MAKING INTELLIGENT TUTORS EMOTIONALLY AWARE

Banda N., Robinson P.

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

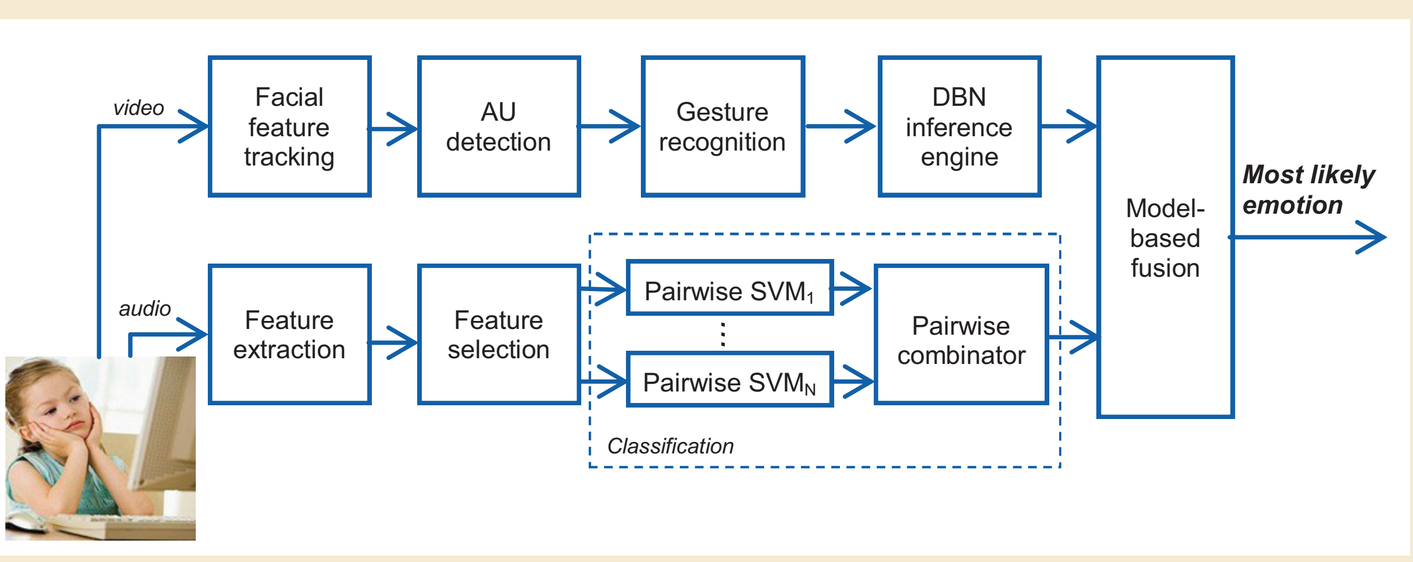
Our research is directed toward equipping intelligent tutoring systems with the ability to infer complex mental states from visual and audio cues. This will allow the tutoring system to maximize the learning potential of a student by adjusting the study material in response to the detected mental state. Inference is achieved by analysing facial expressions, head gestures and audio features.

Multi-modal inference

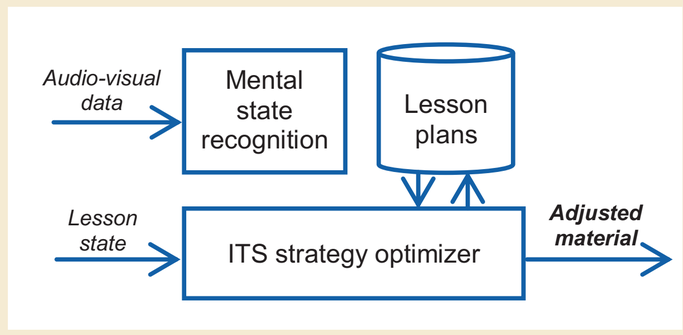
The video and audio channels are processed independently. The facial expression recognition subsystem (developed by Kaliouby [1]) is based on FACS with the action units being encoded as gestures, to allow for complex movements such as head nods and head shakes to be represented. A DBN is trained for each mental state; the ensemble is then used for classification.

The analysis of emotion in speech is achieved by extracting features (related to pitch, duration, etc.) and using a correlation-based method to select features which best represent a mental state. Pairwise support vector machines serve as classifiers with a voting scheme employed to determine the most likely mental states [2].

Model-based late fusion is employed to determine the unfolding emotion. The tutoring system will track the emotions on a turn-basis with a time threshold, and the prevalent emotion will be used to determine if a change in tutoring strategy is required.



Intelligent tutor



The mental state recognition system will interface with a tutoring system aimed at improving children's reading skills. Upon detection of states such as frustration and confusion, it will adjust the reading material for easier comprehension, and provide necessary motivation.

References

[1] R. Kaliouby and P. Robinson, "Real-time inference of complex mental states from facial expressions and head gestures," in *Real-time vision for HCI*, Springer-Verlag, 2005, pp. 181-200.

[2] T. Sobol-Shikler and P. Robinson, "Classification of complex information: inference of co-occurring affective states from their expressions in speech," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1284-1297, Jul. 2010.

RAINBOW RESEARCH GROUP

William Gates Building
15 J J Thomson Avenue
Cambridge UK CB3 0FD

http://www.cl.cam.ac.uk/emotions/