# Mixture Models for Optical Flow

**Basic idea:** Motion constraints come from a mixture of simple probability distributions:

- One simple distribution for each of several different *layers*.
- Each layer corresponds to a smooth motion field $\vec{W}(\vec{x}; \vec{a}_l)$.
- A background distribution for *outliers* is included.
- Each motion constraint is assigned an *ownership probability* for belonging to each layer and to the outlier distribution.

The *Expectation-Maximization (EM) algorithm* is used to obtain a maximum likelihood fit for the parameters of the mixture model, including the warp parameters $\vec{a}_l$ for each layer.

The introduction of an outlier layer is roughly equivalent to using a *robust estimator* for $\rho(e)$ (instead of the squared error $e^2$).

# Specification of a Mixture Model

Probability of observing motion constraint $\vec{c}_k$

$$p(\vec{c}_k|\vec{x}_k, \vec{m}, \vec{a}_1, \ldots, \vec{a}_N) = \sum_{n=0}^{N} m_n p_n(\vec{c}_k|\vec{x}_k, \vec{a}_n),$$

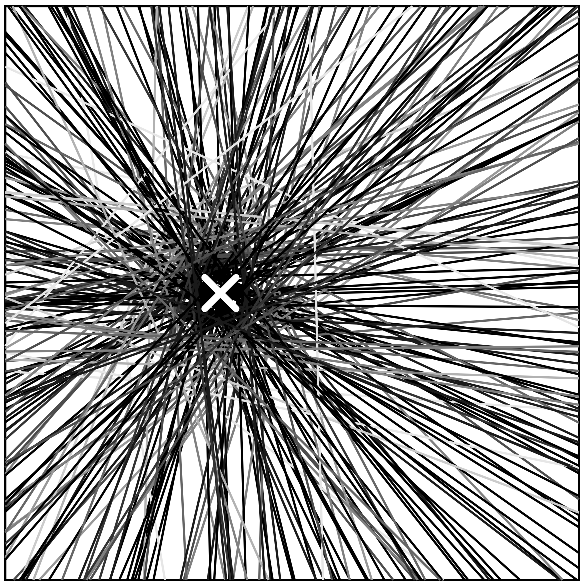
where

- $p_n(\vec{c}_k|\vec{x}_k, \vec{a}_n)$ = probability of observing motion constraint $\vec{c}_k$ given it comes from the $n^{th}$ layer at spatial location $\vec{x}_k$.
- $m_n$ is the mixture proportion for the $n^{th}$ layer
- $\vec{a}_n$ = parameters of $n^{th}$ layer

# Mixtures of Gaussian Distributions

Gaussian distribution of component velocities for $n^{th}$ layer:

$$p_n(\vec{c}|\vec{x},\vec{a}) = \text{Gaussian}(\vec{c}\cdot\vec{v}(\vec{x},\vec{a}) \mid 0, \sigma_v)$$



Models for $n^{th}$ layer:

- Constant velocity within patch:

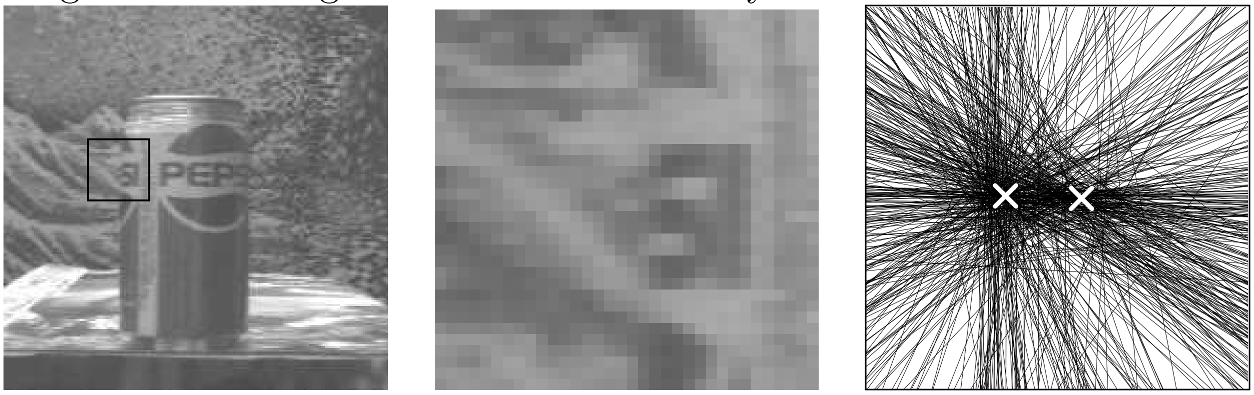
  $\vec{a}$ two dimensional: $\vec{v}(\vec{x},\vec{a}) = \vec{a}$

- Affine motion within patch:

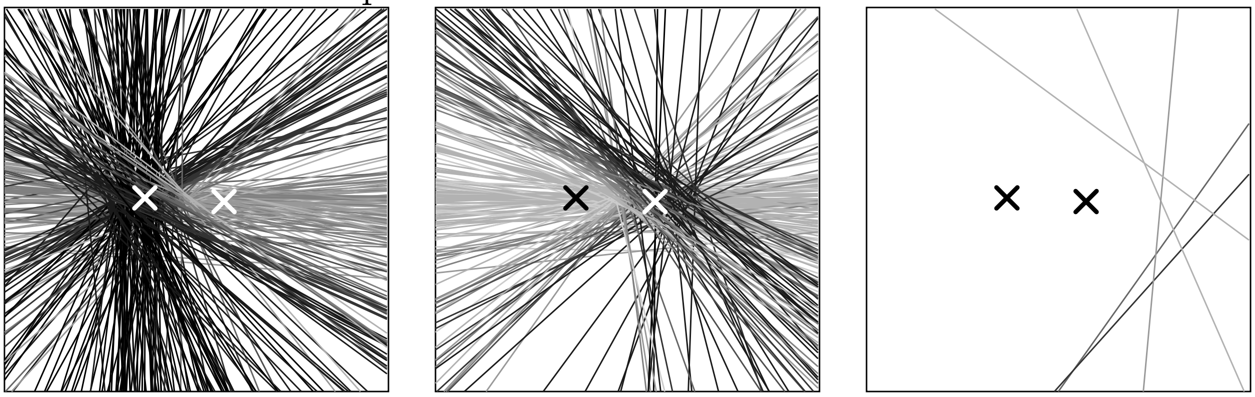
  $\vec{a}$ six dimensional

# Representation for Occlusion Example

A region containing an occlusion boundary:

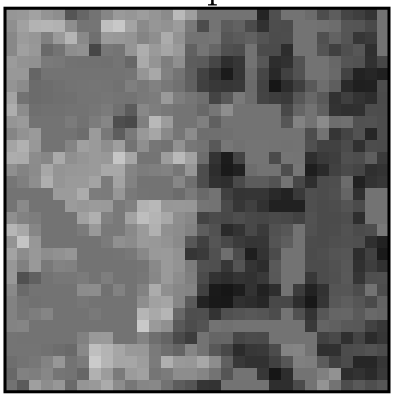



Constraint Ownership:



Layer #1 Layer #2 Outliers

Backprojected Constraint Ownership:



# Fitting Mixture Models to Data

Maximum likelihood fits using the EM-algorithm

Iterate:

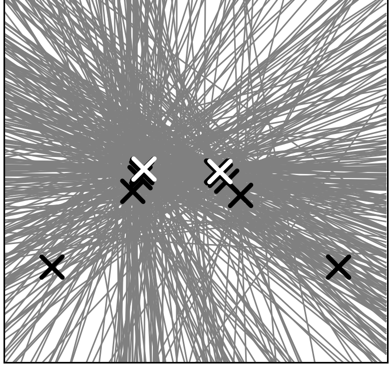
**E-step:**

* For each constraint $\vec{c}_k$, estimate the "ownership probability" for each of the layers.
* Simple computation involving the evaluation of $p_n(\vec{c}_k)$

**M-step:**

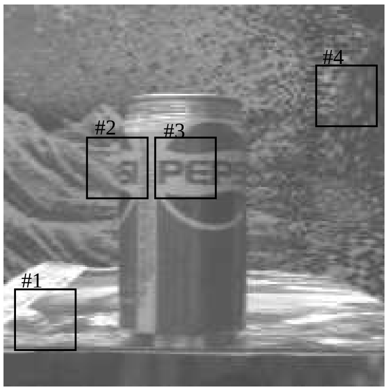
* Use the ownership probabilities as weights in a linear least squares problem for the layer parameters $\vec{a}_n$.

Convergence Behaviour:

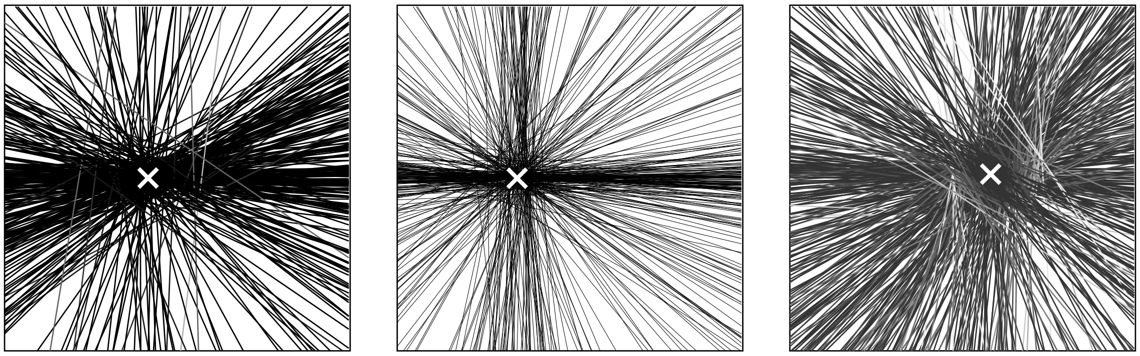


# Results for Additional Test Patches

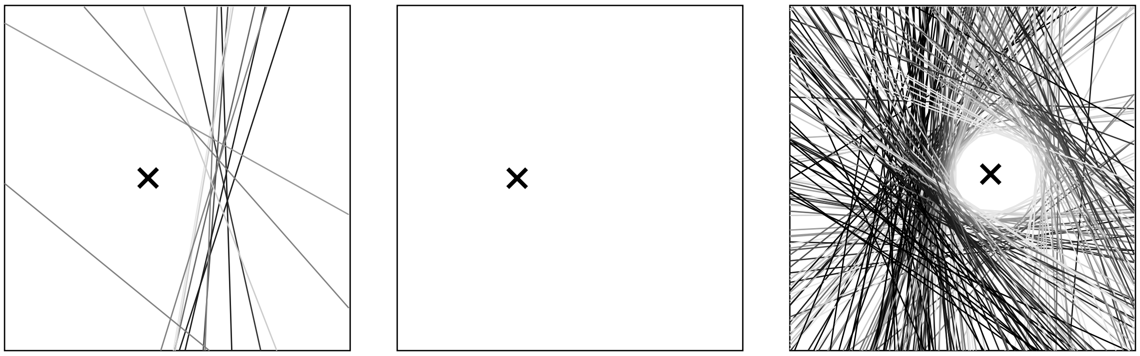
Test Patches:




Motion constraints for regions #1, #3 and #4:



Outliers for regions #1, #3 and #4:



# Mixtures of Affine Models

**Affine Model**: Optical flow $\vec{v}(\vec{x})$ given by a *linear* function of image coordinates $\vec{x}$:

$$\vec{v}(\vec{x}, \vec{a}) = \vec{a}_0 + A\vec{x}$$

**Shallow Structures**

- Small depth variation relative to mean distance.
- Small angular extent of view.

Optical flow for a shallow structure is well approximated by an affine model.

M-Step requires solution of 6D linear system for $\vec{a}$, for each component of the mixture distribution.

# Mixtures of Rational Models

**Rational Model**: Image position in new frame, $\vec{X}(\vec{x})$, given by a *rational* function of image coordinates $\vec{x}$:

$$\vec{X}(\vec{x};\vec{a},\vec{b}) \;=\; \left\{\frac{\vec{a}_0 + A\vec{x}}{1+\vec{b}\cdot\vec{x}}\right\}$$

**Thin Structures**

- Nearly planar 3D structure.
- Deviation from a plane is small relative to distance at that point.
- No restriction on angular extent of view.

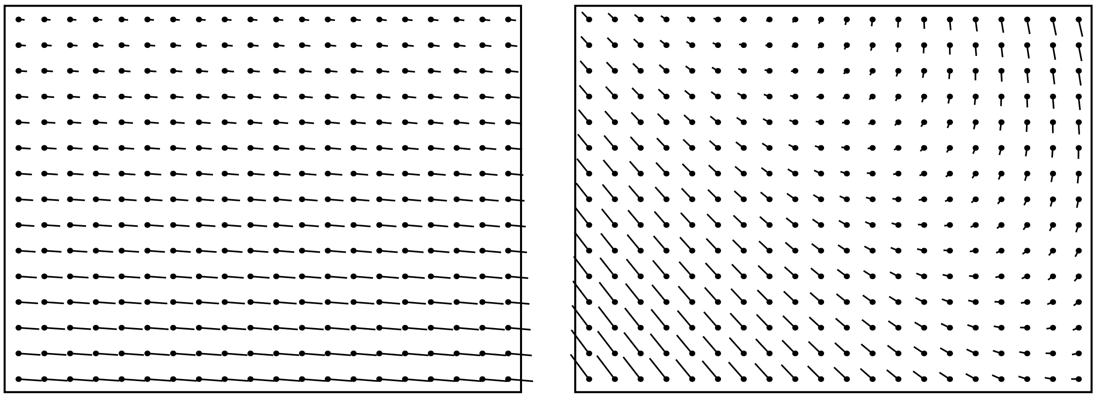
Image displacement mapping for a thin structure is well approximated by a rational model (exact for a planar surface).

M-Step requires solution of 8D nonlinear system for $\vec{a}$ and $\vec{b}$, for each component of the mixture distribution.

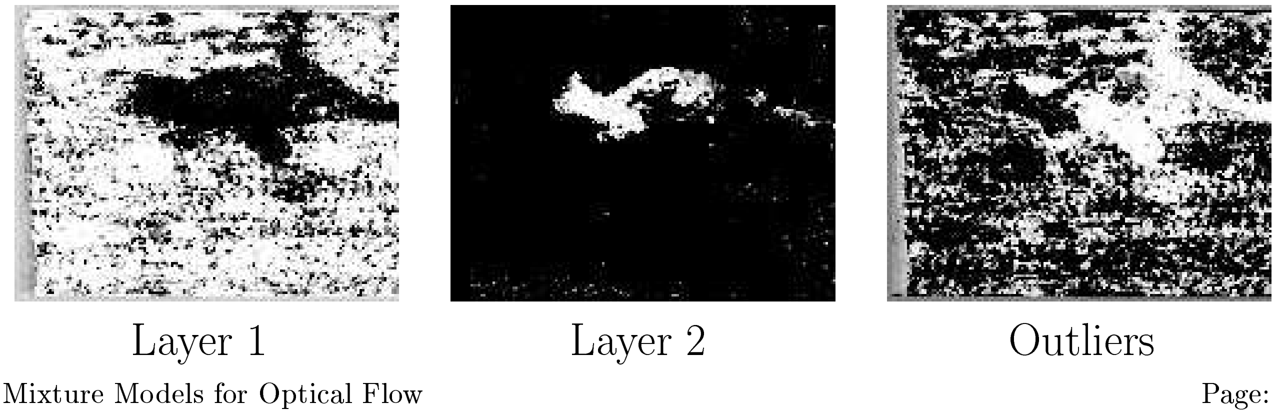
# Layered Motion Model

Two successive image frames:

Two affine layers fit to the dominant motion components:

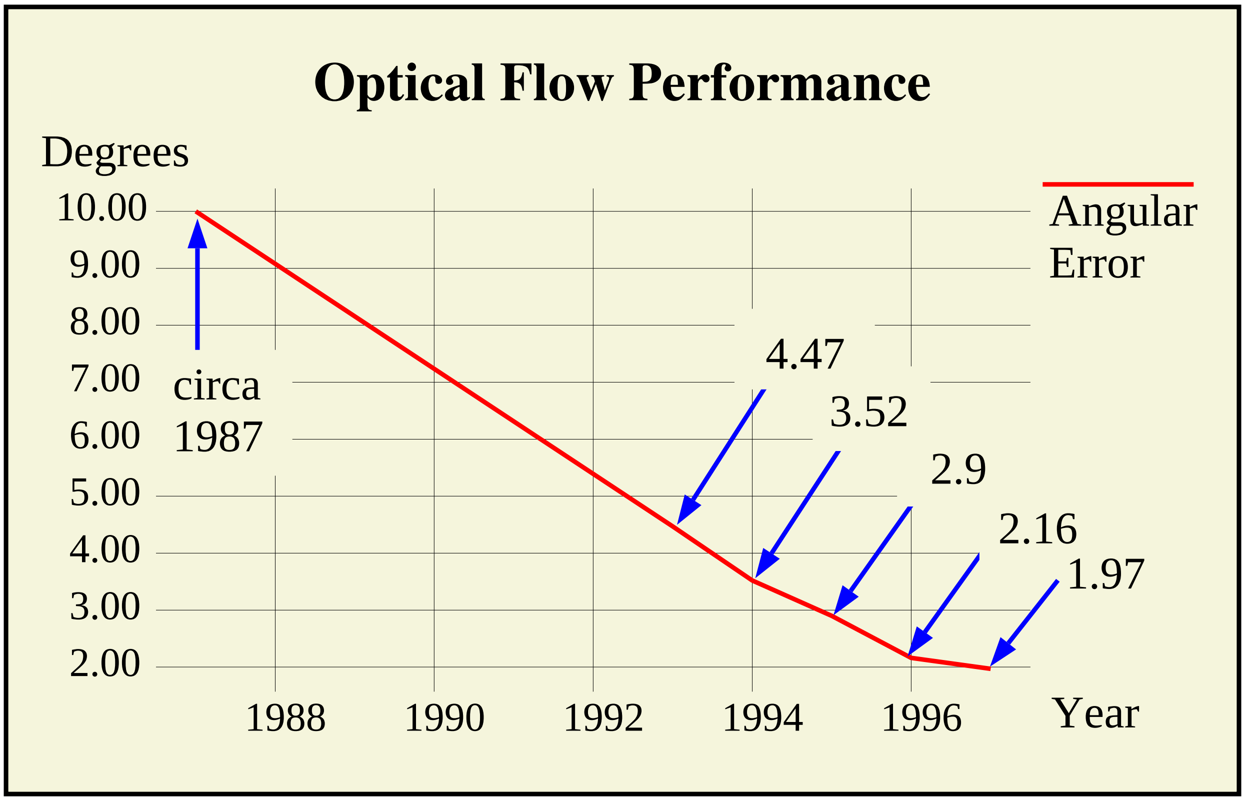


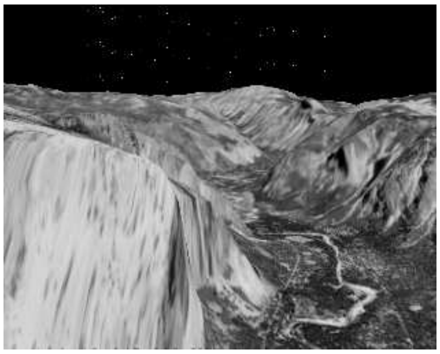
Ownership maps for the two motion layers plus outliers:

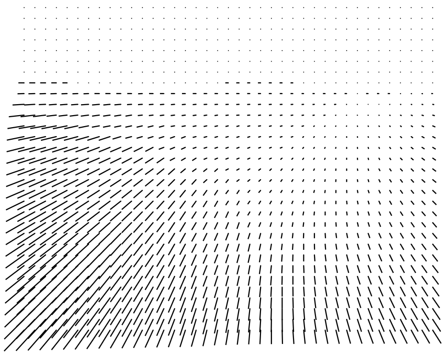


Layer 1 Layer 2 Outliers

# Yosemite Sequence





Ju, Black & Jepson, 1997

# Summary

Mixture models provide a robust and reliable tool for estimating motion or displacement parameters.

The approach:

- Relaxes the single motion assumption.
- Recovers multiple affine or rational motions.
- Detects outliers.

Other applications:

- Layered image models.
  - Stereo disparity.
  - Grey-level and colour.
  - Contour.
  - Depth.
  - 3D motion.
- Parameterized image features.
  - Local features (eg. Stereo disparity edges).
  - View-based object models.

Current research issues:

- Determining the number and type of the layers.
- Dealing with leverage problems.
- Spatial interactions of layer models.
- Spatial priors for pixel ownerships.
- Coupling motion with texture, colour, or stereo information.

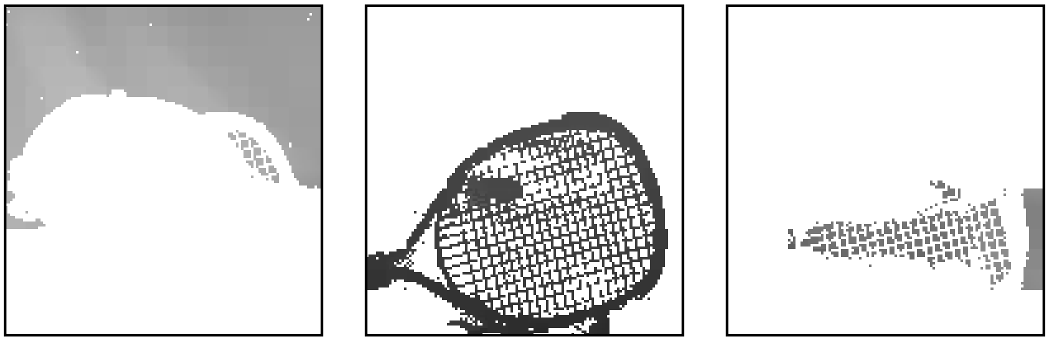
# Layered Image Models

- A natural representation for images which involve overwriting (eg. occlusion or overlays).
- The general task is to identify and separate individual layers given an image (or sequence).

**Intensity Image**

**Depth Image**

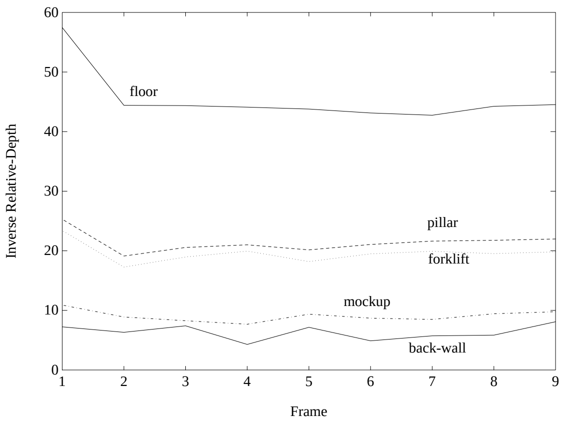
**Example Depth Layers**



- A critical technology which promises to underly much of early visual processing.

# Application: Image Motion and Structure

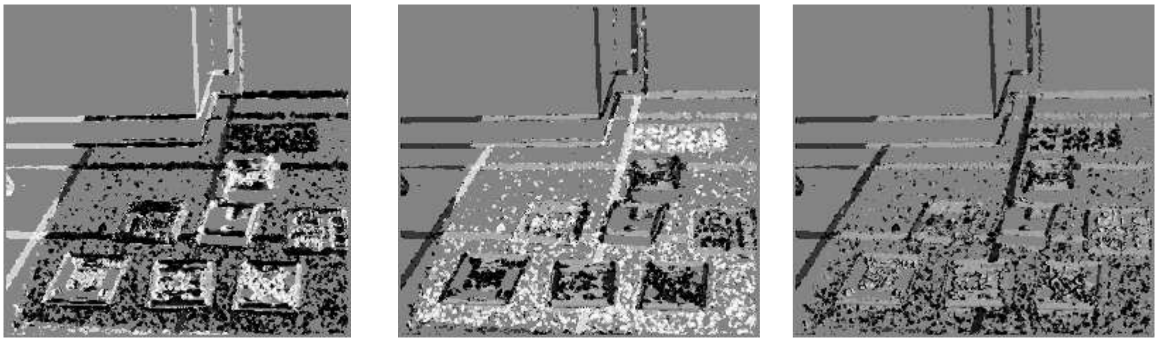
Frame from ‘forklift’ sequence




Recovered range.

# Application: Floor Anomaly Detection (FAD)

A calibrated stereo image pair.



Ownership for outliers, floor, and consistent.