

# 姚焜元

男 | 27岁 | 应届生 | 硕士 | 13120572916 | kunyuan.yao@gmail.com | 现居上海 | 机器学习、深度学习

Github: <https://github.com/beidongjiedeguang/>

## 教育背景

2011.9 - 2015.6

巢湖学院

学士 - 物理学

2016.9 - 2020.?

上海理工大学

硕士 - 光学

曾主修：四大力学，热力学与统计物理，计算物理学，光学原理，傅里叶光学，光散射原理，高等电磁场理论等。

计算机语言：

熟练度由高到低：Python、MATLAB、C、Mathematica、bash/shell、c++、SQL、c# 等

熟悉机器学习框架与工具：Torch、Tensorflow、espnet、sklearn、skimage、OpenCv、Numpy 等

深度学习：熟悉 GANs、word2vec、Encoder-Decoder框架、Attention、Transformer、FasterRCNN等

系统：Linux，容器：docker

## 实践和项目经历

2019.8 - 2020.3

地点：叠纸科技有限公司 部门：技术中心(AI组)

### 1. 文本转语音 (TTS)

基于开源项目[espnet](<https://github.com/espnet/espnet>)，其vocoder使用的是网络是Parallel waveGAN，相比WaveNet或ClariNet，Parallel waveGAN的训练时间明显更短(仅需3天)且其inference速度约为实时的27倍(在TITAN RTX上)。整个项目分四个阶段：

- Front end：使用开源工具speech-aligner或者Montreal-Forced-Aligner获取字素到音素的对齐以及音素规则字典。
- Feature Generation：使用kaldi特征抽取。使用80维的log Mel谱。
- Training Text to Fbank：使用的Transformer框架。训练音素到Mel谱的映射。
- Training vocoder：基于Parallel waveGAN框架。训练Mel谱生成音频。

### 2. AI 作画：

这是一个非常有趣的项目。为了生成一些独一无二的图片，初期准备用GAN进行生成，但DCGAN生成的图片由于转置卷积的原因栅格现象太严重，而BigGAN需要的显存又太大。所以这个项目图像生成过程没有用神经网络。最终我们实现了对给定图片进行油画风格的渲染效果。油画效果的实现得益于以下几点：油画反光效果由不同波长的光线在介质中的吸收不同计算得到。（相关技术源于*Physically-Based Modeling Techniques for Interactive Digital Painting*）；由于绘制时，图像数据是一笔一笔覆盖上去，而两个均有一定透明度颜色的混合显然不是简单色彩叠加，颜色在人眼对众多不同波长光波的信息压缩，不同波长光波对人眼的三种光锥细胞刺激不同激活了大脑对颜色的感知，这也是为何三原色可以模拟出所有颜色的原因，将像素的alpha通道与光学透射率关联，可以得到人眼对颜色正确感知的色彩叠加，所以我们



在计算中模拟了这一过程。效果图：以及使用自定义的浮雕算子实现不同角度的阴影。绘制优化方向是对整个图像的均方根误差的负梯度方向进行。

3. 基于HMM并应用于Unity中的头部姿势识别，其中模型主要复现于论文*Robust Head Gestures Recognition for Assistive Technology*
4. 基于快速傅里叶的一维及二维卷积的实现：对于一维情况，由于FFT的复杂度为 $O(N \log N)$ ，而普通卷积复杂度为 $O(N^2)$

)，而基于FFT的卷积会经历两次FFT以及一次元素乘法，所以复杂度依旧为 $O(N\log N)$ 。目前已经

5. 复现论文Painterly Rendering with Curved Brush Strokes of Multiple Sizes 中笔触绘制模型。

6. 个人项目：（以下项目均可在我的个人Github中找到）

Mie散射模型以及FFT分析——提供对平行光照射下球形颗粒光强分布相关可视化操作，可视化基于streamlit。

guang——主要是一些平时会用到的工具，如机器视觉的一些常用方法，如图像的滤波去噪，直线、圆形检测、条纹频率检测，以及一些NLP中常用方法：音素对齐、音频多进程重采样、HMM等。

pyprobar——好看的进度条模块。地址：<https://github.com/beidongjiedeguang/python-progress-bar>

syncnote——和云端服务器同步的模块

espec——设置特殊日子提醒的模块，支持农历和阳历

2017.9 - 2019.6

地点：上海理工大学

1. 微信小项目：项目整合了人脸识别、聊天机器人、爬虫、分词、词云等。项目中发送和接受微信数据基于itchat；微信上发送照片，使用人脸识别技术判断此人可获得的权限，人脸识别使用的是dlib进行定位然后PCA降维后用CNN进行的识别；聊天机器人自动回复功能：分别使用了两种模式，一是基于训练好的语料模型，二是调用图灵机器人API。词云功能是为方便PPT：给我发送一份文本文件（支持格式PDF、docx、txt），会自动返回一张精美的词云构成的图像。
2. 基于侧向散射光的球形颗粒参数反演：  
根据Mie理论可计算得到球形颗粒的散射光强分布，其分布与两个性质相关——颗粒折射率与粒径。我所做的工作便是根据散射光强分布反演出颗粒折射率与粒径。其中涉及到的技术有：几何光学模型、Debye级数、二分求根、FFT滤波去噪。这里问题求解的角度主要还是通过解析的方式导出的，只是求解过程中涉及到一些数值计算的方法。
1. "华为杯"第15届全国研究生数学建模竞赛（多无人机对组网雷达的协同干扰）：  
将问题抽象为一个0-1整数规划问题。通过构建出一个步长矩阵（就是一个无人机的可行解空间的矩阵），利用遗传算法寻优的方式求得一个较为满意的结果。
1. Richardson外推方法应用至导师的广义Lorenz-Mie理论计算程序中的多维积分中提供了约百分之10的性能加速。
2. 用numpy写过的算法如Logistic回归、K-Means等等。

## ● 奖项以及证书

- "华为杯"第15届全国研究生数学建模竞赛 二等奖 (在团队中担任模型建立以及程序编写角色.)
- 第十七届全国中学生物理竞赛市级一等奖
- 上海理工大学理学院2018年度羽毛球双打第二名
- 电子工程与电气自动化学院“高校师范生教学技能竞赛” 二等奖
- 研究生奖学金三次; 中国颗粒学会证书; CET4; Python二级证书; 教师资格证书;

## ● 个人评价

酷爱编程、物理与数学。从最开始时的计算物理，编程求解线性方程组、微分方程、多维积分等等，也正是这些数值计算方法让我认识到利用计算机在解决这些数学物理问题时的便捷性！研究物理是发现自然中的规律，而计算机中却可以创造规律。在这条道路上可能会遇到很多挫折与阻碍，希望我能坚定走下去。