Model-Based Hand Tracking Using an Unscented Kalman Filter

Björn Stenger and Roberto Cipolla
The Problem
Contributions

• What’s new?
  – Regh & Kanade, ECCV’94
    • Truncated cylinders, no self-occlusion, 10 Hz.
  – Heap & Hogg, F&G’96
    • Point mesh, PCA, invalid motions, 10 Hz.
  – Isard & McCormic, ECCV’00
    • 2D b-spline, partitioned sampling, real-time.
  – Wu, Lin & Huang, ICCV’01
    • Data-glove + PCA, MC, view-dependent.
  – Stenger, Mendonça, Cipolla, CVPR’01
    • Accurate model, self-occlusion, UKF, 12 Hz.
Contributions

- Construction of hand model from truncated quadrics
- Contour generation handling self-occlusion
- Application of Unscented Kalman filter
- Tracking 7 DOF with 12Hz using 2 cameras
Contributions

• Construction of hand model from truncated quadrics

• Contour generation handling self-occlusion

• Application of Unscented Kalman filter

• Tracking 7 DOF with 12Hz using 2 cameras
Contributions

• Construction of hand model from truncated quadrics

• Contour generation handling self-occlusion

• Application of Unscented Kalman filter

• Tracking 7 DOF with 12Hz using 2 cameras
Contributions

• Construction of hand model from truncated quadrics
• Contour generation handling self-occlusion
• Application of Unscented Kalman filter
• Tracking 7 DOF with 12Hz using 2 cameras
System Overview

Model Projection

- 3D Model configuration
- Project quadrics & handle occlusion
System Overview

Model Projection

3D Model configuration → Project quadrics & handle occlusion

Image Processing

Image Sequence → Detect edges
System Overview

Model Projection

3D Model configuration → Project quadrics & handle occlusion

Image Processing

Image Sequence → Detect edges

Unscented Kalman Filtering

Predict State vector → Compute geometric error → Predict Observation vector
System Overview

Model Projection

3D Model configuration → Project quadrics & handle occlusion

Image Processing

Image Sequence → Detect edges

Unscented Kalman Filtering

Predict State vector → Compute geometric error → Update State vector

Predict Observation vector
Quadrics

**Quadric**: Second degree implicit surface defined by points $X$ satisfying $X^TQX = 0$.

- $\text{rank}(Q) = 4$
- $\text{rank}(Q) = 3$
- $\text{rank}(Q) = 2$

Ellipsoid

Cone

Cylinder

Pair of Planes
Shaping Quadrics

The shape of the quadric appear by factorizing \( Q: \)

\[
Q = \begin{bmatrix}
R & 0 \\
0 & 1
\end{bmatrix}^T \begin{bmatrix}
1/a & 0 & 0 & 0 \\
0 & 1/b & 0 & 0 \\
0 & 0 & 1/c & 0 \\
0 & 0 & 0 & 1/d
\end{bmatrix} \begin{bmatrix}
R^T \\
t \\
0^T \\
1
\end{bmatrix}
\]

\text{Shape} \text{ Motion}
Clipping Quadrics

For modelling more general shapes, truncate quadrics by finding points $X$ which satisfy:

$$X^T Q X = 0 \quad \text{and} \quad X^T \Pi X \geq 0$$
Hand Model

• 37 truncated quadrics
• 27 degrees of freedom (currently only 7 are tracked)
Projection of a Quadric

Assuming a normalized projective camera \( \mathbf{P} = [\mathbf{I} \mid 0] \)

Parameterize 3D points \( \mathbf{X}(s) = \begin{bmatrix} \mathbf{x} \\ s \end{bmatrix} \)

\[
\mathbf{X}^T(s) \mathbf{Q} \mathbf{X}(s) = 0
\]
Projection of a Quadric (2)

\[
X^T(s) \quad Q \quad X(s) = 0
\]

\[
\begin{bmatrix} x & s \end{bmatrix} \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = 0
\]

\[
cs^2 + 2b^Tx + x^TAx = 0
\]

Condition for \( X(s) \) to be on the contour generator of \( Q \):

\[
\Delta = 0 \iff x^T(cA - bb^T)x = 0
\]

\[
x^T C x = 0 \quad C = cA - bb^T
\]
Projecting the Hand Model

3D model

Contours
Optimal Filtering

• Given a state-space model

\[ x_k = f(x_{k-1}, v_{k-1}) \]
\[ z_k = h(x_k, w_k) \]

• Maximum likelihood estimator

\[ \hat{x}_k = \text{arg min} \left( -\log L(x_k \mid z_k, \ldots, z_0) \right) \]

• For linear models, Gaussian error model: Kalman filter

\[ \hat{x}_k = \hat{x}_{k|k-1} + K_k (z_k - \hat{z}_{k|k-1}) \]
The Unscented Transform


Given: \( n \)-dimensional random variable \( x_{k-1} \) with mean \( \hat{x}_{k-1} \) and covariance \( P_{k-1} \).

1. Compute \( 2n+1 \) points with associated weights.
2. Apply the nonlinear transformation to each point.
3. Compute mean and covariance of the transformed points.
The Unscented Transform

1. Choose $2n+1$ points with associated weights

$$X_{k-1}^i = \begin{cases} \hat{x}_{k-1} & i = 0 \\ \hat{x}_{k-1} - \sigma_{k-1}^i & i = 1, \ldots, n \\ \hat{x}_{k-1} + \sigma_{k-1}^i & i = n+1, \ldots, 2n \end{cases}$$

where $\sigma_{k-1}^i$ is the $i^{th}$ column of the matrix $\sqrt{(n+1)P_{k-1}}$

The set of points has the same mean, covariance and all higher odd central moments as the Gaussian distribution of $x_{k-1}$. 
2. Apply the nonlinear transformation to each point

\[ X_{k|k-1}^i = f(X_{k-1}^i, k) \quad i = 0, \ldots, 2n \]

3. Compute mean and covariance of the transformed points.

This approximation is correct up to the second order.
Comparison

Actual (sampling)  Linearisation  Unscented Transform
Properties of the UT

• Approximates the distribution rather than the nonlinearity.

• Accurate to at least 2\textsuperscript{nd} order (3\textsuperscript{rd} order for Gaussian distributions).

• No Jacobian or Hessian matrices are needed.

• Efficient “sampling approach”.

• Assumes unimodal distributions.
Unscented Kalman Filter

Prediction Step:
• Prediction of state (and error covariance matrix)
  \[
  \hat{\mathbf{x}}_{k-1} \rightarrow \hat{\mathbf{x}}_{k|k-1}
  \]
• Prediction of observation
  \[
  \hat{\mathbf{z}}_{k-1} \rightarrow \hat{\mathbf{z}}_{k|k-1}
  \]

Measurement Update Step:
• Compute innovation
  \[
  \nu_k = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}
  \]
• Compute Kalman gain \( K_k \)
• Update state estimation (and error covariance matrix)
  \[
  \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + K_k \nu_k
  \]
State & Observation Vector

State Vector $x_k$:
- Global pose parameters (6 DOF)
- Configuration of joints (1 DOF)
- Velocity and acceleration

$$x_k = [x_0, ..., x_l, \dot{x}_0, ..., \dot{x}_l, \ddot{x}_0, ..., \ddot{x}_l]^T$$

Observation Vector $z_k$:
- Local edge detection at contour points

$$z_k = \begin{bmatrix} n_0^T s_0 \\ \vdots \\ n_m^T s_m \end{bmatrix}$$
Single View Tracking
Stereo Tracking 1
Stereo Tracking 2
Conclusions & Future Work

• Construction of hand model from quadrics
• Handling self-occlusion
• Application of Unscented Kalman filtering
• Tracking 7 DOF with 12 Hz
Conclusions & Future Work

• Construction of hand model from quadrics
• Handling self-occlusion
• Application of Unscented Kalman filtering
• Tracking 7 DOF with 12 Hz

• Assess failure modes
• Increase DOFs
• Learn model shape
• Handle cluttered backgrounds
Conclusions & Future Work

• Construction of hand model from quadrics
• Handling self-occlusion
• Application of Unscented Kalman filtering
• Tracking 7 DOF with 12 Hz

• Assess failure modes
• Increase DOFs
• Learn model shape
• Handle cluttered backgrounds

Thanks!