Learning Class Specific Graph Prototypes

Shengping Xia[1] and Edwin R. Hancock[2]

[1] ATR Lab, School of Electronic Science and Engineering,
National University of Defense Technology, Changsha, Hunan, P.R. China 410073
[2] Department of Computer Science, University of York, York YO10 5DD, UK

Abstract. This paper describes how to construct a graph prototype model from a large corpus of multi-view images using local invariant features. We commence by representing each image with a graph, which is constructed from a group of selected SIFT features. We then propose a new pairwise clustering method based on a graph matching similarity measure. The positive example graphs of a specific class accompanied with a set of negative example graphs are clustered into one or more clusters, which minimize an entropy function. Each cluster is simplified into a tree structure composed of a series of irreducible graphs, and for each of which a node co-occurrence probability matrix is obtained. Finally, a recognition oriented class specific graph prototype (CSGP) is automatically generated from the given graph set. Experiments are performed on over 50K training images spanning ∼500 objects and over 20K test images of 68 objects. This demonstrates the scalability and recognition performance of our model.

1 Introduction

Local invariant feature based modeling from multi-view images has become a popular approach to object recognition. Lowe [9] presents a method for combining multi-views of a 3D object into a single model representation using so called SIFT (scale invariant feature transform) features [10]. Rothganger et al. demonstrate how to acquire true 3D affine and Euclidean models from multiple unregistered images, and perform recognition from arbitrary viewpoints [12]. Ferrari et al. present an approach, which includes a mechanism for capturing the relationships between multiple model views, and use the model to effect simultaneous object recognition and image segmentation [6]. Todorovic and Ahuja demonstrate the completely unsupervised extraction and learning of a visual category that occurs frequently in a given set of images [15]. This body of work has demonstrated impressive levels of performance and provide arguably the most successful paradigm for object recognition [4]. Recently, the methods for modeling and recognition have been extended to large image databases. For instance, Nister and Stewenius perform real-time object specific recognition using a database of 40,000 images [11] while Torralba et al. [16] have considered how to scale the methods to the internet.

A limitation of these models is that only small variations in object viewpoint can be accommodated. A more significant limitation is the large number of parameters required and the assumptions imposed (such as the Gaussian nature of the data) which may not apply in practice [8][14]. To avoid these shortcomings, our goal is to derive a canonical model of visual category or class which is totally data driven and does not rely on limiting model assumptions.

One way to overcome these problems is to work with view-clusters and in this way deal with large variations in viewpoint [6][9][12]. For instance Lowe [9] clusters together training images from similar viewpoints into single model views. Each view cluster consists of the complete set of SIFT features extracted from the training views, along with a record of the location, orientation, and scale for each feature within that view. Object models can be incrementally updated and refined as recognition is performed. However, since the view clusters contain the union of the detected SIFT features encountered in the relevant set of training data, two potential pitfalls are encountered. The first of these is that the view-clusters contain an unmanageably large number of feature points. The second problem is that many of the feature points are not salient, and this gives rise to matching ambiguities or errors. Additional problems that remain to be solved are a) how to select the model views and b) how many views suffice. Our method can automatically select an irreducible set of image views, and then merge the information contained within the redundant image views into the most similar model views using a feature co-occurrence probability matrix.

Since the contents of each image can be captured using structured data in the form of SIFT features, attributed graphs provide a versatile and expressive representational tool [2]. There have been previous efforts aimed at generating a prototype graph from a set of training examples. For example, Crandall and Huttenlocher [3] use a graphical structure referred to as k-fans to model geometric constraints based on a statistical model. Here a Gaussian distribution is used to represent geometric relationships between the parts of an object. Jiang, Munger and Bunke [7] use a genetic search algorithm to learn a median graph from a training set. Torsello and Hancock [17] have constructed the class-prototype through tree-union and have performed clustering using a mixture of tree-unions controlled by a description length criterion. These ideas are taken further by Escolano and Lozano [1] who extended the methodology from trees to graphs, and used an EM algorithm for clustering. However, these approaches also only deal with a small range of views. Here we extend the methodology by developing a simple information criterion for learning the prototypes.

We commence by representing each image, or each region of interest in an image, with a graph which is constructed from a group of selected local invariant features. Using graphs alone, it is difficult to construct models from images containing large variations in viewpoint based on only one central prototype. Hence we turn to graph prototypes as a means of representing object models. Specifically, we aim to construct class specific prototype-graph (CSGP) for object recognition. A CSGP model is automatically generated from an arbitrary set of images of the relevant objects under significant or complex variations in imaging

conditions. In a CSGP object model, a group of irreducible object views are selected by minimizing an information theoretic entropy criterion defined on a given training set. Each irreducible graph can be regarded as a central prototype used to seed an object model. The information contained in those graphs that are not selected as prototypes for model construction, and referred to as redundant graphs, is combined with the most similar prototype using node co-occurrence probabilities.

In previous work we have shown how the salient features required to construct such a representation can be extracted from images, and detail the representation of class prototypes [18,19]. The novel contribution of the current work is to show how the representation may be optimised with respect to both representational capacity and performance. The paper is organized as follows. In Section 2, we introduce some preliminaries, including how SIFT features are ranked, and present a pairwise graph matching method and the associated similarity measure required to merge graphs. In Section 3, we present the method used to learn CSGP models. We present experimental results in Section 4 and conclude the paper in Section 5.

2   Preliminaries

For an image, those SIFT [10] features that are robustly matched with the SIFT features in similar images can be regarded as salient representative features. Motivated by this, a method for ranking SIFT features has been proposed in [21]. Using this method, the SIFT features of an image $\mathcal{I}$ are ranked in descending order according to a matching frequency. We select the $\mathcal{T}$ best ranked SIFT features, denoted as $\mathcal{V} = \{V^t,\ t = 1, 2, ..., \mathcal{T}\ \}$, where $V^t = ((\overrightarrow{X}^t)^T, (\overrightarrow{D}^t)^T, (\overrightarrow{U}^t)^T)^T$. Here, $\overrightarrow{X}^t$ is the location, $\overrightarrow{D}^t$ is the direction vector and $\overrightarrow{U}^t$ is the set of descriptors of a SIFT feature. In our experiments, $\mathcal{T}$ is set to 40. If there are less than this number of feature points present then all available SIFT features in an image are selected. We then represent the selected SIFT features in each image using an attributed graph.

Formally, an attributed graph G [3] is a 2-tuple $G = (V, E)$, where V is the set of vertices, $E \subseteq V \times V$ is the set of edges. For each image, we construct a Delaunay graph G using the coordinates of the selected SIFT features. In this way, we can obtain a set of graphs $\mathbb{G} = \{G_l,\ l = 1, 2, ..., N\}$ from a set of images.

We perform pairwise graph matching (PGM) with the aim of finding a maximum common subgraph (MCS) between two graphs G_l and G_q, and the result is denoted as $MCS(G_l, G_q)$. In general, this problem has been proven to be NP-hard. Here we use a Procrustes alignment procedure [13] to align the feature points and remove those features that do not satisfy the spatial arrangement constraints.

Suppose that X_l and X_q are respectively the position coordinates of the selected features in graphs G_l and G_q. We can construct a matrix

$$Z = \arg\min \|X_l \cdot \Omega \cdot X_q\|_F,\ subject\ to\ \Omega^T \cdot \Omega = I. \tag{1}$$

where $\| \bullet \|_F$ denotes the Frobenius norm. The norm is minimized by the nearest orthogonal matrix

$$Z^* = \Psi \cdot \Upsilon^*, \; subject \; to \; X_l^T \cdot X_q = \Psi \; \cdot \; \Sigma \; \cdot \; \Upsilon^*. \tag{2}$$

where $\Psi \; \cdot \; \Sigma \; \cdot \; \Upsilon^*$ is the singular value decomposition of matrix $X_l^T \cdot X_q$. The goodness-of-fit criterion is the root-mean-squared error, denoted as $e(X_l, X_q)$. The best case is $e(X_l, X_q) = 0$. The error e can be used as a measure of geometric similarity between the two groups of points. If we discard one pair of points from X_l and X_q, denoted as $X_{l \to i}$ and $X_{q \to i}$, $e(X_{l \to i}, X_{q \to i})$, $i = 1, 2, ..., \|CS(G_l, G_q)\|$ can be obtained, where $CS(G_l, G_q)\|$ is the number of SIFT features between two graphs initially matched using the matching proposed in [18,19]. The maximum decrease of $e(X_{l \to i}, X_{q \to i})$ is defined as

$$\Delta e(\|CS(G_l, G_q)\|) = e(X_l, X_q) - \min\{e(X_{l \to i}, X_{q \to i})\} \tag{3}$$

if $\Delta e(\|CS(G_l, G_q)\|)/e(X_l, X_q) > \epsilon$, e.g. $\epsilon = 0.1$, the corresponding pair X_l^i and X_q^i is discarded as a mismatched feature pair. This leave-one-out procedure can proceed iteratively, and is referred as the iterative Procrustes matching of G_l and G_q.

Given $MCS(G_l, G_q)$ obtained by the above PGM procedure, we construct a similarity measure between the graphs G_l and G_q as follows:

$$R(G_l, G_q) = \|MCS(G_l, G_q)\| \; \times \; (\; \exp(- \; e(X_l, X_q)) \;)^{\kappa}. \tag{4}$$

where $\|MCS(G_l, G_q)\|$ is the cardinality of the MCS of G_l and G_q, κ is the number of mismatched feature pairs discarded by iterative Procrustes matching, which is used to amplify the influence of the geometric dissimilarity between X_l and X_q.

We use the similarity measure to define a *Positive Matching Graph* (PMG) using the test

$$P(G_l, G_q) = \begin{cases} 1 & R(G_l, G_q) \geq R_{\tau_1} \\ 0 & else \end{cases} \tag{5}$$

where R_{τ_1} is a threshold on the similarity measure. If $P(G_l, G_q) = 1$, the two graphs G_l and G_q are called a PMG pair.

Finally, for the graph set $\mathbb{G} = \{G_q, \; q = 1, 2, ..., N\}$, for each graph $G_l \in \mathbb{G}$, and the remaining graphs in the set ($\forall G_q \in \mathbb{G}$), we obtain the pairwise graph similarity measures $R(G_l, G_q)$ defined in Equation (4). Using the similarity measures we rank in descending order all graphs G_q. The K top ranked graphs are defined as the K-nearest neighbor graphs (KNNG) of graph G_l, denoted as $\mathbb{K}\{G_l\}$.

3   Class Specific Graph Protypes

In this Section, we commence by showing how to construct an initial class specific graph prototype (CSGP). We then describe how to simplify such a CSGP model to obtain a compact model which suffice to capture structural variations from multiple views of objects of a particular class.

3.1   Construction of Initial CSGP

A Class Specific Graph Prototype (CSGP) is defined as 2-tuple $CSGP = (PV, PE)$, where 1) the prototype vertex $PV = \{ \ G_l, l = 1, 2, ..., N \ \}$ is a finite set of graphs forming the nodes of the CSGP; 2) the prototype-edge PE is the edge set of the CSGP. There is an edge between two nodes G_l and G_q, denoted by $Edge(G_l, G_q)$, if and only if $P(G_l, G_q) = 1$.

For efficiency, we use an incremental clustering tree-RSOM proposed in [20] for incrementally learning large corpus of SIFT descriptors and obtaining $K_\tau\{G_l\}$.

3.2   Learning a CSGP Model

Optimal Pairwise Clustering of CSGP. Suppose an initial CSGP model has been trained from a batch of images. In the CSGP model, the relationships of the graphs $K_\tau\{G_l\}$ have been encoded and will be updated with additional graphs. The number of the items in $K_\tau\{G_l\}$ is largely influenced by the threshold R_τ. We wish to obtain a compact representation from the initial CSGP. To this end we provide two basic definitions.

The siblings of G_l are defined as:

$$S\{G_l\} = \{G_q \in K_\tau\{G_l\} \mid P(G_l, G_q) = 1\} \triangleq S_{R_\tau}\{G_l\}. \tag{6}$$

For each graph $G_q \in S\{G_l\}$, the corresponding siblings can also be obtained. In this way, we can iteratively obtain a series of graphs which all satisfy consistent sibling relationships. The graph set, obtained in this way, is called a family tree of graph G_l.

Given a CSGP model, a Family Tree of a Graph (FTOG) of G_l with k generations and denoted as $F\{G_l, k\}$, is defined as:

$$F\{G_l, k\} = F\{G_l, k - 1\} \bigcup_{G_q \in F\{G_l, k-1\}} S_{R_\tau}\{G_q\}. \tag{7}$$

where, if $k = 1$, $F\{G_l, 1\} = F\{G_l, 0\} \bigcup S\{G_l\}$ and $F\{G_l, 0\} = \{G_l\}$; and the process stops when $F\{G_l, k\} = F\{G_l, k + 1\}$. An FTOG, whose graphs satisfy the restriction defined in Equation (6), is a sub-set of graphs in a CSGP model and can be regarded as a cluster of graphs. We refer to this process as pairwise clustering.

Consider a CSGP model, $\forall R_\tau$, $\exists G_{l_i}$,$i = 1, 2, ..., c$, subject to $F\{G_{l_i}, g\} \bigcap F\{G_{l_j}, g\} = \emptyset$, $i \neq j, i, j \in \{1, 2, ..., c \ \}$. We uniquely label these FTOG's as $L_1, L_2, ..., L_c$ and denote the cluster set as follows:

$$C_{R_\tau} = \{ \ F_{R_\tau}\{G_l, \infty\}\} \triangleq \{ \ F_l \mid l \in \{L_1, L_2, ..., L_c\} \ \}. \tag{8}$$

Clearly, the clustering result is influenced by threshold R_τ. Hence, we present a learning approach to obtain the optimal value of R_τ for a given CSGP model.

We can regard the above clustering process as a classification problem which gives rise to four possible outcomes, i.e a) true positive, b) false negative, c) true negative and d) false positive. The number of instances corresponding to

the above four cases are separately denoted as $|TP|$, $|FN|$, $|TN|$ and $|FP|$. Accordingly a confusion matrix can be constructed representing the dispositions of the set of instances. This matrix forms the basis for computing many common performance measures [5], such as precision (p), recall (r) and the F-measure (f).

We denote the positive graph set as g^+, and the negative graph set g^-. We only consider those FTOGs generated from those graphs which belong to g^+. Suppose that $TP = |C_{R_\tau}\{g^+\}|$, $FP = |g^- \bigcap C_{R_\tau}\{g^+\}|$ and $|P| = |g^+|$. We define the precision, recall and the F-measure as $p\{R_\tau \mid g^+, g^-\} = TP/(TP + FP)$, $r\{R_\tau \mid g^+, g^-\} = TP/P$ and $f\{R_\tau \mid g^+, g^-\} = \frac{2}{1/p+1/r}$. We define an entropy on $C_{R_\tau}\{g^+\}$ as follows.

$$E\{C_{R_\tau}\} = -\sum_{l=1}^{L_{N_c}}\{p_l \cdot \log p_l\} \quad where \quad p_l = \frac{\|F_l\|}{\sum_{l=1}^{L_{N_c}} \|F_l\|}, l = 1, 2, ..., L_c. \quad (9)$$

We wish to find the optimal threshold $R_{\tau_1} = \max\{\arg\min_{R_\tau} E\{C_{R_\tau}\}\}$ such that $f\{R_\tau \mid g^+, g^-\} > 1 - \varepsilon$, where the threshold ε is is heuristically set to 0.02 in our experiments. The smaller $E\{C_{R_{\tau_1}}\}$, the better. If $E\{C_{R_{\tau_1}}\} = 0$, then all the graphs in g^+ are clustered into a single FTOG, and the corresponding F-measure is greater than $1 - \varepsilon$. The optimal cluster membership is such that $C_{R_{\tau_1}}\{g^+\} \triangleq \{ F_l \mid l \in \{L_1, L_2, ..., L_c\}\}$.

In the above process, with the decrease of threshold R_τ, then $f\{R_\tau \mid g^+, g^-\}$ may decrease significantly. Namely, if $\exists G_q \in g^+$ and R_τ, subject to:

$$\lim_{\nabla R_\tau \to 0} \frac{\nabla f\{R_\tau \mid g^+, g^-\}}{\nabla R_\tau} \to \infty. \quad (10)$$

4   Experimental Results and Discussions

We have collected 53536 images as a training data set which spans more than 500 objects, including some human faces and scenes. From the training data we have obtained an RSOM clustering tree with 25334 leaf nodes using the method described in [20]. The method was implemented using Matlab 7.2 and run on a 2.14GHz computer with 2G RAM. In the incrementally training process, we have obtained $K_\tau\{G_l\}$ for each of the graphs. Following this, we individually train the CSGP models for the above 68 labeled object classes using the method presented in Section 3.2. For training a single specific object, the positive graphs are those of the desired class and all remaining graphs, labeled or unlabeled, are regarded as the negative set. This method is also adopted for recognition test.

The training process is incrementally carried out commencing from the first batch of 3600 graphs of A to all 50K graphs. For each object, the minimized entropy, and its accompanying F-measure and similarity threshold have been learned. As a second training stage, the similarity threshold for selecting irreducible graphs, the number and percentage of irreducible graphs are also learned. These parameters are part of the generated CSGP model and will be used for recognition and incremental training.

As distinct from most object recognition methods, where many implementation parameters can be varied, our method individually generate optimized models for each concerned object from given training data-sets.

Using the trained model, the F-measures of recognition test for Object 1 to Object 68 are shown in Table 1. It is interesting to note that the test recognition performance of Object B1 to B8 is very close to that obtained in [12] when only SIFT features are used (corresponding to the 8 objects marked magenta in Table 1. However, our results are obtained with large negative sample sets.

Table 1. F-measure f for given test set of Object 1∼68

ID	f	ID	f	ID	f	ID	f	ID	f	ID	f	ID	f	ID	f	ID	f	ID	f
1	.996	2	.955	3	1.0	4	.983	5	1.0	6	1.0	7	1.0	8	.989	9	1.	10	1.0
11	.985	12	1.0	13	.987	14	1.0	15	1.0	16	1.0	17	1.0	18	.984	19	.979	20	1.0
21	.985	22	1.0	23	1.	24	1.0	25	1.0	26	1.0	27	.995	28	1.0	29	1.	30	1.0
31	1.0	32	1.0	33	.992	34	1.0	35	1.0	36	1.0	37	1.0	38	.993	39	.979	40	1.0
41	1.0	42	0.989	43	1.0	44	.995	45	.941	46	1.0	47	1.0	48	1.0	49	1.0	50	.988
51	.625	52	1.0	53	1.0	54	1.0	55	.714	56	1.0	57	.953	58	1.0	59	.992	60	.998
51	*.625*	*52*	*1.0*	*53*	*1.0*	*54*	*1.0*	*55*	*.714*	*56*	*1.0*	*57*	*1.0*	*58*	*1.0*				
61	.992	62	.997	63	.993	64	.983	65	.991	66	.989	67	.994	68	1.0				

The recognition result is largely determined by the training samples. Figure 1 (Left) shows how the F-measure is influenced by training samples for object C10. As a trend, the F-measure increases with the number of training samples. In our model, the F-measure is determined by the irreducible graphs. For objects A1 to A50, the average number of irreducible graphs are about 30% of the initial number of graphs. This means that 1 in 3 of the initial graphs are selected. This verifies the conclusion that an initial model view can be used to match and identify views of the object over a range of rotations of at least 20 degrees in any direction [9]. As an example, the F-measure of object B1 is low. The reasons include 1) B1 is an apple who lacks suitable texture; 2) Several of the test samples are largely occluded; 3) Both the test set (11 samples) and the training set (29 samples) are too small. Hence, for objects B1 to B8, the minimum number of graphs is only 16 and the maximum number of graphs is only 29, while the view variation is very large. Hence basically all graphs are irreducible. If the training and test graphs are obtained under similar imaging conditions (including the camera viewpoint), though the training set is small, we might still obtain good test result. If not, the F-measure tends to be low as shown in Figure 1 (Left).

For objects with complex appearance, we need a large training sample set. For each of the objects in C1 to C9, the number of irreducible graphs is comparatively large. However, for an increasing number of graphs, the number of new irreducible graphs, incrementally learned using our method, will decrease. As an example, only about 5% of the graphs are selected as irreducible graphs for object C10, shown in Figure 1 (Right). Using the redundant graphs, the node co-occurrence probability can be determined and used to locate the the most typical representative sub-graph (i.e. the median) of each irreducible graph. Hence each irreducible graph acts as a central prototype and a test graph, coming from the i.i.d. of an irreducible graph, is likely to be correctly recognized.

For a trained model $CSGP_k$ of object k, we define $\max_{G_q \in CSGP_k} R(G_l, G_q)$ as the similarity of a graph instance G_l to that model. For the above test set, as an example, the probability distribution (histogram) of the similarity of the positive set (the magenta part) and the negative set (the blue part) for object C10 are shown in Figure 1 (Right). This clearly illustrates the distinctiveness between negative and positive instances using our generated CSGP model. Hence we obtain high recognition performance for Object 68, as shown in Table 1.
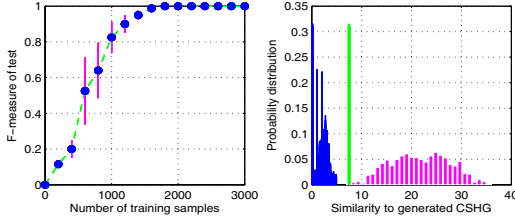

Fig. 1. Test result of object C10. F-measure of recognition test varies with the number of training samples is shown in the left sub-figure. For each number, we randomly select 10 different combinations and obtain the corresponding F-measures. The average of these F-measures is marked as $'\bullet'$, and the standard deviation of F-measure is marked as a vertical magenta line. Similarity distributions of negative and positive test set are shown in right sub-figure. The green line corresponds to the learned threshold R_{τ_1} of object C10.

5   Conclusion

This paper describes a framework for generating structural models from a large corpus of multi-view images. Our model is a comprehensive integration of the global and local information contained in local features from different views. The model is totally data driven. Using our method the precision of recognition can always be kept high while the recall can be improved with incremental training. Directions for future work include exploring the use of the method for image indexing and annotation. The most significant limitation of our approach is that it is best suited for objects that have some texture, much like the other local invariant feature of appearance based recognition schemes [6][12]. Hence we will combine some contour or edge based local invariant feature extractors for further research. We leave such problems as future work.

References

1. Bonev, B., Escolano, F., Lozano, M.A., Suau, P., Cazorla, M.A., Aguilar, W.: Constellations and the unsupervised learning of graphs. GbRPR 14(1), 340–350 (2007)
2. Chung, F.: Spectral graph theory. American Mathematical Society, Providence (1997)

3. Crandall, D.J., Huttenlocher, D.P.: Weakly supervised learning of part-based spatial models for visual object recognition. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 16–29. Springer, Heidelberg (2006)
4. Everingham, M., Gool, L.V., Williams, C., Winn, J., Zisserman, A.: Overview and results of classification challenge, 2007. In: The PASCAL VOC 2007 Challenge Workshop, in conj. with ICCV (2007)
5. Fawcett, T.: Roc graphs: Notes and practical considerations for researchers (2004), http://citeseer.ist.psu.edu/fawcett04roc.html
6. Ferrari, V., Tuytelaara, T., Van-Gool, L.: Simultaneous object recognition and segmentation from single or multiple model views. IJCV 67(2), 159–188 (2006)
7. Jiang, X., Munger, A., Bunke, H.: On median graphs: properties, algorithms, and applications. PAMI 23(10), 1144–1151 (2001)
8. Li, F.F., Perona, P.: A bayesian hierarchical model for learning natural scene categories. CVPR 2(2), 524–531 (2005)
9. Lowe, D.: Local feature view clustering for 3d object recognition. CVPR 2(1), 1682–1688 (2001)
10. Lowe, D.: Distinctive image features from scale-invariant key points. IJCV 60(2), 91–110 (2004)
11. Nister, D., Stewenius, H.: Scalable recognition with a vocabulary tree. CVPR 2(2), 2161–2168 (2006)
12. Rothganger, F., Lazebnik, S., Schmid, C., Ponce, J.: 3d object modeling and recognition using local affine-invariant image descriptors and multi-view spatial constraints. IJCV 66(3), 231–259 (2006)
13. Schonemann, P.: A generalized solution of the orthogonal procrustes problem. Psychometrika 31(3), 1–10 (1966)
14. Sivic, J., Russell, B.C., Efros, A.A., Zisserman, A., Freeman, W.T.: Discovering objects and their location in images. ICCV 1(1), 872–877 (2005)
15. Todorovic, S., Ahuja, N.: Unsupervised category modeling, recognition and segmentation in images. PAMI (in press)
16. Torralba, A., Fergus, R., Weiss, Y.: Small codes and large image databases for recognition. In: CVPR (2008)
17. Torsello, A., Hancock, E.: Learning shape-classes using a mixture of tree-unions. PAMI 28(6), 954–967 (2006)
18. Xia, S., Hancock, E.R.: 3d object recognition using hyper-graphs and ranked local invariant features. In: da Vitoria Lobo, N., Kasparis, T., Roli, F., Kwok, J.T., Georgiopoulos, M., Anagnostopoulos, G.C., Loog, M. (eds.) S+SSPR 2008. LNCS, vol. 5342, pp. 117–126. Springer, Heidelberg (2008)
19. Xia, S., Hancock, E.R.: Clustering using class specific hyper graphs. In: da Vitoria Lobo, N., Kasparis, T., Roli, F., Kwok, J.T., Georgiopoulos, M., Anagnostopoulos, G.C., Loog, M. (eds.) S+SSPR 2008. LNCS, vol. 5342, pp. 318–328. Springer, Heidelberg (2008)
20. Xia, S.P., Liu, J.J., Yuan, Z.T., Yu, H., Zhang, L.F., Yu, W.X.: Cluster-computer based incremental and distributed rsom data-clustering. ACTA Electronica sinica 35(3), 385–391 (2007)
21. Xia, S.P., Ren, P., Hancock, E.R.: Ranking the local invariant features for the robust visual saliencies. In: ICPR 2008 (2008)