Panel

Diversity in Multimedia Information Retrieval Research

James Z. Wang[1] (Chair), Nozha Boujemaa[2], Alberto Del Bimbo[3], Donald Geman[4], Alexander G. Hauptmann[5], and Jelena Tešić[6]

[1] The Pennsylvania State University, University Park, PA, USA
[2] INRIA, Rocquencourt, France
[3] University of Florence, Florence, Italy
[4] Johns Hopkins University, Baltimore, MD, USA
[5] Carnegie Mellon University, Pittsburgh, PA, USA
[6] IBM T. J. Watson Research Center, Hawthorne, NY, USA

ABSTRACT

Multimedia information retrieval is a highly diverse field. A variety of data types, research problems, methodologies are involved. Researchers in the field come from very different disciplines, ranging from mathematical and physical sciences, computational sciences and engineering, to application domains. The panel, consisting of highly visible active researchers from both academia and the industry, opens a discussion on the importance of diversity to the healthy growth of the field. This paper records their opinions expressed at the panel.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms

Algorithms, Design, Experimentation, Human Factors, Performance, Standardization, Theory

Keywords

Benchmarking, Image Retrieval, Video Retrieval, Theory, Applications

1. INTRODUCTION

Multimedia information retrieval is a diverse field. Unlike many other research fields, a wide range of data types have been studied by multimedia researchers. These include text, hypertext, audio, graphics, animation, image, video, rich text, spread sheet, presentation slide, combinations of these, and user interaction. It is a challenge to categorize research and development activities in this field. Roughly, one can consider the current activities residing in a multi-dimensional space, the axes of which can be the following.

- **System:** architecture, enterprise integration, interoperation, design, security, protocols.
- **Content:** metadata, standards, formats, compression.
- **Services:** content creation, crawling, storage, browse, measurements, similarity retrieval, classification, categorization, filtering, clustering, summarization, mining, preservation, decision support, multi-modal fusion, user modeling/personalization, aesthetics ranking.
- **Use:** human-computer interface, interaction, feedback, psychology, hardware devices.
- **Evaluation:** methods, performance metrics, usability, benchmark.
- **Implementation:** indexing, distributed or grid computing, scalability, optimization.
- **Social/Business:** life cycle management, intellectual property, copyright, privacy, legal, social impact, institutional/communities dynamics, policies, user information seeking behaviors, sociology, accounting, marketing, entrepreneurship, strategy.
- **Applications:** Web, biomedicine, art/culture, law enforcement, manufacturing, humanities, government, education, sensors and sensor networks, scientific literature.

While continually developing new techniques for multimedia retrieval, researchers in the field have leveraged mature methodologies developed in related fields including databases, information retrieval, signal and image processing, graphics, vision, document imaging, design research, human-computer interaction, computational sciences, machine learning, statistical modeling, data mining, pattern analysis, artificial intelligence, data fusion, social sciences,

MIR'06, October 26–27, 2006, Santa Barbara, California, USA.

and domain knowledge for applications. Researchers in the field are rarely graduates of a clearly defined multimedia major. Instead, most of them are from various disciplines: computer science, computer engineering, electrical engineering, mathematics, statistics, information sciences and technology, physics, business, humanities, biology, medicine, ...

Because of the inherent heterogeneity of the field, it is natural that different people take very different approaches in their research. If leveraged, this diversity is a powerful advantage because it helps the field maneuver quickly and smoothly in today's highly dynamic environment. On the other hand, this high degree of diversity tends to make breeding grounds for prejudices, biases, or even bigotry which in turn harm the creative potential of the field.

This panel offers an opportunity for researchers to exchange views on the importance of diversity to the prosperity of the field. Recorded in Section 2 are the opinions of the panelists. When the written comments were collected, no length limitation was set. The panel moderator James Z. Wang maintains a neutral stance during the panel discussion. However, in the interest of keeping a record, his personal opinions are incorporated into this article.

2. MAIN ISSUES DISCUSSED

In this section, we list the main issues discussed during the panel and the written comments of the panelists.

2.1 Benchmarking or Not

There are a number of benchmarking initiatives (e.g., TRECVID for video retrieval) in the field of multimedia information retrieval. Does benchmarking kill innovation? Do existing benchmark initiatives cover diverse multimedia retrieval tasks?

Nozha Boujemaa: Besides TRECVID, benchmarking initiatives are becoming numerous: ImageCLEF, Pascal, ImagEval, ... This definitely shows that no existing single initiative could be by itself satisfactory by offering the context to test all the tasks addressed by our community. Also, due to the richness of the scientific objectives of multimedia search engines corresponding to growing and evolutionary use-cases and user needs, benchmark initiatives should be able to follow the field dynamic.

On the other hand, sometimes the work of a researcher could be criticized because of not participating to one given benchmark initiative (that could be important for historical reasons). This yields to an extreme situation when the relevance of researcher work is measured by the participation to these initiatives. In this context, the benchmarking will format all the research work in the community letting people working on the same tasks and necessarily limit the innovation.

Nevertheless, benchmarking remain necessary and valuable for the community as it provides objective reference among the numerous technical academic and industrial solutions. But it should be carefully set with clear and fair rules, and wide consensus of the community regarding definition of tasks, evaluation parameters, performance measures, ground truth setting, conflict of interest avoiding, ... Joining all these conditions remain quite challenging and the bottleneck of some initiatives.

There are many efforts at the European level to investigate these questions via Network of Excellence (e.g., Muscle, Delos) and more recently by a Coordination Action "Chorus". "Chorus" is aiming to analyse and put together not only the European effort but also efforts at the national and the international level around multimedia search engines regarding questions such as benchmarking, standards, interoperability, ..., and to provide technological roadmap and recommendations.

In the ideal conditions, I personally believe that winning a benchmark is worth thousand publications for academia and thousand press releases for industrials and represent a "moment of truth" among all what technology providers (academia/industrial) can argue on their own work and results.

Alberto Del Bimbo: We cannot say benchmark is not useful. But its relevance to the scientific community very much depends on how much goals and objectives on which benchmarks are constructed have been clearly defined and pursued. So benchmarking loses its utility if the databases are simply a miscellaneous collection of more or less difficult test cases. Either test cases are built so as to reflect scientifically challenging problems or are important real cases truly motivated by rewarding applications.

In the first case they can provide some stimulus to improve the quality of the research solutions; in the second they can speed up technology transfer and improve the popularity of the research field.

All in all, I believe they are potential innovation killers, are mandatory when research fields are mature enough, and require careful assessments of the test cases by scientific and industrial subjects. For multimedia retrieval coverage is still very partial.

Donald Geman: Whereas I am not very familiar with the benchmarking initiatives in multimedia information retrieval, I have observed the impact of analogous activities in the area of object recognition in the computer vision community. In my opinion, the results are decidedly mixed. The utility seems to depend on the quality of the database and the maturity of the area.

High quality databases for well-studied tasks seem to be useful. For example, for face detection, which, apart from character recognition, is arguably the most closely examined example of generic object recognition (i.e., searching for all instances of a specific object category) the existence of the MIT-CMU database has been helpful in comparing results, probably because most everybody agrees that the database is "representative". Similarly in handwritten digit recognition (with MNIST).

However, for more ambitious tasks for which the community is still searching for some consensus about what constitutes success, not to mention well-developed methods and some theoretical understanding, the existence of a "standard"

database can impede creativity by overly emphasizing purely performance-driven criteria. For example, for the more ambitious task of recognizing instances from multiple categories, I sometimes feel that the celebrated CalTech 101 database, though constructed with the best intentions, has in fact had some negative consequences. Many questions have arisen about whether or not it is "representative", results are generally poor, papers report incremental gains, and, perhaps most importantly, researchers are criticized for not using it.

Alexander G. Hauptmann: Benchmarking efforts are usually intended to be precise and measure carefully how systems or algorithms perform with respect to a dataset, a task and an evaluation metric. Thus, to be scientifically valid, they have to be specific such that results are unambiguous and measurable. This makes benchmarks necessarily very narrow in focus and they often exclude much research. The goal is to find research questions that are of general interest, where a number of researchers are working on pretty much the same goal, and then evaluate this work.

Benchmarking, per se, does not kill innovation; it provides a focus for research by different groups to be compared. If someone claims success on a task that is sufficiently similar to the benchmark task without using the benchmark, they better explain why their task is different enough to warrant not using the benchmark. In the greater picture however, we find that benchmarks do suppress innovation in a number of ways: Funding agencies cling to the benchmarks as demonstrable measures of success, since these measures are otherwise hard to come by. Short (1-2 year) intervals between evaluations can make it difficult to embark on radically new approaches, that will take longer to get the 'kinks' worked out and prove themselves.

Researchers also tend to merely adopt successful techniques from earlier benchmarks, reducing the diversity of the research - which may also be a good thing for a more mature research area by weeding out spurious claims of successful approaches.

Jelena Tešić: Absolutely not, benchmarking proved to clearly initiate innovation. In the 90s, Text Retrieval Conference (TREC) provided the large-scale support necessary for text-retrieval methodologies. Size of the collection (2 GB of full text) was 1000 times larger than what researchers commonly used before (2MB of abstract). This jump-started the field, resulting in 225 participants, evolving tasks and breakthroughs in the text retrieval research, ultimately enabling the Web search engines to point users to information they seek using statistical techniques that are able to scale to web data.

Multimedia information retrieval is a data driven research field that is gaining a momentum. Parallel can be derived to the text retrieval efforts. Lack of benchmark in multimedia information retrieval community in the late 90s is correlated with no breakthroughs in the field during that period. There was a great need for larger datasets, something more appropriate for 21st century and terabytes of multimedia data created everyday. TRECVID started as a TREC workshop, and evolved to similar forum for encouraging and evaluating new research in video retrieval. TRECVID participants have access to 100s of hours of annotated video datasets, and are able to define diversity of the tasks. Since 2002, TRECVID tasks went through couple of changes: shot-detection challenge reached its maturity, and management of unlabelled video archives emerged as a new challenge. And multimedia community took notice: ACM Multimedia had a tutorial on TRECVID and 2 keynote talks in 2004 and 2005 were on multimedia content retrieval techniques developed and proven using TRECVID benchmark; CIVR 2005 3 out of 4 keynote talks focused on TRECVID; ICME 2006 content track had a plethora of papers on improving semantic content detection in multimedia using context that used TRECVID datasets and common data annotation from that forum. On the other hand, CVPR 2006 became a chaotic benchmark evaluation workshop for image retrieval, where majority of the researchers used Corel, Caltech 101, and UCI datasets, to present the findings.

These recent developments in information retrieval show us that some form of large-scale standardized datasets and metric is driving the innovation and a field. Multimedia community should consider a content track at key multimedia conferences where researchers are mandated to use one of the standardized datasets, metric, and evaluation process to present their work. This might bring more credibility to the research efforts, and drive the future breakthroughs in multimedia information retrieval community.

James Z. Wang: Benchmarking has no doubt been valuable in promoting progress in scientific development for a number of research fields. In image compression for example, the world-famous image (also known as 'Lenna') of Lena Soderberg, as a partial image taken from the November 1972 issue of Playboy magazine, has made it possible to compare a number of dimensions among different compression algorithms. Today, the widespread adoption of the Lenna picture in benchmarking is viewed by many as one driving force behind the fast advancement of the compression field.

In content-based image retrieval, the set of Corel pictures have been used by many researchers since the mid-1990s. In an era in which copyright owners of images eagerly wanted to profit from the use of their pictures, this resource of a large image collection has enabled many researchers to carry out their endeavors. The Corel collection itself is not problem-free. Researchers familiar with the collection would know that every 100 Corel images form a group with some shared semantics or concepts (e.g., Paris, flowers, tigers). But there is sometimes a lack of diversity in the way these 100 images are taken or produced. For example, it is possible to have 100 images of tigers all with the same type of tiger and same grassy background. On the other hand, the diversity of certain Corel image groups (e.g., England) can be so substantial that it is practically impossible for any machine learning or statistical modeling method to conclude anything meaningful from such groups. Even after more than a decade and with the widespread of digital images on the Web, the Corel collection is still probably the largest image research data set with good qualities.

In my opinion, the field needs more copyright-free or copyright-unenforced large-scale data sets readily accessible to researchers. However, benchmarking can potentially kill innovation in a diverse field like multimedia information retrieval. Often times, we see peer-review comments like "The work should be used on so-and-so benchmark to validate its value". As mentioned in Section 1, this field is so large, with so many dimensions. To me, it is inconceivable that anyone can possibly develop a set of data and a set of tasks that would cover all needs. As I have worked in this field for about a decade, I no longer believe that one algorithm can be suitable for many applications. A researcher working on images should not be blamed for not knowing a video benchmarking. A researcher developing a system to analyze ancient paintings should not need to know about benchmarking efforts unrelated to paintings.

I also believe comparisons between approaches are often not a simple matter of being better or worse. Engineering systems are usually assessed from multiple angles, which themselves vary with the targeted applications, or specifically data domains. Benchmarking, unfortunately, has the innate function of limiting the angles from which we evaluate a system. By circumscribing our viewing perspectives, benchmarking diminishes creativity. The issue becomes more serious if we take into consideration scholarly interactions. Researchers who happen to bring in new perspectives are prone to criticisms rationalized by benchmarks and probably not rarely are deprived of opportunities to deepen new ideas. In a sense, the negative effects of benchmarking not only are imposed on those who hold on to benchmarks but also spread to push away those who do not.

Because of these serious potential consequences of its misuse, I advocate for great caution in the development of benchmarking in this field. I would, however, encourage the efforts to make available more data sets from more application domains. I would encourage as wide as possible in the diversity of data sets, problems, and approaches.

2.2 Image or Video

Some may argue that video retrieval is now the focus of the field of multimedia information retrieval. Are there still needs for image retrieval research? What are the interesting and open research problems in image retrieval? What about other multimedia retrieval problems?

Nozha Boujemaa: From time to time, we hear that image retrieval is no more needed and most often from people working on video retrieval who may not be faced with the relevance of image retrieval use-cases and their impact. This also due to the initial big push for TV news application that was trendy application in the early years of content retrieval.

The awareness of the impact of content-based retrieval has spread out later for different application domains. More recently, large image archives present very valuable application with high impact. Let's give some examples within application domains regarding scientific or specific content: botanical applications, remote sensing archives, medical archives, ... Besides image retrieval solution most often provides the bases for many video retrieval problems. Also, we notice that still numerous important problems are not yet solved: summarisation, (semi) automatic annotation, ... great efforts are still needed in this direction that certainly involves cross-modal interaction.

Alberto Del Bimbo: Image retrieval has been the starting point of research on non traditional objects content-based retrieval. Problems are still open from both research and performance viewpoint. Just to mention few of them: effective modelling of spatial relationships, effective graph-based modelling of rich image content, effective graph-based indexing. However, no truly interesting applications have been found yet for image retrieval. On the contrary video retrieval is strongly pushed by a variety of industry operators and promises to be the present and next future major focus of research in content based retrieval. In some sense, several problems of image retrieval are anyway revitalized in video retrieval. As to the other multimedia retrieval problems, 3D object content based retrieval is presumably a new field of investigation with interesting future applications. Here 3D surface representation, local and global similarity measures, effective indexing are challenging problems. Important areas of research with specific open problems can be, among the others 3D face retrieval, 3D CAD objects retrieval, 3D object temporal changes retrieval, ...

Donald Geman: Of course there is still a need for image retrieval. In mathematics, George Polya observed that "if there is a problem you cannot solve, there is always a simpler one that you cannot solve either." In contrast, the modus operandi of computer vision seems to be "if you cannot solve a problem, move to a harder one." (Although video retrieval may not be harder.) Moreover, changing the problem is often driven more by hype and advertising then by scientific considerations. The "interesting and open research problems" are the same ones that have been around for at least a decade, which is not very long. Commercial applications will come when prime-time solutions arrive.

Alexander G. Hauptmann: Video retrieval in practice has been reduced to image retrieval, with slightly different metadata and perhaps some simple motion features thrown in. Essentially it is image retrieval, with low quality, heterogeneous data.

In theory, video and image retrieval should be quite different, but in practice they are not. Video retrieval benchmarks are currently also the only ones with significant collection sizes - it is too easy to get good results on some relatively small data set of less than 100,000 images, by current standards. The real issues are in the billion image datasets. Many other retrieval tasks are possible, but they have not been standardized with any consensus, not do they have large collections of data and truth associated with them.

Jelena Tešić: Need for advanced image retrieval research is great. However, QBIC has shown that image retrieval solves only part of the system requirements to access images. In the absence of real breakthroughs in the traditional areas of image retrieval and computer vision in the past couple of decades, a significant progress has been made in the areas of video retrieval, large-scale image management, coarse

semantic detection and use of context. This argument is parallel to the one of building hybrid cars vs. focusing on alternative fuel development. Hybrid cars were developed as a short term solution to extend limited fuel resources, and hybrid technology is the alternative fuel development.

James Z. Wang: I have heard arguments like "Image retrieval is dead. Multi-modal retrieval leveraging all information sources such as audio, speech transcript, movements in video is the way to go." I believe image retrieval is still and will remain an important problem. In many application areas, a database of unannotated images is all they have. For example, in biological experiments and medicine, high throughput, high resolution, and high dimensional digital imaging equipment have created enormous problems for biologists but golden opportunities for multimedia researchers. Take another example. Google has been able to impact the everyday life of billions of people. May there be one day that an image annotation algorithm can help manage the vast number of digital photos available? May there be one day that an image classification algorithm can tell the potential aesthetic value of a photo to help photographers compose? May there be one day that a computational system can reveal the hidden 'code', or what has made them so special, in artistic paintings?

I think image retrieval is an active field. We are still making good progress in the last few years. A recent survey [1] provides more details. Human brains are not necessarily that complicated. Yet we can be trained to be a pathologist, a satellite image analyst, a photograph annotator, a story illustrator, a painting historian, ... Maybe we can one day build a machine, with the right input devices as a stereo camera and a large amount of computing power, that can be trained to do as well or better than humans in some of these tasks.

There has been real world use of image retrieval. In as early as 1998, the WIPE image filtering system [4] we have developed, based on retrieval, was licensed to a company which later provided parental control functions for large international Internet service providers. Since 1999, more than 100 institutions have requested and obtained the SIMPLIcity image retrieval system [5] for their projects. These include (1) the largest aviation community, airliners.net, to search their database of more than one million images for millions of users, (2) the largest U.S. national parks photography site, terragalleria.com, and (3) the largest mineral research community, mindat.org. We have been working on other interesting problems for potential real-world deployment, including real-time image annotation (ALIP system), Story Picturing Engine, computational aesthetics, image-based CAPTCHA (IMAGINATION system), scientific figure retrieval (collaboration for next-generation CiteSeer), ...

It is debatable if video retrieval is more difficult than image retrieval. A video clip, packed with sound and movements, certainly has more information than a frame extracted from the same video clip. To judge the semantic content of an image frame from the video is like to judge whether a novel is a romance based on reading just one page in the novel.

Similar to my earlier comments regarding benchmarking, I believe we need to encourage a diverse portfolio of research problems to tackle. If in a certain era, we do not make much progress in one problem, we may be making great progress in some other problems. As long as the problem is still open and as long as human brains can be trained to perform this task, I see no reason why we should not continue to explore computational approaches for this challenge.

2.3 Tools or Systems

Is there a need for a wide spectrum of research areas (e.g., from developing core mathematical tools to developing application systems, from users in the loop to fully automatic) in the field of multimedia retrieval?

Nozha Boujemaa: According to my previous comments: of course.

All the spectrum components are needed to contribute to the big framework. Of course the mathematical foundations are essential but also less fundamental contributions could sometimes be crucial to achieve significant technological impact. Putting together in original and efficient way existing but separated know-how that were not able to previously generate the technological jump is also of great benefit and added value.

Note that, the general retrieval problem seems to require most often the participation of the user in the loop. I believe in participatory systems than in fully automatic not only because of the bottleneck due to the semantic gap but also because they have better chance to be adopted by end-users. Hence machine learning issues from few examples (or on-line learning in other words) will be essential in this kind of interaction. In the same time, this topic includes plenty of challenging scientific problems.

Alberto Del Bimbo: Yes of course. I believe that machine learning on the one hand, and users in the loop on the other will be two key research areas. Machine learning will support more effective machine-side content understanding. Users in the loop is the primary distinguishing feature of content-based retrieval; with the advent of new wireless devices and the possibility of more natural modes of interaction interaction paradigms must be redefined and new possibilities for multimedia retrieval will open.

Donald Geman: Absolutely. We should be trying all sorts of things because nobody has the answers. Some researchers should concentrate on developing "theoretical foundations," others on solving specific instances of the problem and everything in between.

Alexander G. Hauptmann: Clearly a wide spectrum of research is beneficial to any field, as long as there is a way to separate the chaff from the truly significant innovations.

Jelena Tešić: First, there is an urgent need to define a multimedia field as a coherent entity of research efforts in the multimodal domain, or in the unimodal domain that includes users. In the past couple of years, multimedia conferences and journals tend to be cluttered with somewhat outliers. If multimedia is defined by majority of the papers

published in the multimedia literature, is that definition true when majority of the papers published in multimedia literature belong to coding and watermarking domain?

Then, there is a great need for new techniques to be developed (a) that scale to the sheer volume and complexity of the multimedia data (b) account for the "noise" level that low-level descriptors introduce, and (c) efficiently deal with the computation bottlenecks. In addition, areas of bioinformatics and medical multimedia data would benefit greatly from the new techniques developed in the field of multimedia retrieval.

James Z. Wang: I remember being asked of an interview question, "Some people develop fundamental tools or theories, and others develop systems using existing tools. In this spectrum of research, where do you position yourself?" This can be a valid question in some disciplines where they take people from one end of the spectrum but not the other. In my opinion, for the field of multimedia information retrieval, this question is not as meaningful for judging the quality of a work. As the prominent Chinese reformer Deng Xiaoping puts it, "It doesn't matter if a cat is black or white, so long as it catches mice." A theory in a practical field as multimedia information retrieval is a good theory as long as it has good practical use. (In mathematics or physics, a good theory may have no practical implications.) A system work is a good work as long as it solves real problems.

In 2005, a researcher of our field wrote an article entitled "We are Sorry to Inform You..." [3] detailing some of the absurd rejection comments for some of the landmark papers of E. W. Dijkstra, E. F. Codd, A. Turing, C. E. Shannon, and others. Some of the quotes "Publishing this would waste valuable paper", "It can be safely rejected", "This is a bizarre paper. It begins by defining a computing device absolutely unlike anything I have seen...", "This paper is poorly motivated and excessively abstract. It is unclear for what practical problem it might be relevant."... Didn't Larry Page and Sergey Brin's paper on PageRank get rejected by an Information Retrieval (IR) conference for being "not IR enough"?

I would choose to work on certain problems but not others and take on certain approaches because of personal interest in the problems and preference of the approaches. But I would not discourage others working on other problems or taking other approaches. As discussed in Section 1, diversity is to our advantage. We need to have people working on different problems using different approaches. It is often difficult to single out work that will eventually generate significant impact. We need to be more tolerant to different schools of thoughts.

2.4 Trends and Creativity

What advice would you offer young researchers in the field on choosing research topics? Should they follow current trends? Or should they try to be creative on selecting problems?

Nozha Boujemaa: My advice is quite simple: follow inspiration and personal conviction.

Sometimes, I can hear from new and young Ph.D. students "I need to work on this topic". When I ask why, most often the answer is because "it's trendy topic, all recent papers are on this subject..." and rarely because "I'm curious about this question" or "I'm convinced that there is a new way to handle this problem and I want to investigate this challenge". This is a frequent behaviour within not yet mature new Ph.D.'s. As researchers are getting experienced they realize that the trendy topics are not a guaranty to have a successful career. The trendy topics could not allow a big scientific impact for a given contributor. On the other hand, fully new topics could not be well understood in the beginning and is a more "risky" choice but more "advantageous" in case of success.

Reciprocally, I would like to warn against some possible supervisors advice to their young students "don't work on this topic, don't waist time on non trendy subject". As far as a problem is not solved it's worthwhile to work on if there is inspiration and originality of course!

Alberto Del Bimbo: I ever believed that a young researcher in engineering should first look at reality and find/define useful real problems to be solved. Then, make the effort of imagining new principles, eventually complex but reasonable, to be used as foundations for new solution. Being acquainted with the literature and current trends is important but a creative step is needed before. Attending key conferences is definitely important.

Donald Geman: A tough choice to be sure. Each research strategy has its risks. Given the absurd current emphasis on numerous publications (off by a factor of ten in my view), a young researcher may take a risk in attempting highly creative work. On the other hand, following current trends in methodology is unlikely to lead to breakthrough results. However, it seems to me that people are generally attracted to illustrating new ideas, or even old ones, on new examples. The exception is again the silly criticism of not working on standard databases.

Alexander G. Hauptmann: There are clearly different schools of thought. Some researchers prefer to find better solutions to existing well-defined problems while others, in the spirit of pioneers, prefer to be the first to venture into uncharted territory, in the hope that others will follow as an interesting and fruitful area of research opens up. The big recognition tends to come for innovative research, not incremental work. At times, work unifying a number of areas can also be very valuable, although, like the pioneering work, failure is somewhat more likely there.

Jelena Tešić: Multimedia information retrieval offers a wide range of problems needed to be solved, ranging from pure theory to feasibility of real-life applications. And most of them are very difficult ones. There is a very fine line between formulating a research problem because it is an interesting direction to pursue and defining research problem for the sake of the problem or short-term paper. Later becomes an "l'art pour l'art" – proved in the modernism art to be a very certain path to further neutralize the content and sometimes noxious effects of progressive thinking. It is up to young researchers not to let this happen. We should also be careful not to re-invent problems and solutions from

other areas and apply it to a very small subset of multimedia retrieval, rather we want to drive the ideas and solutions from applied math and computer vision, and move forward.

James Z. Wang: Having just been through a six-year tenure track, I would suggest young scholars to pursue problems with the most potential impact rather than to pursue incremental gains aimed at publishing more papers. Sometimes, I feel that it is easier to convince paper reviewers when you are working on an existing problem with well-established evaluation criteria. But I can hardly get excited by working toward incremental improvements.

There is a risk you have to take of course. Some times the more novel the work is, the harder it is to get published. It is not uncommon to get strongly opposing comments from some reviewers while others think the work is pioneering. Do expect to go through rounds of rejections or major revisions. If it is gold, it will eventually shine.

My advice for young researchers is to try the best to stay with your passion and forget about the tenure process, or any process. If one cannot afford a bit of risk, one is unlike to produce significant work. No pain, no gain.

2.5 Academia and Industry

Academic researchers and industrial researchers often have different goals. In your opinion, what would be the ideal interaction between the two sides?

Nozha Boujemaa: One positive way is to let industrials draw and prospect the new market needs and imagine the new services and products and let academia project and address the corresponding scientific challenges.

This is a way to get "useful" and "applied" research effort from academia side as they are more expected to investigate more long term and risky scientific issues and less exposed to user-needs or new services that will be deployed by industrial to large or professional public. On the contrary, as industrials are supposed to be committed to short term results and objectives, they could certainly profit from the involvement of academia into the challenging new market needs and services. This consists on a smart way to bridge this gap making these two communities working together for more technology deployment and impact in everyday life.

Alberto Del Bimbo: Innovation transfer is nowadays a keyword for academy in any industrialized country. However it is a complex problem to be solved. On the positive side, innovation transfer is the way in which research groups can raise adequate funding for their research activity and build important infrastructures, while getting visibility also outside the academy. It is also the way in which academy may have important social impact. On the dark side, innovation transfer can easy monopolize the groups scientific investigation so that research will be for its greatest part only performance and application driven.

From the academy side, I believe that new organization and management rules are needed for research. Individual researchers are obsolete. Research groups are needed large enough to show distinct lines of investigation and activity. Group leaders must not ask every researcher to find the funding for his research, but instead let research ready to transfer also to support funding of theoretical investigations. Exchange of know how and experience within the group is mandatory.

In the relationship with industry, it is very important that academy does play the role of a centre of competence and not that of a partner with low-cost skilled manpower.

Donald Geman: Well, I myself have very little interest in the short-term goals that often drive industrial research. And work on long-term objectives may not be supported in industrial labs. This often has the effect of obliging industrial researchers to develop ad hoc methods and to put scholarship aside (e.g., in determining what is really new). So, whereas I may be very old-fashioned, I still see a kind of "oil and water" problem here in that the objectives are just naturally distinct. Hopefully others will have something more constructive to say on this subject.

Alexander G. Hauptmann: Industrial researchers tend to have a 1-2 year horizon, necessitated by company profitability needs. Academic researchers should be encouraged to look at what the world might need/want/have in 5-10 years. To avoid the extremes of only incremental work by industry and pie-in-the-sky work in the academic tower, collaboration and cross-fertilization is necessary.

Jelena Tešić: "Researchers usually fall in love with their tools and start gaining better and better understanding of their tools, completely forgetting why were these tools considered in the first place." – Ramesh Jain, MIR blog 2004. SVM is not an answer to all our problems. Querying Corel dataset is passé - we have hard time finding any web image search system that uses any of the techniques developed using Corel dataset. Is multimedia future in security and surveillance, not in Web data? There are number of issues to be resolved in multimedia research. Academic and industrial researchers need to disagree less, to drive coherent long term goals in multimedia retrieval more, and push for multidisciplinary research in multimedia domains. Together.

James Z. Wang: Ideally, researchers in academia should produce public-domain knowledge and train high quality students who can become productive industrial researchers, academic researchers, or successful entrepreneurs. The industrial researchers work to produce usable and profitable systems to solve real-world problems, leveraging results developed in-house and those available in the public domain. They feedback to the research community new open problems they face.

To achieve this ideal interaction, I feel that graduate students have to be trained to work on innovative problems or to create innovative tools or theories. If, for example, the goal of a Ph.D. student is not to think hard on new initiatives but just to compete in a known benchmarking effort, the student may choose to find the best ways to integrate known tools in order to win in the competition. A student trained in this way may not be able to raise new problems or generate new methodologies at the time of graduation.

Many jobs created for Ph.D.'s are leadership jobs, no matter if in a company or university. Given that today's research environment is so rapidly changing, a successful leader needs to know where to go next and how to take the entire team there. I feel that they need to acquire this capability while still in graduate school. I ask my Ph.D. students to either work on a new problem or develop a novel tool for an existing problem.

I have to admit that this is not always easy to carry out in practice. For a student to graduate and to find a decent job, more than ten publications are often expected nowadays. If a student devotes effort purely to new initiatives, the student may not be able to write enough papers to get a good job. This is an unfortunate reality. As a result, students are often advised to take a portfolio approach, i.e., to have a blend of incremental publications and relatively fundamental ones.

3. CONCLUSIONS

The panelists expressed their view points on issues related to the importance of diversity in the field of multimedia information retrieval. Specifically, they have discussed (1) whether benchmarking kills innovation, (2) is image retrieval still active, (3) to develop tools or to develop systems, (4) to follow trends or to be more creative, and (5) what interaction between academic and industrial researchers would be ideal.

4. ACKNOWLEDGMENTS

The research of James Z. Wang, Donald Geman, and Alexander G. Hauptmann has been funded by the US National Science Foundation. The research of Nozha Boujemaa and Alberto Del Bimbo has been funded by the European Commission DG INFSO. Jelena Tešić is funded by the IBM Corporation. The opinions expressed in this article are their personal views.

Rita Cucchiara, Jia Li, Milind Naphade, Eric Pauwels, and Alan Smeaton participated in discussions related to the theme of this panel with James Z. Wang, Nozha Boujemaa, Alberto Del Bimbo, and Alexander G. Hauptmann during a meeting in San Vincenzo, Italy [2]. Jia Li provided valuable comments to the initial manuscript. Weina Ge helped in editing.

5. REFERENCES

[1] R. Datta, J. Li and J. Z. Wang, "Content-Based Image Retrieval - Approaches and Trends of the New Age," *Proceedings of the 7th International Workshop on Multimedia Information Retrieval*, in conjunction with ACM International Conference on Multimedia, pp. 253-262, Singapore, ACM, November 2005.

[2] Joint DELOS-MUSCLE Summer School on Multimedia Digital Libraries. URL: `http://www-rocq.inria.fr/imedia/DelosMuscleSummerSchool2006`

[3] S. Santini, "We Are Sorry to Inform You ..." *IEEE Computer*, vol. 38, no. 12, pp. 126-128, 2005.

[4] J. Z. Wang, J. Li, G. Wiederhold and O. Firschein, "System for Screening Objectionable Images," *Computer Communications*, vol. 21, no. 15, pp. 1355-1360, Elsevier, 1998.

[5] J. Z. Wang, J. Li and G. Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 9, pp. 947-963, 2001.