分类号： R319 单位代码： 10335

密 级： 无 学 号：21015066



硕士学位论文

**中文论文题目 ：胶片按需打印技术研究及系统实现**

**英文论文题目 ：The Research of On-demand Film Printing**

**Technology and Its Implementation**

申请人姓名： 徐世才

指导教师： 段会龙 教授

专业名称： 生物医学工程

研究方向： 医学信息学

所在学院： 生物医学工程与仪器科学学院

**论文提交日期 2013年 1月 25日**

**胶片按需打印技术研究及系统实现**



**论文作者签名:**

**指导教师签名:**

论文评阅人1： 刘济全 副教授 浙江大学

评阅人2： 吕旭东 教授 浙江大学

评阅人3： 薛凌云 副教授 杭州电子科技大学

答辩委员会主席： 叶枫 教授 浙江工业大学

委员1 ： 宁钢民 教授 浙江大学

委员2 ： 段会龙 教授 浙江大学

委员3 ： 邓宁 副教授 浙江大学

答辩日期： 2013年3月8日

**The Research of On-demand Film Printing**

**Technology and its Implementation**

**Author’s signature:**

**Supervisor’s signature:**

Thesis reviewer 1： Jiquan Liu Associate Professor /ZJU

Thesis reviewer 2： Xudong Lv Professor/ZJU

Thesis reviewer 3： Lingyun Xue Associate Professor/HDU

Committee of Oral Defence：

Committee Chair： Feng Ye Professor/ ZJUT

Committeeman 1： Gangmin Ning Professor/ZJU

Committeeman 2： Huilong Duan Professor/ZJU

Committeeman 3： Ning Deng Associate Professor /ZJU

Date of oral defence： March 8，2013

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得**浙江大学**或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名： 签字日期： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解**浙江大学**有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权**浙江大学**可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名： 签字日期： 年 月 日

导师签名： 签字日期： 年 月 日

# 致谢

从大三时第一次接触实验室，到两年半的研究生学习阶段结束，在这里我认识了很多的老师和同学，在学习和生活中都得到了很多无私的帮助，在此论文完成之际，我要衷心地向他们表达最诚挚的感谢。

感谢我的导师段会龙老师，段老师一丝不苟的治学态度影响了实验室所有的学生，他的每一次教诲都让我受益良多。

感谢吕旭东老师，吕老师对我的悉心指导和谆谆教诲是我在研究生期间的最大收获，他勤勉的工作态度和乐观向上的精神给我带来了深刻的影响，在我未来的生活和工作中仍将起到很大帮助。

感谢赵晨晖、黄正行博士对我学习上的指导和帮助，为我解决了很多的疑问。感谢维科软件公司的傅彬，为我的工作展开提供了良好的环境。感谢朱琴学姐，将我带到了这个实验室；感谢王一宝、王利师兄，为我的技术能力提高提供了很多帮助。感谢实验室的安继业老师，维科软件公司的蔡东旭、吕彩虹、张成学等人对我工作的帮助和支持。我还要感谢我们同届的几位同学，殷琳、俞伟飞、范武、范政昂、王彬，谢谢有你们，两年多一同走过，忧乐与共。

最后，我要特别感谢我的家人，谢谢他们对我一贯的支持，他们的辛勤工作和最无私的关爱，是我一路走来的最大动力。

徐世才

2013年1月

于浙大玉泉

# 摘要

**关键字：**胶片按需打印；DICOM打印；OCR；胶片虚拟打印机

# Abstract

# 目录

[致谢 I](#_Toc350410292)

[摘要 II](#_Toc350410293)

[Abstract IV](#_Toc350410294)

[目录 VI](#_Toc350410295)

[第1章 引言 1](#_Toc350410296)

[1.1 需求分析 1](#_Toc350410297)

[1.2 胶片按需打印 3](#_Toc350410298)

[1.2.1 胶片打印流程分析 3](#_Toc350410299)

[1.2.2 胶片按需打印面临的问题与解决方案 4](#_Toc350410300)

[1.2.3 关键技术问题 8](#_Toc350410301)

[1.3 论文研究目标和内容 10](#_Toc350410302)

[第2章 DICOM打印实现 11](#_Toc350410303)

[2.1 DICOM概述 11](#_Toc350410304)

[2.2 DICOM打印剖析 12](#_Toc350410305)

[2.2.1 DICOM打印的信息模型 12](#_Toc350410306)

[2.2.2 DICOM打印的网络通讯 13](#_Toc350410307)

[2.3 DICOM打印实现 19](#_Toc350410308)

[2.3.1 DICOM打印通讯层实现 19](#_Toc350410309)

[2.3.2 DICOM打印服务端实现 19](#_Toc350410310)

[2.3.3 DICOM打印客户端实现 22](#_Toc350410311)

[2.4 DICOM打印异构性解决方案 23](#_Toc350410312)

[2.4.1 多次打印一次会话 24](#_Toc350410313)

[2.4.2 单次打印多次会话 24](#_Toc350410314)

[2.4.3 单张胶片影像整合 25](#_Toc350410315)

[2.5 本章小结 25](#_Toc350410316)

[第3章 电子胶片信息识别 26](#_Toc350410317)

[3.1 OCR概述 26](#_Toc350410318)

[3.2 Tesseract引擎剖析 27](#_Toc350410319)

[3.3 电子胶片信息识别与验证 29](#_Toc350410320)

[3.3.1 信息识别流程 30](#_Toc350410321)

[3.3.2 图像预处理 31](#_Toc350410322)

[3.3.3 检查-患者匹配方法 34](#_Toc350410323)

[3.3.4 人工匹配方案 38](#_Toc350410324)

[3.4 本章小结 38](#_Toc350410325)

[第4章 胶片按需打印系统 39](#_Toc350410326)

[4.1 胶片按需打印系统概要设计 39](#_Toc350410327)

[4.2 胶片按需打印系统详细设计与开发 41](#_Toc350410328)

[4.2.1 数据库 41](#_Toc350410329)

[4.2.2 文件库 46](#_Toc350410330)

[4.2.3 FilmOCR 47](#_Toc350410331)

[4.2.4 DPHServer 47](#_Toc350410332)

[4.2.5 DPHClient 48](#_Toc350410333)

[4.3 胶片按需打印系统实践 50](#_Toc350410334)

[4.3.1 系统实际应用场景 50](#_Toc350410335)

[4.3.2 系统支持的打印源设备 51](#_Toc350410336)

[4.3.3 系统支持的打印机 51](#_Toc350410337)

[4.3.4 系统实际使用效果 51](#_Toc350410338)

[4.4 本章小结 52](#_Toc350410339)

[第5章 总结与展望 53](#_Toc350410340)

[5.1 总结 53](#_Toc350410341)

[5.2 展望 54](#_Toc350410342)

[作者简介 55](#_Toc350410343)

[参考文献 56](#_Toc350410344)

# 引言

## 需求分析

## 胶片按需打印

### 胶片打印流程分析

### 胶片按需打印面临的问题与解决方案

#### 基于PACS的胶片按需打印

#### 基于虚拟打印机技术的胶片按需打印

#### 结合PACS和虚拟打印机技术的胶片按需打印

#### 胶片按需打印实现方法的选择

### 关键技术问题

#### DICOM打印

#### 电子胶片信息识别

## 论文研究目标和内容

基于上述分析，本论文的研究目标是解决胶片按需打印系统的关键问题并实现该系统，在实际应用中证明其可用性。具体来说，包括以下内容：

* 实现DICOM打印，包括SCU端和SCP端的实现。同时，针对不同设备工作站对DICOM打印的不同优化做出相应的处理，使胶片按需打印系统在接收来自设备的打印任务时，可以作为通用的虚拟DICOM打印机存在。
* 利用OCR技术实现适用于对电子胶片中患者号、检查号识别的算法。这项工作包括识别区域的获取、影像预处理、影像识别、识别结果正误的判断等。本论文需要从整体上优化识别算法，提高识别正确率并且提出一套判断识别结果正误的方法。
* 基于以上两个问题的解决，开发胶片按需打印系统，并在医院放射科应用该系统，验证其可用性和防止胶片浪费的能力。

# DICOM打印实现

## DICOM概述

## DICOM打印剖析

### DICOM打印的信息模型

### DICOM打印的网络通讯

#### DICOM网络层次

#### DICOM通讯的建立

图2‑8 DICOM通讯双方在低层DICOM层的数据交流图

#### DICOM打印的通讯过程

图2‑9 DICOM打印时SCP与SCU端的通讯流程

图2‑10 DICOM传输时SCP与SCU端的通讯流程

## DICOM打印实现

### DICOM打印通讯层实现

### DICOM打印服务端实现

#### 打印数据归档

表2‑1创建打印胶片序列时需要解析并存储的属性

|  |  |  |  |
| --- | --- | --- | --- |
| **属性** | **Tag** | **取值范围** | **含义** |
| Number of Copies | (2000,0010) | 1-99 | 打印份数 |
| Print Priority | (2000,0020) | LOW、MED、HIGH | 打印优先级 |
| Medium Type | (2000,0030) | CLEAR FILM 、BLUE FILM、PAPER、CURRENT | 媒质类型 |

表2‑2创建打印胶片实例时需要解析并存储的属性

|  |  |  |  |
| --- | --- | --- | --- |
| **属性** | **Tag** | **取值范围** | **含义** |
| Image Display Format | (2010,0010) | STANDARD\C,R  ROW\R1,R2,R3  COL\C1,C2,C3 | 单张胶片分格方式 |
| Film Orientation | (2010,0040) | PORTRAIT、LANDSCAPE | 胶片方向 |
| Film Size Id | (2010,0050) | 14INX17IN、14INX14IN、11INX14IN、11INX11IN、85INX11IN、8INX10IN | 胶片尺寸 |
| Magnification  Type | (2010,0060) | REPLICATE、BILINEAR  CUBIC、NONE | 放大类型 |

表2‑3传送单格影像时需要解析并存储的属性

|  |  |  |  |
| --- | --- | --- | --- |
| **属性** | **Tag** | **取值范围** | **含义** |
| Image Position | (2020,0010) | 1-当前胶片分格方式支持的最多分格数 | 当前分格在胶片中的位置 |
| Basic Grayscale  Image Sequence | (2020,0110) | N/A | 影像自身相关信息 |
| >Photometric  Interpretation | (0028,0004) | MONOCHROME1  MONOCHROME2 |  |
| >Rows | (0028,0010) | 1-胶片最大行 |  |
| >Columns | (0028,0011) | 1-胶片最大列 |  |
| >Pixel Aspect Ratio | (0028,0034) |  |  |
| >Bits Allocated | (0028,0100) | 8或16 |  |
| >Bits Stored | (0028,0101) | 8-16 |  |
| >High Bit | (0028,0102) | 7-15 |  |
| >Pixel  Representation | (0028,0103) | 0或1 |  |
| Polarity | (2020,0020) | NORMAL、REVERSE | 是否反向 |
| Magnification Type | (2010,0060) | REPLICATE、BILINEAR  CUBIC、NONE | 放大方式 |
| Smoothing Type | (2010,0080) | 0-15 | 平滑方式 |

### DICOM打印客户端实现

图2‑12按需打印系统作为DICOM打印SCU端的实现流程

## DICOM打印异构性解决方案

### 多次打印一次会话

1. Film Session实例的请求后才开始对接收到的打印数据进行归档，在接单次打印多次会话

### 单张胶片影像整合

## 本章小结

本章对DICOM标准进行了简单地概述之后，对DICOM通讯进行了较详细的剖析，包括通讯中最重要的几个概念以及通讯时网络的分层和数据交互的流程。

之后，本章对DICOM打印进行了深层次的剖析，分析了胶片打印时SCU和SCP之间的通讯和各自需完成的工作。同时，本章在实现了DICOM打印通讯的基础上，对按需打印系统作为SCP端和SCU端分别进行了实现。

最后，结合具体应用，本章分析了不同厂商在实现DICOM打印协议时的差别以及造成这些差别的原因。同时针对这些差别，本章提出了胶片按需打印系统需要采取的相应处理方式。

# 头痛诊断决策支持系统实现

## 系统背景概述

头痛是临床常见的症状,通常指局限于头颅上半部,包括眉弓、耳轮上缘和枕外隆突连线以上的疼痛。头痛主要分为原发性头痛和继发性头痛。原发性头痛，主要包括偏头痛和紧张型头痛(TTH)，在世界各地广为流行，据估计分别影响世界大约11％和42％成年人口。2012年WHO公布的最新数据显示偏头痛已成为第七类致残性疾病。据统计中国在中国原发性头痛发病率为23.8％，其中，偏头痛为9.3％和紧张型头痛为10.8％。这些头痛具有致残性，影响工作，学习和日常活动，降低生活质量，并给人们带来了沉重的负担。

虽然头痛患者众多，但目前我国在头痛疾患的诊断、治疗、教育上仍较为混乱，正确诊断及治疗的比例极低[3]，并且仍然存在很多不规范及医生自己杜撰的头痛诊断，如“血管性头痛”、“神经性头痛”、“血管神经性头痛”等。由于不规范的诊断与治疗会出现很多问题，如药物过量性头痛、慢性每日头痛、滥用仪器检查等。（基金申请）

在国内对于原发性头痛诊断的研究中指出，偏头痛和紧张型头痛就诊医院以社区或乡镇及县区医院为主，因此可见基层医院在中国头痛防控工作的作用尤为重要，基层医院的头痛的诊治水平在很大程度上决定了头痛诊疗的成效。因此，建立面向社区的头痛诊断决策支持系统是具有重要意义的。

## 基于框架的系统开发流程

### 推理引擎选择

目前对于头痛的诊疗知识，目前国内外通行的是2004年国际头痛学会制定的国际头痛疾患分类标准(ICDH-II)。基于规则—CLIPS引擎

首先，参考数字化临床指南的规范化绘制符号，由临床医生将自然语言描述的文本形式的临床指南以流程图的形式进行表达。随后，本研究参照SAGE(standards—based sharable active guideline environment)模型，利用Prot696建模工具，为流程图化的临床指南建立相对应的医学知识表达模型。通过Prot696建立的医学知识模型虽能被计算机表达，但还不能被计算机自动理解与推理。因此，采用文献[5]中的方法，通过规则生成器，由程序自动将医学知识表达模型转换成符合计算机可推理要求的CLIPS规则。流程如下图所示

### 数据模型设计

1. 病人信息集合PatInfo

|  |  |  |
| --- | --- | --- |
| 名 | 数据类型 | 说明 |
| PatName |  |  |
| Sex |  |  |
| Age |  |  |
| Education |  |  |
| Job |  |  |
| Phone |  |  |
| Address |  |  |
| Weight |  |  |
| Height |  |  |

1. 问诊信息集合
2. 病史信息集合

### 问诊界面配置

问诊部分主要是依据头痛诊断专家的意见，建立以下流程：

1. 继发性头痛筛查，主要根据继发性头痛的症状，提供筛查功能，如下图所示

2. 原发性头痛问诊，根据原发性头痛诊断所需的信息，提供问诊导向，如下图所示：

3. 系统辅助诊断，汇总前面问诊的信息，并提供诊断意见，如下图所示：

4. 医嘱处置下达，根据诊断结果，给出用药的建议，并给医生最终医嘱下达的功能，如下图所示：

1. 病历报告预览打印，将问诊的信息和医嘱信息合并，按报告的格式提供预览打印的功能：

### 数据交互层实现

数据交互层是与前端数据录入展示组件交互的部分，通过调用组件提供的接口，将前端输入的头痛问诊数据组织为标准的Json格式的文件传输到服务端以及将服务端传来的Json格式的数据解析后显示在前端的问诊页面。

## 系统实现

## 本章小结

本章首先简要地介绍了OCR的概念以及其常用的算法和流程，接下来本章

套基于检查-患者匹配策略的对识别结果正误进行判断的方法，并通过理论推导证明了该方法的可用性。

对于因检查号或患者号的错误识别导致不能完成检查-患者匹配的电子胶片，本章还提出了一套手动匹配的方法，以弥补本章算法自身的不足。

# 老年痴呆症诊断决策支持系统实现

## 系统背景概述

阿尔茨海默氏病（Alzheimer'sdisease，AD）是一种退化性痴呆特征的记忆丧失和认知功能障碍，目前，AD已成为导致老年人死亡的第四位主要原因，仅次于心脏病、癌症及中风。我国是世界上老龄人口最多的国家，现有老年人约1．42亿(占总人口的11％)[4】。据估计，目前我国约有310万AD患者[5]，疾病终期患者将完全丧失生活自理能力，给家庭和社会增加了沉重的心理和经济负担，成为一个严峻的社会问题。随着我国人口老龄化的日益严重，阿尔兹海默症对卫生和经济等将产生越来越显著的影响。

由于AD在临床还没有可以完全治愈的方法，主要通过延缓病情的发展进行治疗，而且疾病的晚期诊疗方案受到很多的限制，因此早期检测出高风险的人群是治疗AD的关键。目前，AD主要通过神经心理学量表和分子影像学检查进行诊断，其中神经心理学量表费用低廉，方便快速，非有创诊断手段，更易被患者接受，所以神经心理学量表已广泛使用于AD的初步筛查。

## 基于框架的系统开发流程

### 推理引擎选择

由于阿尔兹海默症的发病是多因素、多机制综合作用的结果，不仅受机体衰老和遗传因素的影响，还与环境因素有关【小鼠胆碱能神经系统损伤的干预研究】。在这种疾病的临床研究已经有各种神经心理学量表，如简易精神状态检查（MMSE），临床痴呆评定量表（CDR），蒙特利尔认知评估（MOCA）等。然而量表的使用在不同地区、不同研究机构所采用的临床分界值往往有差异。单个量表的使用也存在一定的限制性，因此采取多量表结合的方式进行诊断。诊断过程主要是针对各个量表的等分情况进行综合考虑。本文中根据经北京某三甲医院的专家筛选的145例典型病例，采取构建贝叶斯网络的初步检测模型来进行诊断

### 数据模型设计

1. 病人信息集合PatInfo

|  |  |  |
| --- | --- | --- |
| 名 | 数据类型 | 说明 |
| PatName |  |  |
| Sex |  |  |
| Age |  |  |
| Education |  |  |
| Job |  |  |
| Phone |  |  |
| Address |  |  |
| Weight |  |  |
| Height |  |  |

1. 问诊信息集合
2. 病史信息集合

### 问诊界面配置

老年痴呆症的问诊过程是医生按照量表，一题一题的询问病人，并将结果记录下来。依据专家的意见，建立以下流程：

1. 基本情况，主要包括体格检查、疾病用药的情况以及相关的检查信息，这个部分主要是了解病人的基本的生理情况，如下图所示：
2. 认知检查，综合多份量表，提供问诊导向，医生可以根据页面提示的问题询问病人。如下图所示：
3. 系统辅助诊断，综合之前的量表各项的得分情况，提供诊断意见，如下图所示：

### 数据交互层实现

数据交互层是与前端数据录入展示组件交互的部分，通过调用组件提供的接口，将前端输入的头痛问诊数据组织为标准的Json格式的文件传输到服务端以及将服务端传来的Json格式的数据解析后显示在前端的问诊页面。

## 系统实现

根据以上步骤系统的开发已完成，下面展示系统的功能

1. 系统的身份验证流程

在页面输入用户名和密码后点击登录按钮，若正确则页面跳转至病人信息页面

1. 系统的病人信息查询流程

在查询病人一栏中，填写所需的查询条件，如病人姓名，填写完毕后点击查询病人按钮，病人列表栏将会显示符合条件的病人的姓名及性别。点击选取病人后，页面跳转至问诊记录页面，自动显示病人的历次就诊记录。点击记录的时间，页面右栏将显示相应的问诊内容。

1. 系统问诊流程

在问诊记录页面点击开始问诊，将跳转到问诊页面，页面中有导航栏，显示目前的问诊进度。页面右下角的【上一步】和【下一步】按钮在基本情况、认知检查和辅助检查三个主要部分之间跳转。

## 本章小结

本章首先简要地介绍了OCR的概念以及其常用的算法和流程，接下来本章

套基于检查-患者匹配策略的对识别结果正误进行判断的方法，并通过理论推导证明了该方法的可用性。

对于因检查号或患者号的错误识别导致不能完成检查-患者匹配的电子胶片，本章还提出了一套手动匹配的方法，以弥补本章算法自身的不足。

# 总结与展望

## 总结

放射科胶片按需打印是随着放射科信息化而产生的现实需求，胶片按需打印系统主要可以有三种实现方式：

* 基于PACS的打印工作站实现方式。
* 基于虚拟打印机技术的实现方式。
* 结合PACS和虚拟打印机技术的实现方式。

为了不影响放射科既有的工作流程和人员分工，也为了不给放射科工作人员添加额外的工作，本论文采用了第二种实现方式。在系统实现过程中，该方法需要解决两个关键问题：

* DICOM打印双向通讯。胶片虚拟打印机不但需接收来自设备工作站的电子胶片数据，也要把数据发送给胶片打印机，因而需同时作为SCP和SCU角色实现DICOM打印的双向通讯。DICOM打印是最复杂的DICOM通讯服务，而且不同设备工作站存在异构性，更增加了实现难度。
* 电子胶片患者信息识别。胶片虚拟打印机接收到电子胶片后，必须用患者号和检查号进行标识，才能实现按需打印。由于以DICOM标准接收到的电子胶片数据中不包含结构化的患者号和检查号，因而需用OCR技术从二进制影像数据中识别出这些信息。

因此，本论文的主要工作包括：

* 在详细分析DICOM打印通讯原理的基础上，提出了一种分层服务的设计方案，不但能实现DICOM打印的双向通讯，而且能针对不同设备工作站的异构性进行自适应处理。
* 提出了一套基于Tesseract识别引擎，在胶片上识别患者号和检查号的方法，该方法可通过检查-患者匹配策略判别识别结果的正确性。
* 基于关键问题的解决，本论文研发了胶片按需打印系统，具备完整的胶片虚拟打印机功能以及按需打印服务功能。系统已在温州医学院附属第二医院等地开展实践应用，电子胶片的正确匹配率达到了96.5%，对胶片的节省率达到了14.8%。结果表明系统具有良好的可用性，能够有效地减少胶片的浪费。
* 为了配合医院在引入按需打印系统后实现管理流程的优化，本论文还做了相关的工作，实现胶片打印与收费的管理集成。

## 展望

本论文实现的胶片按需打印系统虽然能够较好地满足医院的需求，且在实际应用中证明了其可用性和方便性。但该系统仍然有诸多不足，主要表现在以下几个方面：

* 系统目前只支持对灰度胶片的接收和打印，而未实现对DICOM彩色打印的支持，因此本论文接下来的工作需要使系统支持DICOM彩色打印。
* 系统未对刷卡打印进行支持，目前病历卡在医院已得到了广泛的应用，许多医院都具有了刷卡自主取报告的服务。本论文的按需打印系统目前采用的是由护士从web客户端搜索电子胶片点击打印的工作形式，在接下来的工作中，本论文将实现通过刷卡自主取胶片的功能。
* 本论文提出的OCR方法对电子胶片的检查-患者匹配率还有待进一步的提高，尤其是图像的预处理方法和流程还需要进行更多的完善。

DICOM标准每年都在进行修改和完善，随着胶片按需打印需求的增加，在未来，DICOM标准中胶片打印部分可能会要求设备工作站向打印机传输患者号、检查号等属性，那么基于虚拟打印机技术的按需打印系统将不需要利用OCR技术从电子胶片获取患者号、检查号，按需打印系统将更容易实现且能得到更广泛的应用。

# 作者简介

冯冠军，生于1988年10月，湖北武汉人。2006年9月进入华中科技大学控制系，主修自动化专业，2011年7月获得工学学士学位。同年9月，进入浙江大学生物医学工程与仪器科学学院，主修生物医学工程专业，攻读硕士学位。

# 参考文献