

Integrating Boundary Information in Pairwise Segmentation

Andrea Torsello, Marco Di Gesù, and Marcello Pelillo
Dipartimento di Informatica
Università “Ca’ Foscari” di Venezia, Italy

Abstract

Proximity-based, or pairwise, data clustering techniques are gaining increasing popularity due to their versatility and their ability to easily integrate information of different nature. Despite this, most applications to image segmentation incorporate only region-based information, mainly color and texture similarity. In this paper we propose a general approach for integrating boundary information in a pairwise segmentation framework. To this end we propose a measure of distance between pair of pixels that integrates the value of an edge-response function along a path joining the two pixels. Experiments performed using the dominant sets framework show that the proposed approach is competitive with the state of the art pairwise segmentation algorithms even while using boundary information only. Furthermore, we show that the approach can effectively be used when adopting an out of sample approach to pairwise segmentation.

1. Introduction

Proximity-based, or pairwise, data clustering techniques are gaining increasing popularity over traditional central grouping techniques, which are centered around the notion of “feature” (see, e.g., [6, 12, 13, 11]). In many application domains, in fact, the objects to be clustered are not naturally representable in terms of a vector of features. On the other hand, quite often it is possible to obtain a measure of the similarity/dissimilarity between objects. Although such a representation lacks geometric notions such as scatter and centroid, it is attractive as no feature selection is required and it keeps the algorithm generic and independent from the actual data representation and metric involved. This means that similarity information arising from sources of very different nature can be incorporated very easily, often not requiring more than adding together distances or multiplying similarities calculated from different sources. In contrast, integrating information of different nature within the central clustering framework requires an integrated feature model

capable of simultaneously characterizing all information at the feature level. Moreover, pairwise approaches allow one to use non-metric similarities, are applicable to problems that do not have a natural embedding to a uniform feature space, such as the grouping of structural or graph-based representations, and their clustering criteria are intrinsically of a global nature, while feature-based grouping tends to be more “local” in nature.

These approaches have proven very powerful when applied to image segmentation problems [13, 7, 6, 4]. Here, the possibility of easily integrating different sources of information has been used to incorporate color, texture, and proximity information between pair of pixels. Conversely, feature based segmentation algorithms must explicitly integrate all types of information into a single geometrical model which requires a stronger characterization of the geometry of the image.

Despite the promise of ease of integration of inhomogeneous information, most actual implementations of pairwise segmentation only integrate region based information, with color- and texture-based similarities taking the lion’s share over all pairwise measures found in the literature. With few exceptions [7, 5], little work has been done to integrate boundary information in a pairwise setting. In contrast, the feature-based segmentation literature, partly due to the more “local” nature of the approaches, sees an abundance of frameworks integrating region and boundary information [8, 3, 2]. Approaches based on the Mumford and Shah functional [8] and related functional formulation such as Chan and Vese’s active contours [3] integrate region and boundary information by balancing region information obtained through the inside/outside functionals with boundary information imposed by the boundary elasticity prior. Chakraborty and Duncan [2] propose a game-theoretic approach to integrate boundary and region information without a tightly coupled geometric model. A different approach is explored in the graph-cut framework [1]. Here pixel-based boundary information and region-based pairwise information are integrated in a unifying framework. Within the purely pairwise setting, we note Malik and coworkers’ proposal to incorporate boundary information in the

normalized-cut framework by looking for an *intervening contour* [7, 5]. However, their approach only looks for detected edges in the straight line joining two pixels; hence, it is strongly dependant on the quality of the edge extractor and tends to separate pixels belonging to a single region if this is not convex. The normalized cut framework is relatively forgiving about this problem, but it is particularly severe when using pairwise clustering algorithms that favor “compact” globular clusters such as the dominant sets framework [9]. Furthermore, intervening contour information alone is not able to separate regions with fuzzy or unclear boundaries such as regions delimited by relatively smooth gradients.

In this paper, we propose a boundary-based pixel distance measure based on the minimal edge response along *all* paths joining two pixels. The idea underpinning our proposal is similar in nature to the intervening contour; however, the searches for the path with minimal response allows pixels of a connected but concave region to exhibit a low distance. Furthermore, our approach does not need the extraction of the boundary, but only requires the response of any edge filter, hence making the approach more robust. Finally, since we integrate the edge response along the path, the segmentation approach will work even in presence of unclear boundaries.

Experiments performed with the dominant sets framework show that the proposed approach is competitive with the state of the art pairwise segmentation algorithms even when using only boundary information. Furthermore, we show that the approach can be used when adopting an out of sample approach to pairwise segmentation.

2. Path-based Boundary Integration

Let $g(x, y)$ be an edge-response function, a function that has high values on edges and low (zero) values on flat areas. Typically this could be the response of an edge-detection filter, but more sophisticated edge functionals are possible. Further, let $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ be a path from point $a = \gamma(0)$ to point $b = \gamma(1)$. We define the boundary resistance along γ as:

$$\Omega_p(\gamma) = \left(\int_0^1 \|\gamma'(t)\| |g(\gamma(t))|^p dt \right)^{\frac{1}{p}}, \quad (1)$$

where $p \in \mathbb{R}_+$ is a parameter. Clearly, when the path crosses an edge, the boundary resistance assumes high values, while a path through a nearly homogeneous area yields very small boundary resistance. Furthermore, when going through a gradient of image intensity, the path slowly accumulates boundary resistance. The parameter p controls how the response accumulates when crossing gradients or multiple boundaries. The parameter works much in the same

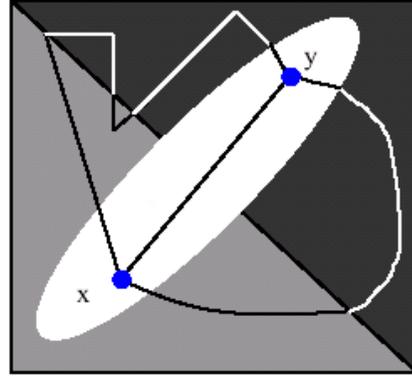


Figure 1. Different path yield different boundary resistances.

way as the parameter p in an l_p norm, indeed, $\Omega_p(\gamma)$ can be seen as the infinite dimensional l_p norm of the function $g(\gamma(t))$ when γ is parametrized by arc-length. This means that, as p increases, Ω_p approximates a max-norm, hence its value is dominated by the maximum edge-response along γ . Conversely, when p is small, the relative differences of the edge responses along γ are smoothed out, resulting in a measure dominated by the length of the path γ .

We define the boundary resistance between two points a and b as:

$$D_p(a, b) = \min_{\gamma} \Omega_p(\gamma) \quad \text{with } \gamma(0) = a, \gamma(1) = b. \quad (2)$$

In order to perform the computation on the discrete image lattice, we abstract the image as a neighborhood graph, where each pixel corresponds to a vertex in the graph, and each vertex is connected by an edge to all vertices corresponding to pixels in its neighborhood. In this paper we are using a 4-neighborhood system, hence, the graph is a mesh, but other neighborhood systems are possible. Further, to each edge linking node i to node j we associate a weight

$$w_{ij} = \left| \frac{(g(x_i, y_i) + g(x_j, y_j))}{2} \right|^p. \quad (3)$$

With this discrete representation, and recalling that $x^{\frac{1}{p}}$ is monotonic for $p > 0$, the computation of the minimum resistance path is reduced to the computation of the shortest path between each pair of points in a bounded-degree graph. This problem can be efficiently solved with the well known Dijkstra algorithm in $O(n^2 \log n)$, where n is the number of pixels.

The proposed approach bears strong similarities with Fischer and Buhmann’s path-based pairwise clustering approach [4], where the distance between two points is calcu-

lated looking for the minimum distance path through the elements to be clustered. However, in their approach, the calculation is performed as a “closure” of the distance matrix and works directly in the pairwise setting. Our approach, on the other hand, works in the image lattice incorporating proximity relations. This makes the initial distance information sparse as only distance between neighboring pixels is needed; hence, making the search for the optimal path more efficient. Furthermore, Fischer and Buhmann only take into account the maximum distance along a path, and use resampling to enforce robustness, while our approach achieves robustness by integrating information along the path in a controlled way.

3. The dominant set framework

The dominant set framework is a pairwise clustering approach [9] 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 pairwise clustering framework.

3.1. Clustering using dominant sets

The dominant set framework is based on a recursive characterization of the weight $w_S(i)$ 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:

1. $w_S(i) > 0$, for all $i \in S$
2. $w_{S \cup \{i\}}(i) < 0$, for all $i \notin S$.

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 [9] provides a one-to-one relation between dominant sets and strict local maximizers of the following quadratic program

$$\begin{aligned} & \text{maximize} && f(\mathbf{x}) = \mathbf{x}^t A \mathbf{x} \\ & \text{subject to} && \mathbf{x} \in \Delta_n \end{aligned} \quad (4)$$

where $A = (a_{ij})$ is the matrix of similarities of the elements to be grouped,

$$\Delta_n = \{\mathbf{x} \in \mathbb{R}^n : x_i \geq 0 \text{ for all } i \in V \text{ and } \mathbf{1}^t \mathbf{x} = 1\}$$

is the standard simplex of \mathbb{R}^n , and $\mathbf{1}$ is a vector of appropriate length consisting of unit entries.

Specifically, in [9] 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

$$x_i^S = \begin{cases} \frac{w_S(i)}{\sum_{i \in S} w_S(i)}, & \text{if } i \in S \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

is a strict local solution of program (4). Conversely, if \mathbf{x} is a strict local solution of program (4) then its support $S = \sigma(\mathbf{x})$ is a dominant set, provided that $w_{S \cup \{i\}}(i) \neq 0$ for all $i \notin S$. Here, the *support* of a vector $\mathbf{x} \in \Delta_n$ is the set of indices corresponding to its positive components, that is $\sigma(\mathbf{x}) = \{i \in V : x_i > 0\}$. The local maxima of (4) is found using the replicator equations, a dynamical systems mutated from game-theory. The approach has proven to be a very effective and robust pairwise clustering approach that has in its speed one of its major selling points.

3.2. Out-of-sample extension

Despite its strong properties, the application of this method to the segmentation of large images, and, indeed, of any pairwise approach, is very hard if not unfeasible due to the scaling behavior with the number of data items. On a dataset containing N pixels, the number of potential comparisons scales with $O(N^2)$, thereby rendering any pairwise approach too demanding, both in terms of computation time and of space. A way of overcoming this drawback is to drastically reduce the number of objects to be clustered and then extend the partition to the full data-set. In [10] Pavan and Pelillo propose an out-of-sample technique for the dominant set framework. The approach creates a virtual cluster centroid as a linear combination of the sample points that belong to the cluster, or, equivalently, gives a measure of distance of a point to a cluster. The weights of the linear combination are taken to be proportional to the weight $w_S(i)$ of sample i with respect to the cluster S . This way, the similarity of an out-of-sample point j to cluster centroid c_S , or simply to cluster S , is:

$$A_{S,j} = \sum_{i \in S} \frac{w_S(i)}{\sum_{i \in S} w_S(i)} a_{ij} = \sum_i x_i^S a_{ij}, \quad (6)$$

where $\mathbf{x}^S = (x_1^S, \dots, x_n^S)^t$ is the maximum of (4) corresponding to the cluster S . The first step of the out-of-sample segmentation is to extract the clusters from the sampled points. With the initial segmentation to hand, each pixel is assigned to the closest cluster according to (6).

Note that using out-of-sample clustering we only need distance from sample points to all other points in the image. This means that we must run the Dijkstra algorithm only from sources within the sample set. Hence, the total cost of computing the distance matrix for the out of sample algorithm is $O(nN \log N)$ where n is the number of samples and N is the total number of pixels. If $n \ll N$ the computing the matrix is essentially log-linear in the number of pixels, which means that it has essentially the same cost as computing the edge response using a derivative of Gaussian or other linear filters using FFT.

4. Experimental Results

In order to assess the ability of the proposed distance measure to represent a stable boundary-based distance for image segmentation, we performed a set of experiments using only our measure. Clearly, the final goal is to incorporate the measure with standard region based similarities, but, to be considered descriptive and stable, the measure must allow for a reasonable segmentation even when applied alone.

Figure 2 shows the result of applying the dominant sets clustering algorithm to the proposed measure with $p = 1$ and the edge response obtained through a Sobel filter. The images were mostly selected from the Berkeley database so that the results could be compared with those published in [9, 13], and were scaled so that their size ranged from 5000 to 36000 pixels. The left column shows the original images, while the right column shows the segmented images, where all pixels belonging to the same cluster are drawn using the average color of the cluster. From the results we can see that the output is competitive with other region-based pairwise segmentation approaches. In particular, note, for example, that grains of the wooden door in the first image, while quite visible in the image, are not enough to split the door into separate clusters, something that would happen looking only at interposing edges.

Our next set of experiments focuses on the effects of the parameter p . Figure 3 displays the segmentations obtained with values of p varying from 0.5 to 8. Again, the segmentation was performed using only the proposed boundary-based measure with the edge response obtained through a Sobel filter. Here we can see that with the lowest value of p the measure is dominated by the distance between the points, with only the strongest edges influencing the segmentation. The extracted clusters tend to be very compact and break only on very strong lines. The compactness imposed by the dominance of the distances prevented the algorithm from extracting thin clusters forcing it to ignore close parallel edges. On the other hand, as p increases, the measure is dominated by the single maximum response along the path, making the measure more sensible to thin lines, but making it less robust to noise as well. In fact we can see that for large p small random noise in the images produced several small clusters with only one or two pixels. A less noisy edge response such as a gradient of Gaussian might alleviate the problem, but ultimately, very high values of p will make the measure unstable even under moderate level of noise. In general the best results were obtained with values of p between 1 and 2, with 1 been the best value on average.

Finally, Figure 4 shows the clusters obtained using our measure with the out-of-sample extension of the dominant set framework. Here the original images were much larger,

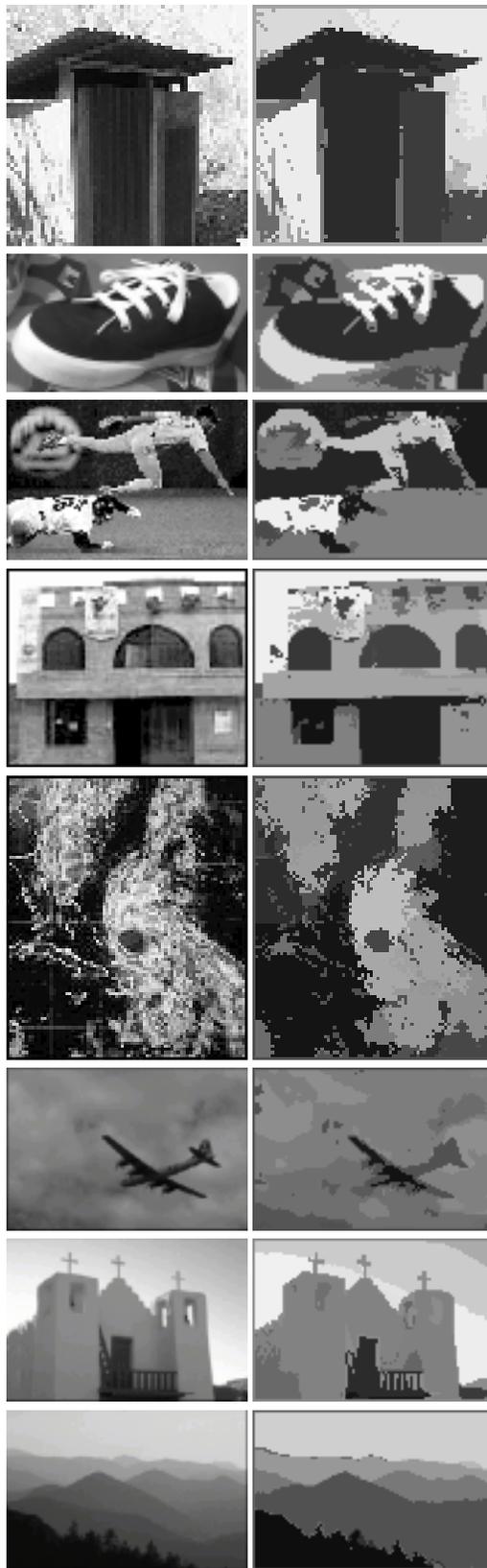


Figure 2. Segmentation using dominant sets and our boundary-based measure.

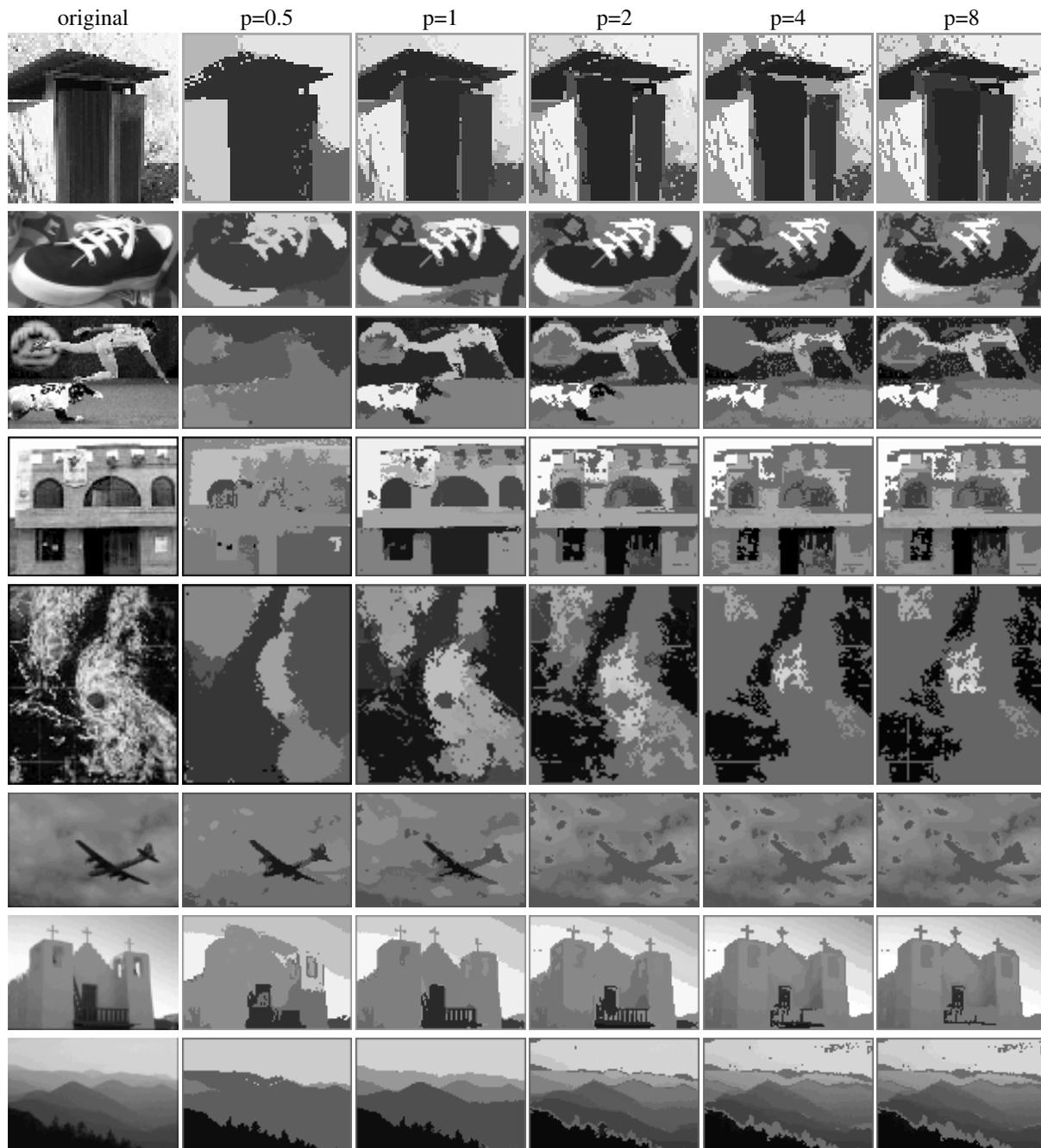


Figure 3. Effects of the parameter p on the segmentation process.



Figure 4. Out of sample segmentation using the proposed measure.

over 150000 pixels in total, with a sampling rate of 1.2%. As we can see, even with the reduced information of the out of sample framework the segmentation works quite well, yielding results comparable with those obtained using the full set of distances.

5. Conclusions

In this paper we presented a novel boundary-based measure of pixel similarity to be used in conjunction of other region-based similarity measures in pairwise segmentation algorithms. The measure integrates an edge-response filter along a path and searches for the minimal-response path between two pixels using the Dijkstra algorithm. A parameter p allows us to control whether the final distance is dominated by the maximal edge-response along the path, by the average response, or by the path length, with all the possibilities in-between being available. Experiments using the dominant sets pairwise clustering framework show that good segmentation can be achieved even when using boundary information alone and that the measure is robust for small values of p , while higher values of the parameter make the approach more sensitive to noise. Finally, experimental results show that the measure is capable of producing good results even when using an out-of-sample extension of the clustering algorithm.

References

- [1] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts." *IEEE Trans. Pattern Anal. Machine Intell.*, 23(11):1222–1239, 2001.
- [2] A. Chakraborty and J. S. Duncan, "Game-Theoretic Integration for Image Segmentation." *IEEE Trans. Pattern Anal. Machine Intell.*, 21(1):12–29, 1999.
- [3] T. F. Chan and L. A. Vese, "Active Contours Without Edges." *IEEE Trans. Pattern Anal. Machine Intell.*, 10(2):266–277, 2001.
- [4] B. Fischer, J. M. Buhmann. Path-based clustering for grouping of smooth curves and texture segmentation. *IEEE Trans. Pattern Anal. Machine Intell.* 25(4):513–518, 2003.
- [5] C. Fowlkes, D. Martin, and J. Malik, "Learning Affinity Functions for Image Segmentation: Combining Patch-based and Gradient-based Approaches", In *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2003.
- [6] T. Hofmann and J. M. Buhmann, "Pairwise data clustering by deterministic annealing." *IEEE Trans. Pattern Anal. Machine Intell.* 19:1–14, 1997.
- [7] J. Maillk, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *Int. J. of Computer Vision* 43(1):7–27, 2001.
- [8] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems." *Commun. Pure Appl. Math*, 42:577–685, 1989.
- [9] M. Pavan and M. Pelillo, "Dominant sets and pairwise clustering," *IEEE Trans. Pattern Anal. Machine Intell.* 29(1):167–172, 2007.
- [10] M. Pavan and M. Pelillo, "Efficient out-of-sample extension of dominant-set clusters." In *Advances in Neural Information Processing Systems 17*, MIT Press, pp. 1057–1064, 2005.
- [11] 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.
- [12] S. Sarkar and K. Boyer, "Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors." *Computer Vision and Image Understanding* 71:110–136, 1998.
- [13] J. Shi and J. Malik, "Normalized cuts and image segmentation." *IEEE Trans. Pattern Anal. Machine Intell.* 22:888–905, 2000.