

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. <ul style="list-style-type: none"> Column 1: MuSE trial number Column 4: Processed spectrum Column 11: Final diagnosis: <ul style="list-style-type: none"> 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). <ul style="list-style-type: none"> 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: <ul style="list-style-type: none"> 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). <ul style="list-style-type: none"> 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: <ul style="list-style-type: none"> 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. <ul style="list-style-type: none"> 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:

		<ul style="list-style-type: none"> ○ 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. <ul style="list-style-type: none"> • Column 1: MuSE trial number • Column 4: Processed spectrum • Column 11: Final diagnosis: <ul style="list-style-type: none"> ○ 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 <ul style="list-style-type: none"> • Column 1: Wavelengths • Column 2: μ_a, $SO_2=98\%$ [mm^{-1}] Column 3: μ_a, $SO_2=0\%$ [mm^{-1}]
Colour Modelling (Fig 5)	NBI_green_olympus.mat	Spectrum of NBI Green Filter Response <ul style="list-style-type: none"> • Column 1: Wavelengths • Column 2: Response
	NBI_blue_olympus.mat	Spectrum of NBI Blue Filter Response <ul style="list-style-type: none"> • Column 1: Wavelengths • Column 2: Response
	light_olympus.mat	Spectrum of Olympus Light Source for NBI <ul style="list-style-type: none"> • 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.	<ul style="list-style-type: none"> • 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.	<ul style="list-style-type: none"> • 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		<ul style="list-style-type: none"> • whole_blood_absorption.m
	whole_blood_absorption.m		None

Colour Modelling (Fig 5)	Colour_Modelling_Figure_5.m	Generates Figure 5.	<ul style="list-style-type: none"> • Prepare_Data.m • AGauss.m • Model_Spectral_Band.m • sRGB2CIEDE.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
	sRGB2CIEDE.m	Calculates colour difference, CIEDE, of a set of sRGB data pairs.	None
	Prepare_Data.m	Prepares the spectral data for modelling.	<ul style="list-style-type: none"> • Data Tables (Reflection)/processed_tissue_spectra.mat • wavelengths.mat • Find_within_class_principal_components.m • Augment_Data.m
	Find_within_class_principal_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