08251 Abstracts Collection
Contextual and Social Media Understanding and Usage
— Dagstuhl Seminar —

Susanne Boll[1], Mohan Kankanhalli[2], Gopal Pingali[3] and Svetha Venkatesh[4]

[1] Universität Oldenburg, D
susanne.boll@informatik.uni-oldenburg.de
[2] National University of Singapore, SGP
mohan@comp.nus.edu.sg
[3] IBM TJ Watson Research Center - Hawthorne, USA
[4] Curtin Univ. of Tech. - Perth, AU
svetha@cs.curtin.edu.au

Abstract. From 15.06. to 20.06.2008, the Dagstuhl Seminar 08251 "Contextual and Social Media Understanding and Usage" was held in Schloss Dagstuhl – Leibniz Center for Informatics. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.

Keywords. Mutable state, Program logics, Semantics, Type systems

08251 Summary – Contextual and Social Media Understanding and Usage

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.

Keywords: Mutable state, Program logics, Semantics, Type systems

Joint work of: Boll, Susanne; Venkatesh, Svetha; Kankanhalli, Mohan; Pingali, Gopal

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2009/2024

Sensing and Using Community

Brett Adams (Curtin Univ. of Tech. - Perth, AU)

This position statement concerns the sensing and use of community from offline and online sources, and in particular, the blogosphere. The blog has matured into a new media genre in its own right and has a number of properties that make it an interesting target for investigation:

Blogs come in various sub-genres that subordinate text to other media, such as images or video; are dynamically inter-related via a range of more or less implicitly typed relations (e.g. blogrolls, cites, critiques); and occupy various positions on the spectrum between mass media and personal communications, typically providing public evidence of social interactions via the ability to leave comments. These properties present different kinds of context that must be incorporated, and corresponding research challenges in (a) truly multimedia document indexing, (b) organic link-content analysis, and (c) inference of underlying social realities (entity resolution, role classification, and sub-community discovery).

We will present some results of using commenting behaviour together with topic models to characterize the groups of readers, together with the best representative of each group, that adhere to a blog.

Keywords: Blog, topic model, affinity propagation, comment, exemplar

Social Media Working Group

Brett Adams (Curtin Univ. of Tech. - Perth, AU)

Recent years have seen the socio-technological phenomenon of Şsocial mediaŤ enter mainstream, public consciousness. Popular Web-based services such as YouTube and Facebook serve not only as media warehouses, but forums in which people express and entertain themselves and ultimately come together as communities. Intuitively we see that social media has a massive impact on underlying social structures /realities, but we are in the infancy of understanding how these communities are formed, grow, and change as a function of the media-types and interactions methods they enable. These forums have already been successful in commercial terms. Research has begun in leveraging community wisdom to augment understanding of media (e.g. through collaborative filtering) and remains an area of massive potential.

However, there has been little quantitative work analyzing the effect of the use of various media types and interaction paradigms in the dynamic evolution of social structures themselves, either within or across the existing forums.

Joint work of: Adams, Brett; Gatica-Perez, Daniel; Wilcox, Lynn; Cesar, Pablo

A Robust Estimator of Image Thumbnail and Video Histogram Representation

Cheng Cai (Universität Oldenburg, DE)

For browsing and retrieval system, images are represented by thumbnails and video shots are represented by content representations. In order to achieve better visual quality and retrieval performance, the representation estimator is expected to be accurate and robust. From the statistical perspective, representation extraction can be treated as central value estimation. In this paper, we propose an adaptive alpha-trimmed average estimator based on Gaussian distribution hypothesis test (AATA-GDHT). For a set of values, this estimator extracts the representation by trimming extreme values and then averaging the rest. The criterion to distinguish between extreme values and useful data is derived from Gaussian distribution hypothesis test on the basis of global statics. Experimental results from standard images and videos show that our proposed scheme outperforms traditional methods.For the browsing and retrieval system, images are represented by thumbnails and video shots are represented by histogram representations. In order to achieve better visual quality and retrieval performance, the representation estimator is expected to be accurate and robust. From the statistical perspective, representation extraction can be treated as central value estimation. In this paper, we propose an adaptive alpha-trimmed average estimator based on the Gaussian distribution hypothesis test. For a set of values, this estimator extracts the representation by trimming extreme values and then averaging the rest. The criterion adopted to distinguish between extreme values and useful data is derived from the Gaussian distribution hypothesis test on the basis of global statics. Experimental results from standard images and videos show that our proposed scheme outperforms traditional methods.

Keywords: Algorithms, Experimentation, Performance

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/2018

Challenges for Next-Generation Video Sharing Systems: Content Modeling, Social Network, and Contextual Setting

Pablo S. Cesar (CWI - Amsterdam, NL)

The purpose of this position paper is to identify key factors for next-generation video sharing systems.

As a discussion starter we highlight three topics: content modeling, the implications and benefits of social networks, and the importance of the contexual setting to distribute and control the video material. The position paper will report on the preliminary work and the results of our group in these directions.

Keywords: Video Sharing, Content Modeling, Social Networks, Contextual Setting

Joint work of: Cesar, Pablo; Bulterman, Dick; Jansen, Jack

Analyzing Flickr Groups

Daniel Gatica-Perez (IDIAP Research Institute - Martigny, CH)

Flickr users follow a large variety of social motivations and themes.
    Flickr Groups is an available system feature that is largely used by people to facilitate the explicit definition of communities sharing common interests, which translates into large amounts of content about specific subjects.
    Little attention, however, has been paid to this social aspect of Flickr, regarding both the understanding of user-to-group content sharing behavior and the development of effective tools to find relevant groups. These are the two open issues that I will discuss in the talk.
    The bottom line is that a clearer (statistical) understanding of the motivations and uses of social media systems might pay off towards identifying and solving new research problems and develop useful applications.

Keywords: Social media, Flickr, photo sharing communities, topic models

GLOW: Organizing and Classifying Photos in a Graph-based Workspace

Andreas Girgensohn (FX Palo Alto Laboratory, US)

We support the organization of personal and shared photo collections by having the system place photos in a workspace by similarity measures based on visual, temporal, and geographic features. Our visual similarity measure uses a classifier trained on tagged photos by a community of users. Users may group photos in stacks that in turn attract neighborhoods of similar photos. We designed an adaptive interactive visual workspace that supports users in browsing and organizing photo collections.

Joint work of: Girgensohn, Andreas; Shipman, Frank; Wilcox, Lynn; Cooper, Matthew

On integration of feature-based analysis and social indexing for improved multimedia content access

Alan Hanjalic (Delft University of Technology, NL)

Many years of research and exciting achievements in the multimedia information retrieval (MIR) field have made the field mature enough to enter a new development phase - the phase in which the technology enabling automated indexing and relevance ranking of multimedia content is steered towards adoption in practical retrieval scenarios and applications. Although the retrieval performance of these machine-level semantic inference solutions will always involve trade-offs in precision vs. recall, many of such solutions are already providing value for users, and many significant hurdles in development of MIR technology have been passed ý[1].

For practical deployment, however, the MIR reliability needs to be further improved. For example, in the case of consumer-oriented MIR solutions integrated in e.g. a personal video recorder, a mobile video retrieval system, a music search framework, or a web search engine, the service they provide in terms of the paradigm content I like, anytime and anyplace will need to be at least as reliable as the button turning a TV set or a mobile device on and off.

An increase in MIR reliability could be achieved in several ways. For instance, one could choose to rely on machine-level semantic inference mechanisms only and to invest more effort to better understand the processes underlying the development of these mechanisms and so to help them forecast, prevent or correct possible inference errors. However, as the average performance of state-of-the-art semantic inference solutions is still too low, insufficiently explainable (predictable), and has improved only slowly over the last couple of years ý[2], machine-level semantic inference solutions are not likely to lead to the desired level of MIR reliability within a reasonable time period.

To make a significant step forward, an existing isolated semantic inference mechanism, e.g. a semantic concept detector, could instead be integrated into a network characterized by distributed and collaborative intelligence ý[3]. A typical example of such a network would be a social network of users interacting with their "local" automated indexing and retrieval devices, individually or collaboratively tagging the content, and seamlessly recommending the content to each other by means of e.g. collaborative filtering.

There are (at least) three benefits of a network-level MIR approach described above. First, an automated semantic inference mechanism can afford to come up with not more than the "best educated gues" regarding the content to be retrieved, which would allow direct practical deployment of currently available (still suboptimal) semantic inference schemes. This initial "guess" can then be improved for future retrieval requests by complementing the signal processing and machine learning intelligence of a semantic inference mechanism by corrective actions of the user and the information stemming from tags and collaborative filtering. Besides, the corrective action of the users can also help the inference

mechanisms to fulfill the individual and context-aware retrieval requests at each network node. Second, local video indexing results obtained with a high confidence can be used to improve the quality of tags, e.g. by de-noising the tag sets and improving the tag consistency, and to enhance and verify collaborative tag/content recommendation processes. Finally, local semantic inference mechanisms taking place at various nodes of a network can learn from each other and benefit from individual user actions and available local context information to improve inference quality at all nodes. In this way, by improving the MIR reliability through a synergetic integration of complementary resources in a user/device network, and by unifying the process of reliably filtering the relevant content to be retrieved with the process of discovering and updating user profile and use context, a considerable improvement of the overall Quality of Experience (QoE) in accessing the multimedia content of interest can be achieved ý[4].

I see the network-level MIR approach described above as a key technological approach marking the research in the field of multimedia information retrieval in years to come. This talk will elaborate on the possibilities for a realization of this approach.

References

[1] A. Hanjalic, R. Lienhart, W.-Y. Ma, J.R. Smith (Eds.): Special Issue on Advances in Multimedia Information Retrieval, Proceedings of the IEEE, Vol.96, No.4, April 2008

[2] A. Smeaton, P. Over: TREC Video Retrieval Evaluation Online Proceedings 2001-2007, (http://www-nlpir.nist.gov/projects/tvpubs/tv.pubs.org.html)

[3] S. Boll: MultiTube - Where Multimedia and Web 2.0 Could Meet, IEEE MultiMedia, January-March 2007

[4] A. Hanjalic: Content You Like Anytime and Anyplace: Multimedia Research for New TV Concepts, Keynote at the Pacific Rim Conference on Multimedia, Hong Kong, December 2007 (http://www-ict.ewi.tudelft.nl/~alan/)

Keywords: Multimedia content analysis, tagging, collaborative filtering

Events for Contextualising media assets

Lynda Hardman (CWI - Amsterdam, NL)

An event can be used as a means of contextualising information that would otherwise be "standalone". For example, people, locations and times can be connected together through an event description.

In human communication, events are useful for describing, for example,

- nuclear reactions between interacting particles;
- births, marriages, deaths;
- politically significant meetings such as G8 summits;
- football matches such as Germany vs. Portugal;
- computer science seminars.

The notions of "who", "where" and "when" can be linked to existing vocabularies, such as

* ULAN <http://www.getty.edu/research/conducting_research/vocabularies/ulan/,> dbpedia http://dbpedia.org/ or FOAF profiles http://www.foaf-project.org/ for who;

* GeoNames <http://www.geonames.org/> for "where";

* OWL time <http://www.w3.org/TR/owl-time/> for "when".

In order to work with events we need an event-based interface. Browsing among the separate "dimensions" (who, where, when) is one way, but browsing relations between events is potentially more informative.

In creating a model of events we need to take a number of issues into account, including:

* concurrent events in multiple places, e.g.

[LiveAid http://en.wikipedia.org/wiki/Live_Aid]

[Live 8 http://en.wikipedia.org/wiki/Live_8],

* frequently occurring relationships between events, for example, causality,

* iterations of similar events, e.g. the G8 summits

[http://en.wikipedia.org/wiki/31st_G8_summit].

We are currently investigating cultural heritage and news as examples of relevant domains which can be used to esablish common factors that can be usefully incorporated into an event model.

[1] Frank Nack Capturing experience: a matter of contextualising events Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence, Berkeley, California pp 53-64 http://doi.acm.org/10.1145/982484.982492

[2] Ryan Shaw Event Model, work in progress

[3] Utz Westermann, Ramesh Jain Toward a Common Event Model for Multimedia Applications IEEE MultiMedia 14(1) (January 2007) pp 19-29

http://dx.doi.org/10.1109/MMUL.2007.23

http://richard.cyganiak.de/2007/10/lod/

Keywords: Events

Different views of online video data

Alexander Hauptmann (Carnegie Mellon University - Pittsburgh, US)

Online video data (e.g. youtube) is growing dramatically, yet its use as annotated data for analysis is just getting to be appreciated. We will present some evaluations of the stability of online video data as a resource, as well as a preliminary analysis of user perspectives through comments on particular video clips.

Content, Context, Metadata ... - What makes a document relevant to a certain user in a certain situation?

Andreas Henrich (Universität Bamberg, DE)

In my position statement I try to address the different possible approaches to exploit and combine the various sources which can be considered when computing a ranking or a result presentation for a query. I will discuss the pros and cons of weighting approaches, interleaving approaches, visualization approaches ...

Difficulties in Image Retrieval

Masashi Inoue (National Institute of Informatics - Tokyo, JP)

The semantic gap is often regarded as a major problem in the field of image retrieval research.

In this paper, I will show that there are other important topics that should be addressed for improving the image retrieval utility.

Among them, the exploitation of limited information and motivating the use of images are considered to be central to the development of image retrieval.

Keywords: Image retrieval, utility, user, scarcity, insufficiency

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2009/2019

How Local is Local? Culture and social context in behaviour analysis

Alejandro Jaimes-Larrarte (Telefónica Research, Madrid, ES)

In this position statement, I will discuss the problems and benefits of the huge amounts of information we generate and collect digitally. In particular, I will discuss why we need to work on new ways to leverage digital data to build user models that lead to personalized products and services. I will argue that human-centered approaches are needed to analyze data and build models that are applicable, and discuss the role of culture and social context in analyzing and modeling human behaviour. I will discuss some applications and identify open issues and future research directions.

Keywords: Culture, context, human behaviour analysis

Location Semantics

Mohan S. Kankanhalli (National University of Singapore, SG)

We will discuss a (very preliminary) multi-modal framework for inferring location semantics from web content.

We focus on the problem of automatic inference of whether the given web content talks about a particular place and further, which aspects of a place it talks about. The approach is designed for a scenario where the ip-addresses or domain names might not be very useful as content is hosted world-wide. Further, we would like to address situations where informal content makes many implicit assumptions about the place and important key-words could hence be missing in the content.

Joint work of: Katti, Harish; Chua, Tat-Seng; Kankanhalli, Mohan S.

Computational discovery of practical wisdom: narrative, community and meaning

Aisling Kelliher (ASU - Tempe, US)

Today's diverse, fast-paced and innovative collaborative environments require knowledge workers that can quickly locate resources, communicate effectively and purposefully reflect on their activities and contributions. The collective articulation, interpretation and understanding of successful workplace processes is therefore vital for fostering creativity and driving innovation. We propose the use of community-centric computational narratives for capturing, analyzing and sharing the everyday experiential 'know-how' and insights accumulated by researchers and workers within the knowledge economy. We will present some preliminary prototypes of dynamic renderings of a diverse workplace community that combines resource selection, message semantics and design aesthetics.

A Mixed-Initiative Philosophy for Human Centered Contextual Media Systems

Andruid Kerne (Texas A &M University, Interface Ecology Lab, US)

Contextual media systems promise to model what the human participant is doing and what s/he is experiencing, and to enhance operations based on this participant model. However, the capture problem is inevitable. No matter how good the model, we can never eliminate the semantic gaps that will sometimes arise between the inferences made on captured data and the human's actual thoughts, feelings, and intentions. Thus, this paper argues for a new mixed-initiative systems design philosophy, which emphasizes making representations of contextual models available to human participants, in the midst of their media discovery tasks, so that, when they choose to, they can adjust the models to express their actual intentions. The representations must employ fluid interfaces, in order to enable participants to adjust the human-in-the-loop models with a minimum of attention and effort. We draw from lessons learned over the past decade of work on the project currently known as combinFormation, the mixed-initiative system that represents information collections as compositions to support information discovery.

Digging out implicit semantics from user interaction

Stéphane Marchand-Maillet (CUI - Geneva, CH)

User interaction may take many forms in multimedia systems. Current systems mainly waste this implicit and natural source of semantic knowledge and rather create tedious and unnatural interaction protocols.

We advocate for a complete integration of natural interaction protocols and semantic knowledge capture, mainly thru mining interaction sessions. We assert that users possess the ability to quickly examine and summarise these documents, even subconsciously. Examples include specifying relevance between a query and results, rating preferences in film databases, purchasing items from online retailers, and even simply browsing web sites. Data from these interactions, captured and stored in log files, can be interpreted to have semantic meaning, which proves indispensable when used in a collaborative setting where users share similar preferences and goals.

Keywords: Multimedia, social media, interaction, crowdourcing

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2009/2020

User context is group context?

Frank Nack (University of Amsterdam, NL)

The challenge of media research will be finding a balance between the exploitation of social behaviour and social networks in and for media usage and addressing the needs of the individual user. The context model differs in complexity for both aspects. The aim of this statement is to provide some thoughts so that we finally do not add the 'context gap' to the 'semantic gap'.

Impact of Visual Context Representation and Control on Learning/Teaching of Physical Activities in Tele-Presence Environments

Klara Nahrstedt (Univ. of Illinois - Urbana, US)

Shared visual context plays a vital role in video-mediated remote learning tasks. However, it remains unclear how to provide such visual context in order to best facilitate learning. We will claim that tele-immersive spaces with multi-view 3D cameras as the most effective video-mediated learning environments due to various recent experiments and results confirming this claim.

Keywords: Semantic visual context, tele-immersive environments, video-mediated remote learning, physical activities

Full Paper:
http://cairo.cs.uiuc.edu/teleimmersion

Context-as-Clustering: Unsupervised Extraction of Social Context from Sensor Data

Dinh Phung (Curtin Univ. of Tech. - Perth, AU)

Automatic extraction of contextual information provides foundational units for context-aware applications and socially-situated media understanding and usage. But what sorts of contexts and how to extract them remain a key challenge. In this talk, we will present a framework where contexts can be viewed as an unsupervised clustering process. In particular we will comment on the appropriateness of clustering approaches including nonparametric and Bayesian clustering methods. Three applications are presented to demonstrate the framework: extraction of socially significant locations, extraction of social rhythms, and extraction of computable social patterns using Bayesian clustering.

Keywords: Social context, nonparametric, Bayesian clustering, latent Dirichlet allocation

Joint work of: Phung, Dinh; Adams, Brett; Venkatesh, Svetha

Usage Driven Understanding of Personal Photo Collections and Implications for Retrieval

Philipp Sandhaus (OFFIS - Oldenburg, DE)

Semantic understanding of digital media is usually driven by the methods available for content- and context-based media analysis and thus the means of accessing and retrieving media are driven and often limited by the kind of metadata and semantics extracted. On the other hand some research has been done at understanding the ways people interact with their photos, how they organize and access them, but the findings of these investigations have had little impact on pure media analysis and therefore the means of retrieving media content. Thus, the goal of my research is to understand personal photo collections from a user's perspective and find implications for photo retrieval and media analysis.

Keywords: Photo understanding, user modelling, personal photo collections, image retrieval

Linking the Semantics Ecosystem with Semantics Derivation Rules for Multimedia Content

Ansgar Scherp (Universität Koblenz-Landau, DE)

Multimedia content exhibits multiple semantics that is influenced by different factors like time, contextual use, and personal background.

With the semantics ecosystem, we find an elegant and high-level description of the different factors that influence the semantics of multimedia content. On the other hand, semantics derivation rules are a concrete means to extract and to derive semantics of multimedia content while authoring it.

These rules are directly applicable in concrete applications and domains.

Thus, there is a gap between the high-level ecosystem and the concrete semantics derivation rules. In this position paper, we propose the use of an ontology-based description of events to combine the high-level description of the semantics ecosystem with the concrete method of semantics derivation for page-based multimedia presentations.

Keywords: Multimedia Semantics, Semantics Ecosystem, Semantics Derivation, Event Ontology

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2009/2021

See also: This article bases mainly on: A. Scherp and R. Jain, Towards an ecosystem for semantics, Workshop on multimedia information retrieval on The many faces of multimedia semantics; Augsburg, Bavaria, Germany, 2007, ACM Press. S. Boll, P. Sandhaus, A. Scherp, and U. Westermann, Semantics, content, and structure of many for the creation of personal photo albums; Augsburg, Bavaria, Germany, Proceedings of the 15th international conference on Multimedia, 2007, ACM Press.

Interaction and User Experiences with Multimedia Technologies-Challenges and Future Topics

Ansgar Scherp (Universität Koblenz-Landau, DE)

In this paper, we investigate future topics and challenges of interaction and user experience in multimedia. We bring together different perspectives from overlapping fields of research such as multimedia, human-computer interaction, information retrieval, networked multimedia, and creative arts.

Based on potential intersections, we define three application domains to be investigated further, as they create high demand and good prospect for long-lasting developments in the future. These application domains are: media working environments, media enter-/edutainment, and social media engagement. Each application domain is analyzed along five dimensions, namely: information quality, presentation quality, ambience, interactivity, and user expectations. Based on this analysis, we identify the most pressing research questions and key challenges for each area. Finally, we advocate a user-centered approach to tackle these challenges and questions in order to develop relevant multimedia applications that best meet the users' expectations.

Keywords: Interaction, User Experience, Multimedia, Humancomputer Interaction, Information Retrieval, Networked Multimedia, Creative Arts

Joint work of: Ansgar Scherp; Frank Nack; Klara Nahrstedt; Masashi In-oue; Andreas Girgensohn; Andreas Henrich; Philipp Sandhaus; Sabine Thieme; Michelle Zhou

See also: Results of the working group on Interaction and User Experience have been published at: Ansgar Scherp, Frank Nack, Klara Nahrstedt, Masashi Inoue, Andreas Girgensohn, Andreas Henrich, Philipp Sandhaus, Sabine Thieme, Michelle Zhou: Interaction and User Experiences with Multimedia Technologies-Challenges and Future Topics, HCC'08, October 31, 2008, Vancouver, British Columbia, Canada, ACM Press.

Learning to exploit social-tagged multimedia

Cees Snoek (University of Amsterdam, NL)

The fundamental problem limiting current multimedia information retrieval methods based on supervised learning is: the lack of a large and diverse set of manually labeled visual examples to adequately model diversity in appearance. Incremental improvements in feature extraction and machine learning will have a positive impact on automatic multimedia understanding, but the decisive leap forward will stem from an increase in the amount of training data of at least two orders of magnitude beyond current practice, for example by exploiting freely available social tagged media.

Social and Context Driven browsing for personal multimedia

Svetha Venkatesh (Curtin Univ. of Tech. - Perth, AU)

We explore systems that harvest readily available context information (GPS, bluetooth, user) when multiple media such as photos, video, audio, or activity streams (eg. from Twitter, Facebook, etc.) are acquired through cell phones and uses it for multimedia navigation, search and sharing. Separate context streams can be recorded on the phone, and related to media captured (on the phone or other devices) based on recorded time. We present a framework that integrates and unifies all time-based media and uses contextual meta-data to construct novel, rich browsers, facilitating the sharing of both data and meta-data across users. This includes location, co-presence and activity, which are used in new ways for navigation and search in the aggregated media. We explore issues and problems in this growing area.

Keywords: Social, context, browser, media

Context-Sensitive Information Seeking and Analysis: What Context?

Michelle Zhou (IBM TJ Watson Research Center - Hawthorne, US)

Imagine the next generation of information portals, where users are able to obtain information through an intelligent multimodal conversation that is tailored to the tasks you are performing, customized to your personal preferences, and adapted to your context and interaction devices.

To realize this vision, we are building a suite of intelligent user interaction technologies, which aims to bridge the gap between what users want and what today's systems can provide.

On the one hand, our work can automatically translate high-level user information requests ("finding cheap homes along Hudson") into low-level system-understandable data operations in context. On the other hand, our work can automatically synthesize multimedia presentation of system-retrieved data, which is tailored to a user's interaction context (e.g., user tasks and preferences).

During the above user input interpretation and output generation process, we also dynamically track and model a user's context, including both user explicit interaction (e.g., a query) and implicit interaction (e.g., clickthrough). Such context is then used to help better understand and satisfy user requests (e.g., more accurate and tailored information retrieval and presentation).

Keywords: Intelligent User Interaction, User Context Modeling, Implicit User Context, Explicit User Context

Full Paper:
 http://www.research.ibm.com/RIA

Geospatial Video Search through Viewable Scene Modeling

Roger Zimmermann (National University of Singapore, SG)

Video sensors are becoming ubiquitous and the volume of captured video material is very large. Therefore, tools for searching video databases are indispensable. Current techniques that extract features purely based on the visual signals of a video are struggling to achieve good results when the content domain is not constrained. By considering video related meta-information, more relevant and precisely delimited search results can be obtained. We propose a novel approach for querying videos based on the notion that the geographical location of the captured scene in addition to the location of a camera can provide valuable information and may be used as a search criterion in many applications. To explore the feasibility of this approach we have investigated an estimation model of the viewable area of a scene for indexing and searching. We have further implemented a prototype system that is being used for testing. Among our objectives is to stimulate a discussion of these topics in the research community as information

fusion of different georeferenced data sources is becoming increasingly important. The initial results we obtained are promising and illustrate the feasibility of the proposed approach.