WELCOME TO *Fabulous* CVPR 2016 LAS VEGAS

THURSDAY June 30

CVPR Daily

Today's Picks

See you all at CVPR2017!

CVPR Expo

Highlights

Women in Computer Vision

Presentations

For Today, Thursday 30

David Fleet is Professor of Computer Science at the **University of Toronto**, and Chair of Computer and Mathematical Sciences at the **University of Toronto Scarborough**. David accepted to share his picks for today with **CVPR Daily**. Here are David's picks for Thursday 30. Don't miss them!

"I have not read these papers in any detail yet, but I am interested in learning more"

Morning:

09:00AM to 10:18 AM
Non-Rigid Reconstruction and Motion Analysis
Oral session O4-1B-17:
Full Flow: Optical Flow Estimation By Global Optimization Over Regular Grids
Page 40 of the Pocket Guide and presentation 17 on CVPR's website

Afternoon:

1:45PM to 2:50PM
Learning and CNN Architectures
Oral session O4-2A-2
Structural-RNN: Deep Learning on Spatio-Temporal Graphs
Page 43 of the Pocket Guide and presentation 2 on CVPR's website

3D Shape Reconstruction
Oral session O4-2B-17
SVBRDF-Invariant Shape and Reflectance Estimation From Light-Field Cameras
Page 43 of the Pocket Guide and presentation 17 on CVPR's website

3:45PM to 2:50PM
3D, Stereo, Matching, and Saliency Estimation
Spotlight session S4-3B-45
FANNG: Fast Approximate Nearest Neighbour Graphs
Page 44 of the Pocket Guide and presentation 45 on CVPR's website

People and Faces
Spotlight session S4-3A-34
Fits Like a Glove: Rapid and Reliable Hand Shape Personalization
Page 44 of the Pocket Guide and presentation 34 on CVPR's website

Dear Reader,

This is the last issue of **CVPR Daily** for this year. Thus, even though **CVPR 2016** is still in full swing, we would like to think for a moment about what we've experienced here in Las Vegas.

Everyone agrees about the **exceptional level of the technical program**, including the award-winning works:

1) Deep Residual Learning for Image Recognition, by Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun;
2) Sublabel-Accurate Relaxation of Nonconvex Energies, by Thomas Möllenhoff, Emanuel Laude, Michael Moeller, Jan Lellmann and Daniel Cremers;
3) Structural-RNN: Deep Learning on Spatio-Temporal Graphs; by Ashesh Jain, Amir R. Zamir, SilvioSavarese and Ashutosh Saxena.

Of course there are many other excellent papers featured at the conference. Through hundreds of presentations, we have discovered the amazing talent of many outstanding researchers. In its 4 first-ever issues, CVPR Daily has done its best to uncover, at least partially, all this talent and expose the work of **more than 100 scientists**.

One of the goals of this magazine is to put some name and stories with the many faces in the crowd. We deliberately chose to highlight the diversity of people who are part of our research community, in particular introducing you to **the many brilliant women at this year's CVPR**.

Finally, one word about CVPR's growth. We had 2800 attendees in 2015; **we are at 3609 attendees this year**. That makes a year over year growth of nearly 30%. This exceptional progression enables us to look forward to the future of our community with full confidence and enthusiasm.

We hope the last days of CVPR 2016 are fruitful for you. We believe this week has created countless channels of communication and strengthened even more the bonds of the **international computer vision community**. We wholeheartedly hope to see you all again at CVPR 2017!

Olga Russakovsky
Press Chair, **CVPR 2016**
Robotics Institute: **Carnegie Mellon University**

Ralph Anzarouth
Editor, **Computer Vision News**
Marketing Manager, **RSIP Vision**

Danna Gurari

"*I love what I do*"

CVPR Daily: Danna, what is your current position?

Danna Gurari: I am currently a post-doctoral fellow with Kristen Grauman at University of Texas at Austin.

CVPR Daily: What is the subject of your work?

Danna: I've been working at lot with microscopy images as well as everyday images. I look at how to actively involve humans to solve tasks such as object segmentation and visual question and answering.
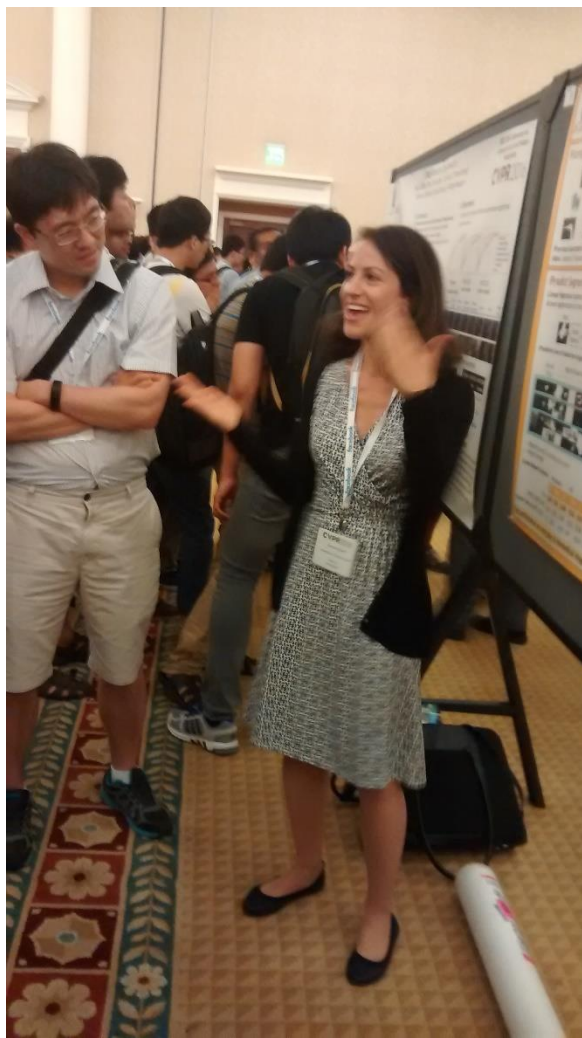
CVPR Daily: Why did you decide to dedicate your career to computer vision?

Danna: I think that's just an instinctual love.

CVPR Daily: Do scientists follow their instincts too?

Danna: I think it better be something you love when it comes to what you do every day when you wake up and what's going to motivate you to do your work.

CVPR Daily: Do you love what you do?

Danna: I love what I do.

CVPR Daily: Is it difficult to be a woman in computer vision? If yes, why?

Danna: I would say it's been a lot easier for me in academia than in industry. I worked in industry for five years, and I found that I was typically alone as a female. That was harder.

In academia, I've been very fortunate because my advisor for my PhD as well as my current postdoc advisor are female. I find, in general, the community really fosters and creates a safe place for people my age who are females.

CVPR Daily: That's very interesting because several times I've heard women saying that being the only woman was a limitation. But if I were the only man among many women, I wouldn't feel isolated. Why is it so different?

Danna: I would say the life challenges are different for men and women. Long term, a lot of us in academia think of our academic career and how to balance that with a home life. I also think there are other female aspects. There's something intangible about having someone you can relate to that way. When I think about a role model, typically I feel more connected to a female. That story matches mine better**.**

As a counterpoint, my father was a faculty in computer science. He was one of my main mentors. He's been a strong male presence in my life, especially in computer science.

"I worked in industry for five years, and I found that I was typically alone as a female. That was harder"

"*I would like to have my work help save lives*"

CVPR Daily: What would you like to achieve in your career?

Danna: I would like to have my work help save lives. I work a lot in the microscopy and the medical imaging space. My greatest wish is that something I do will in some way help save someone's life.

CVPR Daily: What are the chances that you give yourself to be able to fulfil this hope one day?

Danna: I think it's quite possible. As a mentor, I can be a voice that creates a movement for this area. Whether it's me or someone I teach, both feel equally good.

CVPR Daily: You may have noticed that CVPR Daily highlights the diversity of people who are part of our research community, in particular introducing the many brilliant women at this year's CVPR. What do you think of that?

Danna: I trust those that are making the decisions. It's interesting to me, and I don't think I have the full picture. I just trust that there's some value in trying to put a lot of female faces out there. Maybe it's just a small thing of making the younger females coming to CVPR feel like they fit in. It could make a long term impact on females going forward. I'm not really sure of the intention, but I can imagine good outcomes from it.

Qihoo 360

Shuicheng and **Cheong** invited me to see **something very special** at the Qihoo 360 booth and I was not disappointed. Follow the guide!

They showed me a very good **facial analysis system**. Very accurate, stable, and energy efficient. Accurate means that if you stand in front of the camera, the position or the important points will be very accurate.

Second, it will be very stable. Stable also means when you speak, all the points will stick on your face with markers. It is also very robust, in the sense that when only a part of the face is visible, even only one quarter of it, these points are still very accurate and stable.

With this kind of functions, many things become possible: make a face look thinner or enlarge the eyes to make it look more beautiful. In Asian culture, Shuicheng told me, thin face and big eyes make one's aspect beautiful. You can also perform analysis, for example, of the age or gender, whether you wear glasses, how beautiful you are, etc. **Your humble servant tried this out**, see the results in the bottom right image!

The Qihoo/360 AI Institute claims a face SDK which is very stable in regards to head movement, expression changes and pose changes, as well as robust to partial visibility of the face. These features are critical for face manipulations/editing like mask transfer (see image below).

Mask transfer effect
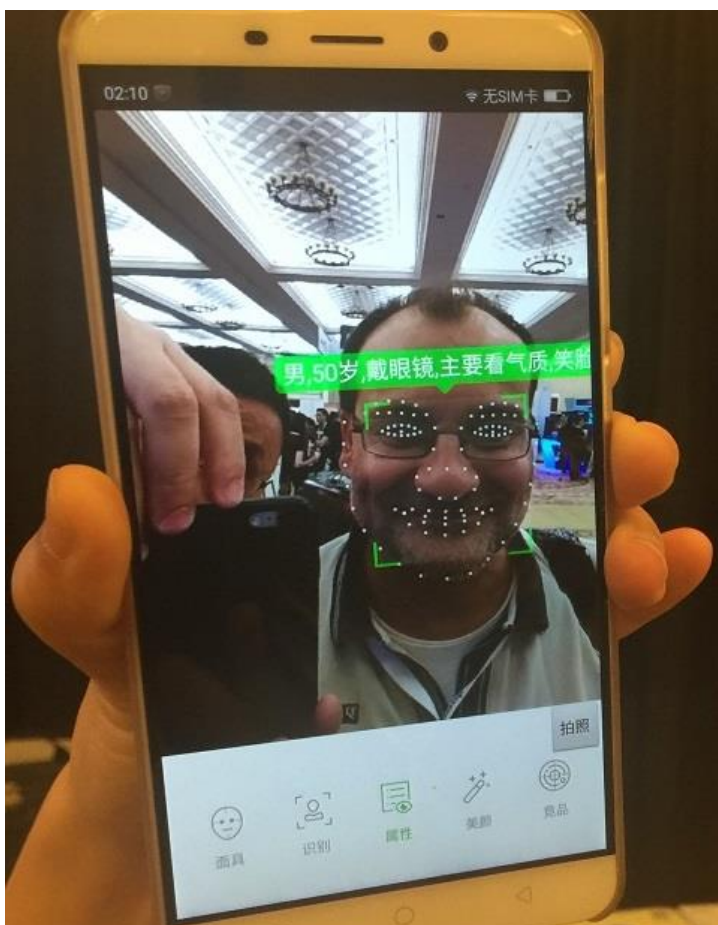
This is what Qihoo's software says about me! I look 2 years younger than I am, I wear glasses, I am a gentleman (of course, I am Italian!) and my face is smiling ☺

Maros Blaha and Audrey Richard

The **Large-Scale Semantic 3D Reconstruction** project was presented both at the orals and as a poster. What it does is automatic generation of semantically annotated **3D city models**, starting from large image collections. The idea of semantic 3D reconstruction is to recover the geometry of an observed scene while at the same time also interpreting the scene in terms of semantic object classes (such as buildings, vegetation etc.).
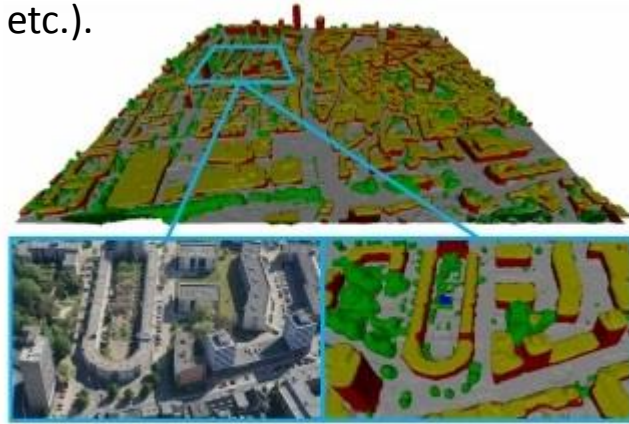

Figure 1: Semantic 3D model of the city of Enschede generated with the proposed adaptive multi-resolution approach.

While jointly inferring 3D shapes and semantic classes delivers appealing results, up-to-date methods are memory-hungry and computationally expensive. This approach demonstrates how to scale semantic 3D reconstruction up to be able to reconstruct large geographic areas such as cities. The resulting superior 3D models allow for advanced reasoning tasks such as urban planning, navigation, physics-based simulations and the like.
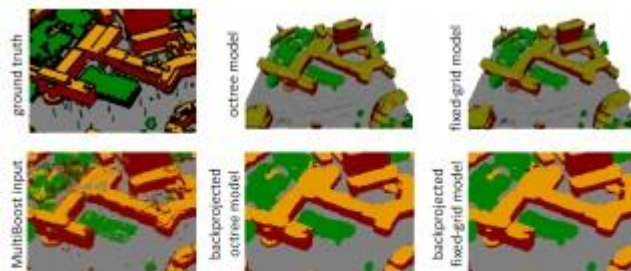

Figure 5: Comparison of the labeling accuracy. Colors indicate ground (*gray*), building (*red*), roof (*yellow*), vegetation (*green*) and clutter (*blue*).

This method is not only limited to 3D city modelling: possible further applications include a number of multi-class segmentation problems, e.g. in the field of medical imaging. This six-teamer work is available in full here.

Is There a Scientist in the Room?

This is what it looks when many of the 3,609 CVPR attendees, coming from countless countries and from all the continents, have lunch together. CVPR Daily owes this lovely image of friendship and comradery to the kindness of Dave Rushton-Smith of Google.

Spare5

Besides distributing T-shirts, at the Spare5 booth they talk about their service enabling machine learning teams to generate accurate and diverse annotations on their datasets to train, validate, and test their algorithms.

Their **Training Data as a Service** platform is a complete solution for acquiring and annotating datasets and building ground truth at scale. I was told there that leaders in machine learning rely on Spare5's automated platform for efficiencies in accuracy, specificity, scale, cost, and reduced level of effort in computer vision and natural language processing applications.

TuSimple

If you ask why the company is called **TuSimple**, this must have something to do with "**Technology made simple**". What they do is **ADAS (Advanced Driver Assistance System)**, autonomous driving, and web-scale image recognition applications.

We asked them about that and **Amanda Song** explained that their ADAS/autonomous driving system incorporates state-of-the-art modules for multi-class object detection, lane detection, depth estimation, and scene analysis. Optimized for complex road conditions in cluttered urban scenes, this technology can accurately assess driving risks from **pedestrian, bicycle, motorcycle, cars, and other types of obstacles**. The solution being 100% vision-based, this means dramatic reduction in hardware cost, comparing to today's LiDAR-based perception solutions. The company is headquartered in **Beijing, China**; with its research institute in **San Diego, USA**.

Before you ask, they are looking for talented minds and innovative hearts in CVPR community, with specializations in computer vision, machine learning, and GPU optimization - both in San Diego and Beijing.

NB: Today (Thursday) is the last day to vote for the People's Choice award. Make your vote count!

We noticed a cute demo at MorpX and decided to go and see it. Tian Li was kind enough to tell us all about it.

Tian Li: I am the founder of MorpX. We are making a vision sensor for toy robots. The vision sensor uses machine learning to detect different objects like faces, bars, and lines using them to control the toy robots.

CVPR Daily: What is novel in it?

Tian Li: The novelty is that it's making all of these algorithms run inside tight, small packaging embedded system. Also, its arm contact sensor with 72 megahertz frequency.

CVPR Daily: What was particularly challenging in doing this?

Tian Li: The challenge was making all the right hardware and software because the existing library cannot fit into these kind of systems.

CVPR Daily: Does it make any difference to work with toys?

Tian Li: There are lots of fun stories because we are working with toys. When the algorithm is defected, then everything falls apart! So it's fun.

This black little robot has learned to hit down only when he is shown a ball. He does not react to any other object, no matter how hard we try!

"everybody plays in the office"

CVPR Daily: So everybody plays in the office?

Tian Li: Yea!

CVPR Daily: What is the next step for your application?

Tian Li: We are trying to push these products to classrooms for kids to have easy access to vision sensors to create some more fun robots. Our product is already in the market and you can reserve it online right now at our website.

Àngels Ratés Borràs

This is our last issue of CVPR daily for this year, so we have decided to host two Women in computer Vision. After Danna Gurari at page 4, here is Àngels Ratés Borràs.

CVPR Daily: Àngels, what kind of work do you do?

Àngels: I am doing a PhD at Northeastern University under the supervision of Octavia Camps. My field of research is computer vision. That's why I'm here.

CVPR Daily: Why did you choose to dedicate your career to computer vision?

Àngels: I first studied electrical engineering back in Barcelona. Then I did an exchange at Northeastern University. I realized that I wanted to work with something related to artificial intelligence because I really enjoy technology as well as science and maths. I met Octavia and I liked her so much that I decided to join the lab.

CVPR Daily: Do you find it more difficult to be a woman in this field?

Àngels: In terms of studying and getting opportunities, I don't think so. I believe that if you're good, they're going to hire you. In terms of your team at work, there might be some differences between a woman and a man.

CVPR Daily: Can you tell us any stories about that?

Àngels: I had a very fun college experience. Back in Barcelona, you have to adapt yourself to work with men because they work differently. Sometimes, they don't know how to work together or how to communicate with each other. Usually the women are the ones that are in charge of organizing everything and chasing them to work…

CVPR Daily: Can you tell me the best teaching that you received from Octavia?

Àngels: Of course, her course at Northeastern. The way she teaches is so lovely. She has a lot of patience. You can see in her face that she's enjoying what she's doing, and that's the most important thing. When she teaches, she makes you feel willing to learn and to listen to her. You always want to pay attention. In general, she's an amazing teacher.

CVPR Daily: Do you see yourself teaching one day?

Àngels: No, not for me.

CVPR Daily: What would you like to achieve in this field?

Àngels: I would like to learn a lot, that's why I'm doing a PhD. I would like to publish someday. I would also like to work in a nice company.

CVPR Daily: Do you think that you'll publish a paper one day?

Àngels: Yes, I would like to publish something in computer vision, for sure.

"I wanted to work with something related to artificial intelligence because I really enjoy technology as well as science and maths"

CVPR Daily: What chance do you think you have of achieving this goal?

Àngels: For sure, I will do it!

Guy Rosman - CSAIL/MIT

This project creates **adaptive 3D sensors based on information maximization**. Instead of treating 3D sensors as a black-box and asking what to do with the 3D data we get or where to move the sensor to learn more about the scene, we are asking: if I already know something about the environment, and I have a specific task in mind, how should the sensor operate?

The **CSAIL / MIT team** focused on structured light scanners, that emit/project light patterns over time and acquire images. The question is: which pattern should the sensor project next? For example, in order to reconstruct a specific object that moves, you may not need to illuminate the whole scene. If you just need to localize in a known environment or help a robot avoid obstacles, you may need radically different patterns than if you want full scene reconstruction for augmented reality.

This research paper explores such options based on information maximization and sensor planning. Taking these concepts from decision theory and robotics, it is shown that with the right probabilistic model, they can be used inside the sensor as well.

Scene | Scene w/ Pattern | Range Image | Pixelwise Information
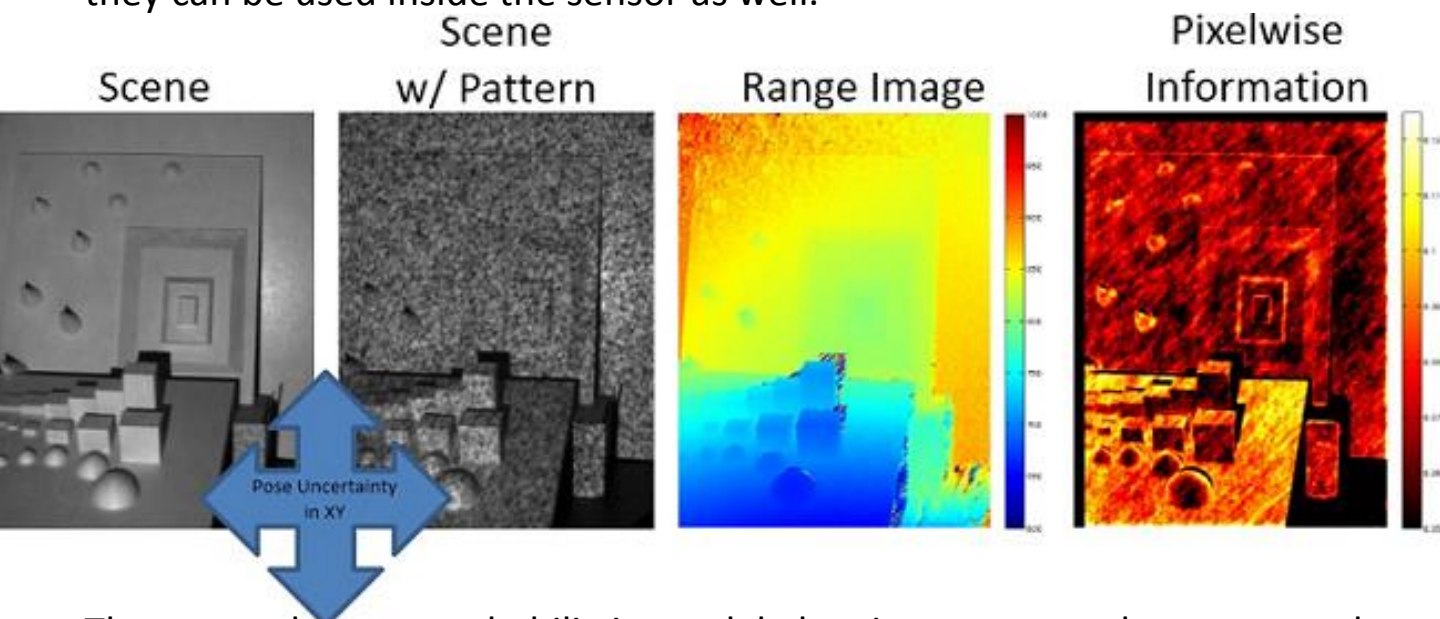
The team chose a probabilistic model that incorporates the scene and sensor pose uncertainty, and yet allows to approximate the information gain between the acquired images and subsets of variables in the scene such as the sensor pose or aspects of the geometry. The model allows to do so in a highly-parallel way, which, it is hoped, will make it useful for real systems.

Adaptive 3D scanners, and the concepts shown are expected to result in more efficient and accurate sensors that are better suited to the kind of multiple roles we expects robots and mobile devices will play in the future.