Grouping with Asymmetric Affinities: A Game-Theoretic Perspective

Andrea Torsello, Samuel Rota Bulò, and Marcello Pelillo
Dipartimento di Informatica, Università "Ca' Foscari"
via Torino 155, 30172 Venezia Mestre, Italy
{torsello,srotabul,pelillo}@dsi.unive.it

Abstract

Pairwise grouping and clustering approaches have traditionally worked under the assumption that the similarities or compatibilities between the elements to be grouped are symmetric. However, asymmetric compatibilities arise naturally in many areas of computer vision and pattern recognition. Hence, there is a need for a new generic approach to clustering and grouping that can deal with asymmetries in the compatibilities. In this paper, we present a generic framework for grouping and clustering derived from a game-theoretic formalization of the competition between the hypotheses of group membership, and apply it to perceptual grouping. In the proposed approach groups correspond to evolutionary stable strategies, a classic notion in evolutionary game theory. We also provide a combinatorial characterization of the stable strategies, and, hence, of the elements that belong to a group. Experiments show that our approach outperforms both state-of-the-art clustering-based perceptual grouping approaches with symmetric compatibilities, and other approaches explicitly designed to make use of asymmetric compatibilities.

1. Introduction

A common approach to grouping is to cast it into an instance of pairwise clustering by assigning to each element a compatibility with the other responses of the detector [13, 11, 6, 5, 15, 12]. This is due to the fact that in many application domains, the objects to be clustered are not naturally representable in terms of a vector of features, but their properties are more naturally described in terms of similarity/dissimilarity between the various objects. The grouping algorithms within this class are very versatile and, in contrast to parametric approaches, do not require prior information of the group distribution, but can work with simple soft priors. There are two fundamental ingredients in this approach to grouping. The first is a similarity or affinity measure that quantifies the compatibility between two elements. The second ingredient is a clustering algorithm that can extract a set of mutually compatible image elements.

Stemming from a natural assumption for central clustering frameworks, pairwise grouping and clustering approaches have traditionally worked under the assumption that the similarities satisfy metric properties, i.e., they are symmetric, non-negative, and satisfy the triangle inequality. However, recently there has been a strong interest in relaxing these requirements [9, 4, 16]. This is due to the fact that in many applications non-metric similarities arise naturally [15, 2]. More fundamentally, some researches argue that human perception does not satisfy metric properties [4]. While the literature presents many approaches that lift the assumption of non-negativity and triangle inequality [9, 4], little progress has been made in relaxing the symmetry constraint. Note, however, that the limited progress in grouping with asymmetric affinities is not due to the lack of interest. Non-symmetric similarities, in fact, arise naturally in many areas. Common examples are the directed Hausdorff distance between sets, and Kullback-Leiber divergence between probability distributions, but several others can be found in many fields.

While it is relatively easy to come up with an asymmetric affinity, it is much harder to define what a cluster is when we lift the symmetry constraint. For this reason, generic clustering approaches using asymmetric affinities have proven elusive. A common method to deal with asymmetric compatibilities is to symmetrize it by transforming the similarity matrix, typically by taking the maximum, minimum or mean of the entries corresponding to the two directions of a binary relation. This approach, however, loses any information that might reside in the asymmetry. Yu and Shi [16] presented an extension of the normalized cut approach to take into account directed attractions and repulsions. Here the affinity matrix is split into its symmetric and antisymmetric components and a Hermitian Matrix is constructed by setting the real part equal to the symmetric component and the imaginary part equal to the antisymmetric component. Spectral analysis then provides the group information.

In this paper, we present a new framework for grouping and clustering derived from a game-theoretic formalization of the competition between the hypotheses of group mem-

Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)
0-7695-2597-0/06 $20.00 © 2006 IEEE

bership. In our approach the group corresponds to the evolutionary stable strategies of a non-cooperative game [14]. The basic idea behind our proposal is as follows: the hypotheses that each object belongs to the figure compete with one-another, each obtaining support from compatible edges and competitive pressure from all the other. Competition will reduce the population of individuals that assume hypotheses that do not receive strong support from the rest, while it will allow populations assuming hypotheses with strong support to thrive. Eventually all inconsistent hypotheses will be driven to extinction, while all the surviving hypotheses will reach an equilibrium where all receiving the same average support, hence exhibiting the internal coherency of a group, while all the extinct hypotheses must have a lower support, hinting to external incoherency. The stable strategies are found using replicator dynamics, a classic formalization of a natural selection process. We apply the resulting grouping algorithm to perceptual grouping.

2. Grouping as a non-cooperative game

Consider the following *grouping* game. Assume a pre-existing set of objects O and a (possibly asymmetric) matrix of affinities A between the elements of O. Two players with complete knowledge of the setup play by simultaneously selecting an element of O. After both have shown their choice, each player receives a payoff, monetary or otherwise, proportional to the affinity that the chosen element has with respect to the element chosen by the opponent. Clearly, it is in each player's interest to pick an element that is strongly supported by the elements that the adversary is likely to choose. As an example, let us assume that our grouping problem is one of figure/ground discrimination, that is, the objects in O consist of a cohesive group with high mutual affinity (figure) and of non-structured noise (ground). Being non-structured, the noise gives equal average affinity to elements of the figures as to elements of the ground. Informally, assuming no prior knowledge of the inclination of the adversary, a player will be better-off selecting elements of the figure rather than of the ground.

Let $O = \{1, \cdots, n\}$ be the set of available elements (*pure strategies* in the language of game theory) and, $A = (a_{ij})$ be the $n \times n$ element-affinity matrix, also called payoff or utility matrix in game theory [14]. Specifically, for each pair of strategies $i, j \in O$, a_{ij} represents the payoff of an individual playing strategy i against one playing strategy j.

A *mixed strategy* is a probability distribution $\mathbf{x} = (x_1, \ldots, x_n)'$ over the available strategies O. From here on, a prime denotes transposition. Clearly, mixed strategies are constrained to lie in the standard simplex of the n–dimensional Euclidean space $\mathbb{R}^n$

$$\Delta = \left\{ \mathbf{x} \in \mathbb{R}^n \ : \ x_i \geq 0 \text{ for all } i \in O, \ \sum_{i \in O} x_i = 1 \right\} .$$

The *support* of a mixed strategy $\mathbf{x} \in \Delta$, denoted by $\sigma(\mathbf{x})$, is defined as the set of elements chosen with non-zero probability: $\sigma(\mathbf{x}) = \{i \in O \mid x_i > 0\}$. The expected payoff received by a player choosing element i when playing against a player adopting a mixed strategy $\mathbf{x}$ is $(A\mathbf{x})_i = \sum_j a_{ij} x_j$, hence the expected payoff received by adopting the mixed strategy $\mathbf{y}$ against $\mathbf{x}$ is $\mathbf{y}' A\mathbf{x}$. The *best replies* against mixed strategy $\mathbf{x}$ is the set of mixed strategies

$$\beta(\mathbf{x}) = \{\mathbf{y} \in \Delta \mid \mathbf{y}' A\mathbf{x} = \max_{\mathbf{z}} (\mathbf{z}' A\mathbf{x})\} .$$

while the *best pure replies* against mixed strategy $\mathbf{x}$, denoted with $\Omega(\mathbf{x})$, is the set of pure strategies that are best replies to $\mathbf{x}$. It can be shown that, if $\mathbf{y}$ is in $\beta(\mathbf{x})$, then each strategy in $\sigma(\mathbf{y})$ is in $\Omega(\mathbf{x})$. A strategy $\mathbf{x}$ is said to be a *Nash equilibrium* if it is the best reply to itself, i.e., $\forall \mathbf{y} \in \Delta, \ \mathbf{x}' A\mathbf{x} \geq \mathbf{y}' A\mathbf{x}$. It is easy to show that this implies that $\forall i \in \sigma(\mathbf{x})$ we have $(A\mathbf{x})_i = \mathbf{x}' A\mathbf{x}$; that is, the payoff of every strategy in the support of $\mathbf{x}$ is constant. Furthermore, note that, in general, we have $\sigma(\mathbf{x}) \subseteq \Omega(\mathbf{x})$.

Within our setting, Nash equilibria abstracts well the main characteristics of a group: internal coherency, that is, a high mutual support of all elements within the group, and external incoherency, or low support from elements of the group to elements that do not belong to the group. In fact, any element $i \in \sigma(\mathbf{x})$ of a Nash equilibrium $\mathbf{x}$ receive from $\mathbf{x}$ the same expected payoff $(A\mathbf{x})_i = \mathbf{x}' A\mathbf{x}$, while elements not in $\Omega(\mathbf{x})$ receive a lower or equal support from the elements of the group. Note, however, that external incoherency is not strict: while strategies that are not in $\sigma(\mathbf{x})$ cannot have higher than average payoff, they can have a payoff equal to $\mathbf{x}' A\mathbf{x}$ like elements in the group. For this reason we will impose a more stringent requirement, namely that $\Omega(x) = \sigma(x)$. Note, however, that this is not enough, we also want the solution to be stable and unambiguous, that is we require the solution to be isolated and unique in $\beta(\mathbf{x})$.

To this end, here we undertake an evolutionary game-theoretic analysis of the possible strategies available to each player. Evolutionary game theory considers an idealized scenario whereby pairs of individuals are repeatedly drawn at random from a large population to play a symmetric two-player game. In contrast to traditional game theoretic models, players are not supposed to behave rationally or to have complete knowledge of the details of the game. They act instead according to a pre-programmed behavior pattern, or mixed strategy, and it is supposed that some selection process operates over time on the distribution of behaviors.

In our grouping-game setting, each player is pre-programmed to select each element in O with a certain probability and the evolutionary selection will allow players that select elements with high average support to thrive, while driving players that choose elements with low support to extinction. In our grouping setup, we expect the selective pressure to drive to extinction the players programmed to select elements of the ground, converging to a population selecting elements of the figure only.

A strategy $\mathbf{x}$ is said to be an *evolutionary stable strategy* (ESS) if it is a Nash equilibrium and

$$\forall \mathbf{y} \in \Delta \quad \mathbf{x}'A\mathbf{x} = \mathbf{y}'A\mathbf{x} \Rightarrow \mathbf{x}'A\mathbf{y} > \mathbf{y}'A\mathbf{y}\,. \tag{1}$$

This condition guarantees that any deviation from the stable strategies does not pay. Further, it implies that A is negative-definite in the face of Δ spanned by the strategies with maximum payoff. In fact, $\forall \mathbf{y} \in \Delta$, $\mathbf{y}'A\mathbf{x} = \mathbf{x}'A\mathbf{x}$, we have, $(\mathbf{y}-\mathbf{x})'A(\mathbf{y}-\mathbf{x}) = \mathbf{y}'A\mathbf{y} - \mathbf{x}'A\mathbf{y} < 0$.

Evolutionary stability provides a constraint that forces the group to be non-ambiguous. In fact, the fact that $\mathbf{x}$ is ESS implies that it is an isolated Nash equilibrium, or that there exists an open set $\mathcal{U}$ containing $\mathbf{x}$ with no other other Nash equilibrium within it. Hence, evolutionary stable strategies with $\sigma(\mathbf{x}) = \Omega(\mathbf{x})$ satisfy all the conditions we posed for a cluster: internal coherency, external disomogeneity, stability and non-ambiguity.

3. Characterization of the equilibria

In this section we will provide a combinatorial characterization of the evolutionary stable points of the two-player game with both binary and weighted affinities. We start with the binary case since it is easier and allows a direct interpretation. However, the results will be sufficient to draw some conclusions about the role of the asymmetry within this framework.

3.1. Binary affinities

In this subsection we will characterize the solutions to our game-theoretic clustering approach in presence of binary 0-1 affinities. A clustering problem with binary affinities can be described as a directed graph where the presence of a directed edge from node i to node j implies a positive compatibility of node j with node i. The intuitive notion of a cluster is a subset of elements that are all mutually compatible and have low external compatibility. The graph theoretic counterpart of mutual-compatibility within a subset is a straightforward extension to directed graphs of the concept of clique. A *clique* is a subset of vertices of a undirected graph that are all mutually adjacent.

Let $G(V,E)$ be a directed graph with vertex set V and edge set $E \subseteq V \times V$, a $S \subseteq V$ is a *doubly-linked clique* if $\forall i,j \in S$, $(i,j) \in E$, and $(j,i) \in E$. Furthermore, if there is no $j \in (V \setminus S)$ such that $\forall i \in S, \quad (i,j) \in E$, the doubly linked clique is said to be *saturated*.

Let $S \subseteq V$, the characteristic vector of S is a vector $\mathbf{x}^S \in \Delta$ defined as

$$x_i^s = \begin{cases} 1/|S| & \text{if } i \in S \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Theorem 1 *Let $G(V,E)$ be a directed graph with adjacency matrix A. $S \subseteq V$ is a saturated doubly-linked clique of G if and only if $\mathbf{x}^S$ is an ESS for a two-player game with payoff matrix $B = A' + \alpha I$, $1/2 < \alpha < 1$*[1].

The reason for using A' instead of A is due to the fact in graph-theory that a_{ij} represents an edge from vertex i to vertex j, while in game theory a_{ij} represents a support from node j to node i. Hence, we transpose the affinity matrix to translate between the two interpretations. Now we can proceed to proving the theorem.

Proof. Suppose S is a doubly-linked clique, $\mathbf{x}^S$ is a Nash point. In fact,

$$\forall i \in S \quad (B\mathbf{x}^S)_i = \frac{|S| - 1 + \alpha}{|S|}\,.$$

Further, let $O_j^S = \{i \in S | (i,j) \in E\}$, we have

$$\forall j \in V \setminus S \quad (B\mathbf{x}^S)_j = \frac{|O_j^S|}{|S|}\,.$$

For the hypothesis, $|O_j^S| \leq |S| - 1$, hence, $(B\mathbf{x}^S)_j < (B\mathbf{x}^S)_i$.

Let B_S be the restriction of B to S. Clearly, $\mathbf{x}^S$ is the single global maximum of $\mathbf{x}'B_S\mathbf{x}$, hence $\forall \mathbf{y} \in \Delta, \sigma(\mathbf{y}) = S$ we have $\mathbf{y}'B\mathbf{y} < (\mathbf{x}^S)'B\mathbf{x}^S$. Yet $(\mathbf{x}^S)'B\mathbf{x}^S = \mathbf{y}'B\mathbf{x}^S$ since $\mathbf{x}^S$ is a Nash point, and $\mathbf{y}'B\mathbf{x}^S = (\mathbf{x}^S)'B\mathbf{y}$ since B_S is symmetric. Hence, $\mathbf{y}'B\mathbf{y} < (\mathbf{x}^S)'B\mathbf{y}$.

Conversely, if $\mathbf{x}$ is an ESS, $\sigma(x)$ is a doubly-linked clique. In fact, let, by absurd, $a_{ij} = 0$ with $i,j \in \sigma(\mathbf{x})$, $i \neq j$. and let $\mathbf{y} = \mathbf{x} + \delta(\mathbf{e}_i - \mathbf{e}_j)$ with $0 < \delta \leq x_j$. Then,

$$\begin{aligned} &\mathbf{y}^T B\mathbf{y} = \mathbf{x}^T B\mathbf{y} + \delta(\mathbf{e}_i - \mathbf{e}_j)^T B\mathbf{y} = \\ &\mathbf{x}^T B\mathbf{y} + \delta(\mathbf{e}_i - \mathbf{e}_j)^T B\mathbf{x} + \delta^2(\mathbf{e}_i - \mathbf{e}_j)^T B(\mathbf{e}_i - \mathbf{e}_j) = \\ &\qquad \mathbf{x}^T B\mathbf{y} + \delta^2(2\alpha - a_{ij} - a_{ji}) > \mathbf{x}^T B\mathbf{y}\,, \end{aligned}$$

which implies that $\mathbf{x}$ is not an ESS, against the hypothesis. Hence, $S = \sigma(\mathbf{x})$ is a doubly linked clique. Furthermore,

$$\forall i \in S\, (B\mathbf{x})_i = (|S| - 1) - x_i(|S| - 1 - \alpha)\,.$$

Hence, since all the supports must be equal, $x_i = 1/|S|$. □

It is interesting to note the role of the asymmetry in the selection of the cluster. First, the elements that belong to a group must all be mutually compatible, hence forcing, in the binary case, a strong symmetry within the cluster. In the general case the affinities within the cluster must not be

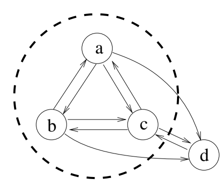

Figure 1. Unstable doubly-linked clique.

[1]The diagonal value α is just a technicality to guarantee unambiguous solutions.

completely symmetric, but there must be a strong mutuality between each pair of elements so that compatibility must be high in both directions. The asymmetry comes into play only in inside/outside relations though the condition that a doubly-linked clique must not be fully connected to an external node to be evolutionally stable. This condition allows the asymmetry to intervene in the selection of an equilibrium by dominating strategies belonging to a clique. See, for an example, Figure 1. Here nodes a, b, c form a doubly-linked clique, however all the nodes link to an external node d, hence, by Proposition 1 $\mathbf{x}^{\{a,b,c\}}$ is not an ESS. On the other hand node d links back only to c. This means that any deviation towards strategy d will reduce the support of a and b with respect to that of c or d. Hence, a and b are not best replies in the new environment and will be driven to extinction, leaving the solution $\{c, d\}$ which is a doubly-linked clique not fully connected to any external node and, hence, an ESS. In this case the asymmetry is able to route the selection process towards a smaller clique.

3.2. Continuous affinities

After providing a complete characterization of the ESS points obtained from binary affinities, now we provide a complete characterization of the ESS with general affinity matrix in terms of a generalization of the dominant set concept to the general case of asymmetric affinities. The dominant set framework has been presented in [8]. While it has been introduced in the context of symmetric similarities, it is straightforward to extend it to the case of asymmetric affinities. Let $S \subseteq O$ be a non-empty subset of elements. The *(average) weighted in-degree* of $i \in O$ w.r.t. S is defined as:

$$\operatorname{awindeg}_S(i) = \frac{1}{|S|}\sum_{j\in S} a_{ij} \tag{3}$$

where $|S|$ denotes the cardinality of S. Moreover, if $j \in S$ we define $\phi_S(i,j) = a_{ij} - \operatorname{awindeg}_S(j)$ which is a measure of how compatible node i is with node j with respect to the average compatibility of node j with elements in S.

Let $S \subseteq O$ be a non-empty subset of vertices and $i \in S$. The *weight* of i w.r.t. S is

$$\mathrm{w}_S(i) = \begin{cases} 1, & \text{if } |S| = 1 \\ \displaystyle\sum_{j\in S\setminus\{i\}} \phi_{S\setminus\{i\}}(i,j)\,\mathrm{w}_{S\setminus\{i\}}(j), & \text{otherwise} \end{cases} \tag{4}$$

while the *total weight* of S is defined as: $\mathrm{W}(S) = \sum_{i\in S}\mathrm{w}_S(i)$. Intuitively, $\mathrm{w}_S(i)$ gives us a measure of the support that vertex i receives from the vertices of $S\setminus\{i\}$ relative to the overall mutual affinity of the vertices in $S\setminus\{i\}$,. Here positive values indicate that i has high affinity to $S\setminus\{i\}$.

A non-empty subset of vertices $S \subseteq O$ such that $\mathrm{W}(T) > 0$ for any non-empty $T \subseteq S$, is said to be *directed-dominant* if:

1. $\mathrm{w}_S(i) > 0$, for all $i \in S$
2. $\mathrm{w}_{S\cup\{i\}}(i) < 0$, for all $i \notin S$.

The two previous conditions correspond to the two main properties of a cluster: the first regards internal coherency, whereas the second regards external incoherency. The above definition represents our formalization of the concept of a cluster in an edge-weighted graph.

The characteristic vector x^S of a set $S \subseteq O$ is defined as

$$x_i^S = \begin{cases} \frac{\mathrm{w}_S(i)}{\mathrm{W}_S} & \text{if } i \in S \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 2 *If $S \subseteq O$ is a dominant set with respect to affinity matrix A, then $\mathbf{x}^S$ is an ESS for a two-player game with payoff matrix A.*

Conversely, if $\mathbf{x}$ is an ESS for a two-player game with payoff matrix A, then $S = \sigma(x)$ is a dominant set wrt A, provided that $\sigma(\mathbf{x}) = \Omega(\mathbf{x})$.

Proof. Here we give only a sketch of the proof, with some details missing.

For all $T \subseteq O$, with A_T we represent the restriction of A to T and with B_T the matrix

$$B_T = \left(\begin{array}{c|c} 0 & \mathbf{1}' \\ \hline \mathbf{1} & A_T \end{array}\right),$$

where $\mathbf{1} = (1, \ldots, 1)'$. Further, with ${}^iB_T$ we indicate a matrix obtained from B_T by substituting the ith column with vector $(1, 0, \ldots, 0)'$.

Clearly, requiring that x^S be a Nash equilibrium is equivalent to saying that there is a $\lambda \in \mathbb{R}$ such that

$$B_S(\lambda, \mathbf{x}'_S)' = (1, 0, \ldots, 0)' \tag{5}$$

Further, we must have $\forall i \in O \setminus S \quad (A\mathbf{x}^S)_i \leq \lambda$.

Following the a proof technique similar to the equivalent in [7], we can prove that for all $T \subseteq O$

$$\begin{array}{l} \mathrm{w}_T(i) = (-1)^{|T|}\det\left({}^iB_T\right) \\ \mathrm{W}(T) = (-1)^{|T|}\det\left(B_T\right). \end{array}$$

By Cramer rule, we get $B_S(\lambda, \mathbf{x}'_S)' = (1, 0, \ldots, 0)'$. Further, since S is a directed-dominant set, for $j \notin S$ we have

$$\begin{array}{l} 0 > \mathrm{w}_{S\cup\{j\}}(j) = \sum_{i\in S}\phi_S(j,i)\mathrm{w}_S(i) = \\ \sum_{i\in S} a_{ji}\mathrm{w}_S(i) - \sum_{i\in S}\operatorname{awindeg}_S(i)\mathrm{w}_S(i) = \\ \mathrm{W}(S)\left[(A\mathbf{x}^S)_j - \frac{1}{|S|}\sum_{i\in S}(A\mathbf{x}^S)_i\right]. \end{array}$$

Hence, $(A\mathbf{x}^S)_j < \lambda$, which proves that $\mathbf{x}^S$ is a Nash equilibrium.

To prove that $\mathbf{x}^S$ is an ESS, we need to prove that $\forall \mathbf{y} \in \Delta$ with $\mathbf{y}'A\mathbf{x} = \mathbf{x}'A\mathbf{x}$, $(\mathbf{y}-\mathbf{x})'A(\mathbf{y}-\mathbf{x}) < 0$, but, using the bounded Hessian test, this can be proven to hold if and only if $\forall T \subseteq \Omega(\mathbf{x})\, \mathrm{W}(T) > 0$, which is true since S is a directed-dominant set.

To prove the converse, we note that, since $\mathbf{x}$ is an ESS, we have $\forall \mathbf{y} \in \Delta, \sigma(\mathbf{y}) \subseteq \Omega(\mathbf{x})$, $(\mathbf{y}-\mathbf{x})'A(\mathbf{y}-\mathbf{x}) < 0$. Hence, $\forall T \subseteq S \quad W(T) > 0$, which implies that $\det(B_S) = (-1)^{|S|}\mathrm{W}(S) \neq 0$. This, in turn, implies that there is only one solution to (5) and, as shown previously, this solution is $\mathbf{x}^S$ and $\forall i \in S \quad \mathrm{w}_S(i) = x_i^S \mathrm{W}(S) > 0$. It remains to be shown that $\forall j \notin S\, \mathrm{w}_{S\cup\{j\}}(j) < 0$, but as shown previously, $\mathrm{w}_{S\cup\{j\}}(j) = \mathrm{W}(S)\left[(A\mathbf{x}^S)_j - \lambda\right]$. Since, by hypothesis, $\sigma(\mathbf{x}) = \Omega(\mathbf{x})$, we have $(A\mathbf{x})_j < \lambda$, hence, $\mathrm{w}_{S\cup\{j\}}(j) < 0$. □

This result has a twofold implication. On the one hand, it supports the idea of using ESS as a notion of group, since in the symmetric case it reduces to a known cluster concept. On the other hand, it provides a generalization of the dominant set framework to asymmetric affinities.

4. Evolving towards a cohesive group

In this section, we describe an algorithmic approach for extracting a coherent group from a set of stimuli. This is done by simulating the evolution of a natural selection process which is guaranteed to converge to Nash equilibria and (hopefully) to ESS's. If successive generations blend into each other, the evolution of behavioral patterns can be described by a set of ordinary differential equations.

A well-known formalization of the selection process is given by the replicator equations [14]:

$$\dot{x}_i = x_i\left((A\mathbf{x})_i - \mathbf{x}'A\mathbf{x}\right). \tag{6}$$

A point $\mathbf{x}$ is said to be a *stationary* (or equilibrium) point of our dynamical system, if $\dot{x}_i = 0$, for all $i = 1 \dots n$. A stationary point $\mathbf{x}$ is said to be *asymptotically stable* if any trajectory starting sufficiently close to $\mathbf{x}$ converge to $\mathbf{x}$.

Theorem 3 *A point $\mathbf{x} \in \Delta$ is the limit of a trajectory of (6) starting from the interior of Δ if and only if $\mathbf{x}$ is a Nash equilibrium. Further, if point $\mathbf{x} \in \Delta$ is ESS then it is asymptotically stable.*[2]

Proof. See [14].

In our approach, we let elements to be grouped compete with each other, each obtaining support from compatible elements and competitive pressure from all the others. At equilibrium, only elements that are mutually compatible should survive and are then taken to form a highly cohesive group.

[2]Note however, that in the asymmetric case it is possible to have asymptotically stable points that are not evolutionary stable.

Furthermore, at equilibrium the distribution of surviving strategies exhibits properties that are similar to those enjoyed by Shashua and Ullman's saliency vector [11]: all elements that do not belong to the group will not be played by any surviving player and, hence, will have a zero probability of being selected, while elements with a strong support, which are more "central" to the group will be chosen by a high number of players, hence having a high probability of being selected.

Thus, in order to find a cohesive group from a set of elements with affinity matrix A, we simply run (6) and at convergence elements with non-zero population will be assigned to the figure, while extinct elements will be assigned to the ground.

5. Experimental evaluation

In this section we apply the proposed clustering framework to the perceptual grouping of edge elements, or edgelets, in a noisy image. The goal of perceptual grouping is to partition detected image elements, like points or edges, into perceptually coherent groups. Here we concentrate on the task of grouping together the responses of an edge detector that belong to the same object boundary. The similarity functions found in the literature of perceptual grouping can be categorized into three main groups. The first consists of the measures that are defined in terms of co-circularity, a quantification of how far the edges are from laying on a common circumference [13, 6, 1]. The second group is defined in terms of curves of least energy [11]. Here the similarity is inversely related to the bending energy of the curve joining the two edges. The third class of measures is based on a probability distribution of shape boundary modeled in terms of a random walk [5, 15],

In this paper we apply the proposed grouping algorithm to two affinity measures. The first is a co-circularity based symmetric affinity proposed by Hérault and Horaud [1]. The second is an asymmetric measure proposed by Williams and Thornber [15], which is based on the transition-probability of a random walk between two directed edgelets, where the asymmetry comes from splitting each edge element into its two directed parts. For each measure we apply two grouping algorithms. Using Williams and Thornber's asymmetric affinities we compare the steady-state extraction proposed in [15] with our game-theoretic approach, while with Hérault and Horaud's co-circularity based affinities we compare our game-theoretic approach with the mean-field annealing approach presented in [1]. Note that the latter affinities are symmetric and hence our approach reduces to dominant-sets. However, we add the results in an attempt to separate the relative effects of the different measures and of the grouping approaches. Further, we create a new measure by rendering the one proposed by Williams and Thornber symmetric by taking the max of the

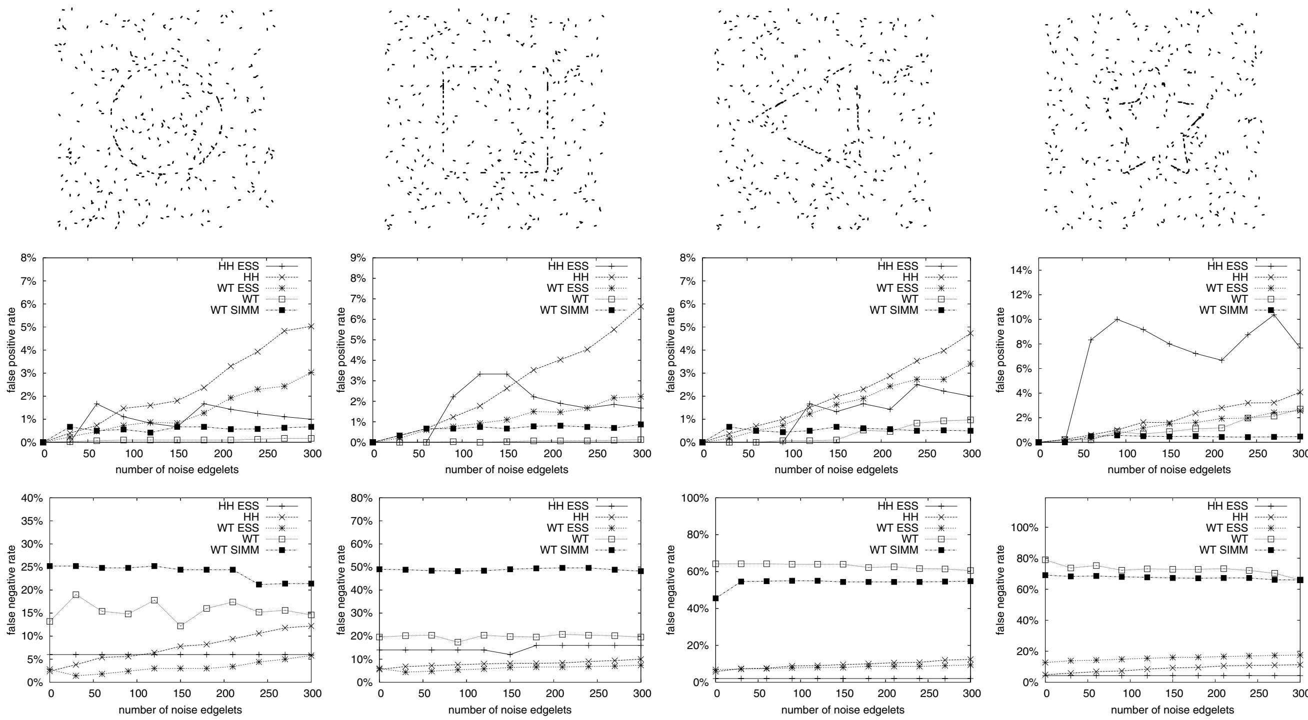

Figure 2. Grouping results on synthetic images.

compatibilities between the opposite directions of the directed edges. This way we have two symmetric matrices obtained in using different approaches that can be used with our approaches. This would allow us to factor out the relative quality of the two measures from our evaluation, and concentrate on the clustering approaches.

5.1. Sensitivity analysis

To begin, we assessed the difference in sensitivity to noise of the four approaches. To this end, we applied the grouping algorithm to images corrupted by an increasing amount of random noise, The ground-truth figures consisted of 4 simple geometrical shapes: a circle, a square, a triangle and a star. 50 edges where selected at random from the boundary of the first 3 shapes, while 70 where selected from the star. Then, we added an increasing amount of randomly positioned and oriented edgelets, from 0 to 300 added edgelets, and applied the four grouping algorithms to the corrupted images. The experiments where repeated 10 times and Figure 2 shows the average false positive (FPR) and false negative (FNR) rates of the experiments. Here HH refers to the results obtained with Hérault and Horaud's mean field annealing approach, WT the algorithm proposed by Williams and Thornber using the asymmetric affinities, while HH ESS and WT ESS refer to our approach applied on the Hérault and Horaud and Williams and Thornber's affinities, respectively. Finally, WT SIMM refers to the proposed algorithm applied to the symmetrized version of Williams and Thornber's measure.

The results show that the proposed game-theoretic approach, with each affinity measure, performs consistently better than the grouping algorithms specifically proposed for the measures. Moreover, our approach outperforms the mean-field annealing algorithm irrespective of the measure used, both in terms of FPR and FNR. The algorithm of Williams and Thornber, on the other hand, constantly selects only a small fraction of the edge elements belonging to the figure and very little belonging to the noise, thereby producing a low FPR, but a FNR that is several times larger than what we obtained with our approach.

It is interesting to note that the symmetric HH measure performs similarly to the asymmetric WT measure. Note, however, that the symmetrized WT measure used in conjunction with the proposed approach, performed much worse than its asymmetric counterpart. This means that the similarity between the performance of the HH and WT measures is due to the relative suitability of the measures to the grouping approach and should not be read as an argument against asymmetric compatibilities, since clearly making use of the asymmetry in the WT measure actually improves the results.

5.2. Textured background

Arguably, random noise is not a good model of the background of real images. In fact, any texture present in the background produces more structured responses from

Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)
0-7695-2597-0/06 $20.00 © 2006 IEEE

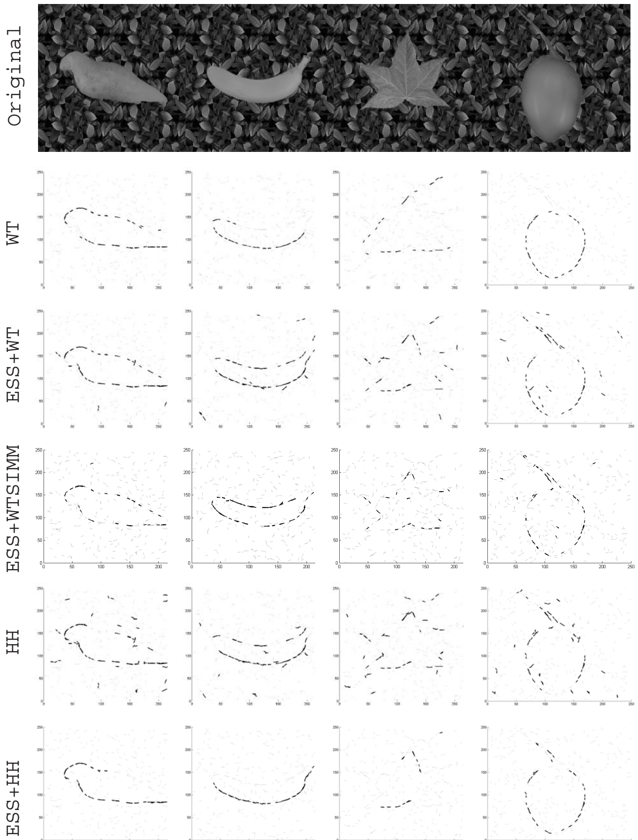

Figure 3. Grouping results on images with textured background.

an edge detector and the presence of structured noise can severely affect the performance of the clustering algorithm.

Figure 3 shows the results of applying the clustering approaches under study to a few images obtained by superimposing four hand-segmented figure over a structured background. In this case, the manual segmentation provides us with ground truth information. Here, noise is added by extracting edge elements from a texture and fusing them with the elements extracted from the figure. For each image, 70 edges where selected from the figure and 300 from the texture, hence reproducing the maximum noise-level of the previous experiment. Note that the top row of Figure 3 shows the hand-segmented images superimposed over the texture, however the edge elements are extracted separately, so that noise edgelets can be found even inside the figure's boundary. Figure 4 shows the FPR's and FNR's obtained with the three algorithms. Williams and Thornber's algorithm confirms its tendency to select only a few very consistent edges, losing many edgelets belonging to the figure and giving few false negatives but having a number of undetected edges belonging to the figure several times higher than our proposed approach. The mean field annealing approach, on the other hand, achieved a FNR comparable to our approach, but suffered from a much higher FPR assigning almost twice as many background edges to the figure. The proposed approach, on the other hand, confirms its tendency to outperform the other approaches both in terms of false positives and false negatives regardless of the affinity measure adopted. The poor performance obtained using symmetrized version of Williams and Thornber's measure confirms our idea that differences in performance using the two measures are to be ascribed to the relative suitability of the measures and not to the utility of asymmetric measures.

Finally, Figure 5 shows the result obtained by the four clustering approaches to real images. The results confirm the observations, with Williams and Thornber's approach selecting only a few edges and missing completely corners, hence approximating a somewhat "round" shape regardless of the actual shape of the figure, while the game-theoretic approach outperforms the other algorithms selecting larger parts of the main figure, and smaller or more structured parts of the background.

6. Conclusions

In this paper, we presented a new framework for grouping with asymmetric affinities derived from a game-theoretic formalization of the competition between the hypotheses of group membership, In this framework the elements belonging to the group are those that survive. We proved a characterization of the evolutionally stable points of the distribution of the group membership hypotheses in the case of both binary and weighted affinities. Interestingly, in the latter case the set of surviving hypotheses are in a one-to-one correspondence to an extension to directed graphs of dominant sets. We applied the proposed algorithm to the grouping of responses of edge detectors. Experiments on both synthetic and natural images showed that our approach outperforms both state-of-the-art clustering-based perceptual grouping approaches with symmetric compatibilities, and other approaches explicitly designed to make use of asymmetric compatibilities.

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. Ho et al., Clustering apperances of objects under varying illumination conditions. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Vol. 1, pp. 11–18, 2003.

[3] T. Hofmann and J. M. Buhmann. Pairwise data clustering by deterministic annealing. *IEEE Trans. Pattern Anal. Machine Intell.* 19:1–14, 1997.

[4] D. W. Jacobs, Q. Weinshall, and Y. Gdalyahu, Classification with nonmetric distances: Image retrieval and class representation, *IEEE Trans. Pattern Anal. Machine Intell.*, 22(6):583–600, 2000.

	WT		ESS with WT		ESS with symmetric WT		HH		ESS with HH	
	FPR	FNR	FPR	FNR	FPR	FNR	FPR	FNR	FPR	FNR
Potato	1%	45%	6%	16%	0.71%	39.71%	13%	17%	13.67%	8%
Banana	0%	45%	6%	9%	1.53%	35.2%	9%	20%	4.33%	8%
Leaf	5%	76%	6%	37%	1%	59.73%	12%	44%	14.67%	26.67%
Tamarillo	0.7%	50.4%	6.7%	19.5%	1.47%	41.33%	12.4%	24.5%	5.3%	16%

Figure 4. FPR and FNR on images with textured background.

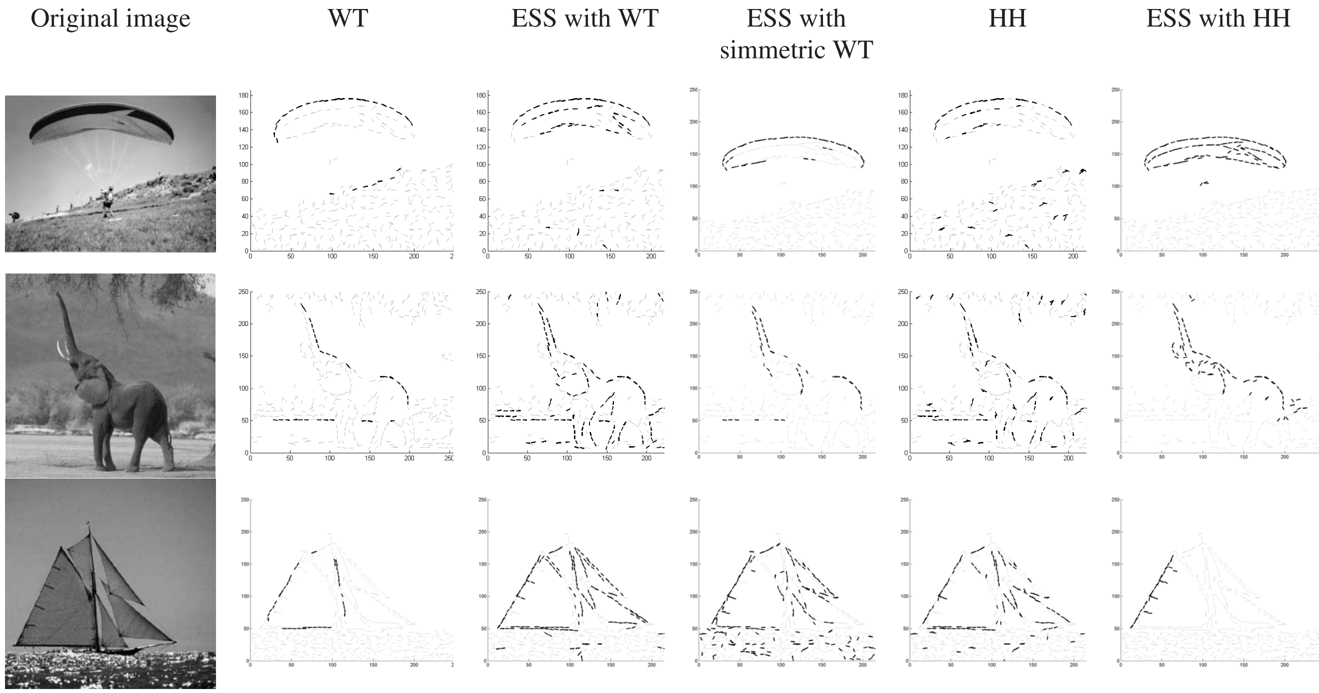

Figure 5. Grouping results on real images.

[5] D. Mumford, Elastica and computer vision. In *Algebraic Geometry and its applications*, Springer-Verlag, 1994.

[6] P. Parent and S. W. Zucker, Trace inference, curvature consistency and curve detection. *IEEE Trans. Pattern Anal. Machine Intell.*, 11:823–889, 1989.

[7] M. Pavan, A new graph-theoretic approach to clustering, with applications to computer vision. PhD thesis, Università "Ca' Foscari", Italy.

[8] M. Pavan and M. Pelillo. A new graph-theoretic approach to clustering and segmentation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 145–152, 2003.

[9] 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.

[10] 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.

[11] A. Shashua and S. Ullman, Structural saliency: The detection of globally salient structures using a locally connected network. In *Intl. Conf on Computer Vision*, 1988.

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

[13] S. Ullman, Filling-in the gaps: The shape of subjective contours and a model for their generation. *Biological Cybernetics*, 21:1–6, 1976.

[14] J. W. Weibull. *Evolutionary game theory*. MIT Press, Cambridge, MA, 1995.

[15] L. R. Williams and K. K. Thornber, A comparison of measures for detecting natural shapes in cluttered backgrounds. *Intl. J. of Computer Vision*, 34(2/3):81–96, 2000.

[16] S. Yu and J. Shi, Grouping with directed relationships. In M. Figueiredo, J. Zerubia, and A. K. Jain (Eds.), Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 283–97, Springer, Berlin, LNCS 2134, 2001

Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)
0-7695-2597-0/06 $20.00 © 2006 IEEE