A publication by

Computer Vision News

RSIP VISION

The magazine of the algorithm community

December 2016

Women in Computer Vision
Zoya Bylinskii - MIT

Spotlight News

Challenge
CAMELYON16 (ISBI 2016)

Research Paper
Shapes from Pixels

Trick of the Month
Web Tools for Visualization

Guest
Prof. Nikos Paragios

Event
Fall Meeting of the NVPHBV

We Tried for You NEW!
Generalized Hough Transform

Application
ICT4Life - Life Improvement for Elderly

Guest
Prof. Nikos Paragios

04

Focus on... LIPNET
Machine lip-reading

09

Women in Computer Vision

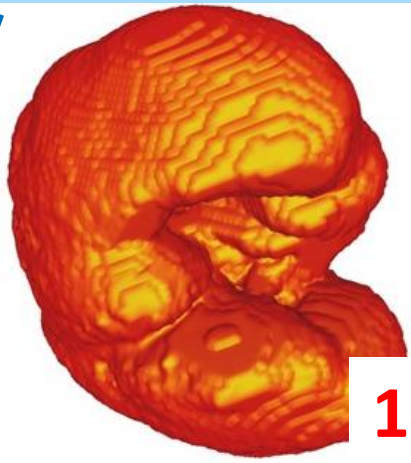
16

We Tried for You
Generalized Hough Transform

NEW!

10

Trick
Web Tools for Visualization



14

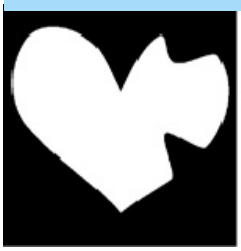
Events

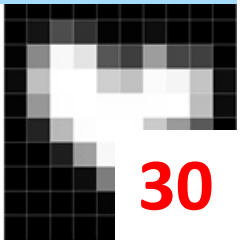
34

Project
Montage of Retinal Images

26

Research
Shapes from Pixels





30

Spotlight News

38

Dear readers,

As usual, this **December issue of Computer Vision News** is dedicated to the people, applications and techniques which help improve people's lives. A good example is **ICT4Life**, a service to support independent life for people suffering of **Alzheimer's**, **Parkinson's** and other dementias, which we have chosen as the **Application of the Month**; a **retinal montage software** is our **Project of the Month**, as it supports ophthalmologists in their quest for complete and meaningful images of their patients.

You will also read stimulating interviews with thought-provoking scientists who accepted to be our guests: on one hand, **Professor Nikos Paragios** of **Ecole Centrale Paris** and **INRIA** will tell us his personal views on how **deep learning** is transforming our community, with visible benefits but also some hidden perils. On the other hand, **Zoya Bylinskii** of **MIT** will tell us her personal views about a host of interesting topics: find her in our **Women in Computer Vision** section. Nikos and Zoya are not our only guests: **Veronika Cheplygina** was so kind to tell us all about the recent Fall Meeting of the **Dutch Society of Pattern Recognition and Image Processing** in Eindhoven.

Once again, our technical content is rich enough to satisfy even the most demanding of our readers: a new section called **We Tried for You**; a great research paper called **Shapes from Pixels**, about a model to recover shape images with smooth boundaries from a set of samples; the **Trick of the Month,** with Web tools for visualization in medical image analysis; and our section about **Image Processing Project Management**: this time, our CEO **Ron Soferman** lectures on his recommendations about **Proof of Concept (POC)**, a crucial step in algorithmic projects.

You will love reading this December magazine. Please keep sharing it with friends and colleagues. And **subscribe**, if you didn't yet.

Enjoy the reading!

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

Computer Vision News
Editor: **Ralph Anzarouth**
Publisher: **RSIP Vision**

Click to contact us

Click to give us feedback

Click for free subscription

Click to read previous magazines

Professor Nikos Paragios

Nikos Paragios is full professor of Applied Mathematics at the **Ecole Centrale Paris** and the scientific leader of an **INRIA** research group called **GALEN** at **Paris-Saclay**. He is also Editor in Chief of the **Computer Vision and Image Understanding Journal**. He wrote some time ago a very intriguing text about deep learning called ***Computer Vision Research: The deep "depression"***. He was so kind to discuss with **Computer Vision News** about scientific research, education, robots, industry and more. He also gives very precious advise to all students…

Computer Vision News: *Nikos, I know that some time ago you wrote a very original text about deep learning. Can you share with our readers your intriguing views on deep learning?*

Nikos Paragios: Deep learning is the hottest topic in computer vision and beyond in our days. We have more and more people jumping in, and the reason why this is happening is because the performance is impressive. We're getting results that were unimaginable with state of the art methods five or ten years back. The reason why this is happening is because we have more data and computing power. We also have methods that were developed in this area, like **neural nets**, that have been around for 40 years and have progressed in terms of how you can optimize and learn these networks.

"Our job is to progress science and the progress of science is understanding what we are doing"

I wouldn't say I'm not fond of it, but the reason why something is bothering me is because I have the impression that the community is actually moving forward building architectures: it's becoming a way of composing a puzzle without having a clear picture of what is happening behind. I think something like 50% of the community is working in this area. Some people are doing great theory stuff. But the rest of the community is actually trying to get the results out of these methods by decomposing, putting layers and adjusting layers. I think that shouldn't be the main objective of academic research. This is something that should happen in industry. **Our job is to progress science** and the progress of science is understanding what we are doing. So I think it's a great tool, its performance is really impressive and it has all the potential to become a very advanced tool in the field. But I prefer that we spend more time on theoretical questions rather than spend time trying to beat results on a benchmark.

CVN: *It does help us create better applications in real life...*

Paragios: Yes, it's already happening. And this is one of the reasons why deep learning is very popular: because even without having a very strong expertise in the area, if you have a dataset you can build very good prediction and recommendation systems. That's the reason for its success, it is a great tool and it contributes to make computer vision so popular in the industry. If you go to conferences now, the industry presence with its demos is as strong as posters. I remember 5 years ago there were only a few companies. Now there are more and more small, medium, and large sized companies in different areas. It's actually helping us a lot to transfer technology because it works.

CVN: *On the other hand, there may be students who do not learn enough about the basics of science, math, physics, geometry, etc. and something is getting lost.*

Paragios: Unfortunately, that is what is happening. Now, the way the papers are evaluated is based on how well they do on benchmarks. The objective becomes to produce the best possible results. If you don't have a principle based on a mathematically rigorous way to understand what you are doing, then you're just going to keep spending time on trying.

It's something that has its own value, but in the long run it's important that every PhD student or researcher knows where the field comes from: deep learning now is hot, but if I look 10 years back, everyone was doing **compressed sensing** and everyone was saying that compressed sensing was the future. What is happening now, is that compressed sensing is still there, but the impact is not there yet.

It's important that people know what has been done before. Perhaps a wide theoretical view of the problem is better than a constrained view of the visual field by just applying these methods to different problems.

CVN: *Scientists are trying to teach computers to mimic how our brains work. I know of a 5 years old kid, who could immediately recognize what another dinosaurs looked like from one toy dinosaur he had. A computer needs to see a massive number of dinosaurs until he starts to recognize them with some kind of certainty. Does it mean that the computer is nowhere near what a little kid can comprehend?*

Paragios: There are two ways of answering your question. Actually, we don't have a real clue on how the brain works. We have ideas on how decisions are made. We have some ideas on the connectivity. We have some ideas on the computing power, but we don't really know on a very fine scale what is exactly happening, because we don't have the tools to visualize how decisions are made.

"There should be something more fundamental that we are not getting yet. That's why we are really far"

Another way to look at this question is that it really depends on the task. **Human intelligence is not only vision, it's a combination of things**. We understand the environment. We use our experience. We use plenty of other sources of information. We have a lot of work that we do in order to actually approach this. I don't think it's only a problem of data. There should be something more fundamental that we are not getting yet. That's why we are really far.

CVN: *Is it a kind of failure of science that we do not know yet how the brain functions? On what grounds are we sending kids to school where they will be taught things, if we still don't know how their brain functions?*

Paragios: Everything is a question of progress. We are getting closer and closer. We now have more clues and we are trying to adapt educational and computer vision systems according to the brain. As I said, it's **a very complex engine and a highly efficient device**. I can give you a very simple example. My brother and I went to the same school and we had the same instructors, but we don't have the same way of thinking. We don't have the same capacity.

CVN: *...which is actually good so that we can be a bit different from one another. Otherwise we would be 7 billion robots in the world.*

Paragios: It might happen.

CVN: *When will there be more robots than people?*

Paragios: Robots will become something that we use in our daily lives in 10 to 15 years from now. What scares me is what will be the position of humans in the world. It's great to build a virtual reality mask which gives you the impression that you are on a sunny beach if you cannot afford it, but it's better if you can be on a real sunny beach. Technology has a huge impact on society. That might cause huge problems in the future. We need to think about how we integrate our technology progress with the way of life of humans.

CVN: *What are the dangers if we don't pay attention now?*

Paragios: There is already a huge gap between the different parts of society. Either you are privileged or not. I think technology is actually going to increase this gap. There are people that are going to be able to afford these kinds of tools. Then there are people whose jobs will go away. You have to think about how society is impacted.

"Robots and technology should improve the lives of all humans, rather than focus only on the few people who can afford that"

CVN: *There is already a lot of wealth, but it is completely unevenly distributed.*

Paragios: Exactly - I think that is what's going to happen. It's already the case, unfortunately. The way that society is progressing, it will be even stronger in the future. Robots and technology should be there to improve the lives of all humans, rather than focus only on the few people who can afford that. They should focus on the whole society and be beneficial to the majority of people living on earth.

CVN: *Do you see your students' mentality evolving and giving less importance to the human side and more importance to the machine side?*

Paragios: Our students in the past wanted to build a career with international companies and explore the possibilities. Now their dream from the very beginning is to make big money as part of a startup, by simply applying some very stupid machine learning tool to make big money. The positions of humans in the future is something that we're not paying a lot of attention to. We are excited about technology because we are scientists, which is good. We should go back and think about ethics and try to **have a better understanding about how this technology affects our lives**.

CVN: *What is the most important thing you learned from your own teachers?*

Paragios: They taught me **hard work and dedication to objectives**. Set your target, keep investing and never give up. Even if you fail, at least you did your best and have no regrets.

CVN: *Failure is certainly part of the life of a scientist. Sometimes you succeed and sometimes you don't. Did you see any of your colleagues or students very deeply affected by failure?*

Paragios: Yes. I had several cases of PhD students whose first paper was rejected and **their whole world collapsed**. You should be able to handle these kinds of things. No one is perfect and failures are part of life.

CVN: *We are not all equally strong. How can we help younger students who will read your interview gain enough confidence to overcome failures?*

Paragios: The important thing is that you do your best. Each of us has different strengths. Some people are good at some things, other people are good at something else. Once you assure yourself that it is possible to achieve that goal and invest the time and the effort to reach the objective, then it's not a failure, since you did your best. Maybe their objective was too high or the timing not right. If you did your best, then you can move on, because you fulfilled the requirement. We can't meet our objectives all the time, but that doesn't stop us from submitting papers to the next conference, even if papers are not all of exceptional quality.

CVN: *The largest companies are already exploring the hottest topics of scientific research. How can small companies or startups find a place in research and make a breakthrough in technology?*

"I had several cases of PhD students whose first paper was rejected and their whole world collapsed"

"If you did your best, then you can move on, because you fulfilled the requirement"

Paragios: There are two different things. **One is technology. The other is the market**. When you have great technology, it does not necessarily mean that the market is there. That's one thing.

The other thing is actually something happening in vision now, which I don't like: the role of academics is to advance science and not trying to beat the performance results on a benchmark. If we do that in the academics, we become like industry where the objective is to reach the target. **Universities should be places where people have the time needed to take a step back from the problem and think about the most appropriate way.** We shouldn't have the pressure to make it better and better. We shouldn't have the pressure of publishing tens of papers each year.

CVN: *Who is applying the pressure?*

Paragios: This is the system. When I was a student, a PhD with two of three major conferences was a fantastic PhD, because then you didn't have a lot of pressure to submit to every conference. It's impossible even for the brightest students to produce four or five great ideas every year. Now, you have something like 4,000 papers submitted at every major conference, only a few months after the previous one.

Academia is a place where startups and industry should look for innovation. In order to do so, **academics should have a really close view of what is happening in industry**. There is nothing wrong with applied research. There is a big difference between applied research and the application of research to solve a problem.

"Keep working hard and set the bar as high as you can. Never give up. Hard work and persistence always pay off!"

CVN: *It sounds like you are saying that academia and industry are not going perfectly hand in hand. What would be needed so that they do?*

Paragios: It depends on where you are. For example, in the US, industry is investing as much as academia in research. Something that might work really well is taking some people from industry and some from academia with funding grants that actually put these two together.

CVN: *If I gave you a microphone to talk to all the students in the world, what would you tell them?*

Paragios: I think they should invest in broad knowledge. It's better to have broader knowledge than to be highly specialized in something. More knowledge will give you access to a different way of thinking and help you to **make the right decisions for yourself and for the rest of us as well**. I say to all students: keep working hard and set the bar as high as you can. Never give up. Hard work and persistence always pay off!

LIPNET - How easy is lip-reading ?

If you are not covering your mouth yet, you probably should. At least, if you don't want anybody else to know what you are saying. This is the one of the consequences of an impressive project called **LipNet**, a model that maps variable-length sequences of video frames to text, making use of spatiotemporal convolutions, a Long Short-Term Memory (LSTM) recurrent neural network, and the connectionist temporal classification loss, trained entirely end-to-end.

LipNet attains a remarkable **93.4% sentence-level word accuracy**, outperforming previous state-of-the-art models, which performed the same tasks with accuracy just below 80%.

An even more striking comparison is that with **hearing-impaired people**, generally considered among the most accurate human expert lipreaders: the authors found that on average this population achieved an accuracy of 52.3%, much below Lipnet's.

Besides Lipnet's notable performance, the main novelty of this work is that it maps sequences of image frames of a speaker's mouth to **entire sentences**, eliminating the need to segment videos into words before predicting a sentence. Here is the PDF of the paper.

"Real-world applications would include hearing aids improvement"

Real-world applications for accurate machine lipreaders would include **hearing aids improvement**, silent dictation in public spaces and speech recognition in noisy environments.

The sentences tested in the video above follow a pattern, which makes the demo slightly less impressive. **Yannis Assael**, one of the authors, explains that sentences without context are important to evaluate the actual performance per word. Furthermore, this is due to the **GRID dataset** (one of the few available), offering 64,000 possible combinations of sentences with a fixed structure.

The Generalized Hough Transform

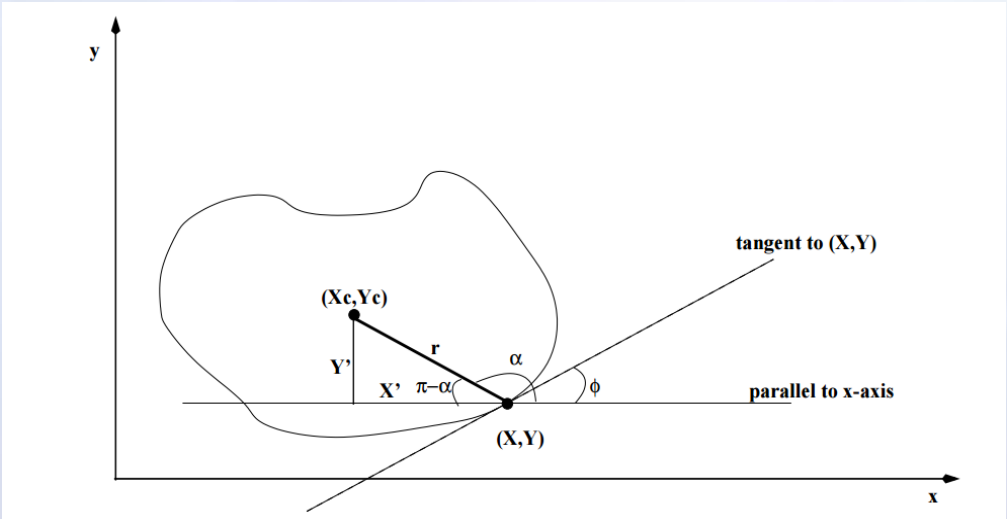
The **Original Hough Transform** was developed to detect analytical shapes (lines, circles, ellipses, etc.). It does so by seeking the parameters defining the analytical representation of the shape obtained as local maxima in a so-called accumulator space (or Hough space). The **GHT (Generalized Hough Transform)** is an extension of the original Hough transform, introduced by **Dana H. Ballard in 1981**, to enable detection of arbitrary shapes (i.e. those shapes that cannot be represented parametrically).

The Matlab code we tried for you for the generalized Hough transform can be downloaded here.

To generalize the Hough algorithm to non-analytic shapes, Ballard constructs the following: given any shape image and a pre-define reference point inside it, instead of a parametric curve in the transform stage, the information provided by the boundary pixels is stored in the form of the **R-table** (to be described in the next section). For object localization, every edge point on the test image is looked upon the R-table, the properties of the point are retrieved and the appropriate cell in a matrix called the **Accumulator matrix** is incremented. Next, we will examine code snippets of those two key elements of the GHT: **Building the R-table** and **Object localization**. Then, we will demonstrate how the GHT localizes an object in a given image.

Building the R-Table

Each entry of the R-table is comprised of two values: the gradient direction and all the boundary points corresponding to that gradient. The table is constructed as follows: given a reference point (*Xc*,*Yc*) inside the shape's boundary (previously detected using an edge detector), the gradient direction – ϕ and the distance from the reference point – *r* (see image below) are computed for each boundary point (X,Y). The *r* values of each boundary point, indexed as a function of ϕ, constitute the R-table. Note that each index of ϕ may have many values of *r*.



```
% devide the angel space to MaxAngelsBins uniformed space bins
MaxAngelsBins=30;
[y x]=find(Itm>0);
% counter for the amount of edge points associate with each
% angel gradient
PointCounter=zeros(MaxAngelsBins,1);

GradientMap = gradient_direction( Itm );
for f=1:1:nvs(1)
    % transform from continues gradient angles to
    % discrete angle bins and one to the number of
    % points in the bin
    bin=round((GradientMap(y(f), x(f))/pi)*(MaxAngelsBins-1))+1;
    PointCounter(bin)=PointCounter(bin)+1;

    % add the vector from the point to the object center to the bin
    Rtable(bin, PointCounter(bin),1)= Cy-y(f);
    Rtable(bin, PointCounter(bin),2)= Cx-x(f);
end;
```

The construction phase may be inverted to detect repeat occurrences of identical objects in the image.

Object localization

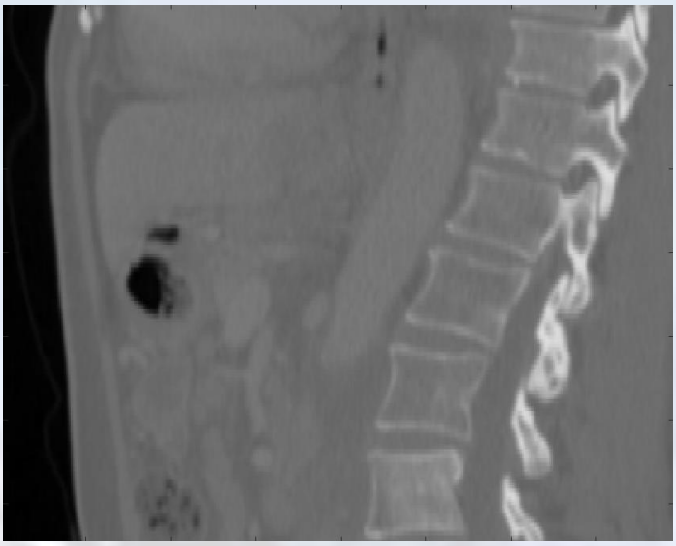
To detect an object in the image, the GTH loops through all the edge-pixels. And for each it computes the gradient ϕ. For each edge pixel (*X,Y*) in the image, increment all the corresponding points in the accumulator array *A* (initialized to a maximum size of the image) where *r* is a table entry indexed by ϕ, i.e. *r(ϕ)*. Maxima in *A* correspond to possible instances of the shape *S*.

Below is the code snippet for this process:

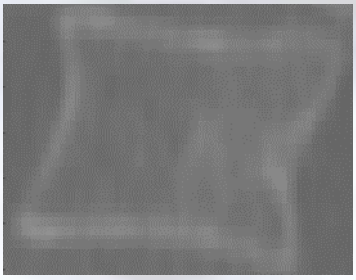
```
% find all edge points in the main image
[y x]=find(Image>0);
% find number of edge points in the edge image
np=size(x);

Ss=size(Image);
houghspace=zeros(size(Image));
for f=1:1:np(1)
    % transform from continue gradient angles to discrete angle bins
    bin=round((GradientMap(y(f), x(f))/pi)*(MaxAngelsBins-1))+1;

    for fb=1:1:PointCounter(bin)
      % add point where the center of the image should be
      % according to the pixel gradient
      houghspace(Rtable(bin, fb,1)+ y(f), Rtable(bin, fb,2)+ x(f))= ...
          houghspace(Rtable(bin, fb,1)+ y(f), Rtable(bin, fb,2)+ x(f))+1;
    end;
end;
```

Example

Let us consider the following test image: a sagittal slice of a CT-scan



and the following template (Spinal vertebrae) to be detected inside the image



The following three-line code snippet invokes the GHT function. As you can see it is easy to use: given the image and a template, the GHT returns the x,y-coordinates of the template inside the image.

```
EDGE_IMAGE    = edge( ORIG_IMAGE    ,'Canny' );
EDGE_TEMPLATE = edge( ORIG_TEMPLATE ,'Canny' );

[score,  y,x ]  ...
    = Generalized_hough_transform(EDGE_IMAGE, EDGE_TEMPLATE );
```

Following are the results of the edge detection in the test image and the template (i.e. the output of the first two lines of code):

Next, you can see the localization of the template within the test image (according the x,y-coordinates return by the GHT function).

FEEDBACK

Dear reader,

How do you like Computer Vision News? Did you enjoy reading it? Give us feedback here:

Give us feedback, please (click here)

It will take you only 2 minutes to fill and it will help us give the computer vision community the great magazine it deserves!

Web Tools for Visualization

Every month, Computer Vision News shows you a nice trick that will make your work easier. This time, our engineers show how **you can share interactive, animated 3D figures, for instance of medical scans, by making them viewable in browser**.

To do this, we will use the x3mesh_deform function written by **Benjamin Irving**, which you can download here.

Let's demonstrate this x3mesh_deform by using our Matlab kidney visualization (follow the link to find it as the **Trick of the Month** in **Computer Vision News of July**); but this time, we will visualize it and animate it via HTML in the browser.

The following Matlab code creates a mesh (comprised of faces and vertices) to construct the object that needs to be visualized; it then calls the x3mesh_deform function to generate the HTML file:

```
1  KidneyImageSmooth = smooth3(KidneyImage);
2  isoVal = 1;
3
4  fv =isosurface(KidneyImageSmooth,isoVal);
5
6  fv2=fv;
7  fv2.vertices(:,2)=0.9*fv2.vertices(:,2);
8  fv2.vertices(:,3)=0.9*fv2.vertices(:,3);
9
10 % running the html creating function
11 x3mesh_deform(fv.faces, fv.vertices, fv2.vertices,
12                    'name','HTML_FILE_NAME')
```

Line 1: the Matlab code takes as input the kidney segmentation from a CT scan (kidneyImage) - which is a 3D matrix with positive **Hounsfield Unit (HU)** values inside the kidney and zero elsewhere.

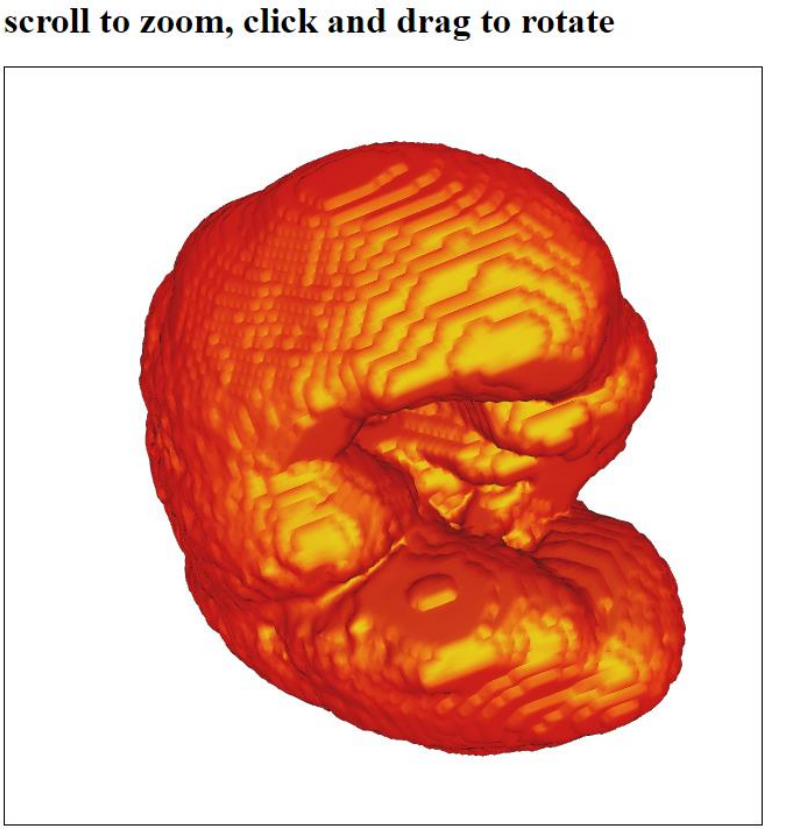
Line 4: we call **isosurface** the Matlab's built-in function for computing the mesh for a given threshold pixel value (isoVal). Setting the isoVal to 1 means that all pixels with positive values will be included in the mesh that envelops the segmentation; this will give us the kidney boundaries, since our kidney segmentation (kidneyImage) has a greater than zero values (HU) inside the kidney and zero outside.

Lines 6-8: so far, we have generated a static mesh; now we will animate the image by generating a second mesh and the animation will morph from the first mesh to the second and back. To do so, we will create a second set of vertices to define the second mesh. This is done by defining a mesh 90% of the size of the first mesh.

Lines 11-12: we call the x3mesh_deform function which takes as its input the faces and vertices of the first mesh and the vertices of the second mesh (i.e. the animation will flow from one set of vertices to the other). HTML_FILE_NAME defines the name of the html output file.

When you open the HTML file in a browser, you will see this nice mesh visualization. In addition, this visualization can be dragged, rotated and zoomed-in.

A few words on the X3D-html technology that generates this visualization: X3D is an open standard file format and run-time architecture to represent and communicate 3D objects using XML. It provides a system for the storage and playback of real time graphics, all within an open architecture to support a wide array of domains and scenarios.



X3D has a rich set of componentized features that can be tailored for use in engineering and scientific visualization, CAD and architecture, medical visualization, training and simulation.

The following snippet gives you the basic outline of the X3D components used to generate the visualization above:

```
<x3d id=''someUniqueId'' showStat=''false'' ... >
    <scene>
        <viewpoint id=''aview'' centerOfRotation=''0 0 0'' ></viewpoint>
        <transform DEF=''airway1'' rotation=''0 1 0 0''>
            <shape>
                <appearance DEF=''App''/>
                <indexedFaceSet  coordIndex=''<YOUR FaceSet>''>
                    <coordinate  point=''<YOUR VerticesSet>''>
                    </coordinate>
                </indexedFaceSet>
            </shape>
        </transform>
    </scene>
</x3d>
```

The *shape* tags tell the browser that we want to put a 3d shape in the scene. The *appearance* tags delineate any alterations that we want to make to the appearance of the mesh. The *indexedFaceSet* tags define the faces to be used in displaying the mesh – listed as a string <YOUR FaceSet>. The *coordinate* tags define the edges to be used in displaying the mesh – listed as a string <YOUR VerticesSet>.

Zoya Bylinskii - MIT

We continue our series of interviews with **women in computer vision**. This new section, which we started with the CVPR Daily at CVPR 2016, hopes to help mitigate the severe gender imbalance in the computer vision community by getting to know better some remarkable female scientists and their career paths. We think that some of what they did might serve as an example for other young females who wish to enter this field. This month, we interview **Zoya Bylinskii**, who is currently pursuing a **PhD in Computer Science at MIT**.

Computer Vision News: *Zoya, when did your interest in science start?*

Zoya Bylinskii: In high school I was interested in psychology and biology. I read books about the brain a lot. I was interested in brain facts and trivia. I read a lot of books about human psychology. At the same time, I was also quite good at math and participated in math competitions.

I wasn't interested in computer science, even though both of my parents are computer scientists. Later, during undergrad, I did decide to pursue a computer science degree in artificial intelligence because I was so interested in studying human psychology and the brain. In this program, I could have access to the right technical tools.

A lot of the courses that I took in my first year were high level courses on cognitive science, linguistics, and intro to psychology. As I went along with my program, I also took basic courses in

computer science in the first few years. Later, it was a mix of computational linguistics, computational vision, and computational cognitive science.

CVN: *What was your experience during your formative years that made you a scientist? What did it teach you about your current approach to education?*

Zoya: I'm going to answer it a little on the side. In terms of my view on educating children about computer science and how that could have influenced me. When I first learned about computer science, I wasn't so interested in it. When I talked to other children my age, the ones that were really interested in it were the ones that played video games. That aspect of it never appealed to me. That was the way that it was mostly taught to children in school. Unless you played video games, there weren't quite the right approaches to teaching children about computer science. I do have strong opinions about that because I think that it ends up cutting off children from understanding that computer science is also a tool and not for the purpose of entertainment or playing games.

CVN: *You were taught that a minute not spent learning is a minute lost. Is that right?*

Zoya: My dad taught me this drive to **pack new knowledge into every minute**. I find myself at home picking up new things and reading them. It doesn't have to be in the computer science field. I have this picture in my head, when my dad took a piece of paper and said: "If you are going to learn something now, it is this dot on a piece of paper. Everything you learn is these dots. Don't you want to fill the sheet?" You want to remove these blank spaces and fill it as much as possible.

My view is that if I fully concentrate on my field and read papers in my field, I will fill a very dense piece of this paper. If I have interest in physics, biology, neuroscience, finance, and so on, it's further along on this piece of paper. If they can connect to the denser parts, I might have new thoughts that I wouldn't have had otherwise. I am a deep believer of really expanding knowledge because it makes you think faster. Maybe two parts of your brain that weren't connected suddenly connect, and you have this quick pathway. It allows creativity to develop.

CVN: *Is a minute spent teaching a minute lost or does it enter into learning?*

"I get very excited when I have a connection between ideas that I didn't have before or that didn't already exist in the literature"

Zoya: It very much enters because **teachers are solidifying pathways that may be very far away**. I taught math and computer science to different ages ranging from 10 years old up to undergrad and some grad students. When I teach, I don't try to teach along one pathway. I'll explain something in maybe five different ways in order to enforce different connections. I also make sure that people with one set of knowledge will understand it, and people with another set of knowledge will also understand it in another way.

"I'm interested in human memory and attention and how they can apply to applications"

CVN: *Do you see yourself teaching in the future?*

Zoya: I do. I would really like to. I think part of how to pick my future career will depend on whether or not there is a way that I could teach. **I would really like an ability to communicate concepts to a large audience** to really make sure that the ideas are connected in interesting ways. I think that I can connect things in interesting ways and give people a higher level of understanding. I start at a higher level and then descend into the details. For many people it's a very helpful approach. I can give these overviews and I would really like to have a chance to do that.

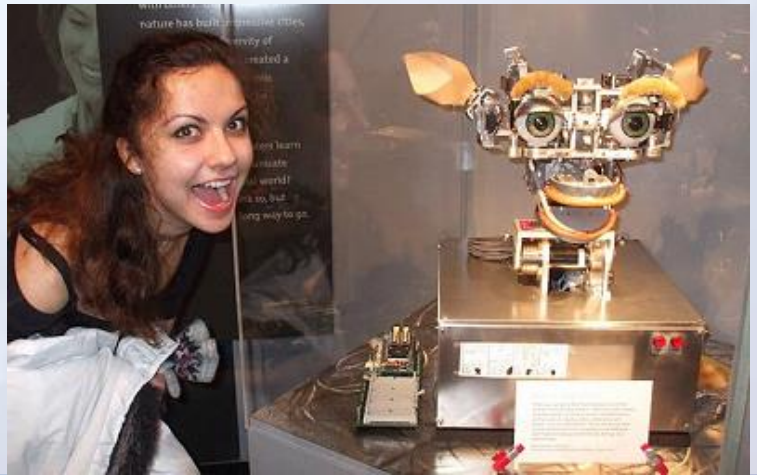
CVN: *Why are you so passionate about that?*

Zoya: I get very excited when I have a connection between ideas that I didn't have before or that didn't already exist in the literature. Things that were described maybe didn't click to me right away or I didn't quite go along with that explanation. If I am able to connect things in a new way, **I have a chance to contribute to this whole field or to a new generation of learners**. If it didn't quite click to me in that particular way, suddenly there is something to contribute. I get excited about that.

CVN: *What is your current work?*

Zoya: I work in an interdisciplinary field with a **cognitive science** professor and a **computer science** professor. A lot of what I do is in between: I'm interested in human memory and attention and how they can apply to applications. This means where people look and how they process images. Those images could be natural images, but they could also be more cognitively demanding tasks like looking at graphs, data visualization, and charts. I'm interested in knowing how people start the process, how they look at this image and how they integrate it in their head to form some kind of understanding.

CVN: *Do you want to reproduce this*



understanding inside robots to try to make them work like we do? Or you only want to learn better the way we work?

Zoya: My work has those two goals exactly, because a cognitive science goal is to understand how it works for humans. From the computer vision side of things, I seek to leverage that knowledge to more efficiently process images or decompose images in the way humans would decompose them, in order to make predictions. Also, we can use that to make smart applications that make use of human cognition as a hack.

CVN: *What do you dream of producing with that?*

Zoya: It goes back to the education domain. It's the ability to customize visual content. I'm not just talking about natural images. I'm talking about content that can be used to teach concepts like graphs and charts. It could also extend to slideshow presentations. I want to be able to **connect that to an individual user to whom this content would adapt**. If I know where this user is looking and how this user is processing this content, I can adapt it to make this user learn more efficiently and process more content during a fixed amount of time.

CVN: *Can you give me a tip for memory?*

Zoya: Whether with using visual slides or presentations, I try to **increase the number of associations** that you could have into a memory. When I say something, I want to evoke a very clear, tangible concept. If I say something at a very abstract level, it's not going to be grounded in your memory. If it can include, let's say, anecdotes or paint a very bright picture in your mind, or show you a very clear picture or diagram to evoke some kind of emotional response, or evoke an association that you might have to something else, **you're more likely to retrieve anything I say** around that time because it's clearly imprinted.

CVN: *I know that you want to break some stereotypes about scientists and engineers. What are these stereotypes?*

Zoya: It goes back a little bit to the conversation we had at the beginning about what kind of personalities or people are suited for particular careers and what is their daily life like; what they do and what kind of mindsets and hobbies they have. Some of the initiatives that I've done before have been to counter it a little bit. This extends beyond showing women in science and computer science. It's important to show children that there is a diversity of people in computer science and the other hard sciences. More than just showing them women, it's important to show them people of very different backgrounds with very different experience levels and interests.

"Show children that there is a diversity of people in computer science and the other hard sciences"

CVN: *You have ideas also regarding what people say about women in science.*

Zoya: I do not personally draw attention to myself as a woman in computer science. I call myself a computer scientist, and I happen to be a woman. I prefer not to put the woman first and foremost. In doing that, I push myself to achieve and represent myself as a scientist.

"I'm very positively charged emotionally - it compensates for any lack of confidence"

CVN: *You seem very confident. Isn't there a bad day when you hear something that you didn't want to hear? How do you get over it?*

Zoya: I don't think my responses are any different than my male colleagues'. As I said, what I think I have is some kind of energy with which I can potentially compensate for the lack of confidence. When it's time to give a talk somewhere, of course, I am worried. At the same time, I'm so energized. I'm very positively charged emotionally and it compensates for any lack of confidence. If I do feel excited about something, I will be able to communicate that concept. Over time, I've been able to focus more about that excitement. Since I want to carry this message across, then nothing else matters.

CVN: *I'd like to talk about your experience with many cultures. You were born in one country, raised in another country, and now you study in another country. What was it like for you to come from Russia? What did it mean for you to come to Canada at such an early age? How did that help you? Or was it an obstacle for you?*

Zoya: It connects a slightly different idea, which is that when I traveled and lived in Russia or Israel, it was during early parts of my life. It's hard to pin down exactly to what extent it influenced me. It definitely influenced my personality. One of the biggest things about moving so often between countries, or once we were in Canada, moving between schools or neighborhoods, was the ability to be constantly surrounded by different people. It has shaped me into the person I am today. I have an ability to communicate with people of all different experience levels and backgrounds. It's very important to me. I try to always surround myself with people from very different fields because I can always learn something from them. This goes back to the piece of paper… I really try to add more dots and connect ideas in my head. Not just ideas, but also thoughts on how the world works and how people think about other things.

"A whole diversity of viewpoints"

My cultural backgrounds, but also my choice of friends and social groups, is very much governed by this idea of wanting to understand a whole diversity of viewpoints and individuals.

CVN: *You told me a few minutes ago that if I want to impress the memory of people, I have to use certain techniques to make my speech more vivid in their memory. What can you tell us now so that you, Zoya, remain impressed in our memory after this talk? What is the message that you want to leave with us?*

"Keeping your eyes very widely open"

Zoya: [*laughs*] That is not an easy question! I think a good image to paint in people's minds that will combine what we've talked about is **the idea of really big eyes. It's the idea of keeping your eyes very widely open** for curiosity, but also for exploration and noticing more elements around you, anything beyond just the computer screen.

CVN: *So you want to be the camera of the computer screen with the capabilities and the CPU of your computer?*

Zoya: [*laughs*] The computer is not quite there yet. Computational processing isn't quite there yet. I want people to have in mind "big eyes" that see widely and really take in the world around them. Recognize that all the different things you see in the world, the experiences that we have, will affect the way we feel about things. That is something that maybe people take for granted. The moment you walk outside the door, the experiences you have on the street will influence your thinking, your problem solving, and the way that you construct your neural maps.

CVN: *What if you see something you don't like?*

Zoya: It still shapes you. It makes you think about why you don't like that. It's not that you don't have to see everything that you like. **You have to be exposed to things you like and don't like to understand better**. It's going to enrich your mind and the way you are able to think about things. Otherwise you are going to be narrow minded. You don't want to put a filter on your eyes. **You really want those eyes to be open and unconstrained**. As a computer scientist, you want to be able to see everything, beyond the monitor of your computer.

ICT4Life - Life Improvement for Elderly

We had the opportunity to speak with **Stylianos (Stelios) Asteriadis**, **Mirela Popa** and **Dario Dotti**, some of the developers behind the upcoming application **ICT4LIFE**. The project provides services for integrated care by employing user-friendly ICT tools, in order to improve the quality of life for patients with **Parkinson's**, **Alzheimer's**, and other dementias.

The team is based at the **University of Maastricht, the Netherlands**, in the Data Science and Knowledge Engineering Department (DKE). Asteriadis works as an assistant professor; Popa is a researcher in Artificial Intelligence, Data Mining, and Human-Computer Interaction; Dotti is a PhD candidate at the university.

Ordinarily, a doctor diagnoses a patient with dementia or Parkinson's according to general symptoms, but the conclusions on the different symptoms are very generic. Because symptoms can manifest differently in each person, the generalizations of different symptoms limit the degree to which doctors can assist people with these diseases, especially when they live by themselves.

From left to right: Stylianos (Stelios) Asteriadis, Mirela Popa and Dario Dotti of the ICT4Life team

> ***"Improving the quality of life for patients with Parkinson's, Alzheimer's, and other dementias"***

Today, in the era of big data, technology is at a mature stage with advanced techniques such as sensor technology in assisted living and data. The ICT4LIFE development team are using these technologies to provide more personalized diagnoses and to diagnosis patients at earlier stages.

Using data analytics, the system learns the behaviors or habits of a person in their home in order to infer patterns. The team assured us that they do this taking special care to respect the person's privacy. They decided to use small sensors and choose proper placement so that the system wouldn't disturb a person's daily life. Then, the system is ready to alert caregivers or doctors of irregular behavior which may indicate gradually forthcoming deteriorations.

For example, a person with Alzheimer's commonly suffers from depression too. If the system recognizes the normal

activity of a person with early stages of Alzheimer's, it can then infer gradual changes such as if the person stays in bed for longer periods of time. Then, the application alerts caregivers of these changes. It can also help doctors prescribe activities such as cognitive games to keep the patient occupied or social activities with their family and friends that help prevent depression.

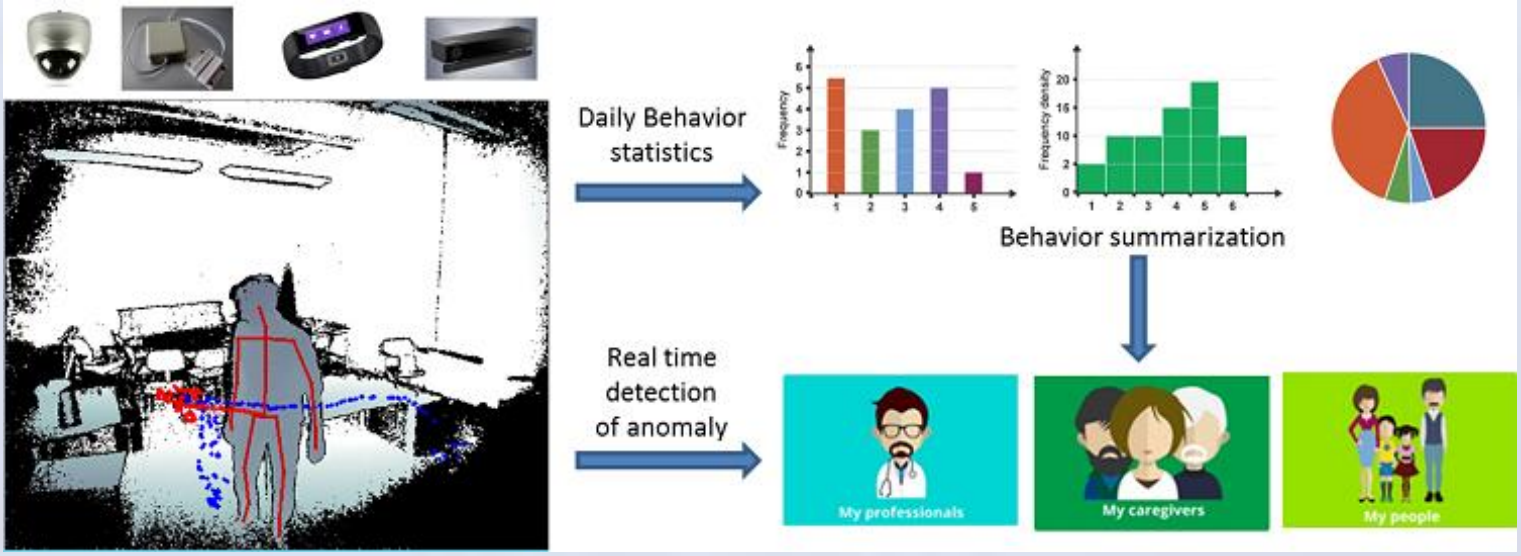
In addition, if a person leaves the house in a confused or agitated state, that might put him or her in danger. In this case, the application alerts the caregivers of the absence. ICT4LIFE has also discussed the possibility of adding a **GPS system** to locate the person, but it hasn't been implemented yet.

Although research and other technology to help people with these conditions already exists, ICT4LIFE enhances the technology to improve the care of these patients. For instance, the computer vision datasets previously available online were mainly non-spontaneous activities of people in indoor environments. Moreover, to the team's knowledge, existing datasets are not shared, in order to preserve the privacy of patients.

The ICT4LIFE development team handles the machine intelligence part of the project: this includes many types of sensors such as magnetic sensors placed on furniture, doors, windows. The project also involves medical data that comes to their platform in real time such as monitoring the heart rate. The medical data provides a static parameter related to the condition of the patient such as if the person has diabetes or other conditions that they need to take into account.

As a **multidisciplinary project**, involving also biomedical tasks, they must coordinate with many different partners from different areas such as between one doctor and another or between the doctors and the engineers, in what can be considered an horizontal system involving many disciplines and skills. This input, coupled with computer vision, leads to a **big data analysis problem**.

Big data is now at its height in the industry, but the applications mainly come from the big players in the field like Google, Facebook, or other Silicon Valley companies. ICT4LIFE seeks to take all of this technology to address



real social needs, beyond just social networking or search engines.

During development, big data analysis proved difficult. They meet these challenges by taking advantage of **deep learning techniques**: in fact, while data is very noisy, deep learning provides a promising direction for this system.

Meanwhile, they wanted to tackle the issue of personalization, especially when diagnosing patients who may exhibit varying symptoms to different degrees. Deep learning is an excellent technique to do that.

From an algorithmic point of view, they want to propose a **multi-modal fusion technique** to the research community. The technique looks at how multi-modality can be associated and mapped with personalization in end-to-end assisted living for elderly people.

In order to create a successful product that can work in real life, it's almost impossible to have **the label or ground truth for every situation**. When they detect anomalies in a person's behavior, the system cannot know what is normal for someone and what is not. Therefore, a supervised approach did not work very well in this case.

Instead, **unsupervised approach** studies a person's behavior without human intervention, using different clustering and classification technique to detect and track the behavior of people in a room, such as where people stand and sit or how fast they move.

ICT4LIFE will make all their datasets and publications open to the public. They continue to work towards making the alpha version of their system available during their first pilot by **June 2017**.

There are also funny moments while working on the project. Stylianos, Dario and Mirela told us that, as a part of their research, they developed data sets of activity recognition in which they recorded people asked to find different objects hidden in a room. Sometimes they would ask people to find objects that weren't there at all. Of course, the person's reaction was always entertaining as they continued to search in all of the drawers and cupboards without any success. Here is Mirela, looking for something that is not.

POC - Proof of Concept

Our CEO **Ron Soferman** has launched a series of lectures to provide a robust yet simple overview of how to ensure that computer vision projects respect goals, budget and deadlines. This month we learn about the **Proof of Concept**.

"Starting a POC with very clear goals, definite timetable and a precise work plan"

The **Proof of Concept (POC)** is a feasibility test. The subject is interesting and useful because in the life cycle of R&D activity we always try to check whether there is any additional technology which could be used to solve problems or provide new features to our software.

I could not stress enough the importance of starting a POC with very clear goals, definite timetable and a precise work plan. I mention this subject because it also belongs to the realm of research, with characteristics of open-end investigation. It is true that when you start to dig into a problem, you don't know yet how you are going to solve it, but if you plan it properly and your **requirements are well defined**, you are in better position to progress efficiently.

What are we going to prove with a POC? Which cases are we going to work on during the phases of the POC? We are not solving anything yet, but it is crucial to define the exact data and the exact device that we are going to work on, as well as the exact pathologies or phenomena that we want to detect.

The **availability status of needed data** and the ground truth with which we will perform the validation process must be very clear at this stage. Timetable is key: when we estimate the resources and the **feasibility of solutions**, we have to make sure that we meet deadlines even at the level of initial research. The developer's strong beliefs regarding his or her next steps might hamper the feasibility, most developers underestimating the time needed to reach a full solution.

Time definitions must therefore be set in advance. If some specific results have already been obtained and we believe that they are meaningful, we can think of modifying the timeline. But if no essential results to lean on are at hand, we might suspect that this proof of concept is not working well. In that case, the options are to **modify part of the definitions** or to **switch to a new plan altogether**.

Sometimes, timelines are influenced by technical problems. Especially when data to be used during research is not available yet. Whatever was the level of confidence in collecting the images, it might happen that the device is not well-prepared, that privacy issues arise or simply that images are not good enough. In this case, it is better to freeze the research temporarily, rather than investing too much effort **before the infrastructure is ready**.

My recommendation to project managers is to **plan the POC very carefully**: by that, I believe that they will get the **best results**.

Montage of Retinal Images

Every month, Computer Vision News reviews a successful project. Our main purpose is to show how diverse image processing applications can be and how the different techniques can help to solve technical challenges and physical difficulties. This month we review software for **montage of retinal images**, developed for a client by **RSIP Vision engineers**. Do you have a project in computer vision and image processing? **Contact our consultants**.

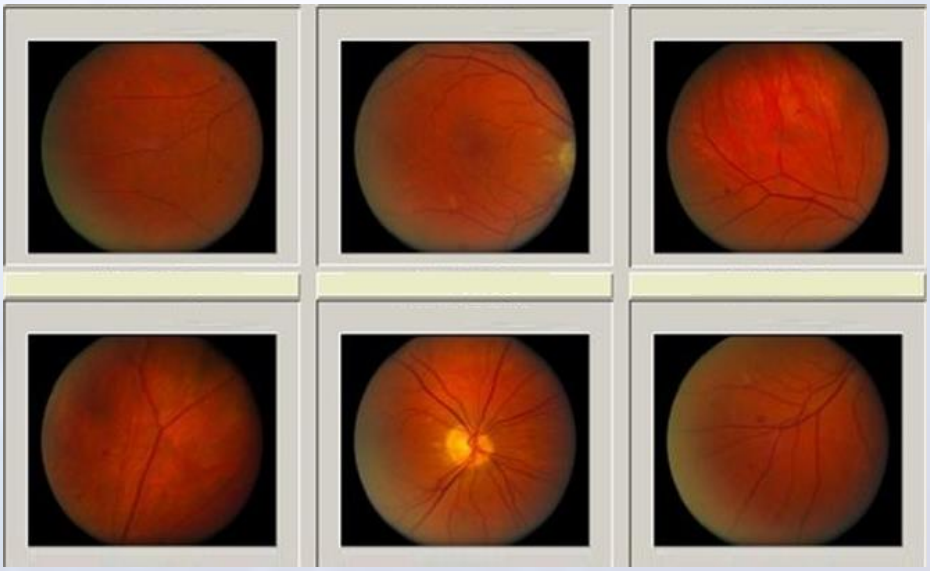
This article relates a very particular project in the field of **ophthalmology**: it implies stitching of fundus (retina) images. **Image stitching** is a very well-known process with many classical applications in computer vision. The most popular one is panorama: after taking images in some sequence, software will arrange them together stitching all images into a full panorama of the whole landscape. Another application involves surveillance, when you stitch together images or other data coming from different cameras with specific fields of view in order to have a **broader vision of the area**.

In eye health, **stitching retina images** is very important: the retina is observed though the pupil of the eye, therefore the field of view is very limited. Even though ophthalmologists often use **dilation agents** to enlarge the size of the pupils, it is still not enough to provide the whole information about the retina.

The common procedure is to ask the patient to look in all four directions, one after the other: this enables to scan the retina from all different angles, so that it is fully included in the images. The purpose is to use the digital images just captured and stitch them together in order to obtain a **full panorama of the retina**: this will enable to study the blood flow in it, understand the status of the vessels and find eventual **leaks or other pathologies**.

Images of the eye are not taken with only one device: they might come from regular cameras, fluorescein (which emphasizes blood vessels), Fundus AutoFluorescence (FAF), IndoCyanine Green (ICG) contrast agent and even Optical Coherence Tomography (OCT). All these modalities offer only a **limited field of view** per every image, hence the need to stitch them one to another if we want to **see the retina as a whole**.

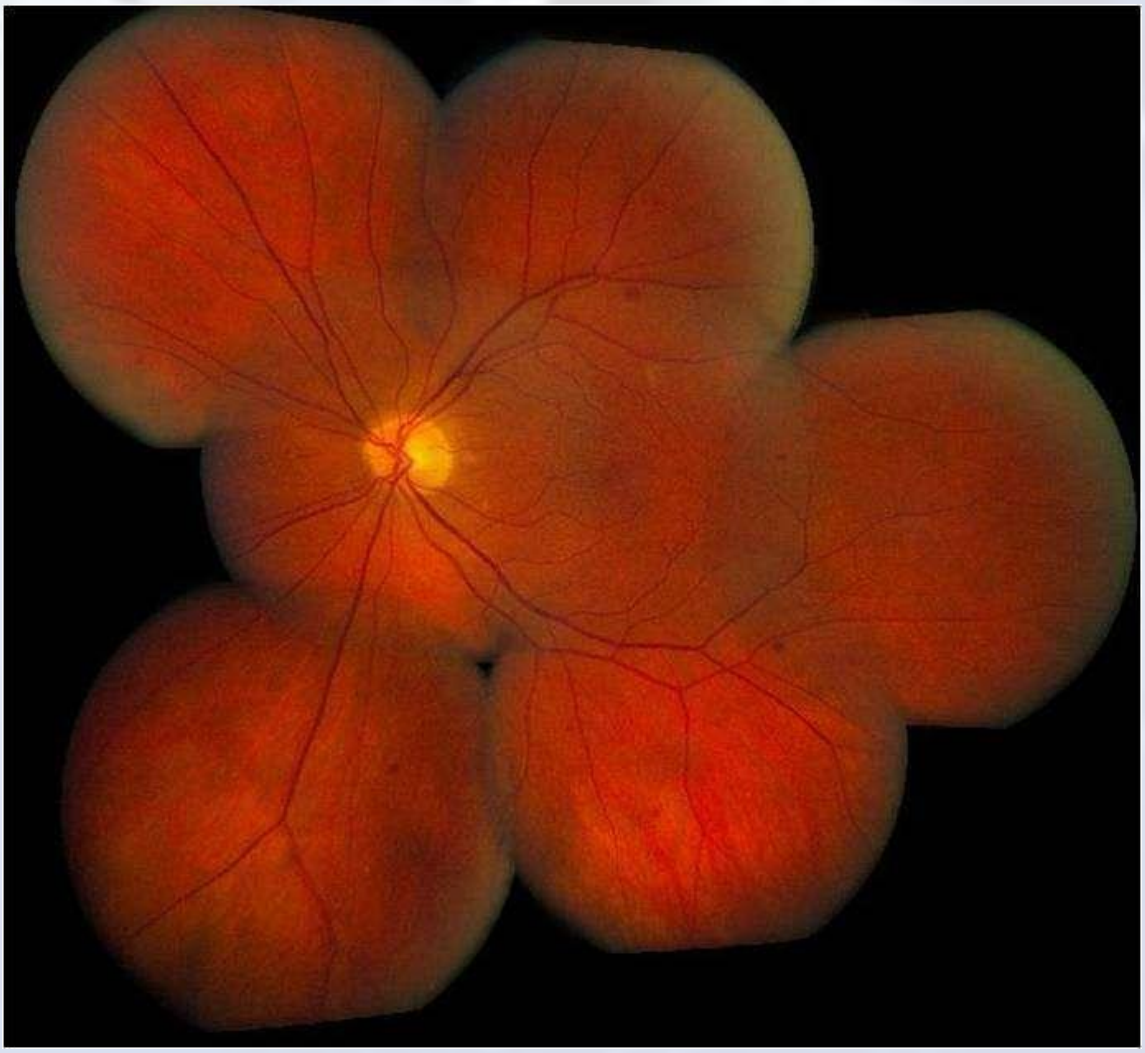
Stitch them one to another to see the retina as a whole



The main technique to do that is features detection: **bifurcations of blood vessels** are taken as landmarks and guide us to stitch the images, just like the ophthalmologist would do if there was no montage software.

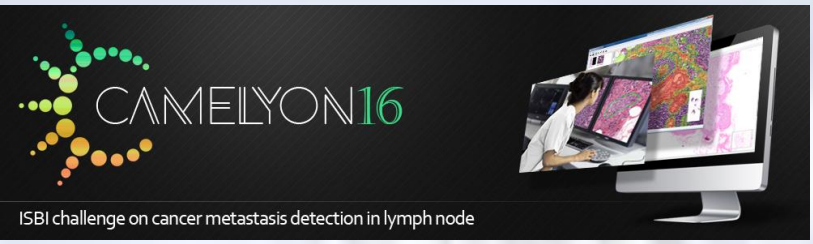
Even after that, not all matchings will be accurate yet: techniques like **RANSAC** are used to add consistency and eliminate incorrect matchings. Points of interest, even though correctly detected in every image, cannot be correctly paired yet and it is interesting to understand the reason for that. Generally, when we want to match and stitch, it is important to understand what transformation we expect: naïve stitching will generate rigid transformation and even some minor rotation. But since the eye is a 3D sphere, its shape requires a more elaborated transformation in order to generate a correct stitching of the retina images in the space: in other words, each image is a projection of parts of a sphere over a 2D detector in the camera. We do the inverse transformation, **from 2D back to 3D images**, which are ready to be stitched only when this process is done correctly. Inaccurate transformation would result in duplication and breaks of blood vessels, for a very inaccurate panoramic view. The task of RANSAC is to make sure that this transformation is uniform and that all the paired matchings are true.

Once this process is completed with the required attention, the exact montage of retinal images will be ready to be presented to the ophthalmologist in a 2D format.



CAMELYON16 - from ISBI 2016

Every month, Computer Vision News reviews a challenge related to our field. If you do not take part in challenges, but are interested to know the new methods proposed by the scientific community to solve them, this section is for you. This month we have chosen to review the **Cancer Metastasis Detection in Lymph Nodes Challenge**, organized around ISBI 2016, which was held earlier this year in Prague: **Cancer Metastasis Detection in Lymph Nodes**. The website of the challenge, with all its related resources, is **here**.



Background

The Cancer Metastasis Detection in Lymph Nodes (CAMELYON16), was designed to evaluate new and existing algorithms for automated detection of micro- and macro-metastases in H&E (hematoxylin and eosin) stained whole-slide images of lymph node sections. This is the first challenge using whole-slide images in histopathology and it will run for two years: 2016 and 2017.

Motivation

This task is of crucial medical importance: **cancer can spread to distant parts of the body**, spreading via the walls of lymph nodes or blood vessels. When this happens, it is called **metastatic cancer**. Lymph node metastases occur in most cancer types (e.g. breast, prostate, colon/rectum). Lymph nodes are small glands that filter lymph, the fluid that circulates through the lymphatic system. The lymph nodes in the underarm are the first place **breast cancer** is likely to spread. Metastatic involvement of lymph nodes is one of the most important prognostic variables in breast cancer. Once cancer spreads, it can be hard to control: chances for successful treatment are lower when cancer has spread to the lymph nodes.

However, the diagnostic procedure for pathologists is tedious, time-consuming and susceptible to misinterpretation. Therefore, a successful solution to this challenge would hold great promise to reduce the workload of the pathologists while at the same time reduce the subjectivity and the costs of diagnosis. The 2016 challenge focuses on sentinel lymph nodes of breast cancer patients: a **sentinel lymph node** is defined as the first lymph node to which cancer cells are most likely to spread from a primary tumor. The challenge provides a large dataset from the Radboud University Medical Center and the University Medical Center in Utrecht (both in the Netherlands).

"Improving the detection of cancer metastasis in lymph node images"

Solution

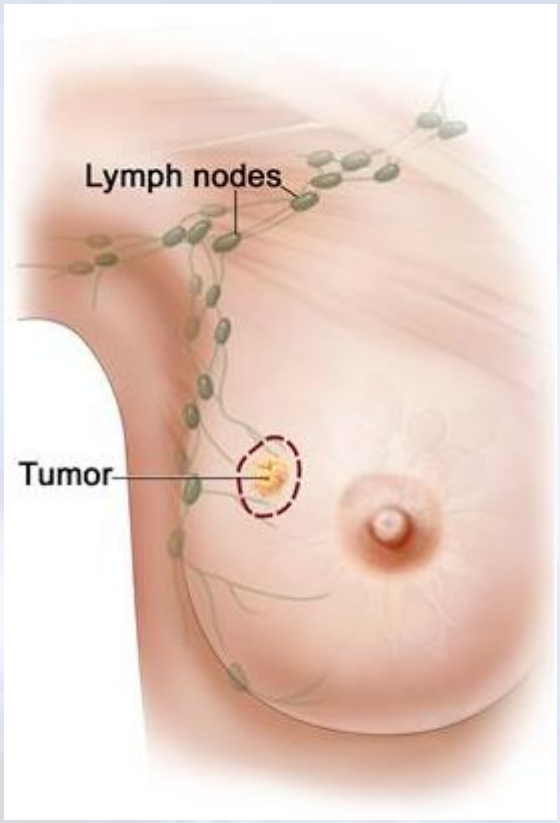
This challenge aims at improving the detection of cancer metastasis in lymph node images: the goal is to develop algorithms for a (fully) automated analysis of whole-slide images to detect or grade cancer, to predict prognosis or identify metastases.

Results

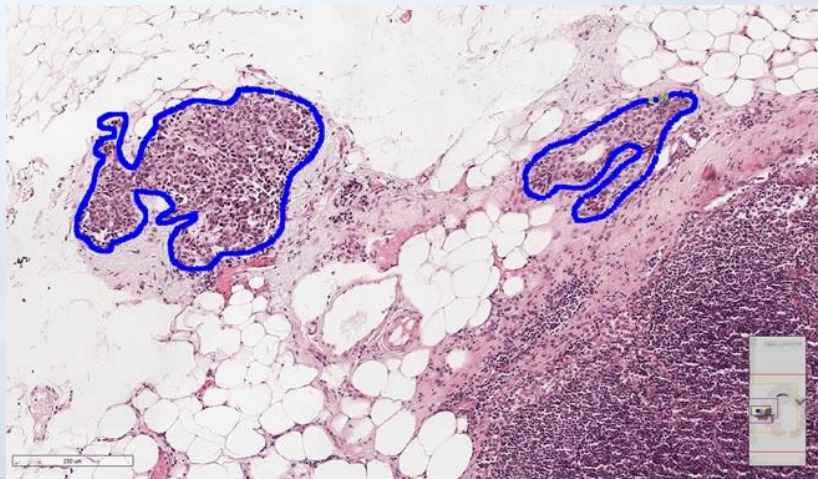
Two strategies were retained for evaluating the performance of the algorithms, giving way to two leaderboards: (A) Slide-based Evaluation: how the algorithms discriminate between slides containing metastasis and normal slides. (B) Lesion-based Evaluation: a free-response receiver operating characteristic (FROC) curve, defined as the plot of sensitivity versus the average number of false-positives per image.

Both leaderboards were led by Harvard Medical School teams. We would like to mention the team formed by the Harvard Medical School (BIDMC) and the Massachusetts Institute of Technology (CSAIL): the authors explain in this page both their methods and results: their area under the ROC curve surpassed the one given by the pathologist in the study.

Sentinel lymph node: the first lymph node to which cancer cells are most likely to spread from a primary tumor



The second year's challenge CAMELYON17 will strengthen the challenge by moving from slide-level analysis to patient-level analysis (that means combining the assessment of multiple lymph node slides into one outcome). This is expected to bring the efforts closer to direct usefulness in clinical setting. Compared to CAMELYON16, the dataset will be significantly extended and will contain images from four medical centers and pathology labs. Registrations are already open at the challenge website and the CAMELYON17 workshop will be held at ISBI 2017 in Melbourne, Australia.



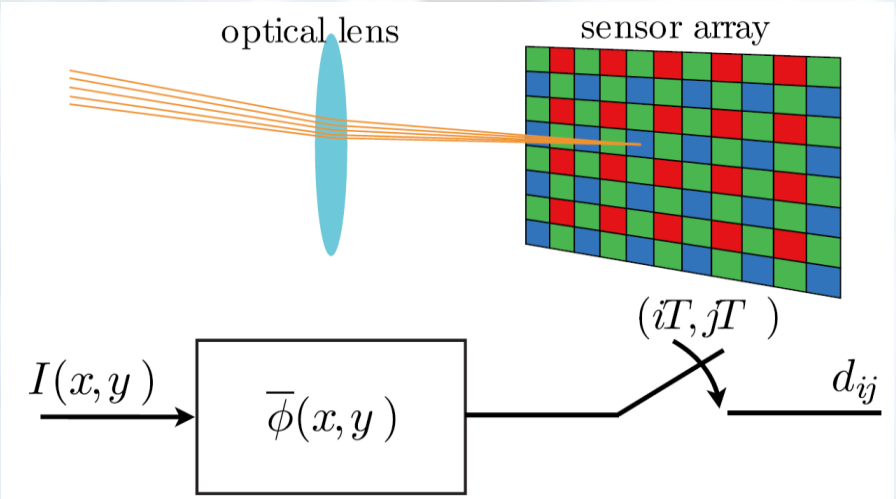
Shapes From Pixels

Every month, Computer Vision News reviews a research from our field. This month we have chosen to review **Shapes From Pixels**, a research paper studying the problem of recovering shape images with smooth boundaries from a set of samples. We are indebted to the authors (**Mitra Fatemi**, **Arash Amini**, **Loic Baboulaz**, and **Martin Vetterli**) for allowing us to use their images to illustrate this review. The full paper is **here**.

Recovering continuous-domain visual signals from their samples

Background:

Sampling is a key procedure in digital signal acquisition. Continuous-domain signals are converted into a sequence of numbers to enable storage and processing of the data. With regards to digital image acquisition, the optical lens and the sensor array are responsible for the sampling process: the optical system determines the impulse response of the involved filter, which is called the **Point Spread Function (PSF)**; the sensor array controls the number of samples and the sampling resolution. The purpose of this paper is to study the problem of **recovering continuous-domain visual signals from their samples**.



Sampling system implemented by a digital camera; the effect of the optical lens and the sensor array on 2D visual signals can be modeled by filtering followed by sampling in space

Motivation:

Recovering a continuous image will enable applications to have: (1) unlimited zooming capability; (2) the capacity to arbitrarily rotate the image; (3) improved feature extraction and object localization.

Challenge:

Continuous-domain visual signals are usually captured as discrete (digital) images. In general, this operation is not invertible, in the sense that the continuous-domain signal cannot be exactly reconstructed based on the discrete image, unless it satisfies certain constraints (e.g. band limitedness). The challenge is how to **recover the shapes directly from the pixel data**, without intermediate curve fitting steps or the band limitedness assumption, usually used for this purpose.

Novelty:

The problem of reconstructing a continuous-domain shape from a gray-scale discrete image is essentially equivalent to the interpolation of pixels in a way that generates a binary image. The authors of Shapes From Pixels formulate this as a **minimization problem**, where the domain is the sampling kernels and the constraints encode the sampling relation. The minimizers will be shapes with minimum perimeter and smooth boundaries. However, this is a non-convex, computationally intractable, problem. The authors introduce a reducibility condition on the sampling of the discrete image and prove that, when it is satisfied, the problem becomes convex.

Method:

I(x,y) denotes the continuous image with pixel values in the range Ω=[1,0]2. D denotes the discrete image (m×m-pixel). In the consistent image recovery problem, we seek the approximation $\tilde{I}$ of the original image that generates the same measurement pixels. We can relate the d_{ij}, 1 ≤ i, j ≤ m of D to the image I(x, y) as:

$$d_{ij} = \frac{1}{T^2}\,\overline{\phi}\left(\frac{x}{T}, \frac{y}{T}\right) * I(x,y)\,|_{(x,y)=(jT,iT)} = \iint_{\Omega} \frac{1}{T^2}\phi\left(\frac{x}{T} - j, \frac{y}{T} - i\right) I(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

where T is the sampling period. The equivalent representation of index of d_{ij} in the vertical raster scan of D is defined as follows:

$$f_k(x,y) = \frac{1}{T^2}\phi\left(\frac{x}{T} - \lceil k/m \rceil, \frac{y}{T} - ((k \mod m) + 1)\right)$$

Where k=(j-1)m+i, $1 \le k \le m^2$ and f_k is the sampling kernel associated with d_k.

Next, the authors define the set of all non-negative-valued images over Ω that are consistent with $[d_k]_{1 \le k \le m^2}$ as:

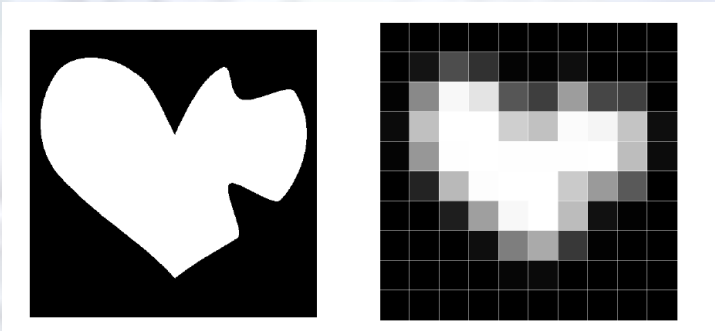
$$\mathcal{C}_{\Omega}(\mathbf{D};\ f_1, \ldots, f_{m^2}) = \left\{ I \in BV(\Omega),\ I \ge 0;\ \iint_{\Omega} I f_k\,\mathrm{d}x\,\mathrm{d}y = d_k,\ 1 \le k \le m^2 \right\}.$$

Where BV(Ω) is the set of functions over Ω with bounded variation. Consistent image recovery is equivalent to finding an element of $C_\Omega(\mathbf{D}; f_1, \cdots, f_{m^2})$.

$\mathcal{S}$ is called a shape if it is the union of a finite number of connected subsets of Ω. $\chi\mathcal{S}(x, y)$ denote the characteristic function of $\mathcal{S}$ over Ω:

$$\chi_{\mathcal{S}}(x, y) = \begin{cases} 1, & \text{if } (x, y) \in \mathcal{S} \\ 0, & \text{if } (x, y) \in \Omega \setminus \mathcal{S}. \end{cases}$$

For the consistent shape recovery problem, the authors limit the permissible solutions to the above shape characteristic functions. In this case, we call $\chi\mathcal{S}$ a shape image. Below is an example of a shape characteristic function and its associated 10×10 discrete image:



Now, the consistent shape reconstruction problem is equivalent to finding a shape image $I = \chi\mathcal{S}(x, y) \in C_\Omega(\mathbf{D}; f_1, \cdots, f_{m^2})$. Solution is the minimizers of the following constraints:

$$\begin{aligned} &\inf_{\mathcal{S}\subset\Omega, \chi_{\mathcal{S}}\in BV} \ \mathrm{Per}(\mathcal{S}), && (P_0) \\ &\text{s.t. } I = \chi_{\mathcal{S}} \in C_\Omega(\mathbf{D};\ f_1, \ldots, f_{m^2}), \end{aligned}$$

where Per($\mathcal{S}$) is the perimeter of $\mathcal{S}$. Problem (P_0) is a variation non-convex problem prone to having many local minima. In the simplest scenario, having only a single pixel is a well-studied topic known as the **Cheeger problem**. There is already a rich literature regarding the existence, uniqueness properties and regularity of such sets for almost arbitrary kernels f.

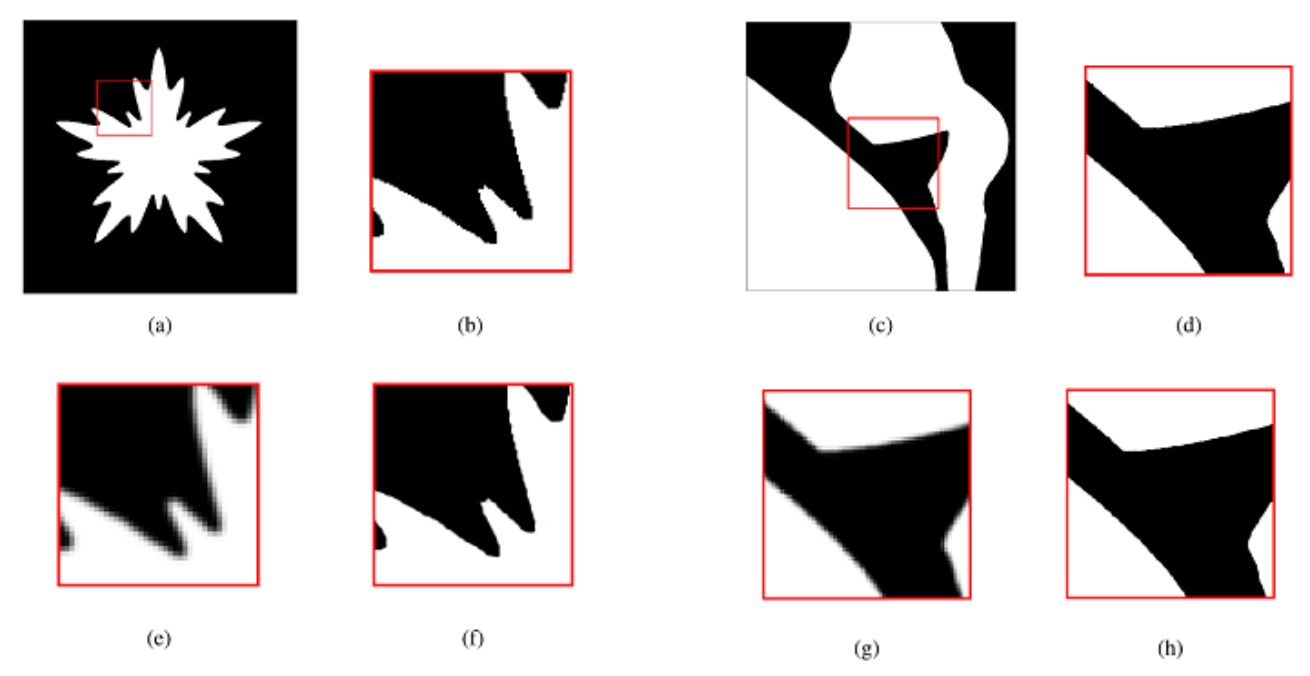
To adapt the Cheeger problem for images with more than one pixel and gray-level values the authors define the reducibility constraint: $(\mathbf{D}; f_1, \cdots, f_{m^2})$ if A (the index set of active pixels) can be partitioned into K_1 and K_2 such that:

$$\begin{aligned} &\text{(i)} \ \ \forall k \in K_1, \boldsymbol{\lambda} \in \Delta_\rho, \lambda_k = 0 : \iint_{\Omega_r} I^{\boldsymbol{\lambda}} f_k \mathrm{d}x\mathrm{d}y < d_k, \\ &\text{(ii)} \ \ \forall k \in K_2, \boldsymbol{\lambda} \in \Delta_\rho : \iint_{\Omega_r} I^{\boldsymbol{\lambda}} f_k \mathrm{d}x\mathrm{d}y \le d_k. \end{aligned}$$

In the paper, the authors prove that as long as the above reducibility constraint is met, the **multi-constraint minimization problem** is equivalent to the Cheeger problem.

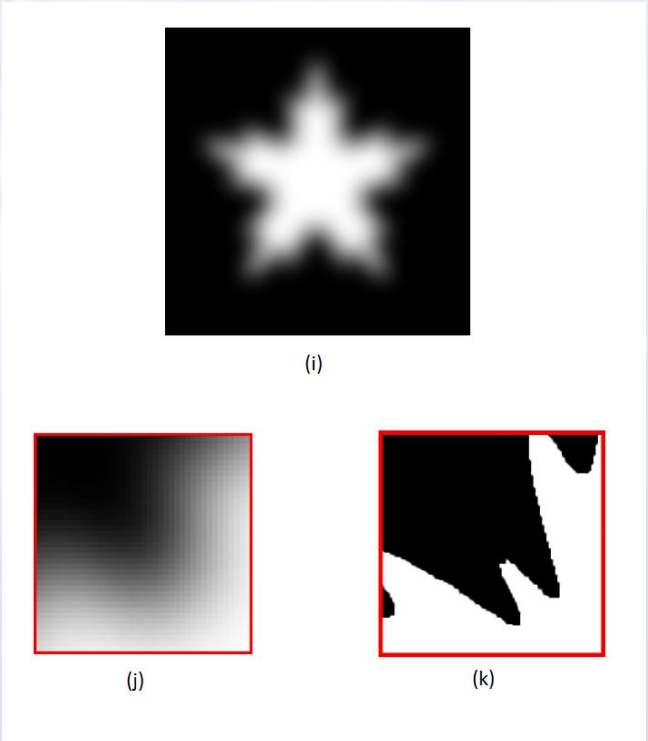
Essentially, the reducibility condition calls for smooth changes in the values of the neighboring pixels. From this perspective, this is an intuitive requirement on the minimum sampling density that is needed for tracking local changes in the shape boundaries.

Results:



To illustrate some of the algorithm's capabilities, the authors took shape images at the resolution 1000x1000, then created 200x200 discrete images of them with biquadratic B-spline sampling kernels; successively, they reconstructed continuous shapes and compared them to the originals. The figure above demonstrates this process: (a) and (c) are the original images; (b) and (d) are an enlarged section of each, on which the rest of the illustration focuses; (e) and (g) are the respective discrete images; (f) and (h) are the reconstructed shapes. In the figure on the right, (i) is a 200x200 discrete image corresponding to (a) above, which was created with sampling kernels that are shifts of a stretched biquadratic B-spline with an effective support of 40x40 pixels; (j) is the corresponding enlarged section, as above, and (k) is the reconstructed image. The recovered image (without any thresholding) has PSNR of 33.8096 dB with respect to the original shape and a measurement PSNR of 75.0489 dB.

The multi-constraint minimization problem is equivalent to the Cheeger problem



The Fall Meeting of the NVPHBV - Dutch Society of Pattern Recognition and Image Processing

Event

Veronika Cheplygina is a postdoctoral researcher at the Biomedical Imaging Group Rotterdam, Erasmus Medical Center, and a member of the NVPHBV board. She tells us about the **Fall Meeting**, held in Eindhoven on November 11.

"I really liked the interaction between the talks, and how each of the subjects linked to the other talks."

Veronika Cheplygina (left) with speaker Diana Mateus

Computer Vision News: *Veronika, what is the Dutch Society of Pattern Recognition and Image Processing?*

Veronika Cheplygina: It's a society with different types of people working in these areas. It includes researchers, but also people in industry. This year is the society's 35th anniversary. There are about 300 members. All of us are also members of the **International Association of Pattern Recognition** which is a worldwide entity. We are the regional part of that. We have a board of five members, and I am on the board. We meet twice a year when we organize a meeting, and then we invite all of the members and anybody else who might be interested.

CVN: *I understand that the conference was free. Who sponsored it?*

Veronika: We sponsor it ourselves. Many of us have academic affiliations so we can get a place for the symposium. Members pay a small contribution each year, so we have a little money to buy lunch and other things for the event.

CVN: *What are the goals of this operation?*

Veronika: It's a place for certain groups

working on similar topics to meet and discuss their latest works. The Netherlands are a small country, so it's easier to travel between cities. It's great for PhD students because not all of them will end up in academic careers. They can see if they would like to start a business or work at a research institution. I'm also trying to have talks at the meetings about this. Not only scientific talks, but people talking about their career.

CVN: *What can you tell us about the fall meeting?*

Veronika: It was great. I think it was one of the most highly attended meetings so far. About 60 people registered. We had three invited speakers, who each gave an hour talk. **Joaquin Vanschoren** talked about a platform, **OpenML**, which recently won the **Dutch Data Prize for Exact and Technical Sciences**. The idea is that of a website where you can share results and algorithms. The website is OpenML.org.

The other speaker was **Diana Mateus** from the **University of Munich**. She talked about two different applications that she works on in segmentation and registration about interactive machine learning, and how we can interact with experts to get more labels. BTW, she organized the LABELS workshop at MICCAI 2016.

The last speaker was **Julian de Wit**. He is not somebody that we knew in the society, but I was told that he ranked very high in the Second National Data Science Bowl, a challenge hosted by Kaggle.com. He is a freelance software engineer, and he participated in the challenge by himself, ending up at the third place. He was only beaten by two very large teams. It was very exciting to hear how he did this and for him to reveal his secret. He just went ahead and labelled some of the data himself without knowledge of heart disease. That tied in very nicely with the other talks on interactive machine learning and sharing data in competitions. I really liked the interaction between the talks, and how each of the subjects linked to the other talks.

Event

"It's a very unique combination of groups that wouldn't typically meet together at a conference"

Godliver Owomugisha from the University of Groningen, presenting her work on classification of cassava leaves

CVN: *What will be the next step for the organization? Are you planning anything between now and the next meeting in the spring?*

Veronika: The way it usually goes is that we just do the two meetings per year. I think it would be nice to organize more things. You've inspired me a bit to look at your newsletters. That would be nice. I think that we should try to get other people from the society involved.

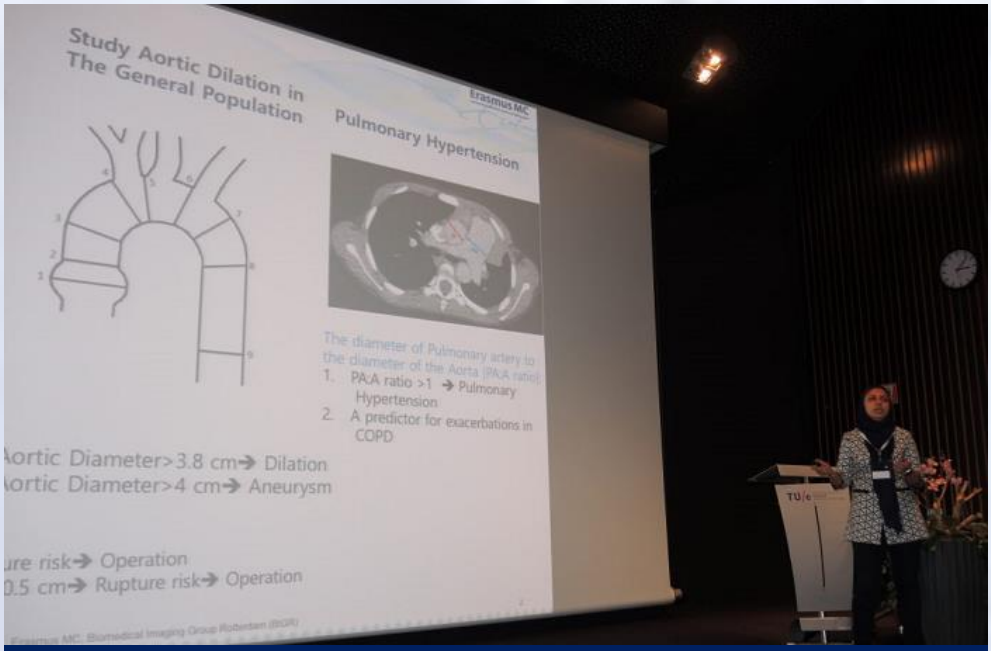
CVN: *What makes these events special in your opinion?*

Veronika: It's a very **unique combination of groups that wouldn't typically meet together at a conference**. People are working on slightly different problems. We have people doing medical imaging, but also people, for example, classifying plant leaves or something entirely different. This event makes it possible for these people to talk to each other.

There were also 10 contributed talks. It's simpler than a conference because they can print up an open problem that they did before. This is quite good for discussion. Now we even ran out of room for more talks, so I had to ask people to talk at the next meeting instead.

CVN: *Can you tell us more about the organization or the meeting?*

Veronika: Maybe you have noticed that we have a very complicated acronym, **NVPHBV** (this is *"Nederlandse Vereniging voor Patroonherkenning en Beeldverwerking"*, the Dutch translation of "*Dutch Society of Pattern Recognition and Image Processing*"). A lot of people complain about it that it's difficult to pronounce.... you can't easily say "are you going to NVPHBV?" in the same way as you can say "are you going to MICCAI?", for example. But at the same time I think it helps people remember the meeting as "the one with that weird acronym". We could change the name to something more popular and include "machine learning", but I think there is also some sentimental or even historical value to it, as it links back to how the field was 30 or 40 years ago - before some of the members were born…



Zahra Sedghi from Biomedical Imaging Group Rotterdam, Erasmus Medical Center, presenting her work on segmentation of the aorta and the pulmonary artery

"It links back to how the field was 30 or 40 years ago - before some of the members were born…"

FREE SUBSCRIPTION

Dear reader,

Do you enjoy reading Computer Vision News? Would you like to receive it **for free in your mailbox** every month?

Subscription Form
(click here, it's free)

You will fill the Subscription Form in **less than 1 minute**. Join many others computer vision professionals and receive all issues of Computer Vision News as soon as we publish them. You can also read Computer Vision News on our website and find in our archive new and old issues as well.

We hate SPAM and promise to keep your email address safe, always.

ICPR 2016: International Conference on Pattern Recognition
Cancun, Mexico Dec. 4-8 Website and Registration

ISVC 2016: International Symposium on Visual Computing
Las Vegas NV, USA Dec. 12-14 Website and Registration

IPTA 2016: Image Processing Theory, Tools and Applications
Oulu, Finland Dec. 12-15 Website and Registration

ICCVISP 2016 : Computer Vision, Image and Signal Processing
London, UK Dec. 15-16 Website and Registration

ICVGIP 2016: Computer Vision, Graphics and Image Processing
Guwahati, India Dec. 18-22 Website and Registration

VLSI Architectures for Signal Processing and Machine Learning
IIT Kharagpur, India Dec. 19-30 Website and Registration

Jerusalem Winter School in Computer Science and Engineering
Jerusalem, Israel Jan. 8-12 Website and Registration

ICCVIP 2017 on Computer Vision and Image Processing
Zurich, Switzerland Jan. 13-14 Website and Registration

ICAIEPR 2017 on Advanced Information Engineering and Pattern Recognition
Bangkok, Thailand Jan. 19-20 Website and Registration

ICISP 2017 on Imaging and Signal Processing
Innsbruck, Austria Jan. 26-27 Website and Registration

RE•WORK Deep Learning Summit
San Francisco CA, USA Jan. 26-27 Website and Registration

ICMLCS 2017 on Machine Learning and Computer Science
Dubai, UAE Jan. 30-31 Website and Registration

ICCVCG 2017 on Computer Vision and Computer Graphics
Melbourne, Australia Feb. 2-3 Website and Registration

ICCVIP 2017 on Computer Vision and Image Processing
Mumbai, India Feb. 7-8 Website and Registration

ICCVISP 2017 on Computer Vision, Image and Signal Processing
Barcelona, Spain Feb. 26-27 Website and Registration

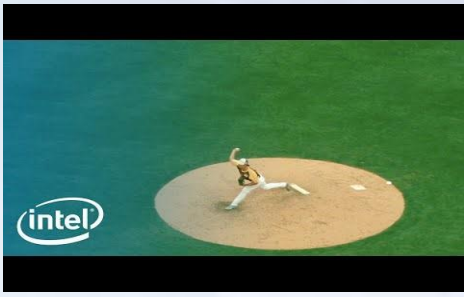
VISIGRAPP on Computer Vision, Imaging and Computer Graphics
Porto, Portugal Feb. 27-Mar 2 Website and Registration

RE•WORK Deep Learning in Healthcare Summit
London, UK Feb. 28-Mar 1 Website and Registration

Did we miss an event? Tell us: editor@ComputerVision.News

Computer Vision News lists some of the great stories that we have just found somewhere else. We share them with you, adding a short comment. Enjoy!

How Machine Vision Is Changing Sports Spectating

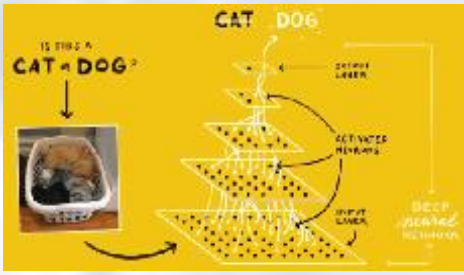


Those who follow sports events are well aware of the increasing use of technology to boost viewer's experience, improve athletes' training and make judging more accurate. Think at Hawk-Eye in tennis and goal-line technology in football. This nice article tells you more about **machine vision**'s impact in this area, including Intel's spectacular 360-degree replay of MLB action. **Read...**

AI Curated a Magazine Using Image Recognition

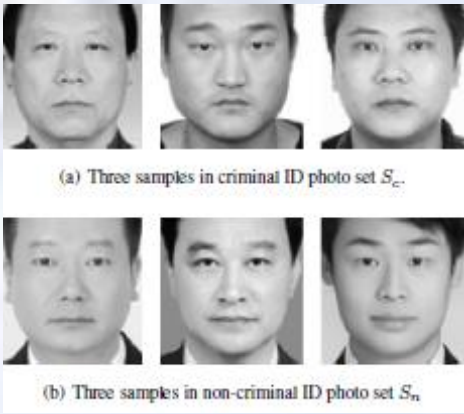
Among the many things that a machine can do with a picture, here is a new one: decide whether it is aesthetic or not. Not like an unskilled viewer would do, but like the 18 million expert photographers members of the EyeEm community, whose magazine was curated by an **AI trained to understand** moods, feelings and beauty. **Read...**

Now Machines Recognize Something After Seeing It Once



We are used to think that **algorithms** need thousands of examples to learn something. Which can be quite annoying when time and data are scarce. It seems that researchers at **Google DeepMind** found a way around that. **Read about it**, **get the paper**, **watch the video**

Automated Inference on Criminality using Face Images



Two **Shanghai Jiao Tong University** researchers claim that their study is the first to offer automated inference on criminality based solely on still face images. They do it exactly as you would expect: **machine learning**, classifiers, a dataset of 1856 real persons controlled for race, gender, age and facial expressions. Half of them were convicted criminals. It seems that variation among criminal faces is much greater than that of the non-criminal faces. **Read...**

Trainspotting via the Caltrain Rider

How many of our readers have taken a ride on a **Caltrain**, between SanFran and the Silicon Valley? Probably most of them. And how many of them think that delay estimates provided by Caltrain are reliable? Probably many less... Engineers at a Silicon Valley firm used **image processing and recognition** to improve arrival predictions. **Read...**

The Motley Fool tells us what is cooking at Apple.

3 revelations found looking into Apple Job Listings! **Read...**