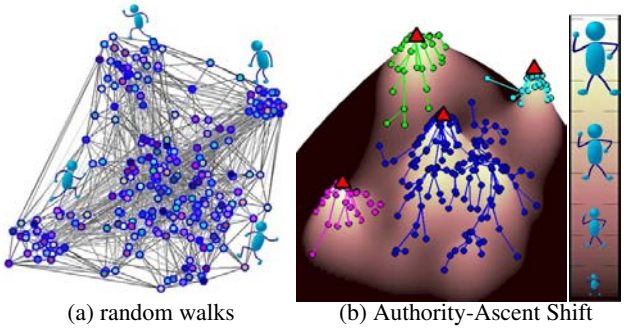
Mode-Seeking on Graphs via Random Walks

Minsu Cho and Kyoung Mu Lee
Dept. of EECS, ASRI, Seoul National University, Korea
chominsu@gmail.com kyoungmu@snu.ac.kr

Abstract

Mode-seeking has been widely used as a powerful data analysis technique for clustering and filtering in a metric feature space. We introduce a versatile and efficient mode-seeking method for "graph" representation where general embedding of relational data is possible beyond metric spaces. Exploiting the global structure of the graph by random walks, our method intrinsically combines mode-seeking with ranking on the graph, and performs robust analysis by seeking high-ranked authoritative data and suppressing low-ranked noise and outliers. This enables mode-seeking to be applied to a large class of challenging real-world problems involving graph representation which frequently arises in computer vision. We demonstrate our method on various synthetic experiments and real applications dealing with noisy and complex data such as scene summarization and object-based image matching.



(a) random walks on a weighted graph (b) Authority-Ascent Shift on the random walk density

Figure 1. Authority-Ascent Shift. Based on the visiting probability of *random walks* on a *graph* of (a), AAS shifts each node to the most frequently visited node among local neighbors on the graph. In (b), the height of each node shows its visiting probability of random walks, called the *authority score*, and the edges represent shifting trajectories of our AAS. Each node converges to its *authority mode* (red peak) and the corresponding clusters emerge.

1. Introduction

Many computer vision tasks involve the extraction of *meaningful* information from real-world images where data arise from a multi-modal distribution with noise and outliers. For those, mode-seeking has been widely used as a powerful data analysis technique for clustering and filtering [5, 26, 25, 23, 22]. The basic idea is to associate each point through a shifting scheme, e.g. shifting to its local centroid in the case of mean shift [5], with a mode of the underlying density function in a feature space. This approach has appealing advantages over other traditional clustering techniques; the structure of the clusters may be rather arbitrary and the number of clusters does not need to be known in advance. Furthermore, clustering by mode-seeking is not severely hindered by noise and even a large amount of outliers, and can effectively filter them out [26]. While recently extended to non-linear manifolds [25], general metric spaces [23], and tree-structured distance spaces [27], mode-seeking is still restricted to traditional data representation as points in a metric feature space.

In this paper, we propose a mode-seeking method for *graph* representation as illustrated in Fig. 1. A graph or a network describes more general relational data than the notion of *features* and provides versatile representation for analysis of non-metric and even directional data. In many applications, given data are not naturally representable in terms of a feature vector, whereas quite often we can readily encode them in a weighted graph (e.g. structural representation of object shape [7], topological relation of visual elements [3], co-occurrence of object entities [12]). Accordingly, graph-based analysis has gained increasing popularity in computer vision and machine learning.

Since graph representation lacks explicit geometric notions such as a centroid or distance as in a metric feature space, mode-seeking on graphs requires a novel approach based on a graph theory. We leverage a graph-theoretical approach of *random walks*, which has been successfully used to analyze complex graphs such as the Web and social networks [20, 11, 6]. Simulating a random walker on a graph, we define *modes* on the *graph*, called *authority modes*, as local maxima of the stationary distribution of "visits". This notion of authority modes naturally relates to centrality analysis [6] and PageRank [20], and mutually cor-

978-1-4673-1228-8/12/$31.00 ©2012 IEEE

responds to the conventional notion of modes [10, 5]. Based on this concept, we propose a nonparametric estimator of the authority modes, the *Authority-Ascent Shift* (AAS). As illustrated in Fig. 1, it is formulated as a mode-seeking procedure that makes the shifts along the steepest ascent direction toward authority modes.

Recently, there have been a few attempts to implement mode-seeking in the graph domain [4, 15, 9]. Jouili *et al*. [9] proposed the median graph shift for clustering a set of graphs, extending the medoid shift [23] based on the concept of the median graph. Liu and Yan [15] presented the graph shift algorithm for dense subgraph detection which alternates graph expansion and contraction. In these two methods, a mode is determined as a graph among a given set of graphs [9] or as a subgraph among the power set of a given graph [15], and their mode-seeking is done by shifting from one (sub)graph to another (sub)graph. On the other hand, Cho and Lee [4] introduced a node-shifting scheme, called Hierarchical Authority Shift (HAS), which iteratively shifts from one node to another node using the personalized PageRank (PPR) matrix of a given graph. Their mode-seeking naturally leads to graph partitioning unlike the other approaches [9, 15]. However, it demands the full PPR matrix of the given graph and its shifting scheme of iterative PPR propagations has an adverse effect of accumulating errors on noisy data. Our AAS tackles these issues by taking different approaches both in defining a mode and in shifting across nodes. Unlike HAS [4], it directly exploits and utilizes the global random walk density distribution in defining the *modes*. It not only enables to suppress contaminative shifting of outliers but also leads to an computatinally an efficient shifting scheme. Furthermore, in this way, AAS frames a graph-based analogue to traditional mode-seeking, relating 'a graph' to 'features in a metric space', 'a node' to 'a feature', 'weights' to 'distances', 'random walk density distribution' to 'kernel density distribution', and so on.

AAS has three principal advantages over previous methods. First, it allows the idea of mode-seeking to be applied to various problems involving general graph representation, where the conventional mode-seeking algorithms [2, 5, 25] are inapplicable. Second, AAS is intrinsically combined with graph ranking [6, 20], effectively exploiting the global structure by random walks, thus performs robust data analysis by seeking high-ranked authoritative data and suppressing low-ranked noise and outlier data. Third, it is computationally efficient and easily scalable to large graphs in real-world problems. We will demonstrate by experimental evaluations and comparisons that AAS is a powerful tool for vision problems dealing with a large amount of noisy data.

2. Graph and its Authority Modes

A relational graph $G = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ consists of nodes $\mathcal{V}$, edges $\mathcal{E}$ and weights $\mathcal{W}$. Weight $w(i,j) \in \mathcal{W}$ is associated with edge $e(i,j) \in \mathcal{E}$ from node i to node j, characterizing similarity or proximity from i to j. If $w(i,j) = w(j,i)$ is satisfied for all weights, the graph is undirected, and otherwise, it is a directed graph. In this paper, N denotes the size of graph G, that is the number of nodes $\mathcal{V}$.

2.1. Authority Modes

The notion of *modes on a graph* is naturally conceptualized by random walks on the graph. A random walk is a stochastic process widely used for graph analysis [6] and Webpage ranking [20]. Suppose a random walker travelling on the given graph G, whose probability of a transition from node i to node j is

$$p(i,j) = w(i,j)/d(i) \tag{1}$$

where $d(i) = \sum_{j \in \mathcal{V}} w(i,j)$ is the out-degree of node i. Hence, its associated Markov transition matrix $\boldsymbol{P}$ is the $N \times N$ matrix whose $(i,j)_{th}$ entry is simply

$$\boldsymbol{P}_{ij} = \begin{cases} p(i,j) & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Let $x_k(i)$ be the probability of node i being visited by the random walker at k_{th} step, then the probability of node j at the next step is formulated by

$$x_{k+1}(j) = \sum_{i \in \mathcal{V}} x_k(i) p(i,j). \tag{3}$$

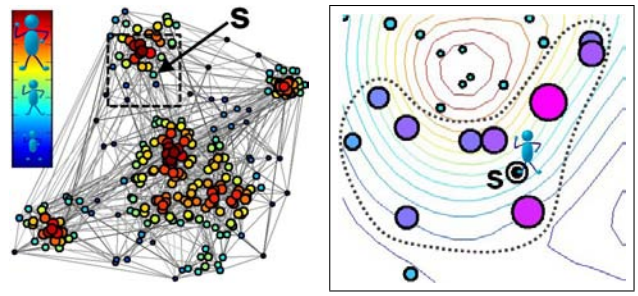
In this model, what happens to $x_k(i)$ when $k \to \infty$? If it converges to a unique stationary distribution π invariant to initial distribution x_0, we can assign $\pi(j)$ to each node j as a time-invariant probabilistic measure that characterizes the probability of visits by the random walker on the graph. Using this model, we define the *authority mode on a graph* as follows.

Definition 1 *Authority modes on a graph are the most frequently visited nodes among their local neighbors by random walks on the graph.*

The most frequently visited nodes in local neighbors correspond to local maxima of underlying probability distribution of random walks over the graph. This definition naturally captures the conventional concept of modes, which is defined as the local maxima of the underlying probability density distribution over a feature space [10, 5].

The probability of visits by random walks is related to the PageRank originally developed by Brin and Page for Web page ranking [20]. In PageRank, random jumps (teleports) are introduced to guarantee convergence to a unique stationary distribution. Adopting this, we call the probability of visits on the graph as the *authority score* π, and define it as the solution to

$$\pi(j) = \alpha \sum_{i \in \mathcal{V}} \pi(i) p(i,j) + (1-\alpha)\frac{1}{N}, \tag{4}$$



(a) authority scores π (b) node relevancy $\Psi(s,\cdot)$

(c) authority increment expectation $p_k(s,\cdot)\nabla_\pi(s,\cdot)$ (d) authority-ascent $\mathcal{A}(s)$ in local neighbors $\mathcal{N}_\varepsilon(s)$

Figure 2. Authority-ascent as the steepest ascent over the authority distribution. In (a), the authority score π is computed by random walks. Based on this, each node s has relevancy $\Psi(s,\cdot)$ as shown in (b) and obtains expectation of authority increment $p_k(s,\cdot)\nabla_\pi(s,\cdot)$ as described in (c). The bigger cicle means the larger absolute value, and $+,-$ denote the signs of the value. Here, the contour lines are used to describe the authority score π and the dotted line delineates local neighbors $\mathcal{N}_\varepsilon(s)$ with a sufficient relevancy. In (d), the authority-ascent of s is determined as the node with the highest expectation of authority increment among its local neighbors. For details, see the text.

where $\alpha \in (0,1)$ denotes the decay factor that represents the probability with which the walker follows edges, while $1-\alpha$ is the probability of jumping to a random node[1]. Usually, α is set between 0.8 and 0.95 in the literature [20, 11]. In this work, we set $\alpha = 0.9$. The authority score π refers to the stationary distribution of the random walk governed by the transition matrix $\alpha\boldsymbol{P} + (1-\alpha)\boldsymbol{e}\boldsymbol{e}^T$ where $\boldsymbol{e}$ is a column of ones. It can be efficiently computed by power iterations or a linear system formulation [11]. Figure 2a represents π through the color and size of each node.

2.2. Local Neighbors by Node Relevancy

The fact that graphs represent general relational data even with directions implies that graphs deviate from explicit geometric notions used in a metric space and simple distance-based neighbors are hard to be applied on them. Thus, the term "*local neighbors*" of Def. 1 needs to be formally described to specify the *authority modes* on a graph.

[1]When G has dangling nodes without any out-going edge, additional edges are constructed. See [11].

To adaptively decide neighbors based on random walks, we establish three criteria: (1) *Reachability*: a random walker from a node has a high probability to reach its neighbors. (2) *Supportability*: the more out-degree a node has, the more neighbors it obtains. This reflects the supports of out-going edge weights in the original graph. (3) *Authority smoothness*: the neighbors of a node should have the similar authority scores to the node. This derives from the fact that relevant nodes are frequently visited from each other by a random walker so that they are likely to have a similar random walk density. Exploiting the global behavior of random walks, this smoothness prevents the scope of neighbors from extending into irrelevant nodes.

The reachability is measured by the probability $p_k(s,t)$ that a random walker originating at node s will visit node t at its k_{th} step. It corresponds to the $(s,t)_{th}$ entry of the k_{th} power of $\boldsymbol{P}$, that is $\boldsymbol{P}^k_{st}$. The supportability of node s amounts to the out-degree $d(s)$. The authority smoothness between node s and t is formulated using the difference of authority scores $|\pi(t) - \pi(s)|$.

We define the *authority increment* $\nabla_\pi(s,t)$ from node s to node t by

$$\nabla_\pi(s,t) = (\pi(t) - \pi(s)) \tag{5}$$

where each authority score is multiplied by N to have the unit mean value. Then, based on the three criterion, the *node relevancy* Ψ from node s to node t is formulated as

$$\Psi(s,t) = d(s)p_k(s,t)\exp(-\gamma\nabla_\pi(s,t)^2) \tag{6}$$

where parameter γ controls the authority smoothness term. $\Psi(s,t) \neq \Psi(t,s)$ and $\Psi(s,t)$ depends on the random walk step k. The first step random walk of $k=1$ is usually used so that $\Psi(s,t) = w(s,t)\exp(-\gamma\nabla_\pi(s,t)^2)$. Figure 2b illustrates the relevancy distribution of node s to other nodes. Finally, adopting the ε-neighborhood on the node relevancy, the set of "*local neighbors*" of node s, $\mathcal{N}_\varepsilon(s)$, is defined as

$$\mathcal{N}_\varepsilon(s) = \{t \in \mathcal{V} \mid \Psi(s,t) > \varepsilon\} \cup s \tag{7}$$

As delineated by the dotted line in Fig. 2b, local neighbors $\mathcal{N}_\varepsilon$ are the nodes with a sufficient relevancy ε and determines the possible shifting scope.

3. Authority-Ascent Shift

The notions in the previous section pave the way to Authority-Ascent Shift. AAS can be explained by the mode-seeking procedure along the *authority-ascent* direction. The authority-ascent $\mathcal{A}(s)$ of node s is formulated as

$$\mathcal{A}(s) = \arg\max_{t \in \mathcal{N}_\varepsilon(s)} p_k(s,t)\nabla_\pi(s,t) \tag{8}$$

which means the neighboring node of s with the highest expectation of authority increment by a random walk. This authority-ascent is the steepest ascent over the authority

score distribution on the graph. It parallels the fact that the standard mode-seeking direction of mean shift is equivalent to the steepest ascent on a underlying density function over a feature space [2, 5]. Figure 2c-d illustrates the authority increment expectation $p_k(s,\cdot)\nabla_\pi(s,\cdot)$ and authority-ascent $\mathcal{A}(s)$, respectively. As shown in Fig. 2d, the authority-ascent is bounded by local neighbors $\mathcal{N}_\varepsilon$ with sufficient node relevancy ε.

Theorem 3.1 *For any starting node, the sequence of successive authority-ascent shifts is finite and converges to an authority mode.*

Proof Local neighbors of any node s includes s itself, which has no authority increment. Thus, the authority scores of the successive shifts keep strictly increasing until it reaches a node whose authority-ascent is itself. Therefore, the length of the sequence is at most N, and the final node has the highest authority score among its local neighbors.

Therefore, starting from a node, successive shifts to authority-ascents progress toward its authority mode. As shown in Fig. 1b, the shifting trajectory of nodes sharing a common authority mode builds a tree, and the clusters corresponding to the modes emerge as disjoint trees of shifting trajectories. Once an authority-ascent has been computed for each node on the graph, the next authority-ascent already exists for any node. It makes AAS computationally efficient so that a single tree traversal can efficiently assign the cluster label to all nodes associated with each tree. Thus, unlike mean updates in mean shift [2, 5], authority-ascent updates $\mathcal{A}$ need to be computed only once per node and no threshold is required to specify the terminating condition. In that sense, AAS has a similar property to the medoid shift [23], but does not require iterative processes to complete the mode-seeking unlike the medoid shift.

Note that by AAS, each cluster has "*cluster authority*" as the sum of the authority scores of its members. It means the probability of a random walker to travel inside the cluster. As an extended concept of authority [20] to a related group, the cluster authority provides a reliable measure for the *inlierness* of data. Noise and outlier data can be easily detected and eliminated using the cluster authority of AAS as demonstrated in our experiments. The AAS algorithm is summarized in Algorithm 1.

AAS Kernel AAS is controlled by random walk step k, authority smoothness factor γ, and relevancy ε. AAS usually uses the first step random walk step of $k = 1$. The further step $k \geq 2$ can be adopted for long-range interaction on the graph as shown in the experiment in Sec. 4.1. Smoothness factor γ of Eq. (6) relates to the authority smoothness inside each emerging cluster. Increasing γ leads to gentle authority-ascents so that each emerging cluster has a smooth authority distribution of its members. We found that $\gamma = 20$ is adequate for most cases, and fixed it in this work. Relevancy ε controls the scope of local neighbors, and determines the granularity of emerging clusters. It is the main control parameter and has similar effect to the size of kernel in conventional mode-seeking.

Algorithm 1: Authority-Ascent Shift

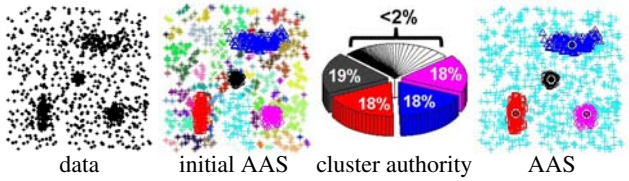
Input: Weight W, relevancy ε
Output: authority mode labels for each node, and cluster authority for each authority mode
Assign transition $p(i,j)$ and out-degree $d(i)$
Compute authority score $\pi(i)$ by solving Eq. (4)
foreach *Node* s **do**
  Find neighbors $\mathcal{N}_\varepsilon(s)$ by Eq. (7)
  $\mathcal{A}(s) = \arg\max_{t\in\mathcal{N}_\varepsilon(s)} p_k(s,t)\nabla_\pi(s,t)$
end
Associate each node s with its authority mode by tree traversals along $\mathcal{A}(s)$, and assign the sum of authority scores to their mode as the cluster authority.

Complexity Issue The computational complexity of AAS is mainly governed by that of the power iterations for computing authority score π. In the case of a fully-connected graph with N nodes, each iteration takes N^2 multiplications. Because the number of iterations needed to achieve a fixed accuracy is independent of the graph [1], the total complexity approximately amounts to $O(N^2)$. In the case of sparse graphs, however, the complexity decreases much further; it is observed in many real applications of the link analysis techniques [1] that the complexity of the power iteration of sparse matrices grows roughly linearly with N. Furthermore, many speed-up techniques motivated by Web networks also enable scaling to large graphs [11].

4. Experiments and Applications

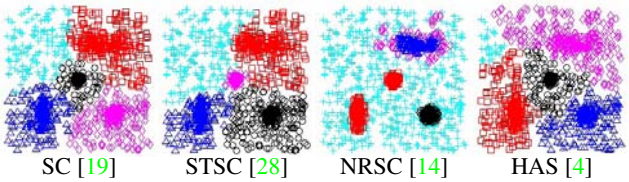
4.1. Synthetic Experiments

To compare the characteristics of AAS with conventional mode-seeking and graph-based clustering algorithms, we produced experiments on synthetic point datasets. To construct the graph for graph-based algorithms including AAS, a standard manner [19] was used as follows. The given points were normalized to occupy the $[-1,1]^2$ space and the graph with edge weight $w(i,j)$ is constructed by $w(i,j) = exp(\frac{-d(i,j)^2}{\sigma^2})$ for $i \neq j$ and $w(i,i) = 0$. $d(i,j)$ denotes the Euclidian distance between point i and j. σ is a scale parameter that we fixed $\sigma^2 = 0.1$. For performance evaluation, the normalized mutual information (NMI) [21], a common metric for clustering accuracy, was adopted. The larger value of NMI $\in [0,1]$ means the better result. In the synthetic experiments, we set $\varepsilon = 0.001$ for AAS.



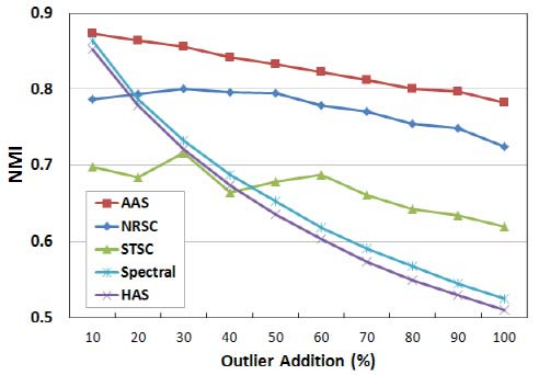
data initial AAS cluster authority AAS

(a) AAS against outliers using cluster authority



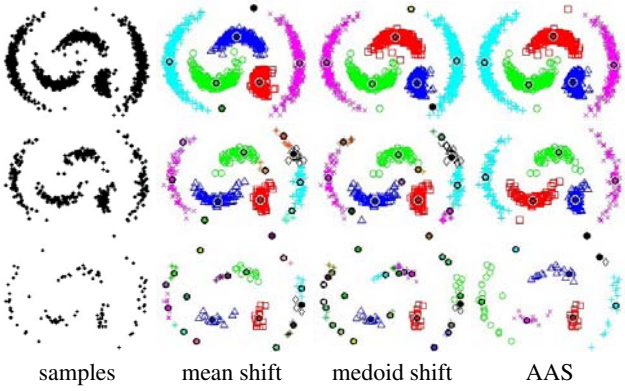
SC [19] STSC [28] NRSC [14] HAS [4]

(b) other graph-based clustering methods



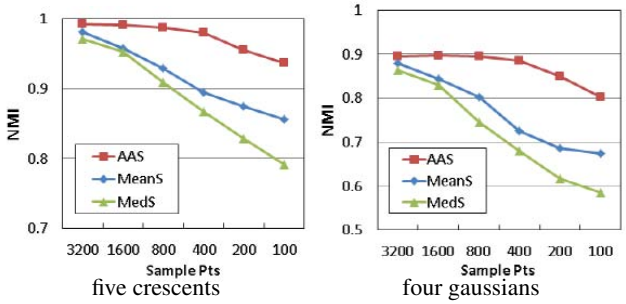
(c) comparative experiments varying the outlier ratios

Figure 3. Clustering comparison on noisy data. In (a), based on *cluster authority* values, AAS distinguishes inlier clusters from outlier ones. Other graph-based clustering methods are compared in (b), and the average NMI results varying the amount of outliers are plotted in (c) (x-axis: # of outliers / # of inliers).



samples mean shift medoid shift AAS

(a) 'five crescents' examples with decreasing sample points



(b) average NMI varying the number of sample points

Figure 4. Mode-seeking comparison on different sample sizes. The example results of 'five crescents' with 1600, 400, 100 pts are shown in (a), where each mode is represented by the thick dot in the cluster. The average NMI results gradually decreasing sampling density from 3200 pts to 100 pts are plotted in (b).

Comparison with Graph-based Clustering Methods Real-world data often contains noisy outliers irrelevant to underlying data sources, and those have posed challenges to the data analysis and clustering community [28, 14]. Most existing graph-based algorithms fail in these problems since the real data distribution is distorted by noise. On the contrary, AAS robustly detects clusters against outliers by seeking high ranked authority modes and suppressing low ranked outlier data. To demonstrate this, we sampled points from four gaussian distributions with different variances, and added uniformly distributed random outliers as shown in Fig. 3a. Applying AAS on the data, inlier elements find their authority modes and form large groups, while outlier elements are isolated in small groups since they do not have enough relevancy with inliers (2nd in Fig. 3a). Thus, simply using *cluster authority* values (3rd in Fig. 3a), we can distinguish inlier clusters from outlier ones (4th in Fig. 3a). In all experiments, we consider the inlier clusters have cluster authority values more than 5% of the total authority. In the example of Fig. 3a, all cluster authorities of outlier clusters are less than 2%. To compare the sensitivities of different algorithms against outliers systematically, we formed a series of noisy datasets by increasing outliers ratio with respect to the number of inliers, and compared ours with Spectral Clustering (SC) [19], Self-Tuning SC (STSC) [28], Noise Robust SC (NRSC) [14], and HAS [4]. Each level of outliers was tested 100 times, and average NMI results are plotted in Fig. 3c. For SC and HAS, we provided the number of clusters as 5 (4 inlier clusters + 1 outlier cluster) STSC and NRSC automatically estimates the number of clusters.

As observed in Fig. 3b-c, SC and HAS are very sensitive to outliers. Although NRSC and STSC improve clustering on highly noisy data, they have adverse effects on data with less outliers. On the other hand, AAS clearly outperforms all the other methods. Especially, note that AAS and HAS corresponds to two extremes in this outlier experiment. While AAS shifting based on node relevancy effectively suppresses merges between inliers and outliers, HAS shifting by the PPR propagations [4] tends to propagate the influences of outliers. This difference can be easily observed in the comparative examples of Fig. 3a-b.

Comparison with Feature Space Mode-Seeking Since AAS is based on the global behavior of random walks, it provides more accurate and robust seeking of the underly-

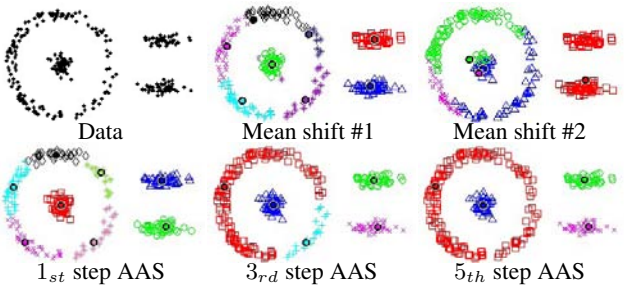


Figure 5. Mode-seeking on manifold data. AAS with 5_{th}-step random walks ($k = 5$) detects a ring-shape distribution as a mode, whereas mean shift fails with any kernel size.

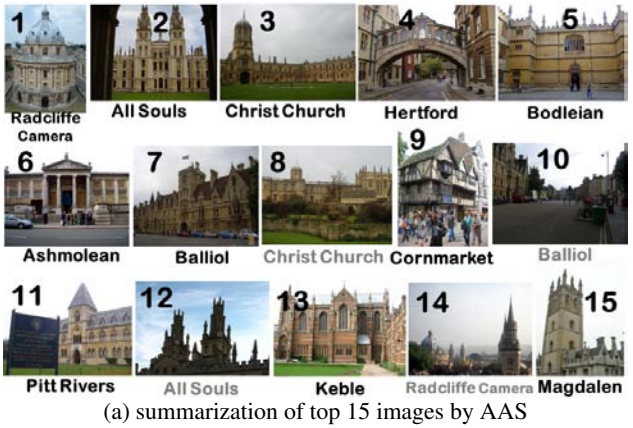
ing modes than the conventional methods based on local kernel density estimation. To demonstrate this, we compare AAS with the mean shift [5] and medoid shift [23] on two datasets as shown in Fig. 4: (1) five crescent shaped distributions, not linearly separable, also tested in [23, 22], and (2) four gaussian distributions with different densities, distributed as shown in Fig. 1a. While the experiments in the previous mode-seeking papers [5, 23, 22] only dealt with data having dense and sufficient samples, we performed extensive tests gradually decreasing sampling density from 3200 pts to 100 pts. The example results of 'five crescents' with 1600, 400, 100 pts are shown in Fig. 4a. Each point set is generated 100 times for each density, and average NMI accuracy is computed for each method. The results are plotted in Fig. 4b. It shows that AAS consistently outperforms two mode-seeking methods in seeking the underlying modes even in sparse data, whereas both mean shift and medoid shift quickly deteriorate from 1600 pts. In this experiment, the kernel sizes of both mean shift and medoid shift are already finely tuned, thus varying the kernel sizes does not improve the performance.

Another advantage of AAS over conventional mode-seeking is that it can perform mode-seeking with long-range interaction by increasing random walk step k as stated in Sec. 3. Comparative examples in Fig. 5 show the effect of step k on the data spanned along a manifold.

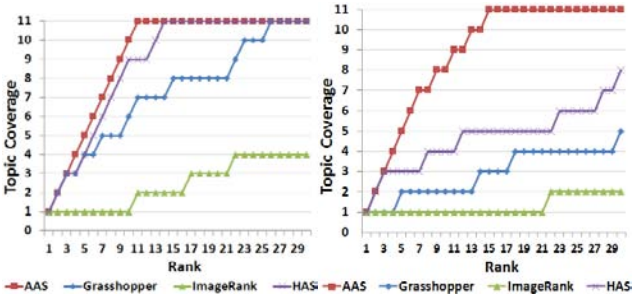
4.2. Scene Summarization

Finding multiple themes and their relative scores in large and noisy data is the goal of data ranking and summarization tasks [24, 8, 29], and it easily relates to seeking the authority modes with their cluster authorities. Hence, we applied AAS to the scene summarization problem, a data summarization task in computer vision [24]. The goal is to select a small set of canonical views from given images, which has representativeness as well as diversity. We adopted the Oxford buildings dataset[2] including 5062 images of 11 different Oxford landmarks together with a large amount

[2] http://www.robots.ox.ac.uk/~vgg/data/



(a) summarization of top 15 images by AAS



(b) coverage on 'Good' images (c) coverage on all images

Figure 6. Scene summarization on the Oxford building dataset. The result of AAS visualized in (a) covers all the high-quality images of 11 landmarks and their major views. The performance is compared with other methods in (b) and (c). See the text.

of distracting outliers. Each image has manually labeled ground truth for its landmark and its visual quality (*Good*, *OK*, *Junk*, or *Absent*). We built up a noisy dataset of total 1145 images including all of the *Good*, *OK*, *Junk* images (845 images) and randomly chosen *Absent* images (300 images). Then, we constructed a graph by treating images as nodes. To define the edge weight $w(i, j)$ between image i and j, we performed feature matching using the Hessian affine detector [18] and the SIFT descriptor [16], and further prune the set of candidates by RANSAC. Then, the number of the remained matches was set to $w(i, j)$. Note that unlike the conventional mode-seeking as in [23], AAS does not require any consideration to endow the similarity of an image pair with a metric structure. We compared ours with data ranking and summarization of ImageRank [8] and Grasshopper [29], and hierarchical clustering of HAS [4].

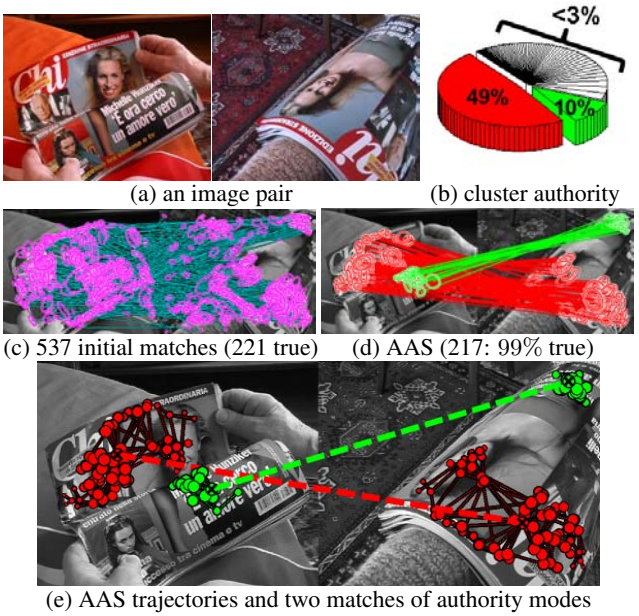
Figure 6a shows a summary of 15 images by AAS, which correspond to the top 15 authority modes with the highest cluster authorities. As shown by its visual quality and ground truth landmark below each image, AAS gives an excellent summary of the dataset. The result includes *Good* images of all 11 landmarks and their different major views. To quantify the performance, we counted the number of landmarks covered by *Good* images among the sum-

mary of top k images with increasing k. Figure 6c represents comparative results of all algorithms using the same graph. To examine the influences of noise and outliers on each method, we did the same experiment on another graph constructed using only all the *Good* images (257 images), which means relatively clean data without outliers and significant noise, and plotted the results in Fig. 6b. For both experiments, AAS robustly estimates ranking by cluster authority and effectively suppress its redundant members under the authority modes. It clearly outperforms all the other methods especially on noisy data. This result was obtained with $\varepsilon = 15$, but any $\varepsilon \in [5, 30]$ also gave similar coverage results with different granularity. The top rank images of ImageRank [8] are dominated by a few most popular landmarks and therefore lack diversity. While Grasshopper [29] improves the diversity by suppressing similar images in its sequential re-ranking, the redundant views still appears in top ranks due to the inaccurate suppression. HAS [4] provides a reliable summary on clean data reflecting the data structures by its hierarchical clustering, whereas it is drastically disturbed by noise and outliers as also observed in the previous synthetic experiments.

4.3. Feature Matching and Clustering

In this experiment we applied AAS to the real image matching problems where multiple deformable objects appear with significant clutter. For this, we construct an association graph by treating candidate matches as nodes and their similarities as edge weights [13, 3]. Here, (i, a) denotes the candidate match from feature i in the first image to feature a in the second image, and corresponds to a node in the association graph. The candidate matches are generated using the Harris affine [18] and the MSER detectors [17] with the SIFT descriptor [16]. In order to measure dissimilarity between two matches (i, a) and (j, b), we adopted the symmetric transfer error $d(ia, jb)$ used in [3], which is adaptive to affine deformation. The association graph has weight $w(ia, jb) = \max(100 - 0.05 \times d(ia, jb)^2, 0)$. In this setting, we used $\varepsilon = 1.0$ for AAS.

Figure 7 illustrates feature matching and clustering by AAS. The ground truths are manually labeled for all candidate matches. Despite deformation and outliers, it successfully detects two inlier match clusters with 99% accuracy. In AAS matching, the matches of authority modes (dashed lines) represents the most principal matches, while the matches with low authority scores (small features) implies deformation or edges of the object. In Fig. 8a, we applied AAS matching on the image pairs used in [3]. It performs comparable to or better than the method of [3] (ACC). Figure 8b demonstrates a comparative experiment on a challenging matching problem with repetitive patterns. To solve this, it is required to consider global structure of patterns with one-to-one matching constraints between features. As shown in the Fig. 8b, the bottom-up matching methods like ACC [3] fails to consider the global structure of the patterns, and graph matching methods like Spectral Matching (SM) [13] cannot distinguish inlier match clusters from outliers. AAS matching can solves this problem effectively by reflecting the matching constraints in each shift. The constrained shifting procedure of AAS successfully detects true matches and distingushes them from outliers. It largely outperforms ACC and SM in both precision and recall as detailed in Fig. 8b.



(a) an image pair (b) cluster authority

(c) 537 initial matches (221 true) (d) AAS (217: 99% true)

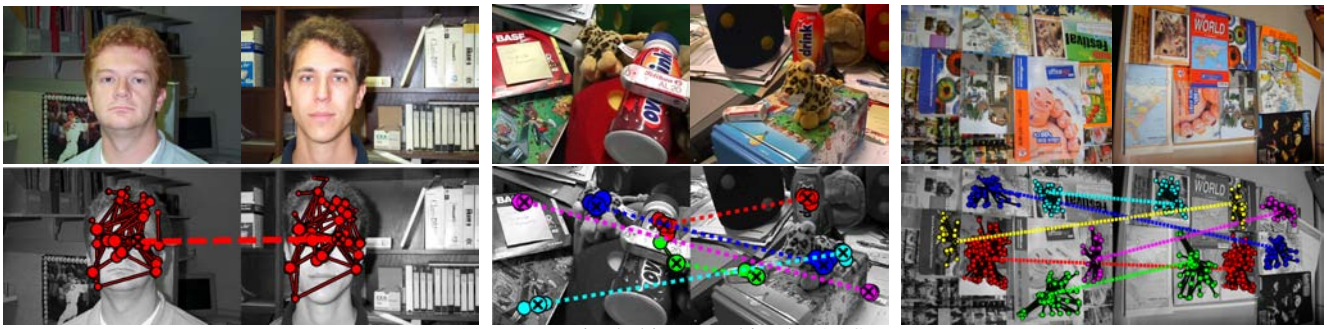
(e) AAS trajectories and two matches of authority modes

Figure 7. Feature matching and clustering by AAS. The association graph is constructed by the candidate matches in (c) as nodes. In (d), AAA finds multiple match clusters against significant outliers using cluster authority in (b). In (e), the size of matching features represents the authority score of each match, and the lines between features shows shifting trajectories. Two dashed lines show the matches corresponding to two authority modes.

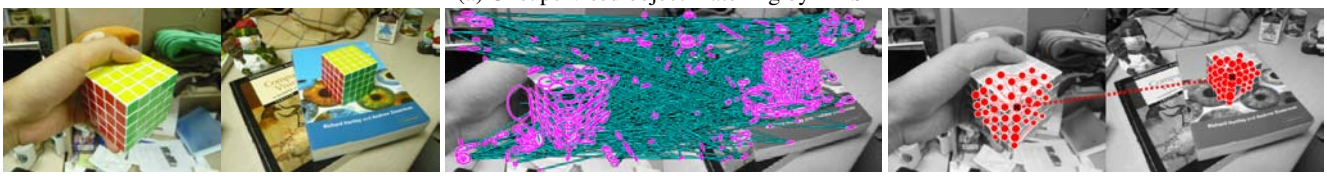
For more details and experiments, refer to our project site: http://cv.snu.ac.kr/research/~AAS/

5. Conclusion

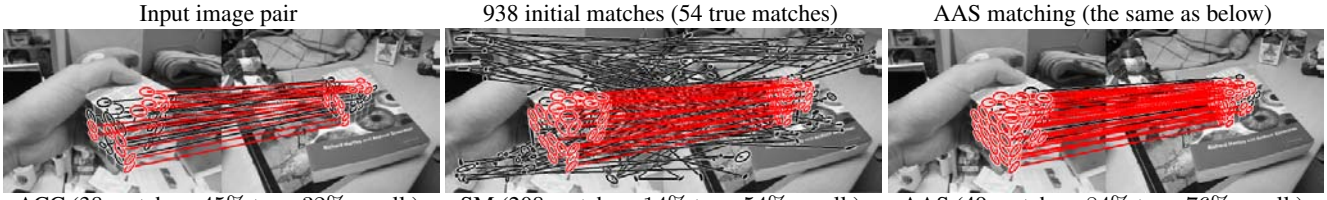
We proposed graph-based mode-seeking, called AAS, which extends mode-seeking to graph domains. Leveraging random walks on graphs, AAS leads to the integration of non-parametric clustering, ranking, and outlier elimination. Experiments demonstrate that it is adequate for challenging real-world problems where complex data arise with significant outliers. AAS is a versatile tool for general graphs not restricted to the applications shown in our experiments. We leave other applications as future work.



(a) Unsupervised object matching by AAS



Input image pair | 938 initial matches (54 true matches) | AAS matching (the same as below)



ACC (38 matches: 45% true, 32% recall) | SM (208 matches: 14% true, 54% recall) | AAS (49 matches: 84% true, 76% recall)

(b) A comparative experiment on an object with repetitive patterns

Figure 8. AAS matching results on various examples. In (a), all the common objects are perfectly detected by AAS matching. In the bottom of (b), each algorithm, the number of true matches, the precision (true matches/detected), and the recall (true matches/ground truths) are captioned. For details, see the text.

References

[1] P. Berkhin. A survey on pagerank computing. *IM*, 2005. 4

[2] Y. Cheng. Mean shift, mode seeking, and clustering. *PAMI*, 1995. 2, 4

[3] M. Cho, J. Lee, and K. M. Lee. Feature correspondence and deformable object matching via agglomerative correspondence clustering. *ICCV*, 2009. 1, 7

[4] M. Cho and K. M. Lee. Authority-shift clustering: Hierarchical clustering by authority seeking on graphs. *CVPR*, 2010. 2, 5, 6, 7

[5] D. Comanicu and P. Meer. Mean shift: A robust approach toward feature space analysis. *PAMI*, 2002. 1, 2, 4, 6

[6] R. Diestel. Graph theory. *Springer, New York, NY*, 2005. 1, 2

[7] P. F. Felzenszwalb and D. P. Huttenlocher. Pictorial structures for object recognition. *IJCV*, 2003. 1

[8] Y. Jing and S. Baluja. Pagerank for product image search. *WWW*, 2008. 6, 7

[9] S. Jouili, S. Tabbone, and V. Lacroix. Median graph shift: A new clustering algorithm for graph domain. *ICPR*, 2010. 2

[10] L. H. K. Fukunaga. The estimation of the gradient of a density function, with applications in pattern recognition. *IT*, 1975. 2

[11] A. N. Langville and C. D. Meyer. Deeper inside pagerank. *IM*, 2003. 1, 3, 4

[12] Y. J. Lee and K. Grauman. Object-graphs for context-aware category discovery. *CVPR*, 2010. 1

[13] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. *ICCV*, 2005. 7

[14] Z. Li, J. Liu, S. Chen, and X. Tang. Noise robust spectral clustering. *ICCV*, 2007. 5

[15] H. Liu and S. Yan. Robust graph mode seeking by graph shift. *ICML*, 2010. 2

[16] D. G. Lowe. Object recognition from local scale-invariant features. *ICCV*, 1999. 6, 7

[17] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. *BMVC*, 2002. 7

[18] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *IJCV*, 2004. 6, 7

[19] A. Ng, M. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. *NIPS*, 2001. 4, 5

[20] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. *TR, Stanford University*, 1998. 1, 2, 3, 4

[21] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. Numerical recipes in c. *Cambridge University Press*, 1988. 4

[22] L. Shapira, S. Avidan, and A. Shamir. Mode-detection via median-shift. *ICCV*, 2009. 1, 6

[23] Y. A. Sheikh, E. A. Khan, and T. Kanade. Mode-seeking by medoidshifts. *ICCV*, 2007. 1, 2, 4, 6

[24] I. Simon, N. Snavely, and S. M. Seitz. Scene summarization for online image collections. *ICCV*, 2007. 6

[25] R. Subbarao and P. Meer. Nonlinear mean shift for clustering over analytic manifolds. *CVPR*, 2006. 1, 2

[26] D. Suter and H. Wang. Robust fitting using mean shift: Applications in computer vision. *ICORS*, 2003. 1

[27] Z. Yu, O. Au, K. Tang, and C. Xu. Nonparametric density estimation on a graph: Learning framework, fast approximation and application in image segmentation. *CVPR*, 2011. 1

[28] L. Zelnik-Manor and P. Perona. Self-tuning spectral clustering. *NIPS*, 2004. 5

[29] X. Zhu, A. B. Goldberg, J. V. Gael, and D. Andrzejewski. Improving diversity in ranking using absorbing random walks. *NAACL-HLT*, 2007. 6, 7