

**Title:** Feature Extraction for Incomplete Data via Low-rank Tensor Decomposition with Feature Regularization

**Abstract:** Multi-dimensional data (i.e., tensors) with missing entries are common in practice. Extracting features from incomplete tensors is an important yet challenging problem in many fields such as machine learning, pattern recognition and computer vision. Although the missing entries can be recovered by tensor completion techniques, these completion methods focus only on missing data estimation instead of effective feature extraction. To the best of our knowledge, the problem of feature extraction from incomplete tensors has yet to be well explored in the literature. In this talk, we therefore tackle this problem within the unsupervised learning environment. Specifically, we incorporate low-rank Tensor Decomposition with feature Variance Maximization (TDVM) in a unified framework. Based on orthogonal Tucker and CP decompositions, we design two TDVM methods, TDVM-Tucker and TDVM-CP, to learn low-dimensional features viewing the core tensors of the Tucker model as features and viewing the weight vectors of the CP model as features. TDVM explores the relationship among data samples via maximizing feature variance and simultaneously estimates the missing entries via low-rank Tucker/CP approximation, leading to informative features extracted directly from observed entries. Furthermore, we generalize the proposed methods by formulating a general model that incorporates feature regularization into low-rank tensor approximation. In addition, we develop a joint optimization scheme to solve the proposed methods by integrating the alternating direction method of multiplier with the block coordinate descent method. Finally, the proposed methods are evaluated on six real-world image and video datasets under a newly designed multi-block missing setting. The extracted features are evaluated in face recognition, object/action classification and face/gait clustering. Experimental results demonstrate the superior performance of the proposed methods compared with the state-of-the-art approaches.