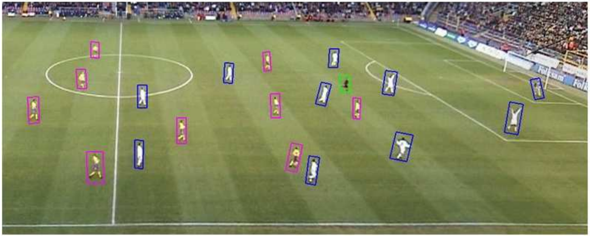
Identity Management on Homogeneous Spaces

Xiaoye Jiang and Leonidas J. Guibas
Stanford University

Algebraic Methods in Machine Learning Workshop for NIPS 2008

December 12, 2008

Ranking, Voting and Tracking



Problem in Identity Management

Problem in Identity Management

Markov Model for Identity Management

$$
\begin{array}{ccccccccc}
\sigma_0 & \xrightarrow{M_0} & \sigma_1 & \xrightarrow{M_1} & \sigma_2 & \xrightarrow{M_2} & \sigma_3 & \xrightarrow{M_3} & \cdots \\
\Big\downarrow{\scriptstyle L(z_0|\sigma_0)} & & \Big\downarrow{\scriptstyle L(z_1|\sigma_1)} & & \Big\downarrow{\scriptstyle L(z_2|\sigma_2)} & & \Big\downarrow{\scriptstyle L(z_3|\sigma_3)} & & \\
z_0 & & z_1 & & z_2 & & z_3 & &
\end{array}
$$

σ: true state; z: observations; M: markov matrix; $L(z|\sigma)$: likelihood function.

- Mixing Model: tracks swapped identities with some probability.
- Observation Model: identity on a particular track is observed.
- Problem: For each timestep, find posterior over σ_t conditioned on all past observations.
- Our Problem: Find posterior over **class characteristics** (red or blue) conditioned on all past observations.

Our Problem

- Define $\sigma^{(t)} \in S_n$ to be a mapping from identities $\{i_1, i_2, \cdots, i_{m+n}\}$ to tracks $T = \{t_1, t_2, \cdots, t_{m+n}\}$.
- After a random permutation among tracks $\tau^{(t)}$. The association of identities with tracks at time $t+1$ is $\sigma^{(t+1)} = \tau^{(t)}\sigma^{(t)}$.
- Assume n of the identities are red and the remaining m identities are blue.
- We care only about the **class characteristics** (red or blue) of identities.

Homogeneous Space

- **Homogeneous Space:** All $k-$subsets of $\{1, 2, \cdots, n\}$.
- Permutation groups act on homogeneous spaces.

Example

- Suppose $n = 3, k = 2$, homogeneous space X is all $2-$subset of $\{1, 2, 3\}$, i.e. $X = \{\{1, 2\}, \{2, 3\}, \{1, 3\}\}$.
- Permutation group S_3 acts on X, e.g., if

$$\tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

then $\tau(\{1, 2\}) = \{2, 3\}; \tau(\{2, 3\}) = \{1, 3\}; \tau(\{1, 3\}) = \{1, 2\}$.

Markov Process on Homogeneous Space

- A probability distribution Q on permutation groups induces a Markov process on the homogeneous space X with transition probability

$$P_x(y) = \sum_{\tau:\tau x = y} Q(\tau)$$

- **Naive Model:** Maintain beliefs on homogeneous space instead of full permutation group.

Running Example

Example (Markov Model on Homogeneous Space)

- Suppose $m = n = 3$ and we are sure that $\{t_1, t_2, t_3\}$ are red, then $f \in L(X)$

$$f(x) = \begin{cases} 1 & \text{if } x = \{t_1, t_2, t_3\} \\ 0 & \text{otherwise} \end{cases}$$

- If a mixing happened among tracks t_3 and t_4, then

$$Q(\tau) = \begin{cases} p & \tau = \text{id} \\ 1 - p & \tau = (t_3, t_4) \\ 0 & \text{otherwise} \end{cases}$$

- The Markov mixing matrix induced from Q would be

	$\{t_1, t_2, t_3\}$	$\{t_1, t_2, t_4\}$	$\{t_1, t_2, t_5\}$	$\cdots$	$\{t_3, t_5, t_6\}$	$\{t_4, t_5, t_6\}$
$\{t_1, t_2, t_3\}$	p	$1 - p$	0	$\cdots$	0	0
$\{t_1, t_2, t_4\}$	$1 - p$	p	0	$\cdots$	0	0
$\{t_1, t_2, t_5\}$	0	0	1	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\{t_3, t_5, t_6\}$	0	0	0	$\cdots$	p	$1 - p$
$\{t_4, t_5, t_6\}$	0	0	0	$\cdots$	$1 - p$	p

Mixing Model

- Suppose Q is a distribution on permutation group S_{m+n}, then the simplest mixing model is

$$Q(\tau) = \begin{cases} p & \tau = \mathrm{id} \\ 1-p & \tau = (t_i, t_j) \\ 0 & \text{otherwise} \end{cases}$$

- Q induces a Markov update of beliefs for $f \in L(X)$

$$f(y) \leftarrow \sum_x P_x(y) f(x)$$

where $P_x(y) = \sum_{\tau:\tau x = y} Q(\tau)$.

Observation Model

- The simplest model for observation consist of receiving information z that with some high probability, target on track t_i is red.
- Likelihood function have the form ($a \gg b$):

$$L(z|x) = \begin{cases} a & \text{if } t_i \in x \\ b & \text{if } t_i \notin x \end{cases}$$

- Posterior by Bayes rule

$$f(x|z) = \frac{L(z|x) \cdot f(x)}{\sum_x L(z|x) \cdot f(x)}$$

Decomposition of Homogeneous Space

- Function space of homogeneous space $M^{m,n}$ decomposes as

$$S^{m+n} \oplus S^{m+n-1,1} \oplus S^{m+n-2,2} \oplus \ldots \oplus S^{m,n}$$

- $S^{m+n-i,i}$ is invariant under actions by S_{m+n}.
- Hierarchical structures: Direct sum of the first j subspaces is a $\binom{m+n}{j}$ dimensional subspace, can be regarded as functions defined on all $j-$subsets (j^{th} order statistics).
-

$$\begin{aligned} M^{m,n} &= S^{m+n} \oplus S^{m+n-1,1} \oplus S^{m+n-2,2} \oplus \cdots \oplus S^{m,n} \\ &= M^{m+n-j,j} \oplus S^{m+n-j-1,j+1} \oplus \cdots \oplus S^{m,n} \end{aligned}$$

  - $M^{m,n}$: all $n-$subsets of $\{1, 2, \cdots, m+n\}$.
  - $M^{m+n-j,j}$: all $j-$subsets of $\{1, 2, \cdots, m+n\}$.

Radon Up Transformations

- For $1 \leq k \leq n$ define the *Radon up transform*

$$R^+ : M^{m+n-k,k} \to M^{m,n} \quad \text{by} \quad R^+ f(s) = \sum_{s \supset r} f(r)$$

where $r \in M^{m+n-k,k}$ is a $k-$subset and $s \in M^{m,n}$ is an $n-$subset.

Example

- Suppose $f^2 \in M^{4,2}$ is

$\{t_1, t_2\}$	$\{t_1, t_3\}$	$\{t_2, t_3\}$	$\{t_1, t_4\}$	$\cdots$	$\{t_4, t_6\}$	$\{t_5, t_6\}$
4	4	4	2	$\cdots$	0	0

- After Radon transformation, $f^3 = R^+_{2,3} f^2$ would be

$\{t_1, t_2, t_3\}$	$\{t_1, t_2, t_4\}$	$\{t_1, t_4, t_5\}$	$\cdots$	$\{t_4, t_5, t_6\}$
4+4+4	4+2+2	2+2+0	$\cdots$	0+0+0

Radon Down Transformations

- If $M^{m,n}$ and $M^{m+n-k,k}$ are given bases consisting of delta functions on $n-$subsets and $k-$subsets. For $1 \leq k \leq n$ define *Radon down transform* $R^- : M^{m,n} \to M^{m+n-k,k}$, the (r,s) element of R^- is

$$\frac{(-1)^{n-k}(n-k)}{(-1)^{|s-r|}|s-r|\binom{m+n-k}{|s-r|}}$$

  where $r \in M^{m+n-k,k}$ is a $k-$subset and $s \in M^{m,n}$ is an $n-$subset.
- Radon transform R^+ and R^- satisfy
  - $R^- R^+ = I$
  - $R^+ R^-$ is an orthogonal projection.

Bandlimited Mixing Model

- Bandlimiting: Maintain k^{th} order statistics $f^k \in M^{m+n-k,k}$, which can be interpreted as the likelihood of a particular $k-$subset being all red.
- Induce mixing model Q to $M^{m+n-k,k}$ and update f^k by

$$f^k(y) \leftarrow \sum_x P_x(y) f^k(x)$$

$$\begin{array}{ccccccccc}
f_0^n & \xrightarrow{\mathrm{M}_0^\mathrm{n}} & f_1^n & \xrightarrow{\mathrm{M}_1^\mathrm{n}} & f_2^n & \xrightarrow{\mathrm{M}_2^\mathrm{n}} & f_3^n & \xrightarrow{\mathrm{M}_3^\mathrm{n}} & \cdots \\
R_{n,k}^- \downarrow & & R_{n,k}^- \downarrow & & R_{n,k}^- \downarrow & & R_{n,k}^- \downarrow & & \\
f_0^k & \xrightarrow{\mathrm{M}_0^\mathrm{k}} & f_1^k & \xrightarrow{\mathrm{M}_1^\mathrm{k}} & f_2^k & \xrightarrow{\mathrm{M}_2^\mathrm{k}} & f_3^k & \xrightarrow{\mathrm{M}_3^\mathrm{k}} & \cdots
\end{array}$$

Theorem

Both R^+ and R^- commute with the Markov mixing matrices induced from probability Q on permutation group S_{m+n}.

Bandlimited Observation Model

- Observation consists of first order statistics (observing the identity on track t_i is red with high probability)
- Lift first order statistics to k^{th} order statistics by Radon up transform.
- Use Bayes update to get posterior.

Classification Criteria

- We project k^{th} order statistics to first order statistics using Radon down transform.
- Predict the tracks with highest n scores as red members.

Real Camera Data

- Real Network with 8 Cameras
- 11 People (5 red, 6 blue)
- Experiments with different number of mixing events and observation events

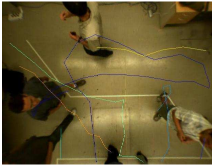


Figure: Sample Image.

Table: Experiments Data Summary

Experiment	#Mixings	#Observations	Explanations
1	8	76	few mix, lots of obs
2	169	184	
3	226	116	
4	261	64	lots of mix, few obs

Energy Distributions

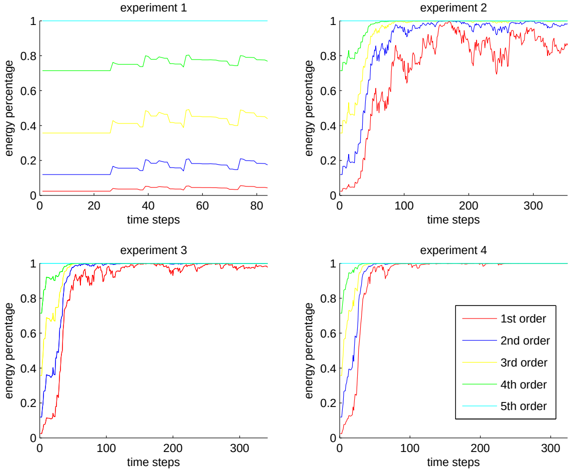


Figure: Energy distributions for four experiments

Classification Accuracy

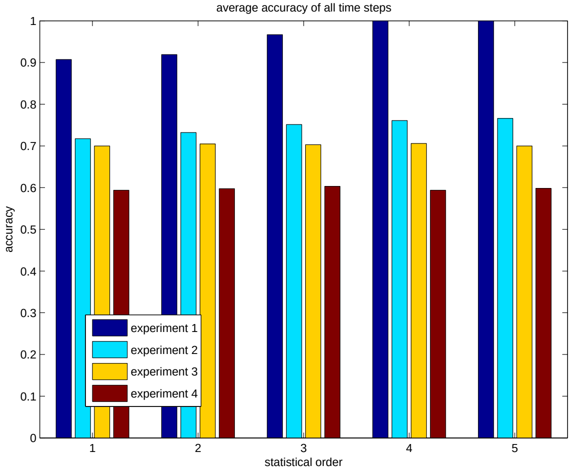


Figure: Classification accuracy of implementation with different statistical order.

Conclusions and Future Work

Conclusions

- Distributions on homogeneous spaces can be compactly summarized.
- Radon transforms useful for mapping distributions between different statistical orders.
- Evaluation of our model on a real camera network.

Future

- Use similar ideas to study other machine learning problems arising from ranking and voting.
- Smarter ways of projecting data on homogeneous spaces to low order statistics.