報告番号	※甲　　第　　号

主　論　文　の　要　旨

論文題目　A Study on Recognition of Students' Multiple Mental States during Discussion Using Multimodal Data

（マルチモーダルデータを用いた議論中の学生の心的状態の認識に関する研究）

氏　　名　　彭　詩朦

論　文　内　容　の　要　旨

Students will experience a complex mixture of mental states during discussion activities, which have been widely acknowledged as crucial components for revealing a student’s learning state. An ideal teacher should be sensitive enough in monitoring the mental states of students during learning, infer the learning troubles students are facing, and make real-time decisions on what kind of adaptive support to provide at what time. Many lines of research have explored the automatic measurement of students’ single mental state in pre-designed learning tasks in a computer environment. It still remains a challenge to monitor the complex mental states of multiple students in a real-world environment of the real teacher-student interactive discussion activity.

In this study, I propose using multimodal data to design an intelligent monitoring agent that can assist teachers in effectively monitoring the multiple mental states of students during discussion. To achieve this goal, an advanced multi-sensor-based data-collection system was developed and applied in a real university’s research lab to record and accumulate a massive amounts of multimodal conversational data in “in-the-wild”. An explorative work was firstly conducted to validate the predictive ability of students’ physiological cues heart rate in automatically recognizing students’ mental confidence toward appropriately giving answers in Q&A session of discussion. Then, I further explored how to effectively design multimodal analytics to take advantages of the supplement and replacement capabilities between different modalities in augmenting human teacher’s perceptual and reasoning capability in monitoring the multiple mental states of students

including concentration, confusion, frustration and boredom during discussion, in which I derived a set of proxy features from facial, heart rate, and acoustic modalities and used them to train several supervised learning classifiers with different multimodal fusion approaches single-channel-level, feature-level, and decision-level fusion to recognize students' multiple mental states in conversations. I explored effectively design multimodal analytics can augment the ability in recognizing different mental states and found that fusing heart rate and acoustic modalities yields better recognize the states of concentration (AUC = 0.842) and confusion (AUC = 0.695), while fusing three modalities yield the best performance in recognizing the states of frustration (AUC = 0.737) and boredom (AUC = 0.810). The results also explored the possibility of leveraging the advantages of the replacement capabilities between different modalities to provide human teachers with solutions for addressing the challenges with monitoring students in different real-world education environments.