INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONE: 2240-1
CABLE: CENTRATOM - TELEX 460392 - I

SMR/302-17

COLLEGE ON NEUROPHYSICS:
"DEVELOPMENT AND ORGANIZATION OF THE BRAIN"
7 November - 2 December 1988

"Motion Detection in Fly, Machine and Human Vision"

Heinrich H. BULTHOFF
Center for Biological Information Processing
M.I.T.
Cambridge, MA
USA

Motion Detection in Fly, Machine and Human Vision

by

Heinrich H. Bülthoff

Max Planck Institute

Biological Cybernetics

TÜBINGEN, FRG

and

Artificial Intelligence Laboratory

Center for Biological Information Processing

Massachusetts Institute of Technology

OUTLINE

- Motion Detection in Biological Systems (Flies and Rabbits) studied at different levels:
  Computational Level: Minimal Models
  Algorithmic Level: Correlation and Veto
  Hardware Level: Shunting Inhibition

- New parallel algorithm implemented on the *Connection Machine* solves

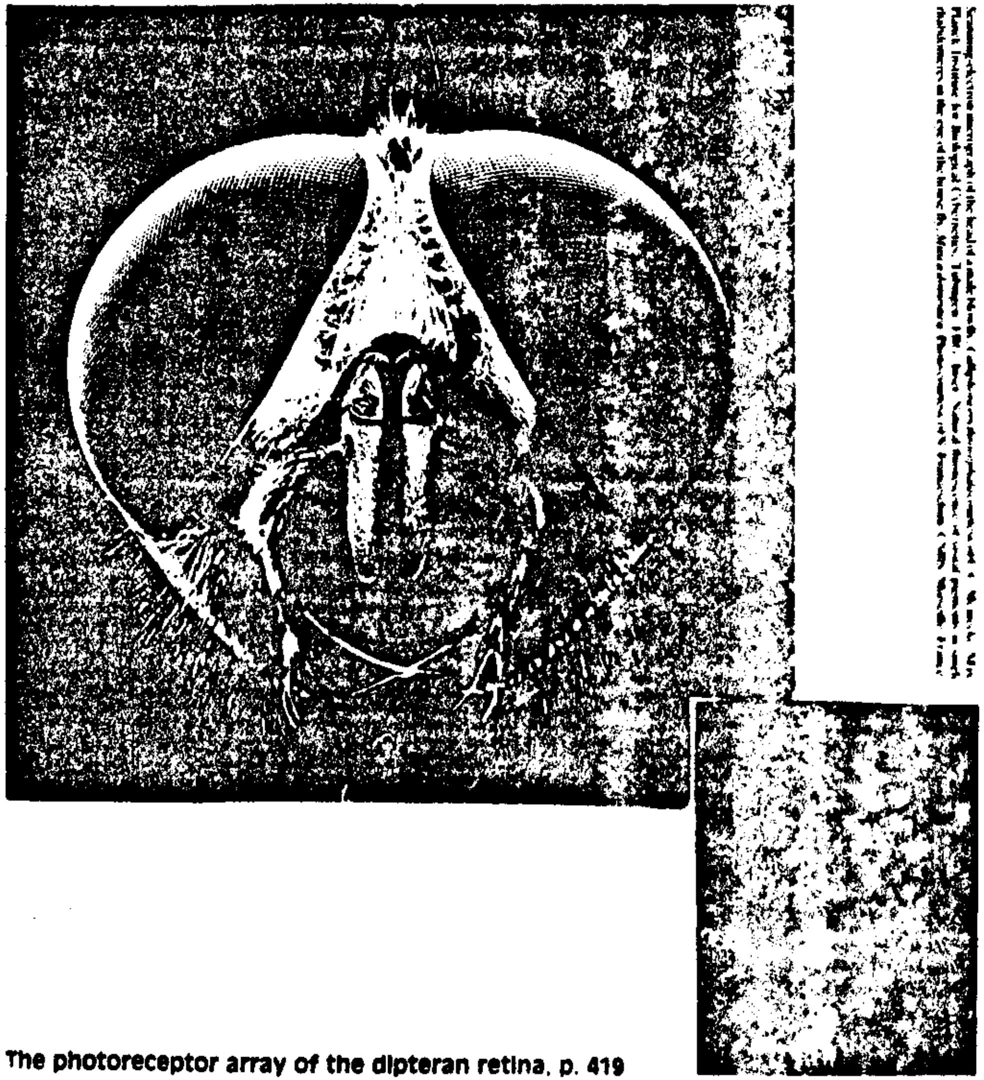
- *Correspondence* and *Aperture Problem*

- explains Psychophysical Illusions
  Barber Pole and Motion Capture, ...

TINS – September 1986 [99]

Volume 9, No. 9

Trends in NeuroSciences



The photoreceptor array of the dipteran retina, p. 419

WHY FLIES

- Model for high performance vision system
- Example for complex task: chasing flies
- 2 modules: **Movement and Position**
- for **Course control** and **Tracking**
- studied at Max Planck Institute, Tübingen
- **at the** behavioral, physiological and computational level

Motion Detection Theory

- Minimal computational requirements for diretion selective motion detection:
- **2 inputs**
- **Non-linearity**
- **Asymmetry**
- many different implementations at **algorithmic level**
- every 2-input system is equivalent to the correlation model under stationary and symmetrical experimental conditions, if non-linearities of order higher than second are negligible (Poggio and Reichardt, 1973)

Motion Detection at Hardware-Level

- many different schemes at algorithmic level
- based in fly-behavior, rabbit-physiology, human-psychophysics
- how is it implemented in neural hardware?
- two classes of neuronal models
- excitatory vs inhibitory models as proposed by Barlow and Levick, 1965

longer than excitation; a definite delay when it is passed laterally is not strictly necessary.

Some evidence favouring the right-hand, inhibitory, scheme has already been given. (1) As shown in Fig. 2 a stationary spot turned on and off elicits a response. If the excitatory conjunction scheme was modified to account for this it would probably still predict a considerably lower threshold for a moving than for a stationary spot. As shown in Fig. 5 of Barlow *et al.* (1964), the thresholds for spots of various sizes moving in the preferred direction differ by small and inconstant amounts from those for the same spot turned on or off. (2) The most striking feature of these directional units is the absence of any impulses when movement is in the null direction. This prompts one to look for a mechanism that inhibits unwanted responses. (3) When testing for directional selectivity in

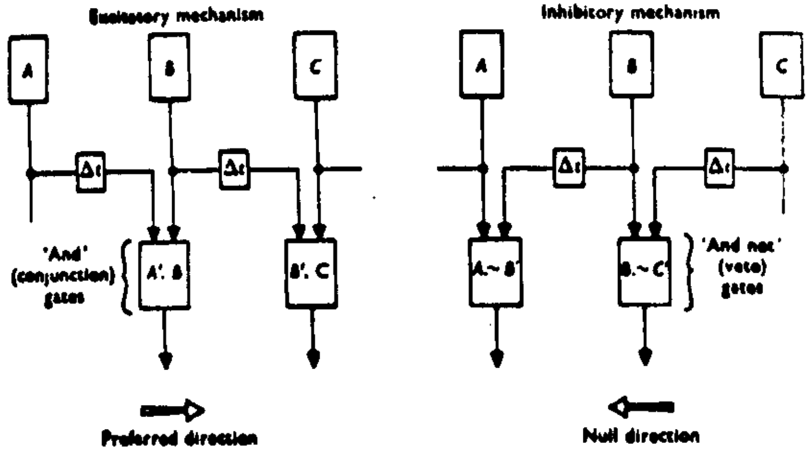


Fig. 7. Two hypothetical methods for discriminating sequence. For both, the preferred direction would be from left to right, null from right to left. In the excitatory scheme activity from the groups of receptors *A* and *B* is delayed before it is passed laterally in the preferred direction to the 'and' (conjunction) gates. If motion is in the preferred direction *A'* (delayed *A*) occurs synchronously with *B*, *B'* occurs synchronously with *C*, and these conjunctions cause the units in the next layer to fire. In the scheme on the right the activity spreads laterally, but in the null direction, from the groups of receptors *B* and *C*, and it has an inhibitory action at the units in the next layer; hence these act as 'and not' (veto) gates. The inhibition prevents activity from *A* and *B* passing through these gates if motion is in the null direction, but arrives too late to have an effect if motion is in the preferred direction. Notice that a special delay unit is not really necessary, for this scheme works if inhibition simply persists longer than excitation and can thus continue to be effective after a lapse of time. The excitatory scheme works by picking out those stimuli with the desired property, whereas the inhibitory scheme works by vetoing responses to unwanted stimuli; the latter is the one favoured by the experimental evidence.

Differentiate between Models at Hardware-Level

joint work with I. Bülthoff and A. Schmid

- Combining Neuropharmacology and Electrophysiology
- GABA-Antogonist Picrotoxinin
- blocks Chlorid-Channel
- blocks shunting
- should remove direction selectivity
- Test: extracelluar recording of large-field movement integrationg neuron (H1-cell)

Motion Detection

For Directional Selectivity

Minimal Requirements

- two inputs
- non-linearity
- asymmetry

Many Motion Algorithms

- Correlation Model (Hassenstein-Reichardt)
- Veto Model (Barlow-Levick)
- Shunting Inhibition (Torre-Poggio)

Motion Detection

New Parallel Algorithm

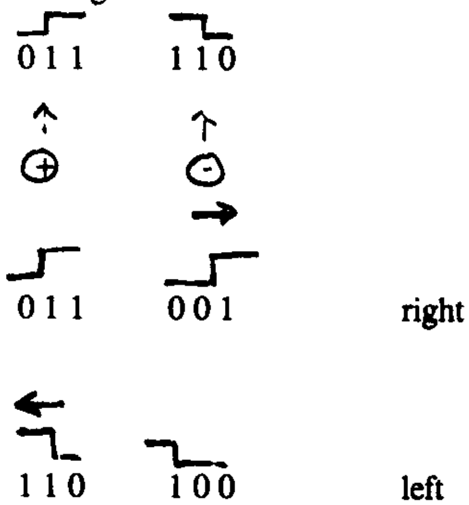
- based on biology
- edge-based Alg. motivated by Veto-Scheme
- intensity-based Alg. motivated by Correlation Model

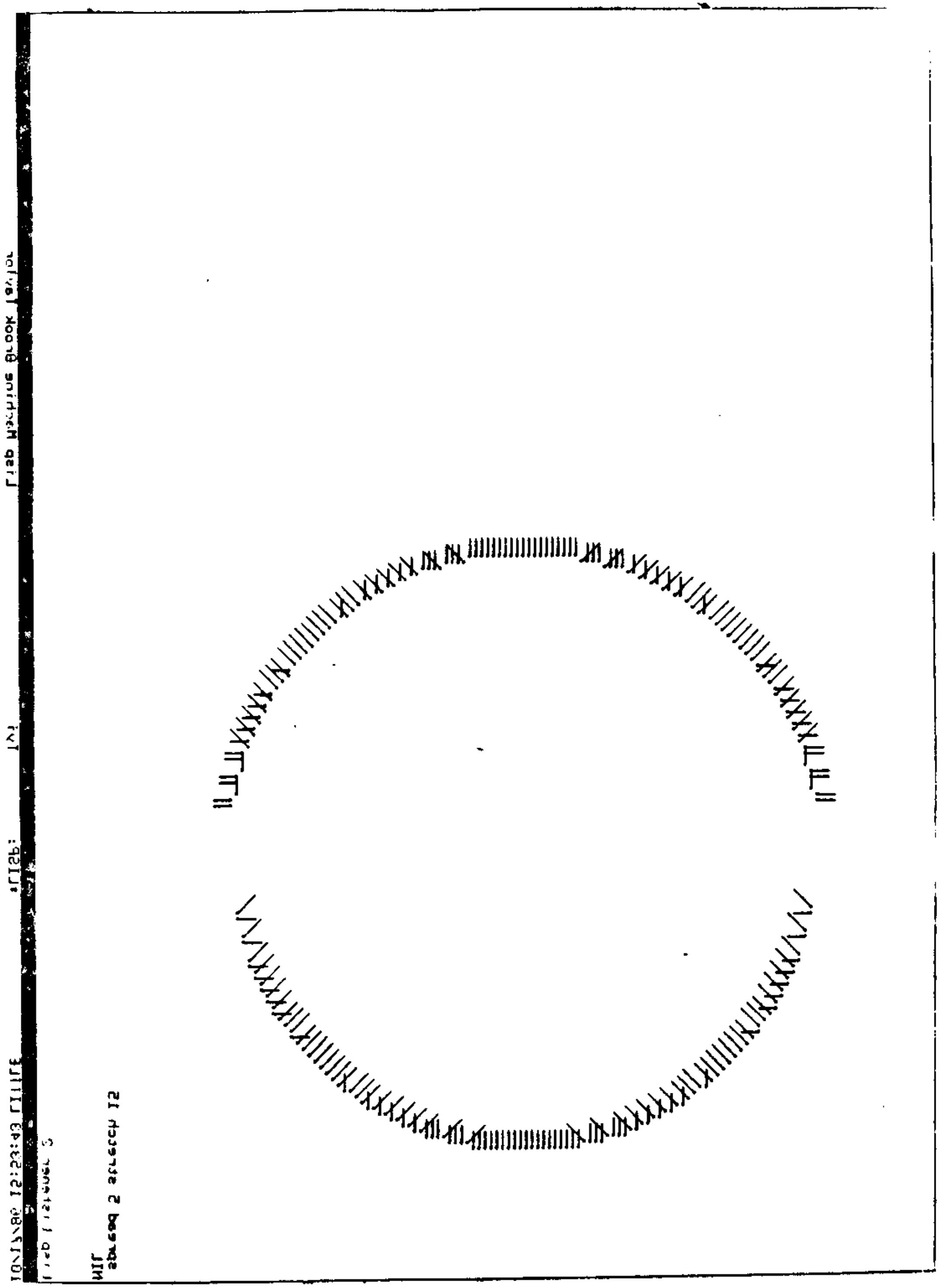
Our parallel algorithm for computing the *optical flow* is based on the simple assumption, that the optical flow is *locally uniform.*

Begin with simple binary detector (Barlow & Levick)

- Drawback: flicker sensitive (local intensity changes)
  small velocity range

- Remedy: broaden support in space because we restrict
  ourselfes to 2 timesteps
  3-input detector
  label edges

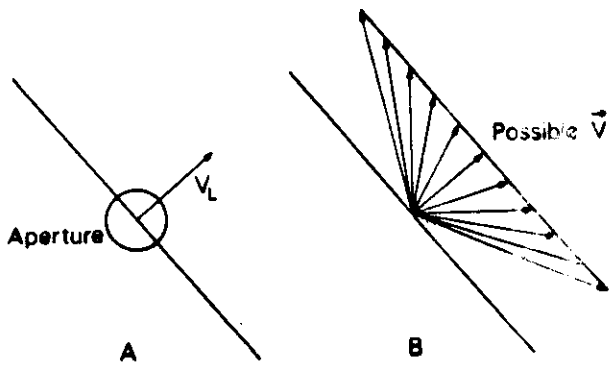




Strong and weak aperture problem

Aperture Problem

Local detectors can only compute normal component of motion (Marr, Ullman)



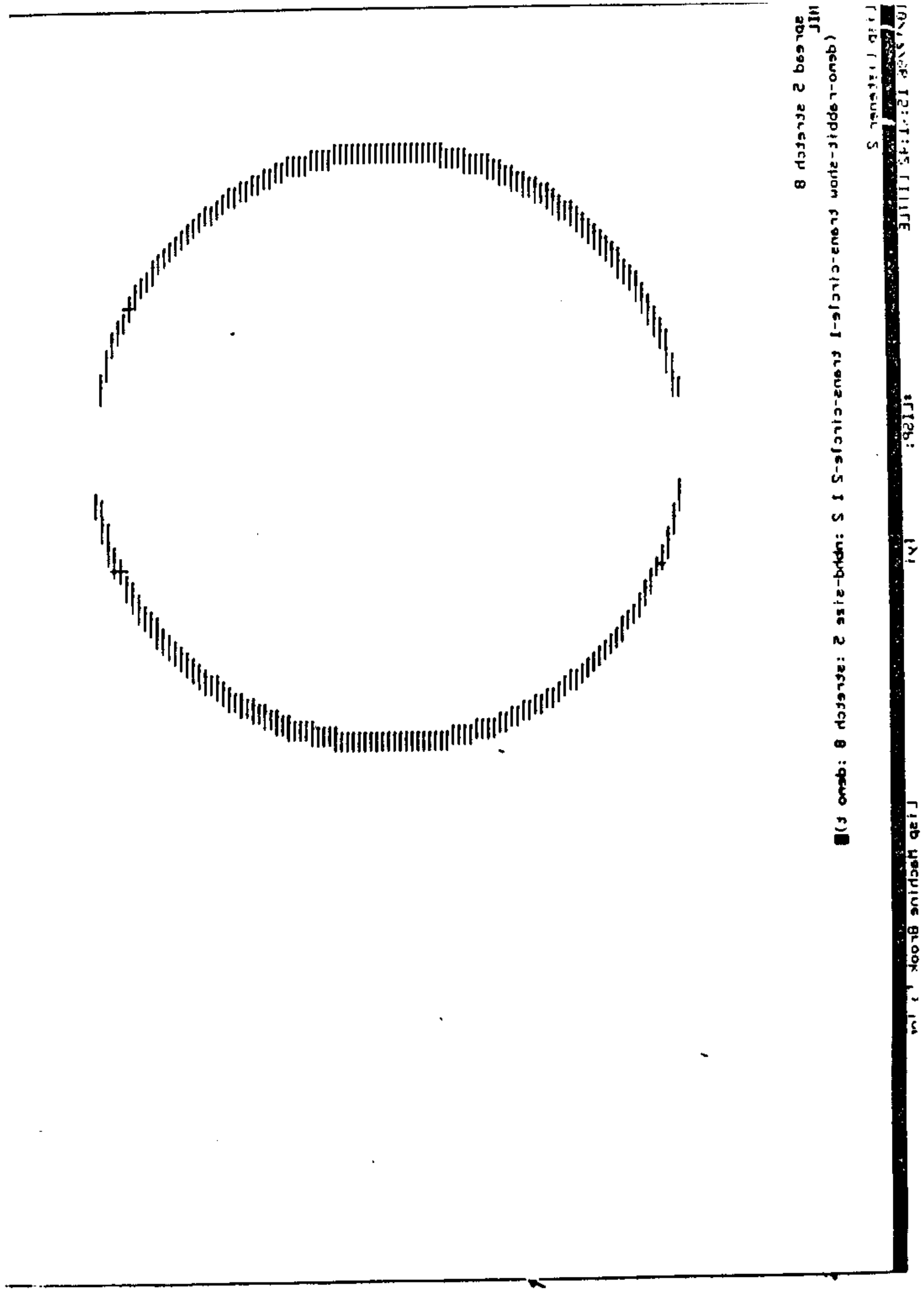
APERTURE PROBLEM

Solved by a variety of methods:

- Smoothest velocity field (Horn, Schunk)
- Smoothness along contours (Hildreth)

Aperture problem is an instance of the correspondence problem

(finding corresponding features in two images)



Voting Motion Algorithm
VMA

Physical constraints on motion limit the spatial variation of the optical flow field.

Constraints

- ***uniqueness,*** **each image point has a unique velocity**
- ***continuity*** **surface are locally smooth**

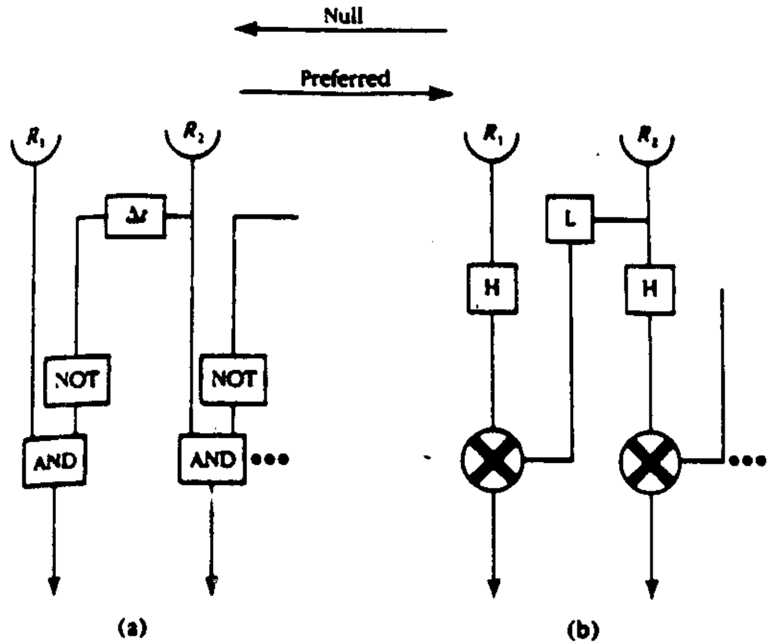


Figure 3–30. (a) Barlow and Levick's (1965) model for directional selectivity connects two detectors to an AND–NOT gate, one via a delay. Thus the network does not respond to stimuli moving with roughly the right speed in the null direction. (b) Hassenstein and Reichardt's (1956) model operates on the same principle except that the delay is replaced by a temporal low-pass filter (L). H = high-pass filter.

Vote for Motion

1. Find and Label Edges
2. Match Edges
   For $(-\delta \leq \Delta x \leq \delta, -\delta \leq \Delta y \leq \delta) = D$,
   shift $edges_2$ over $edges_1$.
3. Find Local Support
   For each $(\Delta x, \Delta y) \epsilon D$, count the matches in a neighborhood N of a pixel.
4. Vote
   At each (x, y), choose motion $(\Delta x, \Delta y)$ which has maximum local support.

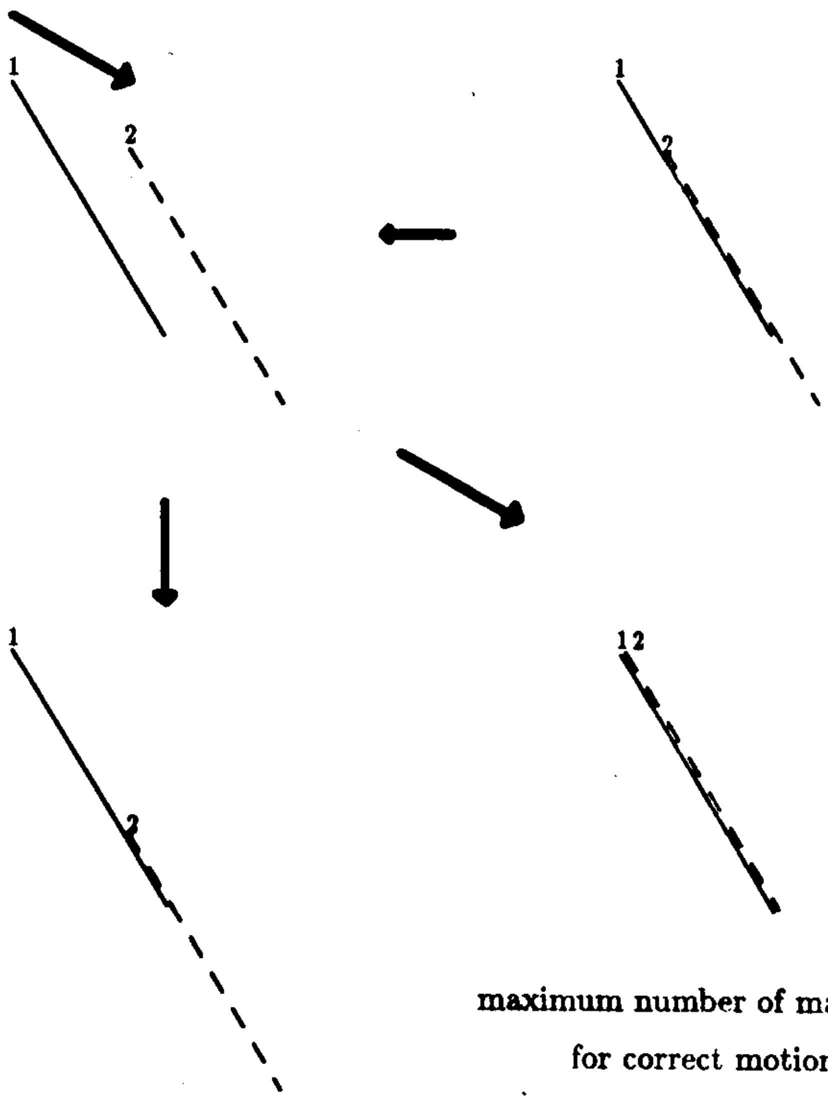
This results in a partial solution to the aperture problem:

All points in neighborhood N of a feature identify the correct motion.
Lines are not disambiguated, since there are no features to match.
Heuristic can select the motion of smallest magnitude, in case of ambiguity.

Edge-based VMA

- local comparison of edge maps over displacement range
- record matches for each displacement (see CM-Implementation)
- gather local spatial support for each displacement (**area-based**)
- find maximum votes for displacement vector
- winner-take-all
- image segmentation by using conflicting votes (relative motion)

SHIFT and MATCH

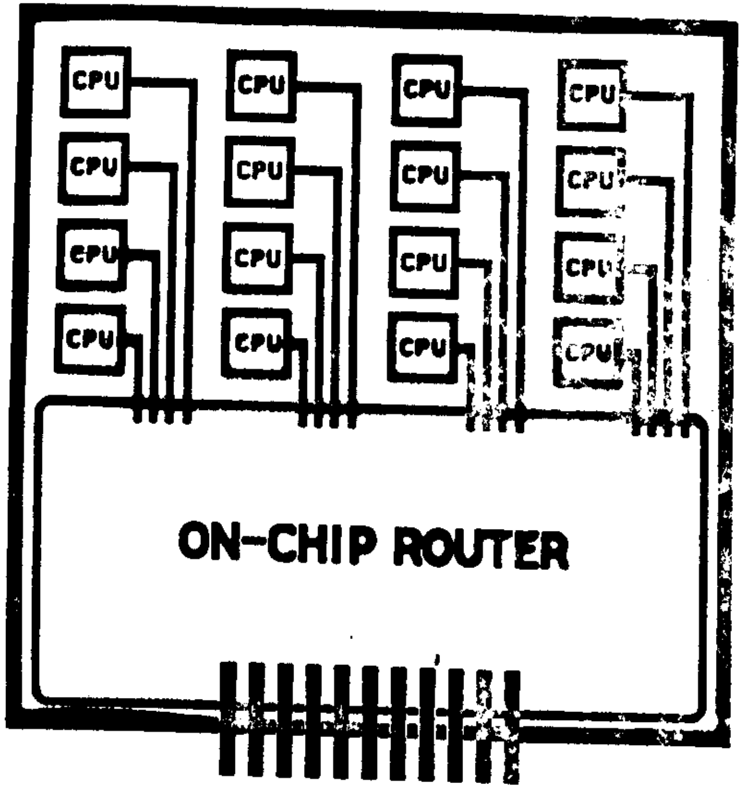


maximum number of matches
for correct motion

Connection Machine Implementation

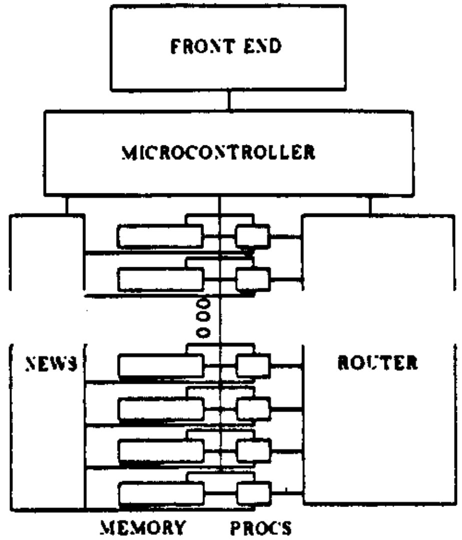
- VMA maps easily into CM-architecture
- retinotopic mapping into CM-memory
- one processor per pixel
- parallel shift and match operation (NEWS)
- each processor keeps record of correct matches
- **vote** for maximum consistency in an area
- **voting scheme is area-based**

The Connection Machine Chip



To Other Chips

Structure of the Connection Machine



SIMD - Single Instruction Multiple Data

Subsets of processors can be *selected*, i.e., made active, based on some logical condition.

Several global operations can be performed rapidly either on *all* processors, or on *selected* processors

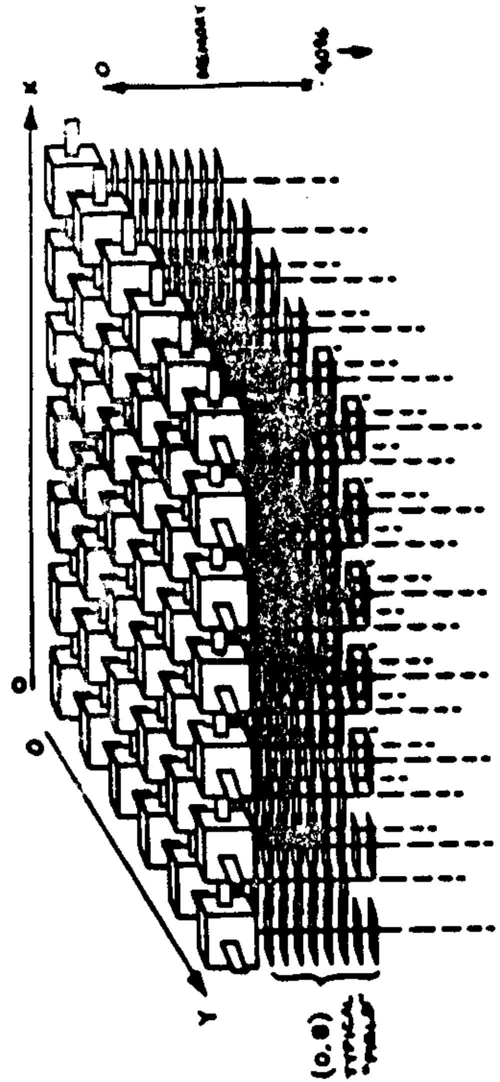
logical *OR*, logical *AND*, *max*, *min*, *sum*, *count*

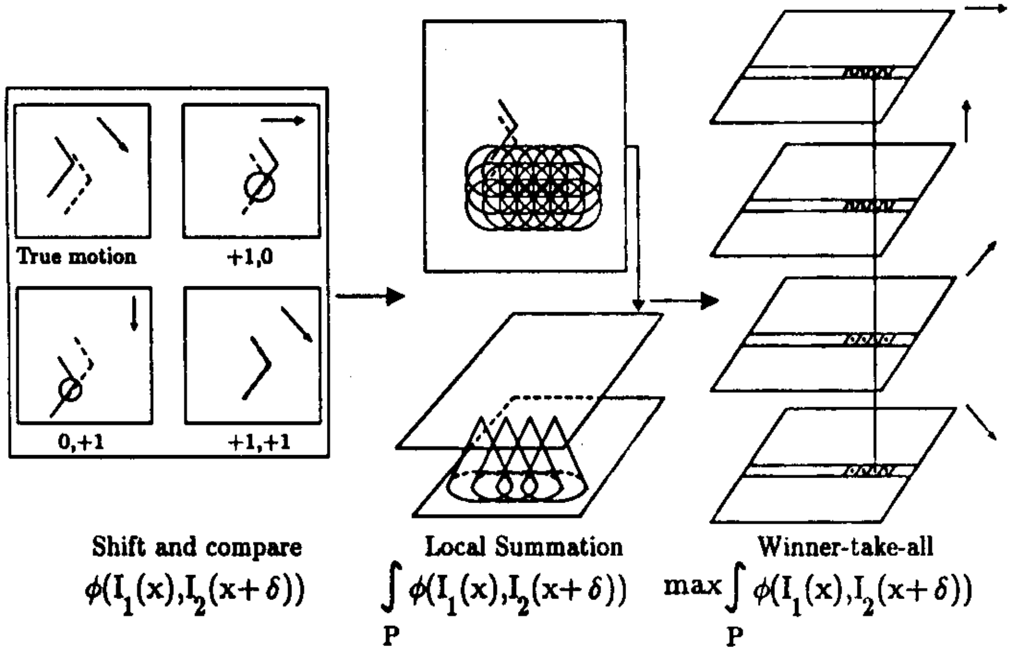
64 4096 Chips

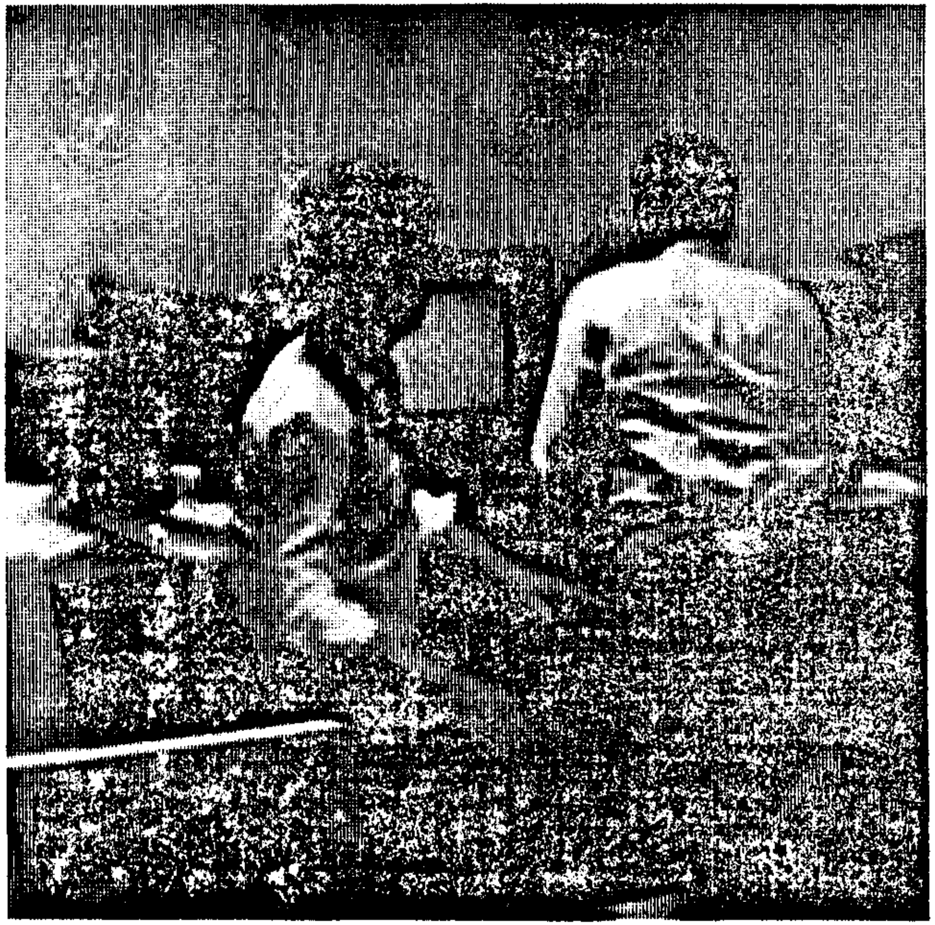
The value is returned to the host.

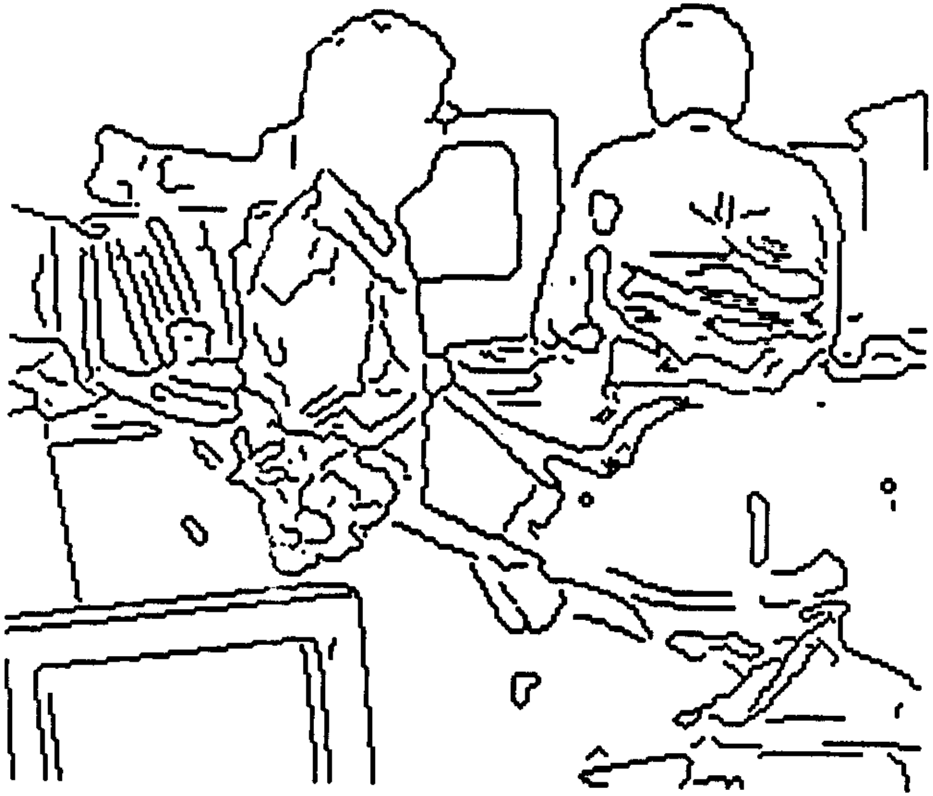
Connection Machine Implementation

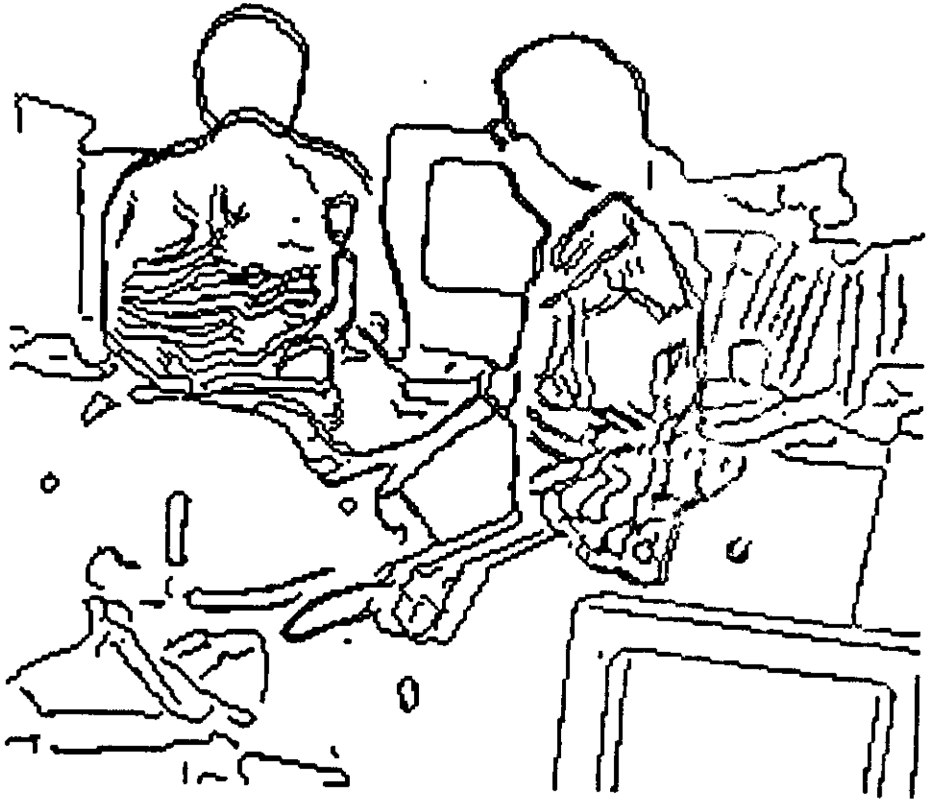
- maps easily into CM-architecture
- retinotopic mapping into CM-memory
- one processor per pixel
- shift and match operation in parallel
- each processor keeps record of correct matches
- vote for maximum consistency in area



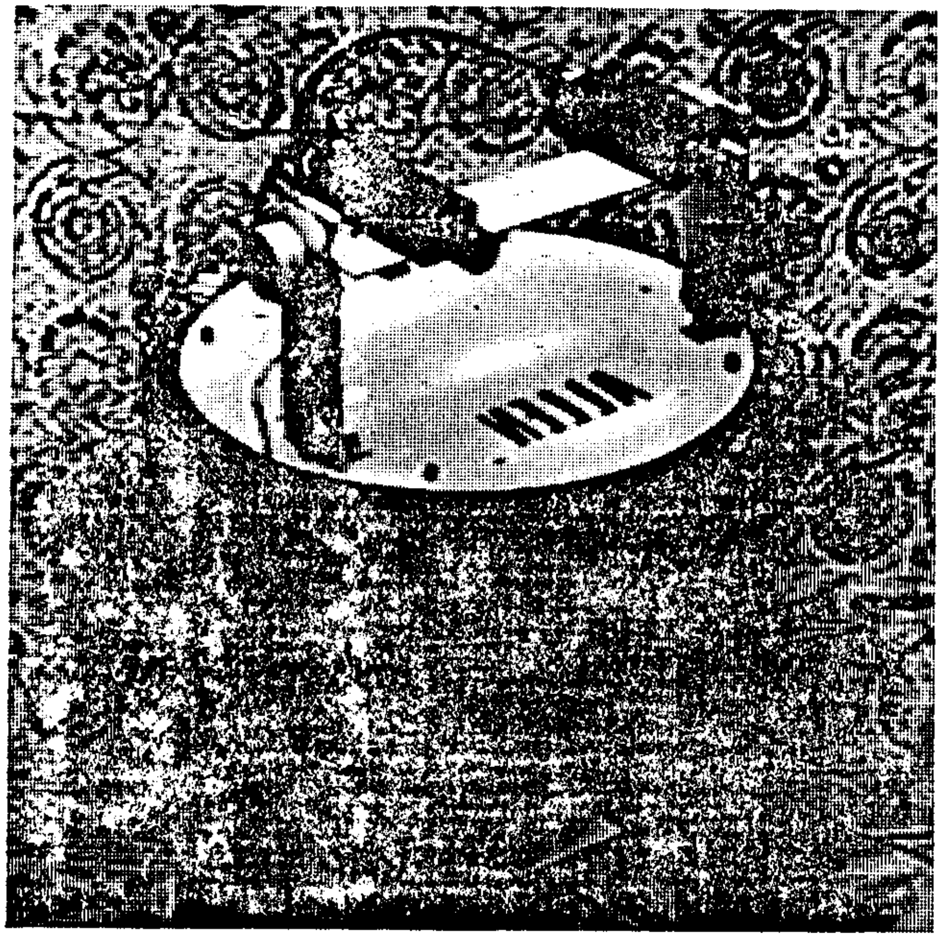

True motion
+1,0
0,+1
+1,+1
Shift and compare
$\phi(I_1(x), I_2(x+\delta))$
Local Summation
$\int_P \phi(I_1(x), I_2(x+\delta))$
Winner-take-all
$\max \int_P \phi(I_1(x), I_2(x+\delta))$

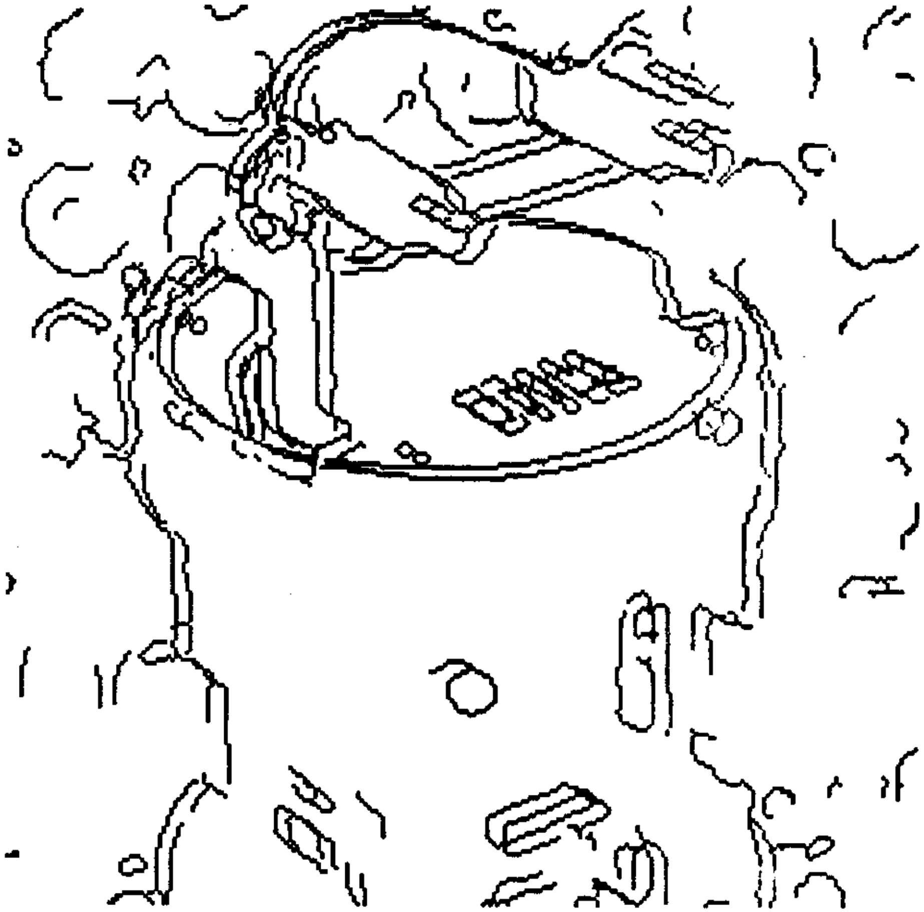


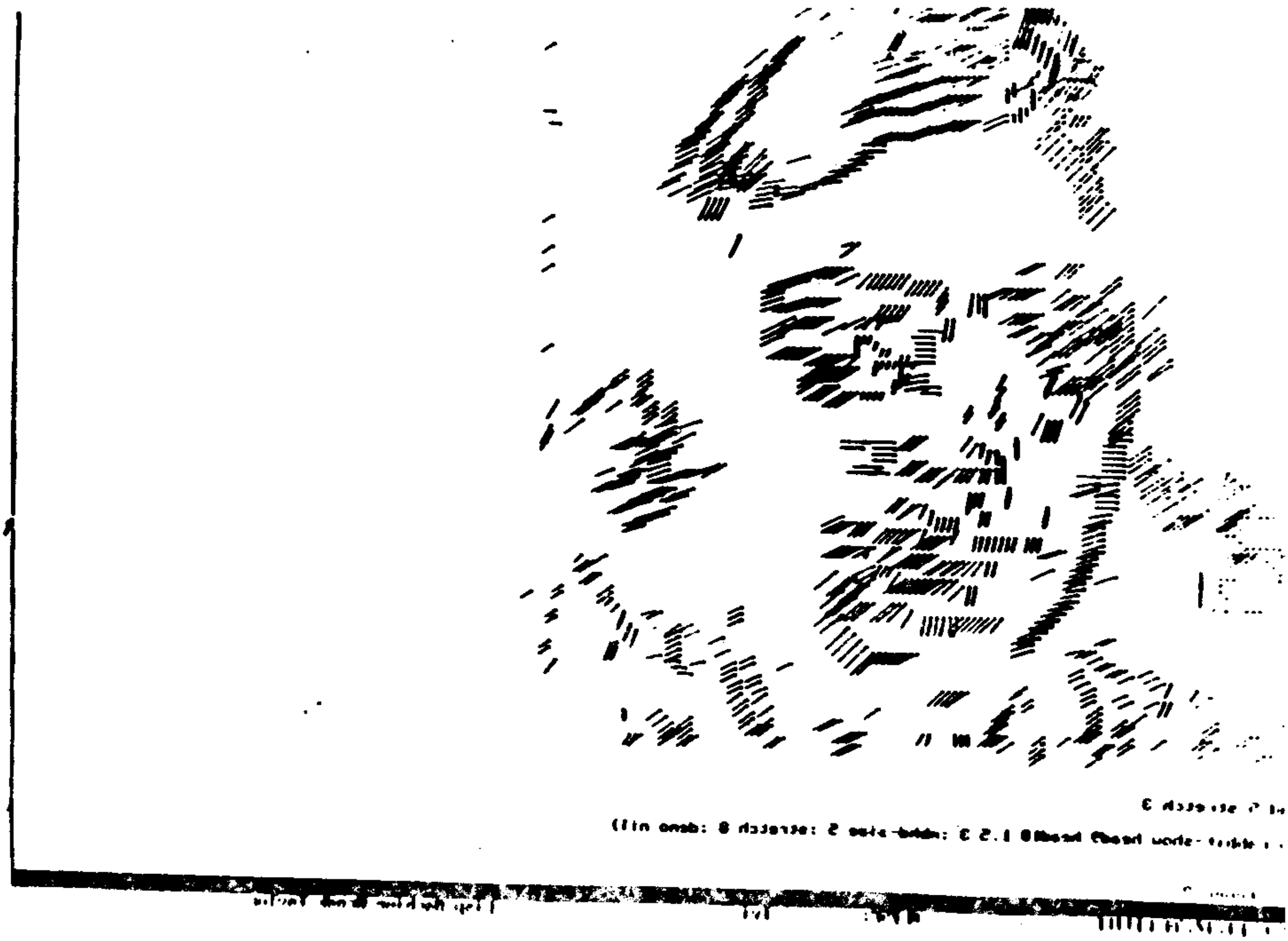





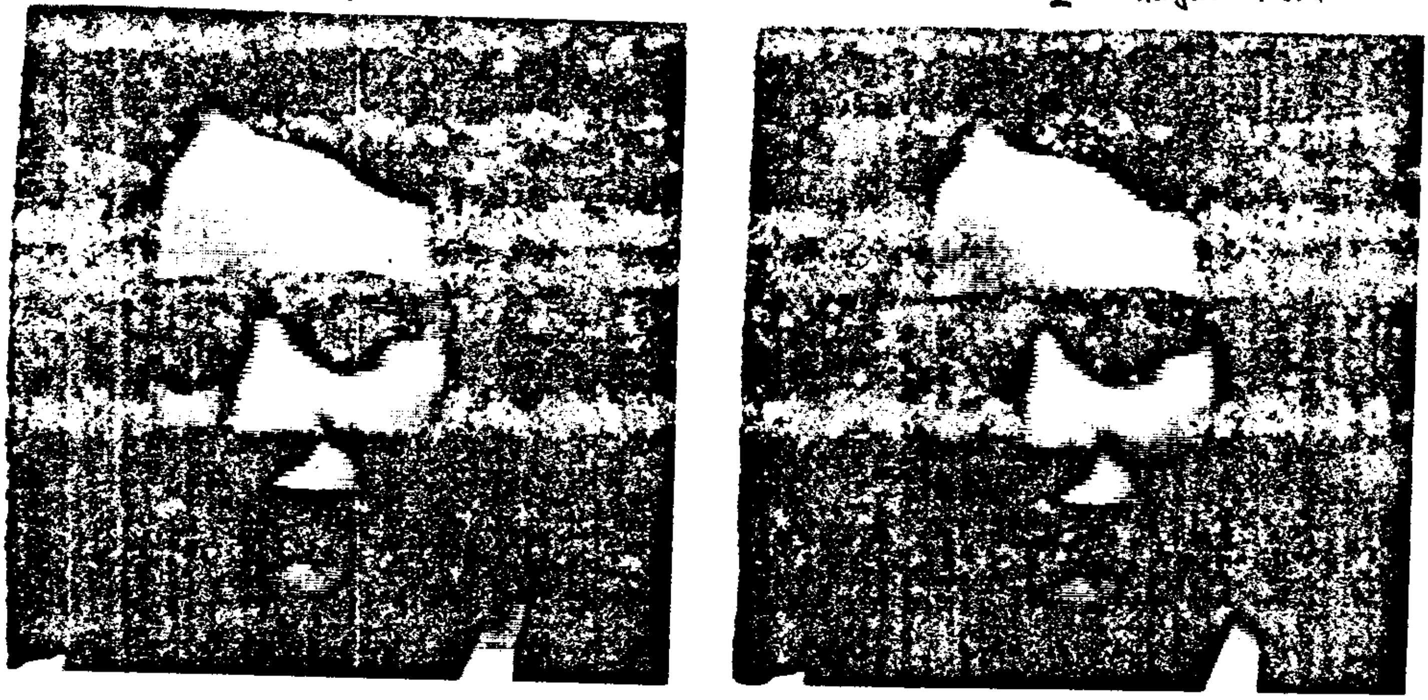
2.







1 → 2 Margaret Fleck



Stereo-Algorithm

Drumheller and Poggio have implemented the Marr-Poggio cooperative stereo algorithm on the *Connection Machine*.

- shift $edges_2$ over $edges_1$ by disparity d
- for each disparity d, count the matches in a neighborhood around each pixel
- select disparity which maximizes the local count of matches

Stereo- vs Motion-Algorithm

- **Stereo:** displacements are restricted to lie along one dimension (epipolar lines). The displacements range can be large.
- **Motion:** the search region is two-dimensional. The displacements range must be restricted and can be very small for high sampling rate (small δt). No Correspondence Problem for $\delta t \rightarrow 0$.
- In both modalities, we choose displacements which maximizes the local count of matches.

Edge-based VMA

- local comparison of edge maps over displacement range
- record matches for each displacement
- gather local spatial support for each displacement (**area-based**)
- find maximum votes for displacement vector
- **winner-take-all** or non-maximum suppression
- image segmentation by using conflicting votes (**relative motion**)

Image Segmentation by Relative Motion Two Algorithms

- **1. algorithm:** based on figure-ground detection in flies (Reichardt et al., 1983)
- inhibit or veto value of optical flow at each point by average value of large field centered at that point, **whenever** the motion is of the same type.
- similar to Land-Retinex-1986
- similar to center-surround operation (Laplacian of a Gaussian with large surround) on the *log* of the optical flow

Image Segmentation by Relative Motion

- **2. Algorithm:** based on the statistics of the voting step

- segmentation where local approximation of constant flow breaks down

- find location that get a minimum number of votes for consistent motion

- region of close "ties"

- scale the number of votes at a location by the total number of features in a local neighborhood

Advantages of VMA

- facilitates image segmentation

- segmentation not based on output, it is internal to the computational mechanism

- non-iterative → fast

- less noise sensitive (patch integration)

- intensity-based VMA produces **dense** motion field

- biological plausible

- explains psychophysical illusions