We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Computational Intelligence Approaches to Brain Signal Pattern Recognition

Pawel Herman, Girijesh Prasad and Thomas Martin McGinnity
Intelligent Systems Research Centre, University of Ulster
Northern Ireland

1. Introduction

Analysis of electrophysiological brain activity has long been considered as one of indispensable tools enabling clinicians and scientists to investigate various aspects of cognitive brain functionality and its underlying neurophysiological structure. The relevance of electroencephalogram (EEG) in particular, due to its inexpensive and most importantly, non-invasive acquisition procedure, has been reflected in the abundance of clinical applications and the diversity of areas of research studies it has contributed to. These studies lie within the realm of brain science understood nowadays in a broad sense embracing and linking interdisciplinary fields of neuroimaging, cognitive psychology and neurophysiology among others. In medical practice, EEG is used more pragmatically to support clinicians in their effort to establish the presence, severity and cerebral distribution of neurological disorders. Epilepsy diagnostic serves as a prime example in this regard (Fisch, 1999). The complex nature of brain signals and the intricacies of the measurement process involved (Fisch, 1999; Niedermeyer & Lopes da Silva, 2004), particularly in the case of EEG, render their analysis and interpretation challenging (Kaiser, 2005). Historically, these signals used to be examined only qualitatively based on routine visual inspection and the experience of responsible technicians or practitioners. With the advent of the era of digital biosignal recordings, computerised quantitative electroencephalography gained notable popularity as a supplementary tool enhancing objectiveness of analysis (Kaiser, 2005). The fast pace of technological advancement, considerable progress in neuroscience and neuroengineering along with growing investments in medical and health sectors among others have opened up new possibilities for automated EEG analysis systems. A continually growing scope for their applications set dominant design trends and imposed requirements regarding their functionality that prevail in today's practice and research. One of the key points in this regard is the need for the increased independence, autonomy and thus the improved reliability of such systems. This has led to a more comprehensive formulation of a computational problem of brain signal analysis within the realm of pattern recognition, which facilitates a more generic description of existing approaches, and development or re-use of suitable pattern recognition methods. In consequence, the notion of brain signal pattern recognition has been introduced to refer to the underlying concept of processing raw data with the aim of acting upon its category (Niedermeyer & Lopes da Silva, 2004; Duda et al., 2001). The objective is to identify patterns in electrophysiological brain activity that are

indicative of various cognitive brain states (Niedermeyer & Lopes da Silva, 2004). The demanding nature of this task is here emphasised due to the spatio-temporal complexity of brain signal dynamics and low signal-to-noise ratio, particularly in the case of EEG. In order to ensure robust recognition of relevant brain signal correlates, as required in automated brain signal analysis, the major challenges should be identified. One of the most urgent needs is to robustly account for uncertain information inherent to biological data sources. The uncertainty effects arise mainly out of stochastic nature of signal acquisition processes and nondeterministic characteristics of the underlying neurophysiological phenomena, which cannot be accurately explained by any biologically plausible model but are attributed to the existence of a general biological tendency to undergo changes and transitions (Fisch, 1999; Wolpaw et al., 2002). In this regard, the multitude of behavioural, cognitive and psycho-emotional or physiological factors play a substantial contributory role. The resultant uncertainty manifestations are rarely dealt with in an explicit and effective manner. It should be realised though that ignoring them or adopting simplistic assumptions may undermine the concept of robust brain signal analysis.

This chapter is concentrated on a particular instance of brain signal pattern recognition, where uncertainty and variability effects manifest themselves with relatively high intensity. More precisely, the problem of classification of spontaneous electrophysiological brain activity when the subject is voluntarily performing specific cognitive tasks is examined. This deliberate control of thoughts provides a scope for a communication channel between the brain and the external environment with brain signals being the carrier of information. Such an alternative form of communication, independent of peripheral nerves and muscles, underpins the concept of the so-called brain-computer interface (BCI). Thus, the outcome of the study reported in this chapter bears direct relevance and has intrinsic implications for a broad area of BCI. This work follows the prevailing trends in BCI and is focused on the discrimination between self-induced imaginations of right and left hand movements, referred to as motor imagery (MI), based on analysis of the associated EEG correlates. The essence of uncertainty manifestations in this challenging brain signal pattern recognition problem and a range of existing approaches adopted to minimise the associated detrimental effects on BCI performance are discussed in section 2. Then, in section 3 a computational intelligence methodology is briefly introduced with emphasis on fuzzy logic (FL) paradigms in the context of pattern recognition under uncertainty. Section 4 describes methods developed and employed in this work to address a given MI-based brain signal pattern recognition problem. It also reveals the details of a BCI experimental procedure allowing for MI related EEG data acquisition. A comparative analysis of the results obtained with novel approaches introduced in this chapter and with more conventional BCI techniques is reported in section 5. Final conclusions and summary of the chapter are included in section 6. The key directions for further work are also suggested.

2. Uncertainty effects in EEG-based BCI

Uncertainty as an inseparable feature of BCI operation needs to be properly addressed in order to develop practical and robust systems (Wolpaw et al., 2002). Effective handling of uncertainty effects, strongly reflected in EEG signals, has been recognised as a key challenge in BCI (Vaughan et al., 2003). These effects have been traditionally associated with inherent changes in the underlying brain dynamics and varying physical characteristics of the signal measurement environment (Wolpaw et al., 2002; Vaughan et al., 2003; Millán et al., 2003).

There are a number of behavioural and neurophysiological factors that determine the character of transitions between cognitive brain states. Consequently, electrophysiological signals display a degree of inconsistency due to a varying level of subject's awareness, mental focus, motivation and fatigue among others (Wolpaw et al., 2002). In addition, the brain plasticity harnessed by the mechanism of neurofeedback involved in BCI operation[*] inevitably produces changes in the brain's behaviour. With regard to signal recording environment, it has been reported that inter-session changes in EEG cap placement (McFarland et al., 1997) or in the impedance of scalp-electrode interface (Sykacek et al., 2004) may affect BCI performance.

Uncertainty in the space of brain state categories poses another challenge in BCI. It arises out of intrinsic ambiguity and vagueness in interpretation of different brain states correlated with specific cognitive tasks, no matter how well they are defined. It is hard to assume that there is a crisp unequivocal association between characteristic patterns of brain's electrophysiological activity and classes of particular mental tasks. As suggested in (Yang et al., 2007), a mixture of some residual correlates of different cognitive processes should always be expected. This facet of uncertainty related to brain state class assignments is perceived as an inherent feature of brain signal pattern recognition.

Regardless of the sources of variability in BCI, it is predominantly reflected in electrophysiological brain signals, particularly in EEG, in the form of nonstationarity effects at different temporal levels. Their manifestations are present in any low-dimensional EEG feature space and are difficult to model analytically due to limitations in today's understanding of the underlying brain phenomena. Thus, their handling is considered as a challenging task and poses an urgent objective in the presence of numerous literature reports on a detrimental impact of EEG nonstationarities on the performance of BCI systems. In (Cheng et al., 2004), significant discrepancies in the distribution of EEG power features, extracted from data sets acquired at different times than the original training data set, were observed to result in a relatively poor accuracy of linear classifiers employed in an MI-based BCI. A similar inter-session deterioration of the performance of a linear discriminant analysis (LDA) classifier was reported in (Obermaier et al., 2001). The authors concluded that the LDA method did not provide capabilities to generalise MI induced spatio-temporal patterns in EEGs. In (Townsend et al., 2006), special attention was paid to inconsistencies in machine learning-based selection of the most relevant discriminative EEG feature components within the same BCI data set. Analogous incoherence in the localisation of optimal electrodes and in the identification of the most reactive EEG rhythms providing the maximum distinguishability of MI related EEG signals was described in (Pregenzer & Pfurtscheller, 1999). The changes were particularly noticeable in feedback sessions. Shenoy et al. (2006) and Vidaurre et al. (2006) made an attempt to graphically illustrate session-to-session nonstationarities in different EEG feature spaces by performing their two-dimensional projections. Although the projection approaches adopted in (Shenoy et al.,

---

[*] In the context of EEG-based BCI, subjects receive mostly visual, auditory or haptic feedback information about their brain activity reflected in the EEG. It conveys the degree of success in voluntary control of the brain activity. Thus, the feedback signal has an important motivational role as it facilitates higher attention levels or otherwise causes frustration and confusion if it is unreliable (McFarland et al., 1998).

2006) and (Vidaurre et al., 2006) were distinct, the conclusions were the same. Namely, inter-session discrepancies between clusters of the features representing the classes of associated MIs were clearly identified. Schlögl et al. (2005) analysed several types of EEG nonstationarities in BCI experiments using the state-of-the-art adaptive autoregressive (AR) features and an LDA classifier. The effect of both short- and long-term variability of the EEG dynamics was reflected in the considerable inconsistency of BCI performance.

There has been some empirical evidence gathered (Millán et al., 2002; Pfurtscheller et al., 2000; Guger et al., 2001; Pfurtscheller & Neuper, 2001) that in the face of the problem of session-to-session performance transfer it is beneficial to update or re-train a BCI classifier on a regular basis using the most recent data from one or a few past sessions. Still, the effectiveness of this method is limited as presented in (Shenoy et al., 2006; Guger et al., 2001) using linear classification approaches. In addition, it appears rather impractical considering the automated nature of BCI systems. The burden associated with their frequent re-calibration can be partly mitigated by computationally efficient algorithms for BCI prototyping and incorporating necessary modifications, as suggested in (Guger et al., 2001). Despite the shortcomings discussed here, this practice of regular BCI update has been a traditional approach to the problem of inter-session variability in BCI and it is still widely utilised.

There has also been considerable research conducted on adaptive BCI classification (Sykacek et al., 2004; Vidaurre et al., 2006; Millán et al., 2003) in the spirit of Wolpaw's principle of adaptive human (brain)–machine interaction (Wolpaw et al., 2002). Unlike the approach involving frequent off-line BCI re-calibration, adaptive systems are updated nearly instantaneously in on-line mode. Some of them have demonstrated the enhanced potential in handling uncertainty effects and have thus led to improved BCI performance than regularly re-trained but static linear, quadratic and probabilistic classifiers (Sykacek et al., 2004; Vidaurre et al., 2006; Shenoy et al., 2006). It should be noted however that the focus of adaptive BCI has been on reducing the effect of spontaneous EEG variations, and thus on handling short-term within-session changes of the signal dynamics. In consequence, the concept of continuous on-line update (Vidaurre et al., 2006; Sykacek et al., 2004) is likely to result in undesirable excessive detuning of a BCI classifier under the conditions of acute variability when handling short-lived transients, as indicated in (Vaughan et al., 2003). Yet, it does not necessarily address the problem of long-term changes in the EEG dynamics, particularly in a session-to-session scenario. Moreover, on-line adaptive classifiers are generally developed under the assumption of a known type of the signal's feature distribution, which may not be satisfied increasing the risk of lower accuracy.

It has become clear that various manifestations of uncertainty inherent to brain signal pattern recognition constitute a serious challenge in BCI research. The problem of maintaining good BCI performance over a reasonably long period in spite of the presence of these effects has not yet been effectively addressed using classical signal processing techniques, statistical pattern recognition methods or machine learning approaches. In the next section, advantages of a different methodological paradigm referred to as computational intelligence, eg. (Gorzalczany, 2002), with emphasis on computing with fuzzy sets (FSs), in application to pattern recognition under uncertainty are outlined. Special attention is given to an emerging type-2 (T2) fuzzy logic (FL) methodology (Mendel, 2001) due to its enhanced uncertainty handling capabilities.

3. Computational intelligence in pattern recognition

As discussed earlier, the uncertainty effects inherent to brain signal pattern recognition have a multi-faceted nature. In the context of EEG-based BCI, nonstationarity of EEG dynamics reveals nondeterministic characteristics of the underlying data generation mechanism, and thus it is not suitable for analytical modelling. It is also difficult to make valid statistical inference about its probabilistic features. In the realm of uncertainty analysis, there appears a group of methods that have demonstrated true potential in dealing with complexity and uncertainty in numerical data without any underlying physical model of their generator. Such a model-free approach can be adopted using computational intelligence paradigms (Mendel, 2001; Gorzalczany, 2002). They allow for data-driven design of computational systems that are capable of generalising knowledge, performing abstract associations and inference using approximate reasoning even in the presence of vague, ambiguous or imprecise information in ill-structured environments, and thus providing robust low-cost solutions to real-world problems (Mendel, 2001). Pattern recognition naturally lends itself as an application domain for computational intelligence. When uncertainty is strongly manifested in a given class of problems, FL methodology and the related FS theory are of special relevance. With a suitable system framework, the transparency of inference methods and mechanisms, and the flexibility of available data-driven design methods, this computational intelligence tool offers considerable potential in the context of uncertainty management in brain signal pattern recognition.

Recently, new directions in FL development have been explored to further enhance uncertainty modelling apparatus of conventional type-1 (T1) FL systems (FLSs), eg. (Karnik et al., 1999; Mendel, 2001). As a result, the notion of an extended type-2 (T2) FS with an additional dimension of fuzziness has received growing research attention and the corresponding T2FL uncertainty calculus has been shown to outperform its classical T1 counterpart in practical applications (Mendel, 2001). Thus, T2FL methodology appears to be a promising approach to the challenging brain signal pattern recognition problem undertaken in this work. Below, fundamental concepts in the area of T2FL related to this work are briefly presented.

At the heart of T2FL lies the definition of a T2FS originally introduced by Zadeh (1975) as an extension or a fuzzy version of a classical T1FS. This additional level of fuzziness is associated with another dimension in the definition of a T2FS. As a result, instead of being two-dimensional, a T2FS $\tilde{A}$ is three-dimensional and the membership grade defined in (1) for any given $x^* \in U_X$ (U_X is a domain, also called a universe of discourse) is an ordinary FS with the membership function $\mu_{\tilde{A}}(x^*, u), u \in J_x \subseteq [0,1]$ (J_x is the primary membership of x), not a crisp number $\mu_A(x^*)$ in [0,1] as in a classical T1FS A (c.f. (2)):

$$\tilde{A} = \left\{ ((x,u), \mu_{\tilde{A}}(x,u)) \mid \forall x \in U_X, \forall u \in J_x \subseteq [0,1] \right\}. \tag{1}$$

$$A = \left\{ (x, \mu_A(x)) \mid \forall x \in U_X, \mu_A \in [0,1] \right\}. \tag{2}$$

The domain of support for membership functions in T2FS representation is two-dimensional and is often referred to as a foot of uncertainty (FOU) (Mendel, 2001). Since it is effectively the union of all $J_x, \forall x \in U_X$, the FOU allows for embedding a range of T1FSs. The resultant

extra degrees of freedom facilitate capturing more information about the represented term than a single T1FS can and thus render FOU particularly important in handling inconsistently varying information content. This enhanced flexibility in modelling the associated uncertainty underlies the potential of T2FLSs to outperform their T1 counterparts in problems where classification or approximation is to be made under uncertain, variable conditions.

On the other hand, T2FLSs are more computationally expensive. This overhead can be reduced by exploiting the so-called interval T2FSs (IT2FSs) (Liang & Mendel, 2000). Their membership functions over the FOU are constant and equal one (Mendel, 2001). This substantially simplifies operations on FSs, which now amount to interval-type operations (Liang & Mendel, 2000; Gorzalczany, 1988) on the associated FOUs, and facilitates transparent flow of uncertainties through a T2FLS. Moreover, the use of IT2FSs has proven to be beneficial in practical applications (Mendel, 2001). FOUs of the two most common Gaussian IT2FS, with uncertain mean, m, but fixed standard deviation, σ, and with fixed mean and uncertain standard deviation, are depicted in Fig. 1a-b. Since they embed conventional T1FSs A_e, as mentioned earlier, these FOUs can be easily parameterised with T1 membership functions, respectively (with m_1, m_2, σ_1, σ_2 defining the ranges of parameter variations):

$$\mu_{A_e}(x) = \exp\left[\frac{(x-m)^2}{-2\sigma^2}\right], \ m \in [m_1, m_2], \ \sigma \text{ fixed}, \tag{3a}$$

$$\mu_{A_e}(x) = \exp\left[\frac{(x-m)^2}{-2\sigma^2}\right], \ \sigma \in [\sigma_1, \sigma_2], \ m \text{ fixed}. \tag{3b}$$
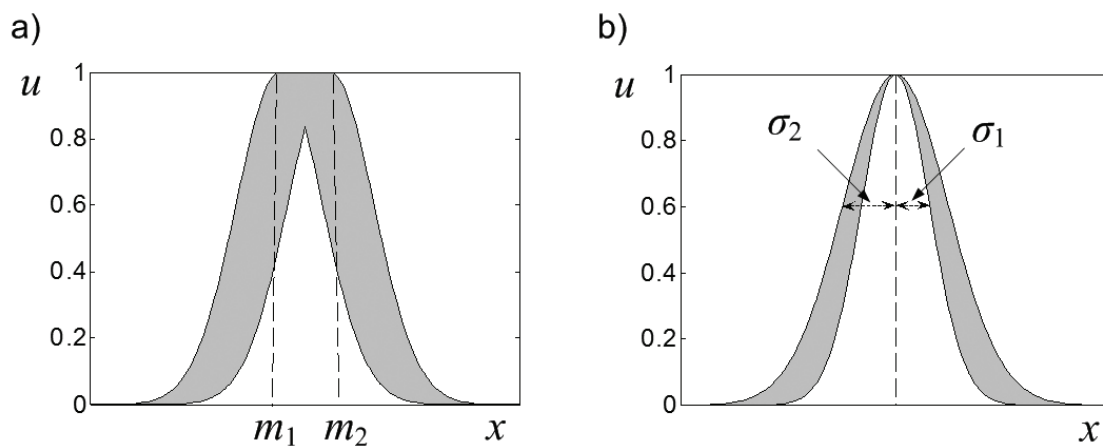

Fig. 1. An illustration of the FOUs of Gaussian T2FSs a) with uncertain mean and fixed standard deviation, b) with fixed mean and uncertain standard deviation.

Architecture of a T2FLS is analogous to that of its T1FLS counterpart with the difference in the type of FS representation in the antecedents and consequents of fuzzy rules, and in FS operators. In consequence, since the result of T2FL inference is a T2FS, the process of obtaining a crisp value from a final FLS output involves an additional step in T2FLSs when compared to T1FLSs. To this end, type reduction is applied to reduce a T2FS to a

T1FS before it is ultimately defuzzified using classical fuzzy methods. Type reduction constitutes the computational bottleneck in interval T2FLSs (IT2FLSs) (Mendel, 2001; Liang & Mendel, 2000). There are a number of type reduction approaches including approximate techniques reported in the fuzzy literature with centre-of-sets and centroid type reduction being the most popular (Karnik et al., 1999; Mendel, 2001). The entire process of information flow through a T2FLS can be summarised by the following sequence (c.f. Fig. 2):

1. Fuzzification (optional) – transforming a crisp input value to a T1FS or a T2FS.
2. Inference using a compositional rule (Mendel, 2001) involving the system input (fuzzified) and fuzzy rule base relations.
3. Aggregation of the resultant T2FSs obtained from different rules in the process of inference (in some cases, aggregation is considered as part of the inference process).
4. Type reduction, eg. by evaluating the centroid or the centre-of-sets of the aggregated output T2FS.
5. Defuzzification of the T1FS obtained in 4) (optional) to extract a crisp output.
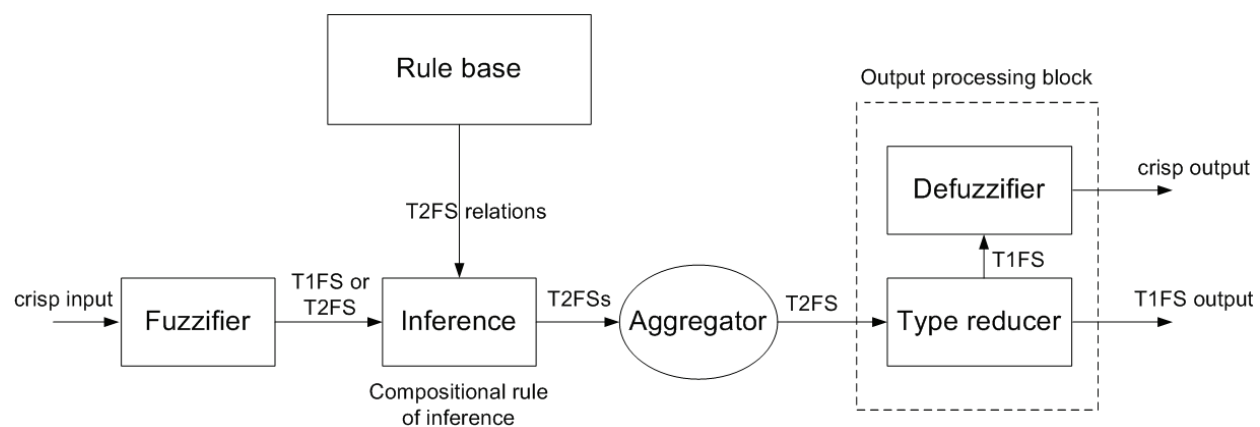

Fig. 2. T2FLS framework.

In the context of the brain signal pattern recognition problem considered in this chapter, it is expected that the increased flexibility of IT2FSs in modelling uncertainty can be effectively utilised to encapsulate the range of possible behaviours of brain signal dynamics correlated with MI and thus robustly account for the associated variability. Consequently, the central objective is to examine the potential of a novel IT2FLS-based approach to dichotomous classification of MI induced EEG patterns. The emphasis is on the classifier's capability to generalise well across a few data sets obtained at different times (exhibiting mainly long-term changes). At the same time, it should be realised that despite the early progress in the domain of applied T2FL, there has been rather limited research done on systematic approaches to data-driven design of IT2FLSs used in pattern recognition. This chapter also outlines some developments that address this emerging need and discusses key issues related to the effective exploitation of IT2FLS's uncertainty handling apparatus in the given instance of brain signal pattern recognition. Automation of the fuzzy classifier design process is intended and to this end, its computationally efficient implementation is proposed. A detailed description of the BCI experimental setup and the pattern recognition methods devised and employed in this work are presented in the next section.

4. Methods and experimental work

4.1 Experimental setup and data acquisition

In the presented work, EEG data acquired in BCI experiments in two different labs were utilised. The first data set was obtained from the Institute of Human–Computer Interfaces, Graz University of Technology. The EEG signals were recorded from three healthy subjects (S1, S2 and S3) in a timed experimental recording procedure where the subjects were instructed to imagine moving the left and the right hand in accordance with a directional cue displayed on a computer monitor (Fig. 3a). Each trial was 8 s in length. A fixation cross was displayed from $t = 0$ s to $t = 3$ s. The beginning of a trial was marked by acoustic stimulus at $t = 2$ s. Next, an arrow (left or right) was displayed as a cue at $t = 3$ s. Therefore the segment of the data recorded after $t = 3$ s of each trial was considered as event related and was used for off-line analysis. The recordings were made with a g.tec amplifier (http://www.gtec.at) and AgCl electrodes over two consecutive sessions, each session consisting of 140 trials for S1 and 160 trials for S2 and S3 with equal number of trials representing two MI classes (Wang et al., 2004). Two bipolar EEG channels were measured over C3 and C4 locations (two electrodes placed 2.5 cm anterior and posterior to positions C3 and C4) according to the international standard nomenclature (10/20 system) (Niedermeyer & Lopes da Silva, 2004). The EEGs were then sampled at a frequency of 128 Hz and band-pass filtered in the frequency range 0.5–30 Hz.

The second EEG data set was acquired at the Intelligent Systems Research Centre (ISRC), University of Ulster using the same g.tec equipment and the location of two bipolar channel electrodes as that used by the Graz BCI group. The EEG data were recorded from six healthy subjects (S_I–S_{VIII}) over ten 160-trial (balanced) sessions with a sampling frequency of 125 Hz. Depending on the subject, first one or two sessions were conducted without neurofeedback, and to this end, a directional cue following a fixation cross was displayed in the form of an arrow pointing to left or right to instruct a subject which MI should be carried out, as in the Graz paradigm. In the subsequent feedback sessions, the game-like basket paradigm was employed. In each trial of 7 s duration, two baskets were displayed at $t = 3$ s at the bottom of the screen in the form of bars – the target basket in green and the non-target one in red. Subjects were asked to perform MI that allowed them through the BCI to direct a ball falling from the top of the screen for the last 3 s of a trial to the target basket. The ball movement was continuously (in real-time) controlled in a horizontal direction from $t = 4$ s to $t = 7$ s utilising the proposed fuzzy classifier's output signal, which served as BCI feedback. The timing and a graph illustrating the concept of this paradigm are presented in Fig. 3b.

Although the EEG data sets under consideration were originally recorded in on-line BCI paradigms with continuous classification, they were also exploited in the context of off-line discrete classification of entire trials. As a result, two separate BCI study cases were investigated in this work, with continuous on-line (only on the ISRC data set) and discrete off-line application of an IT2FLS classifier. Still, it should be emphasised that they share similar characteristics of MI related brain signal pattern recognition with slightly different aspects of the uncertainty effects being exposed in each case (see section 5 for more discussion). From the perspective of a signal processing methodology, the major difference lies in the way that temporal information is handled at the feature extraction stage (c.f. section 4.2). Moreover, on-line verification of BCI classification performance raises additional issues related to instantaneous neurofeedback delivery, which are not taken into account in a post-hoc off-line simulation. They are given more attention in section 4.3.3.
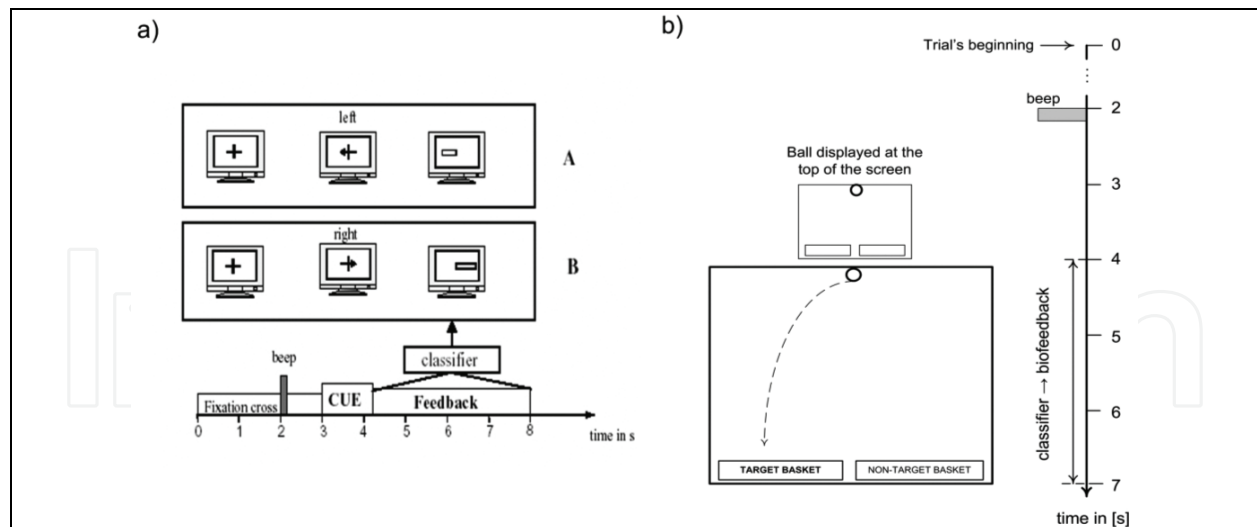
Fig. 3. Data recording in a) Graz BCI paradigm (Haselsteiner & Pfurtscheller, 2000) and b) BCI basket paradigm (Wang et al., 2004).

It should also be mentioned that only a few final sessions when individual subjects acquired an acceptable level of BCI control were closely examined and evaluated in the study reported in this chapter. The data gathered during earlier training sessions were exploited in most cases to pre-calibrate BCI methods and conduct preliminary off-line analyses.

4.2 EEG feature extraction and analysis

Sensorimotor rhythms represent the most discernible and reliable category of EEG correlates of MI induced brain phenomena (Vaughan et al., 2003; McFarland et al., 1997). Thus, brain signal patterns considered in this work are derived from mu (μ) and beta (β) rhythms of spontaneous EEG activity over the specified sensorimotor areas (C3 and C4 locations, c.f. section 4.1). In particular, the imagination of hand movement causes activation of the brain's motor cortex that is usually manifested in the interplay between contralateral attenuation of the μ rhythm and ipsilateral enhancement of the central β oscillations in different phases of MI. These processes occur due to the neurophysiological mechanisms of the so-called event-related desynchronization (ERD) and event-related synchronization (ERS) (Niedermeyer & Lopes da Silva, 2004). The exact sensorimotor EEG patterns and the most reactive frequency bands of ERS and ERD vary from subject to subject. Preliminary analysis performed in this work confirmed that overall, ERD manifestations in the μ range could be observed on the contralateral side and a slight ERS in the central β rhythm on the ipsilateral hemisphere. This hemispheric lateralisation of the oscillatory brain signal patterns underlies discrimination between the left and right MIs. In consequence, methods of spectral analysis played a dominant role in the process of EEG quantification conducted in this work to extract discriminative signal features.

As mentioned in section 4.1, the problem of MI related brain signal pattern recognition was addressed in two modes – with discrete classification of entire EEG trials and instantaneous discrimination within a trial. The main difference between these two BCI approaches lies in the temporal characteristics of a feature extraction protocol. Consequently, handling and quantification of the relevant spatio-temporal EEG patterns requires distinct approaches. They are described in two subsequent sections.

4.2.1 Off-line analysis of spectral EEG patterns

In off-line discrete classification each EEG trial is represented as a single feature vector. To this end, the event-related segment (starting from $t = 3$ s) of length $N = 5*128 = 640$ samples for the Graz data set and $N = 4*125 = 500$ samples for the ISRC data set was divided into rectangular windows depending on the settings of two parameters: window length, *win_len*, and the amount of overlap, *ovl*. Next, the frequency-related information was independently extracted from each of n_{win} windows (c.f. (5)) and the relevant spectral correlates of ERD and ERS phenomena were quantified. In particular, the μ and β bandpower components were merged within each window to constitute a feature vector element, r_i^j ($i=1,..,n_{win}$) given two recording channels, $j \in \{C3, C4\}$. The entire feature vector $\boldsymbol{r}$ representing an EEG trial was composed of $2n_{win}$ such components:

$$\boldsymbol{r} = \left(r_1^{C3}, r_2^{C3}, .., r_{n_{win}}^{C3}, r_1^{C4}, r_2^{C4}, .., r_{n_{win}}^{C4} \right),$$ (4)

where:

$$n_{win} = \left\lfloor \frac{N - ovl}{win\_len - ovl} \right\rfloor.$$ (5)

In the preliminary analysis reported in (Herman et al., 2008a), a wide range of spectral methods such as power spectral density (PSD) estimation techniques (Stoica & Moses, 1997), atomic decompositions including short-time Fourier transform (STFT) (Stoica & Moses, 1997) and S-transform (Assous et al., 2006), quadratic time-frequency energy distributions and wavelet-based methods (Akay, 1997) were thoroughly examined in the given brain signal pattern recognition problem. They were all employed within the same window-based feature extraction framework to obtain signal's bandpower components in the μ and central β ranges. The resultant low-dimension feature representations (c.f. (4)) were assessed in terms of their discriminative properties quantified using the classification accuracy (CA) rate obtained with popular linear and nonlinear BCI classifiers (c.f. section 4.3.2). Since PSD approaches were demonstrated overall to deliver consistently superior performance in within-session and inter-session classification scenarios, this category of spectral quantification methods was exploited in this work. In particular, nonparametric periodogram (Stoica & Moses, 1997) and parametric PSD estimate using Yule-Walker algorithm (Haykin, 1996) were applied depending on the subject. The exact frequency bands within the μ and central β ranges, from which bandpower components were extracted, were tuned individually for each subject to maximise linear separability between the resultant feature vectors representing two-class MI related EEG trials. To this end, linear discriminative analysis (LDA) (Bishop, 1995) was conducted on the initial calibration data within a cross-validation (CV) framework.

In order to demonstrate the problem of variability in BCI, discussed in section 2, session-to-session changes in the distribution of class-specific EEG features acquired from one of the subjects under consideration are presented in Fig. 4. In particular, the feature space was projected on the principal components (PC) axes. PC analysis (PCA) was performed with one session (I) as the reference and the data in the other session (II) were transformed according to this new set of directions of the largest variance. For illustrative purposes, only the first two components accounting for over 70% of the total variance are shown. Apart

from the projected two-dimensional feature samples, their means and standard deviations, estimated in each class after removing the most noticeable outliers, are depicted. The standard deviations presented in the form of ellipses centred at the corresponding means were scaled down to enhance the clarity of the illustrations.

Several relevant observations can be made based on the proposed analysis. Firstly, largely overlapping regions of the projected feature space corresponding to different MI classes are evident. Secondly, the inter-session shifts of the class means for both left MI and right MI groups are strongly manifested in the given data set (c.f. 4a-b). They are indicative of the variability effects inherent to BCI as discussed in section 2. Since there is no underlying model of these changes and due to their inconsistent nature, reported in a multi-session analysis, the issue of uncertainty arises and renders this brain signal pattern recognition problem particularly challenging.
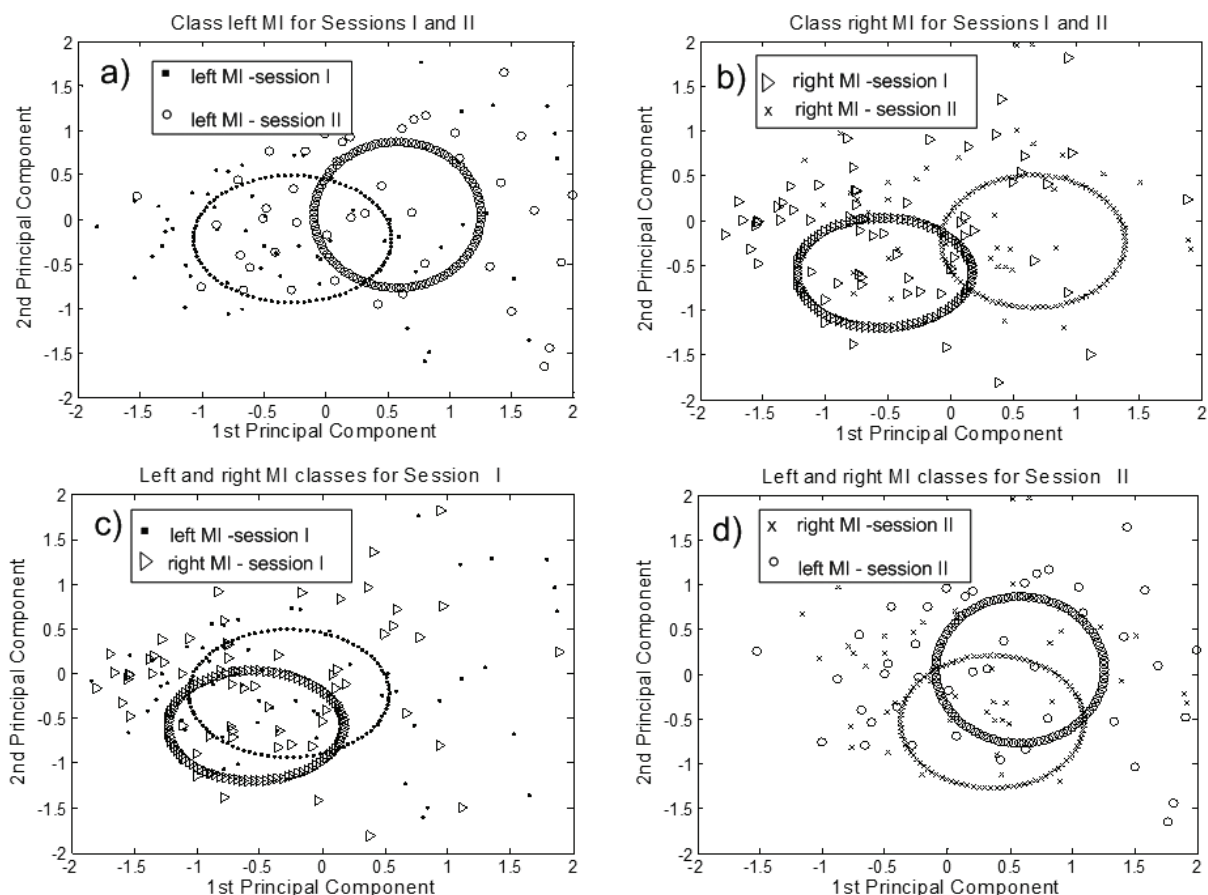

Fig. 4. The distribution of spectral EEG features in two-dimensional normalised PCs' space with their corresponding class means and scaled standard deviations: a) left and b) right MI features in sessions I and II plus within-session feature distribution: c) session I and d) session II.

4.2.2 Feature extraction for on-line BCI

As discussed in section 4.2, on-line BCI was implemented in continuous mode. In other words, EEG features were extracted and classified instantaneously within a trial, which led to as many classifications per trial as the number of its even-related data samples (considering that the length of the event-related part of a trial in a basket paradigm was 4 s,

there were $4 * 125 = 500$ relevant applications of a feature extractor and a classifier). To this end, a sliding window approach was adopted within a causal framework. In consequence, the window acts as a buffer and introduces a delay with respect to the temporal occurrence of relevant MI correlates in the signal examined. The window sizes used in this work were identified with a view to compromising the time resolution of BCI control (reactivity) and the MI related content of spontaneous EEG activity. The delay was found to be acceptable in on-line operation and its effect could only be felt at the trial's onset.

Three alternative techniques of spectral analysis were utilised in this study to suit individual cases. Similarly as in the earlier study involving discrete classification of entire trials, PSD approaches, Welch periodogram and Yule-Walker's parametric PSD estimation in particular, were found to facilitate consistent and robust BCI performance. Additionally, for a small proportion of subjects, STFT was demonstrated in off-line preliminary analyses to lead to higher CA rates than those reported with PSD techniques. Therefore, the identification of an optimal feature type extractor for the on-line use was subject specific.

The spectral methods just mentioned were employed to extract bandpower information from EEGs in the frequency ranges related to the ERD/ERS phenomena. Due to distinct temporal scales of signal representation in continuous feature extraction and in a discrete approach (with an entire trial being represented as a feature vector), the relation between the quantified oscillatory components in the μ and β bands had different characteristics in both cases. Although spectral contributions from the two relevant frequency ranges were proven in the study reported in section 4.2.1 to provide more discriminative feature representation when merged together, in the preliminary off-line simulation of continuous BCI it was demonstrated that treating these ERD and ERS correlates separately, as independent feature components, led in the clear majority of cases to better classification results. Moreover, it was concluded that normalizing the resultant feature vector r (c.f. (5)), extracted on a sample-by-sample basis (the window was shifted at the sampling rate, i.e. every 8 ms), by its Euclidean length facilitated handling the variance of the signal's energy.

$$r = \left(r_\mu^{C3}, r_\beta^{C3}, r_\mu^{C4}, r_\beta^{C4} \right), \tag{5}$$

where $r_{\mu(\beta)}^{C3(C4)}$ corresponds to the spectral feature component extracted from the adjusted μ (or β band) from the EEG channel C3 (or C4). The instantaneous feature extraction procedure is schematically illustrated in Fig. 5.

4.3 Classification of EEG trials

Classification constitutes another phase of recognition of brain signal patterns allowing for a categorical interpretation of EEG relying on its feature representation. In the context of the work reported in this chapter, the aim of BCI classification is to assign signal trials to the classes of the associated mental tasks (MIs). This given instance of brain signal pattern recognition is dichotomous since an imagination of left hand movement is to be distinguished from an imagination of right hand movement. As discussed earlier, the problem is challenging mainly due to strong EEG nonstationarity effects manifested even in low-dimensional feature spaces. The resultant inter-session variability in the feature distributions was demonstrated in section 4.2.1. In consequence, the study on single trial classification in discrete mode was aimed at effective dealing with these long-term changes in
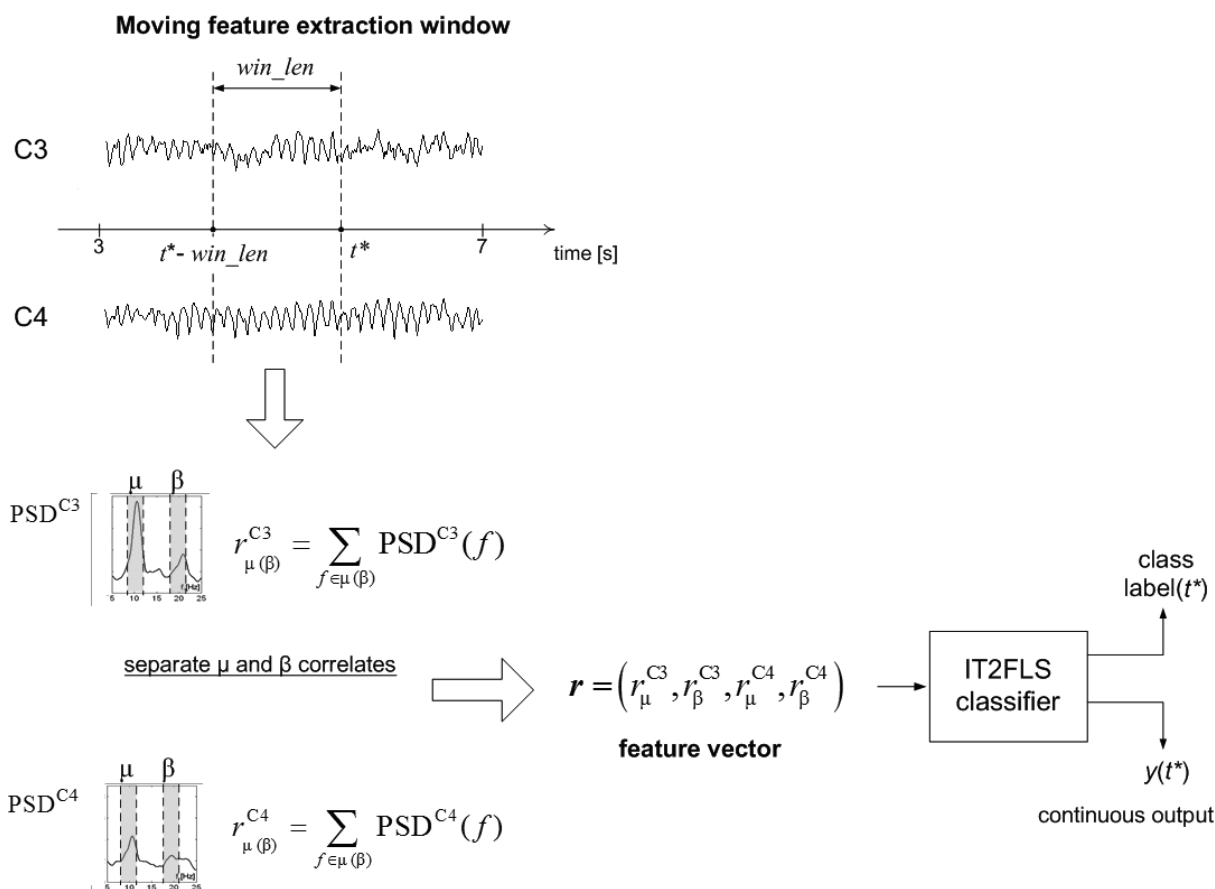
Fig. 5. Graph illustrating the proposed concept of instantaneous BCI feature extraction and classification.

EEG spectral patterns correlated with MI. A successful method is expected to maintain a satisfactory accuracy rate over a few sessions recorded with around one-week break in between without the need for frequent inter-session adjustments. The shorter-term within-trial manifestations are also reported difficult to handle in BCI experiments (Wolpaw et al., 2002; Vaughan et al., 2003; Sykacek et al., 2004). In this work, they could be observed in the study involving instantaneous BCI operation. The intrinsic characteristics of discrete and continuous BCI classification are discussed in section 4.3.3.

In conclusion, the concept of robust brain signal pattern recognition is linked to the key issue of uncertainty in a broader sense, as elaborated in section 2. The emphasis is on handling its multi-faceted manifestations at the classification stage. In order to address this urgent challenge, a novel fuzzy BCI classifier was proposed (c.f. section 4.3.1) and its inter-session performance was compared to that of more traditional BCI approaches: LDA and support vector machines (SVMs) (Cristianini & Shawe-Taylor, 2000) (c.f. section 4.3.2). The CA rate was used as an objective measure in this evaluation.

4.3.1 Fuzzy classification

As elaborated in section 3, T2FLS framework offers more flexibility in handling uncertain information content than its T1 counterpart. It should be emphasised however that in order to appropriately exploit the T2FL apparatus for handling uncertainty without sacrificing its generalisation capability, special care is required in T2FLS development. Therefore,

considerable effort was devoted in this work to devise effective techniques for a fuzzy classifier design. For faster computations, IT2FSs were employed in the construction of a Mamdani-type rule base (Mendel, 2001) (c.f. section 2). The following template of a fuzzy rule was adopted:

$$\text{IF } X_1 \text{ is } \tilde{A}_1 \text{ AND...AND } X_n \text{ is } \tilde{A}_n \text{ THEN } \textit{class} \text{ is } [c_{left}, c_{right}], \tag{6}$$

where fuzzy variables $X_1, ..., X_n$ correspond to the fuzzified components of an input feature vector $r = (r_1, ..., r_n)$, n is their number and $\tilde{A}_1, ..., \tilde{A}_n$ denote IT2FSs with uncertain means (c.f. Fig. 1a, section 3) that serve as the rule antecedents. C is the centroid of the consequent T2FS (in the form of a rectangular T1FS) representing the class that the input feature vector is assigned to. As a result, the rule base models uncertainty related to the variability of EEG features, as discussed in section 4.2.1 (c.f. Fig. 4), and the vagueness or ambiguity of a crisp MI label, i.e. left (associated with numerical value -1) versus right (value 1), c.f. section 2. When $\tilde{A}_i$'s are replaced by T1FSs and C becomes a crisp centroid of a T1FS, the T2 fuzzy rule reduces to the T1 rule format with limited capacity to account for the aforementioned types of uncertain information. The input features to both fuzzy classifiers are represented as T1FSs (fuzzification) to model stationary uniform noise present in the feature space (with standard deviation s_{fuzz_inp}). Gaussian type of FSs was used in the proposed design to facilitate gradient-based tuning. Fig. 6 illustratively juxtaposes the T1FL and T2FL rule pattern (for one-dimensional input) adopted in the reported study.

The IT2FLS classifier was developed in a two-stage procedure, inspired by general FLS design methodology. Firstly, an initial fuzzy rule base was identified and secondly, its parameters were tuned using a global optimisation approach. The design was conducted on a so-called calibration data set, split into a validation and a training subset. The final evaluation was performed on an unseen test data set. In most cases, the calibration and test data sets were taken from independent sessions.
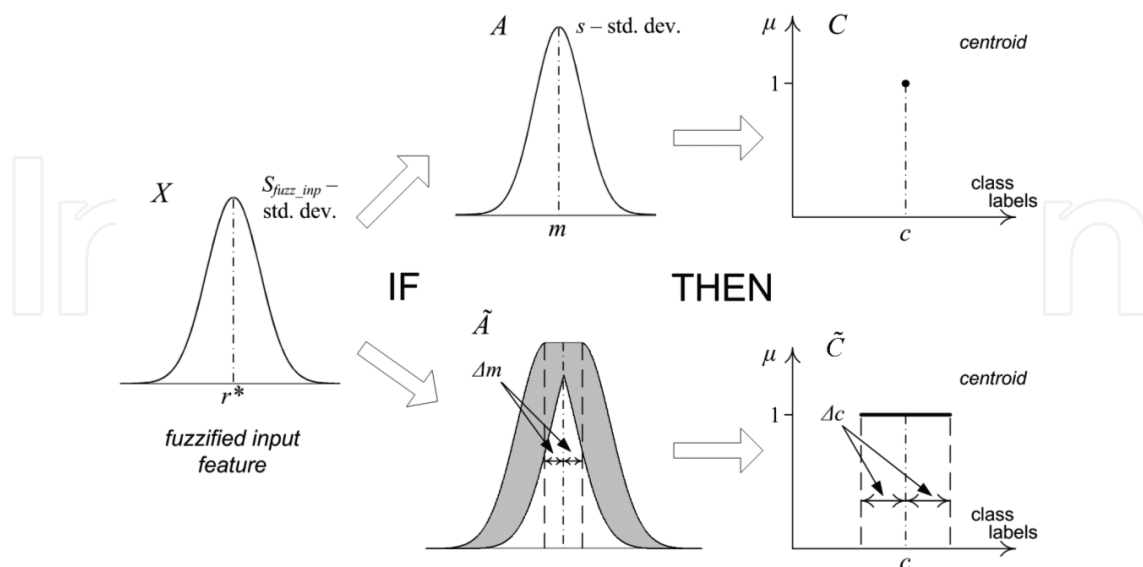

Fig. 6. Illustrative comparison of T1FL and T2FL rule patterns.

An initial fuzzy rule base was identified using a partitioning approach. In other words, the input space was divided into regions accounting for the underlying distribution of a

training set of EEG features with the main objective to obtain a compact data representation that captures their salient characteristics and preserves the inherent input-output relationship (class assignments). Thus, a general clustering approach was adopted to construct a conventional prototype T1FLS rule base that could be extended to serve as an initial T2FLS framework (Herman et al., 2008c). Several clustering methods were examined to identify an optimal design strategy. To this end, a simple heuristic for their initial evaluation was developed. The resultant cluster validity index was primarily used as a criterion for selecting an optimal set of parameters for the initialisation schemes under consideration. It was based on the performance of a prototype (untrained) singleton T1FLS classifier derived directly from the given cluster structure on the calibration data set without any extra parameters, as described later in this section. A final comparative evaluation of the initialisation techniques was conducted within the entire design framework, i.e. in combination with a parameter tuning phase. In consequence, the CA rates obtained with fully trained T1FLSs and with T2FLSs in within-session CV and inter-session classification served as a final performance measure. The outcome of this analysis is discussed in section 5.1. Below, the fuzzy rule base initialisation methods investigated in this work are outlined.

Firstly, a mapping-constrained agglomerative (MCA) clustering algorithm was employed to reinforce the consistency in the mapping from the input to the output space. It has been proven to be robust in the presence of noise and outliers that can affect the input-output relationship (Wang & Lee, 2002). However, due to the excessive susceptibility of an original single-pass (sp) MCA to variations in the input data ordering, a heuristic modification was proposed to alleviate this problem. As a result, a multi-pass (mp) MCA algorithm was developed (Herman et al., 2008c). It relied on iterating the original spMCA several times (controlled by a parameter) with the core input data appended with the data points representing means of clusters found in the previous iteration. The core data were shuffled at each iteration. Moreover, for every iteration the record of the cluster validity index, reported on a separate validation set, serving as a performance measure of the given cluster structure was kept. The maximum of this measure determined the iteration that resulted in the selected cluster structure. The underlying concept of this approach is presented in the form of pseudocode in Fig. 7.

It is worth emphasising that the MCA provides information not only about the cluster position in the multi-dimensional input space (the cluster mean, m_{INP}) but also determines its spread in terms of the standard deviation estimate, s_{INP} (independently along different dimensions). Moreover, the assignment of a class label to each cluster is straightforward due to the consistency in the input-output mapping promoted by the algorithm.

Secondly, the well-established fuzzy c-means (FCM) clustering was examined in this work due to its wide applicability in fuzzy rule base identification (Bezdek, 1981). Although the algorithm requires the prior assumption of the number of clusters, its identification was automated using the above-mentioned cluster validity index as a selection criterion. The input data space was clustered resulting in the specified number of cluster centres m_{INP}. The clusters' width vectors, s_{INP}, were composed of the one-dimensional standard deviations, $s_{INP}^{(i)}$, $i = 1,\ldots, n$, calculated independently for each feature vector component over the subset of the input data points with the membership degree in the corresponding clusters above a certain threshold (controlled by a parameter). Since FCM does not explicitly enforce the consistency in mapping between the input and the output space, the class assignments were uniformly randomised in the interval corresponding to class labels, i.e. [-1,1].

Pseudocode of the mpMCA clustering

1. Divide an ordered data set $D=((r,y)_j)$ into a training D_{tr} and a validation subset D_{vld} (here the proportion of 80% to 20% was used). Iteration counter $l=0$.

2. Perform original spMCA clustering on D_{tr} and find the resultant cluster means $\boldsymbol{m}_{INP}$, their standard deviations $\boldsymbol{s}_{INP}$ and class assignments c.
Update iteration counter: $l=l+1$.

3. Initialise a prototype T1FLS based on the clusters found in 2 and use it for classification of the data set D_{vld}. Evaluate the CA rate as $CA_{vld}(l)$.

4. Shuffle D_{tr} and set $D'_{tr}=D_{tr}$:

$$D_{tr} \xrightarrow{\text{shuffle}} D'_{tr}$$

5. Update D_{tr} by appending the clusters' means to the beginning of the ordered set D'_{tr}:

$$D_{tr} = ((\boldsymbol{m}_{INP},c)\cup D'_{tr})$$

6. IF $l \leq N_{passes}$ THEN go back to 2.

7. Find an iteration where the CA_{vld} was maximum

$$l^* = \max_{1<l\leq N_{passes}} \arg CA_{vld}(l)$$

and restore the clusters means $\boldsymbol{m}_{INP}$ with their standard deviations $\boldsymbol{s}_{INP}$ and class assignments c that were identified at the l^*th iteration. They serve as the cluster output of the procedure.

Fig. 7. Pseudocode of the modified MCA algorithm – mpMCA.

Thirdly, subtractive clustering (Chiu et al., 1994) as a computationally effective implementation of mountain clustering, originally proposed by Yager and Filev (Yager & Filev, 1994), was employed in this study. The selection of cluster centres was based on the density of data points (feature vectors). The density-related measure assumed the form of an iterative combination (for subsequent clusters) of radial basis functions. Analogously to the FCM approach, a certain neighbourhood of each resultant cluster centre was specified to determine the membership status of the clustered data points and then to estimate the corresponding one-dimensional standard deviations. The size of the neighbourhood was controlled by an extra parameter, which facilitated adjustments of the size of overlap between the clusters. The output space assignments were made randomly for the same reasons as in the FCM-based scheme.

A prototype singleton T1FLS rule base was straightforwardly derived from the resultant clusters in the input space and their class assignments. To this end, each multi-dimensional cluster was projected on single input dimensions (feature vector components) to form a fuzzy rule. Its antecedents were modelled using Gaussian T1FSs, whose means, $m^{(i)}$, and widths, $s^{(i)}$, $i = 1,..,n$, were determined as the projections of the cluster's $\boldsymbol{m}_{INP}$ and $\boldsymbol{s}_{INP}$, respectively. The consequent was defined in the output space as centroid centred at the associated class label. For the purpose of easy visualisation, an example of the projection of a two-dimensional cluster of data belonging to class c on the axes corresponding to respective feature vector components (r_j and r_k) and the resulting T1 fuzzy rule are illustrated in Fig. 8.
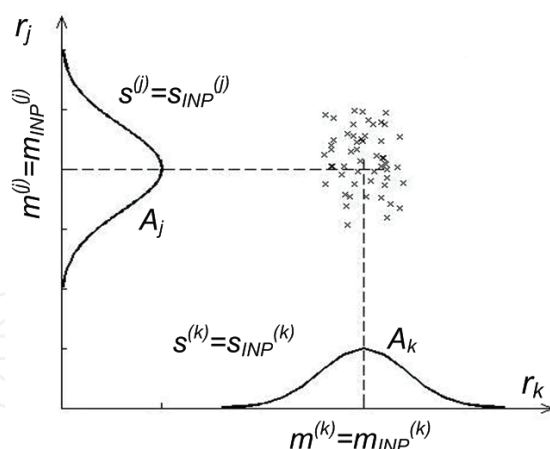
IF r_j is A_j AND r_k is A_k THEN *class* is *c*

Fig. 8. A two-dimensional cluster in the feature space and the corresponding prototype T1 fuzzy rule.

After the identification of the prototype T1FLS, it was extended to serve as a framework for an IT2FLS. As presented above, each T1FL rule was described in terms of its antecedent FSs A_i ($i=1,..,n$), parameterised with vector $\boldsymbol{m}=(m^{(1)},...,m^{(n)})$ of their means and vector $\boldsymbol{s}=(s^{(1)},...,s^{(n)})$ of their standard deviations, and a crisp consequent, *c*. The uncertainty bounds of the FSs defining the antecedent and the consequent part of an IT2FL rule were expressed using additional quantities, Δm and Δc, respectively (c.f. Fig. 6). The resultant formulae for IT2FL rule induction from the classical T1FL rule prototype are as follows:

$$\boldsymbol{m}_1 = \boldsymbol{m} - \Delta\boldsymbol{m} \qquad \boldsymbol{m}_2 = \boldsymbol{m} + \Delta\boldsymbol{m},$$
$$c_{left} = c - \Delta c \qquad c_{right} = c + \Delta c. \tag{7}$$

Vectors $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ refer to the lower and the upper bound of the uncertain means (c.f. Fig. 1a) in the antecedent IT2FSs and c_{left}, c_{right} define the consequent centroid. The standard deviations, $\boldsymbol{s}$, of the prototype T1FSs were kept the same for the resultant IT2FSs. Furthermore, it was found that the constrained parameterization of Δm and s_{fuzz_inp} (used in the description of the fuzzified inputs, c.f. Fig. 6) with multiplicative factors *dm* and *a* in (8) and (9), respectively, led to a more computationally efficient parameter selection procedure.

$$\Delta\boldsymbol{m} = dm\,\boldsymbol{s} , \tag{8}$$

$$\boldsymbol{s}_{fuzz\_inp} = a\,\boldsymbol{\sigma}_r , \tag{9}$$

where $\boldsymbol{\sigma}_r$ is a vector of the standard deviations of the input features $\boldsymbol{r}$ in a training set.

The parameters *dm*, Δc and *a*, assumed to be homogeneous for the entire rule base, determined the initial bounds of the uncertainty captured in the system. They were selected in combination with a training process, described below, with the aim of maximising the performance of the resultant IT2FLS classifier evaluated using a CV approach on the selected calibration data set (within-session classification).

In the second stage of the IT2FLS classifier design, the quantities initialised in the earlier step, m_1, m_2, $\boldsymbol{s}$, c_{left}, c_{right} and s_{fuzz_inp}, were tuned for every rule. A global nonlinear

optimisation approach was adopted to this end. The learning algorithm was based on the concept of steepest gradient descent with the mean square error loss function, L, defined in (10). In the training phase, a continuous defuzzified output, y, of the fuzzy classifier was taken into account whereas in the recall process, simple thresholding was applied to obtain a dichotomous class label.

$$L = \frac{1}{M}\sum_{j=1}^{M}(y_j - label_j)^2, \tag{10}$$

where M is the number of training trials (feature vectors) and $label_j$ is the desired class label (-1 or 1) assigned to the j-th trial (feature vector).

A heuristic training strategy for IT2FLS classifiers was proposed in this work with a view to enhancing their generalisation properties and speeding up the convergence of nonlinear optimisation. It is composed of three stages and combines two approaches known in the domain of IT2FLSs: the conventional steepest gradient descent algorithm developed by Liang and Mendel (Liang & Mendel, 2000; Mendel, 2001), and the method based on the dynamic optimal rate theorem (Wang et al., 2004). This hybridisation was demonstrated to result in more robust and effective search for an optimal configuration of the system parameters than the conventional Liang and Mendel's approach in the given brain signal pattern recognition problem. The three stages were conducted as follows (Herman et al., 2008c):

Stage I)
The conventional steepest descent was applied with learning rates being reduced by a constant factor every 10 epochs. Identification of their initial values was found to play a significant role in the entire optimisation process and it was thus incorporated in the parameter selection scheme. A validation data set was utilised to implement an early stopping criterion. This facilitated an informed decision about terminating the optimisation process and led to the enhancement of the classifier's generalisation capabilities.

Stage II)
An algorithm based on the dynamic optimal rate theorem was applied to accelerate the optimisation of the parameters of the fuzzy rule consequents. In particular, the combination of sample-by-sample training of the standard deviations s_{fuzz_inp} and the antecedent parameters m_1, m_2 and s with a batch update of the consequents c_{left} and c_{right} was adopted. This parameter learning phase was terminated based on the same early stopping criterion as in the first stage.

Stage III)
The T2FLS's parameters were fine tuned using an algorithm similar to that of the first stage with far lower learning rates and the reduced number of epochs. The updated system parameters were accepted only if the classifier's performance in terms of the CA rate improved in comparison with the outcome of the second stage. Otherwise, the parameter configuration was rolled back.

In order to conduct a fair comparative analysis, an analogous learning algorithm was developed for T1FLS classifiers. As mentioned earlier, the results of the examination of the presented design variants are summarised in section 5.1.

4.3.2 Popular BCI classifiers

In a comparative evaluation, more conventional binary classifiers widely used in EEG-based BCI were verified. In the first place, parameter-less LDA, which is Bayes optimal classifier for normally distributed features in each of two classes with the same covariance matrix (Bishop, 1995), was employed. Although the condition mentioned is often violated (it was not met for any of the data sets considered in this study), LDA is commonly perceived as an effective, easy to use and computationally cheap classification method in BCI work (Vaughan et al., 2003; McFarland et al., 1997). In consequence, it has been proven to perform well even with relatively small data sets. SVM classifiers adopt a different approach to identifying a class separating hyperplane in dichotomous problems. Unlike in LDA, where the discriminative boundary is determined as a result of maximising the ratio of inter-class variance to the intra-class variance, SVM hyperplane provides the largest margin between classes without taking any second-order statistics into account (Cristianini & Shawe-Taylor, 2000). This intuitively facilitates generalisation and from an algorithmic perspective, it requires solving a quadratic programming problem with a unique solution. A soft version of SVM (Kecman, 2001) is more popular in real-world problems involving discrimination in feature spaces with overlapping class specific regions. It allows for misclassification of a proportion of data points in a certain neighbourhood of the decision boundary, where overlapping is likely to occur. The size of the neighbourhood is controlled by a corresponding regularisation parameter that decides the trade-off between the training error and the size of the margin between classes. The core assets of SVM classifiers stem from their generalisation power, wide availability of computationally effective approaches to quadratic programming even in the presence of large data sets, and a kernel machine-based structure allowing for straightforward transformation of inherently linear classifiers into nonlinear ones. This latter property, which arises out of implicit nonlinear mapping determined by the so-called kernel function (Cristianini & Shawe-Taylor, 2000), is often exploited in practical applications. In this study, a Gaussian kernel with homogeneous variance was examined due to its successful application in other BCI studies, eg. (Garrett et al., 2005). For clarity, linear SVM (without a nonlinear kernel) is referred to throughout this chapter as SVM_{lin} whereas SVM with the Gaussian kernel – SVM_{Gauss}. Finally, the relevance of parameter selection for optimal SVM performance should be emphasised. In this work, the regularisation constant for SVM_{lin} and the kernel parameter with the regularisation constant for SVM_{Gauss} were identified in a simple off-line grid search based on the performance of the resultant classifiers. It was assessed using a CV estimate of the generalisation error, which was proven reliable in previous studies, eg. (Herman et al., 2008a).

4.3.3 Discrete vs. continuous classification

As emphasised in section 4.2, the main methodological difference between discrete and continuous classification of MI related EEG spectral patterns lies in the design of a feature extraction unit. The classification framework remains essentially the same with the structure determined by the format of a feature vector and the parameter setup reflecting the temporal characteristics of the features to be discriminated. In discrete mode, the number of classifier's inputs was dependent on the number of feature extraction windows, n_{win} (c.f. (4) and (5) in section 4.2.1). Since a feature vector described an entire signal trial, off-line handling of the resultant training data set was rather straightforward. This is more intricate

in continuous classification mode. The dimensionality of the feature space and thus of the classifier's input was set to four at every time point (separate correlates of μ and β for C3 and C4 channels, c.f. section 4.2.2) within an event-related part of the trial. Consequently, in off-line calibration of the continuous BCI classifiers, there were as many data sets as many data points within a trial (a data set is defined here as a collection of feature vectors for all trials). In other words, there could be a new classifier set up at each time point. Since it was not considered to be a practical option, a single classifier to be applied in subsequent BCI sessions was derived using the most separable set of training features extracted at the so-called optimum classification time point (c.f. Fig. 9a). It was expected that the classifier's generalisation capability should secure robust performance over the entire trial's length. Discriminative properties of the given feature sets were quantified based on the average CA obtained in CV analysis. The trained BCI classifier was then continuously applied at every sample point in trials in the following sessions.
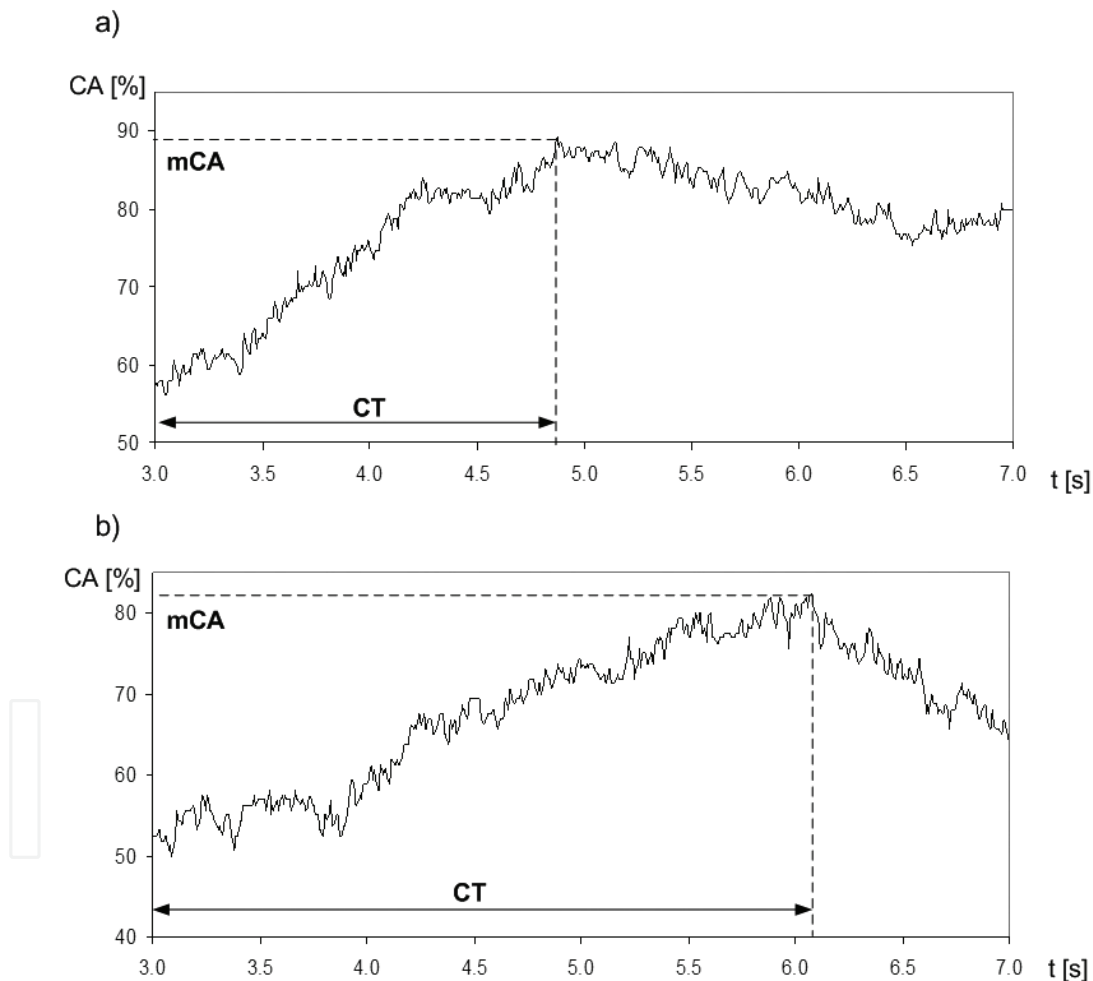

Fig. 9. Instance CA time courses obtained in a) off-line training and b) on-line BCI evaluation for subject S_{III}. Maximum CA and the optimum classification time within the event-related part of the trial are marked as mCA and CT, respectively.

Instantaneous mode also renders the evaluation process and performance assessment more complex than in discrete classification of entire EEG trials. Application of a continuous classifier in an experimental session does not result in a single CA rate but allows for

generating a CA time course. There are several measures commonly employed in the quantification of performance of continuous BCIs such as information transfer rate or mutual information (Wolpaw et al., 2002). Here, the focus is on a pattern recognition aspect and thus CA plays a key role. In particular, the resultant CA time course was quantified by the maximum CA (mCA) rate within the event-related segment of a trial. An example pair of CA time courses, the first one obtained in the CV-based evaluation on a training session and the other one reported on-line in the subsequent test session is depicted in Fig. 9a-b.

With regard to evaluation of an on-line BCI system it is important to realise that any comparative study has to involve retrospective off-line simulation of alternative BCIs since only one closed-feedback system (neurofeedback provider) can be used on-line at a time. To this end, the data recorded during on-line experiments with the original BCI (configured with an IT2FLS classifier) was exploited in the post-hoc examination of other BCI classifiers (c.f. section 4.3.2) in the way that ensured full correspondence in terms of the data handling strategy. Still, it should be noted that a comparative analysis of methods tested on-line and off-line has intrinsic limitations. An on-line BCI is part of a closed feedback loop and is thus coupled with the changes in the on-going MI related EEG activity due to the brain plasticity phenomenon (Herman et al., 2008b). This interaction renders on-line BCI classification particularly challenging. In short-term perspective, especially for subjects with little or no prior BCI experience, on-line BCI systems can either benefit from a facilitatory role of feedback in subject's learning or suffer from inhibitory feedback effects (Herman et al., 2008b). A retrospective off-line examination of BCI methods does not allow for such demanding verification. In spite of that, off-line evaluation still provides a valuable indicator of the potential of the methods being scrutinised to deal with a range of problems inherent to BCI. In this work, it served as the reference for demonstrating uncertainty handling capabilities of the proposed on-line T2FLS classification framework.

5. Results and discussion

5.1 Evaluation of fuzzy classifier design variants

The aim of the extensive examination of several different design variants for T1FLS and IT2FLS was to identify an effective strategy for devising robust fuzzy classifiers with strong uncertainty handling and generalisation capabilities needed in BCI applications. Due to the pattern recognition focus of the presented work and dichotomous nature of the BCI classification, CA was meant to serve as the key comparative criterion. Other aspects such as computational time or sensitivity to initial parameters were assessed qualitatively and were also taken into account. The analysis involving all possible combinations of the proposed clustering-based initialisation methods and the learning algorithms (c.f. section 4.3.1), conventional steepest descent and the enhanced hybrid learning scheme, was performed off-line in discrete classification mode on arbitrarily chosen initial calibration data sets (c.f. section 4.2). In particular, FLSs were set up using each of the design strategies and verified in two types of experimental tests. Firstly, within-session CV (five-fold) study was conducted to illustrate generalisation properties of the developed classifiers and secondly, session-to-session performance transfer was examined, i.e. the classifiers were calibrated including selection of design parameters (c.f. section 4.3.1) on data from one session and tested in single pass on unseen data from the other subsequent one. The inter-session classification helped to gain insight into the issue of dealing with long-term variability effects, highlighted in this chapter.

Despite rather insignificant differences in the classification performance reported with all the design schemes, the comparative evaluation still allowed for several valuable observations to be made. They motivated the final selection of an optimal approach to devising IT2FLS classifiers utilised in this work and were less informative in the case of T1FLS.

In the beginning, the proposed mpMCA-based initialisation was examined with the two parameter tuning approaches. As expected, it significantly alleviated the problem of inconsistency in the cluster structure due to random input data ordering, which manifested itself when the original spMCA was applied. The mpMCA heuristic also led to the improved classification performance of the resultant classifier, particularly IT2FLS in the session-to-session setup. As far as a training algorithm is considered, faster and steadier convergence was reported with the proposed heuristic learning strategy when the fuzzy rule base was initialised using mpMCA (Herman et al., 2008c).

In the next stage, the mpMCA approach was juxtaposed with the other fuzzy structure identification approaches under consideration. At first, a paramount role of the three-stage parameter tuning algorithm was reported in the development of FLSs initialised with the FCM-based technique. This heuristic optimisation approach stimulated a considerable increase of the inter-session CA rates obtained with the resultant classifiers and demonstrated the improved convergence when compared to the conventional steepest descent method. Less pronounced effects of its application were observed when the subtractive clustering-based initialisation was employed. Still, both FCM and subtractive clustering approaches rendered the IT2FLS's performance comparable (insignificantly worse) to that reported with the mpMCA-based fuzzy rule base identification. The differences were even more negligible in the domain of T1FLS. However, what turned out to play a decisive role was the fact that the mpMCA initialisation contributed to the faster convergence of the enhanced gradient descent training, and from the perspective of the derived FLS, mpMCA was found to be less susceptible to its initial parameters, especially considering the strong reliance of FCM upon the number of clusters to be found. In consequence, the combination of the mpMCA initialisation scheme and the proposed three-stage gradient descent-based learning algorithm was identified in this study as the most robust approach to fuzzy classifier design, IT2FLS in particular (Herman et al., 2008c). It was employed in the full evaluation of the proposed fuzzy pattern recognition methods on multi-session EEG data sets in discrete (section 5.2) and continuous (section 5.3) BCI classification.

5.2 Discrete classification

The same types of experimental tests were conducted to analyse the performance of BCI classification frameworks in discrete mode as in the examination of the proposed FLS design variants (c.f. section 5.1). At first, within-session CV was carried out on training sets to estimate the overall efficacy of the classifiers and to select optimal configurations of the system parameters. The main objective of the analysis however was to study the IT2FLS classifier's performance over longer periods across different recording sessions (they were obtained with one-week gaps in between), which were expected to reflect a broad range of non-stationary variability effects in the underlying EEG. To this end, four consecutive sessions were arbitrarily selected for each subject, except for the Graz data set including only two-session recordings, and a strategy for multi-session comparative evaluation involving single-pass tests was devised. In particular, using the initial parameters identified

in the CV analysis the classifiers were calibrated on a training session data set and tested in one pass over multiple subsequent sessions. The overall results in the single pass training-test experimental setup were grouped into three main categories reflecting a temporal relationship between the training and the test session. Next, they were averaged within these categories resulting in three mean CAs for every subject. The first category was generated from three training-test pairs, session I–II, II–III and III–IV. The second one consisted of the test CAs from experiments involving session pairs: I–III and II–IV, and the third category with the largest temporal gap between the training and the test data sets was composed of one CA result obtained with a classifier trained on session I and tested on session IV. This configuration of training-test session pairs is conceptually illustrated in Fig. 10. These collective CAs were then averaged within each category over subjects. The inter-subject means and their standard deviations are presented in Table 1 for every classifier examined in this study, i.e. IT2FLS, T1FLS, LDA, SVM_{lin} and SVM_{Gauss}. In addition, the CV results averaged over three training sessions and subjects are also included in Table 1, in the column 'CV'.
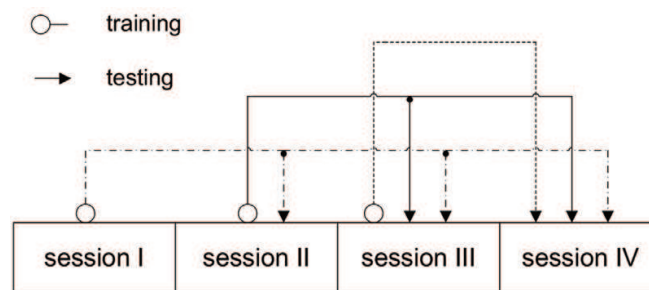

Fig. 10. An illustrative description of the multi-session experimental design for four-session data (lines connect training sessions with the corresponding test sessions).

Classifier	Training Session	Test-Cat. I	Test-Cat. II	Test-Cat. III
	mean CA ± std.dev. [%]			
	CV	Single pass training-test evaluation		
IT2FLS	71.2 ± 8.4	**73.4 ± 9.0**	**64.8 ± 6.7**	**65.4 ± 6.7**
T1FLS	70.4 ± 8.3	71.8 ± 9.1	63.6 ± 6.3	63.9 ± 7.5
LDA	71.5 ± 8.4	67.5 ± 9.3	61.8 ± 8.0	60.7 ± 7.1
SVM_{lin}	71.1 ± 9.3	69.8 ± 9.9	61.7 ± 7.5	60.3 ± 6.4
SVM_{Gauss}	71.0 ± 9.3	69.7 ± 9.8	61.8 ± 7.0	60.4 ±6.9

Table 1. Comparative analysis of the fuzzy rule based classifiers, i.e. T1FLS and IT2FLS, LDA and SVMs: linear (SVM_{lin}) and with Gaussian kernel (SVM_{Gauss}) in terms of the mean CA rates obtained off-line in the multi-session setup with discrete classification. The mean values were calculated across subjects and averaged within given test categories. The standard deviations reflect inter-subject variability (Herman et al., 2008c).

The results in Table I were analysed using one-way ANOVA with repeated measures to test the significance of the differences in the classifiers' performance independently in each category. The null hypothesis could not be rejected for CV results so the focus was on the single pass test CA rates, which reflect the capability of the classifiers to effectively handle

the inherent inter-session variability of the MI induced EEG patterns. The ANOVA test carried out on these sets of results revealed statistically significant (p<0.05) differences in the classifiers' performances. The post test comparison was conducted using Tukey's honestly significant difference criterion (at the significance level of α=0.05) (Maxwell & Delaney, 2004). It showed that IT2FLS outperformed LDA and SVM in every test category, from I to III, whereas T1FLS delivered significantly higher CA rates than LDA only in the first category, i.e. when the classifiers were trained on the session directly preceding the test session.

Although the difference between the mean CA rates obtained with IT2FLS and T1FLS was not found statistically significant, the superior trend of the IT2FLS-based approach was observed consistently for every category of the presented results across all subjects considered in this work. Overall, the analysis demonstrated the potential of the designed IT2FLS in offering enhanced robustness against the inter-session uncertainty effects in MI induced brain phenomena reflected in EEG, especially when compared to common BCI methods.

5.3 Continuous classification and on-line BCI

As discussed in section 4.3.3, the evaluation of the real-time performance of the IT2FLS classifier embedded in the on-line BCI system was accompanied by a comparative off-line study involving other BCI classifiers applied post hoc in continuous mode. The original on-line study was conducted on six subjects performing MI tasks in the ISRC BCI setup over ten sessions (see section 4.1). Still, only four or five sessions when particular subjects showed some reasonable level of BCI control via neurofeedback mechanism are reported here. A strategy for off-line calibration and selection of a reliable BCI system configuration for on-line application was devised individually for each subject. The objective was to maintain a given on-line BCI setup, particularly the IT2FLS in use, over the longest possible time provided that its real-time performance was satisfactory. Otherwise, the classifier was re-trained and an optimal model was selected off-line using the most recent session data as a design (training) data set in the way described in section 4.3.3 (involving within-session CV-based estimate of generalisation error and the identification of the optimal classification time within a trial). This procedure of occasional (every few sessions) calibration of the BCI, with emphasis on the embedded IT2FLS classifier, was aimed at delivering the best possible on-line BCI performance (Herman et al., 2008b). More specifically, the maximum recognition rate in each session was of main concern (c.f. Fig. 9 and section 4.3.3) as it was directly relevant to the desirable effect of the biofeedback facilitating real-time BCI operation and enhancing the consistency of the EEG patterns correlated with the brain activity underlying MI generation.

With regard to the retrospective simulation of a continuous BCI with the other classifiers including T1FLS and the classical methods described in section 4.3.2, the combination of training and test data sets was made analogous to that adopted individually for each subject in on-line experiments. The performance was assessed based on the examination of the resultant CA time courses, as described in section 4.3.3 (c.f. Fig. 9b). In consequence, depending on the subject, four or five mCA rates are considered here as a result of single-pass test evaluation (on-line for IT2FLS and off-line for the rest of the classifiers). Their average is illustrated for every subject in Fig. 11. The mean values across all six subjects are then presented in Table 2.

Classifier	Average performance in test sessions
	mean mCA ± std.dev. [%]
IT2FLS (online)	**69.2 ± 4.6**
T1FLS	66.9 ± 4.3
LDA	66.0 ± 4.4
SVM$_{lin}$	66.8 ± 4.9
SVM$_{Gauss}$	67.1 ± 4.5

Table 2. The average mCA rates obtained in the comparative study with the classifiers evaluated off-line, i.e. T1FLS, LDA, SVM$_{lin}$, SVM$_{Gauss}$, and IT2FLSs applied on-line. The mean values were calculated across all subjects and respective test sessions. The standard deviations reflect inter-subject variability (Herman et al., 2008b).
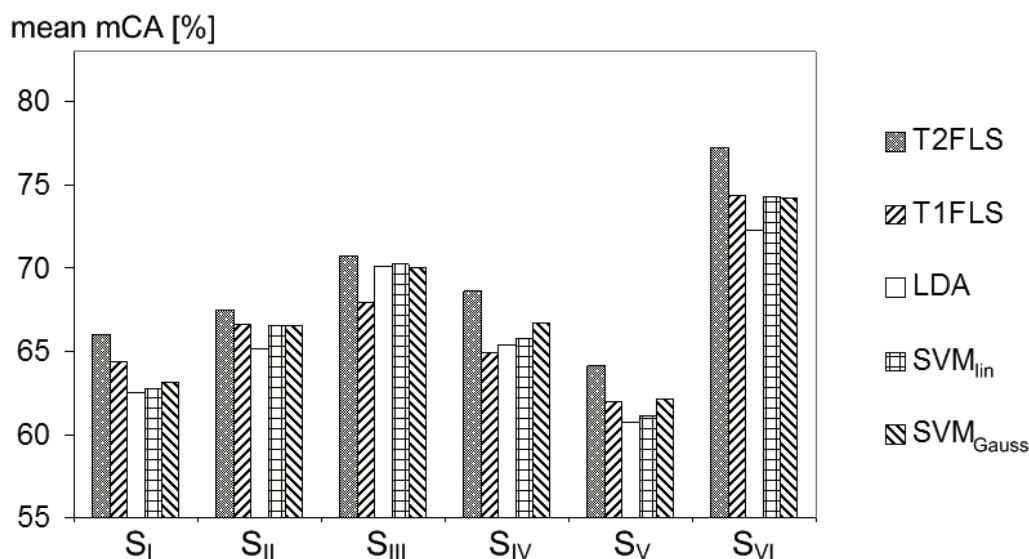

Fig. 11. The average mCA rates (maximum values in the CA time courses) obtained for every subject across BCI test sessions with LDA, SVM$_{lin}$, SVM$_{Gauss}$, T1FLS in off-line and with IT2FLS in on-line mode (Herman et al., 2008b).

Similarly as in the study on discrete classification of MI induced EEG spectral patterns, the experimental results were analysed in the framework of one-way ANOVA with repeated measures. It facilitated accounting for the inter-subject variability. The test conducted at a significance level of α=0.05 allowed for rejecting the null hypothesis about the equality of the grand average of the mCAs reported with different classifiers. The Tukey-Kramer post-ANOVA multiple comparison test demonstrated that the IT2FLS classifier produced overall better classification performance than the other classifiers except SVM$_{Gauss}$, which is generally known for its robustness in the presence of incoherency, noise and nonlinearity effects in the feature space (Cristianini & Shawe-Taylor, 2000; Kecman, 2001). Still, a complementary planned pair-wise comparison between IT2FLS and SVM$_{Gauss}$ with t-statistic and the Bonferroni adjustment helped in manifesting the statistically significant difference between the two mCA means. Fig. 11 illustrates the consistency of the IT2FLS approach across different subjects in delivering superior mCA rates even if the overall degree of the enhancement was rather modest. Despite the limitations of such comparative evaluation of

BCI methods applied on-line and those simulated off-line, as discussed in section 4.3.3, the results obtained in this study reinforce the conclusions drawn in the previous section regarding the potential of the IT2FLS-based classification framework to account for intrinsic uncertainty manifestations more effectively than the other more common BCI classifiers. It should be noted that unlike discrete single trial classification, continuous within trial analysis allows for examining BCI performance at different temporal scales – not only in the presence of inter-session variability effects.

With regard to the juxtaposition of a conventional T1FLS and a new IT2FLS, it should be reminded that the observed superiority trends of the latter method could not be proven statistically in the earlier study reported in section 5.2. However, in continuous recognition of MI induced EEG spectral patterns, the ANOVA test demonstrated a significant advantage of IT2FLS in terms of the average mCA rate. This outcome substantiates the fact that a conventional T1 uncertainty handling framework, unlike the enhanced IT2FLS, does not provide a sufficiently flexible mechanism to account for a range of variability effects observed in MI related EEG at various temporal levels. At the same time, the computational time involved in the IT2FLS optimisation was comparable to that of the T1FLS classifier due to rather small rule base sizes being used.

6. Summary, conclusions and future work

The major theme of this chapter has been centred on the demanding nature of brain signal pattern recognition in the presence of uncertain information intrinsically associated with the biological data source. Specifically, the problem of classification of EEG correlates of MI related brain phenomena has been studied in detail due to its direct relevance to the fast growing field of BCI. Uncertainty effects in EEG-based BCI pose a particularly serious challenge due to their strong manifestations and multi-faceted characteristics. As suggested in the BCI literature and confirmed in the work reported in this chapter, these effects are reflected in the nonstationary changes in the underlying spatio-temporal dynamics of the spectral EEG correlates. Long-term variability have been the main focus of this research due to its particularly adverse influence on inter-session classification performance. In consequence, the research efforts have been concentrated on creating a robust pattern recognition framework capable of more effective handling of the uncertainty manifestations to perform reliable brain signal analysis for BCI purposes. Two instances of a general MI related EEG classification problem have been examined in this chapter – firstly, discrete (single) classification of entire trials and secondly, continuous discrimination of MI induced EEG patterns in a multi-session setup for multiple subjects. The nature of relevant spectral EEG correlates is similar in both cases as they arise out of the same neurophysiological phenomena. However, feature extraction approaches were devised to suit different temporal characteristics of the given problem instances.

A fuzzy rule base classifier proposed here constitutes the core of the developed framework for robust brain signal pattern recognition. The emphasis in this chapter has been put on T2FLS methodology recently enjoying a considerably surge of interest. The extended definition of T2FSs provides more flexibility in modelling inherently uncertain phenomena. In order to effectively exploit this enhanced framework for handling variability effects and boost its generalisation capability, the need to devise a suitable design strategy has been identified and addressed in this chapter. In particular, incremental modifications and hybridisation of existing initialisation and optimisation techniques have resulted in an effective design scheme for IT2FLS classifiers.

The outcome of both empirical studies, conducted with discrete and continuous pattern recognition framework, has led to the overall conclusion that the proposed IT2FLS method lends itself as a more robust alternative to the state-of-the-art BCI classification approaches in the presence of intrinsic variability of the spectral EEG correlates of MI at different temporal scales. The IT2FLS classifier's capacity to embrace inter-trial variations in the EEG feature patterns over a training session and encapsulate them within the FOU for use on subsequent test sessions is considered to be of particular relevance in this regard (Herman et al., 2008c). It should be noted that the applicability of the proposed classification framework in real-world BCI situations was successfully verified in demanding on-line tests, which reflected the potential of IT2FLS in generating consistent and contingent real-time neurofeedback responses correlated with subjects' motor imaginations (Herman et al., 2008b). Consequently, the IT2FLS-based BCI has recently been used in a preliminary study on post-stroke rehabilitation.

In addition, it is worth mentioning that the inherent transparency of a fuzzy inference system was exploited in original work in an attempt to enhance the understanding of MI classification rules derived from the underlying data. Due to a purely qualitative nature of these investigations, they are not reported in this chapter. It suffices to say that they led to the identification of subject specific discriminative trends in spectro-spatio-temporal EEG patterns characteristic of each of the MI classes under consideration. This interpretative approach is still in its early stage. It is aimed to constrain the rule base analysis to gain valuable insight into more general (for a population of subjects) electrophysiological correlates of MI induced brain phenomena through the prism of the qualitative knowledge extracted in the classifier induction process. This may add a new dimension to machine learning-based BCI studies.

The work presented in this chapter is by no means completed. It marks a crucial development and experimental stage, still leaving room for further advancements. In the first place, it is suggested that a general problem of initialising uncertainty bounds for antecedents and consequents of IT2FLSs based on a priori knowledge, if available, could be investigated. In this work, where the fuzzy classifier was applied to MI induced brain signal pattern recognition problems, these bounds were adjusted to account for the inter-trial variability within a BCI session granted that uncertainty associated with session-to-session transfer was in the corresponding range. In this regard, a more systematic approach to quantifying or estimating the effective scope of T2FL uncertainty in data would be more beneficial. In a broader perspective, it could be paired with investigations into alternative methods of an initial IT2FLS structure identification. Research efforts should be then concentrated on the development of new clustering algorithms in the T2FL domain, which would also facilitate the initialisation process of the IT2FLS's uncertainty bounds.

From a more theoretical perspective, it is felt that more insight into the formalism of the underlying T2FL apparatus in comparison with well-established probabilistic approaches is needed. Investigations along these lines would pave the way for a more specific description and quantification of uncertainty effects and thus facilitate more informed IT2FLS design. This work could be inspired by the existing links between a probabilistic and a possibilistic perspective of modelling uncertain information.

In the longer perspective it is envisaged that an application of a generalised T2FLS (unlike simplified IT2FLS methodology employed in this work) in MI related EEG pattern recognition can result in further performance enhancement. It should be realised however

that design of generalised T2FLSs involves several serious challenges, starting from a selection of initial rule bases expressed in terms of generalised T2FSs and ending with heuristic approaches to parameter optimisation. In addition, real-time feasibility is expected to be problematic when the resultant T2FLS classifier is applied within a BCI framework. Approximate algorithms may be needed in this regard. The starting point and inspiration for this work should be the remarkable contribution to the area of generalised T2FL made by the research team led by John, e.g. see a review in (John & Coupland, 2007).

In a broader context, it is envisaged that the integrated framework devised in this work to contend with a specific brain signal pattern recognition problem can address a range of complex nonstationary biological and physical signals with uncertain spatio-temporal characteristics that cannot be handled using a rigorous analytical modelling apparatus. In the presence of such highly variable components exhibiting nondeterministic behaviour, T2FLS methodology would appear particularly advantageous over other more conventional approaches. The proposed data driven design approach would promote then the development of a model-free, yet qualitatively interpretable, system for automated analysis with a classification output.

7. References

Akay, M. (ed.) (1997). *Time Frequency and Wavelets in Biomedical Signal Processing*. IEEE Press Series in Biomedical Engineering, New York.

Assous, S.; Humeau, A.; Tartas, M.; Abraham, P. & L'Huillier, J.-P. (2006). S transform applied to laser Doppler flowmetry reactive hyperemia signals. *IEEE Transactions on Biomedical Engineering*, vol. 53, pp.1032–1037.

Bezdek, J.C. (1981). *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York.

Bishop, C. (1995). *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford.

Cheng M.; Jia, W.; Gao, X.; Gao, S. & Yang, F. (2004). Mu rhythm-based cursor control: an offline analysis. *Clinical Neurophysiology*, vol. 115, pp. 745–751.

Chiu, S.L. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, vol. 2, pp. 267–278.

Cristianini, N. & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press.

Duda, R.O.; Hart, P.E. & Stork, D.G. (2001). *Pattern Classification*, Wiley, New York.

Fisch, B.J. (1999). *Fisch and Spehlmann's EEG Primer: Basic principles of Digital and Analog EEG*, Elsevier, Amsterdam.

Garrett, D.; Peterson, D.A.; Anderson, C.W. & Thaut, M.H. (2005). Comparison of Linear, Nonlinear, and Feature Selection Methods for EEG Signal Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp.141–144.

Gorzalczany, M.B. (1988). Interval-valued fuzzy controller based on verbal model of object. *Fuzzy Sets and Systems*, vol. 28, pp. 45–53.

Gorzalczany, M.B. (2002). *Computational Intelligence Systems and Applications: Neuro-Fuzzy and Fuzzy Neural Synergisms*. Physica-Verlag, Springer-Verlag, Heidelberg, Germany.

Guger, C; Schlögl, A.; Neuper, C., Walterspacher, D.; Strein, T. & Pfurtscheller, G. (2001). Rapid Prototyping of an EEG-Based Brain–Computer Interface (BCI). *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 9, pp. 49–58.

Haselsteiner, E. & Pfurtscheller, G. (2000). Using Time-Dependent Neural Networks for EEG Classification. *IEEE Transactions on Rehabilitation Engineering*, vol. 8, pp. 457–463.

Haykin, S. (1996). *Adaptive Filter Theory*, Prentice Hall, 3rd edition.

Herman, P.; Prasad, G.; McGinnity, T.M. & Coyle, D. (2008a). Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 4, pp. 317–326, August 2008.

Herman, P.; Prasad, G. & McGinnity, T.M. (2008b). Design and On-line Evaluation of Type-2 Fuzzy Logic System-based Framework for Handling Uncertainties in BCI Classification. *Proc. of the 30th International IEEE Conference of Engineering in Medicine and Biology Society*, Vancouver, Canada, August 2008.

Herman, P.; Prasad, G. & McGinnity, T.M. (2008c). Designing a Robust Type-2 Fuzzy Logic Classifier for Non-stationary Systems with Application in Brain-Computer Interfacing. *Proc. of the 2008 IEEE International Conference on Systems, Man, and Cybernetics*, Singapore, October 2008.

John, R. & Coupland, S. (2007). Type-2 Fuzzy Logic: A Historical View. *IEEE Computational Intelligence Magazine*, vol. 2, no. 1, pp. 57–62.

Kaiser, D.A. (2005). Basic Principles of Quantitative EEG. *Journal of Adult Development*, vol. 12, pp. 99–104.

Karnik, N.N.; Mendel, J.M. & Liang, Q. (1999). Type-2 Fuzzy Logic Systems. *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 6, pp. 643–658.

Kecman, V (2001). *Learning and Soft Computing*. Cambridge: MIT Press.

Liang, Q. & Mendel, J.M. (2000). Interval type-2 Fuzzy Logic Systems: Theory and Design. *IEEE Transactions on Fuzzy Systems*, vol. 8, pp. 535–550.

Maxwell, S.E. & Delaney, H.D. (2004). *Designing Experiments and Analyzing Data: A Model Comparison Perspective*. Lawrence Erlbaum, New Yersey.

McFarland, D.J.; McCane, L.M.; David, S.V. & Wolpaw, J.R. (1997). Spatial filter selection for EEG-based communication. *Electroencephalography and Clinical Neurophysiology*, vol. 103, pp. 386–394.

McFarland, D.J.; McCane, L.M. & Wolpaw, J.R. (1998). EEG-Based Communication and Control: Short-Term Role of Feedback. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 6, pp. 7–11.

Mendel, J.M (2001). *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, New York.

Millán, J. del R.; Mouriño, J.; Franzé, M.; Cincotti, F.; Varsta, M.; Heikkonen, J. & Babiloni, F. (2002). A local neural classifier for the recognition of EEG patterns associated to mental tasks. *IEEE Transactions on Neural Networks*, vol. 13, pp. 678–686.

Millán, J. del R. & Mouriño, J. (2003), Asynchronous BCI and Local Neural Classifiers: An Overview of the Adaptive Brain Interface Project. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, pp. 159–161.

Niedermeyer, E. & Lopes da Silva, F. (2004). *Electroencephalography: Basic principles, clinical applications, and related fields*, 5th ed., Williams & Wilkins, Baltimore.

B. Obermaier, C. Guger, C. Neuper, G. Pfurtscheller, (2001). Hidden Markov models for online classification of single trial EEG data. *Pattern Recognition Letters*, vol. 22, pp. 1299–1309.

Pfurtscheller, G.; Neuper, C.; Guger, C.; Harkam W.; Ramoser, H.; Schlögl, A.; Obermaier, B. & Pregenzer, M. (2000). Current trends in Graz Brain-Computer Interface (BCI) research. *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp.216–219.

Pfurtscheller, G. & Neuper, C. (2001). Motor Imagery and Direct Brain-Computer Communication. *Proc. of the IEEE*, vol. 89, no. 7, pp. 1123-1134.

Pregenzer, M. & Pfurtscheller, G. (1999). Frequency Component Selection for an EEG-based Brain to Computer Interface. *IEEE Transactions on Rehabilitation Engineering*, vol. 7, no. 4, pp. 413–419.

Schlögl, A.; Vidaurre, C. & Pfurtscheller, G. (2005). Assessing non-stationarities in BCI data. *BCI 2005 Workshop*, Rensellarville, NY, USA, 14-19 June 2005.

Shenoy, P.; Krauledat, M.; Blankertz, B.; Rao, R.P.N. & Müller, K.-R. (2006). Towards adaptive classification for BCI. *Journal of Neural Engineering*, vol. 3, pp. 13–23.

Stoica, P. & Moses, R.L. (1997). *Introduction to Spectral Analysis*, Englewood Cliffs, USA, Prentice-Hall.

Sykacek, P.; Roberts, S.J. & Stokes, M. (2004). Adaptive BCI based on variational Bayesian Kalman filtering: an empirical evaluation. *IEEE Transactions on Rehabilitation Engineering*, vol. 51, no. 5, pp. 719–727.

Townsend, G.; Grainmann, B. & Pfurtscheller, P. (2006). A Comparison of Common Spatial Patterns With Complex Band Power Features in a Four-Class BCI Experiment. *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 4, pp. 642–651.

Wang, C.H.; Cheng, C.S. & Lee, T.T. (2004). Dynamical optimal training for interval type-2 fuzzy neural network (T2FNN). *IEEE Transactions on Systems, Man and Cybernetics*, vol. 34, pp. 1462–1477.

Wang, J.-S. & Lee, C.S.G. (2002). Self-Adaptive Neuro-fuzzy inference systems for classification applications. *IEEE Transactions on Fuzzy Systems*, vol.10, pp.790–802.

Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J.; Pfurtscheller, G. & Vaughan, T.M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, vol. 113, pp. 767–791.

Vaughan, T.M.; Heetderks, W.J.; Trejo, L.J.; Rymer, W.Z.; Weinrich, M.; Moore, M.M.; Kübler, A.; Dobkin, B.H.; Birbaumer, N.; Donchin, E.; Wolpaw, E.W. & Wolpaw, J.R. (2003). Brain-Computer Interface Technology: a Review of the Second International Meeting. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, pp. 94–109.

Vidaurre, C.; Schlögl, A.; Cabeza, R.; Scherer, R. & Pfurtscheller, G. (2006). A Fully On-line Adaptive BCI. *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1214–1219.

Yager, R. & Filev, D. (1994). Approximate clustering by the mountain clustering. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, no. 8, pp. 338–358.

Yang, B.-H.; Yan, G.-Z.; Wu, T. & Yan, R.-G. (2007). Subject-based feature extraction using fuzzy wavelet packet in brain-computer interfaces. *Signal Processing*, vol. 87, pp. 1569–1574.

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning-1. *Information Sciences*, vol.8, pp.199–249.

http://www.gtec.at (accessed on 10.08.2008).

Pattern Recognition Techniques, Technology and Applications

Edited by Peng-Yeng Yin

A wealth of advanced pattern recognition algorithms are emerging from the interdiscipline between technologies of effective visual features and the human-brain cognition process. Effective visual features are made possible through the rapid developments in appropriate sensor equipments, novel filter designs, and viable information processing architectures. While the understanding of human-brain cognition process broadens the way in which the computer can perform pattern recognition tasks. The present book is intended to collect representative researches around the globe focusing on low-level vision, filter design, features and image descriptors, data mining and analysis, and biologically inspired algorithms. The 27 chapters coved in this book disclose recent advances and new ideas in promoting the techniques, technology and applications of pattern recognition.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Pawel Herman, Girijesh Prasad and Thomas Martin McGinnity (2008). Computational Intelligence Approaches to Brain Signal Pattern Recognition, Pattern Recognition Techniques, Technology and Applications, Peng-Yeng Yin (Ed.), ISBN: 978-953-7619-24-4, InTech, Available from: http://www.intechopen.com/books/pattern_recognition_techniques_technology_and_applications/computational_intelligence_approaches_to_brain_signal_pattern_recognition

INTECH
open science | open minds