Beyond Partitions: Allowing Overlapping Groups in Pairwise Clustering

Andrea Torsello, Samuel Rota Bulò, and Marcello Pelillo
Dipartimento di Informatica, Università Ca' Foscari di Venezia, Italy
{*torsello,srotabul,pelillo*}@*dsi.unive.it*

Abstract

The field of pairwise clustering is currently dominated by the idea of dividing a set of objects into disjoints classes, thereby giving rise to (hard) partitions of the input data. However, in many computer vision and pattern recognition problems this approach is too restrictive as objects might reasonably belong to more than one class. In this paper, we adopt a game-theoretic perspective to the iterative extraction of possibly overlapping clusters: Game dynamics are used to locate individual groups, and after each extraction the similarity matrix is transformed in such a way as to make the located cluster unstable under the dynamics, without affecting the remaining groups.

1 Introduction

The most common approaches to pairwise clustering are centered around the notion of a cut [5], and this makes them intrinsically unable to provide overlapping groups. and, in general, require the number of groups to be known *ab initio*. The idea of overlapping groups is often associated with that of "soft" or "fuzzy" clustering, where each element has a "degree of membership" to each cluster, where this fuzzy membership represents the uncertainty on class membership rather than the ability of one element to belong to a single class. A more interesting approach would be one that can assign deterministically an object to more than one class. The distinction, however, is very fine, and it is worth noting that there are few "soft" clustering approaches that can deal with pairwise data.

Recently, a novel graph-theoretic notion of a cluster has been proposed, i.e., the "dominant set", which captures the two basic requirements of a cluster, namely internal coherency and external incoherency [3]. The approach is capable of automatically detect the number of groups present as it works as a multi figure/ground discrimination algorithm, extracting only cohesive groups, leaving spurious entries unclustered.

In this paper, we make use of the intrinsic properties of dominant sets in an attempt to extract possibly overlapping groups. Following [6], we adopt a game-theoretic perspective to the iterative extraction of the dominant sets: Game dynamics are used to locate individual dominant sets, and after each extraction the similarity matrix is transformed in such a way as to make the located cluster unstable under the dynamics, without affecting the remaining groups. This guarantees that once found, a cluster will not be extracted again.

2 The dominant set framework

The dominant set framework is a pairwise clustering approach [3] that is based on the notion of a *dominant set*, which can be seen as an edge-weighted generalization of a clique. The approach has proven to be a fast and efficient framework for pairwise clustering. The framework is based on a recursive characterization of the weight $\mathrm{w}_S(i)$ of element i with respect to a set S of elements, and characterizes a group as a *dominant set*, i.e., a set that satisfies:

$$\begin{aligned} &\mathrm{w}_S(i) > 0,\ \forall i \in S\,, \\ &\mathrm{w}_{S\cup\{i\}}(i) < 0,\ \forall i \notin S\,. \end{aligned}$$

These conditions correspond to the two main properties of a cluster: the first regards internal homogeneity, whereas the second regards external inhomogeneity.

The main result presented in [3] provides a one-to-one relation between dominant sets and strict local maximizers of $f(\boldsymbol{x}) = \boldsymbol{x}^T A\boldsymbol{x}$ in the standard simplex $\{\boldsymbol{x} \in \mathbb{R}^n\ :\ x_i \geq 0 \forall i \in V \text{ and } \sum_{i=1}^n x_i = 1\}$, where $A = (a_{ij})$ is the matrix of similarities of the elements to be grouped.

Specifically, in [3] it is proven that if S is a dominant subset of vertices, then its (weighted) characteristic vector $\mathbf{x}^S$, which is the vector of Δ_n defined as $x_i^S = \frac{\mathrm{w}_S(i)}{\sum_{i\in S}\mathrm{w}_S(i)}$ if $i \in S$ and $x_i^S = 0$ otherwise, is a strict local solution of $f(\boldsymbol{x})$. Conversely, if $\boldsymbol{x}$ is a strict local solution of $f(\boldsymbol{x})$, then its support $S = \sigma(\boldsymbol{x})$ is a dominant set, provided that $\mathrm{w}_{S\cup\{i\}}(i) \neq 0$ for all $i \notin S$. Here, the *support* of a vector $\boldsymbol{x} \in \Delta_n$ is the set of indices corresponding to its positive components, that is $\sigma(\mathbf{x}) = \{i \in V\ :\ x_i > 0\}$. By virtue of this result, a dominant set can be found by localizing a local solution of $f(\boldsymbol{x})$ and then picking up its support. In order to get

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

a partition of the input data Pavan and Pelillo suggest to iteratively find a dominant set and then remove the elements in its support from V, until all vertices have been grouped. This approach, while effective, is rather heuristic in nature and forces a partition on a framework that by its nature allows overlapping clusters. In fact, according to the framework, all local maxima form cohesive groups, yet the support of the maxima are not, in general, disjoint. Furthermore, by eliminating compact clusters from the similarity matrix we force an implicit change of scale.

3 Enumerating Dominant Sets

In [6] the dominant set framework was reinterpreted in a game theoretic perspective, to extend it to non-symmetric affinities. Within this perspective, groups are related to a particular game equilibrium: evolutionary stable strategies. Here we follow this approach to show how an asymmetric extension of the similarity matrix A can be used to enumerate all the dominant sets without the implicit change of scale. Note, however, that [6] focussed on asymmetric affinities and adopted the peeling strategies previously used by Pavan and Pelillo to extract the clusters, which forces the implicit change of scale. In this paper, on the other hand, we make use the properties of asymetric affinities to eliminate previously extracted solutions, thus enumerating several (possibly overlapping) clusters without the implicit change of scale.

Let $A = (a_{ij})$ the $n \times n$ payoff or utility matrix where a_{ij} is the payoff that a player gains when playing the strategy i against an opponent playing j. A *mixed strategy* is a probability distribution $\boldsymbol{x} = (x_1, x_2, \dots, x_n)'$ over the available strategies in V. The expected payoff that a player obtains by playing the element i against an opponent playing a mixed strategy $\boldsymbol{x}$ is $u(\boldsymbol{e}^i, \boldsymbol{x}) = (A\boldsymbol{x})_i = \sum_j a_{ij} x_j$, where $\boldsymbol{e}^i$ is the vector with all components equal zero except for the i^{th}-component which is equal to 1. The expected payoff received by adopting a mixed strategy $\boldsymbol{y}$ is $u(\boldsymbol{y}, \boldsymbol{x}) = \boldsymbol{y}' A\boldsymbol{x}$. The *best replies* against a mixed strategy $\boldsymbol{x}$ is the set of mixed strategies $\beta(\boldsymbol{x}) = \{\boldsymbol{y} \in \Delta : u(\boldsymbol{y}, \boldsymbol{x}) = \max_{\boldsymbol{z}} u(\boldsymbol{z}, \boldsymbol{x})\}$. A mixed strategy $\boldsymbol{x}$ is a *Nash equilibrium* if it is a best reply to itself, i.e. $\forall \boldsymbol{y} \in \Delta$, $u(\boldsymbol{y}, \boldsymbol{x}) \leq u(\boldsymbol{x}, \boldsymbol{x})$. A strategy $\boldsymbol{x}$ is said to be an *evolutionary stable strategy* (ESS) if it is a Nash equilibrium and for all $\boldsymbol{y} \in \Delta$ such that $u(\boldsymbol{y}, \boldsymbol{x}) = u(\boldsymbol{x}, \boldsymbol{x})$ we have that $u(\boldsymbol{x}, \boldsymbol{y}) > u(\boldsymbol{y}, \boldsymbol{y})$. In [6] was showed that ESSs are in a one-to-one relationship with an asymmetric generalization of the notion of dominant set. This notion reverts to the original one when the affinity matrix is symmetric.

Evolutionary game theory assumes the existence of a selection process that drives strategies with low payoff to extinction, a well-known formalization of this selection process is given by the replicator equations: $\dot{x}_i = x_i(u(\boldsymbol{e}^i, \boldsymbol{x}) - u(\boldsymbol{x}, \boldsymbol{x}))$. If the payoff matrix is symmetric then $\boldsymbol{x}' A\boldsymbol{x}$ is strictly increasing along any non-constant trajectory of any payoff-monotonic dynamics, while for a generic real matrix A, ESS are guaranteed to be asymptotically stable states for the replicator equations.

In order to enumerate dominant sets we iteratively render unstable all previously extracted ESSs by adding new strategies that are best replies to the previous ESSs, but to no other. This way the previous equilibria will no longer be asymptotically stable.

Let Σ be a tuple of ESSs of a game with payoff matrix A. So for example if $\boldsymbol{x}$ and $\boldsymbol{y}$ are ESSs of a doubly symmetric game then $\Sigma = (\boldsymbol{x}, \boldsymbol{y})$ and with Σ_i we select the i-th ESS. We denote the barycenter of the simplex-face spanned by the vertices in C with $\boldsymbol{b}^C \in \Delta$. The Σ-extension $A^\Sigma = (a^\Sigma_{ij})$ of the payoff matrix A is defined as follows.

$$a^\Sigma_{ij} = \begin{cases} a_{ij} & \text{if } i, j \in [1, n] \\ \alpha & \text{if } j > n \text{ and } i \notin \sigma(\Sigma_{j-n}) \\ \beta & \text{if } i, j > n \text{ and } i = j \\ \frac{1}{|\Sigma_{i-n}|} \sum_{k \in \Sigma_{i-n}} a_{kj} & \text{if } i > n \text{ and } j \in \sigma(\Sigma_{i-n}) \\ 0 & \text{otherwise.} \end{cases}$$

where $\alpha > \beta$ and $\beta = \max_{i,j} a_{ij}$.

Let Φ be a two-player doubly symmetric game with payoff matrix A and Φ^Σ be a two-player game with payoff matrix A^Σ. If $\boldsymbol{x}$ is a mixed strategy of Φ then $\bar{\boldsymbol{x}}$ is a mixed strategy of Φ^Σ obtained from $\boldsymbol{x}$ by setting the components relative to Σ-strategies to 0.

Theorem 1 *Let Φ be a two-player doubly symmetric game with payoff matrix A and let Σ be a tuple of ESSs of Φ. Furthermore let Φ^Σ be a two-player game with payoff matrix A^Σ. Then $\boldsymbol{x}$ is an ESS of Φ not in Σ if and only if $\bar{\boldsymbol{x}}$ is an ESS of Φ^Σ.*

For a proof see [7].

We use this result to enumerate the dominant sets in the following way: We iteratively find new dominant sets by looking for an asymptotically stable point using the replicator dynamics. After that, we extend the graph by adding the newly extracted set to Σ, hence rendering its associated strategy unstable, and reiterate the procedure until we have enumerated all the groups and hence are unable to find new dominant sets.

4 Experimental Results

In out first set of experiments we apply the proposed algorithm to the grouping of edge elements, namely, grouping the responses of an edge detector that belong

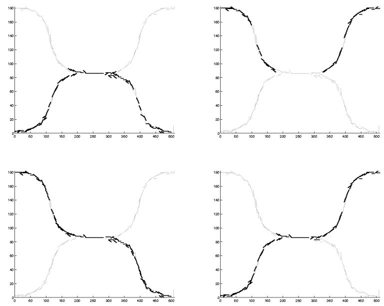


Figure 1. Grouping of line segments with multiple possible assignments.

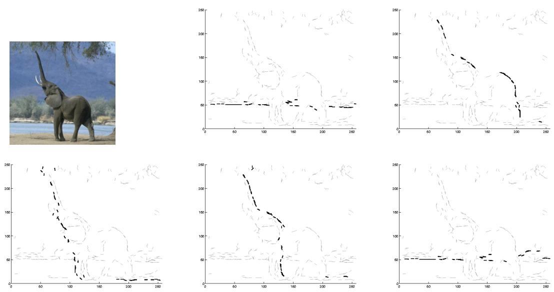


Figure 2. Perceptual grouping of edge elements.

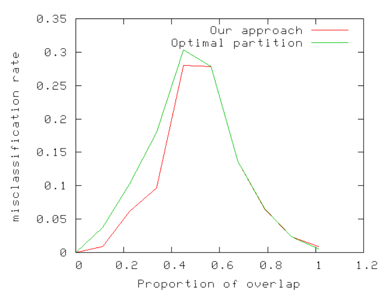
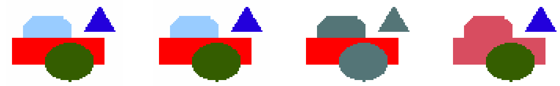


Figure 3. Overlapping Gaussian distributions.

to the same object boundary. To this end, we apply the proposed dominant set enumeration algorithm to the edge affinity measure proposed by Hérault and Horaud [1]. Figure 1 shows the results on a toy example where there is a clear ambiguity as to how the edges should be grouped together, yet our approach can deal with it allowing the central edges to be assigned to multiple groups. Figure 2 shows the result on a real image. Note how the group of edges from the trunk of the elephant is grouped both with the back of the animal and with the front leg.

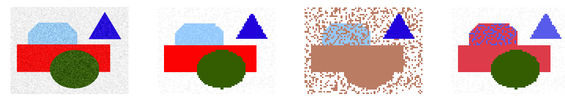
With the second set of experiments we try to assess robustness of the approach with respect to the amount of overlap of the groups. To this end we randomly sampled 300 points from two Gaussian distributions with increasing amount of overlap and assigned each point to a distribution if it was within one standard deviation from the mean, hence, allowing each point to be assigned to both distributions. The point distribution was then clustered using the proposed approach. Figure 3 shows the misclassification rate obtained with our approach and the best possible result that can be obtained with a hard partition. As long as the area of overlap is low enough for the clusters not to be merged, our approach is capable of recognizing the two separate distributions, yielding consistently lower misclassifications than the theoretical best partition.

In the third set of experiments we applied the proposed clustering approach to image segmentation based on color and texture similarities. To measure the similarity/distance between two points in an image we used similarity measure presented in [2] and we compared



Original | Dominant set | Normalized cut (partition) | Minimal-shift embedding (partition)

Figure 4. Synthetic segmentation, no noise.



Original | Dominant set | Normalized cut (partition) | Minimal-shift embedding (partition)

Dominant set

Normalized cut

Minimal-shift embedding

Figure 5. Synthetic segmentation, 30% of Gaussian noise.

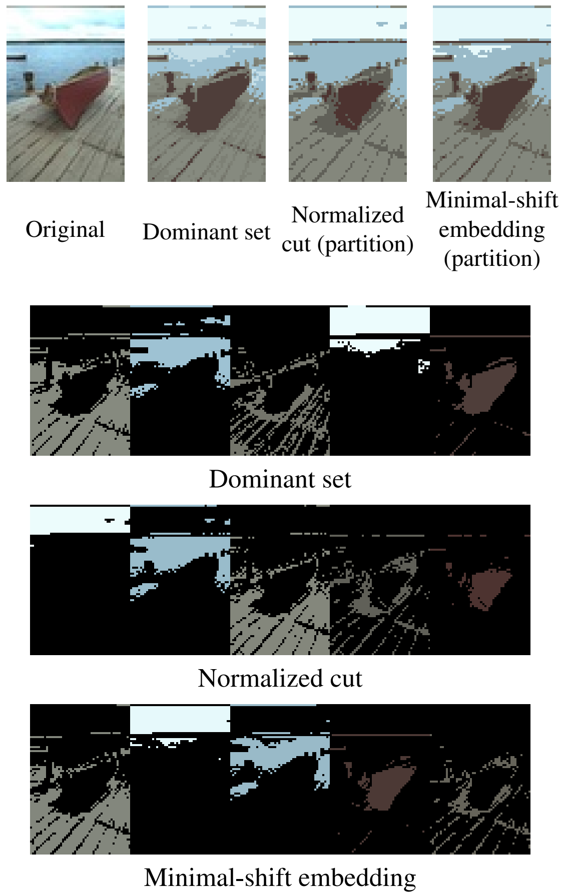


Figure 6. Comparing segmentation results.

our approach to Normalized cut [5] and the minimal shift embedding approach of Roth et al. [4]. Figure 4 shows the result of applying the three algorithms to a toy scene with clear partitions. In each image a point was assigned the average color of the cluster it belonged to, in the case of multiple memberships, the dominant colors of all the groups the pixel belongs to are averaged. The bottom row separately shows the groups extracted using the dominant set enumeration algorithm. Note that both normalized cut and the minimal shift embedding approach require the number of classes to be known. On the other hand our approach was able to recognize the number of objects in the scene and delivered a clear partition. Figure 5 shows the segmentations obtained using the three approaches after 30% Gaussian noise was added to the image. Here we show the extracted clusters for all three algorithms. While our enumeration algorithm did split a few groups, it proved more robust than the other two, never merging groups.

Finally, Figure 6 shows the comparison of the segmentation obtained with the tree approaches on a real image. The proposed approach is capable of respecting the clear boundaries present in the image, but at the same time to capture the ambiguities. Note the reflection of the sun on the sea: Our approach was able to assign that area both to the sea and to the sky, while normalized cut only assigns it to the sea and minimal shift embedding assigns it, at least partly, to the sky.

5 Conclusions

In this paper we presented a pairwise clustering approach based on dominat sets that allows for overlapping groups. We adopted a game theoretic approach to enumerate all the coherent groups, where groups are related to evolutionary stable strategies. The basis of our enumeration approach is a theoretical result allowing us to render unstable selected strategies. Experimental results show that the approach is capable of dealing with overlapping groups when there is ambiguity about class membership, yet delivering crisp partitions when elements clearly belong to a single group.

Acknowledgment

We acknowledge the financial support of the Future and Emerging Technology (FET) Programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open project SIMBAD grant no. 213250.

References

[1] L. Hèrault and R.Houraud, Figure-ground discrimination: A combinatorial optimization approach. *IEEE Trans. Pattern Anal. Machine Intell.*, 15:899–914, 1993.

[2] J. Mailk, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *Int. J. of Computer Vision* 43(1):7–27, 2001.

[3] M. Pavan and M. Pelillo. Dominant sets and pairwise clustering. *IEEE Trans. Pattern Anal. Machine Intell.* 29(1):167-172, 2007.

[4] V. Roth, J. Laub, M. Kawanabe, and J. M. Buhmann. Optimal cluster preserving embedding of nonmetric proximity data. *IEEE Trans. Pattern Anal. Machine Intell.* 25:1540–1551, 2003.

[5] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Machine Intell.* 22:888–905, 2000.

[6] A. Torsello, S. Rota Bulò, and M. Pelillo. Grouping with Asymmetric Affinities: A Game-Theoretic Perspective. In *IEEE Conf. on CVPR*, Vol. 1, pp. 292–299, 2006.

[7] A. Torsello, S. Rota Bulò, and M. Pelillo. A Game-theoretic approach for enumerating dominant sets. Technical report CS-2008-2, Dipartimento di Informatica, Università Ca' Foscari, 2008.

[8] K. Yu, S. Yu ,and V. Tresp. Soft Clustering on Graphs. In *Advances in Neural Information Processing Systems 18*, MIT Press, pp. 1553–1560, 2006.