Table of Contents

**Association for
Computing Machinery**

Advancing Computing as a Science & Profession

SIGMM Records

Volume 4, Issue 3, September 2012 (ISSN 1947-4598)

Editorial

Dear Member of the SIGMM Community, welcome to the third issue of the SIGMM Records in 2012.

As you can see, the format of the Records as changed dramatically with this issue, and the migration is going to be completed in the coming months. The new system is meant to make the Records more valuable and interactive for your benefit, and we hope the long wait for the third issue was worth your while. First of all, your submissions will become visible on the front page of the Records as soon as they have been approved by one of the editors, and they will be included in the following issue. The submission of standard content formats has become much easier than before: select your contribution from the pulld0wn menu, add your information and submit it, and see immediately on the submission page that your submission was successful. You can of course also send your contributions and any questions to enews-contributions@sigmm.org.

We are furthermore inviting a new category: Please tell us about your ongoing research projects! What are your goals and achievements? Who are your partners? How are your publications connected?

Of course, this issue has also some content: SIGMM's 2012 awards have been handed out at ACM Multimedia in October in Nara. In this issue you can read about the SIGMM award for outstanding contributions to multimedia, the best PhD thesis award 2012 and the first ever awarded Nicholas D. Georganas Award for the best TOMCCAP paper.

TOMCCAP announces a major policy chang; you can read about the startup foodQuest, the Open Source project Ambulant, and the latest MPEG meeting. You can read PhD thesis summaries provided by a two candidates who have recently passed their doctoral exams.

Last but most certainly not least, you find pointers to the latest issues of TOMCCAP and MMSJ, and several job announcements.

We hope that you enjoy this issue of the Records.

The Editors
Stephan Kopf
Viktor Wendel

Lei Zhang
Pradeep Atrey
Christian Timmerer
Pablo Cesar
Carsten Griwodz

GameDays & Edutainment 2012

On behalf of the conference co-chairs, we wish to provide a report of the eight GameDays, which have been held from September 18th to 20th at Technische Universität Darmstadt and in the premises of Fraunhofer IGD.

Opening of the GameDays 2012

The GameDays are initiated and mainly organized by Dr. Stefan Göbel, the head of the Serious Games group at the Multimedia Communications Lab at TU Darmstadt. The GameDays take place as a "Science meets Business" event in the field of Serious Games on an annual basis since 2005 in cooperation with Hessen-IT, the Forum for Interdisciplinary Research of TU Darmstadt and other partners from science and industry.

Like in 2010 and 2011, the GameDays 2012 were held both as an academic International Conference and as a "Science meets Business" congress. This year, the conference was held in conjunction with Edutainment 2012, the seventh International Conference on E-Learning and Games. The Edutainment origins from China and has been hold at different countries in Asia so far, this year for the first time in Europe.

On the academic side, 39 papers originated from 18 countries all over the world, among others Australia, China, Hong Kong, Japan, Singapore, Taiwan, and the United States, as well as Germany and different

European countries have been committed to the joint conference of which 21 have been accepted for scientific publication in the Springer LNCS conference proceedings. The acceptance rate has been 50% for full papers. All papers have been reviewed by at least four international reviewers.

The papers have been presented in four tracks focusing on

- Game-based training and learning
- Emergent learning and gaming technologies
- Authoring tools and mechanisms
- Serious Games for health

Keynotes were held by Noah Wardrip-Fruin from the University of Southern California, Santa Cruz, Wolfgang Müller-Wittig from Fraunhofer IDM, Singapore, Henrik Hautop Lund from the Center for Playware, Technical University of Denmark, and Florian 'Floyd' Mueller from the RMIT University, Melbourne, Australia.
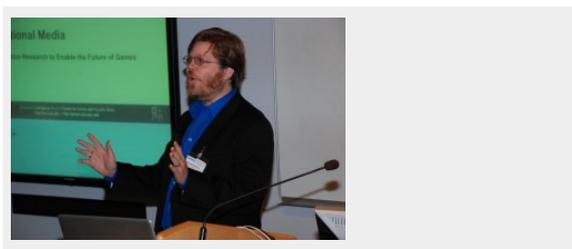
Noah Wardrip-Fruin

In his talk, Noah Wardrip-Fruin pointed out the importance of multidisciplinarity in computer games design (and game research as well) and the limits to handcrafting in computer games (e.g. level design, story, or animations) limiting the level of interactivity. He identified two important challenges to address multidisciplinarity: Authoring and generativity. He further presented several projects for an easy game creation, like Mame-o-Matic, a tool in which users define the game logic via graphs, or Ludocore, an authoring approach for generation of rules, not levels.

Wolfgang Müller-Wittig expressed the necessity to not just hear or see, but to experience through interactive visual real-time systems. He provided an interesting insight in various exhibits of Fraunhofer Singapore, among which were augmented world technologies for the classroom or museums and science centers.

Henrik Hautop Lund presented 'Playware', a modular hardware and software system consisting of interactive tiles which can be freely arranged on the floor or at a wall, connected to at least one other tile, thus creating a playground for various games. Players can step on the colorfully illuminated tiles in order to jump a certain

pattern, as a reaction game, or as a contest with another player. He showed that the system could be successfully used both for training of youth and elderly people with significant effects.

Floyd Mueller presented several innovative Serious Game ideas from his lab, all of them focusing the body. He presented a quadrocopter as a robot companion for joggers, or a bike helmet with a rear display showing the pulse rate only to the one driving behind the biker. One of his main statements were that Serious Games should not compare to traditional learning/working medias like e.g. school books but rather show that they can offer new ways of learning/training as alternatives and supplement.

All of these keynotes provided an interesting and promising multi-perspective view on the future uses of Serious Games technology in the fields of training, education, health and sports. However, they also pointed out the necessity of interdisciplinary future research in these fields.

In addition to the scientific talks, an exhibition area with more than 20 demos and prototypes provided insight into a variety of Serious Games from the fields of learning, training, sports & health. Among those BalanceFit, a game for training of balance for fall prevention of elderly people which is currently in evaluation in several hospitals in Darmstadt or KTexFlex, a collection of games for preservation of physical and mental fitness using a Whiteboard touchscreen in combination with a Microsoft Kinect. With StoryTec, an authoring tool for Serious Games designed for non-programmers has been shown. PEDALE provides a learning environment for peer assessment of creative tasks using StoryTec as authoring environment. 3D Multiplayer Serious Games like Woodment or Escape From Wilson Island represent examples for game-based training in collaborative learning scenarios. Furthermore, among many others, a Serious Game about architecture was exhibited as well as SHORE, a software for recognition of facial expression for interaction- and game design.

On the Science meets Business day, the importance of Serious Gaming in the State of Hesse was pointed out by Christian Flory of Hessen IT and Andreas Gelhard from the "Forum of interdisciplinary research" at TU Darmstadt.

Panel Session

As a highlight, a panel discussion entitled „Use of Serious Games for prevention and rehabilitation" was held. Experts like Dr. Jürgen Richter, chairman of the AWO (public welfare organization) in the state of Hesse, representatives of three medical centers in Darmstadt (S. Becker, Darmstädter Kinderkliniken Prinzessin Margaret; M. H.-D. Pfisterer, Agaplesion Elisabethenstift Ev. Krankenhaus und M. Held, Klinikum Darmstadt, Alten- und Pflegeheim Emilstraße), Ruth Lemmen from the Bundesverband für Unterhaltungssoftware, and Prof. Wiemeyer from the Technische Universität Darmstadt discussed the topic and emphasized the potential of Serious Games for health applications as playful instrument for therapy, rehabilitation and prevention. Also, the necessity for comprehensive, scientific sound evaluation studies has been pointed out. That kind of validated 'proof-of-concept' showcases might provide the entry point to the prospering healthcare market.

In summary, the GameDays 2012 – held in conjunction with the Edutainment 2012 – brought together scientists and practitioners in the multidisciplinary environment of Serious Games and edutainment applications. Key challenges have been addressed which build the ground for further research and the establishment of best practice showcases for the use of Serious Games in education, training or sports and health.

Further information about the GameDays 2012 and Edutainment 2012, its program, exhibits, visual impressions, etc. is available at www.gamedays2012.de.

The Edutainment 2013 will take place at September 30th to October 3rd in Sydney, Australia.

The GameDays 2013 will take place at TU Darmstadt at March 22nd-23rd (Friday, Saturday) as a "Science meets Business" event with an invited academic panel, workshops and an "Open Door Day" for the public. The exhibition will be enhanced by a Serious Games Team Challenge. In 2014, the GameDays will be held as Int'l Conference on Serious Games again.

foodQuest

Enjoy!

TU Darmstadt spin-off „foodQuest" is developing an iPhone app for personalized restaurant recommendations
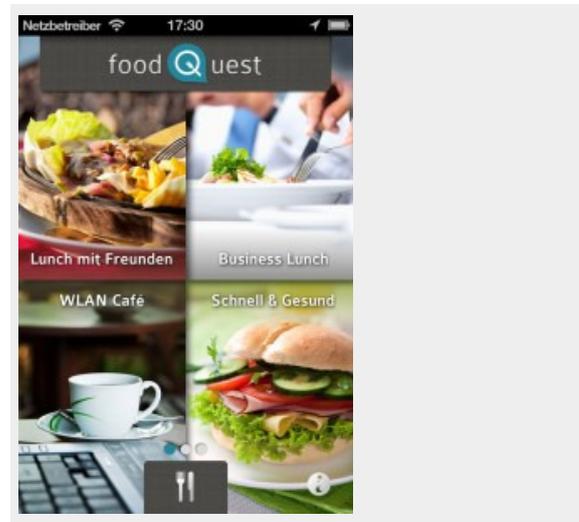
What do a student, a top-manager, a loving couple and a family have in common? Indeed, they all are hungry! And: They always want to discover great restaurants. The requirements, however, vary a lot: Low price, suitable for business talks, romantic or child-friendly. Not every restaurant is suitable for every guest and every occasion. foodQuest is the first app for restaurant recommendations that caters to the individual needs of its hungry users. The suggestions are derived through a hybrid model of automated analysis, crowdsourcing, and editorial content. In the current version, foodQuest supplies culinary assistance for the cities of Hanover and Frankfurt. The next release will feature recommendations for all of Germany. The app is available for free in Apple's App Store.

"foodQuest's core concept is to enable its users to find restaurants that fit their individual needs as well as their current situation", says Chris Chard, co-founder of foodQuest. An unhurried evening with friends, a nice café with a fast Wi-Fi connection for mobile working or a spontaneous lunch with business partners – foodQuest understands the occasion and recommends the best suitable restaurants in the vicinity. Thus, foodQuest is not only about pricy gourmet restaurants. "A great café or a passionately run takeaway can be perfect choices depending on the situation", says Chard. However, the prerequisite for any recommendation is a high quality standard.

To assure the high quality of the recommendations, foodQuest employs a hybrid model of semantics, machine learning, crowdsourcing, and editorial supervision. The technical process considers relevant information from gourmet guides, websites, blogs and rating communities. The analysis does not only rely on hard facts like location, opening hours, cuisine, and price range, but also on "soft factors" such as a pleasant discussion atmosphere, speed of service, clientele, or comfort. The resulting recommendations are further enhanced by foodQuest's community of foodies and double-checked by editorial staff. foodQuest lays great emphasis on reducing the amount of information that the user is confronted with. "For any query, we only recommend a handful of restaurants. Thus, the accuracy and quality of those recommendations is of upmost importance to us," says Ken Knoll, who co-founded the startup out of Darmstadt together with Chris Chard. The entrepreneurs are passionate about creating a great and easy user experience. Vital steps in the process of choosing a restaurant are embedded such as routing directions or booking a table online directly from within the app.

In March 2012, foodQuest was publicly presented for the first time on the CeBIT in Hannover. Ever since, the startup has close ties with its alma mater, the TU Darmstadt, and closely cooperates with the institute of Multimedia Communications. "We greatly benefit not only from the shared office space and infrastructure, but also from invaluable know how: The technologies at foodQuest's core have a significant overlap with the institute's scientific work. The collaboration with Multimedia Communications provides us with an unparalleled environment for our startup," explains Knoll.

The institute explores vital components of the Internet's future: Textual analysis, semantics, distributed sensor networks and peer-to-peer infrastructures. Professor Ralf Steinmetz, head of the institute of Multimedia Communications, recognizes mutual benefits for both sides: "We put great effort into fostering the intensive exchange between research and practice. Having young companies like foodQuest or our second spin-off "wer denkt was" on-site dramatically improves this exchange."

foodQuest is available in German and English language as a free download in the App Store. More information can be found on www.foodquestapp.com or on Facebook via www.facebook.com/foodquestapp
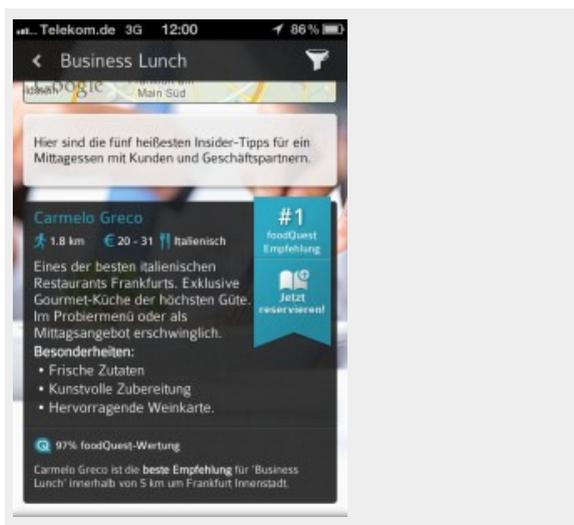
Contact:

Dipl. Wirt.-Ing. Chris Chard

foodQuest GbR

Rundeturmstr. 10

64283 Darmstadt

Email: chris@foodque.st

SIGMM Award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications

Dr. HongJiang Zhang

The 2012 winner of the prestigious Association for Computing Machinery (ACM) Special Interest Group on Multimedia (SIGMM) award for Outstanding Technical Contributions to Multimedia Computing, Communications and Applications is Dr. HongJiang Zhang. He is currently Chief Executive Officer at Kingsoft. He also holds guest professorships at Tsinghua University and Harbin Institute of Technology. The ACM SIGMM Technical Achievement award, given in recognition of outstanding contributions over a researcher's career, cites Dr. Zhang's "pioneering contributions to and leadership in media computing including content-based media analysis and retrieval, and their applications." The SIGMM award will be presented at the ACM International Conference on Multimedia 2012 that will be held Oct 29 – Nov 2 2012 in Nara, Japan.

In the early 1990s, Dr. Zhang began his pioneering work on content analysis and content-based abstraction, browsing, and retrieval of video, when these research areas were about to emerge. He established the foundations of this new research area by his numerous seminal contributions. Dr. Zhang's most noteworthy early works include the first algorithm for reliably detecting gradual video scene transitions and content-based video key-frame extraction, one of the first works on compressed domain video content analysis, as well as his structured video analysis framework and algorithms. These pioneering works have had tremendous impact on the directions, methodologies and advancements of the media computing field.

Dr. Zhang's research contributions also made a profound impact on the establishment of the ISO (International Standards Organization) MPEG-7 standard, which is the international standard that defines multimedia content descriptions.

In addition to his scholarly contributions, Dr. Zhang has significantly shaped the video indexing and editing software industry through his seminars, publications, patents and technology licensing and transfers to a number of companies and successful HP and Microsoft products. Most significant are:

1. Image Bank's video cataloging tools licensed to Image Bank. Inc.(1995);

2. Video structure parsing technologies licensed to Intel in (1996);

3. Media metadata definition and extraction algorithms in Window Imaging Platform;

4. Image search in Microsoft Digital Image Pro (2003) and web search releases; and

5. Automated video editing, a technology breakthrough that gained Microsoft MovieMaker 2.0 a five star rating by About.Com Desktop Video (2003).

In summary, Dr. Zhang's accomplishments include pioneering and extraordinary contributions to media computing and outstanding service to the computing community.

—

ACM is the professional society of computer scientists, and SIGMM is the special interest group on multimedia.

Outstanding PhD Thesis in Multimedia Computing, Communications and Applications
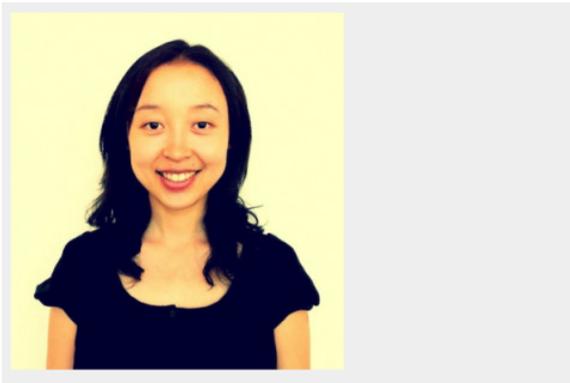
Dr. Wanmin Wu

The **SIGMM Ph.D. Award Committee** is pleased to recommend this year's award for the outstanding Ph.D. thesis in multimedia computing, communications and applications to Dr. Wanmin Wu.

Wu's dissertation documents fundamental work in the area of unifying systems and user-centric approaches to managing information flows for supporting 3D tele-immersive environments. She has developed a theoretical framework for modeling and measuring QoE, and for correlating QoE with Quality-of-Service (QoS) in distributed multi-modal interactive environments. This work has been significant in that it introduced the importance of the user-centric approach to modelling and managing complex three-dimensional data exchanges in time-constrained systems.

The committee considered the main innovations of this work to be:

1. Identifying and incorporating human psycho-physical factors along with traditional QOS to improve experience;

2. Proposing new methods and theory for QOS in interactive multi-camera environments that have served as a catalyst for enabling work in distributed education, medicine and conferencing;

3. The development of new methods for video coding incorporating understanding of users psycho-physical understanding of color and depth.

These new methods have significantly reduced the impact of sharing tele-immersive information and are likely to have a longer-term benefit that is similar to that of selective audio encoding.

The committee has considered this contribution as worthy of the award as it tackles a new problem, proposes new theory and practice as a solution to this problem area, and opens the way for further research into effective distributed three-dimensional immersive systems.

ACM TOMCCAP will accept survey papers

Dear members of the ACM Multimedia community, we are glad to let you know that the Editorial Board of ACM TOMCCAP has decided to accept survey papers. If you have such a paper or are planning to write one, please consider submitting it to our premier journal first.

Ambulant – a multimedia playback platform

Distributed multimedia is a field that depends on many technologies, including networking, coding and decoding, scheduling, rendering and user interaction. Often, this leads to multimedia researchers in one of those fields expending a lot of implementation effort to build a complete media environment when they actually only want to demonstrate an advance within their own field. In 2004 the authors, having gone through this process more than once themselves, decided to design an open source extensible and embeddable multimedia platform that could serve as a central research resource. The NLNet Foundation, www.nlnet.nl, graciously provided initial funding for the resulting Ambulant project.

Ambulant was designed from the outset to be usable for experimentation in a wide range of fields, not only in a laboratory setting but also as a deployed player for end users. However, it was not intended to compete with general end-user playback systems such as the (then popular) RealPlayer, Quicktime or the Windows Media Player. Our goal was to build a glue environment where

various research groups could plug in next approaches to media scheduling, rendering and distribution. While some effort was spent on things like ease of installation, multi-platform compatibility and user interface issues, Ambulant has never hoped to usurp commercial media players. The user interface on three different platforms can bee seen in the figure below.


The first deployment of the platform was during the W3C standardization of SMIL 2.1 and 3.0 [2, 3], when Ambulant was used to test the specification and create an open reference implementation. The fact that Ambulant supports SMIL out of the box means that it is not only useful to "low-level" multimedia researchers who want to experiment with replacing systems components, but also to people interested in semantics or server-side document generation: by using SMIL as their output format they can use Ambulant to render their documents on any platform, including inside a web browser.

Design and Implementation

Ambulant is designed so that all key components are replaceable and extensible. This follows from the requirement that it is usable as an experimentation vehicle: if someone wants to replace the scheduler by one of their own design this should be possible, and have little or no impact on the rest of the system.

To ensure wide deployability it was decided to create a portable platform. However, runtime efficiency is also an issue in multimedia playback, especially for audio and video decoding and rendering, so we decided to implement the core engine in C++. This allowed us to use platform-native decoding and rendering toolkits such as QuickTime and DirectShow, and gave us the added benefit of being able to use the native GUI toolkit on each platform, which makes life easier for end users and integrators. Using the native GUI has been a bit of extra effort up front, finding the right spot to separate platform-independent and platform-dependent code, but by now porting to a new GUI toolkit takes about three man-months. About 8 GUI toolkits have been supported over time (or 11 if you count browser plugin APIs as a GUI toolkit). The current version of Ambulant runs natively on MacOSX, Linux, Windows and iOS, and a browser plugin is available for all major browsers on all desktop platforms (including Internet Explorer on

Windows). Various old platforms (WM5, Maemo) were supported in the past and, while no longer maintained, the code is still available.
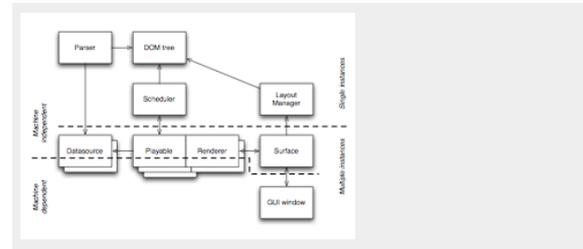
The design of Ambulant is shown in the figure above. On the document level there is a parser which reads external documents and converts them into a representation that the scheduler and layout engine will handle during document playout time. On the intermediate level there are datasources that read documents and media streams and handles them to the playout components. On the lower level there are the machine-dependent implementations of those stream readers and renderers. For each of these components there are multiple implementations, and those can easily be replaced or extended.

The design largely uses factory functions and abstract interfaces, therefore the implementation uses a plugin architecture to allow easy replacement of components at runtime without having to rebuild the complete application.

To make life even more simple, the API to the core Ambulant engine is available not only in C++ but also in Python. The Python bridge is complete and bidirectional: all classes that are accessible from C++ are just as accessible from Python and vice versa, and sending an object back-and-forth through the Python-C++ bridge results in the original object, not a new double-wrapped object. Moreover, not only can C++ classes be subclassed in Python but also the reverse. This means both extending Ambulant through a plugin and embedding Ambulant can be done in pure Python, without having to write any C/C++ code and without having to rebuild Ambulant.

Applications

Over the years, Ambulant has extensively been used for experimentation, both within our group and externally. In this section we will highlight some of these applications. The overview is not complete, but it highlights the breadth of applications of Ambulant.

One of the interests of the authors is maintaining the temporal scheduling integrity of dynamically modified multimedia presentations. In the Ambulant Annotator [4],

we experimented with using secondary screens during playback, allowing user interaction on those secondary screens to modify existing shared presentations on the main screen. The modification and sharing interface was implemented as a plugin in Ambulant, which is also used to drive the main screen. In Ta2 MyVideos [5] we looked at a different form of live modification: a personalized video mashup that was created while the user is viewing it.

Integration of live video conferencing and multimedia documents is another area in which we work. For the Ta2 Family Game project [6] we augmented Ambulant with renderers to do low delay live video rendering and digitizing, and a Flash engine. The resulting platform was used to play a cooperative action game in multiple locations. We are also using Ambulant to investigate protocols for synchronizing media playback at remote locations.

In a wholly different application area, the Daisy Consortium has used Ambulant as the basis of AMIS, www.daisy.org/projects/amis. AMIS is software that reads Daisy Books, which are the international standard for digital talking books for the visually impaired. For this project Ambulant was only a small part of the solution. The main program allows the end user, who may be blind or dyslectic, to select books and navigate them. Timed playback is then handled by Ambulant, with added functionality to highlight paragraphs on-screen as the content is read out, etc.

At a higher level, an instrumented version of Ambulant has also been deployed to indirectly evaluate social media systems. In 2004, it was submitted to the first ACM Multimedia Open Source Software Competition [1].

Obtaining and Using Ambulant

Ambulant is available via www.ambulantplayer.org, in three different forms: as a stable distribution (source and installers), as a nightly build (source and installers) and through Mercurial. Unfortunately, the stable distribution is currently lagging quite a bit behind, due to restricted manpower. We also maintain full API documentation, sample documents and community mailing lists.

Ambulant is distributed under the LGPL2 license. This allows the platform to be used with commercial plugins developed by industry partners who provide proprietary software intended for limited distribution. We

are considering a switch to dual licensing (GPL/BSD), but a concrete need has yet to arise.

The Bottom Line

Ambulant is a full open source media rendering pipeline. It provides an open, plug-in environment in which researches from a wide variety of (sub)disciplines can test new algorithms and media sharing approaches without having to write mountains of less-relevant framework code. It can serve as an open environment for experimentation, validation and distribution. You are welcome to give it a try and to contribute to its growth.

References

[1]Bulterman, D. et al. 2004. Ambulant: a fast, multi-platform open source SMIL player. In Proceedings of the 12th annual ACM international conference on Multimedia (MULTIMEDIA '04). ACM, New York, NY, USA, 492-495. DOI=10.1145/1027527.1027646

[2]Bulterman, D. et al. 2008. Synchronized Multimedia Integration Language (SMIL 3.0). W3C. URL=http://www.w3.org/TR/SMIL/

[3]Bulterman, D. and Rutledge, L. 2008. Interactive Multimedia for the Web, Mobile Devices and Daisy Talking Books. Springer-Verlag, Heidelberg, Germany, ISBN: 3-540-20234-X.

[4]Cesar, P. et al. Fragment, tag, enrich, and send: Enhancing social sharing of video. Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP (2009) vol. 5 (3). DOI=10.1145/1556134.1556136

[5]Jansen, J. et al. 2012. Just-in-time personalized video presentations. In Proceedings of the 2012 ACM symposium on Document engineering (DocEng '12). ACM, New York, NY, USA, 59-68. DOI=10.1145/2361354.2361368

[6]Jansen, J. et al. Enabling Composition-Based Video-Conferencing for the Home. IEEE Transactions on Multimedia (2011) vol. 13 (5) pp. 869-881. DOI=10.1109/TMM.2011.2159369

ACM TOMCCAP Nicolas D. Georganas Best Paper Award

In its initial year the ACM Transactions on Multimedia Computing, Communications and

Applications (TOMCCAP) Nicolas D. Georganas Best Paper Award goes to the paper *Video Quality for Face Detection, Recognition and Tracking* (TOMCCAP vol 7. no.3) by Pavel Korshunov and Wei Tsang Ooi.
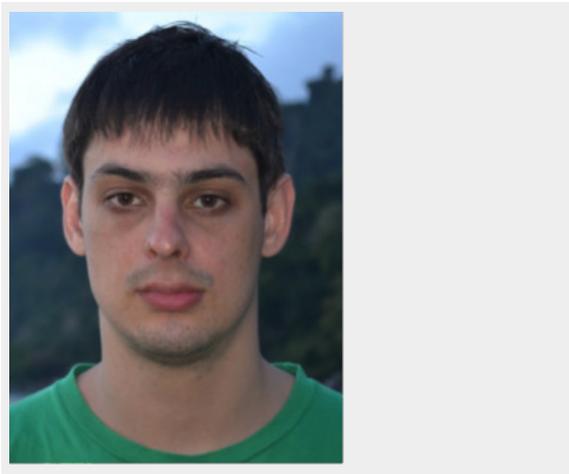
The winning paper is pioneering because it is the very first study which tries to determine an objective quality threshold value for videos used in automated video processing (AVP). The

paper proves that if a video's quality is below a certain threshold (it gives the actual values for this threshold based on video context), it cannot be used in AVP systems. Further, it is shown that

AVP systems still work with reasonable accuracy even when the video quality is low from a human's perspective. This is an important finding because it means we can reduce quality and bit rate of the

video without sacrificing accuracy, leading to reduced costs, greater scalability, and faster processing. What is unique about the paper is that it distinguishes between quality as perceived by humans, versus quality as perceived by AVP systems. In essence, the paper proposes that for AVP systems we should design machine-consumable video coding standards, not human-consumable codes.

The purpose of the award is to recognize the most significant work in ACM TOMCCAP in a given calendar year. The whole readership of ACM TOMCCAP was invited to nominate articles which were published in Volume 7 (2011). Based on the nominations the winner has been chosen by the TOMCCAP Editorial Board. The main assessment criteria have been quality, novelty, timeliness, clarity of presentation, in addition to relevance to multimedia computing, communications, and applications.

The award honors the founding Editor-in-Chief of TOMCCAP, Nicolas D. Georganas, for his contributions to the field of multimedia computing and his significant contributions to ACM. He influenced the research and the multimedia community exceedingly.

The Editor-in-Chief Prof. Dr.-Ing. Ralf Steinmetz and the Editorial Board of ACM TOMCCAP cordially congratulate the winner. The award will be presented to the authors on November 1st 2012 at the ACM Multimedia 2012 in Nara, Japan and includes travel expenses for the winning authors.

MPEG Column: 101st MPEG Meeting

MPEG news: a report from the 101st meeting, Stockholm, Sweden
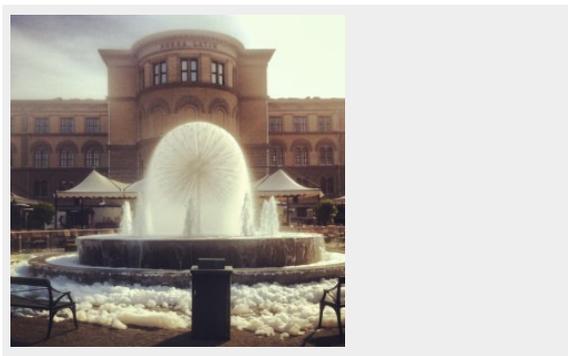
The 101st MPEG meeting was held in Stockholm, Sweden, July 16-20, 2012. The official press release can be found here and I would like to highlight the following topics:

- MPEG Media Transport (MMT) reaches Committee Draft (CD)
- High-Efficiency Video Coding (HEVC) reaches Draft International Standard (DIS)
- MPEG and ITU-T establish JCT-3V
- Call for Proposals: HEVC scalability extensions
- 3D audio workshop
- Green MPEG

MMT goes CD

The Committee Draft (CD) of MPEG-H part 1 referred to as MPEG Media Transport (MMT) has been approved and will be publicly available after an editing period which will end Sep 17th. MMT comprises the following features:

- Delivery of coded media by concurrently using more than one delivery medium (e.g., as it is the case of heterogeneous networks).
- Logical packaging structure and composition information to support multimedia mash-ups (e.g., multiscreen presentation).
- Seamless and easy conversion between storage and delivery formats.
- Cross layer interface to facilitate communication between the application layers and underlying delivery layers.
- Signaling of messages to manage the presentation and optimized delivery of media.

This list of 'features' may sound very high-level but as the CD usually comprises stable technology and is publicly available, the research community is more than welcome to evaluate MPEG's new way of media transport. Having said this, I would like to refer to the Call for Papers of JSAC's special issue on adaptive media streaming which is mainly focusing on DASH but investigating its relationship to MMT is definitely within the scope.

HEVCs' next step towards completion: DIS

The approval of the Draft International Standard (DIS) brought the HEVC standard one step closer to completion. As reported previously, HEVC shows inferior performance gains compared to its predecessor and real-time software decoding on the iPad 3 (720p,

30Hz, 1.5 Mbps) has been demonstrated during the Friday plenary [1, 2]. It is expected that the Final Draft International Standard (FDIS) is going to be approved at the 103rd MPEG meeting in January 21-25, 2013. If the market need for HEVC is only similar as it was when AVC was finally approved, I am wondering if one can expect first products by mid/end 2013. From a research point of view we know – and history is our witness – that improvements are still possible even if the standard has been approved some time ago. For example, the AVC standard is now available in its 7th edition as a consolidation of various amendments and corrigenda.

JCT-3V

After the Joint Video Team (JVT) which successfully developed standards such as AVC, SVC, MVC and the Joint Collaborative Team on Video Coding (JCT-VC), MPEG and ITU-T establish the Joint Collaborative Team on 3D Video coding extension development (JCT-3V). That is, from now on MPEG and ITU-T also joins forces in developing 3D video coding extensions for existing codecs as well as the ones under development (i.e., AVC, HEVC). The current standardization plan includes the development of AVC multi-view extensions with depth to be completed this year and I assume HEVC will be extended with 3D capabilities once the 2D version is available.

In this context it is interesting that a call for proposals for MPEG Frame Compatible (MFC) has been issued to address current deployment issues of stereoscopic videos. The requirements are available here.

Call for Proposals: SVC for HEVC

In order to address the need for higher resolutions - Ultra HDTV - and subsets thereof, JCT-VC issued a call for proposals for HEVC scalability extensions. Similar to AVC/SVC, the requirements include that the base layer should be compatible with HEVC and enhancement layers may include temporal, spatial, and fidelity scalability. The actual call, the use cases, and the requirements shall become available on the MPEG Web site.

MPEG hosts 3D Audio Workshop

Part 3 of MPEG-H will be dedicated to audio, specifically 3D audio. The call for proposals will be issues at the 102nd MPEG meeting in October 2012 and submissions will be due at the 104th meeting in April 2013. At this meeting, MPEG has hosted a 2nd workshop on 3D audio with the following speakers.

- Frank Melchior, BBC R&D: "3D Audio? – Be inspired by the Audience!"

- Kaoru Watanabe, NHK and ITU: "Advanced multichannel audio activity and requirements"
- Bert Van Daele, Auro Technologies: "3D audio content production, post production and distribution and release"
- Michael Kelly, DTS: "3D audio, objects and interactivity in games"

The report of this workshop including the presentations will be publicly available by end of August at the MPEG Web site.

What's new: Green MPEG

Finally, MPEG is starting to explore a new area which is currently referred to as Green MPEG addressing technologies to enable energy-efficient use of MPEG standards. Therefore, an Ad-hoc Group (AhG) was established with the following mandates:

1. Study the requirements and use-cases for energy efficient use of MPEG technology.

2. Solicit further evidence for the energy savings.

3. Develop reference software for Green MPEG experimentation and upload any such software to the SVN.

4. Survey possible solutions for energy-efficient video processing and presentation.

5. Explore the relationship between metadata types and coding technologies.

6. Identify new metadata that will enable additional power savings.

7. Study system-wide interactions and implications of energy-efficient processing on mobile devices.

AhGs are usually open to the public and all discussions take place via email. To subscribe please feel free to join the email reflector.

PhD Thesis Summaries

Mukesh Saini

Privacy-Aware Surveillance System Design

Mukesh Saini

Video surveillance is a very effective means of monitoring activities over a large area with cameras as extended eyes. However, this additional security comes at the cost of privacy loss of the citizens not involved in any illicit activities. Because an adversary can use prior knowledge to infer the identities of individuals in the video even in the absence of facial information, we develop a privacy-aware surveillance framework in which we identify the implicit channels of identity leakage, quantify the privacy loss through non-facial information, and propose solution to block these channels for near zero privacy loss with minimal utility loss. The proposed privacy loss model considers facial as well as non-facial information and is able to consolidate the identity leakage through multiple events and multiple cameras. Privacy loss is modelled as an adversary's ability to correlate sensitive information to the identity of the individuals in the video. Anonymity based approach is used to consolidate the identity leakage through explicit channels of bodily cues such as facial information; and other implicit channels that exist due to 'what', 'when' and 'where' information.

Moreover, any privacy preserving method usually affects the utility of the data; therefore, the choice of data transformation is paramount to ensure an acceptable tradeoff between the privacy and the utility. We propose utility models and privacy preservation framework for the applications of video surveillance and video data publication. Through experiments it is found that current privacy protection methods include high risk of privacy loss while the proposed framework provides more robust privacy loss measures and better tradeoff of security and privacy.

For more details, please email: sainimukesh@gmail.com

Robert Kuschnig

Congestion-Aware Quality-Adaptive Streaming of Scalable Video

Supervisor(s) and Committee member(s): Hermann Hellwagner (1. rapporteur), Wolfgang Effelsberg (2. rapporteur)
URL: http://www.itec.uni-klu.ac.at/bib/files/PhD_thesis_Robert_Kuschnig_2012.pdf

Internet video streaming is a hot topic in multimedia systems. This thesis introduces a new approach to video streaming based on multiple request-response streams. The novelty of this system is that it is able to make use of multiple HTTP-based request-response streams while still providing TCP-friendliness. The results indicate that the streaming system can make good use of the available bandwidth, while the number of quality switches is kept low. Additionally, this client-driven approach responds faster to changing network conditions and enables easy recovery from connection stalls or aborts, because the control loop is at the client.

Summary

Internet video streaming is a hot topic in multimedia systems. A large variety of devices (computers, mobile phones, TVs, etc.) are connected to the Internet via wired or wireless networks and are capable of receiving and decoding HD video content. To enable new services like HD video streaming (e.g., online video rental), the Internet's infrastructure was enhanced. But the Internet is still a best-effort network, which does not implement quality-of-service or admission control, resulting in time-varying bandwidth and packet delay, packet loss and network congestion. Because video streaming accounts for a considerable amount of the Internet's traffic, video streaming needs additionally to be congestion-aware, to avoid a congestion collapse of the Internet. The Transmission Control Protocol (TCP) can adapt to changing network conditions and is currently the de facto standard protocol for congestion-aware and reliable data transmission in the Internet. This fact gave TCP-based video streaming a huge momentum. Consequently, this thesis investigates TCP-based adaptive video streaming for the Internet. The main goal is to provide a solution for congestion-aware video streaming, while still being able to achieve a reasonable performance in error-prone networks. To complement existing work on congestion-aware adaptive streaming, this thesis makes six contributions. (1) The baseline performance of TCP-based adaptive streaming is identified by means of an evaluation of different adaptive streaming approaches. The results represent a reference for further investigations. (2) An investigation on the influence of TCP's behavior in presence of packet loss on the video streaming performance. (3) To overcome the shortcomings of TCP-based video streaming (single TCP connections fail to deliver a good performance in case of packet loss), a new approach to video streaming based on multiple request-response streams was introduced. The novelty of this system is that it is able to make use of multiple HTTP-based request-response streams while still providing TCP-friendliness. (4) A performance model of the HTTP-based request-response streams was developed, to estimate the influence of the system parameters and the network characteristics on the throughput performance. (5) A comprehensive evaluation of the HTTP-based request-response streams under diverse network conditions was conducted, to validate the model's estimations. Additionally, the TCP-friendliness was evaluated, showing that request-response streaming systems can be configured to achieve TCP-friendliness. (6) A cellular network with high bandwidth fluctuations and RTTs was used to investigate the performance of the request-

response streaming system in a mobile video streaming scenario. The results indicate that the streaming system can make good use of the available bandwidth, while the number of quality switches is kept low. While aggregating multiple TCP connections to improve the TCP streaming performance is quite common, usually the improvement comes at the cost of high deployment effort. By placing the streaming logic at the client, request-response streams can avoid this complexity. Additionally, this client-driven approach responds faster to changing network conditions and enables easy recovery from connection stalls or aborts, because the control loop is at the client. To improve the network efficiency and the scalability in terms of number of clients served, HTTP-based request-response streams can utilize HTTP proxies and caches.

Multimedia Communication (MMC)
URL: http://www.uni-klu.ac.at/tewi/inf/itec/mmc/

The research group "Multimedia Communication (MMC)" was founded and is being led by Prof. Hermann Hellwagner. In addition, the group currently has three research assistants, five project staff members, and three administrative and technical staff members.

The research activities of the group are in the areas of

- Multimedia communication and QoS provisioning
- Adaptation of multimedia content w.r.t. network, device, and usage contexts
- Use of Scalable Video Coding (SVC) technology in networks and P2P systems
- Adaptive multimedia applications, e.g., IPTV
- Standardization within ISO/IEC MPEG
- Multimedia in disaster management

The focus of the MMC group is clearly on adaptive delivery of audio-visual contents, taking into account, for instance, fluctuating network and environmental conditions that can occur when users are on the move. In particular, the group is currently investigating the use of Scalable Video Coding (SVC) technology in such networks. The group actively participates in several international and national research projects on all levels, ranging from basic research to application-oriented projects and direct cooperation with industry. In teaching, the MMC group covers the technical courses of the Informatics study programme such as Computer Organization, Operating Systems, Computer Networks, Servers and Clusters, Internet QoS, and Multimedia Coding.

Recently published

MMSJ, Volume 18, Issue 5

Editor-in-Chief: Thomas Plagemann
URL: http://www.springer.de/
Published: October 2012

The Multimedia Systems Journal is the first journal publication sponsored by ACM SIGMM and today published exclusively by Springer Verlag. As a service to Records readers, we provide direct links to Springer Verlag's Digital Library for the papers of the latest MMSJ issue.

- Zhao-Guang Liu, Yang Yang, Xiu-Hua Ji: **Fast macroblock mode decision for H.264/AVC baseline profile video transcoder based on support vector machines**
- Noreen Imran, Boon-Chong Seet, Alvis C. M. Fong: **A comparative analysis of video codecs for multihop wireless video sensor networks**
- Mingfu Li, Tsung-Wei Lin, Shih-Han Cheng: **Arrival process-controlled adaptive media playout with multiple thresholds for video streaming**
- Pejman Goudarzi: **Quality enhancement for scalable video delivery over wireless networks based on particle swarm optimization**
- Navin Sharma, David Irwin, Michael Zink, Prashant Shenoy: **MultiSense: proportional-share for mechanically steerable sensor networks**

MMSJ, Volume 18, Issue 6

Editor-in-Chief: Thomas Plagemann
URL: http://www.springer.de/
Published: November 2012

The Multimedia Systems Journal is the first journal publication sponsored by ACM SIGMM and today published exclusively by Springer Verlag. As a service to Records readers, we provide direct links to Springer Verlag's Digital Library for the papers of the latest MMSJ issue.

- Nicolas Staelens, Jonas De Meulenaere, Lizzy Bleumers…: **Assessing the importance of audio/video synchronization for simultaneous translation of video sequences**
- Mario Montagud, Fernando Boronat, Hans Stokking, Ray van Brandenburg: **Inter-destination multimedia synchronization: schemes, use cases and standardization**

- Naveed Ejaz, Sung Wook Baik: **Video summarization using a network of radial basis functions**
- Ryan P. Hafen, Michael J. Henry: **Speech information retrieval: a review**
- Thomas Haenselmann, Wolfgang Effelsberg: **Optimal strategies for creating paper models from 3D objects**

TOMCCAP, Volume 8, Issue 2S

Editor-in-Chief: Ralf Steinmetz Guest editor: Mohan Kankanhalli
URL: http://tomccap.acm.org/
Published: September 2012
sponsored by ACM SIGMM

The Transactions on Multimedia Computing, Communication and Applications are the SIGMM's own Transactions. This issue is one of TOMCCAP's supplemental online-only issues. Available in ACM's Digital Library, it is a special issue on multimedia security.

- Mohan S. Kankanhalli: **Introduction to special issue on multimedia security**
- Jonathan Weir, Weiqi Yan, Mohan S. Kankanhalli: **Image hatching for visual cryptography**
- Jian Li, Hongmei Liu, Jiwu Huang, Yun Q. Shi: **Reference index-based H.264 video watermarking scheme**
- Xifeng Gao, Caiming Zhang, Yan Huang, Zhigang Deng: **A robust high-capacity affine-transformation-invariant scheme for watermarking 3D geometric models**
- Rui Yang, Zhenhua Qu, Jiwu Huang: **Exposing MP3 audio forgeries using frame offsets**
- Hui Feng, Hefei Ling, Fuhao Zou, Weiqi Yan, Zhengding Lu: **A collusion attack optimization strategy for digital fingerprinting**
- Amit Sachan, Sabu Emmanuel, Mohan S. Kankanhalli: **Aggregate licenses validation for digital rights violation detection**

TOMCCAP, Volume 8, Issue 3

Editor-in-Chief: Ralf Steinmetz
URL: http://tomccap.acm.org/
Published: July 2012
sponsored by ACM SIGMM

The Transactions on Multimedia Computing, Communication and Applications are the SIGMM's own Transactions. As a service to Records readers, we provide direct links to ACM Digital Library for the papers of the latest TOMCCAP issue.

- Haakon Riiser, Tore Endestad, Paul Vigmostad, Carsten Griwodz, Pål Halvorsen: **Video streaming using a location-based bandwidth-lookup service for bitrate planning**
- Victor Valdes, Jose M. Martinez: **Automatic evaluation of video summaries**
- Xinmei Tian, Dacheng Tao, Yong Rui: **Sparse transfer learning for interactive video search reranking**
- Xin Zhang, Tomás E. Ward, Séamus Mcloone: **An information-based dynamic extrapolation model for networked virtual environments**
- Linjun Yang, Bo Geng, Alan Hanjalic, Xian-Sheng Hua: **A unified context model for web image retrieval**
- Paul Patras, Albert Banchs, Pablo Serrano: **A control theoretic scheme for efficient video transmission over IEEE 802.11e EDCA WLANs**
- Xinglei Zhu, Chang W. Chen: **A joint layered scheme for reliable and secure mobile JPEG-2000 streaming**

TOMCCAP, Volume 8, Issue 3S

Editor-in-Chief: Ralf Steinmetz Guest editors: Daniel Gatica-Perez, Gang Hua, Wei Tsang Ooi, Pål Halvorsen
URL: http://tomccap.acm.org/
Published: September 2012
sponsored by ACM SIGMM

The Transactions on Multimedia Computing, Communication and Applications are the SIGMM's own Transactions. This issue is one of TOMCCAP's supplemental online-only issues. Available in ACM's Digital Library, it is a special issue containing the best papers of ACM Multimedia 2012.

- Daniel Gatica-Perez, Gang Hua, Wei Tsang Ooi, Pål Halvorsen: **Introduction to the special section of best papers of ACM multimedia 2011**
- Wanmin Wu, Ahsan Arefin, Gregorij Kurillo, Pooja Agarwal, Klara Nahrstedt, Ruzena Bajcsy: **CZLoD: A psychophysical approach for 3D tele-immersive video**
- Rongrong Ji, Felix X. Yu, Tongtao Zhang, Shih-Fu Chang: **Active query sensing: Suggesting the best query view for mobile visual search**
- Shervin Shirmohammadi, Mohamed Hefeeda, Wei Tsang Ooi, Romulus Grigoras: **Introduction to special section on 3D mobile multimedia**
- Yanwei Liu, Song Ci, Hui Tang, Yun Ye, Jinxia Liu: **QoE-oriented 3D video transcoding for mobile streaming**

- Shujie Liu, Chang Wen Chen: **A novel 3D video transcoding scheme for adaptive 3D video transmission to heterogeneous terminals**
- Hoda Roodaki, Mahmoud Reza Hashemi, Shervin Shirmohammadi: **A new methodology to derive objective quality assessment metrics for scalable multiview 3D video coding**
- Ahmed Hamza, Mohamed Hefeeda: **Energy-efficient multicasting of multiview 3D videos to mobile devices**
- Shu Shi, Klara Nahrstedt, Roy Campbell: **A real-time remote rendering system for interactive mobile graphics**
- Wei Guan, Suya You, Ulrich Newmann: **Efficient matchings and mobile augmented reality**

Job Opportunities

Full Professor in Computer Graphics

The Department of Information and Computing Sciences at Utrecht University has a strong focus on game technology. We offer a special bachelor track in this area with a yearly enrollment of over 100 students and a master program with a yearly enrollment of about 45 students. This focus also shows in the research topics we study. These include computer graphics, simulation, interaction technology, artificial intelligence, and software technology. We take a strong computer science perspective in our teaching and research. We believe in training specialists, not generalists.

Employer: Utrecht University
Expiration date: Thursday, November 15, 2012
More information date: http://www.cs.uu.nl/vacatures/en/440794.html

PhD/Research Position on Video Mining

The Multimedia Information Systems Research Group at the University of Vienna is seeking to recruit a Research Assistant in the field of multimedia retrieval.

You will work on project funded by the Vienna Science and Technology Fund (WWTF) and focused on video mining in media repositories. The main objective of the project is the detection of unusual and, thus, potentially interesting sequences in large video collections. Core components include dynamic feature selection, multimodal sequence representation, and data mining and clustering. Candidates should

be qualified for and willing to perform state-of-the-art research, and to acquire a PhD degree in the above topics.

Employer: University of Vienna, Austria
Expiration date: Friday, November 30, 2012
More information date:
mailto:maia.zaharieva@univie.ac.at

Research Fellow and Postdoctoral positions in computer vision and wearable sensing for bio-motion analysis

As a result of recent research funding success, the CLARITY centre at DCU seeks both a Postdoctoral Researcher and a Research Fellow to work in the areas of computer vision and wearable sensors for motion capture and reconstruction. CLARITY is an SFI-funded CSET (Centre for Science Engineering and Technology), a research centre that focuses on the intersection between two important research areas – Adaptive Sensing and Information Discovery. It is a partnership between University College Dublin, Dublin City University and the Tyndall National Institute. CLARITY employs over 100 researchers including PhD students, postdoctoral researches and administrative staff, across the three sites and has over 40 ongoing research projects funded by SFI, Enterprise Ireland, Marine Institute, Environmental Protection Agency, EU FP7, European Space Agency, and directly by our industry partners.

The position is funded as part of 3-year EU FP7 project that will investigate low-cost approaches to motion capture and synthesis/reconstruction for traditional sports and games using both computer vision and wearable inertial sensors.

The multidisciplinary project consortium features both academic and industrial partners with world-class expertise in computer vision, graphics, motion capture and multimedia information retrieval, along with high-profile end user partners. Working closely with sport science researchers in the School of Health and Human Performance, Dublin City University and collaborating with national sporting organizations around Europe, research will target novel forms of motion capture for both elite and non-elite athletes.

The successful candidate should have completed a PhD in computer vision / pattern recognition or wearable inertial sensing or a closely related

discipline and have an interest in human motion analysis, video tracking, body sensor networks and information fusion. Candidates' publication records will be carefully considered as part of the selection process. Candidates will be expected to work closely with other project partners including traveling to partners' sites to represent DCU's contribution to the project. Experience in international collaboration and research management would be an advantage.

Employer: CLARITY: Centre for Sensor Web Technologies , Dublin City University
Expiration date: Monday, November 12, 2012
More information date: mailto:Noel.OConnor@dcu.ie

Calls for Contribution

CFPs: Sponsored by ACM SIGMM

ICMR 2013

ACM International Conference on Multimedia Retrieval

Submission deadline: 15. October 2012
Location: Dallas, TX, USA
Dates: 16. April 2013 -19. April 2013
More information: http://www.icmr2013.org/
Sponsored by ACM SIGMM

ICMR 2013 is seeking original high quality submissions addressing innovative research in the broad field of multimedia retrieval. We wish to highlight significant contributions addressing the main problem of search and retrieval but also the related and equally important issues of multimedia content management, user interaction, and community-based management.

NOSSDAV 2013

ACM Networks and Operating Systems Support for Digital Audio and Video

Submission deadline: 24. November 2012
Location: Oslo, Norway
Dates: 27. February 2013 -01. March 2013
More information: http://nossdav2013.ndlab.net
Sponsored by ACM SIGMM

The 23nd ACM NOSSDAV will continue to focus on both established and emerging research topics, high-risk high-return ideas and proposals, and future research directions in multimedia networking and systems, in a single-track format that encourages active participation and discussions among academic and industry researchers and practitioners. It will be co-located … Read more →

Back Matter

Notice to Contributing Authors to SIG Newsletters

By submitting your article for distribution in this Special Interest Group publication, you hereby grant to ACM the following non-exclusive, perpetual, worldwide rights:

Impressum