Appl. Comput. Harmon. Anal. 41 (2016) 4–25

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Inducing wavelets into random fields via generative boosting

Jianwen Xie, Yang Lu, Song-Chun Zhu, Ying Nian Wu *

Department of Statistics, University of California, Los Angeles, USA

A R T I C L E   I N F O

Article history:
Received 4 March 2015
Received in revised form 10 August 2015
Accepted 12 August 2015
Available online 24 August 2015
Communicated by Spec. Issue Guest Editor

Keywords:
Generative models
Markov random fields
Simultaneous sparse coding

A B S T R A C T

This paper proposes a learning algorithm for the random field models whose energy functions are in the form of linear combinations of rectified filter responses from subsets of wavelets selected from a given over-complete dictionary. The algorithm consists of the following two components. (1) We propose to induce the wavelets into the random field model by a generative version of the epsilon-boosting algorithm. (2) We propose to generate the synthesized images from the random field model using Gibbs sampling on the coefficients (or responses) of the selected wavelets. We show that the proposed learning and sampling algorithms are capable of generating realistic image patterns. We also evaluate our learning method on a dataset of clustering tasks to demonstrate that the models can be learned in an unsupervised setting. The learned models encode the patterns in wavelet sparse coding. Moreover, they can be mapped to the second-layer nodes of a sparsely connected convolutional neural network (CNN).

1. Introduction and motivations

1.1. Random field models based on wavelets

It is well known that wavelets provide sparse representations of natural images, in that each image can be represented by a linear combination of a small subset of wavelets selected from a given over-complete dictionary. We can exploit this fact to learn statistical models for various image patterns (such as object categories), so that each pattern is represented by a subset of wavelets selected from a given dictionary, while their coefficients (as well as their locations and orientations) are allowed to vary according to a certain probability distribution. Such a model can be written in the form of a Markov random field (or a Gibbs distribution), whose energy function is in the form of a linear combination of rectified filter responses from

* Corresponding author.
E-mail address: ywu@stat.ucla.edu (Y.N. Wu).

http://dx.doi.org/10.1016/j.acha.2015.08.004
1063-5203/

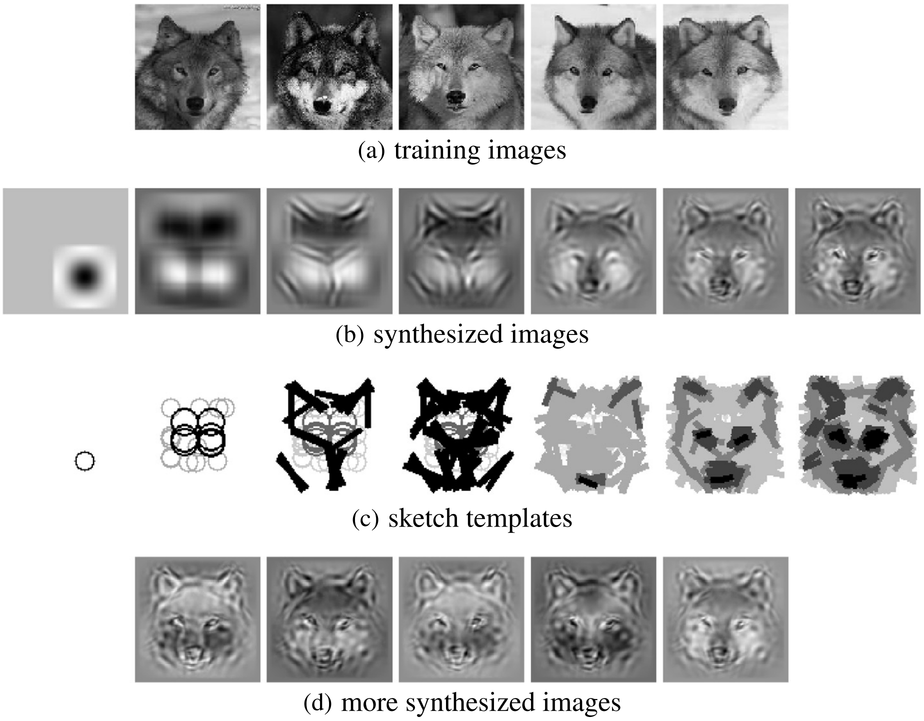

Fig. 1. Learning process by generative boosting. (a) Observed training images (100 × 100 pixels) from which the random field model is learned. (b) A sequence of synthesized images generated by the learned model as more and more wavelets are induced into the model. The numbers of the selected wavelets are $1, 20, 65, 100, 200, 500$, and 800 respectively. (c) A sequence of sketch templates that illustrate the wavelets selected from an over-complete dictionary. The dictionary includes 4 scales of Gabor wavelets, illustrated by bars of different sizes, and 2 scales of Difference of Gaussian (DoG) wavelets, illustrated by circles. In each template, smaller scale wavelets appear darker than larger ones. (d) More synthesized images independently generated from the final learned model.

the subset of selected wavelets [24]. The model is generative in the sense that it is in the form of a probability distribution defined on the space of images, so that we can generate images by drawing samples from the probability distribution.

In this article, we propose a learning method for inducing wavelets into such random field models. It consists of the following two components. (1) We propose to select the wavelets by a generative version of the epsilon-boosting algorithm [10]. We call this process generative boosting because the gradient of the log-likelihood is computed based on Monte Carlo samples generated from the model. (2) We propose to generate synthesized images from the model using a Gibbs sampling algorithm [11] that samples the reconstruction coefficients (or the filter responses) of the selected wavelets, by exploiting the sparse coding form of the model. The proposed learning algorithm identifies important dimensions of variations and generates synthesized images by moving along these dimensions. It also gives a computational justification for sparsity as promoting efficient sampling of the resulting statistical model.

Learning process as a painting process. Fig. 1 illustrates the learning process, which is similar to the way an artist paints a picture by sequentially fleshing out more and more details. (a) displays the training images. (b) displays the sequence of synthesized images generated by the learned model as more and more wavelets are induced into the random field by the generative boosting process. Each wavelet is like a stroke in the painting process. The wavelets are selected from a given dictionary of oriented and elongated Gabor wavelets at a dense collection of locations, orientations and scales. The dictionary also includes isotropic Difference of Gaussian (DoG) wavelets. (c) displays the sequence of sketch templates of the learned model where each selected Gabor wavelet is illustrated by a bar with the same location, orientation and length as