**低质量二维码的图像增强及快速识别**

摘 要

此项目主要用于建立一个快速增强二维码图像的系统,从而增加二维码解码率,减少解码时间。本系统是一个由卷积神经网络构成的自动编码器,通过 TensorFlow 工具集在自生成的图片上进行训

练。它将具有较差分辨率和质量的图像作为输入,并输出相同二维码的清晰,无噪声的图像。然后，CNN 的输出可用于利用任何优选的解码算法解码二维码的数据。该系统已部署在树莓派 3B+上,并在不同的模拟中进行评估,使低质量二维码图像的正确解码成功率提高约100%。

（1）通过CNN，机器学习的优势得以利用，因为图像模糊和图像噪声是图像采集过程中常见的失真。在本文中，我们使用深度CNN图像分类器系统地实现图像失真的效果。第一，then？

（2）在预估了成功率的前提下，本文实现了一种CNN密集层二维重建系统。同时，为了机器学习，本文提出了自动编码的图像。通过应用网络模型的失真变换，从干净的图像中创建人造的噪声图像，可以实现这一点。

实验表明，在缺乏纹理，运动模糊和重复内容等具有挑战性但实际的场景中，该系统在数据集和Keras的CPU（中央处理单元）计算上以25 FPS的速度运行时，具有出色的重建性能。

**关键词：**二维码，二维条，图像增强，CNN，计算机视觉，机器学习数据集，优化

**Image enhancement and fast recognition of low-quality QR codes**

**ABSTRACT**

This work aims to construct a fast system to enhance the quality of QR code images to improve the decoding success rate, hence decreasing the time required to read the data contained in 2D bar codes. The system is a Convolutional Neural Network (CNN) with a structure of an auto-encoder, trained with the TensorFlow toolkit on self-generated images. The algorithm takes an image with a poor resolution and quality as an input and outputs a clear, noise-less picture of the same QR code. The output of the CNN can then be used to decode the QR code's data with any preferred decoding algorithm. The system has been deployed on a Raspberry Pi 3B+ and evaluated in different simulations obtaining an increase in the success rate of about +100% of the correct decoding of low-quality QR code images.

（1）Through a CNN, the advantages of machine learning are utilized, as image blur and image noise are common distortions during image acquisition. In this paper, we systematically implement the effect of image distortions using a deep CNN image classifier. First,

（2）this paper realizes a CNN dense layer 2D reconstruction system, given the estimated success rate. Meanwhile, for the sake of machine learning, this paper proposes auto-encoded images. This is achieved by creating artificial noisy images from clean images through the application of distortion transformations of the network model.

The experiments reveal that the system has a superior reconstruction performance running on 25 FPS on a CPU (Central Processing Unit) computing of datasets and Keras, under challenging, yet practical scenarios including texture-less, motion blur and repetitive contents.

**Keywords**：QR Code, 2D bar, Image Enhancement, CNN, Computer Vision, Machine Learning Dataset, Optimization, CPU

目 录

[1 引 言 4](#_Toc515796651)

[1.1 课题研究背景与意义](#_Toc515796652) 4

[1.2 研究现状](#_Toc515796653) 5

[1.3 问题分析](#_Toc515796657) 8

[1.4 本文工作](#_Toc515796657) 8

[2 图像增强重建系统结构](#_Toc515796658) 10

[2.1 引言](#_Toc515796661) 10

[2.2 堆叠图像数据进行训练](#_Toc515796661) 10

[2.3 CNN模型 1](#_Toc515796661)2

[2.4 本章小结 1](#_Toc515796661)3

[3 视觉二维码联合优化实现 1](#_Toc515796670)4

[3.1 引言 1](#_Toc515796671)4

[3.2 定义符号和优化函数 1](#_Toc515796671)4

[3.3 稀疏去噪自动编码器 1](#_Toc515796671)6

[3.4 二维码预积分实现 1](#_Toc515796674)7

[3.5 平面约束推导 1](#_Toc515796674)9

[3.6 错误回环检测的判断方法](#_Toc515796674) 20

[3.7 二维码在标定方法](#_Toc515796674) 20

[3.8 本章小结](#_Toc515796674) 20

[4 实验测试与结果分析 2](#_Toc515796684)1

[4.1 数据集测试和对比 2](#_Toc515796685)1

[4.2 采集真实数据对比测试 2](#_Toc515796689)1

[4.3 运行效率测试与在线标定测试 2](#_Toc515796689)9

[4.4 本章小结](#_Toc515796689) 31

[5 结论和展望](#_Toc515796693) 32

[5.1 结论](#_Toc515796694) 32

[5.2 展望](#_Toc515796695) 32

[参考文献](#_Toc515796696) 33

[谢 辞](#_Toc515796697) 35

# 1 引 言

* 1. 课题研究背景与意义

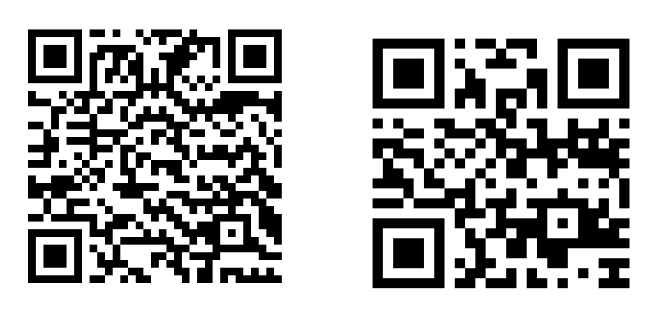
在某些领域，二维码已经取代了传统条形码,在手机上使用二维码可用以支持许多服务，例如运输，市场营销，广告，游戏，社交网络，支付方式，收据管理等。二维码被广泛运用是由于传统的条形码最多只能容纳20位数字，而二维码最多可以容纳7089个字符,提供了强大的多样性和可扩展性，这使得二维码的使用比条形码更具吸引力。从统计学上讲，二维码能够在传统条形码的大约十分之一的空间内对相同数量的数据进行编码。二维码的一个重要特性是它无需从特定角度进行扫描即可识别，这是由于二维码扫描仪角落含有三个特定方块符号和对齐块。二维码最初被汽车制造商用于跟踪零件，一段时间后，公司开始看到二维码的各种不同使用场景。如今二维码最普遍的商业用途是在电信行业中，随着智能手机的不断发展，尤其是在移动互联网访问领域。智能手机的普及极大地推动了二维码的普及，使其成为快速有效地向用户传达网址信息的绝佳工具。这也允许线下媒介（例如杂志，报纸，名片，公共交通工具，标志，T恤衫或任何其他可以接受二维码印刷的媒介）成为线上产品广告的载体。

图 1.1 二维码

如上所述，由于二维码具有大容量、小尺寸、高速扫描、抗损坏性和数据鲁棒性，二维码已被用于各种应用程序中。手机的二维码，例如全方位可读性和纠错能力。因此，智能手机采用二维码来支持许多服务，例如订票、付费和URL读取。传统的方法是从均匀的背景图像中提取条形码。然而，在实际情况中，手机拍摄图像的光线条件不均匀，并且图像的背景也不均匀。因此，传统方案在实际环境中会遇到一些特定的问题。本文将此问题称为“二维码提取问题”。假设用手机摄像头拍摄了背景和光线都不均匀的图像，那么本文的挑战？？？将重点放在卷积神经网络系统上，该系统旨在对二维码图像进行编码，以便为快速阅读器解码出更好的版本。去除不良光线条件并提取二维码以提高识别准确性。据我们所知？？使用TensorFlow技术和Keras构建了卷积神经网络，我们在其中设计了具有失真图像的局部数据集用以清洁图像

由于图像处理系统，虚拟现实和增强现实技术的广泛开发和应用，图像重建需要越来越高的准确性和效率，并且需要应对各种环境和运动状态以实现更干净的重建过程。二维码重构的高精度和高效率的实现是后续快速解码操作的基础。只有智能手机摄像头准确感应到二维码后，才能对二维码进行快速解码。

图 1.2 市场上的深度相机

（a）Python；（b）Tensorflow；（c）Keras

### 11.2 研究现状

当前的二维码重构技术发展迅速，在离线和在线应用场景中有许多不同的实现方法。这些通过图片获得三维模型的不同方法所采用的基本思路是相似的，主要包括以下两个方面：该方法基于自动编码器的结构。使用Dense层，因此使用自动编码器技术，我将自己的噪声图像数据集压缩为3d，我们将其可视化，这称为自动编码器设计。自动编码器是特定于数据的，不适用于完全看不见的数据结构。例如，受过数字训练的自动编码器不适用于字母。另一个限制是自动编码器的压缩是有损的。因此，它不能完美地还原原始信息。我使用嘈杂的图像数据集创建了自己的数据集，该算法的实现是从**嘈杂**的图像中重建**干净**？？的二维码。因此，整个数据集都由**嘈杂**的图像组成，所实现的算法是通过对干净图像应用失真变换来使用我的干净图像中的**嘈杂**数据集图像，在此过程中，我们将指导该研究在实现神经网络中的应用（如图1.3所示），通过了解我们提供给模型的输入数据的种类，我们可以使用该算法的设计来获得这样的结果。二维码重建项目“重建二维码”和“它生成清晰的图像”都使用CNN技术，并且可以通过从各个角度拍摄的图片来重建建筑物的三维模型。首先通过图片的匹配关系获得数据输入，同时计算稀疏点云，然后通过多视角匹配技术获得像素级对应关系，从而获得密集的点云模型，最后通过点云融合与滤波过程，获得高质量的三维纹理模型。这两个项目都可以通过图片获得高质量的三维模型。缺点是计算量大，需要长时间离线运行。

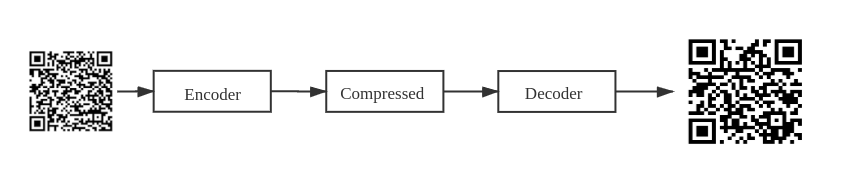


图1.3 自动编码器设计。

（2）编码器和解码器。

到目前为止，我们已经看到了开发人员算法来？？训练我们的神经网络，其中输入和输出是相同的，并且我们的模型的任务是在传递某种信息的同时尽可能地重现输入。回想我提到过的，我们希望我们的自动编码器足够灵敏以重新创建原始观测值，但又对训练数据不灵敏，从而使模型学习通用的编码和解码。开发泛化模型的另一种方法是稍微破坏输入数据，但仍将未破坏的数据作为目标输出。使用这种方法，我们的模型不能简单地开发一个用于存储训练数据的映射，因为我们的输入和目标输出不再相同。相反，该模型学习一个向量场，该向量场用于将输入数据映射到低维流形（从我之前的图形中回忆，流形描述了输入数据集中的高密度区域）；如果此歧管？？准确地描述了自然数据，则我们可以有效地“消除”所增加的噪声。

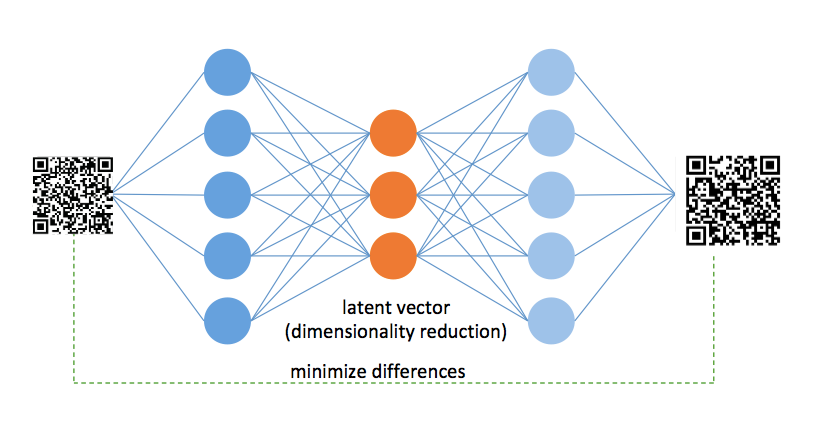


图1.4 自动编码器图

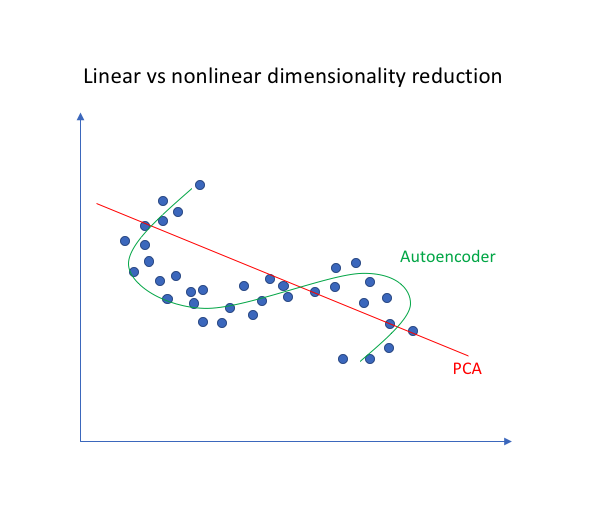
****

图1.5 非线性关系

对于较高维度的数据，自动编码器能够使用数据的复杂表示形式，该数据用于描述较低维度的观测值并相应地解码到原始输入空间中。完全自动编码器下没有明确的正则化术语——我们只需基于重建损失训练我们的模型。因此，确保模型不存储输入数据的唯一方法是确保充分限制隐藏层中的节点数量。对于深层自动编码器，我们还必须意识到编码器和解码器模型的容量。即使“瓶颈层”只是一个隐藏节点，我们的模型仍然可以记住训练数据，只要编码器和解码器模型具有足够的能力来学习一些可以将数据映射到索引的任意函数。鉴于神经网络的事实，我们希望我们的模型能够发现噪声图像数据中的潜在属性，因此重要的是要确保自动编码器模型不只是简单地学习一种有效的方法来记忆训练数据。与有监督的学习问题类似，我们可以对网络采用各种形式的正则化，以鼓励良好的泛化性质。这些技术将在下面讨论。

稀疏自动编码器为我们提供了一种引入信息瓶颈的替代方法，而无需减少隐藏层的节点数量。相反，我们将构造损失函数，以便对层中的激活进行惩罚。对于任何给定的观察，我们将鼓励我们的网络学习仅依赖于激活少量神经元的编码和解码。值得注意的是，这是进行正则化的另一种方法，因为我们通常会正则化网络的权重而不是激活。下方显示了一个通用的稀疏自动编码器，其中节点的不透明度与激活级别相对应。重要的是要注意，激活的训练模型的各个节点是数据相关的，不同的输入将导致通过网络激活不同的节点这一事实的结果是，我们允许我们的网络使各个隐藏层节点对输入数据的特定属性敏感。欠完善的自动编码器将为每次观察使用整个网络，而稀疏的自动编码器将被迫根据输入数据有选择地激活网络区域。结果，我们限制了网络存储输入数据的能力，而没有限制网络从数据中提取特征的能力。这使我们可以考虑网络的潜在状态表示和正则化。

## 1.3 问题分析

在用于通过视觉进行定位的实时三维重建系统中，当快速移动且特征点较少时，通过视觉对摄像机进行跟踪很容易失败。此时，映射和定位无法继续。惯性测量单元（二维码）成本低廉，不受较少的特征点和快速运动的影响，但数据嘈杂且积分具有严重的累积误差。视觉和二维码信息的融合通常使用基于扩展卡尔曼滤波和基于优化的方法的方法。滤波方法效率高但精度差；优化方法的估计较为准确，但效率较低，无法实现实时全局优化。同时，相机和二维码的转换矩阵在信息融合中非常重要。该矩阵通常是通过离线校准或在线初始化过程获得的。

在当前的卷积神经网络系统中，二维码和视觉信息的融合包括基于优化的方法和基于过滤的方法。具体而言，存在以下五种信息融合方法，它们具有以下问题：

（1）卡尔曼滤波：通过滤波融合视觉信息和二维码信息，缺点是这种信息融合不是最优估计，部分信息会丢失，得到的估计不够准确，无法实现全局一致性。

（2）不要反转二维码颜色？？？：当我们测试网络时，我们经常遇到这个错误。在某些设计方案中，黑色或深色背景，白色或较浅前景的二维码在某些设计方案中看起来不错，但无法使用所有二维码扫描仪应用程序进行扫描。许多用于扫描二维码的应用程序都希望该代码与背景形成深色对比，因此，如果您扫描前景为白色的代码，则并非所有人都能使用。在下面的示例中，我们将可视化哪些是好的，哪些则不该做。

（3）二维码为对比度提供足够的颜色：当选择白色背景和黑色前景以外的颜色时，请确保有足够的对比度。将彩色背景用作二维码完全可以，但是应确保与前景形成足够的对比度。如果颜色对比度太相似，则二维码阅读器可能很难在所有光线条件下轻松扫描二维码。有些扫描仪甚至根本无法读取对比度太低的二维码。

（4）模糊的二维码：二维码图像的分辨率足以使它在所有屏幕和打印格式上看起来都清晰。在没有足够分辨率的情况下将代码图像缩放到更大的尺寸会使您的二维码看起来模糊。模糊的二维码会使二维码阅读器应用难以检测到代码的边界并且难以扫描。

## 1.4 本文工作

本文提出了一种构建快速系统的方法，以提高二维码图像的质量，从而提高解码成功率，从而减少读取二维码中包含的数据所需的时间。该系统是具有自动编码器结构的卷积神经网络（CNN），使用TensorFlow工具包对自生成图像进行训练。输入分辨率和质量较差的图像，输出相同二维码清晰，无噪点的图像。然后，CNN的输出可用于通过任何首选的解码算法对二维码的数据进行解码。该系统已部署在Raspberry Pi 3B +上，并在不同的仿真中进行了评估，从而获得了正确解码低质量二维码图像的成功率提高了约100％，具体方面如下：

（1）在最初的视觉和网络信息全局优化方法中，计算复杂度与关键帧数量的平方成正比。随着场景的增加，关键帧的数量也会增加，因此能够在线计算。本文分析了全局优化中主要复杂度的来源，并通过推导来加快计算速度。改进FastGO [6]视觉优化加速方法，使用视觉二次统计和网络预测点[11]建立优化方程，使全局优化计算复杂度与关键帧数量成线性关系，整体计算复杂度大大降低 。实现视觉和网络信息的全局优化，大大加快了优化速度，从而实现了基于CPU的U的实时在线全局优化。

（2）卷积层具有一组矩阵，这些矩阵在称为卷积的过程中乘以上一层的输出，以检测某些特征，该特征可以是基本特征（例如，边缘，颜色等级或图案）或复杂的特征（形状，因此，这些矩阵称为过滤器或内核错误，不使用初始化方法，但将变换矩阵的优化耦合到全局优化函数中。紧密耦合的优化方法使得变换矩阵的估计值将越来越准确，避免了校准误差的积累。存在多种类型的池化层（最大池化，平均池化…），最常见的-今天-？？？是最大池化，因为它提供事务性方差（虽然差，但足以胜任某些任务），并且降低了网络的维数如此便宜？？？（没有参数）max-pooling层实际上非常简单，可以通过预定义一个过滤器（一个窗口），然后在输入中交换该窗口，并以该窗口中包含的最大值作为输出。

（3）反向概率可用于在提出一批数据后查找误差中每个权重的贡献，并且大多数良好的优化算法（SGD，ADAM…）都使用反向传播来找到梯度以进行环路检测的正确性使系统能够准确识别错误并纠正环回；当检测到环回时，可通过全局优化视觉和IMU来实现全局一致性，从而消除了错误的累积。

（4）通过不同的实验验证了算法的有效性和鲁棒性，并在不同的环境下测试了重构的效果。

与现有的二维重建系统相比，本文提出的方法具有以下优点：

（1）实现了高效视觉和CNN信息的全局优化而又不损失准确性，系统成功率高。

（2）紧密耦合的自校准可以得到准确的摄像机和IMU变换矩阵，不存在因校准不准确而引起的累积误差的问题。

（3）反向传播在最近几年一直做得很好，但是由于需要大量数据集，因此并不是一种有效的学习方法

我相信我们可以做得更好？？Is this line related with the (3) above?

本文分为以下几个部分：

第2章：系统结构，介绍系统的多线程结构，然后具体介绍跟踪模块和构建模块。

第3章：卷积神经网络的实现，介绍了深度学习的误差函数的定义，然后介绍了特征点约束，CNN约束和平面约束的最大池化矩阵导出过程。最后介绍了错误环回检测的判断方法和校准的实现。

第4章：实验测试和结果分析，分别介绍了该系统对数据集和您自己？？What`s this “您” used for收集的数据的影响，并介绍了与当前其他最佳卷积神经网络重构系统的比较。

第5章：总结与展望，总结了本文的主要工作，随后介绍了改进思路。

# 2 图像增强重建系统结构

## 2.1 引言

图像失真的类别为图像模糊和图像噪声。它们是由图像获取期间的各种问题引起的。例如相机未聚焦时，会发生散焦模糊。运动模糊则是由相机和物体之间的相对运动引起的，这在基于智能手机的图像分析中很常见。图像噪声通常是由不良的照明和/或高温引起的，这会降低相机内部的电荷耦合器件（CCD）的性能。当我们在实际应用中应用CNN分类器时，输入图像中可能会出现模糊和噪点的情况。这些降级将影响CNN分类器的性能。我们的工作对这个问题做出了一些贡献。

首先，我们研究图像失真对CNN分类器的影响。我们在输入图像的四种失真类型下检查CNN分类器性能：运动模糊，散焦模糊，高斯噪声以及它们的组合。其次，我们考察了两个减轻我们的二维码数据集给定的图像失真影响的方法。在一种方法中，我们用噪点图像重新训练整个网络。我们发现这种方法可以在对失真进行分类时提高准确性。但是，重新训练需要大量的训练数据集。

深度网络。受图像分类来源的启发，以另一种方法，我们用失真的图像微调了网络的前几层。本质上，我们调整CNN的低级滤波器以匹配特征失真的图像。先前的一些研究已经研究了图像失真的影响。针对CNN，提出了一种从深度信任网修改为处理噪声输入的新模型。他们在一个称为n-MNIST的嘈杂数据集上报告了良好的结果，该数据集

与原始MNIST数据集相比，它包含高斯噪声，运动模糊和对比度降低，我们也训练了自己的数据集。最近，道奇（Dodge）和卡拉姆（Karam）报告了由于几个CNN上的各种图像失真导致的性能下降。与这些作品相比，我们进行了统一的研究，以研究图像失真对手写数字分类和自然图像分类的影响。此外，我们研究了对噪声图像使用重新训练和微调来减轻效果的方法。在“干净”图像的分类中（即，没有失真），先前的一些工作试图将噪声引入训练数据。在这些作品中，其目的是使用噪声对模型进行正则化，以防止训练过程中过度拟合。相反，我们的目标是了解在失真图像分类中使用噪声训练数据的好处。我们的结果还表明，在某些条件下，使用噪声图像进行微调可能是一种有效且实用的方法。

如图2.1，该系统使用卷积神经网络结构。数据接收线程获取嘈杂的数据集图像的原始数据，以及我们从单个原始图像（未应用填充）训练CNN的时间，使用32个不同的3 x 3过滤器，此操作的输出是32个图像的集合，每个都来自原始图像中具有不同滤镜的卷积神经网络，其大小略有降低。然后，通过仅将2的最大值保持2乘以组成图像的平方，将这些矩阵的大小减小2倍。此操作就是所谓的最大池操作。这些图像然后用于通过16个不同的内核执行第二轮卷积。同样是3 x3。也是3 x3。结果通过第二个最大池操作图像，并且优化线程基于特征点约束和匹配平面约束执行全局优化。

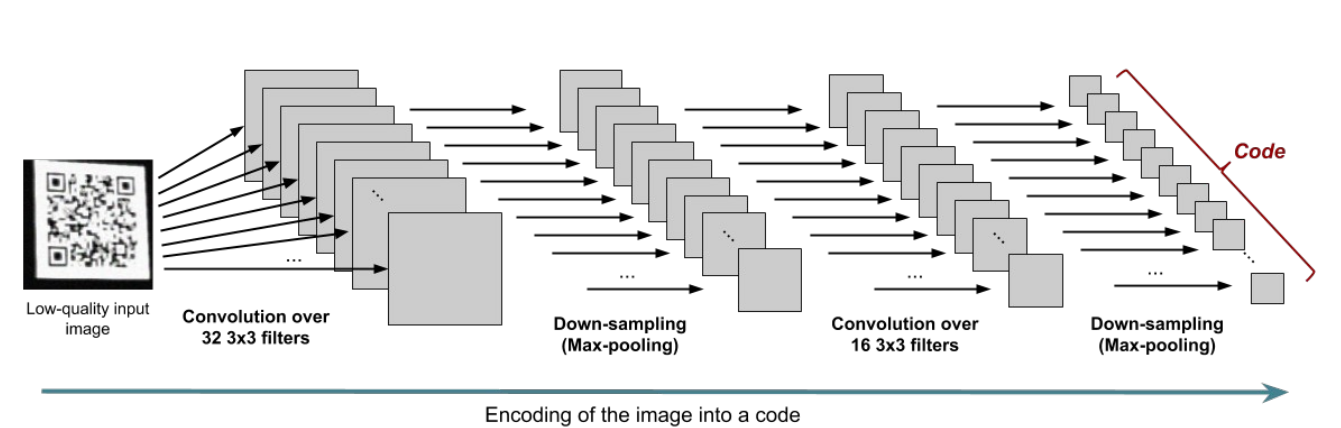
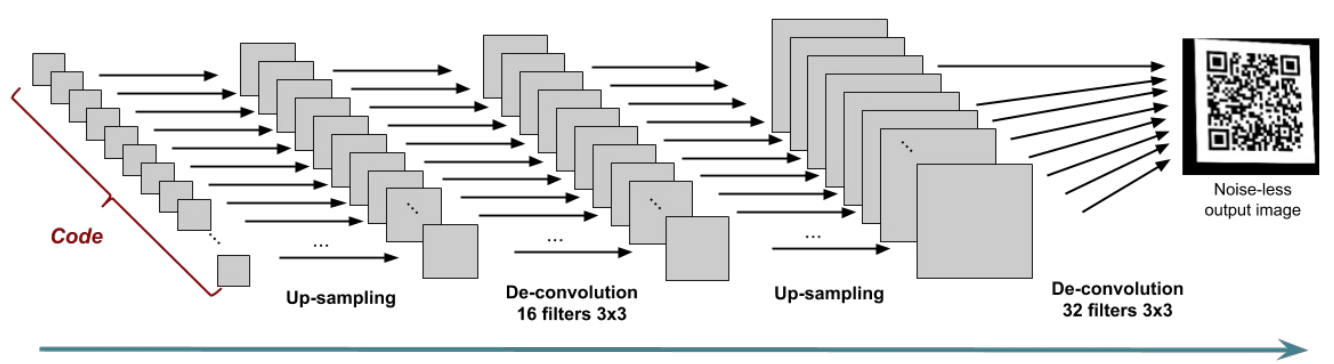


图 2.1 CNN第一部分的图片本文使用的重建系统的结构图

如图2.2，卷积神经网络结构通过并行计算大大提高了系统的运行效率。当前的图像处理在便携式设备中越来越广泛地使用，并且很好地支持多线程应用程序。系统的多线程结构实现了系统的模块化，便于调试和扩展。不同的模块具有自己的输入和输出，这对于与其他三维重建系统的新组合非常方便。此时，图像的尺寸比原始图像小得多。由于它们是通过原始图像中的操作获得的，因此它们代表了自动编码器结构中的代码。此实现帮助我们重新计算了应用的时间填充，从而获得了较大尺寸的输出图像。上采样操作后，尺寸增加了一倍。最终执行了最后的卷积和上采样，也获得了单个输出图像。

图 2.2 CNN第二部分的图片本文使用的重建系统的结构图

## 2.2 堆叠图像数据进行训练

为了使神经网络框架适合模型训练，我们可以将所有28 x 28 = 784值堆叠在一个列中。第一条记录的堆积列如下图所示：（使用x\_train.reshape（1,784））：

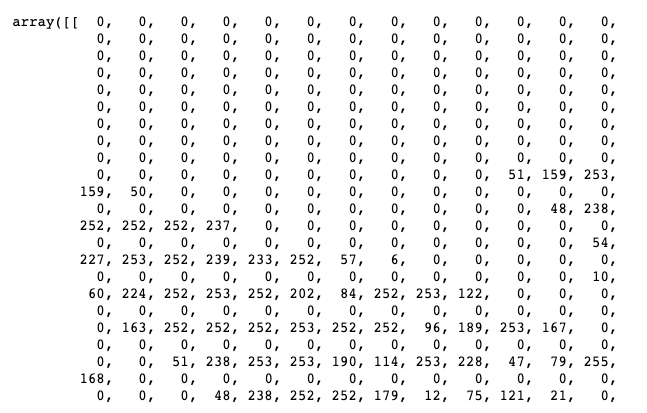
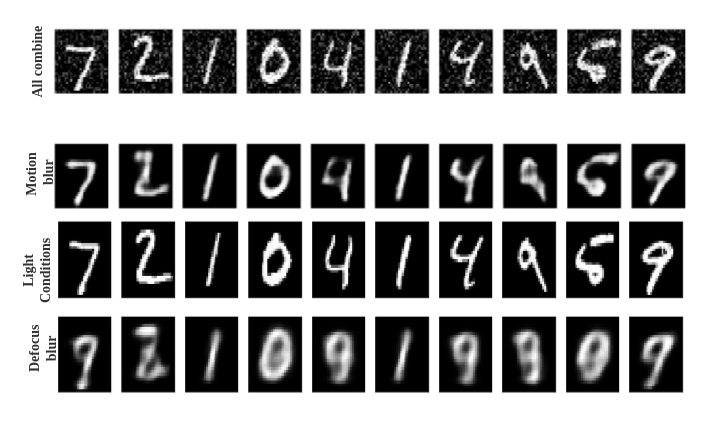


图 2.3 价值的一部分

如图2.2所示，然后，我们可以使用标准神经网络训练模型。784个值中的每个都是输入层中的一个节点。但需要考虑当我们堆叠数据时，是否会存在丢失很多信息的情况。图像中的时空关系已被舍弃，这意味着大量的信息丢失。让我们看看卷积自动编码器如何保留空间和时间信息。同样，在本文中，我们解释了为什么卷积自动编码器是处理图像数据的首选方法。最重要的是，我将演示卷积自动编码器如何减少图像中的噪声。我们使用Keras模块和MNIST以及带有嘈杂QR码图像的大型外部数据集，本文中的数据是使用一些JavaScript编写的。Keras是用Python编写的高级神经网络API，能够在TensorFlow上运行。这篇文章是我以前的文章“什么是图像识别？”的扩展。我鼓励您看看。...

图 2.4 经过不同程度的运动模糊，散焦模糊，高斯噪声后的图像

如图2.4所示，图像降噪的思路是训练一个模型，将噪声数据作为输入，并将它们各自的清晰数据作为输出。这是与上述模型的唯一区别。我们首先将噪声添加到数据中。我们在小型数据集和大型图像数据集上都进行了实验，ImageNet在一些小型数据集上检查了不同的完整训练微调配置，以了解其有效性。然后，我们在ImageNet数据集和经过训练的噪声图像数据集上检查并验证我们的方法

## 2.3 CNN模型

通过此评估，可以获得本地网络，我们考虑了三个以上的知名数据集：MNIST QR CODE数据集和ImageNet MNIST是具有60000个训练图像和10000个测试图像的手写数字数据集。每个图像是一个28 X 28灰度图像，属于从“ 0”到“ 9”的一位数字类别。对于MNIST，我们使用LeNet进行分类。在用于分类观察的LetNet 5的结构中，网络具有6个以上的6层和3个要训练的参数：前两个卷积层，平坦层和完全连接层。我们考虑了两种处理失真图像的方法：微调和带噪图像的再训练。

使用原始数据集训练的模型，我们在失真的数据集上微调了模型的模型的前N个层，同时在其余层中固定了参数。在最后一层固定参数的原因是，图像模糊和噪点被认为对图像的低级特征（例如颜色）有更大的影响，对高级信息（例如图像的语义）有影响，因此 在微调中，我们专注于卷积神经网络的起始层，其中包含更多低层信息。N表示从1降到4。在重新训练中，我们将微调方法表示为从第1到第1至第4的微调方法，我们从头开始使用失真的数据集训练整个网络，而不使用预训练训练模型，我们将训练方法称为再训练。

我们使用的噪声数据集与MNIST的先前方法类似，我们对CIFAR10快速模型使用了微调和重新训练，因此我们将前1到前5作为微调方法。再培训方法称为再培训 模糊和噪声的类型：在本实验中，我们考虑两种模糊运动模糊和散焦模糊，以及一种类型的噪声高斯噪声运动模糊是一种典型的模糊类型，通常是由相机晃动和二维码快速移动引起的。我们在1个像素的随机方向上生成运动模糊内核。均匀抗锯齿光盘的运动模糊内核的大小半径 。

当相机失去图像焦点时会发生散焦模糊，我们通过均匀的抗锯齿光盘生成散焦模糊，从[0，5]采样光盘的半径。在为一个图像生成运动或散焦模糊核之后，我们在整个图像上使用此卷积运算以生成模糊图像。高斯噪声是由于照明不佳和温度过高而导致的，从而导致相机中的CCD无法获得正确的像素值。我们选择高斯噪声为零均值，并且标准差样本为[0，5]，颜色为整数[0，255]。

(2.2)

(2.3)

其中是当前帧观测得到的TSDF值，对每个体素，除了存储一个，还维护了一个不断累计的权重值，对每次测量的置信度进行叠加。为常数。对于体素的颜色，同样通过加权平均来获取融合后的RGB颜色。

在建图线程中，当新的一帧图像判定为关键帧之后，从关键帧获取局部点云，对局部点云进行TSDF融合。在获取TSDF模型之后要从TSDF模型中提取物体表面信息。物体表面通常通过很多相互连接的三角形表示，通过Marching Cubes[15]从TSDF 模型中提取了物体表面的三角形。具体方法如下：

如图2.4(a)所示，通过空间中的相邻8个体素中心组成立方体。在TSDF模型中每个体素中心储存的值是该点到最近平面的距离。如果物体表面经过该立方体，那么在表面两侧的立方体顶点存储的TSDF值一定异号。通过计算立方体每条边两个端点的TSDF值的乘积可以判断是否异号,从而判断是否有表面经过。如果有表面经过，通过顶点的TSDF值的大小可以判断物体表面经过立方体的边的位置。在获取多个物体表面的位置之后可以连接构成三角形，从而获取整体的物体表面。如图2.4(b)所示通过物体表面的顶点来提取三角形。

## 2.4 本章小结

编码器减小输入数据的尺寸，以便压缩原始信息。解码器从压缩数据中恢复原始信息。自动编码器是一个神经网络，可以学习自动编码和解码以生成高质量的新图像。另一个限制是自动编码器的压缩是有损的。因此，虽然降噪效果使用卷积自动编码器，它仍不能完美地还原原始信息。我们向MNIST图像和二维码噪声数据集数据添加随机噪声，并将它们用作训练的输入。我们训练一个新的自动编码器，将嘈杂数据作为输入，将原始数据作为预期输出。在训练过程中，自动编码器学习从输入图像中提取重要特征，并忽略图像噪声，因为标签没有噪声，最后我们将嘈杂的测试图像传递给自动编码器以查看恢复的图像。

# 3 视觉二维码联合优化实现

## 3.1 引言

跟踪模块获取当前帧相对于先前关键帧的系统变化的评估，经过多个帧后会有较大的误差积累。通过全局优化和所有帧状态变量的优化，可以使用所有信息来提高姿势精度。全局优化可以大大减少错误累积。从图3.1可以看出，在帧之间将观察到相同的特征点和相同的平面，并且在帧之间存在系统测量数据。虚线表示视觉和二维码约束，通过它们可以构建帧之间的误差函数以实现所有帧的全局优化，从而获得更准确的帧姿势??。下面将具体介绍全局优化的实现。

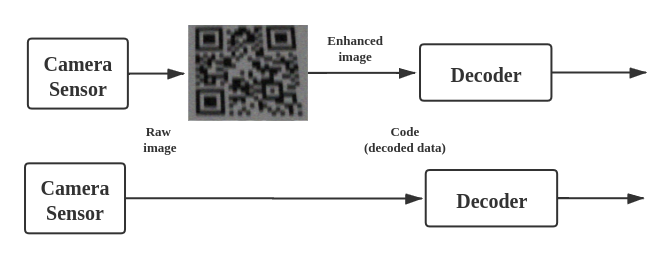


图3.1 来自摄像机的图像的系统在解码之前得到了增强

## 3.2 定义符号和优化函数

在优化中，采用了关键帧策略。仅关键帧的状态量参与全局优化，而其他帧的状态量通过最近的关键帧分配。定义的符号如下：写入第一个关键帧，包括RGBD摄像机的拍摄数据的颜色图和深度图。对于关键帧，需要优化的状态变量为：相机的旋转方向，平移位置，二维码速度，二维码角速度偏差和二维码加速度偏差。是对应二维码的旋转矩阵和平移矩阵，是从摄像机到二维码的转换矩阵。是初始重力值和重力方向的最优化。通过预校准或初始化过程获得的相机二维码转换矩阵和重力方向将导致误差累积，随着路径的增加，误差将逐渐放大。另一方面，通过全局优化得到的相机二维码变换矩阵和重力方向将越来越精确，从而实现全局一致性，避免错误积累。因此，本文在整体全局优化函数中估算了相机的状态，相机的二维码转换矩阵以及重力方向。和是全局变量。为了简化表示，使用表示全局优化中的所有优化量：

, 其中

(3.1)

相邻关键帧之间的IMU数据进行积分可以得到相邻关键帧之间的运动约束。在匹配的关键帧对中，视觉找到的匹配特征点构建了视觉约束。根据视觉找到的匹配平面可以构建平面约束，全部匹配的关键帧对写作。K是所有匹配的特征点对的个数，M是所有匹配平面的个数。构建如式（3.2）的误差函数： \\* MERGEFORMAT

(3.2)

通过最小化误差函数求出的估计值。得到误差函数（3.2）后，通过高斯牛顿优化方法来计算状态量的更新量，实现误差函数的最小化，即通过（3.3）来计算状态量的更新量：

(3.3)

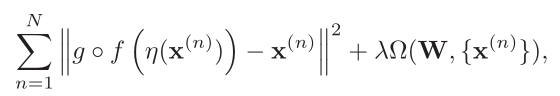
其中是误差函数（3.2）对所有状态量求导得到的雅各比矩阵，是通过误差函数(3.2)得到的误差项，是求出的状态量的更新量。和的计算过程是优化问题中计算复杂度的主要来源，下面将介绍的计算分解过程和加速方法，可以通过同样的方法进行计算加速。

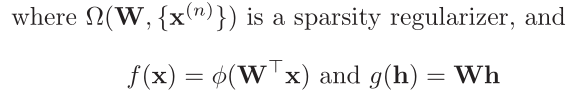
图3.2 的计算分解

如图3.2所示，设共有N个关键帧，每个关键帧需要优化的状态量为15维（分别是旋转向量，平移向量，IMU速度，IMU加速度零偏，IMU角速度零偏），第一个关键帧不参与优化，则参与优化的状态量为15\*（N-1）个，同时在全局优化中耦合9维的全局变量（相机-IMU变换矩阵的旋转向量和平移向量，重力方向），所以需要优化的变量一共是[15\*（N-1）+9]维的。设有K个特征点，每个特征点会提供三维的误差，所以视觉特征点提供的误差项一共是3\*K维的。每两个关键帧之间有9维的IMU误差项（旋转误差、速度误差、平移误差），所以IMU约束一共是9\*（N-1）维的。设观测到的公共平面一共有M个，每个公共平面会提供6维的误差项（法向量，到原点的距离），所以平面提供的误差项一共是6\*M维的。整体的雅各比矩阵是由视觉和IMU误差项分别求出的雅各比矩阵组合而成，组合方式如图3.2所示。和的计算复杂度分别是和，其中N是关键帧的个数，是帧对的个数，k是平均每个帧对的匹配特征点对的个数。关键帧的个数远小于特征点对的个数，所以优化问题的计算复杂度主要存在于，即 的计算耗费主要在图1的 上。加速计算的关键在于加快的计算。其中是特征点对的雅各比矩阵，本文将改进FastGO[6]算法，通过特征点的二阶统计量可以直接得到，在不损失计算精度的前提下把的计算复杂度从降到了，从而显著降低 的计算复杂度，具体实现过程将在3.3中详述。

## 3.3 稀疏去噪自动编码器

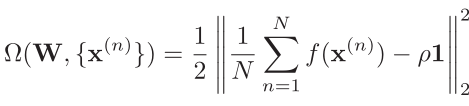
我们从单层更新开始，这是具有单个隐藏层的多层感知器网络的一种特殊形式。该更新尝试通过最小化以下成本函数来学习可最佳重构输入向量的网络：

(3.4)



分别是具有分量非线性函数的编码器和解码器。η明确地将噪声添加到输入样本x（n）。W是输入层和隐藏层之间的权重矩阵，由编码器和解码器共享。为了简化符号，我们忽略了所有单位的偏差。f可见单元和隐藏单元。解码器将Q和P分别作为其域和范围2恰好相反。在这种设置下，稀疏性节器Ω定义或限制学习的编码器f的范围Q。我们应该注意到，对于任何样本x∈R p，如果f（x）∈/ Q，则该样本不属于P。还可以预期，被误差破坏的样本小于η注入的样本仍然是P。编码为Q中的隐藏代码，但任何高噪声或损坏的样本x都会通过编码器f映射到Q之外的区域，尽管不一定。

让我们考虑将平均隐藏激活规则化为预定义的稀疏超参数ρ。在这种情况下，

 (3.7)

假设该模型由随机二元隐藏单元组成，其概率由编码器f给出，我们可以看到f（x）平均具有ρ×q分量，而其他所有分量均不活跃。如果ρ设置为接近0，则稀疏的潜在表示将

由编码器f产生。这导致一个以（近似）范围为基础的编码器f，以数据集P为条件，

由于存在许多匹配关键帧之间的帧对，是总的视觉雅各比矩阵的子矩阵，所以有下面的关系式：

(3.9)

只有两个非零的3\*6的子矩阵，所以 ()只有四个非零的处于在相同位置的6\*6的子矩阵，作为()的组合，有四个6\*6的非零子矩阵，分别写作，，,。下面展示了的计算过程，其他的子矩阵可以通过同样的方法计算得到：

(3.10)

通过式(3.10)可以实现直接由视觉的二阶统计量得到的结果，是的求和。从而可以实现通过视觉的二阶统计量得到的结果，大大降低了视觉部分雅各比矩阵的计算复杂度。

## 3.4 二维码预积分实现

二维码约束的构建借鉴了二维码预积分技术[11]。相机和IMU的变换矩阵是、，可以获得下面的从相机坐标系到IMU坐标系的转换：

(3.11)

二维码获取的角速度和加速度写作 、，真实的角速度和加速度写作、，角速度和加速度的噪声写作、，有下面的关系式：

(3.12)

假设在时刻角速度和加速度保持不变，可以通过时刻的状态获取时刻的状态:

(3.13)

把 (3.12) 带入 (3.13) ，可以得到下面的关系式：

(3.14)

通过多段二维码数据的积分可以得到i到j时刻的位姿关系式：

(3.15)

为了避免重复积分，计算下面的位姿变化量：

(3.16)

通过下面的方式分离中的噪声：

(3.17)

通过同样的方法可以得到 ，；

在通过视觉的方法得到相机位姿之后可以经过转换得到二维码的位姿，从而在相邻的关键帧之间写出下面的位姿变化量：

(3.18)

同时，通过二维码数据的积分获得了关键帧之间的相对位姿变化量和，结合式（3.18）可以得到下面的误差函数：

(3.19)

通过零偏的线性化可以将式（3.19）写成下面的形式：

(3.20)

-

利用式(3.11)把相机位姿带入，并通过来代替，可以得到下面的误差函数：

(3.21)

-

然后可以通过李代数的扰动模型求出误差()相对于需要优化的状态变量()的雅各比矩阵。全局变量(二维码变换矩阵、和重力方向)紧耦合在全局优化函数中，从而实现了在全局优化过程中进行自标定，标定结果不会产生误差积累。

## 3.5 平面约束推导

本文采用了AHC平面检测算法[16]从点云高效地得到了平面，通过最小二乘法可以得到平面参数，得到的平面方程如式（3.21）：

(3.22)

其中是单位法向量，是原点到平面的距离，是在当前的相机坐标系下观测的点的坐标。对于每个平面检测结果而言，通过最小二乘法可以估计上面的平面参数。如果当前相机的位姿为，，在世界坐标系下的平面方程可以通过下面的变换得到：

(3.23)

这里表示平面上的点在世界坐标系下的坐标。

在第i帧相机坐标系下的平面坐标是,,第i帧相机位姿为，。则在世界坐标系下的平面参数为（单位法向量）和（原点到平面的距离）。通过下面的公式得到:

(3.24)

变量的作用在于保证到原点距离的参数是正数。是被第i帧和第j帧同时观测到的第k个平面，通过平面参数的比较可以获得所有的匹配平面。在一个帧对中找到相关平面之后，可以写出下面的残差函数：

(3.25)

把所有相关平面的残差组合起来，就可以得到下面的平面约束：

(3.26)

根据式（3.24）可以推导平面残差对于相机位姿的雅各比矩阵。得到平面误差项的雅各比矩阵，即得到了图1中的。根据计算得到的IMU误差项的雅各比矩阵，即图1中的、、，加上视觉部分已经得到的，可以通过组合的方式计算得到图1中的，通过同样的方法可以得到。得到和后通过稀疏矩阵乘法可以得到状态更新量，从而实现了视觉IMU信息的全局优化。

## 3.6 错误回环检测的判断方法

回环检测可以实现建图的全局一致性，使得二维码的状态量估计的更准确，而错误的回环检测会破坏这种全局一致性，导致建图结果出现混乱，二维码的零偏估计也会发生跳变。在发生正确回环检测的时候，零偏的估计量不会发生大的变化，而发生错误的回环检测时破坏了建图的一致性，为了拟合错误回环会产生轨迹的不合理形变，二维码的零偏估计量的更新量会发生突然的增大。所以根据二维码零偏估计量的变化可以判断当前回环检测的正确性。如果二维码零偏估计量发生了突变，就取消当前的回环检测，这样可以实现准确的错误回环的去除，提高了系统的鲁棒性。

## 3.7 二维码在标定方法

在程序运行一开始的时候不知道二维码的相对位置，这时候的观测信息少，如果直接进行视觉-二维码联合优化，会有比较大的概率发生优化不收敛的情况，这种情况下将无法得到相机的位姿估计。为了避免一开始的优化不收敛，在前15个关键帧只通过视觉信息的优化来得到相机位姿，然后进行保持所有关键帧位姿不变，只优化二维码的状态量以及相机-二维码变换矩阵。在已经得到前15个关键帧较为准确的位姿之后，通过这些位姿信息可以较为准确地得到二维码状态量和相机-二维码变换矩阵的估计值，得到了较为准确的初值之后再进行紧耦合的联合优化就可以避免优化不收敛的发生。

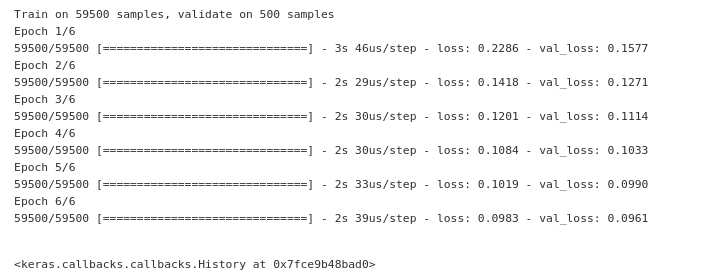
## 3.8 本章小结

通过特征点的二阶统计量和二维码预积分直接构建了优化函数，避免了大量的重复计算，实现了视觉二维码联合优化的复杂度从与关键帧的个数的平方呈正比下降为与关键帧的个数呈正比。将视觉特征点约束、二维码约束、平面约束结合在一起，在高效率的前提下提高了优化位姿的精度。同时通过在全局优化中耦合相机-二维码变换矩阵实现了相机-二维码的在线标定。通过二维码零偏状态的估计量实现了错误回环检测的判断。

# 4 实验测试和结果分析

## 4.1 数据集测试与对比

在本文中，我在带有噪声图像数据集和MNIST数据集的二维码上测试了我的方法。我们训练自动编码器来压缩输入图像，然后恢复到原始大小。这样，我们的训练数据和标签数据都是相同的图像数据。一旦对自动编码器进行了训练，我们就可以使用编码部分对MNIST数字图像进行编码。通过训练自动编码器，我们实际上是在同时训练编码器和解码器，这就是完成上述训练的原因。表中的值是算法获得的路径与实际路径之间的均方根误差。误差越小，定位精度越高。其他比较定位方法包括：可以看出，该方法可以在基于CPU的方法中实现最高的精度，并且与基于GPU的算法相比可以达到相似的精度。

表4.1 本文方法以及其他几种方法在数据集上的测试结果

## 4.2 采集真实数据测试



图4.1 实验设备

除了在数据集中进行测试外，本文还通过jupyter笔记本环境，Keras框架和消费级IMU（SC-AHRS-100D2）收集了数据，并在实际场景中进行了测试。测试设备如上图所示

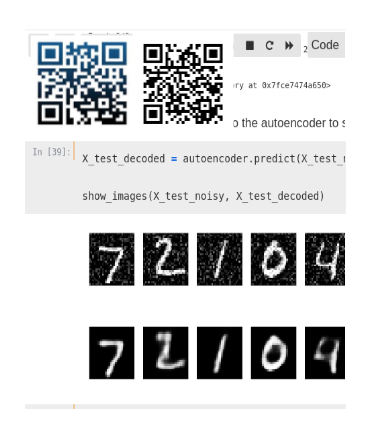


图4.3解码结果和解码图像校正

如图4.3，首先，输入是图像，使用卷积神经网络（卷积网络）作为编码器和解码器是有意义的。在实际设置中，应用于图像的自动编码器始终是卷积自动编码器-它们的性能好得多。我们也对其进行训练，我们将使用原始MNIST和我们自己的数据集数字，其形状为（samples，3、28、28），并且我们将归一化介于0和1之间的像素值。我们可以在图像中看到结果 以上我们训练了这个模型8个纪元。为了演示如何在训练过程中可视化模型的结果，我们将使用TensorFlow后端和TensorBoard回调，然后让我们训练模型。在回调列表中，我们传递TensorBoard回调的实例。在每个时代??之后

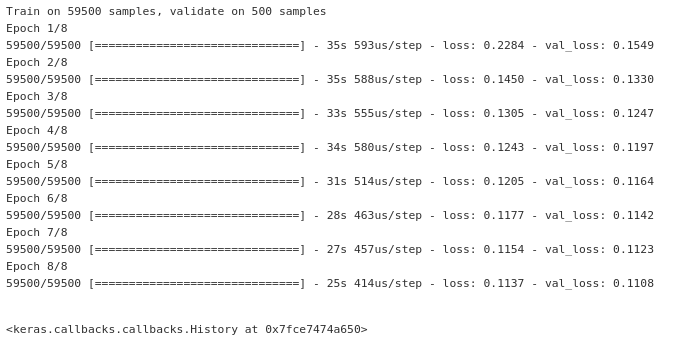


图4.4 自动编码器，其中嘈杂的数据作为输入，原始数据作为预期的输出。

如图4.4如图所示，模型收敛到了0.1137的损失，明显好于我们以前的模型，这在很大程度上是由于编码表示的熵容量更高，即128维与32维相比。我们还可以看一下128维编码 表示形式。这些表示是8x4x4，因此我们将它们重塑为4x32，以便能够将它们显示为灰度图像。

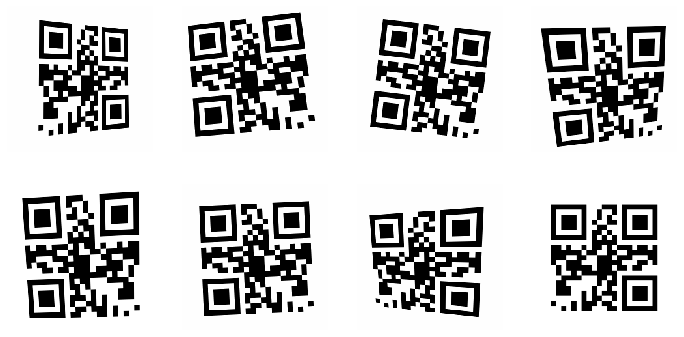


图4.5（a）、（生成干净的二维码）

如图4.5所示，如图所示，在我们的卷积神经网络的最后结果中，与以前的卷积自动编码器相比，为了提高重建的清晰图像的质量，我们可以看到我们的网络的预测和损失是一个略有不同的模型，其中包含更多的滤波器 每层方法可以获得准确的结果卷积参数是通过最小二乘法计算的，我们可以将这些采样的潜在点映射回重构的输入。可以看出，在两个帧之间实现了精确的平面匹配。

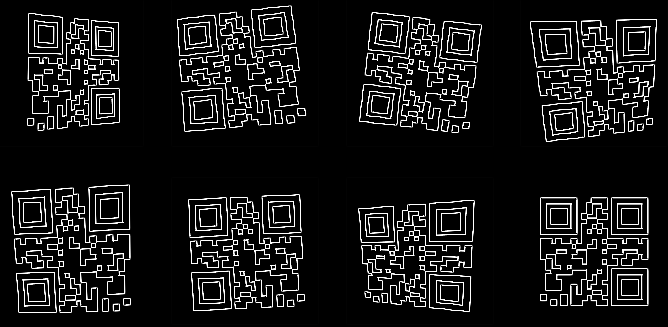


图4.6 二维码检测和过滤器匹配结果

如图4.6，通过网络，我们使用了多帧图像的过滤器，可以获取地面和场景中的二维码。由于二维码和地面的完整性可以构造新的滤波器约束，因此可以防止在多个帧中观察到的同一二维码出现裂纹。

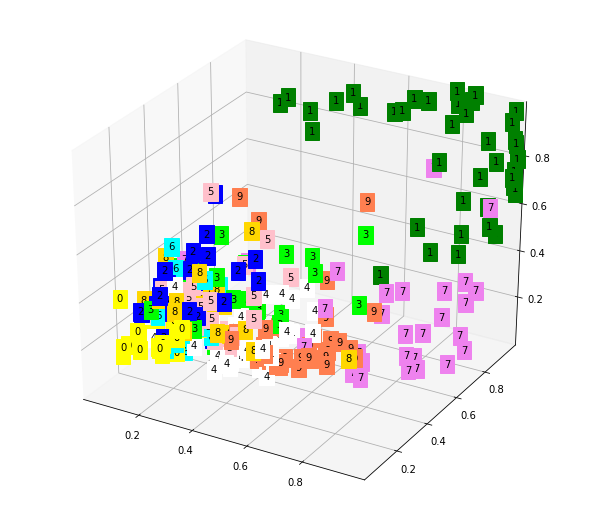


图4.7 数据集3D可视化

如图4.7所示，在通过多种不同的颜色后可以确定3D可视化，以帮助我们可视化数据。数据的选择相对稀疏。回送检测和全局优化在数据上运行。普通帧的姿势是通过最近的数据执行的。匹配计算。数据选择策略可以大大提高运营效率。当相对于先前数据的姿势变化超过阈值时，确定新数据。通过这种动态选择数据的方法，当场景变化很大且移动迅速时，可以选择更多的数据可视化；当场景变化较小时，可以选择较少的数据，并且能够根据建筑地图的需求灵活地选择数据 。



图4.8 网络获得的二维码掩码对的结果

如图4.8所示，通过转换，可以找到观察到相同场景的图像对？？。联合优化由形成的视觉约束，然后将这些转换的结果图像用作模型的输入，同时原始失真的图像和网络生成并提供输出的地面??真实性，该图像从模型中获得了完整的干净图像给定的输入。



图4.9 产生的二维码，（a）干净的图像，（b）图像模糊

图4.9显示大场景重建的效果。在家具店中进行了重建测试，（a）是环境的三维模型，（b）添加了重建的颜色模型。从图中可以看出，经过精确定位后，可以通过TSDF融合获得高质量的图像模型，可以更准确地重建场景细节，包括沙发表面和墙壁之间的间隙，同时还可以确保 高质量的表面颜色； TSDF融合极大地减少了模型的存储空间，因此大型场景的环境重建可以在便携式设备上运行而无需大内存。通过环回检测，在长距离环回之后，可以成功检测到环回，然后通过视觉和模型的联合优化，大大减少了累积误差，因此整个二维码模型不会重叠。



(a) (b)

图4.10 火车模型，（a）输入项，（b）图像输出

图4.10当视觉跟踪失败时，显示输入定位与模糊图像状态优化之间的差异。（A）是输入模型状态优化，可以看到重建结果相对平滑，表明二维码状态的估计值准确，定位精度高。（B）这是没有模糊和清晰图像状态优化的结果，可以看到明显的裂缝。由生成的二维码图像提供的原始数据的噪点非常少。加速度和角速度的偏移为零。可以在第二次加速度积分后获得位移。存在相对较大的错误。图中二维码的定位结果更加准确，表明二维码优化提供了准确的结果。状态量的估计值对定位有很大帮助。

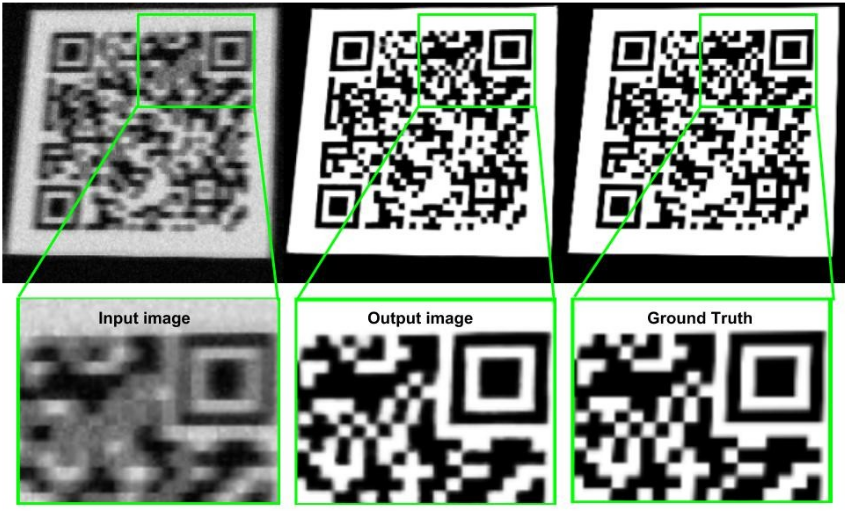
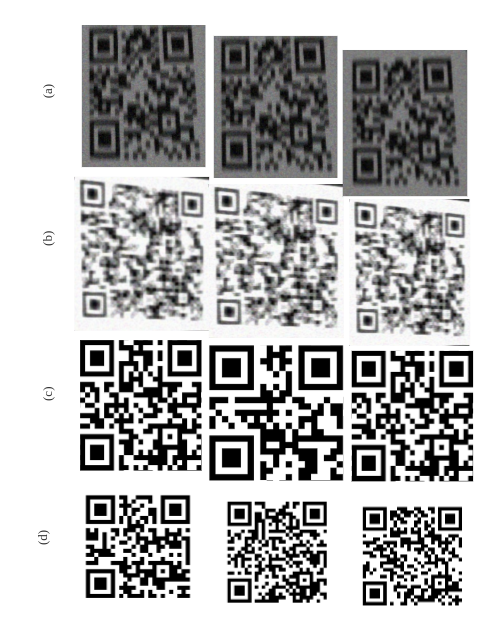


图4.11 输入，输出和基本事实

如图4.11所示，当存在相似的场景而不是相同的场景时，它可以识别错误的循环，并且当返回相同的位置时，它可以识别正确的循环，这可以确保构造的整体一致性并大大减少累积误差，从而使我们可以看到 某些图像似乎对于要重建的图像而言仍然很嘈杂，通常是我们训练的模型的输出图像，我们已经在上图显示了该图像。输出模型看起来比输入模型清晰得多，这意味着该模型的良好性能更加引人注目，如我们在图像细节中所看到的。



本文方法 Image processing

图4.14 比较不同情况下的重建算法解码方法和解码器方法。（A），（b），（c），（d）测试了不同的操作方案

收集数据后，将本文中的方法与目前最好的网络重构方法解码和解码器进行比较，比较结果如图所示。如图4.14所示，（a）（b）使用两个相似的对象作为重复场景。本文中的方法丢弃了错误循环并实现了全局一致性。自动编码器被错误循环所误导，并产生令人困惑的结果。ElasticFusion丢弃实际循环，从而导致累积错误。（c）显示快速运动的场景。该方法跟踪可靠，保证了重建的完整性和一致性。网络跟踪失败，导致重建结果不完整。ElasticFusion的视觉跟踪结果错误，导致重建结果重叠。（d）显示了在快速运动下放大的局部模型和颜色细节。解码具有较粗糙的表面和模糊的颜色外观。解码器表面和清洁器更加稀疏，丢失了许多细节。相反，本文的方法通过融合视觉和惯性观测获得更高的定位精度，从而获得更高的重建质量。4.3 运行效率与在线标定测试

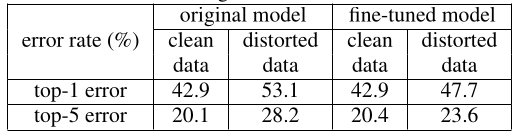


图4.15 预训练模型之间的精度比较

图4.15我们还研究了在大型数据集和非常深的网络上进行微调的效率。对于ILSVRC2012训练和验证集的实验，我们通过组合所有3种类型的模糊/噪声生成了失真的数据。对于每个图像和每种类型的失真，从[0，4]中均匀采样失真级别。获得失真的数据后，我们对预训练的Tensorboard的前3层进行微调。如表所示，原始的预训练Alexnet模型与微调模型之间的精度比较。与原始的预训练模型相比，经过微调的模型提高了失真数据的性能，同时保持了干净数据的性能。当我们想在有限且失真的数据集上使用像Alexnet这样的大型DNN模型时，在前几层进行微调可以提高失真数据的模型准确性，同时保持干净数据的准确性。

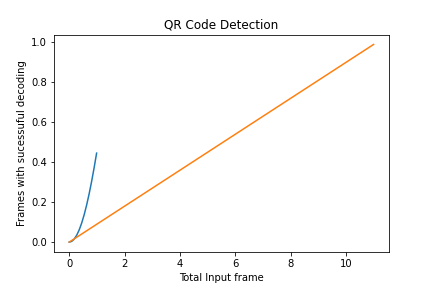


图4.16 松耦合联合优化中最后训练的收敛过程

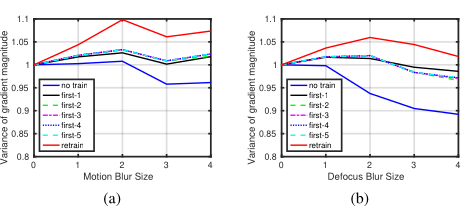


图4.17 转换层的特征图梯度幅度的方差 (a- 运动模糊) (散焦模糊)

在获得特征图f m的梯度幅度g f m之后，我们计算梯度幅度的方差：v f m = var（g f m）。当我们在图像上应用散焦模糊或运动模糊时，清晰的边缘被涂抹成光滑的边缘，因此渐变幅值图变得平滑，并且方差较低。特征具有较高梯度方差值v f m的地图被认为具有更多的边缘和纹理信息，因此对于图像复制更有帮助介绍。而较低的v f m值表示信息。

特征图中的内部空间有限，因此不足以表示图像观点。视觉初始化后，通过优化获得网络变换矩阵。图4.15是通过网络耦合方法来优化损耗的相对位置，即固定输入和层的姿势，并且仅优化变换矩阵。您会看到优化结果不稳定并且准确性很差。卷积自动编码器中的卷积层。在编码部分有三层分别标记为Conv1，Conv2和Conv3。我们将打印出前十个噪点图像以及相应的降噪图像。

## 4.4 本章小结

通过数据集测试和采集真实数据进行实验，验证了以下几点：

（1）本文的方法能够实现高精度的定位结果并获得高质量的重建模型；

（2）全局优化的改进算法实现了优化复杂度和帧数量呈线性关系，大大加快了全局优化的速度，使得算法在只用CPU的情况下能够实时运行在可便携设备上；

（3）成功率获得准确的状态估计，当视觉跟踪失败时可以通过网络数据实现准确定位； 它可以成功应对场景纹理特征和快速移动；

（4）可以成功识别正确的循环并丢弃错误的循环，从而确保全局一致性；

（5）系统实现了大量准确解码，是一般解码系统的两倍

# 5 结论和展望

## 5.1 结论

本文使用卷积神经网络和二维码数据集来实现对成功率重建系统的高质量解码，这是各种应用程序和移动技术的基本组成部分。具体实现了以下部分：

1、首先，从基于完全连接层的简单自动编码器开始，使用二维码噪声图像准备数据集。具有三个隐藏层的层处理编码，而输出层处理解码。其中每个输入图像被展平为784（= 28×28）个数据点的数组。然后，通过完全连接的层将其压缩为32个数据点。然后，将编码数据解码为原始的784个数据点。对于每个像素，S形将返回0到1之间的值。这整个过程成为可训练的自动编码器模型。

2、在卷积神经网络中，架构转换Max-Pooling是耦合的。通过首先执行视觉初始化，然后再执行视觉初始化，max-pooling层非常有用，并且无需学习。他们只是采取一些k×k区域并输出一个值，这是该区域中的最大值。

3、获得高质量的定位精度后，通过网络获得高质量的二维码重构模型。

4、我们用噪声图像对MNIST图像数据和/或自己的数据集进行预处理，以便图像数据在0和1之间归一化。我们还将火车数据分为火车组和验证组。我们训练一个新的自动编码器，将嘈杂数据作为输入，将原始数据作为预期输出。在训练期间，自动编码器学习从输入图像中提取重要特征，并忽略图像噪声，因为标签没有噪声。将嘈杂的测试图像传递到自动编码器以查看还原的图像。内置了一个编码器，并使用它来压缩噪声数字图像。我们确认了784个像素的数据点现在已压缩为32个数据点。

## 5.2 展望

尽管系统已经成功地测试了更少的噪音，但仍需进行部分不公平的光照条件以将模型推广到其他种类的低质量图像，例如玻璃反射等。而且，有可能整合并构建通用模型。该模型适用于一维和二维条形码，因为在许多实际应用中，它们都出现在图像上。验证自动编码器结构是否可以用作解码机制而不是图像增强方法也将很有趣。通过添加额外的卷积并可能增加过滤器数量和训练数据，有可能训练一个自动编码器，该编码器能够解码条形码内的信息。

参考文献

[1] Agarwal S, Furukawa Y, Snavely N, Reconstructing rome[J]. Computer, 2010,43(6):40-47.

[2] Hedman P, Alsisan S, Szeliski R, Casual 3D photography[J]. ACM Transactions on Graphics (TOG), 2017,36(6):234.

[3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, “Ima-genet classification with deep convolutional neural networks,”in Advances in neural information processing systems, 2012,pp. 1097–1105.

[4] Newcombe R A, Izadi S, Hilliges O, Kinectfusion: Real-time dense surface mapping and tracking[C], [IEEE International Symposium on Mixed and Augmented Reality](https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6082471), 2011.

[5] Dai A, Nießner M, Zollhöfer M, Bundlefusion: Real-time globally consistent 3d reconstruction using on-the-fly surface reintegration[J]. ACM Transactions on Graphics (ToG), 2017,36(4):76a.

[6] V Pomponiu, H Nejati, and Ngai-Man Cheung, “Deepmole: Deep neural networks for skin mole lesion classification,” in Proc. IEEE International Conference on Image Processing (ICIP), 2016.

[7] Thomas Whelan, Renato F Salas-Moreno, Ben Glocker, Andrew J Davison, and Stefan Leutenegger, Elasticfusion:Real-time dense slam and light source estimation, The International Journal of Robotics Research[J], 2016:1697–1716.

[8] Matthew Klingensmith, Ivan Dryanovski, Siddhartha Srinivasa, and Jizhong Xiao, Chisel:Real time large scale 3d reconstruction onboard a mobile device using spatially hashed signed distance fields[C], Robotics: Science and Systems, 2015: [XI.040](https://doi.org/10.15607/RSS.2015.XI.040)

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, Deep residual learning for image recognition,” arXiv preprint arXiv:1512.03385, 2015.

[10] Lei Han and Lu Fang, Flashfusion: Real-time globally consistent dense 3d reconstruction using cpu computing[C], Robotics: Science and Systems,2018: [XIV.006](https://doi.org/10.15607/RSS.2018.XIV.006)

[11] Christian Forster, Luca Carlone, Frank Dellaert, and Davide Scaramuzza, Imu preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation[C], Robotics: Science and Systems, 2015: [XI.006](https://doi.org/10.15607/RSS.2015.XI.006)

[12] T.Laidlow, M.Bloesch, W.Li, and S.Leutenegger, Dense rgb-d-inertial slam with map deformations[C], IEEE/RSJ International Conference on Intelligent Robots and Systems, 2017:6741–6748.

[13] Samuel Dodge and Lina Karam, “Understanding how im- age quality affects deep neural networks,” arXiv preprint arXiv:1604.04004, 2016.

[14] Martin Brossard, Silvere Bonnabel, and Axel Barrau, Invariant Kalman filtering for visual inertial slam[C], 21st International Conference on Information Fusion, 2018:pp.2021–2028.

[15] Pharr M, Fernando R. Gpu gems 2: programming techniques for high-performance graphics and generalpurpose computation[M]. Addison-Wesley Professional, 2005

[16] C.Feng, Y.Taguchi, and V.R.Kamat, Fast plane extraction in organized point clouds using agglomerative hierarchical clustering[C], IEEE International Conference on Robotics and Automation, 2014:6218–6225.

[17] Geoffrey E Hinton and Ruslan R Salakhutdinov, “Reducing the dimensionality of data with neural networks,” Science, vol. 313, no. 5786, pp. 504–507, 2006.

[18] M.Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger, Real-time 3d reconstruction at scale using voxel hashing[C], ACM Transactions on Graphics,2013:Volume 32,No.169

[19] Deep Learning documetation: <http://deeplearning.net/tutorial/contents.html>

[20] Lei Han and Lu Fang, Mild: Multi-index hashing for appearance based loop closure detection[C], in Multimedia and Expo, IEEE International Conference on.IEEE, 2017:139–144.

[21] Michal Hradiš, Jan Kotera, Pavel Zemcı́k, and Filip Šroubek, “Convolutional neural networks for direct text deblurring,” in Proceedings of BMVC, 2015, pp. 2015–10.

[22] Lingni Ma, Christian Kerl, Jörg Stückler, and Daniel Cremers, Cpa-slam: Consistent plane-model alignment for direct rgb-d slam[C], Robotics and Automation (ICRA),IEEE International Conference on.IEEE,2016:1285–1291.

[23] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, A benchmark for the evaluation of rgb-d slam systems[C], IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012: pp.573–580.

[24] A. Handa, T. Whelan, J. McDonald, and A. J. Davison, A benchmark for rgb-d visual odometry, 3d reconstruction and slam[C], IEEE International Conference on Robotics and Automation, 2014: pp.1524–1531.

谢 辞

这个学期的完成对我来说是一个巨大的挑战，许多问题使我震惊。我很高兴它可以最终成功完成。非常感谢毕节完成期间陈康力先生的耐心指导。我感谢我周围的学生的建议。没有他们的帮助，我几乎无法解决Bishe中的所有问题。我感谢父母一直为我提供所需的一切支持，无论是物质上的还是精神上的。多亏您周围的朋友，与您的交流是放松身心的最佳方式。

我刚进入同济校区的那一天仍然记忆犹新，转眼已经过去了五年。在过去的四年中，我从一个愚昧的新生开始慢慢成长，在课程和家庭作业中慢慢锻炼，并在考试的压力下激发了我的潜力。在校园的学习氛围和哲学指导下，我逐渐培养了自己的能力。在过去的四年中，生动地记录了老师的精心指导和学生的热情帮助，每个问题都可以耐心地回答，这使我学到了更多的知识和学习新知识的方法。在过去的四年中，我逐渐熟悉了同济大学校园的每个角落，并且合作的理念逐渐融入了我的价值观。衷心感谢刘建老师。首先，我要感谢Alexandra Erfort，卢畅老师，朱东哲和杨文翔

作为学生，我有很多缺点。在过去的五年中，我经历了宽容和耐心。将来我会继续努力。同济带给我知识和幸福。这是美丽的青春和美好的回忆。作为同济人，我感到非常自豪，并且我永远不会忘记自己的身份。

最后，谢谢母校！谢谢您的母校！