三维人脸扫描图像数据预处理

孔德慧，程世铨，尹宝才，胡永利，谷春亮
（北京工业大学 计算机学院 多媒体与智能软件技术实验室，北京 100022）

摘 要：为构造基于知识学习的形变模型所需的规格化三维人脸库，提出一种纹理映射图数据结构及以点为处理单元的分割三维人脸的方法. 该结构将原始数据隐含的信息显性化，为激光扫描仪采集的真实三维人脸中分离的采样点与纹理 2 部分数据建立双向映射关系，以保证数据操作和处理达到实时与同步. 基于此结构的快速映射特点，提出的分割方法以二维纹理点为操作对象能够快速处理大量原始数据.

关键词：三维人脸；采样点；纹理点；双向映射；纹理映射图；分割
中图分类号：TP 391 **文献标识码**：A **文章编号**：0254－0037(2005)03－0323－05

Parke[1] 在 1974 年提出了第 1 个参数化的人脸模型. 20 世纪 90 年代后期 Guenter[2]、Pighin[3] 等提出了基于一般几何模型变形的人脸建模方法. 此类方法获得的特定人脸三维模型真实感不足，方法复杂. Blanz 等[4] 开辟了新方向，借助激光扫描仪获得真实三维人脸数据，其实验结果具有高度真实感. 在此基础上，Blanz、Romdhani、Huang 等[5-7] 在三维人脸识别、鉴定和动画方面的研究取得了很大进展. 该方法异于传统方法：1）采样点密集、精确，具有真实纹理；2）以规格化的三维人脸库作为研究基础. 扫描仪获取的原始数据，由分离的三维几何采样点和纹理 2 部分组成，除研究所需的人脸部位外，还包括对研究无用的部分，须预处理割除这些部分. 预处理中，首先是采样点与纹理点的同步问题，2 部分数据须建立双向映射关系；其次是分割的问题. 针对密集网格的分割，Krishnamurthy 等[8] 提出了一种以网格作为分割基础的方法，但算法复杂. 作者针对扫描数据的特点设计了一种纹理映射图结构，弥补了原有的单向映射关系的不足；并借助扫描数据的分离存储特点，提出以点为处理基础的分割人脸网格与纹理图像的方法.

1 三维人脸数据结构及存储方式

本文使用的三维人脸数据由 CyberWare 的 3030RGB/PS 激光扫描仪获取，通过 1 次柱面扫描采集到人头部精确的空间几何信息和彩色纹理信息. 空间几何信息主体为密集的三维空间几何采样点（下称采样点），用坐标（X，Y，Z）表示，约 20 万；另有描述采样点连接关系的三角网格，约 40 万. 彩色纹理信息包含每个采样点对应的 1 个 24 位彩色（用 R、G、B 表示）纹理像素点（下称纹理点）. 这 2 部分存储在 2 个分离的文件结构中，存储空间信息的文件除采样点与网格（见图 1、图 2），还有采样点在纹理文件中对应的纹理点的归一化坐标. 归一化坐标表明本采样点在纹理文件中对应纹理点位置的索引信息. 彩色纹理信息是采样点纹理柱面投影得到的二维图像，以普通图像格式存储，图像长宽由投影参数、扫描设备硬件与操作平台决定. 本文得到的纹理分辨率为 478×489，且头颈横向，不符合人类视觉的习惯（见图 3）. 纹理图像中像素数是固定的（为长宽的积），而采样点数量不定（由扫描时人的头部大小和其在扫描场中的具体位置决定），其纹理投影到二维图像时不能恰好将所有矩形上的点填充，必然导致在纹理图像中某些区域没有投影点，因此该图像中的某些像素点在三维中没有对应的采样点. 称有对应关系的像素点为有效纹理点，其他的称为无效纹理点.

收稿日期：2003-07-09.
基金项目：国家“九七三”计划资助项目（2001CCA03300）；国家自然基金资助项目（60375007）；北京市自然科学基金资助项目（D070601-01）；北京市科学技术委员会资助项目（N070601-01）；北京市教育委员会资助项目（P070701-01，P070702-01）.
作者简介：孔德慧（1968－），女，陕西吴堡人，副教授，博士.

图 1 三维采样人脸
Fig.1 3D sampling face

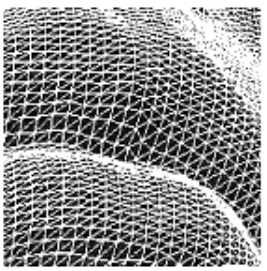

图 2 唇部网格放大图
Fig.2 Magnified mesh of lip

图 3 对应的二维纹理图像
Fig.3 2D correspondent texture

2 纹理映射图

2.1 采样点与纹理点的互定位

在三维人脸的研究中,对数据的处理一般以点为基本单位,如三维数据的旋转、平移等变换需要对每个点进行操作,因此对每个点进行搜索和定位不可避免. 当对数据进行简化、分割等操作时,还需要考虑点和网格的重新组织. 无论使用何种处理,通常都要考虑到采样点和纹理图像中对应的纹理点的同步操作问题. 当给定一个采样点时,寻找对应纹理点的唯一途径是纹理点的归一化坐标,由这个坐标在对应的纹理图像中可以直接找到该点对应的纹理点,时间消耗是 $O(1)$;但是,二维纹理图像中给定一个纹理点时,却无法以 $O(1)$的效率得到它对应的采样点,因为在二维的纹理图像文件中只存储了纹理点的值,并未指明每个纹理点对应的采样点,必须根据给定的纹理点的坐标,全局遍历采样点的存储空间,对比每个采样点所对应的纹理归一化坐标,只有当 X、Y 坐标完全一致时,该采样点才是给定纹理点对应的采样点. 描述这种思想的算法伪代码如下:

```
FOR i := 0 TO points_number    // points_number 采样点数目
BEGIN
    int tx := 四舍五入取整(point[i].vt.x * width);  // width 纹理宽
    int ty := 四舍五入取整((1 - point[i].vt.y) * height);  // height 纹理高
                        // point[i].vt.x/y 为当前采样点的纹理点归一化坐标
    IF (tx = x && ty = y)   // x,y 是图像中给定纹理点的坐标
    RETURN i;
END;
```

用该算法定位 1 个采样点,最快的情况是第 1 次对比即是所求,最慢则遍历至最后一个,因此该算法的时间复杂度是 $O(n)$,定位 1 个采样点的平均查找比较次数为 $n/2$. 此算法在不需实时响应、只需查询少数纹理点对应采样点的情况下是可行的. 现假设采样点有 20 万个,在寻找每个纹理点的对应采样点时,搜索定位 1 个采样点需要平均约 10 万次的查找比较运算,20 万个点都搜索定位,则需要 200 亿次(10 万 × 20 万),即便在不需要实时处理的情况下,这种速度也无法忍受,给很多相关研究和应用带来不便.

2.2 纹理映射图

导致上述算法缓慢的原因是采样点与对应的纹理点的分离存储,这种存储结构只提供了一个单向的直接从采样点到纹理点的定位关系,即采样点中的纹理点归一化坐标(正向映射),而没有提供直接从纹理点定位采样点关系的功能. 因为纹理文件没有相应的采样点索引信息(反向映射),所以数据的相互关系不是一个显式的双向映射关系. 因此,解决这个问题的关键是把正向映射中隐含的关联信息显性化,以此建立反向映射关系. 为此,作者设计了一个附加的纹理映射图结构来解决这个问题,这种方法只需附加一个纹理映射图的存储空间,和一次全局遍历采样点的时间,即能得到反向映射关系,一次性解决数据之间

的映射关系问题(见图 4).

纹理映射图是一个与纹理图像有同样维数的二维数组，数组中每对二维下标所指示的存储空间存储了纹理图像中相同二维下标位置的有效纹理点所对应的采样点的序号(指针). 假设纹理图中 $X = 256$, $Y = 256$ 处是鼻子顶点，则映射图中下标[256][256]处存储鼻子顶点所对应的采样点在文件中的排列序号，从 0 开始计数. 而其他无效纹理点没有对应的采样点的映射图区域则以 -1 填充. 纹理映射图结构是一个反向映射关系，可以快速解决反向定位问题. 若给定纹理图像中的任意纹理点，可以立即找到该纹理点对应的采样点在采样点集合中的确切位置，从而摒弃了上文算法中顺次地对比查找采样点的缺点，其时间复杂度是 $O(1)$. 用 20 万点定位的例子做对比，利用该图，所有点的定位仅需 20 万次的查找. 映射图算法伪代码如下：

```
init map with -1;        //初始化纹理映射图 map
FOR i := 0 TO points_number
BEGIN
    int tx := 四舍五入取整(point[i].vt.x * width);
    int ty := 四舍五入取整((1 - point[i].vt.y) * height);
    map[ty][tx] := i;
END;
```

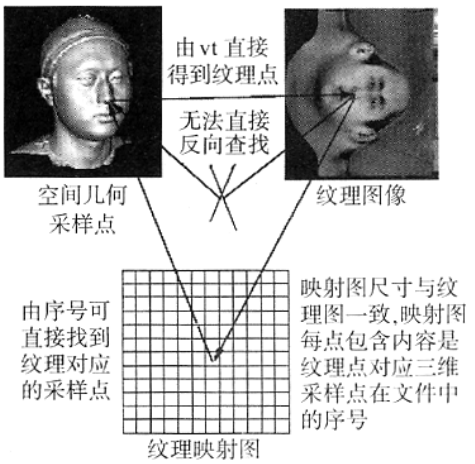

图 4　采样点、纹理及纹理映射图的相互关系

Fig.4　Relationships of sampling points, texture and texture map

3　分割人脸的方法及实例

针对密集网格数据，文献[8]提出了一种人工交互的分割方法，由用户选取一序列点，采用贪心图算法，在网格连线上寻找相临点的最短路径，这些路径则形成分片的边界. 该方法以网格的连接关系为基础进行分割操作，分割的边界由各分片共享，因此可以保证网格分离后没有点和网格的损失，但算法及实现复杂. 考虑此处分割人脸只需将有效人脸部分分离出来，不需保证每个点与网格无损失，可以考虑以点为基础进行分割操作，利用纹理图像在二维空间的便利，不需考虑点之间的连接问题，简化了算法. 分割过程需要解决 4 个问题：①分离边界的确定；②数据初步分离；③初步分离后的数据筛选；④数据重组织.

3.1　边界的确定和初步分离

对于三维的人脸数据，如果采用三维交互，不易实现，因为在图形学中三维交互任务涉及定位、选择和旋转. 用户难以区分屏幕上鼠标选择到的对象的深度值和其他显示对象的深度值[9]. 二维的交互相对容易，鉴于扫描数据包含了三维的采样点和对应的二维纹理图，在纹理图上进行二维交互操作易于实现. 利用纹理映射图，可以快速得到对应的采样点，以点为处理的基本粒度，不考虑网格连接关系，能够大幅度降低交互难度，并简化处理流程.

本文采用文献[4]提出的分离模式，即将前额头发、耳朵以后、脖子以下切除. 为便于处理前额头发及扫描，在扫描时，被扫描者带一顶泳帽将前额的头发拢入，因此分离边界包括泳帽边缘、耳后垂面和脖子下平面. 后 2 者只需交互标定耳朵后 2 点作为竖切面的标准基点，脖子下一点作为横切面的标准基点. 比较复杂的是泳帽边缘的确定，泳帽边缘相对来说是不规则的，且点数众多，因此手动逐点标定不可行，但采用文献[8]中标定一系列点连接的方法需要考虑三维网格的连接关系，比较复杂，且连线不能将泳帽边沿精确描绘. 解决的方法是，在扫描时使用颜色与肤色反差大的泳帽，由此获得的纹理图像泳帽边缘与人脸灰度值差别较大，使用图像处理中常用的边缘检测[10]的方法能够较好地描绘该边缘；但边缘检测方法不能对

所有采集的纹理的帽檐都有一个完整、精确的检测结果,因此借助一定的人工交互来纠正未检测到的和检测错的边缘细节.

在二维纹理图像上确定所有的边界点后,根据纹理映射图直接定位其对应的采样点,由此得到三维空间上的切面和帽檐边界,再以所有采样点为处理对象,判断其与边界的位置关系,不考虑网格关系,将在此边界内的采样点数据初步分离.

3.2 数据的筛选与重新组织

初步分离后生成了一个粗糙的人脸采样点集合,因为,以点为处理对象的方法,舍弃网格的约束关系,致使分离出的采样点边界上某些点还存在,但其原始的网格关系所连接的点却未保留全,导致该点失去邻接网格关系,成为没有存在意义的孤点,需要再一次筛选去除. 如图 5 所示,无连线的点是第 1 次筛掉的点,A 点还在,但它的右方、上方、左方的点已被去除,A 点实际成为一个没有邻接网格的孤点.

二次筛选后,采样点坐标未变化,但其排列顺序与网格关系已改变,需重新组织该数据,依据原网格的连接关系,将采样点的序号(指针)更新,生成最终分割的人脸网格. 对应的纹理图像需要同样的处理,将对采样点的处理映射到二维纹理上,剔除不再包含在更新后人脸内的纹理点. 由于扫描仪及操作平台的原因,采集的纹理图像的头颈是横向的,与视觉习惯不符合,为此,作者在人脸分离后将纹理扭转过来,并平移使之以鼻子顶点作为新纹理图像的中心点. 根据文献[9]中的旋转公式

$$u = x\cos\theta + y\sin\theta, \quad v = -x\sin\theta + y\cos\theta$$

旋转纹理. 因纹理图像的坐标全部为非负数,使用该公式变换会出现负值,这里将它修改并结合平移公式使之符合纹理坐标的要求,设鼻子顶点坐标为 X_{nose}, Y_{nose},原纹理长、宽为 H_{old}、W_{old},新纹理长、宽为 H_{new}、W_{new},且旋转角度为逆时针 90°,则新公式为

$$u = y + (W_{\text{new}}/2 - Y_{\text{nose}})$$
$$v = -X + (W_{\text{old}} - 1) + \{H_{\text{new}}/2 - [-X_{\text{nose}} + (W_{\text{old}} - 1)]\}$$

经过该公式变换,纹理点的坐标改变,因此采样点中原有的纹理点归一化坐标不再有效,需要使用该公式同步变换,得到新的纹理点归一化坐标. 最后,在对数据所有的处理和重组织完成后,根据数据使用的需要,可以采用平面中两不平行的边矢量叉乘的方法[9]来求三角面片(网格)的法向和点的法向. 图 1 ~ 图 3 中数据的分割结果如图 6 所示.

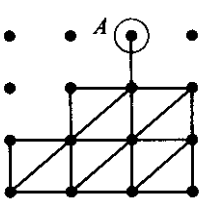

图 5 A 点成为孤点的示意图

Fig.5 Sketch of an outlier point A

(a) (b)

图 6 分割结果

Fig.6 The results of segmentation

4 结束语

作者针对三维扫描数据设计的纹理映射图结构,从本质上提高了三维数据的采样点与纹理点的互定位速度,且该结构简单,便于构造和使用. 分割人脸的方法以点为处理对象,抛弃网格约束分割人脸的方法,有效避免了对网格操作的复杂性 并对分割此种数据常见的 4 个问题进行了分析,提出了有效的解决方案,这些问题不仅在分割人脸中会遇到,在其他处理中也可能会涉及到(处理过程导致点数的增、减,纹理的变换时). 此类三维人脸的研究,海量数据处理不可避免,因此作者提出的纹理映射图以及对数据分

割、再筛选和再组织的方法为这些处理提供了一个可行的解决方案，为建造三维人脸库提供了快速可靠的数据预处理方法. 但在一些需要精确标定特征点的应用中，仅通过在二维图像上标定特征点略显不足，因为当特征点位于三维图像上梯度变化剧烈的区域时，在二维纹理图像中的手工标定容易造成在三维对应位置的较大误差，因此，在今后的应用中，可以结合二维纹理和三维形状来标定特征点，在二维图像上标定，同时在三维上显示标定效果，再根据具体情况进行改正. 在某些分割应用中，需要保证点与网格的无损处理，可以使用文献[8]中的思想，以原有的网格线约束边界线的产生.

参考文献：

[1] PARKE F I. A parametric model of human faces[D]：[dissertation]. Salt Lake City：University of Utah，1974.

[2] GUENTER B，GRIMM C，WOOD D，et al. Making faces[A]. SIGGRAPH '98[C]. Westmont：ACM Press，1998. 55-66.

[3] PIGHIN F，HECKER J，LISCHINSKI D，et al. Synthesizing realistic facial expressions from photographs[A]. SIGGRAPH '98[C]. Westmont：ACM Press，1998. 75-84.

[4] BLANZ V，VETTER T. A morphable model for the synthesis of 3D faces[A]. SIGGRAPH '99[C]. Los Angeles：ACM Press，1999. 187-194.

[5] BLANZ V，BASSO C，POGGIO T，et al. Reanimating faces in images and video[A]. Proceedings of EUROGRAPHICS 2003[C]. Granada：Blackwell，2003. 641-650.

[6] ROMDHANI S，BLANZ V，VETTER T. Face identification by fitting a 3D morphable model using linear shape and texture error functions[A]. Computer Vision-ECCV 2002[C]. Heidelberg：Springer-Verlag，2002. 3-19.

[7] HUANG J，BLANZ V，HEISELE B. Face recognition using component-based SVM classification and morphable models[A]. Proceedings of the First International Workshop on Pattern Recognition with Support Vector Machines[C]. Heidelberg：Springer-Verlag，2002. 334-341.

[8] KRISHNAMURTHY V，LEVORY M. Fitting smooth surfaces to dense polygon meshes[A]. SIGGRAPH '96[C]. New Orleans：ACM Press，1996. 313-324.

[9] 孙家广. 计算机图形学[M]. 第 3 版. 北京：清华大学出版社，1998.
SUN Jia-guang. Computer Graphics[M]. 3rd ed. Beijing：Tsinghua University Press，1998.(in Chinese)

[10] CANNY J F. A computational approach to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，1986，8(6)：679-698.

Preprocessing of 3D Scanning Face Image Data

KONG De-hui，CHENG Shi-quan，YIN Bao-cai，HU Yong-li，GU Chun-liang

(Multimedia and Intelligent Software Technology Lab，College of Computer Science，Beijing University of Technology，Beijing 100022，China)

Abstract：In order to construct normalized 3D face database which knowledge-based morphable model requires，a texture map based on data structure and a point-based segmentation method for 3D face are presented. This structure represents explicitly the implied information of original data，and constructs bidirectional-mapping between the separated sampling points and texture of the real 3D faces scanned by laser scanner. So it can ensure that the two parts of data are processed synchronously and real-time. Using rapid mapping feature of this data structure，the presented segmentation method can process plenty of original data rapidly.

Key words：3D face；sampling point；texture point；bidirectional map；texture map；segment