ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202601_9(1).0009

A Study on the Application of Real-Time Feedback Technology in Classroom Teaching within a Multimodal Learning Environment

Xianzhi Yang
College of Education Sciences, Shenyang Normal University, Shenyang, 110000, China
yangxzhi0221@163.com

Abstract

In the context of educational informatization, multimodal learning environments integrate various types of information such as video, audio, and behavioral data to provide students with a richer learning experience. Real-time feedback technology plays a key role in this process, enabling teachers to quickly understand students' learning states and adjust teaching strategies accordingly. Based on MATLAB, this paper constructs a real-time feedback simulation system for classroom learning states, which incorporates multimodal fusion and temporal smoothing functions. Using simulated classroom data, the feasibility and effectiveness of the system are verified. Experimental results clearly demonstrate that the system can effectively monitor and evaluate students' learning states, providing technical support for the improvement of classroom teaching.

Keywords

Multimodal learning environment; Real-time feedback technology; Classroom learning status; Pattern recognition; Temporal smoothing.

1. Introduction

Multimodal learning environments create immersive learning experiences for students by integrating diverse information such as text, images and sound. Advances in information technology are driving a transformation in educational models from teacher-centred to learner-centred approaches. Within this context, real-time feedback technology has become pivotal for enhancing teaching effectiveness and classroom interaction.

This technology captures students' multimodal data—including speech, facial expressions, and movements—in real time, analysing it to assist teachers in dynamically adjusting teaching strategies. This supports students' self-monitoring and learning optimisation, effectively enhancing classroom personalisation and interactivity.

Through steps including simulated data generation, multimodal feature extraction, fusion of temporal smoothing algorithms, and learning state classification assessment, this study utilised MATLAB to construct a multimodal classroom simulation system. A real-time feedback mechanism was designed to validate its potential for implementing teaching interventions and evaluating learning states, thereby providing a technical foundation for advancing classroom intelligence. The design and implementation of effective feedback models within mobile learning environments also offered solutions to [1].

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202601_9(1).0009

2. Review of Relevant Theories and Technologies

2.1. Overview of Real-Time Feedback Systems in Multimodal Learning Environments

By integrating information from multiple sensory channels such as vision, hearing and touch, multimodal learning environments provide students with a more comprehensive learning experience. Within such settings, real-time feedback systems can rapidly process data derived from different modalities, thereby supporting teachers in adjusting instructional strategies and enabling students to engage in self-regulated learning [2].The role of multimodal reflection in enhancing classroom interaction further underscores the system's potential for improving pedagogical interactivity [3].

The construction of real-time feedback systems typically encompasses the following key stages: collecting students' multimodal data via sensor devices, pre-processing and feature extraction of raw data, integrating information from different modalities using fusion algorithms, generating personalised feedback based on the fusion results, and ultimately applying this feedback to teaching adjustments [4]. The implementation of interactive feedback technology in mathematics classrooms has also demonstrated its potential for enhancing classroom engagement [5].

2.2. Pattern Recognition-Related Technologies

Pattern recognition technology plays a pivotal role in real-time feedback systems within multimodal learning environments, primarily serving to identify students' learning states and behavioural patterns from multi-source data [6].

Feature extraction, feature selection, classifier design, and decision-making processes are typically encompassed within pattern recognition procedures. During the feature extraction phase of raw data, representative feature information is extracted. Feature selection then identifies the most effective subset of features for classification from those extracted. In the classifier design stage, appropriate machine learning algorithms are selected to construct recognition models. Finally, corresponding teaching decisions are formulated based on the classification outcomes [7].

When processing multimodal data within educational contexts, commonly employed classification algorithms such as neural networks, random forests, and support vector machines each exhibit distinct advantages. Random forests are well-suited for handling high-dimensional features, support vector machines demonstrate robust performance under small sample scenarios, whilst neural networks possess strong fitting capabilities for complex non-linear problems [8].

2.3. Time-series Smoothing Techniques

By reducing random fluctuations and noise interference, thereby enhancing data reliability and interpretability, time series smoothing techniques are primarily employed to process continuously acquired data streams. Within educational contexts, this technology can effectively improve the stability and accuracy of learning state monitoring [9].

Common temporal smoothing methods include Kalman filtering, exponential smoothing, and moving averages. Moving averages achieve smoothing by calculating the average of observed values over a past period, expressed mathematically as:

$$S_t = \frac{1}{N}\sum_{i=0}^{N-1} X_{t-i} \tag{1}$$

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202601_9(1).0009

Where X_{t-i} denotes historical observations, N represents the window size, and S_t indicates the smoothed output at time point t. This method is computationally straightforward and well-suited for real-time processing requirements [10].

Exponential smoothing, employing a weighted average approach, assigns greater weight to recent data. Its fundamental formula is:

$$S_t=\alpha X_t+(1-\alpha)S_{t-1} \tag{2}$$

Here, the smoothing coefficient α typically ranges between 0 and 1, regulating the weight distribution between new and historical data [11].

In multimodal learning environments, processing dynamically changing data such as behavioural emotional responses can be effectively managed through temporal smoothing techniques, thereby providing more accurate foundational data support for real-time feedback [12].

3. Real-time Classroom Learning State Feedback System with Multimodal Fusion and Temporal Smoothing

3.1. System Framework Design

Utilising the MATLAB simulation environment, this paper employs a modular design philosophy to construct a real-time feedback simulation system. Five core modules—simulated data generation, multimodal feature extraction, feature fusion, temporal smoothing, and state classification evaluation—collectively form the entire system.

The simulation data encompassing student learning behaviours such as facial expressions, speech, and gestures is generated by the simulation data generation module, laying the groundwork for subsequent processing. Raw data undergoes processing through the multimodal feature extraction module, from which behavioural features, visual features (such as EAR, MAR, HPR), and audio features (MFCC, spectral features) are extracted. Information from different modalities is integrated via the feature fusion module using a weighted fusion strategy, constructing a unified representation of the student's state. The fused data undergoes smoothing processing via the temporal smoothing module to eliminate interference from transient fluctuations. The state classification evaluation module ultimately achieves system performance validation and learning state recognition. Similarly, the classroom state feedback system based on facial pose recognition by Wu Lijuan et al. also provides reference for visual feature extraction in this system [13].

From raw data generation to final state assessment, the complete processing chain is constructed by sequentially connecting all system modules through data flow.

3.2. Construction of Experimental Datasets

This paper constructs a multimodal simulated dataset within the MATLAB environment to validate system performance. The dataset encompasses video, audio, and behavioural data. Through specially designed generative functions, it simulates multimodal behavioural characteristics of students across different attentional states—such as focus, distraction, and confusion—during a 40-minute simulated classroom teaching session.

The dataset construction process comprises four primary stages: data generation, preprocessing, annotation, and storage. The generated data encompasses three modalities: behavioural data (12-dimensional features), audio data (30-dimensional features), and video data (25-dimensional features). Each modality contains 480 samples accompanied by

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202601_9(1).0009

corresponding attention state labels, providing ample data support for algorithm validation, as shown in Table 1:

Table 1. Data Comparison Table

Data type	Sample size	Data source	Feature dimension
Analogue video data	480	generate_40min_video_data	25-dimensional
Analogue audio data	480	generate_40min_audio_data	30-dimensional
Simulated behavioural data	480	generate_40min_behavior_data	12-dimensional

3.3. Multimodal Fusion and Temporal Smoothing Algorithm

The system employs a weighted averaging method during multimodal fusion to integrate information derived from different modalities. Let behavioural modality data be denoted as X_b, visual modality data as X_v, and auditory modality data as X_a, with their corresponding weighting coefficients being α, β, and γ respectively, where α+β+γ=1 holds true. The fused data can be expressed as:

$$X_{fusion}=\alpha X_v+\beta X_a+\gamma X_b \tag{3}$$

To enhance data stability, an exponential smoothing algorithm is introduced for temporal processing. For the t data point in the time series, its smoothed value S_t is calculated as:

$$S_t=\lambda X_t+(1-\lambda)S_{t-1} \tag{4}$$

The smoothing coefficient λ within this formula governs the weighting ratio between the new data point and historical data points.

Through the combined application of the aforementioned algorithms, the system achieves efficient fusion and smoothing processing of multimodal learning data, providing technical support for the precise real-time monitoring of classroom learning states.

Following the construction of a simulated dataset incorporating video, audio, and behavioural data, this chapter proposes a state recognition algorithm combining weighted fusion with exponential smoothing. A real-time feedback system for classroom learning states, based on multimodal fusion and temporal smoothing, has been designed. This lays the groundwork for subsequent simulation validation, demonstrating the technical feasibility of achieving precise learning state monitoring within educational settings through this system framework.

4. Experiments and Analysis

4.1. Experimental Setup

To validate the system's operability and effectiveness, this paper integrates key toolboxes including Statistics and Machine Learning Toolbox, Signal Processing Toolbox, and Image Processing Toolbox to establish a simulation platform within the MATLAB R2020a environment. Simulated datasets were generated using custom functions, encompassing video (25-dimensional), audio (30-dimensional), and behavioural (12-dimensional) features. These datasets simulated multimodal behavioural characteristics of students under varying attentional states. Experimental evaluations of system performance were conducted through 500 repeated simulations, ensuring statistical significance of the results.

4.2. System Validation

With the support of multimodal fusion and temporal smoothing algorithms, the system can effectively identify students' learning states. Across 500 independent simulations, the demonstrated state recognition accuracy validated the feasibility of the proposed method

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202601_9(1).0009

within the simulation environment. Its optimal performance reached 55.21%, maintaining stable results between 40% and 50% throughout the simulation process. By conducting comprehensive analysis of multimodal features and implementing temporal smoothing processing, the system can relatively accurately capture changes in students' learning states, providing a robust and reliable basis for teaching interventions. Research by Sun Yongke et al. on real-time feedback models for university classrooms also supports the feasibility of this system in practical teaching [14]. As shown in Figures 1, 2, and 3:

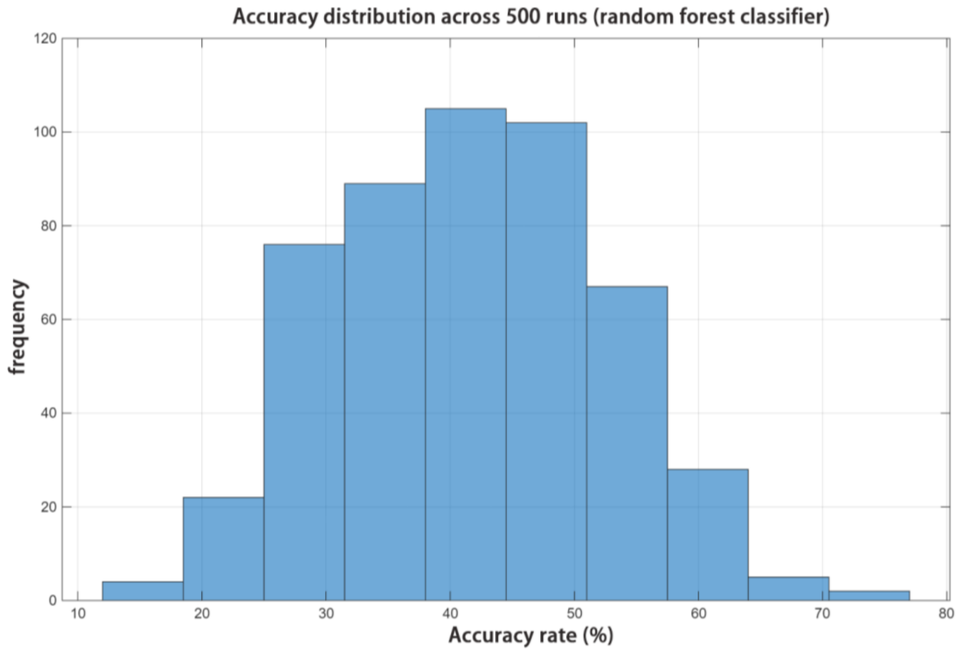

Figure 1. Accuracy distribution across 500 runs (random forest classifier)

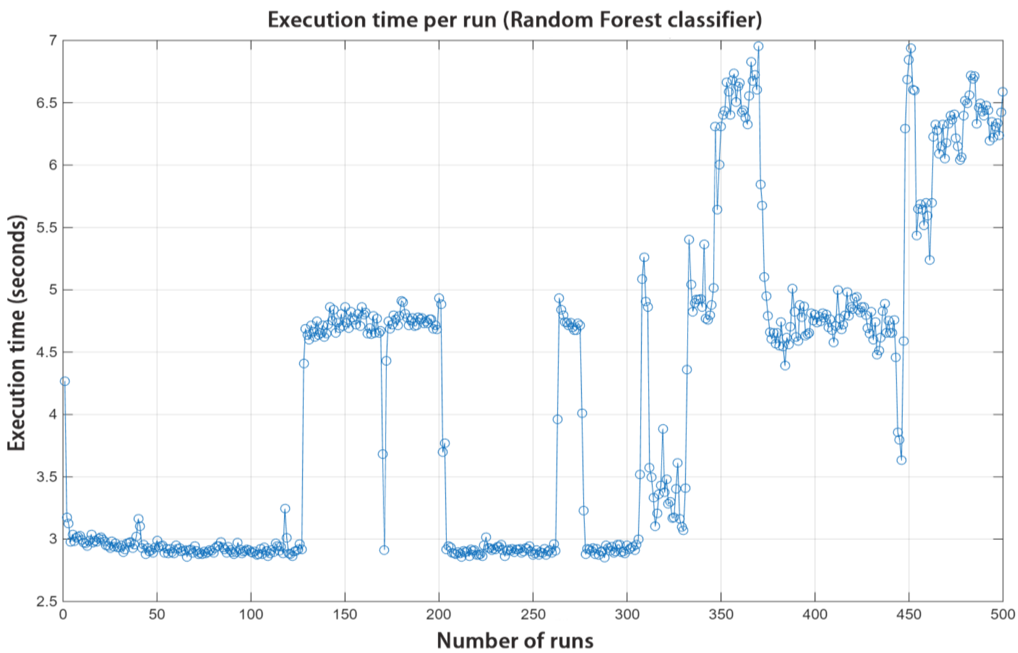

Figure 2. Execution time per run (Random Forest classifier)

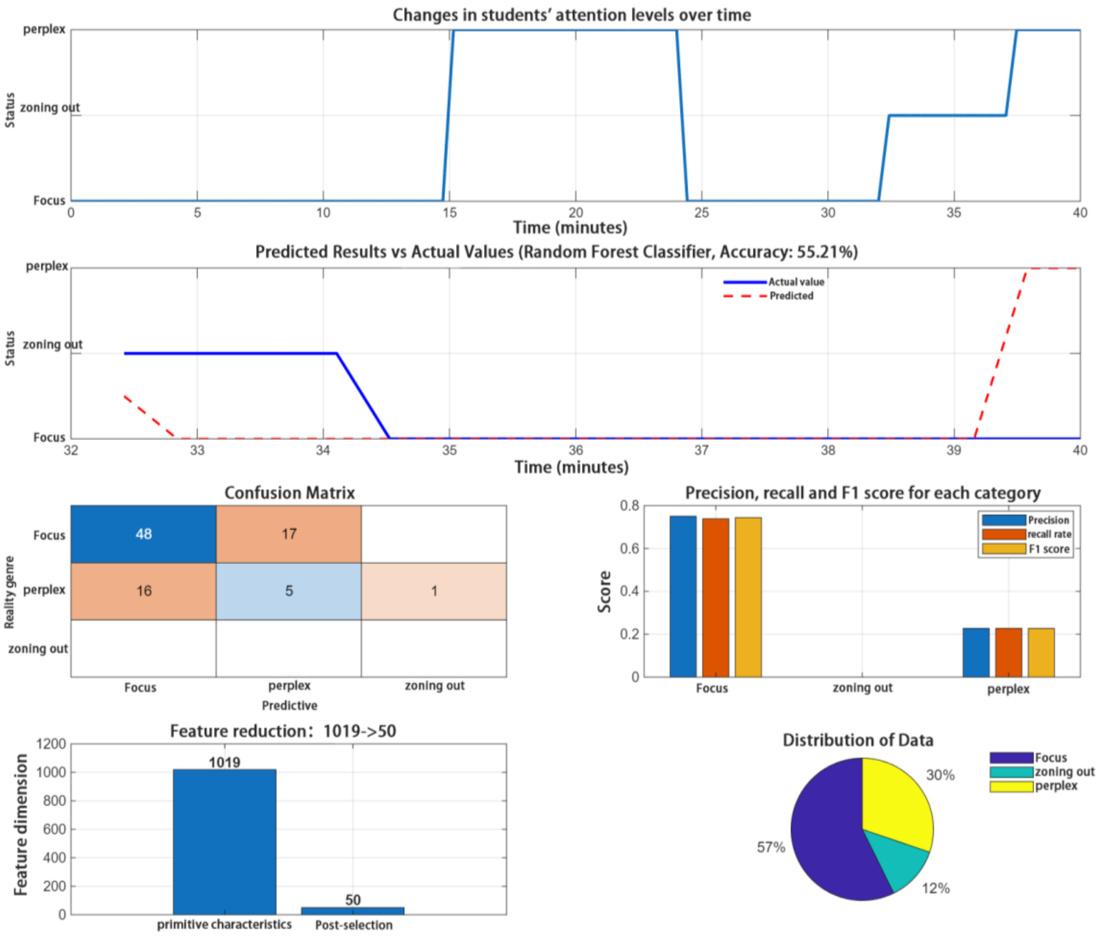

Figure 3. Performance Evaluation of Student Attention State Prediction Models

4.3. System Operational Verification

The system demonstrates excellent practical characteristics at the operational level, achieving complete reproducibility of experimental results under fixed random seed conditions. Key parameters can be flexibly adjusted via interfaces, while the modular code structure facilitates functional modification and expansion. Within educational application scenarios, the system's execution efficiency meets fundamental requirements, with an average processing time of 5.4 seconds across 500 simulation runs, highlighting its potential value.

Through systematic simulation experiments, this chapter validates the operability and efficacy of the proposed real-time feedback system. Supported by multimodal fusion and temporal smoothing algorithms, experimental results demonstrate its capacity for relatively accurate identification of student learning states. The system exhibits commendable stability and practical potential, providing experimental grounds for subsequent practical teaching applications.

5. Conclusions and Outlook

Building upon the MATLAB simulation environment, a real-time feedback system for classroom learning states has been successfully developed, integrating multimodal data with temporal smoothing characteristics. Systematic simulation validation within an artificial teaching environment demonstrates that this approach effectively consolidates multimodal information—including video, audio, and behavioural data—to achieve real-time monitoring and assessment of learning states. Across 500 independent trials, the system achieved an optimal recognition accuracy of 55.21% by combining a temporal smoothing algorithm with a random forest classifier, demonstrating excellent stability. Nevertheless, the research retains certain limitations: real-time performance evaluation in actual environments requires further validation, and discrepancies exist between the simulation environment and authentic classroom scenarios. Additionally, the handling of individual variations remains insufficient.

A comprehensive 'simulation-first' technology validation pathway is proposed herein, offering novel methodological insights for educational technology R&D. Its value lies precisely in this. For subsequent development of real-world systems, multimodal feature fusion methods and temporal processing techniques are introduced, providing crucial reference points. Looking ahead, research may deepen in three directions: exploring the transfer of simulation models to practical applications and advancing technology validation in authentic environments. Establishing state recognition models that account for individual student differences to enhance personalised learning support; optimising technical solutions while respecting educational principles; and deepening interdisciplinary integration with pedagogy. These research directions will powerfully advance the effective application of intelligent educational technologies in actual teaching, ultimately contributing to improved teaching quality and enhanced instructional efficiency [15].

References

[1] Ai, J. (2020) Design and Implementation of an Effective Feedback Teaching Model in Mobile Learning Environments: A Case Study of Senior Secondary Mathematics. Shandong Normal University, Jinan.

[2] Kong, D. (2023) The Impact of Multimodal Reflection on Classroom Interaction Competence: A Case Study of Pre-service Secondary School English Teachers. Zhejiang Normal University, Jinhua.

[3] Lin, X. (2021) The Impact of Multimodal Reflection on Classroom Interaction Competence. Shanghai International Studies University, Shanghai.

[4] Li, J., Zhao, M. (2022) Research on the Application of Classroom Instant Feedback Systems in Instructional Design: A Case Study of the University of Ottawa. Comput. Telecommun., 09: 4–9.

[5] Xu, Z. (2019) An Applied Study on Integrating Interactive Feedback Technology into Mathematics Lessons. Shanghai Normal University, Shanghai.

[6] Zhang, H. (2019) A Study on Teacher Feedback in Online Course Instruction. Theor. Res. Innov. Entrep., 19: 41–42.

[7] Chen, Z. (2019) Real-time Feedback Simulation for Learning Monitoring in Multimedia Sensor Networks. Sci. Bull., 09: 110–118.

[8] Yang, Y. (2019) A Discussion on Teaching Feedback Models in the Online Environment. Mod. Inf. Technol., 16: 111–113.

[9] Yue, H. (2022) An Application Study of Real-time Interactive Feedback Systems in Classroom Teaching Evaluation at Junior Secondary Level. Yan'an University, Yan'an.

[10] Cui, X. (2021) Research on the Current Status of Classroom Teaching Instant Feedback Systems. Comput. Knowl. Technol., 30: 188–190.

[11] Sun, Y., Hu, K., Zhao, Y. et al. (2019) Research on Real-time Classroom Feedback Teaching Models. Comput. Knowl. Technol., 33: 83–84.

[12] Zhang, J. (2020) Research and Development of Real-time Teaching Feedback Information Systems. Comput. Knowl. Technol., 04: 134–135.

[13] Wu, L., Ren, H., Guan, G. et al. (2022) Design and Implementation of a Classroom Learning Status Feedback System Based on Facial Pose Recognition. J. Shenyang Norm. Univ. (Nat. Sci.), 02: 127–132.

[14] Sun, Y., Xiong, F., Cao, Y. (2021) Application and Analysis of Real-Time Feedback in University Classroom Teaching. Comput. Knowl. Technol., 03: 165–167.

[15] Wang, J., Shi, X. (2019) Exploring the Application of Real-Time Feedback Mechanisms in Higher Vocational Education Classrooms. J. Yellow River Water Conserv. Vocat. Tech. Coll., 03: 59–61.