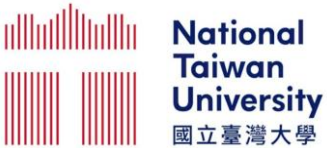
VR Sickness Assessment with Perception Prior and Hybrid Temporal Features

Po-Chen Kuo, Li-Chung Chuang, Dong-Yi Lin and Ming-Sui Lee



Department of Computer Science and Information Engineering,
Graduate Institute of Networking and Multimedia,
National Taiwan University, Taiwan

Speaker: Li-Chung Chuang

VR sickness

Dizziness

Accelerated heartbeat

VR Sickness

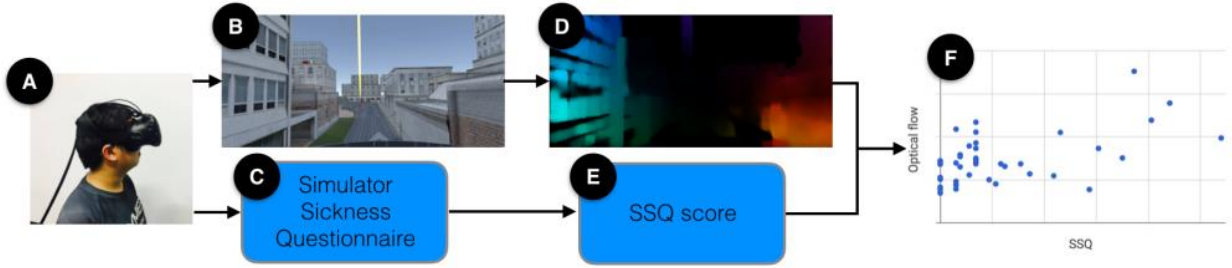
Weakness

nausea

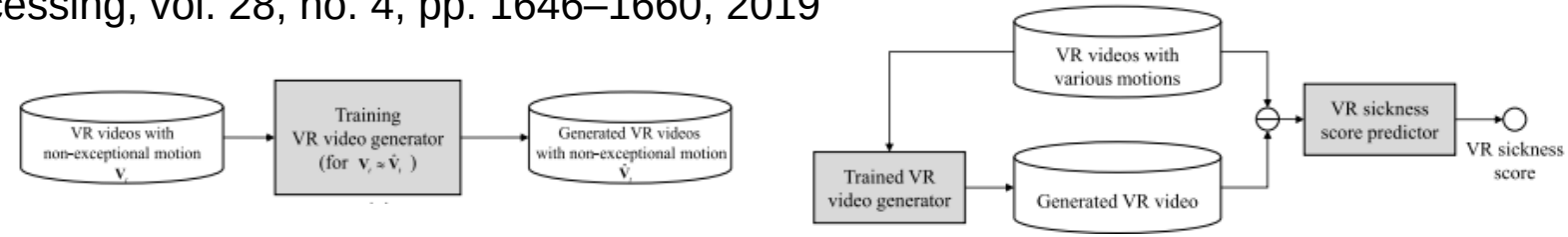
Unsteadiness

Related work

[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



[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



Goal

VR Content

Predict

VR Sickness

Contributions:

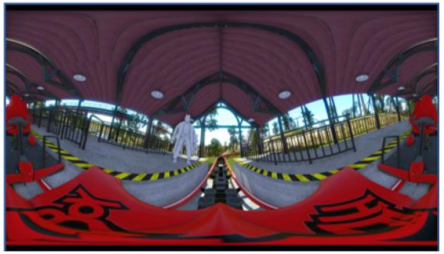
1. A newly created dataset of VR content with per-minute Discomfort Score
2. A novel hybrid temporal feature with perception prior
3. Performance is on par with the state-of-the-art method

Dataset

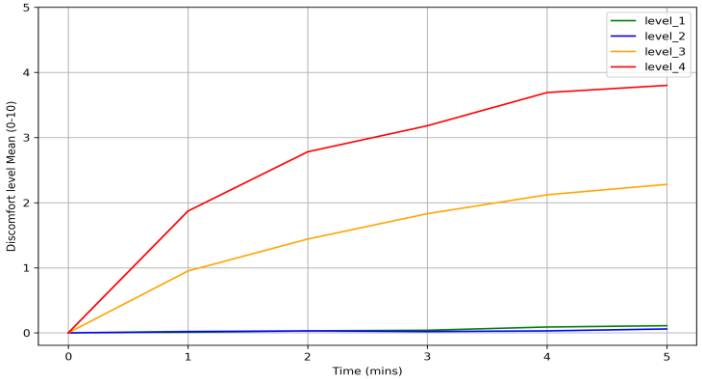
Level 1

Level 2

Level 3



Level 4



- 20 videos, each is 5 minutes
- Every video is watched by 5 testers and report their Discomfort Score every minute
- The average discomfort scores of 5 testers is the DS of that video

Perception prior features

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

- Optical flow based features
- Image area outside 110° FOV are culled out
- Image area inside the FOV are weighted by a Gaussian kernel

Hybrid temporal features

- **Horizontal motion strength**

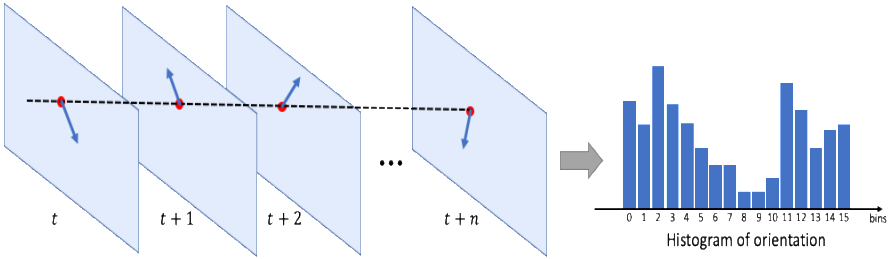
$$f_x = \frac{1}{t}\sum_{k=1}^{t}\sum_{p(i,j)\in V} x^k(i,j) \odot g(i,j);$$

- **Vertical motion strength**

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

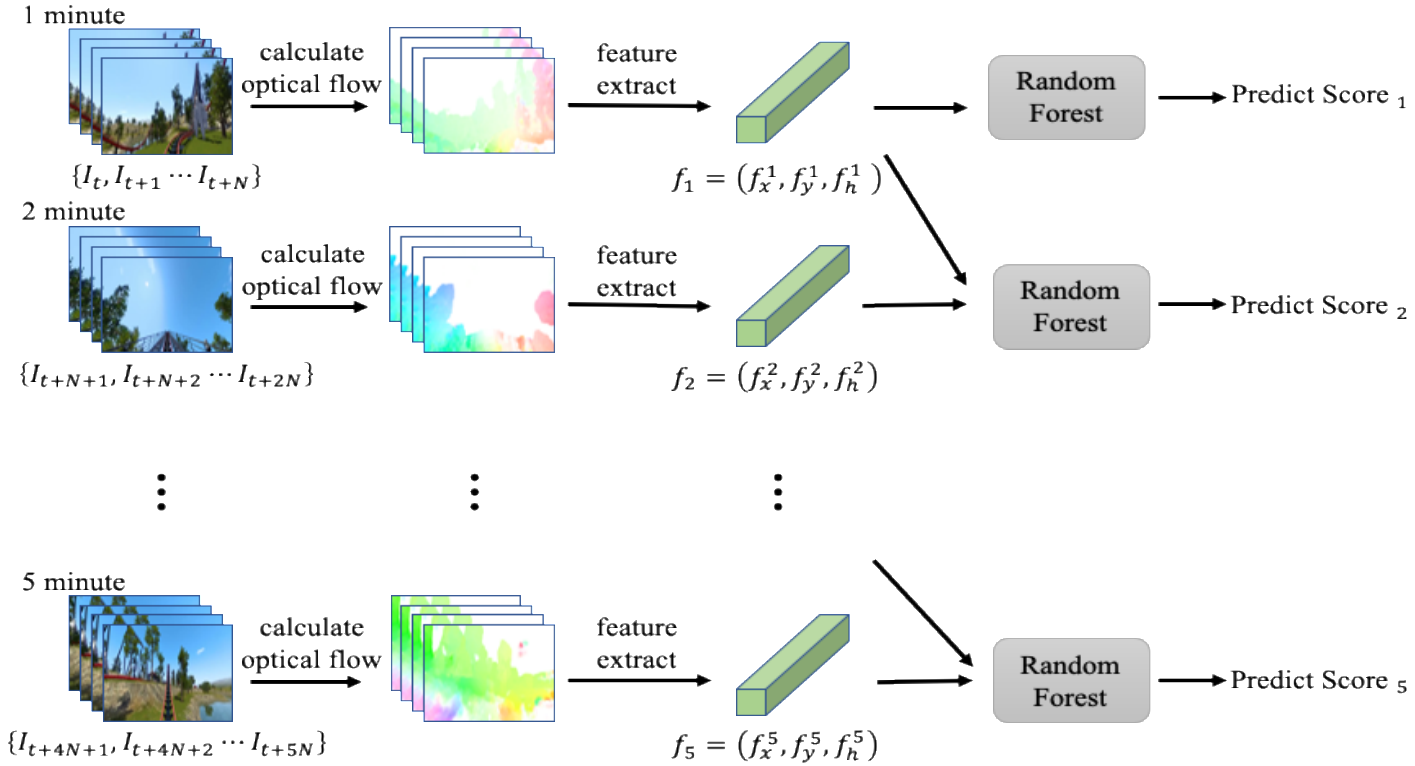
- **Motion anisotropy**

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



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

Training model



Experimental results on KAIST dataset

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

- The proposed method outperforms VRSA and Lee's method on the KAIST dataset.
- The he ablation test demonstrated that the proposed perception prior and hybrid temporal features are effective.

Execution time comparison on KAIST dataset

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

Conclusion

- We collected a dataset of twenty 360 degree videos with per-minute Discomfort Score
- A novel hybrid temporal feature with perception prior was proposed
- The experiment results show that the proposed method is comparable to the state-of-the-art methods