Media on the Web, in Post-Production and Broadcasting

The Practitioner Day of the ACM 2009 International Conference on Image and Video Retrieval [*]

Sébastien Marcel
Idiap research institute
rue Marconi 19
1920 Martigny, Switzerland
marcel@idiap.ch

Roelof van Zwol
Yahoo! Research
Avda. Diagonal 177, planta 8
08018, Barcelona, Spain
roelof@yahoo-inc.com

Ricardo Baeza-Yates
Yahoo! Research
Avda. Diagonal 177, planta 8
08018, Barcelona, Spain
ricardo.baeza@barcelonamedia.org

Oliver Heckmann
Google Switzerland GmbH
Brandschenkestr 110, 8002
Zürich, Switzerland
oliverh@google.com

Jan Erik Solem
Polar Rose AB
Anckargripsgatan 3, 211 19
Malmö, Sweden
janerik@polarrose.com

Johan Oomen
Nederlands Instituut voor Beeld en Geluid
Sumatralaan 45, 1200 BB
Hilversum, The Netherlands
joomen@beeldengeluid.nl

Hans van Gageldonk
Philips Research Laboratories
5656 Eindhoven, The Netherlands
hans.van.gageldonk@philips.com

Jean-Pierre Gehrig
Cinetis SA *Centre du Parc, Rue Marconi 19, 1920 Martigny, Switzerland*
jp.gehrig@cinetis.ch

Xavier Vives
Corporació Catalana de Mitjans Audiovisuals
Cal. G. Fàbregas 81, planta 2a, 08950 Barcelona, Spain
xvives.t@ccma.cat

Baris Sumengen
Like.com
777 Mariner's Island Blvd., Suite 510, San Mateo, CA 94404, USA
baris@like.com

Categories and Subject Descriptors

H.3.1 [**Information Systems**]: Information Storage and Retrieval —*Content Analysis and Indexing*

General Terms

Practitioners, media search, image and video search, face recognition, web 2.0, post-production, broadcasting

1. INTRODUCTION

It has become a tradition that the final day of the International Conference on Image and Video Retrieval is a practitioners day where the researchers are joined by practitioners and will meet peers from multimedia industry. Multimedia search systems, engines and services gained strong interest and support, not only from academia and industry, but also from (inter)national authorities supporting these initiatives. The practitioner day aims to bring together and provide participants from industry (content owners, producers, creators, archives, services, ..), policy makers, academic and industrial researchers with an overview about these areas. The practitioner day will present several already realized commercial solutions, but also will sketch and will share ideas about "Research to Business" road-maps or needs to commercially successful apply research results.

The next sections provide the abstracts of the talks from the speakers of the practitioner day of the ACM 2009 International Conference on Image and Video Retrieval. It includes the keynote given by Dr. Ricardo Baeza-Yates from Yahoo! Research Barcelona on "Mining the Web 2.0 for Improved Image Search" as well as various talks on "Distributed Media on the Web" and "Media in post-production and broadcasting" given by a wide range of speakers including Oliver Heckmann from Google Zurich, Jan Erik Solem from Polar Rose, Johan Oomen from Beeld En Geluid, Hans van Gageldonk from Philips Research Laboratories, Jean-Pierre Gehrig from Cinetis, Xavier Vives from Corporació Catalana de Mitjans Audiovisuals and Baris Sumengen from Like.com (formerly Riya).

2. MINING THE WEB 2.0 FOR IMPROVED IMAGE SEARCH

Dr. Ricardo Baeza-Yates from Yahoo! Research Barcelona
There are several semantic sources that can be found in the Web that are either explicit, e.g. Wikipedia, or implicit,

e.g. derived from Web usage data. Most of them are related to user generated content (UGC) or what is called today the Web 2.0. During the talk, Ricardo Baeza-Yates will show how to use these sources of evidence in Flickr, such as tags, visual annotations or clicks, which represent the wisdom of crowds behind UGC, to improve image search. Presented results are the work of the multimedia retrieval team at Yahoo! Research Barcelona and are part of a larger effort to produce a virtuous data feedback circuit to leverage the Web itself.

3. DISTRIBUTED MEDIA ON THE WEB

3.1 Youtube - a look into the worlds largest video site

Oliver Heckmann from Google Zurich

Every minute, 15 hours of video content are uploaded to Youtube. This talk will give a look behind the scenes of Youtube in terms of the technical infrastructure used and lessons learned. This talk will not be focused on video retrieval.

3.2 Face recognition in the wild: large-scale computer vision for web applications

Jan Erik Solem from Polar Rose

The first part of this talk will detail lessons learned in the building of Polar Rose, a service that uses face recognition for sorting, searching, and sharing digital photos on the web. Successful solutions as well as developments that never made it into production will be discussed together with user reactions and behavior. The second part of the talk will describe the use of object specific descriptors to search large images sets for near-duplicates and derivative works. Focus will be on using descriptors for faces but the procedure can be applied to any object class as long as the descriptor is sufficiently strong. Finally, some future challenges in object recognition will be discussed

3.3 Unlocking value of audiovisual heritage

Johan Oomen from Beeld En Geluid

Audiovisual material is a vital component of our heritage, collective memory and identity – all our yesterdays. But in its current analogue format it is difficult to access. Unlike documents which may last for centuries or even millennia, audiovisual recording materials have a life expectancy best measured in decades. Digitization of analogue carriers will ensure future access. Over the past years, technologies for large-scale migration have matured. The same is also true for thinking about migration projects in terms of their efficiency and the workflow models they could follow. Although the process is far from complete, approximately 10 million hours of European audiovisual material has already been digitized. Recently, the audiovisual production process shifted from analogue to digital. This so-called "born digital" content is directly ingested in asset management systems and will also be kept for posterity as electronic files. Due to digitization and digital production, audiovisual content collections are transforming from archives of analogue material into very large stores of digital data. Digitization is also a driver to establish new services. Distribution over networks, interoperability with other collections and flexible integration in other environments are just a few of many properties in this new era of enormous potential for audiovisual

archives. Therefore, large-scale digitization efforts do not only ensure long-term access, but also have the potential to reveal the social and economic value of the collections. This presentation will focus on the latter: the types of services that can be created as a result of large-scale digitization efforts and the social and economical benefits they bring. Value creation is a key notion, as it determines the factors that legitimize (and determine the level of) investments by the government and funding programs. Two projects the Netherlands Institute for Sound and Vision is involved in will provide examples: (1) Video Active and (2) Images for the Future.

Video Active (www.videoactive.eu) provides free access to European television heritage. Within this three-year project, a large collection of television archive content from 14 archives is selected, digitized and made available through a freely accessible multilingual web portal. The project is funded by the European Community within the eContentplus programme.

Images for the Future is the largest digitization effort in Europe to date. A consortium of 6 partners (including three archives) is migrating a substantial part of the Dutch audiovisual heritage to a digital environment. The project has three objectives: [1] Safeguarding heritage for future generations. [2] Creating social- economical value: migrating large quantities of audio-visual material is a precondition to unlock the social and economic potential of the collections. More professionals and individuals than ever before will be able to access material, and new services and businesses will be launched. [3] Innovation: by digitizing heritage on a massive scale, a completely new infrastructure will have to be built that can strengthen the knowledge economy of the future. Much is to be expected from recent developments in computer science, especially in the areas of data-mining, information retrieval and in the creation of new environments where content can be used.

3.4 Challenges in product search and new developments at Like.com

Baris Sumengen from Like.com (formerly Riya)

In the past, there has been multiple efforts in utilizing content-based image retrieval technologies in commercial applications. Like.com attempted this by applying CBIR to product image search in 2006. Since then it became a company with 15 million dollar annual revenue run rate and it is growing. In addition to technological challenges, I will discuss business challenges we faced during the growth of Like.com. Commercial viability has been the biggest challenge in CBIR startups. In our case, there has been disconnects between engineers and users in terms of tuning our technologies so that in the end user experience is improved and customers buy more products. In addition to product search, we just rolled out a new product we call "Likesense". Likesense is a CBIR-driven advertising platform targeting image page views on social networks such as Myspace and Facebook. Social networking sites have large amounts of image page views that they are not monetizing currently. We have started real world Likesense tests on Myspace network starting June 2009.

4. MEDIA IN POST-PRODUCTION AND BROADCASTING

4.1 User scenarios and user requirements from media professionals

Xavier Vives from Corporació Catalana de Mitjans Audiovisuals

This talk provides a brief summary of user requirements and scenarios in the media industry, and of how R+D collaborative projects have helped CCMA, the Catalan Public Broadcasting Corporation, improve their in-house developed media asset management. European and national co funded projects are valued as positive by industrial media companies as they give the possibility to get in contact with the latest state-of-the-art, to network with relevant European partners and, somehow, to be able to improve their existing technological products. However, the experience also shows that their "Research to Business" road-map is normally diffuse and hard to achieve. On one side, there exists a gap between commercial needs from industrial partners, and research centres and universities objectives and needs. Besides, what normally happens is that we all generally focus too much on invention but hardly manage the innovation, while this is a really stimulating and profitable challenge we should also consider. The talk will also give a brief introduction of CCMA, the Catalan Broadcasting Corporation, and its own-developed Media Asset Management, named Digition. It also shows some considerations about the media sector where this Corporation competes. The most interesting parts, though, are the ones that explain some of the results of SEMEDIA, an FP6 project co-funded by the European Union. This project has shown an example of how the bridge between research and industry can be narrowed with a simple formula: letting user feedback guide technical research. The process of user requirements and scenarios gathering is explained, as well as some of the most relevant requirements from users of the Broadcast, Postproduction, and Social Web Scenarios. These requirements are general and wide enough to be considered whenever dealing with the development of a Multimedia search system, engine or service. They were discovered in SEMEDIA project thanks to a 12 month process that implied up to nearly 2.000 surveys. The paper ends explaining how integration and testing of SEMEDIA technologies tasks were conducted, and the benefits that could be obtained from this effort.

4.2 Multi-Sensor Information Retrieval for Lifestyle Applications

Hans van Gageldonk from Philips Research Laboratories

Information and content is becoming abundantly available. Audio, video, and pictures are instantly accessible and can be enjoyed on a multitude of connected devices. At the same time, it becomes harder for people to organize and retrieve the desired content at the right time and place.

On the other hand, there is a strong trend in society towards individualization and thus the need for personalization. In addition, people live ever busier lives, and wish to have relaxed quality time with their family and friends as compensation.

In the Lifestyle program in Philips Research we explore opportunities in this space between information and content on the one hand, and truly personalized applications on the other. In this presentation I will explain research results in making information easily accessible for end-consumers. Examples include automatic summarization of movies and home video, photo collections, as well as sports highlight detection. Then I will explain new research directions in retrieving information of people and their environments (context), using body signals and interpretation of the context people are living in. I will show some initial results on the combination of this research, resulting in real-time applications where body and other signal information is used to enhance a media experience. Ultimate goals are two-fold: to immerse people in the (AV) experience on the one hand, and to help people actively relax themselves on the other hand.

4.3 Efficient film digitization and restoration: Making of the 20th century film archives accessible

Jean-Pierre Gehrig from Cinetis

During the 20th century, hundreds of millions of films have been shot around the globe by motion pictures enthusiasts. Today, only a few part of these films have been secured in a digital archival system. Our challenge is to convert the physical material in digital data to make it accessible. Cinetis developed an efficient approach to digitize, restore and preserve old films stocks. The process is highly automated and uses advanced image processing techniques such as motion tracking for automatic video stabilization, video segmentation and key frame identification for assisted color grading.

5. BIOGRAPHY

Dr. Sébastien Marcel

Sébastien Marcel (http://www.idiap.ch/∼marcel) is senior research scientist at the Idiap Research Institute. He manages research projects on man-machine interaction and biometric person recognition and supervises a team of reseachers. He has obtained his PhD in signal processing from "Universite de Rennes I" in France (2000) at CNET, the research center of France Telecom. He is currently interested in multi-modal biometric person recognition (face detection, face recognition, speaker authentication, EEG-based authentication), man-machine interaction (hand gesture recognition) and content-based multimedia indexing and retrieval. Since January 2008, he is the main coordinator of the FP7 project MOBIO (http://www.mobioproject.org) on mobile biometric authentication.

Dr. Roelof van Zwol

Roelof van Zwol is a senior research scientist at Yahoo! Research (http://research.yahoo.com/), where he leads the multimedia group in Barcelona. Prior to joining Yahoo! in 2006, he was an assistant professor at Utrecht University working on various topics such as XML retrieval, mulitmedia, and Geographic Information Retrieval. He received his PhD in Computer Science from the University of Twente in the Netherlands in 2002 on the topic of modeling and searching multimedia document collections. He is also the technical coordinator of Semedia, the EU project on Search Environments for Media, and through Yahoo! Iberia involved in the WeKnowIt EU project. His research interests are primarily focussed on multimedia interaction, -mining, and -search and spatial information retrieval.

Dr. Ricardo Baeza-Yates

Ricardo Baeza-Yates is VP of Yahoo! Research for Europe and Latin America (http://research.yahoo.com/), leading the labs at Barcelona, Spain and Santiago, Chile. Until 2005 he was the director of the Center for Web Research at the Department of Computer Science of the Engineering School of the University of Chile; and ICREA Professor at the Dept. of Technology of Univ. Pompeu Fabra in Barcelona, Spain. He is co-author of the book Modern Information Retrieval, published in 1999 by Addison-Wesley, as well as co-author of the 2nd edition of the Handbook of Algorithms and Data Structures, Addison-Wesley, 1991; and co-editor of Information Retrieval: Algorithms and Data Structures, Prentice-Hall, 1992, among more than 150 other publications. He has received the Organization of American States award for young researchers in exact sciences (1993) and with two Brazilian colleagues obtained the COMPAQ prize for the best CS Brazilian research article (1997). In 2003 he was the first computer scientist to be elected to the Chilean Academy of Sciences. During 2007 he was awarded the Graham Medalfor innovation in computing, given by the University of Waterloo to distinguished ex-alumni.

Oliver Heckmann

Oliver Heckmann is an engineering manager at Google in Zurich/Switzerland. He is leading a 30 person Youtube development team working on projects like Youtube Video ID, Youtube Insight, Youtube GData API, and Youtube monetization. Before joining Google, he was working as a research group head in multimedia communications at the Technical University of Darmstadt in Germany.

Jan Erik Solem

Jan Erik Solem (http://www.polarrose.com/) currently works as CTO of Polar Rose and as associate professor at the Mathematical Imaging Group, Lund University, Sweden. He received his M.Sc. in Engineering Physics with focus on computer vision from Lund University in 2001. His PhD-thesis "Variational Problems and Level Set Methods in Computer Vision - Theory and Applications", presented in 2006, received the award for "Best Nordic PhD-thesis in the field of image analysis and pattern recognition in the period 2005-2006" from the Nordic IAPR associations at SCIA 2007. His current research interests include face and object recognition, large-scale vision problems and optimization. Previous research topics include theory and applications of variational level set methods within segmentation, shape modeling and 3D reconstruction. He is the founder of tech startup Polar Rose, a company that works with face recognition for sorting, searching, and sharing digital photos on the web. Polar Rose received a number of awards like "Red Herring Global 100" in 2007 and "World Economic Forum Technology Pioneer" in 2008 for pioneering work on deploying computer vision in large-scale web applications.

Johan Oomen

Johan Oomen (http://portal.beeldengeluid.nl/) is managing the R&D department of the Netherlands Institute for Sound and Vision, part of the Images for the Future project project team. He has been appointed in this position in 2008, but has been working at Sound and Vision since 2000. He is mainly working on externally funded R&D projects and managing the in-house deployment of Digital Asset Management technology.

Xavier Vives

Xavier Vives (http://www.ccma.cat/), MBA and MEng, obtained his Master of Engineering in Electronics from La Salle Engineering School (URL) in 2000, soon after achieving two Bachelors of Engineering, one in Telecommunications and another one in Computer Science. He has received his Master in Business Administration at ESADE Business school in 2009, and before received a Master Degree in Project Management in 2007 (UPC). He started his professional career at CCRTV in 1997 as an analyst/programmer, while also working as an engineer at TVC, the National Television of Catalonia. Five years after, he tackled new challenges in the artificial vision field, working as the RTD Department Manager of Panlab, a company that develops technology for Bioscience. Later, in 2004, he joined Mier Communicaciones, a company devoted to the design, manufacture and installation of professional equipment for TV, Digital Radio and DTT broadcasting. He worked in this firm as the Production Manager of the Terrestrial Communications Division until 2007. Since then, he has been the project manager at CCMA that has leaded several co-funded RTD projects at ICT department. His main interests include analog and digital television, media technologies, audio and video processing, intelligent information management, social networking, human computer interaction, business administration and project management.

Hans van Gageldonk

Hans van Gageldonk is department head of the group Experience Processing in Philips Research (http://philips.com/). This group is allocated in the Lifestyle program of Philips Research. His research interests include body signal processing for lifestyle applications, video analysis and experience creation. Previously he has been a senior researcher in the area of embedded databases for consumer applications, as well as compiler design for embedded DSPs, all for Philips Research. In 1999 he worked for Philips Semiconductors / VLSI Technology in Sophia Antipolis, France, on compiler design for embedded DSPs. Hans holds an MSc degree in computer science (1994) and a PhD degree in energy-efficient asynchronous VLSI design (1998), both from Eindhoven University of Technology.

Jean-Pierre Gehrig

Jean-Pierre Gehrig (http://www.cinetis.ch), CEO and co-founder of Cinetis SA, is an electrical engineer with several year of experience in software development and in embedded system design. Jean-Pierre believes in simple, user friendly, systems and nicely written code. Jean-Pierre holds a B.Sc. in electrical engineering from the University of Applied Sciences in Sion, Switzerland.

Baris Sumengen

Baris received his BS degree from Bogazici University, Turkey in both electrical engineering and mathematics in 1998. He received the MS degree in 2000 and the PhD degree in 2004 in electrical and computer engineering from University of California at Santa Barbara. From 2004 to 2006, he has been a postdoc at UC Santa Barbara Center for Bio-Image Informatics. Since 2006, he is with Like.com (formerly Riya), a

style shopping engine that utilizes content-based image similarity in matching similar products. He has lead the visual search team in 2006-2007. Currently he is leading the search engine marketing team.