

Figure-Ground Segmentation Techniques

Snehal P. Ambulkar¹, Nikhil S. Sakhare²

¹2nd Year Student, Master of Technology, Computer Science & Engineering, Rajiv Gandhi College of Engineering & Research, Nagpur, Maharashtra, India

²Assistant Professor, Department of Computer Science & Engineering, Rajiv Gandhi College of Engineering & Research, Nagpur, Maharashtra, India

Abstract: Image segmentation is the process of dividing an image into multiple segments, such that the representation of an image changes, which makes analysis of image easier. Figure-ground segmentation deals with segmenting out foreground objects from background. This paper describes the different figure-ground segmentation techniques. This paper deals with survey of various figure-ground segmentation techniques and efficient approaches for object recognition.

Keywords: segments, figure-ground segmentation, foreground objects, background, object recognition

1. Introduction

The figure-ground segmentation plays vital role in the field of image processing. The figure-ground segmentation is used to separate foreground objects from background in an image, which partitions an image into several parts or regions depending upon some homogeneity criteria. The figure below shows that how segmentation helps in edge and boundary detection.

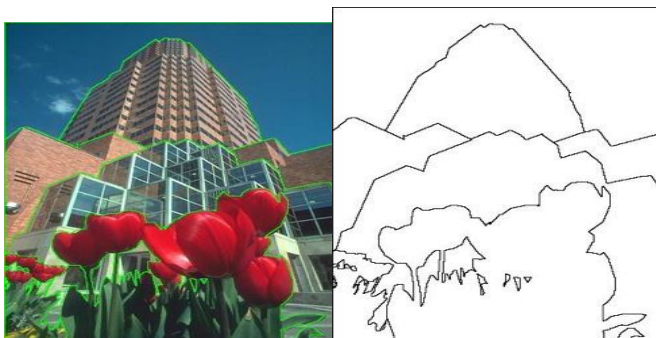


Figure1: Segmentation of an image

This approach is used in many fields like object recognition, image editing, image retrieval, target tracking, medical imaging, etc. The purpose of writing this paper is to provide a literature review in this arena. The structure of paper is as below; section 2 introduces the term figure-ground segmentation. Section 3 describes the current figure-ground segmentation techniques and section 4 concludes the overall study.

2. Figure-Ground Segmentation

Figure-ground segmentation is used to identify objects from its background. It is mainly used for object recognition. Basic idea behind this approach is to collect pixels which satisfy some homogeneity criteria. Pixels showing same texture, intensity, color are grouped together. Thus an object in an image is divided into multiple regions. Each region has its own characteristics. In simple words, separating a figure from its ground is figure-ground segmentation. Figure-ground segmentation has applications in many fields like object recognition, target tracking, image retrieval, etc.

Figure-ground segmentation partitions an image into several regions such that pixels belonging to same region have similar attributes. In order to be used for analysis and interpretation, these regions should closely relate to the depicted object or interested features. Successful image analysis depends on reliability of segmentation, but accurate image partitioning is very challenging task.

3. Current Figure-Ground Segmentation Techniques

Currently, different techniques are available for figure-ground segmentation. Although all of these techniques are based on different criteria and try to achieve accuracy in segmentation, these techniques cannot be applied to all types of images. An approach suitable for one class of images may not be suitable for another class of images. Selection of a technique for a specific class of images is also one of the difficult problems. In spite of research going on from decades, no universally accepted technique is available for figure-ground segmentation. Every approach has some drawbacks.

Existing figure-ground segmentation algorithms are classified into two categories, depending on number of object classes within an image as follows-

Class-specific

In class-specific segmentation [1], input images contain only a single class of objects. Due to lack of adaptively, range of applications of class-specific segmentation is limited.

Class-independent

In class-independent segmentation, segmentation is carried out for any class of objects. There are two approaches to address class-independent segmentation. First approach depends on multiple or hierarchical segmentation. Second approach is exemplar-based segmentation transfer. There are various techniques for figure-ground segmentation.

3.1 Figure-Ground Segmentation by Window Mask Transfer

This approach works on three datasets PASCAL VOC 2010, Weizmann horses, Graz-02. This approach is fully automatic and does not require user input. Basic idea behind this approach is to transfer segmentation masks from supervised training set to test image. Unary potentials of energy functions of test images are derived from transferred masks. Candidate windows having possibility to contain foreground objects are extracted. Masks are transferred from training windows which are similar to test image windows. Windows which are visually similar contain less variability than entire image which leads to better segmentation transfer.

Segmentation masks of similar windows are same. Segmentation masks are transferred using nearest neighbor. Image level neighbours do not have enough good quality. Using image level neighbours, different figure ground segmentations are obtained for most visually similar training images. Considering limitations of this approach, window level neighbours are used. Using this approach, windows containing objects are detected. Windows for all training images and test images are taken into consideration. These windows show very less variations than entire image. Better neighbours are obtained in this way. They are more suitable for segmentation mask transfer. The segmentation masks of the retrieved training windows are transferred to derive the unary potentials for the test image. Location model and appearance models are used for mask transfer. Following figure [2] shows overall idea of this approach.

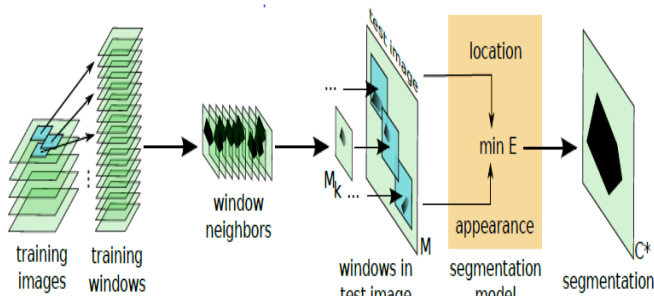


Figure 2: Overview of window mask transfer segmentation approach

Location model gives information about segmentation of a pixel based on its location in image. Location model of a pixel gives the foreground probability of a pixel. Location model is constructed for each pixel. Appearance model is composed of two Gaussian mixture models (GMM), one for the foreground, one for the background. Each GMM is composed of 5 components. Each pixel is represented as a vector of 3 components. Location model works on individual pixels, while appearance model works on large image regions. Compared to location model, appearance model improves segmentation.

For implementation, GIST is computed inside each window. Two unary potentials are obtained using this approach. As initialization area to estimate the appearance models, a box is used in the middle of the image, occupying 50% of its area. Not aligned with the objects in the image. Segmentation

masks of the 5 most similar training images are transferred to the test image, according to the GIST descriptor compute globally on the whole image. Final segmentation is achieved by averaging and thresholding the 5 masks. This results in effective segmentation transfer when supported by windows than by the whole image.[2]

3.2 Figure-ground segmentation using a hierarchical conditional random field

This approach is based on the concept of detection and segmentation of generic object classes that combines three components: generic image segmenter, generic classifier, and generic belief propagation (BP) procedure. Following figure [3] shows basic idea of this approach.

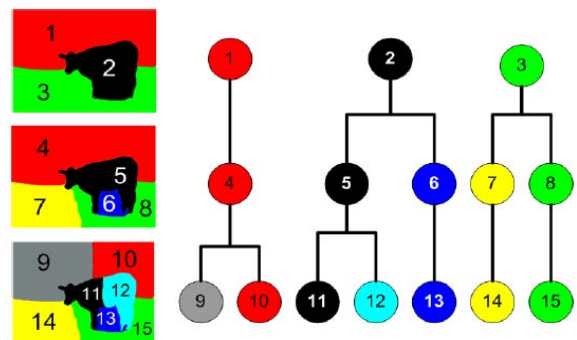


Figure 3: Overview of figure-ground segmentation using hierarchical conditional random field.

This system combines the regions together into a hierarchical, tree-structured conditional random field, applies the classifier to each node (region), and combines all the information together using belief propagation. Classifiers used in this approach depend on color and texture, they can handle deformable objects under severe occlusion and rotation. The drawback of sliding window classifiers is that they work on rigid objects and mostly fill the bounding box. It requires a lot of training data, and cannot easily handle severe rotation. A simple color and texture based classifier could only classify individual points, or regions chosen by a user. Window based approach works only at a single scale.

Hierarchical conditional random field based approach works at multiple scales. It performs histogram equalization to work on images with low contrast. A set of segmented images is obtained using segmenter at multiple scales. For each region, feature vectors are computed. A binary classifier is applied to each region. A tree in which nodes represent regions is constructed. Weights are attached to nodes depending on output of classifier. Belief propagation is used. Thresholds of beliefs are obtained for object segmentation.

This approach works on Pascal VOC dataset. The classifier works efficiently. However the classifier is uncertain about some regions. In this case, tree structure is computed. This technique is not designed for rigid objects for which standard sliding window classifiers works efficiently. It works efficiently for deformed objects. This framework improves the performance of any set of features. A segmenter is run at multiple scales which returns a set of segmented regions. Feature vectors for each region are computed. For this

purpose, color and texture histograms are used. A binary classifier is applied. A tree is built. Tree structure depends on location of regions and not on the image data. Beliefs are threshold to select foreground pixels. Object segmentation is achieved using binary segmentation masks. The technique used in this approach is very simple and mainly used for segmentation of deformable objects like animals. Performance can be improved in many ways such as improving base features, classifier, etc. [3]

3.3 Figure-ground segmentation using global transfer

This approach automatically segment out foreground objects from background. It deals with transferring segmentation masks of globally and locally similar examples into query image. This approach works on three datasets namely Pascal VOC 2010, VOC 2011 segmentation challenges and iCoseg dataset. It can segment out large scale images containing unknown objects. This approach mainly works on three modules: glocal scene retrieval, online prediction and segmentation with Support Vector Machine(SVM). Instead of SVM, random forests [4] can also be used. In glocal scene retrieval, proposed object oriented descriptor concept is used to retrieve glocally nearest neighbours for query image. Online prediction predicts foreground probability of query image.

Segmentation with SVM produces optimal segmentation. It combines SVM map and Markov Random Field (MRF). In global scene retrieval, agglomerative hierarchical clustering is used for object classification. In order to decide which clusters to be merged, a distance function is defined. Hierarchical clustering[5] is used instead of k-means, because k-means does not support required distance function, which is required for computing histogram of bag-of-features (BOF) vector. SVM score vector is normalized to obtain object oriented descriptor. Online prediction first segments query image and k-nearest neighbours into regions. Probability of each region is calculated by using classifier. Contour-based hierarchical segmentation algorithm gPb (globalized probability of boundary) is used for region generation. During segmentation with SVM, standard MRF model is used for segmentation. Basic concept of this approach is shown in following figure[6].

Object oriented descriptor encodes geometric information to highlight to objects in an image. Compared to GIST And PHOG descriptor, object oriented descriptor (OOD) gives more accuracy. Optimal segmentation is obtained by combining online prediction with Markov Random Field (MRF). Computational efficiency of this approach can be improved. Proposed approach can be implemented using thread parallelism.

Most matching images can be retrieved using image tag and context. This approach segments large scale images. OOD is evaluated on global transfer system, where GIST is used for location modeling and PHOG is used for appearance modeling. This approach cannot perform well for images containing multiple objects having scale and appearance variations. Objects with their reflection water cannot be separated using this approach. Small and blurred objects are not properly segmented out using this approach.

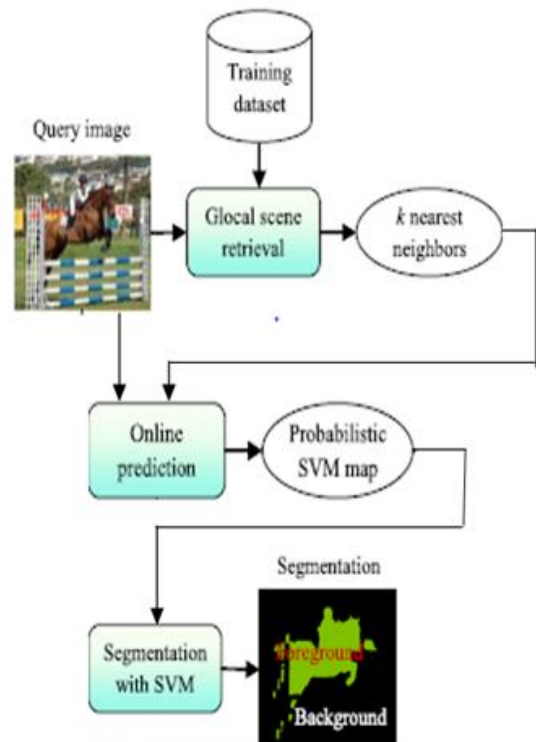


Figure 4: Generic framework of glocal transfer approach

In other cases, this approach performs very well. Accuracy of segmentation obtained using this approach is also very high. Figure-ground segmentation is performed automatically and efficiently using glocal transfer. Currently, this approach is widely used for segmentation. Work is going on to overcome shortcomings in this approach. [6] Following Table [6] shows accuracies:

Table 1: Segmentation accuracies (in percent) on iCoseg dataset

Approach	F-score
Global Transfer	69.5
Window transfer	74.8
Glocal transfer	80.0

3.4 Figure-ground segmentation using Difference of Gaussians

This approach is based on a new algorithm for automatic foreground extraction. Candidate keypoints of an input image are obtained using DoG. Pseudo keypoints are removed and important keypoints are rebuilt by using keypoint filter algorithm. An image is segmented into several regions using normalized-cut. [7] Foreground as well as number of keypoints in foreground are located using N-cut.

Difference of Gaussian (DoG) algorithm is used to select candidate key points. A pixel is compared to its 26 neighbours in 3 by 3 region at adjacent and current scales, in order to detect maxima or minima of difference of Gaussians of images. This pixel is known as candidate point. A filtering function is used to remove candidate points from background and reconstruct new candidate points in foreground. These candidate points are called as key points. Edges of input image are found out using Ncut. Ncut is used to locate edges and to segment an image into multiple regions. Experiments

are carried out on two datasets one consisting of 27 images and another consisting of 26 images.

Regional contrast (RC) method two stage scheme (TSS) method are used in addition to proposed approach for foreground extraction. Regional contrast (RC) evaluates global contrast differences and spatial coherence. Two-Stage Scheme(TSS) is used for bottom-up saliency detection. This approach is best suited for images in which foreground identification is easy by using interactive foreground extraction. In easily in definable foregrounds, this approach performs better than other cases. [8]

3.5 Figure-ground segmentation using foreground/background segmentation algorithm

This approach works on two datasets PASCAL 2009 and 2010 segmentation benchmarks. It extracts foreground masks in order to detect objects using newly proposed foreground/background segmentation algorithm. A trained object classifier is used to classify different image segments. This approach proposes a segregation algorithm which achieves figure-ground segmentation, leaving the objects of interest untouched and whole. Before applying recognition algorithms, image is pre-segmented. A model which segments an image into foreground and background is learnt. For each image, ground-truth pixel wise labeling is given. These pixels provide the visual characteristics of foreground objects. Geometric cues including shape and image location of foreground objects is also given. This approach combines these cues into a single framework. This method calculates foreground mask of an image and this mask can be given as input to any recognition algorithm. Foreground area is used to compute new features. Following figure[5] shows results of experiments on VOC09 dataset. It shows mean overlap score. The experiment is carried out for various number of nearest neighbours.

The proposed foreground/background algorithm effectively and efficiently finds out foreground mask of an input image. Results of the implementation of this algorithm on the given two datasets are also satisfactory. [9]

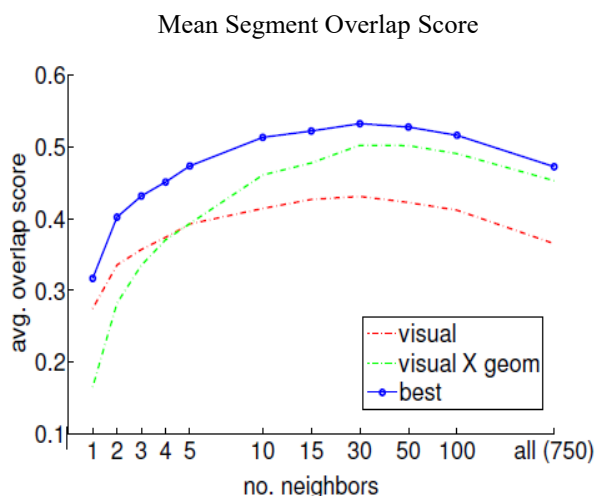


Figure 5: Mean segment overlap score on VOC09 for varying numbers of nearest neighbors

3.6 Figure-ground segmentation using multiple cues

Figure-ground segmentation is addressed through system perspective. 3-D cues are used for figure-ground segmentation. 2-D histograms are segmented by a multiscale approach. Multiple cues are integrated to achieve objective. Classes of data with uniform values can be detected from histogram of data of an image. Peaks of histogram denote clusters. Noisy data may cause problem in identifying peaks. Hence, multiscale approach is needed. 2-D histogram is constructed from 2 independent feature domains. Depth map motion map, confidence maps can be used for 2-D histogram. Multi-scale method is applied to detect peaks at different scales.

Multidimensional clustering method is used for integrating different cues. Integrated multiple cues maintain segmentation under varying circumstances by taking into consideration particular properties in specific condition. Depth and motion can be used for figure-ground segmentation. Analysis of 2-D histogram is done with peak finding algorithm. Object mask is produced by back projecting the most significant peaks in histogram. Cue selection algorithm composed of color algorithm and image motion based algorithm is implemented. When very few things are known about an image and features which are invariant to rotation and scaling are to be extracted, then multiscale representation plays a very important role. This approach mainly focuses on multiscale resolution. Selection of number of feature dimensions for clustering is also difficult problem. Multiple cue approach uses 2 feature dimensions. Integration of multiple cues is an important issue in this approach. [10]

4. Conclusion

In this paper, we have discussed about various figure-ground segmentation techniques. Figure-ground segmentation has its application in many fields like object recognition, target tracking, image retrieval, etc. Every technique uses a different approach. Every approach has its own benefits and drawbacks. An approach suitable for one class of images may not be suitable for another class of images. All techniques do not perform equally well for all types of images. Some techniques work on simple datasets while some techniques work on complex datasets. Single approach can work efficiently on one particular criteria but may fail to overcome certain drawbacks. Many factors affect segmentation. These factors affect differently under different circumstances. Research is going on to find efficient techniques for figure-ground segmentation. Efforts have to be made in direction to remove drawbacks of available approaches so that results of segmentation are very accurate. This will benefit the fields in which figure-ground segmentation is used for various purposes like object recognition, target tracing, etc.

There is no any universal scheme which is perfect for figure-ground segmentation. Every approach has some pros and cons. Combination of these approaches may be helpful for finding efficient approach for segmentation which may yield better results.

References

- [1] E. Borenstein and S. Ullman, "Class-specific, top-down segmentation", in Proc. ECCV, 2002, pp. 639–641.
- [2] D. Kuettel and V. Ferrari, "Figure-ground segmentation by transferring window masks", in Proc. IEEE Int. CVPR, Jun. 2012, pp. 558–565.
- [3] Jordan Reynolds, Kevin Murphy, "Figure-ground segmentation using a hierarchical conditional random field", Computer and Robot Vision, 28 May 2007.
- [4] Bosch, A. Zisserman, and X. Muoz, "Image classification using random forests and ferns", in Proc. IEEE ICCV, Oct. 2007, pp. 1–8.
- [5] M. Maire, P. Arbelaez, J. Malik and C. Fowlkes, "Contour detection and hierarchical image segmentation", Vol. 33, no. 5, IEEE Trans. Pattern Anal. Mach. Intell., pp. 898–916, May 2011.
- [6] Cong Bai, Wenbin Zou, Kidiyo Kpalma and Joseph Ronsin, "Online Global Transfer for Automatic Figure-Ground Segmentation", IEEE transactions on image processing, VOL. 23, NO. 5, May 2014.
- [7] J. Shi and J. Malik, "Normalized cuts and image segmentation", Vol. 22, no. 8, IEEE Trans. Pattern Anal. Mach. Intell., pp. 888–905, Aug. 2000.
- [8] Yubo Yuan, Yun Liu, Guanghui Dai, Jing Zhang and Zhihua Chen, "Automatic Fore-ground Extraction based on Difference of Gaussian", 20 July 2014, The Scientific World Journal Volume 2014
- [9] Amir Rosenfeld and Daphna Weinshall, "Extracting Foreground Masks towards Object Recognition", In Proc. 13th IEEE International Conference on Computer Vision, Barcelona, Spain, November 2011.
- [10] Peter Nordlund, "Figure-Ground Segmentation Using Multiple Cues", Stockholm: Numerisk analys och datalogi, 1998, ix, 68 p.

Author Profile



Snehal P. Ambulkar received B.Tech degree in Information Technology from Government College of Engineering, Amravati in 2012, currently pursuing 2nd Year M.Tech in C.S.E., Rajiv Gandhi College of Engineering and Research, Nagpur, Maharashtra, India.



Prof. Nikhil S. Sakhare, M.E. in Mobile Technology, currently working as an Assistant Professor in Department of Computer Science and Engineering, Rajiv Gandhi College of Engineering and Research, Nagpur, Maharashtra, India.