A PHOTOMETRIC MODEL FOR SPECULAR HIGHLIGHTS AND LIGHTING CHANGES. APPLICATION TO FEATURE POINTS TRACKING.

M. Gouiffès(1) *C. Collewet*(2) *C. Fernandez-Maloigne*(1) *A. Trémeau*(3)

(1) : SIC, University of Poitiers, France (2) : INRIA/IRISA, Rennes, France (3) : LIGIV, Saint-Etienne, France

ABSTRACT

This article proposes a local photometric model that compensates for specular highlights and lighting variations due to position and intensity changes. We define clearly on which assumptions it is based, according to widely used reflection models. Moreover, its theoritical validity is studied according to few configurations of the scene geometry (lighting, camera and object relative locations). Next, this model is used to improve the robustness of points tracking in luminance images with respect to specular highlights and lighting changes.

1. INTRODUCTION

Most computer vision applications based on the computation of correspondences between images are sensitive to illumination changes. Indeed, the robustness of image processing with regard to such a phenomenon remains a crucial problem. Moreover, in luminance images, no global compensation of illumination changes has been defined as far now, contrary to color images [2]. In general, luminance is supposed to be constant in a small window of interest $\mathcal{W}$ during an image sequence [5], but that is correct only for lambertian objects viewed under perfectly constant lighting. Photometric normalization, as in [9] for example, or local photometric models answer partially this issue. The affine photometric model has been widely used, for example in optical flow [7] or tracking [6] applications, but the parameters of this model are supposed to be spatially constant in each point of $\mathcal{W}$. Moreover, it has been shown in [3] that this model is based on the same assumptions as the photometric normalization. In [1] however, the spatial variations of illumination are taken into account. Nevertheless, the assumptions on which these local models are based (according to the physical properties of materials and to the scene geometry) have not been clearly specified, and the definition of each of their parameters has not been given. In this article, we propose a local photometric model adapted to specular highlights and lighting changes. We analyze clearly the assumptions on which it is based, according to the reflection models that are widely used in computer vision and computer graphics. Our second contribution consists in studying theoritically the validity of the proposed model according to some borderline geometry conditions, such as the position of the camera and the light source. To finish, we focus on feature points tracking. The proposed photometric model is used to improve robustness of points tracking with respect to specular highlights and lighting changes, the parameters of the photometric model are computed simultaneously with the parameters of the motion model.

This article is structured as follows. Section 2 focuses on the modeling of luminance changes. Then, the model of photometric changes and its local approximation are proposed in section 3. The section 4 is dedicated to the study on the model validity. To finish, section 5 exposes the tracking method and shows some results obtained on real image sequences.

2. LUMINANCE MODEL

Let us suppose that f and g are respectively the images of an object acquired at the two different times k and k'. A point P of this object projects into image f in p of coordinates (x_p, y_p) and in p' of coordinates (x'_p, y'_p) into image g, after a relative motion between the camera and the scene. The luminance at p depends on the scene geometry. Fig.1 describes the vectors and the angles used in this article. $\boldsymbol{V}$ and $\boldsymbol{L}$ are respectively the viewing and the lighting vectors, which form the angles θ_r and θ_i with the normal $\mathbf{n}$ in P. $\boldsymbol{B}$ is the bisecting line between $\boldsymbol{V}$ and $\boldsymbol{L}$, it forms an angle ρ with the normal $\mathbf{n}$. According to the most widely used reflection models, such as the Torrance-Sparrow [10] and the Phong [8] ones, the luminance at p can be described as follows

$$f(p) = K_d(p)a(p)\cos\theta_i(P) + h_f(p) + K_a \qquad (1)$$

where K_a is the intensity of ambient lighting and K_d a diffuse coefficient corresponding to the direct lighting intensity. These values depend also on the gain of the camera. The term $a(p)$ is related to the *albedo*[1] in P. Function h_f expresses the contribution of the specular reflection, which vanishes in case of a pure diffuse reflection, that is for Lambertian surfaces. Consequently, with such objects and for a given lighting direction $\boldsymbol{L}$, the luminance at p is the same whatever the viewing direction $\boldsymbol{V}$ is. For non-lambertian objects, h_f reaches its

[1] The albedo is the ratio of the amount of light reflected by a small surface in P to the amount of incident light. It only depends on the material and its texture.

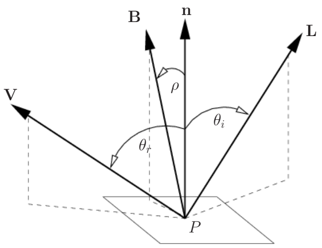

Fig. 1. Vectors and angles involved in the reflection description.

maximum value when $\rho(P) = 0$, that is when $\boldsymbol{B}$ coincides with $\mathbf{n}$ (see for example [8, 10]).

3. LOCAL PHOTOMETRIC MODEL

Now let us consider that some illumination changes expressed as functions t_a, t_d and t_i are respectively provoked on K_a, K_d and θ_i. Therefore, after a lighting change (intensity or location), the luminance g at time k' becomes

$$g(p') = K_d^{'} a(p) \cos\theta_i^{'}(P) + K_a^{'} + h_g(p') \qquad (2)$$

with $K_d^{'} = K_d + t_d$, $\theta_i^{'}(P) = \theta_i(P) + t_i(P)$ and $K_a^{'} = K_a + t_a$. According to (1) and (2), the photometric changes become

$$g(p') = \lambda(p) f(p) + \eta(p) \qquad (3)$$

where $\lambda(p)$ and $\eta(p)$ are given by

$$\lambda(p) \;=\; K_d^{'}(p)\cos(\theta_i{}'(P))/K_d(p)\cos(\theta_i(P)) \qquad (4)$$
$$\eta(p) \;=\; -(h_f(p) + K_a)\lambda(p) + h_g(\delta(p, \mathbf{A})) + K_a + t_a \qquad (5)$$

with θ_r and $\theta_i \in]-\frac{\pi}{2}\ \frac{\pi}{2}[$. Equation (3) has also been obtained in [1]. However, to our knowledge, no formal definition of λ and η had been established.

Let us assume that the surface around P is described by a $\mathcal{C}^1$ function. Consequently, when the light source is sufficiently far from the surface then the angle θ_i and the function t_i vary in $\mathcal{W}$ in a continuous way. The direct light intensity K_d is also supposed to vary smoothly in $\mathcal{W}$. In those conditions, λ is a $\mathcal{C}^1$ function in $\mathcal{W}$.

Now, let us assume that the surface projected on $\mathcal{W}$ has nearly the same roughness parameter in each point. For example, that is correct when each point of $\mathcal{W}$ is located on the same type of material. Then, by assuming a continuous surface projected on $\mathcal{W}$, the functions h_f and h_g are continuous since they depend on the normal in each point of $\mathcal{W}$.

Consequently, $\lambda(m)$ and $\eta(m)$ are continuous in each point of $\mathcal{W}$ and therefore can be expanded in Taylor series at first order around the center of $\mathcal{W}$, p. Therefore (3) becomes

$$g(m') = \mathbf{U}^T \boldsymbol{\lambda} f(m) - g(\delta(m, \boldsymbol{A})) - \mathbf{U}^T \boldsymbol{\eta} \qquad (6)$$

with $\mathbf{U} = (1, x - x_p, y - y_p)^T$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)^T$ and $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3)^T$.

4. VALIDITY OF THE MODEL

This section will focus on the study of the validity of (6). In order to simplify this problem, we consider a moving light source according to a motionless scene. This configuration enables us to study the lighting changes and the specular highlights occurrence. In this purpose we consider the geometry depicted by Fig.2. We call P a physical point of coordinates $(X_p, Y_p, Z_p)^T$ expressed in a frame $\mathcal{R}_c$ related to the camera. P projects in p, at the center of $\mathcal{W}$. Point M, of coordinates $(X, Y, Z)^T$, projects in m in the neighborhood of p in $\mathcal{W}$. In addition, Π is the tangent plane of the surface in P.

When the surface around P is described locally by a C^2 function, it can be expanded in a Taylor series around P. Then the coordinates of M can be approximated by

$$Z \;=\; Z_p + D_X(X - X_P) + D_Y(Y - Y_P) + D_{XX}(X - X_P)^2 + D_{YY}(Y - Y_P)^2 + D_{XY}(X - X_P)(Y - Y_P) \qquad (7)$$

where D_X, D_Y, D_{XX}, D_{XY} and D_{YY} are the first and second derivatives of the surface in P with respect to X and Y. The normal vector in M is expressed by $\mathbf{n} = (D_X, D_Y, -1)^T$. We call $\boldsymbol{S} = (S_x, S_y, S_z)^T$ the position of the direct light source and therefore we have $\boldsymbol{L} = (S_x - X, S_y - Y, S_z - Z)^T$ and $\boldsymbol{V} = (-X, -Y, -Z)^T$. $\boldsymbol{B}$ is the bisecting line between $\boldsymbol{V}$ and $\boldsymbol{L}$ and therefore it can be expressed with respect to (X, Y, Z). In the same way, the angles θ_i, θ_r, ρ, which depend on $\mathbf{n}$, $\boldsymbol{V}$, $\boldsymbol{L}$ and $\boldsymbol{B}$, can also be expressed with respect to X, Y, Z. Consequently, $\lambda(m)$ and $\eta(m)$ can be expressed with respect to X, Y, Z. Then, we assume that both the camera and the scene are motionless between k and k', and that the direct light source undergoes a small motion $\boldsymbol{d}S = (dS_X, dS_Y, dS_Z)^T$ from its initial location $\boldsymbol{S}$. After a perspective projection of M to point m of coordinates (x, y), $\lambda(m)$ and $\eta(m)$ are expanded in Taylor series around p (center of $\mathcal{W}$). At second order, we obtain

$$\lambda(m) \;=\; \lambda_1 + \lambda_2 x + \lambda_3 y + \lambda_4 x^2 + \lambda_5 y^2 + \lambda_6 xy \qquad (8)$$
$$\eta(m) \;=\; \eta_1 + \eta_2 x + \eta_3 y + \eta_4 x^2 + \eta_5 y^2 + \eta_6 xy \qquad (9)$$

The model (6) assumes that $\lambda_4, \lambda_5, \lambda_6, \eta_4, \eta_5, \eta_6$ are null in (8) and (9), but what does it mean in a geometrical and physical point of view? To answer this question, we shall determine the configurations of acquisition for which the model is the most adapted. Therefore, λ_i and η_i for $i = 4..6$ are expanded in Taylor series at first order around $\boldsymbol{d}S \simeq 0$. We call $\widehat{\lambda}_i$ and $\widehat{\eta}_i$ these approximations. We only analyze the expansion of λ_4 and η_4, since almost the same conclusions arise from λ_5, λ_6 and η_5, η_6.

• **Validity of $\boldsymbol{\lambda}$.** Let us consider three interesting configurations.

1) When the lighting and the normal vectors are convergent $\boldsymbol{L} = \alpha\mathbf{n}$ it yields to $\widehat{\lambda}_4 = 0$. Thus, a small motion of the light source does not infer on the coefficients of second order

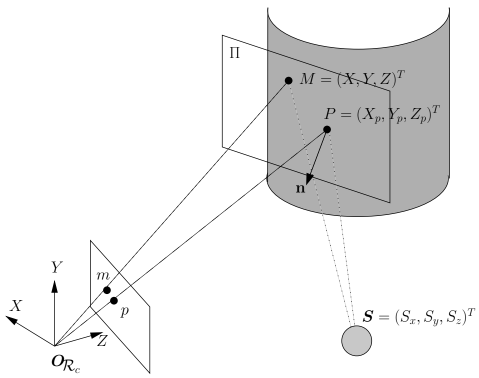

Fig. 2. Model of the scene geometry.

of (8).

2) When the light source is close to the camera $\boldsymbol{S} = \boldsymbol{O}$, we obtain $\widehat{\lambda}_4 = 2(D_Y D_{XX} + D_{XY} D_X)dS_Y - (4D_{XX} + \frac{2}{Z_p})dS_Z + \left(6D_X D_{XX} - \frac{2D_X}{Z_p}\right) dS_X$. We can see from this relationship that $\lambda_4 \simeq 0$ when the surface is nearly planar ($D_{XX} \simeq 0$, $D_{XY} \simeq 0$) and the camera is far from the surface (high Z_P).

3) When the light source $\boldsymbol{S}$ is located at a very small distance ϵ with respect to P, the coefficients of $\widehat{\lambda}_4$ can be expanded in Taylor series around ϵ, and $\widehat{\lambda}_4$ becomes

$$\widehat{\lambda}_4 = \frac{2Z_p(D_X(\epsilon-1)+D_{XX}(2\epsilon^2+Z_p\epsilon))}{\epsilon^3} dS_X + \frac{2Z_p(D_{XX}Z_pD_Y+D_{XY}D_X\epsilon)}{\epsilon^2} dS_Y - \frac{2Z_p(Z_p+D_X^2\epsilon-Z_pD_X^2+2Z_pD_{XX}\epsilon)}{\epsilon^3} dS_Z$$

Since the small distance ϵ appears at the denominator, this is an unfavourable configuration to use the photometric model.

• **Validity of $\boldsymbol{\eta}$.** In order to simplify the analysis, we consider $\lambda(m) = 1$ in (5), such that only specular highlights are supposed to be caused. Because h_f and h_g reach their maximum value when $\rho = 0$, we study the variations of $\boldsymbol{\eta}$ around this specific configuration, by using the Phong model. Therefore, the initial location of the light source $\boldsymbol{S}$ is chosen so as $\boldsymbol{B}$ coincide with $\mathbf{n}$ ($\rho = 0$). We note d the distance between P and $\boldsymbol{S}$. Two configurations are studied:

1) First, when $\boldsymbol{L} = \mathbf{n} = \boldsymbol{V}$, the tangent plane Π is parallel to the sensor plane (D_X and $D_Y = 0$) and $\theta_i = 0$. Since the relations obtained in that case are too complicated to deduce any conclusion, we consider two further constraints. Firstly, if the light source is initially close to the surface, then d is small and $\widehat{\eta}_4$ can be expanded in Taylor series at first order around $d = 0$ so that $\widehat{\eta}_4 = -n\left(2D_{XX} + \frac{1}{Z_p}\right) dSZ$.

According to this relationship, the modeling error is low when the surfaces are rough (n low) and almost planar (low D_{XX}).

Secondly, if the camera is near the surface (low Z_P) then $\widehat{\eta}_4 \simeq 0$. That is a favourable condition to use (6).

2) Now, let us consider that the orientation of the tangent plane Π with respect to the sensor plane is low (low D_X and D_Y). The higher this orientation is, the larger the angle θ_i is. In that case, we expand each term of $\widehat{\eta}_4$ in Taylor series around D_X and D_Y. Because of the complexity of the relations obtained, the surface is assumed to be planar ($D_{XX} = 0$ and $D_{XY} = 0$) and two configurations are considered. Firstly, the camera is supposed to be located near the surface (d small), then

$$\widehat{\eta}_4 = -\frac{nD_X}{4Z_p}(3n+7)\,dS_X - \frac{nD_X}{4Z_p}(n+1)\,dsY - \frac{n}{Z_p}dS_Z$$

According to this expression, the lower the orientation of Π is (low D_X and D_Y), the more (6) is valid.

Secondly, when the camera is close to the surface (Z_p small), we obtain $\widehat{\eta}_4 \simeq 0$.

• **Conclusion.** Consequently, the approximations of λ and η by a Taylor series at first order are quite realistic (the terms of second order of λ and η are almost null) when:
− the orientation of the tangent plane Π of the surface in P is low with respect to the sensor plane;
− the second derivatives D_{XX}, D_{YY} and D_{XY} of the surface are low;
− $\boldsymbol{L}$ coincide with $\mathbf{n}$;
− the camera is close to the surface (low Z_P) and the lighting source is far from the surface;
− the surface is rough (low n).

Let us notice that the affine photometric model [6, 9] is far more restrictive than (6) because $\lambda_2, \lambda_3, \eta_2$ and η_3 are supposed to be null simultaneously in equations (8) and (9).

5. APPLICATION TO FEATURE POINTS TRACKING

We call δ the motion model of a small window of interest $\mathcal{W}$ centered on p. We assume that this motion is parametrized by a vector $\boldsymbol{A}$ so that $p' = \delta(p, \boldsymbol{A})$. The point tracking method consists in computing $\boldsymbol{A}$ and the photometric parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ (see equation (6)) by minimizing the following criterion

$$\epsilon(\boldsymbol{A}, \boldsymbol{\lambda}, \boldsymbol{\eta}) = \sum_{m \in W} \left(\mathbf{U}^T \boldsymbol{\lambda} f(m) - g(\delta(m, \boldsymbol{A})) - \mathbf{U}^T \boldsymbol{\eta}\right)^2 \quad (10)$$

In the following experiments, we choose an affine motion model, which is computed between the first frame and the current one. This technique, that we call M1, is compared on two image sequences (see Fig. 3a and 3b) with the method proposed by Jin and Soatto, based on the computation of an affine photometric model [6]. We call M2 this latter method.

In each case, the objects are motionless and the camera moves. The first sequence shows a book, the cover of which is made of glossy paper. Consequently, the motion of the camera induces some specular highlights variations. No lighting change is caused. This sequence is played from the first frame to the final one and then from the last one to the first one. The second sequence shows a specular object covered by a glossy paper. It is lighted by the daylight and a direct lighting, the intensity of which varies periodically from a low value to a high one each 20 iterations approximately. The points are selected by the Harris detector [4] (52 points in the first sequence and 25 in the second one).

The tracking algorithm integrates an outlier rejection step, based on the analysis of the convergence of residues ϵ: a point is rejected as soon as its residues become greater than a threshold (here, a mean luminance variation of 15 is tolerated for each pixel of $\mathcal{W}$). The size of the window is 15×15. The use of smaller windows can lead to a poor accuracy of the photometric and geometric parameters involved in M1. For small windows, it is more suitable to use the method proposed in [3] which uses only 3 parameters. This latter approach is not used in this article since it is less efficient for large windows. In the first sequence, 18 points are lost by M1 and 34 by M2 (almost twice more points) because of the specular highlights variations. In the second one, only 1 point is lost by M1 and 4 by M2, because of the severe lighting changes. In order to compare the accuracy of the geometric and photometric modeling, Fig. 3c and 3d show the evolution of the residuals (the average of the residuals computed on the points that are correctly tracked by M1 and M2) obtained by each technique with respect to the iterations, respectively for the first and the second image sequences. They are in adequacy with the geometric and lighting changes. Indeed, in Fig. 3c the residues are symetric. In Fig. 3d, the residues vary from a low value to a high one each 20 iterations. In each case, the residuals obtained by M1 are lower than those obtained by M2, which proves that the photometric model is always more adapted to the real illumination changes. In particular, M1 compensates better for specular highlights, as it is shown by the number of points correctly tracked in the first sequence and the residuals obtained. When convergence residuals are low, it is very probable that the accuracy of the points location is good. In addition, the computation times are not significantly increased (5.2 ms for method M1 and 4ms for method M2, on the first sequence).

6. CONCLUSION

This article provided a theoritical explanation of luminance changes and proposed a local photometric model, based on the study on widely used reflection models. We analyzed the validity of this model by considering a locally continuous surface viewed under some borderline configurations of the scene geometry. This model can be used in various computer vision applications based on the computation of correspondences between frames. In this article, it is introduced in a points tracking technique. Compensating the illumination changes during the image sequence improves the robustness of the tracking process when specular highlights and lighting changes occur. Indeed, a larger number of points is correctly tracked and the residuals of the method are low. However it could be interesting to study the accuracy of the points location during the tracking process. That will be part of our future works.

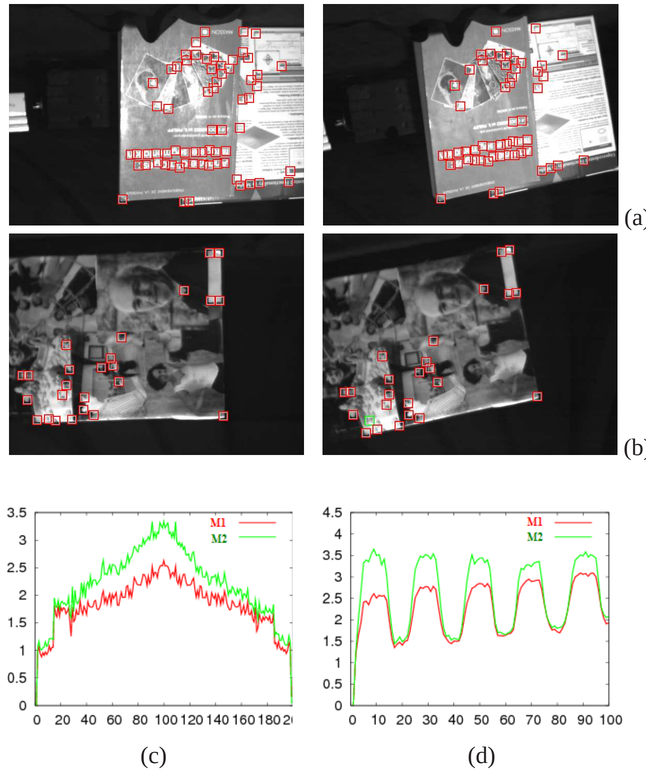

Fig. 3. (a) and (b): two frames of the image sequences. (c) and (d): Comparison of the residuals obtained with M1 and M2: on the first sequence (c) and the second one (d).

7. REFERENCES

[1] M.J. Black, D.J. Fleet, and Y.Yacoob. Robustly estimating changes in image appearance. *Computer Vision and Image Understanding*, 78:8 – 31, 2000.

[2] T. Gevers and A.W.M. Smeulders. Object recognition based on photometric color invariants. In *10th SCIA*, Lappeenranta, Finland, 1997.

[3] M. Gouiffès and C. Collewet and C. Fernandez and A. Trmeau. Feature points tracking : robustness to specular highlights and lighting changes. In *9th European Conference on Computer Vision*, Graz, Austria, 2006.

[4] C.G. Harris and M. Stephens. A combined corner and edge detector. In *4th Alvey Vision Conference*, pages 147–151, 1988.

[5] K.P. Horn and B. G. Schunck. Determinig optical flow. *Artificial Intelligence*, 7:185–203, 1981.

[6] H. Jin, P. Favaro, and S. Soatto. Real-time feature tracking and outlier rejection with changes in illumination. In *IEEE International Conference on Computer Vision*, pages 684–689, Vancouver, Canada, July 9-12, 2001.

[7] S. Negahdaripour. Revised definition of optical flow:integration of radiometric and geometric cues for dynamic scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(9):961–979, September 1998.

[8] B-T Phong. Illumination for computer generated images. *Communications of the ACM*, 18(6):311–317, June 1975.

[9] T. Tommasini, A. Fusiello, E. Trucco, and V. Roberto. Improving feature tracking with robust statistics. *Pattern Analysis & Applications*, 2(4):312–320, 1999.

[10] K.E. Torrance and E.M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9), September 1967.