



北京大学

本科生毕业论文

题目： 基于触摸屏的3D图像
标注工具设计

姓 名： 张吉安

学 号： 00848160

院 系： 信息科学技术学院

专 业： 计算机

研究方向： 计算机应用技术

导师姓名： 王亦洲

二〇一二年五月

基于触摸屏的3D图像 标注工具设计

张吉安 计算机

导师姓名：王亦洲

摘要

由于3D播放设备以及3D电影和电视的日益普及，将普通的视频转换为3D视频也成为了一个极其具有实用意义的研究课题。因为技术水平以及一些固有的难题，目前的3D视频转换仍然不可避免地要加入人工交互的操作使得生成的3D视频尽量逼近真实3D视频。但是现有已知的交互标注的工具，它们对于图像的标注部分无一例外基于传统PC，这对于提高标注效率是不利的。所以我们希望能够开发出一套基于触摸屏设备的标注工具，以提高标注工作的效率。

在毕业设计的过程中，我以已有的标注模型和图像分割算法为基础，完成了基于触摸屏的图像三维信息标注工具的设计，并且基于Android 移动平台开发了一款图像标注工具。

关键词：触摸屏，3D，人机交互，Android

An Design For 3D Image Annotation Tool Based on Touch Screen

Zhangji'an Computer Science

Directed by Prof.Wang Yizhou

Abstract

Since the popularization of 3D playback device and 3D video, converting stereoscopic video from traditional ones has become a practical topic in movie industry. Due to the technical limitations and some immanent problem, we have to do this conversion with the help of human interaction in order to make the result more close to real ones which shot with real stereoscopic camera. But for all the known convert tools now, their image annotation parts is based on traditional PC, which is harmful to raise work efficiency of labeling. So we hope to develop a annotation tool based on touch screen that benefits labeling efficiency.

I design an annotation tool based on touch screen and develop it on Android mobile device as my graduation project.

Keywords: Touch Screen, 3D, Human-Computer Interaction, Android

目录

第一章 项目介绍	1
1.1 视频3D转制现状及前景	1
1.2 项目背景	2
1.3 项目具体工作	2
第二章 系统设计概述	4
2.1 视频标注系统整体概述	4
2.2 视频帧交互式标注模块	5
参考文献	6

第一章 项目介绍

1.1 视频3D转制现状及前景

3D电影技术最早产生于19世纪90年代末期，英国电影先驱威廉·弗里斯格林(William Friese-Green)发明了使用两台播放机放映3D电影的技术，这是可以考证的最早的3D电影的播放技术。而世界上第一部真正的3D长片则是1952年的《非洲历险记》。由于技术手段不足，拍摄成本高昂等原因，3D电影始终没有摆脱在电影工业中的边缘地位。一直到进入21世纪，随着相关技术手段的完善，出现了诸如《极地特快》这样有质量的3D长片，3D视频技术才渐渐重新出现在人们的视野中。

在被称为“3D电影的新元年”的2009年，涌现了诸如《飞屋环游记》《冰河世纪3》以及堪称电影史伟大里程碑的《阿凡达》。2009年底由詹姆斯·卡梅隆(James Cameron)导演的3D电影《阿凡达》的上映则使得原先渐渐回暖的3D电影市场火爆异常，使得大家不再将3D电影视为游乐场中的一种游乐项目，而是真正将其视为电影工业技术中的一支。

而随着《阿凡达》获得巨大成功以来，各个制片公司都纷纷为最新的主打电影采用3D技术。经过不完全统计，2012年在中国上映的3D电影就有22部之多。而随着3D电影带来的热潮，各类3D相关的产品也大行其道。2010年迪士尼公司旗下的美国娱乐和ESPN宣布成立了全球首个3D电视频道ESPN 3D，美国探索传播公司旗下的探索发现频道也在同年同索尼公司和IMAX公司联手成立了3D频道。而我国的首个3D频道也在2012年元旦开始播放节目。在播放设备方面，三星，松下，康佳等国内外公司都推出了3D平板电视，任天堂，LG和HTC等公司推出了支持裸眼3D显示的游戏机和移动电话，英伟达公司推出了基于PC的3D立体显示技术3D Vision。在娱乐方面则出现了诸如《半条命2》，《孤岛惊魂2》等支持3D效

果的游戏。

1.2 项目背景

在我们回顾3D视频相关发展的时候可以注意到，走在前沿的相关技术往往是3D显示技术。而真正的3D视频制作技术的发展却跟不上显示设备的发展。事实上现有的3D视频拍摄的成本是十分昂贵的，《阿凡达》的制作成本高达3亿美金，而《变形金刚3》的3D版本仅仅在制作费的支出上就要高出2D版本3000万美金。所以为了应对各种需要较多3D片源的同时要求控制成本的情况(例如3D电视频道)，发展将普通视频转换为3D视频的技术就显得十分必要。

不仅如此，随着3D电影越来越受到欢迎，将原先的经典电影转换为3D电影也成了一种潮流。最成功的例子就是经典电影《泰坦尼克号》的3D转制版，在2012年4月上映之后在全球获得了3.3亿美元的票房。而我国经典动画片《大闹天宫》的3D转制版也在同年上映。但是《3D 泰坦尼克号》巨大的成功背后却是高达1800万美金的制作成本。所以研发低制作成本的3D转制技术是3D产业发展的关键所在。

目前存在的3D转制技术可以分为全自动和半自动两种，其中全自动技术所产生的效果和真实3D视频观影感受相去甚远。半自动技术虽然转换效果较好，但是由于加入了人工交互的环节所以存在效率不高的问题。在之前开发2D到3D视频交互式转换系统的过程中就发现由于传统PC在图像标注的交互中存在的不够直观，以及频繁的鼠标操作容易给工作人员带来疲劳而导致工作效率下降的问题。

基于触摸屏的标注工具的设计和开发就是针对上述问题，力图使用触摸屏交互直观的特性解决由于传统PC交互所带来的效率低下的缺陷。希望能够通过新的交互方式提高人工标注的效率，从而提高软件整体的工作效率。

1.3 项目具体工作

在本项目中我的工作主要是设计基于触摸屏的单幅图像的标注工具及其实现。

第一，根据视频标注操作的特性结合触摸屏交互的特点，整理出视频交互标注操作的具体模型。依据所整理出的模型为标注工具设计完整的软件模型。

第二，在Android平台的移动设备上实现第一步设计出的软件模型，根据实现的结果对于该软件模型的可行性做出评估。

第二章 系统设计概述

2.1 视频标注系统整体概述

下图是一个成熟的视频标注系统的整体结构。可以看出，在一个视频标注系

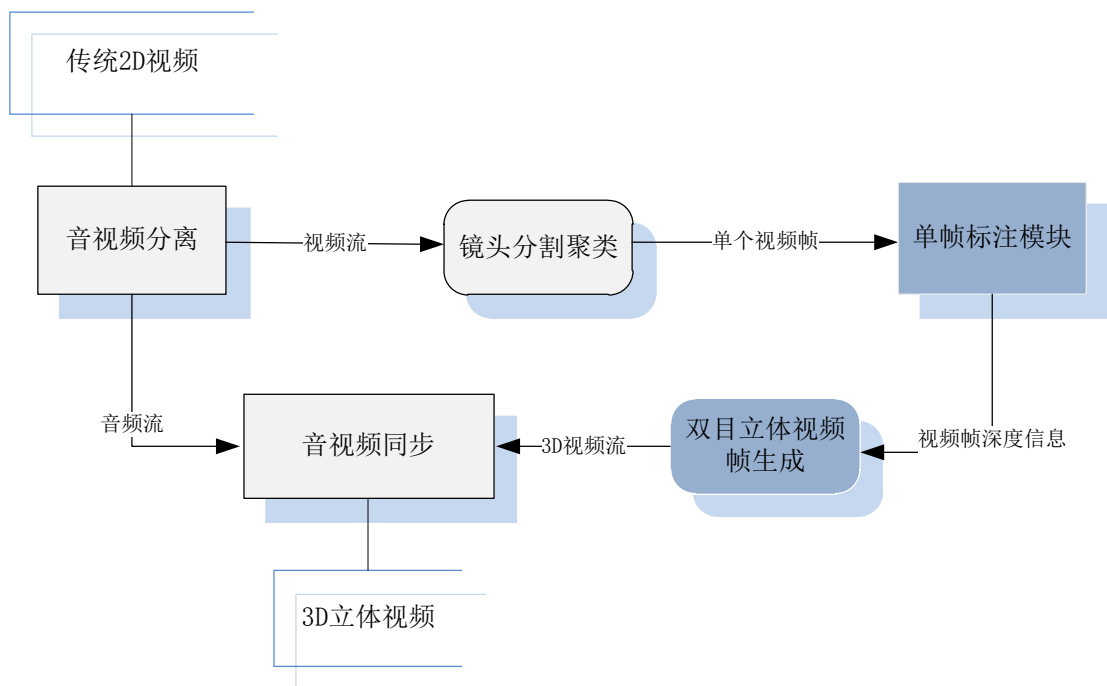


图 2.1: 系统流程

统中，已经存在完善且成熟的技术的有音视频分离和音视频同步部分，而镜头聚类分割和双目立体视频的生成也有了成熟的自动转换的技术。

所以实际上交互式的视频标注系统中的交互标注操作几乎都完全花费在单帧的标注操作上，事实上在我们已有的基于传统PC上的标注工具的使用中发现，除了视频的单帧标注之外，其余的操作可以完全地实现自动化。

2.2 视频帧交互式标注模块

单帧的交互式工具的目的是为了在单幅图像上标记出图像的深度信息，生成对应于图像的depth map。同时这也是本项目工作的重点。

目前针对视频帧的标注模块大致可以分为两个方面，分别为前景部分的深度标注和背景的深度标注。其中背景的深度估计对于估计的准确度和物体互相之间的深度的把握较前景的要求要低，而前景的深度估计则在物体的深度变化和物体之间的深度关系上和观影者的感受较为密切。也就是说，在背景的估计上我们可以使用较为粗略但是效率较高的方法，而在前景的标注上则要求尽量能够反映前景物体的深度信息以带给观影者较为真实的3D观影感受。目前可利用的背景估



图 2.2: 自动/手动深度估计的对比

计的方法有两种，一种是利用Stage Model^[1]

参考文献

- [1] V. Nedovic, A. W. M. Smeulders, A. Redert, J. M. Geusebroek. Stages As Models of Scene Geometry[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010, **32**(9):1673–1687