ISSN 2464-4617 (print)
ISSN 2464-4625 (online)
Computer Science Research Notes - CSRN 3301
http://www.wscg.eu
WSCG 2023 Proceedings

On Unguided Automatic Colorization of Monochrome Images

Andrzej Śluzek[0000-0003-4148-2600]

Warsaw University of Life Sciences - SGGW
Institute of Information Technology
ul. Nowoursynowska 159, Bld. 34
02-776, Warsaw, Poland

andrzej_sluzek@sggw.edu.pl

ABSTRACT

Image colorization is a challenging problem due to the infinite RGB solutions for a grayscale picture. Therefore, human assistance, either directly or indirectly, is essential for achieving visually plausible colorization. This paper aims to perform colorization using only a grayscale image as the data source, *without* any reliance on metadata or human hints. The method assumes an (arbitrary) *rgb*2*gray* model and utilizes a few simple heuristics. Despite probabilistic elements, the results are visually acceptable and repeatable, making this approach feasible (e.g. for aesthetic purposes) in domains where only monochrome visual representations exist. The paper explains the method, presents exemplary results, and discusses a few supplementary issues.

Keywords

Image colorization, decolorization, *rgb*2*gray* models, visual plausibility, color models.

1. INTRODUCTION & MOTIVATION

Image colorization, i.e. reconstructing color images from monochrome ones, is an ill-posed problem due to the infinite number of RGB solutions for a grayscale picture. Nonetheless, this topic holds notable practical and commercial significance, particularly in the restoration of historical photos and movies being the primary application, e.g. [Zeg21], [Sal22].

In the past two decades, many papers have proposed diverse algorithms for reconstructing color images or movies from their monochrome counterparts. Initially, the methods were mainly semi-automatic. Users would provide exemplary images to guide the algorithm in coloring images with similar contents and contexts, primarily using similarly textured patches, e.g. [Iro05] and [Gup12]. Alternatively, monochrome images can be manually "scribbled" to indicate approximate colors over a number of significant locations, e.g. [Lev04], [Lag08].

More recently, advanced machine learning has enabled fully automated image colorization, with coloring patterns learned from relevant images rather than human-provided hints.

.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Typically, the patterns are derived from images of specific domains, e.g. [Des15], [Hwa16], [Zha16].

A more comprehensive system is described in [Iiz16], which learns scene recognition, local priors, and mid-level features from nearly 2.5 million training images. The results are impressive on test images of scenes (if their semantics are correctly recognized). Existing commercial systems (e.g. [Sal22]) generally follow the same concepts.

In an alternative approach, machine learning can be used to identify colorization statistics (instead of direct coloring), as in [Des15] and [Roy17]. Automatic image colorization across multiple domains (transfer learning) is more challenging, and [Lee22] is the first work with limited but convincing results.

In summary, all the methods mentioned above (and many other approaches not discussed here) are assisted by humans, either by providing colorization hints or relevant training data for ML algorithms. Therefore, the proposed objective of this paper seems slightly audacious (if not impossible). Our intention is to develop a mechanism for unguided automatic image colorization *without* additional metadata, assistance, learning processes, or domain identification. In other words, we aim to create an acceptable colored counterpart using only a grayscale image as the data source. By "acceptable," we mean visually attractive results that are statistically repeatable and deliver convincingly rich sensations of colors (excluding pseudo-coloring, as in thermographic cameras).

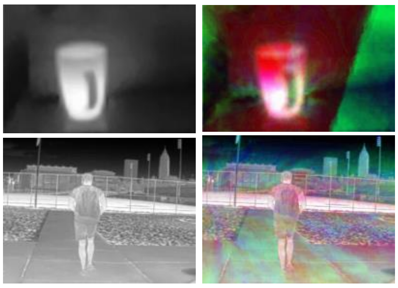


Figure 1. Examples of visually convincing colorizations of infrared images.

In domains of visual-frequency grayscale images, such a problem is rather marginal, but there are areas where only single-channel visualizations actually exist, such as IR/UV/US/MRI/X-ray images. In such cases, we would like to see hypothetical RGB versions of those "gray worlds" for various reasons, even if it is only for aesthetic purposes (see Fig. 1).

Section 2 of the paper discusses a number of assumptions and models adopted in the proposed solution. Further implementation details of the developed algorithms are included in Section 3. Section 4 presents diverse examples of obtained results, with corresponding explanations. The final Section 5 contains conclusions, discusses some supplementary issues, and highlights directions for future work.

2. PROPOSED METHODOLOGY

Using *rgb2gray* models in colorization

Colorization methods generally assume that grayscale image values represent the luminance channel of the colored outputs, requiring reconstruction of only two chrominance channels. However, few papers on image colorization consider the opposite question: *how the original RGB image (real or hypothetical) was decolorized to obtain a grayscale image.*

Standard RGB-to-grayscale models (YUV and YIQ) apply linear functions of primary colors:

$$Y = k_R R + k_G G + k_B B \tag{1}$$

where $k_R = 0.299$, $k_G = 0.587$ and $k_B = 0.114$ (or $k_R = 0.2126$, $k_G = 0.7152$, $k_B = 0.0722$).

Other *rgb2gray* models with arbitrarily assumed k_R, k_G and k_B coefficients (subject to $k_R + k_G + k_B = 1$) produce alternative monochrome images (see Figs 2b,c,d) from which various re-colorizations can be hypothetically reconstructed (Figs 2e,f,g).

Therefore, in the proposed colorization scheme we first assume that:

Monochrome images are derived from (real or hypothetical) color images by an *rgb2gray* model with arbitrarily assumed k_R, k_G and k_B coefficients.

Such an assumption is justified because for problems with only hypothetical existence of color images (as in Fig.1), any *rgb2gray* model can be assumed, as long as the colorization results are visually appealing.

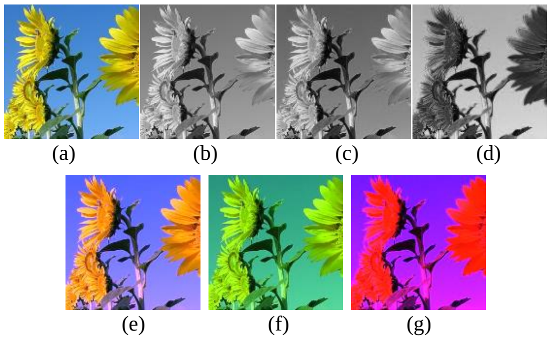


Figure 2. B/w versions of (a) by various *rgb2gray* models (b, c, d). Perfect re-colorizations of (b) using alternative *rgb2gray* models (e, f, g).

Colorization of pixels

2.2.1. Individual pixels

Assume that colorized monochrome images are obtained using known *rgb2gray* model.

Given a single (x,y) pixel with $I(x,y)$ intensity from **[0:255]** discrete range, its colored counterpart should approximately satisfy (subject to color discretization):

$$I(x,y) \approx k_R R(x,y) + k_G G(x,y) + k_B B(x,y) \tag{2}$$

Since the adopted *rgb2gray* model might be inaccurate, we can use reduced numbers of colors (e.g. 32 levels instead of 256) without affecting significantly Eq. 2.

Eventually, all $\mathbf{32^3 = 32768}$ colors are assigned to various intensities, based on the smallest error in Eq. 2. The numbers of colors assigned to a single intensity are non-uniformly distributed. Fig. 3 contains the actual numbers for two exemplary *rgb2gray* models: **[0.299, 0.587, 0.114]** and **[0.69, 0.12, 0.19]**.

It shows the widest selection of color options for mid-range intensities, with the numbers gradually dwindling for darker/lighter values to, eventually, a deterministic choice for extremely dark/light intensities. With no prior information provided, all available colors should be considered equally probable, i.e. $p(C_j \mid I) = 1/N$, where N indicates the number of colors assigned to I value.

Fig. 4 displays the pool of colors (under two *rgb2gray* models) for selected values.

2.2.2. Neighboring pixels

If a pixel at (x,y) has an intensity of I but has not been assigned a color yet, the probabilities of colors that could be assigned to I should be influenced by the presence of a neighboring pixel with an intensity I_1 and its already assigned C_{I_1} color.

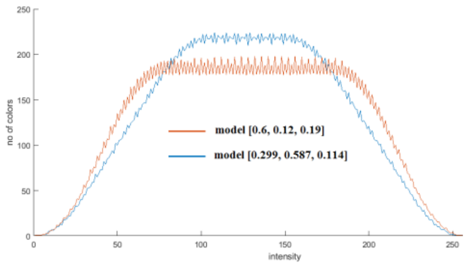
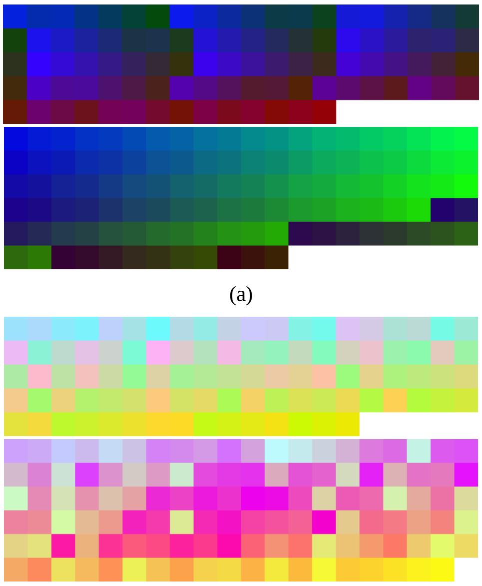


Figure 3. Numbers of RGB colors (out of the total number of 32768) assigned to intensities in two *rgb2gray* models.



(a)

(b)

Figure 4. Colors assigned to (a) 46 and (b) 208 intensities in the [0.299, 0.587, 0.114] and [0.69, 0.12, 0.19] *rgb2gray* models. Note the inconsistencies with human perception of brightness in the second model.

Therefore, we propose a simple heuristic rule:

The greater the difference in brightness between adjacent pixels, the higher the likelihood that their assigned colors will also differ significantly.

Let's consider the pool of colors available for the intensity level I: $\{C_I^1, C_I^2, ..., C_I^N\}$. They are arranged in a monotonically increasing order based on their distance from the color C_{I_1} of the adjacent pixel.

Fig. 4b shows the ordered lists for $I = 208$, assuming that a neighboring pixel of an intensity $I_1 = 46$ was assigned the RGB color $C_{I_1} = [20, 42, 137]$.

Then, we select the color C_I from the list using a uniform distribution which is defined over certain fragments of the list, depending on the difference in intensity levels $abs(I - I_1)$. We tested several options, but eventually implemented a heuristic approach where the C_I color is randomly selected from the $< C_I^{i_{\min}}, C_I^{i_{\max}} >$ range specified by the indexes i_{min} and i_{max} given by Eq. 3.

$$
\begin{aligned}
& i_{\min} = \max\left(1, round\left(\alpha \cdot diff\right)\right), \\
& i_{\max} = \min\left(N, round\left(diff\right)\right), \\
& \text{where} \quad diff = N \cdot \min\left(1, \frac{abs(I - I_1)}{128}\right)
\end{aligned}
\tag{3}
$$

In some cases, the choice is deterministic (formally represented by the $i_{\max} \le i_{\min}$ condition), e.g.:

- White/black pixels are always colored using the brightest/darkest color.
- If neighboring pixels have the same brightness their colors are also the same (this may later change later as discussed below).

In the implementation of the method, images are colored incrementally (see details in Section 3) and it may happen that an uncolored pixel has several already colored neighbors. Then, the color selection can be performed several times for that pixel, and the final choice is a weighted sum of the colors obtained from all colored neighbors.

$$
C_I = \frac{1}{M} \sum_{j=1}^{M} C_{I_j} \quad \text{where} \quad M = 1, 2, 3 \text{ or } 4 \tag{4}
$$

In this way, we can get more colors than a limited pool of **32768** colors initially assumed in Section 2.2.1.

3. IMPLEMENTATION DETAILS

Initialization procedure

Colorization of monochrome images is performed incrementally, starting from a number of initially colored pixels. In the simplest case, it can be even a single pixel.

The proposed options that do not require human assistance for the initial list (queue) of colored pixels are:

a) The darkest/brightest pixel of the image. Because its color is usually deterministic (see Section 2), no human assistance is needed.
b) As in (a), but the list contains all darkest or brightest pixels (or both).

Image colorization

The image colorization method is actually a randomized variant of a popular *flood-fill* algorithm (in the queue-based version).

We randomly select a pixel from the current list ***L*** of colored pixels and colorize its uncolored neighbors using the method outlined in Section 2.2.2. This way, the colored patch grows randomly, avoiding unnecessary regularities in the colorization process (see Fig. 5).

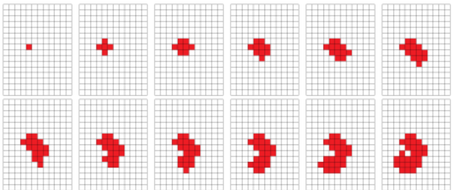


Figure 5. Random growth of the colored patch.

4.3.1. Final touch ups

As highlighted in Section 2.2.1, pixels are initially colored using a limited set of **32768** colors.

However, additional colors can be introduced when the colors assigned to pixels are averaged by Eq. 4 (i.e. a colorized pixel has more already colored neighbors).

We further increase the diversity of colors by projecting them on planes of *rgb2gray* models.

Given a pixel with the original I intensity and the assigned color C_I, we find its closest counterpart $C_{I\text{mod}}$ on the selected *rgb2gray* plane (Eqs 1 and 2).

Such modifications may not noticeably change colorized images when projected onto the plane of the original *rgb2gray* model (Figs 6a and 6b), but they can enhance the visual plausibility of colorized images when projected onto the YUV plane (Fig. 6c).

(a) (b) (c)

Figure 6. The colorization results (a) and their projection on the original *rgb2gray* plane (b) and on the YUV plane (c).

4. EXPERIMENTAL RESULTS

The proposed methodology incorporates several heuristics, arbitrary assumptions, and probabilistic schemes. Therefore, evaluating its performance and practicality requires extensive experimentation. Unfortunately, popular image similarity metrics cannot be used because the ground-truth color images are assumed to be nonexistent.

Therefore, we evaluate the results using subjective criteria as a preliminary approach. We consider two evaluation criteria:

- Visual plausibility, without considering domain-based realities (e.g., grass of any color can be accepted if convincingly rendered).
- Repeatability under the same the same *rgb2gray* model (see Section 2) and the same initialization mode (see Section 3).

Due to page limitations, we include only a selection of results in this section to illustrate the presented conclusions. For example, we only consider three *rgb2gray* models. A more extensive summary of the results is provided in the supplementary materials.

Datasets

We use a diversified collection of monochrome visual-frequency, IR, and other images. For the visual-frequency images, we show their *ground-truth* colors (if available) for information purposes only and do not use them to evaluate colorization quality.

The images are sourced from personal resources and public databases. As no benchmarks (to the best of our knowledge) exist for the discussed topic, we have selected the databases somewhat arbitrarily. The visual-frequency images (converted to monochrome) mainly come from well-known UKBench and SUN databases. The IR images are primarily selected from CAMEL [Geb18] and SMD [Pra17] datasets, while examples of other non-visual images (e.g., MRI, X-ray, etc.) come from various sources.

Parameters of colorization

Altogether, 4851 *rgb2gray* models were considered, corresponding to a sampling of the model coefficients with a 0.001 increment, but in the end, only three models were selected for the experiments reported in the paper, namely:

i. $k_R = 0.299$, $k_G = 0.587$ and $k_B = 0.114$, i.e. the standard YUV model accurately converting colors into a subjective perception of brightness.
ii. $k_R = 0.301$, $k_G = 0.387$ and $k_B = 0.302$, which is similar to a simple mean of primary colors.
iii. $k_R = 0.69$, $k_G = 0.12$ and $k_B = 0.19$, a model with deliberately unrealistic coefficients.

As discussed earlier, three initialization variants are considered, i.e. **(a1)** a single darkest pixel, **(a2)** a single brightest pixel and **(b)** all darkest and brightest pixels. By considering three initialization variants, a single run of the colorization algorithm can produce *nine* results (all possible combinations of *rgb2gray* models and initialization options).

Visual plausibility

The YUV-based *rgb2gray* model produces the best plausibility in the sense that all details from grayscale images are equally clearly seen (sometimes even overexposed) in their colored counterparts. This is not surprising because this model provides the best compatibility between colors and their brightness perception by human eyes.

Results of the second model are still acceptable, but not all details of the original contents can be as clearly seen as in the first model.

For the third rgb2gray model, colors are usually assigned to intensities in such a way that human eyes can hardly identify the image details.

ISSN 2464-4617 (print)
ISSN 2464-4625 (online)
Computer Science Research Notes - CSRN 3301
http://www.wscg.eu
WSCG 2023 Proceedings

Fig. 7 shows exemplary colorization results for selected IR and visual-frequency images by the three models. It can be noticed that the richness of colors is satisfactory in all three models.

Altogether, we can preliminarily conclude that the plausibility of colorization depends strongly on the selected *rgb2gray* model; the more "natural" the model, the better.

Plausibility by repeatability

We found that the plausibility of colorization can be improved by averaging several runs of the algorithm with the same *rgb*2*gray* model and initialization.

Surprisingly, such averaged images do not converge to grayscale, as intuitively expected (since colors are assigned to intensities using probabilistic heuristics with uniform distributions).

Instead, as seen in Fig. 8, we get visually attractive images with diversified coloristics (although the colors are usually less saturated than those from individual runs of the algorithm).

What is even more interesting, the averaged images obtained with the same *rgb*2*gray* model (regardless of the initialization) are usually quite similar. Examples are provided in Fig. 9.

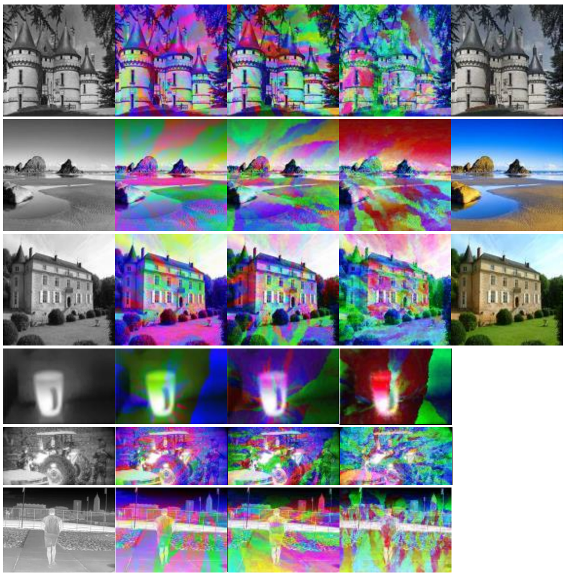
Thus, we cautiously hypothesize (and preliminary theoretical results seem to confirm this hypothesis) that unique image colorizations for the selected *rgb2gray* model might objectively exist. Nevertheless, further experimental and theoretical research on this topic is needed.

5. CONCLUDING REMARKS

In this paper, we attempted to handle the ill-posed problem of colorizing grayscale images without any (direct or indirect) human assistance. We only assume that a hypothetical decolorization model is given. Initially, we use a limited number of $\mathbf{32^3}$ colors, but the algorithm can subsequently use the full sRGB gamut of colors.

The colorization process is performed using a randomized *flood-fill* method, starting from the darkest/brightest pixels for which the choice of color is deterministic. Subsequently, simple probabilistic heuristics are applied to incrementally colorize other pixels.

In spite of the heavy presence of randomizing factors, the results are surprisingly repeatable, depending on the adopted *rgb*2*gray* model and (to a rather insignificant extent) on the applied initialization mode. We even cautiously hypothesize that for the adopted *rgb*2*gray* model, unique optimum colorizations may exist for monochrome images (possibly with some additional limitations).



B/W mod. (i) mod. (ii) mod. (iii) original

Figure 7. Selected b/w images and their colorized samples for three *rgb*2*gray* models. For visual-frequency images, the original color versions are also included.

B/W mod. (i) mod. (ii) mod. (iii) original

Figure 8. Results for Fig. 7 with colors averaged over 10 runs (with the same model and initialization mode).

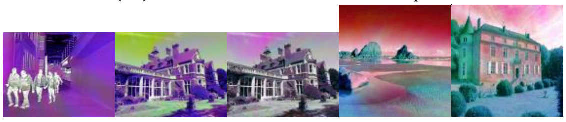
The method is primarily intended for colorizing grayscale images for which there are no physical color counterparts. In other words, we aim to produce convincingly rich colorized versions of "gray worlds". This may be required for various reasons, even if only aesthetic.

Nevertheless, many visual-frequency images are used in the experimental work to better highlight the differences between our approach and the "traditional" re-colorization.

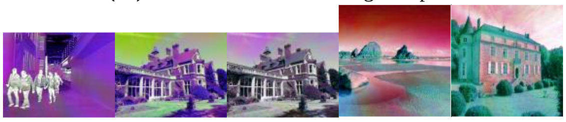
In the future work we intend to focus on the following aspects of the project:

- A formal analysis of the statistical properties of the method (including alternative probability distributions used in the adopted heuristics).
- The development of metrics for objectively estimating the quality of colorization results (e.g. [Has03]).
- Extension of the method to unguided colorization of monochrome movies.

(a1) initialization from the darkest pixel



(a2) initialization from the brightest pixel



(b) initialization from all darkest and brightest pixels

mod. (i) mod. (i) mod. (ii) mod. (iii) mod. (iii)

Figure 9. Colorizations for various initializations using the same *rgb2gray* models. The results are averaged from 75 runs of the algorithm.

6. REFERENCES

[Des15] Deshpande, A., Rock, J., Forsyth, D. Learning large-scale automatic image colorization, 2015 IEEE Int. Conference on Computer Vision (ICCV), pp. 567-575, 2015. DOI: 10.1109/ICCV.2015.72

[Geb18] Gebhardt, E., Wolf, M. CAMEL dataset for visual and thermal infrared multiple object detection and tracking, IEEE Int. Conf. on Advanced Video and Signal-based Surveillance (AVSS), pp. 1-6, 2018. DOI: 10.1109/AVSS.2018.8639094

[Gup12] Gupta, R.K., Chia, A., Rajan, D., Ng, E.S., Zhiyong, H. Image colorization using similar images. In: Proc. 20th ACM International Conference on Multimedia MM'12, pp. 369–378, 2012. DOI: 10.1145/2393347.2393402

[Has03] Hasler, D., Suesstrunk, S.E. Measuring colorfulness in natural images. Proc. SPIE 5007, Human Vision and Electronic Imaging, pp. 87-95, 2003. DOI: 10.1117/12.477378.

[Hwa16] Hwang, J., Zhou, Y. Image colorization with deep convolutional neural networks. Stanford U. Tech. Rep., http://cs231n.stanford.edu/reports/2016/pdfs/219_Report.pdf, 2016.

[Iiz16] Iizuka, S., Simo-Serra, E., Ishikawa, H. Let there be color!: Joint end-to-end learning of global andl image priors for automatic image colorization with simultaneous classification. ACM Trans. Graph., vol. 35(4), Article 110, 2016. DOI: 10.1145/2897824.2925974

[Iro05] Irony, R., Cohen-Or, D., Lischinski, D. Colorization by example. In: Eurographics Symposium on Rendering (eds K. Bala and Ph. Dutre), 2005. DOI: 10.2312/EGWR/EGSR05/201-210

[Lag08] Lagodzinski, P., Smolka, B. Digital image colorization based on probabilistic distance transformation. LNCS 5197 (CIARP 2008) (eds. J. Ruiz-Shulcloper and W.G. Kropatsch), pp. 626–634, 2008.

[Lee22] Lee, H., Kim, D., Lee, D., Kim, J., Lee, J. Bridging the domain gap towards generalization in automatic colorization. LNCS vol. 13677 (ECCV 2022), (eds Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T.), pp. 527-543, 2022. DOI: 10.1007/978-3-031-19790-1_32

[Lev04] Levin, A., Lischinski, D., Weiss, Y. Colorization using optimization. In: ACM SIGGRAPH 2004 Papers (ed. J. Marks), pp.689–694, 2004.

[Pra17] Prasad, D.K., D. Rajan, D., Rachmawati, L., Rajabaly, E., Quek, C. Video processing from electro-optical sensors for object detection and tracking in maritime environment: A survey. IEEE Trans. Intelligent Transportation Systems, vol. 18 (8), pp. 1993-2016, 2017. DOI: 10.1109/TITS.2016.2634580

[Roy17] Royer, A., Kolesnikov, A., Lampert, Ch.H. Probabilistic image colorization. Proc. BMVC'17 Conf., pp. 85.1-85.12, 2017. DOI: 10.5244/C.31.85

[Sal22] Salmona, A., Bouza, L., Delon, J. DeOldify: A review and implementation of an automatic colorization method. Image Processing On Line, vol. 12, pp. 347–368, 2022. DOI: 10.5201/ipol.2022.403

[Zeg21] Žeger, I., Grgic, S., Vuković, J., Šišul, G. Grayscale image colorization methods: Overview and evaluation. IEEE Access, vol. 9, pp. 113326-113346, 2021. DOI: 10.1109/ACCESS.2021.3104515

[Zha16] Zhang, R., Isola, P., Efros, A.A. Colorful image colorization. LNCS vol. 9907 (ECCV 2016), (eds Leibe, B., Matas, J., Sebe, N., Welling, M.), pp. 649–666, 2016. DOI: 10.1007/978-3-319-46487-9_40