Proceedings of the 7th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/2027.38422

Immersive Experimental Teaching Interaction Design Based on Virtual Reality Environment

Zhikai Wang

School of Computer Science, University of Birmingham, Birmingham, UK
wellington589125@gmail.com

Abstract. Immersive virtual reality allows broadening access to laboratory experiments by minimizing the need for tangible devices and repeated procedural learning, but pre-defined instructional cues give little help to learners in case of diverging from complex multi-stage experimental procedures. To solve this problem, an adaptive teaching framework for experimental procedures in acid-base titration was developed using Unity 2022.3 LTS, Meta Quest 3, behavioral event logging, Hidden Markov Model (HMM) and Viterbi decoding. The system detected five states of operation and provided feedback through visual, haptic, and textual feedback in case of illegal transitions, missed steps, or repeated invalid moves. In a controlled experiment conducted on 72 undergraduate students, a model-guided VR condition and a basic-VR condition were compared. The HMM gave an accuracy of 90.8%, Macro-F1 score of 0.908 and mean inference delay of 37.6 ± 6.2 ms. The group with the model guidance made fewer procedural mistakes (1.78 ± 1.35 vs 4.39 ± 1.90), completed more steps (94.82% vs 89.77%), gained more knowledge (17.47 ± 6.83 vs 11.39 ± 5.15) and had lower NASA-TLX scores (37.13 ± 10.01 vs 48.26 ± 10.40). This study proves that adaptive interaction with state awareness allows connecting real-time behavior recognition with improvements in procedural accuracy, cognitive efficiency and learning in laboratories.

Keywords: Virtual Reality, Immersive Experimental Teaching, Hidden Markov Model, Adaptive Interaction, Behavioral Analysis

1. Introduction

Laboratory courses in universities involve precision in procedures, repetition, and correction; however, instrument availability, safety measures, and time constraints imposed by the teacher may limit these requirements. Virtual reality provides immersive 3D environments for active rehearsals and procedural learning in STEM [1]. A meta-analysis has further shown a significant positive impact of virtual reality on STEM learning, although the results depend on the instructional methods and tasks used [2]. Modern virtual laboratories tend to focus on the realism of scenes and immersion as well as post-experiment performance, whereas ongoing detection of user activities is not adequately addressed. Consequently, the capability of the system to differentiate between correct procedure, omission, repetition, and disordering of operations in the process of experiments is limited. Thus, an interaction framework is proposed based on acid-base titration and involving Unity behavior tracking, Hidden Markov Model state recognition, Viterbi decoding, and adaptive

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2027.38422

multimodal feedback. Further, the performance of the state recognition and procedural accuracy are compared along with the cognitive load, task efficiency, and actual learning gains.

2. Literature review

2.1. Virtual laboratory teaching

The modern approaches to the use of virtual laboratory environments treat immersion as a stage preparing the way for procedural learning of science. As to the study of chemistry, there was an investigation demonstrating that the virtual training was efficient in terms of perception of laboratory procedures and proper completion of the assigned task, although the importance of practical lab exercises was perceived by the students anyway [3]. Thus, VR can be regarded as a preparatory stage of procedural learning rather than a complete substitute for hands-on exercises. In the meantime, there was a study of undergraduate chemistry which showed that the pre-training in the context of VR could contribute to laboratory preparedness if digital exercises matched the real-life procedures [4].

2.2. Immersive interaction guidance

The mechanism of immersive guidance affects the distribution of learner's attention during the manipulation of virtual objects. It was found that text-based cues and summary prompts could enhance learning outcomes and help to build mental models in the VR environment without the increase of the cognitive load [5]. In virtual chemistry labs, task-specific hints such as arrows and sound can decrease the cognitive effort required for searching and help to perform better in experiments; however, too much help can cause cognitive overload [6]. This creates a conflict in designing the mechanism of immersive guidance. Persistent guidance can make tasks easier but will undermine the diagnostic value of learner's behavior; meanwhile, delayed guidance might let the mistakes happen. The state-sensitive guidance can solve this conflict.

2.3. Behavioral sequence modeling

Modeling behavioral sequence opens up VR assessment from end results to process-oriented data. Literature reveals the importance of using learning analytics as an essential component alongside immersive technologies since interaction records can provide useful information beyond what traditional assessment metrics can provide [7]. From personalization studies, it becomes clear that any personalized VR must have information on the learner's state in order to tailor instruction rather than provide the same instruction to all learners [8]. Feedback-oriented VR research done through hands-on activities shows how reflective feedback can improve engagement, cognitive skills, and task performance [9]. All these threads point to the computational aspect, whereby time-stamped behavior must first be transformed to a meaningful state for personalized feedback. For procedural laboratory activities, probabilistic sequence modeling is appropriate as transitions occur under temporal dependency constraints.

DOI: 10.54254/2755-2721/2027.38422

3. Experimental methods

3.1. VR environment construction

The virtual titration is coded in Unity 2022.3 LTS through the integration of OpenXR and XR Interaction Toolkit and deployed onto Meta Quest 3. In the virtual scene, the objects include a burette, a pipette, a conical flask, a sodium hydroxide reservoir, a hydrochloric acid sample, a phenolphthalein solution, a clamp, and the measurement interface. C# scripts are used for object grabbing and collision detection, liquid pouring, updating the volume of the burette, changing colors in the endpoint, and procedural constraints. The target states for participants are pipetting, filling, titration, endpoint detection, and reading. All interaction events are recorded into CSV files with data about the participant ID, timestamp, object ID, action code, state code, position of the controller, rotation of the controller, and system feedback. The position and rotation of controllers are sampled at 20 Hz whereas the interaction events will be recorded instantly upon grabbing, releasing, pouring, valve operation, and reading operations. The raw logs will be preprocessed with the removal of duplicated events under 100 milliseconds, synchronization of timestamps, and mappings between low-level actions and observation symbols. The learner's sessions of the tasks will be cut into one single sequence starting from task initiation until its end, which will serve as the input to the HMM training and Viterbi decoding.

3.2. HMM state recognition

In the HMM, there are five hidden states which correspond to the titration phases that are encoded and eight observable symbols which are based on the events grab, move, rotate, pour, release, valve, read, and endpoint check. The HMM decoding is relevant in recognizing behavior in virtual training because sequential observations can be probabilistically associated with latent states of interaction. Thirty demonstrations of experts are gathered before the student's experimentation and are used to estimate the probabilities using maximum likelihood counts and then Baum-Welch algorithm [10]. For an observation sequence O and state sequence S, the joint probability is calculated as shown in Equation (1).

$$P(O,S|\lambda)=\pi_{s_1} b_{s_1}(o_1)\prod_{t=2}^{T} a_{s_{t-1}s_t} b_{s_t}(o_t) \tag{1}$$

Where $O=\{o_1,o_2,\ldots,o_T\}$ denotes the observed VR interaction sequence, $S=\{s_1,s_2,\ldots,s_T\}$ denotes the hidden operational-state sequence, $\lambda=(A,B,\pi)$ represents the HMM parameters, a_{ij} is the state-transition probability, and $b_j(o_t)$ is the emission probability of observation o_t in state j.

The optimal state path is recursively decoded using the Viterbi function, as shown in Equation (2).

$$\delta_t(j)=\max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}]b_j(o_t) \tag{2}$$

Where $\delta_t(j)$ represents the maximum probability of all state paths ending at state j at time t, a_{ij} denotes the transition probability from state i to state j, $b_j(o_t)$ denotes the observation probability, and N is the number of hidden states.

DOI: 10.54254/2755-2721/2027.38422

3.3. Controlled experiment design

G*Power is employed with a medium standardized effect, alpha 0.05, and power 0.80 to determine the recruitment goal to be 72 undergraduates having taken introductory chemistry class but not yet done the VR task. Participants are randomly assigned through computerized block randomization into a model-guided condition or basic-VR condition, 36 subjects each. Both conditions are exposed to identical five minutes' headset tutorial, pretest consisting of ten items, virtual apparatus, titration goal, and task period up to a certain minute. Model-guided only gets HMM-based object highlighting, controller vibration, and brief instructional text; adaptive feedback is limited to those procedural errors identified since learner state-aware assistance can be more effective than static feedback in VR practice [11]. An independent titration is done within 15 minutes before an equivalent ten items posttest and NASA-TLX. Ground truth data collection for Accuracy and Macro-F1 is done through independent annotation of screen recording and event log, and system timestamp for inference latency. Equivalence test on pretests is done. Effect sizes and 95% CI are included in posttest performance analysis by ANCOVA using pretests as covariates.

4. Results

4.1. State recognition performance

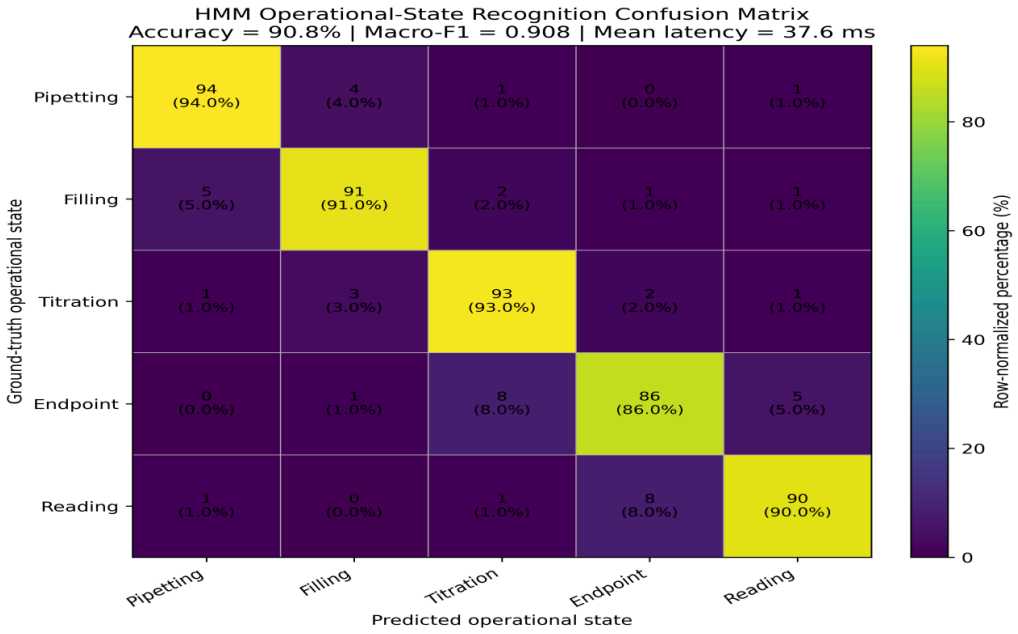


Figure 1. HMM operational-state recognition confusion matrix

HMM reached 90.8% classification accuracy and 0.908 Macro-F1 score for 500 manually annotated operational-state samples. As illustrated in Figure 1, the row-normalized confusion matrix depicts the recognition distribution for pipetting, filling, titration, endpoint decision, and reading. Recall scores for different states amounted to 94.0%, 91.0%, 93.0%, 86.0%, and 90.0%, respectively. Endpoint decision demonstrated the highest level of classification ambiguity, where eight samples were recognized as titration, and five samples as reading, suggesting that errors occurred at the point of transition from titration drop by drop to endpoint decision and measurement. Average inference latency equaled 37.6 ± 6.2 ms and stayed under the 100 ms threshold set up for

DOI: 10.54254/2755-2721/2027.38422

real-time feedback. Precision of recognition varied from 88.6% to 93.1% for all five states, providing an equal contribution to state classification. The procedure guided by the model generated 1.78 ± 1.35 procedural mistakes compared with 4.39 ± 1.90 in the basic-VR group ($p < 0.001$, Cohen's $d = -1.58$) and increased the number of steps completed from 89.77% to 94.82% ($p < 0.001$, $d = 0.98$). Thus, reliable recognition provided opportunities for anomaly detection.

4.2. Learning effectiveness

As shown in Table 1, baseline knowledge was comparable between conditions: the model-guided group scored 68.93 ± 4.92 and the basic-VR group 67.79 ± 6.49 on the pretest, with $p = 0.402$. Posttest means reached 86.41 ± 5.30 and 79.18 ± 5.62, respectively. ANCOVA controlling for pretest performance showed a significant group effect, $F(1,69) = 32.80$, $p < 0.001$, with an adjusted mean difference of 6.77 points and a 95% confidence interval from 4.42 to 9.13. Learning gain was 17.47 ± 6.83 in the model-guided condition and 11.39 ± 5.15 in basic VR, with $p < 0.001$ and $d = 1.01$. Task completion time decreased from 553.34 ± 54.56 s to 465.98 ± 75.86 s, with $p < 0.001$ and $d = -1.32$. NASA-TLX also decreased from 48.26 ± 10.40 to 37.13 ± 10.01, with $p < 0.001$ and $d = -1.09$. The joint pattern associates adaptive feedback with faster execution, lower workload, and stronger post-task learning. The unchanged baseline and covariate-adjusted posttest effect reduce the likelihood that the observed learning difference arose from initial chemistry knowledge imbalance.

Table 1. Learning and task performance under two VR conditions

Measure	Model-guided VR n = 36	Basic VR n = 36	Statistical result
Pretest score	68.93 ± 4.92	67.79 ± 6.49	$p = 0.402$, $d = 0.20$
Posttest score	86.41 ± 5.30	79.18 ± 5.62	ANCOVA $F(1,69) = 32.80$, $p < 0.001$; adjusted difference = 6.77
Learning gain	17.47 ± 6.83	11.39 ± 5.15	$p < 0.001$, $d = 1.01$
Task time, s	465.98 ± 75.86	553.34 ± 54.56	$p < 0.001$, $d = -1.32$
NASA-TLX	37.13 ± 10.01	48.26 ± 10.40	$p < 0.001$, $d = -1.09$

5. Discussion

It is evident that the recognition pattern shows that procedural states characterized by unique object-action relations are simpler to understand than states differing by perceptual continuity transitions. Since endpoint judgment includes events common to late titration and reading afterwards, the entirely discrete HMM reduces chemically relevant visual information in its observation set. The introduction of continuous information like burette flow rate, cumulative dispensed liquid volume, and indicator color characteristics could provide a distinction in this case. The decrease in effort could be understood in terms of selective intervention since feedback comes after deviation identification, preventing unnecessary visual scanning while ensuring independent manipulation on a correct path. This solution also explains why the faster task completion occurs together with learning gain increase as time is spent on recovery from incorrect actions but not on crucial experiment parts. There is one more challenge related to feedback reliance. Instantaneous error correction could lead to better conformity to procedure but reduce chances of error self-diagnosis if the trigger points are too sensitive. Confidence threshold or fading feedback policy could be used to

DOI: 10.54254/2755-2721/2027.38422

give the system ability to intervene only on highly probable errors and eventually give control to the learner. Cross-task validation is also required since transition matrices obtained on titration demonstrate highly structured laboratory activity.

6. Conclusion

A VR experimental teaching environment incorporating behavior tracking, HMM state detection, Viterbi decoding, and adaptive multimodal feedback was developed to teach acid-base titration. The system maps low-level control events to procedural states and engages in tutoring if there is a discrepancy between the detected behavior and expected procedure. Controlled experimentation establishes connections between computational measures, such as recognition performance and reaction time, and procedural mistakes, cognitive effort, task efficiency, and learning outcomes. In this way, the proposed architecture offers a replicable methodology for incorporating sequential intelligence into immersive laboratory education. Existing limitations include a single chemical experiment, a small sample size, short-term performance assessment, and discrete observation codes. Future research should examine long-term retention, more lab experiments, multimodal sensor capabilities, and adaptive confidence levels.

References

[1] Conrad, M., Kablitz, D., & Schumann, S. (2024). Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings. *Computers & Education: X Reality, 4*, 100053.
[2] Cromley, J. G., Chen, R., & Lawrence, L. (2023). Meta-analysis of STEM learning using virtual reality: Benefits across the board. *Journal of Science Education and Technology, 32*(3), 355–364.
[3] Viitaharju, P., et al. (2023). Student experiences from virtual reality-based chemistry laboratory exercises. *Education for Chemical Engineers, 44*, 191–199.
[4] Mojsoska, B., et al. (2024). Virtual reality in an eco-niche undergraduate organic chemistry laboratory course: New practice in chemistry lab teaching. *Journal of Chemical Education, 101*(11), 4686–4693.
[5] Li, W., et al. (2023). Effect of summarizing scaffolding and textual cues on learning performance, mental model, and cognitive load in a virtual reality environment: An experimental study. *Computers & Education, 200*, 104793.
[6] Ali, N., Ullah, S., & Raees, M. (2023). The effect of task specific aids on students' performance and minimization of cognitive load in a virtual reality chemistry laboratory. *Computer Animation and Virtual Worlds, 34*(6), e2194.
[7] Sakr, A., & Abdullah, T. (2024). Virtual, augmented reality and learning analytics impact on learners, and educators: A systematic review. *Education and Information Technologies, 29*(15), 19913–19962.
[8] Marougkas, A., et al. (2024). How personalized and effective is immersive virtual reality in education? A systematic literature review for the last decade. *Multimedia Tools and Applications, 83*(6), 18185–18233.
[9] Huang, Y.-M., et al. (2024). Empowering virtual reality with feedback and reflection in hands-on learning: Effect of learning engagement and higher-order thinking. *Journal of Computer Assisted Learning, 40*(4), 1413–1427.
[10] Amat, A. Z., et al. (2024). Prediction models of collaborative behaviors in dyadic interactions: An application for inclusive teamwork training in virtual environments. *Signals, 5*(2), 382–401.
[11] Chen, C.-Z., Wang, W.-S., & Huang, Y.-M. (2026). Designing adaptive feedback for hands-on learning in VR: Feedback intervention theory. *Education and Information Technologies, 31*(11), 4115–4139.