Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence

Colorization by Matrix Completion

Shusen Wang and **Zhihua Zhang**

College of Computer Science and Technology, Zhejiang University, China
{wssatzju, zhzhang}@gmail.com

Abstract

Given a monochrome image and some manually labeled pixels, the colorization problem is a computer-assisted process of adding color to the monochrome image. This paper proposes a novel approach to the colorization problem by formulating it as a matrix completion problem. In particular, taking a monochrome image and parts of the color pixels (labels) as inputs, we develop a robust colorization model and resort to an augmented Lagrange multiplier algorithm for solving the model. Our approach is based on the fact that a matrix can be represented as a low-rank matrix plus a sparse matrix. Our approach is effective because it is able to handle the potential noises in the monochrome image and outliers in the labels. To improve the performance of our method, we further incorporate a so-called local-color-consistency idea into our method. Empirical results on real data sets are encouraging.

Introduction

For technical reasons, old photos and films are all monochrome, and it is of great interest to colorize those monochrome images and films. Computer assisted colorization has become an important application of artificial intelligence and has been widely applied to free technicians from manual colorization. Many methods have been proposed for the colorization problem in the literature (Horiuchi 2002; Levin, Lischinski, and Weiss 2004; Yatziv and Sapiro 2006; Luan et al. 2007).

One seminal work is the optimization method of Levin, Lischinski, and Weiss (2004). The key idea is based on the assumption that neighboring pixels have similar colors if their intensities are similar. As a result, the colors of unlabeled pixels are estimated by minimizing the difference from the weighted average of the colors at the neighboring pixels. The monochrome pixels are the observations, some of which are labeled with colors and the rest are unlabeled. The task is to learn a function which predicts colors (labels) for the unlabeled pixels. This optimization method uses both labeled and unlabeled pixels for training, thus enjoys semi-supervised learning mechanism (Cheng and Vishwanathan 2007).

However, the local-color-consistency assumption makes the method of Levin, Lischinski, and Weiss (2004) have two major limitations. First, colors are sometimes not local consistent, such as in some complex textures. Second, this local-color-consistency assumption requires each similar-color patch has at least one labeled pixel. Unfortunately, since similar-color patches are sometimes very small, there are numerous such patches, which makes it hard to guarantee each patch to include one labeled pixel.

(a) Monochrome (b) Labels (c) Recovered

Figure 1: Colorization using our method (Low-rank+Local-color-consistency) with 1% pixels labeled with colors.

In this paper we propose a new semi-supervised learning method for tackling the colorization problem. Our work is motivated by the recent advances of matrix recovery and its extensions. Matrix recovery is a class of problems of restoring a matrix corrupted by noises and outliers or a matrix with missing entries. Rank minimization plays a central role in matrix recovery techniques (Candès and Recht 2009; Cai, Candès, and Shen 2010; Mazumder, Hastie, and Tibshirani 2010). In practical applications, as a convex surrogate of the matrix rank, the nuclear norm is typically employed to deal with the NP-hard problem of rank minimization. Recently, Candès et al. (2011) proved that an abitrary matrix can be represented as a low-rank matrix plus a sparse matrix. Accordingly, they proposed the robust principal component analysis (RPCA) model.

Owing to the strong theories and tractable computations, matrix recovery has received wide applications in collaborative filtering (Candès and Recht 2009; Cai, Candès, and Shen 2010), background modeling (Candès et al. 2011), subspace clustering (Liu, Lin, and Yu 2010), image alignment (Peng et al. 2010), camera calibration (Zhang, Matsushita,

and Ma 2011), multi-label image classification (Cabral et al. 2011), etc. To the best of our knowledge, however, matrix recovery has not yet been applied to colorization problem.

Typically, each color image can be represented by a matrix of the three color components. We thus seek to formulate the colorization problem as a matrix completion problem (a kind of matrix recovery problems). Roughly speaking, given a proportion of observed entries from each color component (i.e. labels) and the weighted sum of the three components (the monochrome), colorization amounts to recovering the matrix from the observations. In particular, our work offers several contributions as follows.

1. Our work is the first to formulate colorization as a matrix completion problem. On one hand, this enables us to apply recent advances in matrix completion to the colorization problem. On the other hand, our study brings some new insight for the matrix completion problem.
2. Our approach is reasonable because it is based on the fact that any natural image can be effectively approximated by a low-rank matrix plus a sparse matrix (Cai, Candès, and Shen 2010; Candès et al. 2011). And some recent developments (Candès and Recht 2009; Candès et al. 2011) have even shown that low-rank matrices can be recovered exactly from a small number of sampled entries under some assumptions.
3. We develop a robust formulation which can handle the noises in the monochrome image as well as the outliers in the labels. Moreover, we devise an augmented Lagrange multiplier (ALM) algorithm for solving the model. This algorithm is very efficient; in all our experiments it performs less than 50 times singular value decompositions even under extremely strong convergence criteria.
4. Finally, we show that the local-color-consistency idea can be incorporated into our method, which further improves the performance.

The rest of the paper is organized as follows. We first formulate the colorization problem as a matrix completion problem and seek to solve it in a regularized rank-minimization approach. We relax the rank-minimization problem into a convex nuclear norm minimization problem, providing a robust model as well as an algorithm for solving the model. Then we propose to combine the low-rank and local-color-consistency methods to improve colorization. Finally, we empirically demonstrate the performance of our methods.

Problem Formulation

First of all, we give some notations that will be used in our paper. For a matrix $\mathbf{A} = [A_{ij}] \in \mathbb{R}^{m\times n}$, let $\|\mathbf{A}\|_0$ be the ℓ_0-norm (i.e. the number of nonzero entries of $\mathbf{A}$), $\|\mathbf{A}\|_1 = \sum_{i,j} |A_{ij}|$ be the ℓ_1-norm, $\|\mathbf{A}\|_F = (\sum_{i,j} A_{ij}^2)^{1/2}$ be Frobenius norm, $\|\mathbf{A}\|_\infty = \max_{i,j} |A_{ij}|$, and $\|\mathbf{A}\|_* = \sum_{i=1}^{r} \sigma_i(\mathbf{A})$ be the nuclear norm where $r = \min\{m, n\}$ and $\sigma_i(\mathbf{A})$ is the i-th largest singular value of $\mathbf{A}$. Additionally, let $\mathbf{A} \circ \mathbf{B}$ be the Hadamard product of $\mathbf{A}$ and $\mathbf{B}$, i.e. $\mathbf{A} \circ \mathbf{B} = [A_{ij}B_{ij}]$. Finally, let $\mathbf{I}_m$ denote the $m \times m$ identity matrix.

In this paper we consider RGB color images. Suppose the color image is of size $m\times n$ and has three color components, that is, red $\tilde{\mathbf{R}}$, green $\tilde{\mathbf{G}}$, and blue $\tilde{\mathbf{B}}$, all of size $m \times n$. The corresponding monochrome image is the weighted sum of the three components. There are two kinds of widely used monochrome images: the average $\mathbf{W} = \frac{1}{3}(\tilde{\mathbf{R}} + \tilde{\mathbf{G}} + \tilde{\mathbf{B}})$ and the luminosity $\mathbf{W} = 0.21\tilde{\mathbf{R}} + 0.71\tilde{\mathbf{G}} + 0.07\tilde{\mathbf{B}}$. Without loss of generality, we use the average in our experiments.

Finally, the colorization problem is formally defined as follows.

Definition 1. *Suppose we are given a monochrome image $\mathbf{W} \in \mathbb{R}^{m\times n}$, a partially observed color image $\mathbf{D} \in \mathbb{R}^{m\times 3n}$, and a zero-one matrix $\mathbf{\Omega} \in \{0, 1\}^{m\times 3n}$, where $\Omega_{ij} = 1$ indicates D_{ij} is observed and $\Omega_{ij} = 0$ otherwise. The colorization problem is to obtain the three color components $\mathbf{R}, \mathbf{G}, \mathbf{B} \in \mathbb{R}^{m\times n}$ which best approximate the underlying $\tilde{\mathbf{R}}, \tilde{\mathbf{G}}, \tilde{\mathbf{B}}$, respectively.*

Methodology

Our methodology is motivated by the fact that a natural image matrix can be represented as sum of a low-rank matrix and a sparse matrix. Thus it is intuitive to recover a low-rank matrix from the observations. From this point of view, an optimization problem for colorization is naturally formulated as follows.

Let $\mathbf{R}$, $\mathbf{G}$ and $\mathbf{B} \in \mathbb{R}^{m\times n}$ be the three color components we would like to recover. By stacking them horizontally we form $\mathbf{L} = [\mathbf{R}, \mathbf{G}, \mathbf{B}] \in \mathbb{R}^{m\times 3n}$. Let $\mathbf{S} \in \mathbb{R}^{m\times 3n}$ denote the noises in the labels, $\mathbf{W} \in \mathbb{R}^{m\times n}$ the monochrome image, $\mathbf{D} \in \mathbb{R}^{m\times 3n}$ the labels, and $\mathbf{\Omega} \in \{0, 1\}^{m\times 3n}$ the indices of the observed entries. Assuming that $\mathbf{L}$ is of low-rank and $\mathbf{S}$ is sparse, we have

$$\begin{aligned} \min_{\mathbf{R},\mathbf{G},\mathbf{B},\mathbf{L},\mathbf{S}} \quad & \operatorname{rank}(\mathbf{L}) + \lambda\|\mathbf{\Omega} \circ \mathbf{S}\|_0; \\ \text{s.t.} \quad & \mathbf{L} = [\mathbf{R}, \mathbf{G}, \mathbf{B}]; \\ & \mathbf{L} + \mathbf{S} = \mathbf{D}; \\ & \alpha_1\mathbf{R} + \alpha_2\mathbf{G} + \alpha_3\mathbf{B} = \mathbf{W}, \end{aligned} \tag{1}$$

where $\alpha_1 = \alpha_2 = \alpha_3 = \frac{1}{3}$ for the average monochrome image and $\alpha_1 = 0.21$, $\alpha_2 = 0.71$, $\alpha_3 = 0.07$ for the luminosity. The third constraint is equivalent to $\mathbf{LT} = \mathbf{W}$ where

$$\mathbf{T} = [\alpha_1\mathbf{I}_n, \alpha_2\mathbf{I}_n, \alpha_3\mathbf{I}_n]^T.$$

With these notations, Problem 1 can be equivalently expressed as follows.

$$\min_{\mathbf{L},\mathbf{S}} \operatorname{rank}(\mathbf{L}) + \lambda\|\mathbf{\Omega} \circ \mathbf{S}\|_0; \text{ s.t. } \mathbf{L} + \mathbf{S} = \mathbf{D}; \mathbf{LT} = \mathbf{W}. \tag{2}$$

Since the regularized rank-minimization problem 2 is NP-hard, in most matrix recovery problems the nuclear norm and ℓ_1-norm are often used as surrogates of matrix rank and the ℓ_0-norm, respectively. We relax Problem 2 into Problems 3 of

$$\min_{\mathbf{L},\mathbf{S}} \|\mathbf{L}\|_* + \lambda\|\mathbf{\Omega} \circ \mathbf{S}\|_1; \quad \text{s.t. } \mathbf{L} + \mathbf{S} = \mathbf{D}; \mathbf{LT} = \mathbf{W}. \tag{3}$$

Taking into account the potential noises in the input monochrome image $\mathbf{W}$, we formulate Problem 3 in a more

Figure 2: Combine low-rank method and local-color-consistency method.

Algorithm 1 The ALM Algorithm

Input: $\mathbf{W}$, $\mathbf{D}$ and $\mathbf{\Omega}$, parameters λ and η.
$j = 0$; $\mathbf{Y}_1^{(0)} = \mathbf{Y}_2^{(0)} = \mathrm{sgn}(\mathbf{D}) / \max\{\|\mathbf{D}\|_2, \|\mathbf{D}\|_\infty / \lambda\}$; $\mathbf{L}^{(0)} = \mathbf{0}$; $\mathbf{S}^{(0)} = \mathbf{0}$; $\mathbf{X}^{(0)} = \mathbf{0}$; $\mu_1^{(0)} > 0$; $\mu_2^{(0)} > 0$; $\rho > 1$; $\mathbf{T} = [\alpha_1 \mathbf{I}_n, \alpha_2 \mathbf{I}_n, \alpha_3 \mathbf{I}_n]^T$;
repeat
  $\mathbf{Z}_\mathrm{L} \leftarrow \frac{\left(\mathbf{Y}_1^{(j)} + \mathbf{Y}_2^{(j)} + \mu_1^{(j)}(\mathbf{D} - \mathbf{S}^{(j)}) + \mu_2^{(j)} \mathbf{X}^{(j)}\right)}{\mu_1^{(j)} + \mu_2^{(j)}}$;
  $\mathbf{L}^{(j+1)} \leftarrow \mathcal{S}_{1/\left(\mu_1^{(j)} + \mu_2^{(j)}\right)}\left(\mathbf{Z}_\mathrm{L}\right)$;
  $\mathbf{Z}_\mathrm{S} \leftarrow \frac{1}{\mu_1^{(j)}} \mathbf{Y}_1^{(j)} + \mathbf{D}^{(j)} - \mathbf{L}^{(j+1)}$;
  $\mathbf{S}^{(j+1)} \leftarrow \mathbf{\Omega} \circ \mathcal{D}_{\lambda/\mu_1^{(j)}}\left(\mathbf{Z}_\mathrm{S}\right) + \bar{\mathbf{\Omega}} \circ \mathbf{Z}_\mathrm{S}$;
  $\frac{\partial L_\mathbf{X}}{\partial \mathbf{X}} = \mathbf{X}^{(j)}(\eta \mathbf{T}\mathbf{T}^T + \mu_2^{(j)} \mathbf{I}) - \eta \mathbf{W}\mathbf{T}^T + \mathbf{Y}_2^{(j)} - \mu_2^{(j)} \mathbf{L}^{(j+1)}$;
  $\mathbf{X}^{(j+1)} \leftarrow$ solution to $\frac{\partial L_\mathbf{X}}{\partial \mathbf{X}} = 0$;
  $\mathbf{Y}_1^{(j+1)} \leftarrow \mathbf{Y}_1^{(j)} + \mu_1^{(j)}\left(\mathbf{D} - \mathbf{L}^{(j+1)} - \mathbf{S}^{(j+1)}\right)$;
  $\mathbf{Y}_2^{(j+1)} \leftarrow \mathbf{Y}_2^{(j)} + \mu_2^{(j)}\left(\mathbf{X}^{(j+1)} - \mathbf{L}^{(j+1)}\right)$;
  $\mu_1^{(j+1)} \leftarrow \rho \mu_1^{(j)}$; $\mu_2^{(j+1)} \leftarrow \rho \mu_2^{(j)}$; $j \leftarrow j + 1$;
until convergence
Output: $\mathbf{L}^{(j+1)}$ and $\mathbf{S}^{(j+1)}$.

robust form:

$$\begin{aligned} \min_{\mathbf{L},\mathbf{S}} \quad & \|\mathbf{L}\|_* + \lambda \|\mathbf{S} \circ \mathbf{\Omega}\|_1 + \tfrac{\eta}{2} \|\mathbf{L}\mathbf{T} - \mathbf{W}\|_F^2; \\ \text{s.t.} \quad & \mathbf{L} + \mathbf{S} = \mathbf{D}. \end{aligned} \tag{4}$$

With this robust formulation, we are able to handle the data noises in the monochrome image and the outliers in the labels. Thus, we define colorization as Problem 4 and we are concerned with its solution.

It is worth mentioning that although our model has two tuning parameters, selecting the parameters is not troublesome at all. We will show later that our method is indeed not very sensitive to the parameters. In particular, we prespecify the parameters to be $\lambda = \eta = 10$ in all the experiments.

Solution

In order to solve Problem 4, a slack matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$ is introduced to decouple the terms containing $\mathbf{L}$ in the objective function. The problem is then equivalently defined as follows.

$$\begin{aligned} \min_{\mathbf{L},\mathbf{S},\mathbf{X}} \quad & \|\mathbf{L}\|_* + \lambda \|\mathbf{S} \circ \mathbf{\Omega}\|_1 + \tfrac{\eta}{2} \|\mathbf{X}\mathbf{T} - \mathbf{W}\|_F^2; \\ \text{s.t.} \quad & \mathbf{L} + \mathbf{S} = \mathbf{D}; \\ & \mathbf{L} = \mathbf{X}. \end{aligned} \tag{5}$$

We solve Problem 5 by the augmented Lagrange multiplier (ALM) algorithm (Lin et al. 2009). The corresponding augmented Lagrange function is

$$\begin{aligned} & L(\mathbf{L}, \mathbf{S}, \mathbf{X}, \mathbf{Y}_1, \mathbf{Y}_2) \\ & = \|\mathbf{L}\|_* + \lambda \|\mathbf{S} \circ \mathbf{\Omega}\|_1 + \frac{\eta}{2} \|\mathbf{X}\mathbf{T} - \mathbf{W}\|_F^2 \\ & \quad + < \mathbf{Y}_1, \mathbf{D} - \mathbf{L} - \mathbf{S} > + \frac{\mu_1}{2} \|\mathbf{D} - \mathbf{L} - \mathbf{S}\|_F^2 \\ & \quad + < \mathbf{Y}_2, \mathbf{X} - \mathbf{L} > + \frac{\mu_2}{2} \|\mathbf{X} - \mathbf{L}\|_F^2. \end{aligned} \tag{6}$$

The ALM algorithm solves Problem 5 by alternately minimizing the augmented Lagrange function w.r.t. $\mathbf{L}$, $\mathbf{S}$, $\mathbf{X}$ and maximizing w.r.t. $\mathbf{Y}_1$ and $\mathbf{Y}_2$. This procedure is shown in Algorithm 1, and the derivation is elaborated in Appendix.

In each iteration the ALM algorithm solves a regularized nuclear norm minimization problem w.r.t. $\mathbf{L}$. We resort to the singular value shrinkage operator defined in (Cai, Candès, and Shen 2010). For $\tau \geq 0$, the singular value shrinkage operator $\mathcal{S}_\tau$ is defined as

$$\begin{aligned} \left[\mathcal{D}_\tau(\mathbf{A})\right]_{ij} &= \mathrm{sgn}(A_{ij})(|A_{ij}| - \tau)_+, \\ \mathcal{S}_\tau(\mathbf{B}) &= \mathbf{U}_\mathbf{B} \mathcal{D}_\tau(\mathbf{\Sigma}_\mathbf{B}) \mathbf{V}_\mathbf{B}^T, \end{aligned} \tag{7}$$

where $\mathbf{B} = \mathbf{U}_\mathbf{B} \mathbf{\Sigma}_\mathbf{B} \mathbf{V}_\mathbf{B}^T$ is the singular value decomposition (SVD) of $\mathbf{B}$. The optimality of $\mathcal{D}_\tau$ and $\mathcal{S}_\tau$ is shown in Proposition 2 in Appendix.

Remark: The computation cost of Algorithm 1 in each iteration is dominated by the SVD in Line 5. Our off-line experiments show that the algorithm gets convergence in less than 50 iterations, i.e. performs less than 50 SVDs, for images of all small and large sizes (from 150×200 to 1500×2000 pixels).

Combining Low-rank and Local-color-consistency Methods

In the experiments we notice that our low-rank method works well when there are sufficient labeled pixels, but the performance deteriorates with the decreasing of

(a) Original (b) LR 10% (c) LL 10% (d) LR 1% (e) LL 1%

Figure 3: Figure (b) (d) are computed by the Low-Rank (LR) method and (c) (e) by Low-rank+Local-color-consistency (LL) method. The percentage denotes the proportion of labeled pixels.

(a) R 20% (b) R 30% (c) R 50% (d) LR 10% (e) LL 1%

Figure 4: Figure (a) (b) (c) are computed by RPCA (without exploiting monochrome image) with 20%, 30%, and 50% pixels labeled; Figure (d) (e) are computed respectively by our Low-Rank (LR) method and Low-rank+Local-color-consistency (LL).

labeled-pixel proportion. On the contrary, the local-color-consistency method (Levin, Lischinski, and Weiss 2004) is less sensitive to the labeled-pixel-proportion. This motivates us to combine the advantages of the low-rank method and local-color-consistency method.

Our intuition is to make use of local-color-consistency along with the robustness of our low-rank method to improve colorization ability. First, for each unlabeled pixel (i, j), we find all its neighboring labeled pixels with monochrome intensity close to pixel (i, j), and then label pixel (i, j) with the weighted sum of those labeled neighbors. The weight should be positively correlated with their intensity similarity and negatively correlated with their distance. Notice that some pixels are still unlabeled and some may be incorrectly labeled after this process; this problem can be handled by our robust low-rank method. By taking these labels as well as the initially given labels as inputs, the low-rank method can complete the colorization. This procedure is illustrated in Figure 2.

The comparison of the low-rank method and low-rank+local-color-consistency method is shown in Figure 3. The results clearly show that the low-rank+local-color-consistency method still works well even if very few labeled pixels are given, while the low-rank method fails. Since the low-rank+local-color-consistency method performs significantly better than the low-rank method, in the next sections we mainly compare our low-rank+local-color-consistency with the local-color-consistency method of Levin, Lischinski, and Weiss (2004).

Related Work

In this section we discuss the connection of our method with existing matrix completion methods. We show that the colorization problem is a matrix completion problem with an additional regularization term. With this term, the matrix completion problem can be solved much more accurately.

Candès and Recht (2009) proposed to solve the matrix completion problem via the following convex optimization model:

$$\min_{\mathbf{L}} \ \|\mathbf{L}\|_*; \qquad \text{s.t. } \mathbf{\Omega} \circ (\mathbf{L} - \mathbf{D}) = \mathbf{0}. \tag{8}$$

Later on, Candès et al. (2011) formulated a robust version of Problem 8 by assuming the potential data noises are sparse:

$$\min_{\mathbf{L}} \ \|\mathbf{L}\|_* + \lambda \|\mathbf{\Omega} \circ \mathbf{S}\|_1; \qquad \text{s.t. } \mathbf{L} + \mathbf{S} = \mathbf{D}, \tag{9}$$

which is well known as the RPCA model. With an additional regularization term $\|\mathbf{LT} - \mathbf{W}\|_F^2$, the RPCA model becomes Problem 4.

Given a monochrome image, i.e. the average of the three components in our case, we can add an extra regularization term to the matrix completion problem formulation, with which the matrix recovery accuracy is significantly improved. Figure 4 gives a comparison of matrix completion results with and without exploiting the monochrome image. The results clearly demonstrate that with the information of the monochrome image encoded in our model, the matrix completion problem is solved much more accurately.

Our work suggests that the matrix completion formulation 9 is extensible; with encoding some other knowledge than the given entries, the matrix completion accuracy can be largely improved.

Experiments

In this section we carry out a set of experiments on natural images to demonstrate the performance of our method. We also conduct comparison with the method of Levin, Lischinski, and Weiss (2004), which we will denote by *LLW* for description simplicity. The sample images that we used are shown in Figure 5. Recall that the monochrome image is defined as a weighted sum of the R, G, and B components. Without loss of generality, we average the three components to obtain the monochrome image.

The first set of experiments are conducted on all the sample images in Figure 5 to measure the image recovery accuracy. The monochrome image is obtained by averaging the R, G, and B components. We randomly hold a certain percentage of pixels (uniformly) of the original image as the observed part (labels). Using the monochrome image and the labels, we run our two methods and the LLW method for comparison. We define the relative square error (RSE) to measure the recovery accuracy:

$$RSE = \frac{\|\mathbf{L}^* - [\tilde{\mathbf{R}}, \tilde{\mathbf{G}}, \tilde{\mathbf{B}}]\|_F}{\|[\tilde{\mathbf{R}}, \tilde{\mathbf{G}}, \tilde{\mathbf{B}}]\|_F},$$

(a) (b) (c) (d) (e) (f) (g) (h)

Figure 5: Sample images from the Berkeley Segmentation Dataset (Martin et al. 2001).

(a) LL, 0.1% (b) LLW, 0.1% (c) LL, 1% (d) LLW, 1% (e) LL, 10% (f) LLW, 10%

Figure 6: Results of colorization using our Low-rank+Local-color-consistency (LL) method and the method of Levin, Lischinski, and Weiss (2004) (LLW). The percentage of each figure denotes the proportion of labeled pixels.

where $\tilde{\mathbf{R}}$, $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{B}}$ are the three color components of original image while $\mathbf{L}^*$ is the recovered image obtained by each colorization method. The RSE versus labeled-pixel proportion of each sample image is shown in Figure 7. We also present visual comparisons between our low-rank+local-color-consistency (LL) method and LLW in Figure 6. Although these two methods can recover the color given sufficient labels, a detailed look at the recovered images reveals differences in intensity, illumination, and some other details. The visual results and the RSEs all demonstrate that the images restored by our Low-rank+local-color-consistency method are closer to the original image.

In the second set of experiments, we manually label some pixels of the monochrome images with corresponding colors. We show in Figure 9 the results obtained by each method from two manually labeled images: one with few but relatively correct labels, the other with many noisy labels. From the results we can see that our two methods are more robust than the LLW method.

Finally, we demonstrate that our low-rank method is insensitive to the tuning parameters λ and η of Problem 4. We use the image in Figure 5(a) with 1%, 5%, and 10% pixels being labeled respectively as the input of our method; with one parameter fixed while the other varies, we plot the RSE against the parameter value in Figure 8. Our off-line experiments on a variety of images, small and large sizes, all demonstrated that our method is insensitive to λ and η. From 150×200 small images up to 1500×2000 large images, the results of simply setting $\lambda = \eta = 10$ are always quite near those obtained by carefully tuning the two parameters.

Future Work

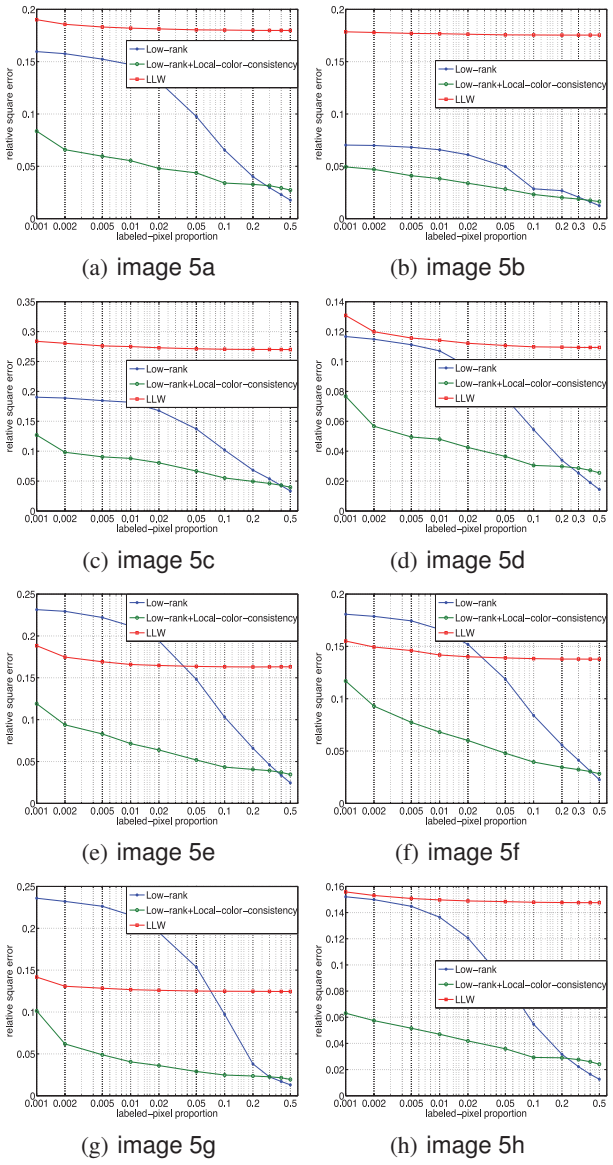
As aforementioned, our low-rank method requires sufficient labeled pixels (say more than 10%) as inputs, thus we seek to improve the performance via a preprocessing step to generate more labels as inputs for our low-rank method. Although the generated labels are potentially noisy, it does not hurt the performance because our low-rank method is very robust. We have successfully employed the local-color-consistency method of Levin, Lischinski, and Weiss (2004) to generate more labels, and the performance has been significantly escalated. Similarly, other approaches can also be incorporated in the same way. For example, we can make use of the color labeling scheme proposed by Luan et al. (2007) which groups not only neighboring pixels with similar intensity but also remote pixels with similar texture.

Moreover, an improvement can be made to deal with the generated labels. Since the generated labels are potentially more noisy, we can discriminate between the generated labels and the original labels by reformulating Problem 4 as

$$\begin{aligned} \min_{\mathbf{L},\mathbf{S}} \quad & \|\mathbf{L}\|_* + \lambda_1\|\mathbf{S}\circ\mathbf{\Omega}_1\|_1 + \lambda_2\|\mathbf{S}\circ\mathbf{\Omega}_2\|_1 \\ & + \tfrac{\eta}{2}\|\mathbf{L}\mathbf{T} - \mathbf{W}\|_F^2; \\ \text{s.t.} \quad & \mathbf{L} + \mathbf{S} = \mathbf{D}, \end{aligned} \tag{10}$$

where the original labels are indexed by $\mathbf{\Omega}_1$ and the generated labels by $\mathbf{\Omega}_2$, and $\lambda_1 > \lambda_2 > 0$.

Finally, our method can also be applied to video colorization and batch image colorization, as well as single image colorization. Such tasks can be fulfilled by stacking the video frames or a batch of images together as a single image followed by colorizing the stacked image. This approach is reasonable because it exploits the similarity among the related frames or images.



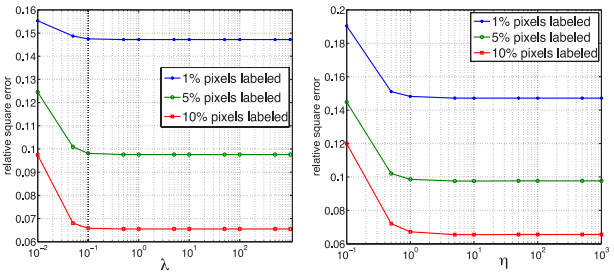
(a) image 5a (b) image 5b

(c) image 5c (d) image 5d

(e) image 5e (f) image 5f

(g) image 5g (h) image 5h

Figure 7: Relative square errors (RSE) of colorization using the three methods on the sample images in Figure 5.



(a) $\eta = 10$ (b) $\lambda = 10$

Figure 8: Relative square errors (RSE) of the low-rank method versus the tuning parameters λ and η of Problem 4.

Conclusions

In this paper we have proposed a matrix completion approach to the colorization problem. Our proposal is built on the robust principal component analysis idea. In particular, we have formulated colorization as a convex optimization problem and devised an augmented Lagrange multiplier algorithm to solve the optimization problem. Our approach is flexible because it can be combined with the local-color-consistency leading to better colorization results. The experiments have demonstrated that our method can produce better performance in comparison with the existing state-of-art method.

Acknowledgments

This work is in part supported by the Natural Science Foundations of China (No. 61070239).

References

Cabral, R.; De la Torre, F.; Costeira, J.; and Bernardino, A. 2011. Matrix completion for multi-label image classification. In *Advances in Neural Information Processing Systems*.

Cai, J.-F.; Candès, E.; and Shen, Z. 2010. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization* 20(4):1956–1982.

Candès, E., and Recht, B. 2009. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics* 9(6):717–772.

Candès, E.; Li, X.; Ma, Y.; and Wright, J. 2011. Robust principal component analysis. *Journal of the ACM* 58(3).

Cheng, L., and Vishwanathan, S. 2007. Learning to compress images and videos. In *Proceedings of the 24th international conference on Machine learning*, 161–168. ACM.

Horiuchi, T. 2002. Estimation of color for gray-level image by probabilistic relaxation. In *Proceedings of the 16 th International Conference on Pattern Recognition*.

Levin, A.; Lischinski, D.; and Weiss, Y. 2004. Colorization using optimization. In *ACM Transactions on Graphics (TOG)*, volume 23, 689–694. ACM.

Lin, Z.; Chen, M.; Wu, L.; and Ma, Y. 2009. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *UIUC Technical Report, UILU-ENG-09-2215*.

Liu, G.; Lin, Z.; and Yu, Y. 2010. Robust subspace segmentation by low-rank representation. In *The 17th International Conference on Machine Learning*.

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

Martin, D.; Fowlkes, C.; Tal, D.; and Malik, J. 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *The 8th International Conference on Computer Vision*.

(a) Monochrome (b) Manual labels (c) LLW (d) LR (e) LL

Figure 9: Colorization with manually labeled color solved by our Low-Rank (LR) method, Low-rank+Local-color-consistency (LL) method, and the method of Levin, Lischinski, and Weiss (2004) (LLW).

Mazumder, R.; Hastie, T.; and Tibshirani, R. 2010. Spectral regularization algorithms for learning large incomplete matrices. *Journal of machine learning research* 11(2):2287–2322.

Peng, Y.; Ganesh, A.; Wright, J.; Xu, W.; and Ma, Y. 2010. Rasl: Robust alignment by sparse and low-rank decomposition for linearly correlated images. In *IEEE Conference on Computer Vision and Pattern Recognition*.

Yatziv, L., and Sapiro, G. 2006. Fast image and video colorization using chrominance blending. *IEEE Transactions on Image Processing* 15(5):1120–1129.

Zhang, Z.; Matsushita, Y.; and Ma, Y. 2011. Camera calibration with lens distortion from low-rank textures. In *IEEE Conference on Computer Vision and Pattern Recognition*.

Derivation of Algorithm 1

The optimality of $\mathcal{S}_\tau$ is shown in Proposition 2, which guarantees the optimality of minimizing the augmented Lagrange function (6) w.r.t. $\mathbf{L}$.

Proposition 2. *For any $\tau \geq 0$, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, $\mathcal{D}_\tau$ and $\mathcal{S}_\tau$ defined in (7) obey*

$$\mathcal{D}_\tau(\mathbf{B}) = \operatorname*{argmin}_{\mathbf{A}} \ \tau\|\mathbf{A}\|_1 + \frac{1}{2}\|\mathbf{A}-\mathbf{B}\|_F^2, \tag{11}$$

$$\mathcal{S}_\tau(\mathbf{B}) = \operatorname*{argmin}_{\mathbf{A}} \ \tau\|\mathbf{A}\|_* + \frac{1}{2}\|\mathbf{A}-\mathbf{B}\|_F^2. \tag{12}$$

Based on Equation (11), we give a similar result in Proposition 3 which guarantees the optimality of minimizing (6) w.r.t. $\mathbf{S}$.

Proposition 3. *Given any $\tau \geq 0$, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, $\mathbf{\Omega} \in \{0,1\}^{m\times n}$. Let $\mathcal{D}_\tau$ be defined in (7) and $\bar{\mathbf{\Omega}}$ be an $m \times n$ matrix whose the (i,j)-th entry is $1-\Omega_{ij}$, then the solution to the following optimization problem*

$$\hat{\mathbf{A}} = \operatorname*{argmin}_{\mathbf{A}} \ \tau\|\mathbf{A}\circ\mathbf{\Omega}\|_1 + \frac{1}{2}\|\mathbf{A}-\mathbf{B}\|_F^2.$$

is $\hat{\mathbf{A}} = \mathbf{\Omega}\circ\mathcal{D}_\tau(\mathbf{B}) + \bar{\mathbf{\Omega}}\circ\mathbf{B}$.

The augmented Lagrange function of Problem 5 is shown in Equation 6. In the augmented Lagrange function the terms containing $\mathbf{L}$ is $L_\mathbf{L}$:

$$L_\mathbf{L} = \|\mathbf{L}\|_* + \frac{\mu_1+\mu_2}{2}\|\mathbf{L}-\mathbf{Z}_\mathbf{L}\|_F^2, \tag{13}$$

$$\mathbf{Z}_\mathbf{L} = \frac{1}{\mu_1+\mu_2}\Big(\mathbf{Y}_1+\mathbf{Y}_2+\mu_1(\mathbf{D}-\mathbf{S})+\mu_2\mathbf{X}\Big) \tag{14}$$

Optimizing $L_\mathbf{L}$ w.r.t. $\mathbf{L}$ lead to the solution $\mathcal{S}_{1/(\mu_1+\mu_2)}(\mathbf{Z}_\mathbf{L})$, as guaranteed by Proposition 2.

The terms containing $\mathbf{S}$ is $L_\mathbf{S}$:

$$L_\mathbf{S} = \lambda\|\mathbf{S}\circ\mathbf{\Omega}\|_1 + \frac{\mu_1}{2}\|\mathbf{S}-\mathbf{Z}_\mathbf{S}\|_F^2, \tag{15}$$

$$\mathbf{Z}_\mathbf{S} = \frac{1}{\mu_1}\mathbf{Y}_1 + \mathbf{D} - \mathbf{L}. \tag{16}$$

Optimizing $L_\mathbf{S}$ w.r.t. $\mathbf{S}$ lead to the solution $\mathbf{\Omega}\circ\mathcal{D}_{\lambda/\mu_1}(\mathbf{Z}_\mathbf{S}) + \bar{\mathbf{\Omega}}\circ\mathbf{Z}_\mathbf{S}$, as guaranteed by Proposition 3.

The terms containing $\mathbf{X}$ is

$$L_\mathbf{X} = \frac{\eta}{2}\|\mathbf{X}\mathbf{T}-\mathbf{B}\|_F^2 + <\mathbf{Y}_2, \mathbf{X}-\mathbf{L}> + \frac{\mu_2}{2}\|\mathbf{X}-\mathbf{L}\|_F^2.$$

The derivative of $L_\mathbf{X}$ w.r.t. $\mathbf{X}$ is

$$\frac{\partial L_\mathbf{X}}{\partial \mathbf{X}} = \mathbf{X}(\eta\mathbf{T}\mathbf{T}^T + \mu_2\mathbf{I}) - \eta\mathbf{W}\mathbf{T}^T + \mathbf{Y}_2 - \mu_2\mathbf{L},$$

setting which to zero leads to the optimal value of $\mathbf{X}$. This linear system can be solved efficiently by the conjugate gradient algorithm.