

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04L 29/06

H04Q 7/22



[12] 发明专利申请公开说明书

[21] 申请号 03822752.5

[43] 公开日 2005 年 10 月 19 日

[11] 公开号 CN 1685686A

[22] 申请日 2003.8.4 [21] 申请号 03822752.5

[30] 优先权

[32] 2002.8.14 [33] IT [31] TO2002A000724

[86] 国际申请 PCT/EP2003/008604 2003.8.4

[87] 国际公布 WO2004/019583 英 2004.3.4

[85] 进入国家阶段日期 2005.3.24

[71] 申请人 意大利电信股份公司

地址 意大利米兰

[72] 发明人 吉安尼·L·古格利厄米

吉安卢卡·弗兰希尼

克劳迪奥·兰德 卢卡·鲍索尼图

[74] 专利代理机构 中国国际贸易促进委员会专利商
标事务所

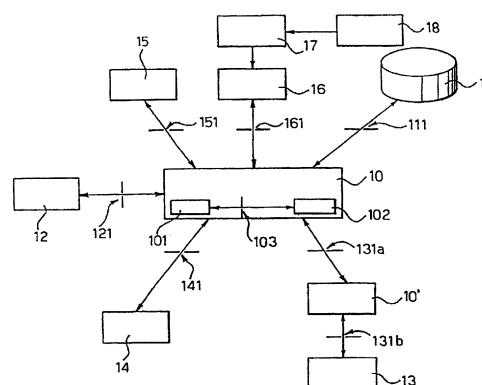
代理人 李德山

权利要求书 5 页 说明书 15 页 附图 4 页

[54] 发明名称 用于在电信网上传输消息的方法和系统及相关发送器终端

[57] 摘要

本方法包括以下步骤：从发送者终端(18)接收(17)文本消息，诸如 SMS 消息，综合(16)所述文本消息与视频内容，以生成多媒体消息，以及把所述多媒体消息以 MMS 消息的形式发送(10)到至少一个接收终端(12, 13, 14)。由此保证传统的移动终端(例如，GSM)与新一代移动终端(例如，UMTS)的可能的共存和互操作。



1. 用于在电信网上传输消息的方法，其特征在于，该方法包括以下步骤：

- 5 - 从发送者终端(18) 接收(17) 文本消息，
 - 综合(16)所述文本消息与视频内容，以生成多媒体消息，以及
 - 把所述多媒体消息以 MMS 消息的形式发送(10) 到至少一个接收终端(12, 13, 14)。

2. 如权利要求 1 所述的方法，其特征在于，它包括接收(17)SMS
10 消息的形式的所述文本消息的步骤。

3. 如权利要求 1 或 2 所述的方法，其特征在于，它包括以下步骤：

- 通过识别所述接收终端的特征而识别能够接收所述多媒体消息的接收终端(12, 13, 14) 的类型，以及
15 - 使所述 MMS 消息适合(16, 326;10)于所述接收终端(12, 13, 14) 的特征。

4. 如权利要求 3 所述的方法，其特征在于，它包括综合所述文本消息与生成的视频内容(326)，以使得所述多媒体消息适合于所述接收终端(12, 13, 14) 的特征的步骤。

20 5. 如权利要求 3 所述的方法，其特征在于，它包括以下步骤：

- 用与接收终端(12, 13, 14) 的特征无关地确定的视频内容补充所述文本消息，以及
 - 使由此得到的多媒体消息适合(10)于所述接收终端(12, 13, 14) 的特征。

25 6. 如以前的权利要求的任一项所述的方法，其特征在于，它包括在由以下项目组成的组内选择所述视频内容的步骤：

- 动画的图像，
 - 背景图像，以及
 - 具有可变的观察点的图像。

7. 如以前的权利要求的任一项所述的方法, 其特征在于, 它包括从所述文本消息合成能够与所述多媒体消息内的所述视频内容相联系的语音信号(V) 的步骤。

8. 如权利要求 7 所述的方法, 其特征在于, 它包括生成所述动画图像(308, 310)作为讲出相应于所述文本消息的合成的语音信号的人物的图像的步骤。

9. 如权利要求 8 所述的方法, 其特征在于, 它包括借助于文本动画系统(308, 310)生成所述人物的图像的步骤。

10. 如以前的权利要求的任一项所述的方法, 其特征在于, 它包括综合(328)所述消息与背景音乐(330)的步骤。

11. 如以前的权利要求的任一项所述的方法, 其特征在于, 它包括把动画的 GIF 图像包括在所述视频内容中的步骤。

12. 如以前的权利要求 6, 8, 9 或 11 的任一项所述的方法, 其特征在于, 所述动画的图像是以 Hz 量级的动画采样速率得到的。

13. 如以前的权利要求的任一项所述的方法, 其特征在于, 它包括鉴于所述文本消息的接收(17), 把用于标识在由以下项目组成的组内选择的所述视频内容的特征的至少一个字段与所述文本消息相联系的步骤:

- 被使用来显示所述文本消息的虚拟人物(P), 以及
- 所述多媒体内容的背景(A) 。

14. 如以前的权利要求的任一项所述的方法, 其特征在于, 它包括在所述发送者终端(18) 提供用于选择所述视频内容和所述接收终端(12, 13, 14) 的脚本功能的步骤。

15. 如以前的权利要求的任一项所述的方法, 其特征在于, 它包括在所述发送者终端(18) 提供用于自动校正可能被包含在所述文本消息中的任何错误的功能的步骤。

16. 如以前的权利要求的任一项所述的方法, 其特征在于, 它包括把用于选择地修正所述视频内容的特征的元信息与所述文本消息相联系的步骤。

17. 如以前的权利要求的任一项所述的方法, 其特征在于, 它包括把用于选择地修正所述视频内容的特征的情绪标记形式的附加信息与所述文本消息相联系的步骤。

18. 如以前的权利要求的任一项所述的方法, 其特征在于, 所述
5 视频内容是从以下项目中选择的:

- 按帧排序的动画的 GIF 图像, 所述文本消息的各个部分与其相联系,
- 伴着压缩的音频的动画的 GIF 图像, 以及
- 连同音频一起的视频剪辑。

19. 用于在电信网上传输消息的系统, 其特征在于, 系统包括:

- 接收模块(17), 用于从发送者终端(18) 接收文本消息,
- 处理组(16), 具有视频信息的至少一个数据库(302, 314, 330)和用于综合所述文本消息与视频内容以生成多媒体消息的至少一个综合模块(326, 328), 以及
- 15 - 发送模块(10), 用于把所述多媒体消息以 MMS 消息的形式发送到至少一个接收终端(12, 13, 14)。

20. 如权利要求 19 所述的系统, 其特征在于, 所述接收模块(17)被配置成从所述发送者终端(18)接收以 SMS 消息的形式的文本消息。

21. 如权利要求 19 或 20 所述的系统, 其特征在于, 它包括:

- 20 - 检测模块(300;10), 用于通过识别所述接收终端的特征(TD)而检测要作为所述多媒体消息的接收者的接收终端(12, 13, 14) 的类型, 以及

- 用于使所述 MMS 消息适合于所述接收终端(12, 13, 14) 的特征的模块(16, 326;10)。

22. 如权利要求 21 所述的系统, 其特征在于, 所述综合模块(326, 328) 被配置成用于综合所述文本消息与生成的视频内容(326), 以使得所述多媒体消息适合于所述接收终端(12, 13, 14) 的特征。

23. 如权利要求 21 所述的系统, 其特征在于, 所述集成模块(326, 328) 被配置成综合所述文本消息和与接收终端(12, 13, 14) 的特征无

关地确定的视频内容,以及系统具有与其相联系的用于 MMS 消息(10)的传输的模块(10),所述 MMS 消息(10)被配置成使得所述多媒体消息经受使其适合(10)于所述接收终端(12, 13, 14) 的特征的步骤(10)。

24. 如以前的权利要求 19 到 23 的任一项所述的系统,其特征在
5 于,它包括:至少一个视频生成器模块(302, 308, 310),用来生成在以下项目内选择的视频内容:

- 动画的图像,
- 背景图像,以及
- 具有可变的观察点的图像。

10 25. 如以前的权利要求 19 到 24 的任一项所述的系统,其特征在于,它包括:语音合成器(306),用来从所述文本消息合成能够与所述多媒体消息内的所述视频内容相联系(326)的语音信号(V)。

26. 如权利要求 25 所述的系统,其特征在于,将至少一个运动生成模块(308, 310)与所述视频生成模块(302, 308, 310)和所述语音
15 合成器(306)相联系,以生成所述动画图像作为讲出相应于所述文本信号的合成的语音信号的人物的图像。

27. 如权利要求 26 所述的系统,其特征在于,所述运动生成模块(308, 310)是诸如 JoeXpress®系统那样的文本动画系统。

28. 如以前的权利要求 19 到 27 的任一项所述的系统,其特征在
20 于,它包括:背景音乐的数据库(330),与所述至少一个综合模块(326, 328)共同运行,以综合所述 MMS 消息与背景音乐。

29. 如以前的权利要求 19 到 28 的任一项所述的系统,其特征在于,所述综合模块(326, 328)被配置成把动画的 GIF 图像包括在所述视频内容中。

25 30. 如以前的权利要求 24, 26, 27 或 29 的任一项所述的系统,其特征在于,所述综合模块(326, 328)被配置成把具有 Hz 量级的动画采样速率的动画的图像包括在所述视频内容中。

31. 如以前的权利要求 19 到 30 的任一项所述的系统,其特征在于,所述接收模块(17)包括信息提取方框(300),用于从所述发送者终

端(18)接收的所述文本消息中提取标识在由以下项目内选择的所述视频内容的特征的至少一个字段:

- 被使用来显示所述文本消息的虚拟人物(P), 以及
- 所述多媒体内容的背景(A)。

5 32. 如以前的权利要求 19 到 31 的任一项所述的系统, 其特征在于, 具有所述至少一个视频信息的数据库(302, 314, 330)和用来综合所述文本消息与视频内容的所述至少一个综合模块(326, 328)的所述处理组(16)被配置成生成在以下项目内选择的多媒体消息:

- 按帧排序的动画的 GIF 图像, 所述文本消息的各个部分与其相
- 10 联系,
- 伴有压缩的音频的动画的 GIF 图像, 以及
- 连同音频一起的视频剪辑。

33. 用于如以前的权利要求 19 到 32 的任一项所述的系统的发送者终端, 其特征在于, 所述发送者终端(18) 被提供有用于选择所述视

15 频内容和所述接收终端(12, 13, 14)的脚本功能。

34. 用于如以前的权利要求 19 到 32 的任一项所述的系统的发送者终端, 其特征在于, 所述发送者终端(18) 被提供有用于自动校正可能被包含在所述文本消息中的任何错误的功能。

35. 用于如以前的权利要求 19 到 32 的任一项所述的系统的发送者终端, 其特征在于, 所述发送者终端(18) 被提供有用于把选择地修正所述视频内容的特征的元信息与所述文本消息相联系的功能。

20

36. 用于如以前的权利要求 19 到 32 的任一项所述的系统的发送者终端, 其特征在于, 所述发送者终端(18) 被提供有用于把选择地修正所述视频内容的特征的情绪标记形式的附加信息与所述文本消息相

25 联系的功能。

用于在电信网上传输消息的方法 和系统及相关发送器终端

5

技术领域

本发明涉及电信网上消息的传输。

背景技术

10

例如按照 UMTS 标准(通用移动通信系统)或 GSM/GPRS 标准(全球移动通信系统和通用分组无线电业务的缩写)的新一代移动终端的引入使得能够传输和在终端上显示具有包括如文本、声音、图像、也可能还有运动图像等的不同要素的多媒体内容的消息。所述消息当前被表示为 MMS, 即, 多媒体消息传送系统的缩写。

15

发送所述消息的能力引起不同类型的问题。

第一方面, 必须保证所述消息可以通过使用如移动电话的设备相对容易地被构建, 该设备由于减小的尺寸和处理能力, 不能理想地适合于生成具有复杂内容的消息。

20

第二方面, 希望具有发送和接收 MMS 消息的能力的终端能够与老的一代终端, 诸如按照 GSM 标准运行的移动终端的并只能生成当前称为 SMS - 短消息业务的缩写 - 的那种文本消息的老一代终端共存和交互。有理由认为在所有当前流行的终端被替代之前的相当长的一段时间内两种技术肯定会共存。

25

发明内容

本发明的目的是支持具有发送如 SMS 消息那样的文本消息的能力的终端与能够接收 MMS 消息的终端之间的共存和交互。

按照本发明, 所述目的是由于其特征为在以下的权利要求中具体地阐述的方法而达到的。本发明还包括相关的系统以及相应的发送器

终端。

实际上,按照本发明的解决方案允许老一代终端--能够发送 SMS 文本消息的--导致具有要到 MMS 终端的多媒体内容的消息的生成。

5 在当前的优选实施例中,按照本发明的解决方案允许提供一种业务,它把纯文本消息自动变换成多媒体消息,从而变换成比起由纯文本构成的开始的"消息"更丰富的"消息"。

在当前的优选实施例中,按照本发明的解决方案提供了使用由同一个申请人制造的、并由注册商标 JoeXpress®标识的基于文本或自然音频来自动化一个三维人物的系统。

10 在这方面,查询文件 EP-A-0 991 023、EP-A-0 993 197、和 WO-A-01/75805 是有用的。所讨论的系统能够把文本或记录的语音变换成讲出处理的句子的人物的运动。所述运动也包括没有与讲的话、面部表情和身体运动联系起来的运动。系统还能够操纵其它要素,诸如人物外貌的个性(例如,头发颜色、眼睛颜色、衣着方式等等)、人物所处的位置,观察点的移动,背景音乐。所有这些都在从所提供的有限数目的输入参数来构建视频剪辑时同时发生。

20 这样,按照本发明的解决方案例如允许根据被包含在起始 SMS 消息中的内容生成要到 MMS 终端的动画。在这种情形下,结果是包括诸如场景描述部分(以"同步的多媒体集成语言"或 SMIL 表示)和包含要被插入消息中的多媒体对象的部分的不同的部分的 MMS 消息,在这些部分之间自动生成动画。

25 第一代 MMS 终端受到对于消息内容的相当严格的约束限制:具体地,不支持视频且消息的最大尺寸是 30 千字节。按照本发明的解决方案的优选实施例因此允许以生成的 MMS 语言引入具有小的尺寸的动画。具体地,视频被变换成按照 GIF 标准(用于图形交换格式的缩写)的图像,这个图像经受使用相当低的动画采样速率的动画,即,约 1Hz。

而且,原始的文本被在序列的各个帧之间子划分。通过这样做,利用尺寸例如约为 100x80 像素(当前市场上的 MMS 终端的显示单元

的尺寸)的动画,可以生成包含具有复杂的模型和场景的持续约 15 秒或在较简单的模型的情况下更长的动画的消息,这允许在动画的 GIF 图像内更高的压缩比。

如果消息的总的尺寸是有限的(例如,30 千字节)而使得同时发送
5 视频和音频是成问题的,则在观看动画的 GIF 图像期间有可能使得终端重现被插入在消息中的旋律,而不是语音消息:这种类型的声音(“振铃声”)能够被包含在非常小数目的字节中。

在对于消息的尺寸存在不太严格的约束时,按照本发明的解决方案允许发送与由例如语音合成器生成的与动画有关的音频,而不是发
10 送在帧里面的文本或甚至与其并行地发送。在这种形下,有可能甚至从自然音频自动生成 MMS 消息,在这种情形下,动画通过由语音识别器实行的处理的结果引导。能够实行上述功能的语音合成器和语音识别器当前在本领域是可得到的。

除了动画以外,MMS 消息还可以有利地考虑旨在于包含对插入
15 例如所谓的“标记”(logo)和/或广告标语有用的更多文本、旋律和图像的部分。

附图说明

现在参照附图,完全作为非限制的例子来描述本发明,其中:
20 图 1 在功能体系结构级别上显示能够按照本发明运行的系统的结构,

图 2 是显示按照本发明的用于发送消息的步骤的流程图,以及
图 3,其包括分别表示为 3A 和 3B 的两个部分,重现显示按照本
发明的系统的安排的可能的形式的功能方框图的两个邻近的部分。

25

具体实施方式

这里提供的说明参照现在至少是对于本发明的可能的使用来说最有吸引力的一个的应用情形,即,把在 GSM 移动终端中作为 SMS 消息生成的文本消息变换成将要在按照 UMTS 标准运行的网络上传

输的 MMS 消息。

无论如何，按照本发明的解决方案也可以应用于不同地生成的、例如电子邮件消息的形式的文本消息，以及它可被使用来在诸如支持这样的传输的任何类型的网络上发送 MMS 消息，因此不限于 UMTS 网络。

在图 1 上，标号 10 统一表示具有 MMS 中继/服务器的功能的模块，以及为此包括：具有中继功能的子模块，被表示为 101；和具有服务器功能的子模块，被表示为 102，通过被表示为 103 的接口被互相连接。自然地，子模块 102 和 103 也可以互相集成。

标注数字 11 代替地表示 MMS 业务的用户的数据库。这基本上 10 是其中为可获得 MMS 业务的每个用户记录电话号码(或等同的指示)和有所讨论的用户采用的终端类型的信息的数据库。

数据库 11 通过接口 111 被连接到模块 10。

标注数字 12 和 13 表示在网络中被连接到模块 10(这可以典型地 15 发生在 UMTS 网络)以便能够接收 MMS 消息的两个用户。

被表示为 12 的用户是被直接包括在其中连结有模块 10 的网络中的用户。相关的连接因此是通过被表示为 121 的接口的直接类型。

而被表示为 13 的用户是被标称地连结到另一个移动网的用户。

在这种情形下，与模块 10 的连接不是直接的，而是借助于被表示为 131a 和 131b 的相应的接口通过基本上类似于接口 10 的附加模块 10' 达到的。 20

用户 12 和用户 13 的不同表示，是为了强调在其中多个电信网络以通用互联网共同运行的情形或漫游情形中也应用按照本发明的解决方案的可能性。

标注数字 14 表示通过各个接口 141 被连接到模块 10，以便能够 25 作为 MMS 消息的接收者运行的服务器，诸如电子邮件服务器。

最后，标注数字 15 表示通过各个接口 151 被连接到模块 10 的用于对 MMS 消息业务的提供进行记账的系统。

这里描述的系统体系结构和各种构成的单元相应于被认为在本

领域中全都知道的解决方案。这些解决方案已能被使用于在电信网络内发送 MMS 消息(这样的新一代移动网络按照 UMTS 标准运行)。这个事实使得在这里提供对于所讨论的体系结构和单元的更详细的说明成为多余的。

5 按照本发明的解决方案的重要的特征是通过以下事实给出的、优选地通过各个接口 161, 将能够把诸如来自 SMS 消息管理中心 17(通常称为缩写 SMSC)的 SMS 消息的纯文本消息变换成具有多媒体内容的消息的模块或子系统 16 与模块 10 相联系。在模块 10 中进行可能的进一步处理后, 所述消息可以由模块 10 以要到诸如图 1 所示的用户
10 12, 13 和 14 那样的用户的 MMS 消息的形式进行广播。

具体地, 模块 10 可被配置成允许确定的消息 MMS 传输到多个接收者或接收者列表。因此, 虽然此后将几乎仅仅参考从 SMS 消息生成被发送到单个接收者的 MMS 消息, 但按照本发明的解决方案容易
15 适合于允许把所讨论的 MMS 消息广播到例如借助于 http 请求或借助于被发送到模块 10 的 ftp 请求规定的接收者列表。

如前所述, 模块 16 的核心是由用于创建由文本或自然的语音产生动画的虚拟人物表示的多媒体内容的系统组成的。这样的系统的例子是上面提到的 JoeXpress[®]系统。

这样的系统使得用户能够选择虚拟人物, 它的背景, 任何个性,
20 要产生的内容的格式。选择的参数被使用来用想要的情景和格式产生动画。

图 2 的流程图显示处理步骤, 通过这些步骤, 按照本发明的系统被 在图 1 上显示为 18 的、起到“发送者”作用的用户访问。用户 18 具有终端, 该终端能够把 SMS 消息发送到能够处理这种类型的消息的
25 相应的中心, 诸如在图 1 上显示为 17 的中心。

从被表示为 200 的起始步骤开始, 标注数字 202 表示其中用户 18 在他/她的终端上编写 SMS 消息(其特征在此后会更好地说明)的步骤: 把它发送到与业务有关的电话号码, 该电话号码在给所述 SMS 消息提供 MMS 特征后转发它。

所讨论的业务主要由被表示为 16 的模块实施，但某些功能可以由模块 10 执行，以及可能由模块 17 执行。

在图 2 上被表示为 204 的步骤中，业务管理功能--因此基本上是模块 16--生成对发送相应于接收的 SMS 消息的 MMS 消息的请求。正如此后更好地解释的，除了消息本身以外，这样的请求还包含用户的识别号以及(可能地)有关接收终端的类型的信息。

在被表示为 206 的步骤中，模块 16 处理接收的请求，生成适合于接收终端的图形和处理能力特征的 MMS 消息。在被表示为 208 的步骤中，所述 MMS 消息被发送到相应的 MMS 中心(诸如模块 10)，相应的 MMS 中心在以后的步骤 208 中，把消息转发到接收终端，诸如终端 12，13 或 14。

步骤 210 表示其中把所述消息按照 MMS 的显示的典型的模式显示给接收终端的步骤。一旦通过读出 MMS 消息而完成传输，系统就移到被表示为 212 的结束步骤。

将要由用户在步骤 202 拨打的与业务有关的电话号码优选地是通常称为“大帐户”的那种专用电话号码。

由用户发送的字符序列，除了消息的文本以外，还包含在报头中的某些信息，诸如 MMS 消息的接收者的电话号码(图 1 上的用户 12、13，14)、将重现消息的虚拟人物和它被插入到的背景。

最后两个信息项是任选的，所以可被忽略。在忽略的情形下，相应的信息由模块 16 自动选择，例如作为随机选择或作为预定的选择(缺省)。自然地，即使仅仅对于所述信息的一部分，这也是可应用的：例如，如果只规定人物，则模块 16 自动选择背景。

被发送到业务的字符序列因此通常具有以下形式：

**<recipient telephone number>[<virtual character(<background>)]
<text message>**

在步骤 202，消息的报头可以人工地或借助于处在终端 18 中的脚本被编写，这允许借助于菜单选择虚拟人物和背景以及从地址簿中选择接收者。

如果消息被人工地拨号，则字符序列会包含错误。例如，用户会规定不存在的虚拟人物或背景的名称。在这种情形下，业务通过自动选择正确的而代替错误信息。

将会看到，所述脚本信息实际上相应于在用于发送 SMS 消息的某些移动电话终端中提供的功能，其有可能由同一业务管理系统远程地把相关的软件加载到各个终端 18(具体地，加载到终端的用户身份模块或 SIM)。

优选地基于以上已经几次提到的 JoeXpress®系统的用于把 SMS 文本格式变换成 MMS 多媒体格式的模块，优选地在被称为“文本动画”的模式中被使用。

在这种情形下，SMS 消息的文本由语音合成器进行处理，语音合成器把文本变换成语音和提供定时的语音序列，然后该语音序列被使用于自动生成选择的虚拟人物的发声活动。

作为到 SMS/MMS 变换模块的输入提供的文本可包含对于最终得到的动画有影响的元信息，这些元信息把表情和手势加到虚拟人物上以及改变合成的语音。

所述元信息作为例如可以具有以下形式的字符的序列被插入到文本：

`<tag><action_type>[<par1>][<par2>]...[<parn>]`

其中：

`<tag>`是对于区分元信息和要被合成的文本所必需的。

`<action_type>`规定要执行哪个动作。动作的例子是：语音音质的改变、面部表情或身体运动的重现、观察点的改变等等。

`<par1-n>`是修正动作的参数，例如，面部表情的持续时间的改变。

较高级别上的另一个表示由所谓的“情绪标记”组成，即，由在文本通信中在互联网中通用的代表情绪状态的字符序列组成。情绪标记的例子是：“;-)”、“":-)”、“":-o”，等等。

情绪标记由系统通过使用上述的表示法被变换成语义上等价的形式。对于情绪标记的支持由于它们被用户所熟悉的和能简单地插入

到文本中而同时具有与低级别表示相同的灵活性的事实而被推动。

像 JoeXpress[®] 系统那样的系统产生三维模型的动画，它们可以由系统转换成不同的格式，这些格式可以取决于三维信息是否被保持而被分类成两类。

- 5 属于第一类的有，例如，MPEG-4 面部和身体动画参数序列、VRML 动画(虚拟现实模型语言的缩写)、3D 演播室最大动画等等。

属于第二类的有，视频编码格式，像 MPEG-1、MPEG-2、MPEG-4 视频、动画的 GIF(虽然它在严格的意义上不是视频编码格式，但 GIF-89a 格式确实允许创建图像序列)。

- 10 动画的音频可以连同视频一起或分开地被编码，正如在 VRML 或动画 GIF 的情形下那样。

由于传输网络的终端的限制，多媒体内容受到约束，诸如消息的最大尺寸、空间分辨率、时间分辨率、和动画的编码的类型。

- 15 为此，除了消息的文本和发送者的识别号以外，必须考虑多媒体消息要被传送到终端的类型。

终端类型实际上标识终端的类别(本质上，诸如存储容量、显示尺寸等的特征)和由于传输网络引起的任何其它约束。

将要在按照本发明的系统中产生的 MMS 消息因此被调节为在施加的约束内最有效地利用可用的资源。

- 20 这个要求可以以至少两个不同的方式被满足。

- 第一种方式规定在步骤 204 生成的创建 MMS 消息的请求除了消息的文本和发送者的识别号以外，还包含表示要被生成的消息必须属于的类别的信息，即 MMS 消息要到的终端的类型以及进而它的性能特征。将要综合 SMS 文本消息的视频内容然后按照接收终端类型被生成，即，被这样地生成以使得 MMS 消息(从通过综合所述视频内容与 SMS 消息而得到的多媒体消息中得出)直接与将要接收多媒体消息的 MMS 终端的特征兼容。
- 25

当这个解决方案被采用时，模块 16 能够根据接收者的识别号搜索被存储在数据库 11 中的终端类型信息。在模块 16 与数据库 11 之间

的连接,按照图 1 所涉及的准则,可以通过模块 10 的直接或间接类型。

得到相同的结果的第二种方式规定多媒体视频内容(将要被加到 SMS 消息)由模块 16 根据作为标准的准则被生成,从而与消息将要被发送到的终端的类型无关。

从由 SMS 文本消息与所述标准多媒体视频内容之间的综合得出的多媒体消息被模块 16 转发到模块 10,该模块 10 从数据库 11 读出有关接收终端的信息后,将从多媒体消息得出的 MMS 消息”专门化”,使它适合于接收终端的特征。

10 采用一个还是另一个解决方案的选择主要是由应用考虑决定的。

第一个解决方案,至少在原理上,具有不需要生成在消息适合于接收终端的要求时将要被丢弃的信息的优点。然而,这个优点被保证模块 16 能够接收位于数据库 11 中的有关终端的类型的信息的需要抵消。

15 第二解决方案具有这样的优点:它利用通常已经为当前的 MMS 应用提供的在模块 10 的级别上的数据库 11 的信息的可用性。在当前的 MMS 应用中,模块 10 已经能够达到按照接收终端的特征转发的 MMS 消息的专门化。然而,上述的优点至少部分地被以下事实减弱:这个解决方案需要由模块 16 生成将要被丢弃的信息。

20 无论采用哪个解决方案,有可能从以下事实获得好处:相同的动画可以在 MMS 消息中以基本上不同的方式表示。

例如,如上所述,可以利用每秒小数目的帧的动画 GIF 图像,在这种情形下,每个帧显示由人物在该时刻发音的消息的文本。这个特别紧凑的表示非常适用于其中消息尺寸约束特别严格或当接收终端不能显示视频时的情形。

25 替换地,可以通过压缩音频利用动画的 GIF 图像。在这种情形下,合成的语音,连同场情景音频一起,也被包括在消息中。当消息的尺寸足够大能包含运动图像和音频轨道时,这是对于不支持视频但能够处理音频的终端的有用的表示。

另外的替换例通过连同音频一起的视频剪辑表示。在这种情形下，得到的动画由于通过视频编码提供的比动画 GIF 图像更高的压缩比和动画中经常使用的更高数目的帧而在其运动方面更具有流动性。这种解决方案可以用于能够支持视频编码的终端。

- 5 应当强调，以上所述的把消息打包的方法仅仅是例子，它们完全不是按照本发明的解决方案提供的全部可能性。

现在参照图 3A 和 3B 描述图 1 上表示为 16 的模块的可能的体系结构安排。

- 10 方框或模块 300 旨在于接收实际上由终端 18 发送的 SMS 消息作为其输入，以及对其执行从报头提取信息的操作。

- 如前所述，文本的第一部分由报头表示，该报头包含接收终端(例如，参照图 1，终端 12、终端 13、或终端 14)的数目以及任选地包含发送者用户想要使用来生成视频内容的人物和背景的指示。这些数据通过分隔符字符从实际的消息中分隔出。消息可包含影响最终得到的
- 15 动画的、低或高级别元信息(例如，所谓的情绪标记)。

 作为这样的文本的例子，可以考虑以下的字符串：

“3356121180 Morpheus Country@H1! I’m at the beach: -) but I’m getting bored without you, \kyawn.150.”

 在该例子中，所使用的分隔符是字符@。

- 20 与所讨论的消息有关的是发送者的识别号以及，可能地，表示接收终端模型的字符串。

- 标注数字 302 表示模块 16 的数据库，它在基于 JoeXpress®系统的优选实施方案中包含诸如可被使用来生成视频内容的字符表、与它们有关的语言、可得到的场景等等的信息。数据库 302 也包含人物和
- 25 背景的三维模型。

 方框 300 与数据库 302 共同运行，从消息中提取报头信息，诸如接收者的识别号，以及被使用来创建视频内容的人物和背景。

 方框 300 然后与包含字符表、语音、可得到的背景的数据库 302 通信，以及如果这些信息在接收的 SMS 消息的报头中是被省略或是错

误的, 则方框 300 自动选择正确的任选项。

方框 300 在它的输出端处输出以下数据/信息:

--将要发送到其功能此后将更容易明白的附加方框 302 的不带有报头的消息文本(**"H1! I'm at the beach: -) but I'm getting bored without you, \kyawn.150."**);

-- 人物 P--动画的主角--的名字(在这里示出的例子中, 所述名字是**"Morpheus"**),

-- 与人物有关的语言 L(例如英语),

-- 相应于其中放置虚拟人物 P 的场景的背景 A(在这里考虑的例子中, 背景是**"乡村"**背景), 以及

-- 接收者 D 的识别号(在所示的例子中, 由号码 **3356121180** 构成)。

从方框 300 接收消息 M 的文本开始, 方框 320 把情绪标记变换成能够被信息系统使用的元信息, 信息系统同时确定哪个文本将被插入到组成构成模块 16 的输出的 MMS 消息的动画的帧中。

所以, 方框 302 的输出由具有低级别信息的文本 TBS, 即其中情绪标记用低级别元信息代替的文本(**"H1! I'm at the beach\ksmile but I'm getting bored without you, \kyawn.150."**), 和其中所有低级别元信息被消除而只保持人物所说的加上情绪标记的文本 TE(**"H1! I'm at the beach: -) but I'm getting bored without you."**)组成。

由方框 302 生成的文本 TBS 被发送到方框 304, 方框 304 旨在提取在文本中包含的动作表以及准备具有由语音合成器 306 使用的形式的文本, 以便还得到与前述的动作有关的定时。

方框 304 把其中低级别元信息用语音合成器的**"标签"**代替(文本到语音)的文本 TAG 发送到合成器 306。所述标签是由合成器标识为专门信息的字符序列, 以及被使用来改变合成的语音或从合成器 306 得到与在合成的句子中的标签有关的时刻。所述时刻被使用来确定动作的时序。

方框 304 也连同任意参数一起生成基本上相应于被包含在文本中

的动作表的信号 TA 作为附加输出。

参照上面几次提到的 SMS 消息，实际上包含有两个动作，即：

- 微笑，和
- 打呵欠，150。

5 参数 150 修正“打呵欠”动作相对于标准持续时间的持续时间。

语音合成器 306 通过使用由方框 300 生成的信号 L 识别的选择的语言把从方框 304 接收的文本 TAG 变换成语音信号。

除了语音信号以外，方框 306 还产生定时的语音序列 FT，被用作构成讲话的运动的的基础。应当想到，定时的语音序列是构成讲出的句子的音素的序列，并被与音素被讲出的时刻综合在一起。

代替地，被表示为 V 的信号是实际合成的语音信号。

用标注数字 308 和 310 表示的方框是监管用于视频内容的人物的讲话的动画和相应的面部和身体动画的机器。

方框 308 接收语音序列 FT 作为输入，把它变换成“面部表情标记 (visemic)”序列，即，变换成在讲话时由面部产生的运动。为了得到现实的运动，动画机考虑相邻的音素的互相影响结果，所述共同清晰度现象。所产生的运动是三维的，以及相关的输出信号 AP 由以三维方式和与人物无关地描述讲话的运动的动画参数构成。这意味着，这样的参数可接连地应用于具有任何形状和复杂性的人物，人等。

20 用作为面部和身体动画机的方框 310，根据相应于由被集成在虚拟相加节点 312 中的方框 304 生成的信号 TA 的动作表用由合成器 306 生成的关于动作的时序的信息运行。

方框 310 以与包含与人物无关的动画参数的形式的面部和身体运动的序列的附加数据库 314 合作的方式运行，该动画参数因此在这方面类似于由方框 308 输出的参数。在本例中，序列“微笑”和“打呵欠”是从数据库 314 取出的两个运动。

面部和身体 310 动画方框联合相应于人物必须执行的各种运动的各个动作，创建单个动画参数序列。各个运动根据与其有关的任何参数进行改变。运动也经受强度、持续时间、镜面反射特征等方面的自

动变化，以增强多样性。最后，也加上由人物执行的、但没有明显地表示的某些运动，诸如眨眼。

方框310的输出由代表以三维方式与人物无关地描述讲话的运动的动画参数的信号AFC组成。所述参数因此可接连地应用于具有任何形状和复杂性的人物，人等，诸如动物。

被表示为316的接连的方框具有把讲话的运动(信号AP)与其它运动(信号AFC)混合，以得到现实的结果的任务。方框316的运行是基于考虑到正在进行对照的运动的优先级的逻辑，诸如讲出的爆破音素(诸如字母“p”)和打呵欠。最终得到的运动是三维的。

方框316的输出信号由代表与人物无关的动画的信号AIP组成。

信号AIP被馈送到方框318，方框318把独立的动画(信号AIP)变换成根据从方框300提取的信号P选择的人物的运动。最终得到的运动取决于模型的拓扑。正如以前看到的，与人物有关的模型被包含在数据库302中。

方框318的输出信号由识别所选择的人物的运动序列的信号ADP组成。

所讨论的信号ADP被馈送到方框320，方框320把具有来自方框300的背景信息A的信号ADP与关于直接从数据库302取出的人物和背景的附加信息合并。

所有这些是为了也将可以在场景中显示(信号A)并可以借助于文本中的元信息驱动的其余的动画，作为物体的运动或拍摄镜头的观察点的改变加到人物的动画中。

方框320的输出信号由最后的三维动画信号TRD组成，该三维动画信号TRD将要发送到任务为提供操作，即，任务为在屏幕上把由人物和背景组成的三维场景表示为像素矩阵的操作的方框322。以规则的时间间隔得到的所述像素矩阵的序列组成所述方框的输出。提供方框322的输出由被表示为FV的动画的视频帧序列组成。视频帧的采样速率是一个参数，在优选方式下它典型地被设置为25Hz。

信号FV被馈送到附加的方框324作为输入，附加的方框也将要

接收具有由方框 302 生成的情绪标记 TE 的文本。

方框 324 在构成产生的视频动画的各个帧之间分发文本。所述操作是任选的以及在生成不带有音频的 MMS 消息时，即，其中 SMS 消息以文本和动画的形式被显示的 MMS 消息时被执行。

- 5 方框 324 的输出由人物和场景的所有运动的组组成。所述信号 FVT，实际上相应于具有文本的视频帧序列，被馈送到视频编码方框 326，该视频编码方框 326 将要除了信号 FVT 以外，还接收有关合成的语音的信号 V 以及有关接收终端的类型的信息 TV 作为它的输入。

- 10 图 3A 和 3B 上显示的实施例涉及其中所述信息在模块 16 的级别上可得到的解决方案。所述信息总地表示接收终端的牌子和模型名称(例如，Sony Ericsson T68i，Nokia 7650，等等)。

方框 326 在这种情形下通过以适合于从所讨论的接收终端观看的格式直接创建视频剪辑而进行。视频剪辑对于确定的终端类型的适配，例如影响帧的空间与时间分辨率，影响音频信道是否被插入等等。

- 15 这里参考的解决方案因此规定把 SMS 消息与这样生成的视频内容进行综合，以使得由模块 16 生成的、最终得到的多媒体消息具有适合于从所述终端观看的格式。

- 20 然而，如前所述，按照本发明的解决方案也可以在其中模块 16(所以，在这里显示的实施例中的方框 326)不实行这种“专门化”的动作的条件下被实施。

在这种情形下，将要补充进入的 SMS 文本消息的视频剪辑，或一般地说，视频内容，以标准格式被生成，即，不考虑接收终端的特征。

- 25 将要使得最后的 MMS 消息实际上是由接收终端可看见的相关的格式变换然后留给具有 MMS 中继/服务器功能的模块 10(图 1)。

在这里显示的实施例中(它事实上是例子)，来自方框 326 的输出信号然后由基本上类似于以压缩格式的视频剪辑的信号 VC 组成。

所述信号被发送到方框 328，该方框将要用来从在它的输入端处承载的多媒体消息出发，构建相应于 MMS 标准的消息。

为了这样进行，除了由方框 326 输出的信号 VC 以外，方框 328 在它的输入端也接收相应于具有由方框 302 生成的情绪标记的文本的信号 TE，有关来自方框 300 的接收者 D 的信号，以及关于发送者 S 的信息：后一个信息是从图 1 的中心 17 按照已知的准则得出的，这里
5 不需要详细说明。

为了生成将要被发送到模块 10 的 MMS 消息，方框 328 把以前计算的视频动画插入在 MMS 消息中。这优选地通过使用场景说明的 SMIL 语言和加入包括多个部分的单一形式的各个多媒体对象来进行。

10 方框 328 也把有关发送者、接收者和对象的信息插入消息报头中。该对象通过使用组成具有情绪标记的文本的第一人物被自动构成。

优选地，方框 328 也将要与由要被插入到 MMS 消息作为“标记”或广告、或作为能够被用作为用于场景的背景音乐或作为广告的叮当声的一系列图像组成的附加数据库 330 共同运行。

15 自然地，在不改变本发明的原理的情况下，本发明的实施方案和实施例的细节可以相对于这里纯粹作为例子描述和显示的内容充分地改变，而不会从而背离本发明的范围。这对于应用本发明以把除了由 SMS 生成的文本消息，例如，电子邮件消息的形式的文本消息，变换成 MMS 消息文本的可能性，以及对于应用本发明以在不同于 UMTS
20 网络上传输 MMS 消息的可能性来说，尤其如此，但并不具体限于此。

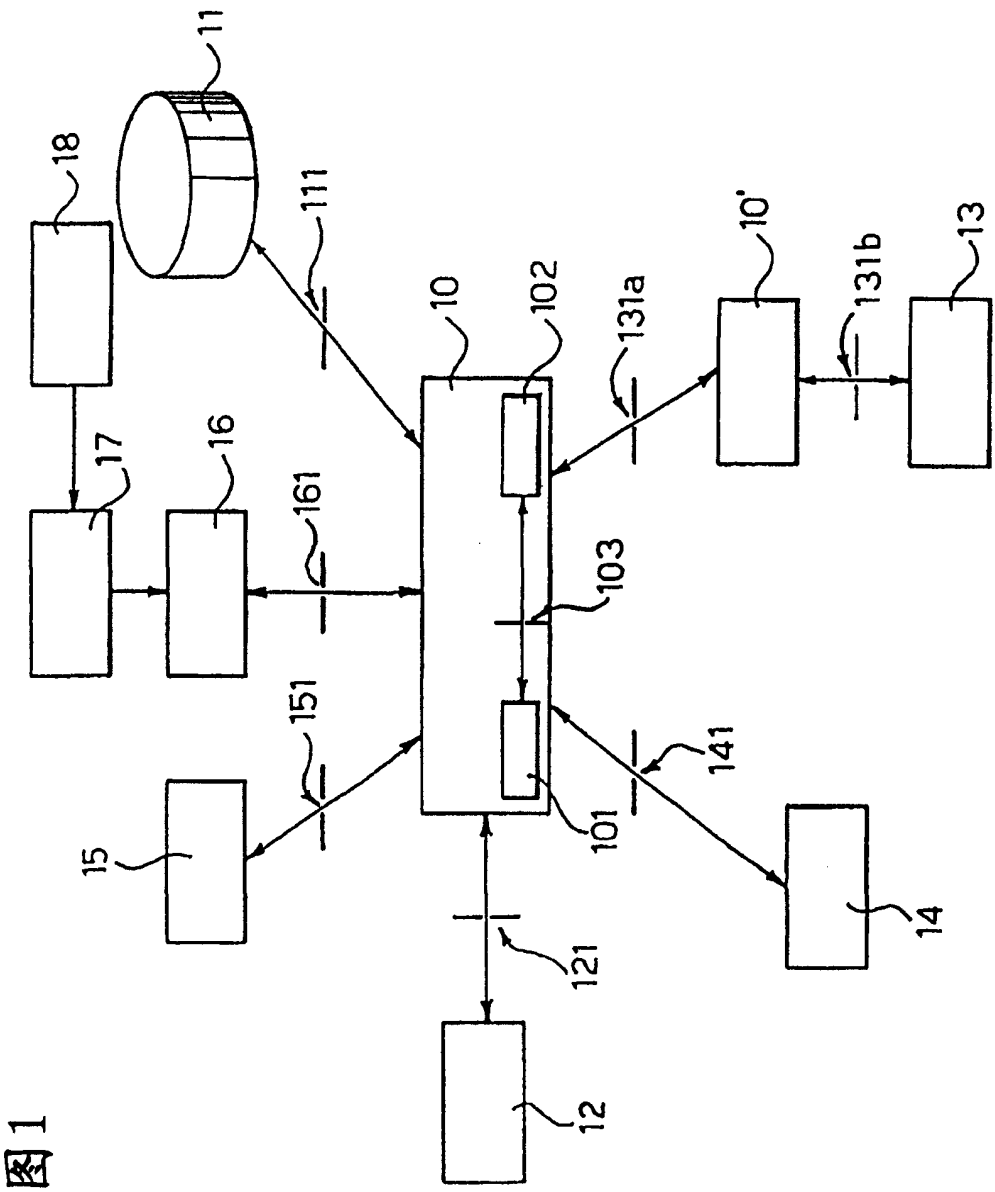


图 2

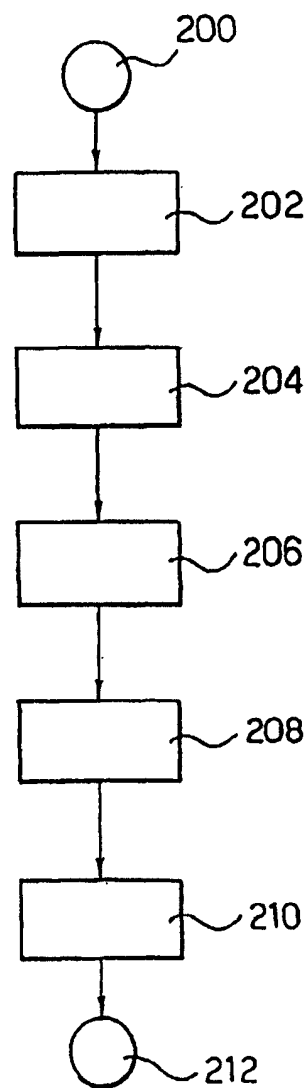


图 3A

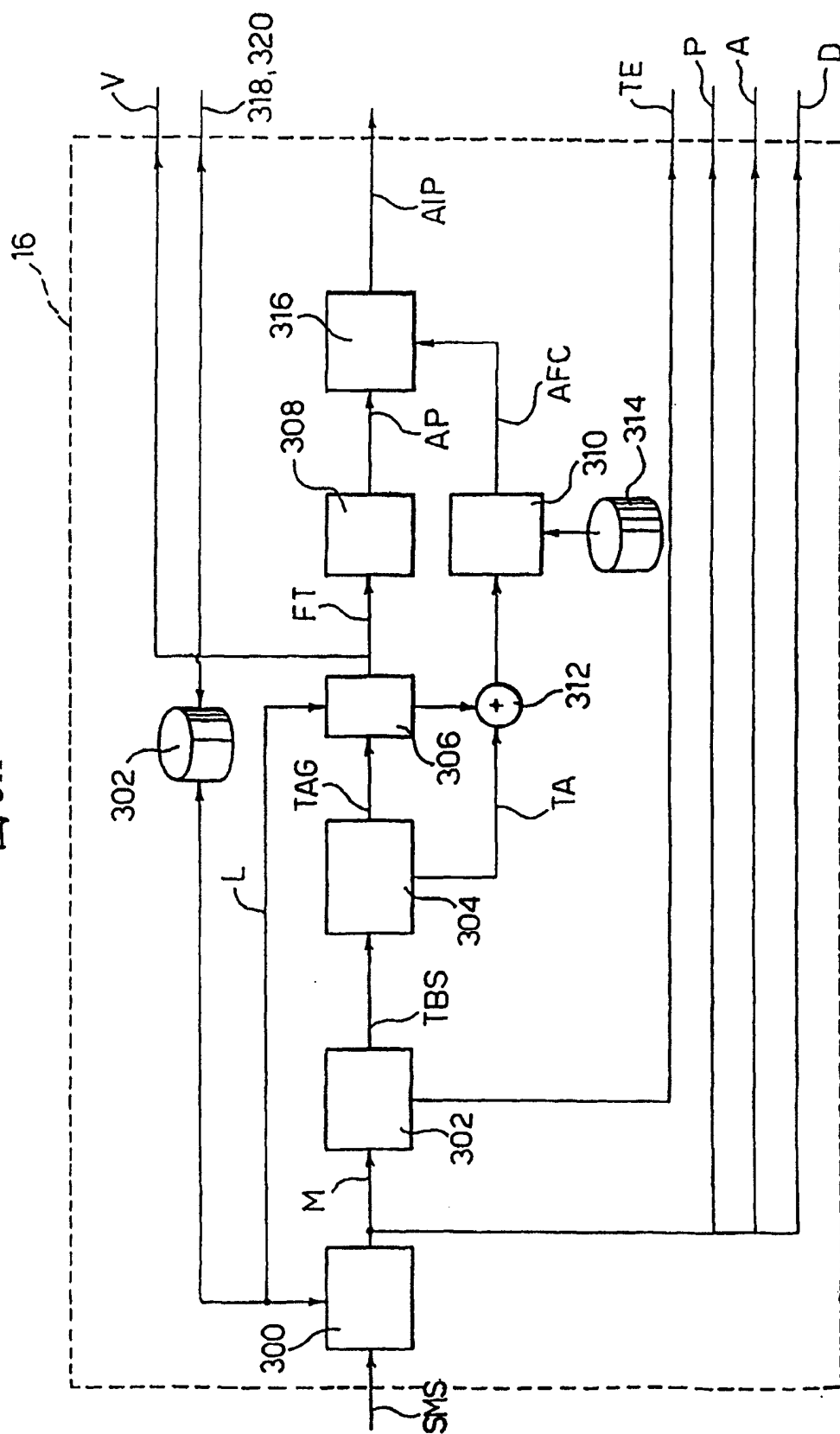


图 3B

