1. INTRODUCTION

In this paper we will describe recent developments in the theory of early vision which lead from the formulation of the motion problem as ill-posed problem to its solution by minimizing certain "cost" functions. These cost or energy functions can be mapped onto very simple analog and binary resistive networks. Thus, we will see how the optical flow can be computed by injecting currents into "neural" networks and recording the resulting stationary voltage distribution at each node. These networks can be implemented in cMOS VLSI circuits and represent plausible candidates for biological vision systems.

2. Motion

There exist two basic methods for computing motion. Intensity-based schemes rely on spatial and temporal gradients of the image intensity to compute the speed and the direction in which each point in the image moves. The second method is based on the identification of special features in the image, tokens, which are then matched from image to image.

The principal drawback of all intensity-based schemes is that the data they use - temporal variations in brightness patterns - gives rise to the perceived motion field, the so-called optical flow. In general, the optical flow and the underlying velocity field, a purely geometrical concept, differ. If strong enough gradients exist in the image, the estimated optical flow will be very nearly identical to the underlying velocity field. In this article, we will assume that such strong gradients exist, as they do for most natural scenes, and consider how the velocity field can be computed using simple neural networks.
Figure 1. The aperture problem of motion.
Any system with finite aperture, whether biological or artificial in origin, can only measure the velocity component perpendicular to the spatial gradient \( I \), here indicated by the heavy line. Motion parallel to the gradient will not be visible, except by tracking salient features in the image.

2.1 Aperture Problem

Let us derive an equation relating the change in image brightness to the motion of the image (for more details, see Horn and Schunck, 1981). We will denote the image at time \( t \) by \( I(x,y,t) \). Let us assume that the brightness of the image is constant over time:

\[
\frac{dI}{dt} = 0.
\]

This transforms into

\[
\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = I_x u + I_y v + I_t = \nabla I \cdot \mathbf{v} + I_t = 0,
\]

if we define the velocity as \((u,v)=(dx/dt,dy/dt)\), and where \( I_x \), \( I_y \) and \( I_t \) are the partial derivatives of the brightness \( I \). Since we assume that we can compute these spatial and temporal image gradients, we are now left with a single linear equation in two unknowns, \( u \) and \( v \), the two components of the velocity vector.

Graphically, this so-called "aperture problem" is illustrated in figure 1. The problem remains even if we measure these velocity components at many points throughout the image, since or a single equation in two unknown is recovered at each location.

2.2 Smoothness Assumption

Formally, this problem can be characterized as ill-posed (Poggio, Torre, and Koch, 1985). How can it be made well-posed, that is, having a unique solution depending continuously on the data? One form of "regularizing" ill-posed problems is to restrict the class of admissible solutions by imposing appropriate constraints (Poggio et al., 1985).