摘要

随着智能手机和移动互联网的发展，用户可以通过智能手机方便地访问数字图书馆内的图书资源。由于数字图书馆蕴藏着大规模数据，一个便捷高效的搜索引擎有助于用户准确定位所需信息。目前大部分数字图书馆采用基于文本关键字的搜索引擎，但是这类搜索引擎存在两个问题：（1）使用不便捷，需要多次交互操作（输入关键字，选择搜索范围等）。（2）用户往往无法用文字准确表达他们的搜索意图。

将移动视觉搜索技术引入数字图书馆将有效解决这两个问题，用户无需思考如何表达搜索意图，也无需进行多次交互操作，仅需用手机拍摄一张书皮，书页或者书中插图等视觉元素的照片上传至服务器即可搜索相关的书籍信息。但是将移动视觉搜索技术应用于数字图书馆遇到三个问题：

（1）如何描述书中插图的视觉特性。书中插图大多是线图，仅由若干线条组成，线条之间是平滑的白色背景。一方面，常用的基于局部兴趣点的描述子（如 SIFT）应用于插图时检测出来的兴趣点不稳定，其尺度和位置会随着拍摄姿态不同和变化，致使无法准确描述书中插图的视觉特性。另一方面，基于形状的描述子只关注物体的外部轮廓，无法区分轮廓相同但内部线条结构不同的两张插图。

（2）如何检索英文文档图像。常用的基于词袋模型（BoW）的检索方法能较好的解决中文文档图像检索（mAP 为 80%），但应用于英文文档图像检索时搜索性能很低（mAP 仅为 16%）。这是因为英文字符仅包含 26 个字母，英文单词之间视觉上较为相似，使得从英文文档图像上提取的 SIFT 特征彼此距离很近，在聚类生成码书的过程中容易出现量化误差，令不匹配的 SIFT 特征被量化为码书上同一个词，最终导致检索性能较低。

（3）如何降低传输延迟。由于移动互联网的带宽较窄并且不稳定，直接传送手机拍摄的查询图像将导致传输延迟，增加了用户等待时间，最终影响用户体验。

本文的主要贡献在于解决了将移动视觉搜索技术应用于数字图书馆过程中遇到这三个问题，具体包括：

（1）提出了局部形状上下文特征，使得对于插图的检索在召回率为 10%时

为准确率达到 90%。

（2） 提出了基于汉明嵌入的 KD-Tree，在维持原始 KD-Tree 对于中英文文档图像检索时的高精确度（90%）的同时，将内存开销降至 1/16 左右。

（3） 提出了基于 JBIG2[22]压缩的低比特视觉搜索方案，使得手机客户端与服务器之间每次查询传输的图像大小控制在 30KB 左右，降低了传输延迟。

关键字：数字图书馆，低比特移动视觉搜索， 形状上下文，KD-树，汉明嵌入

Research and application of low bit rate Mobile Visual Search for Digital Library

Zhang Chen (Computer Applied Technology)

Directed by Ling-Yu Duan

Nowadays, the popularization of smart phones makes it possible for people to access the digital books via wireless network. Since there is great amount information in Digital Library, an efficient search engine is in great need. However, the popular textual keywords-based search engine demands not only the precise expression of users' retrieving intention but also trivial interactions of typing and choosing the searching range.

By bringing the Mobile Visual Search into Digital Library would solve this problem. Instead of figuring out the precise expressions and several interactions, people could initiate a query by a simple capturing of photos. Though Mobile Visual Search has applied successfully in many fields like landmark search, product search etc., directly applying the existing technologies of Mobile Visual Search into Digital Library faces three challenges:

i. The frequently used local visual features, e.g. SIFT and SURF could not describe the visual characteristics of the line-drawings or sketches in books.

ii. The popular Hierarchical clustering algorithm, e.g. Hierarchical K-Means would lead serious quantization errors if applied to generate codebook of the visual features extracted from English Document Images, and consequently influents the retrieval performance. Though some Nearest Neighbor Algorithms, e.g. KD-Tree and LSH function well in Retrieving English Document Images, they suffer high memory expense because they have to store all the features in memory.

iii. The existing Low Bit Rate Mobile Visual Search technologies, e.g. CHoG, PQ-SIFT, LDVC, MCVD fail to achieve the balance between transmit latency and retrieval performance on document images. On the other hand, the method which compress query image in mobile phones could not

achieve high compression ratio.

This paper tries to solve all these three problems, and our contributions include:

i. We propose a novel visual descriptor, named Local Inner-distance Shape Context (LISC), to describe the line-drawings' visual properties. LISC combine the advantages of both local visual descriptor and the descriptive power of Shape Context on shapes. Besides, LISC can be applied seamlessly to inverted index for real-time search on a large scale database of sketches. The experiments on MPEG-7 Shape 1 Part B and the line-drawings dataset from industrial books and Geometry books testify the effectiveness of LISC.

ii. We propose Hamming-Embedding-based KD-Tree (HE KD-Tree) to search Document Images. HE KD-Tree decreases the memory overhead and at the same time guarantees the high retrieval performance and it can also accelerate finding the Nearest Neighbor in KD-Tree. The experiments on both Chinese and English document images demonstrate the advantages of HE KD-Tree.

iii. We also propose a JBIG2-based compress scheme for document images to achieve the low bit rate Mobile Visual Search of Digital Library. This scheme reduces the transmit latency on wireless network without sacrificing the retrieval performance.

iv. Finally, we integrate these three innovations to develop an low bit rate Mobile Visual Search app for Digital Library, which can search the relevant book covers, book pages and line drawings in books.