TRACKING BY COMBINING PHOTOMETRIC NORMALIZATION AND COLOR INVARIANTS ACCORDING TO THEIR RELEVANCE.

Michèle Gouiffès

michele.gouiffes@ief.u-psud.fr, IEF, UMR 8622, Paris XI, Bat 220, 91405 Orsay cedex

ABSTRACT

This paper addresses the problem of robust feature points tracking by using specific color invariants –robust to specular reflections, lighting changes and to some extent to color lighting changes– when they are relevant and photometric normalization in the opposite case. Indeed, most color invariants become noisy or irrelevant for low saturation and/or low intensity. They can even make tracking fail. Combining them with luminance information yields to a more performant tracking, whatever the lighting conditions are. A few experiments on real image sequences prove the efficiency of this procedure.

***Index Terms*—** Tracking, color invariants, photometric changes.

1. INTRODUCTION

In many computer vision applications, such as segmentation, indexing or tracking, the use of color proves to be relevant and decisive. Contrary to luminance, where no comprehensive compensation of illumination changes has been defined as far now, literature provides some color invariants and color constancy [1, 2, 3], which are robust to intensity of lighting, to the scene geometry, shadings, and/or to occurrence of specular reflection, according to the assumptions made on the scene and on the acquisition conditions.

However, such invariants are irrelevant and noise sensitive when saturation and/or intensity of colors is low, for example when lighting intensity hugely lowers during an image sequence. In this condition, the only use of luminance is as relevant as color and is less-time consuming. Then, the question is how to switch properly between these two attributes.

This article addresses this issue in the context of differential feature points tracking, as defined initially by Kanade Lucas and Tomasi[4, 5]. The proposed procedure combines luminance and color invariants, according to the relevance of color.

On the one hand, when color is relevant, the tracking is carried out by using attributes that are to some extent invariant to changes of illuminant color, gain of the camera, acquisition geometry, and variations of specular reflection.

On the other hand, when color is irrelevant, the luminance information is used. In order to improve the robustness of the tracking, a photometric normalization, as in [6] for example, or local photometric models, [7, 8] can answer partially the problem of illumination changes. In this paper, we use a photometric normalization which is, to some extent robust to specular reflection, lighting changes, and to the gain of the camera.

This article is structured as follows. Section 2 focuses on the color and luminance attributes, and defines the relevance functions. Then, the tracking procedure is detailed in section 3. To finish, a few experimental results are reported in section 4.

2. COLOR AND LUMINANCE ATTRIBUTES

2.1. Color attributes

Gevers [1] has proposed a panel of color invariants. Most of them are only available under white illuminant, either for lambertian surfaces ($c_1c_2c_3$ and normalized RGB components) or specular surfaces ($l_1l_2l_3$, the hue). On the other hand, the $m_1m_2m_3$ components are appropriate under color illuminant but for lambertian objects only. They have been used with success in [9] to improve the stability of points detection. However, the invariance problem for specular surfaces viewed under illuminant of varying color has no simple solution, but often requires costly constancy algorithms [2].

As defined by the dichromatic model [10], the color vector $\boldsymbol{C} = (R, G, B)$ in a point p, which is the projection of a physical point P of a specular surface, can be simplified as

$$\boldsymbol{C}(p) = m_b(P)\boldsymbol{C_b}(p) + m_s(P)\boldsymbol{C_s}(p) \tag{1}$$

where $\boldsymbol{C}_b = (C_b^R, C_b^G, C_b^B)$ and $\boldsymbol{C}_s = (C_s^R, C_s^G, C_s^B)$ are respectively the colors of body reflection and specular reflection. The two terms m_b and m_s depend on the scene geometry (lighting angle, viewing angle, orientation of surface). As an example, for the red channel, ${C_b}^R$ and ${C_b}^R$ are written:

$${C_b}^R(p){=}\int_\lambda S_R(\lambda)\mathcal{I}(\lambda, P)\mathcal{R}_b(\lambda, P)d\lambda,\ C_s^R(\lambda, p) = \int_\lambda S_R(\lambda)\mathcal{I}(\lambda, P)d\lambda \tag{2}$$

where $S_R(\lambda)$ is the camera sensitivity which is a function of the wavelength λ, $\mathcal{I}$ is the illuminant spectrum and $\mathcal{R}_b$ the reflectance of the material. Let us assume that $S_R(\lambda)$ is band-limited around the red wavelength λ_R and can be approximated by a constant S_R with respect to λ. By considering this assumption on each sensor, it yields to the following simplification of (1):

$$\boldsymbol{C}(p) = \boldsymbol{a}(p).\boldsymbol{C}_I(p)m_b(p) + \boldsymbol{C}_I(p)m_s(p) \quad (3)$$

with . the Hadamard product. $\boldsymbol{C}_I = (I_R, I_G, I_B)$ is defined as the color of the illuminant and depends on the gains (S_R, S_G, S_B) of the camera. To finish, $\boldsymbol{a}(p) = (a_R, a_G, a_B)$ depends on the reflectance and is supposed to be constant during the time.

In order to define an invariant attribute against lighting color gain changes, we assume that the geometry m_b, m_s and the lighting color vary slowly in the image. In such a context, we assume that the lighting conditions are the same in a 3 × 3 neighborhood $\mathcal{V}_s$ of the point p with $s \neq p$. Consequently, it is easy to show, owing to (3), that any ratio $(R(p) - R(s_0))/(R(s_1) - R(s_2))$ with s_0, s_1 and $s_2 \in \mathcal{V}_s(p)$ does not depend neither on the lighting conditions nor on the gain of the camera but only on the constant parameters (a_R, a_G, a_B). In order to reduce the sensivity to noise, we use the following color attribute:

$$\mathcal{C}_R(p) = [R(p) - \min_{s \in \mathcal{V}_s} R(s)]/\sigma_R(s) \quad (4)$$

The similar transformation is made on G and B. According to the assumptions made on the bandwidth limitation of the camera sensitivity and on the smoothness of illuminant changes in a 3 × 3 neighborhood, $\boldsymbol{C}_{inv}(p) = (\mathcal{C}_R, \mathcal{C}_G, \mathcal{C}_B)$ can only be viewed as a color quasi-invariant available for non-lambertian surfaces viewed under color illuminant.

2.2. Photometric normalization

When saturation of colors is low or when $\boldsymbol{C}_{inv}$ is noisy because of low values of σ in (4), the use of luminance information I can become more appropriate and less time consuming. In order to improve the robustness of the tracking, we consider the following photometric normalization, which has been experimented in [6]. It is a satisfying trade-off between robustness and implementation complexity:

$$[I(p) - \mu_I]/\sigma_I \quad (5)$$

where μ_I and σ_I are respectively the average and standard deviation of I computed in a small window of interest $\mathcal{W}$ centered around p.

2.3. Relevance functions

Relevance of color invariants α. The attributes described by (4) are noise sensitive when standard deviation at denominator is low. In order to answer this issue, we define a relevance function $\alpha_i(p)$ for each component $\mathcal{C}_i$, $i = R, G, B$ by a sigmoid function $\alpha_i(p) = 1/(1+\exp(-(\sigma_i(p)-\sigma_0)u_\alpha))$ where σ is the standard deviation, σ_0 the abscissa of the inflection point and u_α depends on the slope. $\alpha(p)$ is used as a ponderation function of the invariant, so that $\mathcal{D}_i(p) = \alpha_i(p)\mathcal{C}_i(p)$ for $i = R, G, B$ are finally the color attributes used in the tracking procedure: $\boldsymbol{D} = (\mathcal{D}_R \mathcal{D}_G \mathcal{D}_B)$. In that way, relevant invariant values are preserved, whereas noise is reduced.

Relevance of color versus luminance β. As it has been said in introduction, color is relevant only when saturation is significant. Therefore, we define in each point a relevance function $\beta(p)$ as a sigmoid, which is a function depending on the saturation in p: $S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)}$. We call S_0 its inflection point and u_β its slope parameter.

3. TRACKING PROCEDURE

Let us call k and k' two successive times of the image sequence and $\boldsymbol{D}_k$, $\boldsymbol{D}_{k'}$, I_k, $I_{k'}$ the images obtained at these two times. A point P of an object projects into image in p, of coordinates (x, y) in frame k and in p' in frame k', after a relative motion between the camera and the scene. We call δ the motion model of a small window of interest $\mathcal{W}$ centered around the point to be tracked p and we assume that this motion is parameterized by a vector $\boldsymbol{A}$, so that the position of point P in k' $p' = \delta(p, \boldsymbol{A})$. We call q a neighbor of p in $\mathcal{W}$. The tracking procedure consists in computing $\boldsymbol{A}$ by minimization of the following criterion

$$\epsilon(\boldsymbol{A}) = \sum_{q \in \mathcal{W}} (\gamma(q, q') \|\boldsymbol{D}_k(q) - \boldsymbol{D}_{k'}(\delta(q, \boldsymbol{A}))\| + (1 - \gamma(q, q'))\|aI_k(q) - I_{k'}(\delta(q, \boldsymbol{A}) + b)\|)^2 \quad (6)$$

where $\gamma(p, p') = [\beta_k(p)\beta_{k'}(p') \max(\alpha_R, \alpha_G, \alpha_B)]^{1/3}$ is the function used to ponderate the relevance of $\boldsymbol{D}_k$ in the computation of motion.

Let us assume that $\boldsymbol{A} = \widehat{\boldsymbol{A}} + \Delta\boldsymbol{A}$ where $\boldsymbol{a}$ refers to a small variation around an estimate $\widehat{\boldsymbol{A}}$ of $\boldsymbol{A}$. We achieve a Taylor expansion of $\boldsymbol{D}_{k'}(\delta(q, \boldsymbol{A}))$ and $I_{k'}(\delta(q, \boldsymbol{A}))$ at first order around $\widehat{\boldsymbol{A}}$. By neglecting the coefficients of second order, we obtain:

$$\boldsymbol{D}_{k'}(\delta(q, \boldsymbol{A})) = \boldsymbol{D}_{k'}(\delta(q, \widehat{\boldsymbol{A}})) + \boldsymbol{G_D}(\delta(q, \widehat{\boldsymbol{A}}))J_\delta^{\widehat{\boldsymbol{A}}} \Delta\boldsymbol{A} \quad (7)$$

$$I_{k'}(\delta(q, \boldsymbol{A})) = I_{k'}(\delta(q, \widehat{\boldsymbol{A}})) + \boldsymbol{G_I}(\delta(q, \widehat{\boldsymbol{A}}))J_\delta^{\widehat{\boldsymbol{A}}} \Delta\boldsymbol{A} \quad (8)$$

where $\boldsymbol{G_D}$ and $\boldsymbol{G_I}$ are respectively the Jacobian matrices of $\boldsymbol{D}_{k'}$ and $I_{k'}$ with respect to x and y. By injecting (7) and (8) in (6), it finally yields to a linearized system in $\boldsymbol{a}$:

$$\left(\sum_{q \in \mathcal{W}} \boldsymbol{V}_C \, \boldsymbol{V}_C{}^T \right) \Delta\boldsymbol{A} =$$

$$\sum_{q \in \mathcal{W}} \left[\frac{\alpha(q, q')}{3} \left(\sum_{i \in RGB} \left(D_k{}^i(q) - D_{k'}^i(\delta(q, \widehat{\boldsymbol{A}}) \right) \boldsymbol{V}_i \right) + (1 - \alpha(q, q')) \left(I_k(q) - I_{k'}(\delta(q, \widehat{\boldsymbol{A}})) \right) \boldsymbol{V}_I \right]$$

In the case of an affine motion model, $\boldsymbol{V}_i$ for $i = R, G, B, I$ is expressed by:

$$\boldsymbol{V}_i = \left[g_x{}^i, g_y{}^i, xg_x{}^i, xg_y{}^i, yg_x{}^i, yg_y{}^i \right]^T \quad (9)$$

where $g_x{}^i$ and $g_y{}^i$ are the first derivatives of i= R, G, B, I in q. To finish, the vector $\boldsymbol{V}_C$ is expressed as $[g_x, g_y, xg_x, xg_y, yg_x, yg_y]^T$ with $g_x = \left[\sum_i (\gamma/3) g_x^i\right] + (1 - \gamma) g_x^I$ and $g_y = \left[\sum_i (\gamma/3) g_y^i\right] + (1 - \gamma) g_y^I$ for $i = R, G, B$.

4. EXPERIMENT RESULTS

Since the proposed procedure has to track points either by using color or luminance information, the points are detected (with a window $\mathcal{W}$ of size 15 × 15) in I by a Harris detector [11], and in $\boldsymbol{D}$ by its color extension [12]. Of course, points which are selected twice are considered only once. We use an affine motion model, which is computed between the previous frame k and the current one k'. A point is rejected as soon as its convergence residues become greater than a threshold (a mean luminance variation of 15 is tolerated for each pixel of $\mathcal{W}$). The parameters used for the α (relevance of invariants) are $\sigma_0 = 30$ and $u_\alpha = 0.2$ and the parameters used for β (relevance of color) are: $S_0 = 0.1$, $u_\beta = 100$.

The sequences are played from the first frame to the last one and then from the last one to the first one in order to check whether the points have correctly come back to their initial location. The error is then obtained by the euclidian distance between the initial location and the final one.

Fig. $1(a)$ shows a few images of the sequence *Card Game* from the ALOI [1] image data base [13]. First of all, the points detected are drawn in yellow circles in Fig.$1(b)$ for luminance image and in Fig. $1(c)$ for the $\boldsymbol{D}$ image. Obviously, color provides relevant information since a larger number of points is detected by the use of color (50 points) than by only using luminance (22 points).

The first image sequence one corresponds to images from l2c1.png to r100.png in ALOI base, that is to say when temperature of illuminant is varying; a few images deal with object orientation changes. The second sequence deals with illumination directions and intensity changes (images i110.png to l1c3.png in ALOI). Fig.2 deals with the tracking results obtained on the first sequence. Fig. $2(a)$ to $2(c)$ draw yellow circles for the points that have been correctly tracked and red circles for the points that have drifted. They refer respectively to results for photometric normalization (PN), $\boldsymbol{D}$ and the combination PN + $\boldsymbol{D}$. Whereas only one point has been correctly tracked either by PN or $\boldsymbol{D}$, the combination of the two attributes (PN+$\boldsymbol{D}$) correctly trackes 9 points.

Fig.3 shows the results (for frames 1, 15, 31) obtained when different lighting directions are considered. Fig.$3(a)$ deals with PN whereas figure $3(b)$ refers to the combination PN+$\boldsymbol{D}$. The results obtained with the only use of $\boldsymbol{D}$ is not shown since no point has been correctly tracked until the end of the sequence. 50 points are selected in luminance (see Fig .$3(a)1$) and 60 by using both luminance and $\boldsymbol{D}$ (see Fig .$3(b)1$). 13 points have been correctly tracked by PN and 18 points by PN+$\boldsymbol{D}$, which confirms the relevance of this technique to improve the robustness of tracking when illumination changes are caused.

To finish, let us consider the *road* sequence[2] (see Fig. $4(a)$) which is blurred, noisy and where and gain changes occur. The tracking of the road sign has succeed with the three techniques for a window size 61×61. However, the tracking errors are different: 4.87 pixels with PN, 16.72 pixels with $\boldsymbol{D}$, and only 2.60 pixels with PN+$\boldsymbol{D}$. As a conclusion, PN+$\boldsymbol{D}$ has improved the accuracy of the tracking.

5. CONCLUSION

This article has exposed a feature points tracking procedure that uses both color pseudo-invariants and photometric normalization. When saturation of color is sufficiently high, the cost function involved is computed on color. In the opposite case, photometric normalization is considered. This experimental results have shown that the combination color-luminance is more robust than the only use of color or photometric normalization, since a large number of points is correctly tracked, for exemple when temperature and/or direction of illuminant vary. Furthermore, it has proved to be more accurate. Future works will deal with the optimization of the procedure in order to integrate more comprehensive photometric models in this procedure [8].

[1]available on http://staff.science.uva.nl/ aloi/, images of size 768 × 576

[2]available on http://vasc.ri.cmu.edu/idb/html/jisct/index.html.

6. REFERENCES

[1] T. Gevers and A.W.M. Smeulders, "Color-based object recognition," *Pattern Recognition*, vol. 32, no. 1999, pp. 453–464, 1999.

[2] G.D. Finlayson, S. D. Hordley, G. Schaefer, and Gui Yun Tian, "Illuminant and device invariant colour using histogram equalisation," vol. 38, no. 2, pp. 179–190, 2005.

[3] J-M. Geusebroek, R. Van den Boomgaard, A. W. M. Smeulders, and A. W. M. Geerts, "Color invariance," *IEEE Transactions on PAMI*, vol. 23, no. 12, pp. 1338–1350, 2001.

[4] B.D. Lucas and T. Kanade, "An iterative image registration technique," in *Internation Joint Conference on Artificial Intelligence*, 1981, pp. 674–679.

[5] C. Tomasi and T. Kanade, "Detection and tracking of point features," Technical report cmu-cs -91-132, Carnegie Mellon University, 1991.

[6] T. Tommasini, A. Fusiello, E. Trucco, and V. Roberto, "Improving feature tracking with robust statistics," *Pattern Analysis & Applications*, vol. 2, no. 4, pp. 312–320, 1999.

[7] 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*, Canada, 2001, pp. 684–689.

[8] M. Gouiffès, C. Collewet, C. Fernandez-Maloigne, and A. Trémeau, "A photometric model for specular highlights and lighting changes. application to feature points tracking.," in *IEEE Int. Conf. on Image Processing, ICIP'2006*, 2006.

[9] F. Faille, "Stable interest point detection under illumination changes using colour invariants.," in *British Machine Vision Conference*, 2005.

[10] S.A. Shafer, "Using color to separate reflection components," *Color Research and Applications*, vol. 10(4), pp. 210–218, 1985.

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

[12] V. Gouet, P. Montesinos, R. Deriche, and D. Pel, "Evaluation de dtecteurs de points d'intrt pour la couleur," in *Reconnaissance de Formes et Intelligence Artificielle*, Paris, 2000, pp. 257–266.

[13] J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, "The Amsterdam library of object images," *Int. J. Comput. Vision*, vol. 61, no. 1, pp. 103–112, 2005.

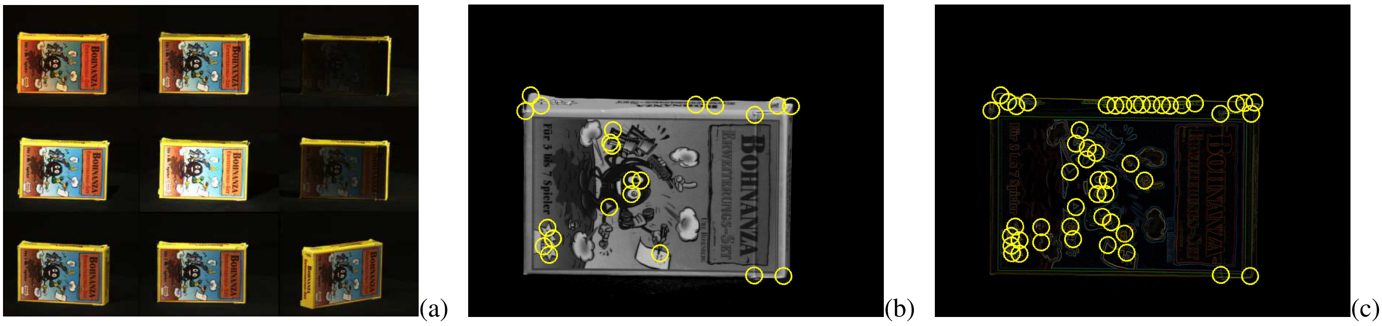
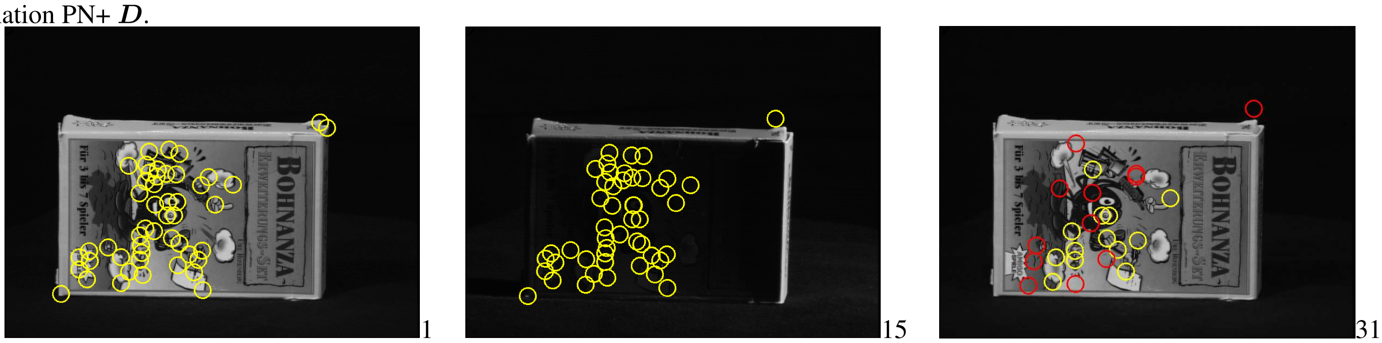
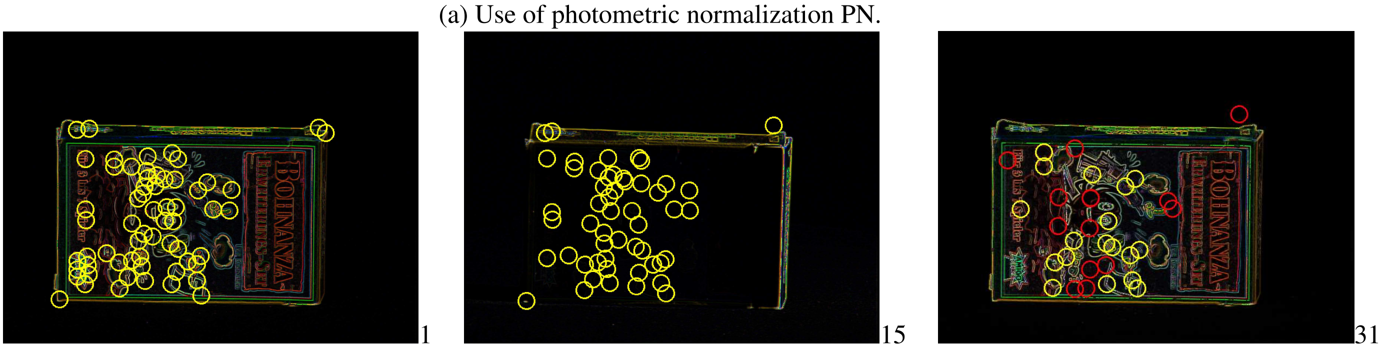


Fig. 1. (a) A few images of the sequence. (b) Points detected in luminance image and (c) in $\boldsymbol{D}$ image.

Fig. 2. Tracking results for color illumination changes. (a) Use of photometric normalization (PN). (b) Use of $\boldsymbol{D}$. (c) Combination PN+ $\boldsymbol{D}$.

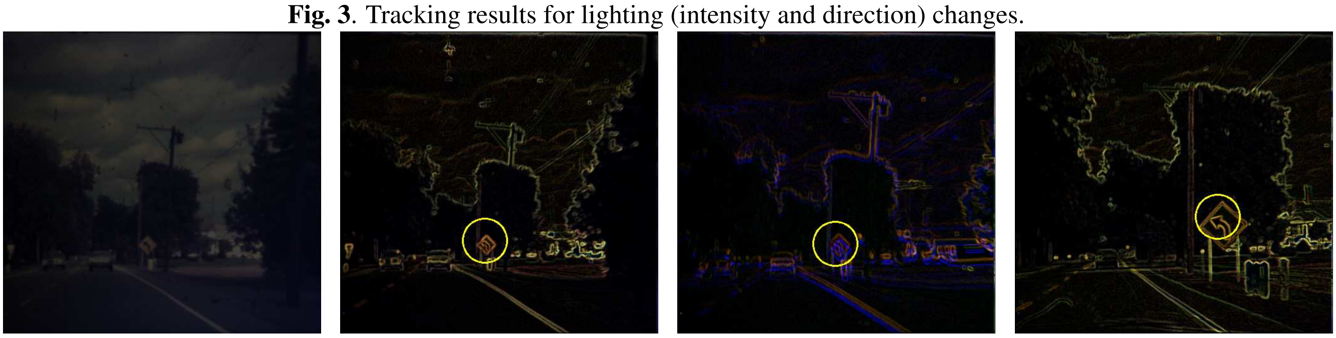


(a) Use of photometric normalization PN.



(b) Joint use of photometric normalization and color attributes PN+ $\boldsymbol{D}$

Fig. 3. Tracking results for lighting (intensity and direction) changes.



(a) RGB | 1^{th} frame | 12^{th} frame | 54^{th} frame

Fig. 4. Tracking of the road sign on *road* sequence by combining photometric normalization and color invariant PN+$\boldsymbol{D}$.