



Lecture 22

How to write papers and give talks.



6.869/6.819 Advances in Computer Vision

Bill Freeman and Phillip Isola

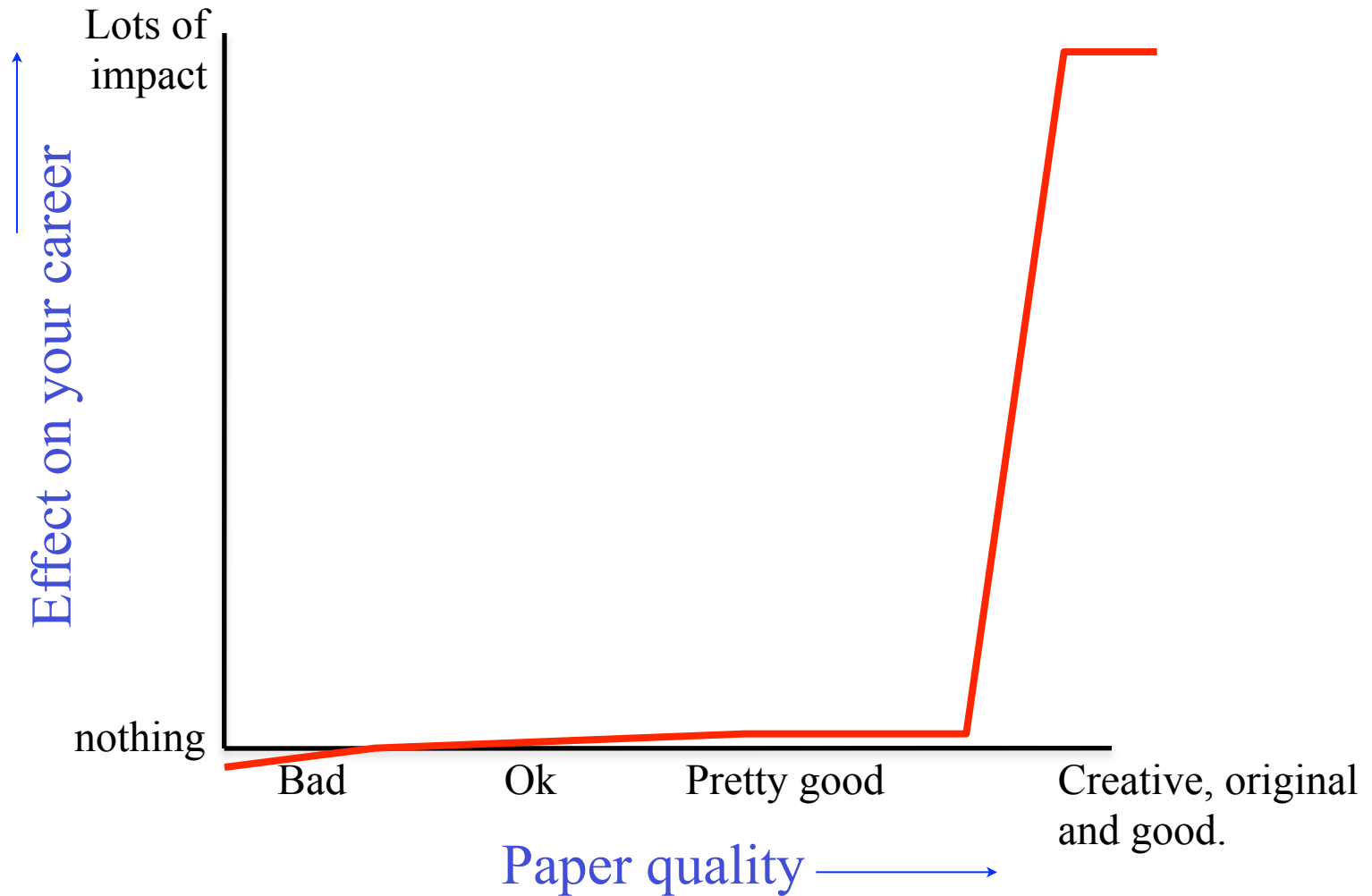
Spring 2021

May 10, 2021

Outline

- writing technical papers
- giving technical talks

A paper's impact on your career



Our image of the research community

- Scholars, plenty of time on their hands, pouring over your manuscript.



The reality:
more like a large, crowded marketplace



<http://ducksflytogether.wordpress.com/2008/08/02/looking-back-khan-el-khalili/>

Ted Adelson on how to write a good paper

- (1) Start by stating which problem you are addressing, keeping the audience in mind. They must care about it, which means that sometimes you must tell them why they should care about the problem.
- (2) Then state briefly what the other solutions are to the problem, and why they aren't satisfactory. If they were satisfactory, you wouldn't need to do the work.
- (3) Then explain your own solution, compare it with other solutions, and say why it's better.
- (4) At the end, talk about related work where similar techniques and experiments have been used, but applied to a different problem.

Since I developed this formula, it seems that all the papers I've written have been accepted. (told informally, in conversation, 1990).

Example paper organization: removing camera shake from a single photograph

1 Introduction

2 Related work

3 Image model

4 Algorithm

Estimating the blur kernel

Multi-scale approach

User supervision

Image reconstruction

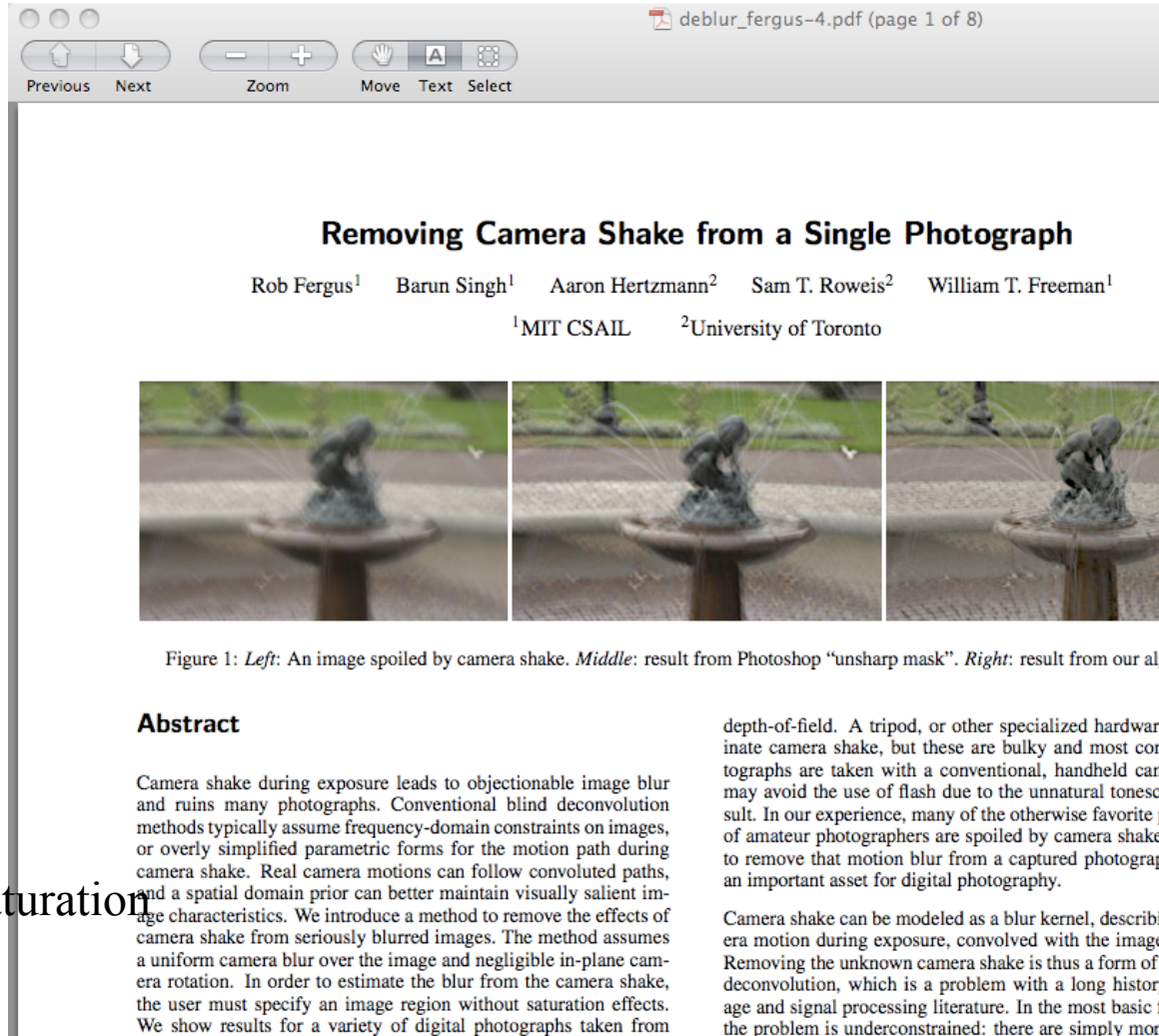
5 Experiments

Small blurs

Large blurs

Images with significant saturation

6 Discussion



The screenshot shows a PDF viewer window titled 'deblur_fergus-4.pdf (page 1 of 8)'. The toolbar includes 'Previous', 'Next', 'Zoom', 'Move', 'Text', and 'Select' buttons. The main content is the title page of a paper:

Removing Camera Shake from a Single Photograph

Rob Fergus¹ Barun Singh¹ Aaron Hertzmann² Sam T. Roweis² William T. Freeman¹

¹MIT CSAIL ²University of Toronto




Figure 1: *Left*: An image spoiled by camera shake. *Middle*: result from Photoshop “unsharp mask”. *Right*: result from our algorithm.

Abstract

Camera shake during exposure leads to objectionable image blur and ruins many photographs. Conventional blind deconvolution methods typically assume frequency-domain constraints on images, or overly simplified parametric forms for the motion path during camera shake. Real camera motions can follow convoluted paths, and a spatial domain prior can better maintain visually salient image characteristics. We introduce a method to remove the effects of camera shake from seriously blurred images. The method assumes a uniform camera blur over the image and negligible in-plane camera rotation. In order to estimate the blur from the camera shake, the user must specify an image region without saturation effects. We show results for a variety of digital photographs taken from

depth-of-field. A tripod, or other specialized hardware, can avoid camera shake, but these are bulky and most commercial photographs are taken with a conventional, handheld camera. One may avoid the use of flash due to the unnatural tones of the result. In our experience, many of the otherwise favorite photographs of amateur photographers are spoiled by camera shake. We introduce a method to remove that motion blur from a captured photograph, an important asset for digital photography.

Camera shake can be modeled as a blur kernel, describing camera motion during exposure, convolved with the image. Removing the unknown camera shake is thus a form of blind deconvolution, which is a problem with a long history in image and signal processing literature. In the most basic case, the problem is underconstrained: there are simply more

The introduction

1 Introduction

2 Related work

3 --Main idea--

4 Algorithm

- Estimating the blur kernel

 - Multi-scale approach

 - User supervision

- Image reconstruction

5 Experiments

- Small blurs

- Large blurs

- Images with significant saturation

6 Discussion

Jim Kajiya: write a dynamite introduction

You must make your paper easy to read. You've got to make it easy for anyone to tell what your paper is about, what problem it solves, why the problem is interesting, what is really new in your paper (and what isn't), why it's so neat. And you must do it up front. In other words, you must write a dynamite introduction.

Underutilized technique: explain the main idea with a simple, toy example.

1 Introduction

2 Related work

3 Main idea

4 Algorithm

Estimating the blur kernel

Multi-scale approach

User supervision

Image reconstruction

5 Experiments

Small blurs

Large blurs

Images with significant saturation

6 Discussion

← Often useful here.

Show simple toy examples to let people get the main idea

From
“Shiftable
multiscale
transforms”

588

IEEE TI

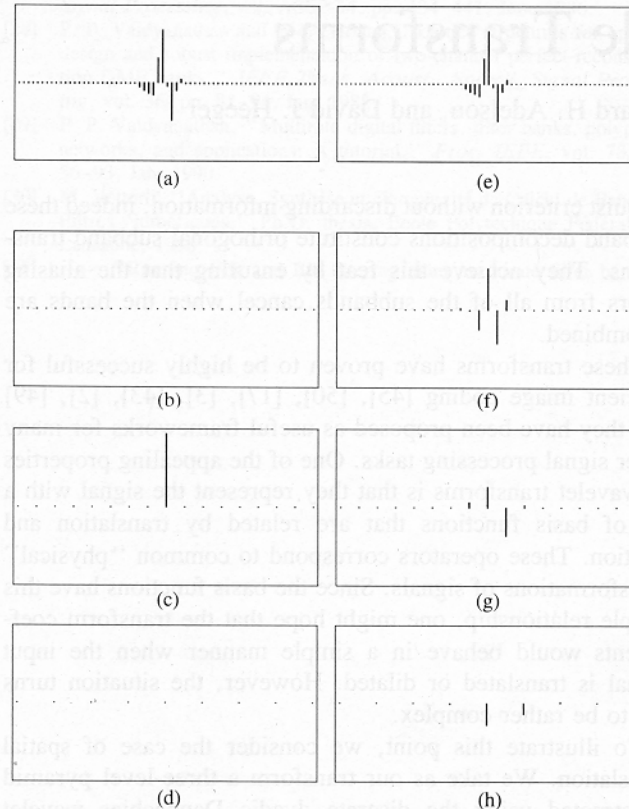


Fig. 1. Effect of translation on the wavelet representation of a signal. (a) Input signal, which is equal to one of the wavelet basis functions. (b)–(d) Decomposition of the signal into three wavelet subbands. Plotted are the coefficients of each subband. Dots correspond to zero-value coefficients. (e) Same input signal, translated one sample to the right. (f)–(h) Decomposition of the shifted signal into three wavelet subbands. Note the drastic change in the coefficients of the transform, both within and between subbands.

Experimental results are critical now at CVPR

1 Introduction

2 Related work

3 Image model

4 Algorithm

Estimating the blur kernel

Multi-scale approach

User supervision

Image reconstruction

5 Experiments

Small blurs

Large blurs

Images with significant saturation

6 Discussion

Methods	Dataset	two-view?	si-full	si-env	si-hum	si-intra	si-inter	RMSE	Rel
Russell <i>et al.</i> [31]	-	Yes	2.146	2.021	2.207	2.206	2.093	2.520	0.772
DeMoN [39]	RGBD+MVS	Yes	0.338	0.302	0.360	0.293	0.384	0.866	0.220
Chen <i>et al.</i> [3]	NYU+DIW	No	0.441	0.398	0.458	0.408	0.470	1.004	0.262
Laina <i>et al.</i> [17]	NYU	No	0.358	0.356	0.349	0.270	0.377	0.947	0.223
Xu <i>et al.</i> [46]	NYU	No	0.427	0.419	0.411	0.302	0.451	1.085	0.274
Fu <i>et al.</i> [7]	NYU	No	0.351	0.357	0.334	0.257	0.360	0.925	0.194
<i>I</i>	MC	No	0.318	0.334	0.294	0.227	0.319	0.840	0.204
<i>IFCM</i>	MC	Yes	0.316	0.330	0.302	0.228	0.323	0.843	0.206
<i>ID_{pp}M</i>	MC	Yes	0.246	0.225	0.260	0.233	0.273	0.635	0.136
<i>ID_{pp}CM</i> (w/o d. cleaning)	MC	Yes	0.272	0.238	0.293	0.258	0.282	0.688	0.147
<i>ID_{pp}CM</i>	MC	Yes	0.232	0.203	0.252	0.224	0.262	0.570	0.129
<i>ID_{pp}CMK</i>	MC	Yes	0.221	0.195	0.238	0.215	0.247	0.541	0.125

Table 2. Results on TUM RGBD datasets. Different si-RMSE metrics as well as standard RMSE and relative error (Rel) are reported. We evaluate our models (light gray background) under different input configurations, as described in Table 1. *w/o d. cleaning* indicates the model is trained using raw MVS depth predictions as supervision, without our depth cleaning method. Dataset '-' indicates the method is not learning based. Lower is better for all error metrics.

Gone are the days of, “We think this is a great idea and we expect it will be very useful in computer vision. See how it works on this meaningless, contrived problem?”

How to end a paper

1 Introduction

2 Related work

3 Image model

4 Algorithm

- Estimating the blur kernel

 - Multi-scale approach

 - User supervision

- Image reconstruction

5 Experiments

- Small blurs

- Large blurs

- Images with significant saturation

6 Discussion

Conclusions, or what this opens up, or how this can change how we approach computer vision problems.

How not to end a paper

- 1 Introduction
 - 2 Related work
 - 3 Image model
 - 4 Algorithm
 - Estimating the blur kernel
 - Multi-scale approach
 - User supervision
 - Image reconstruction
 - 5 Experiments
 - Small blurs
 - Large blurs
 - Images with saturation
 - 6 Discussion
- Future work?

I can't stand "future work" sections.
It's hard to think of a weaker way
to end a paper.

"Here's a list all the ideas we wanted to do but
couldn't get to work in time for the conference
submission deadline. We didn't do any of the
following things: (1)..."

(You get no "partial credit" from reviewers and readers
for neat things you wanted to do, but didn't.)

"Here's a list of good ideas that you should now go
and do before we get a chance."

Better to end with a conclusion or a summary, or you can
say in general terms where the work may lead.

General writing tips

Knuth: keep the reader upper-most
in your mind.

Perhaps the most important principle of good writing is to keep the reader uppermost in mind: What does the reader know so far? What does the reader expect next and why?

Treat the reader as you would a guest in your house

Anticipate their needs: would you like something to drink?
Something to eat? Perhaps now, after eating, you'd like to rest?



Writing style, from the elements of style, Stunk and White

13. Omit needless words.

Vigorous writing is concise. A sentence should contain no unnecessary words, a paragraph no unnecessary sentences, for the same reason that a drawing should have no unnecessary lines and a machine no unnecessary parts. This requires not that the writer make all his sentences short, or that he avoid all detail and treat his subjects only in outline, but that every word tell.

Many expressions in common use violate this principle:

the question as to whether	whether (the question whether)
there is no doubt but that	no doubt (doubtless)
used for fuel purposes	used for fuel
he is a man who	he
in a hasty manner	hastily
this is a subject which	this subject
His story is a strange one.	His story is strange.

Re-writing exercise

The underlying assumption of this work is that the estimate of a given node will only depend on nodes within a patch: this is a locality assumption imposed at the patch-level. This assumption can be justified in case of skin images since a pixel in one corner of the image is likely to have small effect on a different pixel far away from itself. Therefore, we can crop the image into smaller windows, as shown in Figure 5, and compute the inverse J matrix of the cropped window. Since the cropped window is much smaller than the input image, the inversion of J matrix is computationally cheaper. Since we are inferring on blocks of image patches (i.e. ignoring pixels outside of the cropped window), the interpolated image will have blocky artifacts. Therefore, only part of xMAP is used to interpolate the image, as shown in Figure 5.

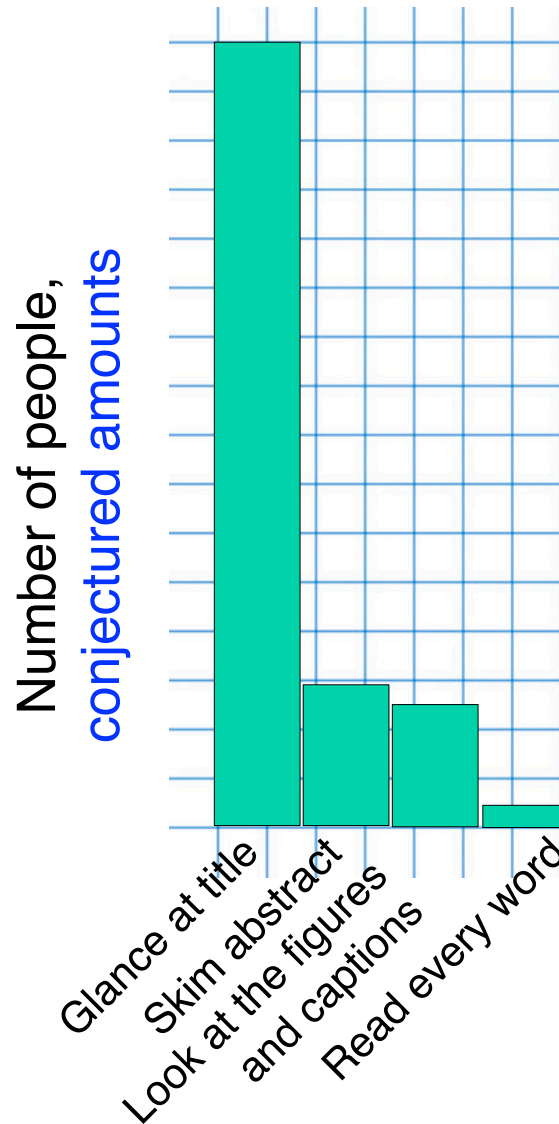
Before

We assume local influence--that nodes only depend on other nodes within a patch. This condition often holds for skin images, which have few long edges or structures. We crop the image into small windows, as shown in Fig. 5, and compute the inverse J matrix of each small window. This is much faster than computing the inverse J matrix for the input image. To avoid artifacts from the block processing, only the center region of xMAP is used in the final image, as shown in Fig. 5.

After

This editing benefits you twice: (1) you have 50% more space to tell your story, and (2) the text is easier for the reader to understand.

The readership of your paper



The “read every word” readers are your most important ones. But you should make the paper “work” for all the other readers, too.

Title?



IEEE TRANSACTIONS ON

INFORMATION THEORY

MARCH 1992 VOLUME 38 NUMBER 2 IETTAW (ISSN 0018-9448)

A Journal Devoted to the Theoretical and Experimental Aspects of Information Transmission, Processing and Utilization
PART II OF TWO PARTS

SPECIAL ISSUE ON WAVELET TRANSFORMS AND MULTIREOLUTION SIGNAL ANALYSIS

<i>I. Daubechies, S. Mallat, and A. S. Willky</i>	Introduction to the Special Issue	529
PAPERS		
Theory and Implementation of Wavelet and Multidimensional Transforms		
<i>J. Kocačević and M. Vetterli</i>	Nonseparable Multidimensional Perfect Reconstruction Banks and Wavelet for \mathbb{R}^n	533
<i>K. Gröchenig and W. R. Madsych</i>	Multiresolution Analysis, Haar Bases, and Self-Similar Tilings of \mathbb{R}^n	556
<i>O. Rioul and P. Duhamel</i>	Fast Algorithms for Discrete and Continuous Wavelet Transforms	569
<i>E. P. Simoncelli, W. T. Freeman, E. H. Adelson and D. J. Heeger</i>	Shiftable Multiscale Transforms	587
<i>N. J. Munch</i>	Noise Reduction in Tight Weyl-Heisenberg Frames	608
Time-Frequency and Event Localization		
<i>S. Mallat and W. L. Hwang</i>	Singularity Detection and Processing with Wavelets	617
<i>N. Delprat, B. Escudé, P. Guillemain, R. Kronland-Martinet, Ph. Tchamitchian, and B. Torrèsani</i>	Asymptotic Wavelet and Gabor Analysis: Extraction of Instantaneous Frequencies	644
<i>B. Friedlander and B. Porat</i>	Performance Analysis of Transient Detectors Based on a Class of Linear Data Transforms	665
<i>R. Wilson, A. D. Calway and E. R. S. Pearson</i>	A Generalized Wavelet Transform for Fourier Analysis: The Multiresolution Fourier Transform and Its Application to Image and Audio Signal Analysis	674
<i>A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo (Palacios)</i>	Localized Measurement of Emergent Image Frequencies by Gabor Wavelets	691
Compression and Efficient Representation		
<i>R. R. Coifman and M. V. Wickerhauser</i>	Entropy-Based Algorithms for Best Basis Selection	713
<i>R. A. DeVore, B. Jawerth, and B. J. Lucier</i>	Image Compression Through Wavelet Transform Coding	719
<i>A. H. Tewfik, D. Sinha, and P. Jorgensen</i>	On the Optimal Choice of a Wavelet for Signal Representation	747
Multiresolution Stochastic and Fractal Models		
<i>M. Basseville, A. Benveniste, K. C. Chou, S. A. Golden, R. Nikoukhah, and A. S. Willky</i>	Modeling and Estimation of Multiresolution Stochastic Processes	766
<i>G. W. Wornell and A. V. Oppenheim</i>	Wavelet-Based Representations for a Class of Self-Similar Signals with Application to Fractal Modulation	785
<i>P. Moulin, J. A. O'Sullivan, and D. L. Snyder</i>	A Method of Sieves for Multiresolution Spectrum Estimation and Radar Imaging	801
<i>A. Barbé</i>	A Level-Crossing Based Scaling-Dimensionality Transform Applied to Stationary Gaussian Processes	814
Application to Wavelet Transforms		
<i>X. Yang, K. Wang, and S. A. Shamma</i>	Auditory Representations of Acoustic Signals	824
<i>D. M. Healy and J. B. Weaver</i>	Two Applications of Wavelet Transforms in Magnetic Resonance Imaging	840

Our title

- Was:
 - Shiftable Multiscale Transforms.
- Should have been:
 - What's Wrong with Wavelets?

Figures and captions

It should be easy to read the paper in a big hurry and still learn the main points. **Probably most of your readers will be skimming the paper.**

The figures and captions can help tell the story.

So the figure captions should be self-contained and **the caption should tell the reader what to notice about the figure.**

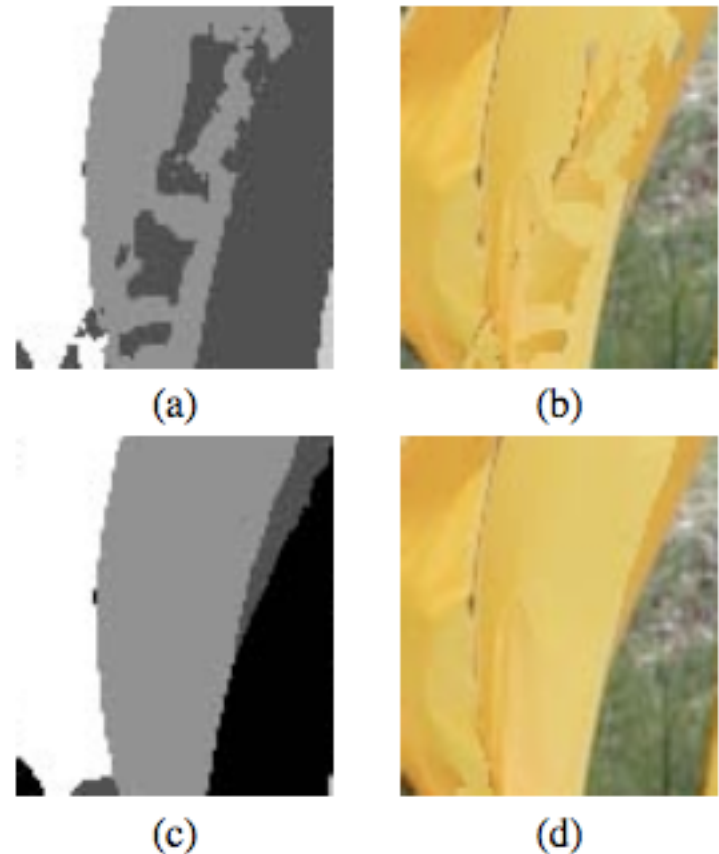


Figure 3: (a) Time-frame assignments for the front-most surface pixels, based on stereo depth measurements alone, without MRF processing. Grey level indicates the time-frame assignment at each pixel. (b) Shape-time image based on those assignments. (c) Most probable time-frame assignments, computed by MRF. (d) Resulting shape-time image. Note that the belief propagation in the MRF has removed spurious frame assignment changes.

Knuth on equations

13. Many readers will skim over formulas on their first reading of your exposition. Therefore, your sentences should flow smoothly when all but the simplest formulas are replaced by “blah” or some other grunting noise.

Mermin on equations

rule in your original manuscript.

Rule 2 (Good Samaritan rule). A Good Samaritan is compassionate and helpful to one in distress, and there is nothing more distressing than having to hunt your way back in a manuscript in search of Eq. (2.47) not because your subsequent progress requires you to inspect it in detail, but merely to find out what it is *about* so you may know the principles that go into the construction of Eq. (7.38). The Good Samaritan rule says: *When referring to an equation identify it by a phrase as well as a number.* No compassionate and helpful person would herald the arrival of Eq. (7.38) by saying “inserting (2.47) and (3.51) into (5.13) . . .” when it is possible to say “inserting the form (2.47) of the electric field \mathbf{E} and the Lindhard form (3.51) of the dielectric function ϵ into the constitutive equation (5.13)”



Tone: be kind and gracious

Efros's comments within our texture synthesis paper about competing methods.

A number of papers to be published this year, all developed independently, are closely related to our work. The idea of texture transfer based on variations of [6] has been proposed by several authors [9, 1, 11] (in particular, see the elegant paper by Hertzmann et.al. [11] in these proceedings). Liang et.al. [13] propose a real-time patch-based texture synthesis method very similar to ours. The reader is urged to review these works for a more complete picture of the field.

Written from a position of security, not competition

Develop a reputation for being clear and reliable

(and for doing creative, good work...)

- There are perceived pressures to over-sell, hide drawbacks, and disparage others' work. Don't succumb. (That's in both your long and short-term interests).
- “because the author was [REDACTED], I knew I could trust the results.” [a conference chair discussing some of the reasons behind a best paper prize selection].

Be honest, scrupulously honest

Convey the right impression of performance.

MAP estimation of deblurring. We didn't know why it didn't work, but we reported that it didn't work. Now we think we know why. Others have gone through contortions to show why they worked.

Author list

- My rule of thumb: All that matters is how good the paper is. If more authors make the paper better, add more authors. If someone feels they should be an author, and you trust them and you're on the fence, add them
- It's much better to be one of many authors on a great paper than to be one of just a few authors on a mediocre paper.
- The benefit of a paper to you is a very non-linear function of its quality:
 - A mediocre paper is worth nothing.
 - Only really good papers are worth anything.

From an area chair's point of view, the
two types of borderline papers...

Quick and easy reasons to reject a paper

With the task of rejecting at least 75% of the submissions, area chairs are groping for reasons to reject a paper. Here's a summary of reasons that are commonly used:

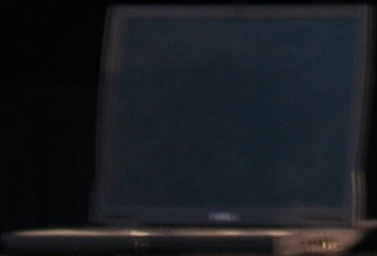
- Do the authors not deliver what they promise?
- Are important references missing (and therefore one suspects the authors not up on the state-of-the-art for this problem)?
- Are the results too incremental (too similar to previous work)?
- Are the results believable (too different than previous work)?
- Is the paper poorly written?
- Are there mistakes or incorrect statements?

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>
- Notes on writing by Fredo Durand, people.csail.mit.edu/fredo/PUBLI/writing.pdf and Aaron Hertzmann, <http://www.dgp.toronto.edu/~hertzman/advice/writing-technical-papers.pdf>
- Three sins of authors in computer science and math, Jonathan Shewchuck, <http://www.cs.cmu.edu/~jrs/sins.html>
- Ten Simple Rules for Mathematical Writing, Dimitri P. Bertsekas http://www.mit.edu:8001/people/dimitrib/Ten_Rules.html

Outline

- writing technical papers
- giving technical talks



How to give talks

- Giving good talks is important for a researcher.
- You might think, “the work itself is what really counts. Giving the talk is secondary”.
- But the ability to give a good talk is like having a big serve in tennis—by itself, it doesn’t win the game for you. But it sure helps. And the very best tennis players all have great serves.
- Researchers as little corporations (see http://people.csail.mit.edu/billf/publications/How_To_Do_Research.pdf).



Sources on giving talks

Patrick Winston's annual IAP talk on how to give talks: <https://vimeo.com/101543862>

Books on speaking.

Suggestions from your advisor or helpful audience members.

Analyzing good talks that others give.

High order bit: prepare

- Practice by yourself.
- Give practice versions to your friends.
- Think through your talk.
- You can write out verbatim what you want to say in the difficult parts.
- Ahead of time, visit where you'll be giving the talk and identify any issues that may come up.
- Preparation is a great cure for nervousness.





David Jacob's 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...



(told to me by David on a beach in Greece, a few hours before my oral presentation at ICCV. That motivated me to leave the beach and go back to my room to work more on my talk, which paid off).

A tip to not be nervous that I found useful

- Get over it. They're not there to see you, they're there to hear the information. Just convey the information to them.

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

- 2-5 minute talks
- Longer talks (10 -20 minute conference presentations, or 30-60 minute colloquia)

Recommendation

- For your five-minute talks, write down:
 - what problem did you address?
 - why is it interesting?
 - why is it hard?
 - what was the key to your approach?
 - how well did it work?

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

- 2-5 minute talks
- Longer talks (10 -20 minute conference presentations, or 30-60 minute colloquia)

Figure out how one part follows from another

Ahead of time, think through how each part motivates the next, and point that out during the talk. If one part doesn't motivate the next, consider re-ordering the talk until it has that feel.

Your audience

- Your image of your audience:
 - Paying attention, listening to every word
- Your audience in reality:
 - Tired, hungry, not wanting to sit through yet another talk at the conference...

Layering the talk. When we read a paper, headings and sections help us follow the paper. You should provide the verbal equivalents of headings to the listener.



http://tbn0.google.com/images?q=tbn:4oWYOjaSp4vpM:http://bakery.grillforallseasons.com/photos/wedding_cake3.jpg

The probability of an observation has three terms to it.

Blah blah blah blah blah blah blah blah blah blah
blah blah blah blah blah blah blah blah blah blah
blah blah blah blah blah blah blah blah blah blah
blah blah blah

So that gives us the objective function we want to optimize. Now, how do we find the optimal value?

There are two approaches you can take. blah blah blah
blah blah blah blah blah blah blah blah blah blah
blah blah blah blah blah blah blah blah blah blah
blah blah blah blah blah

So now, with these tools in hand, we can apply this methods to real images. blah blah blah blah blah
blah blah blah blah blah blah blah blah blah blah

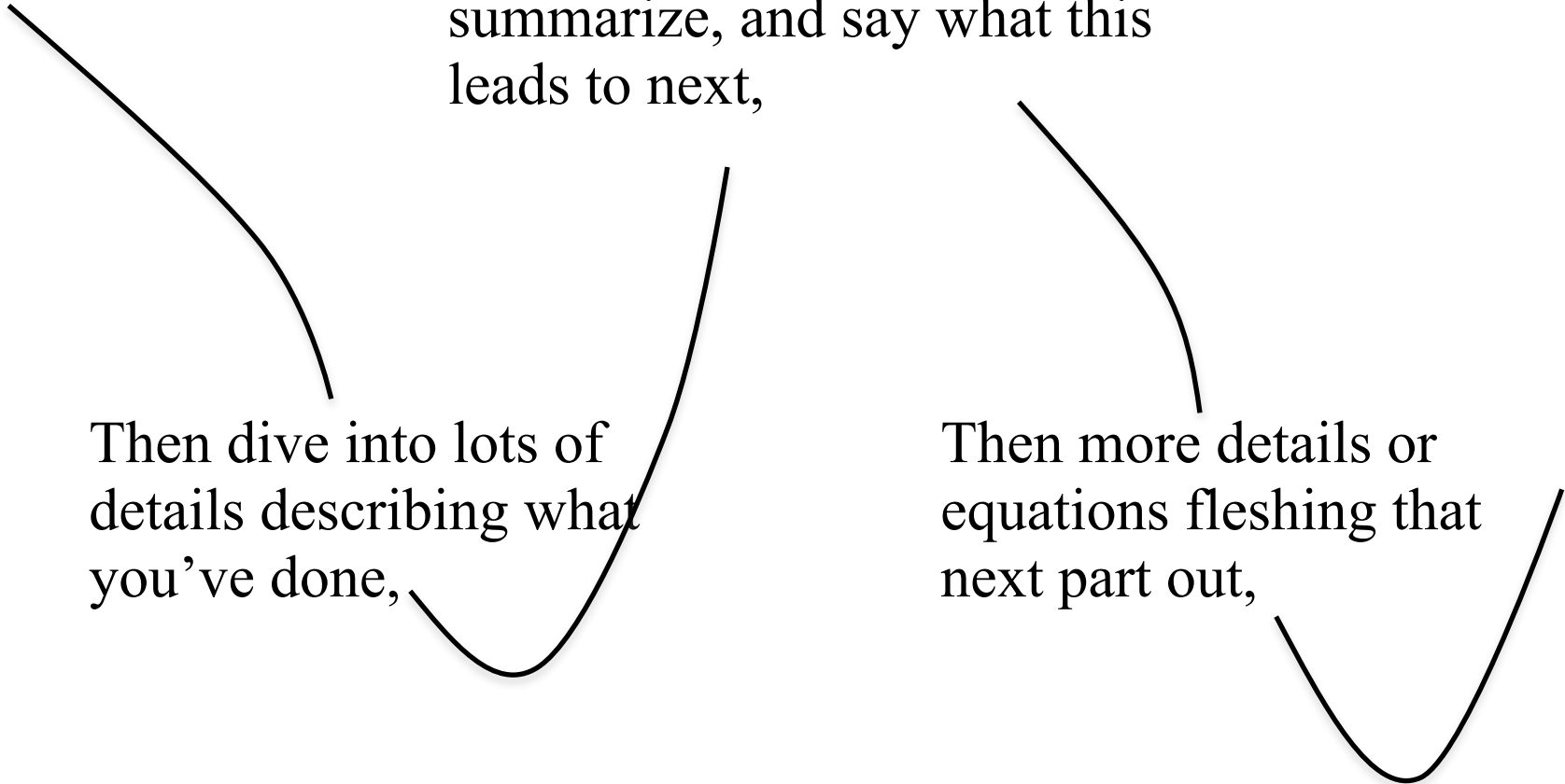
You tell the story at several different levels of detail

The main idea

Then come up for air,
summarize, and say what this
leads to next,

Then dive into lots of
details describing what
you've done,

Then more details or
equations fleshing that
next part out,

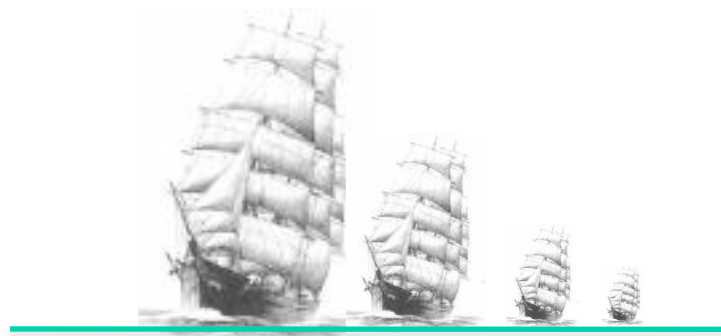


Ways to engage the audience

- So you've been talking on and on. You want to break things up and keep the audience engaged. Can you think of a way to bring the audience into the talk?
- Demos can also help.
- Or add audience participation components to the talk. For human or computer vision talks, you can often present to the audience what the task is that the human or computer has to solve.
- The audience loves to figure things out, to solve puzzles, to make guesses. Feed those desires.

Ted Adelson

- “people like to see a good fight”
- The flat earth theory predicts that ships will appear on the horizon as small versions of the complete ship. Under that theory, you’d expect approaching ships to look like this:



Present a fight

Whereas the round earth theory predicts that the top of the sails will appear first, then gradually the rest of the ship below it.





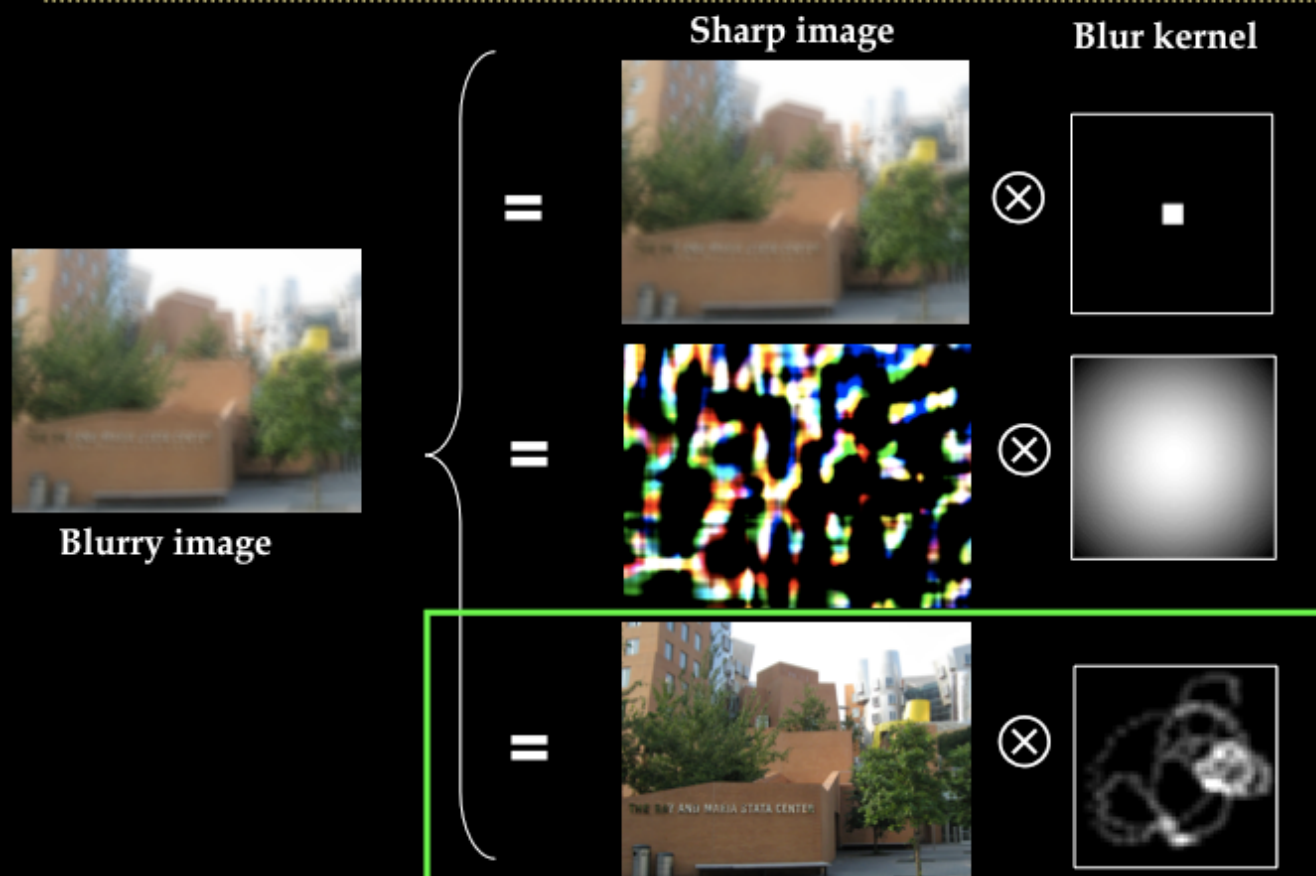
[http://www.flickr.com/photos/mnsomero/
2738807250/](http://www.flickr.com/photos/mnsomero/2738807250/)

Add dynamics to the talk



- A talk is a story. As in a story, there can be different levels of excitement or tension in different parts of the talk. This makes it easier for the audience to pay attention to what you're saying. Perhaps move to another location.
- I like to find some part of the work that really grabs me, that I'm really excited about, and let that show through. (The audience loves to see you be excited. Not all the time, but when appropriate). "I love this problem; it's beautifully underdetermined. There are lots of different ways we can explain the observed blurry image. It could be that that's what was there in the world, and we took a sharp picture of it...."

Multiple possible solutions



What I think the audience of a technical talk wants

To have everything follow and make sense

To learn something

To connect with the speaker, to share their excitement.

They want to watch you love something!

Alan Alda: <https://www.youtube.com/watch?v=j4XgjkXDxss>, and others

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.



How to end a talk

- People often say “are there any questions?” but then people don’t know whether to applaud or to raise their hand.
- If you say “thank you”, then everyone knows that they’re supposed to applaud now. After that is over, then you can ask for questions.