MATH 6380O Final Project: Colorizing Black-White Movies Fastly and Automatically

Chenyang Lei, Yazhou Xing, Yue Wu, and Jiaxin Xie {cleiaa, yxingag, ywudg, jxieax}@connect.ust.hk
Department of Computer Science and Engineering, HKUST

1. Introduction

Gray Input Ours

Early movies are black and white due to the limitation of film technology of the time. With the rapid growth of machine learning, colorizing black and white movies becomes available. In this project, we rethink the video colorization problem and reimplement one of the state-of-the-art colorization model. We also modify their network architecture to speed up model forward process and still achieve a comparable accuracy according to extensive evaluation on DAVIS dataset. Besides, our modification enables testing the model without optical flow data, which is hard to obtain in practice. Perceptual experiments demonstrate that our approach achieves state-of-the-art performance on fully automatic video colorization. Video is available at https://www.youtube.com/watch?v=8Nk-ITnLSXU

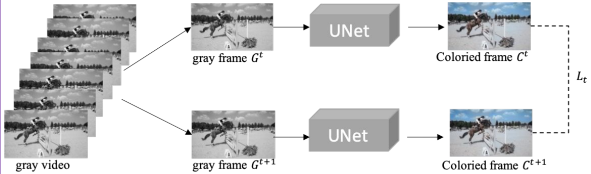
2. Method

Formulation

Given a sequence of grayscale video frames G, the objective is to automatically colorize the given frames as realistic as possible. The illustration of symbols are as follows: G is the input grayscale frame, C represent the colorized results generated by our network. W is the Warp operator to warp frames based on the optical flow F. Y represents the corresponding RGB color image ground truth.

Framework

The backbone of our method is a U-Net network structure. Given two consecutive frames, our network will output the colorized result C. The framework is shown in following figure.



Loss function

For video colorization problem, it's quite important to make the result temporal coherent and diverse. Thus, In our model, we have two loss functions. The first is temporal loss to regularize the temporal consistency.

$$L_t = \|W(C_1, F_{2\to1}) - C_2\|$$

The second is a diversity control loss aims at producing a diverse set of colorized videos . We add the minimum perceptual loss distance and weighted sum of different outputs' perceptual loss distance.

$$L_d = \min\{\|\phi(C) - \phi(Y)\|\} + \sum_d \lambda_d \|\phi(C^d) - \phi(Y^d)\|$$

3. Experiments and Analysis

Dataset

In experiments, we adopt DAVIS dataset to evaluate the performance. The test set consists of 30 videos, each video has 30~100 frames. We convert them to gray frames as input and use the RGB images as ground truth.

	Zhang et al.	Iizuka et al.	Lei et al.	Ours
PSNR	29.07	29.25	**30.35**	29.76

Results

We choose two single image colorization methods and one video colorization methods as baseline.

From Table. 1, we find that our performance outperforms two single image baseline on PSNR and achieve similar performance compared with [1]. However, our method is much faster than [1] because we don't need to compute the optical flow and refine the results again.

Input Ours Zhang et al. [2] Iizuka [3]

This figure shows two comparisons between our methods and single image methods. These two rows are the input and results for two consequent frames. we could see that the colorization of our method is much more stable than single image methods. Single image colorization methods get inconsistent colorization because they are trained single image.

4. Conclusion

In this project, we propose a CNN model to do video colorization. Through modification of our baseline model, we speed up the network forward path by 30% while still keep the comparable accuracy. In addition, our model gets rid of hard-to-obtain optical flow data during testing. Compared with state-of-the-art image colorization works, our results keep temporal coherence and multimodility.

5. References

[1] Chenyang Lei and Qifeng Chen. Fully Automatic Video Colorization with Self-Regularization and Diversity. In CVPR, 2019.
[2] R. Zhang, P. Isola, and A. A. Efros. Colorful image colorization. In ECCV, 2016
[3] S. Iizuka, E. Simo-Serra, and H. Ishikawa. Let there be Color!: Joint End-to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification. ACM Trans. Graph., 35(4), 2016.

6. Contribution

Chenyang Lei: Implementation, poster
Yazhou Xing: Implementation, poster
Jiaxin Xie: Baselines, video
Yue Wu: Poster, video