Fuzzy Encoding For Image Classification Using Gustafson-Kessel Aglorithm

Ashish Gupta, Richard Bowden

Centre for Vision, Speech, and Signal Processing, University of Surrey, Guildford, United Kingdom

UNIVERSITY OF SURREY

Abstract

Feature vectors in visual descriptor vector space do not exist in naturally occurring clusters, but have been shown to exhibit visual ambiguity. 'Bag-of-Features' (BoF), the most popular approach, which utilizes hard partitioning, ignores the existence of a semantic continuum in this space. Kernel Codebooks have been demonstrated to improve upon BoF, by soft-partitioning descriptor space. Building on these results, this paper, to model the visual ambiguity, formulates feature encoding, as clustering with fuzzy logic. The Gustafson-Kessel algorithm, computes hyper-ellipsoid shaped clusters, each with learnt mean and co-variance. The approach is demonstrated to provide better classification performance than BoF for several popular data-sets.
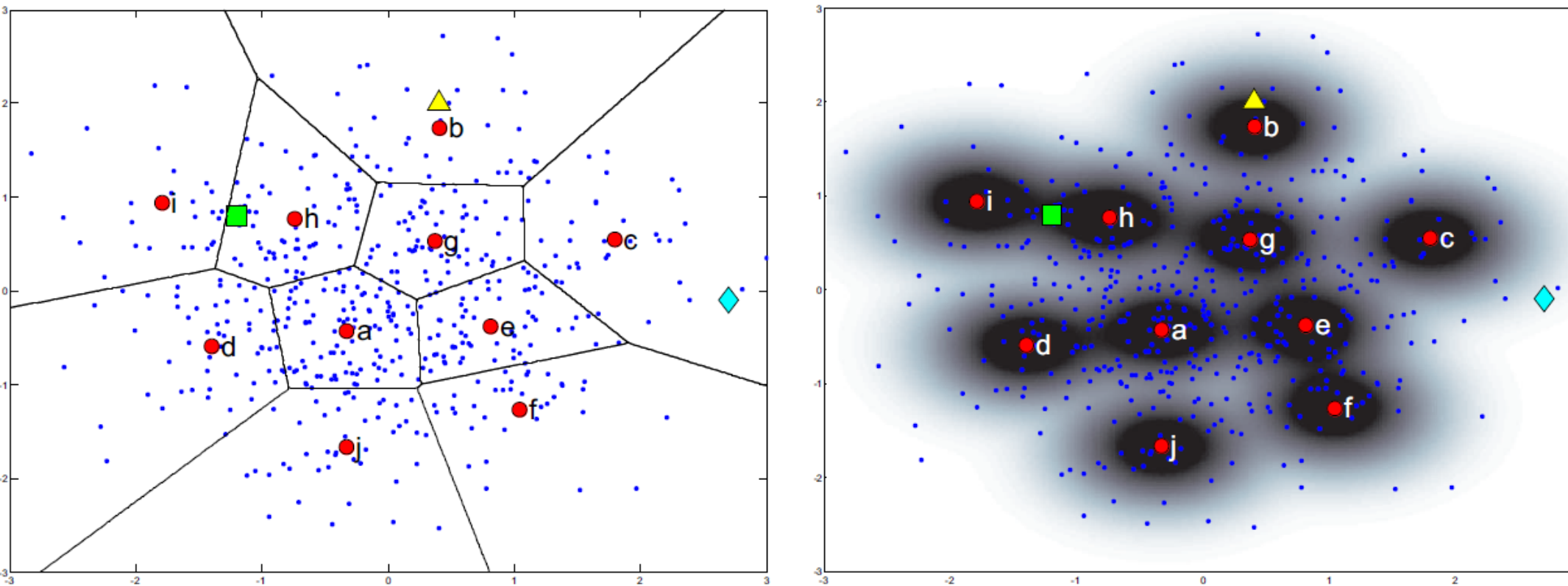
Visual Word Ambiguity

Figure: Hard partitioning leads to: suitable (triangle); uncertain (square); and plausible (diamond) types of assignments. Image from [1]
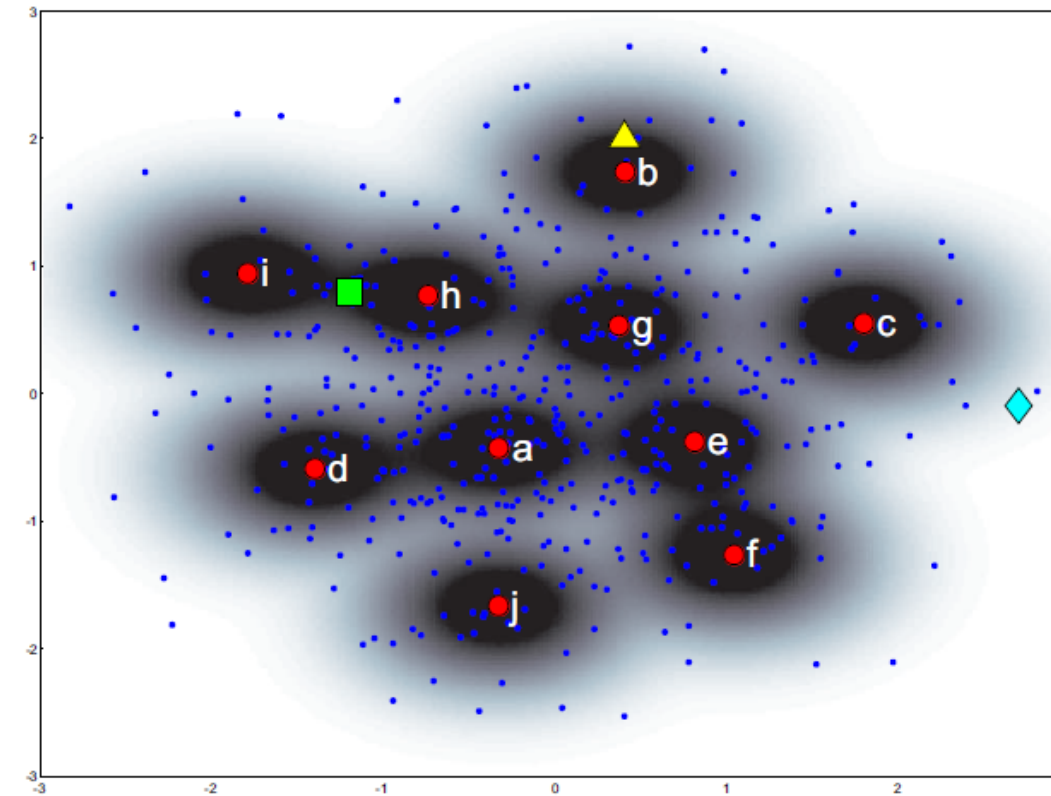
Figure: Kernel Codebook ameliorates issues with uncertainty and plausibility by weighted assignment to words. Image from [1]

The feature descriptor space is a continuum, which dictates that the exclusive assignment of a descriptor to a cluster prototype - crisp logic - can be modelled better by soft-assignment of descriptor to multiple cluster prototypes [1].

Image Classification Pipeline
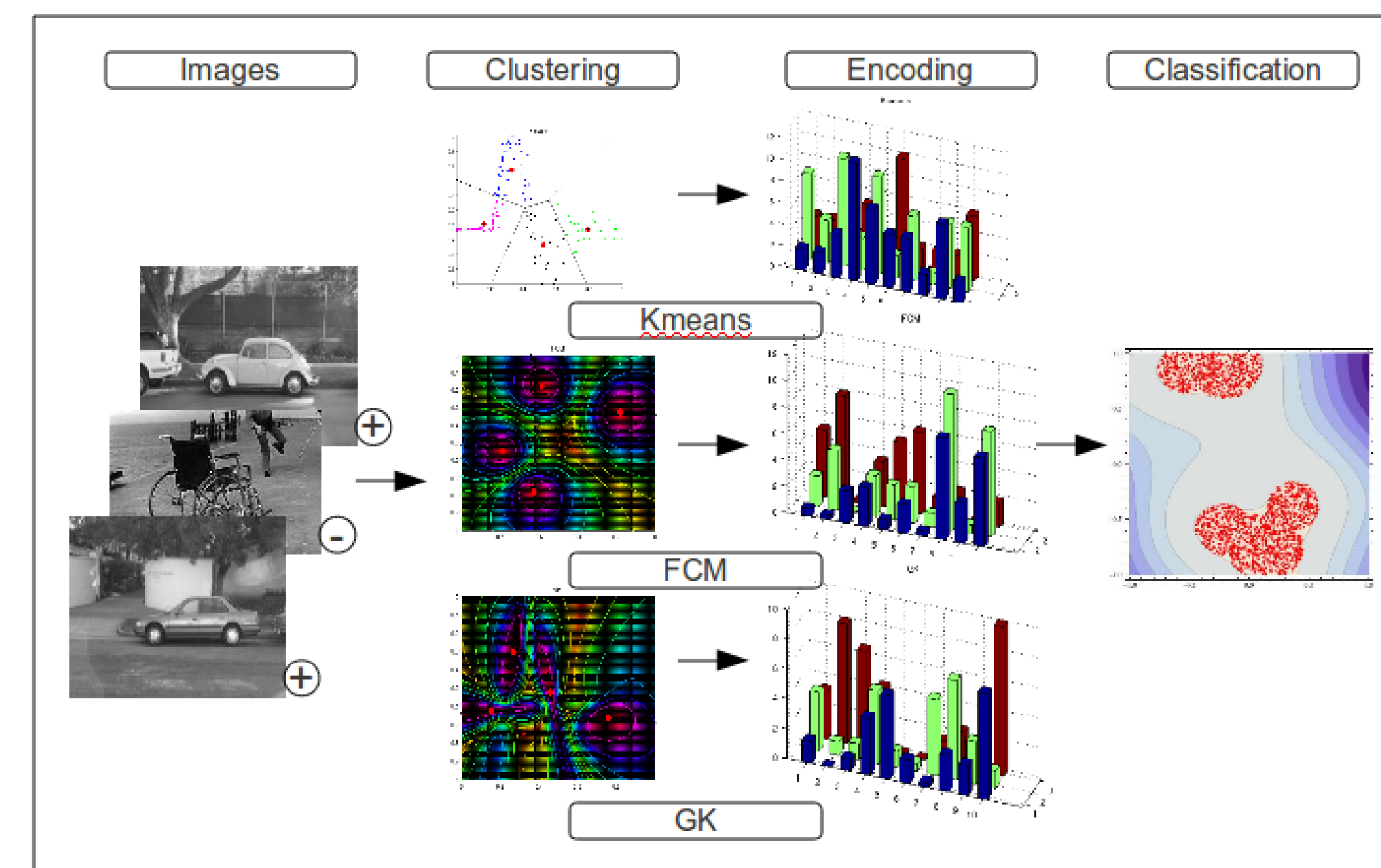
The training data-set, for visual category 'car', consists of positive and negative labelled images. A sample of descriptors computed on these images is clustered using either K-Means (BoF), Fuzzy C-Means (FCM) [2], or Gustafson-Kessel (GK) [3]. The descriptors from each image are encoded according to weighted word assignment(s) to compute histogram feature vectors for each image. A classifier is trained using these feature vectors.

Fuzzy Encoding: Methods
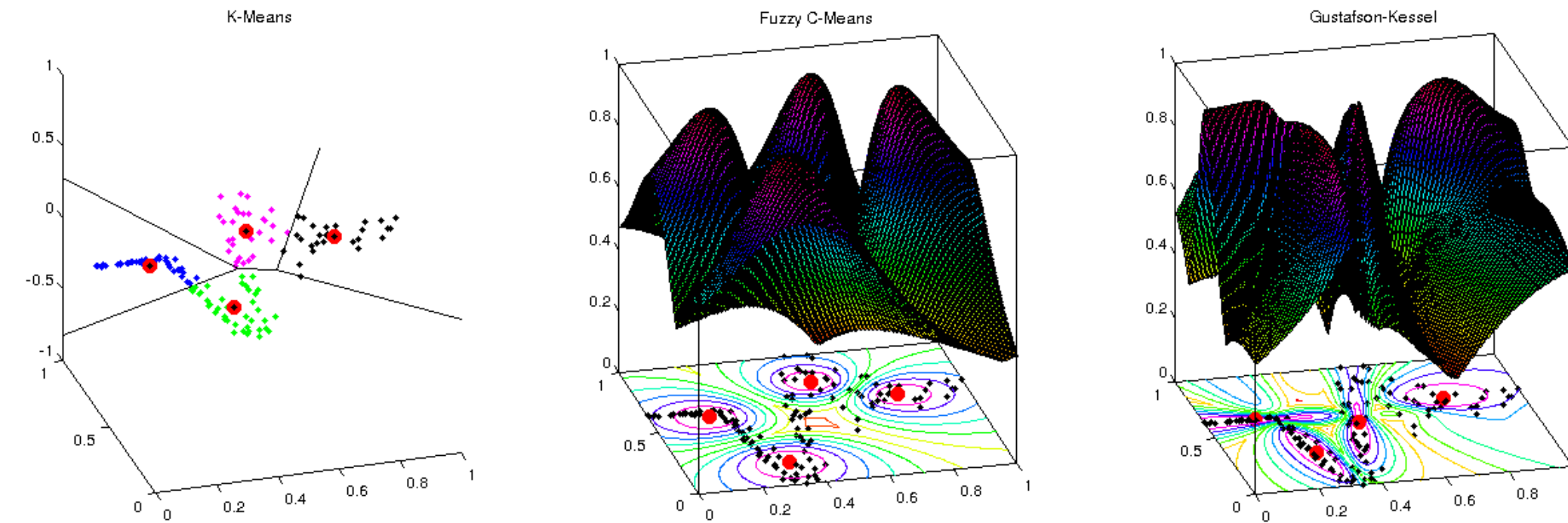
K-means computes words $V = \{v_1, v_2, \ldots, v_K\}$ and assignment of descriptors $X = \{x_1, x_2, \ldots, x_N\}$ to words by minimizing:

$$J(X; V) = \sum_{i=1}^{K} \sum_{k=1}^{N} \mathbb{1}_k^i \parallel x_k - v_i \parallel^2 \;, \; \mathbb{1}_k^i = \begin{cases} 1 & \text{if } x_k \in v_i \\ 0 & \text{otherwise} \end{cases}$$

The encoded feature for an image ν is a discrete valued histogram. The FCM algorithm attempts to minimize :

$$J(X; U, V) = \sum_{i=1}^{K} \sum_{k=1}^{N} \mu_{ik}^m \parallel x_k - v_i \parallel^2 \;, \; 1 \leq m < \infty$$

where $U = [\mu_{ik}]$, and μ_{ik} is degree of membership of x_k to cluster i. m is a measure of 'fuzzification'. ν reflects a degree of membership of descriptor to words. However, FCM in unable to adapt to local distribution of descriptors. The GK algorithm extends FCM, using a metric induced by a positive definite matrix A, learning hyper-ellipsoidal clusters instead of hyper-spherical clusters of FCM. The objective function minimized in GK is:

$$J(X; U, V, A) = \sum_{i=1}^{K} \sum_{k=1}^{N} \mu_{ik}^m D_{ikA_i}^2$$

where,

$$D_{ikA_i}^2 = (x_k - v_i)^T A_i (x_k - v_i) \;, \; 1 \leq i \leq K \;, \; 1 \leq k \leq N$$

Gustafson-Kessel Algorithm

Algorithm 1 Gustafson-Kessel

$\tau \leftarrow 1$
repeat
$v_i^{(\tau)} \leftarrow \frac{\sum_{k=1}^{N} (\mu_{ik}^{(\tau-1)})^m x_k}{\sum_{k=1}^{N} (\mu_{ik}^{(\tau-1)})^m}$
$F_i \leftarrow \frac{\sum_{k=1}^{N} (\mu_{ik}^{(\tau-1)})^m (x_k - v_i^{(\tau)})(x_k - v_i^{(\tau)})^T}{\sum_{k=1}^{N} (\mu_{ik}^{\tau-1})^m}$
$D_{ikA_i}^2 = (x_k - v_i^{(\tau)})^T [\rho_i det(F_i)^{\frac{1}{n}} F_i^{-1}](x_k - v_i^{(\tau)})$
$\phi_k \leftarrow \{i \mid D_{ik} = 0\}$
for $k \leftarrow 1, N$ **do**
  if $\phi_k = \emptyset$ **then**
    $\mu_{ik}^{(\tau)} \leftarrow (\sum_{j=1}^{K} (\frac{D_{ikA_i}}{D_{jkA_j}})^{\frac{2}{m-1}})^{-1}$
  else
    $\mu_{ik}^{(\tau)} \leftarrow \begin{cases} 0 & \text{if } D_{ikA_i} > 0 \\ \frac{1}{|\phi_k|} & \text{if } D_{ikA_i} = 0 \end{cases}$
$\tau \leftarrow \tau + 1$
until $\parallel U^{(\tau)} - U^{(\tau-1)} \parallel < \epsilon$
for $j \leftarrow 1, M$ **do**
$\nu_j \leftarrow \sum \mu \mathbb{1}_{jk}$
$\gamma_j \leftarrow \begin{cases} 1 & \text{if } \mathcal{I}_j \in \mathcal{C} \\ -1 & \text{if } \mathcal{I}_j \notin \mathcal{C} \end{cases}$

Notation:
v_i: centre of ith cluster
F_i: co-variance of ith cluster
x_k: k^{th} descriptor
μ_{ik}: membership of x_k to i^{th} cluster
D_{ikA_i}: inner product norm
ρ_i: $det(A_i)$
m: measure of fuzzification
M: no. of images
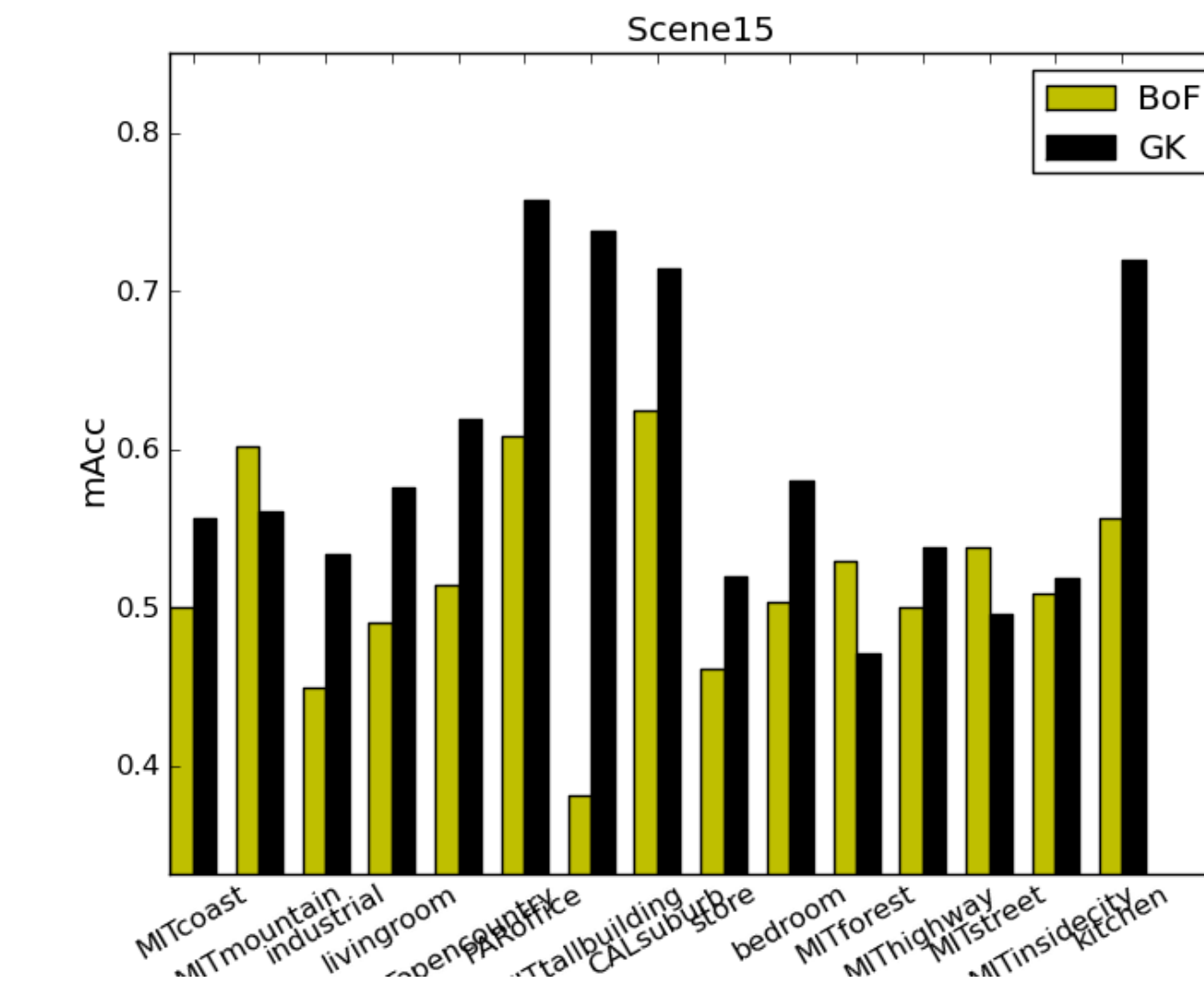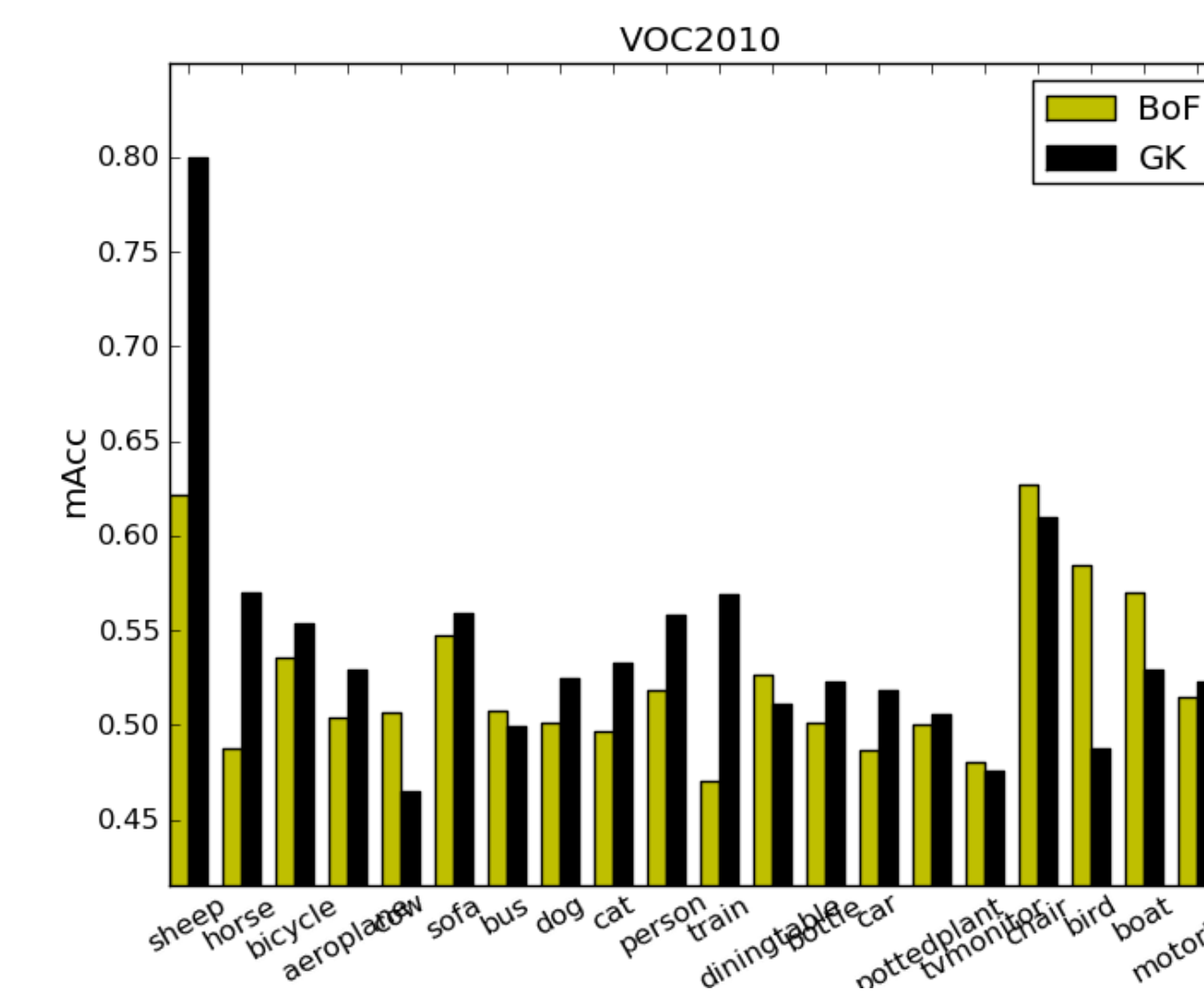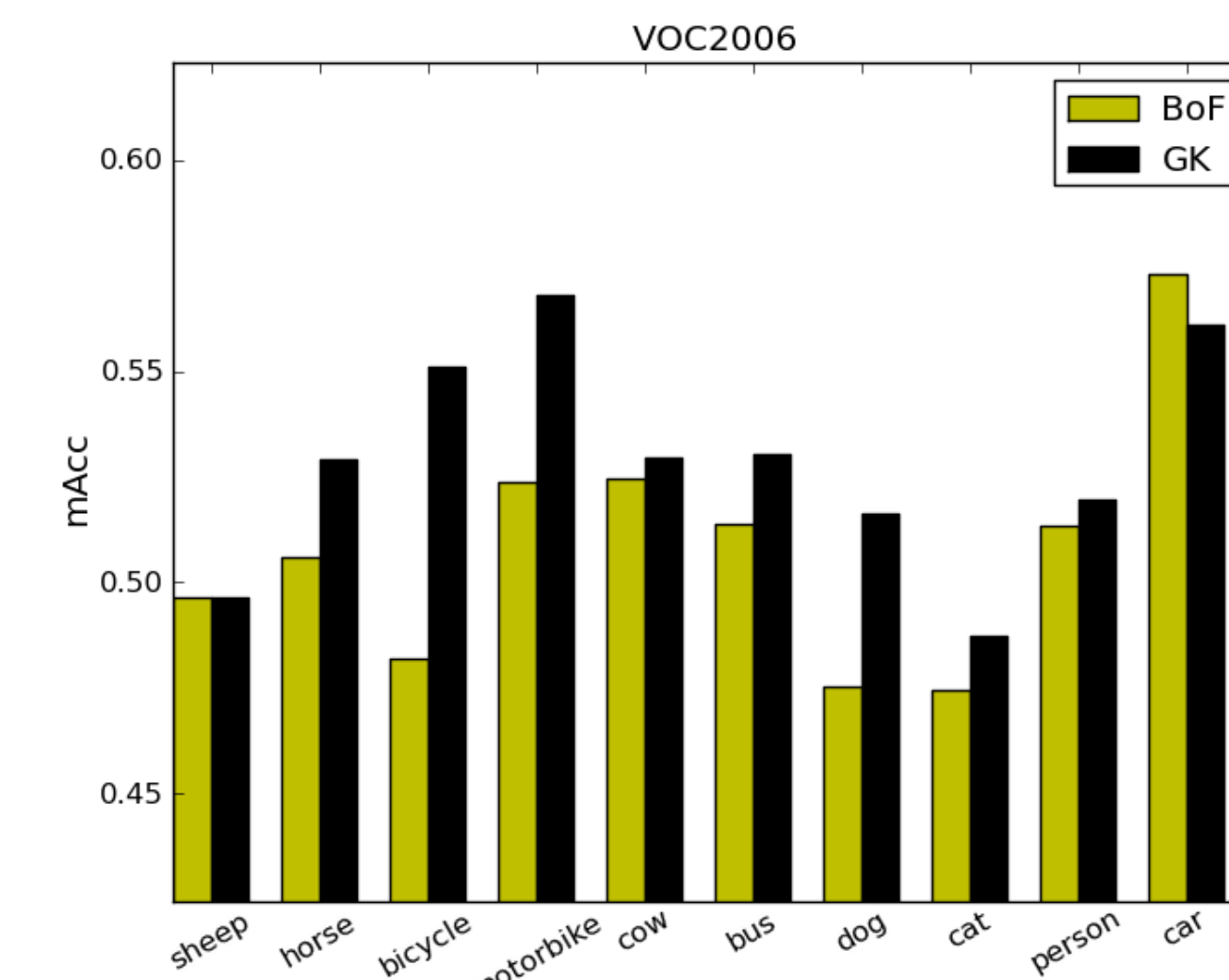$\mathcal{C}$: category
$\mathcal{I}_j$: j^{th} image

Experiments

The comparative classification performance of BoF, FCM, and GK algorithms is analysed across visual categories in a data-set; across set of data-sets; and for different codebook size. The feature descriptor utilized in all experiments is the popular local affine co-variant descriptor SIFT. A classifier is SVM with RBF kernel. The datasets utilized are Caltech-101, Caltech-256, Pascal VOC 2006, Pascal VOC 2010, and Scene-15. These datasets vary in terms of number of categories, number of images within each category, visual domain of categories, inherent difficulty in modelling a category.
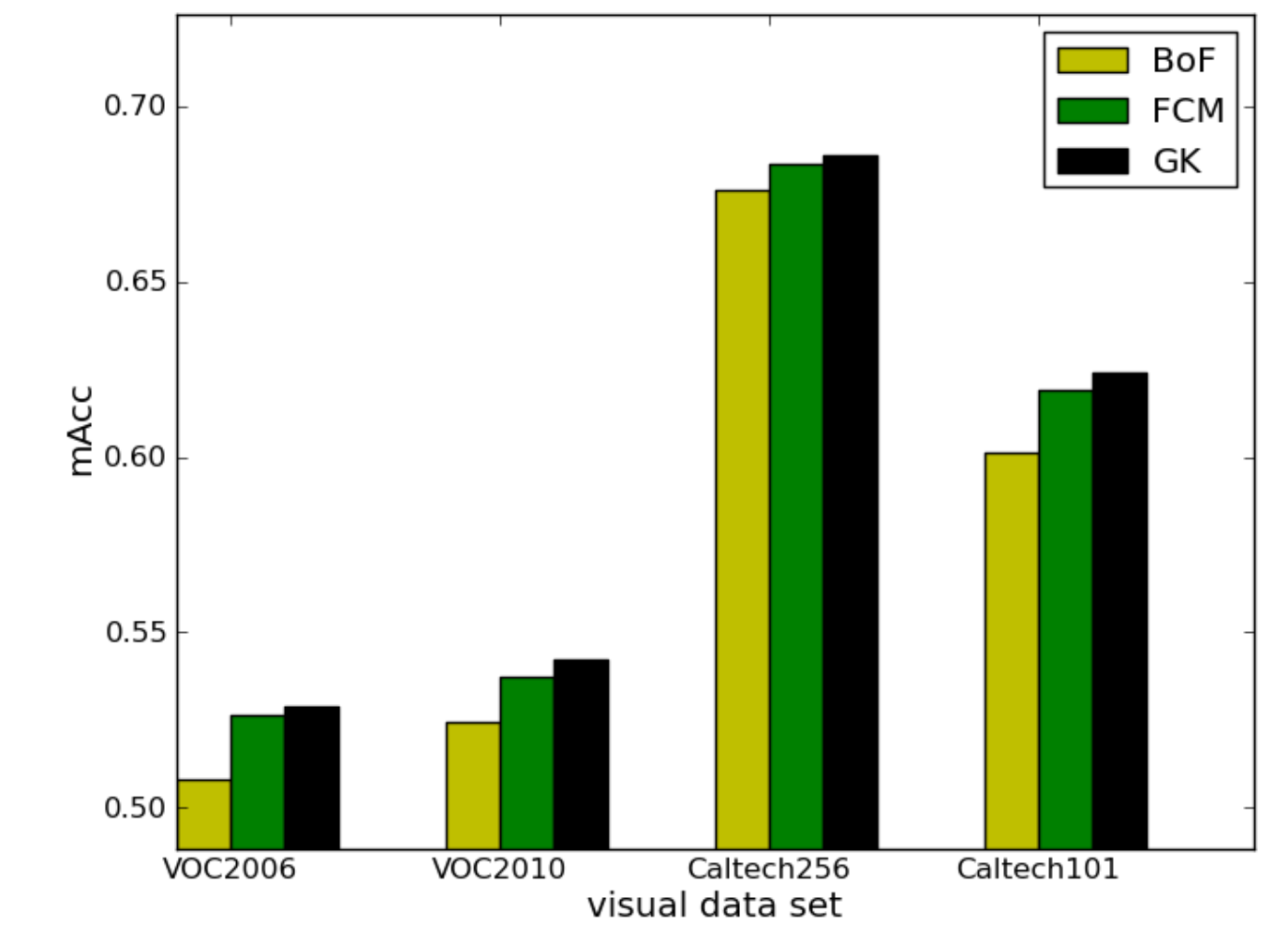
Performance across categories

The graphs in the figures show the comparative mean classification accuracy of BoF and GK approaches for each visual category in the datasets: VOC2006, VOC2010, and Scene15. The absolute and relative performance of both BoF and GK varies across the categories due to the variation in content and complexity of each category.

Performance across datasets

Comparison of the BoF, FCM, and GK approaches in terms of their classification performance for different datasets. The results in the figure show the mean accuracy averaged across all categories of the dataset.

Performance across codebook sizes

Analysis of comparative mean classification accuracy of BoF and GK approaches for different codebook sizes. Graph shows performance for Caltech101 dataset.
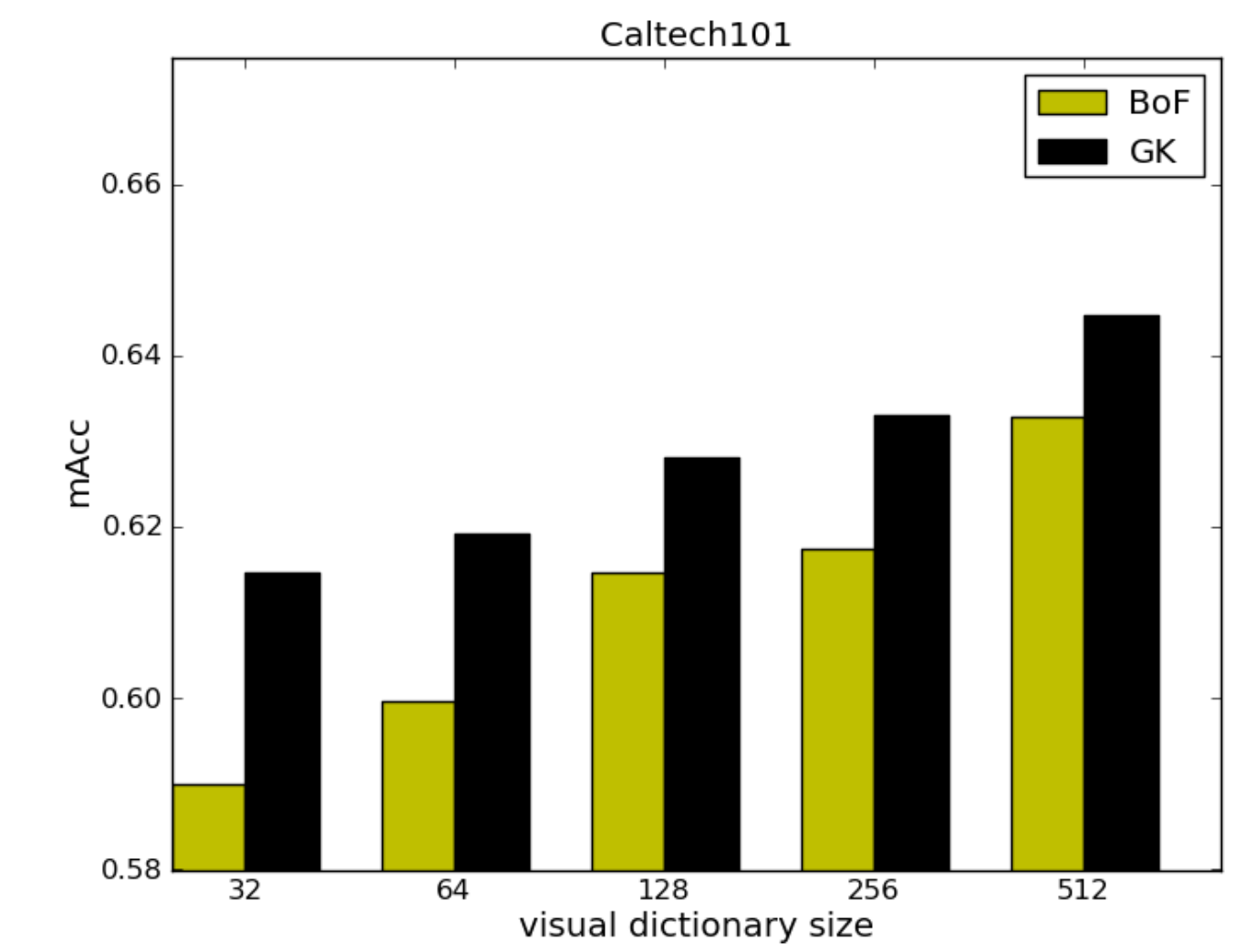
Summary

We have introduced fuzzy encoding technique for image classification using Fuzzy C-Means (FCM) to compute a fuzzy membership function. We extended this work to the Gustafson-Kessel (GK) fuzzy clustering algorithm, which was shown to adapt to local distributions. We demonstrated empirically that our fuzzy encoding approach is consistently better than the BoF model, using several popular datasets. GK algorithm was shown to provide a marginal improvement over FCM, which we expect to improve with optimization of covariance matrices in future.

References

[1] J.C. van Gemert, C.J. Veenman, A.W.M. Smeulders, and J.-M. Geusebroek, "Visual word ambiguity," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 32, no. 7, pp. 1271–1283, july 2010.

[2] A. Baraldi and P. Blonda, "A survey of fuzzy clustering algorithms for pattern recognition. i," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, vol. 29, no. 6, pp. 778–785, dec 1999.

[3] Donald E. Gustafson and William C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," in Decision and Control including the 17th Symposium on Adaptive Processes, 1978 IEEE Conference on, jan. 1978, vol. 17, pp. 761–766.

Acknowledgement