

硕 士 研 究 生 读 书 报 告



题目 非参数后投影表面上的触摸感应： 用于实体医疗培训的物理虚拟头

作者姓名 郭志强＿

作者学号 21651160 ＿

指导教师 李启雷 ＿

学科专业 移动互联网与游戏开发技术 ＿

所在学院 软件学院

提交日期 二○一七 年四月

Touch Sensing on Non-Parametric Rear-Projection Surfaces:  
A Physical-Virtual Head for Hands-On Healthcare Training

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: QiLei Li

By

Zhiqiang Guo

Zhejiang University, P.R. China

2017

摘要

我们展示了一种可扩展的方法，用于使用后投影动画3D面对人头像表面进行统一的多点触摸检测和响应。该方法有助于实际操作具有动态物理虚拟患者行为的触敏训练，它需要一个或多个红外（IR）照相机，一个或多个投影仪，红外光源和后投影表面来推广到其他非参数后投影表面。用人造手指反射的红外光会被具有相匹配的IR通滤波器的相机捕获，它允许多点触摸事件定位。渲染系统将这些事件紧密结合在一起，通过使用投影机上的脸部动态画面产生听觉和视觉反应，从而产生响应式的交互式体验。我们说明了我们的物理原型在医疗训练场景中的适用性。

**关键词**： 人造、增强、虚拟现实

Abstract

We demonstrate a generalizable method for unified multitouch detection and response on a human head-shaped surface with a rear-projection animated 3D face. The method helps achieve hands-on touch-sensitive training with dynamic physical-virtual patient behavior. The method, which is generalizable to other non-parametric rear-projection surfaces, requires one or more infrared (IR) cameras, one or more projectors, IR light sources, and a rear-projection surface. IR light reflected off of human fingers is captured by cameras with matched IR pass filters, allowing for the localization of multiple finger touch events. These events are tightly coupled with the rendering system to produce auditory and visual responses on the animated face displayed using the projector(s), resulting in a responsive, interactive experience. We illustrate the applicability of our physical prototype in a medical training scenario.

**Keywords：**Artificial, Augmented, Virtual Realities

# 1 简介

交互式计算机图形的触摸感应通常在可以通过数学参数化的表面上实现

分析方程，如平面/平面或球体。我们提出了基于带后投影动画内容的非参数曲面的一种统一的多点触摸检测和响应方法，类似于[3]，我们的方法包括红外（IR）相机、投影仪、红外光源和后投影面。外灯从表面的下面通过，从人的手指反射回表面，并被摄像机（表面下方）捕获与之匹配的IR通滤波器。这允许在表面上定位多个手指触摸事件，触摸事件与渲染系统紧密耦合以产生一个高度响应的互动。

虽然该方法可以推广到其他非参数曲面，但我们已经说明了我们的物理原型在使用人头部表面的医疗训练场景中的适用性。在表面上触摸事件提示图形和音频反应，例如，参与者可以通过用手指（图1）拨开嘴唇打开患者的嘴巴，以及通过某些动作让病人说话。另外，我们设计了包含这些功能的医疗培训场景。我们的方法与现有方法相辅相成：与病人扮演者和机器人模特儿一样，我们的方法提供了直接的手势范例，无需头戴式显示器（HWD）或分开的虚拟内容显示。像虚拟病人一样，我们能通过被动的（只是看）或交互式（触摸）的视觉外观诊断来呈现异常体征。

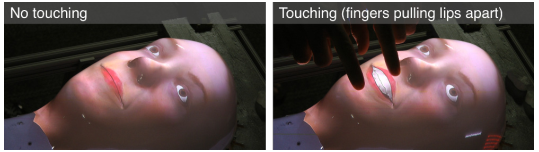


图1我们医疗培训中实际操作的触摸感应物理虚拟头，原型包括带有数字投影仪的半透明头饰外壳、红外光源和下面的照相机。左：未触摸；右：一个正在诊断可能中风的病人的护士使用她的手指拉开嘴唇，检查牙龈。

在第2节中，我们简要介绍现有的触摸非参数表面的感测和投影技术。我们所提出的方法依赖于各个坐标空间中关联点的查找表构造，这在第3节中有详细描述。我们在第4节中提供了我们的原型的描述，其适用于第5节中的中风评估情景。最后，我们将在第6节讨论此方法的潜在扩展并在第7节中提供简要摘要。

# 2 相关工作

存在的用于各种表面上的触摸感测的多种方法中，电容感测方法[6,8,17]是受欢迎的，它们通常假定平坦的表面。导电材料可以模制以适应更复杂的表面[17]，然而与复杂性相比，电容元件的密度必须足够高（空间频率）。无论怎样，电容方法将会干扰投影图像。

计算机视觉方法使用可见光或IR光来检测触摸事件，一类方法使用受抑全内反射[9,19,21,22]，但这种方法无法处理具有任意曲率和不连续性的曲面。 与表面相互作用的手和手指通过材料反射IR光; 这种灯可以直接用于触摸感应[2,3,13,24]。以这种方式，在平面和参数非平面表面上进行了触摸感测，如球体[2,3]。当把这种技术扩展到任意表面时，在检测到的触摸事件和表面上的3D坐标之间找到适当的映射可能具有挑战性。此外，可能很难提供均匀或平滑变化的表面上的红外照明。

已经使用线性方法、非线性方法、参数平面上的几何注册技术和分段线性方法[12]，分片线性方法也被应用于参数非平面显示器。对于非参数显示表面，它是一种涉及找到投影图像和观看者的眼睛之间的映射的方法，使用位于观看者的预期位置的相机来表示。

我们这里的工作是由专门的医学培训驱动，如越来越多的希望充分利用虚拟和物理的一部分研究人员开发了物理虚拟患者（化身和药剂）[11,23,18]。Kotranza and Lok[10]探讨了一种针对真实乳房检查的混合现实（MR）范式，受训者站在一个倾斜的物理人体模型，包括物理乳房模拟器——橡胶乳房，它提供乳房皮肤、组织和底层乳房肿块的触感，并含有十二个压力传感器来检测用户的触摸。通过HWD，受训者看到一个覆盖在人体模型和乳房模拟器上的虚拟患者，通过相机跟踪人体模型的物理衣物，持续呈现动态的面部和衣物Chuah等人[5]已经对现有研究进行了调查，并对体现对话代理的物理意识形式化。在这里，我们提出一种基于计算机视觉的方法，使用红外光在后投影表面上实现紧密耦合的触摸感应和渲染能力。这种方法将支持涉及触摸的各种物理虚拟医疗训练场景，包括诊断、治疗和舒适触摸。

# 3 坐标空间

在传统的触摸系统中，输入装置（触摸）和输出装置（显示器）的坐标空间之间的几何关系通常可以通过诸如6D刚性变换或单应性的参数函数建模。 我们的情况在两个方面有所不同：对应于物理表面的3D触摸和3D图形空间是离散采样的，空间之间的关系不能用参数函数来描述。因此，我们使用查找表直接链接坐标空间，这在时间上具有解耦渲染和触摸的附加益处，使得可以独立地（以比触摸事件更高的速率）执行渲染事件。总共有四种类型的坐标空间：对应于相机的2D空间、对应于投影机的2D空间以及对应于触摸感应和渲染的3D空间3D物理表面。接下来我们来描述这四个空间。

CAM2：在2D相机空间（CAM2）中，使用触摸感应系统（4.2.1节）触摸，事件被检测和定位为2D（x，y）坐标，能够将触摸事件成像的摄像机记录其在坐标空间内的位置。

PRO2：在渲染系统中，检测到的触摸事件必须触发适当的图形响应，需要一个在2D投影机空间（PRO2）中相机坐标与2D（u，v）投影仪坐标之间的映射。 为了获得这些映射，我们在（u，v）投影仪坐标的网格上投影半径为13像素的白色圆圈，每个维度上的像素间距为11像素，黑色的背景，每个红外摄像机分割圆圈并记录其重心。我们总共捕获了11706个这样的圈子，这提供了投影仪像素和所有相机中的像素之间的直接映射。映射的分辨率受到预测的圈数的影响; 我们在CAM2到PRO2通信上执行三次插值，以提供足够的分辨率。

TCH3：在典型的多视图几何场景中，有些特征匹配形式提供了不同视角的对应关系。由于我们的后投影面均匀、颜色流畅，展现的功能很少，所以很难通过特征匹配方法找到对应关系。但我们可以使用上述每个预想圈子作为手动“特征”，将其在每个摄像机图像中的检测位置视为对应点。这些对应关系在三维空间中被三角测量，导致稀疏3D触摸空间（TCH3）中的3D（X，Y，Z）顶点，这与通过校准获得的相机的3D位置一起（在4.1节中描述）提供了整个设置的三维模型，如图4所示。

GFX3：当TCH3是表面的三维三角点云时，提供由相机成像的点和表面上的3D位置之间的对应关系，其分辨率受到投影圆圈数和成像分辨率相机的限制。因为渲染和预测的效果需要要相对密集，我们创建一个单独的可渲染3D模型作为密集3D中的顶点（X'，Y'，Z'）的集合图形空间（GFX3），如第4.1节所述。我们使用TCH3作为CAM2中的相机像素和GFX3中的3D模型之间的中介。具体来说，我们使用迭代最近点算法（Per Bergstrom [4]的实现），估计将GFX3中的表面的高分辨率3D模型与TCH3中的三角形点云的空间相关联的旋转和平移矩阵。从这里，我们对每个相机图像平面（CAM2）中的所有2D像素进行反投影，并在模型（GFX3）上找到相应的3D点。当在特定相机内检测到触摸时，图形模型上相关联的3D位置可用作定时查找。

# 4 原型

我们在3D物理虚拟人类头部的非参数曲面上调整触摸检测和图形响应的方法，我们的原型钻机（图3）包含四个具有红外滤光片（780nm）的点灰色，两个红外照明器（850nm）和一个AAXA P300微型投影机的Blackfly单色相机。我们用具有可移动红外滤光片的单色相机，以便相机可以捕获出现在可见光谱中的表面上的投影图像，在查找表中构建CAM2至PRO2对应关系。一旦该预处理阶段完成，IR滤镜被重新插入，以便相机可以检测IR光谱中的触摸事件。

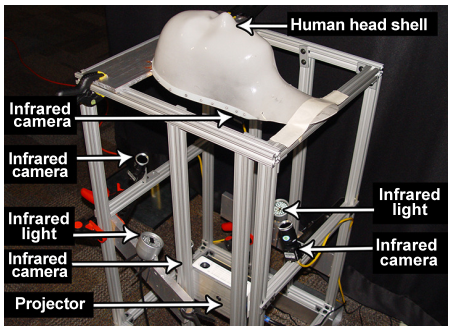


图3 在我们的原型机中，我们使用了四台红外摄像机，两台红外光源和一台投影机。

## 4.1 设置

相机、投影机和红外灯放置：我们将相机和投影机放置在允许足够视觉覆盖头部表面的位置（参见图4）。同样，我们定位红外照明器，为触摸检测提供充足的照明。

相机校准：我们使用OpenCV中的标准棋盘校准方法校准IR摄像机。每台摄像机单独校准；它们的计算内在特性被用作所有成像相机之间随后的立体相机校准的初始估计。最后，我们获得所有相机之间的相对位置信息。为了简化操作，我们将一台摄像机作为稀疏3D触摸空间的起源。

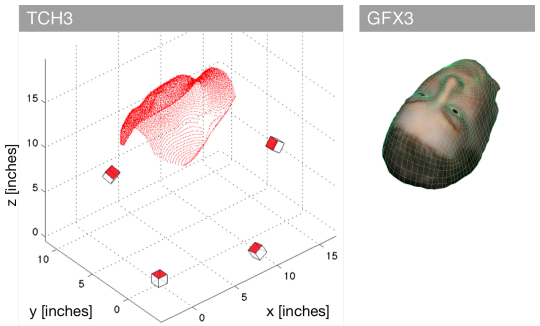


图4 我们的物理设置的稀疏3D触摸空间（TCH3）将触摸事件链接到密集的3D图形空间（GFX3），如右图所示进行比较。红外摄像机显示为立方体;每个立方体的彩色面代表该特定摄像机的方向。

密集的3D信息：使用摄影测量技术，从其周围的所有角度的重叠照片生成表面的密集扫描。为了提供合适的特征，将彩色棋盘投影到表面上（图5b）。这大大提高了扫描的精度，因为外壳是均匀的颜色（图5a）。照片被发送到Autodesk 123D Catch，一个免费的软件包，将一个对象的照片集合转换为一个密集的3D网格[1]。使用钻机的物理测量清洁、缩放和对准所得网。通过123D Catch获得的模型是非常密集的，不容易纹理化或动画化（图5c）。基于密集网格，我们生成一个较低密度的，使用具有边缘流拓扑的四边形而不是三角形，这更适合于动画，因为使用四边形可以最小化平滑过程中由细分产生的假象。四边形的性质使它们成为这里使用的边缘循环建模的基础（图5d）[15]。极点是具有三个或五个或更多边缘的顶点;它们位于特定位置以减少动画期间的伪影。对于我们的头壳，通过使用Maya [14]合并顶点，创建眼球和内口几何并附加到网格上。使用一个展开图像对网格进行纹理化，然后使用骨骼和混合形状的组合对各种动画进行装配。

将动画3D模型导出到Unity3D [20]，并按照制造商手册中的描述将其放置在场景大小与投影机最大分辨率（1920×1080像素）相匹配的场景中。使用来自制造商的规格表和钻机测量的信息，创建其视野、位置和旋转大致与后投影机的内在和外在参数相匹配的虚拟相机。如在物理钻机中它位于表面下方。将相机产生的渲染视图作为图像发送到投影机。

# 5 结语

我们已经提出了使用红外相机在非参数后投影表面上进行触摸检测的可概括的方法,我们开发了一种使用人体头部表面的原型，投影仪在其上显示头部的图形模型。触摸表面上的事件会提示模型中的视觉和听觉响应,我们展示了一个示例应用场景，其中可以使用触摸事件和结果图形更新来训练医疗保健专业人员进行中风评估。

我们的方法的主要目标是非常直接的触摸图形架构/方法，以确保交互式图形应用程序的适当语义响应，同时最小化感知的延迟。为此，我们的方法包括预校准步骤，来填充在运行时间期间使用的查找表、将2D相机空间中的点直接映射到3D图形空间，以及允许每个以适当的速度进行的去耦触摸和渲染空间/过程。

我们使用反向触摸方法分析了我们的原型的准确性，选择了2D投影仪点，并尝试触摸该点，评估了检测到的触摸事件如何映射回投影机空间。结果表明，在相机/投影仪表面法线更接近平行（病态）的区域，有更大的静态误差（不匹配）。我们应该能够通过各种手段减少这个错误，包括更好的相机/投影机放置和使用密集的校正因子。

我们期待通过多台投影机和相机将我们的方法扩展到完整的人体表面，并评估更广泛的医疗训练场景，并结合实际的全身诊断、治疗和舒适感。

参考文献

[1] Autodesk 123D Catch-3d model from photos. http://www.123dapp.com/catch. Accessed: 2014-09-02.  
[2] H. Benko. Beyond flat surface computing: challenges of depth-aware  
and curved interfaces. In Proceedings of the 17th ACM international conference on Multimedi*a*, MM ’09, pages 935–944, New York, NY,USA, 2009. ACM.  
[3] H. Benko, A. D. Wilson, and R. Balakrishnan. Sphere: multi-touch  
interactions on a spherical display. In *Proceedings of the 21st annual  
ACM symposium on User interface software and technology*, UIST’08, pages 77–86, New York, NY, USA, 2008. ACM.  
[4] P. Bergstrom. Iterative closest point method.http://www.  
mathworks.com/matlabcentral/fileexchange/12627-iterative-closest-point-method. MATLAB Central File Exchange. Accessed 2014-09-18.

[5] J. H. Chuah, A. Robb, C. White, A. Wendling, S. Lampotang, R. Kopper, and B. Lok. Exploring agent physicality and social presence for medical team training. Presence: Teleoperators and Virtual Environments, 22(2):141–170, 2013/09/29 2013.

[6] P. Dietz and D. Leigh. Diamondtouch: a multi-user touch technology.In Proceedings of the 14th annual ACM symposium on User interface software and technology, UIST ’01, pages 219–226, New York, NY, USA, 2001. ACM.

[7] M. J. Garside, M. P. Rudd, and C. I. Price. Stroke and TIA assessment training: A new simulation-based approach to teaching acute stroke assessment. Simulation in Healthcare, 7(2):117–122, 2012.

[8] J. Gu and G. Lee. Touchstring: a flexible linear multi-touch sensor

for prototyping a freeform multi-touch surface. In Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology, UIST ’11 Adjunct, pages 75–76, New York, NY, USA, 2011. ACM.