The International Journal of Virtual Reality, 2008, 7(1):53-66

How and Why Affordable Virtual Reality Shapes the Future of Education

Barnabas Takacs[1]

[1]*VHI MTA SZTAKI, Budapest, Hungary/Digital Elite Inc, Los Angeles, CA, USA*

***Abstract*—This paper presents an analysis of how and why a new generation of virtual reality peripherals affect virtual reality in general, and its uses for education in particular. We describe a method to bring psychological principles, such as positive learning and cognitive dynamics into the best utilization of VR in the classroom and present the technique of emotional modulation aimed to help students in absorbing the information they are presented with more efficiently. Edutainment is made available by a range of new devices to support different modalities of interaction. We present taxonomy of low-cost, yet advanced IO devices and show how the future of education will be shaped by these factors. Finally, we describe an architecture called the *Virtual Human Interface* used as a generalized framework for building new educational content and conclude with showing practical examples and psychiatric evaluation of how our system was used to address real-life applications.**

***Index Terms*—Animated agents, edutainment, game-based interfaces, portable virtual-reality, virtual human interface.**

I. INTRODUCTION

Virtual reality offers students experiences they otherwise could not ordinarily have [30, 45]. During the past decades a multitude of VR-solutions in the classroom have been developed where typically a library of inexpensive and independent *VR* programs are used to educate, clarify, and reinforce each subject studied. However, due to the high cost of hardware such as head mounted displays (HMDs), data gloves, tracking devices, CAVES and advanced projection techniques including multi-screen tiled displays [46] and multiple head mounted displays [47], all needed to deliver fully immersive experiences research results have been severly limited from entering the classroom and the lives of every-day students. That “landscape” has recently changed by the broad availability of gaming devices and computing platforms. Therefore it can be argued, that a single and unified framework with advanced input/output capabilities may well suite the needs of many classroom applications of the future and can deliver turly immersive and effective learning experiences using not more than a laptop and a few novel peripherals attached to it. However in orderr to mimic the quality of everyday human dialogues that take place between teachers and their students’ future virtual learning environments must thrive to achieve the ability to *modulate the emotions* of a student in a personalized manner. One approach to this problem combines the *benefits of high visual fidelity virtual-reality environments***,** complete with *animated human agents* while delivering interactive experiences in a controlled and repeatable fashion.

From a computer science perspective to achieve the above requires the ability to create believable *digital tutors* with photo-real faces capable of expressing the finest shades of emotions in a controllable manner. The same system must also be able to *read the users’ emotional reactions* and adapt the behavior of the digital human accordingly in real-time. During our current research, we implemented this concept as a step towards creating a novel VR-based educational solution builds upon many years of *interdisciplinary research* to create a *closed-loop model of interaction* whereas the students’ internal state (emotion, level of attention, etc.) is constantly monitored and driven directly by an animated character with the purpose of creating emotional bonding. *This emotional bonding then acts as a catalyst to help turning information into knowledge*. In other words, our *affective intelligence* interface draws on emotions to help its users in the learning process by intelligently tailoring its workload and constantly adapting its presentation strategies. One of the key points of this research is its notion of blending animation and perception is into a single and unified framework while building upon concise *models of human face-to-face communication.* We draw on principles of psychology and cognitive dynamics driven by algorithmically controlled visual, auditory and olfactory cues. The acute capability of students to read various social cues from faces offers the unique opportunity to *use high fidelity digital faces as a prime interface.* Such a communication layer provides thousands of parallel channels readily understandable even by the least educated or youngest members of our society.

This paper is organized as follows: In Section we describe how the technique of positive learning and cognitive dynamics can be incorporated into the learning process and help increase the efficiency of VR-based learning and training. Our discussion is followed, in Section , by a description of low-cost input and output devices currently available for building novel *VR* applications for the classrooms of the future. In Section we present one such system, called the *Virtual Human Interface* [5], which provides seamless integration of these capabilities and provides a general platform for many practical applications. Section discusses our methodology for emotional modulation and Section describes practical uses, including our *BabyTeach* application as well as evaluation of these tools using psychiatric measures. Finally in Section we present our conclusion.

Manuscript received on April 27, 2006.
E-Mail: BTakacs@digitalCustom.com.

II. EDUTAINMENT: THE ROLE OF POSITIVE LEARNING AND COGNITIVE DYNAMICS

“Information consumes the attention of its recipients, hence the wealth of information creates a poverty of attention, and the need to allocate that attention efficiently among the sources that might consume it.” (Herb Simon, 1995 Nobel Prize Winner).

Virtual Reality offers one of the most powerful means to fully control a student’s attention and therefore maximize his or her ability to learn buy absorbing and later remembering the information presented. While the process of interaction and a game-like personalized point of view in some cases in alone itself is enough to create positive bonding, research needs to focus on evoking emotional reactions by other, subtler means. Emotion is an essential element in human communication. When students interact in virtual world they exhibit emotions, which need to be mirrored and matched by the system’s ability to generate and display a range of emotions autonomously in response. This feature becomes a critical factor of success since emotions also play a vital role in how we learn or more precisely how we absorb the wealth of information surrounding us.

Emotional modulation is a technique that helps transforming *information* to *knowledge* using our own emotions as a catalyst. Our world is filled with *data*. Data in itself is not a piece of information, but rather it holds the potential of becoming one. For data to become useful *information* it first must go through a verification or editorial stage that guarantees and qualifies its value in a given context. The purpose of learning in general, and *VR*-based education in particular is to turn this information into *knowledge*. However, as the amount of information grows the brain finds it increasingly difficult to keep up with this ever-increasing demand and function properly. Thus, one of the greatest challenges of today’s modern society is how we can enable people to increase the efficiency of the learning process itself in general, and their own learning capacity in particular.

The *link between emotional processing and the capacity of the brain to absorb and memorize information efficiently has long been discovered*. We created a complex system, called the *Virtual Human Interface* in the form of a digital person (See Section 4 for more details). The basic idea of emotional modulation stems from a simple every day observation. When in a good mood, we are generally more susceptible to information presented in a positive fashion, and when we are sad or feeling down, we prefer things presented in a more subdued manner. Based on this observation we defined an *Artificial Emotion Space* (*AES*) for our digital tutor who acts as a mediator throughout the learning exercises and employs many strategies to link the real person and the synthetic character together. As an example, the digital human might exhibit layers of emotion that coincide with or in other cases exactly oppose the student’s mood, thus creating the illusion of compassion or “playing the devil’s advocate” respectively. In essence, we argue, that *virtual humans are a key enabling factor to achieve emotional engagement*. They unlock the learning potential in an individual by creating an emotional link between user and digital character. Face-to-face communication with a digital interactive virtual human is therefore one of the most powerful methods for providing personalized and highly efficient information exchange. In the *closed-loop* model of interaction the digital character is capable of perceiving the moods and momentary emotions of the student and express their own feelings in a way to reflect empathy and understanding. The virtual reality system gauges the user’s level of attention, fatigue or emotional state by actively *prompting* them and looking for an appropriate reaction. This process works much the same way as it does between two people taking part in a *dialogue*.

The basic model of human computer interaction applicable to *VR*-based education systems first assumes is that there are minimally two separate *agents* involved, one human and one machine. They are physically separated, but are able to exchange information through a number of information channels as schematically shown in Fig. 1. On the side of the human user there are two basic processes involved, namely *Perception* and *Control*. The *Perceptive process* may be further divided into a) Human *Input Channels* (HIC) and b) *Computer Output Media* (COM) while similarly the *Control process*, comprises of c) *Human Output Channels* (HOC) and d) *Computer Input Modalities* (CIM) [28]. Within both of the agents, a *cognitive* or computational component can be identified, which processes the incoming input information and prepares the output. Also, at this intermediate cognitive level, intentional parameters will influence the processing, either implicitly, such as by design, in the case of non-intelligent agents, or explicitly, as in humans or in more sophisticated agents containing an explicit representation of *“goals and beliefs”*.

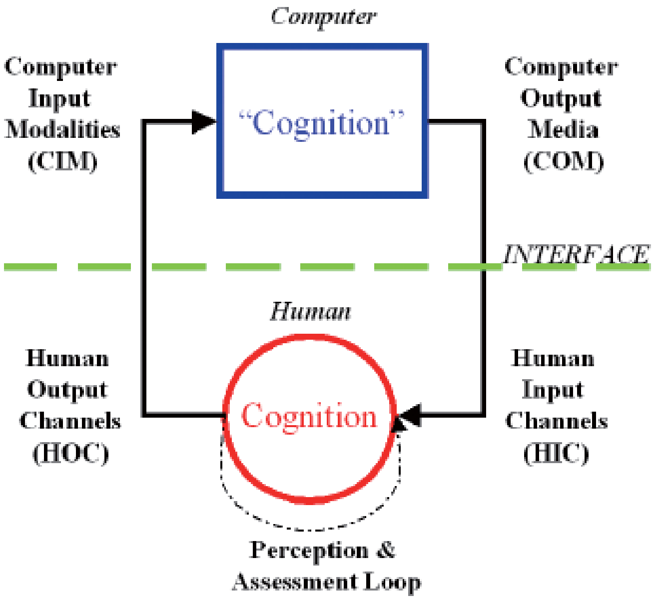


Fig. 1. Model of human-computer interaction (based on [28]).

For a multi-modal *VR* system, several aspects of intelligent agents are especially interesting, e.g. intentional behavior, believe, autonomy, negotiating capabilities, etc. All of them will significantly contribute to an intelligent user interface, which is able to interpret the students' actions, decide for itself the best way to present learning information and cope with the several parallel input modalities existing in the system.

To *create the emotional engagement* that is required to actively modulate our student's emotional and motivational state, our virtual human needs to exhibit high levels of control of its own *body language.* Meta-communication is implemented in the form of subtle *non-verbal signals* that encourage or oppose the students' actions in a given virtual situation, and thus guide him or her to associate positive feelings with the piece of information received. These layers of communication occur at many different subliminal levels and support verbal content by means of facial animation, body movements, and hand gestures. The final result is a process that by using photo-real digital humans and their synthetic 3D environments engages the user in a series of emotional responses that in turn opens the channel to "engrave" knowledge and/or help modify behavior.

The *tool of emotional modulation* and the algorithms we developed not only serve as aids to the learning process, but in some circumstances they affect the very basis of how learning takes place. To explain this better let us consider the classical learning paradigm with its roots in the principal idea of "*learning by failure*". The emotional states of a student here interweave with the *cognitive dynamics of the learning process* itself [17]. In many educational applications of virtual reality, however the above model is simply non-functional. As an example, for a broad range of *VR*-tools designed for special education as well as in rehabilitation the above model needs to be partly revised as the very *experience of failure* may often lead to such a strong negative affect that it *prevents the child from further participation* and efficient learning. Therefore a more gentle method based on positive reinforcement is needed. We offer a solution to this problem by providing visual feedback in the form of encouraging facial displays, expressions and other non-verbal signs that help the student avoid the mistakes they are about to make.

From the point of view of creating practical applications our *VR*-architecture is mostly concerned with what is known as *affective states of learning* (like interest, boredom, excitement, confusion, fatigue etc.). These telltale signs offer brief insight into a students' internal state as well as their relations to the material or information being presented at the time. The affective states of learning are accompanied by different patterns of facial expressions, eye-gaze, had nod, hand movement, gestures and body posture together called *surface level behaviors* which can be interpreted in the context of an application as indicators of the person being "*On-task*" i.e. paying attention, or "*Off- task*" in all cases otherwise (Figure 2). As an example, the *direction of eye gaze* and *head orientation*, are prime indicators of a learner's focus of attention. In an *On-task* state, the focus of attention is mainly toward the problem the student is working on, whereas in an *off-task* state the eye-gaze might wander off. One of the many involuntary reactions a person exhibits is *pupil dilation*, which is also known to have high correlation with the level if interest in a topic. Specifically, when a student finds a particular subject fascinating, his or her pupils are unconsciously dilated, as to opening up to the speaker, in the virtual tutor in our case. Other signs of interest (being On-task) are approving head nods, facial actions like smile, tightening of the eyelids while concentrating, eyes widening or raising eyebrows. On the other hand, i.e. being Off-task boredom may be detected from the withdrawal expressed by the diminished pupils of the person, head shakes, the lowering of eyebrows, nose wrinkling and depressed lower lip corners. These non-verbal signs can be measured directly from sensor output, such as a 6DOF head tracker, gaze tracker and data glove, or inferred from visual and other cues, like face tracking and gesture analysis by means of computer vision or the levels of stress and emotional engagement from biofeedback, etc.

The second important psychological aspect of human learning and adaptation in *VR* how the *students' emotional state interweaves with the cognitive dynamics of the learning process* [14,17]. This is demonstrated in the circular arrangement in Figure 3. On the horizontal axis the emotional state is plotted, and on the vertical axis the user's relationship to the learning that takes place in that moment is represented. Through the course of interaction the learner's state-of-mind moves continuously in this space evolving in a circular pattern. Ideally the process begins in quadrant I. or II; whereas he or she might be curious and fascinated about a new topic of interest (quadrant I) or he might be puzzled and motivated to reduce confusion (quadrant II). In either case, this state of mind will be in the top half of the space whenever the system's main focus is on constructing or testing knowledge. *Movements happen in this space as learning and interaction proceeds.* As an example, the student may get an idea about how to implement a solution. When that fails to work, he realizes that his original idea needs to be deconstructed and analyzed. At this point it is rather common for a user to move to quadrant III. where emotions may be negative. The focus on learning turns into the process of eliminating misconceptions. But as time goes on, the user consolidates his/her knowledge and gradually moves to quadrant IV. from where a fresh idea propels the user back to quadrant I. again [17].

Observation	**ON-Task**	**OFF-Task**
Posture	Leaning forward, sitting upright	Slumping on the chair, Fidgeting
Eye-gaze	Looking towards the problem	Looking everywhere else
Facial Expressions	Eyes tightening, widening, raising eyebrows, smile	Lowering eyebrow, lower-lip depression, winkling nose
Head nod and shake	Up/Down head nod	Head shake sideways
Hand Movement	Typing, clicking, mouse, gestures	Hands not on mouse/keyboard

Fig. 2. Surface level behaviors as indicators of a student's internal state (based on [14]).

While there are clearly different individual learning characteristics and preferences involved in any interactive process the above models allow for *representing those person-dependent traits in a unified and parametric manner.*

To address the needs of VR-based learning we have been developing a system aimed to integrate the above cues into a seamless process of interaction and bi-directional information exchange. The practical implementation of this *closed-loop model* relies on the newly developed availablity of novel input and output devices. We now turn our attention to what makes this revolution possible. In our assessment virtual reality has finally reached a point where it stands at the threshold of entering our everyday lives. As a result of this process *VR* will leave the realm of laboratory equipment and military simulators to enter the world of education [32,45], health care [34,52,53,54] and - of course - personalized entertainment [48].

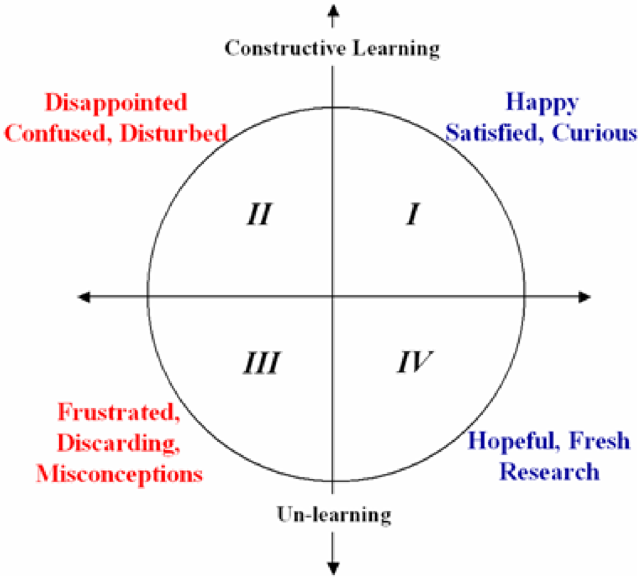


Fig. 3. The relationship of emotions to the learning process [17].

III. RESOURCES FOR VIRTUAL REALITY IN EDUCATION

In the decades of its existence the term “virtual reality” has become a household name. It was featured in many movies and entertainment parks, and practically everybody is at least aware of its existence. Yet, at the same time, only a very small percentage of the general public has ever had the pleasure to experience first hand what it is capable of delivering. The primary reason for this contradiction was, for many years, the high cost of computing power and I/O devices that rendered these systems unavailable to the average individual.

Today the sudden availability of low-cost and consumer-grade input and output devices which may be used in virtual reality class-rooms comes at no surprise. The very existence of these novel devices has spurred the opportunity and created a business landscape that encourages the development of virtual reality applications for the classrooms of the future. Small companies and research groups are now allowed to enter this field without the heavy investment of capital required before. This process eventually fuels creativity and leads to applications never before existing or even though to be possible.

Fig. 4 shows a list of resources for *INPUT* devices that may be used in virtual reality classrooms of the future. Besides the classical input devices, such as the keyboard and mouse, *programmable switches* [18,25,41,49] and *MIDI* [19] devices offer a new way of interaction.

In the *visual domain* a number of broadly available robotic web-cameras offer a new possibility to track and analyze, as well as record the activities of the students’ while participating in a VR classroom exercise [2,6]. Such analysis may include *facial tracking* and the *recognition of internal state* from expressions and displays of spontaneous reactions in addition to computer vision tools that identify key gestures and body postures. A special kind of visual input, eye trackers [9,36,42] can be used to reveal where the student is looking on the screen at any given moment and subsequently identify objects and regions of interest for interaction purposes. The closed-loop model we described above allows for suing this information to change the presented material in a manner to best ensure that all critical pieces of knowledge are presented for learning. Similarly, *desktop motion capture* systems [21,22] that use *IR* illumination to track the 3D trajectory of arbitrary reflective markers attached to objects, props or even body parts offer new means of linking the outside world with the virtual environment. Finally, *immersive video* [26,38,39,40,48] may be used to record and interactively render experiments where virtual tele-presence further enhances the emotional engagement and experience levels of the learning process.

Modality	**Device Category**	**Resources**
I/O	Programmable switches	[18,25,41,49]
	MIDI	[19]
Visual	Robotic cameras	[2,26]
	Eye tracking	[9,36,42]
	Desktop motion capture	[21]
	Panoramic video	[26,28,39,40]
Tactile	Touch screen	[15]
Motion	Trackers	[7,8,13,21,22,49]
	Data glove	[8]
Stress	Biofeedback	[38]
Physical	Acceleration, local forces, air flow, gas pressure, etc.	[44, 49]

Fig. 4. Review of available low-cost *INPUT* devices and solutions for VR-based education systems.

The visual experience may be augmented with a number of other sensory events including *tactile* interaction and feedback with the help of *touch screens* [15] and motion analysis via head trackers [7,8,13,22] and data gloves [8]. In terms of physical interaction the *levels of stress and emotional reactions* may also be recognized during a learning session with the help of *biofeedback* devices [38] and finally, for demonstration purposes or direct control of educational material measuring *devices for various physical characteristics and phenomena* [44,49] may also be incorporated into virtual reality-based educational programs.

Having discussed the input modalities, let us now turn our attention to *OUTPUT* devices and modalities that may revolutionarize how classrooms present virtual reality in the future. Fig. 5 shows a list of resources for *output* devices that may be used to increase the efficiency of learning by emotional engagement and higher-fidelity. Starting with visual output, which is perhaps the most common and accepted form of immersion and engagement the functionality of *2D computer monitors* may increasingly be replaced with portable palm-projectors [20] and mobile devices where each student can receive personalized educational content on their 3G phone or PDA [34].

Modality	**Device Category**	**Resources**
Visual	2D displays, 3G mobile devices	[5,34]
	Palm projectors	[20]
	3D stereo projection with glasses	[4]
	Head Mounted Displays (HMD)	[7]
	Portable CAVE	[10,23,34,43]
	Holographic Monitors	[12]
Acoustic	Multi-channel surround sound	[5,34]
Tactile	Force and Vibro-Feedback	[1, 27,29]
Olfactory	Digital smell	[24,37]

Fig. 5. Review of available low-cost *OUTPUT* devices for VR-based education systems.

Besides using palm-size devices other aspects of projection technology have also advanced to the point where a single computer (instead of a cluster of synchronized computers) can generate four independent views and create a low-cost *CAVE system* for virtual presence and demonstration purposes [10,43]. On the pother hand, for representing *3D stereoscopic projection* in the classroom stereo-pairs of images or video can be now projected using a single and affordable 3D projector (instead of playing with carefully calibrated and aligned, thus very sensitive dual projection system) and each member of the class may enjoy full viewing experience [4]. Another way to present stereo images to a student is using *head mounted displays* or HMDs, whereas a new generation of OLED-based consumer-priced devices are now available with built-in head trackers [7]. Finally, *holographic monitors* [12] present *true 3D information* for multiple observers and from multiple viewing angles at the same time without the use of any special glasses or devices, making them ideal for a number of future applications.

The *acoustic output* of VR education solutions may be based on surround sound systems delivering 5.1 or 7.1 experience. As an example the most advanced 7.1 multichannel sound technology features left, right, center, left surround, right surround, left rear, right rear positions and also have one channel for low frequency effects (LFE) which is usually sent to a subwoofer. Even lower frequency domains are perceived as vibrations and essentially stimulate the student's *tactile sensory* system. Specifically, a number of bass-shakers and vibration devices [1,27]. Finally, force feedback [29] has also become a useful and affordable tool to explore physical processes such as forces between molecules.

To conclude this section we mention a new direction in output devices that are designed to stimulate or senses in yet another way. *Olfactory* interaction employs *digitally dispersed smell* [24,37]. For a more detailed taxonomy of these input-output devices as well as their use in virtual-reality rehabilitation, the reader is referred to [34].

IV. VIRTUAL HUMAN INTERFACE PLATFORM

To demonstrate the concept of closed-loop emotional modulation in a virtual-reality-based educational context we have created a novel application, called *BabyTeach.* To achieve our goals we used an advanced interactive virtual reality platform, called the *Virtual Human Interface* or *VHI* [5,32,33,34,35], we have been developing for the past years. In the sections below we briefly review our system and its components.

The principal operation of the *VHI* real-time interactive system is demonstrated in Fig. 6. It employs a digital animated face (a virtual child) as its principal agent "who" appears on a display in front of the children. A built in camera and microphone deliver images and audio to the *Perception* subsystem which detects and analyzes faces as they appear in front of the computer, performs visual analysis of the scene, and processes the input in order to derive a symbolic representation of the outside world in a format that may be fed the *Behavior* module. The purpose of this module is to control the real-time response of the virtual child via the tools of artificial intelligence, reactive behavior and modeling of the user's own internal state. The generated response of the virtual child then manifests itself in the form of animations, facial expressions, body language or other scripted behavior overlaid on the information content delivered by the virtual character. This endless loop of perception feeds the process of interaction and is designed to mimic a dialogue that takes place between the two parties involved (user and computer).

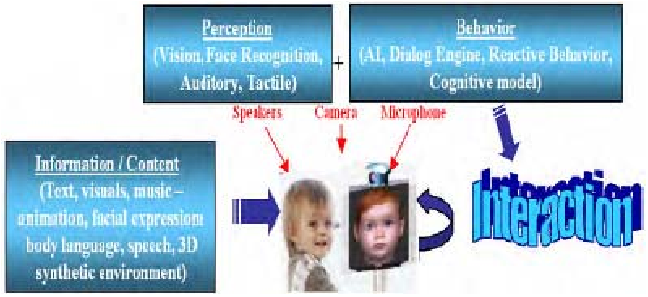


Fig. 6. Functional diagram of the Virtual Human Interface system [5,32,33,34,35].

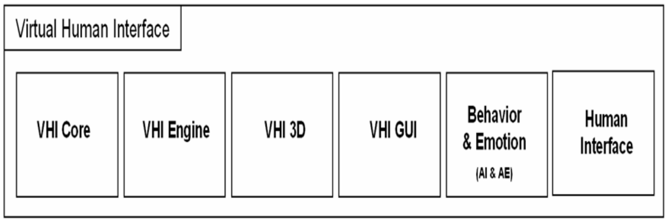


Fig. 7. Basic modules of the VHI architecture.

The above functionality was implemented through the generic overall architecture as shown in Fig. 7 through Fig. 9, respectively. Fig. 7 and Fig. 8 explain the principal modules of the *VHI* system. These are as follows:

- VHI Core: Responsible for constructing, animating and deforming the virtual human. This module implements bone systems, forward and inverse kinematics, muscle systems, skin and shape deformers, hair, etc.)
- *VHI Engine*: Controls the dynamics, motion libraries, physics and low-level reactions of the virtual human, its interactions with synthetic objects and its animated 3D environment.
- *VHI 3D*: Interface layer for 3D graphics standards (Open GL and DirectX) and external graphic environments. The functionality of the VHI is accessible via its own real-time display or from high level professional animation packages.
- *VHI GUI*: Graphical User Interface control layer. The VHI supports platform independent user interfaces and widgets that can control the operation of the system via the scripting layer. Besides its own Windows-based GUI, custom built GUI's may be created using TCL/TK, LUA or HTML + Java scripts.
- *Behavior and Emotion*: Controls the reactions, behavior and internal state, such as emotion of the virtual human as a function of the interaction process and the information being presented. This level contains implementation of the reactive and affective intelligence modules as well as a cognitive engine.
- *Human Interface*: The perceptive capabilities of the virtual human are implemented via an array of visual, acoustic and tactile sensors. These sensors provide the system with a continuous data stream describing the user's reactions and internal state. They are processed with the help of specialized algorithms (e.g. face finder, facial feature analysis, etc.) in order to derive a symbolic representation of the students' world as it applies to the interaction process.

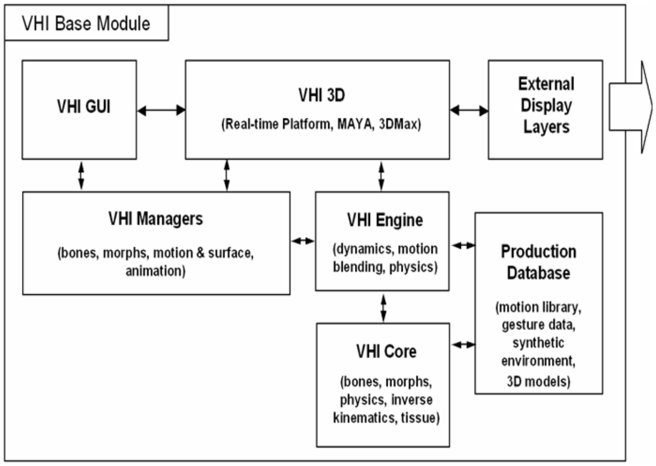


Fig. 8. Functional diagram of the VHI base module.

To further elaborate on the operation and relationship of the above modules Fig. 8 details the functionality of the *VHI Base module*. It contains three additional elements (in addition to the VHI Core, VHI Engine, VHI 3D, and VHI GUI), namely the *VHI Managers* that *schedule and prioritize* how the internal resources (bones, morphs, motion patterns, surface deformers and animation data) are to be used to execute a given task. The *VHI Managers* control the *VHI Engine* directly which in turn accesses information, such as available motions, facial expression repertoire, gesture data, elements of the synthetic environment, and 3D models all stored in the *Production Database*. Finally, driven by *VHI 3D* unit directly an additional module, the *External Display Layer* serves as the primary output interface to the outside world. This latter module implements various *broadcast strategies* that allow the *VHI* to be used not only by being physically in front of the computer screen, but from remote locations via the *Internet* or even digital networks such as *mobile phones*.

The *VHI Base* module is connected to the *Behavior and Emotion* control unit as shown in Figure 9. The complex motion of the virtual human is created by blending four independent sources of motion data. The first such source is the *motion and gesture library* where patterns of human motions (captured or hand animated) are stored and "acted out" as part of the application content. These basic motion patterns are recorded prior to the interaction process and are fixed during production. To appear more natural in a given application the virtual human may use multiple motion patterns to carry out the same task. These motion patterns are randomly selected from a sub-set of available motion data but still result in a deterministic behavior. The *behavior engine* is responsible for prioritizing possible actions and making this selection. When locomotion of the digital human character is involved the *path* of the character need to be *transformed* in real-time so that it reaches its final destination. The second source of motion data is the *artificial emotion space* module that encodes the current emotion and mood of the character and modifies the output of the behavior modules. The third such modifier, called *reactive behaviors*, is driven by the user's interaction (e.g. touching, speaking from behind the character, etc.). This is implemented via short scripts feeding into the *communicative intelligence module* the role of which is to generate acknowledgement signals (head nods, looking into the eye of the user, encouraging smiles, etc.) in keeping with the rules of a closed-loop-dialogue. All the above four sources of motion data are blended together by the unit termed *Temporal Disc Controllers*.

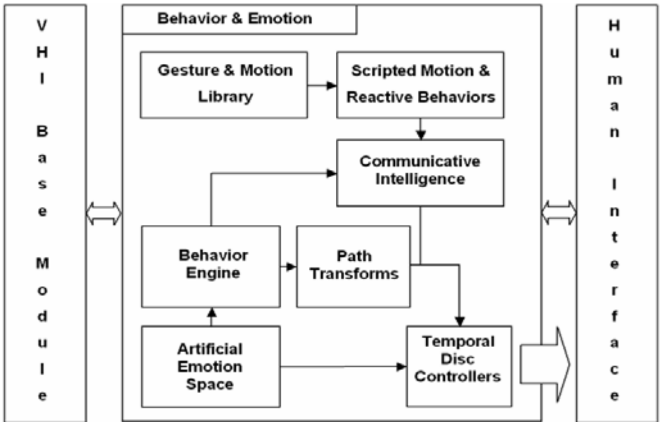


Fig. 9. Functional diagram of the Behavior and emotion module in the VHI system.

The last principal building block of the *VHI* implementation is the *Human Interface* module shown in Fig. 10. The purpose of this module is to implement the perceptive capabilities of the virtual human by extracting information from various sensors attached to the *VHI* system. The perceptive capabilities implemented are similar to those of the basic human senses, namely vision, hearing and touching. The particular external devices used here are connected to the *VHI* system via a set of *device drivers*. The raw *input data* processed by specialized algorithms is categorized into four distinct groups. The most important of these data streams are the images received from one or more video cameras. These are processed by a *computer vision* module, which primarily looks for faces, facial features or other telltale signs of the user's presence, his or her actions or internal state. The second set of algorithms employs *signal processing and voice analysis* to process data captured by the microphones. *Tactile interfaces* are implemented via touch screens, while similarly *physical devices* with no direct perceptual meaning are simply mapped further as raw data. Irrespective what channel the information is arriving from, the purpose of processing is to "recognize" and in some aspect "understand" the its meaning (e.g. to verify the presence of a face in front of the terminal). The output of these mechanisms therefore can be mapped onto attributes of virtual *markers* that carry this information from the physical domain into an abstract symbolic representation suitable for the low-level reactive scripts and the high level *cognitive engine* to operate on.

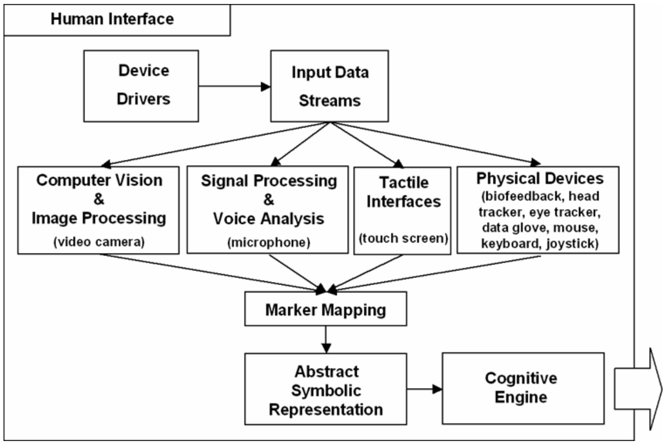


Fig. 10. Diagram of the Human Interface module.

Having discussed the major functional blocks of the *VHI* system implementation we now turn our attention to the implementation of its principal agent, the photo-real virtual child that plays a central role in the emotional modulation process.

V. EMOTIONAL MODULATION

Emotional modulation is a technique that helps transforming *information* to *knowledge* using our own emotions as a catalyst. The *Artificial Emotion Space (AES)* algorithm as part of the *VHI* system is the manifestation of a simple every day observation, that when in a good mood, we are generally more susceptible to information presented in a positive fashion, and on the contrary, when we are sad or down, we prefer things in more subdued manner. This simple notion of empathy supposes i) the recognition of the user's state of mind and ii) a mechanism to replicate it in the animation domain. As an example, the digital character may exhibit layers of its own emotions that coincide or alternatively oppose of that of the user's mood. Such behaviors create the illusion of empathy, compassion. Via this *coupling mechanism* the system is capable of improving the user's mood and relation to the presented material gradually driving him or her towards a happy state from sad or neutral. The *emotional modulation* algorithm serves as a powerful method in increasing the efficiency of a person's ability to absorb information in general and educational materials, in particular. The *AES* algorithm employs *Temporal Disc Controllers (TDC)* [32,33] whereas the resulting transitions from one emotional state to the other can be recorded and/or passed on to the *AI* unit that controls the duration, and the decay of these emotions, as well as it modulates the blending of different sources. The resulting output is a seamless deformation of facial surfaces (different emotional expressions), bone structures (tone and posture associated with moods) and even shading effect (skin blushing, tiredness, etc.) in real-time all happening in response to the students' own action and internal state.

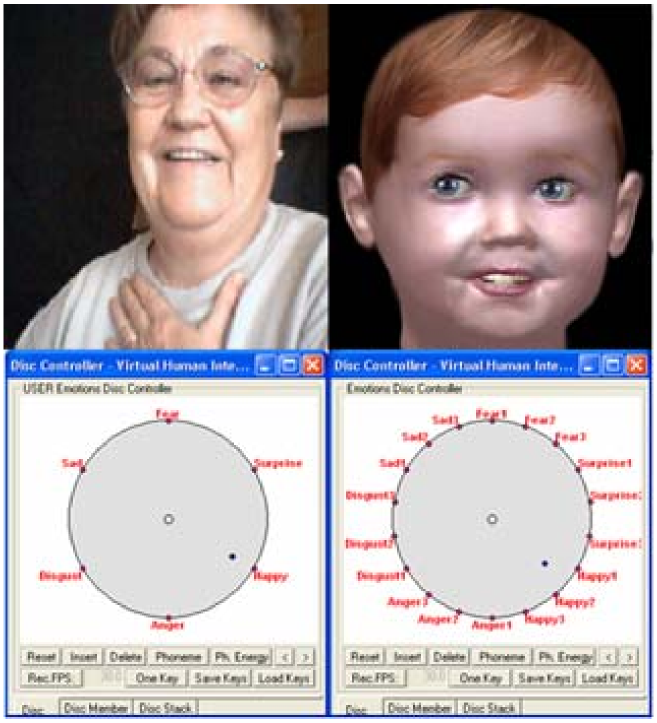


Fig. 11. The artificial Emotion Space (AES) algorithm links the user's internal state to that of the virtual child's animated character.

Fig. 11 demonstrates the simple operation of the *AES* algorithm. The upper row contains the image of the user during one of the sessions in the evaluation tests, with the corresponding image of the virtual human next to it. Below the *AES* representation/estimation of the user's own emotional state (left) is shown next to the *Temporal Disc Controller* of the digital character. Here the user's emotional state is estimated using *visual cues* from the *facial information processing* pipeline as well as input from other *optional devices*, such as a *biofeedback sensor*. Besides the capability to couple and modulate the emotional state of the user with that of the virtual human the *TDC* mechanism is also very efficient in automatically generating an infinite number of spontaneous facial expressions expressing emotions. This feature becomes very important in *avoiding the repetitive behavior* of the virtual character. For real-life practical applications, repetitive behavior becomes one of the most obvious and annoying faults of a digital character. The problem of repetitiveness comes from the limited and finite number of responses a system may have at its disposal to address a situation. Typically, these reactions are stored in a graph and recalled as needed. However, during a prolonged interaction with the user these repertoires quickly become obvious since the reactions are deterministic. Deterministic in this context means that every time a particular emotion or facial expression is required to accompany verbal content, the system produces the same visual output. Clearly this breaks the illusion of life or being a believable character. Therefore any perceptual multi-modal interface must be able to address this problem and avoid repetitive behavior. To demonstrate how this mechanism works we briefly describe how linking mechanisms may be use to implement low-level reactive control. *Linking mechanisms* are responsible for passing data and information from one stage of the interactive process to another thereby controlling various aspects of the animated dialogue process during a learning exercise. The basis of linking is the VHI's marker system that provides an abstraction layer to describe the outside world in the context of the educational VR application. Markers represent the *link* between the physical world and the virtual environment of the digital human character. They can be used as *symbols* to specify tasks and rule-based actions or as *variables* that reflect external physical properties associated with the user. *Markers* are dimensionless representations of 3D positions attached to any object in the scene. They carry unique *hierarchical labels* and *attributes values* encoding physical locations or abstract properties and they can be referred to by their names when describing high level tasks in a rule-based behavior control system. As an example, one or multiple markers may be attached to a camera, a virtual table, the floor, etc. Other objects with markers on them may have special significance, such as a virtual *monitor*, which serves to represent the user and the outside world in the synthetic world of a *virtual digital tutor*, as demonstrated in Fig. 12. This mechanism allows for controlling animated eye contact - a powerful means to support learning material - as driven by the direct interaction of the students. We use multiple discs to control the attributes and motion of all elements in a virtual scene. Due to the layered nature of this architecture each disc can be controller independently via animation channels (as part of the interactive content designed and produced before the interaction takes place), user interaction in real-time, or physical attributes represented by the marker system. Fig. 13 shows how this system works. In this simple example The *Emotion Disc Controller* is to be driven by the *hear rate variability* of the user with data collected using an external *biofeedback device*, capable of measuring stress levels in the user. When the level of stress rises above a preset value (*base*) compute the *distance vector* and use this amount to move towards the label named "*Fear*" (the direction vector is x = -0.99, y = -0.142).

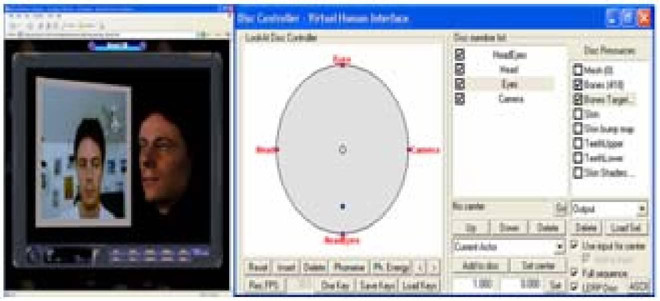


Fig. 12. Controlling eye contact using the marker system and linking mechanism create a level of believability and reactivity in a virtual digital tutor.

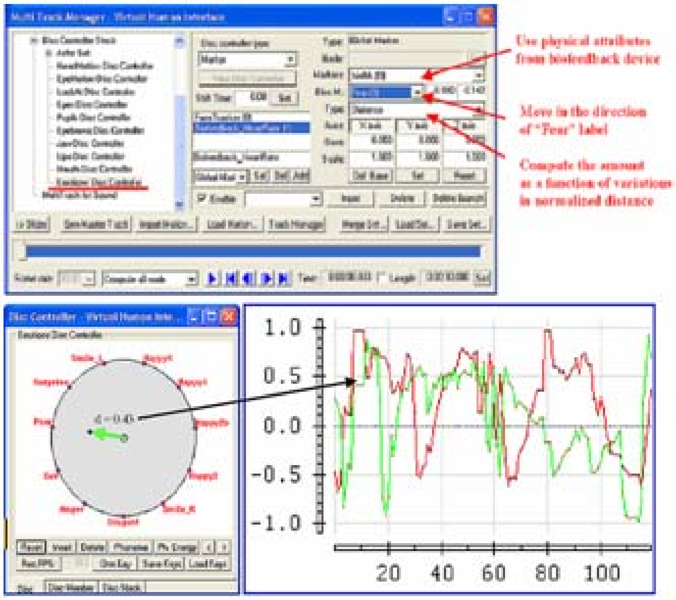


Fig. 13. Linking markers that represent physical measurements and interaction to control the animated facial display of a virtual face in our virtual reality education system.

Fig. 14. Low level visual cues (edges, morphology, etc.) observed on a student's face are extracted and mapped onto an emotion disc as competing 'clues'.

This is equivalent to a simple rule stating:

IF *the biofeedback device shows an increased hear rate* THEN *exhibit to be fear* END

The visual setup of this rule is shown in Fig. 14. The blue arrow refers to the realization of the above rule for normalized hear rate levels of $d = 0.43$ activations. The direction of the vector is defined using the label "*Fear*" and the amount of displacement by the data from the biofeedback device. The *advantage* of using this representation becomes more apparent when *multiple rules* are set up driving the same disc. In the current example, another marker may be controlled by the *facial expression recognition* module indicating the onset of a smile and therefore a "Happy" state (Rule2), while yet another could indicate the sudden motion of the mouse as a telltale sign of "Surprise" (Rule 3). Each of these *rules* adds another vector to the disc (see Fig. 14) but since effects are combined in an additive manner (black dotted arrow), the final facial display will be one single emotion. Thus, the disc acts as a *constraint system* that *interprets and integrates the different indications* of the student's emotional state [50,51,3]. When the constraints system is set up properly – like in the current example the circular arrangement of facial displays of affect verified by psychological studies – the disc operates as an *emotion recognizer*, a feature the vision system relies on heavily when recognizing the facial expressions of the student in front of the computer terminal.

VI. VIRTUAL TUTOR EVALUATION

Fig. 15. Examples of real-life application of our virtual-reality system for education and training.

The basic architecture of the *Virtual Human Interface* system, implemented as described in the above sections has been used in many practical applications in the fields of human computer interaction, education, advanced health care, and training. Specifically, these installations include an *E-hostess* for intelligent information kiosks as a tool to help explain the use and benefits of a particular method (Fig. 15/a) [33, 35], a virtual "Buddy" for children with autism (Figure 15/b) [32], "BabyTeach" an education program that teaches geography using an interactive virtual child while providing non-verbal facial feedback as to the correctness of the answers (Fig. 15/c) [34], virtual patients [16, 31], (Fig. 15/d), a VR-based rehabilitation system (Figure 15/e) [32], and finally psychiatric evaluation (Fig. 15/f) [3]. In the next paragraphs of this paper we present a more detailed description of our *BabyTeach* program and its evaluation.

Fig. 16. Using the BabyTeach program to study the map of Europe.

Nem vagyok szomorú.

Szomorú vagy nyomott vagyok.

Mindig szomorú vagyok és nem tudok kikeveredni belőle.

Annyira szomorú és boldogtalan vagyok, hogy nem bírom tovább.

Fig. 17. Example of using multiple-choice questions in the *BabyTeach* application.

The *BabyTeach* program package builds upon the *emotional modulation* and *non-verbal feedback* techniques described above to *teach geography* and help students of all ages to *take tests*. The system operates by prompting a series of questions and asking the user to indicate the correct answer with moving a pointer using the mouse. The animated face appears in the upper left hand corner and expresses encouraging or disappointed facial expressions according to the answers. As an example, when the mouse moves toward the proper answer it gently leans forward or nods as a form of acknowledgment. Once a selection is made, immediate feedback regarding the choice is displayed as a smile or tweaking mouth. Fig. 16 demonstrates this process by showing a test map of Europe. During learning the student may make many attempts while the baby's face gently guides him or her towards the correct answer. Apart from the positive effects of *emotional modulation* on learning, this application also involves the *motor system* thereby further helping to memorize spatial locations. There are multiple levels of difficulties in which the system offers increasingly more context and details. The figure shows average difficulty, where the lines of different countries are draw. Easier tests would contain cities and other landmarks, while the most difficult task involves only the map with no boundaries drawn. During training, the *BabyTeach* system also offers cues after a preset number of failed correct answers. In particular, the correct answers will be shown for the student to study and the same question later asked. As a result of the positive facial feedback the student may explore the map by continuously holding the mouse down. The facial feedback of the virtual child will indicate how close he or she is to the country of choice. Of course, when in test mode, no facial feedback is available. However, this mechanism proved to be very successful in helping children to learn geography in a fast and pleasurable manner.

In the above example the preset animated behavior of the virtual child is dynamically modified by the answer of the student. Specifically, as the mouse pointer moves over the map area, the emotional state of the student is affected as a function of the distance from the correct location. The timing of the virtual character's reactions is set such that it appears to be responsive, but not immediate. It allows the student to settle on an answer and if that is correct is smiles, and frowns if not. The underlying mechanisms ensure that the virtual child moves constantly and appears to be "alive". With the help of the built-in face recognition system, it is also capable of maintaining eye contact, following the learner's face in front of the computer terminal and apply different strategies to track the motion of the user. Another example of using the *BabyTeach* program is shown in Fig. 17. Here the student is prompted with multiple-choice questions, e.g. while taking a language test or filling out a questionnaire. Each answer is associated with a range of different expression, usually representing a continuous scale from positive towards negative. And the virtual child "rewards" the answers with immediate visual feedback.

Fig. 18. Snapshot from a video sequence recorded during an interactive learning session.

In both cases, the *VHI* system continuously records the performance of the student. Beyond gathering statistics, such as reaction time, number of trials, etc. it also records a video of the student as well as all what happens on the screen. This is achieved with the help of *multiple virtual cameras* that can be directed to view any portion of the 3D test scene. The respective views of these virtual cameras are rendered into texture buffers and placed on the screen in front of the student using *pixel shaders*. In parallel, they are saved to a sequence of image files with a preset frequency. Fig. 18 shows one output frame of such a sequence. The immediate advantage of such representation is that later these videos can be analyzed to study the reactions and emotions of the student during the learning process. This provides a glimpse into the underlying psychological processes and the mechanisms of learning as it is discussed in the following section.

To evaluate the performance of the *BabyTeach* system we designed a set of experiments in order to explore how effective the technique of *emotional modulation* is and how it takes place during the interaction process itself. In particular, we were interested in detecting and confirming two fundamental mechanisms, one is *gating* and the other one is *mirroring*. *Gating* is a mechanism with great importance in learning. It means that the presence of emotion in a subject while studying multiplies or in other cases degrades the efficiency of the learning process. Depending on the positive or negative emotions associated with the learning experience or the subject matter gating can help the embedding of information in the brain. The other mechanism, *mirroring*, refers to how the emotion from one person causes another one to take on the same emotion and facial display. In our case this means that the positive facial expressions on the animated baby's face should be followed by positive emotions on the student's face.

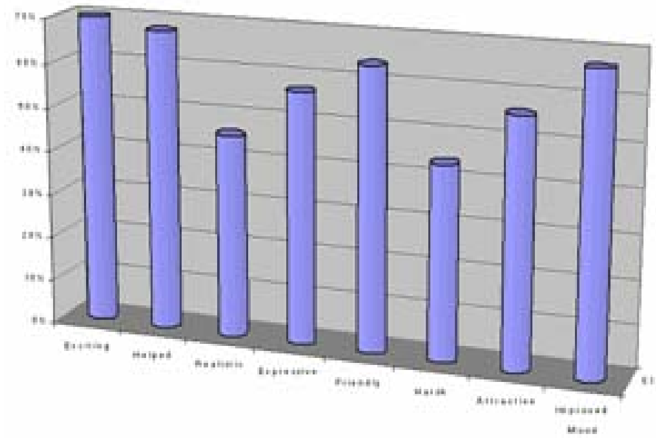


Fig. 19. User scores on using the BabyTeach system.

In the experiments *15 individuals*, most of them attending university or are at least a high school degree, were tested. The group had an average age of 24.1 years (deviation= 4.1 years) and a 2:1 ratio of females vs. males. The learning task selected for these experiments was to locate cities on the map of Europe as shown above. The tests were carried out in two phases. While the baby was continually moving as if it was a real person, its positive *facial expressions were modulated by the selections the student made*. As an example, when the person taking the test clicked on the country, the virtual baby smiled or grinned depending whether the answer was correct or not. The intensity of this expression depended on how close the selected location was to that of the correct answer. Another mechanism to support the learning experience we *positive head nods* which the virtual human character delivered whenever the direction of the mouse movement was such that it indicated a possibly correct answer. This type of visual feedback was provided continuously until the user decided to finalize his or her answer. In each experiment the system randomly asked 10 countries and collected statistical information on the correctness, speed and accuracy of the answers. In some case, when there were more than five errors to any specific question, the *VHI* system showed the map with the name of the countries filled in. During the second portion of this experiment users were asked to carry out the same task as before, but without the virtual child present or providing feedback. We examined how effective the learning process was by prompting them with ten questions again and comparing their responses with against their own previous scores as well as those of others in the test group. In addition to the statistical data being collected using the web camera interface of the *VHI* we also recorded the user during the process of interaction. The purpose of doing so was later to be able to find out if *gating* and/or *mirroring* occurred.

At the end of the experiments each subject was asked to fill out a short questionnaire scoring the performance of the system and expressing their opinion. For each answer we used a sliding scale containing seven different possible choices with the scores ranking from 1 through 3 considered as negative, 4 being neutral and scores above 4 treated as being positive opinions.

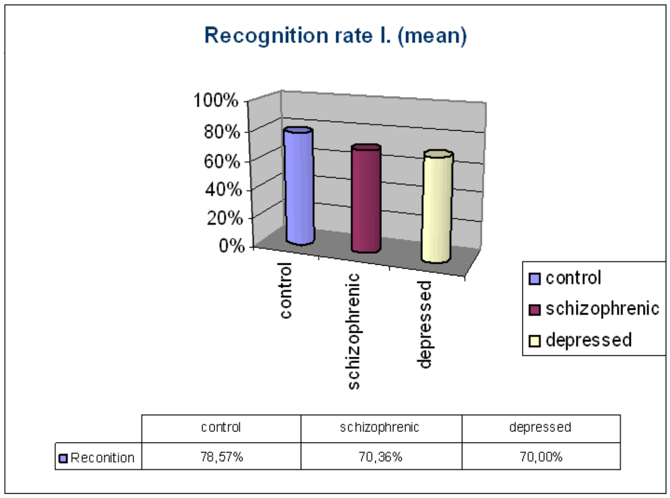


Fig. 20. Statistical results obtained from the static emotion identification experiment (based on a study of 25 people).

The results of our experiment showed the 70% of the subjects tested considered the *BabyTeach exciting* (average=70%, deviation=18%), with 68% saying it *helped* them during the exercises (average=68%, deviation=21%), and another 67% finding that their *mood improved* (average=67%, deviation=20%). People also tended to find the virtual child *friendly* (average=64%, deviation=22%) expressive (average=57%, deviation=25%) and attractive (average=56%, deviation=21%). Finally, the students tended to perceive the digital character as reserved (average=44%, deviation=18%) perhaps because it did not speak, but not too realistic (average=46%, deviation=17%). These results are summarized in Fig. 19. Upon analyzing the recorded videos during the interaction we confirmed that most of the students actually communicated with the virtual child while learning took place. *Mirroring* often occurred indicating that what happens in real life between two people participating in a dialogue (who copy the facial expressions of one another) may also take place in the interactive dialogue with the virtual child. Fig. 18 shows an example of mirroring. Finally, *most subjects agreed that the positive emotional affect* the *BabyTeach* program brought to the learning experience added to their ability in memorizing things.

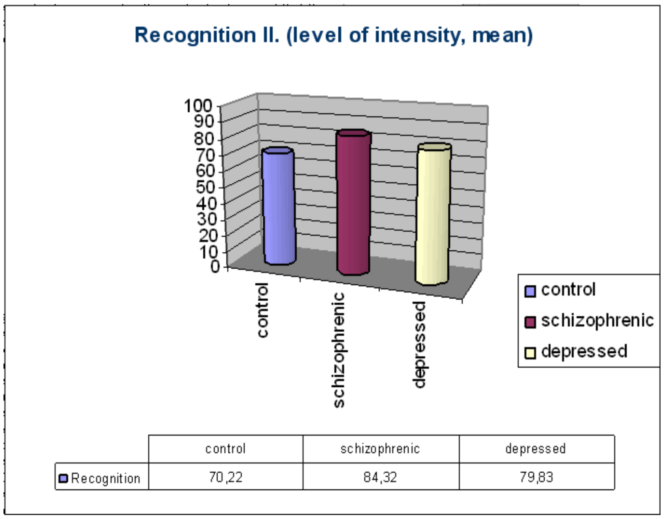


Fig. 21. Statistical results obtained from the dynamic/time-varying emotion intensity recognition experiment (based on a study of 25 people-see text).

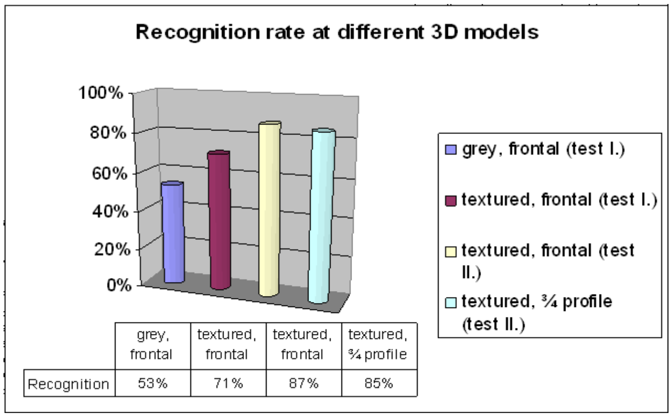
The above results represent the outcome of a relatively small group of people and of a specific age group. Nonetheless these early results have already *demonstrated the positive effects of using a closed-loop interactive dialogue* model with an animated virtual human character with perceptive and reactive capabilities. Another round of human experiments targeting a more specific and younger group, primary school children, is currently on-going. Our future work includes a more detailed evaluation of the learning tool on a representative population of primary school children as well as the implementation of new content to support other subjects of arts, literature and sciences. To characterize the underlying emotion recognition and educational mechanism of using virtual faces as a tutor in VR applications we further investigated the notion of repeatable and parametric facial stimuli are presented interactively to the users of such systems. Specifically, in these experiments high fidelity, photo realistic digital replicas of a male and a female face were used. Capable of expressing fine tones of neutral and six basic emotions (happiness, anger, sadness, disgust, fear, surprise). Their circular arrangement conformed to the findings of describing how transitions occur in emotion space [51]. In the studies a total of 25 people (10 control and 15 with psychiatric disorders) were evaluated in two experiments as shown below:

- In Experiment #1 subjects were asked to *identify emotions* using only the frontal views of the two 3D animated heads as they appeared briefly on the screen. Two separate sequences (7 pictures each) were recorded, one with textured models and the second one without textures on them (gray heads).

- In Experiment #2 *slowly varying emotional displays* we created and shown to the subjects as they gradually changed starting from the neutral position towards their respective 100% activation values during a 5 seconds period. Subjects were asked to start and later stop the animation sequence by first pressing and releasing a button when they recognized any emotion on the digital face.

Fig. 20 through Fig. 22 show the results of the above two experiments. Based on statistical analysis of the data collected the most important finding was that even at this relatively small sample scale of the subject population and the control group *measurable differences in the recognition rate of emotional display were found.* Furthermore there was a *significant difference* in the second experiment as well. In particular, differences were shown in the average *intensity levels of emotions* when subjects recognize them. These results are summarized in Fig.20. They support the notion that significant differences exist between the control and patient groups when recognizing digitally created basic displays of affect with the help of the *Virtual Human Interface*. These studies confirmed *that the digital facial images we generated using the VHI are processed the same way that real faces are.*

The final experiment was designed to find out what other imaging variables (e.g. shape, texture gender, view point, etc.) influence the very process of emotion recognition itself. In particular, in *Experiment #3*, gray head models were used instead of full colored ones with facial texture on them, rotated the head 45 degrees to show a ¾ profile view instead of the frontal, and varied the time of exposure the subject could have seen the images. These results are summarized in Fig. 22.



	grey, frontal	textured, frontal	textured, frontal	textured, ¾ profile
Recognition	53%	71%	87%	85%

Fig. 22. Comparison of recognition rates as a function of 3D facial model quality (See text).

Our general conclusions from this third experiment are as follows:

- Expressions on the gray models (i.e. only shape information without facial texture) are more difficult to recognize than otherwise.
- Recognition rates of particular expressions were almost the same whether they were presented from a frontal or ¾ profile view.
- The shorter an emotion is displayed (Test I.), the more difficult it becomes to recognize it.

Partly, the above results they seem to contradict the outcome of other experiments on facial information processing, which used photographs. Particularly, there *should be* significant differences between the recognition rate of frontal faces vs. ¾ profile views. While in these experiments we managed to show that the former is higher, the differences are not pronounced enough to draw a positive conclusion. On the other hand these experiments *confirmed* that affect recognition from shape alone (gray heads) is more difficult than using textures. Therefore the ability of the *VHI* system to deliver *high fidelity detail in texture (skin tone as well as secondary effects, such as wrinkles, blushing, etc.) is critical for the application of virtual humans in general and digital tutors to support virtual-reality education, in particular*. These results are presented in more detail in [51,52].

VII. APPLICATIONS BEYOND THE CLASSROOM

Besides the classroom application described above our portable virtual reality solution have been utilized in many training scenarios related to Image Guided Diagnosis and Therapy and the Operating Room of the Future [56]. The goal of using virtual reality in anatomically guided ultrasound diagnosis is to aid the medics locating anatomical structures inside the body to detect internal bleeding [57]. Our compact visualization and guidance system is shown in Figures 23 through 25. Medics are using the system to detect internal bleeding of a patient lying on the ground. The entire system is carried in a backpack on his back, his hands hold three outside components, namely the US probe, the reference cube for the tracking system, and the touch-sensitive computer screen which serves as the main interface to access the system's functions. On the right the HW elements packed tightly inside the backpack are shown. From left to right a Portable Ultrasound device (Sonosite C180) is combined with a 6DOF, high precision tracker (Ascension Microbird) to create a tracked sequence of US images. This information is transmitted to a small computer (Apple macMini), the heart of our Anatomical 3D Guidance solution, that combines this information with generic models of the human body, called virtual humans, and outputs guidance information to the medic's hand-held touch screen.

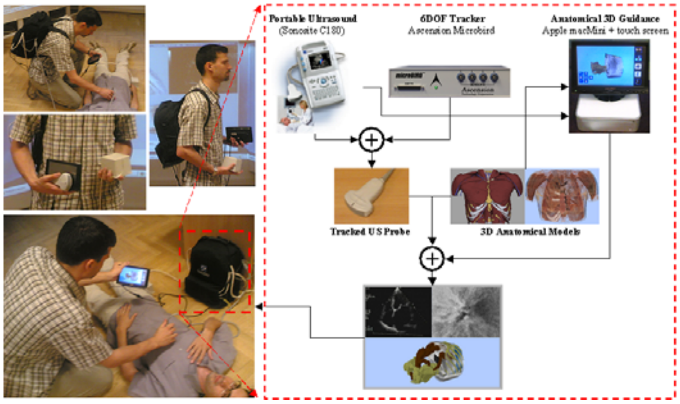


Fig. 23. Overview of our backpack anatomically guided ultrasound system.

Fig. 24. User interface and system in use.

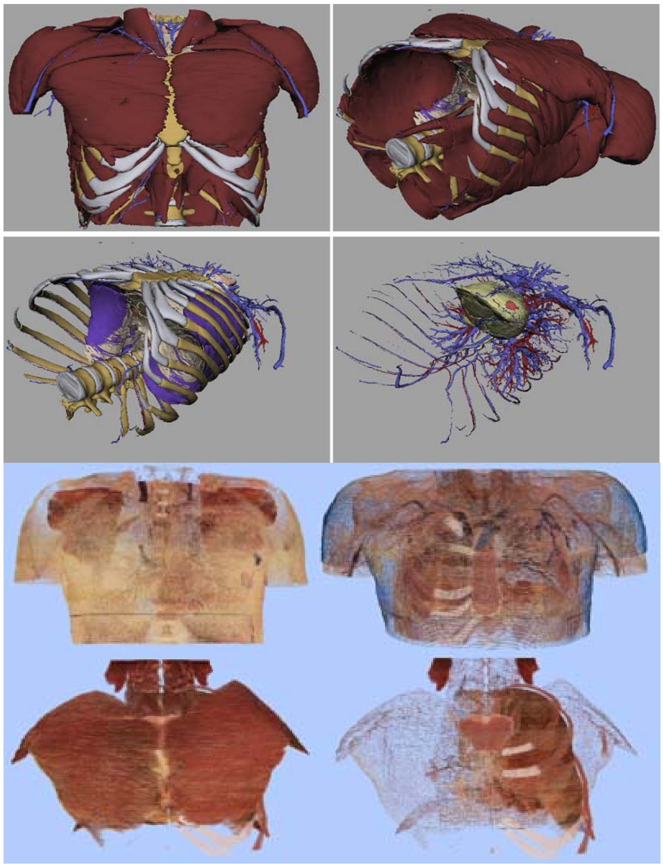


Fig. 25. Segmented and labeled anatomical structures derived from the NIH Visible Male Data set and optimized for real-time viewing and guidance.

Our final figure demonstrates using VR to build the operating room of the future where entire medical procedures are recorded, played back and later optimized with the help of VR tools.

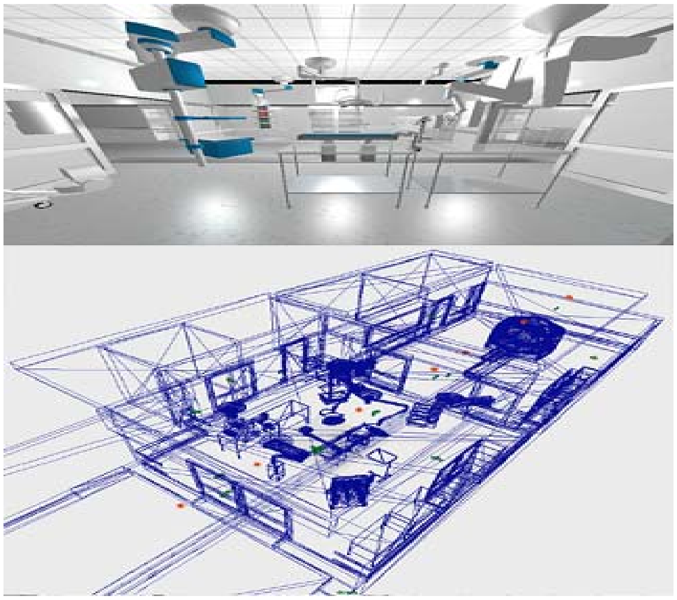


Fig. 26. Virtual reality operating room model to help optimize medical procedures and train surgeons in the Operating Room of the Future [56].

VIII. CONCLUSION

In this paper we presented an overview and a *road-map* to build virtual-reality educational tools using low-cost interactive devices and novel modalities. We discussed in detail, how *cognitive dynamics* play a vital role in the learning process and furthermore how to use these mechanisms and achieve improved results in learning via *closed-loop-interaction*. We presented a number of devices readers may use to construct their own systems and also reviewed the architecture and basic modules of our specific virtual reality environment, called the Virtual Human Interface, which incorporates all these modalities in a single, easy-to-use platform. In addition to the review of some basic devices we presented a framework to use the physical measurements obtained by them to control animated content and reactive behavior. The VHI also includes high fidelity animated faces, used as digital tutors to help students learn. These virtual faces rely on *non-verbal cues* to support educational dialogue and achieve *emotional modulation*, a process vital for learning to take place. Finally, we introduced a number of *real-life applications* including a program called *BabyTeach*, *Anatomical Ultrasound Guidance*, and *Operating Room of the Future*, where our system was successfully deployed in, and carried out detailed psychiatric evaluation of using high fidelity digital faces as a prime interface for VR-based education. Our results indicate, that such methodology provides a valid and effective technical foundation to use VR as an educational and training tool and we predict, that with the advent of low-cost devices which now can be placed in the hand of educators, we are now at the verge of a new wave where VR will realize its potential and go through a life cycle similar to multi-media in the past decade.

REFERENCES

[1] Aura Sound Systems (2007), Bass Shakers, www.aurasound.com.
[2] Cannon (2007) VC-C50i P/T/Z communication camera www.canon-europe.com/For_Home/Product_Finder/Web_Cameras/Web

_Cameras/vc_c50i/

[3] G. Csukly, L. Simon, B. Benedek, G. Szijártó, B. Kiss and B. Takacs. Evaluating Psychiatric Patients Using High Fidelity Animated 3D Faces, *CyberTherapy*, San Diego, California, January 2004.

[4] DepthQ, 3D projector, www.depthq.com, 2005.

[5] Digital Elite, Virtual Human Interface and Cyber Reality System, www.digitalElite.net, 2007.

[6] Eagletron, Tracker Pod Robotic Camera, www.trackerpod.com, 2007.

[7] eMagin, Z800 3D Visor, www.3dvisor.com www.emagin.com, 2007.

[8] Essential Reality, P5 Data Glove, www.essentialreality.com/Alliance_Brand/Products.php, 2005.

[9] Eye Response Technologies, Eye Tracker System , www.eyeresponse.com, 2007.

[10] GigaByte, Dual SLI Graphics Processor, http://www.xbitlabs.com/news/video/display/20050715030609.html, 2005.

[11] R. Gutierrez-Osuna. Olfactory Interaction, Encyclopedia of Human-computer Interaction, *W. S. Bainbridge (Ed.), Berkshire Publishing*, pp. 507-511, 2004.

[12] Holographica, Holographic Display, www.holografika.com, 2007.

[13] Intersense, InterTrax2 and InertiaCube3 Trackers, www.intersense.com, 2007.

[14] A. Kapoor, S. Mota and R. W. Picard. Towards a Learning Companion that Recognizes Affect, Proc. of *Emotional and Intelligent Agents II: The Tangled Knot of Social Cognition*, AAAI Fall Symposium, 2001.

[15] KeyTec. Magic Touch, www.magictouch.com, 2007.

[16] B. Kiss, B. Benedek, G. Szijártó, G. Csukly, L. Simon and B. Takacs. The Virtual Patient: A Photo-real Virtual Human for VR-based Therapy, *Medicine Meets Virtual Reality*, Newport Beach, CA. , 2004.

[17] B. Kort and R. Reilly. Analytical Models of Emotions, Learning and Relationships: Towards and Affect-sensitive Cognitive Machine, in Proc. *Intl. Conf. on Virtual Worlds and Simulation*, San Antonio, TX. , 2002.

[18] Logitech, 3D Mouse, www.vrealities.com/3dmouse.html, 2007.

[19] MIDI, MIDI Manufacturers Association, www.midi.org, 2007.

[20] Mitsubishi, Palm Sized Pocket Projector, www.i4u.com/article2867.html, 2007.

[21] Natural Point, OptiTrack Flex3 Motion Capture Solution, www.naturalpoint.com/optitrack/, 2007.

[22] Natural Point, TrackIR4 6DOF Head Tracker, www.naturalpoint.com/trackir/, 2007

[23] NVidia, SLI Technology, www.slizone.com/page/home.html, 2006.

[24] Olfacom, www.olfacom.com, 2007.

[25] PI Engineering and Tash, X-Keys and Buddy Buttons, www.piengineering.com, www.tashinc.com, 2007.

[26] Point Grey Research, LadyBug2 Spherical Digital Video Camera, http://www.ptgrey.com/products/spherical.asp, 2007.

[27] H. Regenbrech, J. Hauber, R. Schoenfelder and A. Maegerlei. Virtual Reality Aided Assembly with Directional Vibro-tactile Feedback, in Proceedings of *GRAPHITE*, Dunedin, New Zealand, 2005.

[28] L. Schomaker et. al. A Taxonomy of Multimodal Interaction in the Human Information Processing System, http://www.ai.rug.nl/~lambert/projects/miami/taxonomy/taxonomy.html, 1995.

[29] Sensable Technologies, Phantom Desktop Haptic Device, http://sensable.org/products/phantom_ghost/phantom-omni.asp, 2007.

[30] SunRise, Virtual Reality Education Programs, http://www.sunrisevr.com/edvrprog.htm, 2007.

[31] B. Takacs, G. Szijártó and B. Benedek. Virtual Patient for Anatomical Ultrasound Guidance, *3rd Hungarian Conference on Computer Graphics and Computational Geometry*, 2005.

[32] B. Takacs. Special Education and Rehabilitation: Teaching and Healing with Interactive Graphics, *IEEE Computer Graphics and Applications*, Sept/Oct., pp.40-48, 2005.

[33] B. Takacs and B. Kiss. Virtual Human Interface: a Photo-realistic Digital Human, *IEEE Computer Graphics and Applications, Special Issue on Perceptual Multimodal Interfaces*, Sept/Oct., pp.39-45, 2003.

[34] B. Takacs. Cognitive, Mental and Physical Rehabilitation Using a Configurable Virtual Reality System, *Intl. Journal of Virtual Reality*, vol. 5, no. 4, pp.1-12, 2006.

[35] B. Takacs. Affective Intelligence: A Novel User Interface Paradigm, in *1st Int. Conference on Affective Computing & Intelligent Interaction*, Beijing, China, J. Tao, T. Tan, and R. W. Picard (Eds.): ACII 2005, LNCS 3784, pp. 764-771. Springer-Verlag, 2005.

[36] Tobii, Eye tracker system, http://www.srlabs.it/pdf/Tobii1750_PD_2.pdf, 2006.

[37] TriSenx, Scent Dome, www.trisenx.com, 2007.

[38] Wild Divine, www.wilddivine.com, 2007.

[39] Immersive Media, http://www.immersivemedia.com/, 2007.

[40] Remote Reality, http://www.remotereality.com/, 2007.

[41] Review of Immersive Video Solutions, http://www.washedashore.com/projects/surround/, 2007.

[42] Optimus Mini Three, http://www.artlebedev.com/everything/optimus/, http://www.artlebedev.com/everything/optimus-mini/, 2007.

[43] Open Eyes, http://hcvl.hci.iastate.edu/cgi-bin/openEyes.cgi, 2007.

[44] Matrox QID Pro, http://www.matrox.com/mga/workstation/3dws/products/id/qid_pro.cfm, 2007.

[45] Vernier Technologies, http://www.vernier.com/probes/, 2007.

[46] C. Sik Lányi, Z. Geiszt, P. Károlyi, A. Tilinger and V. Magyar. Virtual Reality in Special Needs Early Education, *The International Journal of Virtual Reality*, vol. 5, no. 4, pp. 55-68, 2006.

[47] P. Yin, X. Jiang, J. Shi and R. Zhou. Multi-screen Tiled Displayed, Parallel Rendering System for a Large Terrain Dataset, *The International Journal of Virtual Reality*, vol. 5, no. 4, pp. 47-54, 2006.

[48] H. Kaufmann and M. Csisinko. Multiple Headmounted Displays in Virtual and Augmented Reality Applications, *The International Journal of Virtual Reality*, vol. 6, no. 2, pp. 43-50, 2007.

[49] B. Takacs. PanoMOBI: A Panoramic Mobile Entertainment System, *ICEC07*, Shanghai, China, LNCS 4740, pp. 219-224, 2007.

[50] Nintendo. Wii controller (2007), http://wii.nintendo.com/.

[51] R. Adolphs. Neural Systems for Recognizing Emotions, in *Current Opinion in Neurobiology*, vol. 12, Elsevier Science Ltd. , 2002.

[52] G. Csukly, P. Czobor, L. Simon and B. Takacs (in press). Basic Emotions and Psychological Distress: Association between Recognition of Facial Expressions and Symptom Checklist-90 Subscales, *Comprehensive Psychiatry*, YCOMP-50405.

[53] G. Csukly and B. Takács. Assessing Affect and Facial Expression Recognition Abilities Using Virtual Faces: A Clinical Study, in Proc. *Central European Multimedia and Virtual Reality Conference*, Eger, Hungary, 2006.

[54] C. Sik Lányi. Virtual Reality in Healthcare, in *Intelligent Paradigms for Assistive and Preventive Healthcare*, A. Ichalkaranje, et al. (Eds.), Springer-Verlag, 2006, pp. 92-121. ISBN 3-540-31762-7.

[55] C. Sik Lányi. Multimedia Medical Informatics System in Healthcare, in *Intelligent Paradigms for Assistive and Preventive Healthcare*, A. Ichalkaranje, et al. (Eds.), Springer-Verlag, pp. 39-91, 2006. ISBN 3-540-31762-7.

[56] Harvard CIMIT, Image Guided Therapy, Operating Room of the Future, www.cimit.org/orfuture.html, www.cimit.org/igt.html 2007.

[57] B. Takacs, G. Szijarto, D. Hanak, Constant Rate Visualization of Large Anatomical Models for Medical Guidance, GRAFGEO, Budapest, 2007.

[58] B. Takacs, and K.G. Vosburgh, Compact Anatomically Guided Ultrasound for Casualty Care**,** *ACHI, Sainte Luce, Martinique, 2008.*

Barnabas Takacs (PhD. 97, M.Sci'94) was born in Budapest, Hungary. His research interest includes developing the technological foundation for real-time virtual humans and their applications, such as photo-real human animation and modeling, face recognition, motion tracking, and AI-based intelligent conversational agents as well novel applications of VR in rehabilitation. He has spent much of his carrier in industry where in 1999, as the Director of Research of Virtual Celebrity Productions (Los Angeles, California, USA) he has co-developed a novel animation technology, called Digital Cloning which, for the fist time, used computer animation to create believable CG-humans for the film industry. Subsequently, he founded Digital Elite Inc. (Los Angeles, California, USA) to bring these technologies to the realm of real-time, interactive characters. His current work involves using virtual-reality technology in a variety of applications including education, health care and image-guided diagnosis & surgery.

Dr. Takacs currently heads the Virtual Human Interface Group at SZTAKI, Hungarian Academy of Sciences in Budapest, Hungary. He is also affiliated with Harvard BWH in Boston, MA, USA. He can be reached at BTakacs@sztaki.hu www.vhi.sztaki.hu / www.digitalElite.net