Semi-supervised Concept Detection by Learning the Structure of Similarity Graphs

Symeon Papadopoulos[1], Christos Sagonas[1], Ioannis Kompatsiaris[1],
and Athena Vakali[2]

[1] Information Technologies Institute, CERTH, Greece
{papadop,sagonas,ikom}@iti.gr
[2] Informatics Department, Aristotle University of Thessaloniki, Greece
avakali@csd.auth.gr

Abstract. We present an approach for detecting concepts in images by a graph-based semi-supervised learning scheme. The proposed approach builds a similarity graph between both the labeled and unlabeled images of the collection and uses the Laplacian Eigemaps of the graph as features for training concept detectors. Therefore, it offers multiple options for fusing different image features. In addition, we present an incremental learning scheme that, given a set of new unlabeled images, efficiently performs the computation of the Laplacian Eigenmaps. We evaluate the performance of our approach both on synthetic datasets and on MIR Flickr, comparing it with high-performance state-of-the-art learning schemes with competitive and in some cases superior results.

1 Introduction

Concept detection in images is typically conducted by learning a mapping between a set of features extracted from the input images and a set of concepts. In conventional settings, features extracted from individual labeled images are provided to classifiers (e.g. SVM) in order to learn the features-to-concept mapping. Once a new image appears, concept detection is conducted based on its feature vector, thus considering the image in isoliation. In this work, we tackle concept detection by leveraging the similarity of the unknown image with labeled images of the collection. We call our proposal the Graph Structure Features (GSF) approach. GSF is based on building a similarity graph between the images of the collection and mapping them to low-dimensional features based on the eigenvectors of the graph Laplacian. These features correspond to semantically coherent groups of images, and are thus used to train concept classifiers.

The GSF approach is expressed as a combination of a semi-supervised with a supervised learning algorithm: at the first step the similarity graph and the graph structure features are computed for both the labeled and the unlabeled images, while at the second step the graph structure features are used to train a supervised learning algorithm. The approach leads to high concept detection performance, since it utilizes information on the similarities of unlabeled images with the labeled ones. It can also accommodate a variety of fusion techniques

S. Li et al. (Eds.): MMM 2013, Part I, LNCS 7732, pp. 1–12, 2013.

for combining different sets of features to further improve performance. To make the approach applicable in online settings, we also devise an incremental scheme for computing the graph structure features of unknown images in online mode.

Contributions: The paper proposes a novel semi-supervised concept detection approach that does not rely on features extracted from isolated content items, but from features that capture the similarity structure between the unknown and the labeled items. Thanks to the transformation of original content features into a similarity graph and the extraction of low-dimensional graph structure features, GSF is well-suited to high-dimensional features, and offers several options for performing fusion between multiple feature sets. GSF is tested in comprehensive experiments, in which it is found to outperform or to compete closely with two high performing state-of-the-art approaches. In addition, we propose an incremental implementation of the proposed scheme that achieves similar performance as the batch learning scheme, and enables application of the framework in online learning settings.

2 Related Work

Utilizing the implicit relational structure by computing similarities between images of a collection has been proposed before. In [1], an extended similarity measure is proposed that leverages local neighborhood structures of images, computed from the content and label information of similar images. The extended similarity was used on top of well-known semi-supervised learning methods [2] and shown to improve performance. However, this approach is not amenable to online learning settings, as it relies on global graph computations [2,3].

A different approach is presented in [4], where a sparse similarity graph is constructed based on a convex optimization method. Then, an additional λ_1-norm minimization problem is solved to perform semi-supervised inference on the noisy image tags, and to derive a compact concept space. Online concept detection is possible in the derived concept space. The approach of [4] is shown to yield superior concept detection accuracy in a standard benchmark. However, it suffers from a computationally intensive training step. Another semi-supervised concept detection approach is grounded on the notions of hashing-based λ_1-graph construction and KL-based multi-label propagation [5]. Despite being applicable to large-scale datasets and yielding high performance, this approach is not practical in online learning settings since it relies on a computational scheme that requires 50 iterations (as stated in [5]) for convergence during inference.

Our work is mostly related to [6] that introduces the concept of “social dimensions”, i.e. the top-k Laplacian eigenvectors, as a means to tackle the relational classification problem [7]. We adopt a similar feature representation, but apply it in a different problem, i.e. concept detection in multimedia. Moreover, we consider an extension that renders our approach practical in online settings in contrast to [6] that is only applicable in transductive learning settings.

3 Proposed Approach

The basic formulation of the GSF approach is provided in a transductive learning setting, where both the labeled and the unlabeled samples are available at training time (subsection 3.1). To render our proposal more practical in real-world learning settings, we describe an incremental extension that can predict the concepts of unknown items that were not available during training (subsection 3.2). Finally, we describe a set of possibilities for fusing multiple features with the goal of improving the concept detection performance (subsection 3.3).

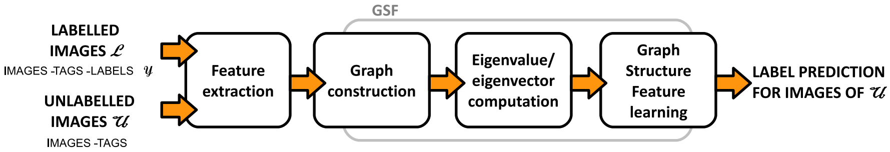

GSF
LABELLED IMAGES $\mathscr{L}$
IMAGES -TAGS -LABELS $\mathscr{Y}$
UNLABELLED IMAGES $\mathscr{U}$
IMAGES -TAGS
Feature extraction
Graph construction
Eigenvalue/ eigenvector computation
Graph Structure Feature learning
LABEL PREDICTION FOR IMAGES OF $\mathscr{U}$

Fig. 1. Overview of proposed approach

3.1 Transductive Learning Setting

The GSF approach is illustrated in Figure 1. Given a set of K target concepts $\mathcal{Y} = \{Y_1, ..., Y_K\}$ and an annotated set $\mathcal{L} = \{(\boldsymbol{x_i}, \boldsymbol{y_i})\}_{i=1}^{l}$ of training samples, where $\boldsymbol{x_i} \in \mathbb{R}^D$ stands for the feature vector extracted from content item i and $\boldsymbol{y_i} \in \{0, 1\}^K$ for the corresponding concept indicator vector, a transductive learning algorithm attempts to predict the concepts associated with a set of unknown items $\mathcal{U} = \{\boldsymbol{x_j}\}_{j=l+1}^{l+u}$, by processing together the sets $\mathcal{L}$ and $\mathcal{U}$.

Based on the features of the input items, a graph $G = (V, E)$ is constructed that represents the similarities between all pairs of items. The nodes of the graph include the items of both sets ($\mathcal{L}$ and $\mathcal{U}$), i.e. $V = V_L \cup V_U$ with $|V| = n$. There are different options for constructing such a graph:

- **Full graph:** All possible edges between the items of V are inserted as edges using weights to determine the degree of similarity for each pair. For two feature vectors x_i, x_j, a popular weighting scheme is the heat kernel:

$$w_{ij} = \exp\left(-\frac{||x_i - x_j||^2}{t}\right) \tag{1}$$

  Full graph is not considered for use with the GSF approach, since GSF requires a sparse adjacency matrix to compute the graph structure features.
- ***k*NN graph:** An edge is inserted between items i and j as long as one of them belongs to the set of top-k most similar items of the other. Two basic variants of this scheme are possible, symmetric and asymmetric, depending on whether both items i, j belong to the similar set of each other or not.
- **ϵNN graph:** A global distance threshold ϵ is defined and then an edge is inserted between items i and j if $||x_i - x_j||^2 < \epsilon$.

Having constructed the similarity graph between the input media items, GSF proceeds with mapping the graph nodes to feature vectors that represent the associations of nodes with latent groups of nodes that form densely connected clusters. In order to extract such features, we first construct the normalized graph Laplacian from the degree (D) and adjacency (A) matrices of the graph:

$$\tilde{L} = D^{-1/2} L D^{-1/2} = I - D^{-1/2} A D^{-1/2}, \tag{2}$$

where $L = D - A$ is the graph Laplacian. Computing the eigenvectors of $\tilde{L}$ corresponding to the C_D smallest non-zero eigenvalues of the matrix results in n vectors of C_D dimensions, which when concatenated form the input matrix $S \in \mathbb{R}^{n \times C_D}$, each row of which is denoted as $S_i \in \mathbb{R}^{C_D}$ and constitutes the graph structure feature vector for media item i. These features, known as Laplacian Eigenmaps [8], are derived by solving the following minimization problem:

$$\underset{S}{\operatorname{argmin}}\, S^T \tilde{L} S, \qquad \text{s.t. } S^T S = I. \tag{3}$$

In the final step, a classifier is trained using the structure feature vectors of the labeled items as input. In our implementation, we opted for the use of SVM. Apart from classification performance considerations, it is important for retrieval applications that the classifier produces real-valued prediction scores for unlabeled items, so that they can be ranked per concept. Producing real-valued prediction scores is also important for performing result fusion when multiple sets of features are available (see subsection 3.3).

3.2 Incremental Learning

The scheme of subsection 3.1 requires both labeled and unlabeled items to be available at train time for constructing a similarity graph from both sets and computing the respective structure features. In case new unlabeled samples were provided as input, it would be necessary to reconstruct the similarity graph for the extended set of items and recompute the eigenvectors of $\tilde{L}$ (Equation 2). This is clearly impractical in settings where new media items are regularly arriving to the indexing system. Thus, it is necessary to devise incremental means for computing the graph structure features of unknown items.

Linear Projection (LP): A simple approach for deriving the graph structure features of a new item $n+1$ is to determine the set N_{n+1} of k most similar items and then to compute the weighted mean of their graph structure features:

$$\hat{S_{n+1}} = \frac{\sum_{j \in N_{n+1}} w_{n+1,j} S_j}{\sum_{j \in N_{n+1}} w_{n+1,j}} \tag{4}$$

where $w_{n+1,j}$ denotes the similarity score between the new item and neighbouring item j, which may be computed by use of the heat kernel (Equation 1).

Submanifold Analysis (SA): A more accurate technique for estimating the graph structure feature vector of item $n+1$ relies on the analysis of the graph submanifold around it [9]. Initially, the $(k+1) \times (k+1)$ similarity matrix W_S is constructed between the new item and the k most similar items. Then, the sub-diagonal and sub-Laplacian matrices are derived as follows:

$$D_S(i,i) = \sum_j W_S(j,i), \qquad L_S = D_S - W_S$$

We compute the eigenvalues $0 = \lambda_S^0 \leq \lambda_S^1 \leq ... \leq \lambda_S^d$ and d eigenvectors $v_1, ..., v_d$ of the non-zero eigenvalues. This computation is lightweight since k is selected to be small (cf. experiments in Section 4). Finally, a weight vector $C = [c_1 ... c_k] \in \mathbb{R}^k$ is determined by minimizing the following reconstruction error:

$$\underset{c_i}{\text{argmin}} \, |v_{k+1} - \sum_{i=1}^{k} c_i v_i|^2 \quad \text{s.t.} \, \sum_i c_i = 1$$

by using a non-linear constaint optimization method [10]. Once the weight vector C is computed, the new feature vector is computed as $\hat{S}_{n+1} = \sum_{i=0}^{k} c_i S_i$.

3.3 Fusion of Multiple Features

GSF offers several options for fusing different sets of input features:

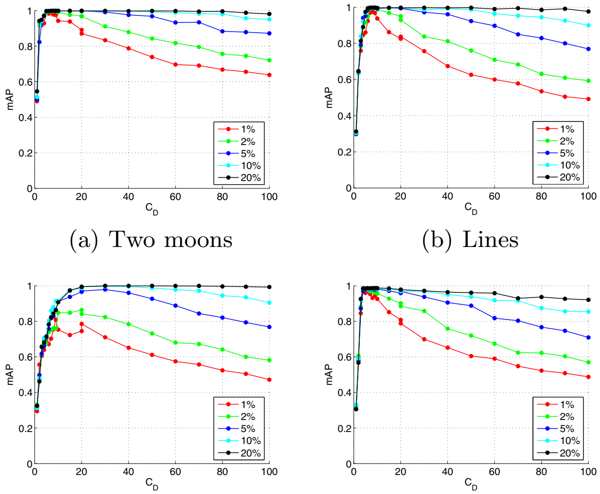
- **Feature fusion (F-FEAT):** This constitutes a common early fusion technique. In its simplest form, this is implemented by simple vector concatenation and by an optional feature normalization step.
- **Similarity graph fusion (F-SIM):** Combining two similarity graphs G_1 and G_2 constructed from two feature sets, this technique produces a fused graph G_F. The combination can be implemented by means of an elementwise additive or multiplicative operation between G_1 and G_2.
- **Graph structure feature fusion (F-GSF):** According to this, graph structure feature vectors are computed separately from each similarity graph and are combined by concatenating the corresponding vectors.
- **Result fusion (F-RES):** This widely used late fusion technique is implemented by training a second-level classifier with the prediction scores of individual feature concept detectors as inputs.

4 Evaluation

The proposed approach was evaluated on both synthetic and real data.

4.1 Synthetic Data

GSF was thoroughly tested using synthetic 2D distributions with limited number of samples to enable visualization of input data and classification results, and to make possible the exploration of a large number of experimental settings and the repetition of each experiment multiple times for deriving reliable performance estimates. Four kinds of distributions were used: (a) Two moons, (b) Lines, (c) Circles, (d) Gaussians. The following performance aspects were investigated:



(a) Two moons

(b) Lines

(c) Circles

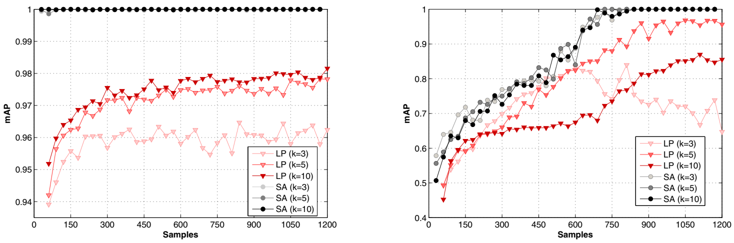
(d) Gaussians

Fig. 2. Role of number of graph structure features (C_D) and training samples (α)

- **Parameters of GSF:** Number of graph structure features C_D, graph construction technique (kNN or ϵNN) and associated parameters (k, ϵ).
- **Inductive learning:** GSF was evaluated in an inductive learning setting with the use of the baseline scheme of subsection 3.2.
- **Fusion method:** Four fusion methods of subsection 3.3.

Parameters of GSF: To test the behaviour of the proposed approach for different number of graph structure features, for each distribution, the number of features C_D was varied from 1 to 100 and three settings were tested: low, medium and high noise. In the case of Two moons and Gaussians distributions, we found that the learning performance is almost insensitive to C_D even when large amounts of noise are added to the input data. In the case of the Lines distribution, a similar behaviour is observed with the exception of the high noise setting, where we found that increasing the number of graph structure features harms performance. Finally, in the Circles distribution, a larger number of graph structure features was found to be beneficial, especially in high noise settings. Furthermore, performance was measured for different training set sizes (Figure 2). When 10% or more of the input data are available for training, performance appears to be insensitive to the number of features C_D. However, for smaller training sizes, GSF appears to be increasingly sensitive to the selection of C_D.

The role of the graph construction technique was examined by graphing the performance of GSF using both kNN and ϵNN for different values of k and ϵ. Out of the two competing options, kNN appears as less sensitive to the selection of k, than ϵNN is to ϵ. Only in the case of the Circles distribution, increasing the value of k seems to have an adversary effect on performance. In case of ϵNN,



(a) Lines (b) Circles

Fig. 3. Inductive learning performance

increasing the value of ϵ seems to harm performance, whereas in the case of the Circles distribution the relation between performance and ϵ is non-monotonic.

Additionally, we investigated the computational requirements of GSF when an increasing number of input features are provided as input. As the number of input features increases, so does the execution time of SVM-RBF and for large number of features it is considerably slower than GSF (for instance for 500 features SVM-RBF was 4x slower than GSF). Given the multitude of features (> 1000) typically used in multimedia concept detection, GSF may be considered a much more efficient option for practical problem settings.

Inductive Learning: We generated 3000 samples from the four distributions and $\alpha\%$ were used for training. The rest were not used in the similarity graph construction. Once they were provided as input, the two incremental schemes of subsection 3.2 (LP and SA) were used to estimate their graph structure features. The estimation of graph structure features was conducted by setting k equal to 3, 5 and 10. The results are presented in Figure 3 (for Lines and Circles). Submanifold Analysis (SA) consistently outperforms Linear Projection (LP). In the case of Two moons and Gaussians, their performance is comparable. However, for Lines and Circles, there is a clear improvement when using SA over LP. An additional benefit of SA is that its performance seems to be only marginally affected by k in contrast to LP that is sensitive to it. The above observation coupled with the fact that the overall performance rates for SA are exceptional demonstrate that the devised incremental scheme can effectively be used in online learning settings. We recognize that online learning with synthetic data is an easier problem compared to real-world settings due to the fact that the training and test samples are drawn from the same distribution.

Fusion Method: To evaluate the fusion techniques of subsection 3.3, we generated feature vectors from two 2D independent distributions. In the first experiment, both distributions were generated according to the Lines pattern, and the free parameter was the size of the training set. In the second experiment, Circle distributions were used and increasing amounts of noise were added to one of the two feature sets. In both experiments, the two late fusion techniques, namely GSF-level fusion (F-GSF) and result fusion (F-RES) led to the best performance, higher than both feature-level performance and early fusion methods, namely feature-level (F-FEAT) and similarity-based fusion (F-SIM).

4.2 MIR-Flickr

MIR-Flickr [13] was used as a real-world benchmark. Ground truth is available for all 25,000 images of the collection in three variants, strict relevance of image to concept (REL), loose relevance (POT) and aggregate annotation (ALL) [13]. REL and POT contain annotations for 14 concepts, while ALL includes annotations for 24 more concepts (38 in total). We compared GSF with two state-of-the-art approaches, the Semantic Spaces (SESPA) approach [14] and Multiple Kernel Learning (MKL) [15], which were selected for two reasons:

- High performance: SESPA reports better than average performance for the ImageCLEF 2009 photo annotation task. MKL reports higher performance compared to the best method in the PASCAL VOC'07 dataset.
- Reproducibility: The authors of both approaches have disclosed all necessary information and data (e.g. visual features of the dataset images) making possible the replication of their experimental settings.

In all experiments that follow, the similarity graphs were built with the kNN technique, setting $k = 2000$ (empirically found to produce slightly higher mAP scores compared to $k = 500, 1000$). We used different k values during the testing of incremental methods (LP, SA). In addition, we experimentally selected $C_D = 500$ due to its higher performance compared to $C_D = 100, 200, 1000$.

GSF vs SESPA. SESPA [14] is based on the concept of creating a high-dimensional space, in which closely positioned images are expected to be semantically related. The authors use three well known visual descriptors based on SIFT. Apart from the SESPA approach, the authors of [14] specify a complete MIR-Flickr-specific evaluation protocol. Following that protocol, we computed the performance of GSF for three different train-test splits and over all three annotation sets (REL, POT, ALL). The performance was quantified in terms

Table 1. GSF vs SESPA

Method	**$\vert\mathcal{L}\vert = 5000$**			**$\vert\mathcal{L}\vert = 10000$**			**$\vert\mathcal{L}\vert = 15000$**		
	REL	**POT**	**ALL**	**REL**	**POT**	**ALL**	**REL**	**POT**	**ALL**
SESPA	0.216	0.313	0.276	0.237	0.324	0.287	0.240	0.327	0.292
GSF-F$_1$	0.202	0.313	0.272	0.225	0.339	0.297	0.237	0.350	0.308
GSF-F$_2$	0.146	0.257	0.216	0.160	0.276	0.233	0.171	0.286	0.244
GSF-F$_3$	0.142	0.265	0.142	0.162	0.285	0.240	0.174	0.295	0.251
GSF-C	0.222	0.332	0.291	0.247	0.361	0.319	0.269	0.380	0.339
GSF-D$_1$	0.227	0.340	0.292	0.256	0.371	0.325	0.275	**0.388**	0.345
GSF-D$_2$	**0.237**	**0.357**	**0.314**	**0.263**	**0.379**	**0.337**	**0.278**	0.384	**0.350**

Table 2. GSF vs MKL ($|\mathcal{L}| = 12500$, ALL)

	MKL	**GSF**	**GSF***
Visual	**0.530**	0.511	0.511
Tag	0.424	**0.433**	**0.474**
Visual+Tag	0.623	**0.624**	**0.641**

of mean Average Precision (mAP) and Area Under the Curve (AUC). Table 1 presents the obtained results (only mAP). Several variants of GSF were evaluated: (a) single feature GSF for each of the three features (GSF-F_1, GSF-F_2 and GSF-F_3), (b) multi-modal fusion (GSF-C) relying on the concatenation of the Laplacian eigenvectors of the three similarity graphs , (c) two result fusion variants (GSF-D_1 and GSF-D_2). Result fusion in GSF-D_1 relied on linear SVM [11], while in GSF-D_2 it relied on SVM-RBF [12].

The results of Table 1 indicate that GSF clearly outperforms SESPA. Across all train-test splits and all annotation sets, one or more of the feature fusion variants of GSF yield higher mAP scores. In terms of AUC, the difference in performance is less pronounced and in one case ($|\mathcal{L}| = 5000$, REL), SESPA outperforms GSF by a small margin. Out of the GSF variants, the highest performing ones are those relying on result fusion. There is a tendency for bigger improvement in performance (compared to SESPA) as the training set size increases, and the performance increase is larger on the POT set than on REL.

GSF vs MKL. The MKL approach by [15] leverages Multiple Kernel Learning for combining a kernel based on image content with a second kernel that encodes tag information. To derive the visual kernel, the authors of [15] make use of 15 features, while the tag-based kernel is computed by selecting the 457 most frequent tags as features. Table 2 presents the comparison of the mAP-based performance between MKL and GSF on a 50-50 train-test split (ALL annotation set). According to it, MKL outperforms GSF by 3.7% when only the visual features are used as input, while GSF outperfoms MKL by 2.1% when tag features are used. When all features are provided as input, GSF outperforms MKL by a very small margin. The third column of Table 2 reports the results obtained by GSF with the use of tag features computed by use of inverse document frequency. Using the MKL approach with this feature would be impractical due to its very high dimensionality (68,894). Instead, GSF could make use of the feature, since its complexity is not significantly affected by its dimensionality.

Figure 4 illustrates the top eight results from GSF using visual and tag features for the first five REL concepts of MIR-Flickr. All returned results (apart from the first in the Baby concept) are highly relevant to the query concepts. Manual inspection of the top 50 results for other concepts of the dataset further confirms their high relevance.

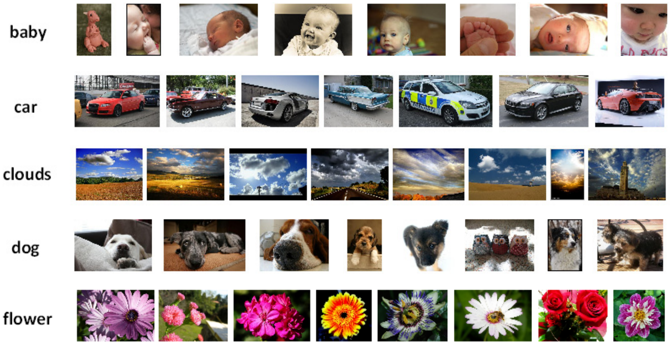
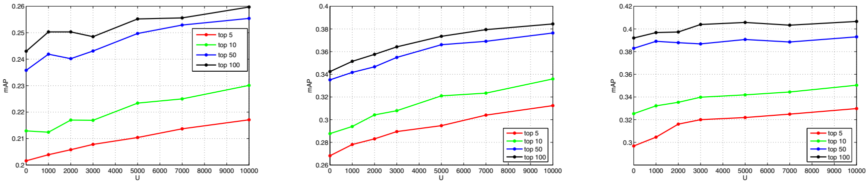


Fig. 4. Examples of top retrieved images per concept using the batch mode GSF

Contribution of Unlabeled Samples. In this experiment, we evaluate the gains in learning performance when more unlabeled samples are available. Out of the 25,000 images, 5,000 were reserved for training and 10,000 for testing. The remaining 10,000 images were used as unlabeled items. Figure 5 presents the results for three different features (GIST, DenseSiftV3H1, TagRaw50). The features are representative of three feature types: global visual, local visual, and text-based. The experiment is repeated for different values of k, i.e. the number of top-k most similar neighbours used by SA.

Adding more unlabeled samples together with the labeled ones leads to significant performance gains. In the case of GIST, the largest performance benefits are observed when $k = 10$: compared to using only the labeled examples (mAP $= 0.21$), the performance of the concept detector rises to 0.235, i.e. a relative increase of 11.9%, when 10,000 unlabeled samples are added during the computation of the graph structure features. For larger values of k, the performance benefits appear to be much less pronounced. In the case of DenseSift3VH1, significant performance gains are observed across all values of k. For instance, for $k = 10$ a 16.3% increase in performance is recorded, while for $k = 50$ the performance benefit still amounts to 11.8%. In case of the tag-based features, the situation is more similar to the one described for GIST. The performance improvements are clearer for smaller values of k and appear to level off after 3000-4000 unlabeled samples are added together with the labeled samples.

In the same experiment, we could also conclude that the performance of the SA online learning scheme is comparable to the one of the transductive learning scheme. For instance, when using $k = 100$ (which leads to the best performance for SA among the tested values), the map score of SA using the GIST features and no unlabeled images is 0.243, while the score achieved by the transductive version of GSF is 0.2456 (1.1% higher). When 5000 unlabeled images are provided together with the labeled ones, then the SA performance (again for GIST features) rises to 0.2552, slightly higher than 0.2522, which was achieved by the transductive implementation of GSF. Similar observations were made when using other features as well.



(a) GIST (b) DenseSiftV3H1 (c) TagRaw50

Fig. 5. Performance gains by adding unlabeled samples

5 Conclusions

We presented GSF, a multimedia annotation approach leveraging the structure of image similarity graphs. GSF relies on the assumption that images with similar positions on the graph tend to carry the same concepts. Concept detection is then conducted by training classifiers using the graph Laplacian eigenvectors as features. Apart from the transductive formulation of the problem, we proposed two incremental versions, one based on Linear Projection and the other on Submanifold Analysis, and described four fusion techniques applicable to GSF. The transductive version of our approach was evaluated on a wide range of synthetic distributions, and was also compared against two state-of-the-art learning approaches on the MIR-Flickr dataset, giving superior or comparable results. The two incremental implementations were compared on synthetic data, with SA method yielding superior performance. SA was also evaluated on MIR-Flickr in a semi-supervised learning setting, resulting in mAP rates very close to the ones achieved with the transductive version. In addition, SA was found to be quite fast; the average time for predicting the concepts of an image using SA with $k = 5$ was measured to be 38.4ms (not including feature extraction). In the future, we plan to further study the computational characteristics of the proposed approach by applying it to larger scale problems.

Acknowledgements. This work is supported by the SocialSensor FP7 project, partially funded by the EC under contract number 287975.

References

1. Wang, M., Hua, X.-S., Tang, J., Hong, R.: Beyond distance measurement: constructing neighborhood similarity for video annotation. TMM 11(3), 465–476 (2009)
2. Zhu, X.: Semi-supervised learning with graphs. PhD Thesis, Carnegie Mellon University (2005) 0-542-19059-1
3. Zhou, D., Bousquet, O., Navin Lal, T., Weston, J., Schölkopf, B.: Learning with Local and Global Consistency. In: Advances in NIPS, vol. 16, pp. 321–328. MIT Press (2004)

4. Tang, J., et al.: Inferring semantic concepts from community contributed images and noisy tags. ACM Multimedia, 223–232 (2009)
5. Chen, X., et al.: Efficient large scale image annotation by probabilistic collaborative multi-label propagation. ACM Multimedia, 35–44 (2010)
6. Tang, L., Liu, H.: Leveraging social media networks for classification. Data Mining and Knowledge Discovery 23(3), 447–478 (2011)
7. Macskassy, S.A., Provost, F.: Classification in Networked Data: A Toolkit and a Univariate Case Study. Journal of Machine Learning Research 8, 935–983 (2007)
8. Mikhail, B., Partha, N.: Laplacian Eigenmaps for dimensionality reduction and data representation. Neural Computing 15(6), 1373–1396 (2003)
9. Jia, P., Yin, J., Huang, X., Hu, D.: Incremental Laplacian eigenmaps by preserving adjacent information between data points. PR Letters 30(16), 1457–1463 (2009)
10. Leyffer, S., Mahajan, A.: Nonlinear Constrained Optimization: Methods and Software. Preprint ANL/MCS-P1729-0310 (2010)
11. Fan, R., Chang, K., Hsieh, C., Wang, X., Lin, C.: LIBLINEAR: A Library for Large Linear Classification. Journal of ML Research 9, 1871–1874 (2008)
12. Chang, C.-C., Lin, C.-J.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2(3), 27:1–27:27 (2011)
13. Huiskes, M.J., Michael, S., Lew, M.S.: The MIR Flickr Retrieval Evaluation. In: Proceedings of ACM Intern. Conference on Multimedia Information Retrieval (2008)
14. Hare, J.S., Lewis, P.H.: Automatically annotating the MIR Flickr dataset. In: ACM ICMR, pp. 547–556 (2010)
15. Guillaumin, M., Verbeek, J., Schmid, C.: Multimodal semi supervised learning for image classification. In: Proceedings of IEEE CVPR Conference, pp. 902–909 (2010)