UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

MASTER THESIS num. 000

Image Colorization Methods

Adi Čaušević

Zagreb, June 2021.

Ljubljana, April 26 2021

MASTER THESIS ASSIGNMENT

Student: **Adi Čaušević**

Title: **Image Colorization Methods**

Description:

Greyscale or 8-bit images take up significantly less memory space compared to RGB images that use 24 bits. Sometimes greyscale images are the only thing available, e.g. in the case of historical photographs. Image colorization is a process in which colours are reconstructed using original grayscale image.

Within the thesis study and describe image colorization methods. Special attention should be paid to methods based on soft computing, such as the use of neural networks (today typically deep models). Implementations of at least three different state of the art methods have to be trained and evaluated. The thesis should encompass the source code and other relevant material for recreating the results. The results have to be explained, supported with arguments. The used literature has to be cited and the obtained help acknowledged.

I agree with publishing the final version of the thesis in electronic form on the website of the Faculty of Electrical Engineering and Computing, University of Zagreb, as well as in the public Croatian Digital Thesis Repository of the National and University Library.

Mentor:

Prof. Dr. Peter Peer

Marko Čupić

Assoc. Prof. Dr. Marko Čupić

Ovo je zahvala

Contents

1. Introduction

Ever since the advent of photography in the early 19th century people had the urge to add color to photographs. Not long after the invention of the daguerreotype by Louis Daguerre in 1839, the Swiss painter and later photographer Johann Baptist Isenring produced the first colored photographic image by hand-applying pigments to the surface of a monochrome photograph [1]. Although first experiments with color photography date back to the middle of the 19th century hand-coloring monochrome photographs, as seen in figure 1.1, was the easiest method to produce full-color photographic images until the arrival of color film in the mid-20th century. With the beginning of the digital age and the increase of computational power the fascination with coloring images reemerged, but this time the problem was tackled from a computational perspective.

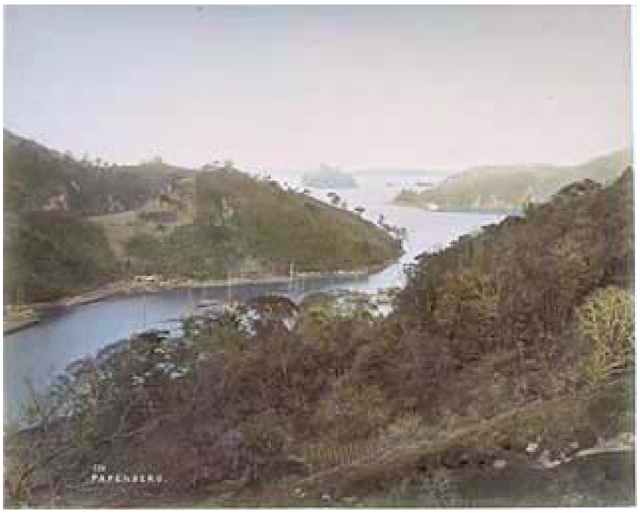

Figure 1.1: Hand-colored photograph by Stillfried & Andersen (made the period between 1862 and 1885)

At first sight, the task of automatic colorization might look impossible to accomplish, as information is lost when taking a monochrome photograph. If our goal is to perfectly colorize an image that assumption is valid most of the time because we encounter the problem of multimodality that is described in section **??**. If, on the other hand, we are satisfied with a plausible colorization, the task becomes achievable as there is still a lot of semantic information ingrained in the grayscale image.

(a) Original **(b)** Grayscale **(c)** Colorized

Figure 1.2: Colorization example on an image of a stone house

For example, if our colorizer is given the grayscale image in figure 1.2 it can assume that the vegetation surrounding the house is green, while the stone walls are probably some shade of grey. All that information is available no matter that there is no color in the image, but the colorizer should be able to deduce a likely color from the shapes and textures present.

That said, in this thesis we will explore the topic of automatic image colorization. Firstly the task of image colorization will be described as well other notable details important for the understanding of that topic. After the basic are covered

2. Methods

The task of automatic colorization can be defined as an image-to-image translation problem [3], where for a monochrome input image we have to propose a plausible colored output. Although grayscale is a subset of monochrome (meaning *having one color*), the two terms will be used interchangeably throughout this thesis. The defining feature of the input images is that they only have one channel representing the perceptual lightness of the given image. That means that each pixel only has a numeric value, usually in the 8-bit range, expressing how light (or dark) it is.

Figure 2.1: Colored image of Alim Khan taken by Sergey Prokudin-Gorsky in 1911 using the three-image color photography method on the left, and the green, red, and blue (top to bottom) filter negatives (shown as positives) on the right.

On the other hand, if we want to render a colored photograph we need at least 3 channels to do, meaning that for each pixel we want to draw, we need at least 3 values associated with it. The specific way of mapping the three values to a color is called a color space. An example of a color space is the RGB color model in which each of the 3 channels depicts the amount of **R**ed, **G**reen, or **B**lue color present, as shown in figure 2.1.

That said, the task of colorization can be summed up as

$$\mathbf{Y} = \mathcal{F}(\mathbf{X})$$

where $\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$ is the input and $\mathbf{Y} \in \mathbb{R}^{H \times W \times 3}$ is the colored output. $\mathcal{F}$ is the function that maps the domain $\mathbf{X}$ to the codomain $\mathbf{Y}$, which in our case represents the colorizer.

2.1. Color space

Even though automatic colorization can be accomplished using any color space, as with any task some tools are better for the job than others. Although the RGB color space would be the easiest to use, as it is the default option for many tools, it doesn't give us an advantage when colorizing. On the other hand, the CIELAB color space has a few useful properties which make it convenient for the task. The CIELAB (or L*a*b*) color space was defined by the International Commission on Illumination in 1976 as an attempt to create a perceptually uniform color space [2].

In perceptually uniform color spaces, a numeric change in any direction would result in a similar perceived color change. Although L*a*b* is not truly perceptually uniform, it is nevertheless useful for that purpose. Perceptual uniformity is not necessary while colorizing, but it is practical. As small changes in value result in small color changes.

- What L*a*b* mean - Plots of lab colorspace - Not as effective for image compression as only a third of the box is used - Quintization errors with ints, best to use 32 bit floats to repersent the colorspace - Covers the whole visible color spectrum - Created as perceptually coherent color space - Not entirely successful at it, but quite successful - Why use LAB colorspace - Good for this task because of: - Perceptual distance close to numerical distance (good at assesing what humans would percieve color differences) - Input diemnsion is grayscale which matches the L channel of Lab colorspace, we only need to predict 2 channels a* and b*

As we are working with images the choice of colorspace is quite important Covers the while human visible spectrum Perceptually close colors are

thats why rather than using the rgb colorspace ... La*b*

-

2.2. Task definition

- That said - The task of colorization is: get ab color channels from the luminance value of the photo image-to-image translation problem where the domain is the grayscale L channel and codomain are the ab color channels Trying to imagine the probable values for the a and b (color) channels of the CIELab colorspace givent the luminance values of a grayscale image Colored images have 3 channels, one channel is already available, so the colorizer has to hallucinate

- It is impossible to perfectly recover the original image as too much information has been lost due to the colorspace - The best the model can do is to offer a plausible colorization - Trying to imagine 2 color dimensions of the lab color space

- Image colorization definition

- The baseline knowledge that is needed to understand - Approaches to the problem - Dataset and dataset preparation (possible improvements like adding noise) - describing

2.2.1. Multimodality

2.3. Dataset

- The model is only as good as the dataset

3. Results

4. Related work

4.1. Related work

- Beginning statistical models - Convolutional networks - Generative Adversarial networks - In the last few years novel approaches: Transformers, cINN - Specialized approaches for different use cases

5. Conclusion

Conclusion.

5.1. Possible improvements

- Add noise to input images when training - Better way to filter dataset -

Bibliography

[1] Heinz K Henisch and Bridget Ann Henisch. *The painted photograph, 1839-1914 : origins, techniques, aspirations*. Pennsylvania State University Press, 1996.

[2] *Image technology colour management — Architecture, profile format, and data structure*. International Color Consortium, 2004. URL `https://www.color.org/icc1v42.pdf`. Profile version 4.2.0.0.

[3] Yingxue Pang, Jianxin Lin, Tao Qin, and Zhibo Chen. Image-to-image translation: Methods and applications, 2021.

Image Colorization Methods

Abstract

This is the abstract

Keywords: Keywords.

Naslov

Sažetak

Sažetak na hrvatskom jeziku.

Ključne riječi: Ključne riječi, odvojene zarezima.