How to write a good research paper

Bill Freeman
MIT CSAIL
Nov. 29, 2018
A paper’s impact on your career

Paper quality

Effect on your career

- Bad
- Ok
- Pretty good
- Creative, original and good.
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
Removing Camera Shake from a Single Photograph

Meh Fergus*, Bernhard Girod†, Anette Wibisono‡, Sam L. Abrams*, William L. Freeman*

†Stanford University,‡University of Toronto

Abstract
Camera shake during exposure leads to objectionable image blur and ruins many photographs. Conventional blind deconvolution methods typically assume frequency-domain images or many cycles of periodic Fourier filtering. As the motion progresses during exposure, the effective spatial frequency of the image decreases, leading to a spatially incoherent blur. An algorithm using a sparse Fourier prior for image restoration can solve this problem. We describe a new method to estimate the effects of camera shake from video-based blurred images. The method assumes a stationary camera and a known motion. An example of a camera shake that is difficult to correct is shown.

Keywords: camera shake, deblurring, image processing.

1 Introduction
Camera shake, in which an unintended camera motion results in blurry photographs, is a common problem in photography. In exposure of a camera the photographic plate moves relative to the position of a small, non-moving, high-resolution image. The movement of the camera is usually not controlled or reproducible with control.

*Supported by the National Science Foundation (NSF) through the Knowledge Capture Program.
2 Related Work

The goal of image restoration is to recover the original image, in the absence of noise, from the distorted image. This is often modeled as a denoising problem, where the goal is to estimate the original image $\hat{I}$ from the degraded image $I$. The degradation model is typically given by $I = D(I') + N$, where $D$ is the degradation process and $N$ is the noise.

The simplest approach to image restoration is to use a point spread function (PSF) model, where the degraded image is considered to be a convolution of the original image and a PSF. This approach is computationally efficient but may not provide good results for images with high levels of noise.

Another approach is to use a variational model, which is based on the principle of minimizing a functional that represents the difference between the original image and the restored image. This approach is more flexible and can handle a wide range of image degradation models.

3 Image model

Our algorithm uses a simple linear model $A$ with a learned weight $W$ to represent the convolution of a blur kernel $B$ with the original image $I$. This model is given by $A = B * I$, where $*$ denotes convolution.

The problem is to estimate the original image $I$ from the degraded image $A$. This can be formulated as an optimization problem, where we seek to minimize a cost function that measures the difference between the degraded image and the convolution of the original image with the estimated blur kernel.

4 Algorithm

There are several algorithms that can be used to solve this optimization problem. One popular approach is to use a gradient-based optimization method, such as the proximal gradient descent. This method iteratively updates the estimate of the original image by solving a series of subproblems that are easier to solve.

Another approach is to use a non-convex optimization method, such as the alternating direction method of multipliers (ADMM). This method iteratively updates the estimate of the original image and the blur kernel, which allows the algorithm to achieve results similar to those obtained with convex optimization.

5.1 Retraining the blur kernel

Given the degraded image $D(I)$, we estimate the blur kernel $B$ by finding the values with highest probability in the gradient of $D(I)$. Once we have estimated the blur kernel, we use it to estimate the original image via the optimization problem described above.
Figure 7: The multi-scale inference scheme operating on the input image in Figure 1. (a) 2nd view: The estimate for the blur kernel is a single scale. (b) 3rd view: The intensity image was reconstructed from the gradient image and was used to compute the final estimated blur kernel.

5 Experiments

We performed an experiment to check that the blur images are ranked due to their actual motion as expressed by the blur kernel, such as in-plane rotation. To do this, we asked 5 people to estimate the blur kernel of a blurred image, and the results were averaged across the blur kernels. The average kernel was then used to estimate the blur image and was saved as a new image. The blur images were then deblurred using our algorithm and compared to the original images using the PSNR metric. The results showed that our algorithm was able to accurately deblur the images.

Figure 8: Best: A case with complex motion. While the blur kernel is not visible, the blur is still evident in the image. The blur images were then deblurred using our algorithm and compared to the original images using the PSNR metric. The results showed that our algorithm was able to accurately deblur the images.

Figure 9: Above: A case with small blur. The blur images were then deblurred using our algorithm and compared to the original images using the PSNR metric. The results showed that our algorithm was able to accurately deblur the images.

Image with significant saturation. Figures 12 and 13 contain areas where the hue and intensity are not visible, due to the significant saturation. This is because the hue and intensity of the image are not preserved, and the image is dominated by the saturation. The edges of the image are still visible, even though the hue and intensity are not.

Figure 10: The blur kernel estimate obtained from the image in Figure 9. The blur kernel is then used to deblur the image. The results show that our algorithm is able to accurately deblur the image.

Figure 11: A case with complex motion. While the blur kernel is not visible, the blur is still evident in the image. The blur images were then deblurred using our algorithm and compared to the original images using the PSNR metric. The results showed that our algorithm was able to accurately deblur the images.

Figure 12: Best: A case with complex motion. While the blur kernel is not visible, the blur is still evident in the image. The blur images were then deblurred using our algorithm and compared to the original images using the PSNR metric. The results showed that our algorithm was able to accurately deblur the images.
6 Discussion

We have introduced a method for removing camera shake effects from photographs. This problem appears highly underdetermined if at all. However, we have shown that by applying several image processing techniques and advanced statistical techniques, plausible results can be obtained. Such an approach may prove useful in other computational photography problems.

Most of our effort has focused on blur estimation, and, visually, the kernels we estimate seem to match the image kernels motion. The results of our method often contain a few more prominent edges in our successful runs and regions of high object motion. We suspect that these artifacts can be blamed primarily on the model's estimation errors. We believe that this is significant since it may improve the performance of applying modern statistical methods to further blur estimation problems.

There are a number of common photographic effects that we do not explicitly model, including saturation, object motion, and cornering artifacts. Incorporating these factors into our model would improve robustness. However, as we assume images have a linear transform, once the gamma correction has been removed. Therefore, cameras typically have a slight vignette that is removed in the inverse space to correct for their dynamic range. Finally, this nonlinearity would be removed, perhaps by minimizing it during inference, or by reestimating the camera from a series of bracketed images.

Figure 7: A scene with a large blur. Bottom: Output of our algorithm. See Figure 1 for a closeup view.

Figure 8: Closeup of the scene in Figure 7. The original image (left) shows a moiré pattern induced by the camera motion. In the deburred image (right), the moiré pattern is much reduced. These images were captured with a low-light camera and processed using our algorithm.

Figure 9: A scene with significant camera motion. The image (left) was captured using a low-light camera while the camera was moving. The output of our algorithm is shown on the right. Note the double exposure type blur kernel.

Figure 10: Top: A scene with significant motion blur. The blur is caused by the camera moving during exposure. Bottom: The output of our algorithm. Note the double exposure type blur kernel.
Acknowledgements

We are indebted to Antonio Torralba, Dan Ceman and Srebro Dixon for their insights and suggestions. We are most grateful to James Smith and David Young, for making their cultural data online. We would like to thank the following people for supplying us with shared images for the paper: Omar Khan, Richard Klein, Richard Lee, Pietro Perona, and Elizabeth Van Genderen. Funding for this project was provided by NSF and YIIE 060368 and the Microsoft Research.

References

[References not visible]

Appendix A

Here we give pseudocode for the algorithm Image Deblur. This code can be referenced in the main text. It's adapted from Nelder and Mead's (1965) optimization algorithm. For clarity, only the basic steps are described. In the actual implementation, the Nelder-Mead method is used. The Nelder-Mead method is a derivative-free optimization algorithm that refines a multidimensional search procedure.

Algorithm 1 Image Deblur

Input: A blurred image $I$, a blur kernel $K$, and a weight $w$

Output: A deblurred image $D$.

1. Define a search space $S$.
2. Initialize the simplex $S_0$.
3. While $S$ is not converged do:
   a. For each vertex $v_i$ in $S$ do:
      i. Compute the function value $f(v_i)$.
   b. Update $S$ according to the Nelder-Mead method.
4. Return $D$. 

Nelder-Mead method:

1. If $f(S)$ is not converged, go to step 2.
2. If $f(S)$ is converged, return $D$.
3. For each vertex $v_i$ in $S$ do:
   a. Compute the function value $f(v_i)$.
4. Update $S$ according to the Nelder-Mead method.
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
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
Show simple toy examples to let people get the main idea

From "Shiftable multiscale transforms"
Steerable filters simple example

Fig. 1. Example of steerable filters: (a) $G_{1}^{0^\circ}$ first derivative with respect to $x$ (horizontal) of a Gaussian; (b) $G_{1}^{90^\circ}$, which is $G_{1}^{0^\circ}$, rotated by $90^\circ$. From a linear combination of these two filters, one can create $G_{1}^{\theta}$, which is an arbitrary rotation of the first derivative of a Gaussian: (c) $G_{1}^{60^\circ}$, formed by $\frac{1}{2}G_{1}^{0^\circ} + \frac{\sqrt{3}}{2}G_{1}^{90^\circ}$. The same linear combinations used to synthesize $G_{1}^{\theta}$ from the basis filters will also synthesize the response of an image to $G_{1}^{\theta}$ from the responses of the image to the basis filters; (d) image of circular disk; (e) $G_{1}^{0^\circ}$ (at a smaller scale than pictured above) convolved with the disk (d); (f) $G_{1}^{90^\circ}$ convolved with (d); (g) $G_{1}^{60^\circ}$ convolved with (d), obtained from $\frac{1}{2}$ (image (c)) + $\frac{\sqrt{3}}{2}$ (image (f)).
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

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?”
Experimental results from Fergus et al paper

Figure 10: Top row: Inferred blur kernels from four real images (the cafe, fountain and family scenes plus another image not shown). Bottom row: Patches extracted from these scenes where the true kernel has been revealed. In the cafe image, two lights give a dual image of the kernel. In the fountain scene, a white square is transformed by the blur kernel. The final two images have specularities transformed by the camera motion, revealing the true kernel.
Experimental results from a later deblurring paper

![Bar chart showing evaluation results for different error ratios.](image)

Figure 9. Evaluation results: Cumulative histogram of the deconvolution error ratio across test examples.
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

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.

12. Motivate the reader for what follows. In the example of §2, Lemma 1 is motivated by the fact that its converse is true. Definition 1 is motivated only by decree; this is somewhat riskier.

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:

<table>
<thead>
<tr>
<th>the question as to whether</th>
<th>whether (the question whether)</th>
</tr>
</thead>
<tbody>
<tr>
<td>there is no doubt but that</td>
<td>no doubt (doubtless)</td>
</tr>
<tr>
<td>used for fuel purposes</td>
<td>used for fuel</td>
</tr>
<tr>
<td>he is a man who</td>
<td>he</td>
</tr>
<tr>
<td>in a hasty manner</td>
<td>hastily</td>
</tr>
<tr>
<td>this is a subject which</td>
<td>this subject</td>
</tr>
<tr>
<td>His story is a strange one.</td>
<td>His story is strange.</td>
</tr>
</tbody>
</table>
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.
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.
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.

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.

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.
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.

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.
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 $\varepsilon$ into the constitutive equation (5.13) . . .”
Tone: be kind and gracious

• My initial comments.
• My advisor’s comments to me.
Image Quilting for Texture Synthesis and Transfer

Alexei A. Efros\textsuperscript{1,2} \hspace{1cm} William T. Freeman\textsuperscript{2}

\textsuperscript{1}University of California, Berkeley \hspace{1cm} \textsuperscript{2}Mitsubishi Electric Research Laboratories

Abstract

We present a simple image-based method of generating novel visual appearance in which a new image is synthesized by stitching together small patches of existing images. We call this process image quilting. First, we use quilting as a fast and very simple texture synthesis algorithm which produces surprisingly good results for a wide range of textures. Second, we extend the algorithm to perform texture transfer — rendering an object with a texture taken from a different object. More generally, we demonstrate how an image can be re-rendered in the style of a different image. The method works directly on the images and does not require 3D information.

Keywords: Texture Synthesis, Texture Mapping, Image-based Rendering

1 Introduction

In the past decade computer graphics experienced a wave of activity in the area of image-based rendering as researchers explored the idea of capturing samples of the real world as images and using them to synthesize novel views rather than recreating the entire
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 Fleet, 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 order

• Some communities use alphabetical order (physics, math).
• For biology, it’s like bidding in bridge.
• Engineering seems to be: in descending order of contribution.
• Should the advisor be on the paper?
  – Did they frame the problem?
  – Do they know anything about the paper?
  – Do they need their name to appear on the papers for continued grant support?

My experiences with having names on papers
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.
Title?
Our title

- Was:
  - Shiftable Multiscale Transforms.
- Should have been:
  - What’s Wrong with Wavelets?
How papers are evaluated

After the papers come in:

• Program chairs assign each paper to an area chair.
• Area chairs assign each of their papers to 3 (or for SIGGRAPH, 5) reviewers.
• Reviewers read and review 5 – 15 papers.
• Authors respond to reviews.
• Area chairs read reviews and author/reviewer dialog and look at paper and decide whether to reject or accept as poster or oral talk. The area chair may have 30 or so papers to handle.
Strategy tips
From an area chair’s point of view, the types of papers in your pile

• About 1/3 are obvious rejects
• In the whole set, maybe 1 is a really nice paper--well-written, great results, good idea. That will be an oral presentation.
• The rest are borderline, and these fall into two camps...
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

- Three sins of authors in computer science and math, Jonathan Shewchuck, http://www.cs.cmu.edu/~jrs/sins.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.
Sources on giving talks

Patrick Winston’s annual IAP talk on how to give talks.

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.
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
- 10-20 minute conference presentations
- 30-60 minute colloquia
Very short talks

- Rehearse it.
- Cut things out that aren’t essential. You can refer to them at a high level.
- You might focus on answering just a few questions, eg: what is the problem? Why is it interesting? Why is it hard?
- Typically these talks are just little advertisements for a poster or for some other (longer) talk. So you just need to show people that the problem is interesting and that you’re fun to talk with.
- These talks can convey important info--note popularity of SIGGRAPH fast forward session.
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
• 10-20 minute conference presentations
• 30-60 minute colloquia
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).
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.

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 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 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 blah blah 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 figures things out, to solve puzzles, to make guesses. Feed those desires.

• The response-meter.
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.
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

Blurry image

\[
\text{Blurry image} \approx \text{Sharp image} \circ \text{Blur kernel}
\]
What I think the audience 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’s comments (see http://mcgovern.mit.edu/video-gallery, starting at 18 minutes in (but earlier is good, too).)
Present to the mean.
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.