



1st PetaMedia SIG 2 Meeting "Processing"

Berlin, 8 October 2008

Minutes

Mission

The special interest group “processing” (SIG 2) deals with any research problem related to processing of (a) multimedia content data and (b) annotation data within a social and peer-to-peer network. Immediate goal of content and annotation processing is to obtain a useful description of multimedia content. Results of processing research will be of influence on or provide tools for the other PetaMedia research areas, such as indexing, social interaction, or content retrieval.

Agenda (revised)

10:45 - 11:00	Arrival & coffee
11:00 - 11:15	Welcome address, meeting overview
11:15 - 12:45	Short presentations by the participants: SIG 2 related work
13:00 - 13:40	Lunch at Café Campus
13:50 - 16:00	Discussion: objectives, challenges, research plans, possibilities of collaboration, mini-projects, future activities
16:00	Coffee & end of the meeting

Venue

Technische Universität Berlin
Communication Systems Group (TUB-NÜ)
Einsteinufer 17
10587 Berlin

Minutes

Presence List

Naeem Ramzan	QMUL (UK)
Ivan Ivanov	EPFL (CH)
Eric Bruno	University of Geneva (CH)
Alan Hanjalic	TUD (NL)
Nicolas Neubauer	TUB-NI (D)
Matthias Zappe	TUB-NI (D)
Thomas Sikora	TUB (D)
Engin Kurutepe	TUB (D)
Sebastian Schmiedeke	TUB (D)
Pascal Kelm	TUB (D)
Andreas Cobet	TUB (D)
Martin Haller	TUB (D)
Florian Kaiser	TUB (D)
Jiande Sun	TUB (D) / Shandong University (China)
Kai Clüver	TUB (D)

Welcome Address and Meeting Overview (T. Sikora)

suggestions for altering the agenda:

- discussion only until 16:00 h (T. Sikora)
- mission statement for the 1st PYTIM (A. Hanjalic)

Presentations of SIG2-Related Work

A. Hanjalic (TUD):

- scope of SIG 2 yet to be defined
- work on content processing at TUD:
 - *video segmentation*: shot-based vs. scene-based - search for theoretical limits of parsing performance
 - *audio content analysis* (e. g. keywords)
 - *MCA in open domains*: smart cameras, training - person-based content access - football highlights (arousal / corrective action / relevance feedback)

S. Schmiedeke, P. Kelm (TUB):

clustering of web video according to semantic similarities

discussion: only audiovisual feature analysis, tags and text currently not used - links to similar videos (e. g. from Youtube) are deemed unreliable, would have to be filtered

A. Cobet (TUB):

video content analysis via text analysis (incl. demo)

M. Haller, F. Kaiser (TUB):

video summarization:

- search for key frames - analysis of objects, backgrounds - looking through clips of key shots better than through fast forward
- A. Hanjalic: possible interface for tagging

E. Bruno (U Geneva):

topics at the Viper group:

- *content processing*: content analysis - multimodal feature construction (i. e. multi-modal fusion)
- *multimedia content indexing and retrieval*
- *multimedia content description* (goal)
- *semantic information propagation*: tag propagation - long-term learning from user interaction log

I. Ivanov (EPFL):

work at EPFL:

- *video segmentation*
- *tracking*: hybrid, i. e. region-based and feature-based object tracking - demo with static camera
- *events detection*

Summary of the presentations (T. Sikora)

additional mentioning of work at TUB Neural Information Processing group (NI):

- *machine learning for tag processing* (N. Neubauer)
- *image search engine with relevance feedback* (M. Zappe)

Results of the Discussion

Objectives of the discussion (T.Sikora):

- scope of SIG 2
- links to other SIGs
- possible collaborations and/or mini-projects

Purpose of SIG 2: *Content / annotation processing for content description*

Research topics:

- Development and optimization of low-level to high-level audio-visual descriptors
 - face descriptors? (high level)
 - text extraction
 - audio / speech descriptions (keywords)
 - image and video descriptions
 - event detection
- Classification of media content
(video + associated audio + images + text / tags)
- Summarization / segmentation of media
 - intra-frame / temporal segmentation
 - speech / audio segmentation
 - applications for media tagging?
- Relevance feedback strategies

- Relevance feedback, multi-modal processing and tagging
- Learning / clustering / classification strategies (theory)
novel algorithms for multi-modal classification problems
- Content ignorant / content aware methods for annotation processing?
tag filtering / de-noising / classification
- ***Holy Grail: Ground Truth Data***

Mini-project ideas (all partners pending)

1. *Media summarization for tagging*
TUD, TUB-NÜ, EPFL, inter-SIG
tagging cues through content analysis
(proposal only, to be presented at PYTIM by Alan)
2. *Youtube video classification using user tags / comments*
TUB-NÜ, U Geneva, TUD, TUB-NI, inter-SIG
classification of specific sub-genres like “travel”
Youtube is taken as a realistic model of an application that could be implemented as a P2P network.
3. *Key frame / key feature extraction from video*
TUB-NÜ, TU NI
both audio and video analysis
image / audio search over the key frames