

ABSTRACT

This application provides a multimodal information-based clustering method, a server, and a storage medium. According to the method in this application, a posterior matching probability function in each single modality is generated through self-learning according to multimodal
5 feature information of to-be-clustered target objects, so that the posterior matching probability function in each modality can be generated through self-learning without using annotation information; a posterior matching probability of any two target objects in multimodality is calculated according to the posterior matching probability function in each single modality; and the target objects are clustered according to the posterior matching probability of any two target
10 objects in multimodality to obtain a clustering result. The posterior matching probability that the target objects belong to a same cluster is calculated based on the multimodal feature information of the target objects by using the posterior matching probability function generated through self-learning, and is used as a metric distance for the similarity between the target objects, which can measure the similarity between the target objects more accurately. Clustering
15 is performed based on the posterior matching probability between the target objects, which can improve accuracy and a recall rate of the clustering.

Fig. 2