Editorial Policy

The Australian Journal of Intelligent Information Processing Systems is an interdisciplinary forum for providing the latest information on research developments and related activities in the design and implementation of intelligent information processing systems.

The areas of interest include, but are not limited to:
- artificial intelligence
- artificial neural networks
- computer science
- fuzzy systems
- human centred computing
- virtual reality

The journal both theoretical and application-oriented papers in diverse areas related to intelligent systems. The journal also publishes survey articles, short notes, Ph.D. thesis abstracts and project reports.

Editorial Board

Tom Gedeon  
Editor in Chief  
*Australian National University, Australia*

Yianni Attikiouzel  
*Murdoch University, Australia*

James C. Bezdek  
*University of Western Florida, USA*

Tony Constantinides  
*Imperial College of Science, Technology and Medicine, U.K.*

Peter Eklund  
*University of Wollongong, Australia*

Nikola Kasabov  
*Auckland University of Technology, New Zealand*

László Kóczy  
*Budapest University of Technology & Economics, Hungary*

Takeshi Furuhashi  
*Nagoya University, Japan*

Marimuthu Palaniswami  
*University of Melbourne, Australia*

M. V. Srinivasan  
*University of Queensland, Australia*

A. Venetsanopoulos  
*University of Toronto, Canada*

Responsibility of the content of the papers rests upon the authors and not with the publishers. Data and conclusions developed by the authors are for information only and are not intended for use without independent substantiating investigation on the part of the potential user.

The Australian Journal of Intelligent Information Processing Systems is published quarterly. All correspondence including manuscripts and advertising inquiries should be addressed to the Editor in Chief: Professor Tom Gedeon, Faculty of Engineering and Information Technology, The Australian National University, Canberra ACT 0200, Australia.

Abstracting is permitted with due credit to the Australian Journal of Intelligent Information Processing Systems.

Subscriptions:
Individual rate: A$90 within Australia, A$99 overseas
Institutional rate: A$310 within Australia, A$308 overseas.

Individual copies can be purchased at A$25 per single copy.

Reprints of technical articles are available, quantities no less than 50 can be ordered. Orders should be addressed to: AJIIPS, Faculty of Engineering and Information Technology, The Australian National University, Canberra ACT 0200, Australia. Email address: ajiips@cs.anu.edu.au.
Special Issue: Computational Intelligence and Human-Centered Interfacing Approaches in Cognitive Infocommunications

Editors: Adam Csapo, Leana Copeland, Joni Jämsä

CONTENTS

Recognizing Complex Head Movements
Kornél Bertók, Attila Fazekas ................................................................. 3

Developing a 3D virtual library model based on the ancient Library of Alexandria
István Boda, Erzsébet Tóth, István Csont, László T. Nagy ............................. 19

The education of disruptive technologies in the innovative engineering training
Ildikó Horváth ......................................................................................... 31

BCI controlled WiFi mobile robot
József Katona, Attila Kovari ................................................................. 61

This is a refereed journal.
Abstracted by INSPEC.

Publication Date December 2016
Humans and the infocommunications network surrounding them are merging together at several levels, ranging from the level of interactions with personal devices to the highest level of collective behaviors such as mass movements, mass habits etc. Consequently, humans and infocommunications will soon coexist as an entangled web, resulting in an augmentation of both natural and artificial cognitive capabilities across multiple modalities. This process of merging is occurring today, and is expected to gain further impact in the near future.

Cognitive infocommunications (CogInfoCom) is a multi-disciplinary field that aims to reflect on this process of merging by investigating the link between the research areas of infocommunications and the cognitive sciences, as well as the various engineering applications that have emerged as a synergic combination of these sciences. A key observation behind the field is that through the long-term co-evolution of humans and ICT, new cognitive entities are formed which cannot, at least from the perspective of cognitive capabilities, easily be broken down into their human and artificial components. The goal of this special issue is to present a cross-section of research results relevant to the application of computational intelligence and human-centered interfaces in support of the formation of new cognitive entities.

Potential topics include, but are not limited to:

- Augmented human capabilities (e.g. in industrial applications)
- Cognitive control
- Cognitive infocommunication channels
- Ergonomics-based aspects of networked CogInfoCom
- Human-oriented aspects of Future Internet
- Interaction capabilities of CogInfoCom systems, including e.g. physiological interfaces (BCIs, BANs, etc.), modality-oriented CogInfoCom (human-X interaction), augmented content management and cognition
- Augmented mathematical capabilities (mathability)
- Modeling cognitive biases
- Augmented speech-based capabilities (speechability)

**Guest Editors**

Adam Csapo  
Szechenyi Istvan University  
Gyor, Hungary  
csapo.adam@sze.hu

Leana Copeland  
Australian National University  
leana.copeland@anu.edu.au

Joni Jämsä  
Centria University of Applied Sciences  
joni.jamsa@centria.fi

**Manuscript Due:**  
31 July, 2016

**First Round of Reviews:**  
15 September, 2016
Recognizing Complex Head Movements

Kornél Bertók
Department of Computer Graphics and Image Processing
University of Debrecen, Faculty of Informatics
Debrecen, Hungary
E-mail: bkornel@gmail.com

Attila Fazekas
Department of Computer Graphics and Image Processing
University of Debrecen, Faculty of Informatics
Debrecen, Hungary
E-mail: attila.fazekas@inf.unideb.hu

Abstract—Developing new approaches of Human-Computer Interactions, which make the interaction logic more natural to humans, is one of the most forward-looking trends of information technology. There is growing demand for infocommunications which are not limited to traditional senses. For example, recognition of human activities or the gesture based control of computers makes the communication between human and computer easier than ever. There is a leading communication channel of the message that is being communicated: the human head and face. Recognizing head movements are crucial in understanding the message as well as making the interaction self-evident to humans. Head movements can be recognized based on their shape and characteristics. In this study, we introduce a general model for describing and understanding complex head movements. A novel spatiotemporal representation is used for head movements, which is based on the extraction of facial features and the estimation of 6 DoF head pose. Our approach uses monocular camera images without the need for sensors or markers on the face. The efficiency of head movement recognition is demonstrated by head-drawn letters in real time environment.

Keywords—infocommunications; human activity recognition; head movement recognition; spatiotemporal head movements; head pose estimation

I. INTRODUCTION

There is a notable influence of information technology on our daily life. Electronic devices, smartphones, and telecommunications became an integral part of our everyday lives. Information processing and content handling are becoming more and more important for almost everybody, and it is one of the most promising and forward-looking trends of the future. For today, it has formed a particular field of science for analyzing the information transmitted through telecommunications, namely the field of infocommunications which is a superset of telecommunication. Infocommunications expand the space with several functions of information processing and content handling, and so create new opportunities in the way of interaction, especially in Human-Computer Interaction (HCI). This fact is important because the degree of connection between humans and infocommunications network are getting more and more crucial. We live in a world where the level of interaction with personal devices (i.e. smartphones, TVs or consoles) no longer restricts on traditional controlling or input methods.

Nowadays, ICT systems with the ability to understand audiovisual information are getting more and popular. There is also an increasing demand for ICT systems which are not limited to the traditional senses. Such systems, for example, can respond to complex human activities, gestures in the 3-D space. The human factor is one of the most important parts of infocommunications. A sensory information can be obtained, transformed and then transferred to the user, in a way that the user can process effectively [1]. It is, therefore, necessary to analyze human behavior during the information processing or to make the whole interaction logic more natural to humans. Among others, this is the purpose of cognitive infocommunications. Cognitive infocommunications establish the link between several research areas (mainly infocommunications and cognitive sciences), in order to expand the content space with cognitive and sensory contents [2]. Niitsuma et al. proposed the concept of spatial memory for describing a memory system in the 3-D space [3]. Users can interact with spatial memory using gesture-like commands (the terminology of cognitive infocommunications call these commands as bodicons [4]). This way, the interpretation of bodicons namely the connection between human activities and segments in spatial memory can be established by cognitive systems.

The ability of human activity recognition (HAR) can be essential for cognitive systems. HAR involves recognizing gestures, interactions or other actions performed by users from a series of observations about them. Nowadays, the importance of HAR, along with infocommunications, is growing rapidly due to the demand of personalized interfaces with new HCI approaches. Fitting the system to the expectations and abilities of users is getting more and more important. HAR provides new and unique solutions and
interfaces to interact computers in a different way as the usual (i.e. controlling applications with head movements). During HAR, several attributes of the user are measured and indexed over the time dimension and mathematically it can be defined as below [5]:

**Definition 1 (HAR problem (HARP)):** Given a set \( S = \{ S_0, ..., S_{k-1} \} \), of \( k \) time series, each one from a particular measured attribute, and all of them defined within a time interval \( I = [t_0, t_w] \). The goal is to find a temporal partition \( \{ I_0, ..., I_{r-1} \} \) of \( I \), based on the data of \( S \), and a \( l_j \in L \) label which represents the activity performed during a given interval \( I_j \) (e.g. head shaking, nodding, etc.). It implies that each \( I_j \) is consecutive, non-empty, non-overlapping, and such that \( \bigcup_{j=0}^{r-1} I_j = I \). In the following term, a HAR problem is an ordered triple \( HARP = (S, I, L) \), comprising a set \( S \) of measured attributes, a set \( I \) of time intervals together with a set \( L \) of labels.

II. RELATED WORK

The information encoded by head movements has a major role in HCI, it is because the recognition of head movements has got a considerable attention in the state-of-art studies (including surveillance-, patient monitoring systems, or any others that involve interactions between humans and computers). It is required from a HAR system to be able to recognize complex activities of multiple actions, which can be performed more than a single person. HAR approaches can be classified by placing a sensor (e.g. a camera or accelerometer) the user, which is often called as anchoring [6]. Anchoring can be relative or absolute to the user. It means, if the sensor moves together with the user then the sensor is anchored relatively. Furthermore, if the sensor is fixed in the users’ environment, then it is anchored absolutely.

A. Relatively-anchored approaches

Relatively-anchored approaches usually use a HAR sensor which is attached to the user while performing complex activities. The signal which is measured by the HAR sensor is typically a movement related signal (i.e. position, orientation or velocity). Recently the number of relatively-anchored approaches have been gradually increased due to the presence of low-cost sensors and smartphones. The idea behind these approaches is to extract features from the measured signal and then analyze their characteristics or classify them with pre-trained classifiers, such as in [7], [8], [9] and [10]. The advantage of relatively-anchored approaches is that the HAR sensor can provide accurate measurements of the activity, and the recognition is actually a direct classification problem in the feature space of the signal. In contrast, the main disadvantage of these approaches is that the users have to wear disagreeable devices on their body and sometimes a cable-based connection is used during performing activities, which restricts the users in their movements.

B. Absolutely-anchored approaches

The latter disadvantage of relatively-anchored approaches can be overcome by using absolutely-anchored approaches. Due to the nature of them, the HAR sensor is placed in the space where the activities are performed. However, this advantage is a disadvantage at the same time because the space of HAR is limited to a small part of the environment – obviously around the HAR sensor. Absolutely-anchored approaches are mostly based on computer vision technology. The approaches can be classified based on modeling variations in time. There are various direct classification methods, where usually local or global features are extracted from the image which is classified by pre-trained classifiers. These methods do not take the time factor into account, and the HAR is performed over each, consecutive frames, such as in [11], [12], [13] and [14]. Additionally, there are many approaches which based on temporal state-space models. Here, the temporality is modeled as a particular dimension, and a suitable image representation is assigned to the observations in time, such as in [15], [16], [17] and [18].

In this paper, we introduce an absolutely-anchored, vision-based approach for HAR. Our method does not need any installation or calibration of HAR sensors, and nothing must be hung to the user. Our HAR sensor is a simple webcam so the activities should be recognized in a fixed location around the screen. We will demonstrate that our approach is efficient enough to run in embedded or mobile environments, so that the activity space can be expanded in the future. The rest of this article is organized as follows. In section 3, we define the basis of HAR, where a face and facial feature tracking scheme is described. Facial features are the input of our 6 DoF head pose estimation approach. In section 4, we give an efficient spatiotemporal representation to the head movements and then in section 5 a DTW-based head movement recognition method will be presented. Finally, we summarize our results in section 6 through detailed test experiments.
III. FROM FACE TO HEAD POSE

A. Face Detection

The recognition of head movements as human activities is implemented by estimating the human head pose in the 3-D space. As it is mentioned, our system uses monocular images so the head pose must be estimated based on the 2-D projection of a 3-D head. First of all, we need several feature points from the face. The geometric relation between the 2-D and 3-D facial features should be considered invariant, which means that we need several fix points from the face (i.e. corners of eyes, or nose tip). The extraction of facial features is overtaken by face detection. The face is detected by the well-known algorithm of Viola and Jones [19], which can be considered as a de-facto standard of the problem. There are three main concepts around their algorithm: a new image representation (integral image); using variants of AdaBoost learning in order to select features and train classifiers; and a cascade architecture of several weak classifiers. The algorithm is relatively simple and allows us to detect faces in real-time but for efficiency purposes, the face detector is not allowed to run on every frame, it runs only at specified intervals, and the face is tracked by a simple template based procedure between the endpoints of the intervals.

B. Facial Feature Extraction

As we already mentioned, the complex head movements will be recognized based on the head pose in the 3-D space. Although the head pose can be estimated by the appearance of faces but we rather chose a feature-based approach because of several stability considerations. Feature based approaches take the advantage of the geometry of facial features, and they often use deformable models to estimate the head pose, i.e. in [20] and [21]. We also chose a similar approach, which was roughly described by us in [22]. The only exception is that a robust Active Shape Model (ASM) is used in this study instead of the set of Haar classifiers. That is, an ASM is used for extracting facial features from the face. The concept of ASMs are described in [23] and [24] in detail, here we give just a short outline for the notation we use. ASM is a statistical model of several faces and consists of two stages: a model training phase and a model fitting phase. The ASM focuses on the shape, which is the external boundary of the face. An arbitrary \( \mathbf{s} \) shape can be represented by its landmarks:

\[
\mathbf{s} = \{ \mathbf{p}_i \in \mathbb{R}^n; i = 1 \ldots N \},
\]

where the contour is given by the concatenation of \( \mathbf{b}_i \) landmarks and \( n = 2 \) in case of planar objects. So landmarks can be given as \( \mathbf{p}_i = (x_i, y_i) \) and an \( \mathbf{s} \in \mathbb{R}^{2N} \) shape can be represented as a vector with \( 2N \) components:

\[
\mathbf{s} = (x_1 \ldots x_N, y_1 \ldots y_N)^T.
\]

Given several \( \mathbf{s} \) shapes from a previously annotated database, the ASM tries to superimpose them in a common coordinate system. The Generalized Procrustes Analysis (GPA) is used to measure shapes to each other [25], then a statistical model is defined based on the shapes in the common coordinate system:

\[
\mathbf{s} \approx \mathbf{s} + \mathbf{b} \Phi,
\]

where \( \mathbf{s} \) is the mean shape; \( \Phi \in \mathbb{R}^{2N \times t} \) contains the eigenvectors of the \( t \) largest eigenvalues, and the \( \mathbf{b} \in \mathbb{R}^t \) column vector is the scale factor of the principal components (the parameter set of the deformable model). Similar shapes can be given by varying the \( \mathbf{b} \) column vector, and this is the main idea behind the fitting phase, which is essentially the minimization of a cost function. The following parameters are taken into account while finding the minimum: \( \mathbf{b} \in \mathbb{R}^t \) parameter vector; \( \mathbf{T} = [T_x \ T_y] \) translation; \( R(\mathbf{\Theta}) \) rotation; and \( S \) uniform scaling. The model which depends on these parameters is called as Point Distribution Model (PDM). PDM ensures that shapes cannot have an arbitrary layout, they can be formed only similarly to the shapes of the training database. Different kind of ASMs differs only on the strategy of PDM fitting over an input image.

C. Head Pose Estimation

Our head pose estimation method is the solution to a Perspective n-Point problem (PnP problem), so it is based on 2-D and 3-D point correspondences. The 2-D points are resulted by the ASM: they are the \( (x_i, y_i), i = 1 \ldots N \) landmarks of the \( \mathbf{s} \in \mathbb{R}^{2N} \) shape. For clarity, we will denote the 2-D landmarks with a \( \mathbf{C} \) upper index, because they are defined in the camera coordinate system (on the image plane). Furthermore, we have the same number of 3-D model points in an appropriate model coordinate system.
which will be denoted with a $D$ upper index. The relation between landmarks and model points is the following: a $p_i = (x_i^C, y_i^C)$ point is the projection of a $P_i = (x_i^M, y_i^M, z_i^M)$ 3-D model point as it can be seen in Fig. 1.

$$p_i = A[R|t]P_i,$$

which can be written as:

$$\begin{bmatrix} x_i^C \\ y_i^C \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x_i^M \\ y_i^M \\ z_i^M \\ 1 \end{bmatrix}.$$  

where $A$ is the camera matrix which contains the intrinsic parameters: the $f_x$ and $f_y$ focal lengths and the $(c_x, c_y)$ principal point. Intrinsic parameters can be estimated by a chessboard-based calibration routine, and they do not depend on the scene viewed, so $A$ should be estimated only once. The role of $A$ is to project the 3-D points of camera coordinate system onto the 2-D image plane. The joint rotation-translation matrix $[R|t]$ means the extrinsic parameters of the model; this matrix transforms points from the 3-D world coordinate system to the 3-D camera coordinate system. Extrinsic parameters greatly depend on the scene viewed. In our case, the world coordinate system is an appropriate model reference frame and POSIT estimates an $R$ rotation and $t$ translation which brings points from the 3-D model reference frame to the 3-D camera coordinate system. In the very first step, we created a 3-D head model for having point correspondences from a stereo face database [31] (see Fig. 2). The accuracy of the depth map is very high; according to the authors, it is 0.32 mm to each direction which is a good precision to estimate the 6 DoF head pose.

---

1 Distortion coefficients are not included in the equation for the simplicity.
We will not describe the whole method, just the main idea behind. First of all, \( A^{-1} \) to get normalized image coordinates on the left side and to eliminate \( A \) from the right side:

\[
\begin{bmatrix}
u_i \\
v_i \\
w_i
\end{bmatrix} = \begin{bmatrix} r_1^T / t_z & t_1 / t_z \\
r_2^T / t_z & t_2 / t_z \\
r_3^T / t_z & 1
\end{bmatrix} P_i. \tag{6}
\]

Defining \( r_i = [r_{i1}, r_{i2}, r_{i3}]^T \) and then dividing each \([R|t]\) elements by \( t_z \) gives us the following:

\[
\begin{bmatrix}
u_i \\
v_i \\
w_i
\end{bmatrix} = \begin{bmatrix} r_1^T / t_z & t_1 / t_z \\
r_2^T / t_z & t_2 / t_z \\
r_3^T / t_z & 1
\end{bmatrix} P_i. \tag{7}
\]

Eq 7 provides us a linear system of equations. We do not deal with the third (which belongs to \( w_i \)) at the moment, but transposing the first two equations lead up to:

\[
[u_i \ v_i] = P_i^T \begin{bmatrix} r_1^T / t_z & r_3^T / t_z \\
t_1 / t_z & t_2 / t_z
\end{bmatrix}. \tag{8}
\]

Now considering all of the 3-D model points we can extend the previous equation as below:

\[
\begin{bmatrix}
u_1 \\
v_1 \\
\vdots \\
u_N \\
v_N
\end{bmatrix} = \begin{bmatrix} x_1^M & y_1^M & z_1^M & 1 \\
x_2^M & y_2^M & z_2^M & 1 \\
\vdots & \vdots & \vdots & \vdots \\
x_N^M & y_N^M & z_N^M & 1
\end{bmatrix} \begin{bmatrix} r_1^T / t_z & r_3^T / t_z \\
t_1 / t_z & t_2 / t_z
\end{bmatrix}. \tag{9}
\]

where the matrix of \( P_i \) points is often called as an \( M \) model matrix. The \( r_1, r_2, t_x, t_y, \) and \( t_z \) pose parameters can be retrieved straightforward by solving the equations. Notice that the \([R|t]\) pose is fully defined, since the \( R \) matrix is orthogonal \((r_3 = r_1 \times r_2)\). Note that, at least four non-coplanar points are required; otherwise, the \( M \) model matrix is singular. This part of POSIT is referred as Pose from Orthography and Scaling (POS). There is a question what we have not answered yet: the role of the third equation in Eq 7 (which belongs to \( w_i \)). POSIT tries to find the pose for fixed values of \( w_i \) and assumes that the image points of perspective projection are the same as scaled orthographic projection so that \( w_i = 1, i = 1 \ldots N \). Then it solves the linear system of equations under this assumption and it results an approximation for the head pose. This approximation is used as an input in the next iteration to refine the pose until it does not converge to the valid head pose.
IV. REPRESENTATION OF HEAD MOVEMENTS

There have been some studies recently for estimating head pose as the basis of a HAR system, i.e. [32], [33], [34] and [35]. Although these works define several action primitives, however, the methods lack HAR and do not allow us to make complex decisions. In this section, we give the spatiotemporal representation of head movements. The representation is based on facial primitives defined previously (landmarks and head pose), and the head movements are considered as time series. The expected length of the movement is around three or five seconds, but there are no limitations. Obviously, if users do not perform two head movements consciously, then it is not expected that they will repeat the action the same way and beat. So head movements can vary in time or speed which indicates the need for a non-linear alignment to determine the degree of similarity. The trajectory of head movements is based on the head pose over the consecutive frames, so the 6 DoF head pose means the spatial dimension of our representation. Moreover, the movements should be segmented in time as well. Therefore, the temporal boundaries of the movements are also estimated by the procedure written below.

A. Trajectory of head movements

Space, where we try to recognize head movements, is fixed in the environment of the screen of a PC or mobile. We assume that users are sitting before the screen, so they are not changing their place during performing activities. Considering the spatial representation, we only deal with the rotations around the $X$ and $Y$ axes, the position and the rotation around $Z$ is not taken into account. Supposing that we have an image sequence of a head movement, together with the estimated \( \mathbf{a}_i \) attributes (observations) over every frame. If the first element of the sequence belongs to the $\tau$ timestamp and the sequence consists of $n$ frames, then we have an $a = (a_\tau, a_{\tau+1}, ..., a_{\tau+n})$ time series about the movement. Here $a \in S$ is an ordered $n$-tuple of the HAR problem ($a \in S$ is a sample of a probability distribution). The $a_i$ attributes are the element of the $\mathcal{F}$ feature space: $a_i \in \mathcal{F}, i \in [\tau, \tau + n]$. Now each $m \in \mathcal{M}$ head movement can be defined as below:

$$m = (l, a_\tau, a_{\tau+1}, ..., a_{\tau+n}) = (l, a), \quad (10)$$

where $l \in L$ is a label of the performed movement (e.g. “head shaking” or “nodding”). For an unknown $m \in \mathcal{M}$ movement, the $l \in L$ label is predicted by the $a_i \in \mathcal{F}$ attributes. The determination $a_i \in \mathcal{F}$ attributes are based purely on the head pose. The $\mathbf{k} = (0,0,1)$ unit vector of the model reference frame is rotated by the orientation of the current head pose. The rotated vector over the $i$th frame is denoted by $\mathbf{k}_i$, which is the direction vector of the head. Then the $a_i \in \mathcal{F}$ attributes are defined as below:

$$a_i = k_i - k_{i-1}. \quad (11)$$

That is $a_i$ is the difference between the current and the previous direction vector, so it represents the variation in orientation over the $X$-$Y$ plane between the neighboring frames. Furthermore, $a_i$ is close to zero if the head is standing still and it has a high value if the head movement is intensive. All of the attributes are shifted by $a_\tau$ to be independent of the start position, so head movements are defined in the final form of below:

$$m = (l, 0, a_{\tau+1} - a_\tau, ..., a_{\tau+n} - a_\tau) = (l, a'). \quad (12)$$

Fig. 3 shows the variation in orientation between 2-2 frames. The arrow around the nose indicates the direction and the amount of variation.

![Fig. 3. Visual representation of the variation in head movements.](image-url)
B. Temporal boundaries of head movements

Now we have spatial representation, but still, we did not localize head movements in time. The attributes should be divided into \( I = [t_{\omega}, t_{\omega}] \) consecutive, non-empty and non-overlapping time intervals. A simple state transition model was defined for this reason:

- If there is no head movement, and the variation in orientation is over of a given threshold value, then the head movement started,
- If the head is moving and the variation in orientation is over of a given threshold value, then the machine remains in moving state,
- If the head is moving, and the variation in orientation is below of a given threshold value, then the head movement ended,
- Otherwise, there is no head movement.

The final algorithm for the spatiotemporal localization of head movements is below:

1. Estimating the variation in orientation compared to the previous timestamp and GOTO 2,
2. Temporal localization of the current head movement based on the variation in orientation:
   - If the head movement started or the head is moving then GOTO 3,
   - If the head movement ended then GOTO 4,
   - Otherwise GOTO 1,
3. Spatial representation of the head movement: producing the \( \alpha \in \mathcal{F} \) attributes for estimating the similarity and GOTO 1,
4. Estimating the similarity between the current attribute set and several previously determined attribute set, then GOTO 1.

V. HEAD MOVEMENT RECOGNITION

In this section, we give an explanation about the matching of two different time series. The method is an extraction of our previous study [36] with some improvements. As it is already mentioned, the head movements can be considered as time series which are performed in a different way, tempo or time. The matching of time series is carried out by the algorithm of Dynamic Time Warping (DTW), see in [37] and [38]. DTW can be considered as a generative model, so we are looking for the \( p(a, l) \) joint distribution, where \( a \in S \) is a time series, and \( l \in L \) is a label (e.g. “nodding”). An \( l \in L \) label can be predicted by an \( a \in S \) time series. So recognizing an \( m \in \mathcal{M} \) head movement means the determination of an \( l \in L \) label based on some time series. DTW determines the optimal (least cumulative distortion or distance) alignment between a pair of time-series. It transforms the time axis such that it optimally maps one of the time series (test) onto the whole or a part of the other one (reference). DTW computes point-by-point correspondences (by a warping function), and it also gives a distance-like quantity between a pair of time series. The so-called DTW distance is semimetric, because it satisfies the positivity, positive definiteness, and symmetry, but not necessarily the triangle inequality. Let us suppose that we have two time series given by their attribute set:

\[
X = (x_1, x_2, ..., x_N) \quad Y = (y_1, y_2, ..., y_M)
\]

where \( X, Y \in \mathcal{S} \) the time series and \( x_i, y_j \in \mathcal{F}, \ i = 1 \ldots N, j = 1 \ldots M \) are the attributes of the \( \mathcal{F} \) feature space. For simplicity, we will use \( i \in [1: N] \) index in case of \( X \) and \( j \in [1: M] \) in the case of \( Y \).

DTW defines a \( c: \mathcal{F} \times \mathcal{F} \rightarrow \mathbb{R}_{\geq 0} \) metric for measuring the similarity between any pair of \( x_i \) and \( y_j \) attributes. The most common choice for \( c \) is the Euclidean distance. DTW aligns the attributes of time-series by minimizing a cost function based on the \( c \) distance between the pairs of \( x_i \) and \( y_j \) attributes. So that DTW computes a \( C \in \mathbb{R}^{N \times M} \) cost matrix between any \( x_i \) and \( y_j \):

\[
C(i, j) = c(x_i, y_j) \geq 0.
\]

Time series can be aligned along multiple paths in the \( C \) matrix and these paths often called as warping path (or warping function) and denoted by \( w(t) \). The goal is to find a path with minimal cost in the \( C \); it is the optimal alignment between the two time series. DTW defines the optimal alignment as the solution to an optimization problem. The \( w(t) \) warping function can be formed by a series of elements:
\[
w(t) = \left( w_x(t), w_y(t) \right),
\]

where the \( w_x(t) \in [1, N] \) and \( w_y(t) \in [1, M] \) functions map the time indices of \( X \) and \( Y \) respectively and \( t \in [1:T] \) is the index range of the common axis. For any \( w(t) \) warping function, the averaged accumulated distortion between two time series can be given as below:

\[
d_w(X, Y) = \frac{1}{M_w} \sum_{t=1}^{T} \mathcal{C} \left( w_x(t), w_y(t) \right) m_w(t),
\]

where the \( \mathcal{C} \left( w_x(t), w_y(t) \right) \) term is the distance between the \( x_{w_x(t)} \) and \( y_{w_y(t)} \) attributes warped over the common time axis, \( m_w(t) \) is a non-negative weighting coefficient, and \( M_w \) is a normalizing constant (usually \( M_w = M + N \)). The \( m_w(t) \) and \( M_w \) are necessary because there are more than one warping paths between two time series, but we need the one with minimal cost. Let \( w(t) \) be the optimal warping function between the time series \( X \) and \( Y \), then \( w(t) \) chooses the indices of \( X \) and \( Y \) which minimize the overall accumulated distance:

\[
D(X, Y) = \min_w d_w(X, Y).
\]

The warping function can be determined by finding paths in \( D \) matrix, and there are multiple techniques for that. Usually, several optimizations are also used, which are mostly based on the characteristics of warping paths: monotonous condition, continuity condition, boundary condition, adjustment window condition and slope constraint condition. The ideas are originally described by Sakoe and Chiba [39]. These conditions allow us to restrict the searching space of warping paths to a narrow space in \( D \).

Fig. 4. shows an alignment between two head movements (drawing letters of \( k \)) performed by two different users. Head movements are very similar to each other because the optimal warping path (red curve in the lower figure) passes through the cost matrix diagonally, in a relatively straight line. 

![Figure 4](image-url)
A. Human Activity Database

At this point, we can align two time series to each other, but we would like to recognize much more type of head movements. A DTW based classifier was created with 26 classes. Let \( M_i \subset \mathcal{M} \), \( i \in [1:26] \) a head movement class, where \( \mathcal{M} \) is the set of all head movements. Each \( M_i \) class contains head movements with similar trajectory of head drawn letters of the English alphabet (i.e. the trajectories of \( a \) or \( b \) characters):

\[
M = \{M_i \subset \mathcal{M}\}_{i \in [1:26]},
\]

where each \( M_i \subset \mathcal{M} \) classes contain exactly 100 head movements, so \( |M_i| = 100 \). The classes are disjoint and non-empty, furthermore \( \bigcup_{i \in [1:26]} M_i = \mathcal{M} \). A kind of ordering can be defined between the head movements in each class, and this ordering is based on the pairwise DTW distance of each \( m_{i}^{a}, m_{i}^{b} \in M_i \) elements (\( a, b \in [1:100], a \neq b \)). For efficiency purposes, a pivot element is appointed in each class, because in this way only 26 cost matrix need to be estimated during the classification. In fact, we use a simple 1-NN classifier. The pivot element \( m_{i}^{p} \in M_i \) of the \( i \) th class can be defined as below:

\[
m_{i}^{p} = \min_{D} D(m_{i}^{a}, m_{i}^{b}),
\]

where \( D(\cdot) \) is the pairwise DTW distance between the head movements of \( M_i \). That is, \( m_{i}^{p} \) is the element which minimizes the distance to all of the elements in its \( M_i \) class. That is, \( m_{i}^{p} \) can be considered as a kind of cluster centroid. Fig. 5. below illustrates several moments of the head movement recognition, especially the visual representation of trajectory which only based on the head pose.

Fig. 5. Several moments of head movement recognition. Head movements are colored by a red-green transition, where the color of the start point is red and the current point of the trajectory is green. The shape of trajectory only depend on the head pose, more accurately the rotation around the \( X \) and \( Y \) axes.
VI. EXPERIMENTAL RESULTS

The performance of the proposed HAR system is evaluated separately regarding the head pose estimation and head movement recognition. It is because the head pose means the basis of our system, so its accuracy fundamentally determines the precision of head movement recognition.

A. Head Pose Estimation

During our head pose test cases we were focused on measuring the yaw (rotation around Y) and pitch (rotation around X) errors, because these coefficients have the highest effect on head movement recognition. Our head pose estimation approach was evaluated over the Biwi database [40], which contains Microsoft Kinect scanned faces. Biwi covers a relatively large interval both for pitch and yaw. Fig. 6. shows the absolute yaw and pitch errors in function of yaw and pitch separately. So the absolute yaw (and pitch) error is illustrated only in function of ground truth yaw (and pitch) values.

The negative values mean the upward direction in the case of pitch and the leftward in the case of yaw. It can be seen that the absolute error of yaw and pitch is increasing along to the ends of the range. In the case of yaw, $[-40^\circ, +40^\circ]$ is the range where the absolute error is below 10 degrees, here the yaw error can be considered low enough to perform head movement recognition. Yaw reaches its minimum in the range of $[-12^\circ, +12^\circ]$, where the absolute error is below 5 degrees, here we have the chance to have a very accurate head movement recognition. Although, we can define a yaw range of 80-100 degrees length, where we expect that head movements can be recognized with high precision.

The pitch is a little bit complicated than yaw. Here the range $[-32^\circ, +40^\circ]$ is that where the absolute pitch error is below 10 degrees and pitch reaches its minimum in the range of $[-6^\circ, +16^\circ]$. The characteristics of the error function are not symmetric to the zero; it is shifted to upwards. It means that we will have a higher error for people who is looking upwards.

![Fig. 6. Absolute yaw and pitch errors in function of yaw and pitch separately. The absolute yaw (and pitch) error is illustrated only in function of ground truth yaw (and pitch) values.](image)

It is also interesting to visualize yaw and pitch error in function of both yaw and pitch. This visualization provides much more information, because, in this way, the absolute errors are dependent on pitch and yaw at the same time. Fig. 7. shows the surfaces of this visualization. An $L$ grid was defined in the yaw-pitch coordinate system with the resolution of $N \times M$. The distance between the grid points is 2, so the absolute errors were accumulated over this environment of each grid points. The empty areas were interpolated. For example, in case of 12° of pitch and -40°of yaw:

- The average pitch error is 7.44°,
- Moreover the average yaw error is 10.06°.

It can be seen that the pitch error begins to grow drastically by raising up the head. Moreover, there is also some correlation between the absolute pitch error and yaw rotation because we measured the highest pitch error in the bottom-left and up-left regions.
Fig. 7. Absolute pitch and yaw errors in function of yaw and pitch at the same time. Errors were represented along the Z axis. The ground truth measurements of pitch and yaw were represented along the X and Y axis.

B. Head Movement Recognition

As the very first step of evaluating head movement recognition, we took a measurement regarding the relationship between the pivot element and the remaining elements of its class. As it is mentioned, pivot elements minimize the distance within their classes, and they were chosen from 100 head movement. Fig. 8. shows the average distance and the corrected standard deviation of pivot elements compared to the other elements in $M_i$, so according to the definition:

$$
\mu_i = \frac{1}{100} \sum_{j=1}^{100} D(m^*_i, m^j_i)
$$

$$
\sigma_i = \sqrt{\frac{1}{99} \sum_{j=1}^{100} (D(m^*_i, m^j_i) - \mu_i)^2}.
$$

Fig. 8. shows that the letters of $E$ or $V$ are very close to the pivot element, so all elements of these classes are in “sphere” with a narrow radius. In contrast of this, we got significant DTW distance for the $Q$ and $T$, where the average DTW distance to the pivot element is three times higher than the previous case. It can be possible to conclude the accuracy of HAR system from Fig. 8: we can believe that the recognition will be less accurate in cases where the average DTW distance is higher, but we will see that this is not always true.

Fig. 8. Average DTW distance and standard deviation from the pivot element in case of every class.
The method was also tested on an independent database. The classes of the test database contained exactly 100 head movements, so the method was tested on 2600 head movement. For each \( m \in M \) element of the test database, we estimated the \( D(m^p, m) \) distance \((i \in [1:26] \text{ as } j \in [1:100])\) and \( m \) was classified to that \( M_i \) class where the distance was minimal. Fig. 9. shows the precision and recall of each class. It can be seen that the classes \( B, E, I, P, U, V, W \) and \( Z \) have the minimum false positive rate. Furthermore, only a few elements of \( F, I, J, K, S, T, V, W, X, Y \) and \( Z \) were misclassified as false negatives. The average precision of our system is 0.81, and the average recall is 0.83 which can be considered quite good results because it provides the guarantee for the low number of false positives and negatives at the same time. However, in the case of \( A, K, L, Q, T \) and \( Y \) classes we will have false positives and at the same time, we will have false negatives in case of \( A, D, E \) and \( H \) classes.

We also analyzed the relationship between the \( F_1 \) measure and the average length of the movements. Fig. 10. shows \( F_1 \) measure for the previous test database. The best recognition is expected for the \( R, U, J, I, W, F, V, S, X \) and \( Z \) classes. It should be highlighted that it was not any error in case of the \( Z \) class, all of the 100 movements of the test database were recognized correctly. The \( A, Q, L, T, D, H, N, K, M \) and \( Y \) classes have the worst recognition. A possible explanation is that some of these classes have similar characteristics considering their shapes (i.e. an \( a \) is written similarly to \( d \) and so \( e \) and \( l \)). We also analyzed the correlation between the \( F_1 \) measure and the average length of the movements, but we did not notice high correlation. In Fig. 10. there is a secondary vertical axis which measures the length of the movements. The measure of this axis is the number of attributes. We can say, that smaller head movements can be recognized on similar accuracy as longer ones and vice versa.
VII. CONCLUSIONS AND PLANS FOR THE FUTURE

In the above, we proposed and described a complete system for human activity recognition which is suitable for recognizing various, complex head movements. Our solution is an absolutely-anchored, vision-based approach which does not need any installation or calibration of HAR sensors, and nothing must be hung to the user. Our HAR sensor is a simple webcam, and the activities should be recognized in a fixed location, so the space where the activities are performed is around of PC’s screen. The HAR is based on a 6 DoF head pose estimation procedure. During our test experiments, we found that the approach could work in wide range of head movements. The typical length of a movement was 2-3 seconds, but there are no limitations regarding this. The average CPU time of the overall procedure is less than 25 milliseconds beside 640x480 resolution on an Intel Core i5 2.4GHz PC. Based on our results, we think that our approach can run in embedded or mobile environments. Our primary goal is bound to the assumption; we are going to expand the activity space by adapting our approach in mobile environments, which can provide ease-of-access to information systems. Furthermore, it strengthens the link between the research areas of info-communications and the cognitive sciences, because we have the intention to represent human activities better, process or recognize them more efficiently, and last but not the least it can be very interesting to separate conscious human activities from unconscious ones which can open new chapters in our researches.

VIII. REFERENCES


Developing a 3D virtual library model based on the ancient Library of Alexandria

István Boda, Erzsébet Tóth, István Csont, László T. Nagy
Department of Library Informatics, Faculty of Informatics
University of Debrecen
Debrecen, Hungary
boda.istvan@inf.unideb.hu, erzsebet.toth@inf.unideb.hu, ist.csont@gmail.com, t.nagy.laszlo@hunep.hu

Abstract—In our paper we describe in detail a three-dimensional virtual library model inspired by the system and presumable content of the ancient Library of Alexandria. In addition, we introduce one the possible implementation of our virtual library model based on the VirCA system and web technology (such as HTML, CSS, PHP etc.). In our model, we used the classification system elaborated by the famous librarian, poet and scholar Callimachus. Focusing primarily on the field of Greek poetry and drama in the 3rd century BC, in the current implementation of our model we selected those ancient and well-known authors whose texts, completed with short biographical data, may be a good starting point for the content of the virtual library. Currently we prefer various English translations of the selected ancient Greek texts with an intention to gradually improve the content of the library with other texts and multimedia materials (other translations, paintings, images of papyri, sculptures, commentaries, selected quotations, dictionary and encyclopedia entries, related texts from the ancient and modern culture etc.) which form background knowledge for the interpretation of the processed texts. We describe a scale-free network model, the so-called Alexandrian model of network, which the organization and future development of the virtual library is based on. Finally, we discuss the main purpose of our virtual library model and its possible relationships with the current cognitive infocommunications research.

Keywords—Library of Alexandria; three dimensional virtual library model; knowledge base of literary content; VirCA system; spatial hypertext; scale-free networks; Alexandrian model of network

I. WHY DID WE CHOOSE THE LIBRARY OF ALEXANDRIA?

The Library of Alexandria is definitely the most prominent and celebrated collection of classical (Greek-Roman) antiquity, but nothing was left from it — it is known only from the contemporary writers’ works. Its fame has survived in the later ages too, for centuries it has been embodied a library ideal representing a symbolic force for modern public libraries. To formulate an appropriate answer to the above mentioned question we can say that it was an important aspect for us that all the ancient written materials collected in this library were eternal values for human culture and erudition. We also took into account that the study of the poetical Greek works remained an open issue till nowadays. It can be also realized that Greek mythology provides a kind of ‘common language’ and symbolic basis for the better interpretation and understanding of the ancient poetical works. Therefore we decided to focus on the field of ancient Greek poetry and drama once available in the Library of Alexandria. In our research we intended to present these ancient masterpieces in new formats and environment. In this sense it was crucial for us to attach various verbal and multimedia metadata to the English translations of the chosen ancient Greek texts in the three dimensional environment.

The Musaeum or Mouseion (“Institution of the Muses”) at Alexandria was a research institution built by Ptolemy I Soter and by his son, Ptolemy II Philadelphus at the turn of 4th and 3rd centuries BC. Ptolemy I Soter initially wanted Theophrastus, Aristotle’s favored pupil and leader of the Peripatetic School, to manage and control the affairs of the Mouseion. Soter intended to establish the Mouseion, at least in part, by transplanting Aristotle’s Peripatetic School from Athens to Alexandria. The prestige of accomplishing this idea would have been enormous, and would have attracted other scholars much easier. Transplanting Aristotle’s school, Ptolemy I Soter would have reinforced Alexandria’s cultural ties to Alexander the Great; not only was the city founded by him, but it would have also participated in his intellectual tradition by continuing the famous institution of his beloved tutor, Aristotle. Although the Peripatetic School did not actually move to Alexandria, it had a great influence on the Mouseion and the Great Library. It is assumed that the Library of Alexandria has received some of the private library of the Peripatetic school’s founder, Aristotle himself.

The Greek word Mouseion means “Seat (Institution, Shrine, Temple etc.) of the Muses” (i.e. a place devoted to the nine Muses) which designated the home of music or poetry, a philosophical institution or a place of contemplation such as Plato’s renowned Academy in Athens. Originally a Greek mouseion was a temple sacred to the Muses, so it was a purely religious establishment. The Alexandrian Mouseion mixed the religious and intellectual features of similar Greek institutions with the religious and bibliophilic characteristics of analogous Egyptian institutions. The latter features imply that the Museion was based on the Egyptian tradition of placing libraries within religious temples [1]. As far as we know, hundreds of studies have emphasized the nearly religious foundation of the Museum, as an expression of the human longing towards the wholeness and the unity of knowledge [2].
We can observe that the physical structure of the Mouseion not only reflected Aristotle’s division of knowledge into observational and deductive topics, but it was planned in a way to express and support Aristotle’s peripatetic ideal of scholarship. The main Mouseion building and the Library building were joined by and surrounded with a network of paths, colonnades, and courtyards. Botanical gardens and zoological displays served the scholars’ entertainment and study needs. In addition, an outdoor amphitheatre called the exedra was the integral part of the Mouseion building.

Use of the Latin form, *museum*, seems to have been restricted in Roman times mainly to places of philosophical discussions. The modern usage of the word “museum” is originated from the Greek version, *Musaeum*. The great Musaeum at Alexandria with its college of scholars and its library can be considered more a prototype university than an institution which has to preserve and explain material aspects of the cultural heritage. Demetrius of Phaleron worked on the library’s initial organization, who had a good knowledge of the achievements of the philosophers’ libraries at Athens. He organized both the museum and the library in faculties, with a president-priest at the head. His tutors were Aristotle and Theophrastus. He is considered to provide a relationship between Aristotle’s Peripatetic school and the Mouseion in Alexandria. According to our opinion, this initial structure created by Demetrius of Phaleron also reflects the above mentioned university concept which has been realized in the great Musaeum at Alexandria [1]. In fact Demetrius received large sums of money for the purchase or copying of Greek literary works, and occasionally even for the acquisition and translation of significant writings in foreign languages [2]. The Ptolemies offered scholars free board, lodging, servants, tax exemptions, and good salaries for life to attract them to the Mouseion. Some of the well known scholars of the Mouseion were Strabo, Zenodotus, Aristophanes, Eratosthenes, Euchid, Archimedes etc. We suppose that the Mouseion Library was somewhat similar to our modern academic libraries because the part of the library placed within the Mouseion was intended only for the use of the scholars who stayed there.

At that time, being as one of the successor states, Egypt belonged to the extended empire of Alexander III of Macedon, whose name is commonly known as Alexander the Great. A subsidiary “daughter library” was founded about 235 BC by Ptolemy III (Euergetes) in the Temple of Serapis. It is believed that it included copies of literary works intended for general use by people who did not have access to the library of the Mouseion. In addition, the Temple of Serapis was a public building and open to all members of the public, so literate Alexandrians were the intended user group of the Serapeum library. It has been estimated that the total number of papyrus scrolls available in the Serapeum was 42 800. Though the Serapeum served a different group of users and was located in a different place than the Library of the Mouseion, the two libraries were headed by the same authority. The same staff and policies served both, and therefore they are considered to be an institution with two branches. In Roman times the “daughter library” was the main literary center. From this we can arrive at the conclusion that the creation of the Library of Alexandria and its attendant institutions were absolutely essential contributions toward making Alexandria into an intellectual and cultural center [1].

It can be somewhat contradictory to point out that the historical sources on the Library of Alexandria are hardly much enough. Therefore we can say almost nothing certain about it; “where and how the papyrus scrolls were stored; what dimension its collections really had; what role the other public library of the town, the Serapeum library had in Alexandrian cultural life; if books continued to be added with the same regularity after the death of Ptolemy III Euergetes etc” [2]. Even the information about the end of the collection covers a period of six centuries, from the age of Caesar to the age of the prophet Muhammad. Given from this fact a flood of publications were filled with hypotheses and speculations about the fate of library.

The Library of Alexandria formed one of the most significant parts of the Mouseion funded by the royal treasury. As far as the Mouseion and its library are concerned, we can state that they played an essential role in enhancing the prestige and influence of the royal house. Note that the main museum and library were placed at this time in the palace precincts, in the district named the Bruchium. Its mission was to gather all the Greek documents, so its ambition was to achieve completeness in written Greek literature. Soon it has amassed several thousands of papyrus rolls in its holdings. During its most flourishing period it is said to have included 490 000, or, according to another authority, involving all duplicates, as many as 700 000 volumes. Furthermore it can be mentioned that the collection included not only all Greek documents but also translations into Greek from the other languages of the Mediterranean countries, the Middle East and India. We are sure that the library of the Mouseion contained mainly Greek documents; the only translation recorded was the Septuagint (the earliest extant Greek translation of the Old Testament from the original Hebrew).

The Ptolemies wished to acquire the best, the most original, the most authoritative copies of works, and they were willing to purchase, borrow, or plunder to obtain them. They employed ethically questionable means also for procuring materials. For example during a famine in Athens ambassadors from the Library of the Mouseion compelled the sale of valuable original manuscripts owned by that city in exchange for food. Ptolemies sent people out to buy books, looking especially for rare texts and libraries which might be bought. Because older versions of the manuscripts were preferred to newer copies (older versions were considered to be more authentic and less likely to include mistakes), a miniature industry sprung up that manufactured “old” texts. In addition to purchasing books, the Ptolemies acquired works through confiscation. It is reported that upon entering the Alexandrian harbor, ships were examined closely, and any books they were carrying were plundered. A copy was made and given to the original owner, but the original manuscript was preserved for the Great Library. Using such arbitrary means the Great Library has created a large collection of standard texts of Greek classics.
The reason for this is that the Ptolemies had a concept of policies, goals, and procedures for obtaining documents, then they actually set that concept into practice and acquired scrolls for their Great Library, arising from these we can conclude that the Great Library of Alexandria showed the characteristics of a true “collection”. The library’s editorial program incorporated various elements which were as follows: namely the establishment of the Alexandrian canon of Greek poets, the division of works into “books” as they are now known (probably they corresponded to the standard length of papyrus rolls), and the gradual introduction of systems of punctuation and accentuation.

Several stories survived about the utter destruction of the Library of Alexandria which could be tracked in various myths. Of course some of the events caused damage in the holdings of the library. For example during the siege of Alexandria by Julius Caesar (in 48 BC) the greater part of the library was damaged badly by fire; but it was later replaced by the collection of Pergamos which was offered to Cleopatra by Mark Antony. The other part of the library was located in the temple of Serapis, where it remained till the time of Theodosius the Great. When this emperor allowed all the pagan temples in the Roman empire to be destroyed, the remarkable temple of Serapis was damaged badly by the Christians in AD 391. We mention that by this time, Christianity has been regarded as the official religion of the Roman Empire. The museum and library continued to exist for many centuries but were destroyed in the civil war that happened under the Roman emperor Aurelian in 272 AD. In 2002 the Egyptian government inaugurated a new library building, the Bibliotheca Alexandrina which was established exactly on the site of the previous ancient institution [1, 2, 4, 5].

II. STRUCTURE OF THE CLASSIFICATION SCHEME

As preliminaries to Callimachus’ hierarchical system we can mention Zenodotus’ “subject/room” and alphabetization methods. He was the first director of the Library of the Mouseion appointed by the royal court. At this time, the director held a priestly position besides doing his secular duties, which is perhaps an influence coming from Egyptian traditions. Arising from the fact that an early Greek mouseion operated also as a religious establishment. In later times the director was appointed by the Roman Emperor, but the duties of the position remained mainly the same. Zenodotus introduced a rudimentary organization system where texts were assigned to various rooms based on their content. Zenodotus first made an inventory of the library’s holdings which he then divided into three major categories. The first category contained history books, edited and standardized literary works, and new works of Ptolemaic literature. The second one included materials used for comparison and in the creation of the standardized works mentioned above. This category also contained letters and maps. The third group comprised original works in foreign languages, many of which had been translated into Greek, and which, in translation were involved in the first category. Within each of these categories, Zenodotus arranged works alphabetically by the first letter of the name of their author. The principle of alphabetic order was first introduced by Zenodotus. He attached each work’s author, title, and subject to the end of each scroll, so that materials could be efficiently reshelved to the area in which they had been classified, but also so that library users did not have to unroll each scroll in order to see what it contained. It seems to be very impressive for us that it was the first recorded use of metadata, which can be considered as a landmark in library history.

However as the size of collection grew, Zenodotus’ methods became less and less effective. A more efficient scheme of classification had to be found out. So while Zenodotus made a good start at organizing the Library of Alexandria, in order to finish this task, a scholar of encyclopedic knowledge and erudition as well as of infinite energy was needed [1]. Such a person was found in Callimachus who was the third director of the library between 260 and 240 BC. Callimachus was born in Cyrene in Libya, but he spent the greatest part of his life at the Ptolemaic court in Alexandria. His patron was Ptolemy II, and Callimachus continued his activity in the era of Ptolemy III (246-222 BC). According to the Byzantine lexicon Suda, Callimachus was a famous grammarian and created more than 800 books in verse and prose. Ancient authors considered him as one of the greatest Hellenistic poets [2]. Under his leadership the library catalogue was created, resulting in the so called pinakes (tablets). Its name originates from the appropriate parts of the catalogue on papyrus leaves which were stuck (probably written) on wooden tablets. Those tablets were placed above the library cases or shelves to support a search for papyrus rolls and to make their reshelving more flexible. The full title of the catalogue was the following: Tables of Those Who Have Distinguished Themselves in Every Form of Culture and of What They Wrote. It was one of the first known documents that listed, identified, and categorized a library’s holdings. Within the pinakes, Callimachus gave a list of works alphabetically by author and genre. The very fact that he took Zenodotus’ organisational principle further emphasizes for us that the Great Library was an organized collection [1, 4].

It is worth mentioning that pinakes as a Greek term can be used in the sense of ‘list’ or ‘register’. In the 5th and 4th centuries BC there was a well-established tradition of compiling lists of priests, victors and dramatists. Aristotle also wrote different works in the form of lists, pinakes that have not survived. Only from quotations by later scholars we know about his lists which were as follows: Winners at the Olympic Games, Winners at the Delphic Games, Victories in the dramatic contests of the Dionysia at Athens, and Dramatic Plays. The latter lists recorded the history of Attic drama, because they provided a list of all tragedies, satyr plays, and comedies performed in Athens during the 5th and 4th centuries at the most significant Dionysian festivals. So the pinakes of Callimachus formed part of this tradition strongly supported by the Aristotelian school [6].

1 King Ashurbanipal of Assyria (668 – c.627 BC) applied similar means like Ptolemies to increase his library’s collection at Nineveh. “In a letter (almost certainly from Ashurbanipal) the king orders the scribe to gather tablets, especially those bearing omen texts, from both private houses and temples for his palace collection” [3].
In Greek the „literature” word refers to education or erudition, so Callimachus, being adjusted to the Egyptian ruler’s objective, intended to cover the complete Hellenistic culture in the index list of rolls [4]. The Ptolemies conceived the Library of Alexandria as a kind of repository for the accumulated knowledge of the humanity and as one in which all the written Greek documents could be found and accessed. They seemed determined to follow Alexander the Great’s plans to establish a universal library. Defining their institution as a “universal” library, modern readers are immediately provided with a sense of the scope and priorities the Ptolemies had for their institution; they wished to collect everything. Using the author’s approach Callimachus added a short biography to each author’s name and he described the originality of each document. In modern sense we can say that he “added metadata” to each entry – writing a short biographical note on each author –, which prefaced that author’s entry within his catalogue. It supported to avoid confusion in the works of authors with similar or identical names, but separating materials of the original author and works of namesakes was extremely difficult. Since the title was seldom clear (if it existed at all), therefore he always added the first line of the work to his catalogue and noted how many lines the given work included. A brief summary of the content was also provided about the work. Variations in author’s names or titles were carefully recorded. Using the pinakes a library user could find out if the library contained a work by a certain author, how it was categorized, and where it might be found [1].

The catalogue was composed of 120 books (rolls), and it had registered approximately 200 000 papyrus rolls. Since the library of the Mouseion was practically complete as regards Greek literary works, the catalogue is thought to be a national bibliographical survey of the most significant Greek works. In Greek the “literature” word refers to education or erudition too. It also appears as an ancestor of bibliography and science history because of its biographical notes [7]. Briefly we summarize that Callimachus’ impressive work was a detailed biobibliographical survey of the most significant Greek works. He tried to provide complete and reliable access to the library holdings, and that work consisted basically of literary criticism. Although now lost, it existed in the Byzantine period as a standard reference work of Greek literature. This reference tool was used by later scholars, as can be inferred from the fact that Aristophanes of Byzantium, the great Alexandrine grammarian, published corrections and supplements to the pinakes [1, 2].

The structure reflecting the contemporary scientists’ and literary men’s intellectual system was separated into two levels:

- At the higher level there was a categorization of the writers that provided a ‘skeleton’ of the division.
- At the other level each category was arranged in alphabetic order by author.

Callimachus divided the library documents into two main parts: poets and prosaists. In both categories he created six subgroups separately:

- poets: epic poets, elegists, satirical poets, lyrical poets, tragedians, comic playwrights;
- prosaists: historians, speakers, philosophers, physicians, natural scientists, miscellaneous literature (comprising also writings on cooking and dining) [7, 8].

III. CONTENT IN THE VIRTUAL LIBRARY MODEL

First we decided to implement Callimachus’ hierarchical structure in our virtual library model. We selected three illustrious Greek ancient authors who would stand for each sub-group within the „poets” main category. We followed the same author’s approach as Callimachus used in his catalogue by connecting the author’s image or his bust with the author’s name, and added a short biography to him. Using this approach we imitated the same the content layout once available in pinakes.

We have integrated the collected and edited content into this library model in the following way:

1. We tried to find images closely related to the authors’ life and poetical works. These images covered a wide spectrum of artistic works: statues, graphics, marble reliefs, paintings and maps.

2. Then we gathered and edited a short biography of the selected authors. Referring to their significance in Greek literature, we provided a short overview about their life and listed the most important masterpieces they created. Entries in Encyclopedia Britannica and in Wikipedia online sources were used and checked for this task.

3. Considering the copyright restrictions, we retrieved proper English translations of Greek poetical works. For this aim we could use the electronic version of literary works which were available in various digitization projects e.g. Perseus Digital Library Project (Tufts University)\(^2\), Project Gutenberg, Google Books Library Project, California Digital Library (University of California). Among them we emphasize the importance of Perseus Digital Library Project which has built a retrievable database from the old Greek texts both in native language and in English. They have also developed 3D data models with an intention to incorporate them into Perseus Digital Library. In this comprehensive project standardized schemas are used for describing ancient Greek texts [9]. Then we edited these English translations in one page (A/4) length and in the suitable format.

4. We loaded all of the collected library content into the VIRCA system in a certain logical order and means according to the special characteristics of the software.

---


\(^3\) http://www.perseus.tufts.edu/hopper/ (et passim) (2016-07-13)
IV. IMPLEMENTATION OF THE VIRTUAL LIBRARY IN A 3D ENVIRONMENT

In the implementation of our three-dimensional virtual library model about three years ago [10, 11] we decided to use the excellent 3D presentation and navigation features of the Virtual Collaboration Arena (VirCA) system developed in the 3D Internet-based Control and Communications Research Laboratory at the Institute for Computer Science and Control (MTA SZTAKI) [12, 13]. In the further development of the library model, the 3D environment of the VirCA system remained the base of the implementation, although we added new presentation features to the model (e.g. browser windows provided by the new version of the VirCA system) [14] and we have been developing alternative web-based presentation interfaces as well [15].

The base of the 3D model is a virtual room of the VirCA system which we decorated and ‘furnished’ according to the library content we selected for the presentation — not mentioning the atmosphere or ‘feeling’ which we intended to recreate imagining the ancient, precious environment of the Library of Alexandria. The back wall of the room contains the ‘main title’ of the room – Mouseion – on a background picture representing the ancient Library of Alexandria (see Fig. 1). On the left there are some posters describing the brief history of the Library of Alexandria (with a map of the respective area); on the right we placed a few additional posters which provide the user or ‘visitor’ with some important information of the system of the library elaborated by the famous librarian, scholar and poet Callimachus.

![Fig. 1 The back wall of the room representing the 3D library model](image1)

This part of the room is obviously important for the visitors to form an initial conception of the environment. However, the selected library content is presented in the front part of the room in carefully arranged boxes or ‘cabinets’. Each cabinet contains a selected text inside and has a lamp on its ceiling to illuminate the inner content properly. The cabinets are usually curtained or ‘veiled’ by an appropriate image identifying the content of the cabinet (e.g. by the title of the presented text, the image of the author etc.), but sometimes, e.g. when the text inside the cabinet is presented in a browser window, the front side of the cabinets can be wide open (i.e. without any ‘curtain’ image) so that their inner content can be seen (see Fig. 2).

![Fig. 2 Cabinets on the front wall of the virtual room with or without ‘curtain’ images](image2)

Each cabinet has a label which identifies a corresponding category (e.g. epic poets, elegists, lyrical poets etc.) selected from the ancient classification system of Callimachus. In addition to the texts which are available inside the cabinets, we attached on both sides a maximum of four relevant images to each cabinet in order to help the users find the cabinet they are looking for more easily (if possible, “at first glance”), or understand the represented content more clearly. We called the sequence of images ‘image corridors’ because they can serve as a kind of ‘navigator’ the main function of which is to direct the user to the corresponding cabinet (see Fig. 3).

![Fig. 3 Cabinets on the front wall of the virtual room with image corridors on their left side](image3)

The images displayed in the image corridors achieve two different purposes. On the left side, each image belongs to only one cabinet illustrating, explaining or completing its content. The images can be ancient maps, paintings, antic sculptures, reliefs, papyruses, book covers etc. on the one hand, and selected quotations, concordances, explanations, commentaries etc. on the other hand. On the right side, the images may occur...
in more than one image corridors. They can present dictionary or encyclopedia entries, lists of selected sources or references, bibliographies etc. which refer to every one of the cabinets (i.e. their represented content) in a particular set of cabinets.

As regards the searchability and accessibility to the relevant content of the three-dimensional virtual library model, the careful arrangement – or juxtaposition – of the cabinets are of vital importance, especially in terms of the presented collage of the various content which the spatial hypertext structure is based on [16]. There is a wide variety of possible visual features which can imply meaning in a spatial hypertext system considered as a special collection of visually grouped elements or (as in our case) objects. Such features include “shapes, colours, borders, lines, arrows, regions of visual space, proximity of visual elements, and labels affixed to any of the above” [17], as well as the various styles or formats used in the presented (hyper)text (e.g. by using CSS in an HTML-based environment). In addition to the spatial arrangement of cabinets by categories of Callimachus, in the present implementation of our virtual library model additional labels, different curtain images and specific image corridors can be used to create an effective spatial hypertext structure in the virtual 3D space. In this space the users can move on purpose, searching for a specific information — or just walk across the room and amble along for pleasure, and explore everything they think is important or interesting. They can have a close look at the images and enter into each cabinet and read (or just have a short glance at) its content. Note that one of the great advantages of reading texts in the three-dimensional virtual space is that the users can freely choose the reading distance from the text they are interested. It is a feature which makes the three-dimensional space, at least concerning readability, an inherently accessible application because moving closer results in bigger fonts which obviously help the user to read the text more easily.

In order to take full advantage of the 3D environment we need a dynamic server-based system which can manipulate a database of the virtual library content, offers an interactive interface for the users to compile a search query from the metadata (categories, keywords etc.) available, and generates spatial maps and multiple views respectively — using the features of both 2D (i.e. web pages) and 3D (i.e. the VirCA system). For example, the 3D view of the virtual room that can be seen in Fig 3 has been generated by a query such as

\[
\text{(SELECT TOP 3 * FROM ... WHERE CATEGORY LIKE 'EPIC POETS')} \\
\text{UNION} \\
\text{(SELECT TOP 3 * FROM ... WHERE CATEGORY LIKE 'ELEGISTS')} \\
\text{UNION} \\
\text{(SELECT TOP 3 * FROM ... WHERE CATEGORY LIKE 'LYRICAL POETS')};}
\]

which provides all relevant information which is necessary to generate the corresponding view. Note that a query like this naturally implies the visual grouping (or the hierarchical order) of the library content to be displayed according to the given categories.

V. THE HYPERTEXT REPRESENTATION OF THE 3D VIRTUAL LIBRARY MODEL

Using HTML, CSS and JavaScript technology, we have created a hypertext representation or projection of our three-dimensional virtual library model (see Fig. 4, 5). In a web browser window, we can navigate in the two-dimensional projection or ‘map’ of our 3D library model. We can see a selector table in Fig. 4 which consists of two parallel rows each containing nine cells symbolizing the selected cabinets. Every cabinet can be selected interactively by simply clicking on the corresponding cell in the table. Note that the content of the table is generated automatically (in fact, by a PHP script) according to the query the user has sent to search the database. To avoid the continuing and unpleasant use of the vertical scrollbar, we focus on only one cabinet displayed in a main cell instead of all the cabinets that can be displayed in the virtual room in the 3D space. The function of the main cell is to display the content of a selected cabinet (see Fig. 5).

![Fig. 4 The hypertext representation of the 3D library model](image-url)

Let us see an example of the use of the 2D environment. After ‘Sappho’ has been selected in the selector table, the main cell would display the text of the curtain image of the selected cabinet (e.g. the title ‘Sappho’ or her picture). The main cell is surrounded on the left side by the thumbnails representing the images which are attached to the left image corridor of the corresponding cabinet in the 3D space. We can select every image by simply clicking on the thumbnail on the left; for example, if we select the image of Sappho (that is, the second thumbnail), the result can be seen in Fig. 5.

![Fig. 5 The image of Sappho in the main cell of the hypertext representation with the left image corridor](image-url)
In the current implementation of the model, the images of both the left and the right image corridors of the cabinets can be displayed at the same time. Fig. 6 displays the two-dimensional projection of the virtual room focusing on the cabinet containing information on Homer.

Fig. 6 The image of Homer in the main cell of the hypertext representation with both left and right image corridors

VI. INFORMATION RETRIEVAL AND CONTENT GENERATION IN THE 3D VIRTUAL LIBRARY MODEL

In order that the collected and processed information could be well-searchable we created a dynamic server-based system based on a MySQL database of the virtual library content [11]. A PHP-based interactive web interface is available for the users to create and modify a search query the functionality and content of which have been continuously improved since its first implementation (see Fig. 7).

![Fig. 7 Selecting and adding new criteria to a search query](image)

The users can select (any combination of) the available items from several lists that contain the proposed values of the different types of library metadata (e.g. author, category, title, language etc.). After the selection procedure has been completed the browser generates the corresponding SQL query statement (using JavaScript). One definite advantage of this solution is that the users can freely modify the query and search for other data in the database (with the natural risk that they will get no results). After all search criteria have been properly given, the users can simply click on the ‘Search’ button and thus can send the accepted SQL query for further processing to the DBMS. Note that in the “SQL Result” text area of the search interface the generated SQL query can be manually modified before sending it.

After the query has been successfully processed the search results can be displayed, and thus can be quickly viewed and checked, in two different ways. In the first place, a specifically designed web page is generated which serves as a kind of 2D projection or hypertext representation of the three-dimensional presentation of the search results as we have described in the previous section.

There is another way of displaying the search results. After clicking on the “Generate VirCA room description file” button just above the selector table on the web page (see Fig. 4) a well-formed XML file would be generated according to the corresponding validity requirements of the VirCA system. Of course, the main function of this XML file is to present the virtual room in the three-dimensional virtual world of the locally installed VirCA system (the corresponding folder of which can be accessed by the PHP program). But the virtual room described by the XML file can also be displayed as a web page using XSL technology [11]. In the current implementation of our virtual library model we also use the jQuery technology to implement some useful dynamic effects on the generated web page, e.g. placing opaque thumbnails representing the content of the image corridors on the left and right sides of the curtain images of the cabinets and, after clicking on them, displaying their contents (see Fig. 8). The inner content of the cabinets is displayed behind the curtain images, and they can be displayed by clicking on them. In case we attached a valid URL to the displayed images the browser can also display the linked pages.

![Fig. 8 The projected cabinets, labels and image corridors (as small opaque thumbnails) displayed on an XSL generated web page](image)

Both presentation forms provide a holistic overview of the objects representing the search results on one page. Note that in both projections all essential constituents of the 3D virtual environment (e.g. cabinets, labels, curtain images, the content of the cabinets as background pictures, image corridors on both sides, main arrangement characteristics of the cabinets...
etc.) are represented on the 2D plane. The search results can be available in two ways:

1. after clicking of the symbols of the cabinets arranged in the selector table serving as a 2D map on top of the generated web page (see Fig. 4). The content of the selected cabinet can be displayed in the main cell of the web page surrounded by image corridors on both sides (see Fig. 5 and 6);

2. the search results are displayed as background pictures of the (symbolic) cabinets arranged in two parallel rows, just like in the 3D environment (see Fig. 8).

If we are satisfied with the results, the generated XML file providing the description of the virtual room could then be used in the locally installed VirCA system to view the results in the cabinets in the virtual 3D space – offering all the great advantages and benefits described before.

VII. THE 3D VIRTUAL LIBRARY MODEL AS A NETWORK AND KNOWLEDGE BASE

The database structure of the current implementation of the 3D virtual library model is based on three tables (called main, links and extras) which specify the structure and function of the elements or entities of the library model as a network. The records of the main table can be considered as the (parent or child) nodes of the network. They describe selected primary texts (poems, fragments, or selected passages of literary works etc.) and the visual representations of the various sources of the texts (e.g. web pages or printed materials). The records are classified into different categories including Callimachus’ original categories (e.g. Elegists, Epic poets etc.).

The current implementation of our model, the records of the extras table have two different functions. First, they refer to images which are closely related to the texts or their sources (e.g. relevant pictures of a web page or illustrations in a book). Second, they refer to images representing commentaries or other verbal content (selected concordances or quotations etc.) which explain and/or emphasize certain parts of the texts of the main table or add relevant information to them. The records of the extras table are displayed in the left image corridors. As we mentioned before, the number of displayed images on each corridor is limited to a maximum of four items. They represent leaf nodes (i.e. nodes that do not have any child nodes) in the network of the library content.

The records of the links table describe various relationships between the texts of the main table including e.g. collections of references and sources, translations, bibliographies, and other related documents. The records of the links table are displayed in the right image corridors. In addition to the hierarchical relationships (and the respective links) formed by categories and authors, the records of the links table form additional links between the nodes of the network.

It is worth mentioning that using “the new science of networks” paradigm [18, 19, 20], the primary texts with the corresponding illustrative images and/or commentaries in the left image corridors might be considered as “small worlds” in the network because they are always displayed together in the 3D space as well as in the spatial 2D hypertext representation (see Fig. 9).

![Fig. 9](image_url) A selected passage from Odyssey by Homer. The primary text with the images on the left side form a “small world” of nodes, while the dictionary entries on the right side act as “authorities”.

In addition, those nodes that convey basic reference knowledge (e.g. dictionary entries about Greek mythology) displayed in the right image corridors (as in Fig. 9) could play the important role of authorities [21, 22]. They link the highly interconnected groups of nodes (i.e. nodes forming “small worlds”) having a reference to the same dictionary entry. Note that in our virtual library model the authors of the texts could play the role of hubs linking those “small worlds” the text of which has the same author.

Because in our virtual 3D library model the selection and organization of materials, and consequently the structure of the network of library entities depend only on the content we would like to represent and offer, we are free to build a network having a specifically designed characteristics. We have decided that our network should be scale-free because scale-free networks seem to have a strong relationship with human language, cognition and behavior [18, 21, 22, 23]. Therefore they seem to be especially suitable and efficient also in library environment. Our research hypothesis is that having been ensuring the scale-free characteristics of the network of nodes in our virtual library model (constructing “small-worlds” of nodes, interconnecting them with appropriate hubs and authorities etc.) would gradually lead to a carefully and efficiently organized network of represented knowledge.

VIII. THE ALEXANDRIAN MODEL OF NETWORK FORMING THE COGNITIVE FRAME OF THE VIRTUAL LIBRARY MODEL

The validity of our research hypothesis is naturally based on the question whether our virtual library model could satisfy the basic requirements of scale-free networks. In the following we shall describe a simple abstract network model called the Alexandrian model of network which seems to comply with all the necessary requirements and, moreover, it provides some essential characteristics of the network with very plausible interpretations. We will show that the concepts of this abstract model can be easily mapped into the main concepts of our virtual library model (i.e. categories, authors, texts, illustrative images and/or commentaries in the left image corridors etc.). Consequently, we might consider the Alexandrian model of
network as a cognitive frame of our virtual library model, and the representation of the knowledge our virtual library model conveys.

The Alexandrian model of network is represented by a directed graph forming a hierarchical (or tree) structure. The model contains four different levels.

(1) For the sake of simplicity, let the first (or top) level have only one node or vertex which might correspond to a specific category of the library model (later we shall see that the number of nodes on the first level has only minor effects on the main characteristics of the model).

(2) There are \( n_1 \) nodes on the second level which are linked to the node on the first level; these nodes, according to Callimachus’ scheme, might correspond to the authors in the virtual library model.

(3) The organization of the third level is similar to that of the second one, except that each node of the second level has exactly \( n_2 \) links to the different nodes of the third level. Consequently, in the third level there are \( n_1 \cdot n_2 \) nodes which might correspond to the selected texts of the authors.

(4) The organization of the fourth level is similar to that of the third one, except that each node of the third level has exactly \( n_3 \) links to the different nodes of the fourth (and last) level. Consequently, in the fourth level there are \( n_1 \cdot n_2 \cdot n_3 \) nodes which might correspond to the illustrative images and/or commentaries in the left image corridors of our virtual library model. Note that in our virtual library model there can be only a maximum of four images in the left image corridors which results in \( n_3 = 4 \).

The Alexandrian model of network represents a scale-free network where the frequency or degree of nodes (that is, in our model, the number of outgoing links from nodes) should follow a power law distribution \( f(x) = c \cdot x^{-\gamma} \) which results in three simple equations:

\[
\begin{align*}
  c \cdot n_1^{-\gamma} &= 1 \\
  c \cdot n_2^{-\gamma} &= n_1 \\
  c \cdot n_3^{-\gamma} &= n_1 \cdot n_2 
\end{align*}
\]

Supposing that the value of the parameters \( n_2 \) and \( n_3 \) has been already set by predefined values (\( n_2 > n_1 \)), the second and third equations lead to the determination of the exponent \( \gamma \) by the formula

\[
\gamma = \log n_2 / (\log n_2 - \log n_3) \quad (A).
\]

Now that the exponent \( \gamma \) can be counted, the first and second equations lead to the determination of the variable \( n_1 \) (which might correspond to the necessary number of authors) and the constant \( c \) by the formulas

\[
\begin{align*}
  \log n_1 &= (\gamma \cdot \log n_2) / (\gamma - 1) \\
  c &= n_1^{\gamma} \quad (B) \quad (C).
\end{align*}
\]

Assuming that \( n_3 = 4 \) the exponent \( \gamma \) depends only on the number of (outgoing) links of each node in the second level pointing to different nodes of the third level, i.e. \( n_2 \). Note that setting the value of \( n_2 \) to 8 which is very close to George Miller’s “magic number seven” [23], we get for the exponent \( \gamma \) the exact value 3 (in fact, \( n_2 = 7 \) results in \( \gamma = 3.47722525... \)).

A simple interpretation of these results can be as follows. As we mentioned before, \( n_1 \) is the number of (outgoing) links of each node of the third level pointing to different nodes of the fourth level which might correspond to the illustrative images and/or commentaries in the left image corridors of our virtual library model. Because the nodes of the third level might correspond to the selected texts of the authors in the virtual library model, and the texts and the corresponding images and commentaries form an inseparable whole which can be plausibly interpreted as a “small world” of the network. Hence, each small world contains \( (n_3 + 1) \) nodes. In addition, \( n_2 \) is the number of (outgoing) links of each node of the second level pointing to different nodes of the third level; the nodes of the second and third level might correspond to the authors and selected texts of the authors in our virtual library model, respectively. So the authors can be considered as hubs linking the small worlds formed around the selected texts of the authors.

According to the Alexandrian model of network the value of the exponent \( \gamma \) of the power law distribution depends only on the degree (i.e. the number of outgoing links) of hubs and the size (i.e. the number of nodes) of the small worlds of the network described by the simple formula (A) which gives a very plausible interpretation for one of the most important characteristics of the described scale-free network.

Finally, we would like to make some additional remarks. First, we mention that using more than one node in the first level (which might correspond to more than one category) only the constant factor \( c \) of the power law distribution will be affected (e.g. using two nodes \( c \) will be twice as much etc.).

Second, adding new nodes to the fourth level of the model as authorities linking the appropriate nodes in the third level (e.g. those texts which have a reference to the same dictionary entry), only the parameter \( n_1 \) will be affected, i.e. its value will be increased by a fixed number. This number, in our virtual library model, is equal to the maximum number of displayed nodes in the right image corridor which results in \( n_3 = 8 \).

IX. THE MAIN PURPOSE AND POSSIBLE APPLICATIONS OF THE VIRTUAL LIBRARY MODEL

The main and unique design and development principle of our virtual library model based on the system and content of the ancient Library of Alexandria is to represent the library content deliberately as a scale-free network. The systematic and efficient organization of data, which is one of the expected results of the scale-free characteristics of the network of the represented library knowledge, reflects the main purpose of the virtual model of the ancient Library of Alexandria. There are a lot of valuable sites in the internet (e.g. the Perseus Digital Library\(^4\)) which provide rich sources for those who are

\(^4\) http://www.perseus.tufts.edu/hopper/ (2015-09-06)
interested in the works and thoughts of the ancient authors. But they offer access to data in the first place and leave the selection and processing of data to the reader who might or might not cope with the enormous quantity and high complexity of the available materials (not mentioning the necessary background knowledge, linguistic competence etc.). For qualified scholars, it may seem to be quite acceptable; but for non-professional users (including students) there is a considerable and increasing need for a kind of efficiently preprocessed, “ready-made” knowledge which is suitable for understanding, interpreting, learning and/or memorizing the processed content “as is”. Which is, in our case, one of the most important cultural heritage of the western civilization – to say nothing of other important “side effects”, e.g. the outstanding educational role of the provided literary texts in language learning (in our case, learning English). During the development and implementation of our virtual library model we have been focusing on a kind of “read and learn at once” strategy which is perfectly supported by the three-dimensional and spatial representation of the provided knowledge (using both verbal and multimedia texts), the built-in limits for the available materials (e.g. a maximum of 18 cabinets in a virtual room, a maximum of four images in each of the two image corridors including illustrative images, appropriate quotations, commentaries etc. on the left side and related materials on the right side, the limited size or segmentation of primary texts etc.), and the conscious organization of the content.

In addition to the possible applications described above, our research work is strongly connected to digitized content management which has a vital role in the infocommunication sector. Here we can speak about an expanding content space where ‘content’ is interpreted in a broader sense, covering voice, data and visual information as “traditional” inputs on the one hand, and smelling, touching, emotions, sensory information, 3D gestures etc. on the other hand. Combining infocommunications and cognitive science, and expanding the content space with cognitive and sensory contents, the central aim of cognitive infocommunications can be “to manage this expanding content space, and to provide opportunities for creating enhanced, richer applications” [24, 25], including the enhanced use of the 3D environment and virtual reality [26]. We would like to add that, by any form and composition of the content space the application is actually using, only the efficient organization and representation of data could effectively lead to the aspired transformation from the represented and manipulated content (e.g. the retrieved data of the application) to real knowledge, so that the cognitive “capabilities of the human brain” would be extended and may “interact with the capabilities of any artificially cognitive system” [26]. Not to mention that the study of the representation of knowledge is one of the main research areas of cognitive science. Besides, it seems to be a question of great interest whether the approach of considering the represented content as a scale-free network could lead to “emergent cognitive properties” of a systematically organized and well managed application – just like in the case of social networks where those properties are the result of “high-level structural and organizational management” which is therefore one of the main issues of interest to CogInfoCom [27]. So we firmly hope that our three-dimensional virtual library model, its implementation and content, and the scale-free Alexandrian model of network which the organization and content development of our virtual library model is based on, will be a really useful and valuable contribution to the cognitive infocommunications research which inspired our work.

REFERENCES


The education of disruptive technologies in the innovative engineering training

Ildikó Horváth
University of Pécs
Faculty of Engineering and Information Technology
Pécs, Hungary. Email: horvath.ildiko@mik.pte.hu

Abstract

The study analyzes the disruptive effects the technical innovations have on the social – students’ and employers’ – requirements, on the educational environment, on the contents of the learning material, on the educational and learning methods, and on the students’ and lecturers’ roles by investigating - from a multidisciplinary point of view - the disruptive effects the disruptive technologies appearing rapidly from the end of the 20th century have on the higher education.

The theoretical background of the study is based on several pillars: besides overviewing the disruptive innovation based on the disruptive technologies, the innovation-spread, the life-cycle analytic background of the technologies and the acceptance models, generational theories and the international educational developments make up the basis of my empiric research which is based on international studies on educational development.

In the scope of our study are the disruptive technologies and the education suiting the digital life of the Z and CE generation students, who make up the CogInfoCom society. The CE generation unfold their personality and social life in an environment that is inseparable from the cognitive ICT, so it is very important to observe their way of thinking and habits because we have to teach them the disruptive technologies.

In the article we show an international questionnaire survey which we carried out concerning the time and the applied methods of the introduction of integrating the disruptive technologies in the educational material of the higher education of engineering. With a following questionnaire survey we searched for reliable information concerning the digital life, the internet using and learning habits of the Z generation students studying in the higher education of engineering. In accordance with the investigation results we see it is necessary to review the educational material contents and the applied educational methods of the higher education of engineering, and the wider and wider application of the modern technologies.

In connection with the integration of the disruptive technologies in the educational material of the higher education, the 21st century teachers’ roles, the educational materials and the educational scenes are also defined.

Based on the above-mentioned points, we can say that the cyber-higher educational environment feasible by the technological innovations appearing in a rapid pace and the methods successfully applicable in the future educational space are in the focus of the article. Our aim is to show – through an educational example - the realization of how we can teach the future technologies to the future engineers by applying innovative educational methods. The cooperative VR learning environment presented in the article, due to its project-based educational opportunity, provides an efficient educational scene for the higher education of engineering, at the same time it facilitates the education of the disruptive technologies appearing in higher and higher number, even right after the appearance of the technology. At the same time, it meets the students’ requirements set for the educational environment and it also provides a perfect scene for developing the competences determined in the employers’ requirements.

Key words: Disruptive technology, VR education
Introduction

The results of science, the innovations of technology and the disruptive technologies will greatly determine the quality and conditions of people’s life in the following years. More and more research deals with predicting the development of informatics and technology following the intense advancement of the informatics and knowledge society in the 21st century. The intensified development of advancement of science and technology is predicted which can already be seen today.

Today many various technologies parallelly exist and change our society, the new generations’ chances and life, the expectations of the labour market, our educational system and each area of our life. Day after day we experience it that the globalizing economy, the mobility of the workplaces, the acceleration of the technical and technological development have put the institutions of the higher education in a competitive situation. The service providing role of the higher education due to the market-oriented character of the higher education of engineering has appeared and become more emphasized. Service towards the students as future employees and towards businesses in the market sector as employers. Nowadays it has become inevitable to provide new solutions for the new types of market challenges in the global system of education, thus in the higher education of engineering, too.

More and more disruptive technologies appear every year on the Gartner’s Hype Graph, so the competitiveness of the higher education of engineering can be maintained if it can provide this extra knowledge for students necessary for innovation in this fast changing world. Understanding technology is the pledge of becoming a global citizen, a global professional of engineering. In the engineering training, the efficient teaching of the disruptive technologies bursting into the market requires the application of new ways of pedagogical methods, efficient time use, the application of technics and technologies suiting the digital life space of the CE generation students studying in the higher education.

At the same time an increasing number of new scientific fields are appearing and one of these is Cognitive Infocommunication (CogInfoCom) (P. Baranyi és A. Csapo, 2012, P. Baranyi, A.Csapo, Gy.Sallai, 2015) which investigates the trends of informatics development and directs attention to the importance of the need for access to infocommunication. The internet will be the network of people who use them and not the network of computers, and telecommunication, informatics and the media will be intertwined. Informatics systems will evolve which will realize the communication between the people and the system control where the people and the ICT are inseparable. It is especially characteristic of the CE generation students who live in this „human computer interaction way of life”, in this human-ICT combination.

According to Castells, the informatics society is a new way of human coexistence in which the networked production, the storing, the processing and the recalling of the information play the most important role (Castells, 1998). The informatics society can be viewed as a new society because not only new things appear in it, but also the old ones start to operate differently, the same way as the disruptive technologies do for the social (educational), economical and technological processes.

Similarly to the informatics society, the real dimensions of the disruptive technologies can be highlighted through education, science, innovation, the (new) economy, content and culture, and in our article we want to present the ideas and research results in connection with these.
Disruptive innovation

From the end of the 20th century, more and more studies have been dealing with the increasing number of disruptive innovations, and within these with the targeted disruptive technologies and their effects made on the society.

Clayton M. Christensen, a professor at the Harvard University created the concept of ‘disruptive’ in 1997 meaning ‘creative destruction’ and published it in his book ‘The dilemma of the inventor’. An expression used for applying solutions in which the new technologies and business models fundamentally transform or replace the existing methods and this way they influence the business value of the products and services.

Innovation is a relatively new concept. Hardly any studies were made about it until the end of the 20th century, however the topic has become significant recently. The most commonly used definition of the word innovation originates from Schumpeter, that is innovation means the new combination of the production factors (Schumpeter, J. A., 1939.). This definition later served the basis for OECD where a more comprehensive definition was created for innovation.

According to the definition of the European Union: „Innovation is the process of knowledge application, reforming and increasing the products and services and their markets, applying new procedures in production, in distribution, in the market labour, in management, in organizations and working conditions, the expansion and reform of the professional knowledge of the labour force” (EC, 2004.). Innovation, by which we may mean satisfying the customers’ demands at a higher and higher level, can mean product development, technology development and organization development, can be a continuous development or skyrocketed, comprehensive strategic innovation (Chikán, 2004. p. 180-181.).

The Innovator’s dilemma (Christensen, 1997) refers to the decision between development based on the known methods or the paradigm shift. According to Christensen, innovation has two forms depending on its circumstances:

- Sustainable innovation which wants to provide a greater performance to the target-oriented customers with high demands, usually at a high price. These innovations are usually developed by big companies. They have financial, technical and human resources for development, for improving the already existing technologies.

- Disruptive innovation is where creating markets to satisfy new customer demands determined by new rules and not creating old markets is happening. Typically the start-up businesses have „specialized” for the disruptive innovations.

Disruptive technologies

Radical technical innovation, especially the so-called disruptive technological development usually evolves with a longer research. Very often a long time elapses between the birth of the theory and its proving and implement. Quite a lot and often costly experiments, test productions and testing prove that the technological solution is really possible. And after a shorter or longer time the disruptive innovative technologies typically change into a sustainable cycle.

Concerning the objectives of the technological development we can distinguish two groups. The first is to maintain sustainable development, that is the better utilization of scarce resources such as the alternative energy resources, waste usage, environmentally-friendly architectural technologies, or the development of the environmentally-friendly informatics devices suiting the criteria system concerning the informatics devices of the EU GPP.
In the second group we can find the innovations aiming at improving the quality of life, the devices serving speed and parallelism (Duma, Erdős, 2008), such as in transport the driver’s navigational, PDA-NDA, robotized systems, nanotechnology, the memristor, the Internet of Things and we could still continue the list with the disruptive technologies grounding the disruptive innovations.

It is clearly visible that today these technologies have a mostly digital characteristic and are expected to ground the technical cycle change.

The life cycle of the technologies

During studying the diffusion they fundamentally investigate the process when innovation spreads in the society in a communicational channel in a given time. The studies of diffusion have made it possible to follow the changes caused by the innovations in the society and the effects of the social forces on innovation and its spread. Everet M. Rogers and Frank Bass are the first representatives of the innovation spread models (Dessewffy - Rét, 2006). According to Rogers’ definition, the spread, diffusion of the innovation is the temporal process during which a new product becomes gradually accepted in the target market by the potential customers and in the entire society (Rogers, 1995). Following their work, numerous technological life cycle models try to measure and then demonstrate the evolution of the new technologies.

The diffusion model describes the spread of innovations with a so-called S-shaped curve. The S curve shows the frequent spread pattern of the innovations. Initially, the number of users increases only in a low pace, then a significant boom comes which is followed by another slow-down after reaching the saturation level. In the initial phase of the spread, the so-called Innovators (2.5%) start to use the new technology. They are usually more open to innovations, and attribute a special value to the innovations. They are followed by the Early Adopters (13.5%), then by the Early Majority (34%) whose joining in the process of diffusion suddenly increases. As groups of different sizes are concerned, the technological spread is not a linear process (Rogers, 1995; Dessewffy-Rét, 2006). The so-called Adoption Curve shows the market adoption of the technology in relation to time. The life cycle of the disruptive technologies and the disruptive innovations is most typically shown in the Gartner’s Hype Graph.

Fig.1. Technological life cycles (Source: Gartner Research, May 2003)

On the horizontal axis of the Hype Graph we can find the maturity or development in relation to which the curve shows the adoption of the given technology. Gartner’s Hype Graph introduces a new dimension in this diffusion model which is different from other life cycles. It
does not only show the people’s attitude towards technology but also the technological maturity. Technology has a significant impact on business and society which depends on a secondary factor, that is how much the technology is preferred and can be used for various applications and services. Thus, it is understandable that for the developing companies, for other service providing companies it is vital to have prepared engineers concerning understanding the more and more numerous disruptive technologies, and measuring their feasibility, or even implementing their own developments.

Each disruptive technology goes through a similar cycle. The particular disruptive technologies do not move at a constant speed on the Hype Graph. The progress of some technologies lasts for decades, for example object orientation which was in the scope of the universities and research centres for decades and after a long preparation phase the development started. In contrast, the so-called „accelerated technologies” find their place in the market in 2-4 years. Simple usage and high quality, for the businesses and the average users alike, the support of some potential producers and the fact that they suit to the recent infrastructure are typical features of the accelerated technologies, such as the SMS. The „Long-fuse” technologies are more and more frequently paid attention to. The sci-fi type of characteristics of this technology is specific to them because they precede their time. The development is very complex and thus it is basically connected to science and technology and highly depend on competences. The new infrastructure is accepted after a long progress which is also followed by relevant changes in the processes of businesses. As the prevailing time of the technologies is not at all uniform, it is justified to integrate them in the training material of the future engineers within the shortest possible time after their appearance.

The effects of the disruptive technologies on higher education and social processes

The next chapter is about the studies which scrutinize the effects the disruptive technologies have on the society, education, and especially on the higher education. Academic literature concerning the acceptance of technologies and the spread of innovation date back to several decades. Publications dealing with the appearance and the spread of ICT devices try to describe the interaction of technology and society with different scientific models from which TAM (Technology Acceptance Model) (Davis, 1989, Schepers-Wetzels,
2007) has an outstanding role. The aim of the TAM model - which investigates the acceptance of new technologies and the effects the digital environment has on the customers and their habits - is to forecast the user objectives and the social acceptance of the technologies, and to determine the necessary changes for the acceptability of the technology. According to Davis’ TAM theory, acceptability is influenced by two factors, utility and simple use, then, if the social intention is given for the use, the use of the system actually becomes a part of every day life after a short time.

Fig3. Source: based on Davis (1989) – personal editing

Many of the methods used in the science of future research are suitable to forecast the social acceptance of the technologies. The modern participative methods appeared in the future research from the middle of the 20th century in many different forms (Glenn, 1994), however their use has only become widespread in the last decades (Glenn, 2003). The Futures Wheel (Glenn, 1994) is suitable to identify and categorize the secondary and even the further effects and consequences of the events. The method, as a technique of the systematization of our thoughts and questions concerning the future can also be interpreted as a structured brainstorming (Hideg-Nováky, 2012). The Futures Wheel can have several objectives, such as considering the possible effects of the existing trends, potential future events, and the systematization of thinking about the future events and trends, making forecasts, presenting complex interrelations, visualization of another kind of future research, elaborating multilateral approaches, strengthening the future-oriented point of view, supporting the team brainstorming (Glenn 2003). During the STEEP - Sociological, Technological, Economic, Ecological or Environmental and Political – analysis, they group the effects to be considered, then based on the analysis they determine the effects of the innovations. According to the professionals dealing with future research, the Futures Wheel can also be successfully applied in case of the lack of preliminary future-methodological knowledge, if we would like to collect and systematize the effects of an event which has already happened or is to be expected to happen. It is one of the popular analyses due to its simplicity and easy interpretability. Its main advantage is that it can be applied at any point of the research to understand trends and further events. It is easily adaptable to investigate different situations.

Technology-based society

In the 21st century, more and more studies deal with forecasting the occurring informatics and technological developments. According to Charlie Gere’s essay about data processing in the Digital Culture, the network culture extends the informatics era of the digital data processing. In his analysis, he considers the digital change just as much a social economical phenomenon
as a technological development. By the 20th century, besides the digital culture the development of the network culture has a more and more significant role. Unlike digital culture, in the network culture information is not so much the result of the relations between the separate processing units, but the result of relations between the humans, the machines, and the humans and the machines. The evolutionary Hype curves after 2010 refer to the fact that the primary technological components shift towards a dynamic technological integration (P. Baranyi, A. Csapo, Gy. Sallai, 2015).

The increase in the number, the diversity and complexity of the technological devices results in the radical subversion of the lifestyle. The complex systems have complex effects which launches more and more complex work processes, thus the technological determinism sees technology as the main driving force of the society.

Producing immaterial information and transmitting it through the network is becoming more and more the dominant organizing principle of the global economy.

In his work „The Rise of the Network Society” Castells suggests that the result of the series of the technical and technological changes is the change of the capital and the change of the personal behaviour.

In accordance with the networking of the society, several international and national organizations deal with the functions and the research of the Future Internet, as the Future Internet among the infocommunicational technologies as the key technology („Key Enabling Technology”) of the 21st century has an effect on the innovation and the efficiency of each area.

The Japanese National Institute of Information and Communications Technology (NICT) presented its vision for the New-generation Networks – NWGN in 2008, which aimed to draw up research objectives and technological requirements to facilitate the implement of future knowledge society. The ITU, the institute of the UN specialized for ICT has made its recommendations in the ITU-TY series for the foundation of the standardization of the future network. The recommendations of the ITU-T Y.3000 (especially the 3011, 3021, 3031) identify objectives which do not receive enough attention during the planning of the present networks to realize the Future Internet.

Fig. 4: Future internet vision, based on the NWGN, FN and FIA visions
(Source: FIRCC Report, 2014)
The European Union pays a special attention to the new generation internet research to which Hungary also joins with an active participation. In 2011, in Hungary the Future Internet National Technological Platform was established and in 2013 for its initiation the Future Internet National Research-Coordination Centre. An extensive interest is shown by the fact that 34 institutions joined the programme and 132 research topics were registered by the beginning of 2014. The 2014 FIRCC report shows the results of 83 research from the almost complete collection of the national internet research where the 3D internet and the cognitive infocommunication, and „The future internet social applications” topics have a prominent role as the 21st century is an age when the human and the ICT are interwoven. This fusion opens a new age in the digital and network life, creates new disruptive technologies, inspires the creation of new scientific areas. It makes the generation groups (Baby-boom, X, Y, Z, α) already known in the classical generation theories more complete by the appearance and defining the newer so-called CE (P. Baranyi, A.Csapo, Gy.Sallai, 2015), and CCE (P. Waldbuesser, L.I.Komlósi, 2015) generations suitable for the spirit of the present era.

The skyrocketed acceleration of the technological development and the rise of human average age – primarily in the developed countries – result in the fact that even 5 generations totally different in technological preparation, skills and even point of views live together which naturally inspires the generation research.

It can be observed that all over the world the researchers deal more and more with the habits and the digital life of the youngsters belonging to the CE, CCE and within that the Z and α generations. The reason is to be searched, on the one hand in the numbers, on the other hand in the lifestyle and habits so significantly different from the earlier generations. As a numerical support, I am relying on the different research results as the present Z generation in the USA will amount to 40% of the population by 2020, and with the α generation following them they mean the larger proportion of the society (15 Mind-Blowing Stats About Generation Z - CMO.com). In Australia, by 2025 more than 35% of the employees will belong to the Z generation (McCrindle Research) and these figures show a similar proportion in Europe, too. The generation which was born between 1995 and 2010 belong to the Z generation who are the students of the present higher education, and are very different from the other age groups.

<table>
<thead>
<tr>
<th>CogInfoCom Society</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE generation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CCE generation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>α generation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Z generation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>• Their personality and social life unfolds in an environment that is inseparable from cognitive ICT.*</td>
</tr>
<tr>
<td>• They adjust their lives in the technologies used by them.</td>
</tr>
</tbody>
</table>

Fig.5.: The CogInfoCom society (Source: personal editing based on P.Baranyi és Á.Csapó)
Known as the „digital integrator” generation, they suit their life to the technologies which they use. The world they live in has no limits, they can find all the necessary information they need with just a click in an „open book” digital environment. They grew up in a modern world where they got used to having the access to all innovative technologies and electronic and smart media devices where the internet mostly ensures the use of short texts, mainly videos and pictures providing information fast, even parallelly and straightforwardly. „While in the life of the earlier generations an easily distinguishable real and online identity existed, for the youngsters today there is only one identity, that is for them the offline and online existence are completely intertwined. For the young generation these two are inseparable, and technology is only the tool of expressing the identity” (Ujhelyi, p. 9., 2013). The members of the older generations are surprised and do not understand this new phenomenon and that is why there are so many studies carried out in this topic worldwide. The generational differences do not only appear in the family but there is a very serious generation gap even in the educational environment. „It also raises the current issue of what skills and abilities are needed to be developed for the more efficient knowledge imaging of the new generations and what methods can lead to success (Pais, 2013).

Concerning the habits of the Z generation, the researchers particularly deal with the internet using and media using habits, but also their relation and requirements to learning and work are in the focus of attention in several international surveys, like the EU Kids online research carried out by the European Union in 2011-2012 and the research questionnaires used by the American organization Common Sense.

The effect of technological development on pedagogy

The technological development undoubtedly has an effect on the social processes, on the labor market requirements, thus on education. The appearance of the new technologies, the more and more widespread use of the internet and its social acceptance have an effect on the educational environment and also on the contents of education. Amid the rapid, continuous expansion of knowledge a significant change can be seen in education (Brown, 2002, Condie and Murano, 2007; Gibson, 2002). The cause of the paradigm shift is that the new technologies have changed the access to the information, thus that is not closed between the walls of the educational institutions. The world is only a few clicks, moves away from the students which they can reach with their mobile smart devices regardless of time and space.

The Schooling for Tomorrow project of the OECD was launched in 1997 at an international conference organized in Hiroshima, then the OECD - in cooperation with the Centre for Educational Research and Innovation (CERI) - with a wider and wider international participation is trying to create a systematized knowledge base and tool park for the decision-makers in education politics for thinking about the school of the future, for the realization of the educational innovations. Pointing out that nowadays the new technological bases facilitate new learning forms, expanding the requirements concerning learning itself. The programme of the OECD helps building out the decision-making strategy and at the same time it widens the chance of comparing the developing trends and experiences with the involvement of international expertise not only in space but also in time.

In 2003 Benő Csapó, one of the Hungarian researchers, wrote in his study that the appearance of the ICT technologies strongly influenced the education, as it set new needs for education, on the other hand it provided an efficient toolkit system to increase efficiency. Affected by these, the learning scenes shifted from the formal education towards informal education.

At the beginning of the 21st century, the „learning society”, or better known „lifelong learning” appears as a political objective worldwide. The lifelong learning is a leading principle according to which the educational system has to be transformed so that the
necessary learning opportunity could be available for every European citizen at any time of their lives (A Memorandum of Lifelong Learning, 2000; Making a European Area of Lifelong Learning a Reality, 2001), which makes it necessary to rethink the forms and dimensions of learning.

The ISTE (International Society for Technology in Education) has worked out so-called standards for the learning and educational processes of the digital age which the competence family called NETS (National Educational Technology Standards) contains. Its members: for teachers the NETS for Teachers – NETS-T, for students the NETS for Students – NETS-S, for informatics teachers the NETS for Computer Science Educators – NETS-CSE, but it also contains standards for the administrators NETS for Administrators – NETS-A and for the educational assistants the NETS for Coaches. The ISTE NETS-S specifies the skills which are necessary in the global, digital world for efficient learning and productive lifestyle.

Fig.6.: Students’ competence system (Source: ISTE NETS-S)

As a result of thinking together about the future of education, many scientific publications have been published. As their summary, some cardinal changes will be highlighted which have a great significance in the future education. The realization of learning in diverse places and time, the need for personalized learning, the students’ chance for free choice becoming stronger, the intensification of the short-term, project-based learning views, involvement of personal experience facilitated by the modern technologies, the ability of successful learning, the change of the pedagogist’s role.
From the modern pedagogical methods developed to achieve the above-mentioned goals the gamification, the flipped classroom, the mobile learning, the virtual learning environment are the „disruptive technologies” – based on informatics development -, of the education today whose position, besides many other technologies, in the Hype curve in 2016 are shown in the Figure xxx.

Among the disruptive technologies of the education today, the Gartner’s Hype curve predicts a longer cycle for the different realization of the technical-based learning, such as the Virtual environment, the Mobile learning, or the Gamification. Technical Based learning – TBL means the electronic technical based learning, including the internet, the intranet, the satellite programmes, the audio-videoconferences, the chat rooms, the web contents and the lecturers’ CD and DVD applications, but the technical based learning also includes the complete range of the learning supported by the ICT devices and the online learning, including the today so fashionable expression of e-Learning. The technical based learning facilitates the opportunity to access learning regardless of space and time by merging technological innovations appearing in higher and higherumber – among them the disruptive technologies -, the realization of individual learning progress, thus it gives space to the spread of the disruptive innovations.

While the studies dealing with the educational appearance of the disruptive technologies primarily write about educational use of the technologies, as new devices, such as in Clayton Cristensen’s book titled „Disrupting Class, How disruptive innovation will change the way the world learns” published in 2008, several studies deal with the future of the higher education, the education-shaping effects of the modern technologies. One of these is the study titled „The future of higher education: How technology will shape learning” published in the Economist Intelligence Unit in October 2008, supported by the New Media Consortium which presented a global online corporate survey and detailed interview. Of the 289 executives responding to the survey, 189 participants came from higher education and 100 came from corporate settings. The US accounted for slightly over one-half (154) of all respondents, with the remainder distributed through Europe (69), Asia-Pacific (43) and the rest of the world (23). Nearly 63% of survey respondents from both the public and the private sectors say that the technological innovation will have a major influence on teaching methodologies. 60% of the respondents said that the technological changes happening nowadays will change the views of the campuses from a one-dimensional, so-called physical concept to a multi-
dimensional (physical and virtual, - online) concept. Thus, online learning will become stronger dominant, more and more universities will react to globalization.

The presentation of the international survey.
I made an international survey based on the questionnaire 'The future of higher education: How technology will shape learning’ published in the Economist Intelligence Unit in 2008 concerning the appearance of the disruptive technologies in the learning material in the higher education.

The objectives of the questionnaire survey:
- To state whether the disruptive technologies appear in the learning material in the higher education and how much after their appearance do the disruptive technologies appear in the learning material in the higher education?
- How much is the introduction of the disruptive technologies accepted among the lecturers of the higher education?
- What educational methods and technologies do the institutions of the higher education which already have experience in the field of teaching disruptive technologies consider efficient to teach disruptive technology?

Hypotheses
- The disruptive technologies do not at all or hardly appear in the learning material in the higher education.
- If a disruptive technology is introduced in the learning material in the higher education, then it happens with a significant delay, it happens in the higher education after the Peak of Inflated Expectations in the Hype graph, in the Trough of Disillusionment.
- The methods applied in teaching the disruptive technologies differ from the traditional educational methods.

I used the IBM SPSS Statistics 23 programme and the Microsoft Excel 2013 programme for the statistical analysis and assessment of the responses.

Presenting the questionnaire

The questions of the questionnaire can be categorized in six clearly separable groups:
1. Details of the institution in higher education
   1. Basic information concerning the answering colleague (gender, age, experience in higher education)
   2. Background information concerning the educational technologies and methods applied in the institution of the higher education (with the opportunity of target-oriented Yes/No answers)
   3. I requested the answers by applying the 5-scale Likert scale concerning the appearance of disruptive technologies in the higher education.
   4. Collecting information concerning the introduction of specific disruptive technologies was carried out in the form of graphs where I asked about the educational level concerning the introduction, the scenes of education, the applied pedagogical methods and the results achieved.
   5. In the multiple-choice questions I tried to find the answer to the advantages and challenges of introducing disruptive technologies, and providing the most efficient methods.

In 2015, I sent the representative survey concerning the introduction of disruptive technologies in the higher education to 30 universities which have contacts with the Faculty
of Engineering and Information Technology of the University of Pécs, to international institutions of higher education ranging from the USA to Europe and China.

Assessing the questionnaire

They answered 93.5% of the questions. They most often did not respond to the questions about the results of the introduction of the disruptive technologies, however in some cases they did not determine the disruptive technologies either.

The respondents’ average age was 46 with a standard deviation of 5.44, the average experience they had was 19.5 years with a standard deviation of 6.5 All the responding institutes of the higher education who took part in the survey and filled in the questionnaires participate in academic research projects. 60% of these involve the business partners in the educational programmes mostly with guests lecturers. 80% of them use cooperative learning technologies and 90 % of the universities responding help their students by providing their own electronic teaching materials. Only 40% use the virtual educational environment.

The responding universities integrated 2 to 5 of the disruptive technologies in their teaching material with an average of 3.4. 70% of the institutes of the higher education integrated introducing the disruptive technologies in the learning material in their BSc courses and 30% of the universities integrated them in their MSc courses.

As for the significance of introducing the disruptive technologies in the higher education, the respondents’ opinions were asked by using a 5-scale Likert scale which is presented in Figure 3.

![Fig.8: Figure prepared by myself](image)

It is clearly seen that the respondents are determined about the significance of integrating new technologies in their education, however in terms of the timing of the introduction they have different views. Some respondents did not think this question was relevant at all. A relevant deviation is clearly visible (p=0.00) at the institutes of the higher education which integrated technological innovations in their MSc courses as they all considered the students’ participation in innovations fundamentally important.

The lecturers’ opinion was that introducing disruptive technologies in their teaching materials at the universities enhances the students’ problem-solving attitude, their job opportunities and their creativity. It also caused that a new way of thinking among lecturers appeared besides the positive influence on the students’ preparation. The rate of the responses are presented in Figure 4.
The respondents considered the integration of the disruptive technologies in their education and time management while teaching the dynamically changing teaching materials the biggest challenges. It is absolutely understandable because learning the new technologies demands to have the knowledge which spans through different fields of sciences and it requires occasional supplementation so it is really difficult to teach the related knowledge in the traditional 45 and 90 minute lessons. It is not easy to find the online internet contents particularly in the phase immediately after the appearance of the technologies, but with the adoption of the disruptive technologies and by publicizing them this problem can be solved. The lecturers who are sometimes strained considered continuously updating the learning material another challenge. The disruptive technologies they most often introduce are the Big Data (2012, 2013), Cloud Computing (2012, 2013, 2013), Private Cloud Computing (2014, 2014), Gamification (2014, 2015), HTML5 (2013, 2013), Internet of Things (2014, 2014, 2015), 3D printing (2013, 2014), and Virtualization (2010, 2011).

The following figure presents the position of the particular technologies in the Hype Graph the first time the institutes of the higher education introduced them in their educational materials.

In the representative questionnaire survey it was shown that the disruptive technologies appear in the teaching materials in the higher education with a delay, often at the end of the Technology Trigger after the well-distinguishable technological boom on the Hype Graph following the Peak of Inflated Expectations, in the Trough of Disillusionment. According to David TAM’s theory, the acceptance of new technologies, the innovative pursuits are influenced by two factors. Utility and simple use. All the respondents agreed with the usefulness of integrating disruptive technologies in the educational material and besides these
the needs coming from the globalizing economy also confirm the agreement with usefulness. We can conclude from this that the causes of the late introduction have to be investigated from the point of view of the simple use, even if here it is not effectively, clearly about the use and distribution of the technological innovation.

According to our supposition, among the causes of the late introduction are the differences between the technologies and technics supporting learning available in the higher education and the online learning environment applied in the students’ practices and the difference between the digital life of the students and their lecturers present in the higher education of engineering.

To justify the hypothesis, from the research strategies – descriptive, correlation exploring, experimental (Falus, 2010) – we applied the descriptive and correlation exploring strategies. The realization was made by carrying out and assessing the questionnaire compiled to investigate the Z generation students’ ICT devices, their internet using, information finding and learning habits, and their attitude towards technological innovations and to the requirements of the labour market.

The questionnaire survey of the Z generation at the Faculty of Engineering and Information Technology at the University of Pécs

To set up my questionnaire, I used the EU Kids online research carried out in 2011-2012 for the initiation of the European Union, plus the research questionnaires used by the American Common Sense organisation. In the survey carried out in October 2015 with the BSc students of Computer Science engineers and electronic engineers studying at the Faculty of Engineering and IT, I asked 100 people, aged 18-20 about the trends of science-communication, their learning habits, about their applied devices and technologies, time use and media use, information finding and internet using habits. As an additional methodological factor, I carried out a detailed interview questioning among the university lecturers to get to know their educational experiences obtained in the circles of the Z generation, the applied devices and technologies. Among the lecturers, I carried out a detailed interview concerning the data analysis.

The objectives of the questionnaire:

- To gain relevant information about the digital life of the Z generation, their information finding habits, about their attitude to technological innovations and research results
- To reveal the effects of the digital life of the Z generation students – the so-called cyber-life- on the abilities
- To state whether the students belonging to the Z generation studying at the Faculty of Engineering and IT create an online learning environment connected to their preparation in the higher education individually
- To claim how much the online educational scenes offered by the university are accepted among the students of the institution of the higher education
- I am trying to find the answer to the question what is the optimal learning environment – based on ICT devices and technologies – aimed at the education of the Z generation students like

Hypotheses

- The ICT devices available for the students are beyond the devices provided by the institution of the higher education.
- The students follow the technical, technological innovations, they require them to be integrated in the educational material of the higher education.
• The digital life of the students has an influence on the abilities of the students (multitasking, abilities for rapid decision-making, highlighting the main points and finding the connections)
• The students’ ability for time management shows a correlation with their other online activities.
• The students create learning teams on their own in connection with their studies, and also online learning environment in which they apply the social network and other internet services.
• There is a difference between the digital living space of the Z generation students and the ICT-based educational abilities of the higher education.

Presenting the questionnaire

Our questionnaire contains 114 items connected to closed-ended questions and possible answers to 4 open-ended questions.

The questions asked in the questionnaire can be categorized in eight well-distinguished groups:
1. The demographic background of the person answering
2. Their status in the labour market, the question of employment
3. ICT device and internet accessibility
4. Questions aiming at being informed and at attitudes
5. Questions concerning the digital life of the students
6. Information finding habits and knowledge
7. Time management and time spent on internet applications
8. The students’ requirements concerning the educational environment

I do not present the results of all the questions, I only present the fields that are significant from the point of view of the topic.

The assessment of the questionnaire:
For the assessment of the answers and their statistical analysis I used the IBM SPSS Statistics 23 programme and the Microsoft Excel 2013 programme.
With a questionnaire survey our intention is to be able to collect correct data in connection with the issue without measurement errors. I measured the potential errors, and corrected with the specified statistical methods.

Descriptive statistics:
100 Hungarian citizens between the age of 18 and 20 took part in the research concerning the Z generation’s learning habits, their use of technical and technological devices. 88 of them were men and 12 were women. The shift in the rate of the sexes suits to proportion of the sexes of the population that participate in the higher education of engineering. The responding students study in the day courses. The confidence interval was determined at a reliability level:95%, for a population of normal distribution. The faults of the measuring were assessed, investigated and corrected by using statistical methods (for example the socially accepted aswers, statistical correlation of the yes answering attitude).
Demographic background: 5% of the respondents come from the capital. This low proportion is due to the „drain“ power of the capital in terms of the applications to the institutions of higher education. Most of the applicants are less well-off and come from the Central and
South-Transdanubian regions. The following figure presents the rate of students taking part in
the survey concerning their permanent place of living.

![Percentage according to the permanent place of residence](image)

Fig.11. Figure prepared by myself

Most students live in dormitories, and 27% live with the parents, while 23% share a flat they
rent and 5% live in their own household while doing their studies.
As for their labour market status and employment:
83%, a quite high rate of the respondents work besides doing their studies in the day courses,
and 44 do professional jobs. In the dual training, 16 work with a students’ working contract in
their own profession, and 10 work in the trainee programmes, 18 do a temporary job in
professional fields.

ICT devices and their use
Concerning the ICT devices students were asked if they have a desktop PC, a laptop, a tablet,
an E-book reader, MP3/MP4 player, DVD/Bluray player or a smart phone. The students’
informatics device supply can be claimed very good, and although they have financial
difficulties they are willing to spend large amounts on modern ICT devices. The high number
of these devices can be explained not only with the fact that they are members of a generation
which is characterized by „device dependence” but also with their professional interests.

![ICT device facilities](image)

Fig.12. Figure prepared by myself

The research results clearly show that the students of the Faculty of Engineering and
Information Technology at the University of Pécs primarily used the smart phones, however
the MP3, MP4, DVD and Bluray players have been effaced.
The responding students have been using mobile phones since the age of 8-10, smart phones
since their appearance and computers for 14 years on average.
Being informed, attitude

A 5-scale Likert-scale, which was named after Rensis Liker who first applied it, was used to rate how the students were informed. It was created with the aim of investigating the attitudes in connection with a concept or a person’s particular activity. Concerning its structure, two extremes are designated at the two ends of the attitude scale that instantiate the total opposition (minimum value) and the total agreement (maximum value) determined concerning the statements in the questionnaire. In the middle of the scale (the median value) the neutral feelings are shown. In our case we used a 5-scaled scale where scale 1 means opposition stating „not at all”, scale 5 means agreement stating „completely”.

Concerning the technological innovations:
How much do you feel you are informed about technical, technological innovations?

![Figure 13. Figure prepared by myself](image)

It is clear that students have an interest in technological innovations. No one student said that they never get informed about the technological innovations daily and 78% of the respondents check on the technical and technological innovations weekly. 90% of the respondents claimed that the main scene where they find information is the internet.

87% of the respondents say that the technical and technological innovations should be integrated in the teaching materials of the higher education as soon as possible after they appear.

Internet using habits

In connection with the internet using habits, the aim of the first question was to find out about the scenes of internet use. 100% of the students said that they have internet access at home, and they also have broadband internet access in the dormitories. In addition, 86% of the students subscribe to mobile internet, and concerning the question of „where do you use the internet” they responded „everywhere”. The rest of the respondents, 14% said that they use the internet at home, in the dormitory and also at the university. They actively use the internet 5-7 hours a day saying „because I do not have more time for it”.

As for being informed about the ethical and legal questions, the average score is 3.3 with a deviation of 0.6.

Multitasking is obviously typical in terms of active activities as the figure below shows that the students spend 100% of their time with dealing with several activities at the same time.

Concerning the respondents’ other abilities, the score was an average of 4.2 points for „I can make decisions quickly”. In connection with the question „I found the significant facts from the information available” the score was 4.3, and they had a positive assessment for their
abilities in terms of the question „I can easily find the correlation between the information gained from different places” giving it a score of 4.5.
- How many people do you keep in touch personally on a daily basis? Average: 15.
- How many people do you keep in touch online on a daily basis? Average: 37.

Time management
We tried to find the answer to what connection there is between the frequency of doing different learning, or freetime or entertainment online activities and their time schedule, the ability for time management.
According to the results of the correlational connectional tests, time management shows a positive significant correlation (r = 0.15-r = 0.632; p < 0.01 - p < 0.05), that is between the online learning activities (49-55;58;75. items) and the efficient time management (53, 57, 62, 69, 73, 106. items), at the same time, a negative significant correlation can be shown in case of students giving high scores for watching a video on video-sharing portals and time management (r = -0.24; p < 0.01).

Assessment of science
Concerning the recent results in science there is a clear interest. 87% of the respondents have an interest in scientific and technical innovations, and explorations, too. Scientific innovations seem to be welcome in the social media and online publications are considered more useful than those in the printed press and journals. 19% of the people asked read scientific journals on a regular basis. 17% would also work as researchers. As being students at the Faculty of Engineering and Information Technology, the respondents have a major interest in informatics firstly and in engineering and innovation secondly.

Learning environment, learning techniques, technologies
The students enjoy cooperating and preparing together and prefer collaborative work on online spaces. Common studying motivates them and they willingly spend more time on preparation and preparing with other students than the time they would need individually to achieve success. 100% of the respondents said that they create online social learning spaces for working together. In terms of studying in teams, they claimed several advantages, for example sharing the tasks, time management, and deepening the knowledge after elaborating the parts of the processes on their own, and the motivation which the success they achieved together gave them. In their opinion when studying in teams the participants are all forced to actively deal with the task, and far fewer external disturbing factors appear. According to the respondents, the most popular internet applications that they apply when writing tests and preparing for the exams are the following.
They use the Skype and the TeamSpeks3 which are applied for communicating with each other, also the Facebook to work and keep in touch in a closed community. They also use TeamViewer for sharing the screen with the great advantage of preparing graphs and making explanations, and apply Google Drive to share the contents and to edit the common documents and tables. The Google applications, with a score of 4.4 on the 5-scaled Likert-scale proved particularly popular among the students.
Based on the responses, it is clear that the students do not use the application that the universities provide them for common work, for example Coospace and the Neptun, or if they do even then only for a limited time and only when they need to.
Detailed interview questionning among the university lecturers

During the detailed interview made with the lecturers I asked the colleagues teaching the Z generation students about the ICT facilities, their internet and device using habits and about their educational applications. The chosen colleagues who teach the Computer Science engineer and the electric engineer students belong to different generations. 3 out of 10 colleagues belong to the „Baby-boom” generation (1946-1964), 4 people are the members of the X generation (1965-1979) and 3 people belong to the Y generation (1980-1995). Due to the salaries – not considering them competitive - the characteristics of the higher education of engineering, this age scale is characteristic of the board of lecturers at the University of Pécs at the Faculty of Engineering and Information Technology.

Presenting the results of the interview

The colleagues belonging to the Baby-boom generation have been working in the higher education for more than 40 years. They have desktops both at the workplace and at home, but they do not have their own laptops or tablets. All three of them use smart phones, none of them use MP3/MP4 or DVD/Bluray players. 2 of them have been using a computer for 25-35 years, and smart phones for about 3 years. From the functions of the phones they primarily use the phoning, the text messaging and receiving emails.

On average they spend 1-2 hours on using the internet, mainly emailing, reading the news, searching for information, and about 15-20% of their time is used up by social applications. They are the members of some social network applications, mainly Facebook, but only to private connections. They do not like using the applications in multitasking forms.

During their work as lecturers, they use presentations and videos to support teaching the material more efficiently. Only one colleague approves of the cooperative work and even he rarely uses it explaining it with the difficulty to assess the students’ performance. They all disapprove of the students using their own ICT devices during the classes saying that they disturb the class with these and that these distract their attention from work.

They spend 4-5 hours at their workplace on average. This is justified by a lower number of compulsory lessons. They rarely deal with talent care.

The colleagues belonging to the X generation did not start to work in the higher education. Some started their career in the industry, some in the state education. Their experience in the higher education is 20 years on average. They are characterized by being equipped much better with ICT devices. They have desktops at the workplace, they all work with laptops and 2 people use tablets for their work. 3 lecturers have MP3/MP4 players, 1 of them has DVD/Bluray player.

They have been using computers for 25 years, so from an earlier period of their lives than the „Baby-boom” generation colleagues. They use the functions of the smart phones, and they willingly download other applications. Managing multitasking applications is accepted by them.

During their work as lecturers, besides presentations, they like experimenting with integrating internet applications in the education. They rarely make a facebook group for teaching a specific subject. Their opinions vary about applying cooperative form of work. 2 people disapproves of it – with an explanation that „it requires too much time and energy”, and 2 of them try to use these in their work. As for the virtual educational environment, the situation is even worse. Only one person said that he has ever used such a solution in education. They disapproved of the students’ use their own ICT devices 3-1. Also because of disturbing the flow of work in class.
The average time spent at the workplace is 7-8 hours a day. As for the talent care, they deal with one student a year.

The ICT device facilities of the lecturers belonging to the Y generation is the best compared to their older colleagues, as expected. Although they only have a desktop PC at their workplace, they are willing to spend more time at the workplace because of research work demanding more informatics. They all have all the other devices mentioned.

They willingly make their teaching work more colourful with the opportunities provided by the internet, often searching for and showing videos from youtube in connection with their topics, they know and apply the google services, the new presentation programmes, they choose from them practically as for the application. They willingly use modelling and simulating softwares for education. They do not disapprove of cooperative work, but do not use them in the daily lives either. However, they are interested in the VR environment, although as they admit, they hardly ever use them in education.

They approve of students use their own devices, but they do not prefer using the smart phones in class.

They spend 7-8 hours at their workplace on average.

Summarization of the Z generation survey:

<table>
<thead>
<tr>
<th>Information</th>
<th>Preferred by the students</th>
<th>Preferred/expected by the lecturers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image, sound video</td>
<td>Text information</td>
<td></td>
</tr>
<tr>
<td>Fast, straightforward information finding</td>
<td>Gradual, limited number of „step-by-step” information transfer</td>
<td></td>
</tr>
<tr>
<td>Relevant, promptly usable information</td>
<td>Information suitable for the curricular directives</td>
<td></td>
</tr>
<tr>
<td>Information helps learning and is entertaining at the same time</td>
<td>Standardized information helping the education</td>
<td></td>
</tr>
<tr>
<td>Random multidisciplinary information „accessible through hyperlinks”</td>
<td>Linear, logically constructed professional information</td>
<td></td>
</tr>
<tr>
<td>Working style</td>
<td>When „we feel like it”, not time-bound working</td>
<td>Keeping the frames of the timetable</td>
</tr>
<tr>
<td></td>
<td>Often „Just in time” performance</td>
<td>Preparation is pre-planned, ready for any eventuality, keeping the requirements in mind</td>
</tr>
<tr>
<td></td>
<td>Team work, cooperative activity</td>
<td>Well distinguishable individual performance</td>
</tr>
<tr>
<td></td>
<td>Real time online communication</td>
<td>Physical presence in class</td>
</tr>
<tr>
<td>Assessment</td>
<td>Prompt confirmation, stepping further</td>
<td>Deferred assessment</td>
</tr>
<tr>
<td>ICT device, software</td>
<td>Smart phone, tablet</td>
<td>Devices recommended by the lecturer</td>
</tr>
<tr>
<td></td>
<td>Use of social media</td>
<td>Use of software strictly connected to the educational material</td>
</tr>
</tbody>
</table>

Table 1. Prepared by myself

The Z generation students and the lecturers teaching them are significantly different concerning their ICT device supply, their habits in terms of using the internet and their digital life. As for the use of smart phones, it is the primary device for the students while their lecturers do not approve of them being used in the lessons. The students use the internet applications to which only their peers in the same social group have access as the common
scene for studying. It provides the opportunity of real-time communication and with that they can also keep the written comments and the explanations. For a better understanding, their useful devices are also available with which they can edit drawings and make demonstrations for example by graphs, and with some applications they can also share the screen. The students create their VR learning environment, although the lecturers rarely know and apply these applications in their teaching.

The Z generation students apply the ICT devices more confidently than their lecturers who are members the other generations. Their cyber-space life integrated in the ICT has a developing effect on the students’ practical information finding and processing abilities. They often hear about the technical innovations earlier than the lecturers teaching them.

According to the results of the above-mentioned surveys, we decided to create an ideal educational environment which suits both the employers’ and the students’ requirements. We found the realization of these requirements in the creation of a virtual 3D educational environment which I will present below.

Today the new technologies facilitate new ways of learning widening our attitude towards learning. According to the connectivist learning theory, „learning is a process where the information exchange, which is informal, networked and supported by the electronic devices, has an increasing role.

In the VR space the educational technological and educational methodological conditions are perfectly suitable for arranging courses with the connectivist learning approach, but to increase the efficiency of learning we extended these with more elements.

The advantages of the created Cyber-learning system

We applied the 3D visualization in the VR learning environment created in the VirCA -an alternative of MaxWhere - platform developed by the MTA SZTAKI which suits to the natural cognitive processes of the human brain better. The 3D visualization also suits to the preferred scened of the students’ digital life, thus it ensures a suitable comfortable feeling for work, at the same time helping to forget about the working and learning features of the activity done here. By merging the cooperation resulting from team work, the project method, the investigating and exploring methods and Gamification we further enhance the students’ motivation. We exploit the opportunity of working together and learning from each other with cooperation, merging theoretical knowledge and practical application with projects, developing the competences essential for the process of gaining knowledge and the practice of setting, proving and disproving the hypotheses with exploring methods, and making learning an experience with gamification. The „Cone of Learnig” based on experiences and created by Edgar Dale, the renowned pedagogist, infers that depending on the type of activity, the retention levels of information will increase or decrease. As seen, Active Learning stimulates higher cognitive processes.
The Cyber Learning solution enforces the activities shown in Edgar Dale’s cone of learning in the most optimal way. As another advantage, we must highlight that the previously made e-learning materials can be used smoothly in the VirCA - an alternative of MaxWhere - environment, the opportunities provided by the google applications (Google Drive, Google Calendar, Translator, News etc.) can easily be integrated in the system, so there is no need to design one’s own website to share the self-developed contents needed for the education. The applications embedded in the cloud, such as chat, video etc. ensure the opportunity for continuous contact and real time communication in the virtual space.

VirCA – Virtual Collaboration Arena

The Computing and Automation Research Institute (SZTAKI), of the Hungarian Academy of Science (MTA), the Széchenyi István University and the Technical University of Budapest commonly created the VirCA - an alternative of MaxWhere - platform with the help of national and international projects. The platform first appeared internationally with grandious VR abilities for which they received the special prize (EU-FET’11 Award Cross Exhibit) in 2010 at the IROS conference and exhibition in Thaiwan, and in 2013 it also won the TUV Rheiland innovation prize. The advantage of VirCA is that it is a platform able to visualize many different VR and AR projects which is also capable of 3D visualization, thus can be fully applied for training the future engineers. To use it, there is no need for special visualizing devices, it can be run and enjoyed on a PC with a suitable resource and on laptop, so it does not mean an extra investment for the students.

Creating the VR learning environment for teaching the memristor

As a pilot programme, we created a VirCa cooperative learning environment to teach the memristor, as a disruptive technology, to the students at the Faculty of Engineering and Information Technology of the University of Pécs. We created a learning team of six people in which three IT engineer and three electrical engineer students took part. We chose students from two different courses because of the topic, the memristor – a new circuit element. Its application in informatics can be expected. We primarily wanted to achieve – by integrating the applications which are suitable for the CE generation’s digital life in their learning space – that the students could work without leaving the learning environment, but meanwhile they
could also use the social media applications, apply search engines, were able to watch videos and to communicate and make the solutions concerning their task. According to our suppositions, with the fact that they can access everything in an environment which they know and prefer, the boundaries between learning, carrying out tasks and leisure time activities are eliminated, so the students are willing to spend a significantly more amount of time and energy on doing their projects. To enhance the effect, we have exchanged the teacher’s role for a coach role suit the coaching technique applied successfully in the company trainings in the market sector. The steps to create the course became significantly simpler provided by the opportunities of the VirCA platform, as we could experience it in the case of the „classical” connectivist courses.

The steps of the course titled „Teaching the memristor in VR space”:
1. Creating the name and goal of the course
2. Assembling the theme
3. Assembling the „recommendation” (related electronic materials, presentations, videos, learning materials, URL addresses) which continuously expanded, dynamically „developed” with the help of students compared to the starting point
4. Defining the tasks
5. Assembling the questions for the specific tasks according to the coaching technique
6. Contents management
7. Assembling the elements of the Cooperative VR learning environment suitig to the topic

We created a futuristic „capsule” as the scene for the research work distant from everything and when the students enter this they can notice at once that with the internet support the digital world opens up its endless opportunities to them. Information – connected to the topic in the focus of the education – arrives on loads of channels, parallel to each other so ensuring to satisfy the CE generation’s hunger for information and to maintain the continuous interest. When we equipped the virtual space we had to care about choosing the suitable equipments and applications which provide the greatest cooperation opportunities. The students were able to move in the room freely and were allowed to see each other’s work, could communicate without limitations with the other students and their lecturer who helped them as a coach.
We put a surface which we called a huge „notice board“ or „presentation space“ in the centre of the round room which they could use to present the information that they wanted to share, revise or prepare with the members of their team in their brainstormings. As Figure 3. shows 5-5 freely changeable surfaces are provided on each side of the board. Here they can actually put any kind of contents, e.g. videos, presentations, plans, publications, even questions which help to carry out the project successfully.

We put a virtual presentation space which was made of four units on the two sides of the board in the opposite part of the room. On each side, one of them was uploaded for the IT engineer students and the other for the electrical engineer students with the initial information of a lecture presentation material, an awareness-raising motivating professional video, a chat board and a surface for browsing.

Figure x. shows a part of the initial surface created for the IT engineer students in the middle of which there is the lively presentation of the HP „The Machine“ project awaiting the students (the fourth screen missing from the picture includes the professional presentation.) In Figure y., on the surface of the electrical engineers Leon Chua, the inventor of the memristor talks about what the first thought was to make him think that the fourth basic circuit element
should also exist. With showing the HP film and Leon Chua’s presentation we had the aim of making the students meet the technological innovations concerning their professions in the original environment first, and at the same time to increase its awareness-raising effect. (The notice surface helping the students’ common thinking is missing from this figure.) Both of the teams could alter these contents occasionally as they carried on with their own projects. We can say that with a range of the cooperative and collaborative activities the professional teams created the materials prepared for the „notice board”.

Both teams were given 2-2 working desks where – besides the online spaces - a clock was also placed for being able to keep the deadlines which had been set for them earlier, and a calculator was also available for them to use it for the calculations.

In a room equipped this way, all the students had the chance to create their working environment according to the task which they shared. The main rule was that 1-1 student in every team in each project worked at the working desks where they had all the equipments seen above available. The most common contents that supported the students were the Facebook, the chat, the corresponding programmes, a browsing surface and a professional programme surface which was accessible online, so they did not need to waste time when they were choosing these applications. The third member of the team who was also usually the project leader coordinated the 4-unit presentation space mentioned above.

In this virtual space the element of the psychologically motivated learning definitions so correct in the every days of the pedagogy was created, according to which learning is not only finding information, but also forming the attitude, as the students learnt in the VR space how to behave during a project. What factors they should consider in distributing the roles, how they should behave in the roles given to them. To take responsibility for their work, to learn the importance of managing time, keeping the deadlines, to learn the competences and skills needed during the team work, to focus primarily on the goal to be achieved. They also successfully developed their presentational skills demanded by their future employers, their communicational skills developed through presenting and defending their professional viewpoint. They experienced the need for professional English knowledge which lead to developing the competences of the professional foreign language.
Presenting the applied educational method

The 3D VR space itself has a remarkable motivating effect on the students and this effect is further enhanced by the application of the edu-coach method which I had already tried in the traditional learning environment in the higher education of engineering, which includes cooperativity, the project method, the investigating and exploring method and the Gamification method supplemented these with the techniques applied in coaching. I think it is important to present the main aspects of the application of the edu-coach method in the VR space because in my view it helps to understand why not to worry about the students not carrying out the tasks they undertook.

1. At the first meeting, as the „amazement” at the VR space greatly helped the intention to participate in the educational project, creating the goal of the course is done commonly in the planning stage. The main question is what is the desired knowledge in the future which we want to have at the end of the project? This, in our case, was to explore and find out the new fields of application of the memristor. The goal in itself included the „excitement of playing” activated by the creative, innovative attitude. According to the students’ accounts it was remarkably inspiring for them to try „whether I am able to create something new, a technically grounded device, development?”

A significant element of the planning stage is to define the time of the meetings when every participant checks in at the same time in the virtual space and participates in the work. Planning, time management and concentrating on the task require the creation of this foreseeable system. These meeting times usually coincide with the closing time of a sub-project.

2. Edu-coaching ensures a theoretical frame for the students to have an opportunity to invent and develop new solutions and as a result of this they could consider the achieved success as their own. This theoretical frame in practice means that the initial information has to be assembled after proper consideration, and a too extended task should not be set at a time. Between the present and set system of knowledge there is a contradiction in most cases. During this process we create hypotheses continuously and then we check them. The final goal is to be able to place a given problem in the framework of a greater, more general system of concepts so that, by this, these could also be managed securely in new situations. Overviewing the background information should be smartly done in a way that the students themselves would realize which of their knowledge earlier gained they could rely on and what they should overview for stepping further.

3. Creating the students’ self-assessment system is also a significant element which is done at every meeting because this helps to maintain the feeling of commitment and greatly contributes to establish responsibility and goal-oriented work. It is usually a frequent question: „On a scale of 10 where 10 point means the maximum completion of the task, where would you assess your present state in the working process?” What should be done to achieve a value of one score higher?” Thus, the students themselves define the next short-term goal to the question being asked. As long as the student cannot say what the next expedient step should be then the members of the team commonly plan the process of stepping further.

4. In the reflection and problem-solving stage which accompanies the learning process, we must concentrate on the solution and the way which leads to the solution and not on the problems, thus if the students get stuck with their task then we are not seeking the cause of the problem with targeted questions (for an experienced lecturer it is well-known for the teacher), but we are trying to find the answer to how we can help
in stepping further so that the students themselves could solve it. These „clarifying” discussions can be done in private talks or in groups.

5. Following the rules of project management naturally the closing of the work cannot be omitted. It is now that it is discussed how the goal has been achieved, what worked, what did not work, what experiences can be carried on and what conclusions can be drawn from the process done. Last but not least, the common joy of successful work positively confirms the joint work in the future projects.

Conclusion:

The 3D VR learning environment stimulated the students’ learning by being suited to the CE generation students’ digital life. They said that they had enjoyed working together, the challenges and that they would go on with their studies in several topics in a similar 3D VR environment.

In the environment which was made suitable for the cooperative work, they could watch not only their own work but also their peers’ work and performance which had an encouraging effect on completing the good quality tasks by the deadlines.

The students suggested new and creative solutions for applying the memristor, for example as a solution for over voltage or flash protection. Obviously it is time-consuming and also tiring for the lecturers to prepare the work and accompany the tasks, however the professional success can compensate them for these.

The cooperative VR learning environment with the advantages of the 3D visualization created by the VirCA platform is completely suitable for training the future engineers. Because of its project-based educational opportunity it ensures an efficient educational scene for the higher education of engineering where the 3D planning and manipulation are part of the every day work.

With the education done in the presented 3D VR environment we can meet the employers’ requirements, that is besides the professional preparation they draw up a strong need for the development of the so-called soft skills, such as creative thinking, critical thinking, problem-solving, innovative thinking, doing project work, cooperative work, development of the communicational and presentational skills, foreign language knowledge.

At the same time the VR space facilitates the education of the disruptive technologies- which appear more and more extensively – even immediately after their appearance and this way help the training of creative and innovative engineers.

In conclusion, we can say that after the cautious preparational work (questionnaire surveys and their assessment) and the comprehensive planning we closed a successful project with the integration of the 3D VR learning environment - which we created - in the higher education of engineering.

Among our further plans we have the creation of a virtual laboratory environment which facilitates the preliminary preparation for the laboratory measurement practices necessary for the engineering profession and then the compiling of the protocols based on the measuring results.
References


Abstract — In this article, a Brain Computer Interface (BCI) based speed controlled mobile robot is introduced. The mobile robot can be controlled via WiFi wireless network by the BCI application. This topic strongly belongs to Cognitive Infocommunication and BCI research areas in the field of psychological, information and collective infocommunication. The implementation of the BCI controlled mobile robot helps to understand the design, and the application, programming and wireless communication use of embedded systems and BCI based cognitive technology. The mobile robot may be a quite useful practical device in education as well, regarding the fields of informatics and mechatronics.

I. INTRODUCTION

The artificial cognitive skills of machines are steadily becoming more complex. The cognitive abilities of humans and machines are increasingly manifested in a co-evolve life form. It has psychological, information and collective infocommunication based connections as well as direct hardware contact levels and their combinations. This research deals with an example of a hardware-BCI-human cognitive chain.

Brain Computer Interface (BCI) is an emergent technology to build a direct communication channel between human brain and the computer. [1] BCI reads the waves produced by the brain and translates these signals into actions and commands, which can control the computer. [2, 3] The BCI can be used in many applications for example playing games [4, 5], social interactions by detecting emotions [6, 7], or to help disabled persons [8-11]. EEG system has been widely used in the field of BCI. [12, 13] Mobile robots are getting more and more important in our everyday life. By the development of electronics, embedded systems and informatics, the price of mobile robots significantly decreased, while their field of use increased. One of the domestic uses of mobile robots is the field of hoovers or lawn mowers. Hardware design of such a robot is not too complicated, so the simplified version of such a robot containing its control system and sensors required for moving can be assembled using cheap parts. In this article, a BCI based speed controlled mobile robot is introduced, which can be controlled via WiFi wireless network by the BCI application.

II. HARDWARE OF THE MOBILE ROBOT

During the development of the mobile robot, simple design, upgradeability, and wireless control were the main goals. In case of educational use, the application of modern ICT based technology and knowledge transfers offer new possibilities [14, 15]. The developed device can be used well in education, because simple and clearly arranged layout, control features have to be achieved. Motor control of the mobile robot and the WiFi wireless communication are usually performed by two separate units; one unit, which performs motor control, is generally a microcontroller, while WiFi communication is ensured by a separate communication unit [16]. In case of a cheap, easy construction mobile robot, such a solution is preferable, where the microcontroller and WiFi communication units are achieved in one IC. For example such a cost-effective IC is the ESP8266 chip based modules (manufactured by Espressif Systems), available since 2014. When creating the mobile robot, this module was chosen. The main features of ESP8266 IC are summarized in Table I.
The device is available in differently designed modules, for development purposes the ESP-12E was chosen, which contains integrated antenna, and important outputs of the IC are available. For the module, a 2nd generation LoLin development board was chosen, which contains the 3.3V voltage controller required for USB power supply, and USB-UART serial converter based on CH340G IC for programming via USB port. (Figure 1). Connection of other external modules to LoLin developer chip is made easier by a separate I/O adapter module (Figure 2).
Programming of the device can be accomplished via UART connection, so for programming, no special (and usually expensive) programmer tool is necessary. For the main purpose of the mobile robot, primarily the simple-design, upgradeable chassis and the control of mechanic movement was preferred. Thus, for performing control mechanism, driven front wheels and a free rear wheel was chosen. For chassis of the mobile robot, such design is expedient, which makes placing other accessories required for further developments, and sensors possible. On the international market, several kits containing the chassis of the mobile robot are available, and such solution was chosen, where the battery holder responsible for power supply can be placed on the lower part of the chassis, leaving more free space for other units on the upper part of the chassis (Figure 3 and 5).

3. 2 Wheel Drive Mobile Robot Kit

4. L9110S motor control module

The mobile robot kit contains a removable coated transparent plexi chassis, two driven wheels, two direct current motors with 1:48 transmission ratio, one free wheel, battery holder and other accessories required for assembly. In the battery holder, 4 pcs AA type 1.5 V battery can be inserted, so the supply current is appr. 6 V. The current of the direct current motors measured during operation are below 500 mA, so for their supply, a module built on L9110S motor controller’s IC was chosen, whose operating voltage range is between 2.5 and 12 V, while its continuous output current is 800 mA (Figure 4). The assembled mobile robot containing modules placed on the chassis, and also the motor and power supply input is shown on Figure 5.
III. SOFTWARE OF THE MOBILE ROBOT

For programming ESP8266 IC, there are several developer environments available. For the development of embedded systems’ software, generally C or C++ programming languages are used, which are also supported by the freeware Arduino development environment, so this one was chosen. In the development environment, the ESP-12E module is not among the supported developer tabs, but with the help of ‘Boards Manager’, it can be easily added. For programming the developer tab, driver of CH340G IC performing USB-UART conversion needs to be installed, which creates a virtual serial port on the computer. For programming, the proper serial port has to be selected in the development environment, as well as the 115.2 kbit/s bandwidth. Programming is supported by the ESP8266WIFI function library. By the function library, connection to WiFi network, creation of server making the connection of external clients possible, and the client-server data connection can be handled easier. The PC control software, sends the speed parameters of the robot motors (in the range of 0...100), connecting to the control unit of the mobile robot (as server) as a client, via TCP/IP connection. The speed of the mobile robot is defined by the speed of the controller-motor unit supplying each wheels, which in case of permanent magnet direct current motor. In case of permanent load the supply voltage of the motor is proportional to the speed. The effective, low-loss control of the motor’s speed is performed by switching operation, the speed control was performed by pulse width modulation (PWM). Digital outputs of ESP8266 IC are capable of creating direct pulse width modulated signal, whose default base frequency is 1 kHz. So in the program code, speed control was performed by the value defining the duty cycle of PWM output sent by the PC control software. The flow chart of the program ensuring speed control of the mobile robot is shown on Figure 6.
6. Flow chart of the mobil robot’s control program

IV. BCI SYSTEM
The brain-computer interface is a direct communication channel between a biosensor based brain signal processing unit and an external device. BCI systems can be divided into two major groups according to the measurement of bioelectrical signals. In case of EEG, where biosensors are directly placed into the brain, we are talking about invasive BCI, otherwise, if the electrodes are placed somewhere on the head, for example on the forehead, we are talking about non-invasive BCI [17-19]. Due to the dangers of invasive process, non-invasive BCI systems are preferred. Information process flow of BCI systems are built up by four main components. Figure 7. shows a typical BCI system process flow.
7. BCI process

The Signal Acquisition component realizes the signal acquisition, filtering, amplification and digitalization functions. The electric signals could be recorded from the electrodes. Because the amplitude of captured signals are low, they need to be amplified, then digitalized to process by computer algorithms. The digitalized EEG signals are pre-processed (i.e. digital filtering) and main features are extracted and classified by complex signal processing algorithms (i.e. Fourier method). These features can be used to control in external devices or computer. To perform BCI based speed control the MindWave EEG headset headset was used to measure and preprocess the brain signals. The MindWave EEG headsets are using single sensor on frontal lobe and neutral points on ears by ear clips. Figure 8 shows the headsets and their main parts. The Neurosky uses the ThinkGear Technology to process and classify the EEG signal based on quantitative approach. The device can calculate the eSense Attention and Mediation values between 0 and 100. The eSense Attention value was used to change the speed of the mobile robot.

8. MindWave EEG headset

V. BCI SOFTWARE

Visual Studio development environment was used to implement the PC based BCI appalication on Windows operating system. The FormOfBrainWaveRead partial class handles events happening during the application, and can be considered as the extention of the Form base class. The FormGraphics describes the summary of data visualization, and depends on the FormOfBrainWaveRead class. The DataSend object is responsible for sending the speed data for the mobile robot. The logical structure of the application can be seen on Fig 9.
The article presents the implementation of a BCI system, which controls the speed of a mobile robot. The device can be well-used as a tool for those, who are interested in this field of cognitive technology, mechatronics and informatics. This cheap, easy-to-build system even for students can be well-used for educational purposes too to support ICT based learning [14] and knowledge transfer [15], which helps to understand the design, and the application, programming and wireless communication use of embedded systems through practical examples. Summarizing the results the article shows a BCI system and its application to control the speed of a mobile robot. Of course, this method has got some disabilities yet, but by further development of EEG technology, more and more accurate controlling will become available, which may lead to a wider-spread use of such systems.

REFERENCES


