Automatic Colorization

Gustav Larsson

TTI Chicago / University of Chicago

Joint work with Michael Maire and Greg Shakhnarovich

NVIDIA @ SIGGRAPH 2016

Colorization

Let us first define “colorization”

Colorization

Definition 1: The inverse of desaturation.

Original

Colorization

Definition 1: The inverse of desaturation.

Original

$\xrightarrow{\text{Desaturate}}$

Grayscale

Colorization

Definition 1: The inverse of desaturation.

Grayscale

Colorization

Definition 1: The inverse of desaturation.

Original

$\xleftarrow{\text{Colorize}}$

Grayscale

Colorization

Definition 1: The inverse of desaturation. (**Note: Impossible!**)

Original

$\xleftarrow{\text{Colorize}}$

Grayscale

Colorization

Definition 2: An inverse of desaturation, that...

Grayscale

Colorization

Definition 2: An inverse of desaturation, that...

Our Method

$\xleftarrow{\text{Colorize}}$

Grayscale

... is plausible and pleasing to a human observer.

Colorization

Definition 2: An inverse of desaturation, that...

Our Method

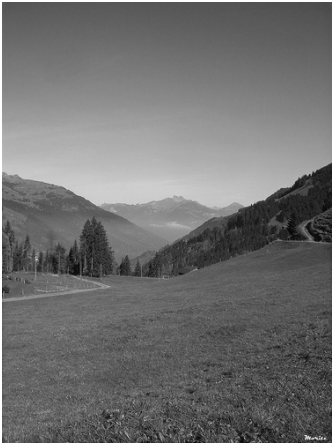
$\xleftarrow{\text{Colorize}}$

Grayscale

... is plausible and pleasing to a human observer.

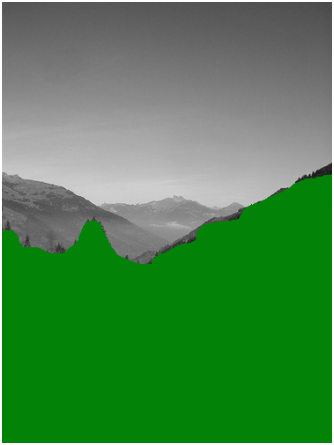
- Def. 1: Training + Quantitative Evaluation
- Def. 2: Qualitative Evaluation

Manual colorization



I thought I would give it a quick try...

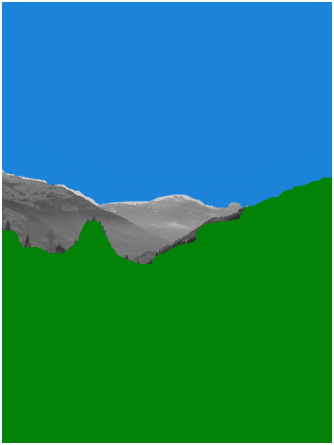
Manual colorization



Grass is green

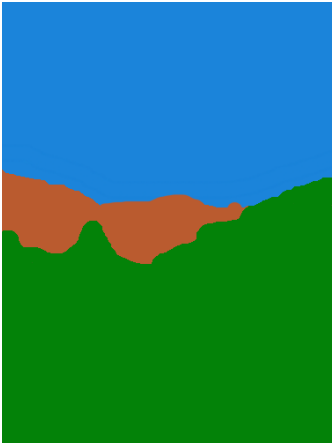
(low-level: grass texture / mid-level: tree recognition / high-level: scene understanding)

Manual colorization



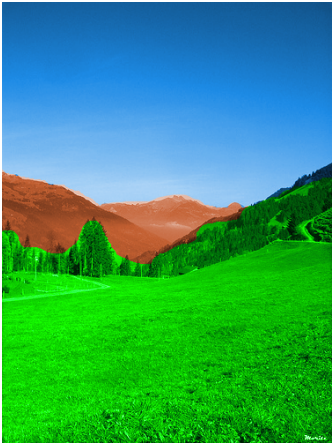
Sky is blue

Manual colorization



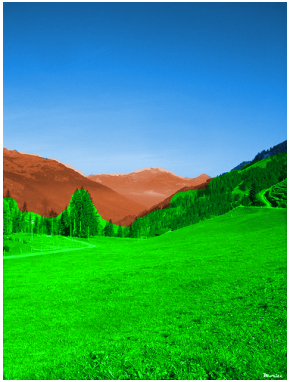
Mountains are... brown?

Manual colorization



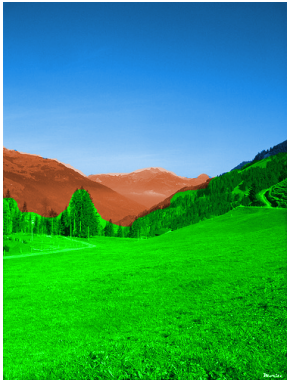
Use the original luminosity

Manual colorization

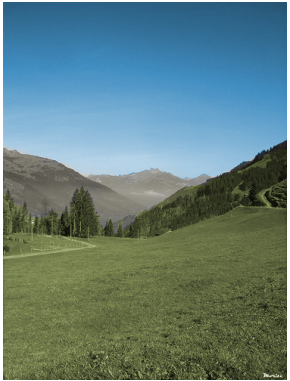


Manual ($\approx$ 15 s)

Manual colorization

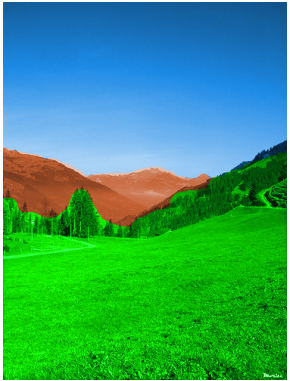


Manual ($\approx$ 15 s)

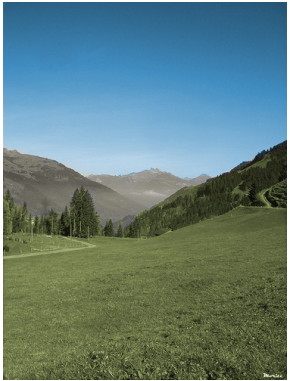


Manual ($\approx$ 3 min)

Manual colorization



Manual ($\approx$ 15 s)



Manual ($\approx$ 3 min)

Automatic (< 1 s)

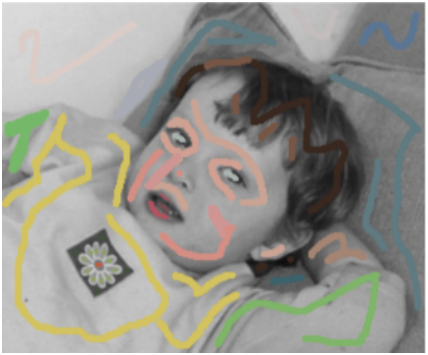
A brief history

The history of computer-aided colorization in 3 slides.

Method 1: Scribbles

Manual ——●—— Automatic

User-defined scribbles define colors. Algorithm fills it in.



Input

Output

Levin et al. (2004)

→ Levin et al. (2004); Huang et al. (2005); Qu et al. (2006); Luan et al. (2007)

Method 2: Transfer

Manual ——●—— Automatic

Reference image(s) is provided. Scribbles are automatically created from correspondences.

Reference

Input

Output

Charpiat et al. (2008)

→ Welsh et al. (2002); Irony et al. (2005); Charpiat et al. (2008); Morimoto et al. (2009); Chia et al. (2011)

Method 2: Transfer

Manual ——●——●——● Automatic

Reference image(s) is provided. Scribbles are automatically created from correspondences.

Reference

Input

Output

Charpiat et al. (2008)

→ Welsh et al. (2002); Irony et al. (2005); Charpiat et al. (2008); Morimoto et al. (2009); Chia et al. (2011)

Method 3: Prediction

Manual ——●—— Automatic

Fully parametric prediction.

`colorize` () = 

Automatic colorization is gaining interest recently:

→ $\underbrace{\text{Deshpande et al., Cheng et al.}}_{\text{ICCV 2015}}$; $\underbrace{\text{Iizuka \& Simo-Serra et al.}}_{\text{SIGGRAPH 2016 (2pm, Ballroom E)}}$ $\underbrace{\text{Zhang et al., Larsson et al.}}_{\text{ECCV 2016}}$

Method 3: Prediction

Manual ——●—— Automatic

Fully parametric prediction.

`colorize_pixel`  $= (60, 87, 44)$

Automatic colorization is gaining interest recently:

→ Deshpande et al., Cheng et al.; Iizuka & Simo-Serra et al. Zhang et al., Larsson et al.

ICCV 2015 SIGGRAPH 2016 (2pm, Ballroom E) ECCV 2016

Model

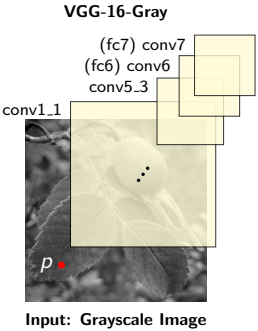
Design principles:

- Semantic knowledge

Model

Design principles:

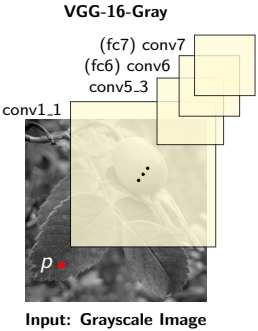
- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier



Model

Design principles:

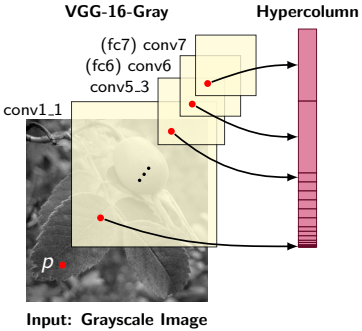
- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features



Model

Design principles:

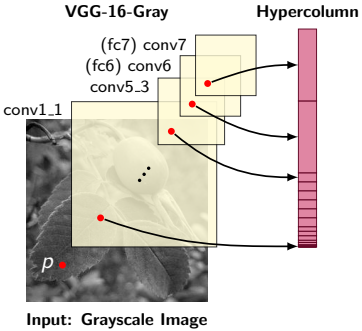
- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features $\rightarrow$ Zoom-out/Hypercolumn architecture



Model

Design principles:

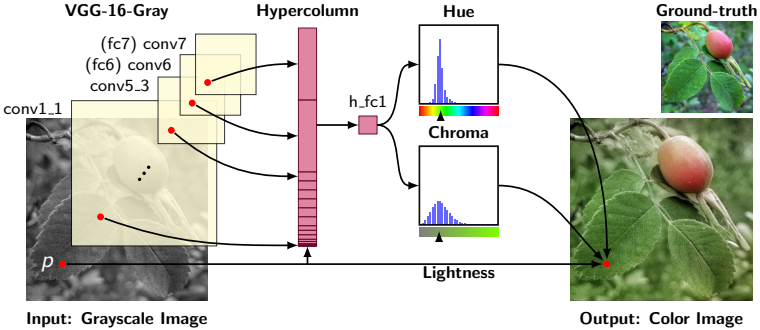
- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features $\rightarrow$ Zoom-out/Hypercolumn architecture
- Colorization not unique



Model

Design principles:

- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features $\rightarrow$ Zoom-out/Hypercolumn architecture
- Colorization not unique $\rightarrow$ Predict histograms



Instantiation

Going from histogram prediction to RGB:

- Sample

Instantiation

Going from histogram prediction to RGB:

- Sample
- Mode

Instantiation

Going from histogram prediction to RGB:

- Sample
- Mode
- Median

Instantiation

Going from histogram prediction to RGB:

- Sample
- Mode
- Median
- Expectation

Instantiation

Going from histogram prediction to RGB:

- Sample
- Mode
- Median $\leftarrow$ Chroma
- Expectation $\leftarrow$ Hue

Instantiation

Going from histogram prediction to RGB:

- Sample
- Mode
- Median $\leftarrow$ Chroma
- Expectation $\leftarrow$ Hue

The histogram representation is rich and flexible:

Instantiation

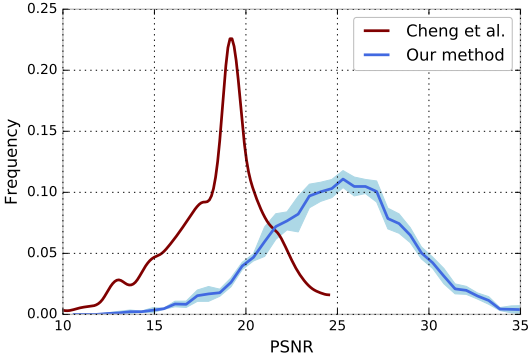
Going from histogram prediction to RGB:

- Sample
- Mode
- Median $\leftarrow$ Chroma
- Expectation $\leftarrow$ Hue

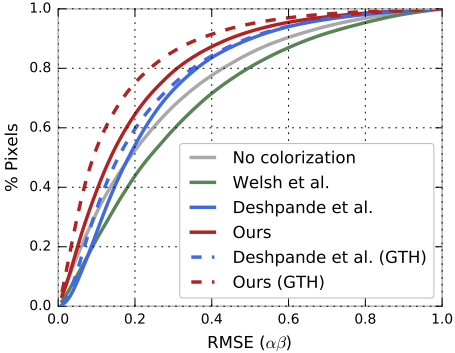
The histogram representation is rich and flexible:

Results

Significant improvement over state-of-the-art:



Cheng et al. (2015)

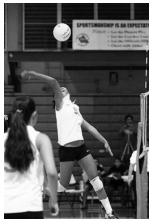


Deshpande et al. (2015)

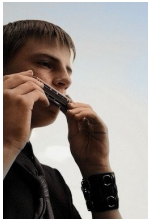
Comparison

Model	AuC CMF non-rebal (%)	AuC CMF rebal (%)	VGG Top-1 Classification Accuracy (%)	Turk Labeled Real (%) mean	Turk Labeled Real (%) std
Ground Truth	100.00	100.00	68.32	50.00	–
Gray	89.14	58.01	52.69	–	–
Random	84.17	57.34	41.03	12.99	2.09
Dahl	90.42	58.92	48.72	18.31	2.01
Zhang et al.	91.57	65.12	56.56	25.16	2.26
Zhang et al. (rebal)	89.50	**67.29**	56.01	**32.25**	2.41
Ours	**91.70**	65.93	**59.36**	27.24	2.31

Table: Source: Zhang et al. (2016)

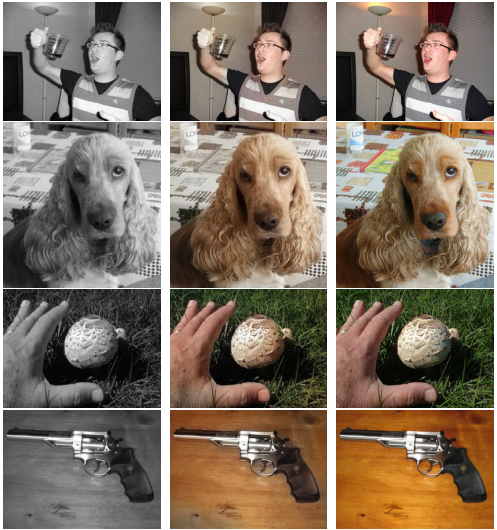


SCHOOL BUS
SPORTSMANSHIP IS AN EXPECTAT
Copyright © 2007 James Markus - All Rights Reserved
Photomatter.com



SCHOOL BUS
Beth Wade Photography
SPORTSMANSHIP IS AN EXPECTATION
Copyright © 2007 James Markus - All Rights Reserved
Photomatter.com

Examples



Input **Our Method** **Ground-truth**

Input **Our Method** **Ground-truth**

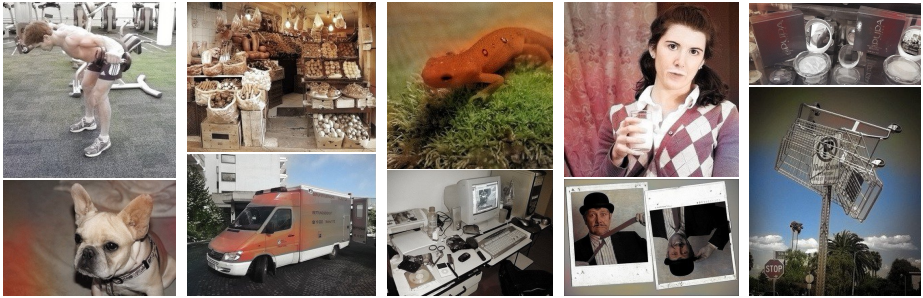


Figure: Failure modes.

Figure: B&W photographs.

Self-supervision (ongoing work)

Colorization as a means to learn visual representations:

1. Train colorization from scratch

Self-supervision (ongoing work)

Colorization as a means to learn visual representations:

1. Train colorization from scratch
2. Use network for segmentation, detection, style transfer, texture generation, etc.

Self-supervision (ongoing work)

Colorization as a means to learn visual representations:

1. Train colorization from scratch
2. Use network for segmentation, detection, style transfer, texture generation, etc.

Initialization	Architecture	X_{ImageNet}	Y_{ImageNet}	Color	mIU (%)
Classifier (ours)	VGG-16	✓	✓		64.0
Random	VGG-16				32.5
Classifier	AlexNet	✓	✓	✓	48.0
Random	AlexNet			✓	19.8

Table: **VOC 2012 segmentation validation set.**

Self-supervision (ongoing work)

Colorization as a means to learn visual representations:

1. Train colorization from scratch
2. Use network for segmentation, detection, style transfer, texture generation, etc.

Initialization	Architecture	X_{ImageNet}	Y_{ImageNet}	Color	mIU (%)
Classifier (ours)	VGG-16	✓	✓		64.0
Random	VGG-16				32.5
Classifier	AlexNet	✓	✓	✓	48.0
BiGAN (Donahue et al.)	AlexNet	✓		✓	34.9
Inpainter (Deepak et al.)	AlexNet	✓		✓	29.7
Random	AlexNet			✓	19.8

Table: **VOC 2012 segmentation validation set.**

Self-supervision (ongoing work)

Colorization as a means to learn visual representations:

1. Train colorization from scratch
2. Use network for segmentation, detection, style transfer, texture generation, etc.

Initialization	Architecture	X_{ImageNet}	Y_{ImageNet}	Color	mIU (%)
Classifier (ours)	VGG-16	✓	✓		64.0
Colorizer	VGG-16	✓			**50.2**
Random	VGG-16				32.5
Classifier	AlexNet	✓	✓	✓	48.0
BiGAN (Donahue et al.)	AlexNet	✓		✓	34.9
Inpainter (Deepak et al.)	AlexNet	✓		✓	29.7
Random	AlexNet			✓	19.8

Table: **VOC 2012 segmentation validation set.**

Questions?

Try it out yourself:
`http://colorize.ttic.edu`

References

Charpiat, G., Hofmann, M., and Schölkopf, B. (2008). Automatic image colorization via multimodal predictions. In ECCV.

Cheng, Z., Yang, Q., and Sheng, B. (2015). Deep colorization. In ICCV.

Chia, A. Y.-S., Zhuo, S., Gupta, R. K., Tai, Y.-W., Cho, S.-Y., Tan, P., and Lin, S. (2011). Semantic colorization with internet images. ACM Transactions on Graphics (TOG), 30(6).

Deshpande, A., Rock, J., and Forsyth, D. (2015). Learning large-scale automatic image colorization. In ICCV.

Huang, Y.-C., Tung, Y.-S., Chen, J.-C., Wang, S.-W., and Wu, J.-L. (2005). An adaptive edge detection based colorization algorithm and its applications. In ACM international conference on Multimedia.

Irony, R., Cohen-Or, D., and Lischinski, D. (2005). Colorization by example. In Eurographics Symp. on Rendering.

Levin, A., Lischinski, D., and Weiss, Y. (2004). Colorization using optimization. ACM Transactions on Graphics (TOG), 23(3).

Luan, Q., Wen, F., Cohen-Or, D., Liang, L., Xu, Y.-Q., and Shum, H.-Y. (2007). Natural image colorization. In Eurographics conference on Rendering Techniques.

Morimoto, Y., Taguchi, Y., and Naemura, T. (2009). Automatic colorization of grayscale images using multiple images on the web. In SIGGRAPH: Posters.

Qu, Y., Wong, T.-T., and Heng, P.-A. (2006). Manga colorization. ACM Transactions on Graphics (TOG), 25(3).

Welsh, T., Ashikhmin, M., and Mueller, K. (2002). Transferring color to greyscale images. ACM Transactions on Graphics (TOG), 21(3).

Zhang, R., Isola, P., and Efros, A. A. (2016). Colorful image colorization. In ECCV.