Video Hyperlinking (LNK)
TRECVid 2017

Maria Eskevich (CLARIN ERIC), Roeland Ordelman (University of Twente), Gareth J. F. Jones (Dublin City University), Benoit Huet (EURECOM)
Video Hyperlinking

• Establish links between an ‘anchor’ video segment and ‘target’ video segment(s) in a repository, based on topical relevance

• Or:
  • Content-based, segment level video recommendation
  • Multimodal search using video segment as a query
TRECVID LNK task

• Given a set of manually defined ‘anchors’ in a video collection, provide for each anchor a ranked list of ‘target’ videos that are ‘about’ the content in the anchor

• Target requirements:
  • Not in same video as the anchor
  • Targets should not overlap
  • Between 10-120 secs in length
Data set

- Blip10000 dataset: 14,838 semi-professional videos (3288h)
- Shot segmentation (2012)
- Visual features (AlexNet): 1000 visual concepts

- 25 anchors, manually selected, defined by video + start/endtime.
- To stimulate multimodal anchors: select segments in which the producer is using both audio and visuals to convey a message
Anchor Example

<anchor>
   <anchorId>anchor_126</anchorId>
   <itemId>item_7</itemId>
   <title>Fixing a bike</title>
   <description>I'm looking for other videos that show how to fix a bike</description>
   <characteristic>Both</characteristic>
   <refId>592496f643316916996d91f7</refId>
   <video>vid07338</video>
   <startTime>53.17</startTime>
   <endTime>53.34</endTime>
</anchor>
Relevance assessment
Crowd for Insight and Evaluation

3 stage approach

Manual Anchors Creation

1. Anchor Verification (2016)

Multimodal Anchors

Video-to-Video retrieval systems

2. Target Vetting

Automatic Targets Creation

Qrel

3. Video-to-Video Relevance Analysis

Detailed descriptions for relevance

Sponsored by:

BEELD EN GELUID
Target Vetting: Forced choice

1) Please make sure that you read all of the descriptions below and then choose the description that you find would be the best match with this video. Note: You might feel your best match is a good or even an excellent match. It is also possible that this question is difficult to answer because none of the choices is a particularly good match. If there is no particular good match, please make a choice the best you can.
   - $\{\text{DescriptionOption1}\}$
   - $\{\text{DescriptionOption2}\}$
   - $\{\text{DescriptionOption3}\}$
   - $\{\text{DescriptionOption4}\}$
   - $\{\text{DescriptionOption5}\}$

2) This question allows us to gather feedback on whether the question was difficult to answer.
   - I felt the video I chose was a relatively good match. The question was easy to answer.
   - I felt that the video I chose was the best possible in the list. It wasn't a particularly good match, and for this reason the question was not easy to answer.

3) Please write 2-3 sentences in the box below to explain your decisions.

<table>
<thead>
<tr>
<th>Case ID</th>
<th>Choice of target description</th>
<th>Feedback on decision making process</th>
<th>Relevance decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Correct</td>
<td>Positive</td>
<td>Relevant</td>
</tr>
<tr>
<td>2</td>
<td>Correct</td>
<td>Negative</td>
<td>Manual Check</td>
</tr>
<tr>
<td>3</td>
<td>Other</td>
<td>Positive</td>
<td>Non-Relevant</td>
</tr>
<tr>
<td>4</td>
<td>Other</td>
<td>Negative</td>
<td>Non-Relevant</td>
</tr>
</tbody>
</table>

1. Choose a description that fits “I'm looking for other videos that show how to fix a bike”
2. Was it difficult to find a fit?
<table>
<thead>
<tr>
<th>Case ID</th>
<th>MTurk worker’s choice of target description</th>
<th>MTurk worker’s feedback on decisionmaking process</th>
<th>Relevance decision</th>
<th>Number of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Correct</td>
<td>Positive</td>
<td>Relevant</td>
<td>547</td>
</tr>
<tr>
<td>2</td>
<td>Correct</td>
<td>Negative</td>
<td>Relevant</td>
<td>3849</td>
</tr>
<tr>
<td>3</td>
<td>Other</td>
<td>Positive</td>
<td>Non-relevant</td>
<td>1021</td>
</tr>
<tr>
<td>4</td>
<td>Other</td>
<td>Negative</td>
<td>Non-relevant</td>
<td>864</td>
</tr>
</tbody>
</table>
Participants

- Number of registrations: 16
- Number of finishers: 3 (12 runs)

- Minimizing risk in video hyperlinking (VIREO - City University of Hong Kong)

- IRISA at TRECVID2017: Beyond Crossmodal and Multimodal models for Video Hyperlinking (IRISA - IRISA, CNRS, INRIA & INSA Rennes, Univ. Rennes 1)

- Eurecom-Polito at TRECVID 2017: Hyperlinking task (Eurecom-Polito)
Evaluation metrics
Official metrics

• Precision at rank 5

• Mean Average *interpolated Segment* Precision (MAiSP)
  • User **effort**: the number of seconds they must spend auditioning content
  • User **satisfaction**: the number of seconds of new relevant content that they can watch starting from a suggested jump-in point.
Result examples

I'm looking for other videos that show how to fix a bike

I'm looking for other videos on martial arts

Practice…
- Makes money
- Makes perfect
- Makes you more confident