neurophysiology* 94(5): 3023–3036.

Bessou, P., P. R. Burgess, E. R. Perl, and C. B. Taylor (1971). Dynamic properties of mechanoreceptors with unmyelinated (c) fibers. *Journal of Neurophysiology* 34(1): 116–131.

Bessou, P. and E. R. Perl (1969). Response of cutaneous sensory units with unmyelinated fibers to noxious stimuli. *Journal of Neurophysiology* 32(6): 1025–1043.

Birbaumer, N. and L. G. Cohen (2007). Brain-computer interfaces: communication and restoration of movement in paralysis. *Journal of Physiology* 579(Pt 3): 621–636.

Björnsdotter, M., L. S. Löken, H. Olausson, Å. B. Vallbo, and J. Wessberg (2009). Somatotopic organization of gentle touch processing in the posterior insular cortex. *Journal of Neuroscience* 29(29): 9314–9320.

Björnsdotter, M., K. Rylander, and J. Wessberg (2009). A monte carlo method for locally-multivariate brain mapping (submitted).

Björnsdotter Åberg, M. and J. Wessberg (2008). An evolutionary approach to the identification of informative voxel clusters for brain state discrimination. *IEEE Journal of Selected Topics in Signal Processing* 2(6): 919–928.

Björnsdotter Åberg, M. and J. Wessberg (2009). Particle swarm voxel clustering for multivariate fMRI mapping. *NeuroImage* 47: S39–S41.

Blomqvist, A. and A. D. Craig (2000). Is neuropathic pain caused by the activation of nociceptive-specific neurons due to anatomic sprouting in the dorsal horn? *The Journal of Comparative Neurology* 428(1): 1–4.

Blum, A. and P. Langley (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence* 97(1-2): 245–271.

Bode, S. and J.-D. Haynes (2009). Decoding sequential stages of task preparation in the human brain. *NeuroImage* 45(2): 606–613.

Bolanowski, S. J. (1984). Intensity and frequency characteristics of pacinian corpuscles. III. effects of tetrodotoxin on transduction process. *Journal of Neurophysiology* 51(4): 831–839.

Bolanowski, S. J. and J. J. Zwislocki (1984a). Intensity and frequency characteristics of pacinian corpuscles. I. action potentials. *Journal of Neurophysiology* 51(4): 793–811.

Bolanowski, S. J. and J. J. Zwislocki (1984b). Intensity and frequency characteristics of pacinian corpuscles. II. receptor potentials. *Journal of Neuophysiology* 51(4): 812–830.

Boser, B. E., I. M. Guyon, and V. N. Vapnik (1992). A training algorithm for optimal margin classifiers. *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory* pp. 144–152.

Brooks, J. C. W., L. Zambreanu, A. Godinez, A. D. Craig, and I. Tracey (2005). Somatotopic organisation of the human insula to painful heat studied with high resolution functional imaging. *NeuroImage* 27(1): 201–209.

Büchel, C., A. P. Holmes, G. Rees, and K. J. Friston (1998). Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. *NeuroImage* 8(2): 140–148.

Buckner, R. L. (1998). Event-related fMRI and the hemodynamic response. *Human Brain Mapping* 6(5-6): 373–377.

Buntine, W. (1989). Learning classification rules using bayes. In *Proceedings of the sixth international workshop on Machine learning*, pp. 94–98, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Burgess, P. R. and E. R. Perl (1967). Myelinated afferent fibres responding specifically to noxious stimulation of the skin. *Journal of Physiology* 190(3): 541–562.

Burke, D., K. Murphy, H. Garavan, and R. Reilly (2004). Pattern recognition approach to the detection of single-trial event-related functional magnetic resonance images. *Medical and Biological Engineering and Computing* 42(5): 604–609.

Burton, H., T. O. Videen, and M. E. Raichle (1993). Tactile-vibration-activated foci in insular and parietal-opercular cortex studied with positron emission tomography: mapping the second somatosensory area in humans. *Somatosensory and Motor Research* 10(3): 297–308.

Carlson, T. A., P. Schrater, and S. He (2003). Patterns of activity in the categorical representations of objects. *Journal of Cognitive Neuroscience* 15(5): 704–717.

Chumbley, J. R. and K. J. Friston (2009). False discovery rate revisited: FDR and topological inference using Gaussian random fields. *NeuroImage* 44(1): 62–70.

Clithero, J. A., R. M. Carter, and S. A. Huettel (2008). Local pattern classification differentiates processes of economic valuation. *NeuroImage* 45(4): 1329–1338.

Coghill, R. C., C. N. Sang, J. M. Maisog, and M. J. Iadarola (1999). Pain intensity processing within the human brain: a bilateral, distributed mechanism. *Journal of Neurophysiology* 82(4): 1934–1943.

Collins, D. L., P. Neelin, T. M. Peters, and A. C. Evan (1994). Automatic 3d intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography* 18(2): 192–205.

Connor, C. E., S. S. Hsiao, J. R. Phillips, and K. O. Johnson (1990). Tactile roughness: neural codes that account for psychophysical magnitude estimates. *Journal of Neuroscience* 10(12): 3823–3836.

Cover, T. and P. Hart (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13(1): 21–27.

Cox, D. D. and R. L. Savoy (2003). Functional magnetic resonance imaging (fMRI) 'brain reading': detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage* 19(2 Pt 1): 261–270.

Craig, A. D. (2002). How do you feel? interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience* 3(8): 655–666.

Craig, A. D. (2003a). Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology* 13(4): 500–505.

Craig, A. D. (2003b). A new view of pain as a homeostatic emotion. *Trends in Neuroscience* 26(6): 303–307.

Craig, A. D. (2003c). Pain mechanisms: labeled lines versus convergence in central processing. *Annual Review of Neuroscience* 26: 1–30.

Craig, A. D., M. C. Bushnell, E. T. Zhang, and A. Blomqvist (1994). A thalamic nucleus specific for pain and temperature sensation. *Nature* 372(6508): 770–773.

Craig, A. D., K. Chen, D. Bandy, and E. M. Reiman (2000). Thermosensory activation of insular cortex. *Nature Neuroscience* 3(2): 184–190.

Craig, A. D., K. Krout, and D. Andrew (2001). Quantitative response characteristics of thermoreceptive and nociceptive lamina i spinothalamic neurons in the cat. *Journal of Neurophysiology* 86(3): 1459–1480.

Craig, A. D. and E. T. Zhang (2006). Retrograde analyses of spinothalamic projections in the macaque monkey: input to posterolateral thalamus. *The Journal of Comparative Neurology* 499(6): 953–964.

Craig, A. D., E. T. Zhang, and A. Blomqvist (1999). A distinct thermoreceptive subregion of lamina i in nucleus caudalis of the owl monkey. *The Journal of Comparative Neurology* 404(2): 221–234.

Craig, A. D. (Bud) (2009). How do you feel – now? The anterior insula and human awareness. *Nature Reviews Neuroscience* 10(1): 59–70.

Craig, A.D. (Bud) (2008). *Handbook of emotions*, chapter Interoception and Emotion A Neuroanatomical Perspective, pp. 272–288. Guilford Press.

Davatzikos, C., K. Ruparel, Y. Fan, D. G. Shen, M. Acharyya, J. W. Loughead, R. C. Gur, and D. D. Langleben (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage* 28: 663–668.

De Martino, F., F. Gentile, F. Esposito, M. Balsi, F. Di Salle, R. Goebel, and E. Formisano (2007). Classification of fMRI independent components using ic-fingerprints and support vector machine classifiers. *NeuroImage* 34(1): 177–194.

De Martino, F., G. Valente, N. Staeren, J. Ashburner, R. Goebel, and E. Formisano (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *NeuroImage* 43(1): 44–58.

deCharms, R. C. (2008). Applications of real-time fMRI. *Nature Reviews Neuroscience* 9(9): 720–729.

Detre, G. J., S. M. Polyn, C. D. Moore, V. S. Natu, B. D. Singer, J. D. Cohen, J. V. Haxby, and K. A. Norman (2005). The multi-voxel pattern analysis MVPA toolbox. *The 11th Annual Meeting of the Organization of Human Brain Mapping, Toronto, Canada* 50.

Disbrow, E., T. Roberts, D. Poeppel, and L. Krubitzer (2001). Evidence for interhemispheric processing of inputs from the hands in human S2 and PV. *Journal of Neurophysiology* 85(5): 2236–2244.

Dostrovsky, J. O. and A. D. Craig (1996). Cooling-specific spinothalamic neurons in the monkey. *Journal of Neurophysiology* 76(6): 3656–3665.

Douglas, W. W. and J. M. Ritchie (1957). Nonmedullated fibres in the saphenous nerve which signal touch. *Journal of Physiology* 139(3): 385–399.

Duda, R. O., P. E. Hart, and D. G. Stork (2000). *Pattern Classification*. Wiley-Interscience Publication.

Dunwei, G., P. Fengping, and X. Shifan (2002). Adaptive niche hierarchy genetic algorithm. In *Proceedings IEEE Region 10 Conference on Computers, Communications Control and Power Engineering TENCON '02*, Vol. 1, pp. 39–42.

Ecker, C., V. Rocha-Rego, P. Johnston, J. Mourão-Miranda, A. Marquand, E. M. Daly, M. J. Brammer, C. Murphy, and D. G. Murphy (2009). Investigating the predictive value of whole-brain structural MR scans in autism: A pattern classification approach. *NeuroImage* (in press).

Edin, B. (2001). Cutaneous afferents provide information about knee joint movements in humans. *Journal of Physiology* 531(Pt 1): 289–297.

Edin, B. B., G. K. Essick, M. Trulsson, and K. A. Olsson (1995). Receptor encoding of moving tactile stimuli in humans. i. temporal pattern of discharge of individual low-threshold mechanoreceptors. *Journal of Physiology* 15(1 Pt 2): 830–847.

Essick, G. K., A. James, and F. P. McGlone (1999). Psychophysical assessment of the affective components of non-painful touch. *Neuroreport* 10(10): 2083–2087.

Evans, A. C., D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters (1993). 3D statistical neuroanatomical models from 305 MRI volumes. *Proceedings of the IEEE-Nuclear Science Symposium and Medical Imaging Conference* pp. 1813–1817.

Fisher, R. A. (1936). The use of multiple measures in taxonomic problems. *Annals of Eugenics* 7: 179–188.

Fodor, J. (1983). *The Modularity of Mind. An Essay on Faculty Psychology*. MIT Press, Cambridge, 5th edition.

Forget, R. and Y. Lamarre (1995). Postural adjustments associated with different unloadings of the forearm: effects of proprioceptive and cutaneous afferent deprivation. *Canadian Journal of Physiology and Pharmacology* 73(2): 285–294.

Formisano, E., F. De Martino, M. Bonte, and R. Goebel (2008). "who" is saying "what"? brain-based decoding of human voice and speech. *Science* 5903(322): 970–973.

Friedman, R. M., P. S. Khalsa, K. W. Greenquist, and R. H. LaMotte (2002). Neural coding of the location and direction of a moving object by a spatially distributed population of mechanoreceptors. *Journal of Neuroscience* 22(21): 9556–9566.

Friman, O., M. Borga, P. Lundberg, and H. Knutsson (2004). Detection and detrending in fMRI data analysis. *NeuroImage* 22(2): 645–655.

Friston, K. (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press.

Friston, K. J., P. Fletcher, O. Josephs, A. Holmes, M. D. Rugg, and R. Turner (1998). Event-related fMRI: Characterizing differential responses. *NeuroImage* 7(1): 30–40.

Friston, K. J., A. P. Holmes, K. J. Worsley, J. P. Poline, C. D. Frith, and R. S. J. Frackowiak (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping* 2(4): 189–210.

Fu, C. H. Y., J. Mourão-Miranda, S. G. Costafreda, A. Khanna, A. F. Marquand, S. C. R. Williams, and M. J. Brammer (2008). Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. *Biological Psychiatry* 63(7): 656–662.

Ganesh, G., E. Burdet, M. Haruno, and M. Kawato (2008). Sparse linear regression for reconstructing muscle activity from human cortical fMRI. *NeuroImage* 42(4): 1463–1472.

Gardner, E. P. and J. H. Martin (2000). *Coding of sensory information. Principles of neural science.* McGraw-Hill.

Genovese, C. R., N. A. Lazar, and T. Nichols (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15(4): 870–878.

Goense, J. B. M. and N. K. Logothetis (2008). Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current Biology* 18(9): 631–640.

Golland, P., F. Liang, S. Mukherjee, and D. Panchenko (2005). *Permutation Tests for Classification*. Springer, Berlin / Heidelberg. 978-3-540-26556-6.

Good, P. I. (2004). *Permutation, Parametric, and Bootstrap Tests of Hypotheses (Springer Series in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Grave de Peralta Menendez, R., M. M. Murray, and S. L. Gonzalez Andino (2004). Improving the performance of linear inverse solutions by inverting the resolution matrix. *IEEE Transactions in Biomedical Engineering* 51(9): 1680–1683.

Guyon, I. and A. Elisseeff (2003). An introduction to variable and feature selection. *Journal of machine learning research* 3(1): 1157–1182.

Haggard, P. (2006). Sensory neuroscience: from skin to object in the somatosensory cortex. *Current Biology* 16(20): R884–R886.

Han, Z. S., E. T. Zhang, and A. D. Craig (1998). Nociceptive and thermoreceptive lamina I neurons are anatomically distinct. *Nature Neuroscience* 1(3): 218–225.

Hanke, M., Y. O. Halchenko, P. B. Sederberg, S. José Hanson, J. V. Haxby, and S. Pollmann (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fmri data. *Neuroinformatics* 7(1).

Hanson, S. J. and Y. O. Halchenko (2008). Brain reading using full brain support vector machines for object recognition: There is no face identification area. *Neural Computation* 20(2): 486–503.

Hanson, S. J., T. Matsuka, and J. V. Haxby (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? *NeuroImage* 23(1): 156–166.

Harel, N., K. Uğurbil, K. Uludağ, and E. Yacoub (2006). Frontiers of brain mapping using MRI. *Journal of Magnetic Resonance Imaging* 23(6): 945–957.

Harlow, H. F. (1958). The nature of love. *American Psychologist* 13(2): 673–685.

Hashemi, R. H., W. G. Bradley, and C. J. Lisanti (2004). *MRI: The Basics*. Lippincott Williams and Wilkins.

Haxby, J. V., M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, and P. Pietrini (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293: 2425–2430.

Haykin, S. (1999). *Neural Networks, A Comprehensive Foundation*. Prentice Hall, 2nd edition.

Haynes, J.-D. and G. Rees (2005a). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience* 8(5): 686–691.

Haynes, J.-D. and G. Rees (2005b). Predicting the stream of consciousness from activity in human visual cortex. *Current Biology* 15: 1301–1307.

Haynes, J.-D. and G. Rees (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience* 7(7): 523–534.

Haynes, J.-D., K. Sakai, G. Rees, S. Gilbert, C. Frith, and R. E. Passingham (2007). Reading hidden intentions in the human brain. *Current Biology* 17(4): 323–328.

Heeger, D. J. and D. Ress (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience* 3(2): 142–151.

Henderson, L. A., S. C. Gandevia, and V. G. Macefield (2007). Somatotopic organization of the processing of muscle and cutaneous pain in the left and right insula cortex: A single-trial fMRI study. *Pain* 128(1-2): 20–30.

Henson, R. N. A. (2003). Analysis of fMRI time series. In Frackowiak, R. S. J., K. J. Friston, C. Frith, R. Dolan, K. J. Friston, C. J. Price, S. Zeki, J. Ashburner, and W. D. Penny, editors, *Human Brain Function*. Academic Press, 2nd edition.

Hofbauer, R. K., P. Rainville, G. H. Duncan, and M. C. Bushnell (2001). Cortical representation of the sensory dimension of pain. *Journal of Neurophysiology* 86(1): 402–411.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, USA.

Hölzel, B. K., U. Ott, T. Gard, H. Hempel, M. Weygandt, K. Morgen, and D. Vaitl (2008). Investigation of mindfulness meditation practitioners with voxel-based morphometry. *Social Cognitive and Affective Neuroscience* 3(1): 55–61.

Howard, J. D., J. Plailly, M. Grueschow, J.-D. Haynes, and J. A. Gottfried (2009). Odor quality coding and categorization in human posterior piriform cortex. *Nature Neuroscience* (in press).

Hsiao, S. S. and S. Bensmaia (2007). *Coding of Object Shape and Texture, in The senses: A Comprehensive Reference*. Elsevier, Amsterdam.

Hua, L. H., I. A. Strigo, L. C. Baxter, S. C. Johnson, and A. D. Craig (2005). Anteroposterior somatotopy of innocuous cooling activation focus in human dorsal posterior insular cortex. *American Journal of Physiology - Regulatory, Integrative, and Comparative Physiology* 289(2): 319–325.

Iggo, A. (1959). Cutaneous heat and cold receptors with slowly conducting C afferent fibres. *Quarterly journal of experimental physiology and cognitive medical sciences* 44: 362–370.

Iggo, A (1960). Cutaneous mechanoreceptors with afferent C fibres. *Journal of Physiology* 152: 337–353.

Iggo, A. and H. H. Kornhuber (1977). A quantitative study of C-mechanoreceptors in hairy skin of the cat. *Journal of Physiology* 271(2): 549–565.

Iggo, A. and A. R. Muir (1969). The structure and function of a slowly adapting touch corpuscle in hairy skin. *Journal of Physiology* 200(3): 763–796.

Johansson, R. S. and I. Birznieks (2004). First spikes in ensembles of human tactile afferents code complex spatial fingertip events. *Nature Neuroscience* 7(2): 170–177.

Johansson, R. S. and Å. B. Vallbo (1979a). Detection of tactile stimuli. thresholds of afferent units related to psychophysical thresholds in the human hand. *Journal of Physiology* 297(0): 405–422.

Johansson, R. S. and Å. B. Vallbo (1979b). Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *Journal of Physiology* 286: 283–300.

Johansson, R. S. and Å. B. Vallbo (1980). Spatial properties of the population of mechanoreceptive units in the glabrous skin of the human hand. *Brain Research* 184(2): 353–366.

Johnson, K. O. (2001). The roles and functions of cutaneous mechanoreceptors. *Current Opinion in Neurobiology* 11(4): 455–461.

Kaas, J. H. (2007). *The Somatosensory Thalamus and Associated Pathways in The senses: A Comprehensive Reference*. Elsevier, Amsterdam.

Kamitani, Y. and F. Tong (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience* 8(5): 679–685.

Kandel, E. R., J. H. Schwartz, and T. M. Jessell (2000). *Principles of Neural Science*. McGraw-Hill.

Kemsley, E. K., G. Le Gall, J. R. Dainty, A. D. Watson, L. J. Harvey, H. S. Tapp, and I. J. Colquhoun (2007). Multivariate techniques and their application in nutrition: a metabolomics case study. *British Journal of Nutrition* 98(1): 1–14.

Knibestöl, M. (1973). Stimulus-response functions of rapidly adapting mechanoreceptors in human glabrous skin area. *Journal of Physiology* 232(3): 427–452.

Knibestöl, M. (1975). Stimulus-response functions of slowly adapting mechanoreceptors in the human glabrous skin area. *Journal of Physiology* 245(1): 63–80.

Kohavi, R. and G. H. John (1997). Wrappers for feature subset selection. *Artificial Intelligence* 97(1-2): 273–324.

Koutsouleris, N., E. M. Meisenzahl, C. Davatzikos, R. Bottlender, T. Frodl, J. Scheuerecker, G. Schmitt, T. Zetzsche, P. Decker, M. Reiser, H.-J. Möller, and C. Gaser (2009). Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Archives of General Psychiatry* 66(7): 700–712.

Kriegeskorte, N. and P. Bandettini (2007). Analyzing for information, not activation, to exploit high-resolution fmri. *NeuroImage* 38(4): 649–662.

Kriegeskorte, N., E. Formisano, B. Sorger, and R. Goebel (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences* 104(51): 20600–20605.

Kriegeskorte, N., R. Goebel, and P. Bandettini (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Science* 103: 3863–3868.

Kriegeskorte, N., M. Mur, D. A. A. Ruff, R. Kiani, J. Bodurka, H. Esteky, K. Tanaka, and P. A. A. Bandettini (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60(6): 1126–1141.

Kriegeskorte, N., K. W. Simmons, P. S. Bellgowan, and C. I. Baker (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience* 12(5): 535–540.

Krubitzer, L. A. and J. H. Kaas (1987). Thalamic connections of three representations of the body surface in somatosensory cortex of gray squirrels. *Journal of Comparative Neurology* 265(4): 549–580.

Ku, S. P., A. Gretton, J. Macke, and N. K. Logothetis (2008). Comparison of pattern recognition methods in classifying high-resolution bold signals obtained at high magnetic field in monkeys. *Magnetic Resonance Imaging* 26(7): 1007–1014.

Kumazawa, T. and E. R. Perl (1977a). Primate cutaneous receptors with unmyelinated C fibres and their projection to the substantia gelatinosa. *Journal of Physiology* 73(3): 287–304.

Kumazawa, T. and E. R. Perl (1977b). Primate cutaneous sensory units with unmyelinated C afferent fibers. *Journal of Neurophysiology* 40(6): 1325–1338.

Kumazawa, T. and E. R. Perl (1978). Excitation of marginal and substantia gelatinosa neurons in the primate spinal cord: indications of their place in dorsal horn functional organization. *Journal of Comparative Neurology* 177(3): 417–434.

LaConte, S., J. Anderson, S. Muley, J. Ashe, S. Frutiger, K. Rehm, L. K. Hansen, E. Yacoub, X Hu, D Rottenberg, and S. Strother (2003). The evaluation of preprocessing choices in single-subject BOLD fMRI using NPAIRS performance metrics. *Neuroimage* 18(1): 10–27.

LaConte, S., S. Strother, V. Cherkassky, J. Anderson, and X. Hu (2005). Support vector machines for temporal classification of block design fMRI data. *NeuroImage* 26(2): 317–329.

LaConte, S. M., S. J. Peltier, and X. P. Hu (2007). Real-time fMRI using brain-state classification. *Human Brain Mapping* 28(10): 1033–1044.

Laubach, M., J. Wessberg, and M. A. Nicolelis (2000). Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. *Nature* 405(6786): 567–571.

Lavine, B. K. and M. N. Vora (2005). Identification of africanized honeybees. *Journal of Chromatography A* 1096(1-2): 69–75.

Leem, J. W., W. D. Willis, and J. M. Chung (1993). Cutaneous sensory receptors in the rat foot. *Journal of Neurophysiology* 69(5): 1684–1699.

Leoni, R. F., K. C. Mazzeto-Betti, K. C. Andrade, and D. B. de Araujo (2008). Quantitative evaluation of hemodynamic response after hypercapnia among different brain territories by fMRI. *NeuroImage* 41(4): 1192–1198.

Lewis, T. (1942). *Pain*. MacMillan, New York, USA.

Li, L., H. Tang, Z. Wu, J. Gong, M. Gruidl, J. Zou, M. Tockman, and R. A. Clark (2004). Data mining techniques for cancer detection using serum proteomic profiling. *Artificial Intelligence in Medicine* 32(2): 71–83.

Li, Z-C., X-B. Zhou, Y-R. Lin, and X-Y. Zou (2008). Prediction of protein structure class by coupling improved genetic algorithm and support vector machine. *Amino Acids* 35(3): 581–590.

Light, A. R. and H. H. Willcockson (1999). Spinal laminae I–II neurons in rat recorded in vivo in whole cell, tight seal configuration: properties and opioid responses. *Journal of Neurophysiology* 82(6): 3316–3326.

Liu, Q., S. Vrontou, F. L. Rice, M. J. Zylka, X. Dong, and D. J. Anderson (2007). Molecular genetic visualization of a rare subset of unmyelinated sensory neurons that may detect gentle touch. *Nature Neuroscience* 10(8): 946–948.

Logothetis, N. K., J. Pauls, M. Augath, T. Trinath, and A. Oeltermann (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412(6843): 150–157.

Löken, L. S., J. Wessberg, I. Morrison, F. McGlone, and H. Olausson (2009). Coding of pleasant touch by unmyelinated afferents in humans. *Nature neuroscience* 12(5): 547–548.

Löken, L. S., J. Wessberg, and H. W. Olausson (2007). Unmyelinated tactile (CT) afferents are present in the human peroneal and radial nerves. *Society for Neuroscience 37th Annual Meeting, San Diego, USA* 827.2.

Lovero, K. L., A. N. Simmons, J. L. Aron, and M. P. Paulus (2009). Anterior insular cortex anticipates impending stimulus significance. *NeuroImage* (in press).

Lu, Y. and E. R. Perl (2003). A specific inhibitory pathway between substantia gelatinosa neurons receiving direct C-fiber input. *Journal of Neuroscience* 23(25): 8752–8758.

Lynn, B. and S. E. Carpenter (1982). Primary afferent units from the hairy skin of the rat hind limb. *Brain Research* 238(1): 29–43.

Marinakis, Y., G. Dounias, and J. Jantzen (2009). Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification. *Computers in Biology and Medicine* 39(1): 69–78.

Marquand, A. F., J. Mourão-Miranda, M. J. Brammer, A. J. Cleare, and C. H. Y. Fu (2008). Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. *Neuroreport* 19(15): 1507–1511.

McGlone, F., Å. B. Vallbo, H. Olausson, L. S. Löken, and J. Wessberg (2007). Discriminative touch and emotional touch. *The Canadian Journal of Experimental Psychology* 61(3): 173–183.

McLeish, D. L. (2005). *Monte Carlo simulation and finance*. John Wiley & Sons, New York, USA.

Miranda, E. R. and J. A. Biles (2007). *Evolutionary Computer Music*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Mitchell, T. M., R. Hutchinson, R. S. Niculescu, F. Pereira, X. Wang, M. Just, and S. Newman (2004). Learning to decode cognitive states from brain images. *Machine Learning* 57(1-2): 145–175.

Mørch, N., L. K. Hansen, S. C. Strother, C. Svarer, D. A. Rottenberg, B. Lautrup, R. Savoy, and O. B. Paulson (1997). Nonlinear versus linear models in functional neuroimaging: Learning curves and generalization crossover. In *IPMI '97: Proceedings of the 15th International Conference on Information Processing in Medical Imaging*, pp. 259–270, London, UK. Springer-Verlag.

Mountcastle, V. B., W. H. Talbot, I. Darian-Smith, and H. H. Kornhuber (1967). Neural basis of the sense of flutter-vibration. *Science* 155(762): 597–600.

Mourão-Miranda, J., A. L. Bokde, C. Born, H. Hampel, and M. Stetter (2005). Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data. *NeuroImage* 28(4): 980–995.

Mourão-Miranda, J., C. Ecker, J. R. Sato, and M. Brammer (2009). Dynamic changes in the mental rotation network revealed by pattern recognition analysis of fMRI data. *Journal of Cognitive Neuroscience* 21(5): 890–904.

Mourão-Miranda, J., E. Reynaud, F. McGlone, G. Calvert, and M. Brammer (2006). The impact of temporal compression and space selection on SVM analysis of single-subject and multi-subject fMRI data. *NeuroImage* 33(4): 1055–1065.

Muehllehner, G. and J. S. Karp (2006). Positron emission tomography. *Physics in Medicine and Biology* 51(13): R117–R137.

Murray, E. A. and M. Mishkin (1984). Relative contributions of SII and area 5 to tactile discrimination in monkeys. *Behavioral Brain Research* 11(1): 67–83.

Naidich, T. P., E. Kang, G. M. Fatterpekar, B. N. Delman, S. H. Gultekin, D. Wolfe, O. Ortiz, I. Yousry, M. Weismann, and T. A. Yousry (2004). The insula: anatomic study and MR imaging display at 1.5 T. *American Journal of Neuroradiology* 25(2): 222–232.

Nichols, T. and S. Hayasaka (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical Methods in Medical Research* 12(5): 419–446.

Niiniskorpi, T., M. Björnsdotter Åberg, and J. Wessberg (2009). Particle swarm feature selection for fMRI pattern classification. In *Proceedings of the Second International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS)*, pp. 279–284. INSTICC - Institute for Systems and Technologies of Information, Control and Communication.

Nordin, M. (1990). Low-threshold mechanorecepetive and nociceptive units with unmyelinated C fibres in the human supraorbital nerve. *Journal of Physiology* 426(310): 229–240.

Norman, K. A., S. M. Polyn, G. J. Detre, and J. V. Haxby (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences* 10(9): 424–430.

Norris, D. G. (2006). Principles of magnetic resonance assessment of brain function. *Journal of Magnetic Resonance Imaging* 23(6): 794–807.

Norrsell, U. and H. Olausson (1994). Spatial cues serving the tactile directional sensibility of the human forearm. *Journal of Physiology* 478(Pt 3): 533–540.

Ogawa, S., T. M. Lee, A. R. Kay, and D. W. Tank (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences* 87(24): 9868–9872.

Olausson, H., J. Cole, K. Rylander, F. McGlone, Y. Lamarre, B. G. Wallin, H. Krämer, J. Wessberg, M. Elam, M. C. Bushnell, and Å. B. Vallbo (2008). Functional role of unmyelinated tactile afferents in human hairy skin: sympathetic response and perceptual localization. *Experimental Brain Research* 184(1): 135–140.

Olausson, H., Y. Lamarre, H. Backlund, C. Morin, B. G. Wallin, G. Starck, S. Ekholm, I. Strigo, K. Worsley, Å. B. Vallbo, and M. C. Bushnell (2002). Unmyelinated tactile afferents signal touch and project to insular cortex. *Nature Neuroscience* 5(9): 900–904.

Olausson, H., J. Wessberg, I. Morrison, F. McGlone, and Å. B. Vallbo (2008a). The neurophysiology of unmyelinated tactile afferents. *Neuroscience & Biobehavioral Reviews* (in press).

Olausson, H. W., J. Cole, Å. B. Vallbo, F. McGlone, M. Elam, H. H. Krämer, K. Rylander, J. Wessberg, and M. C. Bushnell (2008b). Unmyelinated tactile afferents have opposite effects on insular and somatosensory cortical processing. *Neuroscience Letters* 436(2): 128–132.

O'Toole, A. J., F. Jiang, H. Abdi, and J. V. Haxby (2005). Partially distributed representations of objects and faces in ventral temporal cortex. *Journal of Cognitive Neuroscience* 17(4): 580–590.

Pereira, F., T. Mitchell, and M. Botvinick (2009). Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage* 45(1 Suppl): S199–S209.

Phillips, J. R., R. S. Johansson, and K. O. Johnson (1992). Responses of human mechanoreceptive afferents to embossed dot arrays scanned across fingerpad skin. *Journal of Neuroscience* 12(3): 827–839.

Polyn, S. M., V. S. Natu, J. D. Cohen, and K. A. Norman (2005). Category-specific cortical activity precedes retrieval during memory search. *Science* 310(5756): 1963–1966.

Qi, H.-X., T. M. Preuss, and J. H. Kaas (2007). *Somatosensory Areas of the Cerebral Cortex: Architectonic Characteristics and Modular Organization in The senses: A Comprehensive Reference*. Elsevier, Amsterdam.

Reeves, C. R. and J. E. Rowe (2002). *Genetic Algorithms - Principles and Perspectives: A Guide to GA Theory*. Kluwer Academic Publishers, Norwell, MA, USA.

Ritter, P. and A. Villringer (2006). Simultaneous EEG-fMRI. *Neuroscience Biobehavioural Reviews* 30(6): 823–838.

Rosen, B. R., R. L. Buckner, and A. M. Dale (1998). Event-related functional MRI: past, present and future. *Proceedings of the National Academy of Sciences of the U.S.A.* 95: 773–780.

Roy, C. S. and C. S. Sherrington (1890). On the regulation of the blood-supply of the brain. *Journal of Physiology* 11(1-2).

Ruben, J., J. Schwiemann, M. Deuchert, R. Meyer, T. Krause, G. Curio, K. Villringer, R. Kurth, and A. Villringer (2001). Somatotopic organization of human secondary somatosensory cortex. *Cerebral Cortex* 11(5): 463–473.

Salarian, A., H. Russmann, F. J. G. Vingerhoets, P. R. Burkhard, and K. Aminian (2007). Ambulatory monitoring of physical activities in patients with parkinson's disease. *IEEE Transactions on Biomedical Engineering* 54(12): 2296–2299.

Sato, J. R., A. Fujita, C. E. Thomaz, G. Martin Mda, J. Mourão-Miranda, M. J. Brammer, and E. Amaro Junior (2009). Evaluating SVM and MLDA in the extraction of discriminant regions for mental state prediction. *NeuroImage* 46(1): 105–114.

Schölkopf, B. and A. J. Smola (2001). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)*. The MIT Press.

Schwartzman, A., R. F. Dougherty, J. Lee, D. Ghahremani, and J. E. Taylor (2009). Empirical null and false discovery rate analysis in neuroimaging. *NeuroImage* 44(1): 71–82.

Shea, V. K. and E. R. Perl (1985). Sensory receptors with unmyelinated C fibers innervating the skin of the rabbit's ear. *Journal of Neurophysiology* 54(3): 491–501.

Sinclair, D. C. and J. R. Hinshaw (1950). A comparison of the sensory dissociation produced by procaine and by limb compression. *Brain* 78(4): 480–498.

Sirotin, Y. B. and A. Das (2008). Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature* 457(7228): 475–479.

Sitaram, R., A. Caria, and N. Birbaumer (2009). Hemodynamic brain-computer interfaces for communication and rehabilitation. *Neural Networks* (in press).

Sobol, I. M. (1994). *A Primer for the Monte Carlo Method.* CRC Press, Boca Raton, FL.

Spinazzola, L., L. Pia, A. Folegatti, C. Marchetti, and A. Berti (2008). Modular structure of awareness for sensorimotor disorders: evidence from anosognosia for hemiplegia and anosognosia for hemianaesthesia. *Neuropsychologia* 46(3): 915–926.

Spiridon, M. and N. Kanwisher (2002). How distributed is visual category information in human occipito-temporal cortex? an fMRI study. *Neuron* 35: 1157–1165.

Staeren, N., H. Renvall, De Martino F., R. Goebel, and E. Formisano (2009). Sound categories are represented as distributed patterns in the human auditory cortex. *Current Biology* 19(6): 498–502.

Stender, J. (1993). *Parallel Genetic Algorithms: Theory and Applications*. IOS Press, Amsterdam, The Netherlands.

Sterman, A. B., H. H. Schaumburg, and A. K. Asbury (1980). The acute sensory neuronopathy syndrome: a distinct clinical entity. *Annals of Neurology* 7(4): 354–358.

Stokes, M., R. Thompson, R. Cusack, and J. Duncan (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience* 29(5): 1565–1572.

Sugiura, Y., C. L. Lee, and E. R. Perl (1986). Central projections of identified, unmyelinated (C) afferent fibers innervating mammalian skin. *Science* 234: 358–361.

Sugiura, Y., N. Terui, and Y. Hosoya (1989). Difference in distribution of central terminals between visceral and somatic unmyelinated (C) primary afferent fibers. *Journal of Neurophysiology* 62(4): 834–840.

Suykens, J. A. K., T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle (2002). *Least Squares Support Vector Machines*. World Scientific.

Talairach, J. and P. Tournoux (1988). *Co-Planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System : An Approach to Cerebral Imaging*. Thieme Medical Publishers.

Taneda, A. (2008). An efficient genetic algorithm for structural RNA pairwise alignment and its application to non-coding RNA discovery in yeast. *BMC Bioinformatics* 9: 521.

Thachuk, C., J. Crossa, J. Franco, S. Dreisigacker, M. Warburton, and G. Davenport (2009). Core hunter: an algorithm for sampling genetic resources based on multiple genetic measures. *BMC Bioinformatics* 10(1): 243.

Theodoridis, S. and K. Koutroumbas (2006). *Pattern Recognition, Third Edition*, chapter Nonlinear classifiers. Academic Press, San Diego, USA.

Thompson, S. K. (2005). The legality of the use of psychiatric neuroimaging in intelligence interrogation. *Cornell Law Review* 90: 1601–1637.

Turner, R., A. Howseman, and K. Friston (1998). Functional magnetic resonance imaging of the human brain: Data acquisition and analysis. *Experimental Brain Research* 123: 5.

Uğurbil, K., G. Adriany, P. Andersen, W. Chen, M. Garwood, R. Gruetter, P.-G. Henry, S.-G. Kim, H. Lieu, I. Tkac, T. Vaughan, P.-F. Van De Moortele, E. Yacoub, and X.-H. Zhu (2003). Ultrahigh field magnetic resonance imaging and spectroscopy. *Magnetic Resonance Imaging* 21(10): 1263–1281.

Uğurbil, K., L. Toth, and D. S. Kim (2003). How accurate is magnetic resonance imaging of brain function? *Trends in Neuroscience* 26(2): 108–114.

Uttal, W. R. (2003). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. The MIT Press.

Vallbo, Å. B. and R. S. Johansson (1984). Properties of cutaneous mechanoreceptors in the human hand related to touch sensation. *Human Neurobiology* 3(1): 3–14.

Vallbo, Å. B., H. Olausson, and J. Wessberg (1999). Unmyelinated afferents constitute a second system coding tactile stimuli of the human hairy skin. *Journal of Neurophysiology* 81(310): 2753–2763.

Vallbo, Å. B., H. Olausson, J. Wessberg, and N. Kakuda (1995). Receptive field characteristics of tactile units with myelinated afferents in hairy skin of human subjects. *Journal of Physiology* 483 (Pt 3): 783–795.

Vallbo, Å. B., H. Olausson, J. Wessberg, and U. Norrsell (1993). A system of unmyelinated afferents for innocuous mechanoreception in the human skin. *Brain Research* 628(310): 301–304.

Vallbo, Å. B., K. A. Olsson, K. G. Westberg, and F. J. Clark (1984). Microstimulation of single tactile afferents from the human hand. Sensory attributes related to unit type and properties of receptive fields. *Brain* 107 (Pt 3): 727–749.

Vapnik, V. and A. Lerner (1963). Pattern recognition using generalized portrait method. *Automation and Remote Control* 24.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer.

Varney, N. R. and A. L. Benton (1975). Tactile perception of direction in relation to handedness and familial handedness. *Neuropsychologia* 13(4): 449–454.

Wahde, M. and Z. Szallasi (2006). Improving the prediction of the clinical outcome of breast cancer using evolutionary algorithms. *Soft Computing* 10(4): 338–345.

Wang, Z., A. R. Childress, J. Wang, and J. A. Detre (2007). Support vector machine learning-based fMRI data group analysis. *NeuroImage* 15(36): 1139–1151.

Weiskopf, N., F. Scharnowski, R. Veit, R. Goebel, N. Birbaumer, and K. Mathiak (2004). Self-regulation of local brain activity using real-time functional magnetic resonance imaging (fMRI). *Journal of Physiology (Paris)* 98(4-6): 357–373.

Weiskopf, N., R. Veit, M. Erb, K. Mathiak, Grodd W., Goebel R., and Birbaumer N. (2003). Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. *NeuroImage* 19: 577–586.

Wessberg, J., H. Olausson, K. Wiklund Fernström, and Å. B. Vallbo (2003). Receptive field properties of unmyelinated tactile afferents in the human skin. *Journal of Neurophysiology* 89(3): 1567–1575.

Wessberg, J., C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan, and M. A. L. Nicolelis (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* 408: 361–365.

Wiklund Fernström, K. (2004). Physiological properties of unmyelinated low-threshold tactile (CT) afferents in the human hairy skin. Ph.D. diss., University of Gothenburg, Gothenburg, Sweden.

Wild, J. (2005). Brain imaging ready to detect terrorists, say neuroscientists. *Nature* 437(7058): 457.

Willis, W. D. Jr, X. Zhang, C. N. Honda, and G. J. Jr. Giesler (2002). A critical review of the role of the proposed VMpo nucleus in pain. *Journal of Pain* 3(2): 79–94.

Worsley, K. J., S. Marrett, P. Neelin, and A. C. Evans (1992). A three-dimensional statistical analysis for cbf activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism* 12: 900–918.

Worsley, K. J., S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping* 4: 58–73.

Wu, F.-X. (2008). Genetic weighted k-means algorithm for clustering large-scale gene expression data. *BMC Bioinformatics* 9 Suppl 6: S12.

Yanagisawa, T., M. Hirata, Y. Saitoh, A. Kato, D. Shibuya, Y. Kamitani, and T. Yoshimine (2009). Neural decoding using gyral and intrasulcal electrocorticograms. *NeuroImage* 45(4): 1099–1106.

Yoo, S. S., H. M. O'Leary, T. Fairneny, N. K. Chen, L. P. Panych, H. Park, and F. A. Jolesz (2006). Increasing cortical activity in auditory areas through neurofeedback functional magnetic resonance imaging. *Neuroreport* 17(12): 1273–1278.

Zaidi, H. and G. Sgouros (2002). *Therapeutic Applications of Monte Carlo Calculations in Nuclear Medicine*. Taylor & Francis Ltd, London, Great Britain.

Zhang, J., Y. Shang, R. Gao, and Y. Dong (2009). An improved multi-objective adaptive niche genetic algorithm based on pareto front. In *Proceedings of the IEEE International Advance Computing Conference IACC 2009*, pp. 300–304.

Zhu, Z., Y.-S. Ong, and M. Dash (2007a). Markov blanket-embedded genetic algorithm for gene selection. *Pattern Recognition* 40(11): 3236–3248.

Zhu, Z., Y.-S. Ong, and M. Dash (2007b). Wrapper-filter feature selection algorithm using a memetic framework. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics* 37(1): 70–76.

Zotterman, Y. (1939). Touch, pain and tickling: an electro-physiological investigation on cutaneous sensory nerves. *Journal of Physiology* 95(1): 1–28.