4B11 Photonic systems
Prof T D Wilkinson

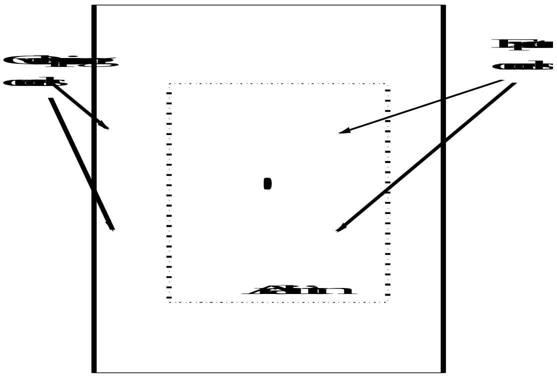
Q1 a) [30%] Limitations in the quality of the RPF are all due to the use of sampling within the SLM and binary phase modulation

1) 180° symmetry - a drawback of both the binary modulation schemes is that the hologram must always be a real function, hence

$$F_T\left[h(x,y)^*\right] = H(-u,-v)$$

If function $h(x,y)$ is real then $h(x,y)^* = h(x,y)$ and we cannot differentiate between $H(u,v)$ and $H(-u,-v)$ which means that both must appear in the replay field. Hence any replay field generated by a binary phase or amplitude hologram will always have 180° rotational symmetry. This symmetry restricts the useful area of the replay field to the upper half plane of the sinc envelope and results in a 50% loss of optical power in the RPF.

2) Outer orders - When an arbitrary pattern is generated by the FT of a hologram, it is contained within a sinc envelope based on the dimension of the pixels. For each lobe in the sinc there is an associated replication of the pattern. There is also a replication of the pattern about each zero in the sinc envelope. With holograms, we are only interested in the central lobe of the sinc envelope The other orders or lobes merely repeat the desired pattern and waste the available power in the central pattern. Within the central lobe, care must be taken with the orders that are replicated on the first sinc zero as they will overlap into the central lobe, leading to unwanted extra information in the replay field. For this reason, the area of interest in the replay field must be limited to half of the area of the central lobe to prevent order crosstalk.



3) Uneven noise floor - due to the limitations of the quantization to binary phase modulation, the distribution of the background power is not uniform and there tends to be small peaks of intensity, which may occur at unwanted positions. Given that most holograms will have a repetition of the spatial frequencies within them, these frequencies are also repeated in the RPF which further enhances the non-uniform nature if the noise floor in the RPF. This becomes less of a problem with large numbers of CGH pixels and careful CGH design.

Candidates may also mention the limits due to apodisation and the finite number of pixels in the hologram which leads to poor spatial resolution in the RPF.

b) [30%] The hologram need to be replicated onto a 2x2 grid to increase the resolution to the 2Nx2N of the SLM. The new hologram should have the form:

$$\begin{bmatrix} h(x,y) & h(x,y) \\ h(x,y) & h(x,y) \end{bmatrix}$$

This is effectively creates a shifted version of the holograms each shifted by an offset of N/2. Using the Fourier shift theorem:

$$f(x - x_0, y - y_0) = F(u,v)e^{-j2\pi(x_0 u + y_0 v)}$$

We have the result that the new replicated RPF H' will have the same overall structure

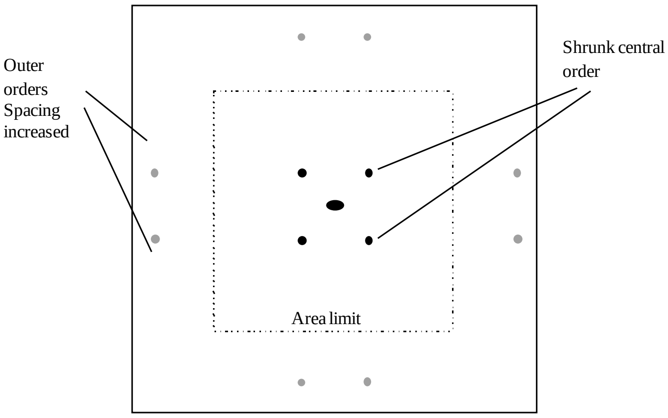
$$H'(u,v) = H\ (u,v)[e^{-j2\pi(uN/2+vN/2)} + e^{-j2\pi(-uN/2+vN/2)} + e^{-j2\pi(uN/2-vN/2)} + e^{-j2\pi(-uN/2-vN/2)}]$$

From this we see that the RPF is in fact unchanged in structure, but the phase contains extra components due to the 4 shifted replications. These phase terms can be summed to give the final result:

$$H'(u,v) = -4H\ (u,v)\sin(Nu/2)\sin(Nv/2)$$

Hence there is a modulation at the spatial frequency of N/2 which is effectively the Nyquist limit due to sampling as we would expect from the overlapping orders. Each spot in the RPF now appears to the sinusoid modulation on it which creates a dark fringe between every adjacent spot in the RPF. The factor of 4 represents the effective increase in intensity due to the replications compare to the original hologram.

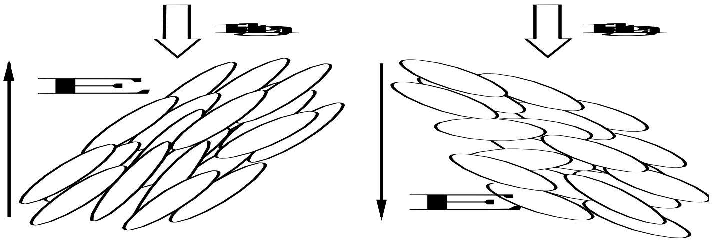
c) [20%] This form of replication is more difficult to express mathematically, but in its simplest form it is replacing one pixel in the hologram with 4 (in a 2x2 sub array). This is effectively increasing the pixel pitch of the hologram by a factor of 2 which has the effect of reducing the size of the RPF by a factor of two. However, the effect of the sampling also changes the relationship between the pitch of each pixel and its associated sampling rate by the same factor of two. This has the effect of pushing the outer orders further away from the central zero order. Hence the chances of an overlap from the outer orders in greatly reduced by this type of sampling. The resulting RPF is as on the right showing the new distribution of the orders in the RPF.



d) [20%] The third technique which could be used is zero padding, where the new hologram is set to a 2N by 2N array of zeros and then the N by N hologram is placed in the central region of this plane. This has two main effects on the RPF.

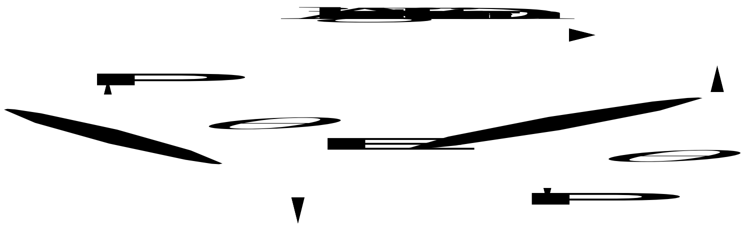
1) The area of the zeros is now a third state, which will block light leading to loss. This is less efficient.
2) The aperture of the hologram has been reduced by a factor of two, hence the effects of apodisation are increased by a factor of two and the spot size in the RPF is doubled. This leads to problems in defining the RPF resolution.

Q2 a) [30%] One of the commonest calimatic liquid crystal mesophases is the nematic phase. This is the least ordered mesophase before the isotropic. Here the molecules have only long range order and no longitudinal order. This means that the molecules retain a low viscosity, like a liquid, and are prone to flow. This can greatly effect the speed at which nematic LCs can modulate the light. The existence of dielectric anisotropy means that we can move the molecules around by applying an electric field across them. This combined with the flow properties means that a nematic molecule can be oriented in any direction with the use of an electric field. This is a very desirable feature as it leads to their ability to perform greyscale modulation of the light.



The calimatic molecular shape also leads to an optical anisotropy in nematic LCs, with the two axes of the molecule appearing as the refractive index. The refractive index along the long axis of the molecules is often referred to as the extraordinary n_e (or fast n_s) and the short axis the ordinary n_o (or slow n_s) axis. The difference between the two is the birefringence. $\Delta n = n_e - n_o$. One of the most useful smectic mesophases is the smectic C (SmC) phase as the molecules are highly ordered and form layers with the molecules tilted within each layer. The smectic C structure can be improved by adding chirality to the molecular structure which adds an extra dipole perpendicular to the molecular axis of the LC material. This is often is referred to as chiral smectic C or SmC*.

If the FLC is restricted to a cell thickness of 2-5μm then the helix of along the cell is suppressed and the molecules are bounded into two stable states either side of the director cone. The angle between these two states is defined as the switching angle θ. This is referred to as a surface stabilised FLC geometry an creates a high degree of ferroelectricity and creates a large birefingent electro-optical effect. The penalty for doing this is that the molecules are only stable in the two states an therefore the modulation will only be binary. The up side to this binary modulation is that it can be very fast (~10μsec) and that the stability can lead to the molecules remaining in the two states in what is known as bistable switching.

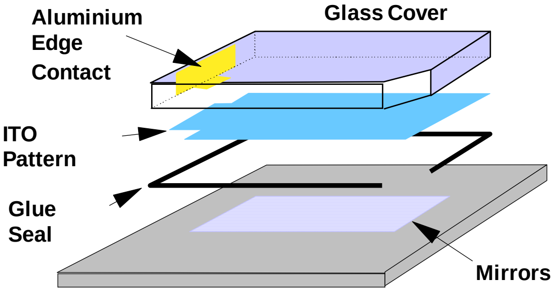


b) [20%] i) A single mode fibre to fibre optical switch – Speed is important, but not as important as loss and crosstalk. Therefore need to use a nematic LC for multi-level phase modulation. But this is polarisation dependent, hence we need an integrated quarter wave plate device. FLC could be used, but at a penalty of loss AND crosstalk.

ii) An optical packet switch – speed is everything. Need to modulate intensity below 1usec, hence an electoclinic modulator is Ok. Optical loss is high, but this is Ok for speed

iii) A matched filter optical correlator – filter SLM must be fast hence FLC is best. Binary phase is ok, but has no asymmetry in the recognised objects. To break symmetry, must use more than binary phase levels. Polarisation sensitivity is not an issue. Input SLM can be videao rate and grayscle so could be a nematic.

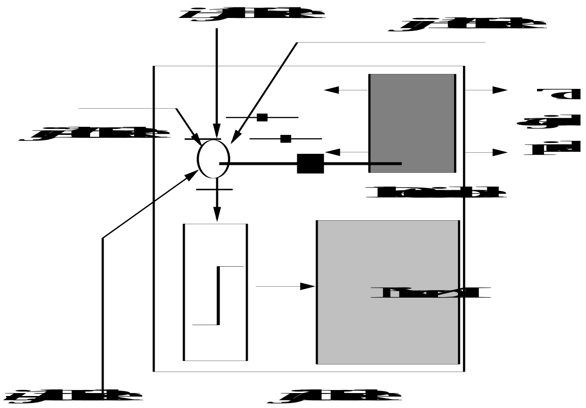
c) [25%] LCOS device has integrated circuitry and pixel mirrors on the silicon backplane. This means the device can be very small, compact and high speed as well as possible intelligence. All of the drive electronics can be integrated into the backplane which makes for a very compact microdisplay. The mirror quality can also be optimised using planarisation and cold evaporated aluminium.



A LCOS backplane is designed for a particular type of modulation or material. FLCs are binary materials which require very simple circuitry to drive the pixels. This means that the circuits can implement complex binary logic, memory, registers and counters which makes the implementation very flexible and sophisticated. Allows for DC balancing and bitplane greyscale type operations.

An LCOS backplane used with nematics is more complicated as it must maintain an analogue RMS voltage across the pixels. This needs bulky analogue/mixed circuitry and also includes circuits such as D to A converters and analogue multiplexers.

d) [25%] A logical advance in the design of a SLM is to combine the optical properties of the SLM with the processing ability of electronics. This is especially important when VLSI silicon backplane SLMs are used as the VLSI processes are identical to those used to create a standard silicon chip. For this reason, electronics is combined with the SLM to create smart pixels capable of processing information presented optically and then modulating light for further processing. Such pixels are very useful for both image processing and for optical neural networks hence they combine the modulating qualities of the LC material with the electronics and the use of silicon photodetectors at the pixel level.

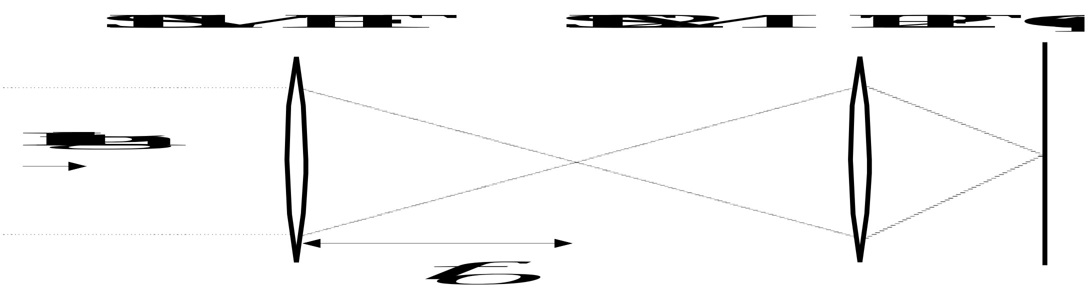


One type of smart pixel is the isophote or intelligent camera, which performs spatial image processing. In this case, the pixel is fed by the four adjacent pixels to the left, right, above and below. The photodiodes from the adjacent pixels are connected to the central pixel in the fashion shown. The functionality of the summation of these signal is set such that we have the left pixel subtracted from the right added to the top subtracted from the bottom.

$$P'_{i,j} = \left|P_{i,j+1} - P_{i,j-1}\right| + \left|P_{i-1,j} - P_{i+1,j}\right|$$

This performs the function of a Roberts Cross filter which is a means of achieving edge enhancement. More complex processes such as Sobel filtering (a combination of all eight nearest neighbours) could also be implemented.

Q3 a) [25%]



The basic optical layout for a BPOMF follows directly from the theoretical expectations. The input light illuminates SLM1 which is used to display the input image. SLM1 is also a FLC SLM, but it is used in intensity mode (black and white). The SLM is an N_1xN_1 array of square pixels, with a pitch of Δ_1, we are assuming that there is no pixel deadspace. The modulated light then passes through lens f_0 which performs the FT of the input image. The FT is formed in the focal plane of the lens and will have a finite resolution. Once the FT of the input (matched in size to SLM2) has passed through SLM2, the product of the input FT and the BPOMF has been formed. This is then FT'ed again by the final lens and the output is imaged by a CCD camera.

- Light source is collimated coherent laser illumination
- FT1 and FT2 are positive focal length lenses
- SLM1 and SLM2 are FLC transmissive SLMs. SLM1 is amplitude, SLM2 is phase.

The main limitation of this system is the alignment and scaling of the FT of the input image on SLM1 and the filter in SLM2. They must be matched and aligned to the nearest pixel and maintained by the opto-mechanics of the system. As the SLM shrinks in size, so does the pixel pitch which makes the alignment even more critical.

b) [25%] The pixel resolution of the SLMs dictate the size of the FT of the SLM1 and therefore the size of the overall BPOMF. This spectrum has to be matched to SLM2 as well. The FT is formed in the focal plane of the lens and will have a finite resolution (or 'pixel' pitch) given by. This comes from the sinc envelope for a pixellated object.

$$\Delta_0 = \frac{f_0 \lambda}{N_1 \Delta_1}$$

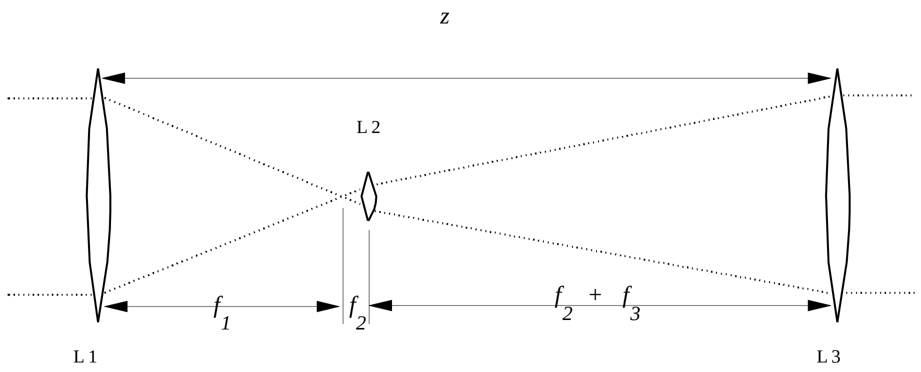
There are N_1 'pixels' in the FT of the input image on SLM1, hence the total size of the FT will be $N_1\Delta_0$. The BPOMF is displayed on SLM2 in binary phase mode. SLM2 is also a FLC device with N_2xN_2 pixels of pitch Δ_2. The FT of SLM1 must match pixel for pixel with the BPOMF on SLM2 in order for the correlation to occur. For this reason we must choose f_0 such that.

$$N_1 \Delta_0 = N_2 \Delta_2$$

Hence we can say.

$$f_0 = \frac{N_2 \Delta_2 \Delta_1}{\lambda}$$

c) [25%] It is possible to shorten the actual length of the optical transform whilst still keeping the effective focal length that is desired by including further lenses in a combination lens. One technique is to combine a positive lens with a negative lens to make a two lens composite. This gives a length compression of around $f_0/5$ which in the example above is still 2m and impractical. Furthermore, the two lenses combine in aberrations which leads to poor correlations due to optical quality.



The three lens combination above provides a much higher length compression of around $f_0/15$. The only drawback is that the more elements you have, the more position errors there will be when the optical system is mounted. The three lens system is an FT lens followed by a two lens telescope to magnify the FT to the size of SLM2. The first lens f_1 generates the FT of the input in its focal plane. The size of this FT will be too small, as f_1 is chosen to be much shorter than f_0. The FT is then magnified by lenses f_2 and f_3 which are configured in a telescope. Lenses f_2 and f_3 are spaced $f_2 + f_3$ apart so that an object placed a distance f_2 in front of lens f_2 will be magnified after lens f_3 by a magnification of f_3/f_2. Hence, if the FT is placed a distance f_2 in front of lens f_2, then it will be magnified after f_3. We can now set the magnification factor so that the input FT matches the size of the BPOMF.

$$Magnification = \frac{f_3}{f_2} = \frac{N_2 \Delta_2 \Delta_1}{f_1 \lambda}$$

d) [25%] If the chosen lenses and SLMs fit this equation, then the FT will be the same size as the BPOMF and the total length of the three lens system will be.

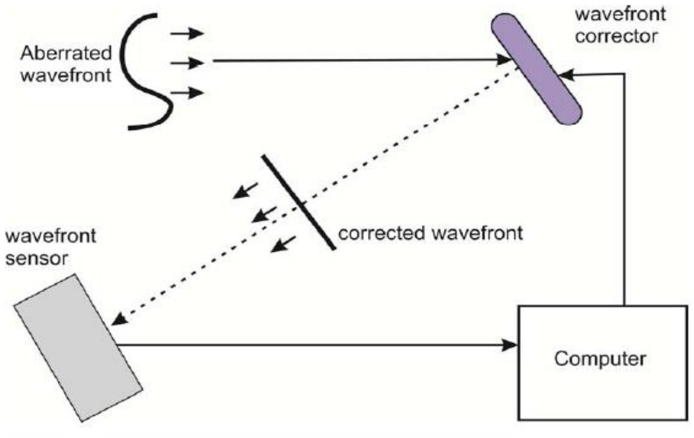
$$z = f_1 + 2 f_2 + f_3$$

When choosing the lenses, care must b taken to find good quality low aberration doublets. Choose a suitable f_1 length. Aim to create an FT of the input image which needs scaling by an integer value of magnification to match SLM2. The choose f_2 and f_3 to make this magnification.

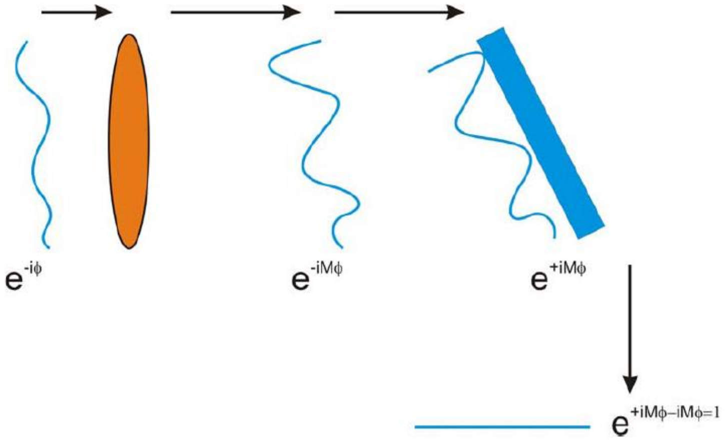
<u>Example</u> For the SLMs in the above example, a first lens f_1 = 250mm was chosen to perform the initial FT. This means a magnification of 39.1 is required for the telescope. From the available catalogue lenses, a combination of f_2 = 10mm and f_3 = 400mm was chosen to give a magnification of 40. A lens of f_3 = 365mm was also available, but was not chosen as the magnification was closer with 400mm. Although the other lens offers a shorter overall optical length, it is best to choose the lens

combination which offers the closest magnification to that which is desired. The overall length of the system was *z* = 670mm, which means that the BPOMF can now be constructed on an optical table.

Q4 a) [25%]



The basic principle is to determine in a rapid timescale how an image is being distorted by the medium or environment through which it passes and from the measurement of the distortions apply a correction to the wavefront and then use a "closed loop" feedback to generate a distortion-free image. Adaptive optics are heavily used in astronomy to correct for the aberrations introduced by the earth's atmosphere. Adaptive optic systems remove the phase aberrations that arise in propagating wavefronts to ensure that a high quality image is obtained. Generally, adaptive optic systems use the principle of phase conjugation to correct the distortions in the wavefronts. This is done by adding aberrations of equal amplitudes of the opposite sign.



b) [30%] Four common sources of aberrations in an optical system:
1) The optical medium which transports the waves and have variations in refractive index due to differing optical path lengths, thermal fluctuations etc.
2) Optical components such as lenses are never perfect and they introduce various aberrations due to the approximations in their design and the tolerances of their fabrication.
3) Many optical components are not optically flat hence they can lead to optical path length differences. Polarisers and SLM cover glass/substrates are good examples
4) If the system contain an optical source such as a laser, this can have a non-uniform phase front which will affect the performance of the system.

Aberrations can be described mathematically in optical systems by characterizing the aberrations in terms of power series expansions. As most optical systems have an axis based on circular symmetry, it is useful to expand the wave aberration in terms of a complete set of basis functions that are orthogonal over the interior of a circle. Zernike polynomials satisfy this condition as they form a complete set of functions or modes that are orthogonal over a circle of unit radius. Other power series descriptions are not suitable as they are not orthogonal over the interior of a circle. For a circular pupil or aperture, the aberrations in a wave can be described by a weighted sum of Zernike polynomials.

In the context of imaging, Zernike polynomials can be used to fit aberration of the wavefront in an optical system and using Fourier transform properties to determine the Point Spread Function (PSF). We use Zernike polynomials because they have a number of beneficial mathematical properties. They are orthogonal over the continuous unit circle and all their derivatives are continuous. They are efficient at representing errors that are common in optics such as coma, spherical aberration, and astigmatism. Finally, they form a complete set, meaning that they can represent arbitrarily complex continuous surfaces given enough terms.

c) [25%] The input wave is assumed to be optically correct. If there were a distortion from the source then there would be an extra term $s'(x,y)$ in the expression below to represent the source aberration.

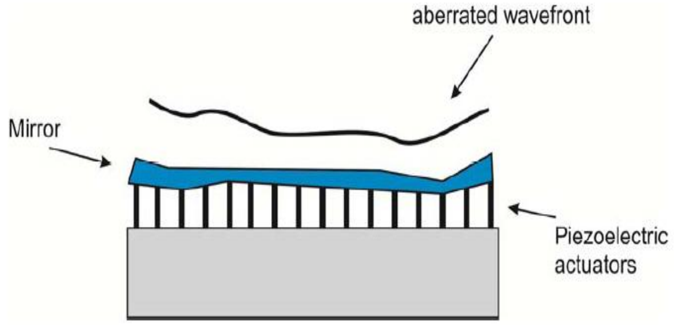
The medium in which the optical wave passes would ideally have a transfer function of 1, hence the presence of $g(x,y)$ must be aberration. This needs to be conjugated out.

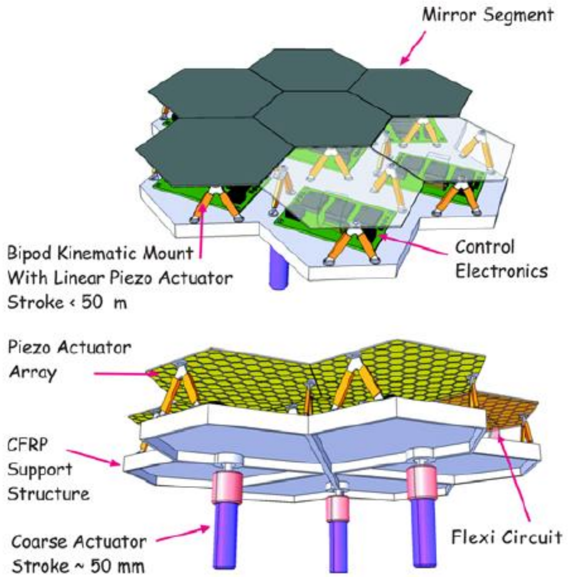
The optical system will have a transfer characteristic which represents the function of the system and will include an aberration term $p'(x,y)$ which should be conjugated out.

The overall conjugation function should take into account the three factors above and can be calculated as $c(x,y) = s'(x,y)*g(x,y)*p'(x,y)*$ where * denotes the complex conjugate of the function. This function can be implemented optically as follows

Deformable mirrors include all those which use a mechanical means of deformation to match the conjugate wavefront. They can be continuous: such as membrane and bimorph mirrors or they can consist of discrete actuators that are perpendicular to the surface or at the edge of the mirror. Right is an example of a deformable mirror that is driven by piezoelectric actuators, which extend or contract when a voltage is applied. The mirror, as a result, deforms to match the conjugate wavefront.

Segmented mirrors are formed by closely spaced small flat mirror segments. Each segment can approximate the average value of the wavefront over the patch area. Considerable improvement of the performance of the segmented mirror can be achieved by





introduction of three degrees of freedom per segment.

d) [20%] In a Fourier system such as a holographic optical projector, the image is generated in the far field, hence the aberration c(x,y) is Fourier transformed to C(u,v) and is then convolved with the ideal replay field image to create a distorted or blurred version of the image. This is very undesirable and even only a few waves of non-flatness in an optical component can lead to significant distortion. Phase conjugation could be applied to this system, but the location of the conjugate in not simple and the effect is still dominant in the final RPF.

A key advantage of computer generated holograms is that they use algorithms to optimise the relationship between the hologram plane and the RPF generated. If there is any known distortion in the system, then this can be added to the algorithm and then the optimisation process removes or minimises the effect of it in the RPF. This requires no additional hardware and requires very little extra processing power when calculating the holograms.