DOCTORAL THESIS

Advances in Hash-Based Cross-Modal Retrieval

HU, Zhikai

Date of Award:
2025

Abstract

Cross-modal retrieval enables searching across data from different modalities. The introduction of hashing techniques significantly enhances cross-modal retrieval speed and reduces storage costs. However, this also complicates the preservation of semantic information within the corresponding Hamming space. This thesis presents three methods to address this issue from different perspectives.

Joint Semantic Preserving Sparse Hashing (JSPSH) addresses the issue of defining more refined semantic relationships between samples. We find that previous methods often overlook the distribution of labels when defining semantic relationships between samples, which easily results in semantically identical sample pairs. Although there are distinctions between these pairs, they cannot be further differentiated. To address this, we introduce the concept of cluster-wise semantic relationships for more detailed differentiation. In learning hash codes, we simultaneously consider the commonly used sample-wise semantic relationships and the newly proposed cluster-wise semantic relationships. Experiments demonstrate that by incorporating more detailed semantic relationships, JSPSH achieves improved retrieval results.

Key Points Centered Sparse Hashing (KPCSH) filters out more critical semantic relationships between samples. We observe that previous methods tend to preserve all semantic relationships between samples, leading to increased computational complexity and forcing semantically dissimilar samples to cluster together, resulting in erroneous retrieval results. To address this issue, we propose first defining key points and selecting

them from the training set. When constructing semantic relationships, we only consider the relationships between training samples and these key points. This approach prunes away unimportant edges in the relationship graph, reducing computational complexity and effectively preventing the clustering of semantically dissimilar samples due to excessive pathways. Experiments indicate that hash codes learned by retaining only key semantic relationships can achieve better retrieval results.

Semantic Channel Hashing (SCH) improves the retention of semantic relationships in Hamming space. We find that due to the differences between Hamming space and real-number space, i.e., space gap, existing methods face issues of solution space compression and loss function oscillation. We design experiments to verify these claims and conclude that the unique properties of Hamming space necessitate a more nuanced treatment of sample pairs with different semantic relationships. Consequently, we propose dividing sample pairs into three categories based on their semantic relationships: fully semantic-positive, partially semantic-positive, and semantic-negative. We then assign them different semantic channels in the Hamming space to avoid issues of solution space compression and loss function oscillation. Experiments demonstrate that bridging the space gap significantly improves retrieval performance.

Keywords: Cross-modal retrieval, Hashing, Semantic relationship exploration, Space gap