



(12) 发明专利

(10) 授权公告号 CN 111048100 B

(45) 授权公告日 2023. 09. 08

(21) 申请号 201911146300.X

(22) 申请日 2019.11.21

(65) 同一申请的已公布的文献号
申请公布号 CN 111048100 A

(43) 申请公布日 2020.04.21

(73) 专利权人 深圳市东进银通电子有限公司
地址 518000 广东省深圳市福田区沙头街
道泰然六路苍松大厦2101A室

(72) 发明人 符运实 贺建楠

(74) 专利代理机构 深圳市顺天达专利商标代理
有限公司 44217
专利代理师 郭伟刚 李闯

(56) 对比文件

- CN 101447188 A, 2009.06.03
- JP 2007052496 A, 2007.03.01
- US 2013289989 A1, 2013.10.31
- WO 2017012496 A1, 2017.01.26
- WO 2018149077 A1, 2018.08.23
- CN 101740031 A, 2010.06.16
- US 2014278416 A1, 2014.09.18
- CN 1556522 A, 2004.12.22
- CN 106209786 A, 2016.12.07
- US 2003009333 A1, 2003.01.09
- US 2016035349 A1, 2016.02.04

审查员 张少文

(51) Int. Cl.

G10L 17/10 (2013.01)

H04L 67/1004 (2022.01)

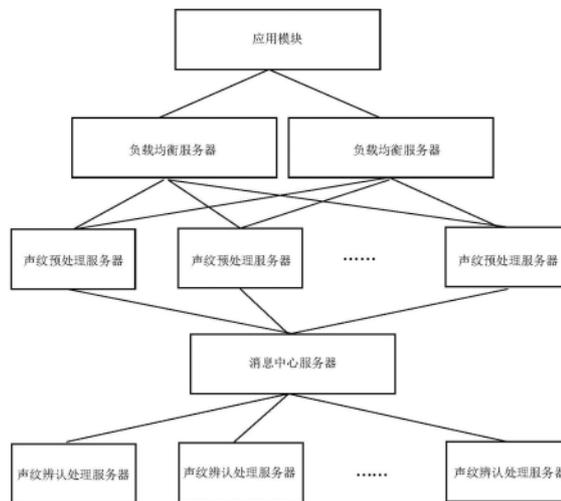
权利要求书2页 说明书5页 附图2页

(54) 发明名称

一种大数据并行化声纹辨认系统和方法

(57) 摘要

本发明公开了一种大数据并行化声纹辨认系统和方法。所述系统包括应用模块、多个负载均衡服务器、多个声纹预处理服务器、多个声纹辨认处理服务器和消息中心服务器；每一个所述负载均衡服务器均与所述应用模块通信连接，每一个所述声纹预处理服务器均与多个所述负载均衡服务器和所述消息中心服务器通信连接，每一个所述声纹辨认处理服务器均与所述消息中心服务器通信连接。所述大数据并行化声纹辨认系统和方法均可以能很好解决声纹辨认系统中大容量扩容难，大并发量提升难，识别响应时间长和可靠性不高的四大难题。



1. 一种大数据并行化声纹辨认系统,其特征在于,所述系统包括应用模块、多个负载均衡服务器、多个声纹预处理服务器、多个声纹辨认处理服务器和消息中心服务器;每一个所述负载均衡服务器均与所述应用模块通信连接,每一个所述声纹预处理服务器均与多个所述负载均衡服务器和所述消息中心服务器通信连接,每一个所述声纹辨认处理服务器均与所述消息中心服务器通信连接;其中,

所述应用模块用于供用户采集待辨认声纹语音数据并发起声纹辨认指令;

所述负载均衡服务器用于甄选出最优的所述声纹预处理服务器,并将所述待辨认声纹语音数据和声纹辨认指令转发到所述最优的所述声纹预处理服务器;

最优的所述声纹预处理服务器用于提取出所述待辨认声纹语音数据的声纹特征,将所述声纹特征通过所述消息中心服务器广播到所有的所述声纹辨认处理服务器;

所述声纹辨认处理服务器用于对接收到的所述声纹特征进行辨认,并将辨认结果通过所述消息中心服务器反馈至最优的所述声纹预处理服务器;

最优的所述声纹预处理服务器还用于根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果,并将所述最终辨认结果通过所述负载均衡服务器反馈至所述应用模块;

每一个所述声纹辨认处理服务器均包括存储有多个声纹模型的分片声纹库,不同的所述声纹辨认处理服务器的分片声纹库所存储的声纹模型均不相同;

采用多个声纹辨认处理服务器形成声纹辨认集群,多个声纹辨认处理服务器的分片声纹库构成一个总声纹库。

2. 根据权利要求1所述的大数据并行化声纹辨认系统,其特征在于,所述辨认结果是:所述声纹辨认处理服务器将所述声纹特征与自有的所述分片声纹库中的所有的声纹模型进行相似度匹配计算后得出的相似度分值最高的一组有序序列结果子集。

3. 根据权利要求2所述的大数据并行化声纹辨认系统,其特征在于,所述最终辨认结果是:所述最优的所述声纹预处理服务器将各个所述声纹辨认处理服务器的有序序列结果子集按照高分优先序列进行重组,计算每一个相似度分值,再度进行全局排序,得出的最终有序结果子集。

4. 一种大数据并行化声纹辨认方法,其特征在于,所述方法通过如权利要求2-3中任一项所述的大数据并行化声纹辨认系统实施,所述方法包括如下步骤:

步骤S1,应用模块采集待辨认声纹语音数据并发起声纹辨认指令;

步骤S2,负载均衡服务器甄选出最优的所述声纹预处理服务器,并将所述待辨认声纹语音数据和声纹辨认指令转发到所述最优的所述声纹预处理服务器;

步骤S3,最优的所述声纹预处理服务器提取出所述待辨认声纹语音数据的声纹特征,将所述声纹特征通过所述消息中心服务器广播到所有的所述声纹辨认处理服务器;

步骤S4,声纹辨认处理服务器对接收到的所述声纹特征进行辨认,并将辨认结果通过所述消息中心服务器反馈至最优的所述声纹预处理服务器;

步骤S5,最优的所述声纹预处理服务器根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果,并将所述最终辨认结果通过所述负载均衡服务器反馈至所述应用模块。

5. 根据权利要求4所述的大数据并行化声纹辨认方法,其特征在于,每一个所述声纹辨

认处理服务器均包括存储有多个声纹模型的分片声纹库,不同的所述声纹辨认处理服务器的分片声纹库所存储的声纹模型均不相同。

6.根据权利要求5所述的大数据并行化声纹辨认方法,其特征在于,在所述步骤S4中,所述的“声纹辨认处理服务器对接收到的所述声纹特征进行辨认”的过程是:所述声纹辨认处理服务器将所述声纹特征与自有的所述分片声纹库中的所有的声纹模型进行相似度匹配计算后得出相似度分值最高的一组有序序列结果子集;所述有序序列结果子集即为所述辨认结果。

7.根据权利要求6所述的大数据并行化声纹辨认方法,其特征在于,在所述步骤S5中,所述的“最优的所述声纹预处理服务器根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果”的过程是:所述最优的所述声纹预处理服务器将各个所述声纹辨认处理服务器的有序序列结果子集按照高分优先序列进行重组,计算每一个相似度分值,再度进行全局排序,得出最终有序结果子集;所述最终有序结果子集即为所述最终辨认结果。

一种大数据并行化声纹辨认系统和方法

技术领域

[0001] 本发明涉及声纹识别领域,具体涉及一种大数据并行化声纹辨认系统和方法。

背景技术

[0002] 声纹识别,生物识别技术的一种,也称为说话人识别,包括说话人辨认和说话人确认。声纹识别就是把声信号转换成电信号,再用计算机进行识别。不同的任务和应用会使用不同的声纹识别技术,如缩小刑侦范围时可能需要辨认技术,而银行交易时则需要确认技术。

[0003] 申请号为CN201010046573.X的中国发明专利公开了一种基于网络动态负载均衡的声纹识别系统,包括用于对用户发出的声音进行采集的声纹客户端,声纹客户端与负载均衡服务器通讯,用于查找最优声纹识别服务器的负载均衡服务器与声纹识别服务器组通讯,声纹识别服务器组由至少两个声纹识别服务器组成,最优声纹识别服务器将识别结果反馈至声纹客户端。在进行语音鉴别(辨认)时,最优声纹识别服务器调用接口模块的鉴别功能,语音处理模块对语音进行预处理,对预处理后的语音按照一定的算法提取特征,将该提取特征与所要鉴别的所有用户的模型进行一一比对,并计算得分排序,若第一名得分高于设定的阈值,则说话人就是第一名的用户,否则,则说明该语音不属于用户中的任何一个人。

[0004] 上述专利的声纹识别系统问题在于,其在进行语音鉴别时,虽然利用了负载均衡服务器查找最优声纹识别服务器实现自动对流量和声纹识别服务器集群进行动态分配,但是,对声纹进行预处理、提取特征、模型比对、算分排序的过程仍然是通过单个的声纹识别服务器来完成的。导致上述专利的声纹识别系统在对大容量、高并发的声纹辨认(也称鉴别)时还是有可能出现辨认响应时间长的问题。

发明内容

[0005] 本发明所要解决的技术问题是提供一种大数据并行化声纹辨认系统,其可以能很好解决声纹辨认系统中大容量扩容难,大并发量提升难,识别响应时间长和可靠性不高的四大难题。

[0006] 本发明所要解决的另一个技术问题是提供一种大数据并行化声纹辨认方法。

[0007] 为解决上述技术问题,本发明所采用的技术方案是提供一种大数据并行化声纹辨认系统,所述系统包括应用模块、多个负载均衡服务器、多个声纹预处理服务器、多个声纹辨认处理服务器和消息中心服务器;每一个所述负载均衡服务器均与所述应用模块通信连接,每一个所述声纹预处理服务器均与多个所述负载均衡服务器和所述消息中心服务器通信连接,每一个所述声纹辨认处理服务器均与所述消息中心服务器通信连接;其中,

[0008] 所述应用模块用于供用户采集待辨认声纹语音数据并发起声纹辨认指令;

[0009] 所述负载均衡服务器用于甄选出最优的所述声纹预处理服务器,并将所述待辨认声纹语音数据和声纹辨认指令转发到所述最优的所述声纹预处理服务器;

[0010] 最优的所述声纹预处理服务器用于提取出所述待辨认声纹语音数据的声纹特征，将所述声纹特征通过所述消息中心服务器广播到所有的所述声纹辨认处理服务器；

[0011] 所述声纹辨认处理服务器用于对接收到的所述声纹特征进行辨认，并将辨认结果通过所述消息中心服务器反馈至最优的所述声纹预处理服务器；

[0012] 最优的所述声纹预处理服务器还用于根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果，并将所述最终辨认结果通过所述负载均衡服务器反馈至所述应用模块。

[0013] 在本发明提供的大数据并行化声纹辨认系统中，每一个所述声纹辨认处理服务器均包括存储有多个声纹模型的分片声纹库，不同的所述声纹辨认处理服务器的分片声纹库所存储的声纹模型均不相同。

[0014] 在本发明提供的大数据并行化声纹辨认系统中，所述辨认结果是：所述声纹辨认处理服务器将所述声纹特征与自有的所述分片声纹库中的所有的声纹模型进行相似度匹配计算后得出的相似度分值最高的一组有序序列结果子集。

[0015] 在本发明提供的大数据并行化声纹辨认系统中，所述最终辨认结果是：所述最优的所述声纹预处理服务器将各个所述声纹辨认处理服务器的有序序列结果子集按照高分优先序列进行重组，计算每一个相似度分值，再度进行全局排序，得出的最终有序结果子集。

[0016] 为解决上述另一个技术问题，本发明所采用的技术方案是提供一种大数据并行化声纹辨认方法，所述方法包括如下步骤：

[0017] 步骤S1，应用模块采集待辨认声纹语音数据并发起声纹辨认指令；

[0018] 步骤S2，负载均衡服务器甄选出最优的所述声纹预处理服务器，并将所述待辨认声纹语音数据和声纹辨认指令转发到所述最优的所述声纹预处理服务器；

[0019] 步骤S3，最优的所述声纹预处理服务器提取出所述待辨认声纹语音数据的声纹特征，将所述声纹特征通过所述消息中心服务器广播到所有的所述声纹辨认处理服务器；

[0020] 步骤S4，声纹辨认处理服务器对接收到的所述声纹特征进行辨认，并将辨认结果通过所述消息中心服务器反馈至最优的所述声纹预处理服务器；

[0021] 步骤S5，最优的所述声纹预处理服务器根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果，并将所述最终辨认结果通过所述负载均衡服务器反馈至所述应用模块。

[0022] 在本发明提供的大数据并行化声纹辨认方法中，每一个所述声纹辨认处理服务器均包括存储有多个声纹模型的分片声纹库，不同的所述声纹辨认处理服务器的分片声纹库所存储的声纹模型均不相同。

[0023] 在本发明提供的大数据并行化声纹辨认方法中，在所述步骤S4中，所述的“声纹辨认处理服务器对接收到的所述声纹特征进行辨认”的过程是：所述声纹辨认处理服务器将所述声纹特征与自有的所述分片声纹库中的所有的声纹模型进行相似度匹配计算后得出相似度分值最高的一组有序序列结果子集；所述有序序列结果子集即为所述辨认结果。

[0024] 在本发明提供的大数据并行化声纹辨认方法中，在所述步骤S5中，所述的“最优的所述声纹预处理服务器根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果”的过程是：所述最优的所述声纹预处理服务器将各个所述声纹辨认处理服务器

的有序序列结果子集按照高分优先序列进行重组,计算每一个相似度分值,再度进行全局排序,得出最终有序结果子集;所述最终有序结果子集即为所述最终辨认结果。

[0025] 实施本发明提供的大数据并行化声纹辨认系统和方法,可以达到以下有益效果:

[0026] 1、采用负载均衡方案将计算耗时最长的提取待辨认声纹语音数据的声纹特征的环节均匀分布于多个声纹预处理服务器上,能应付应用模块较大的并发量变化。

[0027] 2、采用消息中心服务器可以选择成熟的消息中心方案,提高成熟度,降低多个声纹预处理服务器和多个声纹辨认处理服务器之间的耦合度,可以大大简化系统架构的设计难度。

[0028] 3、采用多个声纹辨认处理服务器形成声纹辨认集群,多个声纹辨认处理服务器的分片声纹库构成一个总声纹库,使得总声纹库的规模数量不再受限于单个声纹辨认处理服务器的资源约束,扩容时只要相应增加声纹辨认处理服务器即可。

[0029] 总而言之,采用上述的大数据并行化声纹辨认系统和方法能很好解决现有声纹辨认系统中大容量扩容难,大并发量提升难,识别响应时间长和可靠性不高的四大难题。

附图说明

[0030] 图1为本发明实施例一提供的大数据并行化声纹辨认系统的结构示意图;

[0031] 图2为本发明实施例二提供的大数据并行化声纹辨认方法的流程示意图。

具体实施方式

[0032] 为了对本发明的技术特征、目的和效果有更加清楚的理解,现对照附图详细说明本发明的具体实施方式。显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0033] 实施例一

[0034] 本实施例提供了一种大数据并行化声纹辨认系统。所述系统包括应用模块、多个负载均衡服务器、多个声纹预处理服务器、多个声纹辨认处理服务器和消息中心服务器;每一个所述负载均衡服务器均与所述应用模块通信连接,每一个所述声纹预处理服务器均与多个所述负载均衡服务器和所述消息中心服务器通信连接,每一个所述声纹辨认处理服务器均与所述消息中心服务器通信连接;每一个所述声纹辨认处理服务器均包括存储有多个声纹模型的分片声纹库,不同的所述声纹辨认处理服务器的分片声纹库所存储的声纹模型均不相同,多个声纹辨认处理服务器的分片声纹库构成一个总声纹库。

[0035] 本实施例中,所述应用模块用于供用户采集待辨认声纹语音数据并发起声纹辨认指令。具体的,所述应用模块可以是计算机、智能手机等,其可通过实时采集语音或者语音文件回放方式发起声纹辨认处理请求(声纹辨认指令),并将待辨认声纹语音数据和相关参数发往一个所述负载均衡处理服务器。

[0036] 本实施例中,所述负载均衡服务器用于甄选出最优的所述声纹预处理服务器,并将所述待辨认声纹语音数据和声纹辨认指令转发到所述最优的所述声纹预处理服务器。具体的,所述负载均衡服务器根据最优规则于甄选出最优的所述声纹预处理服务器,最优规则是指从多个所述声纹预处理服务器筛选出最空闲的所述声纹预处理服务器作为最优的

所述声纹预处理服务器。

[0037] 本实施例中,最优的所述声纹预处理服务器用于提取出所述待辨认声纹语音数据的声纹特征,将所述声纹特征通过所述消息中心服务器广播到所有的所述声纹辨认处理服务器。具体的,最优的所述声纹预处理服务器对所述待辨认声纹语音数据进行预处理后,提取出所述待辨认声纹语音数据的声纹特征,将所述待辨认声纹语音数据的声纹特征打包成下一级声纹辨认分片处理指令包,并发往所述消息中心服务器,然后所述消息中心服务器将所述声纹辨认分片处理指令包广播到所有的所述声纹辨认处理服务器。

[0038] 本实施例中,所述声纹辨认处理服务器用于对接收到的所述声纹特征进行辨认,并将辨认结果通过所述消息中心服务器反馈至最优的所述声纹预处理服务器。具体的,所述声纹辨认处理服务器将所述声纹特征与自有的所述分片声纹库中的所有的声纹模型进行相似度匹配计算后得出相似度分值最高的一组有序序列结果子集,所述有序序列结果子集即为所述辨认结果。

[0039] 本实施例中,最优的所述声纹预处理服务器还用于根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果,并将所述最终辨认结果通过所述负载均衡服务器反馈至所述应用模块。具体的,所述最优的所述声纹预处理服务器将各个所述声纹辨认处理服务器的有序序列结果子集按照高分优先序列进行重组,计算每一个相似度分值,再度进行全局排序,得出最终有序结果子集,所述最终有序结果子集即为所述最终辨认结果。

[0040] 实施例二

[0041] 本实施例提供了一种大数据并行化声纹辨认方法,所述方法通过实施一提供的系统实施。所述方法包括如下步骤:

[0042] 步骤S1,应用模块采集待辨认声纹语音数据并发起声纹辨认指令;具体的,所述应用模块可以是计算机、智能手机等,其可通过实时采集语音或者语音文件回放方式发起声纹辨认处理请求(声纹辨认指令),并将待辨认声纹语音数据和相关参数发往一个所述负载均衡处理服务器。

[0043] 步骤S2,负载均衡服务器甄选出最优的所述声纹预处理服务器,并将所述待辨认声纹语音数据和声纹辨认指令转发到所述最优的所述声纹预处理服务器;具体的,所述负载均衡服务器根据最优规则于甄选出最优的所述声纹预处理服务器,最优规则是指从多个所述声纹预处理服务器筛选出最空闲的所述声纹预处理服务器作为最优的所述声纹预处理服务器。

[0044] 步骤S3,最优的所述声纹预处理服务器提取出所述待辨认声纹语音数据的声纹特征,将所述声纹特征通过所述消息中心服务器广播到所有的所述声纹辨认处理服务器;具体的,最优的所述声纹预处理服务器对所述待辨认声纹语音数据进行预处理后,提取出所述待辨认声纹语音数据的声纹特征,将所述待辨认声纹语音数据的声纹特征打包成下一级声纹辨认分片处理指令包,并发往所述消息中心服务器,然后所述消息中心服务器将所述声纹辨认分片处理指令包广播到所有的所述声纹辨认处理服务器。

[0045] 步骤S4,声纹辨认处理服务器对接收到的所述声纹特征进行辨认,并将辨认结果通过所述消息中心服务器反馈至最优的所述声纹预处理服务器;具体的,所述的“声纹辨认处理服务器对接收到的所述声纹特征进行辨认”的过程是:所述声纹辨认处理服务器将所

述声纹特征与自有的所述分片声纹库中的所有的声纹模型进行相似度匹配计算后得出相似度分值最高的一组有序序列结果子集;所述有序序列结果子集即为所述辨认结果。

[0046] 步骤S5,最优的所述声纹预处理服务器根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果,并将所述最终辨认结果通过所述负载均衡服务器反馈至所述应用模块。具体的,所述的“最优的所述声纹预处理服务器根据多个所述声纹辨认处理服务器反馈的所述辨认结果得出最终辨认结果”的过程是:所述最优的所述声纹预处理服务器将各个所述声纹辨认处理服务器的有序序列结果子集按照高分优先序列进行重组,计算每一个相似度分值,再度进行全局排序,得出最终有序结果子集;所述最终有序结果子集即为所述最终辨认结果。

[0047] 综上所述,实施本发明提供的大数据并行化声纹辨认系统和方法,可以达到以下有益效果:

[0048] 1、采用负载均衡方案将计算耗时最长的提取待辨认声纹语音数据的声纹特征的环节均匀分布于多个声纹预处理服务器上,能应付应用模块较大的并发量变化。

[0049] 2、采用消息中心服务器可以选择成熟的消息中心方案,提高成熟度,降低多个声纹预处理服务器和多个声纹辨认处理服务器之间的耦合度,可以大大简化系统架构的设计难度。

[0050] 3、采用多个声纹辨认处理服务器形成声纹辨认集群,多个声纹辨认处理服务器的分片声纹库构成一个总声纹库,使得总声纹库的规模数量不再受限于单个声纹辨认处理服务器的资源约束,扩容时只要相应增加声纹辨认处理服务器即可。

[0051] 总而言之,采用上述的大数据并行化声纹辨认系统和方法能很好解决现有声纹辨认系统中大容量扩容难,大并发量提升难,识别响应时间长和可靠性不高的四大难题。

[0052] 上面结合附图对本发明的实施例进行了描述,但是本发明并不局限于上述的具体实施方式,上述的具体实施方式仅仅是示意性的,而不是限制性的,本领域的普通技术人员在本发明的启示下,在不脱离本发明宗旨和权利要求所保护的范围情况下,还可做出很多形式,这些均属于本发明的保护之内。

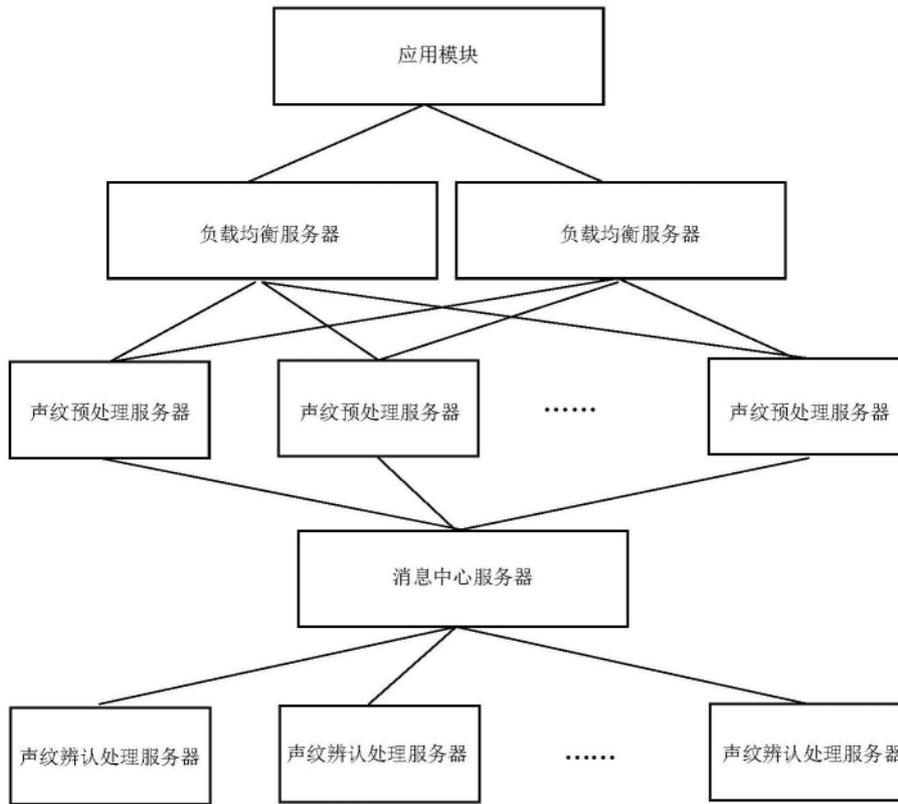


图1

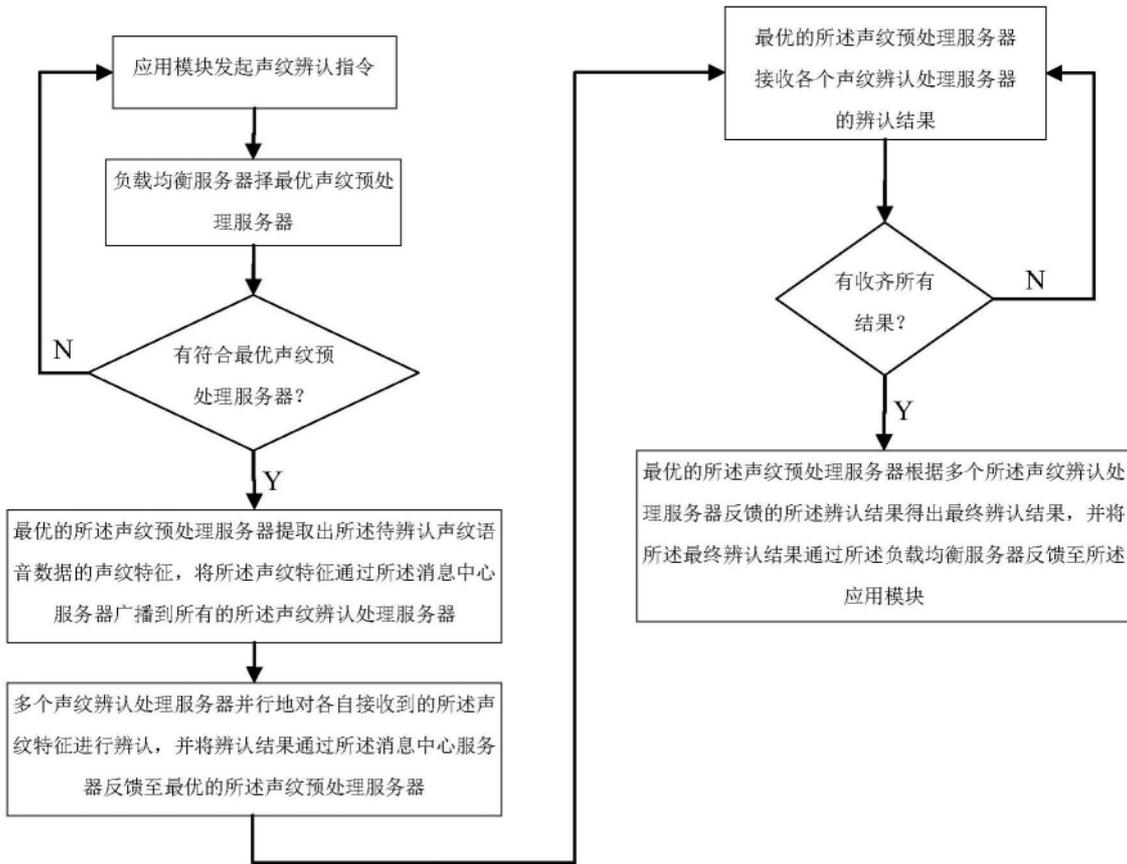


图2