Beyond the Piece of Cardboard: Learning to Adjust Photographs

Sylvain Paris, Adobe (sparis@adobe.com)

with Mathieu Aubry (ENS/INRIA), Soonmin Bae (MIT), Vladimir Bychkovsky (MIT), Eric Chan (Adobe), Sam Hasinoff (Google), Jan Kautz (UCL), Frédo Durand (MIT),
Photos need to be retouched

count print

final print by Ansel Adams
“Straight out of the camera”
Printing photos in the darkroom

- choice of paper
  - grade paper, properties vary with light color

- projecting negative onto paper
  - dodging and burning

- finishing the print
  - choice of chemicals, possibly painting them

- critical process to get top prints
  - tedious, error-prone
The digital era: Photoshop

- Zoom
- Unlimited undo
- Accurate selection
- Preview
- HDR
- Panorama
- And more...
The digital era: Photoshop

- Zoom
- Unlimited undo
- Accurate selection
- Preview
- HDR
- Panorama
- And more...

Photoshop is only a better piece of cardboard.

We need algorithms that help users make better pictures!
Solution #1
Model-based Adjustment

with Soonmin Bae and Frédo Durand

[SIGGRAPH’06]
Scenario

We want with the look of
Tonal Aspects of Look

Ansel Adams

Kenro Izu
Tonal aspects of Look: Global Contrast

Ansel Adams

High Global Contrast

Kenro Izu

Low Global Contrast
Tonal aspects of Look: Local Contrast

Ansel Adams

Variable amount of texture

Kenro Izu

Texture everywhere
Our approach:

Match global and local contrasts
Related Work: Scale/Frequency Manipulation

• Used for audio visual equalizer
  – controls sound ambiance

• Not really used yet for images
  – Exception: Kai’s Power Tools
Related Work: Tone Mapping

- Reduce global contrast
  
  [Pattanaik 98; Tumblin 99; Ashikhmin 02; Durand 02; Fattal 02; Reinhard 02; Li 05...]

- Seeks neutral reproduction

  ❌ Little control over look

In contrast, we want to achieve particular looks.
Pipeline

Input Image

Decompose

Local contrast
(small-scale variations “texture”)

Global contrast
(large-scale variations “input minus the texture”)

Made fast in
[ECCV’06, SIGGRAPH’07]
Pipeline

Global contrast

Input Image

Decompose

Histogram transfer

“Histogram transfer”

Local contrast

User study in [ICCP’11]
Pipeline

Input Image

Global contrast

Histogram transfer

Recombine

“Histogram transfer”

Local contrast

Decompose
Pipeline

Input Image

Global contrast

Histogram transfer

Local contrast

"Histogram transfer"

Recombine

Result
Try it yourself!

• Slightly simplified version available
  – Adobe Photoshop Elements 9
  – On-line demo at:
    http://www.photoshop.com/tools/stylematch

• Inspired “HDR Pro” in Photoshop CS5
Alternative Algorithm
with Mathieu Aubry, Sam Hasinoff, Jan Kautz, Frédo Durand

GLOBAL
(intensity histogram transfer)

LOCAL
(gradient histogram transfer)

Advantage: no image decomposition ➔ faster, more robust
Alternative Algorithm
with Mathieu Aubry, Sam Hasinoff, Jan Kautz, Frédo Durand

GLOBAL
(intensity histogram transfer)

LOCAL
(gradient histogram transfer)

Nontrivial
(cf. tech report)

Advantage: no image decomposition ➔ faster, more robust
input (HDR)
output

model
Summary

• Model-based adjustment is easy
  – only need to have an example of the desired look

• Tonal aspects of “look” well characterized by
  – \(0^{th}\) order statistics: luminance histogram
  – \(1^{st}\) order statistics: “texture” histogram
Solution #2
Machine Learning

with Vladimir Bychkovsky, Eric Chan,
and Frédo Durand

[CVPR’11]
Objective: Fully Automatic Adjustment

• Input: new, unadjusted photo

• Output: adjusted photo

• We focus on global transformations, i.e. no brush, no mask...
Previous Work: Hand-tuned Algorithms

- Easy to understand, total control
- Depends on the photographic skills of developer
- Painstaking to adapt to specific styles
Previous Work:
Flickr-based Restoration [Dale 2009]

• Many images (cheap to get)

• Only output is available
  – Image descriptor must be invariant to the adjustment

• Demonstrated on degraded images
  – Our input photos are ok.
Previous Work:
Machine Learning on Synthetic Data
[Kang et al. 2010]

• Principled
• Limited because of lack of data
   – does not perform better than best hand-tuned algorithm
Our Approach: Supervised Learning

• A dataset of input and retouched photos

• Transformation as labels

• Image descriptors as features

• We learn the mapping from features to labels
Our Dataset

• 5000 photos in RAW format

• 5 students retouched them by hand
  – Trained at the Visual School of Art in New York
  – Paid for their work
    (5000 photos in 2 months)
It is available on-line.

http://groups.csail.mit.edu/graphics/fivek_dataset/

Data

All of our data can be downloaded as a single archive (~50GB, SHA1). This archive includes the following items:

- 5,000 photos in DNG format
  This format can be read by DCRAW, Adobe Lightroom, and many other tools.
- An Adobe Lightroom catalog with renditions by 5 experts
  This includes values of individual sliders and the full history of adjustments for each photo.
- Semantic information about each photo
  Indoor vs. outdoor, time of day, type of light, and main subject (people, nature, man-made objects, etc)

Please join our FiveK Dataset group to get updates about this dataset or to ask questions about the data.

For your convenience we have made available individual input and output files below. Input files are available in DNG format; the five output renditions are available as TIFF (16 bits per channel, ProPhoto RGB color space, lossless compression) to preserve the maximum amount of information. You can browse these files below.
Tags from Mech. Turks
Examples from our dataset

"input photo"

photographer A

photographer B

photographer C

photographer D

photographer E
Examples from our dataset
Remapping Curve

- We focus on luminance, e.g. brightness / contrast.

- output luminance = f(input luminance)
  - extracted from the image only
  - does not assume a specific software

- brightness & contrast explain ~95% of expert’s adjustments
  - brightness alone explains 80 to 90%.
Image Descriptor

• Ideally, correlates with the adjustment.

• We tried many descriptors:
  – Luminance, color, and gradient distributions
  – Global and local (3x3, 5x5, center / surround...)
  – Scene descriptor (GIST)

• The winner is: global luminance distribution + faces
Discussion about Scene Descriptors

• Scene descriptors are meant to recognize scenes, e.g. street or forest.
  – Good with millions of images, e.g. Flickr.

• Time of the day, lighting conditions, mood matter more to us than the type of scene.
  – Scene descriptors do not help us.
  – Complimentary problem: we want to be sensitive to criteria to which scene descriptors are invariant.
Learning Algorithms

• We tested Nearest Neighbor, Least Squares, LASSO (Sparse Least Squares), Gaussian Process Regression (GPR).

• All perform more or less the same
  – GPR is best and safely extrapolates
    • goes back to neutral adjustment instead of exaggerating
More on Extrapolation
More on Extrapolation

outlier photo
(very different from any training photo)
More on Extrapolation

outlier photo
(very different from any training photo)

Linear extrapolation out of range

brightness

-10 training set +10
More on Extrapolation

GPR average of training set

Outlier photo (very different from any training photo)

Linear extrapolation out of range

-10 training set +10

Brightness
Results

• We selected one of the experts and seek to predict his adjustments.
• Dataset split into training and testing sets
• Input: test photo never seen before
• Goal: reproduce the expert’s version (hidden from the algorithm)
  – L2 norm in CIE Lab
Sample result
Sample result (representative performance)
“Failure”, did not recognize it as a night picture, yet does not do anything wrong
Learning User Preferences

1. Train the algorithm off-line on 2500 photos
2. User gives a few examples (3+) and we learn preferences from them.
Learning User Preferences

1. Train the algorithm off-line on 2500 photos
2. User gives a few examples (3+) and we learn preferences from them.

3. Our algorithm adapts and apply the user’s style
Summary

• Photographers are consistent  
  – makes it possible to use machine learning
• Brightness & contrast explain  
  most tonal adjustments
• Appearance-based features are useful
• Failsafe prediction is important in practice
Discussion
Tones vs Colors

- Tones are comparatively easier
  - e.g. mid-day sky can almost be anything from black to white but must be blue
- Colors conflate physical white balance and “subjective temperature”
  - white balance: is this object white?
    - still an open problem
  - temperature: warm (red) or cold (blue) rendition
Are we stealing photographers’ jobs? No!

- Machine learning finds repeated patterns and reproduces them.
  - Does the repetitive tedious task
  - Does not do the one-of-a-kind adjustment
    - It never will. The artist remains in charge.
User Studies

Does this adjustment make the picture look good?

• Different task, we seek to answer “Is this what a photographer would have done?”

• Many ways to make a picture look good
  – not ideal for research, comparisons

• User studies are difficult
  – dependencies on the protocol,
    image content, viewer’s background...
  – biases, e.g. contrasted images stand out

• Yet, it’s a legitimate question for the “I am feeling lucky” button
  – subtle, not easy
Conclusion

• Smarter-than-cardboard tools that reproduce a photographer’s look from a model or from a training set.
• Some insights into photography.
• More needs to be done!