Data Fusion for Difficulty Adjustment in an Adaptive Virtual Reality Game System for Autism Intervention

Lian Zhang[1], Joshua W. Wade[1], Dayi Bian[1], Amy Swanson[3], Zachary Warren[2,3], and Nilanjan Sarkar[1,4]

[1] Electrical Engineering and Computer Science Department
[2] Treatment and Research in Autism Spectrum Disorder (TRIAD)
[3] Pediatrics and Psychiatry Department
[4] Mechanical Engineering Department
Vanderbilt University, Nashville, TN 37212
{lian.zhang,joshua.w.wade,dayi.bian,amy.r.swanson, zachary.e.warren,nilanjan.sarkar}@vanderbilt.edu

Abstract. A virtual reality driving simulator is designed as a tool for improving driving skills of individuals with autism spectrum disorders (ASD). Training at an appropriate driving difficulty level can maximize long term performance. Affective state information has been used for difficulty level adjustment in our previous work. This paper integrates performance with affective state information to predict the optimal difficulty level. The participant's performance data, physiology signals, and eye gaze data are captured. The performance features and affective state features are extracted. Two classification methods, Support Vector Machine (SVM) and Artificial Neural Network (ANN), are implemented to predict difficulty level. The results demonstrate that performance together with affective state information outperform the separated features in difficulty level prediction. A highest accuracy of 83.09% is achieved with the integrated features.

Keywords: autism, pattern recognition, virtual reality, data fusion, cognitive load, difficulty adjustment.

1 Introduction

Researchers are increasingly utilizing virtual reality (VR) technologies as intervention platforms for individuals with autism spectrum disorder (ASD) [1, 2]. The capability for real-time adaptation makes VR games particularly appealing as a potential vehicle for teaching complex skills to individuals with ASD. In adaptive systems, a computer can utilize complex data streams in order to predict a human's cognitive and affective state, and then employ previous and current performance along with current context to modify the game's difficulty level. Such an adaptation to optimize difficulty levels has the potential to maximize the learner's long-term performance within and beyond the system.

C. Stephanidis (Ed.): HCII 2014 Posters, Part I, CCIS 434, pp. 648–652, 2014.

Typically, cognitive load is inferred using model-based data fusion methods [3]. Markov models have been used for an intelligent tutoring system and virtual reality applications [4]. Dynamic Bayesian network model as a data fusion method has the advantage of dealing with temporal information [5]. Artificial Neural Network (ANN) has been used to detect mental workload in real time [6]. The accuracy of these model-based methods, however, is limited by the accuracy of the models.

Additionally, classification methods have been used to recognize cognitive workload. In [7], linear discriminant analysis is applied to determine a game's difficulty level. Fuzzy logic is a simple method for cognitive workload classification [8]. Support Vector Machines (SVM) has a good accuracy in cognitive load recognition [9].

In our previous work, we implemented and evaluated SVM, Bayesian network, k-nearest neighbor and Regression Trees for difficulty level adjustment with affective state in an adaptive system [10]. In this paper, both affective state and performance are integrated using SVM and ANN for difficulty adjustment in a VR based driving task for ASD interventions. The rest of the paper is organized as follows. In Section 2, the VR-based driving task is described. Section 3 discusses the two difficulty adjustment methods used in this work. We present the results in Section 4.

2 VR-Based Driving Task

A VR-based driving simulator was developed to improve the driving skill of individuals with ASD. A set of driving tasks with different difficulty levels was designed to investigate the response of a participant under different cognitive loads. A total of seven ASD teenagers between the ages of 13 and 17 years participated in the driving experiment. Eye gaze data, physiological data, and performance of the participants were captured during the experiment.

The failure, achievement, and driving behavior were recorded as the performance. Biopac MP 150 [11] was used to acquire participant's physiological signals with a sample rate of 1000 Hz in our experiment. The Tobii X120 remote eye-tracker recorded eye gaze behavior [12]. The accuracy of the eye tracker is 0.5°. Considering the variation of both the gaze data and physiological data from person to person, we recorded physiological and eye gaze baseline data for each participant before the experiment.

During the experiment, a therapist observed and reported the participants' affective state every two minutes using a 1-5 Likert scale. The optimal difficulty levels for the driving task were also suggested by the therapist according to her understanding of the participants' affective states and performance. This was done to train and test the fusion methods.

3 Difficulty Level Prediction

We extracted performance features and affective features from the initial data mentioned above offline. Every two minutes of signal was processed in Matlab for the features extraction.

The mean of the driving speed, indicating driving performance, and the number of failures, reflecting the action error, were two essential performance features used in this work. The achievements during driving were reflected by the obtained score, which were the other performance feature.

The mean of the pupil diameter and the blink rate were chosen to be eye gaze features. Unfortunately, we lost some physiological data due to data collection error, so the physiological features were not used in this paper. Instead, the therapist's reports about the participants' affective state were utilized for the difficulty assessment. The consistency between the physiological signal and the therapist report has been studied in our previous work [13-15]. So the reported affective state was a reasonable choice for difficulty level prediction. We chose four affective states in this work: engagement, enjoyment, anxiety, and boredom. Support Vector Machine(SVM) and Artificial Neural Network(ANN) were presented in this paper for the purposes of estimating difficulty level. The Matlab functions of these two methods were used for classification-based fusion.

3.1 Support Vector Machine

Support vector machine is a supervised learning classifier which forms a hyperplane with labeled training data [16]. The defined hyperplane can be used for classification. Given a set of training data $= \{\boldsymbol{x}_i, y_i\}_{i=1}^{n}$, the hyperplane satisfies the equation:

$$\boldsymbol{wx}-b=1 \tag{1}$$

The hyperplane can be found by maximizing (in the Lagrange multipliers $\boldsymbol{\alpha}$) the function

$$L(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i y_j k(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{2}$$

Subject to $0 < \alpha_i < C$ (for $i = 1, 2, \ldots, n$) . $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is the kernel function, which can be replaced by different nonlinear kernel functions. The penalty parameter C is the weight for the misclassification.

3.2 Artificial Neural Network

Neural network model includes an input layer, one or more hidden layers, and an output layer. Each layer is composed of neurons. Neurons in different layers are connected by weighted connections. A two-layer feed-forward network with a sigmoid transfer function in both the hidden layer and output layer was used in this paper[17]. Every neuron receives inputs coming from the previous layers and calculates the sigmoid function of weighted inputs.

$$z_k = \sigma(\boldsymbol{w}_k^T \boldsymbol{x}) \tag{3}$$

$\boldsymbol{w}_k \in \mathbb{R}^m$ is the weight vector pointing to the neuron z_k . $\boldsymbol{x} \in \mathbb{R}^m$ is the input vector of the neuron z_k . The Backpropagation method updates the weights until the error is smaller than a threshold or the iteration number reaches the iteration threshold.

4 Results

ANOVA was used to select the most discriminative features for difficulty level assessment. The result showed that the number of failures from the performance features and the enjoyment level from the affective state features had strong relationship with the difficulty level ($p<0.005$ for each). These two features were selected for data fusion.

The five-scale difficulty levels were clustered into three classes: easier level (level 1 and 2), medium level (level 3), and harder level (level 4 and 5). The accuracies of multiple-class classification were 58.66% for SVM and 52.75% for ANN with selected features. After removing the medium level (level 3), we analyzed the binary classification methods with easier level (level 1 and 2) and harder level (level 4 and 5).

Table 1 shows the binary classification results of SVM and ANN with different features. These accuracies were generated by averaging over 100 computations. Affective state features coming from therapist proved to be most powerful features in difficulty level prediction. Eye gaze features had lower accuracy than other features. The highest accuracy was from the SVM method with the selected features.

Table 1. difficulty level estimation accuracy

	Eye	Performance	Affective	All features	Selected features
SVM	62.36%	65.36%	81.73%	74.82%	**83.09%**
NN	68.50%	68.13%	77.25%	75.38%	78.25%

5 Conclusions

Our previous studies [14, 15] have shown that affective state can be used for difficulty level estimation. This paper improved the game difficulty level assessment by combining multiple signals. Integrating all the features without feature selection did not increase the accuracy; however the selected affective features and selected performance features together lead to the highest classification accuracy.

In the future, we will use this ability to choose an optimal game difficulty level for each individual with ASD to provide a more challenging yet fruitful skill development opportunity.

Acknowledgment. This work was supported in part by the National Institute of Health Grant 1R01MH091102-01A1, National Science Foundation Grant 0967170 and the Hobbs Society Grant from the Vanderbilt Kennedy Center.

References

1. Lahiri, U., Bekele, E., Dohrmann, E., Warren, Z., Sarkar, N.: Design of a Virtual Reality based Adaptive Response Technology for Children with Autism. IEEE Transactions on Neural Systems and Rehabilitation Engineering 21(1), 55–64 (2013)
2. Lahiri, U., Warren, Z., Sarkar, N.: Design of a Gaze-sensitive Virtual Social Interactive System for Children with Autism. IEEE Transactions on Neural Systems and Rehabilitation Engineering 19(4), 443–452 (2011)
3. Novak, D., Mihelj, M., Munih, M.: A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. Interacting with Computers 24(3), 154–172 (2012)
4. Woolf, B., et al.: Affect-aware tutors: recognizing and responding to student affect. International Journal of Learning Technology 4(3), 129–164 (2009)
5. Rowe, J.P., Lester, J.C.: Modeling User Knowledge with Dynamic Bayesian Networks in Interactive Narrative Environments. In: AIIDE (2010)
6. Wilson, G.F., Russell, C.A.: Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. Human Factors: The Journal of the Human Factors and Ergonomics Society 49(6), 1005–1018 (2007)
7. Novak, D., et al.: Psychophysiological measurements in a biocooperative feedback loop for upper extremity rehabilitation. IEEE Transactions on Neural Systems and Rehabilitation Engineering 19(4), 400–410 (2011)
8. Bergasa, L.M., Nuevo, J., Sotelo, M.A., Barea, R., Lopez, M.E.: Real-time system for monitoring driver vigilance. IEEE Transactions on Intelligent Transportation Systems 7(1), 63–77 (2006)
9. Kapoor, A., Picard, R.W.: Multimodal affect recognition in learning environments. In: Proceedings of the 13th Annual ACM International Conference on Multimedia. ACM (2005)
10. Liu, C., Rani, P., Sarkar, N.: An Empirical Study of Machine Learning Techniques for Affect Recognition in Human-Robot Interaction. In: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005), August 02-06, pp. 2451–2456 (2005)
11. Biopac system, `http://www.biopac.com/`
12. Eye tracker X120, `http://www.tobii.com/`
13. Liu, C., Conn, K., Sarkar, N., Stone, W.: Physiology-based affect recognition for computer-assisted intervention of children with Autism Spectrum Disorder. International Journal of Human-Computer Studies 66(9), 662–677 (2008)
14. Rani, P., Sarkar, N., Smith, C., Adams, J.: Affective Communication For Implicit Human-Machine Interaction. In: IEEE International Conference on System, Man and Cybernetics, Washington D. C., pp. 4896–4903 (October 2003)
15. Liu, C., Rani, P., Sarkar, N.: Human-Robot Interaction Using Affective Cues. In: The 15th IEEE International Symposium on Robot and Human Interactive Communication - ROMAN 2006, United Kingdom, pp. 285–290 (September 2006)
16. Boswell, D.: Introduction to support vector machines (2002)
17. Beale, M., Hagan, M.T., Demuth, H.B.: Neural network toolbox. Neural Network Toolbox, The Math Works, 5-25 (1992)