IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 May 2024, accepted 10 June 2024, date of publication 21 June 2024, date of current version 1 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3417814

RESEARCH ARTICLE

A Proposed Workflow for the Restoration of Image Artifacts in Forensic Applications

FABRIZIO ARGENTI[1], (Senior Member, IEEE), STEFANIA BELLAVIA[2], MARCO FONTANI[3], (Member, IEEE), GABRIELE GUARNIERI[3], MARTINO JERIAN[3], (Member, IEEE), ALBERTO LIMONE[4], AND SIMONE REBEGOLDI[5]

[1]Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Firenze, 50139 Firenze, Italy
[2]Dipartimento di Ingegneria Industriale, Università degli Studi di Firenze, 50139 Firenze, Italy
[3]Amped Software, 34149 Trieste, Italy
[4]Dipartimento di Matematica e Informatica "Ulisse Dini," Università degli Studi di Firenze, 50139 Firenze, Italy
[5]Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università degli Studi di Modena e Reggio Emilia, 41125 Modena, Italy

Corresponding author: Fabrizio Argenti (fabrizio.argenti@unifi.it)

The work of Simone Rebegoldi was supported in part by the Istituto Nazionale di Alta Matematica - Gruppo Nazionale per il Calcolo Scientifico (INDAM-GNCS) through Progetti di Ricerca 2024; and in part by PNRR—Missione 4 Istruzione e Ricerca—Componente C2 Investimento 1.1, Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN) funded by European Commission under the NextGeneration European Union (EU) Programme, Project "Inverse Problems in Partial Differential Equations (PDE): Theoretical and Numerical Analysis," code: 2022B32J5C Ministero dell' Università e della Ricerca (MUR) D.D. financing decree n. 973 of June 2023 under Grant CUP B53D23009200006.

ABSTRACT The last years have witnessed significant developments in image acquisition systems and in algorithms for extracting information from them. Nevertheless, in many scenarios, several factors can hinder the recovery of useful data from images. This is especially true and important in forensic applications, where images are often accidentally captured by an imaging system not engineered for that specific acquisition (for example, the surveillance system designed to monitor the entrance of a bank may accidentally capture the license plate of a vehicle passing outside the bank). Therefore, the acquired images often need to be processed to facilitate extracting information from them. When facing a combination of several impairment factors, such as blur and perspective distortion, several image restoration algorithms must be applied. Then, it is necessary to choose a restoration order, which means the order by which single restoration algorithms are chained together to obtain the enhanced image. This study aims to understand whether such an order may impact the final result. Of course, there exists a wide variety of image impairments; in this study, we focus on the case of an image affected by a combination of optical/motion blur, perspective distortion, and additive noise, which are all widespread artifacts in forensic image applications. To answer the question about the importance of choosing one restoration workflow over another, we first model each considered defect and its restoration operator and then analyze and compare the effects of the composition of such operators on the restored output. Such a comparison is made from both a mathematical and experimental point of view, using both images with synthetically generated impairments and pictures with real degradations. The results show that the restoration order can affect significantly the results, especially when the defects are severe.

INDEX TERMS Image enhancement, image forensics, image restoration, video surveillance.

I. INTRODUCTION

In the last century, technology has evolved dramatically, changing the world and our life in almost every aspect. Crime and the fight against crime are two fields where technology brought consistent innovation. Video surveillance is prominent among the various technologies used to promote justice and security. Ashby [1] conducted an extensive analysis covering over 250,000 crimes recorded by the British Transport Police between 2011 and 2015. He concluded that

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo.

Closed-Circuit Television (CCTV) recordings were available in 45% of the investigated cases, and it was considered helpful in 65% of the cases where available. Ashby also found that, when useful, CCTV recordings increased the rate of solved crimes by 25% on average. More recently, Jung and Wheeler [2] analyzed roughly 200,000 crime incidents in Dallas between 2014 and 2018. They compared crime clearance rates before and after the installation of CCTV and concluded that the crime clearance rate raised by a modest 2% after the cameras were installed, with most of the benefits going to investigations about thefts.

The two mentioned studies, probably because of the extensive amount of data, do not provide details about *why* in some circumstances, CCTV is available but not helpful. In this direction, Brookman and Jones [3] recently analyzed in depth 44 homicide cases in the United Kingdom. They found that, from a quantitative point of view, CCTV was the most widely used forensic science; however, from a qualitative point of view, they observed that CCTV is often handled in a sub-optimal way since the involved personnel lacks training or fails to adhere to standards and principles for recovering, storing, enhancing, interpreting, preparing, and presenting evidence. This result raises the importance of how CCTV evidence is processed: it may well happen that footage of a crime is available, but its quality is limited by the presence of multiple artifacts or distortions, as explained in a recent work by Seckiner et al. [4]. In such cases, forensic video enhancement can potentially turn what seemed a useless exhibit into a critical piece of evidence. However, as pointed out in [3], the lack of knowledge and tools is a significant impediment against forensic video enhancement, leading to useless or even dangerous results being produced and, often, to frustration in the operator. Moreover, when dealing with forensic science in general, it is not acceptable to support a conclusion solely on the basis of common sense, as this would expose the analyst to criticism for not using a scientific approach. Therefore, each decision taken during the forensic restoration of images and videos has to be based on solid scientific ground. This work aims to contribute to laying the scientific foundations for choosing the order of various image processing operators.

A. RELATED WORK

Considerable efforts have been put into establishing guidelines and best practices for forensic image and video enhancement, resulting in the publication of several documents by the Scientific Working Group on Digital Evidences (SWGDE),[1] or by the Digital Imaging Working Group of the European Network of Forensic Science Institutes (DIWG-ENFSI).[2] Even though these guidelines provide essential recommendations, they typically aim to remain sufficiently general and not replace the standard operating procedures that each unit should develop internally; therefore, some practical questions remain unanswered. Provided an operator has suitable training and tools, one crucial question that is currently left open is the following: when dealing with a video that presents several artifacts or defects - a detailed list of intrinsic and extrinsic artifacts and distortions is provided by Seckiner et al. [4] - which is the correct order in which they should be compensated for? To cite an example, in the widespread case of video suffering from optical distortion and motion blurring, which issue should be compensated first?

In 2015, Ledesma [5] analyzed this question from an empirical perspective, comparing the results of some competing restoration chains. The conclusions reached by Ledesma are substantially in line with the mathematical principle of inverting defects in reverse order compared to the one they were introduced. This would indeed reflect the well-known rule for inverting a composed function. However, such a rule cannot directly be applied to the case of video enhancement since no video tool employs the actual *inverse operator* of any defect; indeed, video restoration algorithms aim at restoring the original video by approximating the inverse operator, and thus they introduce processing artifacts in their output. Noticeably, Zhang et al. presented a detailed study about the commutability of blur and affine warp transformations [6]; however, the scenario they considered and the rules they derived hardly fit the case of CCTV video enhancement, where the more general perspective transformations have to be considered.

Finally, we acknowledge that several deep learning-based methods for image deblurring have recently been proposed [7], [8], [9], [10]. These methods show excellent results both on synthetically degraded and real images. The use of deep learning complicates the task of including the inversion process in a mathematical framework, which is the approach pursued in this work. Moreover, the suitability of deep learning-based techniques for forensic image enhancement is still disputed, mainly due to their lack of explainability, as recently acknowledged in the European Union Artificial Intelligence Act [11]. For these reasons, our study employs a classical regularization-based deblurring technique.

B. CONTRIBUTION OF THIS STUDY

In this work, we address the proposed question in both an analytical and experimental way. In particular, we focus on cases where the combination of optical/motion blur, noise, and perspective distortion is considered. For our case study, we aim at: (i) providing a mathematical formulation of the problem; (ii) proposing the image restoration workflow that should be implemented along with the theoretical mathematical motivations supporting this choice; (iii) validating numerically the proposed workflow on some image restoration test problems; (iv) showing that the proposed workflow performs best in restoring some real degraded images.

The work is structured as follows. Section II gives a brief introduction to the typical image generation model of a

[1] https://www.swgde.org
[2] https://enfsi.eu

CCTV system and provides a mathematical formulation for the case of deblurring and perspective correction, presenting the proposed general restoration workflow. Section III presents the experimental tests to assess the proposed framework and the obtained results. Finally, Section IV draws some conclusions and lays the basis for future work.

II. PROPOSED WORKFLOW FOR IMAGE RESTORATION

Given the complexity of the image generation model, surveillance videos are often affected by multiple defects and artifacts. In Figure 1, we provide a block diagram showing the usual image generation workflow on the top and a possible restoration workflow at the bottom.

When multiple defects are observed, it may be necessary to compensate for some of them, depending on the following:

- the purpose of the examination; for example, lens distortion could be tolerable when the purpose of the examination is reading a vehicle's license plate, while it must be removed when trying to measure a vehicle's speed;
- the intensity of the defect; for example, a slight optical blur is often tolerable, while a strong one makes interpretation of the imagery problematic;
- the availability of a reliable restoration algorithm; indeed, unreliable restoration procedures may introduce artifacts that make the processed image worse than the original.

For the above reasons, it is generally not recommended to compensate for every artifact introduced during the image generation process. However, it often happens that at least two, or more, artifacts must be addressed, as they hinder the interpretation of events [12]. When this is the case, what is the correct order to be followed when compensating for them?

If we had the ideal inverse operator for each step of the image generation process, the answer would follow trivially from the basic rules of mathematics. Indeed, if we define the ideal, distortion-free representation of the original scene as x, and we model each natural or digital processing step with a function, e.g., $\phi_1, \phi_2, \ldots, \phi_N$, then the final output of the generation model could be written as

$$z = \phi_N(\phi_{N-1}(\ldots(\phi_3(\phi_2(\phi_1(x)))))). \tag{1}$$

Obviously, in equation (1) the innermost function, ϕ_1, models the first defect encountered in the generation model (e.g., perspective), while the outermost function ϕ_N, represents the last defect (e.g., blocking artifacts). Suppose an examiner is provided with z and the ideal inverse operator of each function, namely $\phi_1^{-1}, \phi_2^{-1}, \ldots, \phi_N^{-1}$. In that case, x can be easily recovered by applying the rule for inverting a composed function. Taking into account that z has the form given in (1), and that for any j, $j = 1, \ldots, N$, $\phi_j^{-1}\phi_j$ is the identity operator, i.e. $\phi_j^{-1}(\phi_j(x)) = x$, we obtain

$$\phi_1^{-1}(\phi_2^{-1}(\phi_3^{-1}(\ldots((\phi_N^{-1}(z)))))) = x. \tag{2}$$

As equation (2) shows, the original defect-less signal is obtained by applying the inverse operators *in the reverse order* with respect to the one employed in the generation process.

Going back to the real world, the main issue is that, usually, we do not have the functions ϕ that model each step of the acquisition process and/or their inverse operators. Then, the approach taken in the restoration numerical procedures is to employ approximations of such operators and provide restoration formulas to invert them. To cite an example, optical blur is commonly modeled with convolution with a circular PSF, and the restoration operator for that is obtained with, e.g., Wiener-based deconvolution [13].

It is, therefore, of interest to ask whether the "invert the order" rule should be maintained even in the real world, where forensic video analysts can only access approximated inverse operators. We argue that this is the case, and we aim to justify it analytically and experimentally in a relevant case study.

A. CASE STUDY: DEBLURRING AND PERSPECTIVE CORRECTION

In the following, we will focus on a particular yet insightful case study, which is obtained by simplifying the whole image generation model illustrated in Figure 1. More precisely, we will assume that the original image is (i) distorted by a disadvantageous perspective due to the angle between the camera and the object; (ii) blurred due to either the optics of the acquisition system or the motion of the object; (iii) corrupted by noise, which the sensor can introduce in the conversion from analog to digital. In so doing, we neglect all other possible defects that might occur either in the scene, camera, transmission, or view phases. We remark that, even though our mathematical development focuses on the artifacts mentioned above, the general approach can be extended to other case studies similarly. In the case where several degradations must be restored, establishing the best workflow with our approach would require exploring all possible operators' permutations. Nevertheless, to the best of our knowledge, it is common practice to compensate for no more than a couple of the most "visually impacting" defects while minor degradations are kept. Furthermore, we do not require the existence of an exact inverse operator, since the approach can be applied by using an approximation of the inverse operator, such as the pseudo-inverse.

We assume that our simplified image generation model can be described the following linear system

$$y = HWx + e, \tag{3}$$

where $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ is the unknown image or ground truth, $H \in \mathbb{R}^{n\times n}$ is the $n \times n$ blurring matrix, whose structure is determined by imposing some specific assumptions on the image outside the field of view, i.e., outside the image boundary [14], $W \in \mathbb{R}^{n\times n}$ is an interpolation matrix modelling the perspective distortion, $y = (y_1, \ldots, y_n)^T$ denotes the digital noisy image, and

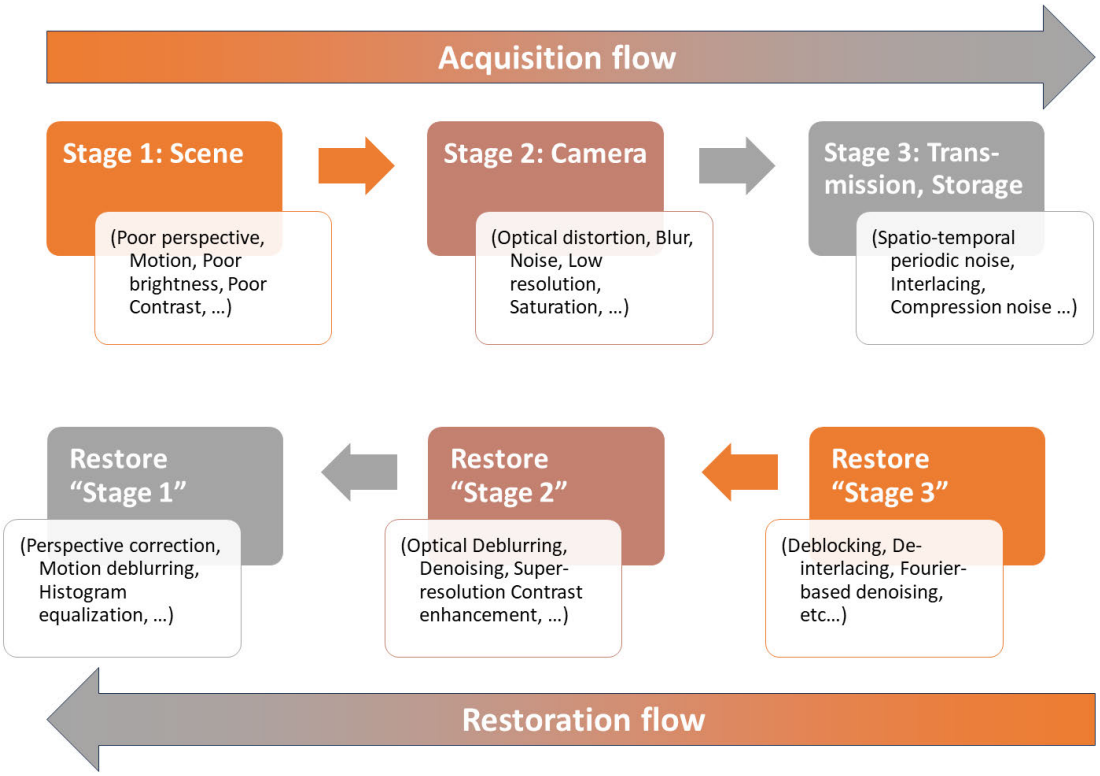


FIGURE 1. A block diagram representation of the typical image generation workflow (top) and a possible image restoration workflow (bottom), which corrects defects in the reverse order.

$e = (e_1, \ldots, e_n)^T \in \mathbb{R}^n$ represents the additive noise contribution.

Based on the simplified model (3), our goal is to recover the original image x from the distorted, blurred, and noisy image y. The difficulty of this task lies in the fact that the linear system (3) can be rarely solved by inverting the matrix HW, either because the inverse does not exist or has a high condition number that may lead to numerical instabilities. Hence, problem (3) is usually solved by employing regularization techniques.

B. PROPOSED WORKFLOW: DEBLUR-UNDISTORT

We now propose our image restoration workflow for solving (3). The proposed workflow, called *deblur-undistort*, is based on the decomposition of problem (3) into the following two sub-problems

$$y = Hz + e \tag{4}$$
$$z = Wx, \tag{5}$$

where z denotes the distorted unblurred image, so that first we correct the blurring defects due to the system optics or motion of the subject/camera, then we correct the distortion due to perspective. In the following, we explain how we perform these two tasks in detail.

Deblurring First, we solve sub problem (4) by following the Maximum a Posteriori (MAP) approach, which has become increasingly popular in the community of image processing in recent years [15], [16]. The MAP approach is based on a statistical formulation of the image restoration problem of interest, where the noise vector e and the image z are assumed to be realizations of two multivariate random variables E and Z, respectively. Regarding the noise, we assume, in our simplified model, that each component e_i of the noise vector e is a realization of a random variable E_i with Gaussian distribution of zero mean and standard deviation $\sigma > 0$. Then the vector e is a realization of the multivariate random variable E, whose probability density is [17, Chapter 7]

$$p_E(e) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|e\|_2^2\right),$$

where $\|\cdot\|_2$ denotes the usual Euclidean norm on $\mathbb{R}^n$. Because of (4), each pixel y_i of the acquired image is also a realization of a random variable Y_i. By setting $Y = (Y_1, \ldots, Y_n)^T$, the modeling of the system is then related to the probability density of the multivariate random variable Y. This density depends on the object z and therefore, we denote it as $p_Y(y; z)$. In conclusion, the statistical model for the detected image is

$$p_Y(y; z) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|y - Hz\|_2^2\right). \tag{6}$$

TABLE 1. Parameters of the PSFs used to distort images in the case of mild and strong perspective degradation, respectively.

PSF type	Mild perspective distortion	Strong perspective distortion
Gaussian	$\sigma = 10$	$\sigma = 5$
Out-of-focus	$r = 18$	$r = 12$
Motion	$\theta = 45°, l = 60$	$\theta = 45°, l = 40$

Regarding the image z, the probability density $p_Z(z)$ of the random variable Z is called *prior* and incorporates some a priori information on the image to be recovered, such as smoothness, sparsity, or presence of edges. For our workflow, we assume that Z has a Tikhonov prior [18], which has the form

$$p_Z(z) = \exp\left(-\frac{\alpha}{2\sigma^2}\|z\|_2^2\right),$$

where $\alpha > 0$ is the so-called *regularization parameter*. Introducing also the marginal probability $p_Y(y)$, we can compute, by means of the *Bayes theorem*, the conditional probability of Z with respect to the given value y of Y:

$$p_Z(z|y) = \frac{p_Y(y|z)p_Z(z)}{p_Y(y)}.$$

Then, the MAP estimate of z under Tikhonov regularization is the point $z_{\text{Tik}}(y)$ that maximizes the a posteriori probability $p_Z(z|y)$, namely

$$z_{\text{Tik}}(y) = \underset{z\in\mathbb{R}^n}{\text{argmax}}\, p_Z(z|y).$$

By taking the negative logarithm of $p_Z(z|y)$, we can reformulate the above maximization problem as the following minimization problem

$$\begin{aligned} z_{\text{Tik}}(y) &= \underset{z\in\mathbb{R}^n}{\text{argmin}} - \ln p_Z(z|y) \\ &= \underset{z\in\mathbb{R}^n}{\text{argmin}} - \ln(p_Y(y|z)) - \ln(p_Z(z)) + \ln(p_Y(y)) \\ &= \underset{z\in\mathbb{R}^n}{\text{argmin}} \|Hz - y\|_2^2 + \alpha\|z\|_2^2. \end{aligned}$$

The function $\varphi(z) = \|Hz - y\|_2^2 + \alpha\|z\|_2^2$ attains its minimum value at any point such that the gradient $\nabla\varphi(z)$ is null. Since the gradient can be written as $\nabla\varphi(z) = 2H^T(Hz - y) + 2\alpha z$, it is easy to see that $z_{\text{Tik}}(y)$ is the unique solution of the following linear system

$$(H^TH + \alpha I_n)z_{\text{Tik}}(y) = H^Ty, \tag{7}$$

where $I_n \in \mathbb{R}^{n\times n}$ denotes the $n \times n$ identity matrix. By inverting the symmetric positive definite matrix $H^TH + \alpha I_n$, we can finally write the deblurred image $z_{\text{Tik}}(y)$ as follows

$$z_{\text{Tik}}(y) = (H^TH + \alpha I_n)^{-1}H^Ty. \tag{8}$$

Perspective correction As a second step, we correct the distorted (yet deblurred) image $z_{\text{Tik}}(y)$. Assuming that the matrix W is invertible, we simply compute the unique solution of the linear system $Wx = z_{\text{Tik}}(y)$ by a linear system solver. Then, the vector $x_{\text{rec}}(y)$ takes the form:

$$x_{\text{rec}}(y) = W^{-1}z_{\text{Tik}}(y) = W^{-1}(H^TH + \alpha I_n)^{-1}H^Ty. \tag{9}$$

FIGURE 2. `License plate` **image (ground truth, size 1309 × 1109).**

The restored image $x_{\text{rec}}(y)$ represents the output of our proposed workflow.

In case W is not invertible we can compute $x_{\text{rec}}(y)$ as the least-squares minimum norm solution to $Wx = z_{\text{Tik}}(y)$. For the sake of simplicity, we assume W invertible, however the following analysis still holds by replacing W^{-1} with the pseudo inverse $W^{\dagger}$.

The good performance of the workflow depends on the choice of the regularization parameter α in the deblurring phase, which needs to be tuned through a trial-and-error procedure. In the following, we show with simple arguments coming from the theory of regularization [14] that it is possible to select α in such a way that the reconstruction error $\|x - x_{\text{rec}}(y)\|_2$ is sufficiently small.

Let $H = U\Sigma V^T$ be the singular value decomposition (SVD) of the matrix H, where $U, V \in \mathbb{R}^n$ are orthonormal matrices, i.e. $U^TU = V^TV = I_n$, and Σ is the following diagonal matrix:

$$\Sigma = \begin{pmatrix} \sigma_1 & \dots & \dots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \sigma_n \end{pmatrix},$$

where the positive scalars $\sigma_1, \dots, \sigma_n$ are the so-called singular values of the matrix. By assuming that the values $\{\sigma_i\}$ are in decreasing order, and letting $r \le n$ be the rank of H, i.e., the maximum number of linearly independent rows (or columns) of H, it holds $\sigma_1 \ge \sigma_2 \cdots \ge \sigma_r > 0$ and $\sigma_i = 0$, $i = r+1, \dots, n$.

We recall that the deblurred image $z_{\text{Tik}}(y)$ prior to the perspective correction is defined as

$$z_{\text{Tik}}(y) = (H^TH + \alpha I_n)^{-1}H^Ty. \tag{10}$$

TABLE 2. Best RRE provided by the three methods *deblur-undistort* (left), *undistort-deblur* (center) and *undistort-deconvblind* (right) for the three different blurs and distortion levels considered. Lower values indicate higher reconstruction quality.

RRE	*deblur-undistort*	*undistort-deblur*	*undistort-deconvblind*
		Mild perspective distortion	
Gaussian	**0.301**	0.359	0.551
Out-of-focus	**0.276**	0.403	0.791
Motion	**0.301**	0.433	1.056
		Strong perspective distortion	
Gaussian	**0.299**	0.331	0.719
Out-of-focus	**0.311**	0.363	0.864
Motion	**0.306**	0.424	1.129

TABLE 3. Best PSNR provided by the three methods *deblur-undistort* (left), *undistort-deblur* (center) and *undistort-deconvblind* (right) for the three different blurs and perspective distortion levels considered. Larger values indicate higher reconstruction quality.

PSNR	*deblur-undistort*	*undistort-deblur*	*undistort-deconvblind*
		Mild perspective distortion	
Gaussian	**18.91**	17.38	13.65
Out-of-focus	**19.67**	16.38	10.51
Motion	**18.91**	15.74	8.00
		Strong perspective distortion	
Gaussian	**18.85**	17.97	11.25
Out-of-focus	**18.53**	17.19	9.65
Motion	**18.66**	15.83	7.32

TABLE 4. Best SSIM provided by the three methods *deblur-undistort* (left), *undistort-deblur* (center) and *undistort-deconvblind* (right) for the three different blurs and distortion levels considered. Larger values indicate higher reconstruction quality.

SSIM	*deblur-undistort*	*undistort-deblur*	*undistort-deconvblind*
		Mild perspective distortion	
Gaussian	**0.578**	0.569	0.129
Out-of-focus	**0.591**	0.554	0.091
Motion	**0.557**	0.514	0.032
		Strong perspective distortion	
Gaussian	**0.613**	0.550	0.067
Out-of-focus	**0.573**	0.542	0.076
Motion	**0.552**	0.501	0.032

By employing the SVD decomposition $H = U\Sigma V^T$ in (10), we get

$$\begin{aligned} z_{\text{Tik}}(y) &= (V\Sigma \underbrace{U^T U}_{=I_n} \Sigma V^T + \alpha I_n)^{-1} V\Sigma U^T y \\ &= (V\Sigma^2 V^T + \alpha I_n)^{-1} V\Sigma U^T y \\ &= \left(V(\Sigma^2 + \alpha I_n)V^T\right)^{-1} V\Sigma U^T y \\ &= V(\Sigma^2 + \alpha I_n)^{-1} \underbrace{V^T V}_{=I_n} \Sigma U^T y. \end{aligned}$$

Then, $z_{\text{Tik}}(y)$ can be expressed in terms of the SVD decomposition of H as

$$z_{\text{Tik}}(y) = V(\Sigma^2 + \alpha I_n)^{-1}\Sigma U^T y. \tag{11}$$

In the following Proposition, we derive an upper bound for the error $\|x - x_{\text{rec}}(y)\|_2$, which measures the quality of the restored image $x_{\text{rec}}(y)$ with respect to the ground truth x.

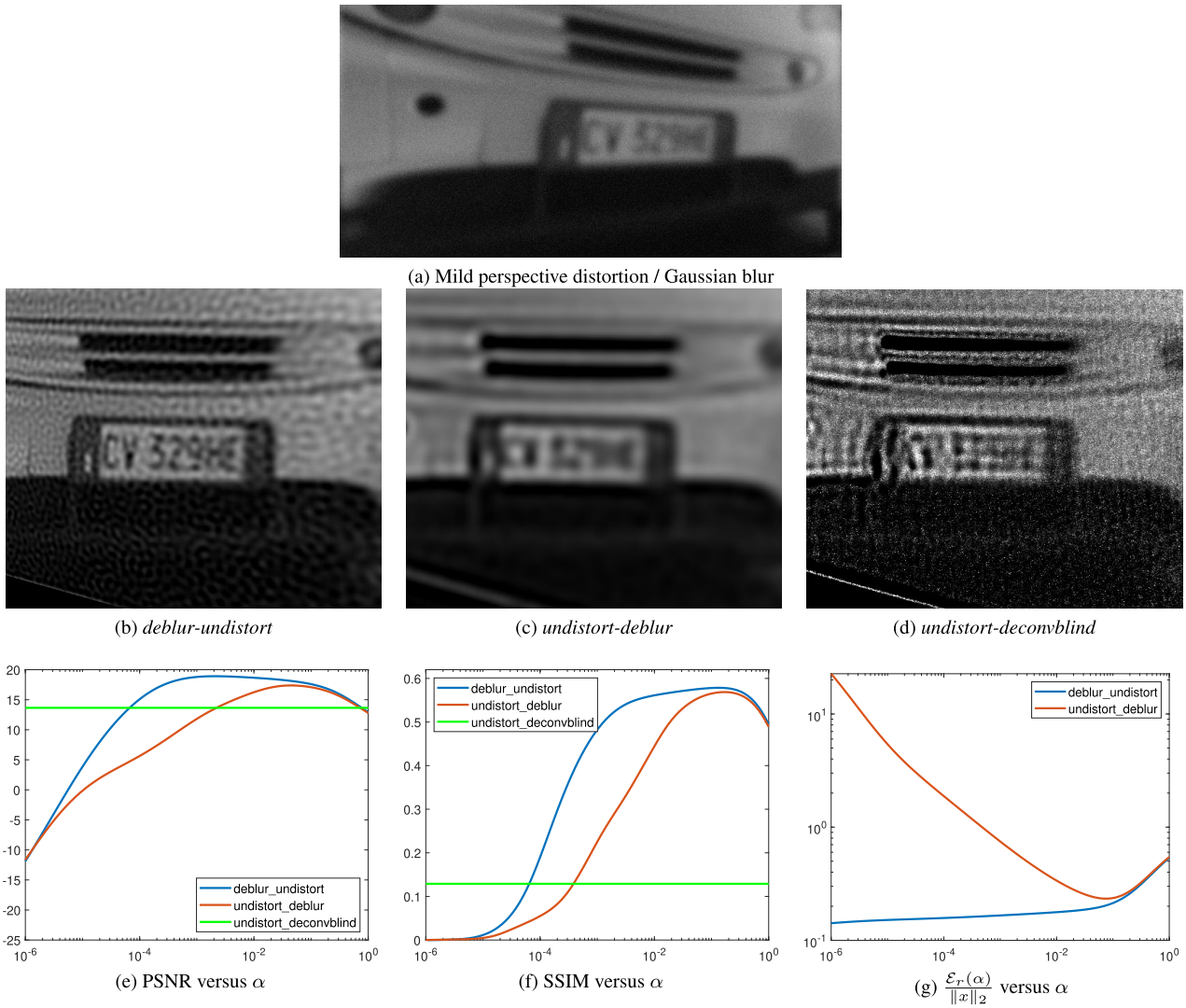
Proposition 1. *Let* $\Phi_1 \in \mathbb{R}^{r\times r}$ *be the diagonal matrix*

$$\Phi_1 = \begin{pmatrix} \frac{\sigma_1}{\sigma_1^2+\alpha} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\sigma_r}{\sigma_r^2+\alpha} \end{pmatrix}, \tag{12}$$

$\Sigma_r \in \mathbb{R}^{r\times r}$ *the diagonal matrix whose diagonal entries are the nonzero singular values* $\sigma_1, \ldots, \sigma_r$, $U_r \in \mathbb{R}^{n\times r}$ *and* $V_r \in \mathbb{R}^{n\times r}$ *the matrices containing respectively the first* r *columns of* U *and* V. *Then*

$$\|x - x_{rec}(y)\|_2 \le \underbrace{\alpha\|W^{-1}V_r\Phi_1\Sigma_r^{-2}U_r^T)\bar{y}\|_2}_{\textit{Regularization error}} + \underbrace{\|W^{-1}V_r\Phi_1U_r^T e\|_2}_{\textit{Perturbation error}}.$$

Proof: See the Appendix. □



(a) Mild perspective distortion / Gaussian blur

(b) *deblur-undistort* (c) *undistort-deblur* (d) *undistort-deconvblind*

(e) PSNR versus α (f) SSIM versus α (g) $\frac{\mathcal{E}_r(\alpha)}{\|x\|_2}$ versus α

FIGURE 3. (a) Image degraded with mild perspective degradation and Gaussian PSF. Reconstructions provided by: (b) *deblur-undistort* with $\alpha^*_{\mathbf{DU}}$; (c) *undistort-deblur* with $\alpha^*_{\mathbf{UD}}$; and (d) *undistort-deconvblind* workflows. Curves of: (e) PSNR versus α, (f) SSIM versus α and (g) relative regularization error versus α.

Therefore, we see that $\|x - x_{\text{rec}}(y)\|_2$ is upper bounded by the sum of two errors, the former related to the use of regularization in our restored image (regularization error), and the latter due to the presence of noise in the observed image (perturbation error). Both errors are controlled by the regularization parameter α. The regularization error $\mathcal{E}_r(\alpha)$ can be made arbitrarily small as α tends to zero as

$$\mathcal{E}_r(\alpha) = \alpha \|W^{-1} V_r \Phi_1 \Sigma_r^{-2} U_r^T) \bar{y}\|_2 \tag{13}$$

$$\leq \frac{\alpha}{(\sigma_r^2 + \alpha)\sigma_r} \|W^{-1}\|_2 \|\bar{y}\|_2, \tag{14}$$

as the matrices U_r and V_r are orthonormal and

$$\Phi_1 \Sigma_r^{-2} = \begin{pmatrix} \frac{1}{(\sigma_1^2+\alpha)\sigma_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{(\sigma_r^2+\alpha)\sigma_r} \end{pmatrix}.$$

Then, $\mathcal{E}_r(\alpha) = O\left(\frac{\alpha}{(\sigma_r^2+\alpha)\sigma_r}\right)$.

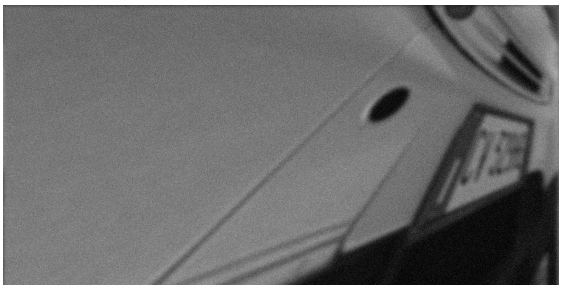
On the other hand, the perturbation error decreases as α increases since Φ_1 tends to the null matrix as $O(\frac{1}{\alpha})$. Consequently, the regularization parameter α has to be chosen in order to properly balance the two errors.

- When α is "too small", the regularization error is small, however, the perturbation error becomes large as it is controlled by the following norm

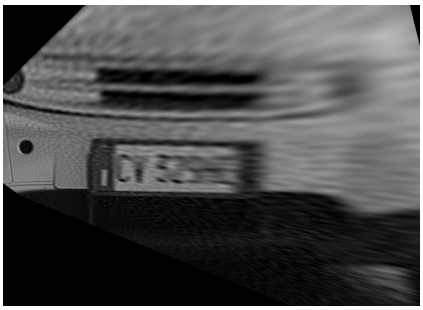
$$\|\Phi_1\|_2 = \sqrt{\max_{i=1,\ldots,r} \frac{\sigma_i}{\sigma_i^2 + \alpha}} \tag{15}$$

  and if $\alpha = 0$, the maximum in (15) is attained at $1/\sigma_r$, which can be large if σ_r is small.
- When α is "too large", the perturbation error is small, but the regularization error is inevitably large as, given the form of $z_{\text{Tik}}(y)$ in (11), the workflow output $x_{\text{rec}}(y) = W^{-1} z_{\text{Tik}}(y)$ is close to the null matrix.

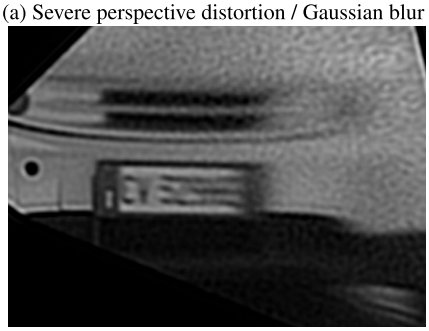
An optimal choice of the regularization parameter α must aim at balancing the two sources of errors. By assuming



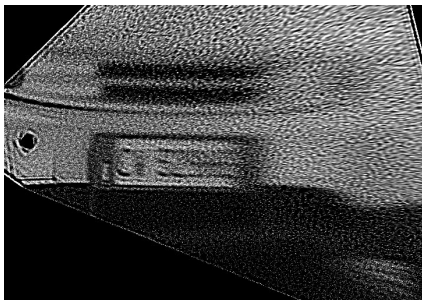
(a) Severe perspective distortion / Gaussian blur



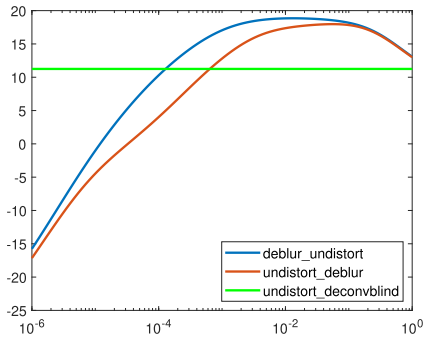
(b) *deblur-undistort*



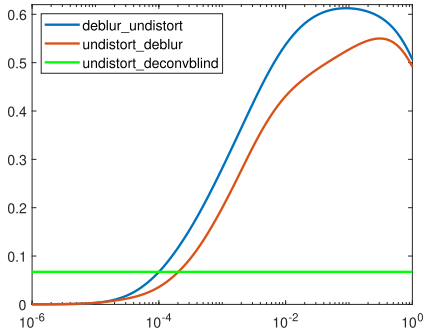
(c) *undistort-deblur*



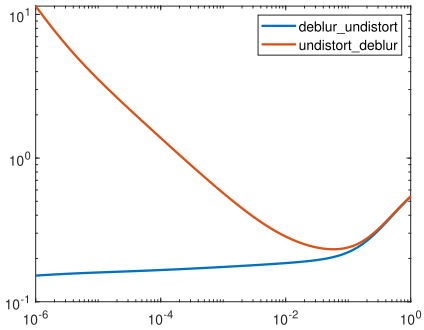
(d) *undistort-deconvblind*



(e) PSNR versus α



(f) SSIM versus α



(g) $\frac{\mathcal{E}_r(\alpha)}{\|x\|_2}$ versus α

FIGURE 4. **(a) Image degraded with severe perspective degradation and Gaussian PSF. Reconstructions provided by: (b) *deblur-undistort* with α^*_{DU}; (c) *undistort-deblur* with α^*_{UD}; and (d) *undistort-deconvblind* workflows. Curves of: (e) PSNR versus α, (f) SSIM versus α and (g) relative regularization error versus α.**

that $\|W^{-1}\|_2$ is moderate, the regularization parameter α can be properly chosen so that both the perturbation and the regularization error remain small, according to parameter choice methods such as the L-curve criterion or the Generalized Cross Validation (see, e.g., [14, Section 6.4] for further details).

C. ALTERNATIVE WORKFLOWS

It is worth noting that one could invert the order with which deblurring and perspective correction is performed and consider an alternative approach, according to which the blurred distorted image is first undistorted, then deblurred. The resulting workflow, named *undistort-deblur*, provides the following restored image

$$\begin{aligned}\tilde{x}_{\text{rec}}(y) &= x_{\text{Tik}}(W^{-1}y)\\ &= (H^TH + \alpha I_n)^{-1}H^TW^{-1}y.\end{aligned}$$

By proceeding as in the previous section, we can derive the following upper bound on the reconstruction error $\|\tilde{x}_{\text{rec}}(y) - x\|_2$.

Proposition 2. *We have*

$$\|\tilde{x}_{rec}(y) - x\| \le \underbrace{\|(V_r\Phi_1U_r^TW^{-1} - W^{-1}V_r\Sigma_r^{-1}U_r^T)\bar{y}\|_2}_{\textit{Regularization error}} + \underbrace{\|V_r\Phi_1U_r^TW^{-1}e\|_2}_{\textit{Perturbation error}}.$$

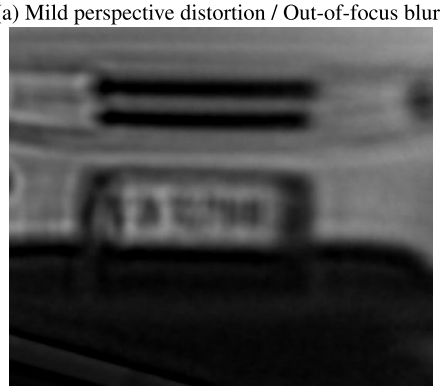
Proof: See the Appendix. □

Again the error can be decomposed in the sum of regularization and perturbation error. However with this approach the regularization error

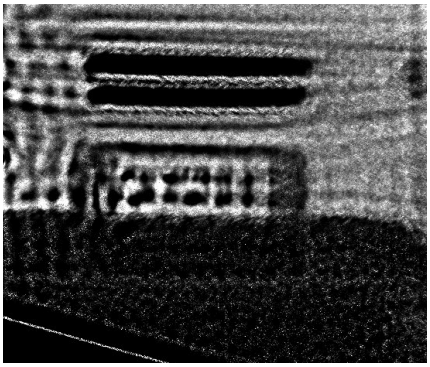
$$\tilde{\mathcal{E}}_r(\alpha) = \|(V_r\Phi_1U_r^TW^{-1} - W^{-1}V_r\Sigma_r^{-1}U_r^T)\bar{y}\|$$

(a) Mild perspective distortion / Out-of-focus blur

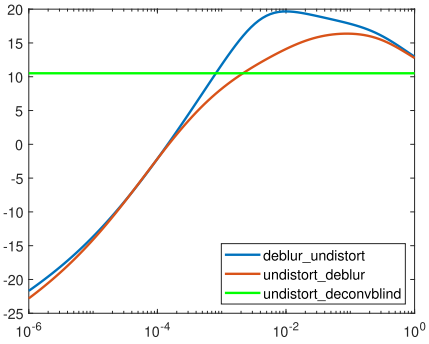
(b) *deblur-undistort*



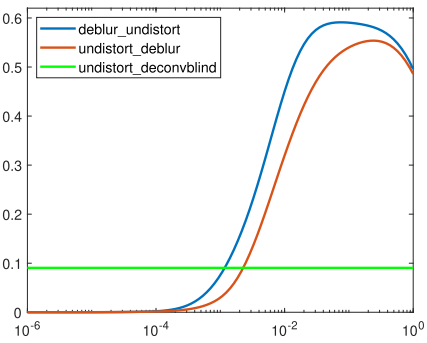
(c) *undistort-deblur*



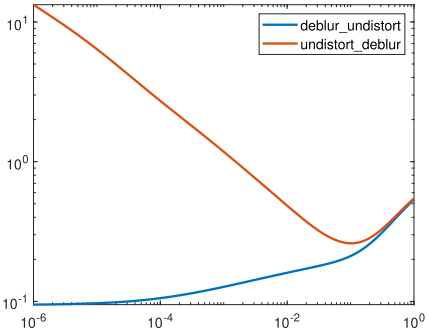
(d) *undistort-deconvblind*



(e) PSNR vs. α



(f) SSIM vs. α



(g) $\frac{\mathcal{E}_r(\alpha)}{\|x\|_2}$ versus α

FIGURE 5. **(a) Image degraded with mild perspective degradation and out-of-focus PSF. Reconstructions provided by: (b) *deblur-undistort* with α^*_{DU}; (c) *undistort-deblur* with α^*_{UD}; and (d) *undistort-deconvblind* workflows. Curves of: (e) PSNR versus α, (f) SSIM versus α and (g) relative regularization error versus α.**

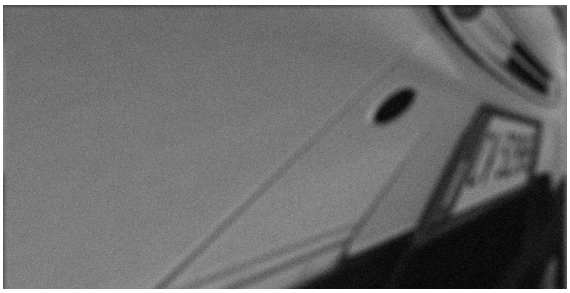
takes for $\alpha = 0$ the following strictly positive value that potentially can be large

$$\tilde{\mathcal{E}}_r(0) = \|(V_r \Sigma_r^{-1} U_r^T W^{-1} - W^{-1} V_r \Sigma_r^{-1} U_r^T)\bar{y}\|.$$

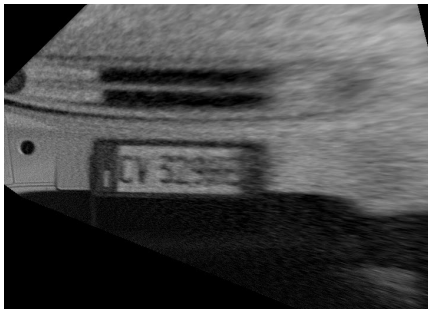
Since both $\mathcal{E}_r$ and $\tilde{\mathcal{E}}_r$ continuously depend on α we can conclude that there exists $\bar{\alpha}$ such that $\mathcal{E}_r(\alpha) \leq \tilde{\mathcal{E}}_r(\alpha)$ for $\alpha \in [0, \bar{\alpha}]$, while the behaviour of $\mathcal{E}_r$ and $\tilde{\mathcal{E}}_r$ is similar for large values of α as both $x_{\mathrm{rec}}(y)$ and $\tilde{x}_{\mathrm{rec}}(y)$ are close to the null matrix. We finally note that the perturbation error in both cases depends on $\|\Phi_1\|_2$ (see (15)).

As a third alternative, one could consider an approach similar to *undistort-deblur*, where the deblurring step via Tikhonov regularization is replaced by a *blind deconvolution algorithm*. We denote this additional workflow by *undistort-deconvblind*. We recall that blind deconvolution is the problem of recovering both the unknown image and the PSF from the observed image. Indeed, when perspective is corrected first, we must take into account that the blur's PSF is also distorted in an unknown way. Therefore, even if the blur's PSF is known, it makes sense to estimate it in this scenario. On the other hand, when compensating for optical blur first, there is no need to estimate the shape of the PSF assuming it is known (for this reason, we do not consider a *deconvblind-undistort* workflow).

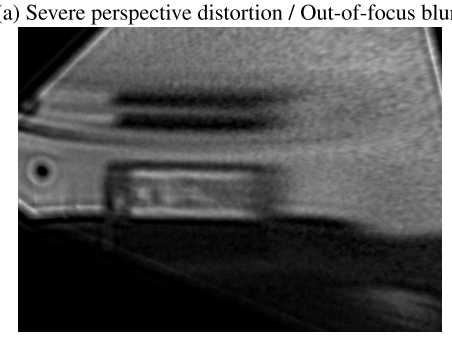
Blind deconvolution is typically reformulated as a non-convex minimization problem through the MAP approach and then solved using alternating minimization algorithms to numerically estimate both the image and the PSF [19], [20]. Therefore, we define an additional workflow, called *undistort-deconvblind*, where the perspective



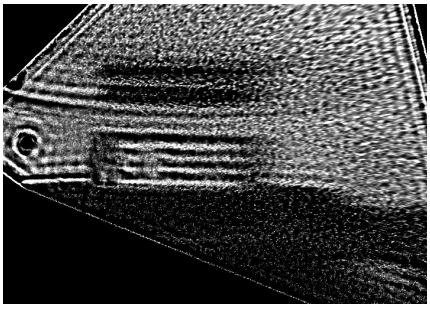
(a) Severe perspective distortion / Out-of-focus blur



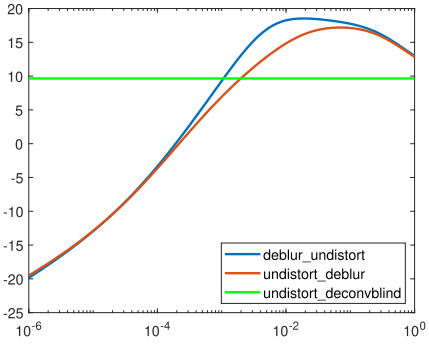
(b) *deblur-undistort*



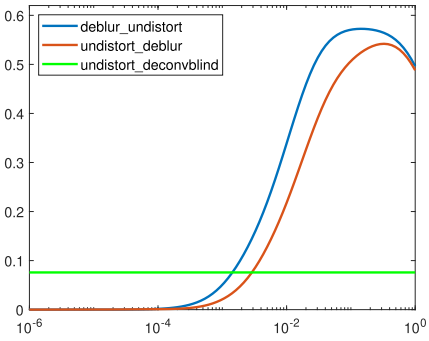
(c) *undistort-deblur*



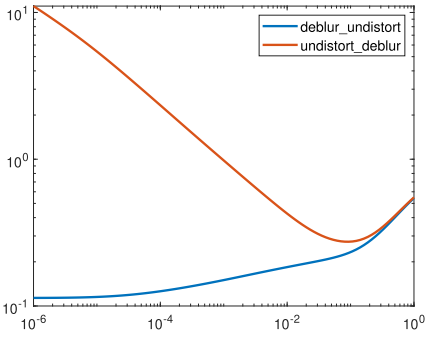
(d) *undistort-deconvblind*



(e) PSNR versus α



(f) SSIM versus α



(g) $\frac{\mathcal{E}_r(\alpha)}{\|x\|_2}$ versus α

FIGURE 6. **(a) Image degraded with severe perspective degradation and out-of-focus PSF. Reconstructions provided by: (b) *deblur-undistort* with α^*_{DU}; (c) *undistort-deblur* with α^*_{UD}; and (d) *undistort-deconvblind* workflows. Curves of: (e) PSNR versus α, (f) SSIM versus α and (g) relative regularization error versus α.**

correction step is followed by a fixed number of iterations of a blind deconvolution algorithm, which is performed in practice by calling the `deconvblind` MATLAB built-in function (see the next Section for more implementation details) [21], [22], [23]. With this additional method, we investigate whether an automatically estimated PSF might improve the performance of the workflow compared to those obtained with *undistort-deblur*.

III. EXPERIMENTAL RESULTS

In this section, we present some experimental results to evaluate the performance of the different restoration workflows in the inversion of problem (3). The experiments refer to the case study identified in the previous section, which is an image affected by perspective and blur distortion. The aim of these tests is to evaluate which restoration workflow, among *deblur–undistort*, *undistort–deblur*, and *undistort–deconvblind*, yields the best results.

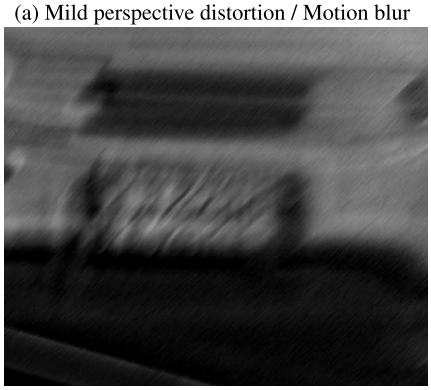
We provide two different experimental setups. In the first one, we start from a ground truth image, synthetically degrade such image, and then restore the degraded version and compute its similarity against the ground truth. This setup allows us to objectively assess and compare the performance of the various restoration workflows. In the second setup, we instead work with real images that are affected by a combination of blur and perspective. In this setup, the ground truth image is not available, so the comparison between the performance of different methods is left to visual inspection.

A. EVALUATION ON SYNTHETICALLY DEGRADED IMAGES

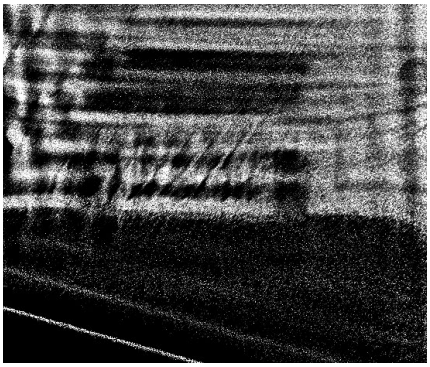
All the numerical experiments hereby presented have been performed in MATLAB on a PC equipped with an Intel

(a) Mild perspective distortion / Motion blur

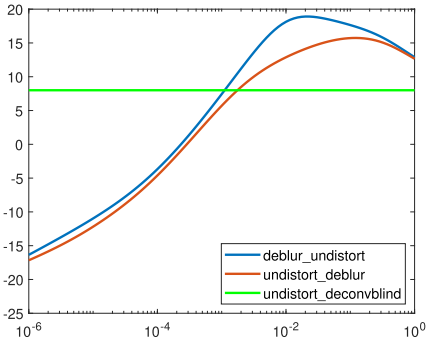
(b) *deblur-undistort*



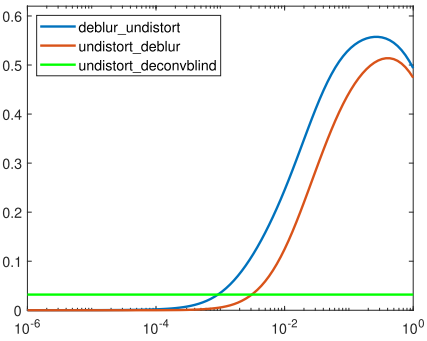
(c) *undistort-deblur*



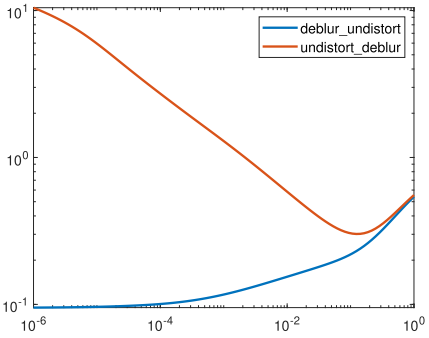
(d) *undistort-deconvblind*



(e) PSNR versus α



(f) SSIM versus α



(g) $\frac{\mathcal{E}_r(\alpha)}{\|x\|_2}$ versus α

FIGURE 7. (a) Image degraded with mild perspective degradation and motion PSF. Reconstructions provided by: (b) *deblur-undistort* with α^*_{DU}; (c) *undistort-deblur* with α^*_{UD}; and (d) *undistort-deconvblind* workflows. Curves of: (e) PSNR versus α, (f) SSIM versus α and (g) relative regularization error versus α.

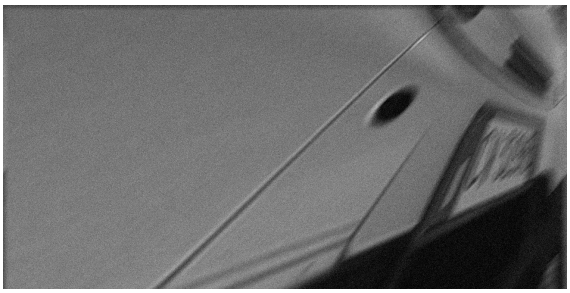
Core i7-6500 CPU@2.5GHz, with 16GB RAM, running Windows 10 Pro.

1) EXPERIMENTAL SETUP

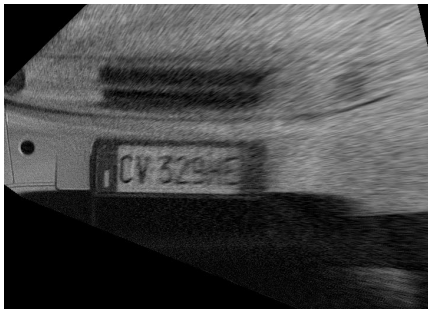
We used a grayscale test image of 1309 × 1109 pixels, representing a frontal view of a car with its license plate (see Figure 2), referred to as `license plate` in the remainder of the paper. The original undistorted image has been synthetically degraded according to the acquisition model in (3), that is by using the following three-steps procedure:

1) the original `license plate` image has been distorted by using the MATLAB built-in functions `fitgeotrans` (to define a projective transformation through some hand-picked control points) and `imwarp` (to apply such a transformation to the ground truth);
2) a blurring kernel is created by using the MATLAB function `fspecial` and then applied to the previously distorted image; the convolution is performed assuming periodic boundary conditions, which enables us to use Fast Fourier Transform (FFT) processing;
3) a white Gaussian noise vector, with zero mean and variance $\sigma^2 = 0.05$, is created with the MATLAB function `randn` and added to the blurred image.

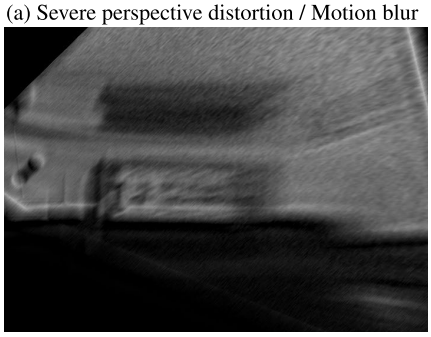
By using the aforementioned procedure and varying the projective transformation and the blur parameters, we generate six different distorted images, in order to simulate light to severe degradations. In particular, we consider two different projective transformations, inducing a mild and a strong perspective distortion, respectively, in order to simulate (in



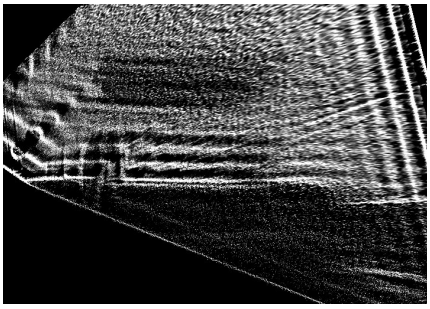
(a) Severe perspective distortion / Motion blur



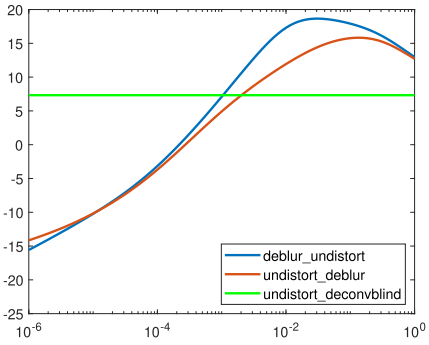
(b) *deblur-undistort*



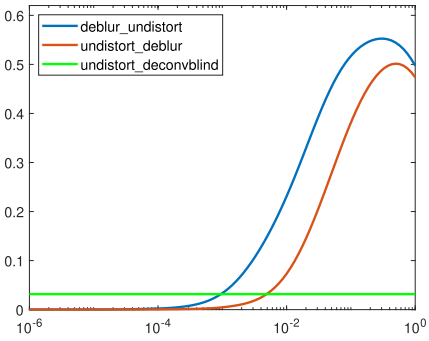
(c) *undistort-deblur*



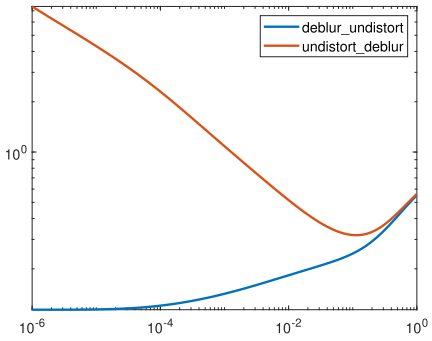
(d) *undistort-deconvblind*



(e) PSNR versus α



(f) SSIM versus α



(g) $\frac{\mathcal{E}_r(\alpha)}{\|x\|_2}$ versus α

FIGURE 8. **(a) Image degraded with severe perspective degradation and motion PSF. Reconstructions provided by: (b) *deblur-undistort* with α^*_{DU}; (c) *undistort-deblur* with α^*_{UD}; and (d) *undistort-deconvblind* workflows. Curves of: (e) PSNR versus α, (f) SSIM versus α and (g) relative regularization error versus α.**

the latter case) also very hard recognition problems. As for the blurring, we consider three different types of kernels: Gaussian blur (with zero mean and standard deviation σ); out-of-focus blur (uniform PSF equal to 1 within a radius r, and 0 elsewhere); motion blur (of length l and angle θ with respect to the horizontal axis). The parameters of the PSFs used to create the six degraded images are reported in Table 1.

Each distorted image is restored according to the three different workflows presented in Section II, using the MATLAB implementation of the Tikhonov deblurring method described in [14, Section 6.2]. Each restored image is then compared against the ground truth; we limit the comparison to the license plate region only, simulating a real user scenario of license plate recognition.

To evaluate the quality of the reconstructed images, we use both qualitative criteria (visual inspection) and the following image quality metrics:

- Relative Reconstruction Error (RRE), defined as:

$$\mathrm{RRE}(x, x_{\mathrm{rec}}) = \frac{\|x - x_{\mathrm{rec}}\|_2}{\|x\|_2};$$

- Peak Signal-to-Noise Ratio (PSNR), defined as:

$$\mathrm{PSNR}(x, x_{\mathrm{rec}}) = 10 \log_{10} \left(\frac{255^2}{\frac{1}{n}\|x - x_{\mathrm{rec}}\|^2} \right),$$

  where 255 is the maximum possible value of the image, whereas the denominator is the mean squared error with n the number of pixels of x;
- Structural Similarity Index Measure (SSIM) [24], defined as

$$\mathrm{SSIM}(x, x_{\mathrm{rec}}) = \frac{(2\mu_x \mu_{x_{\mathrm{rec}}} + c_1)(2\sigma_{x x_{\mathrm{rec}}} + c_2)}{(\mu_x^2 + \mu_{x_{\mathrm{rec}}}^2 + c_1)(\sigma_x^2 + \sigma_x^2 + c_2)},$$

FIGURE 9. Thumbnails of real images considered for restoration. First row: images acquired with a Nikon D50 camera, affected by a combination of optical blur and perspective distortion (first two images) and by motion blur and perspective distortion (third image). Second row, left: image acquired with a Sony Xperia XA1 affected by a combination of optical blur and strong perspective distortion. Second row, right: image acquired with an Apple iPhone XS, affected by a combination of optical blur and mild perspective distortion.

FIGURE 10. Comparison of the three considered restoration workflows on a real image of a book affected by a combination of optical blur and mild perspective distortion. Top-left: original image; top-right: proposed workflow (deblur followed by perspective correction); bottom-left: perspective correction followed by deblur; bottom right: perspective correction followed by blind deconvolution.

where μ_x, $\mu_{x_{\text{rec}}}$ are the pixel sample means, σ_x, $\sigma_{x_{\text{rec}}}$ the standard deviations, $\sigma_{xx_{\text{rec}}}$ the covariance, and c_1, c_2 two variables to stabilize the division with weak denominator. In practice, all the mentioned statistics are

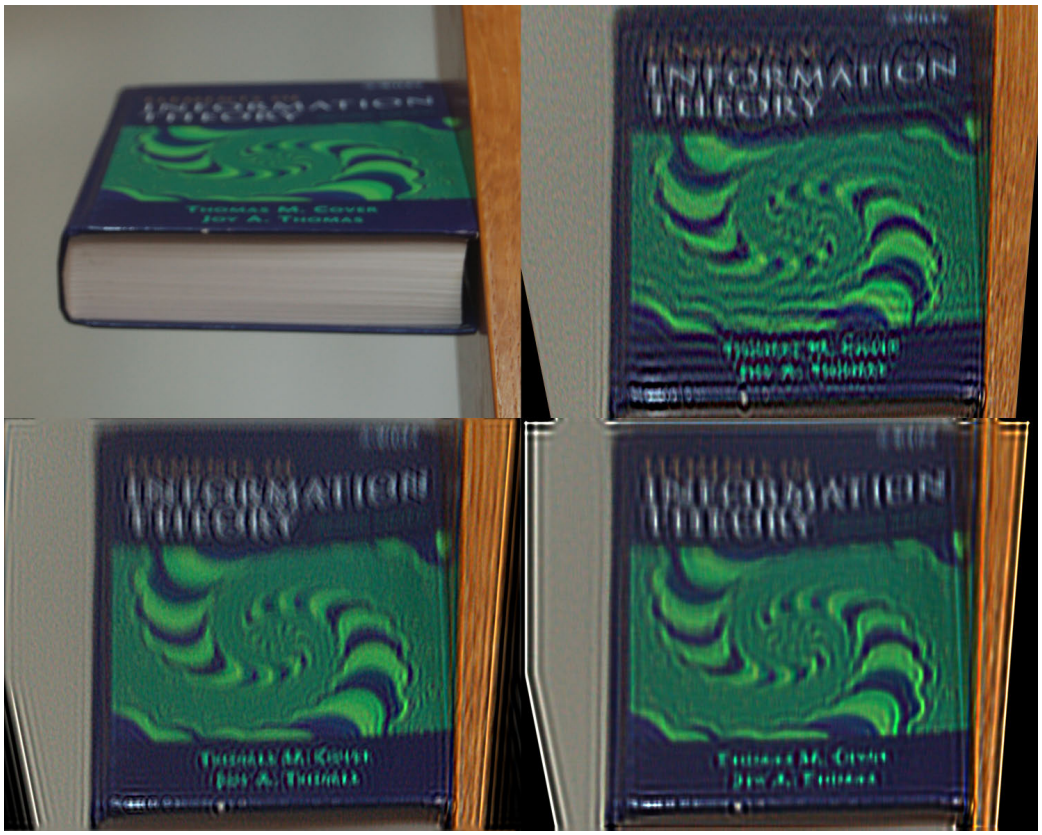


FIGURE 11. Comparison of the three considered restoration workflows on a real image of a book affected by a combination of optical blur and strong perspective distortion. Top-left: original image; top-right: proposed workflow (deblur followed by perspective correction); bottom-left: perspective correction followed by deblur; bottom right: perspective correction followed by blind deconvolution.

computed on windows centered around each pixel, and the global SSIM value is obtained by averaging the local SSIM ones.

We remark that SSIM aims to measure the perceptual difference between images, taking into account the local brightness and texture of images; on the other hand, the RRE and the PSNR calculate an arithmetical similarity of images that, often, is not descriptive of perceptive image quality. Moreover, while RRE lower values indicate a higher reconstruction quality, the opposite holds for both PSNR and SSIM.

2) RESULTS ON SYNTHETICALLY DEGRADED IMAGES

For each distorted image, we run both the *deblur-undistort* and *undistort-deblur* workflows for 100 logarithmically distributed values of α in the interval $[10^{-6}, 1]$. Then, for each of the two workflows, we select the value of α that maximizes one of the quality metrics and consider the corresponding restored image as the ''best reconstruction''.

Tables 2, 3, and 4 report the best RRE, PSNR, and SSIM values, respectively, provided by the two workflows *deblur-undistort* and *undistort-deblur*, together with the corresponding values given by the *undistort-deconvblind* workflow. Figure 3 shows the degraded image obtained with mild perspective distortion and Gaussian blur, the reconstructed images obtained with the three different workflows, the plots of the PSNR and SSIM metrics versus the α parameter, and the plot of the quantity $\mathcal{E}_r(\alpha)/\|x\|_2$ versus the α parameter, being $\mathcal{E}_r(\alpha)$ the regularization error defined in (14). The reconstructed images were obtained with the value of α that maximizes the PSNR metric, denoted as α^*_{DU} for the *deblur-undistort* case, and α^*_{UD} for the *undistort-deblur* case. Analogously, Figures from 4 to 8 show the same data for the other degraded images.

By considering the objective metrics presented in Tables 2-4, we observe that the *deblur-undistort* workflow outperforms the others in all degradation scenarios and according to every similarity metric. This result confirms our claim, based on intuition and analytically explained in Section II-B and II-C, that reversing the order of defects provides the best results.

The same conclusion can also be reached by visually inspecting the reconstructed images reported in Figures 3-8. From the PSNR and SSIM curves shown in the same figures, we can observe that *deblur-undistort* yields the best PSNR/SSIM values. Moreover, it is also more robust to little variations of α, whereas the *undistort-deblur* workflow performance rapidly decreases when moving away from

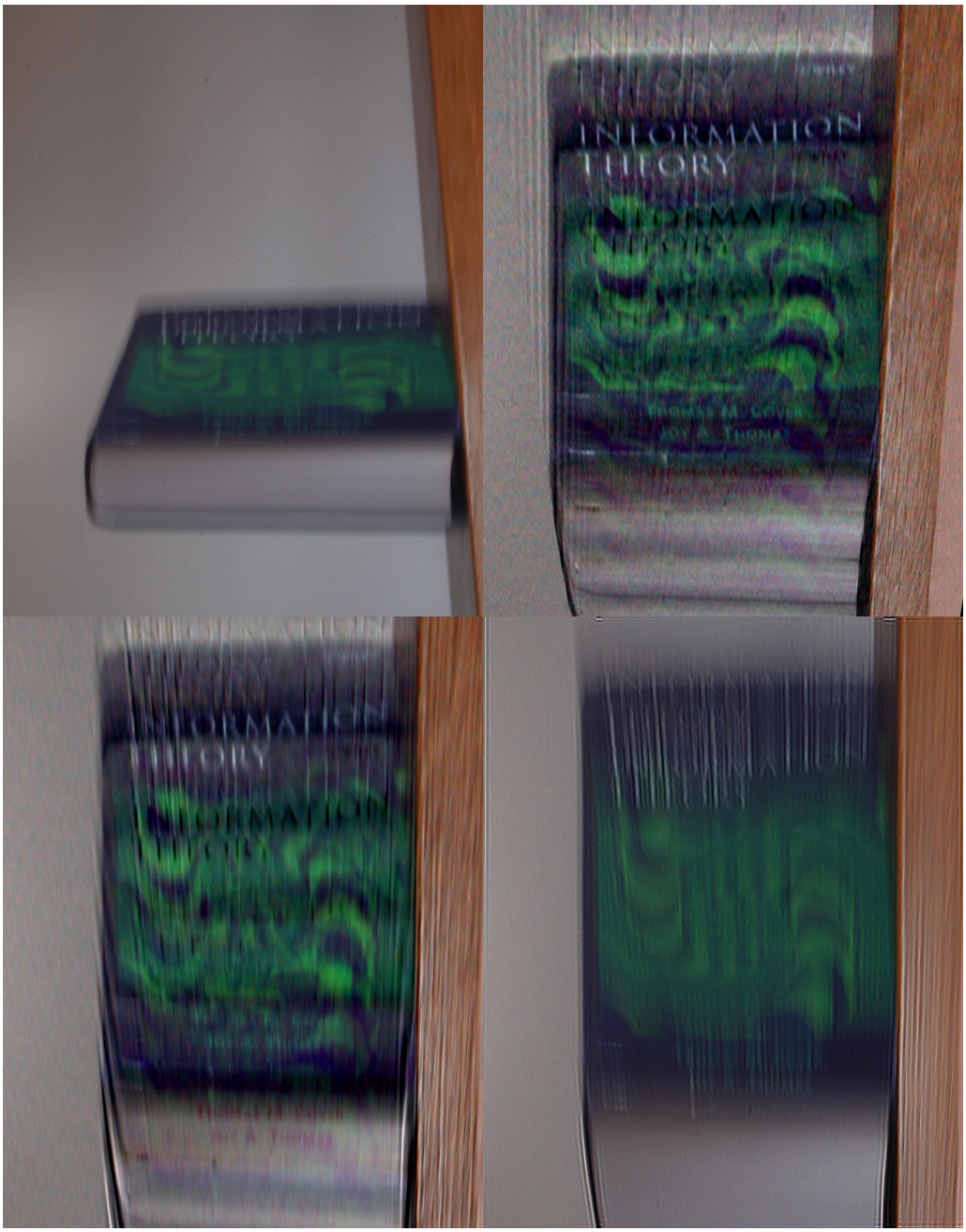


FIGURE 12. Comparison of the three considered restoration workflows on a real image of a book affected by a combination of motion blur and perspective distortion. Top-left: original image; top-right: proposed workflow (deblur followed by perspective correction); bottom-left: perspective correction followed by deblur; bottom right: perspective correction followed by blind deconvolution.

the optimal value. The greater robustness of our proposed workflow is confirmed by the plots of the regularization error, which remains small for a wide range of values of α in *deblur-undistort* while rapidly increasing as α decreases in *undistort-deblur*. This is in agreement with the analysis carried out in Section II-B and II-C, according to which the regularization error vanishes as α tends to zero for *deblur-undistort* but not for *undistort-deblur*. On the other hand, the reconstructions provided by *undistort-deconvblind* are unacceptable in almost all cases, as some serious ringing artifacts prevent any reading of the license plate. This confirms that correcting the defects in the same order as they were introduced in the model does not work, even when an automatically estimated PSF is employed.

B. EVALUATION ON REAL IMAGES

In this section, we aim to compare the performance of the various workflows on real images that are originally

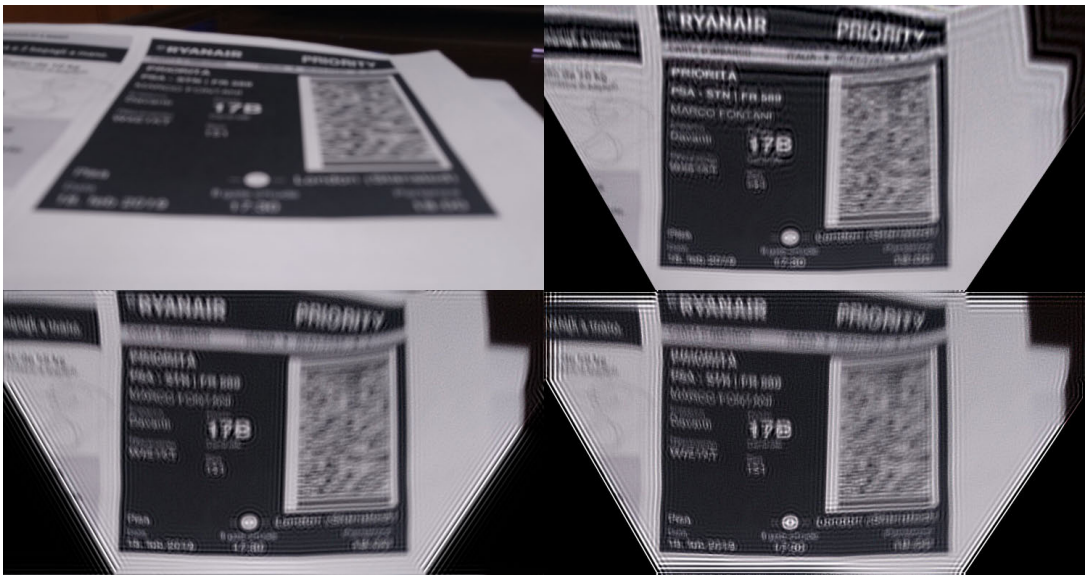


FIGURE 13. Comparison of the three considered restoration workflows on a real image of a boarding pass affected by a combination of optical blur and strong perspective distortion. Top-left: original image; top-right: proposed workflow (deblur followed by perspective correction); bottom-left: perspective correction followed by deblur; bottom right: perspective correction followed by blind deconvolution.

affected by perspective distortion and blur. A thumbnail of the considered images is visible in Figure 9, and their full-resolution versions are available for download.[3] Images have been obtained from various source devices: a Nikon D50 reflex camera was used to create images of a book with different viewpoints in the top row, a Sony Xperia XA1 was used to capture the boarding pass image, and the license plate image was obtained with an Apple iPhone XS. We acknowledge that the selected images do not closely match the typical imagery encountered by forensic video analysts; on the other hand, this approach allowed us to control the strength of the distortion introduced in the image and facilitates the comparison between different workflows.

1) EXPERIMENTAL SETUP

To maximize the realisticness of the evaluation, the experiments presented below have been carried out using Amped FIVE,[4] version 29850, a commercial forensic image enhancement software used by law enforcement agencies worldwide. Specifically, the following filters were used to carry out the processing:

- *Correct Perspective* to correct the perspective distortion. The filter lets the user click on the four vertices of the element of interest (which must be planar) and then drawing the output rectangle.
- *Optical Deblurring* to correct the out-of-focus blur. The user manually sets the radius of the circular PSF, and a slider is provided to estimate the noise-to-signal power ratio: lower values lead to a sharper but noisier image, while larger values lead to a smoother but cleaner image.
- *Motion Deblurring* to correct the motion blur. The filter lets the user manually define the PSF associated with the motion in terms of length and angle. The same noise-to-signal power ratio slider discussed above is presented.
- *Blind Deconvolution* to correct both kinds of blur by automatically determining the shape of the PSF, starting from a radius size provided by the user. The same noise-to-signal power ratio slider discussed above is presented.

We employed the following experimental protocol: each image was loaded into Amped FIVE and possibly rotated by multiples of 90 degrees and cropped when needed. We then applied the filters mentioned above in different orders to implement the three proposed workflows. Each time, we manually optimized each filter's parameters in such a way as to obtain the best visual result. We then used the *Multiview* filter to create a grid with four composite images (Figures 10-14): in each of them, the original image is at the top-left, the image enhanced with the proposed deblur-perspective correction workflow is presented at the top-right, the image enhanced with the perspective correction-deblur is shown at the bottom-left, and the image enhanced with perspective correction followed by blind deconvolution is shown at the bottom right. For each processed image, the detailed settings applied for each filter are provided in a PDF report generated by Amped FIVE, which is available in the same archive containing the

[3] https://www.dropbox.com/scl/fi/rl1rh6ukwfmpw8z4pyaun/DigInv_RestorationWorkflow_RealImagesData.zip?rlkey=nwztqg9a6v02wc2kwcj1bjic1&dl=0

[4] https://ampedsoftware.com/five

FIGURE 14. Comparison of the three considered restoration workflows on a real image of a license plate affected by a combination of optical blur and mild perspective distortion. Top-left: original image; top-right: proposed workflow (deblur followed by perspective correction); bottom-left: perspective correction followed by deblur; bottom right: perspective correction followed by blind deconvolution.

original images; the web address to access it is provided in Section III-B.

2) RESULTS ON REAL DEGRADED IMAGES

Since there is no objective and reliable way to obtain a ground truth reference for real degraded images, we leave the interpretation of results presented in Figures 10-14 to the reader. It is our opinion that, for all the considered examples, the top-right image (corresponding to the proposed deblur-undistort workflow) yields the best result, followed by the undistort-deblur (bottom left), while the image restored with blind deconvolution (bottom right) is visually less appealing. Not surprisingly, the advantage of using the deblur-undistort workflow is more evident for images where a strong perspective distortion is observed (specifically, Figures 11 and 13).

Noticeably, these results are in total agreement with what we observed in the experiments on synthetically degraded images.

IV. CONCLUSION

While the scientific literature is rich in image enhancement and restoration techniques, researchers gave little attention to investigating the *order* by which one should apply these tools. This work deals with such a question, providing several contributions. We argued that the video enhancement workflow should compensate for defects in the reverse order, which means that the last introduced defect should be compensated first, and so on. We then provided an analytical model of a case study that considers three widespread defects (optical and motion blur, perspective distortion, and additive noise). Using that model, we showed mathematically and experimentally that the best restoration order is the reverse order, as initially argued. Even though the examined case study contains three specific degradations operators, it adheres to many practical scenarios (blur, perspective, and noise are among the most common defects in surveillance videos). The presented methodological approach can be applied to other chains of operators, provided that a restoration model can be formulated for them. Instead, for the case of lossy degradations, for which no inverse operator is defined, a broader generalization of the framework is required. These points are left to future work.

APPENDIX

Proof of Proposition 1: Letting $\bar{y} = y - e$ and denoting by u_i and v_i the i-th column of the matrices U and V respectively, we can write $z_{\mathrm{Tik}}(y)$ given in (11) as

$$z_{\mathrm{Tik}}(y) = \sum_{i=1}^{r} \frac{\sigma_i}{\sigma_i^2 + \alpha} u_i^T (\bar{y} + e) v_i.$$

By (4) it holds

$$z = \sum_{i=1}^{r} \frac{1}{\sigma_i} u_i^T \bar{y} v_i + \bar{z}, \tag{16}$$

where $\bar{z}$ is an arbitrary vector of the null space of H and $\bar{z} = 0$ if z is the minimum norm solution to $Hz = \bar{y}$.

Therefore, assuming that z is the minimum norm solution, we get

$$\begin{aligned} z_{\text{Tik}}(y) - z &= \sum_{i=1}^{r} \left(\frac{\sigma_i}{\sigma_i^2 + \alpha} - \frac{1}{\sigma_i} \right) u_i^T \bar{y} v_i \\ &\quad + \sum_{i=1}^{r} \frac{\sigma_i}{\sigma_i^2 + \alpha} u_i^T e v_i \\ &= -\sum_{i=1}^{r} \frac{\alpha}{\sigma_i^2 + \alpha} \frac{u_i^T \bar{y}}{\sigma_i} v_i + \sum_{i=1}^{r} \frac{\sigma_i}{\sigma_i^2 + \alpha} u_i^T e v_i. \end{aligned}$$

Then, using the matrix Φ_1 given in (12) and the matrices Σ_r, U_r and V_r, we can rewrite $z_{\text{Tik}}(y) - z$ as follows:

$$z_{\text{Tik}}(y) - z = -\alpha V_r \Phi_1 \Sigma_r^{-2} U_r^T \bar{y} + V_r \Phi_1 U_r^T e. \tag{17}$$

By using (17), (5) and (9), we conclude that

$$\begin{aligned} \|x - x_{\text{rec}}(y)\|_2 &\leq \alpha \|W^{-1} V_r \Phi_1 \Sigma_r^{-2} U_r^T \bar{y}\|_2 \\ &\quad + \|W^{-1} V_r \Phi_1 U_r^T e\|_2. \end{aligned}$$

Proof of Proposition 2: By using the SVD decomposition of H, we reformulate the output of the workflow *undistort-deblur* as follows

$$\begin{aligned} \tilde{x}_{\text{rec}}(y) &= x_{\text{Tik}}(W^{-1} y) \\ &= (H^T H + \alpha I_n)^{-1} H^T W^{-1} y \\ &= \sum_{i=1}^{r} \frac{\sigma_i}{\sigma_i^2 + \alpha} u_i^T (W^{-1} \bar{y} + W^{-1} e) v_i \end{aligned}$$

Assuming that z is the minimum-norm solution to $Hz = \bar{y}$, by (16) we get $z = \sum_{i=1}^{r} \frac{1}{\sigma_i} u_i^T \bar{y} v_i$ and

$$x = \sum_{i=1}^{r} \frac{1}{\sigma_i} u_i^T \bar{y} W^{-1} v_i.$$

Then, proceeding as in Proposition 1, we write the difference between the ground truth and workflow output as

$$\begin{aligned} \tilde{x}_{\text{rec}}(y) - x &= \sum_{i=1}^{r} \left(\frac{\sigma_i}{\sigma_i^2 + \alpha} u_i^T W^{-1} \bar{y} v_i - \frac{1}{\sigma_i} u_i^T \bar{y} W^{-1} v_i \right) \\ &\quad + \sum_{i=1}^{r} \frac{\sigma_i}{\sigma_i^2 + \alpha} u_i^T W^{-1} e v_i \end{aligned}$$

and by taking the norm of both side we obtain

$$\begin{aligned} \|\tilde{x}_{\text{rec}}(y) - x\| &\leq \|(V_r \Phi_1 U_r^T W^{-1} - W^{-1} V_r \Sigma_r^{-1} U_r^T) \bar{y}\|_2 \\ &\quad + \|V_r \Phi_1 U_r^T W^{-1} e\|_2. \end{aligned}$$

□

REFERENCES

[1] M. P. J. Ashby, "The value of CCTV surveillance cameras as an investigative tool: An empirical analysis," *Eur. J. Criminal Policy Res.*, vol. 23, no. 3, pp. 441–459, Sep. 2017.
[2] Y. Jung and A. P. Wheeler, "The effect of public surveillance cameras on crime clearance rates," *J. Experim. Criminol.*, vol. 19, no. 1, pp. 143–164, Mar. 2023.
[3] F. Brookman and H. Jones, "Capturing killers: The construction of CCTV evidence during homicide investigations," *Policing Soc.*, vol. 32, no. 2, pp. 125–144, Feb. 2022.
[4] D. Seckiner, X. Mallett, C. Roux, D. Meuwly, and P. Maynard, "Forensic image analysis—CCTV distortion and artefacts," *Forensic Sci. Int.*, vol. 285, pp. 77–85, Apr. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0379073818300380
[5] S. A. Ledesma, *A Proposed Framework for Forensic Image Enhancement*. Denver, CO, USA: Univ. Colorado, 2015.
[6] X. Zhang, J. Jiang, and S. Peng, "Commutability of blur and affine warping in super-resolution with application to joint estimation of triple-coupled variables," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1796–1808, Apr. 2012.
[7] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 257–265.
[8] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, and L. Shao, "Multi-stage progressive image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 14816–14826.
[9] J. Koh, J. Lee, and S. Yoon, "Single-image deblurring with neural networks: A comparative survey," *Comput. Vis. Image Understand.*, vol. 203, Feb. 2021, Art. no. 103134.
[10] K. Zhang, W. Ren, W. Luo, W.-S. Lai, B. Stenger, M.-H. Yang, and H. Li, "Deep image deblurring: A survey," *Int. J. Comput. Vis.*, vol. 130, no. 9, pp. 2103–2130, Sep. 2022.
[11] Eur. Union. (Dec. 2023). *Artificial Intelligence Act*. [Online]. Available: https://artificialintelligenceact.eu/
[12] *Best Practices for Digital Forensic Video Analysis*, Scientific Work. Group Digit. Evidence (SWGDE), 2018.
[13] A. K. Jain, *Fundamentals of Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1989.
[14] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring Images: Matrices, Spectra and Filtering*. Philadelphia, PA, USA: SIAM, 2006.
[15] M. Bertero, H. Lantéri, and L. Zanni, "Iterative image reconstruction: A point of view," in *Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT)*, Y. Censor, M. Jiang, and A. K. Louis, Eds. Pisa, Italy: Birkhauser-Verlag, 2008, pp. 37–63.
[16] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
[17] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. Bristol, U.K.: Institute of Physics, 1998.
[18] N. A. Tikhonov and V. Y. Arsenin, *Solution of Ill Posed Problems*. New York, NY, USA: Wiley, 1977.
[19] G. R. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," *Opt. Lett.*, vol. 13, no. 7, pp. 547–549, Jul. 1988.
[20] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1964–1971.
[21] D. S. C. Biggs and M. Andrews, "Acceleration of iterative image restoration algorithms," *Appl. Opt.*, vol. 36, no. 8, p. 1766, Mar. 1997. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-36-8-1766
[22] R. J. Hanisch, R. L. White, and R. L. Gilliland, *Deconvolution of Hubble Space Telescope Images and Spectra*. Millbrae, CA, USA: Academic, 1997, pp. 310–360.
[23] T. J. Holmes, S. Bhattacharyya, J. A. Cooper, D. Hanzel, V. Krishnamurthi, W.-C. Lin, B. Roysam, D. H. Szarowski, and J. N. Turner, *Light Microscopic Images Reconstructed By Maximum Likelihood Deconvolution*. New York, NY, USA: Plenum Press, 1995, pp. 389–402, doi: 10.1007/978-1-4757-5348-6_24.
[24] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

FABRIZIO ARGENTI (Senior Member, IEEE) received the Laurea degree (cum laude) in electronics engineering and the Ph.D. degree in electronics and information engineering from the University of Florence, Florence, Italy, in 1989 and 1993, respectively. Since 1993, he has been with the Department of Information Engineering (formerly Department of Electronics and Telecommunications), University of Florence, where he was initially an Assistant Professor and then has been an Associate Professor, since 2002. His teaching experience includes courses on digital signal processing, estimation theory, and information theory. He has been involved in several research projects in the fields of image processing, multimedia transmission, and remote sensing. His research interests include image processing, statistical signal processing, inverse problems, and machine learning approaches to signal processing.

STEFANIA BELLAVIA received the degree in mathematics from the University of Florence, Italy, in 1993, and the Ph.D. degree in computational mathematics and informatics from the University of Padua, in 1997. Since 1999, she has been with the University of Florence. She is currently a Full Professor in numerical analysis. She is currently an Expert in numerical methods for nonlinear optimization problems. Her recent research interest includes stochastic optimization methods for optimization problems arising in machine learning. She also has a strong expertise in linear algebra methods for large scale problems.

MARCO FONTANI (Member, IEEE) received the M.Sc. degree (summa cum laude) in computer engineering from the University of Firenze, Italy, in 2010, and the Ph.D. degree in information engineering from the University of Siena, Italy, in 2014.

His research interest includes multimedia forensics. His doctoral dissertation, "Digital Forensic Techniques for Splicing Detection in Multimedia Contents," was awarded the Best 2015 Ph.D. Thesis on Signal Processing by Italian National Telecommunications and Information Theory Group (GTTI). He is currently the Forensic Director of Amped Software, a software company developing image and video forensic solutions for law enforcement agencies worldwide. He participated in several research projects funded by the EU and EOARD. He has authored or coauthored over 30 works published in journals or conference proceedings. He also has experience delivering training to law enforcement and provided expert witness testimony on several forensic cases involving digital images and videos.

Dr. Fontani is a Former Member of the IEEE Information Forensics and Security Technical Committee.

GABRIELE GUARNIERI was born in 1982. He received the degree (summa cum laude) in electronics engineering and the Ph.D. degree in information engineering/signal and image processing from the University of Trieste, Italy, in 2005 and 2009, respectively.

He was a Visiting Researcher with the Image Processing Research Laboratory, University of California at Santa Barbara, working with Prof. Sanjit Mitra, and the Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD, USA, working with Dr. Aldo Badano. After graduating, he collaborated in the EU/ARTEMIS research project "CHIRON" and was an Adjunct Professor in digital image processing with the University of Trieste. Since 2014, he has been an Image Processing Researcher with Amped Software. He is the author or coauthor of several journal articles and conference proceedings. His primary research interests include high dynamic range imaging, with special attention to computational efficiency and accurate color management.

MARTINO JERIAN (Member, IEEE) received the Graduate degree (summa cum laude) in electronic engineering from the University of Trieste, Italy. He is currently the CEO and the Founder of Amped Software. His thesis on forensic image processing and founded Amped Software, in 2008. He has extensive software engineering experience designing and starting the development of Amped Software products. He taught as a Contract Professor in a few university courses related to investigations, forensics, and intelligence, has published several scientific articles related to image and video forensics and has also worked as a forensic expert in major judiciary cases.

ALBERTO LIMONE received the master's degree in mathematics from the University of Florence, Italy, in 2022. The topic of his dissertation, which was conducted in collaboration with Amped Software, was the basis of this article. He is currently a Mathematics and Physics Teacher at a scientific high school.

SIMONE REBEGOLDI received the Ph.D. degree in mathematics from the University of Ferrara, Italy, in 2017. His Ph.D. dissertation titled "Variable metric line-search based methods for nonconvex optimization." Since 2023, he has been an Assistant Professor in numerical analysis with the Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia, Italy. His research interests include the design and analysis of continuous optimization algorithms for large-scale applications arising in image restoration and machine learning. Several of his works employ concepts and tools coming from variational analysis, convex analysis, and probability theory.

Open Access funding provided by 'Università degli Studi di Firenze' within the CRUI CARE Agreement