ECCV DAILY

THE DAILY MAGAZINE OF THE EUROPEAN CONFERENCE ON COMPUTER VISION

WEDNESDAY
October 12

Oral Sessions and Posters

Women in Computer Vision

Today's Picks Demos
Yesterday's Highlights

In cooperation with

Computer Vision News

The Magazine of The Algorithm community

A publication by RSIP VISION

Xiaoyong's Picks

For Wednesday 12

We have asked Xiaoyong Shen to recommend his **ECCV 2016** picks for today, **Wednesday 12**. Xiaoyong is Postdoctoral Fellow at the Chinese University of Hong Kong. Here are his picks:

Morning:

09:00 to 10:00 O-2A-2, O-2A-2 and O-2A-4 *Page 68 of the Pocket Guide*

- ℓ^0-Sparse Subspace Clustering
- Normalized Cut meets MRF
- Fast Global Registration

10:00 to 10:30 S-2A-7 and S-2A-9 *Page 68 of the Pocket Guide*

- Efficient Continuous Relaxations for Dense CRF
- A Convex Solution to Spatially-Regularized Correspondence Problems

Afternoon:

14:00 to 15:00 O-2B-1, O-2B-2, O-2B-4 *Page 71 of the Pocket Guide*

- The Fast Bilateral Solver
- Phase-based Modification Transfer for Video [with Simone Meyer]
- Focal flow: Measuring depth and velocity from defocus and differential motion [with Emma Alexander - read about her in the next pages]

15:00 to 15:30 S-2B-5, S-2B-7 *Page 71 of the Pocket Guide*

- An evaluation of computational imaging techniques for heterogeneous inverse scattering
- Fast Guided Global Interpolation for Depth and Motion

Do you have any completely subjective **ECCV 2016** picks for tomorrow? Tell us about it, **anytime today!**

Welcome

ECCV Daily
Editor: **Ralph Anzarouth**
Publisher: **RSIP Vision**

Dear reader,

ECCV 2016 in Amsterdam is in full motion. We have been presented work of great quality and this will continue during the whole week.

This **ECCV Daily** magazine will accompany this progress until Friday, trying to bridge different realities: on one hand, by bridging between the **academia** and the **industry**, as these pages keep hosting both; on the other hand, by allowing people who were not able to be with us at ECCV 2016 to be informed nearly in real time of what is happening here and what innovations are being presented.

This **ECCV Daily of Wednesday** features new interventions which will enrich the reader. First, at page 4, the outstanding (in our opinion) work of the **Focal Flow** team that will be presented to us today at 14:00. Second, at page 10, an exceptionally sensitive and revealing interview for our **Women in Computer Vision** section: we recommend that every one, man and woman, takes the time to read the very sincere and enlightening testimony of **Zeynep Akata**. Don't miss it!

Finally, we are preparing great new stories for the coming days, as well as what will be (in our view) **a major surprise**. Stay tuned on ECCV Daily and enjoy the reading!

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

Our Pick for Today

Focal Flow - Emma Alexander, Harvard University

Among the many projects presented at **ECCV2016**, **RSIP Vision's engineers** picked a few which they found particularly impressive, even for ECCV standards. **Emma Alexander**, a PhD student at **Harvard University**, agreed to tell us about **Focal Flow: Measuring depth and velocity from defocus and differential motion**. Oral Session 2B-04, today 14:00 - 15:00

ECCV Daily: *What is the Focal Flow work you are presenting, Emma?*

Emma: It is a new depth queue that can be measured very efficiently. Our model is just a thin lense camera that either is on a moving platform or observes a moving scene. We show that under Gaussian Blur, and Gaussian Blur only, there's a very simple linear constraint that relates image derivatives to a four-vector which is a similar computation to optical flow with just a 4x4 linear system over each patch.

ECCV Daily: *Why is this work needed?*

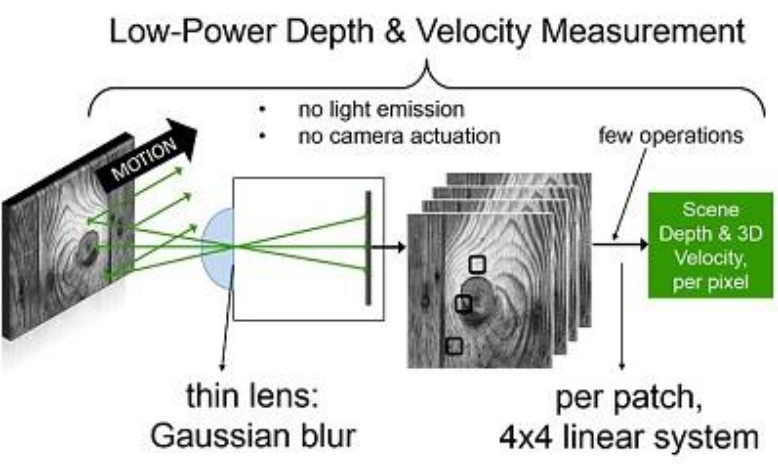
Emma: We were inspired by this new generation of micro-scale platforms that's being developed. Computer vision has done a great job at scaling up. We're great with big data. We're great with big systems with a lot of power and computation. We're starting to find out how to sacrifice computation to work on smaller systems with smaller power budgets.

"Computer vision doesn't have anything to offer these robots"

There are robots being built now that work with fractions of milliwatts, microwatts. Computer vision doesn't really have anything to offer them. We found a computational shortcut for depth measurement that will allow platforms in this class. We hope to be able to understand their scenes.

ECCV Daily: *What is the novelty in this work?*

Emma: I would say the novelty is the extreme efficiency of the computation. There are a lot of ways to measure depth. We've got stereo, Depth from Defocus, and there are a lot of ways to handle motion.



Our cue is monocular. It's passive so we don't change the camera. We don't shine light onto the scene. Then our measurement algorithm is just a small linear system on image derivatives. We have very few adds and multiplies. We don't have to expend electrical power on any other part of the sensor.

ECCV Daily: *What is particularly challenging in doing that?*

Emma: We have two main parts of the project. We have a theoretical side where we prove a theorem that our constraint holds and that it's the only constraint of its kind. To do that, we had to really delve into some deep mathematics. We had to use the theory of distributions rather than function. We had to learn a lot to handle all of the possible cases in a single framework.

Then on the experiment side, to build a camera system that works takes a lot of time in the lab. We messed around with optics and found that there was a gap between the blur that we were expecting and the blur that we measured. We had to find solutions to those day to day problems..

ECCV Daily: *How did you manage to solve that?*

Emma: Actually, we realized that we didn't need to solve a lot of these problems. Instead of refining the optics, we get approximate solutions to our equation. We show experimentally that it's good enough. We do this LeadSquared system. We developed a calibration procedure that optimized our depth measurement, and found that approximate solutions to our constraint are sufficiently robust to provide useful scene information, so that we could use a simple camera made of off-the-shelf, imperfect parts, rather than requiring highly precise, specialized optics.

ECCV Daily: *Who works with you?*

Emma: I work with another graduate student named Qi Guo. He's great. He handled most of the experimental side. Then we worked with three professors; Sanjeev Koppal at the University of Florida, who has a lot of experience in micro-sensing, Steven Gortler at Harvard who helped out a lot on the math, and my advisor, Todd Zickler, who is so important to the whole project.

ECCV Daily: *What is the main teaching you learned from Todd Zickler?*

Emma: I think in terms of scientific work, he really cuts through abstractions and makes sure that we understand everything at a very basic level. For him, anytime you make a statement, he may ask you to go to the board, show the equation, and show how you got there. There's no fear in asking simple questions, and there's no shame in explaining simple things. That has really deepened my understanding and helped me communicate my understanding to other people.

Our Pick for Today

ECCV Daily: *What does he have that you recommend to other advisors?*

Emma: I think he has a willingness to engage with any sort of project. You look at his work, and he has advised so many unique projects that are even unique from each other. It's all tied together in a physics based perspective, but the other guys in my lab do totally different projects from me. That's been true for as long as he has had students. He's very eager to help his students pursue their own ideas, and he's willing to go out on a limb to try very unusual, niche projects. It's very motivating.

ECCV Daily: *Are you going to teach one day?*

Emma: I hope to! I hope that if I ever advise a student or an employee that I can be half the mentor that my advisor is.

ECCV Daily: *What are the next steps in your work?*

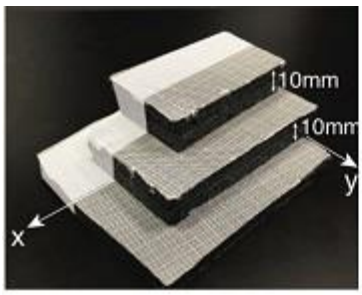
Emma: There's a lot - We found this whole new cue that really hasn't been explored at all. There's decades of engineering work that has been done on similar cues like optical flow. We haven't explored that space at all. That includes everything from adding a regularization term to training a deep net. We really don't know what's out there on a performance side.

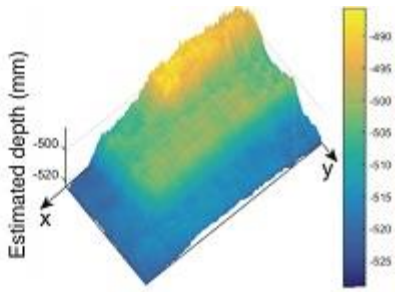
In terms of the math, there's a more general principle behind this camera. We use a combination of motion and defocus to get depth measurements, but in fact a constraint like this will hold for any system with a differential change in defocus. We were inspired by a biological system that seems to be using this principle. We are following up with biologists who confirm or deny that that is the algorithm being used.

We're also categorizing all of the different novel sensors that can use this camera. We're talking fabrication on our next model. We're hoping to get the full picture on a family of cameras, not this cue specifically, but very closely related cues.

ECCV Daily: *Any funny story about this project?*

Emma: Yes! At one point in this project, I was taking care of 25 spiders in a closet in our computer science building. The reason for it is





$$\begin{bmatrix} I_x & I_y & (xI_x + yI_y) & (I_{xx} + I_{yy}) & I_t \end{bmatrix} \vec{u} = 0$$

"Taking care of 25 spiders in a closet in our computer science building. For one year"

"There's a long standing biological mystery as to how they do this"

that, when I mentioned that there was a biological system involved, is that the jumping spider gets its name because it jumps. It will leap several body lengths to catch prey or to get from place to place. It has very accurate depth vision. It also eats about a house fly a week, which works out to a power budget around 700 microwatts. It's in the place that we wanted to build sensors. There's a long standing biological mystery as to how they do this, but a recent behavioral study showed that they are actually using defocus as their depth cue.

From the computer vision literature, defocus tends to be very expensive to process. You wouldn't think that if you wanted to build a very power constraint system that you would use that cue. You would want to use stereo or something simpler.

We saw this and thought that there must be a short cut. We found a short cut. Then we wanted to see if this is what they are doing or is there something else, another short cut. We were going to design an optical illusion and confirm or deny our algorithm. It turns out that the behavioral biology is really hard, and we didn't get any results from that. Still, there was a period of time where I was feeding flies to spiders every week. I learned a lot about that.

They lived out their entire, natural life spans happily which was under a year. We got them as adults so it was only a few months. It turns out you can have spiders sent to you in the mail from, let's say, interesting individuals on the internet. I learned a lot that I didn't expect to in the course of my computer science PhD so far.

"Like lab pets"

ECCV Daily: So as a computer science engineer, you dedicated one year of your life to feeding spiders?

Emma: Eventually, they were more like "lab pets". It was something where I went in and did it for 15 minutes. They were like mascots. It was like having a dog, but easier.

ECCV Daily: Did you give them names?

Emma: No, I didn't. I figured that would just make me sad… [*she laughs*]

Eigen Appearance Maps of Dynamic Shapes

We spoke with **Jean-Sebastien Franco**, Associate Professor at Ensimag and **INRIA Grenoble Rhône-Alpes**, who works with a team interested in retrieving appearance representations for full-sequences of 4D recordings: people that are recorded by multiple camera setups. The project seeks to capture their performance, their geometry, and their appearance (or texture).

Franco works with **Adnane Boukhayma**, a PhD Student, who is the main author, another PhD student, **Vagia Tsiminaki**, and **Edmond Boyer**, who leads the group.

The novelty of their work is mainly the texture in that they are trying to capture the variability of the texture of the person being captured over the whole sequence. In existing works, textures have been constructed per frame or for small time frame window. Here, they are trying to get a presentation that factors all of the variability for multiple sequences for one sequence, but also for multiple sequences of the same person.

It is difficult when you involve a lot of cameras. One of their earlier works computed a high resolution texture in all of these cameras in a single high resolution texture. Getting as much detail as possible for one time instant already proved challenging.




Now that the team went through this step, one of the things that make it difficult when going over all of the sequence is the assumptions when you have one time frame. With one time stamp, the lighting condition and the geometry doesn't change. However, when you go over the full sequence, the frames have micro variations in the geometry along with lighting change because of changes in the body position. The texture is not actually being observed over the same angle with the light sources.

All of this variability can't be captured in just one texture. It requires some kind of representation to capture the variability over all of the sequences to capture these kind of variations.

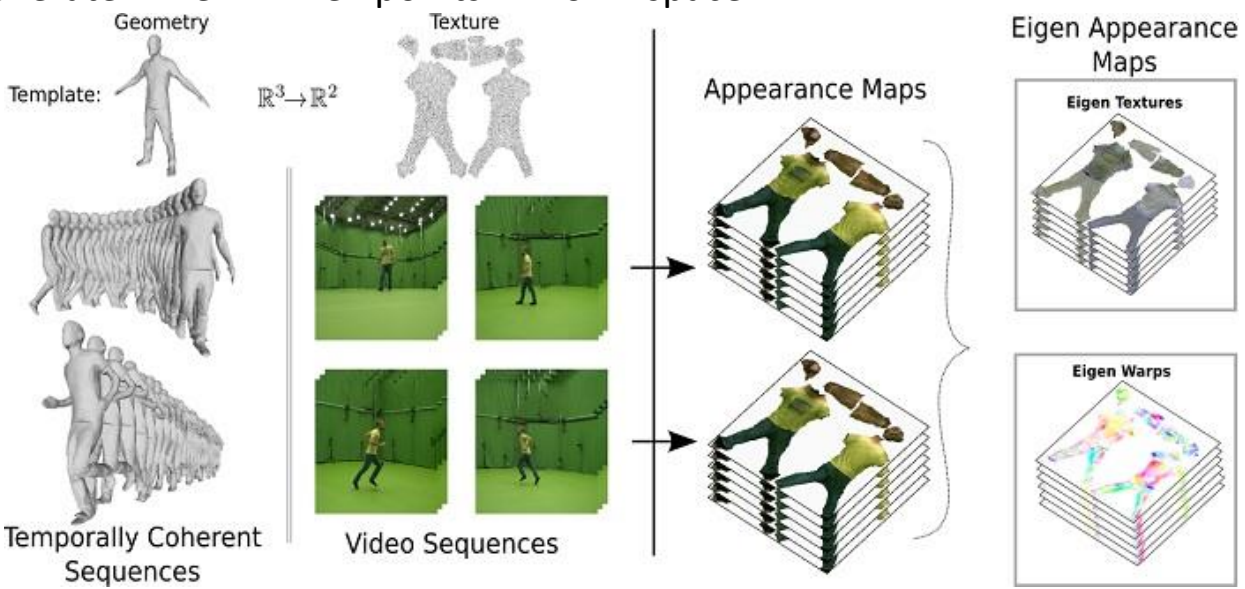
To solve this, they used **Eigen Textures**, a previously existing technique that re-factors a set of texture. In other words, it captures the variation space of a set of textures with a PCA. They employed this kind of technique with some enhancements. They also compute warps to realign textures with small geometric variations within the sequence.

They have factored in some portions of the variation, but they are looking ahead and think more things can be done. For example, while they have already re-factored micro-geometries and recapture variations once the geometries align, they could more explicitly model light variations. They could try to capture one albedo irrespective of where the light source is. They could also separate the lighting component.

In terms of the practical application, their model can generate virtual characters. With several sequences of the same person, it can generate full virtual characters by re-rendering the representation to generate new viewpoints. For example, in recorded sequences in which the character jumps, runs, and moves, it can transition between these different sequences and animations and generate new animations. It could provide a full character model and do renderings of this character model by interpolating between the acquired sequences.

Their technique provides an efficient way to do this because, with Eigen texture representation, it can easily interpolate between sequences by interpolating the coefficients of the Eigen representation between the sequences.

The other application of this work is for texture completion. Let's say you have a sequence of an actor whose performance you are capturing, and some part of his body is being included for that particular view. Since there is a texture space or PCA space that captures the variability, it can reproject that particular frame in the PCA space and generate the missing part of the texture by extrapolating it from the texture space.



Zeynep Akata

ECCV Daily: *Zeynep, what is your position?*

Zeynep Akata: I am doing a postdoc at the **Max Planck Institute for Informatics**.

ECCV Daily: *Where are you originally from?*

Zeynep: I am Turkish from Istanbul.

ECCV Daily: *What is the project you are working on now?*

Zeynep: In general, my research interests have been image classification. When you don't have enough training data for some of the classes, we are still doing classification using some form of sight information, for example, different object properties help us or sentences that we write about objects help us. After having sentences that are very descriptive and detailed, we have also decided to use them to generate images from scratch. I worked on image generation.

There are other ways of conditioning the generated models. Apart from the text, we can also use locations or object key points, for example. This helps us to build larger images that are visibly high quality. With Lisa Anne Hendricks we were working on the opposite problem, generating sentences using images that Lisa is going to talk about.

ECCV Daily: *How did you decide to become a scientist? What happened, as a young girl, that motivated you to pursue science?*

Zeynep: I was interested in math, but not physics. My father is a mechanical engineer, and he wanted me to become a mechanical engineer. Because I was interested in physics, I said I would do computer science. I made the decision right before exams for university around the age of 18.

ECCV Daily: *Before you decided to study computer science, did you feel gifted in technology?*

Zeynep: No, I didn't. I grew up as a normal kid without any special interests. I was doing well in all of my courses. For example, I liked reading and drawing a lot. I chose computer science partially because I could find a job easily afterwards.

ECCV Daily: *What were your dreams when you were a kid?*

Zeynep: I didn't really know what I would become or have any special dreams. It didn't worry me. I thought to myself, whichever subject I choose, in the end I will have to work very hard to become good at it. This can be something that I really, really love so maybe it will be easy to work hard. I also thought that if I select something, one of my hobbies for example, if I have to work really, really hard on that, then maybe it's not going to be my hobby anymore. Therefore, the choice for me was mostly logical. I will be able to find a job, and computer science seems to be an interesting area which has a lot of different subjects. If I don't like computer vision, I can move to computer graphics. If I don't like that, I can move to theoretical computer science. I thought this is a subject that I can change when I want to, if I want to.

ECCV Daily: *Have you always been a hard worker?*

Zeynep: Yes

ECCV Daily: *Are you also a perfectionist?*

Zeynep: Maybe, a little.

ECCV Daily: *When did you decide to leave Turkey?*

Zeynep: After I finished my Bachelors. I decided to go to Europe to do a Masters and to have experience.

ECCV Daily: *What was your choice?*

Zeynep: I applied to a lot of places in Germany because you don't have to pay anything for the education. It's free. I wanted to learn another language. This was another motivation. That's why I chose Germany.

ECCV Daily: *Were you encouraged by teachers, family, and friends?*

Zeynep: Yes, my mom. She said you will try it and there is no question. You will just go.

ECCV Daily: *Did anyone oppose you?*

Zeynep: No, no one opposed me.

ECCV Daily: *Do you think that you grew up in an environment which was supportive of what you would become?*

Zeynep: Yes, of course, my parents gave me direction. They told me their opinions, but they didn't really force me in terms of subject choices.

ECCV Daily: *Is it the same for all young women in Turkey?*

Zeynep: No, I don't think so. I know that I'm quite privileged for that.

ECCV Daily: *When did you decide to take the next step? What happened?*

Zeynep: After, I was writing my Master's thesis. I found out about a computer vision summer school in France. I decided to submit my thesis paper.

ECCV Daily: *To learn another language?*

Zeynep: Partially, maybe - Also good food, sunshine, and good wine.... [she laughs]

ECCV Daily: *Whereabout in France?*

Zeynep: Grenoble - Basically, I went there and presented my work. Then, I got the best poster prize. That gave me courage to talk to the head of the group, Cordelia Schmid. I talked

to her and told her I wanted to do my PhD. I didn't apply to any PhD positions, aside from talking to her. She was interested in hiring me in Grenoble so I went there.

ECCV Daily: *What is it like to work with her? Can you tell our readers what it is like to work with Cordelia Schmid?*

Zeynep: It's very fulfilling. She is very accessible. Whenever I had any problems, even sometimes technical questions, she would have time for me. We would have almost regular meetings, maybe once a week or once every two weeks. She was very interested in looking at the results. She has a very good intuition. When she sees results, she can say ok, there is something wrong, or there is something interesting that we have to explore. Sometimes she can even detect at first sight, that I have a bug in my code, just by looking at the results, because the way the results look, she can see that it is not possible that this can happen, so there must be something wrong in the code.

ECCV Daily: *Did you find anything challenging as a woman in your profession?*

Zeynep: Yes: career advice that I got wasn't always encouraging. People were wondering why I don't go back to Turkey and get married, for example. These kind of comments I heard from people.

ECCV Daily: *Were these comments from people who barely know you or from a closer circle?*

Zeynep: From a closer circle, in Europe

ECCV Daily: *How do you react to these kind of things? Do you just shake it off or do you address the opinion of that person?*

Zeynep: When I was a PhD student, I was much less open about it. I couldn't handle the conflicts very well so I would just ignore, and basically do whatever I wanted to do. I would not reply to those people.

Now, that I have more experience, and I have seen other examples, now I think that it's important to speak your mind, to tell people that this is not the right way of giving advice. This is very discouraging because there is already a lot of male domination in our field. If anyone gives that kind of career advice, this becomes even more discouraging for women. I think that this is maybe one of the reasons why women don't want to continue in postdoc. They suffer so much.

ECCV Daily: *You work now with **Bernt Schiele** at the **Max Planck Institute**. Tell us more about it.*

Zeynep: It has been fantastic. He is very, very supportive of women. I think maybe he treats women differently, in a good way. Because, sometimes we get emotional. He can handle it very well. You know that he is supporting you, even in difficult situations. He doesn't have to, but he supports you.

ECCV Daily: *In which situations do scientist women become emotional?*

Zeynep: Maybe when the student you are working with, usually a male student, becomes very frustrated and becomes aggressive toward you. Then, you have to keep calm, be cool, and not get extremely angry or apologetic.

ECCV Daily: *Do you think young people might not respect your role only because that role is occupied by a woman and not by a man?*

Zeynep: Yes, sometimes. I'm not saying that for everyone this is the case, but there might be some situations. I have had some experiences like that.

ECCV Daily: *If you would give advice to a young man entering this profession, what would you say?*

Zeynep: Just ignore the fact that your supervisor is a woman. Treat her as a man.

ECCV Daily: *Do you think men can ignore that the other person is a woman?*

Zeynep: I don't know, but they have to. They have to force themselves.

ECCV Daily: *What advice would you give to a young woman entering this profession?*

Zeynep: If they want to do it, they should do it because the field really needs more women. I encourage everyone who wants to explore, just to explore.

ECCV Daily: *You said that you need to be strong to address all of these kind of issues. What if that person wants to explore, but she is not strong enough to deal with the pressure?*

Zeynep: Then they can find a mentor, someone to talk to, someone to take advice from. That mentor can come from the computer vision community, among the people that you have interviewed perhaps, or from their

own labs. That person doesn't even need to be a female. It could be a very understanding male boss. That can also work. You just need somebody to discuss these issues.

ECCV Daily: *Do you know any woman scientist who gave up that direction because of this?*

Zeynep: No, I can't think of anyone at the moment.

ECCV Daily: *Do you have any special dreams now?*

Zeynep: Becoming a professor. I want to teach because I like to be around young people and students. It's a sense of community service for me. I feel satisfied and fulfilled. Academia is not only about teaching. I can also continue doing my research.

ECCV Daily: *What would you like to achieve?*

Zeynep: To have taken part in a project like building intelligent machines. I want them to be more intelligent in a general way. Intelligent, not only on this application, but also generalized to many different application.

ECCV Daily: *Do you think you can achieve that?*

Zeynep: Not me alone, but with a good team, I can.

ECCV Daily: *Will machines become more clever than us?*

Zeynep: I don't think so, but you never know.

"I wouldn't like to build a silent robot"

ECCV Daily: *If you could give a feature from a human to a machine, which one would you give it?*

Zeynep: This is a very difficult question. To talk to me, to have a discussion. An autonomous discussion. I wouldn't like to build a silent robot.

ECCV Daily: *That machine could have an autonomous discussion because it has emotions or because it has an intellect?*

Zeynep: Not emotions, intellect. Instead of searching on my own on Wikipedia for what I want, I can just ask a robot, and that robot is going to explain to me what I want.

ECCV Daily: *Which is not so different from what Google does now?*

Zeynep: Yes, but it's still not that unconstrained. I want it to be as unconstrained as possible.

ECCV Daily: *When do you see that happening?*

Zeynep: Maybe in 10 years.

ECCV Daily: *Enough time to work and make that happen. Good luck!*

Microsoft HoloLens

Marc Pollefeys
Partner Director of HoloLens Science, Microsoft

There was a long queue yesterday in front of the Microsoft demo booth. It was just natural to ask **Jamie Shotton of Microsoft** what was going on.

ECCV Daily: *What are you doing here at ECCV, and why are there so many people in front of your booth?*

Jamie: We are demoing the world's first, fully-untethered holographic computer, the **Microsoft HoloLens**. We are allowing participants at the conference to come in, put the HoloLens on their heads, look through, and see the future of mixed reality.

ECCV Daily: *I understand that people are very impressed by what they see. What is the feeling that they have when putting on the HoloLens? Can you explain it?*

Jamie: You put it on, and this whole new world pops into your vision. You can see objects that aren't really there in the real world, but they float out there in front of you. You can move around them, look at them, and interact with them. This is augmented reality. It is already on the market as a developer edition, and it has a price for developers. We're super excited about this!

Salehe Erfanian Ebadi

Foreground Segmentation via Dynamic Tree-Structured Sparse RPCA

Salehe Erfanian Ebadi is a PhD student at Queen Mary University of London, working under the supervision of Professor Ebroul Izquierdo. Together they developed a video recording project with BBC called THIRA and a security project for Europe and the UK called LASIE.

This project seeks to get foreground segmentation and background modeling in general video sequences. In a video, foreground can include a lot of things. For example, if a video contains a moving object such as a person walking across the scene, the person is foreground. Then, if the person puts down a piece of luggage and leaves the scene without the luggage, that's also foreground. The same for cars moving or entering/parking/leaving the scene.

The main goal is to find all of the regions that are interesting for further processing, which can include tracking, segmentation, or recognition. It can recognize the person or detect if the object left on screen contains a bomb. That's the practical application of the project.

However, sensors are very local. In addition, CCTV cameras provide poor images and color definition with high noise. The video may also have low lighting situations making it difficult for computers to detect moving objects. This can prove difficult even for humans, making it even more challenging for a computer to detect all of the parts correctly.

For this work, they use heavily modified and approximated robots principal component analysis; and dynamic tree structured sparsity to get very good, crisp foreground segmentation accuracy. The method was tested on 4 different data sets, which are benchmark background subtraction data sets providing state of the art results for all four.

The next step is making it real-time to make this into a product. The most challenging thing for Salehe was to actually get this published, even without using deep learning (the hype of these days, in her words). Instead, she makes an effort and prove that this model works better than deep learning. Still, she successfully published her work, e.g. with ECCV (find it in the proceedings) and also in in a couple of ICIP papers as well.