Cumulative Attribute Space for Age and Crowd Density Estimation

Ke Chen[1], Shaogang Gong[1], Tao Xiang[1], and Chen Change Loy[2]
1.Queen Mary, University of London; 2. The Chinese University of Hong Kong
{cory, sgg, txiang}@eecs.qmul.ac.uk; ccloy@ie.cuhk.edu.hk

INTRODUCTION

Problems:

Age estimation / Crowd counting / Head pose estimation

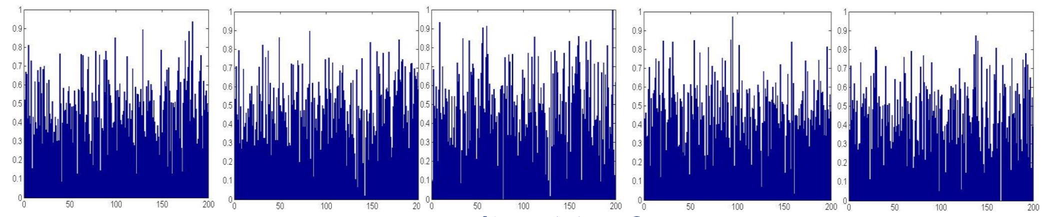
How old? / How many persons? / What is the viewing angle?

Why Challenging:

1. Feature Variation

Age estimation

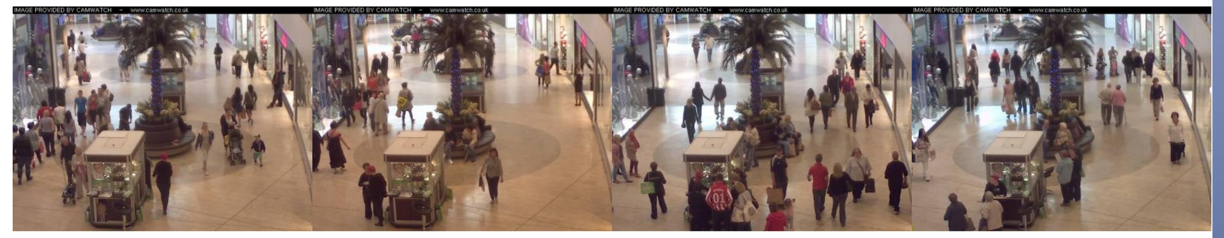
different people with the same age



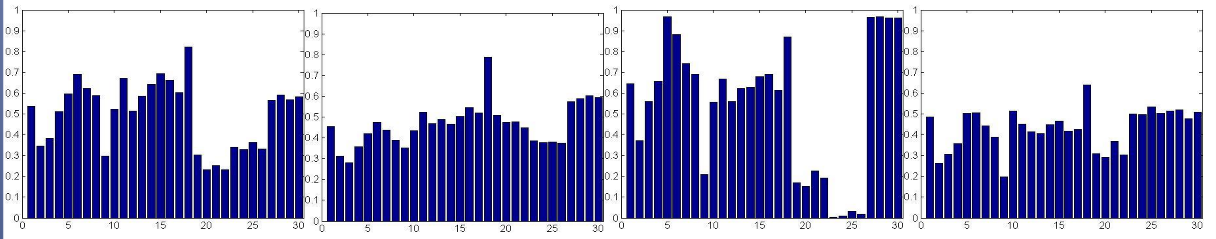
corresponding AAM features

- ✓ Extrinsic condition: Lighting conditions; Viewing angles
- ✓ Intrinsic condition: aging processing of different people glasses, hairstyle, gender, ethnicity

Crowd density estimation



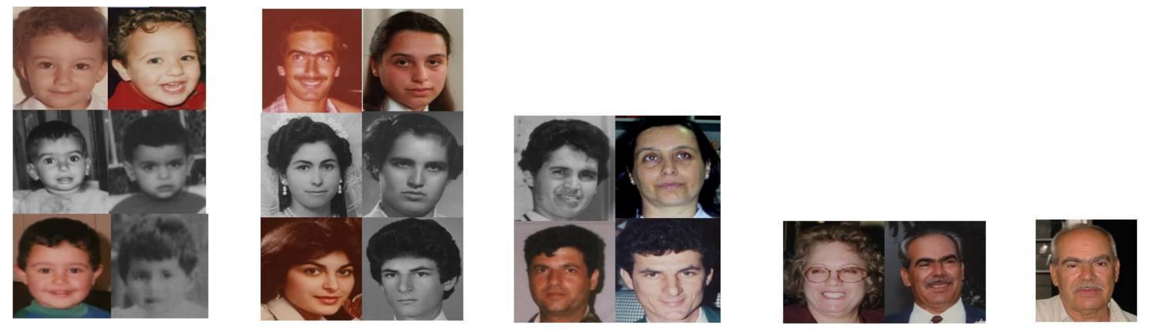
The same person count

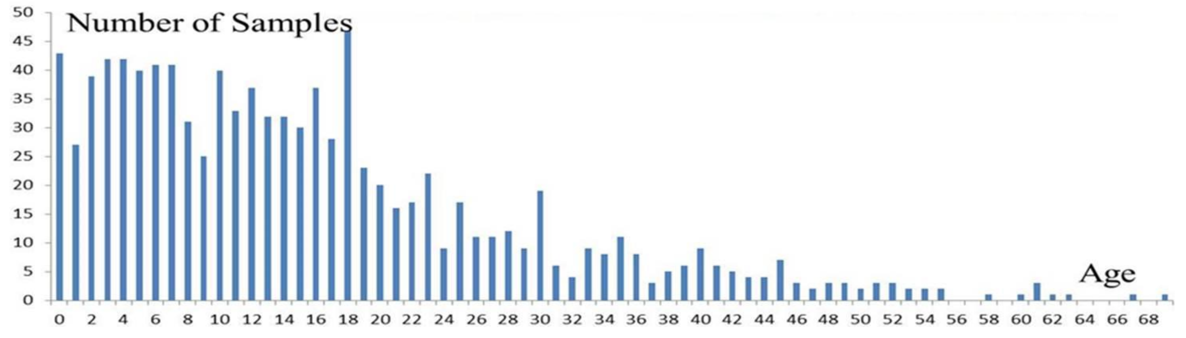


corresponding segment+edge+texture features

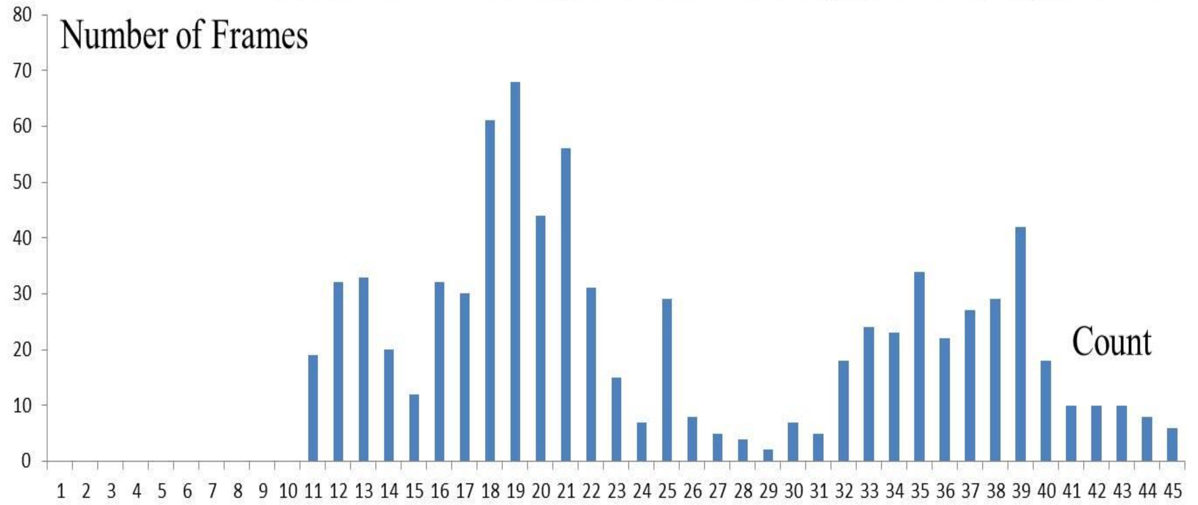
- ✓ Extrinsic condition: Lighting conditions; Viewing angles
- ✓ Intrinsic condition: occlusion; density distribution in the scene

2. Sparse and imbalanced data





Data Distribution of FG-NET Dataset

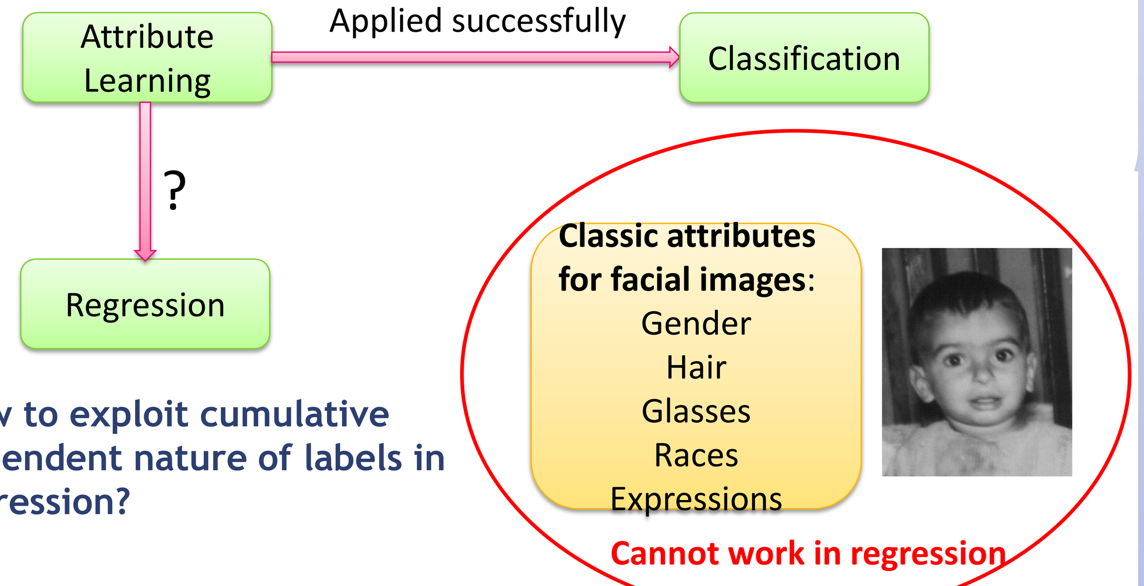


Data Distribution of UCSD Crowd Dataset

CONTRIBUTIONS

- ✓ Propose an attribute representation for regression
- ✓ Our cumulative attributes:
  - a) has clear semantic meaning;
  - b) discriminative;
  - c) do not need additional attribute annotation
- ✓ Address sparse & imbalanced data and feature variation problems
- ✓ Extensive experiments on two applications: facial age estimation and crowd counting

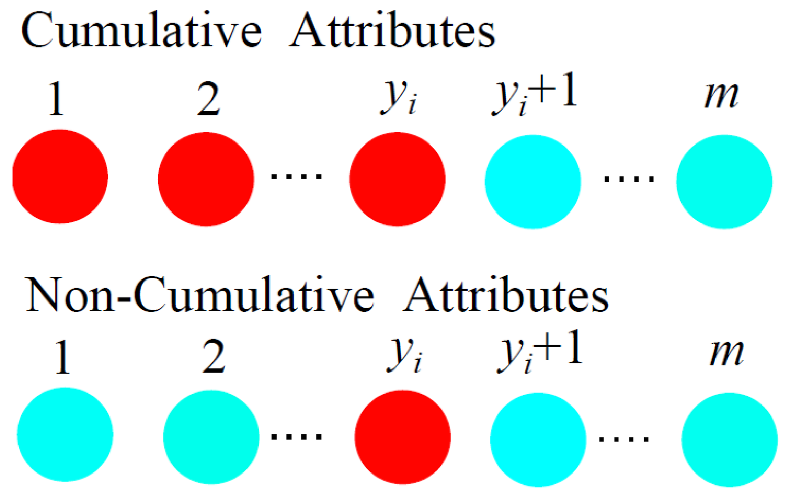
METHODOLOGY



How to exploit cumulative dependent nature of labels in regression?

Cumulative Attributes:

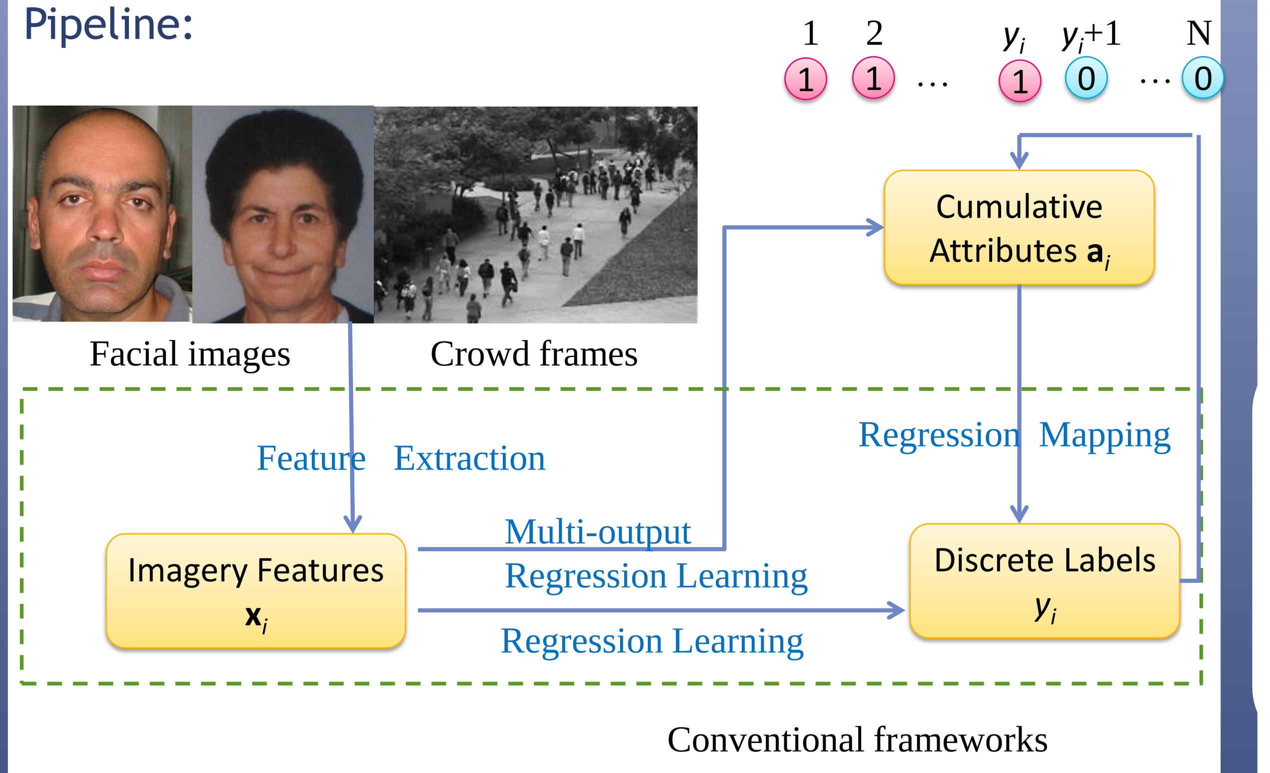
Definition:



$$a_i^j = \begin{cases} 1, & \text{when } j \leqslant y_i, \\ 0, & \text{when } j > y_i, \end{cases}$$

$$a_i^j = \begin{cases} 1, & \text{when } j = y_i, \\ 0, & \text{when } j \neq y_i. \end{cases}$$

Pipeline:



Model Learning:

✓ Joint attribute learning

$$\min \ \frac{1}{2}\|\mathbf{w}^j\|_2^2 + C\sum_{i=1}^{N} \text{loss}(a_i^j, f^j(\mathbf{x}_i)),$$

where loss function is quadratic funtion.

$$\min \ \frac{1}{2}\|\mathbf{W}\|_F^2 + C\sum_{i=1}^{N} \|\mathbf{a}_i^T - (\mathbf{x}_i^T\mathbf{W} + \mathbf{b})\|_F^2,$$

In light of this, the close-form solution of the above formulation can be obtained.

✓ Regression Mapping
- with attribute representation as input
- is not limited to a specific regression model

EXPERIMENTS

Comparative Evaluation:

Method	FG-NET [13]		MORPH [7]	
	mae	cs	mae	cs
AGES [13]	6.77	–	8.83	–
RUN [33]	5.78	–	–	–
Ranking [32]	5.33	–	–	–
RED-SVM [6]	5.24	–	6.49	–
LARR [15]	5.07	–	–	–
MTWGP [35]	4.83	–	6.28	–
OHRank [7]	4.85	74.4%	**5.69**	56.3%
SVR [15]	5.66	68.0%	5.77	57.1%
CA-SVR (Ours)	**4.67**	**74.5%**	5.88	**57.9%**

Age Estimation

Method	UCSD [4]			Mall [8]		
	mae	mse	mde	mae	mse	mde
LSSVR [31]	2.20	7.29	0.107	3.51	18.2	0.108
KRR [1]	2.16	7.45	0.107	3.51	18.1	0.108
RFR [23]	2.42	8.47	0.116	3.91	21.5	0.121
GPR [4]	2.24	7.97	0.112	3.72	20.1	0.115
RR [8]	2.25	7.82	0.110	3.59	19.0	0.110
CA-RR (Ours)	**2.07**	**6.86**	**0.102**	**3.43**	**17.7**	**0.105**

Crowd Counting

Cumulative (CA) V.S. Non-Cumulative (NCA):

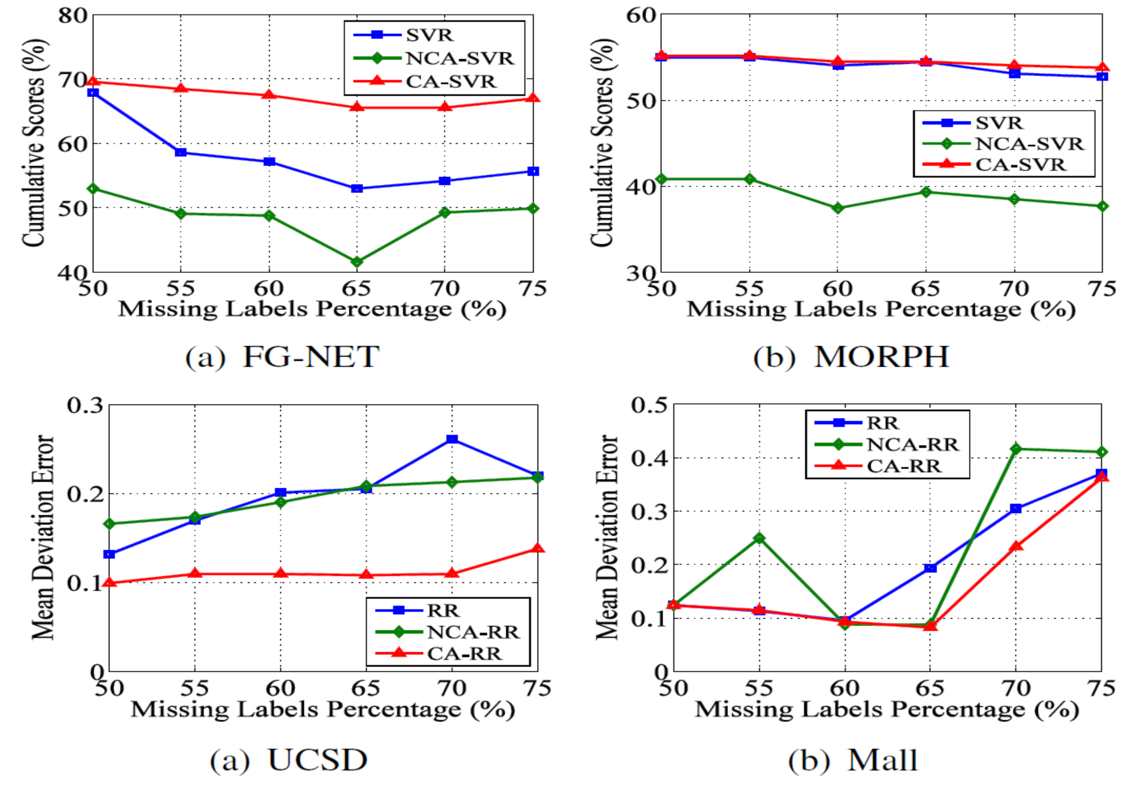
Methods	FG-NET [13]		MORPH [7]	
	mae	cs	mae	cs
NCA-SVR	8.95	41.8%	7.28	44.2%
CA-SVR	**4.67**	**74.5%**	**5.88**	**57.9%**

Age Estimation

Methods	UCSD [4]			Mall [8]		
	mae	mse	mde	mae	mse	mde
NCA-RR	2.85	11.9	0.137	4.31	25.8	0.131
CA-RR	**2.07**	**6.86**	**0.102**	**3.43**	**17.7**	**0.105**

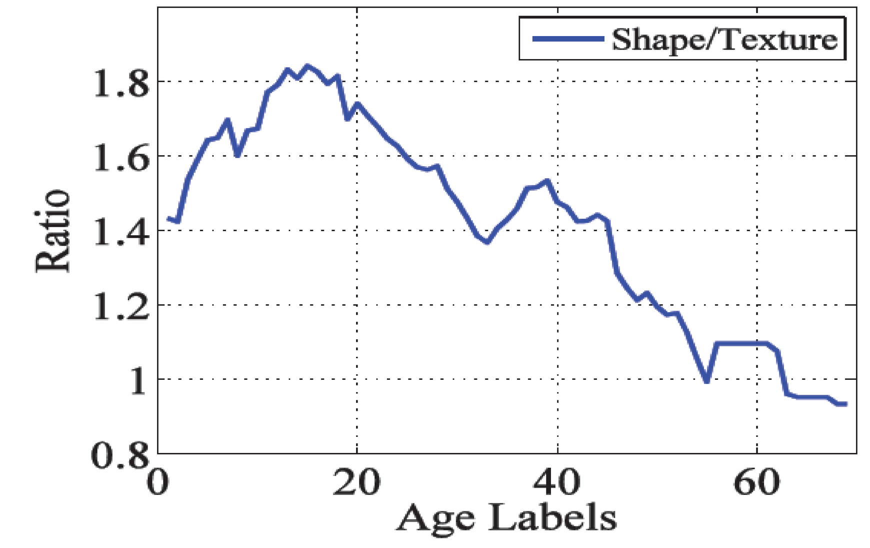
Crowd Counting

Robustness Against Sparse and Imbalanced Data :



(a) FG-NET (b) MORPH
(a) UCSD (b) Mall

What Is Learned by Cumulative Attributes:



Shape plays a more important role than texture when one is younger.