

摘要

人脸的计算机模拟长期以来一直是计算机图形学以及人机交互中非常活跃的研究领域。随着计算机图形学在建模、渲染和动画等方面的发展，人脸表情动画在诸多领域得到了广泛的应用。因此，作为当前计算机视觉和计算机图形学领域的研究热点，人脸表情动画不但具有理论研究意义，同时也具有极大的应用价值。

基于 MPEG-4 标准的表情动画方面，本文主要说明了 MPEG-4 的概念和原理，以及基于此的表情动画制作。

关键字：表情动画，MPEG-4，计算机图形学

ABSTRACT

Computer simulation of human face has been a very active research field in computer graphics and human-computer interaction for a long time. With the development of computer graphics in modeling, rendering and animation, facial expression animation has been widely used in many fields. Therefore, as a research hotspot in the field of computer vision and computer graphics, facial expression animation not only has theoretical research significance, but also has great application value.

Based on the MPEG-4 standard for expression animation, this paper mainly explains the concept and principle of MPEG-4, and the expression animation based on it.

KeyWords: Emoticon animation, MPEG-4, Computer graphics

目录

第 1 章	引言	1
1.1	研究工作的背景与意义	1
1.2	国内外研究历史与现状	1
1.3	主要贡献与创新	2
1.4	结构安排	2
第 2 章	MPEG-4 标准	3
第 3 章	MPEG-4 原理	4
3.1	人脸定义参数 FDP	4
3.2	人脸动画参数 FAP	5
3.3	人脸动画参数单元 FAPU	6
第 4 章	基于 MPEG-4 的面部表情动画	7
第 5 章	总结与展望	8
致谢	9
参考文献	10

第1章 引言

1.1 研究工作的背景与意义

人脸面部表情的多样性和独特性一直是计算机视觉、图形学和模式识别等领域研究的热点问题。面部表情作为人脸的门户，是人们关注虚拟人物表现是否真实的重要来源之一。目前，虚拟动画人物的面部表情制作主要面临着以下几个问题：逼真性：动画人物的面部表情调节存在诸多问题，以至于调节出的表情生硬不自然，人脸表情动画的逼真性直接影响着该技术的推广应用；丰富性：面部表情的变化丰富多样，每个人物的表情也不尽相同。因此突破对面部表情多样化的、精致的微表情制作也成为行业持续关注和努力的方面之一，有效、更逼真地制作各种人脸表情动画的技术是人脸动画获得成功应用的关键所在；直观性：对人脸的面部表情产生的影响的诸多因素涵盖了多个方面，要综合考虑人体结构、解剖学、生物力学、心理学等等，因此控制一个表情的表现难度巨大，研究直观、方便地控制任意人脸表情动画是人脸表情合成动画研究的焦点；快速性：基于MAYA、3DMax 等软件制作三维人脸表情周期长，控制及调试时间长，因此研究简易的、方便地执行的、能够快速调节的任何面部表情的制作是动画动作研究领域的又一个焦点。

由于人脸面部表情十分丰富，喜、怒、哀、乐程度与深浅又不一样，每一种表情使得人脸模型的各个部位有不同程度的变化。人脸面部表情尽管十分丰富，但有必要对人脸各种表情的特征规律进行分析，从而找到各种表情与脸部五官之间的变化规律，并给出一般的形式化描述方法；当人脸产生表情时，脸部表面各点将发生变化，即产生位移，因此需要设计一种模型，用来计算面部各点的位移量，而且这种模型应该能准确地模拟人脸的面部表情；在确定了形变模型后，为使计算简单，还需要研究或选用一种好的计算方法，以提高系统效率。

1.2 国内外研究历史与现状

人脸表情动画是计算机通过调用预先设定的规则，使计算机生成的人脸图像从一种状态变形到另一种状态的过程，从而实现对人脸表情的模拟。目前已有的

方法是：

基于先验模型的人脸表情动画。利用生理学、心理学的研究成果，预先设定基本人脸表情动作单元，然后对人脸表情进行分析，将人脸表情分解为一系列动作单元的线性组合。合成动画时，先确定好始末时刻人脸表情，再利用变形规则，实现动作单元间的平稳过渡，最终合成人脸表情动画。

基于图像序列的人脸表情动画。先在特定人脸上的特征点上贴标签，然后利用摄像机采集此人各种表情变化的图像序列，接着依特征点在图像序列中的变化来修改人脸几何模型，并将改动信息作归一化处理后而建立人脸表情动作库。当需要产生其他人脸表情动画时，调用表情动作库中的相应数据来修改人脸模型，从而生成特定人脸表情动画的图像序列。

基于 2D 网格变形的人脸表情动画。以两幅（或更多）不同人脸的 2D 图像为基础，在起始图像上定义 2D 参数曲面（B6zier、B-Spline、NURBS 曲面，网格顶点越多越好），然后将此曲面覆盖在另外的图像上，拖动网格顶点，使某些顶点覆盖在相同的脸部特征上，完成网格变形；同时采用以时间为参数的插值方法，生成中间网格序列。

上述这些生成特定人脸表情动画的方法都采用了关键帧技术，即首先确定若干个关键的人脸表情画面，其它中间画面由计算机系统自动完成。

1.3 主要贡献与创新

本章主要研究基于 MPEG-4 标准生成人脸表情动画的技术。

1.4 结构安排

本文第一章介绍了面部表情动画的背景意义及历史研究情况；二三章分别介绍了 MPEG-4 的概念及原理；第四章说明了基于 MPEG-4 的面部表情动画。

第2章 MPEG-4 标准

MPEG-4 是一个基于对象的多媒体压缩标准,它允许独立的对视频上不同的可视对象 AVO (Audiovisual Objects) 进行编码。可视对象包括自然的或合成的视频对象,以及一些特殊的合成对象,例如人脸,身体,以及一般的 2D/3D 对象,它们可能是由矩形,球体,或索引面组成。在 MPEG-4 标准中,合成的可视对象是基于虚拟现实建模语言标准的,它使用例如 Transform 的节点,该节点定义了对对象的旋转,缩放和平移操作,以及 Indexed Face Set 节点,它通过索引面集合描述了对对象的 3D 形状。MPEG-4 的编码理念是: MPEG-4 标准同以前标准的最显著的差别在于它是采用基于对象的编码理念,即在编码时将一幅景物分成若干在时间和空间上相互联系的视频音频对象,分别编码后,再经过复用传输到接收端,然后再对不同的对象分别解码,从而组合成所需要的视频和音频。这样既方便对不同的对象采用不同的编码方法和表示方法,又有利于不同数据类型间的融合,并且这样也可以方便的实现对于各种对象的操作及编辑。例如,可以将一个卡通人物放在真实的场景中,或者将真人置于一个虚拟的演播室里,还可以在互联网上方便的实现交互,根据自己的需要有选择的组合各种视频音频以及图形文本对象。

MPEG-4 是世界上第一个基于对象的多媒体压缩标准[ISO/IEC1999],该标准可以对各种自然的或合成的音频/可视对象分别独立编码,将它们有机的集成在同一个场景中。由于人自身在多媒体中占据着十分重要的位置,而人脸又是人体最能够表现情感的部位,所以 MPEG-4 对三维人脸动画定义了国际标准。在这个标准中包括了 FDP 和 FAP 两套参数的定义。人脸动画系统的主要工作包括人脸建模和人脸动画,在基于 MPEG-4 的人脸动画系统中,FDP 同构造人脸模型密切相关,FAP 则是用来表示驱动人脸驱动的人脸运动数据。

人脸表情动画在人机交互、图像和视频编解码尤其是模型基人脸视频编解码以及人脸合成等领域有广泛的应用。MPEG-4 是目前图像和视频编码业界的一个通用标准,它使不同的人脸合成算法之间的互操作成为可能,并为此制定了用于表示、压缩和解释表情运动参数的标准方法。

第3章 MPEG-4 原理

MPEG-4 标准定义了一个中性状态下的面部模型，一系列的面部参数 **FP** 和面部动画参数 **FAP** 的集合，每一个 **FAP** 对应一个特殊的面部动作，用来变形中性状态下的面部模型。在某一时刻对应于一些 **FAP** 值下的面部模型变形可以生成一个动画序列，一个完整的表情动画就是一系列的动画序列。对特定的 **FAP**，给定的 **FAP** 值指示了对应动作的幅度大小，对于一个使用它的面部模型来解释 **FAP** 的 MPEG-4 终端，它必须预先定义动画规则来产生对于每个 **FAP** 的面部动作。终端或者使用它自己的动画规则，或者下载一个面部模型以及所对应的动画定义表 **FAT**。由于 **FAP** 被用来进行不同大小面部模型的动画，**FAP** 被定义以 **FAPU** 为单位，它和具体的面部模型无关，具有模型独立性。**FAPU** 是通过计算中性状态下指定的特征点之间的距离来得到的。

由于人脸动画具有广阔的应用领域，这使得在 MPEG-4 标准中增加了人脸动画标准。MPEG-4 所定义的人脸是一个合成的视觉对象，通过人脸动画参数来控制人脸对象的变形，可以在极低的数据率下获得理想的人脸动画效果。在 MPEG-4 中，人脸对象是用三维网格模型来表示的，它通过一系列的参数来刻画一个人脸。在 MPEG-4 标准中，定义了 84 个人脸特征点和 68 个人脸动画参数。

3.1 人脸定义参数 FDP

FDP 人脸定义参数，描述人脸的几何和纹理信息，有时还可以包括场景的信息，它提供了人脸特征点、网格、纹理、人脸动画定义表等数据的描述格式，有了这些数据就可以把一般人脸模型转化为特定人脸模型。在 MPEG-4 中定义了若干个组，在这些组内有序地定义了个特征点，它们描述了眉、眼、鼻、口、舌、齿等人脸主要特征部位的位置和形状，如图一所示。MPEG-4 中 **FDP** 所定义的特征点提供了很好的参考标准和范围，这些特征点都是经过大量的实验分析得到的，它们能够真实地反应出人脸表情的变化。**FDP** 用来定义接受端的人脸模型，具体可以通过修改通用模型或者发送一个新模型。这个模型通过 **FAP** 来驱动。

在一个 **FDP** 域中包含以下 5 个内容：**Feature Points Coord**，指定网格中所有特征点的坐标；**Texture Coords**，指定所有特征点在纹理上的坐标；**Use Ortho Texture**，

指定纹理的类型，如果 Use Ortho Texture 值为 FALSE，则纹理采用圆柱投影，如果 Use Ortho Texture 值为 TRUE，则纹理采用正投影。知道纹理类型，才可以正确地计算非特征点在纹理上的坐标；Face Def Tables，即人脸动画定义表，描述 FAP 对人脸网格变形的控制方式和参数；Face Scene Graph，包含一张纹理图像或者一个上色的人脸模型。

3.2 人脸动画参数 FAP

FAP 是 MPEG-4 标准中定义的一组人脸动画参数，它是基于对最小可感知动作（Minimal Perceptible Actions, MPA）的研究，MPA 和面部肌肉动作密切相关。FAP 定义了一整套基本的人脸动作，可以再现绝大多数自然的人脸表情和嘴唇运动。FAP 定义了一整套基本的人脸动作，可以再现绝大多数的人脸表情和唇动。FAP 中的前两个被称作高级 FAP，分别是唇形（Viseme）FAP 和表情（Expression）FAP。

MPEG-4 共有 68 个 FAP，每个 FAP 所对应的详细内容可以参考 MPEG-4 关于人脸动画的标准[ISO/IEC1998a]和[ISO/IEC1998b]。其中 FAP1 和 FAP2 被称作高级 FAP，分别是唇形（Viseme）FAP 和表情（Expression）FAP。对于唇形 FAP 来说，可以预先定义好一些基本的、不同的唇形，其他的唇形可以由这些基本的唇形线性组合而成；表情 FAP 也是一样的原理，可以用几种基本的表情线性组合出各种丰富的表情。除高级 FAP 外其他的普通 FAP 分别定义了人脸某一小区域的运动。两个高级 FAP 的作用是更准确、方便地表现一般的唇动和人脸表情，这些唇动和人脸表情用普通 FAP 也可以实现。但是对于复杂、不规则的唇动和表情，则只能用普通的 FAP 来实现。本文的研究对象是普通 FAP。3-68 号 FAP 为低级 FAP，每个 FAP 控制人脸某一区域的运动，在 MPEG-4 标准中定义了它的运动意义，运动方向等，通过一组 FAP 控制面部点的运动来实现网格的变形。

该方法可以对照 MPEG-4 中识别的人脸定义特征点，然后计算这些特征点的位移，最后由人脸表情库实时地组合出新的人脸表情。该方法能实时地和细致地生成具有真实感的人脸表情动画；基于抽象肌肉模型的人脸表情动画的实现过程，在程序实现过程中用到的人脸几何表示、抽象肌肉模型、一般人脸变化到特定人脸等技术。

人脸动画数据库有相应的对应表，主要由三部分内容构成：FAP 的值域分为

哪几段；FAP 控制哪些网格点；FAP 控制的网格点在每段中的运动因子是多少。对于任何一个 FAP，要根据数据库定义表查出三部分的具体内容，然后根据视频图像方面相应提供的算法，大概算出由相对应 FAP 影响的所有数据点的位移旋转和缩放。对于每一个 FAP，都计算出影响数据网的位移的多少，这些点的位移总和就会构建出一个生动的人物面部表情动画。

3.3 人脸动画参数单元 FAPU

FAP 的值都是以 FAPU 为单位的。以 FAPU 为单位的作用是使同样的 FAP 参数应用到不同的模型上，产生的是同样的唇动和表情，而不会因为模型的不同而使唇动和人脸表情走样，这样就使 FAP 具有了通用性。

为了使 FAP 可以应用到不同的模型上，并且能够产生同样的表情，而不使表情由于不同的模型产生走样，MPEG-4 定义了人脸动画参数单元 FAPU。FAP 与模型无关，而 FAPU 则是与模型相关的，FAP 的值是以 FAPU 为单位的，这样 FAP 就可以独立于具体的模型，从而具有通用性，这也是 MPEG-4 追求的一个目标。

FAPU 使同样的 FAP 参数应用到不同的模型上，产生的是同样的唇动和表情，这样就使 FAP 具有了通用性。

第4章 基于 MPEG-4 的面部表情动画

生成人脸表情动画需要三个过程：设计一个人脸通用三维网格模型。人脸通用网格模型的优点是网格模型的拓扑结构总是相同的，模型的驱动与主体对象无关，不同的主体对象之间的唯一不同是一次性传输的 FDP 值，所以一旦定义好了人脸通用网格模型的驱动方法，便可应用于任意一个主体对象。模型的网格点数增多，当然能很好地表达人脸特征，但这将造成模型难以控制、计算量加大，给动画的实时性造成影响，所以模型的最终设计需要在表达力和计算量之间做很好的折中；人脸特定化通用三维网格模型；设计一个计算模型来生成人脸表情动画。

通过手工建模和手工设定 FDP 参数的方法，可以得到兼容于 MPEG-4 的三维虚拟人脸模型而通过采集动作视频并对人脸标记点进行运动分析提取，可以得到来自于真实人脸的 FAP 动作数据。在合成阶段，通过在虚拟人脸模型上利用数据收集得到的 FAP 参数进行驱动，根据 MPEG-4 的人脸动画模式，可以合成渲染出最终的动画序列结果。MPEG-4 标准中，定义了一些基本表情——高兴、悲伤、愤怒、恐惧、厌恶、惊讶。理论上所有的面部表情都可以由这些基本的表情合成出来，这样就可以方便实际的应用。面部表情可以用人脸动画参数的集合来表示，不同的组合可以得到不同的表情。六种基本人脸表情可以由面部局部区域运动组成，而这些面部局部区域的运动可以分解成具体的集合来表示。根据 MPEG-4 定义的人脸特征点，在三维人脸网格模型上选择用作 FDP 的网格顶点，及该网格顶点所影响的一组与其相邻的网格顶点。用户要按照 MPEG-4 定义的人脸定义参数选择与 FDP 对应的网格顶点，并根据各 FDP 点在人脸部的的位置以及与该 FDP 临近的 FDP 的位置，选择受该 FDP 影响的一组网格顶点。MPEG-4 的 FAP 定义表定义了各 FAP 影响的 FDP 以及在该 FAP 的驱动下，FDP 运动的正方向。该表中定义的正方向作为各 FDP 在 FAP 驱动下运动的正方向，每个 FDP 影响的一组网格顶点的运动方向与该 FDP 的运动方向相同。当 FDP 受到 FAP 的驱动时，FDP 影响的各网格顶点随控制该组顶点的 FDP 顶点的运动而运动。在人脸动画的过程中，为了保证自然真实的人脸表情，必须保持人脸几何模型的平滑性。当受 FDP 影响的网格顶点随 FDP 运动时，为了保证人脸的平滑性，离 FDP 网格顶点距离近的点比远的点随 FDP 运动的速度快，离 FDP 越远，运动速度越慢。

第5章 总结与展望

MPEG-4 人脸运动参数的获取在基于 MPEG-4 的人脸图像编码传输中具有极为重要的作用。本文主要研究了基于 MPEG-4 的人脸表情动画的相关内容，阐述了 MPEG-4 的基本内容与原理，给出了面部表情动画的要点。

本文的研究工作中仍存在诸多不足之处，需要进一步改进和完善。如改进表情动画的算法。考虑人脸的对称性，改进运动方向及网络格点的计算。

致谢

感谢蔡洪斌老师认真负责的讲授计算机三维动画技术的相关知识，在课堂上表述准确，思路清晰，逻辑性强，教学内容重点突出，难点清晰，有一定的深度和广度。

参考文献

- [1] 於俊,汪增福.一种基于 MPEG-4 的三维人脸表情动画算法[J].中国科学技术大学学报,2011,41 (4) .
- [2] 姜大龙,王兆其,高文.基于 MPEG-4 的三维人脸动画实现方法[J].系统仿真学报,2001 (S2) :493-496.
- [3] 商圣贺.三维人脸建模与表情动画技术研究[D].哈尔滨工业大学.
- [4] 罗恒希.游戏引擎中三维人脸造型与表情动画的研究与实现[D].电子科技大学.