OBJECT-BASED EVALUATION OF HIERARCHICAL REGION-BASED REPRESENTATIONS
BASED ON INFORMATION THEORY STATISTICAL MEASURES

Felipe Calderero, Ferran Marques

Department of Signal Theory and Communications
Technical University of Catalonia (UPC), Barcelona, Spain
{felipe.calderero, ferran.marques}@upc.edu

ABSTRACT

This work presents an evaluation in terms of object representation of the hierarchical region-based representations created by a family of general statistical region merging algorithms. These merging techniques are based on different versions of information theory statistical measures; concretely, the Kullback-Leibler divergence and the Bhattacharyya coefficient. Additionally, a significance index can be defined together with these techniques to extract the most statistically meaningful partitions at different levels of resolution. The first part of this object-based analysis evaluates the ability of the hierarchy of partitions created by each method to represent the objects in the image. The second set of experiments compares the quality of the object representation into the most significant partitions extracted from the hierarchical representations in two different cases: considering that the object may be represented by a union of regions (object oversegmentation is not penalized), and considering only the best single region representing it (oversegmentation is penalized).

1. INTRODUCTION

Object representation and recognition from images is an immediate and almost innate task for humans, but a challenging and surprisingly difficult task for artificial vision systems and computers. Similarly to the importance of perceptual grouping into human vision, image segmentation is a first and key step into object-based image analysis, as it represents a first level of abstraction, simplifying the image into something more semantically meaningful and easier to analyze. Concretely, it refers to the partitioning of the image into connected and disjoint sets of pixels (known as regions) holding a common pattern.

However, in a large number of cases, a unique solution for the image segmentation problem does not exist, i.e., instead of a single optimal partition, it is possible to find different region-based explanations of an image, at different levels of resolution. To overcome this situation, a hierarchical segmentation approach can be used where, instead of a single partition, a hierarchy of partitions is provided, generating a hierarchical region-based representation of the image.

There exist different approaches to hierarchical region-based image representations, usually associated to tree data structures. In these structures, tree nodes are related to regions in the image whereas tree links represent the inclusion property. For instance, on the literature we found quad-tree image representations [1], min-tree and max-tree representations [2], and component trees [3] or trees of shapes [4, 5] which merge both previous trees (min- and max-tree) into a single representation.

In this work, the binary partition tree (BPT) structure proposed in [6, 7] is chosen as hierarchical representation for its efficiency to represent image regions at different levels of resolution and for its simplicity to be generated by an iterative one-to-one region merging process (an example is shown in Fig. 1). The BPT is a tree-based structure where each node has either two children or no children (a leave). Given an initial partition (that can be as fine as the image pixel domain), and a region merging algorithm, the BPT stores the whole merging sequence that has been generated by the merging process. Setting the regions of the initial partition as leaves of the tree, each merging is represented by a parent node (the new region resulting from the merging) which is linked to its two children nodes (the pair of regions that are merged). Thus, the nodes on the BPT reflect the similarity between neighboring or adjacent regions, hierarchically organized under an inclusion relationship.

Hence, jointly with a region merging technique, the BPT provides a reduced set of the most significant regions into an image (from the similarity criterion point of view). Concretely, the number of nodes into the BPT is $N_{\text{nodes}} = 2N_{\text{reg}} - 1$, with $N_{\text{reg}}$ the number of regions in the initial partition. This represents a compromise between representation accuracy of regions (mainly given by the quality of the initial partition representation) and processing efficiency (mainly given by the number of nodes to analyze in the tree).

In turn, region merging techniques are an important type of hierarchical bottom-up segmentation approaches. These techniques are region-based, in the sense that they consider regions not only as the goal of the segmentation process but also as the mean to obtain a partition of the image. They are based on local decisions, directly on the region features. Starting from an initial partition or from the collection of pixels, regions are iteratively merged until a stopping criterion is reached. Thus, region merging algorithms are specified by: a merging criterion, defining the cost of merging two regions; a merging order, determining the sequence in which regions are merged based on the merging criterion; and a region model that determines how to represent the union of regions.

In this work we focus on a family of unsupervised region merging techniques, using a general region model, and providing a general strategy with no assumptions about the nature of the regions. Under a statistical framework, these approaches lead to merging criteria based on information theory statistical measures: the Kullback-Leibler divergence and the Bhattacharyya coefficient. A first set of methods was proposed in [8], based on an independent and identically distributed (i.i.d.) pixel region model, including an area-weighted and an area-unweighted version. In [9], an extension of
the previous approaches for a first order Markov region model was presented. Additionally, any of these techniques can be used together with a partition selection criteria based on a statistical significant index that can extract the most statistically meaningful partitions from the hierarchy.

Other region merging techniques in the literature lack this generality principle. For instance, approaches in [10, 11] use a criteria combination over a more simple region model and merging order. Most approaches assume either color homogeneous [10, 11] or textured regions [12], or consider simpler region models and merging orders [10, 11]. Under the statistical framework, some approaches consider different parameter-based families of probability distributions as region model [13], or include further assumptions about the nature of the regions, such as the homogeneity property in [14].

The goal of this work is to evaluate the quality in terms of object representation of the hierarchical region-based representations, concretely the BPT structures, generated by the family of general statistical region merging techniques (the four versions introduced in [8] and the four ones introduced in [9]). First, we evaluate the ability of the hierarchy of partitions created by each method to represent the objects in the image. In other words, how well the BPT nodes represent semantic objects.

In turn, we study the performance, in terms of object-based content, of the selection criterion extracting the most significant partitions from the BPT created by each method. The quality of the object representation in the selected partitions is evaluated from two different points of view: considering that the object may be represented by a union of regions (object oversegmentation is not penalized), or considering only the best single region representing the object (thus, penalizing oversegmentation).

The rest of the paper is organized as follows. Section 2 outlines the set of proposed general statistical merging criteria. Concretely, Section 2.1 reviews statistical criteria under the pixel independency assumption, while Section 2.2 presents the extension when statistical dependencies for pixels belonging to the same region are considered. Section 4 presents the evaluation in terms of objective performance of the algorithm.

2. GENERAL STATISTICAL REGION MERGING CRITERIA

Under a statistical framework, a single channel image can be considered as a realization of a 2D stochastic process. Therefore, each pixel is a sample of one of the discrete random variables composing the image process. The extension of this definition is straightforward for the multichannel case under channel independence assumption.

Hence, the image segmentation problem can be formulated as the partitioning of the image into (statistical) regions, i.e. connected and disjoint subsets of pixels that share similar statistical properties or that are more likely to have been generated by the same statistical source.

Depending on the type of assumptions about the statistical properties of image pixels, the region merging problem can be formulated from different perspectives. In Section 2.1, this problem is formally tackled considering a region as a set of independent and identically distributed pixels. Section 2.2 presents a similar development, considering statistical independence only for pixels in different regions and modeling the statistical dependency for pixels belonging to the same region as a first order finite-state Markov model.

2.1. Independent and Identically Distributed Pixel Region Model

Assuming that image pixels are statistically independent, a region can be considered as a set of independent and identically distributed (i.i.d) pixels. Therefore, a region is completely characterized by the probability distribution common to all pixels, that can be estimated by its empirical distribution.

The empirical distribution $P_n$ or type of a sequence $x$ of $n$ samples from an alphabet $X = \{a_1, a_2, \ldots, a_N\}$ is defined as the relative proportion of occurrences of each value of $X$, i.e., $P_n(a) = N(a|x)/n$ for all $a \in X$, where $N(a|x)$ is the number of times the symbol $a$ occurs in the sequence $x \in X^n$.

Using the empirical distribution as region model provides a unified and general framework for image segmentation, as arbitrary discrete distributions are directly estimated from data. Apart from pixel independency, no further assumptions are made. Moreover, this model can be easily computed and, after the union of a pair of regions, updated. In addition, the number of bins into the empirical distribution (applying a uniform quantization in $X$) can be set to optimize the performance of the algorithm.

2.1.1. Area-Weighted Kullback-Leibler Merging Criterion

The first criterion is based on merging at each step the pair of adjacent regions maximizing the probability of being generated by the same statistical distribution. We tackle this problem as a pairwise hypothesis test. Assume $R_1$ and $R_2$ are two adjacent regions with size $n_1$, $n_2$, and empirical distributions $P_1$, $P_2$, respectively, whose union would generate a new region with empirical distribution $P_{1\cup 2}$. Then, the two hypotheses considered are:

- $H_0$: Pixels in the first region, $x_1 \in R_1$, and pixels in the second one, $x_2 \in R_2$, are distributed by $P_{1\cup 2}$;
- $H_1$: Pixels $x_1 \in R_1$ are distributed by $P_1$; and pixels $x_2 \in R_2$ are distributed by $P_2$.

In general, we wish to minimize both probabilities of error. The Neyman-Pearson lemma [15] proofs that the optimal test for two hypotheses, in that sense, is the so-called likelihood ratio test:

$$\frac{P_{H_0}(x_1, x_2)}{P_{H_1}(x_1, x_2)} \geq \frac{1}{T}$$

As shown in [8], particularizing for the probability of the empirical distribution of a sequence of i.i.d. observations, the log-likelihood ratio test can be rewritten as:

$$\log \frac{P_{H_0}(x_1, x_2)}{P_{H_1}(x_1, x_2)} = -n_1 \cdot D(P_1||P_{1\cup 2}) - n_2 \cdot D(P_2||P_{1\cup 2})$$

where $D(P||Q)$ is the Kullback-Leibler divergence between two statistical distributions:

$$D(P||Q) = \sum_{a \in X} P(a) \log \frac{P(a)}{Q(a)}$$

Consequently, at each merging stage, the two adjacent regions (written as $R_i \sim R_j$) with maximum log-likelihood should be merged. We will refer to this statistical criterion as the Kullback-Leibler merging criterion (KL), formally stated as:

$$\{R_1, R_2\} = \arg \max_{R_i \sim R_j} -n_i \cdot D(P_i||P_{R_i}) - n_j \cdot D(P_j||P_{R_j})$$
This criterion is based on measuring the similarity between the empirical distributions of the regions and the empirical distribution of their merging, weighted by the size of the regions.

2.1.2. Area-Weighted Bhattacharyya Merging Criterion

This criterion is based on a direct statistical comparison between the empirical distributions of the regions. In this case, each region is considered as a class in the space defined by all possible empirical distributions for a sequence of \( n \) i.i.d. samples, centered at the point generated by its own empirical distribution. The exponent of the probability of error of such a classifier is bounded by the minimum Chernoff information between the statistical distributions of any pair of classes [15].

Redefining the probability of error of a classifier as the probability of fusion in a clustering method, at each step the pair of regions maximizing this probability can be merged. Nevertheless, instead of computing the Chernoff information (which implies an optimization process), in practice, an easily computable upper bound, known as the Bhattacharyya coefficient, is proposed:

\[
B(P_i, P_j) \triangleq -\log \left( \sum_x P_i^2(x) P_j^2(x) \right)
\]

(4)

In conclusion, a statistical clustering approach leads to the merging of the adjacent pair of regions with maximum (bound of the) probability of fusion, or equivalently, maximizing its exponent:

\[
\{ R_1, R_2 \} = \arg \max_{R_i \sim R_j} -\min(n_i, n_j) \cdot B(P_i, P_j)
\]

This method is based on a size-weighted direct statistical measure of the empirical region distributions, and we will refer to it as the Bhattacharyya merging criterion (BHAT).

2.1.3. Extension to Area-Unweighted Criteria

The previous merging costs depend on the size of the involved regions, establishing, in some sense, the confidence of the estimated empirical models. This approach assures that the resulting partitions are size consistent, meaning that the area of the regions tends to increase as the number of regions in the partition decreases. Additionally, as the fusion of smaller regions is favored with respect to larger regions, even when they are similarly distributed, area-biased methods generally present a good performance in terms of oversegmentation but suffer from oversegmentation.

To provide a better trade-off between under- and oversegmentation, while increasing the size resolution of the partition, a modified region merging method can be introduced. Under the assumption that regions are large enough to have a high confidence on the estimated distribution, the size dependency can be removed from the merging criteria and incorporated to the merging order. Hence, the area-unweighted merging criteria can be written as:

- Area-unweighted Kullback-Leibler merging criterion:

\[
\{ R_1, R_2 \} = \arg \max_{R_i \sim R_j} -D(P_i \parallel P_{(i,j)}) - D(P_j \parallel P_{(i,j)})
\]

- Area-unweighted Bhattacharyya merging criterion:

\[
\{ R_1, R_2 \} = \arg \max_{R_i \sim R_j} -B(P_i, P_j)
\]

In practice, it cannot be always assured that the distributions of all regions are perfectly estimated, specially, in early stages of the merging process. For that reason, both methods are combined with a scale-based merging order, assuring the size consistency of the partitions. The idea is to define a scale threshold for each level of resolution. Regions beyond this threshold are considered as out-of-scale and are merged with the highest priority, fusing them with their most similar region in the partition. Finally, when no out-of-scale regions remain, the algorithm continues merging in-scale regions normally. At each merging step the scale threshold is updated, and normal merging continues till new out-of-scale regions appear.

The scale threshold is defined as:

\[
T_{\text{scale}} = \alpha \cdot \frac{\text{Image Area}}{\text{Number of Regions}}
\]

(5)

i.e., regions that are smaller than a given percentage of the mean region area at the current scale are considered out-of-scale. The \( \alpha \) parameter controls the minimum resolution at each scale. Heuristically, it has been found that values around \( \alpha = 0.15 \) provide a good compromise between under- and oversegmentation.

The benefit of this approach is that the fusion of large regions is not penalized, once out-of-scale regions have been removed. All regions are equally likely to merge despite its size, because the merging cost only measures the statistical similarity of the empirical distributions, without being size biased.

2.2. First Order Markov Pixel Region Model

Under the same statistical framework, the region merging problem can be formally tackled considering that the statistical dependency is restricted to pixels belonging to the same region, which may be true in general for a correct segmentation. To simplify the statistical analysis and with the low complexity constraint required by a region merging approach, in this section the statistical dependency for the pixels belonging to the same region is modeled as a 1D first order finite-state Markov model. The reduction of the dimensionality for the Markov model is based on considering only the average pairwise dependency of a pixel on its four closest neighbors, which can be seen as the (empirical) probability transition matrix of the first order finite-state Markov process characterizing the region. This empirical pairwise pixel distribution leads to a second order statistic extensively used in texture analysis, known as cooccurrence matrix [16].

Formally, given the set of region pixels \( x \) from an alphabet \( \mathcal{X} = \{a_1, a_2, \ldots, a_{|\mathcal{X}|}\} \), their cooccurrence matrix \( P_x \in \mathbb{R}^{|\mathcal{X}| \times |\mathcal{X}|} \) is defined as the relative proportion of occurrences of each pair of pixel values of \( \mathcal{X} \) separated by a given displacement \( \Delta \), i.e.,

\[
P_x(a, b) = \frac{N(x_i = a, x_{i+\Delta} = b \mid x)}{N_p}
\]

(6)

where \( N(x_i = a, x_{i+\Delta} = b \mid x) \) is the number of times the pixel value \( a \) occurs at a given location, while the pixel value \( b \) occurs at a displacement \( \Delta \) from that location; and \( N_p \) is the total number of pairwise pixel occurrences at displacement \( \Delta \) in \( x \). Under the previous assumptions, the considered displacements are \( \Delta = \{(0, 1), (0, -1), (1, 0), (-1, 0)\} \).

Assuming that these stochastic processes can be considered as ergodic and with an equiprobable initial state distribution, a region is completely characterized by the probability transition matrix of the Markov process that generates it, directly estimated by its cooccurrence matrix.

Similarly to empirical distributions in Section 2.1, cooccurrence matrices provide a general and unified framework for image segmentation. Arbitrary discrete distributions are directly estimated from
data, incorporating spatial information not only about the region itself but also about its interactions with adjacent regions (existence of an edge), with no specific assumptions about the nature of the regions (neither homogeneity nor texture). Moreover, this model can be easily computed, and after the union of a pair of regions, updated. Again, the number of bins in each dimension of the coocurrence matrix can be set to optimize the performance of the algorithm.

2.2.1. Area-Weighted Kullback-Leibler Merging Criterion

Similarly to the i.i.d. case in Section 2.1.1, we can define a criterion based on merging at each step the pair of adjacent regions maximizing the probability of being generated by the same first order Markov process. Assuming that \( R_1 \) and \( R_2 \) are two adjacent regions with \( n_1, n_2 \) pixels, and with empirical transition matrices \( P_1, P_2 \), respectively, whose union would generate a new region with empirical transition matrix \( \bar{P}_{1/2} \), the two hypotheses considered are:

- \( H_0 \): Pixels in the first region, \( x_1 \in R_1 \), and pixels in the second region, \( x_2 \in R_2 \), are both distributed by the same first order Markov process, with probability transition matrix \( P_{1/2} \);
- \( H_1 \): Pixels in the first region, \( x_1 \in R_1 \), are distributed by the first order Markov transition matrix \( P_1 \); and pixels in the second region, \( x_2 \in R_2 \), are distributed by the first order Markov process, with transition matrix \( P_2 \).

Similarly to the Neyman-Pearson lemma for i.i.d. observations, in [17] it is proved that the best achievable error exponent for testing between two stationary and irreducible Markov sources (thus, holding for ergodic Markov processes) is given again by the so-called likelihood ratio test in Eq. (1).

Hence, for \( n_1 \) and \( n_2 \) sufficiently large, considering the probability of a first order Markov sequence, the ergodicity property and the equiprobable initial state distribution, the log-likelihood ratio can be approximated in terms of the transition matrices of the candidate regions:

\[
\log \frac{P_{H_0}(x_1, x_2)}{P_{H_1}(x_1, x_2)} \propto -n_1 D(P_1 || \bar{P}_{1/2}) - n_2 D(P_2 || \bar{P}_{1/2})
\]

(7)

where \( D(P || Q) \) is defined as the Kullback-Leibler divergence between two transition matrices \( P = (p_{ij}) \) and \( Q = (q_{ij}) \):

\[
D(P || Q) \triangleq \sum_{i,j} p_{ij} \log \frac{p_{ij}}{q_{ij}}
\]

(8)

Consequently, at each merging stage, the two adjacent regions (written as \( R_i \sim R_j \)) with maximum log-likelihood should be merged. We will refer to this statistical criterion as the **Markov Kullback-Leibler merging criterion** (M-KL), formally stated as:

\[
\{R_1, R_2\} = \arg \max_{R_i \sim R_j} -n_i \cdot D(P_i || \bar{P}_{i/2}) - n_j \cdot D(P_j || \bar{P}_{i/2})
\]

This criterion is based on measuring the similarity between the empirical probability transition matrices of the regions and the empirical transition matrix of their merging, weighted by the size of the regions.

2.2.2. Area-Weighted Bhattacharyya Merging Criterion

Identically to Section 2.1.2, the idea behind this approach is to use the Chernoff information bound on the exponent of the probability of error of a classifier, based on the maximum intersection between two distributions, as a measure of similarity, and consequently, to perform a clustering procedure based on the maximization of this bound.

Proceeding analogously to the classical derivation of the Chernoff bound for the i.i.d. case, a similar result can be obtained for first order Markov sequences [9]. The bound on the minimum exponent error can be defined as the Chernoff information between the probability transition matrices of two first order Markov processes.

Redefining the probability of error of a classifier as the probability of fusion in a merging method, and approximating the Chernoff information by the Bhattacharyya coefficient for a pair of transition matrices, defined as:

\[
B(P_i, P_j) \triangleq -\log \left( \sum_x P_i^1(x) P_j^1(x) \right)
\]

(9)

this approach leads to an analogous expression to the criteria in Eq. (10) for the i.i.d. case:

\[
\{R_1, R_2\} = \arg \max_{R_i \sim R_j} \min(n_i, n_j) \cdot B(P_i, P_j)
\]

This method is based on a size-weighted direct statistical measure of the empirical probability transition matrices, and we will refer to it as the **Markov Bhattacharyya merging criterion** (M-BHAT).

2.2.3. Extension to Area-Unweighted Merging Criteria

In this section, the previous criteria are extended, identically to Section 2.1.3, under the assumption that the number of pixels in each region is large enough to provide an accurate estimation of the region model, i.e., an empirical transition matrix close to the exact transition probabilities of the first order Markov process generating the data. The consequence of this assumption is the removal of the size-dependency from the merging criteria. The resulting methods do not suffer from area-bias but they cannot assure the size consistency of the generated partitions. In order to preserve the size consistency, area-unweighted merging criteria are accompanied by the scale-based merging order detailed in Section 2.1.3.

The expressions of the area-unweighted merging criteria for the first order Markov case are analogous to those defined for the i.i.d., but in function of the transition matrices instead of the probability density functions:

- **Area-unweighted Kullback-Leibler merging criterion:**

\[
\{R_1, R_2\} = \arg \max_{R_i \sim R_j} -D(P_i || \bar{P}_{i/2}) - D(P_j || \bar{P}_{i/2})
\]

- **Area-unweighted Bhattacharyya merging criterion:**

\[
\{R_1, R_2\} = \arg \max_{R_i \sim R_j} -B(P_i, P_j)
\]

3. BPT PARTITION SELECTION CRITERION

In this section a partition selection criterion is presented, i.e., an automatic technique to extract the most statistically significant partitions at different resolutions from the hierarchy of partitions. Although being at an early level of analysis, and without any a priori knowledge on the image, we expect statistical meaningful partitions to contain most human-representative regions, for different levels of analysis.
Fig. 1: Example of BPT image representation. (a) BPT created from the sequence of merging similarities computed by the i.i.d. version of the area-unweighted Bhattacharyya method (using 5 bins). Some correspondences between nodes and the image regions are shown. (b) Original image. (c) Initial BPT partition with 95 regions (dark blue nodes in the BPT). Partition from the set of most significant partitions with the largest number of regions but not reaching the 100 region bound. (d) First partition in importance (white nodes in the BPT), containing 3 regions, given by the significance index in Eq. (10). (e) Second partition in importance with larger number of regions than the first one (light blue nodes). It contains 41 regions.

Experimentally, we have observed that the merging similarity sequences for area-weighted and area-unweighted methods (in this case, without considering the costs of out-of-scale regions) present a similar behavior. For that reason, the proposed selection strategy can be applied to all merging techniques presented in this work.

The proposed strategy relies on selecting the partitions whose creation has lead to a significant decrease into the sequence of merging similarities. Therefore, we consider a non-decreasing version of the sequence of merging similarities \( s(n) \), where \( n \) is the number of remaining regions, defined as \( \hat{s}_{\text{min}}(n) = \min_{n \leq k < \infty} \{ s(k) \} \).

Determining the most important decrements on \( \hat{s}_{\text{min}}(n) \) provides the set of statistically significant partitions. These partitions may be ordered using some significance index. Here, we propose an importance weight based on the relative increase with respect to the current similarity value. Given \( \delta(n) = \frac{d}{dn}(\hat{s}_{\text{min}}(n)) \) and a non-increasing version of \( \delta(n) \), \( \hat{\delta}_{\text{max}}(n) = \max_{n \leq k < \infty} \{ \delta(k) \} \), the importance weight is defined as:

\[
\omega_k = \frac{\hat{\delta}_{\text{max}}(k - 1) - \hat{\delta}_{\text{max}}(k)}{\hat{\delta}_{\text{max}}(k)}
\] (10)

Examples of automatically selected partitions from the BPT using this significance order are shown in Figure 1 and 2. Note that, in general, the first selection corresponds to a coarse partition, whose regions are good approximations of the objects. Usually, the second proposal gives a finer partition with most representative regions in the scene.

4. EVALUATION RESULTS

Experiments in this section were performed over a set of 100 images from the Corel\textsuperscript{©} image database [10]. The set contains 10 images of 10 different complexity classes: tigers, horses, eagles, mountains, fields, cars, jets, beaches, butterflies and roses. The objects in this Corel\textsuperscript{©} subset (160 in total) have been manually segmented in the context of the SCHEMA project (http://www.iti.gr/SCHEMA/), and the resulting object partitions are used as ground truth in the experiments.

The statistical region merging techniques were applied on an initial partition of the original image in HSV color space, computed using the watershed algorithm. From these merging sequences, the BPT image representations were automatically generated. The initial BPT partitions (whose regions correspond to the leaves of the tree) were selected from the set of most statistically meaningful partitions given by the significance index in Eq. (10). Concretely, the statistically significant partition with the largest number of regions but not reaching the 100 region bound was chosen.

To evaluate the similarity with respect to the object partitions, we use the distances defined in [18]. Initially, a symmetric distance is proposed \( d_{\text{sym}}(P, Q) \). This distance is defined in terms of the minimum number of pixels whose labels should be changed between
Table 1: Evaluation of the quality in terms of object representation of the automatically selected BPT initial partitions. Columns (from left to right): mean asymmetric distance with respect to object partitions, and mean value of the minimum symmetric distance presented by a region into the BPT initial partition.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KL Area-weighted</td>
<td>0.1013</td>
<td>0.7567</td>
</tr>
<tr>
<td>2. BHAT Area-weighted</td>
<td>0.1060</td>
<td>0.7384</td>
</tr>
<tr>
<td>3. KL Area-unweighted</td>
<td>0.1134</td>
<td>0.6407</td>
</tr>
<tr>
<td>4. BHAT Area-unweighted</td>
<td>0.1176</td>
<td>0.6167</td>
</tr>
<tr>
<td>5. Markov KL Area-weighted</td>
<td>0.0985</td>
<td>0.7745</td>
</tr>
<tr>
<td>6. Markov BHAT Area-weighted</td>
<td>0.0961</td>
<td>0.7502</td>
</tr>
<tr>
<td>7. Markov KL Area-unweighted</td>
<td>0.1122</td>
<td>0.6292</td>
</tr>
<tr>
<td>8. Markov BHAT Area-unweighted</td>
<td>0.1294</td>
<td>0.6342</td>
</tr>
</tbody>
</table>

Table 2: Evaluation of the quality in terms of object representation provided by the BPT. Mean symmetric distance between manually segmented objects and the best nodes in the BPT.

<table>
<thead>
<tr>
<th>Asymmetric Distance</th>
<th>First Select.</th>
<th>Second Select.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KL Area-weighted</td>
<td>0.2250</td>
<td>0.1701</td>
</tr>
<tr>
<td>2. BHAT Area-weighted</td>
<td>0.2422</td>
<td>0.2195</td>
</tr>
<tr>
<td>3. KL Area-unweighted</td>
<td>0.2452</td>
<td>0.2315</td>
</tr>
<tr>
<td>4. BHAT Area-unweighted</td>
<td>0.2342</td>
<td>0.2294</td>
</tr>
<tr>
<td>5. Markov KL Area-weighted</td>
<td>0.2092</td>
<td>0.1855</td>
</tr>
<tr>
<td>6. Markov BHAT Area-weighted</td>
<td>0.1901</td>
<td>0.1348</td>
</tr>
<tr>
<td>7. Markov KL Area-unweighted</td>
<td>0.2548</td>
<td>0.1588</td>
</tr>
<tr>
<td>8. Markov BHAT Area-unweighted</td>
<td>0.2495</td>
<td>0.1558</td>
</tr>
</tbody>
</table>

Table 3: Evaluation of the quality in terms of object representation given by the first and second (with largest number of regions than the first one) most meaningful partitions from the BPT. Mean asymmetric distance between manually segmented objects and proposed partitions.

To complete the evaluation, the last experiment studies the representation of the object by a single region of the most meaningful partitions. In this case, both under- and oversegmentation are penalized. Hence, for each object, the region of the proposed partition with minimum symmetric distance to the object was selected as its best representative. Results are summarized in Table 5 for the first and second partitions in significance. In general, objects in the first partition suffer from oversegmentation as can be seen from the fact that the mean values of symmetric distance are larger than the unity. This means that the number of pixels that must be relabeled to obtain the same object partition is larger than the size of the object. On the contrary, the objects in the second proposal tend to be oversegmented, and hence, the amount of relabeled pixels is smaller than the object size (regions are smaller in general in those partitions).
Table 4: Evaluation of the quality in terms of object representation given by the first and second (with larger number of regions than the first one) most meaningful partitions from the BPT. Mean number of regions of the selected partitions.

<table>
<thead>
<tr>
<th>Number of Regions</th>
<th>First Select.</th>
<th>Second Select.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KL Area-weighted</td>
<td>6.69</td>
<td>23.91</td>
</tr>
<tr>
<td>2. BHAT Area-weighted</td>
<td>5.48</td>
<td>22.60</td>
</tr>
<tr>
<td>3. KL Area-unweighted</td>
<td>14.28</td>
<td>35.75</td>
</tr>
<tr>
<td>4. BHAT Area-unweighted</td>
<td>9.79</td>
<td>27.20</td>
</tr>
<tr>
<td>5. Markov KL Area-weighted</td>
<td>5.19</td>
<td>21.66</td>
</tr>
<tr>
<td>6. Markov BHAT Area-weighted</td>
<td>4.74</td>
<td>17.93</td>
</tr>
<tr>
<td>7. Markov KL Area-unweighted</td>
<td>14.52</td>
<td>43.68</td>
</tr>
<tr>
<td>8. Markov BHAT Area-unweighted</td>
<td>11.36</td>
<td>33.49</td>
</tr>
</tbody>
</table>

Table 5: Evaluation of the quality of the object representation given by the first and second (with larger number of regions than the first one) most significant partitions from the BPT. Mean value of the minimum symmetric distance presented by a region into the selected partitions.

<table>
<thead>
<tr>
<th>Symmetric Distance</th>
<th>First Select.</th>
<th>Second Select.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KL Area-weighted</td>
<td>1.0535</td>
<td>0.5774</td>
</tr>
<tr>
<td>2. BHAT Area-weighted</td>
<td>1.4238</td>
<td>0.6321</td>
</tr>
<tr>
<td>3. KL Area-unweighted</td>
<td>1.0503</td>
<td>0.5841</td>
</tr>
<tr>
<td>4. BHAT Area-unweighted</td>
<td>1.1291</td>
<td>0.7952</td>
</tr>
<tr>
<td>5. Markov KL Area-weighted</td>
<td>1.4361</td>
<td>0.6170</td>
</tr>
<tr>
<td>6. Markov BHAT Area-weighted</td>
<td>1.6736</td>
<td>0.6413</td>
</tr>
<tr>
<td>7. Markov KL Area-unweighted</td>
<td>0.5672</td>
<td>0.5709</td>
</tr>
<tr>
<td>8. Markov BHAT Area-unweighted</td>
<td>0.8580</td>
<td>0.5785</td>
</tr>
</tbody>
</table>

(iv) of the same example. Consequently, the best region representing the object in that partition corresponds to a region larger than the object, as can be seen in column (v). Nevertheless, a region correctly representing the object and with similar size appears into the second meaningful partition (column (vi)).

Apart from this general behavior, it is remarkable that Markov area-unweighted versions present similar values for first and second meaningful partitions. In other words, the quality of the object representation is well maintained as the number of regions is reduced. Particularly for the case of the Kullback-Leibler version, the performance is even improved, which means that undersegmentation does not increase significantly and oversegmentation is reduced when going up into the hierarchical region-based representation.

6. REFERENCES

Fig. 2: Evaluation of the object representation into the BPT and into the most statistically significant partitions. Columns (from left to right): (i) original image, (ii) object partition, (iii) best BPT node representing the object, (iv) first significant partition, (v) best region representing the object into the first significant partition, (vi) second significant partition. The methods used in each row are (descending order): i.i.d. area-weighted KL (1st row), i.i.d. area-weighted BHAL (2nd), i.i.d. area-unweighted KL (3rd), i.i.d. area-unweighted BHAL (4th), Markov area-weighted KL (5, 6, 7th), Markov area-weighted BHAL (8th), Markov area-unweighted KL (9th row), Markov area-unweighted BHAL (10, 11th rows). In each case, empirical distributions or empirical probability transition matrices were quantized to 5 bins.