Semi-supervised segmentation of images

Michael Beller, Hartmut Gemmeke, Tim. O. Müller, Rainer Stotzka

Forschungszentrum Karlsruhe
Institute for Data Processing and Electronics
Karlsruhe, GERMANY

Abstract

Segmentation of patterns in images is usually based on the processing chain "segmentation – feature extraction – classification". It will be described, how for an arbitrary application an automatic processing stream can be generated with an initial feed back from an expert. The core of this new method is an automatic generator of optimized classifications and an example based interactive feature extractor. The method is demonstrated at various examples, where the features of the patterns to be segmented differ not very much from their environment.

1 Introduction

In many applications of image processing the task of recognizing patterns is a crucial step that is usually carried out by human experts, who rely on knowledge and gathered experience. For large numbers of patterns, this manual recognition becomes quite time-consuming and the results may differ substantially [6, 8, 11]. This makes an automatic recognition preferable. It is often not possible for the expert to explain exactly why a specific pattern is assigned to a certain class, making it difficult to specify rules for an automation. An algorithm that is suitable for all recognition purposes does not yet exist, so there are many algorithms each solving a specific problem.

For most automatic recognition tasks usually a four step image processing chain is constructed (see fig. 1): segmentation, feature extraction, feature selection and classification. The most important step is the segmentation as all subsequent tasks heavily rely on its quality [2, 14].

Our main assumption is that the better the segmentation, the more precisely the regions' features can be computed and the more accurately the regions can be categorized. Yet there is no algorithm that supports a broader band of applications and is able to adapt to environmental changes [2].

In this paper we propose a generic semi-supervised approach for segmentation combining segmentation concepts and machine learning mechanisms and integrating expert knowledge to automatically generate an optimized processing stream. We add a feedback to the common image processing chain in order to automatically find good parameters for the segmentation itself.

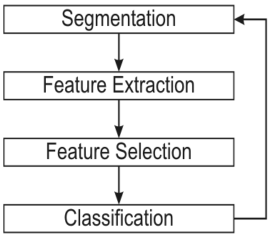

Figure 1: *Simple image processing chain. After the segmentation features from the regions are extracted. Important features are selected and the regions are categorized by a classifier. This chain is executed for each image.*

The remainder of this paper reviews in Section 2 the basic concepts of segmentation, feature extraction/selection, classification and other general approaches to object recognition. Section 3 describes our proposed system. Experimental results are presented in section 4 followed by conclusions and future work in section 5.

2 Background

This section covers the basic concepts used to construct our system.

2.1 Pattern recognition

Pattern recognition is defined as the computerized detection and analysis of patterns in signals. In image processing this usually means the analysis, description, identification and classification of objects or other meaningful regularities [26].

The process to recognize patterns that we will use is shown in fig. 2. It consists of several stages leading from preprocessing over segmentation, feature extraction and feature selection to classification. All data gathered throughout this process will be stored in a database.

We assume that we cannot influence the preprocessing stage. In the other case we would have to incorporate it into the optimization. So the earliest stage to affect the system is the segmentation. Its parameters are controlled by the extraction and classification stages to optimize the segmentation result and the overall recognition quality.

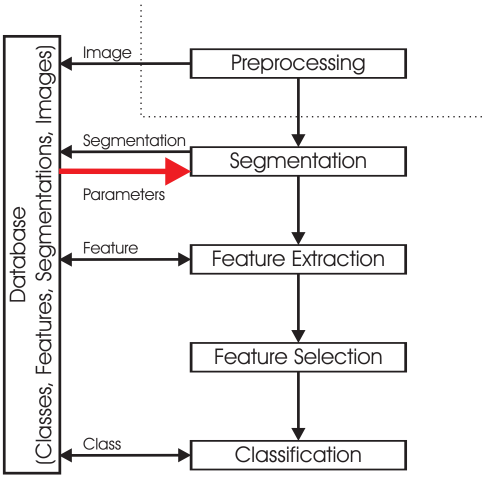

Figure 2: *The main pattern recognition process consists of segmenting the pattern of interest from the image, extracting its features, selecting the most descriptive features and determining the class membership based on those features.*

2.2 Segmentation

The aim of segmentation is to partition the image into contiguous regions, each containing one semantic object. A commonly used segmentation method is the seeded region growing (SRG) [4, 20, 22, 23, 25]. Starting from a seed pixel, regions are formed by aggregating adjacent pixels, whose gradients of greylevel are below a certain threshold. The resulting regions are homogeneous in respect to their gradient. On textured regions the SRG will often fail. The SRG can easily be enhanced to overcome the limitations of this single fixed homogeneity criterion. Therefore the constrained region growing (CRG) has been developed [22, 25] that allows the use of different homogeneity criteria for the segmentation.

Warfield et al. described a CRG in [22]. Additionally to observed features they used expert knowledge from an anatomical atlas.

Pohle et al. [20] estimated the homogeneity with the gradient and simple morphological properties and adjusted the segmentation parameters accordingly.

Paclík et al. [19] used spectral features for an unsupervised texture segmentation. However unsupervised clustering increases the risk of misclassifying pixels.

Tan et al. [10] extracted regions from natural images with little user interaction by outlining the region of interest and applying a global thresholded segmentation based on photometric features. A linear discriminant analysis refined the segmentation locally.

We [13] presented a more general approach where a segmentation system was built up from a database of manually segmented objects. In this approach the constraints were automatically selected from a library; yet it still requires a great amount of expert work for the manual segmentations.

Other commonly used segmentation methods work with boundaries rather than with regions. Ginneken et al. [5] and Cremers et al. [7] constrained their methods by statistical models of the outer shape, that were derived from training sets. Objects whose shape is not noticeable or can not be estimated, cannot be segmented in this manner.

Recent contributions proclaim that combining segmentation and classification together with learning techniques are suited to adapt to specific tasks. Lee et al. [11] used unsupervised learning to find task-specific templates for each class for segmenting with snakes. The templates were generated by averaging over each class of outlines. Bhanu et al. [3] combined a segmentation algorithm with a genetic algorithm that optimizes the segmentation parameters for the segmentation of Synthetic Aperture Radar (SAR) images. The set of segmentation parameters is fixed and may not be suitable for other applications. In another contribution Bhanu et al. [2] presented a general pattern recognition approach that adapts segmentation parameters to changing environmental conditions, but constrained the application to objects whose model is known.

2.3 Features

The deterministic calculation rule for the measurement of the segmented regions' properties is called a feature. The computation of a region's features is known as feature extraction. There exists a wide variety of features, e.g. statistical, morphological and textural features, of different categories, e.g. numerical or nominal. The measured features of an object are called observation.

In statistical pattern recognition, each pattern is represented by a set of d features [1]. The goal is to find those features that enable a grouping of patterns of different categories into disjoint regions in a d-dimensional space.

There exists a combination of features that is better suited to group the patterns than others [9, 21]. Choosing such a combination is called feature selection. Several selection methods are described in [17, 21].

2.4 Classification

For a set of training patterns from each class, the goal is to find boundaries in the feature space that separate clusters of patterns from different classes [1, 12]. Depending on the set of features, the number of observations and the classifier, the correct class will not always be predicted. This is called classification error. The performance of the classifier on unknown patterns can be estimated by the generalization error [17, 21].

3 Used Methods

3.1 Overview

On the image, the expert interactively selects a region exemplary for the object that is to be detected. Additionally he selects a counter-exemplary region that usually marks the background. From these two regions as many local features as possible are automatically extracted. Note that the extracted features do only apply to region elements (e.g. pixels or local neighbourhoods) and not to the whole region. From these features a subset is selected and a classifier is constructed that can differentiate between the two exemplary regions' elements. In a region growing process this classifier is applied to the whole image resulting in a segmentation that is based on the given examples. The expert reviews the segmentation result and interactively corrects misclassified regions. The features will be reselected and the classifier rebuilt to include these changes. This is repeated until the result is satisfying.

3.2 Adaptive Segmentation

We state that most of the constraints used in literature can simply be regarded as extractable features, e.g. the standard region growing criterion "greylevel gradient". Computing the gradient can be interpreted as the extraction of the feature "greylevel gradient" from a highly localized region, i.e. two adjacent pixels. The threshold based decision, to which region the pixels belong, corresponds to a simple classification task.

Any computable feature is a candidate for a constraint. The features most feasible for describing the inner homogeneity of a region should be used as constraints. The term inner homogeneity refers to the similarity of elements of a region [27], including similar brightness, texture, …, while the outer homogeneity refers to the similarity of two different regions.

We suggest that choosing the homogeneity measure – and thus the segmentation constraints – should not be done subjectively. As an example, humans are able to describe textures in fuzzy terms like roughness or granularity, but it's not clear, which combination of computable features truly applies. Yet we'd like to avoid fuzziness if possible.

The main novelty here lies in the use of *freely* selectable parameters for the segmentation such as greyvalue statistics *and* texture measures. These parameters are *automatically* selected from a feature pool in different combinations depending on the regions to be segmented and are optimized by a classification subsystem.

The segmentation in detail works as follows: Given an image $I(x, y)$, the expert selects two exemplary regions, one containing a part of the desired object and the other containing a part of the background (see fig. 3). The exemplary regions define the class-membership of the following observations, i.e. the object will be denoted by class C_1, while the background is referred to as class C_0. 45 features $\mathbf{M}$ from a library of algorithms are extracted from sliding windows of different sizes k around each pixel (x, y). The window sizes range from 3×3 to 64×64, thus forming several observation vectors $\mathbf{Y}_k$. On the $\mathbf{Y}_k$ feature selection is performed, selecting a subset of features $\mathbf{M}^*$ with a bidirectional forward selection, resulting in a reduced observations vector $\mathbf{Y}_k^*$. For each $\mathbf{Y}_k^*$ a classifier f_k is constructed. The f_k with the lowest generalization error E_k – estimated by 10-fold cross-validation – is selected. The features $\mathbf{M}^*$ and the classifier are fed back to the segmentation step as parameters. The selected features are now extracted from an appropriate window around every pixel of $I(x, y)$ and the classifier decides upon the class-membership of each pixel. Following a region growing scheme, this results in a segmentation based only upon the two exemplary regions.

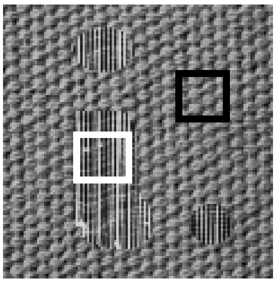

Figure 3: *Image after the expert interaction: two boxes have been dragged with the mouse and contain exemplary information about object and background. The white and black box marks exemplary content of the object and background respectively.*

If the segmentation is unsatisfactory, the expert makes corrections by giving additional example regions from the image. Their features will be calculated and the classifier will be adapted. This iterates until the expert is satisfied with the segmentation and the system is adapted to the current image.

Convergence is given here due to the limited image size:

The expert cannot give more examples than there are pixels present on the image. In the worst case all pixels would have to be labeled manually, which corresponds to manual segmentation.

4 Results

We have tested the method on many images of very different content. A few of these examples are selected in fig. 4,5: several synthetic texture images [16, 18], natural scenes [16] and digital mammogram containing and indicator for breast cancer [15]. At least two exemplary rectangles per image were defined manually. An equal number of observations per class – 100 if possible – were computed from each rectangle for each moving window size so that the computed generalization errors are comparable. To select the features we used a forward-selection [1, 21]. For the classification we used a pruned tree (J48) [17]. The seed point for the region growing scheme is automatically selected inside the rectangles. Only a 4-neighbourhood is considered for the growing process as we want to avoid single-pixel diagonal region connections. Visually evaluated, the resulting segmentations are very good and the effort for the expert is minimal.

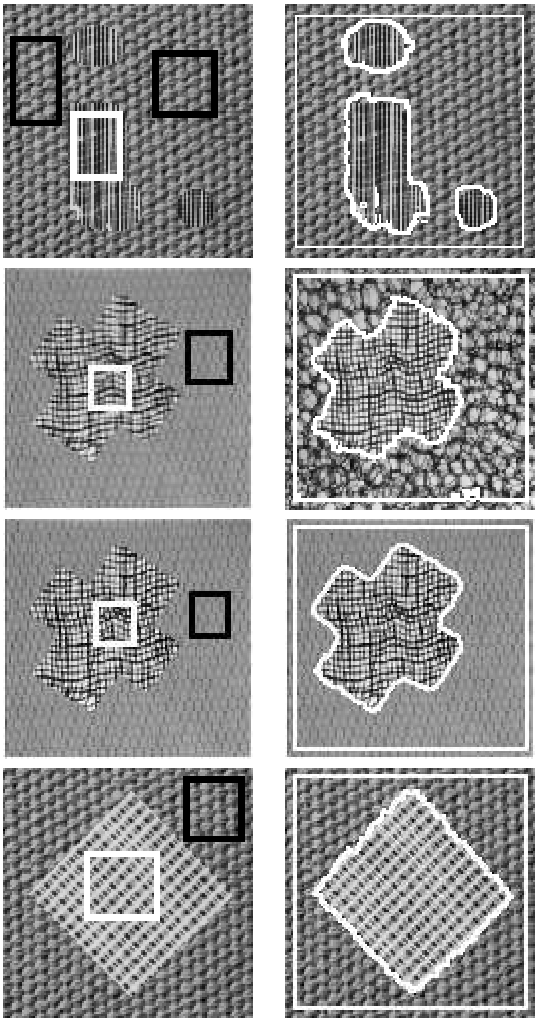

Figure 4: *Example images. Left: the original synthetic textures with exemplary (white) and counter-exemplary (black) interactive selections. Right: the boundaries of the resulting segmentations.*

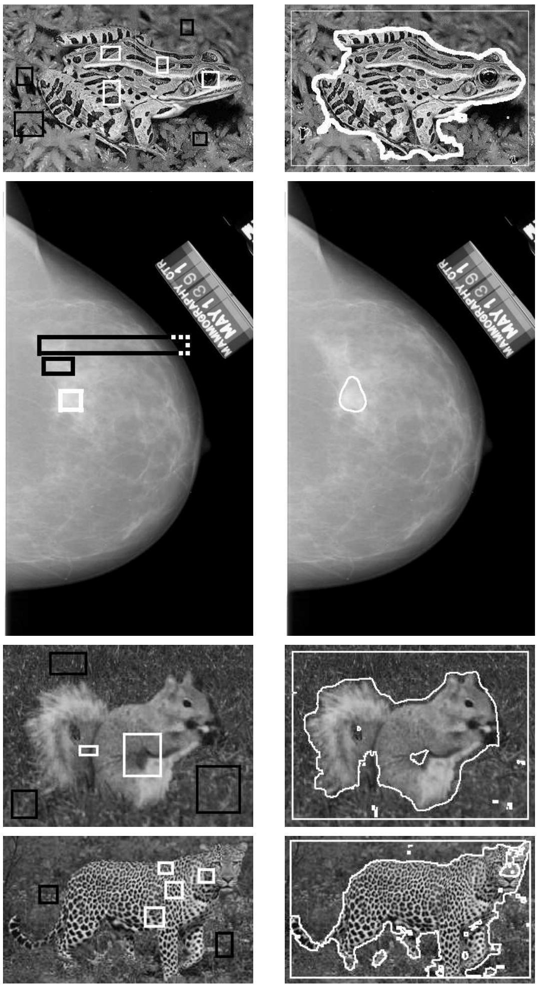

Figure 5: *Example images (top down). Left: the original natural scene of a frog on moss, a digital mammogram with a spiculated mass that indicates breast cancer, a squirrel on meadow and a leopard. Exemplary (white) and counter-exemplary (black) interactive selections. Right: the boundaries of the resulting segmentations.*

Fig. 4 and 5 show the original images and the boundaries of the found segments. The outer white rectangle results from the moving window technique: a fringe of half the window size remains unprocessed as the window would move out of

the image.
The main objects in the images are very well separated from each other. It can be seen that the boundaries of the objects are very clear and fine details have been preserved well.

5 Conclusions and future work

We have described the automatic generation of a processing stream for segmentation in arbitrary pattern recognition problems. As our main assumption was that good recognition results depend on a good segmentation, we have focussed this contribution on an adaptive segmentation. This adaptivity was achieved by adding an automatically generated classification subsystem to the segmentation, adding a feedback to an image processing chain and using a freely selectable parameter set of arbitrary size. The parameter set consists of a classifier and local features. A maximum number of 45 different features is available, which is often sufficient for good results, but can not guarantee good results in all cases. If the system performs badly, eventually better features are needed [21]. The feature selection and classification algorithms were fixed for the presented applications, yet there is a broad variety of 20 selectors and 30 classifiers available. A framework to find a good combination has been provided by Müller et al. [24].
Currently the time to calculate the parameters, i.e. extracting the features and rebuilding the classifier might take up to 20 minutes depending on the number of observations and the complexity of the classifier. The segmentation of a single image might take up to 10 minutes[1].
The applicability was demonstrated on several artificial textures as well as on natural images. The segmentation quality was very good by visual inspection, but the ”ground truth segmentation” may be given in synthetic images. We use the generalization error as an evaluation criterion for the mapping of the pixels to a region. A quantitative measure of the error cannot be evaluated for natural images therefore we plan to set up synthetic images with known segmentation boundaries to quantify the segmentation quality of our method. So far we use the estimated generalization error, which is a good criterion to evaluate the performance of the classification-subsystem. The performance of the segmentation however cannot be clearly assessed by this criterion. We further tend to process publicly available databases of images and use better suited evaluation criteria to compare our results.

References

[1] A.K. Jain, R.P.W. Duin, J. Mao. Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 2000.

[2] B. Bhanu, J. Peng. Adaptive Integrated Image Segmentation and Object Recognition. *IEEE Transactions on Systems, Man, and Cybernetics*, 30(4):427–441, 2000.

[3] B. Bhanu, S. Lee and J. Ming. Adaptive Image Segmentation Using a Genetic Algorithm. *IEEE Transactions on Systems, Man and Cybernetics*, 25(12):1543–1567, 1995.

[4] Bernd Jähne. *Digitale Bildverarbeitung*. Springer, 1997.

[5] B.v. Ginneken, A.F. Frangi, J.J. Staal, B.M.ter Haar Romeny, M.A. Viergever. Active Shape Model Segmentation with optimal Features. *IEEE Transactions on Medical Imaging*, 21(8):924–933, 2002.

[6] D. Comaniciu, P. Meer, D. Foran. Image Guided Decision Support System for Pathology. *Machine Vision and Applications*, 11(4):213–224, 2000.

[7] D. Cremers, T. Kohlberger and C. Schnörr. Nonlinear Shape Statistics in Mumford-Shah Based Segmentation. *7th European Conference on Computer Vision*, 2002.

[8] Oliver Dehning. Wissensbasierte Inspektion industrieller Objekte. Technical report, Institut für Nachrichtentechnik und Informationsverarbeitung, Universität Hannover, 1995.

[9] I. Guyon, A. Elisseeff. An Introduction to Variable and Feature Selection. *Machine Learning Research 3*, 3(3):1157–1182, 2003.

[10] K.-H. Tan, N. Ahuja. Selecting Objects with Freehand Sketches. In *Proc. IEEE International Conference on Computer Vision*, 2001.

[11] K.-M. Lee, W.N. Street. Model-based detection, segmentation and classification for image analysis using on-line shape learning. *Machine Vision and Applications*, 13:222–233, 2003.

[12] Keinosuke Fukunaga. *Introduction to Statistical Pattern Recognition 2nd Edition*. Academic Press, 1990.

[13] M. Beller, R. Stotzka, H. Gemmeke. Merkmalsgesteuerte Segmentierung in der medizinischen Mustererkennung. In *Bildverarbeitung für die Medizin 2004*, pages 184–188, 2004.

[14] M. Beller, R. Stotzka, T. O. Müller. Application of an interactive feature-driven segmentation. In *Biomedizinische Technik*, pages 210–211, 2004.

[15] M. Heath, K.W. Bowyer, D. Kopans et al. Current status of the digital database for screening mammography. In *Digital Mammography*. Kluwer Academic Publishers, 1998.

[16] Mikaël Rousson, Thomas Brox, Rachid Deriche. Active Unsupervised Texture Segmentation on a Diffusion Based Feature Space. Technical report, Institut National de Recherche en Informatique et Automatique, 2003.

[17] University of Waikato. WEKA: Waikato Environment for Knowledge Analysis. http://www.cs.waikato.ac.nz/˜ml/index.html.

[18] P. Haberäcker. *Praxis der Digitalen Bildverarbeitung und Mustererkennung*. Hanser, 1995.

---

[1] exemplary for an image with 400×400 pixels and 40 features to be calculated on a standard PC with 3 GHz

[19] Pavel Paclík and Robert P.W. Duin. Multi-spectral image segmentation algorithm combining spatial and spectral information. *Proceedings of SCIA 2001 conference*, pages 230–235, 2001. http://citeseer.nj.nec.com/513814.html.

[20] Regina Pohle and Klaus D. Toennies. Segmentation of medical images using adaptive region growing. In *Proceedings of SPIE (Medical Imaging 2001)*, volume 4322, pages 1337–1346, 2001.

[21] Sholom M. Weiss, Casimir A. Kulikowski. *Computer Systems that learn*. Morgan Kaufmann, 1991.

[22] S.K. Warfield, J. Dengler, J. Zaers et al. Automatic identification of gray matter structures from mri to improve the segmentation of white matter lesions. In *Journal of Image Guided Surgery*, pages 326–338, 1995.

[23] Thomas Lehmann, Walter Oberschelp, Erich Pelikan und Rudolf Repges. *Bildverarbeitung für die Medizin*. Springer, 1997.

[24] T.O. Müller, R. Stotzka, M. Beller, N.V. Ruiter, V. Hartmann. Schneller Aufbau medizinischer Diagnosesysteme mit der Komponentensoftware ICE. In *Bildverarbeitung für die Medizin 2004*, pages 443–447, 3 2004.

[25] S.Y. Wan and W.E. Higgins. Symmetric region growing. *IEEE Transactions on Image Processing*, 12(9):1007–1015, 2003. http://cobb.ee.psu.edu/publications/Publications/ieee-ip-wan.pdf.

[26] Wikipedia – The free Encyclodedia. Pattern recognition. http://www.wikipedia.org.

[27] Xiaofen Ren, Jitendra Malik. Learning a Classification Model for Segmentation, 2003.