Statistical Textural Distinctiveness (TD) for Salient Region Detection in Natural Images

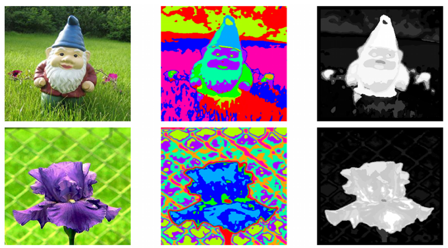
Christian Scharfenberger, Alexander Wong, Khalil Fergani, John S. Zelek, and David A. Clausi
{cscharfenberger, a28wong, kfergani, jzelek, dclausi}@uwaterloo.ca

University of Waterloo
Vision and Image Processing Group
Systems Design Engineering

Abstract

A novel statistical textural distinctiveness approach for robustly detecting salient regions in natural images is proposed. Rotational-invariant neighborhood-based textural representations are extracted and used to learn a set of representative texture atoms for defining a sparse texture model. Based on the learnt model, a weighted graphical model is constructed to characterize the textural distinctiveness between all representative atom pairs. Finally, the saliency of each pixel in the image is computed based on the probability of occurrence of the representative texture atoms, their respective statistical textural distinctiveness based on the constructed graphical model, and general visual attentive constraints. Experimental results using the EPFL dataset and a variety of performance evaluation metrics show that the proposed approach provides promising results when compared to existing saliency methods.

Sample Results



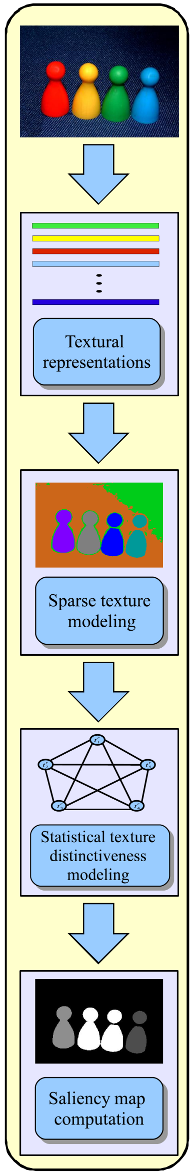
Input | Sparse Texture Model | Saliency Maps

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada and the Ontario Ministry of Economic Development and Innovation.

NSERC CRSNG

Statistical Textural Distinctiveness Model



Rotational-invariant neighborhood-based textural representation *t(x)*

Sparsified representation *t(x)* using u principal components Φ_i of the local representation h(x)

$$h(x) = \langle I(x)\ \ \mathrm{sort}\{I(x_{1,j})\}\ldots\mathrm{sort}\{I(x_{n,j})\}\rangle$$

with the highest variance using PCA.

$$t(x) = \langle \Phi_i(h(x)) \mid 1 \leq i \leq u \rangle$$

Sparse texture modeling via representative texture atom learning

Sparse texture model: Defined as a set of m atoms

$$T^r = \{t_i^r \mid 1 \leq i \leq m\}$$

and learned by minimizing the L_p norm using k-means.

$$T^r = arg\,min \sum_{i=1}^{m} \sum_{t_j^r \in S_i} \|t_j^r - t_i^r\|_p$$

Statistical textural distinctiveness (TD)

t_i^r and t_j^r: Pair of representative texture atoms in the sparse texture model. t_i^r can be seen as a realization of t_j^r in the presence of noise:

$$t_j^r = t_i^r + \eta_{i,j}$$

Textural Distinctiveness: Probability of t_i^r **not** being a realization of t_j^r: $\beta_{i,j} = 1 - P(t_i^r|t_j^r)$

Saliency map computation

Statistical textural distinctiveness

Occurrence probability of the texture atom

$$\alpha_i = \left(\sum_{j=1}^{m} \beta_{i,j} P\left(t_i^r|I(x)\right)\right) \underbrace{\left(\exp\left(-\frac{1}{n_{t_i^r}} \sum_{x \in S_i} \frac{(x-x_c)^2}{\sigma^2}\right)\right)}_{\text{Visual attentive constraint}}$$

$$\boxed{\Psi(x) = \alpha_i,\ \text{if } x \in S_i}$$

Experimental Results

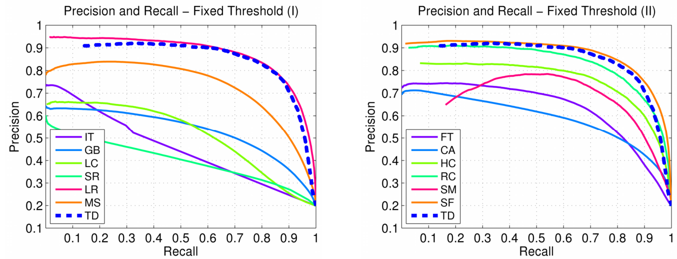


Fig. 1: Precision and Recall rates based on the EPFL database.

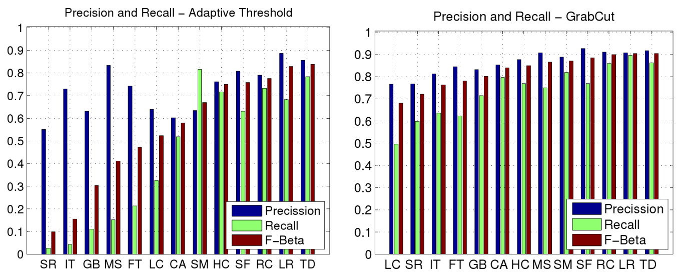
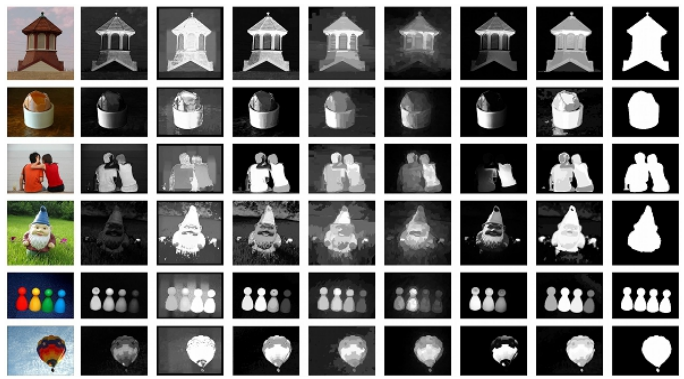


Fig. 2: Precision, Recall and F-measure for adaptive thresholding and for cut-based (GrabCut) segmentation of salient objects, initialized with saliency maps.



Input FT SM HC RC LR SF TD GT

Fig. 3: Visual comparison of our approach (TD) with other saliency approaches and ground truth (GT).