

Galilean Differential Geometry of Moving Images

Daniel Fagerström

Computational Vision and Active Perception Laboratory (CVAP)
Department of Numerical Analysis and Computing Science
KTH (Royal Institute of Technology) SE-100 44 Stockholm, Sweden
`danielf@nada.kth.se`

Abstract. In this paper we develop a systematic theory about local structure of moving images in terms of Galilean differential invariants. We argue that Galilean invariants are useful for studying moving images as they disregard constant motion that typically depends on the motion of the observer or the observed object, and only describe relative motion that might capture surface shape and motion boundaries. The set of Galilean invariants for moving images also contains the Euclidean invariants for (still) images.

Complete sets of Galilean invariants are derived for two main cases: when the spatio-temporal gradient cuts the image plane and when it is tangent to the image plane. The former case correspond to isophote curve motion and the later to creation and disappearance of image structure, a case that is not well captured by the theory of optical flow.

The derived invariants are shown to be describable in terms of acceleration, divergence, rotation and deformation of image structure.

The described theory is completely based on bottom up computation from local spatio-temporal image information.

1 Introduction

The aim of this paper is to describe the local (differential) structure of moving images. By doing this we want to find a set of local differential descriptors that can describe local spatio-temporal pattern much as e.g. gradient strength, Laplacian zero-crossings, blob and ridge detectors, isophote curvature etc describe the local structure in images.

The dominating approach to computational visual motion processing (reviewed in [2, 15]) is to first compute the *optical flow* field, i.e. the velocity vectors of the particles in the visual observer's field of view, projected on its visual sensor area. From this various properties of the surrounding scene can be computed. Ego-motion can, under certain circumstances, be computed from the global shape of the field, object boundaries from discontinuities in the field, and surface shape and motion for rigid objects, can be computed from the local differential structure of the field [12, 13].

Unfortunately the computation of the optical flow field leads to a number of well known difficulties. The input is the projected (gray-level) image of the surroundings as a function of time, i.e. a three-dimensional structure. It is in general not possible to uniquely identify what path through the spatio-temporal image is a projection of a certain object point. Thus, further assumptions are needed, the most common one is the *brightness constancy assumption*, that the projection of each object point has a constant gray level. The brightness constancy assumption breaks down if the light changes, if the object have non-Lambertian reflection, or, if it has specular reflections. However, the problem is still under-determined, generically. Except at local extrema in the gray-level image, points with a certain gray-level lie along curves, and these curves sweep out surfaces in the spatio-temporal image. A point along such a curve can therefore correspond to any point on the surface at later instants of time. This is referred to as the *aperture problem* and is usually treated by invoking additional constraints e.g. regularization assumptions, such as smoothly varying brightness patterns, or parameterized surface models and trajectory models, leading to least-square methods applied in small image regions. Beside the questionable validity of these assumptions they lead to inferior results near motion boundaries, i.e. the regions that carry most information about object boundaries. The behavior when new image structure appears or old structure disappears is also undefined.

An alternative approach for visual motion analysis is to directly analyze the geometrical structure of the spatio-temporal input image, thereby avoiding the detour through the optic flow estimation step [18, 19, 11]. By using the differential geometry of the spatio-temporal image, we get a low level syntactical description of the moving image without having to rely on the more high level semantic concept of object particle motion.

A systematic study of the local image structure, in the context of scale-space theory, has been pursued by Florack [6]. The basic idea is to find all descriptors of differential image structure that are invariant to rotation and translation (the Euclidean group). The choice of Euclidean invariance reflects that the image structures should be possible to recognize in spite of (small) camera translations and rotations around the optical axis. This theory embeds many of the operators previously used in computer vision, such as Canny's edge detector, Laplacian zero-crossings, blobs, isophote curvature and as well enabling the discovery of new ones.

2 Spatio-Temporal Image Geometry

Extending from a theory about spatial images to one about spatio-temporal images it is natural to use the concept of *absolute time* (see e.g. [8] for a more elaborate discussion). Each point in space-time can be designated numeric label describing what time it occurred. The sets of space-time points that occurred at the same time are called *planes of simultaneity* and the temporal distance between two planes of simultaneity can be measured (in the small spatio-temporal regions that seeing creatures, operates in, we see no need for handling relativistic

effects, there are however other opinions, see [10]). The space-time can be stratified in a sequence of planes of simultaneity, and be given coordinate systems that separates time and space, $(t, x) \in \mathbb{R} \times \mathbb{R}^2$. From the consequences of absolute time, we conclude that we only want to allow for space-time transformations that never mixes the planes of simultaneity.

As a spatio-temporal image restricted to a plane of simultaneity can be considered as a still image the reasons for using Euclidean invariance in the image plane applies to moving images as well. Image properties should not be dependent on when we choose to measure them (invariance under time translations). The local average velocity contains only information about the ego motion and no information about the three dimensional structure of the environment, and is therefore natural to disregard. We thus search for properties that are invariant to the 2+1 dimensional Galilean group. The use of Galilean image geometry has been proposed in e.g. [4, 1, 9]. Using parallel projection as image formation model, the Galilean invariants are those properties of the surrounding that cannot be explained in terms of a relative constant translational motion. A Galilean model of the moving image is also implicitly assumed when divergence, curl and deformation are described as flow field invariants [12].

Definition 1 (Galilean group). *The group of Galilean motions Γ_{n+1} :*

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} 1 & v \\ 0 & R \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} + a = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \begin{pmatrix} 1 & v \\ 0 & I \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} + a \quad (1)$$

$x, v \in \mathbb{R}^n, t \in \mathbb{R}, R \in SO(n)$ and $a \in \Gamma_{n+1}$.

Each Galilean motion can be decomposed in a spatial rotation, a spatio-temporal shear (constant velocity) and a space-time translation. It can be shown that planes of simultaneity (constant time) are invariant and has Euclidean geometry, i.e. distances and angles are invariants. The temporal distance between planes of simultaneity is invariant.

3 Moving Frames

The Galilean geometry has no metric in traditional sense. That means that metric based differential geometry cannot be used in its normal formulations. We therefore chose to use a Lie group based approach instead (see [14] for a different approach on a geometry with degenerate metric).

According to Klein's famous Erlangen program, given a space S and a group of transformations G over S , the geometric structure of (S, G) is all structure that is invariant to transformations in G . In the following we will study the differential geometric properties of scalar functions and sub-manifolds (curves and surfaces) in \mathbb{R}^2 and \mathbb{R}^3 subject to Galilean and in some cases Euclidean transformations.

A convenient way to find geometrical structure is to use Cartan theory about moving frames [3, 16]. A *frame field* is a smooth map from the base space to group

elements, $S \rightarrow G$. For a Galilean geometry Γ_{n+1} the frame field is a mapping $\mathbb{R}^n \rightarrow \Gamma_{n+1}$. A frame field can be conceptualized by its action on an arbitrary coordinate system for the tangent space of the base space. For Γ_{n+1} we can e.g. attach a Galilean ON-system at each point.

Definition 2. A Γ_{n+1} coordinate system is an affine coordinate system where n vectors lies in the spatial part. A Γ_{n+1} ON-system is a Γ_{n+1} coordinate system s.t. the spatial part consists of n dimensional ON-coordinate system and the remaining base vector has unit temporal length.

The property of being a Γ_{n+1} ON-system is a Galilean invariant. In the sequel we will use the coordinate system view of frame fields as we find it easier to visualise.

The main idea of Cartans theory about moving frames is to put a frame at each point that is connected to the local structure of the sub-manifold or the function in an invariant way. In this way we get a frame field.

For a function f defined on S , all expressions over mixed derivatives w.r.t. the Cartan frame at a certain point are by construction geometrical invariants. This class of invariants are called *differential invariants*.

On sub-manifolds, we can find the local geometrical structure from how the frame field varies in the local neighborhood.

Let i be any (global) frame and e a frame connected to the local structure s.t. $e = Ai$, where the *attitude transformation* $A \in G$ is a function of position. The local variation of e can be described in an invariant way in terms of e ,

$$de = dAi = dAA^{-1}e = C(A)e, \quad (2)$$

where the *one-form* (see [3]) $C(A)$ is called the *connection matrix*. In a certain sense, the connection matrix contains all geometric information there is.

Scalar invariants can be generated by contracting the coefficients in the connection matrix on the vectors in the Cartan frame, $c_{ij}e_k$. A useful property of the connection matrix is,

$$C(AB) = C(A) + AC(B)A^{-1}, \quad (3)$$

which is a direct consequence of the definition.

The level-sets $f^{-1}(c)$ of smooth scalar functions f are sub-manifolds, the geometric structure of those, the *level-set invariants*, are invariant w.r.t. the group of constant monotonic transformations $g \circ f$, $g : \mathbb{R} \rightarrow \mathbb{R}, g' > 0$.

4 Image Geometry

Now we will study Galilean differential geometry of moving images using Cartan frames. Image spaces can be considered being trivial fiber bundle $S \otimes I$, where S is the base space and the fiber I is log intensity [14]. Most of the time we will discuss the image geometry in terms of an arbitrary section of the fiber bundle i.e. functions $f : S \rightarrow I$. We will start by reviewing differential geometry for

images over E_2 to illustrate the method of moving frames and as E_2 is a subgeometry of I_3 so that we will need these results later anyway. We continue by studying differential geometry of I_2 and, which is our main goal, differential geometry of images over I_3 .

For scalar functions over E_2 there are two typical situations: the gradient is non-zero almost everywhere and it is zero along curves.

4.1 Gradient Gauge

We study the geometry of functions f in E_2 . For points p where $\nabla f \neq 0$ we attach an ON-frame $\{\partial_u, \partial_v\}$ s.t. $f_u = 0$. (u, v) is a gauge coordinate system.

$$\begin{pmatrix} \partial_u \\ \partial_v \end{pmatrix} = \frac{1}{\|\nabla f\|} \begin{pmatrix} f_y - f_x \\ f_x \ f_y \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} = A \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} \quad (4)$$

where $\{\partial_x, \partial_y\}$ is a global ON-frame.

All functions over $\partial_u^i \partial_v^j f$, $i + j \geq 1$ becomes invariants w.r.t. rotations in space and translation in the intensity fibers. From (2) we get the anti-symmetric connection matrix:

$$C(A) = \begin{pmatrix} 0 & c_{12} \\ -c_{12} & 0 \end{pmatrix}, \quad (5)$$

where,

$$c_{12} = \frac{(f_x f_{xy} - f_y f_{xx})dx + (f_x f_{yy} - f_y f_{xy})dy}{f_x^2 + f_y^2} = -\frac{f_{uu}}{f_v} du + -\frac{f_{uv}}{f_v} dv. \quad (6)$$

where the expression is simplified by the use of the $\{\partial_u, \partial_v\}$ coordinate system, and the relation $f_u = 0$. By contracting c_{12} on the components in the Cartan frame we arrive at:

Theorem 1. *A complete set of level-curve invariants for scalar functions on E_2 is the level curve curvature, and the flow line curvature,*

$$\kappa = c_{12} \partial_u = -f_{uu}/f_v, \quad \mu = c_{12} \partial_v = -f_{uv}/f_v. \quad (7)$$

These are invariants w.r.t. rotation in the plane and monotonic transformations in the intensity fibers.

4.2 Hessian Gauge

The ON-frame (4) is not defined on critical points, $\nabla f = 0$, on typical critical points we can instead use an ON-frame $\{\partial_p, \partial_q\}$ that diagonalize the Hessian, i.e. $f_{pq} = 0$ and $|f_{pp}| > |f_{qq}|$.

$$\begin{pmatrix} \partial_p \\ \partial_q \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} = A \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix}, \quad (8)$$

where $\tan 2\phi = f_{xy}/(f_{yy} - f_{xx})$. All functions over $\partial_p^i \partial_q^j$, $i + j \geq 2$, becomes invariants w.r.t. the unimodular isotropic group, i.e. rotation in the image plane and addition of a linear light gradient [14]. The Hessian frame $\{\partial_p, \partial_q\}$ is invariant w.r.t. the isotropic group, i.e. all the motion in the isotropic group as well as scaling in the plane and in the intensity fiber [14].

5 Functions in Γ_2

First let us study the general geometrical situation for Γ_2 . The attitude transformation must be of the form:

$$\begin{pmatrix} \partial_s \\ \partial_u \end{pmatrix} = \begin{pmatrix} 1 & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \partial_t \\ \partial_x \end{pmatrix} = A \begin{pmatrix} \partial_t \\ \partial_x \end{pmatrix}, \quad (9)$$

where v , is a function of the spatio-temporal position and $\{\partial_s, \partial_x\}$ is the adapted frame. We immediately see that $\partial_u = \partial_x$. The connection matrix becomes:

$$C(A) = \begin{pmatrix} 0 & c_{01} \\ 0 & 0 \end{pmatrix} \quad (10)$$

where $c_{01} = v_t dt + v_x dx$. This could be expressed in the adapted coordinate system instead, giving $c_{01} = v_s ds + v_u du$. If the coefficient in the connection matrix is contracted on the vectors in the adapted frame, we get two scalar invariants, $a = c_{01} \partial_s = v_s$, that describe how the spatio-temporal part of the frame changes in the direction of it self, i.e. it describes the acceleration of the structure that the frame is adapted to. The other scalar invariant, $\delta = c_{01} \partial_u = v_u$, describes how the spatio-temporal part of the adapted frame changes in the spatial direction, i.e. the divergence of the vector field ∂_s , restricted to the spatial line.

For scalar functions on Γ_2 , there are three typical situations, the level curves are transverse to the spatial lines almost everywhere, along isolated curves the level curves are tangent to the spatial lines and there are also isolated critical points.

If one uses the constant brightness assumption as binding hypothesis between image patterns and surface motion then the level curves, (or isophotes) corresponds to motion in the traversal case and creation or annihilation of structure in the non-transversal case.

5.1 Spatially Transversal Level Curves

On points where the level curve is transverse to the spatial line, $f_x \neq 0$, we can define a Γ_2 -frame, $\{\partial_s, \partial_x\}$, s.t. $f_s = 0$. Expressed in an arbitrary Γ_2 -frame, $\{\partial_t, \partial_x\}$, ∂_s must be on the form:

$$\partial_s = \partial_t + \gamma \partial_x, \quad (11)$$

using $f_s = 0$ and solving for γ , we get $\gamma = -f_t/f_x$. Hence the attitude matrix becomes,

$$A = \begin{pmatrix} 1 & -f_t/f_x \\ 0 & 1 \end{pmatrix} \quad (12)$$

and for the connection matrix (10), we get:

$$c_{01} = \frac{f_t f_{tx} - f_x f_{tt}}{f_x^2} dt + \frac{f_t f_{xx} - f_x f_{tx}}{f_x^2} dx = -\frac{f_{ss}}{f_x} ds - \frac{f_{sx}}{f_x} dx. \quad (13)$$

Contracting c_{01} on the vectors of the adapted frame we get our scalar invariants, the invariants are summarized in the following theorem.

Theorem 2. *A complete set of level-curve invariants for spatially transversal level-curves on Γ_2 is level-curve acceleration the level-curve divergence*

$$a = c_{01}\partial_s = -f_{ss}/f_x, \quad \delta = c_{01}\partial_x = -f_{sx}/f_x. \quad (14)$$

5.2 Hessian Invariants

On points where $f_x = 0$, there is no tangent gauge. For points where $f_{xx} \neq 0$, we can define a *Hessian gauge*, i.e. an adapted Galilean ON-frame $\{\partial_s, \partial_x\}$ s.t. $f_{sx} = 0$. Repeating the steps from the last section, applying (11) on f_x , using $f_{sx} = 0$ and solving for γ , we get the attitude transformation:

$$\begin{pmatrix} \partial_s \\ \partial_x \end{pmatrix} = \begin{pmatrix} 1 - f_{tx}/f_{xx} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \partial_t \\ \partial_x \end{pmatrix} = A \begin{pmatrix} \partial_t \\ \partial_x \end{pmatrix}. \quad (15)$$

and in the connection matrix (10), we get:

$$c_{01} = -\frac{f_{ssx}}{f_{xx}} ds - \frac{f_{sxx}}{f_{xx}} = a ds + \delta dx. \quad (16)$$

Which we summarize in the following theorem.

Theorem 3. *A complete set of Hessian invariants for points where $f_{xx} \neq 0$ on Γ_2 is Hessian acceleration and Hessian divergence*

$$a = c_{01}\partial_s = -f_{ssx}/f_{xx}, \quad \delta = c_{01}\partial_x = -f_{sxx}/f_{xx}. \quad (17)$$

6 Functions in Γ_3

For Galilean 2 + 1 dimensional geometry, the attitude matrix in general have the form:

$$\begin{pmatrix} \partial_t \\ \partial_u \\ \partial_v \end{pmatrix} = \begin{pmatrix} 1 & v^x & v^y \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \partial_t \\ \partial_x \\ \partial_y \end{pmatrix} = Ai, \quad (18)$$

where v^x, v^y and θ are functions of the spatio-temporal position. It can be shown that the connection matrix expressed in the adapted coordinate system has the form:

$$C(A) = \begin{pmatrix} 0 & a^u ds + \delta^u du + \sigma^u dv & a^v ds + \sigma^v du + \delta^v dv \\ 0 & 0 & \rho ds + \kappa^u du + \kappa^v dv \\ 0 & -(\rho ds + \kappa^u du + \kappa^v dv) & 0 \end{pmatrix} = \begin{pmatrix} 0 & c_{01} & c_{02} \\ 0 & 0 & c_{12} \\ 0 & -c_{12} & 0 \end{pmatrix}. \quad (19)$$

Here c_{01} and c_{02} describes how the spatio-temporal part of the frame moves in different directions, c_{01} describes the motion projected on the $\{\partial_s, \partial_u\}$ plane,

and c_{02} the motion projected on the $\{\partial_s, \partial_v\}$ plane. The form c_{12} describes how the spatial frame $\{\partial_u, \partial_v\}$ rotates when moving in different directions.

Contracting the connection forms on the different vectors in the local adapted frame, we get nine different scalar invariants. We continue by giving these invariants an interpretation. If we consider the integral curves from the vector field $\{\partial_s\}$, then a^u describe the acceleration of the integral curve projected on the $\{\partial_s, \partial_u\}$ plane, and a^v the corresponding acceleration on the $\{\partial_s, \partial_v\}$ plane. ρ describes how much the spatial part of the frame rotates in the ∂_s direction. The invariants κ^u and κ^v describe the curvatures of the integral curves for the vector fields $\{\partial_u\}$ and $\{\partial_v\}$ respectively. The remaining invariants describe how the vector field $\{\partial_s\}$ changes for motions in the spatial plane, δ^u and δ^v describe the divergence in the ∂_u and ∂_v directions respectively. σ_u describes the skew of the vector field in the ∂_u direction while moving in the ∂_v direction and σ_v the skew in the ∂_v direction while moving in the ∂_u direction.

6.1 More Descriptive Invariants

Even if the above discussed set of scalar invariants constitute a complete set of scalar invariants for Γ_3 , they are not necessarily the ones that have largest descriptive value. As any invertible transformation of the scalar invariants give rise to a new complete set of scalar invariants, we will develop a set of invariants that are closer to what have been used in other work about moving images.

The acceleration invariants $\{a^u, a^v\}$ could instead be described in a polar coordinate system:

$$a = \sqrt{(a^u)^2 + (a^v)^2}, \quad a_\theta = \arctan(a^v/a^u), \quad (20)$$

here a is the magnitude of the acceleration, an a_θ the angle relative to the ∂_u direction. The invariants, $\delta_u, \delta_v, \sigma_u, \sigma_v$ describes how ∂_s changes along motions in the spatial plane. Observe that the vectors in the vector field $\{\partial_s\}$ always have unit length in the temporal direction, therefore the vector field restricted to a certain spatial plane can be projected onto that plane without losing any essential information. The matrix:

$$D = \begin{pmatrix} \delta_u & \sigma_u \\ \sigma_v & \delta_v \end{pmatrix} \quad (21)$$

is the rate of strain tensor for that projected vector field and it might be more useful to describe the invariants in terms of the Cauchy-Stokes decomposition theorem [12]:

$$D = \frac{\sigma_u - \sigma_v}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \frac{\delta_u + \delta_v}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \delta_u - \delta_v & \sigma_u + \sigma_v \\ \sigma_u + \sigma_v & \delta_v - \delta_u \end{pmatrix} \quad (22)$$

$$= \frac{\text{curl}D}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \frac{\text{div}D}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\text{def}D}{2} Q(\phi)^{-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} Q(\phi). \quad (23)$$

First the matrix can be decomposed in an anti symmetric and a symmetric part where the coefficient of the anti symmetric part is called the *curl* that describes

the rotational component of the vector field. The symmetric part can in turn be decomposed in a multiple of the identity matrix, the *divergence* part that describe the dilation component of the vector field, and a symmetric matrix with zero trace. The remaining symmetric component of the matrix can be described in terms of the *deformation*, i.e. an area preserving stretching in one direction combined with shrinking in the orthogonal direction, and the direction ϕ of the stretching relative to the direction of ∂_u .

6.2 Choice of Gauge

For Galilean 2 + 1 dimensional geometry isophotes are typically 2 dimensional surfaces. There are two generic cases: points where the isophote surface cuts the spatial surface through the point, and points where the isophote surface is tangent to the spatial surface through the point. The first case can be interpreted as motion of isophote curves in the image, and the second case as creation, annihilation or saddle points.

6.3 Tangent Gauge

Our next task is to define an adapted frame for points where the isophote surface cuts the spatial surface. For the spatial plane we can reuse the tangent gauge for E_2 in Section 4.1. Starting from an arbitrary frame i , we first adapt the spatial sub frame $\{\partial_x, \partial_y\}$, to the gradient and tangent direction in the spatial plane:

$$\begin{pmatrix} \partial_t \\ \partial_u \\ \partial_v \end{pmatrix} = \frac{1}{\sqrt{f_x^2 + f_y^2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & f_y & -f_x \\ 0 & f_x & f_y \end{pmatrix} \begin{pmatrix} \partial_t \\ \partial_x \\ \partial_y \end{pmatrix} = Ai. \quad (24)$$

The spatio-temporal vector ∂_s must have unit length in time to be part of a Galilean frame. By requiring ∂_s to lie in the spatio-temporal tangent plane, i.e. $f_s = 0$, it is constrained in one direction. The adapted spatio-temporal direction must have the form:

$$\partial_s = \partial_t + \beta\partial_u + \gamma\partial_v,$$

in terms of the new frame. Using $0 = f_s = f_t + \gamma f_v$ and solving for γ we get that $\gamma = -f_t/f_v$. Still we have one undetermined degree of freedom $\beta \in \mathbb{R}$. For each choice of β we have a plane spanned by $\{\partial_s, \partial_v\}$. The image restricted to such a plane is a function on I_2 and can be studied by the methods from Section 5.1. From Theorem 2 there are two scalar invariants: acceleration $a = -f_{ss}/f_v$ and divergence $\delta = -f_{sv}/f_v$. We can see that acceleration becomes a quadratic function of β and thus the gauge can be fixed by finding a β s.t. $a(\beta)$ is an extremum, i.e. by solving $\partial_\beta a(\beta) = 0$ for β , which gives:

$$\beta_a = \frac{f_t f_{uv}}{f_v f_{uu}} - \frac{f_{tu}}{f_{uu}}. \quad (25)$$

Which is defined as long as $f_{uu} \neq 0$, i.e. as long as the isophote curvature in the spatial plane is non-vanishing. It can be shown that requirement of an

acceleration extrema is equivalent to requiring $f_{su} = 0$ i.e. finding a β that diagonalizes the Hessian matrix in the $\{\partial_s, \partial_u\}$ plane. We will also see that this choice of gauge makes the direction of the spatial tangent, ∂_u , constant along ∂_s , i.e. $\rho = 0$. From this requirement Guichard [9], derived the same gauge as we use here.

Another choice of spatio-temporal gauge can be found by studying the divergence as a function of β . The divergence is a linear function of β and the disappearance of the divergence, $\delta(\beta) = 0$, is a natural way to fixate the gauge, giving:

$$\beta_\delta = \frac{f_t f_{vv}}{f_v f_{uv}} - \frac{f_{tv}}{f_{uv}}. \quad (26)$$

This is defined as long as $f_{uv} \neq 0$, i.e. when the flow line curvature in the spatial plane is non-vanishing. It can be shown that the disappearance of the divergence is equivalent to requiring that $f_{sv} = 0$, i.e. finding a β such that the Hessian in the $\{\partial_s, \partial_v\}$ plane is diagonalized.

Using (25) and (24) we find the attitude matrix for the acceleration based Γ_3 tangent gauge,

$$\begin{pmatrix} \partial_s \\ \partial_u \\ \partial_v \end{pmatrix} = \begin{pmatrix} 1 & \frac{f_t f_{uv}}{f_v f_{uu}} - \frac{f_{tu}}{f_{uu}} - \frac{f_t}{f_v} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \partial_t \\ \partial_u \\ \partial_v \end{pmatrix} = BAi. \quad (27)$$

The connection matrix can then be found by a tedious but elementary calculation using (3). Using notation from our general discussion about Γ_3 invariants the elements in the connection matrix (19) becomes:

$$c_{01} = a^u ds + \delta^u du + \sigma^u dv, \quad c_{02} = a^v ds + \delta^v dv, \quad c_{12} = \kappa du + \mu dv. \quad (28)$$

Observe that the skew invariant σ^v that describe the skew in the gradient direction while moving in the tangent direction, disappear. The spatio-temporal rotation of the frame in the spatial plane ρ disappears as well. We use the conventional notation $\kappa = \kappa^u, \mu = \kappa^v$, for isophote and flow line curvature. We list the resulting scalar invariants in the following theorem.

Theorem 4. *A complete set of scalar invariants for scalar functions on Γ_3 at points where the gradient and isophote curvature are non-vanishing are acceleration in the tangent and gradient direction,*

$$a^u = \frac{f_{ss} f_{uv}}{f_v f_{uu}} - \frac{f_{ssu}}{f_{uu}}, \quad a^v = -\frac{f_{ss}}{f_v}, \quad (29)$$

divergence in the tangent and gradient direction and skew in the gradient direction while moving in the tangent direction,

$$\delta^u = -\frac{f_{suu}}{f_{uu}}, \quad \delta^v = -\frac{f_{sv}}{f_v}, \quad \sigma^u = \frac{f_{sv} f_{uv}}{f_v f_{uu}} - \frac{f_{svv}}{f_{uu}}, \quad (30)$$

as well as isophote and flow line curvature, (see Theorem 1).

The invariant a^v , is also found in [9] and is denoted *accel*. The reasoning leading to Theorem 4 can be repeated for the divergence based tangent gauge (26).

Theorem 5. *A complete set of scalar invariants for scalar functions on Γ_3 at points where the gradient and flow line curvature are non-vanishing are acceleration in the tangent and gradient direction,*

$$a^u = \frac{f_{ss}f_{vv}}{f_v f_{uv}} - \frac{f_{ssv}}{f_{uv}}, \quad a^v = -\frac{f_{ss}}{f_v}, \quad (31)$$

divergence in the tangent and gradient direction, skew in the gradient direction while moving in the tangent direction,

$$\delta^u = \frac{f_{su}f_{vv}}{f_v f_{uv}} - \frac{f_{su}}{f_{uv}}, \quad \delta^v = -\frac{f_{su}}{f_v}, \quad \sigma^u = -\frac{f_{svv}}{f_{uv}}, \quad (32)$$

and isophote and flow line curvature, (see Theorem 1).

6.4 Hessian Gauge

On points where the isophote surface is tangent to the spatial surface, the tangent gauge is not defined. As long as the Hessian is non-degenerate, which generically is the case, we can define an adapted Γ_3 -frame, $\{\partial_r, \partial_p, \partial_q\}$ that diagonalize the Hessian, i.e. $f_{pq} = f_{rp} = f_{rq} = 0$. Using the fact that the spatio-temporal vector in the adapted frame must be on the form,

$$\partial_r = \partial_t + \beta\partial_x + \gamma\partial_y. \quad (33)$$

Starting by diagonalizing the Hessian in the spatio-temporal direction we get the constraints $f_{rx} = f_{ry} = 0$, and by using (33) and solving for β and γ , we get

$$\beta = \frac{f_{ty}f_{xy} - f_{tx}f_{yy}}{f_{xx}f_{yy} - f_{xy}^2}, \quad \gamma = \frac{f_{tx}f_{xy} - f_{ty}f_{xx}}{f_{xx}f_{yy} - f_{xy}^2}. \quad (34)$$

This gives the first part of the attitude transformation, a spatio-temporal shear A . If we project ∂_r on the spatial plane we get the same vector field as when the optical flow constraint equation is used on the gradient of the image [17]. As the next step the frame must be rotated in the spatial plane s.t. the spatial Hessian is diagonalized. Here we can use the results for the Hessian gauge for E_2 reviewed in Section 4.2. Combining these steps we get,

$$\begin{pmatrix} \partial_r \\ \partial_p \\ \partial_q \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} 1 & \beta & \gamma \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \partial_t \\ \partial_x \\ \partial_y \end{pmatrix} = BAi, \quad (35)$$

where $\tan 2\phi = f_{xy}/(f_{yy} - f_{xx})$. We proceed using (3) and the same reasoning as for the tangent based frames in the preceding section and arrives to the following theorem.

Theorem 6. *A complete set of scalar invariants for scalar functions on Γ_3 at points where the Hessian is non-degenerate are as follow:*

$$\begin{aligned}
a^p &= -f_{rrp}/f_{pp} & a^q &= -f_{rrq}/f_{qq} \\
\delta^p &= -f_{rpp}/f_{pp} & \delta^q &= -f_{rqq}/f_{qq} \\
\sigma^p &= -f_{rpq}/f_{pp} & \sigma^q &= -f_{rpq}/f_{qq} \\
\rho &= f_{rpq}/(2f_{pp} - 2f_{qq}) \\
\kappa^p &= f_{ppq}/(2f_{pp} - 2f_{qq}) & \kappa^q &= f_{ppq}/(2f_{pp} - 2f_{qq}).
\end{aligned} \tag{36}$$

Observe that in contrast to the tangent based gauge systems the Hessian gauge has all the scalar invariants listed in (19).

7 Conclusion and Discussion

In this paper we have developed a systematic theory about local structure of moving images in terms of Galilean differential invariants. We have argued that Galilean invariants are useful for studying moving images as it disregard constant motion that typically depends on the motion of the observer or the observed object, and only describe relative motion that might capture surface shape and motion boundaries. The set of Galilean invariants for moving images also contains the Euclidean invariants for (still) images.

Comparing to using optic flow as the basic element for describing image motion, the above suggested theory is completely bottom up and local, while optic flow is based on trying to directly interpreting the image motion in terms of (the projection of) motion of object surface points. The estimation of optic flow is non-local as it typically is based on gathering statistics about low level features in a small spatio-temporal surrounding. There are also Galilean differential invariants that can capture creation and disappearance of image structure, situations that are not covered by the concept of optic flow.

Experimental work is of course needed for evaluating how useful the suggested theory is for finding structure in real image sequences. Spatio-temporal images derivatives cannot be measured in a point, an integration over a non-vanishing spatio-temporal volume is needed [7], i.e. we need filters for measuring derivatives. As there are no localized filters that are invariant w.r.t. Galilean shear [5], a family of velocity adapted filters is needed. For computing a Galilean differential invariant, the velocity adapted filter used for measuring it should have the same spatio-temporal direction as the spatio-temporally directed gauge coordinate for the invariant. This could either be implemented by searching over a pre-computed set of spatio-temporally directed derivative filters or by iteratively adapt the spatio-temporal direction of the filter. It should be noted that in general, gauge adapted derivative filters can be found for several spatio-temporal directions at a point, i.e. for real image sequences the invariants can be multi-valued. This can be the case for e.g. transparent motion.

References

1. L. Alvarez, F. Guichard, P. Lions, and J. Morel. Axioms and fundamental equations of image processing: Multiscale analysis and p.d.e. *Archive for Rational Mechanics and Analysis*, 123(3):199–257, 1993.
2. J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.
3. H. Flanders. *Differential forms with applications to the physical sciences*. Dover Publications, Inc., 1989.
4. L. Florack, B. ter Haar Romeny, J. Koenderink, and M. Viergever. Families of tuned scale-space kernels. In *Proc. 2nd European Conference on Computer Vision*, pages 19–23, 1992.
5. L. Florack, B. ter Haar Romeny, J. Koenderink, and M. Viergever. Scale and the differential structure of images. *Image and Vision Computing*, 10:376–388, 1992.
6. L. Florack, B. ter Haar Romeny, J. Koenderink, and M. Viergever. General intensity transformations and differential invariants. *J. of Mathematical Imaging and Vision*, 4:171–187, 1994.
7. L. M. J. Florack. *Image Structure*. Series in Mathematical Imaging and Vision. Kluwer Academic Publishers, Dordrecht, Netherlands, 1997.
8. M. Friedman. *Foundation of Space-Time Theories: Relativistic Physics and Philosophy of Science*. Princeton University Press, 1983.
9. F. Guichard. A morphological, affine, and galilean invariant scale-space for movies. *IEEE Transactions on Image Processing*, 7(3):444–456, 1998.
10. W. C. Hoffman. The Lie algebra of visual perception. *J. Mathematical Psychology* 3 (1966), 65–98; errata, *ibid.*, 4:348–349, 1966.
11. B. Jahne. *Spatio-Temporal Image Processing-Theory and Scientific Applications*. Number 751 in Lecture Notes in Computer Science. Springer, 1993.
12. J. Koenderink and A. van Doorn. Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22(9):773–791, 1975.
13. J. Koenderink and A. van Doorn. Local structure of movement parallax of the plane. *J. of the Optical Society of America*, 66(7):717–723, July 1976.
14. J. Koenderink and A. van Doorn. Image processing done right. In A. H. et al., editor, *ECCV 2002*, number 2350 in LNCS, pages 158–172. Springer Verlag, 2002.
15. A. Mitiche and P. Bouthemy. Computation and analysis of image motion: A synopsis of current problems and methods. *International Journal of Computer Vision*, 19(1):29–55, July 1996.
16. M. Spivak. *Differential Geometry*, volume 1–5. Publish or Perish, Inc., Berkeley, California, USA, 1975.
17. S. Uras, F. Girosi, A. Verri, and V. Torre. A computational approach to motion perception. *bc*, 60:79–87, 1988.
18. M. Yamamoto. The image sequence analysis of three-dimensional dynamic scenes. Technical Report 893, Electrotechnical Laboratory, Agency of Industrial Science and Technology, May 1988.
19. C. Zetsche and E. Barth. Direct detection of flow discontinuities by 3d curvature operators. *Pattern Recognition Letters*, 12:771–779, 1991.