# 第一章 绪论

## 1.1 选题的背景和意义

### 1.1.1 选题的背景

近年来，随着人脸识别技术的迅速发展，越来越多的基于人脸识别的应用进入了日常生活领域。从人脸识别考勤系统、人脸识别门禁系统等较为传统的人脸识别应用，到新近流行起来的“刷脸”支付的手机app、“刷脸”取款的ATM，人脸特征凭借着其方便、友好的交互方式，迅速取代了其他的冗余、繁琐的验证方式。但目前，绝大多数的人脸识别系统都把研发重心放置于身份识别的准确性以及算法效率上；却忽视了人脸识别的安全性问题，即——输入的人脸可以是非真实的脸（包括高分辨率的彩色打印照片、3D打印的人脸模型、人脸面具等），使用者以此可以仿冒他人身份通过人脸识别系统，对后续的应用安全性构成了极大地安全隐患。

任何一种验证系统，都会面临仿冒攻击的威胁。应对的方式无外乎以下两种：（1）通过保护验证信息，使仿冒者不易获取；（2）通过验证系统本身来鉴别验证信息是否有效。前者的主要应用场景就是密钥验证系统（包括对称密钥以及非对称密钥），使用此方式来保护验证系统对密钥持有者的要求较高，且密钥本身不易被他人获取。由于人脸的特殊性，它作为日常生活中人类的最重要的身份标示，被广泛地传播，极容易被他人获取；所以对于人脸验证系统而言，通过验证系统本身来鉴别验证信息的可靠性是唯一的解决方案。

因此，用于增强人脸识别系统的安全性的人脸活体验证技术受到了越来越多人的关注。人脸活体验证技术旨在通过人脸信息来验证目标的合法性，更加具体的说，即在传统的人脸识别系统之前，增加一道防火墙，所有试图进行人脸验证的输入都将首先进行人脸活体验证，通过验证的输入才能继续进行人脸识别，否则就认为是非法输入，从而拒绝访问。

### 1.1.2 选题的意义

信息安全的重要性和敏感性以及无需赘述。在人脸识别领域，使用者不仅仅需要方便快捷地进行身份验证，更希望验证系统能够保证其身份的安全性，即不被仿冒者盗用。人脸活体验证系统是一个添加在人脸识别系统之前的认证模块，采用带有人脸活体验证模块的人脸识别系统有助于在保证信息真实有效并且可以使用的情况下大大提高人脸认证的效率。传统的人脸识别中，其可欺骗性会导致系统运行的结果不可靠等非预期性结果。所以在高敏感性的应用环境中，对于机器识别的结果，还需要人工进行筛查纠错，这就直接降低了人脸识别系统的实用性。因为动用人力来进行筛查纠错不仅成本高，而且效率极其低下，也容易发成错误。

本次课题就是针对这一应用环境，通过图像信息来对输入信息的合法性进行验证，旨在通过尽可能少的交互量以及时间，最大限度地提高伪造人脸的识别率（即最大限度地提高伪造人脸成功的成本）。

综上所述，人脸活体验证系统具有非常广阔的应用场景，人脸活体验证技术能够为人脸识别系统提供可靠的输入数据，使人脸识别系统的应用范围得到进一步的拓展。

## 1.2 国内外研究现状

### 1.2.1 人脸活体验证算法研究现状

随着人脸识别技术的成熟，在2000年左右，活体验证技术逐渐被各大领域重视。人脸活体验证技术的目标是通过一幅（组）人脸照片来判断照片中的人脸是否为真实的人脸，而非模型或者照片。

Herbert Bay等人于2006年提出了一种基于SFM（Structure from Motion）的活体检测模型。该模型通过定位眼睛、嘴巴等显著位置来预测特征点的三维深度，从而实现判断活体的方法。Klaus Kollreider等人于2009年提出了利用光流（Optical Flow）来分析人脸各部位的移动量，最终进行活体检测的方法。孙霖于2010年提出了利用人脸识别进行多模活体验证的概念。其思想大多是将多个检测源的数据进行混合，然后计算加权检测结果，来进行评判。

李翼于2011年采用高斯差分（Difference of Gaussian, DoG）特征结合稀疏低秩双线性逻辑斯特回归的方法，在NUAA数据库上获得了很高的准确率。杨健伟于2014年提出了基于运动一致性的活体检测的方法，通过计算微纹理，计算脸部与背景的运动一致性。为人脸活体检测提供了新的方向。曹瑜于2014年提出了局部二值特征（Local Binary Pattern, LBP）结合卡方统计量来判断活体人脸的方法，使用该方法能够有效提高检测的准确率。刘华成于2014年提出了使用奇异值分解（Singular Value Decomposition, SVD）与HSV直方图的方法用来检测活体，具有较高的鲁棒性。

由于活体检测的巨大市场，工业界也推出了很多了里程碑式的产品。例如2012年，由北京旷视科技推出的Face++，在LFW评测集上连续创造世界纪录，并且在人脸活体检测上，也有较为出色的表现。不过目前Face++仅以API调用的形式为开发者提供服务，并没有公开其实现算法。阿里巴巴公司的手机支付宝app于2015年推出了“刷脸”登录的功能，手机支付宝采用的是交互式验证手段，需要用户进行做出眨眼、点头等动作进行验证，其识别率达到90%以上。

### 1.2.2 人脸特征点定位算法研究现状

人脸特征点定位技术的目标是通过一幅人脸图片，得到其中人脸的语义特征点位置的技术。其结果的准确性通过如下误差函数来衡量：

|| S – SGround True||2

其中S是指通过预测得到的人脸特征点位置，SGround True表示实际的人脸特征点位置，但由于在预测阶段，SGround True是未知的，因此不能直接通过最小化误差函数来得到S。现阶段有两类方法，分别从两个不同的角度去得到S，因此绝大多数主流的人脸特征点定位算法都可以被分为如下两类：基于最优化算法的人脸特征点定位算法以及基于回归算法的人脸特征点定位算法。

基于最优化算法的人脸特征点定位需要再设计一个能量函数，该能量函数要求与上述误差函数相关性大，通过最优化该能量函数来减少误差。此类方法的结果误差直接取决于能量函数是否能够反映误差函数的变化以及能量函数能否被最优化。例如由Timothy F. Cootes等人于2001年提出的AAM（Active Appearance Models）算法就是典型的基于最优化算法的人脸特征点定位算法。AAM算法通过训练得到一个表观模型（Appearance Model）；然后在预测阶段，使用该模型最小化纹理误差（Texture Residual）来达到减少误差的目的。但是，由于表观模型的形变能力不强，所以对于训练集之外的人脸，不能做到很好的拟合；并且AAM算法受初始化数据的影响较大，鲁棒性不强。

基于回归的算法并不需要另外设计一个能量函数，此类算法是通过训练得到回归器，在预测阶段直接通过输入映射出输出结果，此类算法的普遍优点是能够有效利用大容量的训练集，从而提高回归质量。David Cristinacce等人与2007年提出的Boosted Regression框架结合ASM（Active Shape Models）算法以及Michel Valstar等人与2010年提出的Boosted Regression框架结合Graph Models的方法中，均采取为每个特征点训练一个回归器的策略。这种策略有两个很明显的劣势：不容易区分一些特征区分度不高的特征点（比如脸颊上的两个特征点）、没有考虑特征点之间的位置关系。Piotr Doll´ar等人于2010年使用所有特征点进行整体回归，并提出了级联回归器的算法框架，该方法通过贝叶斯概率模型逐步对结果进行修正。Xudong Cao改进了级联回归器的框架，并且引入了新的语义特征，在人脸特征点定位上取得了非常好的效果。

## 1.3 本人主要工作

* 研究人脸活体验证系统的定位，了解相关的基本功能需求；并且通过阅读大量文献，掌握国内外有关人脸活体验证的发展现状。
* 根据系统的应用场景，明确系统角色，确定最后的项目功能性需求以及非功能性需求。
* 通过项目的需求分析，确定项目的整体架构以及相关系统功能模块的划分。
* 根据需求分析，尝试进行核心算法的实现并且通过不同算法的组合来达到项目需求，并最终确定核心算法流程。
* 细化项目的各个模块设计，抽象出各个模块之间的通信接口，按照软件工程的项目规范把所有模块实现。
* 搭建系统的测试环境，并根据系统的功能性需求以及非功能性需求进行完善的测试。

## 1.4 论文组织结构

本论文的内容结构安排如下：

第一章为绪论。该章论述了人脸活体检测系统的项目背景和项目意义，并且介绍了现阶段工业界和学术界有关人脸活体检测以及人脸特征点定位的最新成果。

第二章为相关技术介绍。主要介绍了本项目涉及的一系列算法原理以及实现框架，其中包括半朴素贝叶斯模型、随机蕨回归算法、显式形状回归算法、Node.js以及MongoDB。

# 第二章 相关技术介绍

## 2.1 半朴素贝叶斯模型原理

半朴素贝叶斯模型（Semi-Naive Bayes Model）对朴素贝叶斯模型（Naive Bayes Model）的改进。该算法由Geoffrey I. Webb等人于2005年提出。该算法填补了朴素贝叶斯模型的两大缺陷：需要大量训练样本以及独立性假设与大部分特征不相符。假设输入的N维特征向量的形式为：f = (f1, f2, … , fN)，模型的输出类别为C∈{c1, c2, …, cK}。贝叶斯模型是一个通过计算后验概率并求出最大值所属的类别，并作为输出。贝叶斯模型的形式化表示为：

H(f) = argmaxk P(C=ck|f1,f2,…,fN)

使用贝叶斯公式可以将后验概率的形式变化为先验概率的形式：

H(f) = argmaxk P(f1,f2,…,fN | Ck) \* P(Ck)

在朴素贝叶斯模型中，假设所有的特征都是互相独立的，因此上述联合概率分布等于各个特征的先验概率的乘积，即为：

但是这个假设过于严格，在实际的应用中，该假设往往是不成立的，这就直接导致了朴素贝叶斯模型得出的结果与真实情况有较大的偏差。

半朴素贝叶斯模型取消了这个全局独立性假设，而是对特征进行分组，假设各个组内特征不独立，各个组间特征独立。假设N个特征被划分为L个组，每个组内含有S个特征，则上述的联合概率分布可以表示为：

通过计算每个组内的联合概率分布，并通过独立事件的组合来求得该先验概率，再通过最大化先验概率来得到整个模型的最终输出。

半朴素贝叶斯模型通过组内容量的大小来协调算法复杂度与准确度之间的关系。S的量级与算法准确度成正比，而L的量级与算法复杂度成正比。

## 2.2 随机蕨回归算法原理

随机蕨回归器是由Dollár等人在2010年首次提出。该算法将半朴素贝叶斯模型进行了应用到了回归器的设计上。随机蕨回归器采用了级联式回归的设计方式，由多个回归器共同来约束输出结果。更具体地说，随机蕨级联回归框架的每一级回归器都是对上一级回归器所做出的结果的一个修正。该算法不同于之前流行的大部分算法，采用参数化的约束来控制输出结果；在随机蕨回归算法中，最终的输出是所有的训练集的一种线性组合，由此保证了在不依赖参数化约束模型的情况下，依然能够控制输出结果。

以图像的特征点定位的应用为例，随机蕨回归器在特征选择上，采用的是像素差异特征。选取的是随机生成的N对像素对的强度差，组成原始的特征向量。其形式如下：

X = {x1, x2, … , xn}

通过随机采样，将特征的数量缩减至S个，并且在每次迭代时生成S个划分阈值，将S个经过随机采样的特征映射成S个二进制位。映射规则如下：

在训练时，通过将S个二进制位的映射，可以把S维特征映射为0-2S-1之间的一个整数d，将该训练集存于序号为d的桶中，完成第一个回归器的训练。随后用第一个回归器预测出所有训练集图片的特征点，并将该误差作为第二级回归器的回归目标，以此类推，完成所有回归器的训练。

在预测时，对于每个回归器，对输入图片进行取“像素差异特征”后，再进行随机采样，并将映射的到的序号为d的桶中的所有回归目标中，选取一个残差最小的回归目标，作为本次回归器的输出。形式化表示如下：

其中，函数d(Si,S)表示第一个回归目标与序号为d的桶中所有回归目标的残差。

在预测时，对于整个级联回归器，将输入图片依次通过所有的回归器，将每个回归器输出的结果进行累加，即为整个级联回归器最终的输出。形式化表示如下：

RFinal = R0 + R1 + ... + RS

随机蕨回归器的训练，预测速度都非常快，并且由于采用了级联的方式约束结果，随着级联回归器数量的增加，回归器输出的误差粒度逐渐细化，可以得到比较准确的输出结果。由于级联框架中的每一级回归器的回归目标为上一级回归器与真实形状之间的残差，因此采用随机蕨这种较弱的回归器亦能取得较好的回归效果，并且能够极大地提高回归速度。

同样得益于级联回归的设计，经过Dollár等人证明，随机蕨回归器可以在每一个阶段以指数级的速度逐步收敛，因此该算法具有非常强的的鲁棒性，可以在不同的输入的情况下得到稳定的输出结果。

## 2.3 显式形状回归算法原理

显式形状回归（Explicit Shape Regression）算法由Xudong Cao等人于2014年提出。该算法也是以随机蕨回归算法为基础，并且在全局特征选取，整体架构设计，以及初始化数据等方面进行了大量改进，使其能够非常高效准确地进行人脸特征点定位。

在全局特征选取上，Xudong Cao提出了形状索引特征（Shape Indexed Feature）。形状索引特征相较于原始的随机蕨回归算法的随机像素差异特征，能够保持非常良好的语义一致性。更具体地说，随机像素差异特征是在全局坐标系上随机地取点对，并计算像素的强度差作为特征，各个像素对的坐标是以全局坐标系下的绝对坐标来表示的，不能根据输入图像的实际情况进行改变。形状索引特征同样是从像素差异特征改进而来：首先在全局坐标系下随机选取点对，而每个点对并非以绝对坐标的形式保存，而是保存为距离最近的特征点的相对坐标；在实际计算特征之时，再通过上述方式逆向映射会实际的坐标点。因此形状索引特征的各个坐标能根据不同的输入形状进行调整，从而保证特征的语义一致性，增强算法的准确率以及鲁棒性。

在整体架构的设计上，该算法创造性地使用了2层回归器的设计，其结构如下：

其中，高层次的回归器Rt被称为外层回归器，而低层次的回归器rk被称为内层回归器。与上一节提到的级联回归的思想类似，显式形状回归的每一层也采用级联式的设计：每一个回归器的输入是上一个回归器的输出，每一个回归器的回归目标是上一个回归器的的结果与真实形状的残差。但由于单个内层回归器的回归能力很有限，并不能很好地纠正上一层的误差，并且容易造成过拟合的现象。因此，Xudong Cao采用了两层回归器的结构：外层回归器形状索引特征的初始化，内层回归器使用该形状索引特征进行回归，即同一个外层回归器下属的内层回归器均使用相同的形状索引特征，这样能够提升训练与预测的速度，并且使得结果更加准确。

此外，显式形状回归还提出了初始数据扩增来优化回归结果。由于级联式的回归器的特点所致，对于每一个输入，都需要有一个初始化形状作为第一个回归器R0的输入。之前的方法无外乎采用所有训练集的均值形状或者选取某一个全局残差最小的形状作为初始形状输入。这会导致训练得到的回归模型的泛化能力不足，并且由于随机蕨算法的随机性，还会导致输出结果的不稳定（对于同样的输入，回归得到差距较大的输出形状）。Xudong Cao采用了初始化数据扩增的方法：在训练时，随机地采用多个形状作为初始形状进行回归，可以大幅度增强回归模型的泛化能力，有效避免过拟合；在预测时，将通过多个初始形状进行回归得到的结果的均值进行输出，能够提高输出结果的稳定性。

## 2.4 主流人脸数据库介绍

* BioID：BioID数据库由Oliver Jesorsky等人于2001年提出，该数据库包含了1521张人脸图片，每张人脸照片采用20个特征点进行标注，该数据库的人脸均由相同距离的正脸在实验室环境下采集而成。
* LFPW：LFPW（Labeled Face Parts in the Wild）由XX于2011年提出，中包含了1100张训练图片以及300张测试图片，每张人脸图片使用68个特征点来标注。LFPW人脸库中包含大量的复杂表情、姿势、光照等条件下采集的相片，对于人脸识别算法具有一定的挑战性。
* LFW87：LFW87数据库由XX于20008年创建，其中大量图片提取自LFW（Labeled Face in the Wild）数据库，该数据库中包含4002张训练图片以及1716张测试图片，每张人脸图片使用87个特征点来标注。
* Helen：Helen数据库由XX于2012年创建，该数据库采用2330张高分辨率的人脸图片，每张人脸图片使用194个特征点进行标注。由于该数据库的高分辨率、多特征点的特点，使得该数据库训练出来的模型能够更好地利用图像的细节信息，使结果更加准确。

## 2.5 Node.js介绍

Node.js是一款流行的高性能开源的网络服务与应用平台，使用JavaScript作为脚本语言。Node.js的构建基础是Google公司的V8引擎（Chrome JavaScript Runtime V8），而Google的V8引擎亦是目前全球效率最高的JavaScript解释器之一，这使得Node.js拥有非常好的并发性能以及非常快的执行速度。

Node.js能够提供一个使用事件驱动来实现异步开发的优秀解决方案。由于Node.js中绝大多数的API都是采用基于事件触发的异步调用模式设计的，非常适合进行高并发、高IO、低运算的网络开发。

Node.js旨在提供一种可伸缩、可扩展的高性能运行环境，以解决WEB服务器的性能瓶颈问题。由于传统的WEB服务器需要为每一个访问链接创建一个处理线程，会导致服务器只能响应十分有限的请求。而Node.js由于采用了单进程、单线程的事件驱动的设计方式，把每个新的请求加入事件循环（event loop）的轮询队列而非创建新线程，这个方式极大地减少了系统开销，提高了系统的处理并发连接的能力。

Node.js还是用模块化的设计来划分不同的功能包，这一特点也在很大程度上提高了开发效率。例如HTTP模块内封装了大量HTTP以及TCP/IP协议栈的操作函数，采用Node.js可以有效降低服务器的资源消耗，提高脚本的性能。

虽然Node.js与2013年才正式发布，但是已经得到了业界的一致认可，其中国际著名的社交网站LinkedIn、国际著名的开源项目托管网站GitHub和国内的阿里巴巴都把Node.js作为项目的重要组成部分。

## 2.6 MongoDB介绍

MongoDB是一款分布式的非关系型（NoSQL）数据库系统，该数据库系统于2009年发布，采用C++语言实现。MongoDB作为一款非关系型数据库，更加适合分布式应用的场景，与传统的关系型数据库相比，有如下优点：

* 高性能：相比于关系型数据库的查询缓存（针对每一条查询语句设置缓存信息，每当查询的目标发生改写，会导致缓存失败），MongoDB的缓存粒度更加细，每当数据改写后能够及时更新缓存，提高查询性能。此外，由于MongoDB并不支持关系型数据库的事务操作，可以进一步减少操作时间，提升性能。
* 易扩展性：MongoDB采用的是面向文档存储的数据模型，这是其非常容易在多个服务器之间进行数据分割；此外，由于MongoDB的各个集合之间取消了外键的约束，因此也方便数据进行扩展。
* 高灵活性：MongoDB采用键值对的形式进行数据存储，并不要求每个表项拥有相同的字段，可以灵活地存取数据而避免频繁地增加减少字段。

## 2.7 本章小结

本章节从半朴素贝叶斯模型开始，具体介绍了两种应用了该模型思想的级联式回归算法：随机蕨回归算法以及显式形状回归算法。推导了半朴素贝叶斯模型的数学模型，阐明了两个回归算法的历史以及详细的算法流程。除此之外，本章节还介绍了本项目所涉及的两个开源软件，分别是JavaScript运行平台Node.js以及非关系型数据库MongoDB。总的来说，本章节详细介绍了本项目所涉及的算法以及软件框架，为后续提出系统设计打下了良好的理论基础以及应用基础。

# 第三章 系统需求分析

在第一章中，已经说明过人脸活体验证系统对于信息安全的重要意义，即旨在一系列的人脸图片（序列帧）以及其他辅助信息，判别图片中的人脸是否为真实的活体，而非仿冒攻击的手段（比如照片、视频、三维模型等）。本章对人脸活体验证系统确定了项目的设计目的，并且分别进行功能性需求分析和非功能性需求分析。

## 3.1 系统角色划分

|  |  |
| --- | --- |
| 角色 | 职责或功能 |
| 用户 | 系统的用户，使用人脸活体验证系统进行验证的使用者 |
| 管理员 | 系统管理员，可以设置参数以及查询结果 |

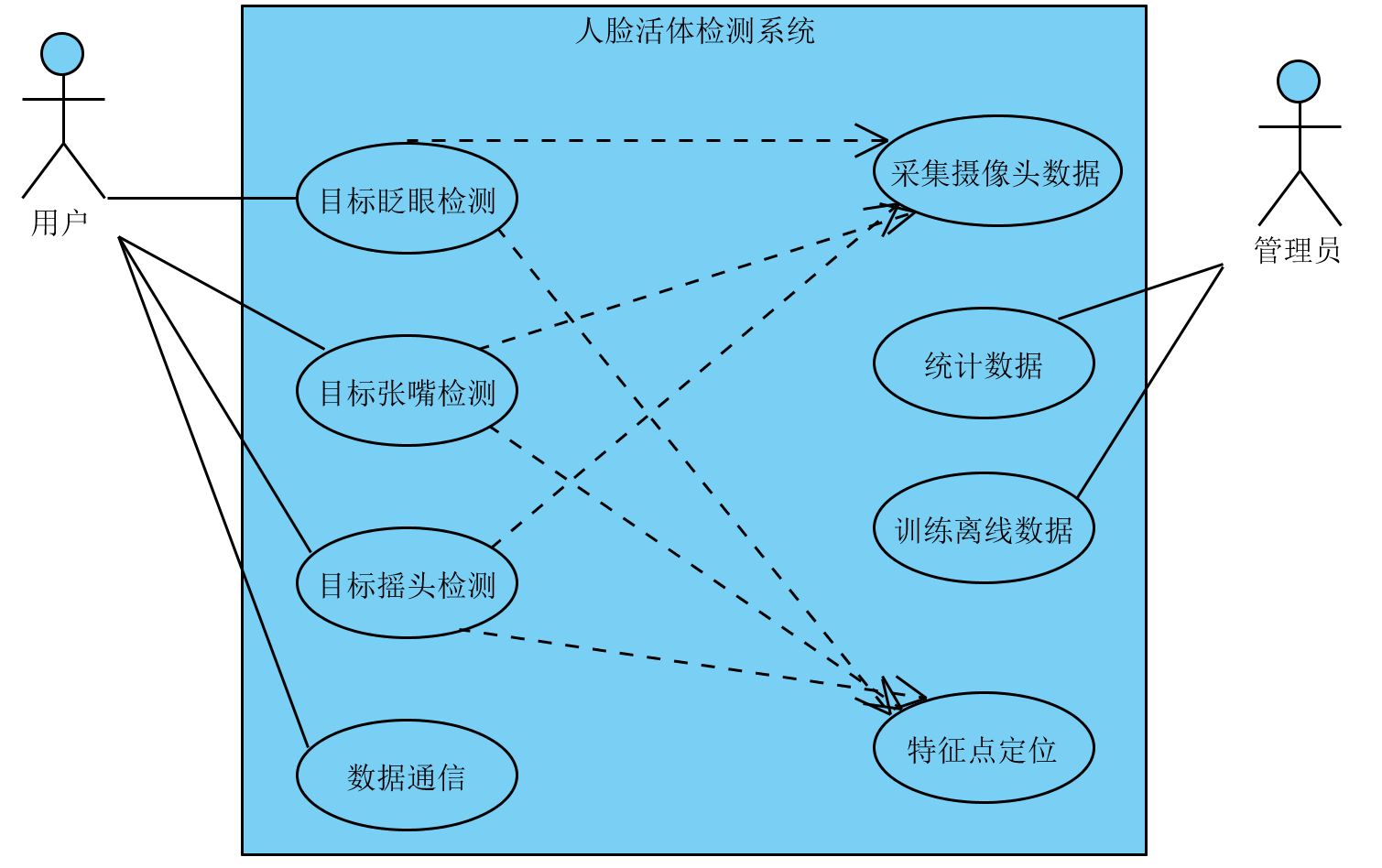
如表X所示，本系统的系统角色有两类，分别为用户和管理员。其中用户是系统前段的使用者，也是被验证者，系统与用户交互并采集用户的图像序列，进行人脸活体验证。管理员是系统的后端管理员，主要负责系统的参数设定，并且可以集中查看整个网络中的系统的验证结果。

## 3.2 系统功能性需求

本系统为一个交互式人脸活体检测系统，因此核心功能为验证输入人脸的有效性，同时还需要与使用者进行简短的交互。系统与使用者的主要交互方式为图形界面指示并伴随语音提示的方式。

从验证方式上，系统可以支持眨眼检测、张嘴检测以及摇头检测。选择此三种验证方式可以在最少的交互下达到比较优秀的识别率。

系统的主要功能性需求如下：离线训练数据、采集摄像头数据、目标眨眼检测、目标张嘴检测、目标摇头检测、数据通信。



人脸活体验证系统

### 3.2.1 离线训练数据

离线训练数据用例对应系统的离线训练数据的功能。由于系统的多个模块需要关联人脸特征点定位的功能。

系统能够自由地让管理员用户选择训练数据的类型，包括训练集的类型、特征点的数量以及位置、回归器的架构以及级联数量等属性；并且系统可以加载对应的训练数据，并且根据管理员所设置的参数组合进行回归模型的训练。

### 3.2.2 采集摄像头数据

采集摄像头数据的用例主要对应摄像头数据的采集以及保存。并且在接收到请求时，及时返回目标图像（或一段序列帧）。

该用例为系统功能单元，其被3个用例所依赖：目标眨眼检测用例、目标张嘴检测用例、目标摇头检测用例。

### 3.2.3 人脸特征点定位

人脸特征点定位的用例主要对应人脸特征点定位的功能。人脸特征点定位可以接受一个图片作为输入，返回一组特征点的坐标作为输出。

该用例为系统功能单元，其被3个用例所依赖：目标眨眼检测用例、目标张嘴检测用例、目标摇头检测用例。

### 3.2.4 目标眨眼检测

目标眨眼检测的用例主要对应人脸眨眼检测的功能。系统可以进行图形界面提示以及语音提示，指导用户进行眨眼动作，同时捕获序列帧。对于用户端采集的一系列序列帧，通过人脸特征点定位确定眼睛的位置，再建立状态机并检测出用户是否眨眼。

### 3.2.5 目标张嘴检测

目标张嘴检测的用例主要对应人脸张嘴检测的功能。系统可以进行图形界面提示以及语音提示，指导用户进行张嘴动作，同时捕获序列帧。系统对于用户端采集的一系列序列帧，通过人脸特征点定位确定嘴巴的位置，能够用过阈值判断检测用户是否张嘴。

### 3.2.6 目标摇头检测

目标摇头检测的用例主要对应人脸摇头检测的功能。系统可以进行图形界面提示以及语音提示，指导用户进行摇头动作，同时捕获序列帧。系统对于从用户端采集的一系列序列帧，通过人脸特征点定位，可以确定人脸的轮廓，再进行前景建模与背景建模，最后能够进行人脸摇头检测，通过阈值判断用户是否摇头。

### 3.2.7 数据通信

数据通信的用例主要对应客户端与服务器算通信的功能。系统能够与服务器进行版本校验、更新本地的数据文件、获取验证指令、返回验证结果。

### 3.2.8 统计数据

人脸活体验证系统为一个联网验证系统，所有的验证结果会存储在数据库中。系统需要提供统计数据并且输出的接口，管理员可以通过web页面的方式访问这些数据，并且进行统计归纳。

## 3.3 系统非功能性需求

上一节简要介绍了系统各个主要功能性需求，本届内容为系统的非功能性需求。软件系统的非功能性需求即软件系统为了满足用户的业务需求或者使用体验而必须满足的且在功能性需求之外的特性。软件系统的非功能性需求不仅仅影响着功能性需求的完整度，同时也影响着整个软件系统的质量。软件系统的非功能性需求贯穿着整个软件生命周期的软过程，结合本项目需求，从以下三个方面进行展开叙述：实时性、健壮性以及可维护性。

### 3.3.1 实时性

本系统设计的初衷是一个交互式的人脸活体验证系统，因此需要与用户进行交互并及时给予反馈。因此，反馈是否及时对用户体验的影响非常大，甚至会影响用户的交互动作能否正常完成。在实际应用中，所有的用户反馈均需要达到实时响应的标准，其中主视频的帧率不能低于12FPS。

### 3.3.2 高效性

本系统作为一个交互式的人脸活体验证系统，其主要应用场景为作为人脸识别系统的过滤器，且在大部分实际应用中，用户对于过于繁琐的交互动作以及长时间的等待是非常反感的。因此，在交互动作的设计上，既要考虑活体验证的成功率，也要兼顾用户体验。在实际应用中，每个交互动作的时间不能超过3秒，整体交互动作时间不能超过10秒。

### 3.3.3 健壮性

本系统作为一个活体验证系统，即需要直接面对各种形式的仿冒攻击。因此，健壮性是最基本的系统特性之一，因为只有保证了系统的健壮性，才能够进一步提升识别的正确率。更具体的说，系统需要应对各种形式的输入而不能崩溃，并且需要有良好的容错性。

### 3.3.4 可维护性

系统具有良好并且详尽的注释以及文档，因为在可以预见的未来，人脸活体验证的方式未来很有可能根据实际使用情况进行扩增，因此系统的检测功能将会越来越多，一个良好的系统架构不仅仅便于系统开发过程中的需求变更以及模块扩展，同时也方便了代码的后期维护以及重构。同时，系统还应具备跨平台的特性，保证用户在不同的操作系统下均能够部署以及使用，并且保证相同或者近似的用户体验。

## 3.4 本章小结

本章从系统的角色划分开始，着眼于系统的功能性需求以及非功能性需求，从两个方面对人脸活体验证系统进行了详尽的需求分析。在功能性需求方面，本章基于系统总体用例图主要阐述了系统的各大功能，并且对每个功能进行了细分描述。在非功能性需求方面，本章从用户体验以及软件工程规范的角度入手，并且对系统需要满足的各大特性进行展开描述。通过对项目进行需求分析，逐步细化项目所需要完成的目标，为后续章节提出系统的概要设计以及详细设计奠定了基础。

# 第四章 系统概要设计

在上一章节，已经详细介绍了人脸活体验证系统的功能性需求以及非功能性需求。本章节将从软件工程的角度提出对整个系统的概要设计，即根据系统的功能性需求以及非功能性需求，确定系统的整体框架以及模块分布。具体内容如下：

## 4.1 系统整体架构

本项目的开发过程采用自顶向下，逐步精化的开发模式。为了兼顾人脸活体检测系统的非功能性需求，系统将实现复杂度较高的模块置于客户机端，可以确定系统的客户机端需要承担较为复杂的功能以及较为繁重的计算量。因此，本项目采用C/S（客户机/服务器）架构进行开发。

相比近几年流行的B/S（浏览器/服务器）架构，虽然B/S架构拥有免安装、部署扩展容易、维护方便的优点；但C/S架构能够充分发挥客户机PC的处理能力，保证客户端的执行效率以及用户体验，并且在正常的检测过程中，客户机与服务器之间的通信次数以及通信量均在可控范围，因此C/S架构非常适合作为人脸活体验证的整体架构。

人脸活体验证系统

客户机端

服务器端

控制器模块

数据通信模块

图像采集模块

人脸特征点定位模块

眨眼检测模块

张嘴检测模块

摇头检测模块

模型模块

控制器模块

视图模块

系统整体架构图

如上图所示，客户机运行在运行在用户的内存空间内，作为一个独立的进程运行在用户态。客户机采用模块化设计，分成7个模块：控制器模块、数据通信模块、图像采集模块、人脸特征点定位模块、眨眼检测模块、张嘴检测模块以及摇头检测模块。其中的核心模块为控制器模块；数据通信模块、图像采集模块以及人脸特征点定位模块为内部调用模块；眨眼检测模块、张嘴检测模块以及摇头检测模块为主要功能模块。为了保证程序执行的高效率，各个模块均采用单独线程的方式进行执行，而线程间通信以及调用通过控制器模块完成。为了满足功能模块的可扩展性，各个功能模块与控制器使用的设计模式为简单工厂模式，通过简单工厂模式隐藏各个功能模块内部的实现细节。

在服务器端，采用成熟的web应用开发框架：基于Node.js的express框架，该框架提供丰富的http快捷方法以及可以任意排列组合的connect中间件，可以迅速实现一个集健壮、友好等优点于一身的web应用。服务器端的架构采用MVC结构，使数据与业务逻辑相分离，简化了开发难度，提高了开发的速度，增强了系统的鲁棒性。

## 4.2 系统功能模块设计

系统整体架构图展示了人脸活体验验证系统的总体模块布局，本节将对各个模块之间的协作方式进行描述。

服务器端

客户机端

图像采集模块

人脸

特征点

定位模块

眨眼检测模块

张嘴检测模块

摇头检测模块

控制器模块

数据通信模块

控制器模块

模型模块

视图模块

系统功能模块间通信示意图

客户机端的控制器模块为客户机的逻辑核心，负责控制客户机端的其他模块并且进行信息交互。控制器模块在客户机端运行之后会首先启动，采用独立线程的方式运行，逐个调用其他模块，并且根据各个模块的返回值，最终输出活体验证的结果。

数据通信模块是客户机端与服务器端通信的模块，主要功能为校验系统版本、获取人脸特征点配准训练文件、获取随机验证指令、返回验证结果以及其他附属信息等。数据通信模块在客户机端启动之时，为之提供初始化数据；在验证完毕以后，向服务器端返回验证结果。

眨眼检测模块负责完成对于输入的人脸进行眨眼检测的功能。该模块接受控制器模块的调用指令，使用独立线程的方式运行。在执行阶段，该模块调用图像采集模块获取用户输入的图像，调用人脸特征点定位模块进行定位，最后对人脸的眼部区域进行眨眼检测，并且向控制器模块返回检测结果。

张嘴检测模块负责完成对于输入的人脸进行张嘴检测的功能。该模块接受控制器模块的调用指令，使用独立线程的方式运行。在执行阶段，该模块调用图像采集模块获取用户的输入图像，再调用人脸特征点定位模块进行定位，最后使用人脸特征点定位的结果进行张嘴检测，并且向控制器模块返回检测结果。

摇头检测模块负责完成对于输入的人脸进行摇头检测的功能。该模块也是接受控制器模块的调用指令，同样使用独立线程的方式运行。在执行阶段，该模块首先通过调用图像采集模块的接口获得人脸的输入图像，并且调用人脸特征点定位模块的接口进行人脸特征点定位，最后通过人脸特征点的坐标划分图像区域，进行摇头检测，并且将结果返回给控制器模块。

人脸特征点定位模块负责完成的功能为人脸图片的特征点定位。该模块接受3个功能模块的调用指令。在程序执行阶段，对于输入的图片，首先进行人脸区域的粗定位，再基于粗定位的人脸区域进行人脸特征点定位，并且将定位的结果返回给调用指令的发起模块。

图像采集模块的主要功能为，将通过摄像头采集到的图像序列提交给功能模块。该模块接受3个模块的调用指令，并且以独立线程的方式运行。在程序执行阶段，该模块以最高帧率获取摄像头所采集的图像，并且保存；等待功能模块的调用指令，并且把图像给调用指令的发起模块。

在服务器端，控制器模块是MVC架构的核心模块，用以处理客户机端发来的请求并且协调模型模块以及视图模块进行响应。模型模块主要负责将服务器端的数据结构与数据库中的表项进行映射，并且封装业务逻辑接口，供控制器模块调用。视图模块由一系列模板页面组成，该模块的主要功能为，接受控制器模块的实例化参数，生成html页面，最后将目标页面返回给控制器模块。

## 4.3 本章小结

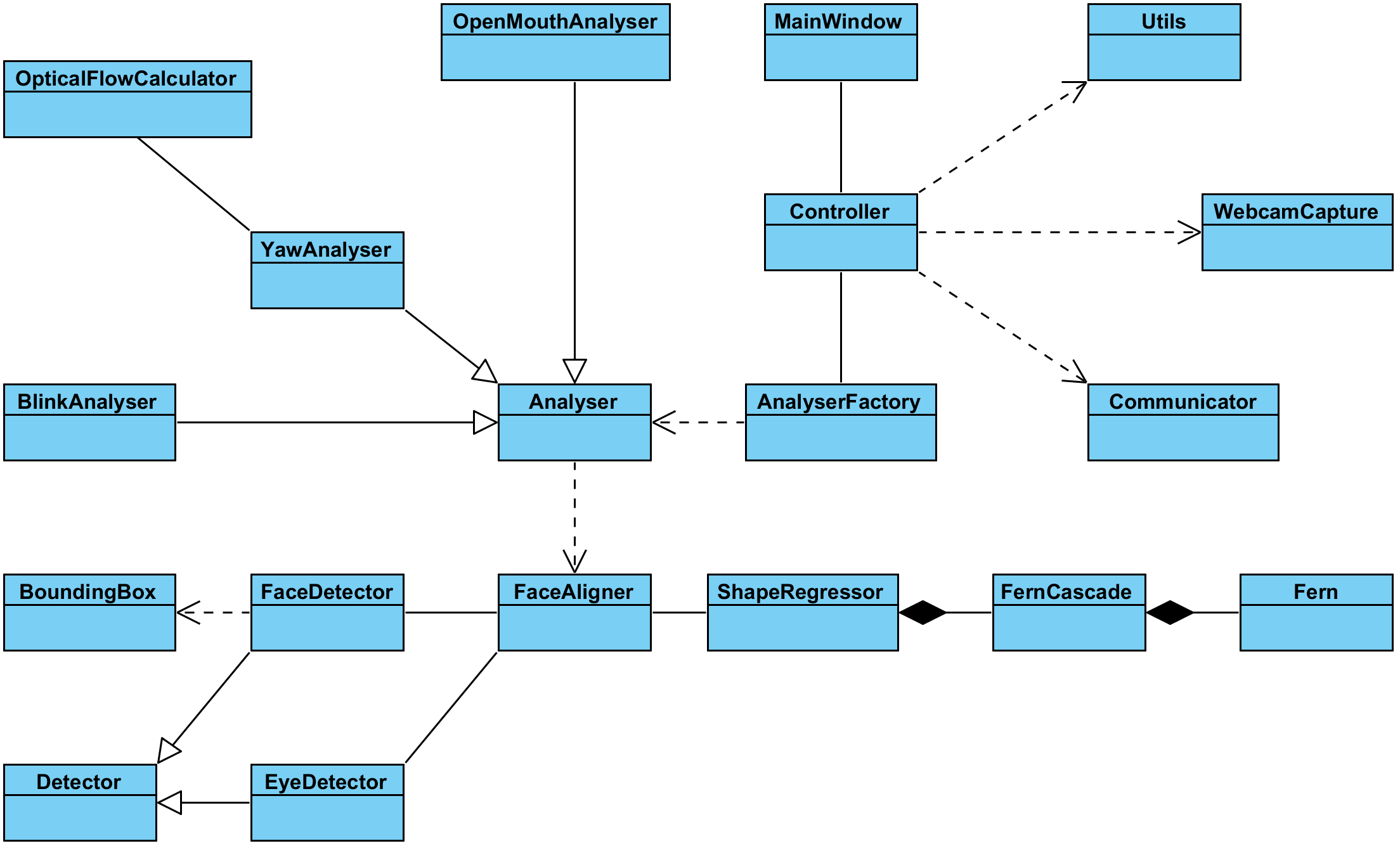
本章节提出了人脸活体验证系统的总体架构以及各个功能模块的划分。首先通过分析项目的需求，确定本项目采用C/S架构，并且使用系统功能模块图对于服务器端以及客户机端的功能模块进行划分。随后使用系统功能模块间通信示意图来明确各个功能模块之间的层次关系、调用顺序以及接口规范等，为后续提出系统的详细设计打下了基础。

# 第五章 系统详细设计与实现

在前两章提出的人脸活体验证系统的需求分析以及概要设计的基础上，本章将会对整个系统及其各个模块进行详细设计。

## 5.1 客户机端详细设计与实现

根据系统总体架构图的设计，进行细化设计，通过抽离出各个模块的通用部分，提出客户机端的系统整体类图，如图所示：



客户机端系统整体类图

客户机端的各个主要功能模块均采用独立线程的方式进行设计，线程间的通信采用Qt的信号槽（Signal and Slot）机制进行。其中控制器模块由3个类组成：MainWindow类、Controller类以及Analyser类；数据通信模块由Communicator类组成；图像采集类由WebcamCapture类构成；人脸特征点定位模块由8个类构成：BoundingBox类、Detector类、FaceDetector类、EyeDetector类、FaceAligner类、ShapeRegressor类、FernCascade类以及Fern类。眨眼检测模块由2个类组成：Analyser类以及BlinkAnalyser类；张嘴检测模块由2个类组成：Analyser类以及OpenMouthAnalyser类；摇头检测模块由3个类组成：Analyser类、YawAnalyser类以及OpticalFlowCalculator类。另外Util类作为全局工具类，包含整个项目的常用函数以及共享变量。

整体设计上，各个模块之间采用高内聚低耦合的设计理念：不同的类的实例之间的函数调用仅限于用同一模块之间，跨模块的信息传递均采用异步的信号槽机制，隐藏了各个模块内部的实现细节，提升了代码的可维护性。

由于系统对于各主要功能模块均采用独立线程方式运行，因此在设计之初就着重考虑了未来功能模块数量的扩展需求，因此引入简单工厂模式来设计控制器模块与各功能模块调用关系：通过建立一个工厂类，来封装所有的功能模块并且提供实例化接口，每个功能模块均采用一致的启动接口；当一个功能模块需要被调用之时，控制器模块通过工厂类提供的实例化接口获得功能模块的实例化对象，再调用该对象的启动接口。通过应用该设计模式，极大地降低了由于添加功能模块而导致的对于现有系统进行修改的工作量，提高了系统的可扩展性。

### 5.1.1 控制器模块设计与实现

#### 5.1.1.1 控制器模块的类结构设计

控制器模块为整个客户机端的核心模块，客户机端的所有其他模块均由控制器负责调度，各个模块的运行结果也汇总至此。控制器模块由3个类组成：MainWindow类、Controller类以及Analyser类。该模块的类图如图所示：



客户机端控制器模块类图

其中，MainWindow类为与用户进行图形界面交互的主窗口类，是客户机端向用户传递信息的主要方式。Controller类为控制器模块的主体，负责响应MainWindow类交互请求、调用其他功能模块的主线程函数、处理返回结果等功能。AnalyserFactory类为简单工厂模式的工厂类，负责根据不同的请求，实例化对应的Analyser类。

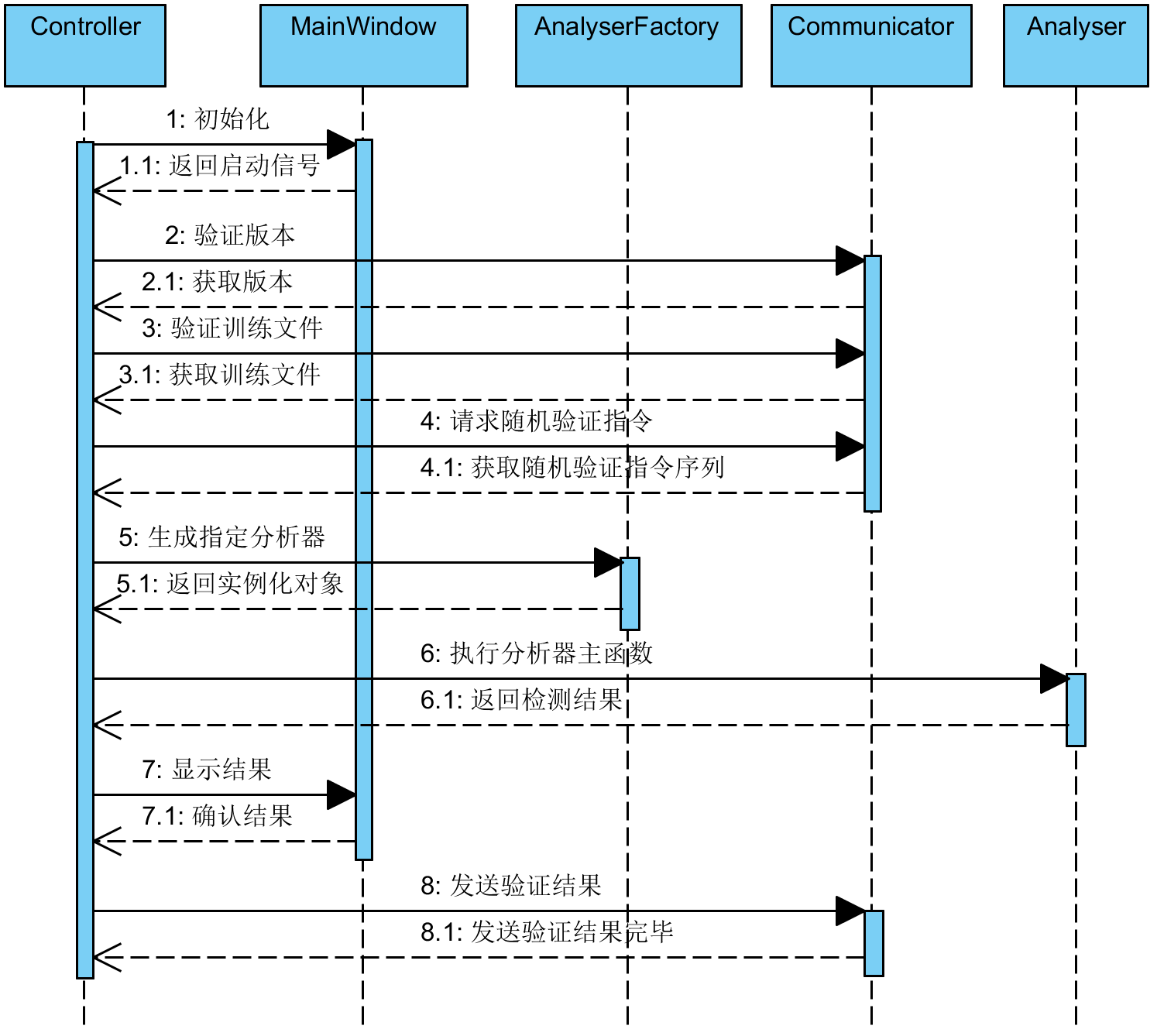
由于Controller类采用单例模式设计，因此ptr2\_controller为指向Controller类的实例化对象的指针；ptr2\_analyserFactory为指向生成Analyser类的实例的工厂类的指针；analyserOrder为通过数据通信模块获取到的随机验证指令序列；currentAnalyserIndex为当前执行的Analyser类的实例的计数；analyserVector为通过AnalyserFactory所生成的实例对象的序列；analyserThread为当前执行的功能模块的主线程；webcamThread为图像采集模块的主线程；faceAlignerThread为人脸特征点定位模块的主线程；ptr2\_opticalFlowCalculator为指向OpticalFlowCalculator类的指针；opticalFlowCalculatorThread为异步光流计算的主线程。在接口设计中，getInstance()方法为Controller单例模型的实例化接口；start()方法为Controller主线程的入口函数；webcamStart()方法为图像采集模块的主线程的入口函数；startNextAnalyserSlot()方法为进入下一轮检测循环的接口，该方法在每个Analyser实例的主函数执行完毕后被调用；receiveAnalyserResultSlot()方法为各个Analyser的实例的运行结果的返回接口；receiveSliderPercentage()方法为Controller与MainWindow中的滑块控件交互的接口。

在MainWindow中，ptr2\_ui为指向UI配置结构体的指针；ptr2\_webcamCaptureinter为指向WebcamCapture类的实例化对象的指针；ptr2\_controller为指向Controller类的实例化对象的指针。在接口设计中，startToRunButtonClicked()方法为用户交互按钮状态改变的接口；updateImage()方法为更新UI中的主视频窗口的接口；receiveSliderPercentage()方法为接受Controller改变UI界面中的控制滑块的参数的接口。

在AnalyserFactory中，analyserType为保存着验证模块信息的序列，每一条记录对应着一个Analyser的实例标识；在接口设计中，createAnalyser()方法为类静态方法，供实例化特定Analyser时调用。

#### 5.1.1.2 控制器模块核心处理流程

Controller类于客户机端启动伊始便首先实例化。为了防止由控制器模块所控制的其他模块的大量计算所造成的“卡顿”现象，Controller类的实例以独立线程的方式运行。其调用关系如图所示：



客户机端控制器模块时序图

如图所示，Controller类是客户机端的入口，在Controller类实例化完毕后，MainWindow类被实例化并且向Controller发送启动信号。随后，Controller调用数据通信模块的Communicator进行版本验证、训练文件初始化以及获取随机验证指令。在数据通信模块初始化完毕之后，Controller实例化AnalyserFactory，获取各个分析器的实例并且执行分析器的主函数，最后通过MainWindow以及Communicator返回验证结果。其主函数活动图如图所示：



客户机端控制器模块主线程活动图

### 5.1.2 数据通信模块设计与实现

#### 5.1.2.1 数据通信模块的类结构设计

数据通信模块是客户机端与服务器端通信的接口模块。该模块由1个类构成：Comminicator类。该模块的类图如图所示：



数据通信模块类图

在Communicator类中，ptr2\_controller为指向Controller类的实例化对象的指针；currentVersion为目前客户机端的版本号；trainingFileMD5为目前在客户机端的训练文件的MD5散列码；analyserOrder保存着从服务器端获取的随机验证指令序列。在接口设计中，verifyVersion()方法是与服务器端进行版本验证的接口；verifyTrainingFile()方法为与服务器端进行训练文件一致性检测的接口；requsetAnalyserOrder()是从服务器获取随机验证指令序列的接口。

#### 5.1.2.3 数据通信接口设计

出于设计与实现的便捷性考虑，客户机端与服务器端之间的通信采用Json作为数据交换的格式。详细接口设计如下：

**{**

"type" **:** "request"**,** //消息类型

"requestType" **:** "verifyVersion" //请求类型

**}**

**{**

"type" **:** "response"**,** //消息类型

"responseType" **:** "verifyVersion"**,** //响应类型

"version" **:** 12 //版本号

**}**

验证系统版本的Json格式

**{**

"type" **:** "request"**,** //消息类型

"requestType" **:** "verifyTrainingFile" //请求类型

**}**

**{**

"type" **:** "response"**,** //消息类型

"responseType" **:** "verifyTrainingFile"**,** //响应类型

//服务器端训练文件的MD5散列码

"trainingFileMD5" **:** "1c0a9bc79c2ce0dd74373a5912e35a76"

//服务器端训练文件的URL

"trainingFileURL" **:** "http://172.21.8.186:8000/trainingFile"

**}**

验证人脸特征点定位的训练文件的Json格式

**{**

"type" **:** "request"**,** //消息类型

"requestType" **:** "getAnalyserOrder" //请求类型

**}**

**{**

"type" **:** "response"**,** //消息类型

"responseType" **:** "getAnalyserOrder"**,** //响应类型

//作为服务器端标识一个请求的ID

"ID" **:** 6235**,**

"analyserCount" **:** 3**,** //随机验证指令的长度

//随机验证指令序列

"analyserList" **:** **[**"yaw"**,**"openMouth"**,**"blink"**]**

**}**

获取随机验证指令序列的Json格式

**{**

"type" **:** "request"**,** //消息类型

"requestType" **:** "sendResult"**,** //请求类型

"ID" **:** 6235**,** //作为服务器端标识一个请求的ID

"result" **:** true**,** //验证结果

//视频截图的base64编码

"base64Image" **:** "data:image/png;base64,iVBQmCC......"

**}**

**{**

"type" **:** "response"**,** //消息类型

"responseType" **:** "sendResult"**,** //响应类型

"ID" **:** 6235**,** //作为服务器端标识一个请求的ID

"flag" **:** "received"

**}**

发送验证结果的Json格式

其中，在由于在验证人脸特征点定位的训练文件阶段，为了保证结果的准确性，需要人脸特征点定位算法所使用的训练文件的一致性，因此作为一种简便快速的验证方法，引入了MD5散列码来验证服务器端的训练文件与客户机端的训练文件的一致性。

在获取随机验证指令序列阶段，可能会出现多个客户机端同时从服务器端获取随机验证指令序列的情况，因此在Json消息中，设置了ID字段来区分不同的请求，保证服务器端可以追踪所有请求的状态。

在发送验证结果阶段，客户机端需要将一张视频截图连带验证结果一起发送至服务器端。出于统一IO接口以及节约服务器系统资源的角度考虑，采用Base64算法将图片编码后，也通过Json消息进行发送；服务器端在收到该Json消息后，通过将Base64的编码字段进行解码，即可得到原图像。

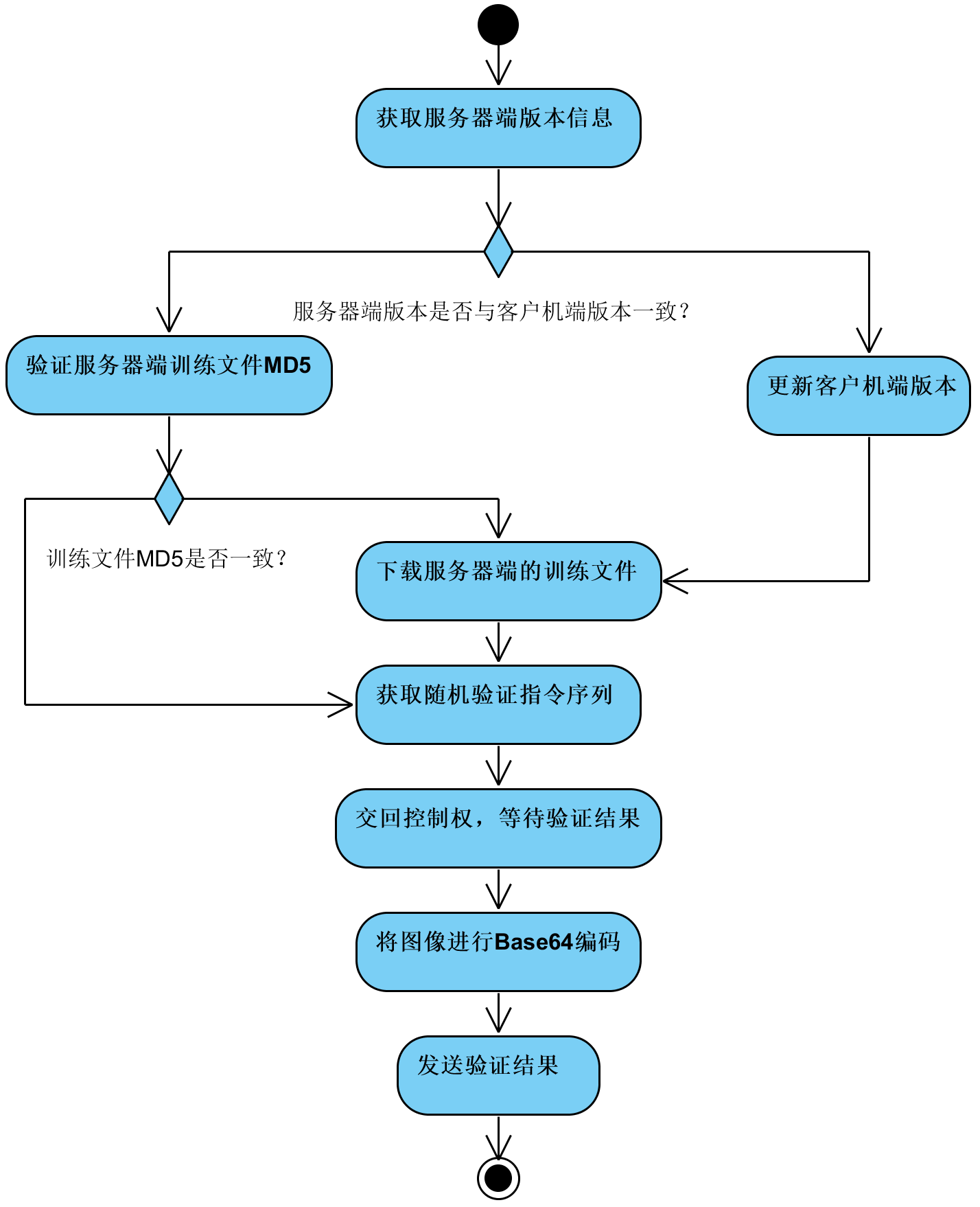
#### 5.1.2.2 数据通信模块核心处理流程

Communicator类由客户机端的控制器模块的Controller负责实例化，与服务器端进行数据通信。该模块的调用关系如图所示：



数据通信模块时序图

在实例化完成后的初始化阶段，该模块会与服务器端进行同步通信，并且完成系统的参数初始化，最后将控制权交回给控制器模块。在发送验证结果阶段，该模块接受从控制器模块传来的参数并将其发送至服务器端，完成数据的传输。其活动图如图所示：



数据通信模块活动图

### 5.1.3 图像采集模块设计与实现

### 5.1.4 人脸特征点定位模块设计与实现

### 5.1.5 眨眼检测模块设计与实现

### 5.1.6 张嘴检测模块设计与实现

### 5.1.7 摇头检测模块设计与实现

## 5.2 服务器端详细设计与实现

### 5.2.1 模型模块设计与实现

### 5.2.2 控制器模块设计与实现

### 5.2.3 视图模块设计与实现

## 5.3本章小结

# 第六章 系统部署与测试

## 6.1 系统部署

## 6.2 系统功能测试

## 6.3 系统性能测试

## 6.4 本章小结

# 第七章 总结

## 7.1 论文总结

## 7.2 个人收获

## 7.3 改善空间

# 参考文献

1. Geoffrey I. Webb, Janice R. Boughton, Zhihai Wang. Not So Naive Bayes: Aggregating One-Dependence Estimators. Machine Learning(ML), 58, 5–24, 2005 提出半朴素贝叶斯定理
2. Piotr Doll´ar, Peter Welinder, Pietro Perona. Cascaded Pose Regression. IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2010 提出随机蕨回归器
3. Mustafa Özuysal, Michael Calonder, Vincent Lepetit, Pascal Fua. Fast keypoint recognition using random ferns. IEEE Transactions on Pattern Analysis & Machine Intelligence(PAMI), 2010 首次提出random fern, 提出随机蕨分类器
4. 蒋秀鹏. 基于NodeJS的数字标牌系统的设计与实现[硕士学位论文]. 天津: 南开大学, 2014 参考了一些Node.js的概念
5. 王越. 基于nodejs的微博系统的设计与实现[硕士学位论文]. 成都: 电子科技大学, 2014 参考了一下nodejs以及mongodb的概念
6. Xudong Cao, Yichen Wei, Fang Wen, Jian Sun. Face Alignment by Explicit Shape Regression. International Journal of Computer Vision(IJCV) 107:177–190, 2014 提出了ESR算法
7. Oliver Jesorsky, Klaus J. Kirchberg, Robert W. Frischholz. Robust Face Detection Using the Hausdorff Distance. Third International Conference on Audio- and Video-based Biometric Person Authentication, 2001 提出了BioID数据库
8. Peter N. Belhumeur, David W. Jacobs, David J. Kriegman, Neeraj Kumar. Localizing Parts of Faces Using a Consensus of Exemplars. IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2011 提出了LFPW数据库
9. Lin Liang, Rong Xiao, Fang Wen, Jian Sun. Face Alignment Via Component-Based Discriminative Search. European Conference on Computer Vision(ECCV), 2008 提出LFW87数据库
10. Vuong Le, Jonathan Brandt, Zhe Lin, Lubomir Bourdev. Interactive facial feature localization. European Conference on Computer Vision(ECCV), 2012 提出Helen数据库
11. 仝义明, 黄蔚, 李戴维. 基于MongoDB的信息集成系统的设计与实现. 信息技术: 1009-2552(2015)02-0125-05, 2015 参考了MongoDB的简介
12. Iain Matthews, Simon Baker. Active Appearance Models Revisited. International Journal of Computer Vision(IJCV), 60(2), 135–164, 2004 AAM参考文献1
13. Patrick Sauer, Tim Cootes, Chris Taylor. Accurate Regression Procedures for Active Appearance Models. British Machine Vision Conference(BMVC), 2011 AAM参考文献2
14. Jason Saragih, Roland Goecke. A Nonlinear Discriminative Approach to AAM Fitting. International Conference on Computer Vision(ICCV), 2007 AAM参考文献3
15. Timothy F. Cootes, Gareth J. Edwards, Christopher J. Taylor. Active Appearance Models. IEEE Transactions on Pattern Analysis and Machine Intelligence,23(6), 681–685, 2001 AAM参考文献4
16. David Cristinacce, Tim Cootes. Boosted Regression Active Shape Models. British Machine Vision Conference(BMVC), 2007 基于回归的传统算法1
17. Michel Valstar, Brais Martinez, Xavier Binefa. Facial Point Detection using Boosted Regression and Graph Models. IEEE Conference on Computeer Vision and Pattern Recognition, 2010 基于回归的传统算法2
18. Herbert Bay, Tinne Tuytelaars, Luc Van Gool. SURF: Speeded Up Robust Features. European Conference on Computer Vision(ECCV), 2006 提出了一种基于SFM（Structure from Motion）的活体检测模型
19. Klaus Kollreider, Hartwig Fronthaler, Josef Bigun. Non-intrusive liveness detection by face images. European Conference on Computer Vision(ECCV), 2009 利用光流来分析人脸各部位的移动量，最终进行活体检测
20. 孙霖. 人脸识别中的活体检测技术研究[博士学位论文]. 杭州：浙江大学, 2010 提出了利用人脸识别进行多模活体验证的概念
21. 杨健伟. 面向人脸识别的人脸活体检测方法研究[硕士学位论文]. 北京：北京邮电大学, 2014
22. 曹瑜. 活体人脸检测技术研究[硕士学位论文]. 北京：北京工业大学, 2014
23. 刘华成, 人脸活体检测关键技术研究[硕士学位论文].  宁波：宁波大学, 2014

# 致谢

|  |  |  |  |
| --- | --- | --- | --- |
| 仿冒人脸输入方式 | 眨眼检测 | 张嘴检测 | 摇头检测 |
| 照片 | 未通过 | 未通过 | 未通过 |
| 三维模型 | 未通过 | 未通过 | 通过 |
| 高分辨率视频 | 通过 | 通过 | 通过 |
| 高分辨率面具 | 通过 | 未通过 | 通过 |

可以发现，在