Difference of Boxes Filters Revisited: Shadow Suppression and Efficient Character Segmentation

Erik Rodner, Herbert Süsse, Wolfgang Ortmann and Joachim Denzler

Chair for Computer Vision, Friedrich Schiller University of Jena,
Ernst-Abbe-Platz 2, 07743 Jena, Germany

`http://www.inf-cv.uni-jena.de`
`{rodner,nbs,noo,denzler}@informatik.uni-jena.de`

seit 1558

Introduction

- Robust character recognition often relies on a good segmentation.
- Difficulties: dirt, non-uniform illumination, shadow, ...
- Our method of character segmentation is **simple, efficient and easy to implement**.
- Algorithm overview:
  1. Shadow suppression using multiple difference of boxes filters
  2. Ternary segmentation using locally estimated thresholds
- Applications:
  - license plate recognition
  - ID card recognition
  - arbitrary document analysis systems

Multiple Difference of Boxes

- Base filter: Difference of Boxes Filter [2]
- Simple interpretation of the idea of Vonikakis et al. [3] (hidden in their formulas)
- Definition for an one-dimensional signal g ($m < M$)

$$\mathrm{DoB}_{m,M}(g) = \frac{1}{m}\sum_{i=1}^{m} g_i - \frac{1}{M}\sum_{l=1}^{M} g_l = \frac{M-m}{M}\left(\frac{1}{m}\sum_{i=1}^{m} g_i - \frac{1}{M-m}\sum_{l=m+1}^{M} g_l\right)$$

- Approximation of a Difference of Gaussians or Mexican-Hat filter
- Runtime independent of filter sizes
- Maximum of the result of several DoB filters with different sizes (m_i, M_i) leads to the final filter output of a Multiple Difference of Boxes (MDoB) Filter

Examples

Our character segmentation framework applied to ID cards and license plates

Local Segmentation

- Ternary segmentation instead of binary segmentation
  - Object
  - Background
  - Unknown
- Local binary decision between object and background is not possible for all pixels (e.g. within homogenous regions)
- Solution: definition of a third label "unknown"
- Local decision depends on maximum and minimum in a neighborhood around each pixel: $g_{\max}(\mathbf{x})$, $g_{\min}(\mathbf{x})$

Original image and ternary local segmentation (white: background, red: object, blue: unknown)

Local Segmentation (Algorithm)

1 `Calculate` $g_{\max}$ `and` $g_{\min}$

2 `For each pixel` $\mathbf{x}$:

  2.1 `If` $g_{\max}(\mathbf{x}) - g_{\min}(\mathbf{x}) < \gamma$ `then`

    2.2 `label point as unknown`

  2.3 `else`

    2.4 $T = \frac{1}{2}\left(g_{\max}(\mathbf{x}) + g_{\min}(\mathbf{x})\right)$

    2.5 `If` $g(\mathbf{x}) > T$ `then label point as object`

    2.6 `else label point as background`

Measuring the Quality of Character Segmentations

- Simple measure of segmentation quality as the distance to a given ground truth segmentation
- Base for parameter optimization and method evaluation
- Distance between two components A and B of segmentations:

$$d_{\mathcal{R}}(A,B) = \frac{|A \setminus B| + |B \setminus A|}{|A| + |B|}. \tag{1}$$

- Distance between two segmentations $\tilde{\mathcal{S}}$ (p components) and $\mathcal{S}$ (q components)

$$d(\tilde{\mathcal{S}}, \mathcal{S}) = \frac{1}{q}\left(\min_{\pi} \sum_{k=1}^{q} d_{\mathcal{R}}(\tilde{\mathcal{S}}_{\pi(k)}, \mathcal{S}_k)\right)$$

- Optimization over all injective maps $\pi : \{1,\ldots,q\} \to \{1,\ldots,p\}$ can be carried out using the Hungarian method

Optimal Parameters of the Segmentation Method

- Given several ground truth segmentations $\mathcal{S}_G^i$, one can search for optimal parameters maximizing segmentation quality
- MDoB parameters $\theta = \{m_1, M_1, \ldots\}$
  - number of DoB filters used
  - sizes of inner boxes m_j
  - sizes of outer boxes M_j
- Parameters of our local segmentation method: $\eta = \{\gamma,\ \mathrm{size}(\mathcal{U}(x))\}$
- Optimization criteria using our segmentation quality measure:

$$\epsilon(\theta, \eta) = \sum_i d(\ \tilde{\mathcal{S}}^i(\theta, \eta),\ \mathcal{S}_G^i\ )$$

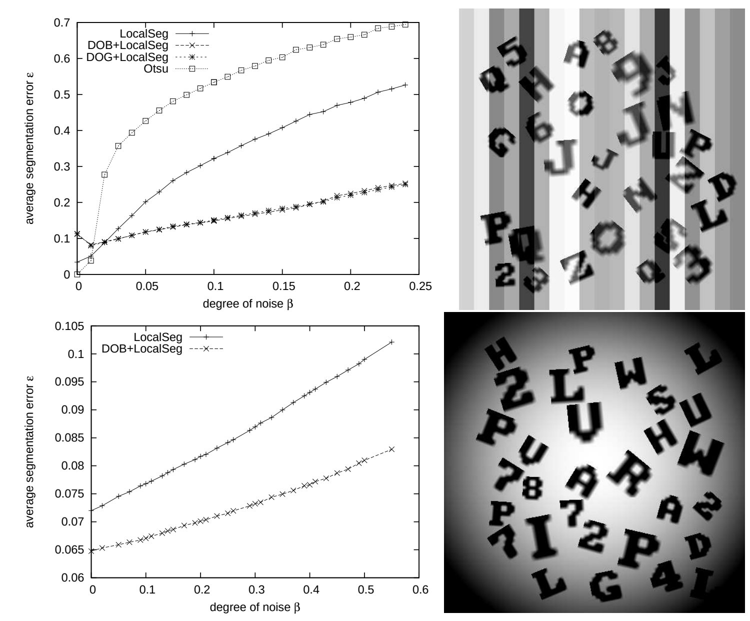
- Optimization performed by cyclic coordinate search [1]
- By iteratively adding a new component to the MDoB filter an optimal number of different DoB filters can be estimated.

Experiments

- Evaluation within a license plate recognition system
  - 6205 test images, fixed set of single letter training images
  - Segmentation framework used to segment an aligned license plate into character regions
  - Recognition performance measured for whole license plates using the complete license plate recognition system

MDoB filters + Local Segmentation	**88.45%**
Local Segmentation	73.47%

- Evaluation using synthetic input images
  - random noise simulating shadow influence parameterized with β
  - left image: analysis of segmentation error with respect to β
  - right image: example of a single synthetic image after applying noise operation



Conclusions

- Simple but robust and efficient method for character segmentation
- Fast computation: combination of basic filter operations
- Proposed measure for segmentation quality can be used for evaluation and optimization
- Optimal parameters of our method can be found with an optimization framework

References

[1] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, August 1999.

[2] A. Rosenfeld and M. Thurston. Edge and curve detection for visual scene analysis. *IEEE Transaction on Computers*, 20:562–569, 1971.

[3] Vassilios Vonikakis, Ioannis Andreadis, Nikos Papamarkos, and Antonios Gasteratos. Adaptive document binarization - a human vision approach. In *Proceedings of the Second International Conference on Computer Vision Theory and Applications (VISAPP), Barcelona, Spain, March 8-11, 2007 - Volume 2*, pages 104–109, 2007.

We would like to thank ROBOT Visual Systems GmbH for financial support and for providing experimental data for large scale evaluation.