IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 February 2024, accepted 25 February 2024, date of publication 29 February 2024, date of current version 7 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3371516

RESEARCH ARTICLE

Exploring Cybersickness Experiences via a Markov Chain Model

SHAMUS P. SMITH, (Senior Member, IEEE)
Institute for Integrated and Intelligent Systems, Griffith University, Brisbane, QLD 4111, Australia
e-mail: shamus.smith@griffith.edu.au

ABSTRACT Users of virtual reality (VR) technology, especially head-mounted displays (HMDs), often experience cybersickness, similar to motion sickness, with feelings of nausea, dizziness and sweatiness. Cybersickness typically increases with duration of wearing a HMD and is commonly evaluated in user studies with physiological measures, in-situ verbal reports and post session questionnaires. However, in addition to being time-consuming, user studies only provide insight into the specific configuration of the VR experience under study and can be limited to participant numbers, duration of VR exposure and the impact of cybersickness on VR experience dropout and completion rates. This paper presents a formal approach to modelling cybersickness. A Markov chain is used to define a general cybersickness model where probabilities represent changes in a user's state of cybersickness. The Markov chain can be populated with historical user study data and interrogated to gain further awareness of the VR experience under evaluation. The approach is exemplified with a custom Markov chain model generated from a public VR experience dataset. The resulting model is shown to be representative of the ground truth user experience from the source material. Examples are presented to demonstrate how the model can be explored to gain insight for (i) scaled up parameters, such as exposure duration and participant numbers, and (ii) acceptance thresholds for minimum/maximum cybersickness. Limitations on the generation of the model and its utility across different user populations and environment types are considered and discussed in the context of future work.

INDEX TERMS Virtual reality, cybersickness, head-mounted display, Markov chain, prediction.

I. INTRODUCTION

There is increasing use of virtual reality (VR) technology. Much of this has been facilitated by the accessibility to high quality and relatively low cost head-mounted displays (HMDs). Also VR technologies support good use cases across entertainment, sport, simulation, education, defence, rehabilitation and psychological therapy [22], [25], [27]. However, users of VR technology often experience cybersickness [15], [21], similar to motion sickness, with reported symptoms of nausea, oculomotor discomfort and disorientation [10]. Poor virtual reality experiences involving cybersickness can deter users from the future use of such technologies [5] and this impedes commercial expansion and wider adoption [7], [11], [19], [24].

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

Research on cybersickness typically involves user studies. However, specific user studies are limited to the number of participants evaluated, the demographics of the participants, the specific virtual environment, and the associated hardware being considered. Although it would be desirable to consider large sample sizes in studies with factorial designs [27], large participant-based studies are time consuming and expensive [26] and proportionally more likely to be impacted by dropout rates [24] than smaller studies.

Also, the measurement of cybersickness is difficult as it is either intrusive, for example attaching physiological biometric devices [7], [18] or use of concurrent verbal protocols [19], or relies on post-session recollection, for example the use of questionnaires, i.e. the Simulation Sickness Questionnaire (SSQ) [13]. This constitutes an additional barrier to running large scale cybersickness evaluations.

An alternative approach is the use of predictive modelling when considering the detection and prediction of

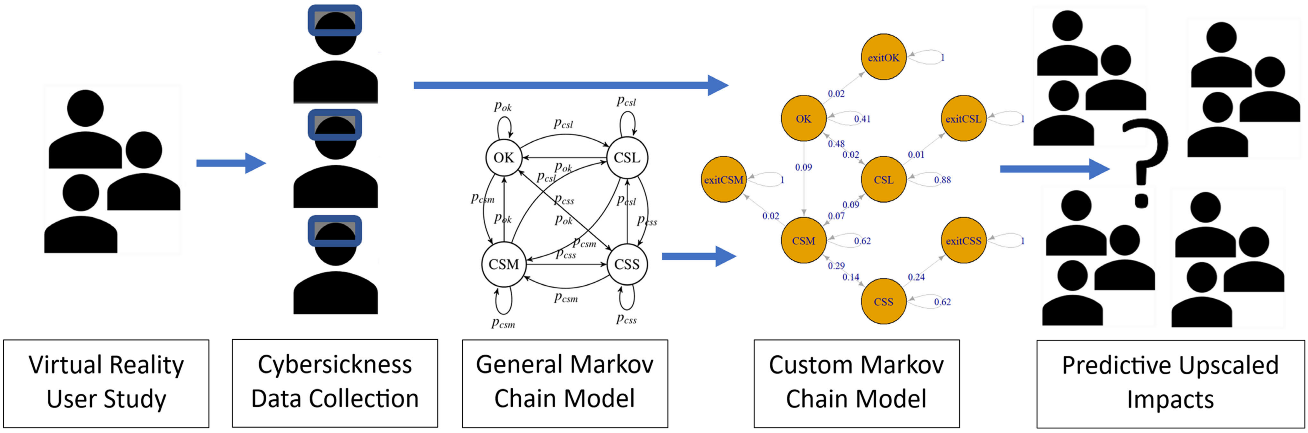

FIGURE 1. **Overview of approach using a cybersickness dataset to populate a custom Markov chain model to support predictions of upscaled impacts.**

cybersickness experiences. Recent work has considered machine learning approaches to cybersickness prediction [10], [11], [20] and there is ongoing interest in the use of formal modelling approaches to explore human experiences, for example to capture erroneous human task behaviour [1].

This paper presents a new approach where user study data on participants' experiences of cybersickness is used to generate a custom Markov chain model. Markov chain models are state-based representations that describe a succession of probable events, with predictions or probabilities for the next state. They can be designed to model real-world processes, and in this case used to model increasing and decreasing states of cybersickness. As a modelling tool, they formalize the relations between the modelled states and the probability of transitions between states. As cybersickness has a cumulative impact, based on exposure duration [26], our work aims to explore how this cumulative impact can be modelled and used to predict future cybersickness states.

Once a Markov chain model has been defined it can be interrogated with a variety of parameters in order to predict the impact of upscaling participant sizes and duration of exposure to a VR experience. An overview of the proposed approach is shown in Fig. 1. There are four contributions to the research described here:

- A new formal approach to modelling cybersickness.
- Development of a general Markov chain representing cybersickness experiences.
- Presentation of a case study using a custom cybersickness Markov chain.
- Examples of insight that can be gained from exploring the custom Markov chain model, including predicting (i) expected participant dropout rates, (ii) limits to exposure duration and (iii) VR experiences across large numbers of participants.

This paper is organised as follows: In Section II, related work on cybersickness prediction and the use of formal models for human experiences are presented. The processes of developing general and customised Markov chain models for cybersickness are defined in Sections III and IV respectively. Section V presents examples of how the customised Markov chain model can be explored with upscaled parameters. The limitations of the approach are considered in Section VI and our conclusions presented in Section VII.

II. RELATED WORK

There is increasing ongoing work in the area of cybersickness with recent literature reviews covering use of the Simulator Sickness Questionnaire (SSQ) [24], individual susceptibility [27], causes and measurements [4] and the adverse effects relating to virtual reality [25]. The following section will focus on the related work for the two main contributions of the research presented here, namely the prediction of cybersickness and the use of formal models to capture human experiences.

A. PREDICTION OF CYBERSICKNESS

Hadadi et al. [10] proposed a machine learning approach to cybersickness prediction based on physiological and subjective data. Their aim was to explore the predictive nature of different physiological measures collected with a wrist sensor. They considered a range of classifier algorithms and the assessment of the classification performance. Topological data analysis [3] was used as the feature extractor to classify participants' multivariate physiological time series during a virtual-navigation experiment (n = 53). Pre and post session SSQs were used to determine cybersickness where participants whose SSQ score was equal to or greater than 20 were assumed to suffer from cybersickness and labeled as "sick", while others were labeled as "non-sick". In comparison to [10], our model is built with in-situ, minute-by-minute, verbal reports of nausea rather than pre-post surveys noted here and has a finer granularity of reporting any cybersickness experiences.

Similarly, Islam et al. [11] observe that the SSQ is not suitable for the automatic detection of cybersickness during

TABLE 1. Summary of cybersickness prediction methods. (FMSS = Fast Motion Sickness Scale [14], SSQ = Simulator Sickness Questionnaire [13].)

Reference	Prediction method	Cybersickness ground truth	Data analysis	Predictive results	Notes
Hadadi et al. [10]	Machine learning with support vector machine classifier	Self-reported survey (SSQ)	Analysis of own laboratory-based user study	71% accuracy in classification performance	Based on physiological data collection
Islam et al. [11]	Machine learning with deep neural network	Self-reported survey (FMSS)	Analysis of own laboratory-based user study	87.38% accuracy for predicting cybersickness severity two minutes into the future	Based on physiological data collection
Padmanaban et al. [20]	Machine learning with bagged decision trees classifier	Self-reported survey (SSQ)	Analysis of own laboratory-based user study	Test set prediction errors range from 4.3% to 57.4%	Head motion constrained to forward-facing viewport
Wang et al. [28]	Deep learning model with a deep long short term memory network	Self-reported survey (SSQ)	Analysis of own laboratory-based user study	Pearson correlations across nausea (0.75), oculomotor (0.77) and disorientation (0.84)	Customized trained networks for each individual user
Jasper et al. [12]	Hierarchical multiple regression models	Self-reported survey (SSQ)	Analysis of own laboratory-based user study	Spearman's correlations with motion sickness history ($p < .001$) and perceived workload ($p < .001$) as significant predictors	Based on individual user characteristics and task characteristics. Use of a custom motion sickness history metric
Our approach	Markov chain models	Self-reported survey (FMSS)	Reuse of existing data set [11]	91.30% accuracy to ground truth source data	Suitable for individual or groups of users

immersions due to the pre-post nature of the data collection. Islam et al. [11] describe the development of a simplified convolutional long short-term memory classifier that can predict cybersickness severity two minutes into the future using the previous two minutes of physiological measurements. Their approach has 97.44% accuracy for detecting current cybersickness severity and 87.38% accuracy for predicting future cybersickness severity. Our work is based on their user study dataset but differs in that our predictions are much broader in scope, in terms of exploring the cybersickness impact space, beyond real-time and two-minute predictions. These approaches are complimentary where Islam et al. [11] can monitor users currently in virtual environments, i.e. to support stopping the experience as needed, and our work provides insight into the impact of upscaled use of an environment, i.e. via increased participant numbers or longer durations of VR exposure.

Wang et al. [28] describe a two phase approach where a first phase is to train a machine learning model based on a deep long short term memory (LSTM) network for each user. The model is trained to identify features when the user does not feel any VR sickness. A second phase uses this model to detect and alert abnormal signals which are obtained when the user is in another physiological state and feels VR sickness. The network prediction was validated with SSQ scores. The use of a pre-generated network is both an advantage, with the potential to be used for real-time cybersickness monitoring, and a disadvantage, with the need for individual networks for users, of the method. In contrast, our approach considers group and large scale predictions, rather that a focus on individual user's experiences. Thus, cybersickness detection and real-time analysis are outside the scope of the current work.

In addition to enabling the comparison of virtual environments for cybersickness [26], VR image data can be used for cybersickness detection. Padmanaban et al. [20] built a dataset of stereoscopic 3D videos and their corresponding sickness ratings via a user study ($n = 12$). A machine learning algorithm was then trained on features (quantifying speed, direction, and depth as functions of time) from each video, learning the contributions of these various features to the sickness ratings. They noted that the predictor generally outperforms a naïve estimate, but is ultimately limited by the size of the dataset. As with [10], cybersickness reports were gathered using the SSQ after participants watched each video. Similar to our work here, this approach explores the VR space without needing new user studies. However, it is limited to the visual aspects of the VR experiences and, as noted by [20], the choice to constrain head motion, i.e. participants' views were fixed in place watching the videos, and, thus, was not a truly interactive VR experience.

An alternative approach is to look at participant and task characteristics as cues to predicting susceptibility to cybersickness [12], [27]. Jasper et al. [12] explored participant demographics, the task to be completed, the virtual environment design and the VR hardware used as the four main factors contributing to the likelihood of cybersickness. Based on a user study ($n = 150$), they constructed hierarchical multiple regression models to examine individual characteristics, such as motion sickness history, previous VR use, gender, age and personality, and task characteristics, particularly workload.

Jasper et al. [12] concluded that who the user is and what they are doing is critical to user-based prediction of cybersickness. However, individual differences is out of the scope of the work presented here. Our aim is to define a general model of cybersickness that can be populated with specific case studies. However, further customisation of any model would benefit the inclusion of user characteristics and the context of any tasks. Thus [12] is complimentary to our future plans to extend the formal modelling presented here.

Table 1 provides a summary of approaches to cybersickness prediction, comparing key aspects of each approach. Our proposed method is included for comparison.

B. FORMAL METHODS FOR HUMAN EXPERIENCES

Formal methods are rigorous, mathematically-based approaches that have demonstrated their ability to verify systems across critical properties such as reliability, safety, security, more generally, dependability and performance [9]. A full overview of formal methods for system analysis is outside the scope of the current work (for recent systematic reviews see [2] and [17]) but our interest is in how formal methods, specifically the use of modelling, can be used to gain insight from systems involving humans.

Human reliability analysis [16] is often combined with formal methods, via probabilistic and statistical modelling, so that the relative likelihood of different outcomes can be determined [1]. This is useful to determine whether predictive behaviours in a system are normal or erroneous behaviours. Also, when no ground truth is available, probabilistic modelling can be used to forecast future states.

Fudolig and Howard [6] use a SIR compartmental model with compartments on susceptible (S), infected (I) and removed (R) to formalise the characteristics of an emerging disease, in this case COVID-19. The model allows simulations to be run across different levels of population vaccination until the system reaches equilibrium. Similar to our work, the aim is to gain predictive insight into the future based on varying initial parameters that drive the model. SIR models, with their simple yet powerful structure, were an initial candidate considered for our work on modelling cybersickness. However, the use of strict compartments in the model was too restrictive and a more general approach was required.

Sánchez et al. [23] consider how human behaviour modelling for welfare technology can be used to recognise an individual's behaviour patterns. The aim is to identify abnormal behaviour so that this can be considered when a safe environment, for example a smart house, is constructed for that person. A Hidden Markov Model (HMM) was used for predicting the behaviour of a person. HMMs are a subclass of Bayesian networks known as dynamic Bayesian networks [8]. The model was found to have accuracy of 72% when trained with an open source real-world dataset. Similar to our work, a public dataset was used to construct the model. However, HMMs require training and function with hidden states and observable states. For our work, a more simple model, with no hidden states, was the starting point for characterising cybersickness impacts. The aim was to build a model that refined itself via probability propagation in a 'chain' of iterative steps. This approach is detailed in the next section.

III. MODELLING CYBERSICKNESS AS A MARKOV CHAIN

A. MARKOV CHAIN MODELS

A Markov chain is a stochastic model which defines a sequence of possible events where the probability of an event is only determined by the current state. For a *finite state* Markov chain there is a discrete, i.e., countable, set of possible states. The Markov chain is *memoryless*, in that the transition to a next state is solely determined by the current state and probabilities associated with any transitions. For example, the Markov model in Fig. 2 has two states, S1 and S2 where each state has a probability of transitioning to the other state (p_1 and p_3) and a probability of staying in the same state (p_2 and p_4). The sum of the probabilities for any state will equal 1.

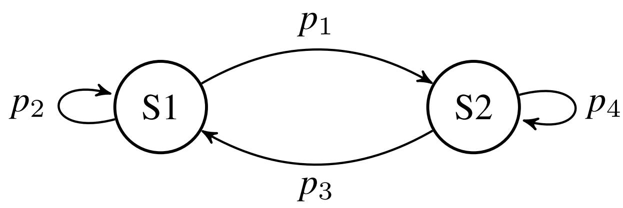

FIGURE 2. Example Markov chain.

More formally, a Markov chain can be defined as:

$$\mathbb{P}(X_{n+1} = x \mid X_n = x_n) \tag{1}$$

The probability of the next state $(n + 1)$ of the random variable X is a specific value x and this depends on the current state of the random variable $X_n = x_n$. This process can be represented as a series of matrix operations. An example *transition matrix*, T, with random probabilities but where the sum of all state probabilities equal 1, would look like the following:

$$T = \begin{bmatrix} 0.30 & 0.70 \\ 0.40 & 0.60 \end{bmatrix} \tag{2}$$

If (2) was the transition matrix for Fig. 2, $T_{1,1} = 0.30$ is the probability of staying in S1, i.e. probability p_2, and $T_{2,1} = 0.70$ is the probability of transitioning to S2, i.e. probability p_1.

A Markov chain can be seeded with an initial configuration vector. Continuing with the current example, the first iteration of the chain could start with 50% in each of S1 and S2 and the transition matrix is then multiplied with the initial vector:

$$X_{n=1} = \begin{bmatrix} 0.30 & 0.40 \\ 0.70 & 0.60 \end{bmatrix} \begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix} = \begin{bmatrix} 0.35 \\ 0.65 \end{bmatrix} \tag{3}$$

Thus, after one iteration, the distribution across the states would be 0.35 in S1 and 0.65 in S2. To calculate $X_{n=2}$, or further iterations, the new result vectors are applied against the original transition matrix (2).

B. A CYBERSICKNESS MARKOV CHAIN

To model cybersickness as a Markov chain, four states are used to represent the cybersickness experiences across no, light, moderate and strong impact [7], [11]. Fig. 3 shows an initial Markov model for cybersickness. Cybersickness states are defined as no cybersickness (OK), light cybersickness (CSL), medium cybersickness (CSM) and strong cybersickness (CSS). The transitions between state nodes are probabilities to another state or for remaining in the current state.

Although Fig. 3 captures the states of an experience across a four state model of cybersickness, it does not consider any exiting criteria. When considering the prediction of cybersickness, the end of a virtual reality experience, and whether this was triggered by cybersickness impact, is of significant interest.

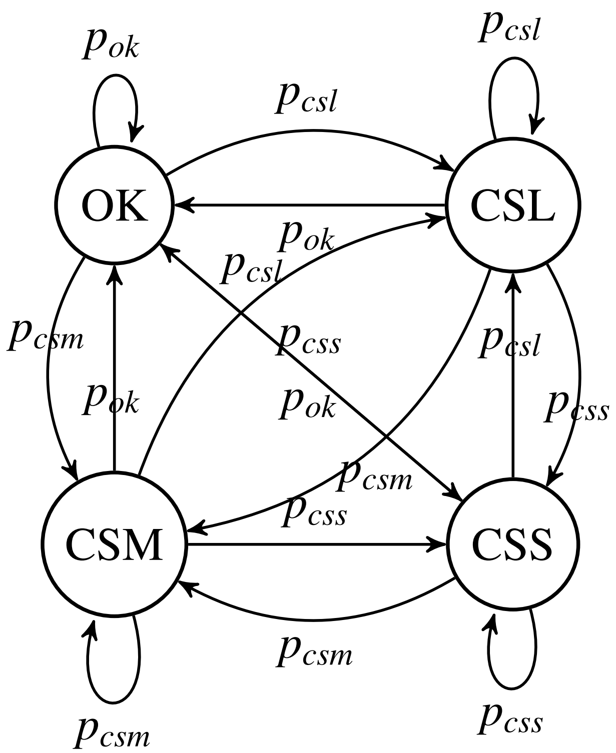

FIGURE 3. Initial cybersickness Markov chain. Cybersickness states are no cybersickness (OK), light cybersickness (CSL), medium cybersickness (CSM) and strong cybersickness (CSS).

Therefore, we can extend the model to have absorbing states, which will capture the probability of users exiting the experience.

C. ABSORBING MARKOV CHAINS

An absorbing Markov chain is a Markov chain that has states that are impossible to leave. Thus there are states in the model that can be reached but the probability of remaining in an absorbing state is 100%, e.g. an absorbing state is a state i in a Markov chain such that $\mathbb{P}(X_{n+1} = i \mid X_n = i) = 1$. In addition to having absorbing states, to be an absorbing Markov chain, all other states must eventually reach an absorbing state.

The transition matrix T for an absorbing Markov chain can be defined as:

$$T = \begin{bmatrix} Q & R \\ 0 & I_s \end{bmatrix} \tag{4}$$

where Q is a $t \times t$ matrix, R is a $t \times s$ matrix (the transitions to the absorbing states), 0 is the $s \times t$ zero matrix (the transitions from the absorbing states, i.e. 0) and I_s is the $s \times s$ identity matrix (the self transitions of the absorbing states). Fig. 4 shows the cybersickness Markov chain extended with absorbing states.

One issue with the models in Figs. 3 and 4 is that each of the main cybersickness states can, theoretically, transition to all other main cybersickness states. In practise, this is very unlikely for some transitions, for example, from a strong cybersickness impact state back to an ok state. Also for the model to be useful for predictive analysis by iterating through the resulting Markov chain, a custom probability matrix is required.

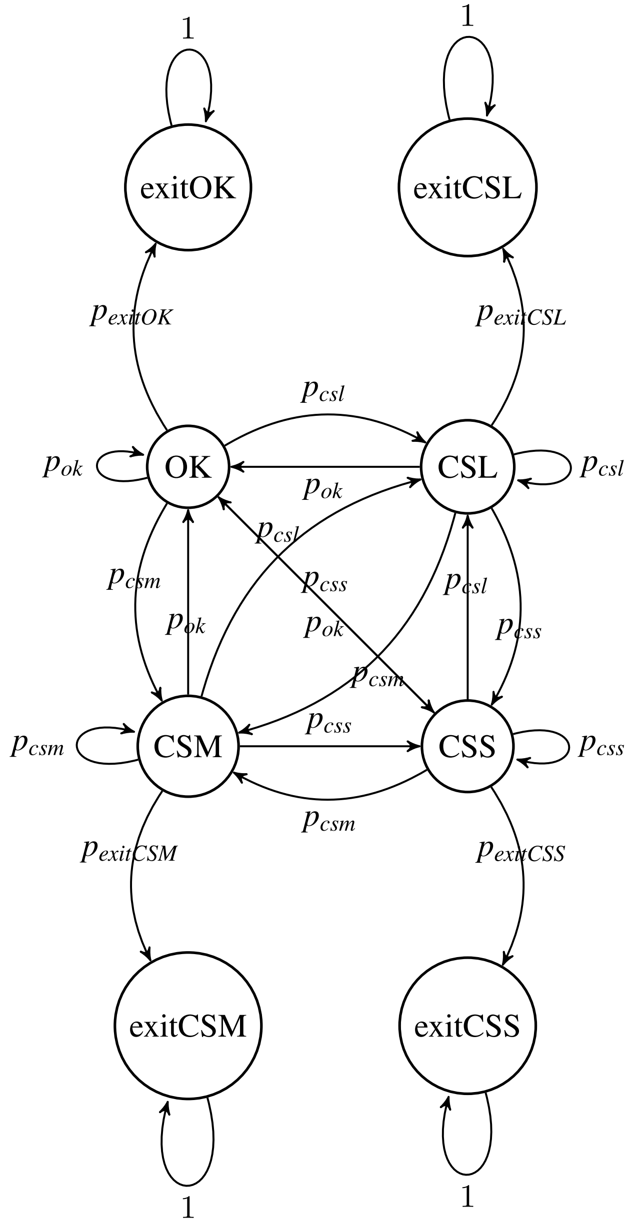

FIGURE 4. Cybersickness Markov chain with absorbing states to represent exiting the virtual reality experience.

IV. BUILDING A CUSTOM CYBERSICKNESS MARKOV CHAIN

Islam et al. [11] conducted a user study to collect data for their work on the detection and prediction of cybersickness severity from a user's physiological signals. In addition to their IEEE publication, they have provided the complete dataset from their study.[1] The dataset includes heart rate, breathing rate, heart rate variability, galvanic skin response and verbal reports of nausea for 23 participants.

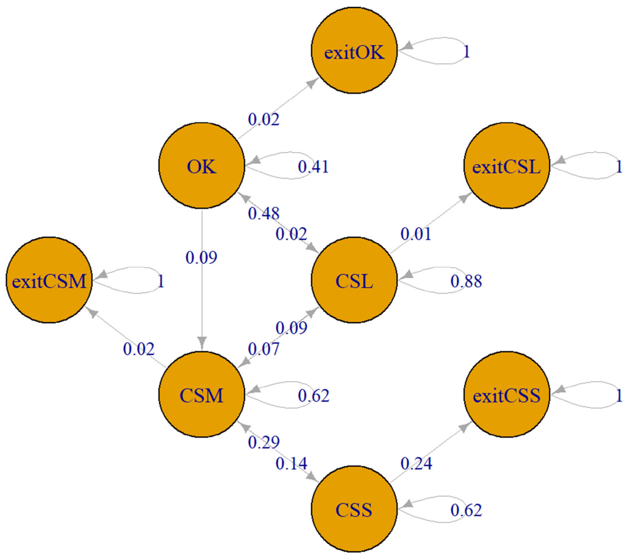

FIGURE 5. Diagrammatic view of the Markov chain transition matrix from the virtual roller-coaster dataset [11].

The verbal responses to nausea were extracted from the dataset to populate the cybersickness Markov chain (Fig. 4.). In [11], the participants rated nausea every minute on a shortened version of the Fast Motion Sickness Scale [14] from 0 (no sickness at all) to 10 (feeling very sick and want to stop) while riding a virtual reality roller-coaster. The verbal nausea scores were then mapped to the definition proposed by [7] across 0 (no nausea), 1-3 (light nausea), 4-6 (medium nausea) and 7+ strong nausea.

Custom R^2 scripts generated probability totals from the nausea reports. Each transition from a changing state, e.g., ok (0 verbal report) to light cybersickness (1-3 verbal report) represented the minute by minute verbal nausea reports across the 8-state model. The transitions from each state were totalled to determine a probability range of 0-1 for each state. The final transition matrix from the [11] data, as a specific encoding of an absorbing Markov chain, i.e. (4), was:

$$T_{Islam2020} = \begin{bmatrix} 0.41 & 0.48 & 0.09 & 0 & 0.02 & 0 & 0 & 0 \\ 0.02 & 0.88 & 0.09 & 0 & 0 & 0.01 & 0 & 0 \\ 0 & 0.07 & 0.62 & 0.29 & 0 & 0 & 0.02 & 0 \\ 0 & 0 & 0.14 & 0.62 & 0 & 0 & 0 & 0.24 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

[1] https://github.com/shovonis/CyberSicknessClassification [last access 9/11/2023]

[2] All R scripts and processing done using RStudio version 2023.09.1+494 and R version 4.3.2 (2023-10-31). All scripts completed in under 3 seconds on a MacBook Pro (14-inch, 2021), Apple M1 Pro chip with 16 GB memory running macOS Monterey version 12.0.1.

TABLE 2. Probabilities of final states with *OK* = 1 seed vector (6).

OK	CSL	CSM	CSS	exitOK	exitCSL	exitCSM	exitCSS
1.3	31	13	11	4.3	8.5	4.3	26

Fig. 5 is the graphic representation of transition matrix (5) and highlights some interesting deviations from the general cybersickness Markov model (Fig. 4). For example, there are no transitions from *medium* or *strong* cybersickness states back to an *ok* state or from the *ok* state directly to the *strong* state. This first observation is indicative of the cumulative nature of cybersickness [26] and the second is likely an indication on how the virtual reality roller-coaster used in [11] was able to gradually induce cybersickness.

V. EXPLORING A CYBERSICKNESS MARKOV CHAIN

With a custom Markov chain defined, it is possible to run iterations through the model by adjusting two core parameters, firstly the number of iterations which in this case are minutes exposed to the virtual environment, and secondly, the seed vector representing the initial configuration of the chained iterations. In the models here, the seed vector can be considered the starting distribution of participant cybersickness, i.e. the likelihood that participants already have some level of nausea. These parameters can then be applied across varying theoretical numbers of participants and VR exposure durations to see how the virtual environment experience scales. However, the accuracy of the Markov model needs to be determined and the model refined to better match the source dataset.

A. REFINING THE MODEL

From [11], an initial assumption is that all participants start in an *ok* state and that they are exposed to the virtual roller-coaster for a maximum of 13 minutes, i.e. 13 iterations through the Markov chain. This would be represented as the seed vector (SV) in (6) and results in the iteration chain shown in Fig. 6.

$$SV1_{t=0} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p_{ok} \\ p_{csl} \\ p_{csm} \\ p_{css} \end{bmatrix} \quad (6)$$

As cybersickness has been found to steadily increase with time spend in an environment [22], Fig. 6 conforms to expected state transitions for a virtual roller-coaster with rapid onset of cybersickness [7] with increasing numbers of participants exiting the virtual environment with strong cybersickness. The chain also provides the final model states, shown in Table 2. The final state probabilities can be multiplied by 23 (n = 23 in [11]) and compared with the ground truth participant experiences extracted from [11] (see Table 3).

A raw measure of the model accuracy can be determined with the edit distance between the ground truth and the seed vector (6) results. In this initial case the edit distance is 4 and gives an overall model accuracy of 82.61%.

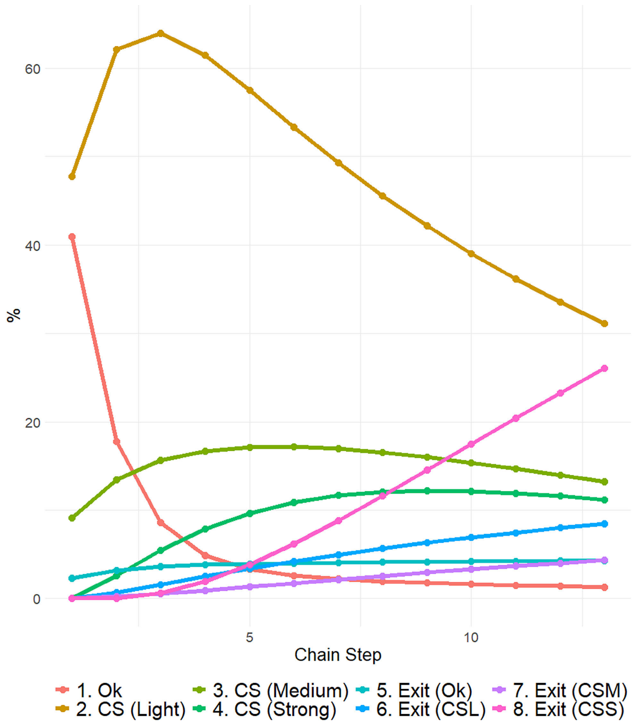

FIGURE 6. Markov chain transitions through 13 iterations from the virtual roller-coaster dataset [11].

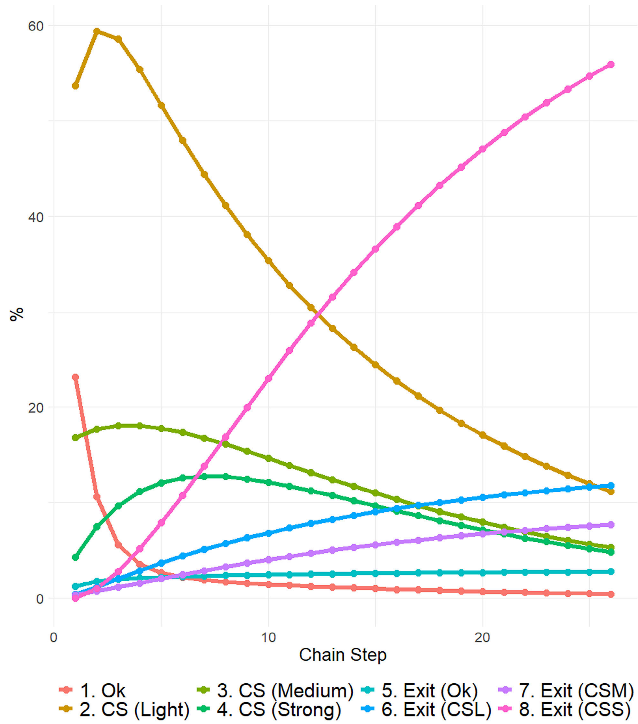

FIGURE 7. Markov chain transitions through 26 iterations.

TABLE 3. Comparing probabilities of final states from the Markov chain to the ground truth (GT) data from [11] (n = 23), with seed vectors (6) and (7).

State / Result	*Ok*	*CSL*	*CSM*	*CSS*	$exit_{OK}$	$exit_{CSL}$	$exit_{CSM}$	$exit_{CSS}$
GT	0	7	4	1	1	2	1	7
SV1 (6)	0	7	3	3	1	2	1	6
SV2 (7)	0	7	3	2	1	2	1	7

However, the assumption here is that all participants started the virtual roller-coaster in an *ok* state, i.e. no feelings of nausea. As the verbal reports are subjective, the degree of *ok* that the participants were feeling will likely have been from varied physical starting points, for example being in different states of alertness or tiredness. Also individual participants will be at different levels of susceptibility to cybersickness impacts. Islam et al. [11] administered a pre-VR exposure simulator sickness questionnaire (SSQ) and found the participants (n = 23) had a mean SSQ score of 8.29 (stdev = 12.71), indicating an initial level of nausea. To capture this and to improve the match to the ground truth experiences from [11], a variety of seed vectors were tested, by trial and error, to improve the initial seed.[3]

$$SV2_{t=0} = \begin{bmatrix} 0.55 \\ 0.3 \\ 0.15 \\ 0 \end{bmatrix} = \begin{bmatrix} p_{ok} \\ p_{csl} \\ p_{csm} \\ p_{css} \end{bmatrix} \tag{7}$$

[3] Automating the optimisation of a seed vector for a known ground truth is the focus of future work.

The results from the refined seed vector (7) are shown in Table 3, with an edit distance of 2. This improves the model accuracy to 91.30%. This result matches 6 of the 8 states to the ground truth and, importantly, matches all the exit states. This seed vector results in one more person with strong, rather than medium, cybersickness at the end of the 13 iterations. This is a minor difference when both medium and strong cybersickness are undesirable outcomes from a virtual environment experience. Thus, unless specifically noted, this is the seed vector that will be used in the following sections as it better represents the ground truth experiences from [11].

B. EXPLORING DROPOUT RATE

The states of the Markov model be can simplified as the states where participants completed the VR experience, i.e., states *ok*, *CSL*, *CSM* and *CSS*, and the four exit states where the participants exited early, i.e. the dropout rate. Also within each of these pairings it is possible to determine the level of cybersickness that the participants reported at the end of their experience.

From Table 3 we can determine, for n = 23 and 13 minutes of VR exposure, that the dropout rate was 48% (agreed by ground truth and the SV2 chain) and for participants that completed the VR experience, 30% had no or only light cybersickness and 22% had medium or strong cybersickness.

However, if the exposure time is doubled, to 26 minutes, the model generates the chain of transitions shown in Fig. 7 and the probabilites in Table 4. The total dropout rate has increased to 78.5% and VR experience completions with no or light cybersickness has decreased to 11.46 %.

TABLE 4. Probabilities of final states with SV2 (7), n = 23 and 26 iterations.

Ok	*CSL*	*CSM*	*CSS*	$exit_{OK}$	$exit_{CSL}$	$exit_{CSM}$	$exit_{CSS}$
0.46	11	5.3	4.8	2.8	12	7.7	56

If a specific dropout rate was required, for example no more that 20% of participants dropout of the VR experience with medium/strong cybersickness, the model can be explored to see at what minute this threshold is broken. In this case, it is 7 iterations where the probability of medium/strong cybersickness dropout is 16.68% and probability of completing the experience with no or light cybersickess is 46.32%.

C. EXPLORING EXPOSURE THRESHOLDS

The model can also be explored for maximum duration to achieve different levels of participants' final cybersickness. For example, the model can show that for 90% of participants to complete the experience with no cybersickness would be impossible. Even after 1 iteration, only 23.1% of participants would complete with no cybersickenss, with 53.7% with light, 16.9% with medium and 4.29% with strong cybersickness. Also 2.04% of participants would have already dropped out (although 0% with strong cybersickness). This is likely due to the model being based on a virtual reality roller-coaster which was used specifically to induce cybersickness [19].

From a health and safety perspective, whether a participant completes or drops out of the VR experience may not be relevant. The concern may be on a participant's level of cybersickness at the end of their experience. For example, if the maximum threshold at end of experience (completion or dropout) was for only 20% of participants to have strong cybersickness, then the maximum number of iterations would be 5, with 19.98% made up of 12.1% CSS completions and 7.88% CSS dropouts (see Fig. 8).

Another example could be to know the duration limit when 80% of participants will likely have medium or higher cybersickness. This would be after 44 iterations with 79.92% probability of medium or higher cybersickness at end of the VR experience. However, at this point only 7.12% would have completed the experience and 67.7% would have already dropped out with strong cybersickness.

D. EXPLORING PARTICIPANT NUMBERS

All the examples in Sections V-B and V-C were based on the participant numbers from the source data, i.e. n = 23. At the Markov chain level, participant numbers are not used as the probabilities for each state are range from [0..1] with the sum of probabilities in a state = 1. However, if participant numbers are required, these can be applied to the final state of the Markov chain, as seen in Table 3 with n = 23 from [11].

If it is useful to consider how a VR experience might scale, then larger numbers of participants can be simulated. However, when considering increasing participant numbers, one limitation is rounding errors. For example with the

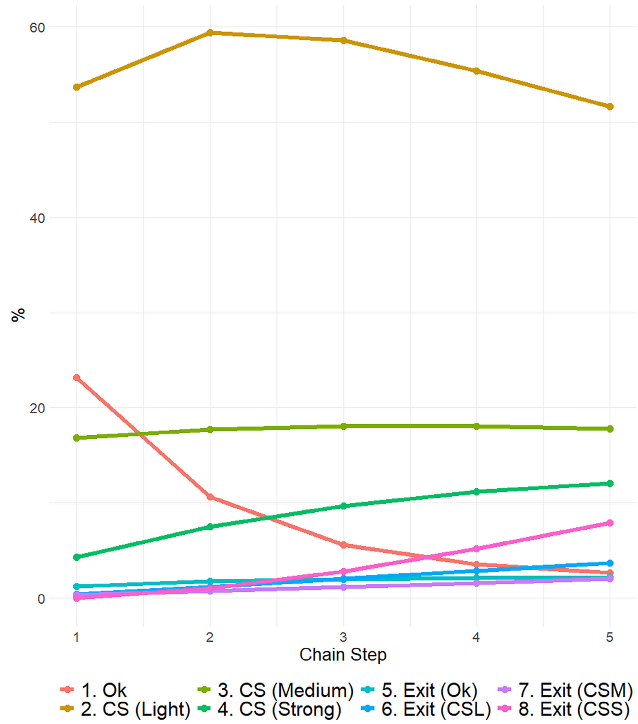

FIGURE 8. Markov chain transitions through 5 iterations.

TABLE 5. Comparing approaches to mapping probabilities to whole participants for n = 100 and the final states from Table (4).

Method	*Ok*	*CSL*	*CSM*	*CSS*	$exit_{OK}$	$exit_{CSL}$	$exit_{CSM}$	$exit_{CSS}$	Total
Round	0	11	5	5	3	12	8	56	100
Floor	0	11	5	4	2	12	7	56	97
Ceiling	1	11	6	5	3	12	8	56	102

double duration chain, with duration set to 26 minutes (Table 4), a simple upscaling would be to consider the impact on 100 participants. As the results have fractional components, these need to be converted to whole participants. Three conversion approaches are to (i) round to the whole person, i.e. a ceiling function, (ii) truncate any fractional component, i.e. a floor function, or (iii) simply use a standard rounding function with > 0.5 rounded up and other values rounded down. These options for the current example can be seen in Table 5, with floor and ceiling functions under and over populating the results respectively. Although the round function has worked well in this example, there is no guarantee that this will always be the case.

Ultimately it comes down to the context in which the participant numbers are to be used and whether there is a preference for underestimating (and using a floor function), for example when considering health and safety issues, or if a more balanced rounding approach is better. As there are only eight states to consider, as a worst case, the error rate will be ±8 participants. When estimating large numbers of participants, the impact may be negligible.

VI. DISCUSSION

The results have shown that a Markov chain model based on a user study dataset can simulate participant cybersickness attributes close to the ground truth of the source material. This can provide insight into dropout rates and completion times on variations of the VR experience. This new approach is in contrast to other work on predicting cybersickness that focused on the use of physiological predictors [10], [11], visual image properties [20] or participant/task characteristics [12].

A formal model, as introduced here, adds utility to data collected in user studies and supports exploring the implications of scaling up any VR experiences, both in terms of participant numbers and/or the duration of exposure. In terms of health and safety, minimum and maximum thresholds for acceptability can be defined and simulated in the model.

However, the work presented here is not without its limitations. Firstly, the general cybersickness Markov chain needs to be populated with representative data to build the custom model. This may still require user studies to gather the initial probabilities. Nevertheless, these user studies may be of a smaller scale and only need to be representative of the target VR experience. The results can then be scaled up in the generated model, without the need for further studies or added risk to participants, i.e. if insight into longer duration of VR exposure was needed.

Alternatively, historical datasets of similar experiences may be used, as described here, to provide insight into the type of environment used, in this case a virtual realty roller-coaster, or provide a baseline for either further user studies or to model simulation of different VR environments. For example, it may be useful to show that a new VR experience is at least as cybersickness inducing as a virtual roller-coaster, if it was being used to habituate VR users.

Secondly, the approaches to optimise model parameters and result thresholds would benefit automation. This was particularly evident with the optimisation of the seed vector (Section V-A) and the exploration of minimum and maximum thresholds for duration (Section V-C). This is a focus for future work.

Finally, the model predictions are based on participants from a specific user study and they may be from a limited demographic. The effect of cybersickness may not be the same for different individuals [12], [27]. It would be more useful if the demographics of the user population could be integrated into the Markov chain model construction. Then, if insight of cybersickness impact was needed from a different demographic, for example older VR users, the probability matrix could be adjusted accordingly to better represent the target demographic. The aim would be to define transformation matrices that either amplify or diminish the original probability matrix so that participants from a demographic with either known susceptibility or resistance to cybersickness could be modelled. A similar approach could be considered for VR experiences of different intensities, with transformation matrices for more or less intense experiences used to further customise the Markov chain probability matrix. A number of datasets and case studies are being explored to feed into this approach and is ongoing work.

VII. CONCLUSION

This paper outlines a new formal approach to modelling cybersickness in virtual reality. A Markov chain is generated from verbal reports of in-situ cybersickness and used to explore a probability model across a number of variations including dropout/completion rates, increased duration of VR exposure, impact by number of participants and thresholds of acceptable VR exposure.

The approach has been demonstrated using an independent dataset from a VR user study. The generated Markov chain has been shown to be representative to the ground truth results of the source dataset. Limitations to the approach have been identified and will become the focus of future work, namely the automation of parameter optimisation and extending the utility of the formal model to participant populations of different demographics and virtual environments of differing intensity. This is ongoing work.

REFERENCES

[1] M. L. Bolton, X. Zheng, and E. Kang, "A formal method for including the probability of erroneous human task behavior in system analyses," *Rel. Eng. Syst. Saf.*, vol. 213, Sep. 2021, Art. no. 107764.

[2] S. Bonfanti, A. Gargantini, and A. Mashkoor, "A systematic literature review of the use of formal methods in medical software systems," *J. Softw. Evol. Process*, vol. 30, no. 5, p. e1943, May 2018.

[3] G. Carlsson, "Topology and data," *Bull. Amer. Math. Soc.*, vol. 46, no. 2, pp. 255–308, 2009.

[4] E. Chang, H. T. Kim, and B. Yoo, "Virtual reality sickness: A review of causes and measurements," *Int. J. Hum.–Comput. Interact.*, vol. 36, no. 17, pp. 1658–1682, Oct. 2020.

[5] A. S. Fernandes and S. K. Feiner, "Combating VR sickness through subtle dynamic field-of-view modification," in *Proc. IEEE Symp. 3D User Interface (3DUI)*, Mar. 2016, pp. 201–210.

[6] M. Fudolig and R. Howard, "The local stability of a modified multi-strain SIR model for emerging viral strains," *PLoS One*, vol. 15, no. 12, Dec. 2020, Art. no. e0243408.

[7] A. M. Gavgani, K. V. Nesbitt, K. L. Blackmore, and E. Nalivaiko, "Profiling subjective symptoms and autonomic changes associated with cybersickness," *Autonomic Neurosci.*, vol. 203, pp. 41–50, Mar. 2017.

[8] Z. Ghahramani, *An Introduction to Hidden Markov Models and Bayesian Networks*. Singapore: World Sci. Publishing Co., 2001, p. 942.

[9] M. Gleirscher and D. Marmsoler, "Formal methods in dependable systems engineering: A survey of professionals from Europe and North America," *Empirical Softw. Eng.*, vol. 25, no. 6, pp. 4473–4546, Nov. 2020.

[10] A. Hadadi, C. Guillet, J.-R. Chardonnet, M. Langovoy, Y. Wang, and J. Ovtcharova, "Prediction of cybersickness in virtual environments using topological data analysis and machine learning," *Frontiers Virtual Reality*, vol. 3, Oct. 2022, Art. no. 973236.

[11] R. Islam, Y. Lee, M. Jaloli, I. Muhammad, D. Zhu, P. Rad, Y. Huang, and J. Quarles, "Automatic detection and prediction of cybersickness severity using deep neural networks from user's physiological signals," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Los Alamitos, CA, USA, Nov. 2020, pp. 400–411.

[12] A. Jasper, N. C. Sepich, S. B. Gilbert, J. W. Kelly, and M. C. Dorneich, "Predicting cybersickness using individual and task characteristics," *Comput. Hum. Behav.*, vol. 146, Sep. 2023, Art. no. 107800.

[13] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *Int. J. Aviation Psychol.*, vol. 3, no. 3, pp. 203–220, Jul. 1993.

[14] B. Keshavarz and H. Hecht, "Validating an efficient method to quantify motion sickness," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 53, no. 4, pp. 415–426, Aug. 2011.

[15] J. J. LaViola, "A discussion of cybersickness in virtual environments," *ACM SIGCHI Bull.*, vol. 32, no. 1, pp. 47–56, Jan. 2000.
[16] S. Massaiu and N. Paltrinieri, "Chapter 14—Human reliability analysis: From the nuclear to the petroleum sector," in *Dynamic Risk Analysis in the Chemical and Petroleum Industry*, N. Paltrinieri and F. Khan, Eds. Oxford, U.K.: Butterworth-Heinemann, 2016, pp. 171–179.
[17] A. D. Mishra and K. Mustafa, "A review on security requirements specification by formal methods," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 5, Feb. 2022, Art. no. e6702.
[18] E. Nalivaiko, S. L. Davis, K. L. Blackmore, and K. V. Nesbitt, "Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time," *Autonomic Neurosci.*, vol. 192, p. 63, Nov. 2015.
[19] K. Nesbitt, S. Davis, K. Blackmore, and E. Nalivaiko, "Correlating reaction time and nausea measures with traditional measures of cybersickness," *Displays*, vol. 48, pp. 1–8, Jul. 2017.
[20] N. Padmanaban, T. Ruban, V. Sitzmann, A. M. Norcia, and G. Wetzstein, "Towards a machine-learning approach for sickness prediction in 360° stereoscopic videos," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 4, pp. 1594–1603, Apr. 2018.
[21] L. Rebenitsch and C. Owen, "Review on cybersickness in applications and visual displays," *Virtual Reality*, vol. 20, no. 2, pp. 101–125, Jun. 2016.
[22] D. Risi and S. Palmisano, "Effects of postural stability, active control, exposure duration and repeated exposures on HMD induced cybersickness," *Displays*, vol. 60, pp. 9–17, Dec. 2019.
[23] V. G. Sánchez, O. M. Lysaker, and N.-O. Skeie, "Human behaviour modelling for welfare technology using hidden Markov models," *Pattern Recognit. Lett.*, vol. 137, pp. 71–79, Sep. 2020.
[24] D. Saredakis, A. Szpak, B. Birckhead, H. A. D. Keage, A. Rizzo, and T. Loetscher, "Factors associated with virtual reality sickness in head-mounted displays: A systematic review and meta-analysis," *Frontiers Hum. Neurosci.*, vol. 14, p. 96, Mar. 2020.
[25] L. Simón-Vicente, S. Rodríguez-Cano, and V. Delgado-Benitoetal, "Cybersickness. A systematic literature review of adverse effects related to virtual reality," *Neurología*, doi: 10.1016/j.nrl.2022.04.009.
[26] S. P. Smith, "Comparing virtual environments for cybersickness using a cumulative optical flow entropy metric," *IEEE Access*, vol. 9, pp. 68898–68904, 2021.
[27] N. Tian, P. Lopes, and R. Boulic, "A review of cybersickness in head-mounted displays: Raising attention to individual susceptibility," *Virtual Reality*, vol. 26, no. 4, pp. 1409–1441, Dec. 2022.
[28] Y. Wang, J.-R. Chardonnet, and F. Merienne, "VR sickness prediction for navigation in immersive virtual environments using a deep long short term memory model," in *Proc. IEEE Conf. Virtual Reality 3D User Interface (VR)*, Mar. 2019, pp. 1874–1881.

SHAMUS P. SMITH (Senior Member, IEEE) received B.Sc., B.Sc. (Hons.), and Ph.D. degrees in computer science from Massey University, New Zealand, in 1992, 1993, and 1999, respectively.

He is currently Senior Lecturer in Immersive Technologies and a member of the Institute for Integrated and Intelligent Systems, Griffith University, Australia. He has published over 100 research articles. His research interests include virtual reality, human–computer interaction, and technology-enhanced learning. He is an Associate Editor of the journal *Entertainment Computing*.