Crowdsourced Time-sync Video Tagging

Bin Wu

Erheng Zhong

Andrew Horner

Ben Tan

Qiang Yang

HKUST

HUawei
Video Tagging
On January 21st, 2008...

Labeler/user

Label/social tagging

- 2008
- president
- bush
- martin
- luther
- king
- protest
- terror
- terror
- rights
- war
- washington
- iraq
- speech

Database
On January 21st, 2008...
Users can ...

- Wait for buffering ...

2min Ad + 30min Buffering + O(video length)
Users can also ...

- Random guess ...

On January 21st, 2008...
Users can also ...

- Random guess ...
Users can also ...

- Random guess ...

*Al Jazeera Exclusive*

The events of September 11 are but a reaction to the continuous injustice and oppression being practiced against our sons in Palestine and Iraq.

2008
- president
- bush
- martin
- luther
- king
- protest
- terror
- rights
- war
- washington
- iraq
- speech
Users can also ...

- Random guess ...

2008
president
bush
martin
luther
king
protest
terror
rights
war
washington
iraq
speech

0’00’’  10’00’’  20’00’’  30’00’’
Users can also ...

- Random guess ...
Can't you do anything right?
The right thing
Ideally

On January 21st, 2008...

0’00”  10’00”  20’00”  30’00”

Terror rights  President Bush  Speech Martin Luther King  Iraq war

2008 president bush martin luther king protest terror rights war washington iraq speech
Ideally

**Time-sync video tagging**
Time-sync video tagging
Time-sync Video Tagging
Time-sync Video Tagging

- Done manually
- Only available for dramas/movies/news
Automatic Time-sync Video Tagging

- Subtitles/webcast texts

- Done manually
- Only available for dramas/movies/news/sports
- Only contain partial information about video contents.
Automatic Time-sync Video Tagging

- Image->Texts

- Requires high recognition accuracy.
- Difficult for intelligent interpretation, e.g., beautiful woman, wonderful scenery.
- Need to scan every single frame.
Automatic **Time-sync Video Tagging**

- Social network comments

- Limited to big events, e.g., sports games, election of the president.
New Data
– Time-sync Comments
Time-sync commented (TSC) videos
Time-sync commented (TSC) videos

- Invented by Niconico (Japan)
Time-sync commented (TSC) videos

- Adopted by major video websites, e.g., Tencent

10X more comments
Using Time-sync Comments
Comments -> independent documents

A BOMB! Are they going to fry the CLAAAAAWWW m home and make ebi tempura FT OF PEACE TO THE SHRIMP KING

w1 w2 w3
d1
d2
d3...

LDA
TF*IDF

w1 w2 w3
bento eat fry king lunch shrimp
Comments -> Non-exchangeable
Comments -> Non-exchangeable

What is it
SHRIMP
Eating the shrimps
Temporal and Personalized Topic Model (TPTM)
Alternating Optimization

Topic prior distribution of Shot $s_c$

Topic prior distribution of comment $c$

Temporal and personalized prior generation

Topic preference prior distribution of User $u_c$

Semantics dependency between shots

Semantics dependency between comments
Alternating Optimization
Alternating Optimization

- Topic prior distribution of Shot $s_c$
- Temporal and personalized prior generation
- Topic preference prior distribution of User $u_c$
- Topic prior distribution of comment $c$

Semantics dependency between shots
Semantics dependency between comments
Experimental Results
Data

- Chinese time-sync commented video. We have made it publicly available.

<table>
<thead>
<tr>
<th>Table 3: Summary of data.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td><strong>Data: Music Videos</strong></td>
</tr>
<tr>
<td>Videos</td>
</tr>
<tr>
<td>Users</td>
</tr>
<tr>
<td>Shots</td>
</tr>
<tr>
<td>Comments</td>
</tr>
<tr>
<td><strong>Data: Fun Videos</strong></td>
</tr>
<tr>
<td>Videos</td>
</tr>
<tr>
<td>Users</td>
</tr>
<tr>
<td>Shots</td>
</tr>
<tr>
<td>Comments</td>
</tr>
</tbody>
</table>
Held-out Log Likelihood

Crowd-sourcing with sequential dependency (AAAI2013)
User study – voting on the best results.

Table 7: Results of user study.

<table>
<thead>
<tr>
<th>Method</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPTM</td>
<td>267</td>
<td>275</td>
<td>240</td>
<td>782</td>
</tr>
<tr>
<td>LDA</td>
<td>207</td>
<td>192</td>
<td>209</td>
<td>608</td>
</tr>
<tr>
<td>Sembler</td>
<td>179</td>
<td>186</td>
<td>204</td>
<td>569</td>
</tr>
<tr>
<td>Shots</td>
<td>TPTM</td>
<td>Sembler</td>
<td>LDA</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>31.3°-45.9°</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>105.8°-116.0°</td>
<td>‘subtitles’</td>
<td>‘...’</td>
<td>‘...’</td>
<td></td>
</tr>
<tr>
<td>149.8°-153.4°</td>
<td>‘like’</td>
<td>‘...’</td>
<td>‘...’</td>
<td></td>
</tr>
<tr>
<td>161.0°-172.0°</td>
<td>‘mine’</td>
<td>‘...’</td>
<td>‘...’</td>
<td></td>
</tr>
<tr>
<td>178.4°-190.8°</td>
<td>‘love her’</td>
<td>‘...’</td>
<td>‘...’</td>
<td></td>
</tr>
<tr>
<td>235.7°-241.6°</td>
<td>‘you’</td>
<td>‘...’</td>
<td>‘...’</td>
<td></td>
</tr>
<tr>
<td>MAP@10</td>
<td>0.187</td>
<td>0.146</td>
<td>0.144</td>
<td></td>
</tr>
</tbody>
</table>
Conclusion
New Data
Time-sync Comments
Temporal
Personalized
Time-sync Video tagging
Looking for job
NY is so expensive
Guitar
Trumpet
Love music & sports
Leaving tomorrow and still in jetlag
Badminton
NY is so expensive
SJTU
HKUST
Hong Kong
Ph.D.
Hong Kong
Q&A
Thank you!