Restoration and Enhancement of Historical Stereo Photos through Optical Flow

Marco Fanfani$^{1[0000-0003-3741-1842]}$, Carlo Colombo$^{1[0000-0001-9234-537X]}$, and Fabio Bellavia$^{2[0000-0002-1688-8476]}$

^{1} Dept. of Information Engineering, Università degli Studi di Firenze, Italy
{marco.fanfani,carlo.colombo}@unifi.it
^{2} Dept. of Math and Computer Science, Università degli Studi di Palermo, Italy
fabio.bellavia@unipa.it

Abstract. Restoration of digital visual media acquired from repositories of historical photographic and cinematographic material is of key importance for the preservation, study and transmission of the legacy of past cultures to the coming generations. In this paper, a fully automatic approach to the digital restoration of historical stereo photographs is proposed. The approach exploits the content redundancy in stereo pairs for detecting and fixing scratches, dust, dirt spots and many other defects in the original images, as well as improving contrast and illumination. This is done by estimating the optical flow between the images, and using it to register one view onto the other both geometrically and photometrically. Restoration is then accomplished by data fusion according to the stacked median, followed by gradient adjustment and iterative visual consistency checking. The obtained output is fully consistent with the original content, thus improving over the methods based on image hallucination. Comparative results on three different datasets of historical stereograms show the effectiveness of the proposed approach, and its superiority over single-image denoising and super-resolution methods.

Keywords: Image denoising · Image restoration · Image enhancement · Stereo matching · Optical flow · Gradient filtering

1 Introduction

Photographic material of the XIX and XX centuries is an invaluable source of information for historians of art, architecture and sociology, as it allows them to track the changes occurred over the decades to a community and its living environment. Unfortunately, due to the effect of time and bad preservation conditions, most of the survived photographic heritage is partially damaged, and needs restoration, both at the physical (cardboard support, glass negatives, films, etc.) and digital (the image content acquired through scanners) levels. Dirt, scratches, discoloration and other signs of aging strongly reduce the visual quality of the photos [1]. A similar situation also holds for the cinematographic material [10].

Digital restoration of both still images and videos has attracted considerable interest from the research community in the early 2000s. This has led to the development of several tools that improve the visual quality. Some approaches rely on the instantiation of noise models, which can either be fixed a-priori or derived from the input images [3,16,18]. Other approaches detect damaged areas of the image and correct them according to inpainting techniques [6]. Self-correlation inside the image, or across different frames in videos, is often exploited in this context, under the assumption that zero-mean additive noise cancels out as the available number of image data samples increases [4,5,7]. A similar idea is exploited by super-resolution techniques, that enhance image quality by pixel interpolation [14,23]. In recent years, the algorithmic methods above have been sided by methods based on deep learning, that can infer the image formation model from a training set in order to inject this information into the final output—a process called image hallucination [19,21,24]. Although the final image may often alter the original image data content, and hence cannot be fully trusted (e.g., in the medical diagnosis domain), from the visual point of view the hallucination methods can give aesthetically pleasing results (see Fig. 1).

Stereoscopy has accompanied photography since its very birth in the nineteenth century, with ups and downs in popularity through time. Notwithstanding the lesser spread of stereo photography with respect to standard (monocular) photography, many digital archives with thousands of stereo images exist, some of which are freely available on the web. Stereo photos have a richer content than standard ones, as they present two different views of the same scene, thus explicitly introducing content redundancy and implicitly embedding information about scene depth. This characteristic can be exploited also in digital noise removal, enhancement and restoration, since a damaged area in one image can be reconstructed from the other image, provided that the correspondences between the two images are known. At a first glance, the above mentioned approach looks similar to that of video restoration from multiple video frames, in which the scene is acquired in subsequent time instants from slightly changed viewpoints. However, stereo images have their own peculiarities, and actually introduce in the restoration process more complications than video frames, which in movies typically exhibit an almost static and undeformed background, differently from stereo pairs. As a matter of fact, although several advances have been recently made in stereo matching and dense optical flow estimation [17], the problem is hard and far to be fully solved, especially in the case of very noisy and altered images such as those generated by early photographic stereo material. To the best of the authors' knowledge, stereo photo characteristics have been employed only for the super-resolution enhancement or deblurring of modern, clean photos [9,22,25]. On the other hand, the image analysis and computer vision approaches developed so far for historical stereo photos mainly aimed at achieving (usually in a manual way) better visualizations or 3D scene reconstructions [8,11,15], with no attempt at restoring the quality of the raw stereo pairs.

(a) I_1 (b) I_2 (c) (I_1, I_2)

(d) BM3D   (e) Remini   (f) proposed

(g) I_1 detail   (h) BM3D detail   (i) Remini detail   (j) proposed detail

Fig. 1. An example of historical stereo pair images, I_1 and I_2, also superimposed as anaglyph in (c). The enhancement of I_1 according to several methodologies is also shown. In particular, although producing visually impressive results, the deep super-resolution method Remini - Photo Enhancer[1] (e) does not preserve the true input image, adding fictitious details. Indeed, a closer look at (i) reveals alterations with respect to the original face expression (g), accentuating the smile and introducing bush-like textures on the left part of the hair. Best viewed in color, the reader is invited to zoom in the electronic version of the manuscript in order to better appreciate the visual differences.

This paper proposes a new approach to clean-up and restore the true scene content in degraded historical stereo photographs, working in a fully automatic way. With respect to existent single image methods, damaged image areas with scratches or dust can be better detected and fixed, thanks to the availability of more sampled data points for denoising. In addition, the correct illumination can be restored or enhanced in a way akin to that of High Dynamic Range

---

[1] https://play.google.com/store/apps/details?id=com.bigwinepot.nwdn.international

Imaging, where the images of the same scene taken at different exposure levels are used in order to enhance details and colors [12]. For this scope the optical flow, estimated with a recent state-of-the-art deep network [17], is used to synthesize corresponding scene viewpoints in the stereo pair, while denoising and restoration is carried out using non-deep image processing approaches. The entire process is superseded by scene content consistency validation, used to check critical stereo matching mispredictions that were previously unresolved by the network. Our approach aims to obtain an output which is fully consistent with the original scenario captured by the stereo pair, in contrast with the recent super-resolution and denoising approaches based on image hallucination.

The rest of the paper is organized as follows: Sec. 2 introduces the proposed approach. An experimental evaluation and comparison with similar approaches is reported in Sec. 3. Finally, conclusions and future work are discussed in Sec. 4.

2   Method Description

Given a pair of stereo images I_1 and I_2, the aim of the process is to output a defect-free version of one image of the pair (referred to as the reference) by exploiting the additional information coming from the other image (denoted as auxiliary). For convenience, the reference is denoted as I_1 and the auxiliary image as I_2, but their roles can be interchanged. Images are assumed to be single channel graylevel and ranging in $[0, 255]$. Code is freely available online[2].

2.1   Auxiliary Image Point-wise Transfer

As first step, the recent state-of-the-art Recurrent All-Pairs Field Transforms (RAFT) deep network [17] is used to compute the optical flow map pair (m_x, m_y), so that a synthesized image based on the content of I_2 and registered onto I_1 can be obtained
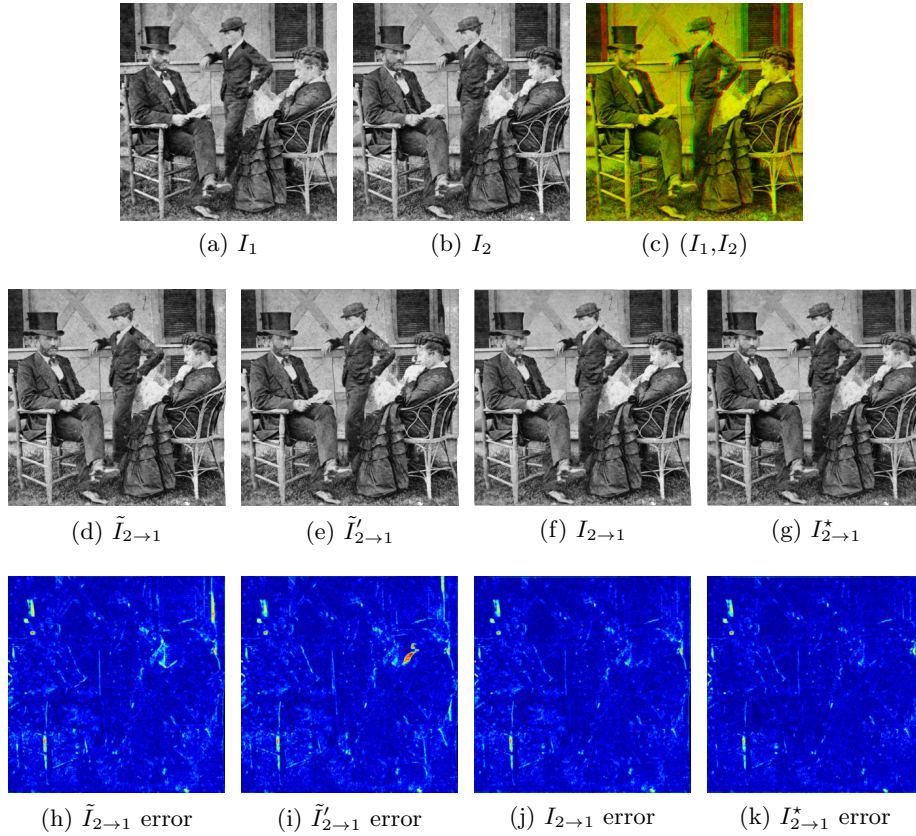
$$\tilde{I}_{2\to1}(x,y) = I_2(x + m_x(x,y), y + m_y(x,y)) \tag{1}$$

by transferring pixel intensity values from I_2 into the view given by I_1. Notice that in the case of perfectly rectified stereo images it holds everywhere that $m_y(x,y) = 0$. However the stereo alignment for the photos under consideration is far from perfect due to the technological limitations of the period, hence both the maps m_x and m_y are considered. RAFT optical flow estimation is not completely accurate, nor it does preserve the map inversion when exchanging the input image order. Hence, a further flow mapping pair (m'_x, m'_y) can be obtained by switching the two input images, which can be employed to synthesize a second image according to

$$\tilde{I}'_{2\to1}(x,y) = I_2(x - m'_x(x,y), y - m'_y(x,y)) \tag{2}$$

The final synthesized image $I_{2\to1}$ is then obtained by choosing the intensity value of each pixel (x, y) as the one from $\tilde{I}_{2\to1}(x, y)$ and $\tilde{I}'_{2\to1}(x, y)$ that minimizes the sum of absolute errors with respect to I_1 on a 5×5 local window centered on the pixel (compare the 2nd, 3th and 4th columns of Fig. 2).

(a) I_1 (b) I_2 (c) (I_1, I_2)

(d) $\tilde{I}_{2\to1}$ (e) $\tilde{I}'_{2\to1}$ (f) $I_{2\to1}$ (g) $I^{\star}_{2\to1}$

(h) $\tilde{I}_{2\to1}$ error   (i) $\tilde{I}'_{2\to1}$ error   (j) $I_{2\to1}$ error   (k) $I^{\star}_{2\to1}$ error

Fig. 2. Auxiliary image point-wise transfer and color correction steps, errors are reported with respect to I_1 while (c) shows the input stereo pair as anaglyph. Best viewed in color, the reader is invited to zoom in the electronic version of the manuscript in order to better appreciate the visual differences.

2.2   Color Correction

Due to the technical limitations of old the photographic instrumentation, illumination conditions between the two stereo images can differ noticeably. For instance, flash lamp and, even more, flash powder did not provide each time uniform and identical illumination conditions, and it was not infrequent that a

single camera was moved in two different positions in order to simulate a stereo setup instead of having two synchronized cameras [8]. Moreover, discoloration of the support due to aging can be present. In order to improve the final result, the state-of-the-art color correction method named Gradient Preserving Spline with Linear Color Propagation (GPS/LCP) presented in [2] is employed to correct the illumination of $I_{2\to1}$ according to I_1, thus obtaining the new image $I^\star_{2\to1}$ (see Fig. 2, last column). This method is able to preserve the image content and works also in the case of not perfectly aligned images. Clearly, if $I_{2\to1}$ presents better illumination conditions than I_1, it is also possible to correct I_1 according to $I_{2\to1}$.

2.3 Data Fusion

Given the reference image I_1 and the synthesized one obtained from the auxiliary view $I^\star_{2\to1}$ after the illumination post-processing, the two images are blended into a new image I_{12} according to the stacked median

$$I_{12} = \nabla(I_1 \cup I^\star_{2\to1}) \tag{3}$$

Formally, the stacked median $\nabla(\{I\})$ of a set of images $\{I\}$ is defined so that the intensity value of the resulting image located at pixel (x, y) is given by the median computed on the union of the 3×3 local windows centered at (x, y) on each image of the set. With this operator, dirt, scratches and other signs of photographic age or damages are effectively removed from I_{12}, but high frequency details can be lost in the process, due to the 3×3 filtering. These are re-introduced by considering a blended version of the gradient magnitude

$$d_{m_{12}} = \nabla(\nabla(M(I_1)) \cup \nabla(M(I^\star_{2\to1}))) \tag{4}$$

obtained as the two-step stacked median of eight possible gradient magnitudes, four for each of the I_1 and $I^\star_{2\to1}$ images. Each gradient magnitude image in the set $M(I)$ for a generic image I is computed as

$$d_m = \left(d_x^2 + d_y^2\right)^{\frac{1}{2}} \tag{5}$$

pixel-wise, where the image gradient direction pairs (d_x, d_y) are computed by the convolution of I with the following four pairs of kernel filters

$$\left\{ \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right), \left(\begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right), \left(\begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \right), \left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \right\} \tag{6}$$

Notice that $d_{m_{12}} \neq \nabla(M(I_{12}))$ in the general case. Consider for now only a single derivative pair (d_x, d_y) of I_{12}: Each pixel intensity $I_{12}(x, y)$ is incremented by a value $v(x, y)$ so that

$$\left(d_x + \frac{v}{2}\right)^2 + \left(d_y + \frac{v}{2}\right)^2 = \frac{d_m^2 - d_{m_{12}}^2}{2} \tag{7}$$

pixel-wise, that gives the a twofold solution

$$v^\star = \pm(2d_x d_y - d_{m_{12}}{}^2)^{\frac{1}{2}} - d_x - d_y \qquad (8)$$

The final solution $v(x, y)$ is chosen as 0 in the case $v^\star(x, y)$ is complex or, otherwise, as the one among the two possible solutions with the minimum absolute value, in order to alter I_{12} as little as possible, thus obtaining the final image $I_{12\uparrow} = I_{12} + v$ (see Fig. 3). Actually, the average of the four allowable derivative pairs is used to get the final $v(x, y)$ values.

(a) I_{12} (b) I_{12} detail (c) $\nabla(M(I_{12}))$

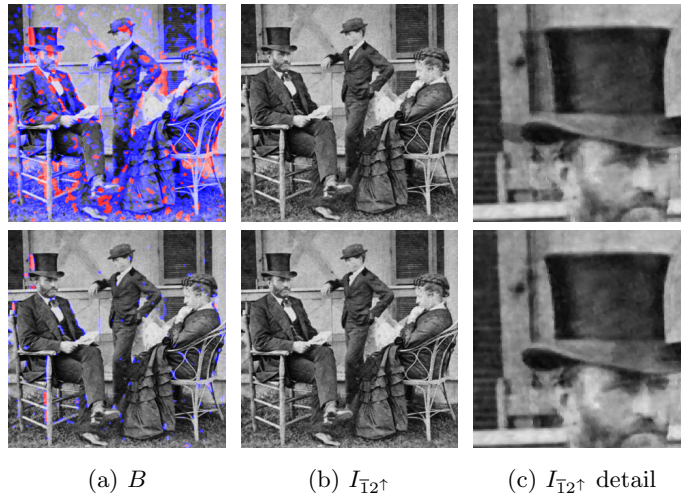(d) $I_{12\uparrow}$ (e) $I_{12\uparrow}$ detail (f) $d_{m_{12}}$
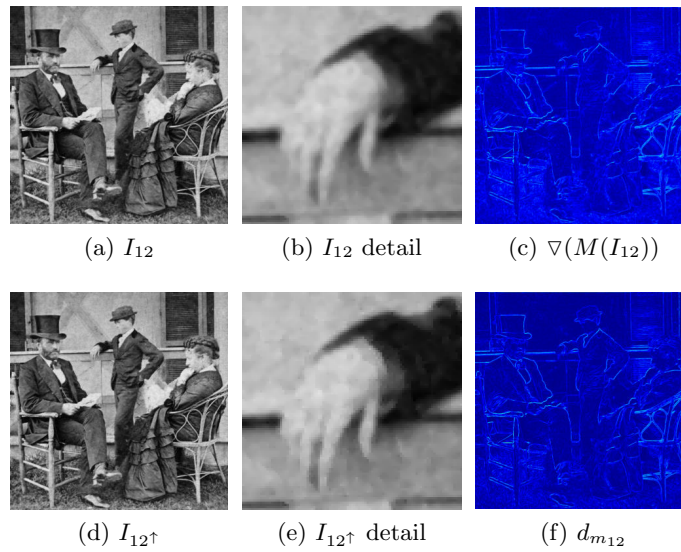
Fig. 3. Data fusion step. Best viewed in color, the reader is invited to zoom in the electronic version of the manuscript in order to better appreciate the visual differences.

2.4 Refinement

As already noted for the first step, the optical flow may be not perfect, causing the presence of wrong data in the image synthesis and hence in the data fusion process described in the previous step. To solve this issue, an error-driven image correction is introduced that consists of two sub-steps. The first one is the detection of the image region that needs to be adjusted, done by considering a 11×11 local window on $(I_1 - I_{12\uparrow})^2$ centered at (x, y). If on this window the root squared average of the values lower than the 66% percentile (i.e., the inliers) is higher than a predefined threshold $t = 16$, pixel (x, y) is selected to be corrected. Optionally, the obtained mask B can be smoothed using a Gaussian kernel followed by binarization. Selection by percentile has shown to be more robust than

working on the whole window. For the second sub-step, data fusion is repeated on the masked pixels by substituting $I_{2\to1}^{\star}$ with $I_{12\uparrow}$ in the input. Since $I_{12\uparrow}$ is a sort of average between I_1 and $I_{2\to1}^{\star}$, repeating data fusion between I_1 and $I_{12\uparrow}$ on masked pixels will push these more towards I_1. If this refined image $I_{\bar{1}2\uparrow}$ still contains regions to be adjusted according to the first sub-step, the second refinement sub-step is repeated again (see Fig. 4). A maximum of four iterations are allowed, as it was verified experimentally that data fusion roughly converges to I_1 within this number of steps. Actually, the current mask is morphologically dilated by a disk with radius proportional to the number of the remaining steps for better results.

(a) B (b) $I_{\bar{1}2\uparrow}$ (c) $I_{\bar{1}2\uparrow}$ detail

Fig. 4. Refinement step. First (top row) and last (bottom row) detection and adjustment sub-steps. The red and blue masks correspond respectively to make the selection with or without the proposed outlier rejection filtering, (c) is the final denoised image. It can be shown by inspecting the detail that the ghosting effect is gradually eliminated. Best viewed in color, the reader is invited to zoom in the electronic version of the manuscript in order to better appreciate the visual differences.

3 Evaluation

3.1 Dataset

In order to evaluate the proposed approach we built a new dataset including historical stereo pairs from different sources. The left frames of the selected stereo pairs are shown as reference in Fig. 5.
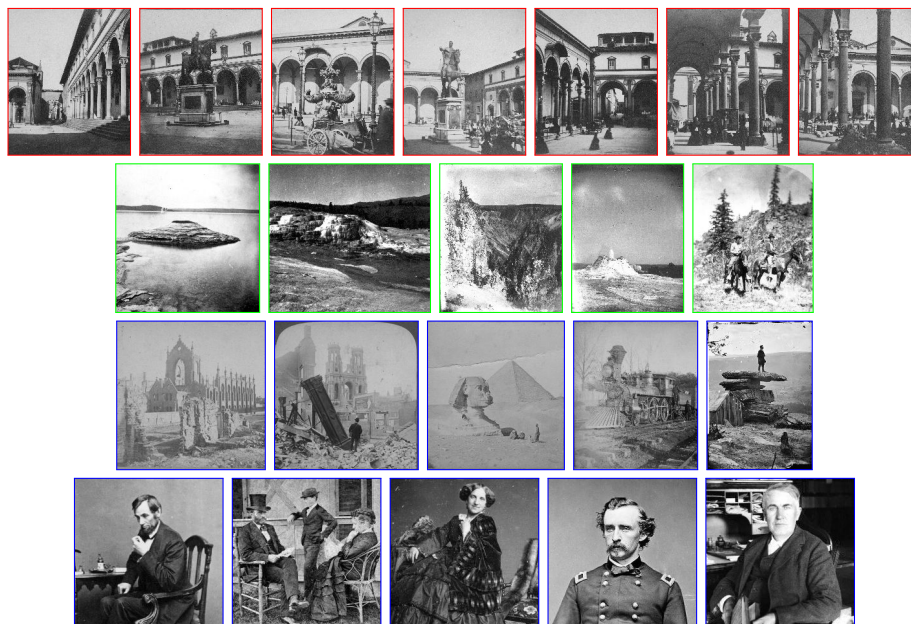
Fig. 5. Left frames of the noisy stereo pair datasets. Image frames for Hautmann's, USGS and Stereoscopic Photos datasets are respectively in red, green and blue. Best viewed in color and zoomed in onto the electronic version of the manuscript.

A first set of seven stereo pairs belongs to the collection of stereograms by Anton Hautmann, one of the most active photographers in Florence (Italy) between 1858 and 1862, courtesy of Photothek des Kunsthistorischen Instituts in Florenz—Max-Planck-Institut. Part of Hautmann's collection is described in [8]. The seven stereo pairs used in this work depict different viewpoints of Piazza Santissima Annunziata in Florence as it was in the middle of the XIX century. Inspecting these photos (see Fig. 5, red frames), it can be noticed that the image quality is very poor: the pairs are quite noisy, have poor definition, low contrast, saturated or blurred areas and also show scratches and stains.

A second set of five images was collected from the U.S. Geological Survey (USGS) Historical Stereoscopic Photos account on Flickr[3], and includes natural landscapes (see Fig. 5, green frames), except for the last one which includes also two horsemen with their mounts. Noise affects these images as with the first set, but also strong vignetting effects are present.

A third set, including ten stereo pairs, was instead gathered from the Stereoscopic History Instagram account[4] (see Fig. 5, blue frames). The first five images are landscape pictures of urban and natural scenes, while the last five pairs are

---

[3] https://www.flickr.com/photos/usgeologicalsurvey/
[4] https://www.instagram.com/stereoscopichistory/

individual or group portraits. This set is the most challenging one, since strong noises affect the images, evidently corrupting the content.

3.2   Compared Methods

The proposed approach was compared against Block Matching 3D (BM3D) [7] and the Deep Image Prior (hereafter denoted DP) architecture [19], two state-of-the-art single image denoising methods. For BM3D, the legacy version was employed, since according to our preliminary experiments the new version including correlated noise suppression did not work well for our kind of images. The BM3D σ parameter, the only one present, was set to 7 and 14, that according to our experiments gave the best visual results. In particular $\sigma = 14$ seems to work better than $\sigma = 7$ in the case of higher resolution images. Besides applying the standard BM3D on the reference image, a modified version of it was deployed in order to allow BM3D to benefit from the stereo auxiliary data. Since BM3D exploits image self-correlation to suppress noise, this modified BM3D generates auxiliary sub-images by siding two corresponding 96×96 patches from I_1 and $I_{2\rightarrow 1}^{\star}$, runs BM3D on each sub-image and then reconstructs the final result collecting those blocks from each sub-image corresponding to the 32×32 central I_1 patches. No difference in the result outputs with respect to the standard BM3D were detected, which plausibly implies that corresponding patches for I_1 and $I_{2\rightarrow 1}^{\star}$ are not judged as similar by BM3D. In the case of DP, the input can be rescaled due to network constraints, and borders were cropped: These missing parts were replaced with the original input image.

3.3   Results

Figures 6 and 7 show some results for a visual qualitative evaluation—the reader is invited to inspect the full-resolution images of the whole dataset included in the additional material[2]. No quantitative evaluations have been reported since, lacking of ground-truth clean data, image quality measurements requiring a reference image such as the Structural Similarity Index (SSIM) [20] cannot be extracted, while non-reference quality measures such as the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [13] are not modeled for old camera model statistics nor are able to judge the removal and restoration of damaged image areas.

According to the results, BM3D and DP seem to often oversmooth relevant details in the image, with BM3D producing somewhat better results than DP. In any case, these methods are not able to detect and compensate for dust, scratches and other kinds of artefacts, that conversely may be enhanced in the process, as one can check by inspecting the dust spots and the sketches on the two images. This is not the case of the proposed method that, exploiting the stereo correspondences, is able to fix these issues. Our method can also enhance the image contrast, as it happens for the dark spot under the right arcade in Fig. 7, 2nd row. In more severe cases, better results can be obtained by forcing the illumination of the auxiliary image into the reference (see Sec. 2.2), as shown

(a) I_1 (b) I_2 (c) BM3D   (d) DP   (e) proposed
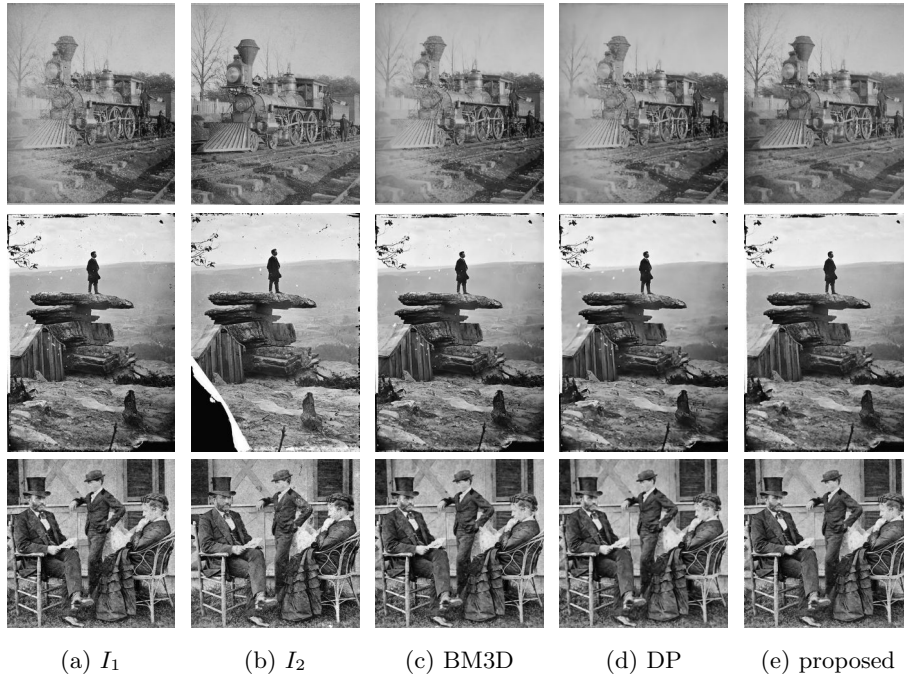
Fig. 6. Visual comparison on the whole image. Best viewed in color, the reader is invited to zoom in the electronic version of the manuscript in order to better appreciate the differences.

in Fig. 6, 1st row. Nevertheless, the proposed method also has some limitations. In particular, finer high-frequency texture details can sometimes disappear in low-resolution images, due to the use of 3×3 windows in the stacked median (see Sec. 2.3) in conjunction with an inaccurate optical flow estimation, as in Fig. 7, 3th row. Indeed, the proposed method strongly depends upon a correct identification of the optical flow, and although its wrong estimation is generally handled well (see Sec. 2.4), as for the tube hat of Fig. 7, last row, some spurious ghosting artefacts can sometimes arise, as those around the light pole in Fig. 7, 1st column.

4   Conclusion and Future Work

This paper proposed a novel method for the fully automatic restoration of historical stereo photographs. By exploiting optical flow clues, the auxiliary view of the stereo frame is geometrically and photometrically registered against the reference view. Restoration is carried out by fusing the data information available from both images according to the stacked median approach followed by gradient adjustments in order to better preserve finer details. Finally, an iter-

ative refinement is performed in order to remove artefacts due to errors in the flow estimation guided by a visual consistency check.

(a) I_1 (b) BM3D   (c) DP   (d) proposed

Fig. 7. Visual comparison on image details. Best viewed in color, the reader is invited to zoom in the electronic version of the manuscript in order to better appreciate the differences.

Results on historical stereo pairs show the effectiveness of the proposed approach, that is able to remove most of the image defects including dust and scratches, without excessive smoothing of the image content. The proposed method works better than its single-image denoising competitors, thanks to the ability of exploiting stereo information. As a matter of fact, single-image meth-

ods have severe limitations in handling damaged areas, and usually produce more blurry results.

Future work will be addressed to investigate novel solutions to refine the optical flow in order to reduce pixel mismatch, but also to consolidate both theoretical and practical aspects of the stacked median. In particular, further and general applications will be explored, as well as new methods for improving the gradient refinement. Additionally, explicit solutions to embed other methods such as BM3D into our approach and to identify damaged image areas on a single stereo frame will be studied in order to let the system better weights the local contributions of a frame in producing the final output. Finally, the proposed method will be extended and adapted to cinematographic images and historical film restoration.

Acknowledgment

References

1. Ardizzone, E., De Polo, A., Dindo, H., Mazzola, G., Nanni, C.: A dual taxonomy for defects in digitized historical photos. In: 10th International Conference on Document Analysis and Recognition. pp. 1166–1170 (2009)
2. Bellavia, F., Colombo, C.: Dissecting and reassembling color correction algorithms for image stitching. IEEE Transactions on Image Processing **27**(2), 735–748 (2018)
3. Besserer, B., Thiré, C.: Detection and tracking scheme for line scratch removal in an image sequence. In: European Conference on Computer Vision (ECCV2004). pp. 264–275 (2004)
4. Buades, A., Lisani, J., Miladinović, M.: Patch-based video denoising with optical flow estimation. IEEE Transactions on Image Processing **25**(6), 2573–2586 (2016)
5. Chen, F., Zhang, L., Yu, H.: External patch prior guided internal clustering for image denoising. In: IEEE International Conference on Computer Vision (ICCV2015). pp. 603–611 (2015)
6. Criminisi, A., Perez, P., Toyama, K.: Object removal by exemplar-based inpainting. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR2003). vol. 2 (2003)
7. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3D transform-domain collaborative filtering. IEEE Transactions on Image Processing **16**(8), 2080–2095 (2007)

8. Fanfani, M., Bellavia, F., Bassetti, G., Argenti, F., Colombo, C.: 3D map computation from historical stereo photographs of Florence. IOP Conference Series: Materials Science and Engineering **364**, 012044 (2018)

9. Jeon, D.S., Baek, S., Choi, I., Kim, M.H.: Enhancing the spatial resolution of stereo images using a parallax prior. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR2018). pp. 1721–1730 (2018)

10. Kokaram, A.C.: Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video. Springer-Verlag (1998)

11. Luo, X., Kong, Y., Lawrence, J., Martin-Brualla, R., Seitz, S.: KeystoneDepth: Visualizing history in 3D (2019)

12. McCann, J.J., Rizzi, A.: The Art and Science of HDR Imaging. Wiley & Sons (2011)

13. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Transactions on Image Processing **21**(12), 4695–4708 (2012)

14. Nasrollahi, K., Moeslund, T.B.: Super-resolution: a comprehensive survey. Machine Vision and Applications **25**(6), 1423–1468 (2014)

15. Schindler, G., Dellaert, F.: 4D cities: Analyzing, visualizing, and interacting with historical urban photo collections. Journal of Multimedia **7** (04 2012)

16. Stanco, F., Tenze, L., Ramponi, G.: Virtual restoration of vintage photographic prints affected by foxing and water blotches. Journal of Electronic Imaging **14**, 043008 (10 2005)

17. Teed, Z., Deng, J.: RAFT: Recurrent all-pairs field transforms for optical flow. In: European Conference on Computer Vision (ECCV2020 (2020)

18. Tegolo, D., Isgrò, F.: A genetic algorithm for scratch removal in static images. In: International Conference on Image Analysis and Processing (ICIAP2001). pp. 507–511 (2001)

19. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Deep image prior. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR2018) (2018)

20. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004)

21. Wang, Z., Chen, J., Hoi, S.C.H.: Deep learning for image super-resolution: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence (2020)

22. Yan, B., Ma, C., Bare, B., Tan, W., Hoi, S.: Disparity-aware domain adaptation in stereo image restoration. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR2020). pp. 13176–13184 (2020)

23. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. IEEE Transactions on Image Processing **19**(11), 2861–2873 (2010)

24. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image restoration. IEEE Transactions on Pattern Analysis and Machine Intelligence (2020)

25. Zhou, S., Zhang, J., Zuo, W., Xie, H., Pan, J., Ren, J.S.: DAVANet: Stereo deblurring with view aggregation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR2019). pp. 10988–10997 (2019)