Deep Learning-Based Approach to Video Colorization

Chester Holtz

Related Work

- Colorful Image Colorization in ECCV 2016
  - (Richard Zhang, Phillip Isola, Alexei A. Efros)
  - Single-network trained to minimize multinomial cross entropy + class rebalancing regularizer

- Learning Representations for Automatic Colorization in ECCV 2016 (*)
  - (Gustav Larson, Michael Maire, Gregory Shakhnarovich)
  - Modified VGG Network to Predict Pixel R G B Channels Via the HCT

- Let there be Color!: Joint End-to-End Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification in SIGGRAPH 2016
  - (Satoshi Iizuka, Edgar Simo-Serra, Hiroshi Ishikawa)
  - Fusion layer merges local information dependent on small image patches with computed using the entire image.

Background

$\rightarrow f$

Greyscale Image: ***L*** Channel

$$X \in \mathbb{R}^{H \times W \times 1}$$

$\rightarrow$

Color Image: ***(L, ab)***

$$(X, \hat{Y})$$

Model Architecture & Training Procedure

- Forward pass through pretrained VGG16
  - Concatenate and interpolate layer activations prior to each maxpool
- Forward pass through “color” network
  - 8 layers, 4 max pooling layers
  - ReLU activation
  - Batch Norm with each convolution
  - SGD (0.1), Dropout (0.2)
- L2 Loss

$$L_2(\hat{Y}, Y) = \frac{1}{2} \sum_{h,w} ||Y_{hw} - \hat{Y}_{hw}||_2^2$$

Frame smoothing

- Optical Flow (Neural Style for video)

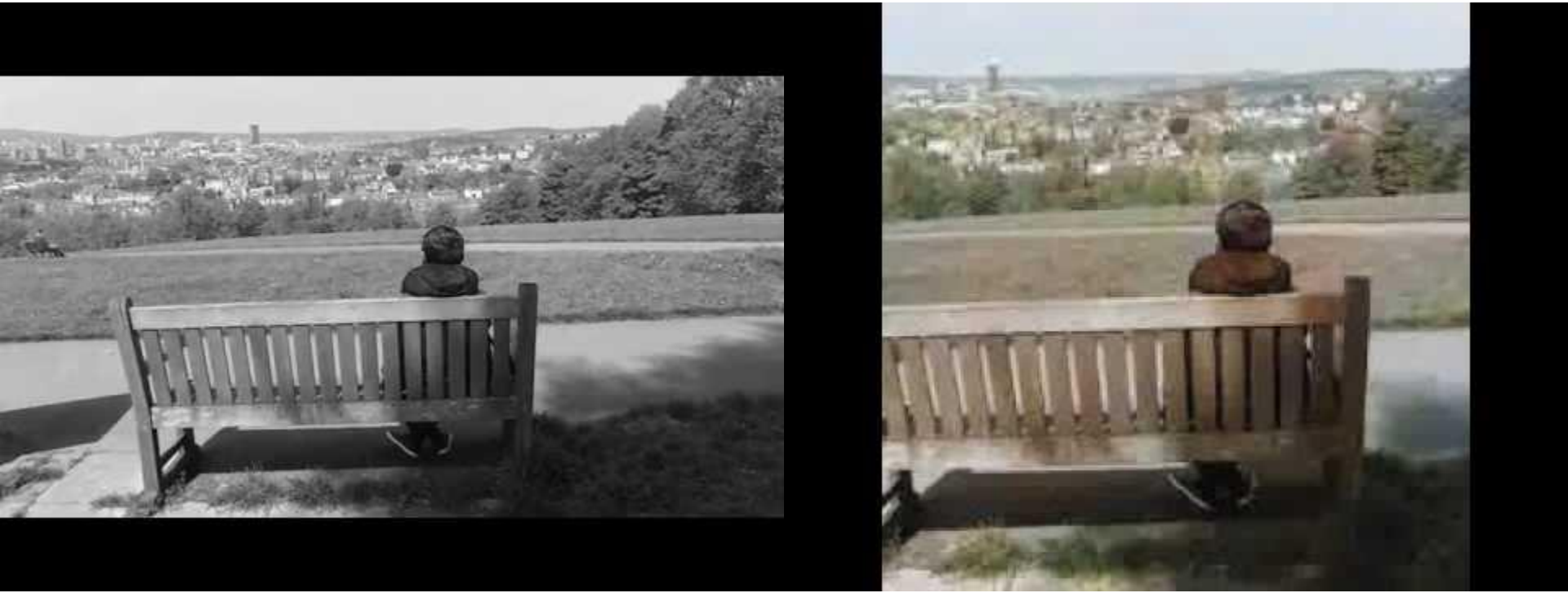
$$\nabla I^T \cdot \vec{V} = -I_t$$

- Total Variation Denoising

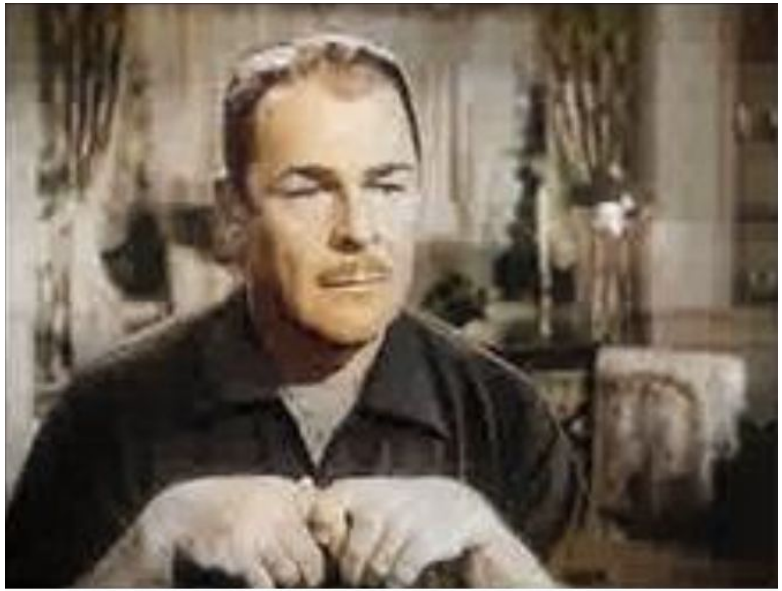
$$\min_x \frac{1}{2}|y_n - x_n|^2 + \lambda \sum |x_n - x_{n-1}|$$

Opencv has functions to solve both of these problems.

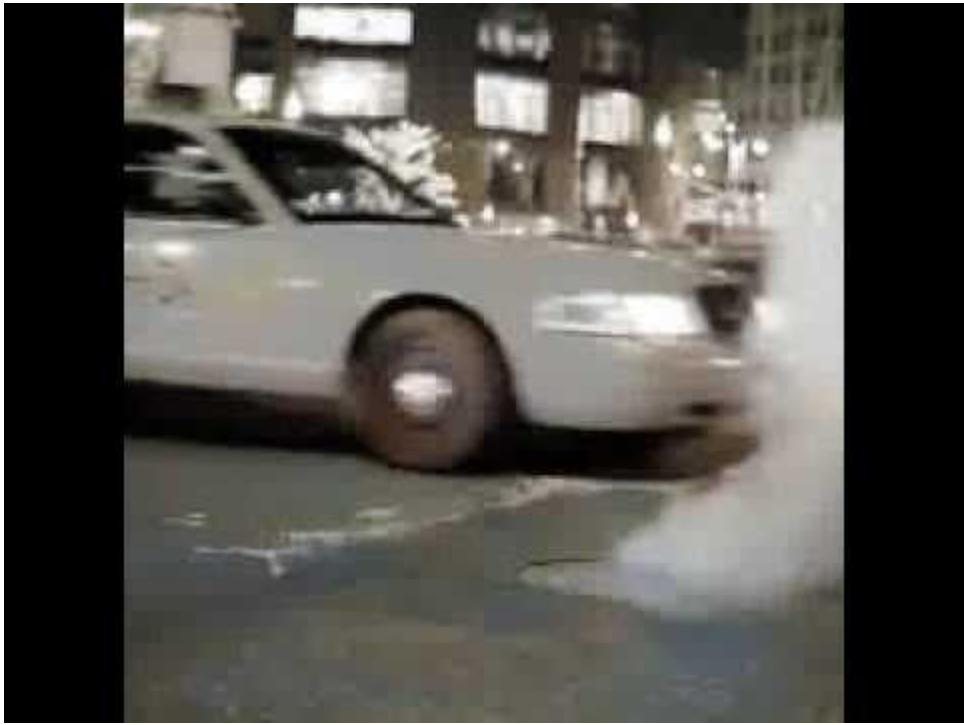
Demo



Failure Cases



Sepia & color averaging



Man-made objects &
dynamic-heavy scene

Future work

- Express post-processing and frame localization in the objective
- Experiment with a deeper and more modern pre-trained network
- Build a dedicated app - finish the imageboard