# 一．研究总要求

## 研究内容

### 3.低分辨率目标状态估计技术

研究移动平台获取的低分辨率视频目标状态估计技术，包含视频目标方位估计，位置估计以及面积/长度等几何特征计算。

## （三）关键技术

### 1.图像的超分辨重建技术

由于海警船载/无人机载移动平台获取目标视频信息时，受制于摄像头传感器参数限制和拍摄距离的影响，比较难以获取满足后续研究要求的较高分辨率的视频/图像信息。因此，需要在原有的低分辨率的视频/图像进行增强预处理，以获得后续所需要的较高分辨率的视频/图像。采用图像超分辨重建实现上述目的。

图像的超分辨率重建是指通过硬件或软件的方法提高原有图像的分辨率，通过一系列低分辨率的图像来得到一幅高分辨率的图像过程。

超分辨重建主要分为两种。Image SR，只参考当前低分辨率图像，不依赖其他相关图像的超分辨率技术，称之为单幅图像的超分辨率。Video SR，参考多幅图像或多个视频帧的超分辨率技术，称之为多帧视频/多图的超分辨率。本研究采用的关键技术为基于深度学习的单幅图像的超分辨率重建。该模块采用主要的关键技术包括：

①基于SRGAN的深度网络模型的设计与搭建

②深度超分辨率网络超参数的选取与模型的训练

③SRGAN损失函数的设计

### 低分辨率目标状态估计技术

根据目标实时分类检测和识别跟踪的结果，对目标进行状态估计。主要包含两部分，通过光电系统和导航定位系统对视频中的目标方位、位置进行估计，得到目标的基本位置信息；通过对视频中的船只图像的轮廓进行提取、分析与计算，对视频中的目标长、宽面等计算，得到目标的基本大小信息。该模块的关键技术包括：

①相机工作模型与坐标系的建立

②像素坐标到实体坐标的映射

③目标定位与地理位置估计

④目标长宽与面积的估计

# 二．研究工作进展情况（重点说明）

## (一)研究内容的完成情况

### 3.低分辨率目标状态估计技术

研究内容完成情况：

根据研究进度的要求，现已经完成无人机平台相机工作模型的构建，图像坐标系与世界坐标系的构建，推导完成像素位置坐标与实际位置坐标的换算，进行了相机非线性畸变矫正，目标位置相对坐标与实际经纬度坐标的换算，目标面积大小的估计。

目标实现情况：

实现了在给定相机主要内参数，和无人平台相关飞行参数（飞行高度，经纬度坐标，航角）的前提下，根据视频目标实时分类检测与跟踪技术输出的结果，对目标的位置信息、方位信息、长度、宽度、面积等基本状态的粗略估计。基本能完成项目所规定的该方面的任务。

## (二)关键技术的攻关情况

### **1.基于深度学习的超分辨率重建技术**

#### 1.1超分辨率研究的基本概念与研究思路

对于端到端的超分辨率网络，其研究的一般的流程如图所示：



图1-1 超分辨率研究的一般流程

评价超分辨重建的效果，选取PSNR作为参数指标。计算公式如下，式中H和W为图像的长和高。





#### 1.2基于SRGAN的超分辨重建网络的搭建

SRGAN是一种经典的超分辨网络，引入对抗生成网络(GAN)更好的拟合超分辨重建的实际效果。由于其采用了对抗式设计，能够更好的生成类似于原图相的纹理细节，使图像显得更为真实。SRGAN设计了Generator网络和Discriminator网络，Generator网络输入低分辨率图像输出高分辨率的图像，Discriminator网络辨识Discriminator网络输出的高分辨率图像和实际的高分辨率图像，给出Generator网络的效果得分来对Generator网络进行反馈。通过这样的依次轮流优化Generator网络和Discriminator网络，下图为GAN网络的基本结构，该结构同样适用于SRGAN。



图1-2 SRGAN总体结构

**1.2.1 Generator网络与Discriminator网络结构**

Generator网络是目的网络，输入低分辨率图片输出高分辨率图片，其基本结构如图所示：

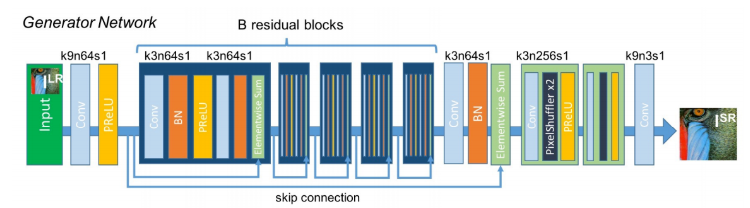


图1-3 Generator网络结构

首先将低分辨率图片进行预处理，裁剪成指定规格大小的块状，之后输入到Generator网络，经由一个Conv和PReLU初步提取网络特点。然后输入至由Conv、Batch-norm和PReLU依次重复连接组成的深层的res-net结构进一步提取图像的特征，最后输入到pixel-shuffler层进行非线性映射并且升采样至高维度，最终输出高分辨率的图片。

Discriminator网络是鉴别网络，输入的是实际高分辨图像或者由Generator网络输出的高分辨率图像，输出一个该网络认为图像的高分辨程度的得分。训练Discriminator网络，使其达到如果高分辨图像的清晰程度越好，得分也越高的效果。Discriminator网络的结构下图所示：

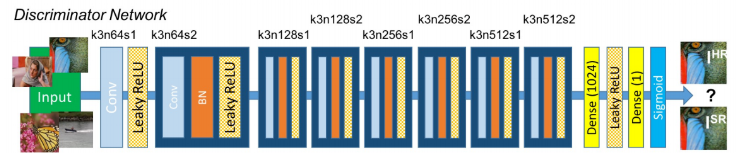


图1-4 Discriminator网络结构

和Generator网络类似，Discriminator首先将输入的高分辨图片输入到Conv和Leaky RelU进行初步特征提取，之后输入到多层的res-net结构进行深度的图像特征的提取。与Generator网络不同，Discriminator网络在最后设计了全连接层，经由sigmoid输出最后的得分。

**1.2.2 SRGAN的交叉训练策略**

由于采用了GAN的网络结构，对网络的训练，采用交叉优化训练的策略，使得Generator网络输出高分辨率的图像在Discriminator的得到一个较高的得分，同时训练Discriminator这个网络使得能够较好区分真实的图片和超分辨重建后的图像，下述公式可以概括训练策略。



式子中，D表示Discriminator网络，G表示Generator网络，、分别表示两个网络的参数。

**1.2.3 SRGAN的损失函数**

为了更好的恢复高分辨率图像的纹理信息，SRGAN改进了常用的均方误差（MSE）损失函数，引入感知损失函数（perceptual loss function），公式如下：



式子中，表示内容损失（content loss），表示对抗损失（adversarial loss）。

内容损失采用采用19层VGG网络预训练的ReLU激活层作为基础的VGG loss，有效解决了常规的MSE损失函数带来的高频细节缺失的问题，公式如下：



式子中，和表示VGG19网络中特征图谱的维度。

对抗损失体现了SRGAN中Generator网络对感知损失函数的影响，一部分损失函数使网络通过“欺骗”判别器从而偏向生成输出更接近自然图像的输出。公式如下：



如图为在公开数据集上SRGAN超分辨重建得到的效果：



图1-5低分辨率图片  


图1-6 SRGAN重建图像

### 4.低分辨率目标状态估计技术

#### 4.1 相机工作模型与坐标系的建立

为了便于进行像素位置与实体位置的映射，首先需要建立坐标系，进而构建相机的工作模型。常用的坐标系包括图像坐标系、相机坐标系、相机坐标系。

图像坐标系包括，坐标系和物理坐标系。如图3-1所示，像素坐标系是以图像左上角的为原点的坐标系，像素点的坐标用表示，和代表该像素点所处位置的行数和列数，像素点的值代表该像素点的颜色信息。物理坐标系即为坐标系，物理坐标系代表图像的物理长度，其中一般位于图像的中心，也就是光轴与成像平面的交点，物理坐标系的单位一般是毫米。在像素坐标系下的坐标为，假设每个像素点在物理坐标系中和轴上的尺寸大小为和，则像素坐标系与物理坐标系之间的转换公式为：



****

图3-1 图像坐标系示意图

相机坐标系，也叫光心坐标系，是一种三维坐标系，其坐标原点定为相机的光心轴与光轴重合，与成像平面垂直，轴和轴与物理坐标系平行。如图3-2所示，其中到的物理距离即为焦距*f*。



图3-2 相机坐标系示意图

图像坐标系到相机坐标系的转换方程为:



世界坐标系为对任何物体进行描述，图3-2中即为点在世界坐标系下的坐标。相机坐标系与世界坐标系之间的变换为刚体变换，其变换关系如下式所示，其中表示旋转矩阵，表示偏移向量。



相机的工作过程即为将三维实景映射为二维图像的过程，理想状态下可以用经典的针孔相机模型来描述这一过程。如图3-3中为相机坐标系，点为相机的光心。相应的，为成像平面，点是点经过光心在成像平面上的投影。



图3-3 针孔成像模型示意图

记点和点在相机坐标系下的坐标分别为和根据相似三角形定理可得



#### 4.2 像素坐标与实体坐标的映射

总结上述公式，可以推导得到



式子中称为相机内参矩阵，称为相机的外参矩阵。

#### 4.3 非线性畸变矫正

假设线性模型下点p的投影点坐标为，畸变后的投影坐标为，则有



畸变误差主要分为径向畸变和切向畸变两种，径向畸变主要来自于透镜形状的误差，切向畸变来自于相机组装过程中出现的误差。如图3-4所示，为切向失真，为径向失真。



图3-4 相机的畸变误差示意图

径向畸变的数学模型可以用泰勒展开式来描述，下式中，代表畸变点在成像平面上的原始位置坐标，代表经过矫正后的位置坐标。





切向畸变由透镜和成像平面不平行而产生的，可以用如下公式来描述





非线性畸变成像模型主要由五个参数来决定的，可以用向量来表示。

#### 4.4 目标的位置估计

将无人机的云台相机设置为始终垂直向下，并且通过传感器信息可以获得实时的相机位置（即无人机位置），从而基于相机模型建立如下定位模型。



图3-5 检测目标位置估计模型

如图3-5所示，相机平面与地面的距离即高度为，根据标定测量得到的相机焦距为，根据上述相机成像模型可得



由此可以得到的坐标，从而即可获得目标位置与中心位置在和轴上的物理距离，其中点和点的坐标均为经过畸变矫正后的坐标。

根据传感器传输的数据可得点的经纬坐标，由此可以得到点的经纬坐标，又根据上述内容得到了点在坐标系下的坐标，因此下一步需要将其转换为经纬坐标。



图3-6 物理距离与经纬坐标转换示意图

如图3-19所示，以为原点建立坐标系，坐标系与经纬度方向保持一致，坐标系代表无人机的飞行方向，轴方向是无人机飞行的方向，与正北方向的夹角为，，由于点**在坐标系中的坐标已知，而由于坐标系相当于对坐标系做旋转变换，因此点**在坐标系中的坐标满足



则点与点在经纬坐标方向上的物理距离为与，点与点之间的距离可由勾股定理求得



因为地球半径为，因此纬度方向上的一米转换为纬度差为



由此可见点的纬度坐标为



又因为已知，根据半正矢公式（Haversine formula）即可计算出经度差，从而计算出经度坐标。





其中，和是点与点的纬度，和和是与点的经度。

#### **4.4 目标大小的估计**

根据目标实时分类检测与识别跟踪技术得到的结果，可得到目标区域的像素坐标信息，利用前面所述方法将区域像素坐标信息转化为实际区域的物理坐标信息，计算实际区域面积即可得到目标面积大小。

## (三)技术指标实现情况

### 3.低分辨率目标状态估计技术

经实验验证，在给定相机内参数、无人机平台主要参数的前提下，能够后实现粗略的估计目标相对于无人机的相对位置信息、经纬度等绝对位置信息和目标的大小信息。

## (四)阶段成果及其应用情况

### 3.低分辨率目标状态估计技术

主要成果：

1. 设计了基于单目相机模型的低分辨率目标状态估计方案
2. 编写了无人机平台低分辨率目标状态估计的算法与程序
3. 基于输入图像验证了低分辨率目标状态估计的程序，能够实现基本的功能

应用方向：

低分辨率目标状态估计技术应用于在较差的侦察环境下，采用计算机视觉方案对所侦察目标进行初步的估计。在军事上可用于战场信息侦察，为敌我双方动态的分析提供帮助，对于构建现代化信息化作战装备具有非凡的意义。民用上，亦可以推广至搜索、救援、科研等相关领域，为社会现代化建设贡献力量。

# 三．下一步工作安排与建议

## (一)下一步工作安排

### 3.低分辨率目标状态估计技术

（1）扩充与优化样本数据集

在设计低分辨率目标状态估计技术方案的前期，比较缺乏专用的海上目标的数据集。因此，在验证低分辨率目标状态估计时，采用航拍地面车辆数据集替代海上目标数据集，先行验证了实验的效果。在下面的工作中，会寻找或者制作专用的数据集，对算法进行更精确更充分的验证，以确保算法满足实际任务要求。

1. 整合目标检测算法与目标状态估计算法

根据任务要求，目标状态估计是在目标实时检测的结果的基础上进行的，但是在任务前期，目标实时检测与目标状态估计是同步进行的，因此之前在设计目标状态估计的算法时，假想已经得到了目标实时检测的结果。在之后的工作中，将整合这两方面的算法与程序，确保目标状态估计的输入参数为实际的目标检测的输出结果，完成项目任务的要求。

1. 优化目标状态估计的程序结构

之前的工作的主要目的是达成任务的基本要求，实现基本的目标状态估计的功能，因此在设计算法与编写程序时存在很多细节的缺失与结构的不合理之处。在之后的工作中，会补充缺失的细节优化算法程序的结构，以提高算法主体的可读性与迁移性。

## (二)建议

### 3.低分辨率目标状态估计技术

主要建议：

①优化与扩充样本数据集

②优化目标状态估计算法

③整合目标检测技术与目标状态估计技术

④优化程序结构，提高其可读性与迁移性