ENHANCING PHOTOGRAPHS USING
CONTENT-SPECIFIC IMAGE PRIORS

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Computer Science

by

Neel Suresh Joshi

Committee in charge:

Professor David J. Kriegman, Chair
Professor Samuel R. Buss
Professor Henrik Wann Jensen
Professor Falko Kuester
Professor Matthias Zwicker

2008
The dissertation of Neel Suresh Joshi is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

________________________________________

________________________________________

________________________________________

________________________________________

Chair

University of California, San Diego

2008
To Kunda and Suresh
All our knowledge has its origins in our perceptions.

—Leonardo da Vinci
# TABLE OF CONTENTS

Signature Page ................................................................. iii
Dedication ................................................................. iv
Epigraph ................................................................. v
Table of Contents ................................................................. vi
List of Figures ................................................................. ix
List of Tables ................................................................. xii
Acknowledgements ................................................................. xiii
Vita and Publications ................................................................. xvi
Abstract of the Dissertation ................................................................. xvii

1 Introduction ................................................................. 1
  1.1 Summary of Original Contributions ................................................................. 3
  1.2 Organization of the Dissertation ................................................................. 5

2 Previous Work ................................................................. 6
  2.1 Deblurring ................................................................. 7
    2.1.1 Image Blur Overview ................................................................. 9
    2.1.2 Image Blur Model ................................................................. 11
    2.1.3 Image Deconvolution ................................................................. 12
    2.1.4 PSF Estimation by Constraining the PSF ................................................................. 16
    2.1.5 PSF Estimation by Constraining the Image ................................................................. 24
    2.1.6 Multi-Image Blur Estimation ................................................................. 34
    2.1.7 Image Deblurring Summary ................................................................. 37
  2.2 Denoising ................................................................. 38
    2.2.1 Basic Filtering Methods ................................................................. 39
    2.2.2 Edge-Preserving Methods ................................................................. 40
    2.2.3 Using Image Priors ................................................................. 41
  2.3 Up-sampling ................................................................. 43
    2.3.1 Image Invariant Filters ................................................................. 44
    2.3.2 Image Dependent Filters ................................................................. 44
    2.3.3 Using Image Priors ................................................................. 45
  2.4 White-Balancing ................................................................. 48
    2.4.1 Estimating Illumination Color ................................................................. 49
    2.4.2 Color Matching ................................................................. 52
3 PSF Estimation using Sharp Edge Prediction . . . . . . . . . . . . . . . . . . . . . . . . . . 55
  3.1 Related Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57
  3.2 Image Formation Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 59
  3.3 Sharp Image Estimation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60
    3.3.1 Blind Estimation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60
    3.3.2 Non-Blind Estimation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 61
  3.4 PSF Estimation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 63
    3.4.1 Computing a Super-Resolved PSF . . . . . . . . . . . . . . . . . . . . . . . . . . . 64
    3.4.2 Computing a Spatially Varying PSF . . . . . . . . . . . . . . . . . . . . . . . . . . 65
  3.5 Chromatic Aberration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 65
  3.6 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 66
  3.7 Discussion and Future Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 70
  3.8 Acknowledgements . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 75

4 Image Enhancement using Color Statistics . . . . . . . . . . . . . . . . . . . . . . . . . . . . 76
  4.1 Related Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 78
  4.2 Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 80
  4.3 Gradient Priors . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 81
    4.3.1 Gaussian and Sparse Gradient Priors . . . . . . . . . . . . . . . . . . . . . . . . . . 81
    4.3.2 Limitations of Gradient Priors . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 82
  4.4 Color Priors . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 83
    4.4.1 The Two-Color Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 84
    4.4.2 Using the Two-Color Model for Deconvolution . . . . . . . . . . . . . . . . . . . . 85
  4.5 Solving for the Final Image . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 87
  4.6 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
    4.6.1 Deblurring . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
    4.6.2 Denoising . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 92
    4.6.3 Up-Sampling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93
    4.6.4 Demosaicing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 94
  4.7 Alpha Distribution Measurements . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
  4.8 Discussion and Future Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 98
  4.9 Acknowledgements . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 101

5 Image Correction using Identity-Specific Priors . . . . . . . . . . . . . . . . . . . . . . . . 102
  5.1 Related Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 103
  5.2 Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 107
    5.2.1 Prior Representation and Decomposition . . . . . . . . . . . . . . . . . . . . . . . . 107
    5.2.2 Enhancement Framework . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 109
    5.2.3 Face Alignment and Mask Computation . . . . . . . . . . . . . . . . . . . . . . . . 110
  5.3 Global Correction Operations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 111
    5.3.1 Image Deblurring . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 112
    5.3.2 Exposure and Color Correction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 115
  5.4 Face-Specific Enhancement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 115
    5.4.1 Modifying Lighting and Texture . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 116
  5.5 Personal Photograph Correction Application . . . . . . . . . . . . . . . . . . . . . . . . . 119
  5.6 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 121
  5.7 Analysis of the Eigenspace Prior . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 127
5.8 Discussion and Future Work ......................... 128
5.9 Acknowledgements .................................. 130

6 Conclusions and Future Work .......................... 131
  6.0.1 Building more “Intelligence” into the Photographic Process .... 132
  6.0.2 Video Enhancement using Content Specific Priors ............... 133
  6.0.3 Enhancement using Images and Video .......................... 133

Bibliography ............................................. 134
Figure 1.1: One of the oldest surviving photographs. This image is believed to be the second photograph ever taken and the first using a camera.  ........................................ 2
Figure 2.1: Examples of Image Blur. .................................................. 7
Figure 2.2: Defocus blur (left) and motion blur (right). .............................. 8
Figure 2.3: Blind deconvolution is under-constrained. ............................... 9
Figure 2.4: Image Formation Model. The imaging model consists of two geometric transforms as well as blur induced by motion, defocus, sensor anti-aliasing, and finite-area sensor sampling. ............................................... 11
Figure 2.5: Sparse Gradient Distributions. ............................................ 15
Figure 2.6: The Richardson-Lucy algorithm shows ringing artifacts, while the Gaussian prior is smoothed, but still contains ringing artifacts. The sparse prior show less noisy and ringing and sharper edges. [From Levin et al. 2007] .......... 17
Figure 2.7: Frequency domain zeros. As the size parameter of a parametric the blur kernel increases (red > green > blue) the spacing and number of zeros changes. By locating these zeros one can recover the blur scale. ............................ 19
Figure 2.8: Selecting the blur kernel size. ........................................... 25
Figure 2.9: Natural Image Statistics. (left) A typical scene. (right) The log distribution of gradient magnitudes within the scene are shown in red. The mixture of Gaussians approximation used by Fergus et al. is shown in green. [From Fergus et al. 2006] ............................................. 29
Figure 2.10: Variational Bayes approximates the full posterior with a joint distribution \( Q(I, K) \) and finds the value of \( K \) such that \( Q(I, K) \) and \( P(I, K|B) \) are most similar as measured by the KL-divergence of the two distributions. [From Fergus et al. 2006] .................................................. 32
Figure 2.11: Result from Fergus et al’s work. (top) The input blurry image and the location on the image that is used to compute the PSF. (bottom) The recovered PSF and deblurred image using Lucy-Richardson. [From Fergus et al. 2006] ............ 33
Figure 2.12: Bascle et al’s Tracking and Temporal deconvolution. (left) Motion is tracked across four input frames. (middle) One of the four input images. (right) Their deblurred output. [From Bascle et al. 1996] ......................... 35
Figure 2.13: Hybrid Imaging. ............................................................. 36
Figure 2.14: Rav-Acha and Peleg’s multi-image deblurring. (left) and (middle) have horizontal and vertical motion blur, respectively. (right) Their recovered image. [From Rav-Acha and Peleg 2005] ................................. 36
Figure 2.15: Yuan et al’s deblurring with a noisy and blurry image pair. ........... 37
Figure 2.16: Comparisons of several denoising method. [From Roth and Black 2005 – top row] and [Liu et al. 2008 – bottom row] ................................. 43
Figure 2.17: Comparisons of several up-sampling methods. [From Fattal 2007] .... 46
Figure 2.18: Comparisons of example-based super resolution [Freeman et al. 2002] and Fattal’s method [2007]. [From Fattal 2007] ................................. 47
Figure 2.19: Comparisons of methods that estimate illuminant color using low-level statistics. 50
Figure 2.20: Estimating illuminant color using the intersection of a dichromatic line with the Planckian locus, i.e., the curve in chromaticity space as specified by Planck’s law for black-body radiation. [Finlayson and Schaefer 2000] ........................................ 51
Figure 2.21: Color matching using linear transformations in RGB and \( l\alpha\beta \) color spaces. .................. 53

Figure 3.1: Sharp Edge Prediction. ............................................. 56
Figure 3.2: Image Formation Model. ............................................ 58
Figure 3.3: Non-Blind Estimation. (left) The tiled calibration pattern, (middle) cropped section of an image of a printed version of the grid, and (right) the corresponding cropped part of the known grid warped and shaded to match the image of the grid. ..................................................................................... 62
Figure 3.4: Recovering Blur Kernels of Different Sizes and Orientations. .................. 66
Figure 3.5: Defocus and Slight Motion-Blur. ......................................... 67
Figure 3.6: Kernel Size and Orientation. Image deconvolved with (left) our kernel, (middle) our kernel scaled 20% larger, and (right) our kernel rotated by 45°. The middle and right images have more ringing (most apparent at the bottom of the word “Leicester”). ..................................................................................... 68
Figure 3.7: Defocus and Slight Motion-Blur. ......................................... 69
Figure 3.8: Motion Blur. ..................................................................... 70
Figure 3.9: 4x Super-Resolution. ............................................................ 71
Figure 3.10: Different Apertures and Focal Lengths. ......................................... 72
Figure 3.11: Sub-Pixel PSFs. ................................................................. 73
Figure 3.12: Blind Chromatic Aberration. ................................................ 73
Figure 3.13: Chromatic Aberration. .......................................................... 73
Figure 3.14: Iterative Blind Deconvolution with Sharp Edge Prediction. .................. 74

Figure 4.1: Deblurring with a two color prior. ............................................ 77
Figure 4.2: (a) There are many sharp edges that can blur to match the observed blurred (and potentially noisy) edge (shown in tan). The sparse prior always prefers the smallest intensity gradient that is consistent with the observation (shown in red). ..................................................................................... 83
Figure 4.3: Over-smoothing and noise texturing with the sparse gradient prior ......... 85
Figure 4.4: Using the two-color prior .......................................................... 86
Figure 4.5: Primary, secondary, and alpha maps for the peppers image in Figure 4.4. .... 87
Figure 4.6: Deblurring Text: Blurred, noisy image (the PSF is 31x31 pixels and \( \sigma = 0.01 \)), deconvolution with Lucy-Richardson, the sparse prior, our result using the two-color prior, and the groundtruth for two images. ................................................................. 89
Figure 4.7: Peppers: Blurred, noisy image (the PSF is 31x31 pixels and \( \sigma = 0.01 \)), deconvolution with Lucy-Richardson, the sparse prior, our result using the two-color prior, and the groundtruth for two images ................................................................. 89
Figure 4.8: Three Colors Meeting at a Point: Even when the two color model does not strictly hold within a neighborhood, the perceptual artifacts in this failure case are minimal. ..................................................................................... 90
Figure 4.9: A dragon and sweater from Yuan et al. and a fountain from Fergus et al. ...... 91
Figure 4.10: Images from the Berkeley Image Database ....................................... 92
Figure 4.11: Denoising: Visual comparison of our denoising results with previous work. The fur in the bear is sharper in our result. ................................................................. 92
Figure 4.12: Deconvolution on an up-sampled grid. We show our method run on a 1× and 2× grid. .......................................................... 94
Figure 4.13: Up-sampling low-resolution images. Our formulation also allows us to perform more traditional up-sampling of low-resolution images. ......................... 95
Figure 4.14: Kodak true color images used for demosaicing experiments. ............ 96
Figure 4.15: Measurements of alpha distributions. ..................................... 99
Figure 4.16: Comparing alpha penalty functions. ........................................... 100

Figure 5.1: Automatically correcting personal photographs. We automatically enhance images using prior examples of “good” photographs of a person. ................. 104
Figure 5.2: Personal image enhancement pipeline. ............................................ 105
Figure 5.3: Mask computation and layer decomposition. We perform our corrections on an “intrinsic image” style decomposition of an image into color, lighting, and texture layers. ................................................................. 108
Figure 5.4: Eigenfaces Constraint. We use linear feature spaces built from an aligned set of good images of a person as a constraint in our image enhancement algorithms. Here we show the eigenfaces used as a prior for the delurring result shown in Figure 5.1. ................................................................. 109
Figure 5.5: Exposure and color correction. Using the same set of prior images our system automatically corrects exposure and white-balance for three different images containing the same person. .................................................. 114
Figure 5.6: Defocus blur. .............................................................................. 117
Figure 5.7: Super-resolution. ......................................................................... 117
Figure 5.8: Removing high-frequency shadows and uneven illumination. .......... 118
Figure 5.9: Personal photograph correction application. Here we show a screenshot of an initial prototype of our personal photograph correction application. .. 120
Figure 5.10: Additional deblurring example. Our method automatically performs blind-deconvolution to recover the blur kernel. ................................. 121
Figure 5.11: Comparison to Fergus et al.’s PSF estimation method. ................ 122
Figure 5.12: Face hallucination comparisons. We compare our result to performing hallucination using an implementation of Liu et al.’s method and to using our enhancement algorithm with a set of generic faces instead of faces of the same person. .................................................. 122
Figure 5.13: Synthetic deblurring experiments. ............................................. 123
Figure 5.14: Synthetic upsampling experiments. ........................................... 124
Figure 5.15: Face hallucination algorithms without using an intrinsic image decomposition and gradient domain editing. .................................................. 125
Figure 5.16: Comparisons to color constancy. We compare our results to the color constancy algorithms discussed by van de Weijer et al. Our results are more consistent across images, appear better white-balanced, and did not require any parameter tuning. .................................................. 126
Figure 5.17: Edge strengths of images in the Eigenspace. ................................. 127
Table 4.1: Denoising PSNR Comparisons: Our PSNR value are consistently higher than those of Portilla et al.’s method, Liu et al.’s 0th order denoising, and the sparse prior in the higher noise case.  

Table 4.2: Demosaicing. With the two color prior. For experiments run with the images in Figure 4.14 our method shows slight improvements for the green channel and no consistent improvements for the red and blue channels.
ACKNOWLEDGEMENTS

A twenty-six year journey through three schools and three universities has led me to this point. Through times of excitement, wonder, joy, stress, strain, and sacrifice, I have been fortunate to have family, friends, and mentors to guide me. Their help and insight has been immeasurable, and I would not be here without them.

First, I would like to thank my committee: David Kriegman, Henrik Wann Jensen, Matthias Zwicker, Sam Buss, and Falko Kuester. I would like to thank Henrik for his enthusiasm over the years and for grabbing my interest in CSE 272, which helped me get involved in research during my first year. I would like to thank Matthias for the energy he brought to the graphics group and for his advice during my last two years at UCSD. I would especially like to thank my advisor, David Kriegman, for his invaluable advice and for providing me with the freedom to find my own way and the guidance to stay on track. His insight has helped me tie together what otherwise could have been a frantic and scattered research path during the past four years.

Over the years, I have been fortunate to have many wonderful teachers. I would particularly like to thank Tony Hughes and Joan Rainier from Miquon and Bob Gordon, Alice Davis, and Doug Uhlmann at Penn Charter, who always provided extra opportunities for me to learn by way of independent studies or just letting me hang around during free periods and after school.

At Brown University, I was introduced to Computer Science and the wonders of Computer Graphics in CS 15 and 123 taught by Andy van Dam. I will never forget the first program I wrote that made a little cloud drop rain every time I clicked on it – I switched from EE to CS soon after. I would also like thank my fellow Brown Alums: Dan Morris, Merrie Morris, and Stephen Chen for many much-needed random acts of fun and advice concerning my life and career during the past decade.

I would not be in computer vision and graphics if it were not for my collaborators at Stanford University. I would like to thank Bennett Wilburn, Mark Horowitz, and Marc Levoy for taking me on as an inexperienced Master’s student on the camera array project and giving me my first research experience in computer graphics and vision. I would like to thank the members of the Stanford Graphics lab, particularly the (original and honorary) GSslackers (Billy Chen, Jim
Chow, Kayvon Fatahalion, Gaurav Garg, Daniel Horn, Jeff Klingner, Chris McGraw, Kimberly McGraw, Dan Morris, Merrie Morris, Ren Ng, Doantam Phan, Augusto Roman, Vaibhav Vaish, Bennett Wilburn, and Ron Yeh) who provided a fun and supportive environment during my Master’s degree.

I would also like to thank the members of the UCSD Pixel Lab (Sameer Agrawal, Neil Alldrin, Kristen Branson, Manmohan Chandraker, Will Chang, Krystle Elaine U. de Mesa, Piotr Dollar, Craig Donner, Toshiya Hachisuka, Wojciech Jarosz, Arash Keshmirian, Ben Laxton, Wan-Yen Lo, Satya Mallick, Iman Mostafavi, Vincent Rabaud, and Josh Wills) who helped ease my transition to UCSD and have provided an enjoyable and stimulating environment for the past four years.

More than a year of my four years at UCSD have been spent in internships. They were a defining part of my Ph.D. – I would like to thank MERL, Microsoft Research, and Adobe Systems and my co-authors and mentors in these labs and other associated labs: Shai Avidan, Wojciech Matusik, Hanspeter Pfister, Richard Szeliski, Larry Zitnick, and Edward H. Adelson. It has been a great privilege to work with each of them. I learned an immense amount from every one of them during my internships and in collaborations afterwards, and this dissertation would be quite different without them.

Lastly, I must thank my family. Education has always been paramount in my family. My family and their educational experiences have given me a strong source of motivation during my Ph.D. and helped me to complete my degree. Now, at last, I can finally say I am no longer the least educated member of my family!!! My sister Hima, being five and a half years older that me, has been a role model as she studied hard in school and went on to get her Ph.D. in Chemistry at UCSD in 2001. My parents, Kunda and Suresh (or Mommy and Daddy as I still call them today) have always encouraged my curiosity, creativity, and need for experimentation. They have always supported me, whether they were supplying me with a huge box of Legos (my Kindergarten graduation present), taking me to weekly art or guitar classes, or kindly accepting my destruction of some household item when my curiosity went too far (such as when I clogged the vacuum cleaner by vacuuming up toy cars). They have taught me to work hard and showed me that true accomplishment is not about what you do and who sees it but how you do your work and how
you conduct yourself. My parents are the most sincere, hard-working, and selfless people I know. I owe them more than I can convey here.

Portions of this dissertation are based on papers which I have co-authored with others. My contributions to each of these papers are listed below.

- Chapter 3 is based on material published in the article:
  
  
  I was the primary investigator and author of this paper.

- Chapter 4 is based on material that is in preparation for submission:
  
  
  I was the primary investigator and author of this paper.

- Chapter 5 is based on material that is conditionally accepted for publication:
  
  
  I was the primary investigator and author of this paper.
VITA

2000  Bachelor of Science, Brown University
2004  Master of Science with Distinction in Research, Stanford University
2008  Doctor of Philosophy, University of California, San Diego

PUBLICATIONS


ABSTRACT OF THE DISSERTATION

ENHANCING PHOTOGRAPHS USING CONTENT-SPECIFIC IMAGE PRIORS

by

Neel Suresh Joshi

Doctor of Philosophy in Computer Science

University of California San Diego, 2008

Professor David J. Kriegman, Chair

The digital imaging revolution has made the camera ubiquitous; however, image quality has not improved at the same rate as the increase in camera availability. Increasingly more cameras are small, with inexpensive lenses, no flash, and lightweight bodies that are difficult to hold steady, and this results in images with blur, noise, and poor color-balance. Consequently, there is a strong need for simple, automatic, and accurate methods for image correction. This dissertation presents work that uses "content-specific" image models and priors for image enhancement.

Image enhancement is a challenge problem – corrections such as deblurring, denoising, and color-correction are ill-posed, where the number of unknown values outweighs the number of observations. As a result, it is necessary to add additional information as constraints. Previous work has focused on using generic image priors that are applicable to a large number of images. In this work, we develop constraints that are tuned to the specific content of an image.

First, we discuss a fast, accurate blur estimation method that models all edges in a sharp image as step-edges. The method predicts the “sharp” version of a blurry input image and uses the two images together to solve for a PSF. Second, we discuss a framework for image deblurring and denoising that uses local color statistics to produce sharp, low-noise results. Even when the blur function is known, deblurring an image is still quite difficult due to information loss during blurring and due to the presence of
noise. In our work, we investigate using local-color statistics of an image in a joint framework for
deblurring and denoising of images.

Lastly, we discuss work in methods that use “identity-specific” priors to perform cor-
rections for images containing faces. These priors provide the guidance needed to perform
high-quality corrections needed for known, familiar faces. Deblurring, super-resolution, color-
balancing, and exposure correction operate independently, so that a user can correct selected
image properties, while still retaining certain desired qualities of the original photo. We have also
developed a prototype application for performing these corrections.