

## 摘 要

随着多媒体技术的飞速发展，数字多媒体数据正以指数数量级的速度在互联网上增长。如何快速，准确地找到自己所需的多媒体数据已经成为用户最关心的问题。传统的基于内容的多媒体分析和检索技术无法跨越低层特征和高层语义特征间的“语义鸿沟”。为解决这个问题，研究人员提出基于语义的多媒体分析和检索技术。

本文以图像为研究对象，首先概述了图像语义分析和检索技术的国内外发展现状以及一些主流的图像检索系统，然后研究了基于上下文的图像语义分析和检索技术中的图像语义标注和检索结果排序等方面的问题，最后讨论了跨媒体搜索引擎中的主要研究问题，并开发了两个原型系统。

在基于上下文的图像语义分析和检索问题中，本文以数字图书中的插图为研究对象，研究了文本-区域匹配模型在插图语义标注问题上存在的问题，并提出了基于多粒度数的插图语义分析模型。在多粒度树模型中，我们建立了图书、章节和插图三层粒度结构，插图所在的图书，章节以及周边文字被视为插图的语义上下文，最后通过概率模型计算粒度内和粒度间的语义相关性来确定标注词和插图的匹配程度。在插图检索问题中，本文提出了基于多样性的检索结果排序算法。与传统的基于相似度的排序算法相比，本文提出的方法能够在保证结果相关性的前提下，在排名靠前的检索集中尽量涵盖更多和用户需求相关的子话题（sub-topic）。这样一来，即使当用户的检索意图并不明确时，用户可以很方便的从广泛的子话题中挑选自己所需的部分。

最后，本文讨论了跨媒体搜索引擎中的几个重要问题：多模态用户接口，统一多媒体数据描述框架以及跨媒体分析模块，并开发了基于关键字查询的插图检索系统 *Illustrator* 和名人检索系统 *MSearch*。

**关键词：** 上下文，图像分析与检索，多粒度树模型，检索结果排序算法，跨媒体

---

## **Research on Context-based Image Analysis and Retrieval**

Song Kai (Computer Application)

Directed By Huang Tiejun

As the multimedia technology develops rapidly, the amount the multimedia data increases through Internet exponentially. Therefore, effective and efficient retrieval for the data become the key issues for users. To overcome the semantic gap between the low-level visual features and high-level semantic features, which cannot be solved by content-based multimedia technology, researchers have proposed semantic-based multimedia analysis and retrieval technology.

This paper first surveys the state of the art in the areas of image analysis and retrieval, and then proposes context-based image analysis and retrieval technologies. Finally, the paper discusses the key problems in cross-media retrieval systems.

In the problem of context-based image analysis and retrieval, this paper mainly focuses on illustration annotation and retrieval results ranking algorithms. The paper proposes multi-granularity tree model to represent the three-tier structure book-chapter-illustration in digital books, in which the books, chapters and collateral texts are considered as contextual information of the illustrations. Then inter-granularity and intra-granularity relations are used for matching probability calculation. In addition, this paper also proposes a novel ranking algorithm which based on topic diversity. Unlike the relevance-based traditional ranking schema, the proposed algorithm aims to cover results with as many different sub-topics as possible in the top retrieval results. Therefore, users can pick up their desired results easily even if their query requirements are ambiguous. Experimental results demonstrate that the proposed method can improve the topic coverage of retrieval results without loss of relevance.

Finally, this paper discusses the key problems of cross-media retrieval systems, which entails multi-modal user interface, unified multimedia description schema and cross-modal analysis. Additionally, a keyword-based illustration retrieval system *Illustrator* and a celebrity retrieval system *MSearch* are implemented.

**Keywords:** Context, image analysis and retrieval, multi-granularity tree, ranking, cross-media