Abstract

Recent work in object categorization often uses local image descriptors such as SIFT to learn and detect object categories. Such descriptors explicitly code local appearance and have shown impressive results on objects with sufficient local appearance statistics. However, many important object classes such as tools, cups and other man-made artifacts seem to require features that capture the respective shape and geometric layout of those object classes. Therefore this paper compares, on a novel data collection of 10 geometric object classes, various shape-based features with appearance-based descriptors such as SIFT. The analysis includes a direct comparison of feature statistics as well as results within standard recognition frameworks, which are partly intuitive, but sometimes surprising.

1. Introduction

Historically, the recognition of geometric objects such as cups and tables has been an important focus of object recognition [24]. In recent work however, the recognition of geometric objects is largely underrepresented with some notable exceptions [6, 25]. Many recent and successful object recognition and categorization approaches are based on local appearance features [5, 4, 17]. These approaches tend to perform well when enough local appearance information can be found, as, e.g., for cars and motorbikes. Incorporating spatial information about feature distributions often proves helpful, even though surprisingly good results have been reported for simple bag-of-words approaches that neglect feature location altogether.

The main question we address in this paper is how these local feature approaches transfer to the recognition of more geometric objects. We therefore compare the statistics of various successful appearance features (such as SIFT [14] and GLOH [20]) with features that are more geometric in nature (such as Geometric Blur (GB) [3] and Shape Context (SC) [1]). We also include the recently proposed k-Adjacent Segments (k-AS) [6] which have been designed to explicitly code information about the geometric layout of an object. The evaluation is performed on two complementary data sets: a new data set containing over 700 images of 10 geometric object classes, and subsets of Caltech-101 [13] containing object classes with important local appearance statistics. One of the surprising results of this evaluation is that performance differences between features on these two rather distinct data sets are less pronounced than one might expect.

Secondly, we analyze the performance of these features in a general object classification setting based on local features, again using the two distinct databases of object classes. As a first baseline we use a Naïve Bayes classifier, where individual features contribute independently to the classification result. As spatial information might be particularly important for the recognition of geometric objects, we use a second baseline, the so called localized bag-of-words representation. It allows to gradually add location information to the object representation, and quantify the contribution of location to classification accuracy.

The results of the first baseline experiments are mostly consistent with what is suggested by the feature statistics evaluation. The introduction of (weak) spatial information however results in a more significant performance boost than caused by differences between individual features alone. This performance boost notably differs between features. Without any spatial information, SIFT and GLOH in combination with the Hessian-Laplace interest point detector perform best on average. When spatial information is added, geometric features, namely Geometric Blur and k-Adjacent Segments, can outperform all other features.

The first contribution of this paper is the introduction of a novel data set of 10 geometric object classes. The second is the evaluation of different appearance as well as geometric features on two distinct data sets. Third, we compare two clustering schemes which have been used in the literature by some of the most successful recognition approaches. Fourth, we give results for two baseline methods for local feature-based recognition with and without spatial information.

2. Related Work

Local appearance-based features have received a lot of attention in the literature. Many general feature evaluations exist, typically focusing on criteria based on corre-
Correspondence matching [20, 21]. Including the notable exception of [18], comparably little work has been done on feature evaluation in the context of object class recognition.

Similarly, the explicit treatment of shape-based local features is non-sufficient. The Shape Context descriptor has been studied in the context of pedestrian detection [27], [22] compares its performance to SIFT, PCA-SIFT [10], Differential Invariants [26] and Steerable Filters in a 3D matching framework. [12] builds upon global image representations, and compares the performance of texture- and contour-cues in a multi-class classification task. Apart from [6], we are not aware of any evaluation including k-Adjacent Segments.

We build upon the work of [18], and base our evaluation on a mid-level clustering representation used by many local feature approaches in object recognition. In fact, we reproduce their major findings by evaluating the best performing features on a subset of Caltech-101. Our evaluation goes beyond their work w.r.t. the following aspects: 1) our evaluation explicitly includes shape-based features such as Geometric Blur, Shape Context and k-Adjacent Segments; 2) we do experiments over two complementary data sets of shape- and appearance-based classes (notably introducing the shape-based data-set), and 3) we report results over two general baseline recognition frameworks with and without spatial information.

3. Data Sets

We shortly present the data sets used in our experiments.

**Shape.** We introduce a novel collection of 724 images showing single objects of 10 geometric object classes, i.e., objects for which shape and geometric layout of object parts determine class affiliation rather than local appearance. Figures 1 (a) and (b) show two examples of each of the classes cap, fork, hammer, knife, mug, pan, pliers, pot, saucepan, and scissors. All images are roughly aligned w.r.t. position and viewpoint. The data set exhibits high intra-class variability, and is challenging for recognition. In order to provide a platform for best-case evaluations, objects are recorded in front of a clean, white background. For experiments, we randomly pick 20 images from each class as a training set, and 10 images of each class for testing.

**Shape2.** Shape2 contains 100 additional images to provide more realistic test data for the above classes (see Figure 1 (c)). While the images contain a single object each the background and image quality vary greatly. Both Shape and Shape2 data sets can be downloaded from our web-page1.

**Caltech-101.** We primarily use a 10 class subset of Caltech-101 [13] that we find is characterized by local appearance statistics, namely accordion, crab, cannon, electric guitar, euphonium, gramophone, inline skate, revolver, watch, and windsor chair (see Figure 1 (d)). We further report results for a larger pool of object classes provided by Caltech-101, but restrict ourselves to two random subsets of 20 and 40 classes, respectively, for computational reasons.

4. Local Features

We briefly introduce the features and interest point detectors used in our comparison. The shape related features are k-Adjacent Segments [6], Geometric Blur [3], and Shape Context [1]; the appearance-based region descriptors are SIFT [14] and GLOH [20]. As interest point detectors, we employ Harris-Laplace [19], Hessian-Laplace [21], and Salient Regions [9].

4.1. k-Adjacent Segments (k-AS)

k-Adjacent Segments have been proposed as an extension to contour segment networks, a graph-based method for template-matching hand-drawings to image databases [7]. [6] demonstrates how k-AS can be incorporated into a general object recognition framework.

We extract k-AS features from an image using the original implementation2. First, edgels are detected via the Berkeley natural boundary detector [15]. Second, neighboring edgels are chained, and further linked to form L-, T- and higher-order junctions. Last, edgel-chains are replaced by straight line approximations (contour segments), and joined into a global contour segment network for the image.

A k-AS descriptor \( d \) then describes the geometric layout of a group of \( k \) adjacent segments in that network. For each such group, one segment is picked as reference, and the layout of the others described relative to that reference segment. In particular, the descriptor encodes relative segment positions \( p_i \), orientations \( o_i \) and lengths \( l_i \). Let segment \( k \) be the reference segment, then the descriptor is

\[
\begin{align*}
    d &= (p_2/N, \ldots, p_k/N, o_1, \ldots, o_k, l_1/N, \ldots, l_k/N),
\end{align*}
\]

where \( N \) is a normalization factor that renders \( d \) invariant to scale. We make descriptors invariant to in-plane rotation by rotating them around \((0,0)\) by \( o_1 \), and excluding \( o_1 \) from the descriptor. The dimensionality of k-AS features is \( n = 4 \times k - 2 \). For the typical choices of \( k \in \{2, 3\} \), \( n \) is 6 or 10, rendering k-AS a comparably low dimensional descriptor.

k-AS features differ from others considered in this paper in several ways. Notably, they do not match the standard scheme of describing local image regions around interest points. Instead, they utilize shape information from the whole image, and thus have the potential to capture the characteristic geometric layout of an object at the cost of sacrificing local appearance information. Although k-AS mostly represent local groups of contour segments, we also observed that descriptors related distant contour segments on opposing object boundaries.

---

1http://www.mis.informatik.tu-darmstadt.de

2http://www.vision.ee.ethz.ch/˜ferrari/release-kas.tgz
4.2. Local Region Descriptors

We briefly present the local region descriptors used in our experiments. Local region descriptors encode information about image patches centered at interest points.

**Geometric Blur (GB).** We use the original implementation of the Geometric Blur [3] region descriptor from [2]. Geometric Blur first extracts \( c = 4 \) channels of oriented edge energy [23] to obtain a sparse signal \( S \). In \( S \), the region centered at interest point location \( x_0 \) is blurred with a spatially-varying Gaussian kernel \( G_d \) to obtain the Geometric Blur \( B_{x_0}(x) = S \ast G_{\alpha x + \beta}(x_0 - x) \). \( B_{x_0}(x) \) is then sub-sampled over all channels at \( n \) distinct locations in a circular grid. The final descriptor is the concatenation of all \( c \times n \) samples. Throughout all experiments, we use the standard values for \( \alpha = 0.5 \), \( \beta = 1 \) and \( n = 51 \), resulting in a descriptor of length 204.

**Shape Context (SC).** Shape Context [1] is originally based on edge information. For a given interest point location, it accumulates the relative locations of nearby edge points in a coarse log-polar histogram. We compute a histogram containing 9 spatial bins over 4 edge orientation channels. Bin size increases w.r.t. distance from the interest point center. Note that this is similar in spirit to spatially varying blur, but results in a smaller descriptor (length 36).

**SIFT.** The Scale Invariant Feature Transform [14] descriptor is a 3D histogram over local gradient locations and orientations, weighted by gradient magnitude. It uses \( 4 \times 4 \) location and 8 orientation bins, i.e., 128 in total.

**GLOH.** Gradient Location Orientation Histograms [20] is an extension of the SIFT descriptor. It uses 17 bins for location and 16 bins for orientation in a histogram over a log-polar location grid, and reduces descriptor dimensionality to 128 by PCA. We use the implementation of [20] for SC, SIFT and GLOH. All descriptors are made invariant to in-plane rotation by aligning the region to the dominant gradient direction before descriptor computation.

4.3. Interest Point Detectors

We compute local region descriptors based on detections of the following interest point detectors: Harris-Laplace (HarLap) [21] is an extension to Harris corners [8]. It selects corners at locations where a Laplacian attains an extremum in scale-space. The Hessian-Laplace (HesLap) [21] detector responds to blob-like structures. It searches for local maxima of the Hessian determinant, and selects a characteristic scale via the Laplacian as for Harris-Laplace. The Salient Regions (SalReg) detector [9] identifies local image regions that are non-predictable across scales by measuring entropy over local intensity histograms. We use publicly available implementations for HarLap/HesLap\(^4\), and SalReg\(^6\).

5. Feature Evaluation

We evaluate the combined performance of feature detectors and descriptors at three different levels. First, we compute statistics over clusterings of local feature descriptors (codebooks), using two different clustering techniques. Second, we represent objects by means of occurrence statistics over codebook matches, and analyze classification performance in a Bayesian framework. Third, we investigate the impact of gradually adding location information.

\(^3\)http://www.cs.berkeley.edu/~aberg/demos/gb_demo.tar.gz

\(^4\)http://www.robots.ox.ac.uk/~vgg/research/affine/descriptors.html

\(^5\)http://www.robots.ox.ac.uk/~vgg/research/affine/detectors.html

\(^6\)http://www.robots.ox.ac.uk/~timork/salscale.html

![Figure 1. Example images from the Shape ((a), (b)), Shape2 (c), and Caltech-101 (d) data sets.](image-url)
to that object representation, by jointly boosting localized histograms of codebook matches over all object categories.

5.1. Cluster Statistics

We follow the argumentation of [18] and base our evaluation on a mid-level representation of image features common to many computer vision techniques. In particular, we analyze the statistics of clusterings of feature descriptors.

A clustering over a set of feature descriptors of a given type is determined by 1) the choice of clustering algorithm, and 2) the choice of a (dis-) similarity measure. We use K-Means as a widely accepted method for 1), and add Reciprocal Nearest Neighbor (RNN) clustering for comparability with [18]. RNN is a centroid-based implementation of Hierarchical Agglomerative clustering [11]. For 2), we consistently use Euclidean distance. We are conscious that various clustering techniques and (dis-) similarity measures have been proposed tailored towards specific feature types. We resort to standard ones for the sake of comparability.

**Cluster precision.** In order to quantify how well a clustering of feature descriptors reflects the separation of object classes, we introduce a refinement to cluster precision [18]. Intuitively, we want to measure to what extent features of a given class $a$ are grouped together by clustering. Original cluster precision therefore memorizes, for each cluster $j$ in which class $a$ dominates, the fraction $p_{ja}$ of features of class $a$, and averages these fractions by the number of clusters $M$ dominated by $a$, i.e., $P_{Ca} = \frac{1}{M} \sum_{j=1}^{M} p_{ja}$. In our experiments, we found that many clusters with high scores according to $p_{ja}$ often contained features from only a single object instance. Because we want to give higher scores to feature descriptors that generalize across multiple instances of an object class, we discount such clusters by summing over the fractions of *objects* of class $a$ in cluster $j$ instead of individual features, and weight these fractions by cluster sizes. We obtain

$$P'_{Ca} = \left( \sum_{j=1}^{N} s_j \right)^{-1} \sum_{j=1}^{N} s_j p_{ja},$$

where $j$ now ranges over all $N$ clusters in which objects of class $a$ dominate, and $s_j$ is the total number of features in cluster $j$. High scores (including weights) are obviously obtained by big clusters with features from many instances of a single object class, and low scores by small clusters with few features, but from multiple classes. The combined score for a descriptor is the average over $P'_{Ca}$ over all classes.

**Matching precision.** In a second experiment, we evaluate the generalization capabilities of a given feature type over an unseen test set. Features are extracted from an independent set of images, and matched against a codebook built from training images. For each test feature belonging to class $a$, we determine the precision $P_{Ca}$ over all *matched* clusters with respect to class $a$. The matching precision is the average precision over all classes.

5.2. Naïve Bayes

The second level of our evaluation represents objects in terms of occurrence statistics (counts over nearest-neighbor matches) over codebook entries and trains a multi-class-classifier on a training set of such representations. We use an analogous approach to *Multinomial Naïve Bayes* [16] for text classification, and model the posterior distribution of an object class, given occurrence statistics over a codebook, as a multinomial distribution. Let $N_{ij}$ denote the number of occurrences of a feature of class $c_j$ that matches codebook entry $w_i$. We estimate the likelihood that this codebook entry originates from class $c_j$ as

$$P(w_i|c_j) = \frac{1 + N_{ij}}{W + \sum_{c=1}^{C} N_{ic}},$$

for $C$ different object classes and a codebook of size $W$, and assuming a simple Laplacian prior.

5.3. Localized Bag-of-Words

While [18] relates the localization properties of a given feature type to entropy over location distributions, we directly measure the impact of adding location information in terms of classification accuracy in a Joint Boosting [28] framework.

The object representation on the third level of our evaluation is based on soft-matched histograms of feature occurrences over a codebook, and inspired by [7]. We divide a rectangular image region into a grid of cells. For each cell, extracted features are matched to a codebook, and a local histogram over soft-matched codebook entries is computed (the inverse distances between feature and all cluster centroids are used). The representation of an object is the concatenation of all local cell histograms. A Joint Boosting algorithm is trained from object representations of a set of training images using a fixed number of boosting rounds, and tested against an independent set of test images. We use *decision stumps* [28] over histogram bins as weak classifiers for boosting. By varying the number of grid cells $g$, we regulate the tradeoff between rich feature statistics (small $g$) and more accurate localization (large $g$).

6. Experimental Results

In the following, we present the results of our experimental evaluation. Due to limited space, we choose to give plots for SIFT, GB, and $k$-AS, since their comparison is a key contribution of this paper. Additionally, we give curves for SC and GLOH descriptors that obtain highest scores with varying detectors as part of the SIFT plots. We also plot non-rotation invariant $k$-AS as a reference, and include the DoG detector proposed for SIFT [14], using the original implementation7. For comparability, all plots in this paper consistently show results for 10 classes of the respective

---

7http://www.cs.ubc.ca/~lowe/keypoints/
data set (Shape, Shape2, Caltech-101). For Caltech-101, we emphasize the major differences to 20 and 40 class subsets in the text, and note that the complete collection of plots is part of supplementary material.

**Cluster precision.** We measure cluster precision for 9 different compression ratios ($\frac{\text{# Features}}{\text{# Clusters}}$) ranging from 4 to 20 in steps of 2 over codebooks generated from 200 training images per data set. We observe that cluster precision changes substantially if we vary detectors for a given descriptor, while it remains relatively stable over varying detectors for a given detector. We give cluster precision plots in Figures 2 (a) to (f) over varying detectors (respectively varying $k$ for k-AS), and fix the descriptors of interest (SIFT, GB, k-AS).

We begin by presenting the results for Caltech-101 (see Figures 2 (a) to (c)), and first consider the appearance-based SIFT and GLOH descriptors with varying detectors. We observe that the ordering of detector performance is in fact consistent across all examined descriptors, including GB and SC. HesLap is best, followed by HarLap and SalReg. SIFT and GLOH descriptors perform equally well, and obtain high scores in our comparison. Both obtain highest scores for HesLap. These results are in line with the results reported in [18].

Surprisingly, these results transfer seamlessly to the Shape data set (see Figures 2 (d) to (f)). Still, appearance-based SIFT and GLOH obtain high scores, and the ordering of detectors remains the same as for Caltech-101. We stress that this stability across data sets is unexpected.

We now examine the performance of the shape-based features GB, SC, and k-AS. Over both data sets, the precision of GB is slightly higher than that of SIFT for HesLap and SalReg detectors (HesLap-GB is best), and about equal for HarLap. SC is best with HesLap, but slightly worse than SIFT over all detectors. k-AS and DoG-SIFT obtain lowest scores. The relative ordering of k-AS follows the intuition that discriminative power increases with $k$, and decreases with rotation invariance.

In summary, appearance- (SIFT, GLOH) and shape-based descriptors (GB, SC) do not show great differences in cluster precision. Both perform comparably well over both appearance-based (Caltech-101) and shape-based (Shape) data sets. k-AS are worst, but still comparable to DoG-SIFT. These results are fully mirrored on 20 and 40 class subsets of Caltech-101.

**Matching precision.** We measure matching precision over an independent test set of 100 images for both data sets, Shape and Caltech-101. We report that cluster precision results transfer to large extents to matching precision over previously unseen data, but omit the plots for brevity.

**Clustering algorithms.** Figures 2 (g) to (i) correspond to plots (d) to (f), but for Reciprocal Nearest Neighbor clustering. The relative ordering of detectors and descriptors is consistent with K-Means, with the notable exception of GB (Figure 2 (h)), where HesLap is inferior to HarLap, and SalReg dominates both for high compression ratios. On average, we obtain higher absolute precisions for K-Means clustering. Reciprocal Nearest Neighbor shows stronger tendencies to yield degenerate clusterings for high compression ratios, where a single large cluster attracts most of the features while leaving others unmatched (singleton). For low compression ratios often used in codebook-based approaches, this is not an issue.

### 6.1. Naïve Bayes

We measure multi-class-classification accuracy ($\frac{\text{# correct Predictions}}{\text{# total Predictions}}$) over fixed numbers of clusters from $n = 50$ to 1600, increasing by powers of 2. Dependent on the detector, $n = 1600$ corresponds to compression ratios of 9 (HarLap), 52 (HesLap), 21 (SalReg), 5 (DoG), 5 (2-AS) and 16 (3-AS) for the Shape data set. For each feature type and data set, we train a Naïve Bayes classifier on a training set of 200 images (20 per category), and test on an independent test set of 100 images (10 per category). Classifiers are trained on bag-of-word representations (hard-matching against a codebook of size $n$ built from the training set).

Again, we start with Caltech-101 (see Figures 2 (j) to (l)). The ordering of detectors over all descriptors is mainly consistent with the results for cluster precision: HesLap is typically best, closely followed by HarLap, and SalReg. Overall, the appearance-based feature combination HarLap-SIFT performs best, followed by HesLap-GLOH and HesLap-SC. While GB performs moderately on 10 object classes with HesLap, it performs generally worse than other descriptors for 20 and 40 classes. Rotation invariant k-AS are consistently worse. SC is competitive to SIFT with HesLap, in particular for 20 and 40 classes.

For the Shape data set (see Figures 3 (m) to (o)), the order of detectors is consistent with the results for Caltech-101; only SalReg gain relative performance in combination with the shape-based GB descriptor. Still, appearance-based HesLap-SIFT and HesLap-GLOH are best. Although GB is comparable with SalReg, it performs worse than SIFT with HarLap and HesLap. Notably, it tends to perform better for lower numbers of clusters, which can be explained by its high dimensionality, bearing the risk of over-fitting. SC is generally inferior to SIFT, but superior to rotation invariant k-AS. Rotation invariant 2-AS are worst. Rotation invariant 3-AS are slightly better than DoG-SIFT, but can not keep up with SIFT in general.

To summarize, the differentiation between descriptors with fixed detectors is more pronounced for Naïve Bayes than for cluster precision. In particular, appearance based features lead on average, over both data sets. GB and SC perform on a comparable level to SIFT and GLOH, but only
for individual detectors (SalReg for GB, HesLap for SC). k-AS exhibit relatively weak discriminative power for Naïve Bayes classification. SC offers a good compromise between strong (GB) and weak (k-AS) discrimination.

6.2. Localized Bag-of-Words

We measure classification accuracy as defined for Naïve Bayes for varying numbers of grid cells \( g \in \{1, 4, 9\} \), using the same training and test images. We assume known bounding boxes for training and test, and use them to anchor histogram grids. We fix the number of clusters to \( n = 200 \), and obtain histograms of length \( g \times n \in \{200, 800, 1800\} \).

Remarkably, for Caltech-101 (see Figures 3 (a) to (c)), the clear ordering of detectors that is present for cluster precision and Naïve Bayes tends to dissolve. For different descriptors, highest scores are achieved for different detectors. Appearance-based SIFT and GLOH perform equally well with HarLap respectively HesLap, and equal to SalReg-SC. Most remarkably, shape-based GB outperforms SIFT and GLOH with HesLap and SalReg, in particular for high values of \( g \). Adding location information boosts the performance of SalReg-GB by 20%. 3-AS gain 26%. Rotation invariant 3-AS perform equally well as HesLap-GLOH for \( g = 9 \). For 20 and 40 classes, k-AS loose performance relative to other features, which may be attributed to the weakness of the location model. The performance boost due to added location information decreases, but remains more important than the choice of descriptor, and at least as important as the choice of detector.

This tendency fully transfers to Shape (Figures 3 (d) to (f)). We observe again that shape-based features benefit to a large extent from location information: 45% boost for rotation invariant 2-AS, 36% for HarLap-GB, 35% for HesLap-SC, compared to 29% for appearance-based HesLap-SIFT. Further, the performance boost for shape-based features in response to increased location information even applies to the more challenging Shape2 data set (Figures 3 (g) to (i)), where we perform the transition from Shape’s best-case scenario to more realistic images. HesLap-GB gains 21%, HesLap-SIFT 7%.

While for Shape, boosting over localized bag-of-words lifts the discriminant power of all feature types to a comparable level for \( g = 9 \) (Figures 3 (d) to (f)), shape-based features win for Shape2: HesLap-GB and rotation invariant 2-AS are best (42% respectively 44% accuracy). The best performing SIFT and GLOH combinations obtain 33% (HesLap-SIFT) and 31% (HesLap-GLOH). HarLap-SC obtains 35%.

To summarize, the localized bag-of-words results suggest two conclusions: first, adding location information can have a much bigger impact on classification accuracy than the choice of detector respectively descriptor. Second, shape features (Geometric Blur, Shape Context, k-Adjacent Segments) can benefit more from location than appearance-based ones. In particular, this boost can be sufficient for outperforming appearance-based features.

7. Summary and Conclusions

In this paper, we have presented an evaluation of local shape- and appearance-based features. Building upon a formerly proposed method [18] based on the comparison of clusterings, we measured local feature statistics over two complementary data sets representing shape- and appearance-based object classes, respectively. We additionally contrasted two clustering techniques, and further evaluated local features as part of two general recognition frameworks with and without spatial information.

The key findings are: Local shape- and appearance-based features do not show great differences in terms of feature statistics over both shape- and appearance-based data sets. The choice of detector is more important on average than the choice of descriptor. Hessian-Laplace with SIFT and GLOH is best on average. Shape-based features (Geometric Blur, k-Adjacent Segments) perform mostly worse than appearance-based ones for classification based on simple occurrence statistics, but benefit more from added location information, and can even overtake appearance-based features on both shape- and appearance-based data.

Acknowledgements This work has been funded, in part, by the EU project CoSy (IST-2002-004250). The authors would like to thank A. Berg and V. Ferrari for providing implementations of Geometric Blur and k-Adjacent Segments, respectively. T. Kadir, D. Lowe and K. Mikolajczyk for making their feature implementations available to the community, and C. Wojek for joint boosting code.

References

Figure 2. Cluster precisions and Naïve Bayes accuracies for SIFT, GB, and $k$-AS. SIFT-plots include the best GLOH and SC curves.
Figure 3. Localized bag-of-words classification accuracies for SIFT, GB, and k-AS. SIFT-plots include the best GLOH and SC curves.