Fusing Subcategory Probabilities for Texture Classification

Yang Song[1], Weidong Cai[1], Qing Li[1], Fan Zhang[1], David Dagan Feng[1], Heng Huang[2]
[1]BMIT Research Group, School of IT, University of Sydney, Australia.
[2]Department of Computer Science and Engineering, University of Texas, Arlington, USA.

Texture provides important information for many computer vision applications, such as material classification and scene and object recognition. Accurate classification of texture images is however quite challenging. Some of the main challenges include the wide variety of natural texture patterns and large intra-class variation caused by illumination and geometric changes, and relatively low inter-class distinction.

Different from the current studies in texture classification, which mostly focus on designing new texture feature descriptors, our aim is to improve the classification accuracy with a new classification model using existing features. In this paper, we propose a sub-categorization model for texture classification. We first design a locality-constrained subspace clustering method to efficiently generate subcategories of individual classes. At the subcategory-level, two probability measures are computed based on between-subcategory distinctiveness and within-subcategory representativeness, to quantify the probability of a test data belonging to each subcategory. The subcategory probabilities are then fused weighted by contribution level and cluster quality together with class-level probabilities to classify the test data. An overview of our method flow is shown in Figure 1.

Subcategory generation. We modify the sparse subspace clustering (SSC) [3] algorithm with locality constraints to enhance the efficiency for obtaining the sparse representation coefficient matrix Z:

$$\min_{\{z_i\}} \sum_{i=1}^{N} \|x_i - X z_i\|^2 + \lambda \|d_i \odot z_i\|^2 \;\; s.t. \;\; \mathbf{1}^T z_i = 1, \;\; \|z_i\|_0 \leq P, \;\; z_{ii} = 0, \;\; \forall i \quad (1)$$

Here the second term is adopted from the locality-constrained linear coding (LLC) [7] to encourage smaller coefficients to be assigned to samples that are more different from x_i. In our formulation, the dataset X contains the training data of one class. The clustering outputs of X correspond to the subcategories of that class.

Subcategory probabilities. We design two types of probability estimates: the between-subcategory distinctiveness and within-subcategory representativeness. The between-subcategory distinctiveness is obtained based on binary classification between a subcategory S_{ck} of class c and all subcategories $\{S_{c'k'}\}$ of the other classes $\forall c' \neq c$ and $k' = 1, ..., K_{c'}$, using linear-kernel support vector machine (SVM). The within-subcategory metric utilizes the training data of a certain subcategory S_{ck} only and describes how well this subcategory represents the test data x. We first obtain an approximated x'_{ck} by averaging the M-nearest neighbors of x from S_{ck}. Next, the center of S_{ck}, i.e. $f_{ck} \in \mathbb{R}^H$, is derived based on the support vector data description (SVDD) [6]. The Euclidean distance between x'_{ck} and f_{ck} then describes the representativeness of x by the subcategory S_{ck}.

Subcategory fusion. We define the probability of test data x belonging to class c as:

$$P(x,c) = P_m(x,c) + \sum_{k=1}^{K_c} w_{ck} q_{ck} \{\beta P_b(x, S_{ck}) + (1-\beta) P_w(x, S_{ck})\} \quad (2)$$

The first term $P_m(x,c)$ is the probability of x belonging to class c obtained using multiclass linear-kernel SVM. The second term is the class-level probability fused from the subcategory-level probabilities. The weight factor w represents the contribution level of subcategory S_{ck}, and is obtained by finding a sparse representation of x from the various subcategories in an LLC construct. The weight factor q_{ck} quantifies the cluster quality, and is computed based on the Dunn index.

Experimental results. For texture descriptors, we use the improved feature vector (IFV) [5] and convolutional architecture for fast feature embedding (Caffe) [4]. The combination of IFV and deep convolutional network activation features (DeCAF) [2] has shown excellent texture classification accuracy in the state-of-the-art [1], achieving about 9% improvement over the previous best result. Experiments are conducted on three challenging datasets: the KTH-TIPS2 database, Flickr Material Database (FMD), and the Describable Textures Dataset (DTD). We obtained better classification performance than the state-of-the-art [1], as shown in Table 1.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

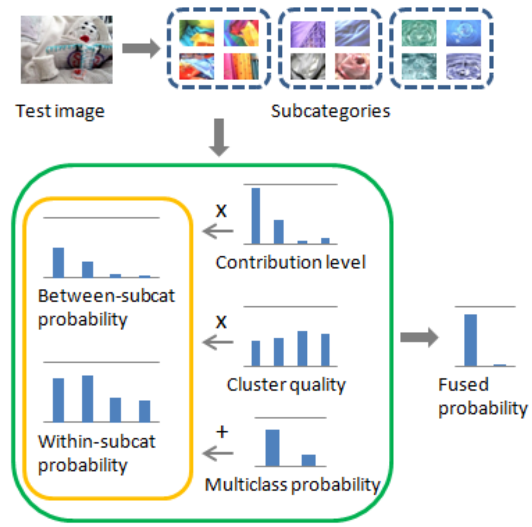

Figure 1: Overview of our proposed method. During testing, the between- and within-subcategory probabilities are computed, then fused based on the contribution levels, cluster qualities and multiclass probability to classify the test image. During training, the training images of each class are subcategorized, subcategory-level models exploring between-subcategory distinctiveness and within-subcategory representativeness are built, the cluster qualities are computed, and multiclass SVM is trained.

Table 1: The classification accuracies (%) compared to the state-of-the-art.

Dataset	IFV+Caffe		State-of-the-art
	SVM	Ours	
KTH-TIPS2	75.4±3.0	**79.3±2.7**	76.0±2.9 [1]
FMD	65.2±1.2	**68.4±1.5**	65.6±1.4 [1]
DTD	65.1±1.4	**67.8±1.6**	64.7±1.7 [1]

[1] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi. Describing textures in the wild. *CVPR*, pages 3606–3613, 2014.

[2] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: a deep convolutional activation feature for generic visual recognition. *arXiv preprint arXiv:1310.1531*, 2013.

[3] E. Elhamifar and R. Vidal. Sparse subspace clustering. *CVPR*, pages 2790–2797, 2009.

[4] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*, 2014.

[5] F. Perronnin, J. Sanchez, and T. Mensink. Improving the fisher kernel for large-scale image classification. *ECCV*, pages 143–156, 2010.

[6] D. M. J. Tax and R. P. W. Duin. Support vector data description. *Machine Learning*, 54(1):45–66, 2004.

[7] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong. Locality-constrained linear coding for image classification. *CVPR*, pages 3360–3367, 2010.