Digital Image Processing – HW1

Submitters:
Amir Avivi – ********
Sagi Bouzaglo – ********

1. Theoretical Questions

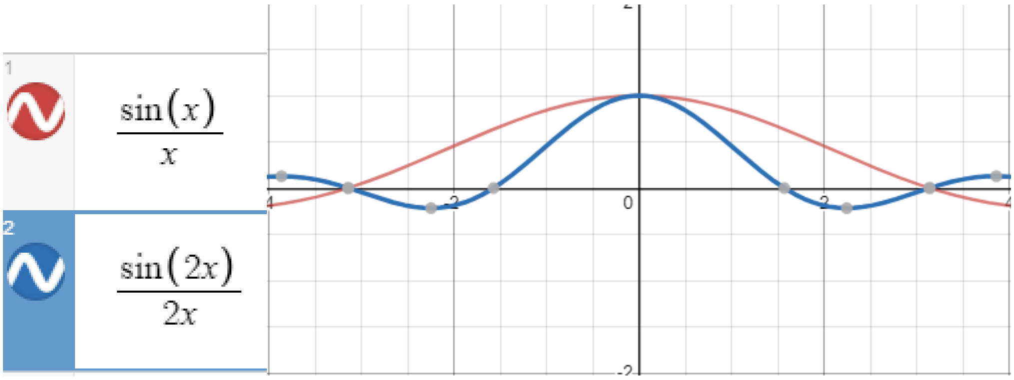
1. Let $f(x)$ be the image that would be formed on the image plane if the camera was still, and $g_k[n] = (f * h_k)(n)$ be the actual formed digital image of the k-th frame.
   In the process of taking the actual image we imply 2 actions – Sampling(digitalizing) and translation(due to the handshakes).
   Mathematically, the convolution kernel describing the above actions would be:
   $h_k = \int_t^{t+1msec} \tau_{o(t)}\, \delta dt$
2. The Fourier transform of h_k is given by:
   $$\mathcal{H}_k[\omega] = \mathcal{F}(h_k)_\omega = \mathcal{F}\left(\int_k^{k+1} \tau_{o(t)}\, \delta\right) = \{\text{Linearty}\} \int_k^{k+1} \phi_{o(t)}\, dt$$
3. Let $G_k[\omega]$ be the spectral decomposition of the real-world image.
   f – Real-world image (calculated via the pixel's cumulative sunlight absorption).
   h_k – Photographer's handshakes.
   $box(t)$ – Camera shutter exposure time describing function.

   Mathematically:

   $$G_k[\omega] = \mathcal{F}(g_k[n] * box(n)) = \mathcal{F}(f(n) * h_k(n) * box(t))_n = \mathcal{F}[\omega] \cdot H_k[\omega] \cdot \text{sinc}(\omega)$$

   The expression for the frequency response of the discrete kernel P_k is $H_k[\omega]$.
4. $|P_k[\omega]| = \left|\int_k^{k+1} \phi_{o(t)}\, dt\right| \le \{\text{triangle inequalty}\} \int_k^{k+1} |\phi_{o(t)}| dt = \int_k^{k+1} \left|e^{-2\pi i w^T o(t)}\right| dt \le \{\text{Euler's identity}\} \le \int_k^{k+1} 1 dt = 1$
5. We'll express $P_k[\omega]$ with respect to the given trajectory:
   $$P_k = \int_k^{k+1} e^{-2\pi i w^T(\tau-k-0.5)ve_1} d\tau \xrightarrow[\text{define } c=-2\pi i w^T ve_1]{} \int_k^{k+1} e^{c(\tau-k-0.5)} d\tau =$$
   $$e^{c(-k-0.5)} \int_k^{k+1} e^{c\tau} d\tau = e^{c(-k-0.5)} \left(\frac{e^{c(k+1)}}{c} - \frac{e^{ck}}{c}\right) =$$
   $$\frac{e^{0.5c} - e^{-0.5c}}{c}$$
   Inserting c:
   $$\frac{e^{-\pi i w^T ve_1} - e^{\pi i w^T ve_1}}{-2\pi i w^T ve_1} = \frac{1}{\pi w^T ve_1} \cdot \frac{e^{-\pi i w^T ve_1} + e^{\pi i w^T vke_1}}{2i} = \frac{\sin(\pi w^T ve_1)}{\pi w^T ve_1} =$$
   $$\mathbf{sinc}(\boldsymbol{v}\mathbf{w}^T\boldsymbol{e}_1)$$

This concludes an expected outcome:
The higher the velocity –> sinc is more narrow This lead to a more concentrated spectrum -> Blurrier images(Which is the main idea of this task 😊)



6. We'll generalize the previous result for $o(\tau) = q + \tau v$:

$$P_k = \int_k^{k+1} \mathrm{e}^{-2\pi i w^T (q+\tau v)} d\tau = \int_k^{k+1} \mathrm{e}^{-2\pi i w^T ((\tau-k-0.5)v + (k+0.5)v+q)} d\tau =$$

$\{$decomposing for integral and consts + using previous result$\}$

$$= sinc(w^T v) \cdot \mathrm{e}^{-2\pi i w^T ((k+0.5)v+q)}$$

$\{$Translating v for its e1 and e2 general cordinates
– based on the assupmtions that the handshakes are parller to the image plane$\}$

We can see that now the sinc's directions are based on the velocity directions + higher velocity will result a more concentrated spectrum, just like before.
The added function, results the same effect of concentrating the spectrum the higher the velocity gets.

7. We invested a lot of time trying to solve this problem.
In our answers we will refer to the following papers:
a.

2762 IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 27, NO. 6, JUNE 2018

Reconstruction of Single Image from Multiple Blurry Measured Images

Tsung-Ching Lin, *Senior Member, IEEE*, Liming Hou, Hongqing Liu, *Senior Member, IEEE*, Yong Li, *Member, IEEE*, and Trieu-Kien Truong, *Life Fellow, IEEE* b.

BlurBurst: Removing Blur Due to Camera Shake using Multiple Images

ATSUSHI ITO
Sony Corporation
and
ASWIN C. SANKARANARAYANAN
Carnegie Mellon University
and
ASHOK VEERARAGHAVAN and RICHARD G. BARANIUK
Rice University

c.

Burst Deblurring: Removing Camera Shake Through Fourier Burst Accumulation

Mauricio Delbracio   Guilermo Sapiro
ECE, Duke University
{mauricio.delbracio, guillermo.sapiro}duke.edu

We will refer to them as **a, b** and **c** from now on.
Our code runs based on paper **c** results which we will explain in a more detailed manner later.
Papers **a and b** modelized the problem pretty much the same, as an optimization problem with the following structure:

Paper a:

$$\textbf{Objective function}: \text{argmin} \quad \sum_l w_l \|\mathbf{Y}_l - (\mathbf{D}_{\mathbf{H}_l} \circ \boldsymbol{\alpha}_{\mathbf{H}_l})(\mathbf{D}_x \circ \boldsymbol{\alpha}_x)\|_2^2 + \beta \|\boldsymbol{\alpha}_x\|_1 + \lambda_l \|\boldsymbol{\alpha}_{\mathbf{H}_l}\|_1$$

Paper b:

$$\arg\min_{\mathbf{x},\mathbf{k}_i} \sum_i \|A(\mathbf{y}_i - \mathbf{k}_i * \mathbf{x})\|_2^2 + \lambda_1 \mathrm{TV}(\mathbf{x}) + \lambda_2 \sum_i \|\mathbf{k}_i\|_1 \quad (2)$$

They both optimized the objective function in a 2 step iterative approach:
1^{st}: Estimating the latent Image given the current Kernel estimators
2^{nd}: Estimating the latent kernels that were used to generate each image by using the current Image estimator

Both papers used a weighted sum (given as w_l in paper **a** and A in paper **b**) and regularization on both the kernel estimators and the image estimator which we will not explain due to the fact that the code we supplied is based on paper **c**.

We tried implementing both method, but via our lack of understanding or the paper's insufficient explanations, both implementations faced a very long(about a day) computation overhead – BUT yielded very nice results, so after investing a lot of time in implementing them we looked for a different approach.

<u>Explanation of paper c</u>:

The paper suggests that each image in the burst will be differently attenuated due to the dissimilarity in the photographer hand movements.

*We saw in section 4 that the blur kernel does not amplify noise, and a proof is also given in the paper is section 2.

The idea is to reconstruct an image whose Fourier spectrum takes for each frequency the value having the largest Fourier magnitude in the burst. Using the above claim(*) it will result the reconstruction having what is less attenuated from each image of the burst.

The weight for a given frequency in the reconstructed image is given by:

$$u_p(\mathbf{x}) = \mathcal{F}^{-1}\left(\sum_{i=1}^{M} w_i(\zeta)\cdot \hat{v}_i(\zeta)\right)(\mathbf{x}),$$

$$w_i(\zeta) = \frac{|\hat{v}_i(\zeta)|^p}{\sum_{j=1}^{M}|\hat{v}_j(\zeta)|^p},$$

- $u_p(x)$ *is the reconstructed image*
- $\hat{v}_i$ is the Fourier transform of the individual burst image v_i.
- w_i is the weight for each frequency ς in the picture $\hat{v}_i$.
  It controls the contribution of the frequency in image Vi to the reconstruction.
- P is an argument for the solution, the higher P is, then we will lower the effect of the lowest value frequencies.
  Note that for P=0 the restored image is just the average of the burst, while if $p \to \infty$ then we get Maximum pooling(take only the image with the maximum value for this frequency)

  <u>Dealing with noise:</u>
  The image in the burst are blurry but they are also contaminated with noise.
  The picture v_i can be smoothed out before computing the weight above to deal with noise, using a low pass Gaussian filter of standard deviation σ (given as a parameter)
  For better sharpening we use a Gaussian sharpening after we computed the weights for the frequencies of the burst image.