The principal limitation to the widespread deployment of X-ray phase imaging in a variety of applications is probably versatility. A versatile X-ray phase imaging system must be able to work with polychromatic and non-microfocus sources (for example, those currently used in medical and industrial applications), have physical dimensions sufficiently large to accommodate samples of interest, be insensitive to environmental disturbances (such as vibrations and temperature variations), require only simple system set-up and maintenance, and be able to perform quantitative imaging. The coded-aperture technique, based upon the edge illumination principle, satisfies each of these criteria. To date, we have applied the technique to mammography, materials science, small-animal imaging, non-destructive testing and security. In this paper, we outline the theory of coded-aperture phase imaging and show an example of how the technique may be applied to imaging samples with a practically important scale.

1. Introduction

Phase contrast imaging has been performed routinely using visible light since well before Zernike received the Nobel prize for his invention of the phase contrast...
microscope in 1953 [1]. Imaging with visible light has led X-ray imaging in the development of phase imaging techniques principally due to the faster development of technology, which is, in general, more difficult in the case of X-ray imaging. As technology and methods mature, it is reasonable to assume that X-ray phase contrast imaging (XPCI) will have a significant impact upon a great number of applications, just as it has on the visible regime.

Quantitative XPCI techniques are particularly important as they are necessary for performing computed tomography. Existing quantitative XPCI techniques fall into three general categories. One category is that of in-line holography in which radiation, perturbed by the sample, interferes with radiation directly incident upon the detector [2,3]. A source of highly spatially coherent radiation is required, thus necessitating the use of a microfocus or synchrotron source. Algorithms have been developed for determining the phase of a sample from its hologram for weakly absorbing [4–6], pure phase [7,8] and homogeneous objects [7,9]. The principal limitation of this technique is that a synchrotron source is required to acquire images with a sufficiently short exposure time so as to be compatible with most applications.

Another category is analyser-based imaging (ABI), which makes use of the rocking curve of an analyser crystal to develop contrast [10]. A number of algorithms have been developed that enable phase and absorption information to be separated [11–17]. This technique requires a highly collimated beam of high spectral purity to be incident upon the sample, thus limiting it to synchrotrons or monochromated laboratory sources [10,18], which appears to be this method’s principal limitation.

The category of techniques that currently appears to possess the greatest potential for widespread application is that of grating interferometry. This technique employs two or three gratings, depending on the source employed, and uses the phenomenon of Talbot self-imaging. In this technique, sample differential phase and absorption information is encoded in the relative shift and mean value of fringes incident upon a detector. This technique has been implemented using synchrotron radiation [19–21] as well as laboratory sources [22,23], in which case a source grating is required to aperture the source, thus resulting in an array of coherent, but mutually incoherent, secondary sources.

The coded-aperture technique presented in this paper fits somewhere between the ABI and grating interferometry categories. On first inspection, the method appears to share similarities with the grating method owing to the presence of two sets of apertures. However, the method does not make use of the Talbot effect, which relaxes the coherence requirements of the source relative to the grating technique [24]. The coded-aperture method is, however, based upon the edge illumination principle [25], which was developed as a synchrotron method analogous to ABI. The edge illumination principle shows how strong differential phase contrast may be achieved by illumination of a sharp boundary between two pixels and forms the basis of a recent quantitative technique based upon the coded-aperture method [26].

As mentioned earlier, the coded-aperture technique is built on the edge illumination principle first demonstrated by Olivo et al. [25]. The first coded-aperture XPCI images acquired with a conventional source were acquired in 2007 [27,28]. This led to a number of applications being investigated, including homeland security [29–31], mammography [32], high-sensitivity quantitative imaging [26] and small-animal imaging [33].

Figure 1 shows a schematic diagram that illustrates the edge illumination principle as deployed using a synchrotron source. The diverging beam depicted is perhaps slightly exaggerated, as the source to sample distance, $z_{so}$, may be very substantial depending on the beamline; however, there will always be a degree of beam divergence. The system may be assumed to extend out of the page, in which case an array of detector pixels would be employed to enable a profile to be obtained with each acquisition. The sample may be scanned in the vertical direction to build up an image. When the sample is in the vicinity of the axis denoted by a dashed line in figure 1, the refraction of photons either into or out of the sensitive region of the detector will result in contrast due to a phase gradient in the vertical direction. Experiments have shown that very strong phase contrast can be achieved using this set-up [25,34]. An important point to note is that $W$ can be increased without degrading spatial resolution. Increasing $W$ causes more
Figure 1. Schematic diagram showing the edge illumination principle. A set of slits is used to form a beam of width $W$, which is incident upon the sample. At least part of the beam is incident upon an absorbing edge positioned in front of the detector. $\Delta l$ is the physical width of the beam that overlaps the sensitive part of the detector; $z_{so}$ is the distance between the beam-forming slit and the source, which, in the case of a synchrotron, can be many metres; $z_{od}$ is the distance between the beam-forming slits and the detector.

Figure 2. Schematic diagram of the coded-aperture system (left). An aperture $A_1$ is used to form a series of beams, which are incident upon absorbing edges provided by the detector aperture $A_2$. The width of each beam after $A_1$ is $W$, but the beams diverge to a width of $MW$ at $A_2$, where $M = (z_{so} + z_{od})/z_{so}$. A flat panel detector with pixel width $P$ is employed, which defines the periodicity of the apertures. The displacement between the centre of each pixel and the centre of each beam, $\Delta P$, is an important quantity, which strongly affects the imaging properties of the system. $\Delta P$ is used to control $\Delta l$, a parameter common to both the coded-aperture and edge illumination systems. A magnified view of part of the coded-aperture system revealing key parameters (top right). The coordinate system employed throughout this paper (bottom right).

background photons to be detected, so any refracted, i.e. signal-forming, photons will cause a perturbation against a higher background. In this sense, increasing $W$ leads to reduced contrast in the image, and so the maximum allowable value for $W$ largely depends on the number of perturbed photons relative to the noise of the system. A reason for having a large $W$ is that it allows the Fresnel number of the diffracting slit at the detector, $W^2/(z_{od}\lambda)$, to be much greater than unity for a wavelength $\lambda$. This means that the system is well described by geometrical optics. This is by no means essential but it enables the forward, and therefore inverse, problems to be described by analytic formulae. This is also a compelling reason for employing the Talbot effect in grating interferometry.

Figure 2 shows a schematic diagram that illustrates the coded-aperture system. The coded-aperture system is closely linked to the edge illumination principle, as is evident from the diagram. The width and separation of the beams are limited by the periodicity of the flat panel.
detector. Off-the-shelf detectors are generally employed to make the system as widely applicable as possible. In particular, we currently employ two detectors in our laboratory, an ANRAD a-Se (SMAM) featuring a pixel size of 85 \(\mu\)m and HAMAMATSU (C9732DK passive pixel CMOS flat panel with directly deposited structured CsI) with a pixel size of 50 \(\mu\)m. The total system length, \(z_{so} + z_{od}\), is generally of the order of 2 m, based on simulations and experiments [32], but a shorter system length is possible and is currently being tested experimentally. This system, too, may be assumed to extend out of the page; however, unlike in the case of edge illumination, sample scanning is not required owing to the array of beams. However, high-resolution images may be obtained by scanning of the sample, by a subpixel amount, if desired.

2. Mathematical analysis of coded-aperture technique and resulting versatility

We provide here a simplified mathematical description of the coded-aperture system. For this purpose, we use the arrangement in figure 3, which contains a simple, representative, wedge object. For simplicity, we consider a system that is uniform in the \(y\) direction of figure 3, although the extension to a two-dimensional system is straightforward. We assume initially a monochromatic point source located a distance \(z_{so}\) from the aperture \(A_1\), which, however, may be located off axis at position \(\tilde{x}\). As is customary, we describe a sample by a transmission function whereby the perturbation to a wave front, at an exit surface very close to the object, is described by a complex function [35], which we denote \(T(\tilde{\xi}) = \exp(-i\phi(\tilde{\xi}) - \mu(\tilde{\xi}))\), where \(\phi\) and \(\mu\), both real valued, are defined as

\[
\phi(\tilde{\xi}) = k \int_\mathcal{O} \delta(\tilde{\xi}, z) \, dz \quad \text{and} \quad \mu(\tilde{\xi}) = k \int_\mathcal{O} \beta(\tilde{\xi}, z) \, dz, \tag{2.1}
\]

\(\mathcal{O}\) is the extent of the object and \(k\) is the wave number. Then, by applying the paraxial approximation to the Fresnel–Kirchhoff diffraction integral, the irradiance of the field a distance \(z_{od}\) from \(A_1\) may be found as [36]

\[
I(x, \tilde{x}, \phi') \approx \frac{|U_0|^2}{k z_{so} z_{od} (z_{so} + z_{od})} \exp(-\mu(0)) \times \int_{-W/2}^{W/2} \exp \left( ik \xi^2 \frac{z_{so} + z_{od}}{2z_{so} z_{od}} - ik \xi \left( x + \frac{z_{od}}{z_{so}} \tilde{x} + \frac{z_{od}}{k^{\prime} \phi'} \right) \right) \, d\xi, \tag{2.2}
\]

where \(\phi\) has been approximated by a first-order Taylor series expansion about \(\xi = 0\), \(\phi'\) represents the gradient of \(\phi\) in the \(\xi\) direction at \(\xi = 0\), \(\mu\) has been assumed to be constant within the beam and the point source is assumed to radiate a spherical wave of the form \(U_0 \exp(i k r)/r\) at a distance \(r\) from the source. This expression is, however, only a partial result because the irradiance is integrated over a region defined by the aperture \(A_2\) and the effect of an extended source must also be incorporated. Suppose, then, that the source focal spot may be described by a function \(\sigma(\tilde{x})\), often assumed to be Gaussian in the literature, and that the combined effect of pixel and aperture leads to a sensitivity function \(Q(x, \Delta P)\), where \(\Delta P\) is incorporated to enable \(A_2\) to be shifted. Then, the pixel signal for a particular position of \(A_2\), after interchanging the order of integration, is given by

\[
S(\Delta P, \phi') = \int_{-\infty}^{\infty} I(x, 0, \phi') K(x, \Delta P) \, dx, \tag{2.4}
\]

where

\[
K(x, \Delta P) = Q(x, \Delta P) \ast \left[ \frac{z_{so}}{z_{od} \sigma} \left( \frac{z_{so}}{z_{od} x} \right) \right], \tag{2.5}
\]
Figure 3. Diagram of a unit of the coded-aperture system or, equivalently, of the edge illumination principle, defining the mathematical quantities used in this paper. Note that the refraction angles have been exaggerated for clarity.

and $*$ is the convolution operator and $K(x, \Delta P)$ is introduced here as the effective pixel sensitivity function. In the case of a point source and a perfectly absorbing detector aperture, for example, $K(x, \Delta P)$ would take the form (figure 2)

$$K(x, \Delta P) = \begin{cases} 
1 & -MW/2 + \Delta P < x < MW/2 + \Delta P, \\
0 & \text{otherwise}.
\end{cases} \tag{2.6}$$

However, this formalism allows parameters, such as detector aperture width and source focal spot size, to be accounted for. The general strategy employed in both grating interferometry and the coded-aperture technique is to develop a means of deriving $\phi'$ from two or more values of $S(\Delta P, \phi')$ for different values of $\Delta P$. In the case of the coded-aperture technique, two values are obtained for two equal and opposite values of $\Delta P$, which results in the inversion formula [26]

$$\phi' \approx \frac{\sqrt{\pi}}{2} \frac{S(-MW/2, \phi') - S(MW/2, \phi') \sigma_d}{S(-MW/2, \phi') + S(MW/2, \phi')} \sigma_{zd} \tag{2.7}$$

where $\sigma_d = \text{FWHM}_d/(2\sqrt{\log(2)})$ and FWHM$_d$ is the full width at half-maximum of the beams projected onto $A_2$. This formalism demonstrates the versatility of the coded-aperture technique. In particular, we note that, within very accommodating limits, the source size, aperture dimensions and system layout may be varied and quantitative phase imaging may still be performed. This allows the system to be adapted to very many applications.

3. Ramifications of versatility

The versatility of the coded-aperture technique has been explained from a mathematical viewpoint in §2 and from an intuitive point of view in §1. This versatility results in the coded-aperture technique possessing a number of practical advantages over other techniques. One important advantage of the coded-aperture technique that seems not to be generally understood is that the sensitivity and spatial resolution are not determined by the periodicity or fill factor of the pre-sample aperture, $A_1$. This is demonstrated by comparing recent experiments on porcine cartilage [37], using a grating interferometer, and on murine cartilage [33], using a coded-aperture system. In particular, the contrast for the two configurations can be directly compared from Marenzana et al.’s [33] fig. 2c (murine cartilage in water) and Stutman et al.’s [37] fig. 9d (porcine cartilage in water), which both demonstrate a phase contrast modulation of the order of approximately 5% at the interface between cartilage and water. In the case of the grating interferometer experiments, it was necessary to place the analyser grating at the third Talbot distance, resulting in an interferometer of length 2.5 m with 1.25 m between sample and detector, in order to be sensitive to the small refraction angles induced by the porcine cartilage in water. The source, analyser and phase gratings each had a periodicity of 10 $\mu$m, with the former having a 50%
duty cycle. Further, the authors of this work explain that a significantly reduced contrast would be achieved at the fifth Talbot distance, thus effectively placing a limit on the system’s sensitivity to small refraction angles. In contrast, the coded-aperture system used a distance of 2 m between the source and the detector and only 0.4 m between sample and detector, while employing a pre-sample aperture with periodicity of 66.8 $\mu$m and opening of width 12 $\mu$m, yet achieved contrast comparable to that of the grating interferometer.

Other advantages of the coded-aperture system are that, by accommodating a range of aperture periodicities, a flat panel detector compatible with industrial or medical applications is able to be used and that standard laboratory sources may be employed without needing to aperture the source. This means that the only portion of source flux not used in image formation is that absorbed by the apertures. We note that the system described above uses approximately up to 18% of the available source flux for imaging. This is the amount of flux incident upon the sample; the amount used in image formation depends on how the pre-sample and detector apertures are aligned. This system was designed specifically to perform low-dose mammography [38], which is one reason why only 18% of source flux is incident upon the sample. Other configurations of the system, however, demonstrate how apertures can be used that transmit 50% of incident flux [27,28,39]. The flux used in image formation by these configurations can thus vary between nearing zero (dark-field configuration), 25% (commonly used XPCI configuration) and nearing 50% (bright-field configuration). This has important ramifications in reducing the cost of systems and ensuring that the exposure time is sufficiently low to be compatible with the majority of applications. The relatively large periodicity and size of the transmitting regions of the apertures represent a further strength of the method in that the apertures are able to be fabricated routinely and with physical dimensions large enough to accommodate mainstream medical and industrial applications. This also has important ramifications for reducing the system cost.

We have not discussed at length in this paper the effect of a polychromatic source. However, this represents another strength of this method, as, because it is not interferometric, it can be used with sources having a very broad spectrum without experiencing image degradation due to lack of temporal coherence [24]. Of course, sensitivity to phase gradients at high photon energies may be lost as the apertures become transparent, but this is common to all techniques employing apertures or gratings. However, it has been shown that the coded-aperture technique still performs well at high photon energies [29].

A final strength that we note is that the system is strongly immune to vibrations as may be experienced in, for example, a hospital due to foot traffic. A manuscript is currently in preparation that demonstrates the robustness of the system to vibrations induced by people walking in close proximity to the instrument.

4. Examples

As the first example, we show a quantitative result obtained using synchrotron radiation at the Elettra synchrotron radiation facility in operation in Trieste, Italy. The experiment is fully described in a recent publication [40] and the result we present here differs in that a photon energy of 25 keV was employed. It is interesting to note that identical systems were used to obtain accurate quantitative results at both 20 and 25 keV photon energies. The experimental system used was similar to that of figure 1 although images were taken for both positive and negative values of $\Delta P$. Figure 4 contains an image of the differential phase gradient of a poly(ether ether ketone) (PEEK) fibre with a nominal diameter of 450 $\mu$m. A single scan line is plotted along with the analytic comparison, which confirms the accuracy of the technique. PEEK has values of $\delta = 4.58 \times 10^{-7}$ and $\beta = 1.51 \times 10^{-10}$ [41].

As another example, we have selected a sample image of a marker pen, which demonstrates the high sensitivity and spatial resolution that may be obtained by using the coded-aperture system. Figure 5 contains retrieved absorption and phase images of the marker pen. The retrieved absorption image is analogous to a conventional absorption radiograph, while the phase image has a grey level in proportion with the refraction angle induced by the sample. The images
Figure 4. (a) Image of $\phi'$ for a PEEK fibre of nominal diameter 450 µm and (b) a profile plot of the differential phase. The experiment profile plot (black) is extracted from the location indicated by the line in the phase image. The analytic value of the differential phase profile is denoted by the grey line. The differential phase image was acquired using synchrotron radiation under the conditions described in this section.

Figure 5. (a) Absorption and (b) refraction images of a marker pen acquired using a coded-aperture XPCI system.

were obtained using a coded-aperture system such as that illustrated in figure 2 with total system length, $z_{so} + z_{od}$, of 2 m, with $z_{so} = 1.6$ m and $z_{od} = 0.4$ m. The pre-sample aperture, $A_1$, had a period of 66.8 µm with an opening of 12 µm, while the detector aperture, $A_2$, had a period of 83.5 µm and an opening of 20 µm. A Rigaku 007HF X-ray tube generator operated at
35 kV_P/25 mA with a rotating Mo target was used. Images were acquired with an ANRAD a-Se (SMAM) flat panel detector featuring a pixel size of 85 µm. Images were taken at a total of four subpixel image positions to increase the resolution of the images. Each such image was acquired using an acquisition time of 70 s, meaning that the images in figure 5 took a total of 560 s to acquire, because images are required to be taken for both positive and negative displacements of the detector aperture.

The phase image in figure 5 contains many details that are easily visible in the phase image but not in the absorption image. Some of these details have been surrounded by grey boxes. It is also interesting to note that the texture of the porous tip is clearly depicted in the phase image. The only obvious region where the phase image shows less detail than the absorption image is the region annotated with an arrow. A vertical band is seen in the absorption image, which is missing in the phase image. This is because this band of material is very close to being perpendicular to the direction in which the system is sensitive to phase gradients. Detail can, however, be seen at the top of this band where the material boundary bends.

Funding statement. This work was financially supported by the UK Engineering and Physical Sciences Research Council (EP/G004250/1 and EP/I021884/1). P.R.T.M. is currently supported by a Discovery Early Career Research Award from the Australian Research Council (DE120101331).

References


