Robust Visual Tracking Using Compressed Sensing

Jainisha Sankhavara (201311002)
Falak Shah (201311024)

MTech, DA-IICT

April 16, 2014

Outline

Introduction

- Just an attempt to explain the implementation of visual tracking as explained in
  [1] Hanxi Li; Chunhua Shen; Qinfeng Shi, "Real-time visual tracking using compressive sensing," Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on , vol., no., pp.1305,1312, 20-25 June 2011
  [2] Xue Mei; Haibin Ling, "Robust visual tracking using l1 minimization," Computer Vision, 2009 IEEE 12th International Conference on , vol., no., pp.1436,1443, Sept. 29 2009-Oct. 2 2009

Introduction

- Tracking object through frames video in real time.
- Assume initial position of object available in the first frame
- Challenges: Occlusion,illumination changes, shadows, varying viewpoints, etc

Particle Filter Definition

- Posterirori density estimation algorithm
- There is some unknown we are interested in called state variable (eg. location of object)
- We can measure something (measurement variable), related to the unknown variable
- Relation between state variable and measurement variable known.

Particle Filter Visualisation

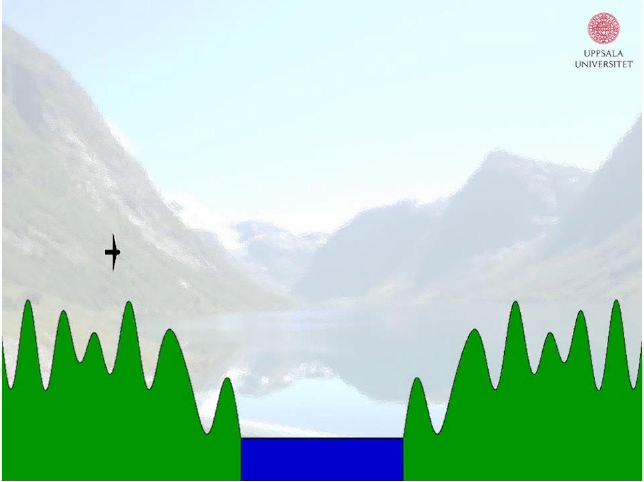

Figure 1: Plane moving- Position unknown [1]

[1] Andreas Svensson, Ph.D Student, Uppsala University

Particle Filter Visualisation

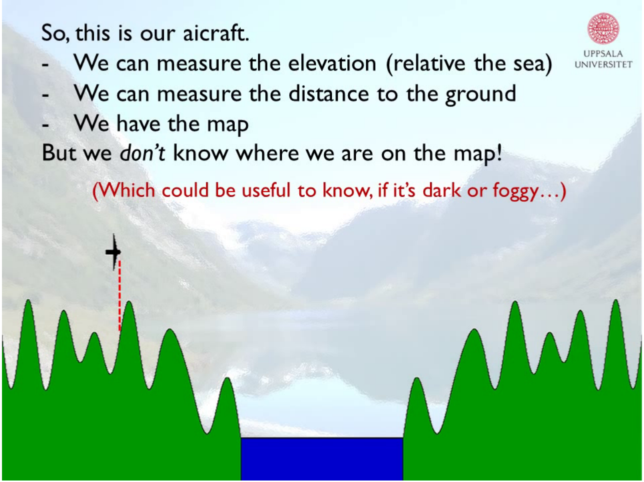

Figure 2: Available data

Particle Filter Visualisation

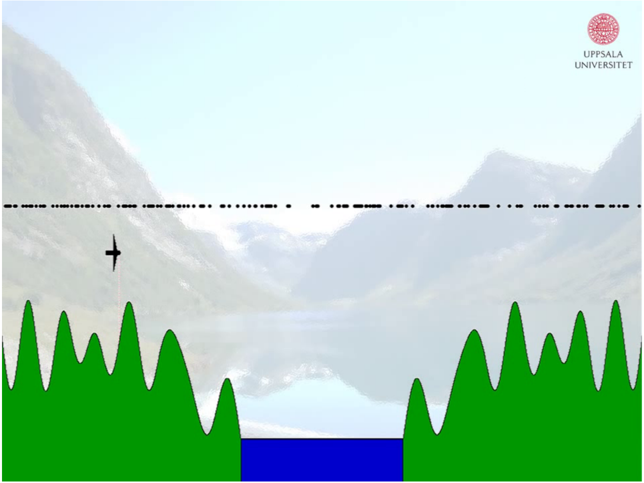

Figure 3: Initial distribution of particles

Particle Filter Visualisation

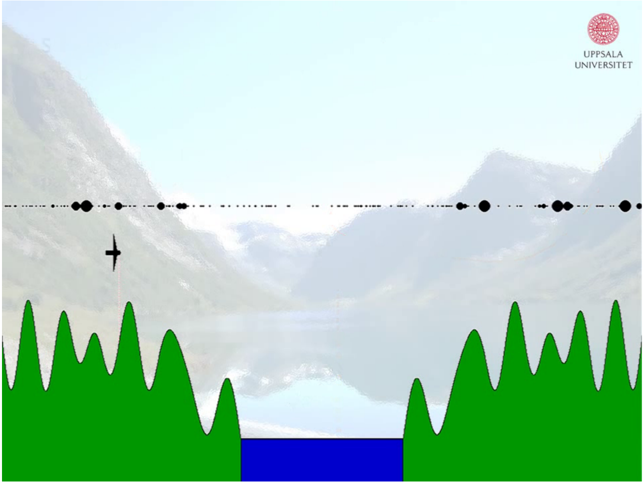

Figure 4: Observation Likelihood

Particle Filter Visualisation

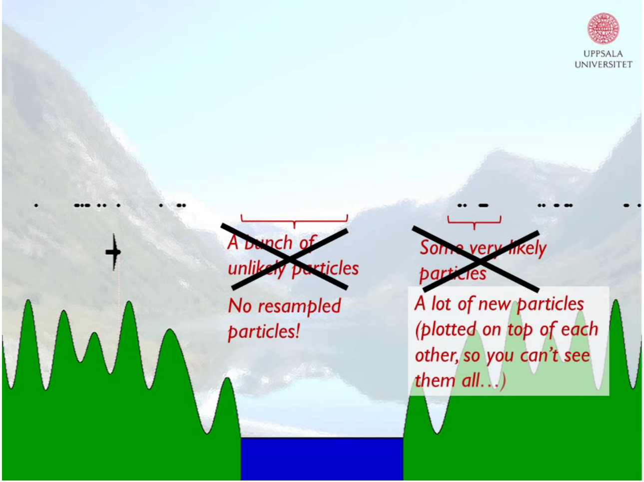

Figure 5: Resampling Step

Particle Filter Visualisation

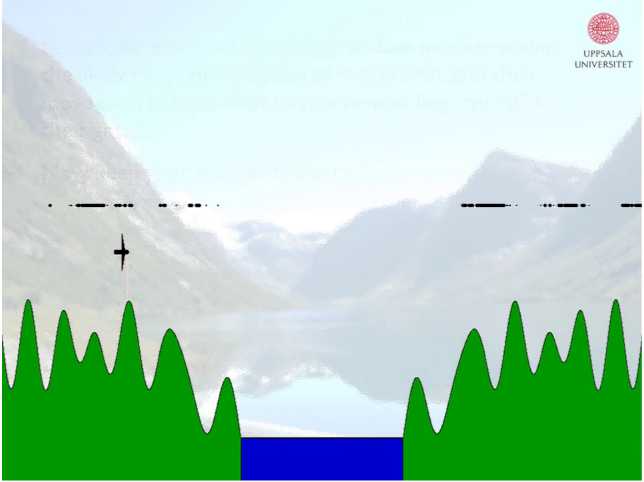

Figure 6: Posteriori estimate

Particle Filter Visualisation

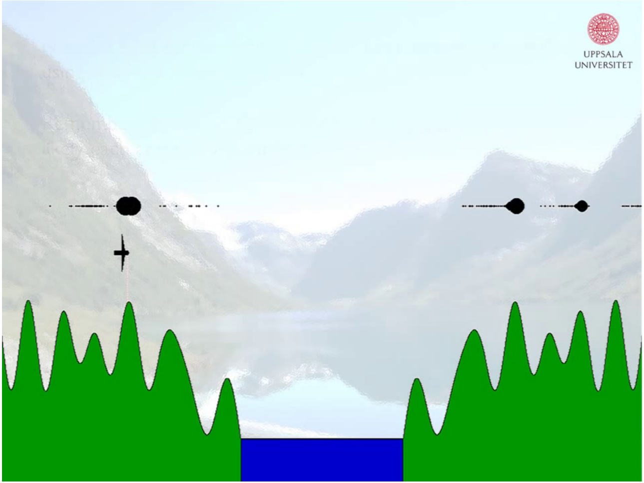

Figure 7: Observation likelihood: step 2

Particle Filter Visualisation

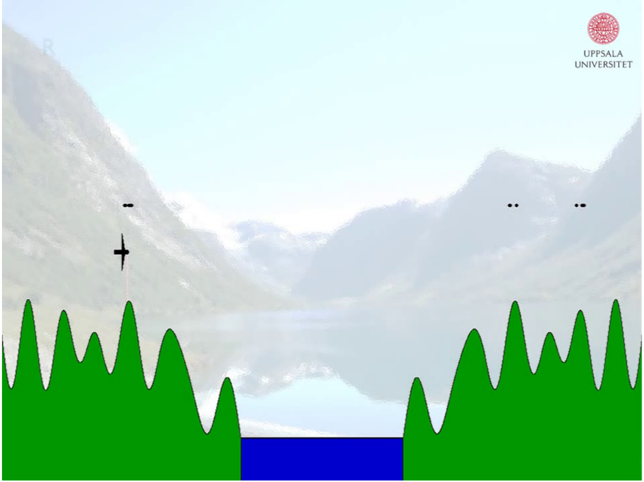

Figure 8: Resample: step 2

Particle Filter Visualisation

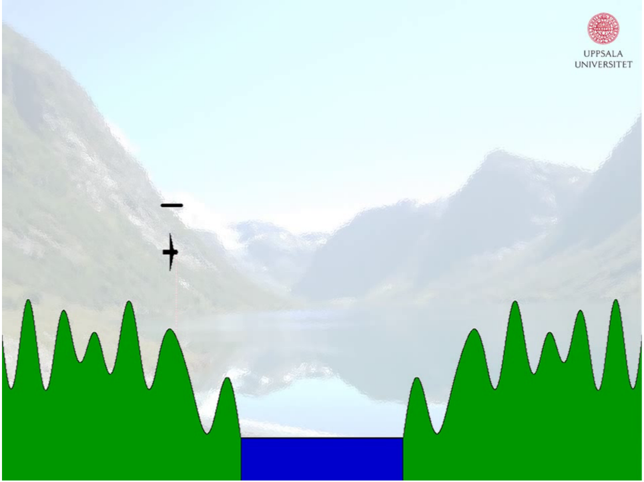

Figure 9: Particles converge very close to object

Particle Filter visulisation

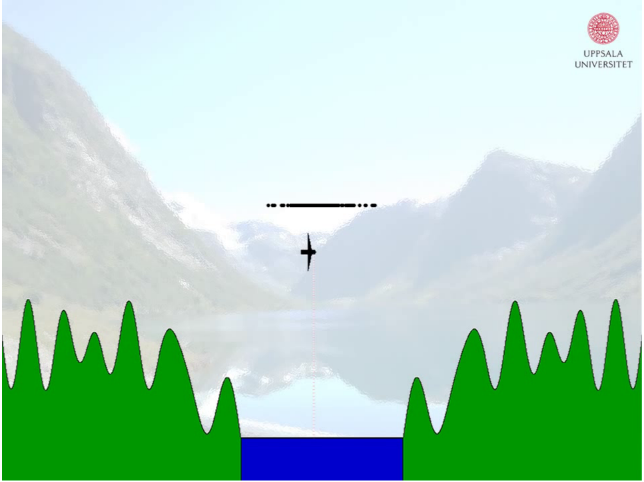

Figure 10: Issues

Particle Filter Visualisation

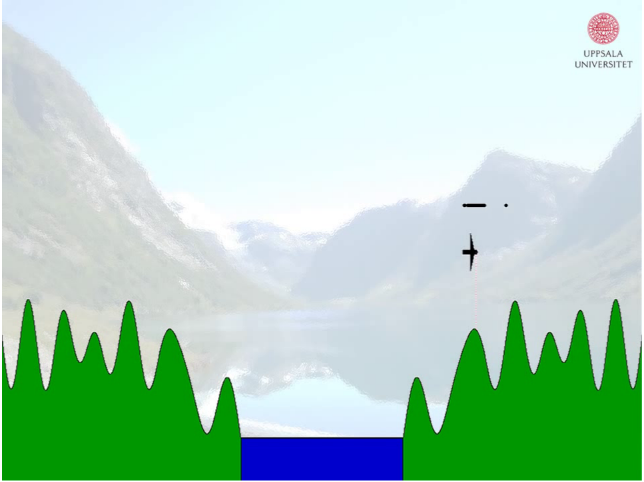

Figure 11: Back to tracking

Particle Filter Equations

- x_t - state variable
- z_t - observation at time t
- x_t is modeled by six parameters of affine transformations. $x_t = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, t_x, t_y)$
- All six parameters are independent.
- State transition model $p(x_t|x_{t-1})$ is gaussian.
- $p(z_t|x_t)$ is also gaussian.

Particle Filter Equations

- state vector prediction

$$p(x_t|z_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|z_{1:t-1})dx_{t-1}$$

- state vector update

$$p(x_t|z_{1:t}) = \frac{p(z_t|x_t)p(x_t|z_{1:t-1})}{p(z_t|z_{1:t-1})}$$

- weight update

$$w_t^i = w_{t-1}^i \frac{p(z_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t|x_{1:t-1}, z_{1:t})}$$

Jainisha Sankhavara (201311002)
Falak Shah (201311024)

Template Dictionary

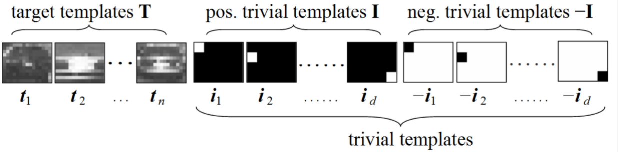

Figure 12: Target and Trivial Templates [2]

- Represent each of the particles as a linear combination of target templates and trivial templates.

Equation

$$y = \begin{bmatrix} T & I & -I \end{bmatrix} \begin{bmatrix} a \\ e^+ \\ e^- \end{bmatrix}$$

where,

- T $= (t_1;\ t_2\ \ldots\ ;\ t_n) \in R^{dxn}$ (d $\gg$ n) is the target template set, containing n target templates such that each template $t_i \in R^d$.
- a $= (a_1;\ a_2\ \ldots\ ;\ a_n)^T \in R^n$ is called a target coefficient vector and
- $e^+ \in R^d$ and $e^- \in R^d$ are called a positive and negative trivial template coefficient vectors.
- A tracking result y $\in R^d$ approximately lies in the linear span of T.

Underdetermined system

- No unique solution
- For a good target candidate, there are only a limited number of nonzero coefficients in e^+ and e^-

$$min \parallel Ax - y \parallel_2^2 + \lambda \parallel c \parallel_1$$

where,

- A = [T, I,-I] $\in R^{d\times(n+2d)}$
- x = [a; e^+ ; e^-] $\in R^{(n+2d)}$ is a non-negative coefficient vector.

Template Update

- Template replacement : If the tracking result y is not similar to the current template set T, it will replace the least important template in T.
- Template updating: It is initialized to have the median weight of the current templates.
- Weight update: The weight of each template increases when the appearance of the tracking result and template is close enough and decreases otherwise.

Dimensionality Reduction

- l_1 tracker

$$min \parallel x \parallel_1^2 \ s.t. \parallel Ax - y \parallel_2^2 \leq \epsilon$$

- Dimensionality reduction if the measurement matrix ϕ follows the Restricted Isometry Property (RIP) [1], then a sparse signal x can be recovered from

$$min \parallel x \parallel_1^2 .s.t. \parallel \phi Ax - \phi y \parallel_2 \leq \epsilon, x \geq 0.$$

where, $\phi \in R^{d_0 x d} \ d_0 \ll d \ and \ \phi_j \sim N(0,1)$

[1] E. Cand'es, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," Communications on Pure and Applied Mathematics, vol. 59, pp. 1207–1223, 2006.

OMP

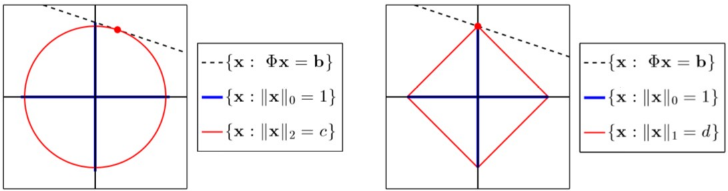

Figure 13: l_1 norm minimization

OMP

begin

  Initialize the residual $\mathbf{r}_0 = \mathbf{y}$, index set $\Lambda_0 = \varnothing$ and selected template set $\Psi_0 = \varnothing$;

  for $t \leftarrow 1$ **to** η **do**

    $\lambda_t = \operatorname*{argmax}_{j=1,\ldots,n} \langle r_{t-1}, \mathbf{a}_j \rangle$;

    $\Lambda_t = \Lambda_{t-1} \cup \{\lambda_t\}$;

    $\Psi_t = [\Psi_{t-1} \;\; \mathbf{a}_{\lambda_t}]$;

    Solve the least-squares problem:

$$\mathbf{x}_t = \operatorname*{argmin}_{\mathbf{x}} \|\Psi_t \mathbf{x} - \mathbf{y}\|_2;$$

    Calculate the new residual:

$$\mathbf{r}_t = \mathbf{y} - \Psi_t \mathbf{x}_t \;;$$

    if $\|\mathbf{r}_t\|_2 < \varepsilon$ **then** break;

  end

  Retrieve signal $\mathbf{x}$ according to $\mathbf{x}_t$ and Λ_t;

end

Figure 14: Customize OMP algorithm [1]

Conclusion

- The RTCST tracker achieves higher accuracy than existing tracking algorithms, i.e., the PF tracker.
- Dimension reduction methods and a customized OMP algorithm enable the CS-based trackers to run real-time.

References

- Hanxi Li; Chunhua Shen; Qinfeng Shi, "Real-time visual tracking using compressive sensing," Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on , vol., no., pp.1305,1312, 20-25 June 2011
- Xue Mei; Haibin Ling, "Robust visual tracking using l1 minimization," Computer Vision, 2009 IEEE 12th International Conference on , vol., no., pp.1436,1443, Sept. 29 2009-Oct. 2 2009
- D. Donoho, "For Most Large Underdetermined Systems of Linear Equations the Minimal l1-Norm Solution Is Also the Sparsest Solution," Comm. Pure and Applied Math., vol. 59, no. 6, pp. 797-829, 2006.
- E. Cande's and T. Tao, "Near-Optimal Signal Recovery from Random Projections: Universal Encoding Strategies?" IEEE Trans. Information Theory, vol. 52, no. 12, pp. 5406-5425, 2006.
- Wright, J.; Yang, A.Y.; Ganesh, A.; Sastry, S.S.; Yi Ma, "Robust Face Recognition via Sparse Representation," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.31, no.2, pp.210,227, Feb. 2009
- E. Cand'es, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," Communications on Pure and Applied Mathematics, vol. 59, pp. 1207–1223, 2006.
- A. Yilmaz, O. Javed, and M. Shah. 'Object tracking: A survey". ACM Comput. Surv. 38(4), 2006.