ABSTRACT

High-dimensional (image) data can often be made more manageable and understandable by identifying underlying causes. Each observation is then explained through the interaction of multiple latent factors. We start by discussing a hierarchical mixture model which is based on a recursive partitioning of the image grid. Individual parts take responsibility for a certain fixed subregion of the image. This architecture allows efficient training and inference, and likelihoods can be evaluated exactly. We then continue with a more flexible model in which the spatial support of the parts is learned. For this task we introduce a new interaction rule which creates competition among the parts. The representation we obtain is very intuitive and disentangles factors of variations. We finish by comparing reconstruction errors with denoising autoencoders, restricted Boltzmann machines and dictionary learning for sparse coding.