Bayesian Filter based Object Traction in Computer Vision

Liang Xu, xul4@vcu.edu

Department of Electrical and Computer Engineering

Virginia Commonwealth University

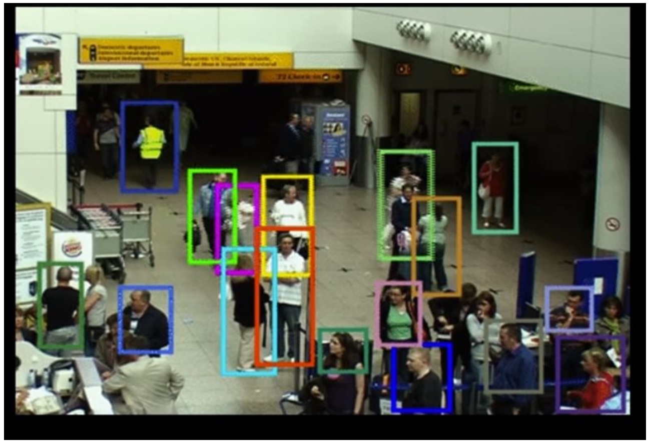
Adviser: Dr. Weijun Xiao

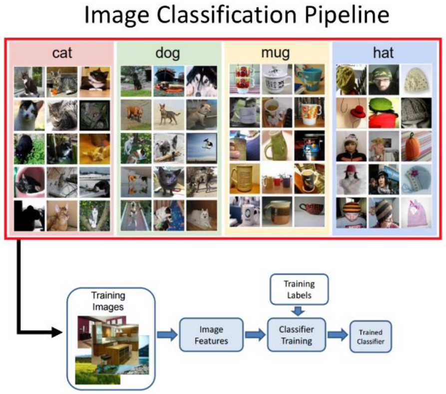
Content

- Introduction
- Particle Filter based Object Traction
- Kalman Filter based Simultaneous Localization and Mapping (SLAM)
- Conclusion

Introduction

- Camera is the most widely used sensor.
- Camera is very cheap
- An Image contain lots of information.

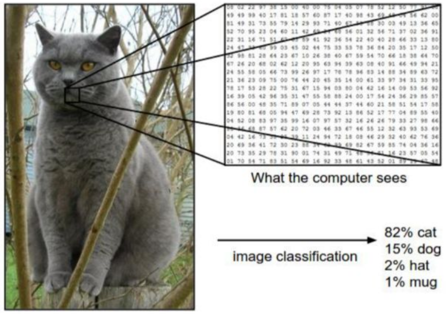




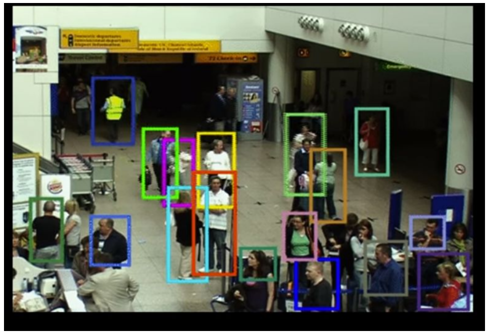
Introduction

- Images are Numbers

Images are Numbers



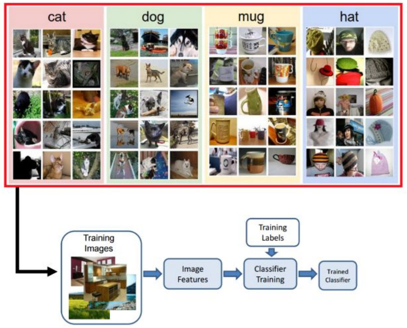
Object Traction



Moving Object: Kalman Filter

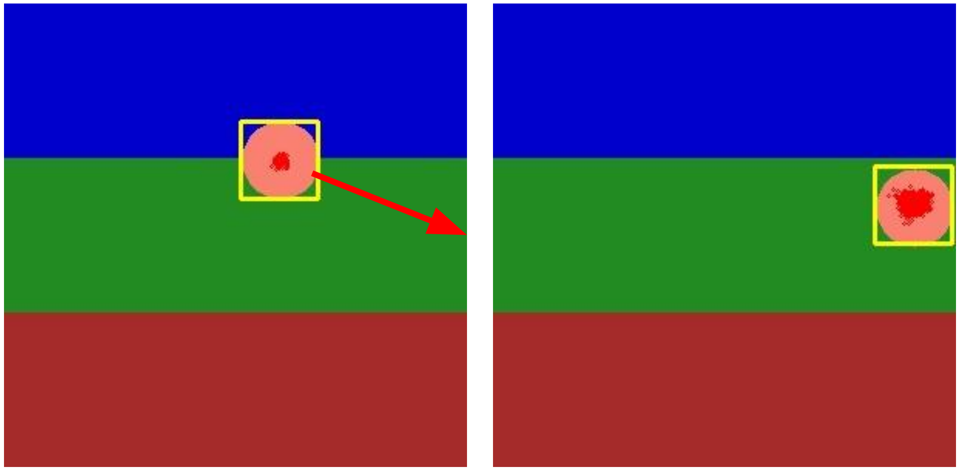
Object Classification/reconsization

Image Classification Pipeline

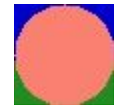


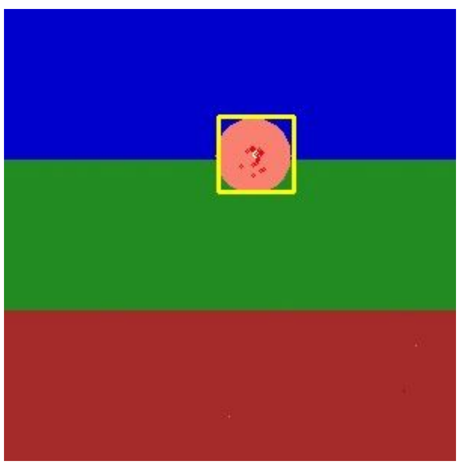
Deep Learning

Particle Filter based Object Traction



Object

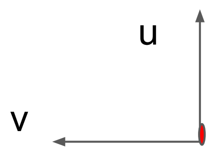




1. Object is the “thing” that is actually being tracked.
2. x(t) represent the state of the model at time t.
3. A dynamic model p(x(t)|x(t-1)) distribution of the state at time t given the state at t − 1
4. a measurement z(t) that somehow captures the data of the current image.
5. a sensor model p (z (t) | x (t)) that gives the likelihood of a measurement given the state.

$$Bel(x_t) \propto p(z_t|x_t)p(x_t|u_t,x_{t-1})Bel(x_{t-1})$$

Particle Filter based Object Traction

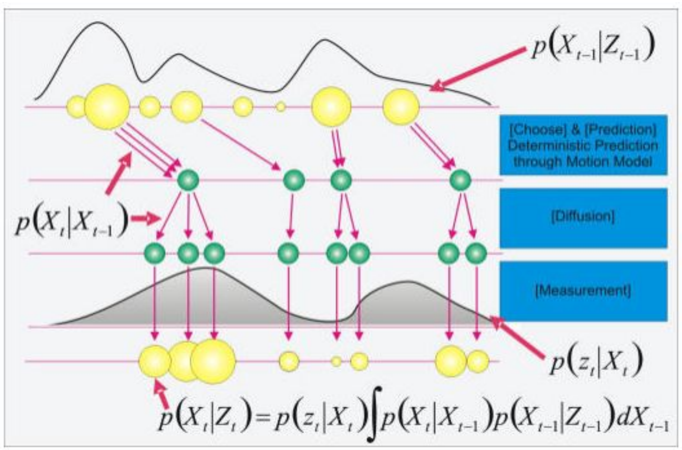


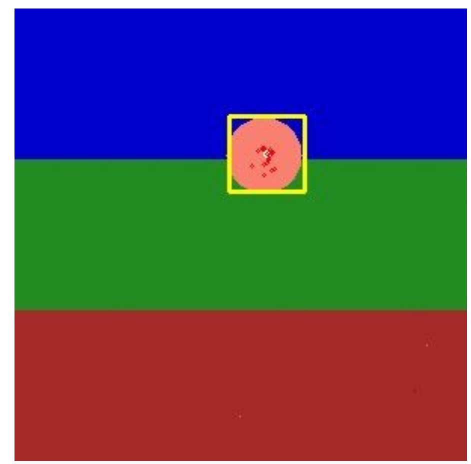
Each Particle have 4 states.
u, v , u speed , v speed.

$$MSE(u_p, v_p) = \frac{1}{mn}\sum_{u=1}^{m}\sum_{v=1}^{n}(Template(u,v) - Image^t(u+u_p-m/2, v+v_p-n/2))^2$$

Important sampling

$$p(z_t|x_t) \propto \exp\left(-\frac{MSE}{2\sigma^2_{MSE}}\right)$$





Particle Filter based Object Traction

Objects are Changing

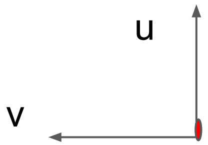
$$Template(t) = \alpha\, Best(t) + (1-\alpha) Template(t-1)$$

where Best(t) is the patch of the best estimate or mean estimate. It's easy to see that by recursively updating this sum, the window implements an exponentially decaying weighted sum of (all) the past windows.

Particle Filter based Object Traction

Particle Filter based Object Traction

Objects are Changing, and the scale of object also change, and occlusion



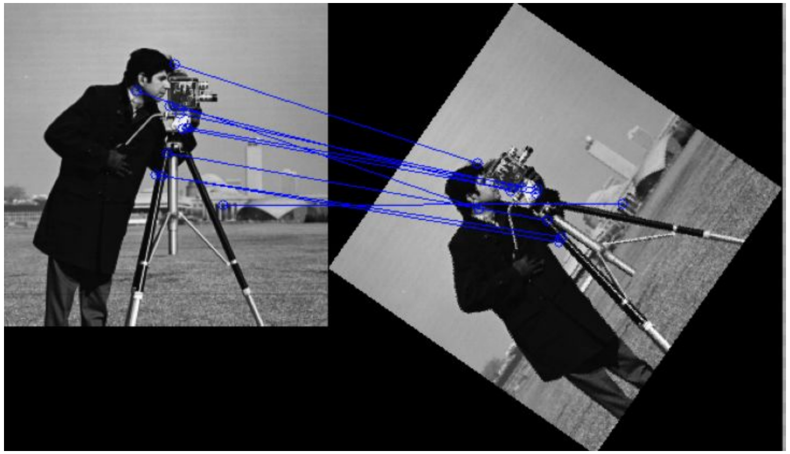
Each Particle have 5 states.
u, v , u speed , v speed, object scale

Particle Filter based Object Traction

Particle Filter based Object Traction

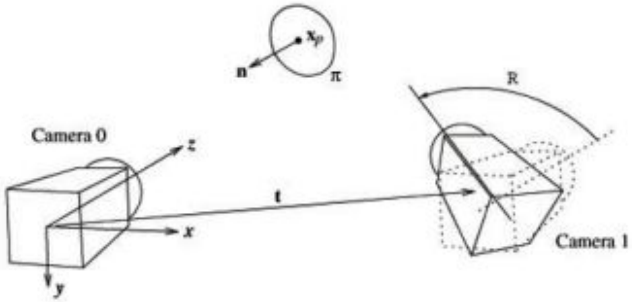
Multi Object Tracking

Kalman Filter based Simultaneous Localization and Mapping (SLAM)

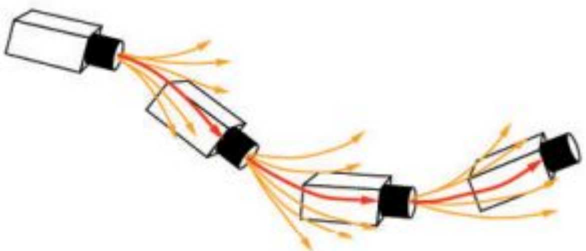


Scale-invariant feature transform (SIFT)

Can we use those Key points to estimate the camera's location?



Kalman Filter based Simultaneous Localization and Mapping (SLAM)

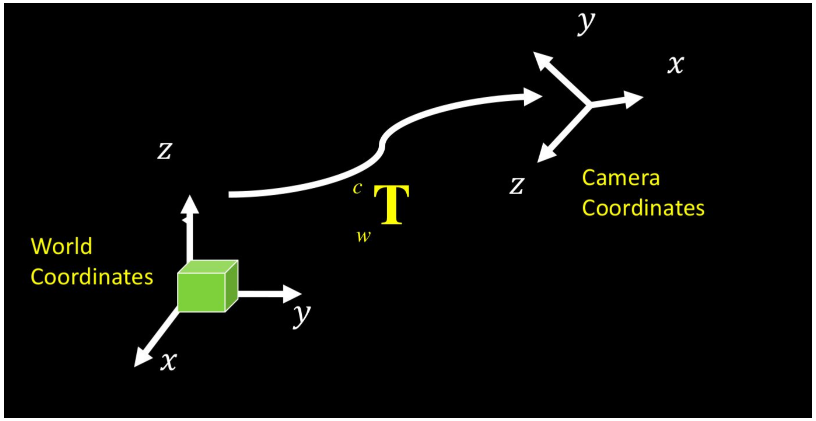


$$\mathbf{f}_v = \begin{pmatrix} \mathbf{r}^W_{new} \\ \mathbf{q}^{WR}_{new} \\ \mathbf{v}^W_{new} \\ \omega^R_{new} \end{pmatrix} = \begin{pmatrix} \mathbf{r}^W + (\mathbf{v}^W + \mathbf{V}^W)\Delta t \\ \mathbf{q}^{WR} \times \mathbf{q}((\omega^R + \Omega^R)\Delta t) \\ \mathbf{v}^W + \mathbf{V}^W \\ \omega^R + \Omega^R \end{pmatrix} .$$

$$\mathbf{x}_v = \begin{pmatrix} \mathbf{r}^W \\ \mathbf{q}^{WR} \\ \mathbf{v}^W \\ \omega^R \end{pmatrix} .$$

$\mathbf{r}^W$	3D position vector
$\mathbf{q}^{WR}$	orientation quaternion
$\mathbf{v}^W$	velocity vector
ω^R	angular velocity vector

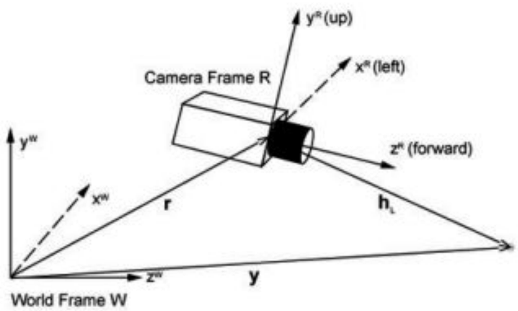
Kalman Filter based Simultaneous Localization and Mapping (SLAM)

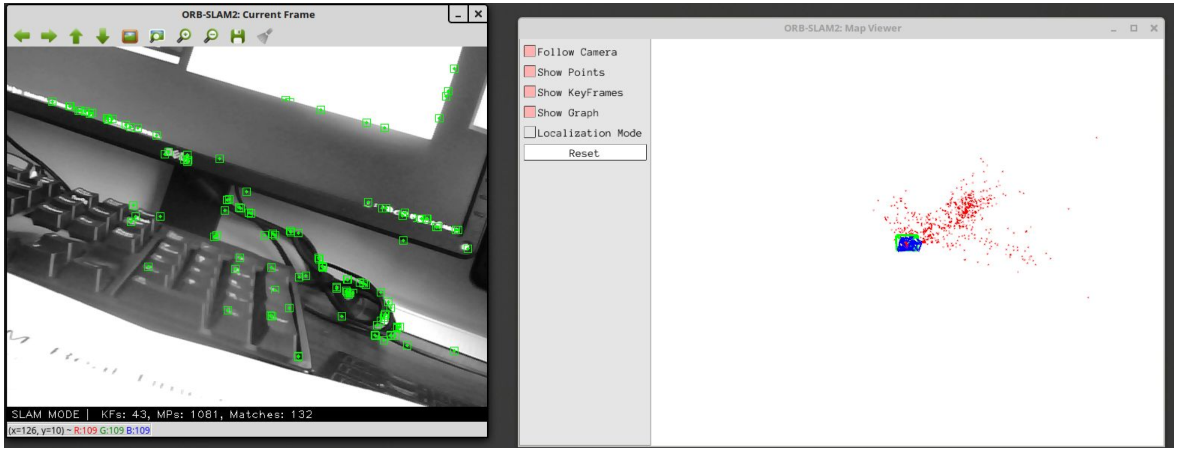


Extrinsic Camera Calibration

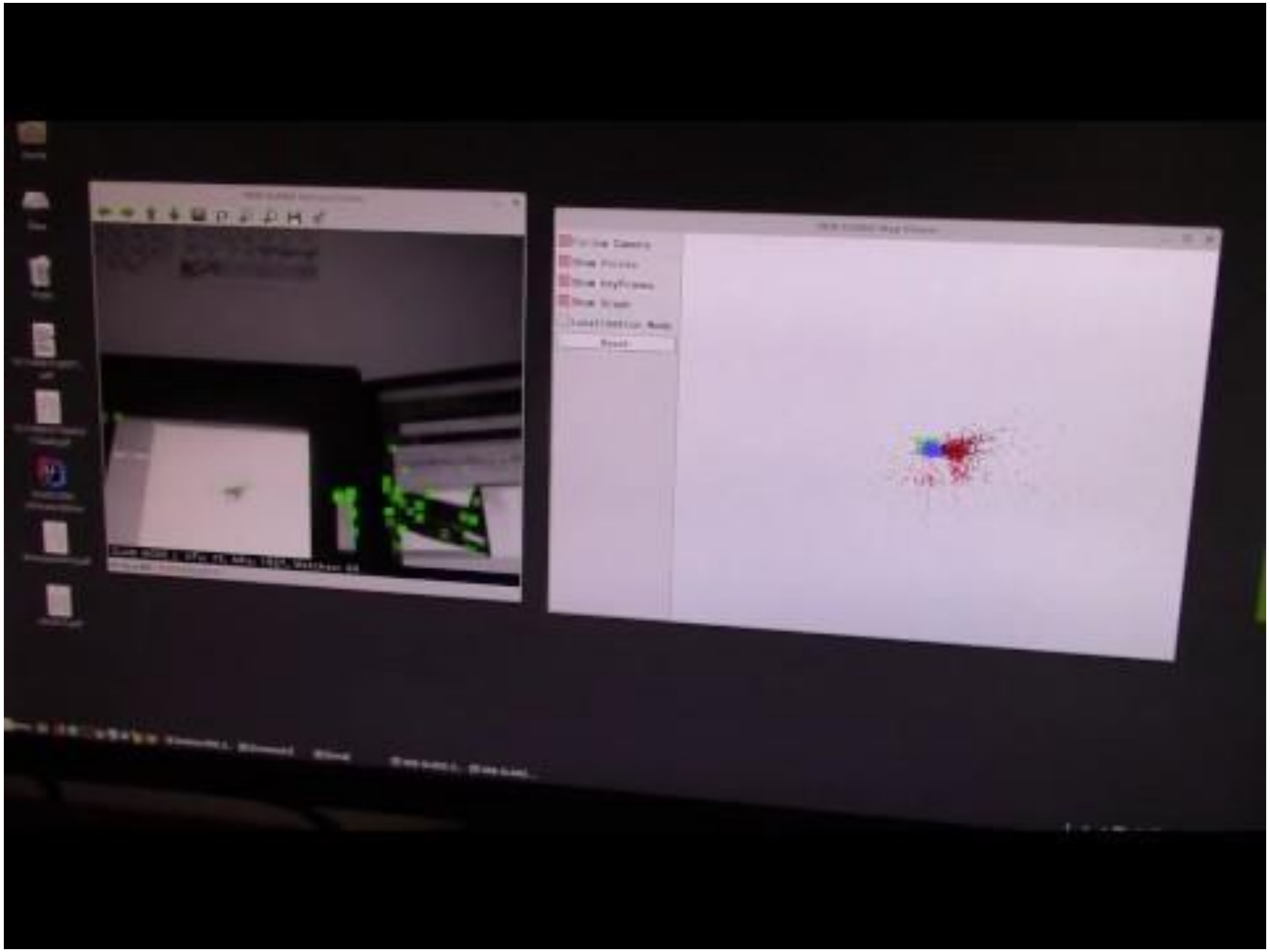
Intrinsic camera calibration

Kalman Filter based Simultaneous Localization and Mapping (SLAM)





Kalman Filter based Simultaneous Localization and Mapping (SLAM)



Questions?