VR Sickness Assessment with Perception Prior and Hybrid Temporal Features

Po-Chen Kuo, Li-Chung Chuang, Dong-Yi Lin and Ming-Sui Lee
Department of Computer Science and Information Engineering,
The Graduate Institute of Networking and Multimedia, National Taiwan University Taipei, Taiwan

Abstract

VR sickness is one of the main reasons that hinder the growth of VR. Since the level of VR sickness is in proportion to the degree of motion in VR content, it would be benefitted if the level of VR sickness can be predicted based on VR contents. In this paper, we propose a novel method with perception prior and hybrid temporal features to predict the occurrence of VR sickness.

Method

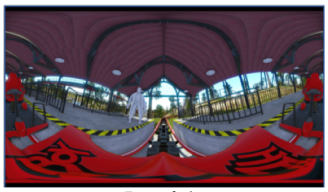
Dataset

20 five-minutes long 360 degree videos from YouTube are downloaded and classified into 4 categories according to the level of motion. Level 1: steady cam. Level 2: steady cam with fast moving object. Level 3: cam with steady movement, e.g. walking, driving. Level 4: fast moving cam, e.g. rollercoaster.

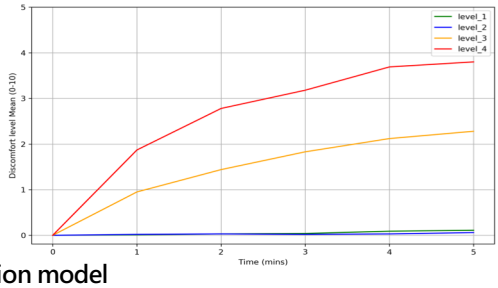

Level 1


Level 2

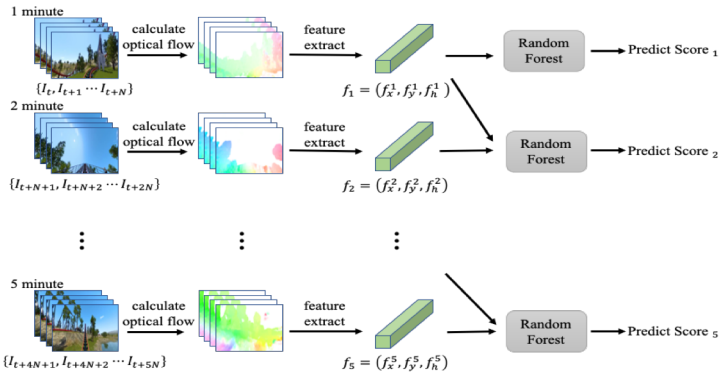

Level 3


Level 4

The testers are asked to report their discomfort score(DS) every minute while watching the video. At the end of video watching, they are asked to answer both the sickness questionnaires (SSQ) and discomfort score (DS). The mean of DS for videos with different levels are shown below.



Prediction model

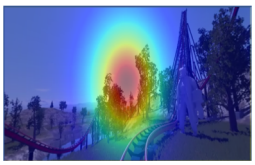


[1] J.-Y. Lee, P.-H. Han, L. Tsai, R.-D. Peng, Y.-S. Chen, K.-W. Chen, Y.-P. Hung, "Estimating the Simulator Sickness in Immersive Virtual Reality with Optical Flow Analysis," in *SIGGRAPH Asia 2017 Posters*, 2017, doi: 10.1145/3145690.3145697
[9] H. G. Kim, H. Lim, S. Lee, and Y. M. Ro, "VRSA Net: VR Sickness Assessment Considering Exceptional Motion for 360° VR Video," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 1646–1660, 2019, doi: 10.1109/TIP.2018.2880509

Perception prior

The extracted optical flow is Gaussian weighted to mimic the human perception system.

width:height	PLCC	SROCC
1:1	0.82	0.81
1:2	0.87	0.83
1:3	0.90	0.76
2:1	0.73	0.75
3:1	0.67	0.58



Hybrid temporal features

Three motion features are derived from the extracted optical flow. The horizontal motion strength and vertical motion strength are used to estimate the spatial movement in the video and the motion anisotropy is adopted to indicate the temporal steadiness of the video content.

Horizontal motion strength :

$$f_x = \frac{1}{t}\sum_{k=1}^{t}\sum_{p(i,j)\in V} x^k(i,j) \odot g(i,j);$$

Vertical motion strength :

$$f_y = \frac{1}{t}\sum_{k=1}^{t}\sum_{p(i,j)\in V} y^k(i,j) \odot g(i,j)$$

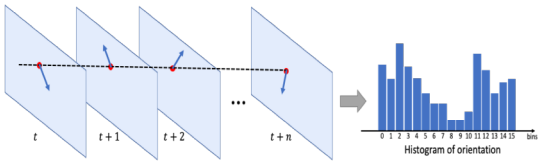
Motion anisotropy :

An orientation histogram is constructed for each optical flow pixel by binning its orientation along the time period.

$$h(i,j) = -\sum_{l=1}^{16} P\left(z^l(i,j)\right)\log_b P\left(z^l(i,j)\right)$$

$$f_h = median\left(h(i,j)\right), p(i,j) \in V$$

The median of the Shannon entropy of all pixels is defined as the motion anisotropy feature.



Experiment results on KAIST dataset

- Execution time comparison

Video id	fps	Lee's [1]	VRSA [9]	Proposed
1	30.00	3.81	249.79	4.57
2	29.97	3.56	249.24	3.98
3	29.97	3.65	249.62	4.10
4	30.00	3.58	249.46	4.13
5	29.97	3.57	249.95	4.07
6	29.97	3.70	250.03	4.34
7	25.00	3.63	208.13	4.07
8	59.94	3.72	500.23	4.25
9	29.97	3.67	248.98	4.23

- Performance comparison and ablation study

Method	Perception Prior	Hybrid temporal feature			PLCC	SROCC
		Horizontal Motion	Vertical Motion	Motion Anisotropy		
Lee's [1]	N/A	N/A	N/A	N/A	0.75	0.80
VRSA [9]	N/A	N/A	N/A	N/A	0.89	0.88
Proposed		✓			0.57	0.40
			✓		0.69	0.66
	✓	✓			0.79	0.75
	✓		✓		0.82	0.82
	✓	✓	✓		0.87	0.83
	✓		✓	✓	**0.91**	**0.92**
	✓	✓	✓	✓	0.90	0.90

Conclusion

Our contributions in this paper are threefold: First, a dataset of VR contents with per minute discomfort score are collected. Second, a novel hybrid temporal feature with perception prior is proposed. Third, the performance to predict the VR sickness is comparable to the state-of-the-art methods.

The 25th International Conference on Pattern Recognition (ICPR 2020)