SCENE AND OBJECT RECOGNITION WITH SUPERVISED NONLINEAR NEIGHBORHOOD EMBEDDING

XIAN-HUA HAN AND YEN-WEI CHEN

College of Information Science and Engineering
Ritsumeikan University
56-1 Toji-in, Kasatsu-shi 603-8577, Japan
hanxhua@fc.ritsumei.ac.jp

ABSTRACT. *Image category recognition is important to access visual information on the level of objects and scene types. In this paper, we develop a Supervised Nonlinear Neighborhood Embedding (SNNE) subspace algorithm of different visual features for object and scene recognition, which learns an adaptive nonlinear subspace by preserving the neighborhood structure of the visual feature space. In the proposed subspace algorithm, we combine the idea of nonlinear kernel mapping and preserving the neighborhood structure of the samples, so it can not only gain a perfect approximation of the nonlinear image manifold, but also enhance within-class neighborhood information. So, the proposed SNNE algorithm models the ensemble of visual features to a more discriminative space for category recognition, and at the same time, can effectively combine several visual features to improve recognition rate. The proposed method is evaluated by using the scene database (SIMPLicity) and object recognition database (Caltech). We confirm that the proposed method is much better than state-of-the-art methods only with simple visual features.*

Keywords: Scene and object recognition, Supervised nonlinear neighborhood embedding, Visual features, Subspace learning

1. **Introduction.** Image category recognition is important to access visual information on the level of objects (motorbikes, cars, etc.) and scene-like types (beaches, mountains, foods, etc.), and it has a wide range of applications, such as intelligent image processing and content-based image indexing and retrieval (CBIR) [1-3]. In CBIR, an efficient and effective classification method can significantly improve the retrieval accuracy by removing the irrelevant images. Image classification has posed a significant challenge to the research community of computer vision due to interclass variability, illumination and scale changes. The rich context of an image makes the semantic-understanding (scene and object recognition) very difficult. In image categorization problem, only category label is given and the location, scaling and pose of interesting part are unknown. Some categories such as cars or horses include ones of various color, position and shapes. Thus, the within-class variation is large, and image categorization problem is more difficult than conventional classification problem such as face recognition and difficult to detect the interesting region as the preprocessing for classification.

Recently, many authors use the feature detectors such as Harris operator [4] or SIFT operator [5] to detect the characteristic local regions (keypoints), and use Support Vector machine (SVM) for image classification with bag-of-feature histogram [6-8]. However, there are still many problems to discuss: how to sample patches, how to describe them, how to characterize the resulting distributions and how to classify images based on the result (Is it optimal to classification with SVM), how to combine other features with the