EdgeMask: An Edge-based Privacy Preserving Service for Video Data Sharing

Samira Taghavi and Weisong Shi

Wayne State University
http://thecarlab.org
Outline

• Research problem
• Literature review
• Background
• Contributions
• High level architecture
• Algorithm
• Evaluation
• Future works
Research problem

- ADAS, SLAM, and smart cities required image or video data from public places
- These data may contain sensitive personal identifying data
- Any using or sharing these raw datasets can have legal implications
Edge Mask: Goals and objectives

- Privacy protection of individuals in video data
  - Blurring sensitive information
  - Undesired object removal
Related works
Background

What is optical flow?

- 2-Dimensional Map
- Color: Direction
- Saturation: Magnitude

Flying chair dataset
### Background


|---------------|--------------|---------------------------|-----------------------|----------------------|------------------|

---

EPE: 6.18  13.11  9.24  7.92
EPE: 7.22  7.35  7.71  4.70
## Background

<table>
<thead>
<tr>
<th>Number of Networks</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture</td>
<td>s</td>
<td>ss</td>
<td>sss</td>
<td></td>
</tr>
<tr>
<td>Runtime</td>
<td>7ms</td>
<td>14ms</td>
<td>20ms</td>
<td>–</td>
</tr>
<tr>
<td>EPE</td>
<td>4.55</td>
<td>3.22</td>
<td>3.12</td>
<td>–</td>
</tr>
<tr>
<td>Architecture</td>
<td>S</td>
<td>SS</td>
<td></td>
<td>–</td>
</tr>
<tr>
<td>Runtime</td>
<td>18ms</td>
<td>37ms</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>EPE</td>
<td>3.79</td>
<td>2.56</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Architecture</td>
<td>c</td>
<td>cs</td>
<td>css</td>
<td>csss</td>
</tr>
<tr>
<td>Runtime</td>
<td>17ms</td>
<td>24ms</td>
<td>31ms</td>
<td>36ms</td>
</tr>
<tr>
<td>EPE</td>
<td>3.62</td>
<td>2.65</td>
<td>2.51</td>
<td>2.49</td>
</tr>
<tr>
<td>Architecture</td>
<td>C</td>
<td>CS</td>
<td>CSS</td>
<td></td>
</tr>
<tr>
<td>Runtime</td>
<td>33ms</td>
<td>51ms</td>
<td>69ms</td>
<td>–</td>
</tr>
<tr>
<td>EPE</td>
<td>3.04</td>
<td>2.20</td>
<td><strong>2.10</strong></td>
<td>–</td>
</tr>
</tbody>
</table>

Contributions

- **System Level**
  - The *first* approach that leverages both object blurring and removal technique
  - An edge-based real-time privacy preserving framework
  - Object segmentation without requiring complex network architectures

- **Algorithm Level**
  - Tracking camera velocity and direction based on optical flow for static object segmentation
  - Generating object mask for video inpainting based on optical flow and Mask R-CNN
High-Level Architecture

- Edge-Intelligence
- Application Programming Interface
- Security Layer
## Algorithm Design

<table>
<thead>
<tr>
<th>Dynamic object segmentation and blurring</th>
<th>Static object segmentation and blurring</th>
</tr>
</thead>
<tbody>
<tr>
<td>• FlowNet2.0</td>
<td>• Mask R-CNN</td>
</tr>
<tr>
<td>• MBD</td>
<td>• MBD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Congested scene analysis</th>
<th>Video inpainting</th>
</tr>
</thead>
<tbody>
<tr>
<td>• FlowNet2.0</td>
<td>• Mask R-CNN</td>
</tr>
<tr>
<td>• MBD</td>
<td>• MBD</td>
</tr>
<tr>
<td></td>
<td>• FlowNet2.0</td>
</tr>
<tr>
<td></td>
<td>• Deep Flow Guided video inpainting</td>
</tr>
</tbody>
</table>
Dynamic Object Segmentation

“Zhang et al. Minimum Barrier Salient Object Detection at 80 FPS”

- Original image

- Optical flow
Dynamic Object Segmentation

\[ S_i^t(F^t) = \sum_{\forall F_j^t \in F_t} d(F_i^t, F_j^t) \]

“d” stands for Minimum Barrier Distance (MBD) Transform

\[ \beta_I(\pi) = \max_{i=0}^{k} I(\pi(i)) - \min_{i=0}^{k} I(\pi(i)). \]
Static Object Segmentation

1- Find the static background region

\[ M_b = \{M_1, \ldots, M_k\} \]

\[ M_o = \overline{M_b} \& MBD \]

The background region \( = \overline{M_o} \)

2- Move the static object masks generated by Mask R-CNN based on the background velocity and direction.
Video Inpainting

- Video inpainting is performed in offline mode
- The required inputs are:
  - Video Frames
  - Object Masks
    - Generated by Mask RCNN and “MBD”
  - Objects’ Optical flows
    - Generated by FlowNet2.0
Video Inpainting

Xu et al. Deep Flow-Guided Video Inpainting, CVPR 2019
Video Inpainting

Object recognition for parallel video inpainting

- Object label
- Color histogram
- Object direction
Video Inpainting
Experiment Setup

- **Dataset**
  - DAVIS 2016

- **Hardware setup**
  - | Hardware | Model | Memory |
  - | GPU | NVIDIA GeForce GTX 1060 | 3G |
  - | VPU | Intel Neural Compute Stick 2 | - |

- **Model setup**
  - | Mask RCNN | Parameter | Value |
  - | | Backbone Pretrained Weights | ResNet101 COCO Dataset |
  - | Inpainting | Backbone Flow Extraction Propagation Kernel Enlarged Mask Enlarged | ResNet101 FlowNet2 Bidirectional 50 or 70 YES |
  - | Flownet2.0 | Backbone Model | ResNet101 FlowNet2.0 |
Evaluation

- The quality evaluation

<table>
<thead>
<tr>
<th>Approach</th>
<th>Kernel 70</th>
<th>Kernel 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xu et al. [17]</td>
<td>27.838</td>
<td>28.011</td>
</tr>
<tr>
<td>EdgeMask</td>
<td>27.851</td>
<td>28.026</td>
</tr>
</tbody>
</table>

- The computational time based on a single frame

<table>
<thead>
<tr>
<th>Approach</th>
<th>Object blurring</th>
<th>Object removal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frome et al. [6]</td>
<td>10.1s</td>
<td>-</td>
</tr>
<tr>
<td>Flores et al. [5]</td>
<td>-</td>
<td>31.6s</td>
</tr>
<tr>
<td>Agrawal et al. [3]</td>
<td>12s</td>
<td>-</td>
</tr>
<tr>
<td>Nodari et al. [12]</td>
<td>-</td>
<td>21.4s</td>
</tr>
<tr>
<td>EdgeMask</td>
<td>80ms</td>
<td>1.16s</td>
</tr>
</tbody>
</table>
Efficiency

- P stands for parallel
- S stands for serial
Evaluation

- GPU power

The GPU power for parallel and serial computing units (#frames == 280)
Discussion & Future work

- Neural network pruning for video inpainting algorithm
- Evaluate in more realistic scenarios
- Generality of EdgeMask

  EdgeMask workflow can generalize to many other identification tasks
thank you

samira.taghavi@wayne.edu