NII team: Query-adaptive asymmetrical dissimilarities for instance search

Cai-Zhi Zhu [1], Hervé Jégou [2], Shin'ichi Satoh [1]
[1] NII, Tokyo Japan [2] INRIA, Rennes, France
Instance Search Task, TRECVID
November 20, 2013

大学共同利用機関法人 情報・システム研究機構
NII 国立情報学研究所
National Institute of Informatics

Outline

- Asymmetrical method
  - Instance search is inherently asymmetric.
  - Query-adaptive asymmetrical dissimilarities.
- INS2013 submission
  - Experimental settings.
  - Performance.

Outline

- Asymmetrical method
  - Instance search is inherently asymmetric.
  - Query-adaptive asymmetrical dissimilarities.
- INS2013 submission
  - Experimental settings.
  - Performance.

Behind this method

- A joint work between NII and Dr. Hervé Jégou at INRIA.
- Accepted to ICCV2013.

The problem

- Instance search is inherently asymmetric
  - the query object is mostly included in the database video, while the converse is not necessarily true.
- However, existing BoW approaches mostly compare query ROIs and database videos with symmetrical measures
  - L_1 and L_2 distance metrics are mostly used, while they are symmetrical.

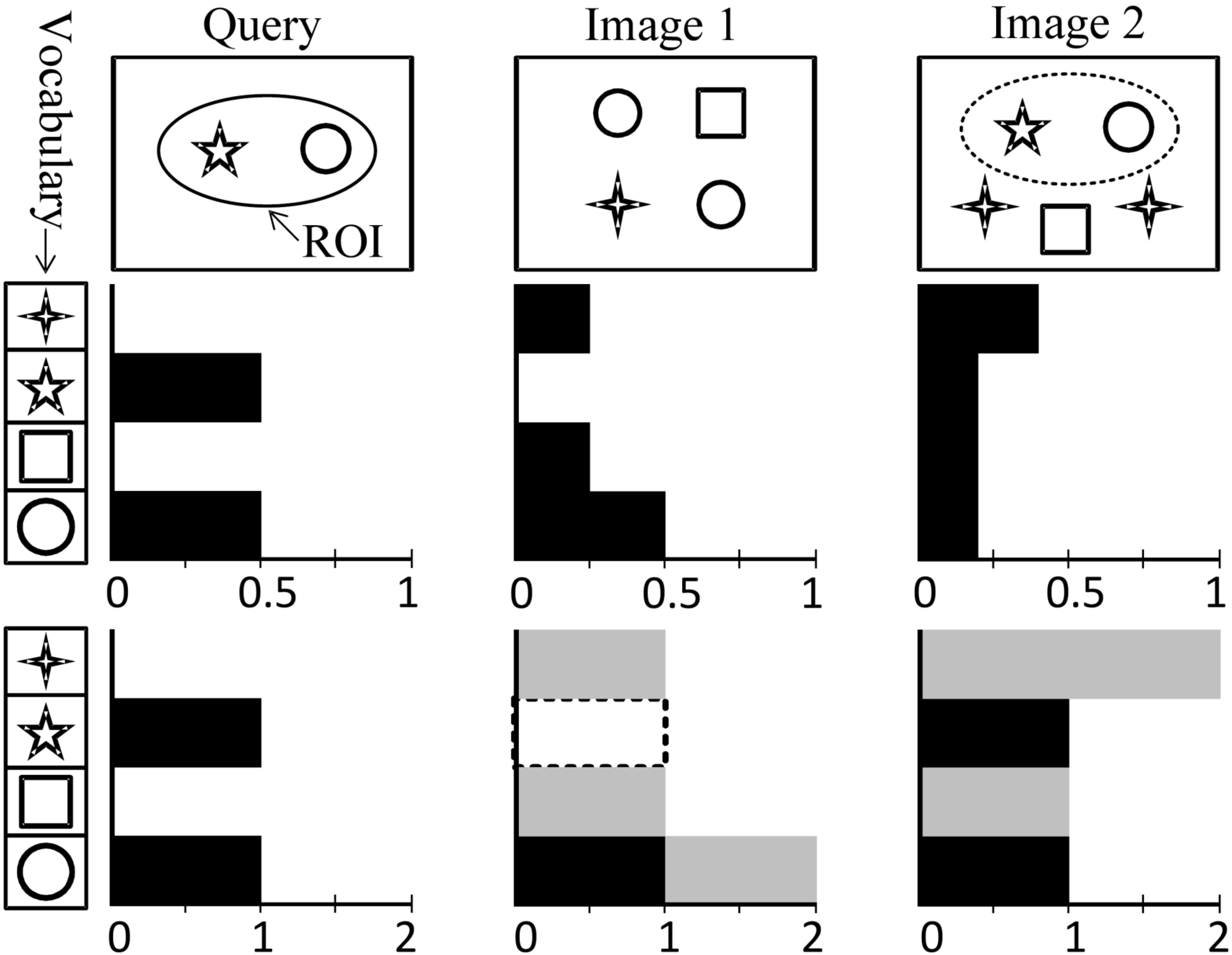
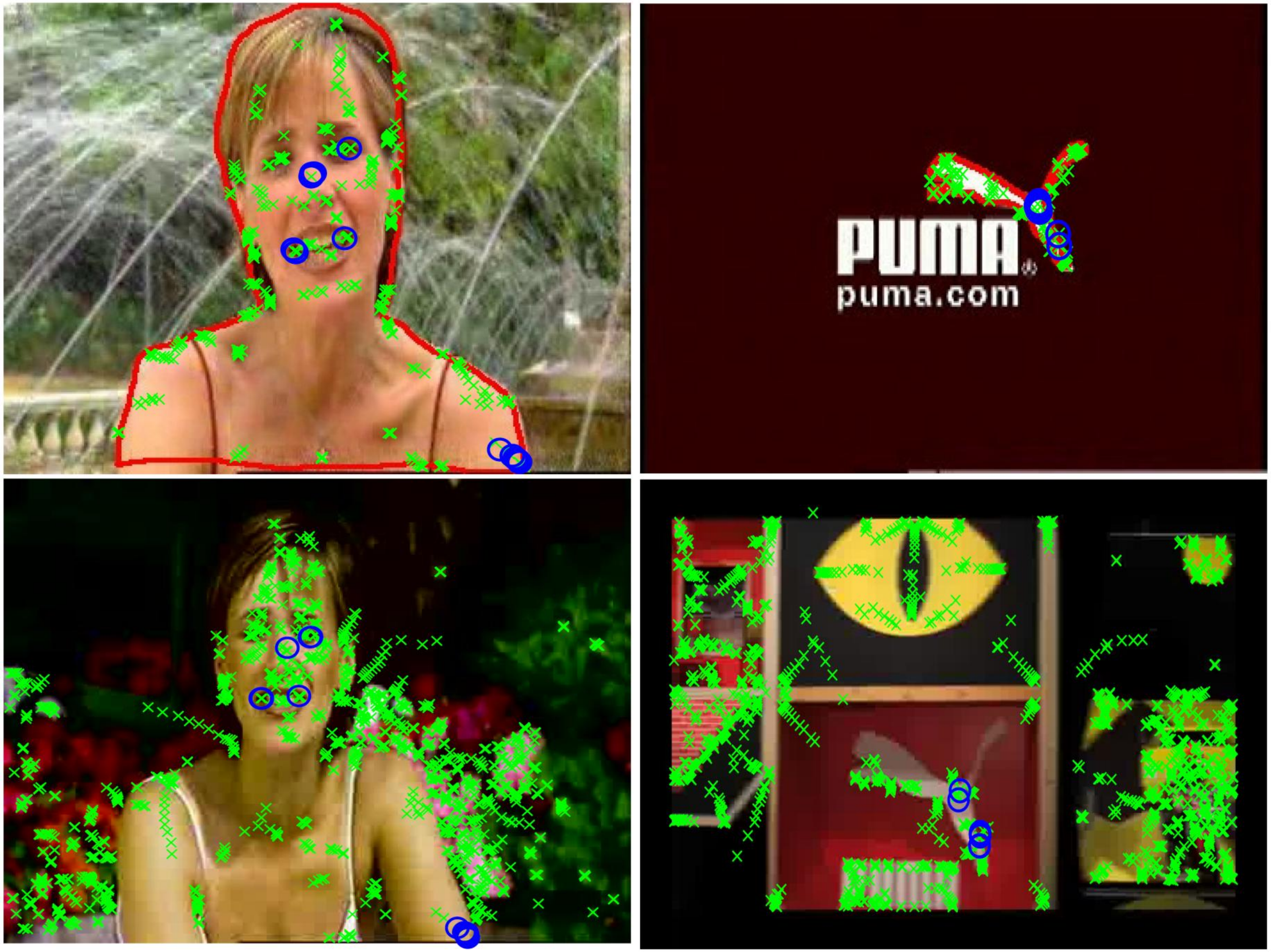


Fig. 1: A toy example comparing the standard scoring method in the second row ($\ell_1(\mathbf{Q}, \mathbf{T}_1) = 1, \ell_1(\mathbf{Q}, \mathbf{T}_2) = 1.2$) with our asymmetrical dissimilarity in the third row ($\delta_1(\mathbf{Q}, \mathbf{T}_1, \infty) = 1, \delta_1(\mathbf{Q}, \mathbf{T}_2, \infty) = 0$).



INS2011 INS2012

Fig. 2: Examples visualizing the asymmetrical inlier/outlier ratio on the query and database side on each benchmark.

Table 1: Symmetrical ℓ_1 *vs.* (query-adaptive) asymmetrical δ_1. Note the asymmetrical methods are compatible with an inverted index.

Symmetrical	$\ell_1(\mathbf{Q}_i, \mathbf{T}_j) = \left\| \frac{\mathbf{Q}_i}{\|\mathbf{Q}_i\|_1} - \frac{\mathbf{T}_j}{\|\mathbf{T}_j\|_1} \right\|_1$
Asymmetrical	$\delta_1(\mathbf{Q}_i, \mathbf{T}_j, w) = \|\mathbf{T}_j\|_1 - w \left\|\min(\mathbf{Q}_i, \mathbf{T}_j)\right\|_1$
Query-adaptive	$\delta_1(\mathbf{Q}_i, \mathbf{T}_j, a) = \|\mathbf{T}_j\|_1 - \boxed{\alpha \frac{\sum_{j=1}^{N} \|\mathbf{T}_j\|_1}{\sum_{j=1}^{N} \|\min(\mathbf{Q}_i, \mathbf{T}_j)\|_1}} \left\|\min(\mathbf{Q}_i, \mathbf{T}_j)\right\|_1$

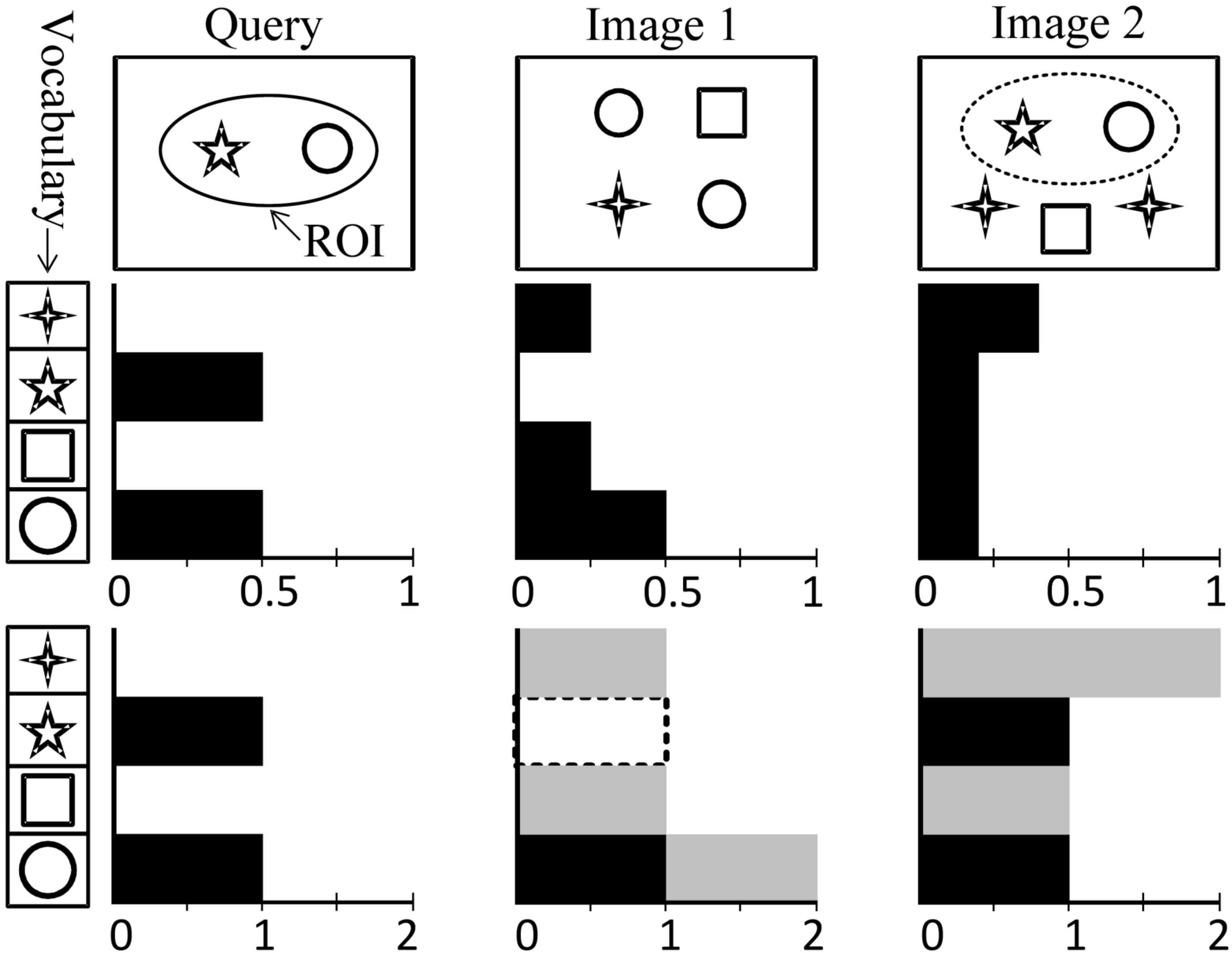


Fig. 1: A toy example comparing the standard scoring method in the second row ($\ell_1(\mathbf{Q}, \mathbf{T}_1) = 1, \ell_1(\mathbf{Q}, \mathbf{T}_2) = 1.2$) with our asymmetrical dissimilarity in the third row ($\delta_1(\mathbf{Q}, \mathbf{T}_1, \infty) = 1, \delta_1(\mathbf{Q}, \mathbf{T}_2, \infty) = 0$).

Table 2: Performance obtained with different parameter w and α.

Configurations	Oxford105K	Oxford105K*	INS2011	INS2012
$\delta_1(\mathbf{Q}, \mathbf{T}, 0)$	0.3	0.3	0.02	0
$\delta_1(\mathbf{Q}, \mathbf{T}, 1)$	2.79	2.78	0.02	0
$\delta_1(\mathbf{Q}, \mathbf{T}, \infty)$	65.29	38.85	44.88	19.51
$\delta_1(\mathbf{Q}, \mathbf{T}, w_{\text{opt}})$	75.38	55.81	47.38	20.88
$\delta_1(\mathbf{Q}, \mathbf{T}, \alpha_{\text{opt}})$	78.14	61.05	48.50	21.73
$\ell_1(\mathbf{Q}, \mathbf{T})$	73.88	54.47	45.16	19.83

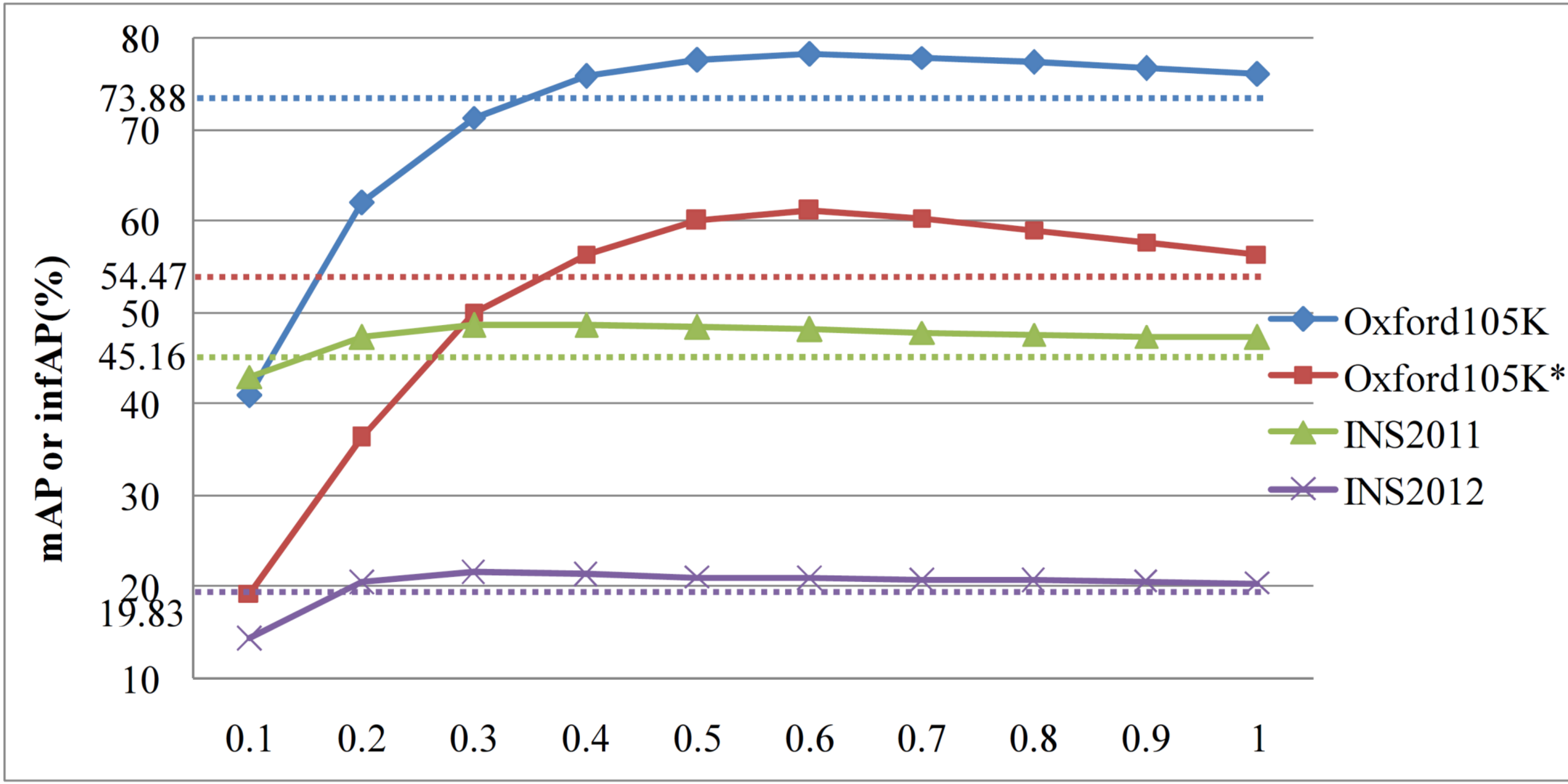


Fig. 3: Impact of the parameter α (horizontal axis) on the performance (vertical axis) of the δ_1 asymmetrical dissimilarity.

More details please refer to our recent paper:

C.-Z. Zhu, H. Jégou, and S. Satoh. Query-adaptive asymmetrical dissimilarities for visual object retrieval. Accepted to ICCV, 2013.

Outline

- Asymmetrical method
  - Instance search is inherently asymmetric.
  - Query-adaptive asymmetrical dissimilarities.
- INS2013 submission
  - Experimental settings.
  - Performance.

INS2013 dataset

- 469,539 shots from 243 videos (~433.5 hours after the shot0_* being excluded)
  - Recall that 20,982 and 76,751 shots are in the INS2011 and INS2012, respectively.
- 30 query topics, 4 images each.

- 1+4 non-rigid objects

- 25 rigid objects
  - 8 ‘big’ objects

  - 7 logos

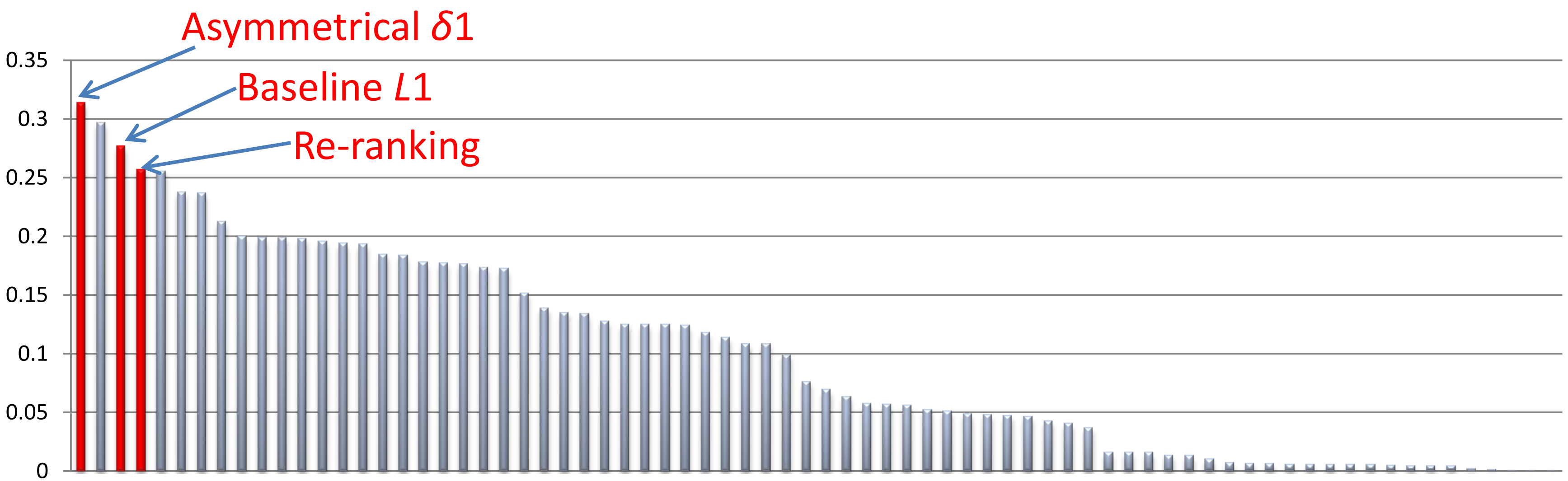
  - 10 other ‘small’ objects

Our submission

- Experimental settings
  - Sample 5 frames/sec.
  - SIFT only
    - 3 detectors: Hessian-affine, Harris-Laplace and MSER.
    - 2 descriptors: Root-SIFT and color SIFT.
- Three BoW based submissions
  - LO-RANSAC re-rank, Hessian-affine Root-SIFT only.
  - Asymmetrical δ1 dissimilarity with multiple SIFTs.
  - Baseline L1 distance with multiple SIFTs.

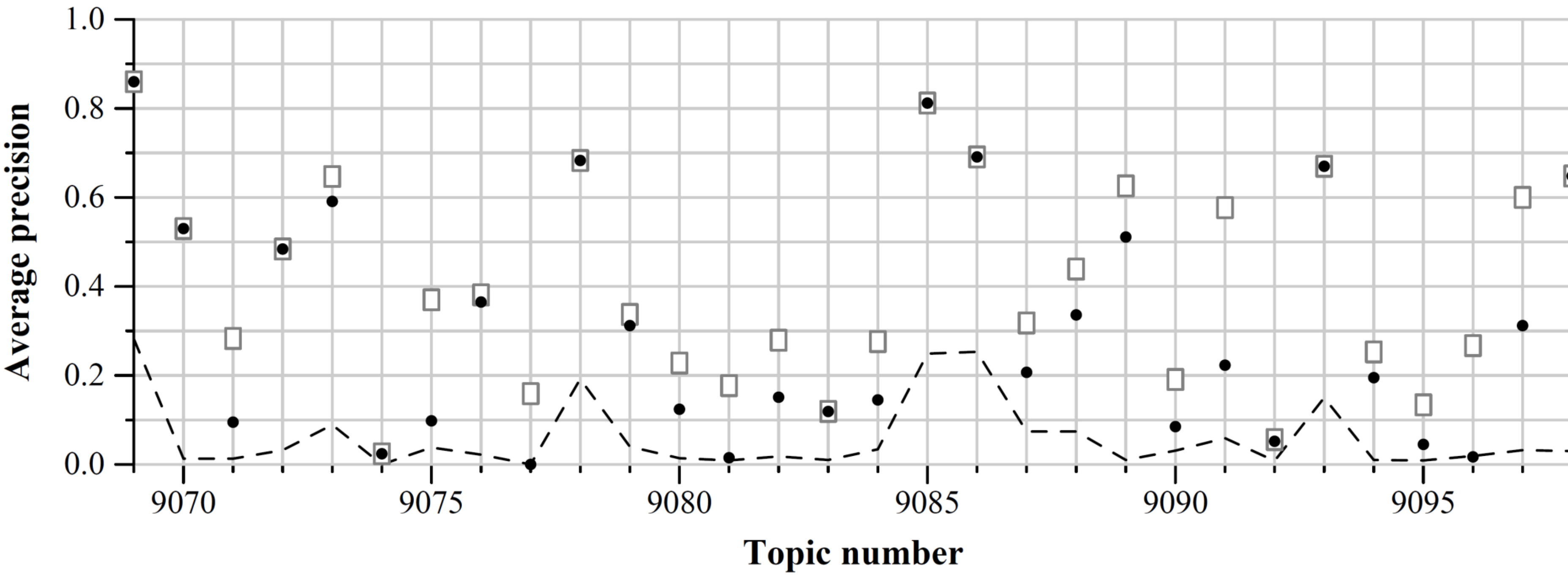
Performance table list

- We are ranked 1st, 3rd and 4th, among 74 submissions.



Performance per topic of the best run

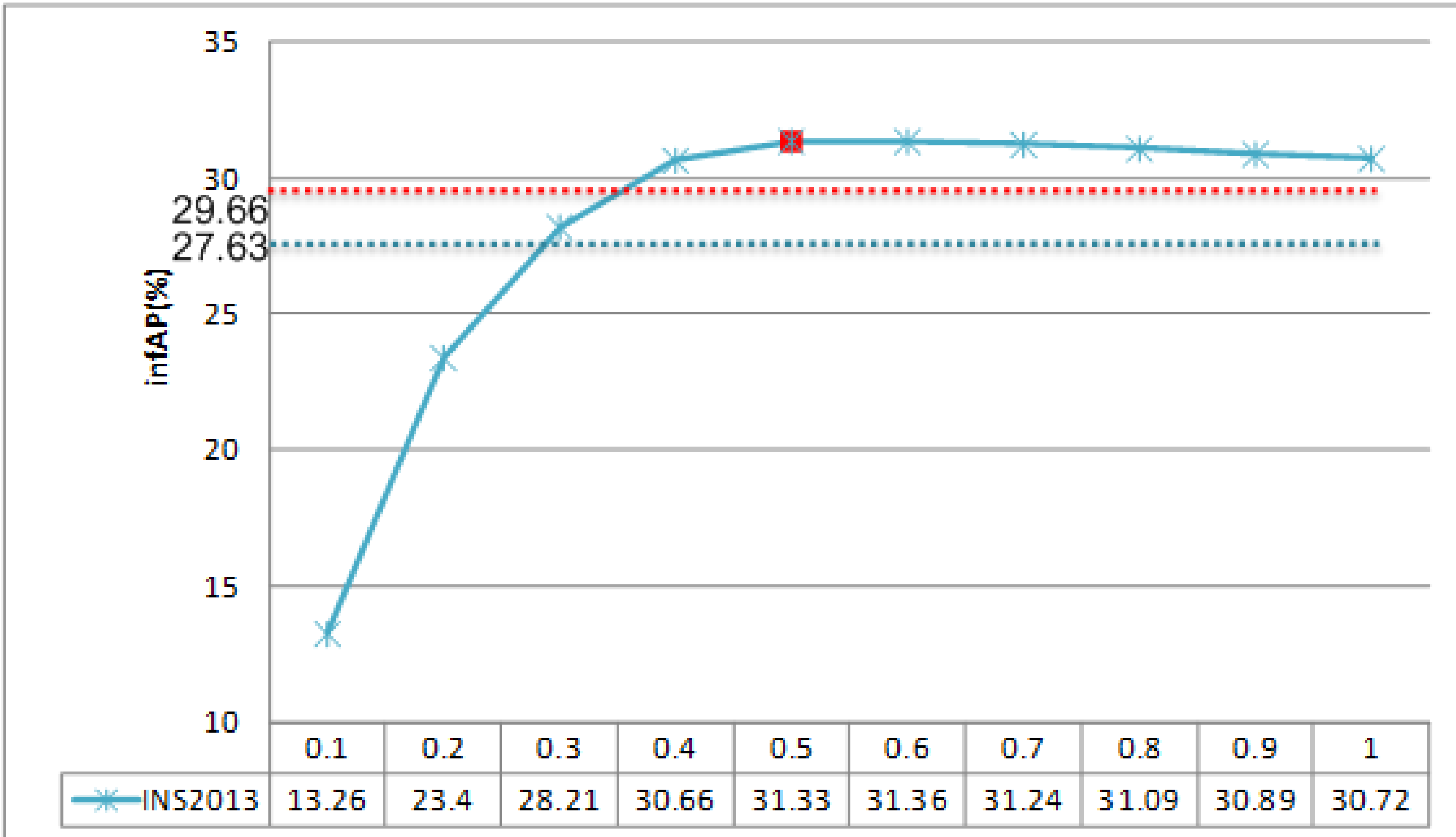
- In total we won 12 topics out of 30, in which the spatial re-ranking method contributes two



Run score (dot) versus median (---) versus best (box) by topic


9069
9070
9071
9072
9073
9074
9075
9076
9077
9078
9079
9080
9081
9082
9083
9084
9085
9086
9087
9088
9089
9090
9091
9092
9093
9094
9095
9096
9097
9098

α *vs.* infAP on INS2013



Thank you!