<u>INTRODUCTION</u>

Hello everyone, today I'm going to be demonstrating how to use the university of Liverpool's Euclid software, a package for performing orthogonal decomposition of data.

Orthogonal decomposition is a technique for reducing the dimensionality of 2D data fields. These data fields can contain any type of data, e.g.: image data, strain fields or displacement fields which are often represented by images. The data is decomposed into a set of coefficients using a set of orthogonal kernel functions. The coefficients are then collated into a feature vector. After this, a reconstruction of the original data can be obtained by multiplying each of the kernels by their associated coefficient in the feature vector. By using more coefficients, a better reconstruction can be obtained. Underneath this video you'll find a manual, which contains links to references describing the mathematics behind orthogonal decomposition in more detail.

This software package can be used to perform orthogonal decomposition on data contained in several different file formats. It can also be used to rapidly perform batch processing, on up to tens of thousands of files. Euclid can be used to perform simple image operations, such as rotating and cropping data, prior to the decomposition process. Decomposition can be performed using different sets of orthogonal kernel functions, but in this package Chebyshev polynomials have been used. I'll now introduce to you what the software looks like.

After installing the software click on the icon on the desktop or start menu to begin. From the offset, a splash screen, as seen here, will display before the main window loads. On the right-hand side of the window, you'll be able to see the programme controls, and on the left hand side is the main display area. At the top, you'll notice the control menu, where you can select which set of controls are displayed in the main window. There are two options to select from, the "view" option, which loads data into the programme and then shows decomposition results, and the "export" control, which determines how the data will be saved once processed. After the programme has opened, only the three import buttons under the "view" control will work, this is because image data must first be loaded into the program.

To import your data, click the "import" button at the bottom of the main window while it is in view mode. The other two buttons are shortcuts used for the batch processing of multiple images, which we will cover later in this demonstration.

After the "Import" button has been clicked, a standard file picker window will appear on the screen. This file picker can be used to select files which are one of five types: (insert images of types of files similar to a slide show set up) image files, Istra4D Export Files, text files (e.g. comma separated value files), Matlab data files and DeltaTherm files (.dt3 only). Multiple files can be selected, allowing the same procedure to be applied to each file, assuming they are of the same type. Whilst different file types can be selected simultaneously, only files of the same type as the first selected file will be processed.

Once the files have been selected, the "Import Data" wizard will be displayed, then the wizard guides the user through each step of the process. The first page of the wizard is different depending on the type of file being imported, but the subsequent pages are identical.

IMAGE FILES

I will begin this demonstration by talking you through the steps using an image file, and then will demonstrate the steps which are different for the other file types. Most of the steps are similar for all file types but differ depending on the file type being used.

Image files store information as a matrix of pixels. Each pixel is coloured to represent a value. To import an image file the relationship between the colour of each pixel, and the value that colour represents must be determined. This relationship is established on the first page of the "Import Data" wizard.

The image file is shown on the left-hand side of the wizard. A colour-bar is shown to the right of the image, this colour-bar indicates the colours the wizard expects to find in the image. Other software packages can produce images that use different sets of colours, called colourmaps, to represent data. The aim of this page is to choose the correct colour-map for the wizard and specify the values those colours represent.

You can change the set of colours using the drop down box shown on the right hand side of the wizard. Ideally, when importing data as images, the image data should be encoded as greyscale. This is because greyscale is a widely available colour set, simple to define, and experiences fewer artefacts when compressed with algorithms such as JPEG. However, this is not always possible, so two other pre-defined colour maps are available to be chosen; "jet" and "parula", from MATLAB. As the colour-map is changed in the drop down box, the colour-bar changes to represent that colour-map. The user must then decide if the colours match. If one of the colour-maps matches the data, then the values represented by the extrema of the colour-map should be written in the two boxes below the drop down box.

When none of the three colour maps are suitable for the image, there's a fourth option which is available via the dropdown box called "Use colour bar". When this option is selected, you'll notice that the colour bar disappears and is replaced by a blue line. This blue line can be positioned by the user along the full length of the colour bar in the image data by dragging the squares at the top and bottom of the blue line. You can also move the line by clicking on the thick blue line and dragging it. Ideally the ends of the line should be as close as possible to the end of the colour-bar without overshooting it. The ends of the blue line are defined as the ends of the thick line at the centre of the squares. To make this task easier the "Import Data" wizard can be enlarged to fill the entire screen. The colours along the length of the line are then sampled to estimate the colour-map for the rest of the image.

Once the import colour-map has been specified, the user can move to the next page of the wizard by pressing the "Next" button in the bottom right-corner of the window. The program will then attempt to find the closest match between the selected colour-map and the colours in the image data. The results of this matching process can be seen on the next page of the "Import Data" wizard. If the image doesn't look as expected, click the "Previous" button in the bottom right corner to return to the first page of the wizard and try the colour-map selection again.

SUBSEQUENT STEPS

Sometimes after importing your data, you'll need to rotate or flip it. To ensure the correct orientation of your images, the second page of the import data wizard allows you to apply simple image operations, such as flipping horizontally or vertically or rotating.

The third page of the insert data wizard allows you to crop your data. The decomposition applied by this program is applied to rectangular areas of data, which can be specified at this point.

To crop, simply drag the corners of the blue rectangle to the outside of your region of interest, data outside of this rectangle will be discarded, so if all of the data in the image is required, drag and drop the rectangle to the corners of your image.

Additionally, the interpolation type can be selected on this page. On the right hand side you'll see a drop down box containing several strategies for interpolating data over "Not a Number" values. There are 6 different interpolation types, None, Nearest Neighbour, Linear, Natural, Cubic and Bi-harmonic Spline - suggestions for when to use which method can be found in the Euclid manual.

The "Save Mask" button is used to save the position of the rectangle in the image. The mask is saved as a text file containing a single line of four integers representing the coordinates of the region of interest rectangle, which can be reloaded when importing a future dataset using the "Load Mask" button.

The "Normalise Data" checkbox, shown here (point) is used to control the magnitude of the data. When selected, the image data in the cropped rectangle is scaled such that the maximum value is equal to 1 and the minimum value is equal to -1. This normalisation scheme is used, for example, with image decomposition of mode shapes, where the magnitude of the measurements or predictions is not important.

DECOMPOSING IMAGES

The fourth page of the "Import Data" wizard is used to specify the parameters of the image decomposition. Currently, only image decomposition using Chebyshev polynomials is available. The "Number of Terms." text box is used to specify the number of coefficients that will be calculated and thus the size of the feature vector produced for each image. The number of coefficients can be any integer greater than zero. The maximum order of the Chebyshev polynomials can be calculated from the number of terms in the feature vector. More information can be found in the manual.

RESULTS

The final page of the import data wizard shows the results of the decomposition. The reconstruction of the image using only the data contained in the feature vector is shown here (point).

The accuracy of the reconstruction is shown in the text on the right of the window. The accuracy of the reconstruction is calculated by comparing the reconstructed image with the original image at the locations where interpolation did not occur. For measurement data, "Error" should ideally be just below the uncertainty of the measurement system, and ideally you don't want the cluster statistic to be any bigger than 0.3%. For more info on error calcs see the manual.

The reconstruction error is a good statistic for determining if the number of terms should be changed. If the calculated error is too high, return to the previous page and specify a higher number of coefficients. If the accuracy of the reconstruction is very high, then it implies that an unnecessarily high number of coefficients might have been used. This means the number of terms could be reduced, resulting in faster computation and greater reductions in data size.

FINISHED

The import data wizard is now complete and you have two options continuing forward. You can click the finish button just here (point), which will apply the same parameters specified in the wizard to each of the images selected using the file picker at the start of the import. The wizard will then close and the results of the decomposition will be displayed in the main window. Or, the second option is to save the recipe, which you'll see right here (point). This option treats all of the parameters defined in the wizard as a set of instructions that can be performed on any similar data. This is useful in situations where data from the same apparatus but different experiments is regularly processed. The "recipe" is saved as a ".mat" file that can then be applied to different data. If the results of the decomposition aren't required, click on the cross at the top-right corner of the "Import Data" wizard to return to the main window.

A quick tip for importing large amounts of data would be to first make a recipe to ensure the time taken to process them isn't too long. If you take the time to make a recipe using a small number of files, this can subsequently be applied to all of the data using the import with recipe button, or the import folder with recipe button, seen underneath. After selecting the recipe, a second file picker will appear, where you can select either multiple files or a whole folder, depending on which import button was clicked. The recipe will then be applied to every file of the chosen file format.

VIEWING DATA

After data has been decomposed it will be displayed in the main window like this. At the side here you'll be able to see where the view controls can be specified. Underneath, the file name drop down menu can be used to specify which file is being displayed, and the display image drop down box can be used to change which image is displayed. Here you'll see a number of options, original image, reconstructed, error, error cluster, mask and bar chart – for more information, see the Euclid manual.

EXPORTING DATA

The last step of this whole process is exporting your data. You can do this by selecting export in the drop down box at the top right corner, here. When the main window is in export mode, the window will look like this (point), and you can then use the controls displayed to save the data in different formats, for example, "Save as .csv" will allow you to save a comma-separated-value file that contains the feature vectors for all of the images. These feature vectors are stored as a table, where each column is a single feature vector describing an image. The order of the feature vectors is the same as the order of the files shown in the "File Name" drop down box when the main window is in "View" mode. "Save as .mat" button saves the data as a Matlab data file.

OTHER FILE TYPES

TEXT FILES

"Text Files" refers to tables of numbers stored in text, such as ".csv" files. Values on each row can be separated with: tabs, spaces, commas, semi-colons, colons, or pipes. Rows are separated using new lines. Files cannot currently contain information other than the pixel values – so any header information needs to be removed before import into Euclid. When you open a text file using the software, you’re taken straight to the window where you crop or rotate images. (Cut to screen grab and load txt files then say subsequent steps are the same as just demonstrated for image files)

MATLAB DATA FILES

Data contained in Matlab data files (.mat) can also be imported using the wizard. When importing Matlab data, the first page of the “Import Data” wizard will look like this (show on video). The wizard searches the data file for matrices of double or single precision floating point numbers but cannot locate matrices which are contained in data structures or cell arrays. The variable names for any suitable matrices located are then displayed in the drop down box here (point). When the drop down box changes, the selected matrix will be displayed. The imported matrices can contain “Not-A-Number” values. This value can be used to represent locations where data is unavailable and can be interpolated over at a later stage of the import.

ISTRA4D

This file type is exported by the Digital Image Correlation software package Istra4D, developed by Dantec Dynamics, and has the extension “*.hdf5”. On the first page of the wizard the desired measured quantity is selected using the drop down box on the right-hand side. Locations where Istra4D data is not available will be represented with “Not-A-Number” values. (point to the missing white data part) at a later stage of the import, the user can choose to interpolate or crop when working with such data fields

When the drop down box selection is changed, the display updates to show the desired data. Istra4D files also contain coordinates for each of the facets. If the checkbox “DIC Coordinates” is selected, then the strain or displacement fields are displayed using the spatial data, shown here in the wizard. When using the coordinates, interpolation is required to map all of the measurements onto a uniform grid, which can drastically increase the computation time when working with large datasets.

DELTATHERM

Thermoelastic stress analysis (TSA) data can be imported from DeltaTherm files (“.dt3” only). When importing TSA data, the first page of the “Import Data” wizard looks like this. The drop down box at the right of the window can be used to select the X, Y, R, phase or DC images encoded in the file. When the selection changes, the display updates to show the desired image.

To clarify anything mentioned in this tutorial, please refer back to the euclid manual, otherwise, thank you for watching.