

## 原 2012,2013年总结：在视音频技术道路上摸索

2014年01月05日 00:09:57 阅读数：7939

2013年已经过去了，看到很多人写了各种的总结，感觉这种习惯很值得提倡。

其实每个人都是一部故事，有自己的时间地点人物事件。过去我一直忽略了这一点，从来不记录任何东西。高中的时候我语文不太好，尤其是作文不太好，于是也就觉得自己并不擅长文字类的记录。后来连图片类的记录也很少，几乎不拍什么照片。后来大学毕业的时候，发现自己的照片也不多。想想这都是不应该的。工作生活应该很精彩，值得我们去记录，回忆，总结，甚至批评与自我批评。

本文大体记录一下这两年我在业余时间都干了什么，至于写项目报告，写论文，汇报，接活儿之类的就不细说了。

总体说来，研究生的生活自由度相当的大，如果自己把握方向，就会无所事事。我们学校工程硕士是两年毕业的，这样时间还算比较紧凑。而我这样的学术型硕士，是三年毕业的。第一年上课，第三年毕业设计，第二年就会有大量的空闲时间。如何利用这闲下来的一年，也是一个大问题。

2012年没有写过总结，在此先简单总结一下2012年的工作：

### 1.2012上半年

我们实验室主要是研究视频质量评价的。此前师哥师姐在研究的过程中，一直在用一个软件：JM。可是我在接触了JM之后，发现它虽然代码很正规很容易懂，可是速度太慢了，与实际应用相差很远。JM确实很适合搞学术研究，可是它和我的想法不一样。我希望可以做出实用的，离实际生活很近的视音频技术应用。因而我没有去深入研究JM，转而去寻找实际环境中使用很多的视音频项目。

此外，由于在做一个网络视音频方面的项目，需要研究网络协议。之前师哥师姐都是在研究RTSP+RTP/RTCP这样的流媒体协议。这样的协议组合很适合IPTV系统，我们曾经去上海测试过实际的IPTV系统，确实是用这样的协议。《计算机网络》这样的教科书上也只介绍了这样的协议。但是对实际互联网视频分析的时候，我却一点也没有找到它们的影子。我比较相信互联网视频才是未来的主旋律。因而想要研究互联网视频使用较多的协议。

导师人很好，她给了我自由的空间，去做自己的研究。只要不偏离项目的主题就行。

然而对一个未知领域的摸索也让我吃了很多苦头。很少有可交流的人，只好自己弄。首先为了找到开源的优秀的视音频项目，我花费了大量的时间。有实际应用价值的很多视音频项目都是不开源的商业项目。我把他们一下载，然后测试，最后收获的都是失落。最后，终于找到了FFmpeg。在发现了它的强大之后，我随即发现了迎面而来的问题：该怎么学习它。FFmpeg项目文档比较少，中文的就更少了。而我上研究生之后开始学习C/C++，基础不太牢固，看代码的水平还略有不足。师哥师姐做这方面的相对较少，有师姐用过FFmpeg中的抽取出来解码H.264的代码，并给我了相关的指导；但是像libavcodec，libavformat这些完整的库就得自己研究了。从网上找了很多的例子，但绝大部分都是不能一次编译通过的；而有些能编译通过的，又比较陈旧。我印象最深的是：有一次下载了一个项目，竟然可以编译通过，很高兴。但是随即发现，它可以解码AVI封装的H.264，却不能解码FLV封装的H.264。而FLV是互联网上相当通用的封装格式，H.264也是相当通用的视频编码标准。如果不能解码这种格式，工程就没有意义了。但是我又实在找不出来其它可以编译通过的例子了，于是把工程的库换成最新版的，然后对着FFmpeg的API文档一个一个的看，把旧的函数换成了新的。比如av\_open\_input\_file()换成avformat\_open\_input()，avcodec\_decode\_video()换成avcodec\_decode\_video2()。最后一次当编译器输出了很多warning却没有输出error的时候，我兴奋了很久。总算实现出来一个最简单的播放器。现在已经总结到自己的文章《100行代码实现最简单的基于FFmpeg+SDL的视频播放器》里面了。后来为了方便以后其他人学习FFmpeg，我把自己学习FFmpeg的资料通通总结到自己的博客上了。

此外，为了寻找互联网视频使用较多的协议，我做了很多的分析工作。比如说，用WireShark抓取正在播放的互联网视频的数据报。然后把数据报的内容取出来进行分析。在多次的尝试之后，我发现：直播使用最多的是RTMP，主要因为它可以实现无插件直播；点播使用最多的是HTTP，主要因为它不丢包，而且节约服务器资源。基于HTTP的点播其实不用多学，因为实际上它和下载文件差不多。直播的RTMP协议则是需要学习的，我的精力主要用在了这里。

为了学习RTMP协议，开始也挺不容易的。最主要的原因是RTMP并不是完全开放的。他是Adobe公司的私有协议。找来找去，大部分文章都是协议破解之类的文章。后来找到了协议规范，但是发现有的地方含糊其辞（也许是Adobe公司出于商业目的的考虑），让人很难彻底掌握。而且也没找到最合适的开源项目。其实RTMP相关的项目倒是有一些，一直没确定应该看哪个项目。最后终于决定，把libRTMP看一看。libRTMP差不多是我看的第一个视音频工程内部的代码了，通过查看libRTMP的源代码，我渐渐的掌握了RTMP。

### 2.2012下半年

这时候的我第一次感觉到了人生的迷茫：以后该做些什么？毕竟，没有一样拿的出手的“东西”，会让人感到不安。学了技术，全都存在脑子里，也没太发挥出来作用。我有一种比较笼统的观念：第一，以后我可能要搞视音频了；第二，我应该搞出个东西来。但是什么都光想不做的话，黄花菜都会凉了。在学校已经一年了，我也基本上了解了学校挣钱的方法：做项目，或者是做产品。做项目是最普遍的选择。做项目有两个优点：一是可以先获得经费；二是项目要求相对宽松。做项目有一个不足：就是每次接下一个项目，都是零的开始，需要重新学习。做产品的特点基本上和做项目是反着的。做产品有两个不足：一是必须得先做出来一个东西，并且有一定的前景，才可能会有收入；二是产品的要求是非常严格的。做产品有一个优点：做出一个东西并成功销售后，就可以简单地复制粘贴提供给很多的客户，可以获得很大的回报，而且以后可以围绕着产品进行研究，有利于技术积淀。我想试一下做产品，尽管难很多。我有点急切，迫切想要做出点东西。这时候我感觉自己有点忙乱，我给自己定下了一个目标：做一个教室直播的系统。我当时这么做只有个简单的想法：可以把学校自习室的监控信号转播出来，让大家方便看每个自习室的人数情况，方便大家找座位。可是这个想法却有着很大的争议。我跟一两个人说过，他们都觉得这有点侵犯别人的隐私。导师让我们想想自己毕业设计做什么，我把这个想法说给了导师，导师也并不认同。我一时间感觉很失落。但是我在想：我应该先把东西搞出来再说。研二的时间很充裕，我就开始搞我这个系统。捡起了本科时候学习的JavaEE，再结合新学的流媒体技术，自己设计前台，后台，各种界面；前台用div + css + jQuery，后台用Struts2 + Spr

ing + Hibernate,流媒体技术采用RTMP相关技术,搞出了一个Demo,可以运转。可是这时候我又想到,还有很多的功能没有实现:按时间直播,按课程表直播等等。渐渐地,慢慢就有点乱了套。想法一直在变动,数据结构变得比较模糊,开发的难度慢慢变大。我想到一个点子,就弄一个点子,而没有提前计划好,最后自己都不想弄了。于是这个想法就搁置了起来。

教室直播系统不成功之后,加之和好朋友聊天得知大家进入社会后诸多不顺利的经历,让我感觉社会仿佛有一层灰色的感觉,产生了一些负能量。十月份去重庆参加了一次学术会议,然后又去上海测试了一下自己实验室开发的软件,相当于出去放放风,溜达了两圈,还是挺愉快的。但是回来后又要面对同样的问题:以后该做些什么呢?一度又听说做项目管理软件可以卖给关系单位,但是前提是必须把软件开发出来。我又埋头开始开发项目管理软件,还是用开发教室直播系统的那套技术。现在看来,我可能真的是有点昏了头了。细细想想,开源的项目管理软件都很多,当时为什么不选择二次开发而是直接去开发呢?再者,就算开发出来,成功卖给了关系单位,那未来又能怎样呢?我做的软件和现有的软件比起来,会更有竞争力吗?项目管理软件的市场真的是已经太成熟了,凭学生的力量已经不可能挤进去了。

结果我在开发方面就遇到问题了,因为项目管理软件的需求还是很多的,不光是项目,人员这些就够了。具体的技术很杂,还包括导入Excel,发邮件,文件上传,文件预览这些功能。这些功能虽然都不复杂,并且有相应的工具的支持,但是这些功能的数量实在是挺大的。简而言之就是类似“搬砖”的工作。搞了好久,感觉自己有点搞不动了。其实软件已经做出来了,功能基本上已经实现了,就是一些细节的问题。但是我自己有些怀疑这个没有什么很大的前景,继续开发不知道意义何在。现在项目工程文件还原封不动的躺在电脑硬盘里面,等到以后把它上传开源了算了。

2013年我主要做了以下几件事:

### 3.2013年上半年

新的一年开始的时候,我还是面对着老的问题:以后该做些什么?寒假在家的时候,突然意识到:自己已经好久没有研究视音频技术了。确实,在自己之前设计开发系统的时候,用的主要是JavaEE,并没有怎么用过C/C++技术。突然很想研究一下FFMPEG内部是怎么一回事。这时候我已经成功把FFPLAY移植到MFC下面了,这样做主要是因为毕业设计的需要。但是我还没有剖析过FFMPEG内部的结构。反正在家闲着也没什么事,所以用Eclipse + MinGW搭建了一个环境后就开始看起了FFMPEG的源代码。

没想到小小的FFMPEG工程里面竟然包含了一个大大的世界。它的代码严谨,命名规范,很多的地方设计很巧妙,而且c语言的各种知识几乎全都被用上了。我沉入其中,仔细研究,体会到了编程的神奇。

而后,我又发现它其实很强大。不仅仅局限于视频解码。事实上,它可以做和视音频相关的方方面面的工作。掌握了它,基本上就掌握了视音频处理技术了。在学习了一段时间之后,我用它实现了基于内容的媒资检索系统中的一个视频关键帧提取模块,发现以前看起来天书般的任务其实很容易就实现了。我心生感慨:FFMPEG真有用。

结合着FFMPEG的学习,我开始做毕业设计的软件。每了解FFMPEG的一个新功能,我都试着把它加入到自己毕设软件中。慢慢的,我发现自己的毕设软件竟然成了一个功能挺全的互联网视音频的播放,码流分析,图像分析的系统。我感觉很有成就感,于是想办法完善自己的软件。我一边开发毕设软件,一边学习视音频技术。很多的视音频技术,都是在这个时期学习的:H.264, AAC, FLV, TS等等。在学完这些技术后,我都会找找看有没有相关的开源工程,实践一下学到的技术。

三月份左右的时候,第一次写了两篇博客。写博客的动因很简单:我CSDN的下载积分没有了,为了挣一些积分用于下载资源,我上传了自己制作的最简单的播放器,并写了博客作为介绍,但是此后又没有写博客了。

除了视音频技术之外,我又学习了视频质量评价技术。这是我们实验室的“老本行”。我本应该早点学的,但是没有找到兴趣点,所以就一直拖着。当时发现不能再拖了,所以就开始学习它。后来发现,视频质量评价技术其实很重要,是衡量视频好坏的关键技术。学习的过程中阅读了很多的英文文献,感觉英语阅读能力有了一些提高。

此外,出于对视音频技术的兴趣,我又分析了Mplayer, Media Player Classic, ffdshow的源代码,他们无一例外的都是和ffmpeg相关的工程。

暑假的时候,突然想出了一个方向:以后就搞互联网视频技术,重点应该是质量评价。

质量评价领域国内比较空白,目前来看挺难盈利,但是谁又知道以后究竟会怎样呢?

### 4.2013年下半年

今年下半年开始的时候,一个直接的问题摆在了眼前:考博还是工作?

“以后该做些什么?”已经不是头等问题了,因为毕竟先要考虑吃饭的问题。

经过一阵子纠结,最终还是决定考博。我觉得学校可以提供一个相对比较自由的空间。而我整体来说还是比较随性的,学校可能更适合我一些。

九月份的时候,试着想把视频质量评价的一些术语加入到百度百科中。结果添加了十条左右,一条也没有通过,各种原因,反正就是审核不过。我百度百科账户的审核通过率一下降到了个位数。当时感觉:也许太多人不了解视频质量评价了吧。

当时还有一股倔劲,想着百度没给过,难道就放弃了?那我自己写的那些名词解释就都要扔掉了?不。这时候又想起了CSDN博客。我要把它们写到博客里,有人需要的时候自然会搜索到的。

于是就开始写起了博客。先把那些名词贴了出来。帖出来之后,发现竟然真的有人看,看来还是有“同道中人”的。后来想想,干脆把这两年研究的经验都写出来算了。毕竟学校第一重要的责任就是:传播知识。作为在学校的一份子,我有义务这样做,尽管力量可能很微小。

于是我翻出了自己这两年收集的学习资料。它们被我统一的存在一个大文件夹里,杂乱无章的,正好借这个机会整理一下。我编写了许多文章,并且上传了自己用过的一些资源,基本上都是标的0分,只是希望能帮助到大家学习。

随着博客文章的不断更新,我认识到了很多的朋友,他们也提供给我很多的机会,帮了我很多忙,非常的感谢他们。他们也许地理上和我不在同一

个位置，但是和我做着相同的工作，我们有完全相同的话题：网络，视频，音频，编解码，质量评价。我猛然感觉到还是有很多人在视音频领域做相关工作的。

## 2014年需要面对以下几件事：

随着2013年的结束，我之前积累的经验也分享得差不多了。我感觉对于我个人来说，还是需要思考：“以后该做些什么？”

以下几件事在新的一年里需要完成：

### 1. 毕业设计

这个是顺利毕业所必需的。

### 2. 考博士

这个是读博士所必须的。

### 3. 建设一个网站

心中的计划，先好好规划一下吧。

版权声明：本文为博主原创文章，未经博主允许不得转载。 <https://blog.csdn.net/leixiaohua1020/article/details/17851977>

文章标签：[视音频](#) [总结](#) [人生](#) [产品](#) [技术](#)

个人分类：[杂谈/非技术](#) [综合/IT与广电](#)

此PDF由spygg生成,请尊重原作者版权!!!

我的邮箱:liushidc@163.com