Photo Clustering of Social Events by Extending PhotoTOC to a Rich Context

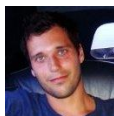Daniel Manchón-Vizuete   Irene Gris-Sarabia   Xavier Giró-i-Nieto

Outline

Photo Clustering of Social
Events by Extending
PhotoTOC to a Rich Context

D. Manchón-Vizuete   I. Gris-Sarabia   X. Giró-i-Nieto

Motivation

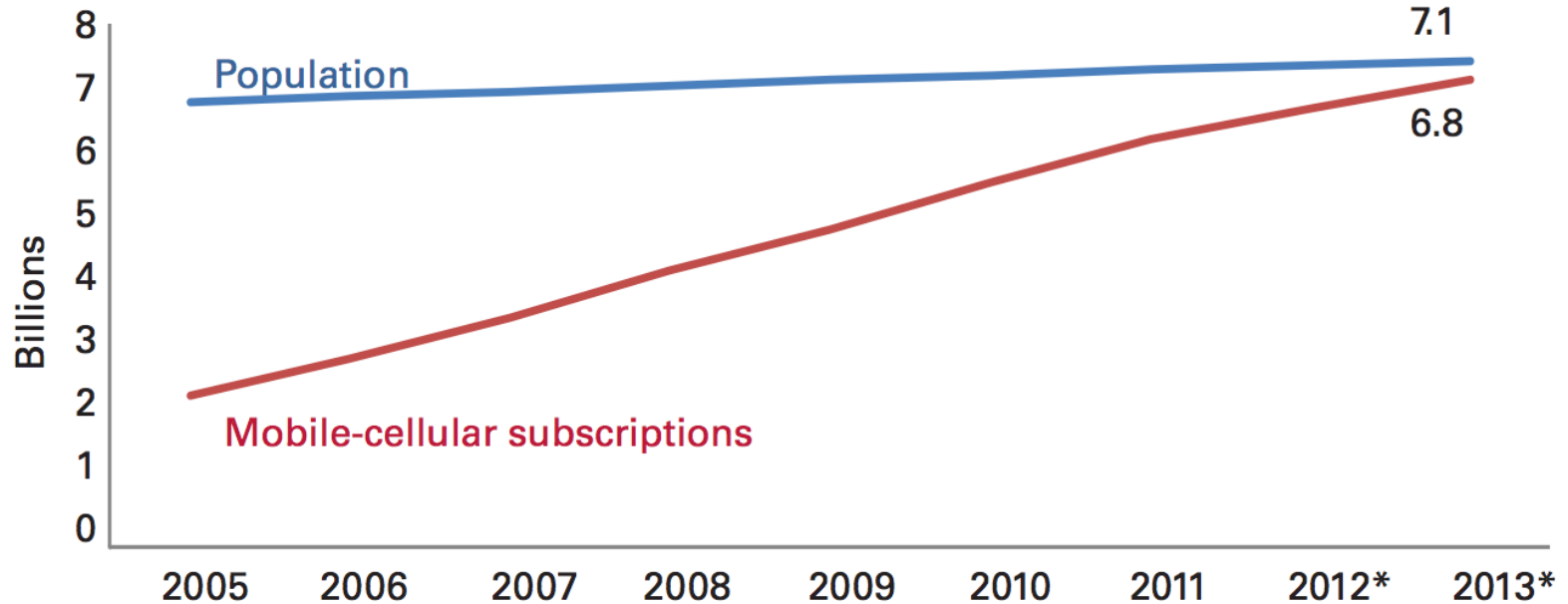Source: ITU World Telecommunication /ICT Indicators database
Note: * Estimate

Motivation
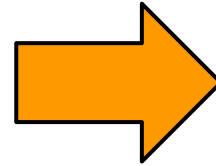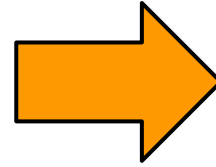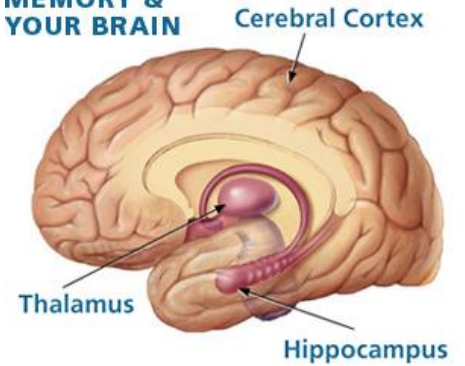
Motivation

Credit: National Institute on Aging and FDA.

Motivation

Motivation

Credit: Alan Smeaton and Cathal Gurrin (Dublin City University, 2013)

Outline

Photo Clustering of Social

Events by Extending

PhotoTOC to a Rich Context

D. Manchón-Vizuete   I. Gris-Sarabia   X. Giró-i-Nieto

Related work

PhotoTOC
[Platt et al, PACRIM 2003]

Related work

Rich context

User-provided annotations → Title Description

Automatically generated information → Creation time Geo-location

[Becker, Naaman & Gravano, ACM WSDM 2010]

Outline

Photo Clustering of Social

Events by Extending

PhotoTOC to a Rich Context

D. Manchón-Vizuete   I. Gris-Sarabia   X. Giró-i-Nieto
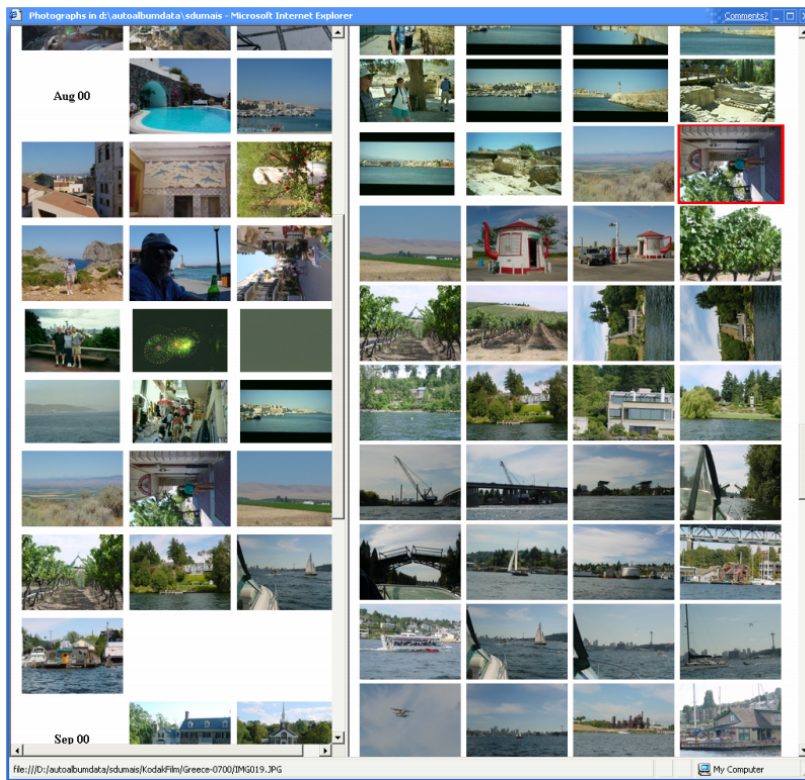
Motivation

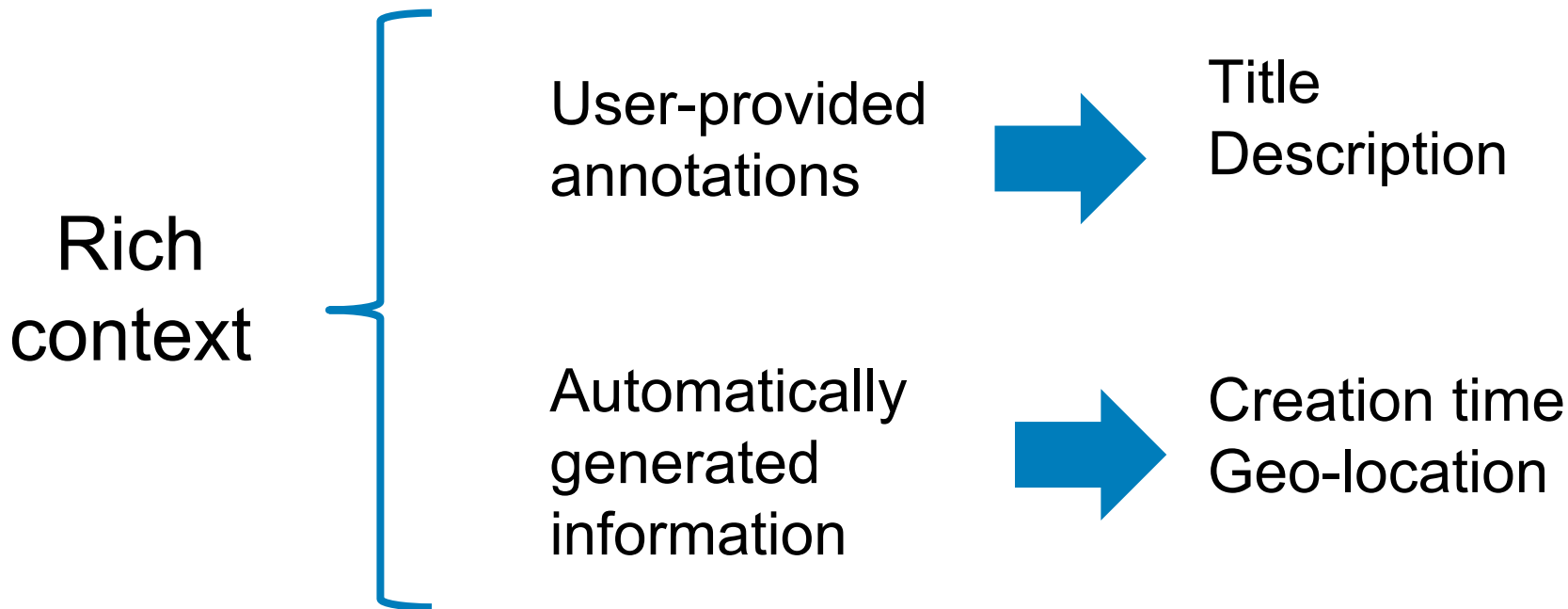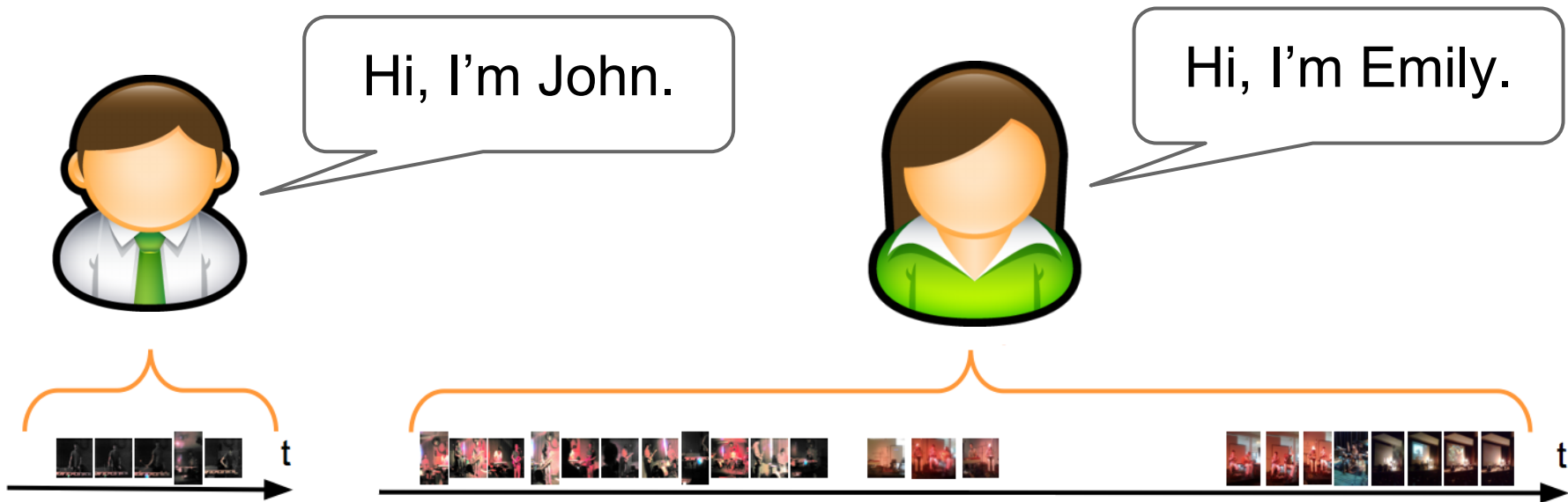Related work

Approach

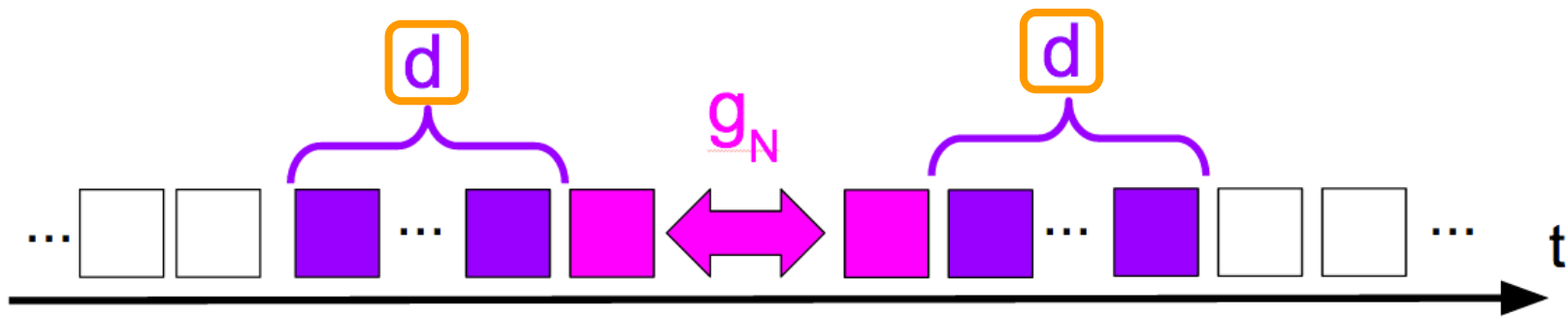Results

Future work

Conclusions

Approach

(a) Temporal sorting by each user independently

Approach

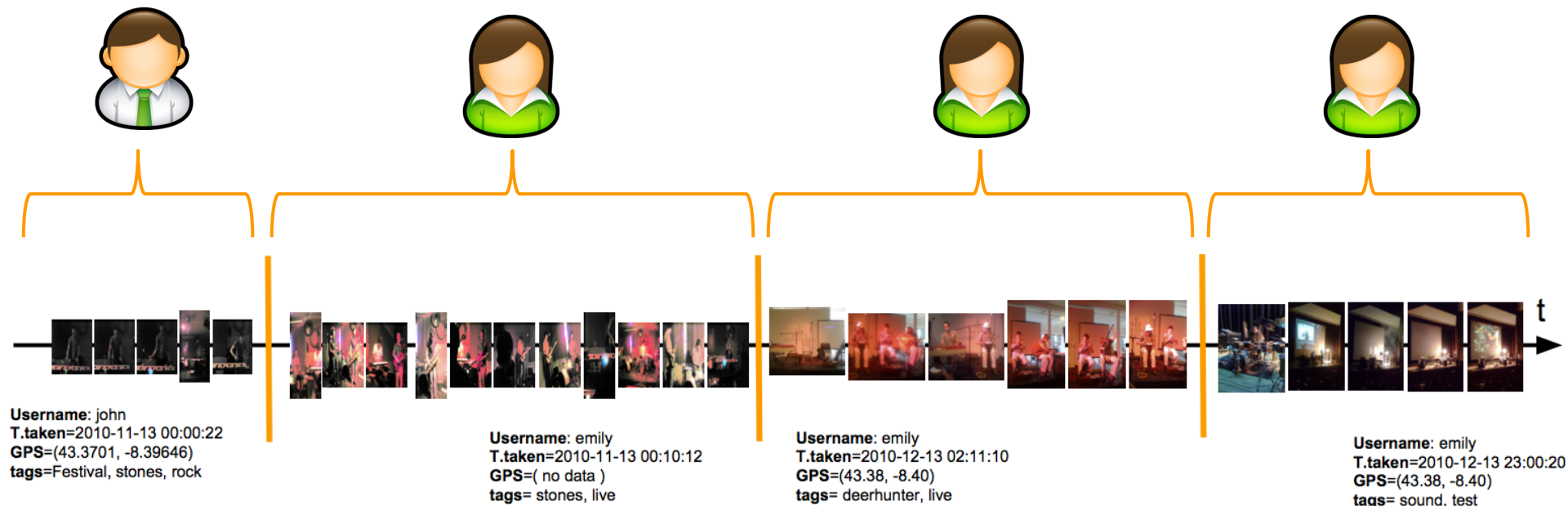## (b) Temporal-based oversegmentation in mini-clusters

$$log(g_N) \geq \boxed{K} + \frac{1}{2d+1} \sum_{i=-d}^{d} log(g_{N+1})$$

PhotoTOC
[Platt et al, PacRim 2003]

Approach

(b) Temporal-based oversegmentation in mini-clusters

Username: john
T.taken=2010-11-13 00:00:22
GPS=(43.3701, -8.39646)
tags=Festival, stones, rock

Username: emily
T.taken=2010-11-13 00:10:12
GPS=(no data)
tags= stones, live

Username: emily
T.taken=2010-12-13 02:11:10
GPS=(43.38, -8.40)
tags= deerhunter, live

Username: emily
T.taken=2010-12-13 23:00:20
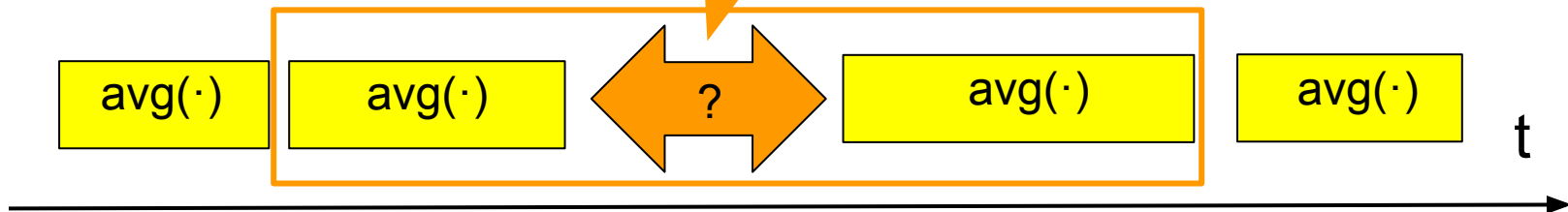**GPS**=(43.38, -8.40)
tags= sound, test

Approach

(c) Sequential merging of mini-clusters

$$x = \sum_i w_i \bar{s}_i \; < \; \gamma$$

Approach

(c) Sequential merging of mini-clusters
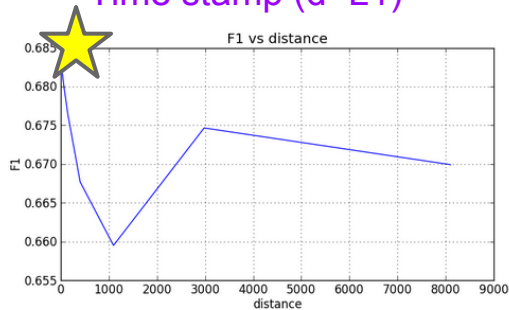
$$x = \sum_i \boxed{w_i} \bar{s}_i < \gamma$$

⬇ Weighted modalities

- creation (or upload) time
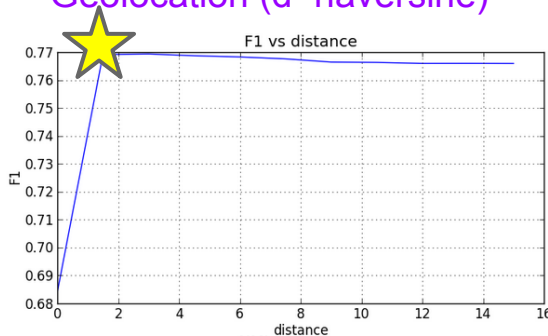- geolocation
- textual labels
- same user

Approach
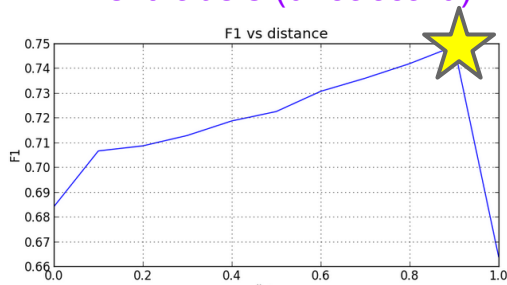
(c) Sequential merging of mini-clusters
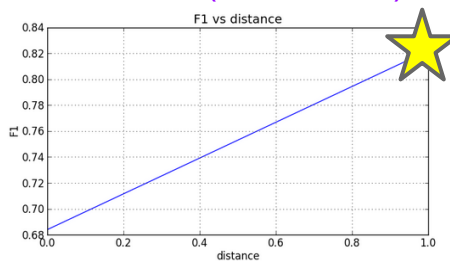

Time stamp (d=L1)

Geolocation (d=haversine)

Text labels (d=Jaccard)

Same user (d=boolean)

$$w_i = \frac{max\ F_1^i}{\sum_j max\ F_1^j}$$

Approach

(c) Sequential merging of mini-clusters

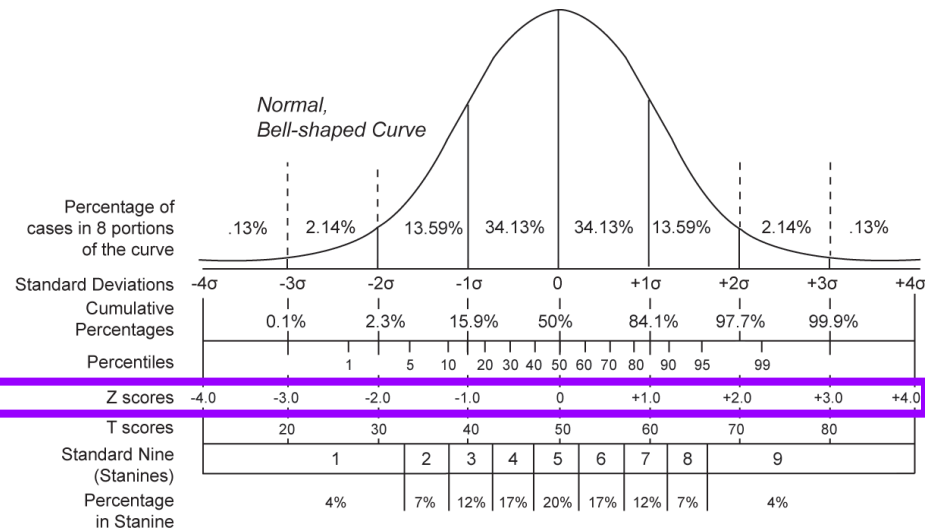|	Geolocated	No geolocated
Geolocation	0.28	-
User ID	0.44	0.60
Tags	0.22	0.30

Approach

(c) Sequential merging of mini-clusters

Mean and std. deviation learned on pairs of photos within the same training event.

$$x = \sum_i w_i \bar{s}_i \ < \ \boldsymbol{\gamma}$$

$$\bar{s}_i = \Phi(d_i, \mu_i, \sigma_i) = \frac{1}{2}\left[1 + erf\left(\frac{d_i - \mu}{\sqrt{2\sigma^2}}\right)\right]$$

Approach

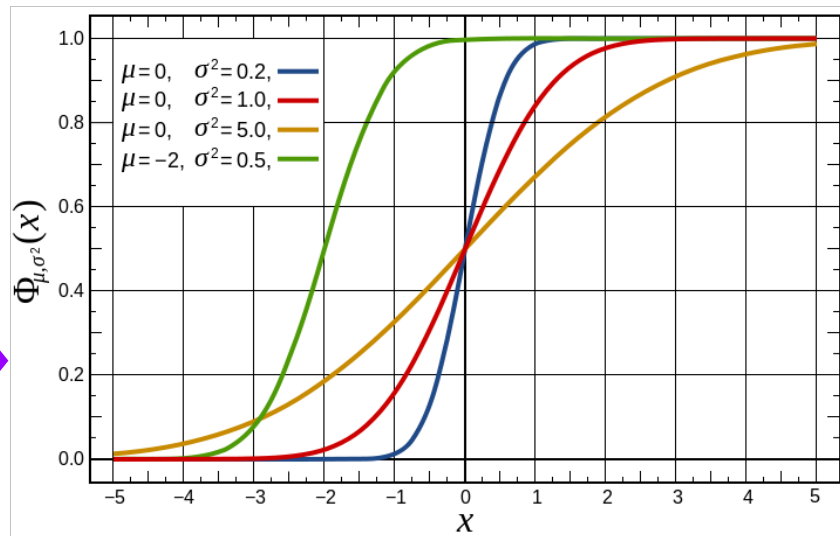## (c) Sequential merging of mini-clusters

$$x = \sum_i w_i \bar{s}_i \ < \ \gamma$$

$$\bar{s}_i = \boxed{\Phi\left(d_i, \mu_i, \sigma_i\right)} = \frac{1}{2}\left[1 + erf\left(\frac{d_i - \mu}{\sqrt{2\sigma^2}}\right)\right]$$
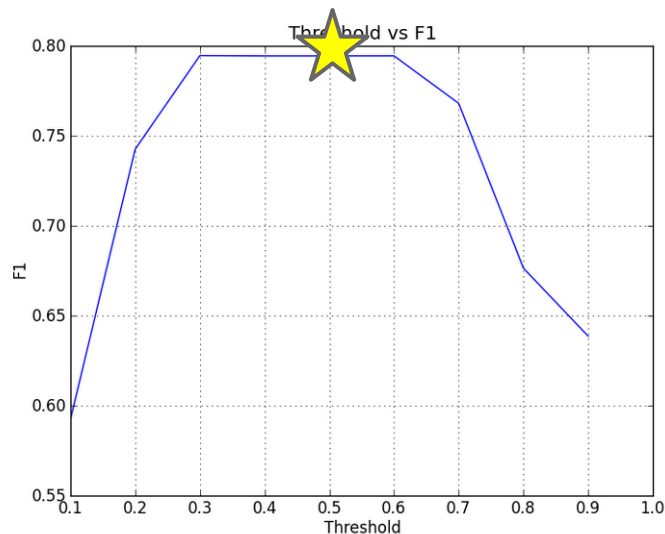
phi function

Approach

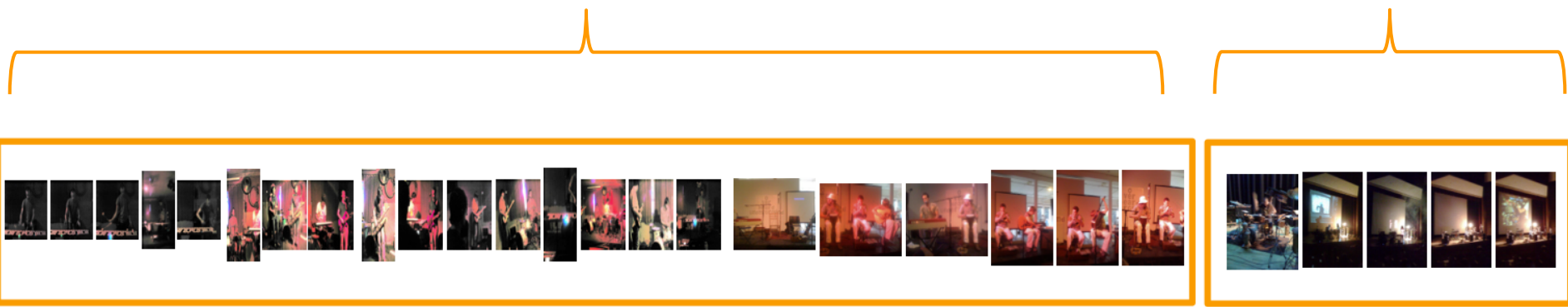## (c) Sequential merging of mini-clusters

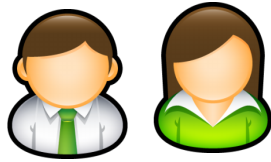$$x = \sum_i w_i \bar{s}_i < \boxed{\gamma}$$

decision threshold

Approach

(c) Sequential merging of mini-clusters

Outline

Photo Clustering of Social
Events by Extending
PhotoTOC to a Rich Context

D. Manchón-Vizuete   I. Gris-Sarabia   X. Giró-i-Nieto
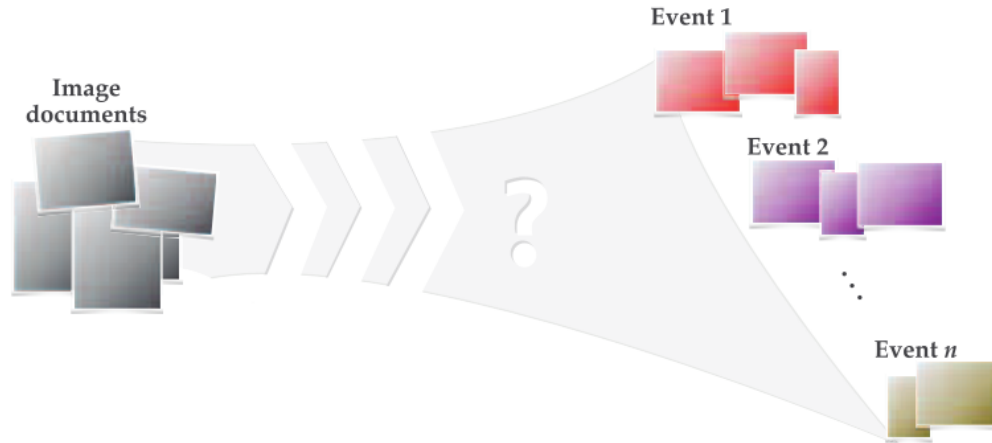
Motivation

Related work

Approach

Experiments

Future work

Conclusions

Dataset

ReSEED - Social Event Detection Dataset

Results

	F1-Score	Precision
Samangooei et al [20]	0.9454	0.96
Nguyen et al [14]	0.9234	0.98
Our work	**0.8833**	**0.96**
Witsuba et al [23]	0.8720	0.91
Sutanto et al [21]	0.8112	0.86

Results

	PhotoTOC [16] K=log(17), d=10	Our work K=log(600), d=14
Time	0.749	0.880
Time+Geolocation	0.802	**0.893**
Time+User ID	0.837	0.875
Time+Tags	0.814	0.883
Time+Fusion	0.822	0.883

Results 👍
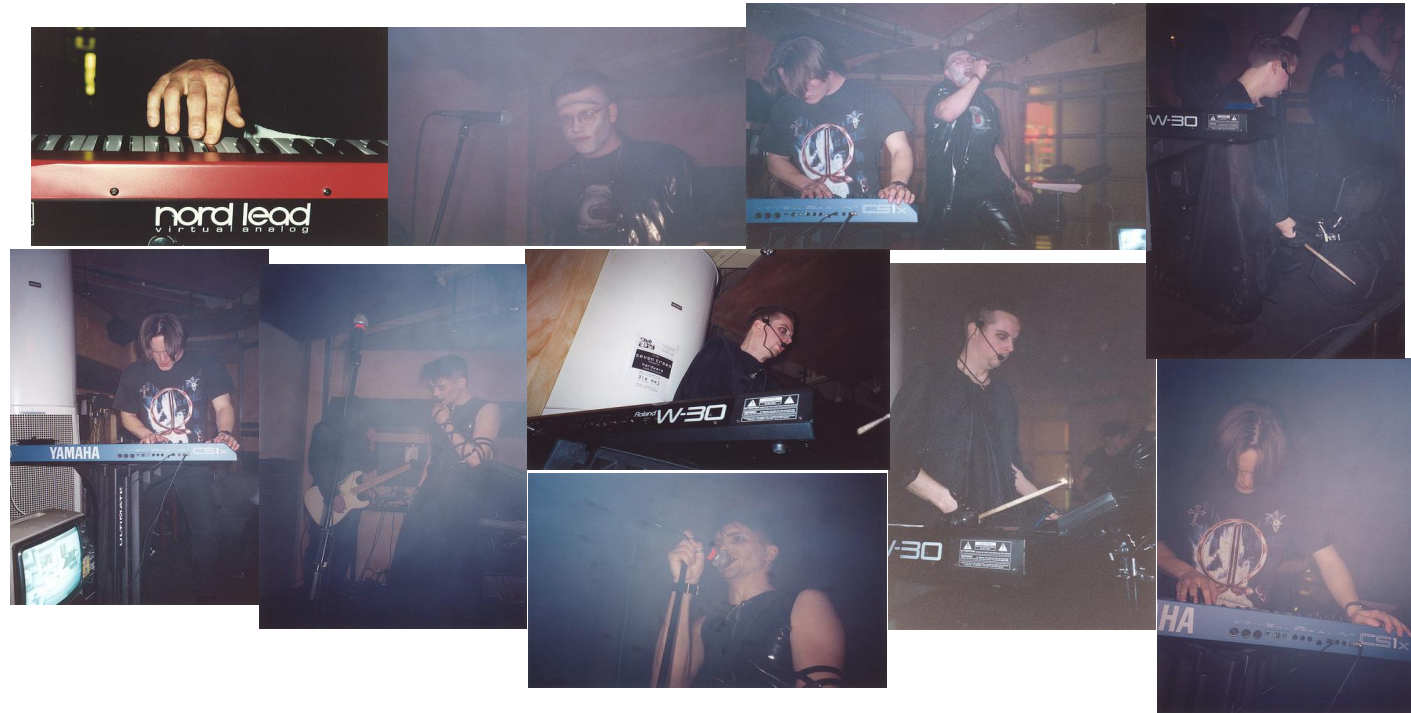
Diversity of cameras (qualities)

Results

Diversity of locations

Results 👍

Visual redundancy (not exploited)

Results

False split

Results 👎

False merge

Outline
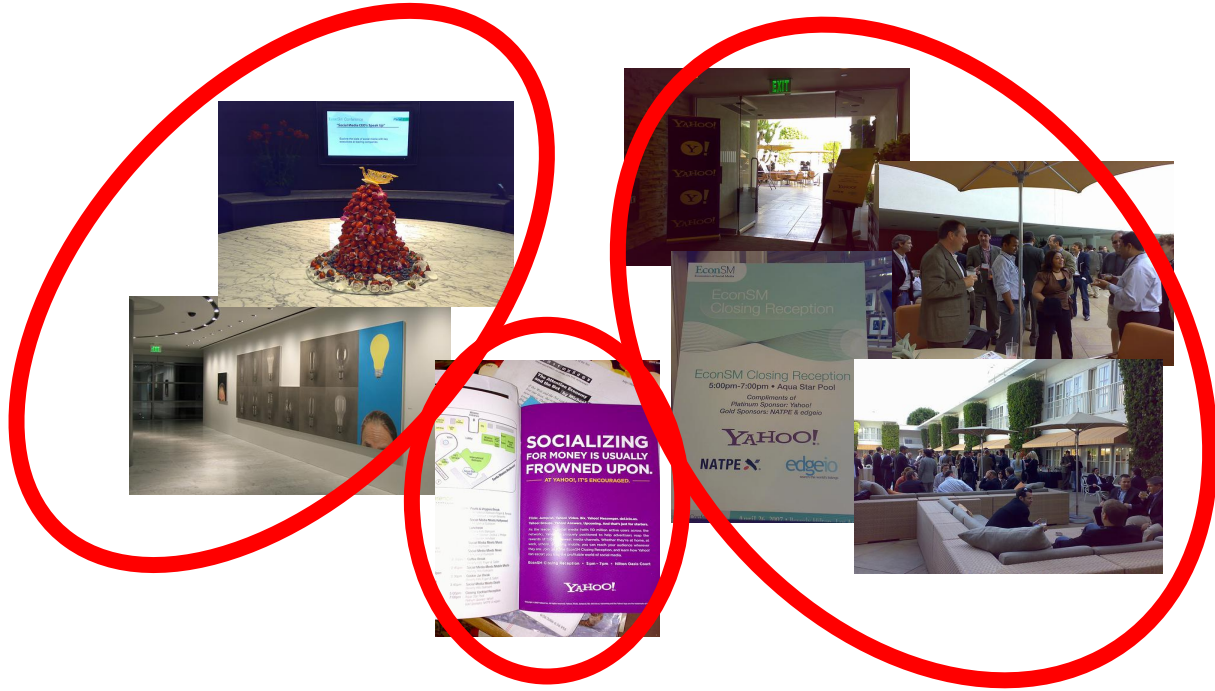
Photo Clustering of Social
Events by Extending
PhotoTOC to a Rich Context

D. Manchón-Vizuete   I. Gris-Sarabia   X. Giró-i-Nieto

Motivation
Related work
Approach
Experiments
Future work
Conclusions

Future work

- Median-based normalisation for robustness against outliers.

- Better optimisation of parameters K and d.

- Improve fusion approach (SVM ?).

- Event-dependent merging criteria.

- Efficient use of visual features (higher computation).

Outline

Photo Clustering of Social

Events by Extending

PhotoTOC to a Rich Context

Motivation
Related work
Approach
Experiments
Future work
Conclusions

D. Manchón-Vizuete

I. Gris-Sarabia

X. Giró-i-Nieto

Conclusions

- All considered context (geo, user, tags) can help.

- Watch out for outliers.

- Divide and conquer… sequentially.

Thank you SEWM !

Thank you !

✉ xavier.giro@upc.edu

🐦 @DocXavi

📝 http://bitsearch.blogspot.com

f /ProfessorXavi

🌐 https://imatge.upc.edu/web/people/xavier-giro