Image Processing and Computer Vision (IPCV)

Prof. J. Weickert
Mathematical Image Analysis Group

Summer term 2019
Saarland University

Example Solutions for Homework Assignment 3 (H3)

Problem 1 (Discrete Fourier Transform)

We show that it holds for all $p, q \in \{0, ..., M-1\}$:

$$\langle \boldsymbol{b}_p, \boldsymbol{b}_q \rangle = \begin{cases} 1, & \text{if } p = q \\ 0, & \text{else} \end{cases}$$

First we consider $p, q \in \{0, ..., M-1\}, q = p$. We have

$$\begin{aligned}
\langle \boldsymbol{b}_p, \boldsymbol{b}_q \rangle &= \sum_{m=0}^{M-1} \frac{1}{\sqrt{M}} \exp\left(\frac{2\pi i p m}{M}\right) \overline{\frac{1}{\sqrt{M}} \exp\left(\frac{2\pi i q m}{M}\right)} \\
&= \frac{1}{M} \sum_{m=0}^{M-1} \exp\left(\frac{2\pi i p m}{M}\right) \exp\left(-\frac{2\pi i q m}{M}\right) \\
&= \frac{1}{M} \sum_{m=0}^{M-1} \exp\left(\frac{2\pi i \overbrace{(p-q)}^{0} m}{M}\right) \\
&= \frac{1}{M} \sum_{m=0}^{M-1} \underbrace{\exp(0)}_{1} \\
&= \frac{M}{M} = 1
\end{aligned}$$

Now we consider $p, q \in \{0, ..., M-1\}, q \neq p$. It holds that

$$\begin{aligned}
\langle \boldsymbol{b}_p, \boldsymbol{b}_q \rangle &= \frac{1}{M} \sum_{m=0}^{M-1} \exp\left(\frac{2\pi i(p-q)m}{M}\right) \\
&= \frac{1}{M} \sum_{m=0}^{M-1} \left(\exp\left(\frac{2\pi i(p-q)}{M}\right)\right)^m \\
&\overset{(1)}{=} \frac{1}{M} \frac{1-\left(\exp\left(\frac{2\pi i(p-q)}{M}\right)\right)^M}{1-\exp\left(\frac{2\pi i(p-q)}{M}\right)} \\
&= \frac{1}{M} \frac{1-\exp\left(\frac{2\pi i(p-q)M}{M}\right)}{1-\exp\left(\frac{2\pi i(p-q)}{M}\right)} \\
&\overset{(2)}{=} \frac{1}{M} \frac{1-\exp\left(2\pi i(p-q)\right)}{1-\exp\left(\frac{2\pi i(p-q)}{M}\right)} \\
&= \frac{1}{M} \frac{1-1}{1-\exp\left(\frac{2\pi i(p-q)}{M}\right)} \\
&= 0
\end{aligned}$$

with

(1) For $r \neq 1$, the sum of the first n terms of a geometric series is

$$\sum_{k=0}^{n-1} r^k = \frac{1-r^n}{1-r}$$

We have

$$0 \leq p, q < M \Rightarrow \frac{p-q}{M} \notin \mathbb{Z} \Rightarrow \exp\left(\frac{2\pi i\,(p-q)}{M}\right) \neq 1.$$

(2) $\quad (p-q) \in \mathbb{Z} \Rightarrow \exp\left(2\pi i\,(p-q)\right) = 1$

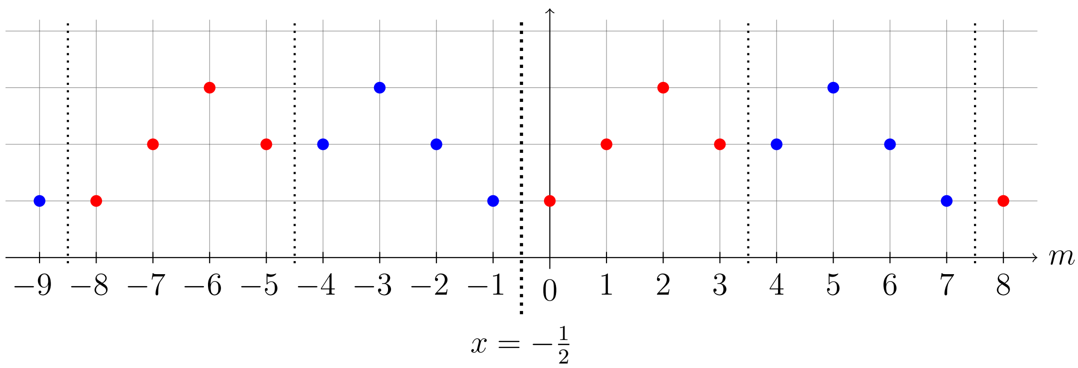
As the set $\{\boldsymbol{b}_0, ..., \boldsymbol{b}_{M-1}\}$ has cardinality M, it follows that it forms an orthonormal basis of the M-dimensional vector space $\mathbb{C}^M$ with respect to $\langle \cdot, \cdot \rangle$.

Problem 2: (Relation between DFT and DCT)

We have the discrete signal $\boldsymbol{g}$ defined as:

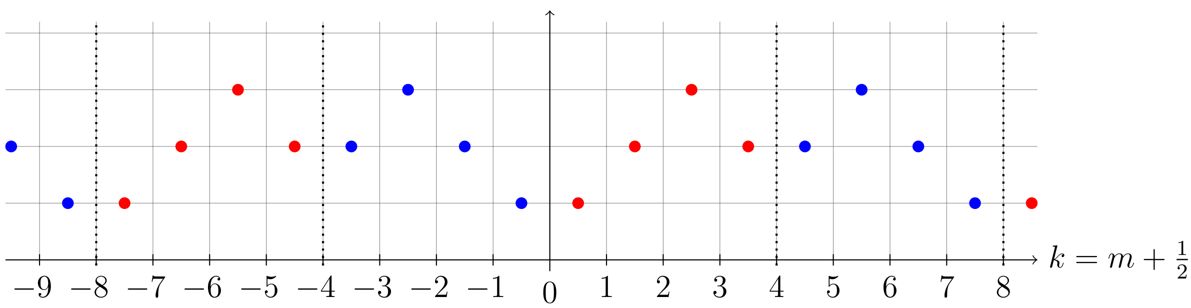
$$g_m := \begin{cases} f_m, & \text{if } 0 \leq m \leq M-1 \\ f_{2M-m-1}, & \text{if } M \leq m \leq 2M-1 \end{cases} \tag{1}$$

This signal is identical to $\boldsymbol{f}$ for $0 \leq i \leq M-1$ and identical to a mirrored version of $\boldsymbol{f}$ for $M \leq i \leq 2M-1$. We now assume this signal to be extended periodically with period $2M$ over $\mathbb{Z}$. Below you can see a sample sketch of the function $\boldsymbol{g}$ with $M = 4$.



As we can see, the whole signal $\boldsymbol{g}$ is symmetric with respect to the point $x = -\frac{1}{2}$. Also note that $g_m = g_{-m-1} = g_{2M-m-1}$.

In order to get a symmetry with respect to $x' = 0$, we have to shift the signal by $\frac{1}{2}$ to the right, which means that we define a new, shifted signal $\boldsymbol{h}$ and a (non-integer) index $k \in \{\frac{2l+1}{2} | l \in \mathbb{Z}\}$ such that $h_k = g_{k-\frac{1}{2}}$. The result of this index-shift can be seen in the figure below.



We can now compute the discrete Fourier transform (DFT) of $\boldsymbol{g}$ at a point $p \in \{0, .., 2M-1\}$ to obtain a slightly modified discrete cosine transform (DCT) of the original signal $\boldsymbol{f}$. However, we have only defined the DFT for the grid points at integer points while the samples of the shifted signal are defined on a intermediate positions $k = m + \frac{1}{2}$. Therefore we use the shift theorem to express $\boldsymbol{h}$ in terms of the unshifted signal $\boldsymbol{g}$ which lives on the original discrete grid $m = k - \frac{1}{2}$.

$$
\begin{aligned}
\mathrm{DFT}\left[h_k\right]_p &= \mathrm{DFT}\left[g_{k-\frac{1}{2}}\right]_p \\
&\overset{(1)}{=} \exp\left(\frac{-i2\pi\frac{1}{2}p}{2M}\right)\mathrm{DFT}\left[g_m\right]_p \\
&= \exp\left(\frac{-i\pi p}{2M}\right)\mathrm{DFT}\left[g_m\right]_p \\
&\overset{(2)}{=} \exp\left(\frac{-i\pi p}{2M}\right)\frac{1}{\sqrt{2M}}\sum_{m=0}^{2M-1} g_m \exp\left(\frac{-i2\pi pm}{2M}\right) \\
&\overset{(3)}{=} \exp\left(\frac{-i\pi p}{2M}\right)\frac{1}{\sqrt{2M}}\left(\sum_{m=0}^{M-1} f_m \exp\left(\frac{-i2\pi pm}{2M}\right) + \sum_{m=M}^{2M-1} f_{2M-1-m}\exp\left(\frac{-i2\pi pm}{2M}\right)\right) \\
&\overset{(4)}{=} \exp\left(\frac{-i\pi p}{2M}\right)\frac{1}{\sqrt{2M}}\left(\sum_{m=0}^{M-1} f_m \exp\left(\frac{-i2\pi pm}{2M}\right)\right. \\
&\qquad\qquad\left. + \sum_{n=M-1}^{0} f_n \exp\left(\frac{-i2\pi p(2M-1-n)}{2M}\right)\right) \\
&= \exp\left(\frac{-i\pi p}{2M}\right)\frac{1}{\sqrt{2M}}\sum_{m=0}^{M-1} f_m \underbrace{\left(\exp\left(\frac{-i2\pi pm}{2M}\right) + \exp\left(\frac{-i2\pi p(2M-1-m)}{2M}\right)\right)}_{(*)}
\end{aligned}
$$

Here, the following properties have been used:

(1) Shift Theorem, rename index to m (completely optional)

(2) Definition of the Discrete Fourier Transform

(3) The definition of g in terms of f

(4) Substitution $n = 2M-1-m$ and the fact that the cosine function is even

For $(*)$ we can calculate:

$$\begin{aligned}
& \exp\left(\frac{-i2\pi pm}{2M}\right) + \exp\left(\frac{-i2\pi p(2M-1-m)}{2M}\right) \\
= \quad & \exp\left(\frac{-i2\pi pm}{2M}\right) + \underbrace{\exp\left(\frac{-i2\pi p2M}{2M}\right)}_{=1} \exp\left(\frac{i2\pi p}{2M}\right) \exp\left(\frac{i2\pi pm}{2M}\right) \\
= \quad & \exp\left(\frac{i\pi p}{2M}\right)\left(\exp\left(\frac{-i\pi p}{2M}\right)\exp\left(\frac{-i\pi p(2m)}{2M}\right) + \exp\left(\frac{i\pi p}{2M}\right)\exp\left(\frac{i\pi p(2m)}{2M}\right)\right) \\
= \quad & \exp\left(\frac{i\pi p}{2M}\right)\left(\exp\left(\frac{-i\pi p(2m+1)}{2M}\right) + \exp\left(\frac{i\pi p(2m+1)}{2M}\right)\right) \\
= \quad & \exp\left(\frac{i\pi p}{2M}\right) 2\cos\left(\frac{\pi p(2m+1)}{2M}\right)
\end{aligned}$$

Plugging this expression in the above equation yields

$$\begin{aligned}
& \exp\left(\frac{-i\pi p}{2M}\right)\frac{1}{\sqrt{2M}}\sum_{m=0}^{M-1} f_m \exp\left(\frac{i\pi p}{2M}\right) 2\cos\left(\frac{\pi p(2m+1)}{2M}\right) \\
= \quad & 2\frac{1}{\sqrt{2M}}\sum_{m=0}^{M-1} f_m \cos\left(\frac{\pi p(2m+1)}{2M}\right) \\
= \quad & \sqrt{\frac{2}{M}}\sum_{m=0}^{M-1} f_m \cos\left(\frac{\pi p(2m+1)}{2M}\right)
\end{aligned}$$

As we can see, we are able express our signal in terms of the basis vectors $\boldsymbol{v}_p$ $(p = 0, .., M-1)$ with

$$\begin{aligned}
v_p &= \sqrt{\frac{2}{M}}\left(\cos\left(\frac{\pi p(2m+1)}{2M}\right)\right)^\top_{m=0,..,M-1} \\
&= \sqrt{\frac{2}{M}}\left(\cos\left(\frac{\pi p}{2M}\right), \cos\left(\frac{\pi p\cdot 3}{2M}\right), \cos\left(\frac{\pi p\cdot 5}{2M}\right), \ldots, \cos\left(\frac{\pi p(2M-1)}{2M}\right)\right)^\top
\end{aligned}$$

A simple computation shows that these vectors are orthogonal to each other. Regarding the norm, one can see:

$$\|v_p\| = \sqrt{\sum_{m=0}^{M-1}\frac{2}{M}\cos^2\left(\frac{\pi p(2m+1)}{2M}\right)} = \begin{cases}\sqrt{2}, & \text{if } p = 0 \\ 1, & \text{if } p = 1, .., M-1\end{cases} \tag{2}$$

To make the transformation orthonormal we can further use the coefficients

$$c_p := \begin{cases} \sqrt{\frac{1}{M}}, & \text{if } p = 0 \\ \sqrt{\frac{2}{M}}, & \text{if } p = 1, .., M-1 \end{cases}$$

It follows that the DCT can be derived via the Discrete Fourier Transform of a shifted and mirrored signal.

Problem 3 (Image Pyramids)

(a) The *Gaussian pyramid* $\left\{\boldsymbol{v}^N, ..., \boldsymbol{v}^0\right\}$ of a signal $\boldsymbol{u} = (u_0, ..., u_{2^N-1})^\top$ is defined as

$$\begin{aligned} \boldsymbol{v}^N &:= \boldsymbol{u}, \\ \boldsymbol{v}^{k-1} &:= R_k^{k-1}\boldsymbol{v}^k \quad (k = N, \ldots, 1), \end{aligned}$$

with

$$R_k^{k-1} \quad := \quad \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \ldots & 0 & 0 & 0 & 0 \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ 0 & 0 & 0 & 0 & \ldots & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad (k \geq 1)\ .$$

Applied to our signal $\boldsymbol{u} := (2, 6, 9, 9, 4, 8, 11, 1)^\top$ we get

$$\begin{aligned} \boldsymbol{v}^3 &= \boldsymbol{u} = (2, 6, 9, 9, 4, 8, 11, 1)^\top \\ \boldsymbol{v}^2 &= R_3^2\boldsymbol{v}^3 \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 6 \\ 9 \\ 9 \\ 4 \\ 8 \\ 11 \\ 1 \end{pmatrix} \\ &= (4, 9, 6, 6)^\top \end{aligned}$$

$$
\begin{aligned}
\boldsymbol{v}^1 &= R_2^1 \boldsymbol{v}^2 \\
&= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 \\ 9 \\ 6 \\ 6 \end{pmatrix} \\
&= (6.5, 6)^\top \\
\boldsymbol{v}^0 &= R_1^0 \boldsymbol{v}^1 \\
&= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 6.5 \\ 6 \end{pmatrix} \\
&= 6.25
\end{aligned}
$$

(b) The *Laplacian pyramid* $\left\{\boldsymbol{w}^N, \ldots, \boldsymbol{w}^0\right\}$ of a signal $\boldsymbol{u} = (u_0, \ldots, u_{2^N-1})^\top$ with Gaussian pyramid $\left\{\boldsymbol{v}^N, \ldots, \boldsymbol{v}^0\right\}$ is defined as

$$
\begin{aligned}
\boldsymbol{w}^k &:= \boldsymbol{v}^k - P_{k-1}^k \boldsymbol{v}^{k-1} \quad (k = N, \ldots, 1), \\
\boldsymbol{w}^0 &:= \boldsymbol{v}^0
\end{aligned}
$$

with

$$
P_{k-1}^k \quad := \quad \begin{pmatrix}
1 & 0 & \ldots & 0 & 0 \\
1 & 0 & \ldots & 0 & 0 \\
0 & 1 & \ldots & 0 & 0 \\
0 & 1 & \ldots & 0 & 0 \\
 & \ddots & \ddots & \ddots & \\
0 & 0 & \ldots & 1 & 0 \\
0 & 0 & \ldots & 1 & 0 \\
0 & 0 & \ldots & 0 & 1 \\
0 & 0 & \ldots & 0 & 1
\end{pmatrix} \qquad (k \geq 1) \; .
$$

Applied to our signal, we obtain

$$\begin{aligned}
\boldsymbol{w}^0 &= \boldsymbol{v}^0 = 6.25 \\
\boldsymbol{w}^1 &= \boldsymbol{v}^1 - P_0^1 \boldsymbol{v}^0 \\
&= \begin{pmatrix} 6.5 \\ 6 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} 6.25 = \begin{pmatrix} 0.25 \\ -0.25 \end{pmatrix} \\
\boldsymbol{w}^2 &= \boldsymbol{v}^2 - P_1^2 \boldsymbol{v}^1 \\
&= \begin{pmatrix} 4 \\ 9 \\ 6 \\ 6 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 6.5 \\ 6 \end{pmatrix} \\
&= (-2.5, 2.5, 0, 0)^\top \\
\boldsymbol{w}^3 &= \boldsymbol{v}^3 - P_2^3 \boldsymbol{v}^2 \\
&= \begin{pmatrix} 2 \\ 6 \\ 9 \\ 9 \\ 4 \\ 8 \\ 11 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 9 \\ 6 \\ 6 \end{pmatrix} \\
&= (-2, 2, 0, 0, -2, 2, 5, -5)^\top
\end{aligned}$$

(c) The Laplacian pyramid allows to *reconstruct* the original signal via the Gaussian pyramid:

$$\begin{aligned}
\boldsymbol{v}^0 &= \boldsymbol{w}^0 \\
\boldsymbol{v}^k &= \boldsymbol{w}^k + P_{k-1}^k \boldsymbol{v}^{k-1} \quad (k = 1, \ldots, N), \\
\boldsymbol{u} &= \boldsymbol{v}^N
\end{aligned}$$

Let us now reconstruct the original signal from the Laplacian pyramid:

$$
\begin{aligned}
\boldsymbol{v}^0 &= \boldsymbol{w}^0 = 6.25 \\
\boldsymbol{v}^1 &= \boldsymbol{w}^1 + P_0^1 \boldsymbol{v}^0 \\
&= \begin{pmatrix} 0.25 \\ -0.25 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} 6.25 = \begin{pmatrix} 6.5 \\ 6 \end{pmatrix} \\
\boldsymbol{v}^2 &= \boldsymbol{w}^2 + P_1^2 \boldsymbol{v}^1 \\
&= \begin{pmatrix} -2.5 \\ 2.5 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 6.5 \\ 6 \end{pmatrix} \\
&= (4, 9, 6, 6)^\top \\
\boldsymbol{u} = \boldsymbol{v}^3 &= \boldsymbol{w}^3 + P_2^3 \boldsymbol{v}^2 \\
&= \begin{pmatrix} -2 \\ 2 \\ 0 \\ 0 \\ -2 \\ 2 \\ 5 \\ -5 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 9 \\ 6 \\ 6 \end{pmatrix} \\
&= (2, 6, 9, 9, 4, 8, 11, 1)^\top
\end{aligned}
$$

(d) The resulting Laplacian pyramids contain redundant information. As one can see from the resulting vectors $\boldsymbol{w}^k$, the vectors contains doubled entries with $w_{2i+1}^k = -w_{2i}^k$. With this knowledge, the size of vector $\boldsymbol{w}^k$ can be reduced by factor $\frac{1}{2}$ by storing only all vector entries with even index. In our example, it suffices to store a vector $\boldsymbol{w} = (w_0^3, w_2^3, w_4^3, w_6^3, w_0^2, w_2^2, w_0^1, w_0^0)^\top = (-2, 0, -2, 5, -2.5, 0, 0.25, 6.25)^\top$.

Problem 4 (Filtering in the Fourier Domain)

(a) From C3, Problem 1 and the lecture we know that wave-like patterns generate 3-point spectra which are oriented in the same way as the wave pattern. Vertical line artefacts as the ones `smoke.pgm` are generated by the superposition of many such wave-patterns. Therefore, the three-point spectra of such wave patterns form a horizontal line of coefficients with a high contribution to the spectrum. Setting the coefficients on this line to zero successfully removes the line artifacts. The overall quality of the image `smoke.pgm` is affected only slighly, since the removal of these frequencies does not remove too many significant structures of the smoke (see Figure 1).

```
/* ------------------------------------------------------------ */
long centre_x = nx/2 + 1;
long centre_y = ny/2 + 1;

for (i=1; i<=nx; i++)
 for (j=1; j<=ny; j++)
     {
       /* smoke.pgm */
       if ( abs(j-centre_y) <=0 && abs(i-centre_x) > 2) {
         ur[i][j] = ui[i][j] = 0.0f;
       }
     }
/* ------------------------------------------------------------ */
```

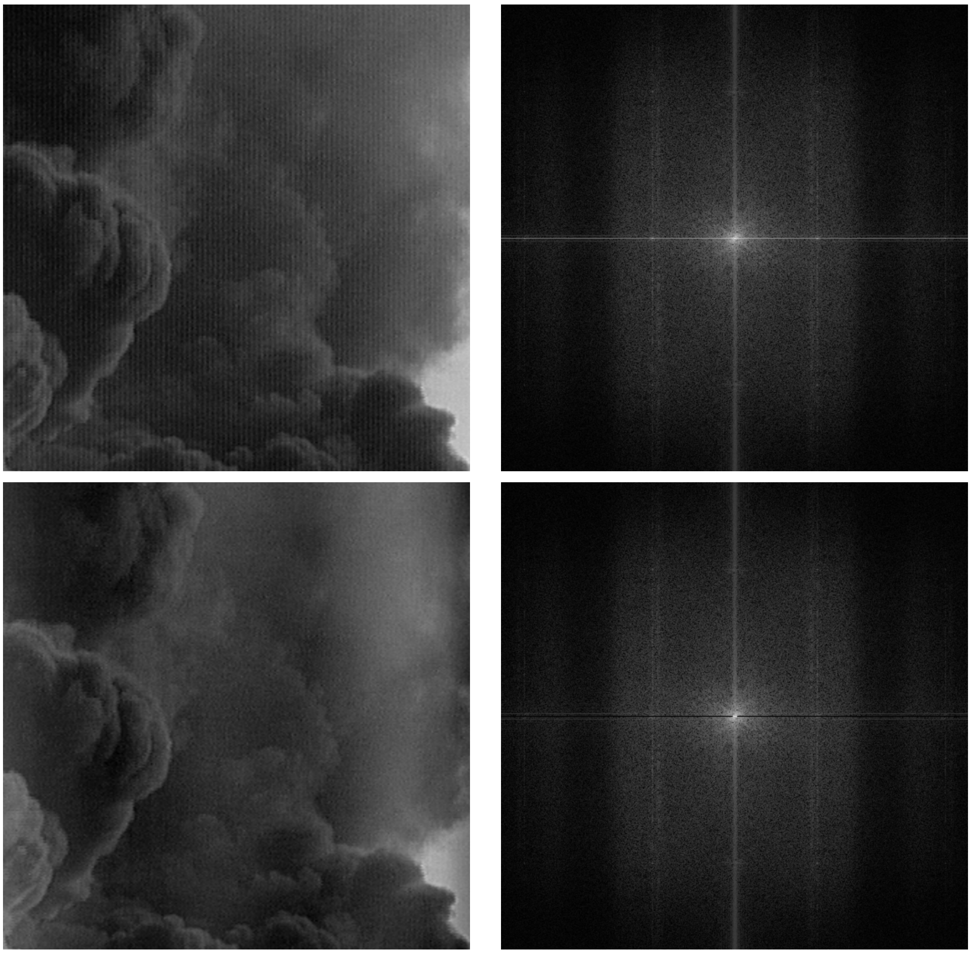


Table 1: Filtering results for `smoke.pgm`. *(a) Top left:* original image. *(b) Top right:* original Fourier spectrum. *(c) Bottom left:* filtered image. *(d) Bottom right:* filtered Fourier spectrum.

(b) For the image `fire.pgm`, a similar approach can be used to remove the artifacts. However, the vertical structures in the corrugated iron sheets of the hut create problems. These vertical structures occupy similar frequencies as the artifacts in the image. Removing the vertical line artifacts by removing the corresponding Fourier coefficients therefore destroys much more of the actual image structure than in the smoke example (see Figure 2).

```
/* ----------------------------------------------------------- */
long centre_x = nx/2 + 1;
long centre_y = ny/2 + 1;

for (i=1; i<=nx; i++)
 for (j=1; j<=ny; j++)
     {
       /* fire.pgm */
       if ( (abs(j-centre_y) <=1 && abs(i-centre_x) > 2) ||
            (abs(j-centre_y) <=10 && abs(i-centre_x) > 60 &&
             abs(i-centre_x) < 100)) {
         ur[i][j] = ui[i][j] = 0.0f;
       }
     }
/* ----------------------------------------------------------- */
```

Table 2: Filtering results for `fire.pgm`. *(a) Top left:* original image. *(b) Top right:* original Fourier spectrum. *(c) Bottom left:* filtered image. *(d) Bottom right:* filtered Fourier spectrum.