

# **A Unified Model for Image Colorization**

Fabien Pierre, Jean-François Aujol, Aurélie Bugeau, Vinh-Thong Ta

# **To cite this version:**

Fabien Pierre, Jean-François Aujol, Aurélie Bugeau, Vinh-Thong Ta. A Unified Model for Image Colorization. CPCV workshop - European Conference on Computer Vision (ECCV), Sep 2014, Zürich, France. pp.1-12. hal-01038055

# **HAL Id: hal-01038055 <https://hal.science/hal-01038055>**

Submitted on 23 Jul 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

- 
- 
- 
- 
- 
- 

- 
- 
- 
- 
- 
- 

# A Unified Model for Image Colorization

Fabien Pierre<sup>1,2,3,4</sup>, Jean-François Aujol<sup>1,2</sup>, Aurélie Bugeau<sup>3,4</sup>, Vinh-Thong  $Ta^{4,5}$ 

<sup>1</sup>Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France. CNRS, IMB, UMR 5251, F-33400 Talence, France. Univ. Bordeaux, LaBRI, UMR 5800, F-33400 Talence, France. CNRS, LaBRI, UMR 5800, F-33400 Talence, France. IPB, LaBRI, UMR 5800, F-33600 Pessac, France.

Abstract. This paper addresses the topic of image colorization that consists in converting a gray-scale image into a color one. In the literature, there exist two main types of approaches to tackle this problem. The first one is the manual methods where the color information is given by some scribbles drawn by the user on the image. The interest of these approaches comes from the interactions with the user that can put any color he wants. Nevertheless, when the scene is complex many scribbles must be drawn and the interactive process becomes tedious and timeconsuming. The second category of approaches is the exemplar-based methods that require a color image as input. Once the example image is given, the colorization is generally fully automatic. A limitation of these methods is that the example image needs to contain all the desired colors in the final result. In this paper, we propose a new framework that unifies these two categories of approaches into a joint variational model. Our approach is able to take into account information coming from any colorization method among these two categories. Experiments and comparisons demonstrate that the proposed approach provides competitive colorization results compared to state-of-the-art methods.

## 1 Introduction

Colorization is an old subject that began with the ability of screens and devices to display colors. The first colorization method [1] tried to map every luminance (gray-scale) level into a color space. Obviously, all the color space can not be recovered without injecting other information. Priors can be added in two ways: with manual interactions or by giving a color image as an example (also called source). In the rest of this paper, we call *target* the gray-scale image to colorize.

 In the first category of methods, a user manually adds points of color (called scribbles) on the target. Figure  $1(d)$  shows an example of result obtained with such priors. Generally, the image is considered as the luminance channel which is not modified during the colorization. Only chrominance channels are computed within the algorithm. Numerous methods have been proposed based on this principle. For instance, the method of Levin *et al.* [2] solves an optimization problem to diffuse the chrominance of the scribbles to the target with the



Fig. 1. Example of exemplar-based, manual and unified colorization. Our method is both able to work with only exemplar or only manual priors. Its novelty is the ability to combine both priors.

 assumption that chrominance must have small variations if the luminance has small variations. Yatziv *et al.* [3] propose a simple and fast method by using geodesic distances to blend the chrominances given by the scribbles. Heu et al. [4] use pixel priorities to ensure that important areas end up with the right colors. Other propagation schemes have been proposed, for instance, probabilistic distance transform [5], random walks [6], discriminative textural features [7], structure tensors [8] or non local graph regularization [9]. As often described in the literature, with these color diffusion approaches, the contours are not well preserved. In [10], the scribbles are automatically generated after segmenting the image and the user only needs to give one color to each scribble. As it is the case with all manual methods, this latter approach suffers from the following drawback: if the target represents a complex scene, then the user's interactions become very important.

 In exemplar-based colorization methods, the color information is extracted from a *source* color image selected by the user. Figure  $1(c)$  shows an example of such a method where the final result is obtained with the approach introduced in this paper. The first exemplar-based method is the one proposed by Welsh et al. [11] (derived from a texture synthesis algorithm [12]) that uses image patch similarities on the gray-scale channels to provide colors. [11] also proposes manual information (called swatches) to specify where to search for similar patches in the source image. Generally, exemplar-based approaches suffer from spatial consistency since each pixel is processed independently. To overcome this limitation, several works use image segmentation to improve the colorization results. For instance, Irony *et al.* [13] propose to determine the best matches between the target pixels and regions in a pre-segmented source image. With these correspondences, micro-scribbles from the source are initialized on the target image and colors are propagated as in [2]. Gupta *et al.* [14] extract different features from the superpixels [15] of the target image and match them with the source ones. The final colors are computed by imposing spatial consistency as in [2]. Charpiat *et al.* [16] ensure a spatial coherency without a segmentation but their method involves many complex steps while Chen *et al.* [17] uses image matting. Without requiring image segmentation or superpixels extraction, Bugeau et al. [18] compute a set of color candidates for each target pixel by matching patches

 with the source image using different features. The final color is obtained by minimizing a functional including a total variation regularization on the chrominance channels. Nevertheless, despite this regularization, the contours are not always well preserved.

 To summarize, the interactivity of manual methods is interesting since the user can add the color information he wants, but this task can be tedious and time-consuming. In contrast, the automatic aspect of exemplar-based methods is also interesting since they try to avoid these user's interventions by using one or several source images. However, in many cases, the user wants to add a color that is not present in the source image or to improve/correct the colorization result in one particular region by adding a scribble. The work proposed in this paper follows this idea and proposes a novel model that unifies both approaches. Up to our knowledge, this is the first method proposing this unification.

 The main contribution of this work is the combination of the two categories of methods into a unified variational model which allows interactivity.

The paper is organized as follows. We first introduce in Section 2 a variational model for image colorization and give an efficient algorithm to compute the solution. In Section 3 we present the unified model for colorization. In section 4 comparisons with state-of-the-art methods demonstrate the improvements of the proposed approach.

# 2 A Variational Model for Image Colorization

In this section we propose a variational framework for image colorization. As done in most existing colorization methods, we consider the target to be a luminance image and we compute chrominances for each pixels in order to recover the final color. First, the variational model including a regularization is introduced. Next, the corresponding algorithm is given.

#### 2.1 Penalized Variational Labelling

The color can be expressed as a value of luminance Y and two values of chrominance  $U$  and  $V$ . With these three values, it is possible to display an image on most devices. The luminance channel being given by the target, the colorization problem consists in recovering the chrominance values for each pixels. Inspired from [18], we suppose that, for each pixel x, C chrominance candidates  $c_i, i = 1...C$ are available. These color candidates can both be given by extraction according texture features, by a manual method or by any colorization method. Our model consists in choosing the best color among the C.

To that end, we propose to minimize the following functional where  $u$  is the vector of chrominances  $(u = (U, V))$  to compute:

$$
F(u, W) := TV_c(u) + \frac{\lambda}{2} \int_{\Omega} \sum_{i=1}^{C} w_i ||u - c_i||^2 + \chi_{\mathcal{R}}(u) + \chi_{\mathcal{E}}(W) . \tag{1}
$$

 To simplify the notations, the dependence of each values to the position of the current pixel is removed. For instance, the second term of (1) is a notation for  $\int_{\Omega} \sum_{i=1}^C w_i(\omega) \|u(\omega) - c_i(\omega)\|^2 d\omega$ . We explain hereinafter the significance of each term of functional (1).

 Consider the minimization of the term  $\int_{\Omega} \sum_{i=1}^{C} w_i ||u-c_i||^2$  with respect to u. The minimum is equal to  $\sum_{i=1}^{C} w_i c_i$ . Thus it corresponds to a weighted average of the candidates. In order to transform this minimization into a problem of labelling, we add the constraint that all weights are equal to 0 except one which corresponds to the selected label. The set of such weights is the canonical basis of  $\mathbb{R}^C$  denoted as  $\mathcal{E}$ . This constrained problem is denoted as  $\int_{\Omega} \sum_{i=1}^C w_i ||u |c_i|^2 + \chi_{\mathcal{E}}(W)$  where  $\chi_{\mathcal{E}}$  is a function which value is equal to 0 on  $\mathcal{E}$  and  $+\infty$ otherwise and W represents the weights  $(w_1, \ldots, w_C)$ .

 The minimization of this term with respect to  $u$  and  $W$  provides a natural labelling where  $c_i$  are labels. Indeed, this minimum is obtained for a binary weight and the minimization of this term with respect to  $u$  is equal to  $c_i$ . Authors of [18] propose a different approach to get binary weights, but [19] shows that it is too much sensitive to the initialization. Every labelling corresponds to a local minimum (which is global) of this term. A penalization term is therefore introduced to encourage a regularized labelling. In this model, we introduce the total variation  $TV_c$  to favors images with contours but few discontinuities.  $\lambda$  is a parameter which controls the influence of the regularization term on the result. Finally, the term  $\chi_{\mathcal{R}}$  requires the values to remain in the standard range for chrominance channels.

 $TV_c$  is the coupled total variation defined in the following.

### 2.2 Coupling the Channels

 The classical total variation  $\int_{\Omega} \sqrt{\partial_x U^2 + \partial_y U^2 + \partial_x V^2 + \partial_y V^2}$  suffers from a lack of coupling and leads to some halo effect in the colorization results [18].

We introduce a TV regularization that is able to couple the chrominance channels with the luminance one. Let  $TV_c$  be a *coupled* total variation defined as

$$
TV_c(u) = \int_{\Omega} \sqrt{\gamma \partial_x Y^2 + \gamma \partial_y Y^2 + \partial_x U^2 + \partial_y U^2 + \partial_x V^2 + \partial_y V^2} ,\quad (2)
$$

 where  $Y, U$  and  $V$  are the classical luminance and chrominance channels.

 This formulation naturally favors images where contours in chrominance channels are at the same locations as the luminance ones. Figure 2 illustrates this effect. For the sake of clarity, assume that there is a vertical contour in the Y channel,  $\partial_x Y = a > 0$  and  $\partial_y Y^2 = 0$ , and another one in the U channel such that  $\partial_x U = b > 0$ ,  $\partial_y U^2 = \partial_x V^2 = \partial_y V^2 = 0$ . If the two contours are in the same location, the value of the total variation is equal to  $TV_c(u) = \sqrt{\gamma a^2 + b^2}$  but if the contours have different locations, the value is equal to  $TV_c(u) = \sqrt{\gamma a^2} + \sqrt{b^2}$ . Since  $\sqrt{\gamma a^2 + b^2} < \sqrt{\gamma a^2} + \sqrt{b^2}$ , the minimization of TV<sub>c</sub> favors the values of U such that the contours in the chrominance channels are in the same location



Fig. 2. Illustration of two situations for contours in luminance and chrominance channels with  $\gamma = 1$ . TV<sub>c</sub> favors coupling of channels.

 as the luminance one. We experimentally chose  $\gamma = 25$  for all the experiments of the paper.

In Figure 3 we compare our model with a version without coupling. This result is provided by replacing  $TV_c$  by the classical total variation on chrominance channels, or by taking  $\gamma = 0$  in our model. These results have been performed with  $3(b)$  as target image and  $3(a)$  as source.



Fig. 3. Comparison of our method with the classical  $TV$ .

Without coupling the results are totally burred and the colorization process fails. Thanks to  $TV_c$ , our new formulation is able to preserve contours by coupling the channel of luminance Y with the chrominance channels U and V.

#### 2.3 Algorithm

In this section, a min-max version of (1) is proposed. A primal-dual algorithm, inspired from [20] is described to provide a local minimum of functional (1).

To this end, we rewrite the term  $TV_c$  from (2) in a dual form:

$$
TV_c(U, V) = \min_{U, V} \max_{p=(p_1, p_2, p_3)} \langle \nabla U, \nabla V | p_2, p_3 \rangle + \langle \nabla Y | p_1 \rangle + \chi_{\|(p_1, p_2, p_3)\| \le 1} \quad (3)
$$

 where  $p \in \mathbb{R}^6$  and  $p_j \in \mathbb{R}^2$ ,  $j = 1 \cdots 3$ . Y is provided from the target. Minimizing (1) becomes equivalent to maximizing the dual model w.r.t. to the dual variable  $p$  and minimizing it w.r.t.  $u$  and  $W$ . Algorithm 1 summarizes the minimization procedure.



 $S(u_n)$  stands for the closest candidate of  $u_n$ , the colorized image at the n<sup>th</sup> iteration of Algorithm 1. Y is the luminance channel of the image to colorize. Parameters  $\tau$  and  $\sigma$  are the time steps. The algorithm of [20] converges when  $16\tau\sigma < 1$ . The operator div stands for the divergence and  $\nabla$  the gradient defined as in [21]. The algorithm requires the projection of the two estimated data  $p$  and u. The projection  $P_u$  is necessary to ensure that the estimated image stays in the standard range of chrominance values  $R$  and is just a projection onto a rectangle. The computation of  $w_i$  used at initialization is described in Section 3. Finally the projection of the dual variable  $P_p$  ensures the constraint  $\chi_{\{(p_1,p_2,p_3)\}\subset\{1\}}$ , by projecting p onto the unit ball:

$$
P_p(p) = \frac{(p_1 - \sigma(\partial_x Y, \partial_y Y), p_2, p_3)}{\left(\frac{1}{p_1 - \sigma(\partial_x Y, \partial_y Y), p_2, p_3}{\sigma(\partial_x Y, \partial_y Y)}\right)}.
$$
(4)

$$
P_p(p) = \frac{\Gamma\left(\frac{p_1}{p_1} - \sigma\left(\frac{\partial_x p_1}{\partial y_2}\right) + \frac{p_2 p_3 p_4}{p_3 p_4}\right)}{\max\left(\left\|\left(p_1 - \sigma\left(\frac{\partial_x p_1}{\partial y_2}\right) + \frac{\partial_y p_2}{\partial y_4}\right)\right\|_2^2, 1\right)}.
$$
\n(4)

#### 3 Unifying Manual and Exemplar-based Colorization

This section presents a simple method to unify both exemplar and manual priors for colorization of image. Although exemplar-based colorization tackles the tedious work of putting scribbles in manual colorization, the choice of the source image is rarely easy and the results contain often errors. The user may prefer to correct the result by adding a manual prior. The integration of this new prior to the exemplar-based result is not obvious. The solution of this problem is the main contribution of the paper.

 Exemplar-based colorization. In our model, the exemplar-based priors are introduced via candidates. When a source image is provided, the first step consists in extracting for each pixel the set of candidates as done in [18]. These candidates can be provided by any colorization method. The weights  $w_i$  are then chosen as  $W = 1/C$  where C is the number of candidates extracted at each pixel. The algorithm can work directly with these data.

 Manual colorization. This section presents how scribbles can be directly introduced into the proposed model. The scribbles can either be given by the user before or added in an interactive and/or an iterative way. The proposed model can use the scribbles alone, the source alone or both.

 The scribble information only affects the weights and the number of candidates. More precisely, for each pixel, a new candidate per scribble is added.

 Its value is the chrominance of the given scribble. When scribbles candidates are present, their initial weights depend on the geodesic distance. The geodesic distance map, denoted by  $D$ , is computed with the fast marching algorithm  $[22]$ and with a potential equal  $(0.001 + ||\nabla u||_2^2)^{-4}$  given in [23]. D is normalized to have a range between 0 and 1. We use the implementation of  $[24]$  to compute the geodesic distance.



(a) Initial scribble (b) Distance map  $D$  (c) Colorization

Fig. 4. Example of color propagation with the geodesic distance. (a) represents the initial scribble, (b) the geodesic distance map from this scribble and (c) the diffusion of the color of the scribble according to this distance map. This diffusion provides a good initialization which is not perfect, but sufficient for our algorithm.

Unified colorization. The unification of the two priors is done at initialization of Algorithm 1. The variable W is initialized with  $1/C$  for the candidates which come from the exemplar-based candidates extraction (in the source image) and with  $1 - D$  for the candidate(s) corresponding to the scribble(s), where D is the normalized distance map. Pixels that have a low geodesic distance to a scribble get its color. At the opposite for pixels having a high geodesic distance, this new candidate will have no influence onto the colorization result.

The variable  $W$  is projected onto the simplex before running the algorithm. The projection is performed with the algorithm of  $[25]$ . The variable u is set to  $\sum_i w_i c_i$  and the functional is minimized using this initialization.

Figure  $4(a)$  shows the initial blue scribble located in the sky. The associated geodesic distance D with respect to the scribble is presented Figure  $4(b)$  and Figure 4(c) is our colorization result. Our method is able to diffuse the color information of the scribble on constant parts of the image and at the convergence of the algorithm the coupled total variation gives more accurate results than the geodesic distance.

# 4 Experimental Results

 In this section, we give some details about the implementation and the parameters of the unified model. We demonstrate the potential of our approach compared to state-of-the-art methods in both categories, i.e., manual and exemplar-based methods.

#### 4.1 Implementation and Parameters Setting

 The implementation of Algorithm 1 has been done on GPU. The convergence takes few seconds and allows interactivity. All the presented experiments use the same set of parameters, *i.e.*,  $\sigma = 0.004$ ,  $\tau = 5$ ,  $\lambda = 7.10^{-3}$ ,  $\gamma = 25$  and 500 iterations. With this choice, a lot of different types of image can be colorized without tuning of parameters which is a practical advantage. For exemplar-based results, we use the candidate extraction described in [18].

#### 4.2 Comparison with State-of-the-art Methods

 

 Figure 5 presents a first example of our proposal. Figures  $5(a)$  and  $5(b)$  show the source and the target images. Figure  $5(c)$  corresponds to the exemplarbased colorization result provided by our model. In this figure, the sky is not correctly colorized since it appears brown instead of blue as in the ruins main door. Moreover, blue colors appear on the floor. Figure 5(d) shows the results performed with only the corrections of the user where 3 scribbles are added in order to correct the first (exemplar-based) colorization result (Figure  $5(c)$ ). Figure 5(e) illustrates the advantage of the proposed unified image colorization since, the user with less effort, obtained the desired result. Finally, this result also illustrates that our model is well adapted to preserve the color contours.

 Figure 6 presents results and illustrates the advantage of using a unified image colorization as compared to only using a source image or some scribbles. Manual colorization results are obtained with [2] and exemplar-based colorization results are obtained with [14]. Colorization results of the last column of Figure 6 are clearly better than the ones obtained with only the source image (fourth column) or with only the scribbles (fifth column). Actually, some objects (for instance the tramway, or the background of the portrait) are not present on the source image, thus the exemplar-based method fails. For the manual method, there is a strong lack of information, because few scribbles are provided. We remark in the fourth image that the method of [2] is not robust to noise, compared to ours. This experiment also highlights that old photographs and faces are hard to colorize as remarked, e.g., in [17]. Indeed, old photographs contain a lot of noise and the texture are usually degraded. Face images contain very smooth



Fig. 5. First example of proposed image colorization model. Neither exemplar-based approach nor the manual one are able to properly colorize this image with the given priors. The unified method gives a suitable result.



Fig. 6. Advantage of the proposed unified model as compared to only exemplar-based or scribble-based model. From left to right: the target image to colorize, the source image, the scribbles added by the user, exemplar result of  $[14]$ , manual result of  $[2]$ , result with the unified approach. The unified model increases the quality of the results by taking advantages in both types of methods.

parts (e.g., the skin) and the background is rarely suitable. Nevertheless, very promising results are obtained with our method. The manual method does not colorize the hair because no scribble is given. The exemplar one does not colorize the background. Finally, the additional prior given by the scribbles of the user does not only have a local effect. Indeed, in the last result of Figure 6, the blue scribble needed to colorize the sky through the arch also improves the sky color at the bottom left of the image. Figure 9 provides additional results.

#### 4.3 Our Model vs. existing Exemplar-based Methods

 Figure 7 provides comparison of our method with existing exemplar-based colorization methods. Figure 8 is a zoom on a particular results. On the left, the source and target images are shown. Our results are in the third column while the other columns are results from  $[14]$ ,  $[18]$ , and  $[11]$ . Due to the lack of regularization, images of [11] present artefacts: areas that were originally homogeneous now present irregularities (see the sky of the first image). Moreover, their method is not reliable on contours (see the third image). Our approach better preserves the contours and the homogeneous parts such as the sky. The results of [18] present halo near contours due to the lack of coupling of the classical total variation on chrominance channels. On the second image, the color are too shiny.



Fig. 7. Comparison with state-of-the-art exemplar-based methods. In the exemplarbased category, our method is competitive with other state-of-the-art approaches.

This is due to the post-processing that does not constraint the hue to be constant. The quality of our results are comparable to [14] whereas our approach is much simpler since local segmentation like superpixels [15] is not needed. On the image with flowers, we remark that their method does not recover suitable colors. The segmentation used by [14] makes their method unable to colorize thin structures, e.g., the trees on the left of the third image.

This comparison with state-of-the-art exemplar-based methods shows the reliability and the efficiency of our colorization method.

## 5 Conclusion

In this paper, a variational model for unified image colorization is proposed. This method combines manual and exemplar-based methods in a simple and intuitive model. Moreover, it can take into account results from any colorization methods to improve results. It opens the way to powerful interactive colorization. Our variational model includes a total variation term which couples luminance and chrominances channels. With this representation, the contours of the colorized



Fig. 8. Zoom on the third line of Figure 7.



Fig. 9. Additional results

images are well preserved. As future work, we plan to improve the results of the exemplar-based methods by studying the features to compare source and target patches. Finally, the extension to video colorization will be addressed.

# Acknowledgment

This study has been carried out with financial support from the French State, managed by the French National Research Agency (ANR) in the frame of the Investments for the future Programme IdEx Bordeaux (ANR-10-IDEX-03-02). J-F. Aujol is a member of Institut Universitaire de France. The authors would like to thank Raj Kumar Gupta for providing the images presented in [14].

## References

- 1. Gonzales, R.C., Wintz, P.: Digital Image Processing (2Nd Ed.). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (1987)
- 2. Levin, A., Lischinski, D., Weiss, Y.: Colorization using optimization. ACM Trans. on Graphics 23(3) (2004) 689–694
- 3. Yatziv, L., Sapiro, G.: Fast image and video colorization using chrominance blending. IEEE Trans. on Image Processing 15(5) (2006) 1120–1129
- 4. Heu, J., Hyun, D.Y., Kim, C.S., Lee, S.U.: Image and video colorization based on prioritized source propagation. In: Proc. of ICIP. (2009)
- 5. Lagodzinski, P., Smolka, B.: Digital image colorization based on probabilistic distance transformation. In: Proc. of ELMAR. Volume 2. (2008) 495–498
- 6. Kim, T.H., Lee, K.M., Lee, S.U.: Edge-preserving colorization using data-driven random walks with restart. In: Proc. of ICIP. (2010) 1661–1664
- 7. Kawulok, M., Kawulok, J., Smolka, B.: Discriminative textural features for image and video colorization. IEICE Trans. on Information and Systems 95-D(7) (2012) 1722–1730
- 8. Drew, M.S., Finlayson, G.D.: Improvement of colorization realism via the structure tensor. Int. Jour. on Image Graphics 11(4) (2011) 589–609

