# 绪论

* 1. 水产养殖饵料投喂控制研究背景
     1. 水产养殖发展历史与现状

我国鱼类的养殖具有悠久的历史，早在公元前五世纪的春秋战国时代我国便出现了第一部的淡水鱼养殖著作《养鱼经》，经过几千年的实践积累了丰富的技术经验。在2000年我国的水产养殖总量为2578吨，占渔业总产量的60.2%，是世界上唯一水产养殖量高于捕捞产量的国家 [1]。如图1.2所示，不论是从世界其他各国还是我国来看，水产养殖的比例都是在逐年升高,而且与世界其他各国总产量不同的是我国已经实现了以水产养殖为主的产业发展，而且在2016年我国水产养殖产量比世界其他各国的水产养殖总量的2倍还多，因此不论是从世界发展趋势还是从我国水产养殖现状来看，对水产养殖进行深入的研究是十分有必要的。

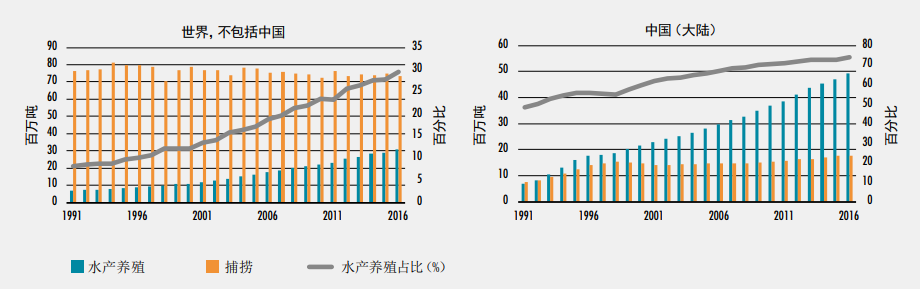


图1.1. 水产养殖与捕捞产量对比图 [2]

在全球的水产养殖部门，需要饲喂的水生生物养殖产量增速高于无饲喂的品种,整个产业逐渐向饲喂品种倾斜世界各国都越来越重视饲喂品种的水生生物养殖（图1.2）。而在饲养生产过程中饲料是最主要的消耗品，在水产养殖的经济投入中占有很大的比例，其投放量对水产养殖的产量有着重要影响，而且其投放量会随着饲养生物的生长阶段发生变化。在传统的水产养殖中，饲料的投放量一直困扰着饲养人员。若饲料投放不足，会使得养殖生物无法最大限度的生长，严重时甚至会引起肉食性养殖生物相互蚕食，降低最终的产量；若投放过量，则不仅会浪费饲料、增大水产养殖的投入产出比，还会污染水质、造成水体富营养化，间接增加养殖生物的患病几率、增大投入成本、降低产量。对饲料进行精准投放，不仅可以保证水产品的产量，还可以减少饲料的浪费、遏制水体富营养化，使得经济效益最大化的同时又保护了环境。所以对鱼类的投饵研究具有重大的意义。

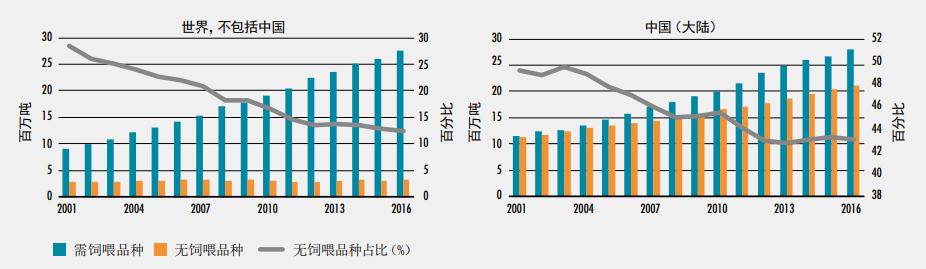


图1.2. 需饲喂和无饲喂鱼类食品养殖产量 [2]

经过不断的发展，我国的水产养殖业已经成为世界水产养殖业的重要组成部分，尤其是集约化养殖因为其单位水体产量高、养殖周期短、饲料转化率高、养殖对象广、便于管理等特点备受青睐[3]。但是集约化水产养殖方案仍然存在一定的局限性，主要表现在以下几个方面：(1)投入成本大，与粗放型水产养殖相比集约化水产养殖需要人工构建鱼类生态环境同时对于生产管理有更高的技术要求使得人力成本也进一步的提高；(2)易污染水源，为使得养殖生物最大限度的生长通常在集约化养殖过程中会大量投喂高蛋白饲料，这造成的一个不良后果就是残余饵料和水生生物的排泄物致使水体富营养化严重污染水体；(3)易滋生鱼病，集约化养殖的高密度环境使得鱼类长期处于紧张状态，特别是肉食性鱼类容易造成同类相食；长期的食用单一的配合饵料也使得鱼类体质脆弱，抗病能力较低容易患病。机器视觉在近几年的蓬勃发展，为集约化水产养殖自动化提供了一种新的解决思路。机器视觉技术因为其高效、经济、非侵害、无损等特点[4][5]已经被研究者在各个领域广泛运用。将机器视觉技术引入到集约化水产养殖中，对于上述集约化养殖所存在的局限性均有所帮助，由于机器视觉技术成本不高自动化程度高所以可以降低生产过程中的人力成本，通过机器视觉技术指导鱼类的饵料投放以及水质检测可以降低对水体的污染，基于机器视觉技术实施监控鱼类健康状态可以及时发现鱼类病害防止传播降低损失。实际上已经有大量的研究者开发了针对水产养殖的机器视觉应用（将在下一节详细介绍），例如基于机器视觉的鱼种类识别[6]、鱼类行为评估[7]、鱼类轨迹跟踪[8]等。

机器视觉技术，除应用于水产养殖之外，在海底勘查、海洋环境监测以及水下目标识别等领域也有重要的研究意义和实用价值，比如水下摄像机可以在海底和海面完成对异常情况的监控[9]，使相关人员及早对一些极端情况进行预防，比如海啸，台风，地震等灾害，也可以监测一些藻类的生长情况，及早对赤潮等灾害采取措施。

* 1. 国内外相关研究现状
     1. 机器视觉在渔业中的研究现状

机器视觉被运用到水产养殖业已有较长历史，在1983年就有研究者通过发光二极管作为光源，红外摄像机作为传感器观察鱼苗的生长过程[10]，而首次通过机器视觉建立一套观测系统的是Huse等[11]人在1990年建立的一套远程监控系统。该系统通过远程控制伺服电机使相机在水平和竖直方向移动并调整对焦观察鳕鱼和比目鱼的鱼苗运动状态。而真正将机器视觉技术用于水产养殖自动化的是从Foster等[12]人提出一种算法通过补光摄像头在水下检测残余饵料并对其进行计数开始，该算法为自动投喂设备的饵料使用量提供了一定的理论指导意见。而在鱼类行为研究方面，Israeli等[13][14]人利用两台摄像机观察在缺氧/氨胁迫环境下鲫鱼的行为特征，研究不同环境压力下鲫鱼的应激行为特征，该系统可作为一个辅助生产系统对养殖鱼类进行监控，提高了水产养殖鱼类福利。

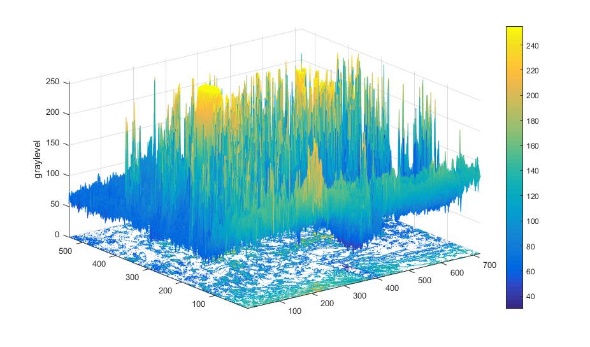
机器视觉技术因其无接触、高效、无损和低成本的优点被广泛的运用于水产养殖的各个领域。比如Zion等人利用不同光照条件下的CCD相机所获取的图像进行不变矩的分析实现了对鲤鱼，圣彼得鱼和鲻鱼的分类[15]，在此基础上为了进一步地提高算法的可用性，他们实现了对运动中的鱼进行分类[16]。但是这一类的研究一直处于实验室状态，算法对环境有较高的要求，直到Spampinato等[17]人提出了一套完整的鱼类检测，跟踪和计数的方法，该方法的所有实验都是在自然环境下进行并且对鱼类没有任何限制。Costa等[18]通过水下双目相机的校准获得了游动鱼的生理尺寸。Hu等[19]人基于多类别支持向量机（MSVM）根据鱼的颜色和纹理特征实现了鱼的分类算法。

目前通过机器视觉技术对鱼类的摄食行为研究主要有两个方向：通过观察残余饵料的方法间接反应鱼类摄食欲望和通过直接观察鱼类的运动行为评估鱼类的摄食欲望。对于残余饵料的研究Atoum Y等[20]人提出一种基于相关滤波的残余饵料检测方法，实现了对水面漂浮残余饵料的检测,但是由于该方法依赖于Bag-of-Words(BoW)模型，所以对采集图像的光环境有较高的要求；而通过鱼类运动行为直接评估摄食欲望的研究有基于前后视频帧差异对鱼类运动行为进行量化[21]。

随着我国集约化养殖尤其是循环养殖池（RAS）的蓬勃发展，国内有大量的研究者基于机器视觉技术针对鱼类的摄食行为进行了研究。周超[22]等人采用近红外摄像头获取鱼类运动图像，首先通过图像分割方法将鱼群从背景中分割出来，对分割后的图片求取每一条鱼的中心点，将这些点作为德劳内三角分割的顶点进行三角分割，基于三角分割后的三角形总周长提出一个描述鱼类摄食行为的鱼群簇拥指数FIFFB(flocking index of fish feeding behavior)。FIFFB指数很好的量化了鱼类的聚散效应，当鱼群聚拢的时候FIFFB指数会降低相反会升高，所以对于抢食性的鱼类而言也是一个鱼类摄食强度的量化指标。但是该方法仍然存在一定的局限性，第一点是图像分割有困难，由于该算法第一步需要对图像进行分割，当实际环境复杂时即便是红外摄像机也会存在分割失败的情况；第二点鱼类中心点难以求取，当养殖密度较大时鱼类容易出现交叠情况导致算法难以提取中心点。于欣等人[23]利用光流法（Lucas-Kanade）检测鱼群运动矢量，得到速度与转角的联合直方图和概率分布，基于标准互信息和局部异尝因子对鱼群进行异常检测，检测准确率达95%。将鱼类的摄食行为作为一种异常状态则该方法也可以用于鱼类摄食行为判单，但是该研究试验对象是作为观赏鱼的斑马鱼，试验条件比较理想，光照充足，成像质量较高，离实际生产环境下的运用还存在一段距离。

近几年得益于深度学习的突破，机器视觉也有了长足的发展，尤其是在图像分类[24]和目标检测[25]方面。大量的基于机器视觉的人工智能方法被运用到渔业各个领域比如识别鱼种类、行为分析、轨迹跟踪、活鱼识别和水质检测等。而对鱼类摄食行为的研究有基于CNN（convolutional neural network）的鱼类摄食强度评估。在文献[26]中，研究者将鱼类的摄食行为分为四类：未摄食、弱摄食、中摄食和强摄食，并人工标记了400张图片，每个类别含有100张，基于标注好的数据训练了一个CNN模型对鱼类的摄食行为分类，实验表明基于CNN的方法在该数据集下达到了90%的准确率。

机器视觉技术在运用到水产养殖虽然仅三十余年，但是得到了大力的发展与研究。最初的机器视觉技术仅仅是作为水产养殖中的监控和观察手段，随着集约化养殖与机器视觉的共同发展研究者们开始将机器视觉技术作为渔业信息自动化的一部分，机器视觉技术开始在渔业生产，鱼类质量监控，鱼产品加工等各个环节被运用。对鱼类摄食行为研究而言，基于机器视觉的研究逐渐从最初的图像差分方法和利用光流和能量模型等方法手工提取鱼类摄食行为特征到目前基于深度学习的卷积神经网络自动感知鱼类摄食行为特征的方向发展，并取得了一定的成果，为进一步的渔业信息自动化奠定了一定的理论基础。但是目前来看还是存在诸多的局限:(1)目前大部分研究就是基于实验室环境或者对工业生产环境的模拟，工业生产环境养殖密度更高，光环境更复杂，干扰因素更多；(2)工业生产环境下自动投饵设备对水面反光影响严重。当摄像机位于RAS正上方时，由于鱼类打破水面引起的反光效应是无法回避的问题，一些研究者想到利用反光效应来对鱼类的摄食行为进行研究[20]。当反光效应仅仅是有鱼类运动引起时的确是一个很好的研究思路，但是在工业养殖厂水面反光效应受到饵料的严重影响。图1.3a展示一个位于水面之上的相机拍摄的RAS投喂场景，可以看到图像中的鱼类几乎都被饵料打破水面引起的反光效应所遮挡，而且鱼类引起的反光效应被严重干扰，图1.3b所展示的灰度3D图则表明反光效应使得局部区域已经出现了过曝；(3)现阶段研究大多都关注鱼类摄食行为的空间特征，也就是从单张图像中提取鱼类摄食行为特征。而经过我们的观察与总结，我们认为对于鱼类摄食行为模式在高密度养殖时其时空特征更为显著。因为当鱼群处于某一模式下的时，其空间特征由鱼群所在相对和绝对位置决定，而当养殖密度增大时鱼类的空间特征也就越不明显（鱼类交叠越严重图像纹理特征越差），与此相反的是其时空特征由鱼群运动决定，当养殖越大鱼群的时空特征会越明显（运动所造成的混乱程度越高）。所以本论文的工作主要集中于对鱼类摄食行为的时空特征研究也就是基于视频的鱼类摄食行为分类，在下一节本论文将对视频分类的研究历史与进展做简要介绍。



(a) (b)

图1.3. (a)工业RAS投喂现场图；(b)灰度3D图

* + 1. 视频分类研究现状

视频分类的研究工作在目前主要分为三大类：通过手工设计的特征结合机器学习算法的传统视频分类、基于视频建模的分类方法和基于深度神经网络的视频分类方法。传统的视频分类算法的性能很大程度依赖于特征提取，特征提取越好分类效果越好。经典的特征提取方法主要有词袋模型（Bag of Words）、局部时空特征的梯度直方图（Histogram of Oriented Gradients）、光流直方图（Histogram of Optical Flow）和运动边界直方图（Motion Boundary Histogram）[27]，将这些方法所提取到的特征输入到机器学习算法中训练分类模型实现对视频的分类。

对于视频建模的方法最为著名的是基于动态混合纹理（Mixture of Dynic Texture）[28]的视频建模。MDT将视频假设为多个动态纹理的混合，每个动态纹理由隐含空间的向量和高斯噪声通过发射矩阵生成，同时不同视频帧的隐含空间向量由状态转移矩阵联系，其中MDT模型参数通过EM算法从视频中自动学习。MDT模型在对烟雾、火焰和水流等视频的建模获得了很好的表现，但是对于更为复杂的场景比如拥挤人群的建模表现欠佳。导致这种现象的根本原因在于MDT模型本质上是一个二阶稳态模型只能对较对变化平缓的纹理建模，对于变化更为复杂的纹理是难以满足的。

随着人工智能的发展，尤其是深度卷积神经网络（CNN）作为图像理解和特征提取的模型有了很大的进步。CNN采用卷积和池化操作，根据数据集图像和标签自动提取复杂特征完成图像的分类、识别、分割等工作。CNN在大量的静态数据集比如MINIST[28]、ImageNet[29]和COCO[30]等取得了空前的成功。基于CNN的研究成果，国内外的研究者开始将研究对象从静态的图像扩展为动态的视频。基于CNN的视频分类研究目前主要为端到端的CNN网络结构，比如3D CNN 和 双流法（Two-Stream）。3D CNN是对静态图像CNN的扩展，将卷积所用的卷积核从原来的二维变为三维，通过对卷积核参数的学习实现对图像的分类。3D卷积最大的局限在于网络参数复杂和无法处理长时间序列视频，相比之下双流法更适合长时间序列的视频分类。双流法的网络输入为视频帧和光流图像，前者为网络提供了空间信息而后者则负责时空信息的提醒。

谈到深度神经网络必须要提及是数据集的问题，为推动视频分类的发展，研究者们陆续建立了相关的视频数据集，如表1.1。小型标注良好的数据集有KTH、Hollywood2和Weizmann；中型规模的数据集有UCF101、Thumos-2014和HMDB51，这些数据集视频类别超过了50类；大规模数据集有Sports-1M、YFCC-100M、FCVID、ActivityNet和YouTube-8M等。

表1.1 视频数据集

| 数据集名称 | 样本数量 | 样本类别 | 发布时间 | 背景 |
| --- | --- | --- | --- | --- |
| KTH | 600 | 6 | 2004 | 静态 |
| Weizmann | 81 | 9 | 2005 | 静态 |
| Kodak | 1358 | 25 | 2007 | 动态 |
| Hollywood | 430 | 8 | 2008 | 动态 |
| Hollywood2 | 1787 | 12 | 2009 | 动态 |
| Olympic Sports | 800 | 16 | 2010 | 动态 |
| HMDB51 | 6766 | 51 | 2011 | 动态 |
| CCF | 9317 | 20 | 2011 | 动态 |
| UCF-101 | 13320 | 101 | 2012 | 动态 |
| THUMOS-2014 | 18394 | 20 | 2014 | 动态 |
| MED-2014 | 31000 | 20 | 2014 | 动态 |
| Sports-1M | 1133158 | 487 | 2014 | 动态 |
| MPII Human Pose | 20943 | 410 | 2014 | 动态 |

续表1.1 视频数据集

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集名称 | 样本数量 | 样本类别 | 发布时间 | 背景 |
| ActivityNet | 27901 | 203 | 2015 | 动态 |
| EventNet | 95321 | 500 | 2015 | 动态 |
| FCVID | 91223 | 239 | 2015 | 动态 |
| YouTube-8m | 8264650 | 4800 | 2016 | 动态 |

在水产领域由于缺乏源于生产环境的视频数据集目前还未有基于深度神经网络的鱼类视频行为分类研究。但是也有一些研究者基于传统视频分类的方法对对鱼类摄食行为进行了研究。Liu[7]等人通过观察鱼类摄食行为运动存在相互遮挡重叠的特点提出一个基于机器视觉的摄食行为指数CVFAI（computer vision-based feeding activity index）。CVFAI计算分两步进行：第一步计算视频前后帧像素值差异，第二步计算视频前后帧未被遮挡鱼的数量之和与养殖池内鱼群总数量的比例。CVFAI的值等于第一步计算的像素值差异与第二步计算的比例系数的乘积。经过实验验证，CVFAI与鱼类养殖专家对鱼类摄食行为的评估值相关性达到了91%。但是由于CVFAI算法对实验环境要求严格，视频图像中不能有其他的干扰因素，但在真是生产环境中肯定存在残余饵料或者其他因素的影响，所以要用于工程实践还是存在一定的困难。但是在另一方面，CVFAI的研究为我们基于视频相关关系的摄食行为研究奠定了理论基础，证明了其可行性。

* + 1. 课题组研究现状

本课题组多年来长期基于机器视觉对水产养殖领域课题进行研究，主要包括水产环境监测、运动中鱼类的尺寸测量和残余饵料计数等，且取得了一定的研究成果为集约化水产养殖信息化自动化做出贡献。

林超峰[31]以游鱼为研究对象，基于双目立体视觉设计了游鱼尺寸测量系统。乐九一[32]将阈值分割的图片进行细化提取鱼群骨架，基于细化后的图片端点对鱼群计数。

高明学[33]针对补光引起的中间亮四周偏暗的水下图像特性，受到局部阈值化方法思想的启发，提出了环带划分的最优阈值分割方法：依据图像的亮区重心定位环带中心，依据目标大小明确环带宽度大小，从而将图像分为一定数量的环状子图像，再采用局部阈值化的方法求得子图像的最佳阈值。作者对将该方法应用于鱼饵检测系统中进行了探索，对采集的饵料图片进行阈值分割，且用基于该算法的软件在广东湛江虾池进行了试验。考虑到环带划分的最优阈值分割方法要求成像平面对光的反射比较均匀、且需要与镜头发现垂直等局限性，刘桓宇[34]结合水下图像光照不均的现象，提出基于邻域灰度值的自适应分类分割算法来探测识别剩余饵料。让一个大小不变的滑动窗口自左向右、自上向下逐步扫描，每次只处理位于滑动窗口中心位置的像素点；结合EM算法和高斯混合模型去拟合滑动窗口内的灰度分布直方图，判断窗口内是否存在前景物体，然后进行分割；为解决“Otsu阈值会偏向方差较大的一类”的问题，又引入阈值补偿；得到分割图像后用基于轮廓跟踪的连通区域计数法统计饵料数量。为了进一步的提高算法在工程实际中的可用性，曹嘉恒等[34]等提出一种水下图像清晰化算法使得上述算法可以在能见度较低的水产养殖场所运用。

2009年，徐盛林、徐立鸿等研发用于鱼饵视频监控的图像解析与自动报警监控基站，第一套样机已经在广东湛江的虾池中试用，并取得初步成功；2010年，徐盛林、徐立鸿等申请软件著作权，基于图像解析算法的水产养殖鱼饵投放自动控制系统软件；2017年，徐立鸿、蔚瑞华等申请发明专利一种水产养殖的自动投饵和水质监测控制系统及方法[35]并进入发明公布；2019年，徐立鸿、蔚瑞华等所申请发明专利一种基于机器视觉的残留鱼饵计数算法[36]进入发明公布阶段。

* 1. 本文主要研究工作及组织结构
     1. 主要研究工作

由于在水产养殖领域目前仍然缺乏对鱼类摄食行为研究的视频数据集，本论文的第一个工作内容就是在实际工业生产环境下采集大西洋鲑鱼摄食行为水下视频数据。该部分的内容又可被具体描述为两个方面：实验数据采集系统的设计和在鱼类养殖专家的指导下对采集数据的标记。本论文第二项工作为基于变分自动编码的视频数据空间变换。如前在背景中所介绍，视频作为动态图像数据量十分巨大，以帧速率为30fps分辨率为1080960的视频为例则每五秒钟包含150张图，每个样本为一个三维矩阵其尺寸为。为了解决这样的问题我们提出一种基于变分贝叶斯的图像空间变换将视频帧映射到隐含空间的多元高斯分布，经过变换后的视频样本成为两个对应的高斯均值矩阵和高斯方差矩阵。本论文第三项工作为提出一种视频帧间关系建模的方法。对于具有不同模式的视频在隐含空间建立贝叶斯估计模型，根据不同模型对视频进行分类。

* + 1. 论文组织结构

本论文选题来自国家重点研发计划：“设施水产养殖智能化精细生产管理技术装备研发”，课题编号：2017YFD0701700。主要以真实集约化养殖环境下的大西洋鲑鱼作为研究对象，通过水下摄像机获取大西洋鲑鱼摄食行为视频数据，将获取数据处理标记并进行研究。本论文共分为五章，各章主要内容如下：

第1章，绪论。简单介绍了本论文的研究背景以及水产养殖中计算机视觉发展。由于本论文是基于视频分类对鱼类摄食行为研究所以我们回顾了在水产养殖领域对鱼类摄食行为的研究，同时也对视频分类的发展进行了简要的介绍。

第2章，介绍实验平台与数据采集与处理。首先介绍实验数据采集基地主要情况，并根据实际情况和数据采集要求设计数据采集方案和制作采集平台。完成数据采集后，在本章我们重点介绍了大西洋鲑鱼摄食行为水下视频数据库的标记和建立。

第3章，基于变分贝叶斯的视频数据空间变换。我们将像素空间下的视频帧数据通过变分贝叶斯网络变换到多元高斯分布的隐含空间，并且可以通过隐含空间下的变量复原为像素空间下的图像。

第4章，基于视频帧间关系的贝叶斯估计建模。我们对于鱼类的两种摄食行为（摄食、未摄食）分别采用贝叶斯估计进行建模，并利用得到的模型对鱼类摄食行为分类。

第5章，总结与展望。对于论文的主要工作进行总结，并分析了目前工作的不足之处，提出改进方向，对未来的研究工作做出展望。

* 1. 本章小结

本章首先对水产养殖中机器视觉的研究背景进行描述，然后回顾了机器视觉在水产养殖中的发展历史，同时对视频分类研究进行综述，提出了本论文的研究意义，最后介绍了本论文的研究工作以及全文的组织结构。