Petr Gronat

petr.gronat@inria.fr

INRIA Paris

23 av. d'Italie   
75013, Paris  
France

Dear Editor / Reviewer,

I would like to take this opportunity to explain the relationship of this manuscript to our previous publication [1] “Learning and calibrating per-location classifiers for visual place recognition”. Most significantly:

1. We have added a new section 5 that describes a special form of affine calibration that does not require construction of empirical cumulative density function.
2. We have added new section 6 that describes new memory efficient classifier representation.
3. We provide additional results in section 8 including: (i) an analysis of recognition accuracy vs. compactness (section 8.3 and figure \*\*\*), (ii) evaluation of the new affine calibration method (table 3), (iii) results using compact Fisher vector descriptors (section 8.2, table \*\*\* and figure \*\*\*), and (iv) new qualitative examples (figure \*\*\*).
4. We have included additional details such as pseudo code for the p-value calibration in Algorithm I and Algorithm II.

The differences are explained in detail below.

**Section 1** This introduction section is an extension of introduction from [1].

**Section 2** This is a new section addressing related work, which addresses state-of-the-art methods for place recognition, per-exemplar support vector machines and classifier score calibration methods.

**Section 3** This section corresponds to Section 2 of [1], a new paragraph (Section 3.2) that motivates the calibrating classifier scores has been added.

**Section 4** This section extends Section 3 of [1]. This section contains a new subsection 4.3 that summarizes the calibration procedure. It also contains new pseudo codes (Algorithm I, Algorithm II) describing the p-value calibration in offline and online stage.

**Section 5** This is anew section that addresses the complexity of p-value calibration and proposes an affine classifier calibration by renormalization that does not need any positive or negative training data.

**Section 6** This is a new section that addresses efficient representation of the classifiers. It first motivates the need for compact classifier representation and then it proposes an efficient representation based on dual form of the learnt SVMs.

**Section 7** This section extends Section 4 of [1].

**Section 8** While [1] includes results only for the bag-of-visual-words representation, we have added new results on the compact Fisher vector descriptors (section 8.2) and an analysis of recognition accuracy vs. compactness (section 8.3).

**Appendix.** We have added an appendix that contains an analysis of the per-exemplar SVM objective that provides additional intuition for the calibration by normalization.

Kind regards

Petr Gronat