[19] 中华人民共和国国家知识产权局

[51] Int. Cl[7]
G06T 15/70

[12] 发明专利申请公开说明书

[21] 申请号 01822832.1

[43] 公开日 2004 年 10 月 13 日

[11] 公开号 CN 1537300A

[22] 申请日 2001.12.21 [21] 申请号 01822832.1
[30] 优先权
[32] 2000.12.22 [33] GB [31] 0031511.9
[32] 2001. 7.20 [33] GB [31] 0117770.8
[32] 2001. 8.10 [33] GB [31] 0119598.1
[86] 国际申请 PCT/GB2001/005719 2001.12.21
[87] 国际公布 WO2002/052863 英 2002.7.4
[85] 进入国家阶段日期 2003.8.22
[71] 申请人 人类技术有限公司
地址 英国伦敦
[72] 发明人 B·J·吉勒特 C·S·维勒斯
M·J·威廉斯 G·M·斯利特

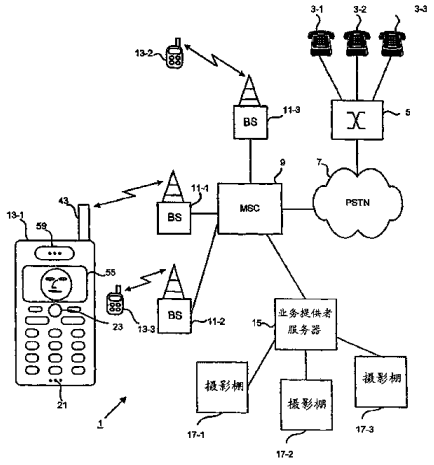
[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 栾本生 罗 朋

权利要求书 10 页 说明书 22 页 附图 17 页

[54] 发明名称 通信系统

[57] 摘要

描述了一种电话系统，其中订户电话为电话呼叫的一方的外貌存储合适的模型，从中其根据从电话网接收的一组外貌参数合成该方的视频序列。外貌参数可从与用户的电话相关的摄像机被生成或者从由该方输入的文本或语音信号被生成。



ISSN1008-4274

1. 一种用于电话网的电话，所述电话包括：

用于存储定义了一个功能的模型数据的存储器，所述功能将一组参数的一个或多个参数与定义对象的形状标准化外貌的纹理数据相关并且其将所述一组参数的一个或多个参数与为所述对象定义形状的形状数据相关；

用于接收表示视频序列的多组参数的装置；

用于对于至少一组接收参数生成定义对象的形状标准化外貌的纹理数据以及用于对于多组接收参数生成对象的形状数据的装置；

用于利用生成的形状数据变形生成的纹理数据以便在视频序列的帧中生成定义对象的外貌的图像数据的装置；以及

用于驱动显示器输出生成的图像数据以便合成视频序列的显示驱动器。

2. 根据权利要求 1 所述的电话，其中从一组参数生成的形状数据包括识别对应于所述接收的一组参数的视频帧中对象上多个预定点的相对位置的一组位置。

3. 根据权利要求 2 所述的电话，其中所述变形装置可操作地识别表示形状标准化对象的所述纹理数据中对象上所述多个预定点的位置，并且可操作地变形纹理数据，以便所述预定点的确定的位置被变形到由所述形状数据所定义的相应点的位置。

4. 根据前面任何一个权利要求所述的设备，其中所述生成装置可操作地对于每组接收到的参数生成定义所述对象的形状标准化外貌的纹理数据以及对于所述对象的形状数据，并且其中所述变形装置可操作地利用从所述一组参数生成的相应形状数据来变形对于每组参数的生成的纹理数据。

5. 根据权利要求 1 到 3 的任何一个所述的设备，其中所述生成装置可操作地对于选择的各组所述接收的参数生成纹理数据，并且如果所述生成装置没有对当前的一组接收参数生成纹理数据，则所述变形装置可操作地利用对于当前一组接收参数的形状数据对于以前的一组参数变形纹理数据。

6. 根据权利要求 5 所述的电话，包括选择装置，用于从所述接收的多组参数中选择各组参数，所述生成装置将为其生成纹理数据。

7. 根据权利要求 6 所述的电话，其中所述选择装置可操作地根据预定规则从接收的多组参数中选择各组参数。

8. 根据权利要求 6 到 7 所述的电话，包括用于将来自当前一组参数的参数值与以前一组参数的参数值相比较的装置，并且其中所述选择装置可操作地根据所述比较的结果来选择所述当前一组参数。

9. 根据权利要求 8 所述的电话，其中如果所述当前组的一个或多个所述参数与以前组的相应参数值差异超过预定阈值，则所述选择装置可操作地选择所述当前组参数。

10. 根据权利要求 6 到 9 的任何一个所述的电话，其中所述选择装置可操作地选择所述生成装置将根据电话的可用处理能力来为其生成所述纹理数据的所述各组参数。

11. 根据权利要求 10 所述的电话，其中每个参数表示对于所述对象的一种纹理变化模式，并且其中所述选择装置可操作地选择能够在可用的处理能力基本上为实时的情况下而转换成为文本数据的那样多的最重要的变化模式。

12. 根据权利要求 1 到 3 的任何一个所述的设备，包括用于将来自当前一组参数的参数值与以前一组参数的参数值比较的装置，并且其中所述变形装置可操作地对于变化最大的 N 个参数值变形纹理数据。

13. 根据权利要求 12 所述的电话，其中根据可用处理能力来确定 N。

14. 根据权利要求 12 或 13 所述的电话，其中所述生成装置可操作地通过对于以前的一组参数利用确定的所述 N 个参数的差值来更新形状标准化纹理数据，从而生成形状标准化纹理的数据。

15. 根据前面任何一个权利要求所述的电话，其中所述模型数据包括将一组接收的参数关联到一组中间形状参数的第一个模型数据以及一组中间纹理参数；其中所述模型数据还包括定义将中间形状参数关联到所述形状数据的功能的第二个模型数据；其中所述模型数据还包括定义将中间纹理参数关联到所述纹理数据的功能的第三个模型数据；并且其中所述生成装置包括用于对于从利用第一个模型数据的电话网发送的每组接收的参数利用第一个模型数据生成一组中间形状和纹理参数的装置。

16. 根据前面任何一个权利要求所述的电话，其中所述接收装置可操作地从电话网接收所述模型数据并且还包括用于在所述存储器中存储所述接收的模型数据的装置。

17. 根据权利要求 16 所述的电话，其中所述接收模型数据被编码并且还包括用于对模型数据解码的装置。

18. 根据权利要求 17 所述的电话，其中通过将预定的各组参数应用到模型数据以便对于每个预定组目标参数推导出相应的纹理数据并且通过压缩从所述组参数生成的确定的纹理数据，从而对于所述模型数据进行编码；并且其中所述解码器包括用于解压缩所述被压缩的纹理数据的装置以及用于利用所述解压缩的纹理数据和预定的各组参数来重新合成所述模型数据的装置。

19. 根据前面任何一个权利要求所述的电话，还包括用于接收与视频序列相关的音频信号的装置以及用于与视频序列同步地向用户输出音频信号的装置。

20. 根据权利要求 19 所述的电话，其中所述音频信号和所述各组参数是彼此交错的。

21. 根据前面任何一个权利要求所述的电话，包括用于接收语音的装置以及用于处理语音以便生成表示所述视频序列的所述多组参数的装置，并且其中所述接收装置可操作地从所述语音处理装置接收所述参数。

22. 根据权利要求 21 所述的电话，其中所述语音处理装置包括用于将接收的语音转换成子字单元的序列的语音识别单元以及用于将所述子字单元序列转换成表示所述视频序列的所述多组参数的装置。

23. 根据权利要求 22 所述的电话，其中所述转换装置包括一个查找表，用于将每个子字单元转换成表示所述视频序列的一帧的相应的一组参数。

24. 根据权利要求 23 所述的电话，其中所述转换装置包括每个与所述对象的不同的情绪状态相关的多个查找表并且还包括用于选择查找表之一以便根据所述对象的被检测的情绪状态来执行所述转换的装置。

25. 根据权利要求 24 所述的电话，其中所述处理装置可操作地处理所述语音以便确定所述对象的情绪状态并且可操作地选择由所述转

换装置使用的相应的查找表。

26. 根据权利要求 1 到 18 的任何一个所述的电话，包括用于接收文本的装置以及用于处理接收的文本以便生成表示对应于说出所述文本的对象的视频序列的各组参数的装置，并且其中所述接收装置可操作地从所述文本处理装置接收所述多组参数。

27. 根据权利要求 26 所述的电话，还包括用于合成对应于文本的语音的文本到语音合成器以及用于输出与相应的视频序列同步的合成的语音的装置。

28. 根据权利要求 26 或 27 的电话，其中所述文本处理装置包括用于将接收的文本转换成子字单元序列的装置以及用于将子字单元序列转换成所述多组参数的装置。

29. 根据前面任何一个权利要求所述的电话，还包括用于存储表示预定视频序列的各组参数的存储器并且还包括用于接收触发信号的装置，所述生成装置响应所述触发信号而可操作地为存储的多组参数生成纹理数据和形状数据。

30. 根据前面任何一个权利要求所述的电话，还包括用于存储定义从一组接收的参数到一组被转换的参数的转换的转换数据的装置以及用于利用所述转换数据改变帧中所述对象的外貌的装置。

31. 根据前面任何一个权利要求所述的电话，还包括:

用于存储定义将第二个对象的图像数据关联到一组参数的功能的第二个模型数据的第二个存储器;

用于为所述第二个对象接收图像数据的装置;

用于利用图像数据和第二个模型数据来为所述第二个对象确定一组参数的装置;

用于将对于第二个对象的确定的一组参数发送到所述电话网的装置。

32. 根据权利要求 31 所述的电话，其中所述图像数据接收装置可操作地接收对应于视频序列的图像数据，其中所述参数确定装置可操作地对于视频序列中的第二个对象确定多组参数，并且其中所述发送装置可操作地将对于所述第二个对象的所述多组参数发送到所述电话网。

33. 根据权利要求 31 或 32 所述的电话，还包括用于感觉来自第

二个对象的光线以及用于从中生成所述图像数据的装置。

34. 根据权利要求 31 到 33 的任何一个所述的电话，其中所述发送装置可操作地向电话网发送所述第二个模型数据以便发送到主叫方或要被呼叫的一方。

35. 根据权利要求 1 到 30 的任何一个所述的电话，包括用于从用户接收语音的麦克风；用于处理被接收的语音以便生成表示用户的外貌的一组参数的装置以及用于向电话网发送表示用户的外貌的参数的装置。

36. 根据权利要求 35 所述的电话，其中所述处理装置包括用于将用户的语音转换成子字单元序列的自动语音识别单元以及用于将子字单元序列转换成表示用户的外貌的所述各组参数的装置。

37. 根据权利要求 36 所述的电话，其中所述转换装置包括查找表，用于将每个子字单元转换成表示用户的外貌的相应的一组参数，同时发出相应的子字单元的音。

38. 根据权利要求 1 到 34 的任何一个所述的电话，还包括用于接收来自用户的文本的装置，用于处理被接收的文本以便生成表示说出所述文本的用户的外貌的各组参数的装置以及用于将表示用户的外貌的参数发送到电话网的装置。

39. 根据权利要求 38 所述的电话，其中所述文本处理装置包括用于将被接收的文本转换为子字单元序列的第一个转换装置以及用于将子字单元序列转换成所述多组参数的第二个转换装置。

40. 根据前面任何一个权利要求所述的电话，其中所述纹理数据定义了对象的形状标准化的颜色外貌。

41. 根据权利要求 40 所述的电话，其中所述纹理数据包括单独的红色纹理数据、绿色纹理数据和蓝色纹理数据。

42. 根据前面任何一个权利要求所述的电话，其中所述对象是表示呼叫的一方的面部。

43. 根据权利要求 42 所述的电话，其中所述生成装置可操作地为面部的眼睛、面部的嘴巴以及面部区域的剩余部分生成单独的纹理数据。

44. 根据权利要求 38 所述的电话，其中每组参数包括各自的参数子集，每个子集与面部的眼睛、面部的嘴巴以及面部区域的剩余部分

相关。

45. 根据权利要求 43 或 44 所述的电话，其中对于面部区域的剩余部分的所述纹理数据是不变的纹理。

46. 一种供电话网使用的电话，所述电话包括：

用于接收来自用户的语音信号的装置；

用于处理被接收的语音信号以便生成表示说出所述语音的用户的外貌的多组参数的装置；以及

用于向电话网发送表示用户的外貌的参数的装置。

47. 根据权利要求 46 所述的电话，其中所述处理装置包括用于将用户的语音转换到子字单元序列的自动语音识别单元以及用于将子字单元序列转换到表示用户的外貌的所述各组参数的装置。

48. 根据权利要求 47 所述的电话，其中所述转换装置包括用于将每个子字单元转换成表示用户的外貌的相应的一组参数同时发音相应的子字单元的查找表。

49. 根据权利要求 48 所述的电话，其中所述转换装置包括多个查找表并且其中所述语音处理装置可操作地从所述的接收语音信号确定用户的情绪并且可操作地选择查找表供所述转换装置使用。

50. 一种与电话网一起使用的电话，所述电话包括：

用于接收来自用户的文本的装置；

用于处理被接收的文本以便生成表示说出所述文本的用户的外貌的多组参数的装置；以及

用于向电话网发送表示用户的外貌的参数的装置。

51. 根据权利要求 50 所述的电话，其中所述文本处理装置包括用于将被接收的文本转换成子字单元序列的第一个转换装置以及用于将子字单元序列转换成所述多组参数的第二个转换装置。

52. 根据权利要求 51 所述的电话，其中所述第二个转换装置包括查找表，用于将每个子字单元转换成表示用户的外貌的相应的一组参数，同时发出相应的子字单元的音。

53. 根据权利要求 52 所述的电话，其中所述第二个转换装置包括多个查找表，每个与用户的各个不同的情绪相关联；并且还包括用于感觉用户的当前情绪以及用于选择相应的查找表供所述转换装置使用的装置。

54. 一种供GSM网使用的GSM电话，该GSM电话包括:

用于对音频数据编码的GSM音频编解码器;

用于接收音频数据和视频数据的装置;

用于混合音频数据和视频数据以便生成音频和视频数据的混合流的装置;

用于利用所述音频编解码器对音频和视频数据的混合流进行编码的装置; 以及

用于将所述被编码的音频和视频数据发送到所述电话网的装置。

55. 一种用于控制第一和第二个订户电话之间的通信链路的电话网服务器，所述电话网服务器包括:

用于为第一个订户存储定义一个功能的模型数据的存储器，所述功能将一组参数的一个或多个参数与定义与第一个订户相关的对象的形状标准化外貌的纹理数据关联并且将所述一组参数的一个或多个参数与定义与第一个订户相关的对象的形状的形状数据关联;

用于接收指示呼叫被在所述第一和第二个订户之间启动的信号的装置;

响应所述信号用于将对于所述第一个订户的所述模型数据发送到第二个订户的电话的装置。

56. 据权利要求55所述的电话网服务器，其中所述存储器还包括用于所述第二个订户的模型数据并且其中所述发送装置可操作地向所述第一个订户的电话发送对于所述第二个订户的模型数据。

57. 根据权利要求55或56所述的电话网服务器，还包括用于生成表示视频序列的多组参数的装置，从中视频序列可被利用所述模型数据来合成，以及用于向所述第一或第二个订户的电话发送所述各组参数的装置。

58. 根据权利要求57所述的电话网服务器，其中所述生成装置可操作地从接收自所述第一个订户的电话的语音信号生成所述多组参数。

59. 根据权利要求58所述的电话网服务器，还包括用于处理所述被接收的语音信号以及用于生成表示被接收的语音的子字单元序列的装置以及用于将所述子字单元序列转换成所述多组参数的装置。

60. 根据权利要求56所述的电话网服务器，其中所述生成装置包

括用于从第一个订户的电话接收文本的装置，用于将被接收的文本转换成子字单元序列的第一转换装置；以及用于将子字单元序列转换成所述多组参数的装置。

61. 根据权利要求 59 或 60 所述的电话网服务器，其中所述转换装置包括将每个子字单元关联到相应的一组参数的查找表。

62. 一种电话网，包括根据权利要求 55 到 60 的任何一个所述的电话网服务器以及根据权利要求 1 到 54 的任何一个所述的多个电话。

63. 一种用于合成视频序列的设备，包括：

用于存储定义了一种功能的模型数据的存储器，所述功能将一组参数的一个或多个参数与定义对象的形状标准化外貌的纹理数据关联并且将所述一组参数的一个或多个参数与定义对象的形状的形状数据关联；

用于接收表示视频序列的多组参数的装置；

用于对于至少一组被接收的参数生成定义对象的形状标准化外貌的纹理数据以及用于对于多组被接收的参数生成对象的形状数据的装置；

用于利用生成的形状数据变形被生成的纹理数据以便在视频序列的一帧中生成定义对象的外貌的图像数据的装置；

用于驱动显示器来输出被生成的图像数据以便合成视频序列的显示驱动器。

64. 根据权利要求 63 所述的设备，其中所述生成装置可操作地对于选择的各组所述被接收的参数生成纹理数据，并且其中如果所述生成装置没有对当前一组被接收的参数生成纹理数据，则所述变形装置可操作地利用对于当前一组被接收的参数的形状数据为以前的一组参数变形纹理数据。

65. 根据权利要求 64 所述的设备，包括选择装置，用于从所述被接收的多组参数中选择各组参数，所述生成装置将为其生成纹理数据。

66. 根据权利要求 65 所述的设备，其中所述选择装置可操作地根据预定规则从被接收的多组参数中选择各组参数。

67. 根据权利要求 65 或 66 所述的设备，包括用于将来自当前一组参数的参数值与以前的一组参数的参数值相比较的装置，并且其中

所述选择装置可操作地根据所述比较的结果来选择所述当前一组参数。

68. 根据权利要求 67 所述的设备，其中如果所述当前一组参数的一个或多个所述参数与以前一组的相应参数值的差异超过预定阈值，则所述选择装置可操作地选择所述当前一组参数。

69. 根据权利要求 65 到 68 的任何一个所述的设备，其中所述选择装置可操作地选择所述生成装置将根据设备的可用处理能力为其生成所述纹理数据的各组参数。

70. 根据权利要求 63 到 69 的任何一个所述的设备，其中所述模型数据包括将被接收的一组参数与一组中间形状参数以及一组中间纹理参数相关联的第一个模型数据；其中所述模型数据还包括第二个模型数据，其定义一个功能，所述功能将中间形状参数与所述形状参数关联；其中所述模型数据还包括第三个模型数据，其定义一个功能，所述功能将所述一组中间纹理参数与所述纹理参数关联；并且其中所述生成装置包括用于对于每组被接收的参数利用第一个模型数据生成一组中间形状和纹理参数的装置。

71. 根据权利要求 63 到 70 的任何一个所述的设备，还包括用于接收与视频序列相关的音频信号的装置以及用于与视频序列同步地向用户输出音频信号的装置。

72. 根据权利要求 63 到 71 的任何一个所述的设备，包括用于接收语音的装置以及用于处理被接收的语音以便生成表示所述视频序列的所述多组参数的装置并且其中所述接收装置可操作地从所述语音处理装置接收所述参数。

73. 根据权利要求 72 所述的设备，其中所述语音处理装置包括用于将被接收的语音转换成子字单元序列的语音识别单元以及用于将所述子字单元序列转换成表示所述视频序列的所述多组参数的装置。

74. 根据权利要求 73 所述的设备，其中所述转换装置包括用于将每个子字单元转换成表示所述视频序列的一帧的相应的一组参数的一个查找表。

75. 根据权利要求 73 所述的设备，其中所述转换装置包括多个查找表，每个查找表与对象的不同情绪相关联并且还包括用于根据检测的对象的情绪状态来选择所述查找表之一供所述转换装置使用的装

置。

76. 根据权利要求 75 所述的设备，其中所述语音识别单元可操作地从所述语音信号检测对象的情绪状态。

77. 根据权利要求 63 到 71 的任何一个所述的设备，包括用于接收文本的装置以及用于处理被接收的文本以便生成表示对应于说出所述文本的对象的视频序列的各组参数的装置，其中所述接收装置可操作地从所述文本处理装置接收所述多组参数。

78. 根据权利要求 77 所述的设备，还包括用于将合成对应于文本的语音的文本到语音合成器以及用于与相应的视频序列同步输出合成的语音的装置。

79. 根据权利要求 77 或 78 所述的设备，其中所述文本处理装置包括用于将被接收的文本转换成子字单元序列的第一个转换装置以及用于将子字单元序列转换到所述多组参数的第二个转换装置。

80. 根据权利要求 79 所述的设备，其中所述第二个转换装置包括用于将每个子字单元转换成表示所述视频序列的一帧的相应一组参数的查找表。

81. 根据权利要求 80 所述的设备，其中所述第二个转换装置包括多个查找表并且还包括用于选择所述查找表之一供所述第二个转换装置使用的装置。

82. 一种计算机可读介质，存储计算机可执行过程步骤，用于导致可编程的计算机设备变成被配置为根据权利要求 1 到 54 的任何一个所述的电话，根据权利要求 55 到 62 的任何一个所述的电话网服务器或者根据权利要求 63 到 81 的任何一个所述的设备。

83. 一种计算机可实现的指令，用于导致可编程的处理器变成被配置为根据权利要求1到54的任何一个所述的电话，根据权利要求55到62的任何一个所述的电话网服务器或者根据权利要求63到81的任何一个所述的设备。

通信系统

本发明涉及视频处理方法和设备。本发明特别地但不是专门地涉及使用陆线或移动通信设备的视频技术、视频会议等。

已有的视频电话系统遇到通信网（例如电话网或互联网）和用户的电话之间可用的有限带宽的问题。结果，已有的视频电话系统使用有效的编码技术（如MPEG）来减少被发送的视频图像数据的量。但是，压缩的图像数据仍相当大并且因此对于实时视频电话应用，仍需要用户的终端和网络之间相当大的带宽。

本发明目标在于提供一种替代的视频通信系统。

根据一个方面，本发明提供一种电话，该电话利用存储的外貌模型将一组外貌参数增加到形状和纹理参数中，将纹理参数组织（morph）在一起以便生成纹理，将形状参数组织在一起以便生成形状并且利用该形状将纹理变形在图像上，从而生成动画序列。通过对接收的各组参数重复地执行这些步骤，动画视频序列可被重新生成并且在电话的显示器上被显示给用户。在优选实施方案中，单独的参数被用于模拟面部的不同部分。这是有用的，因为对于面部的大部分的纹理从帧到帧不改变。在小功率设备中，纹理不需要每帧被计算并且可以每秒或第三帧被重新计算或者当纹理参数改变大于预定数量时，才重新计算纹理。

本发明的各种其他特性和方面将通过参考附图所说明的下列示例实施方案的描述来被理解，其中：

图1是电信系统的示意图；

图2是组成图1中所示的系统一部分的移动电话的示意框图；

图3a是说明被图2中所示的移动电话发送的数据分组的形式的示意图；

图3b示意地说明了被图2所示的移动电话发送的数据分组流；

图4是在像素采样之前训练图像被变形到参考形状的示意说明；

图5a是说明由组成图2中所示电话的一部分的编码器单元执行的处理步骤的流程图；

图 5b 说明由组成图 2 中所示电话的一部分的解码器单元执行的处理步骤;

图 6 是说明组成图 2 中所示电话的一部分的播放器单元的主要部件的示意框图;

图 7 是说明可被用于图 1 所示的系统的替代移动电话的形式的示意框图;

图 8 是说明组成图 1 所示的系统的一部分并且与图 7 所示的电话交互的业务提供者服务器的主要组件的框图;

图 9 是说明在利用图 7 中说明的电话的主叫方和被叫方之间的呼叫连接期间使用的协议的控制时序图;

图 10 是说明根据替代实施方案的移动电话的主要组件的示意框图;

图 11 是说明根据另一个实施方案的移动电话的主要组件的示意框图;

图 12 是说明用于替代实施方案中的业务提供者服务器的主要组件的示意框图;

图 13 是说明根据另一个实施方案的移动电话的主要组件的示意框图;

图 14 是说明播放器单元的另一种形式的示意框图;

图 15 是说明另一个替代播放器单元的主要组件的示意框图;以及

图 16 是说明另一个替代播放器单元的主要组件的示意框图。

发明内容

图 1 示意地说明电话网 1,其包括多个通过本地交换机 5 被连接到公共交换电话网(PSTN)7 的用户陆线电话 3-1、3-2 和 3-3。被连接到 PSTN 7 的还有被链接到多个基站 11-1、11-2 和 11-3 的移动交换中心(MSC)9。基站 11 可操作地接收和向多个移动电话 13-1、13-2 和 13-3 发送通信,并且移动交换中心 9 是可操作地控制基站 11 之间以及基站 11 和 PSTN 7 之间的连接。如图 1 所示,移动交换中心 9 还被连接到业务提供者服务器 15,在这个实施方案中,该业务提供者服务器 15 为移动电话订户生成外貌模型。这些外貌模型模拟订户的外貌或者该订户想要使用的人物的外貌。在外貌模型模拟订户的外貌的地方,订户的数字图像必须被提供给业务提供者服务器 15,以便合适的外貌

模型可被生成。在这个实施方案中，这些数字照片可以从地理地分布在全国的多个摄影棚 17 中被生成。

现在将给出其中可利用订户移动电话 13-1 之一进行视频电话呼叫的简要描述。在这个实施方案中，当主叫方利用订户电话 13-1 发起呼叫时，语音呼叫通过基站 11-1 和移动交换中心 9 按通常的方式被建立。在这个实施方案中，订户移动电话 13 包括用于生成用户的视频图像的视频摄像机 23。但是，在这个实施方案中，从摄像机 23 生成的视频图像不被发送到基站。代替的，移动电话 13 使用用户的外貌模型来用参数表示该视频图像以便生成与外貌模型和音频一起被发送到基站 11 的外貌参数序列。然后这个数据以传统方式通过电话网被发送到被叫方的电话，其中视频图像被利用参数和外貌模型而重新合成。类似地，对于被叫方的外貌模型与由被叫方生成的外貌参数序列一起被在电话网上发送到订户电话 13-1，其中类似的过程被执行以便重新合成被叫方的视频图像。

现在将参考图 2 到 5 对于移动电话 13-1 和移动电话 13-2 之间的示例呼叫描述其中这个实施方案中实现这一点的方式。图 2 是图 1 中所示的每个移动电话 13 的示意框图。如所示，电话 13 包括用于接收用户的语音并且用于将其转换成相应的电信号的麦克风 21。移动电话 13 还包括视频摄像机 23，其包括将来自用户的光线集中在 CCD 芯片 27 上的光学器件，CCD 芯片进而以通常的方式生成相应的视频信号。如所示，视频信号被传递到跟踪器单元 33，其进而处理视频序列的每个帧以便跟踪视频序列里用户的面部移动。为执行这个跟踪，跟踪器单元 33 使用模拟用户的面部的形状和纹理的变化性的外貌模型。这个外貌模型被存储在用户外貌模型存储器 35 中并且被业务提供者服务器 15 生成，并且当用户首次预订系统时被下载到移动电话 13-1。在跟踪视频序列中的用户的面部移动时，跟踪器单元 33 为每帧生成表示当前帧中用户的面部外貌的姿态和外貌参数。然后生成的姿态和外貌参数与从麦克风 21 输出的音频信号一起被输入到编码器单元 39 中。

但是，在这个实施方案中，在编码器单元 39 对姿态和外貌参数以及音频编码之前，其对用户的外貌模型进行编码以便通过收发信机单元 41 和天线 43 发送到被叫方的移动电话 13-2。用户的外貌模型的这个被编码的版本可被存储用于在其他视频呼叫中的后续发送。然后编

码器单元 39 将姿态和外貌参数序列编码并且对于其发送到被叫方的移动电话 13-2 的相应的音频信号进行编码。在这个实施方案中，音频信号被利用 CELP 编码技术进行编码并且被编码的 CELP 参数以与被编码的姿态和外貌参数交织的方式被发送。

如图 2 所示，从被叫方移动电话 13-2 接收的数据被从收发信机单元 41 传递到解码被发送数据的解码器单元 51。最初，解码器单元 51 将接收并且解码被叫方的外貌模型，然后其将其存储到被叫方外貌模型存储器 54 中。一旦这被接收并且被解码，解码器单元 51 就将接收和解码被编码的姿态和外貌参数以及被编码的音频信号。然后被解码的姿态和外貌参数被传递到播放器单元 53，其利用被解码的被叫方的外貌模型生成对应于一系列接收的姿态和外貌参数的一系列视频帧。然后生成的视频帧被输出到移动电话的显示器 55，在那里被重新生成的视频序列被显示给用户。由解码器单元 51 输出的被解码音频信号被传递到音频驱动单元 57，其向移动电话的扬声器 59 输出被解码的音频信号。播放器单元 53 以及音频驱动单元 57 的操作被安排以便在显示器 55 上显示的图像与由扬声器 59 输出的合适的音频信号被时间同步。

在这个实施方案中，移动电话 13 在数据分组中发送被编码的姿态和外貌参数以及被编码的音频信号。分组的通用格式在图 3a 中显示。如所示，每个分组包括头部分 121 和数据部分 123。头部分 121 识别分组的大小和类型。这使得数据格式以向前和向后兼容的方式很容易被扩展。例如，如果旧的播放器单元 53 被用于新的数据流上，则会遇到其不识别的分组。在这种情况下，旧播放器可简单地忽略这些分组并且仍有机会处理其它的分组。每个分组的头 121 包括 16 位（位 0 到位 15）用于识别分组的大小。如果位 15 被设置为 0，则其它 15 位定义的大小是按字节的分组的大小。如果另一方面，位 15 被设置为 1，则剩余的位表示按 32k 块的分组的大小。在这个实施方案中，编码器单元 39 可生成 6 个不同类型的分组（图 3b 所示）。这些包括:

1. 版本分组 125 - 流中第一个被发送的分组是版本分组。版本分组中被定义的数字是整数并且目前被设置为数字 3。由于分组系统的可扩展特性，所以不希望改变这个数字。

2. 信息分组 127 - 要被发送的下一个分组是信息分组，其包括

同步字节：识别视频的每秒平均采样（或帧）的字节；识别用于制作视频短片的每个采样的动画的参数数据短片数的数据；识别每秒音频采样数的字节；识别音频的每个采样的数据字节数的字节以及识别音频是否被压缩的位。目前，这个位对于未压缩的音频被设置为 0 并且对于以每秒 4800 位被压缩的音频被设置为 1。

3. 音频分组 129－对于未被压缩的音频，每个分组包含一秒的音频数据。对于每秒 4800 位被压缩的音频，每个分组包含 30 毫秒的数据，其是 18 字节。

4. 视频分组 131－用于制作视频单一采样的动画的外貌参数数据。

5. 超音频分组 133－这是对于正常音频分组 129 的连续的一组数据。在这个实施方案中，播放器单元 53 通过其大小来确定超音频分组中音频分组的数量。

6. 超视频分组 135－这是对于正常视频分组 131 的连续的一组数据。在这个实施方案中，播放器单元 53 通过超视频分组的大小来确定视频分组的数量。

在这个实施方案中，被发送的音频和视频分组按时间顺序被混合在被发送的流中，其中最早的分组被最先发送。按上述方式组织分组结构还使得分组除了通过 PSTN 7 还能够在互联网上被路由。

外貌模型

这个实施方案中使用的外貌模型类似于由 Cootes 等人开发的外貌模型，并且在例如 1995 年 1 月计算机景象和图像理解，第一卷，No. 1，第 38 到 59 页，名为“Active Shape Models-Their Training and Application（活动形状模型－其训练和应用）”的论文中描述了所述 Cootes 等人开发的外貌模型。这些外貌模型利用某些现有知识是关于面部图像的内容可用的这样一个事实。例如，可以假设人的面部的两个正面图像每个将包括眼睛、鼻子和嘴巴。

如上面所提到的，在这个实施方案中，外貌模型在业务提供者服务器 15 中被生成。这些外貌模型被通过分析各自用户的多个训练图像而生成。为了用户的外貌模型可以模拟视频序列里的用户的面部的可变性，训练图像应该包括面部表情和 3D 姿态具有最大变化的用户的图像。在这个实施方案中，这些训练图像由用户进入摄影棚 17 并且被数

字摄像机拍摄来生成。

在这个实施方案中，所有的训练图像是具有 500x500 像素的彩色图像，每个像素有红、绿和蓝像素值。结果的外貌模型 35 是由训练图像中头部定义的头图像类的外貌的参数表示，因此相当少数量的参数（典型地对于一个人 15 到 40 个）可描述来自该类的头图像的详细（像素级）外貌。

如申请人的较早的国际申请 WO 00/17820（其内容被合并在这里供参考）中所说明的，通过初始地确定模拟训练图像里面部形状的可变性的形状模型以及模拟训练图像中纹理的可变性或者像素的颜色的纹理模型，然后通过组合形状模型和纹理模型，从而生成外貌模型。

为了创建形状模型，训练图像上多个里程碑点的位置被识别，然后其他训练图像上相同的里程碑点的位置被识别。里程碑点的这个位置的结果是对于每个训练图像的里程碑点的表，其识别图像中每个里程碑点的（x,y）坐标。然后这个实施方案中使用的模拟技术检查训练集上这些坐标的统计数字以便确定这些位置如何在训练图像里变化。为了能够从不同的图像中比较等价的点，头必须被关于通用的一组轴而对齐。通过对每个头迭代地旋转、缩放和转换坐标集，以便它们都大致地填充相同的参考帧来实现这一点。每个头的结果坐标集形成形状矢量（x^i），其元素对应参考帧内的里程碑点的坐标。在这个实施方案中，然后通过在形状训练矢量（x^i）集上执行主要组件分析（PCA）来生成形状模型。这个主要组件分析通过下式生成将每个形状矢量（x^i）与形状参数的对应的矢量（P_s^i）相关的形状模型（Q_s）：

$$p_s^i = Q_s(x^i - \bar{x}) \qquad (1)$$

其中 x^i 是形状矢量，$\bar{x}$ 是来自形状训练矢量的均值形状矢量并且 P_s^i 是对于形状矢量 x^i 的形状参数的矢量。矩阵 Q_s 描述了训练头部里形状和姿态的主要变化模式；并且对于给定输入头部的形状参数的矢量（P_s^i）具有与其值将给定输入头部的形状与相应的变化模式相关的每个变化模式相关的参数。例如，如果训练图像包括用户向左和向右看以及向正前方看的图像，则将由形状模型（Q_s）描述的一个变化模式将具有形状参数的矢量（P_s）中尤其影响用户往哪里看的一个相关参数。特别的，这个参数可从-1 到+1 变化，参数值接近-1 与用户向左看相

关，参数值接近 0 与用户向正前方看有关并且参数值接近+1 与用户向右看相关。因此，被需要来说明训练数据里变化的变化模式越多，形状参数矢量 P_s^i 里需要的形状参数就越多。在这个实施方案中，对于使用的特殊的训练图像，形状和姿态变化的 20 种不同的模式必须被模拟以便说明训练头部中被观察的 98％的变化。

除了能够确定对于给定形状矢量 x^i 的一组形状参数 P_s^i 之外，等式（1）可以关于 x^i 被求解以便给出：

$$x^i = \bar{x} + Q_s^T p_s^i \qquad (2)$$

这是因为 $Q_s Q_s^T$ 等于单位矩阵。因此，通过在合适的限制里修改形状参数（P_s^i）组，新的头部形状可被生成，其将类似于训练组中的那些。

一旦形状模型被生成，类似的模型就被生成以便模拟训练面部里的纹理，并且特别是模拟训练面部中的红、绿和蓝等级。为做到这一点，在这个实施方案中，每个训练面部被变形成一个参考形状。在申请人的早期国际申请中，参考形状是均值形状。但是，这导致经过训练面部中所有面像素采样的恒定清晰度。因此，十倍于嘴唇上的小平面的区域的对应于面颊的小平面，将有十倍多的像素被采样。结果，这个面颊小平面将十倍地有助于纹理模型，这并不需要。因此，在这个实施方案中，参考形状通过使围绕眼睛和嘴巴的小平面大于均值形状中的而被变形，因此眼睛和嘴巴区域比面部的其他部分更密集地被采样。在这个实施方案中，通过变形每个训练图像头直到每个图像的里程碑点的位置与描述参考头部（预先被确定）的形状和姿态的相应里程碑点的位置重合来实现这一点。这些形状被变形的图像中的颜色值被用于作为对纹理模型的输入矢量。这个实施方案中使用的参考模型以及参考形状上里程碑点的位置在图 4 中被示意地显示。如从图 4 中所看到的，参考形状中眼睛和嘴巴的大小与面部中特征的其余部分相比被放大。结果，当形状变形的训练图像被采样时，与面部的其他特征相比更多像素采样围绕眼睛和嘴巴被进行。这导致对围绕嘴巴和眼睛以及其中的变化有更大响应的纹理模型，并且因此对于在源视频序列中跟踪用户更好。各种三角测量技术可被用于将每个训练头部变形为参考形状。在上述的申请人的早期国际申请中描述了一种这样的

技术。

一旦训练头部已经被变形到参考模型，通过在形状被变形的头部上以例如，一万个均匀分布的点采样各个颜色等级，对于每个形状被变形的训练面部的红、绿和蓝等级矢量（r^i，g^i和b^i）就被确定。红色等级矢量的主要成分分析生成红色等级模型（矩阵 Q_r），其利用下式将每个红色等级矢量与红色等级参数的相应矢量相关：

$$p_r^i = Q_r(r^i - \bar{r}) \qquad (3)$$

其中 r^i 是红色等级矢量，$\bar{r}$ 是来自红色等级训练矢量的均值红色等级矢量并且 $p^i{}_r$ 是对于红色等级矢量 r^i 的红色等级参数的矢量。绿色和蓝色等级矢量的类似的主要成分分析生成类似的模型：

$$p_g^i = Q_g(g^i - \bar{g}) \qquad (4)$$

$$p_b^i = Q_b(b^i - \bar{b}) \qquad (5)$$

这些颜色模型描述形状标准化训练面部里颜色的主要变化模式。

以与等式（1）关于 x^i 被求解相同的方式，等式（3）到（5）可关于 r^i、g^i和 b^i 被求解，以便给出：

$$\begin{aligned} r^i &= \bar{r} + Q_r^T p_r^i \\ g^i &= \bar{g} + Q_g^T p_g^i \\ b^i &= \bar{b} + Q_b^T p_b^i \end{aligned} \qquad (6)$$

这是因为 $Q_rQ_r^T$、$Q_gQ_g^T$和 $Q_bQ_b^T$ 是单位矩阵，因此，通过在合适的限制里修改颜色参数（P_i，P_g或 P_b）组，新的形状被变形的颜色面部可被生成，其将类似于训练组中的那些。

如上所述，形状模型和颜色模型被用于生成外貌模型（F_a），其集中地模拟其中形状和颜色在训练图像的面部里变化的方式。因为形状和颜色变化之间有相关性，因此组合的外貌模型被生成，其可被用于减少描述训练面部里所有变化所需的参数的数量。在这个实施方案中，通过在对于训练图像的形状和红、绿以及蓝参数上执行进一步的主要成分分析来实现这一点。特别地，形状参数被与对于每个训练图

像的红、绿和蓝参数连接在一起，然后在被连接的矢量上执行主要成分分析，以便确定外貌模型（矩阵 F_a）。但是，在这个实施方案中，在将形状参数和纹理参数连接在一起之前，形状参数被加权，以便纹理参数不支配主要成分分析。通过将加权矩阵（H_s）引入等式（2）可实现这一点，因此：

$$x^i = \bar{x} + [Q_s^T H_s^{-1}] [H_s p_s^i] \qquad (7)$$

其中 H_s 是合适大小的单位矩阵的倍数（λ），也就是：

$$H_s = \begin{pmatrix} \lambda & 0 & 0 & . & . & . & 0 \\ 0 & \lambda & 0 & . & . & . & 0 \\ 0 & 0 & \lambda & . & . & . & 0 \\ . & . & . & . & & & \\ . & . & . & & . & & \\ . & . & . & & & . & \\ 0 & 0 & 0 & & & & \lambda \end{pmatrix} \qquad (8)$$

其中λ是常数。发明人发现在 1000 和 10000 之间的λ值提供好的结果。因此，Q_s^T 和 P_s^i 变成：

$$\hat{Q}_s^T = Q_s^T H_s^{-1}$$
$$\hat{p}_s^i = H_s p_s^i \qquad (9)$$

一旦形状参数被加权，就在对于每个训练图像的修改的形状参数和红、绿、蓝参数的被连接的矢量上执行主要成分分析，以便确定外貌模型，因此：

$$p_a^i = F_a \begin{bmatrix} \hat{p}_s^i \\ p_r^i \\ p_g^i \\ p_b^i \end{bmatrix} = F_a p_{sc}^i \qquad (10)$$

其中 P^i_a 是控制形状和颜色的外貌参数的矢量，并且 P^i_{sc} 是被连接的修改的形状和颜色参数的矢量。

一旦修改的形状模型（Q_s）、颜色模型（Q_r、Q_g和 Q_b）以及外貌模型（F_a）已经被确定，它们就被发送到用户的移动电话 13，其中它们

被存储供后续使用。

除了能够由一组外貌参数（P^i_a）表示输入面部之外，还可能使用那些外貌参数来重新生成输入面部。特别地，通过将等式（10）与上述等式（1）和（3）到（5）合并，对于形状矢量和对于RGB等级矢量的表达式可如下被确定：

$$x^i = \bar{x} + V_s p_a^i \qquad (11)$$

$$r^i = \bar{r} + V_r p_a^i \qquad (12)$$

$$g^i = \bar{g} + V_g p_a^i \qquad (13)$$

$$b^i = \bar{b} + V_b p_a^i \qquad (14)$$

其中从F_a和Q_s获得V_s，从F_a和Q_r获得V_r，从F_a和Q_g获得V_g，并且从F_a和Q_b获得V_b。为了重新生成面部，从颜色参数生成的形状被变形的颜色图像必须从参考形状被变形以便考虑由形状矢量x^i描述的面部的形状。其中变形形状自由的灰色等级图像被执行的方式在上述申请人的早期国际申请中被描述。如本领域的技术人员所理解的，类似的处理技术被用于变形每个形状被变形的颜色成分，然后其被组合以便重新生成面部图像。

编码器单元

现在将参考图5a来描述优选方式，其中图2中所示的编码器单元39对用户的外貌模型编码以便发送到被叫方的移动电话13-2。然后参考图5b来描述其中解码器单元51重新生成被叫方的外貌模型（其以相同的方式被编码）的方式。

最初，在步骤s71，编码器单元39将用户的外貌模型分解成形状（Q_s^{trgt}）和颜色模型（Q_r^{trgt}、Q_g^{trgt}和Q_b^{trgt}）。然后，在步骤s73，编码器单元39对于变化的每个红、绿和蓝模式生成形状被变形的颜色图像。特别地，对于颜色参数的下面的每个矢量利用上面的等式（6）生成形状被变形的红、绿和蓝图像：

$$p_r^i ; p_g^i ; p_b^i = \begin{pmatrix}1\\0\\0\\\cdot\\\cdot\\\cdot\\0\end{pmatrix}; \begin{pmatrix}0\\1\\0\\\cdot\\\cdot\\\cdot\\0\end{pmatrix}; \begin{pmatrix}0\\0\\1\\\cdot\\\cdot\\\cdot\\0\end{pmatrix}; \cdots \begin{pmatrix}0\\0\\0\\\cdot\\\cdot\\\cdot\\1\end{pmatrix} \qquad (15)$$

（尽管等式（6）中使用的均值矢量如果需要可被忽略）。然后在步骤s75，利用诸如JPEG的标准图像压缩算法来压缩这些形状被变形的图像和均值颜色图像（$\bar{r}$，$\bar{g}$和$\bar{b}$）。但是，如本领域的技术人员所理解的，在利用JPEG算法压缩之前，形状被变形的图像和均值颜色图像必须被组合成矩形参考帧，否则JPEG算法将不起作用。因为所有的形状标准化的图像都具有相同的形状，所以其被组合在矩形参考帧中相同的位置。这个位置由模板图像确定，其在这个实施方案中被直接从参考形状（图4中示意说明的）生成，并且其包含1's和0's，模板图像中1's对应背景像素并且模板图像中0's对应图像像素。这个模板图像也必须被发送到被叫方的移动电话13-2中并且在这个实施方案中被利用运转周期编码技术进行压缩。然后编码器单元39在步骤s77输出形状模型（Q_s^{trgt}）、外貌模型（$(F_a^{trgt})^T$）、均值形状矢量$(\bar{x}^{trgt})$并且被压缩的图像用于通过收发信机单元41发送到电话网。

解码器单元

参考图5b，解码器单元51在步骤s81解压缩JPEG图像、均值颜色图像和被压缩的模板图像。然后处理进行到步骤s83，其中被解压缩的JPEG图像被采样以便利用被解压缩的模板图像来恢复形状被变形的颜色矢量（r^i，g^i和b^i）以便识别要被采样的像素。因为被用于生成这些形状被变形颜色图像的颜色参数矢量的选择（参见上面的（15）），所以可利用将相应的形状被变形的颜色矢量叠加在一起来重新构建颜色模型（Q_r^{trgt}，Q_g^{trgt}和Q_b^{trgt}）。如图5b所示，形状自由的颜色矢量的这种叠加在步骤s85被执行。然后处理进行到步骤s87，其中被恢复的形状和颜色模型被组合以便重新生成被存储在存储器54中的被叫方的外貌模型。

在这个实施方案中，利用这种优选的编码技术，颜色模型被发送到其他方比只是独立地被发送的效率高大约10倍。这是因为，在这个实施方案中使用的每个颜色模型典型地是30000x8的矩阵并且每个矩

阵的每个元素需要 3 个字节。因此，每个移动电话 13 需要发送大约 720 千字节的数据以便以未被压缩的形式发送颜色模型矩阵。而是通过生成上述的形状被变形的颜色图像并且利用标准的图像编码技术来对其编码并且发送被编码的图像，发送颜色模型所需的数据的量仅大约是 70 千字节。

播放器单元

图 6 是更详细地说明在这个实施方案中使用的播放器单元 53 的组件。如所示，播放器单元包括参数转换器 150，其在输入线 152 上接收被解码的外貌参数并且在输入线 154 上接收被叫方的外貌模型。在这个实施方案中，参数转换器 150 使用等式（11）到（14）利用线 154 上被叫方的外貌模型输入来将输入外貌参数 P_a^i 转换成相应的形状矢量 x^i 以及形状被变形的 RGB 等级矢量（r^i，g^i，b^i）。RGB 等级矢量在线 156 上被输出到形状变形器 158 并且形状矢量在线 164 上被输出到形状变形器 158。形状变形器 158 操作变形来自参考形状的 RGB 等级矢量以便考虑由形状矢量 x^i 描述的面部的形状。由形状变形器 158 生成的结果 RGB 等级矢量在输出线 160 上被输出到图像组合器 162，其使用 RGB 等级矢量来生成像素值的相应的两维阵列，并且将其输出到帧缓冲器 166 用于在显示器 55 上显示。

修改和替代实施方案

在上述第一个实施方案中，每个订户电话 13-1 包括用于生成用户的视频序列的摄像机 23。然后利用被存储的外貌模型将这个视频序列转换成一组外貌参数。现在描述第二个实施方案，其中订户电话 13 不包括视频摄像机。代替的，电话 13 直接从用户的输入语音中生成外貌参数。图 7 是订户电话 13 的框示意图。如所示，从麦克风 21 输出的语音信号输入到自动语音识别单元 180 和单独的语音编码器单元 182。语音编码器单元 182 对语音进行编码，以便按通常的方式通过收发信机单元 41 和天线 43 发送到基站 121。语音识别单元 180 将输入语音和预存音素模型（被存储在音素模型存储器 181 中）相比较以便生成一系列音素 33，其将这一系列音素输出到查找表 35。查找表 35 为每个音素存储一组外貌参数并且被排列，以便对于由自动语音识别单元 180 输出的每个音素，表示相应的音素发音期间用户的外貌的相应的一组外貌参数被输出。在这个实施方案中，查找表 35 专用于移动

电话 13 的用户并且在一个训练例程期间被预先生成，在所述训练例程中，音素和从外貌模型生成用户的所需图像的外貌参数之间的关系被学习。下列表 1 说明在这个实施方案中查找表 35 具有的格式。

表 1

参数 音素	P_1	P_2	P_3	P_4	P_5	P_6	...
/ah/	0.34	0.1	-0.7	0.23	-0.15	0.0	...
/ax/	0.28	0.15	-0.54	0.1	0.0	-0.12	...
/r/	0.48	0.33	0.11	-0.7	-0.21	0.32	...
/p/	-0.17	-0.28	0.32	0.0	-0.2	-0.09	...
/t/	0.41	-0.15	0.19	-0.47	-0.3	-0.04	...
/s/	-0.31	0.28	-0.02	0.0	-0.22	0.14	...
/m/	0.02	-0.08	0.13	0.2	0.03	0.18	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

如图 7 所示，由查找表 35 输出的外貌参数集 37 然后被输入到编码器单元 39，其对外貌参数进行编码以便发送到被叫方。然后被编码的参数 40 被输入到收发信机单元 41，其将被编码的外貌参数与相应的被编码的语音一起发送。如在第一个实施方案中，收发信机 41 以时间交织的方式发送被编码的语音以及被编码的外貌参数，以便对于被叫方的电话更容易维持合成的视频和相应音频之间的同步。

如图 7 所示，移动电话的接收机端与第一个实施方案中相同并且因此将不再描述。

本领域的技术人员从上面的描述应该理解，在这第二个实施方案中，用户的移动电话 134 不需要有用户的外貌模型以便生成其发送的外貌参数。但是，被叫方将需要有用户的外貌模型以便合成相应的视频序列。因此，在这个实施方案中，所有订户的外貌模型被存储在业务提供者服务器 15 的中心，并且一旦启动了订户之间的呼叫，业务提供者服务器 15 就可操作地将合适的外貌模型下载到合适的电话。

图 8 更详细地显示了业务提供者服务器 15 的内容。如所示，其包括接口单元 191，其提供移动交换中心 9 和摄影棚 17 以及服务器 15

里的控制单元 193 之间的接口。当服务器为新订户接收图像时，控制单元 193 将图像传递给外貌模型生成器 195，其以第一个实施方案中所述的方式建立合适的外貌模型。然后合适的外貌模型被存储在合适的模型数据库 197 中。随后，当订户之间的呼叫被启动时，移动交换中心 9 通知服务器 15 主叫方和被叫方的成分。然后控制单元 193 从外貌模型数据库 197 中检索主叫方和被叫方的外貌模型并且通过接口单元 191 将这些外貌模型发送回移动交换中心 9。然后移动交换中心 9 将主叫方的合适的外貌模型发送到被叫方电话并且将该外貌模型发送到各自订户的电话。

现在将参考图 9 描述这个实施方案的控制定时。最初，主叫方利用键盘输入被叫方的号码。一旦主叫方输入所有的号码并且按下电话 13 上的发送建（未示出），则该号码在空中接口上被发送到基站 11-1。然后基站将这个号码转发到移动交换中心 9，其向业务提供者服务器 15 发送主叫方的 ID 以及被叫方的 ID，以便合适的外貌模型能够被检索。然后移动交换中心 9 通过电话网中合适的连接向被叫方发信号以便引起被叫方的电话 13-2 振铃。当这个发生时，服务提供者服务器 15 向移动交换中心 9 下载主叫方和被叫方的合适的外貌模型，其中他们被存储用于后续下载到用户电话。一旦被叫方电话振铃，移动交换中心 9 就将状态信息发送回主叫方的电话，以便其可生成合适的振铃音。一旦被叫方摘机，合适的信令信息就被发送到电话网返回移动交换中心 9。在响应中，移动交换中心 9 将主叫方的外貌模型下载到被叫方并且将被叫方的外貌模型下载到主叫方。一旦这些模型被下载，各个电话就以与上述第一个实施方案中相同的方式对被发送的外貌参数进行解码，以便合成说话的相应用户的视频图像。这个视频呼叫保持在合适的位置直到主叫方或被叫方结束该呼叫。

上述第二个实施方案比第一个实施方案有多个优点。首先，订户电话不需要有内置的或被连接的视频摄像机。外貌参数被从用户的语音中直接生成。其次，主叫方和被叫方的外貌模型仅在一个限制通信链路上被发送。特别地，在第一个实施方案中，每个外貌模型被从用户的电话发送到电话网，然后从电话网发送到其他人的电话。当电话网中可用的外宽相当高时，从网络到电话的信道中的带宽受到更多限制。因此，在这个实施方案中，因为外貌模型被存储在电话网中心，

所以其仅需在一个受限带宽链路上被发送。如本领域的技术人员所理解的，第一个实施方案可被修改以便以类似的方式在外貌模型存储在电话网中的情况下进行操作。在上述实施方案中，用户的外貌参数被生成并且被从用户的电话发送到被叫方的电话，其中视频序列被合成显示用户在讲话。现在将参考图 10 描述一个实施方案，其中电话基本上有与第二个实施方案相同的结构，但是具有额外的身份（identity）偏移单元 185，其可操作地转换外貌参数值以便修改用户的外貌。身份偏移单元 185 利用在存储器 187 中存储的预定转换来执行转换。该转换可被用于修改用户的外貌或者简单地改善用户的外貌。可能增加将改变用户的被觉察的情绪状态的外貌参数（或者形状或纹理参数）的偏移。例如，向从“中立的”动画的语音生成的所有外貌参数增加轻微笑容的外貌参数的矢量将使用户看上去高兴。增加皱眉的矢量将使其看上去生气。有各种方式其中身份偏移单元 185 可执行身份偏移。一种方式在申请人的早期国际申请 WO00/17820 中被描述。一种替代技术在申请人的共同未决的英国申请 GB0031511.9 中被描述。这个实施方案中电话的剩余部分与第二个实施方案中相同并且因此将不再描述。在上述的第二和第三个实施方案中，电话包括自动语音识别单元。现在参考图 11 和 12 描述一个实施方案，其中自动语音识别单元在业务提供者服务器 15 中而不是在用户的电话中被提供。如图 11 中所示，订户电话 13 比图 7 中所示的第二个实施方案的订户电话简单得多。如所示，由麦克风 21 生成的语音信号被直接输入到语音编码器单元 182，其以传统方式编码语音。然后被编码的语音通过收发信机单元 41 和天线 43 被发送到业务提供者服务器 15。在这个实施方案中，来自主叫方和被叫方的所有语音信号通过业务提供者服务器 15 被发送，其框图在图 12 中被显示。如所示，在这个实施方案中，服务器 15 包括自动语音识别单元 180 和所有的用户查找表 35。

在操作中，当呼叫在主叫方和被叫方之间被建立时，所有被编码的语音通过服务器 15 被发送到另一方。服务器将语音传递到自动语音识别单元 180，其识别该语音和说话人并且将生成的音素输出到合适的查找表 35。然后相应的外貌参数被从查找表提取并且被传递回控制单元 193 用于与被编码的音频一起向前发送到另一方，其中视频序列如以前一样被合成。

如本领域的技术人员将理解的，这个实施方案提供订户电话不需要有复杂的语音识别单元的优点，因为每件事都在业务提供者服务器15中心被完成。但是，缺点是自动语音识别单元180必须能够识别所有订户的语音并且其必须能够识别哪个订户在说什么，以便音素可被应用于合适的查找表。

在上述的第二到第四个实施方案中，对于每个订户提供单一查找表35，其将订户生成的音素映射到相应的外貌参数值。但是，语音识别单元输出的音素和实际外貌参数值之间的关系根据用户的情绪状态来改变。图13是说明替代订户电话的组件的框图，其中查找表数据库205对于用户的不同的情绪状态存储不同的查找表35。查找表数据库205包括对于当用户高兴、生气、兴奋、悲伤等时的合适的查找表。在这个实施方案中，用户的当前情绪状态由自动语音识别单元180通过检测用户语音中的重音等级来确定。在响应中，自动语音识别单元180向查找表数据库205输出合适的指令以便导致合适的查找表35被用于将来自语音识别单元180的音素序列输出转换成相应的外貌参数。如本领域的技术人员所理解的，查找表数据库205中的每个查找表必须从那些情绪状态的每个中用户的训练图像中被生成。再次，这个被预先完成并且合适的查找表在业务提供者服务器16中被生成并且然后被下载到订户电话中。替代的，“中立”查找表被与身份偏移单元一起使用，然后其根据用户的被检测的情绪状态来执行一个合适的身份偏移。

在上述第一个实施方案中，CELP音频编解码器被用于对用户的音频编码。这样的编码器对于每秒大约4.8千比特（kbps）的音频减少了所需的带宽。如果移动电话要在具有7.2kbps带宽的标准GMS链路上发送语音和视频数据，则这对于外貌参数提供了2.4kbps的带宽。但是，大多数已有的GSM电话，不使用CELP音频编码器。代替的，它们使用利用完整7.2kbps带宽的音频编解码器。因此如果CELP音频编解码器在软件中被提供，则上述系统仅能够在已有的GSM电话中工作。但是，因为大多数已有的移动电话没有计算能力来解码音频数据，因此这是不实用的。

但是，上述系统可以被用于已有的GSM电话中以便发送预记录的视频序列。这是可能的，因为在正常交谈中存在沉默，在此期间可用

的带宽没有被使用。特别地，对于典型的讲话者，归因于字或短语之间的小中止，15%到30%之间的时间带宽完全没有被使用。因此，视频数据可以与音频一起被发送以便完全地利用可用的带宽。如果接收机在重新同步视频序列之前要接收所有的视频和音频数据，则音频和视频数据可在GSM链路上以任何顺序和任何序列被发送。替代地，为了允许尽可能快地播放视频序列的更有效的实现，合适大小的视频数据块（如上述的外貌参数）可在相应的音频数据之前被发送，因此音频一旦被接收视频就可以开始播放。在这种情况下在相应的音频之前发送视频数据是最佳的，因为外貌参数数据比音频数据使用每秒更少量的数据。因此，如果播放四秒的视频部分对于音频需要四秒的发送时间并且对于视频需要一秒的发送时间，则整体发送时间是五秒并且视频在一秒之后开始播放。如果音频中的沉默足够长，则这样的系统在接收机处仅需要相当少数量的缓存来缓存在音频之前被发送的被接收的视频数据就可以运行。但是，如果音频中的沉默不够长以便做到这一点，则更多的视频必须更早地被发送，从而导致接收机必须缓存更多的视频数据。如本领域的技术人员所理解的，这样的实施方案将需要音频和视频数据的时间戳，以便其可以被接收机处的播放器单元重新同步。

这些被预记录的视频序列可被生成和存储在服务器上，从中用户可以将序列下载到其电话上用于浏览和后续发送到另一个用户。如果视频序列被用户利用其电话生成，则该电话也需要包括必要的处理电路来识别音频中的停顿以便识别和与音频一起被发送的视频数据的数量，以及合适的处理电路用于生成视频数据以及用于将其与音频数据混合以便GSM编解码器完全利用可用的带宽。

做为从语音直接驱动视频序列的替代，动画序列可被直接从文本生成。例如，用户可向中央服务器发送文本，然后其将文本转换成合适的外貌参数和被编码的音频，其将这些与合适的外貌模型一起发送到被叫方的电话。然后视频序列可以上述的方式被生成。在这样的实施方案中，当用户预订服务并且使用摄影棚之一来提供图像用于生成外貌模型时，用户还通过摄影棚中的麦克风输入某些短语，以便服务器可为该用户生成合适的语音合成器，其将在后来使用从用户的输入文本中合成语音。做为在服务器中合成语音以及生成外貌参数的替

代，这可以直接在用户的电话或在被叫方的电话中被完成。但是，在目前这样的实施方案不实际，因为文本到视频生成是计算密集的并且要求被叫方有具有一个有能力的电话。

在上述实施方案中，模拟用户的面部的整体形状和颜色的外貌模型被描述。在替代实施方案中，单独的外貌模型或者仅仅单独的颜色模型可被用于眼睛、嘴巴和面部区域的剩余部分。因为单独的模型被使用，所以不同数量的外貌参数或者不同类型的模型可被用于不同的元素。例如，对于眼睛和嘴巴的模型可比用于面部剩余部分的模型包括更多的参数。替代的，面部的剩余部分可由没有任何变化模式的平均纹理来简单地模拟。这是有用的，因为对于大部分面部的纹理在视频呼叫期间不显著变化。这意味着在订户电话之间需要发送较少的数据。

图 14 是在实施方案中被使用的播放器单元 53 的示意框图，其中为眼睛和嘴巴以及面部的剩余部分提供单独的颜色模型（但是通用的形状模型）。如所示，除了参数转换器 150 可操作地接收被发送的外貌参数并且生成形状矢量 x_i（其在线 164 上将该矢量输出到形状变形器 158）以及对于各个颜色模型分离出颜色参数之外，播放器单元 53 基本上与第一个实施方案的播放器单元 53 相同。对于眼睛的颜色参数被输出到参数到像素转换器 211，其利用在输入线 212 上提供的眼睛颜色模型将所述参数值转换成相应的红、绿和蓝等级矢量。类似的，嘴巴颜色参数由参数转换器 150 输出到参数到像素转换器 213，其利用在线 214 上输入的嘴巴颜色模型将嘴巴参数转换成对于嘴巴的相应的红、绿和蓝等级矢量。最后，对于面部剩余区域的一个或多个外貌参数被输入到参数到像素转换器 215，其中利用线 216 上输入的模来产生合适的红、绿和蓝等级矢量。如图 14 所示，从每个参数到像素转换器输出的 RGB 等级矢量被输入到面部再现器（renderer）单元 220，其从中重新生成第一个实施方案的形状标准化的颜色等级矢量。然后这些被传递到形状变形器 158，其中他们被变形以便考虑当前的形状矢量 x^i。后续处理与第一个实施方案是相同的并且因此不再描述。

在从外貌参数生成视频图像中最计算密集的操作之一是颜色参数转换成 RGB 等级矢量。现在将描述一个实施方案，其中颜色等级矢量不被每帧重新计算，而是代替地每秒或每第三帧被计算。这个替代实

施方案对于图 15 中所示的播放器单元 53 被描述，虽然其可被用于第一个实施方案的播放器单元。如所示，在这个实施方案中，播放器单元 53 还包括控制单元 223，其可操作地在控制线 225 上输出通用的使能信号，该信号被输出到每个信号到像素转换器 211、213 和 215。在这个实施方案中，当控制单元 223 允许这些转换器这样做时，这些转换器才可操作地将接收的颜色参数转换成相应的 RGB 等级矢量。

在运行中，参数转换器 150 对于要被输出到显示器 55 的视频序列的每帧输出各组颜色参数和形状矢量。形状矢量如以前一样被输出到形状变形器 158 并且各个颜色参数被输出到相应的参数到像素转换器。但是，在这个实施方案中，控制单元 223 仅使能转换器 211、213 和 215 来为每第三个视频帧生成合适的 RGB 等级矢量。对于其参数到像素转换器 211、213 和 215 没有被使能的视频帧，面部再现器 220 可操作地输出对于以前帧生成的 RGB 等级矢量，然后其可由形状变形器 158 利用对于当前视频帧的新的形状矢量予以变形。

作为另一个替代，而不是每秒或每第三个视频帧重新计算一次颜色等级矢量，每当相应的输入参数改变了预定量，颜色等级矢量就能够被计算。在因为只有对应特定部分的颜色需要被更新，所以使用对于眼睛和嘴巴以及面部的剩余部分的单独的模型的实施方案中，这一点特别有用。通过为控制单元 223 提供由参数转换器 150 输出的参数，因此其可以监视从一帧到下一帧的参数值之间的变化，从而实现这样的实施方案。每当这个改变超过预定的阈值，合适的参数到像素转换器将由控制单元到所述转换器的专用使能信号使能。然后面部再现器 220 可操作地将该部分的新的 RGB 等级矢量与其它部分的旧的 RGB 等级矢量组合来为面部生成形状标准化的 RGB 等级矢量，然后其被输入到形状变形器 158。

如上面提到的，本系统最计算密集的操作之一是颜色外貌参数到颜色等级矢量的转换。有时，对于诸如移动电话的低功率设备，每个时间点可用的处理能力的数量将变化。这种情况下，被用来重新构建颜色等级矢量的变化的颜色模式的数量（也就是颜色参数的数量）可根据当前可用的处理能力动态地变化。例如，如果移动电话对于每帧接收 30 个颜色参数，则当所有的处理能力可用时，其使用所有这 30 个参数来重新构建颜色等级矢量。但是，如果可用的处理能力减少，

则只有前 20 个颜色参数（表示变化的最重要颜色模式）可被用于重新构建颜色等级矢量。

图 16 是说明被编程以便按上述方式操作的播放器单元 53 的形式的框图。特别地，参数转换器 150 可操作地接收输入外貌参数以及生成形状矢量 x^i 以及红、绿和蓝颜色参数（P_r^i，P_g^i 和 P_b^i），其将这些参数输出到参数到像素转换器 226。然后参数到像素转换器 226 使用等式（6）将这些颜色参数转换成相应的红、绿和蓝等级矢量。在这个实施方案中，控制单元 223 可操作地根据转换器单元 226 可用的当前处理能力来输出控制信号 228。根据控制信号 228 的等级，参数到像素转换器 226 动态地选择其在等式（6）中使用的颜色参数的数量。如本领域的技术人员所理解的，颜色模型矩阵（Q）的维数不改变但是颜色参数（P_r^i，P_g^i 和 P_b^i）中的某些元素被设置为零。在这个实施方案中，与变化的最不重要模式相关的颜色参数是被设置为零的参数，因为这些将对像素值有最小的影响。

在上述实施方案中，被编码的语音和外貌参数被每个电话接收、解码并且然后被输出给用户。在替代实施方案中，该电话包括用于高速缓存除外貌模型之外的动画和音频序列的存储器。然后这个高速缓存被用于存储预定或“存储”的动画序列。然后一接收来自通信的另一方的合适的指令，这些预定的动画序列就被播放给用户。这样，如果动画序列被重复地播放给用户，则该序列的外貌参数仅需要被发送给该用户一次。

上述实施方案已经描述了多个不同的双向电信系统。如本领域的技术人员所理解的，上述动画技术可以类似的方式被用于向用户留消息。例如，用户可记录一个消息，其被存储在中央服务器上直到由被叫方取回。在这种情况下，该消息包括与被编码的音频一起的相应的外貌参数序列。替代的，对于视频动画的外貌参数可在被叫方取回该消息的时候被服务器或被叫方电话生成。该消息发送可利用用户或某些任意的真实或虚构的人物的预先记录的存储序列。在选择被存储的序列时，用户使用允许其浏览在服务器上可用的存储序列的选择的界面并且在发送该消息之前在他/她的电话上浏览。作为进一步的代替，当用户初始地注册该服务并且使用摄影棚时，摄影棚问用户其是否想要为任何准备的短语记录动画和语音用于后面用做预先记录的消息。

在这种情况下，为用户呈现了短语的选择，其可以从中选择一个或多个。替代的，用户可记录其自己的个人短语。这对于文本到视频发消息的系统尤其合适，因为与仅有文本被用于驱动视频序列时相比，其将提供更高质量的动画。

在上述实施方案中，被使用的外貌模型从一组训练图像的主要成分分析中被生成。如本领域的技术人员所理解的，这些结果应用于可被一组连续的变量参数化的任何模型。例如，矢量量子化和小波技术可被使用。

在上述实施方案中，形状参数和颜色参数被组合以便生成外貌参数。这不是必须的。单独的形状和颜色参数可被使用。而且，如果训练图像是黑白的，则纹理参数可表示图像中除红、绿和蓝等级的灰度等级。而且，代替模拟红、绿和蓝值，颜色可由色度和亮度成分或者由色调、饱和度和值成分表示。

在上述实施方案中，使用的模型是两维模型。如果便携式设备里有足够的处理能力可用，则可使用三维模型。在这样的实施方案中，形状模型可模拟训练模型上里程碑点的三维网格。可利用三维扫描仪或利用摄像机的一个或多个立体象对来获得三维训练例子。

在上述实施方案中，被使用的外貌模型生成各个用户的视频图像。这不是必须的。每个用户例如可选择表示计算机生成的人物的外貌模型，其可以是人物或不是人类的物体。在这种情况下，业务提供者存储多个不同人物的外貌模型，从中每个订户可选择其想要使用的一个人物。仍然是替代的，被叫方可选择用于对主叫方制作动画的身份或人物。选择的成分可以是主叫方的多个不同模型之一或者一些其他真实或虚构的人物的模型。

在上述实施方案中，假设移动电话没有相关外貌模型用来生成另一方的动画序列。但是，在有些实施方案中，每个移动电话可存储多个不同用户的外貌模型以便其不需要在电话网上被发送。在这种情况下，只有动画参数需要在电话网上被发送。在这样的实施方案中，电话网将向移动电话发送请求来询问其是否有对于呼叫的另一方的合适的外貌模型，并且只有其没有时才可操作来发送合适的外貌模型。而且，因为对于当前的移动电话网，在建立连接来发送文件中有大约 5 秒的系统开销，所以如果模型以及参数流被需要，则最好在一个文件

中发送它们。因此在优选实施方案中，服务器存储每个动画文件的两个版本用于发送，一个有模型而一个没有。

在上述第一个实施方案中，主叫方的外貌模型被发送到被叫方并且反之亦然。因此主叫方的电话以及被叫者的电话使用接收的外貌参数来为各个用户生成视频序列。在替代实施方案中，播放器适合于根据谁在说话而在显示被叫方和主叫方的视频之间转换。这样的实施方案对于从语音直接生成视频序列的系统特别合适，因为（i）当被叫方不说话时很难合适地对其制作动画；并且(ii)用户想要看到自己的视频被生成以便验证其可信性。

在上述实施方案中，订户电话被描述为移动电话。如本领域的技术人员所理解的，图 1 中所示的陆线电话也可以适合于以相同的方式运行。在这种情况下，被连接到陆线上的本地交换机需要将陆线电话合适地与业务提供者服务器接口。

在上述实施方案中，摄影棚被提供给用户来为服务器提供图像，以便合适的外貌模型可被生成用于系统。如本领域的技术人员所理解的，其他技术也可被用于输入用户的图像以便生成外貌模型。例如，在上述实施方案中在服务器中提供的外貌模型生成器软件可在用户的家庭计算机上被提供。在这样的情况下，用户可直接从用户从扫描仪或从照片或视频摄像机输入的图像生成其自己的外貌模型。仍然是替代的，用户可简单地将照片或数字图像发送到第三方，然后其可利用它们构建合适的模型用于系统中。

以上描述了多个基于电话系统的实施方案。上述实施方案的许多特性可被用于其他应用中。例如，参考图 14、15 和 16 描述的播放器单元可在任何手持设备或其中有有限的处理能力可用的设备中有利地被使用。类似的，其中视频序列被直接从用户的语音生成的上述实施方案，可被用于本地生成视频序列，而不是将其发送到另一个用户。而且，上述许多修改和替代实施方案可被用于互联网上的通信，其中在例如，用户终端和互联网上的服务器之间有有限的带宽可用。

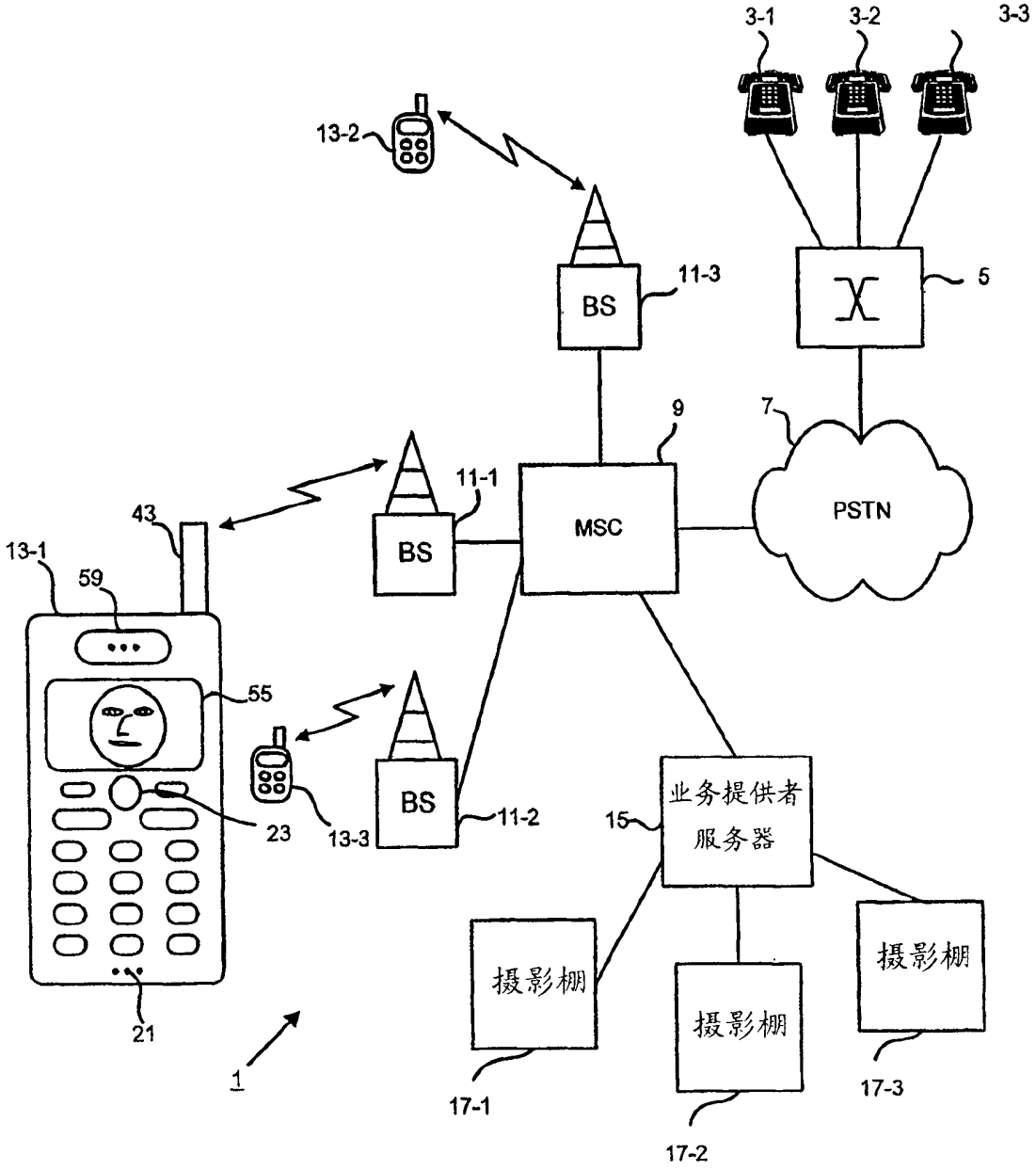

3-1
3-2
3-3
13-2
BS
11-3
5
9
7
11-1
BS
MSC
PSTN
43
13-1
59
55
BS
11-2
23
13-3
15
业务提供者
服务器
摄影棚
摄影棚
摄影棚
21
1
17-1
17-2
17-3

图 1

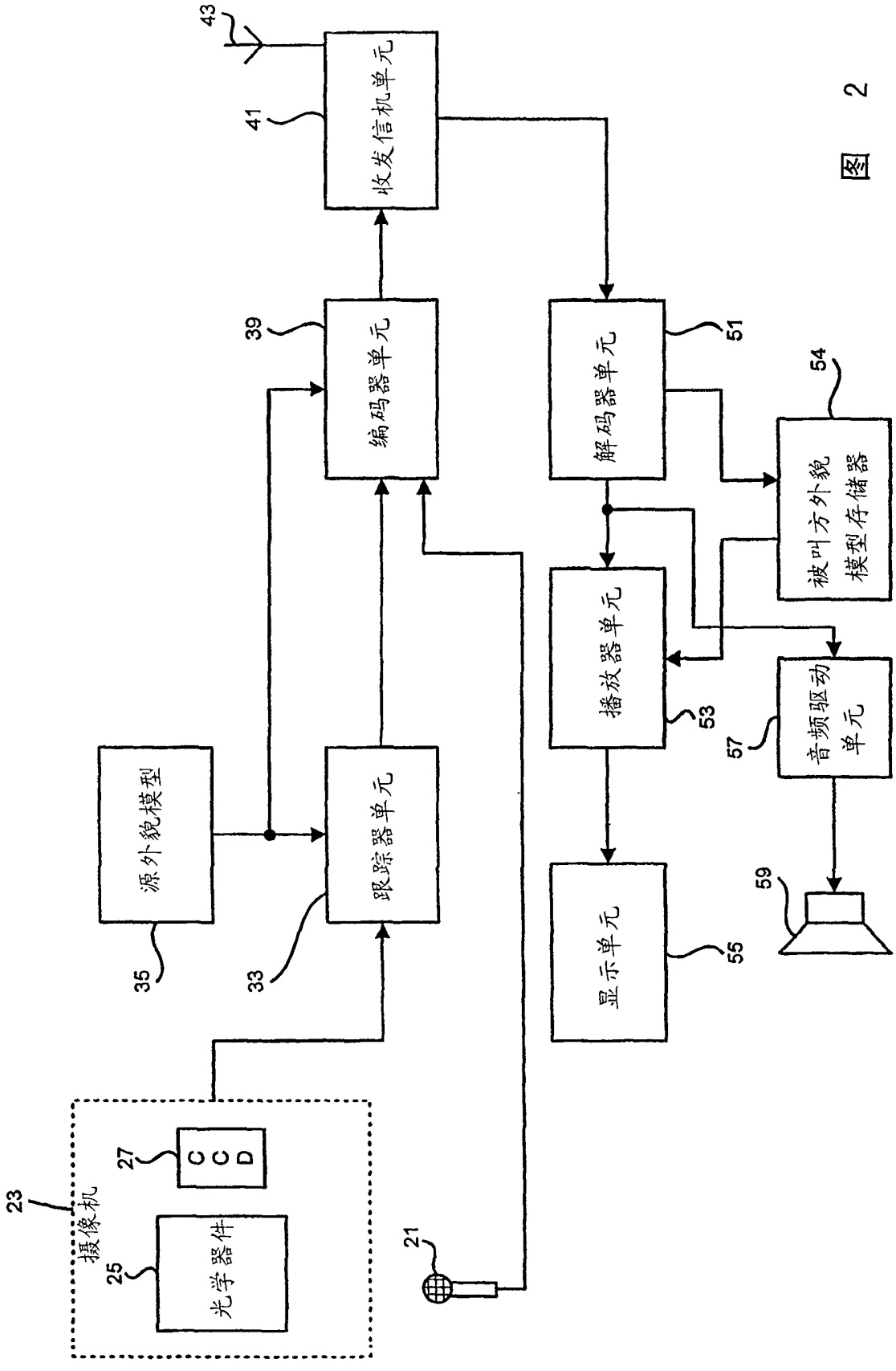

43
41
收发信机单元
39
编码器单元
51
解码器单元
54
被叫方外貌模型存储器
播放器单元
53
57
音频驱动单元
59
显示单元
55
源外貌模型
35
跟踪器单元
33
23
摄像机
27
CCD
25
光学器件
21

图 2

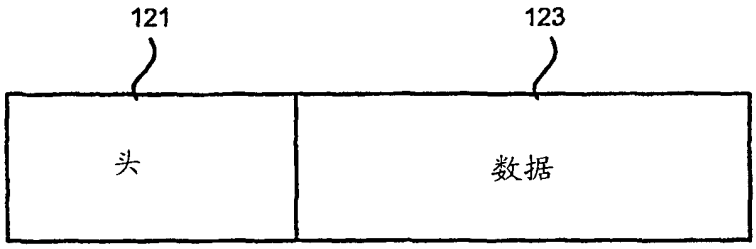


图 3a

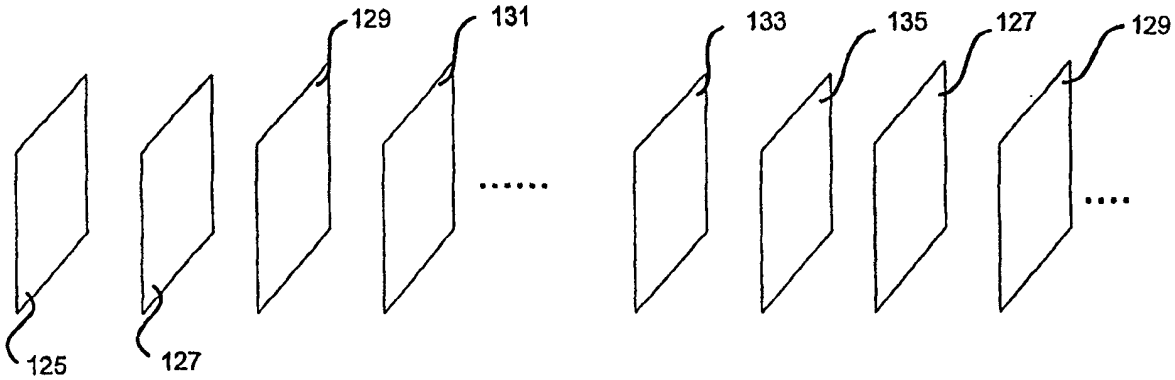


图 3b

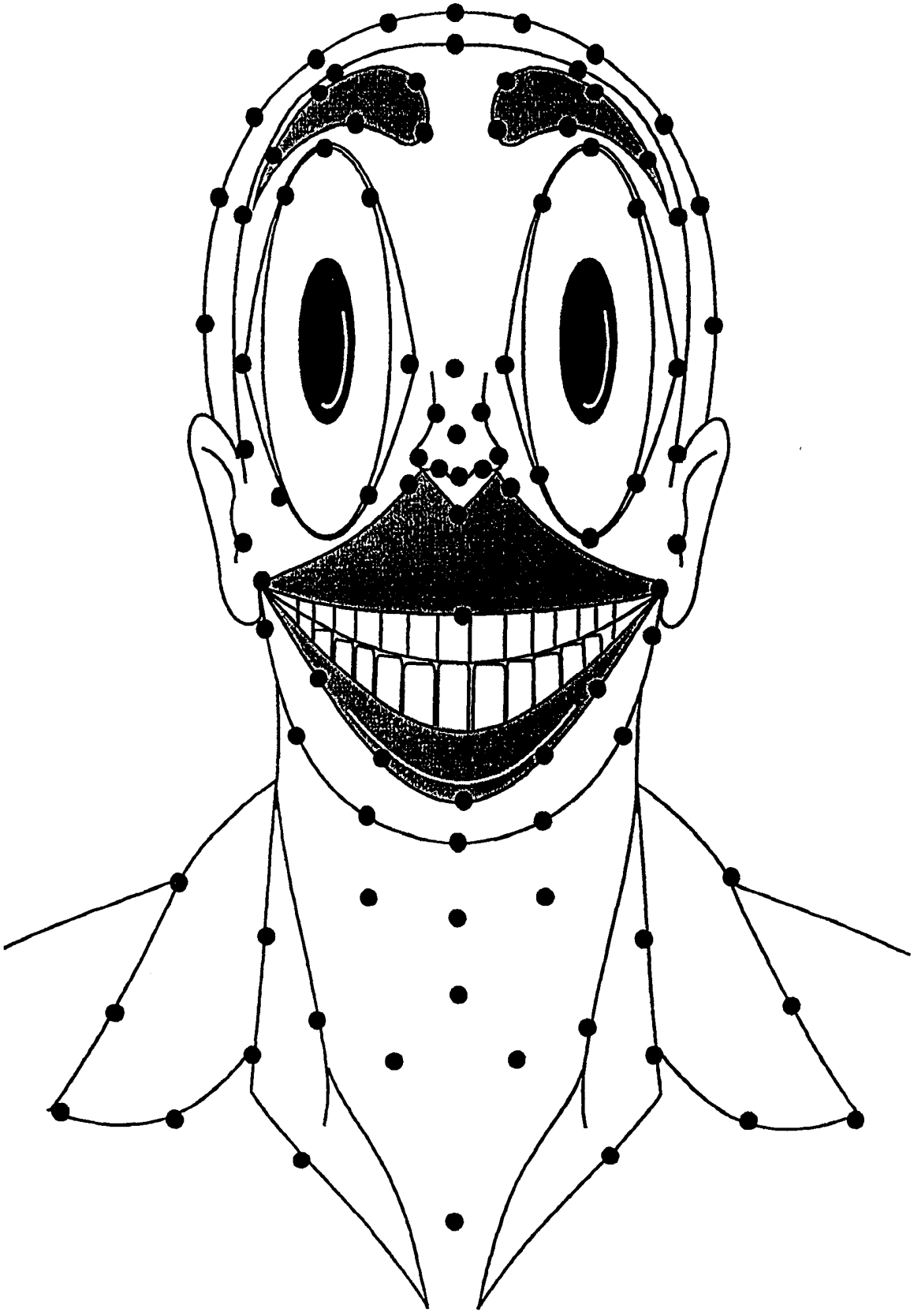


图 4

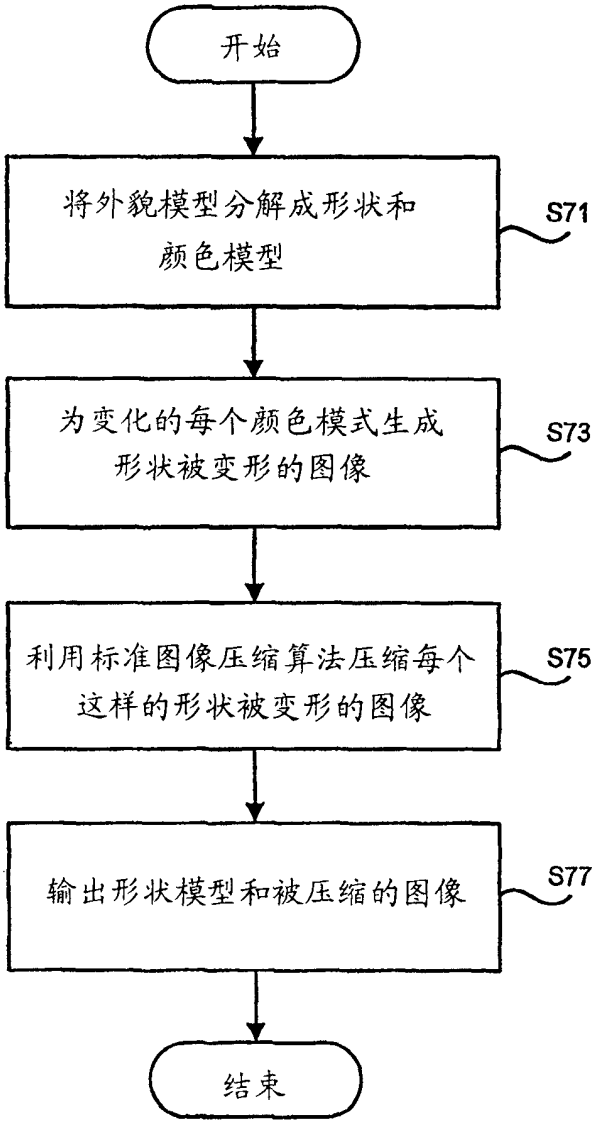

开始
将外貌模型分解成形状和颜色模型
S71
为变化的每个颜色模式生成形状被变形的图像
S73
利用标准图像压缩算法压缩每个这样的形状被变形的图像
S75
输出形状模型和被压缩的图像
S77
结束

图 5a

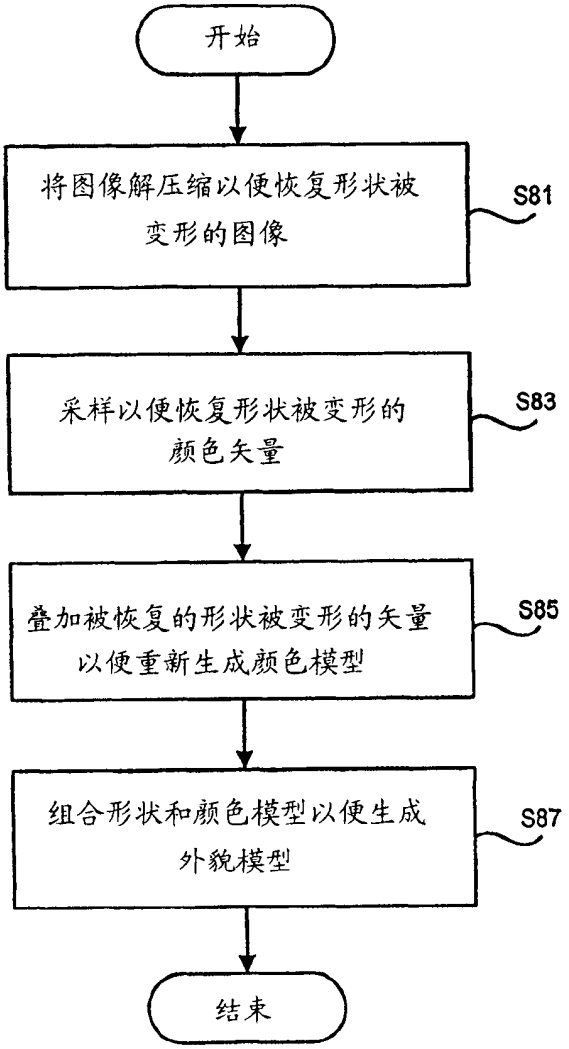

开始
将图像解压缩以便恢复形状被变形的图像
S81
采样以便恢复形状被变形的颜色矢量
S83
叠加被恢复的形状被变形的矢量以便重新生成颜色模型
S85
组合形状和颜色模型以便生成外貌模型
S87
结束

图 5b

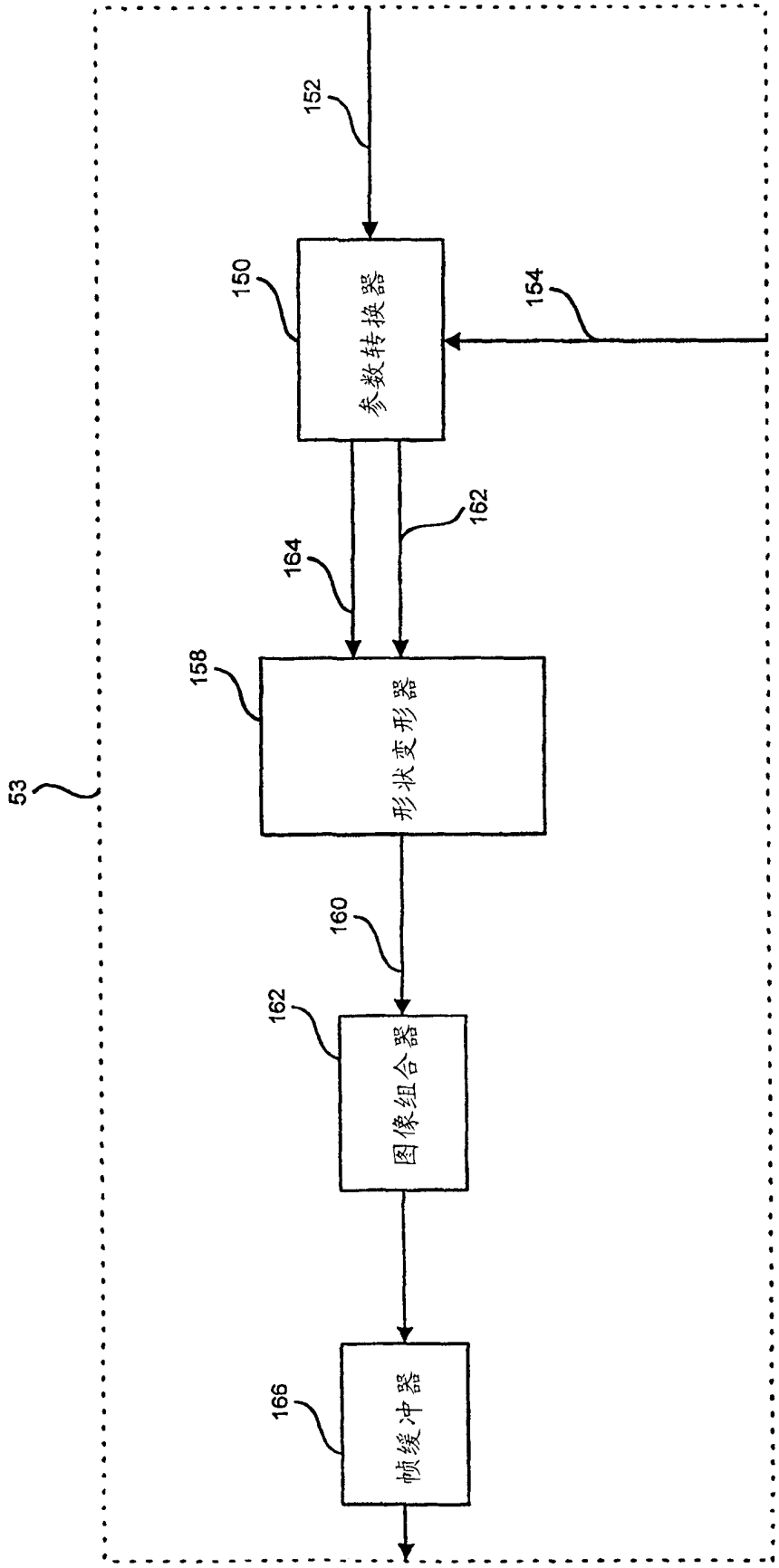


图 6

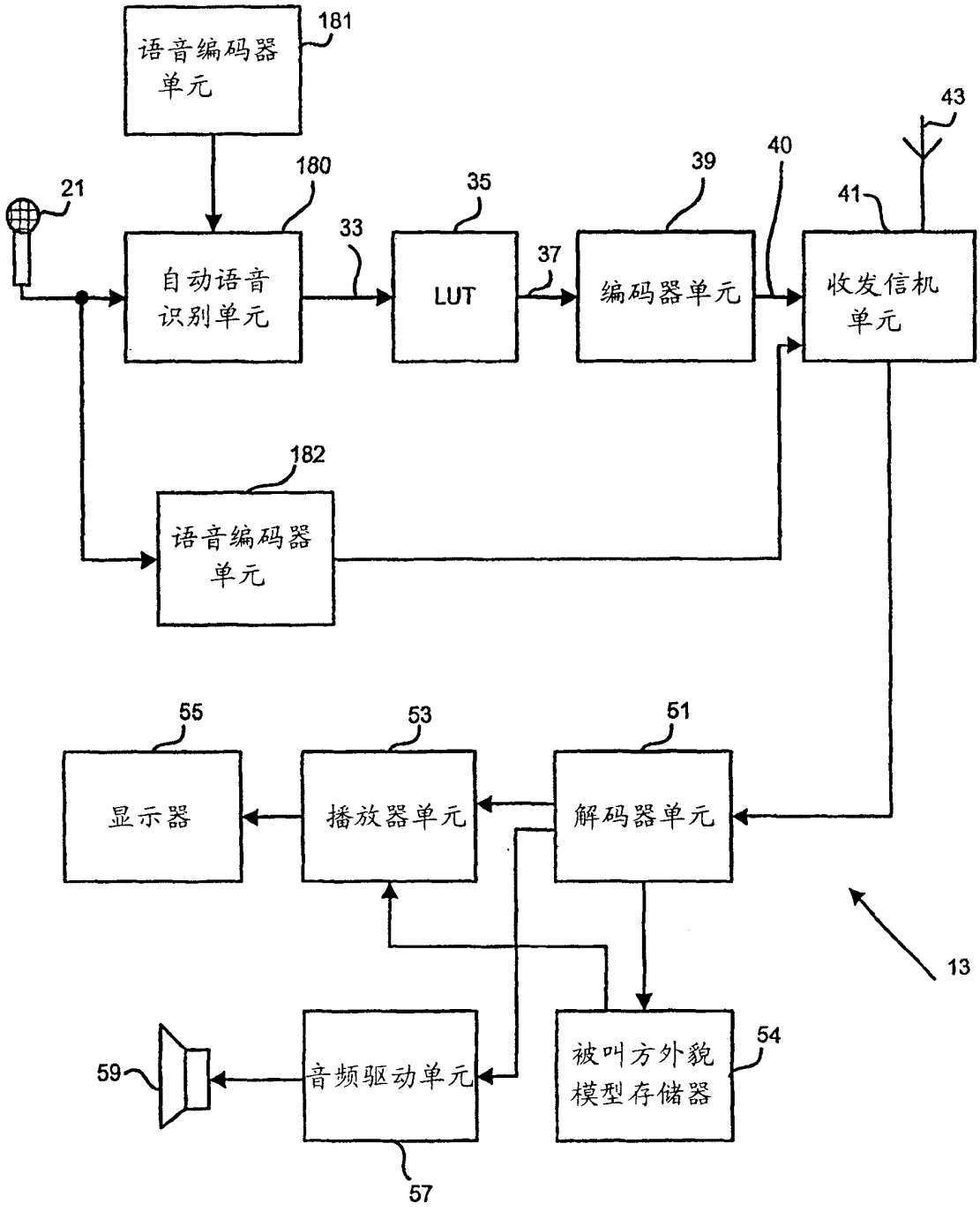

语音编码器单元
181
21
180
自动语音识别单元
33
35
LUT
37
39
编码器单元
40
41
43
收发信机单元
182
语音编码器单元
55
显示器
53
播放器单元
51
解码器单元
13
59
音频驱动单元
57
被叫方外貌模型存储器
54

图 7

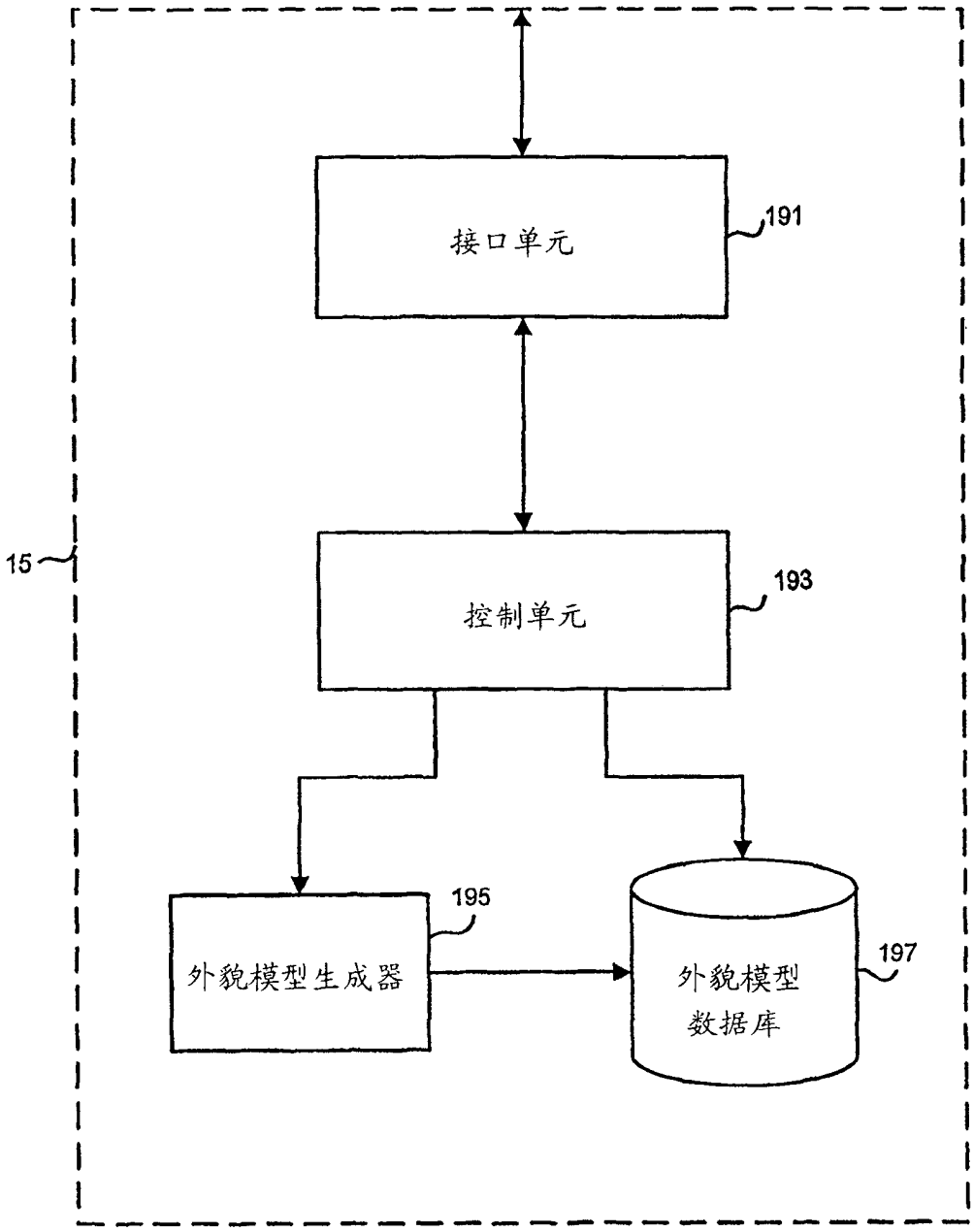


图 8

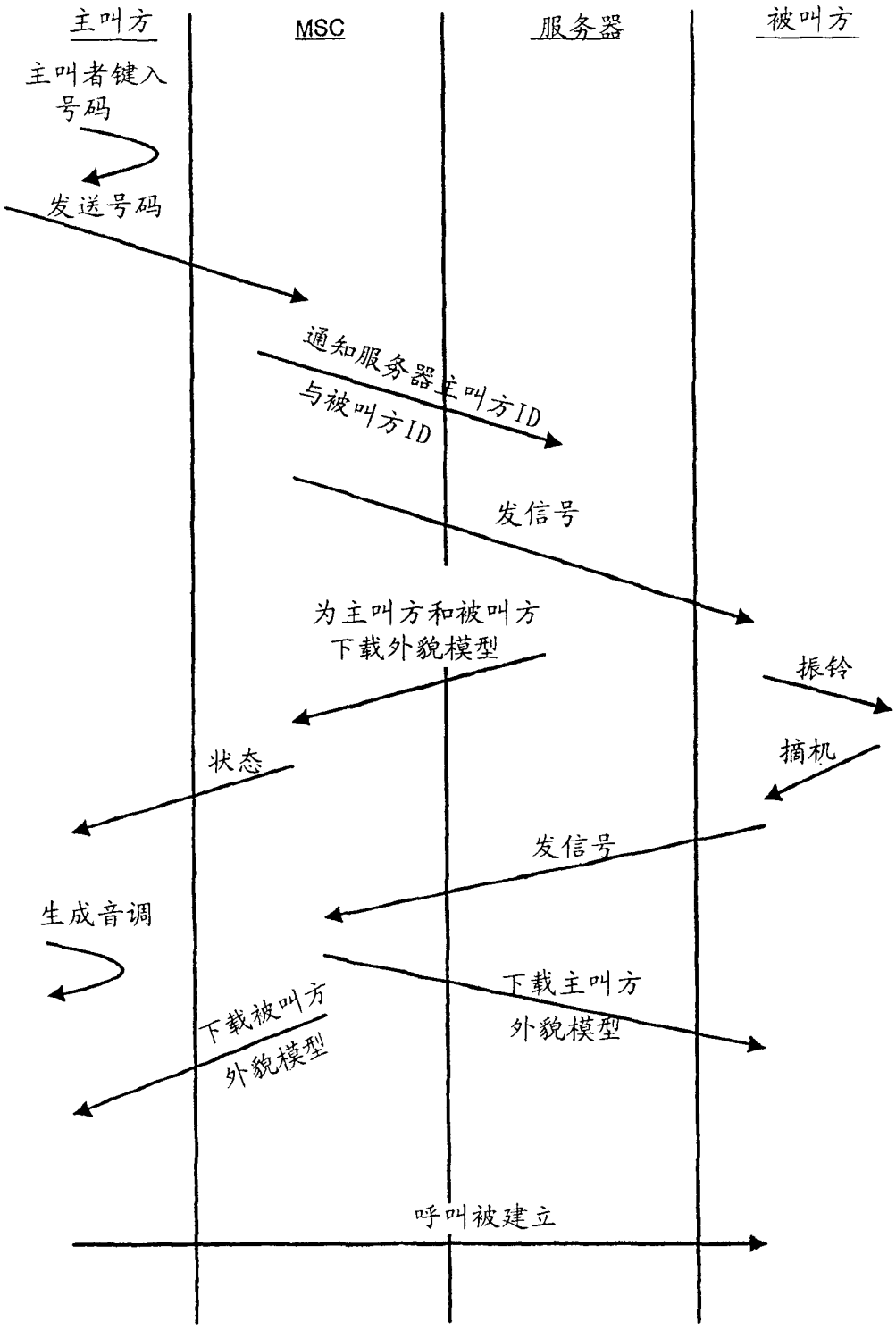

主叫方
MSC
服务器
被叫方
主叫者键入号码
发送号码
通知服务器主叫方ID
与被叫方ID
发信号
为主叫方和被叫方下载外貌模型
振铃
状态
摘机
发信号
生成音调
下载主叫方外貌模型
下载被叫方外貌模型
呼叫被建立

图 9

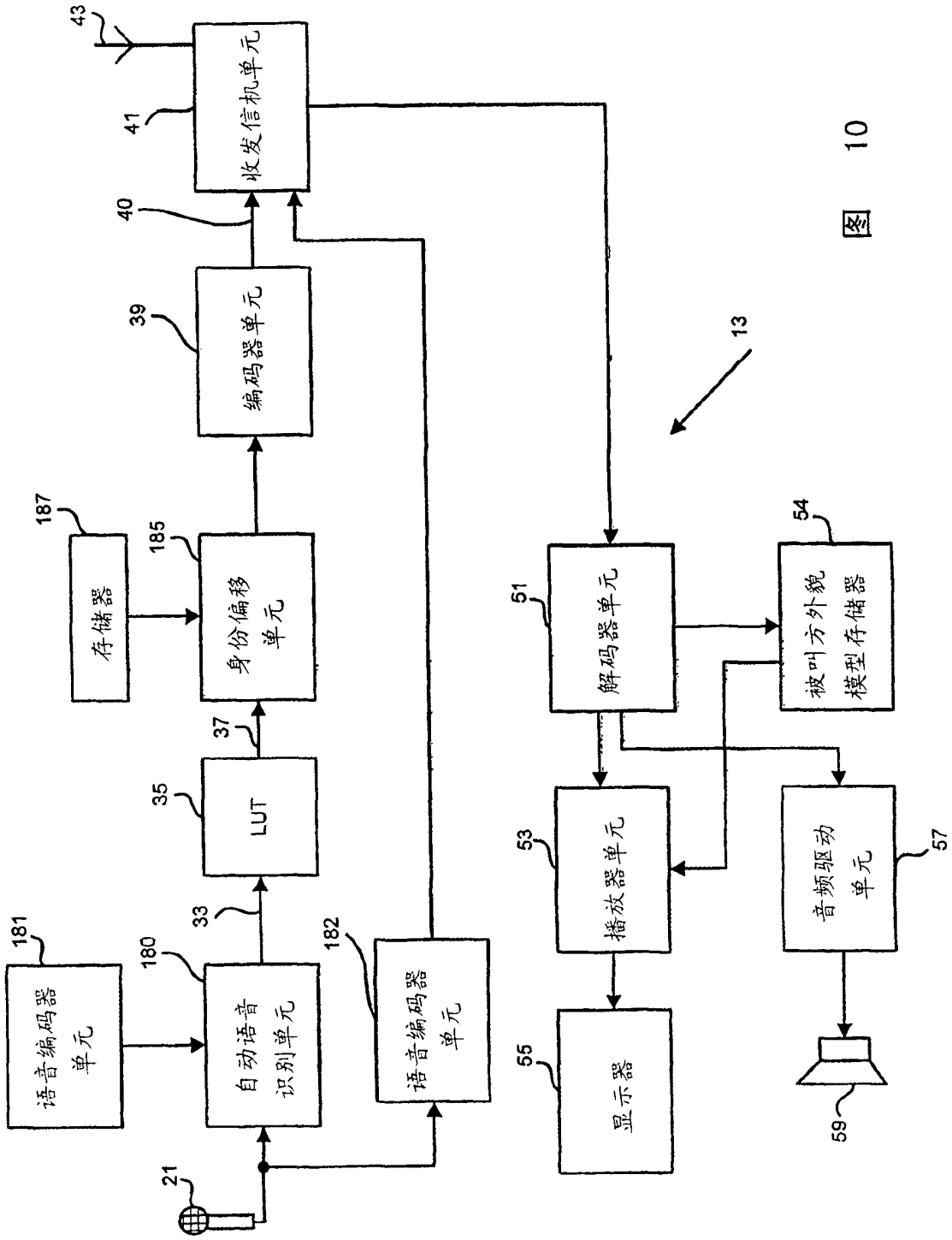

43
41
收发信机单元
40
39
编码器单元
187
存储器
185
身份偏移单元
37
35
LUT
33
181
语音编码器单元
180
自动语音识别单元
182
语音编码器单元
21
13
51
解码器单元
54
被叫方外貌模型存储器
53
播放器单元
57
音频驱动单元
55
显示器
59

图 10

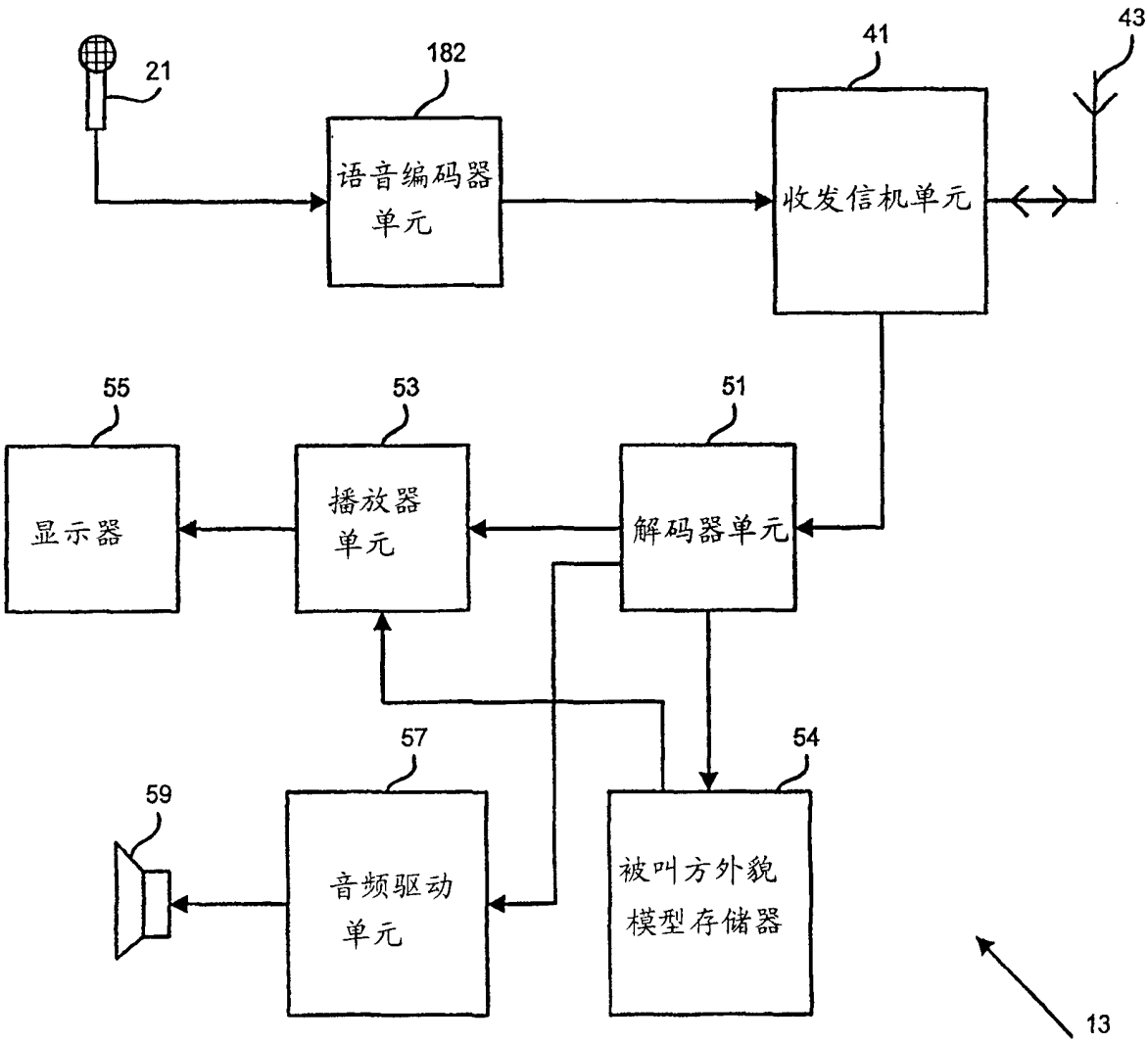

21
182
41
43
语音编码器单元
收发信机单元
55
53
51
显示器
播放器单元
解码器单元
57
54
59
音频驱动单元
被叫方外貌模型存储器
13

图 11

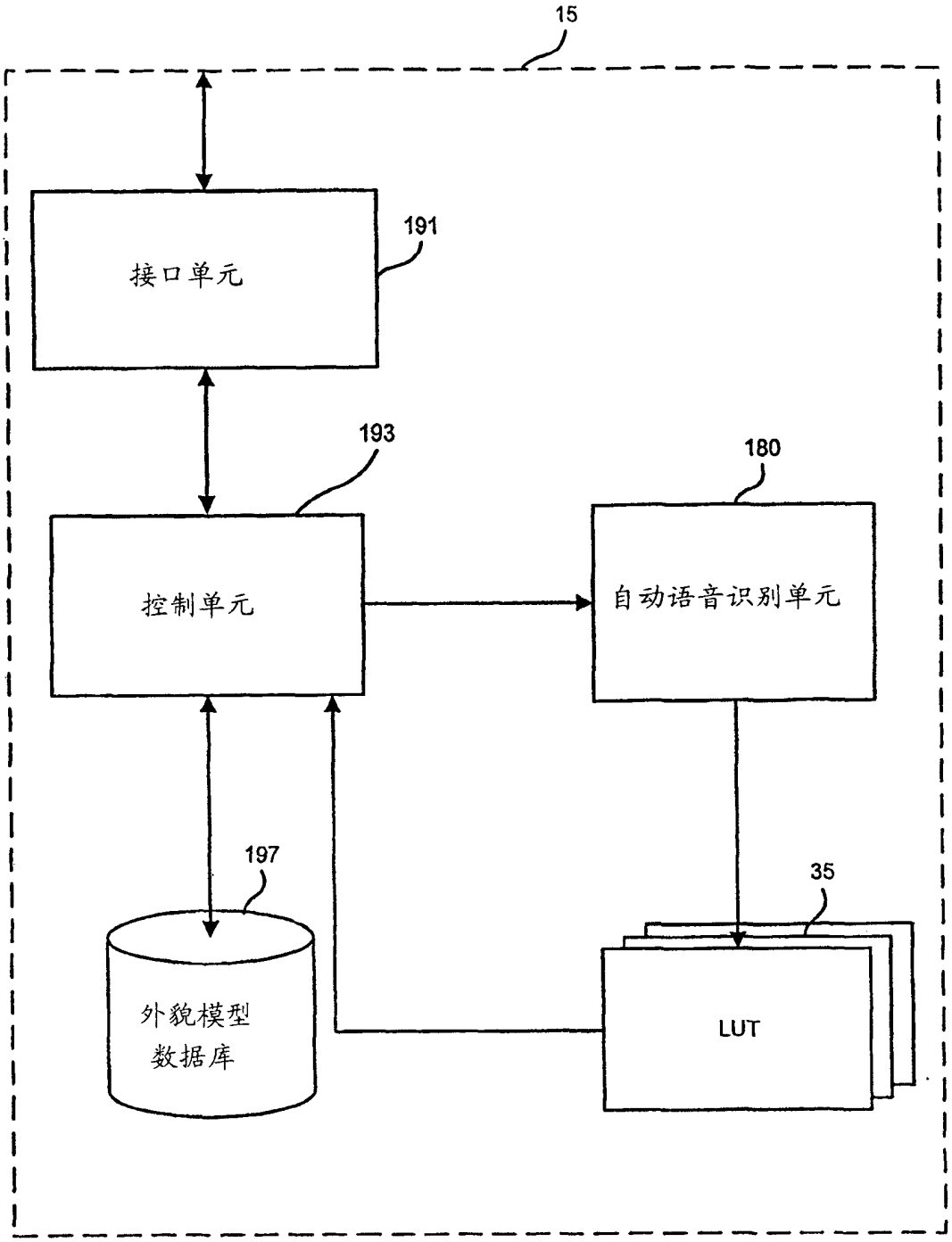


图 12

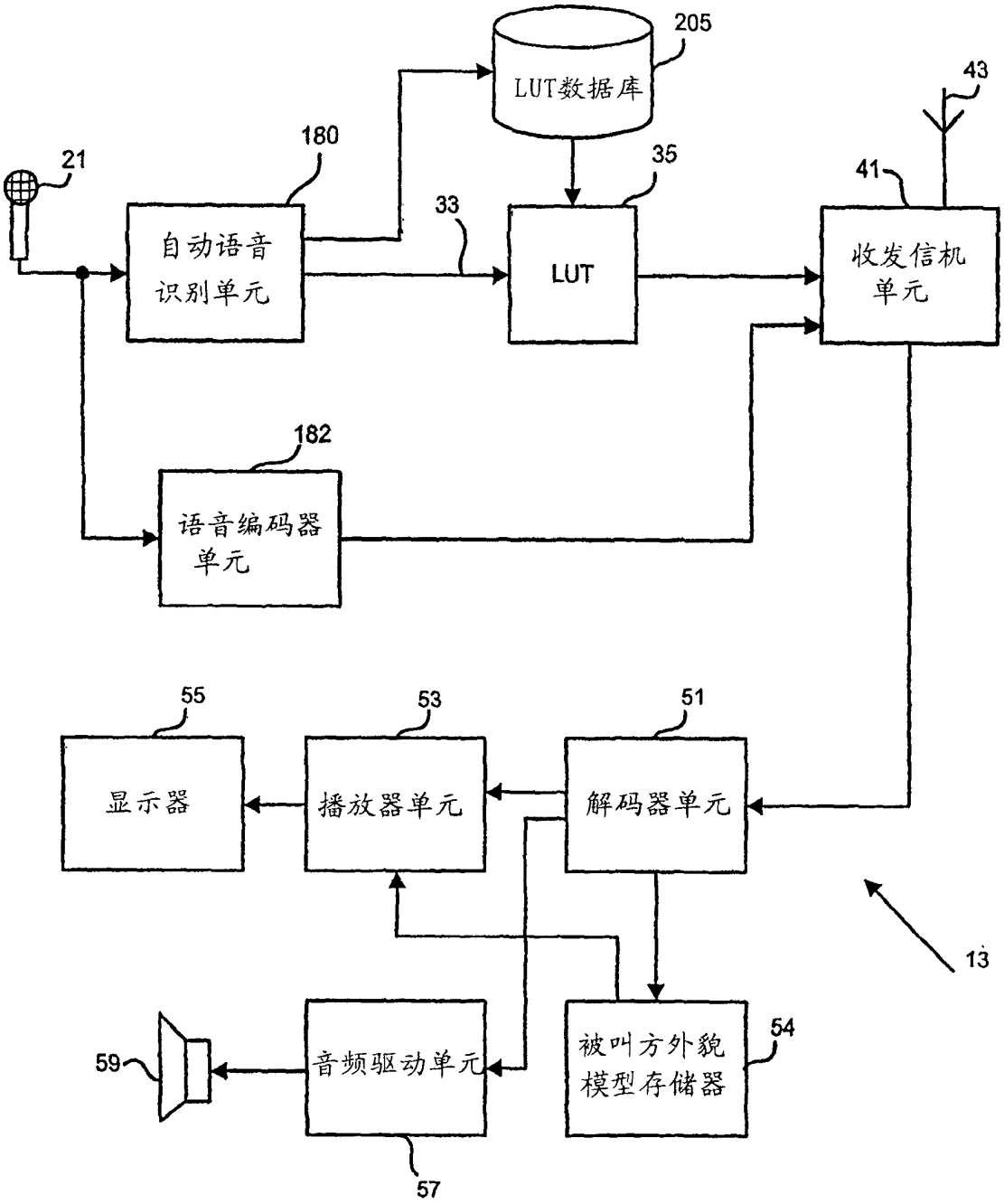

205
LUT数据库
43
180
21
35
41
33
自动语音
识别单元
LUT
收发信机
单元
182
语音编码器
单元
55
53
51
显示器
播放器单元
解码器单元
13
59
音频驱动单元
被叫方外貌
模型存储器
54
57

图 13

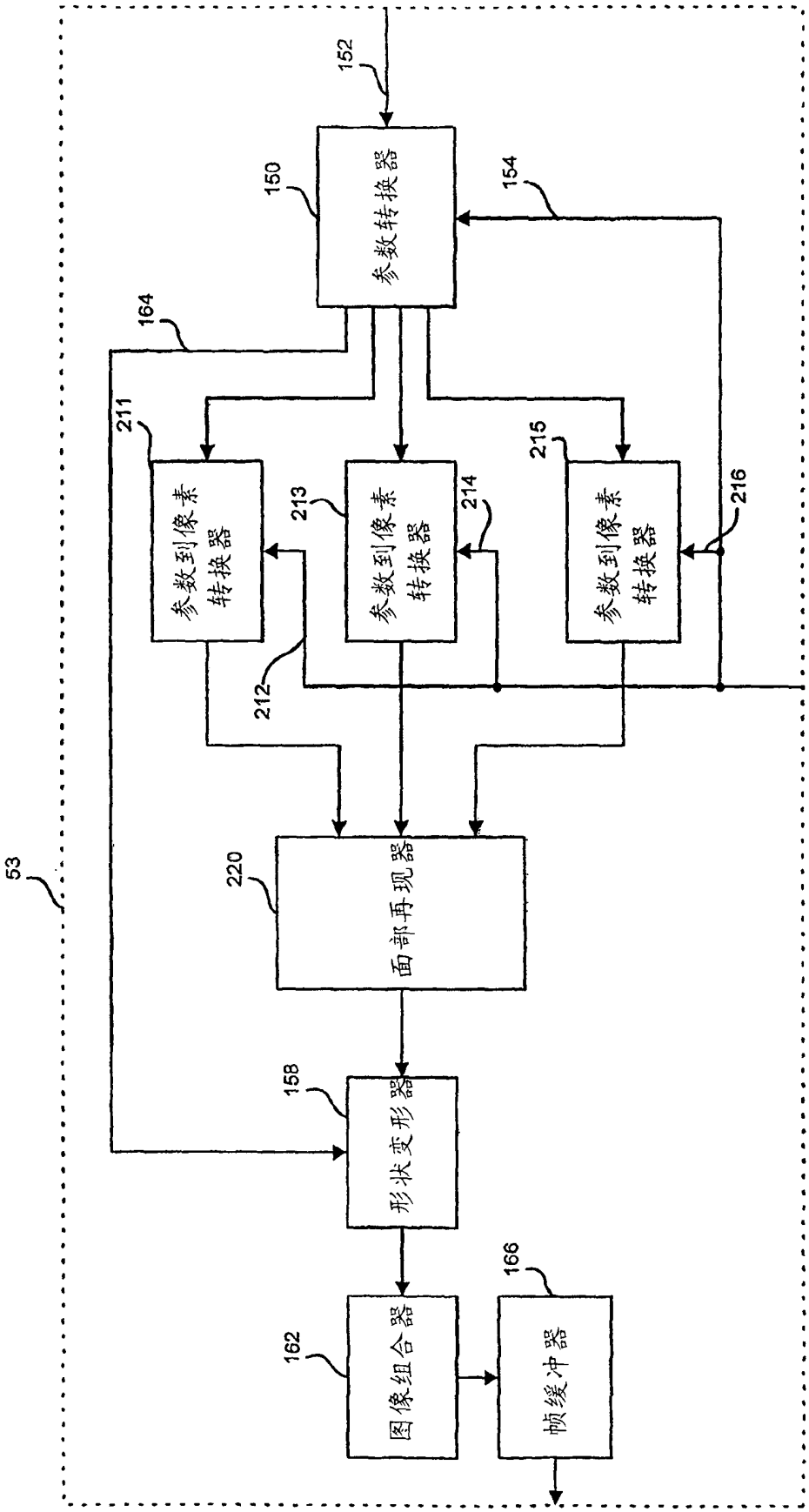

53
152
150
参数转换器
154
164
211
参数到像素转换器
213
214
参数到像素转换器
215
216
参数到像素转换器
212
220
面部再现器
158
形状变形器
162
图像组合器
166
帧缓冲器

图 14

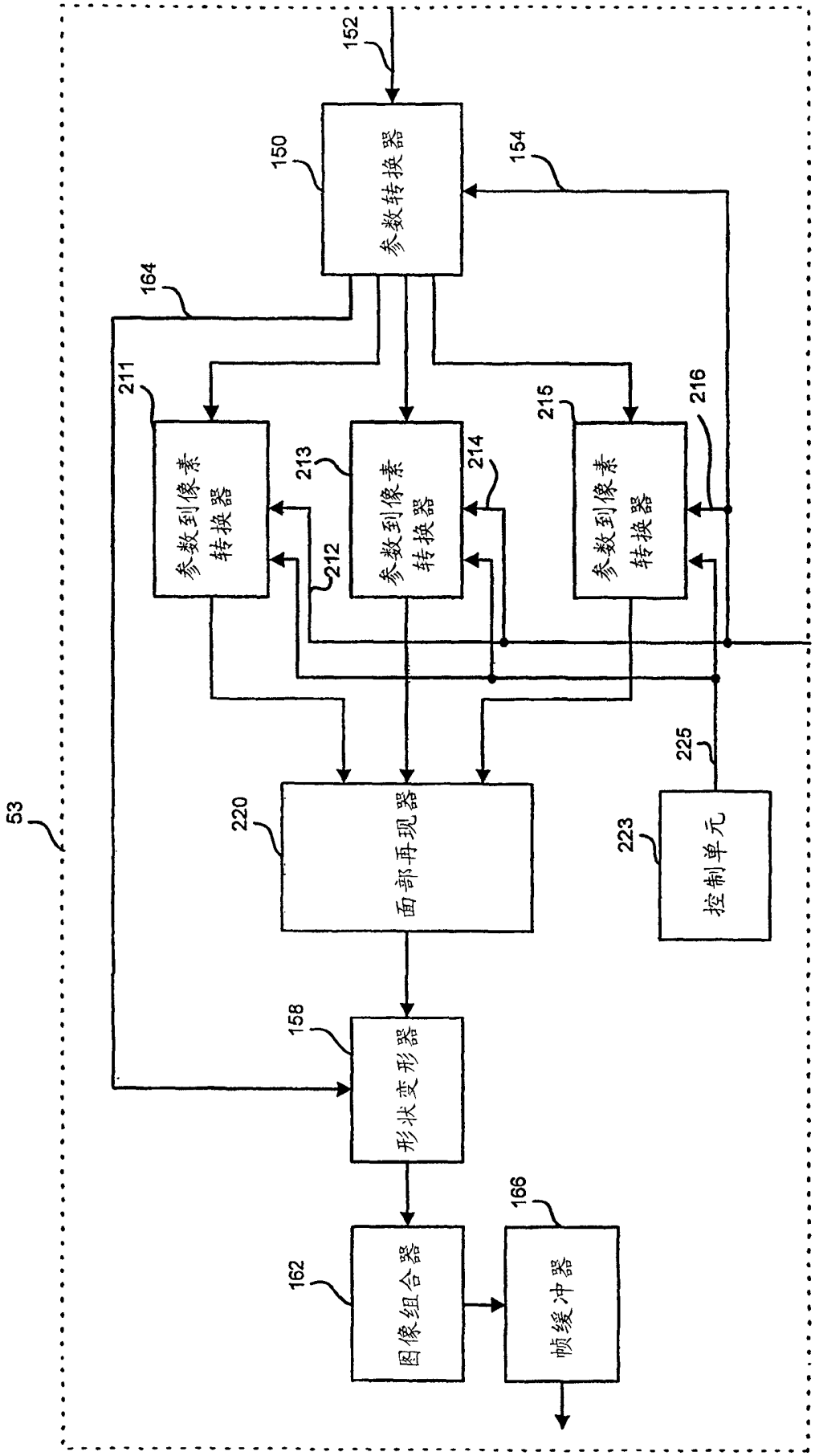

53
152
150
参数转换器
154
164
211
参数到像素转换器
213
212
参数到像素转换器
214
215
参数到像素转换器
216
225
220
面部再现器
223
控制单元
158
形状变形器
166
162
图像组合器
帧缓冲器

图 15

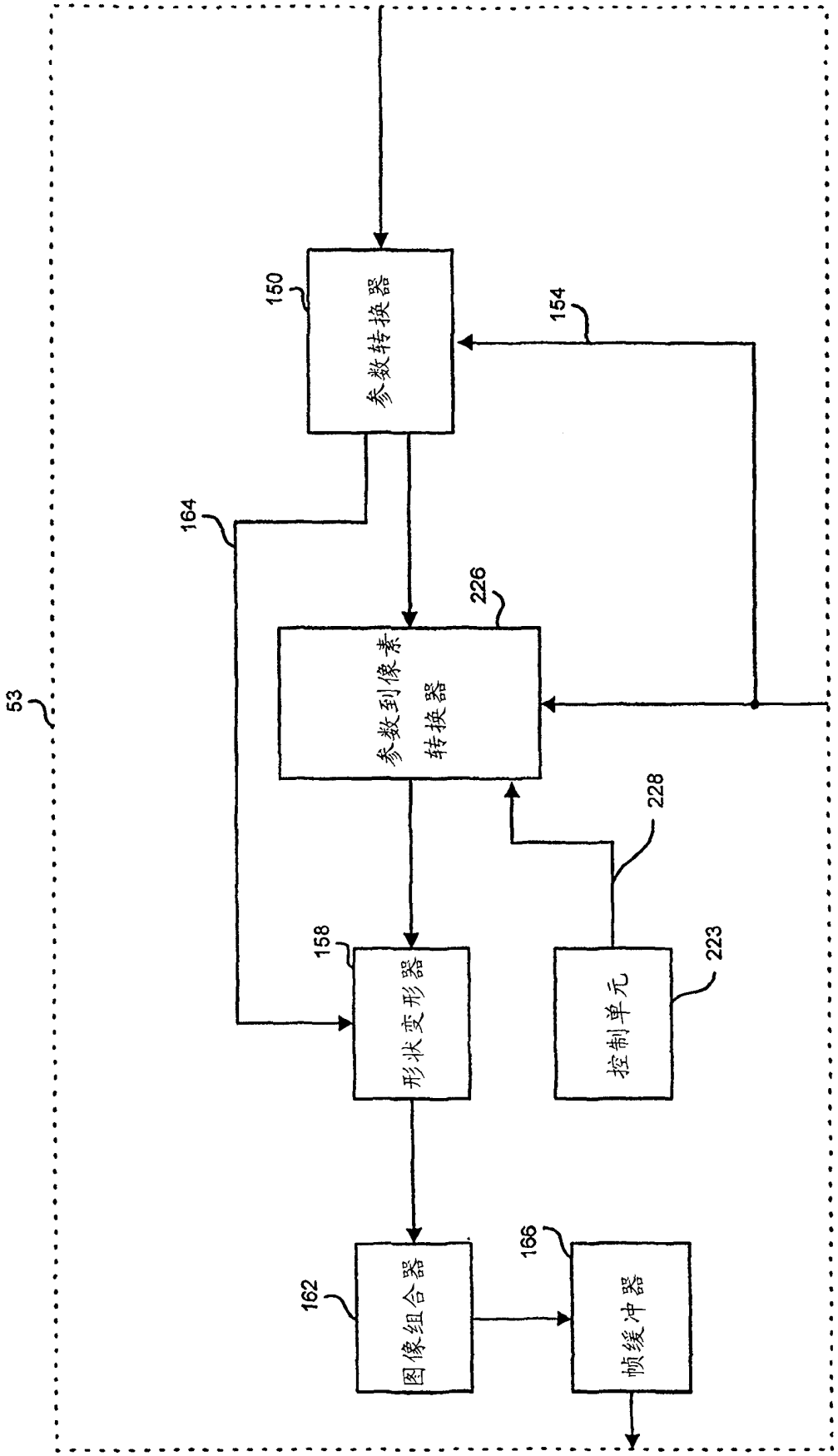

53
150
参数转换器
154
164
226
参数到像素
转换器
228
158
形状变形器
223
控制单元
162
图像组合器
166
帧缓冲器

图 16