A Domain-based Approach to Social Relation Recognition

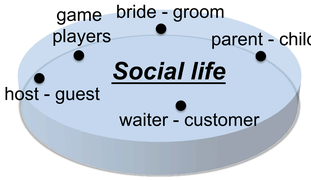
Qianru Sun, Bernt Schiele and Mario Fritz
Max Planck Institute for Informatics, Saarland Informatics Campus, Germany

IEEE 2017 Conference on Computer Vision and Pattern Recognition
CVPR July 21-26 HONOLULU 2017

Motivations

- Understand social relations in daily life photos
- Provide coverage of all social life

In related works:
Ramanathan et al. CVPR13 "*Social role discovery in human events*."
Shu et al. CVPR15 "*Joint Inference of Groups, Events and Human Roles in Aerial Videos*"



In our work:

Attachment: Proximity maintenance during late infancy with specific others in the service of safety

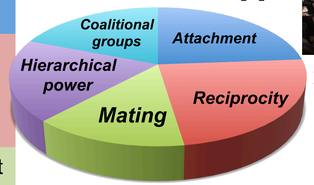
Reciprocity: Maximize joint outcomes for functional equal, require a long-term accounting of the relative benefits and costs

Mating: Select & maintain & protect access to the sexual mate

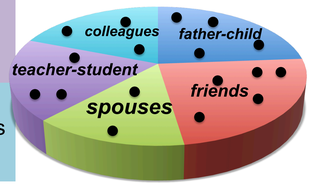
Hierarchical power: Between individuals with unequal control, resources or resource-holding

Coalitional groups: Identification and defense of the lines that divides "us" and "them" in group coalitions

Social domain level [1]



Social relation level



Contributions

- The first domain-based approach that partitions human social life
  5 social relations [1] → *16 social relations*
- The first daily life photo dataset with the hierarchical labeling of social domains and relations
  26,915 person pairs on PIPA dataset [2]
- Semantic attribute models of person appearance and behavior for the interpretable recognition of social domains and relations
  5 head attribute models + 5 body attribute models

5 social domains → 16 social relations

Attachment (d0)

father – child (r0) | mother – child (r1) | grandpa - grandchild (r2) | grandma - grandchild (r3)

Reciprocity (d1)

friends (r4) | siblings (r5) | classmates (r6)

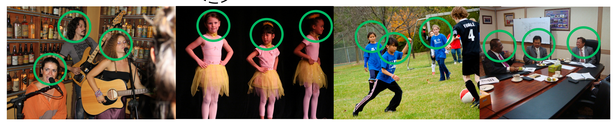
Mating (d2)

lovers / spouses (r7)

Hierarchical power (d3)

presenter – audience (r8) | teacher – student (r9) | trainer – trainee (r10) | leader – subordinate (r11)

Coalitional groups (d4)



band members (r12) | dance team members (r13) | sport team members (r14) | colleagues (r15)

Semantic attribute models

Revisit domain definitions:

e.g. Attachment domain characterized by proximity maintenance within a protective relationship, e.g. kinship between parents and children. Human attributes such as age difference, proximity and the activity of seeking protection are social cues which can be visually recognizable.

Attributes: age, proximity, activity

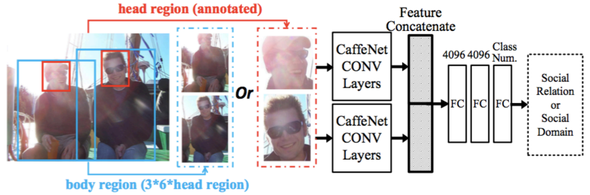
	SEMANTIC ATTRIBUTES	*(DATASET)*
Age	6 classes: senior, young, …	(PIPA, head regions)
Gender	2 classes: female, male	(PIPA, head regions)
Appearance	32 classes: arched eyebrows, …(CelebA Face)	
Pose	5 classes: frontal, left, right, …	(IMFDB)
Emotion	7 classes: anger, happiness, …	(IMFDB)
Age	6 classes: senior, young, …	(PIPA, body regions)
Gender	2 classes: female, male	(PIPA, body regions)
Clothing	8 classes: has t-shirt, has shorts, …(Berkeley P.A.)	
Proximity (pairwise)	7 classes: hold from behind, …	(Immediacy)
Activity	504 classes: adjusting, smiling, …	(ImSitu.)

Build semantic attribute models:
- Feature extraction from 10 attribute models
- Feature concatenation (add location & scale)
- Classifier training, e.g. linear SVM

Advantages:
- Semantic attributes lend themselves for the model interpretability.
- Bridge the gap between social psychology theory and computational models.
- Allow to leverage other large-scale datasets annotated with such attributes.

Baseline vs. Semantic attribute models

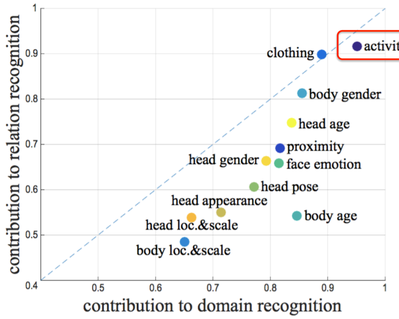
Baseline end-to-end recognition model



Accuracies of social relation recognition, social domain recognition, social domain generalization testing

MODEL	RELATION	DOMAIN	GENERALIZATION
END-TO-END SCRATCH	34.4%	41.9%	–
END-TO-END FINETUNED	46.2%	59.0%	18.5%
PRE-TRAINED, SVM	35.9%	53.3%	27.7%
FINETUNED, SVM	48.6%	63.2%	27.1%
HEAD ATTRIBUTES, SVM	44.8%	59.4%	21.5%
BODY ATTRIBUTES, SVM	57.2%	67.7%	32.8%
ALL ATTRIBUTES, SVM	57.2%	67.8%	33.3%

Go into semantic attributes



top activities

father – child (r0)
- grinning
- giggling
- carrying
- admiring
- hugging

band members (r12)
- singing
- whirling
- dancing
- performing
- twirling

[1] Bugental, D.B.: Acquisition of the Algorithms of Social Life: A Domain-Based Approach. Psychological Bulletin,2000.
[2] Zhang, N., Paluri, M., Taigman, Y., Fergus, R. and Bourdev, L.: Beyond frontal faces: Improving person recognition using multiple cues. CVPR 2015.

code & data