**Data Description**

*Complete raw data is available on request from Sarah Bohndiek (seb53@cam.ac.uk).*

**Folders**

|  |  |
| --- | --- |
| **Folder Name** | **Description** |
| Results/Data Tables (Attenuation) | Contains data tables of attenuation spectra. |
| Results/Data Tables (Reflection) | Contains data tables of reflection spectra. |
| Hb Analysis (Fig 4) | Contains the code used to generate Figure 4 from the data tables. |
| Colour Modelling (Fig 5) | Contains the code used to generate Figure 5 from the data tables. |

**Data Files**

*Acquired data is highlighted in green.*

|  |  |  |
| --- | --- | --- |
| **Containing Folder** | **Filename** | **Description** |
| None | wavelengths.mat | Array of wavelengths. |
| Results/Data Tables (Attenuation) | processed\_tissue\_spectra.mat | Data table of all attenuation spectra. Each row is a spectrum from a frame.   * Column 1: MuSE trial number * Column 4: Processed spectrum * Column 11: Final diagnosis:   + n = first region of this path, n+0.5 = second distinct region of this path   + Neoplasia, n = 3; Barrett's, n = 2; Squamous, n = 1 |
| processed\_tissue\_spectra\_avg\_per\_trial\_per\_region.mat | Data table of mean attenuation spectra per region. Each row is a mean spectrum from a region (per pathology).   * Column 1: MuSE trial number * Column 4: Mean spectrum in region * Column 6: Standard error of spectra within region * Column 8: n spectra * Column 11: Final diagnosis:   + Neoplasia, n = 3; Barrett's, n = 2; Squamous, n = 1 |
| processed\_tissue\_spectra\_avg\_per\_trial\_per\_region\_then\_pooled.mat | Data table of mean attenuation spectra per patient. Each row is a mean spectrum from a patient (per pathology).   * Column 1: MuSE trial number * Column 4: Mean spectrum in patient (per pathology) (mean per region, then over regions within patient) * Column 6: Standard error of spectra over regions * Column 11: Final diagnosis:   + Neoplasia, n = 3; Barrett's, n = 2; Squamous, n = 1 |
| processed\_tissue\_spectra\_avg\_overall\_distinct.mat | Data table of mean attenuation spectra overall. Each row is a mean spectrum from a pathology.   * Column 1: MuSE trial number * Column 4: Mean spectrum over all patients (per pathology)   (mean per region, then over regions within patient, then over all patients)   * Column 6: Standard error from standard deviation of mean-spectra\* from all patients   (\*mean per region, then over regions within patient)   * Column 8: n spectra * Column 11: Final diagnosis:   + Neoplasia, n = 3; Barrett's, n = 2; Squamous, n = 1 |
| Results/Data Tables (Reflection) | processed\_tissue\_spectra.mat | Data table of all reflection spectra. Each row is a spectrum from a frame.   * Column 1: MuSE trial number * Column 4: Processed spectrum * Column 11: Final diagnosis:   + n = first region of this path, n+0.5 = second distinct region of this path   + Neoplasia, n = 3; Barrett's, n = 2; Squamous, n = 1 |
| Hb Analysis (Fig 4) | Hb\_Bosschaart.mat | Data table of whole blood absorption and scattering coefficients from: <https://doi.org/10.1007/s10103-013-1446-7>   * Column 1: Wavelengths * Column 2: μa, SO2 = 98% [mm−1]   Column 3: μa, SO2 = 0% [mm−1] |
| Colour Modelling (Fig 5) | NBI\_green\_olympus.mat | Spectrum of NBI Green Filter Response   * Column 1: Wavelengths * Column 2: Response |
| NBI\_blue\_olympus.mat | Spectrum of NBI Blue Filter Response   * Column 1: Wavelengths * Column 2: Response |
| light\_olympus.mat | Spectrum of Olympus Light Source for NBI   * Column 1: Wavelengths * Column 2: Intensity |

**Analysis/Plotting Scripts**

*All analysis and plotting code is annotated in detail. Please refer to the headers in the code for a more detailed description of function.*

*Primary scripts are highlighted in yellow.* ***These can be run to generate the figures from the paper.***

*Support functions are highlighted in grey.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Containing Folder** | **Filename** | **Description** | **Loaded Data and Called Functions** |
| None | Plot\_Spectra\_and\_Variance\_Fig\_3\_Supp\_Figs\_7\_and\_8.m | Generates Figure 3, Supplementary Figures 7 and 8. | * Results/Data Tables (Attenuation)/processed\_tissue\_spectra.mat * Results/Data Tables (Attenuation)/processed\_tissue\_spectra\_avg\_per\_trial\_per\_region.mat * Results/Data Tables (Attenuation)/processed\_tissue\_spectra\_avg\_overall\_distinct.mat * ANOVA\_regions.m |
| None | ANOVA\_regions.m | Calculate within and between class variance for a cell array of data. | None |
| Hb Analysis (Fig 4) | Hb\_Analysis\_Fig\_4.m | Generates Figure 4. | * Results/Data Tables (Attenuation)/processed\_tissue\_spectra\_avg\_per\_trial\_per\_region\_then\_pooled.mat * wavelengths.mat * Hb\_Bosschaart.mat * Fit\_Mu\_A.m |
| Fit\_Mu\_A.m |  | * whole\_blood\_absorption.m |
| whole\_blood\_absorption.m |  | None |
| Colour Modelling (Fig 5) | Colour\_Modelling\_Figure\_5.m | Generates Figure 5. | * Prepare\_Data.m * AGauss.m * Model\_Spectral\_Band.m * sRGB2CIEDeltaE.m * light\_olympus.mat * NBI\_blue\_olympus.mat * NBI\_green\_olympus.mat |
| Model\_Spectral\_Band.m | Outputs a gaussian spectral band with a centre and height defined by the input. | None |
| AGauss.m | Generates an asymmetric gaussian. | None |
| sRGB2CIEDeltaE.m | Calculates colour difference, CIEDE, of a set of sRGB data pairs. | None |
| Prepare\_Data.m | Prepares the spectral data for modelling. | * Data Tables (Reflection)/processed\_tissue\_spectra.mat * wavelengths.mat * Find\_within\_class\_principle\_components.m * Augment\_Data.m |
| Find\_within\_class\_principle\_components.m | Performs PCA on dataset within each class. | None |
| Augment\_Data.m | Takes a mean spectrum with a standard deviation and generates a noised spectrum based on PCA-based data augmentation. | None |