Grant Agreement No. 611373


FP7-ICT-2013-10

# D3.4 TDNN development and performance for diver behavior interpretation

Due date of deliverable: 31/09/2016
Actual submission date: 27/09/2016

Start date of project: 01 January 2014 Duration: 36 months

Organization name of lead contractor for this deliverable: UNIVIE

Revision (version 3)

| Dissemination level | | |
|---|---|---|
| PU | Public | x |
| PP | Restricted to other programme participants (including the Commission Services) | |
| RE | Restricted to a group specified by the consortium (including the Commission Services) | |
| CO | Confidential, only for members of the consortium (including the Commission Services) | |

# Contents

## 1. OUTLINE OF THE DELIVERABLE

This deliverable documents the development of neuronal network algorithms for diver behavior interpretation. Physiological data – breath rate and heart rate – and motion rate data were combined to train algorithms for the prediction of the diver's internal state in terms of high/middle/low levels of pleasure, arousal and control. In addition, the motion data is used for accurate pose estimation. The resulting algorithms are integrated in the DiverControl Center for real time diver behavior supervision.

## 2. BREATHING THROUGH A REGULATOR

**Aim:** Comparing breathing patterns with and without a regulator

**Experiment:** We collected data from 15 participants – 7 women (mean age: 25.0yrs) and 8 men (mean age: 29.1yrs) – with a breathing belt while they breathed regularly in an upright position with and without the regulator for 3 minutes each. Participants watched a neutral stimulus in both conditions.

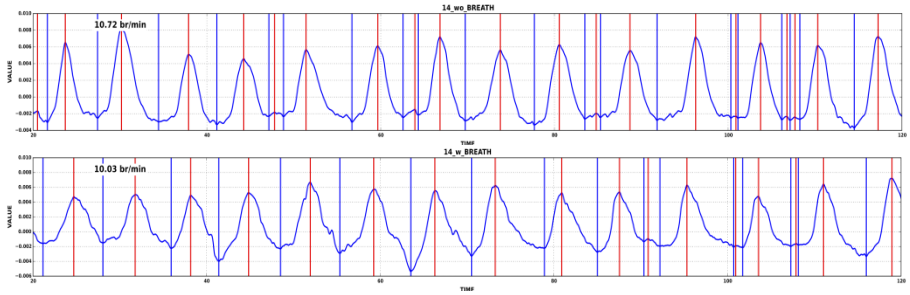


*Figure 2-1 Breathing patterns without and with a regulator*

**Results:** Breathing with a regulator (WR) and free breathing (WOR) are very similar. The form of the breathing curves is unaffected, as are speed of inhalation and exhalation, turbulence and amplitude. However, the rate of breathing was significantly lower when breathing through a regulator (Mean breath rate WR: 11,30 breaths per minute; WOR: 14,48 breaths per minute; paired t-test: t = -3,809; p = 0,003). In addition, breath rate varies between individuals: breath rate WR correlates with breath rate WOR: r = 0,617, p = 0,043).

Breathing through a regulator affects heart rate: It changes more quickly (WR: 3,27; WOR: 2,61, paired t-test: t = 3,359; p = 0,00), and we find that turbulence is lower when breathing through a regulator (WR: 0,546; WOR: 0,815; paired t-test: t =-2,896; p = 0,016). On a physiological level this might suggest that the organism reaches an upper threshold of possible performance when breathing through a regulator. In normal circumstances, physiological variables such as heart rate tend to be irregular as a result of constant adaption to changing internal and external stimuli. Breathing and heart rate are a chaotic dynamical system - when the physiological system is under pressure it reaches its limits to adapt to external circumstances and becomes more regular. We also calculated the correlation between breath rate and heart rate using Granger probabilities: Changes in breathing seem to trigger changes in heart rate. We will use this measure to predict the diver's

physiological stress.

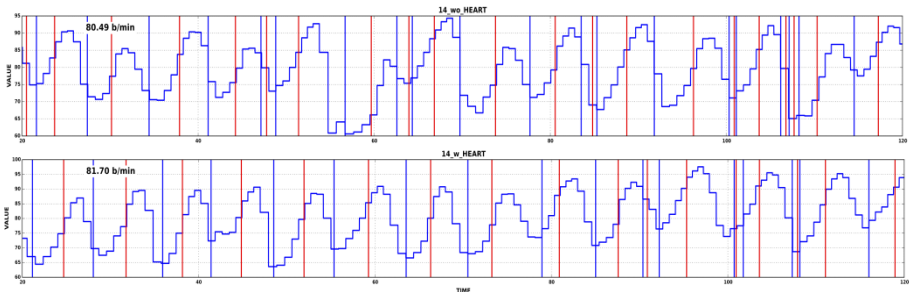


*Figure 2-2 Heart rate patterns without and with a regulator*

## 2.1. DYNAMIC FEATURES OF TIME SERIES

The data from this experiment were also used to develop dynamic descriptors of time series for later statistical analysis In order to extract relevant information from physiological and motion time series data, we need parameters that reflect the dynamic nature of the data. Grammer et al (1999)[1], Grammer et al (2003)[2] and Koppensteiner and Grammer (2010)[3] developed several measure that allow the description of such dynamic features. In the following segment these parameters will be defined.

Peak detection: We used a peak detection algorithm developed by Billauer (ttp://billauer.co.il/peakdet.html) and converted to python by (https://gist.github.com/endolith/250860).

Rate: Both heart rate and breath rate form the basis for the calculation of further parameters. The rate is calculated in breaths per minute and heartbeats per minute respectively by dividing 60 by the difference between two peaks measured in seconds.

The following parameters are computed for overlapping subsets of the entire time series spanning 35 seconds, yielding a continuous

```
def peak_detection_all(self,v, x, window, delta):
    maxtab = []
    mintab = []

    if x is None:
        x = arange(len(v))
    v = np.asarray(v)
    if len(v) != len(x):
        sys.exit('Input vectors v and x must have same length')
    if not np.isscalar(delta):
        sys.exit('Input argument delta must be a scalar')
    if delta <= 0:
        sys.exit('Input argument delta must be positive')

    mn, mx = np.Inf, -np.Inf
    mnpos, mxpos = np.NaN, np.NaN

    lookformax = True
    for i in arange(len(v)):
        this = v[i]
        if this > mx:
            mx = this
            mxpos = x[i]
        if this < mn:
            mn = this
            mnpos = x[i]
        if lookformax:
            if this < mx-delta:
                maxtab.append((mxpos, mx))
                mn = this
                mnpos = x[i]
                lookformax = False
        else:
            if this > mn+delta:
                mintab.append((mnpos, mn))
                mx = this
                mxpos = x[i]
                lookformax = True
    return [maxtab,mintab]
```

*Figure 1-3 peak detection algorithm*

[1] Grammer, K., Honda, M., Juette, A., & Schmitt, A. (1999). Fuzzi
motion energy detection. *Journal of personality and social psychol*

[2] Grammer, K., Keki, V., Striebel, B., Atzmüller, M., & Fink, B. (2003
*aesthetics* (pp. 295-323). Springer Berlin Heidelberg.

[3] Koppensteiner, M., & Grammer, K. (2010). Motion patterns in p
*Journal of Research in Personality*, *44*(3), 374-379.

parameter for the entire duration of the complete time series.

Amplitude: We defined measure of amplitude both for heart rate and breath rate as well as breathing pattern.

If the time series subset starts with a maximum, amplitude is calculated as follows,

$$amplitude = \frac{1}{2n}\sum_{i=0}^{n}\big(value\ of\ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum\big) + (value\ of\ (i+1)^{th}\ maximum - value\ of\ \ i^{th}\ minimum)$$

If the time series subset starts with a minimum, the calculation changes to:

$$amplitude = \frac{1}{2n}\sum_{i=0}^{n}\big(value\ of\ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum\big) + (value\ of\ \ i^{th}\ maximum - value\ of\ \ (i+1)^{th}\ minimum)$$

where

$$n = (Min(number\ of\ maxima, number\ of\ minima) - 1).$$

Amplitude on/off:

The amplitude of the inhalation or the increase in heart rate and breath rate respectively is defined as follows:
If the first maximum in the time series subset precedes all minima, the calculation goes as follows:

$$amplitude_{on} = \frac{1}{n}\sum_{i=0}^{n}\big(value\ of\ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum\big)$$

If the first minimum precedes all maxima, this changes to:

$$amplitude_{on} = \frac{1}{n}\sum_{i=0}^{n}\big(value\ of\ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum\big)$$

The variable amplitude_off describes the amplitude of exhalation and a decrease in breath rate or heart rate respectively.

If the time series subset's first maximum precedes the first minimum, it is calculated as follows:

$$amplitude_{off} = \frac{1}{n}\sum_{i=0}^{n}\big(value\ of\ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum\big),$$

if the minimum comes first this changes to:

$$amplitude_{off} = \frac{1}{n}\sum_{i=0}^{n}\left(value\ of\ i^{th}\ maximum - value\ of\ \ (i+1)^{th}\ \ minimum\right),$$

For all calculations of amplitude,

$$n = (Min(number\ of\ maxima, number\ of\ minima) - 1).$$

Curve length:

The curve length is calculated as follows

$$Curve\ length = \frac{\sum_{i=1}^{n}\sqrt[2]{1 + \ (data[i]) - data[i-1])^2}}{n}$$

$$Curve\ length = \frac{\sum\sqrt[2]{1 + \ (Difference\ between\ two\ consecutive\ data\ points)^2}}{Total\ number\ of\ data\ points}$$

Speed:

The variable speed is used to describe the rate of change for breathing patterns as well as breath rate and heart rate.

If the time series subset begins with a maximum, the speed is calculated as follows:

$$speed = \frac{1}{2n} * \sum_{i=0}^{n}\left[\frac{value\ of\ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum}{time\ difference\ between\ \ i^{th}\ maximum\ and\ \ i^{th}\ minimum} + \frac{value\ of\ (i+1)^{th}\ maximum - value\ of\ \ i^{th}\ minimum}{time\ difference\ between\ (i+1)^{th}\ maximum\ and\ \ i^{th}\ minimum}\right]$$

If the subsets begins with a minimum, the calculation changes to:

$$speed = \frac{1}{2n} * \sum_{i=0}^{n}\left[\frac{value\ of\ \ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum}{time\ difference\ between\ \ i^{th}\ maximum\ and\ \ i^{th}\ minimum} + \frac{value\ of\ \ i^{th}\ maximum - value\ of\ (i+1)^{th}\ minimum}{time\ difference\ between\ \ i^{th}\ maximum\ and\ (i+1)^{th}\ minimum}\right]$$

For both cases *n* is defined as follows:

$$n = (Min(number\ of\ maxima, number\ of\ minima) - 1)$$

Speed on/off: The speed of inhalation and exhalation may differ and similarly the changes in heart rate may be of unequal speed depending on whether the heart rate is decreasing or increasing. Because of this we computed the speed of change in these parameters for inhalation/exhalation and decreasing/increasing heart rate respectively.

Speed on describes the speed of inhalation or increasing heart rate or breath rate. It is computed as follows:

If the timeseries starts with a maximum,

$$speed_{on} = \frac{1}{n} * \sum_{i=0}^{n} \frac{value\ of\ (i+1)^{th}\ maximum - value\ of\ \ i^{th}\ minimum}{time\ difference\ between\ (i+1)^{th}\ maximum\ and\ \ i^{th}\ minimum}$$

If the timeseries subset starts with a minimum,

$$speed_{on} = \frac{1}{n} * \sum_{i=0}^{n} \frac{value\ of\ \ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum}{time\ difference\ between\ \ i^{th}\ maximum\ and\ \ i^{th}\ minimum}$$

where $n = Min(number\ of\ maxima, number\ of\ minima) - 1$

Speed off describes the rate of change during exhalation or a decrease in heart rate or breath rate. If time series starts with max, it is computed as follows:

$$speed_{off} = \frac{1}{n} * \sum_{i=0}^{n} \frac{value\ of\ \ i^{th}\ maximum - value\ of\ \ i^{th}\ minimum}{time\ difference\ between\ the\ \ i^{th}\ maximum\ and\ \ i^{th}\ minimum}$$

if time series starts with min, the calculation changes to:

$$speed_{off} = \frac{1}{n} * \sum_{i=0}^{n} \frac{value\ of\ \ i^{th}\ maximum - value\ of\ (i+1)^{th}\ minimum}{time\ difference\ between\ the\ i^{th}\ maximum\ and\ (i+1)^{th}\ minimum}$$

for $n = Min(number\ of\ maxima, number\ of\ minima) - 1$

Turbulence is a measure of irregularity in time series: The larger the standard deviation of the amplitudes, the more irregular the time series. We computed turbulence as follows:

$$Turbulence = \frac{Standard\ Deviation\ of\ Amplitudes}{mean\ Amplitude} = \frac{\sqrt[2]{\frac{1}{n-1} * \sum(Amplitude - mean\ Amplitude)^2}}{mean\ Amplitude}.$$

## 3. EMOTIONS – BREATHING AND HEART RATE

**Aim:** Relating changes in physiological states to primary and secondary emotions

**Experiments:**

Sample: We ran this experiment in 2014 (99 participants, 50 women) and 2015 (91, 51 women) and achieved an overall sample size of 190 participants.

Methodology: Our approach is principally guided by two different emotion theories. Basic or primary emotions are

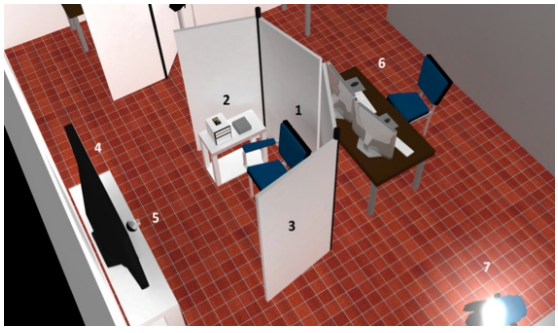


*Figure 0-1 Experimental setup*

short events (4sec) which have discreet corresponding facial muscle contractions. The basic emotions are: Fear, Surprise, Anger, Happy, Sad and Disgust (Ekman and Friesen, 1971[4]). The disadvantage is that they rarely occur as pure emotions. In contrast secondary emotions describe longer lasting emotional states. A circumplex model of emotions (Russel and Mehrabian, 1977)[5] describes an emotional space consisting of three factors: pleasure/valence, arousal and dominance/control. Basic emotions then are subspaces in this model. In this approach we will first verify our emotion induction techniques with basic emotions and then extend the model to the Pleasure-Arousal Dominance model.

Participants were equipped with the breathing belt and heart rate sensor as well as facial electrodes to measure facial muscle activity. At the start of an experimental session, the participant sat down in front of a screen. On this screen we showed them seven movie clips chosen to elicit the emotions anger, anxiety, disgust, happiness, sadness and surprise as well as one neutral clip in a randomized order (a list of the clips is detailed in the reports). After each clip the participant filled in a questionnaire about their emotional state (in the first run, we used a questionnaire based on the basic emotions described by Ekman, in the second run we asked to fill them in an additional Pleasure-Arousal-Dominance/Control questionnaire (PAD) after each clip). From this dataset we derived after elimination of any time series affected by technical difficulties 2100 behaviour segments (heart rate and breathing pattern time series) of 30 second length which can be linked to emotional states.

The facial muscle activity and emotion questionnaires were used to validate the emotion elicitation through movie clips as well as the use of the continuous PAD scale.

**Results:**

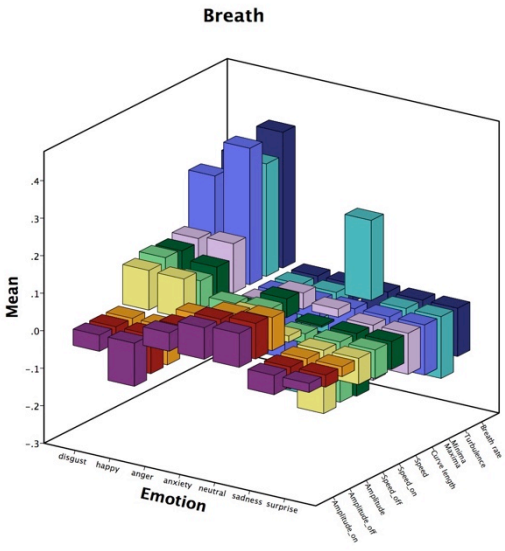


*Figure 2-2 Mean of z-scored breathing parameters for the basic emotions*

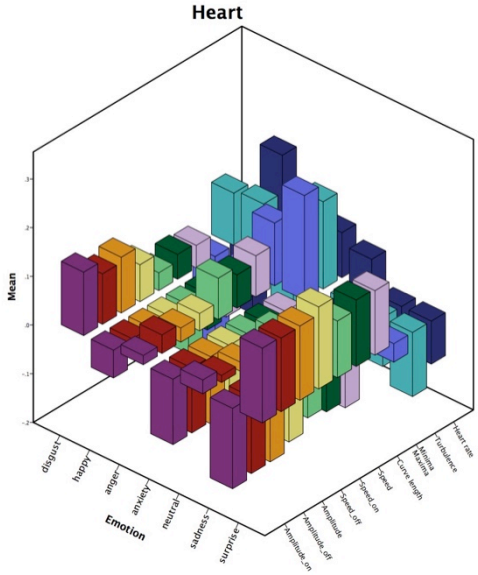


*Figure 2-3 Mean of z-scored heart rate parameters for the basic emotions*

[4] Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, *17*(2), 124. Chicago

[5] Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of research in Personality*, *11*(3), 273-294.

Table 2.1: Discrete primary emotions can be expressed in terms of Pleasure-Arousal-Dominance

| Correlations | Pleasure | Arousal | Dominance/Control |
|---|---|---|---|
| **Happiness** | r = 0.637 (p < 0.001) | r = 0.122 (p = 0.008) | r = 0.132 (p = 0.004) |
| **Anger** | r = -0.557 (p < 0.001) | r = 0.290 (p < 0.001) | r = -0.122 (p = 0.008) |
| **Anxiety** | r = -0.294 (p < 0.001) | r = 0.425 (p < 0.001) | r = -0.160 (p = 0.001) |
| **Sadness** | r = -0.448 (p < 0.001) | r = 0.257 (p < 0.001) | r = -0.032 (p = 0.487) |
| **Disgust** | r = -0.320 (p < 0.001) | r = 0.365 (p < 0.001) | r = -0.060 (p = 0.193) |
| **Surprise** | r = -0.021 (p = 0.653) | r = 0.207 (p < 0.001) | r = -0.091 (p = 0.049) |

For this analysis the original time series from the breathing belt were used and transformed into the basic dynamic features.
We confirmed that the movie clips elicited emotions – the clips for anger, happiness, sadness and surprise caused strong emotions of the same nature, whereas the movie for anxiety caused feelings both of anxiety and of disgust (Table 2.1).

In addition, we demonstrated that the discrete emotions experienced by the participants can be expressed in terms of Pleasure, Arousal and Dominance/Control, a three dimensional continuous model.
Both heart rate parameters and breathing pattern parameters differed significantly between time series of different emotional stimuli (Heart parameters: GLM, Wilk's Lambda = 0.124, F(43,14)=2.297, p = 0.047, ηp=0.162; Breathing pattern parameters: GLM, Wilk's Lambda = 0.110, F(43,14)=2.634, p =0.026, ηp=0.102).

In addition, we conducted linear discriminant analysis for both experimental runs to classify the breathing pattern and heart rate time series according to the stimulus the participant had watched. For all emotions, the highest percentage of classified cases was assorted to the correct category (Table 2.2).

Table 2.2 CLASSIFICATION RESULTS – 1st experimental run | 2nd experimental run

| | Predicted Group Membership: in % | | | | | |
|---|---|---|---|---|---|---|
| | Anger | Anxiety | Disgust | Happy | Sadness | Surprise |
| Anger | 47.4 \| 28.2 | 17.9 \| 20.5 | 7.7 \| 12.8 | 14.4 \| 11.5 | 11.5 \| 9.0 | 1.3 \| 17.9 |
| Anxiety | 21.5 \| 16.5 | 27.8 \| 31.6 | 6.3 \| 8.9 | 17.7\| 15.2 | 24.1 \| 7.6 | 2.5 \| 20.3 |
| Disgust | 2.5 \| 15.0 | 6.3 \| 10.0 | 60.0 \| 47.5 | 8.8 \| 10.0 | 6.3 \| 8.8 | 16.3 \| 8.8 |
| Happy | 17.3 \| 12.3 | 8.6 \| 14.8 | 2.5 \| 23.5 | 53.1 \| 25.9 | 17.3 \| 7.4 | 1.2 \| 16.0 |
| Sadness | 16.7 \| 15.4 | 20.5 \| 14.1 | 2.6 \| 11.5 | 11.5 \| 9.0 | 46.2 \| 35.9 | 2.6 \| 14.1 |
| Surprise | 1.3 \| 11.4 | 5.1 \| 15.2 | 13.9 \| 15.2 | 8.9 \| 10.1 | 5.1 \| 11.4 | 65.8 \| 36.7 |
| Ungrouped | 14.9 \| 4.1 | 40.5 \| 18.9 | 2.7 \| 12.2 | 5.4 \| 5.4 | 29.7 \| 25.7 | 6.8 \| 33.8 |

### 3.1. TIME DELAYED NEURAL NETWORKS AND THE RECOGNITION OF BASIC EMOTIONS FROM BREATHING RATES

The second version of DiverNet used a pressure sensor in the regulator. This method does not record all breathing movements directly, but represents a more reliable method of measuring breathing rate.

We used the SNNS 4.1 (Stuttgart Neural Network Simulator, University of Stuttgart)[6] to simulate a time delayed neural network (TDNN) consisting of three layers. This type of neural network is able to learn time structures in data and extract features translation invariant. The trained patterns can be export as c-code. The organization of the input units allows the manipulation of the time span the network uses and the number of features it tries to extract. In our case we used 12 feature units and 24 time units (= 750 input units). The hidden layer was 12 X 7 units and the output layer consist six units to depict the emotions. This network structure was used successfully by Grammer et al (1999[7], 2003)[8] for the depiction of gender from walking data. Before training, the weights were set randomly and the learning rules used was time-delay back-propagation.
As input we downsampled the breathing rates time series to 10 Hz and uses 336 input values. The number of learning patterns was n=425, test patterns were n=425, and verifications patterns were n=5.
Unfortunately it was not possible to identify basic emotions – only 16.24 % of all cases were classified correctly. We assume that there are several reasons for this. One surely are conceptual problems with emotion theories using discrete emotions. This is why we will rely in the rest of the workpackage on the identification of emotions in a pleasure-arousal-dominance/control space.

### 3.2. TIME DELAYED NEURAL NETWORKS AND THE RECOGNITION OF SECONDARY EMOTIONS: PLEASURE AROUSAL AND DOMINANCE FROM BREATHING RATES

In order to find out if it is possible to predict PAD scores from breathing rates we constructed 2 networks for each dimension. The architecture being basically the same as in the network above but with only three out put units negative, neutral and high.

These networks performed better than those for the recognition of basic emotions. Pleasure was qualified 39% correctly, arousal correct classification 38% and dominance/control with 35%. This basically corresponds to the analysis with classical statistical methods above.

## 4. SECONDARY EMOTIONS – MOTION, BREATHING AND HEART RATE

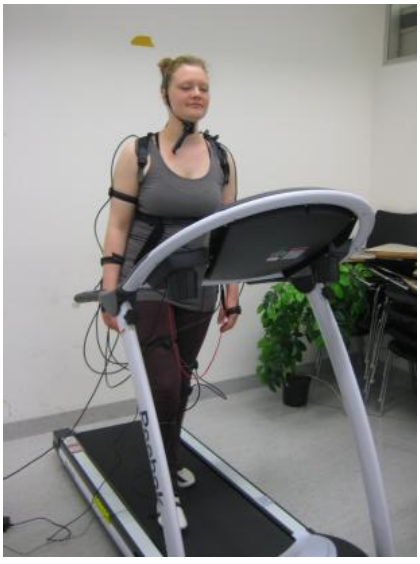

*Figure 0-1 Participant wearing DiverNet on the treadmill*

**Aim:** The aim of this experiment was to provide a large data set of motion data and physiological data in combination with data on emotional state. This dataset was used to train the algorithms for predicting the divers' emotional states from motion and physiological patterns. In addition, we related the motion patterns and physiological patterns to performance in a willpower task.

**Experiment:**

Sample: We collected data from 52 participants who were recruited on university premises and via social networks, of those data of 42

[6] http://www.ra.cs.uni-tuebingen.de/SNNS/
[7] Grammer, K., Honda, M., Juette, A., & Schmitt, A. (1999). Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *Journal of personality and social psychology*, *77*(3), 487.
[8] Grammer, K., Keki, V., Striebel, B., Atzmüller, M., & Fink, B. (2003). Bodies in motion: A window to the soul. In *Evolutionary aesthetics* (pp. 295-323). Springer Berlin Heidelberg.

participants could be used for analysis and algorithm training. The mean age of the 16 men was 26.7 years (S.D.: 4.5 yrs), the 26 women's was 24.4 years (S.D.: 3.4 yrs).

Motion rate calculation: We used the first derivative of the changes in body angles to calculate motion rate for each joint. Composite motion rates for the upper body, the lower body and the entire body were the respective sums.


*Figure 4-1 Experimental setup*

Emotion experiment:
The seven video stimuli described in section 4 were used to elicit the six basic emotions (anxiety, anger, disgust, happiness, sadness, surprise) and a neutral baseline. Participants were equipped with the DiverNet Sensor network, the breathing belt and a heart rate sensor. The participants were asked to select a comfortable walking speed on a treadmill. When they were ready, the video stimuli were displayed in a random order. In between video stimuli, participants filled in a Pleasure-Arousal-Dominance/Control questionnaire and completed a small cognitive task to avoid spill-over effects.

Willpower experiment:
After completion of the first experiment, participants were asked to complete a short breathing exercise after which they had to avoid looking at an attention grabbing video stimulus chosen in a pre-study while focusing their attention on a screen saver. In addition to the motion data and physiological data, we scored the participants' gaze behaviour.
The duration and frequency of gazes to the forbidden screen was used as an inverse measure of willpower. All participants filled in questionnaires about their personality, their exercise habits and other confounding factors that might affect their motion patterns.

**Results:**

Emotion experiment:
We used linear discriminant analysis to predict the video stimulus participants had watched based on the dynamic parameters of motion rate, breath rate and heart rate. The combination of these three groups of parameters yielded good predictive results as described in table 3.1.

Table 3.1 Results of Linear Discriminant Analysis

| Classification results | Predicted Group Membership<br>Percentage of correct classification | | | | | |
|---|---|---|---|---|---|---|
| Stimulus | anger | disgust | happy | sadness | anxiety | surprise |
| anger | 40.0 | 10 | 13.3 | 13.3 | 16.7 | 6.7 |
| disgust | 10.7 | 50.0 | 10.7 | 10.7 | 7.1 | 10.7 |
| happy | 6.9 | 10.3 | 55.2 | 6.9 | 6.9 | 13.8 |
| sadness | 22.2 | 3.7 | 11.1 | 33.3 | 22.2 | 7.4 |
| anxiety | 14.8 | 11.1 | 7.4 | 18.5 | 37.0 | 11.1 |
| surprise | 6.9 | 17.2 | 17.2 | 13.8 | 6.9 | 37.9 |

Willpower experiment:

The fast breathing exercise was associated with more frequent and longer distractions from the willpower task (gaze duration to the distractor stimulus: Mann Whitney U = 182, p =0.45; gaze frequency to the distractor stimulus: t-test, t(45)=-2.38, p =0.21).

We also found that a slow breathing exercise increased heart rate variability (t-test: t = 6.30, p > 0.001), whereas fast breathing decreased it (-2.955, p = 0.004), especially in the low frequency band (0.04-015Hz).

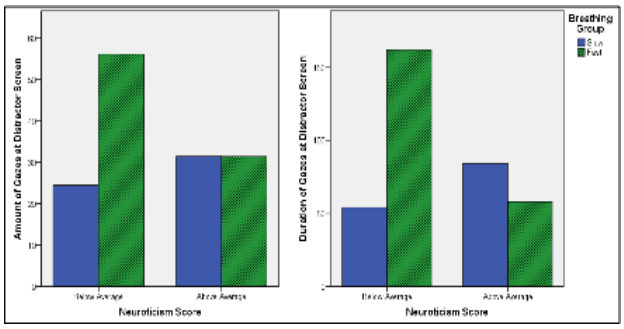


*Figure 3-3 Success at willpower task for participants with above and below average neuroticism scores*

In general, these results support the hypothesis that slower breathing and a greater heart rate variability are connected to self-control and the ability to avoid distractions during cognitively taxing tasks.

However, we found several mediating factors: greater may increase the ability to withstand distractions (Fig. 3.3) when breathing quickly, but decrease it if the participant breathes slowly. In addition, the time distance to the last meal was associated with a greater susceptibility to distractions (duration of gaze: R= 0.35, F (1,36)= 5.02, p = 0.031).

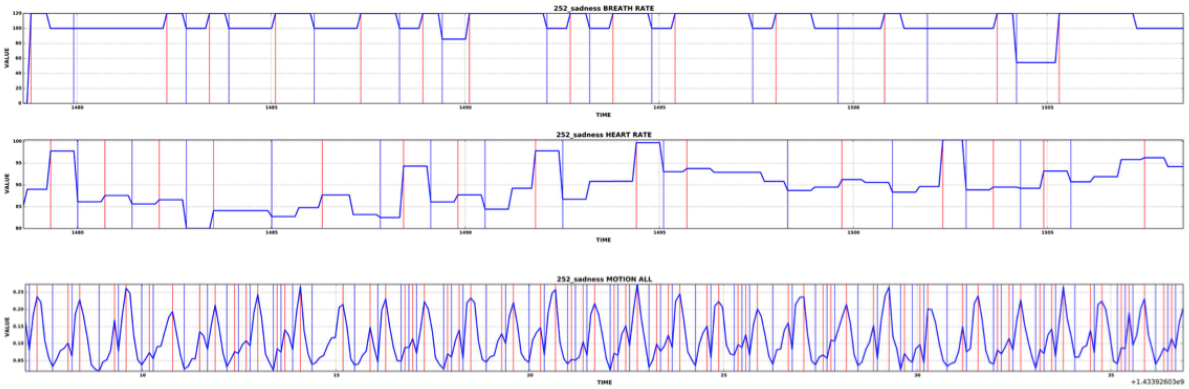


*Figure 3-4 Breath rate, heart rate and motion rate raw data*

### 4.1. MULTILAYER PERCEPTRON: CLASSIFICATION OF SECONDARY EMOTIONS FROM COMBINED SOURCES.

One disadvantage of TDNNs is that they will become very complex for training with combined data sources, like heart rate, breathing rate and motion rate. Another drawback is that folding can not be used in this package. The performance of the network might depend on the randomly selected learning and testing files – but this process can not be automatized and has to be done by hand. Thus we tried a different approach. A first multilayer perceptron was developed and tested in SPSS.

As input data we used the ten dynamic features for breath rate, heart rate and motion rate (n=130 learning data, n=56 testing data). The perceptron thus had 30 input units and a layer of 20 hidden units, nine output units for pleasure, arousal and dominance (each high, neutral and negative). Activation rule was TanH. This approach proved fairly successful with classification rates reaching from 60 % (dominance/control high) to only 18% (pleasure, neutral). Overall classification score was 40 %.


*Figure 3-5 This figure shows a MLP Network generated in SPSS. There are 30 input neurons which hold the dynamic information from breathing rate, heart rate and motion rate. 20 hidden units are activated by a thanH function. The output units denote three states (high.-neutral-negative for each of the three*

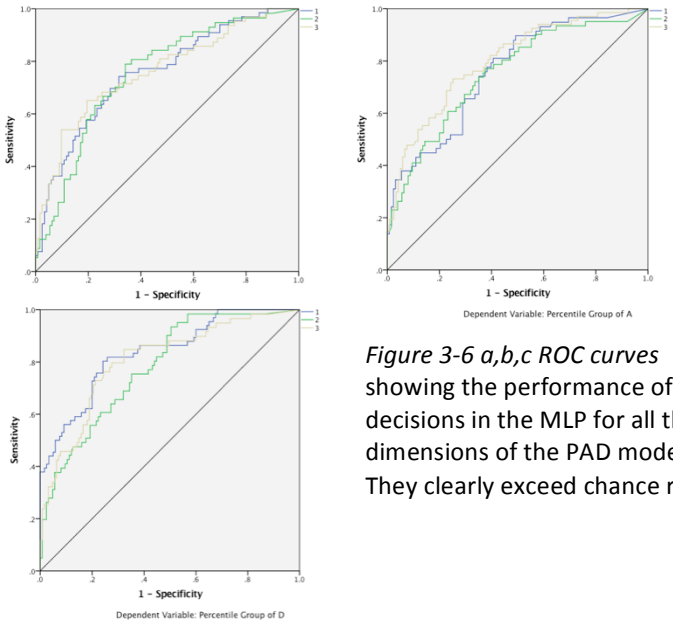


*Figure 3-6 a,b,c ROC curves* showing the performance of decisions in the MLP for all three dimensions of the PAD model. They clearly exceed chance rate.

This first approach then was continued in scikit-neuralnetwork 0.7 in a Python environment. This allows to search for optimal networks by varying learning algorithms with randomized number of hidden layers and units and the use of fold approaches were the training and testing data are selected randomly and used to test new networks – this allows to find the optimal network and assess the real performance. The classification rates are again in the same magnitudes Pleasure (0.49). arousal (0,43) , dominance/control (0.43) .The figures indicate that the results are symmetric – this means that the classification always is highest for the correct classification.

## 5. UNDERWATER EXPERIMENTS

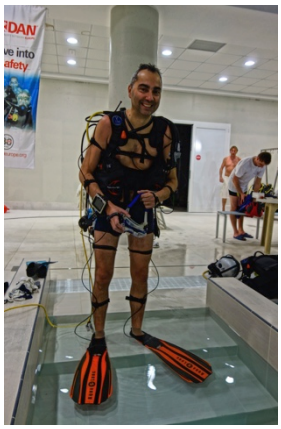

*Figure 4-1 Diver wearing DiverNet*

Aim: Collecting data from divers performing a wide range of tasks. This data was used to test the algorithms for internal state predictions obtained from the dry land data sets.

Experiments: Data collections took place in the Y-40 pool in Padova.

Sample:

Table 4.1. Sample description

| Data collection | Month/Year | Total participants (mean age) | Total Women (mean age) | Total Men (mean age) |
|---|---|---|---|---|
| Padova 1 | 06/2014 | 16 (45,87 yrs) | 4 (39,25 yrs) | 12 (48,08 yrs) |
| Padova 2 | 02/2015 | 19 (45,16 yrs) | 8 (46,75 yrs) | 11 (44,0 yrs) |
| Padova 3 | 06/2015 | 21 (45,60 yrs) | 17 (44,1 yrs) | 4 (52,0 yrs) |
| All | - | 56 (45,53 yrs) | 29 (44,34 yrs) | 27 (47,0 yrs) |

This data set yielded 251 time series segments of 35 seconds length that could be used for TDNN training and testing.

Tasks:
The divers completed the following tasks:

- Breathing underwater/above water
- Decompression stop
- Collecting objects
- Swimming fast/slow
- Moving up and down between markings
- Free behaviour

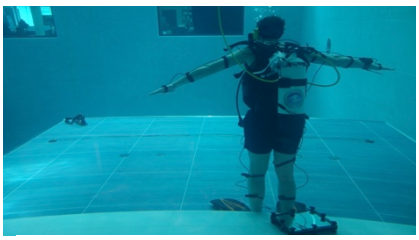

*Figure 4-1 Diver in T-posture*

In addition, divers were asked to fill in Pleasure-Arousal-Control questionnaires to record how they felt in between tasks. T-postures were used for calibration of the motion data.

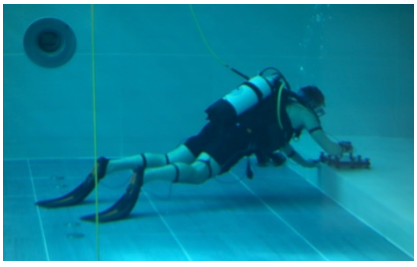

*Figure 4-2 Diver using the tablet*

## 5.1. TDNN AND MULTILAYER PERCEPTRON (MLP) FOR DIVER EMOTION PREDICTION

The data we used in this analysis were 35 seconds before filling out the PAD questionnaire and motion rates were calculated as first derivate from all body joints in a compound measure. We trained a TDNN with motion rates and PAD data.(Learning n= 170, Testing n=50, validation n= 29 patterns.

The overall correct classification for pleasure was 30% correct for Arousal 42 % and for dominance/control 42% . The MLP applied to the same data brought better results. Pleasure (47%)., Arousal 51%) and Dominance/Control %51%)

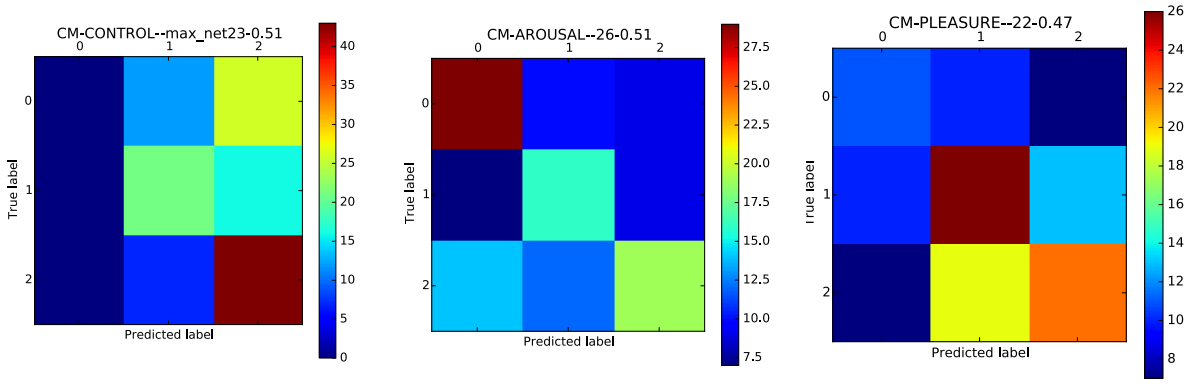


*Fig. 4-4 Classification performance plots of all three dimension of three networks for each dimension in the PAD. With the exception of control they all symmetrical, i.e. the score for the original values is the highest.*

## 5.2. DIVER POSE ESTIMATION

The reconstructed diver posture obtained from DiverNet is used for automatic activity classification, where the system would know what the diver is doing even without a human operator observing.

Two approaches are currently being tested. The first one uses a Dynamic time warping algorithm to align the live data with recorded training data. It classifies current activity based on the distance to aligned training set, choosing the activity with the smallest distance. This method showed good results with static postures, but is not as appropriate for dynamic activity which is important for most diving activities as it cannot cope with huge variances in speed in performing the same tasks. Tests were conducted with 7 poses shown in Figure 4-5. Confusion matrix for initial tests is shown in Figure 4-6.

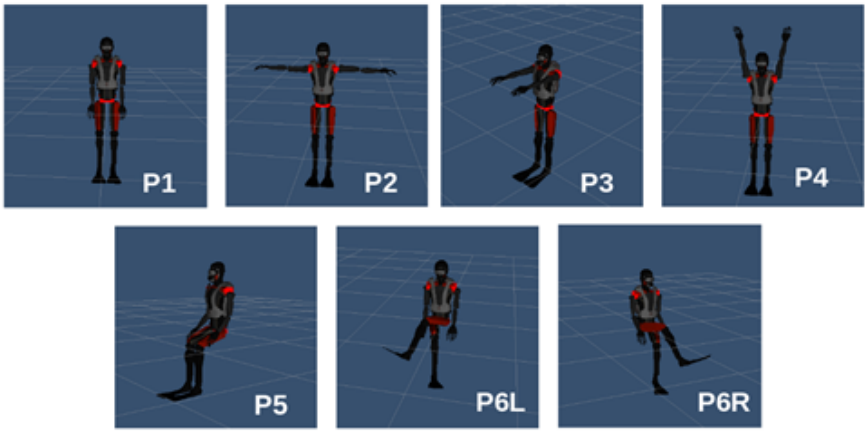


*Fig. 4-5 Poses used in initial static classification tests with Dynamic time warping*

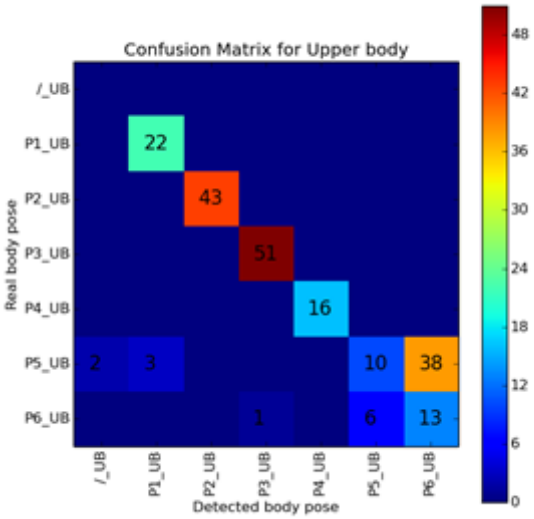


*Fig4-6 Confusion matrix for Dynamic time warping tests*

The second approach that is currently being tested are artificial neural networks. Inputs to the network are calculated joint orientations of the diver, for both current and past frames. This allows time awareness and recognition of dynamic activity. Neural network models are currently being assessed to find a good structure for our task. A simple and shallow (1 hidden layer) feed-forward neural network has been tested so far, and has shown good results on a small dataset with static poses. Provided with reasonably good reconstruction and joint orientation estimate from DiverNet, the algorithm works very well. However, more data needs to be collected to try to also cope with the reconstruction errors. As for now, some improvement was made with artificial data generation. The artificial training data was created by adding a Brownian motion-inspired error to existing training data.

The tests conducted with dynamic poses have shown even stronger need for more data. Even with relatively simple networks, it was hard to prevent the network from overfitting. Initial results, with a single hidden layer network, and using current frame, and three past frames with 0.5s sampling time, have resulted in achieving 100% in the test set, but only 80% in the validation set, showing severe overfitting. After recording some more data and adding artificial training data, the results have improved slightly, to 99% recognition rate on test set and 88% on training set. This still suggest quite a bit of overfitting, but is an improvement. The confusion matrix is shown in the figure 2.5.3.

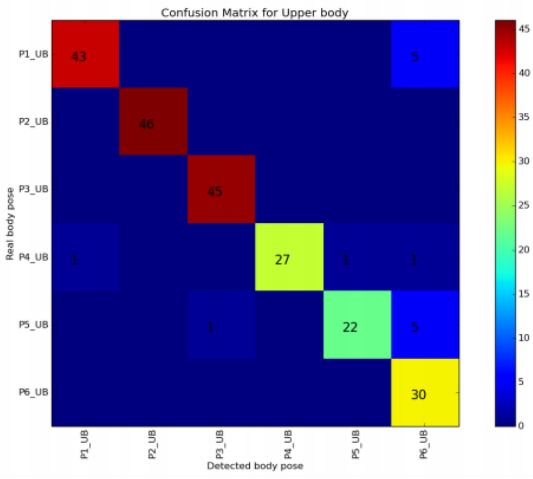


*Fig. 4-7 Confusion matrix for Artificial Neural Networks based dynamic activity recognition*

## 6. DIVER CONTROL ARCHITECTURE

We developed a program which allows real time play back of divernet data synchronized with video. This was used to test various algorithms and to segment divernet data for statistical processing and neural network development. It is necessary to note that this was a major challenge for us. This program is the prerequisite for any development we have done on the diver data. Features of the program include:

- Real time playback of divernet data synchronized with video and behaviour codes
- Calculation of dynamic time series features of heart rate, breath rate and freely combinable motion rates (i.e. whole body or only feet etc)
- Integration of questionnaire data
- Various filters on time series, Fourier analysis and histograms for artefact detection
- Event detection in time series
- Export of any data segmented for behaviours, questionnaires etc.
- Anomaly detection

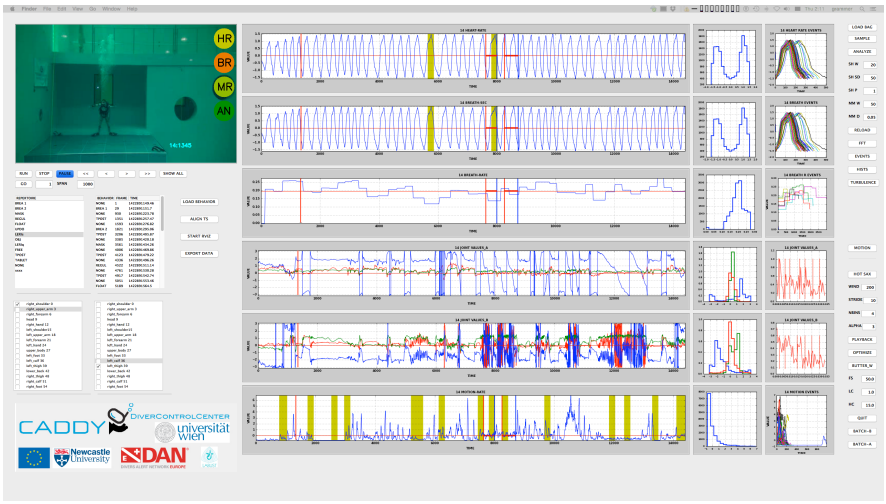


*Figure 6-1 Screenshot of the DiverNet player*

**Anomaly detection:**

From the data in the first experiments in Padua we developed a general procedure for the identification of unusual events in the times series. We implemented a symbolic aggregation approximation algorithm – SAX –(Keogh et al. 2005[9]) which allows us to detect anomalies and motifs in real time with minimal assumptions. SAX does dimensionality reduction and indexing with a lower bounding distance measure. The method is supposed to be equal if not superior to well-known representations such as Discrete Wavelet Transform (DWT) and Discrete Fourier Transform (DFT), while requiring less storage space.

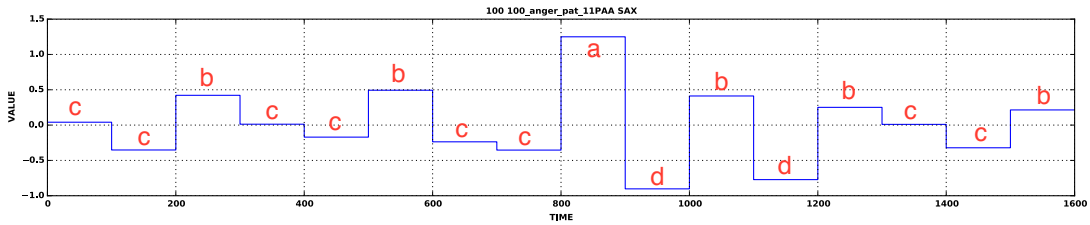


*Figure 6-2 PAA representation of time series segment. The alphabet is abcd and represents 4 bins of equal probability. The time series is then regrouped into 4 words: ccbc, ccbc, adbd, bccb*

[9] Keogh, E., Lin, J and Fu, A. (2005): HOT SAX: Efficiently Finding the most unusual time series subsequence. Proceedings of the Fifth IEEE International Conference on Data Mining (ICDM'05) 1550-4786.

In addition, the representation allows us to avail of the wealth of data structures and algorithms in bioinformatics or text mining, and also provides solutions to many challenges associated with current data mining tasks. The procedure consist of three simple steps:

7. z-score transformation of a time series
7. Piecewise Aggregate Approximation of time series (PAA). The PAA transformation is the discretization of the time-series. The efficiency can be controlled by calculating the distance between the PAA representation and the original time series. This method also can be used to determine alphabetsize (how many values are used for time series transformation), binsize (how many data points are aggregated in PAA) and wordsize (how many values form a word.
7. The PAA then is transformed into SAX letters and this is implemented in a way which produces symbols corresponding to the time-series features with equal probability. The extensive and rigorous analysis of various time-series datasets available to the original algorithm's authors has shown that the values of *z*-normalized time-series follow the Normal distribution. By using its properties it's easy to pick equal-sized areas under the Normal curve using lookup tables for the cut lines coordinates, slicing the under-the-Gaussian-curve area and assign a letter to a value.

The following graph shows the results. On the x –axis are the video frames, the blue line is the motion rate, the green line is the simple distance and the red line represents the Euclidian distance of a word to all other words in the window. As you can see the t-postures from the behaviour file which are presented as vertical red lines all fall on depressions in the similarity. This method allows the identification of certain behaviours and anomalies from motion rate and physiological measurements.

For the detection of anomalies in the string sequence we used the sequitur algorithm developed by Nevill-Manning and Witten (1997)[10]. This algorithm constructs a hierarchical grammar by substituting repeated phrases in a series with rules and produces an exact representation oft the sequence. Phrases which can not be substituted in the rule scheme than are anomalies.
Motion rates from divers form a special case – because the repetition rates are high only for special cases like swimming. In our case the experiment provides a collection of anomalies, like T-posture, fast swimming, tablet use etc

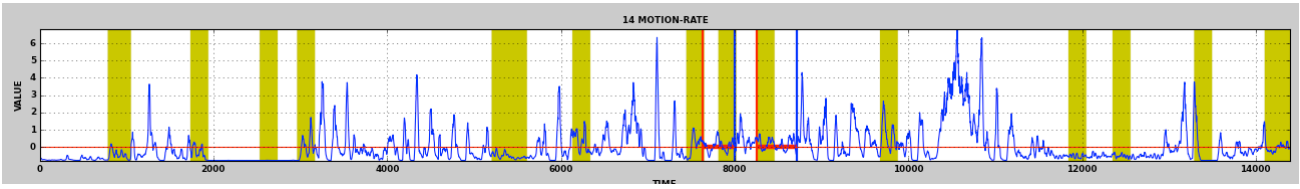


*Figure 6-1 SAX applied to one single dive. SAX detects 14 anomalies – two of them are situated in Left-Right swim slowly.*

[10] Nevill-Manning, Craig G., and Ian H. Witten (1997). Identifying hierarchical structure in sequences: A linear-time algorithm. J. Artif. Intell. Res.(JAIR) 7 (1997): 67-82.

Motion rate is calculated from the first derivate of joint values. The results show that this is possible in principle – but in our view it could be applied to the highly repetitive breathing curves and where it could be used to survey the functioning of the pressure valve.

**Final implementation of diver control module**

The flowchart depicts the final implementation we suggest for a diver control program: There are three main data sources (heart rate, breath rate and motion rate) .All three data streams have the basic dynamic features calculated over a 30 sec. window. These data are fed into the respective MLP. The determination of risk from heart rate and breathing is not trivial, because both depend from age of the diver and task. The most promising approach is to have a thirty second breathing and heart rate recording at the beginning of the dive under resting conditions and then determine the outliers as a function of the mean and the standard deviation of the resting period.
We also suggest to generate a fuzzy logic system which determines a general risk in an additive way.

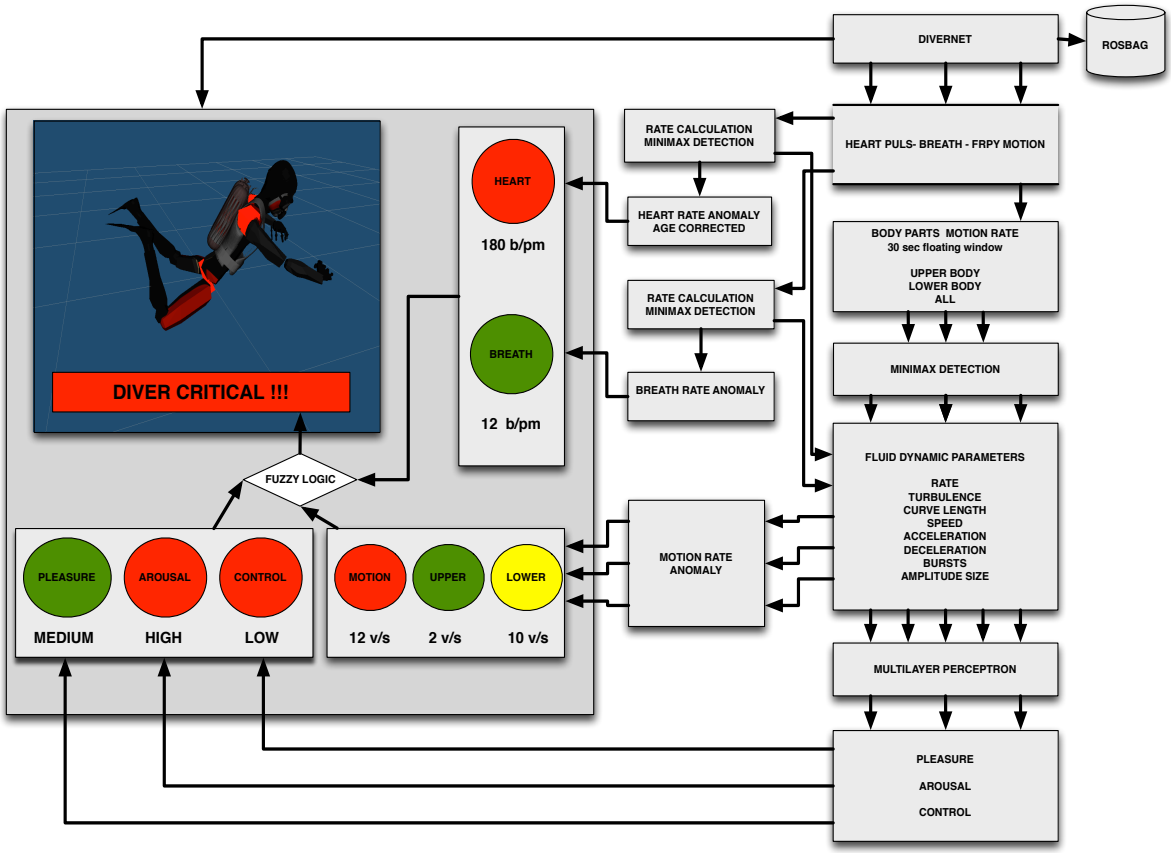


*Fig 5-4 Flowchart of suggested control parameters and variables for diver control. The flowchart outlines the procedures which are necessary to monitor the diver. We have implemented heart rate and heart rate anomaly based on age data, the same is done for breath rates. Red light depicts abnormal rates. We also used motion rates in this system and pleasure, arousal, dominance/control values from the neural networks. We plan to implement a Fuzzy logic module for decision making if the diver mission is critical or not.*

## 7. CONCLUSION

This deliverable documents the successful development of algorithms for diver behaviour interpretation. Based on dry land experiments (section 2-4) and experiments with divers (section 5) neural networks were designed to predict diver posture and the diver's psychological state from diver motion, breath rate and heart rate (variables are described in section 2). In addition, we developed software to monitor the diver's state in real time and to display relevant information to a surface crew. The results of this deliverable will be tested in trials in October 2016.