

基于 WebRTC 技术的诊后医疗系统实时通信研究^①

李鹏树, 范 辉

(北京交通大学 计算机与信息技术学院, 北京 100044)

通讯作者: 李鹏树, E-mail: 553121984@qq.com

摘 要: WebRTC 技术的出现提出了实时音视频处理的解决方案, 所具有的开放标准为研究实时音视频的处理流程提供了技术依据. 通过对 WebRTC 的核心音视频操作库进行研究, 将 WebRTC 的技术方案应用于诊后医疗系统的在线问诊模块, 用以提升线上问诊服务的音视频质量. 通过对实时通信进行研究并依托 WebRTC 协议的通信体系结构, 设计了信令和媒体代理方案并最终实现了诊后医疗系统的实时通信.

关键词: WebRTC; P2P; 实时音视频; 背靠背用户代理服务器 (B2BUA); 会话边界服务器 (SBC)

引用格式: 李鹏树, 范辉. 基于 WebRTC 技术的诊后医疗系统实时通信研究. 计算机系统应用, 2018, 27(6): 75-82. <http://www.c-s-a.org.cn/1003-3254/6403.html>

Research on Real-Time Communication of Post-Diagnosis Medical System Based on WebRTC Technology

LI Peng-Shu, FAN Hui

(College of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

Abstract: With the advent of WebRTC technology, the solution of real-time audio and video processing was provided. Its open standards provide a technical basis for the further research of real-time audio and video processing. Through the research of the core audio and video operation library of the WebRTC, we apply the technical solutions of WebRTC in the online interrogation module of the post-diagnosis medical service system, to improve the quality of the audio and video of online consultation service. Through analyzing the communication of the core audio and video of the WebRTC and depending on the real-time communication architecture of WebRTC protocol, we design the scheme of signaling agent and media agent, and finally realize the real-time communication of the post-diagnosis medical service system.

Key words: WebRTC; Peer-to-Peer (P2P); real-time audio and video; Back-to-Back User Agent (B2BUA); Session Border Controller (SBC)

WebRTC (Web Real-Time Communication) 是一项网页实时通信技术并将逐渐成为下一代音视频通信的标准, 它提供了实时音视频处理的解决方案, 并在当前 Android 系统内置浏览器和 PC 端主流浏览器对该项技术提供支持^[1]. 现如今我国的移动医疗领域发展迅速, 移动问诊的兴起和发展拓宽了患者的就医渠道, 但现阶段移动问诊在实时通信方面的用户体验度不好.

本文对 WebRTC 技术进行分析后, 研究诊后医疗系统中实时问诊业务下的实时通信模块, 主要对实时通信中所涉及信令协商、网络穿越和音视频处理等处理流程进行研究实现, 将 WebRTC 的音视频处理技术应用于移动医疗领域中, 结合出院患者住院信息提出诊后医疗系统以应对慢性病患者出院后期再次就医问题, 建立患者与其主治医师间的在线绑定, 开发出具有高

① 收稿时间: 2017-09-25; 修改时间: 2017-10-18; 采用时间: 2017-11-17; csa 在线出版时间: 2018-05-28

质量实时音视频通信的在线问诊医疗服务系统。

实时通信平台的设计包括信令控制、网络穿越、音视频媒体数据处理与传输控制,考虑到跨平台的应用本系统应用会话初始化协议 (Session Initiation Protocol, SIP)^[2]和会话描述协议 (Session Description Protocol, SDP)^[3]实现会话协商,设计背靠背用户代理服务 (Back-to-Back User Agent, B2BUA) 和会话边界服务器 (Session Border Controller, SBC)^[4]实现信令与媒体代理进而实现网络穿越功能,依托 WebRTC 协议体系并应用 WebRTC 音视频处理技术实现实时通信。

1 WebRTC 技术的引入

1.1 WebRTC 技术简介

WebRTC 是一项支持跨平台的实时音视频处理的解决方案,它提供了音视频处理的核心技术,功能结构框架主要分为两部分: Web API 和核心库^[5],如图 1 所示。Web API 主要供 Web 开发者使用以方便进行浏览器端开发,核心库是内嵌到浏览器中的,这是其核心内容。WebRTC 的核心类库主要包括 Native API、会话控制、音/视频处理引擎和网络传输控制等。

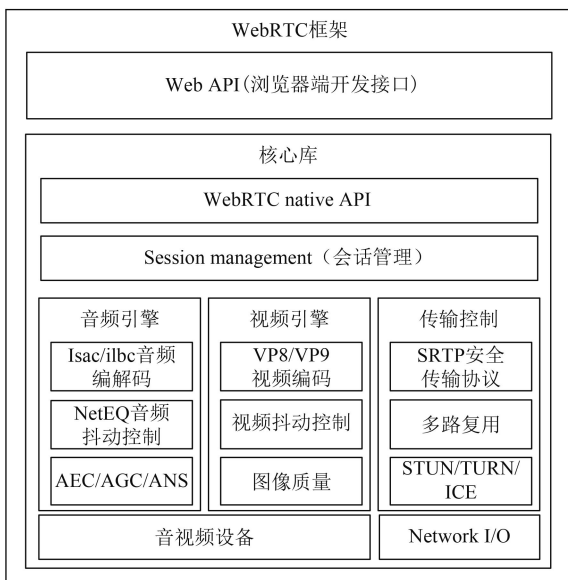


图 1 WebRTC 框架图

在 WebRTC 核心库中音频引擎、视频引擎和传输控制是 3 个重要的模块。音频引擎是 WebRTC 最突出的技术,主要是对语音数据进行处理,包括回声消除、噪声抑制、增益控制和静音检测等功能,用以提

升语音质量,实现了整个音频媒体链的技术框架,包括对音频的声学处理和音频编解码的实现。NetEQ 模块^[6]功能的加入减小了由于网络抖动和语音数据包分组丢失对语音质量的影响,从而在尽可能低的时延下提供高品质的语音通话。视频引擎实现了从采集、传输到展示的视频控制的框架,包括视频图像的增强处理和 VP8^[7]编解码的实现。在传输控制中,WebRTC 应用 STUN、TURN 和 ICE^[8]机制相结合来解决不同类型的 NAT 穿越问题,以实现实时音视频传输。

1.2 实时通信系统架构模式比较

实时通信平台的架构模式主要分为:混合处理结构 (Mixer)、路由转接结构 (Router) 和网络拓扑结构 (Mesh),前两种架构模式需要中间服务器转接以实现流媒体信息实时交互,两者的主要区别是服务端是否承接音视频处理功能。Mesh 结构主要特点是实现实时通信系统的 P2P (Peer-to-Peer) 在线传输方案,去掉了中间转接媒体信息的服务器,系统只需要部署信令协商^[9]和 NAT 穿越服务器,以完成终端用户的会话协商和获取相应网络地址,架构方案的比较如下:

(1) 混合处理结构 (Mixer): 服务端负责音视频处理,包括转码、混音、合屏等,当前多人视频会议基本上是种结构。

方案特点:

- ① 终端负载最小,理论上支持多人同时音视频互通。
- ② 可与现有产品无缝对接,最大程度利用硬件能力。
- ③ 服务端负载很大,系统部署成本很高。
- ④ 多延迟问题,服务端负责工作,如解码、合屏、混音、编码,因此会带来延迟。

(2) 路由转接结构 (Router): 服务端负责媒体包转发但不负责转码处理。

方案特点:

- ① 系统容易扩展,相比 Mixer 结构服务端压力减小。
- ② 低延迟,与 SVC 结合提升用户体验度。
- ③ 服务端要处理不同终端的接收能力问题,整体部署难度并未大幅下降。
- ④ 在一定程度上限制了通信系统的规模。

(3) 网络拓扑结构 (Mesh): 媒体数据包不需要经过服务端,终端直接 P2P 交互传输^[10],是最简单的实时通信系统的架构模式,WebRTC 主要应用于此架构模式。

方案特点:

- ① 服务端压力小,理论上不需要媒体转接服务器.
- ② 低延迟,终端直接 P2P 传输,减小数据转接延迟.
- ③ 终端负载较大,音视频的处理工作主要交于终端来处理,对终端要求相对较高.

2 WebRTC 技术要点分析

2.1 音频处理分析

WebRTC 技术的音频引擎提供了音频控制的整套解决方案,它不仅包括音频中非常重要的编解码模块 (Audio coding) 和音频处理模块 (Audio processing),还包括混音模块 (Audio conference mixer) 和采样频率控制模块 (Bitrate controller) 等. 音频引擎的整体结构如图 2 所示. 音频处理模块主要是对采集到语音数据进行处理,以降低语音数据中由于回声、噪声、声音不平稳等而造成通信效果不理想问题. 因此对捕获到的音频数据都会进入音频处理模块对音频数据进行相应的处理如回声消除 (AEC), 噪声抑制 (ANS), 增益控制 (AGC) 和静音检测 (VAD) 等.

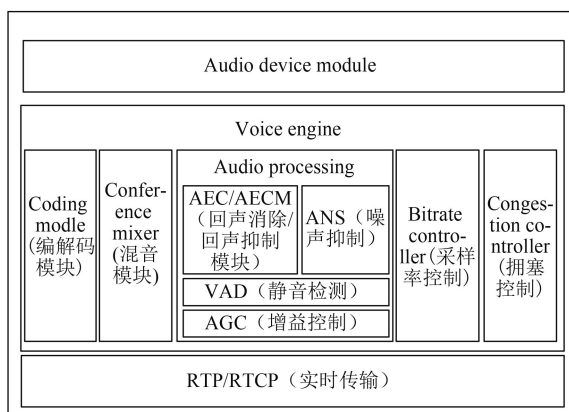


图 2 音频引擎结构图

WebRTC 中比较突出的是其对音频信息的处理,包括音频数据的采集、组包、处理、编码和发送等. 下面音频处理流程进行说明.

- (1) 调用音频设置进行音频数据采集,将采集到的数据交给音频引擎进行处理.
- (2) 媒体数据经过音频引擎处理后形成音频数据单元 AudioFrame^[11].
- (3) 将 AudioFrame 送入音频处理模块进行处理,进行回声和噪声等处理.
- (4) 处理完成后将语音数据送入静音检测模块,确定该音频数据是否为有效语音数据,决定是否进行传输.
- (5) 若通过静音检测,音频数据会经过混音处理,将音频数据分发到各声道,最终将音频数据送入 RTP 传输模块. 反之则重回步骤 (1) 进行音频采集.

其中,接收端接收到音频数据后,由于媒体数据可能是多声道传送,接收端需要将音频数据再次通过混音模块^[11]进行合音处理.

2.2 视频处理分析

WebRTC 对视频的处理主要提供了 5 个功能模块,主要包括视频数据采集 (Video capture)、数据处理 (Video processing)、视频编解码 (Video coding)、视频数据渲染 (Video render) 和实时传输模块 (RTP/RTCP),各模块间协作关系如图 3. 各模块的主要功能是,视频采集: 负责捕获视频数据; 视频处理: 对视频数据进行处理,包括降噪、抽取、采样等; 视频编解码: 对视频数据进行编解码处理; 视频渲染: 用于接收端对接收到的视频数据渲染处理并实现缓存,并投递给显示设备; 实时传输模块: 接收数据包和进行包解析以提供实时传输.

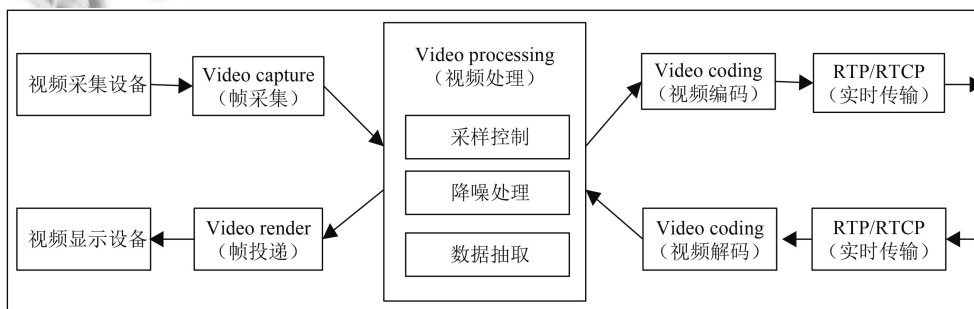


图 3 视频模块协作关系图

相应的视频发送流程是:当发送源端产生视频数据,会调用视频采集设备,将视频数据封装成视频帧,按照设定好的速率传到视频帧采集模块;视频采集模块收到数据后,对画面进行处理,如视频画面亮度调整等操作,处理完成后将视频帧数据送入处理模块进行噪声、抽取和采样等处理,处理完成后将数据送入编码模块进行编码;最后送入实时传输模块.在视频接收处理过程中,由于视频解码相对于编码处理时间要长,为了保证视频显示质量,一般在终端开辟单独的视频解码线程来处理^[12].

3 实时通信系统架构

将 WebRTC 技术应用到医疗领域,用来实现医生和出院患者之间的实时音视频通信,并结合患者住院信息建立诊后医疗问诊系统,本文主要对实时通信模块进行研究与实现.

3.1 实时通信架构设计

诊后医疗系统的实时通信平台的主要功能是实现患者与医生在线实时问诊,实时通信平台的主要架构如图 4 所示,在通信架构中主要应用会话控制服务器、B2BUA 信令服务器、SBC 服务器来实现终端实时音视频通信.

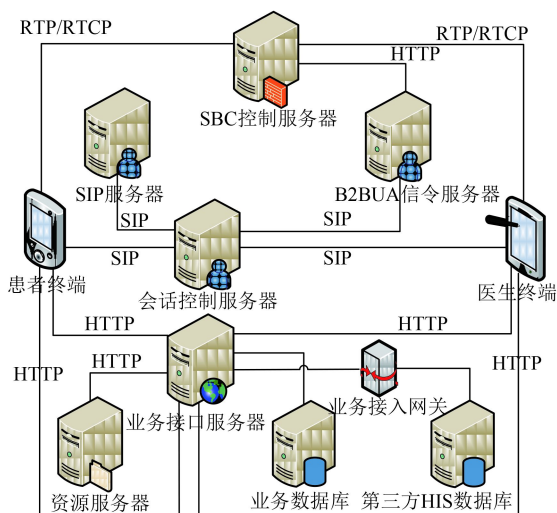


图 4 实时通信架构图

会话控制服务器:负责接收 SIP 信令消息,对其进行处理并最终将 SIP 信令^[13]转发给 B2BUA 服务器.

B2BUA 信令服务器:负责对收到的 SIP 消息解析,SIP 改写、会话保持等.主要是对系统中与 INVITE 相

关的 SIP 消息进行处理,B2BUA 结构的信令服务器在主叫端与被叫端之间分别建立并维持会话^[13],在本实时通信平台中用其实现信令代理功能.

SBC 媒体控制服务器:负责分配通信地址与端口,作为终端会话媒体的中继节点并实现通信媒体备份存档,B2BUA 服务器对其进行控制,以实现网络穿越.

业务接口服务器:为系统提供服务接口控制、实现与医院现有系统的数据对接.

3.2 实时通信服务器设计

通信平台主要设计了会话控制服务器、B2BUA 信令服务器和 SBC 服务器,相互协作完成终端实时通信.

会话控制服务器负责接收终端发送的请求,将请求发送到 SIP、B2BUA 服务器,会话控制服务器接收到 SIP 消息后将消息解析,其中主要包括 REGISTER 注册消息、INVITE 呼叫消息和 OPTION 会话保持消息和通知消息,其中主要的是对 INVITE 呼叫消息进行处理,用以完成终端实时通信的会话协商.SIP 服务器主要实现终端用户的 SIP 注册功能.B2BUA 信令服务器主要是对 SIP 消息进行解析,负责信令代理功能并与终端分别建立并维持会话,与 SBC 服务器协作实现网络穿越.SBC 服务器主要实现媒体代理,完成媒体转接,在终端协商成功后,终端发送实时媒体流到 SBC 服务器进行处理,平台服务器的架构如图 5.

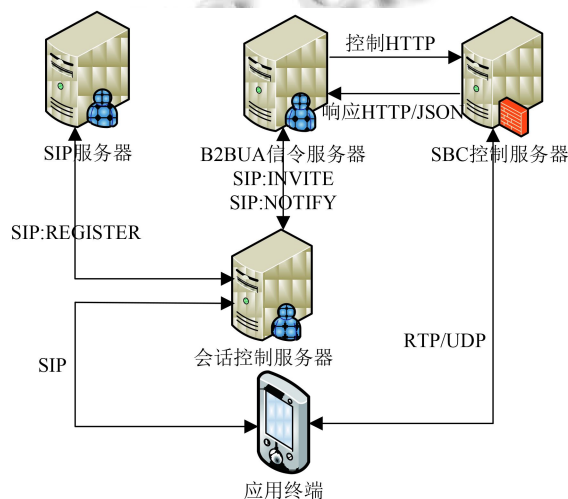


图 5 服务器架构图

3.3 实时音视频通信架构

问诊平台主要是依托实时通信而实现的,用户应用智能终端作为服务入口,通信终端通过业务接口进

行前期的业务绑定,建立医生和患者的对应关系,然后进行终端会话信令协商,最后应用 WebRTC 的音视频控制模块进行媒体数据处理、编码和传输控制等。平

台通过代理服务器完成网络穿越,最终实现终端会话,通信架构分为业务对接、信令处理、音视频控制模块和服务代理模块,各个模块的关系如图6所示。

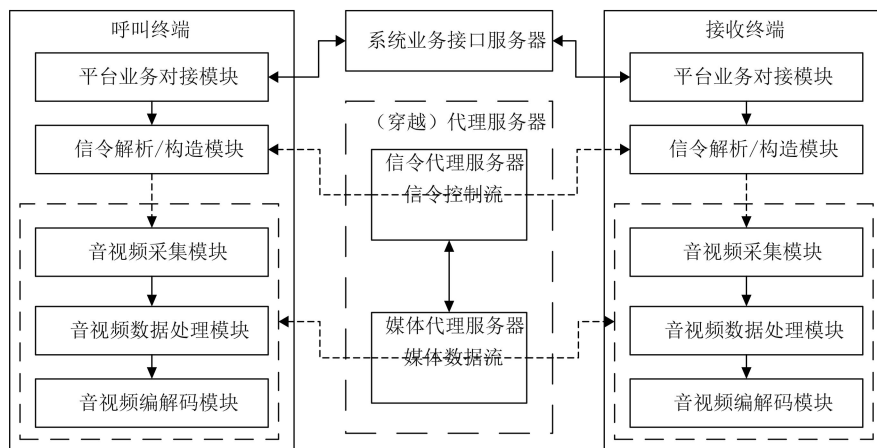


图6 实时通信架构图

3.4 会话流程设计

医生或患者通过智能设备登录终端,调用业务控制接口,查询相关医生、患者及医院信息并进一步进行相应操作。下面是问诊平台会话流程的处理过程。

(1) 出院患者通过智能终端进行登录,首次登录用户需进行医院选择、录入患者住院号和身份证号等,并调用业务接口查询主治医师,建立关系绑定。

(2) 医生登录终端并调用业务接口,查询建立医患绑定的患者基本情况和病历信息。

(3) 患者通过终端与绑定的医生发起问诊邀请消息。

(4) 医生登录终端,选择问诊患者,建立会话呼叫。

(5) 医生终端组装 INVITE 信令,发往会话控制服务器,经分析处理后转接给信令代理服务器处理。

(6) B2BUA 信令服务器接收 INVITE 信令后,向 SBC 服务器发送地址申请并对信令重写,向患者终端发送 INVITE 信令。

(7) 患者终端接收信令后,组装应答信令,发往 B2BUA 信令服务器,经处理后发往医生终端。

(8) 医生终端接收到响应信令,进行信令解析设置,协商完成后,进行媒体通话。

(9) 终端调用 WebRTC 中的音/视频采集组件进行数据采集。

(10) 音频模块设置音频状态、速率、带宽、通道,调用采集设备进行数据采集、数据控制处理,如降噪、回声、静音检测等处理等、数据编码等。

(11) 视频模块设置视频显示参数如,大小、帧率、载荷等,调用采集设备采集视频数据,对数据进行画面处理和数据编码等。

(12) 经过采集编码后的媒体数据,经 SBC 服务器所建立的地址映射关系,建立实时传输。

(13) 患者终端接收到音视频数据后,经 WebRTC 音视频处理模块处理。

(14) 经处理后的媒体数据流调用 WebRTC 解码模块进行解码处理后,调用显示设备呈现给医生。

其中,会话的建立是步骤(1)–(8),会话建立后,双方可同时应用终端向对方发送音视频媒体流,如步骤(10)–(14)。

3.5 系统运行展示

图7是问诊系统医生的排班情况,图8患者与医生在线问诊的运行截图。

4 实现要点分析

4.1 信令代理实现

平台中应用 B2BUA 结构的信令服务器实现信令代理功能,通过 UAS (User Agent Server)、UAC (User Agent Client) 完成信令的接收、修改和转发。以信令协商为例介绍信令代理的实现,信令协商的主要功能是处理呼叫端(医生)UAC 发起的 INVITE 呼叫的相关消息,通过与代理服务器进行交互,最终将信令呼叫消息

发送到被呼叫端(患者)UAS, 以对 INVITE 信令进行处理, 完成会话终端呼叫协商并建立通话, 图 9 是呼叫信令的代理模式处理过程.



图 7 医生就诊时间



图 8 患者在线问诊

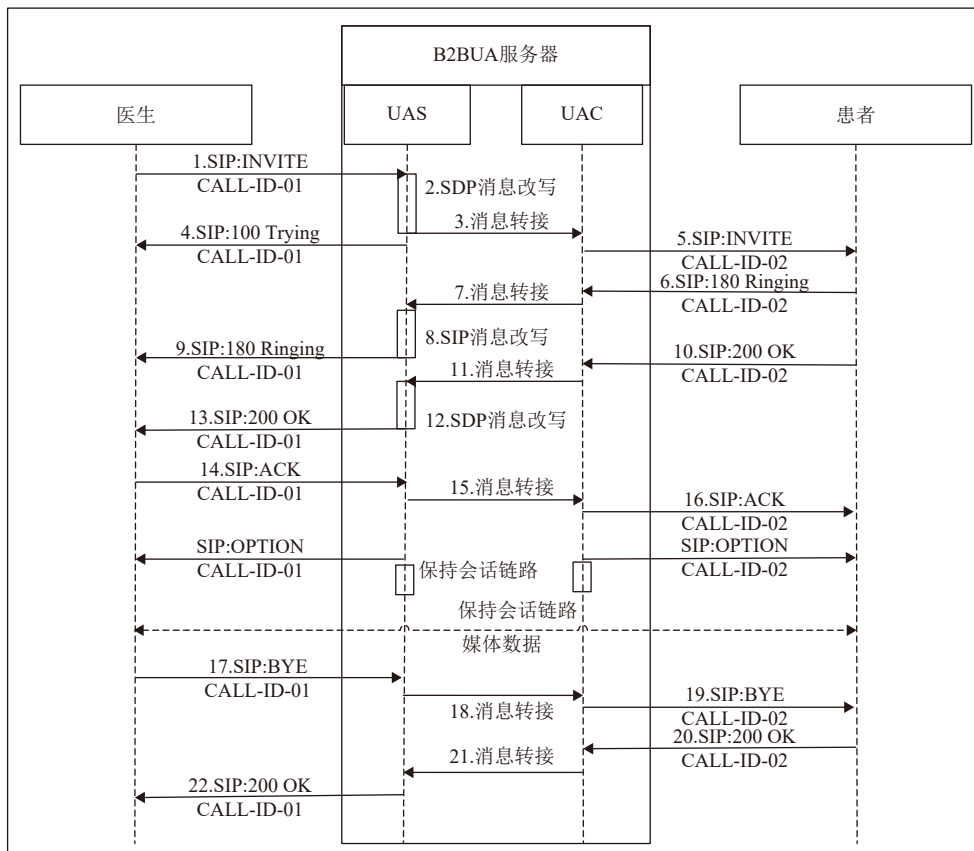


图 9 呼叫信令处理过程

(1) 医生发起呼叫请求, 终端 UAC 向会话管理服务服务器发送 INVITE 信令, 经解析处理后, 将信令转发到 B2BUA 信令服务器。

(2) 代理服务器的 UAS 收到 INVITE 信令后, 向 SBC 服务器发送控制请求, 最终对信令进行改写, 并通过 UAC 发送到患者终端, B2BUA 信令服务器在通信终端分别维持会话。

(3) 患者收到 INVITE 消息后, 构造响应消息如 180 Ringing、200 OK, 发送给信令代理服务器, 控制模块对其改写并发送给医生终端, 完成媒体协商。

4.2 媒体代理实现

本实时通信系统中通过 B2BUA 和 SBC 结构的服务器实现信令和媒体的代理功能, SBC 服务器受信令服务器的控制, 用以处理实时通信的媒体流, 是通信终端媒体数据的中转站, 主要是实现地址映射和通信媒体流的控制。

地址映射: 在信令协商时, 信令代理服模块向 SBC 服务器发送控制请求, 请求为本次会话分配公网地址和端口号, 代理模块分配并存储该地址和端口。在信令协商成功后, 通信双方的实时音视频流会发往本次分配的地址和端口, 终端先进行音频哑包发送, 以打通媒体网络通道, 服务器对到达的实时数据包所处网络建立地址映射, 实现后期通信媒体流的转接。

媒体流控制: 考虑到实时问诊系统的业务背景, 需要对通信数据进行备份转存, 媒体服务器对会话终端所通过的媒体流进行存储备份, 并能对媒体流进行控制, 统计终端通话时长等。

4.3 终端媒体数据处理

在本实时通信终端, 采用 WebRTC 技术来实现终端音视频处理与控制, 在 WebRTC 中对实时音视频的处理是分开的, 各自调用设备采集并通过信道进行传输。但是对于音视频发送流程不同的是, 音频数据需要混音处理操作, 多路视频画面接收后没有类似多路音频混音的操作, 而是分别进行渲染。应用终端音视频处理流程如图 10 所示。

5 结论与展望

基于 WebRTC 的实时音视频通信应用具有良好的跨平台性, 它的提出为全平台的音视频互通提供了技术可行性, 为研究实时音视频解决方案提供了技术基础, 将 WebRTC 技术与移动医疗领域相结合并与现有医院信息系统相对接以设计和实现出针对出院患者

再就医需求的线上问诊的实时通信服务系统, 不仅有效提升了在线问诊的用户体验度和服务质量, 还解决了慢性病患者的长期就医难问题。目前 WebRTC 技术也存在着一些不足, 如对音视频处理的设置参数多数是取自经验值且没有提供多人会话方案, 并且服务器集群管理方面都需要我们对其进行进一步研究以此更好的实现移动医疗领域的实时音视频通信。

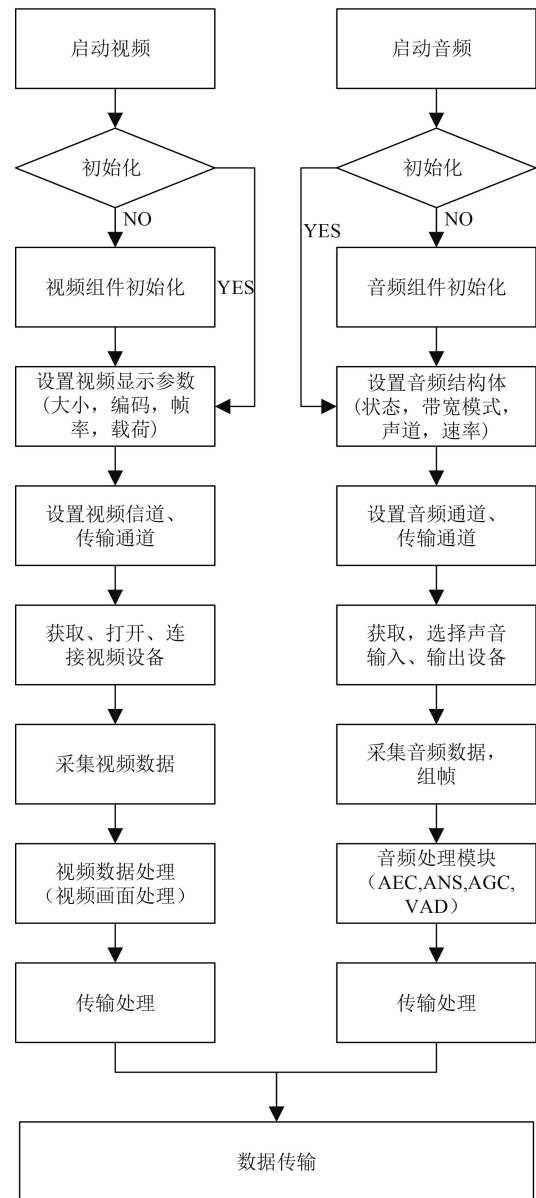


图 10 音视频数据处理

参考文献

- 张向辉, 黄佳庆, 吴康恒, 等. 基于 WebRTC 的实时视音频通信研究综述. 计算机科学, 2015, 42(2): 1-6, 32. [doi:

- 10.11896/j.issn.1002-137X.2015.02.001]
- 2 张永强, 张捍东, 赵金宝. SIP 协议栈研究. 计算机技术与发展, 2007, 17(11): 49–51, 56. [doi: 10.3969/j.issn.1673-629X.2007.11.015]
 - 3 王荣生. SDP 协议在视频点播系统中的应用. 计算机应用与软件, 2005, 22(1): 74–76.
 - 4 潘平. 会话边界控制设备 SBC 应用的相关研究. 广东通信技术, 2010, 30(5): 14–17.
 - 5 林鸿, 王松, 杨鑫, 等. 基于 WebRTC 技术的应用及平台技术开发与设计. 电信科学, 2013, 29(9): 20–25, 36.
 - 6 吴江锐. WebRTC 语音引擎中 NetEQ 技术的研究[硕士学位论文]. 西安: 西安电子科技大学, 2013.
 - 7 Bankoski J, Wilkins P, Xu YW. Technical overview of VP8, an open source video codec for the web. Proceedings of 2011 IEEE International Conference on Multimedia and Expo. Barcelona, Spain. 2011. 1–6.
 - 8 朱光, 张云华, 卢娟. 基于 ICE 的 VOIP 穿越 NAT 方案的研究. 计算机应用与软件, 2011, 28(10): 222–224, 234. [doi: 10.3969/j.issn.1000-386X.2011.10.064]
 - 9 Amirante A, Castaldi T, Miniero L, *et al.* On the seamless interaction between WebRTC browsers and SIP-based conferencing systems. IEEE Communications Magazine, 2013, 51(4): 42–47. [doi: 10.1109/MCOM.2013.6495759]
 - 10 刘亚杰, 王晖, 郭波. P2P 流媒体数据调度研究综述. 计算机应用, 2008, 28(4): 829–831, 848.
 - 11 王亚辉. 基于 WebRTC 语音引擎的会议混音技术研究[硕士学位论文]. 西安: 西安电子科技大学, 2013.
 - 12 熊雨新. 基于 WebRTC 引擎的音频视频交互系统设计与实现[硕士学位论文]. 成都: 电子科技大学, 2014.
 - 13 雷为民, 李伟. 基于 SIP B2BUA 的媒体服务器的集成. 小型微型计算机系统, 2008, 29(11): 2060–2064.