Location-based Image Viewing System Synchronized with Video Clips

Maho Nishizawa (Kwansei Gakuin University, Japan)
Yuanyuan Wang (Yamaguchi University, Japan)
Yukiko Kawai (Kyoto Sangyo University, Japan)
Kazutoshi Sumiya (Kwansei Gakuin University, Japan)
Various video clips, such as educational, drama, travel videos

→ want to know more information about the area
→ want to know other areas which is related the area

SNS image sharing sites: Instagram, flickr

A huge number of location-based image data with hashtags
  • location names are often used as hashtags

Goal

Supporting and arousing user interests & user knowledge
Our Approach

- Supporting and arousing user knowledge & user interests
- Showing deep and relevant information for video clips by using location-based images and hashtags from Instagram

Deep information • • • detail information based on location names in CC
Relevant information • • • information based on location names not included in CC

INPUT (video clips) ➔ OUTPUT (images from Instagram)
Deep and Relevant Information

Deep Information
- Kobe City / The Mount of Rokko
- Kobe City / Sannomiya
- Kobe City / Sannomiya / Chinese town

Relevant Information
- Kobe City = Port town in Japan
- Kobe City / Yokohama City / Nagasaki City
Image Viewing System with Video Clips

- Showing of **deep information** with images from Instagram about video clips
- Showing of **relevant information** with relevant hashtags from Instagram about video clips

If a user clicks a relevant hashtag, this system will show new relevant images and new relevant hashtags from Instagram.

※User screen shot

- A Video Clip Play Window
- A Image Window
- A Map Window

Relevant Topics of Fukuoka Castle
- Castles located in Fukuoka
  - Najima Castle
  - Bessyo Castle
  - Akizuki Castle
  - Mizuki

Relevant Topics of Fukuoka Castle
- Castles located in Fukuoka
- #Najima castle
- #Bessyo Castle
- #Akizuki Castle
- #Mizuki

Images of 「Nagima Castle」
System Overview

1. Detecting location information and scenes of video clips

2. Extracting images from Instagram by using location information

3. Extracting relevant hashtags from Instagram by matching relevant topics in the semantic structure from Wikipedia
System Overview

1. Detecting location information and scenes of video clips

2. Extracting images from Instagram by using location information

3. Extracting relevant hashtags from Instagram by matching relevant topics in the semantic structure from Wikipedia
1. Detecting Location Information and Scenes of Video Clips (1/2)

- Extracting location information (location names and spots) and their appearance time in **Closed Captions (CC)**
- Dividing scenes of each location information, if the next one appears in **CC**

---

<table>
<thead>
<tr>
<th>Video Clip</th>
<th>Fukuoka city (15 sec)</th>
<th>Fukuoka castle (38 sec)</th>
<th>Hakata area (8 sec)</th>
<th>Fukuoka castle (20 sec)</th>
<th>Nakasu area (14 sec)</th>
</tr>
</thead>
</table>

95 seconds
1. Detecting Location Information and Scenes of Video Clips (2/2)

“New discovery! Wondering”* Caption Data

<table>
<thead>
<tr>
<th>Time</th>
<th>Extracted location information from CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0’05”</td>
<td>Fukuoka</td>
</tr>
<tr>
<td>0’20”</td>
<td>Fukuoka Castle</td>
</tr>
<tr>
<td>0’31”</td>
<td></td>
</tr>
<tr>
<td>0’46”</td>
<td></td>
</tr>
<tr>
<td>0’58”</td>
<td>Hakata</td>
</tr>
<tr>
<td>1’09”</td>
<td>Fukuoka</td>
</tr>
<tr>
<td>1’18”</td>
<td>Fukuoka Castle</td>
</tr>
<tr>
<td>2’14”</td>
<td></td>
</tr>
<tr>
<td>2’28”</td>
<td>Nakasu</td>
</tr>
<tr>
<td>3’35”</td>
<td>Ganso Nagahamaya</td>
</tr>
<tr>
<td>3’45”</td>
<td></td>
</tr>
</tbody>
</table>

“Good Place”** Caption Data

<table>
<thead>
<tr>
<th>Time</th>
<th>Extracted location information from CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0’14”</td>
<td>Lake Biwa</td>
</tr>
<tr>
<td>0’19”</td>
<td></td>
</tr>
<tr>
<td>0’24”</td>
<td>Yogo Lake</td>
</tr>
<tr>
<td>1’15”</td>
<td>Nagahama</td>
</tr>
</tbody>
</table>

If the same location information continuously appear in short time, they are determined in one scene ex. “Fukuoka Castle”, “Lake Biwa”

One scene is too short if the next location information appears within $T$ seconds this next location information will be eliminated ($T = 3$seconds)

* : https://www.youtube.com/watch?v=jI0UBH2SwDg&index=10&list=WL
** : https://www.youtube.com/watch?v=7_QP2BaaBAg&index=18&list=WL
System Overview

1. Detecting location information and scenes of video clips

2. Extracting images from Instagram by using location information

3. Extracting relevant hashtags from Instagram by matching relevant topics in the semantic structure from Wikipedia
2. Extracting Images from Instagram by Using Location Information

- Using extracted location information as **object hashtags**, and search for images from Instagram.

![Video Clip]

- Fukuoka
- Fukuoka Castle
- Hakata
- Fukuoka Castle
- Nakasu

**Number of “Like” Presenting top-P images**

#Fukuoka Castle  Search
System Overview

1. Detecting location information and scenes of video clips

2. Extracting images from Instagram by using location information

3. Extracting relevant hashtags from Instagram by matching relevant topics in the semantic structure from Wikipedia
3. Extracting Relevant Hashtags from Instagram

- Constructing the **tree structure** of the video clip by using *Wikipedia* categories, and extracting relevant topics
- Extracting relevant hashtags from Instagram by matching the extracted relevant topics

<**Problems**>
1. A tree structure has a lot of nodes
2. Same location information may exist in different tree structures

<**Solutions**>
1. Filtering nodes of a tree structure if the reference pages of each node more than N pages (N=5)
2. Selecting one tree structure of a location information by comparing the number of reference pages
Problem 1. A Tree Structure Has A Lot of Nodes

The tree structure of “New discovery! Wondering”*
※ blue text = location information

[Diagram]

Fukuoka Prefecture
- 19 sub categories, 7 pages

Castle at Fukuoka
- 53 pages

- Fukuoka Castle
- Najima Castle
- Bessyo Castle
- Akizuki Castle
- Mizuki

Municipality
- 61 sub categories, 63 pages

- Fukuoka City
  - 15 sub categories, 9 pages
  - Hakata-ku
    - 6 sub categories, 74 pages
    - Hakata
    - Nakasu
    - ...
Problem 2. Same Location Information May Exist in Different Tree Structures

Two patterns of tree structure about “Fukuoka Castle”

The number of Reference Pages

53 pages (Castle) > 28 pages (History)
3. Extracting Relevant Hashtags from Instagram

Recommending categories and pages which are parallel relationships with location information in the tree structure of the video clip.

Presenting top-M pages, the Number of referred categories $\rightarrow$ Presenting as Relevant Hashtags
User Study

• Purpose
To evaluate the effectiveness of presenting relevant hashtags and images (Deep information, Relevant information) with video clips

• Subject
6 college students (have experience to live in Kansai, Shikoku area)

• Data (travel video ×2)
“New discovery! Wondering (Fukuoka Prefecture)” * (2014.10.25 broadcast) outside
Domicile

“Good place (Shiga Prefecture)”** (2015.2.20 broadcast) Domicile

• Estimating categories from the tree structure that have referred pages more than N (N=5)
• Presenting top-M referred pages (M=4)
<table>
<thead>
<tr>
<th>Time</th>
<th>Extracted location information from CC</th>
<th>Extracted relevant hashtags from Wikipedia</th>
<th>Extracted images from Instagram (Top-8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0’05”</td>
<td>Fukuoka City</td>
<td>Kita-Kyushu City, Izuka City, Kurume City, Miyawaka City, Munakata City</td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>0’20”</td>
<td>Fukuoka Castle</td>
<td>Najima Castle, Bessyo Castle, Akizuki Castle, Mizuki, Takatori Castle</td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>0’31”</td>
<td></td>
<td></td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>0’46”</td>
<td></td>
<td></td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>0’58”</td>
<td>Hakata</td>
<td>Dazaifu, Munakata City, Neko Castle, Chikuuen-Kokubun temple, Genko</td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>1’09”</td>
<td>Fukuoka City</td>
<td>Osaka City, Kobe City, Yokohama City, Saitama City</td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>1’18”</td>
<td>Fukuoka Castle</td>
<td>Kinryu temple, Koun Shrine, A War of Tenjin, Mizukagami Shrine, Nissan gallery</td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>2’14”</td>
<td></td>
<td></td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>2’28”</td>
<td>Nakasu</td>
<td>Hakata river, Nakasu, Kinkuma, Higashi park, Mino Island</td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
<tr>
<td>3’35”</td>
<td>Ganso-Nagahamaya</td>
<td>Ichiran, Kaedama, Hakata Ippudo, Hakata Tenjin, Furyu</td>
<td><img src="image1.jpg" alt="Images" /> <img src="image2.jpg" alt="Images" /> <img src="image3.jpg" alt="Images" /> <img src="image4.jpg" alt="Images" /> <img src="image5.jpg" alt="Images" /> <img src="image6.jpg" alt="Images" /> <img src="image7.jpg" alt="Images" /> <img src="image8.jpg" alt="Images" /></td>
</tr>
</tbody>
</table>
“Good place (Siga Prefecture)”**  (Using part 0’00” ~ 2’30”)

<table>
<thead>
<tr>
<th>Time</th>
<th>Extracted location information from CC</th>
<th>Extracted relevant hashtags from Wikipedia</th>
<th>Extracted images from Instagram (Top-8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0’14”</td>
<td>Lake Biwa</td>
<td>Yogo Lake, Konoe basin, Torimaru peninsula, Tankai Lake, Nishi pond</td>
<td><img src="https://example.com" alt="Images" /></td>
</tr>
<tr>
<td>0’19”</td>
<td></td>
<td></td>
<td><img src="https://example.com" alt="Images" /></td>
</tr>
<tr>
<td>0’24”</td>
<td>Lake Yogo</td>
<td>Kinpun mountain, Sugaya hot-spring, Takeu Island, Nagahama port, Nagahama science park</td>
<td><img src="https://example.com" alt="Images" /></td>
</tr>
<tr>
<td>1’15”</td>
<td>Nagahama City</td>
<td>Oichigo-chan, Kitakonoe-resort, Nagahama police office, Nagahama-Dora company, Nagahama-Kita Biwa Lake firework festival</td>
<td><img src="https://example.com" alt="Images" /></td>
</tr>
</tbody>
</table>
Experiment

(a) location information synchronized with scenes (Deep information : text)
(b) Relevant hashtags synchronized with scenes (Deep + Relevant information : text)
(c) Images of location information synchronized with scenes (Deep information : images)
(d) Images of relevant hashtags synchronized with scenes (Deep + Relevant information : Images)
(a) location information synchronized with scenes

(Deep information: text)

Relevant words of “Hakata”
• Fukuoka city
• Fukuoka castle
• Nakasu
• Ganso-
• Nagahamaya
• Nakajima shop
Relevant hashtags synchronized with scenes

(Deep information + Relevant information: text)

Relevant hashyags of “Hakata”

# Hakata river
# Nakasu
# Kinkuma
# Higashi park
# Gokusho
(c) Images of location information synchronized with scenes (Deep information: images)
(d) Images of relevant hashtags synchronized with scenes

(Deep information + Relevant information : Images)
Questionnaire

- **Q1**: Could understand the video clips
- **Q2**: Felt spread your interests
- **Q3**: Write down relevant hashtags and images that are not related to the video clips
- **Q4**: Write down relevant hashtags and images that you are interested in
Q1: Could understand the video clips
(a) Deep information: text
(b) Relevant information: text
(c) Deep information: Image
(d) Deep information: Image

Q2: Felt spread your interests

- ✔️ The rating of images is higher than the text (relevant hashtags) with video clip as relevant information.
- ✔️ To present relevant information: the text is the most effective for understanding the video clips.
### Q3

- Subjects felt several unknown topics are not related to the video clips.
- General images about people, flower, or sky.

**Images that are not related to the video clips**

### Q4

- These images and relevant topics are not related to locations in the video clips.

**Relevant topics that they have interest**

- Need to present the relevance of locations in the video clips and their relevant hashtags.

<table>
<thead>
<tr>
<th>Q3</th>
<th>Wondering</th>
<th>Nissan gallery, Gasnso Nagahamaya, Nakajima Shoten</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good place</td>
<td>Nagahama police office, Nagahama, Oichigo-chan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q4</th>
<th>Wondering</th>
<th>Neko Castle, Ichiran, Ganso Nagahamaya, Mizuki</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good place</td>
<td>Sugaya hot spring, Oichigochan, Nagahama science park</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q4</th>
<th>Wondering</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good place</td>
<td></td>
</tr>
</tbody>
</table>
Conclusions and Future Work

◆ Proposed an image viewing system synchronized with video clips
  ● Supporting **Deep information** and **Relevant information** for video clips
    deep information: detail information based on location names in CC
    relevant information: information based on location names except CC

  ● User Study using travel videos (Wondering: 7scenes / Good place: 3scenes)
    ✔ Good point : Presenting of images and relevant topics is good for understanding
      the video clips and spreading interests about the video clips.
    ✔ Bad point    : Some images and relevant topics are not related to locations in the
      video clips

◆ Future work
  To extract relevant hashtags of semantic structures in Instagram
  To extend the system with other media: geographic map, web page, street view, etc...
THANK YOU FOR LISTENING