How to do research in Computer Vision?

Jianxiong Xiao

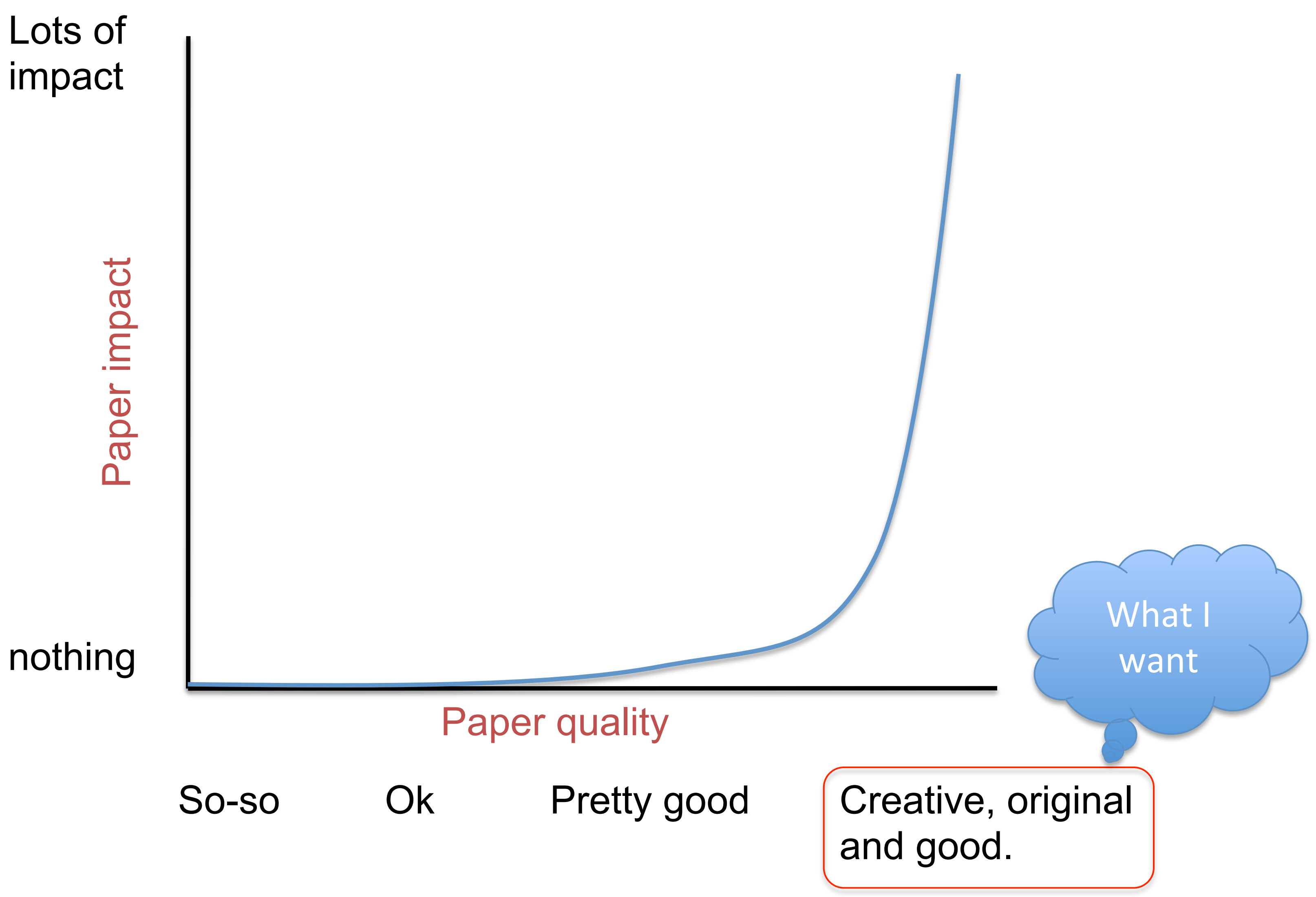
PRINCETON
VISION GROUP

Introduction

- This course has prepared you technical background for research.
- This lecture talks about the general **practice** of doing research (at least in my group).

The paper impact curve

Slide source: Bill Freeman



Math ≠ Science

- A lot of people confuse math with science. Math is not equal to science. Math itself is a discipline of science. But there are many science outside math.
- Most of vision science is outside math (at least for now), with a very small overlapping. Therefore, most of time, ignore the fancy math, at least for most projects.
- Anything can be true in Math. But in real world, everything is very restricted, corrupted, and dirty in some sense. That is what makes CV hard.
- NP-hard is not considered to be hard in CV.

Difficulty ≠ Goodness

- A good research is not necessarily a difficult research.
- Focus on important problems, not the difficult ones.
- Especially, solving a difficult math problem or writing an elegant math proof is not useful for most of computer vision research.

Experimental Science

- Computer Vision is mostly an experimental science now.
- Because it is only 50 years old
- Think about what physics is like when physics is 50 years old

Look at the data

- You know computer vision only when you label enough data manually.

- http://people.csail.mit.edu/torralba/publications/memories.pdf
- *We are under the illusion that seeing is effortless, but frequently the visual system is lazy and makes us believe that we understand something when in fact we don't. Labeling a picture forces us to become aware of the difficulties underlying scene understanding. Suddenly, the act of seeing is not effortless anymore. We have to make an effort in order to understand parts of the picture that we neglected at first glance.*

Visualization: Look at the pixels!

- It is crucial to look at the data carefully and make sure you know the data
- Not just numbers, curves and tables. But visualize it in all possible ways.

Always do the right thing

- how to do a good research? just “always” do the right thing
- never try to be lazy by short-cutting important steps
- always do what our heart tell us that we should do
- carefully analyzing the result and visualization
- play with the algorithms/codes to figure out the best way to go
- We never cut any corners for research. We always do the right thing, no matter how difficult it is. We don’t take the easy and cheap solution.
- So work harder and keep the quality bar high.

Make an point in a paper

- make an point, not just here is the software and it is 2% better than before
- why it is better?
- how do people borrow this idea and carry on further?

Aim Higher: You are a Leader

- We are among the very top research institutes around the world. We have a lot of great resources that others don't, including money and IQ. Therefore, it is our responsibility to lead the field to go towards the right research direction.
- We are the leaders and all of us should act like a leader! The destiny of computer vision research is in our hand (the advance of humanity is in your hands), and what we do significantly affect the progress in vision knowledge for the whole humanity.
- Therefore, we care only about important, creative, novel research.
- Most papers in top conferences/journals is lowered than this standard. And we always aim for the 1% paper that can significantly change the field.
- We don't publish something because we can publish it.
- We publish something because we want to direct the other 99% of researchers to go towards a certain direction that we believe it will produce a breakthrough for research.
- If you cannot even convince yourself, you will not be able to convince others.

Simple is better

- Don’t be pretentious
- Albert Einstein: Everything Should Be Made as Simple as Possible, But Not Simpler

Be creative!

- Computer vision is not solved.
- Which means that all existing algorithms are wrong.
- If you follow the existing works exactly, you are never going to solve the problems successfully.
- The only way is to be creative and struggle to be different.
- All boring research is bad research.
- Don’t believe in authority. Make yourself become authority.

How to do research

from Bill Freeman

http://people.csail.mit.edu/billf/www/papers/doresearch.pdf

Slow down to speed up

- In classes, the world is rigged. There's a simple correct answer and the problem is structured to let you come to that answer. You get feedback with the correct answer within a day after you submit anything.
- Research is different. No one tells you the right answer, we don't know if there is a right answer. We don't know if something doesn't work because there's a silly mistake in the program or because a broad set of assumptions is flawed.
- How do you deal with that? Take things slowly. Verify your assumptions. Understand the thing, whatever it is–the program, the algorithm, or the proof. As you do experiments, only change one thing at a time, so you know what the outcome of the experiment means. It may feel like you're going slowly, but you'll be making much more progress than if you flail around, trying different things, but not understanding what's going on.

don’t tell me “it doesn’t work”

- Of course it doesn’t work. If there’s a single mistake in the chain, the whole thing won’t work, and how could you possibly go through all those steps without making a mistake somewhere?
- What I want to hear instead is something like, “I’ve narrowed down the problem to step B. Until step A, you can see that it works, because you put in X and you get Y out, as we expect. You can see how it fails here at B. I’ve ruled out W and Z as the cause.”
- Please don’t report to me, “This instance doesn’t work”. Why doesn’t it work? Why should it work? Is there a simpler case we can make it work? Do you think it’s a general issue that affects all problems of this category? Can you think of what’s not working? Can you contort things to make an example that does work? At least, can you make it fail worse, so we understand some aspects of the system?

“This sounds like hard work.”

- Yes. It’s no longer about being smart. By now, everyone around you is smart.
- In graduate school, it’s the hard workers who pull ahead. This happens in sports, too. You always read stories about how hard the great players work, being the first ones out to practice, the last ones to leave, etc.
- “How do I get myself to work hard enough to do research well?” It all plays out if you love what you’re doing. You become good at it because you spend time at it and you do that because you enjoy it. So pick something to work on that you can love. If you’re not the type who falls in love with a problem, then just know that working hard is what you have to do to succeed at research.

TSTMTCTMI

- simple toy models always help me. With a good one, you can build up intuition about what matters, which is a big advantage in research.

Here's an example. The color constancy problem is to estimate surface reflectance colors when we only get to observe the wavelength-by-wavelength product of the each surface reflectance spectrum and the unknown illuminant spectrum. A toy model for that problem is to try to estimate the scalars a and b from only observing their product, $y = ab$. There's a surprising richness even to this simple problem, and thinking about it allows you to think through loss functions and other aspects of Bayesian decision theory. I co-authored a paper that discusses $y = ab$ for much of the manuscript. Another toy model: as a proxy for complicated shaded surfaces, a single bump. You get the idea. Having the intuitions from working

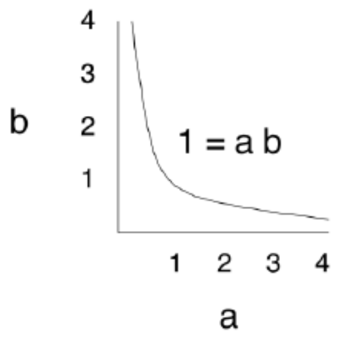


Figure 2: A toy model for the color constancy problem: y = ab

with toy problems gives you a big advantage in the research, because you can figure out what will work by thinking it through with your toy model.

A weak and a strong student

- There is a weak and a strong graduate student. They are both asked by their advisor to try a particular approach to solving a research problem.
- The weak student does exactly what the advisor has asked. But the advisor's solution fails, and the student reports that failure.
- The strong student starts doing what the advisor has asked, sees that it doesn't work, looks around within some epsilon ball of the original proposal to find what does work, and reports that solution.
- The terrible student: I don't have experience doing this at all. Give me N years to learn. After that, I will be TA and do may homework, and I will be too tired and I need to go for vacation.

Literature

- Sometimes it's useful to think that everyone else is an idiot. This lets you do things that no one else is doing. It's best not to be too vocal about that. You can say something like "Oh, I just thought I'd try out this direction".
- It's also sometimes useful to remember that many smart people have worked on this and related problems and written their thoughts and results down in papers. Don't be caught flat-footed with a large body of closely related literature that you aren't familiar with.

final note about doing research

- I hope you love it. I certainly do.
- The research community is a community of people who are passionate about what they do, and we welcome you to it!

How to give a talk

(from Antonio Torralba)

http://www.cs.berkeley.edu/~messer/Bad_talk.html

http://www-psych.stanford.edu/~lera/talk.html

First, some good/bad news

The more you work on a talk, the better it gets: if you work on it for 3 hours, the talk you give will be better than if you had only worked on it for 2 hours. If you work on it for 5 hours, it will be better still. 7 hours, better yet…

All talks are important

There are no unimportant talks.

There are no big or small audiences.

Prepare each talk with the same enthusiasm.

How to give a talk

Delivering:

Look at the audience! Try not to talk to your laptop or to the screen. Instead, look at the other humans in the room.

You have to believe in what you present, be confident… even if it only lasts for the time of your presentation.

Do not be afraid to acknowledge limitations of whatever you are presenting. Limitations are good. They leave job for the people to come. Trying to hide the problems in your work will make the preparation of the talk a lot harder and your self confidence will be hurt.

Let the audience see your personality

- They want to see you enjoy yourself.
- They want to see what you love about the work.
- People really respond to the human parts of a talk. Those parts help the audience with their difficult task of listening to an hour-long talk on a technical subject. What was easy, what was fun, what was hard about the work?
- Don’t be afraid to be yourself and to be quirky.

The different kinds of talks you'll have to give as a researcher

- 2-5 minute talks
- 20 -30 minute conference presentations
- 50-60 minute colloquia / job talk

How to give a talk

Talk organization: here there are as many theories as there are talks. Here there are some extreme advices:

1. Go into details / only big picture
2. Go in depth on a single topic / cover as many things as you can
3. Be serious (never make jokes, maybe only one) / be funny (it is just another form of theater)

Corollary: ask people for advice, but at the end, if will be just you and the audience. Chose what fits best your style.

What everybody agree on is that you have to practice in advance (the less your experience, the more you have to practice). Do it with an audience or without, but practice.

The best advice from Yair Weiss:

"just give a good talk"

Sources on writing technical papers

- How to Get Your SIGGRAPH Paper Rejected, Jim Kajiya, SIGGRAPH 1993 Papers Chair, http://www.siggraph.org/publications/instructions/rejected.html
- Ted Adelson's Informal guidelines for writing a paper, 1991. http://www.ai.mit.edu/courses/6.899/papers/ted.htm
- Notes on technical writing, Don Knuth, 1989.

  http://www.ai.mit.edu/courses/6.899/papers/knuthAll.pdf

- What's wrong with these equations, David Mermin, Physics Today, Oct., 1989. http://www.ai.mit.edu/courses/6.899/papers/mermin.pdf
- Ten Simple Rules for Mathematical Writing, Dimitri P. Bertsekas http://www.mit.edu:8001/people/dimitrib/Ten_Rules.html

Efros' Talk

- http://amps-web.amps.ms.mit.edu/csail/2012-2013/Big_Data/csail-lec-mit-kiva-2012oct24-1600.html

A good abstract

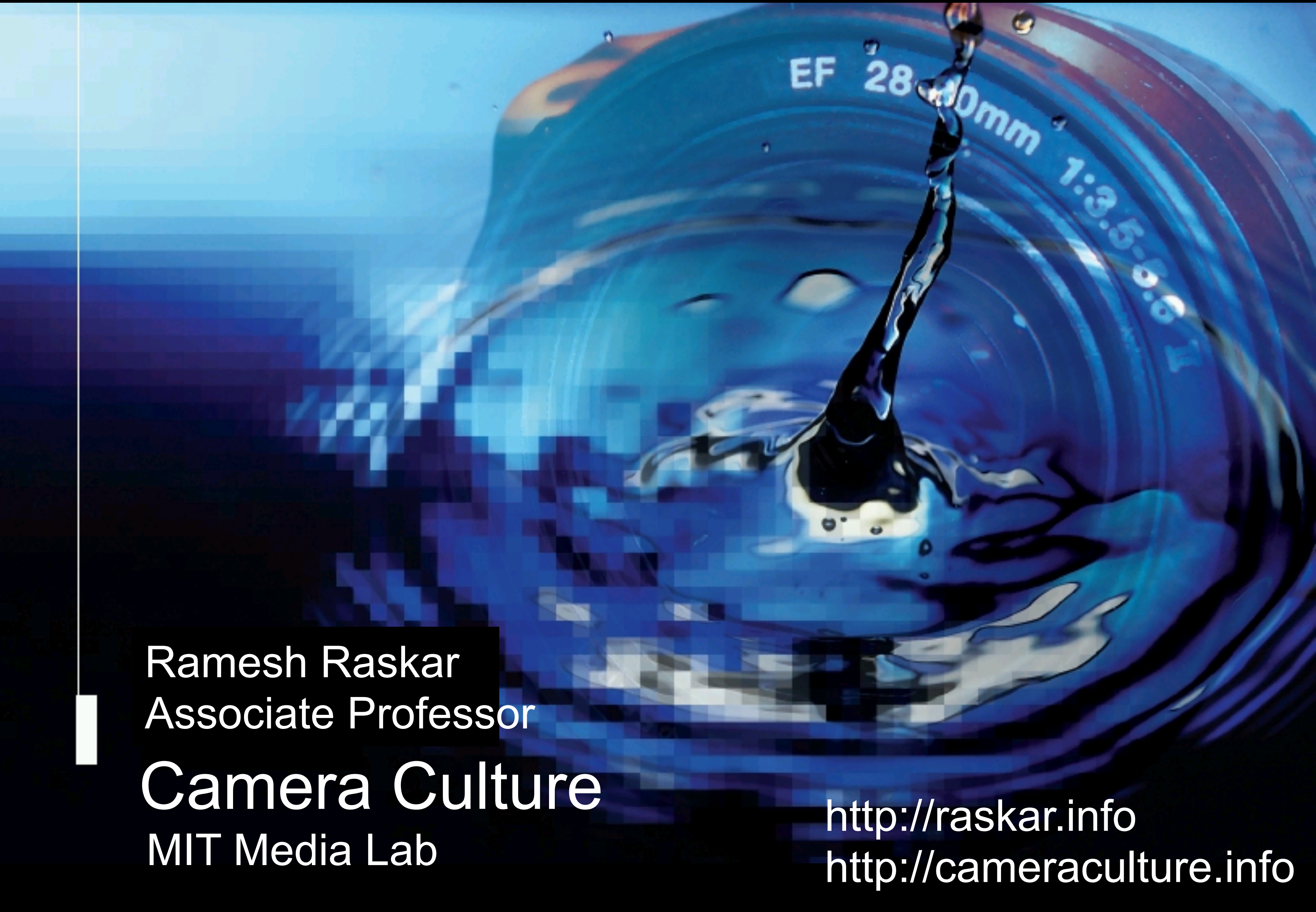
There are an estimated **3.5** trillion photographs in the world, of which **10%** have been taken in the past 12 months. Facebook alone reports 6 billion photo uploads per month. Every minute, **72** hours of video are uploaded to YouTube. Cisco estimates that in the next few years, visual data (photos and video) will account for over **85%** of total internet traffic. Yet, we currently lack effective computational methods for making sense of all this mass of visual data. Unlike easily indexed content, such as text, visual content is not routinely searched or mined; it's not even hyperlinked. Visual data is Internet's "digital dark matter" [Perona,2010] -- **it's just sitting there! (pitch the problem)**

In this talk, I will first discuss some of the unique challenges that make Big Visual Data difficult compared to other types of content. In particular, **I will argue that the (make an point, not just here is the software)** central problem is the lack a good measure of similarity for visual data. I will then present some of our recent work that aims to address this challenge in the context of visual matching, image retrieval and visual data mining. As an application of the latter, we used Google Street View data for an entire city in an attempt to answer that age-old question which has been vexing poets (and poets-turned-geeks): "What makes Paris look like Paris?"

How to Invent?

After X, what is neXt

Ramesh Raskar, MIT Media Lab

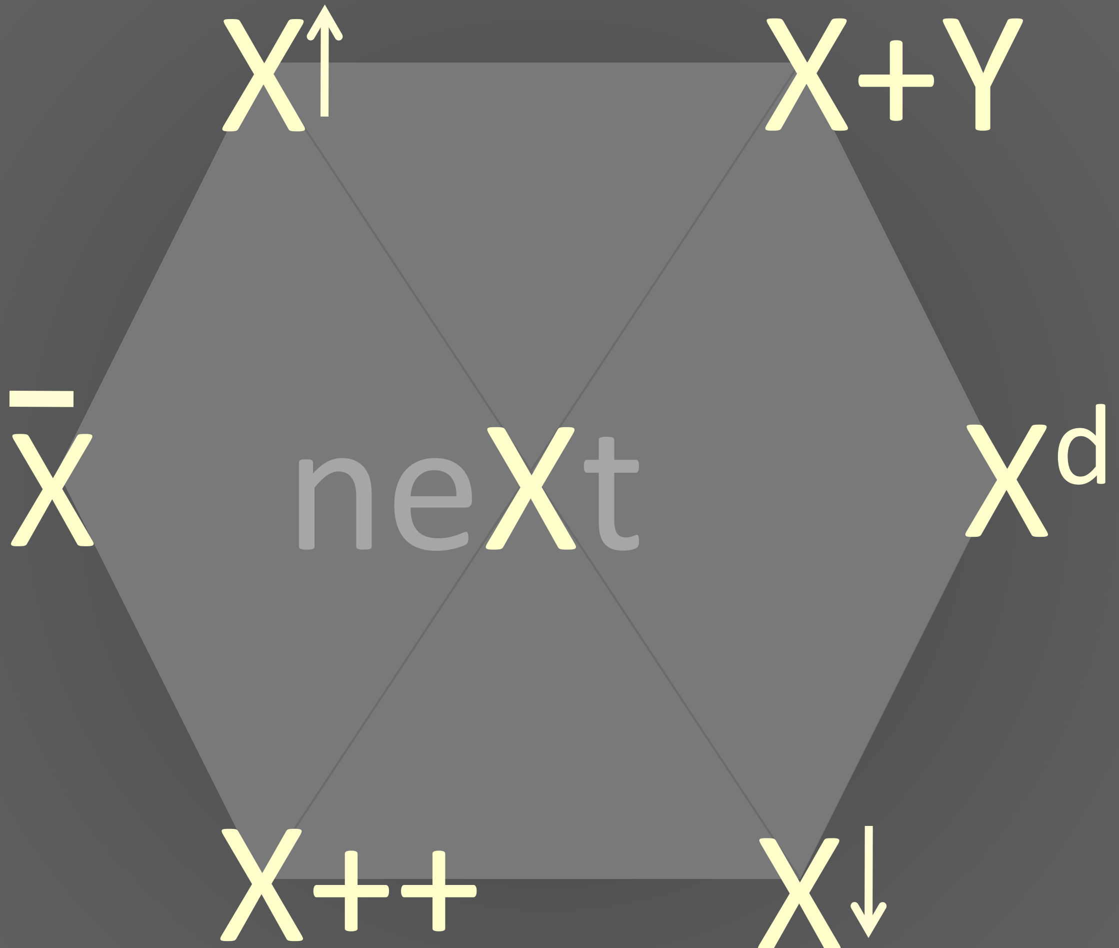


Ramesh Raskar
Associate Professor

Camera Culture

MIT Media Lab

http://raskar.info
http://cameraculture.info


$X\uparrow$
$X+Y$
$\bar{X}$
neXt
X^d
$X{+}{+}$
$X\downarrow$

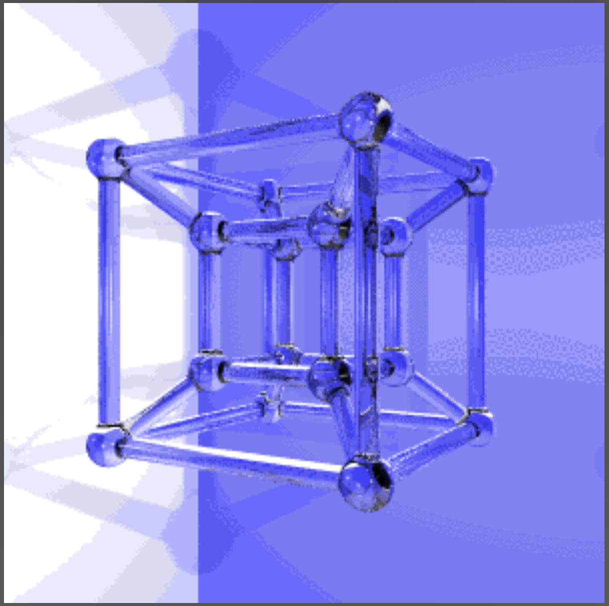
Simple Exercise ..

- Image Compression
  - Save Bandwidth and storage

What is neXt

Strategy #1: X^d

- Extend it to next (or some other) dimension

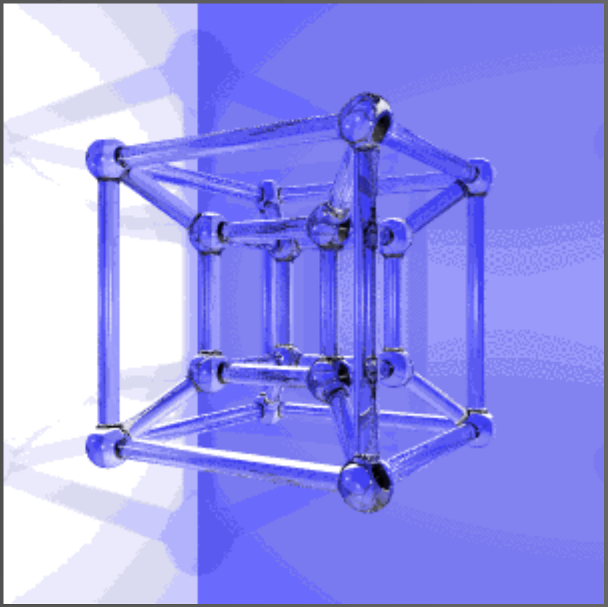


X =

- Idea you just heard
- Concept
- Patent
- New Product
- Product feature
- Design
- Art
- Algorithm

Strategy #1: X^d

- Extend it to next dimension (or some other) dimension
  - Flickr to Youtube
  - Wikipedia to .. ?



- Generalize the concept
  (common in patent applications)

- Text, Audio (Speech), Image, Video .. Whats next ?
  - CD ..
- Images to infrared, sound, ultrasound to EM spectrum
- Macro scale to microscale
  - Airbag for car to airbag for .. ?

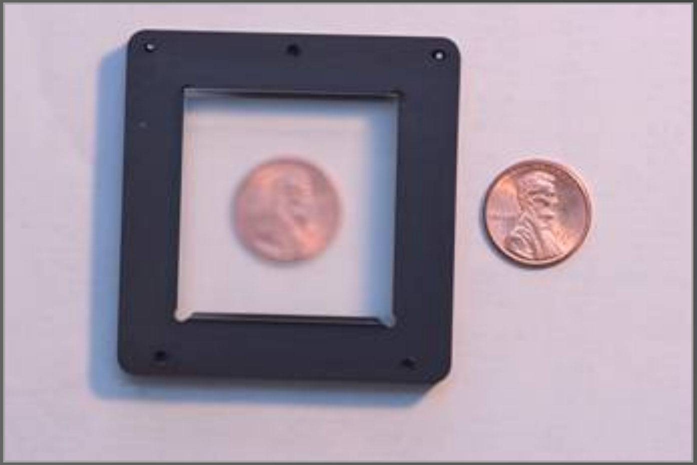
Strategy #2: X+Y

- Fusion of the dissimilar
  - More dissimilar, more spectacular the output
- Example
  - Scientific imaging + Photography
    - Coded aperture
    - Tomography

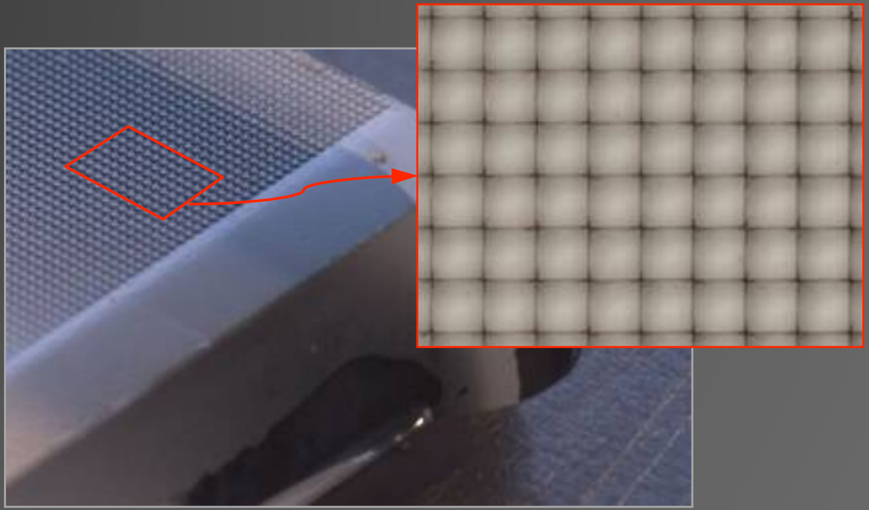
Prototype camera

Contax medium format camera

Kodak 16-megapixel sensor



Adaptive Optics microlens array



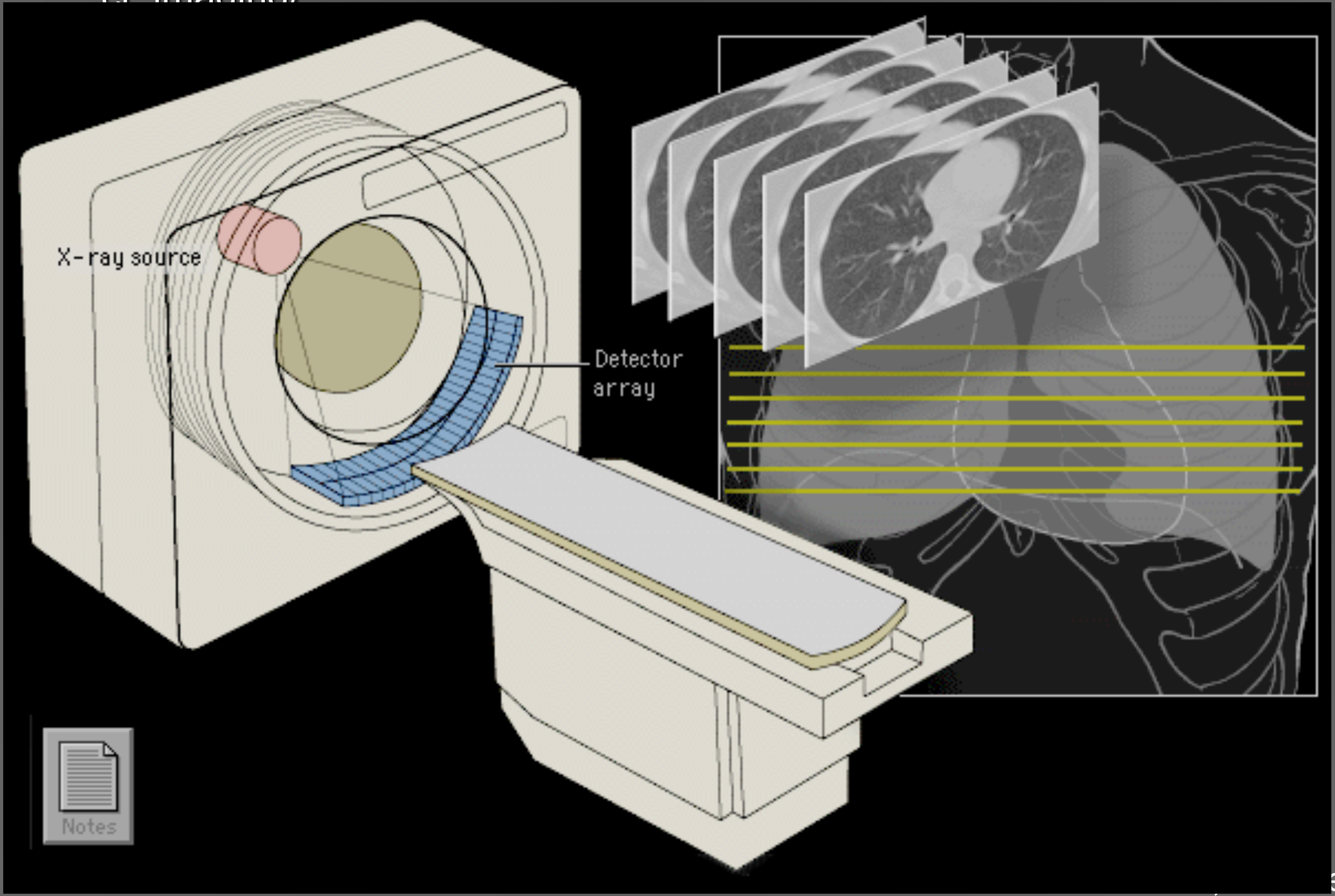
125µ square-sided microlenses

4000 × 4000 pixels ÷ 292 × 292 lenses = 14 × 14

Example of digital refocusing

Imaging in Sciences: Computer Tomography

- http://info.med.yale.edu/intmed/cardio/imaging/techniques/ct_imaging/



Self Evaluation of Eye

Your result is
Spherical: -2.00
Cylindrical: -1.00
Axis: 18.45
spherical
cylindrical

Health Screening Tools

Blood Oxygenation

Blood Pressure

Visual Acuity (our device)

Blood Glucose

Body Temperature

Strategy #3: X
Do exactly the opposite

Strategy #3: $\bar{X}$
Do exactly the opposite

- Processing, Memory, Bandwidth
  - In Computing world, in any era, one of this is a bottleneck
  - But overtime, they change. You can often take an older idea and do exactly the opposite.
  - E.g. bandwidth is now considered virtually limitless
- Business Process Reengineering (BPR)
  - Michael Hammer, James Champy, 1990s
- In imaging:
  - SLR: Faster mirror flip or no mirror flip
    - Companies spent years improving mirror flip speed
    - Why not just remove it?
- More computation
- Less light

Fosbury Flip

Bundesarchiv, Bild 183-S0305-0030
Foto: Mittelstädt, Rainer | 5. März 1977

http://en.wikipedia.org/wiki/File:Bundesarchiv_Bild_183-S0305-0030,_Rolf_Beilschmidt.jpg

Straddle Method for High Jump

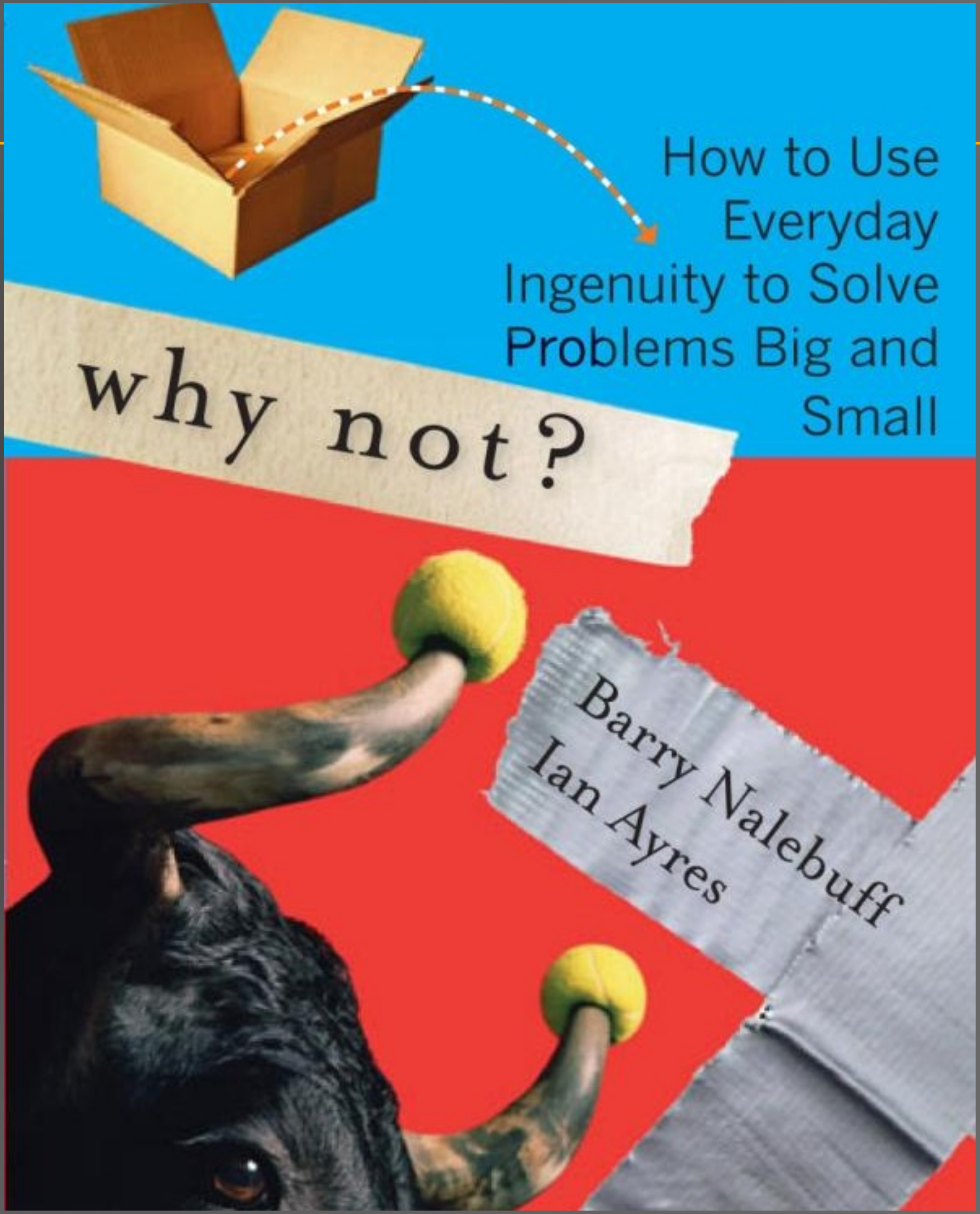
Replacement of landing surfaces with foam rubber

Fosbury Method

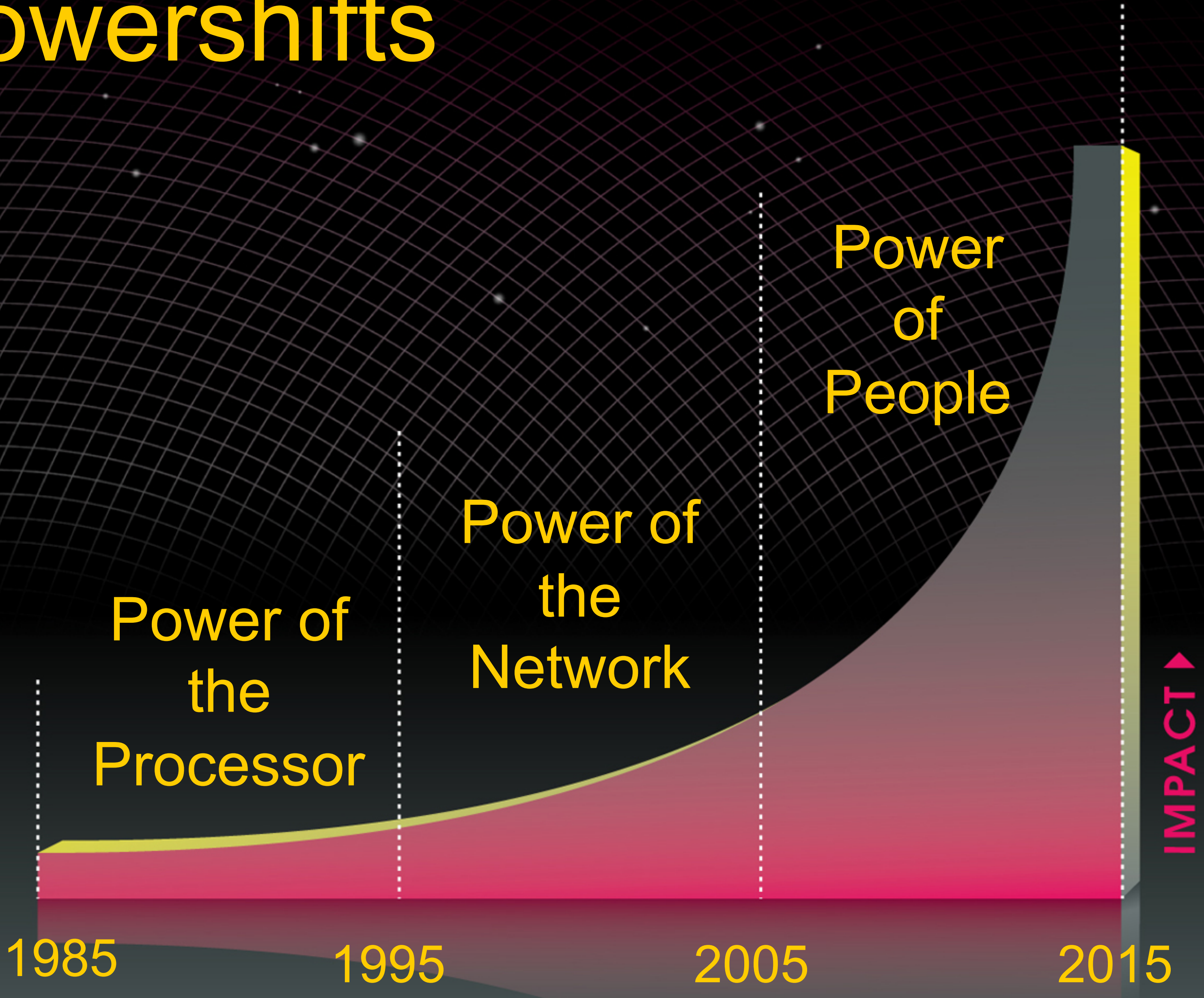
1968 Olympics: 2.24m

http://en.wikipedia.org/wiki/Dick_Fosbury

- Toll Free calls
- Reverse Auction

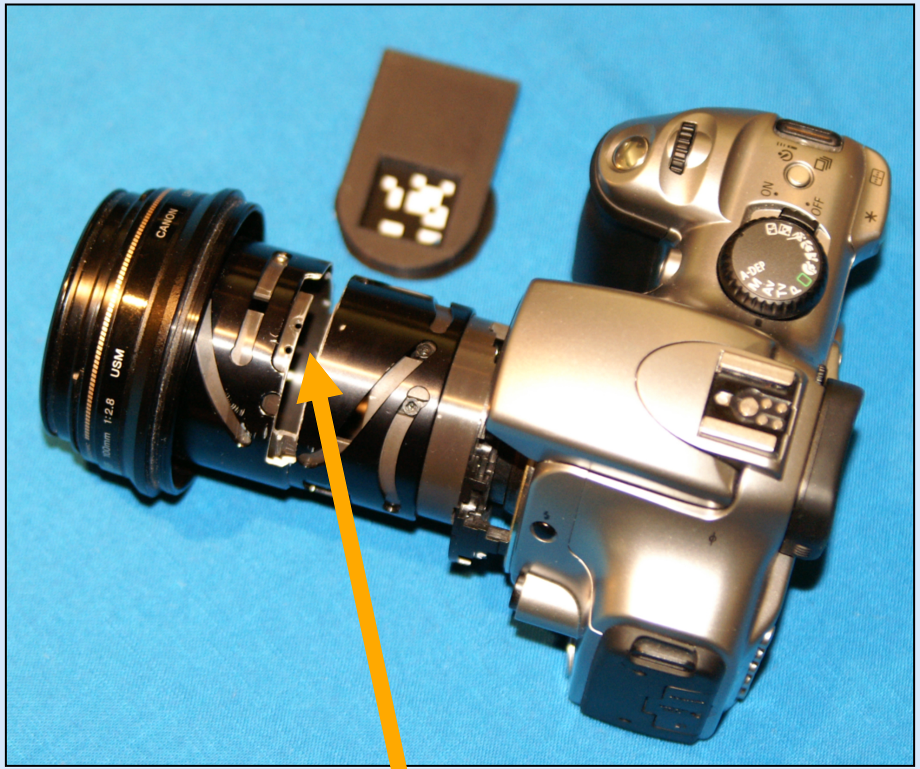

How to Use Everyday Ingenuity to Solve Problems Big and Small
why not?
Barry Nalebuff
Ian Ayres

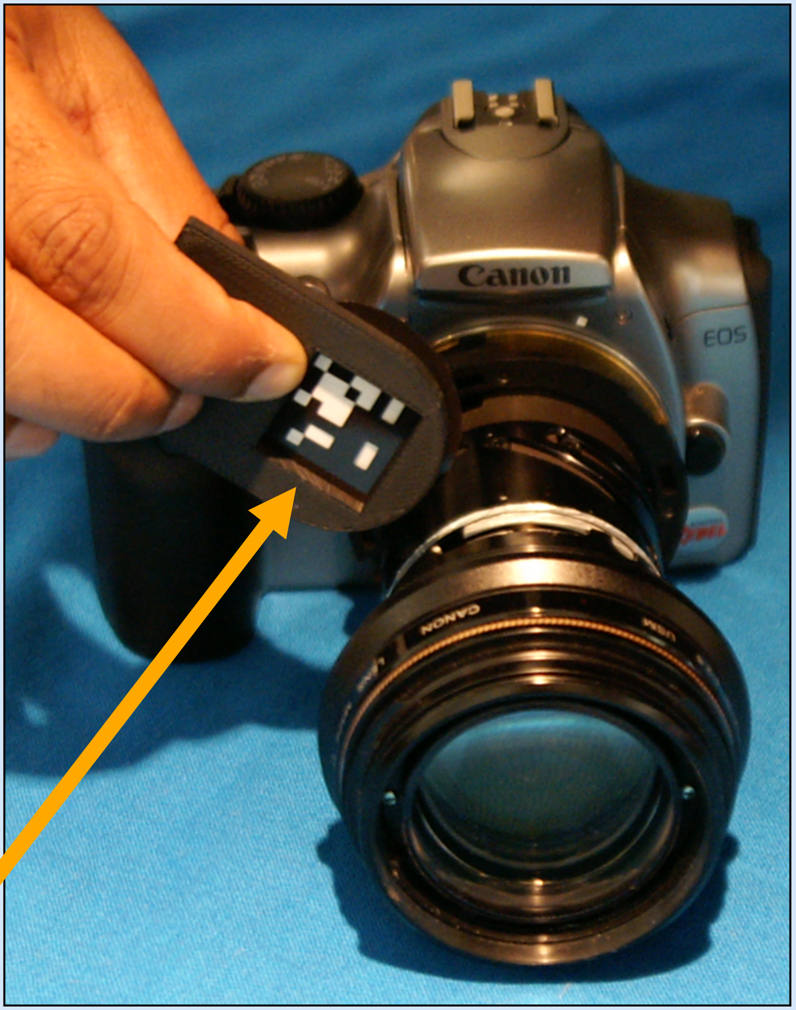
Powershifts



Frank Moss, MIT Media Lab

Coded Aperture Camera



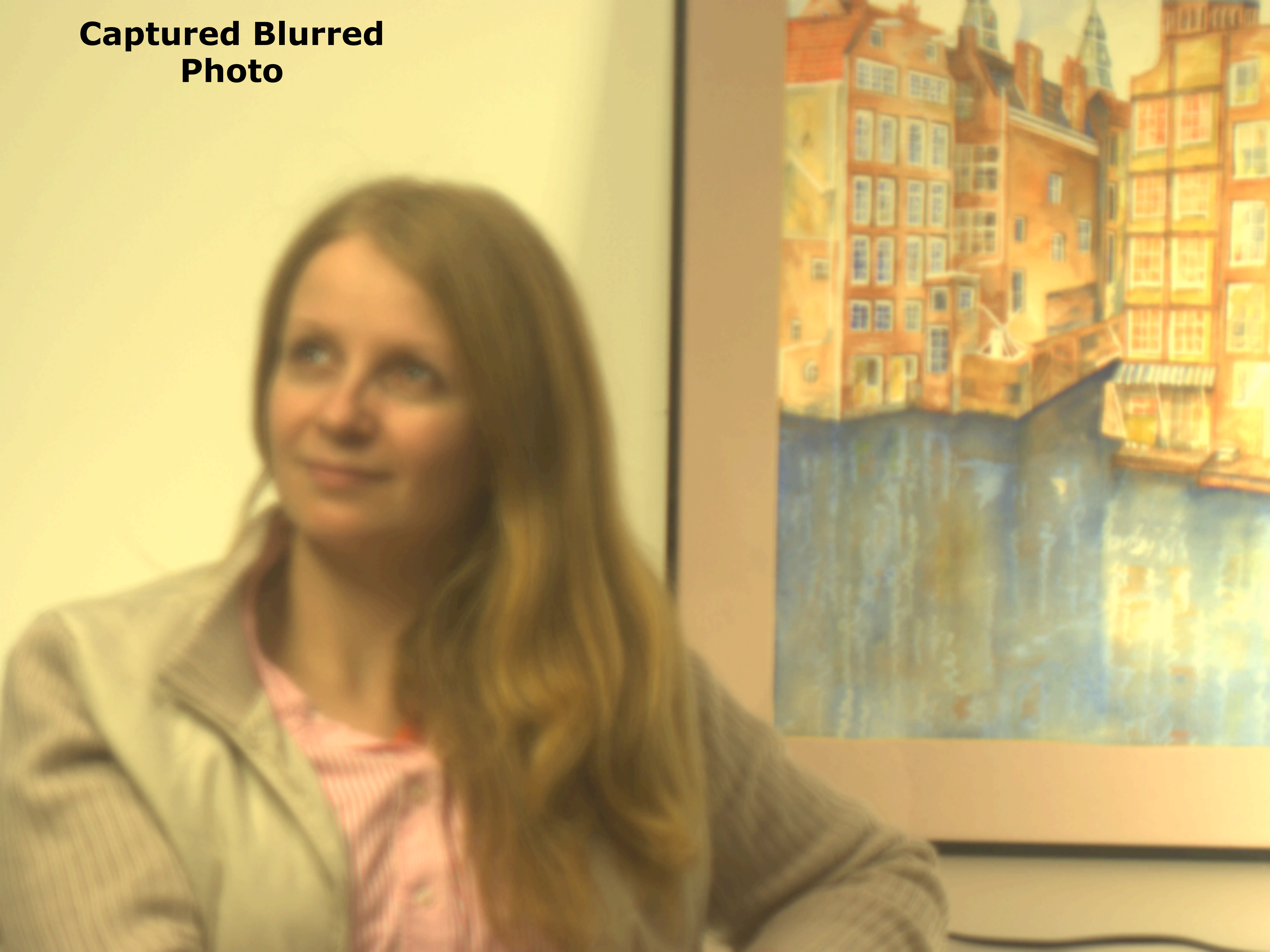


The aperture of a 100 mm lens is modified

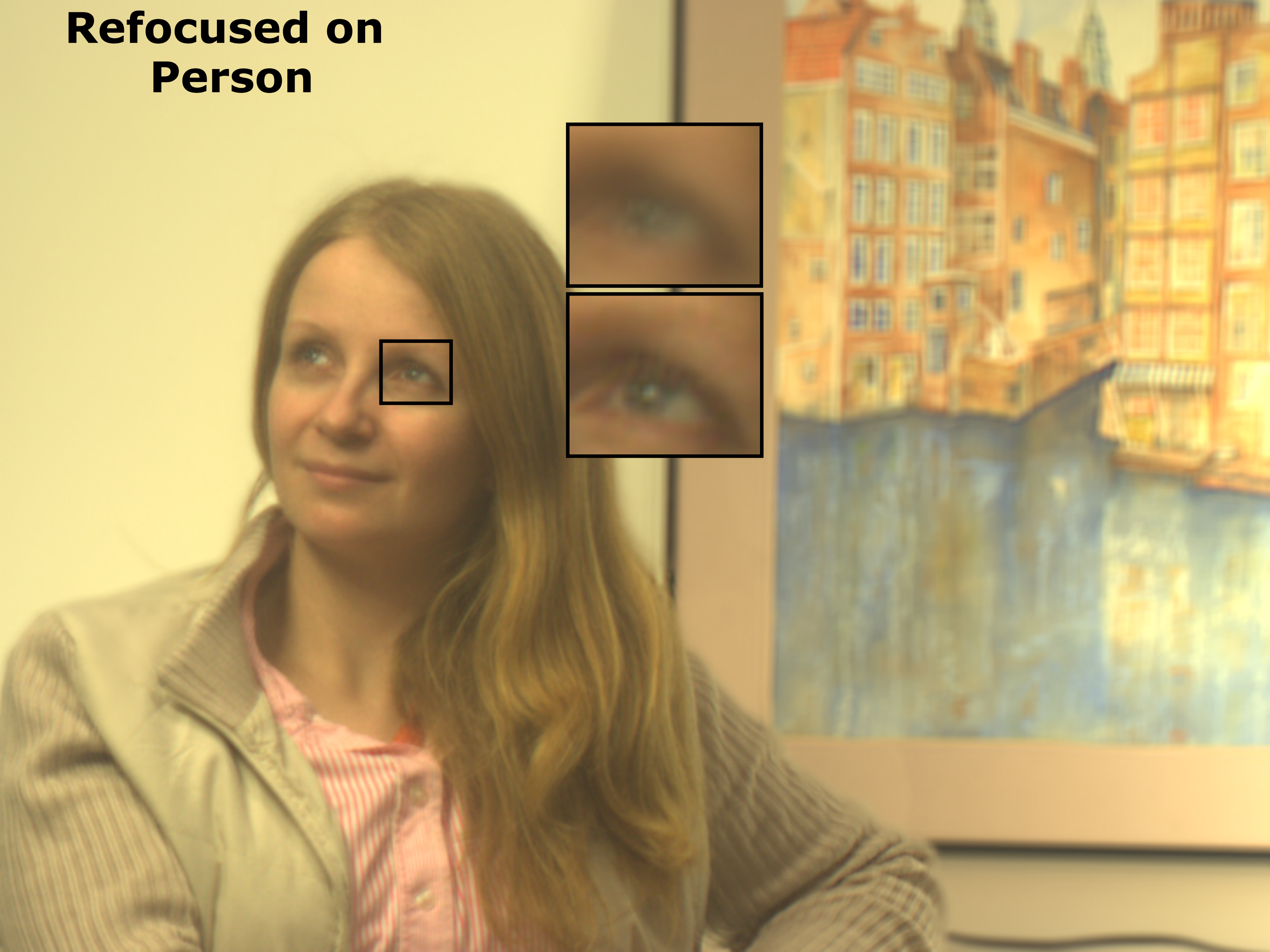
Insert a coded mask with chosen binary pattern

Rest of the camera is unmodified

Captured Blurred Photo



Refocused on Person

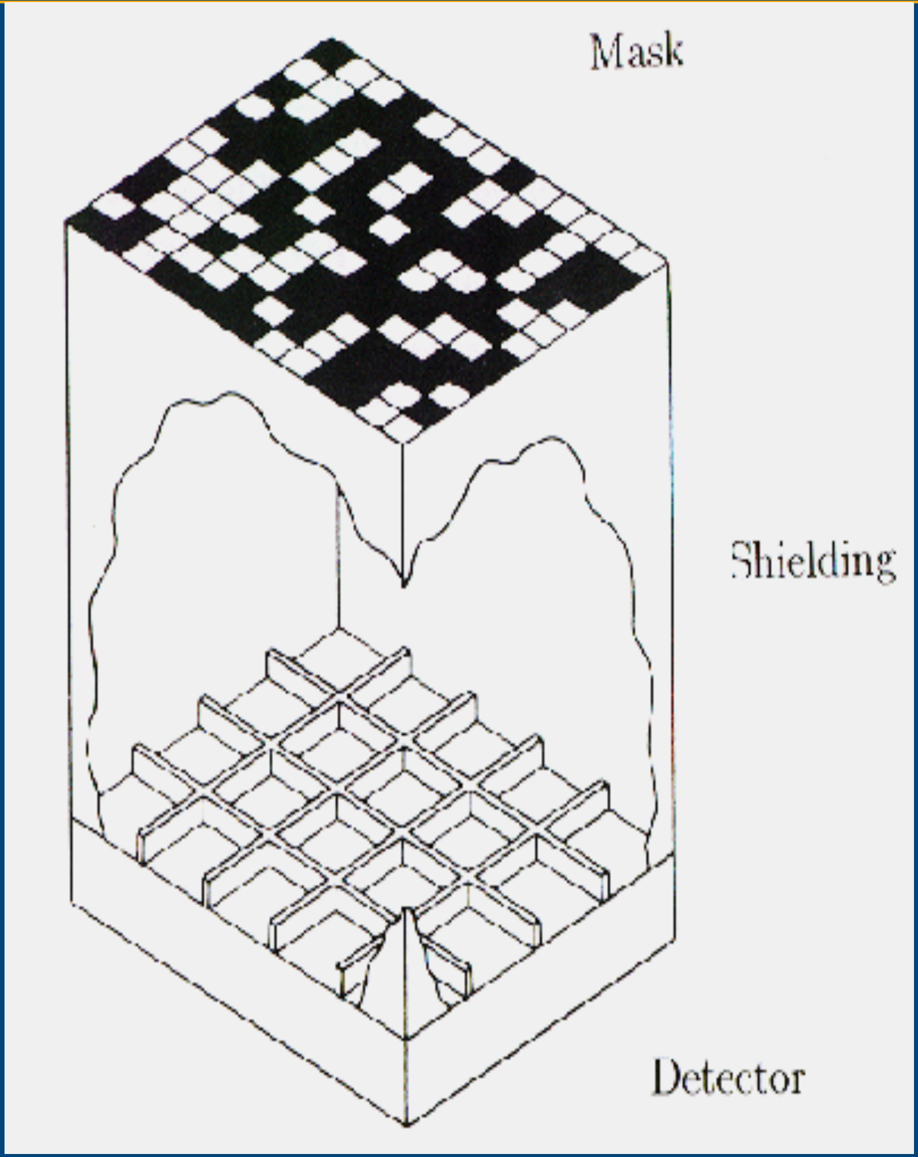


Coded-Aperture Imaging

- Lens-free imaging!
- Pinhole-camera sharpness, without massive light loss.
- No ray bending (OK for X-ray, gamma ray, etc.)

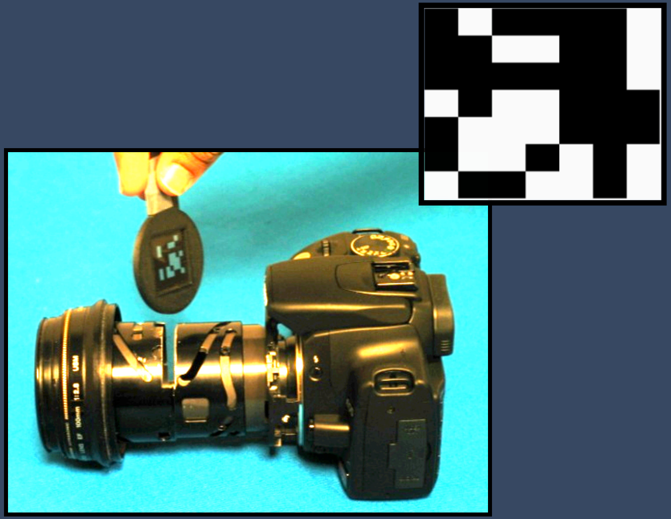
- Two elements
  - Code Mask: binary (opaque/transparent)
  - Sensor grid

- Mask autocorrelation is delta function (impulse)
- Similar to MotionSensor ?



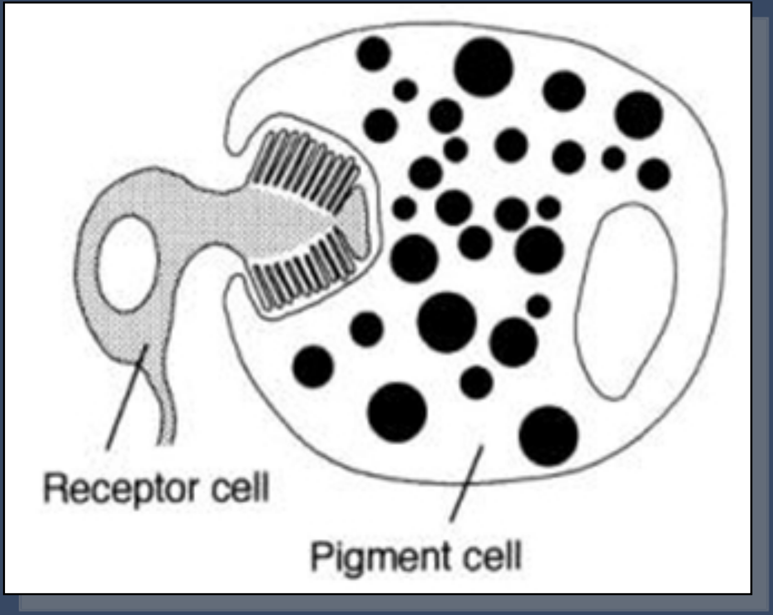
Less is More

Blocking Light == More Information

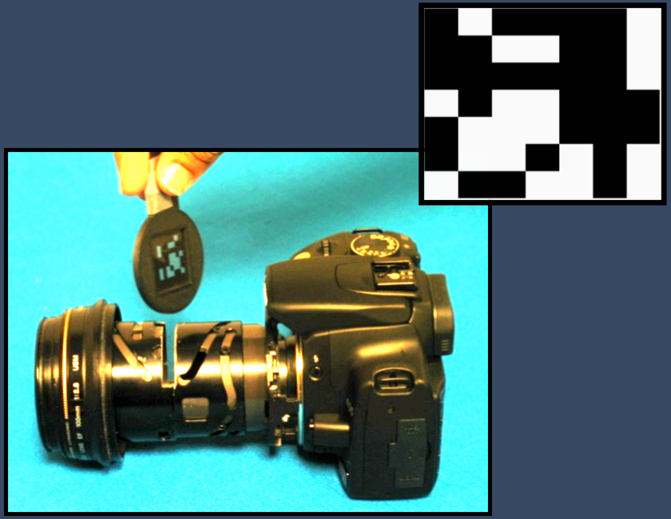


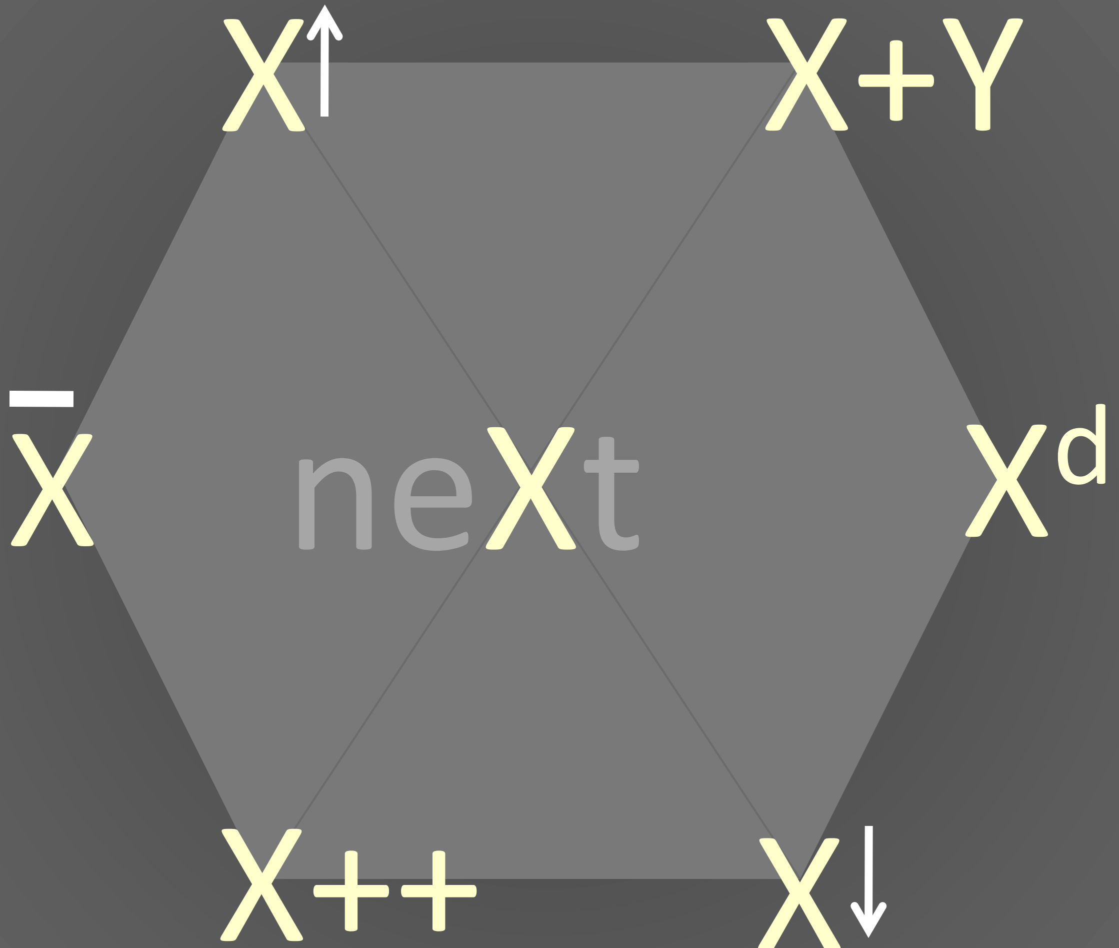
Coding in Time

Coding in Space



Larval Trematode Worm




$X\uparrow$
$X+Y$
$\bar{X}$
neXt
X^d
$X{+}{+}$
$X\downarrow$

Strategy #4: X↑

- Given a Hammer ..
  - Find all the nails
  - Sometimes even screws and bolts

Strategy #4: X↑

- Given a Hammer ..
  - Find all the nails
  - Sometimes even screws and bolts
- Given a cool solution/technique/Opportunity
  - Find other problems
  - (Where to find them?)
- Examples
  - Peltier effect:
    - Create a jacket that keeps you warm or cold
  - Mobile phone opportunity

Strategy #5: X↓

- Given a nail,
  - Find all hammers
  - Sometimes even screwdrivers and pliers may work
- Given a problem,
  - Find other solutions
  - (Where to find them?)
- Examples
  - App store (Apple) .. Open platform for all devices
  - ..

Strategy #5: X↓

- Given a problem, find other solutions

- High Dynamic Range Tone Mapping
  - Started with Jack Tumblin's LCIS
  - Gradient domain
  - Bilateral filter
  - Filter banks etc ..
  - About 6 years of heavy machinery
  - Btw, the topic is done to death but continues to enthuse

Strategy #6: X++

- Pick your adjective ..
- Making it faster, better, cheaper

neXt = adjective + X

X++ : Add your favorite adjective

- Context aware,
- Adaptive
- (temporally) Coherent,
- Hierarchical,
- Progressive
- Efficient
- Parallelized
- Distributed
- Good example: Image or video compression schemes

- Personalized/Customized
- Democratized

X++ : Add your favorite adjective

- Good example: Image or video compression schemes
- But X++ is a sign
  - The field is maturing in terms of research but booming in business impact

- Kaizen
  - Small incremental changes
  - Japanese Management styles (6sigma, Kanban)
  - Mainly to save money/time/resources. Not everyone can do it. GM, 0.84 suggestions per employee vs Toyota 18. GM accepted 23%, Toyota 90%

Hexagon Corners for Different Stages

[Bruce Tuckman, 1965]

- Forming
- Storming
- Norming
- Performing
- Adjourning

Other Techniques

- Ideation
  - Including idea mgmnt software
  - Mind-mapping
- Brainstorming
  - Randomization
  - Follow-on triggers
- Problem Solving
  - (Have I done everything)
  - How to solve it (1945)
    - Mathematician George Pólya
    - Convert to a known problem

‘Reducing’ a problem
Analogy
Generalization
Induction
Variation of the Problem
Auxiliary Problem
Pattern Matching (related problem solved before)
Specialization
Decomposing and Recombining
Working backward
Draw a Figure
Auxiliary Elements

Other Techniques

- Problem Solving

- TRIZ 40
  - Mechanical engineering problems, e.g. how to increase volume w/o extra weight
  - Overcome constraints by using transform of existing solutions
  - TRIZ matrix: database of known solutions to overcome constraints
- Advanced Systematic Inventive Thinking (ASIT)
  - Unification: Multiplication: Division: Breaking Symmetry: Object Removal
  - How to make incremental changes to solve an engineering problem
-

http://www.triz40.com/

Where to find the ‘X’

- Annual Awards (best product, researchers)

- Talks abstract (no need to attend)
- Network and talk to people
- Avoid small-talk .. Ask ‘what is the latest X’

- Patents
- Table of Contents
- Index pages

Pitfalls

- These six ways are only a start
- They are a good mental exercise and will allow you to train as a researcher
- Great for projects
- But
  - Maynot produce radically new ideas
  - Sometimes a danger of being labeled incremental
  - Could be into ‘public domain ideas’

What are Bad ideas to pursue

- X then Y (then Z)
  - X+Y is great with true fusion, fusion of dissimilar is best
  - But avoid a 'pipeline' systems, where the output of one is THEN channeled into the input of the next stage, and non of the components are novel (idea is easy to scoop)

http://www.flickr.com/photos/rsilfver/178134761/

What are Bad ideas to pursue

- X then Y (then Z)
  - X+Y is great with true fusion, fusion of dissimilar is best
  - But avoid a ‘pipeline’ systems, where the output of one is THEN channeled into the input of the next stage, and non of the components are novel (idea is easy to scoop)
- Follow the hype (too much competition)
- Do because it can be done
  - (Why do we climb a mountain? because it is there!)
  - But only the first one gets a credit.
  - May make you strong, and give you a sense of achievement but not a research project.

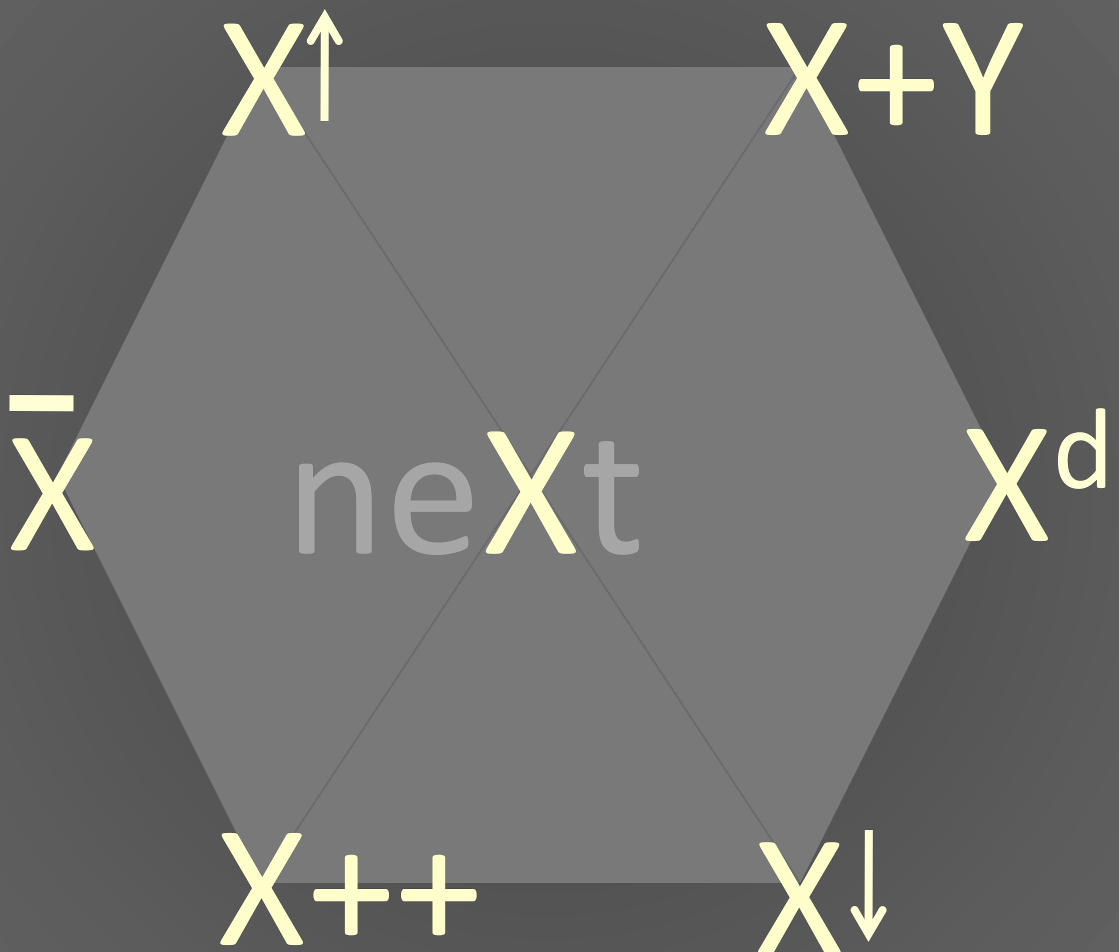
There is more ..

- How to decide if the idea is worth pursuing
  - My personal triangle of criteria
  - Intersection of interests, skills, demand
  - Maybe another talk ..

Acknowledgements

- Members of Camera Culture, MIT group

- Vitor Pamplona
- Kari Pulli, Nokia
- Asmita Joshi
- Rupesh Nasre, IISc
- Mark Bolas, USC
- Rajiv Narayan, Broad Institute, MIT
- Joost Bonsen, MIT


$X^{\uparrow}$
$X+Y$
$\bar{X}$
neXt
X^{d}
$X{+}{+}$
$X^{\downarrow}$