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 and recent publications “Learning and calibrating per-location classifiers for visual place recognition”, presented at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2013 and “Fisher vector places: learning compact descriptors for place recognition”, submitted to the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015.

The manuscript is a fusion and extension of two above mentioned papers. We have expanded several sections of the papers, most significantly we have (i) interpreted the CVPR 2015 method as a special form of calibration (Section 5), (ii) we have added new section about Memory efficient classifier representation (Section 6) and (iii) we provide analysis of recognition accuracy vs. compactness (Section 8). Please find the CVPR 2013 and CVPR 2015 papers included with this submission. The differences are explained in detail below:

**Section 1** This introduction section is an extension of both introduction sections of CVPR2013 and CVPR2015 submissions.

**Section 2** This related work section mostly corresponds to Section 2 of the CVPR 2015 paper, however the section is significantly modified.

**Section 3** This section corresponds to Section 2 of the CVPR 2013 paper, a new paragraph (Section 3.2) motivates for calibrating classifier scores.

**Section 4** This section extends Section 3 from CVPR 2013 paper. The structure of the section and text are modified and we have also added two pseudo codes (Algorithm I, Algorithm II) to explain the p-value calibration in offline and online stage.

**Section 5** This is a new section explains the main drawbacks of the p-value calibration procedure and motivates for usage of simpler calibration model. We take an idea of CVPR 2015 paper and explains the method in the context of affine calibration function.

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

**Section 7** This section is a fusion of Section 4 from CVPR 2013 and Secitons 4.1 and 4.2 from CPVR 2015 submission.

**Section 8** This section presents the experimental validation of the proposed methods. Results correspond to Section 4.3 of the CVPR 2015, but for the e-SVM FV128 descriptors perform marginally better as we have found better regularization parameters parameters.

* + Section 8.3 In this section we analyze accuracy of the learnt representations vs. their compactness measured by their memory footprint.

**Figure 4** The figure shows a recall curves for bag-of-visual-words baseline and learnt representations.

**Figure 6** A new figure with examples with correctly and incorrectly localized queries learnt for the bag-of-visual-word representation.

**Figure 8** A new figure with examples with correctly and incorrectly localized queries learnt for the Fisher vector representation.

Kind regards

Petr Gronat