

An improved plant identification system by fuzzy c-means bag of visual words model and sparse coding

ABSTRACT

Current work adopts the Fuzzy c-means Bag of Visual Words model and sparse coding for plant identification. Plant identification has become a significant research area in the botany field in recent years. SIFT features are distinctive invariant features based on scale-space because of the situation of its robust identical matching capabilities. Bag of visual words (BoVW) model and its variants are used effectively for the retrieval of images by many researchers. Classic bag of visual words algorithm is based on k-means clustering and every SIFT features belongs to one cluster and it leads to decreasing classification results. Data entities may belongs to further than one cluster in the fuzzy clustering (soft clustering), and a set of membership levels are allied with each group. This demonstrate the intensity of the correlation between that aspect of data and a specific cluster. In the classic Bag of visual words model, the Fuzzy c-means algorithm is replaced with K-means and the accuracy of SIFT matching is increased. Moreover, sparse coding has been commonly used in recent years for the purposes of retrieving and identifying images. The pure picture patch computes the atoms in an over-complete dictionary by adding them sparsely. Sparse representation prevents over-fitting in the classifier by eliminating redundancies and evaluating high-frequency patterns between feature vectors. Performance of proposed methods surpass the classic bag of words algorithm for plant identification tasks.

Keyword: Plant identification; Fuzzy C-means algorithm; SIFT Descriptor; Bag of Visual Words algorithm; Sparse coding