From people to groups

In Enviromental and Egocentric Cameras

Rita Cucchiara @Imagelab

Imagelab, http://www.Imagelab.ing.unimore.it

DIPARTIMENTO DI INGEGNERIA «Enzo Ferrari» Università di Modena e Reggio Emilia, Italia

Director **SOFTECH-ICT**
Research Center in ICT for Enterprise , Modena Technopole

UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

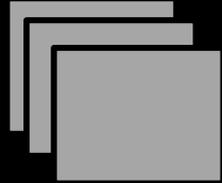Open position for Phd-students and post-doc in Italy

What we are doing..

Computer Vision for Surveillance
Tracking, tracking, tracking
Egocentric Vision
Sensing in IoT
Human behavior/gesture undertsanding
3D interaction for cultural heritage
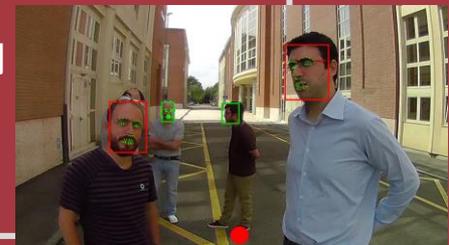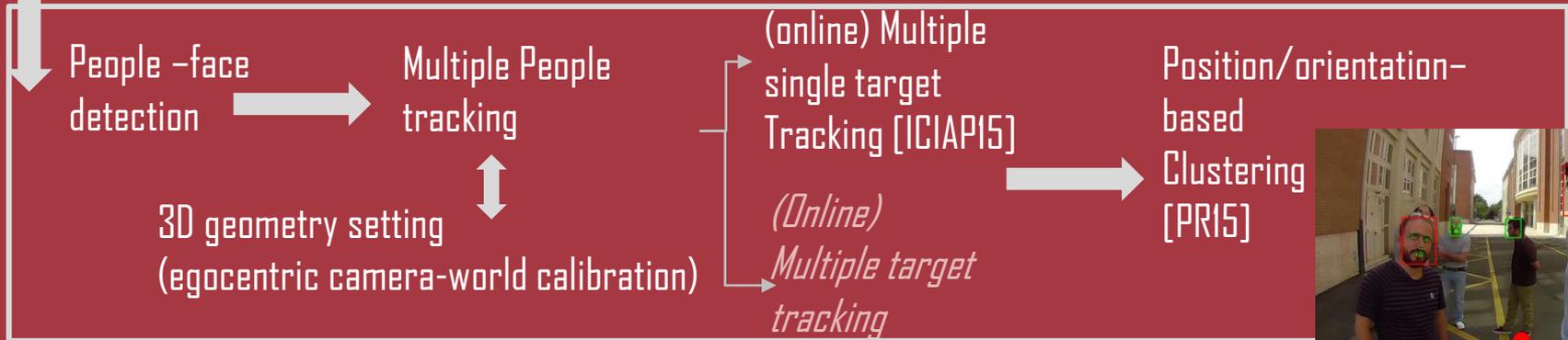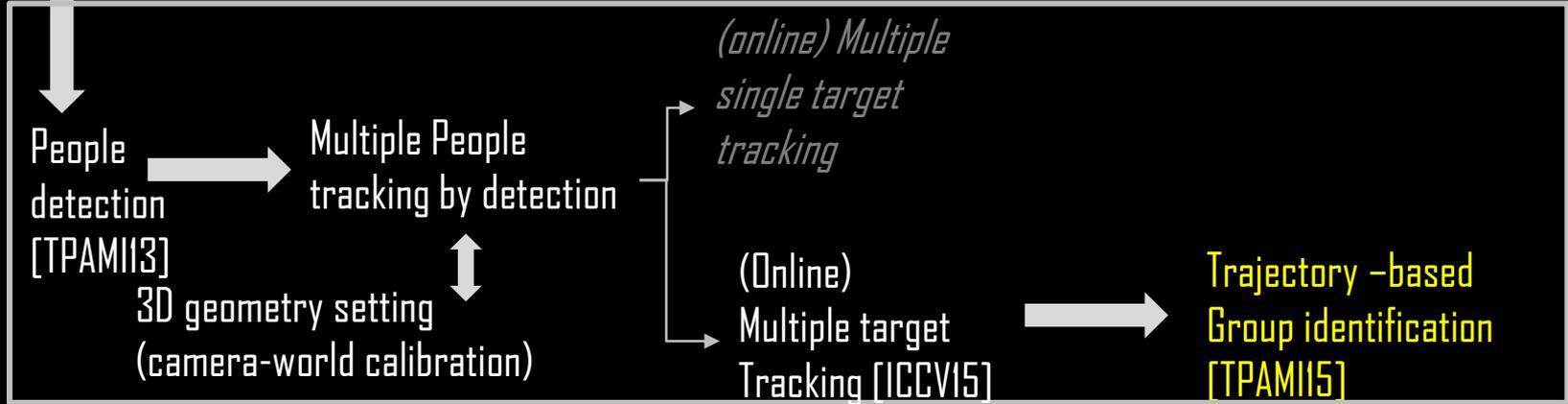Multimedia: deep learning for video annotation

People group in crowd analysis @**Imagelab**

The pipeline

Group detection as a whole
In image and video
 a classification problem

People
detection
[TPAMI13] → Multiple People
tracking by detection

3D geometry setting
(camera-world calibration)

(online) Multiple single target tracking

(Online)
Multiple target
Tracking [ICCV15] → **Trajectory –based Group identification [TPAMI15]**

People –face
detection → Multiple People
tracking

3D geometry setting
(egocentric camera-world calibration)

(online) Multiple
single target
Tracking [ICIAP15]

(Online) Multiple target tracking

Position/orientation–
based
Clustering
[PR15]

WE NEED ALL PIPELINE BLOCKS...

In Enviromental Vision

- People detection in images

- MTT (online) multiple target tracking :

    - Learning to Divide and Conquer for Online Multi-Target Tracking [Solera et al ICCV2015]
    - Evaluating tracking; the importance of FRAG and IDS [Solera et al AVVS2015 BPA]
    - Matlab toolbox (imagelab.unimore.it/tbd_evaluation)
- Group detection with LSSVM [Calderara et al TPAMI2015]

- Leader analysis [Solera et al WS@CVPR2015]
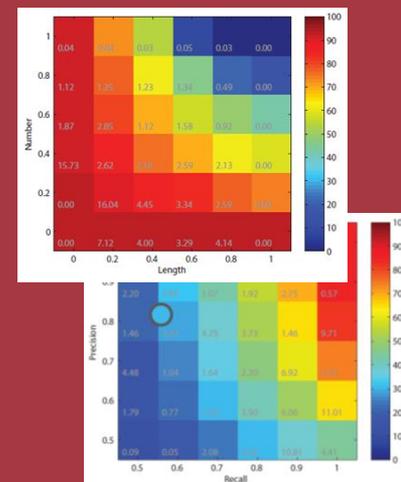
In Egocentric Vision

- Face and orientation detection by egocentric view

- Multiple Single Tracking (online) with odometry constraints

    - Group detection with LSSVM [Alletto et al PR2015]

Group detection

Understanding physical and social features

(SR. Musse, D Thalmann,
A model of human crowd behavior, CAS, 1997]

Defining physical and social **features**
integrated in metric learning

<u>learn to partition pedestrian</u> into groups
observing pairwise and transitive relations.

PHYSICAL IDENTITY	SOCIAL IDENTITY	SOCIAL IDENTITY	COMMON GOAL
Proxemics	Dynamic Time Warping	Granger Casuality	Heat map

[F.Solera, S.Calderara, R. Cucchiara "Socially Constrained Structural Learning for Groups Detection in Crowd" IEEE TPAMI, 2015)
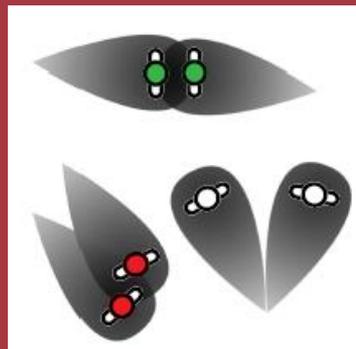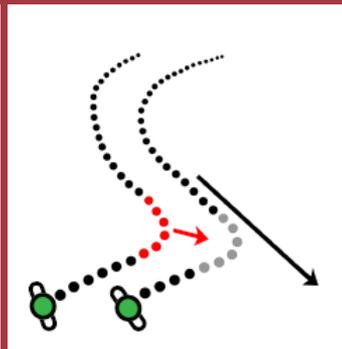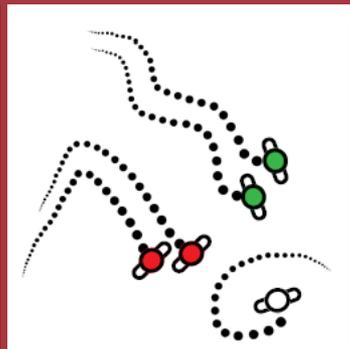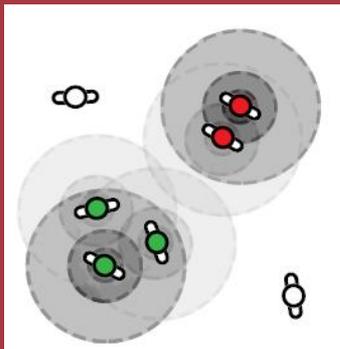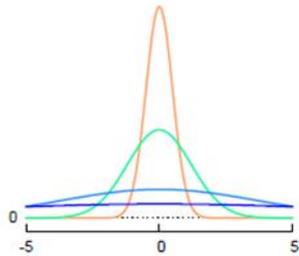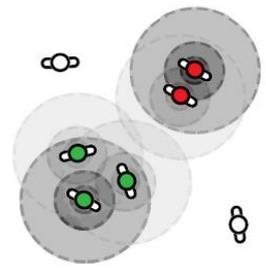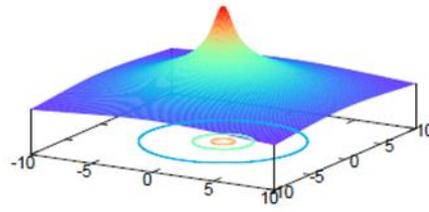
http://imagelab.unimore.it/group-detection/

Physical identity: Proxemics

(a)   (b)

$$\mathrm{GMM}(\mathbf{p}_a^t - \mathbf{p}_b^t) = \frac{1}{4}\sum_{z=1}^{4}\mathcal{N}(\mathbf{p}_a^t - \mathbf{p}_b^t | 0, \Sigma_z)$$

$$d_{\mathrm{ph}}^k(a,b) = \frac{1}{|\overline{\mathcal{T}}|}\sum_{t\in\overline{\mathcal{T}}}\mathrm{GMM}(\mathbf{p}_a^t - \mathbf{p}_b^t)$$

Social Identity: Trajectory Similarity
Dynamyc Time Warping
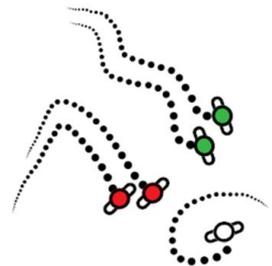
$$\gamma_{ab}(i,j) = D_{ab}^{ij} + \min\{\gamma_{ab}(i\text{-}1,j), \gamma_{ab}(i\text{-}1,j\text{-}1), \gamma_{ab}(i,j\text{-}1)\}$$

$$d_{\mathrm{sh}}(a,b) = \gamma_{ab}(A,B)/\max(A,B)$$

Social identity: Granger Causality
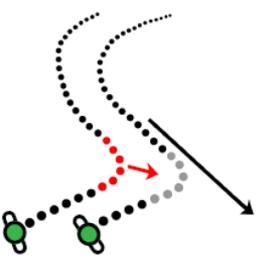
The influence of a trajectory to predicting another in a Fisher-Snedecor Statistics

$$S_{b \to a} = \frac{(RSS_c - RSS_u)/m}{RSS_u/(K - 2m - 1)}$$

$$RSS_c = \sum_{t=1}^{K} \varepsilon_t (T_a | \bar{T}_a(t-m))^2$$

$$RSS_u = \sum_{t=1}^{K} \varepsilon_t (T_a | \bar{T}_a(t-m), \bar{T}_b(t-m))^2$$

F(m, K-2m-1)
critical confidence
S
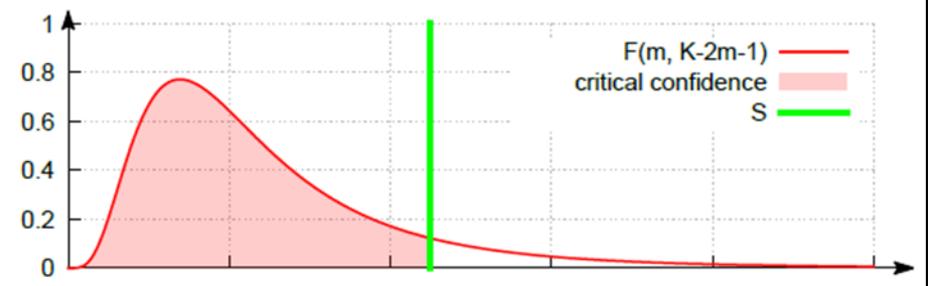
Common Goal: the heat map

0

0

LATENT STRUCTURAL SVM FOR LEARNING GROUPS

Correlation clustering ....

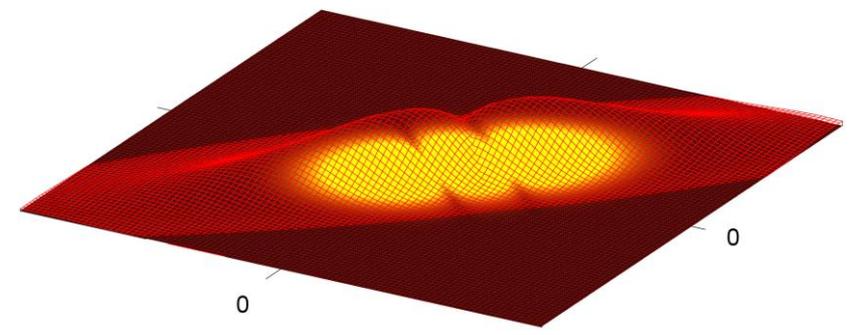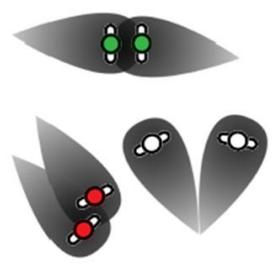$$CC = \arg \max_{\mathbf{y} \in \mathcal{Y}(M)} \sum_{y \in \mathbf{y}} \sum_{a \neq b \in y} W_{\mathbf{d}}^{ab}$$

The affinity matrix is parameterized by a combination of similarity and dissimilarity features ....

$$W_{\mathbf{d}}^{ab} = \boldsymbol{\alpha}^T (1 - \mathbf{d}(a,b)) - \boldsymbol{\beta}^T \mathbf{d}(a,b)$$

The feature vector is defined in terms of distances.................

$$\mathbf{d}(a,b) \stackrel{\text{def}}{=} \mathbf{d}^k(a,b) = [d_{\text{ph}}, d_{\text{sh}}, d_{\text{ca}}, d_{\text{he}}]_{a,b}^k$$

LS SVM with Group-Mitre loss function.......................

$$\min_{\mathbf{w}, \xi} \quad \frac{1}{2}\|\mathbf{w}\|^2 + \frac{C}{n}\sum_{i=1}^{n} \xi_i$$

$$\text{s.t.} \quad \forall i : \xi_i \geq 0,$$

$$\forall i, \forall \mathbf{y} \in \mathcal{Y}(\mathbf{x}_i) \backslash \mathbf{y}_i : \mathbf{w}^T \delta \Psi_i(\mathbf{y}) \geq \Delta(\mathbf{y}, \mathbf{y}_i) - \xi_i,$$

(18)

Hierarchical Coherence   Density Invariance   Transitivity
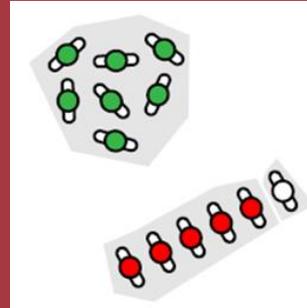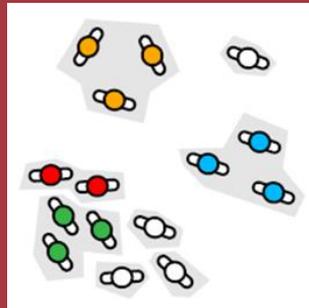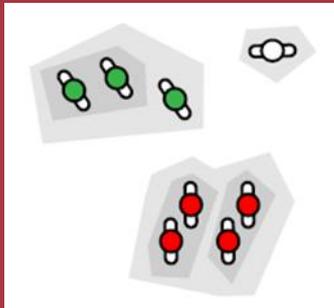
DATASET
HTTP://IMAGELAB.UNIMORE.IT/GROUP-DETECTION/

In Tab. 3 we report some measures useful to characterize the spatial complexity of the datasets:

- d_{in} is the *group compactness*, computed as the mean distance between members of the same groups;
- d_{out} is the *group isolation* or the mean distance between each member and its closest unrelated pedestrian;
- the ratio $d_{i/o} \stackrel{\text{def}}{=} d_{in}/d_{out}$ measures *crowd collectiveness*: small values mean compact groups in a sparse crowd.

TABLE 3

Datasets: number of pedestrians (#p), groups (#g) and density metrics.

|	#p	#g	d_{in} (m)	d_{out} (m)	$d_{i/o}$
CBE					
student003	406	108	0.41	0.70	0.59
BIWI					
eth	117	18	0.99	2.79	0.35
hotel	107	11	0.75	2.00	0.38
VEIIG					
gal1	630	207	0.77	1.66	0.46
MPT-20x100 (averaged)					
20 sequences	82	10	0.63	1.45	0.48

Some results
Groups

The influence of the datasets (results with GT trajectories!!!!)

TABLE 2

Comparison with other state of the art methods on publicly available dataset.

	our method (Δ_{GM})		baseline		[2]		[9]		[26]	
	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{P}$	$\mathcal{R}$
BIWI hotel	**97.3 ± 0.7**	**97.7 ± 1.5**	71.0 ± 8.1	69.6 ± 7.4	86.9	85.5	91.3	95.9	87.0	91.0
BIWI eth	**91.8 ± 1.2**	**94.2 ± 0.9**	72.4 ± 4.4	65.2 ± 3.4	87.0	84.2	83.0	80.2	79.0	80.0
CBE student003	**81.7 ± 0.2**	**82.5 ± 0.2**	59.9 ± 2.9	53.5 ± 6.8	77.2	73.6	80.5	77.0	70.0	74.0

The influence of the features
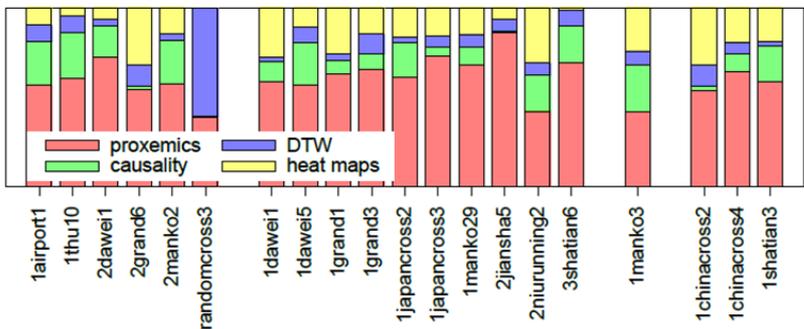
[2] W. Ge, R. Collins, and R. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, pp. 1003–1016, May 2012.

[9] K. Yamaguchi, A. Berg, L. Ortiz, and T. Berg, "Who are you with and where are you going?" in *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 1345–1352.

[26] M. Zanotto, L. Bazzani, M. Cristani, and V. Murino, "Online bayesian non-parametrics for social group detection," in *Proc. British Machine Vision Conference (BMVC)*, 2012, pp. 111.1–111.12.

AT THE END

What are the performance when
- People detection
- People tracking
- Feature extraction
- Trajectory Clustering
- Leader detecion

have errors?

... there is room

to improve

imagelab

group-detection
structured learning for social group detection in crowds

zip Download .zip   G2 Download .tar.gz   View on GitHub

Modern crowd theories agree that collective behavior is the result of the underlying interactions among small groups of individuals. In this work, we propose a novel algorithm for detecting social groups in crowds by means of a Correlation Clustering procedure on people trajectories. The affinity between crowd members is learned through an online formulation of the Structural SVM framework and a set of specifically designed features characterizing both their physical and social identity, inspired by Proxemic theory, Granger causality, DTW and Heat-maps. To adhere to sociological observations, we introduce a loss function (G-MITRE) able to deal with the complexity of evaluating group detection performances. We show our algorithm achieves state-of-the-art results both in presence of complete trajectories and with tracklets provided by available detector/tracker systems.

... ECCV 2016 (oct 8-16)
... ACM MM2016 (oct 15 -19) in Amsterdam
.... IAPR 2016 (dic 4-8) in Cancun

We look for new TUTORIALS!
Open calls,
contact me

ICCV 2017 organization is starting

Call for helping in ICCV 2017

ICCV 2017 Venice, Italy
International Conference on Computer Vision
22-29 October 2017

St. Marc, Rialto bridge,
The lagoon

Festival Cinema
Palace

Lido

The contemporary art
"Biennale di Venezia",
And Art museums

The Venue

Tourism

Casino Palace

-Venice smart city: WIFI everywhere
-Ferryboats and food discounts for students

ICCV 2017

Program Chairs:
R. Cucchiara, Y. Matsushita,
N. Sebe, and S. Soatto

Honorary Chair
J. Koenderink

Facilities

Join us..
IN VENEZIA
22-29 oct '17

Travel
Intern. Venice Airport
at 20 min.

General Chairs:
K. Ikeuchi, G. Medioni, M. Pelillo

Organization

Finance Chair :
O. Camps

The Italian(and Venetian) food

To contact me

Rita Cucchiara

Email: Rita.cucchiara @unimore.it
Skype: Rita.cucchiara
Tel: +39 059 205 6136

Imagelab
Dipartimento di Ingegenria Enzo Ferrari
Università di Modena e Reggio Emilia, Italia

Thanks.