Analog signal processing

Textbook: Goodman (chapters 8,9)

Overview:

Spatial filtering: Abbe-Porter experiment, aberration compensation, phase contrast, Schlieren

Holography: Basics, holographic filters, the Vanderlugt filter, the matched filter

Thick holograms: Bragg conditions, acousto-optic deflectors

Digital holography: Types of SLMs, the Gerchberg-Saxton algorithm.

Spatial filtering

Spatial filtering relates to the use of filters in Fourier space. This can be an extremely potent tool in manipulating real-space data. In the following we discuss some simple variations on spatial filters. In particular, this can be related to a variety of topics, including microscopy, digital holography and Fourier domain optical pulse shaping.

The general optical; setup for this is a 4f correlator, where a phase and/or amplitude mask (possibly also polarization) is placed in the Fourier plane of a coherently illuminated object.

A simple (historic) demonstration of this is the Abbe-Porter experiments, where an amplitude filter is applied at the Fourier plane of a 2D periodic object, such as a binary amplitude grating:

$$t_a(\xi,\eta) = \sum_{n,m} rect\left(\frac{\xi - nL}{l}\right) rect\left(\frac{\eta - mL}{l}\right)$$

The FT of this is:

$$F\{t_a(\xi,\eta)\} = l^2 \sin c(f_x l) \sin c(f_y l) \sum_n e^{-2\pi i f_x nL} \sum_m e^{-2\pi i f_y mL}$$

The latter terms are equal to:

$$\sum_n e^{-2\pi i f_x nL} = \frac{1}{L} \sum_n \delta\left(f_x - \frac{n}{2}\right)$$

Yielding eventually a periodic array of sinc functions in both x and y.

Let us now filter only the central row of this in the Fourier plane.

This, instead of having:

$$\left(\frac{l}{L}\right)^2 \sin c(f_x l) \sin c(f_y l) \sum_n \delta\left(f_x - \frac{n}{2}\right) \sum_m \delta\left(f_y - \frac{m}{2}\right)$$

we only get:

$$\left(\frac{l}{L}\right)^2 \sin c(f_x l) \sin c(f_y l) \sum_n \delta\left(f_x - \frac{n}{2}\right) \delta(f_y)$$

Which, when FT, yields an image:

$$u_i(x,y) = \left(\frac{l}{L}\right)^2 \sum_n rect\left(\frac{x - nL}{l}\right)$$

Keeping just vertical lines in the grating pattern.

The first important application of such techniques are in imaging and microscopy. In these, there are two commonly used schemes – for conversion of weak phase objects to intensity modulations, and for image enhancement and restoration following aberrations. Let us discuss these first.

Suppose we have an aberrated transfer function. In particular, these typically lead to a reduced response at high spatial frequencies. By inserting an amplitude filter (and possibly a phase filter) which is the inverse of the transfer function we can "flatten" the frequency response and enhance the high frequency components to their original level.

A good example of this is the case of defocus, for which

$$\breve{H}(\rho) \approx \frac{J_1(\pi a \rho)}{\pi a \rho}$$

In this case, the combination of an amplitude filter reducing the central lobe and a phase shifting filter to inverse the alternating signs of the jinc function generates a better contrast. The problem is that this cannot be rendered perfect because of practical limitations:
1. The presence of noise
2. The existence of zeros in the transfer function, where information is irreversible lost
Note that this is coherent optical improvement of an incoherent image aberration.

A somewhat more rigorous analysis therefore requires consideration of these two. It can be shown that an optimal filter in this case is the Wiener filter, which takes into account the effect of noise. For a transfer function H and a noise spectrum Φn, we get the following result:

$$S(f_x, f_y) \approx \frac{H^*(f_x, f_y)}{\left|H(f_x, f_y)\right|^2 + \frac{\Phi_n(f_x, f_y)}{\Phi_o(f_x, f_y)}}$$

Which converges to the trivial result wherever noise in no present.

The second important application involves conversion of a weak phase object to an intensity image.

Generally, phase objects are not visible in bright field, except for some diffraction of light outside the optical system. Let us consider a weak phase object of the form:

$$t_a(\xi, \eta) = e^{i\Delta\phi(\xi,\eta)} \approx 1 + i\Delta\phi(\xi, \eta)$$

In this approximation, the phase object is, to first order, invisible, since it is in quadrature with the strong non-diffracted background, such that:

$$I_i(x,y) \approx |1 + i\Delta\phi(\xi,\eta)|^2 = 1 + \Delta\phi(\xi,\eta)^2 \approx 1$$

We can, however, by spatial filtering convert this phase object to an intensity modulation by inducing interference with the strong non-diffracted background. Two common techniques to do this are Zernike's phase contrast and Schlieren imaging.

In the first, a phase shift is introduced between the non-diffracted components and the diffracted ones. This is done by putting a π/2 phase plate (or 3π/2 phase plate) on the undiffracted beam in Fourier plane. The resulting response is now an interference pattern between the two:

$$I_i(x,y) \approx |1 \pm \Delta\phi(\xi,\eta)|^2 = 1 \pm 2\Delta\phi(\xi,\eta)$$

In the Schlieren technique, amplitude rather than phase filtering is used in the Fourier plane. Half of the Fourier plane is blocked by a knife edge, corresponding to an amplitude transmittance of the form:

$$H = \frac{1}{2} + \frac{1}{2}\operatorname{sgn}(f_x)$$

with a corresponding Fourier transform of the form:

$$h = \delta(y)\left[\delta(x) - \frac{1}{i\pi x}\right]$$

Thus, an object ta is converted, after filtering to a convolution with h:

$$u_i(x,y) = \frac{1}{2}\iint \delta(y-\eta)\left[\delta(x-\xi) - \frac{1}{i\pi(x-\xi)}\right] t_a(\xi,\eta)\,d\xi\,d\eta$$

or:

$$u_i(x,y) = \frac{1}{2}\left[u_o(x,y) + \int d\xi \frac{t_a(\xi,\eta)}{i\pi(x-\xi)}\right]$$

For $t_a(\xi,\eta) = 1 + i\Delta\phi(\xi,\eta)$ we get:

$$u_i(x,y) = \frac{1}{2}\left[1 + i\Delta\phi(x,y) + \int d\xi \frac{1 + i\Delta\phi(\xi,\eta)}{i\pi(x-\xi)}\right]$$

The unity in the integral vanishes, yielding finally a real term, linear in the phase shift from the integration. The total intensity is then to first order:

$$I_i(x,y)=\frac{1}{4}\left[1+\frac{2}{\pi}\int d\xi\,\frac{\Delta\phi(\xi,\eta)}{(x-\xi)}\right]$$

This gives an intensity dip and an intensity peak around the phase object, depending on location.

Holographic filters

In a holographic filter, the amplitude of the filter is due to the interference of a reference (plane wave) arm with an object of interest. Basically, this closely relates to the general field of holography, where phase information of an object can be stored by using the interference pattern with a known reference

Let us assume a general object wave Uo, which interferes with a reference plane wave Ur. If the two are coherent, then the total intensity of the summed field is:

$$I=\left|U_0+U_r\right|^2=\left|a(x,y)e^{i\varphi(x,y)}+A_r e^{i\psi(x,y)}\right|=\left|a(x,y)\right|^2+A^2+a(x,y)A\cos(\varphi-\psi)$$

Let us now assume that this light illuminates a film, for which the amplitude transmittance drops linearly with the total illumination intensity. We have thus generated a spatial mask of the form:

$$t_a(x,y)=t_0-k\left[\left|U_o\right|^2+\left|U_r\right|^2+U_oU_r^*+U_rU_o^*\right]=t_0-k\left[\left|a(x,y)\right|^2+A^2+a(x,y)A\cos(\varphi-\psi)\right]$$

Now lets illuminate this mask with a coherent plane wave similar to the incoming plane wave Ur. We then get that the multiplication of the excitation field by the mask gives:

$$U_ct_a(x,y)=\left[t_0-k\left(\left|U_o\right|^2+\left|U_r\right|^2\right)\right]U_r-U_r^2U_o^*-\left|U_r\right|^2U_o$$

Note that the last term is just the original object we had, undistorted. This means that a part of the diffraction from the recorded mask will be an identical reconstruction of the object. The only problem we are now faced with is how to differentiate the various contributions, so that the remaining readout beam and the conjugate component will not interfere with the restored image.

This is done with an offset recording and readout beam (using a large angle between the object and the recording beam). Assuming this corresponds to a phase function ψ which is linear in y.

Writing then the previous expression we get:

$$U_ct_a(x,y)=\left[t_0-k\left(\left|a(x,y)\right|^2+\left|A_r\right|^2\right)\right]A_re^{2\pi i\alpha y}-A_r^2e^{4\pi i\alpha y}a(x,y)e^{-i\varphi(x,y)}-A_r^2a(x,y)e^{i\varphi(x,y)}$$

Such that the first term propagates at an angle α, the second is diffracted at 2α, and the last (the one term we want to maintain) is on axis.

How large should the angle α be?

For this we should consider the angular spectrum of all the components. The first term contains two contributions: the constant one (Ar squared and t0) has, in fact, no angular spread. The second one is an absolute value of a, and thus from the autocorrelation theorem has a bandwidth of twice the bandwidth of a. The second term, as well as the third term have a bandwidth of a. Since we need to separate the first from the third, the general condition is that:

$$\alpha > 3B$$

Where B is the bandwidth of a, corresponding to:

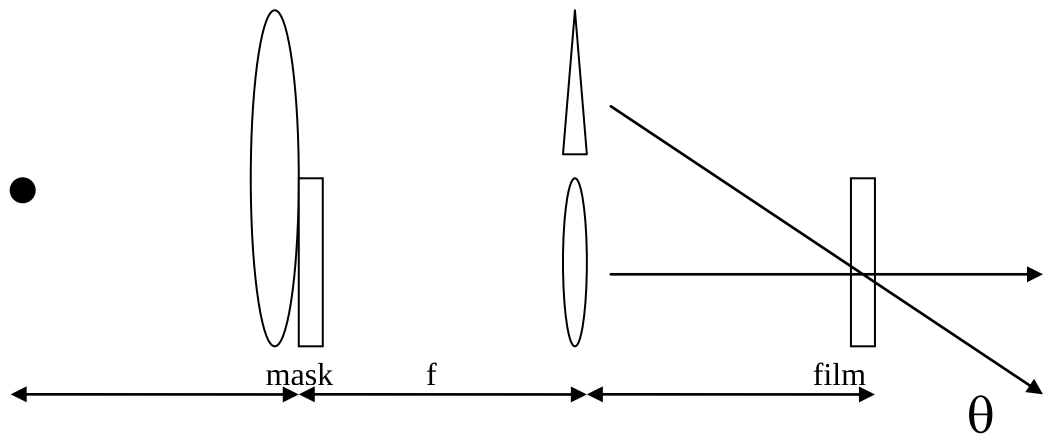
$$\sin(\theta) > 3B\lambda$$

If this is hard to fulfill, we can use a weak object illumination, such that the a squared term can be neglected. In this case we only require:

$$\sin(\theta) > B\lambda$$

<u>The Vanderlugt filter</u>

Now let us then first consider what happens when we take a pattern corresponding to the Fourier transform of some desired transfer function, and record the interference of this with an oblique plane wave (this corresponds to the recording conditions of the Vanderlugt filter).



Thus we have an interference pattern between an amplitude distribution (which has the form of some desired impulse response):

$$F\{h(\xi,\eta)\} = \frac{1}{\lambda f} H\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right)$$

and a plane wave:

$$U_r(x,y) = r_0 e^{2\pi i\alpha y}$$

Where:

$$\alpha = \frac{\sin\theta}{\lambda}$$

The total intensity distribution at the mask plane is therefore:

$$I(x,y) = \left|U_r + \frac{1}{\lambda f}H\right|^2 = r_0^2 + \frac{1}{\lambda^2 f^2}|H|^2 + \frac{r_0}{\lambda f}e^{2\pi i\alpha y}H + \frac{r_0}{\lambda f}e^{-2\pi i\alpha y}H^*$$

Similarly to the above description, we record a filter with this amplitude transmittance, and insert it at the Fourier plane of a 4f system.

Now consider what happens when we coherently illuminate an object g whose Fourier transform fulfills:

$$F\{g(\xi,\eta)\} = \frac{1}{\lambda f}G\left(\frac{x}{\lambda f},\frac{y}{\lambda f}\right)$$

Similarly to the previous case of the hologram, we get several contributions to the Fourier plane field:

$$U_2 = \left|U_r + \frac{1}{\lambda f}H\right|^2 G = \frac{r_0^2}{\lambda f}G + \frac{1}{\lambda^3 f^3}|H|^2 G + \frac{r_0}{\lambda^2 f^2}e^{2\pi i\alpha y}HG + \frac{r_0}{\lambda^2 f^2}e^{-2\pi i\alpha y}H^* G$$

Which we can now Fourier transform back to real space by the final lens. We see three contributions propagating in different directions, corresponding to three shifted images on the observation screen. The first two terms simply correspond to a replica of g and one convolved with h and h*:

$$U_3(x_3,y_3) = r_0^2 g + \frac{1}{\lambda^2 f^2}\left[h(x_3,y_3)\otimes h^*(-x_3,-y_3)\otimes g(x_3,y_3)\right]$$

And do not contain significant interesting information. The other two terms are:

$$U_3(x_3,y_3) = \frac{r_0}{\lambda f}\left[h(x_3,y_3)\otimes g(x_3,y_3)\otimes \delta(x_3,y_3+\alpha\lambda f)\right]$$

and

$$U_3(x_3,y_3) = \frac{r_0}{\lambda f}\left[h^*(-x_3,-y_3)\otimes g(x_3,y_3)\otimes \delta(x_3,y_3-\alpha\lambda f)\right]$$

Corresponding to the convolution of g with h and the cross-correlation of g with h.

Similarly, if the angles α is chosen to be large enough, the three images are completely separated in space.

One very important application of this is the matched filter. Suppose we are not interested in looking at the frequency response of the object or at its image, but are rather interested whether it is close or identical to some known object s (as in character recognition). In this case, we want to get the optimal SNR for this yes/no question. This is optically achieved by choosing, for the Vanderlugt function h:

$$h(x,y) = s^*(-x,-y)$$

In this case, the output of the convolution term in Fourier space is just SS* for the "correct" object, corresponding to a real function. This represents a plane wave after the filter, which would focus down to a diffraction limited spot. For other inputs, the outgoing correlation will generally be a complex function, resulting in a broad spatial distribution at the screen. The diffraction limited spot is optimal in the SNR sense, as noise is accumulated over a small area.

Holographic filters are also very useful in mode conversion of laser beams, since both the diffraction efficiency (corresponding to intensity) and the phase of the diffracted wave can be directly controlled in the recording process.

<u>A brief description of thick phase holograms (volume gratings, also relevant for acousto-optic modulators) – coupled mode theory</u>

For simplicity we can consider the case where the thick hologram contains reflective centers after development, but this is in fact equally relevant to a thick phase object, such as we get in an acousto-optic modulator.

Since the object is thick, we need to consider multiple reflections (or multiple coupling with the grating lines). We then get that the coupling of two spatial frequencies is greatly enhanced when they satisfy the Bragg condition (i.e. that reflections from various planes in the thick hologram are in phase). Mathematically, this corresponds to:

$$\sin\alpha = \pm\frac{\lambda}{2\Lambda}$$

where Λ is the grating period. This corresponds, in practice, to the angle between the two beams used to write the holographic grating.

To simplify the analysis under these conditions, let us assume a phase only grating with no absorption, such that

$$n = n_0 + n_1\vec{K}\cdot\vec{r}$$

and the two Bragg coupled waves:

$$U(r) = R(z)e^{i\vec{\rho}\cdot\vec{r}} + S(z)e^{i\vec{\sigma}\cdot\vec{r}}$$

with

$$\vec{\sigma} = \vec{\rho} - \vec{K}$$

We can now try to solve the Helmholtz equation inside the thick grating:

$$\nabla^2 U + k^2 U = 0$$
$$k^2 = k_0^2\left(n_0 + n_1 \cos\vec{K}\cdot\vec{r}\right) \approx (k_0 n_0)^2 + 2k_0^2 n_0 n_1 \cos\vec{K}\cdot\vec{r} = B^2 + 4\kappa B \cos\vec{K}\cdot\vec{r}$$

We then get:

$$\frac{d^2R}{dz^2}e^{i\vec{\rho}\cdot\vec{r}} + 2i\rho_z\frac{dR}{dz}e^{i\vec{\rho}\cdot\vec{r}} - |\rho|^2 \mathrm{R}e^{i\vec{\rho}\cdot\vec{r}} +$$
$$\frac{d^2S}{dz^2}e^{i\vec{\sigma}\cdot\vec{r}} + 2i\sigma_z\frac{dS}{dz}e^{i\vec{\sigma}\cdot\vec{r}} - |\sigma|^2 Se^{i\vec{\sigma}\cdot\vec{r}} +$$
$$B^2\left(\mathrm{R}e^{i\vec{\rho}\cdot\vec{r}} + Se^{i\vec{\sigma}\cdot\vec{r}}\right) + 2\kappa B\left(e^{i\vec{K}\cdot\vec{r}} + e^{-i\vec{K}\cdot\vec{r}}\right)\left(\mathrm{R}e^{i\vec{\rho}\cdot\vec{r}} + Se^{i\vec{\sigma}\cdot\vec{r}}\right)$$

The second derivative can be assumed small (slowly varying envelope approximation). The "standard" k2 terms for R cancel each other by definition, so we are left only with a mismatch for the scattered wave S, and from the last term we may keep only the coupling terms between R and S (which are strongly coupled under near Bragg conditions). Equating all terms with σ in the exponential to zero, as well as all terms with ρ, we are left with two coupled equations:

$$\frac{\rho_z}{B}\frac{dR}{dz} = i\kappa S$$
$$\frac{\sigma_z}{B}\frac{dS}{dz} + \frac{B^2 - |\sigma|^2}{B}S = i\kappa R$$

If the Bragg mismatch is zero (i.e. the incident wave is exactly the same wavelength and the same angle as the writing wave used in writing the grating) it can be easily seen that full conversion fro R to S and back occurs.

This can be realized either by holographic recording or in an acousto-optic device.

Digital holography

One big limitation of holographic applications such as the ones described above is in the tedious process of recording the holograms. The existence of reconfigurable SLMs opens a way to generate the "hologram" or the Fourier plane mask completely without the need for a film (or, for that matter, for resorting to intensity only masks). Most SLMs available today are 2D phase-only arrays, based on either LCD technology or on MEMS (micromirror arrays). Usually, both operate in the reflection mode (which doubles the dynamic range also for LCD masks, and facilitates manufacture). These devices suffer, however, from two main drawbacks:

1. artifacts due to the pixellation of the device.
2. A limit on the angular frequency of the diffracted wave due to the pixel size (and the sampling limit).- typically 10-20 microns for both.

Perhaps the most commonly used application of digital holography is in the generation of particular intensity patterns by phase modulation in the Fourier plane. This correspond to the solution of a very simple yet fundamental problem. On the SLM, we have a specific field amplitude pattern A(fx,fy) which cannot be manipulated. We have to find the spatial phase function which would generate the "optimal" (in some sense) intensity pattern I(x,y). This mathematical problem generally does not have a solution, so an approximate solution has to be found by some optimization algorithm. The most commonly used is the Gerchberg-Saxton one:

Take a random phase pattern, and IFT to the image plane.
Throw away all the intensity information, and multiply the phase pattern by the required image,
FT back to the Fourier plane.
Throw away all the intensity information, and multiply the phase pattern by the excitation source amplitude.

Repeat until convergence.