2023 4B11 Exam crib

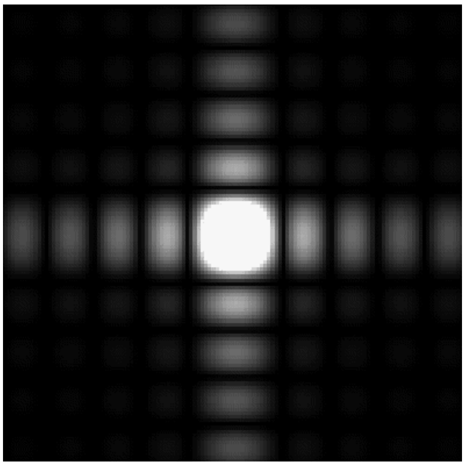
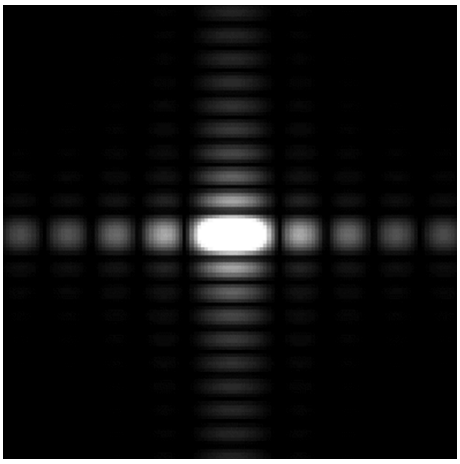
Q1 a) [30%] The exact structure of the far field depends on the shape of the 'fundamental' pixel and the number and distribution of these pixels in the CGH. The pattern we generate with this distribution of pixels is repeated in each lobe of the sinc function from the fundamental pixel. The far field of a single square pixel (Δ = a) is its Fourier transform:

$$F(u,v) = \iint_{\pm\infty} f(x,y) e^{2\pi j(ux+vy)} dxdy = \int_{-a/2}^{a/2}\int_{-a/2}^{a/2} A e^{2\pi j(ux+vy)} dxdy$$

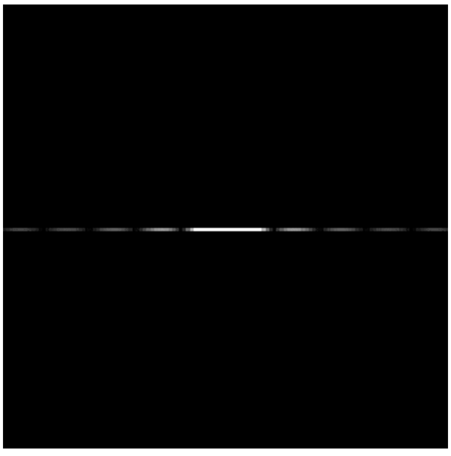
$$= \int_{-a/2}^{a/2} e^{2\pi j(ux)} dx \int_{-a/2}^{a/2} A e^{2\pi j(vy)} dy = \frac{A}{2\pi j}\left[\frac{e^{2\pi j(ux)}}{u}\right]_{-a/2}^{a/2} \frac{1}{2\pi j}\left[\frac{e^{2\pi j(vx)}}{v}\right]_{-a/2}^{a/2}$$

$$= \frac{A}{2\pi j}\left[\frac{e^{\pi jau} - e^{-\pi jau}}{u}\right] \frac{1}{2\pi j}\left[\frac{e^{\pi jav} - e^{-\pi jav}}{v}\right] = Aa^2 \operatorname{sinc}(\pi au)\operatorname{sinc}(\pi av)$$

Assume infinite plane wave illumination, a perfect FT from the lens or free-space, no apodisation, dead space. The result is a sinc function which forms the envelope for the whole replay field of a CGH made with square pixels or apertures. If we then expand the length of the pixel in the y direction (say to 2Δ), then the resulting *v* dimension in the replay field will be twice the original spatial frequency and the sinc function will get thinner by a factor of 2.

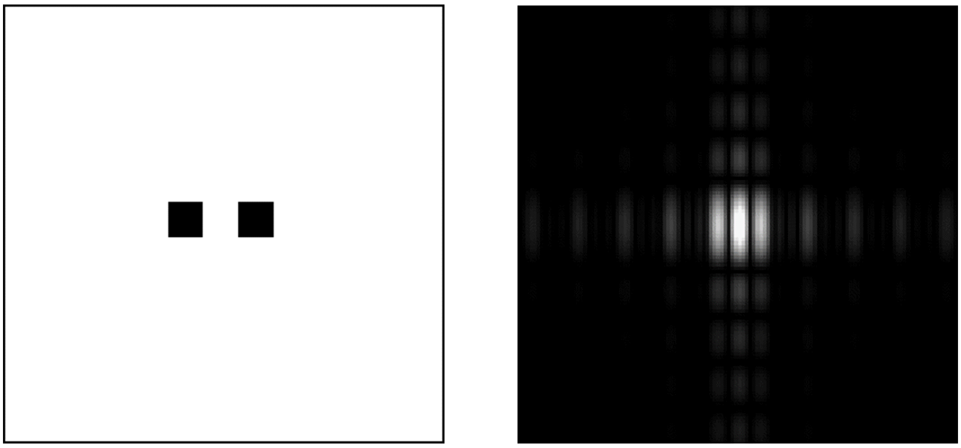

becomes


As we extend the pixel in the vertical direction, we will shrink the width of the sinc function in the v dimension of the replay field. Note that the u dimension does not change and remains as the fundamental sinc function. Eventually the single pixel will ne extended to a long thin vertical strip of width Δ as shown in Figure 1(a). If the strip is infinitely long, then the v dimension in the replay field becomes infinitely thin (in its limit a delta function.

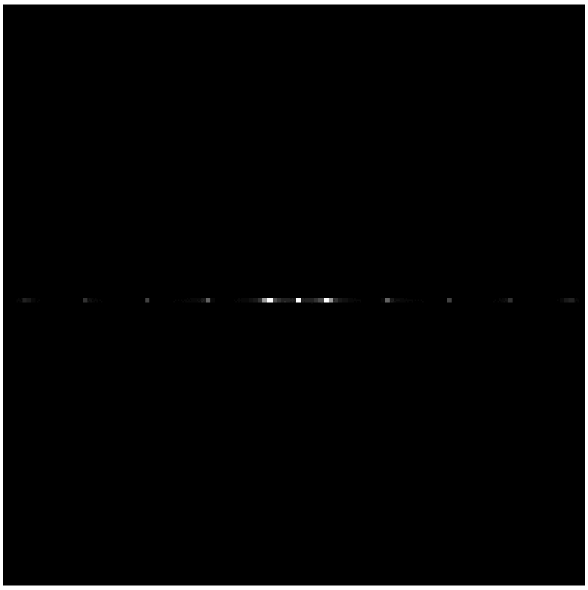


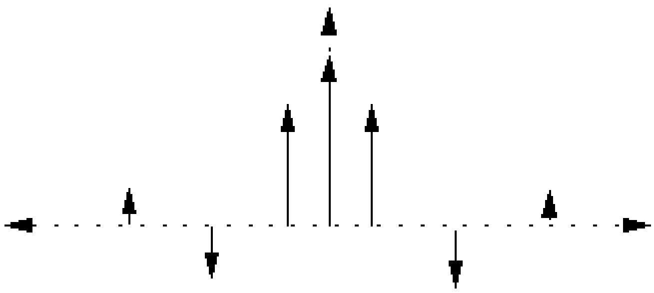
Note that the pixel wide strip is in fact shifted to the right by a distance d. This will not effect the intensity of the plot above but it will add a phase term in the Fourier plane.

b) [20%] The CGH in Figure 1(b) is now a grating of periodicity Δ, with 50/50 mark space ratio. This is effectively shifting the strip in part (a) and repeating it on a pitch of Δ. If we took the original pixel and copied and shifted it by a distance 2D then which has the effect of adding odd harmonics into the structure of the replay field in the same dimension as the shifted pixel.



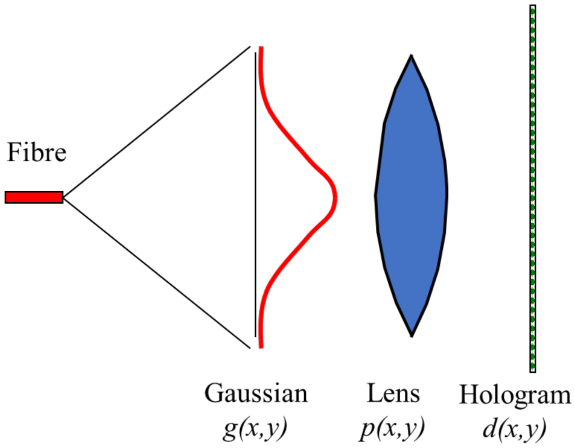
The same principle as in part (a) applies. As we add more shifted pixels, the width of the resulting sinc function decreases and ends up as a series of delta functions that have the same odd harmonic structure and have an envelope given by the original single pixel sinc function. If the shifted pixels are in fact the pixel wide strips shown in Figure 1(a), then the periodic structure in the replay field will be made from delta functions in both the u and the v dimensions.





This would be the side view of the central line given by the u axis.

c) [30%] The effect of the limited area of illumination given by the fiber laser is two fold. Firstly, only 2/3 of the area of the CGH will be covered by a circular aperture and the overall profile of this illumination will be Gaussian and not uniform.

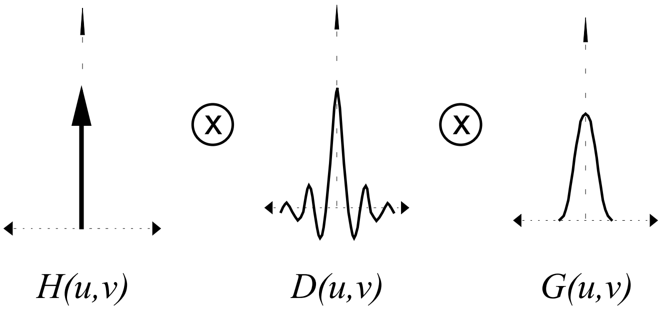


The entire illumination system can be modeled as a sequence of multiplied functions. The input illumination distribution *g(x,y)* times the hologram aperture *d(x,y)* times the total aperture of the FT lens (if it has a smaller diameter than the hologram) *p(x,y)*. Hence effect of the FT on these functions results in a convolution of their transforms.

$$Input = h(x,y)[g(x,y)p(x,y)d(x,y)]$$

$$RPF = H(u,v)\otimes[G(u,v)\otimes P(u,v)\otimes D(u,v)]$$

The ideal hologram replay field *H(u,v)* is designed as an array of delta functions in desired positions. The lens aperture *p(x,y)* is a large circular aperture, so the FT *P(u,v)* will be a first order Bessel function (like a circular sinc function). The hologram aperture is a large square of size $N\Delta$ and its FT, *D(u,v)* will be a sharp sinc function. The effect of the FT of the illumination *G(u,v)* is to add a Gaussian profile to each spot. Hence, the profile of the spots in the hologram replay field will not be delta functions, they will be delta functions convolved with a Bessel function convolved with a sinc function convolved with a Gaussian function. This means that the replay field will not look exactly as expected, spots which are placed next to one another will interfere due to the tails of the Gaussian, sinc and Bessel functions and the individual desired sharp 'spots' become ringed blobs with finite width..



d) [20%] The effects of the apodisation will have a strong impact on the performance of the single mode fibre optical switch.

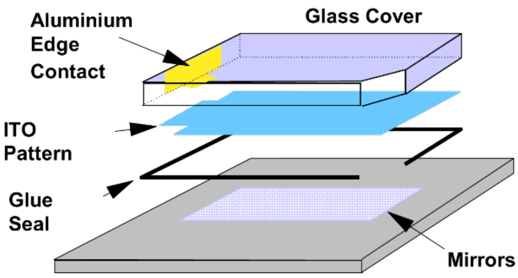
i) The expansion of the CGH spots from idealised delta functions to finite sized spots with Gaussian profiles and sidelobes will create optical loss. The use of single mode fibres means that the mode of the fibre must be matched as closely as possible to the profile of the CGH spot. Any deviation from the ideal Gaussian of the single mode fibre will lead to insertion loss. With 2/3 aperture this will great quite a wide spot profile that will not launch efficiently into the single mode fibre.

ii) The width of the CGH spots and also their sidelobes also means that the fibres in the switch cannot be placed too close together or they will pick up light from neighboring spots which will form crosstalk. This is very undesirable and given a typical target crosstalk figure of 40dB will severely limit the spacing and number of ports in the switch.

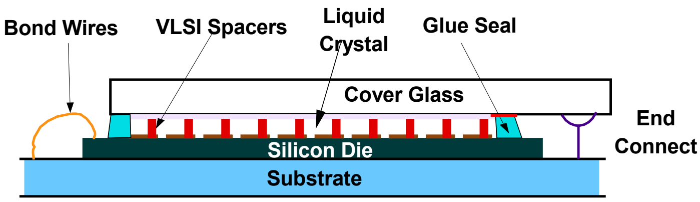
The effects of anodization can be mitigated by better optical design to maximize the area of the CGH illuminated. To avoid the sidelobes, it is best to avoid hard edges to the apertures. Also can exploit the symmetry in the replay field to try and minimize manty of these limitations.

A little disappointingly answers considering it was mostly bookworm. The thin strip example was coveted in lecture 2, but nobody managed to use the right analysis. A few tried the mathematical derivation and made good attempts. Even the section on apodisation, which was a fairly standard question was not answered that well with only a few understanding the relationship between the hologram and the aperture.

Q2. a) [35%] There are two main classes of SLM, transmissive and reflective. LCOS SLMs are reflective as they use a silicon backplane to support the mirror which reflect light through the LC layer. The LC is sandwiched between glass cover and the silicon backplane.



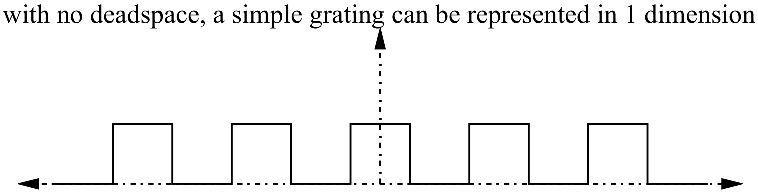
The use of the Si backplane means that the power of electronics can be i=used to interface with the SLM. This allow pixel arrays with very small pitches (<5um) and also allows them to be addressed at high speeds. Other features such as processing, demultiplexing and compression can be added to the electronics on the backplane.



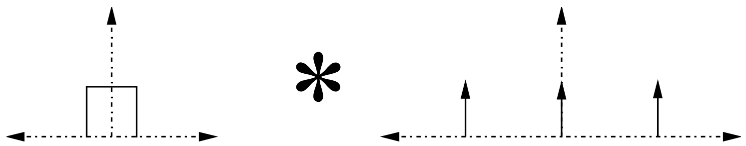
Some of the key parameters that result from the design of the Si backplane include the pixel density (number), size, pitch and corresponding deadspace. This effects the overall size and aperture of the SLM. Another key parameter is the thickness of the cell gap, or liquid crystal layer. This will impact on the type and degree of modulation possible with the choice of LC layer. The uniformity of the LC layer is also vey important as any variation in the thickness will appear like an aberration in the optical path which will then effect the apodisation of the replay field of the CGH.

If a NLC LCOS device is made, then the LC layer will have to be thicker to give the required 2pi phase depth of most multilevel phase modulation scheme (often more than 2pi is required). The modulation scheme is also grayscale which means that the pixel circuitry must be able to provide different RMS voltage levels to move the LC molecules.
If a FLC LCOS device is made, then the thickness of the LC layer is very thin which makes it harder to fabricate a uniform layer across the device. The modulation scheme is binary which much better suits digital electronics, but the LC must be DC balanced which adds more complexity to the addressing modes. FLCs are also capable of kHz frame rates, so the backplane has to be capable of handling very high f=date rates, of the density of pixels ifs high.

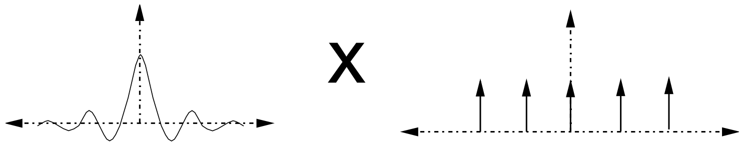
b) [30%] For a SLM with no deadspace, a simple grating can be represented in 1 dimension as:



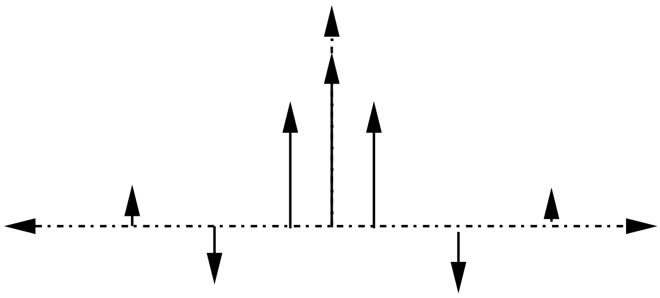
Which can be expressed as a convolution of two functions.



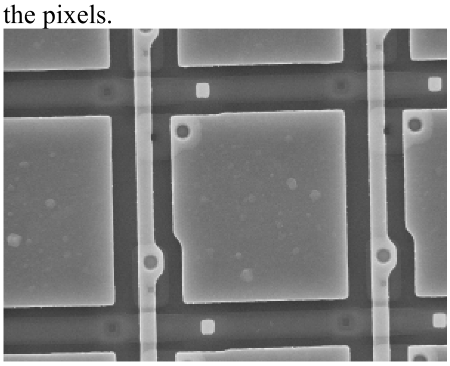
After the Fourier transform the convolution becomes a multiplication of the Fourier transforms of each function..



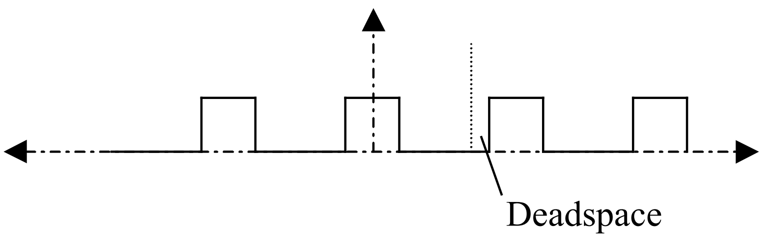
Gives the final result.



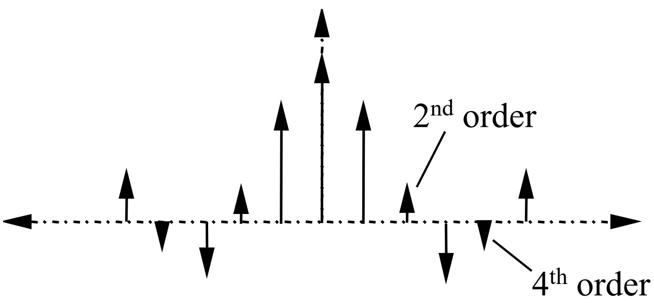
The key to LCoS is the silicon backplane which drives the LC pixels. As a result of this active backplane there will always be deadspace to isolate the pixels.



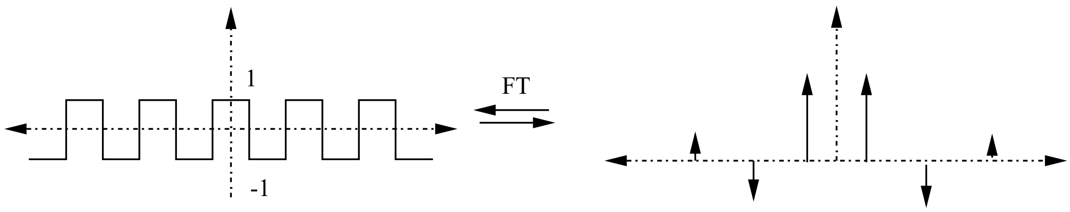
Above we assumed that the pixel pitch matched the finite width of the pixels themselves. This is not the case in reality as each pixel must have a gap between itself and its neighbour to prevent short circuits and to allow for transistors etc to be used to drive the pixels. This region between the pixels is referred to as the deadspace and the ratio of pixel size to pixel pitch is called the fill factor.



When we calculate the replay field of the grating with deadspace, we can no longer assume that the zeros of the sinc envelope due to the fundamental pixel shape will be in the same spatial frequencies as the delta functions due to the repetition of the pixels, hence the second order is no longer suppressed by the sinc envelope. This can lead to unwanted noise and crosstalk in the replay field.



c) [20%] When a hologram is designed with binary phase modulation, it is possible to suppress the zero order by having the same number of +1 as -1 phase states.



When the modulator has deadspace this is not guaranteed so there may be some form of imbalance in the phase states creating an unwanted zero order. The root of this problem is that the phase state of the deadspace is unknown or may depend on the state of the pixels either side.

The problem could possibly be avoided if the state of the deadspace was known or in some way related to the pattern on the SLM. This could be done by either a measurement of the SLM deadspace states or through clever design of the pixels to make sure the deadspace phase state is fixed. If the modulation state of the deadspace is known then it can be put into the hologram design algorithm and compensated for by the other phase states of the hologram pixels.

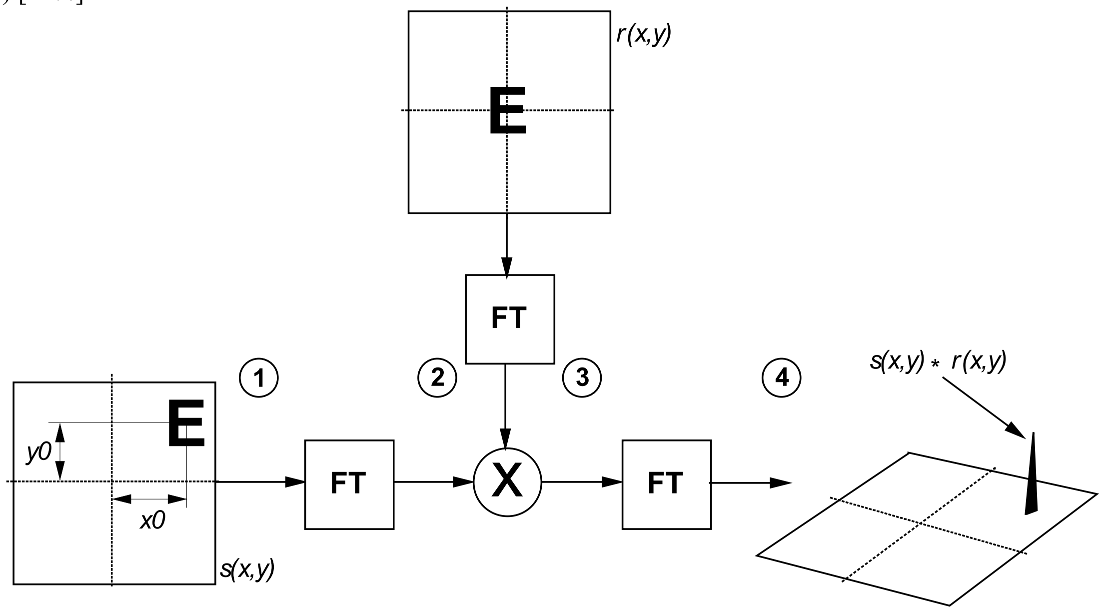
d) [15%] One of the most fundamental characteristics of an optical interconnect is its optical crosstalk. If the switch is configured to route light to the kth fibre in an array of n, then the crosstalk is the *ratio of light launched down the desired fibre to the light launched down one of the other fibres, which are not being routed*. If there is any significant unwanted orders in the replay field, then they will lead to crosstalk.

The effect of deadspace in a hologram replay field is the non suppression of the second order which also will contain other orders due to the multiple periodicities of the hologram. There could all end up launched down an unwanted fibre. This also applies to any extra zero order present which will also lead to significant increase in the crosstalk level. Both noise sources will effectively limit the number of low crosstalk positions in the replay field which limits the number of ports that can be switched.

The loss of light into unwanted orders due to deadspace also contributes the overall insertion loss of the optical switch.

Well answered overall as it was mostly bookwork. a) a lot did not realise LCOS required circuitry on the backplane. A few spotted that NLCs are RMS responding whereas FLCs are AC responding. Everyone assumed that FLCs were binary which is not true (only SSFLCs). Some mentioned DC balancing which was good. B) was well answered bookwork. c) and d) were mostly answered well with many realising that the DC could be compensated for in the CGH calculation.

Q3 a) [25%]

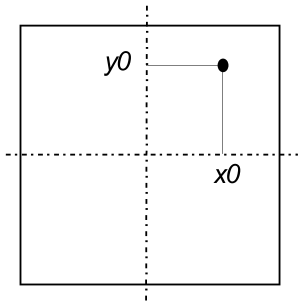


The input image *s(x,y)* is displayed in plane 1 before the FT into plane 2.

$$S(u,v)e^{-j2\pi(x_0u+y_0v)}$$

The FT of *s(x,y)* is then multiplied by the FT of the reference *r(x,y)*.

$$R(u,v)S(u,v)e^{-j2\pi(x_0u+y_0v)}$$



The FT of the reference is done off line on a computer and is defined as the <u>matched filter</u> *R(u,v)* for that particular reference *r(x,y)*. The product of the input FT and the filter then undergoes a further FT to give the correlation in plane 4. Advantages: the object in the reference *r(x,y)* is centred in the process of generating the filter *R(u,v)*, so that if a correlation peak occurs, its position is directly proportional to the object in the input image, with no need for any decoding. Unlike in the JTC, there is only one correlation peak and there are no DC terms to degrade the correlator output. Disadvantage: optomechanics and alignment of filter and FT are difficult.

b) [25%] The filter is matched to the reference via a Fourier transform such that.

$$F(u,v) = F_T[r(x,y)]$$

The best test for any matched filter is to perform an autocorrelation with the filter that has been generated. If the matched filter above is used for the reference image of a letter E, then the autocorrelation will have optimum SNR. The autocorrelation peak is very broad and has a huge SNR, as there is no appreciable noise in the outer regions of the correlation plane. Such a filter is not very useful for pattern recognition. Such a broad peak could lead to confusion when the position of the peak is to be determined. Also, similar shaped objects (such as the letter F) will correlate well with the filter leading to incorrect recognition. Another identical E which is placed in the input along with the original one will also cause problems as the correlation peak will take an extremely complex structure. Finally, the filter is a complex function and there is no technology available to display the filter in an optical system.

Great improvements can be made to the usefulness of the correlation peak, by using a phase only matched filter (POMF). The matched filter *F(u,v)* is stripped of its phase information (i.e. the phase angle of the complex data at each pixel) and this is used as the filter in the correlator.

$$F(u,v) = F_{amp}(u,v)e^{i\phi(u,v)}$$

Hence

$$F_{POMF}(u,v) = e^{i\phi(u,v)}$$

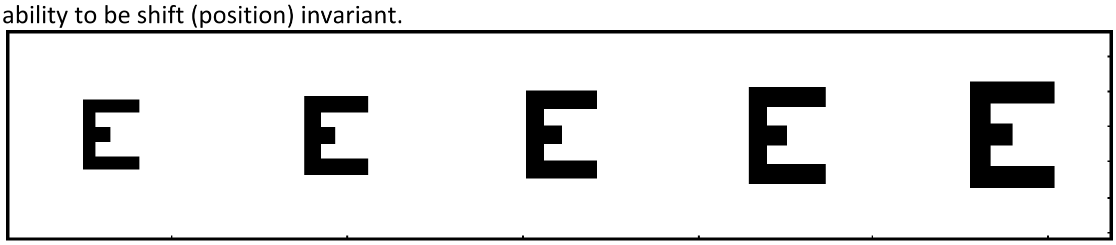
The autocorrelation for the POMF is much more desirable even though there is a reduction in the SNR due to the increase in the background noise. The correlation peak is much narrower which is due to the information which is stored in the phase of the matched filter. The POMF is the most desirable filter to use as it has good narrow peaks and still remains selective of similar structured objects (like Fs). The POMF is also a complex light modulation scheme, so the problems associated with binary phase (180º symmetry) will not occur. However, the continuous phase structure of $\phi(u,v)$ means that it cannot be easily displayed in an optical system.

The penalties associated with going to binary phase are greatly out weighed by the advantages gained by using FLC SLMs in the optical system. The binary phase is selected from the POMF by two thresholds δ_1 and δ_2.

$$F_{BPOMF} = \begin{cases} 0 & \delta_1 \le \phi(u,v) \le \delta_2 \\ \pi & \text{Otherwise} \end{cases}$$

1) The SNR is up to 6dB worse than the case of the POMF.
2) The filter cannot differentiate between an object and the same object rotated by 180º (due to the fact that the BPOMF is a real function).
3) The BPOMF is not as selective as the POMF due to the loss of information in the thresholding.

c) [25%] The SDF provides a way of achieving a limited form of invariance to the effects of image rotation and scaling. The idea is to take a series of reference images $r_1(x,y)$ to $r_n(x.y)$ and combine them by a linear summation. The final composite SDF image contains all of the references and so will correlate with the input image. Unlike mathematical transform techniques, the SDF does not destroy information and still retains the ability to be shift (position) invariant.



If we define a *1xn* vector **a**, whose elements are the weighting coefficients for each respective reference image, then the SDF is defined as.

$$h_{SDF}(x,y) = \sum_{n}^{i=1} a_i r_i(x,y)$$

We can easily calculate the weight coefficients from the cross correlations between all the reference images and the correlation peak value desired for each reference.

$$\mathbf{a} = \mathbf{R}^{-1}\mathbf{c}$$

Where **c** is a *nx1* vector whose elements are the desired correlation peak values for each image correlated with the SDF and **R** is the *nxn* correlation matrix. The elements of **R** are calculated from the cross correlation of the images in the reference set. Hence the element $R_{i,j}$ of the matrix **R** is the correlation between the reference images $r_i(x.y)$ and $r_j(x.y)$.

When the SDF is used in a matched filter and displayed as continuous phase only filter then the correlation peaks are reasonably close to the expected value of 1.0. If the phase only filter is then thresholded to form a binary phase only filter then the results are much worse. The correlator cannot distinguish between similar objects. This is the main drawback with the SDF, as the more references that are included, the more difficult it is to display the SDF on a modulator in a realistic optical correlator.

d) [25%] The choice of a FLC SLM is an attractive one as it has the highest frame rate of any LC electro-optical effect and large arrays of data can be processed optically at kHz frame rates. The problem with FLCs is that they are only capable of binary modulation. If a FLC SLM is to be used as the input SLM, it can be used in binary intensity mode. This is not such a limitation as the matched filter often performs better with binary inputs rather than grayscale ones. This comes from the Fourier transform which gives better selectivity when there are hard edges in the images rather than continuous intensity variations.

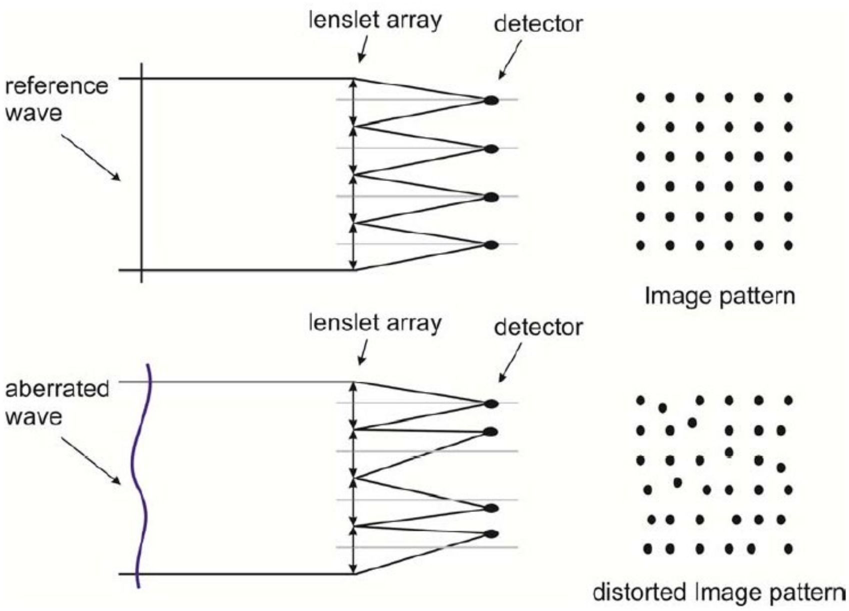
Ther is a bigger problem with using the FLC SLM as the filter SLM as this will most likely be in binary phase mode. While the SNR penalty in the correlation for using binary phase is only around 3-6dB, there is a loss of selectivity in the phase quantisation process which often means that the ability of the correlator to differentiate between similar object is reduced. The E vs F scenario is a good example. When a filter such as the SDF is used this is even worse, as the resulting filter is inherently grayscale and the quantisation process will lose important data in the correlation process. There are invariant filter techniques which can calculate the filter in the Fourier domain, but there is still a loss of data in the binary quantisation.

Another limitation that arises from the use of FLCs for the matched filter is the fact that use of a binary scheme means that complex phase values cannot be utilized. In the filter domain this means that the filter $F(u,v)$ will always be the same as it complex conjugate $F^*(u,v)$ and in the Fourier domain this means that a filter can not differentiate between an object $s(x,y)$ and its rotated version $s(-x,-y)$. In some applications this is useful and in some applications not.

One way to minimize the effect of the binary quantisation is to use an algorithm which directly generates a binary phase filter such as direct binary search or possible the dreaded machine learning.

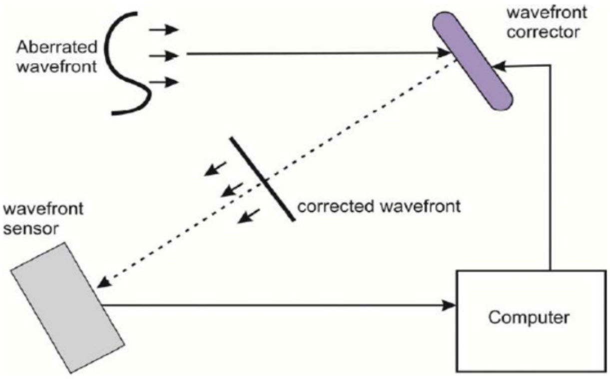
A fairly standard book work question that was answered pretty well by most candidates. a) was the standard introduction, answered well other than only a few pointed out the opto-mechanical limitations of the matched filter. b) well answered by most with a few pointing out the possible use of optimisation algorithms. c) well answered as a whole. d) some struggled with this part as they did not realise that the use of the FLC SLM meant binary phase which means the FT of the DSDF has to be quantised and this effects the scale invariance of the filter

Q4 a) [40%] One method for measuring the slope of the wavefront, is the Shack-Hartmann wavefront sensor. This is made by attaching a lens array to the front of a camera, spaced by the focal length of the lenslets. For a plane wave, a spot will be focussed on the optical axis of each corresponding lenslet in the array, as shown in the diagram below as the reference wave. For a distorted wave each focussed spot is displaced and this displacement is proportional to the slope of the wavefront. The incoming beam, whether it is the reference or the measurement beam, passes through the Hartmann screen which divides the wavefront into many subapertures. These are then focused on the Hartmann grid by the lenslet array. By comparing the difference in the coordinates between the expected and measured beams, we can obtain the slope of the wavefront.



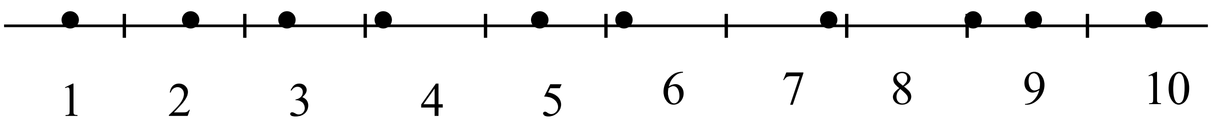
The main limitation comes from the zones defined by the lenslet array. Each lenslet samples a region of the incident wave. The spatial frequencies sensed by each zone will depend on its aperture. The more zones, the higher the frequencies of aberration detected, but also the higher the resolution of the camera needed and the more processing required. Each zone also limits the stroke of the sensor. If the slope of the wavefront in each zone is more than 2π, then it will push the centroid of the spot outside of the corresponding zone on the

camera leading to a false detection for its neighbour. Also, if there is too much information in each centroid then it will distort the spot making it hard to track.
The Shak Harmann wavefront sensor is used to detect the flatness of the optical wavefront in an adaptive optical system shown in the diagram below. It forms a closed loop system, where the function on the wavefront corrector is set to be the conjugate of the aberration. When this occurs, the resulting wavefront on the sensor will be flat.



b) [20%] Looking at the line scan (ignoring the other dimension), assume that the centroid of the focus spot shifts laterally with the slope of the phase across each lenslet zone. A slope of more than 2pi will lead to the centroid crossing into the adjacent zone. From figure 2, we can estimate the approximate slope across each zone.

Zone	1	2	3	4	5	6	7	8	9	10
Slope(pi)	0	0	0.5	1.5	0	1.5	-1.5	-2	0	0



c) [20%] The filter is matched to the reference r(x,y) via a Fourier transform such that.

$$F(u,v) = F_T[r(x,y)]$$

This filter is then converted to a phase only matched filter (POMF). The matched filter *F(u,v)* is stripped of its phase information (i.e. the phase angle of the complex data at each pixel).

$$F(u,v) = F_{amp}(u,v)e^{i\phi(u,v)}$$

Hence

$$F_{POMF}(u,v) = e^{i\phi(u,v)}$$

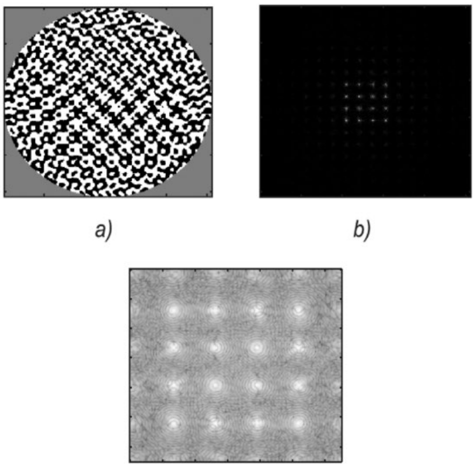
.At this stage , the phase function $\phi(u,v)$ can be modified top take into account effects of the aberration measured in the previous section. This is done by adding the conjugate of the aberration function to the filter phase so that it will cancel out with the actual aberration from the input optics. The resulting phase filter is then quantised to give the binary phase filter.

$$F_{BPOMF} = \begin{cases} 0 & \delta_1 \leq \phi(u,v) \leq \delta_2 \\ \pi & \text{Otherwise} \end{cases}$$

The problem with this technique is that the conjugated aberration function still has to undergo the quantisation process to binary phase and this will result in a loss of information and a loss in the performance of the aberration correction. The technique could also be applied to the output optics if the aberration is known using the same overall technique.

d) [20%] The function of the lenslet array can be performed by the hologram as Fresnel elements, These can be focussed onto a camera or a camera can be focused onto the replay field. If an SLM is used to display the hologram, then it can be programmed to display an hologram of an array of spots, one for each zone. Each zone is then imaged, and corrections are added to the hologram to see how the spot in that zone moves. Once the spot has been centred in the zone, the correction factor for the hologram can be read off as an estimate of the aberration.

A CGH can also be used to break down an aberration into its Zernike components.



a) b) c)

A well answered question which was based on one from a previous year. Most got part a) ok as it was mostly book work, although quite a few did explain how the Shack Hartmann sensor was used in an adaptive optics system. b) was well answered, but some of the data plots were a bit messy and unclear. c) most got that the aberrations could be included in the matched filter calculation, although only a few addressed the final lens. d) some interesting answers with a few suggesting a CGH, but that would have to use something like Zernike. Only a few suggested an array of holograms to mimic the lenslets