Shape from Shading for the Digitization of Curved Documents

Frédéric Courteille, Alain Crouzil, Jean-Denis Durou, Pierre Gurdjos

Shape from shading for the digitization of curved documents

Frédéric Courteille · Alain Crouzil · Jean-Denis Durou · Pierre Gurdjos

Abstract Document digitization is faster and more affordable using digital cameras than scanners. On the other hand, if we aim at extending the basic digital camera functionalities for such a purpose, post-processings will be of first importance, at least to improve the text legibility. In this paper, we address the specific prob-lem of the virtual flattening of curved documents, as for example the pages of an opened book lying on its spine. In order to compute the document shape, we use the shape from shading technique and discuss why, in some cases, it is more suitable than other 3D single-view reconstruction techniques. We extend the seminal work by Wada et al. (Proceedings of the IAPR Workshop on machine vision and applications, Tokyo, Japan, pp. 591–594, 1992) and consecutive papers, reformulating the problem in terms of perspective shape from shad-ing. Finally, we design a complete post-processing algorithm and test it on real images. Even if the documents are much curved, it is shown that the restored images are almost identical to scanned images of the flattened documents.

1 Introduction

Document digitization currently encounters increasing popularity, because of the development of Internet browsing. Our work falls under the line of many public and private research programs related to the creation of digital libraries based on images: DLI[1] (Digital Library Initiative) in the USA, HUMI[2] (HUmanities Media Interface) in Japan, "Mémoire du Monde"[3] at UNESCO, Gutenberg Digital[4] in Germany, etc. The traditional process, which consists in using a flatbed scanner, is satisfactory for common situations. Nevertheless, it has two weaknesses when dealing with a book: repetitively turning the pages and forcing the book sticking to the scanner glass can become rather tedious; in addition, in the case of a thick book, some defects will appear in the digitized image (blurred or deformed characters, effects of parallax, non-inked area having a non-uniform graylevel, see [2]). In order to avoid these two weaknesses, several specific systems have been designed, but of course such systems are not consumer equipments. An alternative consists in replacing the scanner, whose optical system is difficult to model, by a digital camera and in virtually flattening the document.

Our work takes place within this framework. In order to compute a geometrically and photometrically corrected image of a curved document, it is helpful to compute the shape of the document from its initial image, then use this knowledge to simulate the image of the flattened document. Several traditional techniques of 3D reconstruction can be used. Within the framework of this work, we use the well-known technique of shape from shading (SFS), which does not require sophisticated material and can be applied to most of the documents. We show in this paper that the only use of SFS makes it possible to obtain very satisfactory results on

F. Courteille (✉) · A. Crouzil · J.-D. Durou · P. Gurdjos Institut de Recherche en Informatique de Toulouse,
Université Paul Sabatier, 118 route de Narbonne,
31062 Toulouse Cedex 9, France
e-mail: courteille@irit.fr

[1] http://www.dli2.nsf.gov.

[2] http://www.humi.keio.ac.jp.

[3] http://www.unesco.org/webworld/mdm/fr/index_mdm.html.

[4] http://www.gutenbergdigital.de.

real images. The craze for SFS from 1980 until 1995 has declined between 1996 and 2000, probably because the results obtained on real images were disappointing, as shown in two recent surveys [3,4]. Since 2001, a new interest occurred since three research groups simultaneously proposed a more realistic modelling of SFS [5–7] and practical applications could therefore be considered such as human face reconstruction [8], organ reconstruction using medical images [9] or virtual document flattening [7], which is in the heart of our contribution. The graylevel of the photograph of an open book can be simply interpreted in terms of surface orientation, without ambiguity. In particular, the traditional "concave/convex ambiguity" of SFS, which causes optical illusions, is generally avoided, since each page is convex. Consequently, if the relation between graylevel and surface orientation is reliable, then SFS can be used as an accurate 3D reconstruction method.

In Sect. 2, we review different systems of digitization. In Sect. 3, we review the computer vision techniques which can help enhancing the images of digitized documents, and we show why the use of SFS is suitable. In Sect. 4, we throw the traditional modelling of SFS back into question. In Sect. 5, we detail a full processing chain allowing the digitization of curved documents. Experimental results are presented in Sect. 6. Finally, Sect. 7 concludes our study and states several perspectives.

2 Systems of digitization

Even if strong similarities exist between their optical sensors, it must be noted that scanners and cameras do not aim at digitizing the same kinds of scenes.

2.1 Scanners

Classical scanners are specifically designed to digitize flat documents. Their 1D sensors sweep over all the document surface in order to produce the digital image, which makes them rather slow. However, for a perfectly flat document, the image is a perfect replica of the original, i.e., there are neither geometric defect (the original and its digital copy are perfectly superposable, up to a scale factor), blur (stigmatism is perfectly verified), nor photometric defect (the colours of the copy are faithful to those in the original). Thus, the quality of the copy is only limited by the resolution, which can even be more than 1,200 dpi for classical scanners. Note that the digitization of a book of 100 pages requires the repetition of 50 times the same process, which consists in turning a page and carefully forcing the book sticking to the scanner glass but, since the copies will be quasi-identical to the originals, the relative slowness of the process is acceptable.

Apart from this common use of a scanner, i.e., if the document is no longer flat but curved, as for example a thick book, then three types of defects simultaneously appear, particularly in the neighbourhood of the binding (cf. image I_0^1 in Fig. 1a): geometric defect (the lines of text are no longer parallel and some characters are deformed), blur and shading (the non-inked area does not have a uniform graylevel). We will return further in detail to the a posteriori corrections of these defects. Before that, let us mention several systems of digitization which have been specifically designed to correct them a priori and the principle of which is to force the document sticking to the scanner glass.

A first solution consists in reducing the sensor size and running it over the document as an iron. This explains why such systems are named "hand scanners". This process removes indeed the above-mentioned defects. Nevertheless, it could become tiresome when digitizing a book in its whole. Moreover, it would not be tractable in the case of an old document, because of possible damage.

Other systems, called "book scanners", are specifically dedicated to book digitization purposes: the opened book lies on its spine, the pages of which not necessarily forming a 180° angle in order to avoid damaging the binding; a glass moderately presses each page, in order to flatten it. Since the sensor has to be raised after each digitization, this process is rather slow. Nevertheless, the results are of very good quality. The "Parc Book Scanner"[5] system, based on this principle, was used within the framework of the already mentioned DLI program. A number of companies sell such scanners, as for example I2S.[6] These specific systems resemble more digital cameras than scanners since, most of the time, their sensors are CCD matrices. A main problem is their very high cost. Therefore, we will rather consider images captured using common sensors and the way of correcting their defects.

2.2 Digital cameras

Digital cameras contain CCD or CMOS matrix sensors. How non-specific sensors can be worthwhile for document digitization when specific sensors (scanners) exist? In the case of flat documents, the speed of digitization is highly improved using a camera but, because of optical aberrations, the reliability of the copy is less than using

[5] http://www2.parc.com.

[6] http://www.i2s-bookscanner.com.

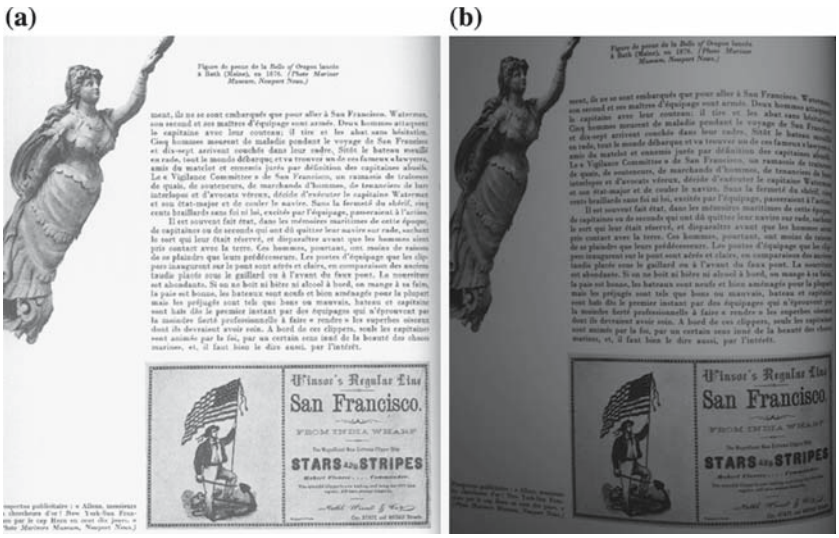

Fig. 1 Left page of a book: **a** scanned image I_0^1 and **b** photograph I_1^1. During the digitization: the document is quasi-plane in **a**; it is curved in **b**

a scanner. On the other hand, when dealing with curved documents, the digital image contains the same defects than using a scanner. In Figs. 1a, b, the scanned image I_0^1 and the digital photograph I_1^1 of the same document are displayed. The geometric defect in image I_1^1 is clearly higher than in image I_0^1, because the book lies on its spine in I_1^1, whereas in I_0^1 it is forced sticking to the scanner glass. As a consequence, the book slopes are clearly stronger in I_1^1 than in I_0^1. Despite that, image I_1^1 is sharp, because decreasing the objective aperture increases the depth of field. On the other hand, the shading defect is much more significant in I_1^1 than in I_0^1. Let us notice that this last defect could be avoided by illuminating the page with diffuse light, instead of using the camera flash. But, as for scanned images, digital photographs of curved documents can be a posteriori corrected and we will show hereafter how the shading "defect" is very useful.

3 Image enhancement techniques

Our aim is the following: design a digitization system of curved documents which does not use specific material but yet reaches the same performances as book scanners which really flatten the documents. The only way of processing consists in digitizing the documents using a classical scanner or a camera, then correcting their defects. These two stages being completely independent, it is possible to digitize a whole book, and to carry out the image enhancement afterwards, when computer resources are available.

3.1 2D–2D transformations

It is possible to carry out a purely 2D image processing technique, which consists in stretching it using interpolation, according to an a priori layout model of the flat document. In [10,11], several nonlinear transformations aim at correcting the geometric distortions of the images. Unfortunately, no test on real images containing text is carried out. In [12–14], the orientation of the characters is evaluated, in order to straighten them out, but here again the results are disappointing because, if it is rather easy to force the text lines to be parallel, it is much more difficult to correctly stretch the characters near the binding. In [15], a judicious 2D deformation makes the page rectangular. The results are good, but a "paper checkerboard pattern" has to be slipped behind the document, which requires some user interaction.

As said in [16], "Image processing algorithms cannot, without a priori information, determine document shape, which is necessary to solve the restoration problem". Instead of 2D–2D transformations, all other contributions consider indeed 2D–3D–2D transformations: first, the 3D reconstruction is carried out using classical computer vision techniques (2D–3D); second, the document flattening is simulated (3D–2D). Let us review these two steps.

3.2 3D reconstruction

Approaches using classical computer vision techniques of 3D reconstruction have been reported. They fall into two categories:

- Geometric methods: use of two photographs taken under two different viewing angles, projection of structured light, study of the deformation of the text lines, study of the contour deformation.
- Photometric methods: use of two photographs taken under two different lights, analysis of the non-inked area graylevel.

3.2.1 Use of two photographs taken under two different viewing angles

Stereoscopy is a well-known technique of computer vision, which consists in reconstructing the shape of the scene using two (or more) photographs taken under different viewing angles. The tricky step of this technique is the matching process, which consists in seeking, for each pixel of one image, the matching pixel in the other image, which corresponds to the same point in the scene. More and more reliable measures of correlation are proposed [17], which make it possible to obtain very good estimates of the shape, but the required computing time also considerably increases. In [18], not less than 6 h are required for a $2,048 \times 1,360$ image. In addition, this technique requires an accurate estimate of the two camera poses.

3.2.2 Projection of structured light

This technique consists in taking two photographs using two different lights, given that, for one of the photographs, a structured light containing a known pattern is projected onto the document [19,20]. Analysing the deformation of this pattern allows to reconstruct the shape. A second photograph taken under homogeneous light is required: otherwise, the simulated image would contain the photometric artefacts due to the pattern. In order to properly analyse the deformation of the pattern, it is necessary to precisely locate it in the image. In [19], this task, which is carried out using wavelet transform, seems to be the main difficulty of the method. Let us also mention the Minolta 3D1500 camera, which projects coloured light bands onto the scene and automatically computes its surface shape. In [21], this camera is used for the digitization of damaged manuscripts, within the framework of the "Digital Atheneum" project.[7] Finally, the results obtained in [22] are quite impressive, but the price to pay is a rather sophisticated system of acquisition.

3.2.3 Study of the deformation of the text lines

The text lines of a document may be a priori supposed parallel and equidistant for most of the documents. Of course, documents without text or with non-horizontal lines of text, like the calligram displayed in Fig. 2, render this technique unusable. It was recently exploited in [23] and provides results of suitable quality in a few seconds. The crucial stage consists in extracting the lines of text, then in computing the shape using shape from texture techniques. In [24], this technique is extended to much more general situations than in [23], using the so-called "texture flow-field analysis".

3.2.4 Study of the contour deformation

It is also possible to use the contours of a document to compute its shape, if it is supposed to be rectangular (an old document could violate such an assumption). This technique was tried out in [25] within the framework of the HUMI project which has already been cited. The proposed method uses the bottom contour *or* the top contour of the document only. The same technique was also exploited in [15], but the bottom contour *and* the top contour of the document are required. In [26], we extended this technique to the case of uncalibrated cameras. Let us also quote a recent work [27], in which the document surface is only supposed to be applicable, and not necessarily cylindrically warped. Using differential geometry leads to an elegant reformulation of the problem as partial differential equations, but unfortunately the results appear to be a little bit disappointing.

3.2.5 Use of two photographs taken under two different lights

Photometric stereo consists in taking two (or more) photographs under the same viewing angle, but under different lights. A photometric model linking the graylevel to the slope of the surface is needed. This technique has been used in [28]. The results are of average quality, probably because the model that is used is not realistic, since as the photographs are taken at short distance, it is essential to take perspective into account, as will be seen further.

3.2.6 Analysis of the non-inked area graylevel

The first paper dealing with the simulation of document flattening [1] uses this technique, named SFS. It was applied to images taken with a traditional scanner. Wada et al. thought up of using the shading defect of the image, in order to compute the shape of the document. Indeed, this defect is linked to the slope of the surface at each non-inked point, and thus allows to correct the geometric, as well as the shading defects (cf. also [29]). This idea has recently been taken up and improved in [30], with rather convincing results.

Our contribution takes up the same idea and improves the results, compared to existing methods. Like in [1,29,30], we suppose that the document shape is a generalized cylinder. But, contrary to [1,29,30], we deliberately use a digital camera instead of a scanner, because the optical

[7] http://www.digitalatheneum.org.

system of a camera is easier to model than that of a scanner; this choice makes it possible to digitize more quickly and without touching the document. Moreover, we propose a new modelling of SFS which takes perspective projection into account. Finally, we use the most powerful method of interpolation among several existing ones. A less accomplished version of our method was already published in [7].

The simultaneous use of several techniques of 3D reconstruction would undoubtedly improve the results, but we show in this paper that the only use of SFS makes it possible to obtain very satisfactory results on real images. Moreover, this technique makes it possible to deal with images without any visible contour. Finally, it can also work for documents containing no text, or containing no horizontal line, as that displayed in Fig. 2. The digitization of such a document is particularly difficult when it is curved, since the lines of text are not rectilinear and different kinds of fonts are mixed: some of the techniques cited above would be unusable.

3.3 Virtual flattening

Once its shape has been computed, the document has to be virtually flattened, which consists in fact in two steps: unwarping and interpolation. As soon as the document surface is a generalized cylinder, it is easy to unwarp it. We use the same unwarping method as that described in much of the papers dealing with cylindrical document digitization (see for example [18,25]). Moreover, let us mention two techniques [31,32] which are well adapted to the case of non-cylindrical documents, and the applicable surface representation used to [27], which is adapted to unwarping in a natural way. Finally, a more general technique has been proposed in [16], which can also unwarp non-applicable surfaces like old or damaged documents.

The second step i.e., interpolation, is more difficult. Contrary to certain authors [25,28] who developed original techniques of interpolation, we address classical existing methods, as will be described further.

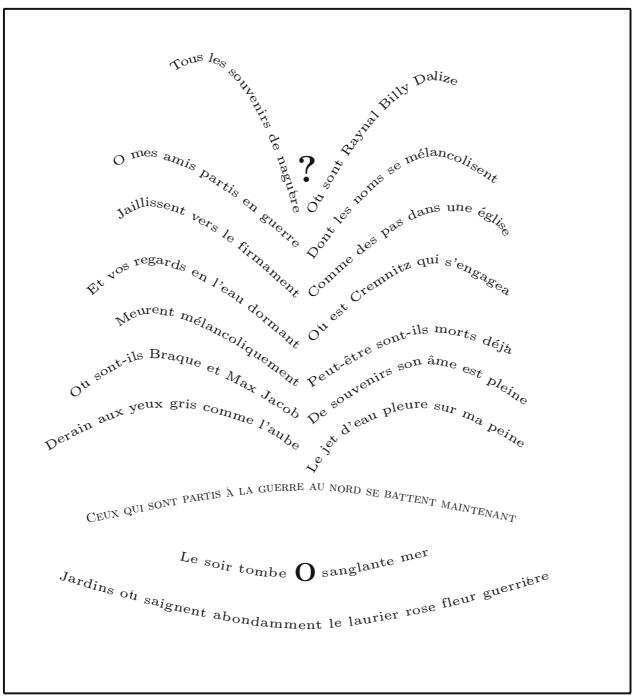

Fig. 2 Guillaume Apollinaire's calligram: *La colombe poignardée et le jet d'eau*

4 Shape from shading modelling under realistic photographic conditions

The SFS problem consists in recovering the shape of a scene from a single image, by means of the graylevel analysis.

4.1 General description of a photographic Bench

In Fig. 3, the three fundamental entities which occur when a photograph is taken are drawn: the scene, the light and the camera, made up of the objective and the sensor. We choose a 3D coordinate system $Cxyz$ related to the camera, whose origin is the optical centre C and such that axis Cz coincides with the optical axis. It is well known that each plane orthogonal to the optical axis, of equation $z = -d$, is conjugated with a plane orthogonal to the optical axis, of equation $z = d'$, such that distances d and d' verify Descartes' relation:

$$\frac{1}{d'} + \frac{1}{d} = \frac{1}{f'}, \tag{1}$$

where f' is the image focal length[8] of the objective (which is positive). The sensor of the camera is located in a plane Π' of equation $z = d'$ and is conjugated with a plane Π of equation $z = d$, called the "focus plane", which contains the object points that have a perfectly sharp image. Eventually, we attach to the image plane Π' the 2D coordinate system Oxy, where O is the principal point, i.e., the intersection of the optical axis with plane Π'.

[8] In accordance to optics terminology, by "focal length", we refer to the distance between the optical centre and the focus. This is *not* the usual definition that can be found in the computer vision literature.

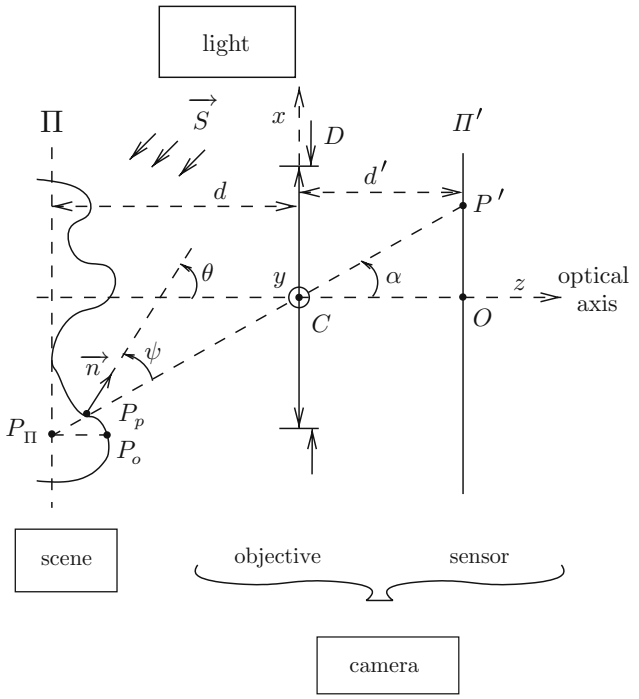

Fig. 3 Relation between scene and image

4.2 Image irradiance equation

SFS being rather complex, it is used to do the following assumptions:

- A_1: The scene reflectance, which finely describes the light re-emission, is known and uniform.
- A_2: The scene material is Lambertian.
- A_3: The surface contains neither edge nor hidden part, and therefore the unit outgoing normal is well defined at any point.
- A_4: The scene is illuminated by a parallel and uniform light beam which is described by vector $\overrightarrow{S}$.
- A_5: The secondary light reflections are negligible.
- A_6: The image is sharp and the possible aberrations of the objective are negligible.
- A_7: The sensor response is linear.
- A_8: Angle α (cf. Fig. 3) is such that $\alpha \ll 1$.

Under assumptions A_1, A_5 and A_6, SFS is modelled through an equation known as "image irradiance equation" [33]:

$$t\,\frac{\pi}{4}\,\cos^4\alpha\,\frac{D^2}{d'^2}\,L(P,\overrightarrow{PC}) = E(P'), \tag{2}$$

where t designates the transmission ratio of the objective, D the diameter of its entrance pupil, $L(P,\overrightarrow{PC})$ the luminance of the scene surface at object point P, in the direction of optical centre C, and $E(P')$ the irradiance at point P' conjugated with P. Assumption A_7 says that there exists a coefficient r such that graylevel $I(P')$ and irradiance $E(P')$ at a point P' are linearly linked:

$$I(P') = r\,E(P'). \tag{3}$$

Assumption A_8 allows us to neglect factor $\cos^4\alpha$ in (2). Assumptions A_2, A_3 and A_4 lead to a particularly simple analytical luminance model, which is independent of the direction of re-emission:[9]

$$L(P,\overrightarrow{PC}) = -\frac{\rho}{\pi}\,\overrightarrow{S}\cdot\overrightarrow{n}\,(P), \tag{4}$$

where ρ designates the albedo (supposed to be uniform, according to assumption A_1) of the surface and $\overrightarrow{n}\,(P)$ the unit outgoing normal at point P. Using all these assumptions, Eq. (2) can be rewritten:

$$-r\,t\,\frac{\rho}{4}\,\frac{D^2}{d'^2}\,\overrightarrow{S}\cdot\overrightarrow{n}\,(P) = I(P'). \tag{5}$$

This equation can be solved only if the geometric relation between P and P' is known, i.e., if the model of projection is known. We will first consider orthogonal projection, and then perspective projection. Before that, let us make a ninth assumption, which is very often used and much simplifies the notations:

- A_9: The light is frontal, $\overrightarrow{S} = (0,0,-S)$.

Nevertheless, the equations that will be obtained thereafter could easily be generalized to any light direction.

4.3 Eikonal equation

Under the assumption of orthogonal projection, image point P' is conjugated with object point P_o, whose orthogonal projection P_Π on Π is conjugated with image point P', through the central projection of centre C on Π' (cf. Fig. 3). The datum of the problem is function i such that $i(x,y) = I(P')$, where P' has (x,y) as coordinates. The unknown of the problem is function u_o such that $u_o(x,y)$ is the height at point P_o. As the coordinates of P_o are $(x/g, y/g, u_o(x,y))$, where $g = -d'/d$ is the transverse magnification, $\overrightarrow{n}\,(P_o)$ can easily be expressed according to the coordinates $(\partial_x u_o(x,y), \partial_y u_o(x,y))$ of $\nabla u_o(x,y)$:

$$\overrightarrow{n}\,(P_o) = \frac{(-g\,\partial_x u_o(x,y), -g\,\partial_y u_o(x,y), 1)}{\sqrt{g^2\,\|\nabla u_o(x,y)\|^2 + 1}}. \tag{6}$$

[9] The exact writing of (4) is $L(P,\overrightarrow{PC}) = \max\{-\frac{\rho}{\pi}\,\overrightarrow{S}\cdot\overrightarrow{n}\,(P), 0\}$, because the graylevel of a non-illuminated point is equal to zero.

Using this expression of $\overrightarrow{n}(P_o)$ and assumption A_9, Eq. (5) can be rewritten:

$$g^2 \, \|\nabla u_o(x,y)\|^2 = \left[\frac{r\,t\,\rho\,D^2\,S}{4\,d'^2\,i(x,y)}\right]^2 - 1. \tag{7}$$

This equation, known as "eikonal equation", is the most frequently found in the SFS literature. It is a non-linear partial derivative equation of first order (Hamilton–Jacobi's equation of first order). Its first member is positive or null and vanishes for object points where $\overrightarrow{n}(P_o)$ is parallel to Cz, which are called "singular points". The graylevel at these points reaches its maximal value i_{max}:

$$i_{max} = \frac{r\,t\,\rho\,D^2\,S}{4\,d'^2}. \tag{8}$$

Eikonal equation (7) can thus be rewritten:

$$g^2 \, \|\nabla u_o(x,y)\|^2 = \left[\frac{i_{max}}{i(x,y)}\right]^2 - 1. \tag{9}$$

Parameter i_{max} can be directly measured in the image, if it contains at least one singular point.

4.4 Perspective eikonal equation

Until recently, very few papers dealing with SFS [34–37] had considered perspective projection. Moreover, this preliminary work had modelled perspective SFS only in the case of scenes made up of flat facets. Recently, three groups of authors simultaneously (and independently of each other) established a more general modelling of perspective SFS [5–7]. The object point conjugated with image point P' is now object point P_p, whose image is P' through the central projection of centre C on Π' (cf. Fig. 3). Therefore, the coordinates of point P_p are $(x\,u_p(x,y)/d', y\,u_p(x,y)/d', u_p(x,y))$, where function u_p is the new unknown of the problem. A straightforward calculus gives the following expression of the coordinates of $\overrightarrow{n}(P_p)$:

$$\overrightarrow{n}(P_p) = \frac{(-\hat{g}(x,y)\,\partial_x u_p(x,y), -\hat{g}(x,y)\,\partial_y u_p(x,y), 1)}{\sqrt{\hat{g}(x,y)^2\,\|\nabla u_p(x,y)\|^2 + 1}}, \tag{10}$$

where $\hat{g}(x,y) = d'/\widehat{u_p}(x,y)$ and:

$$\widehat{u_p}(x,y) = u_p(x,y) + x\,\partial_x u_p(x,y) + y\,\partial_y u_p(x,y). \tag{11}$$

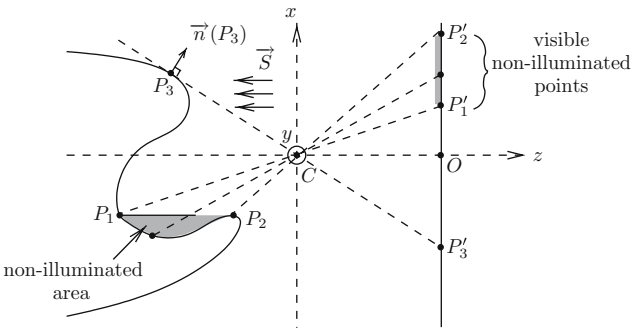

Fig. 4 Existence of visible non-illuminated points under perspective projection and assumptions A_4, A_5 and A_9

Thus, Eq. (5) can be rewritten:

$$\hat{g}(x,y)^2 \, \|\nabla u_p(x,y)\|^2 = \left[\frac{i_{max}}{i(x,y)}\right]^2 - 1. \tag{12}$$

This equation, that we call "perspective eikonal equation", is very similar to (9). Once again, it is a non-linear Hamilton–Jacobi's equation of first order. Actually, Eq. (12) is invalid for visible non-illuminated points, a situation that could not occur under orthogonal projection, but that could occur under perspective projection and assumptions A_4, A_5 and A_9, as illustrated in Fig. 4.

Apparently, the only difference between the two eikonal equations (9) and (12) is that constant g is replaced by $\hat{g}(x,y)$, but this has several notable consequences. Whereas unknown u_o appears in (9) through its gradient only, this is not the case for unknown u_p in (12), which appears explicitly (but indirectly) in expression (11) of $\widehat{u_p}(x,y)$. Knowing a solution u_o of (9), one can thus find an infinite number of solutions $u_o + K$, K being any real number, which is not the case for Eq. (12). On the other hand, it is worthy of mention that Eq. (12) is homogeneous, which means that knowing a solution u_p of (12), one can find an infinite number of solutions $K'\,u_p$, K' being any real number. In other words, the resolution of eikonal equation (9) can be done up to an additive constant only, whereas the resolution of perspective eikonal equation (12) can be done up to a multiplicative constant only.

There are two other differences, quite crucial, between Eqs. (9) and (12). On the one hand, variables x and y implicitly appear in (9), whereas they explicitly appear in (12). On the other hand, a change in d' does not have the same effect on the resolution of both these equations. For example, if d' is multiplied by 2, then so is g, and u_o is simply divided by 2, according to (9), which only rescales the reconstructed surface, since $P_o = (x/g, y/g, u_o(x,y))$. According to (12), the induced transformation of u_p is more complicated, i.e.,

the reconstructed surface is not simply rescaled, since $P_p = (x\,u_p(x,y)/d', y\,u_p(x,y)/d', u_p(x,y))$. The resolution of Eq. (12) thus requires the knowledge of the position of principal point O, as well as the knowledge of distance d'. All these values can be estimated through geometric camera calibration. But it must be kept in mind that, if the calibration is not accurate, the resolution of (12) will not be correctly carried out.

One can fear that solving (12) will be more difficult than solving (9). In [5], (12) is solved through the search for its viscosity solutions. A numerical scheme approximating these solutions, as well as the proof of convergence of the scheme, are provided, which are generalizations of results in [38]. We consider the resolution of Eqs. (9) and (12) from a different point of view.

4.5 Pseudo-eikonal equation

As many authors [39], we add normal $\overrightarrow{n}(P)$ as a new unknown. This unit vector having two degrees of freedom, it is like introducing two scalar unknown functions, for example p and q such that:

$$\overrightarrow{n}(P) = \frac{(-p(x,y), -q(x,y), 1)}{\sqrt{p(x,y)^2 + q(x,y)^2 + 1}}. \tag{13}$$

Let us notice that functions p and q are defined without ambiguity, according to assumption A_3. Equation (5) can then be rewritten:

$$p(x,y)^2 + q(x,y)^2 = \left[\frac{i_{\max}}{i(x,y)}\right]^2 - 1. \tag{14}$$

Contrary to (9) and (12), this equation is not a partial derivative equation. For this reason, we propose to call it "pseudo-eikonal equation". Let us stress that it is valid for orthogonal, as well as for perspective projection. Of course, its resolution is not a well-posed problem, since at each pixel (x,y), it consists in one equation for two unknowns $p(x,y)$ and $q(x,y)$, but several strategies[10] can be considered to render the problem well posed [40].

Once p and q are known, the scene shape has to be computed via integration. Under the assumption of orthogonal projection, it comes from (6) and (13):

$$\nabla u_o(x,y) = \left(\frac{p(x,y)}{g}, \frac{q(x,y)}{g}\right). \tag{15}$$

[10] In the framework of our application, we will see that a priori knowledge on the scene surface enables us to compute p and q without ambiguity.

Under the assumption of perspective projection, it comes from (10) and (13), following a similar reasoning:

$$\nabla u_p(x,y) = \left(\frac{p(x,y)}{\hat{g}(x,y)}, \frac{q(x,y)}{\hat{g}(x,y)}\right). \tag{16}$$

This last equality is equivalent to the following linear system [in order to simplify the notations, the dependences in (x,y) have been omitted]:

$$\begin{cases} \left[x\,p - d'\right]\,\partial_x u_p + y\,p\,\partial_y u_p = -p\,u_p, \\ x\,q\,\partial_x u_p + \left[y\,q - d'\right]\,\partial_y u_p = -q\,u_p. \end{cases} \tag{17}$$

The determinant δ of this linear system of two equations with unknowns $\partial_x u_p$ and $\partial_y u_p$ is equal to

$$\delta = d'(d' - x\,p - y\,q). \tag{18}$$

Using the notations of Fig. 3, the scalar product between $\overrightarrow{n}(P_p)$ and $\overrightarrow{CP'}$ is equal to

$$\overrightarrow{n}(P_p) \cdot \overrightarrow{CP'} = (d' - x\,p - y\,q)\cos\theta, \tag{19}$$

using (13) and noticing that the third coordinate of $\overrightarrow{n}(P_p)$ equals $\cos\theta$. On the other hand,

$$\overrightarrow{n}(P_p) \cdot \overrightarrow{CP'} = \frac{d'\,\cos\psi}{\cos\alpha}, \tag{20}$$

using the fact that $\|\overrightarrow{CP'}\| = d'/\cos\alpha$ and introducing angle $\psi = (\overrightarrow{CP'}, \overrightarrow{n}(P_p))$. Consequently,

$$\delta = \frac{d'^2\,\cos\psi}{\cos\alpha\,\cos\theta}, \tag{21}$$

which vanishes only if $\psi = \pi/2$. If not, i.e., if P does not lie on a "limb" (for example, point P_3 of Fig. 4 lies on a limb), then the solution of (17) can easily be computed:

$$\nabla u_p(x,y) = \frac{u_p(x,y)}{d' - x\,p(x,y) - y\,q(x,y)}\,(p(x,y), q(x,y)). \tag{22}$$

Equation (15) is a system of two partial derivative equations in u_o, as Eq. (22) is a system of two partial derivative equations in u_p. But, contrary to eikonal equations (9) and (12), these two equations are linear and thus will be much simpler to solve. The same differences can be stated between (15) and (22) as between (9) and (12), that is to say: (15) can be solved up to an additive constant only, whereas (22) can be solved up to a multiplicative constant only; the resolution of (22) requires the knowledge of the principal point coordinates, because variables x and y explicitly appear in the

Fig. 5 Photographic Bench

denominator of its right member, which is not the case of (15); the resolution of (22) also requires the knowledge of distance d', contrary to that of (15), for which a change in d' only rescales the reconstructed surface.

5 Application to the digitization of curved documents

Once a more realistic modelling of SFS has been stated, we can address the application which we are concerned with, i.e., the simulation of curved document flattening.

5.1 Description of the photographic Bench

Figure 5 shows the photographic Bench that is used in the experiments. The document lies on a table that is beforehand covered with black cloth. This precaution aims at avoiding possible secondary reflections on the book, which could considerably degrade the quality of the 3D reconstruction using SFS, and consequently that of the flattening. A digital reflex camera (Canon EOS 300D, CMOS sensor of 6.5 Mp) is used, which is fixed under a photographic stand, approximately 400 mm away from the table. Let us remind the reader that the use of a camera makes it possible to speed up the process, so that about 40 pages can be digitized per minute.

As previously said, the use of Eq. (22) requires that the internal parameters of the camera have been estimated (distance d' and the position of principal point O). To do that, we use Bouguet's calibration package.[11] In addition, the scene is lighted using the camera flash. This photographic bench is thus particularly simple. The digitization is carried out page after page, but black cloth has to be laid on the opposite page. In accordance with assumption A_5, this avoids possible secondary reflections near the binding. It might be taken such reflections into account. Nevertheless, the method described in [41] hugely increases the CPU time, which leads us to dodge this problem for the moment.

In order to satisfy assumption A_6, we choose the objective aperture as small as possible, which avoids blur. Radial and tangential distortions are estimated through Bouguet's calibration and then corrected. On the other hand, assumptions A_2, A_4, A_7 and A_8 are not realistic. A way to overcome this problem (and also to take the vignetting effect into account) consists in carrying out a simple photometric correction.

5.2 Photometric correction

Figure 6 describes the full processing chain that we use. The datum is image I_1 (taken with the photographic bench described before). Our aim is to compute an image I_6 in which flattening is simulated. To do that, the first stage consists in applying an elementary photometric correction to image I_1, which is carried out taking a photograph I_{ref} of a white flat page (larger than the open book), lying on the table, then dividing I_1 by I_{ref} (after correction of radial and tangential distortions), which provides a new image I_2. This photometric correction is of primary importance (see [42] for more details) since, as can easily be noted, the graylevel of image I_{ref} is not uniform, as it should if assumptions A_2, A_4, A_7 and A_8 were perfectly satisfied.

In addition, the SFS equations are based on the implicit assumption of a uniform albedo (assumption A_1). However, the albedo of a printed document is obviously non-uniform: the albedo ρ of the non-inked area is close to 1, whereas the inked area has albedos which can take all values in $[0, \rho[$. The only way to ensure A_1 is to eliminate the inked area. Aiming at computing an image I_3 in which all the pixels corresponding to the inked area are masked, i.e., set to 0, we adaptively threshold image I_2 (in our experiments, each threshold is locally computed in a 5×5 window).

5.3 3D reconstruction

Now, we can compute the shape of the document using SFS. As already said, the computation of functions p and q is much simplified since the document shape is supposed to be cylindrical, i.e., two pixels located on a line parallel to the binding in the image are conjugated with two points in the scene having the same height. Choosing axis Cy parallel to the binding, this implies

[11] http://www.vision.caltech.edu/bouguetj.

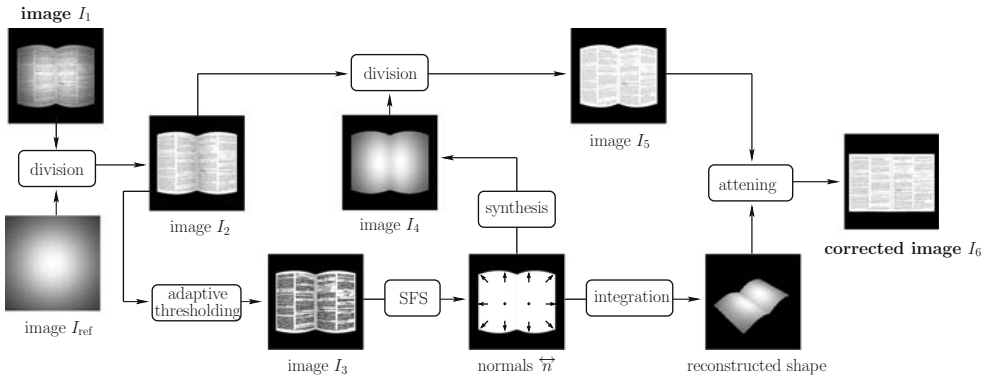

Fig. 6 Full processing chain

that q is uniformly zero and functions p and u_p depend on variable x only. Then, the problem is highly over-constrained and Eq. (14) makes it possible to estimate $p(x)$, for instance through the average $\bar{i}(x)$ of $i(x,y)$ in y, computed for the non-inked pixels of the column:

$$p(x)^2 = \left[\frac{\bar{i}_{\max}}{\bar{i}(x)}\right]^2 - 1, \tag{23}$$

where $\bar{i}_{\max}$ is the maximal value of $\bar{i}(x)$, reached at $x = x_{\text{sing}}$. Moreover, since each page is convex, p is positive for $x \leq x_{\text{sing}}$ and negative for $x \geq x_{\text{sing}}$. Consequently, we can compute u_p by integrating Eq. (22).

5.4 Shading removal

The knowledge of the document shape is useful twice: of course, it allows us to simulate document flattening, but it also makes it possible to remove the shading. First, we simply solve Eq. (14) for i, from known (p,q), assuming a uniform albedo ρ, which yields the "shading image" I_4 of the estimated shape (cf. Fig. 6). Second, using Eqs. (8) and (14), it derives:

$$i(x,y) = \frac{r\,t\,\rho\,D^2\,S}{4\,d'^2\sqrt{p(x,y)^2 + q(x,y)^2 + 1}}. \tag{24}$$

As the shapes corresponding to images I_2 and I_4 are the same, the non-uniform albedo $\rho_2(x,y)$ in image I_2 can be computed by dividing the two equations (24) corresponding to $i_2(x,y)$ and $i_4(x,y)$, which gives:

$$\rho_2(x,y) = \rho\,\frac{i_2(x,y)}{i_4(x,y)}. \tag{25}$$

In other words, we obtain the "albedo image" I_5, in which the shading has been removed, by dividing I_2 by I_4 (cf. Fig. 6). Each pixel (x,y) in I_5 is thus associated with an albedo $\rho_2(x,y)$ and with a 3D point in the scene of coordinates $(x\,u_p(x,y)/d', y\,u_p(x,y)/d', u_p(x,y))$. Now, the remaining task consists in simulating the image I_6 of the flattened document.

5.5 Virtual flattening

Since the documents are supposed to be cylindrical, we use the same unwarping method as that described in many other papers [18,25]. On the other hand, we tested three classical existing methods of interpolation. The first method is based on Smythe's algorithm [43]. It consists in applying two successive 1D interpolations to the image (along both the lines and columns). This algorithm is publicly available through a free software named `xmorph` which allows, amongst others, morphing and texture mapping. The second method consists in using the function `griddata` of `Matlab`, which carries out Delaunay's triangulation, and then an image interpolation technique which uses the barycentric coordinates of the points approximating the scene. The third method consists in using `OpenGL` (programming interface towards graphic hardware). Table 1 summarizes the advantages and disadvantages of these three methods. Visually, the results are very similar. Therefore, as it is the quickest and the most powerful, we choose `OpenGL`.

6 Experiments

All the results shown in this section have been obtained using the processing chain of Fig. 6.

6.1 Qualitative analysis of the results

Figure 7 shows the result I_6^1 computed using our method, when starting from image I_1^1. Let us recall that only the

Table 1 Comparison of three methods of interpolation

	`xmorph`	`griddata`	`OpenGL`
Advantages	Fast (about 10 s) Free software Used for a long time	Very accurate	Very fast (about 1 s) Free software Easy implementation Use of the graphic hardware
Disadvantages	Difficult to use	Slow (about 1 min) Commercial software Useless triangulation	

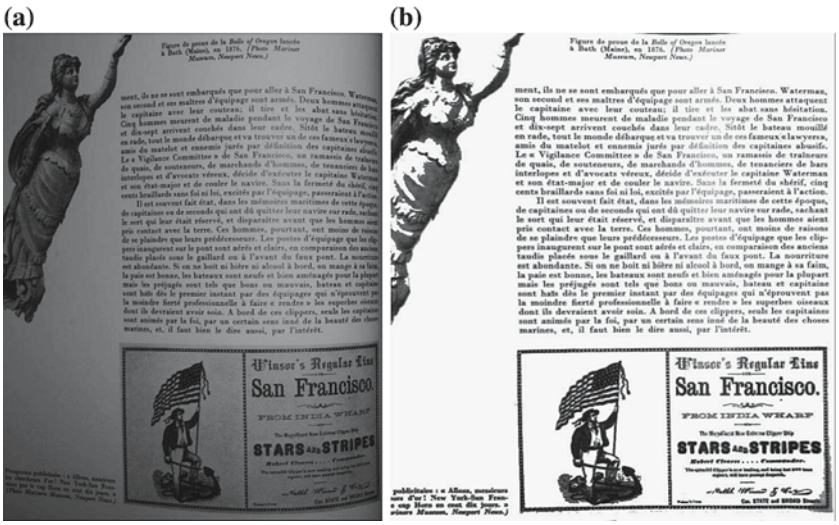

Fig. 7 Test 1: **a** photograph I_1^1 and **b** corrected image I_6^1

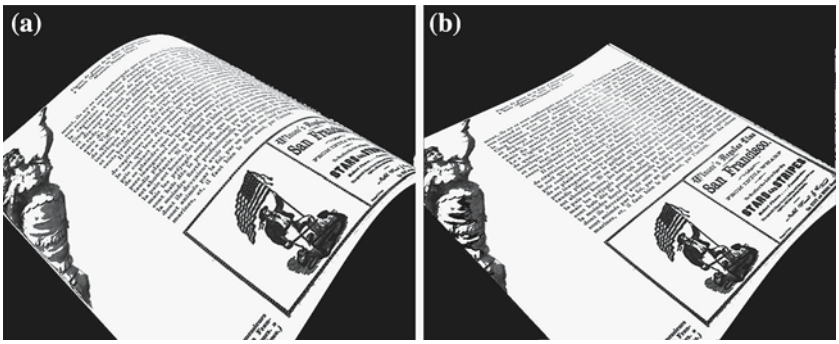

Fig. 8 3D view of the reconstructed page: **a** before and **b** after flattening

non-inked area of I_1^1 is used. Thus, the quasi-parallelism of the text lines in I_6^1 is not used as a priori knowledge. On the other hand, the frame located at the bottom of image I_6^1 has retrieved its rectangular form. Thus, our document flattening technique seems rather satisfactory. From a photometric point of view, it appears that some characters in the corrected image, particularly those near the area with weak slope, are slightly less contrasted than the others. This probably means that the SFS assumptions are not perfectly satisfied, and that our photometric correction (which is supposed to limit the departure from assumptions A_2, A_4, A_7 and A_8) is less efficient there. It would be interesting to use another light (as for example an "annular flash") or to refine the light model. As can be seen in Fig. 8, in which the 3D surface of the document is displayed before and after flattening, the estimated shape using SFS contains strong slopes, which underlines that the result of Fig. 7b is of good quality.

Figure 9 shows two sub-images extracted from images I_1^1 and I_6^1, which correspond to the same text area. To the naked eye, it seems that the correction is of very

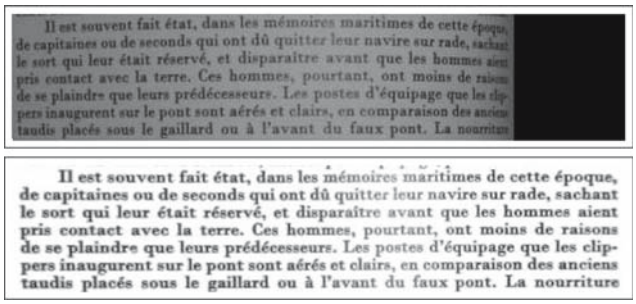

Fig. 9 Sub-images extracted from images I_1^1 and I_6^1

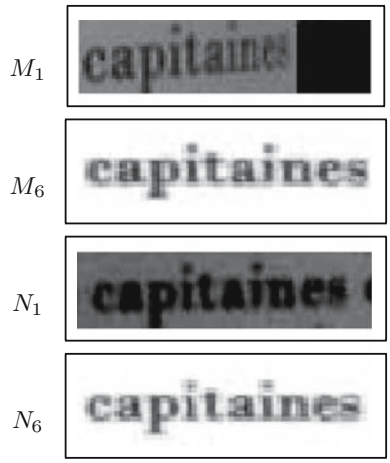

Fig. 10 Four occurrences of the word *capitaines*, extracted from images I_1^1 and I_6^1: M_1 becomes M_6 after flattening; N_1 becomes N_6 after flattening

good quality. Figure 10 shows four occurrences of the same word *capitaines* extracted from different places in I_1^1 and I_6^1: M_6 is the corrected version of M_1, which is near the binding (M_1 is much distorted since the surface slope is strong there); N_6 is the corrected version of N_1, which is in an area where the surface slope is rather weak. We can conclude that the correction is of good quality, since the letters in M_6 and N_6 are horizontal and have all the same size, and since two occurrences of the same word, which look very different in the original image I_1^1, are very similar after flattening.

Figure 11 shows another result. The photographed document is a ship mast plan. It appears in image I_6^2 that the straight lines have been correctly restored.

The last two results clearly show that our algorithm uses the non-inked areas in the document only, independently of its geometric structure or of possible figures of known geometry. In particular, the other techniques (deformation of the text lines, deformation of the contour) using one image only could not correctly work in these last two situations. The result of Fig. 12 represents three images of the calligram *La colombe poignardée et le jet d'eau*: (a) scanned image, (b) photograph and (c) corrected image. The already mentioned geometric and shading defects occur in the photograph. In particular, the word *Jardins* (first word on the last line) is much deformed and quite illegible, whereas the same word becomes perfectly legible after flattening, as regards its geometric appearance, as well as its contrast with the non-inked background. However, it appears that this result is not perfect, comparing the scanned image of the flat document with the corrected image obtained using our method: on the one hand, the text is slightly deformed (various occurrences of the same character do not have the same size); on the other hand, the non-uniform contrast which had already been noticed in the result of Fig. 7 is present again.

Figure 13 shows a last result, which contains only drawings. The bright area in the image in Fig. 13b suggests a non-Lambertian surface, which is in violation of assumption A_2. Yet, the end result in Fig. 13b shows that it works despite this violation (because this bright area is masked in image I_3^4). The two figurines located at the left and right edges of the image before flattening (cf. Fig. 13b) are much damaged, compared to the scanned image (cf. Fig. 13a). Nevertheless, the flattening gives a corrected image of very good quality: the three figurines are well restored, as regards their geometric appearance, as well as their colours (cf. Fig. 13c). This result is not perfect yet, since the contrast is not the same in the scanned image I_0^4 as in the corrected image I_6^4. Finally, let us indicate that the system described in [30] cannot handle colour documents.

6.2 Quantitative analysis of the results

In order to quantitatively evaluate the performances of our method, it is convenient to use an Optical Character Recognition (OCR) software. We applied the free OCR `gocr` to two images I_1^1 and I_6^1 of the same document, which contains 1,709 characters. The results are reported in Table 2. The success rate is multiplied by about 30 between I_1^1 and I_6^1 (61% instead of 2%). The failure rate is divided by about 3 and the abstention rate is divided by about 2. We can thus conclude that the geometric and photometric corrections which are carried out allow the OCR to decide more often and more reliably, even if the success rate is still far from what is commonly expected for digital text, which is, according to our experiments, a bit more than 90%.

6.3 Limits of the proposed method

We have just seen that, if the SFS assumptions are satisfied, the results of our method are of good quality.

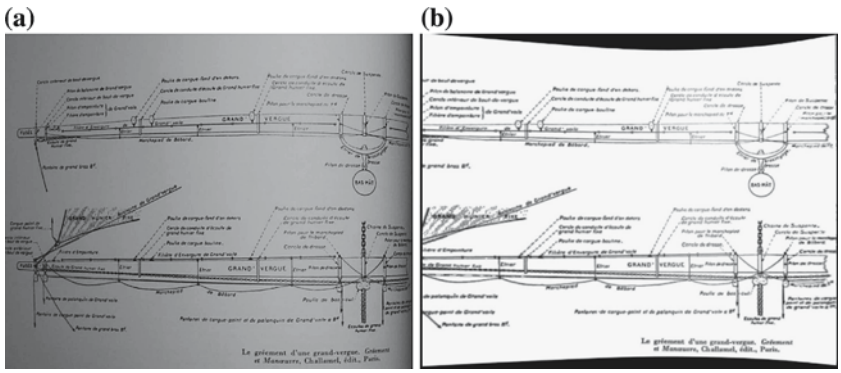

Fig. 11 Test 2: **a** photograph I_1^2 and **b** corrected image I_6^2

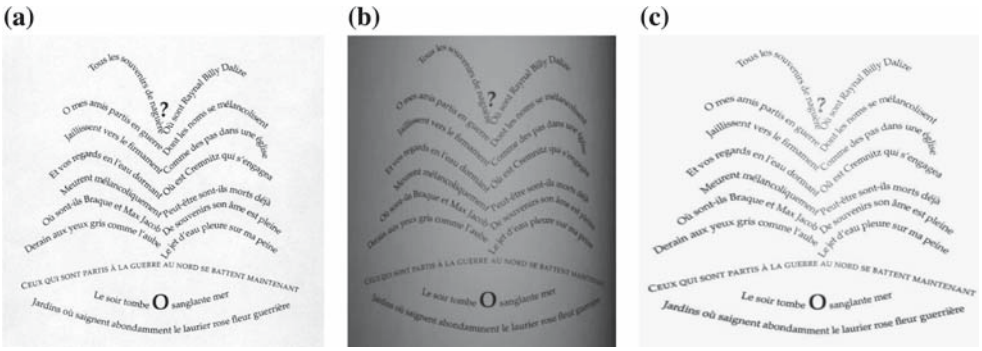

Fig. 12 Test 3: **a** scanned image I_0^3, **b** photograph I_1^3 and **c** corrected image I_6^3

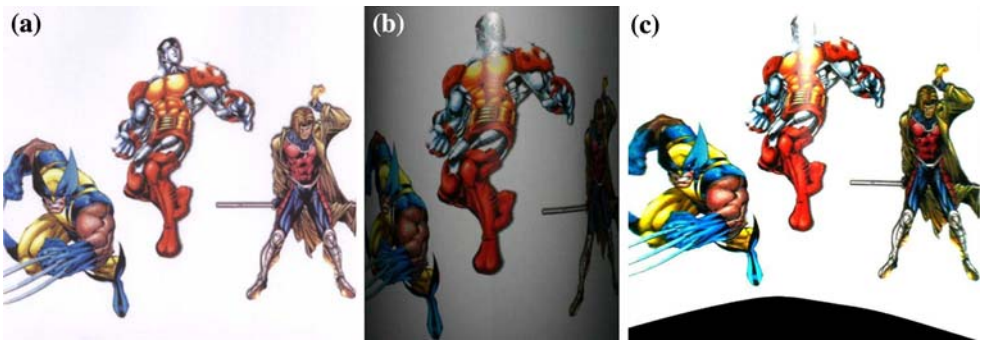

Fig. 13 Test 4: **a** scanned image I_0^4, **b** photograph I_1^4 and **c** corrected image I_6^4

However, it is no longer usable as soon as one of these assumptions is not satisfied, as is the case, for example, for the two photographs of Fig. 14. Indeed, the photograph of Fig. 14a was taken without control on the light (neither on its direction nor on its intensity). Thus, assumptions A_4 and A_9 are not satisfied. In addition, the document whose photograph is displayed in Fig. 14b is not Lambertian and its albedo, which is not uniform, is unknown. Thus, assumptions A_1 and A_2 are not satisfied.

7 Conclusion and perspectives

In this paper, we propose a full processing chain allowing the fast and cheap digitization of curved documents. The suggested method is original and consists in: using a digital camera for the acquisition of the image, which contain geometric and shading defects, if the document is non-flat; then, correcting these defects by simulating the document flattening. The results obtained on real images show that our method is potentially very effi-

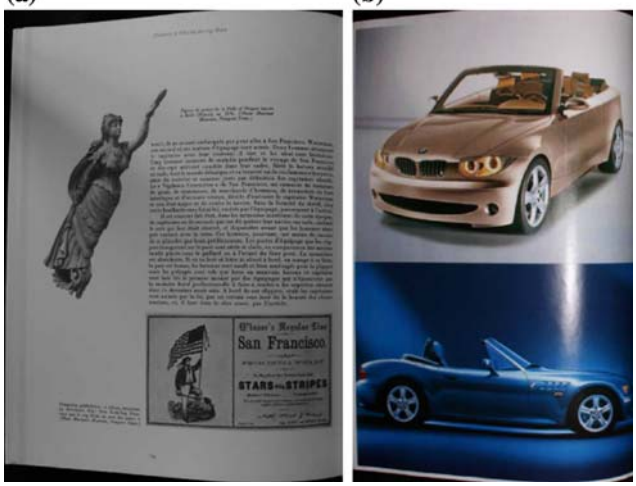

Fig. 14 Examples of images to which our method cannot be applied: **a** photograph for which assumptions A_4 and A_9 are not satisfied; **b** photograph for which assumptions A_1 and A_2 are not satisfied

Table 2 Results of the OCR on images I_1^1 and I_6^1

	I_1^1 (%)	I_6^1 (%)
Success rate	2	61
Failure rate	70	24
Abstention rate	28	15

cient. Even if the success rate of an OCR is still better when applied to a scanned image of the flat document, it is noteworthy that our method is fast and non-intrusive, compared to the classical process, which consists in forcing the document sticking to a scanner glass.

Many improvements, which were already evoked in the paper, must be considered, particularly:

- Our SFS resolution method would have to be generalized to documents of any shape, i.e., which are not necessarily generalized cylinders.
- The joint use of the contour of the document, which is supposed to be applicable to a rectangle, would better constrain the 3D reconstruction problem.
- The light system could be improved by replacing the camera flash with an annular flash.
- Our rather naive method of photometric correction allows to simultaneously correct the departure from assumptions A_2, A_4, A_7 and A_8, but could be still improved while more finely modelling, for instance, the scene reflectance.
- Taking secondary reflections into account would allow to simultaneously digitize the two pages of an opened book. As things are at present, our method correctly works only if one of the two pages is covered with black cloth.

At the end, we plan to add our method to camera functionalities, as is already the case for certain proccessings such as the creation of panoramic images.

Acknowledgements This work has partially been financed by EGIDE, within the framework of the franco-italian Galileo project named "PLATONOV".

References

1. Wada, T., Matsuyama, T.: Shape from shading on textured cylindrical surface restoring distorted scanner images of unfolded book surfaces. In: Proceedings of the IAPR Workshop on Machine Vision and Applications, pp. 591–594. Tokyo, Japan (1992)
2. Kanungo, T., Haralick, R.M., Phillips, I.: Global and local document degradation models. In: Proceedings of the 2nd International Conference on Document Analysis and Recognition, pp. 730–734. Tsukuba, Japan (1993)
3. Zhang, R., Tsai, P.S., Cryer, J.E., Shah, M.: Shape from Shading: a survey. IEEE Trans. Patt. Anal. Mach. Intell. **21**(8), 690–706 (1999)
4. Durou, J.D., Falcone, M., Sagona, M.: A survey of numerical methods for shape from shading. Rapport de Recherche 2004-2-R, Institut de Recherche en Informatique de Toulouse, Toulouse, France (2004)
5. Prados, E., Faugeras, O.: Perspective shape from shading and viscosity solutions. [44] 826–831
6. Tankus, A., Sochen, N., Yeshurun, Y.: A new perspective [on] shape-from-shading. [44] 862–869
7. Courteille, F., Crouzil, A., Durou, J.D., Gurdjos, P.: Towards shape from shading under realistic photographic conditions. In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 2, pp. 277–280. Cambridge,UK (2004)
8. Prados, E.: Application of the theory of the viscosity solutions to the shape from shading problem. Thèse de doctorat, Université de Nice—Sophia Antipolis, Nice, France (2004)
9. Tankus, A., Sochen, N., Yeshurun, Y.: Reconstruction of medical images by perspective shape-from-shading. In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 3, pp. 778–781 Cambridge, UK (2004)
10. Tang, Y.Y., Tang, C.Y.: Image transformation approach to nonlinear shape restoration. IEEE Trans. Syst. Man Cybern. **23**, 155–172 (1993)
11. Weng, Y., Zhu, Q.: Nonlinear shape restoration for document images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 568–573. San Francisco, California, USA (1996)
12. Zhang, Z., Tan, C.L.: Recovery of distorted document images from bound volumes. In: Proceedings of the 6th International Conference on Document Analysis and Recognition, pp. 429–433. Seattle, Washington, USA (2001)
13. Lavialle, O., Molines, X., Angella, F., Baylou, P.: Active contours network to straighten distorted text lines. In: Proceedings of the IEEE International Conference on Image Processing, vol. 3, pp. 748–751. Thessaloniki, Greece (2001)
14. Wu, C.H., Agam, G.: Document image de-warping for text/graphics recognition. In: Proceedings of the Joint IAPR International Workshops on Syntactical and Structural

Pattern Recognition and Statistical Pattern Recognition, pp. 348–357. Windsor, Canada (2002)
15. Tsoi, Y.C., Brown, M.S.: Geometric and shading correction for images of printed materials: a unified approach using boundary. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 240–246 Washington, D.C., USA (2004)
16. Brown, M.S., Seales, W.B.: Image restoration of arbitrarily warped documents. IEEE Trans. Patt. Anal. Mach. Intell. **26**, 1295–1306 (2004)
17. Chambon, S., Crouzil, A.: Dense matching using correlation: new measures that are robust near occlusions. In: Proceedings of the British Machine Vision Conference, Norwich, Royaume-Uni, pp. 143–152 (2003)
18. Yamashita, A., Kawarago, A., Kaneko, T., Miura, K.T.: Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system. In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 1, pp. 482–485. Cambridge, UK (2004)
19. Doncescu, A., Bouju, A., Quillet, V.: Former books digital processing: image warping. In: Proceedings of the IEEE Workshop on Document Image Analysis, pp. 5–9 San. Juan, Puerto Rico (1997)
20. Brown, M.S., Seales, W.B.: Document restoration using 3D shape: a general deskewing algorithm for arbitrarily warped documents. In: Proceedings of the 8th IEEE International Conference on Computer Vision, vol. 1, pp. 367–375, Vancouver, Canada (2001)
21. Brown, M.S., Seales, W.B.: Beyond 2D images: effective 3D imaging for library materials. In: Proceedings of the 5th ACM conference on digital libraries, pp. 27–36 San Antonio, TX, USA (2000)
22. Sun, M., Yang, R., Yun, L., Landon, G., Seales, W.B., Brown, M.S.: Geometric and photometric restoration of distorted documents. In: Proceedings of the 10th IEEE International Conference on Computer Vision, vol. 2, pp. 1117–1123. China (2005)
23. Cao, H., Ding, X., Liu, C.: Rectifying the bound document image captured by the camera: a model based approach. In: Proceedings of the 7th International Conference on Document Analysis and Recognition, vol. 1, pp. 71–75. UK (2003)
24. Liang, J., DeMenthon, D., Doermann, D.: Flattening curved documents in images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 338–345. San Diego, California, USA (2005)
25. Kashimura, M., Nakajima, T., Onda, N., Saito, H., Ozawa, S.: Practical introduction of image processing technology to digital archiving of rare books. In: Proceedings of the International Conference on Signal Processing Applications and Technology, Toronto, Canada, pp. 1025–1029 (1998)
26. Courteille, F., Durou, J.D., Gurdjos, P.: Transform your digital camera into a flatbed scanner. In: Proceedings of the 9th European Conference on Computer Vision, 2nd Workshop on Applications of Computer Vision, Graz, Austria, pp. 40–48 (2006)
27. Gumerov, N., Zandifar, A., Duraiswami, R., Davis, L.S.: Structure of applicable surfaces from single views. In: Proceedings of the 8th European Conference on Computer Vision, vol. 3. Lecture notes in Computer Science, vol. 3022, Prague, Czech Republic, pp. 482–496 (2004)
28. Cho, S.I., Saito, H., Ozawa, S.: A divide-and-conquer strategy in shape from shading problem. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 413–419. San Juan, Puerto Rico (1997)
29. Wada, T., Ukida, H., Matsuyama, T.: Shape from shading with interreflections under a proximal light source: distortion-free copying of an unfolded book. Int. J. Comput. Vis. **24**, 125–135 (1997)
30. Tan, C.L., Zhang, L., Zhang, Z., Xia, T.: Restoring warped document images through 3D shape modeling. IEEE Trans. Patt. Anal. Mach. Intell. **28**, 195–208 (2006)
31. Pilu, M.: Undoing paper curl distortion using applicable surfaces. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 67–72. Kauai, Hawaii, USA (2001)
32. Brown, M.S., Pisula, C.: Conformal deskewing of non-planar documents. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 998–1004. San Diego, California, USA (2005)
33. Horn, B.K.P: Understanding image intensities. Artif. Intell. **8**, 201–231 (1977)
34. Penna, M.A.: A shape from shading analysis for a single perspective image of a polyhedron. IEEE Trans. Patt. Anal. Mach. Intell. **11**, 545–554 (1989)
35. Lee, K.M., Kuo, C.C.J.: Shape from shading with perspective projection. Comput. Vis. Graph. Image Process. Image Underst. **59**, 202–212 (1994)
36. Hasegawa, J.K., Tozzi, C.L.: Shape from shading with perspective projection and camera calibration. Comput. Graph. **20**, 351–364 (1996)
37. Samaras, D., Metaxas, D.N.: Coupled lighting direction and shape estimation from single images. In: Proceedings of the 7th IEEE International Conference on Computer Vision, vol. 2, pp. 868–874. Kerkyra, Greece (1999)
38. Lions, P.L., Rouy, E., Tourin, A.: Shape-from-shading, viscosity solutions and edges. Numer. Math. **64**, 323–353 (1993)
39. Horn, B.K.P., Brooks, M.J. (eds.): Shape from shading. MIT, Cambridge (1989)
40. Horn, B.K.P., Brooks, M.J.: The variational approach to shape from shading. Computer Vision, Graph. Image Process. **33**, 174–208 (1986)
41. Nayar, S.K., Ikeuchi, K., Kanade, T.: Shape from interreflections. Int. J. Comput. Vis. **6**, 173–195 (1991)
42. Fisher, R., Perkins, S., Walker, A.E.W.: Pixel Division. http://www.homepages.inf.ed.ac.uk/rbf/HIPR2/pixdiv.htm
43. Smythe, D.B.: A two-pass mesh warping algorithm for object transformation and image interpolation. Technical Memo 1030, Industrial Light and Magic, Computer Graphics Department, Lucasfilm Ltd (1990)
44. Proceedings of the 9th IEEE International Conference on Computer Vision, vol. 2. In: Proceedings of the 9th IEEE International Conference on Computer Vision, vol. 2, Nice, France (2003)

Author Biographies

Frédéric Courteille received the Ph.D. degree in computer sciences at Paul Sabatier University, in Toulouse in 2006. He is currently an assistant rechearcher and a member of the TCI group (Traitement et Compréhension d'Images) in IRIT (Institut de Recherche en Informatique de Toulouse). He is interested in Shape From Shading, Shape From Contour and 3D reconstruction.

Alain Crouzil received the Ph.D. degree in computer sciences at Paul Sabatier University, in Toulouse in 1997. He is currently an assistant professor and a member of the TCI group (Traitement et Compréhension d'Images) in IRIT (Institut de Recherche en Informatique de Toulouse). He is interested in stereovision, motion analysis and Shape From Shading.

Jean-Denis Durou received the Ph.D. degree from the Université Paris XI, Orsay, France, in 1993. He is currently working in the field of Computer Vision at the Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier, Toulouse, France.

Pierre Gurdjos is with TCI (Traitement et Compréhension d'Images) at the Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier and Institut National Polytechnique de Toulouse, France. His research focuses on the problems of camera self-calibration, Structure From Motion using multiple views and Shape from Shading.