Speech and visual information

Detecting Object Teaching Utterances (Online)

Speech and image preprocessing

Utterance classification (SVM)

A set of segmented object image features and sentence pairs

Learning Object Image Concept and Word Pairs (Batch)

Clustering object and sentence pairs (MHDP)

Selecting object image concept and word pairs (tf-idf)

A set of object image concept and word pairs