

# 前 言

基于内容的视觉信息检索（主要包括基于内容的图象检索和基于内容的视频检索）是当前信息领域的重要研究热点。它是随着科学技术的进步发展和推广应用，在数据快速增长，信息急剧膨胀情况下为满足人们快速提取有用视觉信息的需求而逐步被重视起来的。

基于内容的视觉信息检索是一个挑战性的研究，涉及许多学科和领域。从研究内容来看，它与数学、物理学、生理学、心理学等许多学科密切相关，从技术基础来看，主要包括图象技术、计算机技术、数据库技术等信息技术，而从发展应用来看，它与工业、通信、娱乐、医学、遥感等许多领域也是不可分割的。

基于内容的视觉信息检索从正式提出至今应该说已有 10 个年头，简单回顾一下与本书写作有关的几个时间标记点也很有意义。

1992 年，基于内容的图象检索（CBIR）一词开始在国际上正式使用，同时基于内容的视频检索（CBVR）研究也已开展。那时许多人还基本将多媒体仅与存储和快速传输各种媒体信息的技术联系在一起，而基于内容的视觉信息检索从信息加工的角度指出，对多媒体信息进行自动查询和搜索也是非常重要的。自己有机会参加了一些相关会议，听了一些有关报告，使自己对图象处理、分析和理解有了较全面的了解和更深刻的认识。

1994 年，国内开始关注这方面的研究，自己先后在一些刊物和杂志上概括通俗地介绍了有关的原理方法和研究应用的特点，以唤起国内科技人员的注意和重视。并在为清华大学电子工程系新开研究生课程“图象工程专题（英语授课）”的规划中加入了相关内容（该课 1995 年正式开出）。同时，自己也结合在相关领域已有的工作基础，开始了基于内容视觉信息检索方面的研究工作，这个研究方向的选择现在看来是很值得的。

1996 年，自己在这方面的研究工作得到了国家自然科学基金项目（69672029）的支持，其后几年又陆续得到国家高技术发展计划（863）基金项目（863-317-9604-05），清华大学创新基金项目（EE9906），国家自然科学基金项目（60172025）等的支持。非常感谢这些基金项目的支持，自己得以在这个方向开展了一系列研究工作，并先后在国内外发表了相关研究论文约 80 篇，为本书的写作打下了坚实的基础。

1998 年，自己应邀在中国图象图形学会主办的第 9 届全国图象图形学术大会上做了有关基于内容视觉信息检索的特邀报告（其后还在多个大学和会议上介绍过相关内容，包括在第 10 届全国图象图形学术大会上），受到与会人员的欢迎。与同行的交流，使自己萌生了写一本有关基于内容视觉信息检索专门书籍的念头，并于 1998 年和 1999 年在清华大学电子工程系博士生前沿课中介绍国际标准 MPEG-7 和基于内容的视觉信息检索技术（2000 年开始转为国际标准 MPEG-21 和数字图象水印技术）。

2000 年，自己在每年一次（从 1996 年开始）撰写的中国图象工程系列综述中增列了基于内容的图象和视频检索文献类（由于当年开始有相当数量的有关文献发表）。后来，该类的文献统计量逐年上升，已成为中国图象工程近年发展的主要热点之一。同年自己在清华大学为研究生正式开出了一门新的专业课——“基于内容的视觉信息检索”。

2002 年，自己在多年积累的基础上，正式开始了本书的写作，期望用一年时间，将基于内容的图象检索和视频检索 10 年来的一些发展和成果进行讨论综合，进一步推动相关研究和应用的开展。10 年弹指一挥间，这期间基于内容的视觉信息检索的研究和应用已经发生了巨大的变化和取得了瞩目的进展，这既为本书的写作提供了广泛的素材，也对本书的选材是一个挑战。

本书在编写方针上，基本上从一本专著和科研的角度出发，期望结合理论性、实用性、系统性和实时性。本书既参考了许多他人的有关文献，更结合了作者近年在该方面的研究工作，不仅有较系统全面的原理介绍，还结合科研成果给出了许多实例。本书不仅选取了比较典型和成熟的技术方法，还注意选取了近年来国际上的最新研究动向。本书从内容上试图尽可能地覆盖基于内容视觉信息检索的方方面面，希望读者能借助本书独立地和全面地了解该领域的基本理论、实用技术、应用情况和发展前景，并可以投身到有关研究和应用工作中来。同时，本书也可作为研究生专业课教材和教学参考书。

本书从结构上包括4部分。第1部分（包含第1, 2, 3章和附录A, B）概括介绍主要概念和基本原理以及典型系统的特点，并介绍国际标准MPEG-7和提供进一步参考资源；第2部分（包含第4, 5, 6, 7, 8, 9章）介绍基于特征的图象检索方法和系统，讨论了采用不同图象特征（包括颜色、纹理、形状和空间关系）的技术；第3部分（包含第10, 11, 12, 13, 14, 15, 16章）介绍基于内容的视频检索方法和应用，涉及到数据库管理和模型、镜头检测、聚类描述、组织索引、表达查询、分析检索、浏览显示等不同技术；第4部分（包含第17, 18, 19, 20, 21, 22章）介绍一些当前研究动态和方向，如压缩域检索，网络搜索、语义检索、交互反馈、性能评价，并讨论最新研究和应用发展趋势。

本书从篇幅上看，共有22章正文和2个附录，在与此对应的24个一级标题下有108个二级标题（节），再之下还有260个三级标题（小节），从内容上大体覆盖了基于内容视觉信息检索的各个重要分支和基本技术。全册书共有文字（也包括图象、绘图、表格、公式等）约70多万字。本书共有编了号的图285个（包括几百幅图片）、表格55个、公式507个。另外书中列出了直接引用的约1000篇有代表性参考文献的目录和一些相关的刊物会议名称及网站、网页地址，希望对读者进一步深入学习有关内容，了解技术细节和开展科研工作有所帮助。

本书基于作者和许多学生的共同研究结果。多年来，先后有刘忠伟、姚玉荣、陆海斌、徐寅、杨翔英、高永英、罗运、李娟、俞天利、李勃、黄翔宇、戴声扬、姜帆、徐丹等直接参与了有关的研究工作，本书也使用了他们学位论文工作的一些内容和他们收集与实验得到的图片。另外，还有许多其他学生（包括作者所在的图象工程研究室的学生）也为本书以不同方式做出了贡献，特别是3年来参与“基于内容的视觉信息检索”研究生专业课的学生，他们的提问、作业、项目训练，以及与他们的讨论都为本书的选材和写作提供了许多有益的启发。

本书在整个编写过程中得到王淑兰编审的热心支持，她认真组织了对本书的审阅修改，既使全书大为增色也使作者受益匪浅。

最后要指出，作者妻子何芸、女儿章荷铭和家人在各方面的理解和支持，是作者决心写作并最终能在今年的最后一天把本书写到这最前一页的重要动力和主要保证！

**章毓晋**

2002年12月31日于蓝旗营

通信：北京清华大学电子工程系，100084

电话：(010) 62798540

传真：(010) 62770317

电子邮件：[zhang-yj@tsinghua.edu.cn](mailto:zhang-yj@tsinghua.edu.cn)

个人主页：[oa.ee.tsinghua.edu.cn/teachers/zhangyujin](http://oa.ee.tsinghua.edu.cn/teachers/zhangyujin)

(已更新)