申请上海交通大学工程硕士学位论文

基于Gstreamer的Chromium音视频播放系统的设计与实现

|  |  |
| --- | --- |
| 学校代码： | 10248 |
| 作者姓名： | 胡济豪 |
| 学 号： | 1130379118 |
| 第一导师： | 姚建国 |
| 第二导师： | 王圣浩 |
| 学科专业： | 软件工程 |
| 答辩日期： | 2017年1月12日 |

上海交通大学软件学院

2017 年 1 月

A Dissertation Submitted to Shanghai Jiao Tong University

for Master Degree of Engineering

**THE DESIGN AND IMPLEMENTATION**

**OF CHROMIUM AUDIO AND VIDEO PLAYER**

**BASED ON GSTREAMER**

|  |  |
| --- | --- |
| University Code： | 10248 |
| Author： | Hu Jihao |
| Student ID: | 1130379118 |
| Mentor 1： | Yao Jianguo |
| Mentor 2: | Wang Shenghao |
| Field： | Software Engineering |
| Date of Oral Defense： | 2017.1.12 |

School of Software

Shanghai Jiaotong University

Jan. 2017

**上海交通大学**

**学位论文原创性声明**

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：

日期： 年 月 日

**上海交通大学**

**学位论文版权使用授权书**

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权上海交通大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

**保密**□，在 年解密后适用本授权书。

本学位论文属于

**不保密**🗹。

（请在以上方框内打“**√**”）

学位论文作者签名： 指导教师签名：

日期： 年 月 日 日期： 年 月 日

基于Gstreamer的Chromium音视频播放系统的设计与实现

摘 要

随着汽车的普及和生活节奏的加快，人们在日常生活中越来越多的使用汽车。对于车载多媒体系统，广大消费者已经不再仅仅满足于本地音视频的播放，更希望在车载终端上随时随地享受互联网多媒体服务，例如打开车载终端的浏览器直接联网浏览网页，播放喜欢的音视频，或打开音乐电台在线聆听最新音乐。本课题希望通过对当前网络浏览器的发展规律和未来趋势做简要分析，结合目前车载多媒体的发展现状，对常见构建多媒体应用的开源框架及浏览器音视频播放体系结构的相关技术进行学习研究，完成基于Gstreamer的Chromium浏览器音视频播放系统的设计与实现。

本文的主要工作包括：

（1）通过对当前主流浏览器进行了解，概括其基本特性。系统性地阐述浏览器将HTML、CSS、JavaScript文本文档转变成可视化图像界面的基本流程和工作原理。在此前提下，以渲染引擎WebKit为例介绍其内部设计，并对以Webkit为内核的浏览器Chromium做整体框架分析。

（2）以Chromium浏览器多媒体播放为重点，详细分析相关模块的设计和实现细节，对涉及到的类和流程做详细的说明与解析，了解并掌握其内部实现机制。根据产品需求，对现存的框架进行重构与优化。

（3）了解和分析Gstreamer基本的框架结构，掌握其设计理念和基本原理。简要介绍GStreamer的系统集成方法、调试方法、以及调用机制。基于Gstreamer框架，设计和实现音视频文件的播放，并对实现细节进行封装，以便第三方使用和二次开发。

（4）基于对Gstreamer多媒体播放底层的接口封装和Chromium多媒体部分的重构，重新设计和实现中间层模块，将Chromium与Gstreamer进行对接。

（5）测试和验证设计构想和实现方案，讨论存在的不足之处，为下一步的改进工作指出方向和提出目标。

Gstreamer框架是当前多媒体设计领域较成熟和常用的框架，在此框架下，开发人员可以根据不同的需求灵活地对多媒体格式进行增加，对传输协议的更换也更为便捷。基于Gstreamer的Chro­m­ium音视频播放系统不仅便于后期的扩展，而且被封装成共享库的插件所提供的元件可以被多个程序共享使用，降低开发成本。另外，此系统支持WebAPP的使用，对相关的第三方软件可以无缝对接。基于GStreamer 架构的Chromium音视频播放系统不仅能给用户提供更多、更好的音视频享受，还简化了集成和复用，快速响应市场需求，提高产品竞争力。

关键词 Chromium，Webkit，多媒体框架，Gstreamer

**THE DESIGN AND IMPLEMENTATION**

**OF CHROMIUM AUDIO AND VIDEO PLAYER**

**BASED ON GSTREAMER**

ABSTRACT

With the popularity of cars and the accelerated pace of life, people in their daily lives more and more use of cars. For automotive multimedia systems, the majority of consumers are no longer satisfied with the local audio and video playback. They want to enjoy Internet multimedia services anytime, anywhere in the vehicle terminal, such as openning the car terminal browser directly to browse web, playing favorite music and movie, or listenning to the latest music online. This subject analyzes the present situation and trend of web browser and vehicle multimedia, studies the open source framework and the related technology of the browser audio and video playing architecture, completes the research on the multimedia and multimedia terminal development platform, and lastly realizes Design and Implementation of Chromium Browser Based on Gstreamer for Audio and Video Playing System.

The main work of this paper includes:

1. On the base of having an understading about the current mainstream browsers, summarize its basic features. Systematically describes how the browser converts HTML, CSS, and JavaScript text document into a visual image interface, including the basic flow and working principle. Under these premises, take the rendering engine WebKit as an example to introduce its internal design, and make an overall analysis on Chromium that takes WebKit as the kernel.
2. Provides a detailed analysis of Media in Chromium, including its modules and realization. Give a detailed description and analysis about the related classes and sequences, and understand and master its internal implementation mechanism. Reconstruct and optimize the existing framework, according to the demand of the product.
3. Learn and study of Gstreamer framework and impaortant parts, Master the design concept and basic principles, and briefly introduce its system integration methods, debugging methods, and call mechanism. Design and realize andio and video playing, and package the internel details, in order to provide it to the third party users and to develop secondary.
4. Connect Chromium and Gstreamer through inserting a middle layer, on the base of reconstructing Chromium Media and Gstreamer media playing.
5. Test and vertify design concept and implementation, show the existing problems, and put forward the direction and goal to improve next.

Gstreamer framework is a mature and common framework for multimedia design. It can be used to develop streaming media data processing applications those need to add media formats and transmission protocols. Based on the Gstreamer, Chromium audio and video playback system is easy to be extend later. Multiple programs can share plug-ins packaged as a shared library provided by the components, reducing development costs. In addition, this system supports the use of web application, and the relevant third-party software can be used without obstacles. GStreamer architecture-based Chromium audio and video playback system can not only provide users with more and better audio and video enjoyment, but also simplifies the integration and reuse, rapid response to market demand, and improve product competitiveness.

KeywordsChromium, WebKit, Multimedia framework, Gstreamer

**目 录**

[1 绪论 1](#_Toc472850058)

[1.1 研究背景及意义 1](#_Toc472850059)

[1.2 国内外研究现状 2](#_Toc472850060)

[1.3 本文工作 5](#_Toc472850061)

[1.4 论文组织结构 5](#_Toc472850062)

[2 相关技术 7](#_Toc472850063)

[2.1 浏览器 7](#_Toc472850064)

[2.1.1 浏览器特性 7](#_Toc472850065)

[2.1.2 浏览器基本工作原理 8](#_Toc472850066)

[2.1.3 WebKit内核及Chromium浏览器总体架构 11](#_Toc472850067)

[2.2 Gstreamer基础 14](#_Toc472850068)

[2.2.1 基本概念 15](#_Toc472850069)

[2.2.2 总线（Bus） 17](#_Toc472850070)

[2.2.3 Gstreamer使用方法 18](#_Toc472850071)

[2.3 本章小结 20](#_Toc472850072)

[3 需求分析 21](#_Toc472850073)

[3.1 功能需求 21](#_Toc472850074)

[3.1.1 一般性功能需求 21](#_Toc472850075)

[3.1.2 特殊性功能需求 25](#_Toc472850076)

[3.2 性能需求 27](#_Toc472850077)

[3.2.1 对于网络音视频文件 27](#_Toc472850078)

[3.2.2 对于本地音视频文件 27](#_Toc472850079)

[3.3 可靠性需求 27](#_Toc472850080)

[3.4 资源使用需求 27](#_Toc472850081)

[3.5 设计要求 27](#_Toc472850082)

[3.5.1 设计原则 28](#_Toc472850083)

[3.5.2 编码要求 28](#_Toc472850084)

[3.6 本章小结 28](#_Toc472850085)

[4 设计与实现 29](#_Toc472850086)

[4.1 Chromium中音视频框架分析 29](#_Toc472850087)

[4.2 设计 33](#_Toc472850088)

[4.2.1 概要设计 33](#_Toc472850089)

[4.2.2 Chromium Media Portability Implement分析与设计 34](#_Toc472850090)

[4.2.3 Media Service分析与设计 43](#_Toc472850091)

[4.2.4 AVChannel Manager分析与设计 45](#_Toc472850092)

[4.3实现 46](#_Toc472850093)

[4.3.1 Chromium Media Portability Implement相关的实现 46](#_Toc472850094)

[4.3.2 Media Service相关的实现 48](#_Toc472850095)

[4.3.3 AVChannel Manager相关的实现 50](#_Toc472850096)

[4.4 本章小结 50](#_Toc472850097)

[5 测试与优化 51](#_Toc472850098)

[5.1 功能测试 51](#_Toc472850099)

[5.1.1 测试项目与测试结果 51](#_Toc472850100)

[5.2 性能测试 55](#_Toc472850101)

[5.2.1 测试项目与测试结果 55](#_Toc472850102)

[5.3 优化 57](#_Toc472850103)

[5.3.1 占用资源的优化 57](#_Toc472850104)

[5.3.2 视频描画效率的优化 58](#_Toc472850105)

[5.4 本章小结 58](#_Toc472850106)

[6 总结与展望 59](#_Toc472850107)

[6.1 工作总结 59](#_Toc472850108)

[6.2 问题和展望 60](#_Toc472850109)

[参考文献 62](#_Toc472850110)

[致 谢 64](#_Toc472850111)

[攻读学位期间发表的学术论文目录 65](#_Toc472850112)

1 绪论

本章是本文的绪论部分。在这个章节中，首先对课题的研究背景和意义进行了简要的介绍，随后分析了与课题相关的当前国内外研究现状，接着阐述了本文的主要工作内容，最后对论文的组织结构进行简单说明。

1.1 研究背景及意义

在人们的日常生活中使用互联网的场景和频率越来越高，它已成为人们生活的重要组成部分。伴随着计算机通信技术迅猛发展，以及新的多媒体技术不断涌现，之前静态页面和简单的文字早已无法满足消费者对网络信息的需求，人们往往通过网络来获取包含音频和视频的更加丰富多彩的信息。因此，网络多媒体技术，这种对动态图文信息进行加工处理然后直观地呈献给用户的技术，依靠其集成性良好、交互体验友好以及实时性，结合音频和视频的特性[1]，给人们展示了一个色彩斑斓的网络世界，现在网络多媒体技术已经得到广泛的应用，不管是在工作中还是生活中都给人们带来很大便利。

当下，大部分主流浏览器已经对HTML5技术和多种音视频媒体格式进行了良好的支持，Flash和HTML页面一般会同时被某些音视频网站使用，尽管现在Flash的支持不如以前。Google公司发行的Chrome浏览器可以说是对HTML5最铁杆的支持者[2]，它采用了新的渲染引擎架构，并且引入了时下比较流行的新技术。在Chrome浏览器上打开音视频文件会感受到不同寻常的流畅性和易用性，而且功耗更低。如果我们能将这些优秀的技术应用于嵌入式设备中，那么将会有其独特的优势。同时，对于我们的应用层软件来说，更易于开发跨平台、高性能、兼容性好的产品。

车载多媒体终端作为一款嵌入式移动设备， 被广泛搭在应用消费型轿车上[3]。随着移动互联的大力发展， 消费者已经不再满足于传统本地和外部存储介质（例如U盘）音视频文件播放，等等，更希望在车载移动终端上随时随地使用网络音视频播放服务。

在车载多媒体终端上搭载浏览器就显得很有必要，开源的Chromium浏览器（chr­om­e是Chromium的稳定版，不开源）以其精简的界面、快速的响应、流畅的体验和开源特性获得众多厂商的青睐。

然而，硬件平台和系统平台的差异性和多样性给多媒体系统的开发带来了不小的挑战，开发者往往需要针对不同的平台做定制类的开发。当前多媒体系统的功能也越来越复杂，如果不能妥善解决，对开发成本也造成不同程度的提高。在开发应用程序时候，工作人员需要投入大量的工时去对底层的多媒体处理进行资源整合和恰当使用[4]。

Gstreamer 框架基本的设计思想是基于管道[5]，在这种架构下我们可以方便的添加对新的多媒体格式的支持、以及传输协议的更换等 [6]。我们可以有选择性的把Gstrea­m­er提供的功能性元件安装到管道程序中，即插即用，这也是Gstreamer框架最显著的优点之一。如果使用Gstreamer框架，开发人员可以把多媒体处理相关的处理细节集中管理，并对所有的内部处理进行封装，以接口或者共享库的形式提供给应用程序的开发者，这样上层应用开发者编写多媒体应用程序的时候就不必要再花精力考虑多媒体底层实现的事情，将主要的精力投放到前端的交互体验上面 [7]，以此来节省开发时间，省去一些播放器代码的重复编写。基于GStreamer 架构的播放器不仅能给用户提供更多、更好的音视频享受，而且简化了集成和复用，可以快速响应市场需求，提高产品竞争力。

综上，在嵌入式车载多媒体设备上实现基于GStreamer框架的Chromium音视频播放在当今移动互联时代的工程应用上很有现实意义。

1.2 国内外研究现状

浏览器是一种软件，它可以把文本样式的HTML文档转换成可视化界面，这些文档可以是网页服务器上的，也可以是本地文件系统中存储的，它可以依照一定的规则来解析这些规范性的文档，然后以可视化的图形的形式渲染出来呈献给阅读者，并且可以让用户与这些文件交互和互动。互联网的革命浪潮带动了众多技术的快速发展，网络浏览器作为互联网的最重要的终端接入口之一在短短的二十多年时间里日新月异。经过近二十年的发展，现在的浏览器种类已经比较繁多。在全球市场上占有率比较高的主要有微软的IE、火狐社区的Firefox、谷歌亲生的Chrome、苹果公司的Safari、欧朋的Opera 等[9]。特别是在进入21世纪后，越来越多的功能被加入到浏览器中来。在W3C等标准组织的积极推动下逐步形成HTML5技术，更成为了浏览器发展的火箭推进器。

当前，已经有很多浏览器成功扩展或移植到嵌入式平台的案例。较多的表现在互联网电视领域，例如基于Widget的Yahoo ConnectedTV，增加了 Web 页面扩展功能，它能够用运行在基于Linux系统的电视一体机上，同时在机顶盒中也得到了使用。除了具备了传统电视的播放功能外，还具备网页浏览的互联网功能。比如还有谷歌公司推出的GoogleTV，它是基于Android系统的一款娱乐体验系统，以Android和Chrome为基础，把互联网服务和传统电视服务集合起来，融于一体。GoogleTV 的方案要复杂于Yahoo ConnectedTV，因为GoogleTV除了在系统软件层面做了体系般的定制，还把硬件相关的东西整合进来，并且包括电视周边外围设备，其目的就是为了使生态系统更加完整。

通常，浏览器都有一个核心的东西，就是所谓的渲染引擎，它负责解析网页文档并把解析结果以图像的方式比较直接的表达出来。基于渲染引擎来开发浏览器，是常见的做法。浏览器的渲染引擎有很多种，但目前采用的比较多的只有以下几个[10,11]。

Windows系统原装的浏览器Internet Explorer（IE）使用的渲染引擎是Trident。跨平台的、Mozilla Firefox浏览器的引擎Gecko，Gecko对系统的支持比较全面，它不仅可以运行于 Microsoft Windows、而且可以运行于Linux系统，就连Mac OS X操作系统也支持。Konqueror 是 KDE 桌面系统的一部分，主要用于 Linux 和 BSD家族的操作系统，其使用的引擎是KHTML，速度较快，但是容错能力不是太强，对语法的要求比较严格。开源浏览器引擎WebKit，是目前被采用最广泛的渲染引擎。苹果公司的Safari浏览器和谷歌公司的Chrome浏览器都采用的是WebKit渲染引擎，由于近年来移动互联网的发展以及手机销量的爆炸式增长，这两家公司的浏览器可以说占有了绝对的市场优势，WebKit自然也成了移动设备中市场份额占有量的佼佼者。就连现在Chromium声称的Blink渲染引擎，其实本质上还是WebKit的衍生品。

2008年，Google公司以苹果开源项目WebKit作为内核，创建了一个新的项目Chromium，Chromium是谷歌开源网页浏览器计划的代号[12]。该项目的目标是创建一个快速的、支持众多操作系统的浏览器，包括对桌面系统和移动操作系统的支持。Chromium使用了同Safari一样的浏览器内核。在Chromium的基础上，Google发布了自己的浏览器产品Chrome。不同于WebKit之于Safari浏览器，Chromium本身就是一个浏览器，而不是Chrome浏览器的内核，Chrome浏览器一般会选择Chromium的稳定版本作为它的基础。Chromium是开源试验场，它会尝试很多创新的并且大胆的技术，当这些技术稳定之后，Chrome才会把它们集成进来。也就是说Chrome的版本会落后于Chromium；其次，Chrome还会加入一些私有的编码解码器以支持音视频等；再次，Chrome还会整合Google众多的网络服务[13]。自推出以，以其清爽精炼的用户界面、快速的渲染和脚本执行效率迅速壮大成为最流行的浏览器[14]。

时下对于嵌入式系统来说是一个前景广阔的发展时机，移动互联及物联网的旺盛需求是催生嵌入式市场繁荣的重要因素[15]。随着芯片制造工艺的提升和CPU性能的指数级提升，加上最近几年嵌入式系统的新兴发展，越来越多的功能在嵌入式设备上被是实现，个人电脑的一些功能也在逐渐地向嵌入式设备上移植[16]。有些以前需要在计算机上完成的工作，现在也可以在嵌入式设备中完成，比如办公文件的编写以及图形图像的处理等，都可以很方便的在手机端完成。同时，无线技术的发展也是日新月异，从最初的2G到现在的4G、5G以及随处可见的wifi，无线网络的覆盖和速度正在快速的发展，人们已经不再满足于以往单调的媒体服务。近两年迅速崛起的在线直播、网络视频等新兴业务的增长，对嵌入式多媒体的处理能力提出了新的要求。

各种各样的移动网络业务被中国移动、中国联通、中国电信三大运营商不断推出，流量费用的不断降低对移动网络的大规普及起到了强大的推动作用，移动多媒体的市场也越来越大[17]。就当前来说，多媒体应用几乎存在于每一个智能终端设备，它已经成为了不可或缺的功能之一。要想播放音视频文件，多媒体引擎是必不可少的，不管是本地的文件播放还是网络形式的流媒体文件播放，都需要强大的引擎做支撑。只有多媒体引擎足够强大，才能从根本上满足用户的良好的使用体验。

每种嵌入式移动设备上都有各种不同的播放器或者浏览器软件，不同种类的播放器可能对格式的支持不尽相同。目前有各种各样的多媒体文件格式，每种文件格式的编解码方式又有不同。面对如此多的种类和不同的需求，嵌入式设备的多媒体引擎解码库也越来越复杂[18]。同时，解码库依据不同软件平台和硬件平台又有不同的支持，安装也受到一定的限制，也需要许多的技术支持。所以，开发一个播放器用于嵌入式设备中是比较繁琐和复杂的，当然如果是对每一种多媒体文件去编写新的解码程序就会变得工作量巨大，这也是没有必要的。

针对以上繁琐的问题，Gstreamer提供了完美的解决方案，它的框架简洁，使用简单，思路清晰，开发便捷[19]。工程师们可以使用这种框架结构完成多种多样的音视频处理。Gstreamer的开发资料是比较少的，对开发来说会有一定的难度，但是它的开发机制比较简洁，结构也比较合理，对于多媒体开发从业人员来说是一个良好的工具，相信随着时间的积累，在Gstreamer方向的多媒体开发会变得越来越容易，Gstreamer也会越来越多的应用在各个领域。

Gstreamer多媒体应用程序开发框架多流行于Linux系统下。它的体系结构是基于元件和管道，在Gstreamer框架中几乎可以把所有的功能模块做成可插拔的组件，这些功能模块在需要的时候可以被很容易地安装到目标管道上。利用管道机制可以对几乎所有的插件进行统一的数据管理[20] 。基于以上我们可以看出，使用Gstreamer框架可以快速的使用插件拼出一个具有功能相对完整的多媒体应用程序。

提到多媒体，就不得不提到多媒体文件的编码格式。目前国际上已经形成了比较标准的技术规格和协议，这也为促使嵌入式多媒体的发展做出了推动作用。嵌入式系统的资源是十分有限的，微处理器、存储器等都不如个人计算机那么强大，运算能力和功耗受到制约。而多媒体标准技术的实现对资源的消耗和硬件的要求又是比较高的，这是一个必须要面对的挑战。不过随着MPEG系列H.26X系列等主要国际标准的硬件级实现，多媒体系统的发展被加快。AVS是我国具备自主知识产权的第二代信源编码标准，是《信息技术先进音视频编码》系列标准的简称，它在编码效率方面有一定的提高[21]，这个也在某种程度上解决了软硬件编解码上的效率。

1.3 本文工作

本文首先对当前市场占有率较高的常见浏览器进行了解和总结，概括了其基本特性；随后对具有代表性的浏览器的工作流程做了系统的阐述，并对工作原理做了进一步解析；然后以WebKit为对象对浏览器渲染引擎做了深入的模块解析，在此基础上对Chromium浏览器整体框架分析和了解。

其次，深入分析和研究Gstrea­me­r多媒体开发框架，对一些重要的组成部分，如元件、管道、箱柜、衬垫和总线做详细介绍，同时对Gstreamer的使用方法、运行机制做必要说明。另外对Chromium多媒体部分的设计与实现细节做了详细的了解和分析，针对核心部分和重要环节做出类图和序列图分析。在此基础上对Chromium的Media进行重构与改写，利用当前比较成熟的Gstrea­me­r体系对Chromium的多媒体部分进行新的设计与实现。包括需求分析、总体设计、详细设计、代码实现的部分细节。

最后，编写测试用例对本文所提到的设计方案和实现结果从功能和性能等角度进行了测试与验证，提出尚且存在的不足和下一步的工作方向。

1.4 论文组织结构

论文共分为六个章节，各章节的主要内容安排如下：

第1章，绪论部分。主要介绍了浏览器的发展、各浏览器引擎的基本概况、Chro­m­ium的历史背景、Gstreamer技术的基本思想，并阐述了该文的背景和意义以及相关技术研究成果和未来发展趋势。

第2章，基本理论和相关技术的介绍。介绍了浏览器的特性，并以主流的渲染引擎WebKit和开源的Chromium工程为重点分析了浏览器的工作原理，以及浏览器的架构的设计思想；比较全面的阐述了Gstreamer的基本概念、基本元件、框架理念、使用方法。

第3章，需求分析。对车载多媒体终端要完成的功能、性能，和资源消耗等做了必要的需求分析。

第4章，设计与实现。基于第2章和第3章的系统介绍，对Chromium的源代码分析，进行了新的构建与布局，将Gstreamer应用于Chromium音视频播放系统。从源代码分析、到设计、到实现，完成整个工程的实现。

第5章，测试与优化。对第4章的成果进行测试与验证，并提出在嵌入式系统资源有限的情况下对系统进行优化的方案和实现。

第6章，总结与展望。总结了本文的研究成果与不足，并针对其功能扩展和应用领域提出了进一步的研究方向。

2 相关技术

在本文的绪论部分已经介绍了浏览器的发展历程与当前的状况，随着移动互联网的发展，浏览器作为用户接入互联网的入口将越来越多的受到各大厂商的关注，其竞争也将越来越激烈。本章我们将对浏览器的特性进行概括和总结，提炼浏览器的基本特征；对浏览器的工作原理进行必要的了解和探讨，以便为后文的展开做必要的知识准备；以当前最主流的渲染引擎WebKit为例阐述其总体架构，在此基础上再来分析Chromium的总体架构；本章也对Gstreamer的基本概念、设计理念、运行机制和使用方法做了一定的了解，让我们对Gstreamer的有个初步的认识。

2.1 浏览器

网页浏览器（英语：Web Browser，常被称为浏览器（Browser），之后将简称为浏览器）是一种用于检索并展示万维网信息资源的应用程序。这些信息资源可为网页、图片、影音、或其他内容，他们由统一资源标志符（Uniform Resource Identifier，缩写URI）标志。信息资源中的超链接可使用户方便地浏览相关信息。浏览器虽然主要用于使用万维网，但也可用于获取专用网络中网页服务器之信息或文件系统内之文件[22]。当前，微软IE、Mozilla火狐和Google Chrome成了桌面系统上最流行的三款浏览器，三者一起占据了该市场超过90%的浏览器份额。浏览器作为用户访问互联网最重要的接口，也难怪获得如此众多巨头的关注，未来，必将还是浏览器继续高速发展和竞争激烈的场景[23]。

2.1.1 浏览器特性

大体上来讲，浏览器的这些功能包括网络、资源管理、网页浏览、多网页管理、插件和扩展、书签、历史记录、设置、下载、账户和同步、安全机制、隐私管理、外观主题、开发者工具等。下面是对它们之中一些重要功能的简要介绍。

网络是第一步，浏览器通过网络模块来下载各种各样的资源，例如HTML文本，JavaScript代码、样式表、图片、音视频文件等。网络部分其实非常重要，因为它耗时比较长而且需要安全访问互联网上的资源。从网络上下载或者本地获取的资源，需要资源管理功能将他们管理起来，这需要高效的管理机制，例如如何避免重复下载资源、缓存资源等，都是它们需要解决的问题。网页浏览是浏览器的核心也是最基本、最重要的功能，它通过网络下载资源并从资源管理器获得资源，将它们转变为可视化的结果，这也是后面介绍的浏览器内核最重要的功能。很多浏览器支持多页面浏览，所以需要支持多个网页同时加载，这让浏览器变得更为复杂。同时，如何解决多页面的相互影响和安全等问题也非常重要，为此，一些浏览器做了大量的工作，例如可能使用线程或是进程来绘制网页。插件和扩展是现代浏览器的一个重要特征，他们不仅能显示网页，而且能支持各种形式的插件和扩展。插件是用来显示网页特定内容的，而扩展则是增加浏览器新功能的软件或压缩包。目前常见的插件有NPAPI插件、PPAPI插件、ActiveX插件等，扩展则跟浏览器由密切关系，常见的有Firefox扩展和Chromium扩展。账户和同步将浏览的相关信息，例如历史记录、书签等信息同步到服务器，给用户一个多系统下的统一体验，这对用户非常友好，是浏览器易用性的一个显著标识。安全机制本质是提供一个安全的浏览器环境，避免用户信息被各种非法工具窃取和破坏。这可能包括显示用户访问的网站是否安全、为网站设置安全级别、防止浏览器被恶意代码攻破等。开发者工具对普通用户来说用处不大，但对网页开发者来说意义却非比寻常。一个优秀的开发者工具可以帮助审查HTML元素、调试JavaSc­ript代码、改善网页性能等。

2.1.2 浏览器基本工作原理

以浏览器内核WebKit为例，根据数据的流向，可以将渲染分为三个阶段，第一个阶段是从网页的URL到构建完DOM树，第二个阶段是从DOM树到构建完成WebKit的绘图上下文，第三个阶段是从绘图上下文到生成最终的图像。为了描述这个过程，下面会将WebKit中的一些细节展示出来。图2-1描述的是从网页URL到构建完DOM树这个过程，数字表示的是基本顺序，当然也不是严格一致，因为整个过程可能重复并且可能交叉。

当用户在地址栏输入URL的时候，WebKit调用资源加载器加载该URL对应的网页，对应图中序号1；加载器依赖网络模块建立连接，发送请求并接收答复，对应序号2；WebKit接收到各种网页或者资源的数据，其中某些资源可能是同步或异步获取的，对应序号3；网页被交给HTML解释器转变成一系列的词语（Token），对应序号4；解释器根据词语构建节点（Node），形成DOM树，对应序号5；如果节点是Java­Script代码的话，调用JavaScript引擎解释并运行，对应序号6；解释器根据词语构建节点Node，形成DOM树，对应序号7；如果节点需要依赖其他资源，例如图片、CSS、视频等，调用资源加载器来加载它们，但是它们是异步的，不会阻碍当前DOM树的继续创建；如果是加载JavaScript资源URL（没有标记异步方式），则需要暂停当前DOM树的创建，直到JavaScript的资源加载并被JavaScript引擎执行后才继续DOM树的创建，对应序号8。

在上述的过程中，网页在加载和渲染过程中会发出“DOMContent”事件和DOM的“onload”事件，分别在DOM树构建完成后以及DOM树建完并且网页所依赖的资源都加载完之后发生，因为某些资源的加载并不会阻碍DOM树的构建，所以这两个事件多数时候不是同时发生的。接下来就是WebKit利用CSS和DOM树构建Render­Object树直到绘图上下文，如图2-2所示的过程。

资源加载器

URL

网页内容、JS、CSS等

DOM树

网络模块

HTML解释器

JavaScript引擎

1

2

3

4

5

6

7

8

图2-1 从网页URL到DOM树

Fig 2-1 The process of generating DOM tree

DOM树

CSS解释器

CSS

RenderObject树

RenderLayer树

绘图上下文

布局计算

1

2

3

图2-2 从CSS和DOM到绘图上下文

Fig 2-2 From CSS and DOM to Drawing context

这一阶段的具体过程如下：CSS文件被CSS解释器解释成内部表示结构，对应序号1；CSS解释器工作完之后，在DOM树上附加解释后的样式信息，这就是Render­O­bject树，对应序号2；RenderObject节点在创建的同时，WebKit会根据网页的层次结构创建RenderLayer树，同时构建一个虚拟的绘图上下文，对应序号3。RenderObj­ect树的建立并不表示DOM树会被销毁，事实上，上述图中的四个内部表示结构一直存在，直到网页被销毁，因为它们对于网页的渲染起了非常大的作用。

最后就是根据绘图上下文来生成最终的图像，这一过程主要依赖2D和3D图形库，如图2-3所示。图中这一阶段对应的具体过程如下：绘图上下文是一个与平台无关的抽象类，它将每个绘图操作桥接到不同的具体实现类，也就是绘图具体实现类，对应标号1；绘图实现类也可能有简单的实现，也有可能复杂的实现。在Chromium中，它的实现相当复杂，需要Chromium的合成器来完成复杂的多进程和GPU加速机制，对应标号2；绘图实现类将2D图形库或者3D图形库绘制的结果保存下来，交给浏览器来同浏览器界面一起显示，对应标号3。这一过程实际上可能不像图中描述的那么简单，现代浏览器为了绘图上的高效性和安全性，可能会在这一过程中引入复杂的机制。

绘图上下文

2D图形库

最终的图像

3D图形库

绘图具体实现类

3

1

2

图2-3 从绘图上下文到最终的图像

Fig2-3 From Drawing final image

2.1.3 WebKit内核及Chromium浏览器总体架构

WebKit的一个显著特征就是它支持不同的浏览器，因为不同浏览器的需求不同，所以在WebKit中，一些代码可以共享，但是另外一部分是不同的，这些不同的部分称为WebKit的移植（Ports）。图2-4是一个较为详尽的的WebKit架构图。

图中的WebKit架构，虚线框表示该部分模块在不同浏览器使用的WebKit内核中实现是不一样的，也就是它们不是普遍共享的。用实线框标记的模块表示它们基本上是共享的。之所以没有说的那么绝对，是因为它们中的一些特征性可能并不是共享的，而且可以通过不同的编译配置改变它们的行为。这里面有很多的不同，所以，很多使用WebKit的浏览器可能会表现出不同的行为。

图中最下面是“操作系统”，WebKit可以在不同的操作系统上工作。不同浏览器可能会依赖不同的操作系统，同一个浏览器使用的WebKit也可能依赖不同的操作系统，例如，Chromium浏览器支持Windows、MacOS、Linux、Android等系统。

在“操作系统”之上的就是WebKit赖以工作的第三方库，这些库是WebKit运行的基础。通常来讲，它们包括图形库、网络库、视频库等，加载和渲染网页需要它们不足为奇。WebKit是这些库的使用者，如何高效地使用它们是WebKit和各种浏览器厂商的一个重大课题，主要是如何设计良好的架构来利用它们以获得高性能。现代浏览器的功能越来越强大，性能要求也越来越高，新的技术不断被引入浏览器和Web平台，这也大大增加了WebKit和浏览器的复杂性。

WebKit

WebCore

WebKit嵌入式接口

WebKit2嵌入式接口

WebKit绑定

WebKit2绑定

WebKit Ports

网络栈

音视频

文字

硬件加速

图片解码

. . .

JavaScriptCore

CSS

SVG

布局

渲染树

HTML

DOM

Inspector

. . .

2D图形库

t

3D图形库

t

网络库

t

存储

t

音频库

t

视频库库

t

. . .

t

操作系统

图2-4 WebKit架构

Fig 2-4 The Architecture of WebKit

在它们二者之上的就是WebKit项目了，图中已经把它细分为两层，每层包含很多个模块，由于图的大小限制，略去了其中一些次要模块。

WebCore部分包含了目前被各个浏览器所使用的WebKit共享部分，这些都是加载和渲染的网页的基础部分，他们必不可少，具体包括HTML解释器、CSS解释器、SVG、DOM、渲染树（RenderObject树、RenderLayer树等），以及Inspector（Web Inspector、调试网页）。当然，这些共享部分有些是基础框架，其背后的支持也需要各个平台的不同实现。

JavaScriptCore引擎是WebKit中的默认JavaScript引擎，也就是说一些WebKit的移植使用该引擎。刚开始，它的性能不是很好，但是随着越来越多的优化被加入，现在的性能已变得非常不错。之所以说它还是默认的，是因为它不是唯一并且是可以替换的。事实上，WebKit中对JavaScript引擎的调用是独立于引擎的。在Google的Chromium开源项目中，它被替换为V8引擎。

WebKit Ports指的是WebKit中的非共享部分，对于不同浏览器使用的WebKit来说，移植中的这些模块由于平台差异、依赖的第三方库和需求不同等方面的原因，往往按照自己的方式来设计和实现，这就产生了移植部分，这是导致众多WebKit版本的行为并非一致的重要原因。这其中包括硬件加速架构、网络栈、视频解码、图片解码等。

在WebCore和WebKit Ports之上的层主要是提供嵌入式编程接口，这些嵌入式接口是提供给浏览器调用的（当然也可以有其他使用者）。图中有左右两个部分分别是狭义WebKit的接口和WebKit2（WebKit2 is a new API layer for WebKit to support a split process model, where the web content lives in a separate process from the application UI[24]）的接口。因为接口与具体的移植有关，所以有一个与浏览器相关的绑定层。绑定层上面就是WebKit项目对外暴露的接口层。实际上接口层的定义也是与移植密切相关的，而不是WebKit由是统一接口。

以上，便是对WebKit架构的介绍。

Chromium也是基于WebKit的，而且在WebKit的移植部分中，Chromium也做了很多有趣的事情，所以通过Chromium可以了解如何基于WebKit构建浏览器。另一方面，Chromium也是很多技术的创新者，它将很多先进的理念引入到浏览器领域。Chromium的代码非常复杂，模块非常多，结构也不是特别清晰，非常容易让人迷惑。为了方便理解，我们某块结构上来进行剖析。

首先要熟悉的是Chromium的架构及其包含的模块。图2-5描述了Chromium的架构和主要的模块。从图中可以看到，Blink只是其中的一块，和它并列的还有众多的Chromium模块，包括GPU/CommandBuffer（硬件加速架构）、V8 JavaScript引擎、沙箱模型、CC（Chromium Compositor）、IPC、UI等（还有很多并没有在图中显示出来）。在上面这些模块之上的就是著名的“Content模块”和“Content API（接口）”，它们是Chromium对渲染网页功能的抽象。“Content”的本意是指网页的内容，这里是指用来渲染网页的模块。这里或许有个疑问，WebKit不就是渲染网页内容的吗？是的，没有Content模块，浏览器开发者也可以在WebKit的Chromium移植上渲染网页内容，但是没有办法获得沙箱模型、跨进程的GPU硬件加速机制、众多的HTML5功能，因为这些功能很多是在Content层实现的。

“Content模块”和“Content API”将下面的渲染机制、安全机制、和插件机制等隐藏起来，提供一个接口层。该接口目前被上层模块或者其他项目使用，内部调用者包括Chromium浏览器、Content Shell等，外部包括CEF（Chromium Embedded Framework）、Opera浏览器等。

Chromium浏览器

Content Shell

Android WebView

Content接口

Content模块

Blink(WebKit)

GPU/Command Buffer

V8

沙箱模型

CC(Chromium合成器)

IPC、

UI、

PPAPI、

. . .

图2-5 Chromium模块结构图

Fig 2-5 Modules of Chromium

“Chromium浏览器”和“Content Shell”是构建在Content API之上的两个“浏览器”，Chromium具有浏览器完整的功能，也就是我们编译出来能看到的浏览器式样。而“Content Shell”是使用Content API来包装的一层简单的“壳”，但是它也是一个简单的“浏览器”，用户可以使用Content模块来渲染和显示网页内容。Content Shell的作用很明显，其一可以用来检测Content模块很多功能的正确性，例如渲染、硬件加速等；其二是一个参考，可以被很多外部的项目参考来开发基于“Content API”的浏览器或者各种类型项目。

以上，是我们对浏览器基本技术的简单介绍。当然，浏览器所涉及的技术不仅仅只有我们介绍的这些，它是一个庞大而复杂的项目。考虑到篇幅所限，我们只针对本论文可能涉及的部分进行了简单的描述，以便于后文的展开。

2.2 Gstreamer基础

Gstreamer是一个多媒体框架库，并且它是开源的。我们可以采用Gstreamer方便的开发多媒体模块，比如简单编码格式的音频文件的回放、比如网络流媒体、比如音视频的非线性编辑，等等。解码和过滤技术对于应用程序来说是透明的，上层并不关系具体的实现细节，新的解码器和滤镜可以被开发者使用简单通用的接口编写的插件来添加[25]。Gstreamer对DirectShow的设计思想进行了借鉴，它主要源自视频管道的创意，这些创意是由俄勒冈的一个研究生学院提出的 [26]。插件是Gstreamer的重要特征，种类繁多的编解码器以插件的形式提供，当然也有其他功能。在数据流管道中，我们可以任意链接已经提供好的插件。上面说到Gstreamer只是一个处理多媒体的框架，在这个框架中插件、数据流、媒体操作被核心库函数调度处理。另外，Gstreamer的核心库提供了开放给开发者的API，使用这些API开发者可以方便的使用其他插件，以便于快速的开发应用程序。

2.2.1 基本概念

Gstreamer的基本概念和术语并不是太多，对重要的概念和术语做基本理解是使用Gstreamer的前提。下面我们对此进行简单的说明。

最重要的一个概念就是元件（elements）。要想使数据流顺利的传输，需要创建一系列的元件，并将这些元件按照一定规则连接起来（当然连接元件涉及到例外一个概念，后边会介绍。），然后数据流就可以在这些被连接起来的元件间传输。想象一下一段段的自来水管，为了把洁净的水输送到每家每户，工作人员从供水中心铺设管道，利用大小不一各种各样的接口、水管和其他元件一段衔接一段接入到每个家庭。元件的类型有多种，一个元件它的类型可能是sources、filters、muxers、demux­ers、codecs或者sinks。根据元件类型的不同，它们具有不同的功能，比入source表示的是源元件，它只有输出没有输入，只产生数据不处理数据；而sink代表的接收元件，它只有输入端没有输出端。这些元件都有函数接口提供，有的用于读取文件数据，有的用于解码数据，有的用于对数据分流，比如demuxers就能对视频文件文件中的声音和图像进行分流，将分流出来的数据分别交于声卡和显卡独立处理。我们将这些元件像水管一样对接起来创建一个管道，就能构建一个完整的播放器。很多默认的元件在Gstreamer中已经安装了，如果不是构建特别复杂的多媒体处理软件，这些元件基本够用。当然，在某些情况下，开发人员可以根据项目的需要编写一些特别的新的元件。一般来说，元件像是一个黑盒，数据从黑盒的一端流入，在黑盒内部进行一些加工处理，然后从另外一端流出被处理过的数据。例如：对于一个mux­er元件，流入的是视频文件、音频文件、可能还会有字幕文件，流出的是某一格式的视频文件。如，可将1.avi、1.mp3、1.srt用muxer合并为mkv格式的视频文件。

插件（plugins）本质上是一段可被执行的代码，在程序运行的过程中根据需要来加载。一般来说会把插件的代码编译成动态库或者是共享库，在程序运行的时候被load进去[27]。单个的插件可以只包含一个元件，也可以包含多个元件，我们完全可以把多个元件放入到一个插件之中，使之支持新的功能。在Gstreamer中到处都是插件的概念，哪怕只是开发一个简单的多媒体应用。因为Gstreamer的基本函数是比较少的，它依靠插件实现诸多的功能。插件在使用的时候会有留下注册信息被保存起来，一般来说是一个XML文件，在这个文件中将记载所有注册过的插件的信息。因此，基于Gs­tr­­­e­amer的多媒体应用程序可以在需要的时候进行插件的加载，而没必要在一开始的时候就把所有的插件全部加载进来。

另外一个重要的基本概念就是衬垫（Pad）。衬垫在Gstreamer中起到协商连接的作用，它类似于水管之间的连接头，可以看做是元件之间的连接口，数据在衬垫上流入流出[28]。衬垫具有它自己的特性和独特的数据处理能力，它可以对流经它的数据进行筛选和选择，如果和衬垫所指定的数据类型不符，那么这些数据就无法通过衬垫。比如有解码插件的入口元件上会有一个衬垫来指明自己能够解析的数据类型，当满足于这个数据类型的时候，数据才流向这个插件。也就是说如果一个衬垫的数据想要流向另一个衬垫，那和这两个衬垫所指定的数据类型必须是一致的。

以上是基本概念和术语的介绍，下面我们介绍几种常见的元件：

1. 源元件（source elements）

源元件处于整个管道的最前面，它负责从磁盘或者其他存储介质中将原始的数据读出来，是数据的产生者。与其他元件不同，source element有且只有一个源衬垫，它只产生数据，并不接收数据。也就是说，它是数据的源头，不能对数据做任何处理。

1. 过滤器（filters）与类过滤元件（Filter-like elements）

这两类过滤元件都具有输入端与输出端，过滤器和类过滤元件对从输入衬垫流入的数据进行处理，完毕后经由输出衬垫将处理过的数据留给下一个元件。这类的元件种类是比较多的，比如常见的音量元件，以及ogg分流器，或者视频转换器、还有部分编码格式的解码器，等等。类过滤元件能够拥有一到多个的元衬垫和接收衬垫。比如视频分流器，它或许会拥有一个接收衬垫，但是可能会有多个源衬垫。而解码器一般来说只会拥有一个接收衬垫和一个源衬垫。

1. 接收元件（sink elements）

接收元件是管道的“终结者”，处在管道的最末端，它仅有接收衬垫。因为它是不产生数据的，只用来接收数据。比如使用显卡显示图像或者视频，通过声卡出声，还有写磁盘操作，这些都是由接收元件实现的。

1. 箱柜（bins）和管道（pipelines）

箱柜是可以看做是一个容器，里面可以装载各种元件。管道可以看做箱柜的一个特殊子类型，包含在管道内部的所有元件都能被管道操作。使用箱柜可以降低程序开发的复杂度，我们可以像操作元件一样操作箱柜，因为箱柜本来就是元件的子集。如果我们改变了箱柜的状态，那么箱柜内的所有元件的状态也将被改变。箱柜可以给它的子集元件发送总线消息（bus messages）。管道是一个高级箱柜，如果一个管道处于停止状态，给它设成播放状态，那么数据就开始在整个管道中流动，此时多媒体数据的处理也同时开始了。一旦开始，Gstreamer会使用一个独立线程承载和维护管道的运行，当数据处理完毕或者状态被设置成停止的时候，运行才会结束。

2.2.2 总线（Bus）

我们可以把总线看成一个简单的系统，它自己有一套独特的线程机制，使用管道的这套线程机制可以把消息方便的分发到应用程序中。当也正是因为管道拥有这一套线程处理的机制，在我们基于Streamer进行多媒体应用程序开发的时候，就不用太过关注线程的区分和识别问题，不管Gstreamer是加载了一个线程还是多个线程，我们只需要专注于监听总线消息就好了。

管道被创建的时候，有一个总线被默认包含。所以对于开发者来讲，就不用再创建总线，给总线设计一个消息处理器就好，当主循环run起来的时候，总线会不断的查看消息处理器，看看有没有消息需要处理，如果总线发现有新的消息发来并且需要处理，那么总线便立即调用相关的回调函数进行处理。

有两种使用总线的方法[29]。开启GLib/Gtk+的main loop，然后设置监听器对总线进行消息侦听，这是第一种方法。如果使用这种方法，Glib的main loop将轮询总线上是否有新消息发来，如果侦听到新的message，总线会立即发出通知。关于这种case，有两个函数必不可少，就是对监测进行添加的函数gst\_bus\_add\_watch()以及gst\_bu­s\_ad­d\_signal\_watch()两个函数。如果pipeline发出一个message到bus，就会触发这个message处理器，message处理器就着手分辨消息信号种类以此判断哪些将处理哪些事件。如果不采用前边讲到的第一种方法，那么还有一种方法可以使用，那就是自己对总线消息进行监听，这会用到两个函数gst\_bus\_peek()和gst\_bus\_poll()，此处不再展开。

对于总线传递的消息类型，在Gstreamer中有几种常见的是事先被预定义的。当然，开发者也可以扩展这些消息类型，根据具体的工作需求在插件中定义新的消息类型[30]。这些消息中都会携带info以标识每个信息，比如消息来哪里、消息的类型、以及消息发出的时间戳。如果我们想知道某个消息的消息源在哪里，通过读取info就可以识别出此消息出自哪个element。

2.2.3 Gstreamer使用方法

下面，我们以一个真实的案例来说明如何基于Gstreamer框架的组件创建一个功能不那么复杂的MP3播放器。在这里我们假设MP3音乐文件存储在本地磁盘上，source element从磁盘中将该MP3音频数据读取出来，然后使用filter element对数据进行decode，最后经sink element将解码后的数据写入声卡。

首先我们在main函数中调用gst\_init()来进行初始化工作，以便Gstreamer函数库能够都到用户传递进来的参数。一个经典的基于Gstreamer的多媒体处理程序的初始化过程可参看如下代码[31]：

#include < gst/gst.h >

int main ( int argc, char \*argv[] )

{

gst\_init( &argc, &argv );

}

然后我们创建三个元件并把他们连接成一个管道。GstElement是Gstreamer的所有元件的基类，因此我们在声明各类型元件的时候都可以使用基类的数据类型：

GstElement \*pipeline, \*filesrc, \*decoder, \*audiosink;

前面说过pipeline在Gstreamer框架是特殊的箱柜，可以用来容纳和管理元件，现在我们创建管道，用变量pipeline来引用我们的管道：

pipeline = gst\_pipeline\_new(“my-new-pipeline”);

source element是数据的产生者，他有一个属性叫location，可以把准备好的MP3音频文件所在的路径设置给这个属性。如下filesrc是创建的数据源元件，argv[1]表示用户输入的MP3文件路径：

filesrc = gst\_element\_factory\_make( “filesrc”, “disk\_source” );

g\_object\_set( G\_OBJECT(filesrc), “location”, argv[1], NULL );

创建filter element：

decoder = gst\_element\_factory\_make ("mad", "decoder");

创建sink element：

audiosink = gst\_element\_factory\_make(“audiosink”, “play\_audio”);

三个元件已经创建完毕，它们依次是源元件filesrc，过滤器元件decoder，接收器元件audiosink，下面使用函数gst\_bin\_add\_many把这三个元件添加到管道中，并使用gst\_element\_link\_many把他们按顺序连接起来：

gst\_bin\_add\_many(GST\_BIN(pipeline), filesrc, decoder, audiosink, NULL);

gst\_element\_link\_many(filesrc, decoder, audiosink, NULL);

以上，便完成了基本的准备工作，下面就是直接通过切换管道状态的方式，开始整个流程，我们把状态切换成PLAYING：

gst\_element\_set\_state(pipeline, GST\_STATE\_PLAYING);

若是要监听总线消息，可调用函数gst\_pipeline\_get\_bus(GST\_PIPELINE(pipeline))获取到bus，然后通过gst\_bus\_add\_watch添加总线消息的监听，回调函数是callback：

gst\_bus\_add\_watch(bus, callback, loop)

callback可以定义如下：

static gboolean bus\_call(GstBus \*bus, GstMessage \*msg, gpointer data)

{

// TO DO

}

播放结束的时候，我们将状态设为NULL终止管道，然后接触对pipeline资源的引用：

gst\_element\_set\_state(pipeline, GST\_STATE\_NULL);

gst\_object\_unref(GST\_OBJECT(pipeline));

至此，利用Gstreamer创建一个简单的MP3播放器示例介绍完毕，其过程并不复杂，其目的是帮助我们理解Gstreamer基本概念和熟悉它的使用方法，为后文的展开做铺垫。

2.3 本章小结

本章首先对浏览器的历史沿革做了简单的介绍，以便了解开发浏览器的初衷与目的；其次以现代浏览器为主，描述浏览器具备的特性；在对浏览器充分了解的基础上，随后展开对浏览器工作原理的进行阐述；然后以WebKit和Chromium为重点，详细分析了浏览器的核心架构与设计。此外，还对Gstreamer由浅入深，由简单到复杂地展现了其基本概念、使用方法、工作机制、等等，以便对Gstreamer有全面的认识。以上，为下文的展开，做了系统性的铺垫。

3 需求分析

上一章我们对浏览器的特性、工作原理、WebKit和Chromium的总体框架，及Gstreamer的基础知识进行了介绍。结合车载多媒体的特性及常见的交互逻辑，本章对将要设计和实现的系统进行必要的需求分析。我们所设计的基于Gstreamer的Chromi­u­m音视频播放系统是搭载在车载多媒体终端的，车载多媒体终端是一个嵌入式设备，其硬件性能及内存资源有限，在分析的时候要考虑到内存的占用和性能的优化工作，及时关闭不必要的进程或线程和动态释放不使用的内存资源。

3.1 功能需求

3.1.1 一般性功能需求

1 管理控制

基于Gstreamer的Chromium音视频播放系统应具备基本的播放管理控制功能，即常见的播放、暂停、停止、跳转、监测、调节音量、错误处理等，相关用例图如图3-1所示。

在图3-1中，最左侧表示Chromium浏览器使用者，包括自然人和其他被Chrom­ium浏览器支持的应用等；中间的表示Chromium的音视频播放子系统，它对右边的Media Service和AV Channel Manager存在一定的依赖关系。关于图3-1中涉及到的用例的简要分析如下。

播放（play）：浏览器用户打开带有音频或视频的页面，按下播放按钮或者网页音视频被设置为自动播放时，浏览器加载音视频文件进行播放。其基本事件流如下：首先，用户打开带有<audio>或<video>标签的网页，网页加载完毕；然后，用户点击播放按钮；当用户点击播放按钮后，页面开始加载音视频文件；当数据加载满足播放条件，会看到时间进度发生改变，播放开始。当然，若果页面内的音视频被设置成自动播放，当页面打开的时候，无需点击播放按钮，音视频文件会进行自动播放，这和网站的交互设计相关。另外，在网速较慢的情况下，如果不能立即播放，应该在加载的时候显示加载进度，以便用户了解当前的情况；如果遇到网络连接不稳定，应反馈网络连接不稳定的消息给用户。这里的用户不仅仅是指Chromium浏览器的使用者，也指点三方的WebAPP调用。总之，要将音视频播放系统播放过程中的状态信息或是发生的异常及时地通知给使用者或调用者。



图3-1 管理控制用例图

Fig3-1 Management & control use case diagram

暂停（pause）：当此播放系统系统已经进入到正常播放状态后，播放进度开始更新，用户可以听到音频或是看到视频在播放，这时候点击网页上的暂停按钮，此系统会进入暂停状态中。另外一种情况是，当用户点击播放按钮，由于网速较慢，音视频文件并没有立即播放，而是一直在加载数据，这时候如果点击暂停按钮，系统的状态切换为暂停状态，但是其加载过程不并暂停，默默的在后台运行。注意在这种情况下，等待了一段时间，数据加载完成，并不会进入到播放状态，而是进入到暂停状态。只有用户点击了播放按钮，才会进入到播放；否则，则停留在暂停状态，等待用户操作。

停止（stop）：当播放系统已经进入到播放或暂停状态中，点击停止按钮，此系统会释放掉缓存的资源，并进入到停止状态。一般来说，对于常见的HTML页面，很少会有停止按钮。现在常见的包含有andio或video标签的页面，一般只包含播放和暂停按钮。但是对于某些传统的播放器来说，可能会有停止按钮。由于我们设计的这款基于Gstreamer的Chromium音视频播放系统除了要支持Chromium浏览器本身的播放需求外，还需要支持其他播放器的使用，所以“停止”这个功能还是不可或缺的。这个“停止”功能和“暂停”功能的区别是：如果单击了停止，那么时间轴将会立即回到播放最初的0的位置，而暂停则停在正在播放的时刻；另一个重要的区别是，暂停不会释放已经缓存的资源，但是停止会在某个时间点后释放已经缓存的资源，以减少对内存空间的占用。

静音（mute）：当音视频播放开始后，我们会听到有声音流出，这时候如果点击静音按钮，音量会立即调整为0；如果已经在静音状态下，这时候再点击静音按钮，音量将恢复到之前的大小。这里需要注意，如果是在暂停或停止状态下点击静音按钮，其音量的变化应该同播放状态下一致。即，只要audio或video标签能够被正常的解析出来，静音功能必须是可用的，与系统所处的播放状态无关。

调节音量（change volume）：此功能和静音功能有重叠的部分，其本质都是对音量的控制。当音视频在播放的过程中，拖动音量调节条，可听到音量高低的变化，音量的数值越大表示音量越高，数值越小表示音量越低，其大小范围是从0到100。注意，当调节音量到最小值0的时候，其效果等同于静音，此时，被静音的状态应该显示出来。如果是在暂停或停止状态下，滑动音量调节条，这时候需要给出系统默认的短促的提示音来表示音量的大小。即，如果是在音视频播放的过程中，通过音视频文件的声音反馈音量的大小；如果当前音视频不在播放状态下，应使用系统默认的提示音来表达音量的大小。

跳转（seek）：这里的跳转指的是选时播放，即当系统在播放或暂停状态下，点击或拖动进度条上的滑块到某个时间点，那么播放系统就从该处开始播放。如果遇到网络情况不好的时候，则显示正在加载的提示，提示用户正在进行数据缓冲。这里的跳转包括向前和向后，如果用户拖动滑动条到0的位置则从头开始播放，如果用户拖动进度条到最后一秒，首开手后则进入到播放结束状态。

获取播放进度（get current position）：当系统进入到播放状态后，可以获取到当前的播放进度，即当前播放到几分几秒。对于播放进度的通知，一般来说要求是每秒通知三次。如果用户需要主动获取当前的播放时间，则可通过该功能获取。如果是在暂停状态下，则每次获取到的进度都同样为暂停时刻的时间进度。

获取总时长（get duration）：指获取到文件信息后，首先解析出文件的总时长，并显示在播放界面上。并不是所有的文件都能获取到duration，有些时候在流媒体中可能会缺失总时长的信息，这种情况下系统反馈的总时长信息将会是-1，表示没有找到总时长信息，这时播放进度的变化也就不能按照总时长信息进行显示，可根据Chromium的默认效果处理，如果是定制级的UI则可根据需求自行决定显示策略。

获取播放状态（get state）：这里的播放状态主要用于描述播放系统各种状态的变化，包括：正在加载状态、播放中状态、暂停状态、停止状态、出现错误状态。从一个页面加载出来到音视频文播放结束，其经过的几个状态要能全部获取，并根据需要在网页界面中反馈出来。

获取文件信息（get info）：这个功能主要时用来获取音视频文件的相关信息，比如文件的标题、歌手姓名、专辑名称、年代、比特率等等。对于网络音视频文件，可能这些信息不一定能从流媒体总获取到，因为有些是在页面的其他位置显示的，这些信息没有整合到文件流里；但是对于本地音视频文件来说，一般都有对这些信息的描述，比如某个mp3里的歌手姓名，音乐类型等等。

声音通道抢占通知（AVChannel occupied）：在车载多媒体终端，不可避免的会遇到许多紧急的情况，比如严重的交通信息、车身故障的警报、或者电话的接入，等等。这些事件的优先级都高于普通的音视频文件播放。当这些高优先级的需要占用声音通道的事件发来的时候，需要给用户发通知，暂停当前音视频文件的播放，优先切到级别高的事件中去。

错误处理（error occur）：当发生错误的时候，要能明确的侦测到具体是什么样的错误，并将错误信息反馈给用户。错误的类型有很多，主要的包括：网络连接失败、资源获取失败、格式错误、解码错误、等待超时。

2 格式支持

原生的Chromium浏览器对音视频格式的支持是非常有限的，而车载多媒体系统中存在的文件格式又是多种多样的，所以，必须对格式的支持进行扩充。在音频格式中，MP3因其文件尺寸比较小、音质又好，该格式直到现在作为主流音频格式的地位仍难以撼动；在视频格式中，MP4因其体积小、画质好，非常流行于移动终端设备，而我们的车载多媒体终端显示频像素大小是800\*480，非常适合配套MP4。音视频播放系统的数据处理是系统能否正常工作的前提，本系统对音视频文件格式及编码格式的支持需求如下表（表3-1）所示。

表3-1 支持文件格式与编码格式

Table 3-1 The supported file format and encoding format

|  |  |  |
| --- | --- | --- |
| 类型 | 支持文件格式 | 支持编码方式 |
| 音频 | MP3，AAC，OGG | AAC，MP3，Vorbis |
| 视频 | MP4，AVI，MKV，WEBM，3GP，TS，OGG | H.264，MPEG-4 |

3.1.2 特殊性功能需求

1 Web App的支持

系统除了要支持Chromium浏览器的本地及网络音视频的播放外，还需支持Web APP相关音视频的播放。

当车载多媒体终端的WebAPP启动的时候，会发相应的通知事件给系统，此时系统需要调起Chromium浏览器对WebAPP进行描画和响应。从本质上来讲，WebAPP也是普通的网页文件，只不过其表现形式为一个简单的APP。打开WebAPP和在浏览器地址栏中输入相应的网址打开对应网站没有本质的区别，只是表现形式的不同。所以，如果基于Gstreamer的Chromium浏览器音视频播放系统对前边提到的功能有良好的支持，那么对WebAPP相关的支持也不会有什么问题。

2 传统播放器的支持

系统除了要支持Chromium浏览器的本地及网络音视频的播放及Web APP相关音视频的播放，还需要对本地传统播放器支持，比如本地独立播放器只用于播放插入的USB或者SD卡内的音视频资源文件。其功能包括：获取文件列表、上一首或下一首跳转、快进或快退、倍速播放，其用例图如图3-2所示。



图3-2 对传统播放器的支持用例图

Fig3-2 Supporting traditional player use case diagram

其主要功能简介如下。

获取文件列表（get media file list）：主要用于在给定路径的前提下，扫描该路径下存在的音视频文件资源，并把结果返回给使用者。像多数的嵌入式系统一样，车载多媒体终端也能浏览本地存储的文件，当用户浏览文件的时候，把路径设置给音视频文件播放系统，此系统将扫描该路径下存在的音视频文件。其判断音视频文件的方法是读取文件的后缀名，如果是该系统支持的文件格式则显示出来，如果不是该音视频播放系统支持的文件格式则无视。对于扫描到的文件列表中的音视频文件不一定能够全部解码，因为此处是根据文件的后缀名来判断的。在播放的时候，有可能出现解码失败，如果出现这种情况，要及时同时给使用者。

上/下一首（Pre/Next）：如果已经获取到音视频文件列表，在展示的文件列表下，如果选择了Pre会Next，则系统将自动根据扫描的结果播放上一个或者下一个音乐或视频。

快进/快退（FF/REW）：如果是在正常的播放情况下，使用了快进快退功能，则音视频文件执行快进快退操作。如果是在暂停或停止状态下，该功能不能使用。

倍速播放（Multi speed playback）：在正常播放状态下，点击倍速按钮，则进入倍速播放状态。用户点击一次则2倍速，第二次点击则4倍速，3第三次点击则8倍速，第四次点击则16倍速，第五次点击则变为原速播放，以此循环。

3.2 性能需求

3.2.1 对于网络音视频文件

（1）当发起网络加载时，加载响应时间不超过1秒。

（2）当网络信号较弱或网速较慢时，依据情况进行缓冲，并每隔10毫秒更新一次缓冲进度。

（3）若发生网络中断或遇到解码失败的情况，立即通知Error。

（4）若要求暂停时，软件响应时间最长不超过1秒。

（5）要求停止后，10秒后释放缓存所占资源空间。

3.2.2 对于本地音视频文件

（1）从发出播放请求到音视频文件开始播放，其时间间隔不超过2秒。

（2）从发出暂停请求到音视频文件播放状态切换为暂停，其时间间隔不超过1秒。

（3）播放过程中，要求解码迅速、播放不卡顿、音质良好、画面流畅。

3.3 可靠性需求

（1）连续运行一周无故障概率99.5%。

（2）如果发生Crash，能够立即复归。

3.4 资源使用需求

运行期间，占用最大内存不高于60M。

3.5 设计要求

3.5.1 设计原则

在将软件需求转换成设计模型的过程中，不仅要考虑如何实现功能需求，而且要考虑和关注软件的可扩充性、可维护性、可重复性和通用性，采用抽象、分解和模块化、封装和信息隐藏、高内聚和低耦合的原则进行设计。

3.5.2 编码要求

依照规范的开发流程，编码前要首先完成基本设计，程序的结构设计要合理，对模块进行有效的封装，增强内聚性，耦合程度要低。代码的可读性强，接口定义规范。代码要易于维护，易于扩展、可重用性强、移植性好；注意代码的执行效率和运行时内存空间的占用。

3.6 本章小结

本章对需求做了言简意赅的精要分析，提出了在功能、性能、可靠性、资源占用等方面需满足的标准和交付要求，对第4章的设计与实现提出了具体的约束，具有指导性意义。

4 设计与实现

在前一章需求分析的基础上，本章将针对基于Gstreamer的Chromium音视频播放系统进行总体和详细的分析、设计及代码实现。首先对Chromium中音视频播放相关的代码进行了分析，找出音视频接口移植的关键部分；接着根据需求总体上将系统划分为三个模块，分别针对各模块进行设计；最后，在实现的小节中给出关键的编码。

4.1 Chromium中音视频框架分析

WebKit提供了支持多媒体规范的基础框架，如音视频元素、JavaScript接口和视频播放等[32]。图4-1所示是WebKit支持视频的基础类和关系。



图4-1 WebKit支持视频的基础类和关系

Fig 4-1 The base classes supported by WebKit and their relationship

图中左侧HTMLMediaElement和HTMLVideoElement是DOM树中的节点类，分别对应W3C标准中的定义，包含众多接口，这些接口可以被JavaScript代码访问。

其次是MediaPlayer和MediaPlayerClient两个类，他们的作用非常明显。MediaPlayer是一个公共标准类，被HTMLMediaElement类使用来播放音频和视频，它本身只是提供抽象接口，具体实现依赖于不同的WebKit移植。同时，一些播放器的状态信息需要通知到HTMLMediaElement类，这里使用MediaPlayerClient类来定义这些有关状态信息的接口，HTMLMediaElement类需要继承MediaPlayerClient类并接收这些状态消息。根据前面的描述，规范要求将事件派发到JavaScript代码中，而这一实现在HTMLMediaEle­ment类完成。

然后是是不同移植对MediaPlayer类的实现，其中包括MediaPlayerPrivateInterface类和WebMediaPlayerClientImpl类。前者是除了Chromium移植之外使用的标准接口，也是一个抽象接口，由不同的移植来实现。后者是Chromium移植的实现类。为什么会这样？因为Chromium将WebKit复制出Blink最后就将MediaPlayerPrivateInterface类直接移除了，而在MediaPlayer类中直接调用它。WebMediaPlayerClientImpl类会使用Chromium移植自己定义的WebMediaPlayer接口来作为实际的播放器，而真正的播放器则是在Chromium项目的代码中来实现[33]。

最后是同渲染有关的部分，这里面包含RenderObject树和RenderLayer树，图中RenderMedia类和RenderVideo是RenderObject的子类，用于表示Media节点和Video节点。图4-1是采用硬件加速机制的视频播放所使用的类，WebKit也支持使用软件渲染方式来播放视频，当需要绘制的时候，由RenderVideo类使用HTMLMediaElement类获取MediaPlayer对象，调用它的paint方法来让MediaPlayer对象将解码后的图像绘制在2D图像上下文接口中。

因为视频资源相对其他资源而言，一般比较大，当用户播放视频的时候，需要连续性播放以获得较好的体验，但是网络可能不是一直都稳定和高速，所以资源的获取对用户体验很重要，需要使用缓存机制或者其他机制来获取视频资源[34]。

图4-2是Chromium中缓存资源类。BufferDataSource类表示资源数据，它是一个简单的数据表示类，内部包含一个较小的内存空间，实际的缓冲机制由BufferResourceLoader类来完成。



图4-2 Chromium的视频资源缓存类

Fig 4-2 The cache classes of video supported by Chromium

根据多进程架构的设计原则，Chromium的媒体播放器的实现应该在Renderer进程，而对于资源的获取，应该在Browser进程[35]。Chromium支持媒体播放器的具体实现相当复杂，而且涉及到不同的操作系统，目前Chromium在不同的操作系统上实现的媒体播放器也不一样。图4-3是Chromium基础类，为了方便理解这些类和图4-1中类之间的关系，图4-3标注了一些WebKit中同Chromium直接相关的类。



图4-3 Chromium支持视频的基础类和关系

Fig 4-3 The base classes supported by Chromium and their relationship

图4-3上半部分是WebKit和WebKit的Chromium移植中的相关类，下半部分是Chromium中使用硬件加速机制来实现视频播放的基础设施类。而从左往右看，左边部分是播放器的具体实现类，右边部分则是支持视频在合成工作中的相关类。

首先来看一看这些类和对象的创建过程。WebMediaPlayerClientImpl类是WebKit在创建HTMLMediaElement对象之后创建MediaPlayer对象的时候，由MediaPlayer对象来创建的。当视频资源开始加载的时候，WebKit创建一个WebMediaPlayer对象，当然就是Chromium中的具体实现类WebMediaPlayerImpl对象，同时WebMediaPlayerClientImpl类也实现了WebMediaPlayerClient类，所以WebMediaPlayerImpl在播放视频的过程中需要向该WebMediaPlayerClient类更新各种状态，这些状态信息最终会传递到HTMLMediaElement类中，最终可能成为JavaScript事件[36]。之后，WebMediaPlayerImpl对象会创建一个WebLayerImpl对象，还会同时创建VideoLayer对象，根据合成器的设计，Chromium还有一个LayerImpl树，在同步的时候VideoLayer对象对应的VideoLayerImpl对象会被创建。之后Chromium需要创建VideoFrameProviderClientImpl对象，该对象将合成器的Video层同视频播放器联系起来并将合成器绘制一帧的请求转给提供视频内容的VideoFrameProvider类，这实际上是调用Chromium的媒体播放器WebMediaPlayerImpl，因为它就是一个VideoFrameProvider类的实现子类。

然后是Chromium如何使用这些类来生成和显示每一帧的。当合成器调用每一层来绘制下一帧的时候，VideoFrameProviderClientImpl::AcquireLockAndCurrentFrame()函数会被调用，然后该函数调用WebMediaPlayerImpl类的GetCurrentFrame函数返回当前一帧的数据。VideoLayerImpl类根据需要会将这一帧数据上传到GPU的纹理对象中。当绘制完这一帧之后，VideoLayerImpl调用VideoFrameProviderClientImpl::PutCurrentFrame来通知播放器这一帧已绘制完成，并释放掉相应的资源。同时，媒体播放器也可以通知播放器这一帧已绘制完成，并释放掉相应的资源。同时，媒体播放器也可以通知合成器有一些新帧生成，需要绘制出来，它首先调用播放器的VideoFrameProvider::DidReceiveFrame()函数，该函数用来检查当前有没有VideoLayerImpl对象，如果有对象存在，需要设置它的SetNeedsRedraw标记位，这样，合成器就知道需要重新生成新的一帧。

最后是上述有关视频播放对象的销毁过程。有多种情况使Chromium需要销毁媒体播放器和相关资源，如video对象被移除或者设置为隐藏等，这样视频元素对应的各种层对象，以及WebKit和Chromium中的这些设施就会被销毁。

在桌面系统中，Chromium使用了一套多媒体播放框架，而不是直接使用系统或者第三方库的解决方案[33]。图4-4是Chromium在桌面系统中采用的多媒体播放引擎的工作模块和过程，这一框架称为多媒体管线化引擎，图中主要的模块是多路分配器(Demuxer)、音视频解码器、音视频渲染器。这些部分主要是被WebMediaPlayerImpl类调用。



图4-4 Chromium的多媒体管线化引擎

Fig 4-4 The multimedia pipeline engine of Chromium

在处理音视频的管线化过程中，需要解码器和渲染器来分别处理视频和音频数据。它们均采用一种叫做“拉”而不是“推”的方式进行，也就是说由视频或者音频渲染器根据声卡或者时钟控制器，按需要来请求解码器解码数据，然后解码器和渲染器又向前请求“拉”数据，直到请求从视频资源文件读入数据。根据之前的多进程架构和Chromium的安全机制，整个管线化引擎虽然在Renderer进程中，但是由于Renderer进程不能访问声卡，所以图中渲染器需要通过IPC将数据或者消息同Browser进程通信，由Browser进程来访问声卡。

4.2 设计

4.2.1 概要设计

根据需求分析，在整体上可以分为三个部分：

Chromium Media Portability Implement

Media Service

AVChannel Manager

硬件

图4-5 整体模块划分

Fig 4-5 The overall modules

Chromium Media Portability Implement：它是与Chromium关联最强的部分，通过他联通了Chr­­­­­om­iu­­­m和Media Service的桥梁。当Chromium需要创建一个播放器来播放音视频文件的时候，此处的播放器被创建，与Media Service通信。

Media Service是随系统启动的一个服务进程，等待使用者的连接和播放控制，是真正的音视频播放的执行者。

AVChannel Manager是多媒体通道资源管理者及调配者，比如管理着声音类型的优先级，各种中断，等等。

4.2.2 Chromium Media Portability Implement分析与设计

当Chromium渲染引擎遇到<video> tag时，会创建对应的Element对象HTML­VideoElement；同样地，当遇到<audio> tag的时候，会创建HTMLAudioElement对象。HTMLVideoElement和HTMLAudioElement都是从HTMLMediaElement派生，他们之间的类关系如图4-6所示。



图4-6 Chromium中媒体类派生关系

Fig 4-6 Media derived relationships in Chromium

从上图（图4-6）可以看出，在HTMLMediaElement是HTMLAudioElement和HTMLVideoElement的父类，中拥有一个Media播放器，当音视频需要播放的时候，会通过MediaPlayer进行。从上图中我们可以看到，WebMediaPlayerClient­Impl是Media­Player的子类，负责Media播放的实现。而其实，WebMediaPlayerClientImpl也不是Media播放的真正实现者，WebMediaPlayerClientImpl又把播放的任务委托给了WebMedia­Player，如下图（图4-7）所示，WebMediaPlayer其实也只是一个接口类，根据不同的移植有不同的实现，图4-7中WebMediaPlayer右侧的WebMediaPlaye­rClient是播放状态的监听者，由图4-6中的WebMediaPlayerClientImpl继承实现，从而把播放状态层层向上反馈，最终在浏览器的页面上体现出来，比如播放进度的更新，等等。

对于图4-7中的各个类做如下说明：

WebMediaPlayer：接口类，被WebMediaPlayerSmartAuto实现，提供给Chromium使用。

WebMediaPlayerClient：接口类，由Chromium实现，主要用于监测player的状态及通知事件，并做出相应动作。

图4-7 Chromium Media的实现

Fig 4-7 The implementation of Chromium Media



WebMediaPlayerSmartAuto：WebMediaPlayer的实现者，实现WebMediaPlayer的所有接口。

RenderFrameImpl：是Chromium源代码中的类，我们在这个类里创建了Web­M­ediaPlayerSmartAuto。

RendererMediaPlayerManager：对WebMediaPlayerSmartAuto进行管理的管类

IPC：进程间通信相关，Chromium源生代码。

MediaWebContentsObserver：接收和过滤Render进程发出的Media消息。

BrowserMediaPlayerManager：对MediaPlayerSmartAuto进行管理。

MediaPlayerSmartAuto：Browser进程中的Media播放器。

WebMedia：接口类，被WebMediaImpl实现。

WebMediaImpl：实现WebMedia，中间件，负责Chromium和本地libMedia（NMCl­assicMediaPlayer）的衔接。

NMClassicMediaPlayer：可以看做Media Service提供的proxy类，给WebMediaImp使用。

PlayerListener：监听NMClassicMediaPlayer的状态变化及事件通知。

WebMediaCb：接口类，被WebMediaCbImpl实现。

WebMediaCbImpl：实现WebMediaCb，向MediaPlayerSmartAuto通知各种事件及状态变化。

AVChannelClient：资源管理的监听者，当监听到声音通道被抢占、视频交替、后者高优先级Source切换的变化后通知给Player。

NCSubChannelManager：对资源管理的子集类

AVChannelManager：管理各Channel，声音及视频通道的打开、关闭、切换、状态维护等。

在上图中，WebMediaSmartAuto是针对Chromium的WebMediaPlayer的实现者。在RenderFrameImpl的函数createMediaPlayer()中，完成了对播放引擎（在Chromium中有一套完善的播放引擎创建的规则，这里的引擎可以理解为播放器）的创建。关于WebMediaSmartAuto的I/F接口一览，我们可以参看以下表4-1。

表4-1 WebMediaSmartAuto I/F一览

Table4-1 WebMediaSmartAuto interface list

|  |  |  |  |
| --- | --- | --- | --- |
| **FunctionName** | **Param** | **Return** | **备注** |
| load | WebURL | void | 加载文件 |
| play | void | void | 播放 |
| pause | void | void | 暂停 |
| paint | Rect | void | 视频描画 |
| paused | void | bool | 判断是否暂停 |
| currentTime | void | double | 获取当前播放进度 |
| duration | void | double | 获取音视频文件总时长 |
| networkState | void | NetworkState | 获取网络状态 |
| readyState | void | ReadyState | 获取当前准备状态 |
| OnTimeUpdate | double | void | 进度更新通知 |
| OnDurationChanged | double | void | 总时长更新通知 |
| UpdateReadyState | ReadyState | void | 准备状态变更通知 |
| UpdateNetworkState | NetworkState | void | 网络状态变化通知 |
| playbackStateChanged | void | void | 播放状态变更通知 |

以上各函数的详细说明如下所示

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | load | |
| 処理概要 | | 加载文件 | |
| 参数 | 1 | WebURL | 网络音视频资源文件的路径 |
| 返値 | | void |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | play | |
| 処理概要 | | 播放 | |
| 参数 | 0 | void |  |
| 返値 | | void |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | pause | |
| 処理概要 | | 暂停 | |
| 参数 | 0 | void |  |
| 返値 | | void |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | paint | |
| 処理概要 | | 描画视频 | |
| 参数 | 1 | Rect | 描画区域 |
| 返値 | | void |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | paused | |
| 処理概要 | | 判断是否暂停 | |
| 参数 | 0 | void |  |
| 返値 | | bool | true:暂停 false:非暂停 |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | currentTime | |
| 処理概要 | | 获取当前播放进度 | |
| 参数 | 0 | void |  |
| 返値 | | double | 当前播放时间,单位sec |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | duration | |
| 処理概要 | | 获取音视频文件总时长 | |
| 参数 | 0 | void |  |
| 返値 | | double | 音视频文件总时长, 单位sec |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | OnTimeUpdate | |
| 処理概要 | | 进度更新通知 | |
| 参数 | 1 | double | 当前进度 |
| 返値 | | void |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | OnDurationChanged | |
| 処理概要 | | 总时长更新通知 | |
| 参数 | 1 | double | 音视频文件总时长 |
| 返値 | | void |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | UpdateReadyState | |
| 処理概要 | | 准备状态变更通知 | |
| 参数 | 1 | ReadyState | 准备状态定义  ReadyStateHaveNothing：  没有关于音频/视频是否就绪的信息  ReadyStateHaveMetadata：  关于音频/视频就绪的元数据  ReadyStateHaveCurrentData：  关于当前播放位置的数据是可用的，但没有足够的数据来播放下一帧  ReadyStateHaveFutureData：  当前及至少下一帧的数据是可用的  ReadyStateHaveEnoughData：  可用数据足以开始播放 |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | readyState | |
| 処理概要 | | 获取准备状态 | |
| 参数 | 0 | void |  |
| 返値 | | ReadyState | 准备状态的定义参看函数UpdateReadyState中ReadyState的定义 |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | UpdateNetworkState | |
| 処理概要 | | 网络态变更通知 | |
| 参数 | 1 | NetworkState | 网络状态  NetworkStateEmpty：  音频/视频尚未初始化  NetworkStateIdle：  音频/视频是活动的且已选取资源，但并未使用网络  NetworkStateLoading:  浏览器正在下载数据  NetworkStateLoaded：  加载完毕  NetworkStateFormatError：  格式错误  NetworkStateNetworkError：  网络错误  NetworkStateDecodeError：  解码错误 |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | networkState | |
| 処理概要 | | 获取网络状态 | |
| 参数 | 0 | void |  |
| 返値 | | NetworkState | 网络状态的定义参看函数UpdateNetworkState中NetworkState的定义 |

|  |  |  |  |
| --- | --- | --- | --- |
| 関数名称 | | playbackStateChanged | |
| 処理概要 | | 播放状态变更通知 | |
| 参数 | 0 | void |  |
| 返値 | | void |  |

从播放引擎被创建到开始播放的Sequence如图4-8所示，首先在播放器WebMe­dia­PlayerSmartAuto在Render­Frame­Impl被创建，WebMe­dia­PlayerSmartAuto所在的进程是Render进程，通过IPC通信，将把创建播放器的请求发送到Browser进程，在Browser进程的MediaPlayerSmartAuto中load函数被调用的时候，会通过动态加载函数创建WebMediaImpl，然后对WebMediaImpl进行初始化操作，在初始化函数initPlayer中创建NMClassicMediaPlayer，并调用其加载函数。然后，WebMediaImpl调用AVChannelManager的active函数激活通道，开始play。当播放器进入paly状态后，把PLAYING状态通过回调函数playStateChanged()通知给Chromium浏览器，这时候用户在界面上看到的音视频文件在播放。

注意，图4-8所示的序列图只是一个示意图，在WebMediaPlayerSmartAuto和WebMediaImpl之间实机上还间隔的有进程间通讯及其他相关类，具体的可参看图4-7所示的类图关系。

在音视频被播放的过程中，如果有电话接入、或者有交通警报等高优先级声音切入，就会发生声音通道被抢占的情况，此时，需要暂停Media的播放，其Sequence图如图4-9所示。首先，NCSubChannelManager将声音通道被抢占的通知发送给WebMediaPlayer；然后WebMediaPlayer调用pause函数暂停播放，并把暂停状态通知给WebMedia­Player­­SmartAuto，WebMediaPlayer­SmartAuto将pause状态通知给Chromium浏览器，用户看到播放器被暂停，高优先级的声音切入。

图4-9的序列图也是示意图，实际的情况要比图示复杂，所涉及到的类也比较多，为便于问题的说明，在序列图中不再一一展示。

图4-8 播放器从被创建到开始播放的序列图

Browser进程

Render进程

Fig4-8 The sequence of Media player being created and start



图4-9 声音通道被抢占的序列图

Fig4-9 The sequence diagram of the sound channel occupied

4.2.3 Media Service分析与设计

我们希望所设计出的音视频播放系统除了支持Chromium音视频的播放外，还能支持本地等其他播放应用的需求。所以，我们把Media播放的业务实现做成一个Service运行于独立的进程，当客户端需要播放音视频的时候，就使用进程间通讯的方式和Serv­ice进行交互，图4-10展示了Media Service的类图结构。

在4.2.2小节中，我们提到WebMediaImpl使用了一个类叫NMClassicMediaPlay­er，这个类可以看做Media Service对外的client，当客户端需要使用播放的时候，可以使用此类和Service进行交互，进程间的通讯对于使用者来说是透明的。

在NMClassic­MediaPlayer中封装Service的proxy，即NMMediaPlayerProxy，它负责通过进程间通讯和Media Service的业务实现进行交互。进程间通讯的实现在不同的平台有不同的实现，例如在Android系统中可以使用binder完成。

NMMediaPlayerService是进程间通讯的另外一端，接收和处理NMMediaPlayer­Prox­y发来的各种请求，并把结果再通过NMClassicMediaPlayerListener通知给调用者。

NMMediaPlayerClientManager是Service端播放器的管理者，真正播放器的创建，是由NMMediaPlayerClientInst完成。NMUnifiedPlayer会根据不同的需求创建NMStr­ea­m­PlayerImpl或NMGstPlayerImpl。

在NMStream­PlayerImpl和NMGstPlayerImpl中使用Gstreamer插件完成对音视频流媒体和文件的解码播放。类图的最下方是外部依赖的库，由硬件供应商提供。



图4-10 Media Service 类图

Fig4-10 The class diagram of media service

4.2.4 AVChannel Manager分析与设计

AVChannel Manager模块控制与裁决source优先级的管理模块，根据不同的情况调整哪个source播放声音，哪个source暂停播放，这里的source指的是音频源或视频源。图4-11展现的是AVChannel Manager模块的类关系。



图4-11 AVChannel Manager 类图

Fig4-11 The class diagram of AVChannel Manager

在上图（图4-11）中，AVChannelManager是提供给用户使用的类，被Chromium模块的WebMediaImpl调用，假设WebMediaImpl要开始播放一段视频资源，首先要使用AVChannelManager的函数void activate(MediaTypeMask types, OutputMask outputs)来激活声音通道，当申请的声音通道被激活后，会在notifyStatusChangedAsyn­c(c­­­onst std::string &audioName, const std::string &status)通知状态变化，然后就可以开始播放了。

在AVChannelManager中封装了NCSubChannelManager，AVChannelManager将调度的实现委托给是NCSubChannelManager，NCSubChannelManager是进行source调度的实现者。

NCSubChannelListener是NCSubChannelManager的回调类，将状态的变化以及对NCSubChannelManager函数调用的结果反馈给调用者，调用者在使用NCSubChannel­Manager前需要首先实现NCSubChannelListener。

AVChannelClient是NCSubChannelListener子类，实现对NCSubChannelManager的通知回调处理，在AVChannelManager被创建的时候，AVChannelManager会创建一个AVChannelClient，然后把它注册到NCSubChannelManager。

NCSubChannelPluginCtrl是更底层的类，和硬件通讯，控制音视频通道的开启与关闭。它依赖于第三方库libnmavmplugin，为便于使用，是对依赖库的二次封装。

4.3实现

上一节介绍了基于Gstreamer的Chromium音视频播放系统的分析与设计，本节主要论述其实现过程。关于Chromium源码的获取及移植不是本文重点，这里不再展开，具体可参考Chromium Project官方网站<https://www.chromium.org>，有比较详细的说明。

4.3.1 Chromium Media Portability Implement相关的实现

在图4-7中，我们看到在类RenderFrameImpl中有一个函数是createMediaPlayer，这里便是Chromium的Media被创建的地方。注释掉原来创建函数的代码，改为我们自己的创建播放器代码，此函数中我们调用自己创建player的函数CreateSmartAuto­Web­­Media­Player() ，在此函数中，创建WebMediaPlayerSmartAuto对象，并把它交给Renderer­MediaPlayerManager对象进行管理。WebMediaPlayer­SmartAuto是WebMedi­aPlayer的子类，函数返回值是WebMediaPlayer的指针。

当WebMediaPlayerSmartAuto对象被创建完成，它的load函数将会被调用。在WebMediaPlayerSmartAuto的load函数中，完成初始化操作，此过程是一个跨进程的操作。首先，我们上面看到的RenderFrameImpl对象和WebMediaPlayerSmartAuto对象，它们均运行于Render进程，这里的WebMediaPlayerSmartAuto本子上来讲只是一个proxy，它没有真正的业务逻辑的实现，所有动作都是抛到另外一个进程，即Browser进程完成。

在load被调用的时候，内部函数InitializePlayer被调用，在InitializePlayer通过Renderer­MediaPlayerManager发送一个初始化的消息：

Send(new MediaPlayerHostMsg\_Initialize(routing\_id(), media\_player\_params))

此消息在另外一个独立进程Browser中被MediaWebContentsObserver接收，Media­WebContentsObserver接到初始化消息后调用BrowserMediaPlayerManager的OnInitialize函数。

在BrowserMediaPlayerManager的OnInitialize函数中，创建MediaPlayerSmartAuto对象，然后调用MediaPlayerSmartAuto的load函数。MediaPlayerSmartAuto被创建时，在其构造函数中，通过动态函数加载动态库libWebMedia.so：

dlopen("libWebMedia.so", RTLD\_NOW)

利用这个动态库，创建WebMediaImpl对象：

m\_createWebMedia = (createObjFunc \*) dlsym(m\_dlHandle, "createObj");

createObj的实现在WebMediaImpl.cpp文件里实现：

extern "C" nutshell::web::WebMedia \*createObj(nutshell::web::WebMediaCb \*cb)

{

return new nutshell::web::WebMediaImpl(cb);

}

至此，完成了从Render进程到Browser进程Media Player的创建过程。通过上述介绍可以看出，Render进程的Player只是一个影子，它通过Chromium独特的进程间通信的方式，在Render进程中将一切的函数请求转换成消息抛给Browser进程，Browser进程收到消息后对消息进行处理，在Browser进程里创建了另外一个Meidia Player来响应Render进程的对应请求。

之所以这样做，一是因为Chromium的多进程架构设计，关于Chromium的多进程架构可以参看官方网站，里面以形象的图文形式对Chromium的多进程架构做了直观的说明。二是因为一个Chromium实例可能会有多个Render进程，但只会有一个Bro­w­­­s­er进程，将Media Play的业务实现放在Browser进程便于管理。

4.3.2 Media Service相关的实现

Media Service是一个独立的服务，在系统启动时候被创建，然后等待客户端来连接。在MediaServiceMain的实现中，首先创建NMMediaServerProc进程：

NMMediaServerProc process(argc, argv)

然后对其进行初始化处理：

process.initialize(NMMediaServiceFac)

开启进程：

process.start()

进入到进程loop中，等待连接通信：

process.enterloop()

如果进程遇到终止的条件，则跳出loop，调用stop函数：

process.stop()

最后是反初始化：

process.deinitialize()

Media Service运行起来后，就处于等待处理的循环中。

在上一节提到，MediaPlayerSmartAuto会创建WebMediaImpl对象，当WebMedia­Impl被创建，就会使用NMClassicMediaPlayer中的NMMediaPlayerProxy与NMMedia­Player­Service通信。NMMediaPlayerService通过NMMediaPlayerClientManager创建NMUnifiedPlayer，在NMUnifiedPlayer构造函数中根据feature的不同，来创建不同的Player进行处理。

如果检测到是流媒体的feature，就创建NMStreamPlayerImpl进行播放。在NMStream­PlayerImpl的prepareImpl函数中创建gloop和上下文：

m\_glibLoop->initLoop()

创建管道类：

createPipelineClass(&m\_pipelineClass, m\_mediaPath.c\_str(), NULL,

"pulsesink", "iautovideosink", NULL, elementCreateCallBack, this)

连接bus总线：

attachBusCall(GST\_ELEMENT\_CAST(m\_pipelineClass->pipeline()), m\_glibLoop->getGlibContext(), (GSourceFunc)busCall, this)

开启loop：

m\_glibLoop->startLoop()

如果检测到不是流媒体的feature，就创建NMGstPlayerImpl进行播放。在NMGst­PlayerImpl的prepareImpl函数中首先调用gstPrepare函数，gstPrepare完成gloop和上下文的创建：

m\_glibLoop->initLoop()

创建管道类：

createPipelineClass(&m\_pipelineClass, m\_mediaPath.c\_str(), NULL,

"pulsesink", "iautovideosink", NULL, elementCreateCallBack, this)

获取管道引用：

m\_pipeline = GST\_ELEMENT\_CAST(m\_pipelineClass->pipeline())

连接bus总线：

m\_busCallSource->attachBusCall(m\_pipeline, m\_glibLoop->getGlibContext(), (GSourceFunc)busCall, this)

开启loop：

m\_glibLoop->startLoop()

当调用到gstPlay()的时候，首先查看当前的播放状态，如果是PLAYING状态就直接返回，否则就调用setGstState改变状态开始播放：

setGstState(GST\_STATE\_PLAYING, NC\_TRUE, STATE\_CHANGE\_TIMEOUT, NC\_TRUE)

setGstState函数则是对gst\_element\_set\_state(m\_pipeline, state)的封装。

4.3.3 AVChannel Manager相关的实现

AVChannelManager在WebMediaImpl函数调用，用于通道的申请。当AVChannel­Manager的init函数被调用的时候，会通过调用NCSubChannelManager的registerAvm函数完成source的注册，即：

mNCSubChannelManager->registerAvm(NCString(WEBMEDIA\_AUDIO\_NAME), mClient)

当需要激活通道的时候，调用函数：

mNCSubChannelManager ->reqAudioSource(NCString(WEBMEDIA\_AUDIO\_NAME), true)

在NCSubChannelManager内部，会调用NCSubChannelPluginCtrl来实现，其本质是通过uComLib与硬件进行交互，uComLib是硬件供应商提供的动态库，此处不再展开。

4.4 本章小结

本章提出了基于Gstreamer的Chromium音视频播放系统的总体架构，将总体架构进行模块划分，并提出个模块的详细设计方案，对于比较重要的部分用类图及序列图做了必要的说明，在实现中对重要的代码和逻辑进行了代码级讲解，并阐述了某些关键函数的实现方法，对整个开发过程做了提纲挈领的描述。

5 测试与优化

本章是对上一个章节成果的检验，从功能和性能的角度测试所设计和实现的系统，最后对存在的问题提出优化和整改方案。

5.1 功能测试

5.1.1 测试项目与测试结果

1. 网络音频播放

这个测试项目主要用于测试在网络连接状态良好的情况下，使用Chromium播放网络音频文件的效果。其测试的流程可以描述为：首先保证网络状况连接良好，然后打开浏览器，输入常见的音频网站的网址，比如虾米音乐、百度音乐；进入到歌曲的播放列表界面，点击其中的一首音乐，等待音乐的声音流出。我们同时测试了虾米音乐网站和百度音乐网站，在这个两个网站上都能听到音乐正常的播放。

相关的测试截图如图5-1所示。

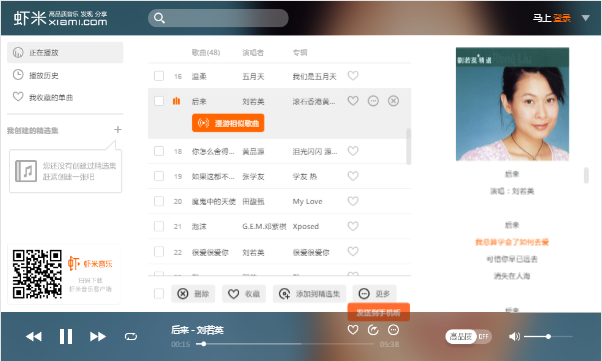
 

图5-1 虾米和百度音乐播放截图

Fig5-1 The screenshot of xiaomi and baidu music net

2. 本地音频播放

这个测试项目用于测试使用Chromium打开本地音频文件播放的效果。首先我们将准备好的test.mp3放置到车载多媒体中终端内，然后点击Chromium选择打开本地文件test.mp3，等待音频测试文件的声音流出。

在Chromium浏览器中播放本地视频文件的测试效果截图如图5-2所示。



图5-2 播放本地音频的截图

Fig5-2 The screenshot of playing local audio

3. 网络视频播放

此项目用于测试在网络连接状态良好的情况下，使用Chromium播放网络视频问价的效果。在这里，我们选取了国内外排名前十的网站进行测试，篇幅所限，就不在一一列举。我们以youtube和搜狐视频为例进行说明。首先输入要测试的视频网站的网址，进入到播放列表界面；然后选择其中的一个视频进行点击；等待视频的播放。如果能看到视频流畅的播放，同时声音正常的播出，则表明测试结果OK。

图5-3所示，是我们测试youtube和搜狐视频的截图。

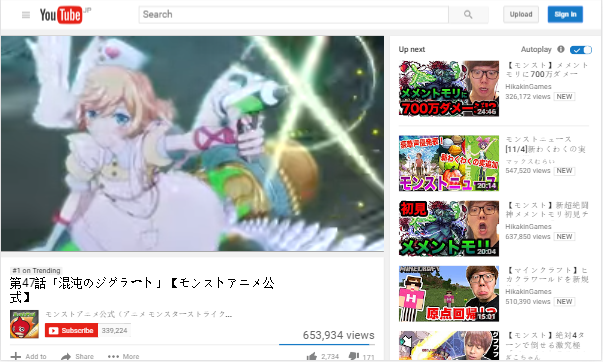


图5-3 播放youtube视频和搜狐视频的截图

Fig5-3 The screenshot of playing youtube and sohu video

4. 本地视频播放

与测试播放本地音频的方法一样，提前将准备好的视频文件放入车载多媒体终端，使用Chromium浏览器打开准备好的视频文件movie.ogg，点击播放按钮进行视频的播放。

效果图如图5-4所示。



图5-4 播放本地视频的截图

Fig5-4 The screenshot of playing local video

5. 播放的管理控制

这个测试项目主要用于测试Chromium在播放音视频的过程中，点击暂停、调节音量、静音按钮的控制效果。首先，打开一个音频或者视频，点击播放按钮进入到播放状态；然后，点击暂停，观察是否进入到暂停状态；点击音量调节滑块拖动，观察音量的大小是否有变化；反复点击静音按钮，观察声音是否有消去和恢复。

经测试，播放的管理控制功能良好。

6. 跳转

跳转用于测试在Chromium播放音视频过程中，拖动进度条到任意时间播放的效果。在正常状态下，用浏览器打开音频或视频文件，进入到正常播放状态；拖动进度条按钮到任意时间点；松开进度条按钮，观察是否从该时间点开始播放。经测试，该系统从指定的时间点开始播放音视频。

7. 事件通知

此项用于测试在Chromium播放音视频过程中，当有优先级较高的事件发生的时候的Notification。首先，我们启动Chromium浏览器，正常的播放音视频文件；这时候发生电话接入事件；观察播放状态是否变成暂停状态。经测试，确认在有电话事件、紧急广播事件、交通信息事件通知的时候，音视频播放系统进入到暂停状态，将声音通道让给高优先级事件响应。

8. 错误处理

在Chromium播放音视频的过程中，可能会遇到错误的发生，比如网络的中断、解码的错误、获取资源的失败，等等。当这种现象发生的时候，需要及时的反馈给使用者。打开Chromium浏览器，在网络状况良好的情况下播放音频或视频文件，在正常播放的过程中，断开网络连接，观察有无“网络连接失败”的警告弹出；准备一个文件中间有损坏的视频文件，使用Chromium打开进行播放，观察在出错的地方是否有解析失败的弹窗提示；准备一个不不支持的视频或音频文件，使用Chromium打开进行播放，观察Chromium是否有解码失败的弹窗提示。

经测试，以上Error信息都能准确的反馈给用户。

9. WebAPP的播放支持

在第3章我们提到过要对WebAPP进行支持，这个项目就是对WebAPP音视频播放相关部分进行的测试。首先，我们可以安装第三方开发的WebAPP或者自己写的测试版WebAPP；然后，打开APP进行音视频文件的播放，观察播放是否正常，各管理控制是否可用。

经测试，对WebAPP音视频播放支持良好。

10. 传统播放器的支持

前边我们的测试都是基于Chromium做客户端进行的测试，现在我们测试使用车载多媒体终端的其他播放器是否可以播放音视频文件。首先我们将准备好的音视频文件拷贝入U盘；然后打开车载多媒体终端的本地播放器，浏览U盘内的文件，可以看到在给定的路径下，播放器可以快速的获取该路径下的音视频文件列表；然后选择其中的一个进行播放，可以看到播放顺利进行；点击上一首/下一首，观察是否有正确的跳转；点击快进/快退，观察是否发生快进快退；点击倍速播放，观察播放的速度是否有相应的倍数增加。

经测试，以上各项功能正常。

11. 封装格式和编码格式的支持

依照需求分析章节表3-1列出的文件封装格式和编码格式，分别准备对应的网络和本质音视频文件进行测试，观察其解码的正确性和效率。

经测试，表3-1中所列格式都能得到良好支持。

5.2 性能测试

5.2.1 测试项目与测试结果

1. 网络加载响应时间

此项目用于测试Chromium在播放网络音视频的时候，加载响应时间。首先，我们准备一个网络状况不好的情况，比如给wifi限速；然后，打开Chromium登陆到音频或视频播放网站，打开某一个音频或视频资源；点击播放按钮，观察加载响应时间是否在1s之内，在这里我们可以看到加载缓冲界面的百分比。

经过输出log，我们检测到在1s之内就发起了资源的加载，测试通过。

2. 网络加载进度更新频率

在网速较慢的情况下，加载的时间可能很长，加载进度需要及时的通知给上层调用者，通知的频率要求10ms一次。这里我们通过输出log的方式来验证通知频率。首先，对测试环境的网络进行限速；使用Chromium打开含有音视频资源的网页，单击播放按钮进行播放，此刻将看到加载进度图标在不停的旋转，同时显示加载百分比；此刻，我们观察log，由时间戳可以发现其更新进度在10ms左右。

经测试可知，加载进度更新频率基本符合目标。

3. 暂停网络音视频的响应时间

该测试项目用来测试Chromium在播放网络音视频的过程中要求暂停时的响应时间。首先，打开Chromium正常播放音视频；在播放的过程中点击暂停按钮，观察播放按钮是否在1s的时间内变成暂停状态，同时观察输出的log的时间戳响应时间是否在1s之内。

经测试，符合要求。

4. 停止网络音视频播放后资源释放

在第3章有一个需求是如果客户决定stop掉某个音视频文件，那么在10s该资源将被清空。此测试case要测试的就是stop后的资源是否被释放。首先，写一个带有停止按钮的测试页面，打开浏览器使用此页面进行正常的音视频文件播放，在播放过程中点击停止按钮，通过log和内存观察在stop点击10s后资源是否被释放。如果在10s内点击了播放，则资源不需释放。

经测试，符合期望。

5. 本地音视频资源文件播放的响应时间

本地资源文件由于存储在本地的媒体介质中，无需使用网络资源，不管其文件的大小，本地播放要能迅速的相应，最差的情况从点击播放到进入到播放状态，时间间隔不能大于两秒。使用Chromium浏览器打开已经准备好的画质较好体积较大的视频文件进行播放，通过时间计时器观察用户体验，同时通过log输出的方式检查系统状态的切换时间，观察响应时间是否控制在2s内。

经测试，使用多种格式的音视频文件，其响应时间都控制在2s内。

6. 本地音视频资源文件暂停响应时间

此项目用来测试Chromium在正常播放本地音视频的情况下，押下暂停按钮，系统切换到暂停状态所花费的时间。用Chromium打开准备好的本地音视频文件，让其进入到播放状态；然后，单击暂停按钮，观察记录系统由播放状态切换到暂停状态的时间间隔。

经计时观察和输出log分析，状态切换控制在1s内。

7. 本地音视频文件播放品质

当播放本地音视频文件的时候，不能出现画面和声音的卡顿，不能出现画面的色块和声音的杂音。该测试项目将会对我们支持的所有格式的音频和视频文件进行长时间的播放观察，每组测试人员对不同的格式分开验证。

经测试，画面流畅，音质良好。

8. 长时间运行测试

循环播放网络音视频列表及本地音视频列表，接通电源长时间使用此系统播放音视频文件，观察是否有死机现象、是否有重启现象、是否存在黑屏现象。我们搭建了测试服务器，将测试用音视频文件1000多个上传至服务器后，在页面内设置循环播放，连续测试一周，没有发现有死机、重启、黑屏现象。

经抽查测试，10台机器连续播放两周，没有出现死机、重启、黑屏现象。

5.3 优化

5.3.1 占用资源的优化

在嵌入式系统中，由于硬件资源十分有限，如果同时打开过个页面进行音视频的播放，则会占用大量的内存空间及处理器的消耗，经过测试，如果同时打开的播放音视频的页面过多，会造成播放卡顿、功耗过大、机器发烫现象。一般来说，在可移动的设备上，不太赞成支持多个音视频文件同时播放，以节省资源。所以，需要对各个页面的播放请求进行管理。

在第4章中我们知道，所有的播放工作都是在Browser进程进行的，而Browser进程在Chromium中是唯一的。所以，每当有播放请求的时候，可以在Browser进程的BrowserMediaPlayerManager对所有播放请求进行管理，控制同时播放的视频的个数。因此，我们可以在新的音视频播放请求加载的时候，暂停之前所有的播放，即每次只播放最新的请求暂停之前的音视频，从而降低资源消耗。

5.3.2 视频描画效率的优化

当前视频的播放方式是先由解码器进行解码，然后将解码后的数据存入到一块内存里，Chromium读取内存资源进行上屏描画，这是一个跨进程的过程。为了提升效率，可将其改为，由Media Service直接描画，Chromium浏览器在页面上挖洞的方法，透到下一层的Media Service播放视频的Layer中，在体验上来说，用户看到的依然和之前的效果一样，但是减少了进程间通信的频率，描画效率得到了提升。

5.4 本章小结

本章介绍了本论文的测试及测试结果，是对第2、3、4章整体分析、设计与实现的成果的检查与验证，然后对设计和实现提出了进一步的优化方案。总体来说，系统设计与实现与初中目标一致，基本完成了基于Gstreamer的Chromium音视频播放系统的设计与实现。

6 总结与展望

本章是该论文的最后一章，在本章中我们对前面的工作进行了总结，陈述本文完成的主要内容，指出尚且存在的问题，并对未来的研究做出规划和展望。

6.1 工作总结

本文通过对嵌入式设备多媒体数据应用的现状和趋势的了解，结合车载多媒体终端软硬件开发平台技术，对常见构建流媒体应用的开源多媒体框架及浏览器音视频播放体系结构的相关技术进行学习研究。根据开发需求和开源项目Chromium的分析，设计并实现了基于Gstreamer架构的Chromium音视频播放系统。

本课题完成的主要内容包括以下几个方面：

1. Chromium浏览器多媒体播放相关的架构分析与重构设计

Chromium是谷歌公司发布的开源浏览器，具有高效、稳定等特点，它的渲染引擎选用的是主流的WebKit，由于采用了沙盒机制，所以其安全性得到保证。Chro­m­ium是一个开源项目，任何组织和个人都可以在其工程的官网上获取源代码。对于多媒体播放相关的模块，Chromium本身提供了一个简单的实现方式，但是非常具有局限性，不能共通使用。在资源十分有限的嵌入式车载多媒体终端产品中，不可能为浏览器单独写相应的多媒体应用模块，一般来说，是多个应用共享一个音视频播放动态库，Chromium也调用此共享库进行音视频的播放。因此需要对Chromium音视频播放相关的部分代码进行重构，添加中间层并抽象出比较通用的接口。这一部分的主要工作是：

（1）分析Chromium音视频播放相关的代码结构及实现流程；

（2）根据需求，分析和抽象出中间层用于共通接口的适配；

（3）在原来的基础上重构优化现有的结构和代码。

1. 掌握Gstreamer架构的原理和使用方法

Gstreamer是一个多媒体开发框架，基于此框架，开发者能够快速的开发出满足需求的多媒体应用程序。其提供的功能元件能够巧妙的接入到任何应用管道程序中，是这种框架最吸引开发者的优点，这种优点可以使设计和开发的复杂度大大降低。Gstr­eamer是基于插件的，它可以把多个元件封装起来组成一个具有特定功能的模块，以插件的形式提供给第三方使用，开发者进行开发时可以引用这些功能模块设计特定的处理程序。几个比较重要的基础结构是：元件(Elements)、衬垫（Pads)、管道(pipelines)、箱柜(bin)。Gstreamer通过把若干elements链接在一起构成pipeline实现对媒体内容的处理，element通过plug-in的方式提供。bin是一种特殊的element，是由多个其它elements组成的。Gstreamer的核心实现了对plug-in的注册和加载等功能，可以通过编写plug-in方式对gstreamer的功能进行扩展，包括编码方式，封装格式等各种功能。

这一部分的主要内容是：研究并设计基于Gstreamer开发框架播放器的后台处理模块架构方案，阐述后台播放通道的创建过程，Gstreamer架构对于事件的处理机制和方法，处理流程等等。

1. Gstreamer音视频播放的实现与接口封装

这一部分的主要内容是：基于Gstreamer开发框架后台处理模块架构方案编程完成后台模块的应用程序；将Gstreamer的实现用接口封装起来，做成一个服务，可以被整个车载系统的任一应用程序加载调用，例如本地播放USB存储器中的音乐、视频。这些接口主要包括：播放，暂停，停止，快进，快退，选时播放等等。

1. Chromium浏览器多媒体播放接口与Gstreamer封装接口的适配

在1和3中已经分别对接口进行抽象和封装，这一部分的主要内容就是二者接口的适配与调试，这里考虑采用适当的设计模式，使在使用上更加灵活。还要考虑线程与进程的问题，因为Chromium使用了较多的进程与线程。

最后，根据需求分析提到的各点进行测试验证，基本达到本课题所立的目标。

6.2 问题和展望

目前对于基于Gstreamer的Chromium的音视频播放系统来说，虽然实现了其基目标，但是还存在一些问题和未来亟待完成的工作。

1. 在视频播放过程如果遇见网页嵌套层次过多的情况下，快速滑动当前播放视频的网页会出现视频播放窗口和网页界面坐标位置不能同步的现象，后续可通过修改视频窗口位置坐标的算法进行改善。
2. 丰富Gstreamer的解码功能，实现更多音视频格式的解码播放。
3. WebRTC（Web Real Time Communication）的实现，即网页及时通讯。

参考文献

[1] 阮胜才.视频点播系统的设计与实现[M].安徽：中国宇航出版社, 2008: 634-635.

[2] 高睿鹏，刘佳玲.基于 FFMPEG 的通用视频插件[D].内蒙古：内蒙古工业大学, 2010, 27(1)：2474-2476.

[3] 赛迪顾问半导体咨询事业部.2005年中国汽车电子市场规模达624 亿元.世界电子元器件，2006年第3期.

[4] 李淑荣.基于Gstreamer的网络媒体播放系统的开发[D]. 中国海洋大学: 中国海洋大学,2012

[5] wikipedia.http://en.wikipedia.org/wiki/GStreamer[J/OL]. last modified on 15 May 2015

[6] Richard John Boulton,Erik Walthinsen,Steve Baker,Leif Johnson,Ronald S. Bultje,Ronald S. Bultje,Tim-Philipp Muller,Wim Taymans.GStreamer PluginWrit-er's Guide ­(1.­5.0­.1­)­[­ J­/O­­L]. http://gstreamer.freedesktop.org/data/doc/gstreamer/head/pwg/html/inde­x.html. 2015.5.21

[7] Wikipedia, http://en.wikipedia.org/wiki/Dynamic-link\_library[J/OL].last modified on 16­ May 2015

[8] Wikipedia, Web Brwoser[EB/OL]. http://en.wikipedia.org/wiki/Web\_browser

[9] Tim Berners-Lee. WorldWideWeb, the first Web client[EB/OL/].http://www.w3.org/People/Berners-Lee/WorldWideWeb.html, 2015-6-28.

[10] 赵经纬,周余,王自强等.基于WebKit的嵌入式浏览器的研究与实现.电子测量技术,200934(3),135-137

[11] 叶卿.嵌入式微浏览器的设计与实现:[D],北京:北京邮电大学.2005,22-25

[12] 西村贤. [EB/0L/].http://www.atmarkit.co.jp/news/200903/30/chrome.html, 2009-3-30.

[13] 朱永盛.WebKit技术内幕[M].北京:电子工业出版社,2014:2-4

[14] 肖梦华.面向智能电视的嵌入式浏览器平台的研究与设计[D].上海:复旦大学.2013

[15] 许晔.发展嵌入式系统是我国后PC时代的战略选择[J].中国科技投资,2010,2:2-3

[16] 梁融凌,余昌嵩. 基于嵌入式Linux电子书的设计与实现[J].牡丹江师范学院学报(自然科学版),2014,4:13-14

[17] 冯礼哲,高声荣,罗怀琴,李逸.发展中的4G网络[J].科技视界,2015,5:23

[18] 罗锐.基于Blackfin平台的嵌入式播放器的研究与实现[D].成都:电子科技大学.2009

[19] 仇昊.基于GStreamer的嵌入式多媒体系统的研究与实现[D].北京:北京邮电大学.2011

[20] Taymans W, Baker S, Wingo A, etal.Gstreamer Application Development Manual (0.10.35.1)

[21] 黄铁军.视频编码国家标准AVS与国际标准MPEG的比较[DB/OL]. http://www.a-vs.org.cn/，2005.

[22] [https://en.wikipedia.org/wiki/Web\_browser, last modified on 7 June 2016](https://en.wikipedia.org/wiki/Web_browser,%20last%20modified%20on%207%20June%202016)

[23] 中国互联网信息中心.中国互联网报告[M].北京.2007

[24] http://trac.webkit.org/wiki/WebKit2. Hosted by Apple. 2016.6.15

[25] <https://gstreamer.freedesktop.org> 2016.8.15

[26] Taymans Wim, Baker Steve, Wingo Andy, et a1. GStreamer Application Development Manual(0.10.33)[EB/OL],https://gstreamer.freedesktop.org/data/doc/gstreamer/0.10.33/manual, 2011-05-10

[27] 张治忠.基于SEP6200处理器的GStreamer媒体播放器优化与实现[D].南京:东南大学.2013

[28] 贺志强.基于Gstreamer媒体播放器的研究与设计[D]. 成都:电子科技大学.2006

[29] 张剑锋.基于GStreamer的STB多媒体播放系统设计与实现[D].上海:上海交通大学.2009

[30] 代坤娟.基于自主Soc媒体播放系统中的音视频同步优化与实现[D].南京:东南大学.2011

[31] 李亚飞.分布式视频转码系统的设计与实现[D].哈尔滨:哈尔滨工业大学.2014

[32] The WebKit Open Source Project [EB/OL]. http://www.webkit.org/

[33] Sten H, Tiger G. Web Operating System for Modern Smart phones[J]. 2011.

[34] 段虎才,倪宏,邓峰等.WebKi t内核的嵌入式浏览器磁盘缓存方法[J].计算机科学与技术.2015,(3):17-19

[35] Google.Multi-process Architecture[EB/OL].http://www.chromium.org/developers/desi­gn­­­-­­­documents­/multi-process-architecture, 2016.9.10

[36] S Alimadadi, S Sequeira, A Mesbah, K Pattabiraman. Understanding JavaScript Event-Based Interactions with Clematis[J]. Acm Transactions on Software Engineering & Methodology. 2016, 25(2):1-38

致 谢

阔别校园已经4年，而今有机会再次步入大学殿堂开始一段新的学习生活，我感到特别的荣幸和感恩。在这几年的时间里，我不仅系统地学习了软件工程相关的知识，而且丰富了分析问题的方法和思路。更让我印象深刻的是敬爱的老师们的谆谆教诲、同学们真诚的友谊。真的非常感谢，感谢良师益友，感谢陪伴的每时每刻，感谢无法追回的岁月，感谢留下的美好回忆。

在这篇论文的完成过程中，我的导师姚建国老师给予了我很大的帮助，从选题到开题报告，从写作纲要到认真地指出每版存在的不足，严格把关，循循善诱，在此表示衷心的感谢。姚老师治学态度严谨，工作态度认真，待人和蔼、朴实无华，不仅带领我树立了自己的学术目标、掌握了基本的研究方法，还使我明白了许多待人接物与为人处世的道理。本论文从开始到完成，每一步都离不开导师的指点和陪伴。再次，向导师表示崇高的敬意和衷心的感谢！

感谢王圣浩博士，在论文的写作期间给予我许多中肯的意见和建议。此外，本文研究工作还得到了上海商泰软件有限公司、杭州微纳科技股份有限公司的各位领导与同事的热情帮助，他们在专业领域的通晓给了我极大的帮助，在此向他们表示真诚的谢意！

还要感谢学校所有辛苦工作、为我们精心安排每次学习与活动的老师们，感谢他们给予我学业上无私的教诲和生活上亲切的关怀！感谢陪伴我的同学，感谢他们在学习和论文写作过程中给予的帮助和敦促。

同时，尤其感谢多年来一直给予我鼎力支持和无私奉献的父母以及支持和照顾我学业和生活的妻子，感谢他们默默的付出。

# 攻读学位期间发表的学术论文目录

[1] 胡济豪，Android系统智能电视HAL层Sensor数据传输的一种实现[J/OL]，上海交通大学软件学院网站公示http://elearning.se.sjtu.edu.cn/announcement/index.asp?ran­g­­e=al­l&courseid=1073&page=6，2016.12