

Nonlinear Spectral Image Fusion

Martin Benning^{1,*}, Michael Möller^{2,*}, Raz Z. Nossek^{3,*}, Martin Burger⁴,
Daniel Cremers⁵, Guy Gilboa³, and Carola-Bibiane Schönlieb¹

¹ University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, UK
{mb941, cbs31}@cam.ac.uk

² Universität Siegen, Hölderlinstraße 3, 57076 Siegen, Germany
michael.moeller@uni-siegen.de

³ Technion IIT, Technion City, Haifa 32000, Israel
{nossekr@campus, guy.gilboa@ee}.technion.ac.il

⁴ Westfälische Wilhelms-Universität, Einsteinstrasse 62, 48149 Münster, Germany
martin.burger@wwu.de

⁵ Technische Universität München, Boltzmannstrasse 3, 85748 Garching, Germany
cremers@tum.de

Abstract. In this paper we demonstrate that the framework of nonlinear spectral decompositions based on total variation (TV) regularization is very well suited for image fusion as well as more general image manipulation tasks. The well-localized and edge-preserving spectral TV decomposition allows to select frequencies of a certain image to transfer particular features, such as wrinkles in a face, from one image to another. We illustrate the effectiveness of the proposed approach in several numerical experiments, including a comparison to the competing techniques of Poisson image editing, linear osmosis, wavelet fusion and Laplacian pyramid fusion. We conclude that the proposed spectral TV image decomposition framework is a valuable tool for semi- and fully-automatic image editing and fusion.

Keywords: Nonlinear spectral decomposition, total variation regularization, image fusion, image composition, multiscale methods

1 Introduction

Since the rise of digital photography people have been fascinated by the possibilities of manipulating digital images. In this paper we present an image manipulation and fusion framework based on the recently proposed technique of nonlinear spectral decompositions [13, 14, 5] using TV regularization. By defining spectral filters that extract features corresponding to particular frequencies, we can for instance transfer wrinkles from one face to another and create visually convincing fusion results as shown in Figure 1.

*These authors contributed equally to this work.



Fig. 1: Example of a facial image fusion result obtained by the proposed framework. The left image of Reagan and right image of Obama were used as an input, while the two middle images are synthesized using nonlinear spectral TV filters.

Classical multiscale methods such as Fourier analysis, sine or cosine transformations, or wavelet decompositions represent an input image as a linear superposition of a given set of basis elements. In many cases, these basis elements are given by the eigenfunctions of a suitable linear operator. For instance, the classical Fourier representation of a function as a superposition of sine and cosine functions arises from the eigenfunctions of the Laplace operator, i.e. from functions v_λ with $\|v_\lambda\| = 1$ and $\Delta v_\lambda = \lambda v_\lambda$, with periodic boundary conditions. Interestingly, the condition for v_λ being an eigenfunction can be written in terms of the regularization functional $J(v) = \frac{1}{2}\|\nabla v\|_2^2$ as

$$\lambda v_\lambda \in \partial J(v_\lambda), \quad (1)$$

where $\partial J(v) = \{p \in \mathcal{X}^* \mid J(u) - J(v) - \langle p, u - v \rangle \geq 0\}$ denotes the subdifferential of the functional $J : \mathcal{X} \rightarrow \mathbb{R}$, with \mathcal{X} being a suitable function space, typically a Banach space. Since inclusion (1) makes sense for arbitrary convex regularization functions (e.g. for TV regularization), it provides a natural definition for generalizing the concept of eigenfunctions, cf. [3].

The idea of nonlinear spectral decompositions [13, 14, 5] (which we will recall in more detail in Section 3.1) is built upon the idea that an eigenfunction in the spatial domain, i.e. an element meeting (1), should be represented as a single peak in the spectral domain. Decompositions with respect to TV regularization have been shown to provide a highly image-adaptive way to represent different features and scales, see [13].

In this paper we will demonstrate that the nonlinear spectral image decomposition framework is very well suited for several challenging image fusion tasks. Our contributions include

- Proposing a nonlinear spectral image editing and fusion framework.
- Providing a robust pipeline for the automatic fusion of human faces, including face and landmark detection, registration, and segmentation.
- Illustrating state-of-the-art results evaluated against Laplacian pyramid fusion, wavelet fusion, Poisson image editing, and linear osmosis.
- Demonstrating the flexibility of the proposed framework beyond the fusion of facial images by considering applications such as object insertion and image style manipulation.

Copyright remark: All photographs used in this paper were taken from the Wikipedia Commons page, <https://commons.wikimedia.org/>, or from the free images site <https://commons.pixabay.com/>. The photo of Barack Obama was made by Pete Souza - Creative Commons Attribution 3.0 Unported license, see <https://creativecommons.org/licenses/by/3.0/deed.en>

2 Image Fusion

The most common image fusion techniques use a multiscale approach such as wavelet decompositions [30] or a Laplacian pyramid [9] to decompose two or more images, combine the decompositions differently on different scales, and reconstruct an image from the fused multiscale decomposition. Applications of the aforementioned fusion techniques include generating an all-in-focus image from a stack of differently focused images (e.g. [20]), multi- and hyperspectral imagery (cf. [1]), or facial texture transfers [28].

It was shown in [5, 15] that the nonlinear spectral decomposition framework actually reduces to the usual wavelet decomposition when the TV regularization is replaced by $J(u) = \|Wu\|_1$, where W denotes the linear operator conducting the (orthogonal) wavelet transform. We, however, are going to demonstrate that the image-adaptive nonlinear decomposition approach with TV regularization is significantly better suited for image manipulation and fusion tasks.

Several other sophisticated nonlinear image multiscale decompositions have been proposed including techniques based on bilateral filtering (e.g. [12]), weighted least-squares [11], local histograms [19], local extrema [27], or Gaussian structure–texture decomposition [26]. Applications of the aforementioned methods include image equalization and abstraction, detail enhance or removal, and tone mapping/manipulation. While [26] briefly discusses applications in texture transfer, the potential of a complete image fusion by combining different frequencies of different images has not been exploited sufficiently yet.

For various image editing tasks related to inserting objects from one image g into another image f , the seminal work of Perez, Gangnet and Blake on Poisson image editing [23], provides a valuable tool. The authors proposed to minimize $E(u) = \|\nabla u - \nabla g\|^2$ subject to u coinciding with f outside of the region the object is to be inserted into.

Recent improvements of the latter have been made with osmosis image fusion, cf. [18, 17]. Linear osmosis filtering for image fusion is achieved by solving a drift-diffusion PDE; here the drift vector field is constructed by combining the two vector fields $\nabla \ln(g)$ and $\nabla \ln(f)$; parts of $\nabla \ln(g)$ are inserted into $\nabla \ln(f)$, and averaged across the boundary. The initial value of the PDE is set to f , or the mean of f . A detailed description of the procedure is given in [18, Section 4.3]. For a general overview of image fusion techniques in different areas of application we also refer the reader to [25].

3 Nonlinear Spectral Fusion

The starting point and motivation for extending linear multiscale methods such as Fourier or wavelet decompositions into a nonlinear setting are basis elements, which often originate as eigenfunctions of a particular linear operator. As shown in Section 1, Fourier analysis can be recovered by decomposing a signal into a superposition of elements v_λ meeting the inclusion (1).

As mentioned in the introduction, the disadvantage of conventional decomposition techniques is the lack of adaptivity of the basis functions. In the following, we recall the definition of more general, nonlinear spectral transformations that allow to create more adaptive decompositions of images.

3.1 Nonlinear Spectral Decomposition

The idea of nonlinear spectral decompositions of [13, 14, 5] is to consider (1) for one-homogeneous functionals J (such as TV) instead of quadratic ones, which give rise to classical multiscale image representations. Since eigenvectors of one-homogeneous functionals are difficult to compute numerically (cf. [3]), the property one aims to preserve is that input data given in terms of an eigenfunction is decomposed into a single peak when being transformed into its corresponding (nonlinear) frequency representation.

Let us consider an eigenfunction $f = v_\lambda$, $\|v_\lambda\|_2 = 1$, obeying (1), and consider the behavior of the gradient flow

$$\partial_t u_{GF}(t) = -p_{GF}(t), \quad p_{GF}(t) \in \partial J(u_{GF}(t)), \quad u_{GF}(0) = f, \quad (2)$$

for a one-homogeneous functional J . It follows almost directly from [3, Theorem 5] that the solution to this problem is given by

$$u_{GF}(t) = \begin{cases} (1 - t\lambda)f & \text{if } t\lambda \leq 1, \\ 0 & \text{else.} \end{cases} \quad (3)$$

Since $u_{GF}(t)$ behaves piecewise linear in t , one can consider the second derivative to obtain a δ -peak. One defines

$$\phi_{GF}(t) = t\partial_{tt}u_{GF}(t) \quad (4)$$

to be the spectral decomposition of the input data f , even in the case where f is not an eigenfunction of J . The additional normalization factor t admits to the reconstruction formula

$$f = \int_0^\infty \phi_{GF}(t) dt + \bar{f}, \quad (5)$$

with $\bar{f} := \min_{\tilde{f} \in \text{kernel}(J)} \|\tilde{f} - f\|_2$, for arbitrary f . We refer the reader to [14] for more details on the general idea, and to [7] for a mathematical analysis of the above approach.

As we can see in (3), peaks of eigenfunctions in ϕ_{GF} appear at $t = \frac{1}{\lambda}$, i.e. earlier the bigger λ is. Therefore, one can interpret ϕ_{GF} as a wavelength decomposition, and motivate wavelength based filtering approaches of the form

$$\hat{u} = \int_0^\infty H(t) \phi_{GF}(t) dt + \overline{H} \overline{f}, \quad (6)$$

where the filter function H (along with the weight \overline{H}) can enhance or suppress selected parts of the spectrum.

As discussed in [5], there exists an alternative formulation to the gradient flow representation defined in (2). One can also consider the inverse scale space flow (see [8, 6])

$$\partial_t p_{IS}(t) = f - u_{IS}(t), \quad p_{IS}(t) \in \partial J(u_{IS}(t)), \quad p_{IS}(0) = 0. \quad (7)$$

For certain regularizations J , the two approaches are provably equivalent (cf. [7]); hence, we use the approaches interchangeably based on the numerical convenience, as we also empirically observe very little difference between the numerical realisations of (2) and (7).

Note that we use the total variation as the regularizer J throughout the remainder of this paper; however, other choices for J are possible (see [5]).

3.2 Numerical Implementation

Spectral Decomposition For the numerical implementation of our spectral image fusion we use both the gradient flow as well as the inverse scale scale flow formulation. The former is implemented in the exact same way as described in [14]. Formulation (7) is discretized via Bregman iterations (cf. [22]). More precisely, we compute

$$u^{k+1} = \arg \min_u \frac{\tau^{k+1}}{2} \|u - f\|_2^2 + (TV(u) - \langle p^k, u \rangle), \quad (8)$$

$$p^{k+1} = p^k + \tau^{k+1}(f - u^{k+1}), \quad (9)$$

starting with $p^0 = 0$. We then define

$$\psi^k = \begin{cases} u^1 & \text{if } k = 1, \\ u^k - u^{k-1} & \text{else,} \end{cases} \quad (10)$$

to be the frequency decomposition of the input data f .

From the optimality condition of equation (8) we conclude that $p^k \in \partial TV(u^k)$ for all k . Furthermore, note that equation (9) can be rewritten as

$$\frac{p^{k+1} - p^k}{\tau^{k+1}} = f - u^{k+1} \quad (11)$$

and can therefore be interpreted as the discretization of the inverse scale space flow. In our numerical implementation we use the adaptive step size $\frac{1}{\tau^k} = 30 \cdot$



Fig. 2: Example of a nonlinear frequency decomposition. The left image is the input image to be decomposed, the following images illustrate selected spectral components with increasing associated frequencies. This type of decomposition is the main tool for our proposed image fusion framework.

0.6^{k-1} to better resolve significant changes of the flow. With this adaptation, we found 15 iterations to be sufficient to approximately converge to $u^{15} = f$ and to still obtain a sufficiently detailed frequency decomposition. Figure 2 illustrates a generalized frequency representation using the above method on an input image of a bee.

To solve the minimization problem of equation (8) numerically we use the primal-dual hybrid gradient method with diagonal preconditioning [24] and the adaptive step size rule from [16].

Image Fusion The general idea of the spectral image fusion is to apply the nonlinear spectral image decomposition to two images or regions therein, combine the coefficients at different scales, and reconstruct an image from the fused coefficients.

Let v be a registration function that aligns a part of the second image with the location in the first image where the object is to be inserted into. Given the corresponding spectral decompositions ϕ^1 and ϕ^2 , we compute the fused image u_{fused} via

$$u_{\text{fused}}(x) = \int_0^\infty H^1(x, t)\phi^1(x, t) + H^2(x + v(x), t)\phi^2(x + v(x), t) dt, \quad (12)$$

where the two filter functions H^1 and H^2 determine the amount of spectral information to be included in the fused image. Finally, we add a weighted linear combination of the constant parts \bar{f}^1 and \bar{f}^2 of the two input images f^1 and f^2 to u_{fused} . Note that – opposed to the original spectral representation framework from [13, 14, 5] – we are considering x -dependent, i.e. *spatially varying filters*, to adapt the filters in different regions of the images.

4 Results

4.1 Automatic Image Fusion of Human Faces

To illustrate the concept of using nonlinear spectral decompositions for image editing, we consider the problem of fusing two images of human faces. The latter

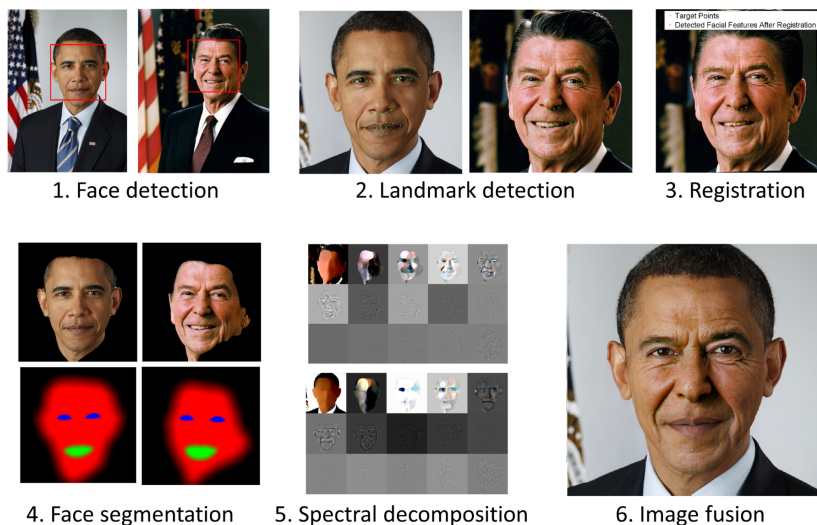


Fig. 3: Illustration of the pipeline for facial image fusion using nonlinear spectral decompositions.

has attracted quite some attention in the literature before, see e.g. [4, 28]. Note that in contrast to [4, 28] our fusion process does not depend on a 3d model of a face (which naturally means our framework does not handle changes of perspective).

For the presented image fusion, we have developed a fully automatic image fusion pipeline illustrated in Figure 3. It consists of face detection using the Viola-Jones algorithm [29], facial landmark detection using [2], determining the non-rigid registration field v that has a minimal Dirichlet energy $\|\nabla v\|_2$ among all

possible maps that register the detected landmarks, a face segmentation using the approach in [21] with additional information from the landmarks to distinguish between the face, mouth, and eye region, and finally the decomposition and fusion steps described in Section 3, where we restrict the decomposition to the regions of interest to be fused. Upon acceptance of this paper we will make the source code available in order to provide more details of the implementation.

The segmentation into the subregions allows us to define spatially varying spectral filters that treat the eye, mouth, and remaining facial regions differently, where fuzzy segmentation masks are used to blend the spectral filters from one region into the next to create smooth and visually pleasing transitions. Effects one can achieve by varying the spectral filters in the eye and mouth regions are illustrated in Figure 4.

Figure 5 shows the filters we used to fuse the faces of the presidents Obama and Reagan for the introductory example in Figure 1. As illustrated, the spectral filters may also differ for each of the color channels and can therefore also be



Fig. 4: The sub-segmentation of each image into a face, a mouth and an eye region allows to define spatially varying filters. The above images illustrate effects of incorporating eyes or the mouth from one or the other image.

applied to images decomposed into luminance and chrominance channels. In our examples we used the $(L, C^{G/M}, C^{R/B})$ color transform which has shown a promising performance e.g. for image demosaicking in [10]. As we can see in Figure 5, one might want to keep more chrominance values of the target image to retain similar color impressions. Furthermore, the filter responses do not have to sum to one. In the high frequencies we keep a good amount of both images, which – in our experience – leads to sharper and more appealing results with skin-textures from both images.

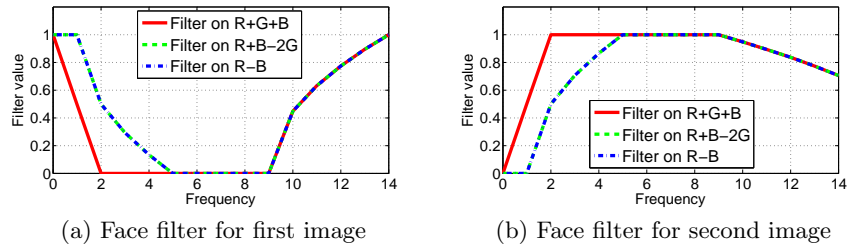


Fig. 5: Illustration of the fusion filters to generate the images in Figure 1.

To illustrate the robustness of the proposed framework, we ran the fully automatic image fusion pipeline on an image set of US presidents gathered from the Wikipedia Commons page. The results are shown in the supplementary material accompanying this manuscript. The proposed nonlinear image fusion approach is robust enough to work with a great variety of different images and types of

photos. The supplementary material contains further examples of fusing people with statues, and fusing a bill with a painting.

Finally, we want to highlight that the nonlinear image fusion framework has applications beyond facial image manipulation. Similar to Poisson image editing [23], one can insert objects from one image into the other by keeping low frequencies (colors and shadows) from one image and using higher frequencies (shapes and texture) from another image. Figure 6 shows an example of fusing the images of a shark and a swimmer.

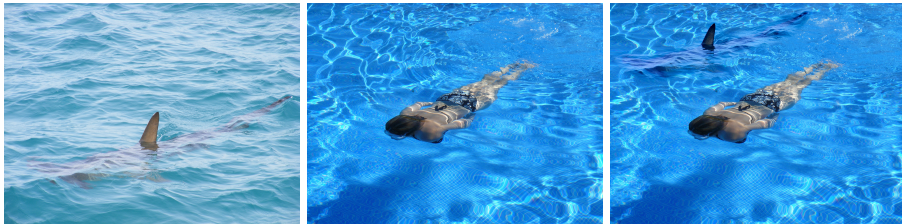


Fig. 6: Inserting the shark from the left image into the middle image via spectral image fusion yields the image on the right: by keeping low frequencies from the middle image, one obtains highly believable colors in the fusion result. A smooth transition between the inserted object and the background image by a fuzzy segmentation mask (alpha-matting) was used to further avoid fusion artifacts.

4.2 Comparison to Other Techniques

To illustrate the advantages of the image-adaptive nonlinear spectral decomposition we compare our algorithm to the classical multiscale methods of wavelet fusion, Laplacian pyramid fusion, and to the fotomontage techniques of Poisson image editing [23] as well as linear osmosis image editing [18, 17]. We compare all methods on the challenging example of fusing a photo of Reagan with the painting of Mona Lisa, see Figure 7. All methods use the identical registration- and segmentation-results from the automatic fusion pipeline described in Section 4.1. As we can see, Poisson and osmosis imaging transfer too many colors of the reference images and require more sophisticated methods for generating a guidance gradient field to also incorporate fine scale details of the target image such as the scratches on the painting. Wavelet image fusion generates unnatural colors and the Laplacian pyramid approach contains some halos. In particular, the texture of Reagans cheeks makes the Laplacian pyramid fusion look unnatural. By damping the filter coefficients of the nonlinear spectral decomposition, one can easily generate a result which is subtle enough to look realistic but still have clearly visible differences.

4.3 Artistic Image Transformations

Another application that demonstrates the variety of possibilities using nonlinear spectral decompositions for image manipulation is transforming an image such

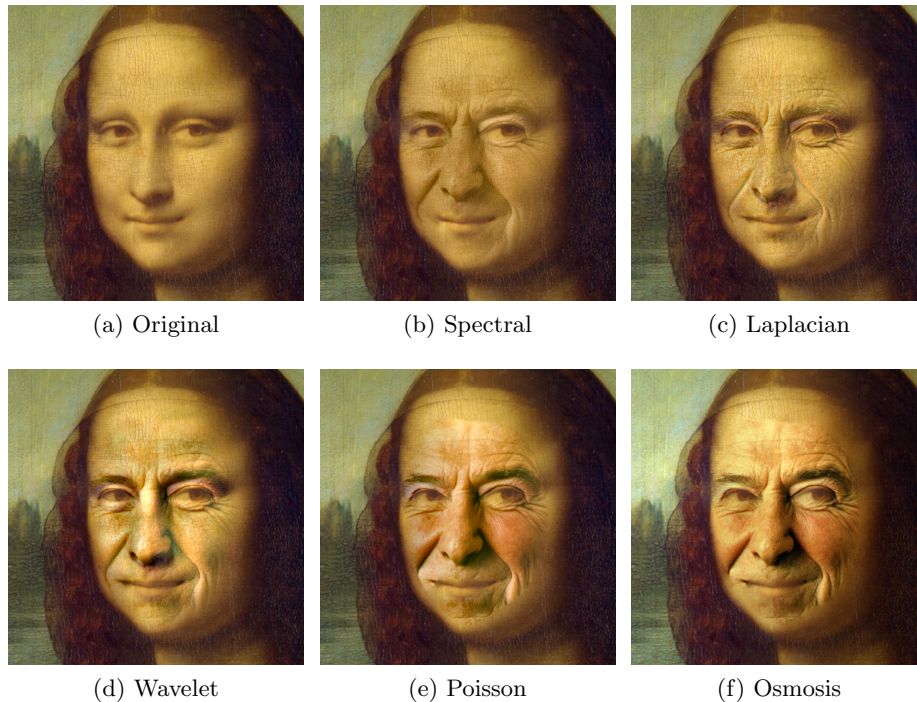


Fig. 7: Comparing the spectral image fusion (b) for the example of injecting details from the Reagan image into the image of Mona Lisa (a) with Laplacian pyramid fusion (c), wavelet image fusion (d), Poisson editing (e), Osmosis (e).

that the transformed image has a new look and feel. This means the image still keeps the same salient objects or features of the original image after the manipulation process, but they now seem as if they were composed in a different way.

As a first example we consider transferring an image of a real world scene into a painting. To accomplish the latter, we extensively enhance medium frequency bands to acquire some characteristics associated with oil paintings: a small smearing effect and high contrast between different objects. To further increase the painting effect we borrow brush stroke qualities from an actual painting (Figure 8 left) and combine them with the original photo. The right image in Figure 8 illustrates the result of such a procedure.

Figure 9 demonstrates a different type of manipulation enabled by nonlinear spectral decomposition. In this case we keep only very low frequencies from a fish image, and import all other frequencies from a mosaic image, leading to the impression of a fish-mosaic in the fused image.



Fig. 8: Example of transforming a photo such that it gives the impression of being an impressionist painting. Using the spectral decomposition, we extract the brush stroke features of the painting (very left) from high frequency bands at a certain area (marked in red) and embedded them into the photo image. The result is shown on the right.

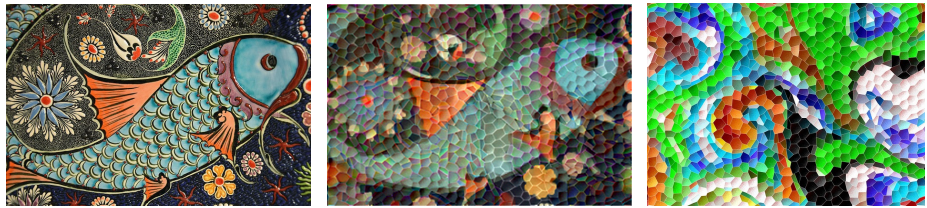


Fig. 9: Example of transforming ceramic art such that it gives the impression of being a mosaic. Low frequencies from the ceramic art (left) are extracted and combined with the high frequencies from the mosaic image (right).

5 Conclusions and Future Research

In this paper we demonstrated the potential of nonlinear spectral decompositions using TV regularization for image fusion. In particular, our facial image fusion pipeline produces highly realistic fusion results transferring facial details such as wrinkles from one image to another. It provides a high flexibility, leading to results superior to methods such as Poisson image editing, osmosis, wavelet fusion or Laplacian pyramids on challenging cases like the fusion of a photo and a painting. Furthermore, it easily extends to several other image manipulation tasks, including inserting objects from one image into another as well as transforming a photo into a painting.

Note that the proposed image fusion framework is not only complementary to other image fusion techniques, but can also be combined with those, e.g. by applying them on individual bands of the spectral decomposition, which is a direction of future research we would like to look into. Further directions of future research include learning a regularization that is possibly even better suited at separating facial expressions and wrinkles from the image than the total variation.

Data Statement: the corresponding programming codes are available at <https://doi.org/10.17863/CAM.8305>

Acknowledgements

MBe and CBS acknowledge support from EPSRC grant EP/M00483X/1 and the Leverhulme Trust project Breaking the non-convexity barrier. MBe further acknowledges support from the Leverhulme Trust early career fellowship "Learning from mistakes: a supervised feedback-loop for imaging applications" and the Newton Trust. MM acknowledges support from the German Research Foundation (DFG) as part of the research training group GRK 1564 Imaging New Modalities. RN and GG acknowledge support by the Israel Science Foundation (grant 718/15). MBu acknowledges support by ERC via Grant EU FP 7 - ERC Consolidator Grant 615216 LifeInverse. DC acknowledges support from ERC Consolidator Grant 3D Reloaded. CBS further acknowledges support from EP-SRC centre EP/N014588/1, the Cantab Capital Institute for the Mathematics of Information, and from CHiPS (Horizon 2020 RISE project grant).

References

1. K. Amolins, Y. Zhang, and P. Dare. Wavelet based image fusion techniques an introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62:249263, 2007.
2. A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic. Incremental face alignment in the wild. In *CVPR Proceedings*, pages 1859–1866, 2014.
3. M. Benning and M. Burger. Ground states and singular vectors of convex variational regularization methods. *Methods and Applications of Analysis*, 20(4):295–334, 2013.
4. V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194. ACM Press/Addison-Wesley Publishing Co., 1999.
5. M. Burger, L. Eckardt, G. Gilboa, and M. Moeller. Spectral representations of one-homogeneous functionals. In *Scale Space and Variational Methods in Computer Vision*, pages 16–27. Springer, 2015.
6. M. Burger, K. Frick, S.J. Osher, and O. Scherzer. Inverse total variation flow. *Multiscale Modeling & Simulation*, 6(2):366–395, 2007.
7. M. Burger, G. Gilboa, M. Moeller, L. Eckardt, and D. Cremers. Spectral decompositions using one-homogeneous functionals, 2015. Submitted. Online at <http://arxiv.org/pdf/1601.02912v1.pdf>.
8. M. Burger, G. Gilboa, S. Osher, and J. Xu. Nonlinear inverse scale space methods. *Comm. in Math. Sci.*, 4(1):179–212, 2006.
9. P.J. Burt and E.H. Adelson. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics (TOG)*, 2(4):217–236, 1983.
10. L. Condat and S. Mosaddegh. Joint demosaicking and denoising by total variation minimization. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 2781–2784, 2012.
11. Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. In *ACM SIGGRAPH 2008 Papers*, SIGGRAPH '08, pages 67:1–67:10, New York, NY, USA, 2008. ACM.
12. R. Fattal, M. Agrawala, and S. Rusinkiewicz. Multiscale shape and detail enhancement from multi-light image collections. In *ACM SIGGRAPH 2007 Papers*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.

13. G. Gilboa. A spectral approach to total variation. In *A. Kuijper et al. (Eds.): SSVM 2013*, volume 7893 of *Lecture Notes in Computer Science*, pages 36–47. Springer, 2013.
14. G. Gilboa. A total variation spectral framework for scale and texture analysis. *SIAM J. Imaging Sciences*, 7(4):1937–1961, 2014.
15. G. Gilboa, M. Moeller, and M. Burger. Nonlinear spectral analysis via one-homogeneous functionals: Overview and future prospects. *J. Math. Imaging Vis.*, 56(2):300–319, October 2016.
16. T. Goldstein, E. Esser, and R. Baraniuk. Adaptive Primal-Dual Hybrid Gradient Methods for Saddle-Point Problems. *arXiv Preprint*, 2013.
17. K. Hagenburg, M. Breu, J. Weickert, and O. Vogel. Novel schemes for hyperbolic pdes using osmosis filters from visual computing. In *Proceedings SSVM 2011*, volume 6667, pages 532–543. Springer, 2011.
18. M. Breu, J. Weickert, K. Hagenburg and O. Vogel. Linear osmosis models for visual computing. In *Proceedings EMMCVPR 2013*, volume 8081, pages 26–39. Springer, 2013.
19. M. Kass and J. Solomon. Smoothed local histogram filters. In *ACM SIGGRAPH 2010 Papers*, SIGGRAPH '10, pages 100:1–100:10, New York, NY, USA, 2010. ACM.
20. S. Li and B. Yang. Multifocus image fusion by combining curvelet and wavelet transform. *Pattern Recogn. Lett.*, 29(9):1295–1301, 2008.
21. C. Nieuwenhuis and D. Cremers. Spatially varying color distributions for interactive multi-label segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(5):1234–1247, 2013.
22. S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation based image restoration. *SIAM Journal on Multiscale Modeling and Simulation*, 4:460–489, 2005.
23. P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. In *ACM SIGGRAPH 2003 Papers*, SIGGRAPH '03, pages 313–318. ACM, 2003.
24. T. Pock and A. Chambolle. Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In *ICCV*, pages 1762–1769, 2011.
25. T. Stathaki. *Image Fusion: Algorithms and Applications*. Academic Press, 2008.
26. Z. Su, X. Lu, and A. Artusi. A novel image decomposition approach and its applications. *The Visual Computer*, 29:1011–1023, 2013.
27. K. Subr, C. Soler, and F. Durand. Edge-preserving multiscale image decomposition based on local extrema. In *ACM SIGGRAPH Asia 2009 Papers*, SIGGRAPH Asia '09, pages 147:1–147:9, New York, NY, USA, 2009. ACM.
28. J. Thies, M. Zollhöfer, M. Nießner, L. Valgaerts, M. Stamminger, and C. Theobalt. Real-time expression transfer for facial reenactment. *ACM Transactions on Graphics (TOG)*, 34(6), 2015.
29. P. Viola and M.J. Jones. Robust real-time face detection. *Int. J. Comput. Vision*, 57(2):137–154, 2004.
30. P.M. Zeeuw. Wavelets and image fusion. *CWI, Amsterdam, March*, 1998.