

作词作曲家识别调用以及测试说明文档

作词作曲家识别---测试说明文档（归档文件）

算法：郭辉铭

2023.3.8

一）需求---即代码功能

识别当前曲谱页可能存在的作曲家以及作词家并返回给用户

二）前端调用接口

ocr_infer.h中的std::vector<ppredictor::OCRPredictResult> infer(cv::Mat &original, ppredictor::OCR_PPredictor *orc_predictor, string label_ch_path, string label_fri_path)方法

1) 输入：

略（与识别类似，前已有相关文档说明）

2) 前端需要的数据：

- 该方法的返回值：results
- 该方法内调用infer_ocr()中的地址传递的作词家以及作曲家信息：composition; 作词家信息lyrics。（结果为空则表示不存在作词作曲家信息，或者未正确识别）

三）测试要求

标题识别情况类似

1) 由于曲谱复杂多样，作曲作词的表现形式不一，位置丰富。因此测试量要达到一定要求；

2) 相比于标题而言，作词作曲家部分字体较小，更容易在去阴影二值化后变得模糊不清，因此拍摄时防止抖动为宜。拍摄页面尽量完整。光线尽量均匀（目前应该不影响）。手机太好的，可以忽略以上情况。

3) 较老的手机机型（3年以前）由于cpu以及相机芯片等问题，可能会导致图片处理后更模糊，失真更多，耗时也会更长。

四) 测试样例

- 1) 测试不同类型的曲谱：钢琴谱，吉他谱。
- 2) 测试文字信息含量不同的曲谱：中文较多，英文较多，术语较多。
- 3) 测试作词作曲家位置不同的曲谱：作词作曲在一起，同行分隔有远有近。作词作曲分列两行。
- 4) 测试不同形式的作曲家信息表达：xxx 曲，xxx编曲，xxx 作曲，作曲：xxxx，xxx 编配，XXX改编。
- 5) 测试同文字干扰的作曲家信息：标题下跟着一些副标题：《xxx》主题曲，片尾曲，金曲，歌曲，舞曲，英国儿歌；标题后有大段文字说明解释的。
- 6) 测试并无作曲家信息出现的作曲家识别。比如：周杰伦，莫扎特（mozart），俄罗斯舞曲，英国儿歌，朝鲜族民歌等。
- 7) 测试不含作曲家作词家信息的曲谱：含标题与否。

说明：

- 1) 如果同时出现xxx 曲 xxx 编配 或者xxx 曲xxx改编。则认为是作曲家是xxx（曲）；
如果只有xxx 改编，xxx配，xxx编配。认为xxx为作曲家；
- 2) 右侧出现的信息不含作曲家关键字的，而有xxx民歌，xxx舞曲，xxx儿歌的，则认为这些都是作曲家。

五) 当前作词作曲家识别存在的一些问题----即本身包含作曲家信息却无法识别的情况

- 1) 拍摄模糊导致去阴影，二值化后文字信息损失严重而使得识别信息不准确，可能无法识别作词作曲家内容；
- 2) 拍摄弯曲，边框与页面框相差较大，导致位置扭曲；
- 3) 没有作曲家关键字信息，且出现在标题下方中间位置，或者标题位置的无法识别出作曲家；
- 4) 标题太大，导致去阴影后缺口较多而无法识别标题，也就无法识别作词家以及作曲家信息；
- 5) 肉眼可以分辨为作曲家信息比如：【德】约翰.赛巴斯蒂安.巴赫，周杰伦。但由于该位置可能在标题下方等其他一些非常规位置，导致无法识别。

六) 待改进处

- 1) 时间紧迫，代码需要优化，提高可读性
- 2) 变量创建较多，消耗内存较大
- 3) 一些细节

七) 2023.3.15更新

产品提出bug:

1. 无作曲家关键字，却包含之前已经去除的内容。需要添加进判别中
2. 有些曲谱，边缘留白较多者，无法被识别。

解决