Final automatic learning principles allowing the enhancement of sense of presence and the increase of motivation

Deliverable D3.2

List of Partners:	ETH Zurich, CH	(ETH Zurich)
	Hocoma AG, Volketswil, CH	(HOCOMA)
	University of Ljubljana, SI	(UL)
	Universitat de Barcelona, ES	(UB)
	Neurological Clinic Bad Aibling, DE	(NKBA)
Document Identifier:	MIMICS-D3.2-pu.pdf	
Version:	5.0	
Date:	2010-06-30	
Organisation:	UL	
Deliverable:	3.2	
Milestone:	3.2	
Workpackage:	3	
Task:	3.1, 3.2, 3.3 & 3.4	
Dissemination:	public	
Authors:	Domen Novak, Jaka Ziherl, Andrej Olenšek, Alexander Koenig, Jason Kastanis, Matjaz Mihelj, Marko Munih	
Approved by:	Mel Slater and Lars Lars Lünenburger	

Project Number: 215 756

Table of Contents

1 Introduction 3

2 Real-time sensing of user motor actions 4

3 Real-time sensing of psychophysiological state 5

4 Fusion of multi-sensory data streams and real-time user state assessment 6

4.1 Introduction to fusion of multi-sensory data streams and real-time user state assessment 6
4.2 Detection of haptic parameters for the upper extremities 7
4.2.1 Method background description 7
4.2.2 Experimental verification 8
4.2.3 Conclusions 13
4.3 Haptic sensor fusion for adaptive haptic assistance 13
4.4 Fuzzy logic 14
4.4.1 Method background description 14
4.4.2 Experimental verification 15
4.4.3 Conclusions 16
4.5 Model of heart rate response to physical load 16
4.5.1 Method background description 16
4.5.2 Experimental verification 16
4.5.3 Conclusions 19
4.6 Principal component analysis and neural networks 19
4.6.1 Method background description 19
4.6.2 Experimental verification 21
4.6.3 Conclusions 24
4.7 Discriminant analysis 24
4.8 Offline assessment of electroencephalography 25
4.8.1 Method background description 25
4.8.2 Experimental verification 25
4.8.3 Conclusions 26

5 Automatic learning system 27

5.1 Introduction to the automatic learning system 27
5.2 Adaptive haptic assistance 27
5.3 Reinforcement learning 30
5.3.1 Method background description 30
5.3.2 Experimental verification 30
5.3.3 Conclusions 31
5.4 Adaptive discriminant analysis 32
5.4.1 Method background description 32
5.4.2 Experimental verification 35
5.4.3 Conclusions 39

6 Conclusions 40

7 Bibliography 41

1 Introduction

Work package 3 is designed to provide the theoretical basis, procedures, and computer algorithms for acquisition, processing (fusion) and interpretation of multi-sensory information arising from the interaction of the human with the robot and the virtual environment. The work package is split into four tasks:

- Task 3.1: Real-time sensing of user motor actions,
- Task 3.2: Real-time sensing of psychophysiological state,
- Task 3.3: Fusion of multi-sensory data streams and real-time user state assessment,
- Task 3.4: Automatic learning system.

Tasks 3.1 and 3.2 were active in project months 3-7. The work done during these two tasks was described in Deliverable D3.1 and will be briefly summarized here for reasons of clarity. The purpose of this deliverable, „Final automatic learning principles allowing the enhancement of sense of presence and the increase of motivation“, is to describe the work accomplished in Tasks 3.3 and 3.4 at the conclusion of Work package 3 in month 30.

Within Task 3.3, several approaches to user state assessment via motor actions and psychophysiology are described. Several potential data fusion methods were evaluated, from classical statistical methods to modern machine learning principles. The automatic learning systems of Task 3.4 build on these approaches, not only identifying the patient’s current state but also guiding the patient into an optimal state in order to maximize motivation and effort. While some of the methods developed are specific to either the upper or lower extremities, others are relatively general and could prove useful in fields other than rehabilitation.

2 Real-time sensing of user motor actions

This work was successfully completed in 2008 and reported extensively in D3.1.

First, the hardware and software required for measurement and processing of force, torque and movement data were implemented. Then, algorithms were developed for multisensory integration and recognition of the subject's voluntary motor activities and intentions (intended movement directions) in real time using Kalman filtering and inverse kinematic models of the human limbs.

3 Real-time sensing of psychophysiological state

This work was successfully completed in 2008 and reported extensively in D3.1.

An extensive literature review was conducted, identifying the most important psychophysiological signals, the features that can be extracted from them, and the connections between psychophysiological signals and psychological state. Algorithms were developed to import these psychophysiological signals from the g.USBamp signal amplifier into MATLAB/Simulink and automatically extract all relevant features. Initial experiments were performed to estimate psychological effects on the psychophysiological features. After that, four physiological measurements were chosen for use in the minimal configuration (intended for patients): electrocardiography, skin conductance, respiration and skin temperature.

4 Fusion of multi-sensory data streams and real-time user state assessment

4.1 Introduction to fusion of multi-sensory data streams and real-time user state assessment

It has been demonstrated by previous studies that motivation and engagement are very positive to the success of rehabilitation and that encouraging unmotivated patients improves the likelihood of their eventual recovery (Robertson & Murre, 1999; Maclean, Pound, Wulfe, & Rudd, 2002). However, estimation of the user's psychological state has not previously been a part of robot-aided motor rehabilitation. Thus, this task is concerned with the **design of algorithms required to fuse motor actions and psychophysiology into a real-time, objective estimate of user state**. It is split into four subtasks:

- detect the patient's activities within an immersive virtual environment,
- detect the patient's psychological state and sense of presence,
- detect the patient's physiological state (muscle effort),
- off-line assessment of sense of presence based on EEG measurements.

Movement data is the primary source of information regarding the patient's activities while **force and torque** sensors provide the most important information about the patient's muscle effort. For the upper extremities, the first and third subtasks were handled with similar data fusion approaches that are detailed in sections 4.2 and 4.3. For the lower extremities, the patient's physiological state and muscle effort was also evaluated through heart rate, and this approach is described in section 4.5.

Detecting the patient's **psychological state** proved to be more complex since psychophysiological features exhibited a high degree of interindividual variability and were difficult to correlate with psychological state. In order to reduce all the calculated psychophysiological features into a smaller number of psychological variables, several methods were applied: fuzzy logic (section 4.4), principal component analysis (section 4.6), neural networks (section 4.7) and discriminant analysis (section 4.8). In addition to psychophysiological measurements, these methods can be expanded to include motor actions or virtual-scenario-specific parameters (e.g. success rate).

Finally, **offline analysis of EEG measurements** was performed by ETH and already reported in D3.1. A brief summary is reported in section 4.9.

It should be noted that two of the approaches developed for multisensory data fusion and user state assessment were extensively upgraded with automatic learning

capabilities. For these two approaches, the basic offline principle is briefly described in sections 4.3 and 4.8 while a detailed description is available in section 5.

4.2 Detection of haptic parameters for the upper extremities

4.2.1 Method background description

The virtual scenario is a dynamic environment where movements are not fully predictable, requiring the subject to be focused and perform considerate physical activity. In order to examine the influence of different features of the virtual environment on the subject's activities and performance, several haptic parameters were defined. These represent the **first step of sensor fusion** for the upper extremities, **combining raw signals from multiple force and position sensors into relevant indicators of patient motor ability**. These parameters can be applied to either the ball-catching scenario or the river scenario developed in WP2.

4.2.1.1 Catching and placing efficiency.

The catching efficiency (CE) is the percentage of caught objects (n_c) divided by the number of all objects (n_a). The placing efficiency (PE) is the percentage of the objects which were successfully placed in the basket (n_p) divided by the number of caught objects (n_c).

$$CE = \frac{n_c}{n_a} \quad (1)$$

$$PE = \frac{n_p}{n_c} \quad (2)$$

4.2.1.2 Mean Reaching Forces.

The mean reaching forces at the end-effector sensor can provide information about the direction of the intended movement. These forces were assessed from the time the object reached the center of the table or river to the time the object was caught. The sign of the force is set with respect to the position of the object. The positive sign represents the force toward the object while the negative sign represents the force away from the object. Only the horizontal component of the force was observed since this component represents the left-right movement of the subject's arm. As the reaching movements primarily occur in a horizontal plane, vertical components are largely irrelevant.

4.2.1.3 Deviation Error.

This is the percentage of the maximal deviation ($e_{\max}$) of the measured movement trajectory perpendicular from the reference line, normalized by the length of the reference line (L_r). The reference line is located in the center of the tunnel.

$$DE = \frac{e_{max}}{L_r} \quad (3)$$

4.2.1.4 Mechanical Work.

The mechanical work is computed from the measured forces at the end-effector and the end-effector positions. The computed work evaluates the interaction between the subject and the HapticMaster. Therefore it is not only the mechanical work performed by the subject. The interaction work toward the target per trial (A_{tt}) and away from the target per trial (A_{at}) were distinguished. The work away from the target represents the resistive work, when the guidance assistance is enabled.

$$A_{tt} = \int_{C_t} |F_t| \cdot ds; \quad ds > 0 \tag{4}$$

$$A_{at} = -\int_{C_t} |F_t| \cdot ds; \quad ds < 0 \tag{5}$$

4.2.1.5 Correlation between the grasping force and the load force.

The grasping forces measured during a single pick-and-place movement are divided into three phases: grasping phase, transport phase and release phase. The characteristic point of the grasping phase is when the grasping force reaches the rising time end-point. Rise time is the time required for the grasping force to change from 10 % value to 90% value. The characteristic point of the transport phase time is the central point between the grasp and the release. The characteristic point of the release phase is the falling time end-point. Fall time is defined as the time required for the grasping force to change from 90 % value to 10% value. The load force is the vertical component of the end-effector force applied by the subject. Pearson correlation coefficients were computed between the grasping force and the load force for each grasping phase and for each trial.

4.2.2 Experimental verification

After defining the various haptic parameters, it was necessary to detect these parameters in the MIMICS virtual scenarios and verify that this approach to haptic sensor fusion actually provides relevant information about user motor ability.

4.2.2.1 Goal

The haptic parameters were used to **study the influence of different modes of haptic assistance on the user** in the ball-catching task developed in WP2. Three non-adaptive modes of haptic assistance developed early in the project were examined:

- **Catching assistance.** The catching assistance helps the subject to reach the catching point. It is realized by the use of impedance controller that moves the subject's arm in a frontal plane. The assistance generates the forces when the ball reaches the center of the table, thus giving the subject the time to reach the catching point. The force is increasing when the ball is getting closer to the robot end-effector.

- **Grasping assistance.** Instead of the manual grasping, the grasping assistance causes the ball to stick to the virtual gripper. When the subject reaches the basket, the ball is dropped. If the grasping assistance is deactivated, the grasping force produced by the subject needs to be higher than a reference force. The reference force can be changed during the task according to subject's grasping ability.
- **Tunnel assistance.** The haptic trajectory tunnel enables movement from the catch point to the placing point along a defined trajectory in the virtual haptic environment. An impedance controller prevents the subject to deviate largely from the desired trajectory. The bisector of the tunnel is generated using B-splines and control points. The control points are approximated by using B-splines from trajectories measured in healthy subjects' movements. The guidance assistance provides a force in the direction of the haptic trajectory tunnel. An impedance controller leads the subject's arm along the desired trajectory.

These modes of haptic assistance should enable subjects in both the subacute and chronic phase of stroke to reach the same efficiency as healthy subjects while maximizing their effort.

4.2.2.2 Experiment protocol

The **haptic data collected during the 2009 MIMICS clinical HapticMaster study** ("Psychophysiology in post-stroke upper limb rehabilitation", full methodology described in WP4.2) was used to examine the effects of the different modes of haptic assistance.

23 subacute stroke subjects, 10 chronic stroke subjects and 23 healthy controls trained for 6 minutes in the **ball-catching scenario** using various modes of haptic assistance. These modes were activated if the subject was unable to perform a particular component of the task. Specifically, 7 subacute subjects had grasping assistance, 5 subacute subjects had catching assistance and 7 subacute subjects had tunnel assistance. Seven chronic subjects had grasping assistance, 4 chronic subjects had catching assistance and 5 had tunnel assistance. The control group performed the task without any assistance.

The parameters listed in section 4.2.2.1 were analyzed using t-tests to determine significant differences between subjects using the assistance and the ones without it. The ANOVA test was conducted to compare the three groups (subacute, chronic, control) without the assistance. Differences were expected in the t-test but no difference between groups in the ANOVA test.

4.2.2.3 Results

Results of exercise with the virtual scenario in subacute subjects showed that catching efficiency was greater when the catching assistance was applied ($p < 0.001$). The difference is not significant in the chronic group ($p = 0.507$). The control group, which performed the task without any assistance, caught more balls than the

subacute dCA (deactivated catching assistance) group ($p < 0.001$) while the difference with the chronic dCA group was not significant. There is a statistical difference in mean reaching forces between the subacute dCA subjects and the control group ($p = 0.004$). There was also significant difference between control group and the chronic dCA group ($p = 0.003$). Significant differences were found between the subacute dCA and subacute CA (catching assistance) groups ($p = 0.022$) as well as between the chronic dCA and chronic CA groups ($p = 0.030$).

Table 1: The results of observed catching efficiency (CE) and mean forces (MF) during the catching phase of the task. The subacute and chronic subjects are divided into the groups with catching assistance (CA) and without catching assistance (dCA). n is the number of subjects.

	Subacute dCA (n = 18)	Subacute CA (n = 5)	Chronic dCA (n = 6)	Chronic CA (n = 4)	Control dCA (n = 23)
CE [%]	63 ± 17	86 ± 14	62 ± 21	78 ± 27	86 ± 13
MF [N]	0.26 ± 0.26	-0.28 ± 0.51	0.11 ± 0.15	-0.42 ± 0.43	0.03 ± 0.07

Table 2: The results of placing efficiency (PE), deviation error (DE), work performed toward the target (WTT) and work performed away from the target (WAT). The subacute and chronic subjects are divided into the groups with tunnel assistance (TA) and without tunnel assistance (dTA). n is the number of subjects.

	Subacute dTA (n = 16)	Subacute TA (n = 7)	Chronic dTA (n = 5)	Chronic TA (n = 5)	Control dTA (n = 23)
PE [%]	79 ± 14	98 ± 6	78 ± 16	100 ± 0	91 ± 9
DE [%]	37.9 ± 16.4	6.9 ± 1.8	29.4 ± 18.2	7.4 ± 3.4	39.4 ± 26.8
WTT [J]	1.39 ± 0.65	0.12 ± 0.38	1.87 ± 1.55	0.01 ± 0.17	1.23 ± 0.91
WAT [J]	0.02 ± 0.40	0.18 ± 0.28	0.19 ± 0.38	0.66 ± 0.83	0.03 ± 0.27

The control group performed pick-and-place movements more successfully than the subacute dTA (deactivated tunnel assistance) and the chronic dTA group during the virtual rehabilitation task ($p = 0.001$). The subacute TA (tunnel assistance) and chronic TA subjects were unable to do any pick-and-place movements by themselves. The subacute TA group efficiency was greater than the efficiency of the subacute dTA group ($p < 0.001$). The same holds for chronic dTA and chronic TA group ($p < 0.001$). Figure 1 shows the deviation error of the stroke subjects with and without tunnel assistance as well as the deviation error of the control group. The deviation error was greater in the subacute dTA group than in the subacute TA group ($p < 0.001$) as well as in the chronic dTA group ($p < 0.001$) with respect to the chronic TA group. The chronic dTA group had lower deviation error than the subacute dTA group and control group ($p < 0.001$). Figure 2 shows the work performed toward the target for single pick-and-place movements. The subacute dTA group performed more work toward the target than the subacute TA group ($p < 0.001$). We can observe the same in chronic subjects ($p < 0.001$). The chronic dTA group performed more work than the control group and more than the subacute dTA group ($p < 0.001$). Figure 3 shows the work performed away from the target for single pick-and-place movements. The subacute dTA subjects performed less work than the subacute TA subjects ($p < 0.001$). The chronic dTA group also performed less work than the chronic TA group ($p < 0.001$). The control group performed significantly more work than the subacute dTA group and the chronic dTA group ($p = 0.005$).

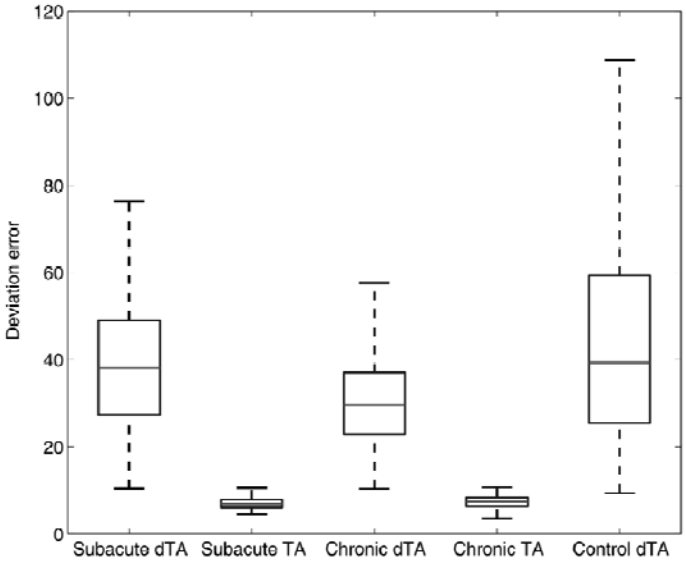

Figure 1: Deviation error of the pick-and-place movement with respect to the predefined central curve line. The results are shown for subacute, chronic and control group without tunnel assistance (dTA) as well as for subacute and chronic group with tunnel assistance (TA).

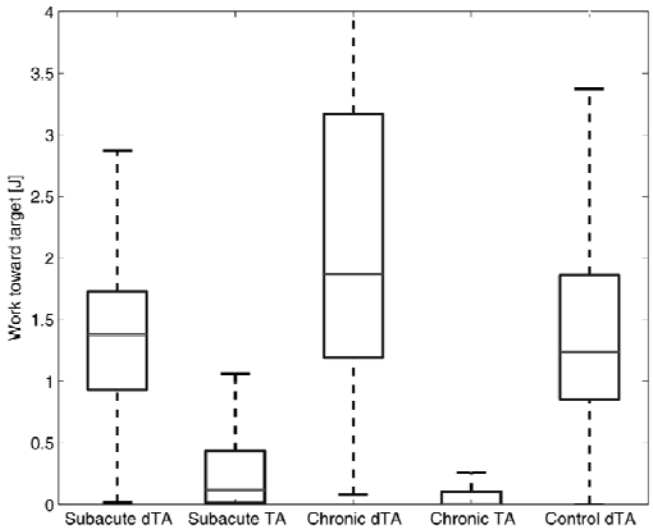

Figure 2: Comparison of the performed work toward the target during pick-and-place movement for the subacute, chronic and control group with deactivated tunnel assistance (dTA). The results of chronic and subacute group with tunnel assistance (TA) are also shown.

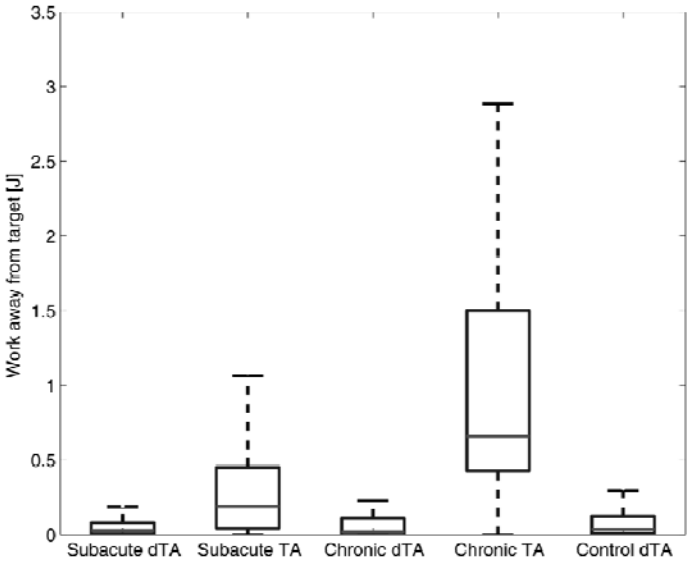

Figure 3: Comparison of the performed work away from the target during pick-and-place movement for the subacute,chronic and control group with deactivated tunnel assistance (dTA). The results of chronic and subacute group with tunnel assistance (TA) are also shown.

Table 3: Results for grasping force rise time (RT), grasping force fall time (FT), correlation between grasp force and load force for grasping phase (CGP), transport phase (CTP) and release phase (CRP). These groups had grasping assistance deactivated (dGA). n is the number of subjects.

	Subacute dGA (n = 16)	Chronic dGA (n = 3)	Control dGA (n = 23)
RT [s]	0.14 ± 0.45	0.47 ± 0.40	0.17 ± 0.34
FT [s]	0.33 ± 0.30	0.54 ± 0.15	0.29 ± 0.39
CGP	0.03 ± 0.58	0.23 ± 0.58	0.12 ± 0.58
CTP	0.01 ± 0.51	-0.36 ± 0.59	0.41 ± 0.58
CRP	0.90 ± 0.40	0.88 ± 0.42	0.89 ± 0.30

The Pearson correlation coefficient is computed for each movement in each grasping phase. The correlations for the subacute, chronic and control groups are shown in Figure 4. Only the subjects who had grasping assistance deactivated are considered. While the correlations are widely spread from -1 to 1 in the grasping and transport phase, the correlation between grasp force and load force exists in release phase. These results are shown for subacute, chronic and control groups, respectively. While there are no differences among groups in grasping and release phases ($p = 0.210$; $p = 0.218$), there is a significant difference in control group during transport phase ($p < 0.001$). There is only a small difference in grasping rise time between the subacute and control groups ($p = 0.004$). The rise time of the grasping force is longer in the chronic group than in control ($p < 0.001$) or subacute group ($p < 0.001$). These relationships are similar for the fall time of the grasping force. There are no differences between stroke and control group ($p = 0.481$) while the chronic group had a longer fall time compared to subacute ($p < 0.001$) and control ($p < 0.001$) groups.

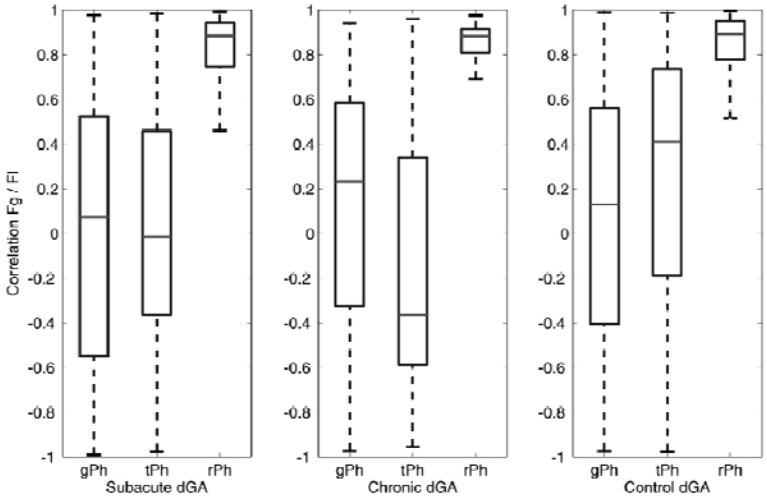

Figure 4: Correlation between the grasping force and the load force for each phase separated: grasping phase (gPh), transport phase (tPh) and release phase (rPh). The load force is the vertical component of the measured force on the end-effector. The results are shown for subacute, chronic, and control group who had no grasping assistance (dGA).

4.2.3 Conclusions

The different haptic parameters highlighted **several important effects of haptic assistance in virtual rehabilitation**, showing the relevance of these parameters as a method of sensor fusion and estimation of user motor ability.

Haptic assistance raises the catching efficiency and pick-and-place efficiency. The results also showed that the stroke subjects without assistance and the control subjects performed similarly. Thus, **haptic assistance is one possible way to improve subjects' performance**. However, **effective work greatly decreases when assistance is enabled**. In most cases, the subjects are not able to perform the movement by themselves, so some assistance is definitely needed. However, it **must be applied properly** or it will not encourage the subject to be more active.

The lessons learned from this study were applied to the creation of an adaptive haptic assistance system that aids the subject only as much as necessary, therefore maximizing the effort applied by the subject and the benefit gained. This system is described in section 4.3.

4.3 Haptic sensor fusion for adaptive haptic assistance

Given the complexity of the upper limbs due to the large number of degrees of freedom which makes the human arm a highly redundant mechanism, a robot-assisted rehabilitation device should allow variability in redundant dimensions. Building on research conducted by ETH and UL prior to MIMICS (Mihelj, Nef, & Riener, 2007), the haptic parameters described in the previous section were expanded to provide the basis for an adaptive haptic assistance system, which, instead of enforcing a desired trajectory, **uses feedback and force fields** to **correct deviations that interfere with task performance**. In this way, the feedback

controller's only concern is task performance, not the precise position of the user's hand. The patient is allowed to select the most comfortable movement trajectory and follow this trajectory as much as possible given their muscle coordination capabilities. The **feedback controller enhances the movement** by minimizing the movement jerk index. This support is active only when the hand's distance from the target is greater than a certain optimal distance from the target. This optimal distance decreases with time, so the robot's assistance will increase if it recognizes that the patient is unable to make effective movements.

While the sensor fusion provided the groundwork for implementation of adaptive haptic assistance, the system also utilized several automatic learning principles. For this reason, it is fully described in section 5.2.

4.4 Fuzzy logic

4.4.1 Method background description

Fuzzy logic is a type of mathematical logic that can be used for pattern recognition. In fuzzy logic, the connections between inputs and outputs are defined using relatively **simple IF-THEN rules**. This makes the pattern recognition system very easy to understand. Following experiments carried out at UL in 2008, it was decided to create a fuzzy system that would **reduce measured motor actions and psychophysiological features into a three-dimensional user state** consisting of:

- the level of physical effort,
- the level of psychological arousal,
- the level of emotional valence (positive or negative emotions).

This definition of psychological state was based on the arousal-valence model defined by (Russell, 1980). This model was expanded with the physical effort dimension in order to be better suited for a motor rehabilitation setting. The IF-THEN rules mapping heart rate, skin conductance, respiration, skin temperature and motor actions to user state were defined based on studies by other authors and the initial studies performed within MIMICS. A similar fuzzy model had previously been created by (Mandryk & Atkins, 2007), but could not be used directly since different psychophysiological measurements were used in that study.

A graphical representation of the connections between physiology and the three dimensions of user state is shown in Figure 5.

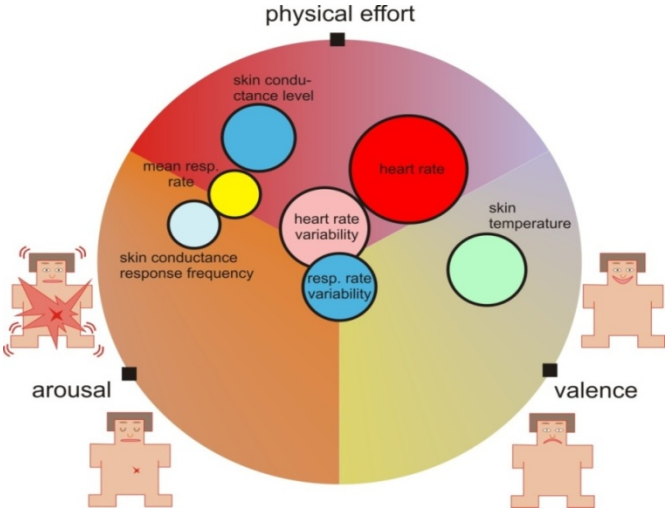

Figure 5: Relationships between various psychophysiological parameters and arousal, valence or physical effort. Each parameter is represented by a circle. The distance of a circle's centre from each of the three black squares represents the effect that square has on it; the shorter the distance, the greater the effect. The radius of a circle represents variability between subjects; a larger radius means that physiological responses vary more strongly between individuals.

The actual implementation of fuzzy mapping is illustrated in Figure 6. Force and movement data were used to estimate physical effort, skin conductance and respiration were used to estimate arousal, and respiration and skin temperature were used to estimate valence.

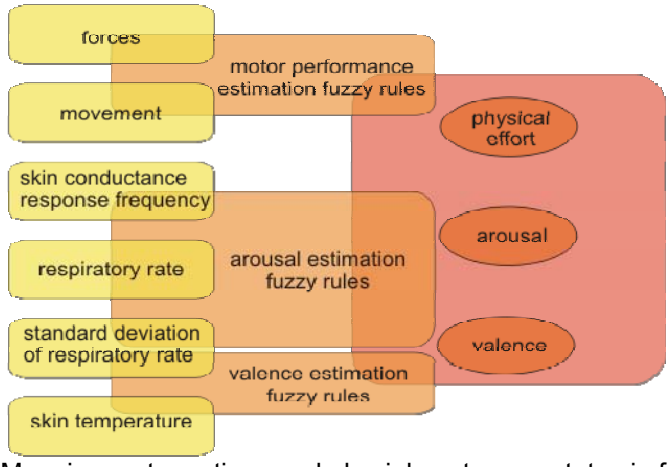

Figure 6: Mapping motor actions and physiology to user state via fuzzy logic.

4.4.2 Experimental verification

The fuzzy system realisation was used to estimate arousal, valence and physical effort in the data set obtained from the second HapticMaster experiment at UL ("Arousal and valence in the presence of physical activity", detailed in D4.2) where physiology and force/movement data were available together with results of arousal-valence questionnaires. However, **results proved to be difficult to interpret and validate**. While the fuzzy system did output estimates of psychological state, it was very difficult to verify these results due to discrepancies between physiological and self-reported indicators of arousal and valence that had been previously noted by other authors (Schwerdtfeger, 2004) and later documented in our own clinical studies (Novak, et al., 2010).

4.4.3 Conclusions

Due to **difficulties in verifying the fuzzy system**, it was decided to shift focus away from the arousal-valence psychological model at this stage and focus on other definitions of psychological state (e.g. underchallenged, challenged, overchallenged). Additionally, we felt that manually defined fuzzy rules are unreliable and that **a more quantitative method of defining connections between psychophysiological measurements and psychological state that did not require expert knowledge was required**. Thus, three well-known pattern recognition methods were explored: principal component analysis (section 4.6), neural networks (4.7) and discriminant analysis (4.8).

4.5 Model of heart rate response to physical load

4.5.1 Method background description

Heart rate is influenced by biomechanical activity as well as by psychological changes as fear or mental concentration (Borg, Hassmen, & Lagerström, 1987; Bernardi, Valle, Coco, Calciati, & Sleight, 1996; Carter, Banister, & Blaber, 2003). The biomechanical activity, i.e. the power that subjects have to expend when walking in the Lokomat, can alter the physiological signals compared to resting condition. The biomechanical dependency between power expenditure and heart rate can be modelled and taken into account during gait training.

Once a model of heart rate response as a function of gait training parameters is available, it is possible to subtract the predicted heart rate (which is only dependent upon power exchange) from the recorded heart rate. The changes in heart rate compared to baseline which are left should then be caused solely by changes on the psychological state of subjects, assuming a perfect model. There will be an error term dependent on the precision of the model.

4.5.2 Experimental verification

Heart rate was recorded in five healthy individuals to define a model for the cardiovascular process of subjects during Lokomat walking (3 m and 2 f, 25.0 yr pm 2.3, 77.2 kg pm. 8.0). Mean heart rate was extracted from the electrocardiogram using an RR interval detection algorithm. The study protocol was approved by local ethics committees and subjects gave informed consent. We varied the three Lokomat parameters: treadmill speed, guidance force and body weight support.

Subjects walked at three walking speeds [1, 2 and 3 km/h], three guidance forces [0%, 50% and 100%] and three different levels of body weight support [0%, 30% and 60%]. Note that a guidance force of 100 % meant a maximally stiff impedance controller and 0 % a fully transparent orthosis. Note that walking speed was limited to

3.2 km/h, as 3.2 km/h is the maximal speed which is allowed in the Lokomat for patient safety. The dependency between walking and heart rate of healthy subjects has been previously investigated (Vokac, Bell, Bautz-Holter, & Rodahl, 1975; Borg, Hassmen, & Lagerström, 1987; Baum, Essfeld, Leyk, & Stegemann, 1992; Arts & Kuipers, 1994; Bernardi, Valle, Coco, Calciati, & Sleight, 1996; Achten & Jeukendrup, 2003). Increases in treadmill speed were shown to linearly increase heart rate. This can be interpreted as lowpass-like reaction to a sudden increase of oxygen demand, which we modeled as a first order delay (PT) element. Treadmill acceleration resulted in an overshoot in heart rate before steady state was reached. An undershoot was observed after a negative acceleration. Heart rate overshoot might be caused by a first overreaction of the cardiovascular system to compensate for the blood pressure drop. Feroldi et al. (Feroldi, Belleri, Ferretti, & Veicsteinas, 1992) argued that the overshoot might be a result of changes in the balance between sympathetic and parasympathetic activity. The overshoot and undershoot behavior was modeled as a second order derivative (DT) element. The power expenditure of the human was taken as a linear input parameter modeled as a first order PT element. After longer training durations a fatigue effect, which resulted in increased heart rate at steady state, was observed and described by several researchers. We modeled this as a first order lowpass element. This resulted in a model with five scaling factors and six parameters. Estimation of the the scaling factors and parameters for each subject were done using a genetic algorithm in combination with gradient descent optimization, as we wanted to explore the whole parameter space for solutions to find the global minimum.

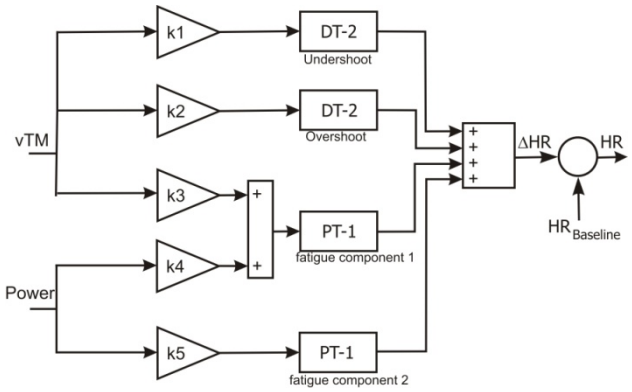

Figure 7: Simulink chart of the heart rate model

Validation of the model was performed with a velocity profile different from the one used for model identification. The goodness of fit was assessed with the coefficient of determination, R^2. R2 is computed from the sum of squares of the whole output data. First, SS_{tot}, the total sum of squres needs to be computed.

$$SS_{tot} = \sum_i (y_i - \bar{y})^2 \quad (6)$$

Where y_i are the output data points at time i and $\bar{y}$ is the mean of the output data. Then, the sum of squared residuals is computed as

$$SS_{err} = \sum_i (y_i - f_i)^2 \quad (7)$$

Where f_i are the model values at data point i. The coefficient of determination R^2 is then

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}} \quad (8)$$

We performed an online identification of subject specific parameters, optimizing only over the first 13 minutes of the speed profile. After the first 13 minutes, the model parameters were fixed and used for heart rate prediction during training sessions of 37 minutes.

The results show that prediction of heart rate is possible in subjects. The method only failed in subject 3, who did not show an increase of heart rate for any condition. These results were encouraging, as they showed that linear changes in heart rate can be predicted.

Table 4: Prediction quality of the heart rate model with optimization of 4 parameters. Comparison between an optimization over the whole data with an optimization which opitmizes over the first 13 minutes and then predicts the rest of the dataset. Subject 1 and 2 were both recorded three times.

Subject	Recording number	R^2 of the whole dataset for optimization over the entire dataset	R^2 of the whole dataset for optimization over the first 13 minutes
1	1	0.59	0.47
	2	0.82	0.81
	3	0.73	0.68
2	2	0.90	0.89
	3	0.73	0.67
	1	0.91	0.91
3	1	0.68	0.28
4	1	0.75	0.72
5	1	0.86	0.81
Average	-	0.78	0.70

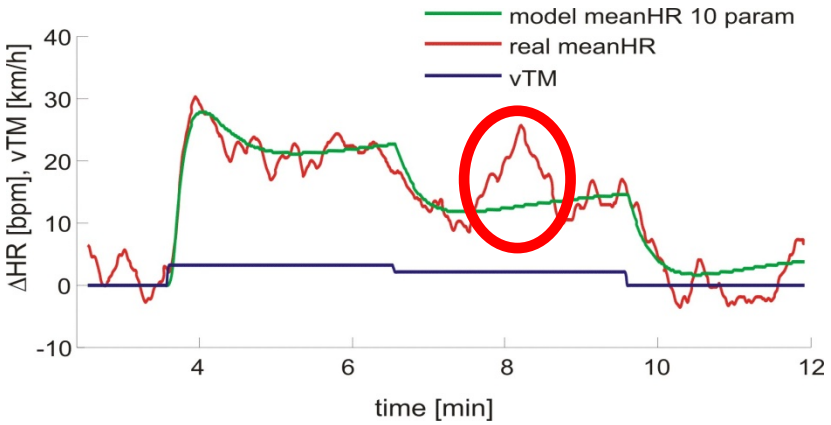

Figure 8: Predicted vs. recorded heart rate of one example session of an unimpaired subject. The red circle indicates a moment during training where the subject was exposed to an exciting stimulus. Due to the psychological change, heart rate increased. As this change was not caused by a change in biomechanical activity, it was not caught by the model. The subtraction between recorded and predicted heart rate would then show changes in heart rate caused by psychological influences.

4.5.3 Conclusions

Predicting heart rate is possible based on changes in power expenditure in the Lokomat and changes in treadmill speed. The drawback of this approach is that 12 minutes of initial identification are necessary to identify the model parameters. We therefore conclude that this model can be used in laboratory settings to investigate the effects of power expenditure on heart rate compared to the effects of psychological changes. However, a clinical application on a day-to-day basis will not be practical due to the long identification phase of the model.

4.6 Principal component analysis and neural networks

We performed experiments with 17 healthy subjects and 10 patients and used a virtual environment scenario in the lower extremity setup to elicit three different states of mental engagement: bored, challenged and over-stressed. Using questionnaires, we verified that we were able to induce these different mental states in subjects.

We evaluated whether all six physiological signals (heart rate, heart rate variability, skin conductance level, number of skin conductance responses per minute, breathing frequency and skin temperature) were needed to classify the mental engagement, and which sensors could possibly be omitted without significantly decreasing the classifier performance. This was of interest to improve clinical applicability of our approach: attaching the sensors to the patient's body was time consuming and reduced the time left for training. Furthermore, sensors such as the breathing sensor were perceived to be disturbing by the patients and patient comfort during the training could be increased without this sensor.

4.6.1 Method background description

4.6.1.1 Principal component analysis (PCA)

PCA was used to identify, which physiological signals and their combination explained most of the variance of the recordings and could therefore be seen as the major markers for changes in psychological states. Inputs to the PCA were heart rate (HR), a discrete time series of heart rate variability (HRV), a discrete time series of the number of skin conductance response (SCR) events, skin conductance level, the skin temperature and a discrete time series of the breathing frequency. PCA is a linear, orthogonal transformation, which projects the original data into a new coordinate system. In this new coordinate system, the first axis (or first principal component (PC)) is that one that it explains most of the variance. The second PC explains the second most important variance, etc. The PCA is computed such that

$$D^{(n \times m)} = A^{(n \times m)} * F^{(m \times m)} \quad (9)$$

where D is the original data, A the activation coefficients and F the loading factors. In our case, the original data consisted of a time series with n data points and m = 6 physical recordings (dimensions). The activation coefficients were again a time series

with 6 dimensions. The matrix of loading factors was the rotation matrix and defined the new coordinate system.
We computed, how many factors k were necessary to explain more than 80% of the variance in all subjects ($k \in [1,n]$), where n is the dimensionality of original data, i.e. 6). A factor rotation on these first k PCs was performed to obtain a clearer picture, which input signals provided the largest variance. Factor rotation is a mathematical transformation that does not alter the subspace spanned by the PCs, but shifts the weight of an input e.g. from the first PC to the second, while maintaining the orthogonality between the components. All data was normalized in amplitude before applying PCA.

4.6.1.2 Neural networks

The term 'neural networks' summarizes several nonlinear methods often used in pattern recognition, particularly in automatic handwriting recognition. Neural networks have an input layer with an input dimensionality of the raw data that is to be processed (Figure). Often, a hidden layer of neurons is the directly connected to an output layer. This structure is known as a perceptron and was introduced by Rosenblatt (Rosenblatt, 1962). The weights between the layers are the tunable parameters which are adjusted during an initial training period. After the training, the weights are set to be fixed.

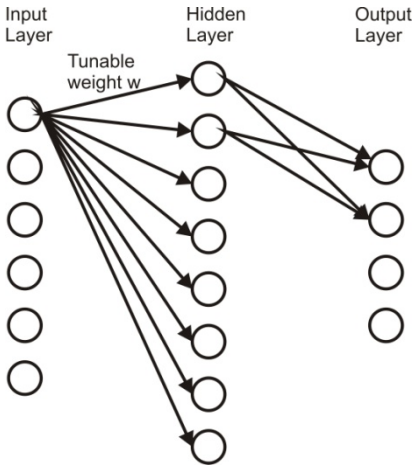

Figure 9: General setup of a neural network

For automatic classification of the mental state of a subject from physiological recordings, we used a data fitting perceptron containing 30 hidden layer neurons. The implementation was performed in Matlab (www.mathworks.com). There were eleven input layer neurons (Heart rate, heart rate variability, skin temperature, skin conductance level, breathing frequency, skin response events, the change in skin temperature, the change in skin conductance, the coefficient of low over high frequencies in the power spectrum of heart rate, the mean biofeedback and the standard deviation of the biofeedback) and four output layer neurons, one for each class (baseline, underchallenged, challenged and overchallenged).

The neural network provided an estimation of the current state of mental engagement, based on the physiological recordings. 20% of the data was taken as

training data, 20% as validation and 60% as testing data. Learning was performed with the Levenberg-Marquardt back-propagation algorithm. We chose the data necessary for training of the network such that an identification phase would not take longer than 20% of the whole training time in the robot.

We compared the performance of the classifier by means of the mean squared classification error for two sets of input data. The neural network classified the mental engagement from all n=6 raw physiological data streams. We then performed classification of the current psychological state with a reduced dataset that contained only the physiological recordings that were dominant in the first k PCs (k<n). We selected the number k of PCs by means of variance explained and decided for a value above 80%.

4.6.2 Experimental verification

The first two PCs explained more than 70% variance in all subjects, healthy as well as patients; the first three PCs explained more than 80% in all but one healthy subject and all patients (Figure 10).

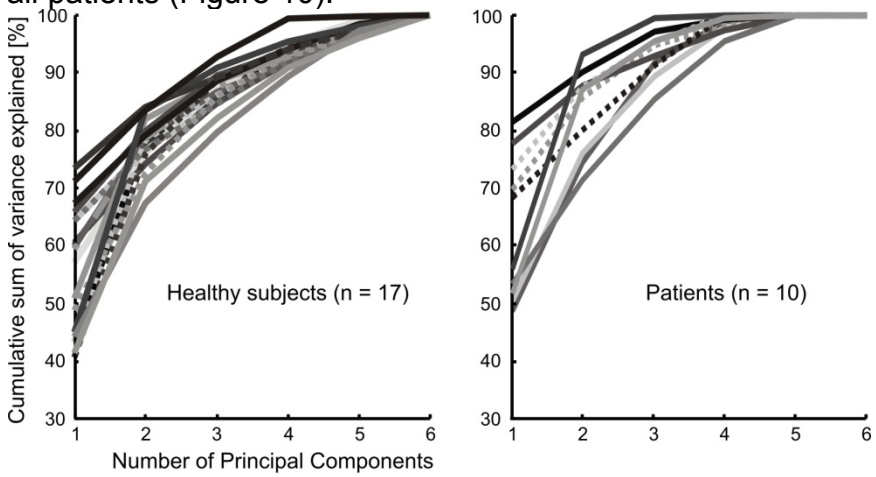

Figure 10: Variance explained by PCs for healthy subjects and patients. As PCA performs without loss of information, the correlation between original signal and PCA decomposed signal must be 100% if all components are added up

In contrast to the direct statistical analysis of the physiological data, the PCA allowed direct visual separation of the four conditions baseline, bored, excited and over-stressed. A typical example is shown in Figure 11 for healthy subject 17). Loading factor one of this particular plot relies mainly on skin temperature, while loading factor two is a combination of HR and skin conductance level (Table 5). The different conditions can be visually distinguished amongst each other (as depicted by ellipsoidal boundaries in Figure 11). The temporal evolution of the PC activation coefficients are displayed via changes in color. The lighter the color of each condition, the earlier in the condition the data was recorded. The black arrows mark the general evolution of the activation coefficients over the time course of one condition. Although we performed classification on three PCs, we only plotted two dimensions, as a 3-D plot would be difficult to display.

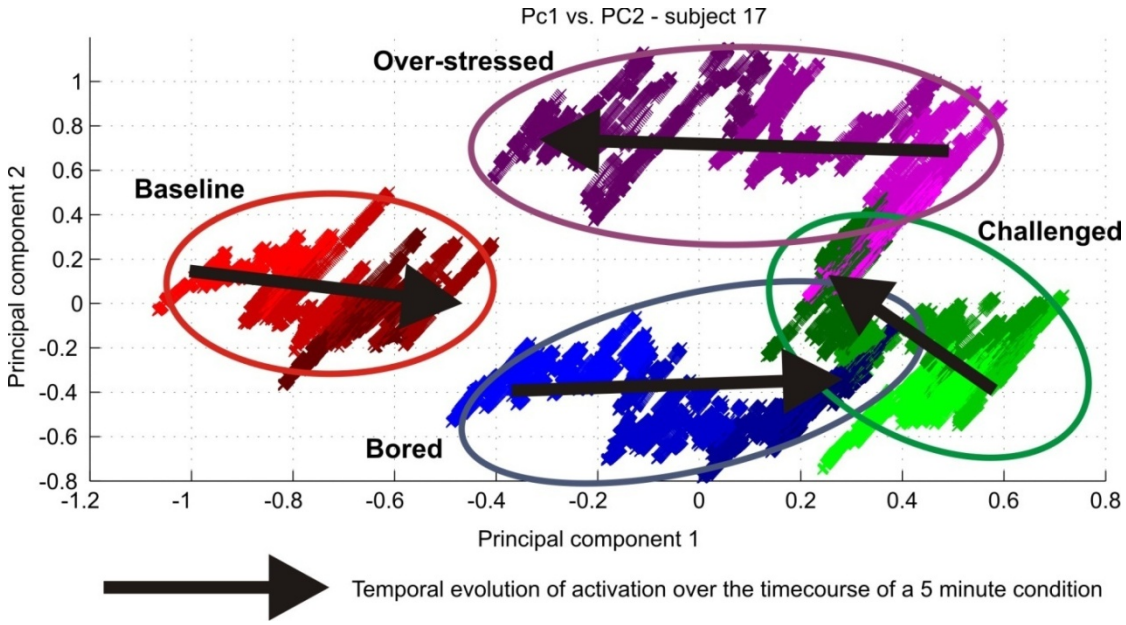

Figure 11: First two activation coefficients of the PCA for subject 17, separated for the four conditions "baseline", "bored", "challenged" and "over-stressed", plotted for the whole length of each condition (five minutes). Within one color, the darkness of the color symbolizes a later time during the condition.

Table 5: Loading factors of the first PCs of subject 17

	Healthy subject 17	
	PC 1	PC 2
HR	0.2471	0.6025
HRV	0.0168	0.0585
SCR	0.0002	0.0607
SCL	-0.1987	-0.7271
Breathing	-0.0093	-0.0668
Skin Temp	-0.9482	0.3111

Evaluation of the loading factors (i.e. the transformation matrix obtained from the PCA) of all healthy subjects and all patients revealed that PC 1 was dominated by skin temperature and skin conductance level in both groups (Figure 12). HRV played a minor role in the first three PCs for both groups. While breathing frequency was not dominant in healthy subjects, data variability of patients showed a major dependency of breathing frequency in PC 3.

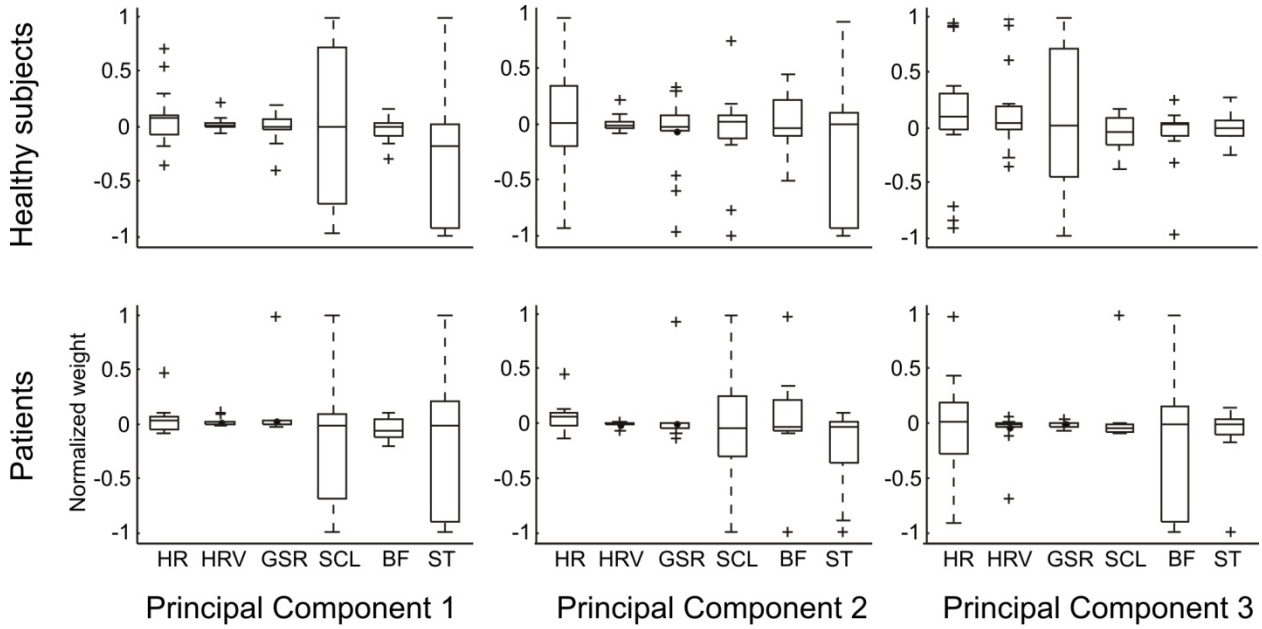

Figure 12: Comparison between the loading factors of the first three PCs of healthy (n=17) subjects and patients (n=10). The physiological recordings used to extract PCs are heart rate (HR), heart rate variability (HVR), galvanic skin responses (GSR), skin conductance level (SCL), breathing frequency (BF) and skin temperature (ST).

Although we only found two significant differences in all physiological recordings over all conditions of patient data, the classification of the different psychological states using a neural network was possible for healthy subjects and patients. As described in the methods section, we evaluated the classification results or a neural network for two different sets of input data: on the one side with six physiological parameters extracted, on the other side using only the physiological signals dominant in the first three PCs. Mean classification error was 1.4% for the full and 2.5% for the reduced dataset in healthy subjects and 2.1% for the full and 4.7% for the reduced dataset for patients (Figure 13).

In 10 out of 17 healthy subjects, the three different task level conditions and the baseline condition could be visually classified by plotting the first two PCs. This is exemplarily shown Figure . In these subjects, automatic classification via a neural network was possible with very little classification error measured by the root mean square error (RMSE). In five of 17 subjects, a graphical separation as in Figure was not possible, while the neural network still performed with less than 3% RMSE. Classification of the remaining two of 17 healthy subjects was possible with 5% and 13% RMSE respectively (Figure 13).

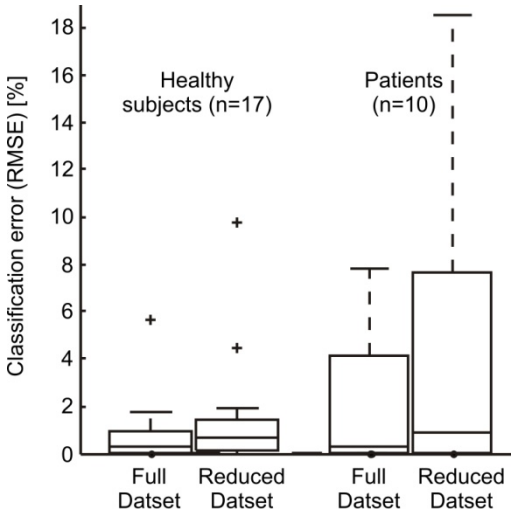

Figure 13: Classification error by means of the root mean square error (RMSE) for all 17 healthy subjects (left) and all 10 patients (right), using all six physiological signals (full dataset) and using only the three signals that were dominant in the first three PCs (reduced dataset)

In 8 out of 10 patients, classification was possible with less than 5% error when using the full set of all physiological signals. The classification results for the different sets of inputs are displayed in Figure 7. In patients, the classification was successful, but to a lesser extent. Particularly in subject ten, the classification error was 18% for a reduced set of input data. Using all recorded physiological data, an overall classification error of less than 7% was possible for all patients. Training of the neural network was possible in less than one minute on a standard Pentium processor (1.4GHz, 2GB RAM).

4.6.3 Conclusions

Neural networks can successfully separate four different classes of mental engagement during robot-assisted rehabilitation, but there are several drawbacks to this approach: first, the neural networks are non-adaptable. Once trained, the weights are set fixed and cannot easily be adjusted. Adaptive neural networks exist, but their convergence at runtime is not guaranteed. Second, the network performed poorly in generalizing. When trained on n-1 subjects, classification performance of data of the n-th patient dropped significantly and averaged around 70%. This means that neural networks are a good option when each therapy session starts with initial network training.

4.7 Discriminant analysis

Discriminant analysis is a popular **statistical method for the classification of multidimensional data into two or more distinct classes**. Based on a training data set with annotated class labels, one or more discriminant functions are generated that optimally classify the training data. The discriminant functions can be linear (linear discriminant analysis – LDA) or quadratic (quadratic discriminant analysis – QDA). Fundamentally, discriminant analysis is similar to principal component

analysis (section 4.6). Both look for combinations of variables that best explain the variability of the data, but discriminant analysis takes class labels into account and is thus more suitable for classification. Discriminant analysis is thus a popular tool for multisensory fusion, since it allows a large number of different inputs to be reduced into a class estimate based purely on training data, without any need for knowledge of the system's actual internal model.

Within MIMICS, discriminant analysis can be used for **fusion of motor actions and psychophysiological responses into classes** such as "task is too easy/task is too hard", "user is not sufficiently challenged/user is overchallenged/user is optimally challenged". This can be done offline, but all experimental work within MIMICS was done in the context of closed-loop, adaptive systems where discriminant analysis is expanded with automatic learning algorithms. Thus, detailed information about methods, results and conclusions is provided in section 5.4, "Adaptive discriminant analysis".

4.8 Offline assessment of electroencephalography

4.8.1 Method background description

Electroencephalography (EEG) is the method of measuring the brain's electrical activity using electrodes placed on the head. In MIMICS, it was considered in a limited number of trials for offline analysis and investigation of possible correlations between the sense of presence, engagement and motivation on one hand and brain activity reflected in EEG signals.

4.8.2 Experimental verification

EEG measurements were tested in an early experiment at ETH (reported in Deliverable D3.1). However, there were two major problems with these measurements. First of all, although the cables were actively shielded, head movement caused cable dragging in some subjects (see Figure 14). Obvious artifacts were removed using independent component analysis (ICA) or not included in the analysis if the amplitude exceeded ±80µV.

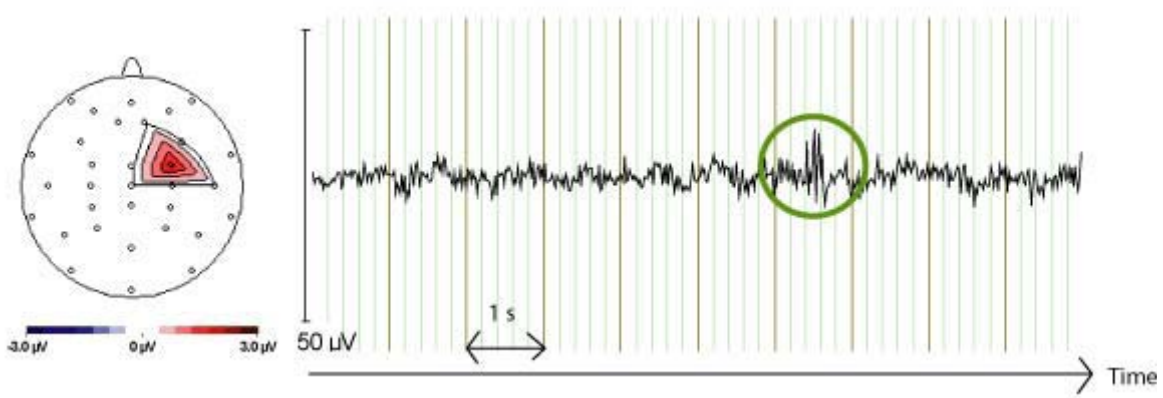

Figure 14: Cable dragging on FC4-channel (see green circle) in one subject, visible as an independent component in ICA.

Second, eye movement artifacts were a major problem in the Lokomat in combination with virtual reality. Some artifacts remained even if subjects were asked to close their eyes. Eye movement that was not linked to the gait cycle was successfully reduced offline using ICA. However, not all of the eye movement artifacts could be removed since this movement was temporally closely related to the cortical activity of interest. Removing all eye movement artifacts would have resulted in the removal of the activity of interest (Figure 15).

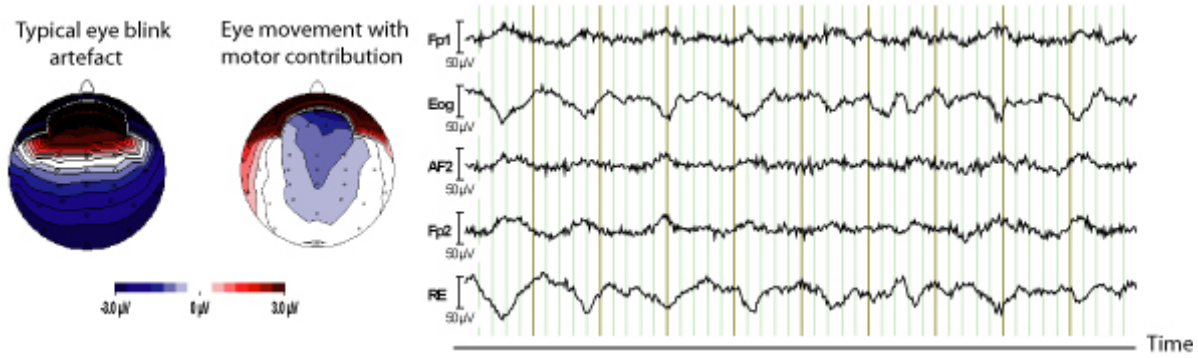

Figure 15: Eye movement artifacts in filtered EEG data of one subject (right side). Eye movement artifacts on Eog (left eye) and RE (right eye) channel were characterized through slow oscillations and eye movement coupled on central activity in one subject, visible as an independent component in ICA (left side). Another independent component (representing isolated eye blink) is shown for comparison.

4.8.3 Conclusions

Our results showed that EEG is not yet an appropriate method to detect brain activity during walking in virtual reality. It is possible to record EEG during walking, but only when the subject has his/her eyes closed. Walking induces rhythmic up and down movements of the eyes which appear as artifacts in the EEG. When walking at a constant velocity with closed eyes, the eyes still move, but the frequency of the walking movement is constant and can be filtered out. While walking in VR, subjects have to move their eyes to capture the whole scenario additionally to the movements induced by walking. This additional eye movement results in artifacts which can not be completely removed without also removing significant information from the signal, as the movement frequency is not constant and unknown. EEG measurements will not be used in further psychophysiological measurements.

5 Automatic learning system

5.1 Introduction to the automatic learning system

Within this task, machine learning systems were designed that learn the associations between environment cues and participant responses (behavioural and physiological). The ultimate goal is to **adapt the virtual environment so that the user is always optimally challenged** – not bored but also not frustrated.

The virtual environment has to adapt autonomously, based on the user's motor performance as well as his/her physiological and psychological state. The learning system will thus be used to generate a set of rules describing the reaction of the patient to the changes in the virtual environment. This set of rules will provide instructions for dynamic adaptation of the virtual environment to the patient's performance and psychopysiological state. Based on these rules, decisions will be made to bring the patient into the optimal state (level of challenge).

Similarly to Task 3 (Fusion of multi-sensory data streams and real-time user state assessment), work was first performed on an **automatic learning system that utilizes only motor actions** (movement, force and torque data). This algorithm is described in section 5.2. Then, **reinforcement learning** was used to learn the set of rules describing the reaction of the patient to the changes in the virtual environment. It is described in section 5.3. While useful for a small number of input signals (in our study, one physiological and one performance input were used), it proved unsuitable for real-time learning when a large number of inputs is involved. Thus, **adaptive discriminant analysis**, which learns from a prerecorded data set, was introduced for situations where a large number of variables (from both motor actions and psychophysiology) must be analyzed. It is described in section 5.4.

5.2 Adaptive haptic assistance

It has been shown that, when moving the hand between two points, a healthy person tends to follow a straight line, minimizing the movement jerk (Flash & Hogan, 1985):

$$J(x) = \int_{t_0}^{t_f} \left(\frac{d^3 x}{dt^3} \right)^2 dt \qquad (10)$$

Often, the patient cannot exert such optimal movement due to muscle coordination impairment. Instead they might generate sufficient voluntary movement towards the target by following the path that meets their muscle coordination demands/capabilities. Therefore to allow the patient to arbitrarily select the most comfortable movement trajectory, the feedback controller should support the

movement of the patient's hand towards the target without predefining the trajectory. To motivate the patient to move along the optimal path, the maximal time allowed is predefined.

The controller developed in this section is built on previous (pre-MIMICS) work by ETH and UL (Mihelj, Nef, & Riener, 2007) and focuses on simple point-to-point reaching position-based movement (orientation of the hand is not considered). Given the starting and target positions, the optimal time course is determined by minimizing the cost function (10). The result is a time-based trajectory, which only determines the optimal distance to the target at certain time, but not the actual reference hand position. Therefore the position distance indicates the distance between the hand position $p_h(t_k)$ and the target position p_t:

$$d_h(t_k) = \|p_t - p_h(t_k)\| \tag{11}$$

Such definition directly leads to the following definitions for the hand distance to the target and the hand velocity:

$$\begin{aligned} {}^{\Psi}x(t_k) &= \begin{bmatrix} 0^{1\times 2} & d_h(t_k) \end{bmatrix}^T \\ {}^{\Psi}\dot{x}(t_k) &= \Psi^T(t_k) v_h(t_k) \end{aligned} \tag{12}$$

In (12) $v_h(t_k)$ is the hand linear velocity in the reference frame, Ψ is the transformation between local and reference coordinate frames and the superscript ${}^{\Psi}$ indicates the vector expressed in the local coordinate frame. Solving the optimization problem (10) defines a minimum jerk trajectory to be followed:

$$\begin{aligned} {}^{\Psi}x_r(t_k) &= {}^{\Psi}x(t_0)\left(1 - 6\tau^5 + 15\tau^4 - 10\tau^3\right) \\ {}^{\Psi}\dot{x}_r(t_k) &= {}^{\Psi}x(t_0)\frac{30\tau^4 - 60\tau^3 + 30\tau^2}{t_f - t_0} \end{aligned} \tag{13}$$

where t_0 and t_f are initial and final time respectively, and $\tau = t_k / (t_f - t_0)$ is the dimensionless time (Figure 16).

Equation 13 determines how a system should optimally move from rest to target position in a desired time without accounting for changing the target position or limb perturbations during the movement. However the same result may be obtained if we use the following feedback controller for monitoring the target and hand positions to ensure that the current desired change in hand position is always such that it brings the hand in a minimal jerk trajectory to the target:

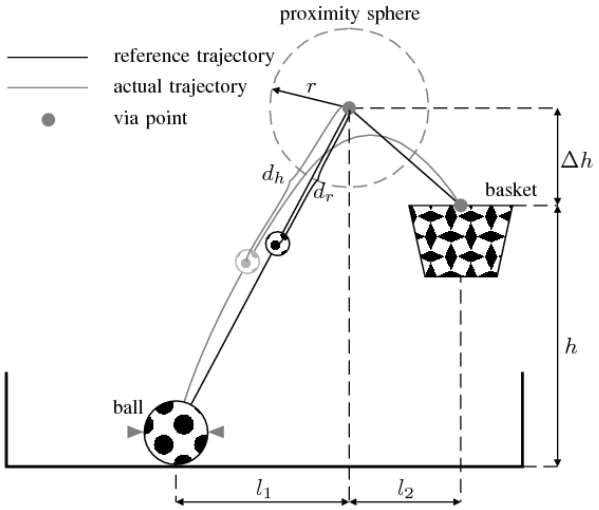

Figure 16: The adaptive haptic support system allows the user to deviate from a reference trajectory and acts only when the user is outside a certain optimal area.

$$\dot{Q}(t_k)=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{60}{(t_f-t_k)^3} & -\frac{36}{(t_f-t_k)^2} & -\frac{9}{(t_f-t_k)} \end{bmatrix} Q(t_k)+\begin{bmatrix} 0 \\ 0 \\ \frac{60}{(t_f-t_k)^3} \end{bmatrix} d_h(t_0) \qquad (14)$$

$$^{\Psi}x_r(t_k)=\begin{bmatrix} 0^{1\times 2} & \left(d_h(t_0)-Q^T(t_k)\begin{bmatrix}1 & 0 & 0\end{bmatrix}^T\right)\end{bmatrix}^T$$

$$^{\Psi}\dot{x}_r(t_k)=\begin{bmatrix} 0^{1\times 2} & \left(Q^T(t_k)\begin{bmatrix}0 & 1 & 0\end{bmatrix}^T\right)\end{bmatrix}^T$$

In (14), Q is a vector, defining the relative distance from the hand starting position and its first and second derivatives.

The actual movement of the patient's hand is compared to the optimal path after which the supporting force/torque on the hand is determined based on the error between the reference and the actual distance to the target. The generalize force on the hand $^{\Psi}h(t_k)=^{\Psi}F(t_k)$ may be formed as:

$$^{\Psi}h(t_k)=K_{dx}\left(^{\Psi}\dot{x}_r(t_k)-^{\Psi}\dot{x}(t_k)\right)-\begin{cases} K_x\left(^{\Psi}x_r(t_k)-^{\Psi}x(t_k)\right) & if\ ^{\Psi}x(t_k)>^{\Psi}x_r(t_k) \\ 0 & if\ ^{\Psi}x(t_k)\le ^{\Psi}x_r(t_k) \end{cases} \qquad (15)$$

where $K_x(t_k)$ and $K_{dx}(t_k)$ are the controller stiffness and damping diagonal matrices respectively. The robot support is only active when the hand distance to the target is greater than optimal distance to the target. The force on the hand expressed in the reference coordinate frame is then:

$$h(t_k)=\Psi(t_k)^{\Psi}h(t_k) \qquad (16)$$

The supporting force provided by the robot is adaptive in a sense that adaptively adjusting the impedance matrix $K_x(t_k)$ in a feedback loop based on the error

between the reference and the actual distances to the target adjusts support force according to patient's performance. A simple proportional controller K_0 is used to adjust $K_x(t_k)$ when the error is larger than ε, whereas a decrease of the controller stiffness is introduced via parameter η to reduce active support when the patient performs well:

$$K_x(t_k) = (1-\eta)K_x(t_{k-1}) + \begin{cases} K_0 diag\left(\left|{}^{\Psi}x_r(t_k) - {}^{\Psi}x(t_k)\right|\right) & if\ {}^{\Psi}x(t_k) - {}^{\Psi}x_r(t_k) < \varepsilon \\ 0 & if\ {}^{\Psi}x(t_k) - {}^{\Psi}x_r(t_k) \le \varepsilon \end{cases} \quad (17)$$

The adaptive nature of robot-supported reaching device may be extended if we adaptively change the desired time as well. If the patient performs well, currently selected completion time may become longer as necessary which may decrease patient's motivation. For this reason and to gradually stimulate the patient to increase voluntary effort as well, desired completion time is computed as the average of the completion times in the last *n* successfully completed reaching attempts:

$$t_f(t_k) = \frac{t_f(t_{k-1}) + t_f(t_{k-2}) + \ldots + t_f(t_{k-n})}{n} \quad (18)$$

In this way the controller completely meets a particular patient's performance capabilities.

5.3 Reinforcement learning

5.3.1 Method background description

The reinforcement methods are based on the algorithms described in (Sutton & Barto, 1998). Algorithms for both discrete and continuous state space have been implemented. The algorithms use the concept of eligibility traces which allows the user of the method to choose parametrically between Monte Carlo and temporal difference methods, thus providing solutions that can optimally learn a variety of tasks. The implementation of the algorithms is efficient in terms both of computational cost and data transmission. Due to the small size of data that needs to be transmitted in and out of the RL module, it can be based on a remote computer.

5.3.2 Experimental verification

As described in previous reports an experiment has been performed to demonstrate (Kastanis & Slater, 2010) the validity of the application of reinforcement learning to problems where a state of a human is to be learned. This proof of concept experiment proved that it is possible to manipulate the state of a human in an Immersive Virtual Environment (IVE) using the response of people to interpersonal distances (proxemics).

A collaborative experiment has been designed by UB and UL to use the upper extremity system to perform a pick and place task where the RL algorithm controls the difficulty level. The aim is to keep the user in a motivated state by maintaining a high performance as well as a desired skin conductance level or frequency of skin conductance responses (SCRs). After initial pilots were run, it was found that the RL module was capable of controlling the performance of the user with relative ease. It was also capable of controlling skin conductance level (Figure 17). However, when aiming to control both performance and SCR, it was discovered that the design of the system required adjustment. After discussions and investigations on these issues the experiment as well as the methodology for RL are being re-designed.

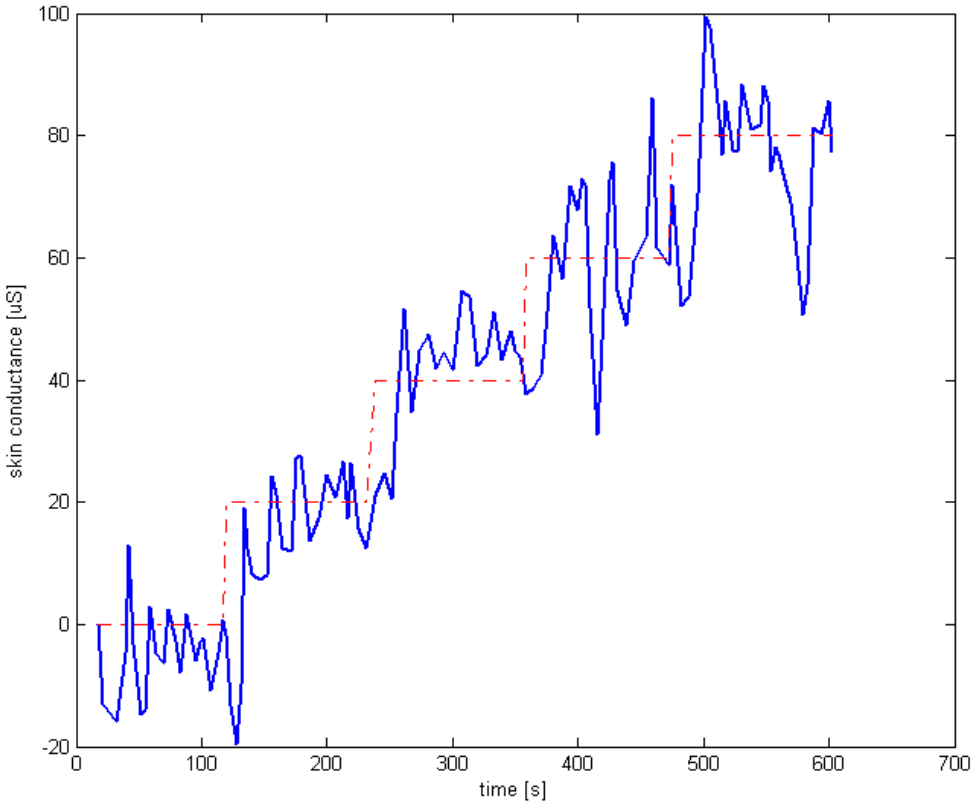

Figure 17: Skin conductance level controlled by reinforcement learning in the ball-catching scenario. The dashed red line represents the reference while the solid blue line represents the actual value.

An experiment examining the application of RL methods in the lower extremity case has been designed by UB and will be run in collaboration with partners in Bad Aibling and Zurich. The experiment will examine the motivation of a user in an navigation task within a city environement. Motivation will expressed as the amount of effort provided by the user of the Lokomat system. The aim of the RL algorithm will be exactly this one, to increase the applied effort during the run of the experiment.

5.3.3 Conclusions

The majority of RL applications aims to optimise directly states of the RL agent. We have reformulated the traditional RL separation between environment and agent in

order to control the state of the human participant. This concept has been investigated and proven to be valid. Further experiments have been planned to apply these ideas in to both upper and lower extremity rehabilitation scenarios. After the initial attempts in the upper extremity pick and place scenario, new methodological ideas have arised to overcome issues typically found in RL, such as long learning times and adaptation to individuals. Experimental evaluation of these ideas is expected to investigate their validity, as well as their applicability in the problem of neurorehabilitation.

5.4 Adaptive discriminant analysis

5.4.1 Method background description

Discriminant analysis is a popular statistical method for the **classification of multidimensional data into two or more distinct classes**. Based on a training data set with annotated class labels, one or more discriminant functions are generated that optimally classify the training data. The discriminant functions can be linear (linear discriminant analysis – LDA) or quadratic (quadratic discriminant analysis – QDA). Recently, variants have been developed where the **discriminant functions adapt to a particular user as new data become available** (Vidaurre, Schlögl, Cabeza, Scherer, & Pfurtscheller, 2007), and these form the basis of the adaptive discriminant analysis system.

5.4.1.1 Linear discriminant analysis

The equation for the LDA classification of data of two different classes can be written as:

$$D(x) = b + w^T \cdot x \quad (19)$$

$$b = -w^T \cdot \frac{1}{2} \cdot (\mu_1 + \mu_2) \quad (20)$$

$$w = (\Sigma_1 + \Sigma_2)^{-1} \cdot (\mu_2 - \mu_1) \quad (21)$$

where Σ_i is the covariance matrix for class *i*, μ_i is the mean value for class *i*, and x is the feature vector. If $D(x)$ is greater than zero, the input is classified as class 2, and if the output is equal to or less than zero, the input is classified as class 1.

5.4.1.2 Quadratic discriminant analysis

Quadratic discriminant analysis (QDA), is in essence, a more general version of LDA. As the name implies, QDA separates data points into classes using quadric combinations of features (unlike LDA, which uses a linear combination). LDA can be derived from QDA by assuming that the covariance of each of the classes is identical.

QDA can be computed using the Mahalanobis distance (d_i) of the feature vector $\boldsymbol{x}$ to each class *i*:

$$d_i(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu}_i)^T \cdot \boldsymbol{\Sigma}_i^{-1} \cdot (\boldsymbol{x} - \boldsymbol{\mu}_i) \tag{22}$$

$$D(\boldsymbol{x}) = \sqrt{d_1(\boldsymbol{x})} - \sqrt{d_2(\boldsymbol{x})} \tag{23}$$

where $\boldsymbol{\Sigma}_i$is the covariance matrix for class *i* and $\boldsymbol{\mu}_i$ is the mean value for class *i*. If $D(\boldsymbol{x})$ is greater than zero, the input is classified as class 2, and if the output is equal to or less than zero, the input is classified as class 1. While other forms of QDA calculation are possible, this form is important since it is the one used as the basis for the adaptive information matrix (described later in this section).

5.4.1.3 Kalman adaptive linear discriminant analysis

Kalman adaptive linear discriminant analysis (KALDA) is an adaptive version of the LDA in which the weighs of the discriminant function $(b, \boldsymbol{w}^T)$ are recursively estimated online using a Kalman filter as new data becomes available. The Kalman gain varies the update coefficient and changes the adaptation speed depending on the properties of the data. The update equations can be summarized as follows (Vidaurre, Schlögl, Cabeza, Scherer, & Pfurtscheller, 2007).

$$\boldsymbol{H}_k = [1, \boldsymbol{x}_k^T] \tag{24}$$

$$e_k = y_k \quad \boldsymbol{H}_k \cdot \hat{\boldsymbol{w}}_{k-1} \tag{25}$$

$$v_k = 1 - UC \tag{26}$$

$$Q_k = \boldsymbol{H}_k \cdot \boldsymbol{A}_{k-1} \cdot \boldsymbol{H}_k^T + v_k \tag{27}$$

$$\boldsymbol{k}_k = \frac{\boldsymbol{A}_{k-1} \cdot \boldsymbol{H}_k^T}{Q_k} \tag{28}$$

$$\hat{\boldsymbol{w}}_k = \hat{\boldsymbol{w}}_{k-1} + \boldsymbol{k}_k \cdot e_k \tag{29}$$

$$\tilde{\boldsymbol{A}}_k = \boldsymbol{A}_{k-1} - \boldsymbol{k}_k \cdot \boldsymbol{H}_k \cdot \boldsymbol{A}_{k-1} \tag{30}$$

$$\boldsymbol{A}_k = \frac{\mathrm{trace}(\tilde{\boldsymbol{A}}_k) \cdot UC}{p} + \tilde{\boldsymbol{A}}_k \tag{31}$$

where e_k is the one-step prediction error, y_k is the current class label, $\boldsymbol{x}_k$ is the current feature vector, $\hat{\boldsymbol{w}}_k$ is the state vector ($\hat{\boldsymbol{w}}_k$ = $[b, \boldsymbol{w}^T]$, the estimated weights for the LDA), Q_k is the estimated prediction variance, $\boldsymbol{A}_k$ is the a priori state error correlation matrix, $\tilde{\boldsymbol{A}}_k$ is an intermediate value needed to compute $\boldsymbol{A}_k$, v_k is the variance of the innovation process, $\boldsymbol{k}_k$ is the Kalman gain, UC is the update coefficient and p is the number of elements of $\hat{\boldsymbol{w}}_k$. The starting values $\boldsymbol{A}_0$ and $\hat{\boldsymbol{w}}_0$ are computed from the training data set.

While it would be possible to simply add the new data to the training set and recalculate the weights of the discriminant function, this is often impractical for online use. Since the training set often consists of dozens or hundreds of measurements, one new measurement is unlikely to strongly affect the weights of the discriminant function. Kalman filtering, on the other hand, can adapt faster since it is a recursive approach and is thus more suitable for online use. Rapid adaptation is especially crucial for psychophysiological responses. Since these vary greatly from subject to

subject, data obtained online must be integrated into the discriminant analysis quickly so that the classifier can better adapt to the subject currently being measured.

5.4.1.4 Adaptive information matrix

The adaptive information matrix (ADIM) is an adaptive version of the QDA in which the estimate of the inverse of the covariance matrix of each class is recursively updated. Thus, the two QDA equations are expanded with the following recursive equation (Vidaurre, Schlögl, Cabeza, Scherer, & Pfurtscheller, 2007):

$$\sum_{k,i}^{-1} = (1 + UC) \cdot \sum_{k-1,i}^{-1} - \frac{1+UC}{1-UC \cdot v \cdot x_{k,i}} \cdot v \cdot v^T \quad (32)$$

where $v = \sum_{k-1,i}^{-1} \cdot x_{k,i}$, UC is the update coefficient for the adaptation, $x_{k,i}$ is the current class *i* feature vector, *k* is the current sample and *i* is the class label. The mean vector μ_i is also needed for the computation of $D(x)$ and needs to be estimated. This mean vector was incorporated as additional row and column data to the matrix $\sum_i$ for its automatic estimation and to avoid an extra algorithm, resulting in an "extended" covariance matrix. The starting matrix $\sum_0$ was computed from the training data set.

5.4.1.5 Classifier fusion

While it is possible to simply use adaptive discriminant analysis to build a single classifier using all the input variables at once, another possibility would be to build a classifier for each input variable and then perform classifier fusion. While the accuracy of any single classifier would be low, fusing the large number of classifiers may result in much higher accuracy. This approach is similar to naïve Bayes classification, where all inputs are treated as independent.

Fusion of the multiple univariate classifiers is performed using a weighted majority vote where contribution of each classifier is weighted by its estimated accuracy. This accuracy can be estimated from the training data. For instance, a classifier that was able to correctly classify 100% of the training data would be weighted with a factor of 1, a classifier that was able to correctly classify 75% of the training data would be weighted with a factor of 0.5, and a classifier that was able to correctly classify 50% of the training data (and thus performed no better than a random guess would be weighted with a factor of 0. Classifier fusion is then done using the following formula:

$$D(x) = \sum_i k_i \cdot C_i \quad (33)$$

where x is the feature vector, C_i is the class assigned to the feature vector by classifier *i*, and k_i is the weighting factor of classifier *i*. If C_i can either be -1 (class 1) or +1 (class 2), the feature vector is assigned to class 1 if $D(x)$ is equal to or less than zero and assigned to class 2 if $D(x)$ is greater than zero.

5.4.2 Experimental verification

Adaptive discriminant analysis was implemented in the ball catching task developed in WP2 in order to **estimate and adjust task difficulty based on a combination of task performance, motor actions and psychophysiology**.

5.4.2.1 Ball-catching task

The ball-catching task was developed in WP2 and is described in detail in D2.2. Although the original version of the ball catching task offers active haptic support, none was provided for verification of adaptive discriminant analysis. Seven different difficulty levels were implemented specifically for adaptive discriminant analysis, with each higher difficulty level featuring smaller and faster balls so that they were harder to catch.

5.4.2.2 Subjects

While the experiment is still ongoing, twenty-three healthy subjects (19 male, 4 female, mean age 31.1 years, st. dev. 10.9 years) and four patients (2 male, 2 female, ages 32, 35, 44 and 45) have participated so far.

5.4.2.3 Experiment procedure

Upon arrival, the task was demonstrated to the subject and the procedure was described. Then, the physiological measurement equipment was attached and turned on. The subject first rested for two minutes, then performed the task for six two-minute periods (12 minutes total). Within each period, the subject tried to catch several balls at a constant difficulty. At the end of a period, the subject was asked whether he or she would prefer the difficulty of the task to increase or decrease. The difficulty of the task then changed randomly by one or two levels in the selected direction. Thus, subjects were able to control whether the difficulty will increase or decrease, but not by how much. Subjects were not given the option to stay at the same difficulty level, as we wanted only two different possibilities for simpler classification.

The randomness in the difficulty change was introduced so that psychophysiological responses from a wider range of difficulty levels. In pretesting, subjects tended to stay within a narrow range of difficulty levels (usually 3-5), if there was no random element in the change of difficulty.

After the experiment had been completed and the subject had left, the collected psychophysiological, haptic and performance data were processed and several relevant variables were extracted for each time period (including baseline). These data were then used to develop a classifier that would predict whether a subject would choose to increase or decrease task difficulty at a particular moment based only on the information collected up to that moment. For healthy subjects, the classifiers were tested using leave-one-out cross-validation. Each classifier was built using training data from all but one subject, then tested on the final subject. This was

repeated as many times as there were subjects (with each subject left out of the training data exactly once) in order to obtain the overall classification accuracy rate. The classification accuracy rate was defined as the number of correctly classified measurements divided by the total number of measurements. For patients, the classifiers were trained using data from all healthy subjects then tested directly on the patient data.

5.4.2.4 Performance variables

The variables used for classification were divided into three groups: performance variables, psychophysiological variables, and haptic variables. Performance variables described how well a subject did during a particular time period and how long he had been performing the task. The variables were:

- difficulty level,
- number of time period (1 – first time period, 6 – last time period),
- percentage of balls successfully caught,
- percentage of caught balls placed into the basket,
- percentage of balls successfully both caught and placed into the basket.

5.4.2.5 Psychophysiological variables

The psychophysiological variables were extracted from the electrocardiogram, skin conductance, respiration and skin temperature signals. They were:

- mean heart rate,
- shortest NN interval (interval between two successive normal R-peaks in the ECG),
- longest NN interval,
- difference between the shortest and longest NN interval,
- standard deviation of NN intervals (SDNN),
- square root of the mean square differences of successive NN intervals (RMSSD),
- the percentage of interval differences of successive NN intervals greater than 50 ms (pNN50),
- total power in the high-frequency (HF – 0.15 Hz to 0.4 Hz) band of the heart rate signal,
- total power in the low-frequency (LF – 0.04 Hz to 0.15 Hz) band of the heart rate signal,
- the ratio of the two above variables (LF/HF ratio),
- mean skin conductance level,
- skin conductance level at the end of the period,
- difference between the skin conductance level at the beginning of the period and the end of the period,
- skin conductance response frequency,
- mean skin conductance response amplitude,
- standard deviation of skin conductance response amplitude,
- mean respiratory rate,
- standard deviation of respiratory rate,
- mean temperature,
- temperature at the end of the period.

Due to large differences in baseline values between subjects, actual values of these variables were not used for classification. The variables were calculated during the two-minute initial rest period, and then the values for the rest period were subtracted from the values for each task period. These 'relative' values were then used as classifier inputs.

5.4.2.6 Haptic variables

Haptic variables described a subject's motor actions and were extracted from the HapticMaster's force and movement sensors. The eight variables used were:

- mean absolute force,
- mean absolute velocity,
- mean absolute acceleration,
- total work,
- mean frequency of the position signal,
- mean frequency of the velocity signal,
- mean frequency of the acceleration signal,
- mean frequency of the force signal.

5.4.2.7 Our classification procedure

While the classification methods have already been described, it is worth restating here exactly what classifiers were tested. First, we tested the possibility of single classifiers with no classifier fusion:

- LDA, QDA, KALDA and ADIM with all performance variables as inputs,
- LDA, QDA, KALDA and ADIM with all psychophysiological variables as inputs,
- LDA, QDA, KALDA and ADIM with all haptic variables as inputs,
- LDA, QDA, KALDA and ADIM with all variables as inputs.

Then, we tested the possibility of building a separate classifier for each input variable and fusing these individual classifiers using a weighted vote into a single classifier. As previously mentioned, while the accuracy of any single classifier would be low, we expected that fusing the large number of classifiers would result in much higher accuracy. Once again, all four methods of discriminant analysis were tested:

- LDA, QDA, KALDA and ADIM on each performance variable, then weighted vote,
- LDA, QDA, KALDA and ADIM on each psychophysiological variable, then weighted vote,
- LDA, QDA, KALDA and ADIM on each haptic variable, then weighted vote,
- LDA, QDA, KALDA and ADIM on each variable, then weighted vote,

To clarify, all LDA classifiers were fused with each other, all QDA classifiers were fused with each other and so on. There was no fusion of, for example, LDA and QDA classifiers together.

5.4.2.8 Results – healthy subjects

Table 6 shows the accuracy rates (percentage of correctly classified samples) for multivariate classifiers (no fusion) with different input data. Table 7 shows the accuracy rates for weighted-vote fusion of univariate classifiers (as described in section 5.4.1.5) with different input data. As previously mentioned, LDA and QDA are nonadaptive classification methods while KALDA and ADIM are adaptive methods.

Table 6: Accuracy rates for different discriminant analysis methods and different types of input data, using a single multivariate classifier for each method.

	performance	haptics	psychophysiology	performance + haptics	all
LDA	82.8 %	74.5 %	54.5 %	80.0 %	78.6 %
QDA	82.1 %	76.6 %	54.5 %	76.6 %	68.3 %
KALDA	82.8 %	75.2 %	71.0 %	78.6 %	75.2 %
ADIM	82.1 %	76.6 %	64.8 %	79.3 %	68.3 %

Table 7: Accuracy rates for different discriminant analysis methods and different types of input data, using weighted vote fusion of univariate classifiers.

	performance	haptics	psychophysiology	performance + haptics	all
weighted LDA	82.1 %	71.7 %	62.1 %	79.3 %	81.4 %
weighted QDA	81.4 %	70.3 %	60.7 %	77.2 %	80.0 %
weighted KALDA	**82.8 %**	**80.0 %**	**77.9 %**	**82.8 %**	**84.8 %**
weighted ADIM	81.4 %	70.3 %	66.2 %	77.2 %	77.2 %

From Tables 6 and 7, it is evident that **nonadaptive classifiers (LDA and QDA) based on performance or haptics perform better than those based only on psychophysiology**. This is the case for both multivariate classification and fusion of univariate classifiers. Interestingly, in the case of multivariate classifiers (Table 6), classification using multiple types of data has a lower accuracy rate than classification using only one type of data. This may be because the data dimensionality is large and the available data set is small, so it is difficult to find a single robust discriminant function. Weighted vote fusion of univariate classifiers should be more robust since each individual discriminant function covers only one dimension.

For both haptic and psychophysiological data, **adaptive methods perform better than nonadaptive ones** (haptics: best nonadaptive result is 76.6 % while best adaptive result is 80.0%; psychophysiology: best nonadaptive result is 62.1 % while best adaptive result is 77.9 %). The advantage of adaptive methods is especially notable for psychophysiological data.

Naturally, perfect classification accuracy should not be expected. Several subjects occasionally expressed the desire to stay at the same difficulty level, but this option was not available. Additionally, it is not certain whether subjective choices are always perfectly reflected in physiological responses and thus whether completely accurate classification is even theoretically possible.

5.4.2.9 Results – patients

Though patient measurements are still ongoing and a full analysis of the results can not yet be given, **initial results from four patients are encouraging**. When using adaptive methods, performance yields an accuracy rate of 87.5 %, haptics yield an accuracy rate of 66.7%, and psychophysiology yields an accuracy rate of 75.0 %. Using all three types of data increases the classification rate to 88.9 %.

5.4.2.10 Lokomat implementation

In late June 2010, researchers from UL and ETH jointly implemented the adaptive discriminant analysis algorithms in the Lokomat platform. In the Lokomat version, adaptive discriminant analysis has been expanded to include three user states (underchallenged/suitably challenged/overchallenged). Initial tests have been promising, and a full experiment is planned for the second half of 2010.

5.4.3 Conclusions

Adaptive discriminant analysis is **able to determine whether a subject finds a task too easy or too hard** with a high degree of accuracy. This method can be **directly used online in task difficulty adaptation**. If the task is too hard, the system should decrease the difficulty. If the task is too easy, the system should increase the difficulty. Methods of increasing or decreasing difficulty must be predefined in advance and can be very simple (e. g. increase the speed of the ball) or more complex (e. g. increase the artificial intelligence of the computer opponents in a complex virtual scenario). Initial tests indicate that adaptive discriminant analysis can also differentiate between multiple user states (underchallenged/suitably challenged/overchallenged).

Additionally, this method seems to be **suitable for psychophysiological studies on patients**, as preliminary patient results have shown that classification accuracy based on psychophysiology is not much lower in patients (75.0 %) than in healthy subjects (77.9 %). The lower classification accuracy based on haptics (66.7 % for patients vs. 80.0 % for healthy subjects) is easy to explain. When presented with very fast balls at high difficulties, all healthy subjects tried to move very fast around the entire table to catch the balls. Three out of the four patients, on the other hand, became frustrated and mainly waited in one area of the table (to catch only those balls that moved through that area), resulting in completely different movement patterns for patients and healthy subjects. Of course, a larger sample size will be required for proper evaluation, and this is scheduled to be completed in WP4.

An important finding was that **combining psychophysiological data with motor actions and task performance does not yield a much higher classification accuracy than simply using task performance**. This raises the question of whether psychophysiological measurements are truly useful in motor rehabilitation, especially in rehabilitation robotics where quantitative measurements of performance and motor ability are available.

6 Conclusions

Within workpackage 3, we developed and tested a number of procedures and algorithms required to acquire, process and interpret multi-sensorial information arising from the interaction of the human with the robot and the virtual environment. In D3.1, the groundwork was laid by creating the algorithms required to sense user motor actions and psychophysiological responses. In this deliverable, we continued our work by fusing both motor actions and psychophysiological responses into a unified estimate of user state through different approaches: fuzzy logic, mathematical models, neural networks and statistical methods.

Furthermore, we developed several automatic learning systems that act on the user's state and learn from it. The adaptive haptic assistance developed for the HapticMaster detects the user's intentions and supports his or her motor actions. Reinforcement learning is a rapidly learning method that can control both the user's performance and psychophysiological state if the number of input variables is relatively low. If many different inputs are required, adaptive discriminant analysis can be used to both learn from previously recorded data and adapt to the current user in real time. Reinforcement learning and adaptive discriminant analysis are applicable to both the upper and lower extremities.

7 Bibliography

Achten, J., & Jeukendrup, A. E. (2003). Heart rate monitoring: applications and limitations. *Sports Medicine, 7*, 517-538.

Arts, F. J., & Kuipers, H. (1994). The relation between power output, oxygen uptake and heart rate in male athletes. *International Journal of Sports Medicine, 15*, 228-231.

Baum, K., Essfeld, D., Leyk, D., & Stegemann, J. (1992). Blood pressure and heart rate during rest-exercise and exercise-rest transitions. *European Journal of Applied Physiology and Occupational Physiology, 64*, 134-138.

Bernardi, L., Valle, F., Coco, M., Calciati, A., & Sleight, P. (1996). Physical activity influences heart rate variability and very-low-frequency components in Holter electrocardiograms. *Cardiovascular Research, 32*, 234-237.

Borg, G., Hassmen, P., & Lagerström, M. (1987). Perceived exertion related to heart rate and blood lactate during arm and leg exercise. *European Journal of Applied Physiology and Occupational Physiology, 56*, 679-685.

Carter, J. B., Banister, E. W., & Blaber, A. P. (2003). Effect of endurance exercise on autonomic control of heart rate. *Sports Medicine, 33*, 33-46.

Feroldi, P., Belleri, M., Ferretti, G., & Veicsteinas, A. (1992). Heart rate overshoot at the beginning of muscle exercise. *European Journal of Applied Physiology and Occupational Physiology, 65*, 8-12.

Flash, T., & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *Journal of Neuroscience, 5*, 1688-1703.

Maclean, N., Pound, P., Wulfe, C., & Rudd, A. (2002). The concept of patient motivation: A qualitative analysis of stroke professionals' attitudes. *Stroke, 33*, 444-448.

Mandryk, R. L., & Atkins, S. M. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies, 65*, 329-347.

Mihelj, M., Nef, T., & Riener, R. (2007). A novel paradigm for patient-cooperative control of upper-limb rehabilitation robots. *Advanced Robotics, 21*, 843-867.

Robertson, I., & Murre, J. (1999). Rehabilitation of brain damage: brain plasticity and principles of guided recovery. *Psychological Bulletin, 125*, 544-575.

Rosenblatt, F. (1962). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms.* Spartan.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161-1178.

Schwerdtfeger, A. (2004). Predicting autonomic reactivity to public speaking: don't get fixed on self-report data! *International Journal of Psychophysiology, 52*, 217-224.

Sutton, S., & Barto, A. (1998). *Reinforcement Learning.* MIT Press.

Vidaurre, C., Schlögl, A., Cabeza, R., Scherer, R., & Pfurtscheller, G. (2007). Study of on-line adaptive discriminant analysis for EEG-based brain computer interfaces. *IEEE Transactions on Biomedical Engineering, 54*, 550-556.

Vokac, Z., Bell, H., Bautz-Holter, E., & Rodahl, K. (1975). Oxygen uptake/heart rate relationship in leg and arm exercise, sitting and standing. *Journal of Applied Physiology, 39*, 54-59.

Publications by members of the MIMICS consortium:

Kastanis, I., & Slater, M. (2010). Reinforcement learning utilizes proxemics: An avatar learns to manipulate the position of people in immersive virtual reality. *Submitted for publication*.

Novak, D., Ziherl, J., Olenšek, A., Milavec, M., Podobnik, J., Mihelj, M., et al. (2010). Psychophysiological responses to robotic rehabilitation tasks in stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, in press.