Vision Interface '99, Trois-Rivières, Canada, 19-21 May

Direct Motion Interpretation and Segmentation Based on the Robust Estimation of Parametric Models

X. Qian, A. Mitiche
INRS-Télécommunications, 16,Place du Commerce, Ille-des-soeurs, Verdun, P. Q., H3E 1H6, Canada

Abstract
A direct motion interpretation and segmentation method based on a motion model and robust statistics is proposed: estimation of optical velocities is consistent with the assumption of rigidity of environmental objects and accounts for motion discontinuities. The method has been tested both on several synthetic and real image sequences.

I. INTRODUCTION

One of the primary goals of motion analysis is to determine the 3-dimensional (3-D) structure of environmental objects and their movement relative to the viewing system. Usually, optical velocities are first estimated from the image spatio-temporal changes, and then interpreted in terms of the 3-D variables of motion and depth. However, one can combine these two steps and determine depth and 3-D motion "directly" from the spatio-temporal changes [1]. Once depth and 3-D motion determined, motion segmentation can be sought that is consistent with the assumption of environmental rigidity. For correct motion segmentation, motion estimation must account for motion discontinuities.

Motion estimation is an ill-posed problem [2]. A well posed problem is often obtained by regularization [3, 4]. Regularization can take a deterministic form [3] or a stochastic form [5]. In either case, if motion discontinuities are not accounted for, blurring of motion estimates occurs at these boundaries. To account for motion discontinuities, a variety of constraints have been proposed. For instance, the oriented-smoothness constraint [6] attenuates smoothing across strong intensity edges, which are identified with motion edges. Simplified, computationally more efficient versions of the oriented smoothness constraint have been developed [7], but at the cost of significant computational complexity. Motion discontinuities can also be processed in stochastic regularization by introducing variables to designate motion edges [5, 8]. This also increases significantly computational complexity which is already inherently high. Simpler, but still efficient methods use implicitly or explicitly outliers detection and rejection, outliers appearing inherently at motion boundaries [9–11].

Schunck [9] argued strongly in favor of robust regression for the outlier detection and rejection. A robust algorithm [9, 12] should be able to cope with noise, other data distortions and outlier occurrence. Proposed robust algorithms are of various complexity. Some can be of significant computational complexity (e.g., clustering methods [13]). Others, such as M-estimators, have low computational cost, but have a sensitivity to outliers that is proportional to the number of unknown [14]. Conspicuously, the least median of squares ($LMedS$) method has low sensitivity to outliers, can account for noise and other data distortions, and be implemented efficiently [12, 15].

In this paper, we present a direct motion interpretation and segmentation algorithm based on robust statistical estimation of a rigidity-based parametric motion model. Moving objects are subsequently segmented by adaptive K-means clustering.

II. MODEL AND ALGORITHM

A. Motion Estimation and Parametric Model

The Horn and Schunck gradient equation:

$$f_x u + f_y v + f_t = 0, \tag{1}$$

relates the spatial (f_x and f_y) and the temporal (f_t) derivatives of the image brightness function to optical velocity, (u, v) at each point. A reflection of the aperture problem, each such equation determines only the component of optical velocity in the direction of the gradient.

For both u and v, we substitute their expression in terms of depth and the parameters of motion in space, the environment assumed rigid:

$$\begin{cases} u = -xy\omega_x + (1+x^2)\omega_y - y\omega_z + \frac{\tau_x - x\tau_z}{Z} \\ v = -(1+y^2)\omega_x + xy\omega_y + x\omega_z + \frac{\tau_y - y\tau_z}{Z} \end{cases} \tag{2}$$

where (x, y) are the image coordinates of point (X, Y, Z) in space. τ_x, τ_y and τ_z are the translation components and ω_x, ω_y and ω_z are the rotational components of rigid motion, respectively [16]. If we substitute this expression of optical velocity into Eq.(1), we obtain an expression relating the brightness pattern spatial and temporal derivatives to the kinematic screw and depth: