

Matching Hierarchies of Segmentations *

R. Glantz⁺, M. Pelillo⁺, and W. G. Kropatsch[#]

⁺ Dipartimento di Informatica

Università Ca' Foscari di Venezia

Via Torino 155, 30172 Mestre (VE), Italy

e-mail: {glantz, pelillo}@dsi.unive.it

[#] Pattern Recognition and Image Processing Group 183/2

Institute for Computer Aided Automation, Vienna University of Technology

Favoritenstr. 9, A-1040 Vienna, Austria

e-mail: krw@prip.tuwien.ac.at

Abstract

We propose to match two hierarchies of segmentations by many-to-many mappings between the regions of the two hierarchies. The mappings preserve the order of the regions (w.r.t. set inclusion) in both hierarchies. The matching involves weights for the significance of individual regions within a hierarchy and similarity measures for the comparison of regions from different hierarchies. Irregular pyramids, in which each level consists of an attributed plane graph and an attributed dual graph are well suited to represent the hierarchies and to provide the information for computing the weights and the similarity measures.

keywords: many-to-many matching, segmentation, pyramid, graphs.

*This work is supported by the Austrian Science Foundation (FWF) under grant P14445-MAT and by MURST under grant MM09308497.

1 Introduction

Hierarchies of segmentations can be obtained from an image by a sequence of criteria for merging neighboring regions. When criterion n cannot be applied anymore, the n -th segmentation is attained. Consider two hierarchies $\mathcal{H}(I_1)$ and $\mathcal{H}(I_2)$ of segmentations with respect to the images I_1 and I_2 , respectively. We assume that the hierarchies have been constructed according to the same sequence of criteria. In this paper the *structural* similarity of I_1 and I_2 is grasped by a hierarchy-preserving many-to-many mapping between the regions of $\mathcal{H}(I_1)$ and $\mathcal{H}(I_2)$.

If we assume that the highest level of $\mathcal{H}(I_j)$ ($j = 1, 2$) contains only one region, i.e. the whole image, the partial order of the regions in each hierarchy may be described by a rooted tree, the vertices of which represent the regions (the root represents the whole image), and the edges of which represent set inclusion. Thus, we may focus on many-to-many mappings between two rooted trees that preserve the orders imposed by the rooted trees.

We use a tree matching algorithm that is based on a maximum clique formulation in a derived association graph [7]. Alterations of the region properties are taken into account by a similarity measure between regions and structural alterations are balanced by means of weights that indicate the relevance of the regions for the hierarchy.

The paper is organized as follows: In Sec. 2 we present a graph-based concept for calculating and representing nested morphological segmentation. The tree matching algorithm is explained in Sec. 3. Sec. 4 is devoted to the weights and the similarity measures for matching nested morphological segmentations. Experimental results are presented in Sec. 5.

2 Nested Morphological Segmentation

Morphological segmentation methods rely on the intuitive idea of flooding a topographic surface in order to find the watersheds and to determine the catchment basins [5]. The idea of flooding is also used to derive hierarchies of catchment basins [6]. We will first sketch how to derive the watersheds and the catchment basins by dual graph contraction [4]. Then we will construct the hierarchy of the catchment basins.

Let the topographic surface be defined by the modulus of the gradient image as in [6]. We represent the topographic surface by a dual pair (\overline{G}_0, G_0) of graphs, G_0 being plane. The vertices and edges of G_0 represent the pixels and the 4-neighborhood of the pixels, respectively. For each vertex v of G_0 let $alt(v)$ denote the altitude (modulus of the gradient) at v . The vertices and the edges of the graphs $G_0 = (V_0, E_0)$ and $\overline{G}_0 = (\overline{V}_0, \overline{E}_0)$ are equipped with attribute values $att(\cdot)$ as follows [3] (Figure 1):

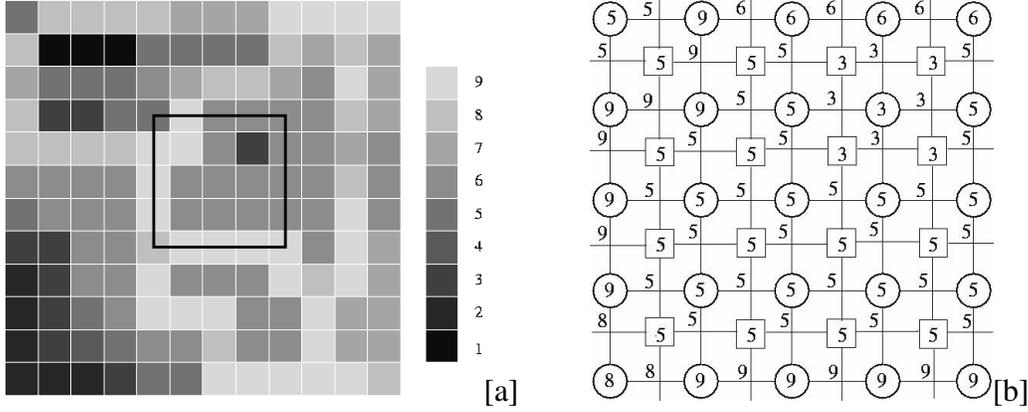


Figure 1: (a) Representation of a topographic surface by means of pixels whose gray values indicate the altitude. (b) Initial pair (\overline{G}_0, G_0) of attributed dual graphs restricted to the square in (a). The circular and square vertices belong to G_0 and \overline{G}_0 , respectively.

- $att(v) := alt(v) \quad \forall v \in V_0$,
- $att(e) := \min\{att(v) \mid v \text{ is end vertex of } e\} \quad \forall e \in E_0$,
- $att(\overline{e}) := att(e)$ for all pairs of dual edges $(e, \overline{e}) \in E_0 \times \overline{E}_0$,
- $att(\overline{v}_0) := \min\{att(\overline{e}) \mid \overline{e} \text{ has } \overline{v}_0 \text{ as an end vertex}\} \quad \forall \overline{v}_0 \in \overline{V}_0$.

A sequence of monotonic dual graph contractions [3] transforms the dual pair (\overline{G}_0, G_0) into the dual pair $(\overline{G}_{2n}, G_{2n})$. The dual pair $(\overline{G}_{2n}, G_{2n})$ obtained from Figure 1a is depicted in Figure 2a. The vertices and the edges of \overline{G}_{2n} represent the catchment basins and the neighborhood relations of the catchment basins, respectively. In accordance with [3] the contraction of G is done in a way which ensures that G_{2n} may be embedded on G_0 .

Coarser segmentations are derived from the catchment basins by unifying the basins. The unification of neighboring basins b_1 and b_2 is achieved by contracting the edge in \overline{G}_{2n} that connects the vertices represented by b_1 and b_2 . As pointed out in [6], a variety of criteria can be used for the choice of the basins to be unified first. The criteria are usually formulated by means of the basin sizes, their depths or the minimal altitude on the common border of the basins. In [3] it is proven that the altitude of the deepest point in basin b is given by the attribute of the vertex representing b in \overline{G}_{2n} . It is also shown that the attribute of each edge \overline{e} with end vertices representing b_1 and b_2 indicates the minimal altitude along that part of the border line between b_1 and b_2 which is represented by e (e and \overline{e} being a dual pair of edges).

Contracting the edges of \overline{G}_{2n} according to increasing values of $att(e_{2n})$ (e_{2n} edge of \overline{G}_{2n}) yields a hierarchy of regions as the one depicted in Figure 2b, where unification of R_i and R_j is denoted by $R_i + R_j$.

The graph \overline{G}_{2n} is contracted in subsequent parallel steps, until there exists but one vertex. The hierarchy of the regions obtained forms a so called *irregular pyramid* [4]

$$(\overline{G}_{2n}, G_{2n}), (\overline{G}_{2n+1}, G_{2n+1}), \dots, (\overline{G}_{2n+2m}, G_{2n+2m}). \quad (1)$$

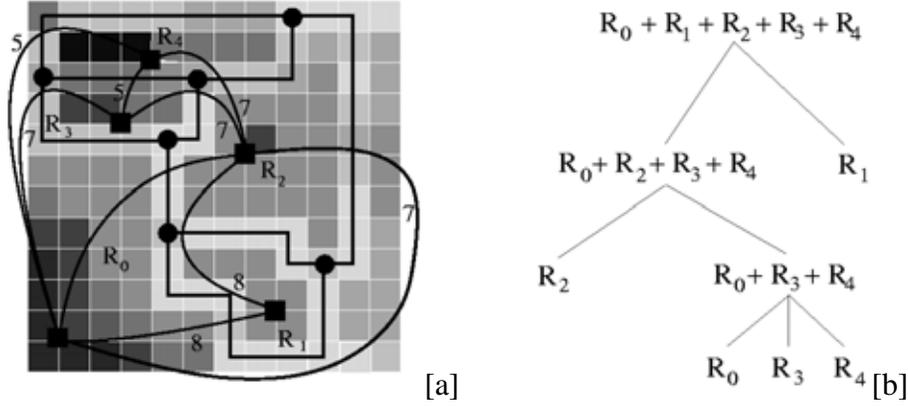


Figure 2: (a) The dual pair $(\overline{G_{2n}}, G_{2n})$. The circular vertices belong to G_{2n} and the square vertices belong to $\overline{G_{2n}}$. (b) The hierarchy of the regions from the pyramid on top of $(\overline{G_{2n}}, G_{2n})$.

The vertices of $\overline{G_{2n+2i}}$ represent the regions of the nested morphological segmentation. The order of these regions with respect to set inclusion defines the hierarchy of the regions.

3 Many-to-many Matching of Attributed Trees

To match hierarchies of segmentations we used a framework recently introduced in [8], which expands on previous work developed in [7]. The basic idea behind this approach is to cast the tree matching problem as an equivalent maximum weight clique problem. This is in turn mapped onto an equivalent quadratic program which is then (approximately) solved by simple dynamics arising in evolutionary game theory and related fields.

Formally, an *attributed tree* is a triple $T = (V, E, \alpha)$, where (V, E) is the “underlying” rooted tree and $\alpha : V \rightarrow \mathcal{A}$ is a function which assigns an attribute vector $\alpha(u)$ to each node $u \in V$. Two nodes $u, v \in V$ are said to be *adjacent* (denoted $u \sim v$) if they are connected by an edge. We shall also consider a function $\delta : \mathcal{A} \rightarrow \mathbb{R}_+$ which assigns to each set of attributes (and therefore to each node in the tree) a real positive number. This will be interpreted as the negligibility of the corresponding node in the tree. Specifically, a node will be declared “negligible” if the value of the function δ corresponding to its attributes is smaller than a fixed threshold ϵ . Clusters of nodes that contain only one non-negligible node (w.r.t. ϵ) are called ϵ -clusters. For a formal definition see [7]. We associate an ϵ -cluster of negligible nodes in the first subtree to an ϵ -cluster of negligible nodes in the second tree, thereby defining a many-to-many mapping from the first to the second tree.

A relation $M \subseteq V_1 \times V_2$ is called a *subtree ϵ -morphism* if it preserves the hierarchies of the ϵ -clusters in each of the trees. A formal definition is given in [8].

Clearly, in realistic applications, it would be desirable to find a subtree ϵ -morphism which pairs nodes having “similar” attributes. To this end, let σ be any similarity measure on the attribute space, i.e. any (symmetric) function which assigns a positive number to any pair of

attribute vectors.

If M is a subtree ϵ -morphism between two attributed trees $T_1 = (V_1, E_1, \alpha_1)$ and $T_2 = (V_2, E_2, \alpha_2)$, the overall similarity between the matched structures can be defined as follows:

$$S(M) = \sum_{(u,w) \in M} \sigma(\alpha_1(u), \alpha_2(w))$$

The ϵ -morphism M is called a *maximal similarity subtree ϵ -morphism* if we cannot add further matchings to M , while retaining the morphism property. It is called a *maximum similarity subtree ϵ -morphism* if $S(M)$ is the largest among all ϵ -morphisms between T_1 and T_2 .

The weighted ϵ -tree association graph (ϵ -TAG) of two attributed trees $T_1 = (V_1, E_1, \alpha_1)$ and $T_2 = (V_2, E_2, \alpha_2)$ is the graph $G_\epsilon = (V, E, \omega)$ where $V = V_1 \times V_2$ such that for any two nodes (u, w) and (v, z) in V the level of u in the hierarchy of T_1 equals the level of v in the hierarchy of T_2 and the same applies to the vertices w and z . Again, the levels are the levels of the corresponding clusters [7]. The following result establishes a one-to-one correspondence between the attributed tree morphism problem and the maximum weight clique problem.

Proposition 3.1 *Any maximal (maximum) similarity subtree ϵ -morphism between two attributed trees induces a maximal (maximum) weight clique in the corresponding weighted ϵ -TAG, and vice versa.*

Once the tree morphism problem has been formulated as a maximum weight clique problem, any clique finding algorithm can be employed to solve it (see [1] for a recent review). In the work reported in this paper, we used an approach recently introduced in [7, 2], which is summarized below.

3.1 Matching via game dynamics

Let $G = (V, E, \omega)$ be an arbitrary weighted graph of order n , and let S_n denote the standard simplex of \mathbb{R}^n :

$$S_n = \{ \mathbf{x} \in \mathbb{R}^n : \mathbf{e}'\mathbf{x} = 1 \text{ and } x_i \geq 0, i = 1 \dots n \}$$

where \mathbf{e} is the vector whose components equal 1, and a prime denotes transposition. Given a subset of vertices C of G , we will denote by \mathbf{x}^c its *characteristic vector* which is the point in S_n defined as

$$x_i^c = \begin{cases} \omega(u_i)/\Omega(C), & \text{if } u_i \in C \\ 0, & \text{otherwise} \end{cases}$$

where $\Omega(C) = \sum_{u_j \in C} \omega(u_j)$ is the total weight on C .

Now, consider the following quadratic function

$$f_G(\mathbf{x}) = \mathbf{x}'(\gamma\mathbf{e}\mathbf{e}' - A_G)\mathbf{x} \tag{2}$$

where $A_G = (a_{ij})$ is the $n \times n$ symmetric matrix defined as follows:

$$a_{ij} = \begin{cases} \frac{1}{2\omega(u_i)} & \text{if } i = j, \\ 0 & \text{if } i \neq j \text{ and } u_i \sim u_j, \\ \frac{1}{2\omega(u_i)} + \frac{1}{2\omega(u_j)} & \text{otherwise} \end{cases} \quad (3)$$

and $\gamma = \max a_{ij}$. The following result allows us to formulate the maximum weight clique problem as a quadratic program, thereby switching from the discrete to the continuous domain (see [2] for proof).

Proposition 3.2 *Let C be a subset of vertices of a weighted graph $G = (V, E, \omega)$, and let A_G be defined as in (3). Then, C is a maximum (maximal) weight clique of G if and only if $\mathbf{x}^C(\mathbf{w})$ is a global (local) maximizer of f_G in S_n . Moreover, all local (and hence global) maximizers of f_G on S_n are strict.*

We now turn our attention to a class of simple dynamical systems that we use for solving our quadratic optimization problem. Let W be a non-negative real-valued $n \times n$ matrix, and consider the following dynamical system:

$$\dot{x}_i(t) = x_i(t) [(W\mathbf{x}(t))_i - \mathbf{x}(t)'W\mathbf{x}(t)], \quad i = 1 \dots n \quad (4)$$

where a dot signifies derivative w.r.t. time t , and its discrete-time counterpart

$$x_i(t+1) = x_i(t) \frac{(W\mathbf{x}(t))_i}{\mathbf{x}(t)'W\mathbf{x}(t)}, \quad i = 1 \dots n. \quad (5)$$

It is readily seen that the simplex S_n is invariant under these dynamics, which means that every trajectory starting in S_n will remain in S_n for all future times. Both (4) and (5) are called *replicator equations* in evolutionary game theory, since they are used to model evolution over time of relative frequencies of interacting, self-replicating agents [9].

If $W = W'$ then the function $\mathbf{x}(t)'W\mathbf{x}(t)$ is strictly increasing with increasing t along any non-stationary trajectory $\mathbf{x}(t)$ under both continuous-time (4) and discrete-time (5) replicator dynamics. Furthermore, any such trajectory converges to a stationary point. Finally, a vector $\mathbf{x} \in S_n$ is asymptotically stable under (4) and (5) if and only if \mathbf{x} is a strict local maximizer of $\mathbf{x}'W\mathbf{x}$ on S_n .

The previous result is known in mathematical biology as the fundamental theorem of natural selection [9] and, in its original form, traces back to R. A. Fisher. Motivated by this result, we use (as in [7, 8]) replicator equations as a simple heuristic for solving our attributed tree matching problem. Let $T_1 = (V_1, E_1, \alpha_1)$ and $T_2 = (V_2, E_2, \alpha_2)$ be two attributed trees, and let $G = (V, E, \omega)$ be the corresponding association graph. By letting

$$W = \gamma \mathbf{e}\mathbf{e}' - A_G \quad (6)$$

we know that the replicator dynamical systems (4) and (5), starting from an arbitrary initial state, which is usually taken to be the simplex barycenter, will iteratively maximize the function $\mathbf{x}'W\mathbf{x}$ over the simplex and will eventually converge to a strict local optimizer which will then correspond to the characteristic vector of a maximal weight clique in the association graph. This will in turn induce a maximal similarity subtree ϵ -morphism between T_1 and T_2 .

4 Weights and Similarity Measures

Matching nested morphological segmentations in a robust way we have to take into account that there are catchment basins which are sensitive to changes of the topography and others that are more stable. The same distinction makes sense for regions obtained by unifying catchment basins. In the following we will define *weights* for regions that reflect the reliability of the regions for the matching. Due to the one-to-one correspondence between the regions of the hierarchical segmentation and the vertices in all $\overline{G_{2n+2i}}$, we may identify the regions with the vertices. The minimal attribute $Att^{min}(r)$ of all edges incident to region/vertex r , i.e.

$$Att^{min}(r) := \min\{att(e) \mid r \in \bar{t}(e)\} \quad (7)$$

indicates the next higher level of the flood that unifies r with a neighboring region. The maximal attribute Att_{max} of all edges $e \subset r$ (e and r both are subsets of \mathbb{R}^2), i.e. $Att_{max}(r) := \max\{att(e) \mid e \subset r\}$ indicates the lowest level of the flood at which all sons of r were merged. Let $size(r)$ denote the size of region r , i.e. the number of pixels in r . We define the *weight* of a region r to be

$$weight(r) = (Att^{min}(r) - Att_{max}(r)) * size(r). \quad (8)$$

The *similarity measure* will depend on the application. It can be derived from topological measurements (genus of the regions), geometric measurements of the regions (area, shape) or of the boundaries (perimeter, curvature), or the colors (gray values) of the regions.

5 Experimental Results

To check the algorithms we generated the test images depicted in Fig. 3a-c. The images are composed such that there are different pairs of neighboring regions with the same contrast. Thus, the unification of neighboring basins as defined in Sec. 2 is not unique. In these cases the choice is made by using a random generator. However, each ambiguous unification yields a region of zero weight (Sec. 4) and after contracting the negligible edges (with respect to $\epsilon = 0$) the hierarchies of the images in Fig. 3a-c should be pairwise isomorphic (trees) again. Indeed, we obtained perfect matches between the contracted hierarchies.

We also performed tests on real images. The hierarchies computed from the subimages *l-eye*, *r-eye*, *mouth*, and *nose* in Fig. 3d had 27, 31, 37, and 41 vertices, respectively. Since there is no preferred value for ϵ , we covered a wide range by choosing ϵ such that the number of clusters in the hierarchy of *l-eye* amounted to 24 (all regions with weight greater than 0), 20, 15, and 10, respectively. The corresponding values for ϵ are between 0 and 250. We did not want unreliable regions to contribute to the weights of the cliques in the ϵ -TAG. Hence, we set the weight of a vertex (u, v) in the ϵ -TAG to zero whenever the weight of u or v was smaller or equal to ϵ . In general, the number of clusters for the same ϵ is different in hierarchies from different images. Thus, we have to compensate for the different numbers of clusters if

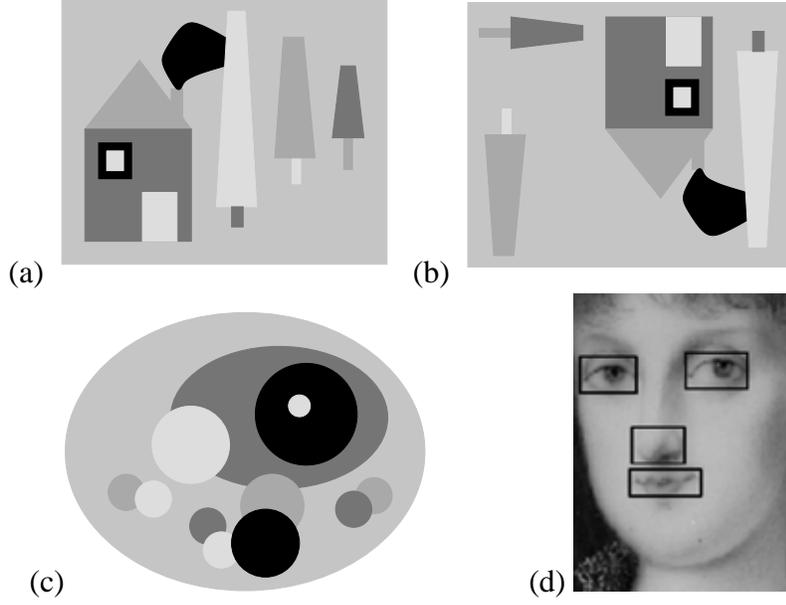


Figure 3: (a-c) Perfectly matched images. (d) The four images *l-eye*, *r-eye*, *mouth*, *nose*.

quantifying the quality of the matches. We calculated the normalized distance ¹

$$dist_{\epsilon}(T_1, T_2) := 1 - \frac{W_{\epsilon}(C_{12})}{M \max(n_{\epsilon}(T_1), n_{\epsilon}(T_2))}, \text{ where} \quad (9)$$

- T_1 and T_2 are the attributed trees of the subimages,
- $W_{\epsilon}(C_{12})$ is the weight of the maximal weight clique C_{12} in the ϵ -TAG of T_1 and T_2 ,
- $n_{\epsilon}(T_1)$ denotes the number of ϵ -clusters in T_1 , and
- M denotes the upper bound of the similarity function σ .

The similarity function is a linear function on the mean gray levels (normalized to $[0, 1]$). Tab. 5 shows the results for $\epsilon = 0$. As for all other ϵ -values tested, the two eyes are most similar, followed by the pair *l-eye* and *mouth*.

Analogous experiments were performed with the images in Fig. 4. For ϵ -values between 3000 and 5000 (see Tab. 5) the two images *pot-0* and *pot-180* have been the most similar ones. Note that the light intensities of the two images are distributed differently and that our method does not make use of shapes.

¹without proof that the metric axioms are fulfilled.

Table 1: Normalized distances of graphs from subimages of Fig. 3d for $\epsilon = 0$.

$\epsilon = 0$	l-eye	r-eye	mouth	nose
l-eye	0.00	0.27	0.42	0.58
r-eye	0.27	0.00	0.45	0.52
mouth	0.42	0.45	0.00	0.51
nose	0.58	0.52	0.51	0.00

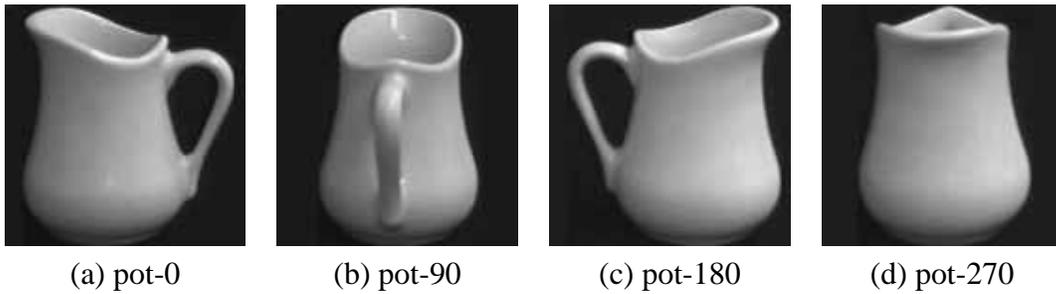


Figure 4: Images of a pot from the COIL-database.

Table 2: Normalized distances of graphs from pots in Fig. 4 for $\epsilon = 4000$.

$\epsilon = 4000$	pot-0	pot-90	pot-180	pot-270
pot-0	0.00	0.32	0.21	0.57
pot-90	0.32	0.00	0.44	0.53
pot-180	0.21	0.44	0.00	0.61
pot-270	0.57	0.53	0.61	0.00

6 Conclusions and Outlook

We proposed a combination of hierarchical segmentation followed by a many-to-many matching of the regions. This combination is well suited to detect structural similarities between images. Robustness is achieved through a weight function and a similarity function for the regions. Our method is invariant to geometrical transformations of homogeneous regions as long as the topological relations between the regions are unchanged. First experiments on real images showed that the matching results correspond to human intuition. In the future we will extend the concept such that the calculation of the hierarchy, as well as the weight and the similarity function may depend on the shape of the regions.

References

- [1] I. M. Bomze, M. Budinich, P. M. Pardalos, and M. Pelillo. The maximum clique problem. In D.-Z. Du and P. M. Pardalos, editors, *Handbook of Combinatorial Optimization (Suppl. Vol. A)*, pages 1–74. Kluwer, Boston, MA, 1999.
- [2] I. M. Bomze, M. Pelillo, and V. Stix. Approximating the maximum weight clique using replicator dynamics. *IEEE Trans. Neural Networks*, 11(6):1228–1241, 2000.
- [3] R. Glantz and W. G. Kropatsch. Plane embedding of dually contracted graphs. In *Discrete Geometry for Computer Imagery, DGCI'2000*, volume 1953 of *Lecture Notes in Computer Science*, pages 348–357. Springer, 2000.
- [4] W. G. Kropatsch. Building Irregular Pyramids by Dual Graph Contraction. *IEE-Proc. Vision, Image and Signal Proc.*, 142(6):366 – 374, 1995.
- [5] A. Meijster and J. Roerdink. A Disjoint Set Algorithm for the Watershed Transform. In *Proc. of EUSIPCO'98, IX European Signal Processing Conference*, pages 1665 – 1668, Rhodes, Greece, 1998.
- [6] F. Meyer. Graph based morphological segmentation. In Walter G. Kropatsch and Jean-Michel Jolion, editors, *2nd IAPR-TC-15 Workshop on Graph-based Representation*, pages 51–60. OCG-Schriftenreihe, Band 126, Österreichische Computer Gesellschaft, 1999.
- [7] M. Pelillo, K. Siddiqi, and S. W. Zucker. Matching Hierarchical Structures Using Association Graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11):1105–1120, 1999.
- [8] M. Pelillo, K. Siddiqi, and S. W. Zucker. Many-to-many matching of attributed trees using association graphs and game dynamics. In Carlo Arcelli, Luigi P. Cordella, and Gabriella Sanniti di Baja, editors, *Visual Form 2001*, volume 2059 of *Lecture Notes in Computer Science*, pages 583–593. Springer, 2001.
- [9] J. W. Weibull. *Evolutionary Game Theory*. MIT Press, Cambridge, MA, 1995.