Overview of the Tramline Map Generation Code for 2dfdr

m. birchall

June 19, 2015

1 Introduction

This document focuses on the code specified for the "make tlm other" option and the reasoning behind them.

The traditional approach to the tramline development has been:

1) Take a spatial slice in the middle of the flat image, with some sort of combination of the neighbouring columns to reduce noise

2) Analyse that slice for peaks that conform to the spatial psf of a typical fibre

3) Apply a distortion model upon horizontal lines that passes through the identified fibre trace peaks of the central column.

4) Optionally determine an extra shift and or rotation required to improve the fit. This is typically done by repeating the above slice analysis for two other columns one left and one right of the central column and scuccefully matching the identified fibers thereby determining how much of a shift and rotation needs to be added.

This approach shows success in a majority of cases, but will ocassionaly fail either in the fibre position identification of the middle spatial slice usualy due to noise or in the determination of the shift and rotation factors again usually due to noise. The cause of these failures can be deemed a consequence of a lack of redundancy.

A secondary level of analysis termed "Ridge Tracking", was developed intended as a slow but sure means of tracking where fibre ridge peaks were and optionally fitting smooth polynomials to these tracks. This method was found to be slow but effective as long as it was given starting points sufficiently close to the fibre traces. Again, however, noise from a low redundancy method of deriving starting points was enough to cause this method to go awry.

In order to improve the success rate of tramline mapping we need to increase redundancy, i.e. take information from as many on the image columns as is practical and apply robust methods that can track and trace with the assumption of incorrect or missing values in a minority of cases.

The steps for the tramline generation code is as follows:

1)Locate traces across the image data

2) Match the traces found to fiber identifiers

3) Review any important fibers missing from the identification process

4) Interpolate missing fibers

We review each in turn

2 Locate traces across the image data

Note, this code resides in the file

`locate_traces.F95`

Locate traces consists of the following steps:

1. Sweep image spectrally for signs of fibre trace peaks at specific columns

2. Link peak points found into a set of traces across the image

3. Use those traces to define a first order empirical distortion map

4. Perform a basic extraction of the image for each spatial pixel utilising the distortion model

5. Average spectrally the extraction to produce a noise suppresed "representation" spatial slice.

6. Analyse the flux peaks of the "representation" slice to identify all fibre traces present in the image

7. From the fibres fluxes identified in the reprsentation slice and the distortion modle, create a first order tramline map for all identified fibre trace fluxes.

8. From the first ordre tramline map, determine the fibre trace peaks for each fibre trace at each column.

9. Perform a model fit for the located peak positions for each trace at each column.

2.1 Locate Traces: Sweep Image

Locate traces makes a spectral sweep across the image data analysing spatially for what appears most likely to be peaks of fiber traces. Ideally each column would be investigated but such a sweep for AAOmega data was found to take about 10 minutes which is deemed to long so steps of between 10 to 100 columns is being used. The purpose of this sweep is to derive a list of peak positions (as a set of reals representing pixel coordinates) with column indices which can be percieved as a peak map:

```
|----------------------------------------------|
|      *                                       |
|          *   *   *                          |
|                      *                       |
|      *   *   *   *                           |
|                      *      Spatial Pos      |
|                                 ^            |
|      *   *   *   *              |            |
|      *   *       *   *          ---> Spectral |
|      *   *   *   *   *               Column  |
|                      *                       |
|----------------------------------------------|
```

2.1.1 Identifying Fibre Trace Peaks in a spatial column

Identification of peaks related to fibre traces across a single column slice is hampered by noise generated local peaks. Previous attempts looked for any local peaks that matched a set of criteria using instrument specific parameters. In general this would work for the majority of cases but had a tendency to fail unpredicatbly when noise conditions were just right. For our puproses we do not require all fibre peaks at this stage, but we do want consistency in selection and high confidence in what has been selected are fibre trace peaks. It is for this reason that we use the watershed method of emergence to locate the fibre peaks. Emergence assumes that data represents a sufrace total immesred in water and then raises it slowly. At any level of emergence there will be a number of islands defined across the water and each island has it's own single peak. The purpose of emergence is to choose an appropriate cutoff level and use the peaks of the emerged islands as the identified values. The general fundamental issue with emergence is choosing the appropriate cutoff level. Too high and we only obtain a few peaks as the rest are still immersed. Too low and we only obtain a few to one peak as all the islnads merge together. However, for a well defined flat field profile of good fibers, such as found with AAOmega data, there is a good range of levels that return all fibre trace peaks which can be determined via histogram analysis. With flat field data that appears to follow multi tier trends, such as early Sami data that had fiber traces from some blocks providing strong traces and some from blocks providing weak traces and also with 6df data that has great variation in troughput strengths there is no ideal single level. Under these circumstances the best returns can be achieved by using historgram analysis to return the maximum number of peaks but it will be a compromise as adding lower strength peaks will cause islands surrounding the higher strength peaks to merge. However, there will be consistency based on consistency of the relative throughputs as this method will return the highest peak of consistent subsections, i.e. there is a high likelyhood that the peaks returned and the peaks missed will be the same for each column. Further, the peaks missed will be close to the peaks selceted therby providing a level of homogeiaty in the placements of the returned peaks. Note homogeniaty cannot be guaranteed as there will always be flux strength distribution that can defy the analysis. Given that it is the aim to use trace paths to consruct an empirical distortion model, it is preferable to be able to select fibre traces both at the upper and lower regions of the image. To insure this, we therfore subdivide the spatial pixel range into 3 slightly overlapping regions and perform emergence on each and then merge the results. Thi sensures that we can find traces in all 3 regions provided that there are traces in these regions. Thus we choose emergence peakfind as it demonstartes consistency in selection, provides high confidence in what is returned is a fibre trace peak and is free of instrument specific paramaters.

It is this authors opinion, however, that if modifcations are later required that remove the overall simplicity of the emerge method, then other methods may be more efficacious.

2.2 Locate Traces: Link Peak Points into Traces

The next step is to link all the peaks into a set of trails

Note, it is assumed from the prvious step that not all fibre traces are located at each column, but that the peaks found are most likely from fiber traces. It is also assumed that there is some consistancy in which fibre peaks are located in each spectral slice but not total.

```
|----------------------------------------------|
|      *-\                                      |
|          *---*---*\                           |
|                     -*                        |
|      *---*---*---*\                           |
|                     -*      Spatial Pos       |
|                                ^              |
|      *---*---*---*\            |              |
|      *---*-------*\ -*         ---> Spectral  |
|      *---*---*---*\ -*              Column    |
|                     -*                        |
```

```
|--------------------------------------------|
```

This problem is equivalent to the multi target tracking problem often used in video suravalence and most particulary in defence, e.g. given a series of snapshots of an area with a number of objects identified as enemy tanks at specific coordinates, determine the tracks of those tanks given the assumption that some may be obscurred from view, i.e. behind some other structure, in some of the snapshots. Though it is difficult to find actual source code for this method, there are many publications presenting the outline which is a simple repetitive usage of a "linear assignment problem" solver for which there are many source codes available including a Fortran 90 version.

2.3 Locate Traces: Use found traces to define a first order empirical distortion map

Given a list of tracks consisting of a list of points associated to those tracks and the assumption that some of those points are incorrect either in position due to noise or due to mis-assignment, we need to determine in a robust sense the curves that adaquately describes these paths. In our experiance, we know that a quadratic is sufficient to describe these paths at least to within one pixel. Thus the curves can be determined via robust quadratic fitting in the sense that we find a quadratic with minimum sum of absolute deviations to the fitted points. The benefit of such a fitting is that its solution is dependent on the majority of points that collectivly describe a single quadratic and is independant of the minority of points that do not no matter how large their inconsistancies are.

Given the assumption that we have empirical curves both in the upper and lower regions of the image and that the number of curves derived are significantly greater than 3, we are then able to analyse how these curves vary according to a relative spatial positioning and thereby create an empirical 2D distortion model. We achieve this in a simple manner by representing the distortion model as measures of by how much a tramline passing through a specific point in the central column would deviate from horizontal on the leftmost side and the righmost side of the image.

2.4 Locate Traces: Representation Spatial Slice

Given the empirical 2D distortion model, we are able to perform a combination of each column corrected for ditortion to create a single spatial slice representation that has strong a signal to noise ratio. The simplest way to acheive this is to generate a pseudo tramline map for curves that pass through the mid points of each pixel in the central column, perform a "TRAM" extraction for this pseudo tramline map and then average the resultant columns. Note, give that the 2D distortion model is only expected to be correct to about 1 pixel, we can expect a "blurring" effect of the fibre psfs by around the same amount.

2.5 Locate Traces: Analyse the flux peaks of the "representation" slice

Use Peak Analysis routine to find all peaks that are of sizes at least 10.0 percent of the maximum peak and any other that shares the same shape. Insist that we obtain at least the traces found but no more than the number of fibers of the instrument.

2.6 Locate Traces: Create a first order tramline map

From the fibre trace peak locations found in the representation slice and the empirical 2D distortion model, we can create a "first order" tramline map.

2.7 Locate Traces: Determine the fibre trace peaks per column

Given that the "first order" tramline map is correct to around 1 pixel, we are able to utilise the section of code of the "Ridge Tracker" module that determines the peak positions of the fibre traces at each column subject to bad pixels and loss of signal.

2.8 Locate Traces: Determine the fibre trace peaks per column

From the array of fibre trace positions, we are able to fit models of the tramline to derive smooth tramlines. This can either be per fibre basis or for all fibers collectively. As a first cut we choose the simplest case of modelling each fibre trace as a polynomial.

3 Match the traces found to fiber identifiers

Traditional approaches to this problem has been to assume a nominal gap size between all fibres, sometimes kludging to add non existant fibers to derive a relatively uniformly spaced model which is then . Again this approach has been found to work the majority of times specifically because sperate association code has been made for each instrument. We find that the majority of cases can be generalised by matching gaps is we provide nominal values for each fibre gap, if nothing else this removes the need for pseudo fibers and kludges for the odd gap that is significantly larger than the rest and makes the code's behaviour significantly more predictable.

The other problamatic issue with fibre identification is that fibers that certain fibers eg those that are parked may or may not leave either a weak or a strong trace If the edgemost fibers were always in use, i.e. either a P type or S type, then identifications of the inbwteen fibres would be a stright forward process of interpolation of the gap model. However this is never a guarantee. The new approach is to utilise all header information with regards the fiber types and to assume that if a fibre type is P or S, then it's trace will be present, if it is of type F (fiducial guide star) then it is not, but if it is of any other type then it is a maybe. From this we can count the number of the "maybe" types that trail at the edgmost Thus ignoring physical gap distances, there would be a possible (n+1)*(m+1) arrangments of the trace association for the first and last fibers officially in use. In principle this could be up to almost the total number of fibres of the instrument squared eg for KOALA 1000**2. However, in practice the majority of these combinations would predict either ridiculously high or low values for the gaps. Thus we can minimise the number of combinations to consider by asserting scale limits on the gap size eg 50 percent of the nominal values. This reduces possible end fibre combinations from millions to the order of 10s.

Given an edge fibre identification, we can scale the nominal gaps to derive a model position for each fibre. We can then apply a "linear assignment problem" solver to find the best matching of found traces to the fiber position model using the Euclidean distance between the two points as the cost function. The solver will derive the matches and a list of what could not be matched and a cost for the solution matching. We can perform this analysis on all of the edge matched combinations deemed physically viable. In an ideal world there will only be one combination that matches all the P and S type fibres. In cases where there are more than one such combinations, the one with the minimum cost is the best match possible, but it is worth highlighting that there is a small chance of a mis-match. In cases where there are more than one combinations that match all the P and S type fibers and have the same minimal cost, then it is clear that there is no unique solution and the user should be informed. Finally, in the event that there is no combination that matches all P and S type fibres, then a solution can be chosen by excepting the combination with mininal failed matches and minimal cost, but the user needs to be informed that a fibre trace that is expected to be present could not be found.

4 Review any important fibers missing from the identification process

Given that there are P or S type fibers that could not be identified through the previous section, it is feasible to use the gap model, the empirical distortion model and the representation slice to check if it is feasible that this trace does exist but was too weak to detect. At this point in time the code does not do this as we need to check on how rare a case this is.

5 Interpolate missing fibers

With the exception of missing "P" or "S" type fibers, interpolation of missing fibers has no impact on the end result science data. It does, however, have impact on image displays and on parts of the code that does not check for fibre types but does need data within valid ranges, and finally on the regression tests as the tests has yet to conseider fibre types and does not ignore data from irrelevent fibres. The simplest and asthetic solution is to scale the gap position models to interpolate and extrapolate fibres from the nearest non-missing fibres. However, at the moment, due to implementation expediency, interploation and extrapolation is performed based on the position of the nearest non-missing fibres.

6 Issues to consider

IFU only has U type fibres by default. We must readdress how IFU is handled correctly.

Problems with AAOmega Nod and Shuffle and AAOmega Mini Shuffle

Subdivion of spatial range during emerge peak find

Handling sparse fiber data