Automation in Construction 20 (2011) 289–298

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Using affective human–machine interface to increase the operation performance in virtual construction crane training system: A novel approach

Iman Mohammad Rezazadeh [a,*], Xiangyu Wang [b], Mohammad Firoozabadi [a,c], Mohammad Reza Hashemi Golpayegani [a,d]

[a] Department of Biomedical Eng., Science and Research Branch, Islamic Azad University, Tehran, Iran
[b] Senior Lecturer, Faculty of Built Environment, The University of New South Wales, Australia, and International Scholar at Department of Housing and Interior Design, Kyung Hee University, Seoul, Korea
[c] School of Medical Sciences, Tarbiat Modares University, Tehran, Iran
[d] Department of Biomedical Eng., Amir Kabir University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Accepted 24 August 2010
Available online 26 November 2010

Keywords:
Virtual Reality
Facial bioelectric-signals
Affective computing
Affective measures
Construction training

ABSTRACT

In the construction industry, some progress have been achieved by researchers to design and implement environments for task training using VR technology and its derivatives such as Augmented and Mixed Reality. Although, these developments have been well recognized at the application level, however crucial to the virtual training system is the effective and reliable measurement of training performance of the particular skill and handling the experiment for long-run. It is known that motor skills cannot be measured directly, but only inferred by observing behaviour or performance measures. The typical way of measuring performance is through measuring task completion time and accuracy, but can be supported by indirect measurement of some other factors. In this paper, a virtual crane training system has been developed which can be controlled using control commands extracted from facial gestures and is capable to lift up loads/materials in the virtual construction sites. Then, we integrate affective computing concept into the conventional VR training platform for measuring the cognitive load and level of satisfaction during performance using human's forehead bioelectric-signals. By employing the affective measures and our novel control scheme, the designed interface could be adapted to user's affective status during the performance in real-time. This adaptable user interface approach helps the trainee to cope with the training for long-run performance, leads to gaining more expertise and provides more effective transfer of learning to other operation environments. The detailed methodology of the affective control is presented in the paper. The results and future applications of the proposed method for disabled users, especially from neck down are discussed.

1. Introduction

Construction equipment operators operate one or several various types of powered construction equipment. Virtual technologies afford new opportunities for effectively training novices with lower cost and fewer hazards. Seidel and Chatelier [1] have suggested, for example, that the use of virtual environments (VEs) may be “training's future”. Virtual environments can be especially valuable where training in real-world situations would be impractical because a real field scenario may be unduly expensive, logistically difficult, dangerous, or too difficult to control. Compelling virtual environments could lead human participants to feel somehow present, for purposes of training. Virtual environments involve structure of a human participant existing within some kind of an artificial interaction environment. Part of the components involved in the interactive virtual training environment are simulated, the operator nevertheless can experience a similar sense of being present and interacting with real/virtual objects via visual, auditory or force displays. Virtual environments can also assist with the delivery of equipment operation training during inclement weather conditions and novices have much more time to practice their skills without the pressure of costs.

Research has shown that the key to acquiring the necessary motor skills to control complex systems, such as a backhoe excavator, is hands-on and coached training [2]. This approach is envisaged to facilitate progress along what is a steep learning curve and enable effective rehearsal of future operations in actual construction sites. The promise of effectiveness is supported by evidence from mental health research revealing that a virtual experience can evoke the same reactions and emotions as a real experience [3]. These technologies should become the bridge connecting the ideal training objective to the current reality of training programs.

Research efforts in VR training for construction have been predominantly focused on a proof-of-concept level of implementing

* Corresponding author.
E-mail addresses: I_rezazadeh@yahoo.com, i.rezazadeh@srbiau.ac.ir (I.M. Rezazadeh).

0926-5805/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.autcon.2010.10.005

computer-based training applications to enhance the training process. However, crucial to the virtual training system is the effective and reliable measurement of training performance of a particular skill. It is known that motor skills cannot be measured directly, but only inferred by observing behaviour or performance measures. The typical way of measuring performance is to through measuring task completion time and accuracy, but can be supported by indirect measurement of some other factors. Examples of indirect methods include quality of output, consistency of operation, smoothness of bucket motion, proactive thinking, adaptability to changing environments, capability for skill improvement, and oral interview with expert.

It has been shown that for a successful acceptance of a new Human–Machine Interface (HMI), the usability (ease of use) of the interface is an important factor. The interface could be rigid with no adaptability over time, means its users MUST adapt themselves to it. Because of this unilateral information flow from the machine to the human, more cognitive pressure (overload) imposes on the user, and this cause to linger the process of gaining expertise, and learning transfer to the other operation environments. One approach to overcome the above drawback is to establish a bi-lateral information flow by designing an adaptable and real-time interface. In this case, the core of the interface (controller) adapts itself to its user's status and the user does the same, simultaneously. Thus, the imposed cognitive pressure is reduced and it facilitates the training process and consequentially, the transfer of learning to the other contexts.

When disabled users are the main target for using a HMI, they may have more limitations and difficulties for manipulating it. In this case, more concerns about usability factor should be dedicated. Thus, the designed interface should be collaborative, interactive, and motivate them to cope with the task and make them ready to get back to daily activities of life.

This paper introduced a new approach to design and implement an adaptive HMI based on the human subject's affective measures. This study is dedicated essentially to the creation of compelling virtual environments within which human participants are led to feel somehow present, for purposes of training. More specifically, a novel facial multi-channel bioelectric-signals processing approach was developed to extract affective measures and control commands. Then, the extracted commands are applied on a designed virtual crane to test and validate our approach. The main reason for using the facial gestures is to investigate the feasibility of the proposed method on disabled users for future applications. It has been shown that, by using this approach, the user can cope with training for a longer period of time and also gets better performance results and expertise factor compare to traditional methods.

The results of this research could enable better development, implementation, and assessment of HMIs, especially for neck-down disables. Given the size of the construction industry and other related industries (e.g., manufacturing), and by considering the social and mental effects of the disabled getting back to the working sites, the results of this research are expected to directly impact on the workforce and economy. Section 1.1 presents a thorough state-of-the-art review for the applications of virtual technologies in training. Section 1.2 introduces the relevant work of affective computing in virtual environments. Section 2 presents the complete methodology of a novel facial multi-channel bioelectric-signals processing approach to extract affective measures. Section 3 discusses the results from an experimentation based on the approach.

1.1. Virtual Reality for construction training

Construction training research has begun to explore computer simulators, tele-operation, and Virtual Reality as training vehicles. The term Virtual Reality (VR) is used for a great deal of different situations. The difference between VR systems is the amount of immersion which stretches from those displayed on an ordinary computer monitor to systems using head-mounted-display (HMD) with a large field of view and a body suit with sensitivity feedback. Immersive VR environment is defined as one in which the user is totally immersed in a completely synthetic world, which might mimic some properties of the real-world but it can also exceed the bounds of physical reality. Non-immersive VR is referred as desktop Virtual Reality or windows on the world, which does not provide full sense of immersion but can be readily configured on an ordinary PC with relatively high mobility.

Virtual Reality training systems have already provided added benefits to many training packages. VEs embody many of the characteristics of an ideal training medium [4,5]. Also, the use of VEs allows the trainer total control of the stimulus, environment situation and the nature and pattern of feedback, and also allows comprehensive monitoring of performance. Sometimes, by combining an educational simulation, the use of VEs may also engender increased levels of motivation [5]. VEs have already been developed for training of drivers [6], fire-fighters [7], pilots [8], console operators [9], naval officers in ship manoeuvres [10], soldiers in battlefield simulations [11,12], and ground control team for familiarity with the operability of the Hubble Space Telescope [13]. In one demonstration, a virtual robotic manufacturing line was created for worker training [14,15]. This study included a control group and provided conclusive evidence of VR's ability to facilitate worker training. Another related example is that Wilson et al. [16] have developed a Virtual Environment Crane Training Simulator (VECTS) for overhead crane operator training. This simulator provides a trainee with a 3-D virtual environment for learning and practicing the skills needed for productive crane operation in the actual factory, and eliminates the need to take a productive crane out of service. A number of industrial training applications of virtual reality are discussed in [17].

Even though virtual environments are extensively used in industrial training, there have been few rigorous investigations of its applications in equipment operator training in construction. Such established uses of VR are becoming both more widespread and more compelling, thus the obvious choice to explore VR in construction heavy equipment operator training. As an effort toward this direction, Wakefield et al. [18] developed an interactive excavator VR simulator for operator training. A more recent and similar training system developed by Simlog Company [19] is a desktop VR-based personal tower crane simulator. Fig. 1 shows two screenshots from Simlog's Tower Crane Simulator, one from outside the cab and the other from inside the cab. The first image is a view from outside the cab showing the tower, the jib, the counter jib, the operator cab, the trolley, the hook-block, and a sample load. The second image is a view from inside the cab showing a wooden box and load target position. This system tests load chart knowledge while the operator is executing lifts using various loads and hook-block assemblies. The target position for the load is shown in red wire-frame. The load-moment indicator in green color is showing that operation is well within the load chart limits for this crane.

1.2. Affective computing in training

Apart from the noted existing VR research in training, emotion recognition (such as joy, teasing, fear, sadness, disgust, anger, surprise, and neutral) becomes an important research issue in the robotics and automation area. A number of studies have been done in this area, ranging from psycho-physiological measures such as heart rate, Electro Dermal Analysis (EDA), pupillometrics or facial muscle activities (fEMG) to speech or video analysis. Different types of evaluation studies show the success rates are 50–60% in emotional speech recognition and 80–90% in facial expressions. Indeed, physiological factors are useful indices to evaluate emotions, since they can be measured and can be applied to engineering approaches. One of the approaches which have been gaining many attentions in this area is affective computing.

Affective computing means computing that relates to, arises from, or deliberately influences emotions. It focuses on creating personal computing systems having ability to sense, recognize and understand

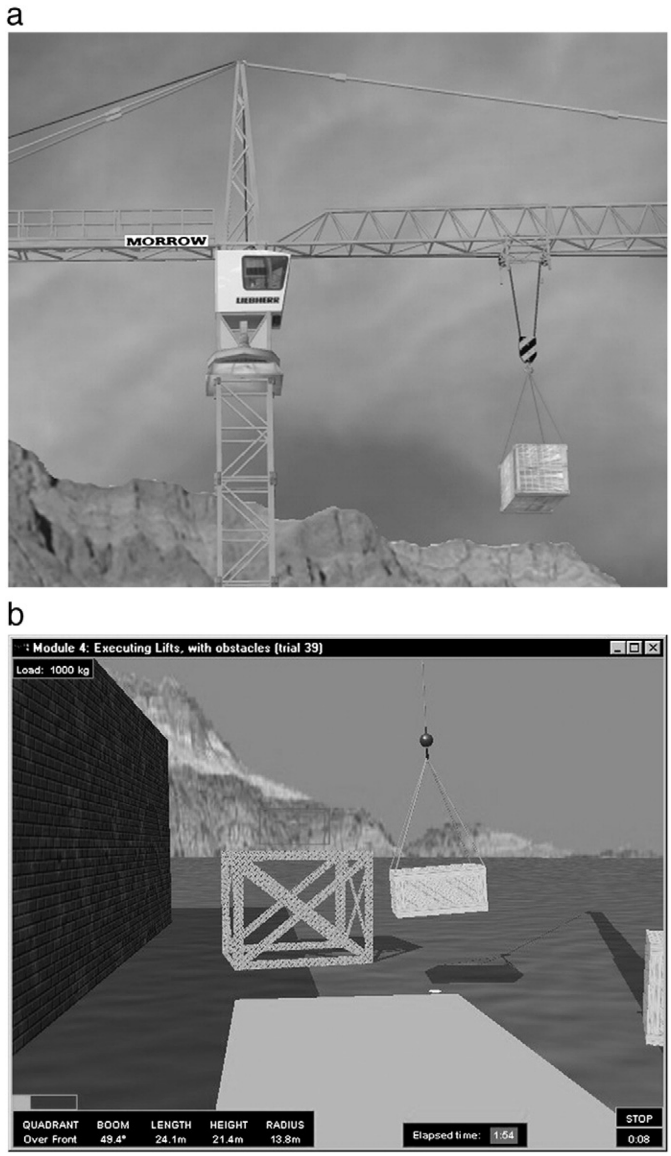

Fig. 1. Simlog's tower crane simulator: (a) from outside the cab and (b) from inside the cab [19].

human emotions, together with the skills to respond in an intelligent, sensitive and respectful manner toward the user and his emotions [20–23].

Brezillon et al. [24] have proposed a decision making system to provide a driver with a system for a self training and self evaluation of his behaviour in different pre-critical situations. They considered time constant for the driver's answers to predefined questions as an essential parameter the context's situation and driver's behaviour when experimenting with the provided scenarios.

Ang et al. [25] stated that facial muscle movements and (fEMG) can be corresponding to certain facial expressions and are the most important visual representation of a person's physical emotional states. Mahlke and Minge [26] used emotional states which were extracted from fEMG to discriminate between usable and unusable computerized contexts. They placed two pairs of electrodes on *zygomaticus* major and *corrugators supercili* to detect positive and negative emotional states, respectively. They concluded that the frowning activity is significantly higher in the unusable system condition than in the usable one. Neimenlehto et al. [27] studied the effects of affective interventions using fEMG in a Human–Computer Interaction (HCI) and concluded that the frowning activity attenuated significantly after the positive interventions than the conditions with no intervention. Nasoz et al. [22] have conducted a study to model user emotional state. Three physiological measurements were used and data was collected from 31 participants (male and female) from student population. This study used normalized signals instead of statistical features and employed Discriminant Function Analysis (DFA). They achieved 90% discrimination ratio for fear, 87.5% for sadness, 78.58% for anger, 56.25% for surprise and 50% for frustration. Healey and Picard [21] used four physiological signals to detect the intensity of stress in automobile drivers. Sequential Forward Floating Search (SFFS) was used to recognize patterns of stress whereby the intensity of stress was recognized with 88.6% accuracy. Mohammad Rezazadeh et al. [28] implemented a novel facial multi-channel bioelectric interface for controlling a virtual crane. They discriminated 5 different facial gestures using subtractive fuzzy clustering method (SFCM) and achieve 92.6% average discrimination ratio among them. Also, they introduced new affective channel from the subject's forehead, which could mirror user's affective states during the experiment.

2. Methods and materials

As mentioned in the Introduction, the goal of this study is to develop an adaptable human–machine interface based on affective measures mainly for subjects with the difficulties to control and manipulate the interface using their upper or lower limbs. Thus, the facial muscles are used in this study for generating command signals and manipulating the virtual crane. The main reason to use the face as an interface in our approach is that the human face is a rich resource of information from syntactic to semantic and pragmatic aspects, and most facial gestures are natural and voluntary and can be easily generated by many individuals even whose motor impediments are mainly from the neck down. The physical state (such as gestures or fatigue) of a user can be monitored by recording and processing the facial muscles activity signals characteristics (such as amplitude, power, frequency spectrum and so forth). In addition, the user's eye movement (EOG) and brain activity signals (EEG) characteristics (amplitude, entropy, and phase space for example) can be also used to indicate the user's affective states. Thus, the proposed methodology is a multimodal approach for improving and enhancing the interface's situational awareness. It prevents saturation or overload on the physical states which could bring physical and mental fatigue, indeed.

This section presents the novel facial multi-channel bioelectric-signals processing approach for extracting affective measures. A virtual crane training system is developed as the substrate for testing this affective approach through integrating multi-channel bioelectric-signals processing approach. Other equipment training platforms can also be used such as VEs for backhoe, excavator, forklift, etc. The general block diagram of our method to design the training interface is depicted in Fig. 2. As shown in Fig. 2, the interface can be divided into the following units: *Virtual Crane Manipulation* (*VCM*), *Affective Cues Extraction* (*ACE*) *and the Control Unit*. VCM uses the captured signals from the two pairs of electrodes located on *Temporalis* muscles for manipulating the virtual crane. Meanwhile, the Control Unit modifies and updates the VCM's inference system (decision machine) by using the human subject's forehead signals which are recorded for extracting the affective measures based on ACE scheme. The following subsections elaborate more details of the components in the block diagram shown in Fig. 2.

2.1. Site selection and placement of electrodes

Before placing any electrode, the selected skin area should be cleansed from dust, sweat and fat layer to reduce the effects of motion artifacts. As illustrated in Fig. 3, three pairs of rounded pre-gelled Ag/AgCl electrodes were placed on the human subject's facial muscles in a

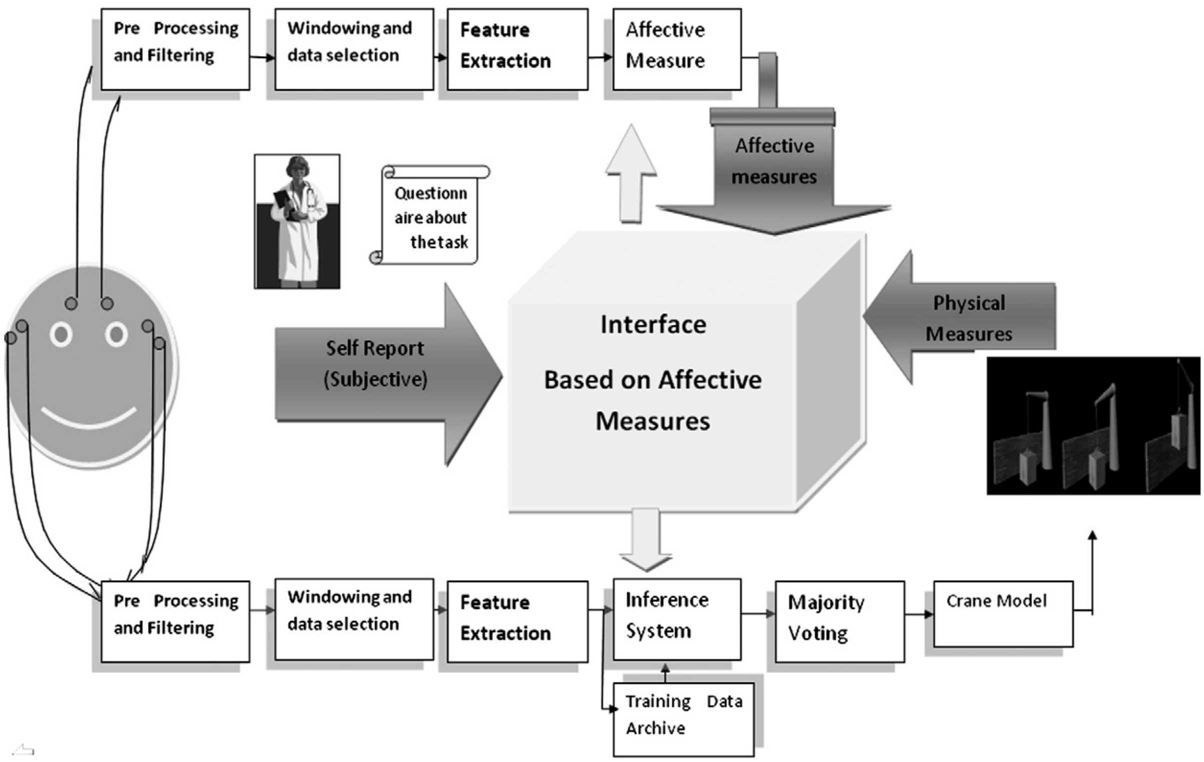

Fig. 2. General block diagram of proposed system.

differentiation configuration to harness the highest amplitude signals [28–30].

One pair is placed on the subject's *Frontalis* muscle: above the eyebrows with 2 cm inter-electrodes distance (Channel 2). This channel is employed for ACE unit which is responsible for extracting the affective measures and cues during operations.

- Two pairs are placed on left and right *Temporalis* muscles (Channels 1 and 3) which are responsible for gathering manipulating signals for VCM unit.
- One ground electrode is placed on the bony part of the left wrist.

2.2. Data acquisition setup

The Biopac system (MP100 model and ack100w software version) [40] was used to acquire bioelectric-signals. It can accurately collect bioelectric-signals with the selected sampling frequency and store them in its own or PC's memory. The sampling frequency and amplifier gain are selected at 1000 Hz and 5000, respectively. The low cut-off frequency of the filter is chosen to be 0.1 Hz to avoid motion artifact. In addition, a narrow band-stop filter (48 Hz–52 Hz) is used to eliminate line noise.

2.3. Data recording protocol for initial training of classifier

The recording protocol for obtaining initial fuzzy inference system has been designed to record 5 different facial gestures as the control commands for manipulating a virtual crane. Ten healthy human subjects have been chosen for this study (male and aged 23 ± 2 year old, all from the school of Biomedical engineering, Azad University, Tehran, Iran).

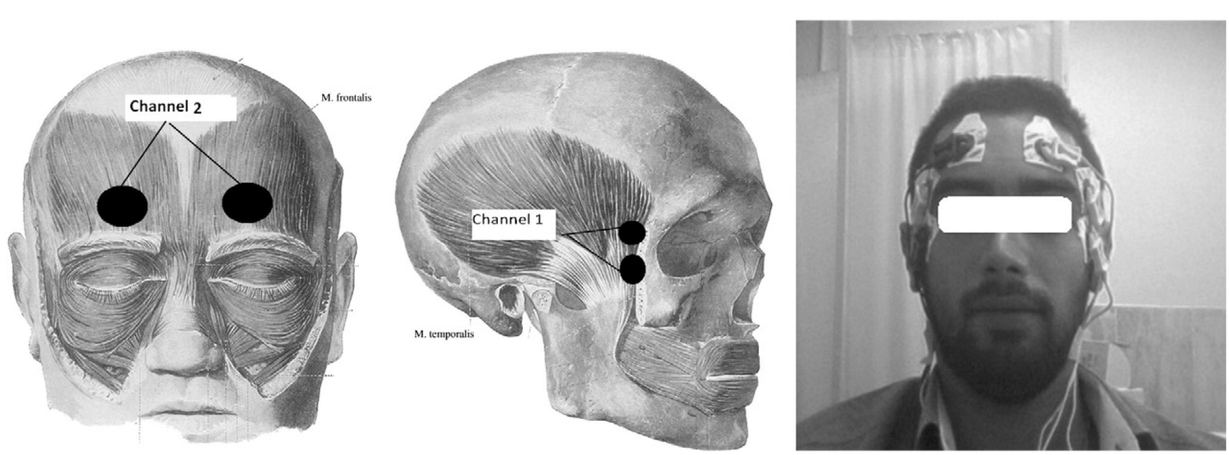

Fig. 3. Illustration of the electrodes configuration over Frontalis and Temporalis facial muscles [28–30]. The configuration of the Channel 3 is the same as Channel 1, but on the opposite side of the face.

2.3.1. Protocol

Each subject rested for 5 min prior to recording session and his *quiescent* signal was captured for a period of 1 min after the rest period. This signal determines the level of relaxation and is used as a threshold for determining the active state for the facial muscles. Then, the subject is asked to *moderately* generate each of the facial gestures as in Table 1 for 10 times.

The gesture generation for each of the mentioned gesture starts 1 s prior to recording and lasts until the end of the recording trial (2 s). The data from all the three channels is recorded for this period. There is a 10-s gap between each trial in order to eliminate the effects of physical fatigue. After 10-s interval, the subject is asked again to repeat the gesture and the above movement-rest task is cycled 10 times. Then, the subject should perform another gesture according to above protocol.

2.4. Feature extraction for VCM unit

The acquired data from Channels 1 and 3 were passed through a band-passed butterworth filter ranges 30–450 Hz which covers the most significant spectrum related to facial electrical activity [30,31]. Then the data was divided into non-overlapped 256 ms time slots. For each slot, the root mean square (RMS) value was calculated as a manipulator feature (R_i). (Eq. (1))

$$R_i = RMS(EMG_i) = \sqrt{\frac{\int_0^T EMG_i^2 dt}{T}} \tag{1}$$

Those manipulation features, whose values were greater than the threshold, were considered as active features (T_i) (Eq. (2)).

$$T_i = \{R_i \geq 3Mean(RMS(EMG_{Quiescent})) + 3std(RMS(EMG_{Quiescent}))\} \tag{2}$$

Finally, the normalized manipulator features were achieved using Eq. (3):

$$S_i = \frac{T_i - Mean(RMS(EMG_{Quiescent}))}{\sum_{i=1}^{K}(T_i - RMS(EMG_{Quiescent}))} \tag{3}$$

Furthermore, to have a more separable feature space, all extracted features were transformed to a non-linear simple feature space using a *log* transform to spread the concentrated data points while condensing the highly scattered points. [29].

2.5. Classification

Extracted features need to be classified into distinctive classes for the recognition of the desired gesture. In addition to inheriting variation of bioelectric-signal over time, there are external factors, such as changes in electrodes position, fatigue, and sweat which may cause changes in a signal pattern over time. A classifier should be able to cope with such varying patterns optimally, as well as prevent over fitting. Classification should be adequately fast to meet real-time processing constraints. A suitable classifier has to be efficient in classifying novel patterns. On-line training can maintain the stability of classification performance over a long-term operation [31].

The idea of fuzzy clustering is to divide the data space into fuzzy clusters, each representing one specific part of the system behaviour. Fuzzy c-means is one the fuzzy clustering methods which is a supervised algorithm, because it is necessary to tell it how many clusters *c* to look for. If the number of centres is not known before, it is necessary to apply an unsupervised algorithm.

Subtractive fuzzy-means clustering (SFCM) is based on the measurement of the density of data points in the feature space. The idea is to find regions in the feature space with high density of data points. The point with the highest number of neighbours is selected as centre for a cluster. The data points within a pre-specified, fuzzy radius are then removed (subtracted), and the algorithm looks for a new point with the highest number of neighbours. Subtractive clustering uses data points as the candidates for cluster centres, instead of grid points as in mountain clustering. This means that the computation is now proportional to the problem size instead of the problem dimension. Based on the above description, the *k-folds* algorithm was applied to the training set; where *k* equals to 10. The k-1 folds were used to train the classifier and applied to SFCM to derive fuzzy inference system and the rest 1 fold used to validate it [32,33].

2.6. Affective feature extraction for ACE unit

The acquired data from Channel 2 was passed through a 7–12 Hz band-passed filter for collecting α sub-band of the subject's brain electrical activity. This sub-band was chosen because research has identified its relationship to different cognitive functions. The α sub-band activity is modulated by semantic memory processes and also related to attentional task demands, level of alertness, expectancy, mental relaxation and satisfaction during the performance. It was reported that, less complex EEG patterns have been observed in more intelligent individuals and in more creative ones. It is assumed that the displayed reduction of the complexity of neural dynamics in high intelligent individuals is due to the inhibition of irrelevant and competitive activity. Alternatively, less intelligent individuals are characterized by more diffuse neural dynamics when performing the same task. One interpretation of such increases in amplitude in high intelligent individuals is a reduced neural network activity in regions not relevant for task performance — the neural efficiency hypothesis.

In EEG signal, the measures of dimensional complexity reflect the complexity of neural generators — the relative number of concurrently oscillating neuronal assemblies and degrees of freedom in the competitive interaction between them. Dimensional complexity appears to be relatively independent on EEG spectral power and inversely related to coherence, suggesting that it may qualitatively reflect different aspects of brain dynamics which cannot be detected by traditional spectral power measures. *Statistical entropy* could appear as a measurement of the system dimensional complexity and the degree of order/disorder of a signal. Therefore, it can provide useful information about the underlying dynamical process associated with the signal [34].

According to the above discussion, one can conclude that, when the subject is satisfied with the environment while performing the requested task or the interface is usable for him, the degree of disorder of the brain's signal will be reduced. Thus, the entropy feature could be considered as a good affective measure (cue) which mirrors user's comfort/discomfort during the performance. Thus, the filtered data was divided into non-overlapped 128 ms time slots. Then, the *log energy entropy* (*LogEn*, hereafter, entropy or statistical entropy) value for each time slot has been calculated using the method described in Aydin et al. [35].

The reason to choose Channel 2 for acquiring the affective measures is that it can capture forehead bioelectric-signals and according to [29],

Table 1
Gestures' name and related movements.

Gesture no.	Gesture name	Related command	Most relative physical data channel
1	Smiling	Move forward	Channels 1 & 3
2	Pulling up right lip corner	Move right	Channel 3
3	Pulling up left lip corner	Move left	Channel 1
4	Opening mouth (like to say ‘a’ in ‘apple’)	Move backward	Channels 1 & 3
5	Clenching Molar teeth	Lift/Release the load	Channels 1 & 3

there was no source of cross talks in the manipulation commands from Channels 1 and 3 in Channel 2.

2.7. Control unit scheme

Providing an *on-line/real-time training* approach, in which a classifier is trained and updated continuously using new patterns during operation is an inevitable need for retaining the rate of the accuracy stable and having a robust control system. There are two critical issues in on-line training: 1) the recognition and updating of valid training data, and 2) applying the training algorithm during operation. Training data requires clarification as to whether classified patterns coincide with user intention. Therefore, input–output pairs must be monitored and evaluated continuously to update training data. In addition, applying a training algorithm to a classifier during operation requires a distinguished method [31].

The control scheme of our approach is depicted in Fig. 4. It is based on a novel affective and adaptable control method described hereafter. As it shows, after acquiring initial inference system from the offline training, the control unit, which is the core of our proposed HMI, modifies the achieved inference system of the VCM unit based on ACE's outputs. It monitors the average entropy measures within a predefined period of time (*TTM: Time to Monitor*) while the subject is trying to undertake the requested task protocol (Section 3) using his/her facial muscles. Then, the Algorithm I is applied to the valid datum.

As stated in the algorithm I, if the average entropy measure is beyond the predefined threshold, it means the level of cognitive load is within acceptable range and the current input–output pairs could be considered to be valid. Then, if the old inputs are within the predefined distance from the new inputs in the feature space and the outputs are the same, then, the old data will be appended into new inputs–output pairs. This task is necessary to retain the *dynamics and backgrounds* of VHM inputs' patterns within the control unit. However, it should be mentioned that, appending old data to new ones may cause redundancy and require more storage space. Since huge training data from the new data set leads to more training time for the inference system, the number of data in the new data set ($\mathbb{P}$) should not be too large to cope with the limitations for on-line training. Thus, a heuristics method has been applied on each data in $\mathbb{P}$ set. Each data has its own importance factor, and as the time goes on, its importance factor reduces, exponentially using *the forgetting factor*. Then, the data in $\mathbb{P}$ with high importance factor will be chosen for forming active $\mathbb{P}$ set ($\mathbb{P}_{active}$). The SFCM + ANFIS method will be applied to $\mathbb{P}_{active}$ to obtain new fuzzy inference system (FIS). If the new FIS outperforms old FIS for the testing set, then it will be substituted for old FIS, otherwise the last FIS will be used for the next TTM too

Algorithm I

// Previous (Last) Inputs-Outputs Pair from Channel 1 and 2; I: Input; O: Output

$\mathbb{P}=\{I_P\text{-}O_P\}$;

// New (Current) Inputs-Outputs Pair from Channels 1 and 2

$\mathbb{N}=\{I_N\text{-}O_N\}$;

IF the average entropy $_{Channel\ 3} \leq$ *Entropy Validation Threshold (EVT) THEN*

{

DO // *For Every New Input–Output pair*

{

IF $((\|I_{Ni}-I_{Pj}\| \leq$ *Valid Distance Threshold*$))$ *AND* $((O_{N\,i} == O_{P\,j}))$ *THEN*

// $\|I_{Ni}-I_{Pj}\|$ *: Euclidean distance between New and Previous input–output pairs*

{

// Append those OLD Input–Output pairs to $\mathbb{N}$ *set which meet the above criteria*

$\mathbb{N}=\mathbb{N}\oplus\{I_{P\,j}\text{-}O_{P\,j}\}$;

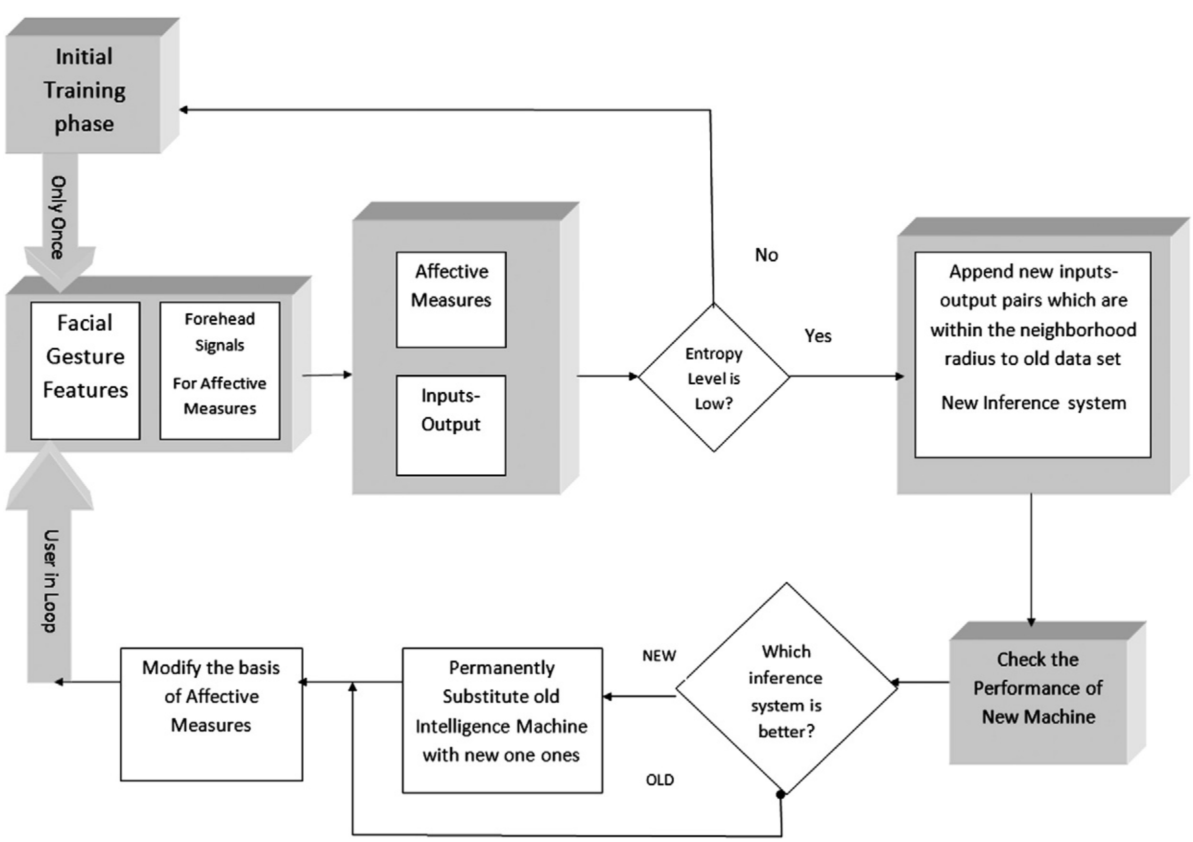

Fig. 4. Affective and adaptable control method block diagram.

// Remove those Input–Output pairs from $\mathbb{P}$ which meet the above criteria
$\mathbb{P}=\mathbb{P}\ominus\{I_{P\,j}\text{-}O_{p\,j}\}$;
}
END IF
$\mathbb{P}=\mathbb{N}$; *// SET $\mathbb{P}$ main set which contains all valid datum*
// Updating the importance factor for each input—output pair in $\mathbb{P}$
// Select Data in $\mathbb{P}$ which has importance factor higher than predefined threshold ($\mathbb{P}_{active}$)
// The training algorithm will be applied to $\mathbb{P}_{active}$ to obtain new fuzzy inference system (FIS) based on SFCM
IF (the new FIS outperforms old FIS for the testing set from $\mathbb{P}_{active}$) THEN
SET new FIS as main FIS of the system
END IF
}
END IF

End of Algorithm I

2.8. Virtual crane

Two different virtual training environments have been created. The first one (VE1) was an in-house virtual crane built by MAYA and it could lift, move and release a virtual load according to manipulation commands as shown in Fig. 5. The second virtual environment (VE2) was also a portal crane from *HUMUSOFT s.r.o.* and the *MathWorks, Inc.* (Fig. 6) [36]. It could lift the virtual load and be controlled via manipulation commands from facial gestures to make the desired trajectory of movement.

The reaction time (speed) on each environment could be tuned to different speed levels to evaluate the effect of difficulty on our proposed methods. If the speed level is so high, then controlling the crane will be tough and many corrective movements should be performed by the user to undertake the task correctly. On the other hand, if the speed of the crane is so slow, then the user should exert more facial activities to undertake the task. This brings muscular and consequentially mental fatigue for the subject.

It should be noted that, the two different VEs have been chosen for our experiment for evaluating the effect of VE's variation on our proposed method.

3. Experimental results and analysis

Each of the 10 human subjects was asked to perform data collection protocol. Table 2 shows the average discrimination ration for the mentioned facial gestures using SFCM. It is clear that the classifier has good power to discriminate between facial gestures. It should be noted that, the generating of these gestures is natural and that is why this gesture combination could be used as a good interface for human machine applications.

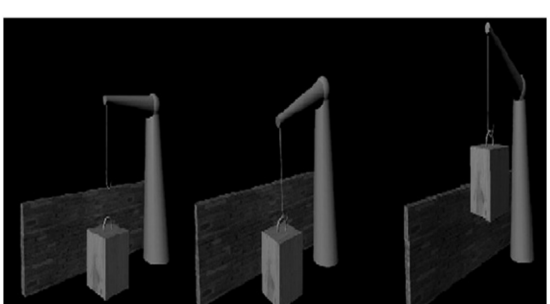

Fig. 5. VE1: Virtual crane could be used to place the load on the other side of the wall [28].

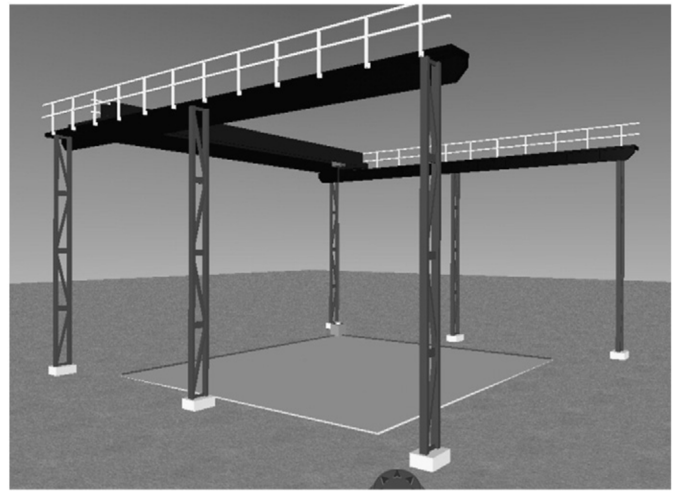

Fig. 6. VE2: Virtual portal crane could be used to place the load in different locations [28].

As stated above, the affective measures could mirror the level of difficulties and satisfaction during performing the requested tasks. In this experiment, the performance of our method and designing scheme have been evaluated using two VEs and with different level of difficulties. The Control unit can be set into two different controlling conditions: *Control unit ON/OFF*. If the control unit is in ON state then, the inference system will be updated using the affective measures and controlling scheme described above. Otherwise, the inference system will not be modified according to user emotional status and it remains in its initial states achieved in the training phase.

Thus, each subject was asked to control VE1 and VE2 according to following protocols:

- *VE1: he should move the crane to the load coordination, lift the load, pass it through the wall and release it in the other side of the wall*
- *VE2: he should move the crane to the load coordination, lift the load, pass it to the opposite corner of the ground and release it.*

Both VE1 and VE2 were set to be performed in three different levels of difficulties:

- *Level 1(Easy): normal crane movement speed*
- *Level 2 (Difficult): Fast crane movement speed (hard for precise control)*
- *Level 3 (Difficult): Slow crane movement speed (takes more time and bring early fatigue)*

3.1. Task performance protocol

After initial training for the classifier (Section 2.3), each subject should undertake 6 experiments and each takes 20 min with a 10-min rest interval between them. He first tries with ON state of Control Unit and then, after performing the three experiments, he tries OFF state —

Table 2
Average discrimination ratio for mentioned facial gestures for 10 users.

Gesture index no.	Gesture name	Discrimination ratio %
1	Smiling	94
2	Pulling up right lip corner	91
3	Pulling up left lip corner	88
4	Opening mouth (like to say ‘a’ in ‘apple’)	97
5	Clenching molar teeth	96

with the same initial inference system achieved by the training protocol (Fig. 7).

1– For each of the described conditions, when he performs the requested task correctly, he achieves a positive score.
2– Then, the crane and load position are set randomly for the next trial.
3– After 10 min. rest, the above scenario repeats for the next experiment (different situation) until all the possible VE are experienced by the subject.
4– The average of obtained score within each experiment time is considered as *Expertise Factor* (*EF*) which is an indicator of the subject's achieved expertise in the performed experiment.

Table 3 shows the achieved expertise factors for each subject and their average. It is shown that the average expertise factor has decreased 15.01% when switching from controlled situation to uncontrolled. It is despite of the transfer of learning effect in the first 3 experiments (when the control unit is ON) to the next three experiments. The one-way analysis of variance between group tests (ANOVA) has been performed on Table 3 to find out whether data from several groups have a common mean. That is, to determine whether the groups are actually different in the measured characteristic. The achieved *p-value* is less than 1% which is a strong indication that the experimental conditions are not the same.

Also, according to above discussion about α sub-band entropy capability to reveal the mental disorganization for a given task, the average slope of α sub-band entropy for each subject was calculated in order to measure the cognitive stress over the subjects within the experimental period. The obtained results show that the average entropy slope increased from 0.12 for affective control method to 0.27 (for traditional controlling scheme). It means more cognitive stress was burdened by the subject when the control unit is in OFF state.

It should be noted that, reducing cognitive overload is a very important issue for increasing the performance measures, especially during training phase. The user who is unsatisfied by using the interface, will burden extra cognitive load for undertaking/continuing the requested task. As a result, the performance measures (or the quality of performance) decreases as a consequence of this cognitive overload. Thus, the user might reject to continue the task for longer running period.

According to the filled out questionnaires from users after the experiment, and achieved results, it is found that, by using the proposed method the users felt more comfortable during the performance, and achieved better performance measures by using the proposed affective control compare to traditional controlling scheme's users (affective control: OFF). This is because, the proposed affective human–machine interface is trying to adapt itself to the system's ecology using the user's emotional status, which is mirrored

Fig. 7. A subject while performing the experiment.

Table 3
The expertise factor over different experimental conditions for 10 subjects (p-values for L_1V_1[a] <0.02; $L_2V_1<0.03$; $L_3V_1<0.02$; $L_1V_2<0.05$; $L_2V_2<0.03$ and $L_3V_2<0.03$).

Subject	VE no.	Affective control = ON			Affective control = OFF		
		EF[b] level 1	EF level 2	EF level 3	EF level 1	EF level 2	EF level 3
# 1	1	57	49	41	51	42	32
	2	46	42	32	42	39	28
# 2	1	61	55	48	54	42	36
	2	53	46	38	46	39	31
# 3	1	64	52	42	52	45	33
	2	54	51	42	43	41	31
# 4	1	71	64	58	62	56	49
	2	63	58	52	55	51	43
# 5	1	56	53	44	51	42	31
	2	50	44	41	41	37	35
# 6	1	73	68	59	65	58	46
	2	70	64	57	61	55	44
# 7	1	64	52	42	52	44	33
	2	58	52	41	51	42	33
# 8	1	71	61	57	62	52	46
	2	66	58	51	57	49	38
# 9	1	62	57	54	55	49	43
	2	58	52	49	51	46	33
# 10	1	55	48	44	47	39	37
	2	48	42	38	35	29	29
Average EP for each level		60	53	46	51	45	36
Total average EP for all difficulty levels		53			44		

[a] L_mV_n: difficulty level (m) in virtual environment (n).
[b] EP: expertise factor.

by α band entropy; hence, it leads to decreasing or retaining the α band entropy growing rate. In other words, by using the bi-directional information flow in the system (including: user, computer and controlling scheme) and the presented adaptive methodology, the cognitive overload during performance is reduced. This cognitive overload reduction helps the users to cope with the experiment for long-running period, enhance the user physical performance, and increases the performance measures consequentially.

It should be noted that our proposed method has some privileges compare with traditional methods such as image-based solution. For example, regarding to our electrodes placement in this study, five different facial gestures could be discriminated with high accuracy (93.2%). Also by employing the proposed configuration and Adaptive Neuro-Fuzzy Inference System (ANFIS) eight different facial gestures (including: smiling, pulling up right and left lip corners, frowning, and eyes movement horizontally and vertically) could be discriminated with average accuracy of 96.99% [29]. It should be considered that, these discrimination accuracies were achieved using low computation cost, easy to implement setup, and robust method. However, several problems were observed during experiments using video-based gesture recognition systems, such as drifts, loss of communication, and slow communication rates. For subjects with insufficient muscles control or for the movements with small changes in facial gesture (such as clenching molar teeth) the camera-based methods could become quite ineffective because the sensitivity of them is low. In addition, the video-based gesture recognition requires a high-speed image processing hardware, so the overall cost of system becomes higher than our proposed method. In addition, it requires fixed or predefined backgrounds and camera positions for calibration and is suitable for a small set of gestures [37] compare with our wide range of gestures which could be recognized using only three pairs of electrodes. Indeed, up to now, there is not any proposed study on adaptive video-based gesture recognition system based on users' affective status. Also, if the affective based adaptation method applied to video-based gesture recognition, using at least one electrode to record signals from brain, or heart, or skin is inevitable. In this case the computation cost of the method will be increased more.

The recommended block diagram of an affective controller is depicted in Fig. 8. A human machine interface can be divided into physical layer, cognitive level, and its interface which controls the interaction process between human and machine. In the previous studies, the interface adapts itself with the physical performance indices eliciting from the physical level. But, the effect of the physical layer on the cognitive layer are mostly ignored or missed in the interface designing. The cognitive level could modify the physical level and the system's performance as well. So, the interface (control system) cannot cope with changes in the system, if it only be adapted to physical layer status, because the physical interactions within the interface have tight correlations with the affective status in the cognitive level.

Here, we propose that, the changes in cognitive level which are mirrored by affective measures could be considered as an important factor for modifying the interface. In Sections 2.6 and 2.7 the described method has been employed as our approach to HMI designing for a virtual forearm prosthesis and crane, respectively.

4. Conclusion and future works

This paper integrates affective computing concept and approach into the conventional virtual reality training platform. In order to test and justify this approach, two virtual crane training environments which can be controlled using commands extracted from facial gestures (facial bioelectric-signals) and is capable of lifting up load/materials at virtual construction sites was created. A preliminary experiment implemented by inviting 10 human subjects and data was collected. These signals were captured when facial gestures were generated by users. By using signals from Channels 1 and 3, the classifier average discrimination ratio was 93.4% for discriminating between 5 mentioned facial gestures.

In this study, it was also shown that Channel 2 could mirror user emotional state. The entropy of this channel rises up when the level of difficulties is increased. Also, it should be noted that the generation source of the mentioned gestures are far enough from Channel 2 and thus this channel can be responsible to extract affective measure without facial gestures interference.

We have proposed a real-time affective and adaptable controlling scheme and implement on a human machine interface for controlling a virtual crane on two different operation sites. This interface could adapt and modify itself based on the affective cues extracted from the facial EEG α sub-band signals. Also, by using our proposed control scheme the average entropy slope of the α band is lower comparing to uncontrolled one. Thus, this bi-directional modification (the machine changes the human α band entropy; the human modifies the machine's control system) could lead to a self-organizing system as a whole. The results showed that, the proposed method could be used over a long-running period, different operation sites, and different levels of difficulty with high performance measures. According to filled out questionnaires, the system's users were more satisfied with the affective interface compared to traditional interface.

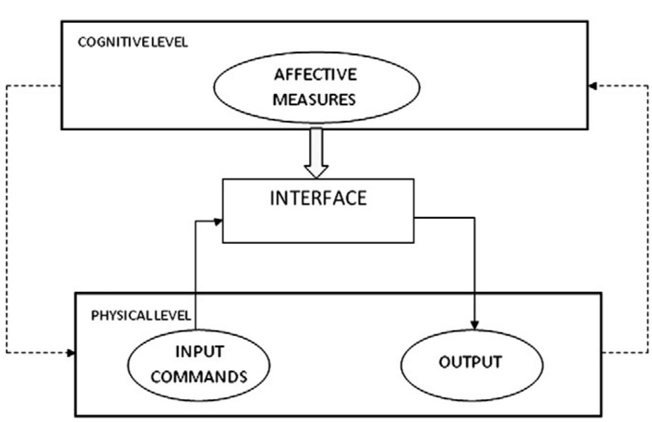

Fig. 8. The affective controller designing schematic. Dashed lines shows indirect influence of one part to another.

This real-time control scheme could be tailored for many applications for disabled users who their facial muscles have enough strength to generate mentioned gestures. The gestures are easy to produce respect to other facial gestures. In addition, all electrodes could be mounted on a sport headband for easy installation. Thus, this method could be a window to get the disabled users back to their activities of daily life and be more involved in the social activities by providing them the required confidence and dexterity. The presented interface could be also used in rehabilitation phase for people who are suffering from facial muscles impediment. These people could use this system as a training environment for practicing to generate facial gestures and it could help them out to reactive their weak muscles. Furthermore, the interface could increase the context's awareness in working sites and prevent possible dangers might come for the user and its working environment [38]. This can be done by detecting his/her feeling and mental states and then performing proper action (help, alarm, or stop for example) for pursuing and accomplishing the task. Also, the proposed HMI has potential to be applicable on more complicated working sites. In addition, this controlling scheme is a good solution when *usability engineering* and *user-centered designing* approaches are concerned [39] and could be use as an assistive strategy which brings effectiveness, efficiency, and satisfaction for the designed HMI. Thus, in our future work, we are going to expand our experiment on a more complicated virtual environments and tasks, and on disabled population who could be the main users of our system.

Acknowledgements

We would like to appreciate the help from Rui Wang from the University of Sydney and eager volunteer participants in the school of Biomedical Engineering, Science and Research Branch, Islamic Azad University.

References

[1] R.J. Seidel, P.R. Chatelier, Virtual Reality Training's Future?: Perspectives on Virtual Reality and Related Emerging Technologies, Plenum Press, New York, 1997.
[2] V.G. Cuqlock-Knopp, C.A. Wilkins, W.S. Torgerson, Multiple cue probability learning and the design of information displays for multiple tasks, in: D.L. Damos (Ed.), Multiple-Task Performance, Taylor&Francis, London, 1991.
[3] M.J. Schuemie, P. Van Der Straaten, M. Merel Krijn, C. Van Der Mast, Research on presence in virtual reality: a survey, Cyberpsychology & Behavior 4 (2001) 183–201.
[4] J. Psotka, Immersive training systems: virtual reality and education and training, Instructional Science 23 (1995) 405–431.
[5] R. Schroeder, Learning from virtual reality applications in education, Virtual Reality 1 (1995) 33–40.
[6] D.P. Mahoney, Defensive driving, Computer Graphics World 20 (1997) 71.
[7] J.P. Bliss, P.D. Tidwell, M.A. Guest, The effectiveness of virtual reality for administering spatial navigation training to firefighters, Teleoperators and Virtual Environments 6 (1997) 73–86.
[8] G. Lintern, S.N. Roscoe, J.M. Koonce, L.D. Segal, Transfer of landing skills in beginning flight simulation, Human Factors 32 (1990) 319–327.
[9] J.W. Regian, W.L. Shebilske, J.M. Monk, Virtual Reality: an instructional medium for visual–spatial tasks, The Journal of Communication 42 (1992) 136–149.
[10] L.E. Magee, Virtual Reality Simulator (VRS) for training ship handling skills, in: R.J. Seidel, P.R. Chantelier (Eds.), Virtual Reality, Training' s Future? Plenum, New York Plenum, 1997, pp. 19–29.
[11] S.L. Goldberg, B.W. Knerr, Collective training in virtual environments: exploring performance requirements for dismounted soldier simulation, in: R.J. Seidel, P.R. Chantelier (Eds.), Virtual Reality, Training' s Future? Plenum, New York, 1997, pp. 41–51.
[12] T.W. Mastaglio, R. Callahan, A large-scale complex virtual environment for team training, Computer 7 (1995) 49–56.
[13] R.B. Loftin, R.T. Savely, R. Benedetti, C. Culbert, L. Pusch, R. Jones, et al., Virtual environment technology in training: results from the Hubble space telescope mission, in: R.J. Seidel, P.R. Chantelier (Eds.), Virtual Reality, Training' s Future? Plenum, New York, 1997, pp. 93–103.
[14] Anon, Immersive VR tests best, CyberEdge Journal 4 (1994).

[15] N. Adams, A Study of The Effectiveness of Using Virtual Reality to Orient Line Workers in a Manufacturing Environment, DePaul University, Chicago, IL, 1996.
[16] B.H. Wilson, R.R. Mourant, M. Li, W. Xu, A virtual environment for training overhead crane operators: real-time implementation, IIE Transactions 30 (1998) 589–595.
[17] Anon, VR in industrial training, Virtual Reality 5 (1996) 31–33.
[18] R.R. Wakefield, J.B. O'Brien, J. Perng, Development and application of real-time 3D simulators in construction training, Paper presented at the the International Conference on Construction Training, Hong Kong, 1996.
[19] Simlog, Podcast retrieved from http://www.simlog.com/index.html., Accessed on 16 March 2010.
[20] R. Fernandez, Stochastic modelling of physiological signals with hidden markov models: a step toward frustration detection in human–computer interfaces. M.Sc. Thesis, MIT (1997).
[21] J. Healey, R.W. Picard, SmartCar: Detecting Driver Stress, Paper presented at the ICPR.00, Barcelona, Spain, 2000.
[22] F. Nasoz, C.L. Lisetti, K. Alvarez, N. Finkelstein, Emotion recognition from physiological signals for user modelling of affect, Paper presented at the 3rd Workshop on Affective and Attitude user Modeling, Pittsburgh, PA, 2003.
[23] R.W. Picard, Affective Computing, MIT Press Cambridge, MA, 1997.
[24] P. Brézillon, J. Brézillon, J. Pomerol, Decision making at a crossroad: a negotiation of contexts, Joint International Conference on Computing and Decision Making in Civil and Building Engineering, Montréal, Canada, 2006.
[25] L. Ang, E. Belen, R.A. Bernardo, E.R. Boongaling, G.H. Briones, J. Coronel, Facial expression recognition through pattern analysis of facial movements utilizing Electromyogram sensors, IEEE Tencon (2004) 600–603.
[26] S. Mahlke, M. Minge, Emotions and EMG measures of facial muscles in interactive contexts, Conf. in Human factor in Eng., Montreal, Canada, 2006.
[27] P. Neimenlehto, M. Juhola, V. Surakka, Detection of electromyographic signal from facial muscles with neural networks, ACM Transaction of Applied Perception 3 (2006) 48–61.
[28] I. Mohammad Rezazadeh, X. Wang, R. Wang, M. Firoozabadi, Toward affective handsfree Human–Machine Interface approach in virtual environment-based equipment operation training, 9th International Conference on Construction Applications of Virtual Reality, Sydney, Australia, 2009.
[29] I. Mohammad Rezazadeh, M. Firoozabadi, H. Hu, M.R. Hashemi Golpayegani, Quantative evaluation of the efficiency of facial biopotential signals based on forehead three-channel electrode placement for facial gesture recognition applicable in a human–machine interface, Iranian Journal of Medical Physics 7 (2010) 65–79.
[30] S.M.P. Firoozabadi, M.A. Oskoei, H. Hu, A Human–Computer Interface based on forehead multichannel biosignals to control a virtual wheelchair, Paper presented at the ICBME08, Tehran, Iran, 2008.
[31] M.A. Oskoei, H. Hu, Myoelectric control systems — a survey, Journal of Biomedical Signal Processing and Control 2 (2007) 275–294.
[32] V. Moertini, Introduction to five data clustering algorithm, Integral 7 (2002).
[33] A. Priyona, M. Ridwan, A. Alias, R. Atiq, O.K. Rahmat, A. Hassan, M. Ali, Generation of fuzzy rules with subtractive clustering, Universiti Teknologi Malaysia, Jurnal Teknologi 43 (2003) 143–153.
[34] N. Jausˇovec, N. Jausˇovec, I. Gerlic, The influence of Mozart's music on brain activity in the process of learning, Clinical Neurophysiology 117 (2006) 2703–2714.
[35] S. Aydin, H. Saraog, S. Kara, Log energy entropy-based EEG classification with multilayer neural networks in seizure, Annals of Biomedical Engineering (2009) 2626–2630.
[36] Mathworks. www.mathworks.com, Accessed on 16 March 2010.
[37] M. Ahsan, M. Ibrahimy, O. Khalifa, EMG signal classification for Human Computer Interaction: a review, European Journal of Scientific Research 33 (2009) 480–501.
[38] J. Teizer, F. Bosche, C. Caldas, T. Haas, Real-time spatial detection and tracking of resources in a construction cnvironment, Joint International Conference on Computing and Decision Making in Civil and Building Engineering, Montréal, Canada, 2006.
[39] J. Abraham, I. Parmee, User-centric evolutionary design systems — the visualisation of emerging multi-objective design information, Xth International Conference on Computing in Civil and Building Engineering, Weimar, Germany, 2004.
[40] Biopac, http://www.biopac.com 2010, Accessed on 16 March 2010.