# Class-Specific, Top-Down Segmentation

Eran Borenstein and Shimon Ullman*

Dept. of Computer Science and Applied Math
The Weizmann Institute of Science
Rehovot 76100, Israel
{boren, shimon}@wisdom.weizmann.ac.il

**Abstract.** In this paper we present a novel class-based segmentation method, which is guided by a stored representation of the shape of objects within a general class (such as horse images). The approach is different from bottom-up segmentation methods that primarily use the continuity of grey-level, texture, and bounding contours. We show that the method leads to markedly improved segmentation results and can deal with significant variation in shape and varying backgrounds. We discuss the relative merits of class-specific and general image-based segmentation methods and suggest how they can be usefully combined.
**Keywords:** Grouping and segmentation; Figure-ground; Top-down processing; Object classification

## 1 Introduction

A major goal of image segmentation is to identify structures in the image that are likely to correspond to scene objects. Current approaches to segmentation mainly rely on image-based criteria, such as the grey level or texture uniformity of image regions, as well as the smoothness and continuity of bounding contours. In this work we describe a segmentation method that is guided primarily by high-level information and the use of class-specific criteria. The motivation for using such class-based criteria to supplement the traditional use of image-based criteria in segmentation has two parts. First, it stems from the fact that although recent image-based segmentation algorithms provide impressive results, they still often fail to capture meaningful and at times crucial parts. Second, evidence from human vision indicates that high-level, class-based criteria play a crucial role in the ability to segment images in a meaningful manner (e.g. [11],[10],[9],[8]), suggesting that the incorporation of such methods will help improve the results of computer vision segmentation algorithms.

Figure 1 demonstrates some of the major difficulties encountered by image-based segmentation algorithms. An appropriate segmentation algorithm should group together the dark and light regions of the horse (left), and separate the man from the horse, despite the grey level similarity (right). Figure 2 shows

* This research was supported by the Israel Ministry of Science under the Scene Teleportation Research Project and by the Moross Laboratory at the Weizmann Institute of Science.

A. Heyden et al. (Eds.): ECCV 2002, LNCS 2351, pp. 109–122, 2002.

the results obtained when applying a state of the art image-based segmentation algorithm [15] to these and other images. The segmentations demonstrate some of the inherent difficulties of an image-based approach, including the splitting of object regions and the merging of object parts with background regions. These shortcomings are due to unavoidable ambiguities that cannot be solved without prior knowledge about the object class at hand. This paper presents an approach that addresses these difficulties, in which a simple representation of object classes in memory is used to guide the segmentation process, leading to markedly improved segmentation of images containing familiar objects (Fig. 3).

The overall structure of the paper is as follows. Section 2 briefly reviews past approaches. Section 3 provides an overview of our approach, and how information about object shapes is represented in memory and used for segmentation. Section 4 describes the approach in detail: 4.1-4.3 describe the segmentation criteria used by our method. 4.4 describes the algorithm that segments class-images according to these criteria. Section 5 shows results, Sect.6 contains a final discussion and conclusions.

**Fig. 1.** Segmentation difficulties: the same object can contain markedly different regions (left), while neighboring objects may contain regions that are similar in color and texture (right).

## 2 Brief Review of Segmentation Approaches

Most of the current approaches to segmentation rely primarily on image-based criteria, such as color, grey level, or texture uniformity of image regions (e.g. [16],[4],[2]); the smoothness and continuity of their bounding contours (e.g. [7]); or a combination of these (e.g. [6]). The region-based approaches merge and split image regions according to specific criteria. Merging approaches recursively merge similar regions (e.g. [14],[1]). "Divide & Conquer" approaches recursively split regions into distinct sub-regions (e.g. [12],[15]). Contour-based approaches emphasize the properties of region boundaries, such as continuity, smoothness, length, curvature, and shape.

Somewhat closer to our goal is the work on deformable templates (e.g. [18]), where the template parameters are used to match a model with an object image. This approach usually assumes approximate initial correspondence between

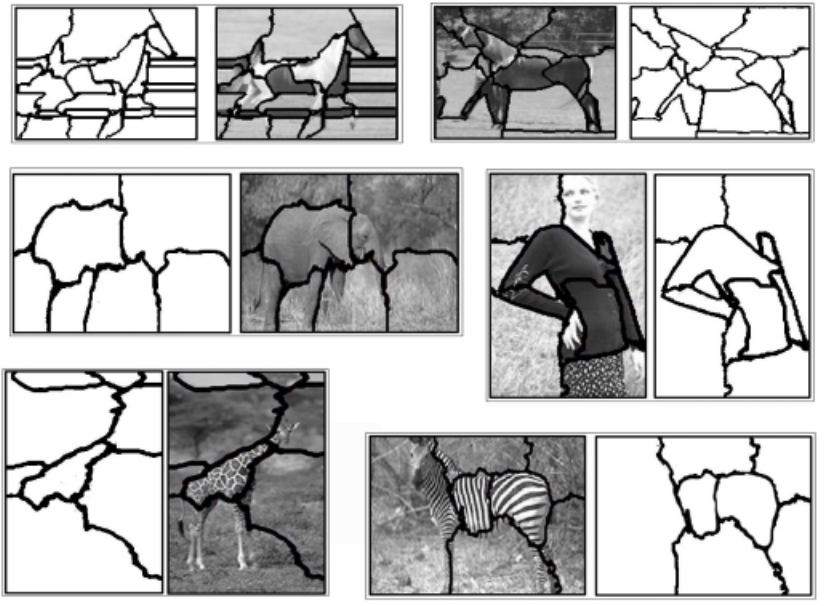

**Fig. 2.** Typical results of low-level segmentation. Objects are broken into sub-regions, and some of these sub-regions are merged with the background. (More results can be seen at `http://www.cs.berkeley.edu/~doron/software/ncuts/results/`)

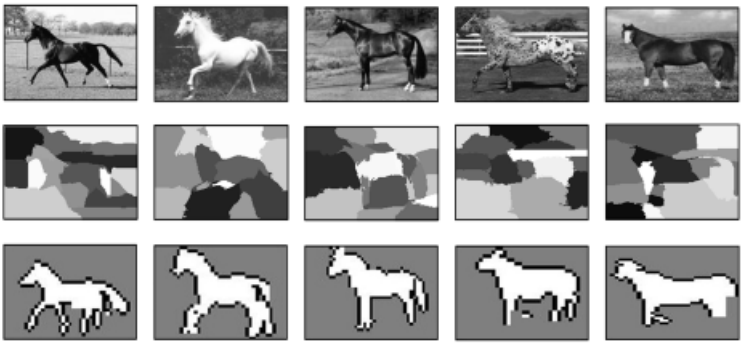

**Fig. 3.** Low-level vs. class-specific segmentation of horse images. Top: input images ($170x120$ pixels). Middle: low-level segmentation into sub-regions, as given by the normalized cuts algorithm [15]. Bottom: figure-ground segmentation map given by the algorithm described in this paper (input resolution reduced to $40x30$).

the model and its image and therefore it can be used as a final stage in our segmentation but not as the main segmentation process.

Recent psychological studies of image segmentation by humans indicate strong effects of prior familiarity with specific shapes on image segmentation. For example, a number of behavioral studies have shown that subjects are more likely to regard a familiar region as "figure" than a less familiar region, indicating that object recognition facilitates segmentation [11],[10]. Developmental studies [9],[8] show that figure-ground segregation in four month-old infants is also affected by the presence of already familiar shapes.

## 3 Overview of the Approach

The general thrust behind our class-based segmentation approach is to use known shape characteristics of objects within a given class to guide the segmentation process. The main difficulty in this approach stems from the large variability of shapes within a given class of objects. We would like to somehow capture the common characteristics of a set of shapes within a class, for example horse images, and then use this information to segment novel images. To address this problem our approach uses a fragment-based representation of object classes. Similar fragments have been used in the past for object classification [13][17] but not for segmentation. Given an image containing a certain object, we use fragments previously extracted from images of the same object class to produce a consistent cover of the novel object. This cover defines a figure-ground map that associates each pixel in the input image with the likelihood of belonging to an object or background. (Fig. 4).

The construction of an object by fragments is somewhat similar to the assembly of a jigsaw puzzle, where we try to put together a set of pieces such that their templates form an image similar to a given example. A common strategy is to start with the easiest pieces (e.g. corners) and proceed by connecting additional pieces that match in shape, color, edges, texture, etc. In some cases, as information accumulates along this process, pieces must be replaced: locally these pieces provide good matches, but the global structure adds constraints that reject the local matches.

The next several sections describe our segmentation algorithm in detail. We first describe the fragment representation — how the class fragments are represented in memory, and how they are extracted from sample images.

### 3.1 Fragment Representation in Memory

In this section we describe the fragment-based representation used for segmentation. The goal of this representation is to cover as closely as possible the images of different objects from a given class, using a set of more primitive shapes. We therefore need to identify useful "building blocks," a collection of components that can be used to identify and delineate the boundaries of objects in the class. To find such common components we look for image fragments that are strongly correlated with images containing the desired object class — they show a high

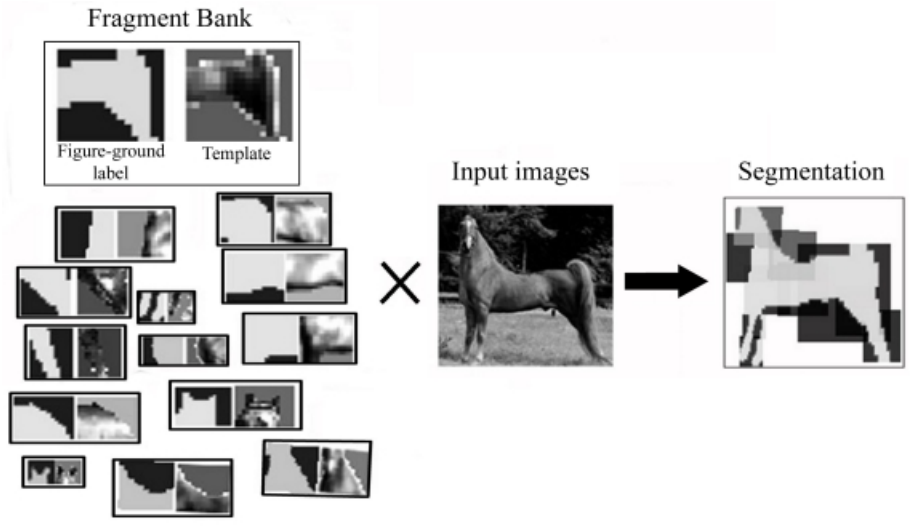


**Fig. 4.** Overview of the approach. Given an input image containing an object from a given class (in this case, a horse) we use class fragments stored in memory to produce a cover of the target image. Each fragment consists of a template and a figure-ground label. The cover defines the figure-ground segmentation.

similarity measure with regions from images containing this desired object class but not with others. Based on the Neyman-Pearson decision theory, optimal fragments can be defined as fragments with maximal frequency (hit rate) within the class, subject to the constraint that the frequency of false detection in non-class images (false alarms) does not exceed a fixed limit (e.g. [3]).

Our search for optimal fragments therefore proceeds in three stages. Stage 1 starts from a set of training images divided into class images (C) and non-class images (NC) and then generates a large number of candidate fragments. We simply extract from the images in C a large number of rectangular sub-images, these sub-images can vary in size and range from $\frac{1}{50}$ to $\frac{1}{7}$ of the object size. In stage 2, which is the crucial step, we compare the distribution of each fragment in the class and non-class training images. For a given fragment $F_i$, we measure the strength of the response $S_i$ in C and NC. $S_i$ is defined in a standard way: we correlate $F_i$ with each image $I$ in C and NC (normalized correlation) and take the maximum value over $I$. To reach a fixed level of false alarms $\alpha$ in non-class images we determine a threshold $\theta_i$ for $F_i$ by the criterion:

$$p(S_i > \theta_i | NC) \leq \alpha \tag{1}$$

This has the advantage of automatically fixing an optimal detection threshold for each fragment. In stage 3 we order the fragments by their hit rate $p(S_i > \theta_i | C)$ and select the $K$ best ones where $K$ determines the size of the fragment set.

To be used for segmentation, we add two factors to each fragment: a *figure-ground label* and a *reliability value*. The figure-ground label marks each pixel

in the fragment as figure or ground. The figure label is learned at present by comparing the fragment to the source image in the data base from which it was derived. We assume that in this limited set the figure-ground information has been estimated, for example by relative motion of the figure with respect to background. An alternative is to label pixels in the fragment according to their grey level variability in the class database. Figure pixels are similar across images and have low-variability, while background pixels show high variability.

The reliability of a fragment measures the extent to which the fragment is class-specific, measured by its hit rate $p(S_i > \theta_i|C)$. These two factors are essential to our segmentation process, as will be later demonstrated.

**Fig. 5.** Fragment representation in memory: Each fragment is represented by its grey level template (left) and figure-ground label (right).

## 4 Segmentation by Optimal Cover

The main stage of the class-based segmentation algorithm consists of covering an image with class-based fragments and using the cover to delineate the figure boundaries. To accomplish this we seek an optimal *cover* of the image in terms of the fragments. A cover is an assignment of fragments to positions in the image $I$, with each fragment being defined as either "present" in the image along with its designated position $p_i$ in the cover or "absent" from the image:

$$I \longrightarrow^{\text{cover}}_{i=1\ldots K} \begin{cases} f_i = 1, p_i = (x_i, y_i) \\ f_i = 0, p_i = \phi \end{cases} \tag{2}$$

Given a cover, we can compute the quality of the cover, which is a function of the individual match of fragments with the image, the consistency of the cover, and the reliability of the participating fragments. The following sections describe each of these factors in turn and an algorithm to find a locally optimal cover according to these criteria.

### 4.1 Individual Match

The individual match measures the similarity between fragments and the image regions that they cover. We use a similarity measure that combines region correlation with edge detection. This combination which is not commonly used,

is designed to meet the special requirements of segmentation as illustrated in Fig. 7.

A key feature of our fragment representation is the elimination of background noise. Using the figure-ground label it is possible to exclude background pixels from the similarity measure, thereby reducing background noise. Since this exclusion results in the loss of the contour features, an edge detector is added to capture the transition from figure to ground (Fig. 7). The edge presence is computed by using a normalized correlation between the figure-ground boundary and edges in the image (Fig. 6). The similarity measure $s_i\,(p, I)$ between fragment $F_i$ at image position $p = (x, y)$ and an image $I$ is defined in (3) and consists of two factors. The first factor – $N_{\mathrm{cor}}$ is the standard normalized correlation between the fragment and the image region, restricted to template pixels labeled as figure. The second term – $E_{\mathrm{dge}}$ is the edge detector response.

$$s_i(p, I) = w \cdot N_{\mathrm{cor}}(p, I)|_{\substack{\mathrm{Object}\\ \mathrm{Pixels}}} + (1 - w) \cdot E_{\mathrm{dge}}(p, I) \tag{3}$$

**Fig. 6.** The edge template (right) is derived from the boundary of the figure-ground label. The response to this template is the edge part $E_{\mathrm{dge}}$ in (3).

### 4.2 Consistency

In covering the image by shape fragments, the fragments should not only provide good local matches, but should also provide a consistent global cover of the shape. We therefore use a consistency criterion for the cover in the segmentation process. Since the fragments are highly overlapping, we define a consistency measure $c_{ij}$ between a pair of overlapping fragments $F_i$ and $F_j$ that is proportional to the fraction of pixels labeled consistently by the two fragments (4). The maximum term in the denominator prevents overlaps smaller than a fixed value $\mu_{ij}$ from contributing a high consistency term. This value is set to $\frac{1}{10}$ of the maximum possible overlap size between the two fragments. Fig. 8 demonstrates two cases of overlapping horse fragments : one in which all the pixels are labeled consistently and one in which some of the pixels are labeled inconsistently.

$$c_{ij} = \frac{\#\ \text{Consistent Overlapping Pixels}}{\max(\text{Total Overlap}, \mu_{ij})} \tag{4}$$

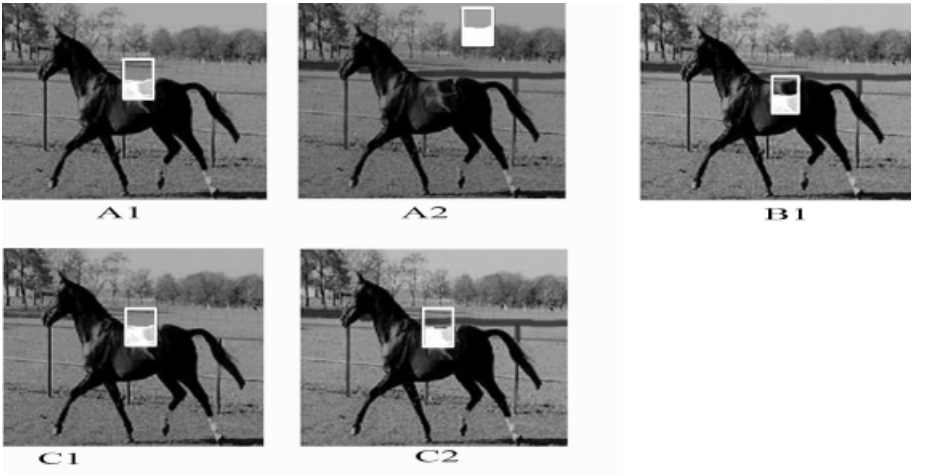


**Fig. 7.** The examples show why a combination of templates and boundary similarity are useful for segmentation. A1 – An image of a horse and a template of its back. A2 - Changes in the background reduce the similarity measure (measured using both the figure and ground parts of the template), causing the template to be more similar to another image region. B1 – In this example the template fit was measured using the figure part of the template only. This reduces background effects but also results in the loss of boundary information leading to inaccurate matches. C1,C2 – Adding an edge detector to the similarity measure yields a more stable similarity measure resulting in an accurate placement of the fragment regardless of background noise.

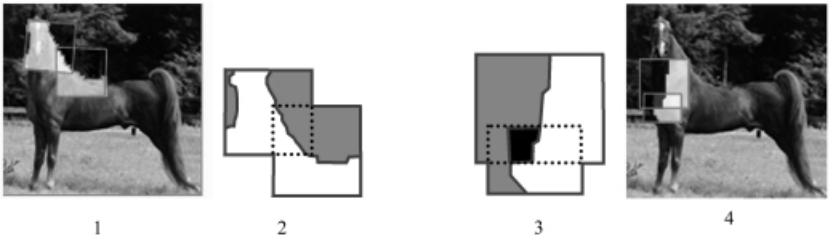


**Fig. 8.** Consistent (left) and inconsistent (right) cover by overlapping fragments. 1,4: the fragments, placed over the images, 2,3: figure-ground assignment of the fragments. Figure pixels are marked white, background pixels are grey. The inconsistent region is marked in black.

### 4.3 Fragment Reliability

The final criterion in determining an optimal cover is that of fragment reliability. Similar to a jigsaw puzzle, the task of piecing together the correct cover can be simplified by starting with some more "reliable" fragments, or anchor fragments, and then proceeding with less reliable fragments that are consistent with the initial cover (Fig. 9). Reliable fragments typically capture some distinguishing features of the shapes in the class and are unlikely to be found anywhere else. A fragment's reliability is therefore evaluated by the likelihood ratio between the detection rate and the false alarm rate. As explained, we set the minimal threshold such that the false alarm rate does not exceed $\alpha$. We can therefore express this ratio using the detection rate and $\alpha$:

$$r_i = \frac{p(S_i > \theta_i|C)}{p(S_i > \theta_i|NC)} = \frac{\text{detection rate}}{\alpha} \tag{5}$$

**Fig. 9.** Reliable fragments guide the covering. Reliable fragments are used first (middle), and subsequently completed by less reliable ones (right).

### 4.4 The Cover Algorithm

A cover on an image by shape fragments determines a segmented figure. Among all possible covers we seek a cover (2) that maximizes the three criteria above, namely, individual match quality (3), consistency (4) and reliability (5). These three factors are therefore combined in the cover score:

$$cs = \underbrace{\sum_i r_i \cdot s_i \cdot f_i}_{\substack{\text{Individual match} \\ \text{and reliability}}} + \underbrace{\frac{1}{\lambda} \sum_{i,j} \beta_{ij} \cdot f_i \cdot f_j}_{\text{Consistency}} \tag{6}$$

The first term combines the match quality and reliability of the fragments, and the second penalizes inconsistent overlapping pairs. We define the interaction term $\beta_{ij}$ between overlapping fragments as $(c_{ij} - \beta) \cdot (r_i s_i + r_j s_j)$ where $c_{ij}$ is the pairwise consistency defined above (4) and $\beta$ a global constant that determines the magnitude of the penalty for insufficient consistency. For non-overlapping

pairs $\beta_{ij}$ is defined as 0. The contribution of a single fragment $F_{\mathrm{k}}$ in this expression is obtained by summing up all the terms in (6) for which $i = \mathrm{k}$. When a fragment contribution is negative, the score is improved by removing the fragment from the cover. Negative contribution indicates poor consistency of the fragment with other fragments and can happen only when $c_{kj} < \beta - \lambda$ for at least one $j$. In our implementation $\beta = 0.65, \lambda = 0.1$.

The algorithm is iterative, but a small number of iterations (typically 2-3) are used. It is described in the Appendix, but the main stages are summarized next. At each stage, a small number $M$ of good candidate fragments are identified. A subset of these $M$ fragments, that maximally improve the current score, are selected and added to the cover. In addition, existing fragments that are inconsistent with the new match are removed. We use a small number of candidates ($M = 15$) that allows us to check all $2^M$ subsets and select the one with the highest score. The algorithm is guaranteed to converge to a local maximum since the score is bounded and increases at each iteration. To initialize the process, we select a sub-window within the image with the maximal concentration of reliable fragments. The similarity of all the reliable fragments is examined at 5 scales at all possible locations – giving a complexity which is linear in the number of reliable fragments, the number pixels of each image scale, and the number of scales. Given this information it is possible to pick the most reliable window and use the matched fragments inside as the initial $M$ candidates for the cover. If the combined evidence from the reliable fragments falls below a classification threshold, the process terminates without producing a cover. In a system containing multiple classes that compete for segmentation (rather than just horse images), the class with the highest evidence will initiate the cover.

## 5 Experiments

We tested the algorithm on a database containing horse images. A bank of 485 fragments was constructed from a sample library of 41 horse containing images of size ($40x30$) for which the figure-ground information was manually segmented. For each fragment we estimated $p(S_i|C)$ and $p(S_i|NC)$ by measuring the distribution of the fragments' similarity measure with 193 low-resolution images of horses and 253 low-resolution images of non-horses. Using these estimated distributions, the fragments were assigned their appropriate threshold and classified to 146 reliable and 339 non-reliable fragments. The algorithm was then tested on 176 novel horse images ($40x30$ pixels). Examples are shown in Fig. 10 and compared to the results of a normalized-cuts segmentation algorithm [15] tested on the same images but with higher resolution ($170x120$). The algorithm obtains high-quality segmentations of figure from background for a variety of images. The algorithm can deal successfully with shape variations using a fixed repertoire of fragments extracted from the training set. The generalization to novel shapes is based in part on the use of multiple alternative fragments for the same object region, and in part in the flexibility in the fragments' arrangement. We also compared qualitatively the agreement between the figure regions produced by the algorithm and the figure region judged by humans. This can be expressed

by evaluating the ratio $r = \frac{|S \cap F|}{|S \cup F|}$ where $F$ is the human-segmented figure and $S$ is the algorithm's segmented figure. $|F|, |S|$ are the sizes of $F, S$ and the size of the entire image $|I| = 1$. The maximal value of $r$ is $r = 1$ obtained only for perfect segmentation. The average score for the current algorithm was $r = 0.71$. The normalized-cuts algorithm, segmenting the images into two segments – figure and ground, gives much lower average score ($r = 0.31$). The last value can also be compared with random segmentation (where $|S|$ pixels are chosen randomly to be figure) , which gives an average of $r = 0.23$. The most problematic figure regions were the horse legs, where variability is high. The initialization step was the most time consuming (about 30 seconds), where the similarity of every reliable fragment was examined on 5 different scales of the target image. Usually the algorithm converged after 2-3 iterations, with each iteration taking a few seconds, giving a total time of about 40 seconds per target image. We used Matlab 6.0 program on a Windows-NT, Pentium-600Mhz platform. The complexity of the algorithm is linear in the number of scales, number of fragments, and size of the fragment bank.

## 6 Discussion

The approach proposed in this paper emphasizes the role of high-level information or class-specific criteria in image segmentation. We present a class-specific segmentation method that successfully addresses ambiguities inherent to segmentation schemes based exclusively on image-based criteria.

Segmentation is obtained by covering the image with a subset of class-specific fragments and using this cover to delineate the figure boundaries. These fragments serve as class-specific shape primitives, or "building blocks," and are used to handle a large variety of novel shapes within the class. The fragments are represented by a template together with a figure-ground label, and a reliability score.

The last two factors in this fragment representation are novel characteristics of our approach and are essential for segmentation. The figure-ground label is essential for: (a) constructing the figure-ground segmentation map; (b) defining a robust similarity measure that reduces background noise; and (c) defining the consistency between fragments. The fragment reliability enables us to detect key fragments that serve as initial anchors in the cover construction. In order to construct a complete cover we use both reliable and less reliable fragments. The less reliable fragments are constrained by the consistency relations of the cover and can therefore be used to complete the segmentation of difficult regions.

Compared with other schemes that apply segmentation and classification in sequence, in our scheme the two processes are intimately linked. The initial evidence from reliable class fragments is used to select the most likely class that serves to initiate the cover, and the final classification decision may depend on the segmentation result.

The algorithm results in the segmentation of images into two regions, figure and ground, in contrast with image-based segmentation algorithms that usually segment the image into multiple regions. The extraction of objects from these

**Fig. 10.** Segmentation results arranged in 3 groups of 4 rows. First row in each group: input images. Second row: results obtained from low-level segmentation. Third row: class-based segmentation to figure and ground. Fourth row: segmentation superimposed on the input images (boundary in red).

images is challenging because objects are often broken into sub-regions, and some of these sub-regions merge with the background. The main difficulty faced by the current algorithm lies in covering highly variable parts, such as the horse's legs and tail. This difficulty is due in part to image processing limitations and the use of relatively low-resolution images. In the future we intend to use higher resolution images or a pyramid of image segments at different scales.

Compared with the class-based segmentation, traditional image-based segmentation methods have two advantages. First, when they detect the correct figure boundaries, they can determine these boundaries with higher accuracy since they are guided directly by image discontinuities. Second, image-based algorithms are general and do not require class-specific information. The relative merits of class-specific and image-based segmentation methods suggest that they can be usefully combined into an integrated scheme. For example, image-based segmentation can be used to identify salient regions and direct class-based segmentation to these regions. At the final segmentation stage, figure boundaries produced by class-based segmentation could be refined by image-based methods (e.g. [5]) resulting in a robust and accurate delineation of object boundaries that cannot be achieved by either method alone.

## 7 Appendix: The Class-Based Segmentation Algorithm

Pre-processing:

- For all reliable fragments $F_i$, compute $s_i(p_j, I)$ for all image positions $p_j$ and all scales (see(3)).
- Set $s(p_j) = \max_i [r_i \cdot s_i(p_j)]$ (best fragment at $p_j$).
- For each image window $W$:
  Pick in $W$ at most $M$ positions $p_j$ with maximal values of $s(p_j)$. Define the score of the window $W_{\text{score}} = \sum_{p_j} s(p_j)$.

Initialization:

- Choose window $W_{\max}$ to be the window with maximal $W_{\text{score}}$ together with its fragments. These fragments compose the initial covering candidates $B_{\text{cand}}$.
- Set the current cover $B_{\text{cover}}$ to be empty.

Choosing the new covering fragments:

- With all fragments ($F_j \in B_{\text{cover}}$) fixed to $f_j = 1$, assign the candidate fragments $F_i \in B_{\text{cand}}$ to $f_i = \{0, 1\}$ such that (6) is maximized.
- Add to $B_{\text{cover}}$ all candidate fragments with $f_i = 1$ and remove all the fragments that reduce the score of (6).
- Use $B_{\text{cover}}$ to construct the figure ground segmentation map.
- If $B_{\text{cover}}$ did not change, then stop.

Updating the candidate fragments sets:

- From all fragments in $W_{\max}$ select the $M$ with the highest score (individual match and consistency with $B_{\text{cover}}$).
- Go to choosing the covering fragments.

## References


1. K. Cho and P. Meer. Image segmentation from consensus information. *Computer Vision and Image Understanding: CVIU*, 68(1):72–89, 1997.
2. J.M.H. du Buf, M. Kardan, and M. Spann. Texture feature performance for image segmentation. *Pattern Recognition*, 23:291–309, 1990.
3. R. Duda, P. Hart, and D. Stork. Pattern classification, 2001.
4. A.K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24:1167–1186, 1991.
5. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1:321–331, 1987.
6. T. Leung and J. Malik. Contour continuity in region based image segmentation. In *Fifth Euro. Conf. Computer Vision*, Freiburg, Germany, 1998.
7. D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *IEEE Conf. on Computer Vision and Pattern Recognition*, San Francisco, 1985.
8. A. Needham. Object recognition and object segregation in 4.5-month-old infants. *Journal of Experimental Child Psychology*, 78:3–24, 2001.
9. A. Needham and R. Baillargeon. Effects of prior experience in 4.5-month-old infants' object segregation. *Infant Behaviour and Development*, 21:1–24, 1998.
10. M.A. Peterson. Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science*, 3:105–111, 1994.
11. M.A. Peterson and B.S. Gibson. Shape recognition contributions to figure-ground organization in three-dimensional displays. *Cognitive Psychology*, 25:383–429, 1993.
12. M. Pietikainen, A. Rosenfeld, and I. Walter. Split and link algorithms for image segmentation. *Pattern Recognition*, 15(4):287–298, 1982.
13. E. Sali and S. Ullman. Combining class-specific fragments for object classification. In *Proc. 10th British Machine Vision Conference*, volume 1, pages 203–213, 1999.
14. E. Sharon, A. Brandt, and R. Basri. Fast multiscale image segmentation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 70–77, South Carolina, 2000.
15. J. Shi and J. Malik. Normalized cuts and image segmentation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 731–737, 1997.
16. A. Tremeau and N. Borel. A region growing and merging algorithm to color segmentation. *Pattern Recognition*, 30, No. 7:1191–1203, 1997.
17. S. Ullman, E. Sali, and M. Vidal-Naquet. A fragment based approach to object representation and classification. In *Proc. of 4th international workshop on visual form*, Capri, Italy, 2001.
18. A. Yuille and P. Hallinan. Deformable templates. In *A. Blake and A. Yuille, editors, Active Vision*, pages 21–38, MIT press, 1992.