IN: *PROCEEDINGS OF THE 10TH INTERNATIONAL CONFERENCE ON SIGNAL AND IMAGE PROCESSING*, KAILUA-KONA, USA, AUGUST 18-20, 2008.

CONTENT-BASED SELECTION OF METHODS FOR IMAGE SEGMENTATION

Gabriele Peters
University of Applies Sciences and Arts
Computer Science, Visual Computing
Emil-Figge-Str. 42
Dortmund, Germany
gabriele.peters@fh-dortmund.de

Matthias Petke
Technical University Dortmund
Computer Science, Graohical Systems
Otto-Hahn-Str. 16
Dortmund, Germany

ABSTRACT
Many different methods for image segmentation have been developed. Each of them usually has its advantages for one single class of images only. An unsolved problem in the field of image segmentation, however, consists in the fact that there does not exist a single approach which can be applied to several classes of images with the same success. As an idea for a solution of this problem we propose an autonomous selection of a segmentation method by a neural network depending on the specific content of the image to be segmented. The neural network is trained with statistical image measures of classical modern paintings and is able afterwards to select the most appropriate of four standard methods for image segmentation for unknown test images. We evaluate our approach in the field of edutainment by an automatization of Ursus Wehrli's idea of "Cleaning up artworks".

KEY WORDS
Image Segmentation, Neural Networks, Edutainment, Digital Art

1 Introduction

Many different methods for image segmentation have been developed. Each of them usually has its advantages for one single class of images only. An unsolved problem in the field of image segmentation, however, consists in the fact that there does not exist a single approach which can be applied to several classes of images with the same success. As an idea for a solution of this problem we propose an autonomous selection of a segmentation method by a neural network depending on the specific content of the image to be segmented. In section 2 we introduce our approach, in section 3 we describe the data we used to train and test the neural network, and in section 4 we present the results. Finally, we introduce an edutainment interface in section 5 which utilizes the proposed selection of segmentation methods depending on the image content.

2 Our Approach

We considered four of the most widely-used methods for image segmentation. They are summarized in subsection 2.1. Our intent was the training of a neural network to be capable of selecting those of these methods which is the best choice in terms of a proper segmentation of the image at hand. To generate ground truth data for the training of the neural network we determined the best method of segmentation for each training image by visual inspection of segmentation results. The selection of the most appropriate method by the neural network is carried out on the basis of the image content. To quantify the image content we consult the values of seven statistical image measures for each image. They are described in subsection 2.2. The neural network is described in subsection 2.3.

2.1 Considered Methods for Image Segmentation

The four methods of image segmentation we consider for our approach are as follows:

THRES: Simple Global Thresholding

We convert the image to gray-values and apply a rotationally symmetric 3x3 Gaussian lowpass filter with standard deviation 0.5. Then we partition the histogram of the image intensities by using a single global threshold. To determine this threshold we utilize Otsu's method [1].

RGROW: Region Growing Using Color

First we convert an RGB image to HSV color space and quantize this space using the method proposed in [2]. This results in 18 bins for the H component and 3 bins for the S and V components each. Then we apply the region growing algorithm described in [3] using random seed pixel. The similarity criterium for the growth process is the deviation of no more than one level of quantisation for each of the H, S, and V components.

COLOR: Advanced K-Means-Clustering Using Color

We utilize the color segmentation proposed in [4]. Again, the RGB image is converted to HSV first. Based on its saturation value a pixel is either characterized by its hue or by its intensity as its dominant property. A subsequent k-means clustering utilizes then either the hue or the intensity to determine the similarity of pixels.

SNAKE: Snakes Using Gradient Vector Flow

We utilize the snake approach with the gradient vector flow algorithm proposed in [5]. The start contour is generated by the Canny edge detector followed by a distance transform to fill the gaps in the contour. After each iteration of the snake algorithm (maximum 20 iterations) a snake is rejected if it either becomes too large or too small. Finally, the the pixels are assigned to the segments according to the resulting snake polygons.

The size of the longer image edge is 900 pixels for all segmentation methods. Further details on the implementations of these segmentation methods can be found in [6].

2.2 Used Statistical Image Measures

The decision of the neural network about the most appropriate method of segmentation is based on the values of seven statistical image measures of first and second order, which characterize the content of the image at hand. For the choice of these measures we leaned on the approved standard image parameters summarized, e.g., in [7]. They are as follows:

1. Entropy
2. Standard deviation of the H, S, and V component of the image (i.e., 3 values)
3. Standard deviation of the gradient
4. Contrast (horizontal and vertical, thus 2 values)
5. Correlation (horizontal and vertical, thus 2 values)
6. Energy (horizontal and vertical, thus 2 values)
7. Homogeneity (horizontal and vertical, thus 2 values)

This results in a vector with 13 values to characterize the content of an image.

2.3 Neural Network

In [6] we have compared several different feedforward neural networks. Here we will report on the one which performed best. It has 120 neurons in a single layer. We employ classical backpropagation [8]. The input per training image consists in the 13-vector of the image measures, and the output consists in the four-vector which encodes the appropriateness of the four methods of segmentation for this training image. One training step consisted of 100 repetitions for each of 21 training images. After 14 steps the network performed without error on the training images. The ground truth data to train the network is described in section 3.1

3 Training- and Testdata

For the evaluation of the developed technique images are appropriate which can be segmented easily based on only a few outstanding features such as color. Such images can be found frequently in classical modern painting. Thus, we chose classical modern paintings for training as well as for testing the neural network.

3.1 Trainingdata

As training images we chose 21 classical modern paintings, 20 of them are depicted in table 1. To generate ground truth data for the training of the neural network we applied all four methods of segmentation to each of the 21 training images and assessed the best method via visual inspection. We also determined the best and second best method for the 10 test images by visual inspection to evaluate the results described in section 4. Figure 1 shows examples for three paintings.

3.2 Testdata

To test the trained neural network we chose those 10 classical modern paintings which are depicted in the leftmost column of tabel 2.

4 Results

The results are summarized in tabel 2. In the second and third column it lists the ground truth data determined for the test images, i.e., the first and second best methods of segmentation for each test image, respectively. For example, SNAKE was identified as best method, RGROW as second best for Malevich's "The reaper". The last two columns show the results: the method chosen by the neural network and a rating of this choice. The rating is "Perfect" if the best method was chosen, "Good" if the second best method was chosen, and "Poor" if any other method was chosen. Summarizing, in 60% we obtain a "Perfect" choice by the neural network. If we are satisfied with the second best method we obtain in 90% a "Good" result, and in 10% a "Poor" choice. To assess these results one may consider the probabilities of choosing 6 times "Perfect", 9 times "Good", and 1 time "Poor", which would correspond to a random choice of a segmentation method. These probabilities are 1.6%, 0.9%, and 0.9%, respectively.

5 Application

The Swiss comedian and artist Ursus Wehrli has attended to the problem of image segmentation in his performances and books on "cleaning up artworks" [9, 10]. He used scissors and glue to segment famous artworks and reorder the segments according to, e.g., color and form (figure 2). This

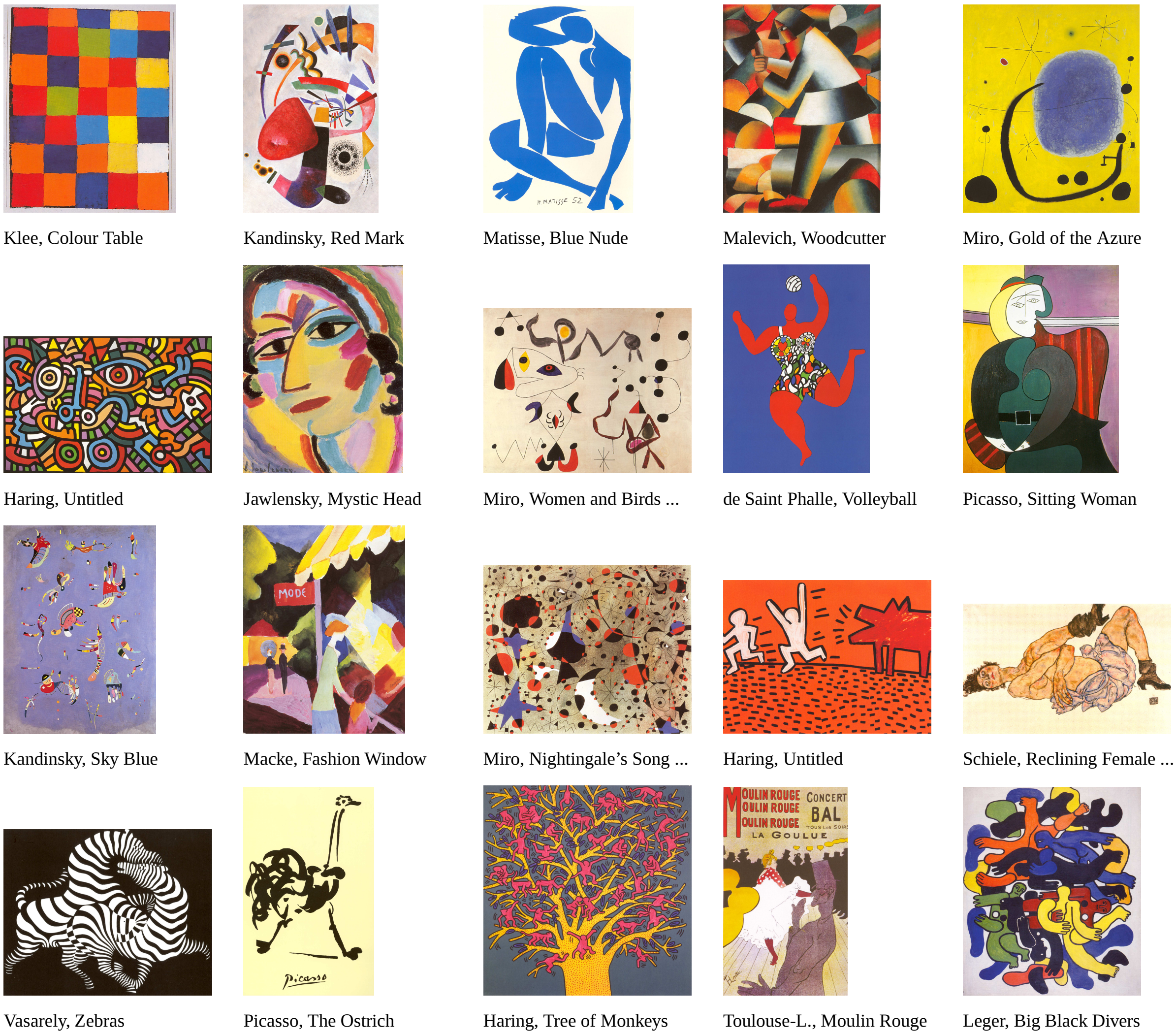


Table 1. Training images

idea was motivation and framework for us to develop an interface which allows for the automatic clean up of artworks in the Wehrli fashion. This interface applies the content-based selection of methods for image segmentation proposed above, and thus is a first application and test case of this approach. Table 3 faces the versions cleaned up by Wehrli and those cleaned up by our approach for two representative sample artworks. For the large majority of tested sample images our results look rather similar to those of Wehrli, which means that our approach is capable of selecting the proper segmentation method depending on the image content. Many more examples and the details of reordering the identified segments are shown and described in [6].

6 Conclusion

An unsolved problem of image segmentation consists in the fact that there does not exist a single segmentation approach which can be applied to several classes of images with the same success. For a solution of this problem we proposed the idea to autonomously choose an appropriate segmentation method depending on the content of the image. This possibility was demonstrated by means of a feedforward neural network which was trained on easy-to-segment images from classical modern painting and which is able to select an appropriate method of segmentation for 90% of the tested images. The interface introduced in section 5 is a first application of this approach and can be employed in multimedia systems or edutainment applications, e.g., for arts education or for interactive museum exhibitions.

References

[1] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man*

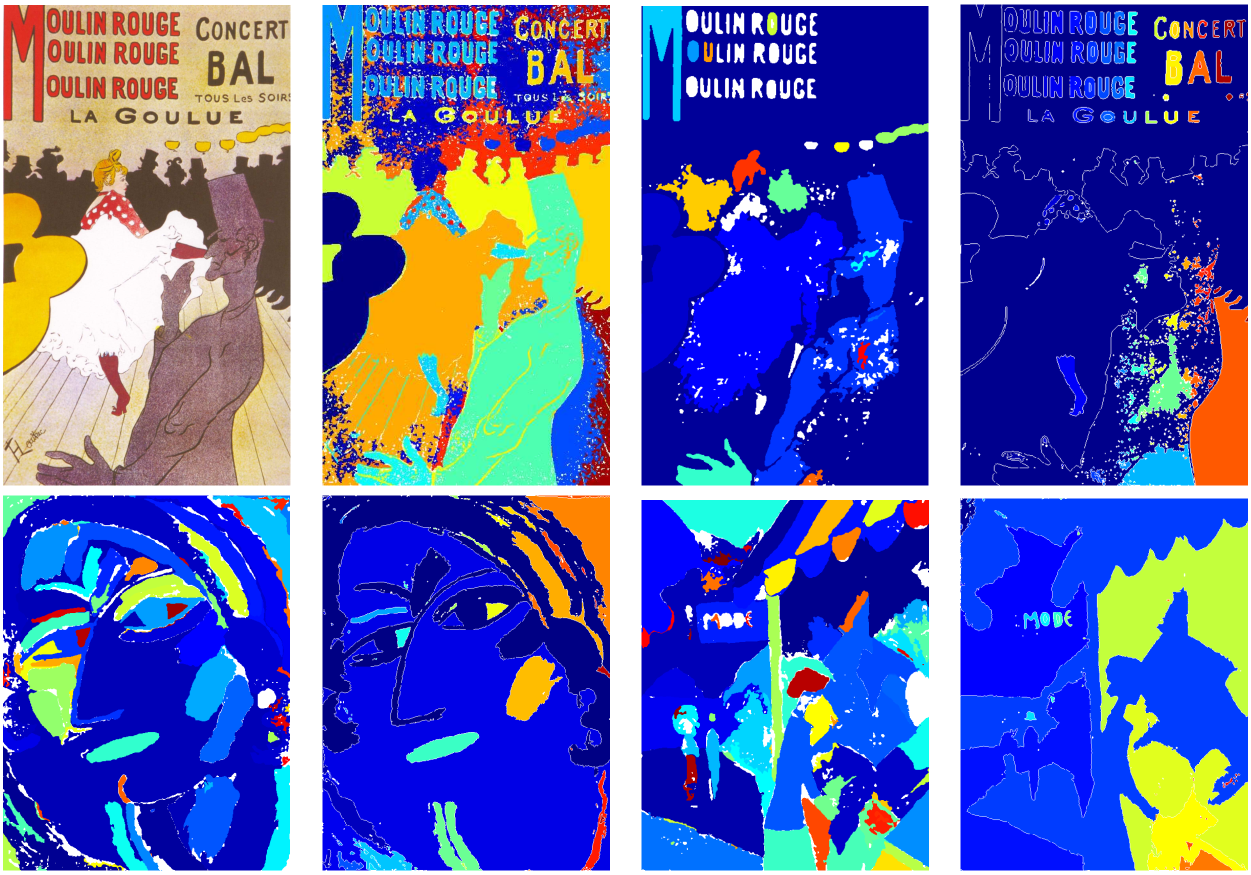


Figure 1. Ground truth. First row, from left to right: original painting, segmented with method COLOR, segmented with method RGROW, segmented with method THRES. Each segment is labeled by an own color. By visual inspection we assessed the segmentations of all four methods. For this image we rated method COLOR as best, RGROW as second best, and THRES as worst method. Second row, first two images: best method RGROW and worst method THRES for Jawlensky's painting (original in table 1). Second row, last two images: best method RGROW and worst method THRES for Macke's painting (original in table 1). To comprehend our ratings, please compare the segmentations with the original paintings.

and Cybernetics, 9(1):62–66, 1979.

[2] J. R. Smith and Shih-Fu Chang. Single color extraction and image query. In *ICIP '95: Proceedings of the 1995 International Conference on Image Processing (Vol. 3)-Volume 3*, page 3528, Washington, DC, USA, 1995. IEEE Computer Society.

[3] R.C. Gonzalez, R.E. Woods, and S.L. Eddins. *Digital image processing using MATLAB*. Prentice Hall, 2004.

[4] S. Sural, G. Qian, and S. Pramanik. Segmentation and histogram generation using the hsv color space for image retrieval. In *Proceedings of the International Conference on Image Processing*, pages 589–592, 2002.

[5] C. Xu and J. L. Prince. Gradient vector flow: A new external force for snakes. In *IEEE Proc. Conf. on Comp. Vis. Patt. Recog. (CVPR'97)*, pages 66–71, 1997.

[6] M. Petke. *Kunst aufräumen á la Ursus Wehrli - Bildinhaltsgesteuerte Auswahl von Segmentierungsverfahren*. Diploma Thesis, Technical University Dortmund, 2007.

[7] R. M. Haralick, Dinstein, and K. Shanmugam. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3:610–621, 1973.

[8] J. P. Werbos. *The roots of backpropagation: From ordered derivatives to neural networks and political forecasting*. John Wiley, New York, 1996.

[9] U. Wehrli. *Kunst aufräumen*. Kein und Aber, Zürich, 2002.

[10] U. Wehrli. *Noch mehr Kunst aufräumen*. Kein und Aber, Zürich, 2004.

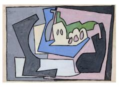
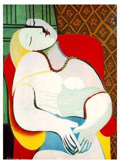
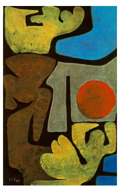
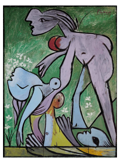
Test Images	Ground Truth		Chosen by NN	
	Best Method	Scnd. Best M.	Method	Rating
 K. Malevich The Reaper	SNAKE	RGROW	SNAKE	Perfect
 Juan Gris Painter's Window	SNAKE	COLOR	SNAKE	Perfect
 Jackson Pollock Blue (Moby Dick)	RGROW	THRES	RGROW	Perfect
 Henri Matisse Music	SNAKE	THRES	SNAKE	Perfect
 Pablo Picasso Still Life, Basket and Fruit	THRES	COLOR	THRES	Perfect
 Pablo Picasso The Dream	SNAKE	COLOR	SNAKE	Perfect
 Roy Lichtenstein Girl with Tear	THRES	COLOR	COLOR	Good
 Juan Gris Landscape with Houses at Ceret	RGROW	SNAKE	SNAKE	Good
 Paul Klee Park of Idols	SNAKE	COLOR	COLOR	Good
 Pablo Picasso The Rescue	SNAKE	THRES	COLOR	Poor

Table 2. Test images and results

Figure 2. Swiss comedian and artist Ursus Wehrli in a performance with a cleaned up version of Miro’s ”Gold of the Azure”

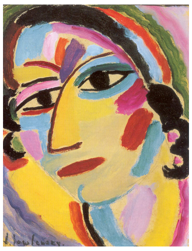
Original Artworks	**Cleaned up by Ursus Wehrli**	**Cleaned up by Our Approach**
		
		

Table 3. Two artworks cleaned up by Ursus Wehrli and by our approach. For both images the best method was RGROW, and the neural network was able to select it for segmentation. After segmentation the segments are arranged as shown in the last column. The technical details of this arrangement are described in [6]. This use case is well suited for edutainment contexts, for example, for children’s arts education.