

Preface

With the rapid development of Internet technology and intelligent mobile devices, multimodal data in the network is growing rapidly. Different types of multimodal data are growing rapidly on various Internet platforms such as social networking sites, video sites, music platforms and news information. Multimodal data are increasingly fused together to form a complex organizational structure and correlation. Due to the explosive growth of multimodal data, the demand of users for data and the efficiency of data retrieval are also improved.

The single modal mutual search includes following methods, such as text retrieval, image retrieval, audio retrieval and video retrieval. Nowadays, such retrieval can no longer meet the diverse retrieval needs of users. Cross media retrieval technology can effectively mine the relationship between different data modes and complete the migration between different modes. It also enables the computer to simulate a series of information processing processes such as human brain's cognition, learning, reasoning and decision-making of different types of multimodal data. It has a wide application prospect in many scenarios such as machine learning, pattern recognition and so on. Users can create and share different modes of data (text, image, audio, video, etc.) anytime and anywhere, upload them to social networking sites such as Facebook, twitter, WeChat, and evaluate the content published by others with different modes of data.

The object of cross media retrieval is multi-modal data with heterogeneous underlying features. The characteristics of multi-modal data are as follows:

(1) Multi-modal data (text, image, audio, video and 3D model) coexist. The organizational structure and correlation of data is complex, and semi-structured and unstructured data account for the majority. It is very difficult to store multimodal data directly and structurally and retrieve it effectively.

(2) Multi-modal data express the same semantic information through complementarity and enhancement.

(3) It is necessary to realize the leap from one mode to another according to the various relationships between modes.

(4) With the rapid growth of multimodal data, the existing human data processing ability is difficult to cope with, and the underlying multimodal data is mixed with a lot of noise. The correlation between different modal data is relatively hidden, so it is not easy to mine the correlation.

Under the background of the rapid development of cross modal data, we should deeply understand and mine the information contained in the data and establish the correlation between cross modal data, so as to effectively retrieve the modal information we need from the massive multi-modal data, which can effectively provide better services for people, improve the efficiency of people's study

and work, and facilitate people's production and life.

The publication of this book has been helped and encouraged by the editors of Scientific Research Publishing. Special thanks to Professor Hu Ning, Professor Wang Jing, Professor Jia Guangyu and Professor Wang Yihua from Shandong Jianzhu University, as well as all experts who put forward valuable opinions. I hope this book could help readers understand cross media retrieval technology to a certain extent.