

# 基于 **Lire** 的分析与扩展 项目计划

Version 2.0

小组成员：

刘少凡

宋昱材

吴沂楠

黄飞

## 版本变更记录

版本	变更时间	修改人	审核人	备注
0.5	20170315	宋昱材	黄飞	完成前言和项目概述
1.0	20170315	黄飞	宋昱材 刘少凡	完成实施计划
1.1	20170316	宋昱材	刘少凡	修改 3.5 和 3.13 节部分内容
1.2	20170316	刘少凡	吴沂楠	修改与完善细节
1.3	20170320	刘少凡	宋昱材	依照同学评审修改部分内容
2.0	20170427	刘少凡		依据现阶段完成情况进行更新

# 目录

1. 前言.....	4
1.1 目的.....	4
1.2 术语与缩略语.....	4
2. 项目概述.....	4
2.1 简介.....	4
2.2 Lire 源代码简要分析.....	5
3. 实施计划.....	6
3.1 工作内容分解.....	6
3.2 参加人员.....	6
3.3 完成项目最后期限.....	7
3.4 本计划的批准者与批准日期.....	7
3.5 项目分工.....	7
3.6 组织形式.....	7
3.7 交付的产品.....	7
3.8 进度安排.....	8
3.9 时间预算.....	8
3.10 遵循标准.....	8
3.11 硬件环境.....	8
3.12 软件环境.....	8
3.13 各阶段实验要点.....	8
参考资料.....	9

# 1. 前言

## 1.1 目的

为了便于协调组内成员进行后期的工作，对项目进行跟踪和监控，对任务的进度进行安排与调控，故对后期工作进行计划。

## 1.2 术语与缩略语

CBIR	Content Based Image Retrieval， 基于内容的图像检索。
Lucene	一个基于 Java 的全文信息检索工具包。
MPEG-7	MPEG-7 标准被称为“多媒体内容描述接口”，为各类多媒体信息提供一种标准化的描述，这种描述将与内容本身有关，允许快速和有效地查询用户感兴趣的资料。
PHOG	Pyramid Histogram of Oriented Gradients， 分层梯度方向直方图，一种描述空间形状的特征向量，在不同层次上统计边缘图像的梯度方向直方图分布情况。
CEDD	颜色和边缘的方向性描述符
FCTH	模糊颜色和纹理直方图

# 2. 项目概述

## 2.1 简介

Lire（Lucene Image Retrieval）是开源项目 Caliph and Emir（项目主页：<http://www.semanticmetadata.com>）的子项目。Lire 提供一种构造基于内容的图像检索系统（Content Based Image Retrieval System）的简单方式。图 1 展示了一个 CBIR 系统的一般框

架。

Lire 可以为 CBIR 系统创建图像特征的 Lucene 索引库。Lire 支持多种不同的底层图像特征，如 MPEG-7 标准的视觉描述符，以及 PHOG，CEDD，FCTH 等。

Lire 提供简单且易扩展的索引搜索的方法。

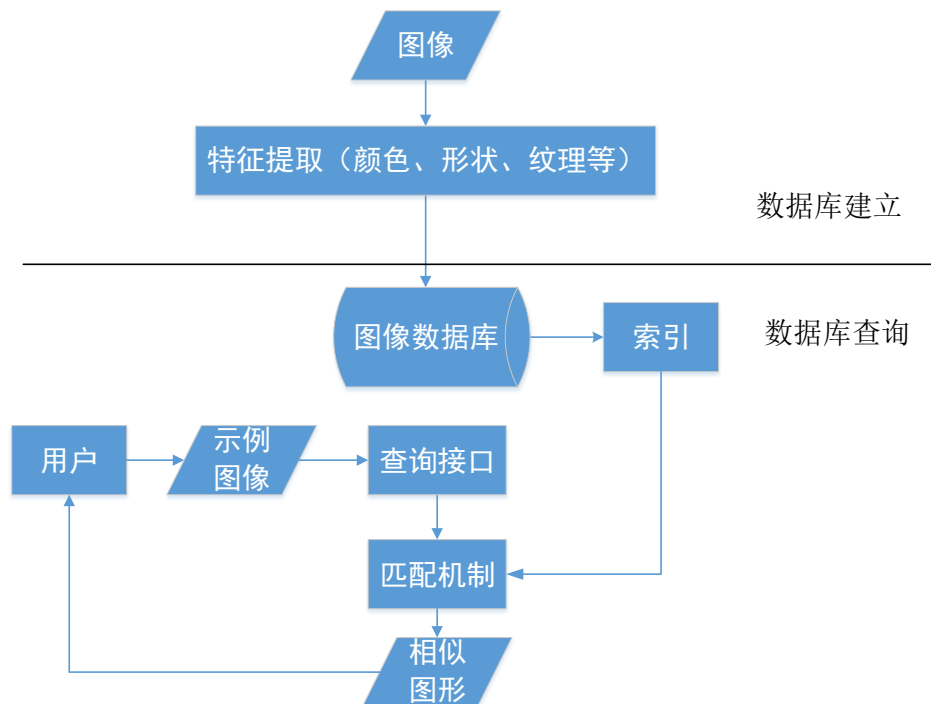


图 1 CBIR 系统的一般框架

## 2.2 Lire 源代码简要分析

### ➤ 根目录：net.semanticmetadata.lire

根目录下提供了 Lire 的基本接口，即生成索引和图像检索。生成索引就是根据图像提取特征向量，然后存储特征向量到索引的过程，以 `DocumentBuilder` 为关键词的文件均实现该类接口。图像检索，就是根据输入图像的特征向量到索引中查找相似图像的过程，以 `ImageSearcher` 为关键词的文件均实现该类接口。

### ➤ 图像特征包：net.semanticmetadata.lire.imageanalysis

`Imageanalysis` 包实现各种图像特征的提取功能。`LireFeature` 类为各种图像特征类的父类，该类定义图像特征类必须实现的各类方法，如特征提取、特征距离计算、特征数据格式转换等。

### ➤ 图像索引与图像检索包：net.semanticmetadata.lire.impl

该包针对不同的特征提取方法实现相应的索引建立方法，每一种方法对应一个 `DocumentBuilder` 和 `ImageSearcher`。

- 索引建立的具体算法包：`net.semanticmetadata.lire.indexing`  
针对索引建立的具体算法实现相应类，如基于哈希的 `LSH` 类。
- 矩阵包：`net.semanticmetadata.lire.matrix`  
定义了距离矩阵、相似度矩阵等接口，用于检索时的相似度计算。
- 搜索结果过滤包：`net.semanticmetadata.lire.filter`  
实现对搜索结果进行过滤的类。
- 常用工具包：`net.semanticmetadata.lire.util`  
实现一些常用的工具，如数据转换、文件操作、图像操作、索引创建优化、计算距离度量、序列化、采样数据等。
- `KMeans` 方法实现包：`net.semanticmetadata.lire.clustering`  
实现 `KMeans` 方法和多线程并行 `KMeans` 方法。

## 3. 实施计划

### 3.1 工作内容分解

首先，了解 `Lire` 的工作流程和主要功能。

然后，利用源代码、开发文档等相关资源，进一步了解 `Lire` 的功能需求，对需求进行分析和归类，细致划分，并且将其对应到源代码的具体模块。

最后，针对项目分析过程中发现的问题，对 `Lire` 进行相应的测试，条件允许情况下对 `Lire` 进行进一步的优化，或者是基于 `Lire` 进行上层应用的开发。

### 3.2 参加人员

刘少凡

黄飞

宋昱材

吴沂楠

### 3.3 完成项目最后期限

本学期第十七周

### 3.4 本计划的批准者与批准日期

刘超老师与任建老师

校历第三周

### 3.5 项目分工

项目已完成任务的分工情况与未来计划任务的分工情况详见项目计划.mpp。

### 3.6 组织形式

- a)每周一次会议，讨论当前工作进展以及各自的问题。
- b)建微信群，保持实时的交流。
- c)资源的共享通过 github 进行实现。

### 3.7 交付的产品

序号	产品名称	交付时间
1	软件项目计划书	第 3 周
2	项目计划.mpp	每周
3	个人工作日志	每周
4	会议记录	每周
5	需求规格说明书	第 4~9 周
6	需求检查单	第 6~8 周
7	软件产品改进与展示	第 9~11 周
8	软件测试分析报告	第 12 周
9	软件测试评审报告	第 13~14 周
10	软件进度分析报告	待定
11	工作量分析报告	待定
12	配置管理	待定

### 3.8 进度安排

项目的进度安排详见项目计划.mpp。

### 3.9 时间预算

已完成任务的时间花费与未来任务的计划时间详见项目计划.mpp。

### 3.10 遵循标准

本项目遵从以下标准：

GB/T 13702-1992 计算机软件分类与代码

GB/T 19003-2008 软件工程

GB/T 9386-2008 计算机测试文档编制

GB/T 9385-2008 计算机软件需求规格说明

GB/T 5532-2008 计算机软件测试规范

GB/T 8567-2006 计算机软件文档编制规范

### 3.11 硬件环境

小组每名成员在实验室都有自己的工作电脑，同时也可使用自己的个人电脑。使用 PC 即可进行代码的阅读、开发、测试、ppt 的制作演示等等。

### 3.12 软件环境

操作系统	Windows 10， Mac Os X 10.12 及以上
源码分析	IntelliJ IDEA
项目管理	Github

### 3.13 各阶段实验要点

实验 1 软件需求分析

获取并说明软件需求，细化并严谨地定义指定的软件需求；进行需求验证和评审，并依



据评审意见对需求进行修改和完善。

#### 实验 2 软件需求评审

采用规范的评审方法，确认软件需求，主要包括网上互评审和会议评审

#### 实验 3 软件产品改进与展示

改进、扩展或应用展示软件特性，进行任务定义、设计与实现。并进行软件更新和追踪分析，软件测试。

#### 实验 4 软件测试

进行测试需求定义和测试用例设计，以及对测试需求和测试用例进行评审。选择并学习使用测试工具，生成并执行测试脚本，进行充分性分析和增强测试，并对测试结果进行评审。

#### 实验 5 软件测试评审

采用规范的评审方法，确认测试需求、测试方法和测试结果，主要包括网上互评审和会议评审。

#### 实验 6 软件进度计划与控制

进行实验任务分解与分配、完成标准和进度计划。对进度进行监控，任务细化与调整。并通过会议研讨、网上交流等形式进行小组协同。

#### 实验 7 配置管理

在软件研发过程中，采用有效方法进行软件变更控制和版本管理，主要的过程要点包括配置管理计划、变更控制、版本管理和配置管理分析。

#### 实验 8 工作量估计与统计分析

通过对实验项目的跟踪，分析工作量差异以及原因。过程要点包括任务工作量估计、任务执行工作量记录、任务调整与控制、工作量统计与影响因素分析。

## 参考资料

[1] <http://www.semanticmetadata.net/lire/>

[2] <http://blog.csdn.net/camu7s/article/details/49611823>