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

Source: Pattern Recognition Techniques, Technology and Applications, Book edited by: Peng-Yeng Yin,
ISBN 978-953-7619-24-4, pp. 626, November 2008, I-Tech, Vienna, Austria

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):

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

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

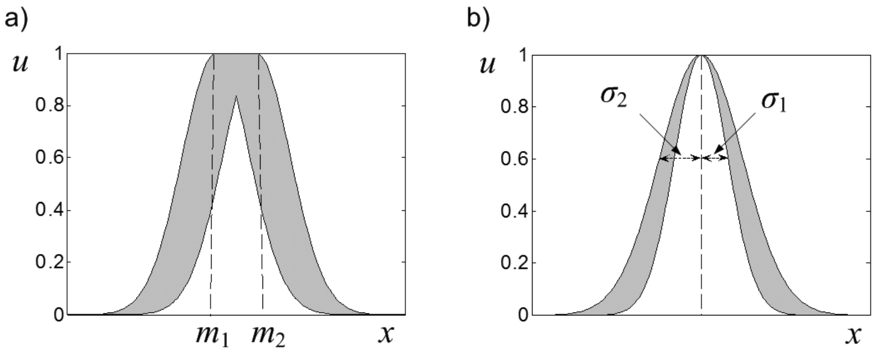

Fig. 1. An illustration of the FOUs of Gaussian T2FSs a) with uncertain mean and fixed standard deviation, b) with fixed mean and uncertain standard deviation.

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.

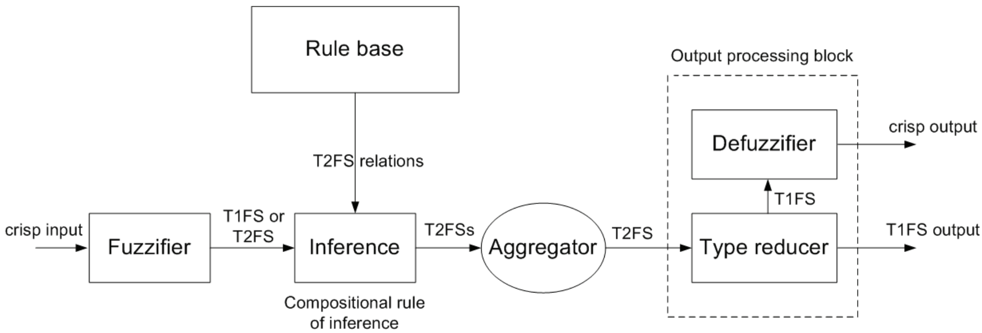

Fig. 2. T2FLS framework.

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.

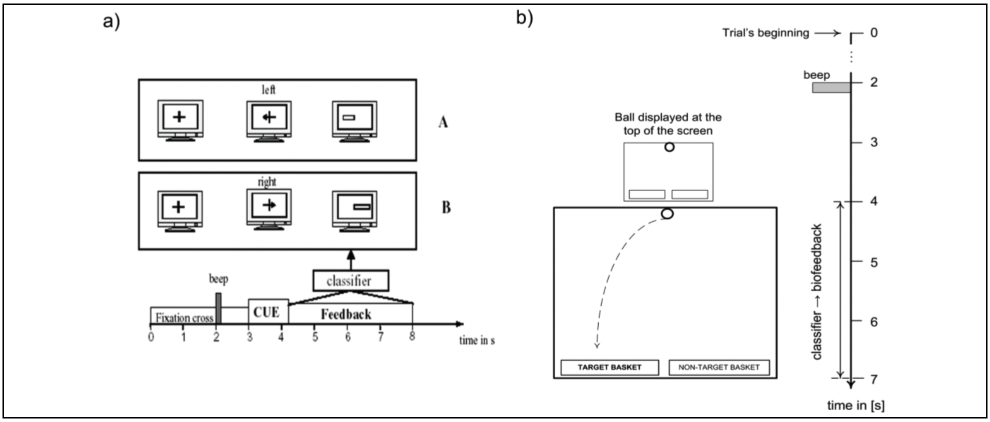

Fig. 3. Data recording in a) Graz BCI paradigm (Haselsteiner & Pfurtscheller, 2000) and b) BCI basket paradigm (Wang et al., 2004).

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:

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

where:

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

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.

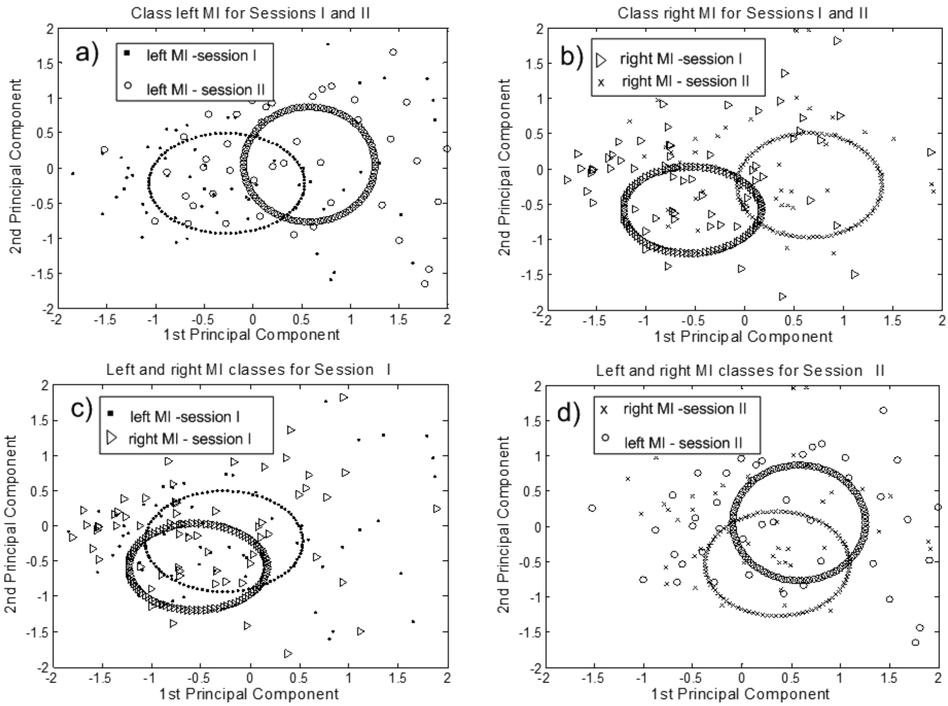

Fig. 4. The distribution of spectral EEG features in two-dimensional normalised PCs' space with their corresponding class means and scaled standard deviations: a) left and b) right MI features in sessions I and II plus within-session feature distribution: c) session I and d) session II.

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 $\boldsymbol{r}$ (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.

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

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

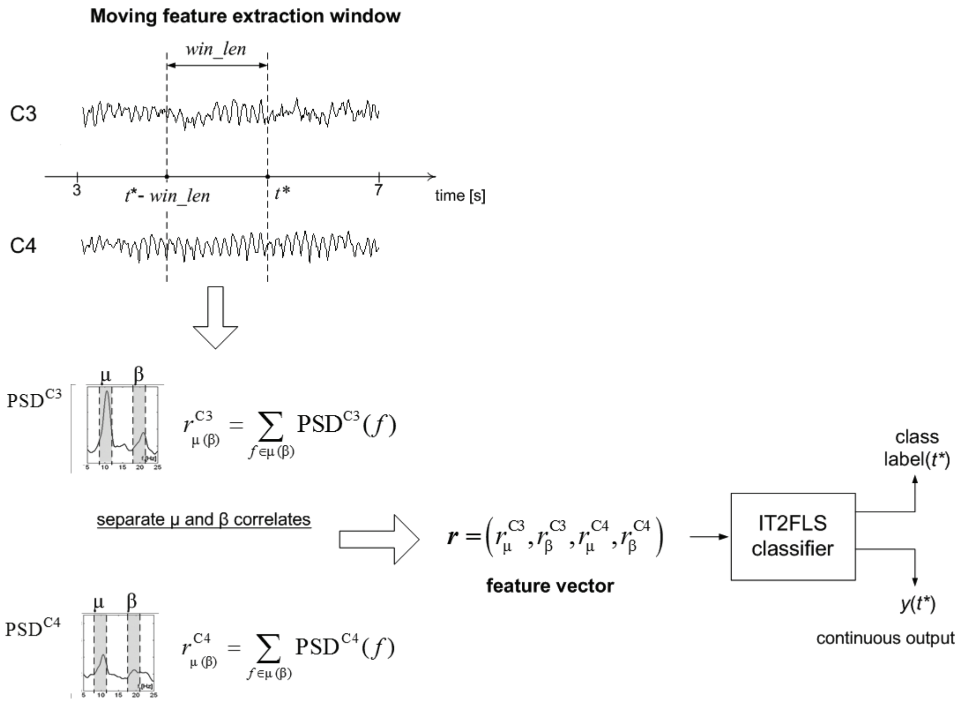

Fig. 5. Graph illustrating the proposed concept of instantaneous BCI feature extraction and classification.

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:

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

where fuzzy variables $X_1,...,X_n$ correspond to the fuzzified components of an input feature vector $\boldsymbol{r}=(r_1,\ldots,r_n)$, n is their number and $\tilde{A}_1,\ldots,\tilde{A}_n$ 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.

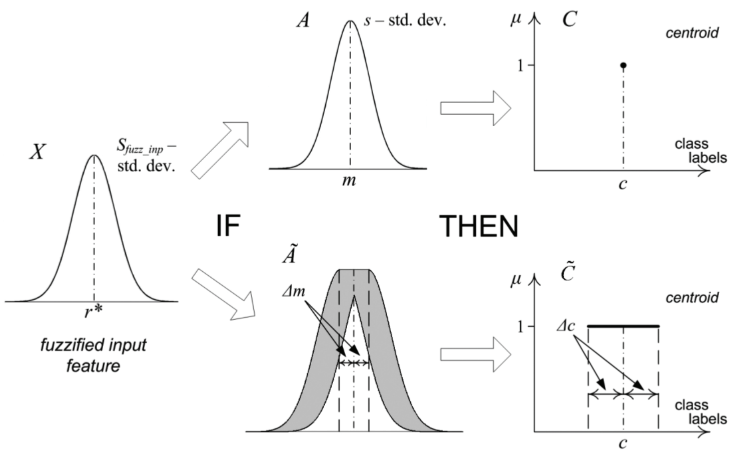

Fig. 6. Illustrative comparison of T1FL and T2FL rule patterns.

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, $\boldsymbol{m}_{INP}$) but also determines its spread in terms of the standard deviation estimate, $\boldsymbol{s}_{INP}$ (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 $\boldsymbol{m}_{INP}$. The clusters' width vectors, $\boldsymbol{s}_{INP}$, 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 $D=((\boldsymbol{r},y)_j)$ 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 $\mathrm{CA}_{\mathrm{vld}}(l)$.

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 $\mathrm{CA}_{\mathrm{vld}}$ was maximum

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

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.

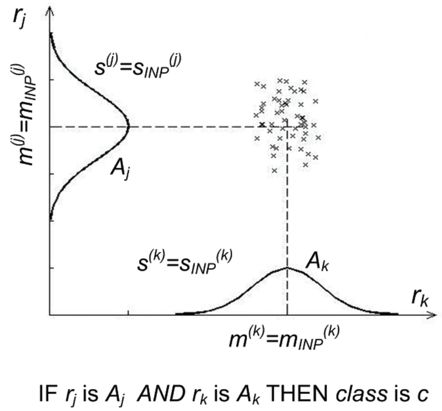

Fig. 8. A two-dimensional cluster in the feature space and the corresponding prototype T1 fuzzy rule.

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 $\boldsymbol{m}=(m^{(1)},\ldots,m^{(n)})$ of their means and vector $\boldsymbol{s}=(s^{(1)},\ldots,s^{(n)})$ 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, $\boldsymbol{\Delta m}$ and Δc, respectively (c.f. Fig. 6). The resultant formulae for IT2FL rule induction from the classical T1FL rule prototype are as follows:

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

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 $\boldsymbol{\Delta m}$ and $\boldsymbol{s}_{fuzz\_inp}$ (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.

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

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

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, $\boldsymbol{m}_1$, $\boldsymbol{m}_2$, $\boldsymbol{s}$, c_{left}, c_{right} and $\boldsymbol{s}_{fuzz\_inp}$, were tuned for every rule. A global nonlinear