

计算机视觉大作业报告

——Use of Laplacian Pyramid: Enhancement & Coring

南开大学软件学院软件工程专业 2019 级 2 班 韦诗睿 1913184

1 Introduction

图像金字塔常用于图像的多尺度表达，其为一系列以金字塔形状排列的、分辨率逐步降低且来源于同一张原始图的图像集合。有两种类型的图像金字塔常常出现在实际运用中：

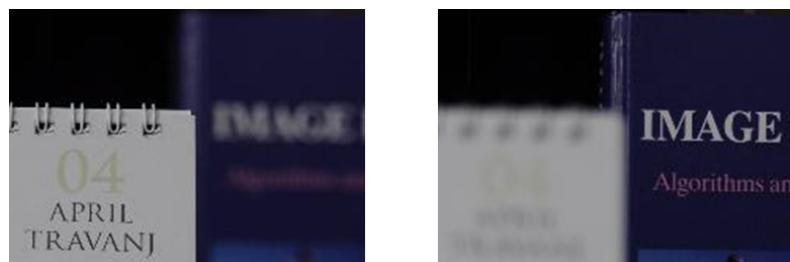
- 高斯金字塔（GP）：最基本的图像金字塔。将原图作为最底层图像并利用高斯核对其进行卷积，再对卷积后的图像进行下采样得到上一层图像；将此图像作为输入，反复迭代以上操作，形成一个金字塔形的图像数据结构，即 GP。
- 拉普拉斯金字塔（LP）：在 GP 的运算过程中，图像经过卷积和下采样操作会丢失部分高频细节信息。用 GP 的每一层图像减去其上一层图像上采样并高斯卷积之后的预测图像，得到一系列的差值图即为 LP。

LP 由于保存了图像的高频部分/残差，因此常配合 GP 完成图像的重建的工作。但这并不是 LP 仅有的应用场景，针对高频部分做文章，LP 还可以应用于图像的融合、压缩、增强以及降噪等方面。本次大作业实验即探究 LP 在图像增强（Enhancement）和降噪（Coring）两方面的效果。

2 Enhancement

2.1 Background

在摄影时，相机的焦距不同则所拍摄下的图像对焦的区域也不同，相机的快门速度不同则所拍摄下的图像的对比度和亮度也有所不同。如何从一组具有不同焦距的图像中重建一副高保真度的图像/全区域聚焦的图像是本节研究的重点。¹



¹ Enhancement 实验中数据集来源：<https://dsp.etfbl.net/mif/>

2.2 Strategy

实际上，这个问题和图像的融合问题没有本质区别。图像的融合问题即输入 n 幅图像，通过掩膜选取每个图像想要保留的部分，不同层的掩膜与相应层的 LP 相结合并重建后，即可在同一幅图像上融合多幅图像的各个部分，且融合效果平滑不突兀。

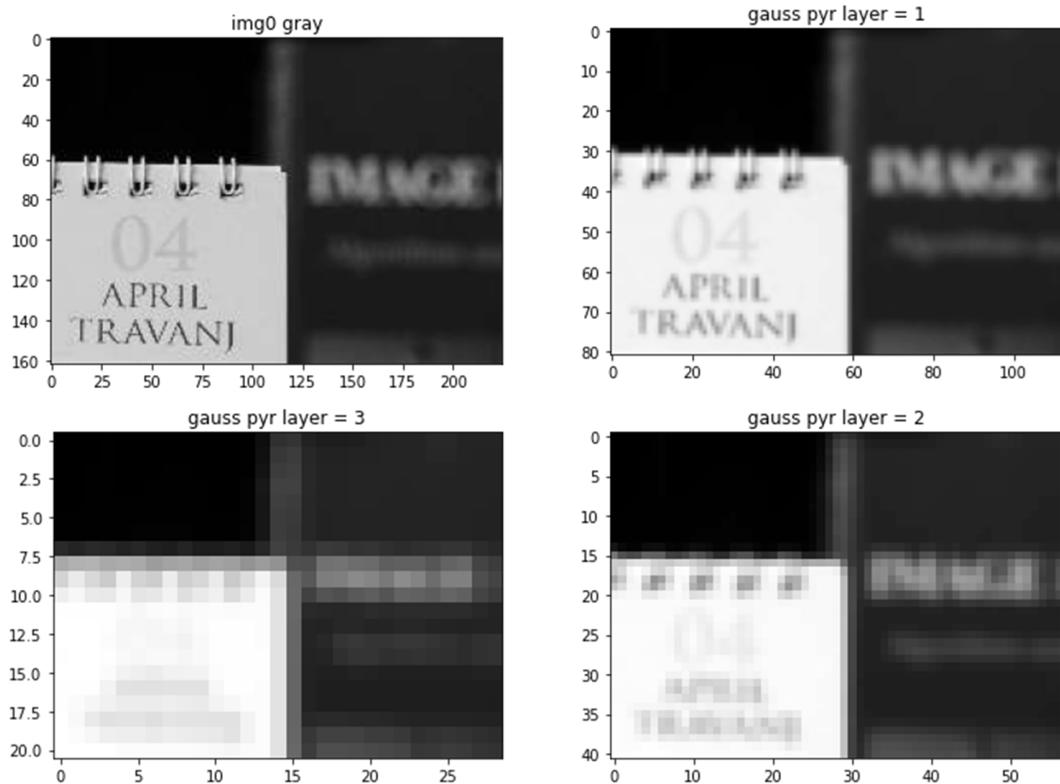
在这个问题上，将每幅图像各自对焦的区域结合在同一张图像上，便可得到全区域对焦的图像。而判断图像上某个区域是否对焦，即衡量两张图像的 LP 相应层相应位置的转换系数（Transform coefficients） b 。 b 存储的是图像的高频信息，高频就代表了图像的边缘、细节或噪音， b 的绝对值越大（ b 可以为负数）就代表该位置的细节越丰富。若有两张拍摄同一场景的图片 A 和 B，当在 LP 的第 j 层的某位置 n 处存在 $|b_{A,j}(n)| > |b_{B,j}(n)|$ ，则代表 A 图像对比 B 在此处对焦， $b_{A,j}(n)$ 需要留下作为最终图像的组成。而对于 LP 的最顶层即保存了图像的低频部分，反应的是基本的图形轮廓和信息。若拍摄的同一场景相同，则图像的低频部分大致相同，取 A 和 B 的 LP 在该层的平均值即可。最后对融合了两张图信息的 LP 进行重构，则可得到全区域对焦的图像。

2.3 Implementation

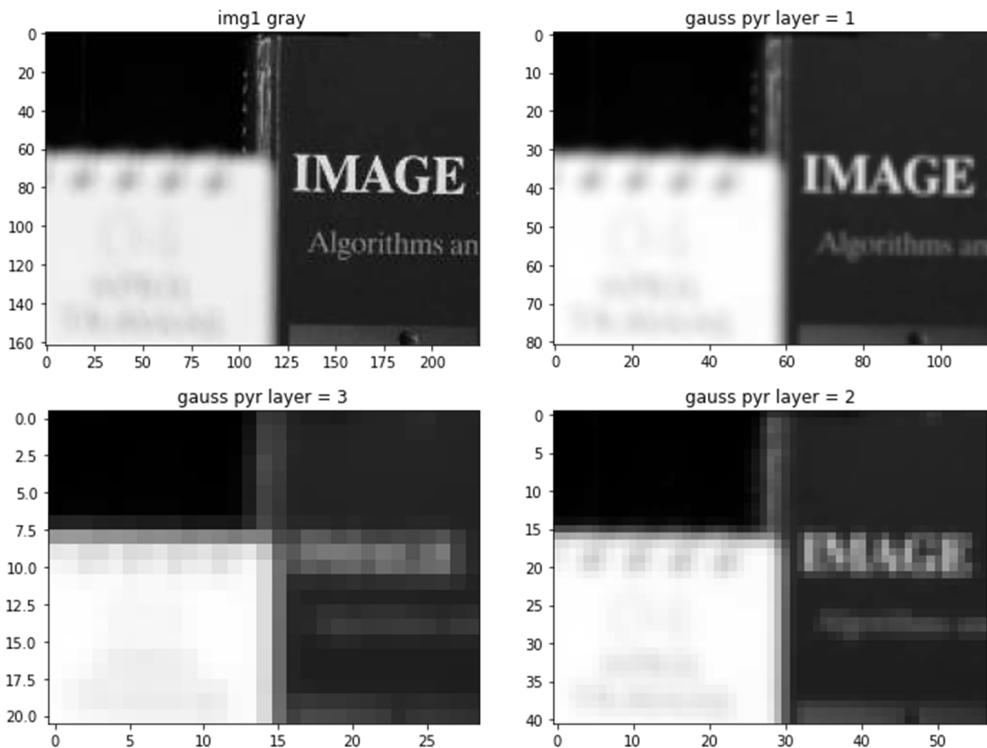
由于彩图可以分成三通道的灰度图分别处理，因此先以灰度图作为实验主体。

2.3.1 输入两张图像 A 和 B，分别进行高斯金字塔的构建。

Img0 灰度图以及三层高斯金字塔（不包括原图）：



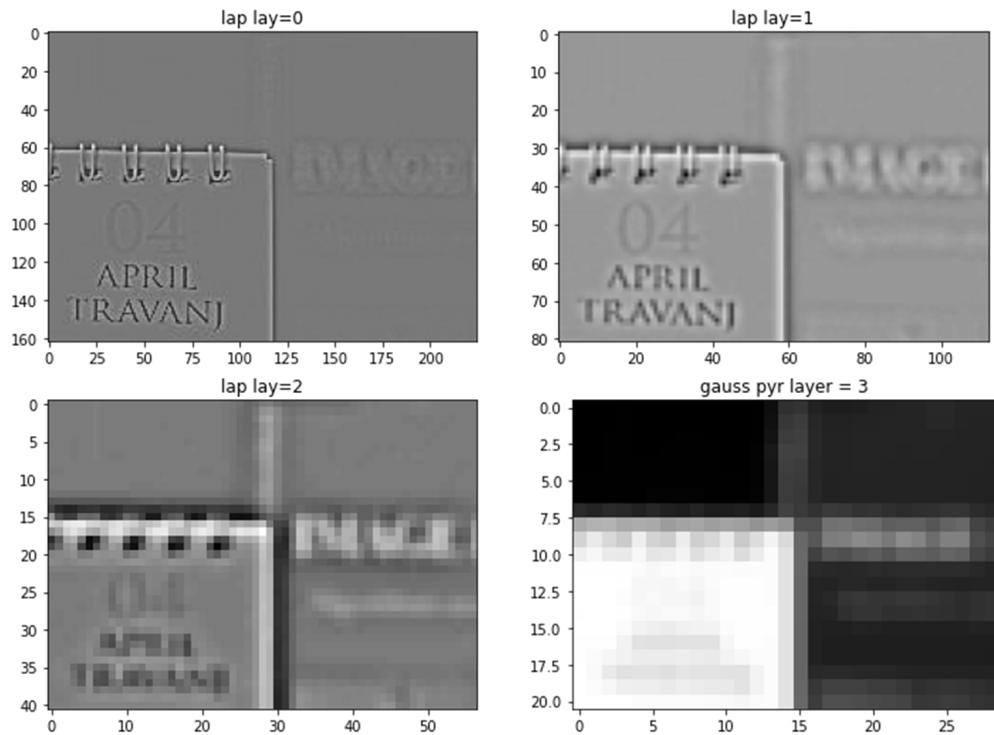
Img1 灰度图以及三层高斯金字塔：



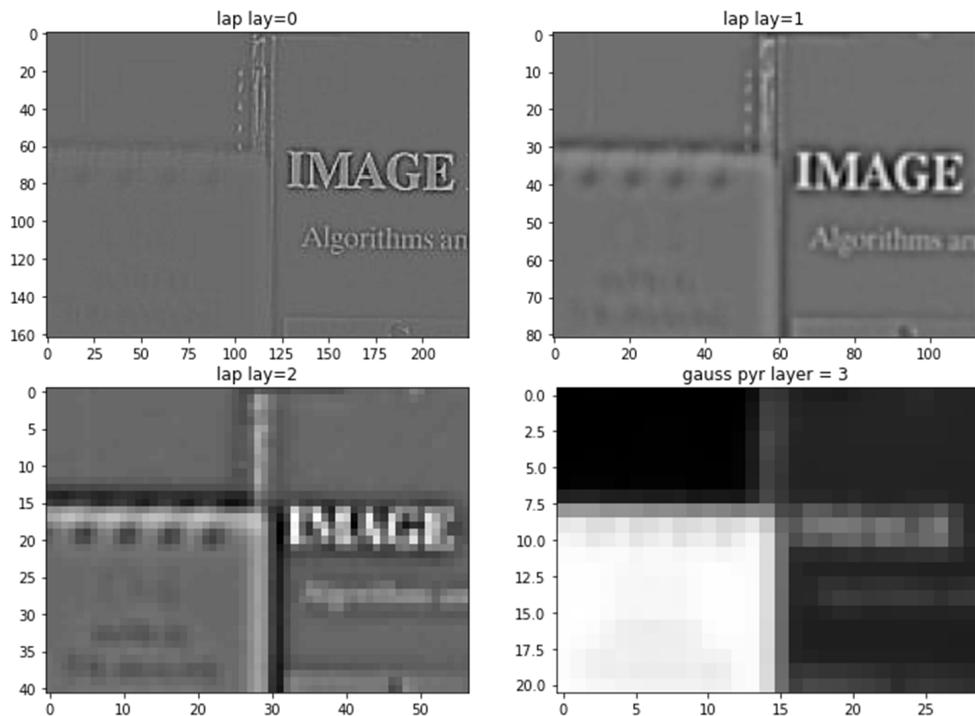
可以看到 Img0 对焦在左前方日历；而 Img1 对焦在右后方书本。随着 GP 层级的增长，图片越来越模糊，因为应用了多次降采样后留下的低频信息相对上一层也越来越小。

2.3.2 对两张图片分别做拉普拉斯金字塔的构建：

Img0 的拉普拉斯金字塔（包括最顶层的低频部分）：



Img1 的拉普拉斯金字塔：

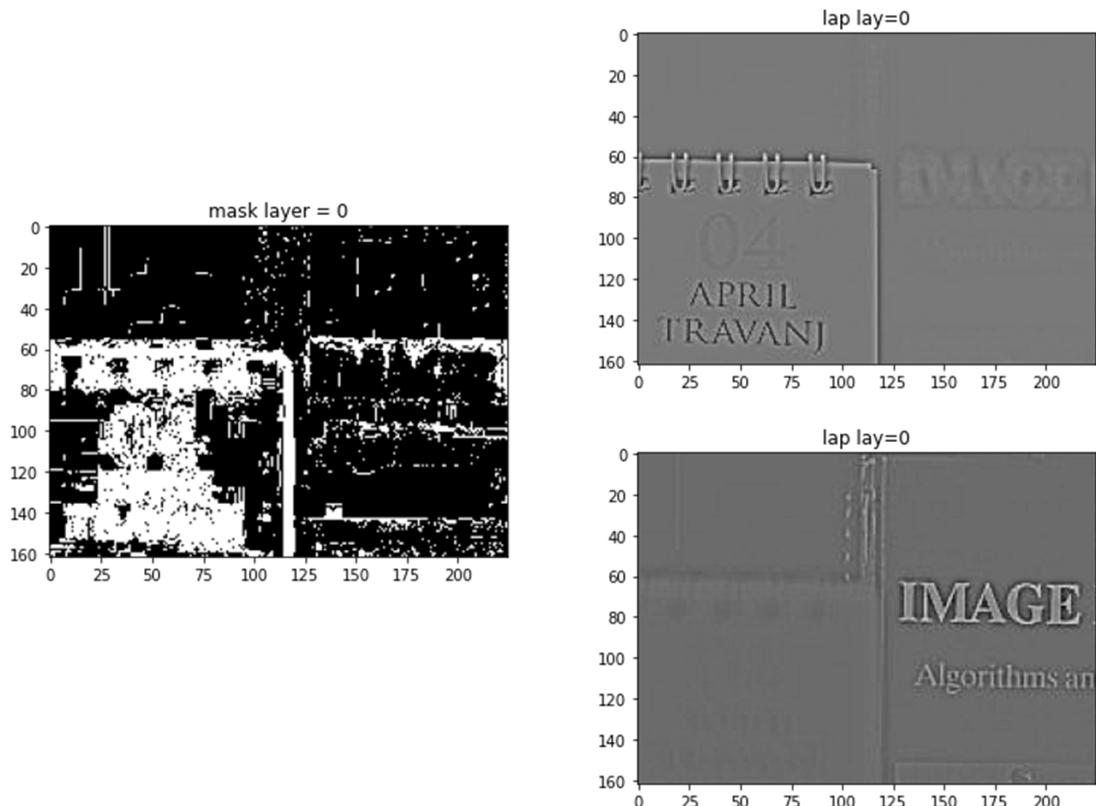


可以看到，随着金字塔层级由低到高，图像越来越模糊，因为高频的信息相比上一层也越来越小（绝对值）。

在这里需要注意，由于转换系数 b 可以为负数，因此在构建的时候需要将数组类型转为 int 而不能保持 uint。

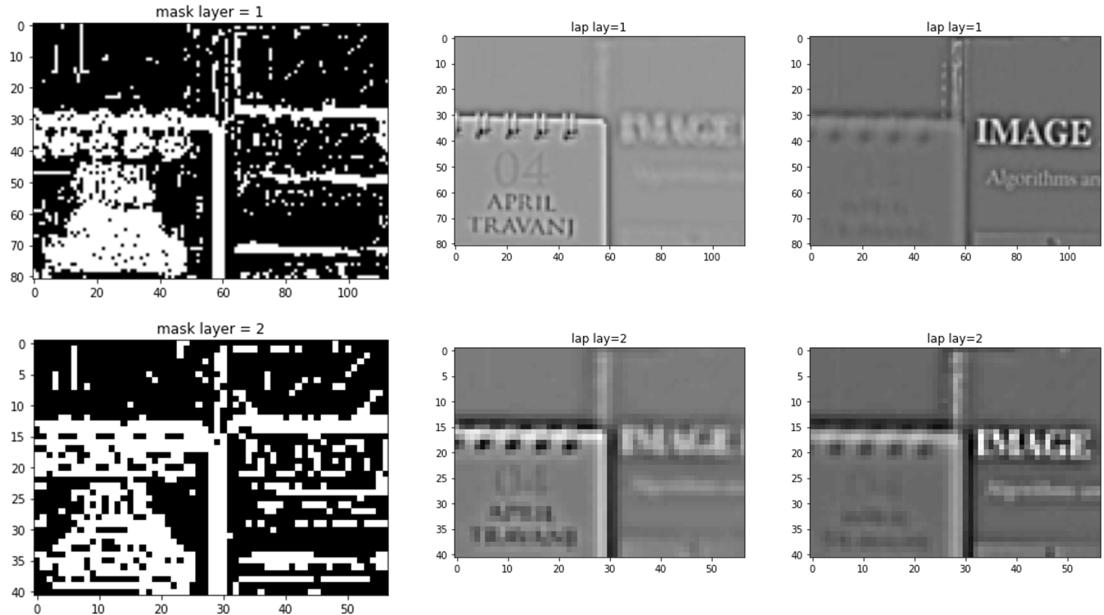
2.3.3 对两个拉普拉斯金字塔进行转换系数的选择：

对两个 LP 的同一层的每一个位置的转换系数进行遍历，并使用掩膜 mask 将 b 值保存下来。若在某处 A 的系数大于 B，则掩膜在该处的值为 1；反之，则为 0。下图为对 LP 的最底层进行掩膜计算：



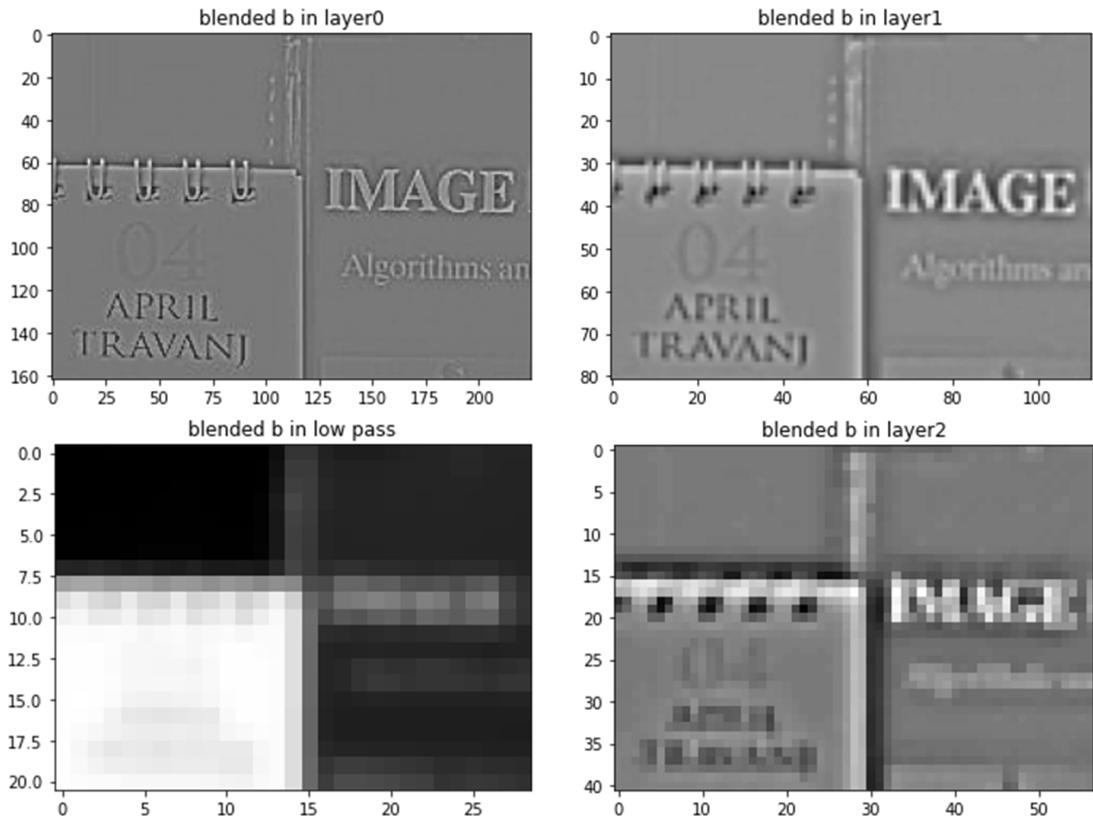
可以看到，右上方 A 图在日历处对焦，因此掩膜在此处较亮；右下方 B 图在书本处对焦，因此掩膜在此处较暗；

其他层同样符合以上规律：



2.3.4 结合掩膜融合两个拉普拉斯金字塔：

掩膜保留了每个位置最高频的信息（除顶层外），得到的融合图像如下：

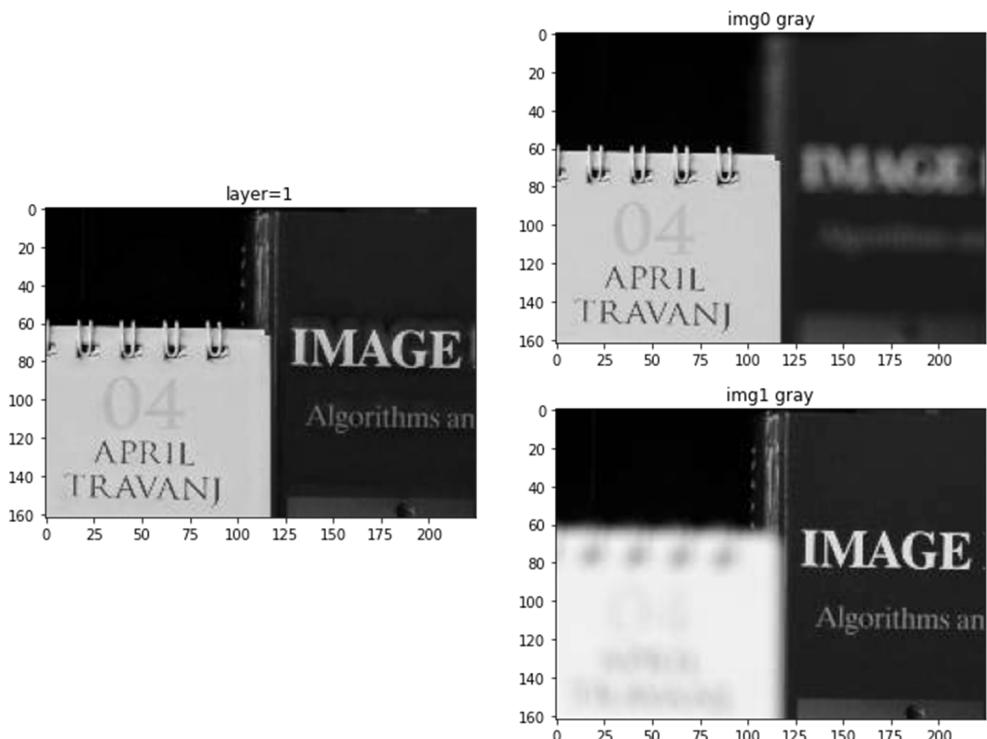


可以看到两张图的高频信息已经融合在了一起，一幅图上既有日历处的对焦也有书本处的对焦。左下角为图像低频部分的平均值融合。

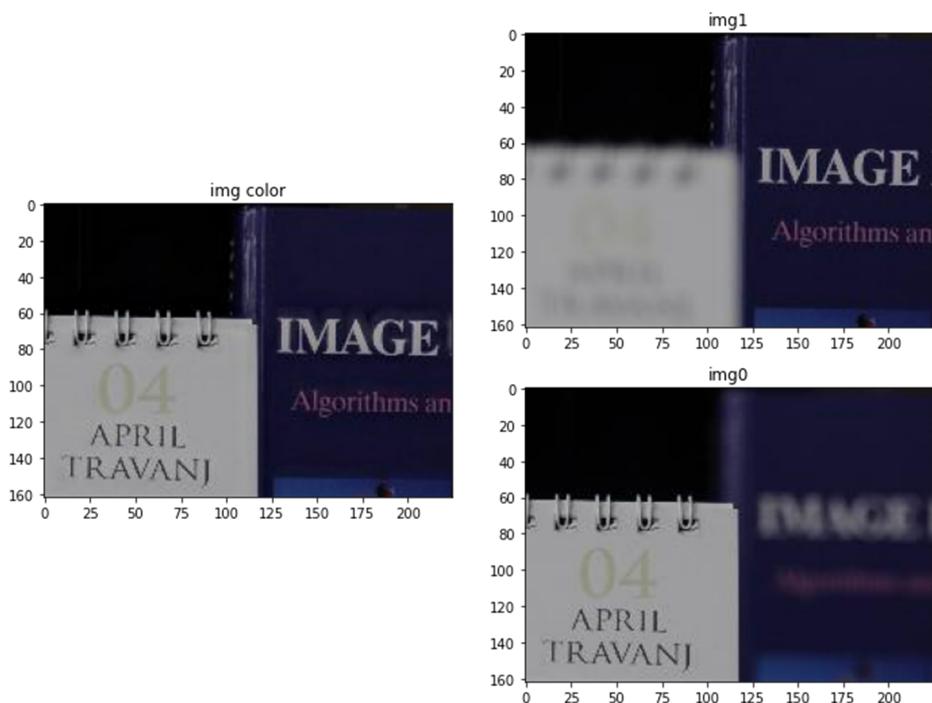
需要注意的是，在融合平均值时也需要先将图像转为 int 格式方能进行加操作。

2.3.5 图像重构：

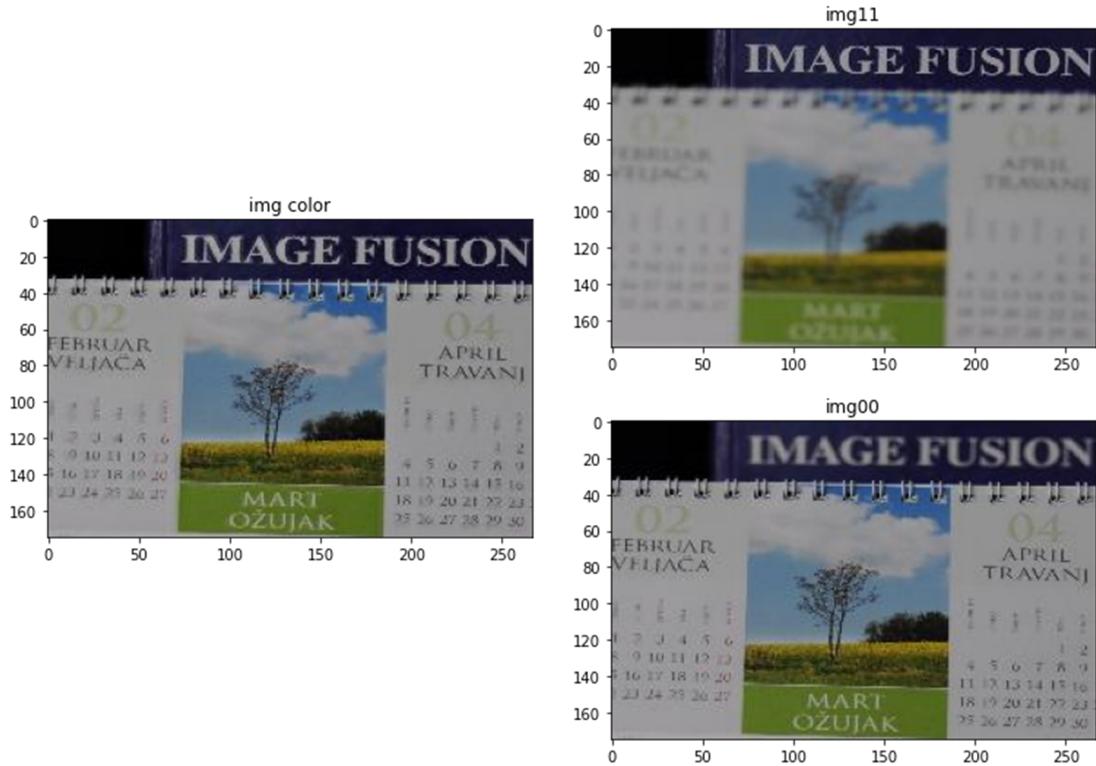
对融合金字塔进行重构，得到了最终的全区域对焦图：



对该图的三通道进行如上操作，得到的效果如下：



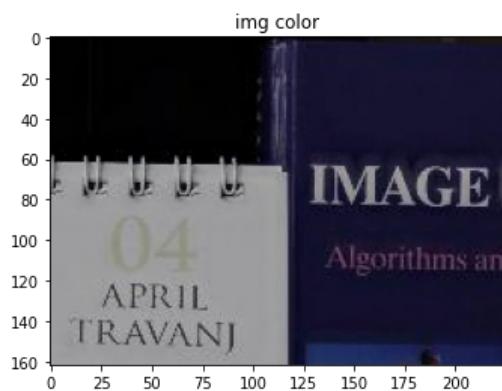
对另一个场景的图进行实验也得到了很好的效果：其中 img11 对焦区域为后方书本；img00 对焦区域为前方日历。



2.4 Discussion

2.4.1 鬼影 (ghost effect)

在融合两张拍摄了同一场景的图像时，我们默认两张图像的像素点得到了对齐，即两张图片相同位置拍摄的是相同的现实点；但实际上不同时间拍摄的图像，会因为相机的抖动和拍摄物体的移动导致内容有所差别，使得最后的融合效果会出现鬼影，即最后像素点出现局部的移动²。可以看到，处理的图像的后方书本“IMAGE”字样右上方存在鬼影。



² 图像处理 exposure fusion 图像曝光融合[EB/OL].https://zhuanlan.zhihu.com/p/455674916

因此在融合时还需要对像素点进行对齐处理，尽可能减少两张图片的差别。可以参考的一种方法或许是基于区域选取而不是基于每一个像素点，如果 A 图的某 $n \times n$ 的区域内每个转换系数 b 的绝对值的最大值大于 B 图，那么选择 A。³这种方法或许能够降低对像素点对齐的严格要求，减少融合像素的错误选取，也能够一定程度上减少噪声的影响（因为噪声也属于高频信息，也有可能被选中）。

2.4.2 彩色图像处理

对于彩色图像，本文采取的方法为三个颜色通道分开处理。但有可能每一个通道每一个像素点的高频或低频部分来自不同的来源，或许在 R 通道某像素处选取了图像 A 的信息；而在 G 通道同样的像素点处却选取了图像 B 的像素点信息，融合在一起出现异常的彩色图像。可以采取的方法为使用灰度图进行判断，若在灰度图处理时在某点处选取了图像 A 的信息，则三个通道在该点处都需要选取 A 的信息，保证三通道的一致性。⁴由于本文处理图片时未遇到颜色异常问题，因此不做过多处理。

3 Coring

3.1 Background & Strategy

当得到一个被污染的图像且仅知道图像上所添加的噪声的模型时，如何还原得到干净的图像，这是本节重点研究的问题。而要解决这个问题，首先需要说明几个概念。

3.1.1 稀疏表示与去噪

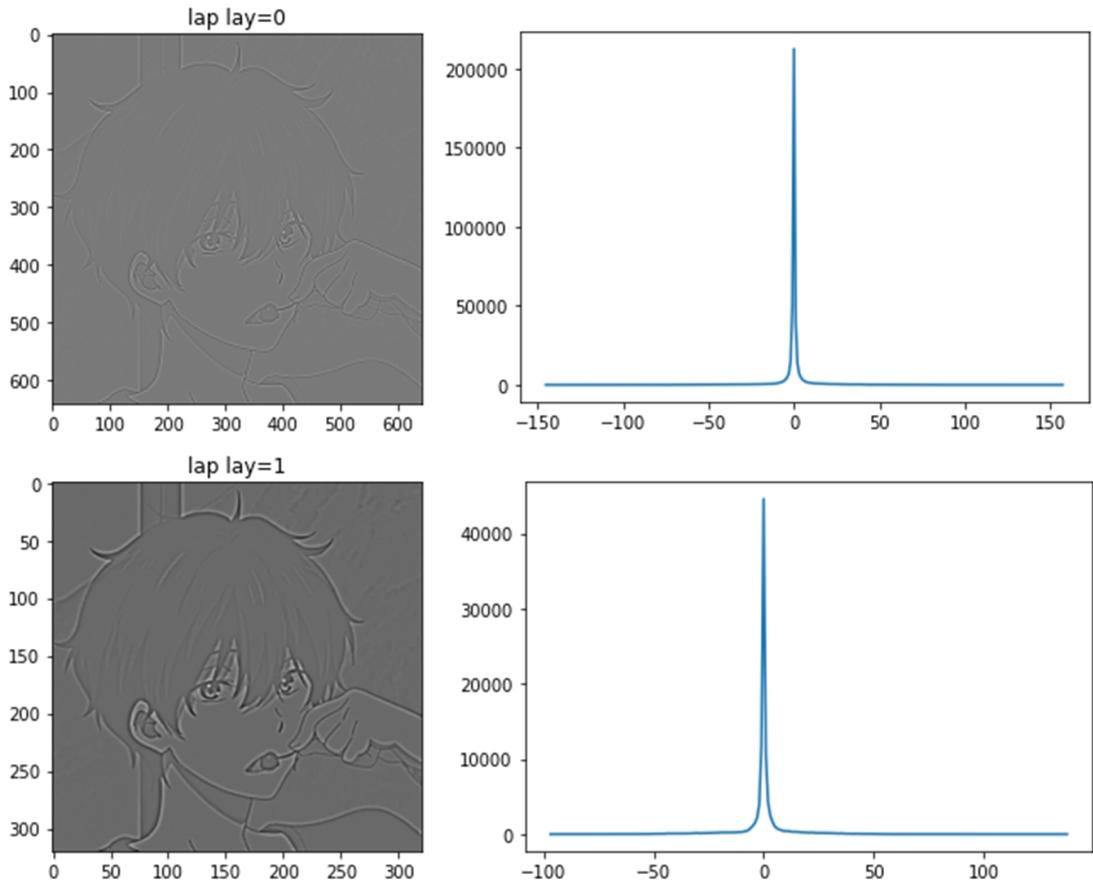
稀疏信号定义为：若信号仅有有限非零采样点，而其他采样点均为零或接近于零，则称该信号是“稀疏的”(sparse)。自然图像信号中，可以稀疏表示的情况较少，很多采样点值很小但并不为零，因此提出了另一种概念“可压缩的”，即如果在不丢失大部分信息的前提下，信号经过任何变换后是稀疏的，那么可以称之为可压缩信号。自然图像信号多数是可压缩信号，因为在经过了计算拉普拉斯金字塔后，得到的转换系数大多数值在零附近。如下图，在对图片进行 LP 计算后对每一层转换系数计算直方图以观察其分布，可以发现直方图/概率密度函数顶峰在 0 处，且尾部很宽（sharply peaked at zero with broad tails.）⁵

³ 图像融合的规则[EB/OL]. [<https://wenku.baidu.com/view/c8ae11adf61fb7360b4c65c4.html>]

⁴ 基础的拉普拉斯金字塔融合用于改善图像增强中易出现的过增强问题

[EB/OL].<https://www.cnblogs.com/Imageshop/p/10620935.html>.

⁵ E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," Proceedings of 3rd IEEE



可以认为含噪的图像是无噪的原图和噪声两者合成的图像，原图即自然图片被认为是可稀疏的，而噪声是随机的不可稀疏的。因此我们可以提取图像的稀疏成分，用这些稀疏成分来重构图像；而在这个过程中，噪声被处理为原图和重构图像之间的残差，在重构过程中残差被丢弃，从而达到去噪的效果。⁶

3.1.2 Coring

Coring 操作是图像去噪领域很常用的一种方法，大致概念为：图像信号被分为多个带宽（bands），高通带宽受到阈值的影响，低幅（low-amplitude）的值被抑制而高幅（high-amplitude）的值被保留。之所以能够使用 Coring 操作去噪的原因有以下几点：

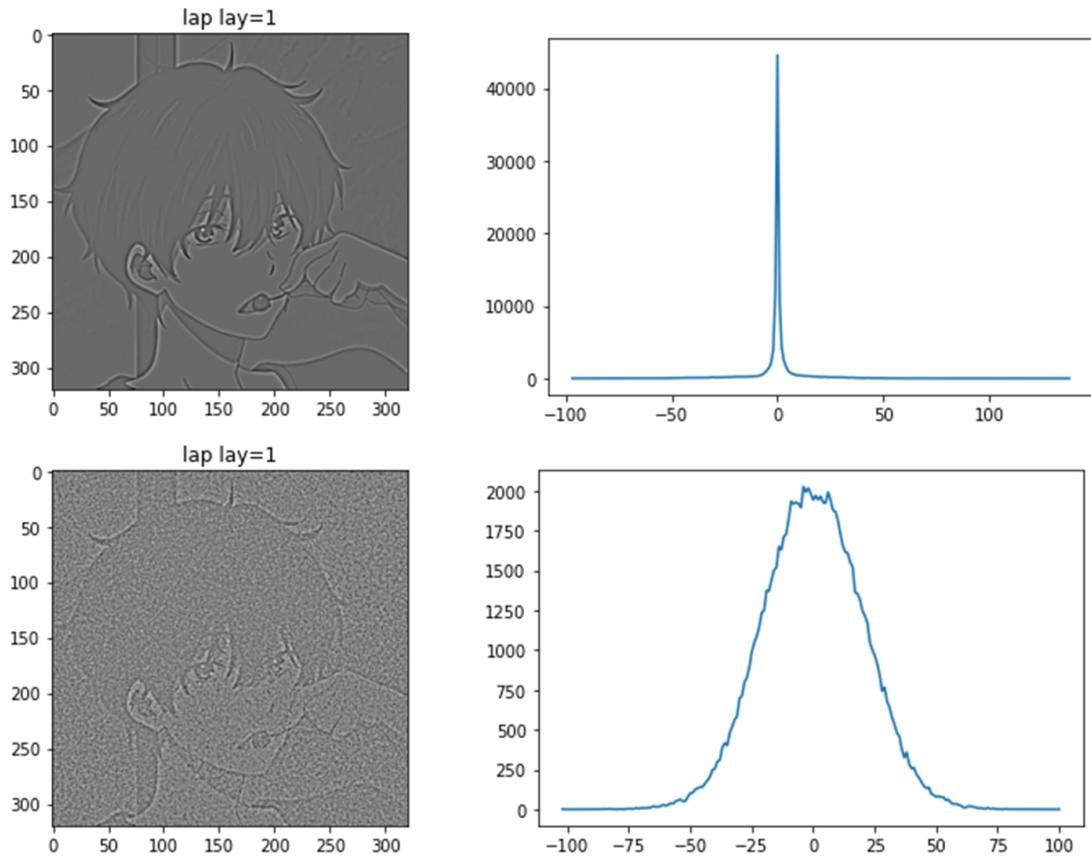
- 自然图像的样本峰度（概率密度在均值处峰值高低的特征）和加了噪声的图像样本峰度存在明显差异，而 Coring 操作正是需要依赖这种显著差异。⁷下图为同一张图像加高斯噪声前后 LP 同一层所计算的直方图，其中上方为未受污染的图

⁶ International Conference on Image Processing, 1996, pp. 379-382 vol.1, doi: 10.1109/ICIP.1996.559512.

⁶ 稀疏表示去噪的理解[EB/OL]. <https://blog.csdn.net/tongdanping/article/details/79162547>

⁷ E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," Proceedings of 3rd IEEE International Conference on Image Processing, 1996, pp. 379-382 vol.1, doi: 10.1109/ICIP.1996.559512.

像，下方为加了高斯噪声的图像，可以看到两者的直方图存在明显差异，加了高斯噪声的直方图虽仍在零点处顶峰，但分布函数“宽度”比起未加噪声明显增加。



- Coring 的原理在于接近 0 的系数被认为可能表示噪声，因此会被抑制/衰减；而如果值比较大，则认为它可能代表原本的图像信息，需要保持不变或略微放大。这样的理解非常直观：由于在 $x=0$ 处给定的信号概率分布函数非常大，因此很小的 y 值（0 附近的值）会被认为是从 $x=0$ 处产生的（given the substantial signal probability mass at $x = 0$, small values of y are assumed to have arisen from a value of $x = 0.$ ）⁸

- 也可以换种方式理解：原本自然的图像的概率密度函数顶峰非常尖，而加了噪声（高斯噪声后）0 处顶峰高度降低，而分布函数形状“变宽”。因此为了维持或尽可能从受噪声污染的分布回复到原本的分布形状，则将零附近某个阈值内的值设为 0，其他数值维持分布，不仅顶峰高度回升，信号宽度也有所“变窄”。

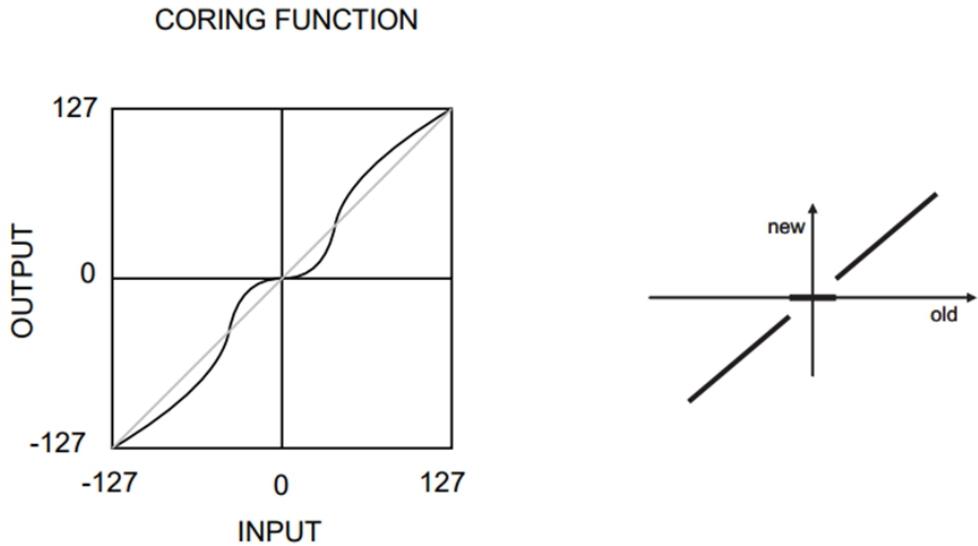
综合以上原因，Coring 操作可以通过如下图的 Coring Function 表示⁹，其中

⁸ E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," Proceedings of 3rd IEEE International Conference on Image Processing, 1996, pp. 379-382 vol.1, doi: 10.1109/ICIP.1996.559512.

⁹ 图片来源：左图：Edward H. Adelson, Subband coring for image noise reduction. Internal Report, RCA David Sarnoff Research Center, Nov. 26 1986. 右图：<https://www.cs.toronto.edu/~jepson/csc320/notes/pyramids.pdf>

零值附近 (small amplitude) 的值被抑制而较大的值 (large amplitude) 被保持。

本次实验将会采取右图的方式处理。



3.1.3 与拉普拉斯金字塔

根据 3.1.1 稀疏表示与去噪中所述，可以在 LP 重构时结合 Coring 完成降噪，流程如下：对 LP 中 bandpass 图像进行 coring，而 lowpass 图像保持即继续携带原有信息包括噪声；重构得到的图像能够减少噪音。示意图¹⁰如下：

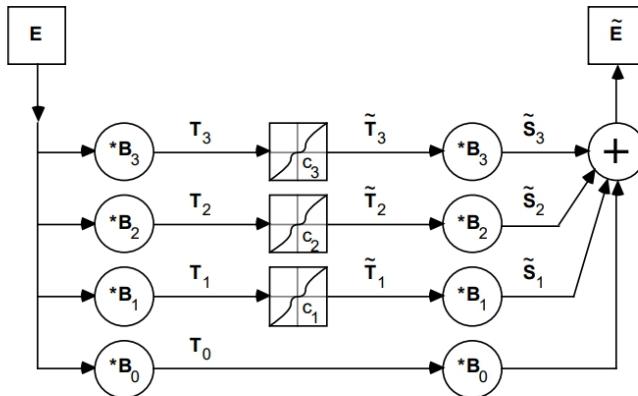


Figure 4: Flow diagram for coring an image. The original, E , is transformed to the images T_0, \dots, T_3 . The three bandpass images, $T_1 \dots T_3$, are then cored with the static non-linearities, c_1, \dots, c_3 . Reconstruction proceeds normally. The final cored image, \tilde{E} , should have less noise.

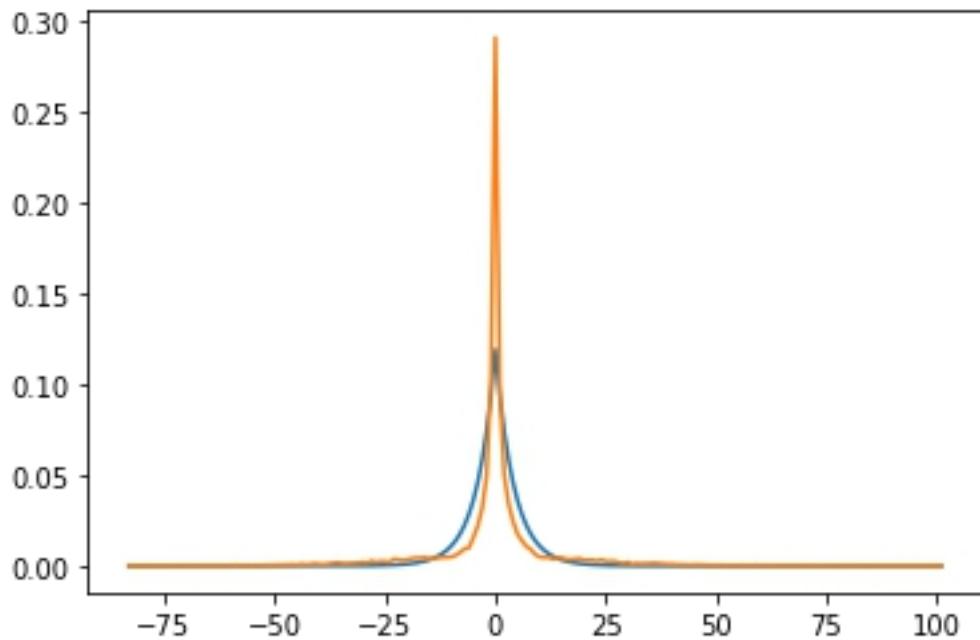
3.1.4 阈值

根据如上思路，为了抑制零值附近的噪音，采取设置阈值的方式，若转换系

¹⁰ Edward H. Adelson, Subband coring for image noise reduction. Internal Report, RCA David Sarnoff Research Center, Nov. 26 1986.

数小于阈值，则设为 0；否则则保持。关于阈值的设定有方法指出：自然图像的转换系数分布符合两参数的广义拉普拉斯分布（two-parameter generalized Laplacian distribution）： $P_x(x) = ce^{-|x/s|^p}$ ，该分布均值为 0，参数 s 和 p 与样本二阶矩（方差）和四阶矩（峰度）相关。给出直方图的方差和峰度信息即可求出 s 、 p 两个参数的值。其中 s 参数为 `scale` 参数，控制方差。论文提出的方法将使用贝叶斯估计法对干净的图像进行计算。

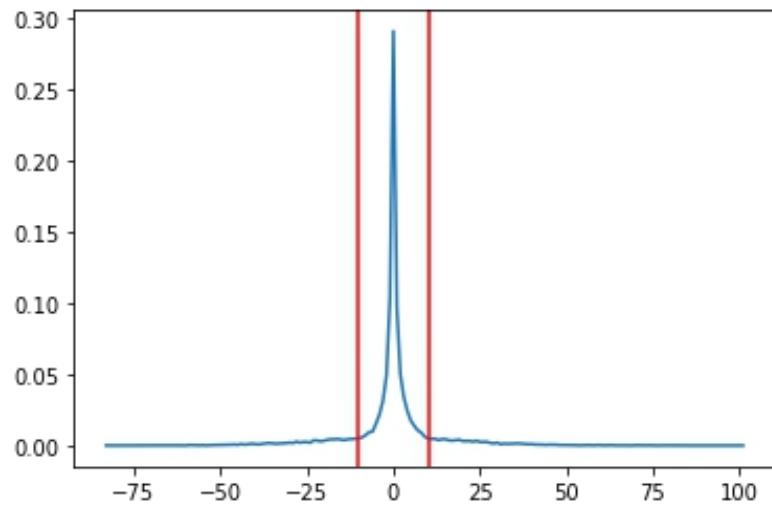
本次实验首先尝试了如上操作：首先对 LP 的变换系数进行拉普拉斯分布拟合（`laplace.fit`），计算得到 `loc` 和 `scale` 参数，其中 `loc` 参数控制均值的位移。但在使用拟合好的 `loc` 和 `scale` 参数逆推计算拉普拉斯分布时，发现计算得到的零点峰值存在较大差异，如图所示：其中橙色为原本的 `pdf`；而蓝色为计算拟合得到的 `pdf`。



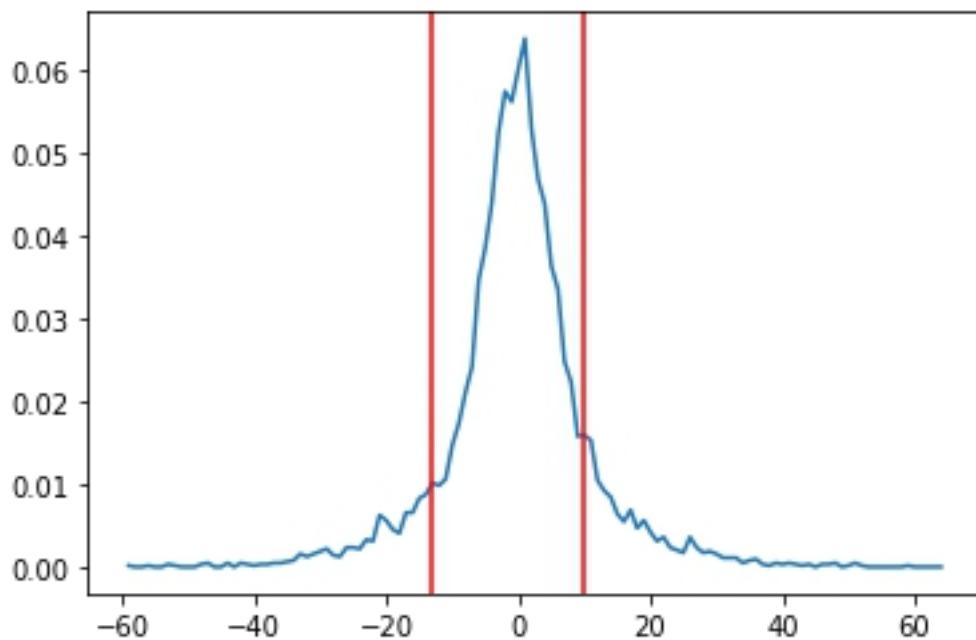
在经过多次尝试但拟合收效甚微后，考虑到使用贝叶斯估计与阈值设置的关联性较弱，因此本文采取了一种新的方法确定阈值——计算半高全宽（FWHM）。半高全宽为峰高一半的峰宽度，即通过峰高的中点作平行于峰底的直线，此直线与峰两侧相交两点之间的距离。相比起使用随机的阈值尝试，更倾向于使用半高全宽这一信号常用概念作为阈值。

虽然拉普拉斯分布拟合结果在峰值处拟合不准，但对于方差的估计即信号“宽度”较为准确；因此在拟合后使用 `laplace.stats` 计算在 `loc` 和 `scale` 给定下的

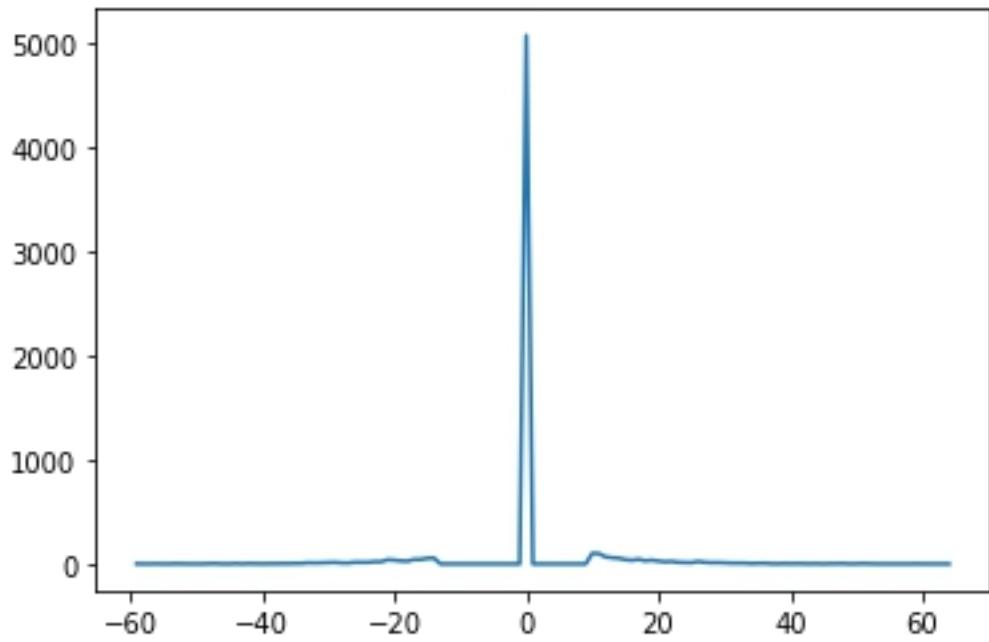
方差 var ，并根据方差计算得到标准差 std 。根据高斯分布的半高全宽公式有 $\text{FWHM} = 2\sqrt{2\ln 2}\text{std}$ ，得到 FWHM 。由于信号均值在 0 点，因此窗口宽度设置为 $\text{FWHM}/2$ 。如下图则为根据半高全宽计算得到的加噪声前的窗口宽度，两条红线代表窗口。



由于图像加噪声前后的 LP 参数分布的半高全宽不同，因此并不能使用加噪声前图像上确切的值作为阈值，而是计算窗口内的值的累积概率，在加噪声后的 LP 分布上找到相应累积概率数值大小对应的窗口，此即为需要设置的最终阈值。下图即为上图加噪声后对应的 LP 分布以及计算得到的左右两个阈值。

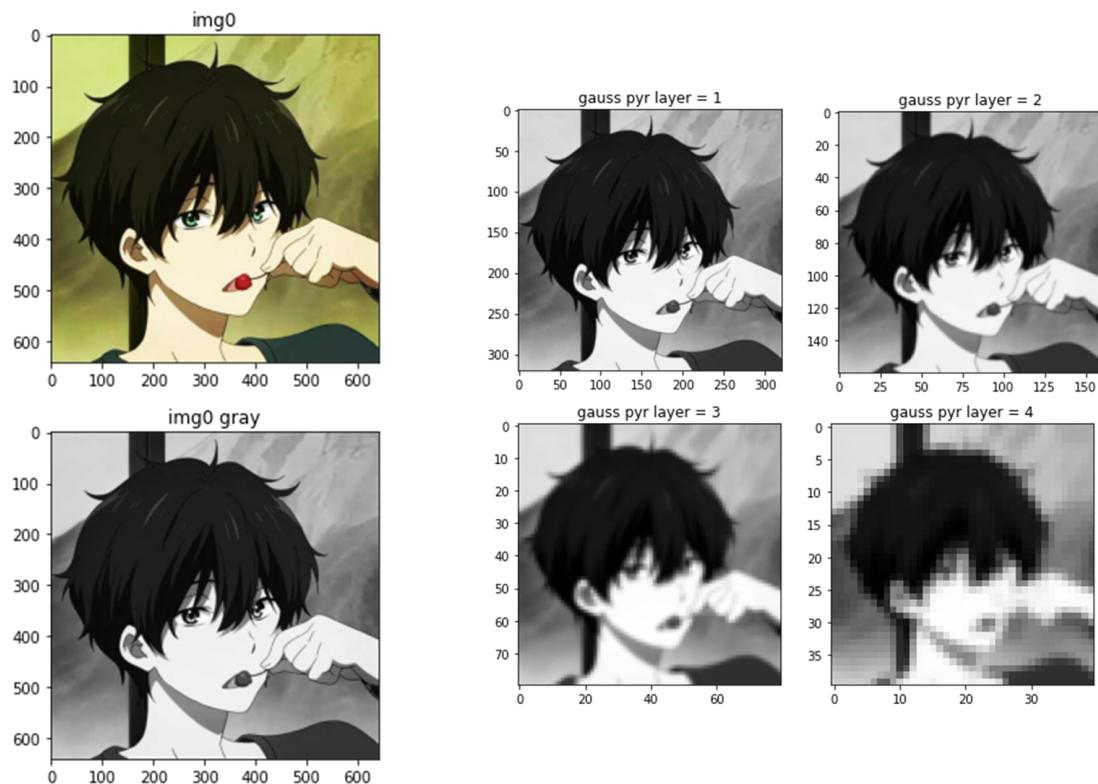


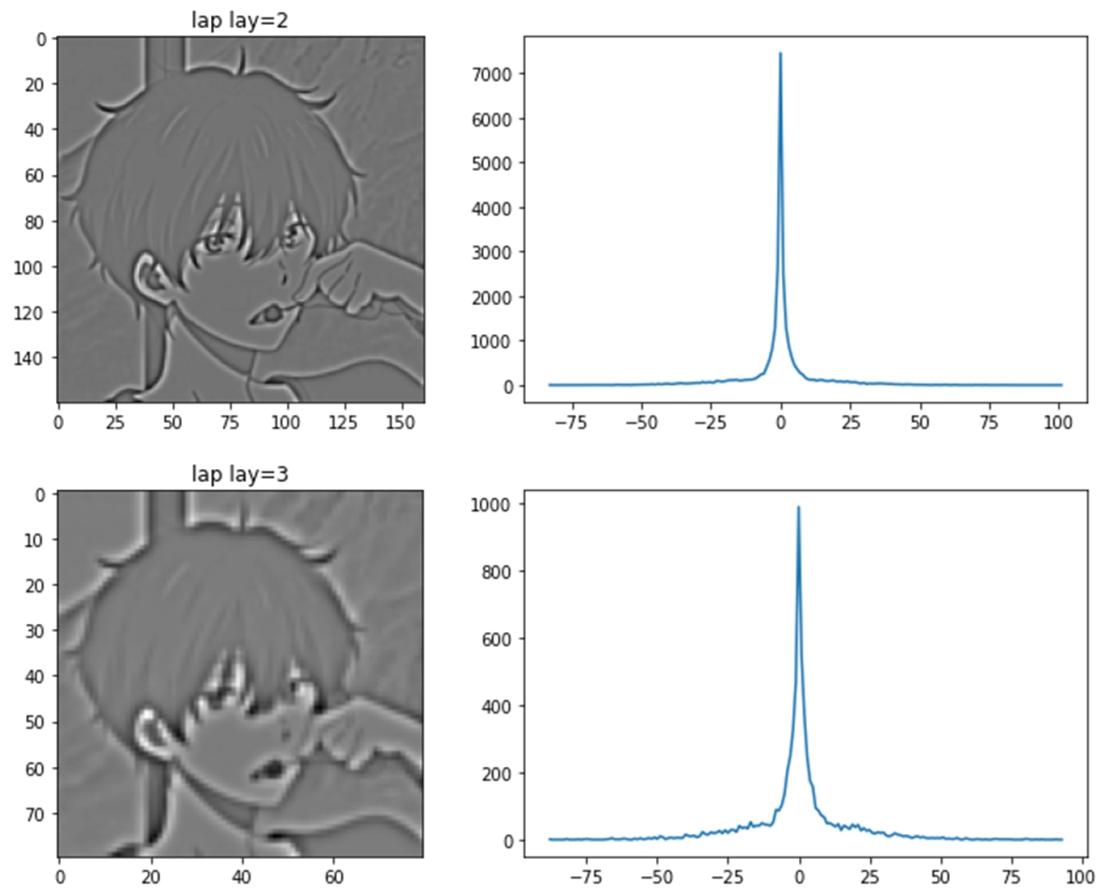
阈值内所有的值设为 0，阈值外的值保持不变，得到了如下的分布，可以看到在 20 附近有明显的断续，此即为上图所对应的阈值。



3.2 Implementation

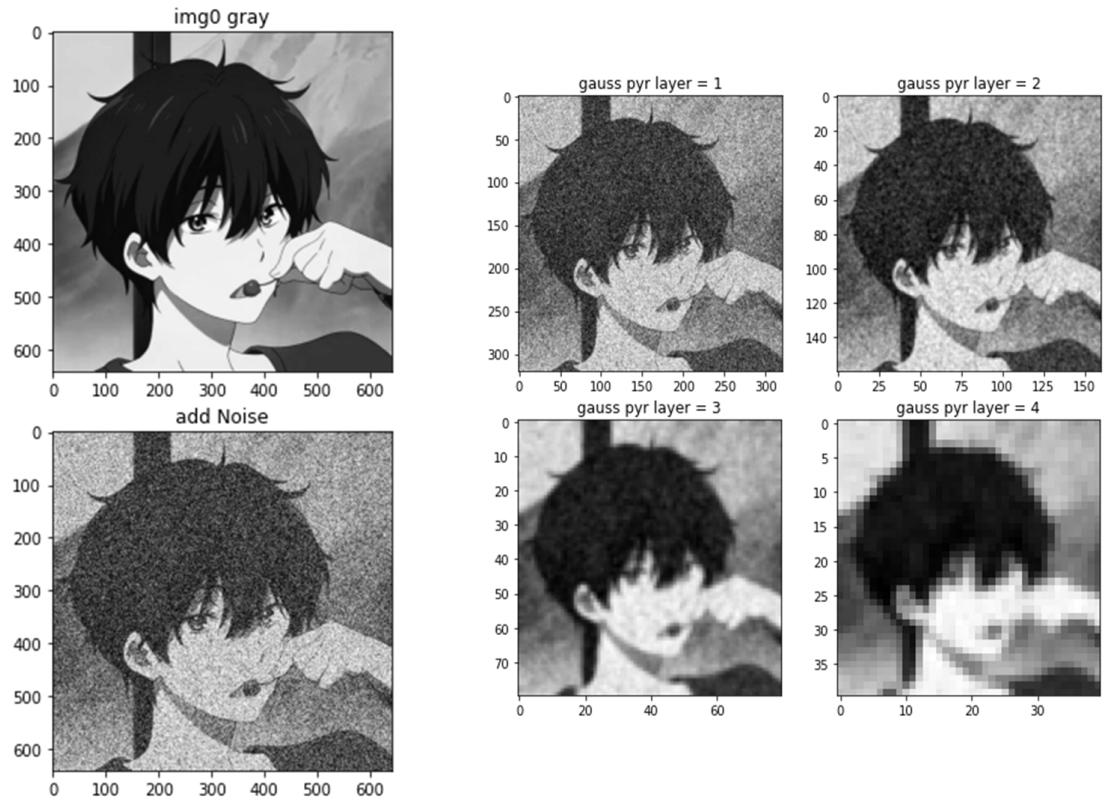
3.2.1 读入图像，构建高斯金字塔、拉普拉斯金字塔并计算直方图

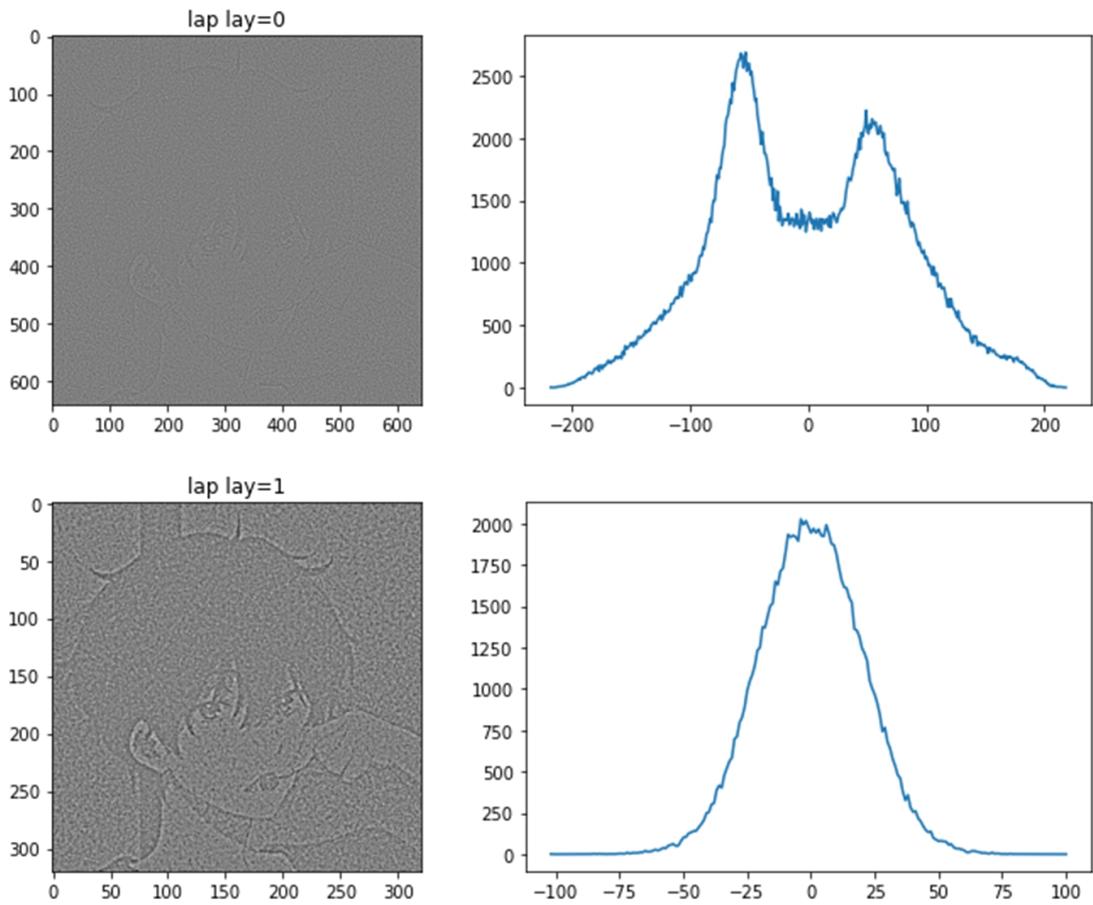




由于上文已涉及，因此只选取其中几层的图像进行展示。

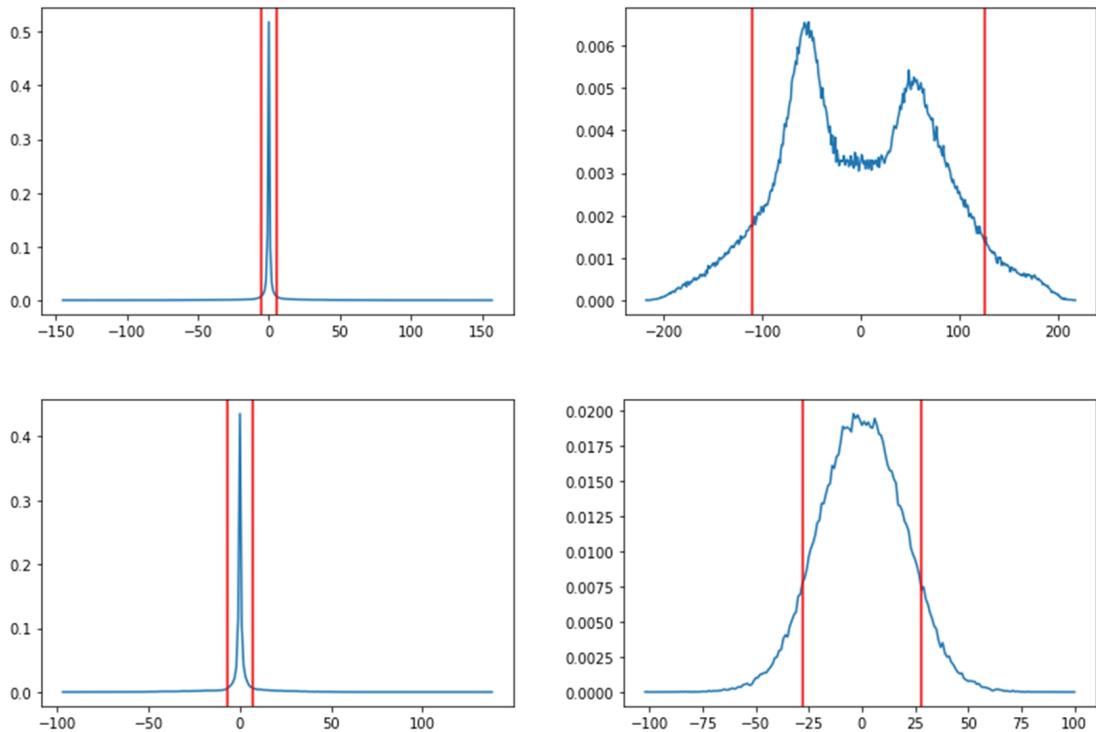
3.2.2 为图像加入高斯噪声，重复 3.2.1 的操作





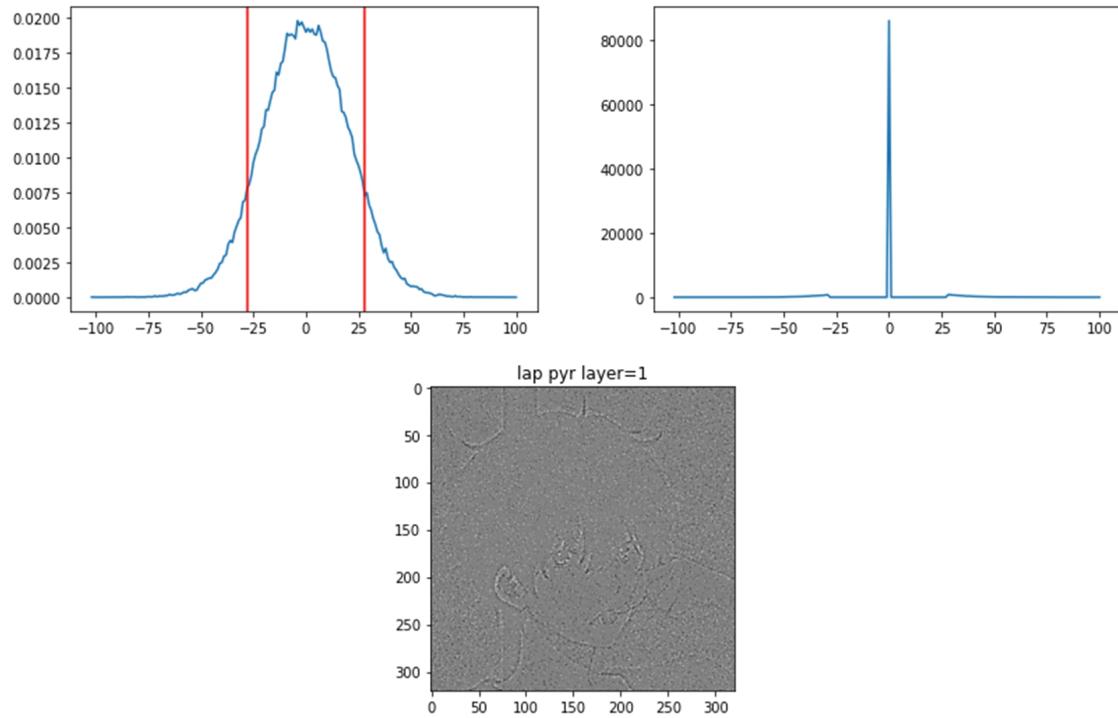
本次实验所加的高斯噪声均值为 0，标准差为 0.5，可以看到图像有很明显的噪声。

3.2.3 计算加噪声前图像 LP 分布的半高全宽以及相对应加噪声后的阈值

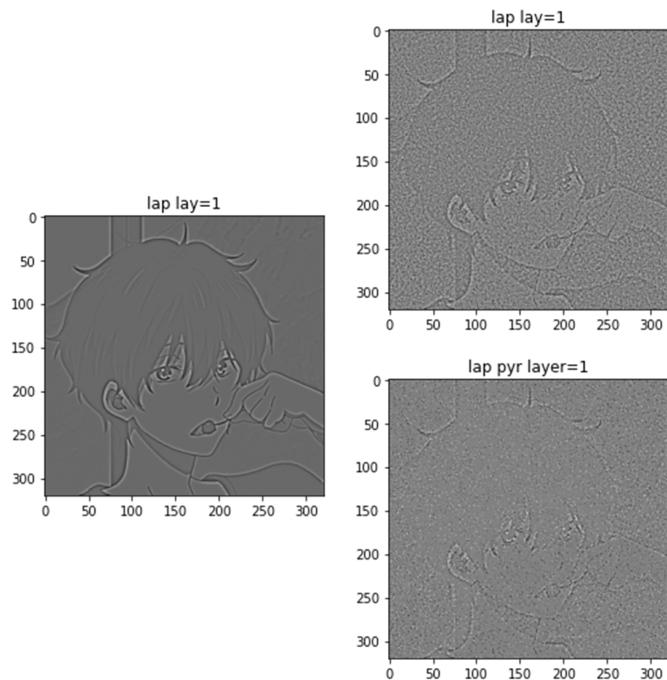


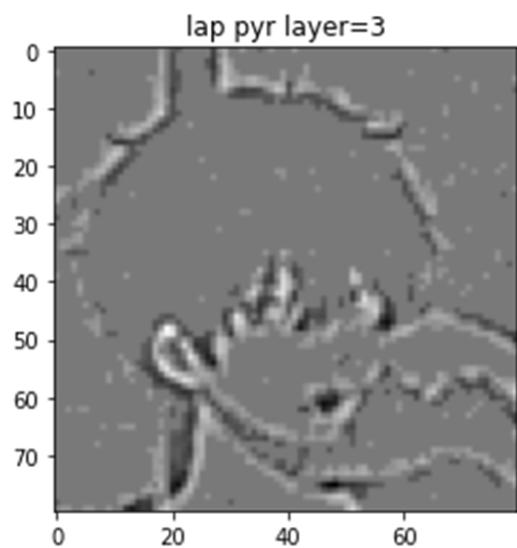
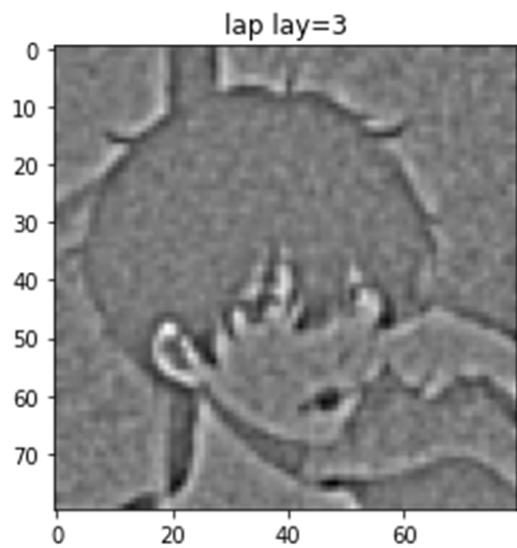
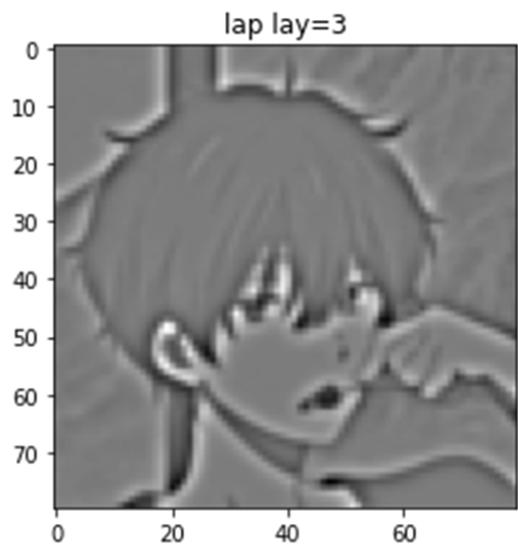
3.2.4 根据计算的阈值进行零值设定

下图为重新进行零值设定后的结果，其中左上图为对应的阈值，右上图为零值设定后的直方图，下方为零值设定后的 LP。



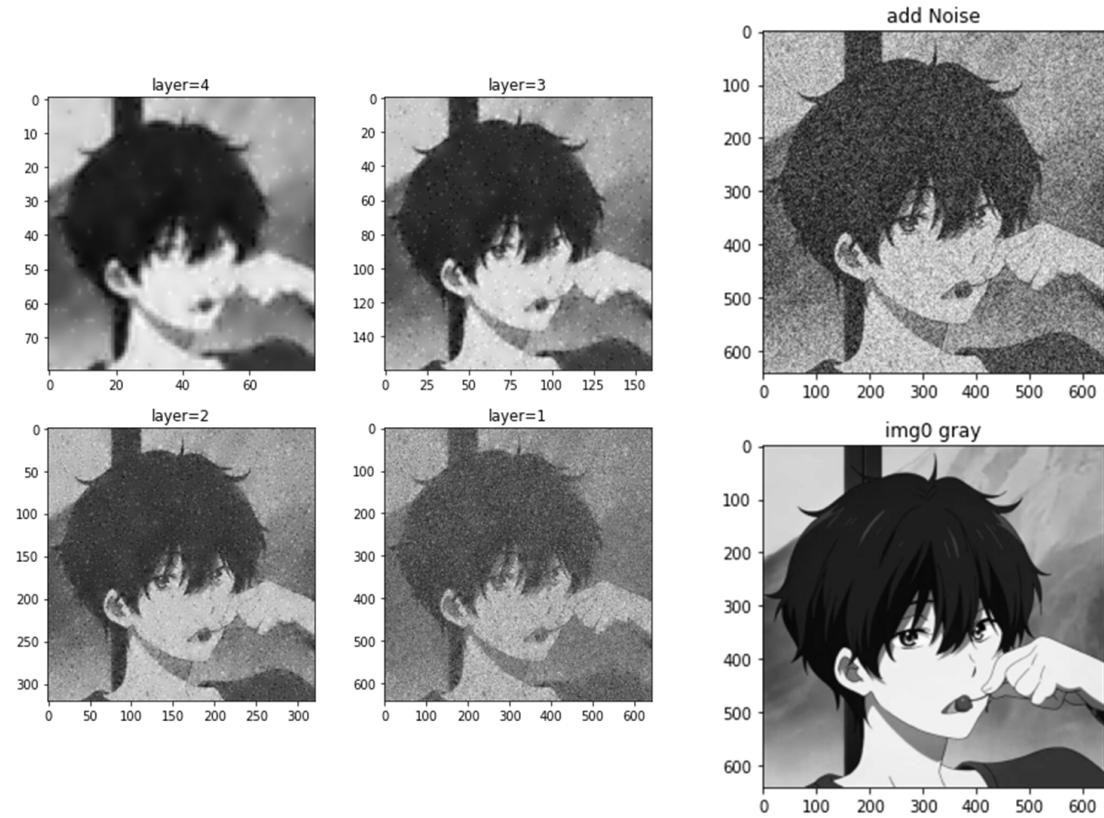
由下图可以看到，在 LP 层面上噪音已有部分减少，其中左图为未加噪声的 LP，右上为加了噪声的 LP，右下为经过处理的 LP。





可以看到，随着 LP 层数越来越往上（低频），噪音减少效果越明显。

3.2.5 重构



对图像进行重构，可以看到在最后的 layer=1 层面上，图像比起 add Noise 图像有噪音部分减少，重构的层数越少噪声去除的越多，但比起干净的原图仍有较大差距。

3.3 Discussion

对于重构金字塔后得到的图像实现效果不好，主要有两方面原因：

- 由于拉普拉斯分布拟合不准确且对于贝叶斯估计的知识掌握不熟练，导致最后采用了 FWHM 进行阈值的计算；而采用 FWHM 虽然比随机选取阈值合理，但仍然与最佳阈值点有一定差距，并没有能够完全的去除或者减少噪声。
- 在实现时注意到，LP 的 lowpass 部分并没有进行 coring 处理，而是保持了原有的信息，包括信号值与噪音，因此一定程度上保留了噪音。我们可以将 lowpass 作为第二阶段子带宽变换的输入，进行进一步分解后重新应用 coring，得到的图像与第一阶段的图像结合并重构，即可以对低频噪音进一步减少。实现的流程如下¹¹：

¹¹ Edward H. Adelson, Subband coring for image noise reduction. Internal Report, RCA David Sarnoff Research Center, Nov. 26 1986.

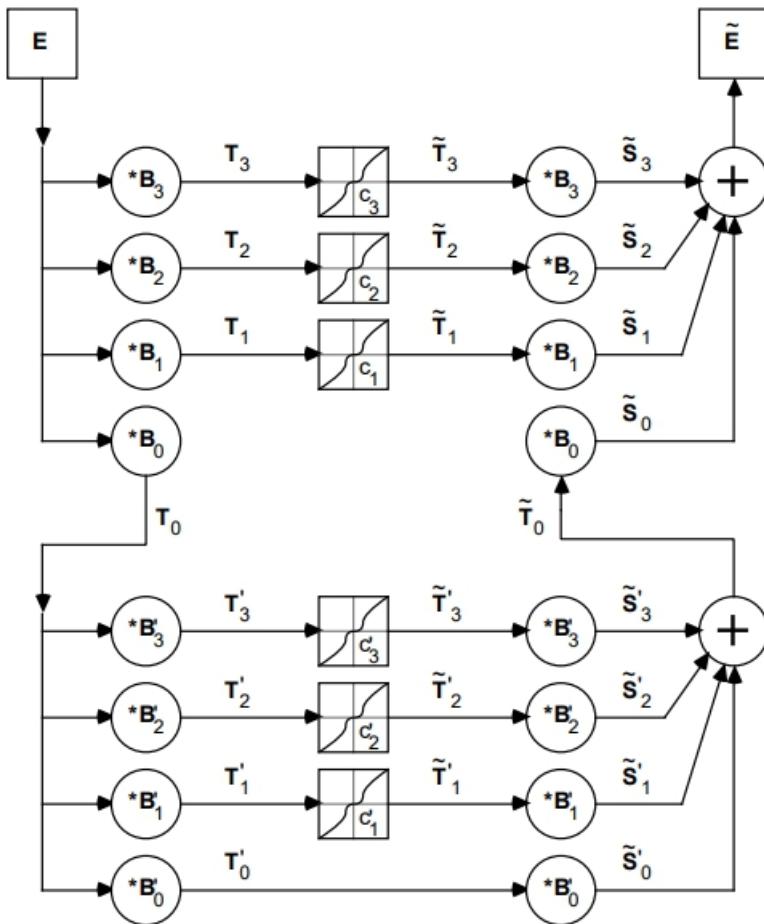


Figure 5: Two stage coring. The low-pass transform image T_0 undergoes a second stage of coring by a set of kernels B'_0, \dots, B'_3 . These kernels are tuned to lower spatial frequency than the first set, and so core out noise in a lower frequency band.

4 Conclusion

本次大作业加深了我对于 LP 的理解，原来很多图像处理任务都可以使用 LP 完成。创建一副全聚焦的图像看似很难，但实际上只需要使用掩膜加判断即可搞定，以后可以应用于自己拍摄的难以对焦的图像了:)。去噪操作有很多，高斯滤波、中值滤波都可应用，但 LP 由于掌握了高频的信息，自然也可以完成去噪的任务。遗憾的是对于拉普拉斯分布的知识掌握得并不熟练，虽然查阅了很多资料，但最后实现的结果并不是非常理想。对于数学、对于计算机视觉以及对于代码编写，我还有很多需要学习的地方。

本次实验中的代码可见于 <https://github.com/RuiNov1st/NKUComputerVision>.

5 References

- [1] Szeliski, R., Computer Vision: Algorithms and Applications. London: Springer, 2010.
- [2] Edward H. Adelson, Subband coring for image noise reduction. Internal Report, RCA David Sarnoff Research Center, Nov. 26 1986.
- [3] <https://www.cs.toronto.edu/~jepson/csc320/notes/pyramids.pdf>
- [4] E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," Proceedings of 3rd IEEE International Conference on Image Processing, 1996, pp. 379-382 vol.1, doi: 10.1109/ICIP.1996.559512.