Optical Flow Invariants in Models of Driver's Visual Perception

Valeriy Eremin, Maxim Koltyapin
Department of Informational Security
Moscow State Industrial University, Moscow, Russia
eremin_valerii@mail.ru, koltyapin@atgs.ru

Abstract

According to J. Gibson's theory of visual perception (Gibson's ecological approach), human beings use optic flow invariants to estimate the environment structure and self motion in it [1]. In author's works the correspondence of the optic flow invariants to the instant velocities field invariants (divergence, curl and deformation) was shown as a result [2][3]. This work extends the suggested approach by involving the real time computer evaluation of these invariants in models of driver's visual perception.

Keywords: *Ecological approach to visual perception, Optic flow invariants.*

1. INTRODUCTION

The human beings obtain the major part of the environment information by means of the sight. It is important information about space-time characteristics of the environment: extension, form and arrangement of environment elements, position and velocity of its objects.

The great modern psychologist J. Gibson has suggested so called ecological approach to visual perception. According to this approach, reflected and ambient light serves as a basis for vision. Thus, reflected light forms structured optic array, i.e. beam of light with an apex at the point of observation, in other words two dimensional set of viewing directions. During observer's motion, an optic array continuously changes, producing an optic (motion) flow or instant local velocities field on observer's eye retina.

The analysis of existing models of visual perception shows that such ideas usually implicitly supposed, but rarely formulated.

2. PROBLEM FORMULATION

The work [2] contains the formalization of ecological approach and the application of the field theory to it. Following coordinate systems were introduced for description of time-dependent optic flow produced by observer's motion (Figure 1):

- OXYZ – observer connected coordinate system (let's assume it placed at the eye central point – center of projection);
- O_1pq – observer's eye retinal coordinate system;
- $O_2\mu\nu\lambda$ – fixed world (scene) coordinate system, connected with the ground surface.

The image (projection) g of the texture element G moves along the retina with the following velocity:

$$D_p = \frac{dp}{dt}, \quad D_q = \frac{dq}{dt} \tag{1}$$

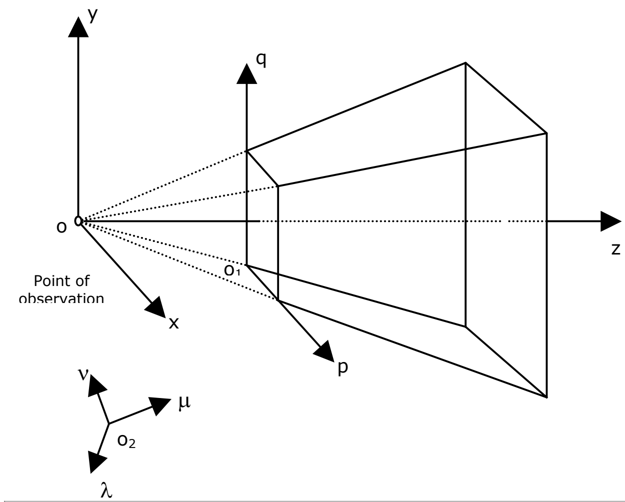

Figure 1: Representation of optic array as a frustum.

A whole set of such retinal images forms the optic flow and motion (instant local velocities) field. Divergence, curl and deformation of this field are invariants, independent of the rotation of coordinate system:

$$div\vec{D} = \frac{\partial D_p}{\partial p} + \frac{\partial D_q}{\partial q}, \tag{2}$$

$$curl\vec{D} = \frac{\partial D_q}{\partial p} - \frac{\partial D_p}{\partial q}, \tag{3}$$

$$def\vec{D} = \sqrt{\left(\frac{\partial D_p}{\partial p} - \frac{\partial D_q}{\partial q}\right)^2 + \left(\frac{\partial D_q}{\partial p} + \frac{\partial D_p}{\partial q}\right)^2}. \tag{4}$$

Divergence field describes isotropic expansion or compression, curl field describes the rotation of small parts of the optic flow and deformation field describes compression in some direction and expansion in the orthogonal direction.

Psychophysiological experiments have proved that human being's visual system has neuronal channels sensitive to these invariants, which are used to extract self motion information and the environment structure information from optic flow pattern [4].

Thus, the invariants may serve as the criterions of human psychophysiological motion preference estimation. Until now, expert judgments usually used for psychophysiological motion pictures influence estimation. Such judgments are often subjective, conflicting and incomparable.

Until recent times, there were no successful tries of the real time computer evaluation of optic flow invariants. The goal of this work is the real time computer evaluation of the motion field and its invariants in model of a road 3D environment.

3. PROBLEM SOLUTION

The above mentioned invariants expressions are taken in suggestion of space and time continuity. But existing hardware and 3D visualization algorithms work with discrete pixels and time slices.

Let's take following admissions for the transition from infinitely small pieces of space and time to pixels and time slices:

- an infinitely small area of the human eye retina corresponds to a pixel – irreducible color element created by an output device on its display surface;
- an infinitely small time slice corresponds to a time slice equal to or less than the time interval between two consequent frames in cinematograph

The human eye retina is approximated by the surface (plane) of some output device.

Let the arbitrary scene be central projected on a display surface at some instant t of time (Figure 1). Let (p', q') be coordinates of projection g of some texture element G having scene space coordinates (μ, ν, λ).

Let the position of center of projection and/or the orientation of optic array be changed at the instant $t+\Delta t$ relative to their values at t. At the same time, the texture element projection may be moved relative to its original position on display surface.

Let (p'', q'') be coordinates of projection g at instant $t+\Delta t$. Then the modulus of vector (p''-p', q''-q') is a distance covered by image g for Δt time period. So vector $\vec{D}$ is an instant velocity of image g at instant t:

$$\vec{D} = \left(\frac{p'' - p'}{\Delta t}, \quad \frac{q'' - q'}{\Delta t} \right) \qquad (5)$$

A set of such vectors throughout the display surface forms the velocity field sought.

In order to evaluate vector $\vec{D}$ it's necessary to know the correspondence of texture element G to its projections at instants t and $t+\Delta t$. The correspondence of G coordinates (μ, ν, λ) and its projection at instant t is determined by matrix A' transforming scene coordinates to display surface coordinates. The inverse matrix $(A')^{-1}$ determines the inverse transformation. However, such transformation is ambiguous, until we know only two coordinate components (p', q'). It's possible to get only a beam crossing the display surface at the point (p', q'), with a beginning at point of observation. It is known, that world coordinates of the texture element are located at the point of the intersection of the beam with the closest opaque scene object. These coordinates can be found by tracing the beam into the scene. Knowing three coordinate components of the projection and the matrix of inverse transformation at instant t it is possible to calculate world coordinates of the texture element:

$$(\mu, \nu, \lambda) = (p', q', z') \cdot (A')^{-1} \qquad (6)$$

where z' is the depth value of texture element expressed in display coordinates.

The display coordinates (p'' q'') of the projection g at instant $t+\Delta t$ are:

$$(p'', q'', z'') = (\mu, \nu, \lambda) \cdot A'' \qquad (7)$$

where A'' is the transformation matrix at instant $t+\Delta t$

Thus, the correspondence of texture element G and its projections at instance t and $t+\Delta t$ is derived (Figures 2-3):

$$(p'', q'', z'') = (p', q', z') \cdot B \qquad (8)$$

where $B = (A')^{-1} \cdot A''$

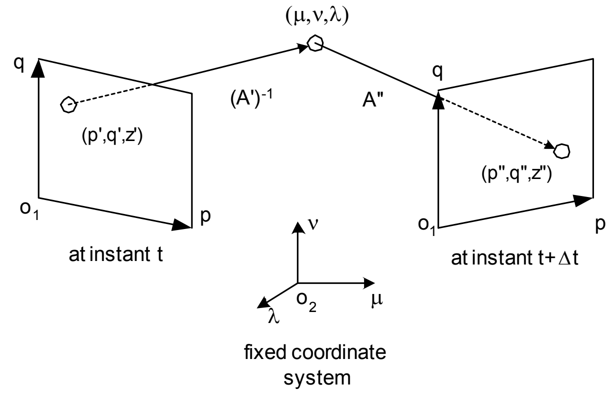

Figure 2: Deriving the projection coordinates at $t+\Delta t$.

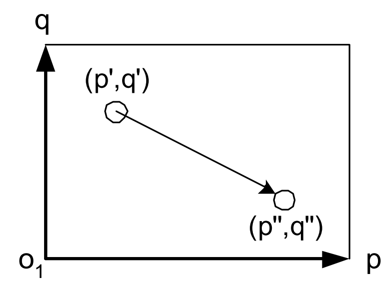

Figure 3: The instant velocity vector of the projection g.

Substituting p', q' and p'', q'' into (5) throughout the display surface, it is possible to calculate the velocity field sought.

Let's examine display surface in the neighborhood of pixel (p_1, q_1). The image g_1 of texture element G_1, represented by this pixel, has instant velocity vector $\vec{D}_1$ with projections D_{1p} and D_{1q} to axes p and q. The image g_2 of the element G_2, of pixel (p_1, q_1+1) has velocity vector $\vec{D}_2$ with projections D_{2p} and D_{2q}. And the image g_3 of the element G_3, of pixel (p_1+1, q_1) has vector $\vec{D}_3$ with projections D_{3p} and D_{3q} (Figure 4).

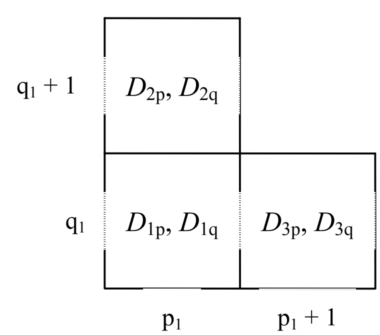

Figure 4: Velocity vectors of texture element projections adjacent to (p_1,q_1).

Considering the admissions taken under velocity field evaluation, the expressions (2), (3) and (4) will have the following form:

$$div\vec{D}_1 = (D_{3p} - D_{1p}) + (D_{2q} - D_{1q}); \quad (9)$$

$$curl\vec{D}_1 = (D_{3q} - D_{1q}) - (D_{2p} - D_{1p}); \quad (10)$$

$$def\vec{D}_1 = \sqrt{((D_{3p} - D_{1p}) - (D_{2q} - D_{1q}))^2 + ((D_{3q} - D_{1q}) + (D_{2p} - D_{1p}))^2} \quad (11)$$

The invariant values for pixels are in range [p=1..(n-1), q=1..(m-1)], where n, m – display surface dimensions.

4. CONCLUSION

The suggested approach has been implemented in the computer program visualizing observer's motion in an arbitrary 3D scene and evaluating resultant motion field and its invariants.

Let's model the vision of the vehicle driver moving in the road environment. The width of the road is 7 meters; it is a two-way road. The driver moves along the centre of the right lane up to the beginning of the right turn. At the beginning of the turn there is a billboard to the right of the road (dimensions: 6x3 meters, position of the billboard centre: 6 meters above the ground, 5 meters from the road edge) (Figure 5). Driver's eyes are 1.2 meters above the road. Driver's field of view is a regular pyramid with an apical angle of 12 degrees.

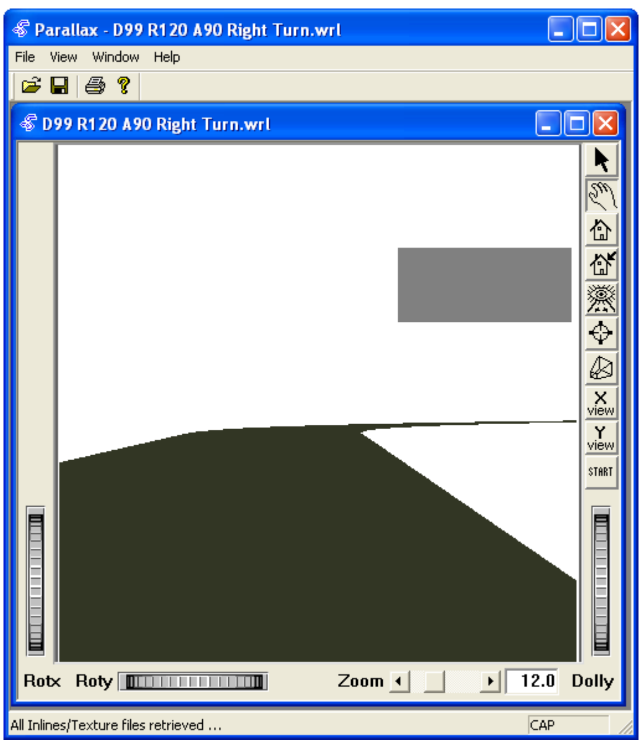

Figure 5: The road environment.

Let's consider the combinations of following options:

1. Turn radiuses: 40 and 120 meters (R40/R120);
2. Turn angles: 60 and 90 degrees (A60/A90);
3. Driver's view oriented to: the centre of the turn or to the centre of the billboard (RightTurn/RightBoard).

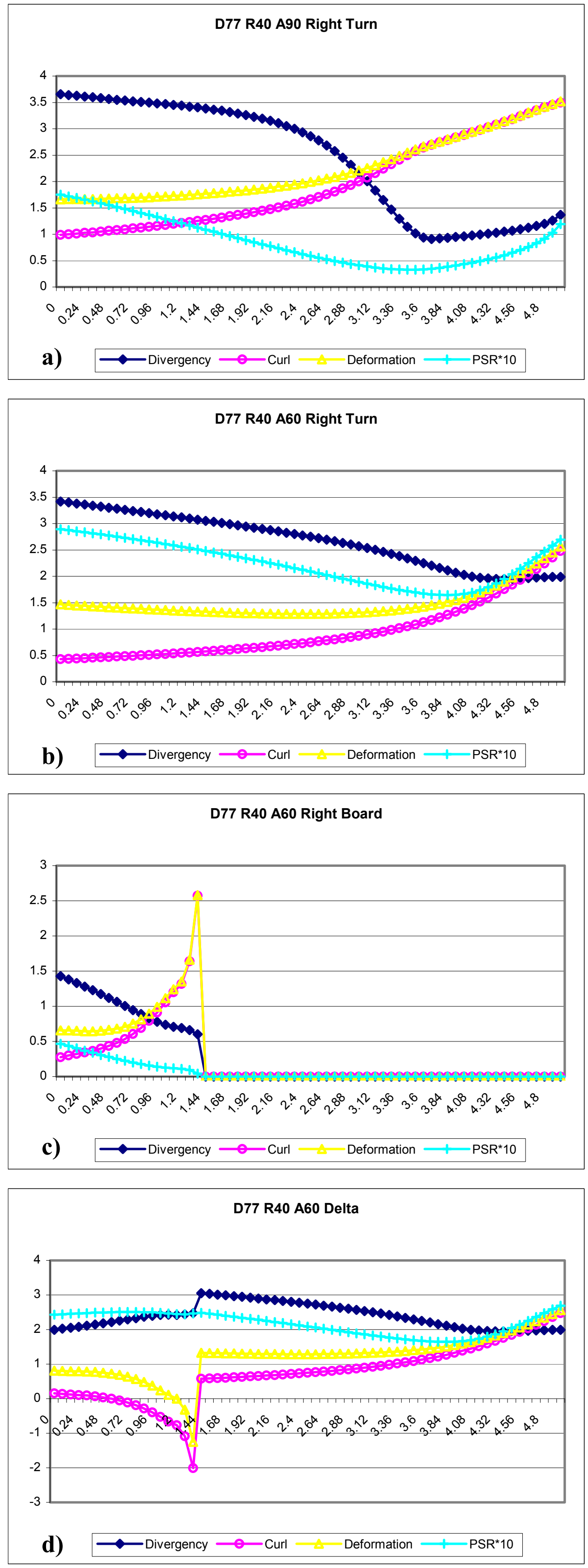

Figure 6: The road invariants dynamics at the turn radius = 40m.

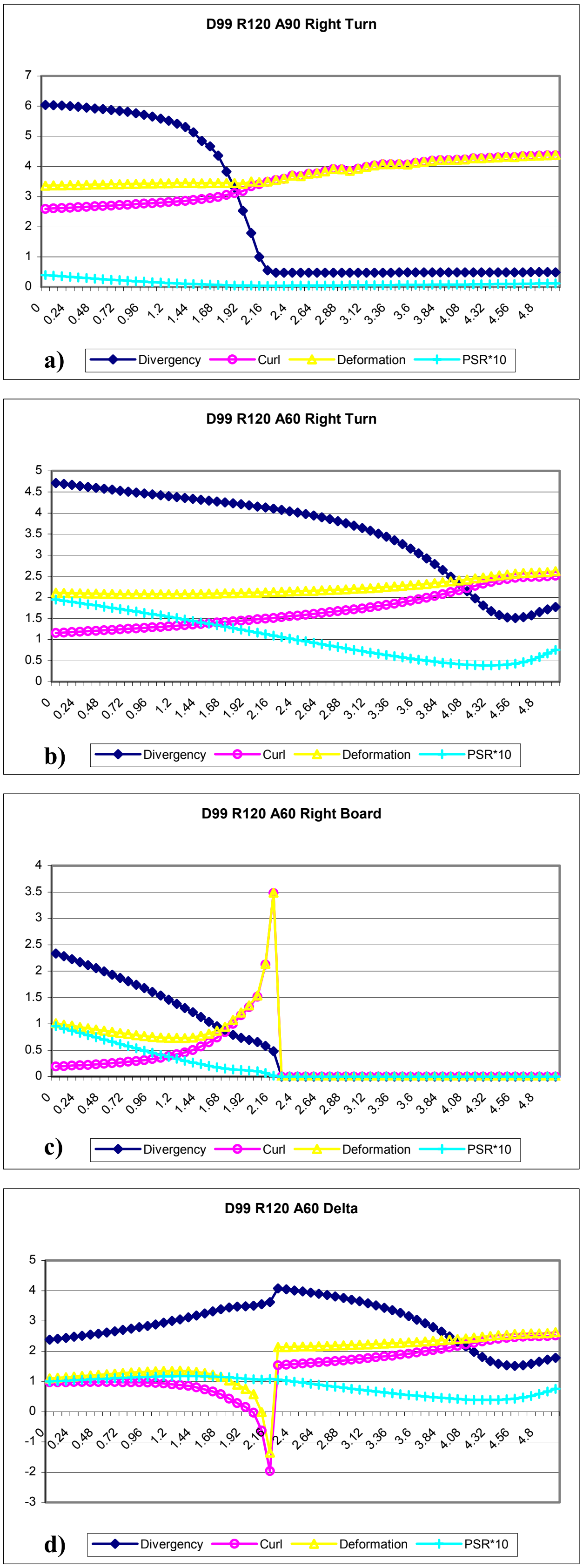

Figure 7: The road invariants dynamics at the turn radius = 120m.

If the turn radius is 40 meters, then at the distance of 77.8 meters (D77) to the turn, the vehicle velocity is 70 km/h. The driver uniformly decreases the speed, and at the beginning of the turn it comes to 42 km/h.

If the turn radius is 120 meters, then at the distance of 99.3 meters (D99) to the turn, the vehicle velocity is 90 km/h. The driver uniformly decreases the speed, and at the beginning of the turn it comes to 53 km/h.

The figures 6-7 are showing the time history of the optic flow invariants of the road up to 5 seconds to the turn reaching. PSR is the projection square to rendering surface square ratio. Diagrams **d)** represent the differences in the corresponding time histories of **b)** and **c)**.

Such data is of great importance both for evaluation of billboards influence on traffic safety and for billboard optimal placement.

5. REFERENCES

[1] Gibson J.J. *The ecological approach to visual perception*. New York, Boston. 1979.

[2] Еремин В.М., Уткин А.В. Зрительное восприятие движения и управление автомобилем. *Офтальмоэргономика: Итоги и перспективы, Тезисы докладов международного семинара*. Москва 13-15 февраля 1991. – М.: МНИИГБ им. Гельмгольца, 1991, -с. 12-13.

[3] Уткин А.В., Оптический поток и управление автомобилем, *Актуальные вопросы повышения безопасности дорожного движения*, Сб. науч. тр. МАДИ, 1988, - с. 36-46.

[4] Regan D., Beverley K.I., Visual fields for frontal plane motion and for changing size. *Vision Res*., v.23, N 7, p. 673-676, 1983.

About the authors

Valeriy Eremin is a professor at Moscow State Industrial University, Department of Informational Security. His contact email is eremin_valerii@mail.ru.

Maxim Koltyapin is a Ph.D. student at Moscow State Industrial University, Department of Informational Security. His contact email is koltyapin@atgs.ru.