segmentation fault
(UNIX OS)

Appendix E

Color Segmentation Using Region Merging

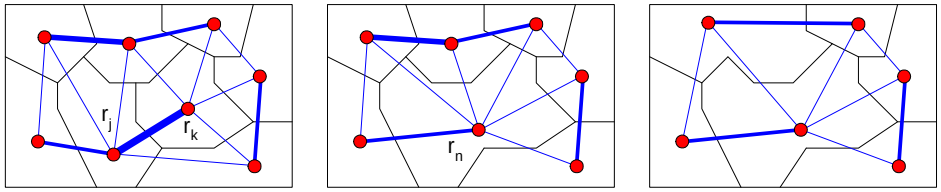
E.1 Introduction

Color segmentation is the process of segmenting an image into homogeneously-colored regions. In most cases, a color segmentation alone cannot yield a semantically good segmentation, but it can be useful as a preprocessing step to transform an image into a set of regions that are processed further in successive stages. We apply region merging in the automatic segmentation step of our object-model detection algorithm (Chapter 9). Furthermore, in Chapter 10, we integrate object-model knowledge and motion information in the region-merging algorithm to obtain a semantically meaningful result.

In this appendix, we briefly introduce the region-merging algorithm and present several merging criteria and evaluate their performance in terms of noise robustness and subjective segmentation quality. Furthermore, we introduce a new merging criterion yielding a better subjective segmentation quality, and propose to change to merging criterion during processing to further increase the overall robustness and segmentation quality.

E.1.1 The region-merging algorithm

The objective of region merging is to group image pixels to *regions* which are similar with respect to a predetermined criterion. The algorithm proceeds by sequentially merging the two most similar neighbouring regions.



(a) Initial regions. (b) Merging step 1. (c) Merging step 2.

Figure E.1: *Two steps of a region-merging process. Thicker edges represent more similar regions.*

The merging process stops when no more regions are found with sufficient similarity, or the minimum number of regions is reached.

Region merging can be viewed as an algorithm working on a graph, where the *nodes* represent regions of pixels and the *edges* indicate a neighbouring relationship. We assign *edge weights* to the edges to represent the dissimilarity between adjoining regions. Let $P = \{p_i\}$ be the set of pixels in the input image with corresponding luminance $I(p_i)$. Furthermore, let a neighbouring-relation $n(p_i, p_j)$ be true iff p_i and p_j are neighbours; however note that $n(p_i, p_i) =$ false.

The input of the region-merging algorithm is a set of regions $R = \{r_i\}$ with $r_i \subset P$, $\bigcup r_i = P$ and $r_i \cap r_j = \emptyset$ for $i \neq j$. The initial set of regions can be obtained in a several different ways. The regions can be the result of a preceding segmentation step such as watershed segmentation, they can be chosen arbitrarily (e.g. blocks of fixed size), or in the extreme case, each input pixel can be considered as a separate region.

The algorithm first builds a neighbourhood graph $G = (R, E)$ with edges $E = \{(r_i, r_j) \,|\, \exists p_k \in r_i, p_l \in r_j : \; n(p_k, p_l) = \text{true}\}$. Additionally, we define an edge weight w on the edges $w\colon E \rightarrow \mathbb{R}$ which describes a measure of dissimilarity of the regions connected by the edge. The definition of the edge weights (i.e. the merging criterion) is the crucial part of the algorithm, directly affecting the quality of the segmentation result.

Region merging is a greedy algorithm following the intuitive process to continuously merge the two most similar regions into a single region. Merging stops when the lower bound of regions $\#r_{min}$ is reached or the minimum edge weight exceeds a threshold w_{max}. The algorithm is outlined in Algorithm 3 and illustrated in Figure E.1.

Algorithm 3 Basic merging algorithm

```
while |R| > #r_min do
  e_min = (r_j, r_k) ← argmin_{e∈E} w(e)
  if w(e_min) > w_max then
    STOP
  else
    Join regions r_n ← r_j ∪ r_k
    Update edges E ← E∪(E_new = {(r_n, r_i)|(r_j, r_i) ∈ E∨(r_k, r_i) ∈ E})
    Remove old edges E ← E ∩ {R \ {r_j ;r_k}} × {R \ {r_j ;r_k}}
    Remove regions r_j and r_k  R ← R \ {r_j ;r_k}
    for all e ∈ E_new do
      Update edge weight w(e)
    end for
  end if
end while
```

E.2 Merging criteria

A merging criterion consists of two parts: a *region model*, describing each image region with a set of features, and a *dissimilarity measure*, defining a metric on the features of the region model. The range of possible region models reaches from simple models like uniform luminance up to texture, shape or motion parameters. In the following, we will concentrate on low-level features which are applied at early stages of the algorithm. Furthermore, we only consider greyscale images. However, all presented critera can be readily generalized to work on color images.

The better a region model matches the real image-data, the longer the minimum edge-weights remain small and the steeper is the relative increase in region dissimilarity as soon as the segmentation has reached its final state. This makes the segmentation process more robust to the selection of the fixed threshold for the stopping condition.

E.2.1 Mean luminance difference

The simplest region model is to describe each region r_i by its mean luminance μ_i. A straightforward possibility to define a dissimilarity measure for this model is to use the squared difference, from now on referred to as the *Mean*-criterion

$$w_{ij}^{M} = \big(\mu_i - \mu_j\big)^2. \tag{E.1}$$

E.2.2 Ward's criterion

Another measure which operates on the mean-luminance model is the *Ward*-criterion [193]. The idea is to consider the model error for a region r_i, defined as $\mathcal{E}_i = \sum_{p \in r_i} \big(I(p) - \mu_i\big)^2$. The dissimilarity associated with a pair of regions is defined as the additional total error that is introduced by merging the two regions: $w_{ij}^W = \mathcal{E}_{ij} - \mathcal{E}_i - \mathcal{E}_j$ (with $\mathcal{E}_{ij}$ being the error after a hypothetical merge of r_i and r_j). After elementary simplifications, this can be expressed as

$$w_{ij}^W = \frac{|r_i| \cdot |r_j|}{|r_i| + |r_j|} \big(\mu_i - \mu_j\big)^2. \tag{E.2}$$

E.2.3 Mean/Ward mixture

As will become clear in the following section, neither the Mean-criterion nor the Ward-criterion produce a subjectively approporiate segmentation. A better criterion may be a compromise between the characteristics of Mean and Ward. For this reason, we introduce the geometrical mean of both criteria ($w_{ij}^G = (w_{ij}^M \cdot w_{ij}^W)^{1/2}$) as a new *Mean-Ward* criterion. Since the absolute value of the criterion is not important, the square-root can be ignored, resulting in

$$w_{ij}^G = \frac{|r_i| \cdot |r_j|}{|r_i| + |r_j|} \big(\mu_i - \mu_j\big)^4. \tag{E.3}$$

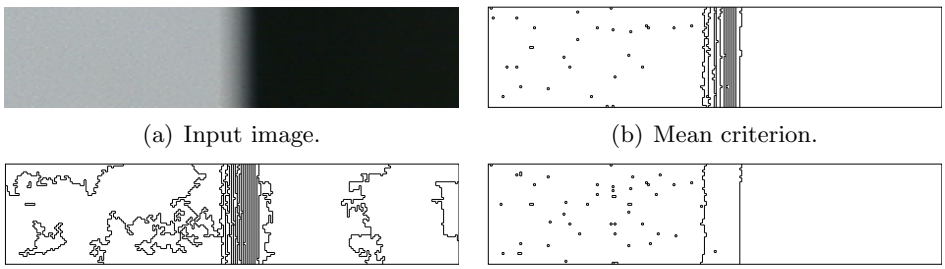
E.2.4 Linear-luminance model

Because of illumination effects, natural images seldomly consist of completely homogeneous regions. Almost all regions that we perceive as homogeneous, contain a small luminance gradient. Therefore, it is sensible to use a region model that is capable of describing slowly varying luminance gradients. A possible region model defines the luminance distribution as $I'(x, y) = \alpha + \beta x + \gamma y$ with the three parameters α, β, γ. For each individual region, these parameters are estimated from the image data, using a least-squares approach. Comparable to the Ward-criterion, we define the model error as $\mathcal{E}_i^L = \sum_{p \in r_i} \big(I(x, y) - I'(x, y)\big)^2$ and the region dissimilarity as

$$w_{ij}^L = \mathcal{E}_{ij}^L - \mathcal{E}_i^L - \mathcal{E}_j^L. \tag{E.4}$$

E.2.5 Border criterion

Although the linear-luminance model handles well most regions occurring in natural images, the model has two main drawbacks: it is rather compu-



(a) Input image. (b) Mean criterion.

(c) Ward criterion. (d) Linear-Luminance criterion.

Figure E.2: *Performance of several criteria for an unsharp edge (50 regions remaining).*

tationally intensive, and it still cannot handle all cases of small luminance variations. Especially curved surfaces have complicated luminance distributions. Both problems can be circumvented by using the following *Border* criterion.

Let $B_{ij} = \{(p_k, p_l)\}$ be the set of pairs of pixels along the common boundary between region r_i and r_j (with $p_k \in r_i$ and $p_l \in r_j$). We define the *Border*-criterion as the sum of squared differences along the boundary

$$w_{ij}^B = \frac{1}{|B_{ij}|} \sum_{(p_k,p_l)\in B_{ij}} \big(I(p_k) - I(p_l)\big)^2. \tag{E.5}$$

Note that this criterion only considers how the regions fit together along the border, not considering the interior of the region area.

E.3 Criteria properties

E.3.1 General behaviour

Figure E.2 depicts a detail view of an image containing an unsharp edge. As the image is part of a real-world image, it contains a hardly visible luminance gradient in the "flat" image regions and some camera noise. The image has been segmented independently with the Mean, Ward, and Linear-Luminance criterion until only 50 regions were remaining.

It is easily visible that the Ward criterion favours the removal of small noisy areas instead of combining large, but only slightly different regions. This occurs because the Ward criterion considers the total error and small differences in very large regions outweigh larger differences in very small regions. The Mean criterion does not show this effect, because it does not take region size into account. Similarly, the Linear-Luminance criterion

criterion / problem class	Mean	Ward	Mean-Ward	Linear-Lum.	Border	Water-shed
noise	−	++	+	−	−	−
blurred edges	−	+	+	+	−−	+
double edges	−	+	+	−	−	+
illumination	−	−−	+	++	++	+
subjective eval.	+	−−	++	+	+	−
stopping criterion	−−	+	++	+	+	N/A
comp. complexity	+	+	+	−−	+	++

Table E.1: *Comparison of the performance of different criteria on a number of typical problem classes.*

can adapt its model to approximate the gradient with sufficient accuracy. Furthermore, it is also capable to model the unsharp edge itself and does not lead to the oversegmentation with many narrow regions, as is the case with the other two criteria.

E.3.2 Comparison

In natural image segmentation, several classes of commonly occuring difficulties can be identified. The robustness of each criterion[1] on the problem classes was evaluated and is depicted in Table E.1. In the following, some problem classes are described in more detail.

- **Noise.** Camera noise has a well visible effect at the beginning of the segmentation process. Dissimilarity measures which are normalized to their region sizes, like Ward's criterion, give superior results, because single noisy pixels introduce no large overall error.

- **Blurred edges and double edges.** Objects which are out of camera focus appear with blurred edges in the image. This can lead to an oversegmentation into many thin rings around the object boundary. The Linear-Luminance criterion can approximate the blurred edge with a single region if the object boundaries are straight lines. Curved boundaries can be handled by the border criterion.

  However, the more general model of the Border criterion has the disadvantage to ignore the pixels inside a region. Thus, it is possible for an object to grow along its unsharp border until it is completely merged with the background (see Figure E.4(c)).

[1] See section E.4.1 for more information on the watershed-presegmentation column.

Almost all sharp edges in real-world images consist of pixels midway between the colors of the two regions. After a segmentation with a too low threshold, objects seem to have double edges. Because of its tendency to merge small regions, the Ward-criterion can remove these double edges well.

- **Illumination effects.** Large regions in natural images usually have differing brightness because the lighting is not strictly uniform. If the region model assumes a uniform region color, these areas are split into several pieces (see Fig. E.5). This is not the case if the region model allows brightness gradients, or if only the border between regions is considered.

- **Subjective evaluation.** As can be seen in Figure E.3, the Ward criterion has a tendency to split large regions into several segments, whereas the Mean criterion removes large regions equally likely as small regions. Figure E.3(d) shows the segmentation using the Mean-Ward criterion. The result is much more subjectively pleasing as the large regions are preserved, and much of the text is kept. The same effect is shown in the natural image in Figure E.6.

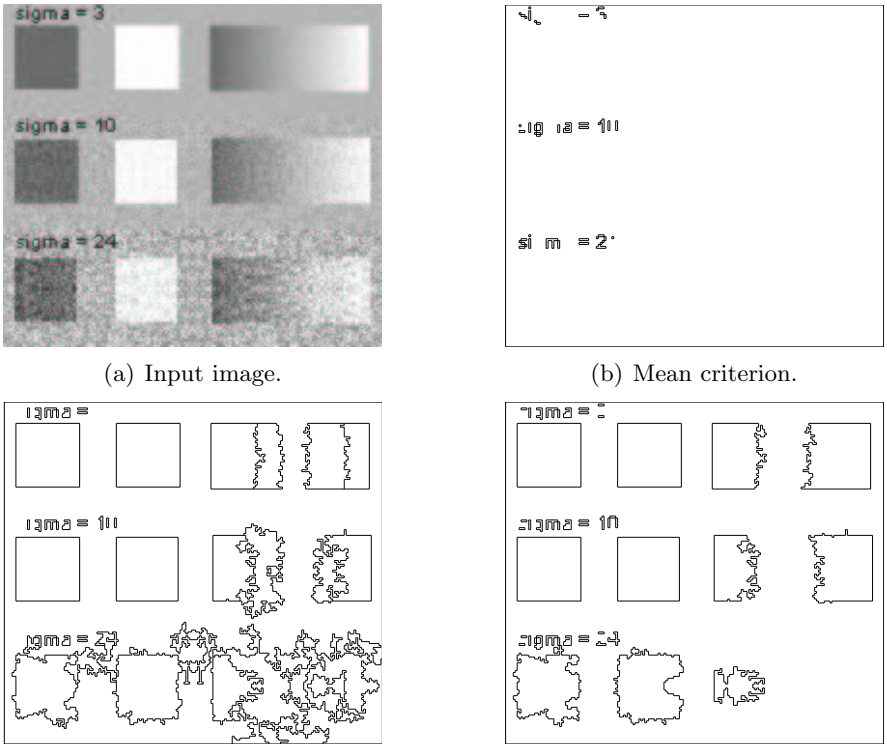
E.4 Multi-stage merging

As discussed in the last sections, each criterion shows both advantages and disadvantages. Choosing a single criterion for the complete segmentation process results in a dissatisfactory segmentation. This motivates a multi-stage approach. A criterion is used as long as it can well handle the current configuration. Afterwards, the criterion is exchanged by another one. By using several stages, the selection of an appropriate threshold in the stopping criterion is not critical. The threshold should be chosen sufficiently low to ensure that control is passed on to the next criterion, before the situation exceeds the capabilities of the criterion's region model. A sequence of criteria that produced good segmentation results was:

1. **Ward**, removing much of the image noise and eliminating double edges,

2. **Mean-Ward**, which does the main work, before finally

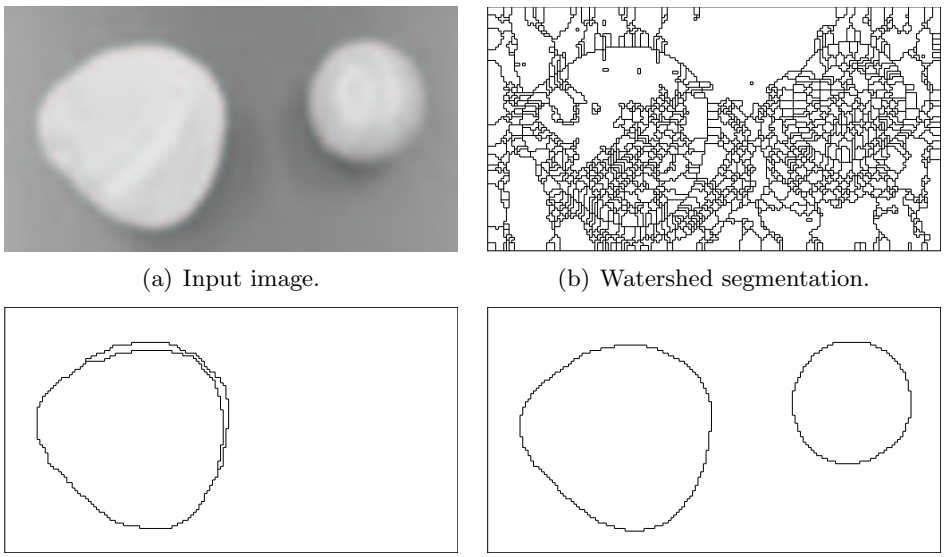
3. **Border** merges regions in which illumination effects play a central role.



(a) Input image. (b) Mean criterion.

(c) Ward criterion. (d) Mean-Ward criterion.

Figure E.3: *Segmentation results for three criteria, 50 regions remaining.*



(a) Input image.

(b) Watershed segmentation.

(c) Without watershed. The right object is removed by the Border criterion.

(d) With watershed. The right object is correctly segmented due to a lower threshold.

Figure E.4: *Effect of applying a watershed presegmentation.*

E.4.1 Applying a watershed presegmentation

Instead of starting the algorithm with single-pixel regions, it is possible to perform a presegmentation with the watershed algorithm on a gradient map of the input image.

This presegmentation has the advantage to considerably reduce the computational complexity, as the watershed transform is a fast algorithm and reduces the amount of input regions. Furthermore, it alleviates the problem of the Border criterion to destroy complete objects having unsharp boundaries (see Figure E.4). The watershed transform splits these blurred areas at the object boundaries along the position of the maximum gradient into only two regions. For this reason, the threshold in the stopping condition for the Border criterion can be set to a lower value.

The disadvantage of applying this presegmentation is that small image structures may be deteriorated or even vanish. Additionally, in the presence of camera noise, smooth edges in the image can become "fuzzy" in the segmentation.

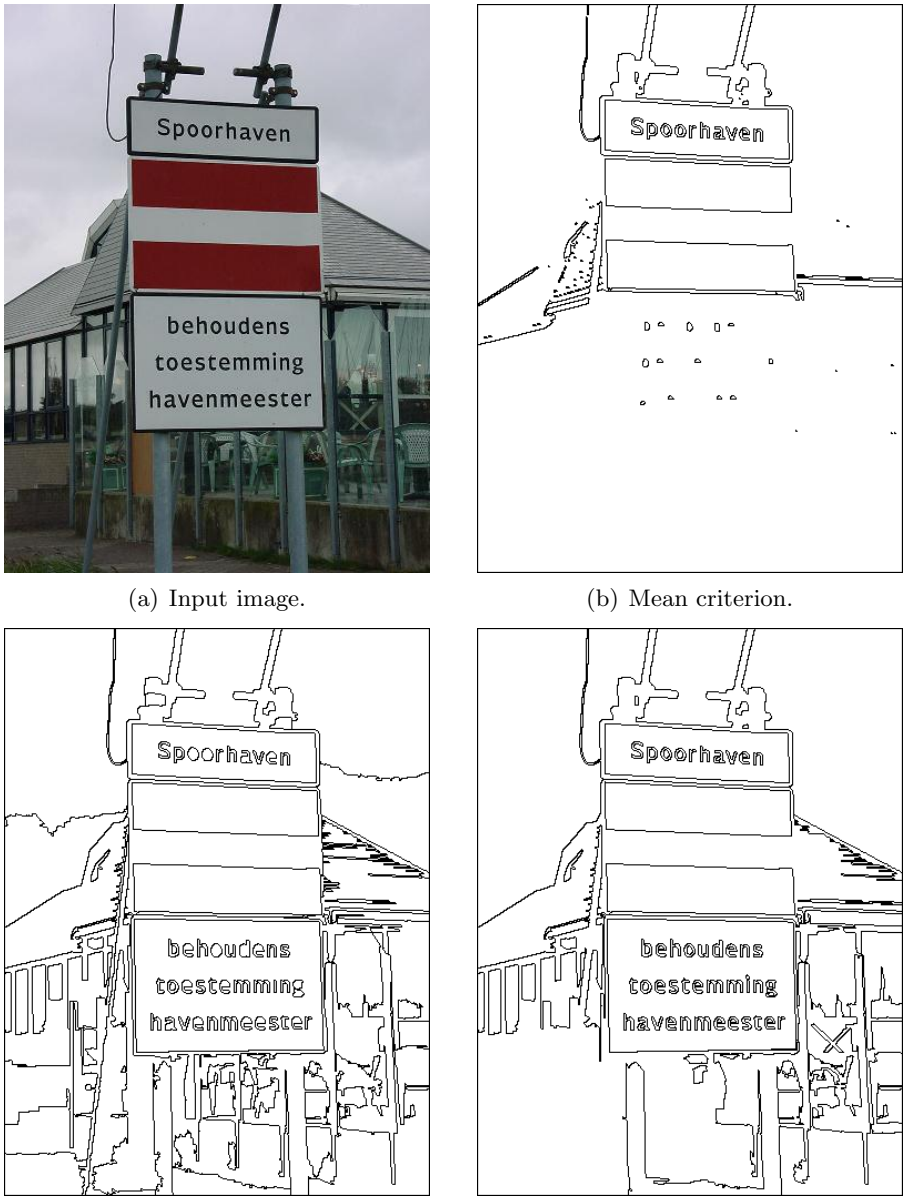
(a) Input image. (b) Segmentation with Ward criterion.

Figure E.5: *Even though the sky looks like a homogeneous color, it is actually a gradient that is split into several regions by the Ward criterion.*

E.5 Results and conclusions

We have described region-merging as an image-segmention algorithm where merging criteria play a key role for improving the segmentation result. Several low-level merging criteria have been evaluated for application in natural image segmentation. Based on the properties of the criteria, a multi-stage approach has been presented. The Ward criterion is used in the first stage to reduce the influence of image noise. The subsequent Mean-Ward stage performs the actual color segmentation, and finally, the Border criterion reduces oversegmentation due to illumination effects.

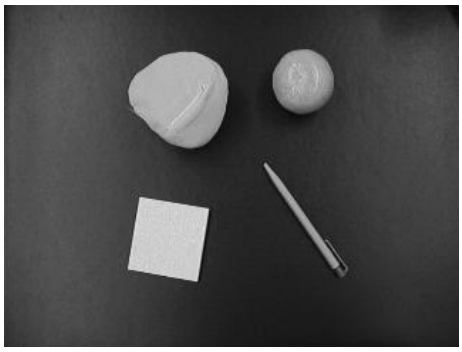
Figure E.7 shows a sample image with the results of the multi-stage segmentation algorithm. Neither the Mean criterion, nor the Ward criterion alone achieves acceptable segmentation results. Only by using a multi-stage approach of Ward, Mean-Ward, and Border (Figure E.7(d)), a subjectively superior object separation is obtained.



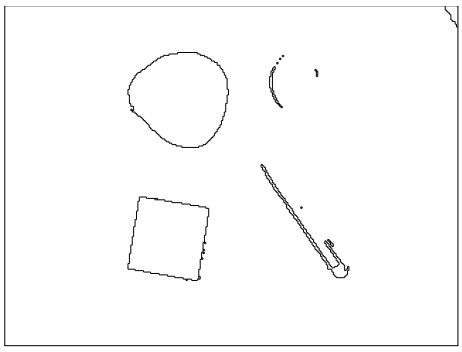
(a) Input image. (b) Mean criterion.

(c) Ward criterion. (d) Mean-Ward criterion.

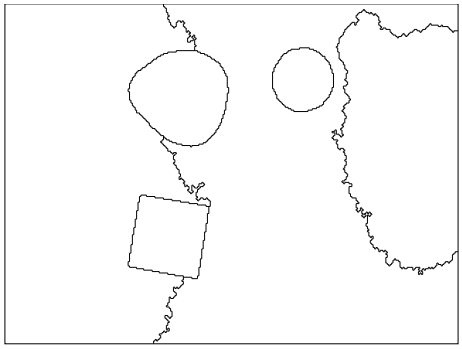
Figure E.6: *Example segmentation result with various merging criteria.*



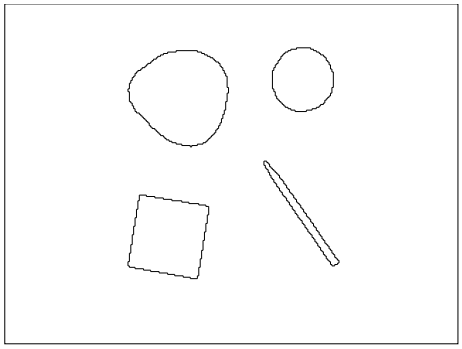
(a) Input image.



(b) Mean criterion.



(c) Ward criterion.



(d) Multi-stage.

Figure E.7: *Segmentation results for three different criteria using an image of several objects on a table.*