Analysis of Energy Dispersive X-Ray Diffraction Profiles for Material Identification, Imaging and System Control

Emily Jane Cook

University College London

July 2008

Submitted for the Degree of Doctor of Philosophy

University of London

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Emily Jane Cook, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis presents the analysis of low angle X-ray scatter measurements taken with an energy dispersive system for substance identification, imaging and system control.

Diffraction measurements were made on illicit drugs, which have pseudo-crystalline structures and thus produce diffraction patterns comprising a series of sharp peaks. Though the diffraction profiles of each drug are visually characteristic, automated detection systems require a substance identification algorithm, and multivariate analysis was selected as suitable. The software was trained with measured diffraction data from 60 samples covering 7 illicit drugs and 5 common cutting agents, collected with a range of statistical qualities and used to predict the content of 7 unknown samples. In all cases the constituents were identified correctly and the contents predicted to within 15%.

Soft tissues exhibit broad peaks in their diffraction patterns. Diffraction data were collected from formalin fixed breast tissue samples and used to generate images. Maximum contrast between healthy and suspicious regions was achieved using momentum transfer windows 1.04-1.10 and 1.84-1.90 nm^{-1}. The resulting images had an average contrast of 24.6% and 38.9% compared to the corresponding transmission X-ray images (18.3%). The data was used to simulate the feedback for an adaptive imaging system and the ratio of the aforementioned momentum transfer regions found to be an excellent parameter. Investigation into the effects of formalin fixation on human breast tissue and animal tissue equivalents indicated that fixation in standard 10% buffered formalin does not alter the diffraction profiles of tissue in the momentum transfer regions examined, though 100% unbuffered formalin affects the profile of porcine muscle tissue (a substitute for glandular and tumourous tissue), though fat is unaffected.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank Professor Robert Speller and Dr Gary Royle, my supervisors, for their advice and support. Special thanks also go to Dr Silvia Pani for her help with my work and her friendship throughout. I am also grateful to; Dr Marinos Metaxas, for providing the soundtrack to my work; Walid Ghoggali and Mohammed Faizal for being on LN_2 duty; Dr Jennifer Griffiths, Sarah Bohndiek and all the Radiation Physics group at UCL; Leah George, David Wilkinson and Sheila Hardwick at HOSDB for their help; Drs Julie Horrocks and Ruby Fong from the St Bartholomew's Clinical Physics group for their support; Nicholas Skipper, Nirmal Perera, Setarah Chavoushi and all the other undergraduates that have grappled with X-ray diffraction over the past few years. Finally, thanks go to Thomas Allen.

Chapter 1

Introduction

1.1 Introduction

This thesis presents work undertaken to investigate the analysis of energy dispersive X-ray diffraction profiles for material identification, imaging and as a control function for adaptive X-ray imaging.

Diffraction patterns are known to be characteristic of the scattering substance producing them, as the positions of the peaks are governed by the inter-atomic and inter-molecular spacings within the material. Energy dispersive X-ray diffraction (EDXRD) techniques utilise this, using polychromatic X-ray sources and thus photons with energies high enough to penetrate extended objects.

This work examines different methods of extracting information from diffraction profiles using a range of techniques. The method of analysis depends on the information required, which is driven by the application of the system. Three applications are considered here.

The identification of a material from a range of possible materials using its diffraction profile requires complex analysis. Interference peaks can occur over a wide range of momentum transfer values and in order to produce a specific system these must all be taken into account. This requires a significant amount of data to be processed, leading to long computation times and large errors. However, as in each profile there are some regions that are essential in the charactersation of the substance, there are also regions that contain no pertinent information, thus could be discarded, reducing the amount of data. Multivariate analysis can do just that, examining the latent structure of the data and using techniques such as principle components analysis (PCA) and partial least squares (PLS) regression to identify regions of the profile that are important and correlated to the substance content, those which are structured but unimportant and those which contain merely noise.

For imaging the information that needs to be extracted from the diffraction profile is very different. Using data from a large number of peaks would

lead to a confusing image that would be difficult to interpret (e.g. by a radiologist). Instead a more continuous parameter is required that can be assigned to a grey or colour scale, producing a clear image with easily distinguishable elements. Thus simple features in the profiles such as the intensity of a single, or at most two, peaks could be used.

System control has similar requirements to those of imaging. Control requires a simple parameter, or combination of parameters, to guide an automated function. The parameters must be easily calculated, to minimise calculation times for real time control, and relate to physical quantities such as sample thickness and content. Thus continuous parameters are required such as in imaging, but more complex functions can be used as the result is to be output to an automated system, not to be visualised.

This thesis considers these techniques and applies them to real world applications.

For the problem of material identification illicit drugs are considered. They are pseudo crystalline materials that produce diffraction patterns with multiple peaks. An example of a real-world scenario which could use this technique is in the detection of illicit drugs in parcels, which is a problem faced by customs officials. In this work the diffraction profiles of a range of narcotics commonly imported into the UK are measured, their characteristics noted and multivariate analysis used to build mathematical models to predict the content of future samples. The results of some predictions of 'unknown' samples are presented and the effect of packaging material on the predictions considered.

Breast tissue is used to explore the imaging capabilities of diffraction techniques. It is suitable as breast tissue has been shown to produce broad peaks characteristic of tissue pathology (Kidane et al. (1999)), with adipose tissue having a sharp peak at 1.1 nm^{-1} and tumourous tissues producing broader peaks at 1.6 nm^{-1}. Thus simple images of different tissue types may be built by considering these regions. Five tissue samples are imaged

here, and the resulting profiles examined to gain maximum contrast in the resulting images between different tissue regions.

The results from the breast tissue imaging are developed to investigate feedback in an adaptive imaging system. The proposed 'I-Imas' system aims to produce optimal mammographic images by using feedback in real time to control exposure parameters throughout the image. Here the possibility of using diffraction information as a feedback parameter is investigated and the results compared to those from other statistical functions.

The remainder of this chapter describes X-ray interactions, the fundamentals of energy dispersive x-ray diffraction and outlines previous investigations in this field.

Chapter 2 describes the EDXRD systems used in these investigations, their alignment and the momentum transfer resolution achieved. An introduction into multivariate analysis techniques is included.

Chapter 3 contains the results of the EDXRD measurements of 7 illicit drug samples, 5 common cutting agents and a range of packaging materials, explains the multivariate anaylsis used to describe these data and make predictions on unknown samples. A comparison of the predicted values and actual components of the unknown samples is also presented and the compromise between counting statistics and momentum transfer resolution considered.

Chapter 4 investigates tissue suitable for imaging studies. The equivalence of pork tissue to human breast tissue is investigated and the effects of formalin fixation on the energy dispersive diffraction profiles of tissue substitute materials (pork adipose and muscle) and human breast tissue considered using both a laboratory based X-ray diffraction system and synchrotron ra-

diation.

Chapter 5 describes the tissue images obtained by diffraction measurements made on fixed breast tissue samples, the optimisation of the contrast in the images, the corrections applied to the images to account for attenuation effects, and analysis of the images obtained.

Chapter 6 investigates the potential of using diffraction data as a feedback parameter in an intelligent mammographic system and compares the results to those from other statistical functions.

Chapter 7 Contains the conclusions from this study and details some of the future work necessary to take these ideas forward.

1.2 X-ray interactions

As X-rays pass through a material some of them will interact with its constituent atoms. The possible interactions that lead to attenuation of the primary beam are coherent scattering, incoherent scattering, photoelectric absorption and pair production. As pair production can only occur at energies above 1.02 MeV, it is not considered here as this is outside the range of X-ray energies used in diagnostic and industrial radiology. As an X-ray beam passes through a material it is attenuated by the above processes and the intensity of a beam that has travel a distance x through a material with a mass attenuation coefficient (μ/ρ) is given by equation 1.1.

$$I(x) = I_o e^{-\rho(\frac{\mu}{\rho})x} \tag{1.1}$$

where ρ is the density of the material and μ/ρ is dependent on the energy of the incident X-ray and the material the beam is passing through.

1.2.1 The photoelectric effect

The photoelectric effect does not contribute to the scattered radiation as the photon's energy is deposited during the interaction, though it contributes to the total attenuation. The X-ray photon interacts with a bound electron, ejecting it from the atom. If an electron from a higher shell drops down to fill the hole its energy is released either by the emission of an Auger electron or a characteristic X-ray. If this characteristic X-ray is able to escape the medium, it could be detected and mistaken as scatter.

The cross section for photoelectric interactions (τ) strongly depends on the atomic number of the material and the energy of the incident photon, with the atomic cross sections having the measured dependences described in equations 1.2 and 1.3.

At high energies

$$\tau \propto \frac{Z^5}{(h\nu)} \tag{1.2}$$

At low energies

$$\tau \propto \frac{Z^4}{(h\nu)^3} \tag{1.3}$$

For the incident photon energies used in diagnostic medical and industrial radiology, equation 1.3 is applicable.

1.2.2 Incoherent scatter

Incoherent, or Compton scatter occurs when the incident photon interacts with a loosely bound electron, and in doing so imparts some of its energy to the electron. The photon undergoes a change in both direction and energy and its new direction and energy are related by equation 1.4,

$$h\nu' = \frac{m_o c^2}{(1 - \cos\theta) + (\frac{m_o c^2}{h\nu})} \tag{1.4}$$

where the incident photon has a frequency ν and the scattered photon has a frequency ν' and is traveling at an angle θ to its initial path, m_o is the mass of the electron, c the speed of light, and h, Planck's constant. The differential cross section for a Compton interaction is given by

$$\left(\frac{d^2\sigma_{incoh}}{\delta\Omega\delta E}\right) = \left(\frac{d\sigma(\theta)}{\delta\Omega}\right)_{KN} S_m(x, Z) \tag{1.5}$$

where $\left(\frac{d\sigma(\theta)}{\delta\Omega}\right)_{KN}$ is the Klein-Nishina cross section (equation 1.6), x the momentum transfer function (equation 1.11), Z the atomic number of the scattering atom and S_m the incoherent structure factor which corrects for atomic binding energies and have been tabulated for all elements (Hubbell and Overbo (1979)).

$$\left(\frac{d\sigma_{KN}(\theta)}{\delta\Omega}\right) = \frac{r_o^2}{2}[1+k(1-\cos\theta)]^{-2}\left[1+\cos^2\theta+\frac{k^2(1-\cos\theta)^2}{(1+k(1-\cos\theta))}\right] \quad (1.6)$$

where r_o is the classical electron radius and $k = \frac{h\nu'}{m_o c^2}$.

1.2.3 Coherent scatter

Coherent, or Rayleigh scatter occurs when the incident photon interacts with an electron that is tightly bound, so the whole atom recoils. The energy transferred to the atom is negligible so the photon is scattered without a change in energy, and therefore wavelength. The differential cross section for coherent scatter is given by equation 1.7.

$$\left(\frac{d\sigma_{coh}}{\delta\Omega}\right) = \left(\frac{d\sigma_T(\theta)}{\delta\Omega}\right)F_m^2(x,Z) \quad (1.7)$$

The differential Thompson cross section $\left(\frac{d\sigma_T(\theta)}{\delta\Omega}\right)$ gives the angular distribution of scattered photons.

$$\left(\frac{d\sigma_T(\theta)}{\delta\Omega}\right) = \frac{1}{2}r_o^2(1+\cos^2\theta) \quad (1.8)$$

The molecular form factor F_m depends on the distribution of the electrons in the scattering atoms, which can theoretically be calculated from the sum of the elemental atomic form factors f_i, that are related to the Fourier transform of the atomic charge density and have been tabulated for elements of atomic number 1-100 (Hubbell and Overbo (1979)), multiplied by the electronic abundance (n_i).

$$F_m^2 = \sum_i n_i f_i^2(x) \quad (1.9)$$

This is known as the free atom model (or Independent atomic model

(IAM)) and describes the experimental results well at large values of momentum transfer (see equation 1.11). However at low values of momentum transfer (< 5 nm^{-1}) interference effects due to the intramolecular and intermolecular electron interactions cause the scattering to become dependent on angle creating a diffraction pattern. These form factors cannot be calculated theoretically but have been measured experimentally for several materials (Narten and Levy (1971) Morin (1982), Kozanetsky et al. (1987), Peplow and Verghese (1998), Poletti et al. (2001), Johns and Wismayer (2004)).

The interaction cross sections of both coherent and incoherent scattering are dependent on both angle and photon energy. The scatter at small angles to the direction of the incident X-rays is predominantly coherent (Johns and Yaffe (1983), Neitzel et al. (1985)), with incoherent scatter becoming dominant at larger angles. It is those photons that are coherently scattered at small angles that are important in the technique of energy dispersive x-ray diffraction described in this work.

Photons that are coherently scattered in a material have a phase relationship and thus can interfere with each other. The interference effects that occur due to the scatter from materials can be constructive or destructive. The condition for constructive interference which produces a peak in a diffraction pattern is given by Bragg's Law

$$n\lambda = d \sin \frac{\theta}{2} \tag{1.10}$$

where n is the order of diffraction, λ the wavelength of the radiation, d the spacing between scattering planes in the material and θ the angle through which the photon is scattered.

Crystalline materials with long range order produce several sharp diffraction peaks, whereas amorphous materials, which have only short range order, produce one or more broad peaks. The shape, intensity and position of these peaks can be used to characterise the scattering material.

1.3 Energy Dispersive X-Ray Diffraction

X-ray diffraction has been used to probe the structure of materials since it was discovered by Laue in 1912, though traditionally a monochromatic beam is used and the scattered intensity measured as a function of angle. This technique is referred to as angular dispersive X-ray diffraction (ADXRD). The beam energies of such sources are typically low, such as 8.047 keV for the K_α line of Copper, produced by filtering a copper spectrum with nickel to remove the K_β characteristic lines and the low energy bremsstrahlung. Due to attenuation affects, radiation of this energy has only a small penetration depth, so although it can be used with small crystalline samples, it is not suitable for imaging larger objects. X-ray sources that produce higher energy X-rays, such as the tungsten target sources used in most diagnostic radiology, have polyenergetic spectra. A polyenergetic source cannot be used in conventional X-ray diffraction, as the conditions set by Bragg's law will be met for a range of angles leading to a blurring out of the diffraction peaks. Instead of the diffraction profile being measured using a single energy over a range of angles, it is equally valid to measure the range of photon energies detected at a fixed angle. This is the basis of energy dispersive X-ray diffraction, using a detector with good energy resolution to measure the diffraction profile of a material at a fixed angle, or a narrow angular range. The benefits of using coherently scattered X-rays in place of the transmitted ones are in the imaging of materials with very similar atomic numbers (Z). Transmission X-ray images can clearly distinguish between materials of high Z and low Z, but the contrast between media with similar Zs (e.g. fibroglandular tissue and carcinoma in mammography) is low. X-ray diffraction techniques can be used to examine the region of the diffraction spectrum where the differences between two media are maximised, allowing much higher contrasts to be achieved.

Many illicit drugs present pseudo-crystalline scattering centres, hence the electron distribution is periodic and thus constructive interference occurs

when the conditions of Bragg's law are met resulting in a series of sharp peaks as the wavelength of the radiation is changed. The position of each peak is governed by the spacing between the scattering planes (d), and the combination of position and intensities of all the peaks provides a unique pattern, or diffraction profile, that can be used to identify the substance.

Amorphous materials such as water and soft tissues do not have the long range order of crystalline structures, but have some short range order and produce broad diffraction peaks at low scatter angles.

In order to compare results from different systems, the diffraction profile is expressed as a function of momentum transfer, x, which depends on both the photon energy (E) and scatter angle (θ) as follows.

$$x = \frac{E}{hc} \sin \frac{\theta}{2} \tag{1.11}$$

where h is Planck's constant and c the speed of light. The ability of the system to resolve the peaks in momentum transfer space depends on the angular resolution of the diffraction cell and the energy resolution of the detector.

1.4 Previous work

Low angle scatter has been used for a wide variety of applications in fields as diverse as medicine and homeland security and for a range of industrial uses, some of which are outlined below.

1.4.1 Medical

Bone

Quantitative analysis of the constituents of trabecular bone using EDXRD has been carried out by Royle and Speller (1995). The diffraction spectra of dried bone and animal fat (a marrow equivalent) were found, and by comparing the relative intensities of the peaks from fat (1.9 nm^{-1}) and powdered bone (1.7 nm^{-1}) the proportions of the two tissue types in the sample could be calculated. This can be used to provide information on the osteoporotic state of the bone, as osteoporosis is caused by the loss of bone tissue which is replaced by adipose tissue, leading not only to a reduction in bone strength but also to a change in the ratio of the intensities of the bone and fat peaks. Coherent scatter is particularly suited to this application as often in osteoporotic patients the overall bone density is not necessarily much different from that of healthy subjects.

Batchelar et al. (2000, 2006) have used coherent scatter computed tomography (see section 1.4.1) to quantitatively analyse bone composition, by measuring the collagen-hydroxyapatite ratio in bone-mimicking phantoms and intact bone specimens. This is required in the detection and monitoring of metabolic bone diseases such as osteomalacia, which arises from a vitamin D deficiency and results in poorly mineralised bone tissue and other bone related investigations. Batchelar et al. (2006) found their results to be accurate to within 2%.

Farquharson et al. (1997) used bone phantoms consisting of a mix of hydroxyapatite powder and animal fat in a dural sleeve to simulate trabecular bone mineral loss and cortical bone thinning. They used multivariate calibration to build a model using the EDXRD data, which was then used to predict the hydroxyapatite content of nine test phantoms, which had an accuracy of $\pm 3\%$ for long measurement times (250s) which fell to $\pm 8\%$ for shorter measurements (5s).

Calculi

Caliculi are kidney and gall stones, the treatment of which depends upon the chemical composition of the stone. The established conventional method of analysis is by monoenergtic X-ray diffraction, though the low energies of these X-rays requires samples to be removed and powdered, rendering a pre-surgical examination impossible. This also destroys all knowledge of the structure and growth of the crystal, as they are typically deposited in concentric lamina around a central core, or nucleus. A powdered sample contains no information pertaining to the order of the layers or the composition of the nucleus, which provides knowledge of the genesis of the calculus and therefore is important in the prevention of reoccurrence. The development of a system able to deduce stone composition and structure in situ would aid both treatment planning and help prevent reoccurrence of urinary calculi. Dawson et al. (1996) used EDXRD to classify urinary stones, using a tungsten target X-ray source and a germanium detector. Batchelar et al. (2002) used a polychromatic diagnostic X-ray source to examine the seven most common components of urinary calculi, and found they could discriminate between them by their coherent scatter profiles. Davidson et al. (2005) have also investigated the possibility of investigating whole calculi tomographically based on the techniques suggested by Harding et al. (1990), and have produced images of a stone-mimicking phantom. Fong (2002) used a combination of EDXRD data from known samples (30 powdered samples and one urinary stone of known composition) and multivariate analysis to identify the composition of 51 stones of unknown composition.

Mammography

Breast cancer is by far the most frequent cancer amongst women with an estimated 1.15 million cases diagnosed worldwide annually (Parkin et al. (2005)). Early in the disease most breast cancers are asymptomatic. When

breast cancer is detected in the localized stage the 5-year survival rate is 97%. If the cancer has spread regionally to underarm (axillary) lymph nodes, the rate drops to 79% and if the cancer has spread (metastasized) to distant organs such as the lungs, bone marrow, liver, or brain, the 5-year survival rate drops to only 23% (Greenlee et al. (2001)). Thus early detection is important and affects the prognosis of the patient. Currently mammography can detect tumours which are about 10mm in diameter, and smaller micro-calcifications, whose pattern can be indicative of the tissue pathology. The current National Health Service breast screening programme (NHSBSP) calls women between the ages of 50-64 for a screen every three years, with older women able to be screened on request. Since 2003, two views of each breast have been taken, mediolateral oblique and craniocaudal, and given that the number of women screened in the two year period 2001-2002 was 1.5 million (Young (2005)), the number of mammographic images taken in the UK can be estimated at over 3 million per annum.
Mammography used in screening currently has a sensitivity of 40-95% depending on the stage of the cancer, a specificity of 95% (Harris and Vogel (1997)) and can be expected to reduce mortality by about 30% (Tabar et al. (2000)).

Mammography is traditionally conducted using a screen film combination, a molybdenum or rhodium X-ray source with molybdenum, rhodium or aluminium filtration. Mammographic film detects the transmitted X-rays, and therefore produces a map of the attenuation coefficients of the tissue the X-rays have passed through, with highly attenuating tissues producing regions of lower optical density on the film.

However, the linear attenuation coefficients of tissue types found in the breast are all similar, leading to low contrast in transmission images. This can be seen from figure 1.1, which indicates that the maximum contrast achievable between fibrous tissue and carcinoma at the standard mammo-

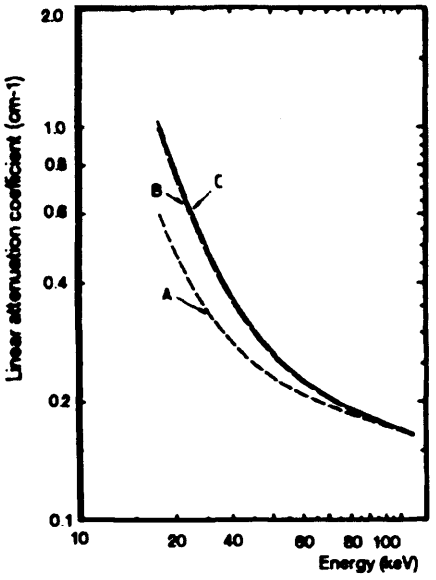


Figure 1.1: Mean values of the linear attenuation coefficients of the three main types of breast tissue: A) fat, B) Fibrous and C) Infiltrating ductal carcinoma (Johns and Yaffe (1987))

graphic energy of 18keV is only 5% (Johns and Yaffe (1987)). This makes it especially difficult to detect lesions in dense breasts, which contain a high proportion of fibroglandular tissue.

X-rays scattered in the tissue can cause an additional source of noise, which further reduces the SNR of the image, though it can be minimised using a suitable grid, or other method of scatter reduction or rejection. These methods have the disadvantage that they also absorb some primary photons and therefore the number of photons incident on the patient must be increased in order to produce an image of good statistical quality. The dose to the patient is thus increased, which is undesirable as breast tissue is radiosensitive and cumulative dose can lead to an increased risk of developing a malignancy. The ratio of scattered radiation to primary radiation can be as high as 1 (Boone and Cooper (2000)), however the X-rays that are coherently

scattered by the tissue can be used to provide additional information about its structure (Speller and Horrocks (1992)). This arises as different molecular arrangements produce diffraction peaks at different values of momentum transfer, and thus by measuring the diffracted signal at small angles as well as the transmitted signal, valuable information can be obtained.

Work has previously been carried out to investigate the diffraction profiles of breast tissue samples, tissue equivalent materials and various animal tissues.

Kozanetsky et al. (1987) measured the differential cross sections of water, some plastics and biological samples, using monochromatic radiation (6.935keV) and a powder diffractometer. They found that adipose tissue has a sharp diffraction peak, and hypothesized that it is due to the high degree of order in the alignment of the fat cells contained in it. Peplow and Verghese (1998) determined and tabulated the molecular coherent scattering form factors, of some plastics, water, pork and beef tissues and fixed breast tissue using monochromatic synchrotron radiation. Evans et al. (1991) examined water, olive oil, blood, Perspex and nineteen breast tissue samples, and recorded the position of the main peak of each.

The linear differential scattering coefficients of various materials have been measured by Kidane et al. (1999), who examined 100 breast tissue samples, both normal and neoplastic, and by Poletti et al. (2001), who studied adipose and glandular breast tissue, water, some plastics and four breast equivalent materials. Both authors found that the momentum transfer values at which adipose and fibrous tissues most strongly diffract are 1.1 nm^{-1} and 1.6 nm^{-1}, respectively.

Cancerous tissue was shown to produce diffraction peaks more characteristic of fibroglandular tissue, with a peak at the momentum transfer value of 1.6 nm^{-1} and a low recorded scatter signature at 1.1 nm^{-1} indicating a low

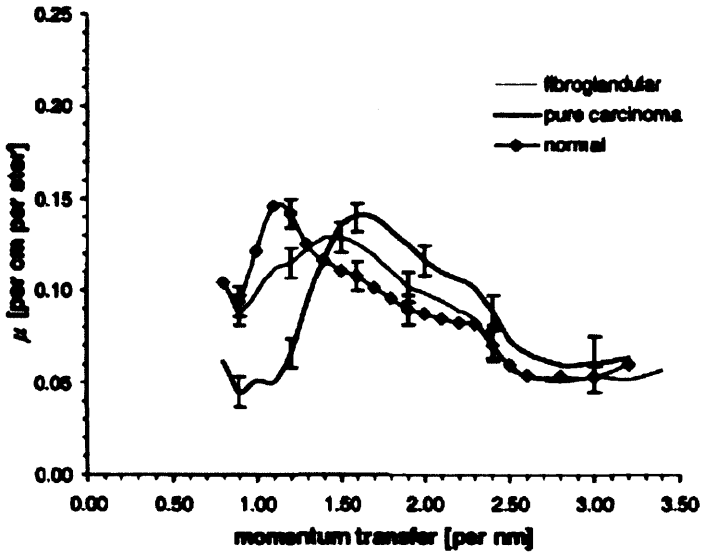


Figure 1.2: Differential linear coherent scattering coefficients for pure carcinoma, normal healthy breast tissue and healthy fibroglandular tissue, from Kidane et al. (1999)

fat content. Harris et al. (2003) used a low light level CCD with synchrotron radiation to obtain both transmission and diffraction (at 1.1 nm^{-1}) images of a phantom made from healthy breast tissue with cancerous inclusions in a perspex holder, finding a contrast improvement of 60-100% between the transmission and diffraction images.

The diffraction peak for water also occurs at 1.6 nm^{-1} (Evans et al. (1991)) and therefore so does that of blood, which contains 90% water. Changizi et al. (2005) extended this to find the peak heights and positions of adipose and fibroglandular tissue, carcinoma and tissue with benign fibrocystic changes. They found the peaks of these tissue types to be at 1.09, 1.25, 1.53 and 1.38 nm^{-1} respectively, indicating that EDXRD as an imaging modality can be used to differentiate not only between healthy and diseased tissue, but also between benign and malignant neoplasms. This increased specificity could be used in mammography to reduce the number of

false positive results from mammograms and thus the number of unnecessary biopsies performed on women.

The differences in the SAXS patterns between diseased and healthy breast tissue biopsy samples were measured with synchrotron radiation over the range of momentum transfer values 0.08 to 0.33 nm^{-1} were analysed using multivariate analysis by Round et al. (2005), and the combination of techniques shown to classify the tissues well, though with a significant 'rogue sample' and some doubt as to how many principle components should be used. Multivariate analysis was also used by Ryan and Farquharson (2007) who used 18 variables extracted from both coherent and compton scatter data to classify tissue samples into the groups fibroadenoma, malignant, fibrocystic change, normal fibrous and adipose. The best model had a sensitivity of 54% and a specificity of 100%, but was unable to distinguish between the different fibrous tissue types.

Other Biological Materials

Low angle x-ray scattering (LAXS) from some biological samples (water, muscle, fat and blood) was modeled by Elshemay et al. (1999), who found that although the position of the peaks could not be used to characterise tissues of similar water contents (water, muscle and blood), the FWHM of the peaks of each sample showed clear differences. Desouky et al. (2001) continued this by examining the LAXS profiles of a range of biological samples (including red blood cells, haemoglobin, fat, muscle, albumin and DNA), finding that protein rich tissues produced an additional characteristic peak that could be used to monitor protein levels in tissue. Changes in the structure of the protein alters the shape, height and position of the peak and has been suggested as a method for detecting hepatocellular carcinoma using the LAXS profiles of serum, as all serum proteins are produced in the liver (Elshemey et al. (2003)). Lazarev et al. (2000) examined samples of

breast, brain and prostate tissue using a diffractometer and found differences between normal and pathological samples.

There is also some discussion as to whether or not the X-ray diffraction patterns of hair can be used as an early indicator of breast cancer as the presence of a diffuse ring in the pattern has been found to provide 100% sensitivity to the disease (James et al (1999)). However, other authors have been unable to replicate these results (Briki et al. (1999), Laaziri et al. (2002), Howell et al. (2000) and Meyer et al. (2000)), though Meyer later reevaluated his results and consequently found a correlation between the appearance of the ring and the presence of breast cancer (Meyer and James (2001)). Hair diffraction has also been suggested as a possible tool to screen for insulin dependent diabetes (James et al. (1997)) and colon cancer (James (2003d)).

Computed Tomography

Several workers have investigated the field of EDXRD computed tomography (Harding et al. (1999), Kleuker et al. (1998), Batchelar et al. (2002) and Davidson et al. (2005b)) to examine urinary stones. This uses 1st generation CT geometry, but collects the X-rays scattered at small angles as opposed to those transmitted. The advantages of this technique over transmission CT is that contrast between different materials varies as a function of momentum transfer so a suitable angle can be selected to best differentiate between the media present.

Another use of X-ray diffraction is in micro CT (Barroso et al. (2001)), which can be used to evaluate the structure of tumours. It can also be used as a means of extracting the diffraction profiles of single tissue types to reduce the error due to overlapping tissues and has been suggested as an appropriate method to examine tissue biopsies due to its relative speed compared with histopathological analysis. Griffiths et al. (2004) analysed excised fresh breast tissue samples using both synchrotron radiation and both transmission and diffraction CT techniques, finding an average increase

in contrast of 25% in the diffraction CT image compared to the transmission CT. The technique was adapted to a laboratory X-ray system (Griffiths et al. (2007)) on which 19 fresh samples were examined and classified into three types: fat, tumour and connective tissue. By examining the ratio of the counts in the regions 1.1 nm^{-1} and 1.5 nm^{-1} it was possible to discrimiinate between the three tissue types. Pani et al. (2006) used EDXRD-CT (both synchrotron and laboratory based) to examine 11 fresh breast tissue samples and extract the diffraction signatures from regions of interest, eliminating the risk of superposition of tissues inherent to transmission imaging. Five distinct patterns were found, indicating that the samples contained 5 different tissue types, though these tissue types have not yet been identified by a pathologist.

1.4.2 Industrial

Food industry

It has been suggested that EDXRD could be used in the food industry as a means of checking for contaminants (Luggar and Gilboy (1999)). Transmission X-ray images can be used to identify dense impurities such as metal or stone, but adulterants with a low Z are difficult to detect. As much of the packaging of foods uses low Z plastics, these are a major contaminant and a possible method of detecting them is using the coherently scattered photons. The detection of nylon inclusions in chocolate can be improved using scatter imaging of 22keV X-rays at 8.5°. The contrast between the regions with and without inclusions is 8% when considering the linear attenuation coefficients compared to 95% when using the aforemention scatter parameters.

Multiphase flow measurements

Another industrial use of EDXRD is in the metering of oil and water pumped from oil reservoirs (Luggar and Gilboy(1999)). The ratio of the peak scattered intensities of oil (S_1 at 0.7 nm^{-1}) and water (S_2 at 1.6 nm^{-1}) normal-

ized by the transmitted X-ray intensities (T_1 and T_2) gives the ratio (R) of oil:water in the emulsion.

$$R = \frac{S_1}{S_2}\frac{T_1}{T_2} \tag{1.12}$$

This technique for the multiphase flow measurement of water content in an oil/water mixture shows promise for the future. The width of the peaks of oil and water also allows CdZnTe detectors to be used, which do not have as high an energy resolution as germanium detectors, but have a higher bandgap (1.6eV compared to 0.67eV for germanium) thus have the benefit of not requiring cryogenic cooling as they have a leakage current of only a few nA at room temperature, making them more suitable for industrial use (see section 7.2.3).

1.4.3 Security

Explosives

EDXRD has been suggested as a technique suitable for the detection of illicit materials in airport security. This includes the examination of suitcases for explosives and parcels and packages for drugs. The diffraction profiles of several explosives have been measured (Speller et al. (1993)), and using multivariate analysis the principle components of the spectra were identified (Luggar et al. (1998)) and a system calibrated to detect semtex in a case containing cotton and wool. Hamburg airport had a system in operation (RayScan, Philips, Strecker (1998)) that used a High Purity germanium detector with its germanium crystals in an annular arrangement to simultaneously scan voxels at different depths in the baggage. The profiles measured were compared to a library of spectra and any matches indicated on the display, with a score indicating the match between the measured spectra and the library profile. Jupp et al.(2000) used an ADXRD system, using balanced filters to define an energy window from a tungsten anode X-ray spectrum, an image intensifier and CCD, together with single value decomposition maxi-

mum likelihood analysis to predict the composition of samples with the aim of identifying explosives in baggage. Plasticene was used as an explosive substitute and was correctly identified in 100% of cases, though with some false positive results from other substances. CdZnTe detectors have also been used in the detection of sheet explosives in baggage (Malden and Speller (2000)), and summing the counts in a small window around an atomic spacing of $3.3\mathring{A}$ was found to provide a positive or negative result in the detection of the exposives semtex and SX2(semtex type 2) in cases containing common suitcase contents, substances known to be strongly scattering and materials difficult to distinguish from explosives in standard transmission scanners.

Narcotics

Due to the crystalline structure of some illicit drugs, they also produce diffraction patterns with sharp peaks, thus X-ray diffraction is a promising technique to detect and identify narcotics (Harding and Schreiber (1999)). The spectra of Heroin and Cocaine (types unspecified) have been taken using the prototype system for explosives detection described in section 1.4.3. Traditionally X-ray diffraction is carried out using a monochromatic beam, and peaks can be seen over the full range of scatter angles. X-ray powder diffractometers (e.g. Debye-Scherrer camera,) use a thin collimated beam directed at a polycrystalline sample, which produces diffraction rings due to the crystals in the sample having all possible orientations. Using such systems, a database of diffraction patterns of known samples has been built up and currently contains over 118,000 entries. This Power Diffraction File (PDF), collated by the Joint Committee on Powder Diffraction Studies (JCPDS), lists the d-spacings of each substance and is used to identify the components of unknown samples by the three most intense peaks in their diffraction spectra. X-ray diffraction is widely used in forensic science as it is a non-destructive testing method and can identify a substance from a very small quantity. It is employed by the Metropolitan Police in drugs analysis to identify the precise

chemical form of the drug (salt, base or acids), to identify any adulterants, and to compare seizures (Rendle (2003)). However, powder diffractometry requires small samples prepared in glass capillary tubes, whereas EDXRD can be used to detect samples in a range of containers and packaging.

1.4.4 Summary

Although the literature presents a range of diffraction data taken from a variety of scenarios, the above review indicates some areas that need to be investigated in more depth.

The energy dispersive diffraction profiles of heroin and cocaine have been measured (Harding and Schreiber (1999)), though the forms of the drugs (free base or salt) were not stated, and data from further narcotics not presented. A useful drug detection system would need to cover as many illicit substances as possible, including a range of class A drugs such as Ecstasy (MDMA) and Speed (amphetamine sulphate), as well as legal substances, thus further measurements need to be made. A technique to analyse the profiles also needs to be considered, as it has been in the detection of explosives, and tested on measured diffraction data.

In the case of mammographic applications, it has been shown that the X-ray diffraction of breast tissue produces profiles which are characteristic of tissue pathology (Kidane et al. (1999)). Measurements on small samples and phantoms using EDXRD and EDXRDCT have produced promising results, offering an improvement in contrast between cancerous and healthy tissue compared to X-ray transmission methods. The imaging capabilities of a diffraction system needs to be fully investigated, and the momentum transfer region, or combination of regions, that gives the optimal information, found. Whether formalin fixation of tissue for preservation affects the measured diffraction profiles is also unclear, with conflicting results published in the literature which are discussed later (see chapter 4).

1.5 Focus of this work

This study expands the work mentioned in the summary above. It focuses on the analysis of diffraction profiles for use in two applications; the identification of illicit drugs and the detection of breast cancer.

Measurements of the diffraction profiles of a range of narcotics are presented, along with some common cutting agents and packaging materials, all of which produce characteristic EDXRD profiles. The high degree of order in the structure of illicit drugs makes the peaks in the measured profiles sharp and numerous, thus requiring complex analysis. Multivariate analysis is used to identify the principle components of the measured profiles and regression models for the prediction of future unknown profiles built and tested. The effects of the system geometry and counting statistics on the accuracy of the analysis are also investigated, and a model built to simulate profiles from angular dispersive data to predict the measured profiles for different geometries.

Breast tissue type can be determined using EDXRD as adipose, glandular and cancerous tissue have distinctly different diffraction profiles (see section 1.4.1). The imaging capability of diffraction information is investigated and the regions of the profile used optimised to produce images with the maximum possible contrast between healthy and cancerous tissue. These images are corrected for attenuation effects and compared to the transmission images created by attenuation alone. The application of diffraction data to an intelligent mammographic imaging system is also considered, and functions based on the diffraction data compared to other possible 'intelligent' algorithms. The effects of formalin fixation on the diffraction profiles of tissue is investigated, as this has implications for tissue storage and comparisons between previous authors' work.

Chapter 2

EDXRD System

Energy dispersive X-ray diffraction (EDXRD) utilises the X-ray photons coherently scattered through small angles in a medium. This chapter outlines the equipment used to produce, collimate and detect these photons and some of the methods used to analyse the diffraction profiles obtained.

2.1 EDXRD System Design

2.1.1 X-ray Source

The source used was a tungsten target industrial x-ray tube (AGO installations, Philips), with a maximum output potential of 125kV and 30mA. It was generally operated at a potential no higher than 70kV to avoid the characteristic K-lines of tungsten becoming significant, and 15mA as the source was used in fluoroscopy mode for long periods of time.

2.1.2 Detector and MCAs

The scattered photons were detected using a high purity germanium detector with a planar crystal of diameter 26mm and depth 10mm (model GLP16/95/16PS, EG& G Ortec, Oakridge, TN, USA). The detected spectra for 2D scans were analysed using an MCA with change sample and sample ready connections (921 EG& G Ortec) whilst the static sample measurements used a computer based MCA card (TRUMP PCI, EG& G Ortec). The PC software (Maestro, EG & G Ortec) was used and the detected photons assigned to one of 512 energy channels. The system was calibrated using an Americium-241 source and had an energy resolution of 0.54keV at 59.5keV.

2.1.3 Collimation Design

Primary beam collimation

The collimation of both the incident beam and the scattered photons was provided by a purpose built brass diffraction cell. The primary collimation was provided by 20mm long brass plugs that slotted into the first plate of the cell and were available in a range of diameters: 0.5-2mm, providing pseudo-pencil beams of X-rays.

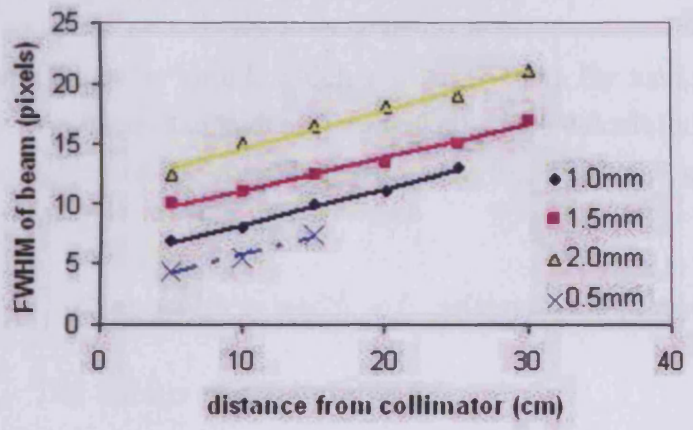


Figure 2.1: Plot of width of beam (in pixels) against distance of CCD from collimators of widths 0.5, 1.0, 1.5 and 2.0 mm.

The nominal focal spot size of the X-ray source was 3mm x 3mm and was approximately 30cm away from the primary collimator. The divergence of the beam was measured using a CCD camera (Grundig FA87 with a GADOX screen and a lens system) to image the beam at various distances from the collimator. The width, in pixels, of the beam in each position was measured and a linear fit acquired from the plot of width against the distance of the front of the camera from the collimator (figure 2.1). The width was measured where the intensity dropped to 10% of the maximum intensity. By extrapolating the linear fit, and finding the point where the widths of the beams are in the correct ratios, the distance at which the beam size equals the slit

Plug diameter (mm)	Gradient	Intercept	Calculated Pixel width (mm)	Divergence (°) θ_{div}
0.5	0.4	1.6	**	0.34
1.0	0.42	3.32	0.301	0.37
1.5	0.48	4.89	0.307	0.42
2.0	0.55	6.60	0.304	0.48

Table 2.1: Data from linear fit of plots of beam width against CCD distance from collimator.** data not included as insufficient number of points due to low beam intensities at large collimator-detector distances

width can be found, which was at -7.1cm. By using this value in the linear fit equations the width of a pixel could be calculated. This can then be used to calculate the angular divergence θ_{div} from the gradient of the linear fit (table 2.1).

$$Beamwidth = width\ \ of\ \ collimator + (2tan\theta_{div} \times\ \ Distance)$$

The data is shown in table 2.1.

Scatter collimation

Three designs for scatter collimation were considered and are described in the following sections. The angular resolution of the system was governed by the cell and is given by the relation

$$\frac{\Delta\theta}{\theta} = \frac{\theta_{max} - \theta_{min}}{\theta} \tag{2.1}$$

where θ is the nominal scatter angle, θ_{max} the maximum angle through which a photon can be scattered and still be detected, and θ_{min} the minimum angle.

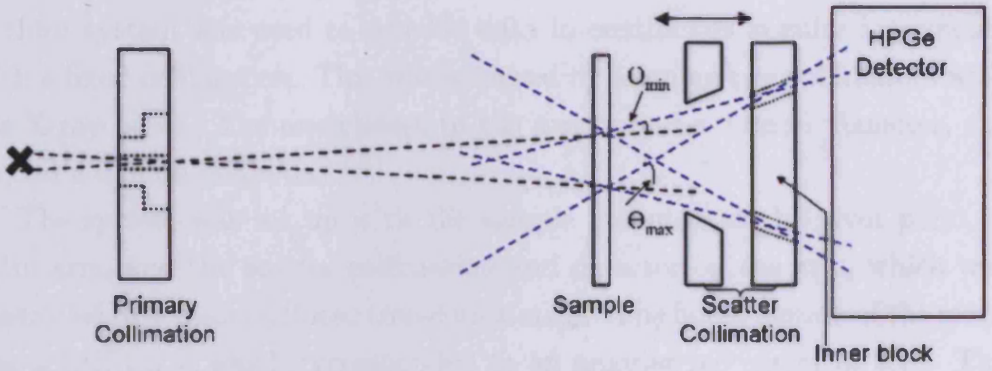


Figure 2.2: Diffraction cell design 1

Diffraction Cell design 1

Cell design 1 consisted of the primary collimator and two plates making up the scatter collimation. The middle plate had a central hole of diameter 10mm and could be moved parallel to the x-ray beam thereby altering the scatter angle. The final plate had a 26mm diameter hole and an inner block to make an annular acceptance ring. A range of inner blocks were available in a range of diameters and were used to alter the angular acceptance of the collimation.

Diffraction Cell design 2

Cell design 2 used the same primary collimation as design 1 but the scatter collimation was provided by a choice of brass blocks with removable conical inserts at a range of angles (4-7°) that were spaced from the block with thin wires. Using wires with larger diameters increased the angular acceptance of the scatter collimation.

Moving arm system

A third system was used to provide data in continuous angular increments with a fixed collimation. This was achieved by aligning two collimators with the X-ray beam. The one closest to the source had a 1.0mm diameter, the second a 0.5mm diameter.

The system was set up with the sample mounted at the pivot point of a 1m arm, and the scatter collimation and detector on the arm, which was controlled by a manual linear translation stage. The finest motion of the stage was a half turn, which corresponded to an angular increment of 0.1°. The collimation was aligned using the CCD camera method described in section 2.1.3. The angular acceptance of the scatter collimation system could be increased by widening the aperture of the pinhole or moving the pinhole and detector closer to the sample. The pinhole was a plug based on the same design as the primary collimator so plugs with holes of differing diameters could easily be exchanged.

Motion control and alignment

Cell designs one and two were mounted on a series of computer controlled stages, allowing linear (Parker, UK) and rotational (Edmund optics, UK) motion using actuators (Zaber inc., Canada) and goniometrical motion (Ko-

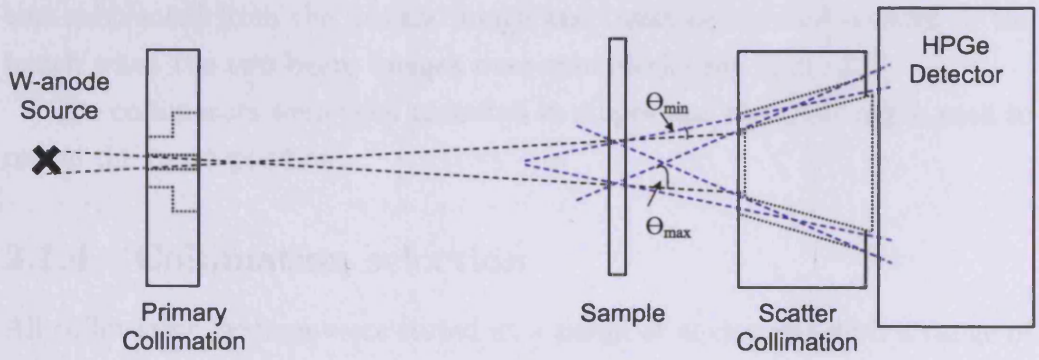


Figure 2.3: Diffraction cell design 2

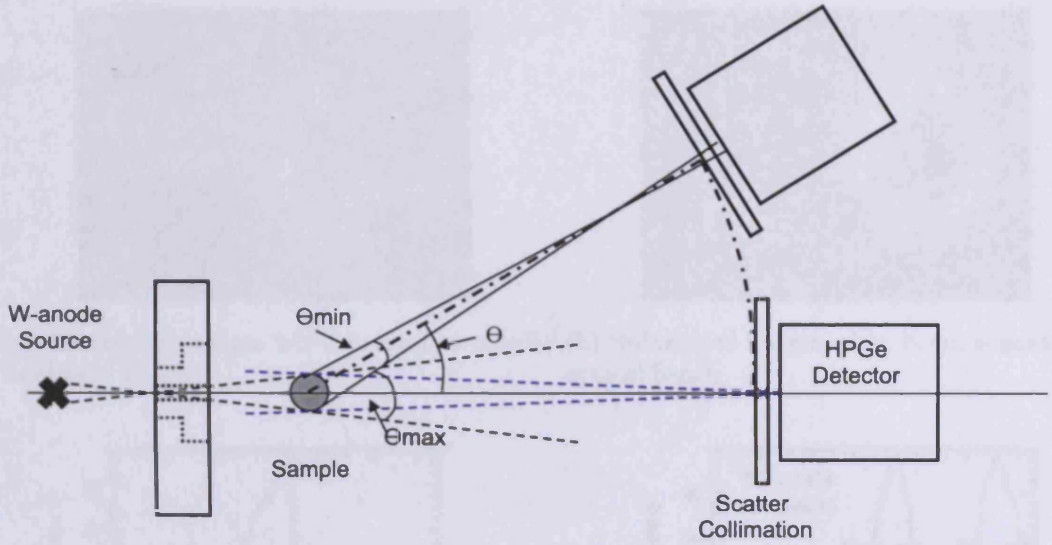


Figure 2.4: Schematic of Moving Arm Diffraction System (plan view)

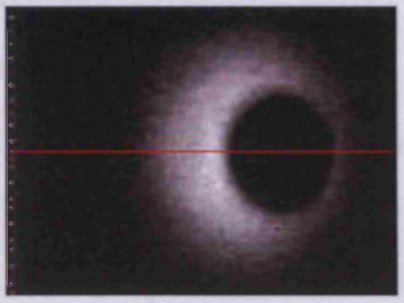
hzu, Japan) using a connector block, motion control PCI card and software (National Instruments, UK). Alignment was assisted using the CCD camera to ensure that the centre of the X-ray beam corresponded to the central axis of the cell. In the moving arm system the primary collimator was mounted on linear stage under PC control via an actuator as above.

First the X-ray beam was adjusted so that it was parallel to the optical bench. This was done by mounting the CCD on a rail and obtaining images of the beam at distances of 25cm and 75cm from the source. The near image was subtracted from the distant image and beam considered parallel to the bench when the two beam images were concentric (see figure 2.5).

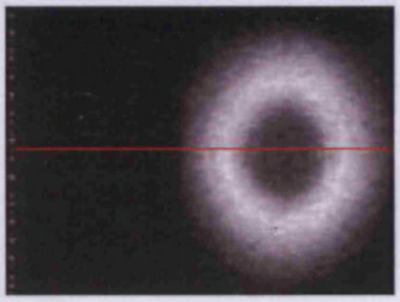
The collimators were then mounted in stages and the CCD again used to image the beam position.

2.1.4 Collimation selection

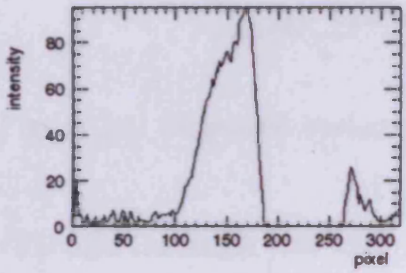
All collimation systems were tested at a range of angles and with a range of incident beam diameters.



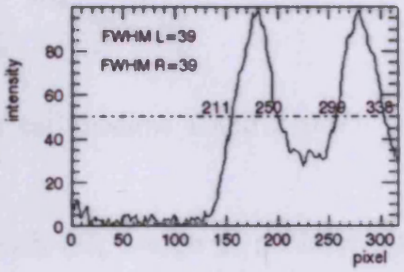
(a) Subtracted images when beam not parallel to optical bench



(b) Subtracted images when beam is parallel to optical bench



(c) Profile through a)



(d) Profile through b)

Figure 2.5: System alignment using CCD camera

Each system was tested in respect to the criteria of amount of background scatter, ease of altering the angle and resolution of the system and the counting times required to get good statistical data.

Cell design 1, with the variable geometry due to the ability to move the second plate, allowed the nominal diffraction angle to be altered easily and with the sample in the calculated position gave peak positions that agreed with theoretical calculations. However the incident plug diameter was limited to below 1mm as at larger dimensions photons could pass through the system without scattering. The acceptance angle of the system could, in theory, be altered by changing the inserts on the third plate, though it varied with the scatter angle, which became principally defined by the position of the sample rather than that of the second plate. This made it difficult to

make a fair comparison of scatter angles, thereby defeating the point of the movable plate design.

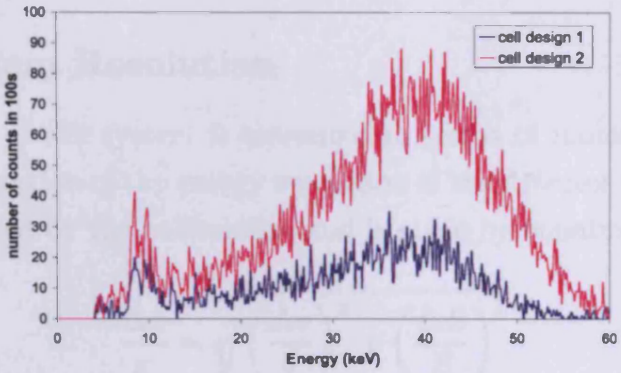


Figure 2.6: Measured background counts for cell designs 1 and 2 at 7°

Though the angle was less easily altered with cell design 2, no incident radiation could reach the detector without scattering and the scattering volume was better defined. The spacing between the conical insert and the block was also easily changed using wire spacers of varying diameters thus producing a system that was easily adapted and therefore suitable for this study.

However, when cell design 1 was correctly adjusted, the background scatter with no sample or sample holder in place was less that detected with cell design 2 (see figure 2.6). This may be due to the internal X-ray scatter within the brass block reaching the detector. This suggests that when all required parameters are identified, a system similar to cell design 1 (i.e. an air gap as opposed to a solid block) is preferable.

The moving arm system provided the least background counts, but as it measured only a single point rather than an annular ring, the majority of the photons scattered through the required angle were not detected leading

to an inefficient system requiring long integration times to collect data with adequate counting statistics. Thus this system was used when small angular steps were required, otherwise cell design 2 was selected.

2.1.5 System Resolution

The resolution of the system is measured in terms of momentum transfer. This is a combination of the energy resolution of the detector and the angular resolution defined by the collimation and is given by equation 2.2

$$\frac{\Delta x}{x} = \sqrt{\left(\frac{\Delta\theta}{\theta}\right)^2 + \left(\frac{\Delta E}{E}\right)^2} \tag{2.2}$$

The energy resolution of the germanium detector is very good (540 eV at 59.5 keV, giving $\frac{\Delta E}{E} = 9.1 \times 10^{-3}$).

The angular resolution can be calculated geometrically. All quantities are defined in figure 2.7.

$$\theta_{min} = tan^{-1}\left[\frac{L_2 tan(\theta) - \frac{d_2}{2}cos(\theta) - \frac{h}{2}}{L_2 + \frac{d_2}{2}sin(\theta)}\right] - \theta_{div} \tag{2.3}$$

and

$$\theta_{max} = tan^{-1}\left[\frac{L_2 tan(\theta) + \frac{d_2}{2}cos(\theta) + \frac{h}{2}}{L_2 - \frac{d_2}{2}sin(\theta)}\right] + \theta_{div} \tag{2.4}$$

where

$$h = d_1 + 2L_1 tan(\theta_{div}) \tag{2.5}$$

Thus for the moving arm system with a primary collimator plug diameter 1.0mm, a scatter collimator plug of diameter 0.9mm, a primary collimator to sample distance of 6cm and a sample to scatter collimation distance of 28cm, for a nominal angle of 5.5°, $\frac{\Delta\theta}{\theta}$ would be (5.89-5.11)/5.5 which is 0.14. This is approximately 15 times greater than the energy resolution of the detector stated above thus it is the angular acceptances of the collimators

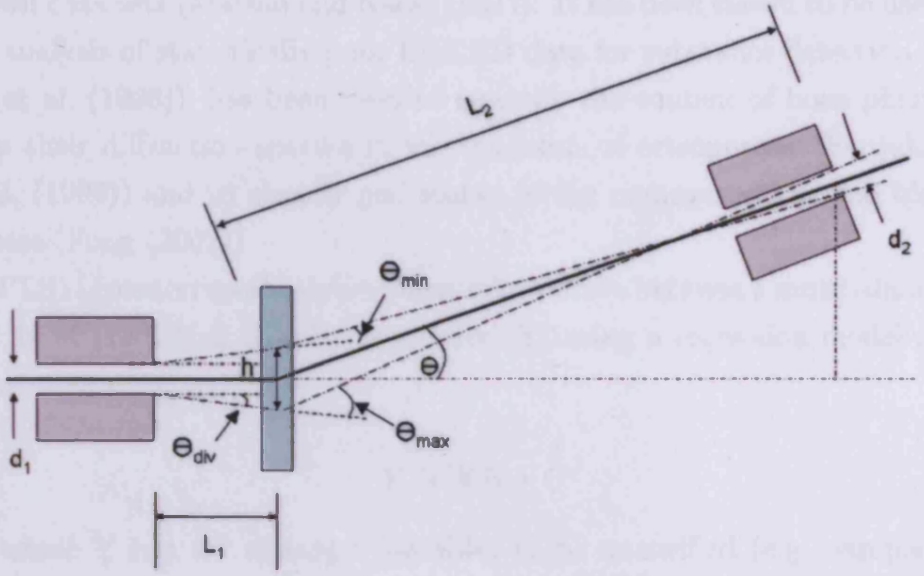


Figure 2.7: Schematic of general diffraction system

and the collimator-sample-detector distances that dominate the resolution of the system.

The angular resolution quoted here describes the maximum range of angles that a photon can be scattered through and enter the detector. In truth not all scatter angles will be equally probable, with those at the extremities being less likely. This produces a gaussian-like distribution of angles (see later section 3.5), the outer edges of which are given by the calculations above, though the FWHM of the distribution will be lower.

2.2 Multivariate analysis

The technique used to analyse the spectra was multivariate partial least squares (PLS) regression, which determines the principle components of sets of known data and uses this information to predict the response from un-

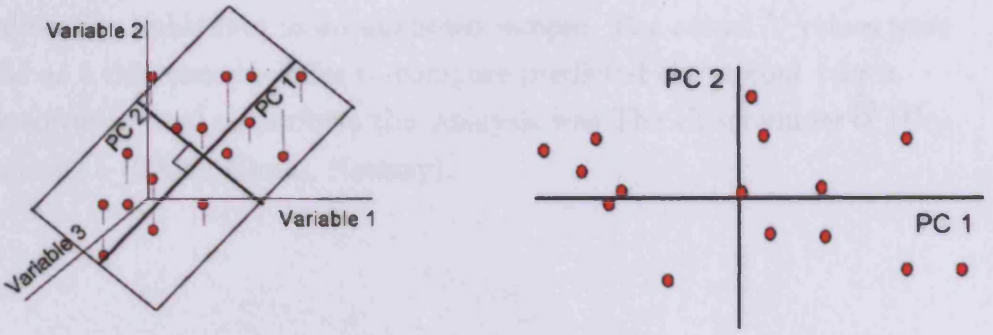
known data sets (Marten and Naes (1991)). It has been shown to be useful in the analysis of statistically poor EDXRD data for substance detection (Luggar et al. (1998)), has been used to quantify the content of bone phantoms from their diffraction spectra in the diagnosis of osteoporosis (Farquharson et al. (1997)) and to classify gall stones in the management of gall bladder disease (Fong (2002)).

PLS regression establishes a linear relationship between a multi -dimensional matrix of predictors (X) and responses (Y) using a regression model of the form

$$Y = XB + C \tag{2.6}$$

where Y is a set of target variables to be quantified (e.g. sample concentration), X a set of many variables (e.g. counts in each detector channel from the measured diffraction profile of a sample), B a regression coefficient matrix and C a matrix of residuals, the data that cannot be explained by the model. The regression coefficients B are determined by the process of calibration, where known X and Y data are used to build a mathematical model.

The X matrix in the calibration set used for this work contained 512 variables, one for each energy channel. If each x variable was used, the model would be extremely complex, inaccurate and would require a large computation time. Instead, the x-variables are reduced to their principle components (PCs), where each PC is linearly related to the x-variables. Principle components are new variables that describe the main difference between objects (samples). The first PC describes the greatest variation between objects, the second PC the next greatest and so on (see figure 2.8(a) for an illustration). There can be as many PCs as there are x variables, though after the main differences have been described, further PCs begin to describe only noise, which is undesirable in a model and leads to inaccurate predictions. A suitable number of PCs is selected by taking the number that produces the first



(a) Finding the principle components. The dotted lines give the residual of each object

(b) A scores plot of PC2 against PC1

Figure 2.8: An illustration of principle components analysis of objects with 3 x-variables

local minimum in the residual variance, so that the data is well explained but noise is not modeled. The residual variance is a measure of how different the objects are from their modeled behavior.

The scores plots (figure 2.8(b)) are plots of the values of each sample projected onto the plane described by the principle components. These plots are useful for examining groups in the data, seeing how different variables affect the principle components and in looking for outliers, samples that are substantially different from the others.

The Y matrix in this work contained the percentage of each of the calibration substances found in each sample. For example, if the model was built with samples containing 12 substances in various proportions, the model would have 12 y variables.

In PLS, Principle components analysis (PCA) is used to extract the information from the X matrix that is related to the Y matrix, with the aim of maximising the covariance between X and Y (Marten and Naes (1991)).

A set of X and Y data were used as a calibration set to train the program. The software was then used to predict Y values from a previously unseen set

of X values, thus giving a quantitative estimate of the relative amounts of each calibration substance in an unknown sample. The actual Y values were available as a reference in order to compare predicted and actual values.

The software used to perform this analysis was The Unscrambler® (Unscrambler v9.5 (2006), Camo, Norway).

Chapter 3

Illicit Drugs Identification

This chapter describes the characterisation of illicit drugs using their measured EDXRD profiles, the use of multivariate analysis to identify the presence of a drug by its diffraction profile and the optimisation of the geometry of the EDXRD system collimation for drugs detection. Predictions of the content of blind samples are made using a combination of these techniques and the compromise between counting statistics and momentum transfer resolution considered.

3.1 Introduction

Illicit drugs enter the United Kingdom through a variety of routes, a major one being the postal system. Given that thousands of parcels from abroad pass through both Royal Mail and private sorting offices daily, manual searching of all, or even a significant proportion, of them, is not viable. An automated non-destructive testing method is required, which can rapidly identify suspicious parcels requiring further investigation. Energy dispersive X-ray diffraction (EDXRD) is a suitable technique as the shapes and positions of diffraction peaks in the measured diffraction profiles are characteristic of the scattering substance and it uses X-rays of sufficiently high energy to penetrate parcels containing other attenuating materials. A fully operational scanning system requires a high throughout of parcels, therefore the time taken to collect and analyse the diffraction spectra must be minimised. The data collection time can be reduced by optimising the geometry of the system to collect as many scattered photons as possible whilst maintaining a sufficiently high momentum transfer resolution to allow the spectra from different substances to be clearly identified. Data collection time can also be reduced by altering the source parameters or by reducing the time spent collecting data at each point in the package. Reducing the counting time leads to fewer overall photons being detected and therefore to statistically noisy data. Multivariate analysis has been shown to be able to identify the

presence of a substance from an EDXRD spectrum, even with statistically poor data (Luggar et al. (1998)), and so is used in this study to identify the presence of illicit drugs.

3.2 Experimental Methods

The X-ray source (section 2.1.1) was operated for these experiments at 70kV, as this gave an adequate number of photons over the energy region required whilst being sufficiently low to prevent the K-lines from the tungsten target becoming prominent. All readings were taken with a tube current of 10mA.

The collimation used was either cell design 2 (figure 2.3) or the moving arm system (figure 2.4) as specified. Scatter angle selection is discussed in section 3.3.1 and the maximum angular acceptance investigated in section 3.4.2.

3.2.1 Samples

Substances investigated

All samples were prepared and supplied by the Home Office Scientific Development Branch (HOSDB). The illicit drugs investigated were: amphetamine sulphate, cocaine (free base), cocaine (hydrochloride), heroin (free base), heroin (hydrochloride), MDMA and methamphetamine hydrochloride. Five common cutting agents (caffeine, glucose, lignocaine, paracetamol and phenacetin) were also investigated. A summary of the physical properties of the samples used along with the peaks used to identify them are contained in table 3.1, along with information about some of the packaging materials investigated (see section 3.14).

With the exception of MDMA, which was only available as a 100% pure sample, samples of each drug were investigated at a range of purities 100%, 80%, 60% 40% and 20% by mass, adulterated with caffeine in each case.

Further samples of each drug with purities of 60% were examined, with the remaining 40% consisting of one or two of the cutting agents commonly found with that drug, in either 40% or 20% concentrations. This gave 62 samples in total, each with one, two or three constituents from 12 possible substances.

Substance	Density (g/cm^3)	Chemical Formula	d-spacings(relative intensity)		
			$D_1(\AA)$	$D_2(\AA)$	$D_3(\AA)$
AmphetamineSO_4	1.15	$C_9H_{13}N.SO_4$	17.1(100)	5.65(19)	4.17(7)
Cocaine free base	1.25	$C_{17}H_{21}NO_4$	4.40(100)	7.35(75)	9.72(55)
Cocaine HCl	1.34	$C_{17}H_{21}NO_4.HCl$	5.88(100)	10.7(95)	4.74(65)
Heroin free base	1.32	$C_{21}H_{23}NO_5$	5.23(100)	6.99(75)	5.07(46)
Heroin HCl		$C_{21}H_{23}NO_5.HCl$	9.03(100)	5.17(90)	3.76(80)
MDMA	1.28	$C_{11}H_{15}NO_2$	4.83(100)	4.58(94)	4.17(68)
Methamphetamine HCl		$C_{10}H_{15}N_4.HCl$	5.11(100)	3.48(65)	6.07(45)
Caffeine	1.2	$C_8H_{10}N_4O_2$	7.52(100)	3.38(17)	3.30(15)
Glucose	1.54	$C_6H_{12}O_6$			
Lignocaine		$C_{22}H_{14}N_2O$	7.06(100)	4.51(95)	8.82(70)
Paracetamol	1.26	$C_8H_9NO_2$	3.65(100)	3.36(77)	5.71(71)
Phenacetin	1.21	$C_{10}H_{13}NO_2$	4.02(100)	3.42(70)	3.68(60)
Polythene	0.93	$(C_2H_4)_n$	4.1(100)	3.6(50)	2.5(40)
Polyurethane		$C_{25}H_{42}N_2O_6$			
Polystyrene	1.05	$(CHC_6H_5 - CH_2)_n$	4.81(100)	5.41(100)	10.9(50)
PTFE	2.15	$(C_2F_4)_n$	4.90(100)	2.42(9)	2.19(8)
PVC	1.4	$(C_2H_3Cl)_n$	4.62(100)	5.26(60)	5.07(60)
Aluminium	2.7	Al	2.34(100)	2.02(47)	1.22(24)
Air	0.0012				

Table 3.1: Illicit drug, cutting agent and packaging material information, with d-spacings and relative intensities of the three most intense lines of each substance from Joint Committee on Powder Diffraction Studies (1980)

Sample holders

The sample holders were plastic cylinders of external diameter 5mm and thickness 0.15 mm. They were sealed at both ends for safety and mounted onto the system by slotting into a holder with a hole of correct diameter, which was in turn mounted on translation stages that could move the sample perpendicular to the beam both horizontally and vertically by computer control, and parallel to the beam manually.

3.2.2 Multivariate analysis

Partial least squares regression with the Unscrambler® software, as described in section 2.2, was used to analyse the data, build mathematical models, and predict the content of new samples. The X variables were the counts in each detector channel from the measured diffraction profile of each sample.

There were two possibilities for the Y variables based on two levels of models.

1. A model to positively identify the sample as containing a narcotic

2. A model to quantify the amount of each substance in the sample

Model level 1 was built using the data from pure samples only. A single y variable was specified which had a value of 1 if the sample was a narcotic and 0 if a cutting agent not to be flagged as dangerous.

Model level 2 required the percentage of each substance present in each sample included in the model to be specified. In this work that meant a variable was required for each of the seven drugs and five cutting agents, giving a total of twelve y variables in the Y matrix.

The regression models built were used to predict the drug content of further samples, both of known content with and without packaging, and of unknown content (without additional packaging). In order to examine

the effect of the statistical quality of the data on the accuracy of the predictions, all data sets were measured with a range of count times to give diffraction profiles containing an integrated number of counts ranging from 200 to 200,000, corresponding to count times with this system of 0.1 to 100 seconds.

3.3 Results

3.3.1 Preliminary Angle selection

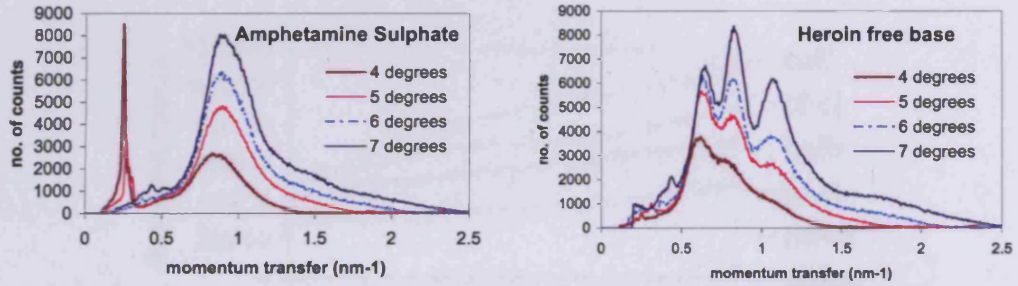


Figure 3.1: Measured diffraction profiles of amphetamine sulphate and heroin free base measured 4°,5°,6° and 7°

In the selection of the scatter angle there are several criteria to be considered. These include the momentum transfer values of interest for the substances to be detected, the number of photons scattered through each angle (with the source potential as set), the FWHM of the peaks and the energy at which the peaks occur, since low energy photons are attenuated more than higher energy photons. These factors are investigated here and used to select an optimum scatter angle for drug detection.

The diffraction profiles of amphetamine sulphate and heroin free base are shown in figure 3.1, measured at angles $4 - 7°$. It can be seen that a scattering angle of 7° provides the highest count rate in the detector and also the clearest peak-to-valley separation. There is little difference in peak

resolution in the profiles measured at 5° and 6° though, for both, the peaks are less sharp than at 7° and the intensities are slightly less at 5°. At 4° the heroin free base profile shows no clear peaks and thus this angle is not suitable for diffraction measurements on illicit drugs. Figure 3.2 shows the integrated counts for all drugs as a function of scatter angle and the FWHM of the sharp peak at 0.66 nm^{-1} in the profile of caffeine and the broad peak of amphetamine sulphate at around 1.1 nm^{-1}. These peaks were selected as they could clearly be distinguished at all angles. The plots indicate that in moving from 4° to 7° the number of counts is approximately doubled and the FWHM of the peaks is halved. This would suggest that an angle of 7°, or even higher, would be preferable.

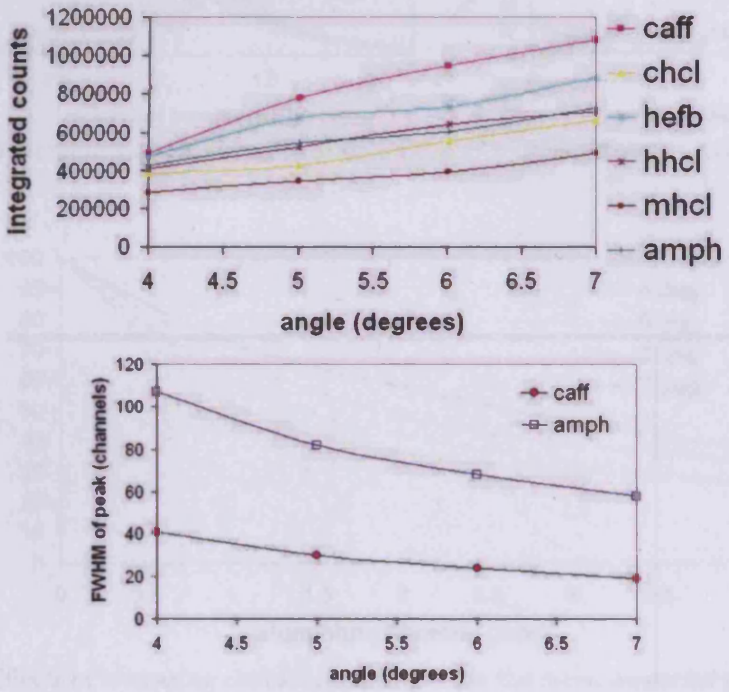
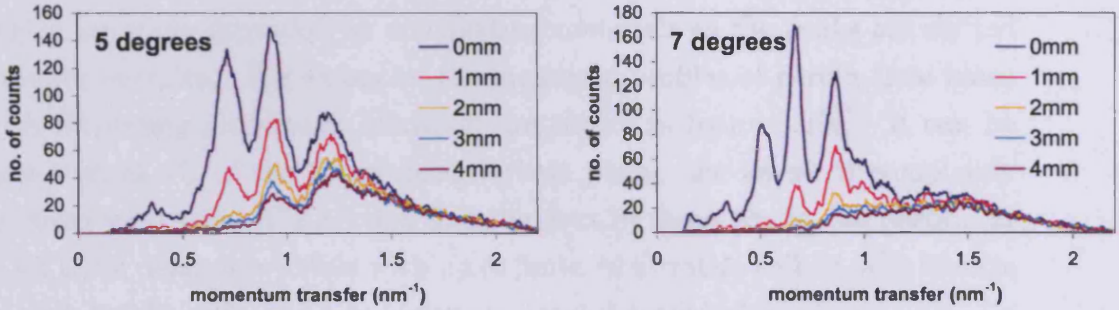


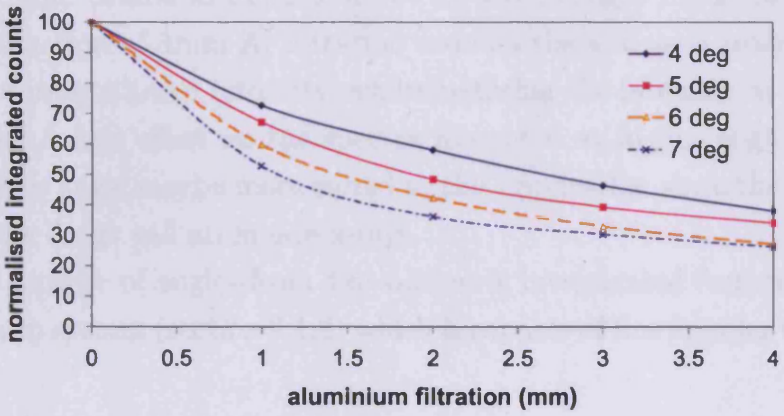
Figure 3.2: Plots of integrated counts in profiles and FWHM of peaks as a function of scatter angle

However, equation 1.11 indicates that for a given value of momentum

transfer, high scattering angles mean lower energy photons are creating the features in the spectra. In a realistic scenario, packaging and containers might remove many of the low energy photons and hence the use of smaller scattering angles should also be considered. As the profiles at 5° are similar to those at 6° no information was lost in moving to the lower angle but the peaks at 5° would occur at higher energies leading to lower attenuation effects.



(a) Effects of increasing aluminium filtration on the measured diffraction profiles of heroin free base at 5° and 7°



(b) Effects of increasing aluminium filtration on the mean measured diffraction profiles of all pure drugs at 4-7°, normalised to the integrated counts with inherent filtration only.

Figure 3.3: Effects of increasing aluminium filtration on diffraction profiles

A parcel in a final system would not contain only the illicit substances of interest, but also a range of other materials for packaging and disguise that would attenuate the x-ray beam, either before or after scatter. To measure the effect of attenuation, the diffraction profile of each pure sample was measured with increasing 1mm layers of aluminium, placed between the X-ray source and the incident collimation. This procedure was repeated at angles 4°, 5°, 6° and 7°. The energy dependence of the attenuation coefficient of aluminium means that diffraction profiles measured at increasingly higher angles are more degraded by attenuating materials as the peaks are shifted to lower energies . For example, the measured profiles of heroin (free base) with increasing aluminium filtration are shown in figure 3.3(a). It can be seen that at 7°, of the three characteristic peaks, the lowest is completely removed with 1mm Al filtration and the next by 2mm Al. In comparison at 5° all three peaks are visible with up to 3mm Al filtration, albeit with change in peak height ratio and a reduction in total detected photons.

Figure 3.3(b) shows the effect of the increasing aluminium filtration on the integrated counts in the profiles of all seven drugs. It can be seen that the introduction of 1mm Al filtration reduces the scattered intensity at 7° to 53% of its unfiltered intensity, whilst reducing the intensity at 4° to only 73%. This larger effect on the spectra measured at higher angles suggests that a lower angle maybe more suited to this application since the packaging around the drugs will attenuate x-rays.

Thus a range of angles from 4.6°-6.2° were investigated further using the moving arm system (section 2.1.3) which is capable of fine angular resolution.

3.3.2 Fine Angle Selection

Diffraction profiles of all drugs except MDMA were measured at a range of angles to produce energy-angle diagrams (figure 3.4), to be able to select the optimal angle for identifying each drug. The moving arm system was used with collimator diameters providing an acceptance angle of 0.1°. The

system was scanned through the angular range of 4.6-6.2° in 0.2° steps. The data was corrected for background and scatter from the sample holder by subtracting the corresponding measurements from an empty sample holder.

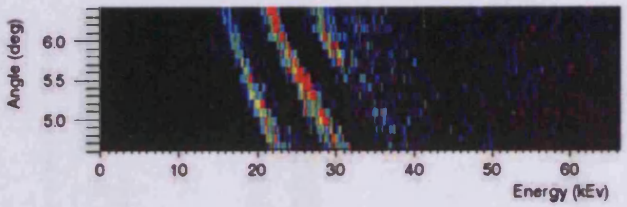
These data were then used to model the diffraction profiles of all these materials using the measured source outputs with no filtration, 1mm Al and 2mm Al (figure 3.6). The attenuation coefficient data for aluminium was taken from Berger and Hubbell (1987). To validate this model the profile of heroin free base was measured at 5.8° for 1,000s with no additional filtration, 1mm aluminium and 2mm aluminium and the measured profiles compared to those simulated from the data with no filtration and the measured incident X-ray spectra (figure 3.5).

The shapes, heights and positions of the peaks were the same in both cases so the model was considered to describe the data well and data was not collected for all angles with all filtration combinations.

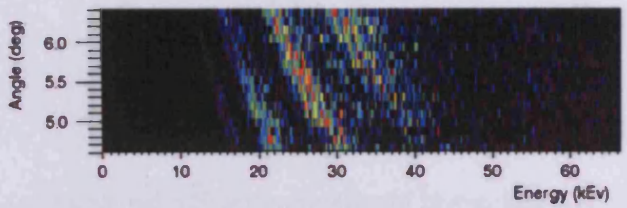
The selection of an optimum scatter angle to work at was based on two parameters, the number of counts in the required energy range and width of the peaks at each angle.

The energy range considered was 20-60keV in order to model a system for drugs detection in packages containing attenuating materials. The scatter profile was integrated over this range and plotted as a function of angle.

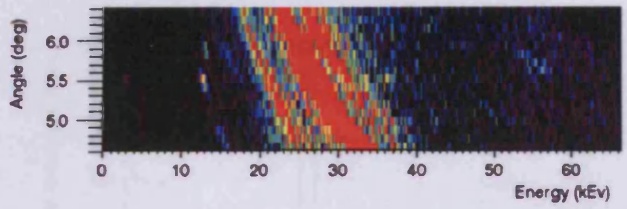
IDL software was used to fit polynomials to the curves, and the fit used to calculate the angle at which the number of counts in the required energy range is maximum for each substance and filtration combination (table 3.2). Caffeine always gave an optimum angle of 4.5° as this was the only angle at which its low energy peak (0.66 nm^{-1}) and its higher energy peaks were all in the energy range being considered. The optimum angle for the other substances changed very little with the addition of filtration. Cocaine free base was not included as the profile did not change smoothly with angle due to the size of the crystals within the sample (see section 3.3.3). The



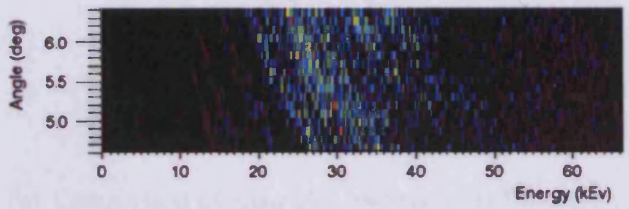
(a) Heroin Free Base



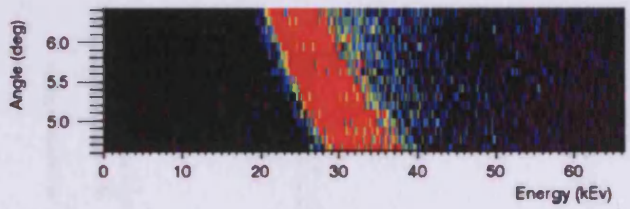
(b) Heroin Hydrochloride



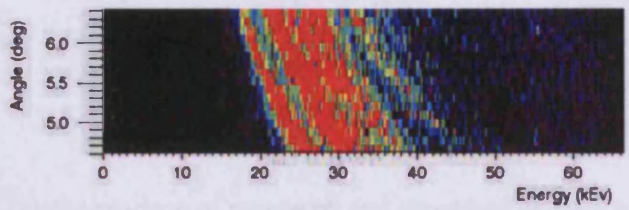
(c) Cocaine Free Base



(d) Cocaine Hydrochloride

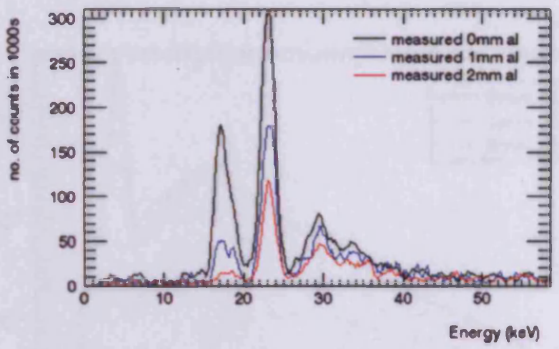


(e) Amphetamine Sulphate

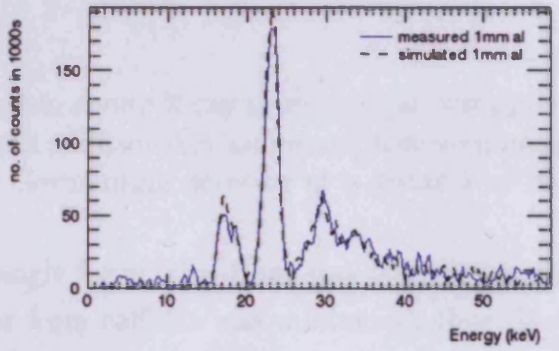


(f) Methamphetamine Hydrochloride

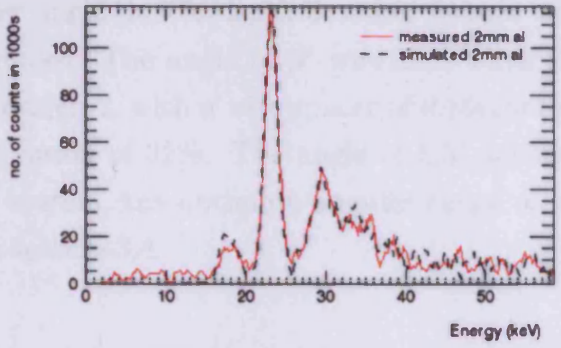
Figure 3.4: Energy/Angle diagrams for illicit drugs measured in 0.1° increments with a source potential of 70kV. Red indicates the most intense scatter, black/purple the least.



(a) Measured diffraction profiles of heroin free base with a range of aluminium filtration



(b) Comparison of measured and simulated profiles with 1mm aluminium filtration



(c) Comparison of measured and simulated profiles with 2mm aluminium filtration

Figure 3.5: The profiles, measured and simulated, of heroin free base at 5.8° counted for 1000s at 70kv and 15mA

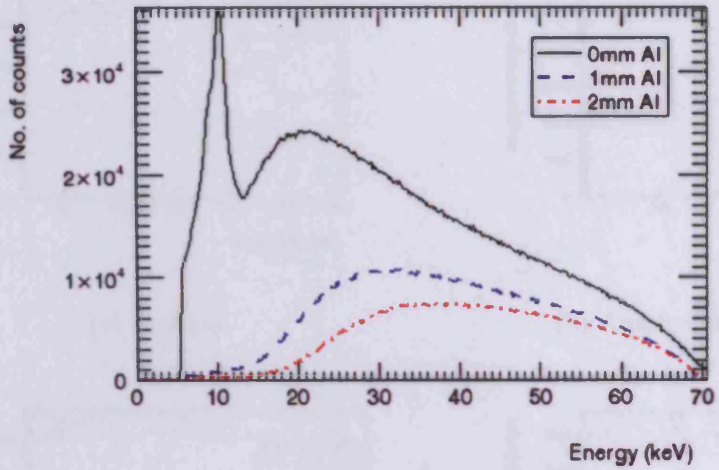
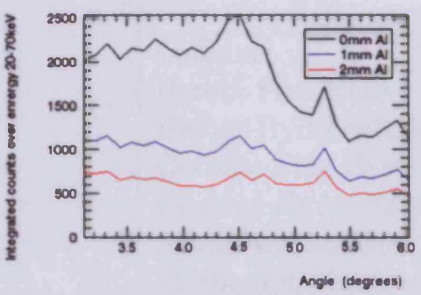


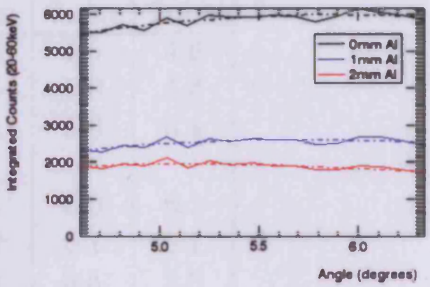
Figure 3.6: Tungsten anode X-ray source output with inherent filtration only and 1mm and 2mm additional Aluminium, measured using a $30\mu m$ Platinum pinhole with the Germanium detector at a distance of 35cm from the source

mean optimum angle for all the drugs was 5.5°, which was also the angle at which the scatter from caffeine was minimised, thus this angle was selected as optimum for this source potential (70kV).

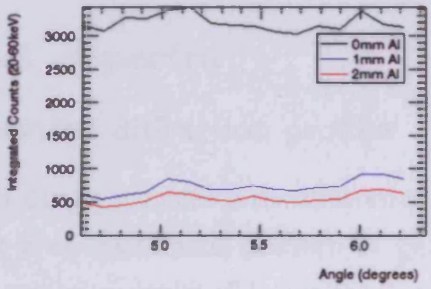
Since the FWHM of the peaks and the number of counts varies smoothly and slowly as a function of angle (see figure 3.2), the precise angle selection is not of critical importance which bodes well for a practical system, as small changes in scatter angle due to depth or other factors will not greatly alter the measured profiles. The angle of 5° was used when the collimation was provided by cell design 2, with a wire spacer of 0.16mm leading to a momentum transfer resolution of 32%. The angle of 5.5° was selected when using the moving arm system, the optimum angular range of which is considered in detail later in section 3.4.



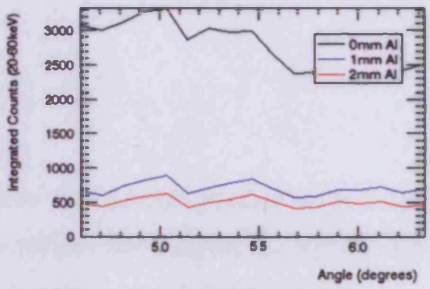
(a) Caffeine



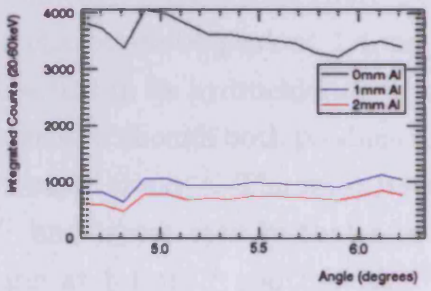
(b) Amphetamine Sulphate



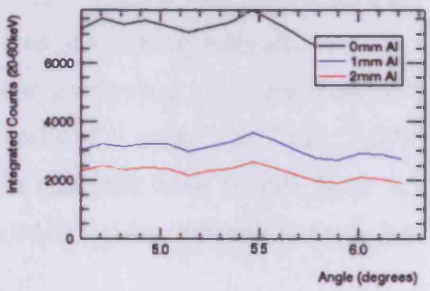
(c) Cocaine Hydrochloride



(d) Heroin Free Base



(e) Heroin Hydrochloride



(f) Methamphetamine Hydrochloride

Figure 3.7: Integrated counts in energy range 20-60keV measured with inherent filtration only and simulated for 1mm and 2mm additional Aluminium filtration

Angle °	Al Filtration		
	0mm	1mm	2mm
Caffeine	4.5	4.5	4.5
Heroin Free Base	5.0	5.0	5.4
Heroin Hydrochloride	4.9	5.0	6.1
Cocaine Free Base			
Cocaine Hydrochloride	5.1	6.0	6.1
Amphetamine $S0_4$	6.0	6.1	6.1
Methamphetamine HCl	5.5	5.5	5.5

Table 3.2: The angles maximising the number of counts in the energy range 20-60keV

3.3.3 Spectra

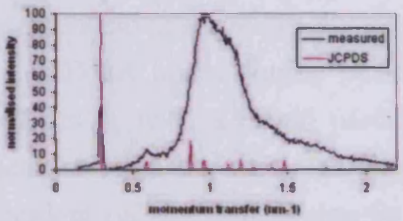
Measured diffraction profiles

Each drug produces a distinct energy dispersive diffraction profile which exhibit a characteristic pattern of peaks. All profiles are shown in figure 3.8 and are an average of ten measurements each with a count time of 10s.

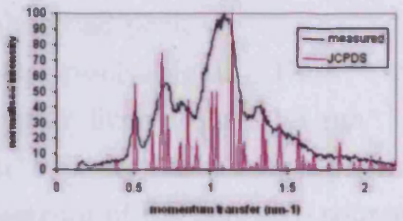
In the measured diffraction profiles the following features can be seen: Heroin produces sharp peaks at 0.5nm^{-1} and 0.7 nm^{-1}, with the free base form having an additional sharp peak at 1.2 nm^{-1} whilst the hydrochloride has a characteristic peak at 1.4 nm^{-1} and is less scattering overall.

Cocaine in its hydrochloride form is also less scattering than its free base counterpart, though both produce a series of peaks at 0.5 nm^{-1},0.7 nm^{-1},0.85 nm^{-1} and 1.0 nm^{-1}. The most intense peak in the free base profile is at 1.1 nm^{-1} and is not seen in that of cocaine hydrochloride, which instead has features at 1.4 nm^{-1} and 1. 6 nm^{-1}.

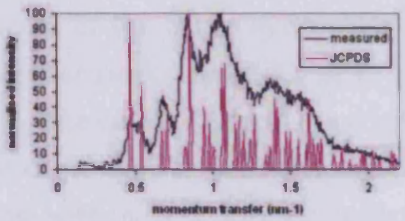
The most intense peak in the diffraction profile of amphetamine sulphate is at 0.32 nm^{-1}, which at an angle of 5° corresponds to an energy of only 9keV, which would easily be lost if there are any absorbing media present. The peaks between 0.9 nm^{-1} and 1.2 nm^{-1} overlap due to the system's



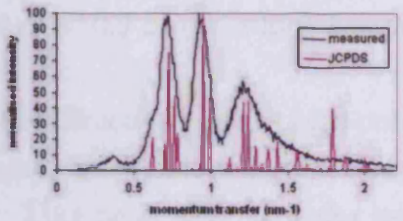
(a) Amphetamine Sulphate



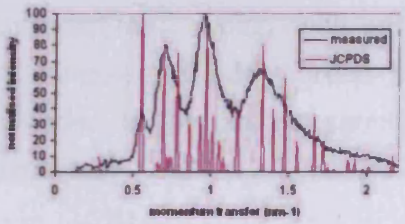
(b) Cocaine free base



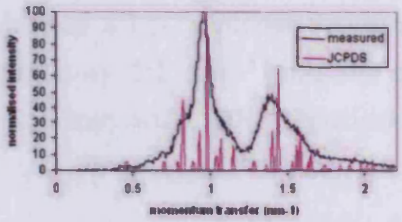
(c) Cocaine Hydrochloride



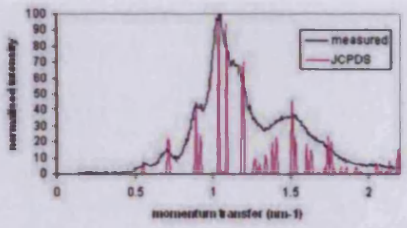
(d) Heroin free base



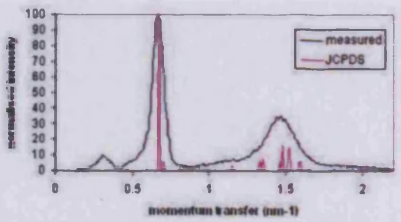
(e) Heroin hydrochloride



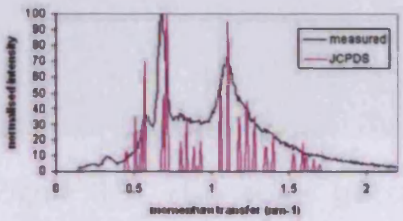
(f) Methamphetamine Hydrochloride



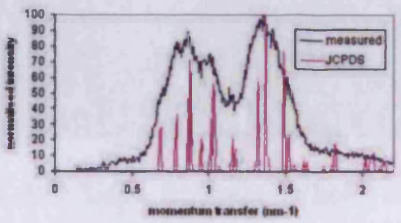
(g) MDMA



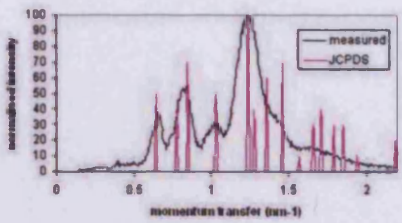
(h) Caffeine



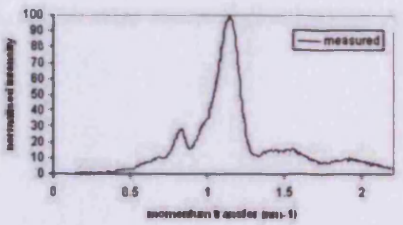
(i) Lignocaine



(j) Paracetamol



(k) Phenacetin



(l) Glucose

Figure 3.8: Measured diffraction profiles of all drugs and cutting agents measured at a nominal angle of 5° compared to the peak positions according to PDF data(JCPDS (1980))

momentum transfer resolution to give a single broad peak.

MDMA has a double peak that is not quite resolvable at 1.1 nm^{-1} and 1.2 nm^{-1}, with a broad peak due to overlapping lines around 1.4 nm^{-1}, a peak at 1.8 nm^{-1} and a small peak at 2.2 nm^{-1} where there are few photons incident on the sample due to the energy spectrum of the tungsten source.

Methamphetamine hydrochloride has the peak at 1.4 nm^{-1}, which appears in the other hydrochlorides, and another at 0.9 nm^{-1}, which has two smaller peaks on either side.

The cutting agents examined also produced diffraction profiles with peaks of characteristic position and intensity. Caffeine was the most highly scattering with a very intense peak at 0.7 nm^{-1}. This peak also coincides with one seen in the cocaine profile, which could be problematic, though the Unscambler® model dealt with it well (see section 3.12). Glucose produced an intense but fairly broad peak at approximately 1.1 nm^{-1} and the remainder of the cutting agents were less scattering, with the exception of lignocaine.

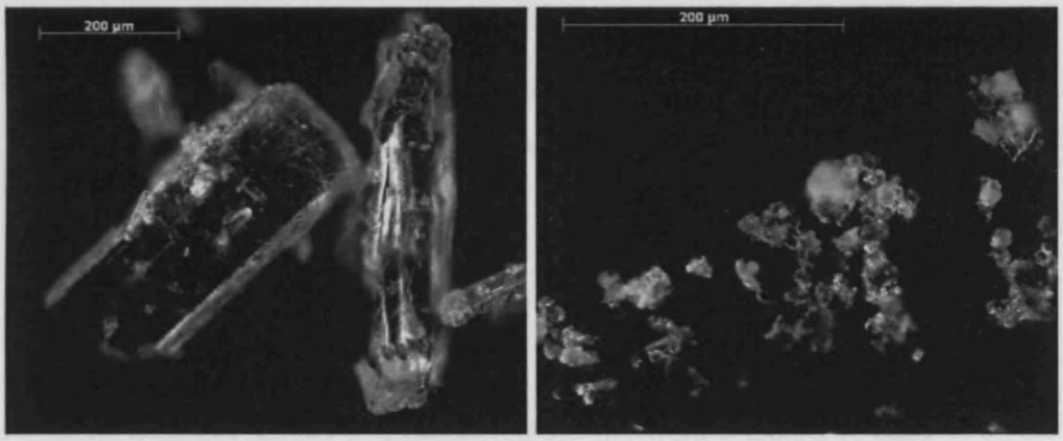


Figure 3.9: Microscope pictures of cocaine free base crystals, whole and crushed

Lignocaine was interesting as when the sample was removed, shaken and

replaced in the system, the relative heights of the peaks at 0.7 nm^{-1} and 1.1 nm^{-1} changed significantly. The profile of cocaine free base also altered when repeat measurements were taken, with the relative intensities of the peaks at 1.2 nm^{-1}, 0.9 nm^{-1} and 0.7 nm^{-1} all changing. When visually inspecting the samples this could be seen to be due to the fact that, whilst the other samples were fine powders, cocaine free base and lignocaine consisted of micro crystals. Figure 3.9 shows the crystals in cocaine free base before and after crushing. Before the crystals have a width of about $200\mu m$ and are approximately $500\mu m$ long. Therefore not all possible crystal orientations were presented in each measurement meaning some atomic planes contributed to the interference at the selected angles whilst some did not. After crushing however, the crystals were reduced to diameters of $5-50\mu m$ eliminating this problem.

The positions and relative intensities of the drugs measured previously by angular dispersive methods are collated in the Powder Diffraction File (JCPDS (1980)). A comparison between the measured profiles at a nominal angle of 5° and the predicted peak positions and intensities can be seen in figure 3.8. The theoretical and measured intensities are not in exact agreement as the spectral shape of the incident X-ray beam and the momentum transfer resolution of the system, multiple and compton scatter and attenuation in the sample holder have not been taken into account. This is considered later in section 3.5.

3.3.4 Multivariate analysis

Building and testing models

A range of Unscrambler® models were built with a range of data and parameters to optimise the predictions.

Model level 1. was built using the data from pure samples only. A single y variable was specified which had a value of 1 if the sample was a narcotic

and 0 if a cutting agent, not to be flagged as dangerous.

Variations of model level 2. were built using a combination of data, ranging from all samples with all count times to only pure samples leaving out the statistically poor data. A y variable was specified for each substance that could be present, giving a total of 12 y variables, with the value of the variable the percentage of that substance present in the sample.

All models were built on measured spectra from which the measured background (the scatter from an empty sample holder) had been subtracted. Each spectrum was then mean-normalised, where the equation for the mean-normalisation of each point is:

$$X(i) = X(i) \div \frac{\sum_{i=1}^{n} X(i)}{n} \tag{3.1}$$

Where n is the number of x-variables in the matrix, in this case then number of MCA channels.

Each model was used to predict the drug content of a second data set. The success of the model could be analysed in several ways:

1. RMSEP (root mean squared error on prediction)
   - This measures the average difference between the predicted content of each substance and the actual content in the original units (i.e. in percentages).

2. Slope - This is the gradient of the best fit line relating predicted and actual contents. A perfect model would have a slope of 1.

3. The sensitivity and specificity of the drugs detection system with different detection thresholds

Model level 1 gave a prediction coefficient for each sample which was an indication of how likely the sample was to contain a drug (0 being not likely, 1 very likely), though with no information about which particular narcotic.

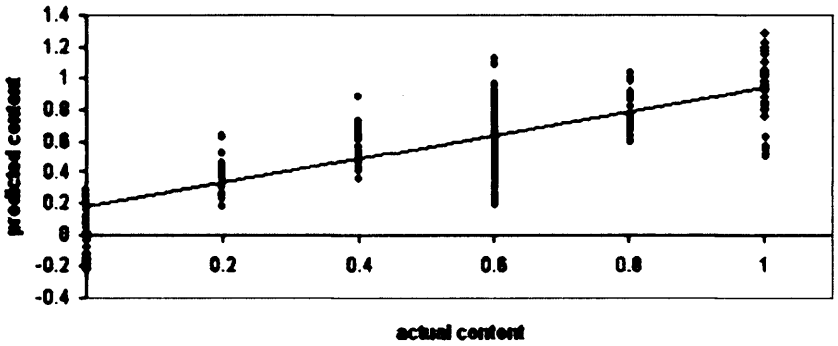


Figure 3.10: Plot of the predicted coefficients against actual for the level 1 model for 62 samples, 4 repeats of each with a 4s count time.

Prediction threshold	sensitivity (%)	specificity(%)
0.10	100	53
0.15	100	63
0.20	99	73
0.25	98	87
0.30	96	100
0.35	92	100
0.40	87	100

Table 3.3: The sensitivity and specificity of Unscrambler® model (level 1) for different prediction threshold values

There was some overlap in the value of the prediction coefficients between samples containing some drug and samples of pure cutting agent (figure 3.10), thus a prediction threshold was required, above which the sample was marked as a drug and below which it was considered to not be suspicious. A range of thresholds were considered and the sensitivity and specificity of the system calculated at each (Table 3.3).

Table 3.3 indicates that a suitable threshold for the detection of drugs is 0.3, as at this value there were no false positive results giving a specificity of 100% and very few missed drugs, with a sensitivity of 96%, including samples which contained only 20% drug to begin with. However, if such low

Data model on which model built	Filtration	Mean slope	Mean RMSEP
all	none	0.51	13.6
all samples (2,000 counts and above)	none	0.68	13.0
all samples (10,000 counts and above)	none	0.78	13.3
all samples (20,000 counts and above)	none	0.77	13.6
pure samples	none	0.50	13.6
pure samples (2,000 counts and above)	none	0.77	14.3
pure samples ((10,000 counts and above)	none	0.80	15.0
pure samples (20,000 counts and above)	none	0.77	13.6
all	1mm Al	0.28	16.6
all samples (2,000 counts and above)	1mm Al	0.55	14.3
all samples (10,000 counts and above)	1mm Al	0.67	16.0
pure only	1mm Al	0.35	16.3
pure samples (2,000 counts and above)	1mm Al	0.57	15.2
pure samples (10,000 counts and above)	1mm Al	0.68	13.2

Table 3.4: The properties and performance of Unscrambler® models.RMSEP = root mean squared error of prediction

concentrations were not considered to be priorities, the threshold could be set higher to further reduce the risk of false alarms.

The properties of each level 2 model and the measures of their success at predicting data are shown in Table 3.4.

In all cases the mean slope was less than 1, indicating that the drug content was always under-predicted. It can be seen from this that whilst the model using only the pure samples with integrated counts of 10,000 and above gives the slope nearest to one, the model using all samples with integrated counts of 2,000 and above has the lowest RMSEP. However, by compromising slightly on both measures, the best overall model was the one built using all samples, but only the integrated counts higher than 10,000. This gave a mean slope of 0.78 and an RMSEP of 13.3, indicating that any predictions should be accurate to within $\pm 13\%$.

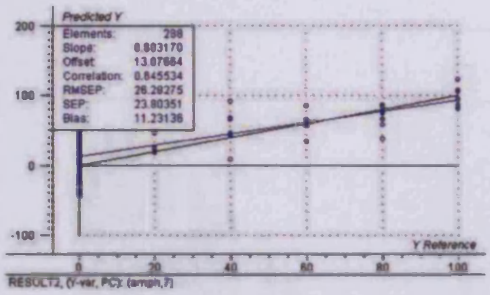
Predictions

The predictions made of all Y variables are shown in figure 3.11.

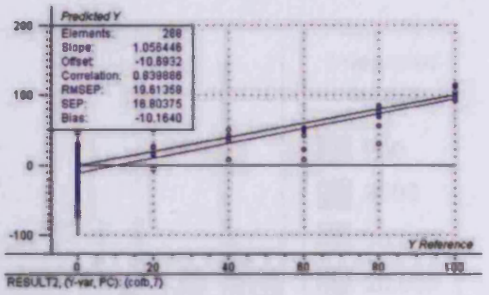
The underestimation of drug content can clearly be seen, as the slope of the linear fit has a gradient less than unity for all drugs. In a system with the requirement of only detecting the drug, this may be sufficient if an appropriate detection threshold is set. It can also be seen that some Y variables can be predicted better than others, with cocaine hydrochloride having the largest RMSEP (25.5) and therefore being the least well predicted. The caffeine content of each sample is the best predicted (RMSEP=7.4), most likely as it has the most distinctive diffraction profile. The most accurately predicted drug is Heroin free base with a RMSEP of 11.7.

By examining the the predictions of the drugs in more detail (e.g. that of heroin free base, shown in figure 3.12), it can be seen that the predictions on profiles containing integrated counts of 10,000 - 200,000 give virtually the same predicted values, with those on profiles containing only 2,000 counts also being similar. This suggests that a system where the profile from each point in an object contained a 2,000 counts would be no less sensitive than a system where the profile from each point contains 200,000 counts. However, the large majority of the false positive results are from the statistically poorer spectra thus a balance between sensitivity, specificity and the number of counts in the profile must be found.

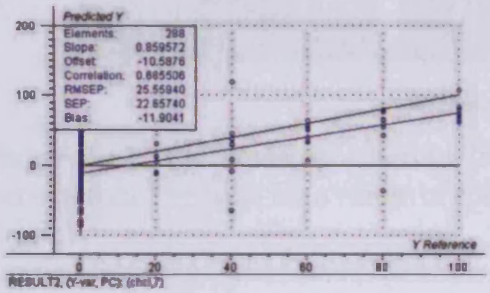
The number of photons detected is dependent on a range of factors such as the angular acceptance of the system collimation, the current and potential settings on the X-ray source and the counting time of the system. If a maximum count time at each point within a parcel is set by a minimum parcel throughput, the other parameters may be adjusted to provide the required number of detected counts for the sensitivity and specificity of detection necessary. This is considered further in section 3.4.



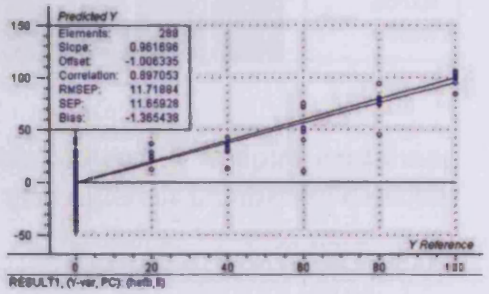
(a) Amphetamine Sulphate



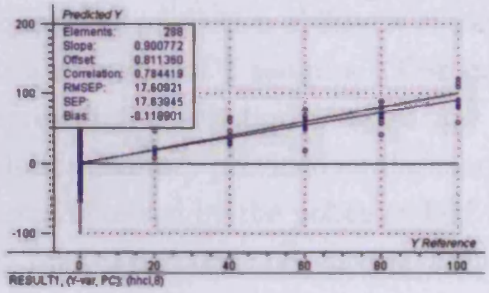
(b) Cocaine free base



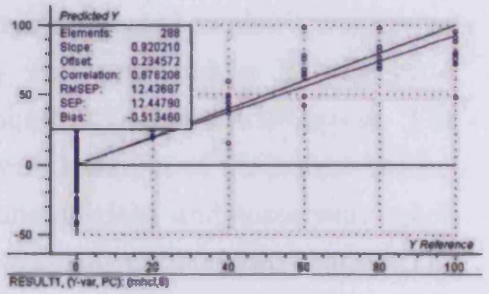
(c) Cocaine hydrochloride



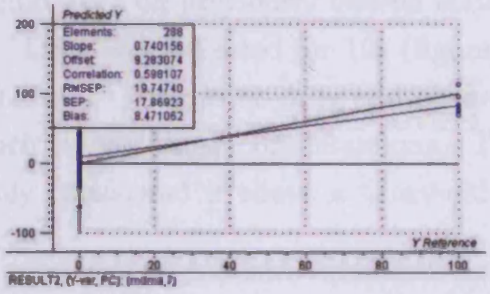
(d) Heroin free base



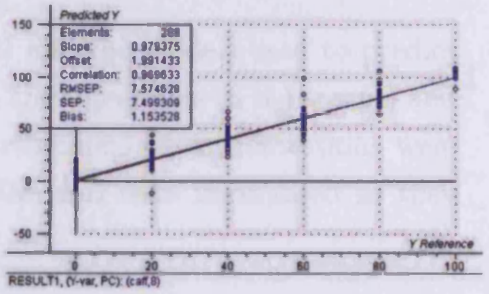
(e) Heroin hydrochloride



(f) Methamphetamine hydrochloride



(g) MDMA



(h) Caffeine

Figure 3.11: Prediction against actual content for model predictions.
The black line indicates the ideal relationship, blue the actual linear fit.

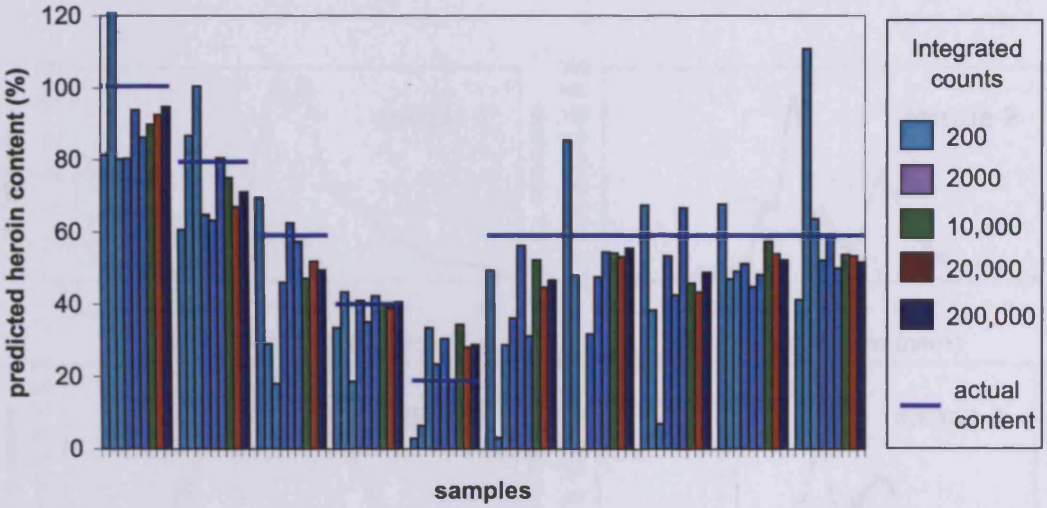


Figure 3.12: Predictions of heroin free base content in samples containing some heroin free base for a range of spectra with different numbers of detected photons

3.3.5 Predictions of 'Blind' spectra

To test the prediction ability of the system and selected model it was tested on a selection of 7 samples. The samples were provided by HOSDB as a representative selection of what may be found in real world scenarios. The information they provided on the samples was that some of them were 'street' samples, seized by the police or HM customs officials, and some were made up from the drugs and cutting agents previously used in the study. The street samples were selected to contain the drugs that had been examined in this study, but were mixed with unknown cutting agents to test if the system could work on previously unseen materials.

Data was collected for 10s (figure 3.13) and the models used to predict a) level 1 - drug or no drug and b) level 2 - the substances in the sample and their approximate concentrations. The predictions of concentrations were only considered if above a threshold of 20% and were normalised so they

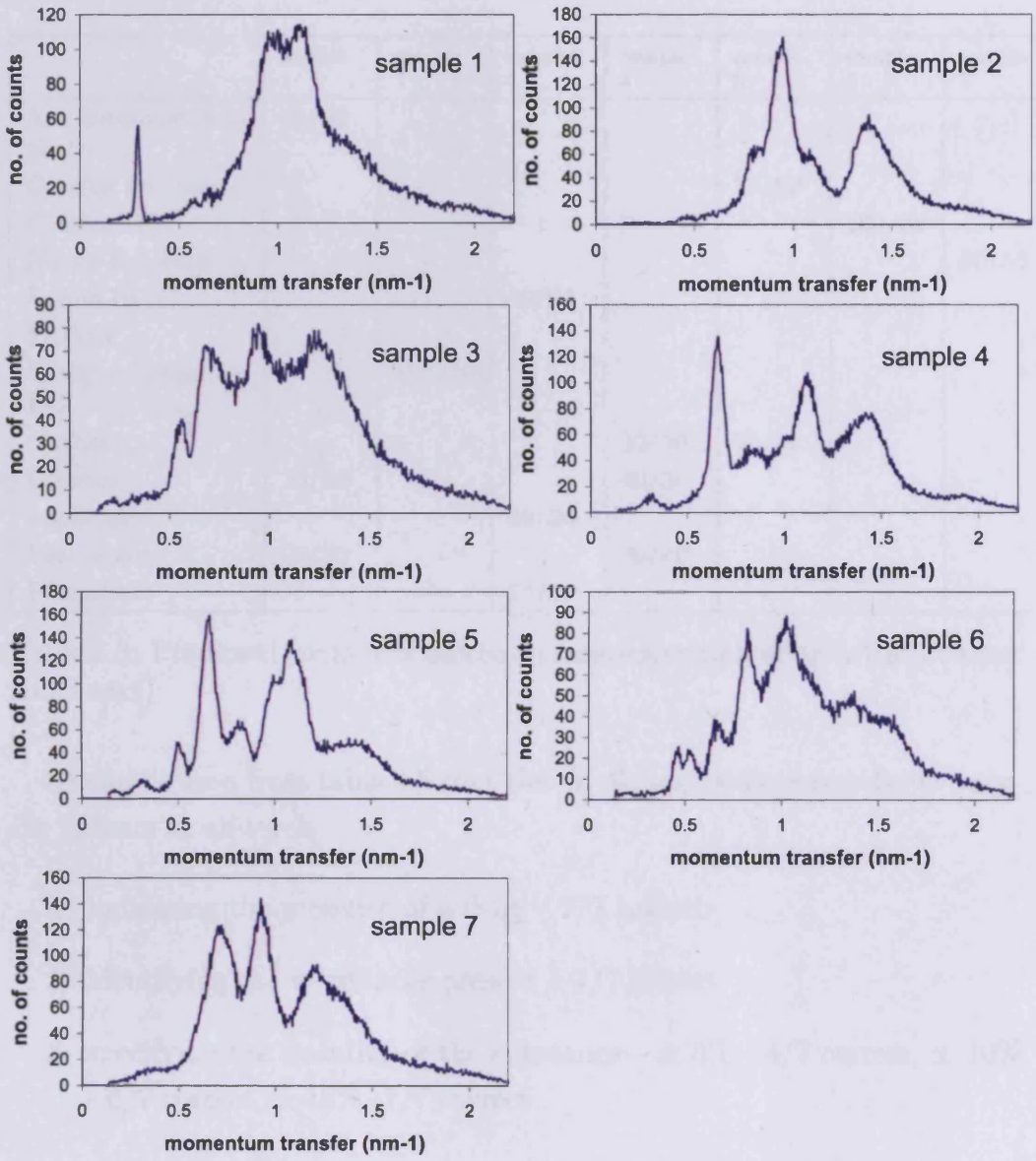


Figure 3.13: Measured diffraction profiles of 'blind' samples at 5° with a count time of 10s

added to give 100% content for each sample where appropriate.

	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6	sample 7
Amphetamine Sulphate	60/**60**						
Cocaine free base					75/**80**		
Cocaine HCl						100/**85**	
Heroin free base							50/**55**
Heroin HCl			50/**60**				
MDMA							
Methamphetamine HCl		100/**100**					
Caffeine				30/**30**	25/**20**		
Glucose	20/**20**			40/**30**			
Lignocaine			30/**20**				
Paracetamol	20/**20**			30/**40**			
Phenacetin			20/**20**				

Table 3.5: Predicted content of unknown samples compared to actual content (bold text)

It can be seen from table 3.5 that the predictions were successful to varying extents at all levels.

1. indicating the presence of a drug – 7/7 correct
2. identifying the substances present – 7/7 correct
3. specifying the quantity of the substance – ± 5% - 4/7 correct, ± 10% - 6/7 correct, ± 15% -7/7 correct

These results show that the model is able to predict the drug, its percentage purity and the cutting agents used well, thus would make an excellent detection system. Blind samples 3, 6 and 7 were 'street' samples and therefore may have contained impurities that were not previously tested on this

system and had not been incorporated into the model. This is why the model could not analyse the 45% of sample 7 that was not heroin free base and over-predicted the cocaine content of sample 6. Sample 3 was predicted exactly as methamphetamine hydrochloride is usually found in its pure form, thus any contaminants present would be from the manufacturing only process and would merely be trace amounts, though the street sample tested was a brown powder while the calibration samples prepared by HOSDB were white. The contents of the other samples were predicted accurately and the ability to predict all drug contents to within 15% is consistent with the RMSEP of the model, which was 13.3%. If all known drug samples were characterised on this system and incorporated into the model, the model should identify the content of a blind sample with confidence.

3.3.6 Effects of Packaging

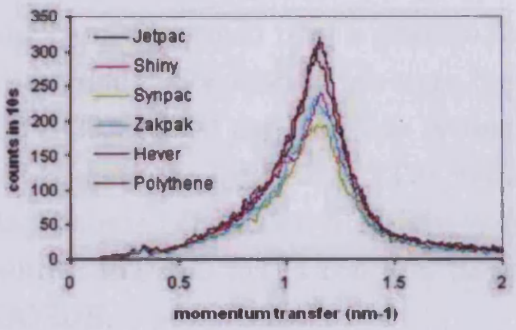
The effects of packaging materials on the spectra measured and the corresponding effects on the prediction accuracy of the unscrambler models have been investigated. The effects of packaging materials can be broadly divided into two categories : attenuating and scattering.

Attenuation

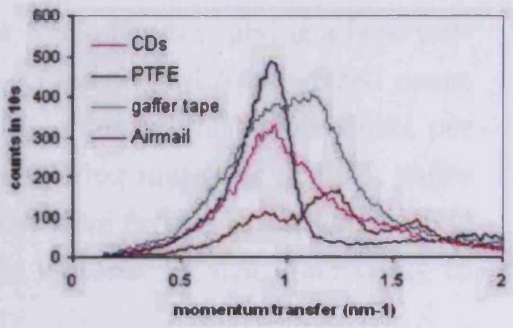
The effects of attenuation on the measured diffraction profiles of drugs was considered in section 3.3.1 when selecting the optimum scatter angle, so are not discussed further here.

Scattering

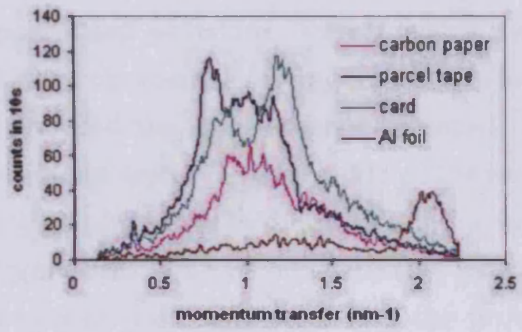
To measure the scattering contributions from packaging, the diffraction profiles of a range of materials, suggested by HOSDB as commonly used to wrap dugs, were investigated.



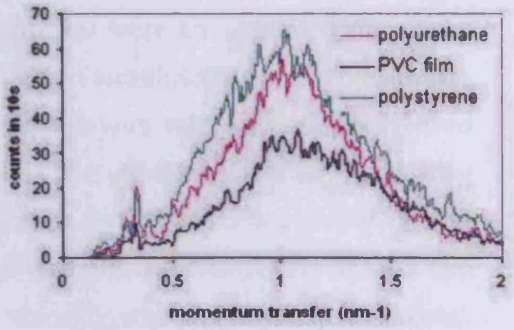
(a) Packaging provided by SOCA



(b) Highly scattering packaging materials



(c) Moderately scattering packaging materials



(d) Weakly scattering packaging materials

Figure 3.14: Measured diffraction profiles of a range of packaging materials at 5° with the source operating at 70kV.

For sheet materials such as laminate polythene, foil, tapes and paper, strips were cut, rolled to the diameter of the sample holder and placed in the same position as the sample. Rigid materials such as CDs were mounted on top of the sample holder with their centre over the centre of the sample holder.

Figure 3.14 shows the diffraction profiles of a range of packaging materials. A range of materials were provided by SOCA (serious organised crimes agency) as examples of packaging used for drugs, and all produced diffraction profiles extremely similar to that polythene, even a material with a metallic appearance. Though none of them produced sharp peaks in their diffraction profiles, polythene appeared to overlap with the profile of amphetamine

sulphate, thus could pose a problem in a detection system giving a false positive result. Some materials were highly scattering (with integrated count rates 3000 -8000 per s for this system), others less so (500-2000 counts per s) as shown in figure 3.14. The highly scattering materials (PTFE, gaffer tape, airmail envelopes, polythene and CDs) were further investigated along with aluminium foil, a common packaging material for drugs according to HOSDB.

Measurements were made using the pure drug samples. The samples were positioned as before. Two layers of the material were then placed on either side of the sample (giving four layers in total) to simulate a double packaging layer and the measurement repeated. Further layers were added until there were ten layers on either side of the sample. The effects of adding increasing layers of polythene around a sample of cocaine hydrochloride can be seen in figure 3.15. The broad scattering profile of the plastic completely masks the peaks of the drugs, rendering the profile visually unrecognisable. However, the Unscrambler ® software model (as described in section 3.3.4) is still able to identify the drug up to a certain point, though the predicted content with 0.8mm of polythene drops by up to one third of that with no packaging.

Though the addition of packaging reduces the predicted drug content, the effects seen here are extreme as the volume of the sample is small (diameter 5 mm), and is of the same order of thickness as the packaging, as ten layers of polythene has a thickness of 0.8 mm, giving 1.6 mm of polythene when on both sides of the sample, thus occupying approximately a third of the scattering volume. A detection system investigating parcels would in general be looking for larger quantities of drugs, therefore this modeled scenario is pessimistic. Different drugs were most disguised by different packaging materials, though in general aluminium foil, paper and PTFE tape had little effect on the predictions, whilst polythene and gaffer tape greatly reduced the predicted drug contents.

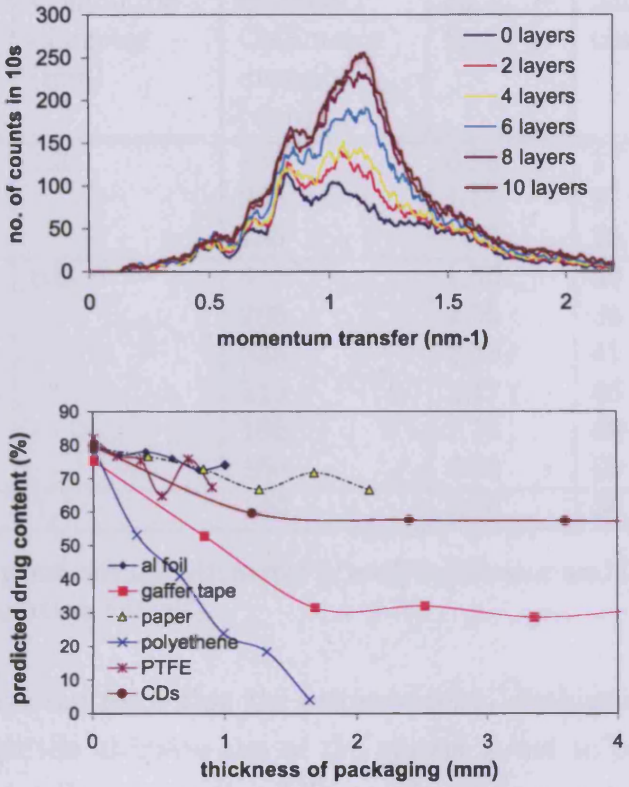


Figure 3.15: a) Effects of increasing layers of polythene packing on the measured profile of cocaine hydrochloride b) Effects of increasing layers of packaging on the predicted content of a pure cocaine hydrochloride sample

3.4 Count rate / Momentum transfer resolution

The next stage in considering using EDXRD for substance detection was to investigate the trade-off between momentum transfer resolution and number of detected counts.

This was done by measuring the diffraction profiles of the pure drugs with a range of momentum transfer resolutions and integrated counts, using the multivariate analysis software on the profiles, and seeing for which combi-

Incident collimation plug diameter (mm)	Collimator diameter (mm)	Sample-Collimator distance (mm)	$\theta_{max} - \theta_{max}(°)$	Angular resolution (%)
0.5	1.3	430	0.93	17
		235	1.14	21
		135	1.47	26
0.5	6.0	385	1.66	30
		285	2.00	36
		235	2.28	41
		210	2.47	45
		185	2.71	49
		160	3.02	55
1.5	1.3	140	2.00	36

Table 3.6: The positions and diameter of each collimator and the corresponding angular resolution values

nations of counts and resolution the software could distinguish between the drugs. Although the ultimate aim of the system is not to be able to identify one drug from the others, the ability to do so gives an indication of the discriminatory power of the system.

For these measurements an additional filtration of 1mm aluminium was used to filter out the lower energy photons that would be most attenuated in a package and the moving arm system set at a nominal scatter angle of 5.5° as determined in the section 3.3.2.

Different plugs were used as the aperture for scatter collimation, with diameters 6mm and 1.3mm. To alter the angular resolutions available using these collimators, the sample-collimator distance was altered, allowing the range $\pm 0.5° - \pm 1.6°$ to be investigated.

The Maestro software was used to select a ROI from 20-60keV, and set the counting limits to an integrated number of counts in this range. The integrated counts 100, 300, 500, 1,000, 3,000, and 10,000 were measured. In order to correct for background scatter from the sample holder and other

sources, before each data set was collected two 100s live time measurements were made, one with the sample in place and one with an empty sample holder. The percentage of counts in 100s second from the sample alone was calculated and used to scale the integrated counts collected so the numbers stated above would be the counts for the sample alone. Once these data had been measured, the live time required for each was noted and scatter with the empty sample holder collected for the same time and subtracted from the sample measurements.

The measured spectra were imported to Unscrambler®, mean normalised and PCA used to identify groups in the data. The scores plot (see section 2.2) was used for this as it shows each sample in the model projected onto axes representing two PCs.

3.4.1 Integrated Counts PCA Results

PCA analysis was conducted on subsets of the measured data to establish the minimum counts and maximum resolution which produce spectra that could be correctly grouped by the Unscrambler software.
Cocaine free base was excluded from the analysis as its diffraction profile varied if the sample was moved and thus its diffraction profiles were not always grouped together, regardless of statistical quality and resolution (see section 3.3.3).
Each set of data with each number of integrated counts was analysed separately and the scores plots displayed in figure 3.16. The PCA scores for the profiles containing 10,000 counts clearly shows that the different drugs can be separated by PCA, though there is some overlap in groups at 49% and 55% angular resolution between the two different types of heroin. The scores plot for 3,000 counts shows no appreciable difference to that for 10,000 counts. At 1,000 counts the groups look the same but with an increased overlap around the origin, indicating more samples are 'average' and cannot be discriminated between. At 500 counts the groups are apparent as for 1,000, though with

a more overlapping group boundaries. For 300 integrated counts the groups have been compressed down to the centre and though visible are not sufficiently distinguishable to be useful, with only the points from high resolution data being separated. For 100 integrated counts, the scores show no clear groups thus having only 100 integrated counts in a profile is insufficient for PCA to be successful, whatever the resolution. These results indicate that a minimum of 300, 500 or 1,000 counts is required in a diffraction profile in order to successfully use PCA to analyse the sample, the exact limit of which depends on the momentum transfer resolution.

3.4.2 Momentum Transfer Resolution PCA Results

Principle components analysis was run on each set of data taken at each momentum transfer resolution for all integrated counts. Since the profiles containing only 100 counts were shown to be inadequate for multivariate analysis, the measured diffraction data collected for 300 counts or more were used to examine the groupings in the PCA scores from a range of collimation geometries, spanning a range of momentum transfer resolutions from 17-55%.

It can be seen from figure 3.17 that with a momentum transfer resolution of 17% the groups are very clear and separate, but converge and begin to overlap at greater angular acceptances. At 21% and 27% the points for 300 counts are within the groups but separated from the other 4 points, suggesting that the data is not sufficiently similar. At 30% the 300 count amphetamine sulphate profile is grouped with the cocaine hydrochloride spectra indicating at this resolution at least 500 counts are necessary. Likewise at 36% the 300 count measurements of heroin free base and methamphetamine hydrochloride also have nearly identical scores on PC1 and PC2. By 45% the groups are concentrated in the centre of the scores plot, indicating that the profiles are all close to the average with the 300 count points at the periphery showing that noise now produces the greatest variation within the samples. At 55%, though some grouping is still apparent, the clusters are too close

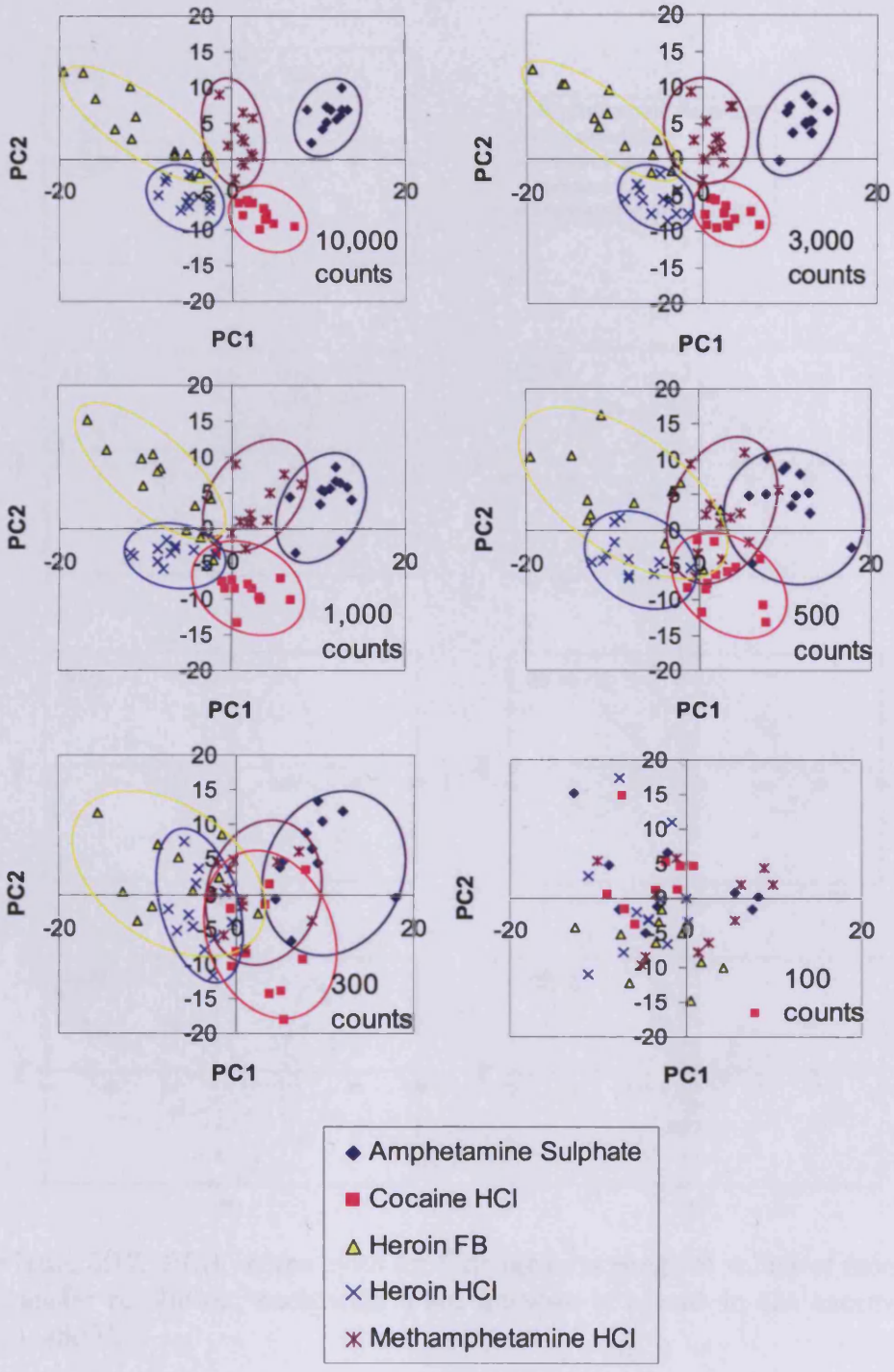


Figure 3.16: PCA scores plots for 5 drugs for a range of values of momentum transfer resolution, each with a set number of counts in the energy region 20-60keV.

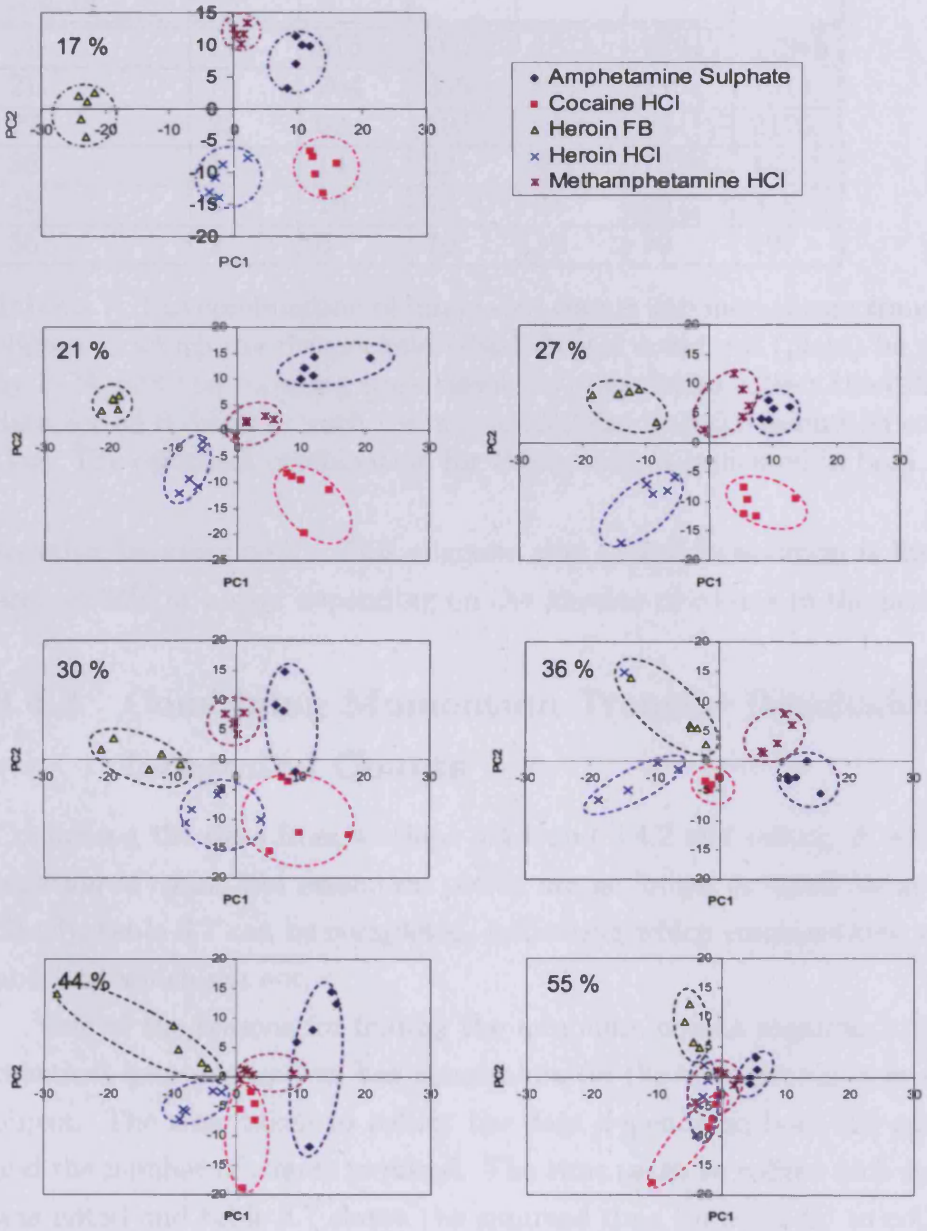


Figure 3.17: PCA scores plots for 5 drugs for a range of values of momentum transfer resolution, each with a set number of counts in the energy region 20-60keV.

res.(%) – counts	100	300	500	1000	3000	10000
17	177	515	705	1813	5404	17945
21	73	204	349	691	2102	7014
27	22	62	108	216	644	2176
36	5	14	22	**44**	132	443
45	3	10	16	31	95	318
55	2	6	10	19	59	197

Table 3.7: The combination of integrated counts and momentum transfer resolution in which the drugs could (shaded) and could not (plain) be grouped by PCA and the counting time taken (in seconds) to collect the diffraction data for all 6 drugs at each count/ momentum transfer resolution combination. The optimum combination for this system is indicated in bold.

together be meaningful. This suggests that angular resolution is limited to around 36% or better depending on the number of counts in the profile.

3.4.3 Combining Momentum Transfer Resolution and Integrated Counts

Combining the data from sections 3.4.1 and 3.4.2 and noting at what combination of count and resolution points are no longer grouped correctly and clearly, table 3.7 can be completed, indicating which combinations are suitable and which are not.

One of the reasons for finding the minimum counts required is that any practical scanning system has constraints on the time available to scan an object. The time taken to collect the data depends on both the resolution and the number of counts required. The time taken to collect each spectrum was noted and table 3.7 shows the summed time (in seconds) to collect the data for all six illicit drugs at each resolution/counts combination. It can be seen that a reduction in counts and resolution vastly reduces the time taken to collect the data. Of the resolution/counts combinations that could

successfully be used for PCA, that of 36% and 1,000 counts has the lowest data collection time and is thus gives the optimum combination for this system and task.

3.4.4 PCA of data

The selected limiting parameters selected as 1,000 counts with a momentum transfer resolution of 36% were then used to collect more data. With the moving arm system this corresponded to a primary collimator diameter of 1.5 mm, and scatter collimator diameter of 1.3 mm at a distance of 140 mm from the sample. An empty sample holder in place gave rise to 14,000 detected counts in 100s. The lowest scattering of the pure samples was Heroin Hydrochloride, which gave 40,000 counts in 100s, thus implying that a count time of 4s was required for each sample to scatter at least 1,000 photons. Diffraction data were collected from each sample for 4s with 5 repeats of each, and PCA conducted on the data. The scores plot is shown in figure 3.18. New samples of cocaine free base and lignocaine were prepared which were more finely crushed, eliminating the irreproducibility discussed previously. It can be seen that each substance is grouped well by PCA, and that the illicit drugs are separated from the cutting agents (caffeine, glucose, lignocaine, paracetamol and phenacetin).

Only PCs 1 and 2 are plotted here, and it should be noted that higher PCs contain more information, e.g. PC 3 separates heroin free base, heroin hydrochloride and methamphetamine hydrochloride. The clear groups and the fact that the drugs are separate from the cutting agents indicate that these data have a suitably high resolution with sufficiently high statistics to be used with this multivariate software in the identification of illicit materials.

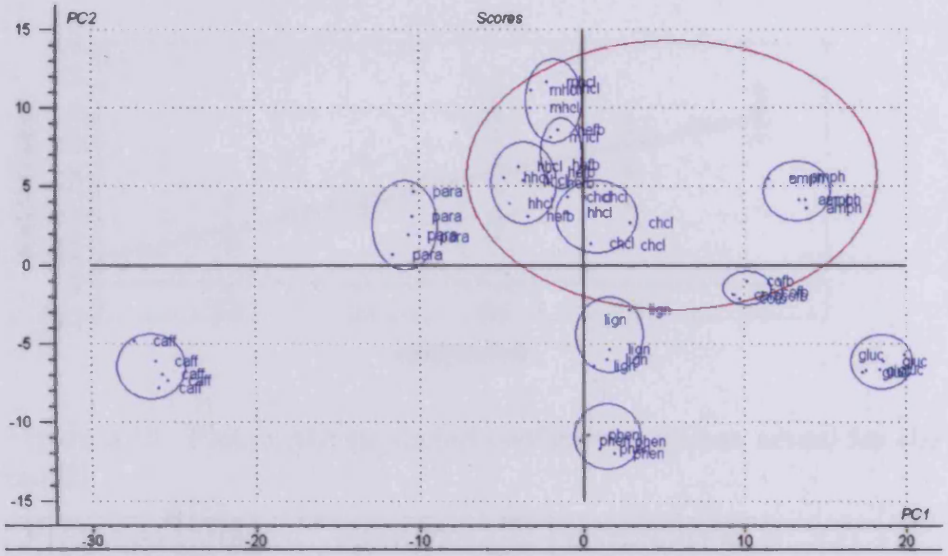


Figure 3.18: A scores plot for 6 drugs and 5 cutting agents grouping 5 sets of diffraction data each containing 1,000 counts taken with a momentum transfer resolution of 36%. The purple line separates the drugs from the cutting agents.

3.4.5 Building and testing new models

To test the system a complete set of data for all 62 samples provided by HOSDB was collected, with five repeats of each sample measured all with a momentum transfer resolution of 36% and a count time of 4s. The spectra were divided into calibration (4 x repeats) and validation (1 x repeat) samples and a regression model built using the calibration set and tested using the validation set.

Two models were built, one with the percentage content of each substance used as Y-variables (11 in total (MDMA was not included)- level 2), and one with only one Y-variable, the percentage of illicit drug in the sample with no distinction between drugs (level 1).

The predictions made by model level 1 are shown in figure 3.19. It can be

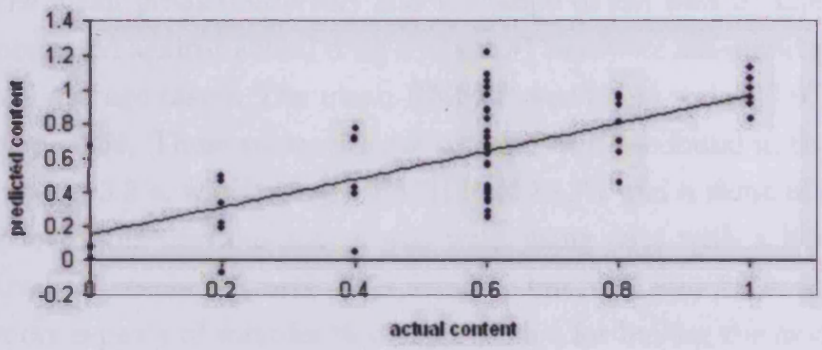


Figure 3.19: Plot of the predicted coefficients against actual for the level 1 model

Prediction threshold	sensitivity (%)	specificity(%)
0.10	96	80
0.15	96	80
0.20	96	80
0.25	93	80
0.30	92	80
0.35	83	80
0.40	81	80

Table 3.8: The sensitivity and specificity of Unscrambler® model (level 1) for different prediction threshold values at 5.5° with a momentum transfer resolution of 36%

seen that some samples are predicted too low and some too high. Applying a detection threshold leads to some missed drugs and some false positive results from the cutting agents (3.8). A detection threshold of 20% gave a sensitivity of 96% and specificity of 80%. These values are comparable to those of the best model from section 3.3.4, which used data with a better resolution (32% compared to 36%) and more counts (up to 200,000 compared to 1,000). This model had a sensitivity of 99% of and a specificty of 73% with a 20% detection threshold.

The success of the level 2 model was measured using the RMSEP (giving

the mean prediction error) and the slope of the best fit line of the plots of predicted against actual drug content. These were measured for all drugs and the average taken. The mean RMSEP was found to be 17.9% and the mean slope 0.56. These values are not as good as those found in the model used in section 3.3.4, which gave a RMSEP of 13.3% and a slope of 0.78. As stated above, this previous model was built using data with a better momentum transfer resolution and more counts, but was also built and tested using more repeats of samples (6 compared to 4 for builing the model, 6 compared to 1 for testing), which may explain the reduced prediction accuracy of the new model.

3.5 Modeling the system response

In order to a) validate the experimental data and b) be able to predict the future response of the system to new substances with various geometries, a simulation was developed which used the JCPDS data. In order to model the system using the JCPDS data, several factors needed to be taken into account. These include the angular resolution of the collimation system, the energy resolution of the detector and its dependence on energy, the shape of the X-ray spectrum from the tungsten anode source, attenuation in the sample and sample holder, partial attenuation and scatter from the collimation system and the contribution from Compton and multiple scatter. The angular resolution of the collimation system and energy resolution of the detector combine together to give the momentum transfer resolution of the system (equation 2.2).

The distribution of angular contributions from the scattering volume was calculated for an EDXRD system by Luggar et al. (1996). They found that the peak of the angular distribution was slightly below the nominal angle and the distribution asymmetric (see fig 3.20). Since this distribution was calculated for a system with a nominal angle of 5° and an angular resolution

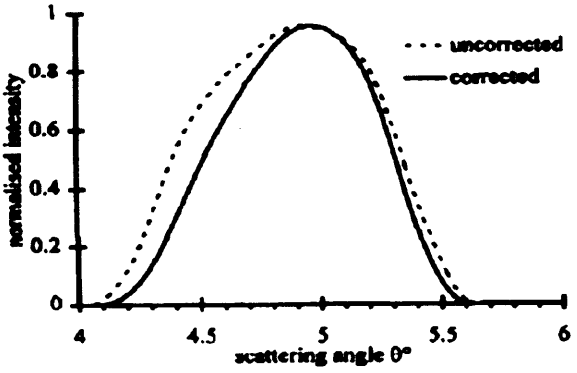


Figure 3.20: Uncorrected and focal spot corrected angular blurring distributions (Luggar et al. (1996)), calculated by using the elemental analysis technique for an EDXRD system with a nominal scatter angle of 5°

of 30% in the horizontal plane, it could be used to model the system used to make the measurements described above.

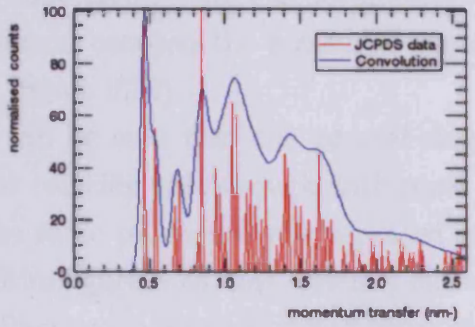
The JCPDS data was used to simulate a measured profile using a program written in IDL. The required input parameters were the angle (in range 1.7-10°), momentum transfer resolution (range 15-70%) and JCPDS data (d-spacings and intensities for each line).

The JCPDS peak positions and intensities in momentum transfer space were convolved with the angular blurring distribution whose width was governed by the input momentum transfer resolution. The resulting profile was then multiplied by the intensity of the incident spectrum at that point, which was measured with the detector a distance 35cm from the source and a $30\mu m$ platinum pinhole in front of the germanium detector (figure 3.6). The simulated data is then corrected for energy dependent attenuation due to

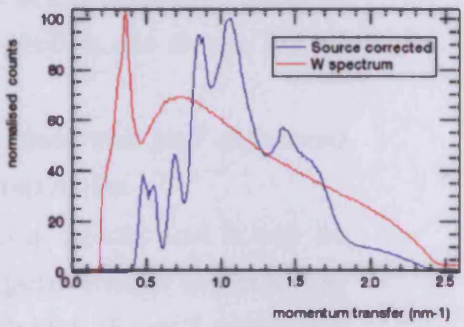
a The thickness of the sample

b The polythene walls of the sample holder and

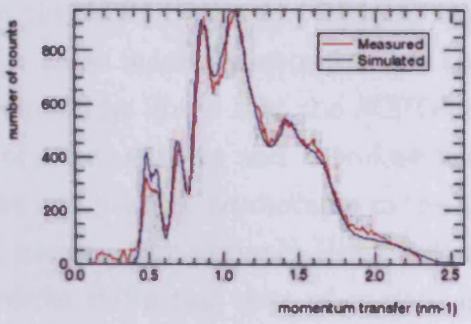
c The air gap between the sample and the detector.



(a) JCPDS data as stated and convolved with system blurring function



(b) Convolved JCPDS data corrected for and compared to incident spectral shape



(c) Attenuation corrected simulated profile compared to measured data

Figure 3.21: The stages of simulating energy dispersive diffraction profiles from angular dispersive data for Cocaine Hydrochloride

The attenuation coefficients were found from XCOM (Berger and Hubbell (1987)), which calculates the attenuation coefficients for a range of materials stored in its database or by elemental composition.

The stages of modeling are demonstrated in figure 3.21. The maximum of the simulated profile was normalised to the peak of the measured data. A comparison between the measured and simulated profiles are shown for all drugs (figure 3.22).

It can be seen that the general shapes of the measured and simulated profiles coincide well, though with some small discrepancies.

The same procedure was repeated for the cutting agents and it can be seen from figure 3.23 that caffeine, lignocaine and paracetamol are modeled well. Phenacetin is modeled well apart from the peak at 1.25 nm^{-1} which was measured to be narrower than the simulated data and the peak at 1.75 nm^{-1} which is missing altogether in the measured data. The simulated profile of glucose is not included as there are many different powder diffraction data sets in the JCPDS files for different types of glucose and the specific type used for these measurements was not known.

It should be noted that the JCPDS data is taken from the experimental work of other authors and therefore has errors associated with it, and the purities and possible adulterants in the samples used for those measurements is not known. For example Hrynchuk et al. (1983) found discrepancies in the powder diffraction data of cocaine free base published by other authors whose data included peaks that could not be accounted for with the known crystal structure of the drug. Canfield et al. (1979) found the same with heroin.

Simulated profiles can be used not only to compare measured diffraction data with that of other authors, but can also be used to predict the profiles measured from the system with a different scatter angle and momentum transfer resolution, or for a substance which has not previously been measured. This could be useful when building up a database of diffraction profiles

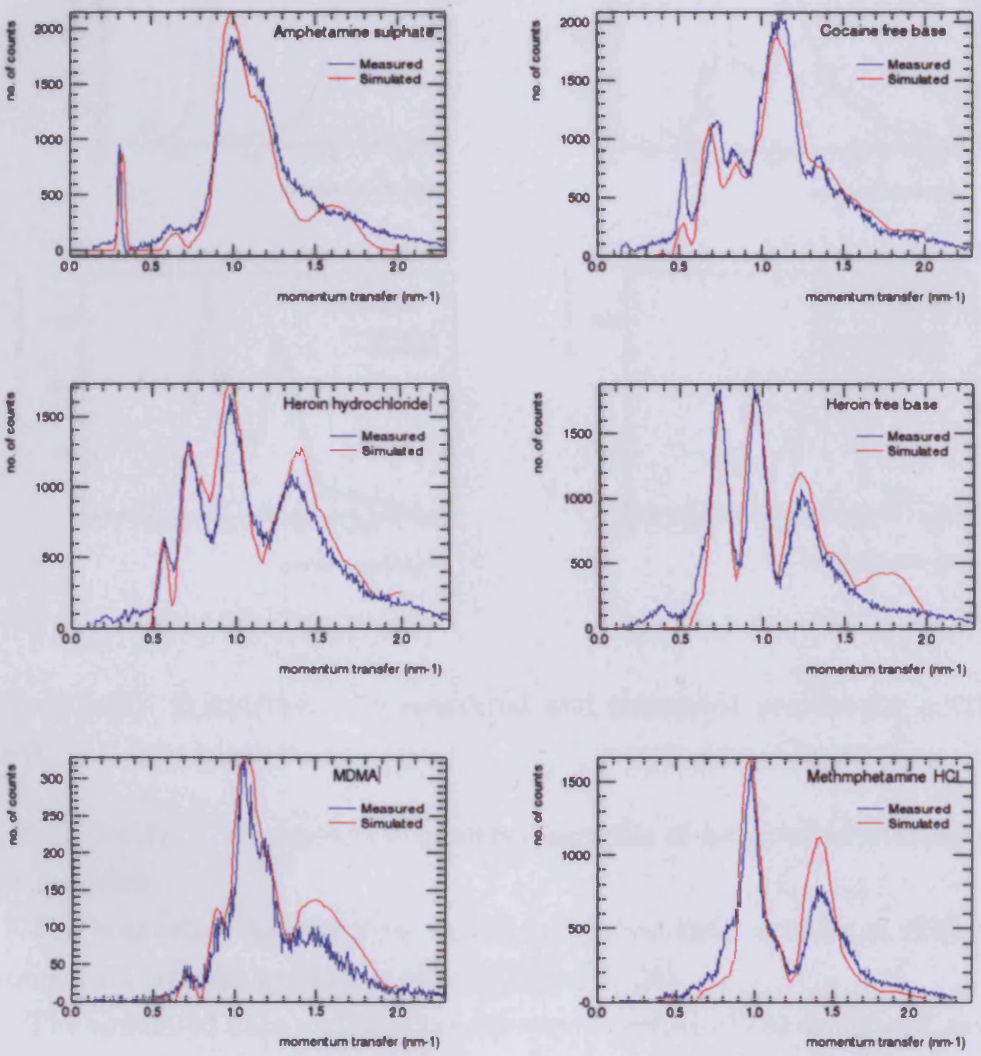


Figure 3.22: Comparisons of measured and simulated profiles for drugs

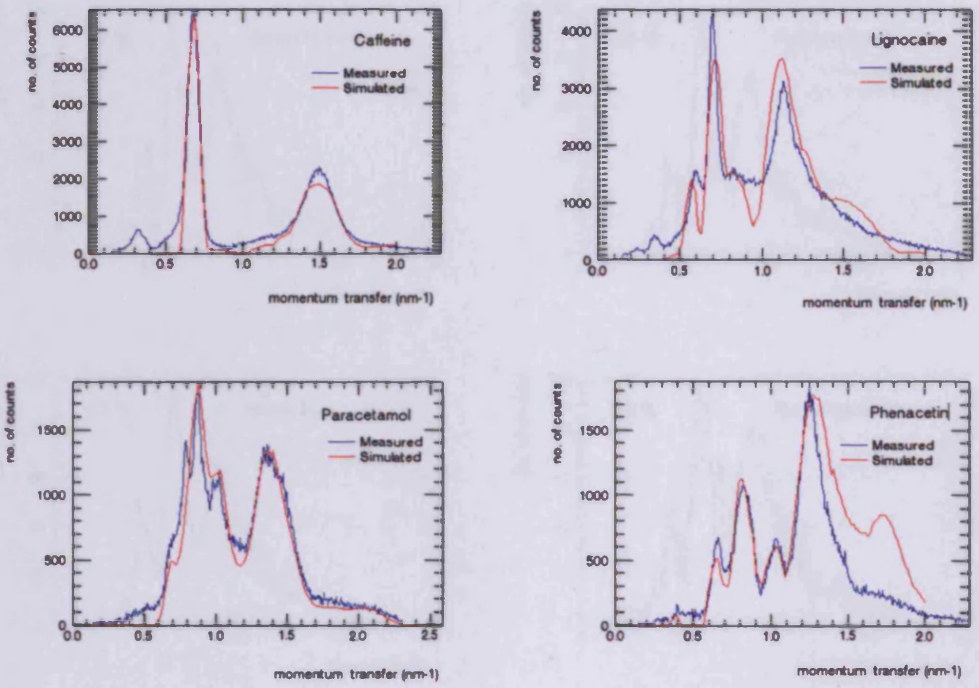


Figure 3.23: Comparisons of measured and simulated profiles for cutting agents

to build further Unscrambler® models if samples of required substances are not available.

The simulation was used to compare the measured profiles at different momentum transfer resolution values (Figure 3.24).

The simulated data matched the measured profiles of the drugs well, noise excluded, with the peak widths and positions being consistent for all values of momentum transfer resolution considered. Where there were variations, such as the additional peaks in the heroin hydrochloride profile at 37 keV and 52 keV predicted from the JCPDS data but not measured, these are most likely due to differences in the sample used for the JCPDS measurements such as impurities. The model otherwise predicts the measured profiles well

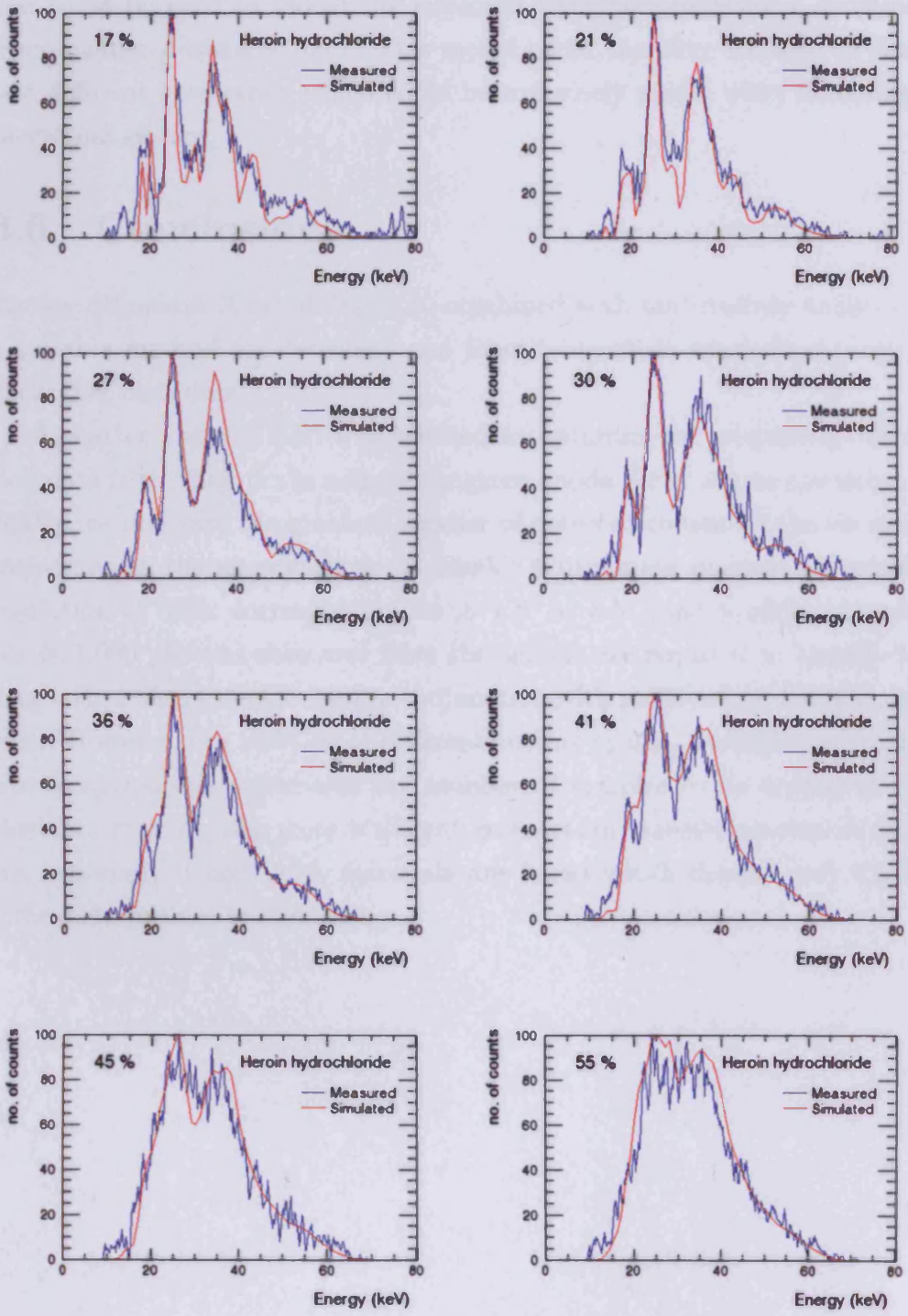


Figure 3.24: Comparisons of measured and simulated profiles for heroin hydrochloride at a scatter angle of 5.5° and a range momentum transfer resolution settings

and could be used to model the expected EDXRD profile for a substance from angular dispersive data. This model could therefore be used to simulate different geometries which could be extremely useful when designing a functional system.

3.6 Conclusions

Energy dispersive X-ray diffraction, combined with multivariate analysis, is a suitable method for detecting and identifying illicit drugs in a range of packaging materials.

A scatter angle of 5.5° was selected as optimum for measuring diffraction data from illicit drugs using a tungsten anode X-ray source operating at 70kVp, as this gave the greatest number of detected counts for the six drugs considered in the energy range 20-60keV. A minimum momentum transfer resolution of 36%, corresponding to ± 1.0° at 5.5°, and a minimum number of 1,000 photons scattered from the sample are required to identify the drugs when the system is used in conjunction with multivariate analysis software. However, the addition of different cutting agents, packaging and other scattering materials increases the number of samples to be separated and therefore may require more stringent momentum transfer resolution criteria, especially if non illicit materials are found which display very similar diffraction profiles to illicit drugs.

Chapter 4

Tissues for diffraction Imaging Studies

The use of multivariate analysis has been demonstrated in the analysis of diffraction profiles for illicit drug detection. The next challenge is to use information from diffraction profiles for imaging, which demands different analysis techniques. Breast tissue was chosen as a suitable application, as there is an inherently low contrast in traditional mammographic images leading to poor discrimination between cancerous and healthy tissues (Johns and Yaffe). As the diffraction profiles of these tissue types have been shown to be substantially different (Kidane et al. (1999)), the additional use of diffraction information could substantially increase contrast, thereby improving the diagnostic power of the technique. However, there are problems associated with imaging tissues. Living tissues begin to decay as soon as they are removed, thus some method of preservation must be found. Freezing can be used, however the samples to be scanned must be defrosted as ice crystals alter the diffraction patterns produced rendering the diffraction data meaningless. The tissue can be kept cool, for example using a peltier cooler, which prevents the tissue decaying but puts constraints on the experimental arrangements. A third option is fixing the tissue, for example using formalin, so that it can be stored and used a room temperature.

This chapter presents work carried out to investigate suitable tissues for use in the diffraction imaging of breast tissue. The equivalence of porcine tissue to human breast tissue is investigated, and is demonstrated to be good. The diffraction profiles of porcine tissues and human breast tissues before and after formalin fixation are then investigated, using both a laboratory X-ray source and synchrotron radiation.

4.1 Formalin fixation

Formalin is used to fix tissues so their structure remains unaltered after they have been removed from the patient. Formalin fixation causes links between proteins, forming a gel in which structures are maintained as close to living

interrelation as possible. Formalin also inactivates enzymes that would start the process of cell self-digestion (autolysis) and prevents decomposition (putrefaction) as bacteria cannot fit around new cross-linked compounds (Hopwood (2002)). The effects of formalin fixation on the diffraction profiles of different breast tissue types need to be investigated and quantified. If there is little or no effect, formalin fixed tissues could be used for a range of X-ray diffraction experiments, removing the need to freeze samples and hold them at a low temperatures whilst measurements are being taken, making the experimental procedure both simpler and less costly. If there is a large or inconsistent discrepancy between the profiles of fresh and formalin fixed tissue, then the measurements made here and in some other low angle X-ray scatter studies which have used fixed tissue require some re-evaluation. It is also important to develop a protocol for tissue fixation for diffraction studies, as fixation is a major cause of inter laboratory variability of data (Werner et al. (2000)).

Investigation of the effects of formalin fixation of tissue on its diffraction profile is important for several reasons:

1. Fixation alters the inter and intra-molecular structure thus may change their relative positions in the tissue, shifting the peaks in the diffraction profile

2. Fixation may enhance / decrease the intensities of peaks and lead to false conclusions about the relative heights of peaks from different tissues

3. Fixation may result in the formation of structures that do not exist in the living cell (fixation artifacts)

4. The presence of excess formalin itself may alter the measured profile

4.1.1 Theoretical effects

Formaldehyde is a gas that is soluble in water and saturates at 37% by mass. This gives a 100% formalin solution. Methanol (10-15%) is often mixed with 37% formaldehyde to inhibit the oxidation of the solution to give formic acid. Formalin is usually used as a 10% solution, therefore containing around 4% formaldehyde, buffered with phosphate to produce a neutral solution. Unbuffered formalin is acidic with a pH of around 3-4 (Puchtler and Meloan (1985)).

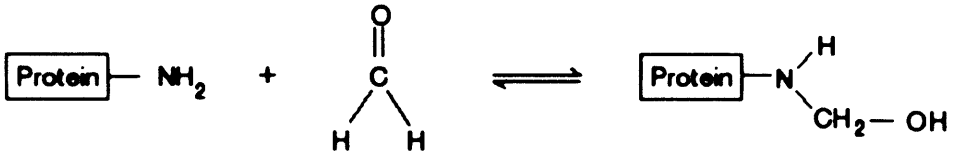
Formalin fixation works by the formaldehyde forming a methylol group ($-CH_2OH$) with amine (NH_2) groups on a protein molecule. The methylol group is highly reactive and binds to other sites on protein molecules (figure 4.1). This produces methylene bridges which link protein molecules either to themselves, giving mechanical stability, or to neighbouring protein molecules forming crosslinks (Metz et al. (2004)). Lipid content of tissues has been shown to not be affected by formalin fixation (Hickey and Hukins (1979) and Boskey et al. (1982)), thus adipose tissue, which consists of 80% lipid, should not be affected by fixation.

Formalin fixation is slow, though the formalin penetrates the tissue rapidly. The distance of penetration of the formaldehyde/ methylene glycol into tissue (D) follows the relationship

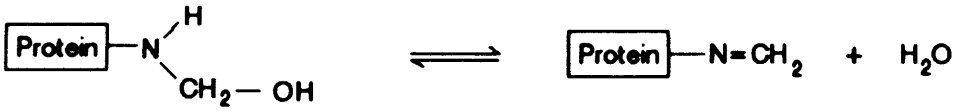
$$D = K\sqrt{T} \tag{4.1}$$

where K is a constant depending on the tissue and fixative, T is the time in hours and D is given in mm (Medawar (1946)). K has been estimated by different authors to be 0.78 (Hopwood (2002)), 3.6 (Baker (1958)) and 5.5 (Medawar (1946)) for formaldehyde in human tissue. Using the middle value of these figures suggests it would take 2 hours for the formalin to penetrate 5mm into tissue.

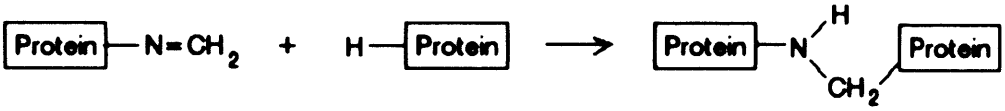
The fixation time is not however governed by penetration time. Instead it is controlled by the series of reactions in which the cross links are formed,



(a) The formaldehyde forms a methyl adduct on amino or thiol groups of a protein (a reversible reaction)



(b) The methylene adducts are then partially dehydrated to give a Schiff base (a reversible reaction)



(c) The Schiff bases can form cross links with one of several amino acid residues on either the same or a different protein (a non-reversible reaction)

Figure 4.1: Chemical reactions involved in formaldehyde fixation

which are shown in figure 4.1.

Formaldehyde, when dissolved in water, rapidly undergoes a reversible hydration reaction forming methylene glycol. Although this reaction is reversible, the equilibrium lies far in favour of the methylene glycol. As formaldehyde binds to proteins in the tissue, methylene glycol is dehydrated to produce more to restore equilibrium (equation 4.1.1).

$$CH_2(OH)_2 \rightleftharpoons HCHO + H_2O$$

$$Methylene\ glycol \rightleftharpoons Formaldehyde + Water \tag{4.2}$$

The rate of this reaction along with those shown in figure 4.1 control the time taken to completely fix the tissue. It has been shown using C-14 labeled formaldehyde that formaldehyde continues to bind to tissue for up to 24 hours at room temperature, thus this gives a minimum fixation time (Fox et al. (1985)). As this experiment used a $16\mu m$ tissue slice it was clearly the fixation reactions that affected the time taken and not the formaldehyde penetration time. Helander (1994) conducted a similar experiment using larger tissue sections ($4mm^3$), and found that the isotope binding reached a plateau after approximately 25 hours, thus the total time required to fix a piece of tissue is 24 hours plus the time taken for the formalin to fully penetrate the tissue.

Formalin fixation may have a range of effects on the structure of tissue that alter the measured diffraction pattern. Firstly water in the tissue is replaced to some extent by formalin, which as the peak at 1.56 nm^{-1} is predominantly due to water, could reduce the height of the peak in the measured profile. Secondly it can alter the alpha helix structure of proteins in the tissue, which may affect the positions of molecules within the tissue, and thirdly the cross links may in themselves scatter photons. These effects need to be quantified if the diffraction measurements made on fixed tissue are to be used to give information on relative peak heights in healthy and

diseased tissue.

4.2 Previous X-ray diffraction work using fixed tissue samples

The effects of formalin fixing on diffraction profiles are not well known, with there being some conflicting results in the published literature. Peplow and Verghese (1998) measured the diffraction profiles of samples of human breast tissue fixed in formalin and of formalin alone. They found the molecular form factors of formalin to be very similar to those of water, and assumed that fixing tissue in this manner would have no significant effect on the diffraction profile of the breast tissue as formalin merely replaces the water content of the tissue with formaldehyde. However the fixing process alters the supra-molecular structure of the proteins (Fox et al. (1985)) and therefore the diffraction pattern produced by the tissue. It has been shown that the fixation in buffered formalin of breast tissue obtained from biopsies alters the SAXS patterns of the tissue so as to render the signals from carcinoma and normal tissue almost indistinguishable in the momentum transfer region 0.03-0.06 nm^{-1} (Changizi et al. (2005b)). These are much lower momentum transfer values than are being considered in this work, and are caused by larger structures such as the structure of the collagen fibrils in the tissue. Formalin fixation is known to cause collagen molecules to become more tightly packed within a fibril though it does not affect the arrangement of fibrils (Hickey and Hukins (1979)). Fernandez et al. (2005) found that the difference in d-spacing between healthy and invaded tissue was 0.2% in samples fixed in formalin for 2 weeks, compared to 0.76% in fresh samples, though for samples fixed for 10 weeks the difference returned to 0.7%. This suggests fixation time is also an important factor which needs to be considered.

Castro et al. (2005b) have performed diffraction CT on formalin fixed breast tissue, so as not to introduce a preferred orientation into the samples

by freezing, though they do not compare it to fresh tissue. They measured the peaks in the diffraction profiles to be 1.11 nm^{-1} for adipose tissue and 1.57 nm^{-1} for infiltrating ductal carcinoma, agreeing well with the literature and data taken in this study. Their data is shown in figure 4.2. It can be seen that the relative intensity of the peak at 1.57 nm^{-1} relative to that at 1.11 nm^{-1} is lower than that measured in other studies of fresh tissue (e.g. Kidane et al. (1999)). This suggests that formalin fixation has not affected the adipose content of the tissue as theorised, whilst it has reduced the intensity of the carcinoma peak, though not its position.

A comparison between fresh and formalin fixed tissue in the momentum transfer region 0.5-3.0 nm^{-1} has not so far been reported in the literature. It is this region that is of interest as it encompasses the momentum transfer values which show the greatest differences between healthy and neoplastic tissues, thus is investigated here.

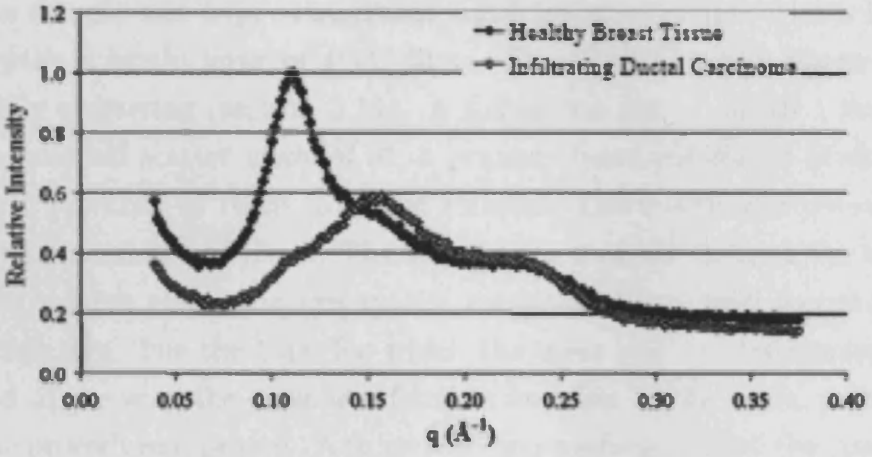


Figure 4.2: Experimental scatter signatures for healthy adipose breast tissue and infiltrating ductal carcinoma, both fixed in formalin (Castro et al. (2005b))

4.3 Experimental methods

Two sets of measurements were taken, one laboratory based using the tissue scanning system described in detail later in section 5.1.3, on pork samples, and a second using a synchrotron source on pork and human breast tissue samples.

4.3.1 Laboratory measurements

Pork was investigated as a breast tissue substitute material, since Peplow and Verghese (1998) suggested that the diffraction coefficients of porcine tissues are similar to those of breast tissue, with pork fat being similar to human adipose tissue and pork muscle similar to glandular or tumourous tissue. A section of a pork chop containing subcutaneous fat, a fatty region inside the tissue, and some muscle tissue was selected. A 2cm x 2cm slice was cut which was 1.4cm thick, giving a penetration time of 6.4 hours using equation 4.1.

The sample was kept refrigerated until mounted, where it was held in place with a single layer of PVC film. The PVC film was known to be negligibly scattering (section 3.14). A diffraction cell of design 1 was used with a nominal scatter angle of 6°, a primary beam collimator of diameter 0.5mm a distance of 16cm from the sample. The source was operated at 60kV with a current of 10mA. The sample was scanned through the system, in 2mm x 2mm steps, the low spatial resolution being used to reduce the scan time and thus the time for which the meat was not refrigerated. At the end of the scan the meat was fixed in formalin for 24 hours, patted dry and the procedure repeated. A third scan was performed after the tissue had been stored in formalin for 90 hours and comparisons between the diffraction profiles made. The system was re-aligned between 0 and 24 hours as it had been moved, but was not altered between 24 and 90 hours.

4.3.2 Synchrotron measurements

These were conducted at station 16.3 at Daresbury laboratories (Warrington, UK) with Dr Silvia Pani. The samples were visually inspected by a pathologist and found to be:

1. Pork muscle
2. Human Breast Tissue (fat)
3. Human Breast Tissue (fat)
4. Human Breast Tissue (tumour)
5. Human Breast Tissue (tumour)
6. Human Breast Tissue (benign lesion)
7. Human Breast Tissue (fibrosis)
8. Human Breast Tissue (fibrosis)

They were contained in straws and kept at -80° until use, when they were defrosted and mounted in a brass holder held at -1°C by a peltier cooler. The tissue was measured to be at a temperature about 3 degrees above that of the holder, thus this held the tissue at a temperature above freezing, but cold enough to prevent tissue decomposition. The equipment was a standard diffractometer set-up with a germanium detector (Canberra, CT, USA) calibrated with Americium - 241. The beam energy was 15keV and the detector scanned in 0.5 ° steps between 3 and 19 degrees, and 1 ° steps between 20 and 60 degrees. The data at high momentum transfer values (4-6 nm^{-1}) was used to normalise the data to the IAM values (see section 1.2.3), with tissue compositions approximated using XmuDat (Nowotny (1998)), equating fibrosis, benign lesion and tumour to soft tissue, and pork meat to muscle (Pani (2007)).

The profiles were created by summing the counts in the channels around 15keV and 5keV, the lower of the two energies being necessary due to the escape peak of germanium. The samples were held in plastic cylinders and a complete set of measurements were made with an empty sample holder which were subtracted from the normalised sample measurements to correct for scatter from the plastic. All profiles were normalised in intensity using the beam currents at the start, middle and end of each measurement (Pani (2007)).

4.3.3 Results

Laboratory measurements

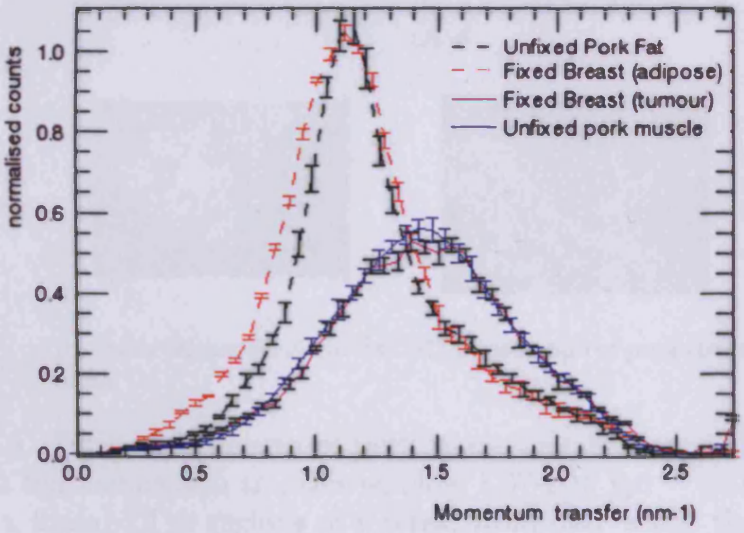
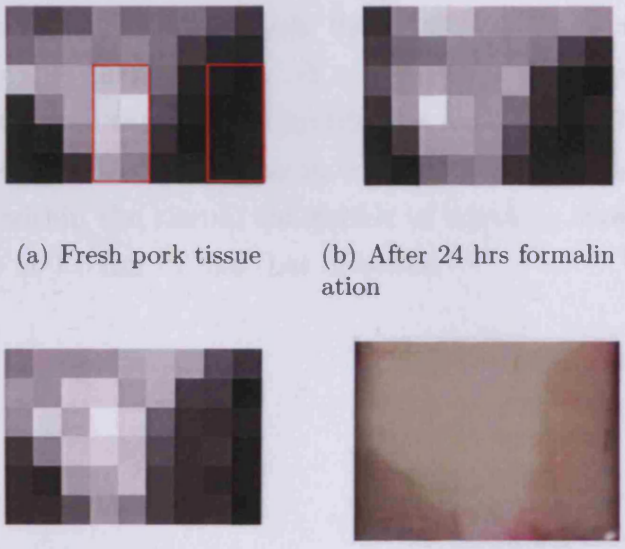


Figure 4.3: Comparison or the diffraction profiles of unfixed pork and fixed human breast tissue

The first stage was to investigate if porcine tissue could indeed be used as a substitute for human breast tissue. A plot comparing the measured

diffraction profiles from fixed human breast tissue and fresh pork fat and muscle is shown in figure 4.3. It can be seen that the diffraction profile of tumourous tissue exactly matches that of pork muscle, and that of human adipose tissue is very similar to pork fat, though the pork adipose peak is sharper, most likely due to the fact that the human breast tissue contains other tissue types thus is less well ordered. This comparison indicates that porcine tissues can usefully be used as substitutes for human breast tissue.



(a) Fresh pork tissue

(b) After 24 hrs formalin fixation

(c) After 90 hrs formalin fixation

(d) Photograph of pork tissue

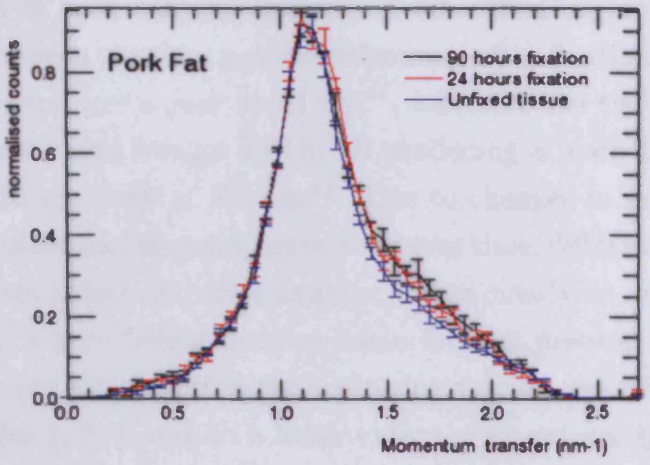
Figure 4.4: Diffraction images of pork tissue (created using the integrated counts in the momentum transfer window 1.07-1.13 nm^{-1} with a pixel size of 2mm x 2mm). The regions of interest from over which the spectra are averaged are indicated in a)

Diffraction images of the pork samples were created by summing the counts in the momentum transfer region 1.07 - 1.13 nm^{-1} (around the adipose peak, see figure 4.3) and assigning the total a proportional grey level. The resulting images are shown in figure 4.4, with the bright regions indicating

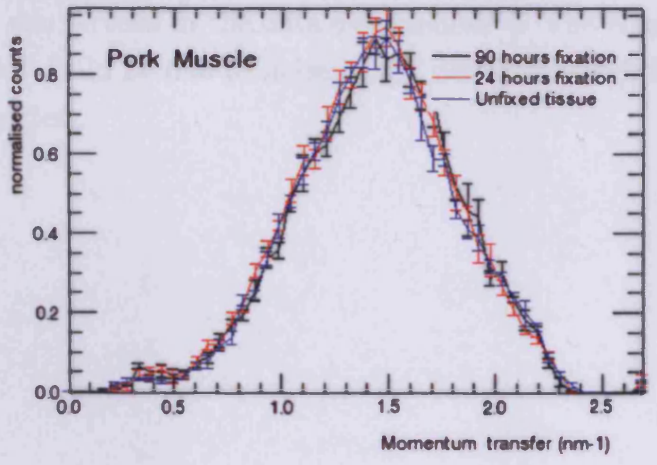
a high fat content and the dark regions low fat, corresponding to muscular regions in this case.

As the sample was removed and remounted in between measurements, there were some small discrepancies in the positions of the different regions in each image. However, the muscular and fatty regions are clearly identifiable in all three images. Regions of interest were selected, one in a fatty region, and one in a muscular region, and the profiles from these regions averaged to give a mean region profile.

The profiles measured when fresh, fixed for a 24 hours and fixed for 90 hours are compared in figure 4.5. It can be seen that they all have the same shape, though in the momentum transfer region 1.4-2.0 nm^{-1} the fixed samples give rise to a slight increase in coherent scatter, possibly due to the formalin itself within the tissue, the profile of which is shown in figure 4.9 and has a peak at 1.5 nm^{-1}, like that of water.



(a) Pork adipose tissue tissue when fresh and after 24 hrs and 90 hrs formalin fixation



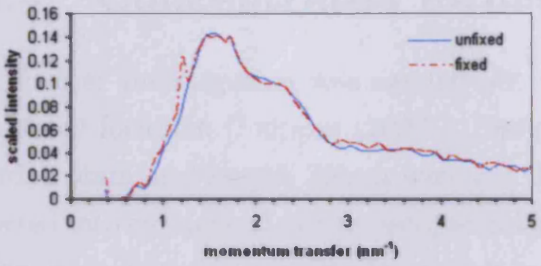
(b) Pork muscle tissue tissue when fresh and after 24 hrs and 90 hrs formalin fixation

Figure 4.5: Comparison of experimental scatter signatures (6°) between fresh and formalin fixed pork adipose and muscle tissue. The error bars are the standard deviation across the number of points in the region considered

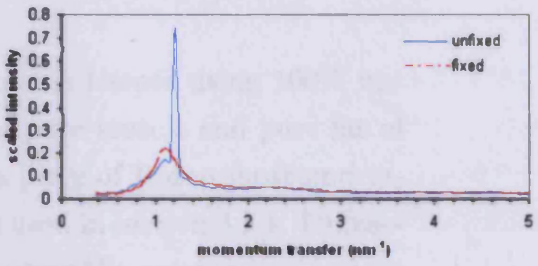
Synchrotron measurements

The profiles of each sample, before and after fixation are shown in figure 4.6. It can be seen that the profiles before and after fixation are very similar. Fatty tissue produces a peak at 1.1 nm^{-1}, with all other tissues (pork muscle, tumour, fibrosis and benign lesion) all producing a main peak at 1.6 nm^{-1} with a subsidiary peak at 2.1 nm^{-1}. Due to changes in sample orientation within the holder and the samples settling over time, different sample volumes were examined before and after fixation, as can clearly be seen from the plots of sample 6, where before fixation some fat was present in the scattering volume whereas after fixation the scattering volume was filled by the lesion.

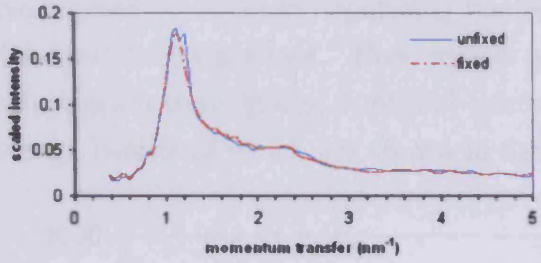
In samples 1, 2, 7 and, to a lesser extent, 4 an extremely sharp peak can be seen at 1.2 nm^{-1}. This is present in the unfixed profile only for samples 2, 4 and 7 and in the fixed profile only in 1. This indicates the peak is not due to the fixation process. This peak can be distinguished from the fat peak at 1.1 nm^{-1} as can be seen in the data from sample 4. The origin of this peak is unknown, but could be due to noise in the detector, which started vibrating at certain angles.



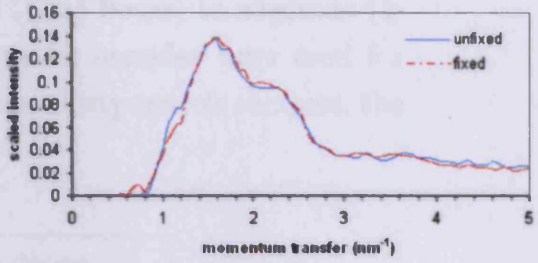
(a) Sample 1. Pork muscle



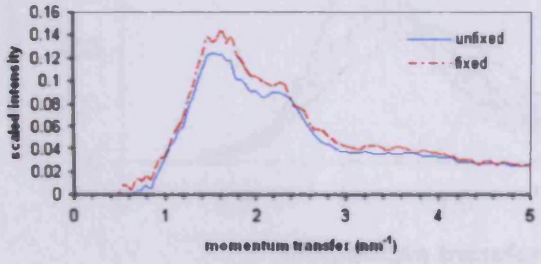
(b) Sample 2. Fat



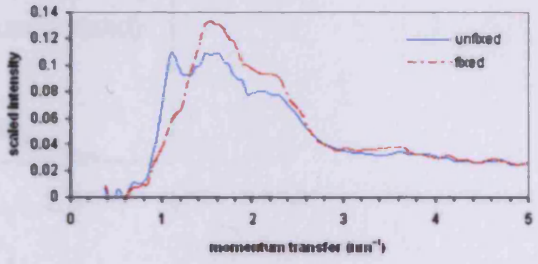
(c) Sample 3. Fat



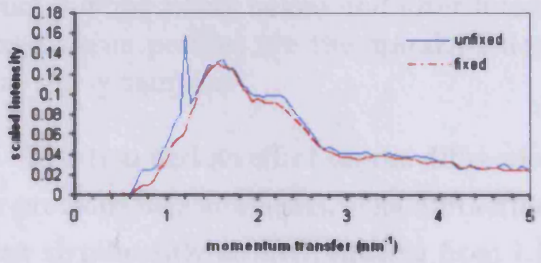
(d) Sample 4. Tumour



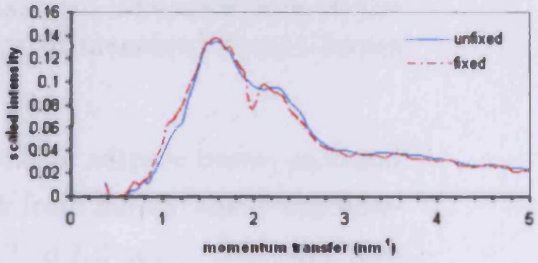
(e) Sample 5. Tumour



(f) Sample 6. Benign lesion



(g) Sample 7. Fibrosis



(h) Sample 8. Fibrosis

Figure 4.6: Comparison of tissue samples before and after 24 hours formalin fixation, measured using synchrotron radiation

4.4 Concentrated formalin

A further investigation was carried out on porcine tissues using 100% unbuffered formalin (Skipper (2007)). Samples of pure muscle and pure fat of width 5mm and length 20mm were cut from a piece of frozen meat and inserted into cylindrical plastic sample holders as used in section 3.2.1. Diffraction measurements were made on each sample using the moving arm system set at 5.5° both before and after fixation. A range of fixation times were investigated (4-48 hours) spanning inadequate ($\leq$ 24 hours) to adequate ($>$ 24 hours) fixation times. Five muscle and five fat samples were used for each fixation time, giving a total of thirty fat and thirty muscle samples, the average results of which are shown in figure 4.7.

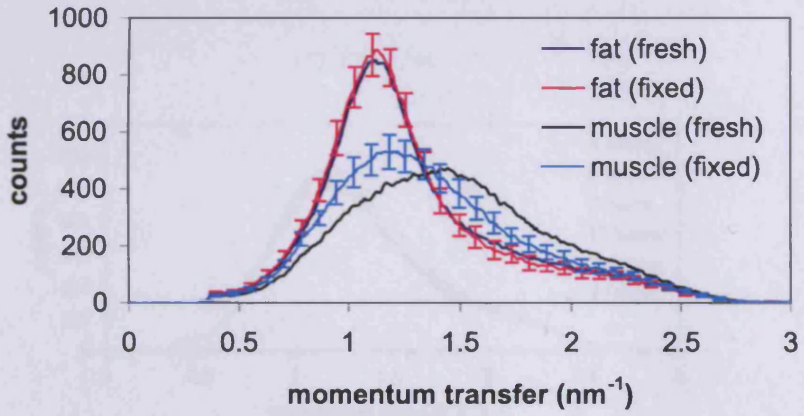
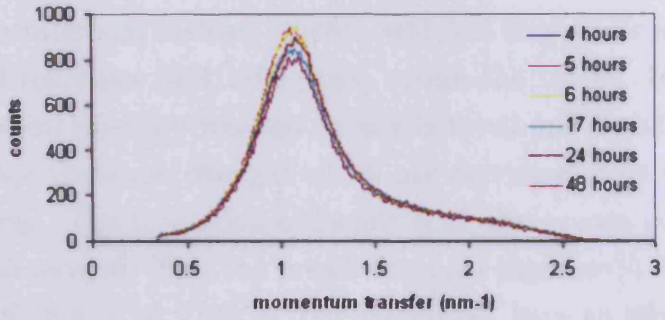
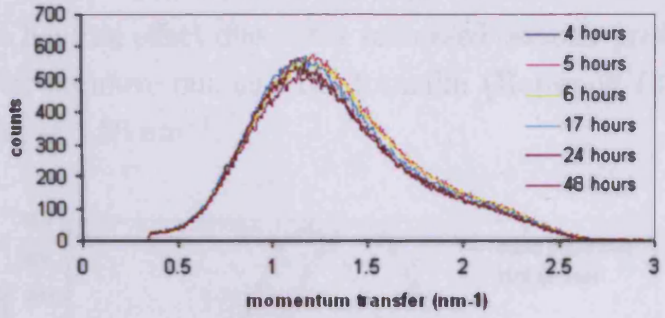


Figure 4.7: Average diffraction profiles of thirty muscle and thirty fat sample from porcine tissue before and after formalin fixation. The error bars on the fixed tissue profiles are the standard deviations in measured counts across the thirty samples.

Fixation had no effect on the diffraction profile of adipose tissue, as found in previous measurements. The diffraction peak from muscle tissue was however significantly altered, shifting from 1.5 nm^{-1} to 1.2 nm^{-1}. The peak also has a slightly higher intensity and the intensity in the region above the peak is diminished.



(a) Pork fat



(b) Pork muscle

Figure 4.8: Comparison or the diffraction profiles of fixed pork fat and fixed pork muscle with fixation times of 4-48 hours

There are several possible causes of this. The fixation time did not affect the shape or position of the peaks, as can be seen in figure 4.8, though the intensities are not the same due to variations in sample volume. Since the tissues would have ranged from unfixed to completely fixed, the similarity in the profiles suggests that the shift after fixation is due to another factor, such as the pH of the fixative. Unbuffered formalin has a pH of around 3-4 (Puchtler and Meloan (1985)), which is very acidic and could alter the chemical reactions that take place within the tissue. Hopwood (2002) suggests that satisfactory fixation occurs between pH 6 and 8 and that outside this range there are changes which are detrimental to the ultrastructure of the tissue. The formation of formic acid also occurs when formaldehyde reacts with oxygen. This can create formalin pigment in blood rich tissues such as liver (Fox et al. (1985)), but should not have an effect on the tissues studied here. However, it can lower the pH and is present in formalin, especially if old as oxidation occurs naturally over time. The high concentration could also have an effect due to the increased osmotic pressure on the cells, causing water to move out into the formalin (Hopwood (2002)), thus reducing the peak at 1.56 nm^{-1}.

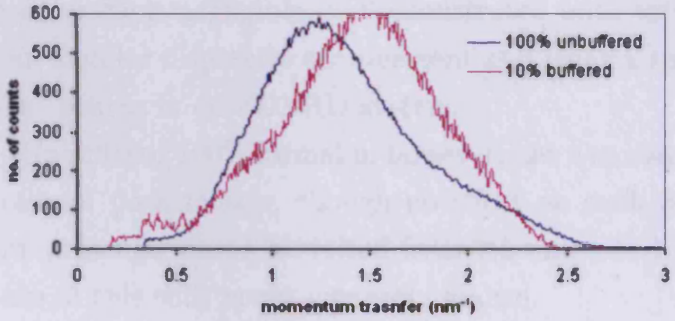


Figure 4.9: Comparison of the diffraction profile of 100% unbuffered formalin measured using the moving arm system with the source operating at 70kV with the profile of 10% buffered formalin taking on the block collimator system with the source operating at a potential of 60kV.

The profiles of formalin when in a 10% buffered solution and an unbuffered 100% solution are also substantially different, with the 10% buffered formalin producing a peak similar to water at 1.56 nm^{-1} and the 100% unbuffered formalin producing a peak at 1.25 nm^{-1} (figure 4.9). However, it is unlikely that a contribution to the diffraction profile from the formalin itself caused the shift, as there was no difference in the corresponding fat profiles, and the reduction in intensity in the region above 1.5 nm^{-1} in the muscle profiles indicates it is a change in the structure of the tissue and not just additional scatter from the formalin that causes the shift.

4.5 Conclusions

Porcine tissues provide suitable breast tissue substitute materials for energy dispersive X-ray diffraction measurements in the momentum transfer region 0.5 - 2.5 nm^{-1}. Pork fat provides a close approximation of human breast adipose tissue whilst pork muscle produces the same diffraction profile as breast carcinoma.

The effect of buffered 10% formalin fixation on the diffraction profile of pork muscle is negligible, as demonstrated both using synchrotron radiation in an angular dispersive arrangement and using a tungsten anode laboratory X-ray source in an EDXRD system.

Unbuffered 100% formalin, however, has a marked effect on the diffraction profile of pork muscle, though no effect on pork fat. The diffraction peak from muscular tissue is shifted from 1.5 nm^{-1} to 1.2 nm^{-1} when fixed. The cause of this shift is not currently known.

Thus the adipose peak in all tissues is unaffected by formalin, regardless of concentration. Glandular, tumourous and fibrotic breast tissue and porcine muscle are unaffected by fixation in 10% buffered formalin, though the diffraction peak from porcine muscle was affected by the 100% unbuffered

formalin solution.

Since fixation in 10% buffered formalin did not affect the measured diffraction profiles, this method of tissue fixation could be used in future EDXRD studies in the momentum transfer range 0.5-2.5 nm^{-1}.

The effect of pH, formalin concentration and fixation time on tissues need to be measured to develop a protocol for correctly fixing samples for EDXRD measurements.

Chapter 5

Breast Tissue Imaging

This chapter concentrates on the use of X-ray diffraction as an imaging modality, both on its own and combined with traditional X-ray transmission images. Six pieces of formalin fixed breast tissue (labeled A-F) were scanned through the EDXRD system and and images generated using different regions of the measured diffraction profiles.

Three large regions of tissue were examined and seven small regions from the six pieces. The images were corrected for attenuation using transmission images. The contrast between healthy and tumourous regions in the images are optimised and possible combinations of diffraction data with standard mammographic images considered.

5.1 Experimental methods

5.1.1 Experimental set up

Cell design 2 was used for scatter collimation, with a nominal angle of 6° which was found to be optimal by Kidane et al. (1999). The X-ray source was operated at 60kV.

5.1.2 Large regions

For the large area images an X-ray tube current of 3mA was used and the incident X-ray beam was collimated to a diameter of 1mm, with the collimation providing a momentum transfer resolution of 32%. The air kerma was measured with the tissue removed and an ionisation chamber in its place attached to an electrometer (Keithley, 35050A). It was found to be 16.3 μGy per second at 60kV and 3mA. Data was collected for 5s at each point in the tissue.

The samples were held at their edges between two sheets of Perspex, on a stage that could be raised and lowered manually, which in turn was fixed to a translational stage that was motorised and controlled by the PC using LabView software. The sample was moved by 1mm increments both

vertically and horizontally. Three large regions in three different tissues were imaged in this way (Tissues A, B and C).

5.1.3 Small regions

For imaging the small regions, a lead mask was attached to the tissue and mounted with the lead on the side facing the source. This mask defined the area to be imaged, which was a square 15mm x 15mm. The samples were scanned in 1mm steps both horizontally and vertically through the beam, with data collected 1mm into the lead to aid the registration of the diffraction image with the images used for correction (see section 5.4.1). The motion control was entirely PC based using the translation stages and actuators described in section 2.1.3.

The addition of motorised scanning vertically as well as horizontally required the primary collimation to sample distance to be increased from 2 to 8 cm. The collimation was thus selected to limit the sample area illuminated by the beam so that the tissue could still sensibly be scanned in 1mm steps. The beam diameter was selected to be 0.5 mm giving a width of beam at the sample of 1.4 mm compared to 1.3 mm for the large region scanning set-up. Altering the collimation in this way improved the momentum transfer resolution of the system slightly to 25%. The reduction in beam diameter by half caused a reduction in flux by a of factor 4, thus the tube current was increased to 10mA so enough photons were collected to give statistically good data without increasing the data collection time.

5.1.4 Tissue samples used

The samples used were taken from patients who had undergone mastectomies, fixed with formalin, and sealed in polythene packs. The samples were of variable thickness, with a maximum thickness of 1cm. For details of the effects of formalin fixation refer to chapter 4.

5.2 Diffraction profiles

Examples of the raw spectra obtained can be seen in figure 5.1, which show the difference in the spectra from a region of tissue of normal appearance and region visibly containing a tumour. The amount of scatter from the cancerous regions was much lower overall. The reduction in total coherent scatter from the diseased tissue suggests a reduction in the structural order of the tissue. This has been suggested to be due to the decrease in structural order of collagen molecules in malignant tissues (Lewis et al (2000)). It can be seen that the healthy tissue has a peak in its diffraction profile at 1.1 nm^{-1}, which is due to adipose tissue. Invaded tissue is often marked by an absence of adipose tissue. The tumorous tissue has a peak at 1.5 nm^{-1}, which is lower than that measured by other authors (1.6 nm^{-1}, Kidane et al. (1999) and Castro et al. (2004)), though the data here has not been corrected for incident spectral shape which may account for this. Geraki et al. (2004) also found the EDXRD peak from tumourous tissue to be at 1.5nm^{-1}, though they attributed this to inaccuracies in their selection of scatter angle.

5.3 Diffraction images

Diffraction images were constructed by summing the counts in a momentum transfer window and assigning the point in the images a grey level accordingly. Appropriate momentum transfer regions were selected by maximising the contrast between regions in the images. This was done by selecting ROIs in both healthy and tumourous regions, altering the number of channels summed and their central point, and calculating the contrast between regions in each case. Figure 5.2 shows the results for summations of 10 -100 channels.

Contrast between regions is calculated using the equation

$$C = \left[\frac{\chi_h - \chi_d}{\chi_h + \chi_d}\right] \tag{5.1}$$

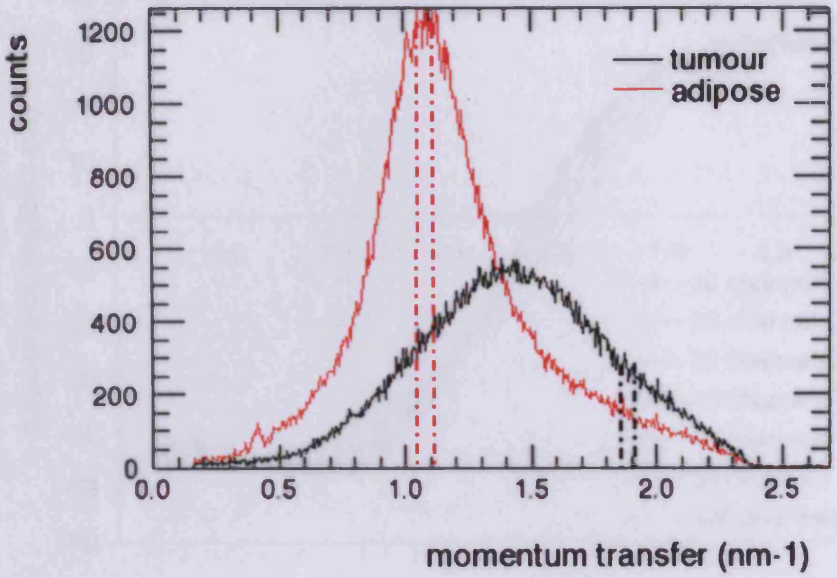
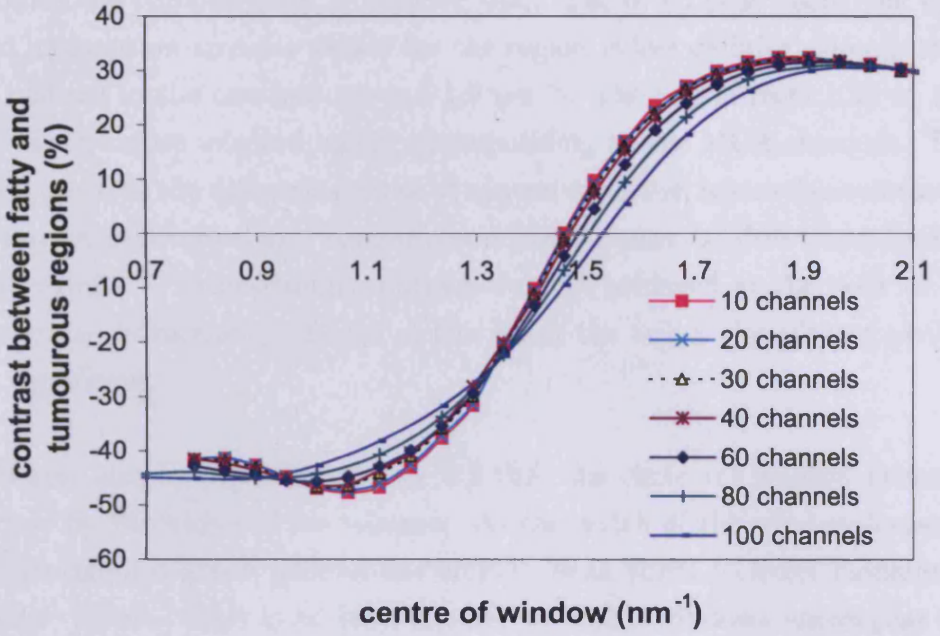


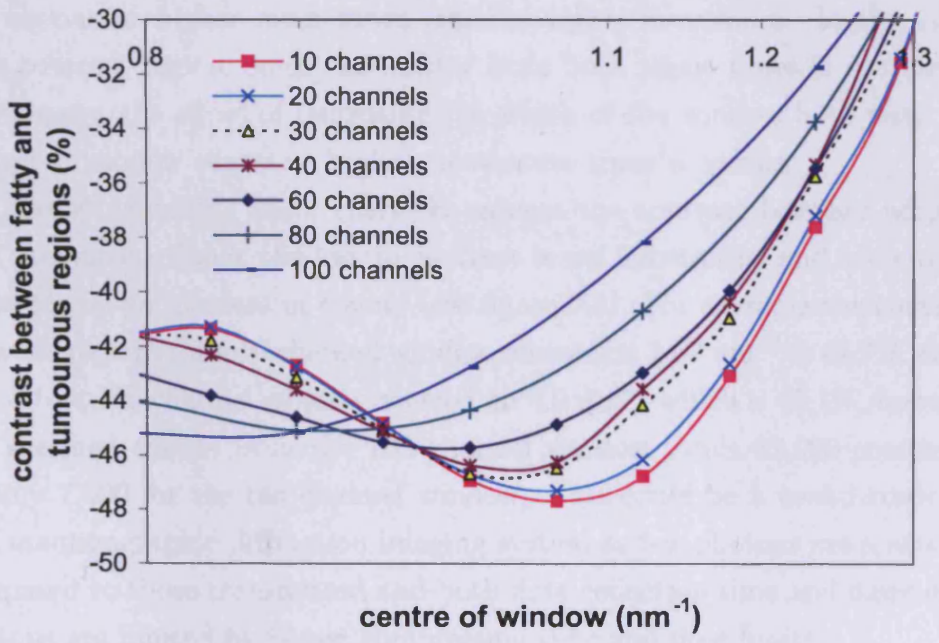
Figure 5.1: Diffraction profiles of fatty and tumourous formalin fixed human breast tissue with momentum transfer windows 1.04-1.10 nm^{-1} and 1.84-1.90 nm^{-1} indicated

where χ_h and χ_d are the transmitted or diffracted intensities of healthy and diseased tissue respectively. The mean intensity values of regions of interest were found using IDL software.

Figure 5.2 indicates that the maximum contrast between healthy and cancerous regions is achieved when a ten channel window is selected with a centre at 1.07 nm^{-1}, corresponding to a momentum transfer window from 1.04 to 1.10 nm^{-1}. This is the region just below the diffraction peak from adipose tissue. The optimum contrast is achieved just below rather than at the peak as fewer photons are scattered by tumourous tissue there. Using this region of the diffraction profile produces images in which healthy, fatty regions are bright and tumourous regions and areas of thinner tissue, which give less total scatter, are darker.



(a) The image contrast between fatty and tumourous tissue as a function of the central momentum transfer value of window when summing 10-100 MCA channels



(b) a) enlarged to show the momentum transfer region 0.8 - 1.3 nm^{-1}

Figure 5.2: The image contrast between fatty and tumourous tissue as a function of the central momentum transfer value of window

Since the tumour peak is broader than that of adipose tissue the optimum momentum transfer values for the region is less definite, though there is a plateau in the contrast around 1.9 nm^{-1}. The region from 1.84 to 1.90 nm^{-1} is therefore selected, again corresponding to ten MCA channels. This region is above the diffraction peak of cancerous tissue, where the contrast in the images is reversed and tumourous regions appear brighter than healthy fatty regions. The maximum contrast was not achieved at the peak of the tumourous diffraction profile as at this point the tail of the adipose peak is still significant.

It can also be seen from figure 5.2 that the choice of window centre is affected by the width of the window. As the width of the window increases the maximum contrast around the adipose peak shifts to lower momentum transfer values. This is to be expected as wider windows encompass the regions around the scatter peak and the regions at lower momentum transfer values the adipose peak have less scatter from tumourous tissue, whereas the regions at higher momentum transfer values have more. In the other high contrast region, since the scatter from both tissue types is decreasing in intensity, the effect of increasing the width of the window is to push the optimum window centre to higher momentum transfer values.

Though summing more channels reduces the contrast between adipose and tumourous tissue, the loss in contrast is not substantial and is compensated for by an increase in counts (see figure 5.3). For example the contrast between regions for a 10 channel window centred on 1.07 nm^{-1} is 47.7%, compared to a 100 channel window centred on 0.9 nm^{-1} which is 45.1%, however the summed counts from the 100 channel window totals 81,000 compared to only 7,500 for the ten channel window. This could be a useful trade off in a mammographic diffraction imaging system as few photons are scattered compared to those transmitted and both data collection time and number of photons are limited by breast compression time and dose limits.

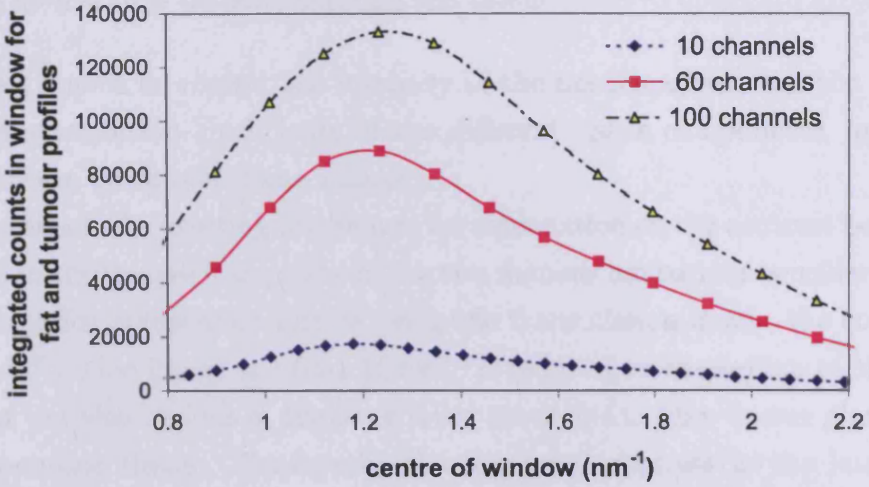


Figure 5.3: The sum of the integrated counts in the momentum transfer windows for the diffraction profiles of adipose tissue and tumourous tissue as a function of window centre for window widths 10-100 MCA channels.

In addition to the two momentum transfer windows, two further parameters were calculated for each point to create images: the ratio between the counts in the two aforementioned momentum transfer regions, and the sum of the counts. The ratio image was intended to improve contrast as healthy tissue had high counts in one region and low in the other, whilst tumourous tissue gave the opposite, and the sum image gave an indication of tissue thickness.

5.4 Correcting diffraction images

In order to examine the true contrast of diffraction images the effects of attenuation need to be removed from the images. Attenuation of the radiation arises as the photons pass through the tissue both before and after scattering. The attenuation is dependent on two variables (see equation 1.1).

1. The attenuation coefficient of the tissue

2. The distance traveled through the tissue

Thus a need to correct the intensity of the images arises from the difference in attenuation coefficients of the different tissue components, and the non-uniform thickness of the sample.

The effect of correcting the images for attenuation on the contrast between regions in the images is opposite in the two momentum regions considered. In correcting for attenuation effects using the transmission image, the contrast in the diffraction image at 1.04-1.10 nm^{-1} is reduced, as the regions of high fat content are also regions of tissue of lower attenuation than denser glandular or tumourous tissue. Conversely, the measured contrast in the image at 1.84-1.90 nm^{-1} is increased upon attenuation correction.

5.4.1 Correction using transmission images

X-ray transmission information has been used to correct for attenuation effects in SPECT (Blankespoor et al. (1996)), PET (Kinahan et al. (1998)) and XRDCT (Kleuker et al. (1998)) using CT data. Since the diffraction images obtained here are 2D rather than 3D, 2D transmission images were used to provide the attenuation corrections.

Transmission images of each tissue sample were taken using a fine focus molybdenum target source (XTek, Tring, UK), operated at 33kV and 0.7 mA, and an amorphous silicon flat panel imager (Dpix Flashscan 30) which had a pixel size of 127 x 127μm and a phosphor coating (Kodak, GADOX screen). ViVA software (Varian Medical Imaging inc.) was used to perform corrections using dark field and flat field images. The source to sample distance was approximately 50cm and the sample to detector distance 10cm.

The tissue images were divided into boxes whose sizes were selected to match the spatial resolution of the diffraction images (approximately 10 by 10 pixels for the 1x1mm diffraction image pixels), and the mean count from each box used to provide the grey level for the pixellated transmission image

(figure 5.5). The pixellated transmission image was then used to correct both diffraction images.

Consider the situation shown in figure 5.4. I_{diff} is the X-ray intensity measured by the scatter detector, $I_o(E)$ the incident photon energy distribution and σ_s the probability of scatter into the required solid angle.
For photons scattered at A

$$I_{diff} = \sigma_s I_o(E_{diff}) e^{-\mu(E_{diff})\times l}$$

For photons scattered at B

$$I_{diff} = I_{trans}(E_{diff}) e^{-\mu(E_{diff})\times t} \sigma_s$$

For the transmitted photons

$$I_{trans} = I_o(E_{trans}) e^{-\mu(E_{trans})\times t}$$

Given that the angle through which the photons are scattered (θ) is small, it can be assumed that the two scattered photons, and also the transmitted photons effectively travel the same length and that all photons traverse the same tissue types, i.e. the values of μ are the same and l=t.

Thus for a photon scattered anywhere along the line AB

$$I_{diff} = \sigma_s I_o(E_{diff}) e^{-\mu(E_{diff})\times t}$$

Which can be rearranged to give

$$\sigma_s = \frac{I_{diff}(E_{diff})}{I_o(E_{diff})} e^{\mu(E_{diff})\times t} \tag{5.2}$$

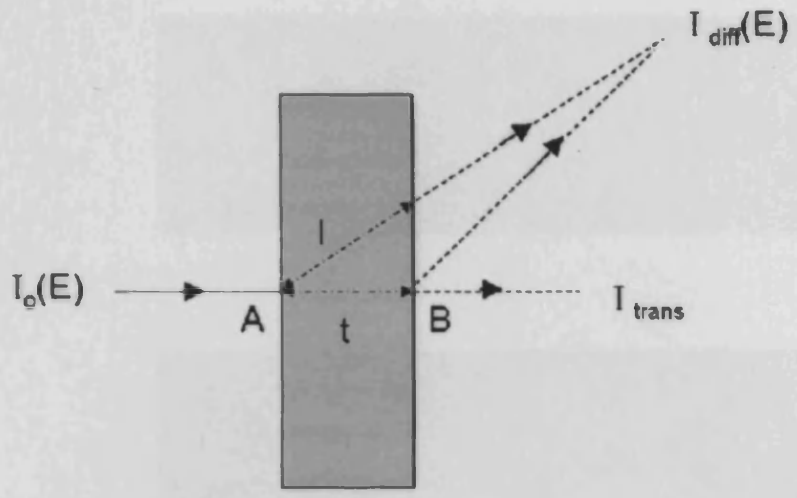


Figure 5.4: A schematic of the paths of scattered and transmitted photons through the tissue. The angle drawn here is exaggerated for clarity and was 6° for the measurements made in this chapter.

Substituting the value of t from the transmitted beam attenution

$$t = -\frac{1}{\mu(E)} ln \frac{I_{trans}(E_{trans})}{I_o(E_{trans})}$$

$$\sigma_s = \frac{I_{diff}(E_{diff})}{I_o(E_{diff})} \cdot \frac{I_o(E_{trans})}{I_{trans}(E_{trans})} e^{\frac{\mu(E_{diff})}{\mu(E_{trans})}}$$

Thus

$$\sigma_s \propto \frac{I_{diff}}{I_{trans}}$$

since all other factors are constant for each image. The images are therefore corrected for attenuation effects by dividing each pixel in the diffraction image by the corresponding value from the pixellated transmission image.

5.4.2 Corrected Large area images

The large area images, corrected by transmission data are shown in figures 5.6, 5.7, and 5.8. For comparison, photographs of the tissues are shown in section 6.5 with the tumourous regions indicated.

(a)

(b)

Figure 5.5: Tissue B a) Transmission image b) Pixellated transmission image created by averaging 10x10 pixel boxes of transmission image

The contrast achieved between regions in the diffraction images is generally superior to that in the transmission images (see table 5.1).

Tissue A contains an irregularly shaped ductal carcinoma with clear spiculations (needle like projections) as identified by a radiologist (Olivo (2008)). It also has small satellite masses, such as the one which can be seen clearly in the bottom left hand corner of the images in figure 5.6. The tumour does not fully penetrate the whole tissue thickness, except at its very centre, the remainder of the volume consisting of fatty tissue. For this reason the smaller masses and the edge of the tumour have less contrast with the fatty tissue than the tumour centre.

Tissue B contains a ductal carcinoma with well defined edges which can be seen in figure 5.7 at the bottom of the images towards the right. Above

(a) diffraction image (1.04-1.10nm^{-1})

(b) diffraction image (1.84-1.90nm^{-1})

(c) transmission image

(d) pixellated transmission image

(e) transmission corrected diffraction image (1.04-1.10nm^{-1})

(f) transmission corrected diffraction image (1.84-1.90nm^{-1})

Figure 5.6: Tissue A transmission and diffraction images. Diffraction data were collected for 5s per pixel.

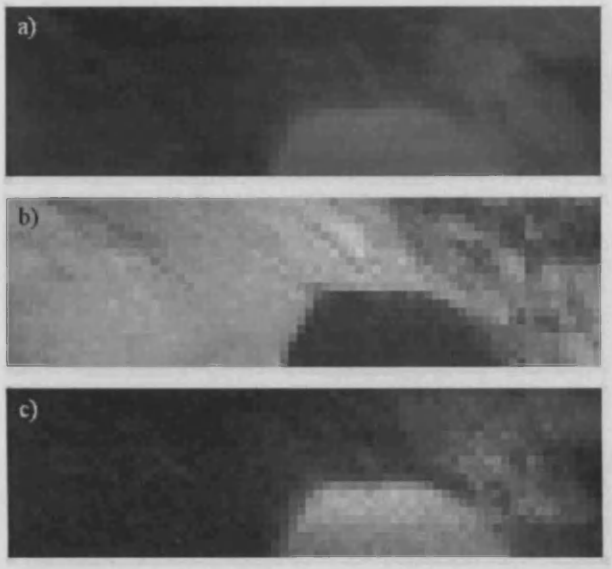


Figure 5.7: Tissue B a)Pixellated transmission image b)diffraction image from window 1.04-1.10nm^{-1} c) diffraction image from window 1.84-1.90nm^{-1}

the tumour on the far right of the image is a dense region of glandular tissue in which a smaller mass (tumour) is embedded.

The region imaged in tissue C is at the edge of the tissue and contains a small mass next to a vey thin tissue region. As with tissue A, the tumour does not occupy the whole tissue thickness but is embedded in fatty tissue. This makes the contrast in all images much reduced compared to those of the other tissues. The thin tissue next to it shows up as a dark region in all images as there is little attenuation and little scatter st any momentum trasnfer value.

The mean contrast between healthy and diseased regions is higher in the diffraction images, 24.6% and 38.9% for 1.04-1.10 nm^{-1} and 1.84-1.90 nm^{-1} respectively, than in transmission images (18.3%). These values indicate a 34% and 113% increase in contrast compared to the transmission images. Harris et al. (2003) found that measuring scattered intensity around 1.1 nm^{-1} gave a mean contrast between regions of 54% compared to transmission data

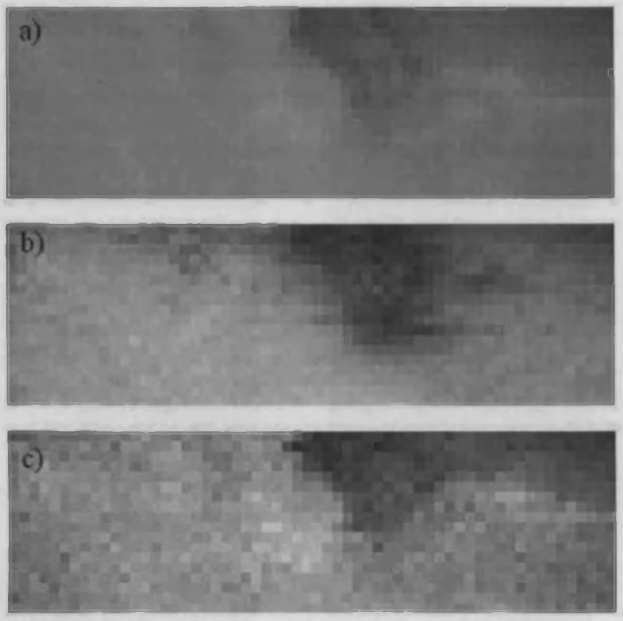


Figure 5.8: Tissue C a) Pixellated transmission image b) Diffraction image from window 1.04-1.10nm^{-1} c) Diffraction image from window 1.84-1.90nm^{-1}

	Contrast (%)		
Tissue	Transmission	Diffraction (1.04-1.10 nm^{-1})	Diffraction (1.84-1.90 nm^{-1})
A	38.2	41.6	42.1
B	22.1	38.4	55.6
C	2.9	2.1	8.3
region 1 (E)	*	*	*
region 2 (D)	24.9	20.8	51.8
region 3 (D)	19.5	26.1	47.0
region 4 (B)	12.3	49.4	43.5
region 5 (B)	8.6	23.2	24.5
region 6 (F)	22.8	14.0	34.8
region 7 (A)	13.3	35.7	42.7
mean	18.3	24.6	38.9

Table 5.1: Contrast between regions. * whole region fatty, therefore no contrasting regions.

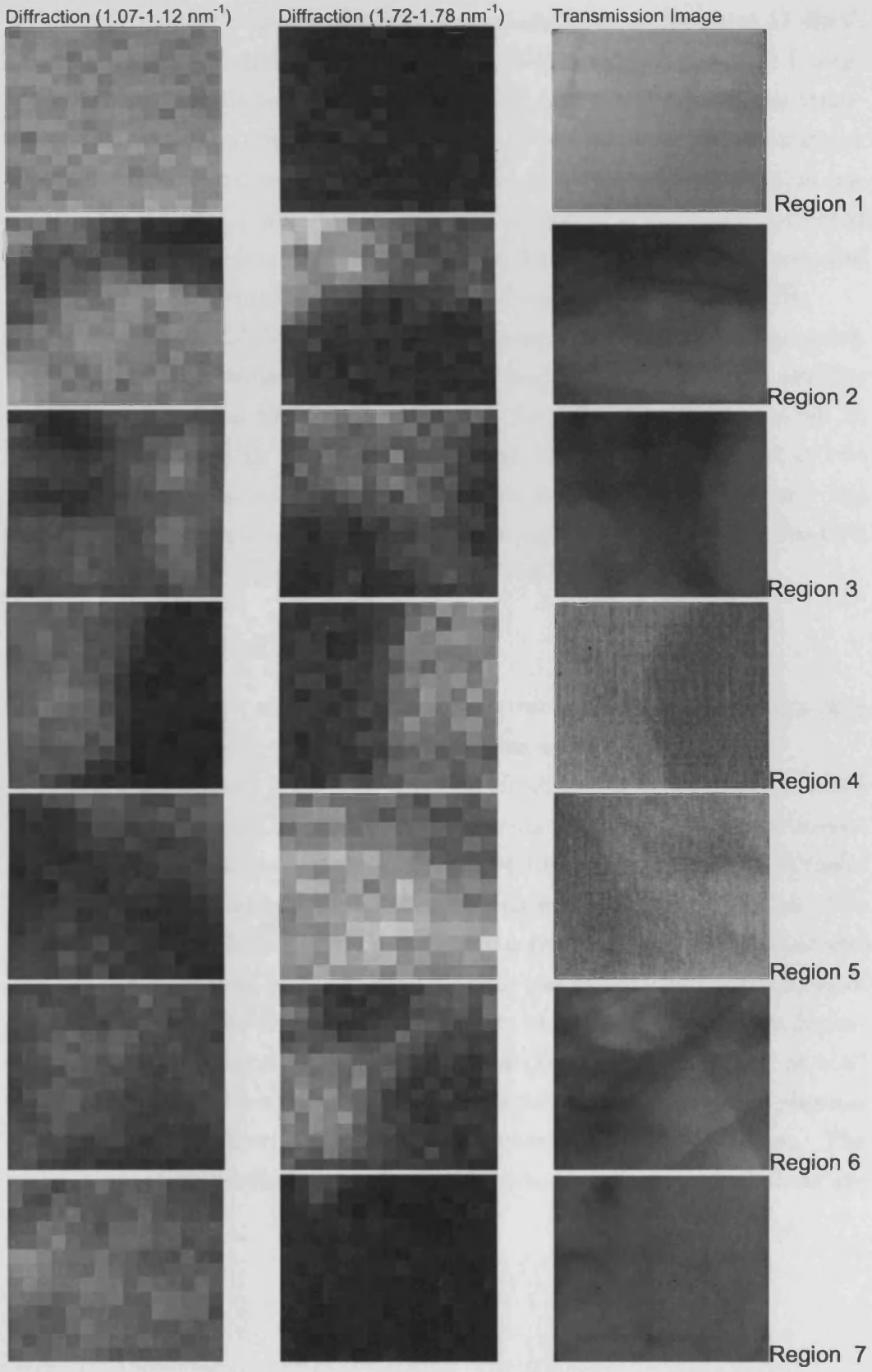


Figure 5.9: Images of all small tissue regions constructed using the counts in the momentum transfer windows 1.04-1.10 nm^{-1} and 1.84-1.90 nm^{-1}.

which gave a mean contrast of 31% using synchrotron radiation at 17.4keV. These values are higher than those found in this work for several resons. Monochromatic radiation was used producing higher contrast in the transmission images. The effective sample thickness was 33.6mm and attenuation effects were not corrected for, leading to a higher contrast in diffraction images measured at 1.1 nm^{-1} than would be found in attenuation corrected images. The momentum transfer resolution for the measurement was also better than those achieved in this work, 22% compared to 25 and 32%.

Griffiths et al. (2003b) used micro CT to examine breast tissue samples, using synchrotron radiation at 18keV. The mean contrast between healthy and diseased regions was found to be 38% for diffraction images taken at 1.1 nm^{-1} compared to 18% for transmission images. The contrast in the transmission images measured here matches well with this data, and the discrepancy between the diffraction data can again be attributed to the lack of attenuation correction.

5.4.3 Corrected ratio images

The ratio images are created by taking the ratio between two regions of a energy dispersive x-ray diffraction profile from each tissue voxel.

To correct the ratio images for attenuation effects each image was divided into regions containing adipose and tumourous tissue by applying a threshold based on the number of counts in the 1.84-1.90 nm^{-1} momentum transfer window. The adipose and tumourous regions in the ratio image were then multiplied by the relevant correction factor to give the corrected ratio image. The correction factors were calculated using the attenuation coefficients of each tissue type. The attenuation coefficients of materials are energy dependent. Since the window around the adipose peak occurs at 26keV at a 6° scatter angle, and the window above the tumour peak at 45keV, the photons contributing to each window will be attenuated to different degrees. The mass attenuation coefficients used for the correction factor calculations are

shown in table 5.2 and are taken from Nowotny (1998), with water used as a substitute for glandular tissue and tumour (Pani et al. (2006)). Data for whole breast tissue is included for comparison.

Tissue	μ/ρ at 26 keV (cm^2/g)	μ/ρ at 45keV (cm^2/g)	Density (g/cm^3)
Adipose	0.37	0.23	0.95
Water	0.47	0.24	1.00
Breast	0.45	0.23	1.02

Table 5.2: Mass attenuation coefficients for breast tissue and water (Nowotny (1998))

Using equation 5.2 for each energy

$$\sigma_s(1.04 - 1.10\text{nm}^{-1}) = \frac{I_{diff}(26keV)}{I_o(26keV)} e^{\mu(26keV)\times t}$$

and

$$\sigma_s(1.84 - 1.90\text{nm}^{-1}) = \frac{I_{diff}(45keV)}{I_o(45keV)} e^{\mu(45keV)\times t}$$

The ratio image is built from the ratio of these scatter probabilities

$$I_{rat} = \frac{\sigma_s(1.04 - 1.10\text{nm}^{-1})}{\sigma_s(1.84 - 1.90\text{nm}^{-1})}$$

Therefore

$$I_{rat} = \frac{\frac{I_{diff}(26keV)}{I_o(26keV)} e^{\mu(26keV)\times t}}{\frac{I_{diff}(45keV)}{I_o(45keV)} e^{\mu(45keV)\times t}}$$

Since $I_{diff}(26keV)$ and $I_{diff}(45keV)$ are the measured intensities and $I_o(26keV)$ and $I_o(45keV)$ are constant within each image, the attenuation

correction factor is given by

$$Correction factor = \frac{e^{\mu(26keV)}}{e^{\mu(45keV)}}$$

Thus substituting the values of mass attenuation coefficients and densities shown in table 5.2 in adipose tissue

$$\begin{aligned} Correction factor &= \frac{e^{\mu(26keV)_{adipose}}}{e^{\mu(45keV)_{adipose}}} \\ &= 1.14 \end{aligned}$$

and in glandular or tumourous tissue

$$\begin{aligned} Correction factor &= \frac{e^{\mu(26keV)_{water}}}{e^{\mu(45keV)_{water}}} \\ &= 1.26 \end{aligned}$$

This means that the ratio image needs to be corrected by a factor of 1.14 in regions of adipose tissue and 1.26 in regions of tumourous tissue.

The images of the large regions are shown in figure 5.10 and the small regions in figure 5.11.

The mean ratio values in regions of interest in all samples was measured and the results shown in table 5.3.

Tissue C shows a ratio more typical of a healthy region, due to partial volume effects as the tumour does not occupy the whole tissue thickness, only a region on the X-ray beam entrance surface. This also explains the large standard deviation in pixel values in the region. The contrast (13.8%), though lower than the other tissues, is still an improvement on that obtained by tranmission imaging (2.9%) or the diffraction regions individually (2.1%

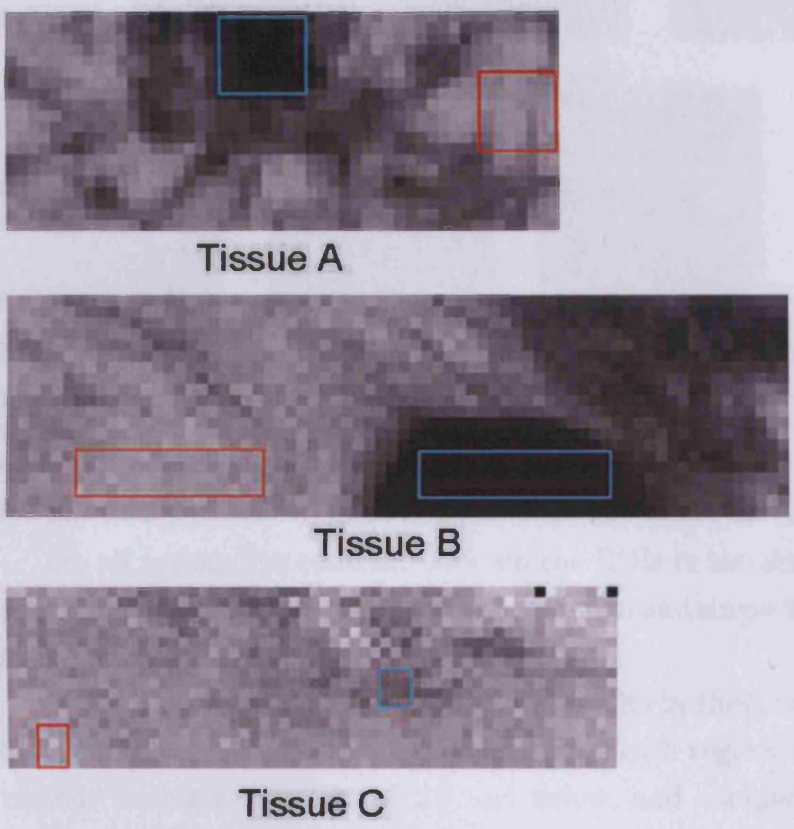


Figure 5.10: Images of all large tissue regions constructed using the ratio of counts in the momentum transfer windows 1.04-1.10 nm^{-1} and 1.84-1.90 nm^{-1}. The regions of interest used to measure contrast are also indicated (red box for healthy tissue, turquoise for tumour).

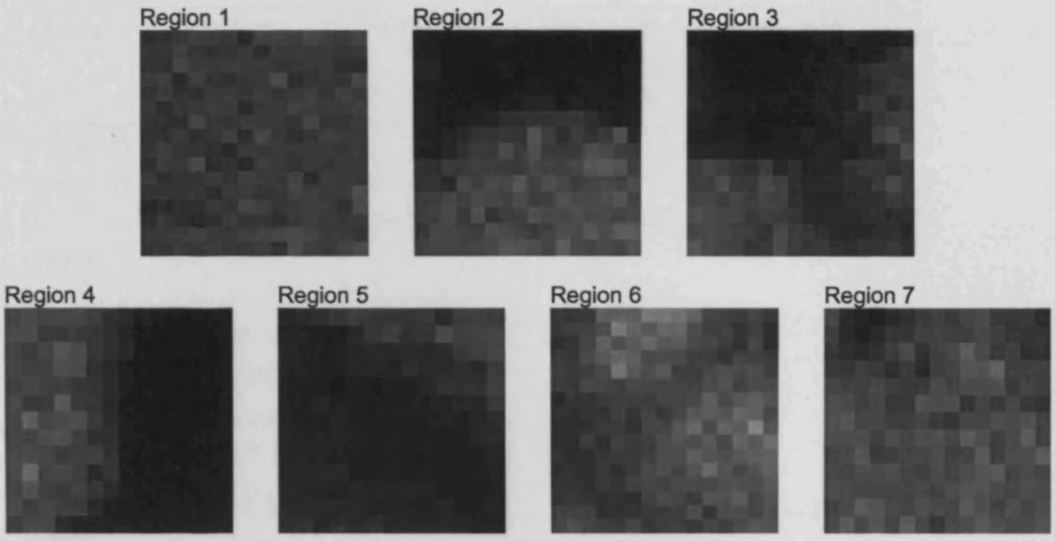


Figure 5.11: Images of all small tissue regions constructed using the ratio of counts in the momentum transfer windows 1.04-1.10 nm^{-1} and 1.84-1.90 nm^{-1}.

for the 1.04-1.10 nm^{-1} image and 8.3% for the 1.84-1.90 nm^{-1} image).

For all regions the contrast between the ROIs in the diffraction ratio images is increased compared to both transmission and single window diffraction images.

These data show that the ratio of the counts in these two regions gives a clear quantitative measure of the tissue type, with regions containing purely tumour having coefficients of 2.5 and below, and completely fatty regions having values of 4.8 and above. A measure such as this could be used to quantitatively identify the content of breast tissue.

An additional region in the middle of these two could also be used if further tissues required identification, e.g. fibrocystic tissue, which had a peak in its diffraction profile in between adipose and tumerous tissue (Kidane (2001), Changizi et al. (2005a)). This would be useful in differentiating between malignant and benign lesions, increasing specificity and reducing the number of biopsies performed on women.

Tissue	Healthy	Tumour	Contrast between adipose tissue and tumour (%)
tissue A	8.4 ±1.0	2.1 ±0.8	60.0
tissue B	9.0 ±0.6	1.3 ±0.1	74.8
tissue C	9.9 ±0.7	7.5 ±1.0	13.8
region 1	6.0 ±0.8		
region 2	7.6 ±0.9	1.8 ±0.3	61.7
region 3	6.8 ±0.6	1.8 ±0.4	58.1
region 4	7.0 ±0.9	1.3 ±0.5	68.7
region 5	5.6 ±1.3	1.7 ±0.2	53.4
region 6	4.8 ±0.7	2.7 ±0.5	
region 7	8.0 ±0.6	3.2 ±0.7	42.9
mean	6.8	2.4	53.0

Table 5.3: Mean values of ratios. The errors stated are the standard deviations of pixel values over the ROI.

5.5 Combining transmission and diffraction information

The information from diffraction data has been shown to produce images with both a high contrast between healthy and diseased regions and with tissue specificity. However the need for collimation of the scattered photons lowers the achievable spatial resolution, especially as a wide angular acceptance is needed to give adequate counting statistics in a short acquisition time.

It would be useful to combine both the transmission and diffraction data to utilise the excellent spatial resolution of the former with the increased contrast and tissue specificity of the latter. This can be done in a range of ways.

One possibility is to use the diffraction information to give markers or contours on the transmission image which could be used to alert Radiologists to regions that may be suspicious, as is done by CAD programs.

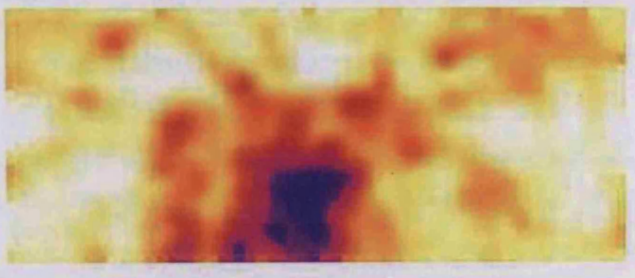
Alternatively the two images could be combined as is done with PET / CT

images, where the excellent spatial resolution and anatomical information of the CT image is combined with the lower resolution functional information from the PET scan. In this the PET scan is assigned a colour scale and superimposed on the CT image. An example of this with the diffraction and transmission images of tissue A is shown in figure 5.12. The diffraction images are 80% transparent to allow the detail from the transmission image to be seen clearly and are smoothed before combination so the edges of the larger pixels do not detract from the transmission information.

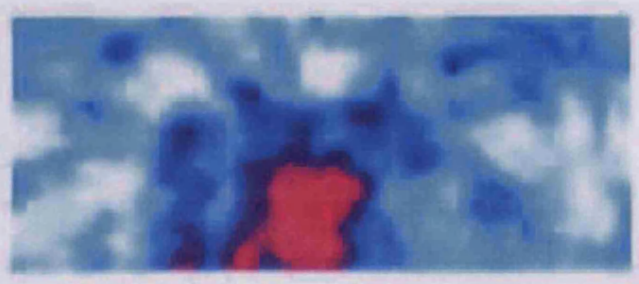
It can be seen here that there are some slight image registration problems, as the two images do not appear to coincide exactly, having some slight differences in the top left hand corner of the images. This is not surprising as the diffraction and transmission images were taken separately and the tissue may have been twisted slightly in between. This could either be compensated for using computer based image registration techniques, or eliminated by acquiring both images simultaneously.

5.6 Conclusions

Six breast tissues samples were imaged by scanning them though a diffraction system that collected coherent scatter at an angle of 6°. Adipose tissue produced a diffraction peak at 1.1 nm^{-1} and tumourous tissue at 1.5 nm^{-1}. Images produced using windows of the diffraction profile were found to give maximum contrast between healthy and tumourous tissue when using momentum transfer regions around of 1.04-1.10 nm^{-1} and 1.84-1.90 nm^{-1}. The images were corrected for attenuation effects using transmission images. The corrected images had mean contrasts between healthy and tumourous tissue of 18.3% in transmission images, compared to 24.6% and 38.9% in the diffraction images. This shows that diffraction images have an increased contrast compared to conventional imaging methods. A further increase in



(a) False colour fat diffraction image



(b) False colour tumour diffraction image

(c) Combined transmission and colour fat diffraction image

(d) Combined transmission and colour tumour diffraction image

Figure 5.12: Smoothed, false colour diffraction images and the results of their combination with a transmission image.

contrast can be obtained by taking the ratio between these two regions (mean contrast 53.0%). This ratio can also can be used to provide a quantitative measure of tissue pathology. Diffraction information could be combined with transmission images to utilise the excellent spatial resolution of conventional techniques with the improved contrast and specificity of diffraction imaging.

Chapter 6

Intelligent Imaging

6.1 Introduction

The advent of digital imaging devices in diagnostic radiology presents an opportunity for optimising the image acquisition process with myriad novel methods.

This chapter considers the potential of using tissue diffraction information as a feedback parameter in an adaptive digital mammographic imaging system.

6.2 The I-Imas System

I-ImaS is a European project aiming to develop intelligent imaging systems for medical applications using MAPS (monolithic active pixel sensor) detectors. The on-chip processing capabilities of MAPS technology combined with an FPGA (field programmable gate array) allows data analysis in real time, thus during the acquisition of an image (Longo et al. (2007)). The I-ImaS system consists of two linear array detectors, the 'scout' sensor and the 'I-ImaS' sensor. Information from the scout scan is used to manipulate the beam intensity for the second scan, thus providing an adaptive imaging system with the exposure parameters depending on features in the tissue. Preliminary results suggest that by redistributing dose to tissue regions indicated by the scout scan, dose savings of 30-70% can be made without loss of image quality compared to conventional mammographic systems (Schulerud et al. (2007)). The system is currently being designed for use in mammography and dental encephalography, though it has many other applications in both diagnostic radiology, security and industry.

Currently Automatic Exposure Control (AEC) is used to set system parameters for optimal images. In AEC the image parameters are set globally, which can lead to areas within the same image being underexposed and overexposed. AEC works in screen-film mammography by having a detector under the film cassette measuring the X-ray flux transmitted through the

breast, terminating the X-ray exposure when a predetermined threshold is reached. Since the detectors used in digital mammography absorb the X-ray beam, conventional AEC cannot be applied and methods such as a low-dose prescan are instead used to select the global exposure parameters for the main image (Elbakri et al. (2005)).

A line-scanning system such as the proposed I-Imas system, where each block of data is processed and the output used to provide information to optimise the second scan, has the potential to capture images with each region containing the maximum diagnostic information with the minimum dose.

The feedback used to provide the intelligence to drive the optimised scan is crucial to the operation of the system. The function or combination of functions used must behave in certain ways (I-Imas (2005)). Feedback algorithms must:

- Be suitable for texture analysis of mammographic and dental images (if using transmission image as feedback)
- Highlight areas of diagnostic interest
- Vary smoothly with exposure (mA and kV)
- Have a low complexity in terms of storage requirements and processing
- Exhibit stability with size over a range of sampling sizes

Any feedback parameter would be used to control one or more exposure parameters of the system (current, voltage or exposure time) thus the feedback parameter used must vary smoothly as a function of these.

The low computational complexity is necessary as the processing time is limited if the feedback is to occur in real time during image acquisition, especially if the optimisation process requires mechanical movement, and on chip processing power is limited. The time required to take a mammogram

is limited due to the large force required to compress the breast which causes extreme discomfort to the patient.

A range of feature functions were selected according to their suitability according to the above criteria.

6.3 Possible algorithms

For the I-Imas investigation six 1st order statistical functions were investigated; minimum, maximum, mean, standard deviation, skewness and kurtosis, with a further five features to estimate signal power, entropy and surface 'roughness'. A further nine synthetic features were created by normalising and combining the existing features to give a set of twenty features in total (I-Imas (2005)).

The performance of the textural feature functions was tested with regard to their capability to track image quality fluctuations when the exposure level was changed with respect to either kVp or mAs. The functions should be able to produce smooth and consistent transitions both in the detection of tissue structures over an image, and with the exposure level.

Nine of the possible feedback parameters considered were found to vary smoothly both with exposure and over an image, making them suitable properties to use in a feedback algorithm. These were:

minimum

$$I_{min} = min\{I(x,y)\} \tag{6.1}$$

maximum

$$I_{max} = max\{I(x,y)\} \tag{6.2}$$

mean

$$\mu = \frac{1}{XY}\sum_{i=1}^{X}\sum_{j=1}^{Y} I(x,y) \tag{6.3}$$

standard deviation

$$\sigma = \left(\frac{1}{XY}\sum_{i=1}^{X}\sum_{j=1}^{Y}(I(x,y)-\mu)\right)^{\frac{1}{2}} \tag{6.4}$$

standard deviation (normalised)

$$\sigma_n = \frac{\sigma}{\mu} \tag{6.5}$$

signal power

$$P_{XY} = \sum_{i=1}^{X}\sum_{j=1}^{Y}\|I(x,y)\|^2 \tag{6.6}$$

volume

$$V_{XY} = \sum_{i=1}^{X}\sum_{j=1}^{Y} I(x,y) \tag{6.7}$$

synthetic feature 19 (normalised power)

$$SF_{19} = \frac{P_{XY}}{XY} \tag{6.8}$$

synthetic feature 20 (normalised exposure)

$$SF_{20} = log\left(1 - \frac{SF_{19}}{255^2}\right) \tag{6.9}$$

The standard deviation alone cannot be used as a feedback parameter without being normalised to the mean value of the region, since this would result in regions with no sample being falsely assigned high doses due to the quantum noise being high because of the large number of photons reaching the image receptor. In contrast the standard deviation normalised to the mean obeys Poisson statistics thus the noise is proportional to the inverse of the square root of the number of photons.

Analysis of a database of mammographic images indicated that the minimum, maximum and mean functions perform well at high exposure levels, but have little discriminatory power at low exposures. Power, volume, syn-

thetic function 19 and synthetic function 20 cope well at low exposures but are unable to perform well close to tissue boundaries. Standard deviation could recognise tissue features against a background but has no sensitivity to overall exposure.

Thus no single function is adequate to track both the information content of image texture and changes in the exposure level. Therefore, it is most probable that more than one of these candidate functions will be used in combination as the final feedback for the sensor intelligent control (I-Imas (2005)).

6.4 Diffraction information as intelligence

An alternative, or indeed additional, feedback parameter could be diffraction information. The current I-Imas sensor does not have energy resolution capabilities thus to use diffraction information an additional sensor array would be required at an angle to the incident beam to investigate the required momentum transfer region. If information from more than one momentum transfer region was required, more than one sensor could be used at slightly different angles providing a primitive ADXRD arrangement. An alternative would be to use instead a sensor with some energy resolving capabilities such as a CdZnTe detector (see section 7.2.3). Since scatter is always present this could be done with no additional dose to patient, though the number of photons counted would be low due to the reduced number of coherently scattered photons compared to transmitted photons.

The parameters that maximised the contrast in the tissue sample images (section 5.3) were also considered as feedback parameters. These were the diffraction count around adipose peak (1.04-1.10 nm^{-1}), diffraction count around tumour peak (1.84-1.90 nm^{-1}), the ratio between the counts in the two regions and the ratio normalised to the total number of counts.

diffraction (fat)

$$D_{fat} = \sum_{x=1.04}^{1.10} N(x) \tag{6.10}$$

diffraction (tumour)

$$D_{tumour} = \sum_{x=1.84}^{1.90} N(x) \tag{6.11}$$

Diffraction (ratio)

$$D_{ratio} = \frac{D_{fat}}{D_{tumour}} \tag{6.12}$$

Diffraction (normalised ratio)

$$D_{rn} = \frac{D_{fat}}{D_{tumour}} \times (D_{fat} + D_{tumour}) \tag{6.13}$$

These parameters fulfil the requirements stated in section 6.2 as they have been shown to be tissue specific (chapter 5), the number of counts detected varies smoothly with exposure and comparing the total number of counts detected in a region, and even the ratio of counts in two sensors, to a threshold value requires only simple computations.

6.5 Testing Feedback Parameters

A low dose reference transmission image (28kv 3mA) was used to calculate the image parameters for feedback using a box of 10x10 pixels, corresponding to a single pixel in the relevant diffraction images (1mm x 1mm). All the statistical functions were tested excepting synthetic function 19 which behaved in the same way as the power function.

For each parameter the maximum and minimum values of that parameter were found for the whole low dose image, and used to generate three equally spaced threshold values. These values described the boundaries of four levels that were used as feedback. The feedback maps for tissue A are shown in

figure 6.1, with the most suspicious regions shown in white and the least suspicious regions shown in black. With the levels selected to be equally spaced between the maximum and minimum values of each parameter for each tissue, the volume algorithm produced exactly the same results as the mean, thus volume is excluded from the following feedback maps, data and discussion.
Photographs of the tissues are included for comparison, with the tumourous regions indicated within the black lines. However, due to tissue movement and deformation between the measurements and the photographs being taken, these do not correspond exactly to the feedback images, though they can be used to give an indication of which regions are suspicious.

Tissue A contains an irregularly shaped tumour with some peripheral smaller lesions, which are indicated on the photograph in figure 6.1. Since the tissues are not uniform througout their depth, the photographs do not show all masses as some are embedded within the tissue.

Figure 6.1 shows that, for tissue A, the functions minimum, maximum, mean, volume, power and synthetic variable 20 all provide the same feedback information. The main body of the tumour, the spicules and the smaller satellite tumours are all flagged as suspicious, though much of the areas in between are assigned only one level lower, leading to a large overall relative dose. The diffraction ratio feedback map is similar to the aforementioned statistical parameters, but indicates more of the regions between lesions as healthy tissue, as does the normalised diffraction ratio image. The two single window diffraction images both highlight the main body of the tumour as suspicious, but are not so sensitive to the smaller features. In contrast, the standard deviation feedback map most strongly indicates these small regions as they are areas in the image with a large variation in pixel values. Looking at the standard deviation feedback maps however it is not possible to clearly distinguish the main tumour, though when comparing the map

(a) Photograph

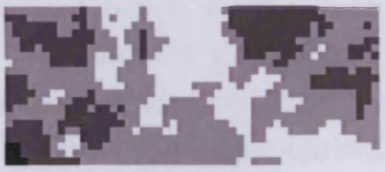
(b) Minimum

(c) Maximum

(d) Mean

(e) Standard Deviation

(f) Normalised Standard Deviation



(g) Power

(h) Synthetic Variable 20

(i) Diffraction (fat)

(j) Diffraction (tumour)

(k) Diffraction (ratio)

(l) Diffraction (normalised ratio)

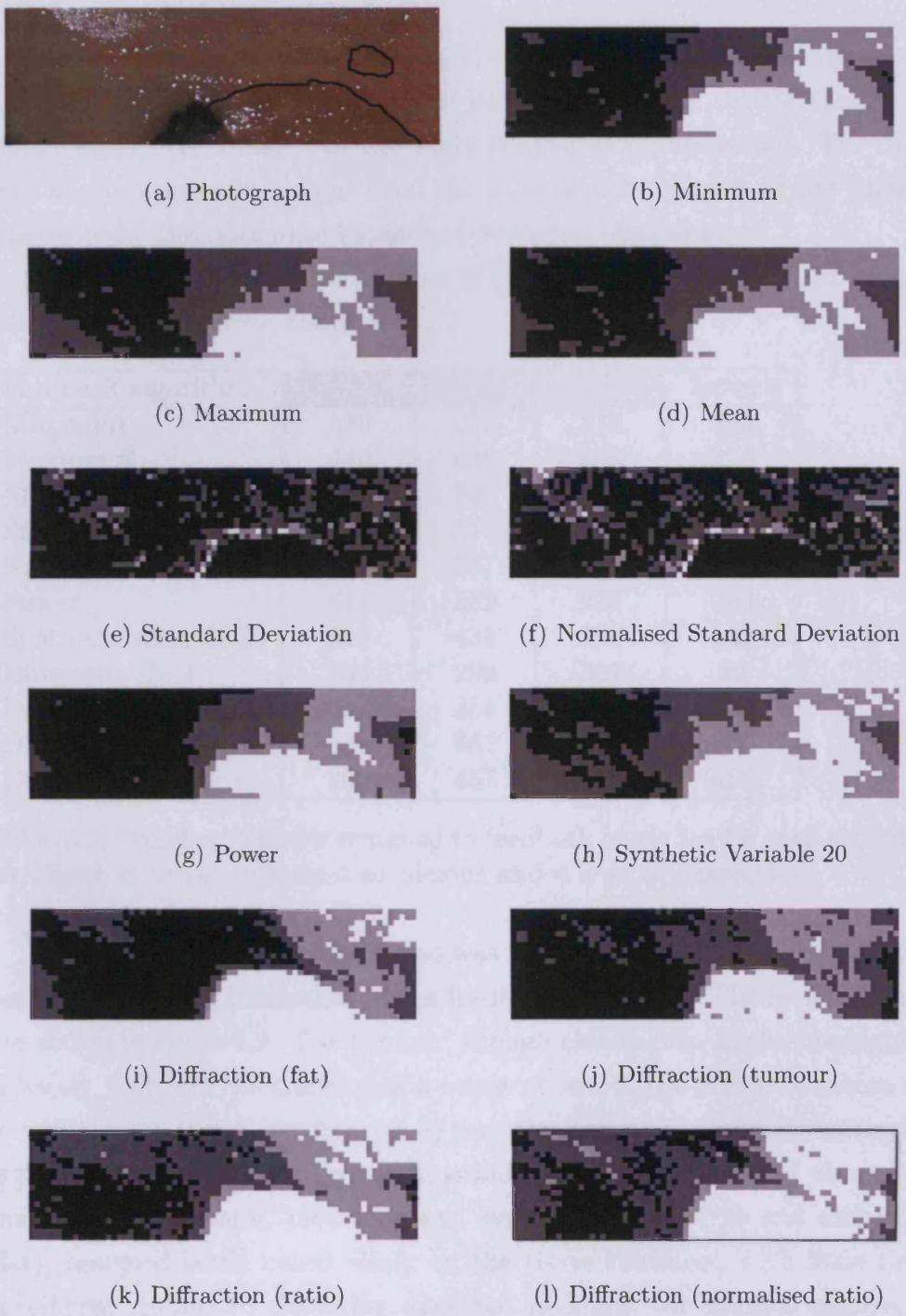
Figure 6.1: Feedback maps for tissue A using statistical functions and diffraction information

to the other maps it is evident that the edges are suggested by some bright pixels corresponding to high levels. Table 6.1 summarises how many pixels of each level make up each feedback map, with level 1 being the least suspicious black areas and level 4 being the very suspicious, white areas.

Feedback algorithm	Level 1	Level 2	Level 3	Level 4
Minimum	23	370	457	200
Maximum	11	418	442	179
Mean	15	368	446	221
Standard Deviation	678	314	53	5
St. Dev. normalised	639	341	61	9
Power	7	243	507	293
Synthetic variable 20	37	474	388	151
Diffraction(fat)	517	333	119	81
Diffraction(tumour)	471	356	142	81
Diffraction(ratio)	144	443	288	175
Diffraction(norm. r)	246	359	265	180

Table 6.1: Number of pixels assigned to feedback levels 1-4 for each algorithm on Tissue A, where 1 is least suspicious and 4 is most suspicious

Tissue B contains a well defined mass with distinct edges with a dense glandular region above and to the right with a smaller mass embedded in it. The feedback maps shown in figure 6.2 again show that the functions minimum, maximum, mean, volume, power and synthetic variable 20 all provide the same information, though for this tissue the diffraction (fat) and diffraction (tumour) images provide extremely similar information as well. All the aforementioned functions identify the main tumour as highly suspicious as well as the density above it, with the rest of the dense region being highlighted as one level down from the most suspicious. The large fatty region shown to the left of the tumour is marked in all cases as uninteresting. The diffraction ratio feedback map is also similar, though with fewer regions marked as not at all suspicious. The normalised diffraction ratio feedback map shares this feature, but also marks a larger dense region next to the



(a) Photograph

(b) Minimum

(c) Maximum

(d) Mean

(e) Standard Deviation

(f) Normalised Standard Deviation

(g) Power

(h) Synthetic Variable 20

(i) Diffraction (fat)

(j) Diffraction (tumour)

(k) Diffraction (ratio)

(l) Diffraction (normalised ratio)

Figure 6.2: Feedback maps for tissue B using statistical functions and diffraction information

main mass as highly suspicious.

Again the standard deviation map is completely different to all the others, with the distinct edges of the tumour being marked very clearly, though the region within the tumour is uniformly flagged as uninteresting. The dense area is given no more weight than the fatty area to the left of the tumour, especially by the normalised standard deviation algorithm.

Table 6.2 summarises the number of pixels assigned to each level for each feedback parameter for tissue B.

Feedback algorithm	Level 1	Level 2	Level 3	Level 4
Minimum	549	275	273	303
Maximum	440	426	306	228
Mean	505	340	287	268
Standard Deviation	1032	311	48	9
St. Dev. normalised	925	387	76	12
Power	614	282	295	209
Synthetic variable 20	329	433	296	342
Diffraction(fat)	765	299	307	29
Diffraction(tumour)	618	384	231	167
Diffraction(ratio)	345	462	288	305
Diffraction(norm. r)	306	457	155	482

Table 6.2: Number of pixels assigned to feedback levels 1-4 for each algorithm on Tissue B, where 1 is least suspicious and 4 is most suspicious

The region of tissue C considered was at the edge of a piece of tissue, thus presented different challenges to the feedback algorithm. The feedback maps are shown in figure 6.3. The tumour, though clearly seen in the photograph, is barely visible in the transmission image shown in the previous section due to very poor contrast (see figure 5.8) thus the statistical parameters based on regions of such an image do not function well. The majority of parameters (minimum, maximum, mean, power, synthetic variable 20 and diffraction (fat)) assigned levels based solely on the tissue thickness, with little tissue specificity. Standard deviation does not highlight the gradual changes in

(a) Photograph

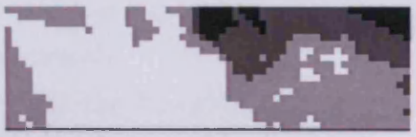
(b) Minimum

(c) Maximum

(d) Mean

(e) Standard Deviation

(f) Normalised Standard Deviation



(g) Power

(h) Synthetic Variable 20

(i) Diffraction (fat)

(j) Diffraction (tumour)

(k) Diffraction (ratio)

(l) Diffraction (normalised ratio)

Figure 6.3: Feedback maps for tissue C using statistical functions and diffraction information

tissue thickness as the other parameters do. Instead it highlights only the sharp tissue boundaries at the edges.

The diffraction feedback parameters functioned much better on this piece of tissue than the statistical functions.

Diffraction (around fat peak) did not work as well as in the previous tissues as it highlights areas with a low fat content as suspicious, thus assigns a greater dose in thinner tissue regions and outside the tissue if not combined with another function. However it did select the tumourous region which was completely missed in all of the previous algorithms. Diffraction (around tumour peak) functions well at tissue boundaries and highlights the relevant regions, though the tissue thickness also has some effect on the amount of scatter detected.

The diffraction ratio provided a more successful algorithm, highlighting the suspicious regions round the tissue edges rather than simply the thicker regions. This is ideal for a feedback algorithm as denser tissue needs a higher exposure for optimal image quality as well as the suspicious regions. The normalised diffraction ratio best indicated the suspicious areas and did not increase the exposure in thicker, healthy adipose regions, producing a lower dose overall.

Thus the function that worked best on this piece of tissue was the normalised diffraction ratio (equation 6.13). It best identified the tumour by the tissue boundary and did not provide feedback that was dominated by tissue thickness. Again, a summary of pixel levels is provided in table 6.3.

Overall, examining the feedback maps indicates that the statistical functions minimum, maximum, mean, power, volume and synthetic variable 20 all provided much the same information, highlighting the denser regions in the tissue.

Standard deviation (both normalised and unnormalised) highlighted the boundaries of the tumours well, as can best be seen in the feedback maps of tissue B (figure 6.2). However the feedback from the region within the tumour

Feedback algorithm	Level 1	Level 2	Level 3	Level 4
Minimum	23	111	275	671
Maximum	22	123	255	680
Mean	28	113	260	679
Standard Deviation	975	90	13	2
St. Dev. normalised	1001	62	15	2
Power	55	169	371	485
Synthetic variable 20	12	55	201	812
Diffraction(fat)	281	398	196	205
Diffraction(tumour)	49	180	636	215
Diffraction(ratio)	96	536	347	101
Diffraction(norm. r)	658	205	153	64

Table 6.3: Number of pixels assigned to feedback levels 1-4 for each algorithm on Tissue C, where 1 is least suspicious and 4 is most suspicious

margin is remarkably uniform and assigned the lowest suspicion feedback level. Thus, though standard deviation provides useful information, it would need to be combined with another function to identify both boundaries and large uniform regions of tumour.

The functions that best represented the tissue type at tissue edges were the diffraction functions based on the ratio of the counts in the two momentum transfer regions. These provided good specificity, even in very thin regions of tissue, as well as highlighting densities in the more central regions of tissue.

6.6 Optimised Image Generation

To simulate an image produced by an intelligent scanning system the feedback maps were used to generate 'optimised' images.

To make the changes more visible the pixels from the feedback map were binned into 2 x 2 boxes, corresponding to 2mm by 2mm tissue regions.

The feedback parameter normalised diffraction ratio was used to provide a feedback level for each 2mm squared region. The feedback map generated

is shown in figure 6.4.

Figure 6.4: Feedback map of tissue A generated using 2mm x 2mm pixels with the feedback parameter normalised diffraction ratio.

Since the levels indicate how interesting a region is, and how suspicious its contents are, an intelligent system would deposit a higher dose in the more suspicious regions to obtain the highest signal to noise ratio. Thus the four feedback levels need to to correspond to some exposure parameter of the system which could be altered to produce the optimised image.

The current I-Imas system has a series of aluminium and/or PMMA wedges that can be moved in and out of the beam to provide different beam filtration according to the information extracted from the scout scan (Longo et al. (2007)). To simulate the effects of altering the exposure a set of images of the tissues with different exposures were required.

Images of the tissue were taken using the DPix detector described in section 5.4.1 at a range of tube settings (kV rage 28-35kV, current range 8-125mAs) with a mammographic X-ray source (Siemens Mammomat B). The entrance dose was measured for each kV/ mA combination when taking each image using an ionisation chamber and electrometer (Keithley, 35050A), with the ionisation chamber on the surface of the tissue to include backscatter as described in section 5.1.2.

The feedback map was used to set the mAs for the optimised image, to simulate the effects of altering the X-ray tube current during image acquisition, with level 1 being assigned the lowest current and level 4 the highest.

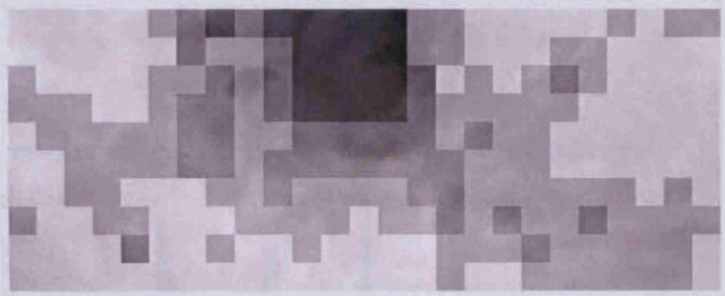
The setting used for the sample images displayed in figure 6.5 were 30kV for all regions, with level 1 regions being imaged with 8mAs, level 2 regions with 20mAs, level 3 with 32mAs and level 4 with 50mAs. These mA values were chosen as they were the most equally spaced of the settings available. Each region of the optimised image was then built from the corresponding region of the image with the required kV / mAs combination as determined by the feedback map.

The optimised image thus contained regions with different exposure parameters and required exposure equalisation for optimal viewing. This was achieved by smoothing each 20x20 box of the optimised image doing the same to the whole of a mid dose reference image (30kV 20mA), finding the ratio between each region in each of the smoothed images, and using these ratios to correct the optimised image. The smoothing procedure suppresses local variations in the image due to tissue composition but preserves the signal variations that reflect the change in exposure (Byng et al. (1997)). The smoothing function applied in IDL used a moving average filter of kernel width 3.

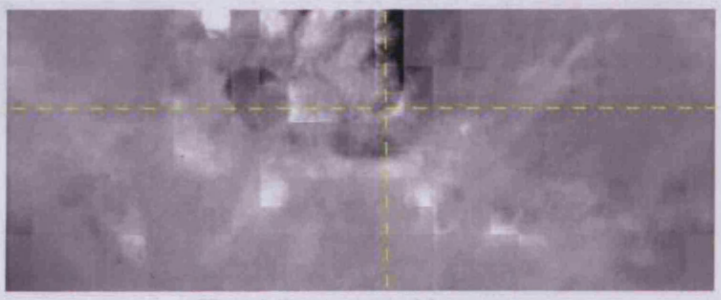
Examples of the correction procedure are shown for tissue A in figure 6.5.

The optimised images have a higher exposure in the dense regions around the tumour, which after normalisation using the reference image lead to a greater contrast in these areas compared to the edges of the image which appear less dense and are assigned a lower exposure. This can be seen in figure 6.6, where the profiles through the reference and optimised image are the same in the regions considered not suspicious (pixels 0-30 in a))or have an even lower intensity (pixels 360-420 in b)), but in the regions marked as suspicious the optimised image has sharper peaks (e.g. around pixel 235 in b)), greater contrast between tissues, which for the peak around 235 in b) is

Reference Image

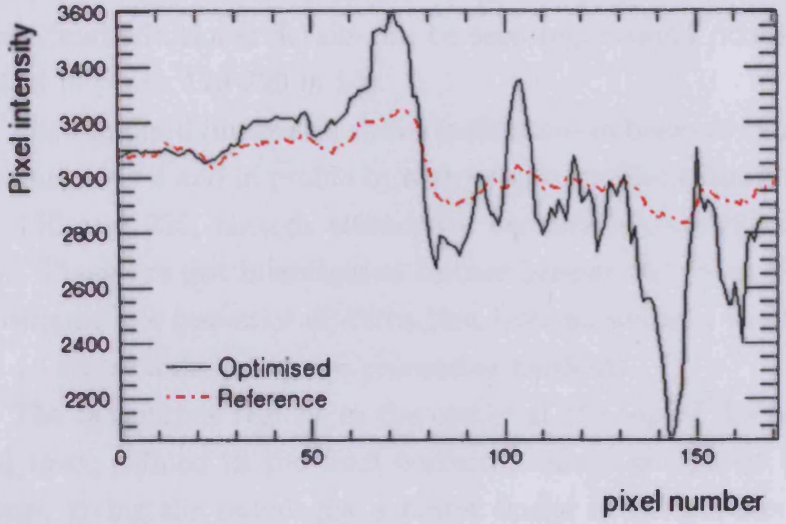


Optimised Image

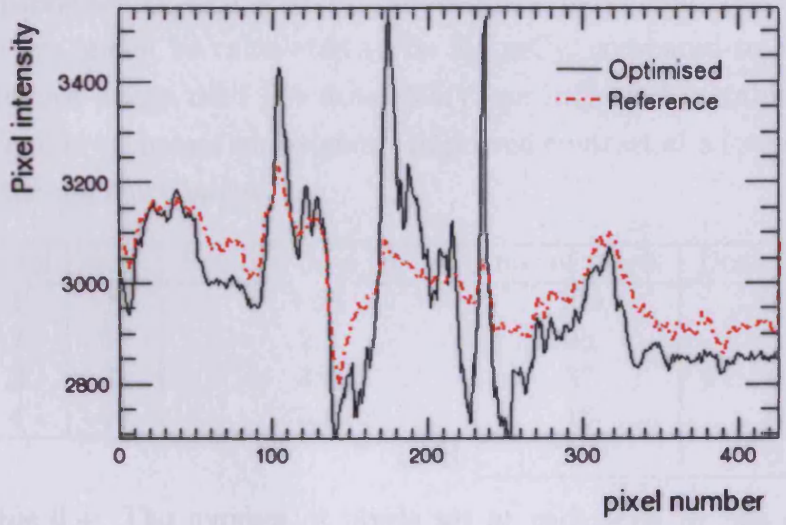


Final Corrected Image

Figure 6.5: Normalising the optimised image, with feedback based on the diffraction counts around the adipose peak. The dotted lines indicate from where the plots shown in figure 6.6 are taken.



(a)



(b)

Figure 6.6: Profiles through the reference and optimised images a) vertically (pixel zero at the bottom of the image) and b) horizontally (pixel zero at the left of the image). The positions in the image are indicated in figure 6.5

12.9% in the corrected optimised image compared to 2.5% in the reference image, and additional details can be seen (eg. around pixels 120 and 150 in a) and in pixels 170-220 in b)).

The equalised image still shows indications of boxes as can be seen in both the image itself and in profile b) where there are discontinuities around pixels 50, 150 and 255, though alternative equalisation algorithms could reduce this. These are not investigated further here as the focus of this work is to investigate the potential of diffraction information as a feedback parameter, not to assess different image processing methods.

The tumourous regions in the centre at the top of the image are clearer and more defined in the final corrected image compared to the reference image, giving the radiologist a better image of the borders of the tumour, which is an important factor in the diagnosis of lesions.

Using the doses measured for each mA/kV combination and the number of pixels set at each level by the feedback parameter, the relative dose for the image can be calculated to be 2.4 mGy, compared to 2.5 mGy for the reference image used (20 mAs, 30kV) as indicated in table 6.4. Thus the corrected optimised image shows improved contrast at a lower dose compared to the reference image.

Level	mAs	Relative dose (mGy)	no. of pixels	Dose fraction
1	8	1.25	100	0.50
2	20	2.5	95	1.0
3	32	3.6	37	0.5
4	50	5.4	18	0.4
		total	250	2.4

Table 6.4: The number of pixels set at each level by the diffraction ratio feedback algorithm and the associated relative dose for the image

6.7 Comparing parameters

The entrance dose was measured for each combination of source parameters as described in the previous section (see table 6.4). Using these values the relative dose for each optimised image could be calculated by multiplying the number of regions at each combination with the corresponding surface dose and dividing by the total number of pixels in the image. These values do not give a measure of the actual dose absorbed, but can be used to compare the different parameters. The numbers of pixels assigned to each level are shown in tables 6.1, 6.2 and 6.3, and a comparison of the relative dose of the 1mm x 1mm pixel size optimised images for all tissues is shown in figure 6.7.

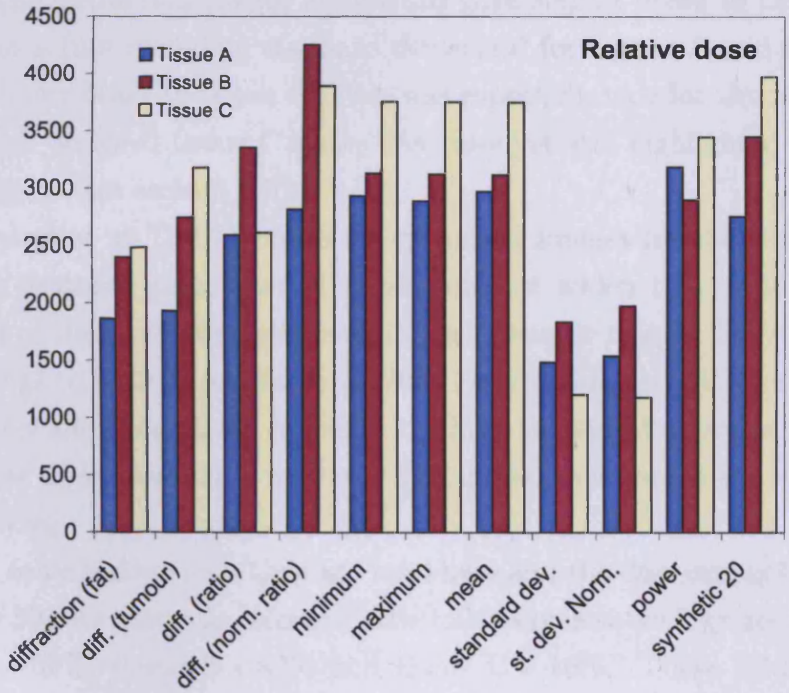


Figure 6.7: The calculated relative doses for each feedback function investigated for each of the tissues. Relative dose is the sum of the products of the alert level dose with the number of pixels with that level.

All three tissues have the same pattern of relative dose given by each parameter, with standard deviation (normalised and unnormalised) being the lowest, though as was stated in section 6.5, these parameters provided very different information from all the others.

The functions minimum, maximum and mean all provided both the same information about the tissue and identical doses for all tissues, with the dose being generally higher than that assigned by the other feedback parameters. Power and synthetic variable 20 were also very similar to the minimum, maximum and mean.

Diffraction (fat) gave the lowest dose for all functions excepting those based on standard deviation, with diffraction (tumour) giving only slightly higher. The diffraction (ratio) algorithms gave similar doses to the statistical functions (not including standard deviation) for tissues A and B, though assigned lower doses to tissue C. This was especially true for the normalised ratio, which assigned tissue C a very low dose yet still highlighted the interesting regions (see section 6.5).

Schulerud et al. (2007) produced optimised images for the same tissues using the feedback parameter of maximum, for which they took the 95th percentile of the maximum grey level in the reference image. They found the dose savings to be 33% for tissue A, 70% for tissue B and 64% for tissue C. This figures are significantly higher than those suggested above as they used the highest dose image as a reference (125mAs), therefore there was always a dose saving.

If the same is done with the data used here and the dose saving calculated using the 50mAs image as reference, the following dose savings are obtained: tissue A - 36%, tissue B - 47% and tissue C - 16%. These values are on average lower than those by Schulerud et al. (2007) as they looked at whole tissue which had large healthy regions whereas here only small regions around the main tumour are considered in this work as that was what was imaged in the diffraction study (chapter 5). This means the overall dose savings are

less as a lower proportion of low dose healthy regions are considered.

The mean dose savings, calculated using the 50mAs reference image, are shown for all parameters in table 6.5, averaged over the three tissues.

Algorithm	Mean Dose saving (%)
Minimum	32.6
Maximum	33.2
Mean	32.4
Standard Deviation	70.4
St. Dev. Normalised	69.5
Power	34.8
Synthetic variable 20	30.3
Diffraction(fat)	55.1
Diffraction(tumour)	47.8
Diffraction(ratio)	43.4
Diffraction(norm. r)	48.2

Table 6.5: Mean dose savings in the optimised images compared to the 50mAs reference image.

It can be seen that the diffraction parameters lead to greater dose savings than the statistical parameters such as maximum, excluding those based on standard deviation which have the highest dose savings.

6.8 Conclusion

The algorithm that best identified suspicious regions in all of the tissues investigated here was the normalised diffraction ratio. This used the ratio between the counts in the window around the fat peak (1.04-1.10 nm^{-1}) to those around the tumour peak (1.84-1.90 nm^{-1}) multiplied by the sum of counts in the two regions. It therefore has an inherent correction for tissue thickness allowing it to function well even at the thin edges of tissue. It can discriminate well between healthy tissue and cancerous tissue. The other algorithms considered (mean, maximum, minimum, power, volume and a synthetic variable) all provided very similar feedback information, with dense

tissue regions being highlighted well except at tissue edges. The standard deviation gave distinctly different feedback, as it identified sharp boundaries in the tissue and small masses. It would also function well at identifying microcalcifications which would not be detected by examining the diffraction windows used here. Thus using the ratio between the counts in the two momentum transfer regions corresponding to fat and tumour combined with the standard deviation of the pixels covering the same region would give the maximum information about the tissue or breast being examined. These parameters could be useful in driving the intelligence of an adaptive mammographic imaging system.

Chapter 7

Conclusions

7.1 Conclusions

The aim of this work was to demonstrate the use of EDXRD in three areas, material identification, imaging and system control. Each application required different approaches and different information to be extracted from the diffraction profiles, thus the method of analysis of the profile was driven by the application. The analysis techniques used and the results obtained are outlined below.

7.1.1 Material Identification

The problem of material identification was investigated using the application of detecting illicit drugs in packages. The diffraction profiles of seven narcotics commonly smuggled into the UK were measured, along with EDXRD diffraction data from five common cutting agents and from some packaging materials. The optimum scatter angle for the measurement of EDXRD drug data was found to be between 5.0° and 5.5° when using a tungsten anode X-ray source operated at 70kVp with 1mm additional aluminium filtration.

The diffraction profiles of the narcotics were distinct and characteristic and clearly different from those of the cutting agents and packaging materials.

In order to analyse the measured diffraction profiles as required in an automated system, multivariate analysis software was used, as suggested for the similar problem of explosives detection (Luggar et al. (1998)). Diffraction profiles of all the drugs mixed in concentrations of 20-100% with the cutting agents were measured with a range of counting statistics.

Partial least squares regression was used to build a statistical model of the data that could then be used to predict the content of future samples from their measured diffraction profiles. Two models were built, one to indicate simply if the sample contained an illicit substance, and a second to identify each substance in the sample and its concentration.

Testing the first model on a second data set indicated that the presence a

drug could be identified down to the 20% level with a sensitivity of 96% and a specificity of 80%. The second model had a root mean squared error on prediction (RMSEP) of 13%, indicating that future predictions of substance content should be accurate to approximately ± 13%.

The addition of packaging in the scatter volume reduced the accuracy of the predictions, with polythene having the most noticeable effect with 1.6 mm of packaging around a sample of 5mm diameter reducing the predicted drug content by two thirds. This suggests that the system would be less accurate in identifying small quantities of narcotic hidden within large amounts of polythene wrapping, but this scenario is extreme and the system would easily identify larger quantities.

The EDXRD system and PLS models were tested on seven samples supplied by HOSDB whose content was not revealed at the time. All seven were correctly identified as having a drug or not, the constituents correctly identified, where known, and the percentage of identified components correct to within 15%, which agreed well with the predicited RMSEP.

The trade off between the statistical quality of the diffraction data and the momentum transfer resolution of the system was also investigated. A minimum momentum transfer resolution of 36%, corresponding to ± 1.0° at 5.5°, and a minimum number of 1,000 integrated counts are required to identify the drugs when the system is used in conjunction with multivariate analysis software. This combination of resolution and counts gave the shortest data collection time thus would maximise throughput in a scanning system.

EDXRD combined with multivariate analysis was therefore demonstrated to provide a suitable system for the detection of illicit drugs in packages.

7.1.2 Imaging

The imaging capabilities of EDXRD were investigated using human breast tissue, as traditional mammographic imaging techniques have an inherently

low contrast between tissue types due to only small difference in the attenuation coefficients, where as the diffraction profiles of healthy and tumourous tissue have been shown to be very different over the momentum transfer range 0.5-2.5 nm^{-1} (Kidane et al. (1999)).

Fixation of tissue in 10% neutral buffered formalin was shown not to affect the measured diffraction profiles of tissue over this momentum transfer range, thus fixed tissues can be used in imaging investigations, eliminating the need for freezing tissues for storage and keeping them cool during the data acquisition. The equivalence of porcine tissues to human breast tissues was also demonstrated, with pork muscle producing an identical diffraction profile to a tumour and the profile of pork fat being very similar to that of human breast adipose tissue.

Three large regions and seven small regions from six tissue samples were studied. The regions were scanned in 1mm intervals and diffraction data collected at each point. Adipose tissue produced a profile with a sharp peak at 1.1 nm^{-1}, and tumour a broader and less intense peak around 1.5 nm^{-1}.

Windows from the measured profiles were selected and used to generate diffraction images of the samples. The optimal windows were selected based on the maximum contrast achieved in the resulting images between a healthy region and a tumourous region. The maximum contrast was found using a window 10 MCA channels wide centred on 1.07 nm^{-1}, corresponding to 1.04-1.10 nm^{-1}. A second region was also selected from 1.84-1.90 nm^{-1}, the region above the tumour peak, which also provided a high contrast. These two regions were then used to generate tissue images, which were corrected for attenuation affects using transmission images. The resulting diffraction images were analysed by measuring the contrast between regions of interest, and the contrast was found to have an mean value of 24.6% for the 1.04-1.10 nm^{-1} window image and 38.9% for the 1.84-1.90 nm^{-1} window image, compared to 18.3% for the transmission images. Images produced using the ratio of counts in these two windows produced even high contrast images,

with an average value over the tissue regions examined here of 53.0%. This quantity also provides a quantitative measure of tissue pathology.

Thus the improved image contrast possible using diffraction techniques was demonstrated, and it was suggested that this could be combined with the high spatial resolution of transmission techniques to provided maximum diagnostic information.

7.1.3 System control

Diffraction information from breast tissue, optimsed as described above, was considered as a feedback parameter in a real time adaptive mammographic imaging system, and compared to a range of other statistical functions, measured from low dose transmission images.

Standard deviation provided excellent discrimination at the edges of features, but required normalization to prevent large doses being recommended at thin tissue edges. It also failed to highlight uniform regions within tumours. Minimum, maximum, mean, and the synthetic functions considered provided the same information as each other, and highlighted the most attenuating regions well whilst still allotting high exposures to large adipose tissue regions.

Diffraction (1.04-1.10 nm^{-1})indicated well the tumourous regions of tissue, but broke down at tissue edges where there was less scatter as it assigned these regions higher doses. Diffraction (1.84-1.90 nm^{-1})functioned very well in all tissues, though both single window diffraction images were not very sensitive to smaller tumours.

The ratio of counts in the two diffraction windows gave a good indication of tissue pathology, but was also less successful in thin tissue regions at the edges of tissues.

The diffraction ratio normalised to sum of counts in the windows was the algorithm that best identified suspicious regions in all of the tissues investigated here. The normalisation gave the algorithm an inherent correction for tissue thickness allowing it to function well even at the thin edges of tissue.

7.2 Future Work

There are many stages to be carried out in the investigation and design of both a mammographic imaging system that uses diffraction information to best effect and a detection system for illicit drugs. Some steps in carrying forward these ideas are outlined below.

7.2.1 Breast Tissue

Samples

The improved contrast achieved in breast imaging using diffraction data has been demonstrated, however using only a small number of samples. To use diffraction imaging quantitatively many more samples would need to be studied to have statistically reliable measures. More other pathologies such as benign conditions such as fibroadenoma and fibrocystic disease which were not investigated here also need to be studied to produce a system capable of recognising specific tissue types leading to a higher specificity than is available currently in mammography.

Fixation

It has been shown that 10% buffered formalin has no affect on the diffraction profiles on the animal and human tissues studied. However stronger concentrations and the lower pH of unbuffered formalin produced a significant shift in the diffraction peak of pork muscle. The effects of different formalin concentrations and pHs on tissue must be examined and understood and standard protocols for tissue fixation for EDXRD measurements established.

Fresh tissue and in vivo measurements

All work presented here has been conducted on fixed tissue, though this has been shown for fixation in 10% buffered formalin to be equivalent to using unfixed tissue. However excised tissue will not behave in the same was as breast tissue in vivo, thus this would be a step forward for the future. The correlation between the diffraction patterns observed on excised tissue and within the breast itself have yet to be demonstrated.

7.2.2 Illicit Drug Detection

Other materials

The materials in the parcel that are not illicit can have two effects on the measured diffraction profile and the subsequent content prediction. The materials will both attenuate the X-ray beam and produce diffraction profiles of their own. This can lead to the profiles of drugs becoming too attenuated to be recognised, to other material masking the drug signature where their diffraction patterns are superimposed or even to materials giving false positive results if their diffraction profiles are similar to those of the illicit drugs.

In order to reduce the likelihood of false positive results, a large number of confounding substances must be examined and built into the software model. Any common materials with diffraction profiles similar to those of the illicit drugs to be detected may lead to the re-evaluation of the limits on counting times and momentum transfer resolution. A database of materials that could be found in packages is required, though since this is an exhaustive list of materials only those commonly sent is practical, requiring a survey of package contents. Special attention should be given to those materials with similar chemical structures to the illicit drugs (e.g. other hydrochlorides) which may exhibit similar diffraction peaks.

The attenuating effects of packaging materials and common contents also

need to be considered in order to select the photon energy range with which to work so that the diffraction profiles are not degraded to such an extent that the software can no longer identify them.

Analysis

Further optimisation of the Unscrambler model is required, investigating factors such as are models built using data of better statistical quality perform better on lower statistical quality profiles, and how smoothing and other pretreatment of data effects the accuracy of predictions. How well the model can distinguish between substances of similar chemical structure as described in the previous section will also need to be investigated.

7.2.3 System design

ADXRD/EDXRD/Hybrid

An investigation into whether EDXRD, ADXRD or a combination of the two techniques (hybrid XRD) would be most suitable for the detection of illicit materials in packages should be undertaken.

The energy dispersive techniques used here utilised the whole tungsten spectrum, however by using a combination of suitable filters, the tungsten spectrum could be modified to give a narrower band width, which could be used as part of an angular dispersive system. The optimum kVp of the source also needs investigating, and will effect the choice of scatter angle.

Dectectors

It may be possible to use a detector with a lower energy resolution than a Germanium detector, such as a Cadmium Zinc Telluride detector, which has been used for explosives detection in baggage (Malden and Speller (2000)). Cadmium Zinc Telluride (CZT) is a semi conductor material used for detectors. Its advantage over germanium is its large band gap (approximately 1.6

eV) which allows it to be used at room temperature, rather than requiring cooling to liquid nitrogen temperatures. It has a lower energy resolution than that achievable with state of the art LN_2 cooled germanium detectors, 5-6% at 60keV (Bertolucci et al. (1998), Takahashi and Watanabe (2001) and Limousin (2003)) compared to $\leq 1\%$ for germanium , though pixellated detectors are available which offer spatial resolution in addition to energy resolution. Its use in place of germanium would be suitable as it has been shown in this work that a momentum transfer resolution of 36% is sufficient for the purpose of drugs identification, which is easily achievable using a CZT detector with appropriate collimation. The additional spactial resolution could provide a hybrid EDXRD/ADXRD system with further improved resolution and faster acquisition times.

With the introduction of filtration narrowing the range of X-ray energies, simple CMOS sensors or CCDs could be used in a purely angular dispersive system. The system would need to be modeled and factors such as the DQE of the system taken into account to investigate its potential for a fast, substance specific detection system. The advent of active pixel sensor technology (APS) also introduces the capabilities on on chip data processing and selective data readout. Combined with appropriate beam filtration and such a sensor could be used to build a detection system.

Practicalities

How the package would be moved through the system must be considered, as well as requirements such as parcel dimensions and weight, required spatial resolution and minimum quantities to be detected and whether depth information on the package is required.

For the system to function automatically, software must be developed that can automatically read in the spectra, give a prediction and alert staff to suspicious packages. It may also need to raise an alert if the X-rays are unable to penetrate the package.

A comprehensive survey of system users is also required to establish information such as acceptable scanning times, quantities and purities to be detected and acceptable false positive and false negative rates.

Bibliography

Avtandilov, G., Dembo, A., Komardin, O., Lazarev, P., Paukshto, M., Shkolnik, L., and Zayratiyants, O. (2000). Human tissue analysis by small-angle x-ray scattering. *Journal of Applied Crystallography*, 33(3 Part 1):511–514.

Baker, J. R. (1958). *Principles of Biological Microtechnique*, pages p37–40. Methuen & Co.

Barroso, R. C., Lopes, R. T., and de Jesus, E. F. O. (2001). X-ray diffraction microtomography using synchrotron radiation. *Nuclear Instruments and Methods in Physics Research A*, 471(1-2):75–79.

Batchelar, D. L., Chun, S. S., Wolin, T. A., Tan, J. K., Beiko, D. T., Cunningham, I. A., and Denstedt, J. D. (2002). Predicting urinary stone composition using x-ray coherent scatter: a novel technique with potential clinical application. *Journal of Urology*, 168:260–265.

Batchelar, D. L., Davidson, M. T. M., dabrowski, W., and Cunningham, I. A. (2006). Bone-composition imaging using coherent-scatter computed tomography: Assessing bone health beyond bone mineral density. *Medical Physics*, 33(4):904–915.

Berger, M. J. and Hubbell, J. H. (1987). XCOM: Photon cross sections on a personal computer. https://srmors.nist.gov/tables. National Institute of Standards and Technology, report NBSIR 87-3597.

Bertolucci, E., Chirco, P., Conti, M., Rossi, M., and Russo, P. (1998). Imaging performance of single-element cdznte detectors for digital radiography. *IEEE Transactions on nuclear science*, 45(3):406–412.

Blankespoor, S. C., Wu, X., Brown, J. K., Tang, H. R., Cann, C. E., and Hasegawa, B. H. (1996). Attenuation correction of spect using x-ray ct on an emission transmission ct system: Myocardial perfusion assessment. *IEEE Transactions on nuclear science*, 43(4):2263–2274.

Bohndiek, S. E., Cook, E. J., Avanitis, C. D., Olivo, A., Royle, G. J., Clark, A. T., Prydderch, M. L., Turchetta, R., and Speller, R. D. (2008). A cmos active pixel sensor system for laboratory- based x-ray diffraction studies of biological tissue. *Physics in Medicine and Biology*, 53:655–672.

Boone, J. M. and Cooper, V. N. (2000). Scatter/primary in mammography: Monte carlo validation. *Medical Physics*, 27(8):1818–1831.

Boskey, A. L., Cohen, M. L., and Bullough, P. G. (1982). Hard tissue biochemistry: A comparison of fresh-frozen and formalin-fixed tissue samples. *Calcified Tissue International*, 34(1):328–331.

Briki, F., Bussin, B., Sacicru, B., Esteve, F., and Doucet, J. (1999). Breast cancer diagnosis using hair. *Nature*, 400:226–227.

Burnett, M. G. (1982). The mechanism of the formaldehyde clock reaction: Methyleneglycol dehydration. *Journal of Chemical Education*, 59(2):160–163.

Byng, J. W., Critten, J. P., and Yaffe, M. J. (1997). Thickness-equalization processing for mammographic images. *Radiology*, 203(2):564–568.

Canfield, D. V., Barrick, J., and Giessen, B. C. (1979). Structure of diacetylmorphine (heroin). *Acta Crystallography B*, 35:2806–2809.

Castro, C. R. F., Barroso, R. C., Anjos, M. J., Lopesand, R. T., and Braz, D. (2004). Coherent scattering characteristics of normal and pathological human breast tissues. *Radiation Physics and Chemistry*, 71:649–651.

Castro, C. R. F., Barroso, R. C., de Oliviera, L. F., and Lopes, R. T. (2005a). Coherent scattering x-ray imaging at the brazilian national synchrotron laboratory:preliminary breast images. *Nuclear Instruments and Methods in Physics Research A*, 548:116–122.

Castro, C. R. F., Barroso, R. C., and Lopes, R. T. (2005b). Scattering signatures for some human tissues using synchrotron radiation. *X-ray spectrometry*, 34:477–480.

Changizi, V., Oghabian, M. A., Speller, R. D., Sarkar, S., and KheradmandA, A. A. (2005a). Application of small angle x-ray scattering (saxs) for differentiation between normal and cancerous breast tissue. *International Journal of Medical Sciences*, 2:118–121.

Changizi, V., Wilkinson, S., Hall, C. J., and Grossman, G. (2005b). A study of the effect of formalin preservation on normal and cancerous breast tissues using small angle x-ray scattering. *Radiation Physics and Chemistry*, 75:932–935.

Davidson, M. T. M., Batchelar, D. L., Velupillai, S., Denstedt, J. D., and Cunningham, I. A. (2005a). The analysis of urinary stone components by x-ray coherent scatter: Characterizing composition beyond x-ray diffractometry in the laboratory. *Physics in Medicine and Biology*, 50:3373–3386.

Davidson, M. T. M., Batchelar, D. L., Velupillai, S., Denstedt, J. D., and Cunningham, I. A. (2005b). Laboratory coherent-scatter analysis of intact urinary stones with crystalline composition:a tomographic approach. *Physics in Medicine and Biology*, 50:3907–3925.

Dawson, C., Horrocks, J. A., Kwong, R., and Speller, R. D. (1996). Low-angle x-ray scattering signatures of urinary calculi. *World Journal of Urology*, 14(1):S43–7.

Desouky, O. S., Elshemey, W. M., Selim, N., and Ashour, A. H. (2001). Analysis of low-angle x-ray scattering from lyophilized biological samples. *Physics in Medicine and Biology*, 46:2099–2106.

Elbakri, I. A., Lakshimarayanan, A. V., and tesic, M. M. (2005). Automatic exposure control for a slot scanning full field digital mammography system. *Medical Physics*, 32(9):2763–2770.

Elshemey, W. M., Desouky, O. S., Mohammed, M. S., Elsayed, A. A., and El-Houseini, M. (2003). Characteristics of cirrhosis and hepatocellular carcinoma using low-angle x-ray scattering signatures of serum. *Physics in Medicine and Biology*, 48:N239–N246.

Elshemey, W. M., Elsayed, A. A., and El-Lakkani, A. (1999). Characteristics of low-angle x-ray scattering from some biological samples. *Physics in Medicine and Biology*, 44:2907–2915.

Evans, S. H., Bradley, D. A., Dance, D. R., Bateman, J. E., and Jones, C. H. (2001). Measurement of small-angle scattering for some breast tissue and tissue substitute materials. *Physics in Medicine and Biology*, 36:7–18.

Farquharson, M. J., Luggar, R. D., and Speller, R. D. (1997). Multivariate calibration for quantitative analysis of edxrd spectra from a bone phantom. *Applied radiation and Isotopes*, 48(8):1075–1082.

Fernandez, M., Keyrilainen, J., Serimaa, R., Torkkeli, M., Karjalainen-Lindsberg, M. L., Leidenius, M., von Smitten, K., Tenhunen, M., Fiedler, S., Bravin, A., Weiss, T. M., and Suortii, P. (2005). Human breast cancer in vitro: matching histopathology with small angle x-ray scattering and

diffraction enhanced imaging. *Physics in Medicine and Biology*, 50:2991–3006.

Fernandez, M., Keyrilainen, J., Serimaa, R., Torkkeli, M., Karjalainen-Lindsberg, M. L., Tenhunen, M., Thomlinson, W., Urban, V., and Suortii, P. (2002). Small-angle x-ray scattering studies of human breast tissue samples. *Physics in Medicine and Biology*, 47:577–592.

Folen, V. A. (1975). X-ray powder diffraction data for some drugs, excipients, and adulterants in illicit samples. *Journal of Forensic Science*, 20(2):348–372.

Fong, R. (2002). *Characterisation of urinary stones using low angle x-ray scattering.* PhD thesis, University of London.

Fox, C. H., Johnson, F. B., Whiting, J., and Roller, P. P. (1985). Formaldehyde fixation. *Journal of Histochemistry and Cytochemistry*, 33(8):845–853.

Geraki, K., Farquharson, M. J., and Bradley, D. A. (2004). X-ray fluorescence and energy dispersive x-ray diffraction for the quantification of elemental concentrations in breast tissue. *Physics in Medicine and Biology*, 49(1):99–110.

Griffiths, J. A., Hanby, A. M., Horrocks, J. A., Bohndiek, S. E., and Speller, R. D. (2007). Correlation of energy disperrsive diffraction signatures and microct of small breast tissue samples with pathological analysis. *Physics in Medicine and Biology*, 52:6151–6164.

Griffiths, J. A., Metaxas, M. G., Royle, G. J., Venanzi, C., Esbrand, C., van der Stelt, P. F., Verheij, H., Li, G., Turchetta, R., Fant, A., Gasiorek, P., Theodoridis, S., Georgiou, H., Cavouras, D., Hall, G., Noy, M., Jones, J., Leaver, J., Machin, D., Greenwood, S., Khaleeq, M., Schulerud, H., Ostby,

J. M., Trantis, F., Asimidis, A., Bolanakis, D., Manthos, N., Bergamaschi, A., and Speller, R. D. (2006). A multi-element dectector system for intelligent imaging:i-imas. *IEEE Transactions on Nuclear Science.*

Griffiths, J. A., Royle, G. J., Speller, R. D., Horrocks, J. A., Olivo, A., Pani, S., Longo, R., Droessi, D., Spencer, S. H., Robbins, M. S., Clifford, D. P., and Hanby, A. M. (2003a). Diffraction enhanced breast imaging: assessment of realistic system requirements to improve the diagnostic capabilities of mammography. *IEEE Transactions on Nuclear Science*, 5:3230–3234.

Griffiths, J. A., Speller, R. D., Royle, G. J., Horrocks, J. A., Olivo, A., Pani, S., Longo, R., Droessi, D., Spencer, S. H., Robbins, M. S., Clifford, D. P., and Hanby, A. M. (2003b). X-ray diffraction ct of excised breast tissue sections: First results from elettra. *IEEE Transactions on Nuclear Science*, 4:2330–2334.

Harding, G. (2004). X-ray scatter tomography for explosives detection. *Radiation Physics and Chemistry*, 71:869–881.

Harding, G., Kozanetsky, K., and Neitzel, U. (1999). X-ray diffraction computed tomography. *Medical Physics*, 14:515–525.

Harding, G., Newton, M., and Kozanetsky, K. (1990). Energy-dispersive x-ray diffraction tomography. *Physics in Medicine and Biology*, 35:33–41.

Harding, G. and Schreiber, B. (1999). Coherent x-ray scatter imaging and its applications in biomedical science and industry. *Radiation Physics and Chemistry*, 56(12):229–245.

Harris, E. J., Royle, G. J., Speller, R. D., Griffiths, J. A., Kidane, G., and Handby, A. M. (2003). Evaluation of a novel (l3 vision) CCD technology for application to diffraction enhanced breast imaging. *Nuclear Instruments and Methods in Physics Research A*, 513(1):27–31.

Harris, K. M. and Vogel, V. G. (1997). Breast cancer screening. *Cancer and Metastasis Review*, 16(3-4):234–262.

Helander, K. G. (1994). Kinetic studies of formaldehyde binding in tissue. *Biotechnic and Histochemistry*, 69:177–179.

Helander, K. G. (1999). Kinetic studies of formaldehyde binding in tissue. *Journal of Histotechnology*, 22(4):317–318.

Hickey, D. S. and Hukins, D. W. (1979). Effect of methods of preservation on the arrangement of collagen fibrils in connective tissue matrices: an x-ray diffraction study of annulus fibrosus. *Connective Tissue Research*, 6(4):223–228.

Hopwood, D. (2002). *Theory and Practice of Histological Techniques*, chapter 5, pages 63–84. Churchill Livingston, fifth edition.

Howell, A., Grossman, J. G., Cheung, K. C., Kanbi, L., Gareth, D., Evans, R., and Hasnain, S. S. (2000). Can hair be used to screen for breast cancer? *Journal of Medical Genetics*, 37:297–298.

Hrynchuk, R. J., Barton, R. J., and Robertson, B. E. (1983). The crystal structure of free base cocaine, $C_{14}H_{21}NO_4$. *Canadian Journal of Chemistry*, 61(6):481–487.

Hubbell, J. M. and Overbo, I. (1979). *Journal of Physics and Chemical reference data*, 8:69–105.

Hubbell, J. M. and Seltzer, S. H. (1996). Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients. *www.nist.gov*.

I-Imas (2005). Translating information signatures to a sequence of well-defined processing functions. *Workpackage 3, Deliverable 8*.

James, V. (2003a). False-Positive Results in Studies of Changes in Fiber Diffraction of Hair From Patients With Breast Cancer May Not Be False. *J. Natl. Cancer Inst.*, 95(2):170–171.

James, V., Corino, G., Robertson, T., Dutton, N., Halas, D., Boyd, A., Bentel, J., and Papadimitriou, J. (2005). Early diagnosis of breast cancer by hair diffraction. *International Journal of Cancer*, 114(6):969–972.

James, V. J. (1999). Using hair to screeen for breast cancer. *Nature*, 398:33–34.

James, V. J. (2003b). Changes in the diffraction pattern of hair resulting from mechanical damage can occlude the changes that relate to breast cancer. *Physics in Medicine and Biology*, 48:L37–41.

James, V. J. (2003c). False-positive results in studies of changes in fiber diffraction of hair from patients with breast cancer may not be false. *Journal of the National Cancer Institute*, 95(2):170–171.

James, V. J. (2003d). Fibre diffraction from a single hair can provide an early non-invasive test for colon cancer. *Medical Science Monitor*, 9(8):MT79–84.

James, V. J. (2003e). The traps and pitfalls inherent in the correlation of changes in the fiber diffraction pattern of hair with breast cancer. *Physics in Medicine and Biology*, 48(2):L5–L9.

James, V. J. and Cookson, D. (2000). The possibilty of using x-ray diffraction with hair to screen for pathologic conditions such as breast cancer. *AIP Conference Proceedings, X-ray and Inner Shell Processes*, 506:590–594.

James, V. J., Yue, D. K., and McLennan, S. V. (1997). Changes in the molecular structure of hair in insulin-dependent diabetes. *Biochemical and Biophysical Research Communications*, 233:76–80.

Johns, P. C. and Wismayer, M. P. (2004). Measurement of coherent x-ray scatter form factors for amorphous materials using diffractometers. *Physics in Medicine and Biology*, 49:5233–5250.

Johns, P. C. and Yaffe, M. J. Coherent scatter in diagnostic radiology. *Medical Physics.*

Johns, P. C. and Yaffe, M. J. (1987). X-ray characterization of normal and neoplastic breast tissue. *Physics in Medicine and Biology*, 32:675–695.

Joint Committee on Powder Diffraction Studies, editor (1980). *Powder Diffraction File.* International Center for Diffraction Data, Pennsylvania.

Jupp, I. D., Durrant, P. T., Ramsden, D., Carter, T., Dermody, G., Pleasants, I. B., and Burrows, D. (2000). The non-invasive inspection of baggage using coherent x-ray scattering. *IEEE Transactions on Nuclear Science*, 47(6):1987–1994.

Kidane, G. (2001). *Breat Tissue Characterisation Using Low Angle X-ray Scattering.* PhD thesis, University of London.

Kidane, G., Royle, G. J., Hanby, A., and Speller, R. D. (1999). X-ray scatter signatures for normal and neoplastic breast tissues. *Physics in Medicine and Biology*, 44:1791–1802.

Kiernan, J. A. (1999). *Histological and Histochemical Methods:Theory and Practice.* Oxford: Butterworth-Heinemann.

Kinahan, P. E., Townsend, D. W., Beyer, T., and Sashin, D. (1998). Attenuation correction for a combined 3d pet/ct scanner. *Medical Physics*, 25(10):2046–2053.

Kleuker, U., Suortti, P., Weyrich, W., and Spanne, P. (1998). Feasibility study of x-ray diffraction computed tomography for medical imaging. *Physics in Medicine and Biology*, 43(10):2911–2923.

Kozanetsky, J., Knoerr, B., Harding, G., and Neitzel, U. (1987). diffraction measurements of some plastic materials and body tissues. *Medical Physics*, 14(4):526–532.

Laaziri, K., Sutton, M., Ghadirian, P., Scott, A. S., Paradis, A. J., Tonin, P. N., and Faulkes, W. D. (2002). Is there a correlation between the structure of hair and beast cancer or brca1/2 mutations? *Physics in Medicine and Biology*, 47:1623–1632.

Lazarev, P., Paukushto, M., Pelc, N., and Sakharova, A. (2000). Human tissue x-ray diffraction: Breast, brain and prostate. In *Proc. of the* 22^{nd} *annual EMBS Int. Conference*, pages 3230–3233.

LeClair, R. J., Boileau, M. M., and Wang, Y. (1999). A semianalytic model to extract differential linear scattering coefficients of breast tissue from energy dispersive x-ray diffraction measurements. *Medical Physics*, 33(4):959–967.

Leclair, R. J. and Johns, P. C. (2002). Optimum momentum transfer arguments for x-ray forward scatter imaging. *Medical Physics*, 29:2881–2890.

Lewis, R. A., Rogers, K. D., Hall, C. J., Towns-Andrews, E., Slawson, S., Evans, A., Pinder, S. E., Ellis, I. O., Boggis, C. R. M., Hufton, A. P., and Dance, D. R. (2000). Breast cancer diagnosis using scattered x-rays. *Journal of Synchrotron Radiation*, 7:348–352.

Limousin, O. (2003). New trends in cdte and cdznte detextor for x- and gamma ray applications. *Nuclear Instruments and Methods in Physics Research A*, 504:24–37.

Longo, R., Asimidis, A., Cavouras, D., Esbrand, C., Fant, A., Gasiorek, P., Georgiou, H., Hall, G., Noy, M., Ostby, J. M., Psomadellis, F., Rokvic, T., Royle, G. J., Schulerud, H., Speller, R. D., vand der Stelt, P. F., Theordoridis, S., Triantis, F., Turchetta, R., and Venanzi, C. (2007). A scanning

system for intelligent imaging: I-imas. *Proceedings of SPIE. Medical Imaging*, 6510.

Luggar, R. D., Farquharson, M. J., Horrocks, J. A., and Lacey, R. J. (1998). Multivariate analysis of statistically poor edxrd spectra for the detection of concealed explosives. *X-ray spectrometry*, 27(23):87–94.

Luggar, R. D. and Gilboy, W. B. (1999). Recent developments in industrial applications of elastic scatter x-ray inspection. *Radiation Physics and Chemistry*, 56(2):213–227.

Luggar, R. D., Horrocks, J. A., Speller, R. D., and Lacey, R. J. (1996). Determination of the geometric blurring of an energy dispersive x-ray diffraction (edxrd)system and its use in the simulation of experimentally derived diffraction profiles. *Nuclear Instruments and Methods in Physics Research A*, 383:610–618.

Luggar, R. D., Horrocks, J. A., Speller, R. D., and Lacey, R. J. (1997). Low angle x-ray scatter for explosives detection: a geometry optimization. *Applied Radiation and Isotopes*, 48(2):215–224.

Malden, C. H. and Speller, R. D. (2000). A cdznte array for the detection of explosives in baggage by energy-dispersive x-ray diffraction signatures at multiple scatter angles. *Nuclear Instruments and Methods in Physics Research A*, 449:408–415.

Marten, H. and Naes, T. (1991). *Multivariate Calibration.* John Wiley, New York.

Medawar, P. B. (1946). The rate of penetration of fixatives. *JR Microscope society*, 61:46.

Metz, B., Kersten, G. F. A., Hoogerhout, P., Brugghe, H. F., Timmerman, H. A. M., de Jong, A., Meiring, H., ten Hove, J., Hennink, W. E., Crommelin,

D. J. A., and Jiskoot, W. (2004). Identification of formaldehyde-induced modifications in proteins. *Journal of Biological Chemistry*, 278(8):6235–6243.

Meyer, P., Goergl, R., Botz, J. W., and Fratzl, P. (2000). Breast Cancer Screening Using Small-Angle X-ray Scattering Analysis of Human Hair. *J. Natl. Cancer Inst.*, 92(13):1092–1093.

Meyer, P. and James, V. J. (2001). Experimental Confirmation of a Distinctive Diffraction Pattern in Hair From Women With Breast Cancer. *J. Natl. Cancer Inst.*, 93(11):873–875.

Morgan, R. H. (1946). An analysis of the physical factors controlling the diagnostic quality of roentgen images part iii, contrast and the intensity distribution function of a roentgen image. *American Journal of Roentgenology*, 55:67–89.

Morin, L. R. M. (1982). Molecular form factors and photon coherent scattering cross sections of water. *Journal of Physics and Chemical reference data*, 11:1091.

Narten, A. H. and Levy, H. A. (1971). Liquid water; molecular correlation functions from x-ray diffraction. *Journal of Chemistry and Physics*, 55:2263–2269.

Neitzel, U., Kozanetsky, J., and Harding, G. (1985). Coherent scatter in radiographic imaging: a monte carlo simulation study. *Physics in Medicine and Biology*, 30:1289–1296.

Newton, M., Hukins, D. W. L., and Harding, G. (1992). Bone composition measure by x-ray scattering. *Physics in Medicine and Biology*, 37:1339–1347.

Nowotny, R. (1998). Institut fur Biomediz. Technik und Physik, Univeristy of Vienna. Attenuation coefficient calculation software.

Olivo, A. (2008). private communication.

Pani, S. (2007). private communication.

Pani, S., Jones, L., Griffiths, J., Royle, G., Horrocks, J., and Speller, R. (2006). Breast tissue characterization using energy-dispersive x-ray diffraction computed tomography. In *EXRS 2006*.

Parkin, D., Bray, F., and Pisani, J. F. P. (2005). Global Cancer Statistics, 2002. *CA Cancer Journal Clin*, 55(2):74–108.

Peplow, D. E. and Verghese, K. (1998). Measured molecular coherent scattering form factors of animal tissues, plastics and human breast tissue. *Physics in Medicine and Biology*, 43:2431–2452.

Poletti, M. E., Goncalves, O. D., and Mazzaro, I. (2001). X-ray scattering from human breast tissues and breast equivalent materials. *Physics in Medicine and Biology*, 47:47–63.

Poletti, M. E., Goncalves, O. D., and Mazzaro, I. (2002). Coherent and incoherent scattering of 17.44 and 6.93kev photons scattered from biological and biological equivalent samples: characterization of tissues. *X-Ray Spectrometry*, 31:57–61v.

Puchtler, H. and Meloan, S. N. (1985). On the chemistry of formaldehyde fixation and its effects on immunohistochemical reactions. *Histochemistry and Cell Biology*, 83(3):201–204.

Puumalainen, P., Uihmarituhta, U., Olkkonen, H., and Alhave, E. M. (1982). A coherent/compton scattering method employing an x-ray tube for measurement of trabecular bone content. *Physics in Medicine and Biology*, 23:425–429.

Rendle, D. F. (2003). X-ray diffraction in forensic science. *The Rigaku journal*, 19(2):11–22.

Round, A. R., Wilkinson, S. J., Hall, C. J., Rogers, K. D., Glatter, O., Wess, T., and Ellis, I. O. (1995). A preliminary study of breast cancer diagnosis using laboratory based small angle scattering. *Physics in Medicine and Biology*, 50:4159–4168.

Royle, G. J., Farquharson, M. J., Speller, R. D., and Kidane, G. (1999). Applications of x-ray diffraction analysis in crystalline and amorphous body tissues. *Radiation Physics and Chemistry*, 56:247–258.

Royle, G. J. and Speller, R. D. (1995). Quantitiative x-ray diffraction analysis of bone and marrow volumes in excised femoral head samples. *Physics in Medicine and Biology*, 40:1487–1498.

Ryan, E. A. and Farquharson, M. J. (2004). Angular dispersive x-ray scattering from breast tissue using synchrotron radiation. *radiation Physics and Chemistry*, 71:971–972.

Ryan, E. A. and Farquharson, M. J. (2007). Breast tissue classification using x-ray scattering measurements and multivariate data analysis. *Physics in Medicine and Biology*, 52:6679–6696.

Schulerud, H., Thielemann, J., Kirkhus, T., Kaspersen, K., Ostby, J. M., Meteaxas, M. G., Royle, G. J., Venanzi, C., vand der Stelt, P. F., Li, G., Turchetta, R., Theordoridis, S., Georgiou, H., Hall, G., Noy, M., Jones, J., Triantis, F., Longo, R., Bergamaschi, A., and Speller, R. D. (2007). Adaptive image content-based exposure control for scanning applications in radiography. *Proceedings of SCIA*.

Skipper, N. (2007). The effect of formalin fixation on edxrd studies of tissue. Project report.

Speller, R. D. (1999). Tissue analysis using x-ray scattering. *X-Ray spectrometry*, 28:244–250.

Speller, R. D. and Horrocks, J. A. (1991). Photon scattering- a 'new' source of information in medicine and biology. *Physics in Medicine and Biology*, 36:1–6.

Speller, R. D., Horrocks, J. A., and Lacey, R. D. (1993). X-ray scattering signatures for material identification. *SPIE Vol. 2092 Substance Detection Systems*, pages 366–377.

Strecker, H. (1998). Automatic detection of explosives in airline baggage using elastic x-ray scatter. *Medicamundi*, 42:30–33.

Strecker, H., Harding, G., Bomsdorf, H., Kanzenbach, J., Linde, R., and Martens, G. (1994). Detection of explosives in airport baggage using coherent x-ray scatter. In Harding, G., Lanza, R., Myers, L., and Young, P., editors, *Proc. SPIE Vol. 2092*, pages 399–410.

Suhonen, H., Fernandez, M., Serimaa, R., and Suortii, P. (2005). Simulation of small-angle x-ray scattering from collagen fibrils. *Physics in Medicine and Biology*, 50:5401–5416.

Tabar, L., Duffy, S. W., Vitak, B., Chen, H.-H., and Prevost, T. C. (2000). The natural history of breast carcinoma: what have we learnt from screening? *Cancer*, 86:449–462.

Taibi, A., Royle, G. J., and Speller, R. D. (2000). A monte carlo simulation study to investigate the potential of diffraction enhanced breast imaging. *IEEE Transactions on Nuclear Science*, 47:1581–1586.

Takahashi, T. and Watanabe, S. (2001). Recent progress in cdte and cdznte detectors. *IEEE Transactions on nuclear science*, 48(4):950–959.

Tartari, A., Casnati, E., Bonifazzi, C., and Baraldi, C. (1997). Molecular differential cross sections for x-ray coherent scattering in fat and polymethyl methacrylate. *Physics in Medicine and Biology*, 42:2551–2560.

Theodorakou, C. and Farquharson, M. J. (2008). Human soft tissue analysis using x-ray or gamma-ray techniques. *Physics in Medicine and Biology*, 53:R111–R149.

Trounova, V. A., Vazina, A. A., Lanina, N. F., Sergienko, P. M., and Kondratyev, V. I. (2002). Correlation between element concentrations and x-ray diffraction patterns in hair. *X-Ray Spectrometry*, 31:314–318.

Unscrambler v9.5 (2006). Camo inc., Norway, Multivariate analysis software.

Werner, M., Chott, A., Fabiano, A., and Battifora, H. (2000). Effect of formalin tissue fixation and processing on immunohistochemistry. *The American Journal of Surgical Pathology*, 24(7):1016–1019.

Westmore, M. S., Fenster, A., and Cunningham, I. A. (1996). Angular-dependent coherent scatter measured with a diagnostic x-ray image intensifier-based imaging system. *Medical Physics*, 23:723–733.

XCOMP. Robert Nowotny, Institut fur Biomediz. Technik und Physik, Univeristy of Vienna. Tungsten anode x-ray spectrum generating software.

Young, K. C., Burch, A., and Oduko, J. M. (2005). Radiation doses received in the UK breast screening programme in 2001 and 2002. *The British Journal of Radiology*, 78:207–218.