

Article

Exemplar-Based Face Colorization Using Image Morphing

Johannes Persch ¹, Fabien Pierre ^{2,*} and Gabriele Steidl ^{1,3}

¹ Department of Mathematics, Technische Universität Kaiserslautern, Paul-Ehrlich-Str. 31, 67663 Kaiserslautern, Germany; persch@mathematik.uni-kl.de (J.P.); steidl@mathematik.uni-kl.de (G.S.)

² Laboratoire Lorrain de Recherche en Informatique et ses Applications, UMR CNRS 7503, Université de Lorraine, INRIA projet Magrit, 54500 Vandoeuvre-les-Nancy, France

³ Fraunhofer ITWM, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany

* Correspondence: fabien.pierre@univ-lorraine.fr

Received: 30 May 2017; Accepted: 19 October 2017; Published: 31 October 2017

Abstract: Colorization of gray-scale images relies on prior color information. Exemplar-based methods use a color image as source of such information. Then the colors of the source image are transferred to the gray-scale target image. In the literature, this transfer is mainly guided by texture descriptors. Face images usually contain few texture so that the common approaches frequently fail. In this paper, we propose a new method taking the geometric structure of the images rather their texture into account such that it is more reliable for faces. Our approach is based on image morphing and relies on the YUV color space. First, a correspondence mapping between the luminance Y channel of the color source image and the gray-scale target image is computed. This mapping is based on the time discrete metamorphosis model suggested by Berkels, Effland and Rumpf. We provide a new finite difference approach for the numerical computation of the mapping. Then, the chrominance U,V channels of the source image are transferred via this correspondence map to the target image. A possible postprocessing step by a variational model is developed to further improve the results. To keep the contrast special attention is paid to make the postprocessing unbiased. Our numerical experiments show that our morphing based approach clearly outperforms state-of-the-art methods.

Keywords: image colorization; morphing; face images

1. Introduction

Colorization consists in adding color information to a gray-scale images. This technique is used for instance by the cinema industry to make old productions more attractive. As usual we can consider a gray-scale image as luminance channel Y of an RGB image [1–4]. The Y channel is defined as a weighted average of the RGB channels, see, e.g., [5]:

$$Y = 0.299R + 0.587G + 0.114B.$$

In addition to the luminance channel, two chrominance channels, called *U* and *V*, enable to recover the RGB image. Recovering an RGB image from the luminance channel alone is an ill-posed problem and requires additional information [3,6]. This information is provided in the literature in two different ways, namely by *manual* [1,4,7] or *exemplar-based* [3,6,8] methods. In the first one the user augments the image by some color strokes as basis for the algorithm to compute the color of each pixel. The colorization of a complex scene by a manual prior can be a tedious work for the user [2]. In the second approach a color image is used as a source of information. Here the results strongly depend on the choice of the image. Therefore it is often called *semi-automatic* [1].

In this paper, we focus on the exemplar-based methods. A common back-bones of these techniques is the matching of the images. First, the color *source* image is transformed to a gray-scale image, referred to as *template* which is compared with the input gray-scale image called *target*. The main issue of the exemplar-based colorization consists in matching the pixels of the template and the target gray-scale images. The basic hypothesis in the literature is that the color content is similar in similar texture patches [8]. Then the main challenge is the choice of appropriate texture descriptors.

In the literature, the exemplar-based methods come along with various matching procedures. For instance, Gupta et al. [6] use SURF and Gabor features, Irony et al. [2] use DCT, Welsh et al. [8] use standard-deviation, etc. Some of them are done after an automatic segmentation, whereas others use local information. The seminal paper on exemplar-based colorization by Welsh et al. [8] was inspired by the texture synthesis of Efros and Leung [9] and uses basic descriptors for image patches (intensity and standard-deviation) to describe the local texture. Pierre et al. [3] proposed an exemplar-based framework based on various metrics between patches to produce a couple of colorization results. Then a variational model with total variation like regularization is applied to choose between the different results in one pixel with a spatial regularity assumption. The approach of Irony et al. [2] is built on the segmentation of the images by a mean-shift algorithm. The matching between segments of the images is computed from DCT descriptors which analyse the textures. The method of Gupta et al. [6] is rather similar. Here, an over-segmentation (SLIC, see, e.g., [10]) is used instead of the mean-shift algorithm. The comparison between textures is done by SURF and Gabor features. Chen et al. [11] proposed a segmentation approach based on Bayesian image matching which can also deal with smooth images including faces. The authors pointed out that the colorization of faces is a particular hard problem. However, their approach uses a manual matching between objects to skirt the problem of smooth parts. Charpiat et al. [12] ensured spatial coherency without segmenting, but their method involves many complex steps. The texture discrimination is mainly based on SURF descriptors. In the method of Chia et al. [13], the user has manually to segment and label the objects and the algorithm finds similar segments in a set of images available in the internet. Recently a convolutional neural network (CNN) has been used for colorization by Zhang et al. [14] with promising results. Here the colorization is computed from a local description of the image. However, no regularization is applied to ensure a spatial coherence. This produces "halo effects" near strong contours. All the described methods efficiently distinguish textures and possibly correct them with variational approaches, but fail when similar textures have to be colorized with different colors. This case arises naturally for face images. Here the smooth skin is considered nearly as a constant part. Thus, as we show in this paper, when the target image contains constant parts outside the face, the texture-based methods fail.

In this paper, we propose a new technique for the colorization of face images guided by image morphing. Our framework relies on the hypothesis that the global shape of faces is similar. The matching of the two images is performed by computing a morphing map between the target and the template image. Image morphing is a generic term for smooth image transition which is an old problem in image processing and computer vision. For example in feature based morphing only specific features are mapped to each other and the whole deformation is then calculated by interpolation. Such method [15] was used by the film industry for example in the movie *Willow*. For an overview of this and similar techniques see also [16,17]. A special kind of image morphing, the so-called metamorphosis was proposed by Miller, Trouvé and Younes [18–20]. The metamorphosis model can be considered as an extension of the flow of diffeomorphism model and its large deformation diffeomorphic metric mapping framework [21–25] in which each image pixel is transported along a flow determined by a diffeomorphism. As an extension the metamorphosis model allows the variation of image intensities along trajectories of the flow. Shooting methods for the metamorphosis model were developed e.g., in [26]. For a metamorphosis regression model and corresponding shooting methods we refer to [27]. A comprehensive overview over the topic is given in the book [28] as well as the review article [29], for a historic account see also [30]. In this paper, we build up on the time discrete metamorphosis model by Berkels, Effland and Rumpf [31]. In contrast to these authors we apply a finite difference approach for

the computation of the morphing maps. This involves the solution of a chain of registration problems in each iteration step. There exists a rich literature on registration problems, e.g., [32–35], and we refer to the books of Modersitzki [36,37] for an overview.

Having the morphing map available, the chrominance channels can be transported by this map to the target image, while preserving its luminance channel. This gives very good results and outperforms state-of-the-art methods. For some images we can further improve the quality by applying a variational post-processing step. Our variational model incorporates a total variation like regularization term which takes the edges of the target image into account. This was also proposed by one of the authors of this paper in [3]. The method is accomplished by adapting a procedure of Deledalle et al. [38] to keep the contrast of the otherwise biased total variation method.

The outline of the paper is as follows: in Section 2 we sketch the ideas of the morphing approach. In particular, we show how the morphing map is computed with an alternating minimization algorithm and describe our finite difference approach. Details of the computation are shifted to the Appendix. Section 3 deals with the color transfer. Having the morphing map at hand we explain the transfer of the chrominance values. In particular we address the necessary scalings. Sometimes it is useful to apply a variational model with a modified total variation regularization as post-processing step to remove possible spatial inconsistencies. Such a procedure is developed in Section 4. Numerical experiments demonstrate the very good performance of our algorithm in Section 5. The paper ends with conclusions in Section 6.

2. Image Morphing

Our colorization method is based on the time discrete image metamorphosis model [31]. We briefly explain the model used also in the numerical part of [31] in Section 2.1 and describe our numerical realization for digital images by a finite difference approach in Section 2.2. Note that in [31] a finite element approach for the spatial discretization was proposed without a detailed description. Finite element methods are highly flexible and can be also applied, e.g., for shape metamorphosis. However, having the rectangular structure of the image grid in mind, we propose to use finite differences for the spatial discretization. Then, in Step 1 of our alternating algorithm, we can build up on registration methods proposed e.g., by Haber and Modersitzki [34].

2.1. Morphing Model Based on [31]

Let $\Omega \subset \mathbb{R}^2$ be an open, bounded domain with Lipschitz continuous boundary. We are given a gray-value template image $I_{\text{temp}} : \Omega \rightarrow \mathbb{R}$ and a target image $I_{\text{tar}} : \Omega \rightarrow \mathbb{R}$ which are supposed to be continuously differentiable and compactly supported. For $K \geq 2$ set

$$I_0 := I_{\text{temp}}, \quad I_K := I_{\text{tar}}.$$

In our application, the template image will be the luminance channel of the color source image and the target image the gray-scale image we want to colorize. We want to find a sequence of $K - 1$ images \mathbf{I} together with a sequence of mappings $\boldsymbol{\varphi}$ on Ω , i.e.,

$$\mathbf{I} := (I_1, \dots, I_{K-1}), \quad \boldsymbol{\varphi} := (\varphi_1, \dots, \varphi_K),$$

such that

$$I_k \approx I_{k-1} \circ \varphi_k \quad \text{for all } x \in \Omega,$$

see Figure 1, and the deformations φ_k have a small linearized elastic potential defined below. To this end, we suppose for $k = 1, \dots, K$ that φ_k is related to the displacement v_k by

$$\varphi_k(\mathbf{x}) = \mathbf{x} - v_k(\mathbf{x}) = \begin{pmatrix} x - v_{k,1}(\mathbf{x}) \\ y - v_{k,2}(\mathbf{x}) \end{pmatrix}, \quad \mathbf{x} = (x, y)^T \in \Omega, \tag{1}$$

and set $\mathbf{v} := (v_1, \dots, v_K)$. The (Cauchy) strain tensor of the displacement $v = (v_1, v_2)^T : \Omega \rightarrow \mathbb{R}^2$ is defined by

$$\varepsilon(v) := \frac{1}{2}(\nabla v + \nabla v^T) = \begin{pmatrix} \partial_x v_1 & \frac{1}{2}(\partial_y v_1 + \partial_x v_2) \\ \frac{1}{2}(\partial_y v_1 + \partial_x v_2) & \partial_y v_2 \end{pmatrix},$$

where ∇v denotes the Jacobian of v . The linearized elastic potential is given by

$$\mathcal{S}(v) := \int_{\Omega} \mu \operatorname{trace}(\varepsilon^T(v)\varepsilon(v)) + \frac{\lambda}{2} \operatorname{trace}(\varepsilon(v))^2 \, dx, \tag{2}$$

where $\mu, \lambda > 0$. Then we want to minimize

$$\mathcal{J}(\mathbf{I}, \mathbf{v}) := \sum_{k=1}^K \int_{\Omega} |I_k - I_{k-1} \circ \varphi_k|^2 + \mathcal{S}(v_k) \, dx, \quad \text{subject to } I_0 = I_{\text{temp}}, I_K = I_{\text{tar}}. \tag{3}$$

This functional was used in the numerical part of [31]. Note that the term $\mathcal{S}(v)$ may be accomplished by a higher order derivative

$$\int_{\Omega} |D^m v_k(x)|^2 dx, \quad m > 2, \tag{4}$$

which ensures in the time continuous setting that φ is indeed a diffeomorphism, see [22]. In the time discrete setting the additional term does not guaranty that $\varphi_k, k = 1, \dots, K$, is a diffeomorphism. Hence we do not include (4) into our functional (3), as the linearized elastic potential \mathcal{S} is sufficient for our purposes.

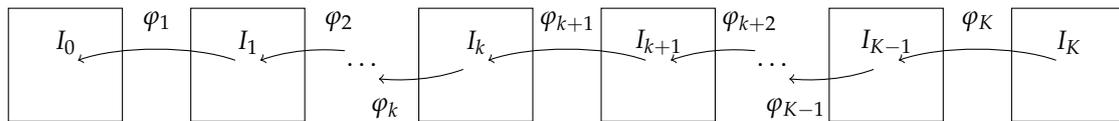


Figure 1. Illustration of the image path and the diffeomorphism path, where $I_0 := I_{\text{temp}}$ and $I_K := I_{\text{tar}}$ are the given template and target images.

The minimizer (\mathbf{I}, \mathbf{v}) of the functional provides us with both a sequence of images \mathbf{I} along the approximate geodesic path and a sequence of displacements \mathbf{v} managing the transport of the gray values through this image sequence. For finding a minimizer of (3) we alternate the minimization over \mathbf{I} and \mathbf{v} :

1. Fixing \mathbf{I} and minimizing over \mathbf{v} leads to the following K single registration problems:

$$\arg \min_{v_k} \mathcal{J}_{\mathbf{I}}(v_k) := \int_{\Omega} |I_k - I_{k-1} \circ \varphi_k|^2 + \mathcal{S}(v_k) \, dx, \quad k = 1, \dots, K, \tag{5}$$

where φ_k is related to v_k by (1).

2. Fixing \mathbf{v} , resp., φ leads to solving the following image sequence problem

$$\arg \min_{\mathbf{I}} \mathcal{J}_{\varphi}(\mathbf{I}) := \sum_{k=1}^K \int_{\Omega} |I_k - I_{k-1} \circ \varphi_k|^2 \, dx. \tag{6}$$

This can be done via the linear system of equations arising from Euler-Lagrange equation of the functional which we describe in the Appendix A.

Note that Miller et al. [39] considered for two given image sequences $\mathbf{I} := (I_j)_{j=1}^N$ and $\mathbf{J} := (J_j)_{j=1}^N$ (e.g., related to corresponding image patches) the registration problem

$$\arg \min_v \int_{\Omega} \sum_{j=1}^N |I_j - J_j \circ \varphi|^2 + \mathcal{S}(v) \, dx,$$

where φ is related to v by (1). In contrast to our problem these authors search for the same mapping v and the N template-target pairs are known.

2.2. Space Discrete Morphing Model

When dealing with digital images we have to work in a spatially discrete setting. Let

$$\mathcal{G}_p := \{1, \dots, n_1\} \times \{1, \dots, n_2\}$$

be the (primal) image grid, i.e., $I_{\text{ref}} : \mathcal{G}_p \rightarrow \mathbb{R}$ and $I_{\text{temp}} : \mathcal{G}_p \rightarrow \mathbb{R}$. We discretize the integrals on the integration domain $\bar{\Omega} := [\frac{1}{2}, n_1 + \frac{1}{2}] \times [\frac{1}{2}, n_2 + \frac{1}{2}]$ by the midpoint quadrature rule, i.e., with pixel values defined on \mathcal{G}_p . For discretizing the operators in (2) we work as usual on staggered grids. For the application of mimetic grid techniques in optical flow computation see also [40]. Let

$$\mathcal{G}_d := \{\frac{3}{2}, \dots, n_1 - \frac{1}{2}\} \times \{\frac{3}{2}, \dots, n_2 - \frac{1}{2}\}.$$

be the (inner) dual grid, i.e., \mathcal{G}_p shifted by $\frac{1}{2}$ in each direction, and

$$\mathcal{G}_1 := \{\frac{3}{2}, \dots, n_1 - \frac{1}{2}\} \times \{1, \dots, n_2\}, \quad \mathcal{G}_2 := \{1, \dots, n_1\} \times \{\frac{1}{2}, \dots, n_2 - \frac{3}{2}\}.$$

We start by considering the registration problems in the first alternating step (5) and turn to the image sequence problem in the second step (6) with (A2) afterwards.

Solution of the Registration Problems

Let us fix $k \in \{1, \dots, K\}$ in (5) and abbreviate $T := I_{k-1}$, $R := I_k$ as template and reference image, resp., and $v := v_k$. Then we have to find a minimizer of the continuous registration problem

$$\mathcal{J}(v) := \int_{\Omega} |R(\mathbf{x}) - T(\mathbf{x} - v(\mathbf{x}))|^2 + \mathcal{S}(v) \, dx \tag{7}$$

where the elastic potential in (2) can be rewritten as

$$\mathcal{S}(v) = \mu \int_{\Omega} (\partial_x v_1)^2 + \frac{1}{2} (\partial_y v_1 + \partial_x v_2)^2 + (\partial_y v_2)^2 \, dx + \frac{\lambda}{2} \int_{\Omega} (\partial_x v_1 + \partial_y v_2)^2 \, dx.$$

For the spatial discrete setting we assume that $T, R : \mathcal{G}_p \rightarrow \mathbb{R}^s$ are given. Further, we consider $v = (v_1, v_2)^T$ with $v_1 : \mathcal{G}_1 \rightarrow \mathbb{R}$ and $v_2 : \mathcal{G}_2 \rightarrow \mathbb{R}$. In contrast to [34] we assume no flow over the image boundary and set

$$\begin{aligned} v_1(\frac{3}{2}, y) = v_1(n_1 - \frac{1}{2}, y) = 0, \quad y \in \{1, \dots, n_2\}, \\ v_2(x, \frac{3}{2}) = v_2(x, n_2 - \frac{1}{2}) = 0, \quad x \in \{1, \dots, n_1\}. \end{aligned}$$

See Figure 2 for an illustration. We approximate $\partial_x v_1$ for $\mathbf{x} = (x, y)^T \in \mathcal{G}_p$ by

$$(D_{1,x} v_1)(x, y) := \begin{cases} 0, & x = 1, \\ v_1(x + \frac{1}{2}, y) - v_1(x - \frac{1}{2}, y), & x = 2, \dots, n_1 - 1, \\ 0, & x = n_1, \end{cases}$$

and $\partial_y v_1$ for $x \in \{1, \dots, n_1 - 1\}$ and $y \in \{1, \dots, n_2 - 1\}$ by

$$(D_{1,y}v_1)(x + \frac{1}{2}, y + \frac{1}{2}) = \left\{ v_1(x + \frac{1}{2}, y + 1) - v_1(x + \frac{1}{2}, y), \quad x = 1, \dots, n_1 - 1, \right.$$

and similarly for the derivatives with respect to v_2 . Then we obtain

$$\begin{aligned} \mathcal{S}(v) = & \sum_{\mathbf{x} \in \mathcal{G}_p} \mu \left((D_{1,x}v_1)^2(\mathbf{x}) + (D_{2,y}v_2)^2(\mathbf{x}) \right) + \frac{\lambda}{2} \left((D_{1,x}v_1)(\mathbf{x}) + (D_{2,y}v_2)(\mathbf{x}) \right)^2 \\ & + \sum_{\mathbf{x} \in \mathcal{G}_d} \frac{\mu}{2} \left((D_{1,y}v_1)(\mathbf{x}) + (D_{2,x}v_2)(\mathbf{x}) \right)^2. \end{aligned} \tag{8}$$

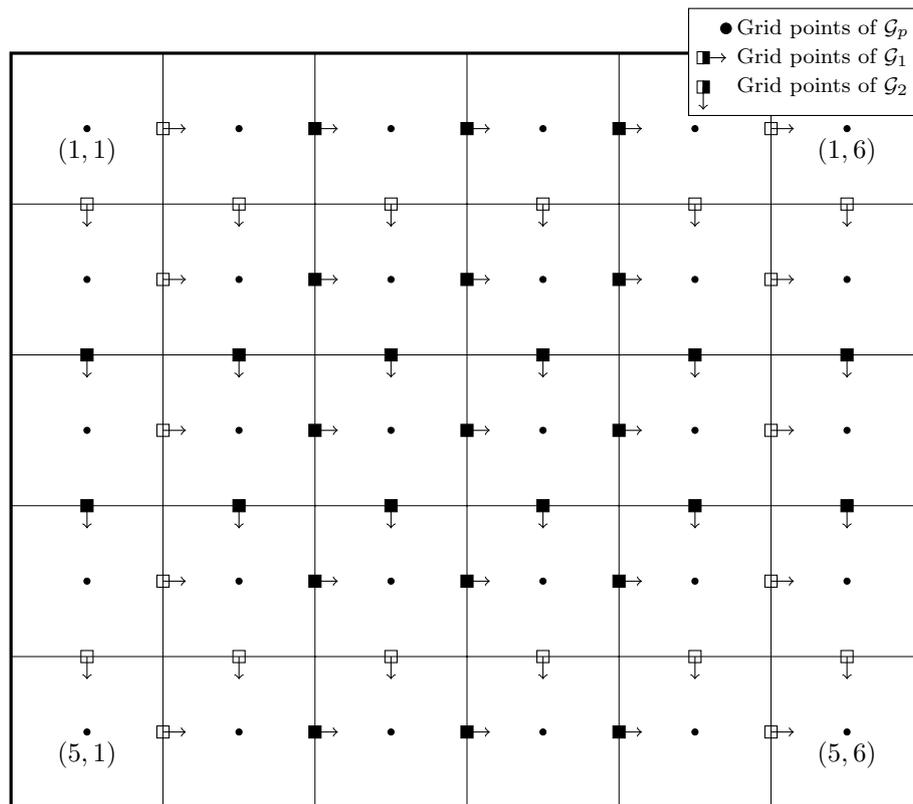


Figure 2. Illustration of the grids, where empty boxes mean zero movement.

To discretize the data term in (7) we need to approximate $T(\mathbf{x} - v(\mathbf{x}))$. Since v is not defined on \mathcal{G}_p we use instead of v its averaged version $Pv := \mathcal{G}_p \rightarrow \mathbb{R}^2$ given by

$$(P_1v_1)(x, y) := \begin{cases} 0, & x = 1 \\ \frac{1}{2} \left(v_1(x - \frac{1}{2}, y) + v_1(x + \frac{1}{2}, y) \right), & x = 2, \dots, n_1 - 1, \\ 0, & x = n_1 \end{cases}$$

and similarly for P_2v_2 in y -direction. Let the continuous function \tilde{T} on $\Omega := [1, n_1] \times [1, n_2]$ be constructed from given values of T on \mathcal{G}_p by interpolation. Here we restrict our attention to bilinear interpolations, see (A3), but more sophisticated approaches are possible. In summary, the discrete registration problem reads

$$\arg \min_v \mathcal{J}(v) := \sum_{\mathbf{x} \in \mathcal{G}_p} |R(\mathbf{x}) - \tilde{T}(\mathbf{x} - (Pv)(\mathbf{x}))|^2 + \mathcal{S}(v),$$

where $\mathcal{S}(v)$ is given by (8). We solve this nonlinear least squares problem by a Gauss-Newton like method described in the Appendix B.

Multilevel Strategy

As usual in optical flow and image registration we apply a coarse-to-fine strategy. First we iteratively smooth and downsample our given images. On the coarsest level we perform a registration to obtain a deformation on the coarse grid. We apply a bilinear interpolation v of this deformation to construct intermediate images on the finer level by

$$I_k(x) = \left(T + \frac{k}{\tilde{K}}(R \circ \varphi^{-1} - T) \right) \left(x - \frac{k}{\tilde{K}}v(x) \right), \quad \varphi(x) = x - v(x), \quad (9)$$

where T, R are the start and end images at the new level. Then we use the alternating algorithm on this level with $\tilde{K} \leq K$ images to obtain deformations and intermediate images on this level. Going to the finer level by bilinear interpolation, we construct more intermediate images by interpolating between neighboring images with (9). We repeat this process until we reach the maximal number K of images and finest level. The multilevel strategy of the algorithm is sketched in Algorithm 1.

Algorithm 1 Morphing Algorithm (informal).

- 1: $T_0 := T, R_0 := R, \Omega_0 := \Omega$
 - 2: create image stack $(T_0, \dots, T_l), (R_0, \dots, R_l)$ on $(\Omega_0, \dots, \Omega_l)$ by smoothing and downsampling
 - 3: solve (3) for T_l, R_l with $K = 1$, for \tilde{v}
 - 4: $l \rightarrow l - 1$
 - 5: use bilinear interpolation to get v on Ω_l from \tilde{v}
 - 6: obtain \tilde{K}_l images $\mathbf{I}_l^{(0)}$ from T_l, R_l, v by (9)
 - 7: **while** $l \geq 0$ **do**
 - 8: find image path $\tilde{\mathbf{I}}_l$ and deformation path \tilde{v}_l minimizing (3) with initialization $\mathbf{I}_l^{(0)}$
 - 9: $l \rightarrow l - 1$
 - 10: **if** $l > 0$ **then**
 - 11: use bilinear interpolation to get \mathbf{I}_l and \mathbf{v}_l on Ω_l
 - 12: **for** $k = 1, \dots, \tilde{K}_l$ **do**
 - 13: calculate \tilde{K}_l intermediate images between $I_{l,k-1}, I_{l,k}$ with $v_{l,k}$ using (9)
 - 14: $\mathbf{I} := \mathbf{I}_0$
-

3. Face Colorization

In this section, we describe a method to colorize a gray-scale image based on the morphing map between the luminance channel of a source image (template) and the present image (target). The idea consists in transferring of the chrominance channels from the source image to the target one. To this end, we work in the YUV color space. For the importance of the chrominance channels in color imaging, see, e.g., [41,42]. While in RGB images the color channels are highly correlated, the YUV space shows a suitable decorrelation between the luminance channel Y and the two chrominance channels U, V . The transfer from the RGB space to YUV space is given by (see, e.g., [5])

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ -0,14713 & -0,28886 & 0,436 \\ 0,615 & -0,51498 & -0,10001 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \quad (10)$$

Most of the colorization approaches are based on the hypothesis that the target image is the luminance channel of the desired image. Thus, the image colorization process is based on the computation of the unknown chrominance channels.

3.1. Luminance Normalization

The first step of the algorithm consists in transforming the RGB source image to the YUV image by (10). The range of the target gray-value image and the Y channel of the source image may differ making the meaningful comparison between these images not possible.

To tackle this issue, most of state-of-the art methods use a technique called *luminance remapping* which was introduced in [43]. This affine mapping between images which aims to fit the average and the standard deviation of the target and the template images is defined as

$$I_{\text{temp}} := \sqrt{\frac{\text{var}(I_{\text{tar}})}{\text{var}(I_Y)}} (I_Y - \text{mean}(I_Y)) + \text{mean}(I_{\text{tar}}),$$

where mean is the average of the pixel values, and var is the empirical variance.

3.2. Chrominance Transfer by the Morphing Maps

Next we compute the morphing map between the two gray-scale images I_{temp} and I_{tar} with model (3). This results in the deformation sequence φ which produces the resulting map from the template image to the target one by concatenation

$$\Phi = \varphi_1 \circ \varphi_2 \circ \dots \circ \varphi_K.$$

Due to the discretization of the images, the map Φ is defined, for images of size $n_1 \times n_2$, on the discrete grid $\mathcal{G} := \{1, \dots, n_1\} \times \{1, \dots, n_2\}$:

$$\Phi : \mathcal{G} \rightarrow [1, n_1] \times [1, n_2], \quad x \mapsto \Phi(x),$$

where $\Phi(x)$ is the position in the source image which corresponds to the pixel $x \in \mathcal{G}$ in the target image. Now we colorize the target image by computing its chrominance channels, denoted by $(U_{\text{tar}}(x), V_{\text{tar}}(x))$ at position x as

$$(U_{\text{tar}}(x), V_{\text{tar}}(x)) := (U(\Phi(x)), V(\Phi(x))). \tag{11}$$

The chrominance channels of the target image are defined on the image grid \mathcal{G} , but usually $\Phi(x) \notin \mathcal{G}$. Therefore the values of the chrominance channels at $\Phi(x)$ have to be computed by interpolation. In our algorithm we use just bilinear interpolation which is defined for $\Phi(x) = (p, q)$ with $(p, q) \in [i, i + 1] \times [j, j + 1]$, $(i, j) \in \{1, \dots, n_1 - 1\} \times \{1, \dots, n_2 - 1\}$ by

$$U(\Phi(x)) = U(p, q) := (i + 1 - p, p - i) \begin{pmatrix} U(i, j) & U(i, j + 1) \\ U(i + 1, j) & U(i + 1, j + 1) \end{pmatrix} \begin{pmatrix} j + 1 - q \\ q - j \end{pmatrix}.$$

Finally, we compute a colorized RGB image from its luminance $I_{\text{tar}} = Y_{\text{tar}}$ and the chrominance channels (11) by the inverse of (10):

$$\begin{pmatrix} R(x) \\ G(x) \\ B(x) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1,13983 \\ 1 & -0,39465 & -0,58060 \\ 1 & 2,03211 & 0 \end{pmatrix} \begin{pmatrix} Y_{\text{tar}}(x) \\ U_{\text{tar}}(x) \\ V_{\text{tar}}(x) \end{pmatrix} \tag{12}$$

for all $x \in \mathcal{G}$. Figure 3 (Image segements with unified background of George W. Bush <https://commons.wikimedia.org/wiki/File:George-W-Bush.jpeg> and Barack Obama https://commons.wikimedia.org/wiki/File:Barack_Obama.jpg both in public domain.) summarizes our color transfer method.

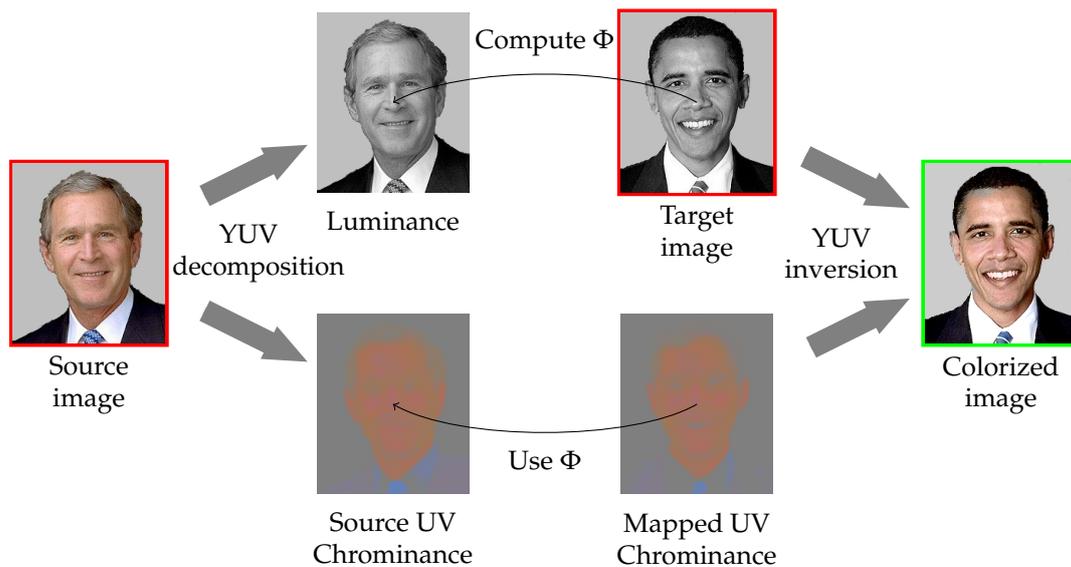


Figure 3. Overview of the color transfer. The mapping Φ is computed from Model (3) between the luminance channel of the source image and the target one. From this map, the chrominances of the source image are mapped. Finally, from these chrominances and the target image the colorization result is computed.

4. Variational Methods for Chrominance Postprocessing

Sometimes the color transfer computed from the morphing map can be disturbed by artifacts. To improve the results, post-processing steps are usually applied in the literature.

Variational approaches are frequently applied in image colorization either directly or as a post-processing step, see, e.g., [1,3,44]. For instance, the technique of Gupta et al. [6] uses the chrominance diffusion approach of Levin et al. [1].

In this paper, we propose a variational method with a total variation based regularization as a post-processing step to remove possible artifacts. We build up on the model [3] suggested by one of the authors. This variational model uses a functional with a specific regularization term to avoid “halo effects”. More precisely, we consider the minimizer of

$$\hat{u} = (\hat{U}, \hat{V}) = \arg \min_{(U,V)} \text{TV}_{Y_{\text{tar}}}(U, V) + \alpha \int_{\Omega} |U(x) - U_{\text{tar}}(x)|^2 + |V(x) - V_{\text{tar}}(x)|^2 dx, \quad (13)$$

with

$$\text{TV}_{Y_{\text{tar}}}(U, V) := \int_{\Omega} \sqrt{\gamma |\nabla Y_{\text{tar}}|^2 + |\nabla U|^2 + |\nabla V|^2} dx.$$

The first term in (13) is a coupled total variation term which enforces the chrominance channels to have a contour at the same location as the target gray-value image. The data fidelity term is the classical squared L_2 -norm of the differences of the given and the desired chrominance channels. Note that the model in [3] contains an additional box constraint.

The parameter γ manages the coupling of the chrominance channels with the luminance one. It has been shown in [3] that a parameter value around 25 can be used for most images. The parameter α is related to the object size in the images. If some colors on large objects leak on small ones, this parameter has to be increased. On the other hand, if some artifacts are still present in the result, the value of α has to be decreased. We apply the primal-dual Algorithm 2 to find the minimizer of the strictly convex model (13). It uses an update on the step time parameters τ and σ , as proposed by Chambolle and Pock [45], as well as a relaxation parameter θ to speed-up the convergence. Here we use the abbreviation $b := (U_{\text{tar}}, V_{\text{tar}})$ and $u := (U, V)$. Further, p is the dual variable which is pixel-wise

in \mathbb{R}^6 . The parameters τ and σ are intern time step sizes. The operator div stands for the discrete divergence and ∇ for the discrete gradient. Further, the proximal mapping $P_{\mathcal{B}}$ is given pixel-wise, for $p \in \mathbb{R}^6$ by

$$P_{\mathcal{B}}(p) = \frac{\hat{p}}{\max\left(1, \|\hat{p}\|_2^2\right)}, \quad \text{where } \hat{p} := p - \sigma \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \partial_x Y \\ \partial_y Y \end{pmatrix}. \tag{14}$$

Algorithm 2 Minimization of (13).

- 1: $u^0 \leftarrow b, \bar{u}^0 \leftarrow u^0$
 - 2: $p^0 \leftarrow \nabla u^0$
 - 3: $\sigma \leftarrow 0.001, \tau \leftarrow 20$
 - 4: **for** $n > 0$ **do**
 - 5: $p^{n+1} \leftarrow P_{\mathcal{B}}(p^n + \sigma \nabla \bar{u}^n)$
 - 6: $u^{n+1} \leftarrow \frac{u^n + \tau (\text{div}(p^{n+1}) + \alpha b)}{1 + \tau \alpha}$
 - 7: $\theta = 1/\sqrt{1 + \tau \alpha}$
 - 8: $\tau = \theta \tau \quad \sigma = \sigma/\theta$
 - 9: $\bar{u}^{n+1} \leftarrow u^{n+1} + \theta(u^{n+1} - u^n)$
 - 10: $\hat{u} \leftarrow u^{+\infty}$.
-

As mentioned in the work of Deledalle et al. [38], the minimization of the TV- L_2 model produces a biased result. This bias causes a lost of contrast in the case of gray-scale images, whereas it is visible as a lost of colorfulness in the case of model (13). The authors of [38] describe an algorithm to remove such bias. In this paper, we propose to modify this method for our model (13) in order to enhance the result of Algorithm 2.

The CLEAR method of [38] relies on the *refitting estimator* $\mathcal{R}_{\hat{x}}(y)$ of the data y from the biased estimation $\hat{x}(y)$:

$$\mathcal{R}_{\hat{x}}(y) \in \arg \min_{h \in \mathcal{H}_{\hat{x}}} \|h(y) - y\|_2^2 \tag{15}$$

where $\mathcal{H}_{\hat{x}}$ is defined as the set of maps $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$ satisfying, $\forall y \in \mathbb{R}^n$:

$$h(y) = \hat{x}(y) + \rho J_{\hat{x}(y)}(y - \hat{x}(y)), \text{ with } \rho \in \mathbb{R},$$

where $J_{\hat{x}(y)}$ is the Jacobian of the biased estimator with respect to the data y :

$$J_{\hat{x}(y)}d = \lim_{\varepsilon \rightarrow 0} \frac{\hat{x}(y + \varepsilon d) - \hat{x}(y)}{\varepsilon}.$$

$J_{\hat{x}(y)}$ contains some structural information of the biased estimator \hat{x} such as the jumps. A closed formula for ρ can be given:

$$\rho = \begin{cases} \frac{\langle J_{\hat{x}(y)}(\delta) | \delta \rangle}{\|J_{\hat{x}(y)}(\delta)\|_2^2} & \text{if } J_{\hat{x}(y)}(\delta) \neq 0, \\ 1 & \text{otherwise.} \end{cases}$$

where $\delta = y - \hat{x}(y)$. With Equation (15), the refitting process is as closer as possible to the original data.

The final algorithm is summarized in Algorithm 3. Note that it uses the result \hat{u} of Algorithm 1 as an input. The proximal mapping $\Pi_p(\tilde{p})$ within the algorithm is defined pixel-wise, for variables $p \in \mathbb{R}^6$ and $\tilde{p} \in \mathbb{R}^6$, as

$$\Pi_p(\tilde{p}) = \begin{cases} \hat{p} & \text{if } \|\tilde{p}\| < 1 \\ \frac{1}{\|\hat{p}\|} \left(\hat{p} - \frac{\langle \hat{p}, \tilde{p} \rangle}{\|\hat{p}\|^2} \hat{p} \right) & \text{otherwise,} \end{cases}$$

where \hat{p} and \hat{p} are defined as in (14).

Algorithm 3 Debiasing of Algorithm 2.

- 1: $u^0 = b, \bar{u}^0 = b$
 - 2: $\delta \leftarrow b - \hat{u}$
 - 3: $\tilde{u}^0 = \delta, \bar{\tilde{u}}^0 = \delta$
 - 4: $p^0 \leftarrow \nabla u, \tilde{p}^0 \leftarrow \nabla \tilde{u}$
 - 5: $\sigma \leftarrow 0.001, \tau \leftarrow 20$
 - 6: **for** $n \geq 0$ **do**
 - 7: $p^{n+1} \leftarrow P_{\mathcal{B}}(p^n + \sigma \nabla \bar{u}^n)$
 - 8: $\tilde{p}^{n+1} \leftarrow \Pi_{p^n + \sigma \nabla \bar{u}^n}(\tilde{p}^n + \sigma \nabla \bar{\tilde{u}}^n)$
 - 9: $u^{n+1} \leftarrow \frac{u^n + \tau (\text{div}(p^{n+1}) + \alpha b)}{1 + \tau \alpha}$
 - 10: $\tilde{u}^{n+1} \leftarrow \frac{\tilde{u}^n + \tau (\text{div}(\tilde{p}^{n+1}) + \alpha \delta)}{1 + \tau \alpha}$
 - 11: $\theta = 1/\sqrt{1 + \tau \alpha}$
 - 12: $\tau = \theta \tau \quad \sigma = \sigma/\theta$
 - 13: $\bar{u}^{n+1} \leftarrow u^{n+1} + \theta(u^{n+1} - u^n)$
 - 14: $\bar{\tilde{u}}^{n+1} \leftarrow \tilde{u}^{n+1} + \theta(\tilde{u}^{n+1} - \tilde{u}^n)$
 - 15: $\rho \leftarrow \begin{cases} \frac{\langle \tilde{u}^{+\infty}, \delta \rangle}{\|\tilde{u}^{+\infty}\|_2^2} & \text{if } \tilde{u}^{+\infty} \neq 0 \\ 1 & \text{otherwise.} \end{cases}$,
 - 16: $u_{\text{debaised}} \leftarrow \hat{u} + \rho \tilde{u}^{+\infty}$.
-

The results obtained at the different steps of the work-flow are presented for a particular image in Figure 4. First we demonstrate in Figure 4c that simply transforming the R, G, B channels via the morphing map gives no meaningful results. Letting our morphing map act to the chrominance channels of our source image and applying (12) with the luminance of our target image we get, e.g., Figure 4d which is already a good result. However, the forehead of Obama contains an artifact; a gray unsuitable color is visible here. After a post-processing of the chrominance channels by our variational method the artifacts disappear as can be seen in Figure 4e.

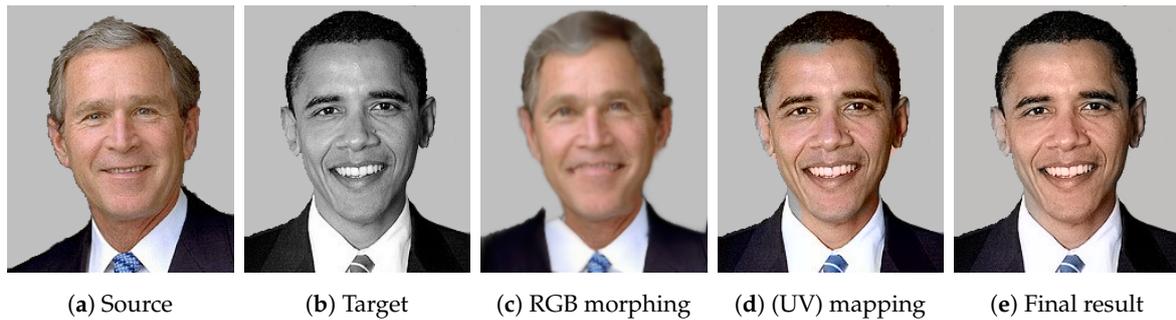


Figure 4. Illustration of the colorization steps of our algorithm. (a) The color source image; (b) The gray value target image; (c) The transport of the R, G, B channel via the morphing map is not suited for colorization; (d) The result with our morphing method is already very good; (e) It can be further improved by our variational post-processing.

5. Numerical Examples

In this section, we compare our method with

- the patch-based algorithm of Welsh et al. [8],
- the patch-based method of Pierre et al. [3], and
- the segmentation approach of Gupta et al. [6].

We implemented our morphing algorithm in Matlab 2016b and used the Matlab intern function for the bilinear interpolation. The parameters are summarized in Table 1. Here $\lambda = \mu$ and K are the parameters for the morphing step. The parameters γ and α appear in the variational model for post-processing and were only applied in three images. The number of deformations K depends on the size and details of the image, i.e., large detailed images need a larger K as small image having not so much details. Usually $K = 24$ was a good choice for our experiments. The parameter λ controls the smoothness of the deformation, i.e., a small value leads to large movements which might not be reasonable, while for large values the deformations become zero and the image metamorphosis is just a blending. In our experiments, a value of magnitude 10^{-2} lead to good results.

Table 1. Parameters for our numerical experiments.

Image	μ	K	γ	α
Figure 5-1. row	0.025	24	50	0.005
Figure 5-2. row	0.05	24	25	0.005
Figure 5-3. row	0.05	12	-	-
Figure 5-4. row	0.05	24	-	-
Figure 6-1. row	0.005	32	-	-
Figure 6-2. row	0.0075	18	50	0.05
Figure 6-3. row	0.04	24	-	-
Figure 7	0.0075	18	-	-
Figure 8	0.01	25	-	-
Figure 9-1. row	0.005	25	-	-
Figure 9-2. row	0.01	25	-	-
Figure 9-3. row	0.01	25	-	-



Figure 5. Comparison of our approach with state-of-the-art methods on photographs. In contrast to these methods our model is not based on texture comparisons, but on the morphing of the shapes. Therefore it is able to handle faces images, where the background has frequently a similar texture as the skin.

First we compare portraits in Figure 5 (Image segments of Steve Jobs https://commons.wikimedia.org/wiki/File:Steve_Jobs_Headshot_2010-CROP.jpg image courtesy of Matthew Yohe, Albert Einstein https://commons.wikimedia.org/wiki/File:Einstein_Portr_05936.jpg in public domain, Catherine Deneuve https://commons.wikimedia.org/wiki/File:Catherine_Deneuve_2010.jpg image courtesy of Georges Biard and Renée Deneuve https://commons.wikimedia.org/wiki/File:Ren%C3%A9e_Deneuve.jpg in public domain.) starting with the modern photographs in the first row. The approach of Welsh et al. [8] is based on a patch matching between images. The patch comparison is done with basic texture descriptors (intensity of the central pixel and standard-deviation of the patches). Since the background, as well as the skin are smooth, the distinction between them is unreliable if their intensities are similar. Moreover, since no regularization is used after the color transfer, some artifacts occur. For instance, some blue points appear on Obama's face, see Figure 5, first row. The approach of Pierre et al. [3] is based on more sophisticated texture features for patches and applies a variational model with total variation like regularization. With this approach the artifacts mentioned above are less visible. Nevertheless, the forehead of Obama is purple which is unreliable. The method of Gupta et al. [6] uses texture descriptors after an over-segmentation, see, e.g., SLIC [10]. The texture

descriptors are based on SURF and Gabor features. In the case of the Obama image, the descriptors are not able to distinguish the skin and other smooth parts, leading to a background color different from the source image. Our method is able to colorize the second image in a more reasonable way, i.e., face and background color are different and the tie gets a blue color. However, our method is not perfect so far. For example part of the forehead of Obama becomes gray is due to the gray hair of Bush, which has a gray value closer to the forehead as to the hair of Obama.

The second and the third rows of Figure 5 focus on the colorization of old photographs. This challenging problem is a real application of image colorization which is sought, e.g., by the cinema industry. Note that the texture of old images are disturbed by the natural grain which is not the case in modern photography. Thus, the texture comparison is unreliable for this application. This issue is visible in all the comparison methods. For the portrait of Einstein the background is not colorized with the color of the source. Moreover, the color of the skin is different from those of the person in the source image. For the picture of Deneuve, the color of her lips is not transferred to the target image (Deneuve's mother) with the state-of-the-art texture-based algorithms. With our morphing approach, the shapes of the faces are mapped. Thus, the lips, as well as the eyes and the skin are well colorized with a color similar to the source image. In the last row we have two images of the same person. Here the state-of-the-art-texture-based methods give unlikely results, especially the methods of Welsh et al. and Gupta et al. lead to a non smooth colorization of the background and the face, respectively. Our method provides a reasonable result with only small artifacts around the glasses.

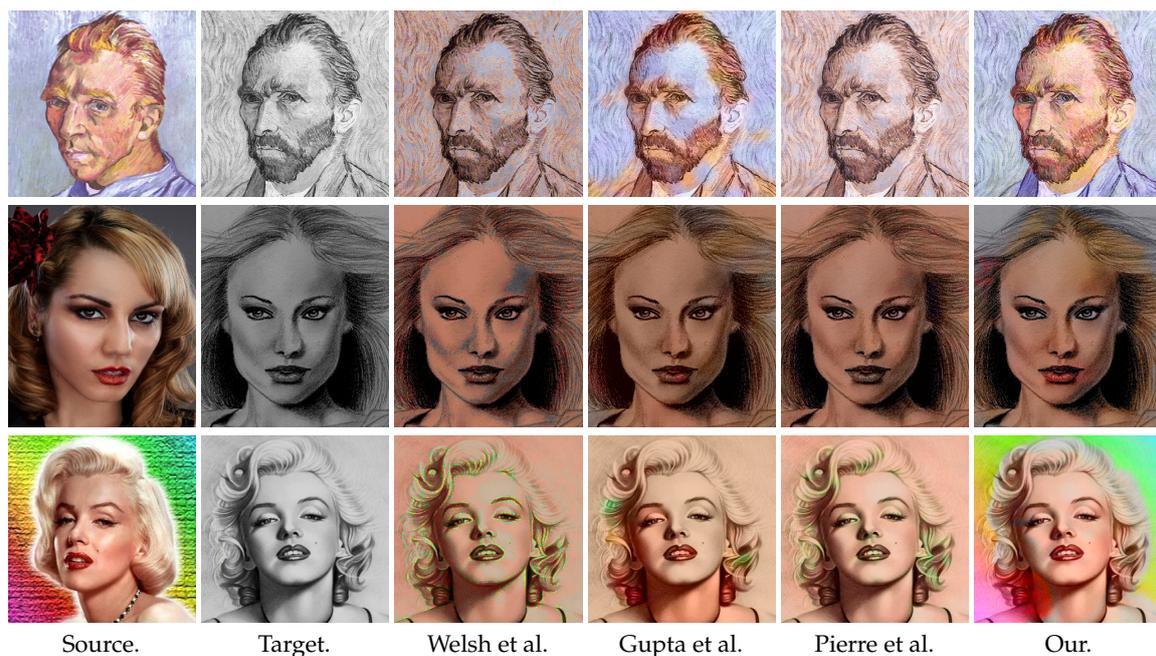


Figure 6. Results including painted faces. Only our morphing method is able to colorize the target images in an authentic way.

In Figure 6 (Image segments of self-portraits of Vincent van Gogh https://commons.wikimedia.org/wiki/File:Vincent_Willem_van_Gogh_102.jpg and https://commons.wikimedia.org/wiki/File:Vincent_van_Gogh_-_Self-Portrait_-_Google_Art_Project.jpg both in public domain, a photography of a woman <https://pixabay.com/en/woman-portrait-face-model-canon-659352/> licensed CC0, a drawing of a woman <https://pixabay.com/en/black-and-white-drawing-woman-actor-1724363/> licensed CC0, a color image of Marilyn Monroe <https://www.flickr.com/photos/7477245@N05/5272564106> created by Luiz Fernando Reis, and a drawing of Marilyn Monroe <https://pixabay.com/en/marilyn-monroe-art-draw-marilyn-885229/> licensed CC0), we provide results including painted images. Note that we use the same Van Gogh self-portraits as in [31]. Due to the low contrast of the

ear and suit to the background we add here the same segmentation information as in [31], which means our images are two dimensional during the calculation for the results shown in the first row of Figure 6. In these examples the similarity of the shapes between the source and target images is again more reliable than the matching of the textures so that only our morphing approach produces suitable results. Consider in particular the lips of the woman in the second and third row. A non post-processed result for the woman in the second row is shown in Figure 7. Comparing the two images we see that only small details change but most of the colorization is done by the morphing.

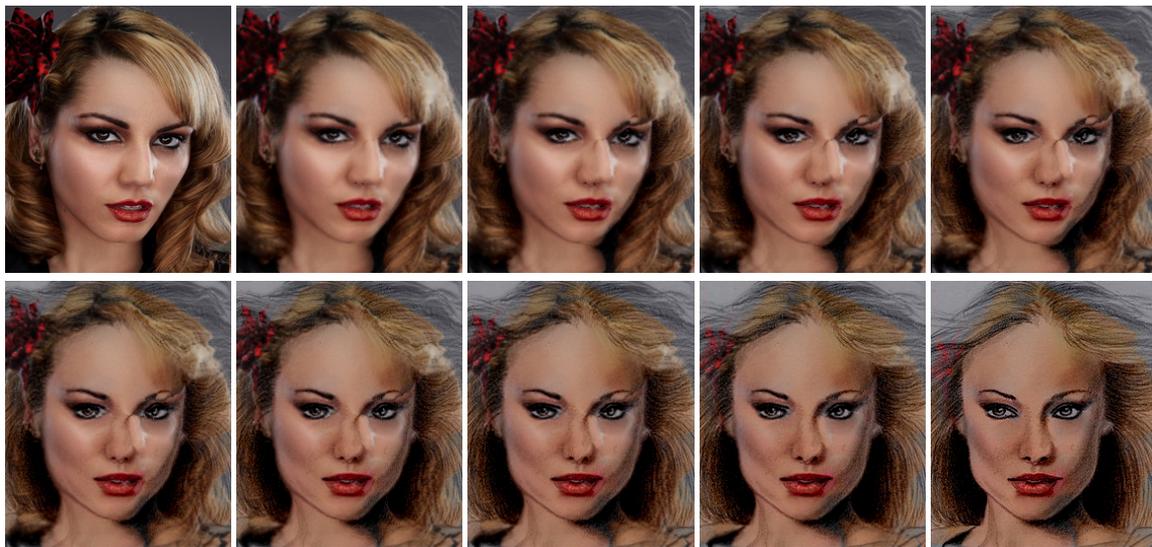


Figure 7. Color transport along the image path.

Figure 7 visualizes the path of the color transfer. As the morphing approach calculates at the same time an image and mapping path, we can not only show the final colorization result, but also the way the color is transported along this path. We illustrate this by the image in second row of Figure 6. The UV channels are transported via the mapping path of the Y channels of the source and the target images, where every second image along the path is shown. We see that even though the right part of the images undergoes large transformations, the eyes and the mouth are moved to the correct places. Note that the final image does not contain a post-processing, in contrast to those in the second row of Figure 6. However, the result is already quite good.

Figure 8 is included to evaluate the quality of our colorization method. We took two photographs of the same person in different poses. Assuming that the color of the person does not change due to the movement, we have a ground truth target image and can do a quantitative comparison using the PSNR. We see that the visual superiority of our result is confirmed by the much higher PSNR.

In Figure 9 we considered three RGB image pairs. Using the luminance of one of the images as target and the other one as source and conversely we colorized the target image and computed the PSNR. In Table 2 the PSNR values of the colorization with our method and the state-of-the-art methods are shown. The performance of our method is very convincing for all these example images.

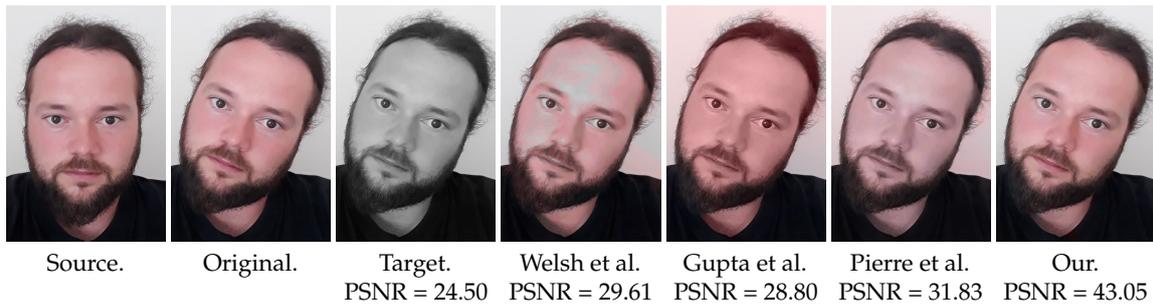


Figure 8. Results on a color image turned into a gray-scale one for a quantitative comparison. The qualitative comparisons with the state-of-the-art methods are confirmed by the PSNR measures.

Table 2. Comparison of the different PSNR values for the image pairs in Figure 9.

	Gray	Welsh et al. [8]	Gupta et al. [6]	Pierre et al. [3]	Our
Figure 9-1. row 1. pair	24.8023	20.0467	26.3527	33.7694	44.7808
Figure 9-1. row 2. pair	24.5218	23.9513	25.9457	34.4231	45.4682
Figure 9-2. row 1. pair	24.3784	22.6729	27.5586	32.0119	41.1413
Figure 9-2. row 2. pair	23.7721	23.2375	25.9375	30.1398	39.4254
Figure 9-3. row 1. pair	24.5950	30.3985	24.3112	31.5263	42.3861
Figure 9-3. row 2. pair	24.3907	27.7816	25.6207	31.8982	42.4092

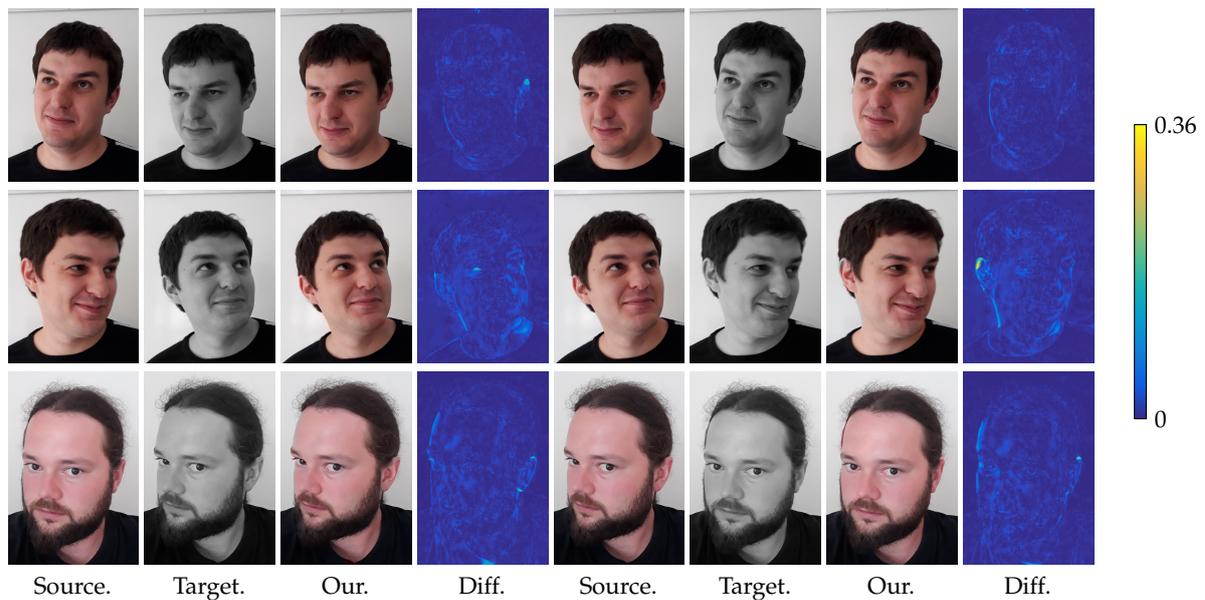


Figure 9. Multiple colorizations of known RGB-images with difference plots measured in Euclidean distance in \mathbb{R}^3 .

Remark (Limitations of our approach). *In all our experiments our method leads to good results and outperforms the state-of-the-art colorization methods. However, the method is not perfect and the face images can not be arbitrarily deformed. Since the morphing is only done in the Y channel, the illumination of the faces should be similar. For example there is a shadow on the left side of the images in the second row of Figure 5. Mirroring one of the images would lead to bad matching results. The same holds true for the features themselves, so we can not expect to match bright hair perfectly to dark hair, as we see for example in Figure 4d.*

Since the deformations are close to diffeomorphisms, features should not be excluded or appear. For example, the left eye in the last row of Figure 9 starts in the middle of the face and moves to the edge, so a part of the

skin-color stays in the eye. Similarly in the second matching here the eye moves from the edge to the inside, so the skin obtains the color from the background.

By a proper initialization the faces should be aligned in such a way that there are no similar features atop of each other, e.g., a left eye in the source should not be atop of the right eye of the target. However with more prior knowledge we could use the given features as landmarks and perform a landmark matching, e.g., [46], for the initialization on the coarsest level.

6. Conclusions

In this paper, we propose an finite difference method to compute a morphing map between a source and a target gray-value image. This map enables us to transfer the color from the source image to the target one, based on the shapes of the faces. An additional variational post-processing step with a luminance guided total variation regularization and an update to make the result unbiased may be added to remove some possible artifacts. The results are very convincing and outperform state-of-the-art approaches on historical photographs.

Let us notice some special issues of our approach in view of an integration into a more global framework for an exemplar-based image and video colorization. First of all, our method works well on faces and on object with similar shapes, but when this hypothesis is not fulfilled, some artifacts can appear. Therefore, instead of using our algorithm on full images, a face detection algorithm can be used to focus on the face colorization. Let us remark that faces in image can be detected by efficient, recent methods, see, e.g., [47]. In future work, the method will be integrated into a complete framework for exemplar-based image and video colorization.

Second, the morphing map has to be computed between images with the same size. This issue can be easily solved with a simple interpolation of the images. Keeping the ratio between the width and the height of faces similar, the distortion produced by such interpolation is small enough to support our method.

Acknowledgments: Funding by the German Research Foundation (DFG) within the project STE 571/13-1 is gratefully acknowledged.

Author Contributions: All the three authors have substantially contributed to the paper. The paper is the outcome of many joint and fruitful discussions J. Persch and G. Steidl had with F. Pierre during the months the later was a Postdoc at the University at Kaiserslautern (Germany). Therefore it is not possible to split the contributions. However, J. Persch was the leading expert concerning morphing and F. Pierre concerning color image processing.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Solution of the Image Sequence Problem

Fixing \mathbf{v} , resp., φ leads to the image sequence problem (6). In the following we show how this problem can be solved via the linear system of equations arising from Euler-Lagrange equation of the functional. We mainly follow [31].

The terms in $\mathcal{J}_\varphi(\mathbf{I})$ containing I_k are

$$\int_{\Omega} |I_k(\varphi_{k+1}(\mathbf{x})) - I_{k+1}(\mathbf{x})|^2 + |I_k(\mathbf{x}) - I_{k-1}(\varphi_k(\mathbf{x}))|^2 \, d\mathbf{x}.$$

Assuming that φ_{k+1} is a diffeomorphism, the variable transform $\mathbf{y} := \varphi_{k+1}(\mathbf{x})$ gives

$$\begin{aligned} & \int_{\Omega} |I_k(\mathbf{x}) - I_{k+1}(\varphi_{k+1}^{-1}(\mathbf{x}))|^2 |\det \nabla \varphi_{k+1}^{-1}(\varphi_{k+1}^{-1}(\mathbf{x}))| + |I_k(\mathbf{x}) - I_{k-1}(\varphi_k(\mathbf{x}))|^2 \, d\mathbf{x} \\ &= \int_{\Omega} |I_k(\mathbf{x}) - I_{k+1}(\varphi_{k+1}^{-1}(\mathbf{x}))|^2 |\det \nabla \varphi_{k+1}(\varphi_{k+1}^{-1}(\mathbf{x}))|^{-1} + |I_k(\mathbf{x}) - I_{k-1}(\varphi_k(\mathbf{x}))|^2 \, d\mathbf{x} \end{aligned}$$

Setting

$$a_k(\varphi_{k+1}^{-1}(\mathbf{x})) := |\det \nabla \varphi_{k+1}(\varphi_{k+1}^{-1}(\mathbf{x}))|^{-1},$$

the regularizing term can be rewritten as

$$\mathcal{S}(\mathbf{v}) = \|\mathbf{S}\mathbf{v}\|_2^2, \quad \mathbf{S} := \begin{pmatrix} \sqrt{\mu} \mathbf{D}_{1,x} & 0 \\ 0 & \sqrt{\mu} \mathbf{D}_{2,y} \\ \sqrt{\frac{\mu}{2}} \mathbf{D}_{1,y} & \sqrt{\frac{\mu}{2}} \mathbf{D}_{2,x} \\ \sqrt{\frac{\lambda}{2}} \mathbf{D}_{1,x} & \sqrt{\frac{\lambda}{2}} \mathbf{D}_{2,y} \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}.$$

The nonlinear data term in the registration problem with the columnwise reshaped image $\mathbf{T}(\mathbf{v})$ of $(\tilde{T}(\mathbf{x} - (Pv)(\mathbf{x})))_{\mathbf{x} \in \mathcal{G}_p}$ and \mathbf{R} of $(R(\mathbf{x}))_{\mathbf{x} \in \mathcal{G}_p}$ reads

$$\mathcal{D}(\mathbf{v}) := \|\mathbf{R} - \mathbf{T}(\mathbf{v})\|^2.$$

We linearize the term at a previous iterate $\mathbf{v}^{(r)}$, i.e.,

$$\mathcal{D}(\mathbf{v}) \approx \|\mathbf{R} - \mathbf{T}(\mathbf{v}^{(r)}) + \underbrace{\nabla_{\mathbf{v}} \mathbf{T}(\mathbf{v}^{(r)})}_{\mathbf{G}^{(r)}} (\mathbf{v} - \mathbf{v}^{(r)})\|_2^2.$$

We comment on the computation of $\mathbf{G}^{(r)}$ below. Then we consider

$$\|\mathbf{R} - \mathbf{T}(\mathbf{v}^{(r)}) + \mathbf{G}^{(r)} (\mathbf{v} - \mathbf{v}^{(r)})\|_2^2 + \|\mathbf{S}\mathbf{v}\|_2^2.$$

The gradient is given by

$$\left((\mathbf{G}^{(r)})^T \mathbf{G}^{(r)} + \mathbf{S}^T \mathbf{S} \right) \mathbf{v} - (\mathbf{G}^{(r)})^T \mathbf{G}^{(r)} \mathbf{v}^{(r)} - (\mathbf{G}^{(r)})^T (\mathbf{R} - \mathbf{T}(\mathbf{v}^{(r)})).$$

Setting the gradient to zero we obtain for \mathbf{v}

$$\begin{aligned} \mathbf{v} &:= \mathbf{v}^{(r)} - \left((\mathbf{G}^{(r)})^T \mathbf{G}^{(r)} + \mathbf{S}^T \mathbf{S} \right)^{-1} \left((\mathbf{G}^{(r)})^T (\mathbf{T}(\mathbf{v}^{(r)}) - \mathbf{R}) + \mathbf{S}^T \mathbf{S} \mathbf{v}^{(r)} \right) \\ &= \mathbf{v}^{(r)} - \left((\mathbf{G}^{(r)})^T \mathbf{G}^{(r)} + \mathbf{S}^T \mathbf{S} \right)^{-1} \nabla_{\mathbf{v}} \mathcal{J}(\mathbf{v}^{(r)}) \end{aligned}$$

As we want to minimize \mathcal{J} , the next iterate is given by

$$\mathbf{v}^{(r+1)} = \mathbf{v}^{(r)} - \left(\frac{1}{2}\right)^l \left((\mathbf{G}^{(r)})^T \mathbf{G}^{(r)} + \mathbf{S}^T \mathbf{S} \right)^{-1} \nabla_{\mathbf{v}} \mathcal{J}(\mathbf{v}^{(r)}),$$

where $l \in \mathbb{N}$ is the smallest number such that $\mathcal{J}(\mathbf{v}^{(k+1)}) < \mathcal{J}(\mathbf{v}^{(k)})$, if $\nabla_{\mathbf{v}} \mathcal{J}(\mathbf{v}^{(r)})$ is not zero.

Finally we comment on the computation of

$$\mathbf{G}^{(r)} := \nabla_{\mathbf{v}} \mathbf{T}(\mathbf{v}^{(r)}) = \begin{pmatrix} -P_1^T \partial_x \tilde{T}(x - (P_1 v_1)(\mathbf{x}), y - P_2 v_2(\mathbf{x})) \\ -P_2^T \partial_y \tilde{T}(x - (P_1 v_1)(\mathbf{x}), y - P_2 v_2(\mathbf{x})) \end{pmatrix}.$$

Since \tilde{T} is computed by linear interpolation from $(T(\mathbf{x}))_{\mathbf{x} \in \mathcal{G}_p}$, Then we have for an arbitrary point $(p, q) \in [i, i+1] \times [j, j+1]$, $(i, j) \in \{1, \dots, n_1 - 1\} \times \{1, \dots, n_2 - 1\}$ that

$$\tilde{T}(p, q) := (i+1-p, p-i) \begin{pmatrix} T(i, j) & T(i, j+1) \\ T(i+1, j) & T(i+1, j+1) \end{pmatrix} \begin{pmatrix} j+1-q \\ q-j \end{pmatrix}. \tag{A3}$$

Then

$$\begin{aligned}\partial_x \tilde{T}(p, q) &= (j + 1 - q) (T(i + 1, j) - T(i, j)) + (q - j) (T(i + 1, j + 1) - T(i, j)), \\ \partial_y \tilde{T}(p, q) &= (i + 1 - p) (T(i, j + 1) - T(i, j)) + (p - i) (T(i + 1, j + 1) - T(i + 1, j)).\end{aligned}$$

Hence the derivatives can be calculated from the forward differences of T in x - resp. y -direction and appropriate weighting.

References

1. Levin, A.; Lischinski, D.; Weiss, Y. Colorization using optimization. *ACM Trans. Graph.* **2004**, *23*, 689–694.
2. Irony, R.; Cohen-Or, D.; Lischinski, D. Colorization by example. In Proceedings of the 16th Eurographics Conference on Rendering Techniques, Konstanz, Germany, 29 June 29–1 July 2005; pp. 201–210.
3. Pierre, F.; Aujol, J.F.; Bugeau, A.; Papadakis, N.; Ta, V.T. Luminance-chrominance model for image colorization. *SIAM J. Imaging Sci.* **2015**, *8*, 536–563.
4. Yatziv, L.; Sapiro, G. Fast image and video colorization using chrominance blending. *IEEE Trans. Image Process.* **2006**, *15*, 1120–1129.
5. Jack, K. *Video Demystified: A Handbook for the Digital Engineer*; Elsevier: Amsterdam, The Netherlands, 2011.
6. Gupta, R.K.; Chia, A.Y.S.; Rajan, D.; Ng, E.S.; Zhiyong, H. Image colorization using similar images. In Proceedings of the 20th ACM International Conference on Multimedia, Nara, Japan, 29 October–2 November 2012; pp. 369–378.
7. Horiuchi, T. Colorization algorithm using probabilistic relaxation. *Image Vis. Comput.* **2004**, *22*, 197–202.
8. Welsh, T.; Ashikhmin, M.; Mueller, K. Transferring color to greyscale images. *ACM Trans. Graph.* **2002**, *21*, 277–280.
9. Efros, A.A.; Leung, T.K. Texture synthesis by non-parametric sampling. In Proceedings of the 7th IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; Volume 2, pp. 1033–1038.
10. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282.
11. Chen, T.; Wang, Y.; Schillings, V.; Meinel, C. Grayscale image matting and colorization. In Proceedings of the Asian Conference on Computer Vision, Jeju Island, Korea, 27–30 January 2004; pp. 1164–1169.
12. Charpiat, G.; Hofmann, M.; Schölkopf, B. Automatic image colorization via multimodal predictions. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 126–139.
13. Chia, A.Y.S.; Zhuo, S.; Kumar, R.G.; Tai, Y.W.; Cho, S.Y.; Tan, P.; Lin, S. Semantic colorization with internet images. *ACM Trans. Graph.* **2011**, *30*, 156.
14. Zhang, R.; Isola, P.; Efros, A.A. Colorful image colorization. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 1–16.
15. Smythe, D.B. *A Two-Pass Mesh Warping Algorithm for Object Transformation and Image Interpolation*; Technical report; ILM Technical Memo Department, Lucasfilm Ltd.: San Francisco, CA, USA, 1990.
16. Wolberg, G. *Digital Image Warping*; IEEE Computer Society Press: Los Alamitos, CA, USA, 1990; Volume 10662.
17. Wolberg, G. Image morphing: A survey. *Vis. Comput.* **1998**, *14*, 360–372.
18. Miller, M.I.; Younes, L. Group actions, homeomorphisms, and matching: A general framework. *Int. J. Comput. Vis.* **2001**, *41*, 61–84.
19. Trounev, A.; Younes, L. Local geometry of deformable templates. *SIAM J. Math. Anal.* **2005**, *37*, 17–59.
20. Trounev, A.; Younes, L. Metamorphoses through Lie group action. *Found. Comput. Math.* **2005**, *5*, 173–198.
21. Christensen, G.E.; Rabbitt, R.D.; Miller, M.I. Deformable templates using large deformation kinematics. *IEEE Trans. Image Process.* **1996**, *5*, 1435–1447.
22. Dupuis, P.; Grenander, U.; Miller, M.I. Variational problems on flows of diffeomorphisms for image matching. *Q. Appl. Math.* **1998**, *56*, 587–600.
23. Trounev, A. An infinite dimensional group approach for physics based models in pattern recognition. *Int. J. Comput. Vis.* **1995**, *28*, 213–221.

24. Trouvé, A. Diffeomorphisms groups and pattern matching in image analysis. *Int. J. Comput. Vis.* **1998**, *28*, 213–221.
25. Beg, M.F.; Miller, M.I.; Trouvé, A.; Younes, L. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. J. Comput. Vis.* **2005**, *61*, 139–157.
26. Richardson, C.L.; Younes, L. Metamorphosis of images in reproducing kernel Hilbert spaces. *Adv. Comput. Math.* **2016**, *42*, 573–603.
27. Hong, Y.; Joshi, S.; Sanchez, M.; Styner, M.; Niethammer, M. Metamorphic geodesic regression. *Med. Image Comput. Comput. Assist. Interv.* **2012**, *15*, 197–205.
28. Younes, L. *Shapes and Diffeomorphisms*; Springer: Berlin, Germany, 2010.
29. Miller, M.I.; Trouvé, A.; Younes, L. Hamiltonian systems and optimal control in computational anatomy: 100 years since D’Arcy Thompson. *Annu. Rev. Biomed. Eng.* **2015**, *17*, 447–509.
30. Miller, M.I.; Trouvé, A.; Younes, L. On the metrics and Euler-Lagrange equations of computational anatomy. *Annu. Rev. Biomed. Eng.* **2002**, *4*, 375–405.
31. Berkels, B.; Effland, A.; Rumpf, M. Time discrete geodesic paths in the space of images. *SIAM J. Imaging Sci.* **2015**, *8*, 1457–1488.
32. Christensen, G.E.; Johnson, H.J. Consistent image registration. *IEEE Trans. Med. Imaging* **2001**, *20*, 568–582.
33. Fischer, B.; Modersitzki, J. Curvature based image registration. *J. Math. Imaging Vis.* **2003**, *18*, 81–85.
34. Haber, E.; Modersitzki, J. A multilevel method for image registration. *SIAM J. Sci. Comput.* **2006**, *27*, 1594–1607.
35. Han, J.; Berkels, B.; Droske, M.; Hornegger, J.; Rumpf, M.; Schaller, C.; Scorzin, J.; Urbach, H. Mumford–Shah Model for one-to-one edge matching. *IEEE Trans. Image Process.* **2007**, *16*, 2720–2732.
36. Modersitzki, J. *Numerical Methods for Image Registration*; Oxford University Press: Oxford, UK, 2004.
37. Modersitzki, J. *FAIR: Flexible Algorithms for Image Registration*; SIAM: Philadelphia, PA, USA, 2009.
38. Deledalle, C.A.; Papadakis, N.; Salmon, J.; Vaiteer, S. CLEAR: Covariant LEAsT-square Refitting with applications to image restoration. *SIAM J. Imaging Sci.* **2017**, *10*, 243–284.
39. Miller, M.I.; Christensen, G.E.; Amit, Y.; Grenander, U. Mathematical textbook of deformable neuroanatomies. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11944–11948,
40. Yuan, J.; Schnörr, C.; Steidl, G. Simultaneous higher order optical flow estimation and decomposition. *SIAM J. Sci. Comput.* **2007**, *29*, 2283–2304.
41. Bertalmio, M. *Image Processing for Cinema*; CRC Press: Boca Raton, FL, USA, 2014.
42. Nikolova, M.; Steidl, G. Fast Hue and Range Preserving Histogram Specification: Theory and New Algorithms for Color Image Enhancement. *IEEE Trans. Image Process.* **2014**, *23*, 4087–4100.
43. Hertzmann, A.; Jacobs, C.E.; Oliver, N.; Curless, B.; Salesin, D.H. Image analogies. In Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles, CA, USA, 12–17 August 2001; pp. 327–340.
44. Peter, P.; Kaufhold, L.; Weickert, J. Turning diffusion-based image colorization into efficient color compression. *IEEE Trans. Image Process.* **2016**, *26*, 860–869.
45. Chambolle, A.; Pock, T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **2011**, *40*, 120–145.
46. Joshi, S.C.; Miller, M.I. Landmark matching via large deformation diffeomorphisms. *IEEE Trans. Image Process.* **2000**, *9*, 1357–1370.
47. Chen, D.; Hua, G.; Wen, F.; Sun, J. Supervised transformer network for efficient face detection. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2016; pp. 122–138.

