Creation of Real Images which are Valid for the Assumptions Made in Shape From Shading

Pascal Daniel Jean-Denis Durou
IRIT - Institut de Recherche en Informatique de Toulouse
UPS, 118 route de Narbonne, 31062 Toulouse Cedex 04, France
daniel@irit.fr, durou@irit.fr

Abstract

Iterative methods showed until now encouraging results to resolve shape from shading. This kind of methods generally work on synthetic images, and occasionally on real images, even if such images do not agree in general with the hypotheses of shape from shading. In this article, we describe the assumptions mentioned above, and propose an original methodology, which enables the production of real images and the process allowing to correct most of their defects in order to make them correspond to most of these assumptions. Moreover, we propose a descent gradient iterative scheme, thanks to which we prove that shape from shading can work as well on real images as on synthetic images.

1 Introduction

Shape from shading recovering the shape starting from a single image of a surface is an inverse problem considered being difficult to resolve, insofar that no general method of resolution has been found to date. Only iterative methods obtain encouraging results such as, for example: the method of *combined gradients* [6], methods resulting from the theory of *viscosity solutions* [5, 1] or those of the resolution of Euler's equations [4, 3]. However, one should note that they are generally shown to be valid only on synthetic images. All of these methods attempt to solve the problem of shape from shading by making a rather large number of assumptions which allow the formulation, and hence the resolution of the problem, to be simplified. However, the validation of these methods can only be completely satisfactory when they are tested on real images, even if such images do not agree in general with the hypotheses of shape from shading. Sometimes, when the same iterative method is tested initially on synthetic images and then on real images [4, 7], the quality of the obtained results decreases noticeably.

In this article, we propose an original methodology which enables the production of real images corresponding to most of the assumptions which can be necessary for shape from shading. A detailed presentation of these assumptions is made in Section 2. In Section 3, the optical arrangement used is described. In section 4, processes which eliminate the residual defects from the obtained images are detailed. In Section 5, a comparison is carried out between one of the obtained images representing a sphere, and the corresponding synthetic image, and a standard method of reconstruction is implemented on these images.

2 The Assumptions of Shape From Shading

The physical space is mapped onto an orthogonal reference mark $Oxyz$, so that Oz coincides with the optical axis of the observer and is directed towards him. One supposes that the scene is reduced to a single object, whose surface can be described by the height $z(x,y)$. The image obtained is a grey scale image which can be described by a positive function $E(x,y)$. The calculation of the function $z(x,y)$, starting from knowing the function $E(x,y)$, the characteristics of the light sources, the characteristics of the observer and the photometric characteristics of the surface is called *shape from shading problem*, for which various mathematical formulations are possible. However, most of the methods of resolution which provide relatively satisfactory results make a rather large number of implicit assumptions described hereafter.

2.1 Minimal Assumptions of Shape From Shading

The image is sharp, and can be considered, to within about a scale factor called *magnification*, to be like the result of the orthographic projection of a 3D scene, on the focal plane of the observer. One thus neglects the phenomenon of perspective projection. The photosensitive receiver is linear. There is only one point light source, placed sufficiently far from the scene so that it is considered as being *at infinity*, and that it has very low angular dimensions. The lighting of the scene is thus related to a parallel and uniform light

beam, which can be described by a single vector $\vec{S}$. The surface presents neither hidden part nor edge, consequently the function $z(x,y)$ is continuous and derivable. The surface is opaque and presents the same photometric characteristics at each point (*homogeneous* surface). One can finally neglect secondary light beams, resulting from the surface itself, with respect to the principal light beam $\vec{S}$.

Under these assumptions, one can show that the grey level $E(x,y)$ depends only on the normal $\vec{N}$ of the surface (outgoing normal, of norm 1), on $\vec{S}$, on the observer (magnification, factor of linearity of the receiver, diaphragm aperture and exposure time) and on the photometric characteristics of the surface. Among all these elements, only $\vec{N}$ varies from one point of the surface to another one [2], leading to the following equation:

$$R(\vec{N}) = E(x,y) \tag{1}$$

where R, which is called the *reflectance map*, depends on $\vec{S}$, on the characteristics of the observer, and on the photometric characteristics of the surface. This equation is usually called the *image irradiance equation*. It constitutes one of the three major equations of shape from shading. In (1), the unknown $z(x,y)$ does not appear directly, but only through its two first derivatives $p = \partial z/\partial x$ and $q = \partial z/\partial y$, since the vector $\vec{N}$ has as coordinates:

$$\vec{N} = \frac{1}{\sqrt{1+p^2+q^2}} \begin{pmatrix} -p & -q & 1 \end{pmatrix}^T \tag{2}$$

Nevertheless, the equation (1) remains very general, insofar as the function R (which depends on many parameters) is *a priori* unknown. For this reason, in the majority of the articles discussing shape from shading, two additional assumptions are added to the minimal assumptions.

2.2 First Additional Assumption

This assumption consists in supposing that the light beam is parallel to the Oz axis (*front lighting*). In this case, due to symmetry, it is obvious that the function R varies from one point of the surface to another only via the slope [5], which has as value: $\|\vec{\nabla} z\| = \sqrt{p^2+q^2}$. The reflectance map can thus be written, in this case:

$$R(\vec{N}) = r(\sqrt{p^2+q^2}) \tag{3}$$

where r is a function which depends on the norm S of $\vec{S}$, on the characteristics of the observer and on the photometric characteristics of the surface. The function r is decreasing for all common materials, which means that, the darker a point appears on the image, the more the slope at this point is large. According to (1) and (3), the image irradiance equation can be rewritten:

$$p^2 + q^2 = \left(r^{-1} \circ E(x,y)\right)^2 \tag{4}$$

since the function r, which generally decreases, can be inverted. The equation (4), which is called the *eikonal equation*, is the second major equation of shape from shading. Just like the image irradiance equation, it is a partial derivative equation in $z(x,y)$ of the first order of derivation, non linear. The right hand side of this equation is *a priori* unknown, insofar as the function r is *a priori* unknown.

2.3 Second Additional Assumption

This assumption originates from the supposition that the surface perfectly diffuses (lambertian assumption), *i.e.*, the grey level of a surface element is constant, when the direction of observation varies (all other factors being constant otherwise). The reflectance function can be expressed [2] therefore in the following particularly simple form:

$$R(\vec{N}) = k \left| \vec{S}.\vec{N} \right| \tag{5}$$

where k is a constant which depends on the characteristics of the observer and on the albedo of the surface. According to the first additional assumption, $\vec{S}$ has the following coordinates:

$$\vec{S} = S \begin{pmatrix} 0 & 0 & -1 \end{pmatrix}^T \tag{6}$$

Using (2) and (6), the expression (5) can be rewritten:

$$R(\vec{N}) = \frac{kS}{\sqrt{1+p^2+q^2}} \tag{7}$$

This expression is definitely of the form $r(\sqrt{p^2+q^2})$, with r decreasing, since S is considered to be a constant. According to (1) and (7), the maximum grey level E_{max} is equal to kS, and the first derivative p and q of the function $z(x,y)$ thus agree with the following equation:

$$p^2 + q^2 = \left(\frac{E_{max}}{E(x,y)}\right)^2 - 1 \tag{8}$$

The equation (8), which is a particular form of the eikonal equation (4), constitutes the third and last major equation of shape from shading. Its right hand side is known, and lies between 0 and $+\infty$. The difficulty of its resolution comes from its non linearity.

We will now describe a methodology which allows us to obtain real images which are valid for almost all the assumptions of shape from shading.

3 Description of the Optical Arrangement

In order to create digital images which are valid for the assumptions of shape from shading, we set up an optical arrangement which is represented schematically on Figure 1. The following describes the various stages which have led to its realization.

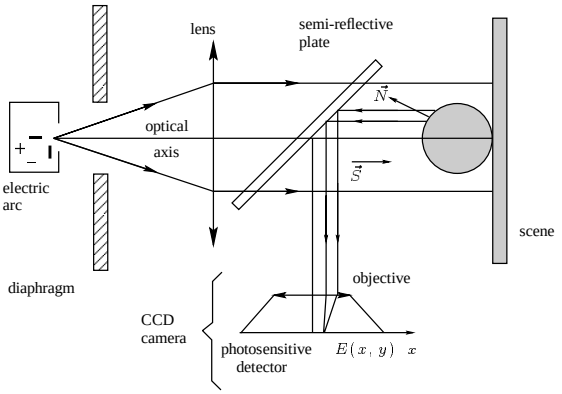


Figure 1. Optical arrangement used.

3.1 The Light Source

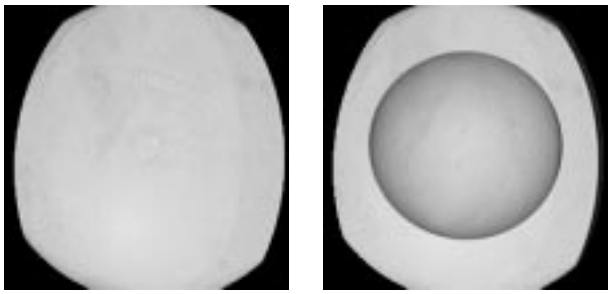
The aim concerning the light source is to produce a parallel and uniform light beam. However, the only light sources producing a parallel beam are lasers, but their light present the "defect" of being too coherent, which causes an interferential phenomenon (*speckle*) on the images which results in a graininess. Therefore an electric arc is used. The crater of its positive electrode, which is the most emissive part, is placed at the focus of an achromatic lens $\mathcal{L}$, of diameter $d_1 = 10\,cm$ and of focal distance $f_1 = 1\,m$. As this crater has a diameter of about $1\,mm$, the divergence of the light beam does not exceed $10^{-3}\,rad$, thus one can regard the beam obtained in this manner as almost parallel. A residual defect of this beam, called D_1, is that it is not uniform. Another defect of this lighting, called D_2, is the appearance of a parasitic luminous background, caused by the elements of the arrangement and their environment.

3.2 The Observer

The principal difficulty regarding the observer is to be able to check the assumption of front lighting. Even whilst moving the observer away from the scene, one can at most decrease the angle between Oz and $\vec{S}$, but never cancel it, *i.e.*, one always sees shadows. In order to make Oz and $\vec{S}$ coincide, a semi-reflective plate of large size (diameter $d_2 = 15\,cm$) is used, tilted with a $45°$ angle compared to $\vec{S}$. A light beam crossing this lens has a cross-section which is the intersection of a circle and an ellipse, of which the smallest dimension is about $9\,cm$. The size of the initial beam is thus only slightly altered, whereas its intensity is decreased by half. The numerical camera used is a *CANON Power Shot 600*, of focal length $f_2 = 7\,mm$ and aperture equal to 2.5. Its optical axis is placed orthogonally to $\vec{S}$, so that, thanks to the semi-reflective plate, a perfect colinearity between this optical axis and $\vec{S}$ is created. A first residual defect of this observer is the phenomenon of multiple reflections between the two faces of the semi-reflective lens. The split of the light beam due to the semi-reflective plate accentuates its non uniformity, and thus constitutes one of the causes of the defect D_1. Moreover, the image of the observed object will be also duplicated, and we agree to call this new defect D_3. Another defect, called D_4, comes from the possible non linearity of the receiver.

3.3 The Scene

According to events, various objects have been used to produce a scene: a plane, a sphere (billiard ball of diameter $5.7\,cm$) and a moulding representing a stag's head (of size $\approx 4\,cm$). None of these objects has neither an edge nor a hidden part inside its silhouette. For each one of them, the variation in height does not exceed $5\,cm$ thus, while placing the scene within approximately $d_3 = 1\,m$ of the camera, one can consider that the perspective effect is completely negligible. The distance d_3 being large in comparison with f_2 and the aperture of the objective being rather weak ($f_2/2.5$), the image is sharp. In order that the surfaces of these three objects are homogeneous, they were all covered with the same white mat painting. However, this painting does not constitute a rigorously lambertian material, introducing a defect called D_5. Lastly, another residual defect concerning the scene, called D_6, is that there can exist secondary light beams resulting from the surface itself, in the case of non convex objects. The images of the three objects used are represented on Figures 2.a, 3.a and 4.a.



(a) Image of the plane. (b) Image of the sphere on a white background.

Figure 2. Images of reference.

In Section 4, we shall see that the image analysis of the plane and of the sphere allow us to correct some of the six defects which have been just raised.

4 Correction of the Defects

The images which were created, whilst using the arrangement described previously, are valid for all the assumptions of shape from shading, within the defects D_1, D_2, D_3, D_4,

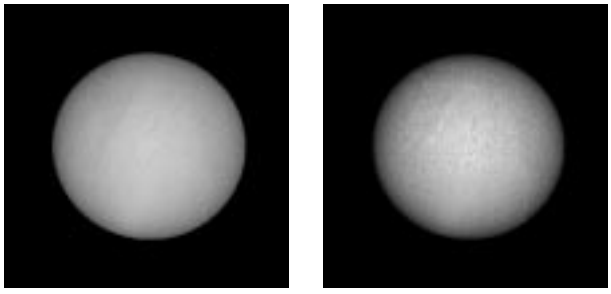
D_5 and D_6. Among these six defects, three will not be corrected within the scope of this work:

- The defect D_4 can be corrected only if the response curve of the receiver is known, which is not the case for the camera used.
- The correction of the defects D_3 and D_6 presents difficulties which would alone constitute a whole field of research. We do not correct these defects, and accept that they represent two additional sources of noise.

4.1 Correction of the Defect D_2

The presence of the defect D_2 results, for example in Figure 3.a, in the fact that the grey level $E_S(x,y)$ is non null on the silhouette of the sphere, which should be the case, whatever the material used. The correction of this defect can thus be made by seeking the minimum $E_{S,min}$ of the function $E_S(x,y)$ for the points of the silhouette of this object. This value corresponds to the parasitic ambient light. To determine precisely these points, the image of the sphere on a white background is used, which is represented on Figure 2.b and on which the silhouette is clearly detached. One finds $E_{S,min} = 53$. It is then enough to subtract this value from any grey level $E(x,y)$ in the image, which gives a new image, of grey level $\bar{E}(x,y) = E(x,y) - E_{S,min}$.

4.2 Correction of the Defect D_1



(a) Initial image. (b) After three corrections.

Figure 3. Images of the sphere.

To carry out the correction of the defect D_1, one places the plane object at the orthogonal of $\vec{S}$. If the light beam were uniform, the grey level would be uniform. However this is not the case since, on the image of Figure 2.a, the ratio between the smallest grey level and the largest one is: $\frac{\bar{E}_{P,min}}{\bar{E}_{P,max}} \approx 0.62$. Let us suppose that $\bar{E}_{P,max}$ corresponds to a point of the plane where the norm of $\vec{S}$ is S_{max}. More generally, let us suppose that $\bar{E}_P(x,y)$ represents the grey level of the image of the plane, after correction of the defect D_2, and that $S(x,y)$ represents the norm of $\vec{S}$ at the point (x,y).

For the illuminated points of the scene, one has:

$$\frac{\bar{E}_P(x,y)}{S(x,y)} = \frac{\bar{E}_{P,max}}{S_{max}} \tag{9}$$

while supposing that the grey level is proportional to the incident luminous flux. That is to say an image with the (corrected) grey level $\bar{E}(x,y)$, and the same light beam $S(x,y)$ as that for the plane. One can calculate the corrected grey level $\tilde{E}(x,y)$ of the image which would be obtained if the luminous flux were uniform and equal to S_{max}, thanks to the following relation:

$$\tilde{E}(x,y) = \frac{S_{max}}{S(x,y)}\bar{E}(x,y) \tag{10}$$

Whilst using (9):

$$\tilde{E}(x,y) = \frac{\bar{E}_{P,max}}{\bar{E}_P(x,y)}\bar{E}(x,y) \tag{11}$$

On the image of Figure 4.a, one notices the trace of a split of the light beam, on the right of the stag's head, whereas this trace has disappeared after corrections (Figure 4.b).

(a) Initial image. (b) After three corrections.

Figure 4. Images of the stag's head.

4.3 Correction of the Defect D_5

If the painting used were lambertian, then, according to (7), its reflectance function would have the following form:

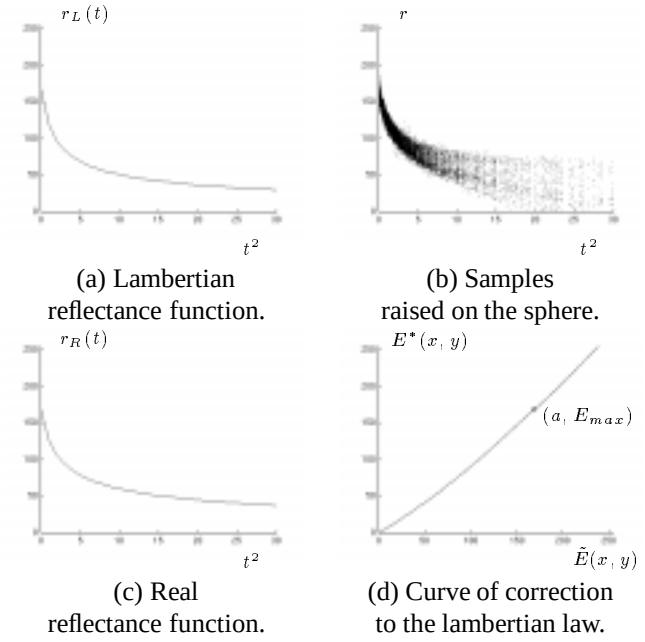
$$r_L(t) = \frac{E_{max}}{\sqrt{1+t^2}} \tag{12}$$

where t denotes the slope. A function of this type is represented on Figure 5.a (t^2 has been used as x-coordinate rather than t, in order to have a better spreading out of the curves). The function $r_R(t)$ denotes the reflectance function of the used painting. By using the image (twice corrected) of the sphere, of grey level $\tilde{E}_S(x,y)$, one can calculate, for each point (x,y) located inside the silhouette, the slope t (computed thanks to the knowledge of the sphere shape) and the value $r_R(t)$, which is equal, according to (1), to $\tilde{E}_S(x,y)$. It is then enough to defer, in a reference mark (t^2, r), all the

couples of values thus obtained as samples (Figure 5.b). One notes a certain analogy between the function $r_L(t)$ and the general bearing of this scatter plot. It is thus legitimate to seek an approximation of the function $r_R(t)$ in the following form, which takes as a starting point the lambertian law:

$$r_R(t) = \frac{a}{\sqrt{1+bt^c}} \tag{13}$$

Thanks to a least squares estimation, one finds the following numerical values: $a = 169.2$, $b = 0.797$ and $c = 1.873$.



(a) Lambertian reflectance function.
(b) Samples raised on the sphere.
(c) Real reflectance function.
(d) Curve of correction to the lambertian law.

Figure 5. Variation to the lambertian law.

This approximation of the function $r_R(t)$ is represented on Figure 5.c. If $\tilde{E}(x,y)$ represents the grey level (after correction of the defects D_1 and D_2) of the image of a scene covered with the same painting as for the sphere, then, according to (1), at each point (x,y) of the image, one must have:

$$\tilde{E}(x,y) = r_R(t) \tag{14}$$

However the function $r_R(t)$ being decreasing, can be inverted, which enables (14) to be rewritten as:

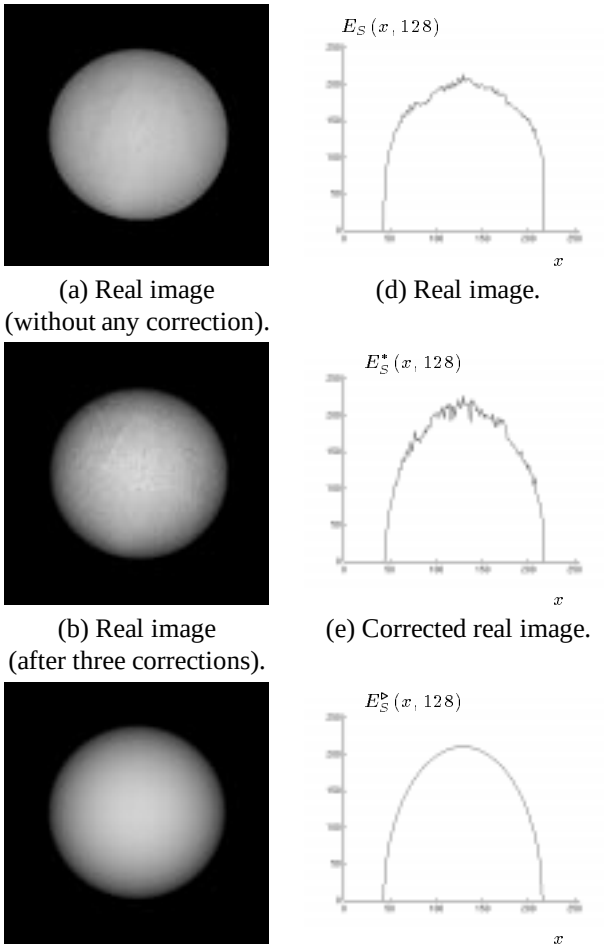
$$t = r_R^{-1}(\tilde{E}(x,y)) \tag{15}$$

Thus, to correct the defect D_5 on an image of grey level $\tilde{E}(x,y)$, it is sufficient to calculate the grey level $E^*(x,y)$ as defined by $E^*(x,y) = r_L(t)$ or, according to (15):

$$E^*(x,y) = r_L \circ r_R^{-1}(\tilde{E}(x,y)) \tag{16}$$

The relation (16) is applicable only for $\tilde{E}(x,y) \in [0,a]$. Indeed, one sees on Figure 5.c that no slope corresponds to a grey level higher than a. It is thus necessary to prolong the relation (16) beyond this limiting value. On Figure 5.d, we graphically represented the relation (16), which we prolonged analytically, beyond the point represented by a small circle, up to the value 255. On this figure, it appears that the relation (16) is almost linear, which means that the painting used was almost lambertian! In choosing $E_{max} = a$, the correction of the defect D_5, never makes the grey level of a point vary by more than 1, *i.e.*, this correction is invisible to the naked eye.

5 Purpose of the Given Corrections



(a) Real image (without any correction).
(d) Real image.
(b) Real image (after three corrections).
(e) Corrected real image.
(c) Synthetic image.
(f) Synthetic image.

Figure 6. Comparison to a synthetic image.

5.1 Real and Synthetic Images Comparison

To highlight the validity of the methodology that we have proposed, it is interesting first to compare the images obtained with synthetic images of the same objects, calculated with all the assumptions of shape from shading. This is what

has been done for the sphere (we do not unfortunately have the height function of the stag's head). Figure 6 represents the real images of the sphere before correction (Figure 6.a), after correction of the defects D_2, D_1 and D_5 and an increase in its contrast (Figure 6.b), and the synthetic image of a sphere of the same diameter and of grey level $E_S^{\triangleright}(x, y)$ (Figure 6.c). One should note that there is a very strong similarity between the images represented on Figure 6.b and Figure 6.c, rather than between Figure 6.a and Figure 6.c, which tends to prove that the suggested corrections bring closer the real images to the corresponding synthetic images. A way of better appreciating the very strong similarity between these two last images consists in plotting the grey levels along the horizontal diameter of these three images. The functions $E_S(x, 128)$, $E_S^*(x, 128)$ and $E_S^{\triangleright}(x, 128)$ have been traced on Figures 6.d, 6.e and 6.f, and the grey scale comparison allow to make the same observation again, even if the corrections made increase the noise.

5.2 3D Reconstruction on Real Images

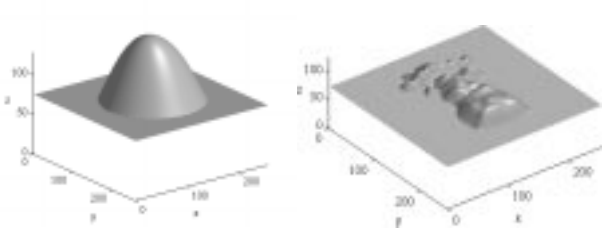
We will now be interested in an iterative method of resolution of shape from shading based on the minimization of an error $\varepsilon(\omega)$ similar to that of Horn and Ikeuchi [4]:

$$\varepsilon(\omega) = \sum_{(i,j)\in D}\sum \left(E_{i,j} - \frac{E_{max}}{\sqrt{1 + {p_{i,j}}^2 + {q_{i,j}}^2}} \right)^2 + \lambda \sum_{(i,j)\in D}\sum \left(\left\| (\overrightarrow{\nabla}\ p)_{i,j} \right\|^2 + \left\| (\overrightarrow{\nabla}\ q)_{i,j} \right\|^2 \right) \qquad (17)$$

where $\omega = (p_{i,j}\ q_{i,j})^T_{(i,j)\in D}$ and where D denotes the domain on which the shape have to be reconstructed. The height is fixed to 0 on the outline of D. The parameter λ is positive and called the *smoothing factor*. The problem requires that an absolute minimum of this error is found, *i.e.*, commencing from an arbitrary starting configuration ω^0, one wishes to find a particular configuration ω^* which is an absolute minimum of the error ε. From a given configuration ω^k, one moves towards a new configuration ω^{k+1}, in the direction of steepest descent, which ensures a decrease of the error, if the length of the displacement is correctly chosen. A method based on gradient descent is defined by an iterative scheme of the type:

$$\omega^{k+1} = \omega^k - dl^k \frac{(\overrightarrow{\nabla}\ \varepsilon)^k}{\|(\overrightarrow{\nabla}\ \varepsilon)^k\|} \qquad (18)$$

with dl^k *a priori* unspecified. The statement of the suggested method is the following: at each iteration k, one calculates the displacement dl^k for which the error is minimal, which ensures an optimal decrease of the error and guarantees the convergence of the method whatever the value of λ, at least toward a relative minimum of ε, which constitutes a trap for all of the methods based on gradient descent. For the images on Figures 3.b and 4.b, we have implemented the method which has been just described with $\lambda = 10^5$. On the one hand, we checked that the reconstructed shape obtained from the real image of the sphere (Figure 7.a) is very close to that corresponding to the synthetic image represented on Figure 6.c. On the other hand, one can observe that the reconstructed shape of the stag's head (Figure 7.b) seems to be entirely satisfactory. This leads to the following conclusion.



(a) From Figure 3.b. (b) From Figure 4.b.

Figure 7. 3D reconstructions obtained.

6 Conclusion

One has now a methodology enabling the creation of real images which are semi perfectly valid for the many assumptions of shape from shading. One succeeded in proving that shape from shading can work as well on real images as on synthetic images, insofar as with a standard method of reconstruction, one obtains results of equivalent quality. The various methods of resolution which have been proposed in the literature will be able to be tested on real images, without having to attribute the bad quality of the reconstructed shape to imperfections in the quality of the images.

References

[1] M. Falcone and M. Sagona. An algorithm for the global solution of the shape-from-shading model. In *ICIAP'97 (IEEE International Conference on Image Analysis and Processing, Florence, Sept 1997)*, pages 596–603, 1997.

[2] B. K. P. Horn. Obtaining shape from shading information. *The Psychology of Computer Vision*, 1975.

[3] B. K. P. Horn. Height and gradient from shading. *International Journal of Computer Vision*, 5(1):37–75, aug 1990.

[4] K. Ikeuchi and B. K. P. Horn. Numerical shape from shading and occluding boundaries. *Art. Intell.*, 17:141–194, 1981.

[5] P. L. Lions, E. Rouy, and A. Tourin. Shape-from-shading, viscosity solutions and edges. *Num. Math.*, 64:323–353, 1993.

[6] R. Szeliski. Fast shape from shading. In O. Faugeras, editor, *Proceedings of the 1st European Conference on Computer Vision (Antibes, France, April 23-27, 1990)*, volume 427 of *LNCS*, pages 359–368, Berlin, apr 1990. Springer.

[7] Q. Zheng and R. Chellapa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Transactions on PAMI*, 13(7):680–702, jul 1991.