VideoForest: Interactive Visual Summarization of Video Streams Based on Danmu Data

Zhida Sun†, Mingfei Sun†, Nan Cao‡, Xiaojuan Ma†
†Department of Computer Science and Engineering, Hong Kong University of Science and Technology
‡College of Design and Innovation, Tongji University
Outline

• Motivation
• System
  • Architecture
  • System overview
  • Preprocessing
  • Analysis
  • Visualization
• Evaluation
• Conclusion
Video Summarization on Youtube

Title: Presidential Election 2016 LIVE | ABC News FULL BROADCAST
Description: Get the latest updates and state-by-state analysis of the 2016 US Presidential Election.
Screenshot: Screenshot of the video.
Summary: The long-awaited 2016 election day is finally here, and it could be a historic one for a number of different reasons. The United States could elect its first female president Hillary Clinton; Republicans could take back the White House with Donald Trump after two terms of Democratic control; there could be record voter turnout and the leading campaign election night parties will both be in New York.
Category: News & Politics
License: Standard YouTube License
Review: Song Master 1 day ago
At the end she says “this will test American Democracy”, perhaps. Electing Hillary Clinton would have ended Democracy as we know it. Congratulations to President Trump and to the preservation of the Nation.
Summarization of Video Streams

- Internal summarization techniques
  - Video, audio, linguistic features
- External summarization techniques
  - Audience’s preferences and interests

Video co-summarization [Chu et al. 2015]  User preferences based [Kannan et al. 2013]
• Summarization representation
  • Matrix form [Lu and Grauman 2013]
  • Illustration: Visual storyline & Imagehive

*Visual storylines [Chen et al., 2012]*

*Imagehive [Tan et al., 2011]*
Live Streaming Platforms

Watch and Interact with Messages and Hearts

Create a live conversation that takes place on Twitter.

2/24/2017

VideoForest, HCI Initiative, HKUST
Video Platforms with Danmu

Danmaku screen capture

Danmaku screen capture

2/24/2017

VideoForest, HCI Initiative, HKUST
Goal: Video + Danmu + Interaction $\rightarrow$ Summarization

Preprocessing
- Visual Feature Extraction
- Textual Feature Extraction
- Preview Assembly

Analysis
- Meta-Frame Analysis
- Session Analysis
- Cluster Analysis

Visualization
- Timeline View
- Session Summarization View
Video Demo
Processing

**Preprocessing**
- Visual Feature Extraction
- Textual Feature Extraction
- Preview Assembly

**Analysis**
- Meta-Frame Analysis
- Session Analysis
- Cluster Analysis

**Visualization**
- Timeline View
- Session Summarization View
Feature Extraction

- Visual feature extraction
  - Key frames extraction from compressed video

Hierarchical structure of a video sequence [Mendi and Bayrak 2010]
• Textual feature extraction
• syntactic and semantic analysis

Similarity
Sentiment
Topic
Function
Symbol
Pattern
Analysis

Preprocessing
- Visual Feature Extraction
- Textual Feature Extraction
- Preview Assembly

Analysis
- Meta-Frame Analysis
- Session Analysis
- Cluster Analysis

Visualization
- Timeline View
- Session Summarization View
Meta-frame, Session and Scene Cluster

Meta-frame
- Visual attributes: image content
- Textual attributes: audience’s reaction
Meta-frame, Session and Scene Cluster

Session

- Color palette of the entire session
- Synoptic linguistic attributes of danmu in the time range
Scene cluster
• Danmu commentary density
Interactions

Smart summarization  Scene preview  Highlighting
New session generating  View switching
Visual Metaphor and Encoding
Meta-frame aggregation

(a) First component of PCA vs. time (t/s)

(b) First component of PCA with connections between nodes
Circle Image Packing

(a)

(b)
Storyline threading

(a) [Diagram showing a complex network of nodes and connections]

(b) [Diagram showing a simplified network with labeled nodes and connections]

(c) [Diagram showing a storyline threading example]
Case study

• 205 seconds long
• 2052 danmu posts
• 61 I-frames
• 1804 P-frames
• 61 sessions
• 9 scenes
User observation

- Thirteen experts
- Live demo
- Interview & open discussion
  - Insights
  - Applications
Future Work and Conclusion

- Conclusion

- Future work
  - Long video stream scalability
  - Hierarchical aggregation
  - Advanced video analysis techniques
Thank you
http://hci.cse.ust.hk/
Zhida Sun
zhida.sun@connect.ust.hk