

New Research Methods for Media and Cognition Experiment Course

Yi Yang^(✉), Shengjin Wang, and Liangrui Peng

Department of Electronic Engineering, Tsinghua University, Beijing, China
{yangyy, wgs gj}@tsinghua.edu.cn,
plr@ocrserv.ee.tsinghua.edu.cn

Abstract. With the development of human-brain cognition and signal processing techniques, there is more attention on media and cognitive disciplines, especially focus on human-computer interaction and human's brain function analysis. Electronic media is a new expression of human civilization, culture and arts. Media and cognition experiment course is to complete the goal of training talents through a large number of state-of-the-art methods. This paper describes the understanding of the new practical engineering projects on media and cognition course. Students were asked to complete several sets of practical engineering courses. Some optional contents are also included. After this training, we were able to select and train more high-level talents further. In fact, this kind of practical engineering course can improve the students' ability to grasp related knowledge points. Eventually they will have the ability to plan projects and solve practical problems.

Keywords: Media and cognition · Analysis of human brain · Human-computer interaction · High-level talents · Investigation of project programming

1 Introduction

Electronic information science and technology serve the people as the electronic media. The electronic media is a new human civilization carrier which will give birth to new culture and arts. Media is defined as the information carrier and can be classified as three parts [1–3]:

- Substances in materials or substances entity;
- Fluctuations signals of matter and energy;
- Symbol carrier, exist and have an effect by the means of two types of carrier mentioned above;

Media information technology research topic included three parts:

- Text: Text retrieval, text classification, text summarization, machine translation;
- Image and video: Video encoding, video summary, target detection, tracking, identification, 3DTV;
- Voice: Speech coding, speech synthesis, speech recognition;

Bill Gates first proposed the concept of “natural user interface” in 2008, and he predicted that human-computer interaction will have a big change in the next few years which means keyboard and mouse will be gradually replaced by more natural module such as touch, vision and voice. At the same time, “Organic User Interface” began quietly rising which includes biometric sensor, skin display, and even directly connection between brain and computer. These technologies will undoubtedly give a significant impact on human’s life. With the application of computer technology and sensors, the real world has gradually emerged its “Digital Edition” side, and natural human-computer interaction is bridge between real and virtual world.

Media and cognition course is the latest created course of the Department of Electronic Engineering of Tsinghua University. This course is to complete the goal of training talents through a large number of state-of-the-art methods. The implementation of all the projects will allow students to deeply understand the basic signal processing methods on media and cognition course. Students were asked to complete several sets of fundamental engineering projects to establish their modelling method and the algorithm programming skills through these practices. Some elective contents are also included to inspire their related fields of research and analysis capabilities. After this training, we were able to select and train more high-level talents further. In fact, this kind of practical engineering course can improve the students’ ability to grasp related knowledge points. Eventually they will have the ability to plan projects and solve practical problems.

2 The Curriculum Contents of Media and Cognition

When confronted with another person, the brain immediately focus on him and identified his identity based on the experience. This process is not through hundreds of layers of decision tree to realize. The human brain is to know. A little baby is difficult to distinguish two different people, but adults can do it through years of study and training. In fact, the human brain may also be able to accurately guess their age, gender, mood, or personality. The purpose of the course is to create a human-like cognitive technology equipment and methods. The purpose of the course is to create a human-like cognitive technology equipment and methods. This technology will observe the world around it and operate and interact with a human user. It can conduct its independent study, and even affect humans to produce some new culture and art. It revolutionized human’s knowledge and means by learning from and interaction with the outside world and other human beings.

To cultivate the high-level talents on this field, we design four kinds of fundamental projects as: three somatosensory entertainment or games based on human-machine interaction; Android-based human face recognition system.

2.1 Somatosensory Entertainment and Games

We designed a variety of entertainment and games somatosensory topics for students to choose and develop. The development platform is Kinect device and its SDK toolkit.



Fig. 1. Kinect device

Kinect is a motion sensing input device by Microsoft for the Xbox 360 video game console and Windows PCs which is shown in Fig. 1. Based around a webcam-style add-on peripheral for the Xbox 360 console, it enables users to control and interact with the Xbox 360 without the need to touch a game controller, through a natural user interface using gestures and spoken commands. Our projects are developed with Kinect Software Development Kit released by Microsoft for Windows 7. This SDK will allow developers to write Kinect apps in C++/CLI, C#, or Visual Basic .NET [4–6]. And the parameter index of Kinect device is:

- The output video frame rate of 30 Hz
- 8-bit VGA resolution (640 × 480 pixels)
- The best recognition region 1.2-3.5 m, 0.7-6 m extended area
- Visual area: horizontal 57° vertical 43°
- Up to track 20 individuals body node

By using multi-channel media interface technology, virtual reality technologies become the future development trend of human-computer interaction. To achieve the objectives of natural human-machine interaction and multi-dimensional information space interaction which is known as “human-machine’s harmony”, we need to use a variety of media to identify human’s body posture, gestures and voice, etc. and to determine person’s intention. Somatosensory entertainment and games are good topics to bring a new awareness of students’ experience which included:

1. Gymnastic Posture Correction and Scoring System:

Students need to design multiple gymnastics pose [7] by Kinect’s interactive features for users to guide user’s gymnastic posture by voice commands which is shown in Fig. 2. The system will compare the degree of difference between the standard and the user’s skeleton node data and give the corresponding scores. According to the degree of difference, the voice interaction wrong posture correction and scoring errors is announced. This system can correct user’s yoga action and correct user’s body shape to keep health.

The core of Kinect skeleton track processing is CMOS sensor to perceive the environment no matter how ambient lighting conditions. Firstly, the sensor generates the depth image stream at a rate of 30 frames per second and the real-time 3D reproduction of the surrounding environment. Next, Kinect will evaluate the depth image on pixel-level to identify the different parts of human’s body. Next, Kinect will evaluate the depth image on pixel-level to identify the different parts of human’s body. The final step is to use these results to generate a skeleton system by tracking human’s joints.

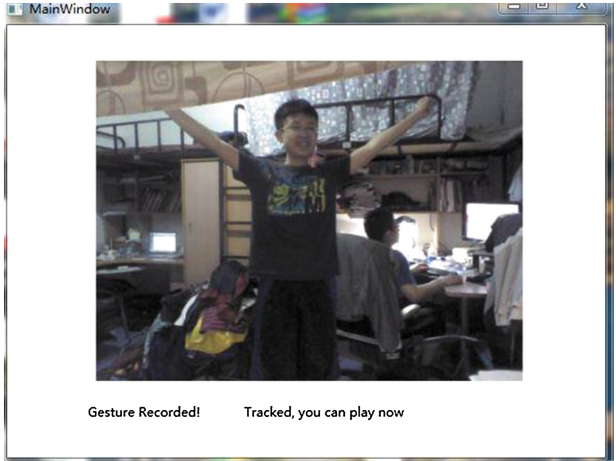


Fig. 2. Gymnastic posture correction and scoring system

2. Motorcycle Driving Games System:

With Kinect device SDK toolkit, students design a human-computer interaction motorcycle driving game [8], which is shown in Fig. 3. Students need to design a menu operation and interface operation mode for the game. The importance is the “two can always switch the operating mode”:

- In the beginning, program is shown as the menu operation interface in the default mode. The user can change the gesture to select the menu’s item included entry, exit and other operations;



Fig. 3. Motorcycle driving games

- After entering the game, user can do the gesture “hands together” to achieve the operating mode switch to enter the somatosensory game mode. Then the user can use his body position to play the game.

According to the body and gestures by the user to simulate the driving motor of the acceleration, deceleration and stopping. They also design the driving the process overturned and overtaking other skills.

3. Music Knocking Drum Games:

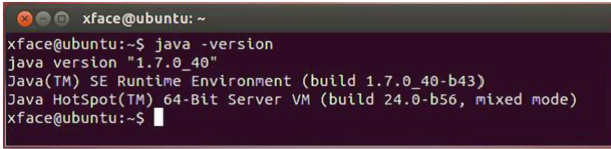
The main problem of the music rhythm interaction through PC’s keyboard is that person have to imitate the “Drumming” action by pressing a key, the realistic action is too low to form a good user’s experience. Realization of gestures by Kinect equipment can enable users to directly operate by imitating drumming gestures which will greatly enhance the game’s experience degree. Another benefit is the exercise effect. This game is designed on the existed music knocking drum games platform, which is shown in Fig. 4. Simulating knocking drum by musical rhythm matching according to the rhythm of the music where the user data is from the human-computer interaction device. The final ranking and achievements is announced by the synthesis voice.



Fig. 4. Music knocking drum games

2.2 Android-Based Human Face Recognition System

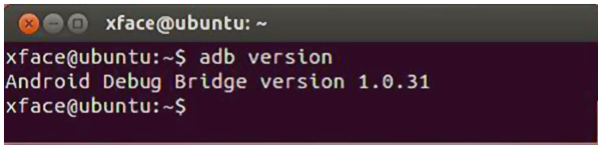
Smart phones and other mobile devices are operating in increasingly rich settings that include both nearby sensors and machines [9]. The android-based human face recognition system is developed on the Linux environment [10, 11]. The Android-based human face recognition system is optional item. But what’s interesting is that many students chose this topic. The Linux configuration environment is shown in Fig. 5:



```
xface@ubuntu: ~
xface@ubuntu:~$ java -version
java version "1.7.0_40"
Java(TM) SE Runtime Environment (build 1.7.0_40-b43)
Java HotSpot(TM) 64-Bit Server VM (build 24.0-b56, mixed mode)
xface@ubuntu:~$
```

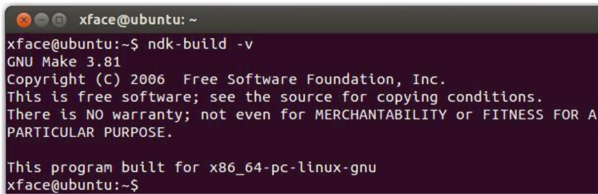
Fig. 5. Linux configuration environment

After finishing the configuration of Android SDK, The Linux environment is shown in Fig. 6:



```
xface@ubuntu: ~
xface@ubuntu:~$ adb version
Android Debug Bridge version 1.0.31
xface@ubuntu:~$
```

Fig. 6. Linux configuration environment



```
xface@ubuntu: ~
xface@ubuntu:~$ ndk-build -v
GNU Make 3.81
Copyright (C) 2006 Free Software Foundation, Inc.
This is free software; see the source for copying conditions.
There is NO warranty; not even for MERCHANTABILITY or FITNESS FOR A
PARTICULAR PURPOSE.

This program built for x86_64-pc-linux-gnu
xface@ubuntu:~$
```

Fig. 7. Linux configuration environment

The Project included the following modules as shown in Fig. 8: Training and Testing. Training module included Face data input, Pre-processing, Feature Extracting, Feature Database; testing module included Face data input, Pre-processing, Feature Extracting, Feature matching. The project is based on principal component analysis (PCA) algorithm. (Figure 7).

- Training module: The training set is 40 individuals and each person 10 kinds of gestures selected from AT & T Laboratories Cambridge ORL face database [12]. Each two-dimensional face gray-scale image is converted into a row vector and calculate the feature vector set by saving all the row vector into one matrix. Then compute the eigenvector and eigenvalues of covariance matrix to produce the Eigen-face. Finally, the selected principal components of Eigen-face are obtained to identify the training and testing face images.
- Testing module: In the testing phase, the testing face image is projected to the Eigen-face subspace and use nearest neighbor classifier with Euclidean distance as a decision. The minimum distance between training image and test image is the criterion of matching.

The final results running in the Android platform is shown in Fig. 9. The registered users' face will be identified and achieved 80 % recognition rate above.

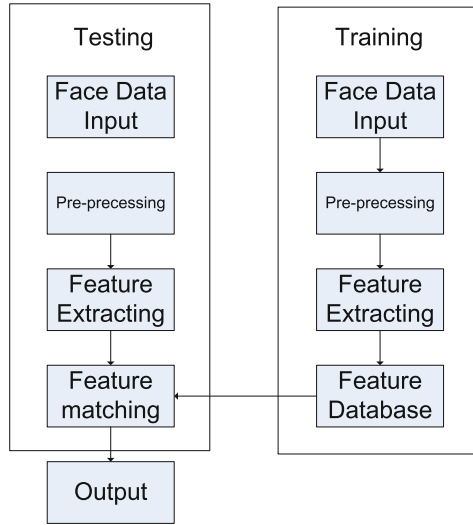


Fig. 8. Flowchart of face recognition based on Android platform

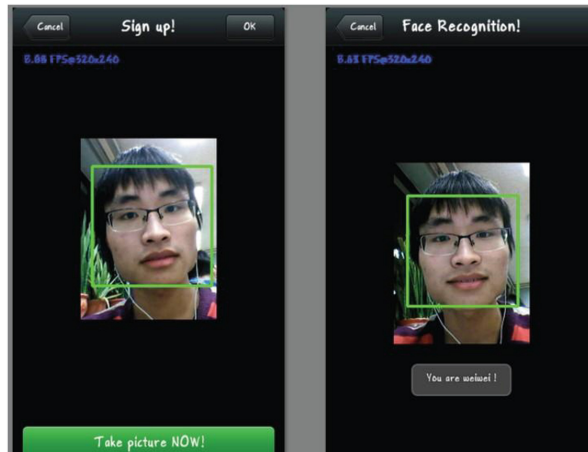


Fig. 9. Android-based human face recognition system

3 Summary

These lively and attractive media and cognition projects will encourage students to broaden their thinking and explore the unknown information researching fields. Students wrote their scientific papers and patent after learning the latest scientific and technological achievements. In this process, they will have deeper understanding of human-computer interaction and pattern recognition. In addition, the course will provide independent practical subjects for some excellent students. These students will have more discussion and development in this field which greatly stimulate their interest.

Acknowledgements. Thanks to Tsinghua University Laboratory Innovation Funding.

References

1. Nick, P., Yonghuai, L., Peter, B.: 3D Imaging. Analysis and Applications. Springer, New York (2012)
2. Wenjun, Z.: Introduction to the new digital media. Fudan University Press, Boston (2009)
3. Zheng, S., Fang, F., Jiongjiong, Y.: Cognitive Neuroscience Introduction. Peking University Press, New York (2010)
4. <http://en.wikipedia.org/wiki/Kinect>
5. Fabian, J., Young, T., Jones, J.C.P., Clayton, G.M.: Integrating the microsoft kinect with simulink: real-time object tracking example. IEEE/ASME Trans. Mechatron. **19**(1), 249–257 (2014)
6. Tao, G., Archambault, P.S., Levin, M.F.: Evaluation of kinect skeletal tracking in a virtual reality rehabilitation system for upper limb hemiparesis. In: 2013 International Conference on Virtual Rehabilitation (ICVR), pp. 164–165. 26–29 Aug (2013)
7. Nakamura, T., Nishimura, N., Asahi, T., Oyama, G., Sato, M., Kajimoto, H.: Kinect-based automatic scoring system for spasmodic torticollis. In: 2014 IEEE Symposium on 3D User Interfaces (3DUI), pp. 155–156, 29–30 March (2014)
8. Chaperot, B., Fyfe, C.: Improving artificial intelligence. In: 2006 IEEE Symposium on A Motocross Game, Computational Intelligence and Games, pp. 181–186. May 2006
9. Jang, M., Schwan, K., et al.: Personal clouds: sharing and integrating networked resources to enhance end user experiences. In: 2014 Proceedings IEEE INFOCOM, pp. 2220–2228. April 27–May 2, 2014
10. Kemp, R., Palmer, N., Kielmann, T., Bal, H.: Cuckoo: a computation offloading framework for smartphones. In: International Conference on Mobile Computing, pp. 59–79 (2010)
11. Ra, M.R., Sheth, A., Mummert, L., Pillai, P., Wetherall, D., Govindan, R.: Odessa: enabling interactive perception applications on mobile devices. In: Proceedings the 9th International Conference on Mobile Systems, Applications, and Service, pp. 43–56 (2011)
12. <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>