Photos, Objects and Computer Vision

Maurizio Falcone and Alessandra Seghini

In his novel "Let's stay together" of 1956, Isaac Asimov has described a critical situation: an officer of the US was interviewing the head of the Department of Robotics in Cheyenne to know more about the research in that area. He was scared by the fact that ten humanoids built by the enemy where walking around the States. At that time this was completely unrealistic, but it seems that the times have changed.

Nowadays robots can walk, read, speak, play a piano and behave as humans in many circumstances although their "intelligence" is still rather limited. However, recent advancements in computer vision can make them recognize a scene or an object and walk safely in an unknown environment avoiding obstacles. This is due to the findings in many research areas, including computer vision. Here we will describe in particular the Shape-from-Shading (SfS) problem, an inverse problem which amounts to reconstruct the 3D structure of an object given one 2D grey value image. For this task, the SfS process relies on informations on the illumination and the light reflectance in the scene. This problem, first introduced by Horn [9] in his PhD thesis, is a classic inverse problem in computer vision with many potential applications beyond robotics, see e.g. [6, 7, 11, 16, 20].

The Classical Shape from Shading Model

Let us give a brief outline of the SFS problem and introduce the basic assumptions for the classical model based on a single image. We attach to the camera a three-dimensional coordinate system ($Oxyz$), such that Oxy coincides with the image plane and Oz coincides with the optical axis, orthogonal projection is usually considered to simplify the model. Under this simplification, the visible part of the scene is, up to a scale factor, a graph $z = u(\mathbf{x})$, where $\mathbf{x} = (x, y)$ is an image point. The datum in this problem is our image which is represented by the *irradiance (or brightness)*

Maurizio Falcone, Alessandra Seghini
Dipartimento di Matematica, SAPIENZA - Università di Roma, Roma, Italy

function I. The function *I* is measured at each point (pixel) of the image, for example in terms of a greylevel (from 0 to 255). However, to construct a continuous model, we will assume that *I* takes real values in the interval [0, 1] where 0 corresponds to black and 1 to white passing through infinitely many gray levels. The SFS problem is characterized by the *image irradiance equation*:

$$R(\mathbf{n}(\mathbf{x})) = I(\mathbf{x}), \tag{1}$$

where $I(\mathbf{x})$ is the irradiance (or brightness) measured in the image at point $\mathbf{x}$ and $R(\mathbf{n}(\mathbf{x}))$ is the reflectance function, which gives the value of the light re-emitted by the surface as a function of its orientation i.e., of the unit normal $\mathbf{n}(\mathbf{x})$ to the surface at the point $(\mathbf{x}, u(\mathbf{x}))$. This normal can be expressed as:

$$\mathbf{n}(\mathbf{x}) = \frac{1}{\sqrt{1 + p(\mathbf{x})^2 + q(\mathbf{x})^2}} (-p(\mathbf{x}), -q(\mathbf{x}), 1), \tag{2}$$

where

$$p = \partial u/\partial x, \ q = \partial u/\partial y \text{ so that } \nabla u(\mathbf{x}) = (p(\mathbf{x}), q(\mathbf{x})) \tag{3}$$

(here $\partial u/\partial x$ and $\partial u/\partial y$ are the partial derivatives uf u with respect to the x and y coordinates and ∇u is the gradient of u). The height function u, which is the unknown of the problem, has to be reconstructed on a compact domain $\Omega \subset \mathbb{R}^2$, called the "reconstruction domain". Assume that there is a unique light source at infinity whose direction is indicated by the unit vector $\omega = (\omega_1, \omega_2, \omega_3) \in \mathbb{R}^3$. and let us assume, for simplicity, that ω is given. In order to simplify the model, it is usually assumed that the surface has uniform reflectance properties, so the light which is reflected by the surface just depends on the angle between the normal to the surface and the direction of the light ω scaled by a positive factor γ (the albedo). A surface with these properties is called a *Lambertian surface* and for it we can finally write (1) as $R(\mathbf{n}(\mathbf{x})) = \gamma \omega \cdot \mathbf{n}(\mathbf{x})$. Setting the light source at infinity guarantees that all the light rays are parallel. Then, setting the uniform albedo equal to 1 and applying definition (2), we obtain

$$I(\mathbf{x})\sqrt{1 + |\nabla u(\mathbf{x})|^2} + (\omega_1, \omega_2) \cdot \nabla u(\mathbf{x}) - \omega_3 = 0 \qquad \text{for } \mathbf{x} \in \Omega, \tag{4}$$

which is a first order non-linear partial differential equation of the Hamilton-Jacobi type (here $\cdot$ is the scalar product). We will also consider the simplified equation which appears in most of the literature and corresponds to a vertical light source at infinity i.e., $\omega = (0, 0, 1)$. Then (4) reduces to the *eikonal equation*

$$|\nabla u(\mathbf{x})| = f(\mathbf{x}) \qquad \text{for } \mathbf{x} \in \Omega, \tag{5}$$

Fig. 1 Object with occluding boundary, which might be used to define the silhouette Ω (the *white region* on the *right*)

where

$$f(\mathbf{x}) = \sqrt{\frac{1}{I(\mathbf{x})^2} - 1}.\qquad(6)$$

Since this equation just depend on ∇u, if u is a solution also $u+constant$ is a solution so one has (at least) to add the value at the boundary of the silhouette Ω in order select a unique solution (we will see later that this is not enough to get uniqueness). Equation (4) or (5) must be complemented with boundary conditions on $\partial\Omega$ and with additional information to select a unique solution. For an image containing an "occluding boundary", it is usual to consider this boundary as $\partial\Omega$. For example, in Fig. 1, the part of the image representing the object in greylevels is the *silhouette* is Ω and $\partial\Omega$ coincides with the occluding boundary.

A natural choice is to consider Dirichlet type boundary conditions in order to take into account (at least) two different possibilities. The first corresponds to the assumption that the surface is standing on a flat background i.e., we set:

$$u(\mathbf{x}) = 0 \qquad \text{for } \mathbf{x} \in \partial\Omega.\qquad(7)$$

The second possibility occurs when the height of the surface on the occluding boundary is known, e.g. this happens when we know (or assume) that the object is a surface of revolution around a given axis. This situation leads to the more general condition:

$$u(\mathbf{x}) = g(\mathbf{x}) \qquad \text{for } \mathbf{x} \in \partial\Omega.\qquad(8)$$

The solution of the above Dirichlet problems (4)–(7) or (4)–(8) will characterize a surface corresponding to the brightness function $I(\mathbf{x})$ measured in Ω which is the
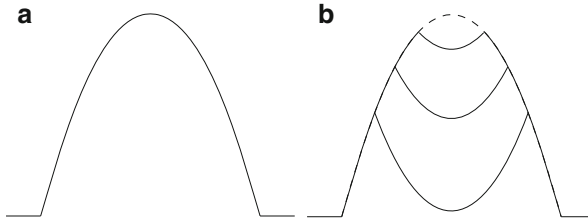
Fig. 2 Illustration of the concave/convex ambiguity: (**a**) maximal solution and (**b**) a.e. solutions giving the same image

datum of the problem. Points $\mathbf{x} \in \Omega$ such that $I(\mathbf{x})$ is maximal correspond to the particular situation where ω and $\mathbf{n}(\mathbf{x})$ point in the same direction: these points are usually called "singular points" and they play an important role to characterize a unique solution.

The first remark is that nonlinear partial differential equation such as (1) do not have in general regular solutions (i.e. solutions which are differentiable at every point). So the natural framework to set the problem is that of weak solutions which do not satisfy the equation everywhere but are at least continuous with (possibly) discontinuities in the gradient. Starting from the paper by Rouy and Tourin [18], the most recent approach to the resolution of SFS uses the notion of *viscosity solutions* to first order PDEs, see e.g. [1]. This is also natural from the point of view of computer vision applications: if you look around you can see many objects whose surface is not smooth. Unfortunately, the Dirichlet problem (4)–(8) can have several "weak solutions" in the viscosity sense and also several classical solutions. As an example, all the surfaces represented in Fig. 2 are viscosity solutions of the same equations (5)–(7), which is a particular case of (4)–(8). The solution represented in Fig. 2a is the maximal solution and is regular (differentiable). All the non-smooth a.e. solutions which can be obtained by a reflection with respect to a horizontal axis, are still admissible weak solutions (see Fig. 2b). In fact the value of the scalar product $\omega \cdot \mathbf{n}(\mathbf{x})$ corresponding to any of these surfaces is the same for every $\mathbf{x}$ and coincides with the same value computed on the regular solution (Fig. 2a). This is the *concave/convex ambiguity* of the SfS model. In this example, the lack of uniqueness of the viscosity solutions is due to the presence of a singular point, where the right hand side of (5) vanishes. An additional effort is then needed to define our solution since the lack of uniqueness is also a big drawback when trying to compute a numerical solution. In order to circumvent those difficulties, the problem is usually solved by adding some informations such as the height at each singular point [14] or the complete knowledge of a level curve [12, 13]. More recently, an attempt has been made to eliminate the need for a priori additional information. Always in the framework of viscosity solutions an effort has been made to characterize the *maximal solution* without additional informations and to build approximation schemes which converge to that solution [3, 8]. Let us finally mention that Eq. (4) is not the most general equation for the SfS. We made along

the way many assumptions but, in general, real materials are not Lambertian, the albedo is not constant, the direction ω is not known a priori. The situation is even more complex if interreflections are taken into account, but will not treat these cases in detail.

In order to obtain a *numerical solution of the SfS problem* two main strategies are available. The first is to solve directly the partial differential equation combined with the boundary conditions, this method will give an approximation of u provided additional conditions on the approximation scheme are satisfied. The second is the *variational approach* where we work first on an optimization problem in the unknown (p, q) at every pixel of the image. Then, using the field (p, q) and recalling their definition (3) we can reconstruct u by an integration along a path, made for example by a an horizontal plus a vertical step. We will not give the details here but they can be found in the survey [7].

The first approach is rather standard and the difficulty is to make the discretization of the continuous equation converge to the right solution. In the variational approach the main point is the choice of the functional to be minimized in the first step. In this approach, some regularity of the surface is assumed, and the approximate solutions computed are typically local minima of the functional. But u appears, in the image irradiance equation, only through its first derivatives p and q, which are two non-independent functions. In fact, if u is regular and twice differentiable

$$\partial p/\partial y = \partial q/\partial x. \tag{9}$$

and this condition guarantees integrability along a path. The only problem with these unknowns is that p or q become infinite at every point $\mathbf{x}$ belonging to the occluding boundary. This problem is not a cause for concern if no point $\mathbf{x}$ in the reconstruction domain Ω is such that $I(\mathbf{x}) = 0$. As Eq. (9) is a hard constraint on p and q, the most natural functional associated with Eqs. (1) and (9) is

$$\mathscr{F}_1\,(p, q, \mu) = \int_{\mathbf{x}\in\Omega} \left[r(p(\mathbf{x}), q(\mathbf{x})) - I(\mathbf{x})\right]^2\, d\mathbf{x} \\ + \int_{\mathbf{x}\in\Omega} \mu(\mathbf{x})\left[\partial p/\partial y(\mathbf{x}) - \partial q/\partial x(\mathbf{x})\right]\, d\mathbf{x}, \tag{10}$$

(here the function r is such that $R(\mathbf{n}(\mathbf{x})) = r(p(\mathbf{x}), q(\mathbf{x}))$) Since the minimization of $\mathscr{F}_1$ is rather slow, Horn and Brooks have also proposed the functional,

$$\mathscr{F}_2\,(p, q) = \int_{\mathbf{x}\in\Omega} \left[r(p(\mathbf{x}), q(\mathbf{x})) - I(\mathbf{x})\right]^2\, d\mathbf{x} \\ + \lambda_{\text{int}} \int_{\mathbf{x}\in\Omega} \left[\partial p/\partial y(\mathbf{x}) - \partial q/\partial x(\mathbf{x})\right]^2\, d\mathbf{x}. \tag{11}$$

where the constraint term for integrability becomes a penalty term and the positive constant weight λ_{int} called "integrability factor" (note that the choice of λ_{int} is arbitrary).

The PSFS Model and Related Equations

Several improvements with respect to the original problem presented in the previous section have been studied. One of the most relevant is the introduction of perspective deformations in the model as proposed in [5, 16, 19], also taking into account a light attenuation factor [17].

Let us sketch the model for PSFS with point light source located at the optical center O and light attenuation term (see Fig. 3). Let $\mathbf{x} = (x, y)$ be a point in the image domain Ω, where Ω is an open bounded subset of $\mathbb{R}^2$. Let $I = I(x, y) > 0$ be the normalised brightness function. We have $I = \frac{E(x,y)}{\sigma}$, where E is the greylevel of the given image and σ is the product of the surface albedo (which tells us to which extent the surface reflects light) and the light source intensity. Moreover, f denotes the focal length, i.e. the distance between the optical center C of the camera and the two-dimensional plane to which the scene of interest is mapped.

M will represent a generic point on the surface Σ to be reconstructed. We choose as unknown of the problem the function $u : \Omega \to \mathbb{R}$ such that

$$M = M(x, y) = u(x, y)\, m'\,, \tag{12}$$

where

$$m' = \frac{\mathsf{f}}{\sqrt{x^2 + y^2 + \mathsf{f}^2}}\, m \quad \text{and} \quad m = (x, y, -\mathsf{f})^\top. \tag{13}$$

Note that, according to these notations, $u > 0$ holds as the depicted scene is in front of the camera. Here we denote by $r(x, y)$ the distance between the point light source and the point $M(x, y)$ on the surface. It holds $u(x, y) = r(x, y)/\mathsf{f}$, since the light source location coincides with the optical center.

The model associated to the PSFS problem is obtained again by the image irradiance equation (1)

We denote by $\omega(x, y)$ the unit vector representing the light source direction at the point $M(x, y)$ (note that in the classic SFS model this vector is constant but this is not the case here):

$$\omega(x, y) = \frac{(-x, -y, \mathsf{f})^\top}{\sqrt{x^2 + y^2 + \mathsf{f}^2}}. \tag{14}$$

Adding the assumptions of a *light attenuation term* and of a *Lambertian surface*, the function R is defined as

$$R(\mathbf{n}(x, y)) = \frac{\omega(x, y) \cdot \mathbf{n}(x, y)}{r(x, y)^2}, \tag{15}$$
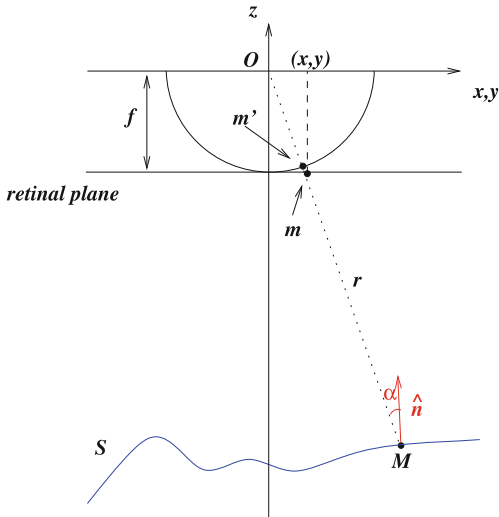
Fig. 3 Notations for the PSFS model with point light source at optical center O

with an attenuation factor which is equal to the inverse of the squared distance from the source. Expression (15) would still hold for any location of the point light source, but the same would not be true for the equality $u(x, y) = r(x, y)/\mathsf{f}$ nor for (14). The case of a light source coinciding with the optical center corresponds more or less to endoscopic images [16] and to photographs taken at short distance with the camera flash [17]. Another notable advantage of the PSFS model using a point light source at the optical center is that there is no shadow in the image.

Finally, by the irradiance equation and (15) we obtain the PSFS equation

$$\frac{\omega(x, y) \cdot \mathbf{n}(x, y)}{r(x, y)^2} = I(x, y). \tag{16}$$

In order to write down the corresponding PDE, it is useful to introduce the new unknown $v = \ln(u)$ (we recall that $u > 0$). Finally, Eq. (16) can be written as

$$H(x, y, v, \nabla v) := \frac{I(x, y)}{Q(x, y)} \mathsf{f}^2 \, W(x, y, \nabla v) - e^{-2v(x,y)} = 0, \quad (x, y) \in \Omega \tag{17}$$

where

$$Q(x, y) := \frac{\mathsf{f}}{\sqrt{x^2 + y^2 + \mathsf{f}^2}} \tag{18}$$

and

$$W(x, y, \nabla v) := \sqrt{\mathsf{f}^2 \|\nabla v\|^2 + (\nabla v \cdot (x, y))^2 + Q(x, y)^2}, \tag{19}$$
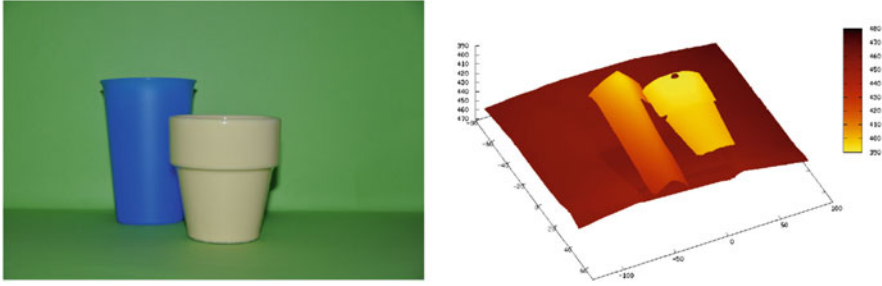
Fig. 4 (*Left*) Real image; (*right*) 3D reconstruction

(here $\| \cdot \|$ denotes the Euclidean norm). Figure 4 shows the result for this model on a real image representing two cups (see [2] for more details on these models).

Recent Developments on Shape from Shading Models

More recently the classic model based on Lambertian assumptions and on orthographic projection has been improved in order to deal with more realistic assumption and improve the results on real images. For example, one can consider the light source at finite distance from the object so that the rays are not parallel and/or include black shadows (i.e. regions where $I(x) = 0$) in the image. The black shadows are usually forbidden for two main reasons. The first reason is technical, in the classical model the eikonal equation can be applied only where $I(x) > 0$ because I appears at the denominator of the right-hand side term (6). The second is the remark that if the image is black we do not have informations on the real surface so we can not reconstruct it. Every surface whose height is below the "black ray" separating light and shadow could be a reasonable solution. Note that "black shadow" regions are typically projected by the object on the background and/or on its own surface and appear when the light source moves away from the vertical. This situation often appears in real images taken at sunset or at dawn. Since it is impossible, for lack of information, to reconstruct the real surface corresponding to these regions we can follow two different strategies. Either we apply the model only to the regions where $I > 0$ and deal with all the boundary conditions corresponding to these regions, or we solve the problem on the whole image and reconstruct a "virtual surface" on the black shadow regions. It is reasonable to define this virtual surface as the one described by the rays separating light and shadows.

The model has been extended in order to include in the solution the flat surface separating light and shadow (a result is shown is Fig. 5).

In Stereo Shape from Shading, one can consider two images *from different point of view under the same light conditions*. This situation mimics our vision system
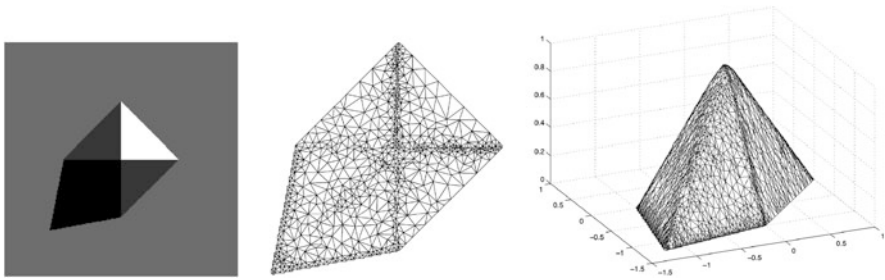
Fig. 5 From *left to right*: the real image with black shadows, the grid discretization of the domain Ω, the 3D reconstruction

based on two eyes separated by a small distance. It has been shown that for this specific problem there is just one solution and the problem is well-posed [4].

Another, recent improvement is the *stereophotometric approach* which is now considered a very promising technique. In this approach, we use several images *from the same point of view but taken under different light source directions*. This is the laboratory situations where one can move the light source around an object while taking several pictures. Clearly, for every image we can use one of the previous models for SfS (orthographic, perspective, light attenuation) and this gives rise to a system of nonlinear partial differential equations. For example, by applying the orthopraphic model and taking two images I_1 and I_2 which correspond respectively to two different light source directions ω^1 and ω^1, we get the following system of nonlinear partial differential equations

$$\begin{cases} I_1(\mathbf{x})\sqrt{1 + |\nabla u(\mathbf{x})|^2} + (\omega_1^1, \omega_2^1) \cdot \nabla u(\mathbf{x}) - \omega_3^1 = 0, \text{ for } \mathbf{x} \in \Omega \\ I_2(\mathbf{x})\sqrt{1 + |\nabla u(\mathbf{x})|^2} + (\omega_1^2, \omega_2^2) \cdot \nabla u(\mathbf{x}) - \omega_3^2 = 0, \text{ for } \mathbf{x} \in \Omega \end{cases} \quad (20)$$

Note that the number of equations is determined by the number of images and in some cases the system has a unique solution with just two images although, in general, to select a unique solution three images are necessary. For example, starting from the three images of the Beethoven bust (Fig. 6) we can compute the solution of the system and get back the 3D reconstruction of the surface (Fig. 7) (see [15] for more details).

As we mentioned at the beginning, the above techniques allow for a rather accurate 3D reconstructions of the objects appearing in a scene. Then, this technology can be applied in medical surgery where the image acquisition is obtained via an endoscope and is important to have a precise description of the area. At a macroscopic level the same inverse problem appears in space investigations when a satellite takes several pictures of the same area allowing for the construction of its altitude map which can be used later for robot explorations in the area.

Another interesting applications of the SfS models is related to the preservation of cultural heritage. The advanced techniques for 3D reconstruction are based on
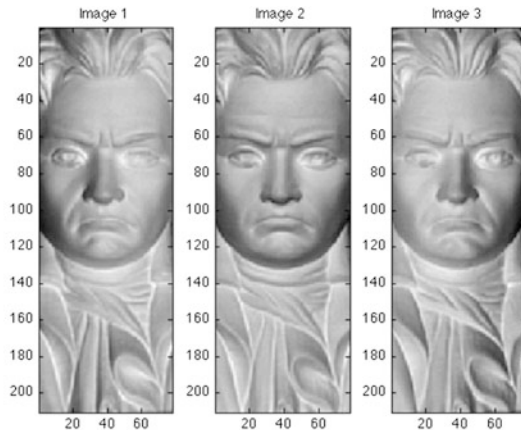
I_1 for ω' : $\varphi_1 = 0.263, \theta_1 = -0.305$ I_2 for ω'' : $\varphi_2 = 0.2, \theta_2 = 3.226$

**Fig. 6** Photometric stereo: Three real images of Beethoven bust
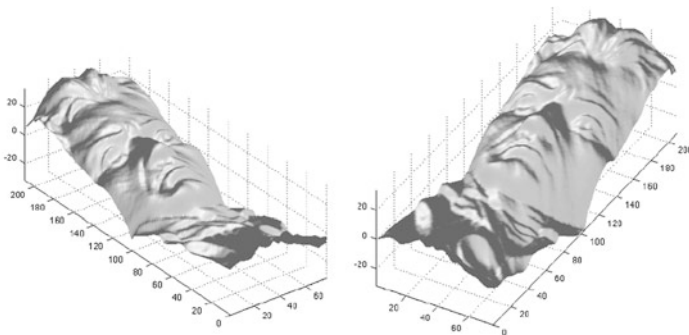
I_1 for ω' : $\varphi_1 = 0.263, \theta_1 = -0.305$ I_2 for ω'' : $\varphi_2 = 0.2, \theta_2 = 3.226$

Fig. 7 The 3D reconstruction of the Beethoven bust starting from the data in Fig. 6

laser or optical measurements but 3D scanners are rather expensive since a specific hardware for the data acquisition is required (see for example the movie presentation [21]). An accurate reconstruction based on SfS model could produce a digitization of a statue just using low cost cameras in a controlled environment where one can adjust the light sources. This is clearly an advantage since the method can be easily applied with low investments.

Another example comes from the digitization of ancient manuscripts. These books are very delicate and their reproduction requires an expensive hardware in order to preserve the volumes. Moreover, the copy of the manuscript via a standard camera has a big perspective distortion which is due to the curvature of the pages
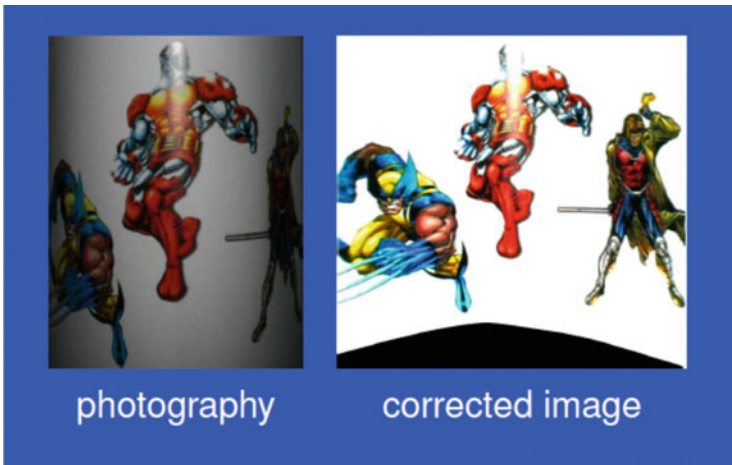
Fig. 8 An example of corrections for book reproduction via SfS models

near the center of the book. Typically the text is not horizontal due to this distortion and classical OCR (Optical Character Recognition) systems cannot work on it, so an automatic classification or analysis of the text is not feasible. Via SfS technique one can compute the real surface of the pages and then reapply the content of the page on it correcting the perspective distortion as in Fig. 8 (more details can be found in [6]).

So computer vision can really help robots to make a walk and, finally, the science fiction novel imagined by Asimov becomes reality.

References

1. G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*. Mathématiques & Applications, vol. 17 (Springer, Berlin, 1994)
2. M. Breuß, E. Cristiani, J.-D. Durou, M. Falcone, O. Vogel, Numerical algorithms for perspective shape from shading. Kybernetika **46**, 207–225 (2010). http://www.kybernetika.cz/content/2010/2/207
3. F. Camilli, L. Grüne, Numerical approximation of the maximal solutions for a class of degenerate Hamilton-Jacobi equations. SIAM J. Numer. Anal. **38**, 1540–1560 (2000)
4. A. Chambolle, A uniqueness result in the theory of stereo vision: coupling shape from shading and binocular information allows unambiguous depth reconstruction. Ann. Inst. Henri Poincaré Analyse Non Linéaire **11**, 1–16 (1994)
5. F. Courteille, A. Crouzil, J.-D. Durou, P. Gurdjos, Towards shape from shading under realistic photographic conditions, in *Proceedings of 17th International Conference on Pattern Recognition*, Cambridge, vol. II (2004), pp. 277–280
6. F. Courteille, A. Crouzil, J.-D. Durou, P. Gurdjos, Shape from shading for the digitization of curved documents. Mach. Vis. Appl. **18**, 301–316 (2007)

7. J.-D. Durou, M. Falcone, M. Sagona, Numerical methods for shape-from-shading: a new survey with benchmarks. Comput. Vis. Image Underst. **109**, 22–43 (2008)
8. M. Falcone, M. Sagona, An algorithm for the global solution of the shape-from-shading model, in *Proceedings of 9th International Conference on Image Analysis and Processing*, Florence, vol. I (1997), pp. 596–603
9. B.K.P. Horn, Shape from shading: a method for obtaining the shape of a smooth opaque object from one view, Ph.D. thesis, Department of Electrical Engineering, MIT, Cambridge, 1970
10. B.K.P. Horn, Obtaining shape from shading information, in *The Psychology of Computer Vision*, ed. by P.H. Winston, chap. 4 (McGraw-Hill, New York, 1975), pp. 115–155
11. B.K.P. Horn, M.J. Brooks, *Shape from Shading*. Artificial Intelligence Series (MIT Press, Cambridge, 1989)
12. R. Kimmel, A.M. Bruckstein, Global shape from shading. Comput. Vis. Image Underst. **62**, 360–369 (1995)
13. R. Kimmel, K. Siddiqi, B.B. Kimia, A. Bruckstein, Shape from shading: level set propagation and viscosity solutions. Int. J. Comput. Vis. **16**, 107–133 (1995)
14. P.-L. Lions, E. Rouy, A. Tourin, Shape-from-shading, viscosity solutions and edges. Numer. Math. **64**, 323–353 (1993)
15. R. Mecca, M. Falcone, Uniqueness and approximation of a photometric shape-from-shading model. SIAM J. Imaging Sci. **6**, 616–659 (2013)
16. T. Okatani, K. Deguchi, Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center. Comput. Vis. Image Underst. **66**, 119–131 (1997)
17. E. Prados, F. Camilli, O. Faugeras, A unifying and rigorous shape from shading method adapted to realistic data and applications. J. Math. Imaging Vis. **25**, 307–328 (2006)
18. E. Rouy, A. Tourin, A viscosity solutions approach to shape-from-shading. SIAM J. Numer. Anal. **29**, 867–884 (1992)
19. A. Tankus, N. Sochen, Y. Yeshurun, A new perspective [on] shape-from-shading, in *Proceedings of 9th IEEE International Conference on Computer Vision*, Nice, vol. II (2003), pp. 862–869
20. R. Zhang, P.-S. Tsai, J.E. Cryer, M. Shah, Shape from shading: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **21**, 690–706 (1999)
21. 3D scanner of a statue of Venus, 2011, http://www.youtube.com/watch?v=wCwommLtMxg