Processing of Spectral X-ray Data Using Principal Components Analysis

Anthony Butler (1,2,3), Nick Cook (2), Nannette Schleich (1), Jochen Butzer (1), Phil Bones (1), Phil Butler (1,3)

1) University of Canterbury, New Zealand
2) University of Otago Christchurch, New Zealand.
3) European Centre for Nuclear Research (CERN), Switzerland.

The goal of this work was to develop a technique to enhance contrast within the spectral domain so that variation in spectral characteristics within an image could easily be identified. The motivation is that using x-ray photon counting detectors, such as Medipix-3, it is now possible to obtain large data sets of spectroscopic data [1]. In particular there has been growing interest in spectral (multi-energy) computed tomography (CT) in the field on biomedical imaging.

Principal components analysis (PCA) was used to enhance the spectral images in the energy domain to identify of the number of independent patterns of spectral variation. The method differs from other forms of spectral analysis in that it requires no a priori knowledge of the materials being imaged, it can be applied to large data sets, and it can give an estimate of the total variance in the spectral domain.

PCA is a statistical technique for multi-dimensional data analysis [2]. It seeks to find a few variables that describe a majority of the variance within a multi-variate system. For spectral image enhancement, PCA is applied in the energy domain to identify the number of independent attenuation profiles within the data [3,4]. In PCA formulation, each measured energy is considered as one variable, while each pixel or voxel is one measurement of possible spectra. PCA then seeks to find a few derived vectors (often called eigen-spectra) which describe the majority of variance within the measured spectra. Each voxel in spectral CT data is transformed from being described by its attenuation over a range of energies to being described by the relative contribution of eigen-spectra, with the first few eigen-spectra containing a majority of the variance in the original data.

Calculating the eigenvectors from the full spectral data set can be done but the computational expense limits this approach to relatively small data sets. For this study PCA methods were adapted from face recognition. In particular, rather than finding eigen-spectra from the full data set which is computationally slow, eigen-spectra were found from a much smaller representative subset of the complete data.

Having identified variance in the spectral domain it is possible to represent this as a series of images each representing one pattern of spectral variance. Alternatively it is possible to produce images which combine the spectral data with traditional intensity images. In this approach the colour space (chroma) is chosen so that vectors that describe the majority of variance are coded as individual colours (eg. Red, Green, Blue). The intensity (brightness/luma) is given by non-spectroscopic data.

Using a Medipix based micro-CT scanner images were obtained of a mouse's thorax containing both iodine and barium (K-edge 2.5 keV apart). A pharmacological preparation of iodine was infused into the vascular system and a barium preparation was levaged into the bronchi. Four energy bin CT data was obtained. Initially, intensity (non-spectroscopic) data was used to perform CT reconstruction. Then each of the four energy bins were independently CT reconstructed to produce a 3 dimensional spectroscopic data set. PCA was applied in the energy domain. The three eigen-spectra that represent the most variance were used to produce colour (combined chroma and luma) images. This demonstrated that the calcium (bones), iodine (vascular), and barium (bronchi) are all distinguishable using this technique.
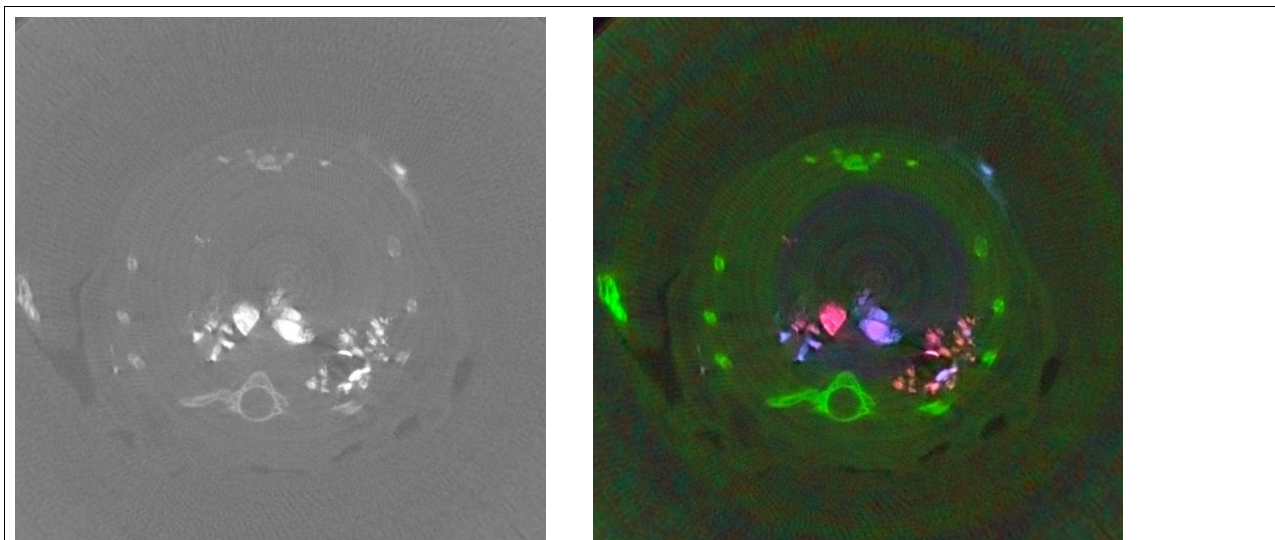
Left: Standard intensity CT of a mouse's chest containing bone, iodine and barium.

Right: PCA coloured spectral CT image. Bones are green, iodine in the vascular system is blue, and barium in the lungs is red.

[1] R. Ballabriga, M. Campbell, E. H. M. Heijne, X. Llopart, and L. Tlustos. The Medipix3 Prototype, a Pixel Readout Chip Working in Single Photon Counting Mode with Improved Spectrometric Performance. IEEE Trans. Nucl. Sci., Vol. 54, No. 5, October 2007.

[2] I Jolliffe. Principal Component Analysis. Published by Springer-Verlag, 2nd Edition, 2002.

[3] A. Kalukin, M. Van Geet and R. Swennen. Principal Component Analysis of Multienergy X-Ray Computed Tomography of Mineral Samples. IEEE Trans. Nuc. Sci., Vol. 47, No. 5, October 2000.

[4] J. Butzer, A. Butler, P. Butler, P. Bones, N. Cook, and L. Tlustos. Medipix Imaging - evaluation of datasets with PCA. Image and Vision Computing New Zealand, 2008 IVCNZ 23rd International Conference proceedings, 26-28 Nov, p1-6.