GRADUATE SCHOOL AND RESEARCH CENTER AT THE HEART OF THE DIGITAL SOCIETY

EURECOM
Sophia Antipolis

MeMAD
Methods for Managing Audiovisual Data

Using Fan-Made Content, Subtitles and Face Recognition for Character-Centric Video Summarization

Ismail Harrando, **Alison Reboud,** Jorma Laaksonen, Pasquale Lisena, Anja Virkkunen, Mikko Kurimo, Raphaël Troncy

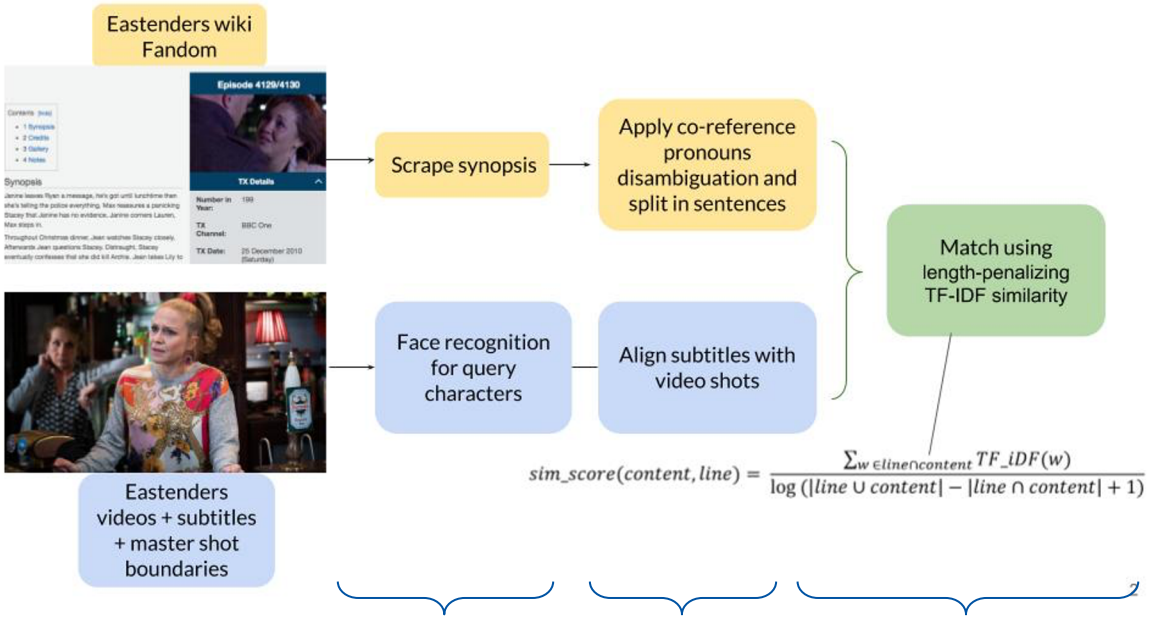
MeMAD Team - Takeaways

- **Contribution**: Showing that matching fan-written synopses with episodes transcripts is an efficient starting point for character-centric video summarization

- **Results:** 32 % in average (scoring over 3 variables + 5 qualitative questions)

- **Observation**: Except from the redundancy score, runs with more shots rarely yield better results

Approach: Leveraging on fan synopses



$$sim_score(content, line) = \frac{\sum_{w \in line \cap content} TF_iDF(w)}{\log\left(|line \cup content| - |line \cap content| + 1\right)}$$

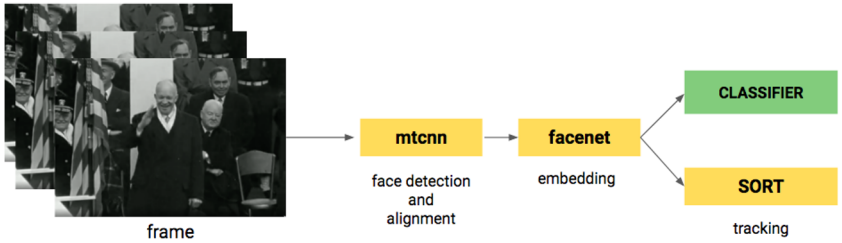
Corpus Creation

Alignment and Preprocessing

Matching and runs generation

Corpus Creation

- **Scraping synopses from Fandom EastEnders Wiki [1]**
  Hypothesis: every sentence represents an equally important event to be added to the final video summary
- **Shots Selection with Face Recognition for query characters :**
  - Crawl the web with name of actor + “EastEnders” (e.g.: Janine = Charlie Brooks)
  - Perform face detection and recognition [2] with one model for each of the 3 characters (Janine, Ryan, Stacey)



[1] https://eastenders.fandom.com/wiki/EastEndersWiki
[2] https://github.com/D2KLab/Face-Celebrity-Recognition

EURECOM
Sophia Antipolis

Alignment and preprocessing

- **Synopses:**
  - Concatenate text from weekdays episodes to longer “EastEnders Omnibus” episodes
  - Split into sentences and perform coreference resolution [3] to explicit character mention
- **Transcript:** From XML file to corresponding shot using timestamps
- **Synopses and Transcript:** lower case, stop words removal and lemmatization

[3] https://github.com/huggingface/neuralcoref

Matching and runs generation

- **Synopses sentence / shot transcript pairwise comparison by generating a similarity score**

$$sim_score(content, line) = \frac{\sum_{w \in line \cap content} TF_iDF(w)}{\log\left(|line \cup content| - |line \cap content| + 1\right)}$$

- **Length penalizing** to avoid matching with unrelated scenes
- **TFiDF** to avoid having only short sentences matching with the highest scores
- **Order shot by similarity:** only keep best match for each shot
- **Choose the N most matching shots in chronological order**

EURECOM
Sophia Antipolis

Average score for each run

TeamRun	Max # shots	Max Summary length	Score
MeMAD1	5 shots	150 sec	31%
MeMAD2	10 shots	300 sec	31%
MeMAD3	15 shots	450 sec	**35%**
MeMAD4	20 shots	600 sec	32%

Individual scores

Team_Run_Query	Tempo	Contextuality	Redundancy	Q1	Q2	Q3	Q4	Q5	final_score
MeMAD_1_Janine	**6**	4	5	No	No	No	No	Yes	29
MeMAD_2_Janine	5	5	6	No	No	No	No	Yes	28
MeMAD_3_Janine	5	5	6	No	No	No	No	Yes	28
MeMAD_4_Janine	5	5	7	No	No	No	No	Yes	27
MeMAD_1_Ryan	4	5	3	No	No	No	No	Yes	30
MeMAD_2_Ryan	5	5	3	No	No	No	No	Yes	31
MeMAD_3_Ryan	3	4	5	No	No	No	Yes	Yes	42
MeMAD_4_Ryan	2	3	5	No	No	No	Yes	Yes	40
MeMAD_1_Stacey	**6**	5	**2**	No	Yes	No	No	No	33
MeMAD_2_Stacey	**6**	5	**2**	No	Yes	No	No	No	33
MeMAD_3_Stacey	**6**	**6**	**2**	No	Yes	No	No	No	**44**
MeMAD_4_Stacey	4	5	4	No	Yes	No	No	No	29

Discussion

- For **Tempo** we decided to include the next shot if sentence was cut. Rather good results despite no additional attempt to specifically optimise for **Tempo, Contextuality and Redundancy**
- **Qualitative Questions**: not more than two (at best) out of five questions
- When generating the runs, having read the summaries but not having watched the videos
  1. We had the impression that the shots chosen were important
  2. We found out it is very challenging for humans too to choose between two potentially interesting moments without knowing beforehand the questions included in the evaluation set

Thank you for your attention!

Any questions?

MeMAD

Methods for Managing Audiovisual Data

INSTITUT CARNOT Télécom & Société numérique

EURECOM Sophia Antipolis