Deep Learning-based VR Sickness Assessment

IEEE SA WG3079 Meeting
: HMD based 3D Content Motion Sickness Reducing Technology

Hak Gu Kim
hgkim0331@kaist.ac.kr

Image and Video Systems Lab.
School of Electrical Engineering
Korea Advanced Institute of Science and Technology (KAIST)

July. 2018
Virtual Reality (VR)

VR displays

360 cameras

Training

Entertainment

Education

Healthcare
Immersive 360-degree VR Content
VR Sickness (≈Cybersickness)

Concern about viewing safety of VR

Interest of VR
Main Factors of VR Sickness

- **Visual-Vestibular Conflict**
  - Mismatches between simulation motion of VR content and viewer’s motion

VR Sickness Assessment (VRSA)

Objective assessment

Subjective assessment
Objective VRSA: Physiological Measurements
Objective VRSA: Physiological Measurements

Physiological measurements

EEG

EGG
Objective VRSA: Physiological Measurements

Physiological measurements:
- EEG
- EGG
- HR/HRV
Subjective VRSA: Subjective Questionnaires

Physiological measurements
- EEG
- EGG
- HR/HRV

Subjective questionnaires
- SSQ
- MSSQ
Proposed Framework for VRSA

Conventional

Proposed

Human Perception

Physiological measurement

Subjective score prediction

Impractical

Inaccurate

Virtual Human Perception (AI Perception Model)

Content analysis

Accurate subjective score prediction

Practical

Accurate
Challenges

- **Lack of labeled datasets**
  - It is difficult to collect a large-scale of fully labeled datasets
    - VR content and the corresponding subjective scores

A large number of target images

Regression

Corresponding subjective score
Proposed Method (1/5)

- **Visual-vestibular conflict**
  - Caused by exceptional motion of VR content

**Simulation motion of VR content**

**Physical motion of viewer**

**Motion mismatch**

**Visual cue**

**Vestibular cue**

**VR Sickness**
Main Idea of Our Research

- Deep learning based generative model
- Non-excessive VR sickness feature representation (instead of excessive VR sickness feature representation)
Proposed Method (3/5)

- **Overall Procedure of the Proposed VRSA Framework**
  - In training, the convolutional autoencoder is trained to reconstruct original VR video sequences with non-exceptional motion such as slow and moderate motion velocity.
  - In testing, by measuring the reconstruction error of the motion information in VR video content, the exceptional motion of VR video content can be detected and measured.
Proposed Method (3/5)

- **Overall Procedure of the Proposed VRSA Framework**
  - In training, the convolutional autoencoder is trained to reconstruct original VR video sequences *with non-exceptional motion* such as slow and moderate motion velocity.
  - In testing, by measuring the reconstruction error of the motion information in VR video content, the exceptional motion of VR video content can be detected and measured.
Proposed Method (4/5)

- **Deep Convolutional Autoencoder for Normal Motion Patterns Learning**
  - Encoder for representing the latent spatio-temporal feature of input sequence
  - Decoder for reconstructing the original sequence from the encoded features
Proposed Method (4/5)

- **Deep Convolutional Autoencoder for Normal Motion Patterns Learning**
  - Encoder for representing the latent spatio-temporal feature of input sequence
  - Decoder for reconstructing the original sequence from the encoded features

![Diagram of the autoencoder process](image)
Proposed Method (4/5)

- Deep Convolutional Autoencoder for Normal Motion Patterns Learning
  - Encoder for representing the latent spatio-temporal feature of input sequence
  - Decoder for reconstructing the original sequence from the encoded features
Proposed Method (5/5)

- Exceptional Motion Pattern Score
  - Reconstruction error at \( t \)-th frame of test dataset
    \[
    e(t) = \sum_{i=1}^{W} \sum_{j=1}^{H} \| I(i, j, t) - \hat{f}_W(I(i, j, t)) \|^2
    \]
    \( I(i, j, t) \): Original \( t \)-th frame
    \( \hat{f}_W(I(i, j, t)) \): Reconstructed \( t \)-th frame
    \( W \) and \( H \): width and height of the frame
  - Proposed exceptional motion pattern score
    \[
    s_m(t) = \frac{e(t)}{\sqrt{W \times H}}
    \]
    \( e(t) \): Reconstruction error at \( t \)-th frame
    \( W \) and \( H \): width and height of the frame
Datasets for Training

- UCSD Ped 1 and Ped 2
  - Ped 1: 34 training video clips (200 frames)
  - Ped 2: 16 training video clips (120~180 frames)
- Avenue
  - 16 training video clips (180~360 frames)
- KITTI benchmark
  - 61 video clips
Datasets for Test

- Three 360-degree VR video contents, collected from YouTube[3],[4],[5]
  - Video 1: Slow velocity
  - Video 2: Moderate velocity
  - Video 3: Fast velocity

[3] https://www.youtube.com/watch?v=JEr3-FzSgzk
[4] https://www.youtube.com/watch?v=wECZs7hewjY
[5] https://www.youtube.com/watch?v=wfNvZwN87Hg
Experiments and Results (3/7)

- **Subjective Assessment Experiment**
  - Equipment for displaying VR content
    - Oculus Rift CV1 HMD
      - 2160 x 1200 pixels @ 90 Hz
      - FoV: 110 degree
    - Intel Core i7-4770@3.4 GHz, 32GB RAM, and NVIDIA GTX 1080TI
  - Subjects
    - 15 subjects, ranging between 20 to 30 years old
      - Normal or corrected-to-normal vision
      - Minimum stereopsis: 60 arcsec.
Experiments and Results (4/7)

Subjective Assessment Experiment

- 16-item SSQ[6]

Procedure

- A week before the actual subjective assessment experiments, we had subjects experience a variety of VR contents with Oculus Rift in order to allow them familiarize with VR environment.
- Every VR video content was displayed for 2 minutes through Oculus Rift CV1.
- After watching the VR video contents, subjects rated their perception of the VR sickness for each symptom in SSQ sheet.

### Subjective Assessment Results

- Degree of VR sickness subjects felt was proportional to the motion magnitude.
- The subjective results show that subjects felt some symptoms of VR sickness when watching VR video 2 and 3. In particular, VR video 3 could lead to excessive VR sickness.

#### Graph

<table>
<thead>
<tr>
<th>VR video contents</th>
<th>Total SSQ score</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR video1</td>
<td>20</td>
</tr>
<tr>
<td>VR video2</td>
<td>40</td>
</tr>
<tr>
<td>VR video3</td>
<td>60</td>
</tr>
</tbody>
</table>

- **Excessive**
- **Perceivable**
- **Unnoticeable**
Experiments and Results (6/7)

- **Performance of the Proposed Network**

Reconstruction results of the proposed deep convolutional autoencoder network for VR video 1 with slow motion (left) and VR video 3 with exceptional motion (right)
Experiments and Results (7/7)

- **Performance of the Proposed Network**

  360 degree video with slow motion pattern

  ![Original seq.](image1)
  ![Reconstructed seq.](image2)
  ![Recon. error](image3)

  360 degree video with fast motion pattern

  ![Original seq.](image4)
  ![Reconstructed seq.](image5)
  ![Recon. error](image6)
Conclusions

- This paper presented a novel measurement of exceptional motion using deep convolutional autoencoder network for assessing the VR sickness of VR video content.

- The convolutional autoencoder learned by normal datasets with slow and moderate motion could reconstruct the non-exceptional motion patterns but it could not recover VR video content having exceptional motion.

- Based on the fact that exceptional motion led to high reconstruction errors in the deep autoencoder network, the level of VR sickness of the input VR video content due to exceptional motion could be predicted.

- The results of our subjective assessment experiments showed that the proposed objective measure strongly had a high correlation with human subjective quality scores, SSQ of our test datasets (PLCC was 0.92).
Thank you

http://ivylab.kaist.ac.kr
hgkim0331@kaist.ac.kr
Appendix

Total SSQ score
• 1. Nausea score
• 2. Oculomotor score
• 3. Disorientation score
• Total SSQ = 3.74 x (1 + 2+ 3)

Table 1: SSQ used in our subjective assessment

<table>
<thead>
<tr>
<th>SSQ Symptoms</th>
<th>Nausea</th>
<th>Oculomotor</th>
<th>Disorientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>General discomfort</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Fatigue</td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Headache</td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Eye strain</td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Difficulty focusing</td>
<td></td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Increased salivation</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Sweating</td>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nausea</td>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty concentrating</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Fullness of head</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blurred vision</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Dizzy (Eyes open)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dizzy (Eye closed)</td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Vertigo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stomach awareness</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Burping</td>
<td>O</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>