Coupling Variational Method with CNN for Image Colorization

Fabien Pierre.
University of Lorraine (France), LORIA, INRIA team MAGRIT.

Variational methods and optimization in imaging.
To the memory of our dear friend and colleague Mila Nikolova
2019.02.04

Joint work with :
Marie-Odile Berger and Thomas Mouzon.
General problem of colorization

Input.

Output.
The YUV color space

Definition of the gray-scale channel from RGB:

\[ Y = 0.299R + 0.587G + 0.114B. \]

Chrominance channel:

- \( U \) and \( V \), enable to recover the RGB image;
- invertible linear map between \( YUV \) and \( RGB \).

Challenge.

Recovering an RGB image from the luminance channel alone is an ill-posed problem and requires additional chrominance information.
The manual colorization

Two approaches:

- fully manual (polygonal masks);
- automatic diffusion.

Input  Levin et al. SIGGRAPH 2004.  Colorization with masks
The manual colorization
Exemplar-based colorization

Source

Target

Research of the "closest" patch.

Extract the color.

Welsch et al. 2002.
Exemplar-based colorization

Gupta et al. 2012.
Exemplar-based colorization

## Image database-based colorization

<table>
<thead>
<tr>
<th>Authors</th>
<th>Color space</th>
<th>Structure of the network</th>
<th>Data base</th>
<th>Cost function</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. Larsson et al. 2006</td>
<td>HCL</td>
<td>VGG</td>
<td>ImageNet</td>
<td>Cross Entropy</td>
</tr>
<tr>
<td>R. Zhang et al. 2016</td>
<td>Lab</td>
<td>VGG</td>
<td>ImageNet</td>
<td>Cross Entropy</td>
</tr>
<tr>
<td>S. Iizuka et al. 2016</td>
<td>Lab</td>
<td>U-Net + Classifier</td>
<td>MIT Places</td>
<td>L2 + Cross Entropy</td>
</tr>
<tr>
<td>S. Guadarrama et al. 2017</td>
<td>YCbCr</td>
<td>Pixel CNN</td>
<td>ImageNet</td>
<td>L1 + Cross Entropy</td>
</tr>
<tr>
<td>Y. Cao et al. 2017</td>
<td>YUV</td>
<td>cGAN</td>
<td>LSUN</td>
<td>Adv</td>
</tr>
<tr>
<td>A. Royer et al. 2017</td>
<td>Lab</td>
<td>Pixel CNN</td>
<td>ImageNet CIFAR10</td>
<td>Cross Entropy</td>
</tr>
<tr>
<td>A. Deshpande et al. 2017</td>
<td>Lab</td>
<td>VAE + MDN</td>
<td>ImageNet</td>
<td>MD</td>
</tr>
<tr>
<td>F. Baldassarre et al. 2017</td>
<td>Lab</td>
<td>U-Net + Classifier</td>
<td>ImageNet</td>
<td>L2</td>
</tr>
<tr>
<td>R. Zhang et al. 2017</td>
<td>Lab</td>
<td>U-Net + LHN + GHN</td>
<td>ImageNet</td>
<td>Huber Loss</td>
</tr>
<tr>
<td>P. Isola et al. 2017</td>
<td>Lab</td>
<td>cGAN</td>
<td>ImageNet</td>
<td>L1</td>
</tr>
<tr>
<td>Y. Xiao et al. 2018</td>
<td>Lab</td>
<td>U-Net + LHN + GHN</td>
<td>MIT Places</td>
<td>Huber Loss</td>
</tr>
<tr>
<td>Z. Su et al. 2018</td>
<td>YUV</td>
<td>VGG</td>
<td>MIT Places</td>
<td>L1 + L2</td>
</tr>
<tr>
<td>K. Nazeri et al. 2018</td>
<td>Lab</td>
<td>cGAN</td>
<td>MIT Places CIFAR10</td>
<td>L1</td>
</tr>
</tbody>
</table>
Lack of regularization with CNN

Target (input)  Result of Zhang et al., 2016  Our model

Limitation of Zhang et al., 2016.
- halo effects;
- mixing of colors.

Based on a variational model, our method is able to remove such artifacts.
Naive approach

CNN method
Zhang et al, 2016

TV-L2 color
post-processing

Too simple!!
This CNN computes color distribution on each pixel.
Set of the “labels”.

$RGB(a,b|L=50)$
Definition of annealed-mean

\[ w_i^* = \frac{\exp(\log(w_i)/T)}{\sum_j \exp(\log(w_j)/T)} \]
A CNN computes color distribution on each pixel that feeds a variational method.
Coupled total variation

Inspired of Pierre et al. 2015 SIAM journal of Imaging Sciences.

Color regularization.

\[ \hat{u} = (\hat{U}, \hat{V}) = \arg\min_{(U, V)} TV_{Y_{\text{data}}}(U, V) + \alpha \int_{\Omega} |U(x) - U_{\text{data}}(x)|^2 + |V(x) - V_{\text{data}}(x)|^2 \, dx, \]

with

\[ TV_{Y_{\text{data}}}(U, V) := \int_{\Omega} \sqrt{\gamma |\nabla Y_{\text{data}}|^2 + |\nabla U|^2 + |\nabla V|^2} \, dx. \]
1D interpretation

\[ \text{TV}_{Y_{\text{data}}} = \sqrt{\gamma a^2 + b^2} \leq \sqrt{\gamma a^2 + \sqrt{b^2}} \]
\[ \hat{u} = (\hat{U}, \hat{V}) = \arg\min_{(U, V)} TV_{Y_{\text{data}}}(U, V) + \alpha \int_{\Omega} M \left( |U(x) - U_{\text{data}}(x)|^2 + |V(x) - V_{\text{data}}(x)|^2 \right) \, dx, \]

with

\[ TV_{Y_{\text{data}}}(U, V) := \int_{\Omega} \sqrt{\gamma |\nabla Y_{\text{data}}|^2 + |\nabla U|^2 + |\nabla V|^2} \, dx. \]

\( M \) a mask, and \((U_{\text{data}}, V_{\text{data}})\) some color scribbles given by the user.

Scribbles  No coupling.  With coupling.
Intuition about coupling

Parameter influence.

$\gamma$ small: chrominance contours have low perimeters.
Let us minimize the following functional with respect to \((u, w)\):

\[
F(u, w) := TV_\varepsilon(u) + \frac{\lambda}{2} \int_\Omega \sum_{i=1}^C w_i \| u(x) - c_i(x) \|_2^2 \, dx \\
+ \chi_R(u(x)) + \chi_\Delta(w(x)).
\]

The central part of this model is based on the term

\[
\int_\Omega \sum_{i=1}^C w_i(x) \| u(x) - c_i(x) \|_2^2 \, dx.
\]
Assume that $u^*$ is a uniform real-valued random variable over the set $[0, 255]^2$. Let us denote $E$ the canonical basis of $\mathbb{R}^C$. The set of minimizers of

$$\int_\Omega \sum_{i=1}^C w_i \| u^* - c_i \|_2^2 + \chi \Delta(w)$$

is reduced to a point $w^*(u^*)$ almost everywhere (a.e.). Moreover, the one of:

$$\int_\Omega \sum_{i=1}^C w_i \| u^* - c_i \|_2^2 + \chi E(w)$$

is reduced to a point $w^{**}(u^*)$ a.e.. When these two minimizers are unique then $w^{**}(u^*) = w^*(u^*)$. 
Minimization algorithm

Primal-dual algorithm inspired by Chambolle and Pock 2011.

1: \[ u^0 \leftarrow \sum_{i=1}^{C} w_0^{n+1} c_i \]
2: \textbf{for } n > 0 \textbf{ do}
3: \[ p^{n+1} \leftarrow P_B (p^n + \sigma \nabla u^n) \]
4: \[ w^{n+1} \leftarrow P_\Delta (w^n - \rho \lambda (\|u^{n+1} - c_i\|_2^2) i) \]
5: \[ u^{n+1} \leftarrow P_R \left( u^n + \tau \left( \text{div}(p^{n+1}) + \lambda \sum_{i=1}^{C} w_i^{n+1} c_i \right) \right) \]
6: \[ \overline{u}^{n+1} \leftarrow 2u^{n+1} - u^n \]

Parameters \( \rho, \tau \) and \( \sigma \) are the time steps.

No proof of convergence
Regularize the regularizer

Introducing some regularity for the total variation:

\[
TV_{Y_{\text{data}}}(U, V) := \int_{\Omega} \sqrt{\max\{1, \gamma |\nabla Y_{\text{data}}|^2\} + |\nabla U|^2 + |\nabla V|^2} \, dx.
\]
Minimization algorithm

Inertial Bregman-based proximal gradient descent for image colorization

1: \( u^0 \leftarrow \sum_{i=1}^C w_0^{n+1} c_i \)
2: \( u^1 \leftarrow u^0 \)
3: \textbf{for } n > 0 \textbf{ do}
4: \( \bar{u}^n \leftarrow 2u^n - u^{n-1} \)
5: \( \hat{u}^n \leftarrow 2u^n - u^{n-1} \)
6: \( u^{n+1} \leftarrow P_{\mathcal{R}} \left( \frac{\hat{u}^n - \tau \nabla TV_C(\bar{u}^n) + \tau \lambda \sum_{i=1}^C w_i^{n+1} c_i}{1 + \tau \lambda} \right) \)
7: \( w^{n+1} \leftarrow \frac{w_i^n \exp \left( -\sigma \lambda \sum_{j=1}^C \|u^{n+1} - c_j\|_2^2 \right)}{\sum_{i=1}^C w_i^n \exp \left( -\sigma \lambda \sum_{j=1}^C \|u^{n+1} - c_j\|_2^2 \right)} \)

- Convergence guaranteed.
- No need of projection onto simplex.
Energy comparison

Energy
(yellow : ASAP, Primal-dual : red)

Zoom

Weights, n=500 (ASAP)

Weights, n=500 (Primal-dual)
Results of Zhang et al. 2016 vs our (histogram of the saturation).

Average of saturation: our=0.4228; Zhang et al.=0.3802.
Halo removal

Target (input)  Result of Zhang et al. 2016  Our model

Toy example

- proof of concept;
- ability to remove the halo effects of Zhang et al. 2016.
Target (input)  
Result of Zhang et al. 2016  
Our model
Limitation.

The results depend on the database.

Zhang et al. ECCV 2016, ImageNet (1.3 millions images)
Larsson et al. ECCV 2016, ImageNet
Iizuka et al. SIGGRAPH, 2016, Places (2.5 millions images)
Conclusion and future works:

**Conclusion:**
- System able to colorize images without user intervention;
- Coupling of CNN and variational model.

**Further improvement:**
- Convergence for standard total variation with primal-dual approach and biconvex functions;
- Debiasing of the results.
Coupling Variational Method with CNN for Image Colorization

Fabien Pierre.
University of Lorraine (France), LORIA, INRIA team MAGRIT.

Variational methods and optimization in imaging.
To the memory of our dear friend and colleague Mila Nikolova
2019.02.04

Many thanks for your attention.