A Game Theoretic Approach to Learning Shape Categories and Contextual Similarities

Aykut Erdem and Andrea Torsello

Contextual Similarities

- **Psychology research of the '60s and '70s showed that perception of similarities is strongly influenced by the underlying category structure**
  [Tversky, Gati, “Features of Similarity.” Psychological Review, 1977]

  - Variations that are common to a class are perceived as less important for objects of the class than for other objects
  - Multiple classes are possible for each element and the ensemble of observed objects determine the class used for comparison.

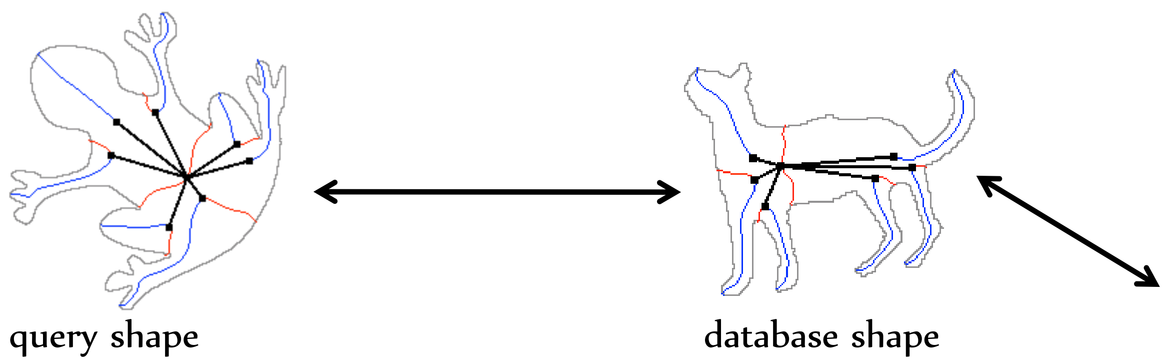
- **Knowledge about the underlying class structure is required**

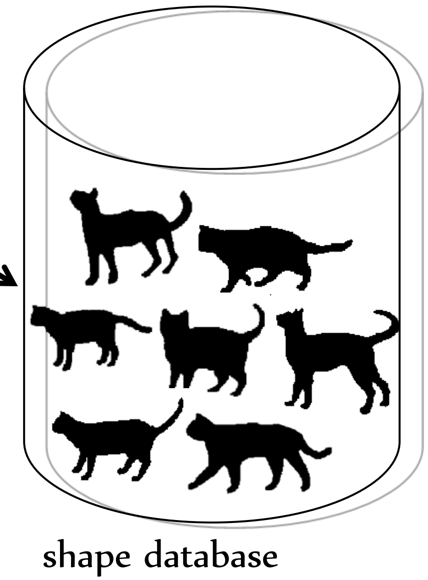
- **Categories are learned through experience**
  - experience comes before task => categories are known

- There is a chicken and egg problem
  - Similarities must be known in order to estimate class structure
  - Class structure is needed to estimate similarities
- Data clustering and contextual similarities should be learned simultaneously
- When comparison is based on matching subparts, similarities matching make sense only on the correct or similar contexts

Category-Influenced Matching

- Disconnected skeletons





- A contextual distance measure that depends on the geometric characteristics of the existing shape categories (edit costs)

- Costs of edit operations depend on variability of skeletal branch within category

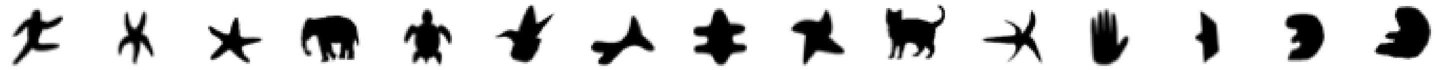
- The roles of the shapes in comparison are <u>asymmetric</u>
- Distance is computed within the context of the database shape (Asymmetric)

Proposed Approach

- Class membership can be considered a latent variable => EM on a similarity space

- Alternate between
  - class estimation (pairwise clustering)
    - Need measure of membership (posterior of latent variable)
    - Use game-theoretic clustering approach
  - Similarity update (class model and contextual similarity)
    - Learn part variability and related edit costs
    - Asymmetric because of difference in role between query and target shape (context of the database shape)

Extracted Shape Categories



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Precision	1.00	0.50	0.64	1.00	0.95	1.00	0.51	1.00	1.00	0.87	0.56	1.00	0.89	0.71	0.50
Recall	1.00	0.50	0.35	0.90	0.95	0.95	0.95	0.90	0.90	0.65	0.25	0.95	0.80	0.60	0.35
Payoff	0.96	0.94	0.91	0.91	0.94	0.95	0.95	0.96	0.97	0.91	0.89	0.94	0.95	0.91	0.90
Entropy	2.86	2.86	2.25	2.70	2.50	2.81	3.12	2.77	2.89	2.41	2.05	2.66	2.82	2.51	2.18

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Precision	0.95	0.84	0.40	1.00	0.79	0.81	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.76	0.95
Recall	0.95	0.80	0.20	1.00	0.75	0.85	0.45	0.30	0.90	1.00	0.45	0.25	0.20	0.95	0.95
Payoff	0.94	0.94	0.87	0.97	0.92	0.90	0.93	0.83	0.94	0.94	0.87	0.87	0.80	0.95	0.96
Entropy	2.65	2.82	1.82	2.98	2.80	2.51	2.08	1.77	2.82	2.69	2.13	1.61	1.38	2.96	2.75

	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
Precision	0.92	1.00	0.79	0.86	0.60	0.63	0.50	1.00	0.91	0.91	0.95	1.00	1.00	0.94	1.00
Recall	0.60	0.25	0.95	0.30	0.45	0.60	0.20	1.00	0.50	0.50	0.95	0.90	1.00	0.85	0.95
Payoff	0.93	0.86	0.94	0.89	0.89	0.93	0.88	0.96	0.94	0.93	0.96	0.93	0.95	0.96	0.97
Entropy	2.36	1.60	2.81	1.89	2.45	2.72	2.06	2.92	2.37	2.37	2.88	2.78	2.81	2.75	2.92

	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Precision	1.00	1.00	0.95	1.00	1.00	1.00	0.63	0.33	0.88	1.00	1.00	0.80	1.00	1.00	0.52
Recall	0.90	0.30	1.00	0.75	0.20	0.90	0.60	0.20	0.70	0.20	0.85	0.40	0.60	0.40	0.75
Payoff	0.95	0.85	0.97	0.95	0.84	0.95	0.93	0.87	0.92	0.86	0.96	0.86	0.92	0.89	0.95
Entropy	2.70	1.77	2.93	2.70	1.38	2.81	2.71	2.13	2.58	1.39	2.83	1.87	2.33	1.94	3.20

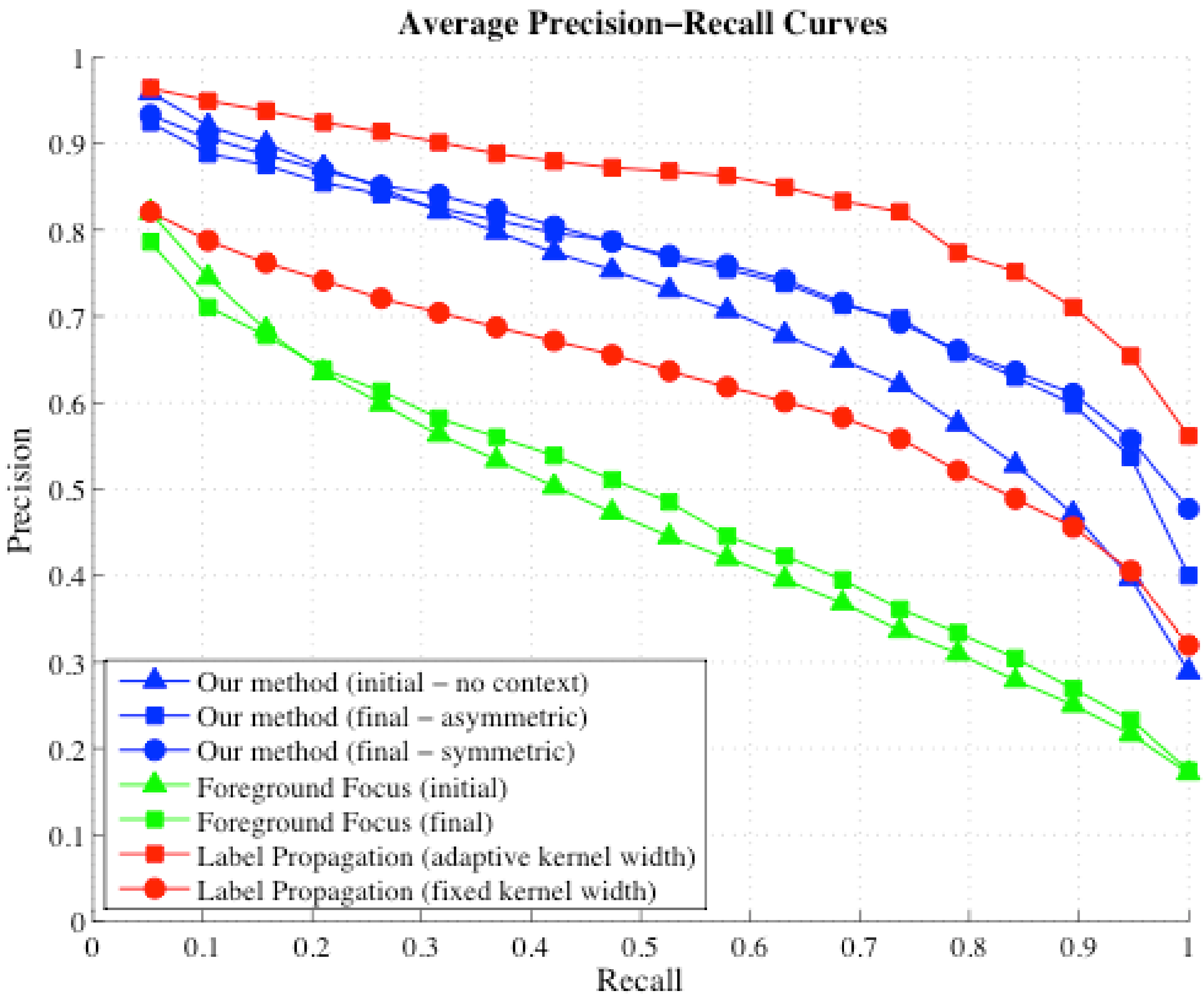
Final shape categories . 65 shapes remain unclassified

Evaluation of Clustering Results

The Method	Rand Index	Corrected Rand Index	Normalized Mutual Information
Our method – at t=0	0.9818	0.9929	0.8517
Our method – asymmetric case	0.9863	0.9916	0.8619
Our method – symmetric case	0.9870	0.9900	0.8680
Normalized Cut [17] – # of classes=50	0.9833	0.9836	0.8493
Normalized Cut [17] – # of classes=51	0.9832	0.9833	0.8381
Normalized Cut [17] – # of classes=61	0.9848	0.9854	0.8380
Foreground Focus [10] – # of classes=50	0.9748		0.7329

- Proposed approach gives a better clustering result when compared to Normalized Cut and Foreground Focus

Retrieval Performances



- The contextual similarities based on the modified category-influenced matching scores are much better than the original ones.

- Foreground Focus performs badly.

- Label Propagation gives the best retrieval performance
  - Its performance degrades when we use a fixed kernel width

Questions

- Is prior knowledge of category necessary to bridge the semantic gap?
- Does similarity come after category evaluation?
- Is shape similarity the main representation or should it be derived from some feature model that characterizes the categories?
- Is the asymmetry introduced by the context important?