

数字图像处理报告八：小波变换与深度学习的结合

姓名：鲁国锐

学号：17020021031

专业：电子信息科学与技术

2020 年 5 月 7 日

目录

1	题目描述	2
2	问题背景 ¹	2
2.1	傅里叶变换的局限性	2
2.2	解决方案：小波变换	3
3	小波变换与深度学习的结合：构建超分辨率图像	3
3.1	用深度学习的方法来预测小波系数	4
3.2	用小波变换来分离“客观质量”和“感知质量”	4
4	总结	4

1 题目描述

在之前的作业中我们探讨过如何将频率域知识运用到深度学习当中，本次作业请大家继续探索如何将小波变换与深度学习相结合进行图像处理。

2 问题背景¹

2.1 傅里叶变换的局限性

傅里叶变换作为一个提取信号频率信息的工具，有着广泛的用途。通过对一个信号进行傅里叶变换，我们可以非常清楚地看到该信号中包含了哪些频率成分。然而，很多时候，我们不仅仅想知道一个信号中包含了哪些频率，还需要知道这些频率出现的时间。

于是这里引出了两个概念：**平稳信号**和**非平稳信号**。简而言之，这两种信号的区别就是：**平稳信号中各频率始终存在，不随时间变化**；**非平稳信号则刚好相反**。显然，对于平稳信号，由于各频率成分不随时间改变，只需要用傅里叶变换得出其中有哪些分量及其大小相位就足以全面地描述该信号了。但是对于非平稳信号，除了知道各频率成分外，我们还需要知道每一个成分是在什么时候出现的，否则无法唯一确定一个非平稳信号。

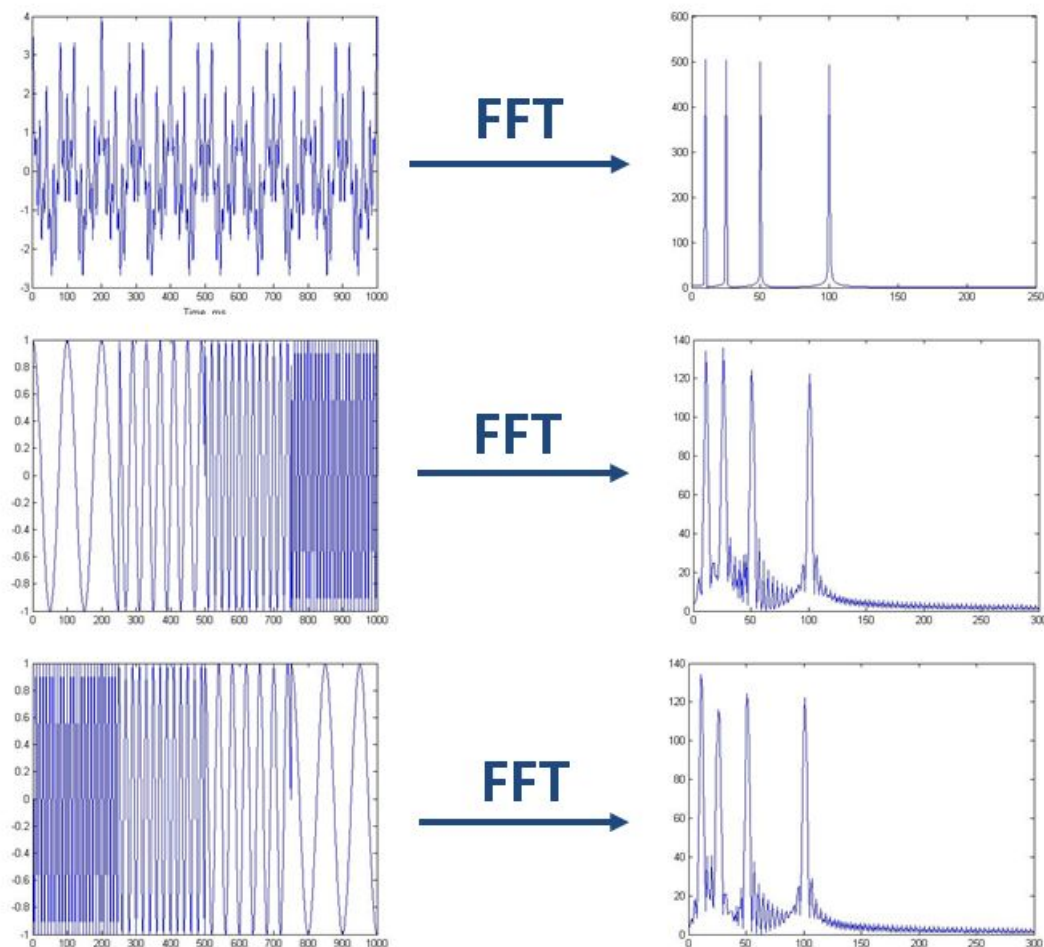


图 1: 平稳信号和非平稳信号及其傅里叶变换幅频特性

如图1所示，最上面的一行表示的是一个平稳信号及其傅里叶变换，下面两行是两个不同的非平稳信号及

¹本节的写作参考及图像来源：

<https://www.zhihu.com/question/22864189/answer/40772083>

<http://users.rowan.edu/~polikar/WTtutorial.html>

其傅里叶变换。首先对比平稳信号和非平稳信号的傅里叶变换，可以看出非平稳信号的傅里叶变换除了有一些类似于噪声的成分外（根据[第 2 条参考链接](#)中的解释，这是由信号从一个频率转换到另一个频率造成的），其余大体相同。接下来再对比一下两个非平稳信号的傅里叶变换，可以看出，尽管从时域上看两个信号明显不一样，但它们的傅里叶变换却几乎没有差别。显然，这样的情况使我们所不愿意看到的。

2.2 解决方案：小波变换

针对2.1节中所说的问题，人们想了各种解决方案，其中包括短时傅里叶变换（STFT），即给信号“加窗”，分段做傅里叶变换。但 STFT 中的窗口大小是不变的，显然这样是无法满足分析大多数非平稳信号的要求的。于是人们又做出改进，提出了小波变换（WT）。只不过小波变换没有在“加窗”这条路上继续走下去，而是将无限长的三角函数基换成了有限长的、会衰减的小波基（如图2所示）。

小波变换

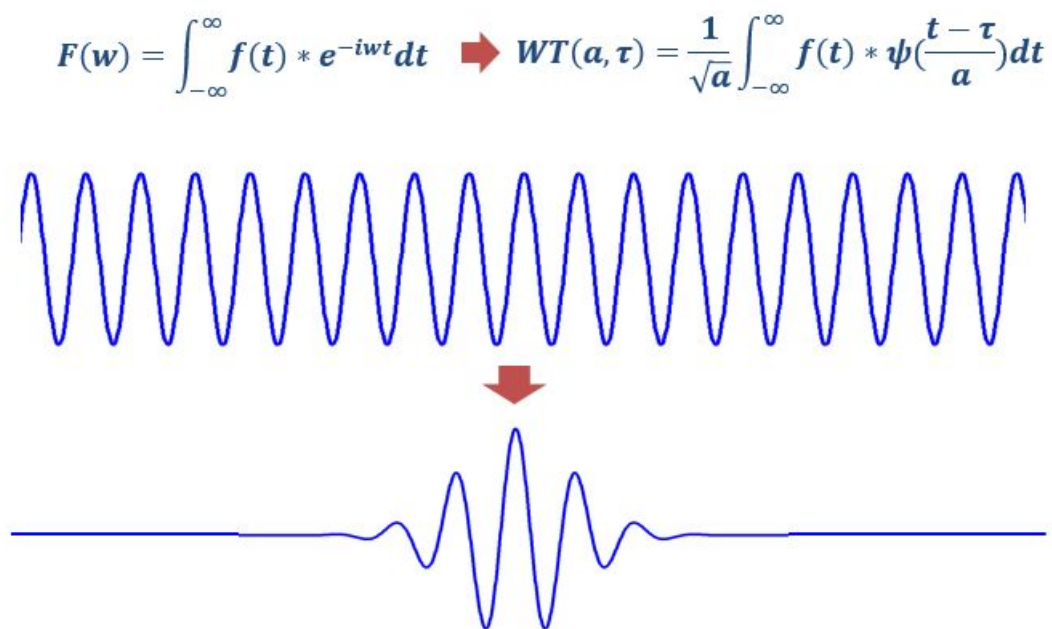


图 2: 小波变换将无限长的三角函数基换成了有限长的小波基

通过对尺度 a 和平移量 τ 的调整，就可改变基函数的频率（与尺度成反比）和作用的时间（对应于平移量）了。相应的，变换后的结果 $WT(a, \tau)$ 也就可以反映出信号所包含的各频率成分及其出现的时间了。

个人理解，如果用一首乐曲来比喻的话，傅里叶变换会给出这首曲子中出现了哪些音符（频率），而小波变换则能够直接给出这首曲子对应的乐谱（频率对应音符，时间对应音符出现的顺序及节奏）。

3 小波变换与深度学习的结合：构建超分辨率图像

关于小波变换与深度学习结合的文献似乎并不多，而在完成本次报告的过程中，搜集到的文献主要又都集中在超分辨率图像的构建上 [\[1-4\]](#)，可见小波变换与深度学习的结合在该领域上确实有优势。故在这一小节中，我们将专门介绍小波变换与深度学习在构建超分辨率图像中的应用。

另外再记录一下本次报告中文献的搜集方式：

1. 首先在网上找到了一篇 ICCV2019 的论文；

2. 然后在论文中搜索关键词 “*wavelet*”，在其参考文献中找到所有包含该关键词的论文；
3. 因为深度学习是在 2012 之后开始变得广为人知的，所以剔除所有早于 2012 年的文献；
4. 再依次去读剩下论文的摘要，判断是否符合本次报告的要求；
5. 找到符合要求的文献后，再对该文献重复上述步骤，最终找到本次报告中用到的 4 篇论文。

3.1 用深度学习的方法来预测小波系数

在 [3,4] 两篇文献中，都采取了将低分辨率图像作为输入，用深度学习的方法预测其对应的高分辨率图像的小波系数来达到提升结果质量的目的。仔细阅读之后发现，这两篇论文在大体的流程设计上都比较相似，具体可以描述为：

1. 输入一张低分辨率图像，得到其一组特征映射 (*feature maps*)；
2. 将得到的特征映射输入到一组子网中，用于预测出高分辨率图像的小波系数；
3. 根据得到的小波系数重建出高分辨率图像。

[2] 中也用到了通过神经网络生成小波系数的方法，但其具体方法我并没有看懂，只能介绍到这里。

3.2 用小波变换来分离“客观质量”和“感知质量”

在 [1] 中，作者指出，在过去的超分辨率图像构建任务中，GAN 的对抗损失函数虽然可以促使网络生成高质量的细节，但是这些细节不在正确的位置上，最终会导致图像中出现很多伪影，出现失真。同时文中也指出，由于客观质量 (*objective quality*) 和感知质量 (*perceptual quality*) 受不同因素影响，如果将它们作为一个整体进行优化，将很难在两者之间找到平衡。于是作者通过使用小波变换来分离出影响客观质量和感知质量的因素，然后对其分别进行提升。

具体而言，可以用小波变换将图像分解为低频子带和高频子带，其中低频子带主要影响客观质量，高频子带主要影响感知质量。然后用一个增强网络来优化低频子带，用小波域风格迁移 (*wavelet domain style transfer*) 方法来提升高频子带的质量。这样做可以使结果在客观质量和感知质量之间达到一个更好地平衡。

4 总结

本次报告大概是自学这门课以来难度最大的一次了，尽管将大多数时间都花在了学习小波变换上，但在阅读论文时还是遇到了不小的阻力。一方面小波变换本身理论性很强，难度较大，通过自学只是了解了一些皮毛，论文中用到的很多知识点都未曾接触过；另一方面，目前对于深度学习更深入的知识掌握得还相当有限，也使得阅读论文的难度加大。

不过，虽然有很多地方不明白，但在阅读这些文献的过程中，也能明显感受到小波变换能够在很大程度上反映出图像质量，尤其是一些细节上的信息，这对于构建超分辨率图像有着相当重要的意义。同时也可以想见，小波变换在其它相关领域，如图像修复、图像增强等方向上，也应当能起到非常重要的作用。

参考文献

- [1] Xin Deng, Ren Yang, Mai Xu, and Pier Luigi Dragotti. Wavelet domain style transfer for an effective perception-distortion tradeoff in single image super-resolution. In Proceedings of the IEEE International Conference on Computer Vision, pages 3076–3085, 2019. 3, 4

- [2] Xing Gao and Hongkai Xiong. A hybrid wavelet convolution network with sparse-coding for image super-resolution. In 2016 IEEE International Conference on Image Processing (ICIP), pages 1439–1443. IEEE, 2016. [3](#), [4](#)
- [3] Huaibo Huang, Ran He, Zhenan Sun, and Tieniu Tan. Wavelet-srnet: A wavelet-based cnn for multi-scale face super resolution. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 1689–1697, 2017. [3](#), [4](#)
- [4] Zhisheng Zhong, Tiancheng Shen, Yibo Yang, Zhouchen Lin, and Chao Zhang. Joint sub-bands learning with clique structures for wavelet domain super-resolution. In Advances in Neural Information Processing Systems, pages 165–175, 2018. [3](#), [4](#)