Usability of facial expressions and gestures in recognizing learning states in the process of e-learning

Yingyue Xu
Department of Computer Science and Engineering
University of Oulu
yixu@mail.student.oulu.fi

STATEMENT OF THE PROBLEM
Nowadays distance e-learning has become one of the most widely used approaches in the delivery of education. Unlike those courses given in a real classroom, distance courses are not delivered face-to-face, which hinder the delivers to receive accurate feedbacks. However, with the development of facial expressions and gestures recognition studies in recent years, it might be possible to apply such techniques to distance learning, so that data about how listeners master the knowledge of the courses can be collected by analysing their behaviours during the learning process.

The purpose of this study is to evaluate users' learning states during the distance learning process by using facial expression and gestures recognition methods to provide firsthand feedbacks to providers.

The research will focus on recognition of facial expression combining hand gestures [1]. Front cameras of computers will be used to record videos of each individual's learning process. Methodologies having been proved to be effect in previous researches, such as Hidden Markov Models (HMM) [2] and Local Binary Patterns from Three Orthogonal Planes (LBP-TOP) [3,4], will be used to process and analyse video data.

RESEARCH QUESTIONS AND/OR HYPOTHESES
The main purpose of the study is to test whether facial expressions together with gestures are effective in recognizing people's learning states. However, learning states are various in real life. In this study, we assume that there are four learning states: understandable, difficult but understandable, confused and distracted. The study will test whether these four learning states can be recognized by machine vision methodologies.

METHODS AND PROCEDURES
The study is separated into three steps.

The first step is to build a learning states database. We collect video data of learning process and categorize the videos by different learning states.

The second step is to extract the features of the videos for each category and classify them into the classifier using support vector machine (SVM) method.

The last step is the testing procedure. We use testing samples to evaluate whether the classifier we build is effective.

Variables
Two groups of volunteers are needed. Group 1 consists of 30 volunteers to finish a learning states experiment for building database; Group 2 consists of 25 people to test the performance of the recognizing system.

A 5 minutes long teaching video will be prepared. The whole instruction will include five teaching points with increasing difficulties and each teaching point takes around 1 minute.

The participants from Group 1 will be invited to watch the teaching video separately. Their learning process will be recorded by cameras. After watching the video, they are required to fill in an questionnaire. The questionnaire is consist of five questions corresponding to the five teaching points. Each question has five options about how the participants think of the teaching point, including understandable, difficult but understandable, confused, distracted and other reason, which are regarded as learning states. In addition, the option 'other reason' will not be taken into consideration as a learning state in the follow-up study.

Based on the participants' answers, their videos will be clipped into 1 minute long clips corresponding to each teaching point and categorized into groups of four learning states.

Participants of Group 2 will be invited to do the same experiment as participants of Group 1 do. Their video clips will be used to test whether the newly built learning state classifier is effective. An accuracy percentage can be used to evaluate the usability.

Sampling
Participants will be invited around the university regardless of gender and occupation. However, this study mainly concentrates on adults, participants should be at least 18 years old.
**Instrumentation**

30 laptops will be sufficient for this study. We just need to confirm that all the cameras in the laptops work and the recording function is available.

Also the testing teaching video should be prepared. If any e-learning providers are willing to provide a video for this experiment, the study will be more significant.

**Data collection**

The experiment involving participants from group 1 will create 150 video clips with different learning states for building database. It is better to keep the numbers of the clips for each learning state at least 15. If the video data is not sufficient, the data collecting experiment will be carried out again with a new teaching video or new participants.

The experiment involving participants from group 2 will create 125 video clips with different learning states for the final evaluating session.

**Data analysis**

According to the questionnaires the participants submit, the video clips will be categorized into four learning states. To analyse the data, the features of each collected video clips will be extracted by using Local Binary Patterns from Three Orthogonal Planes (LBP-TOP) method. The Hidden Markov Models (HMM) will be used to build the dynamic gesture sequence model. Then a support vector machine (SVM) [5,6] will be utilized to classify the features of a sequence of pictures extracted from video clips into the four learning states. The video clips collected from group 2 will be regarded as testing samples.

**LIMITATIONS AND DELIMITATIONS**

This study is a primary exploration of utilizing facial expression together with gestures to recognize people's learning states. It is possible that the samples of the four learning states have no recognizable difference in features, which means that analysing learning states by using facial expressions and gestures recognition is not feasible. Also it is possible that some of the learning states are recognizable while the rest are not. Otherwise, the present algorithms such as LBP-TOP, HMM, SVM may not be effective for this specific study field.

This study concentrates only on adult participants. Children and juveniles may act different from adults. Considering that adults are more stable in behaviour, we study only on adult participants to make the study less complex. In addition, the study only focuses on one-to-one teaching. Meanwhile, there are only four learning states for this study, which are simple but distinguishable. If the study is proved to be effective under the present delimitation, more learning states and participants of different age groups are worth studying in the follow-up studies.

**SIGNIFICANCE**

This study tests learning states with a combination of facial expressions and gestures while previous studies usually focus on these two fields separately.

This study can make distance education more interactive. If the study turns out to be effective, the gap between learners and course providers can be narrowed to some extent. The distance education providers are able to collect more accurate feedbacks from users, which will be an important reference to their improvement and will also help them maintain their users.

**REFERENCES**


