

Spatial Summarization of Image Collections

Diego A. Ballesteros Villamizar

ETH Zürich

February 8th, 2016

- 1 Augmented features
- 2 Sampling the distribution
- 3 Extending the location set

Leftover question

- Does using a feature matrix $\mathbf{X}' = \mathbf{X} \parallel \mathbb{I}$ improve the results?

		K			
		0	2	5	10
L	0	17.38 ± 1.81	18.75 ± 2.95	18.82 ± 2.58	18.91 ± 2.40
	2	22.66 ± 4.58	28.53 ± 4.36		
	5	25.40 ± 4.77		31.59 ± 2.38	
	10	31.13 ± 2.92			30.49 ± 3.51

- Not really, the best score so far is 34.35 ± 2.15 with $\mathbf{X} = \mathbb{I}$.
- Running time is significantly slower, because of the increased number of features $M = N + 4$.

Outline

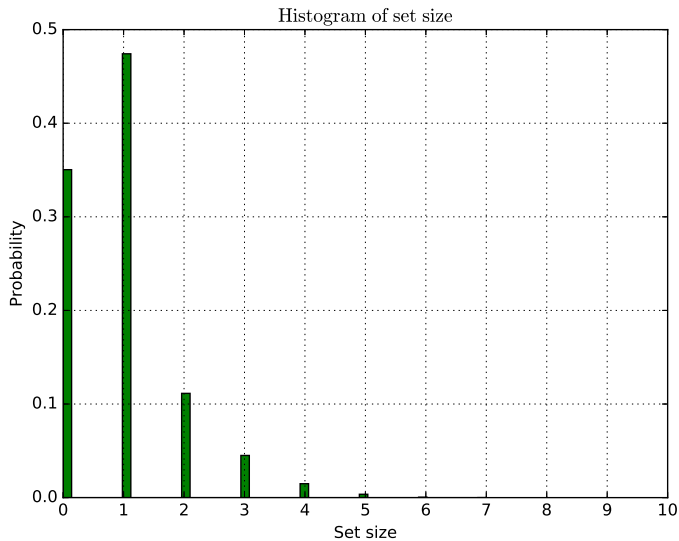
- 1 Augmented features
- 2 Sampling the distribution**
- 3 Extending the location set

Sampling from the model

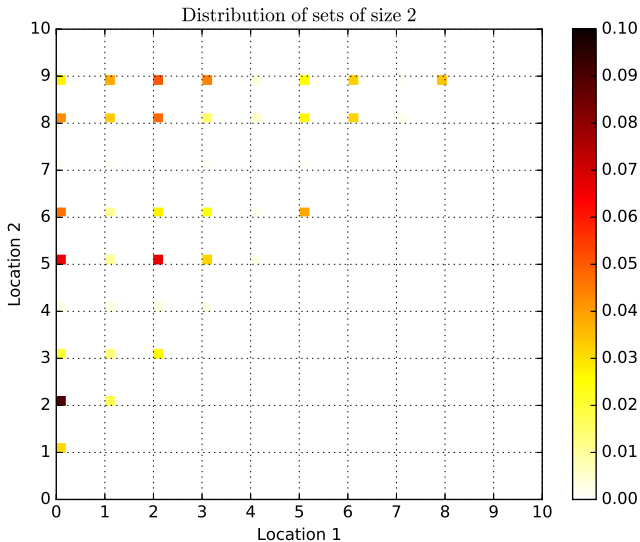
- Using the best model, i.e. without features and with $L = 5, K = 5$.
- How does the resulting distribution look?
- How to use the distribution to recommend sets?

- With $N = 10$, it is possible to calculate the probabilities from the model for all $2^{10} = 1024$ possible sets.
- Evaluating the model on all sets $S \subseteq V$ and then normalizing the probability distribution.
- Takes only seconds to evaluate.

Distribution of set size ($100k$ samples)

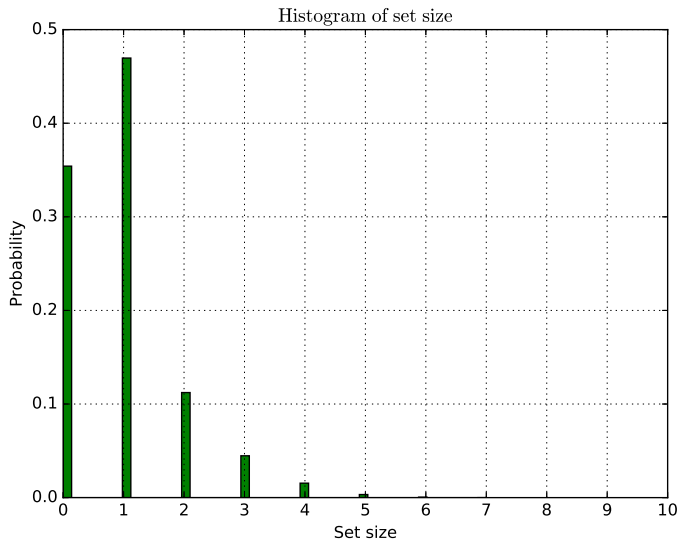


Distribution of sets with $|S| = 2$ (100k samples)

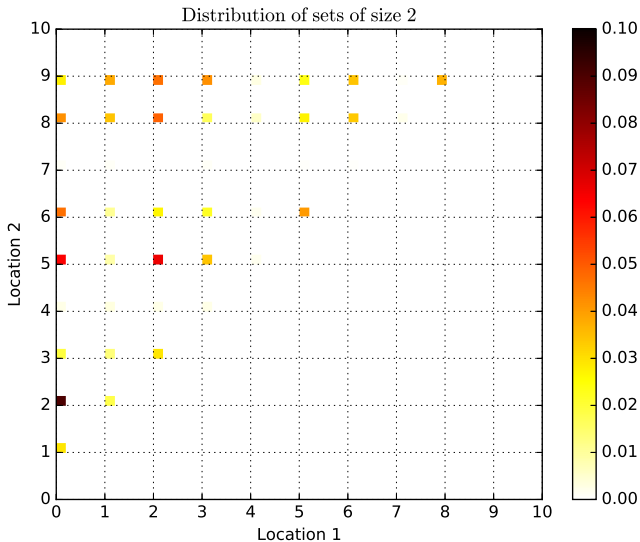


- What about a method that scales? For example if $N = 30$, then there are $2^{30} = 1073741824$ sets.
- Gibbs sampling as presented in [1].
- Run for $1M$ iterations, remove the first half of iterations are burn-in.
- Running time is a couple of minutes.

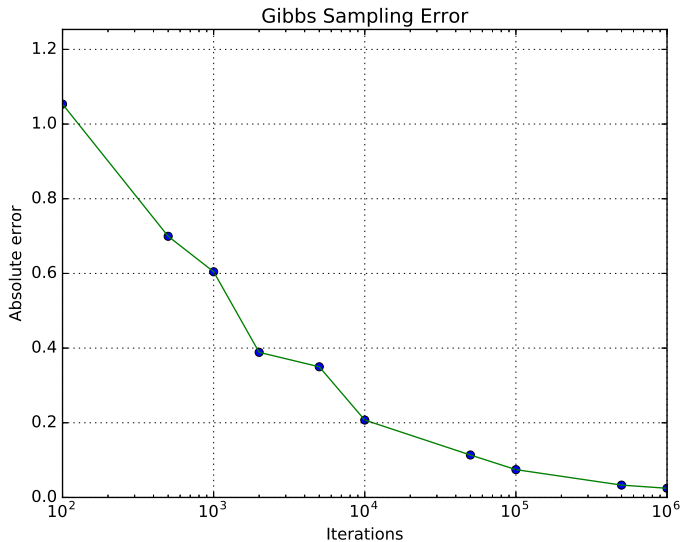
Distribution of set size ($100k$ samples)



Distribution of sets with $|S| = 2$ (100k samples)



Gibbs Sampling Performance



Outline

- 1 Augmented features
- 2 Sampling the distribution
- 3 Extending the location set

More mean-shift clusters

- Original dataset has over 160k photos.
- When clustered using mean-shift with a bandwidth of approximately 100m, there are over 2k clusters.
- Previous tests were done using the 10 top clusters according number of photos per cluster. This covered only over 30k photos.
- Does the approach scale if the number of clusters is increased?
- 50 clusters cover over 100k photos.
- 12k paths are present, in comparison there were 8k paths with 10 clusters.

Model	Accuracy	MRR
Modular	9.21 ± 1.02	27.00 ± 1.01
Markov	19.72 ± 1.23	34.50 ± 1.00
Markov with rejection	22.36 ± 1.41	38.65 ± 1.09
Proximity	12.76 ± 0.70	27.71 ± 0.88
Proximity with rejection	14.74 ± 0.64	31.34 ± 1.14

- Similar trend as with $N = 10$, Markov with rejection is the best model and it's significantly better than the modular model.

- Use the best model from the case for $N = 10$.
- The model with a diversity term, i.e. submodular, and a coherence term, i.e. supermodular. Without features, i.e. $\mathbf{X} = \mathbb{I}$.
- Running on $k = 10$ folds, with a noise factor of 50.
- Latent dimensions are: $0 \leq L \leq 20, 0 \leq M \leq 20$.

		K			
		0	2	10	20
L	0	9.21 ± 1.02	16.24 ± 1.43	16.12 ± 1.29	16.44 ± 1.77
	2	12.09 ± 1.36	17.05 ± 0.91		
	10	12.94 ± 1.39		19.28 ± 0.73	
	20	9.17 ± 2.27			21.53 ± 1.24

- Model with diversity and coherence term has performance close to the markov model with rejection.

Distribution of set size ($100k$ samples)

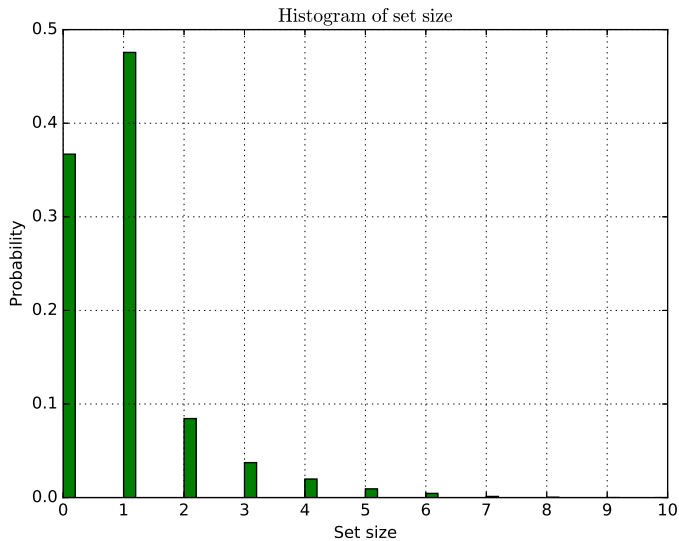


Table: Frequency of Item Sets

Set	Frequency
[0, 7]	61
[27, 28]	39
[13, 7]	36
[11, 14]	27
[0, 13]	25
[2, 7]	23
[10, 2]	22
[25, 34]	22
[2, 15]	21
[10, 15]	21

Table: Locations

Index	Location
0	Grossmünster
2	Bürkliterrasse
7	Fraumünster
10	Quaibrücke
11	Hauptbahnhof
13	Rathaus
14	Urania-Sternwarte
15	Bellevue
25	Frau Gerolds Garten
27	Zoo
28	Restaurant Masoala (Zoo)
34	Stadion Letzigrund



Gotovos, A., Hassani, S. H., and Krause, A.
Sampling from probabilistic submodular models.
In *Neural Information Processing Systems (NIPS)* (December 2015).