

Summary of the Dagstuhl Seminar on

Contextual and social media understanding and usage

Seminar 08251, 15.06.08 - 20.06.08

Many years of multimedia research has achieved interesting results in the field of the semantic understanding of media content. Signal analysis, i.e., video, audio, image analysis has achieved some initial results to (semi-)automatically understand the semantics. However, the ability to really understand the media e.g. tell a sundown from a sunflower in an image is still difficult. It is becoming clear that “one way to resolve the semantic gap comes from sources outside of the image by integrating other sources of information” and that we should “exploit the synergy between the various media, including text and context information”. Flickr brought a new approach of communities, sharing, and tagging of photos. While tagging does not really solve the problem, it brings a new perspective of the situational usage of media, the co-presences of things and persons. Perhaps context and socio-sphere of media will bring us closer to surmounting the semantic gap. Indeed, we are now at the threshold of a new decade of contextual and social understanding of media content.

In this seminar, we aimed to see how content, context, and social community are influencing media understanding. Where do we find the sources to understand media and how does this affect and drive our methods? We examined the potentials and challenges, research directions, and looked into to different application domains ranging from personal media and news to biomedicine, and robotics. The seminar brought the research community (from academia, labs and industry) together to help make progress towards the vision that “capturing, storing, finding, and using digital media should be an everyday occurrence in our computing environment.” We had stimulating discussions on the role of context and social aspects for tomorrow’s multimedia content understanding and media usage.

The Structure of the Seminar

This seminar was structured to examine the understanding of images, video, and audio using a rich array of contextual information - social, spatial, and temporal. After the first two days of presentations of position statements by the individual participants, there was an open brainstorming session organized to find the significant issues to be discussed. The brainstorming session led to the setting up of the following working groups to examine the issues in greater details:

- Fundamentals: Rainer Lienhart, Mohan Kankanhalli, Dinh Phung, B Prabhakaran
- Social Media: Brett Adams, Pablo Cesar, Daniel Gatica-Perez, Lynn Wilcox, Peter Hart, Andreas Girgensohn, Svetha Venkatesh
- Evaluation, Benchmarking and Datasets: Stephane Marchand-Maillet, Alexander Hauptmann, Gareth Jones, Cees Snoek, Mohan Kankanhalli
- Modeling and Context: Daniela Nicklas, Lynda Hardman, Peter Hart, Andruid Kerne, Alejandro Jaimes, Roger Zimmermann, Trevor Darrell
- Challenges and issues in user experience: Ansgar Scherp, Andreas Henrich, Masashi Inoue, Frank Nack, Daniela Nicklas, Sabine Thieme

Upon some detailed discussions, it was decided that the modeling and context group issues were best discussed in conjunction with the rest of the working groups. As a result, the members of that group joined the rest of the groups.

These groups did their deliberations separately and came up with their findings and recommendations. Finally, there was an overall plenary session in which the findings of all of the groups were presented:

- Group 1: Rainer Lienhart spoke about formalizing the notion of context especially with a rigorous and mathematical foundation. Also, many basic problems in bridging the signal to symbol gap were discussed.
- Group 2: Brett Adams presented a call for funding presented a call for papers, “How will social media change the way we communicate and socialize? What is the pathology of social media driven communities”.
- Group 3: Modeling and Context: They stressed that context is the **implicit information** needed to augment the **explicit information** (i.e., information supplied by the user) that is needed by the system in order to help the user achieve her goal. Some questions raised were:
 - What are the components of context-driven information discovery?
 - How do you model them?
 - How do you represent them to people?
 - How do you operate on them (semi-)automatically?
 - How do you give users' control?
 - How do you evaluate systems that do this?
- Group 4: Challenges and issues in user experience were presented in detail by Ansgar. The following exciting topics were offered for further research:
 1. Explore the use of new form of multiple media and interactive

- environments to facilitate learning and training e.g., gaming
 - metaphors and the mixed use of physical + cyber space
 2. Explore the use of existing multiple media (e.g., video) in combination

- with new technologies (e.g. smart home devices & sensors)
 - for new applications (e.g., assisted living)
 3. Develop systems that can automatically detect the user's context

- For example, working vs. entertaining and properly act upon it (e.g. helping users to switch to a different context)
- Group 5: Stephane presented the need for common datasets, the accessibility of common metadata for these datasets, common queries and benchmarking mechanisms. He suggested that leading conferences in the area can serve as valuable forums for propagating these ideas like in the case of Video Olympics at CIVR. ACM SIGMM can also be roped into many of these activities. There can be call for data-sets along with other calls.

The Outcomes of the Seminar

The following issues, plans and action items emerged at the end of the workshop:

1. Establishment of a common repository for multimedia data and benchmarking
2. Organization of workshops
 - a. Along the lines of social media, and aim at a short term collaboration on an initial short paper on the ideas.
 - b. Organization of workshop in modeling and context, with emphasis on context driven information recovery and discovery.
 - c. Challenges and issues in user experience

Susanne Boll (University of Oldenburg)

Svetha Venkatesh (Curtin University of Technology)

Mohan Kankanhalli (National University of Singapore)

Gopal Pingali (IBM Research)