

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04N 7/16 (2006.01)

G06F 17/30 (2006.01)



[12] 发明专利说明书

专利号 ZL 03808260.8

[45] 授权公告日 2008 年 2 月 6 日

[11] 授权公告号 CN 100367794C

[22] 申请日 2003.3.20 [21] 申请号 03808260.8

[30] 优先权

[32] 2002.4.12 [33] JP [31] 110259/2002

[32] 2002.6.19 [33] JP [31] 178169/2002

[86] 国际申请 PCT/JP2003/003450 2003.3.20

[87] 国际公布 WO2003/088665 日 2003.10.23

[85] 进入国家阶段日期 2004.10.12

[73] 专利权人 三菱电机株式会社

地址 日本东京

[72] 发明人 守屋芳美 西川博文 关口俊一

浅井光太郎 山田悦久 乙井研二

黑田慎一 小川文伸

[56] 参考文献

US6370316B1 2002.4.9

JP2001-320693A 2001.11.16

WO01/97044A1 2001.12.20

JP2002-41541A 2002.2.8

JP2001-167099A 2001.6.22

审查员 慈 雪

[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 吴丽丽

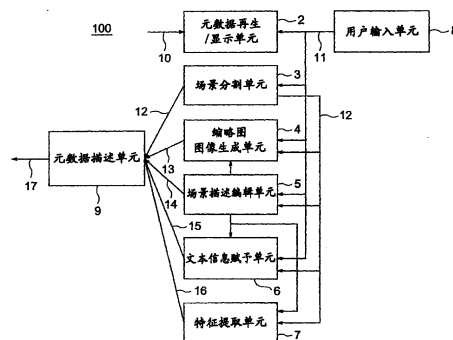
权利要求书 1 页 说明书 29 页 附图 23 页

[54] 发明名称

元数据编辑装置

[57] 摘要

将包含运动图像和音频的多媒体“内容”分割成多个场景，对每个分割的场景生成元数据。除了场景的区间信息和标题外，还能够生成包含描述上述“内容”的分层结构的场景结构信息元数据的元数据。又，为了再生成更适当的元数据，用描述包含在上述元数据中的描述符的种类的元数据最佳化提示信息，进行元数据的再生成。



1.一种元数据编辑装置，其特征在于：它备有，

将包含运动图像、音频中的至少一个的多媒体“内容”分割成多个场景，对每个分割的场景生成表示场景的开始位置和结束位置的场景的区间信息元数据的场景分割单元；

根据来自上述场景分割单元的场景的区间信息元数据，进行上述多媒体“内容”的各场景的分层编辑，生成描述上述多媒体“内容”的分层结构的场景结构信息元数据的场景描述编辑单元；和
统合上述场景的区间信息元数据和上述场景结构信息元数据，生成按照预定格式描述上述多媒体“内容”的内容和结构的元数据的元数据描述单元，

进一步还备有根据来自上述场景分割单元的场景的区间信息元数据，提取上述多媒体“内容”的各场景的视觉特征量，生成特征描述元数据的特征提取单元；

上述元数据描述单元统合上述场景的区间信息元数据、上述场景结构信息元数据和上述特征描述元数据，生成按照预定格式描述上述多媒体“内容”的内容和结构的元数据。

2.根据权利要求1所述的元数据编辑装置，其特征在于：它进一步备有根据帧间差分检测上述多媒体“内容”的场景变化点的场景变换检测单元；

上述场景分割单元根据上述场景变换点将上述多媒体“内容”分割成多个场景。

3.根据权利要求1所述的元数据编辑装置，其特征在于还包括：
场景变化检测单元，基于帧间差分检测多媒体“内容”的每个场景变化点；

其中场景分割单元基于场景变化点将多媒体“内容”分割为多个场景。

元数据编辑装置

技术领域

本发明涉及将包含运动图像和音频的多媒体“内容”分割成多个场景，对每个分割的场景生成元数据的元数据编辑装置、元数据再生装置、元数据分发装置、元数据检索装置、元数据再生成条件设定装置、“内容”分发装置和元数据分发方法。

背景技术

在已有的运动图像管理装置中，在将图像分割成多个场景后，对各场景再生所需的区间信息、场景号码和作为代表场景的图像的集合的索引进行编辑，由生成大于等于一个的索引的部件、将表示检索目的的标题赋予各索引的部件、当检索时，用标题检索索引，按照场景号码的顺序一个接一个地再生索引的场景的部件构成，以并排必要场景的方式编辑索引，能够只再生必要的场景（例如，请参照日本特开2001-028722号专利公报（第1页，图1））。

但是，在上述已有的运动图像管理装置中，因为只不过是由各场景再生所需的区间信息、场景号码和代表场景的图像生成索引，所以存在着不能够管理到视频数据具有的分层性等的视频数据的结构那样的问题。

又，当检索登记的图像时，因为用赋予索引的标题进行，所以存在着为了得到适当的检索结果，必须输入适当的标题那样的问题。

本发明就是为了解决上述问题提出的，本发明的目的是得到除了场景的区间信息和标题外，还能够生成作为视频数据等的“内容”具有的结构等的索引信息的元数据的元数据编辑装置。

又，本发明的目的是得到能够用由元数据编辑装置生成的元数据，只汇集并再生用户想看的场景，并用在元数据中描述的特征量等，检索想看的场景的元数据再生装置、元数据分发装置、元数据检索装置、元数据再生成条件设定装置、元数据分发装置和元数据分发方法。

发明内容

本发明提供一种元数据编辑装置,其特征在于:它备有,将包含运动图像、音频中的至少一个的多媒体“内容”分割成多个场景,对每个分割的场景生成表示场景的开始位置和结束位置的场景的区间信息元数据的场景分割单元;根据来自上述场景分割单元的场景的区间信息元数据,进行上述多媒体“内容”的各场景的分层编辑,生成描述上述多媒体“内容”的分层结构的场景结构信息元数据的场景描述编辑单元;和统合上述场景的区间信息元数据和上述场景结构信息元数据,生成按照预定格式描述上述多媒体“内容”的内容和结构的元数据的元数据描述单元,进一步还备有根据来自上述场景分割单元的场景的区间信息元数据,提取上述多媒体“内容”的各场景的视觉特征量,生成特征描述元数据的特征提取单元;上述元数据描述单元统合上述场景的区间信息元数据、上述场景结构信息元数据和上述特征描述元数据,生成按照预定格式描述上述多媒体“内容”的内容和结构的元数据。

与本发明有关的元数据编辑装置设置了将包含运动图像、音频中的至少一个的多媒体“内容”分割成多个场景,对每个分割的场景生成表示场景的开始位置和结束位置的场景的区间信息元数据的场景分割单元、根据来自上述场景分割单元的场景的区间信息元数据,进行上述多媒体“内容”的各场景的分层编辑,生成描述上述多媒体“内容”的分层结构的场景结构信息元数据的场景描述编辑单元、和统合上述场景的区间信息元数据和上述场景结构信息元数据,生成按照预定格式描述上述多媒体“内容”的内容和结构的元数据的元数据描述单元。

又,与本发明有关的元数据分发装置设置了解析描述元数据中包含的描述符的种类和内容的元数据最佳化提示信息的提示信息解析单元、通过根据上述被分析后的元数据最佳化提示信息和关于元数据再生成的条件,解析描述包含运动图像、音频中的至少一个的多媒体“内容”的内容和结构的元数据,再生成第2元数据的元数据解析/再生成单元、和将由上述元数据解析/再生成单元再生成的第2元数据分发给客户终端的元数据分发单元。

进一步,与本发明有关的元数据分发方法包含着解析描述元数据中包含的描述符的种类的元数据最佳化提示信息的步骤、通过根据上

述被分析后的元数据最佳化提示信息 and 关于元数据再生成的条件, 解析描述包含运动图像、音频中的至少一个的多媒体“内容”的内容和结构的元数据, 再生成第 2 元数据的步骤、和将由上述再生成的第 2 元数据分发给客户终端的步骤。

附图说明

图 1 是表示与本发明的实施例 1 有关的元数据编辑装置的构成的方框图。

图 2 是表示作为与本发明的实施例 1 有关的元数据编辑装置的编辑对象的一个例子的新闻视频的图。

图 3 是表示与本发明的实施例 1 有关的元数据编辑装置的场景分割单元的场景的区间信息元数据的一个例子的图。

图 4 是表示与本发明的实施例 1 有关的元数据编辑装置的场景描述编辑单元的场景结构信息元数据的一个例子的图。

图 5 是表示与本发明的实施例 1 有关的元数据编辑装置的“内容”再生/显示单元、用户输入单元的画面图像例的图。

图 6 是表示与本发明的实施例 2 有关的元数据编辑装置的构成的方框图。

图 7 是用于说明与本发明的实施例 2 有关的元数据编辑装置的工作的图。

图 8 是表示与本发明的实施例 3 有关的元数据再生装置的构成的方框图。

图 9 是用于说明与本发明的实施例 3 有关的元数据再生装置的工作的图。

图 10 是表示与本发明的实施例 4 有关的“内容”分发系统的构成的方框图。

图 11 是表示从与本发明的实施例 4 有关的元数据分发服务器的元数据解析单元输出的“内容”(新闻视频例)的结构信息的图。

图 12 是表示由与本发明的实施例 4 有关的元数据分发系统的元数据再生单元再构成后的“内容”的结构例的图。

图 13 是表示与本发明的实施例 5 有关的元数据分发服务器的构成的方框图。

图 14 是表示用于说明由与本发明的实施例 5 有关的元数据分发服务器产生的元数据最佳化提示信息的视频“内容”的一个例子的图。

图 15 是表示当用由与本发明的实施例 5 有关的元数据分发服务器产生的 MPEG-7 时的元数据的描述例的图。

图 16 是表示与本发明的实施例 5 有关的元数据分发服务器用的元数据最佳化提示信息的格式例的图。

图 17 是表示与本发明的实施例 5 有关的元数据分发服务器用的元数据最佳化提示信息的图。

图 18 是表示与本发明的实施例 5 有关的元数据分发服务器的元数据解析/再生成单元的工作的操作程序图。

图 19 是表示与本发明的实施例 5 有关的元数据分发服务器的元数据解析/再生成单元的工作的操作程序图。

图 20 是表示与本发明的实施例 6 有关的元数据检索服务器的构成的方框图。

图 21 是表示与本发明的实施例 6 有关的元数据检索服务器的元数据解析单元的工作的操作程序图。

图 22 是表示与本发明的实施例 7 有关的客户终端的构成的方框图。

图 23 是表示与本发明的实施例 8 有关的“内容”分发服务器的构成的方框图。

具体实施方式

下面，我们根据附图说明

与本发明的实施例 1 和 2 有关的元数据编辑装置、

与实施例 3 有关的元数据再生装置、

与实施例 4 有关的“内容”分发系统、

与实施例 5 有关的元数据分发服务器、

与实施例 6 有关的元数据检索服务器、

与实施例 7 有关的客户终端、和

与实施例 8 有关的“内容”分发服务器。

实施例 1

在实施例 1 中,我们说明将包含运动图像和音频的多媒体“内容”分割成多个场景,生成场景的分层结构描述和包含各场景的特征量的元数据(索引信息)的元数据编辑装置。

我们一面参照附图一面说明与本发明的实施例 1 有关的元数据编辑装置。图 1 是表示与本发明的实施例 1 有关的元数据编辑装置的构成的方框图。此外,在各图中,相同的标号表示相同或相当的部分。

在图 1 中,元数据编辑装置 100 备有“内容”再生/显示单元 2、场景分割单元 3、缩略图图像生成单元 4、场景描述编辑单元 5、文本信息赋予单元 6、特征提取单元 7、用户输入单元 8 和元数据描述单元 9。

“内容”再生/显示单元 2 再生/显示由视频数据和音频数据等构成的编辑对象的多媒体“内容”10。场景分割单元 3 将“内容”分割成多个场景。缩略图图像生成单元 4 提取场景的代表帧作为缩略图图像。场景描述编辑单元 5 通过由场景分割单元 3 分割的场景的分组化、场景的结合、场景的删除、场景的关系信息的生成等分层地编辑场景。文本信息赋予单元 6 将各种文本信息附加在各场景上。特征提取单元 7 提取场景的特征。

又,用户输入单元 8 将来自用户的指示信息作为用户输入信息 11 输出到“内容”再生/显示单元 2、场景分割单元 3、缩略图图像生成单元 4、场景描述编辑单元 5 和文本信息赋予单元 6。

进一步,元数据描述单元 9 将由场景分割单元 3、缩略图图像生成单元 4、场景描述编辑单元 5、文本信息赋予单元 6 和特征提取单元 7 输出的场景的区间信息元数据 12、场景的缩略图图像信息元数据 13、场景结构信息元数据 14、文本信息元数据 15 和特征描述元数据 16 统合起来,生成按照规定的格式描述多媒体“内容”的内容和结构的元数据 17。

下面,我们一面参照附图一面说明与本实施例 1 有关的元数据编

辑装置的工作。图 2 是表示作为与本实施例 1 有关的元数据编辑装置的编辑对象的一个例子的新闻视频的构成的图。

我们将编辑图 2 所示的构成的新闻视频的情形作为例子进行说明。

首先，元数据编辑装置 100 的“内容”再生/显示单元 2，当通过网络等输入存储在“内容”存储单元（图中未画出）中的视频“内容”等的多媒体“内容”10 时，为了用于编辑地再生/显示它们。

当元数据编辑装置 100 的用户一面观看该再生图像，一面通过用户输入单元 8 输入作为场景的切出位置，即场景开始位置和结束位置时，场景分割单元 3 生成显示从用户输入的场景开始位置和结束位置的场景的区间信息元数据 12。

图 3 是表示与本实施例 1 有关的元数据编辑装置的场景分割单元的场景的区间信息元数据的一个例子的图。

该图 3 所示的区间信息元数据 12 表示从图 2 所示的新闻视频生成的例子。如图 3 所示，由场景分割单元 3，对于从新闻视频“内容”切出的“新闻摘要”、“国内新闻”和“国际新闻”等的各场景，生成表示场景开始位置和结束位置的区间信息的场景的区间信息元数据 12。

在场景描述编辑单元 5 中，当具有通过用户输入单元 8 来自用户的场景编辑的指示时，根据来自场景分割单元 3 的场景的区间信息元数据 12，进行由场景分割单元 3 连续地切出的场景的分层编辑，输出场景结构信息元数据 14。所谓的场景的分层编辑指的是例如场景的分组化、场景的再分割、场景的结合、场景的删除。所谓场景的分组化指的是，例如，从图 2 所示的新闻视频，例如图 4 所示地将以“国内新闻”、“国际新闻”、“经济新闻”等的某种特定特征关联的场景汇集成一个“新闻”组。又，所谓场景的再分割指的是将一个场景分割成多个场景。所谓的场景的结合指的是汇集多个场景成为一个场景。

图 4 是表示与本实施例 1 有关的元数据编辑装置的场景描述编辑单元的场景结构信息元数据的一个例子的图。

该图 4 所示的场景结构信息元数据 14 描述作为在场景描述编辑

单元 5 中的编辑结果，生成的视频“内容”的分层结构。在图 4 中，在场景描述编辑单元 5 中，通过场景的分组化、场景的再分割、场景的结合等的场景的编辑，将称为“新闻”的场景编辑成“新闻摘要”、“新闻”、“特辑”、“体育”等，进一步将该“新闻”分层地编辑成“国内新闻”、“国际新闻”、“经济新闻”显示出来。

而且，将在场景描述编辑单元 5 中生成的例如图 4 所示的元数据 14 输出到元数据描述单元 9。

另一方面，在缩略图图像生成单元 4 中，根据来自场景分割单元 3 的场景的区间信息元数据 12，从在场景分割单元 3 中切出的各场景中生成代表帧作为缩略图图像，将生成的缩略图信息作为缩略图图像信息元数据 13 输出到元数据描述单元 9，登记在元数据描述单元 9 中。这里，用户也能够通过用户输入单元 8，进行缩略图选择，但是也能够自动地将起始帧和固定时间间隔的多个帧作为代表帧，或自动检测场景变换点，将这些帧作为代表帧。缩略图图像信息元数据 13 成为视频“内容”中的缩略图的位置信息（帧号码或时间）、或者，缩略图图像的 URL 等的所处位置信息。

又，在特征提取单元 7 中，根据来自场景分割单元 3 的场景的区间信息元数据 12，从各场景中，提取运动、颜色、或包含在场景内的物体的形状等场景具有的视觉特征量。将提取的特征量作为特征描述元数据 16，输出到元数据描述单元 9，并登记在元数据描述单元 9 中。

又，在文本信息赋予单元 6 中，根据来自场景分割单元 3 的场景的区间信息元数据 12，用户将标题、简介、关键字、评述、场景的重要度等的各种文本信息赋予各场景。文本信息的赋予是当用户通过用户输入单元 8 输入时，通过解析“内容”中包含的音频信息和字幕，自动地赋予的情形。将文本信息，作为文本信息元数据 15，输出到元数据描述单元 9，并登记在元数据描述单元 9 中。

图 5 表示与本实施例 1 有关的元数据编辑装置的“内容”再生/显示单元和用户输入单元 8 的画面图像例。在图 5 中，视频再生画面 G1，与在“内容”再生/显示单元 2 中的画面图像例相当，在该视频再

生画面 G1 中，再生/显示编辑用的“内容”。在图 5 中，虽然没有明示，但是备有“再生”、“停止”、“回卷”，“快进”、“发送帧”再生用的指示按钮等通常的视频再生装置中备有的用户接口。而且，在该视频再生画面 G1 的下方显示出场景分割指示画面 G2。该场景分割指示画面 G2，例如，具有幻灯片形式，一面观看在视频再生画面 G1 中显示的图像，一面用户能够指示在视频再生画面 G1 中显示的图像场景的开始位置和结束位置。又，场景分割指示画面 G2 能够同时在场景的开始位置和结束位置之间，指示缩略图的位置。这里，当由场景分割指示画面 G2 指定缩略图的位置时，缩略图图像生成单元 4 从视频“内容”的指定位置的帧生成缩略图图像。

又，由场景分割指示画面 G2 指定位置的缩略图图像，作为场景的分割信息，在场景分割信息显示画面 G3 中显示出来。在该场景分割信息显示画面 G3 中，除了缩略图图像外，如图 3 所示，也能够对每个场景显示表示场景的开始位置和结束位置的信息。

接着，在树形结构生成指示/显示画面 G4 中，向用户指示场景编辑。一面用户观看在场景分割信息显示画面 G3 中显示出来的缩略图图像等的场景分割信息，一面生成表示视频“内容”具有的分层结构的树形。

作为操作方法，例如，当进行场景的分组化时，在树形上追加新的节点，将想要分组化的场景追加在该节点上。场景的追加操作，可以考虑在场景分割信息显示画面 G3 上选择想要追加的场景，通过拖拽，在节点上追加场景等的方法。通过在场景分割信息显示画面 G3、树形结构生成指示/显示画面 G4 中提供用户输入单元 8 作为用于选择场景、经过文本信息赋予单元 6 将文本信息赋予该场景的用户接口，能够输入对于场景的文本信息。

元数据描述单元 9 将由场景分割单元 3、缩略图图像生成单元 4、场景描述编辑单元 5、文本信息赋予单元 6 和特征提取单元 7 输出的各种元数据统合起来，生成按照规定的描述格式描述的元数据文件。元数据的规定的描述格式也能够用独自规定的格式进行描述，但是在

本实施例 1 中, 使用由 ISO 标准化的 MPEG-7。该 MPEG-7 规定描述“内容”的结构和特征的格式, 具有 XML 文件格式和二进制格式。

从而, 如果根据本实施例 1 的元数据编辑装置 100, 则因为设置了分层编辑场景的场景描述编辑单元 5 和从场景提取特征的特征提取单元 7, 所以能够生成描述视频数据等的“内容”具有的分层结构和各场景的特征量的元数据。

此外, 假定输入到“内容”再生/显示单元 2 的多媒体“内容”10, 有从在网络上的“内容”服务器(图中未画出)取得的情形、从在元数据编辑装置 100 内的“内容”存储单元(图中未画出)取得的情形和从 CD 和 DVD 等的存储媒体(图中未画出)取得的情形等的各种情形。同样, 假定将从元数据描述单元 9 输出的元数据, 有存储在网络上的“内容”服务器(图中未画出)中的情形、存储在元数据编辑装置内的元数据存储单元(图中未画出)中的情形和与“内容”一起存储在 CD 和 DVD 等的存储媒体(图中未画出)中的情形等。

又, 在本实施例 1 中, 设置了场景描述编辑单元 5 和特征提取单元 7 两者并进行了说明, 但是不限于此, 当然既可以只设置场景描述编辑单元 5, 也可以只设置特征提取单元 7。

实施例 2

在上述实施例 1 中, 我们说明全部通过手动进行场景分割的情形, 但是在本实施例 2 中, 我们说明以备有自动检测场景变换点的场景变换检测单元为特征的数据编辑装置。

我们一面参照附图一面说明与本发明的实施例 2 有关的元数据编辑装置。图 6 是表示与本发明的实施例 2 有关的元数据编辑装置的构成的方框图。

在图 6 中, 元数据编辑装置 100A 备有“内容”再生/显示单元 2、场景分割单元 3、缩略图图像生成单元 4、场景描述编辑单元 5、文本信息赋予单元 6、特征提取单元 7、用户输入单元 8、元数据描述单元 9 和场景变换检测单元 39。此外, 标号 40 是自动检测的场景开始位

置信息。

下面，我们一面参照附图一面说明与本实施例2有关的元数据编辑装置的工作。

图7是用于说明与本发明的实施例2有关的元数据编辑装置的工作的图。

除了场景变换检测单元39和场景分割单元3以外的工作都与上述实施例1相同。这里我们说明实施例2特有的工作。

场景变换检测单元39自动地进行场景变换、切去点检测。场景变换检测是，例如，根据帧间象素差分、帧间颜色和亮度的直方图差分等进行的。场景分割单元3根据在场景变换检测单元39中检测出的场景变换点，决定场景开始位置和结束位置。

这里，我们将编辑对象的“内容”是新闻视频的情形作为例子详细述说场景变换检测单元39和场景分割单元3的处理。

我们以用颜色直方图作为用于场景变换检测的特征量的情形为例进行说明。

在场景变换检测检测单元39中，对每一个帧算出颜色直方图。作为表色系，具有HSV、RGB、YCbCr等，但是这里，用HSV色空间。该HSV色空间由色调(H)、色饱和度(S)、亮度(V)这样3个要素构成。算出各要素的直方图。接着，从求得的直方图，例如根据下列的(公式1)，算出帧间的直方图差分。我们假定从场景的开始帧算起的N个帧(例如N=3)属于同一场景，即，没有场景变换点。此外，作为场景的初期特征量，根据下列的(公式2)，求得最初的N个帧间的直方图差分的平均(mean)和标准偏差(sd)。

$$sum_i = \sum_{k=1}^{bin_H} |H_i(k) - H_{i-1}(k)| + \sum_{k=1}^{bin_S} |S_i(k) - S_{i-1}(k)| + \sum_{k=1}^{bin_V} |V_i(k) - V_{i-1}(k)|$$

(公式1)

其中，

sum_i ：帧i和帧i-1的直方图间差分和

$H_i(h)$ ：色调的直方图， bin_H ：直方图的要素数

$S_i(h)$: 色饱和度的直方图, bin_S : 直方图的要素数

$V_i(h)$: 亮度的直方图, bin_V : 直方图的要素数

$$\text{mean} = \frac{1}{N-1} \sum_{i=1}^{N-1} \text{sum}_i, \text{sd} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (\text{sum}_i - \text{mean})^2}$$

(公式 2)

其中,

mean : 帧间的直方图差分的平均

sd : 帧间的直方图差分的标准偏差

而且, $N+1$ 帧及其以后, 将直方图的帧间差分比 $\text{mean} + \lambda \cdot \text{sd}$ 大的帧作为场景变换点, 作为新场景的开始位置候补。

现在考虑如果得到多个场景的开始位置候补, 则接着, 如新闻视频那样, 插入由新闻的切换等决定的模式的图像的情形。

新闻视频, 在新闻的切换中, 例如, 插入由播音员和背景的演播室设置、说明的文字(字幕)构成的图像等、决定的模式的图像的情形是很多的。从而, 预先登记这些决定的模式的图像(称为样板图像)、或描述样板图像的特征量的元数据。所谓样板图像的特征量可以举出样板图像的颜色直方图、或者, 运动图案(在新闻切换中播音员写的部分具有少量的运动等)等。

在预先登记样板图像的情形中, 例如图 7 所示, 当与场景变换点对应的图像与样板图像匹配, 类似度高时, 将该场景变换点作为场景的开始位置进行登记。作为类似度的匹配, 具有帧间差分 and 帧间的颜色直方图差分等。

又, 在预先登记样板图像的特征量的情形中, 当从与场景变换点对应的图像提取特征量, 与样板图像的特征量匹配, 类似度高时, 将该场景变换点作为场景的开始位置进行登记。将场景的开始位置的信息输出到场景分割单元 3。

在场景分割单元 3 中, 根据由场景变换检测单元 39 自动检测出

的场景开始位置信息，决定场景开始位置和结束位置。此外，在本实施例 2 的场景分割单元 3 中，与上述实施例 1 同样，也能够根据来自用户的指示，决定场景开始位置和结束位置。

在场景分割单元 3 中，也能够将描述场景的开始位置和结束位置的场景的区间信息元数据 12 输出到场景变换检测单元 39，用该场景变换检测单元 39，检测在该场景中包含的场景变换点。

在场景描述编辑单元 5 中，能够根据来自场景分割单元 3 的场景的区间信息元数据 12，再分割和统合由场景变换检测单元 39 自动检测出的场景。此外，场景描述编辑单元 5 的详细情况与上述实施例 1 相同。

从而，如果根据与本实施例 2 有关的元数据编辑装置 100，则与上述实施例 1 相同，能够生成描述视频数据等的“内容”具有的分层结构和各场景的特征量的元数据，并且通过设置场景变换检测单元 39，可以自动检测“内容”的场景变换点。

实施例 3

在本实施例中，我们说明利用由上述实施例 1 和 2 的元数据编辑装置生成的元数据，进行图像的概要再生和检索等的元数据再生装置。

我们一面参照附图一面说明与本发明的实施例 3 有关的元数据再生装置。图 8 是表示与本发明的实施例 3 有关的元数据再生装置的构成的方框图。

在图 8 中，元数据编辑装置 200 备有元数据解析单元 19、结构显示单元 20、缩略图图像显示单元 21、用户输入单元 22、检索单元 23、检索结果显示单元 24、概要生成单元 25、概要结构显示单元 26 和“内容”再生单元 27。

元数据解析单元 19 对描述了关于“内容”具有的分层的场景结构和各场景的缩略图的信息、各场景的特征量等的元数据 28 的解析。结构显示单元 20 显示从元数据解析结果取得的场景结构 29，即“内

容”的分层结构。缩略图图像显示单元 21 显示从元数据解析结果取得的缩略图图像信息 30。

用户输入单元 22 进行检索和再生等的指示。检索单元 23 根据来自用户的检索指示(检索条件 31)和从元数据取得的场景的特征量与文本信息 32 进行检索。检索结果显示单元 24 显示检索结果 33。概要生成单元 25 根据来自用户的概要生成指示(概要生成条件 34)生成概要。概要结构显示单元 26 显示生成概要的“内容”结构 38。“内容”再生单元 27 根据概要信息 35、“内容”再生指示 36、再生的“内容”37 再生/显示“内容”。

下面,我们一面参照附图一面说明与本实施例 3 有关的元数据再生装置的工作。

首先,元数据解析单元 19 将描述关于“内容”具有的分层的场景结构和各场景的缩略图的信息、各场景的特征量等的元数据 28 作为输入,进行元数据的解析。

在本实施例 3 中,因为用上述实施例 1、2 的元数据描述单元 9 生成的由 MPEG-7 规定的格式描述该元数据 28,所以元数据既可以用 XML 描述的文本文件,也可以是用二进制格式进行编码的二进制文件。

因此,元数据解析单元 19,如果用 XML 描述元数据 28,则具有进行 XML 文件解析的 XML 服务器的功能。又,如果用二进制格式对元数据 28 进行编码,则具有进行元数据 28 解码的解码器功能。

结构显示单元 20,输入元数据解析单元 19 的解析结果,显示“内容”的分层的场景结构 29。例如,如图 4 所示,与各场景的标题一起树形显示“内容”的场景结构。

缩略图图像显示单元 21 输入元数据解析单元 19 的解析结果(缩略图图像信息 30),显示“内容”的缩略图图像一览表。

检索单元 23,根据通过用户输入单元 22 的来自用户的检索指示,检索“内容”中包含的场景。这时,由用户输入单元 22,根据关键字和取样图像的提示等,输入检索条件。由检索单元 23,根据元数据中

描述的場景的特征量和場景标题等的文本信息 32, 检索与由用户提示的检索条件 (关键字和取样图像的特征) 31 一致的場景。

当由检索单元 23 进行的检索结束时, 检索结果显示单元 24 输入检索单元 23 的检索结果 33, 显示出检索结果。作为检索结果的显示方法, 例如, 显示与检索条件一致的場景的缩略图图像。

又, 概要生成单元 25 根据通过用户输入单元 22 输入的来自用户的概要生成指示, 生成“内容”的概要。这时, 由用户输入单元 22, 输入生成概要的“内容”的再生时间和用户嗜好等的信息。例如, 当“内容”是新闻视频时, 输入将新闻中的体育作为主要的用户想看的消息、或者, 将 1 小时新闻概括成 20 分钟的用户想看的消息等的用户嗜好信息。又, 由概要生成单元 25, 根据元数据中描述的場景的再生时间和場景的标题等的文本信息 32, 生成与概要条件一致的概要信息 35。该概要信息 35 是, 例如, 在生成概要的“内容”中包含的場景的再生清单, 是记述了描述“内容”的 URL 等的位置信息以及该“内容”中的想要再生的場景的开始位置和结束位置的清单。

又, 由“内容”再生/显示单元 27, 根据概要信息 35 中包含的“内容”的位置信息, 特定对象“内容”, 根据概要信息 35 中包含的場景清单, 进行再生場景的取得/再生/显示。又, 作为别的例子, 也可以考虑分层地描述概括概要信息的場景结构的情形。

图 9 是表示分层地描述生成概要的場景结构的一个例子的图。图 9(a) 表示原始“内容”的場景结构的例子。对于各場景, 附加在 0.0~1.0 范围内的重要度。1.0 意味着重要度最高, 0.0 意味着重要度最低。重要度是, 例如, 根据用户嗜好计算的。例如, 当预先登记关于球队 A 的足球賽的場景, 特别是必然要看的比賽結果和得分場景那样的用户嗜好时, 在各場景上附加反应用户嗜好的重要度。

而且, 在图 9(a) 中, 只用重要度最高的場景生成概要, 生成概要的場景结构如图 9(b) 所示。此外, 各場景具有包含該場景的“内容”的 URL 等的位置信息和該場景的“内容”内的位置信息 (开始位置和结束位置) 等的元数据。将关于生成概要的場景结构 38 的信息

传送给概要结构显示单元 26, 该概要结构显示单元 26 例如以图 9(b) 所示的树形形式显示生成概要的场景结构。

又, 当用户通过用户输入单元 22 从在结构显示单元 20 或概要结构显示单元 26 中显示的场景结构、以及缩略图图像显示单元 21 和检索结果显示单元 24 中显示的各场景的缩略图中选择想要再生的大于等于一个的场景时, 能够在“内容”再生/显示单元 27 中再生/显示“内容”中包含的场景。

从而, 如果根据与本实施例 3 有关的元数据编辑装置 200, 则能够用由上述实施例 1、2 中说明的元数据编辑装置生成的元数据, 只汇集并再生用户想看的场景, 或者用在元数据中描述特征量, 检索想看的场景。

又, 在本实施例 3 中, “内容”再生/显示单元 27 是在元数据编辑装置 200 内, 但是“内容”再生/显示单元也可以在别的装置内。这可以考虑为, 例如, 关于场景结构与缩略图图像的显示等、元数据再生的操作和显示是在便携式电话和便携式信息终端等中进行的, 关于多媒体“内容”的再生的处理和显示是在通过网络与便携式电话和便携式信息终端等连接的终端(例如 PC 等)中进行的情形。

实施例 4

在本实施例 4 中, 我们说明向客户终端分发“内容”的元数据的元数据分发服务器(元数据分发装置)和与客户终端的终端能力相应地可伸缩地构成并分发该“内容”的“内容”分发服务器。

我们一面参照附图一面说明与本发明的实施例 4 有关的“内容”分发系统。图 10 是表示与本发明的实施例 4 有关的“内容”分发系统的构成的方框图。

在图 10 中, “内容”分发系统 300 备有元数据分发服务器 400、各种客户终端 481~48n 和“内容”分发服务器 500。

元数据分发服务器 400 由元数据存储单元 41、元数据解析单元 42、终端能力判定单元 43、元数据再生单元 44 和元数据分发单元 45

构成。

在元数据存储单元 41 中,存储着例如上述实施例 1、2 的元数据编辑装置生成的元数据。元数据解析单元 42 对描述“内容”的结构和特征的元数据 49 进行解析。终端能力判定单元 43 根据关于客户终端的性能的信息 51,判定客户终端的终端能力。元数据再生单元 44 根据元数据解析结果 50,与客户终端的终端能力相应地再构成“内容”,再生成描述该“内容”的元数据 52。元数据分发单元 45 将在元数据再生单元 44 中再生成的元数据 53 分发给各种客户终端 481~48n。

此外,也可以将元数据存储单元 41 设置在本实施例 4 的元数据分发服务器 400 的外面。这时,元数据分发服务器 400 通过网络(图中未画出)等从元数据存储单元 41 输入元数据 49。

又,“内容”分发服务器 500 由“内容”存储单元 46 和“内容”分发单元 47 构成。

“内容”存储单元 46 存储“内容”55。“内容”分发单元 47 与来自客户终端 481~48n 的“内容”分发要求 54 相应地分发“内容”56。

与上述元数据分发服务器 400 的情形相同,也可以将“内容”存储单元 46 设置在“内容”分发服务器 500 的外面。这时,“内容”分发服务器 500 通过网络(图中未画出)输入“内容”数据 55。

下面,我们一面参照附图一面说明与本实施例 4 有关的“内容”分发系统的工作。

首先,在元数据分发服务器 400 一侧,元数据解析单元 42 对存储在由元数据存储单元 41 中的元数据进行解析。元数据解析单元 42 的工作与上述实施例 3 的“内容”再生装置 200 的元数据解析单元 19 相同。元数据解析单元 42 通过对元数据进行解析,取得关于各“内容”的结构和特征的信息。

图 11 是表示从与本实施例 4 有关的元数据分发服务器的元数据解析单元输出的“内容”(例如新闻视频)的结构信息的图。在图 11 中,用树形表示“内容”的分层的场景结构。树形的各节点,与各个场景对应,将各个场景信息对应地赋予各节点。所谓的场景信息指的

是作为场景的标题、摘要、场景开始位置和结束位置的时间信息、场景的缩略图、代表帧、缩略图镜头、代表镜头、颜色和运动等视觉上的特征量等的场景特征。此外，在图 11 中，在各种场景信息中，只图示场景的标题。

这里，假定客户终端为终端能力不同的各种信息家电设备。所谓的终端能力指的是通信速度、处理速度和可以再生/显示的图像格式、图像分辨率、用户输入功能等。例如，假定客户终端 481 是在通信速度、处理速度、显示性能、用户输入功能上具有充分性能的 PC（个人计算机）。又，假定客户终端 482 是便携式电话，假定其它的客户终端是 PDA 等。从各客户终端 481~48n 发送与各个终端性能有关的信息。

终端能力判定单元 43 解析与从各客户终端 481~48n 发送的终端性能有关的信息 51，决定可以分发的图像格式、最大图像分辨率、“内容”的长度等，输出到“内容”再生单元 44。例如，在原始“内容”是用 MPEG-2 进行编码的图像分辨率大的视频“内容”的情形中，在具有充分性能的客户终端 481 中，能够再生原始“内容”。又，该客户终端 481 具有上述实施例 3 中说明的可以进行图像的概要再生和检索的功能。另一方面，在客户终端 482 中，只可以再生用 MPEG-4 进行编码的短的视频镜头，并且可以显示的最大图像分辨率也很小。

在“内容”再生单元 44 中，与来自终端能力判定单元 43 的各客户终端 481~48n 的终端性能相应地再构成“内容”，再生成描述其结构和内容的元数据 52，输出到元数据分发单元 45。例如，因为将原始的元数据原封不动地分发给客户终端 481，所以不进行“内容”的再构成。另一方面，因为对于客户终端 482，只具有短的视频镜头的再生功能，所以不能够再生全部场景。从而，用重要场景的短的视频镜头再构成“内容”。

图 12 是表示由与本实施例 4 有关的元数据分发系统的元数据再生单元再构成后的“内容”的结构例的图。如图 12 所示，从各新闻场景中，提取重要的场景，只用该场景的代表镜头或代表帧进行构成。

又，因为客户终端 482 不具有上述实施例 3 中说明的检索功能，所以在场景的各种信息中，不需要将用于检索的场景的特征量描述在元数据中。因此，元数据再生单元 44 再生成只描述再构成的场景结构和该场景的代表镜头或代表帧的位置信息的元数据，并输出到元数据分发单元 45。

该元数据分发单元 45 将由元数据再生单元 44 生成的元数据 53 分发给客户终端 481~48n。

各客户终端 481~48n 对由元数据分发单元 45 分发的元数据 53 进行解析，取得“内容”的树形结构信息。各客户终端 481~48n 的用户，当选择想要再生的场景时，从各客户终端 481~48n 将选出的场景的位置信息发送给“内容”分发服务器 500 的“内容”分发单元 47。

由“内容”分发服务器 500 的“内容”分发单元 47，取得从各客户终端 481~48n 发送的场景的位置信息，从“内容”存储单元 46 取得对应的“内容”55，分发给客户终端 481~48n。在客户终端 481 的情形中，发送场景的开始位置和结束位置，分发与原始“内容”的对应的场景。又，在客户终端 482 的情形中，发送场景的代表镜头的所在位置信息（URI 等）。此外，在客户终端 482 中不可能再生/显示代表镜头的图像格式、图像分辨率和图像文件大小等的情形中，由“内容”分发单元 47，进行格式变换和图像分辨率变换、减小文件大小的“内容”概要等，并进行发送。

从而，如果根据本实施例的元数据分发服务器 400，则可以与各客户终端 481~48n 的能力相应地再生成元数据，分发给各客户终端。

此外，在图 10 中，表示分别构成的元数据分发服务器 400 和“内容”分发服务器 500，但是本发明不限于此，也可以将“内容”分发服务器设置在元数据分发服务器中，或者将元数据分发服务器设置在“内容”分发服务器中。又，当然也可以将元数据分发服务器和“内容”分发服务器设置在同一服务器内。如果这样做，则因为“内容”分发单元 47 能够从顾客判定单元 43 简单地知道各客户终端 481~48n 的能力，所以可以与各客户终端 481~48n 的能力相应地再构成格式变换等

的“内容”，分发给各客户终端 481~48n。

又，在该实施例 4 中，我们说明了存储在元数据存储单元 41 中的元数据是例如上述实施例 1、2 的元数据编辑装置生成的元数据的情形，但是不限于此，当然也可以存储上述实施例 1、2 的元数据编辑装置以外的装置生成的元数据。

实施例 5

在本实施例 5 中，我们说明上述实施例 4 中说明的元数据分发服务器的别的例子。由上述实施例 4 的元数据分发服务器，根据从客户终端发送的终端信息，进行元数据的再生。在本实施例 5 中，我们说明以备有为了进行更适当的元数据的再生成，用作为用于元数据的再生成的提示信息的元数据最佳化提示信息，进行元数据的再生成的元数据解析/再生成单元为特征的元数据分发服务器(元数据分发装置)。

我们一面参照附图一面说明与本发明的实施例 5 有关的元数据分发服务器。图 13 是表示与本发明的实施例 5 有关的元数据分发服务器的构成的方框图。

在图 13 中，元数据分发服务器 400A 备有提示信息解析单元 61、元数据解析/再生成单元 63 和元数据分发单元 45。

提示信息解析单元 61 解析元数据最佳化提示信息 60 并输出它的结果。元数据解析/再生成单元 63，根据被解析出的元数据最佳化提示信息 62、关于客户终端的性能的信息、或者，关于用户嗜好等的元数据再生成的条件 65，解析描述“内容”的结构和特征的元数据 49，输出再构成的元数据 64。元数据分发单元 45 将元数据 53 分发给客户终端。

在元数据存储单元 41 (请参照图 10) 中，存储着描述“内容”的结构和特征的元数据 49 和用于再生成该元数据 49 的提示信息的元数据最佳化提示信息 60。所谓用于再生成元数据 49 的元数据最佳化提示信息 60 指的是在该元数据 49 中，描述包含多少什么种类的信息，或元数据 49 的概要和复杂度的信息。

下面,我们一面参照附图一面说明与本实施例 5 有关的元数据分发服务器的工作。

我们将具有图 14 所示的结构的视频“内容”作为例子,详细述说元数据最佳化提示信息 60。

视频“内容”(Root)(Soccer game program(足球赛节目))大致由 2 个场景(Scene1、Scene2)的前半场和后半场构成,前半场的场景进一步由多个场景(Scene1-1、Scene1-2、..... Scene1-n)(球门场景、角球场景等)构成。在图 14 中,用树形结构表示场景间的时间上的分层结构。

在对应的元数据 49 中描述着这种“内容”的时间上的分层结构,即场景间的时间关系和各场景的开始时间与长度。又,对于各场景,除了与分层的层次对应地该场景具有的特征(例如,颜色直方图和运动的复杂性)外,还描述着标题、简介、类型、注释等的文本信息、重要度等等。此外,在本实施例 5 中,作为元数据的描述格式,用由 ISO 标准化的 MPEG-7。

图 15 表示用 MPEG-7 时的元数据的描述例。在 MPEG-7 中,将各场景描述在称为“视频段”的单元中。在各视频段中,描述着时间信息(场景的开始点和长度)、标题、概要、类型等。此外,存在着与视频段的分层相应地描述在视频段内的信息不同的情形。在图 15 的例子中,在层次 2 和层次 3 的视频段中描述着重要度,但是在层次 4 中不描述重要度。又,只在层次 4 的视频段内描述着颜色和运动的特征量。

通过递归地描述视频段能够表现场景间的时间上的分层关系。在图 15 的描述例中,通过称为“时间分割”的描述,描述着由时间上分割一个视频段的多个视频段构成的情形。在 MPEG-7 中,也能够同样地描述“内容”具有的空间上的分层结构。这时,代替称为“时间分割”的描述,用表示由空间上分割一个视频段的多个视频段构成的“空间分割”的描述。

用于再生元数据 49 的元数据最佳化提示信息 60 是描述该元数

据 49 中包含的信息（描述符）的种类和内容的信息。从而，在元数据最佳化提示信息 60 中，对于图 15 的元数据，包含着表现“内容”具有的时间上的分层结构的描述符（“时间分割”）、表现颜色直方图和运动的复杂性的描述符、表现标题、简介、类型、重要度的描述符。又，作为表示描述“内容”和复杂性的指标，视频段的分层结构的深度最大为 4（层次 1~层次 4）。重要度取 5 个离散的值（{0.0, 0.25, 0.5, 0.75, 1.0}）。作为该重要度的观点，描述着以“TeamA（球队 A）”的观点观看时的重要度和以“TeamB（球队 B）”的观点观看时的重要度。又，也包含着描述重要度的分层位置（描述在视频段的哪个层次上）。

图 16 表示元数据最佳化提示信息 60 的格式例。在图 16 表示的元数据最佳化提示信息 60 中，包含着元数据文件信息和元数据构成要素信息。

元数据文件信息描述元数据文件的所在位置、元数据文件大小、元数据文件格式（表示 XML 格式、二进制格式等的文件格式）、语法文件信息（规定元数据语法的语法文件的所在位置）、表示包含（出现）在元数据中的要素数的出现要素数等、用于预测处理元数据所需的资源（进行元数据的存储/解析所需的存储器大小和解析元数据所需的处理系统（S/W）等）的信息。此外，例如，当用 XML 描述元数据时，规定该描述格式（语法）的 DTD 文件和 schema 文件等与规定元数据文件的格式的语法文件相当，而语法文件信息描述 DTD 文件和 schema 文件等的语法文件的所在位置。

所谓的元数据构成要素信息是描述构成元数据的描述符的种类及其内容的信息。在该元数据构成要素信息中，包含着在元数据中包含的描述符的名称、该描述符在元数据内出现的频度（次数）、该描述符是否包含在语法上可以包含的全部描述符的描述（完全描述性）、或当递归地描述该描述符时，在该描述符持有的时间或空间中的分层性（深度最大值）。例如，在图 15 所示的元数据描述例中，因为“视频段”是递归地描述的描述符，最大具有 4 分层的结构，所以“视频段”

描述符具有的分层深度最大为 4。

进一步,关于在递归地描述的描述符内包含的描述符,该描述符出现的出现位置(分层次)也是提示信息。例如,“重要度”是包含在“视频段”内的描述符,但是当包含在小于等于层次 3 的“视频段”内,即,不包含在层次 4 的视频段内时,“重要度”出现的位置最大达到 3。此外,这样,也能够用分层次指定出现位置,但是当将 ID 分发给包含“重要度”的“视频段”或“视频段”自身时,也可以描述出现位置作为 ID 的清单。又,在具有值的描述符的情形中,描述符的类型和描述符可以取得的值的范围也是一个提示信息。例如,当分别从“TeamA”、“TeamB”的观点出发,用 5 个离散的值({0.0, 0.25, 0.5, 0.75, 1.0})表现重要度时,重要度可以取得的值是具有浮动小数点的类型的清单{0.0, 0.25, 0.5, 0.75, 1.0}。对于每个作为元数据的构成要素的描述符重复进行以上的描述。

图 17 表示按照图 16 的格式描述的元数据最佳化提示信息的一个例子。我们看到在图 17 所示的元数据最佳化提示信息 60 的一个例子中,包含着元数据文件信息、和称为“视频段”与“标题”的元数据构成要素信息。

下面,我们用图 13 说明使用元数据最佳化提示信息 60 进行元数据的再生成的方法。

由提示信息解析单元 61,解析用规定格式描述的元数据最佳化提示信息 60。由元数据解析/再生成单元 63 中,使用从提示信息解析单元 61 输出的被分析后的元数据最佳化提示信息 62,进行元数据 49 的解析,根据关于元数据再生成的条件 65 输出再生成的元数据 64。

图 18 表示由使用被解析的元数据最佳化提示信息 62 的元数据解析/再生成单元 63 实施的元数据的解析方法的一个例子。在该例子中,只从原始的元数据 49 提取具有重要度大于等于 0.5 的特征的视频段,再生成只用关于提取的视频段的描述构成的元数据。

首先,元数据解析/再生成单元 63 根据用于元数据再生成的条件 65,特定再生成所需的元数据(步骤 S1)。这里,因为只提取具有重

要度大于等于 0.5 的特征的视频段，所以“重要度”和“视频段”是再生成所需的描述符。

其次，根据被分析后的元数据最佳化提示信息 62，判定在元数据 49 中是否包含在步骤 S1 特定的描述符（以下，将描述符“重要度”作为例子进行说明）（步骤 S2）。

当在元数据中包含“重要度”描述符时，进行元数据的解析（步骤 S3），当不包含“重要度”描述符时，结束元数据的解析处理（步骤 S4）。

又，当在被分析后的元数据最佳化提示信息 62 中指定“重要度”的出现位置为直到分层结构的层次 3 时，在结束直到层次 3 的视频段的解析的时刻（步骤 S5），不进行大于等于层次 4 的分层的解析，结束解析处理（步骤 S6）。

此外，需要时为了进行别的元数据 49 的解析，重复步骤 S1 以后的处理。又，当在元数据最佳化提示信息 62 中指定“重要度”描述符的出现次数为 20 时，在结束 20 个“重要度”描述符的解析的时刻（步骤 S5），结束元数据的解析（步骤 S6）。进一步，在步骤 S4 或步骤 S6 结束元数据的解析处理后，需要时为了进行别的元数据的解析，重复步骤 S1 以后的处理。

图 19 表示使用被分析后的元数据最佳化提示信息 62 的元数据的解析方法的别的例子。在该例子中，只提取包含“标题”描述符的视频段，再生成元数据。在元数据中是否包含“标题”描述符的判定与图 18 的例子相同。

元数据解析/再生成单元 63，当在元数据中包含“标题”描述符时，判定是否是和在元数据最佳化提示信息中描述的出现位置的 ID 一致的视频段（步骤 S13）。

当不与 ID 一致时，因为是不包含“标题”描述符的视频段，所以跳过该视频段的描述的解析（步骤 S16）。

当与 ID 一致时，为了取得“标题”描述符，进行该视频段的描述的解析（步骤 S15）。

接着，当结束对与出现位置的 ID 一致的全部视频段的解析时（步

骤 S17)，因为在此以上，在元数据内不存在包含“标题”描述符的视频段，所以结束解析处理（步骤 S18）。

此外，为了需要时进行别的元数据的解析，重复步骤 S11 以后的处理。输出用通过以上的解析处理提取的描述符再构成的元数据 64。

由元数据分发单元 45，将再构成的元数据 64 分发给各种客户终端。

此外，虽然没有图示出来，但是因为在元数据再生后，元数据文件的所在位置和元数据文件大小、出现在元数据中的要素数、关于元数据构成要素的信息也发生变化，所以也可以再生成与再生成后的元数据对应的元数据最佳化提示信息。

如上所述，至今为了再生成元数据，必需解析元数据中包含的全部描述符，但是在本实施例 5 中，因为使用描述在元数据 49 中包含的描述符的清单和描述符的出现位置、出现次数等的元数据最佳化提示信息 60，解析元数据 49 的描述符，所以能够省去用于再生成元数据的元数据 49 自身的解析，又，因为根据出现位置和出现次数，能够省去与再生成条件不一致的描述符的解析，所以能够减少伴随着元数据的解析和再生成的处理费用（处理量和存储器使用量等）。

实施例 6

在上述实施例 5 中，我们说明了用为了元数据再生成的元数据最佳化提示信息 60，减少伴随着元数据的解析和再生成的处理费用的元数据分发服务器，但是在本实施例 6 中，我们说明用元数据最佳化提示信息，减少伴随着元数据的检索的处理的元数据检索服务器（元数据检索装置）。

我们一面参照附图一面说明与本发明的实施例 6 有关的元数据检索服务器。图 20 是表示与本发明的实施例 6 有关的元数据检索服务器的构成的方框图。

在图 20 中，元数据检索服务器 600 备有提示信息解析单元 61、元数据解析单元 71 和检索单元 73。

因为提示信息解析单元 61 与上述实施例 5 相同，所以省略对它的说明。元数据解析单元 71，用被分析后的元数据最佳化提示信息 62 和检索条件 70，能够用很少的处理费用高效率地解析描述“内容”的结构和特征的膨大的元数据 49。检索单元 73 利用元数据的解析结果 72，检索适合于检索条件的“内容”。

下面，我们一面参照附图一面说明与本实施例 6 有关的元数据检索服务器的工作。

图 21 是表示与本实施例 6 有关的元数据检索服务器的元数据解析单元的工作的操作程序图。

元数据解析单元 71，用与各元数据对应的元数据最佳化提示信息 62 进行大于等于一个的元数据的解析。所谓元数据的解析，这里指的是从元数据提取检索所需的特征描述。例如，当给出视频段的颜色特征量作为检索条件，检索具有接近该视频段的特征的视频段时，需要提取具有与颜色有关的特征描述的视频段。在图 15 所示的元数据例子中，因为在层次 4 的视频段上附加颜色的特征描述（“颜色直方图”），所以提取与层次 4 的视频段有关的描述。

元数据解析单元 71，解析检索条件 70，特定对检索有效的描述符（步骤 S21）。存在着作为检索条件给出按照在 MPEG-7 中规定的描述的特征量的情形和给出图像与关键字的情形等。当作为按照 MPEG-7 的描述的特征量（例如颜色配置信息）给出检索条件时，该描述符（颜色配置信息）成为对检索有效的描述符。又，作为关键字给出检索条件时，文本形式的描述符（标题、简述、注释等）成为对检索有效的描述符。

下面，参照元数据最佳化提示信息 62，判定选出的描述符是否包含在元数据 49 中（步骤 S22）。当用于检索的描述符不包含在元数据 49 中时，结束该元数据 49 的解析处理（步骤 S24），需要时进行别的元数据 49 的解析。

当选出的描述符包含在元数据 49 中时，进行该元数据的解析（步骤 S23）。关于元数据的解析方法，与上述实施例 5 的情形相同，用

元数据最佳化提示信息 62 高效率地进行图 18 和图 19 所示的元数据解析处理（步骤 S25~S26）。通过以上的处理，由元数据解析单元 71 提取检索所需的特征描述。

由检索单元 73，利用由元数据解析单元 71 输出的元数据的解析结果（检索所需的特征描述），检索适合于检索条件的“内容”。在上述例子中，因为由元数据解析单元 71 输出与具有颜色的特征描述（“颜色直方图”）的视频段有关的描述，所以判定与作为检索条件给出的颜色特征量（直方图）的适合性，作为检索结果 74 输出适合的视频段的信息（例如“时间信息”）。

如以上那样，在本实施例 6 中，因为使用元数据最佳化提示信息 60，解析元数据 49，所以能够省去用于元数据再生成的元数据 49 自身的解析。又，因为根据出现位置和出现次数，能够省去检索不需要的描述符的解析，所以能够减少伴随着元数据的检索的处理费用（处理量和存储器使用量等）。

实施例 7

在上述实施例 5 或实施例 6 中，我们说明了利用元数据最佳化提示信息的服务器一侧，但是在本实施例 7 中，我们说明利用元数据最佳化提示信息的客户终端（元数据再生成条件设定装置）。

我们一面参照附图一面说明与本发明的实施例 7 有关的客户终端。图 22 是表示与本发明的实施例 7 有关的客户终端的构成的方框图。

在图 22 中，客户终端 48A 备有提示信息解析单元 80、元数据再生成条件设定单元 82。

此外，在图 22 中，在客户终端 48A 中备有的功能中，只表示与利用元数据最佳化提示信息 60 设定用于元数据再生成的条件的部件有关的部分。

下面，我们一面参照附图一面说明与本实施例 7 有关的客户终端的工作。

提示信息解析单元 80 进行以规定的格式描述的元数据最佳化提示信息 60 的解析。因为该提示信息解析单元 80 与上述实施例 5 相同，所以省略对它的详细说明。

元数据再生成条件设定单元 82，根据由提示信息解析单元 80 输出的解析结果 81，进行元数据再生成的条件设定 83。所谓的条件设定指的是，例如，从元数据最佳化提示信息 60 中包含的描述符的种类中选择客户终端 48A 不需要的描述符。当客户终端 48A 不备有用特征量的检索功能时，不需要表现颜色直方图和运动的复杂性等的特征量的描述符。

作为条件设定的别的例子，当描述“内容”的场景关系的分层结构越深，元数据的复杂性越增加时，根据在元数据最佳化提示信息 60 中的描述的分层结构的深度的最大值，在客户终端设定可以处理的分层结构的深度。又，在别的例子中，根据元数据最佳化提示信息 60 中描述的重要度可以取得的值的信息，设定用户着眼的观点和选择的场景的重要度的阈值。

如上所述，当分别从“TeamA”、“TeamB”的观点出发，重要度取 5 个离散的值（{0.0, 0.25, 0.5, 0.75, 1.0}）时，能够进行只选择从“TeamA”的观点具有大于等于 0.5 的重要度的场景等的设定。

将由元数据再生成条件设定单元 82 设定的用于元数据再生成的条件 83 发送给元数据分发服务器。在元数据分发服务器一侧，根据用于元数据再生成的条件和客户终端的终端性能，再构成元数据。例如，当原始的元数据的分层结构的深度的最大值为 4，在元数据再生成条件中，设定在客户终端可以处理的分层结构的深度为 2 时，以使分层结构的深度的最大值为 2 的方式，再构成元数据的结构。

又，当在元数据再生成条件中，只选择并设定从“TeamA”的观点，重要度大于等于 0.5 的场景时，再生只由与同一条件一致的场景构成的元数据。与上述实施例 5 相同能够用元数据最佳化提示信息高效率地进行元数据再生成。

如以上那样，在本实施例 7 中，因为能够使用元数据最佳化提示

信息 60 设定用于元数据再生成的条件，所以能够与客户终端和应用相应地生成适合的元数据。

实施例 8

在上述实施例 5 或实施例 6 中，我们说明了用元数据最佳化提示信息再生成元数据，分发再生成的元数据的服务器，但是在本实施例 8 中，我们说明用元数据最佳化提示信息解析元数据，利用解析结果再生成适合于客户终端和用户嗜好的元数据并进行分发的“内容”分发服务器（“内容”分发装置）。

我们一面参照附图一面说明与本发明的实施例 8 有关的“内容”分发服务器。图 23 是表示与本发明的实施例 8 有关的“内容”分发服务器的构成的方框图。

在图 23 中，“内容”分发服务器 500A 备有提示信息解析单元 61、元数据解析单元 86 和“内容”再构成/分发单元 88。

下面，我们一面参照附图一面说明与本实施例 8 有关的“内容”分发服务器的工作。

因为提示信息解析单元 61 的工作与上述实施例 5 相同，所以省略对它的说明。

元数据解析单元 86，使用从提示信息解析单元 61 输出的被分析后的元数据最佳化提示信息 62，进行元数据 49 的解析，提取与客户终端有关的信息、或者，与用户嗜好等的“内容”再构成有关的条件 85 一致的描述。使用提示信息的解析方法具有与上述实施例 5 相同之处，但是与实施例 5 不同之处不是使用提取的描述再生成元数据，而是再构成“内容”这一点。将由该元数据解析单元 86 提取的描述，即被分析后的元数据 87 输出到“内容”再构成/分发单元 88。

“内容”再构成/分发单元 88，根据由元数据解析单元 86 提取的描述，再构成“内容”89。这里，我们用上述实施例 5 中述说的例子进行说明。在实施例 5 的例子中，只从元数据 49 提取具有重要度大于等于 0.5 的特征的视频段，再生成只由与提取的视频段有关的描述构

成的元数据。

在本实施例 8 中，只从元数据 49 提取具有重要度大于等于 0.5 的特征的视频段，再生成并分发只由与提取的视频段对应的场景构成的“内容”90。因为在与提取的视频段有关的描述中，描述了对应的“内容”的所在位置和在该视频段的“内容”内的位置（时间信息），所以也能够从“内容”切出对应的场景，再构成一个“内容”90，并分发再构成的“内容”90，但是也能够从“内容”切出对应的场景，顺次地分发切出的场景。

如以上那样，在与本实施例 8 有关的“内容”分发服务器 500A 中，因为使用描述元数据 49 中包含的描述符的清单和描述符的出现位置、出现次数等的元数据最佳化提示信息 60，解析元数据 49，所以能够省去用于元数据再生成的元数据 49 自身的解析，又，因为根据出现位置和出现次数，能够省去与再生成条件不一致的描述符的解析，所以能够减少伴随着当再生成并分发适合于客户终端和用户嗜好的“内容”时的元数据的解析和“内容”的再构成的处理费用（处理量和存储器使用量等）。

本发明，如以上说明的那样，因为将包含运动图像和音频的多媒体“内容”分割成多个场景，编辑分割成多个的场景，生成描述多媒体“内容”的分层结构的场景结构信息元数据，所以能够生成描述包含视频数据等的多媒体“内容”具有的分层结构的元数据。

图1

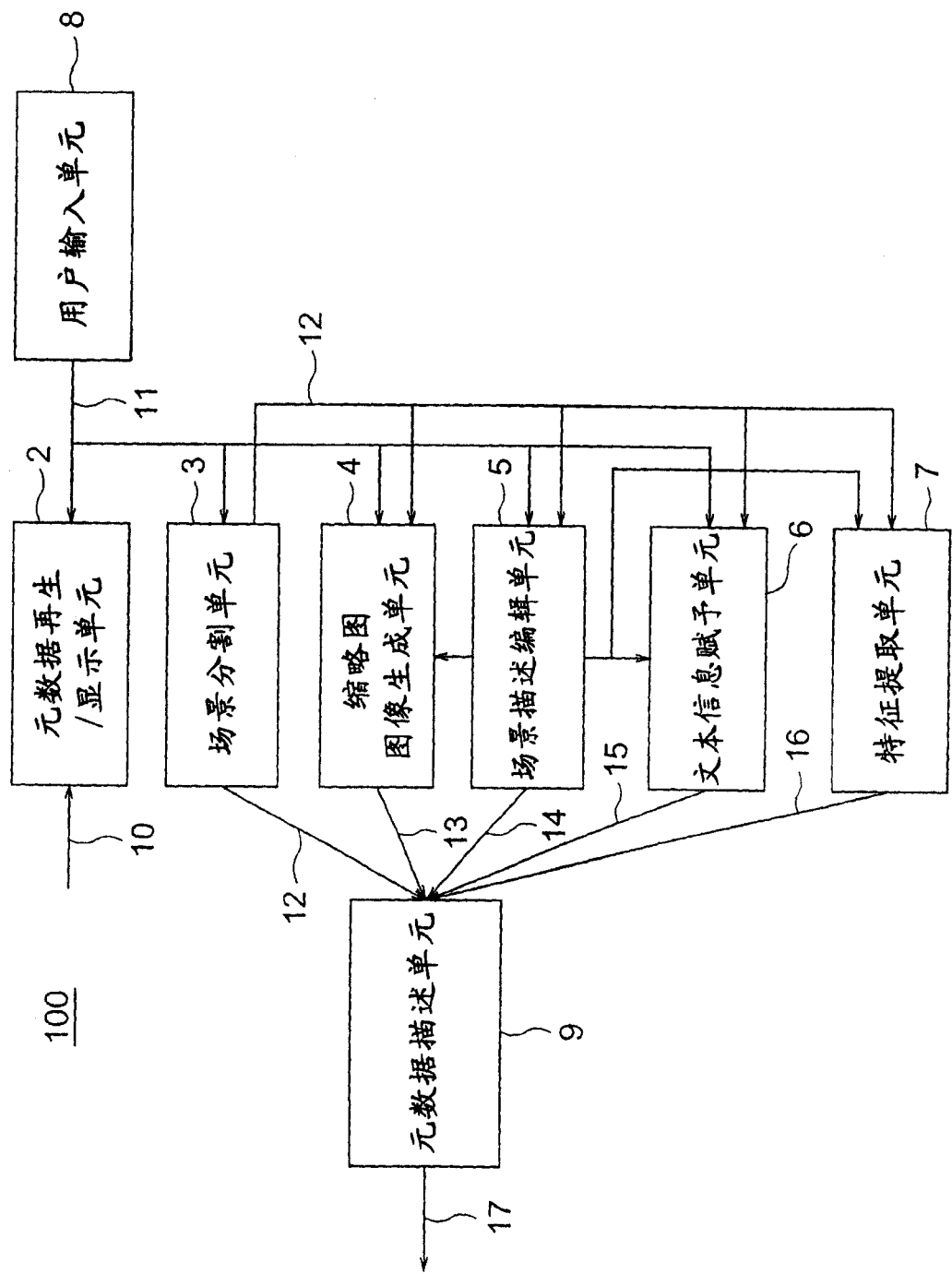


图2

新闻摘要	国内新闻	国际新闻	经济新闻	特辑	体育足球	体育棒球	天气预报
------	------	------	------	----	------	------	------

图3

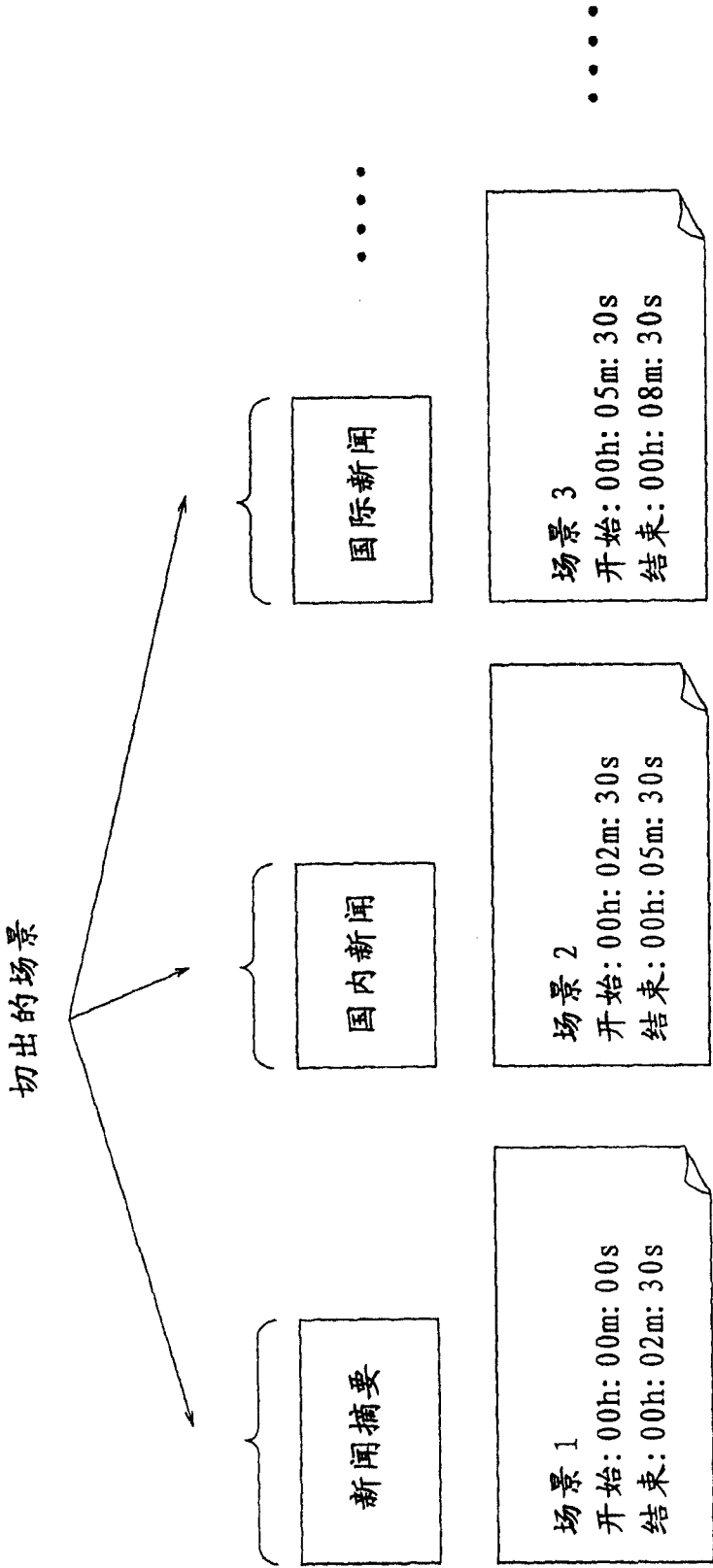


图4

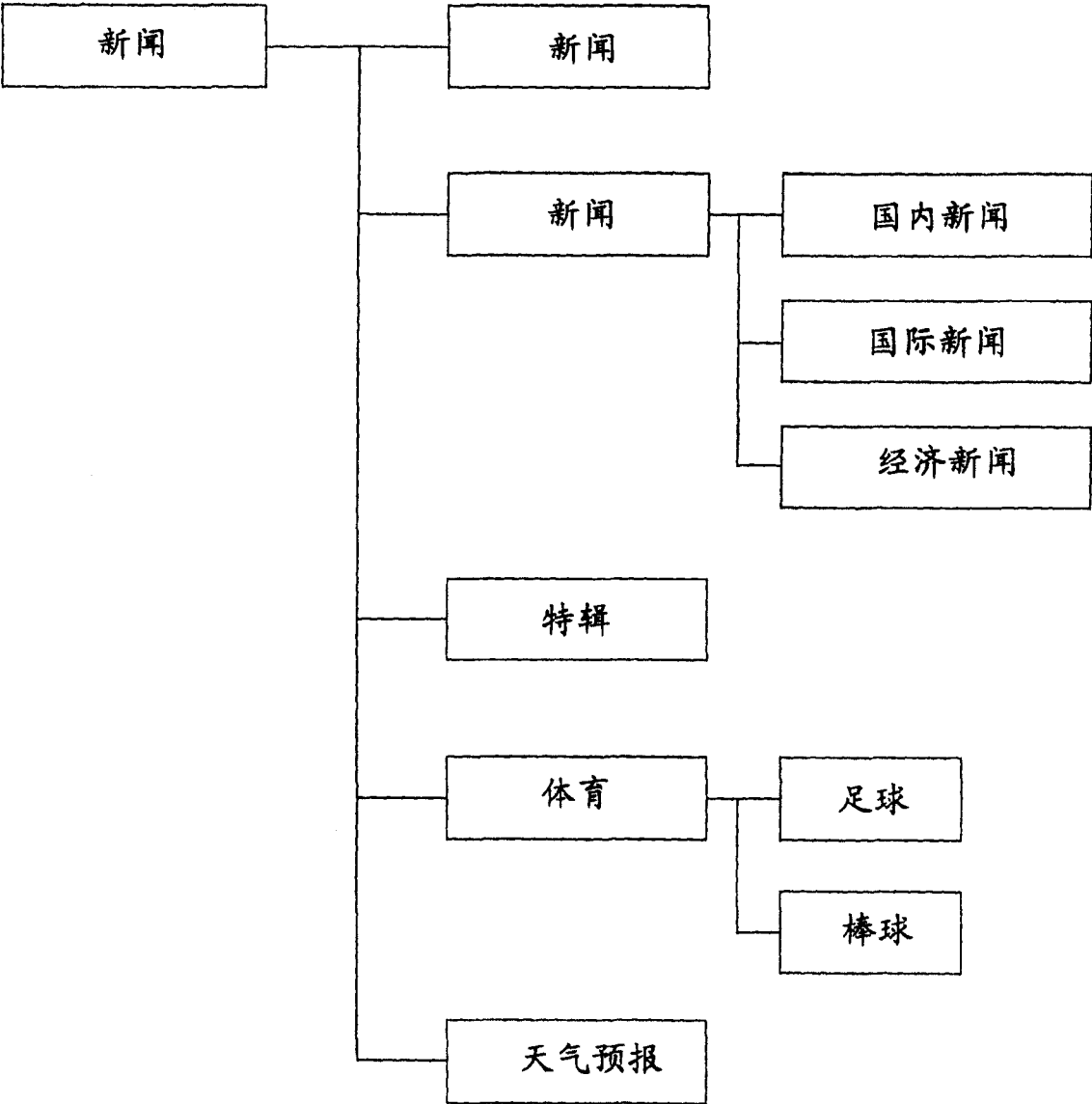


图5

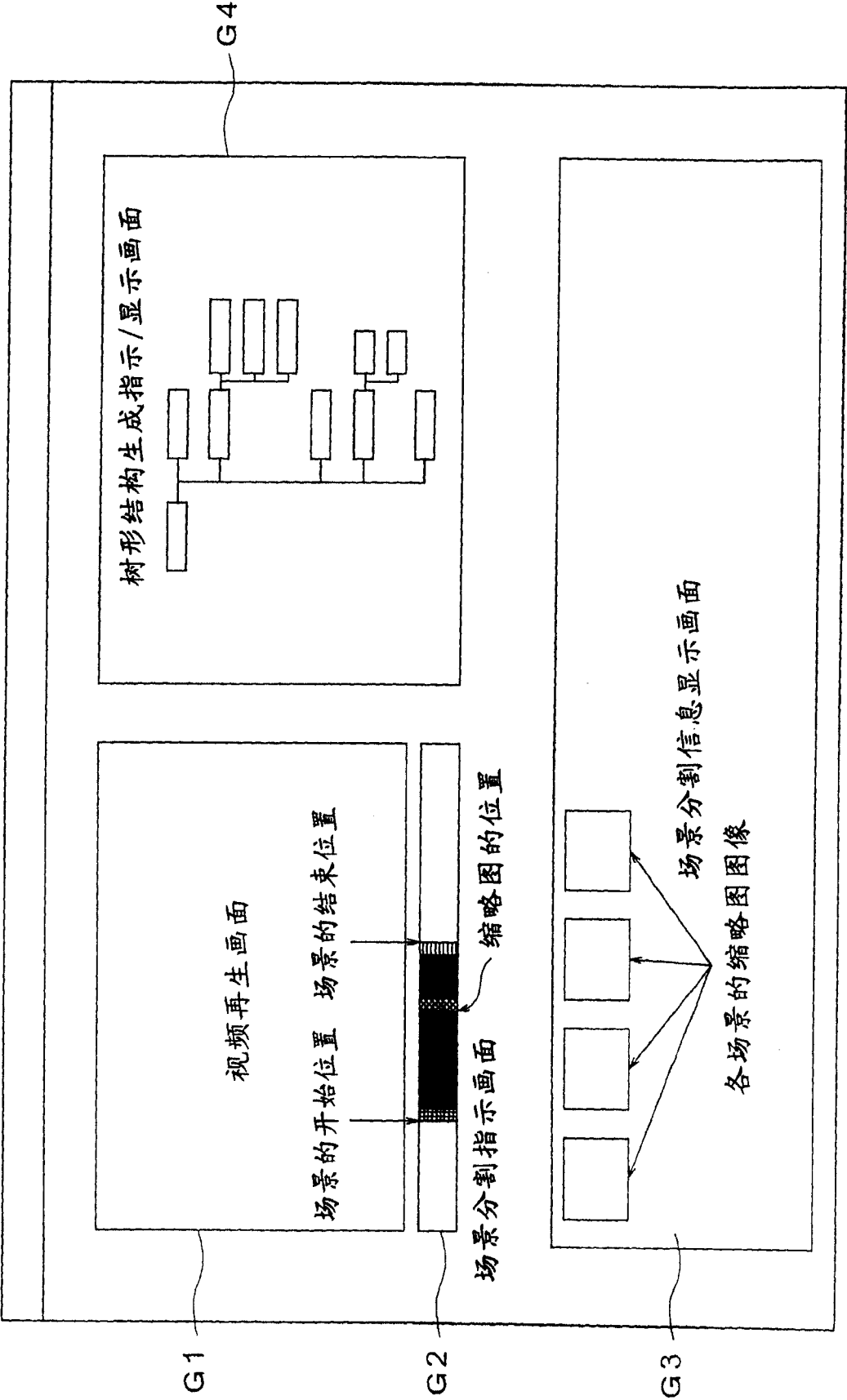


图6

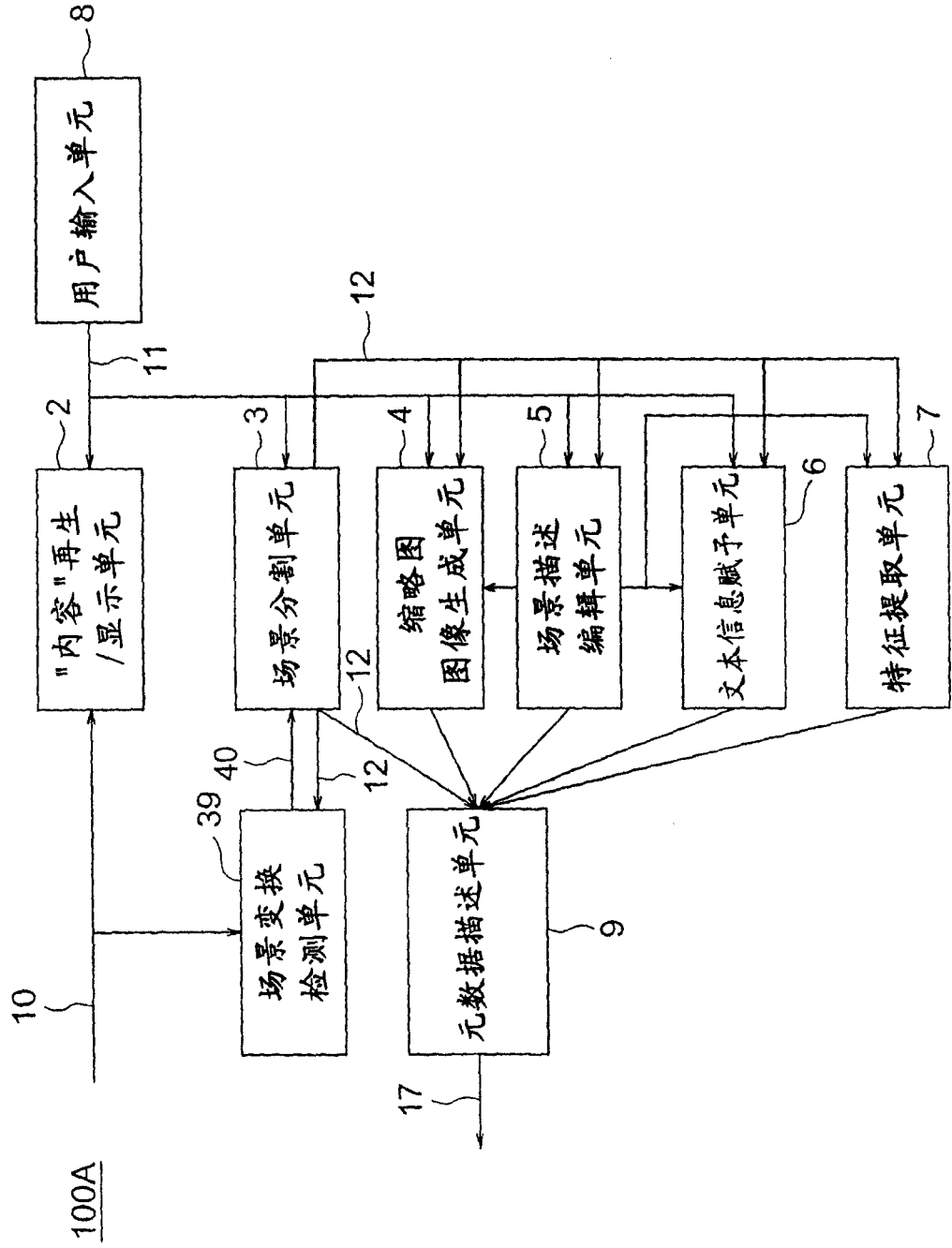


图7

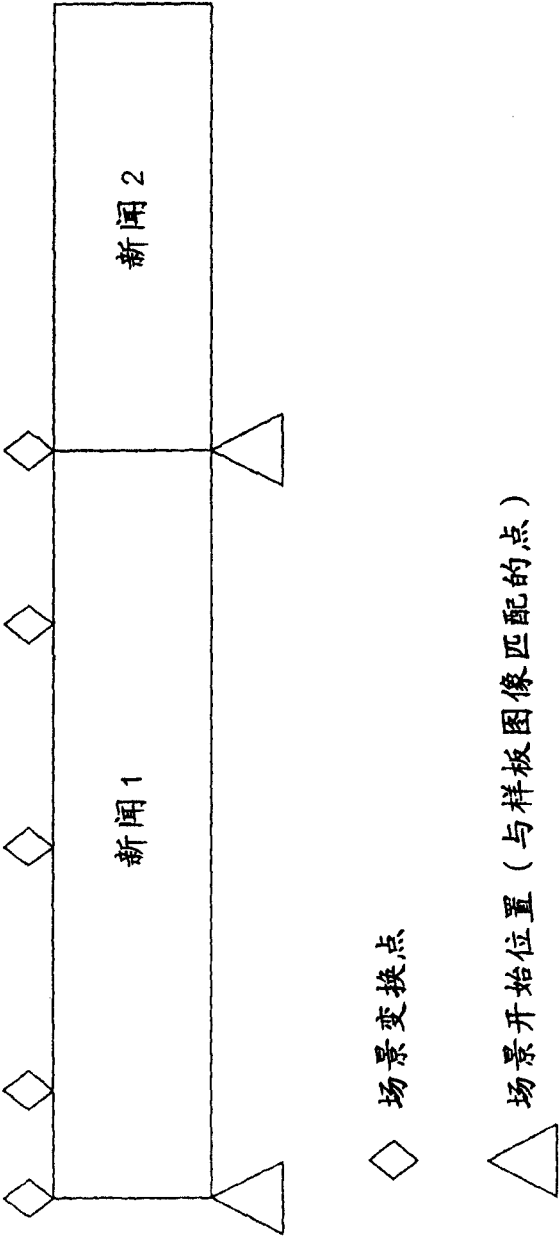


图 8

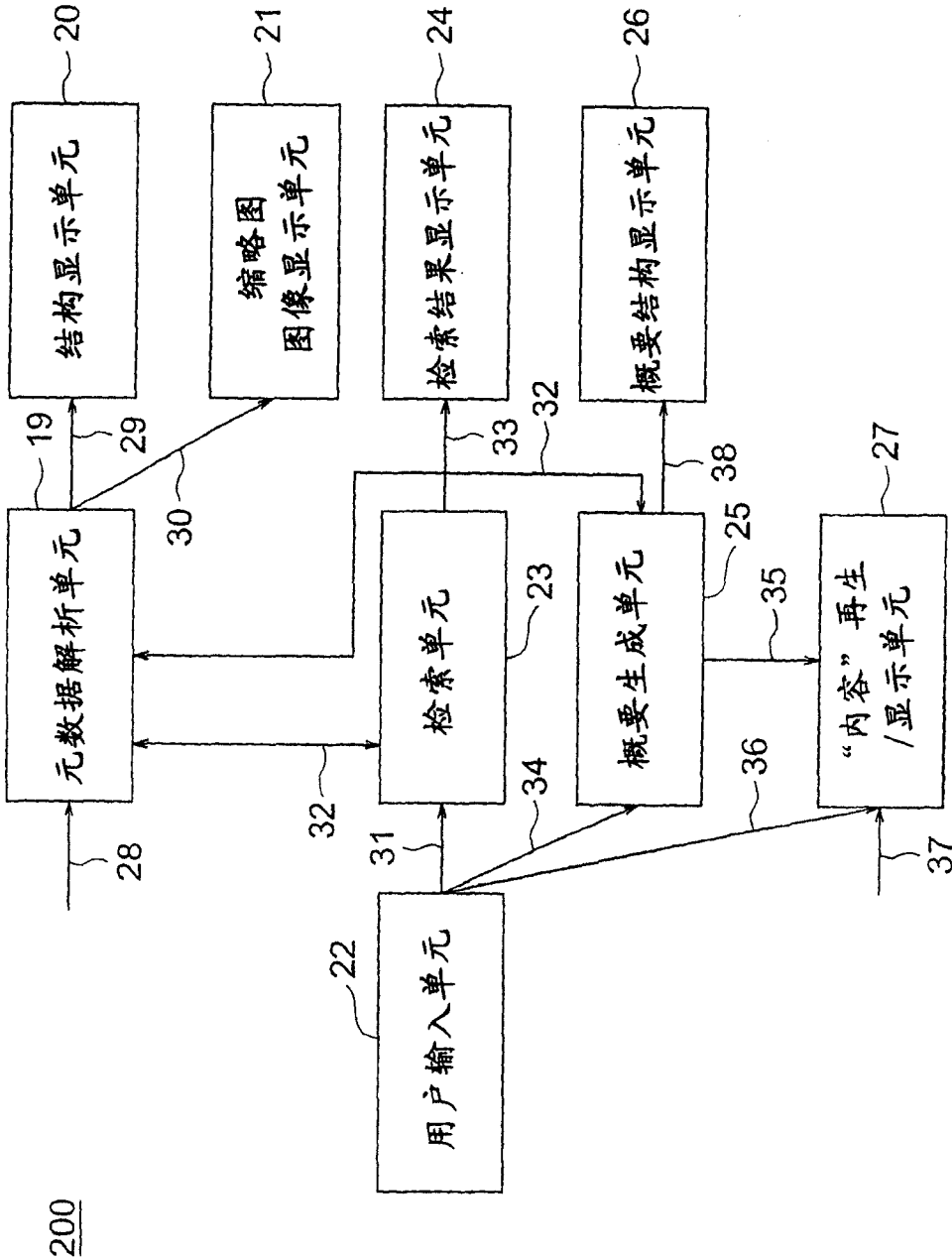
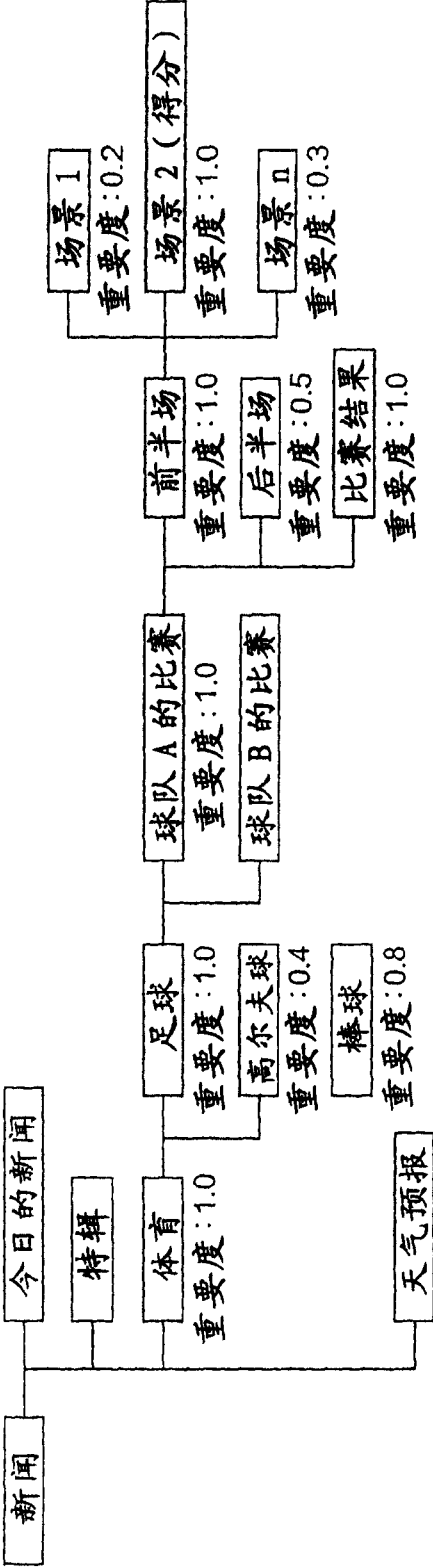
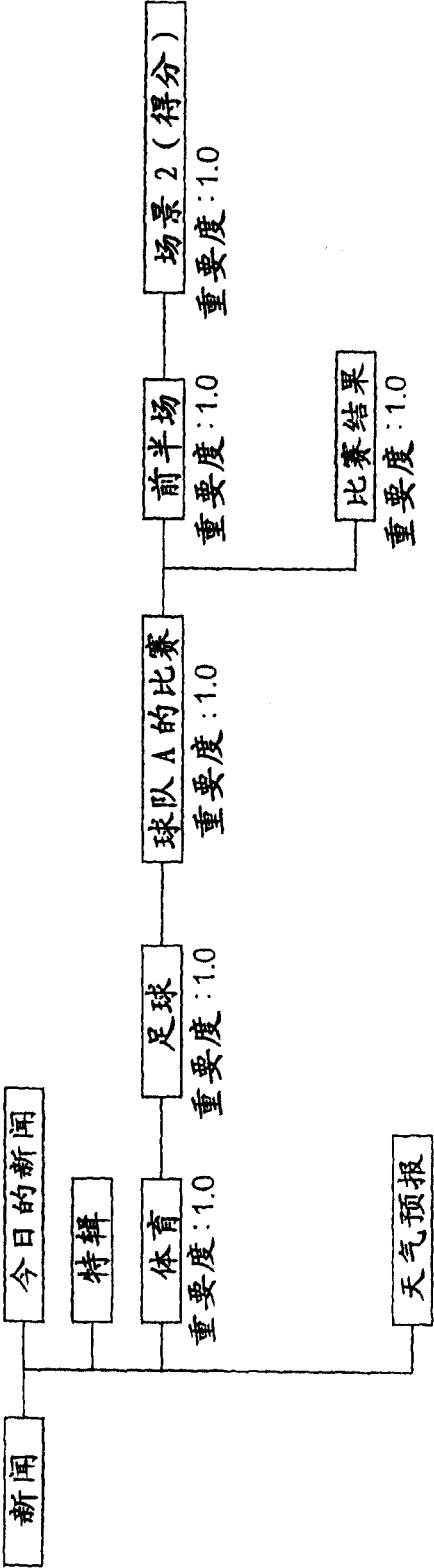


图9



(a) 原始“内容”的场景结构



(b) 生成概要的“内容”的场景结构

图10

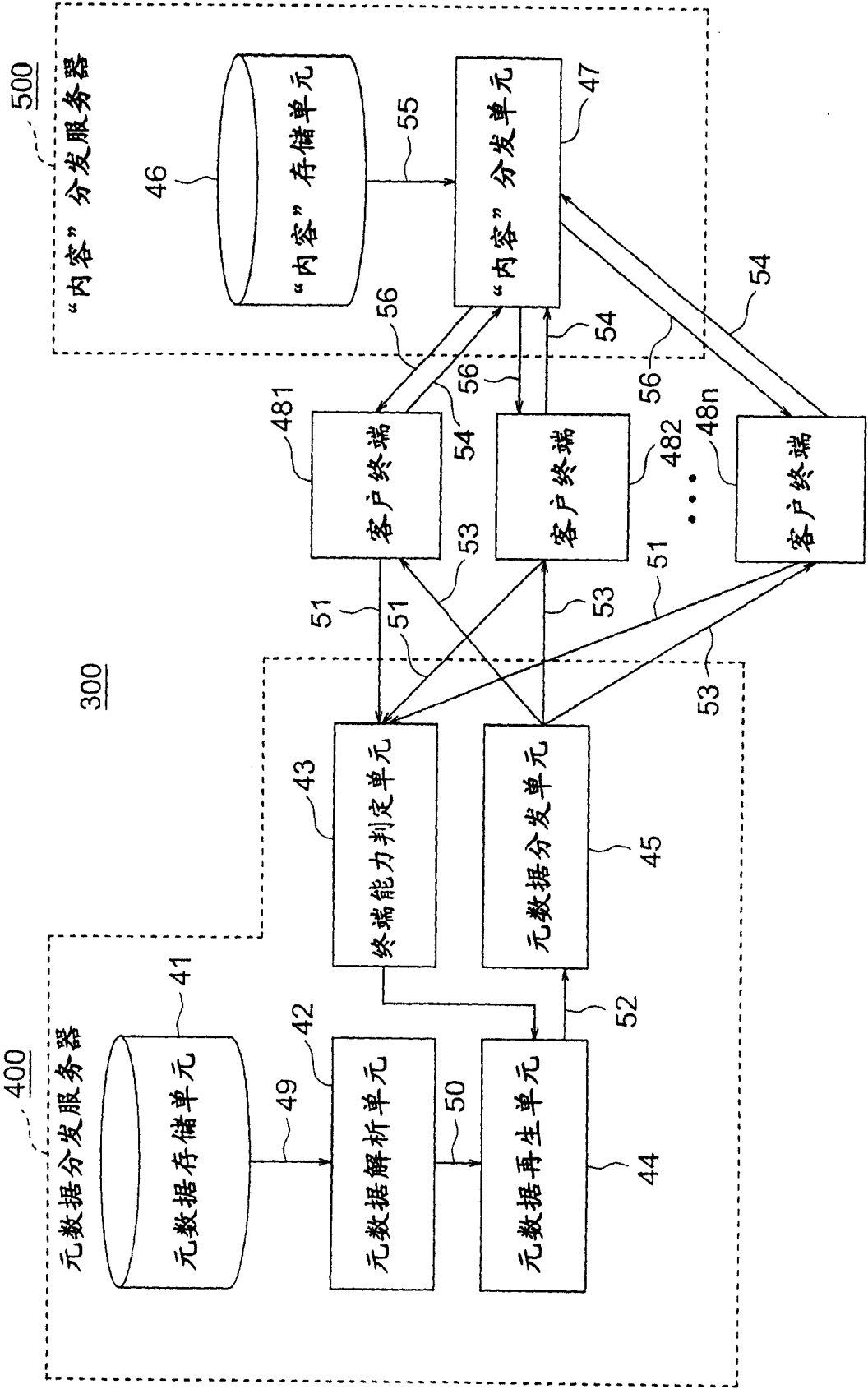


图 11

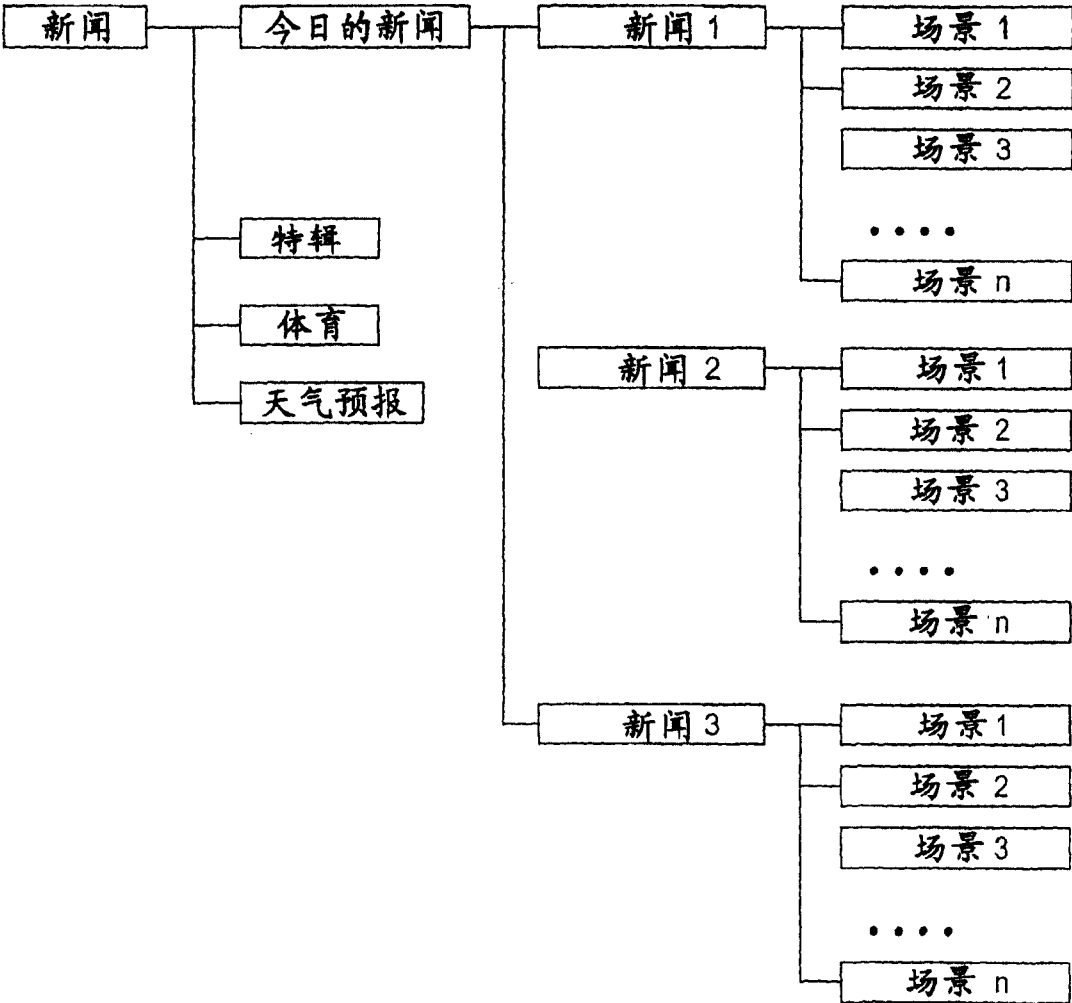


图12

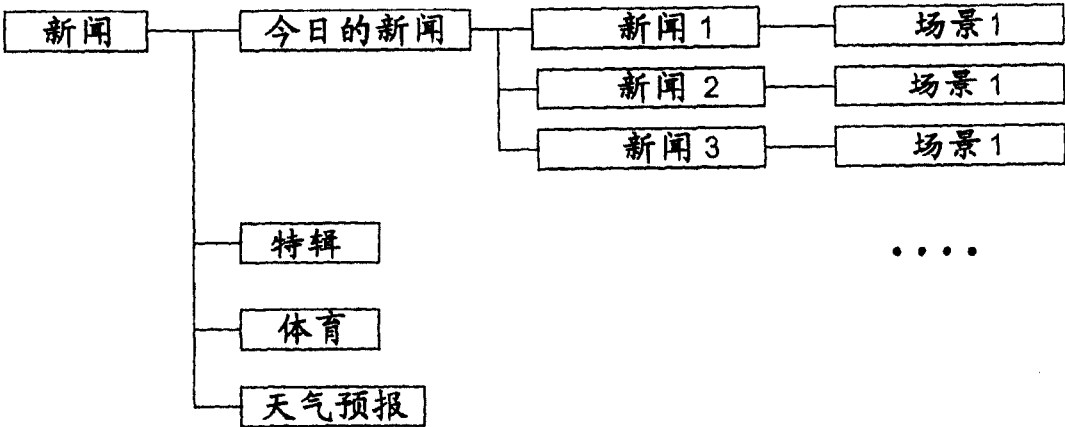


图13

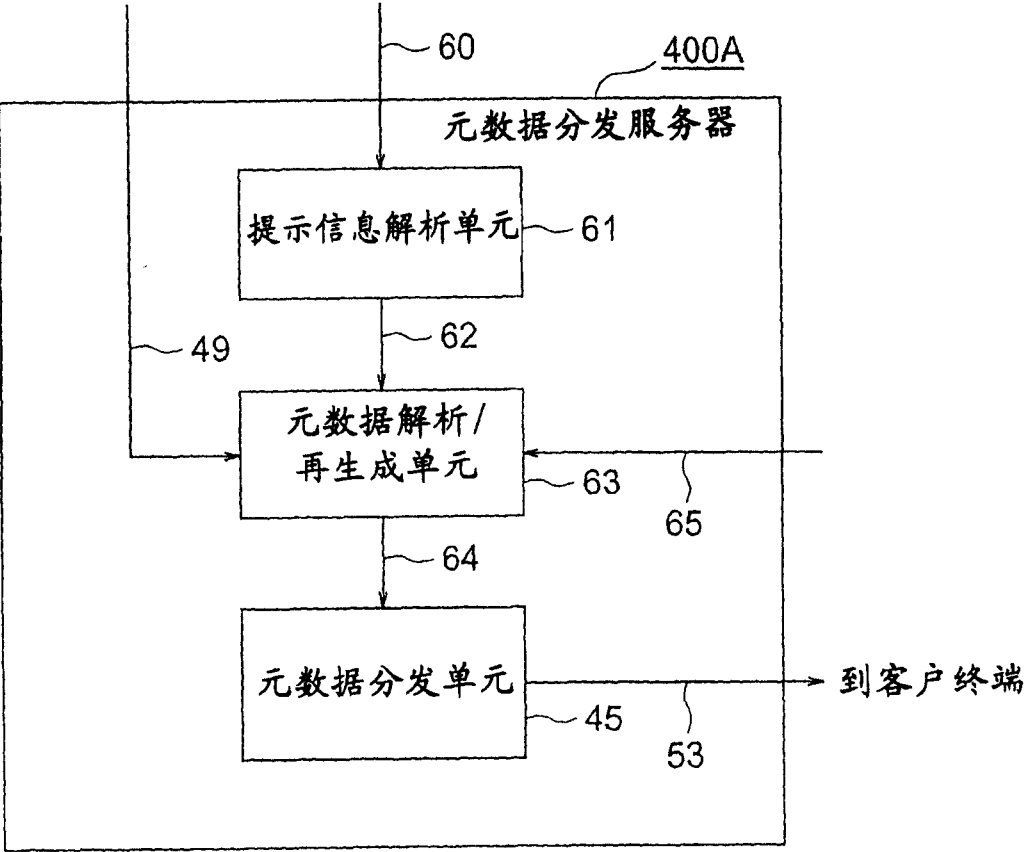


图14

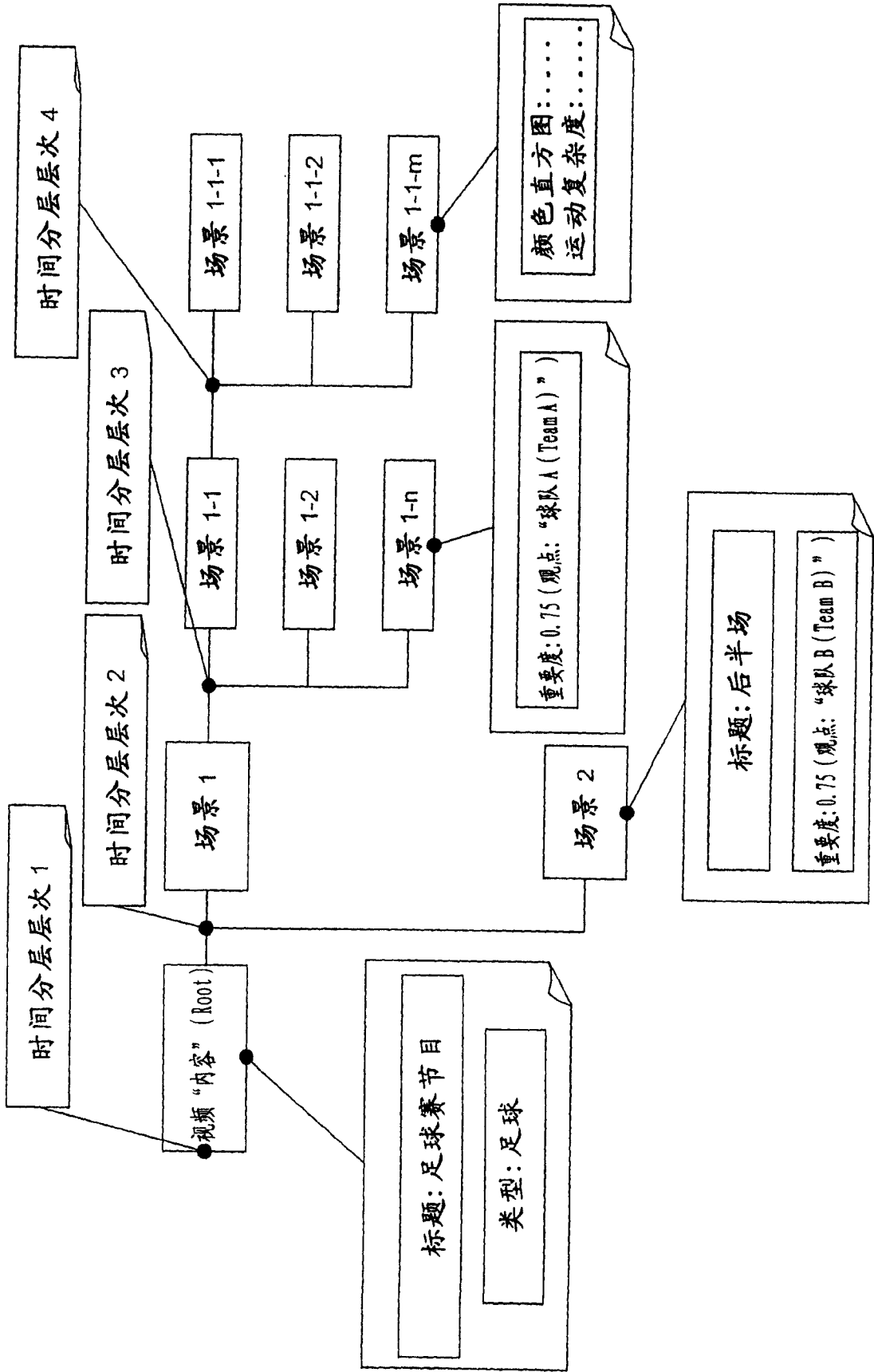


图15

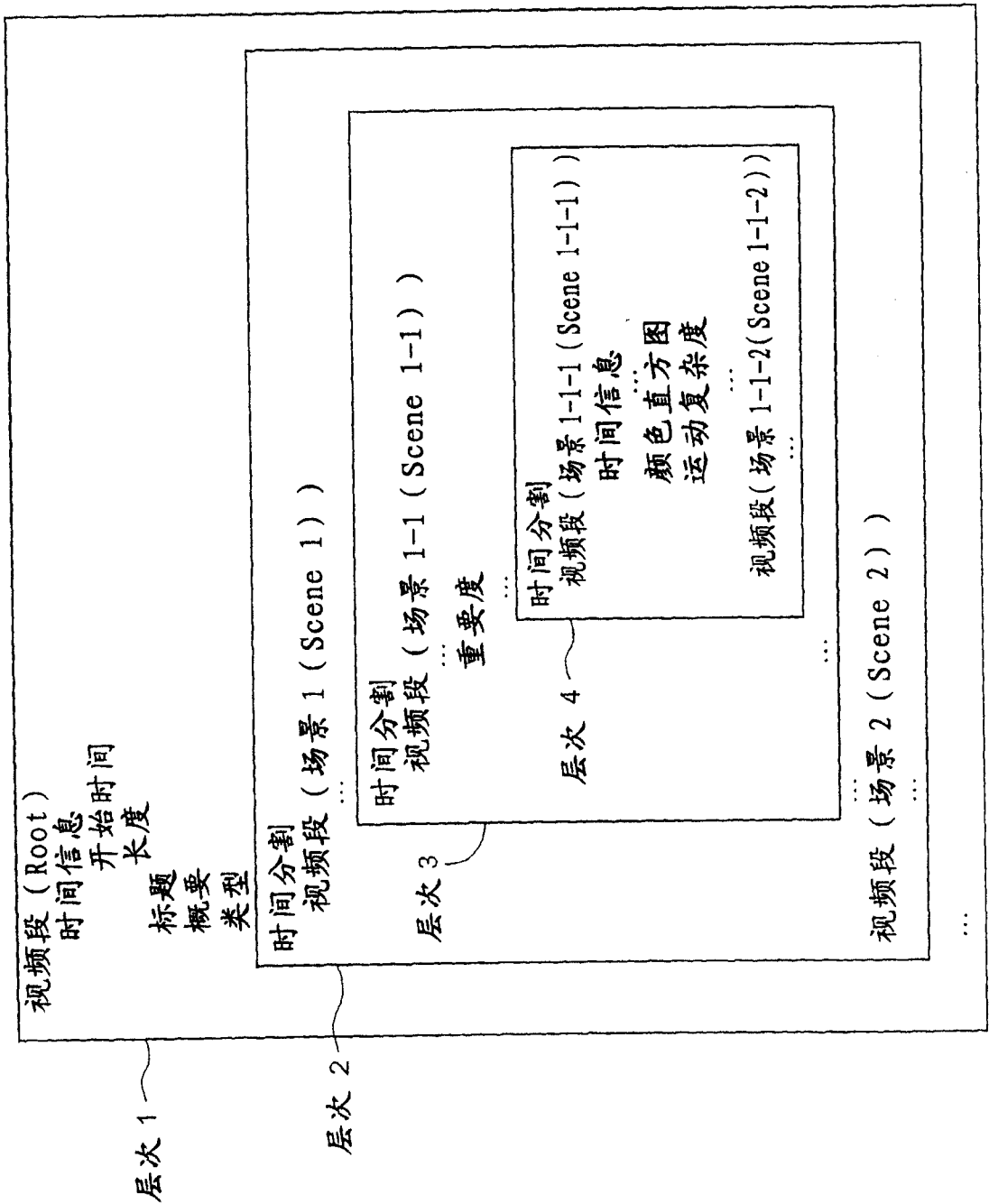


图16

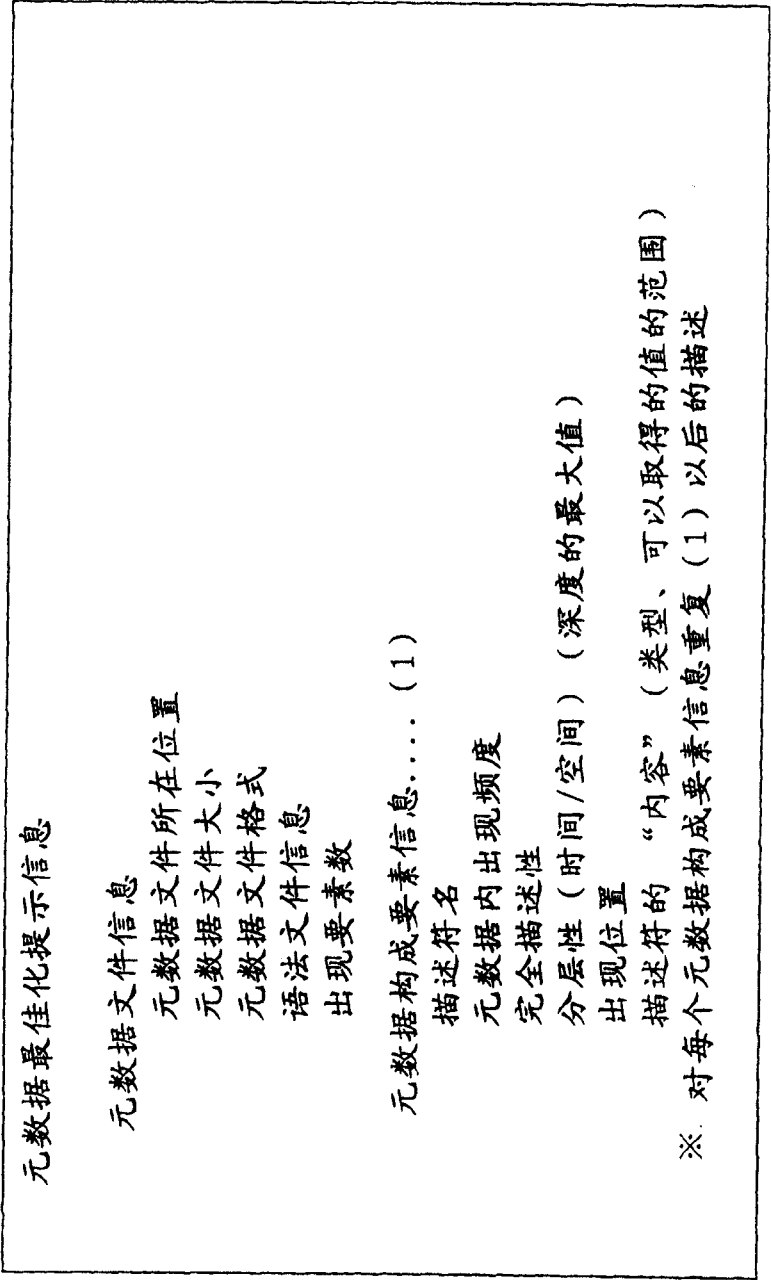


图17

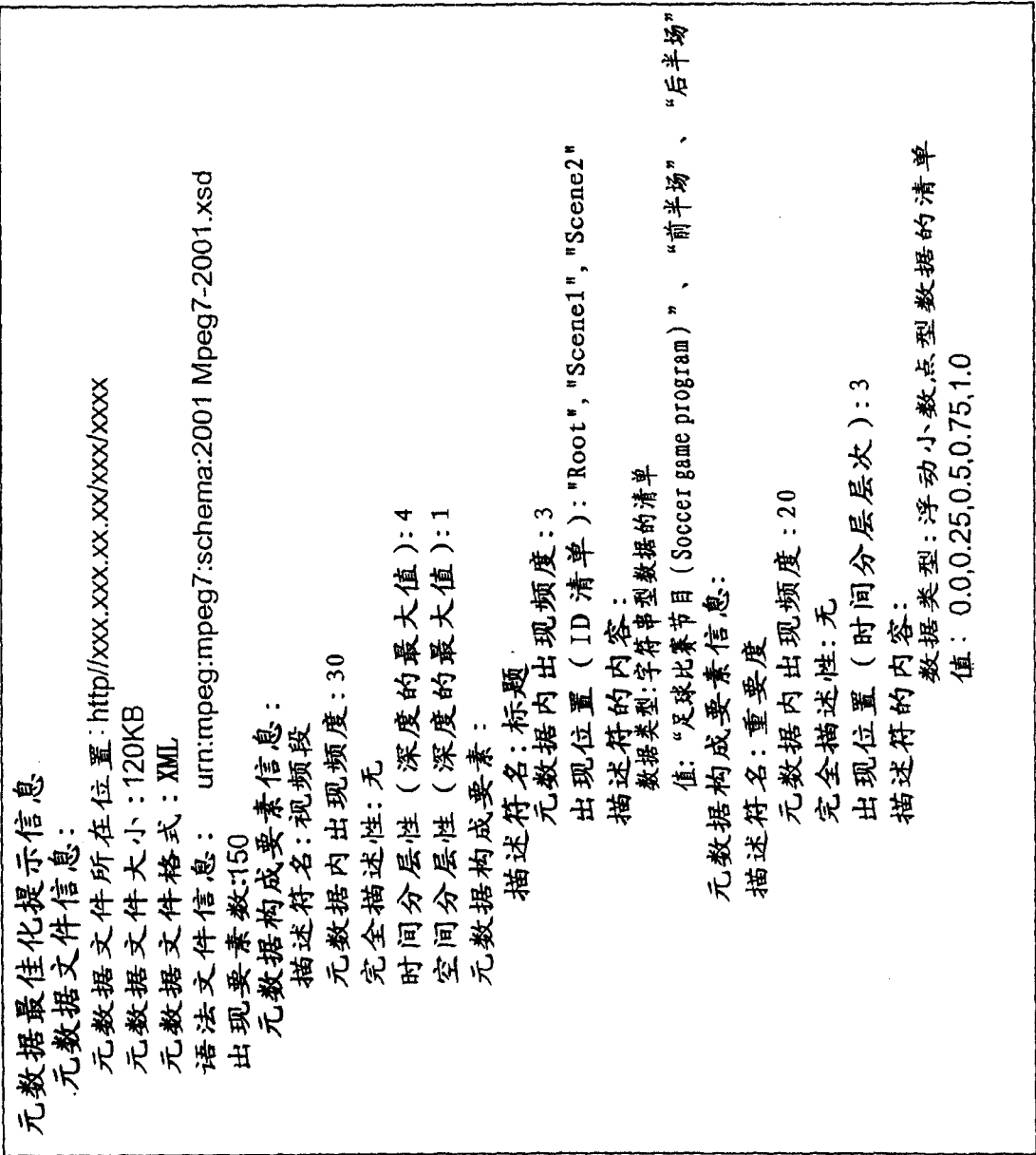


图18

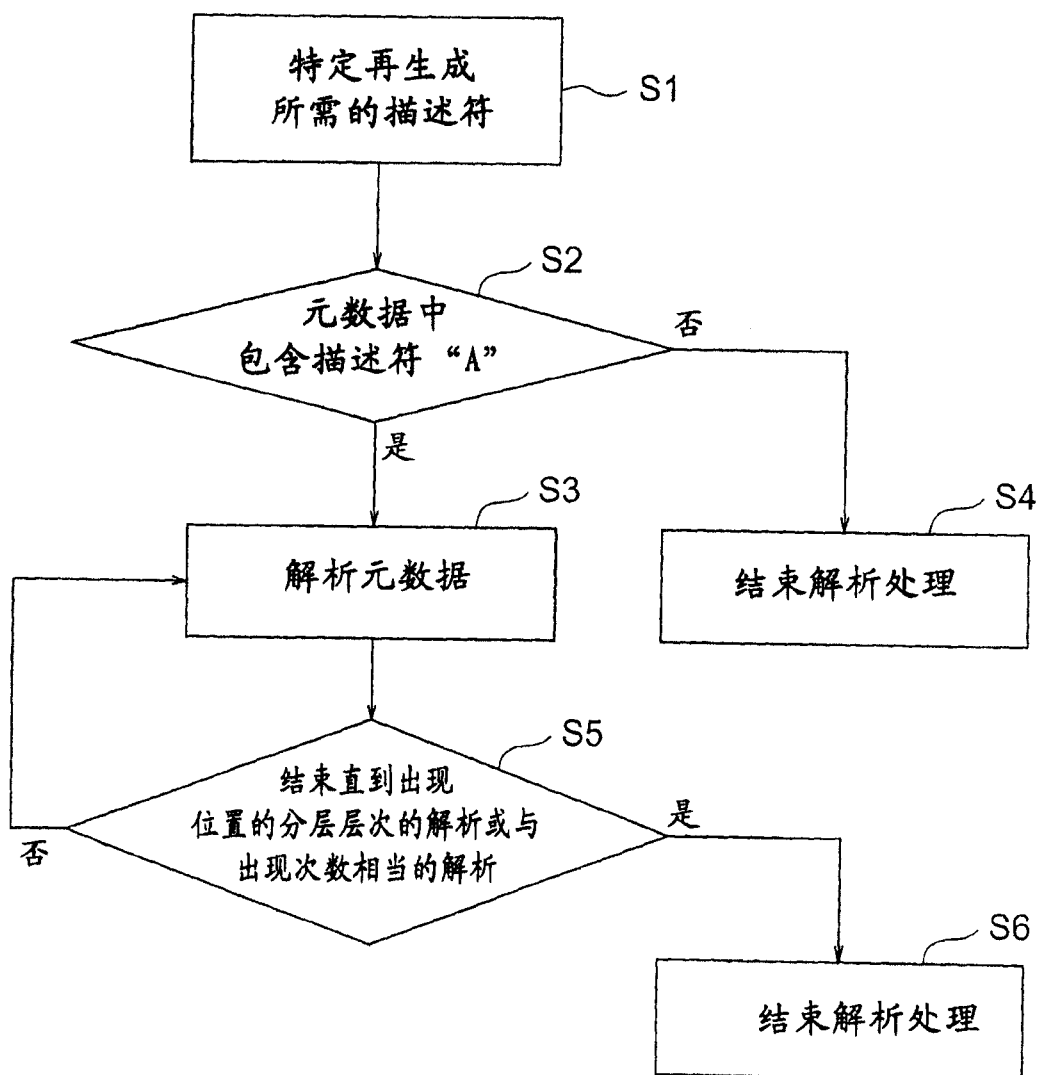


图 19

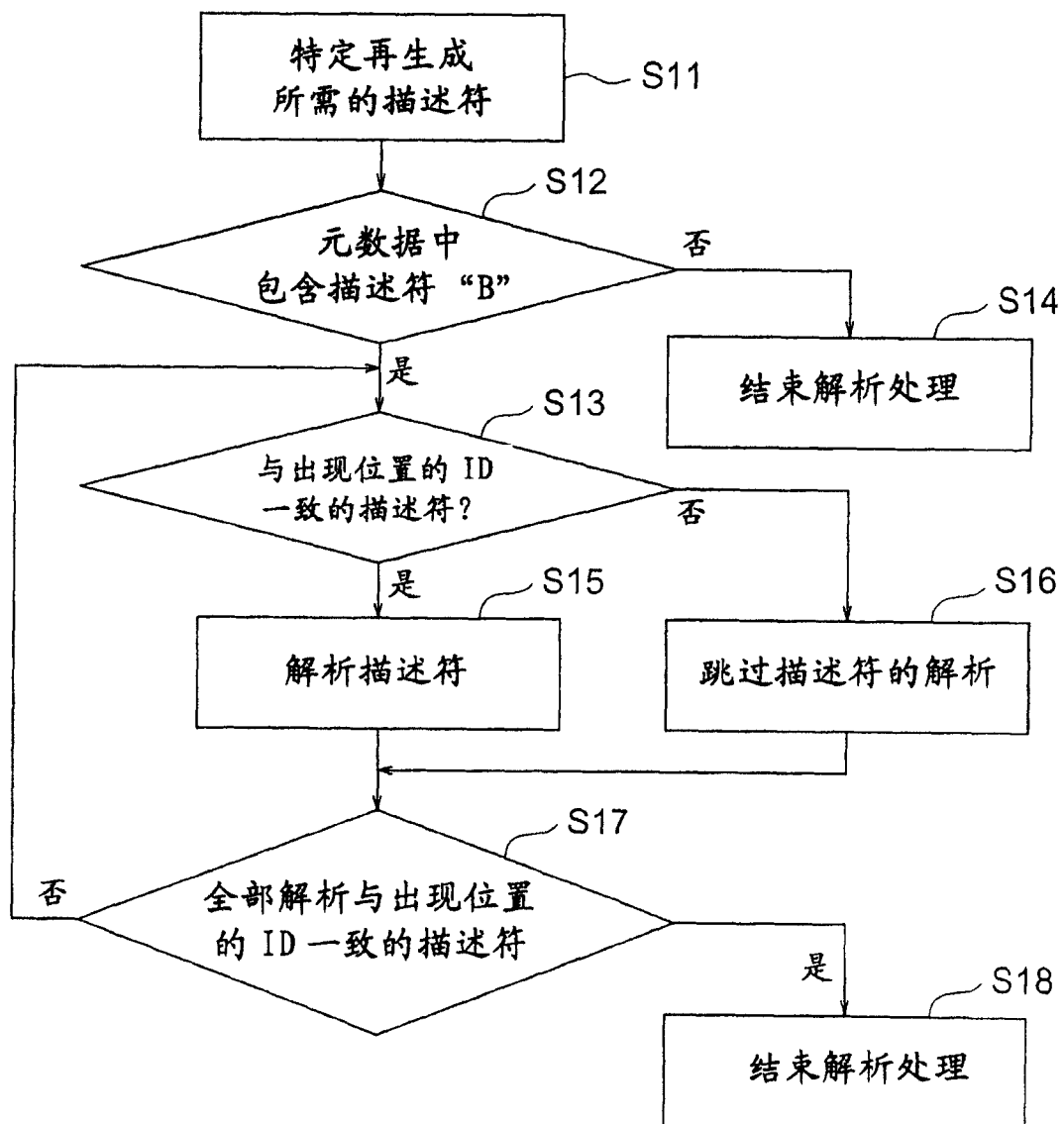


图 20

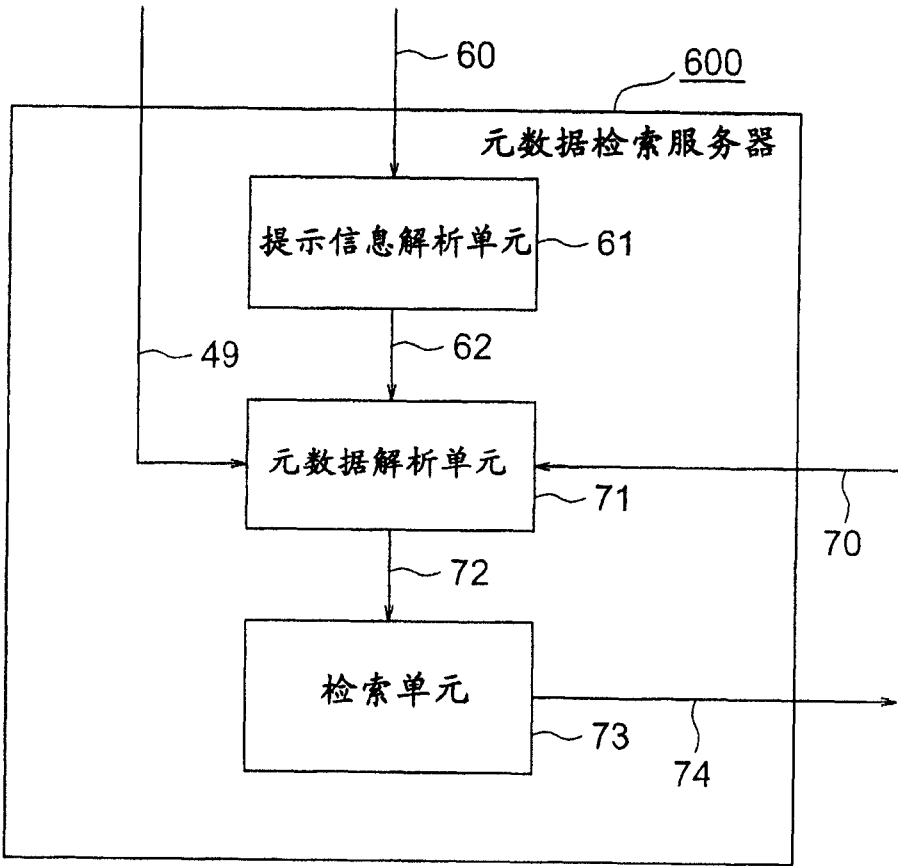


图 21

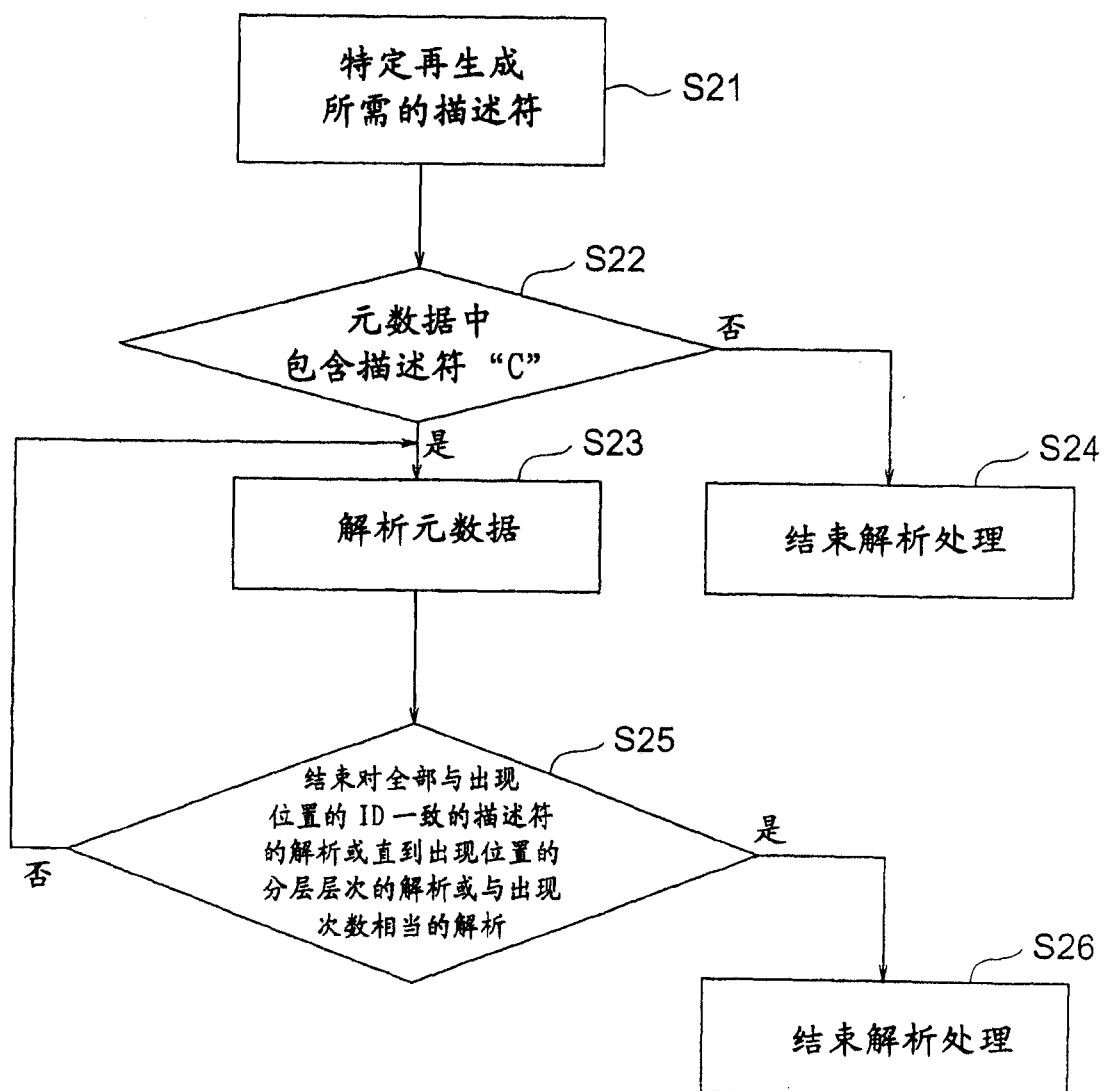


图 22

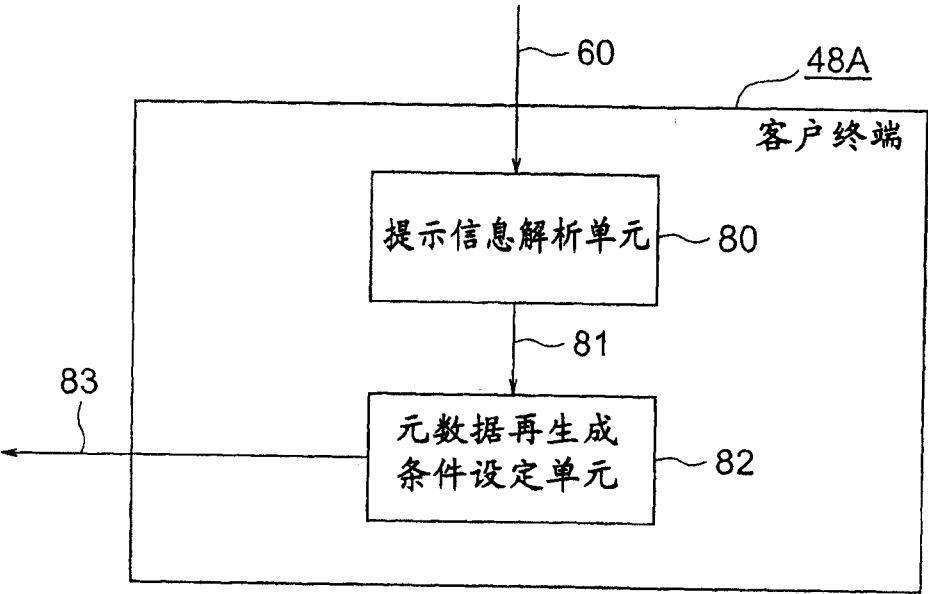


图 23

