

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G10L 19/00 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200580040670.5

[43] 公开日 2007年10月31日

[11] 公开号 CN 101065797A

[22] 申请日 2005.10.28

[21] 申请号 200580040670.5

[30] 优先权

[32] 2004.10.28 [33] US [31] 60/622,922

[32] 2004.10.28 [33] US [31] 10/975,841

[86] 国际申请 PCT/US2005/038961 2005.10.28

[87] 国际公布 WO2006/050112 英 2006.5.11

[85] 进入国家阶段日期 2007.5.28

[71] 申请人 诺伊拉尔音频公司

地址 美国华盛顿州

[72] 发明人 罗伯特·W·里姆斯

杰弗里·K·托姆普森

阿伦·瓦尔纳

[74] 专利代理机构 北京集佳知识产权代理有限公司  
代理人 杨生平 朱 胜

权利要求书 15 页 说明书 19 页 附图 11 页

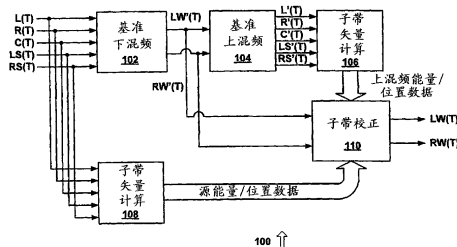
[54] 发明名称

音频空间环境引擎

[57] 摘要

提供了一种用于在音频数据的不同格式之间进行转换的音频空间环境引擎。音频空间环境引擎(100)允许在N声道数据和M声道数据之间进行灵活转换,并且允许从M声道数据转换回N'声道数据,其中,N、M和N'为整数,并且N不必等于N'。例如,这样的系统能够在整个为立体声数据设计的基础设施或网络上传输或存储环绕声音数据。音频空间环境引擎由于先进的动态下混频单元(102)和高分辨率频带上混频单元(104)而在不同的空间环境之间提供了改进和灵活的转换。动态下混频单元包括智能分析和校正环(108,110),其能够校正许多下混频方法所共有的频谱、时间和空间的不准确性。上混频单元利用在整个高分辨率频带上提取并分析重要的声道间的空间提示来得到不同频率元素的空间布置。当下混频和上混频单元单独或

作为系统使用时,提供了改进的声音质量和空间区别。



1. 一种用于从 N 声道音频系统向 M 声道音频系统转换的音频空间环境引擎，其中，M 和 N 为整数，并且，N 大于 M，包括：

基准下混频器，其接收 N 个声道的音频数据并将所述 N 个声道的音频数据转换成 M 个声道的音频数据；

基准上混频器，其接收所述 M 个声道的音频数据并将所述 M 个声道的音频数据转换成 N' 个声道的音频数据；以及

校正系统，其接收所述 M 个声道的音频数据、所述 N 个声道的音频数据和所述 N' 个声道的音频数据，并且基于所述 N 个声道的音频数据和所述 N' 个声道的音频数据之间的差来校正所述 M 个声道的音频数据。

2. 如权利要求 1 所述的系统，其中，所述校正系统进一步包括：

第一子带矢量校准单元，其接收所述 N 个声道的音频数据，并且生成第一多个子带的音频空间图像数据；

第二子带矢量校准单元，其接收所述 N' 个声道的音频数据，并且生成第二多个子带的音频空间图像数据；以及

所述校正系统接收所述第一多个子带的音频空间图像数据和所述第二多个子带的音频空间图像数据，并且基于所述第一多个子带的音频空间图像数据和所述第二多个子带的音频空间图像数据之间的差来校正所述 M 个声道的音频数据。

3. 如权利要求 2 所述的系统，其中，所述第一多个子带的音频空间图像数据和所述第二多个子带的音频空间图像数据中的每一个具有相关的能量值和位置值。

4. 如权利要求 3 所述的系统，其中，所述位置值中的每一个表示二维空间中音频空间图像数据的相关子带的中心的视在定位，其中所述中心的坐标由与 N 个扬声器中的每一个相关能量值和所述 N 个扬声器中的每一个的坐标的矢量和确定。

5. 如权利要求 1 所述的系统，其中，所述基准下混频器进一步包括多个分数希尔伯特级，其中每个接收所述 N 个声道的音频数据中的一个，并且向相关声道的音频数据施加预定的相移。

6. 如权利要求 5 所述的系统，其中，所述基准下混频器进一步包括

多个求和级，其被耦合到所述多个分数希尔伯特级，并且以预定方式组合来自所述希尔伯特级的输出以生成所述M个声道的音频数据。

7. 如权利要求1所述的系统，其中，所述基准上混频器进一步包括：

时域到频域转换级，其接收所述M个声道的音频数据，并且生成多个子带的音频空间图像数据；

滤波发生器，其接收所述M个声道的多个子带的音频空间图像数据，并且生成N'个声道的多个子带的音频空间图像数据；

平滑级，其接收所述N'个声道的所述多个子带的音频空间图像数据，并且平均化每个子带与一个或多个相邻子带；

求和级，其被耦合到所述平滑级，并且接收所述M个声道的多个子带的音频空间图像数据和所平滑的N'个声道的所述多个子带的音频空间图像数据，并且生成经缩放的N'个声道的所述多个子带的音频空间图像数据；以及

频域到时域转换级，其接收所述缩放的N'个声道的所述多个子带的音频空间图像数据，并且生成所述N'个声道的音频数据。

8. 如权利要求1所述的系统，其中，所述校正系统进一步包括：

第一子带矢量校准级，其包括：

时域到频域转换级，其接收所述N个声道的音频数据，并且生成第一多个子带的音频空间图像数据；

第一子带能量级，其接收所述第一多个子带的音频空间图像数据，并且生成用于每个子带的第一能量值；以及

第一子带位置级，其接收所述第一多个子带的音频空间图像数据，并且生成用于每个子带的第一位置矢量。

9. 如权利要求8所述的系统，其中，所述校正系统进一步包括：

第二子带矢量校准级，其包括：

第二子带能量级，其接收第二多个子带的音频空间图像数据，并且生成用于每个子带的第二能量值；以及

第二子带位置级，其接收所述第二多个子带的音频空间图像数据，并且生成用于每个子带的第二位置矢量。

10. 一种用于从N声道音频系统向M声道音频系统转换的方法，其中，

N 和 M 为整数，并且，N 大于 M，包括：

将 N 个声道的音频数据转换成 M 个声道的音频数据；

将所述 M 个声道的音频数据转换成 N' 个声道的音频数据；以及

基于所述 N 个声道的音频数据和所述 N' 个声道的音频数据之间的差来校正所述 M 个声道的音频数据。

11. 如权利要求 10 所述的方法，其中，将所述 N 个声道的音频数据转换成所述 M 个声道的音频数据包括：

用分数希尔伯特函数处理所述 N 个声道的音频数据中的一个或多个，以向相关声道的音频数据施加预定相移；以及

在用所述分数希尔伯特函数处理后，组合所述 N 个声道的音频数据中的一个或多个以产生所述 M 个声道的音频数据，以便所述 M 个声道的音频数据中的每一个中的所述 N 个声道的音频数据中的一个或多个的所述组合具有预定的相位关系。

12. 如权利要求 10 所述的方法，其中，将所述 M 个声道的音频数据转换成所述 N' 个声道的音频数据包括：

将所述 M 个声道的音频数据从时域转换成频域的多个子带；

滤波所述 M 个声道的所述多个子带以生成 N 个声道的多个子带；

通过平均化每个子带与一个或多个相邻带来平滑所述 N 个声道的多个子带；

使所述 N 个声道的多个子带中的每一个与所述 M 个声道的一个或多个相应子带相乘；以及

将所述 N 个声道的多个子带从所述频域转换到所述时域。

13. 如权利要求 10 所述的方法，其中，基于所述 N 个声道的音频数据和所述 N' 个声道的音频数据之间的差来校正所述 M 个声道的音频数据包括：

为所述 N 个声道的音频数据的多个子带中的每一个确定能量和位置矢量；

为所述 N' 个声道的音频数据的多个子带中的每一个确定能量和位置矢量；以及

如果用于所述 N 个声道的音频数据和所述 N' 个声道的音频数据的相

应子带的所述能量和所述位置矢量的差大于可允许的容限,则校正所述M个声道的音频数据的一个或多个子带。

14. 如权利要求13所述的方法,其中,校正所述M个声道的音频数据的一个或多个子带包括:调整用于所述M个声道的音频数据的所述子带的能量和位置矢量,以便所调整的所述M个声道的音频数据的子带被转换成调整的N'个声道的音频数据,与所述N'个声道的音频数据的多个子带中的每一个的未调整的能量和位置矢量相比,调整的N'个声道的音频数据具有更加接近于所述N个声道的音频数据的子带的所述能量和所述位置矢量的一个或多个子带能量和位置矢量。

15. 一种用于从N声道音频系统向M声道音频系统转换的音频空间环境引擎,其中,M和N为整数,并且,N大于M,包括:

下混频器装置,用于接收N个声道的音频数据,并将所述N个声道的音频数据转换成M个声道的音频数据;

上混频器装置,用于接收所述M个声道的音频数据,并将所述M个声道的音频数据转换成N'个声道的音频数据;以及

校正装置,用于接收所述M个声道的音频数据、所述N个声道的音频数据和所述N'个声道的音频数据,并且基于所述N个声道的音频数据和所述N'个声道的音频数据之间的差来校正所述M个声道的音频数据。

16. 如权利要求15所述的系统,其中,所述校正装置进一步包括:

第一子带矢量校准装置,用于接收所述N个声道的音频数据,并且生成第一多个子带的音频空间图像数据;

第二子带矢量校准装置,用于接收所述N'个声道的音频数据,并且生成第二多个子带的音频空间图像数据;以及

所述校正装置用于接收所述第一多个子带的音频空间图像数据和所述第二多个子带的音频空间图像数据,并且基于所述第一多个子带的音频空间图像数据和所述第二多个子带的音频空间图像数据之间的差来校正所述M个声道的音频数据。

17. 如权利要求15所述的系统,其中,所述下混频器装置进一步包括多个分数希尔伯特装置,用于接收所述N个声道的音频数据中的一个,并且向相关声道的音频数据施加预定相移。

18. 如权利要求15所述的系统,其中,所述上混频器装置进一步包

括:

时域到频域转换装置,用于接收所述M个声道的音频数据,并且生成多个子带的音频空间图像数据;

滤波发生器装置,用于接收所述M个声道的所述多个子带的音频空间图像数据,并且生成N'个声道的多个子带的音频空间图像数据;

平滑装置,用于接收所述N'个声道的所述多个子带的音频空间图像数据,并且平均化每个子带与一个或多个相邻子带;

求和装置,用于接收所述M个声道的所述多个子带的音频空间图像数据和所平滑的N'个声道的所述多个子带的音频空间图像数据,并且生成缩放的N'个声道的所述多个子带的音频空间图像数据;以及

频域到时域转换装置,用于接收所述缩放的N'个声道的多个子带的音频空间图像数据,并且生成所述N'个声道的音频数据。

19. 一种用于从N声道音频系统向M声道音频系统转换的音频空间环境引擎,其中,N和M为整数,并且,N大于M,包括:

一个或多个希尔伯特变换级,其中每个接收所述N个声道的音频数据中的一个,并且向相关声道的音频数据施加预定的相移;

一个或多个常数乘法器级,其中每个接收经所述希尔伯特变换的声道的音频数据中的一个,并且每个生成缩放的希尔伯特变换的声道的音频数据;

一个或多个第一求和级,其中每个接收所述N个声道的音频数据中的所述一个和所述缩放的希尔伯特变换的声道的音频数据,并且每个生成分数希尔伯特声道的音频数据;以及

M个第二求和级,其中每个接收所述分数希尔伯特声道的音频数据中的一个或多个以及所述N个声道的音频数据中的一个或多个,并且组合所述分数希尔伯特声道的音频数据中的所述一个或多个和所述N个声道的音频数据中的所述一个或多个中的每一个,以生成M个声道的音频数据中的一个,其具有每个所述分数希尔伯特声道的音频数据中的所述一个或多个和所述N个声道的音频数据中的所述一个或多个之间的预定相位关系。

20. 如权利要求19所述的音频空间环境引擎,其包括希尔伯特变换级,用于接收左声道的音频数据,其中所述希尔伯特变换的左声道的音频数据乘以常数并被加到所述左声道的音频数据以生成具有预定相移的左

声道的音频数据,所述相移的左声道的音频数据乘以常数并被提供给所述M个第二求和级中的一个或多个。

21. 如权利要求19所述的音频空间环境引擎,其包括希尔伯特变换级,用于接收右声道的音频数据,其中所述希尔伯特变换的右声道的音频数据乘以常数并从所述右声道的音频数据减去以生成具有预定相移的右声道的音频数据,所述相移的右声道的音频数据乘以常数并被提供给所述M个第二求和级中的一个或多个。

22. 如权利要求19所述的音频空间环境引擎,其包括接收左环绕声道的音频数据的希尔伯特变换级和接收右环绕声道的音频数据的希尔伯特变换级,其中所述希尔伯特变换的左环绕声道的音频数据乘以常数并被加到所述希尔伯特变换的右环绕声道的音频数据以生成左右环绕声道的音频数据,所述相移的左右环绕声道的音频数据被提供给所述M个第二求和级中的一个或多个。

23. 如权利要求19所述的音频空间环境引擎,其包括接收右环绕声道的音频数据的希尔伯特变换级和接收左环绕声道的音频数据的希尔伯特变换级,其中所述希尔伯特变换的右环绕声道的音频数据乘以常数并被加到所述希尔伯特变换的左环绕声道的音频数据以生成右左环绕声道的音频数据,所述相移的右左环绕声道的音频数据被提供给所述M个第二求和级中的一个或多个。

24. 如权利要求19所述的音频空间环境引擎,包括:

希尔伯特变换级,其接收左声道的音频数据,其中所述希尔伯特变换的左声道的音频数据乘以常数并被加到所述左声道的音频数据以生成具有预定相移的左声道的音频数据,所述左声道的音频数据乘以常数以生成缩放的左声道的音频数据;

希尔伯特变换级,其接收右声道的音频数据,其中所述希尔伯特变换的右声道的音频数据乘以常数并从所述右声道的音频数据减去以生成具有预定相移的右声道的音频数据,所述右声道的音频数据乘以常数以生成缩放的右声道的音频数据; 以及

接收左环绕声道的音频数据的希尔伯特变换级和接收右环绕声道的音频数据的希尔伯特变换级,其中所述希尔伯特变换的左环绕声道的音频数据乘以常数并被加到所述希尔伯特变换的右环绕声道的音频数据以生成左右环绕声道的音频数据,并且所述希尔伯特变换的右环绕声道的音频

数据乘以常数并被加到所述希尔伯特变换的左环绕声道的音频数据以生成右左环绕声道的音频数据。

25. 如权利要求 24 所述的音频空间环境引擎, 包括:

M 个第二求和级的第一个, 其接收所述缩放的左声道的音频数据、所述右左声道的音频数据和缩放的中心声道的音频数据, 并且相加所述缩放的左声道的音频数据、所述右左声道的音频数据和所述缩放的中心声道的音频数据, 以形成左水印声道的音频数据; 以及

M 个第二求和级的第二个, 其接收所述缩放的右声道的音频数据、所述左右声道的音频数据和所述缩放的中心声道的音频数据, 并且相加所述缩放的声道的音频数据和所述缩放的中心声道的音频数据并从所述和中减去所述左右声道的音频数据, 以形成右水印声道的音频数据。

26. 一种用于从 N 声道音频系统向 M 声道音频系统转换的方法, 其中, N 和 M 为整数, 并且, N 大于 M, 包括:

用分数希尔伯特函数处理所述 N 个声道的音频数据中的一个或多个, 以向相关声道的音频数据施加预定相移; 以及

组合用所述分数希尔伯特函数处理之后的所述 N 个声道的音频数据中的一个或多个以产生所述 M 个声道的音频数据, 以便所述 M 个声道的音频数据中的每一个中的所述 N 个声道的音频数据中的所述一个或多个的所述组合具有预定相位关系。

27. 如权利要求 26 所述的方法, 其中, 用分数希尔伯特函数处理所述 N 个声道的音频数据中的一个或多个包括:

对左声道的音频数据执行希尔伯特变换;

将所述希尔伯特变换的左声道的音频数据乘以常数;

将所述缩放的、希尔伯特变换的左声道的音频数据加到所述左声道的音频数据, 以生成具有预定相移的左声道的音频数据; 以及

将所述相移的左声道的音频数据乘以常数。

28. 如权利要求 26 所述的方法, 其中, 用分数希尔伯特函数处理所述 N 个声道的音频数据中的一个或多个包括:

对右声道的音频数据执行希尔伯特变换;

将所述希尔伯特变换的右声道的音频数据乘以常数;



将所述缩放的、希尔伯特变换的右声道的音频数据从所述右声道的音频数据中减去，以生成具有预定相移的右声道的音频数据；以及

将所述相移的右声道的音频数据乘以常数。

29. 如权利要求 26 所述的方法，其中，用分数希尔伯特函数处理所述 N 个声道的音频数据中的一个或多个包括：

对左环绕声道的音频数据执行希尔伯特变换；

对右环绕声道的音频数据执行希尔伯特变换；

将所述希尔伯特变换的左环绕声道的音频数据乘以常数；以及

将所述缩放的、希尔伯特变换的左环绕声道的音频数据加到所述希尔伯特变换的右环绕声道的音频数据，以生成具有预定相移的左右声道的音频数据。

30. 如权利要求 26 所述的方法，其中，用分数希尔伯特函数处理所述 N 个声道的音频数据中的一个或多个包括：

对左环绕声道的音频数据执行希尔伯特变换；

对右环绕声道的音频数据执行希尔伯特变换；

将所述希尔伯特变换的右环绕声道的音频数据乘以常数；以及

将所述缩放的、希尔伯特变换的右环绕声道的音频数据加到所述希尔伯特变换的左环绕声道的音频数据，以生成具有预定相移的右左声道的音频数据。

31. 如权利要求 26 所述的方法，包括：

对左声道的音频数据执行希尔伯特变换；

将所述希尔伯特变换的左声道的音频数据乘以常数；

将所述缩放的、希尔伯特变换的左声道的音频数据加到所述左声道的音频数据，以生成具有预定相移的左声道的音频数据；

将所述相移的左声道的音频数据乘以常数；

对右声道的音频数据执行希尔伯特变换；

将所述希尔伯特变换的右声道的音频数据乘以常数；

将所述缩放的、希尔伯特变换的右声道的音频数据从所述右声道的音频数据中减去，以生成具有预定相移的右声道的音频数据；

将所述相移的右声道的音频数据乘以常数;

对左环绕声道的音频数据执行希尔伯特变换;

对右环绕声道的音频数据执行希尔伯特变换;

将所述希尔伯特变换的左环绕声道的音频数据乘以常数;

将所述缩放的、希尔伯特变换的左环绕声道的音频数据加到所述希尔伯特变换的右环绕声道的音频数据,以生成具有预定相移的左右声道的音频数据;

将所述希尔伯特变换的右环绕声道的音频数据乘以常数; 以及

将所述缩放的、希尔伯特变换的右环绕声道的音频数据加到所述希尔伯特变换的左环绕声道的音频数据,以生成具有预定相移的右左声道的音频数据。

32. 如权利要求 31 所述的方法, 包括:

对所述缩放的左声道的音频数据、所述右左声道的音频数据和缩放的中心声道的音频数据进行求和, 以形成左水印声道的音频数据; 以及

求和所述缩放的声道的音频数据和所述缩放的中心声道的音频数据并从所述和中减去所述左右声道的音频数据, 以形成右水印声道的音频数据。

33. 一种用于从 N 声道音频系统向 M 声道音频系统转换的音频空间环境引擎, 其中, N 和 M 为整数, 并且, N 大于 M, 包括:

希尔伯特变换装置, 用于接收所述 N 个声道的音频数据中的一个, 并且向相关声道的音频数据施加预定相移;

常数乘法器装置, 用于接收所述希尔伯特变换的声道的音频数据中的一个, 并且生成缩放的、希尔伯特变换的声道的音频数据;

求和装置, 用于接收所述 N 个声道的音频数据中的所述一个和所述缩放的、希尔伯特变换的声道的音频数据, 并且每个生成分数希尔伯特声道的音频数据; 以及

M 个第二求和装置, 用于接收所述分数希尔伯特声道的音频数据中的一个或多个和所述 N 个声道的音频数据中的一个或多个, 并且用于组合所述分数希尔伯特声道的音频数据中的所述一个或多个和所述 N 个声道的音频数据中的所述一个或多个中的每一个, 以生成 M 个声道的音频数据中

的一个,其具有每个所述分数希尔伯特声道的音频数据中的所述一个或多个和所述 N 个声道的音频数据中的所述一个或多个之间的预定相位关系。

34. 如权利要求 33 所述的音频空间环境引擎,包括:

希尔伯特变换装置,用于处理左声道的音频数据;

乘法器装置,用于将所述希尔伯特变换的左声道的音频数据乘以常数;

求和装置,用于将所述缩放的、希尔伯特变换的左声道的音频数据加到所述左声道的音频数据,以生成具有预定相移的左声道的音频数据;以及

乘法器装置,用于将所述相移的左声道的音频数据乘以常数,其中,所述缩放的、相移的左声道的音频数据被提供给所述 M 个第二求和装置中的一个或多个。

35. 如权利要求 33 所述的音频空间环境引擎,包括:

希尔伯特变换装置,用于处理右声道的音频数据;

乘法器装置,用于将所述希尔伯特变换的右声道的音频数据乘以常数;

求和装置,用于将所述缩放的、希尔伯特变换的右声道的音频数据加到所述右声道的音频数据,以生成具有预定相移的右声道的音频数据;以及

乘法器装置,用于将所述相移的右声道的音频数据乘以常数,其中,所述缩放的、相移的右声道的音频数据被提供给所述 M 个第二求和装置中的一个或多个。

36. 如权利要求 33 所述的音频空间环境引擎,包括:

希尔伯特变换装置,用于处理左环绕声道的音频数据;

希尔伯特变换装置,用于处理右环绕声道的音频数据;

乘法器装置,用于将所述希尔伯特变换的左环绕声道的音频数据乘以常数;以及

求和装置,用于将所述缩放的、希尔伯特变换的左环绕声道的音频数据加到所述希尔伯特变换的右环绕声道的音频数据,以生成左右声道的音频数据,其中,所述左右声道的音频数据被提供给所述 M 个第二求和装置

中的一个或多个。

37. 如权利要求 33 所述的音频空间环境引擎, 包括:

希尔伯特变换装置, 用于处理左环绕声道的音频数据;

希尔伯特变换装置, 用于处理右环绕声道的音频数据;

乘法器装置, 用于将所述希尔伯特变换的右环绕声道的音频数据乘以常数; 以及

求和装置, 用于将所述缩放的、希尔伯特变换的右环绕声道的音频数据加到所述希尔伯特变换的左环绕声道的音频数据, 以生成右左声道的音频数据, 其中, 所述右左声道的音频数据被提供给所述 M 个第二求和装置中的一个或多个。

38. 一种用于从 N 声道音频系统向 M 声道音频系统转换的音频空间环境引擎, 其中, N 和 M 为整数, 并且, N 大于 M, 包括:

时域到频域转换级, 其接收 M 个声道的音频数据, 并且生成多个子带的音频空间图像数据;

滤波发生器, 其接收所述 M 个声道的所述多个子带的音频空间图像数据, 并且生成 N' 个声道的多个子带的音频空间图像数据; 以及

求和级, 其耦合到所述滤波发生器, 并且接收所述 M 个声道的多个子带的音频空间图像数据和所述 N' 个声道的所述多个子带的音频空间图像数据, 并且生成缩放的 N' 个声道的所述多个子带的音频空间图像数据。

39. 如权利要求 38 所述的音频空间环境引擎, 进一步包括频域到时域转换级, 其接收所述缩放的 N' 个声道的所述多个子带的音频空间图像数据, 并且生成所述 N' 个声道的音频数据。

40. 如权利要求 38 所述的音频空间环境引擎, 进一步包括:

平滑级, 其耦合到所述滤波发生器, 所述平滑级接收所述 N' 个声道的所述多个子带的音频空间图像数据, 并且平均化每个子带与一个或多个相邻子带; 并且

所述求和级被耦合到所述平滑级, 并且接收所述 M 个声道的所述多个子带的音频空间图像数据和所述平滑的 N' 个声道的所述多个子带的音频空间图像数据, 并且生成缩放的 N' 个声道的所述多个子带的音频空间图像数据。

41. 如权利要求 38 所述的音频空间环境引擎, 其中, 所述求和级进一步包括左声道求和级, 其使所述 M 个声道中的左声道的多个子带中的每一个乘以所述 N' 个声道中的左声道的音频空间图像数据的相应多个子带中的每一个。

42. 如权利要求 38 所述的音频空间环境引擎, 其中, 所述求和级进一步包括右声道求和级, 其使所述 M 个声道中的右声道的多个子带中的每一个乘以所述 N' 个声道中的右声道的相应多个子带的音频空间图像数据中的每一个。

43. 如权利要求 38 所述的音频空间环境引擎, 其中, 所述求和级进一步包括中心声道求和级, 其对于每个子带满足以下方程:

$$(G_c(f) * L(f) + ((1 - G_c(f)) * R(f)) * H_c(f))$$

其中

$G_c(f)$  = 中心声道子带缩放因数;

$L(f)$  = 所述 M 个声道中的左声道子带;

$R(f)$  = 所述 M 个声道中的右声道子带; 以及

$H_c(f)$  = 所述 N' 个声道中的滤波的中心声道子带。

44. 如权利要求 38 所述的音频空间环境引擎, 其中, 所述求和级进一步包括左环绕声道求和级, 其对于每个子带满足以下方程:

$$(G_{LS}(f) * L(f) - ((1 - G_{LS}(f)) * R(f)) * H_{LS}(f))$$

其中

$G_{LS}(f)$  = 左环绕声道子带缩放因数;

$L(f)$  = 所述 M 个声道中的左声道子带;

$R(f)$  = 所述 M 个声道中的右声道子带; 以及

$H_{LS}(f)$  = 所述 N' 个声道中的滤波的左环绕声道子带。

45. 如权利要求 38 所述的音频空间环境引擎, 其中, 所述求和级进一步包括右环绕声道求和级, 其对于每个子带满足以下方程:

$$((1 - G_{RS}(f)) * R(f) + (G_{RS}(f)) * L(f)) * H_{RS}(f)$$

其中

$G_{RS}(f)$  = 右环绕声道子带缩放因数;

- $L(f)$  = 所述 M 个声道中的左声道子带;
- $R(f)$  = 所述 M 个声道中的右声道子带; 以及
- $H_{rs}(f)$  = 所述 N' 个声道中的滤波的右环绕声道子带。

46. 一种用于从 M 声道音频系统向 N 声道音频系统转换的方法, 其中, M 和 N 为整数, 并且, N 大于 M, 包括:

接收 M 个声道的音频数据;

生成用于所述 M 个声道中的每个声道的多个子带的音频空间图像数据;

滤波所述 M 个声道的所述多个子带的音频空间图像数据, 以生成 N' 个声道的多个子带的音频空间图像数据; 以及

将所述 M 个声道的所述多个子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据, 以生成缩放的 N' 个声道的所述多个子带的音频空间图像数据。

47. 如权利要求 46 所述的方法, 其中, 将所述 M 个声道的所述多个子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据进一步包括:

将所述 M 个声道的所述多个子带的音频空间图像数据中的一个或多个乘以子带缩放因数; 以及

将所述缩放的 M 个声道的所述多个子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据。

48. 如权利要求 46 所述的方法, 其中, 将所述 M 个声道的所述多个子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据进一步包括将所述 M 个声道的所述多个子带中的每一个乘以所述 N' 个声道的音频空间图像数据的相应子带。

49. 如权利要求 46 所述的方法, 其中, 将所述 M 个声道的所述多个子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据包括将所述 M 个声道中的左声道的多个子带中的每一个乘以所述 N' 个声道中的左声道的相应多个子带的音频空间图像数据中的每一个。

50. 如权利要求 46 所述的方法, 其中, 将所述 M 个声道的所述多个

子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据包括将所述 M 个声道中的右声道的多个子带中的每一个乘以所述 N' 个声道中的右声道的相应多个子带的音频空间图像数据中的每一个。

51. 如权利要求 46 所述的方法, 其中, 将所述 M 个声道的所述多个子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据包括对于每个子带满足以下方程:

$$(G_c(f) * L(f) + ((1 - G_c(f)) * R(f)) * H_c(f))$$

其中

$G_c(f)$  = 中心声道子带缩放因数;

$L(f)$  = 左声道子带;

$R(f)$  = 右声道子带; 以及

$H_c(f)$  = 滤波的中心声道子带。

52. 如权利要求 46 所述的方法, 其中, 将所述 M 个声道的所述多个子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据包括对于每个子带满足以下方程:

$$(G_{LS}(f) * L(f) - ((1 - G_{LS}(f)) * R(f)) * H_{LS}(f))$$

其中

$G_{LS}(f)$  = 左环绕声道子带缩放因数;

$L(f)$  = 左声道子带;

$R(f)$  = 右声道子带; 以及

$H_{LS}(f)$  = 滤波的左环绕声道子带。

53. 如权利要求 46 所述的方法, 其中, 将所述 M 个声道的所述多个子带的音频空间图像数据乘以所述 N' 个声道的所述多个子带的音频空间图像数据包括对于每个子带满足以下方程:

$$((1 - G_{RS}(f)) * R(f) + (G_{RS}(f)) * L(f)) * H_{RS}(f)$$

其中

$G_{RS}(f)$  = 右环绕声道子带缩放因数;

$L(f)$  = 左声道子带;

$R(f)$  = 右声道子带; 以及

$H_{RS}(f)$  = 滤波的右环绕声道子带。

54. 一种用于从 M 声道音频系统向 N 声道音频系统转换的音频空间环境引擎, 其中, M 和 N 为整数, 并且, N 大于 M, 包括:

时域到频域转换装置, 用于接收所述 M 个声道的音频数据, 并且生成多个子带的音频空间图像数据;

滤波发生器装置, 用于接收所述 M 个声道的所述多个子带的音频空间图像数据, 并且生成 N' 个声道的多个子带的音频空间图像数据; 以及

求和级装置, 用于接收所述 M 个声道的所述多个子带的音频空间图像数据和所述 N' 个声道的所述多个子带的音频空间图像数据, 并且生成缩放的 N' 个声道的所述多个子带的音频空间图像数据。

55. 如权利要求 54 所述的音频空间环境引擎, 进一步包括频域到时域转换级装置, 用于接收所述缩放的 N' 个声道的所述多个子带的音频空间图像数据, 并且生成所述 N' 个声道的音频数据。

56. 如权利要求 54 所述的音频空间环境引擎, 进一步包括:

平滑级装置, 用于接收所述 N' 个声道的所述多个子带的音频空间图像数据, 并且平均化每个子带与一个或多个相邻子带; 以及

其中, 所述求和级装置接收所述 M 个声道的所述多个子带的音频空间图像数据和所述平滑的 N' 个声道的所述多个子带的音频空间图像数据, 并且生成缩放的 N' 个声道的所述多个子带的音频空间图像数据。

57. 如权利要求 54 所述的音频空间环境引擎, 其中, 所述求和级装置进一步包括左声道求和级装置, 用于使所述 M 个声道中的左声道的多个子带中的每一个乘以所述 N' 个声道中的左声道的相应多个子带的音频空间图像数据中的每一个。



## 音频空间环境引擎

### 相关申请

本申请要求 2004 年 10 月 28 日提交的名称为 “2-to-N Rendering” 的美国临时申请 60/622,922、2004 年 10 月 28 日提交的名称为 “Audio Spatial Environment Engine” 的美国专利申请 10/975,841、同此提交的名称为 “Audio Spatial Environment Down-Mixer” 的美国专利申请 11/261,100 (代理人案号 13646.0014)、以及同此提交的名称为 “Audio Spatial Environment Up-Mixer” 的美国专利申请 11/262,029 (代理人案号 13646.0012) 的优先权, 其中的每一个都被共同拥有并为了全部的目的而通过引用结合于此。

### 技术领域

本发明涉及音频数据处理领域, 并且尤其涉及一种用于在音频数据的不同格式之间进行转换的系统和方法。

### 背景技术

用于处理音频数据的系统和方法在现有技术中是已知的。大多数这样的系统和方法用于处理已知音频环境的音频数据, 该已知音频环境诸如双声道立体声环境、四声道立体声环境、五声道环绕声环境 (也被称为 5.1 声道环境) 或其它适当的格式或环境。

格式或环境数目增加引发的一个问题是: 在第一环境中为了最佳音频质量而处理的音频数据通常不能够容易地在第二音频环境中使用。这个问题的一个例子是在整个为立体声数据设计的基础设施或网络上传输或存储环绕声音数据。因为用于立体双声道传输或存储的基础设施可能不支持环绕声格式的音频数据的附加声道, 所以难以或者不可能用现有的基础设施传输或利用环绕声格式的数据。

### 发明内容

根据本发明，提供了一种用于音频空间环境引擎的系统和方法，其通过在空间音频环境之间进行转换而克服了已知的问题。

具体地，提供了一种用于音频空间环境引擎的系统和方法，其允许在N声道数据和M声道数据之间进行转换，并且允许从M声道数据转换返回N'声道数据，其中，N、M和N'为整数，并且N不必等于N'。

根据本发明示范性的实施例，提供了一种音频空间环境引擎，用于从N声道音频系统转换到M声道音频系统并回到N'声道音频系统，其中，N、M和N'为整数，并且N不必等于N'。音频空间环境引擎包括动态下混频器，其接收N个声道的音频数据并将N个声道的音频数据转换成M个声道的音频数据。音频空间环境引擎还包括上混频器，其接收M个声道的音频数据并将M个声道的音频数据转换成N'个声道的音频数据，其中N不必等于N'。这个系统的一个示范性应用是在整个为立体声数据设计的基础设施或网络上传输或存储环绕声音数据。动态下混频单元将环绕声音数据转换成立体声音数据以用于传输或存储，并且上混频单元将立体声音数据恢复成环绕声音数据以用于回放、处理或某些其它适当的用途。

本发明提供了许多重要的技术优点。本发明的一个重要技术优点是这样一种系统，其由于先进的动态下混频单元和高分辨率频带上混频单元而在不同的空间环境之间提供了改进和灵活的转换。动态下混频单元包括智能分析与校正环，用于校正许多下混频方法所共有的频谱、时间和空间的不准确性。上混频单元利用对整个高分辨率频带提取并分析重要的声道间的空间提示(cue)，来得到不同频率元素的空间布置。下混频和上混频单元当单独或作为系统使用时，提供了改进的声音质量和空间区别。

本领域技术人员在结合附图阅读随后的详细说明时，会进一步意识到本发明的优点和优越特征以及其它的重要方面。

## 附图说明

图1是根据本发明的示范性实施例的利用分析和校正环进行动态下混频的系统的示图；

图2是根据本发明的示范性实施例的用于从N个声道向M个声道进行下混频数据的系统的示图；

图3是根据本发明的示范性实施例的用于从5个声道向2个声道进行下混频数据的系统的示图；

图 4 是根据本发明的示范性实施例的子带(sub-band)矢量计算系统的示图;

图 5 是根据本发明的示范性实施例的子带校正系统的示图;

图 6 是根据本发明的示范性实施例的用于从 M 个声道向 N 个声道进行上混频数据的系统的示图;

图 7 是根据本发明的示范性实施例的用于从 2 个声道向 5 个声道进行上混频数据的系统的示图;

图 8 是根据本发明的示范性实施例的用于从 2 个声道向 7 个声道进行上混频数据的系统的示图;

图 9 是根据本发明的示范性实施例的用于提取声道间空间提示并产生用于频域应用的空间声道滤波的方法的示图;

图 10A 是根据本发明的示范性实施例的示范性左前声道滤波图的示图;

图 10B 是示范性右前声道滤波图的示图;

图 10C 是示范性中心声道滤波图的示图;

图 10D 是示范性左环绕声道滤波图的示图; 以及

图 10E 是示范性右环绕声道滤波图的示图。

## 具体实施方式

在随后的说明中, 贯穿本说明书和附图用相同的附图标记来标志相同的部分。附图可能没有按比例, 并且某些组成部分能够以概括或示意性的形式显示, 而且为了清楚和简明起见用商品名称来标识。

图 1 是根据本发明的示范性实施例的利用分析与校正环从 N 声道音频格式向 M 声道音频格式进行动态下混频的系统 100 的示图。系统 100 使用 5.1 声道声音(即,  $N=5$ ), 并且将 5.1 声道声音转换成立体声(即,  $M=2$ ), 但是其它适当数目的输入和输出声道也能够或代替地被使用。

使用基准下混频 102、基准上混频 104、子带矢量计算系统 106 和 108 以及子带校正系统 110 来实现系统 100 的动态下混频过程。通过基准上混频 104、子带矢量计算系统 106 和 108 以及子带校正系统 110 实现分析和校正环, 其中, 基准上混频 104 模拟上混频过程, 子带矢量计算系统 106

和 108 计算模拟的上混频和原始信号的每一频带的能量和位置矢量,以及子带校正系统 110 比较模拟的上混频和原始信号的能量和位置矢量,并且调整下混频信号的声道间空间提示以对任何不一致进行校正。

系统 100 包括静态基准下混频 102, 其将接收到的  $N$  声道音频转换成  $M$  声道音频。静态基准下混频 102 接收 5.1 声音声道左  $L(T)$ 、右  $R(T)$ 、中心  $C(T)$ 、左环绕  $LS(T)$  和右环绕  $RS(T)$ , 并且将 5.1 声道信号转换成立体声声道信号左水印  $LW'(T)$  和右水印  $RW'(T)$ 。

左水印  $LW'(T)$  和右水印  $RW'(T)$  立体声道信号随后被提供给基准上混频 104, 其将立体声声道转换成 5.1 声音声道。基准上混频 104 输出 5.1 声音声道左  $L'(T)$ 、右  $R'(T)$ 、中心  $C'(T)$ 、左环绕  $LS'(T)$  和右环绕  $RS'(T)$ 。

从基准上混频 104 输出的上混频的 5.1 声道声音信号然后被提供给子带矢量计算系统 106。来自子带矢量计算系统 106 的输出是用于上混频的 5.1 声道信号  $L'(T)$ 、 $R'(T)$ 、 $C'(T)$ 、 $LS'(T)$  和  $RS'(T)$  的多个频带的上混频能量和图像位置数据。同样地, 原始 5.1 声道的声音信号被提供给子带矢量计算系统 108。来自子带矢量计算系统 108 输出的是用于原始 5.1 声道信号  $L(T)$ 、 $R(T)$ 、 $C(T)$ 、 $LS(T)$  和  $RS(T)$  的多个频带的源能量和图像位置数据。子带矢量计算系统 106 和 108 计算的能量和位置矢量包括对于每一频带的总能量测量和 2 维矢量, 其指示在理想收听条件下对于收听者的给定频率元素的感觉强度和源位置。例如, 使用适当的滤波器组能够将音频信号从时域转换到频域, 该滤波器组诸如有限脉冲响应 (FIR) 滤波器组、正交镜像滤波器 (QMF) 组、离散傅立叶变换 (DFT)、时域混叠消除 (TDAC) 滤波器组或其它适当的滤波器组。滤波器组输出被进一步处理以确定每一频带的总能量和每一频带的标准化图像位置矢量。

从子带矢量计算系统 106 和 108 输出的能量和位置矢量值被提供给子带校正系统 110, 其分析用于原始 5.1 声道声音的源能量和位置与当从左水印  $LW'(T)$  和右水印  $RW'(T)$  立体声道信号生成时用于 5.1 声道声音的上混频能量和位置。然后, 在产生  $LW(T)$  和  $RW(T)$  的左水印  $LW'(T)$  和右水印  $RW'(T)$  上对每一子带识别和校正源和上混频的能量与位置矢量之间的差, 以便提供更加准确的下混频的立体声道信号, 并且当立体声道信号随后被上混频时提供更加准确的 5.1 再现。校正的左水印  $LW(T)$  和右水印  $RW(T)$  信号被输出用于传输、由立体声接收机接收、由具有上混频功能的接收机接收或者用于其它适当的用途。

在工作中，通过包括整个下混频/上混频系统的模拟、分析和校正的智能分析和校正环，系统100将5.1声道声音动态地下混频到立体声声音。这种方法通过以下步骤完成，即，生成静态下混频立体声信号  $LW'(T)$  和  $RW'(T)$ ；模拟随后的上混频信号  $L'(T)$ 、 $R'(T)$ 、 $C'(T)$ 、 $LS'(T)$  和  $RS'(T)$ ；以及分析那些信号与原始5.1声道信号以在子带的基础上识别和校正任何能量或位置矢量差，其能够影响左水印  $LW'(T)$  和右水印  $RW'(T)$  立体声信号或随后上混频的环绕声道信号的质量。产生左水印  $LW(T)$  和右水印  $RW(T)$  立体声信号的子带校正处理被执行，以便当  $LW(T)$  和  $RW(T)$  被上混频时，作为结果而得到的5.1声道声音以改进的准确性匹配原始输入的5.1声道声音。同样地，附加处理能够被执行，以便允许任何适当数目的输入声道被转换成适当数目的水印输出声道，诸如7.1声道声音转换成水印立体声、7.1声道声音转换成水印5.1声道声音、定制声音声道（诸如用于汽车音响系统或电影院之类）转换成立体声或者其它适当的转换。

图2是根据本发明的示范性实施例的静态基准下混频200的示图。静态基准下混频200能够作为图1的基准下混频102或者以其它适当的方式使用。

基准下混频200将N声道音频转换成M声道音频，其中，N和M为整数，并且N大于M。基准下混频200接收输入信号  $X_1(T)$ 、 $X_2(T)$  到  $X_N(T)$ 。对于每个输入声道i，输入信号  $X_i(T)$  被提供给希尔伯特(Hilbert)变换单元202到206，其引入  $90^\circ$  相移信号。诸如实现  $90^\circ$  相移的希尔伯特滤波器或全通滤波器网络之类的其它处理也能够或代替地被使用而代替希尔伯特变换单元。对于每个输入声道i，希尔伯特变换信号和原始输入信号然后通过第一级乘法器208到218分别与预定比例常数  $C_{i11}$  和  $C_{i12}$  相乘，其中，第一个下标表示输入声道号i，第二个下标表示第一级乘法器，而第三个下标表示每级的乘法器号。乘法器208到218的输出然后被加法器220到224求和，生成分数希尔伯特信号  $X'_i(T)$ 。相对于相应的输入信号  $X_i(T)$ ，从乘法器220到224输出的分数希尔伯特信号  $X'_i(T)$  具有可变数量的相移。相移量取决于比例常数  $C_{i11}$  和  $C_{i12}$ ，其中， $0^\circ$  相移可以对应于  $C_{i11} = 0$  和  $C_{i12} = 1$ ，而  $\pm 90^\circ$  相移可以对应于  $C_{i11} = \pm 1$  和  $C_{i12} = 0$ 。使用  $C_{i11}$  和  $C_{i12}$  的适当值，任何中间量的相移都是可能的。

用于每个输入声道i的每个信号  $X'_i(T)$  然后通过第二级乘法器226到242与预定比例常数  $C_{i2j}$  相乘，其中，第一个下标表示输入声道号i，

第二下标表示第二级乘法器,而第三个下标表示输出声道号  $j$ 。乘法器 226 到 242 的输出然后被加法器 244 到 248 适当求和以生成用于每个输出声道  $j$  的相应输出信号  $Y_j(T)$ 。通过每个输入声道  $i$  和输出声道  $j$  的空间位置确定用于每个输入声道  $i$  和输出声道  $j$  的比例常数  $C_{i2j}$ 。例如,用于左输入声道  $i$  和右输出声道  $j$  的比例常数  $C_{i2j}$  能够被设置为近似零以保持空间区别。同样地,用于前输入声道  $i$  和前输出声道  $j$  的比例常数  $C_{i2j}$  能够被设置为近似 1 以保持空间布置。

在工作中,在接收机接收输出信号时,基准下混频 200 以允许输入信号中的空间关系被任意管理和提取的方式将  $N$  个声道组合成  $M$  个声道。进而,如所示的  $N$  声道声音的组合生成  $M$  声道声音,其对于在  $M$  声道音频环境中收听的收听者是可接受的质量。因此,基准下混频 200 能够被用于将  $N$  声道声音转换成  $M$  声道声音,其能够被  $M$  声道接收机、具有适当上混频器的  $N$  声道接收机或其它适当接收机使用。

图 3 是根据本发明的示范性实施例的静态基准下混频 300 的示图。如图 3 所示,静态基准下混频 300 是图 2 的静态基准下混频 200 的实现,其将 5.1 声道时域数据转换成立体声道时域数据。静态基准下混频 300 能够作为图 1 的基准下混频 102 或以其它适当的方式使用。

基准下混频 300 包括希尔伯特变换 302,其接收源 5.1 声道声音的左声道信号  $L(T)$ ,并且对时间信号执行希尔伯特变换。希尔伯特变换引入信号的  $90^\circ$  相移,然后其通过乘法器 310 与预定比例常数  $C_{L1}$  相乘。诸如实现  $90^\circ$  相移的希尔伯特滤波器或全通滤波器网络之类的其它处理也能够或代替地被使用而代替希尔伯特变换单元。原始左声道信号  $L(T)$  通过乘法器 312 与预定比例常数  $C_{L2}$  相乘。乘法器 310 和 312 的输出被加法器 320 求和,以生成分数希尔伯特信号  $L'(T)$ 。同样地,来自源 5.1 声道声音的右声道信号  $R(T)$  由希尔伯特变换 304 处理,并且通过乘法器 314 与预定比例常数  $C_{R1}$  相乘。原始右声道信号  $R(T)$  通过乘法器 316 与预定比例常数  $C_{R2}$  相乘。乘法器 314 和 316 的输出被加法器 322 求和,以生成分数希尔伯特信号  $R'(T)$ 。相对于相应的输入信号  $L(T)$  和  $R(T)$ ,从乘法器 320 和 322 输出的分数希尔伯特信号  $L'(T)$  和  $R'(T)$  分别具有可变量的相移。相移量取决于比例常数  $C_{L1}$ 、 $C_{L2}$ 、 $C_{R1}$  和  $C_{R2}$ ,其中,  $0^\circ$  相移可对应于  $C_{L1} = 0$ ,  $C_{L2} = 1$ ,  $C_{R1} = 0$  和  $C_{R2} = 1$ ,而  $\pm 90^\circ$  相移可对应于  $C_{L1} = \pm 1$ ,  $C_{L2} = 0$ ,  $C_{R1} = \pm 1$  和  $C_{R2} = 0$ 。使用  $C_{L1}$ 、 $C_{L2}$ 、 $C_{R1}$  和  $C_{R2}$  的适当值,任何中间量的相移都是可能的。来自源 5.1 声道声音的中心声道输入被提供给乘

法器 318 作为分数希尔伯特信号  $C'(T)$ ，意味着没有对中心声道输入信号进行相移。乘法器 318 使  $C'(T)$  与预定比例常数  $C_3$  相乘，如衰减 3 分贝。加法器 320 和 322 以及乘法器 318 的输出被适当地求和成左水印声道  $LW'(T)$  和右水印声道  $RW'(T)$ 。

来自源 5.1 声道声音的左环绕声道  $LS(T)$  被提供给希尔伯特变换 306，而来自源 5.1 声道声音的右环绕声道  $RS(T)$  被提供给希尔伯特变换 308。希尔伯特变换 306 和 308 的输出是分数希尔伯特信号  $LS'(T)$  和  $RS'(T)$ ，意味着在  $LS(T)$  和  $LS'(T)$  信号对以及  $RS(T)$  和  $RS'(T)$  信号对之间存在满  $90^\circ$  相移。 $LS'(T)$  然后通过乘法器 324 和 326 分别与预定比例常数  $C_{LS1}$  和  $C_{LS2}$  相乘。同样地， $RS'(T)$  通过乘法器 328 和 330 分别与预定比例常数  $C_{RS1}$  和  $C_{RS2}$  相乘。乘法器 324 到 330 的输出被适当地提供给左水印声道  $LW'(T)$  和右水印声道  $RW'(T)$ 。

加法器 332 接收从加法器 320 输出的左声道信号、从乘法器 318 输出的中心声道信号、从乘法器 324 输出的左环绕声道信号以及从乘法器 328 输出的右环绕声道信号，并且相加这些信号以形成左水印声道  $LW'(T)$ 。同样地，加法器 334 接收从乘法器 318 输出的中心声道信号、从加法器 322 输出的右声道信号、从乘法器 326 输出的左环绕声道信号以及从乘法器 330 输出的右环绕声道信号，并且相加这些信号以形成右水印声道  $RW'(T)$ 。

在工作中，在接收机接收左水印声道和右水印声道立体声信号时，基准下混频 300 以允许 5.1 输入声道中的空间关系被维持和提取的方式组合源 5.1 声道。进而，如所示的 5.1 声道声音的组合生成立体声声音，其对于使用没有进行环绕声上混频的立体声接收机的收听者是可接受的质量。因此，基准下混频 300 能够被用于将 5.1 声道声音转换成立体声的声音，其能够被立体声接收机、具有适当上混频器的 5.1 声道接收机、具有适当上混频器的 7.1 声道接收机或其它适当接收机使用。

图 4 是根据本发明示范性实施例的子带矢量计算系统 400 的示图。子带矢量计算系统 400 提供用于多个频带的能量和位置矢量数据，并且能够用作图 1 的子带矢量计算系统 106 和 108。尽管示出了 5.1 声道声音，但是能够使用其它适当的声道配置。

子带矢量计算系统 400 包括时间-频率分析单元 402 到 410。5.1 时域声道  $L(T)$ 、 $R(T)$ 、 $C(T)$ 、 $LS(T)$  和  $RS(T)$  分别被提供给时间-频率分析单元 402 到 410，其将时域信号转换成频域信号。这些时间-频率分

析单元能够是适当的滤波器组，诸如有限脉冲响应（FIR）滤波器组、正交镜像滤波器（QMF）组、离散傅立叶变换（DFT）、时域混叠消除（TDAC）滤波器组或其它适当的滤波器组。对于  $L(F)$ 、 $R(F)$ 、 $C(F)$ 、 $LS(F)$  和  $RS(F)$ ，从时间-频率分析单元 402 到 410 输出每一频带的幅度或能量值。这些幅度/能量值包括对每个相应声道的每个频带分量的幅度/能量测量。幅度/能量测量被加法器 412 求和，该加法器 412 输出  $T(F)$ ，其中， $T(F)$  是每一频带的输入信号的总能量。这个值然后被除法单元 414 到 422 分成声道幅度/能量值中的每一个，以生成相应的标准化声道间级差（ICLD）信号  $M_L(F)$ 、 $M_R(F)$ 、 $M_C(F)$ 、 $M_{LS}(F)$  和  $M_{RS}(F)$ ，其中，这些 ICLD 信号能够被视为对每个声道的标准化子带能量估计。

5.1 声道声音被映射到标准化位置矢量，如用由横轴和纵轴组成的 2 维平面上的示范性定位所示。如图所示，用于  $(X_{LS}, Y_{LS})$  的定位值被分配到原点， $(X_{RS}, Y_{RS})$  的值被分配到  $(0, 1)$ ， $(X_L, Y_L)$  的值被分配到  $(0, 1 - C)$ ，其中， $C$  是 1 和 0 之间的值，表示左右扬声器离房间后面的退后距离。同样地， $(X_R, Y_R)$  的值为  $(1, 1 - C)$ 。最后，用于  $(X_C, Y_C)$  的值为  $(0.5, 1)$ 。这些坐标是示范性的，并且能够被改变以反映扬声器相对于彼此的实际标准化定位或配置，如扬声器坐标基于房间的大小、房间的形状或其它因素而不同。例如，在使用 7.1 声音或其它适当声道配置时，能够提供附加的坐标值，其反映扬声器在房间周围的定位。同样地，这样的扬声器定位能够基于扬声器在汽车、房间、礼堂、剧场或如其它适当地方中的实际分布来定制。

估计的图像位置矢量  $P(F)$  能够以如下矢量方程中阐述的那样对每一子带来计算：

$$P(F) = M_L(F) * (X_L, Y_L) + M_R(F) * (X_R, Y_R) + M_C(F) * (X_C, Y_C) + \\ i. M_{LS}(F) * (X_{LS}, Y_{LS}) + M_{RS}(F) * (X_{RS}, Y_{RS})$$

因此，对于每个频带，提供总能量  $T(F)$  的输出和位置矢量  $P(F)$ ，它们被用来为该频带限定视在频率源的感觉强度和位置。以这种方式，频率分量的空间图像能够被定位，诸如用于子带校正系统 110 或用于其它适当的目的。

图 5 是根据本发明示范性实施例的子带校正系统的示图。子带校正系统能够用作图 1 的子带校正系统 110 或用于其它适当的目的。子带校正系统接收左水印  $LW'(T)$  和右水印  $RW'(T)$  立体声道信号，并且对水印信



号进行能量和图像校正,以便为每个频带补偿作为基准下混频或其它适当方法的结果而可能产生的信号不准确性。子带校正系统对每个频带接收并利用源的总能量信号  $T_{SOURCE}(F)$  和随后的上混频信号的总能量信号  $T_{UMIX}(F)$  以及用于源的位置矢量  $P_{SOURCE}(F)$  和随后的上混频信号的位置矢量  $P_{UMIX}(F)$ , 诸如图 1 的子带矢量计算系统 106 和 108 生成的那些。这些总能量信号和位置矢量用于确定将要进行的适当校正和补偿。

子带校正系统包括位置校正系统 500 和频谱能量校正系统 502。位置校正系统 500 接收用于左水印立体声道  $LW'(T)$  和右水印立体声道  $RW'(T)$  的时域信号,其分别由时间-频率分析单元 504 和 506 从时域转换成频域。这些时间-频率分析单元能够是适当的滤波器组,诸如有限脉冲响应(FIR)滤波器组、正交镜像滤波器(QMF)组、离散傅立叶变换(DFT)、时域混叠消除(TDAC)滤波器组或其它适当的滤波器组。

时间-频率分析单元 504 和 506 的输出是频域子带信号  $LW'(F)$  和  $RW'(F)$ 。在信号  $LW'(F)$  和  $RW'(F)$  中对于每一子带调整声道间级差(ICLD)和声道间相干性(ICC)的相关空间提示。例如,这些提示能够通过操纵  $LW'(F)$  和  $RW'(F)$  的幅度或能量(所示为  $LW'(F)$  和  $RW'(F)$  的绝对值)以及  $LW'(F)$  和  $RW'(F)$  的相角来调整。通过由乘法器 508 将  $LW'(F)$  的幅度/能量值与以下方程所生成的值相乘来进行 ICLD 的校正:

$$[X_{MAX} - P_{X,SOURCE}(F)] / [X_{MAX} - P_{X,UMIX}(F)]$$

其中

$X_{MAX}$  = 最大 X 坐标边界

$P_{X,SOURCE}(F)$  = 相对于源矢量的估计的子带 X 位置坐标

$P_{X,UMIX}(F)$  = 相对于随后上混频矢量的估计的子带 X 位置坐标

同样地,用于  $RW'(F)$  的幅度/能量通过乘法器 510 与以下方程所生成的值相乘:

$$[P_{X,SOURCE}(F) - X_{MIN}] / [P_{X,UMIX}(F) - X_{MIN}]$$

其中

$X_{MIN}$  = 最小 X 坐标边界

通过由加法器 512 将用于  $LW'(F)$  的相角与以下方程所生成的值相加来进行 ICC 的校正:

$$+/- \pi * [P_{Y,SOURCE}(F) - P_{Y,UMIX}(F)] / [Y_{MAX} - Y_{MIN}]$$

其中

$P_{Y,SOURCE}(F)$  = 相对于源矢量的估计的子带 Y 位置坐标

$P_{Y,UMIX}(F)$  = 相对于随后上混频矢量的估计的子带 Y 位置坐标

$Y_{MAX}$  = 最大 Y 坐标边界

$Y_{MIN}$  = 最小 Y 坐标边界

同样地，用于  $RW'(F)$  的相角通过加法器 514 与以下方程所生成的值相加：

$$-/+ \pi * [P_{Y,SOURCE}(F) - P_{Y,UMIX}(F)] / [Y_{MAX} - Y_{MIN}]$$

注意，相加到  $LW'(F)$  和  $RW'(F)$  的角分量具有相等的值但具有相反的极性，其中，作为结果的极性由  $LW'(F)$  和  $RW'(F)$  之间的超前相角来确定。

校正的  $LW'(F)$  幅度/能量和  $LW'(F)$  相角被加法器 516 重新组合以形成用于每个子带的复值  $LW(F)$ ，然后通过频率-时间综合单元 520 转换成左水印时域信号  $LW(T)$ 。同样地，校正的  $RW'(F)$  幅度/能量和  $RW'(F)$  相角被加法器 518 重新组合以形成用于每个子带的复值  $RW(F)$ ，然后通过频率-时间综合单元 522 转换成右水印时域信号  $RW(T)$ 。频率-时间综合单元 520 和 522 能够是能将频域信号转换回时域信号的适当综合滤波器组。

如这个示范性实施例所示，通过使用适当调整 ICLD 和 ICC 空间提示的位置校正 500，能够校正用于水印左和右声道信号的每个频谱分量的声道间空间提示。

频谱能量校正系统 502 能够用于确保下混频信号的总频谱平衡与原始 5.1 信号的总频谱平衡相一致，因此，例如就补偿了由梳状滤波造成的频谱偏移。分别使用时间-频率分析单元 524 和 526 来将左水印时域信号  $LW'(T)$  和右水印时域信号  $RW'(T)$  从时域转换到频域。这些时间-频率分析单元能够是适当的滤波器组，诸如有限脉冲响应 (FIR) 滤波器组、正交镜像滤波器 (QMF) 组、离散傅立叶变换 (DFT)、时域混叠消除 (TDAC) 滤波器组或其它适当的滤波器组。从时间-频率分析单元 524 和 526 输出的是  $LW'(F)$  和  $RW'(F)$  频率子带信号，其通过乘法器 528 和 530 与  $T_{SOURCE}(F) / T_{UMIX}(F)$  相乘，其中

$$\begin{aligned}
 T_{\text{SOURCE}}(F) &= |L(F)| + |R(F)| + |C(F)| + |LS(F)| + \\
 &\quad |RS(F)| \\
 T_{\text{MIX}}(F) &= |L_{\text{MIX}}(F)| + |R_{\text{MIX}}(F)| + |C_{\text{MIX}}(F)| + \\
 &\quad |L_{\text{SUMIX}}(F)| + |R_{\text{SUMIX}}(F)|
 \end{aligned}$$

来自乘法器 528 和 530 的输出然后被频率-时间综合单元 532 和 534 从频域转换回时域以生成  $LW(T)$  和  $RW(T)$ 。频率-时间综合单元可以是能够将频域信号转换回时域信号的适当综合滤波器组。以这种方式, 位置和能量校正能够被施加到下混频立体声道信号  $LW'(T)$  和  $RW'(T)$ , 以便产生忠实于原始 5.1 信号的左和右水印声道信号  $LW(T)$  和  $RW(T)$ 。 $LW(T)$  和  $RW(T)$  能够以立体声回放或上混频回到 5.1 声道或其它适当数目的声道, 而不显著改变原始 5.1 声道声音中存在的任意内容元素的频谱分量位置或能量。

图 6 是根据本发明示范性实施例的用于从  $M$  个声道向  $N$  个声道上混频数据的系统 600 的示图。系统 600 将立体声时域数据转换成  $N$  声道时域数据。

系统 600 包括时间-频率分析单元 602 和 604、滤波生成单元 606、平滑单元 608 以及频率-时间综合单元 634 到 638。通过允许高分辨率频带处理的可伸缩频域结构, 并且通过提取并分析每频带的重要声道间的空间提示以得到上混频  $N$  声道信号中频率元素的空间布置的滤波生成方法, 系统 600 在上混频过程中提供了改进的空间区别和稳定性。

系统 600 在将时域信号转换成频域信号的时间-频率分析单元 602 和 604 处接收左声道立体声信号  $L(T)$  和右声道立体声信号  $R(T)$ 。这些时间-频率分析单元能够是适当的滤波器组, 诸如有限脉冲响应 (FIR) 滤波器组、正交镜像滤波器 (QMF) 组、离散傅立叶变换 (DFT)、时域混叠消除 (TDAC) 滤波器组或其它适当的滤波器组。从时间-频率分析单元 602 和 604 输出的是一组频域值, 覆盖了人类听觉系统的充分的频率范围, 如 0 到 20kHz 的频率范围, 其中分析滤波器组子带带宽能够被处理以逼近心理声学的临界频带、等效矩形带宽或某种其它知觉特性。同样地, 能够使用其它适当数目的频带和范围。

来自时间-频率分析单元 602 和 604 的输出被提供给滤波生成单元 606。在一个示范性实施例中, 滤波生成单元 606 能够接收关于应当为给定的环境而输出的声道数目的外部选择。例如, 能够选择有两前和两后扬

声器的 4.1 声道,能够选择有两前、两后和一个前中心扬声器的 5.1 音响系统,能够选择有两前、两侧、两后和一个前中心扬声器的 7.1 音响系统,或者能够选择其它适当的音响系统。滤波生成单元 606 在频带的基础上提取并分析诸如声道间级差 (ICLD) 和声道间相干性 (ICC) 之类的声道间空间提示。那些相关的空间提示然后用作参数以生成适应的声道滤波,其控制上混频声场中频带元素的空间布置。通过平滑单元 608 在整个时间和频率上均使声道滤波平滑以限制滤波可变性,所述滤波可变性如果被允许变化太迅速的话,能够造成讨厌的波动效应。在图 6 显示的示范性实施例中,左和右声道  $L(F)$  和  $R(F)$  频域信号被提供给滤波生成单元 606,其产生  $N$  个声道滤波信号  $H_1(F)$ 、 $H_2(F)$  到  $H_N(F)$ ,它们被提供给平滑单元 608。

平滑单元 608 在整个时间和频率维数上为  $N$  个声道滤波中的每个声道平均化频域分量。在整个时间和频率上的平滑有助于控制声道滤波信号中的快速波动,因此就减少了能够令收听者讨厌的抖动现象 (jitter artifact) 和不稳定性。在一个示范性实施例中,通过对根据当前帧的每个频带和根据以前帧的相应频带应用一阶低通滤波,能够实现时间平滑。这具有逐帧减少每个频带的可变性的效果。在另一个示范性实施例中,能够对整个频率槽 (bin) 的组进行频谱平滑,所述频率槽被模拟成近似人类听觉系统的临界频带间隔。例如,如果使用具有均匀间隔的频率槽的分析滤波器组,则对于频谱的不同分区能够对不同数目的频率槽分组和平均化。例如,从 0 到 5kHz,能够平均化 5 个频率槽,从 5kHz 到 10kHz,能够平均化 7 个频率槽,并且从 10kHz 到 20kHz,能够平均化 9 个频率槽,或者能够选择其它适当数目的频率槽和带宽范围。从平滑单元 608 输出  $H_1(F)$ 、 $H_2(F)$  到  $H_N(F)$  的平滑值。

用于  $N$  个输出声道中的每一个的源信号  $X_1(F)$ 、 $X_2(F)$  到  $X_N(F)$  被生成为  $M$  个输入声道的适应组合。在图 6 显示的示范性实施例中,对于给定的输出声道  $i$ ,从加法器 614、620 和 626 输出的声道源信号  $X_i(F)$  被生成为  $L(F)$  乘以适应的缩放信号  $G_i(F)$  和  $R(F)$  乘以适应缩放信号  $1 - G_i(F)$  之和。乘法器 610、612、616、618、622 和 624 所使用的适应缩放信号  $G_i(F)$  是通过输出声道  $i$  的预期空间位置和每一频带的  $L(F)$  和  $R(F)$  的动态声道间相干性估计来确定的。同样地,被提供给加法器 614、620 和 626 的信号极性通过输出声道  $i$  的预期空间位置来确定。例如,在加法器 614、620 和 626 处的适应缩放信号  $G_i(F)$  和极性能够被设计来为前中心声道提供  $L(F) + R(F)$  组合,为左声道提供  $L(F)$ ,为

右声道提供  $R(F)$ ，以及为后声道提供  $L(F) - R(F)$ ，如传统矩阵上混频方法中通用的那样。适应缩放信号  $G_i(F)$  能够进一步提供一方法以动态调整输出声道对之间的相互关系，无论它们是横向还是纵向的声道对。

声道源信号  $X_1(F)$ 、 $X_2(F)$  到  $X_N(F)$  分别通过乘法器 628 到 632 与平滑的声道滤波  $H_1(F)$ 、 $H_2(F)$  到  $H_N(F)$  相乘。

来自乘法器 628 到 632 的输出然后被频率-时间综合单元 634 到 638 从频域转换到时域以生成输出声道  $Y_1(T)$ 、 $Y_2(T)$  到  $Y_N(T)$ 。以这种方式，左和右立体声信号被上混频到  $N$  个声道信号，其中，自然存在的或者如通过图 1 的下混频水印过程或其它适当过程那样的被有意编码到左和右立体声信号中的声道间空间提示，能够用于控制系统 600 产生的  $N$  声道声场之内的频率元素的空间布置。同样地，能够使用输入和输出的其它适当组合，诸如立体声到 7.1 声音、5.1 到 7.1 声音或者其它适当的组合。

图 7 是根据本发明的示范性实施例的用于从  $M$  个声道向  $N$  个声道上混频数据的系统 700 的示图。系统 700 将立体声时域数据转换成 5.1 声道时域数据。

系统 700 包括时间-频率分析单元 702 和 704、滤波生成单元 706、平滑单元 708 以及频率-时间综合单元 738 到 746。通过允许高分辨率频带处理的可伸缩频域结构的使用，并且通过提取并分析每一频带的重要声道间的空间提示以得到上混频 5.1 声道信号中频率元素的空间布置的滤波生成方法，系统 700 在上混频过程中提供了改进的空间区别和稳定性。

系统 700 在将时域信号转换成频域信号的时间-频率分析单元 702 和 704 处接收左声道立体声信号  $L(T)$  和右声道立体声信号  $R(T)$ 。这些时间-频率分析单元能够是适当的滤波器组，诸如有限脉冲响应 (FIR) 滤波器组、正交镜像滤波器 (QMF) 组、离散傅立叶变换 (DFT)、时域混叠消除 (TDAC) 滤波器组或其它适当的滤波器组。从时间-频率分析单元 702 和 704 输出的是一组频域值，覆盖了人类听觉系统的充分的频率范围，如 0 到 20kHz 的频率范围，其中分析滤波器组子带带宽能够被处理以近似心理声学的临界频带、等效矩形带宽或某种其它知觉特性。同样地，能够使用其它适当数目的频带和范围。

来自时间-频率分析单元 702 和 704 的输出被提供给滤波生成单元 706。在一个示范性实施例中，关于应当为给定环境而输出的声道数目，滤波生成单元 706 能够接收外部选择，诸如能够选择有两前和两后扬声器

的 4.1 声道,能够选择有两前、两后和一个前中心扬声器的 5.1 音响系统,能够选择有两前和一个前中心扬声器的 3.1 音响系统,或者能够选择其它适当的音响系统。滤波生成单元 706 在频带的基础上提取并分析诸如声道间级差 (ICLD) 和声道间相干性 (ICC) 之类的声道间空间提示。那些相关的空间提示然后用作参数以生成适应的声道滤波,其控制上混频声场中频带元素的空间布置。通过平滑单元 708 在整个时间和频率上使声道滤波平滑以限制滤波可变性,所述滤波可变性如果被允许变化太迅速的话,能够造成讨厌的波动效应。在图 7 显示的示范性实施例中,左和右声道  $L(F)$  和  $R(F)$  频域信号被提供给滤波生成单元 706,其产生 5.1 声道滤波信号  $H_L(F)$ 、 $H_R(F)$ 、 $H_C(F)$ 、 $H_{LS}(F)$  和  $H_{RS}(F)$ ,它们被提供给平滑单元 708。

平滑单元 708 在整个时间和频率维数上为 5.1 声道滤波中的每个声道平均化频域分量。在整个时间和频率上的平滑有助于控制声道滤波信号中的快速波动,因此,减少了能够令收听者讨厌的抖动现象和不稳定性。在一个示范性实施例中,通过对根据当前帧的每个频带和根据先前帧的相应频带应用一阶低通滤波能够实现时间平滑。这具有逐帧减少每个频带的可变性的效果。在一个示范性实施例中,能够对整个频率槽组进行频谱平滑,所述频率槽被模拟成近似人类听觉系统的临界频带间隔。例如,如果使用具有均匀间隔的频率槽的分析滤波器组,则对于频谱的不同分区能够分组和平均化不同数目的频率槽。在这个示范性实施例中,从 0 到 5kHz 能够平均化 5 个频率槽,从 5kHz 到 10kHz 能够平均化 7 个频率槽,并且从 10kHz 到 20kHz 能够平均化 9 个频率槽,或者能够选择其它适当数目的频率槽和带宽范围。从平滑单元 708 输出  $H_L(F)$ 、 $H_R(F)$ 、 $H_C(F)$ 、 $H_{LS}(F)$  和  $H_{RS}(F)$  的平滑值。

用于 5.1 输出声道中的每一个的源信号  $X_L(F)$ 、 $X_R(F)$ 、 $X_C(F)$ 、 $X_{LS}(F)$  和  $X_{RS}(F)$  被生成为立体声输入声道的适应组合。在图 7 所示的示范性实施例中, $X_L(F)$  被简单地提供为  $L(F)$ ,意味着对于全部频带都有  $G_L(F) = 1$ 。同样地, $X_R(F)$  被简单地提供为  $R(F)$ ,意味着对于全部频带都有  $G_R(F) = 0$ 。作为从加法器 714 输出的  $X_C(F)$  被计算为信号  $L(F)$  乘以适应缩放信号  $G_C(F)$  与  $R(F)$  乘以适应缩放信号  $1 - G_C(F)$  的和。作为从加法器 720 输出的  $X_{LS}(F)$  被计算为信号  $L(F)$  乘以适应缩放信号  $G_{LS}(F)$  与  $R(F)$  乘以适应缩放信号  $1 - G_{LS}(F)$  的和。同样地,作为从加法器 726 输出的  $X_{RS}(F)$  被计算为信号  $L(F)$  乘以适应缩放信号  $G_{RS}(F)$  与  $R(F)$  乘以适应缩放信号  $1 - G_{RS}(F)$  的和。注意,如果对于全部频带都有  $G_C(F) = 0.5$ 、 $G_{LS}(F) = 0.5$ , 以及  $G_{RS}(F) = 0.5$ , 那么前中心声

道源自  $L(F) + R(F)$  组合, 并且环绕声道源自经缩放的  $L(F) - R(F)$  组合, 如传统矩阵上混频方法中通用的那样。适应缩放信号  $G_C(F)$ 、 $G_{LS}(F)$  和  $G_{RS}(F)$  能够进一步提供一方法以动态调整相邻输出声道对之间的相关性, 无论它们是横向还是纵向的声道对。声道源信号  $X_L(F)$ 、 $X_R(F)$ 、 $X_C(F)$ 、 $X_{LS}(F)$  和  $X_{RS}(F)$  分别通过乘法器 728 到 736 与平滑的声道滤波  $H_L(F)$ 、 $H_R(F)$ 、 $H_C(F)$ 、 $H_{LS}(F)$  和  $H_{RS}(F)$  相乘。

来自乘法器 728 到 736 的输出然后被频率-时间综合单元 738 到 746 从频域转换到时域以生成输出声道  $Y_L(T)$ 、 $Y_R(T)$ 、 $Y_C(F)$ 、 $Y_{LS}(F)$  和  $Y_{RS}(T)$ 。以这种方式, 左和右立体声信号被上混频到 5.1 声道信号, 其中, 自然存在的或者如通过图 1 的下混频水印过程或其它适当过程那样的被有意编码到左和右立体声信号的声道间空间提示, 能够用于控制系统 700 产生的 5.1 声道声场之内的频率元素的空间布置。同样地, 能够使用输入和输出的其它适当组合, 诸如立体声到 4.1 声音、4.1 到 5.1 声音或者其它适当的组合。

图 8 是根据本发明的示范性实施例的用于从 M 个声道向 N 个声道上混频数据的系统 800 的示图。系统 800 将立体声时域数据转换成 7.1 声道时域数据。

系统 800 包括时间-频率分析单元 802 和 804、滤波生成单元 806、平滑单元 808 以及频率-时间综合单元 854 到 866。通过允许高分辨率频带处理的可伸缩频域结构, 并且通过提取并分析每一频带的重要声道间的空间提示以得到上混频 7.1 声道信号中频率元素的空间布置的滤波生成方法, 系统 800 在上混频过程中提供了改进的空间区别和稳定性。

系统 800 在将时域信号转换成频域信号的时间-频率分析单元 802 和 804 处接收左声道立体声信号  $L(T)$  和右声道立体声信号  $R(T)$ 。这些时间-频率分析单元能够是适当的滤波器组, 诸如有限脉冲响应 (FIR) 滤波器组、正交镜像滤波器 (QMF) 组、离散傅立叶变换 (DFT)、时域混叠消除 (TDAC) 滤波器组或其它适当的滤波器组。从时间-频率分析单元 802 和 804 输出的是一组频域值, 其覆盖了人类听觉系统的充分的频率范围, 如 0 到 20kHz 的频率范围, 其中分析滤波器组子带带宽能够被处理成近似心理声学的临界频带、等效矩形带宽或某种其它知觉特性。同样地, 能够使用其它适当数目的频带和范围。

来自时间-频率分析单元 802 和 804 的输出被提供给滤波生成单元 806。在一个示范性实施例中, 关于应当为给定的环境而输出的声道数目,

滤波生成单元 806 能够接收外部选择。例如，能够选择有两前和两后扬声器的 4.1 声道，能够选择有两前、两后和一个前中心扬声器的 5.1 音响系统，能够选择有两前、两侧、两后和一个前中心扬声器的 7.1 音响系统，或者能够选择其它适当的音响系统。滤波生成单元 806 在频带的基础上提取并分析诸如声道间级差 (ICLD) 和声道间相干性 (ICC) 之类的声道间空间提示。那些相关的空间提示然后用作参数以生成适应的声道滤波，其控制上混频声场中频带元素的空间布置。通过平滑单元 808 在整个时间和频率上使声道滤波平滑以限制滤波可变性，所述滤波可变性如果被允许变化太迅速的话，能造成讨厌的波动效应。在图 8 所示的示范性实施例中，左和右声道  $L(F)$  和  $R(F)$  频域信号被提供给滤波生成单元 806，其产生 7.1 声道滤波信号  $H_L(F)$ 、 $H_R(F)$ 、 $H_C(F)$ 、 $H_{LS}(F)$ 、 $H_{RS}(F)$ 、 $H_{LB}(F)$  和  $H_{RB}(F)$ ，它们被提供给平滑单元 808。

平滑单元 808 在整个时间和频率维数上为 7.1 声道滤波的每个声道平均化频域分量。在整个时间和频率上的平滑有助于控制声道滤波信号中的快速波动，因此，减少了能够令收听者讨厌的抖动现象和不稳定性。在一个示范性实施例中，通过对根据当前帧的每个频带和根据先前帧的相应频带应用一阶低通滤波，能够实现时间平滑。这具有逐帧减少每个频带的可变性的效果。在一个示范性实施例中，能够在整个频率槽的组上进行频谱平滑，所述频率槽被模拟成近似人类听觉系统的临界频带间隔。例如，如果使用具有均匀间隔的频率槽的分析滤波器组，则对于频谱的不同分区能够分组和平均化不同数目的频率槽。在这个示范性实施例中，从 0 到 5kHz 能够平均化 5 个频率槽，从 5kHz 到 10kHz 能够平均化 7 个频率槽，并且从 10kHz 到 20kHz 能够平均化 9 个频率槽，或者能够选择其它适当数目的频率槽和带宽范围。从平滑单元 808 输出  $H_L(F)$ 、 $H_R(F)$ 、 $H_C(F)$ 、 $H_{LS}(F)$ 、 $H_{RS}(F)$ 、 $H_{LB}(F)$  和  $H_{RB}(F)$  的平滑值。

用于 7.1 输出声道中的每一个的源信号  $X_L(F)$ 、 $X_R(F)$ 、 $X_C(F)$ 、 $X_{LS}(F)$ 、 $X_{RS}(F)$ 、 $X_{LB}(F)$  和  $X_{RB}(F)$  被生成为立体声输入声道的适应组合。在图 8 显示的示范性实施例中， $X_L(F)$  被简单地提供为  $L(F)$ ，意味着对于全部频带都有  $G_L(F) = 1$ 。同样地， $X_R(F)$  被简单地提供为  $R(F)$ ，意味着对于全部频带都有  $G_R(F) = 0$ 。作为从加法器 814 输出的  $X_C(F)$  被计算为信号  $L(F)$  乘以适应缩放信号  $G_C(F)$  与  $R(F)$  乘以适应缩放信号  $1 - G_C(F)$  的和。作为从加法器 820 输出的  $X_{LS}(F)$  被计算为信号  $L(F)$  乘以适应缩放信号  $G_{LS}(F)$  与  $R(F)$  乘以适应缩放信号  $1 - G_{LS}(F)$  的和。同样地，作为从加法器 826 输出的  $X_{RS}(F)$  被计算为信号  $L(F)$  乘以适应



缩放信号  $G_{RS}(F)$  与  $R(F)$  乘以适应缩放信号  $1 - G_{RS}(F)$  的和。同样地, 作为从加法器 832 输出的  $X_{LB}(F)$  被计算为信号  $L(F)$  乘以适应缩放信号  $G_{LB}(F)$  与  $R(F)$  乘以适应缩放信号  $1 - G_{LB}(F)$  的和。同样地, 作为从加法器 838 输出的  $X_{RB}(F)$  被计算为信号  $L(F)$  乘以适应缩放信号  $G_{RB}(F)$  与  $R(F)$  乘以适应缩放信号  $1 - G_{RB}(F)$  的和。注意, 如果对于全部频带都有  $G_C(F) = 0.5$ 、 $G_{LS}(F) = 0.5$ 、 $G_{RS}(F) = 0.5$ 、 $G_{LB}(F) = 0.5$  和  $G_{RB}(F) = 0.5$ , 那么前中心声道源自  $L(F) + R(F)$  组合, 并且侧和后声道源自成比例的  $L(F) - R(F)$  组合, 如传统矩阵上混频方法中通用的那样。适应缩放信号  $G_C(F)$ 、 $G_{LS}(F)$ 、 $G_{RS}(F)$ 、 $G_{LB}(F)$  和  $G_{RB}(F)$  能够进一步提供一方法以动态调整相邻输出声道对之间的相关性, 无论它们是横向还是纵向的声道对。声道源信号  $X_L(F)$ 、 $X_R(F)$ 、 $X_C(F)$ 、 $X_{LS}(F)$ 、 $X_{RS}(F)$ 、 $X_{LB}(F)$  和  $X_{RB}(F)$  分别通过乘法器 840 到 852 与平滑的声道滤波  $H_L(F)$ 、 $H_R(F)$ 、 $H_C(F)$ 、 $H_{LS}(F)$ 、 $H_{RS}(F)$ 、 $H_{LB}(F)$  和  $H_{RB}(F)$  相乘。

来自乘法器 840 到 852 的输出然后被频率-时间综合单元 854 到 866 从频域转换到时域以生成输出声道  $Y_L(T)$ 、 $Y_R(T)$ 、 $Y_C(F)$ 、 $Y_{LS}(F)$ 、 $Y_{RS}(T)$ 、 $Y_{LB}(T)$  和  $Y_{RB}(T)$ 。以这种方式, 左和右立体声信号被上混频到 7.1 声道信号, 其中, 自然存在的或者如通过图 1 的下混频水印过程或其它适当过程那样的被有意编码到左和右立体声信号的声道间空间提示, 能够用于控制系统 800 产生的 7.1 声道声场内的频率元素的空间布置。同样地, 能够使用输入和输出的其它适当组合, 诸如立体声到 5.1 声音、5.1 到 7.1 声音或者其它适当的组合。

图 9 是根据本发明示范性实施例的用于生成用于频域应用的滤波的系统 900 的示图。滤波生成过程使用  $M$  声道输入信号的频域分析和处理。为  $M$  声道输入信号的每个频带提取相关声道间空间提示, 并且为每个频带生成空间位置矢量。对于理想收听条件下的收听者, 这个空间位置矢量被解释为用于该频带的感觉源定位。然后生成每个声道滤波, 以便用于上混频  $N$  声道输出信号中的该频率元素的作为结果的空间位置与声道间提示一致地再生。声道间级差 (ICLD) 和声道间相干性 (ICC) 的估计用作声道间提示以产生空间位置矢量。

在系统 900 中所示的示范性实施例中, 子带幅度或能量分量被用于估计声道间级差, 并且子带相角分量用于估计声道间相干性。左和右频域输入  $L(F)$  和  $R(F)$  被转换成幅度或能量分量和相角分量, 其中, 幅度/能量分量被提供给加法器 902, 其计算总能量信号  $T(F)$ ,  $T(F)$  然后被

用于通过除法器 904 和 906 分别为每个频带标准化左  $M_L(F)$  和右声道  $M_R(F)$  的幅度/能量值。然后, 根据  $M_L(F)$  和  $M_R(F)$  计算标准化的横向坐标信号  $LAT(F)$ , 其中, 用于频带的标准化横向坐标被计算为:

$$LAT(F) = M_L(F) * X_{MIN} + M_R(F) * X_{MAX}$$

同样地, 标准化深度坐标根据输入的相角分量而被计算为:

$$DEP(F) = Y_{MAX} - 0.5 * (Y_{MAX} - Y_{MIN}) * \sqrt{[\cos(\angle L(F)) - \cos(\angle R(F))]^2 + [\sin(\angle L(F)) - \sin(\angle R(F))]^2}$$

标准化深度坐标基本上根据相角分量  $\angle L(F)$  和  $\angle R(F)$  之间的经缩放的和移位的距离测量结果来计算。当相角  $\angle L(F)$  和  $\angle R(F)$  在单位圆上相互接近时,  $DEP(F)$  的值接近 1, 并且当相角  $\angle L(F)$  和  $\angle R(F)$  接近单位圆的相对侧时,  $DEP(F)$  接近 0。对于每个频带, 标准化的横向坐标和深度坐标形成 2 维矢量 ( $LAT(F)$ ,  $DEP(F)$ ), 其被输入到 2 维声道图中, 如以下图 10A 到 10E 中所示的那样, 以产生用于每个声道  $i$  的滤波值  $H_i(F)$ 。从诸如图 6 的滤波生成单元 606、图 7 的滤波生成单元 706 和图 8 的滤波生成单元 806 之类的滤波生成单元输出这些用于每个声道  $i$  的声道滤波  $H_i(F)$ 。

图 10A 是根据本发明示范性实施例的用于左前信号的滤波图的示图。在图 10A 中, 滤波图 1000 接受从 0 到 1 范围的标准化的横向坐标和从 0 到 1 范围的标准化的深度坐标, 并且输出从 0 到 1 范围的标准化的滤波值。灰度阴影用于指示幅度从最大 1 到最小 0 的变化, 如通过滤波图 1000 的右手侧的标度所示。对于这个示范性的左前滤波图 1000, 接近 (0, 1) 的标准化的横和深度坐标将输出接近 1.0 的最高滤波值, 而从近似 (0.6, Y) 到 (1.0, Y) 范围的坐标, 其中 Y 为 0 和 1 之间的数, 将基本上输出 0 的滤波值。

图 10B 是示范性右前滤波图 1002 的示图。滤波图 1002 接受与滤波图 1000 相同的标准化的横向坐标和标准化的深度坐标, 但是输出的滤波值偏向标准化的布局的右前部。

图 10C 是示范性中心滤波图 1004 的示图。在这个示范性实施例中, 用于中心滤波图 1004 的最大滤波值发生在标准化的布局的中心, 随着坐标从布局的前部中心离开朝向布局的后部移动, 幅度显著下降。

图 10D 是示范性左环绕滤波图 1006 的示图。在这个示范性实施例中，用于左环绕滤波图 1006 的最大滤波值发生在标准化布局的后左坐标附近，并且随着坐标向布局的前右侧移动而幅度下降。

图 10E 是示范性右环绕滤波图 1008 的示图。在这个示范性实施例中，用于右环绕滤波图 1008 的最大滤波值发生在标准化布局的后右坐标附近，并且随着坐标向布局的前左侧移动而幅度下降。

同样地，如果使用其它的扬声器布局或配置，那么能够调整现有的滤波图，并且能够生成对应于新的扬声器定位的新的滤波图，以反映新的收听环境的变化。在一个示范性实施例中，7.1 系统将包括两个另外的滤波图，其具有在深度坐标维度上向上移动的左环绕和右环绕，并且具有左后和右后定位，具有分别类似于滤波图 1006 和 1008 的滤波图。能够改变滤波因数下降的速率以适应不同数目的扬声器。

尽管在此已详细地描述了本发明的系统和方法的示范性实施例，但是本领域技术人员还会认识到，能够对系统和方法进行各种替换和修改，而不背离附加权利要求的范围和精神。

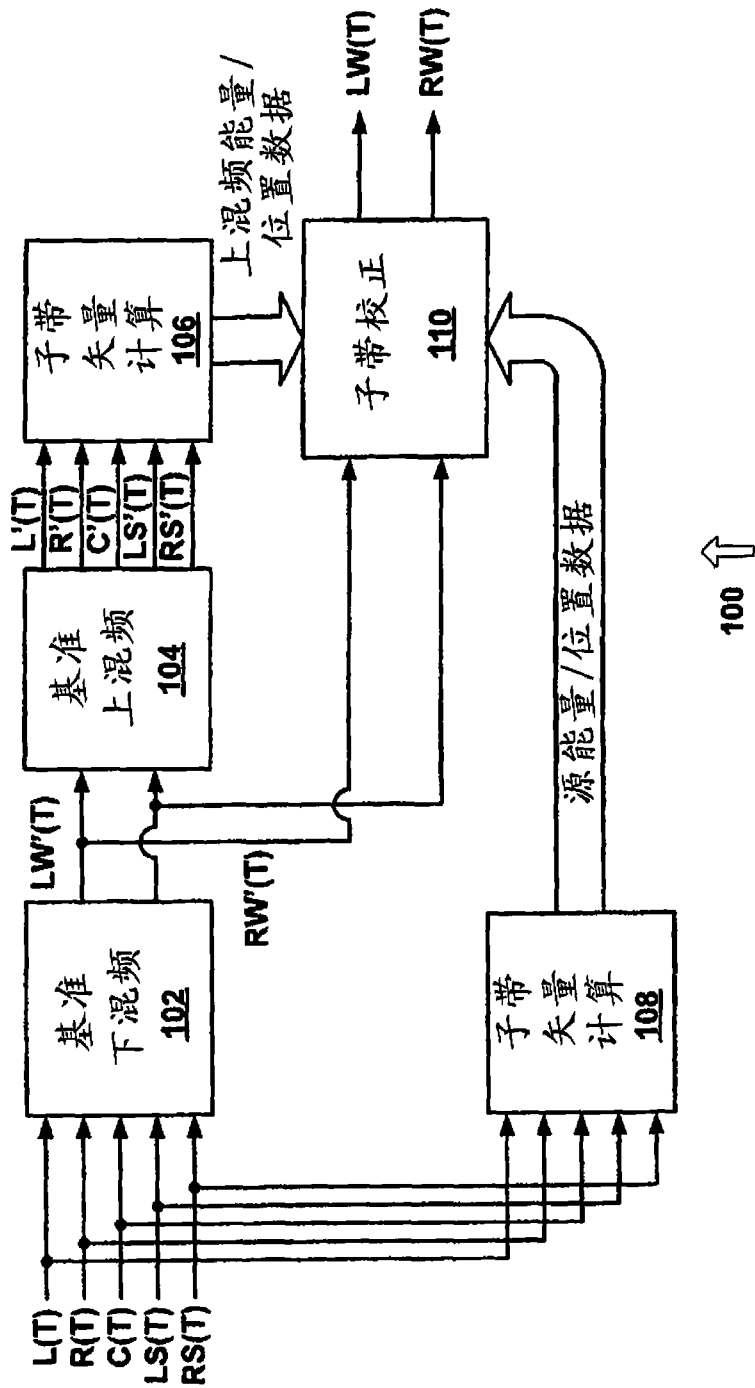


图1

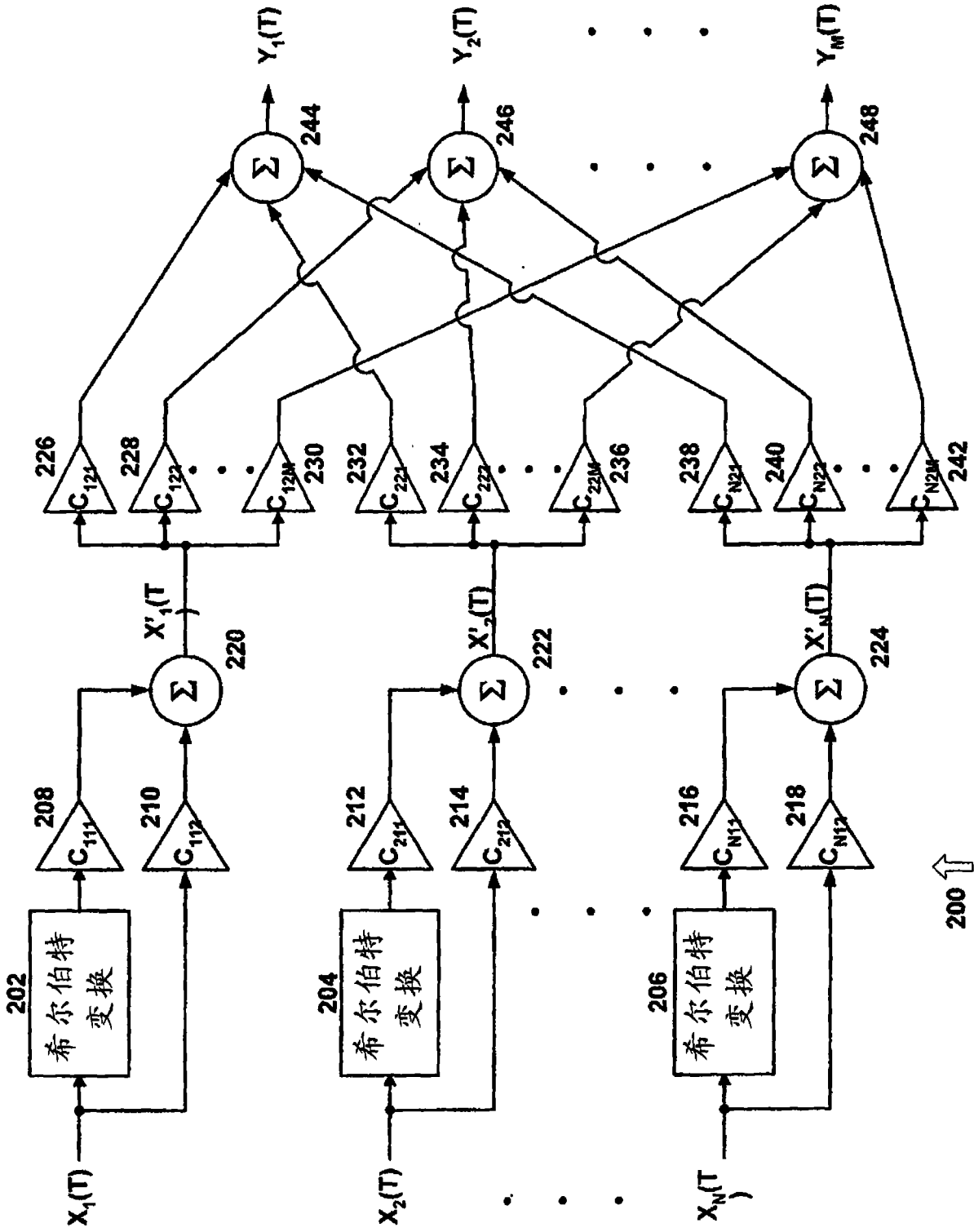


图2

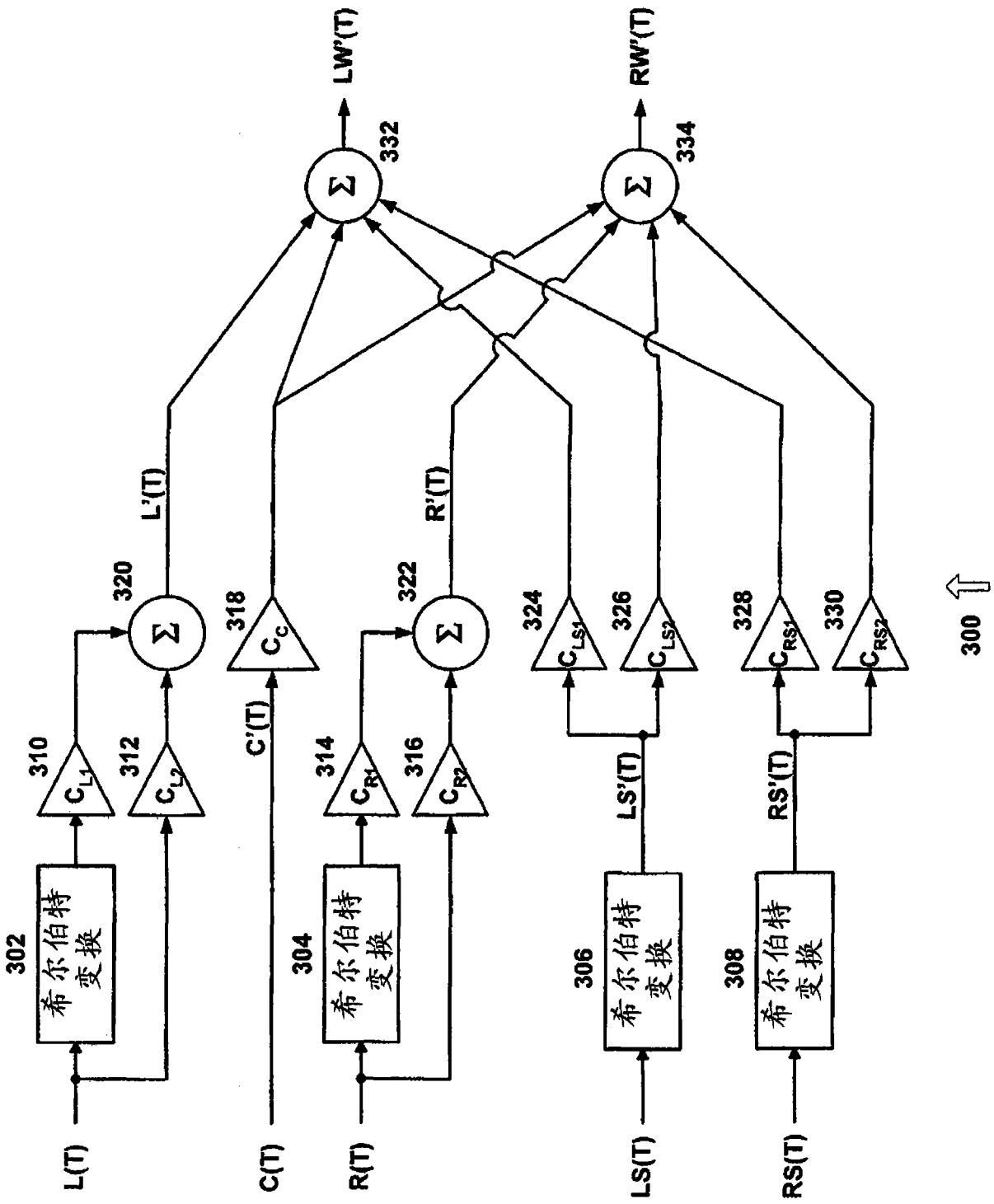
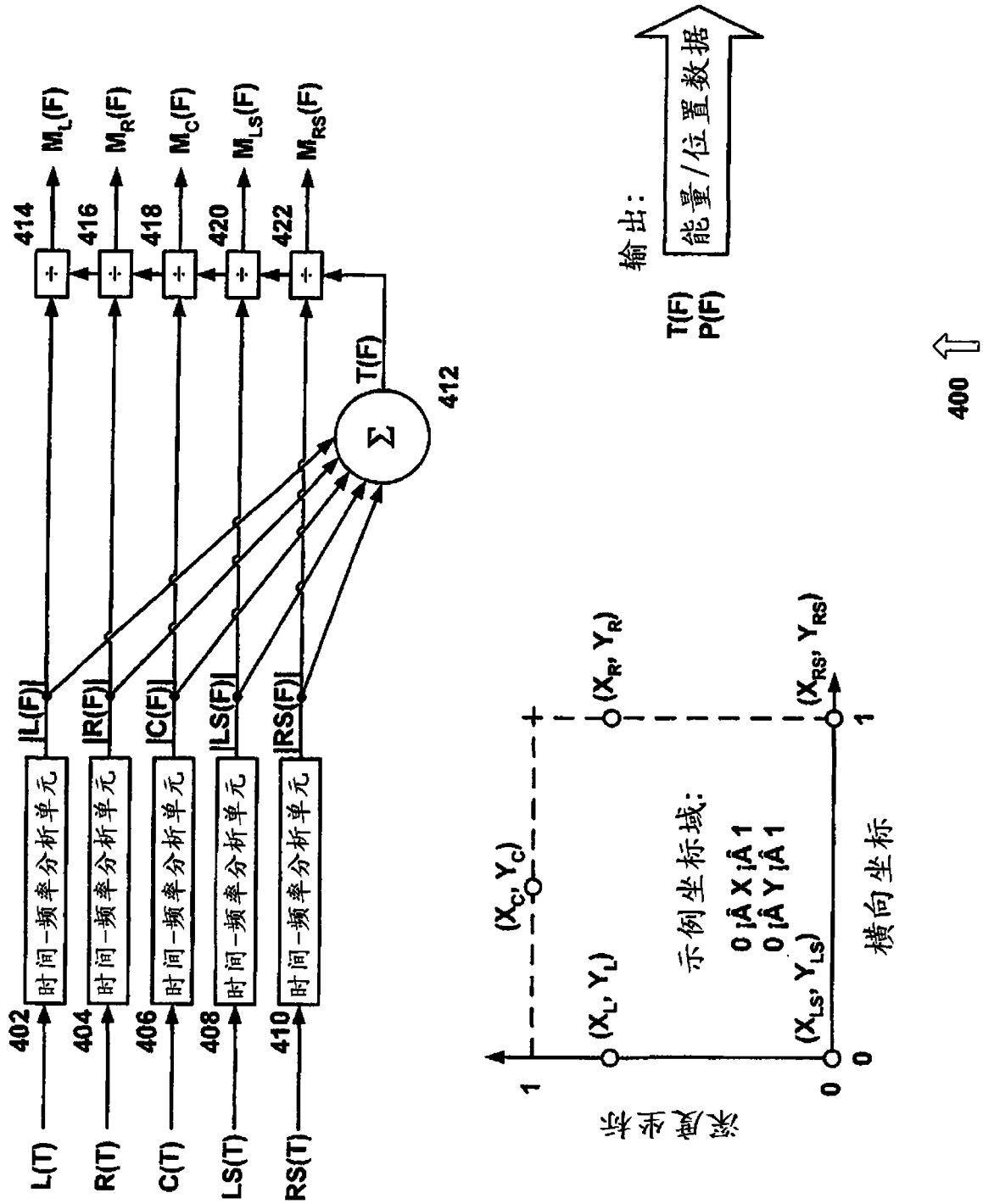


图 3



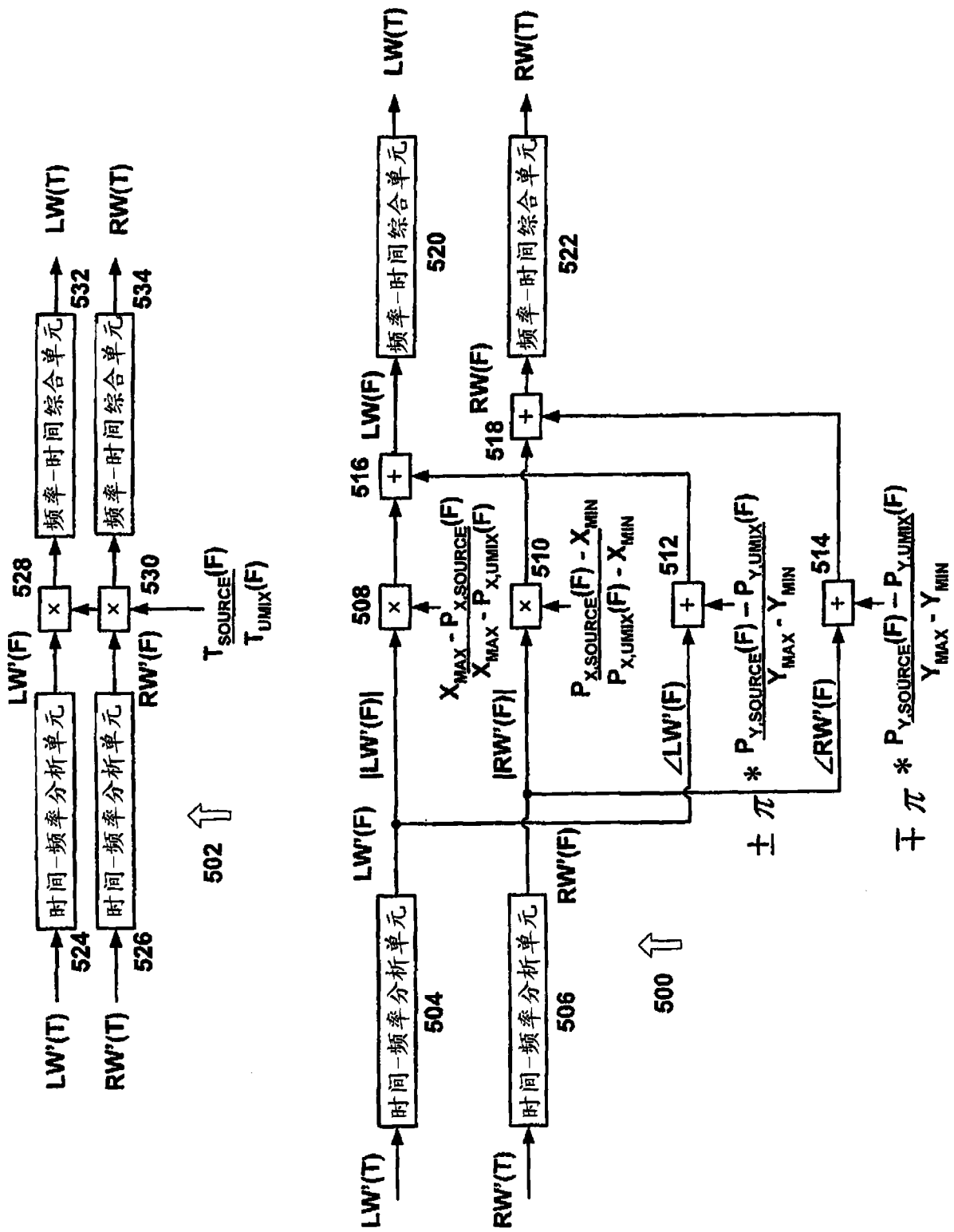


图5



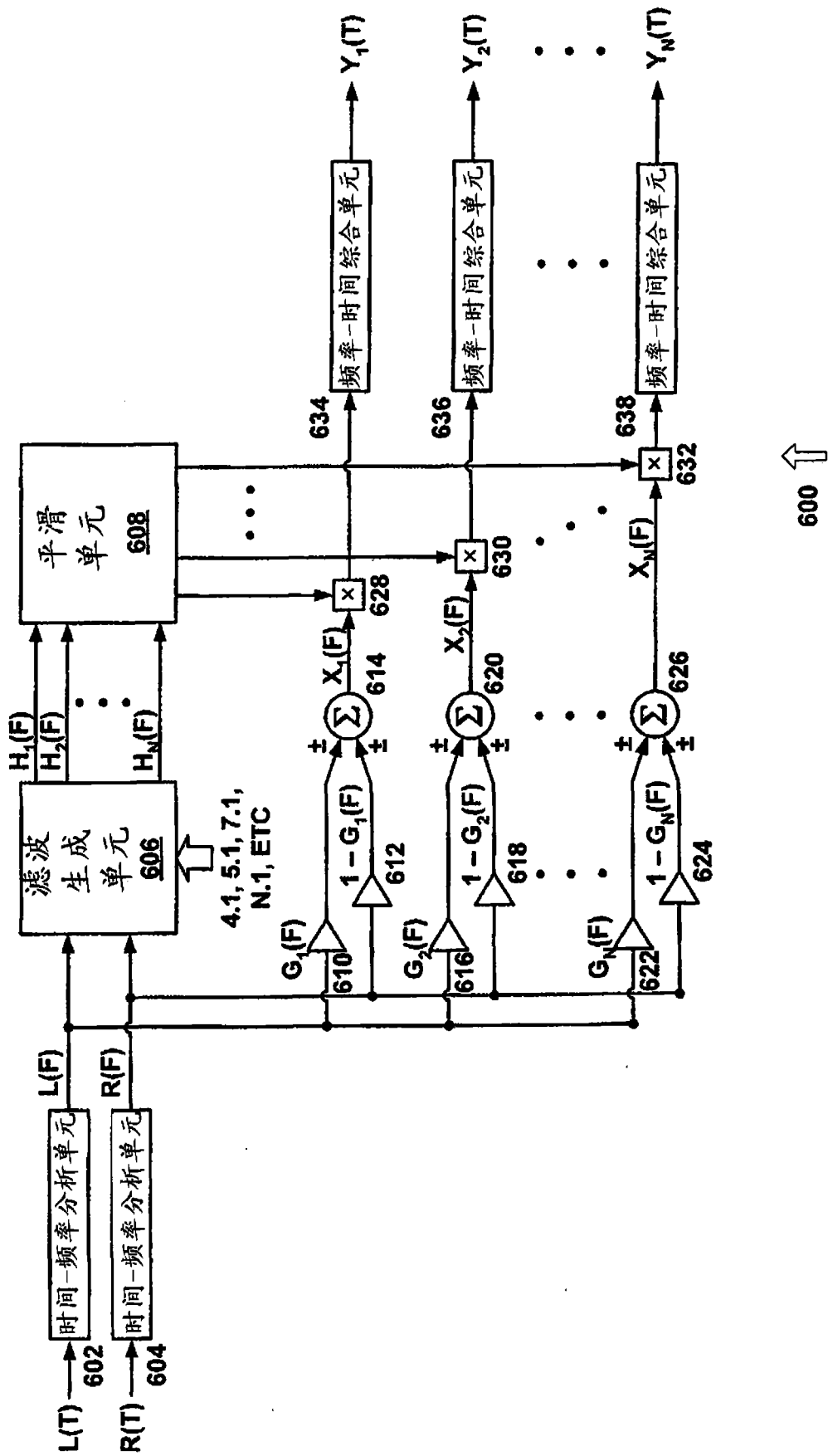


图 6

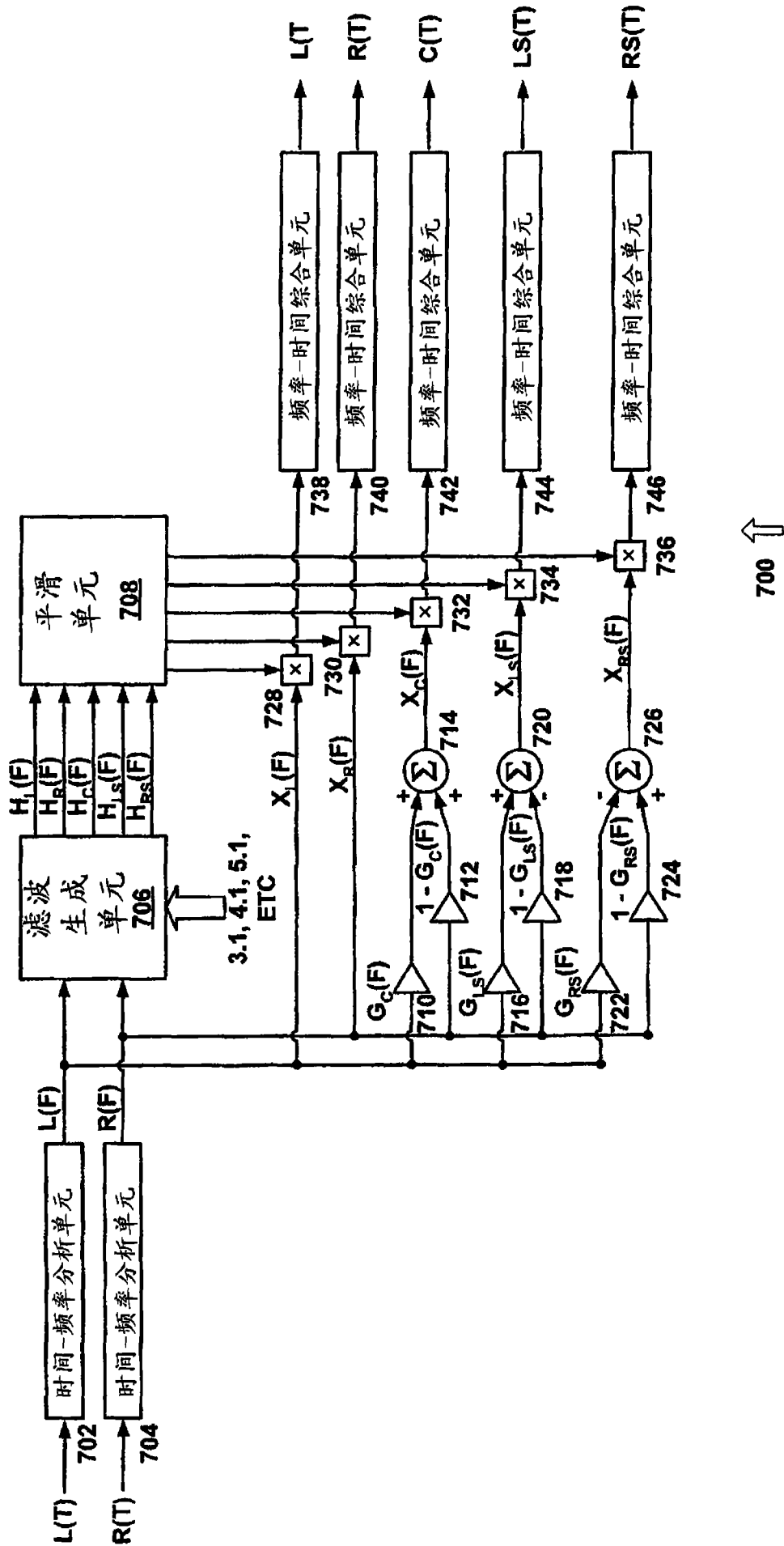
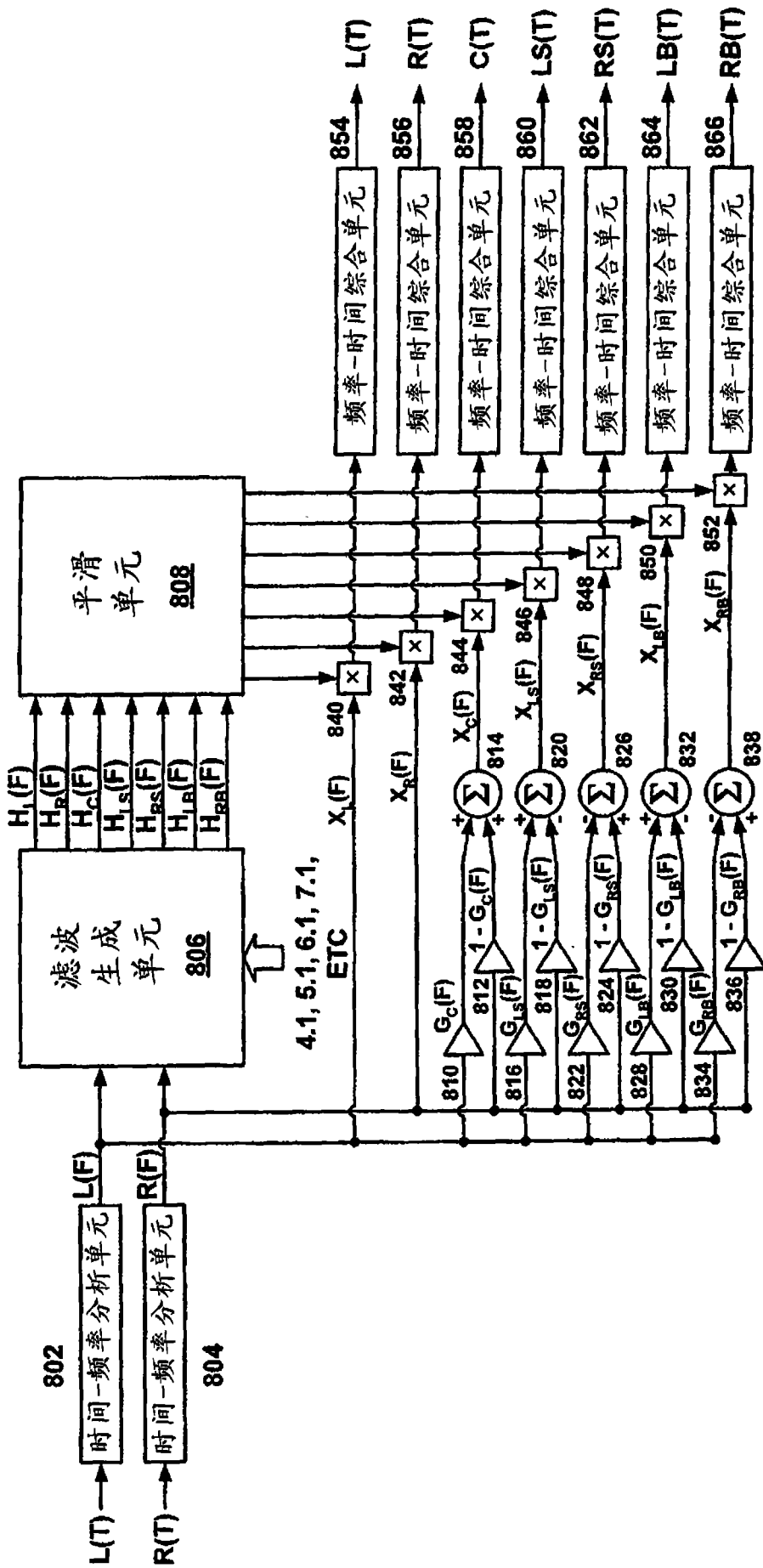


图7



800

图 8

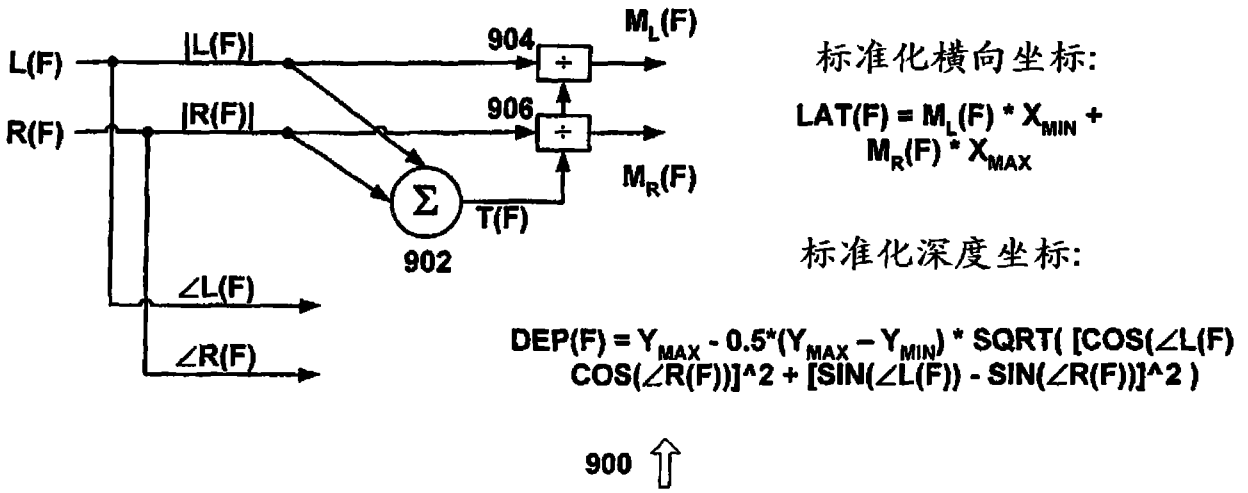


图9

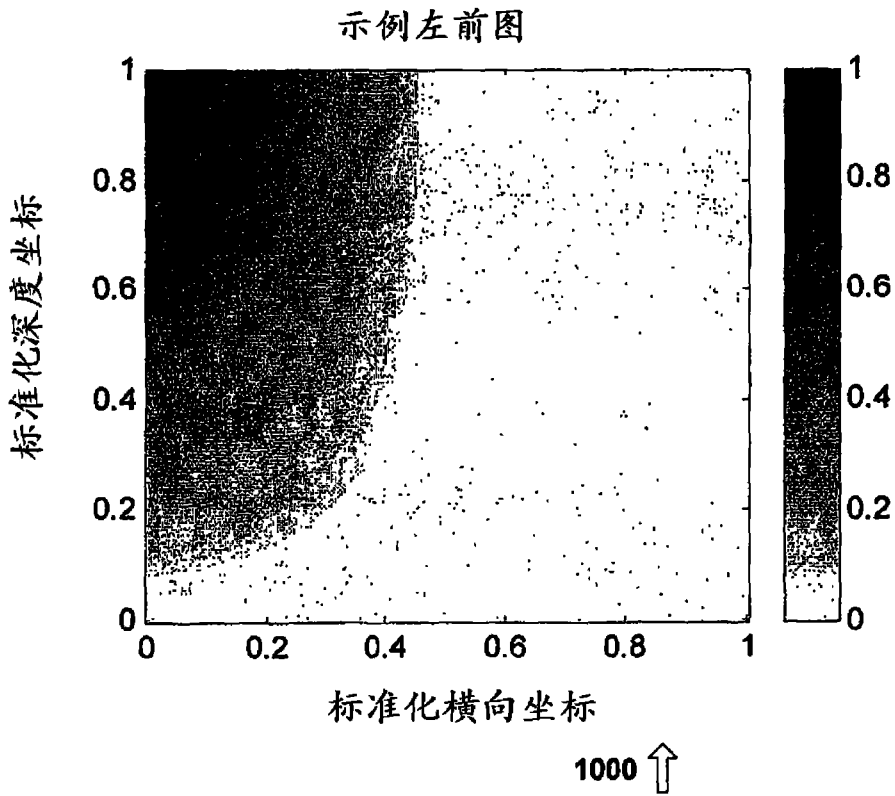


图10A

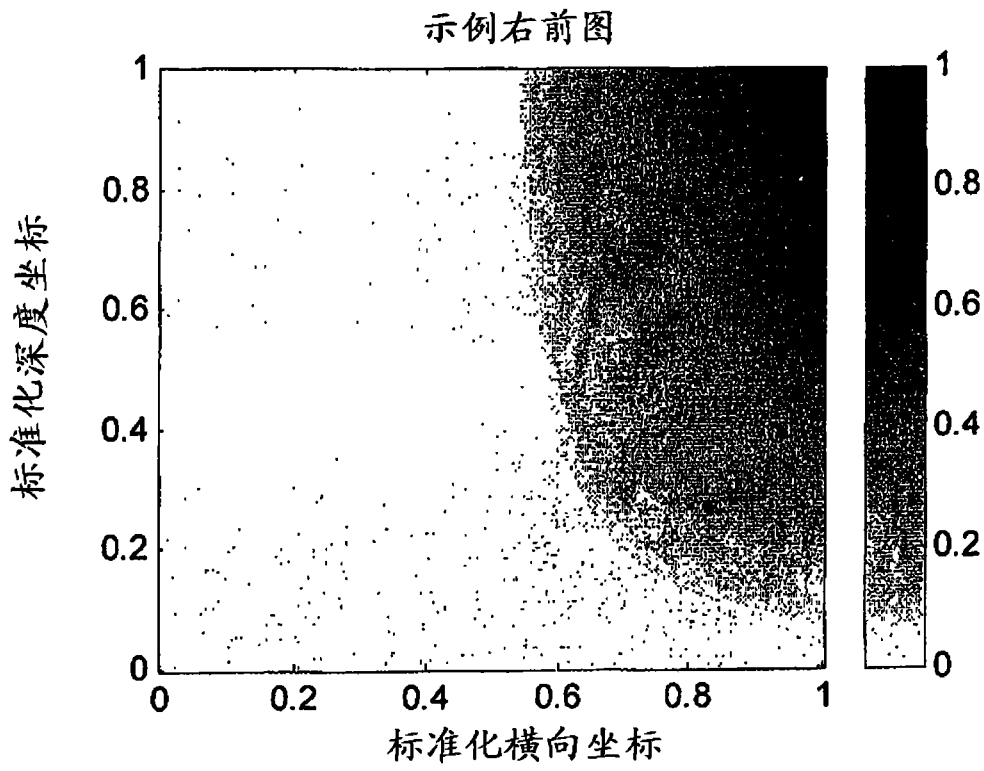


图10B

1002 ↑

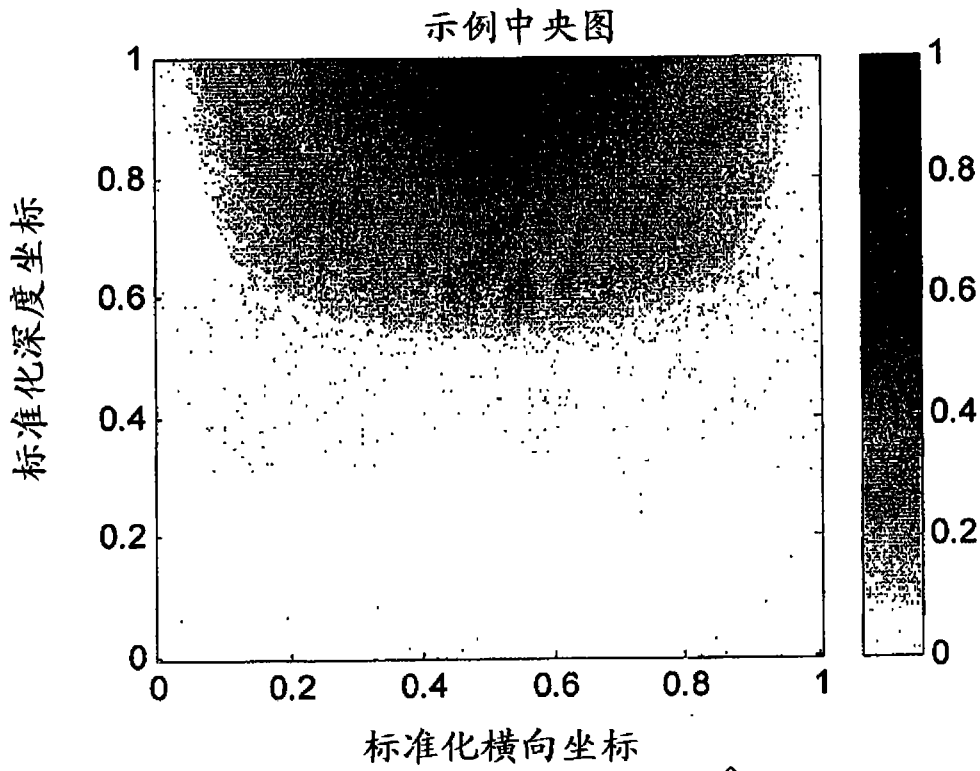


图10C

1004 ↑

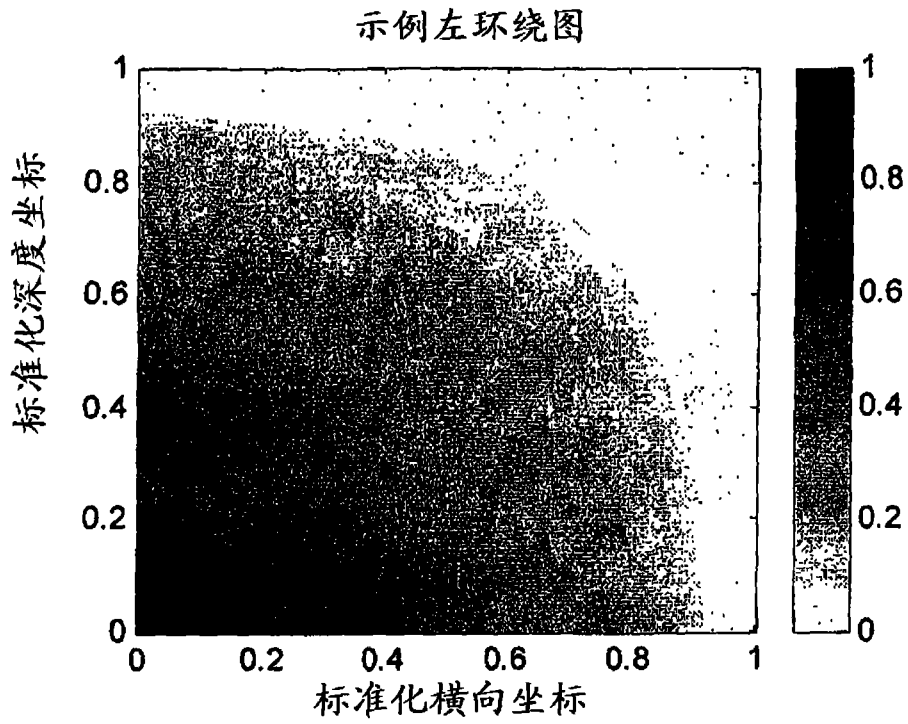
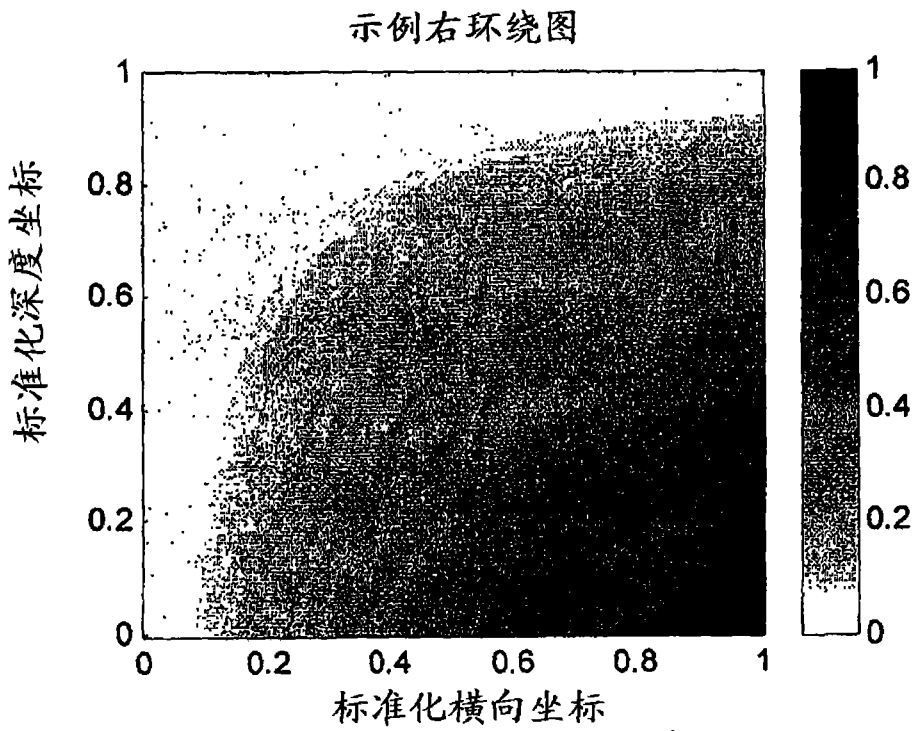


图10D 1006 ↑



1008 ↑  
图10E