

**OPERATING INSTRUCTIONS FOR  
COPA (Computer Aided Photoelastic Analysis) PROGRAMME:  
VERSION 3**

Philip Siegmann & Eann Patterson

**Introduction**

This document accompanies a new piece of software initially produced in the Experimental Stress Analysis Laboratory in the Department of mechanical Engineering at the University of Sheffield. COPA version 3 is written in matlab p-files and uses an improved unwrapping algorithm with respect to previous versions. It is based on an algorithms which are described in [Siegmann P., Beckman D. and Patterson E.A. A robust approach to demodulate and unwrapping phase-stepped photoelastic data,” Exp. Mech. 45(3), pp. 278-289, 2005] and are executed in MatLab. The objective is to process phase-stepped photoelastic images acquired in either a transmission or reflection polariscope. The input for the software is .six files generated by the ‘Catchsix’ software also produced by the Experimental Stress Analysis Laboratory. Data output is in several forms as described below including a new file suffixed cpa. A paper describing the algorithms and providing examples and validation is in preparation for submission for publication in the open literature. For more information contact Philip Siegmann (philip.siegmann@uah.es).

**Installation**

The current COPA program is written in Matlab®, therefore you need to have the Matlab program installed in your computer.

Copy the COPA directory into any directory of your choice in your computer. Open Matlab and select as ‘Current Directory:’ the COPA directory, or set a path to it (press ‘File’, ‘Set path...’, ...). Write COPA in the Matlab command line and press enter.

**Operation**

The COPA window will display a front page and then move the main window. To process photoelastic images, start from the top left and work leftwards and down using at least one button in each dark grey zone. The results of each operation will be shown in the figure space

in the bottom left of the window and a red caption below the figure will give an instruction on what to do next.

### **Data Acquisition**

Data should be acquired using the 'Catchsix' program to capture six phase-stepped images and package them in a .six file. Use the 'Open' button to identify the appropriate .six file. Note that other file formats can be read, please consult the authors for more information. After selecting the file to be opened the six images will be displayed in a new window labelled as shown in table 1 in the Appendix using pseudo-colour to represent intensity. This window is removed automatically after a few seconds, and the wrapped isoclinic map corresponding to equation (1) in the Appendix is displayed in the main window.

### **Saving data**

Once a file has been opened the data displayed in the figure in the main window can be saved using the 'Save' button after selecting the appropriate file format using the drop-down menu to the left of the button. If no selection is made the default file type is txt.

### **Background Removal / Applying a Mask**

The dark grey box in the top right corner of the main window offers two choices: importing a mask or creating a new one. A mask that has been created by 'Catchsix' can be imported by opening a .bng file via the 'Import mask' button. The mask is applied automatically and the background in the image will be removed in the figure displayed in the window. The data saved as .cpa save the six phase shifted image and the created mask in 'Crate Mask' option. By open a .cpa file also the corresponding mask is loaded automatically. (The cpa file, like .six, is an adapted .bmp file, but with additional mask included).

Alternatively, a new mask can be created by using the 'Create Mask' button. A new window will appear with the image corresponding to a dark field circular polariscope on the left and a series of buttons on the right. You can choose to define a mask using a circle, a polygon or any combination of these shapes. Select the required shape and follow the on-screen instructions. It is possible to cut out and enlarge an area of the image using the 'Cut out' button to define the area of interest. Once a shape has been defined you must indicate whether the area to be masked is inside or outside the shape using the buttons: 'Mask inside' or 'Mask outside' respectively. Press 'Finished Mask' to return to the main window. It is possible to use the functions repeated to build a complex shape for the mask. If a mistake has been done by

creating mask you can either click the right mouse button during the input of points, or close the 'create mask' window without pressing 'Finished Mask'.

### **Isoclinic Demodulation**

This function is not essential but is recommended when the isochromatic fringe range is greater than half a fringe, because the half order fringes will cause errors in the isoclinic map as a result of modulation. Also  $45^\circ$  isoclinics cause errors in the isochromatic map via the same modulation. The process described in the Appendix in lines 43 to 124 is applied when the 'isoclinic demodulation' button is used. This process is almost entirely automatic except the operator being asked to select the location from which to start the unwrapping of the areas described starting at line 64 in the Appendix. It is best to choose the largest area displayed in the temporary window because this will provide the fastest solution. The demodulated isoclinic map is displayed in the figure in the main window when the process is finished.

It is possible to adjust the window size,  $w_2$  used in the demodulation process using the slider bar below the button. However, the default value is usually satisfactory. More details about this window size are given in lines 108 to 118 in the Appendix.

The default is for a Wiener filter to be applied to the isoclinic data prior to demodulation as described in lines 91 to 99 of the Appendix. This filtering can be omitted by removing the tick from the box beneath the 'Demodulation' button. It should be noted that regardless of the use or otherwise of this filter, unfiltered isoclinic data is always used in the calculation of the relative retardation data.

### **Relative retardation/Isochromatic Phase Map**

On selecting the 'Isochromatic Phase Map' button the wrapped relative retardation is calculated using equation (4) in the Appendix and displayed in the figure. The default calculation is to use the demodulated isoclinic map, but the raw isoclinic data can be employed instead by removing the tick in the box beneath the 'Isochromatic Phase Map' button.

### **Isochromatic Fringe Order**

The relative retardation or isochromatic phase map (same data but alternative names) is periodic or wrapped and must be unwrapped to generate a continuous map of fringe order. This process is performed when the 'Unwrap isochromatic' button is pressed and is described in lines 125 to 137. The process can be performed completely automatically with the seed point being selected as the pixel having the highest quality. To select the seed point manually

tick the box under the 'Unwrap isochromatic' button and a temporary window will appear. In this window the quality map is displayed in grey scale for each pixel, darker pixels have less quality value.

The unwrapped (continuous) map of isochromatic fringe order is displayed in the figure in the main window at the end of the process.

The unwrapping process automatically sets the lowest fringe order in the map to zero so that no calibration may be needed, otherwise see below. There is some ambiguity in the definition of isoclinic angle and hence in the solution for the isochromatic fringe order, such that the resultant map may be inverted as explained in line 140 to 144 of the Appendix – simply press the 'Invert' button to reverse the distribution.

Select '5 bins Unwrapping' for a faster unwrapping, for more precise but slower unwrapping select '10 or 100 bins Unwrapping'. The respective quality maps can be seen when the option of 'Select Initial Pixel' is activated. For noisier data, more bins should be considered for the unwrapping.

### **Calibration**

The unwrapped isochromatic map may require calibration if there is no zero order fringe present. Enter the value of the fringe order you intend to use for calibration in the box in the dark grey zone in the bottom right corner of the main window. The use of half order or integer fringes is recommended. Then, press the 'Calibration in N' button and a temporary window will appear with a display of  $\frac{1}{4}$  order contours. Select a point on an appropriate contour with the mouse and right click. The process can be performed as many times as necessary.

Subsequently a map of maximum shear stress or strain (or difference in principal values) can be generated by performing a further calibration. Again enter the value of the stress or strain metric in the same box, but also type in the appropriate units in the box below, then press the 'Calib. Stress/strain' button and select the point at which the stress or strain is known. This might be the same or a different point to the one used for the fringe order calibration.

### **Data Display**

The figure (including the colour bar) can be saved using the 'Save current figure as .eps, .tif, .jpg' button. The data in the current figure can be saved using 'Save current figure as .txt' (ascii file), the data will be saved in matrix format (same size as the figure) and in the head line of the

created ascii file are given the units. The button 'Save \*.cpa file' do not save the processed image, only the six phase shifted images and the mask.

Profiles across the data displayed in the figure can be obtained by pressing the 'Profile' button and then left clicking on a start point in the map followed by a left click of the end point of the desired profile followed by pressing the 'Enter' key. A temporary window is displayed with the plot of the data as a function of pixel location along the profile.

After data processing the map displayed in the figure can be selected using the drop-down menu marked 'Show' below the 'Profile' button. Similarly, the units of the data can be changed using the adjacent drop-down menu marked 'Units'.

The colour scheme for the figure can be selected using the drop-down menu marked 'Color Map'.

## Appendix: Description of the algorithm

### Nomenclature

$A(i)$	Initial area for processing
$A(i+1)$	Next area for processing
$h$	Threshold height used in spike removal process applied to partial derivative data
$i_1...i_6$	Intensity measured in polariscope orientated as shown in Table 1
$i_m$	Intensity term to account for stray light entering the polariscope
$i_v$	Intensity observed in a polariscope when the fast axes of all the elements are aligned
$I_c$	$= i_4 - i_6$
$I_d$	$= i_1 - i_2$
$I_s$	$= i_5 - i_3$
$n,m$	Co-ordinates of pixel of interest
$N,M$	Number of pixels in $x, y$ direction in data map
$Q_{\theta}(n,m)$	Quality value for the pixel located at $(n,m)$
$w_1$	Window dimension used in evaluating quality map
$w_2$	Window dimension used in reducing number of areas to be processed.
$\alpha$	Relative retardation
$\Delta\psi_x(i,j), \Delta\psi_y(i,j)$	Difference in partial derivative of the phase data with respect to the $x, y$ directions
$\overline{\Delta\psi_x}(n,m), \overline{\Delta\psi_y}(n,m)$	Mean of the differences in partial derivative of the phase data with respect to the $x, y$ directions within the window
$\theta$	Isoclinic angle
$\rho$	Percentage area left unwrapped
$\sigma_1, \sigma_2$	Maximum and minimum principal stresses

### 5 Introduction

The aim is to generate a continuous map of isochromatic fringe order that is independent of the isoclinic angle and similarly to produce a modulo  $2\pi$  map of isoclinic angle defining the direction of one identified principal stress that is independent of the isochromatic fringe order. Possession of the latter enables a relatively straightforward calculation of the former using the phase-stepping relationships. Ramesh and Ganapathy [1996] identified the phase-stepping algorithm given by Patterson and Wang [1991] as the most robust so it is used here as the fundamental step to obtain phase data. Six phase steps are employed in a circular polariscope as provided in Table 1 such that the isoclinic angle is obtained as:

$$15 \quad \theta = \frac{1}{2} \arctan \frac{I_s}{I_c} \quad (1)$$

where  $I_s = i_5 - i_3$  and  $I_c = i_4 - i_6$ . The relative retardation can be obtained using many alternative formulae but will always contain the modulation introduced by the ambiguity in the isoclinic data. The relative retardation is usually defined as:

$$\alpha = \arctan \frac{I_c}{I_d \cos 2\theta} \quad (2)$$

20 where  $I_d = i_1 - i_2$ . Patterson and Taroni [2003] have recently shown that if the usual arctangent operator is replaced by an arccotangent function by inverting expression (2) then the period boundaries in the relative retardation are translated from  $-\pi/4, \pi/4, 3\pi/4 \dots$  to  $0, \pi/2, \pi \dots$ . This is more convenient for unwrapping and interpretation of the data because the period boundaries correspond to half order and integer fringes  
 25 which are readily observed in a manual polariscope; and because fewer unwrapping operations may be necessary. To further reduce the number of unwrapping operations, a simple logic operation has been employed to effectively extend the range of relative retardation from  $\pi/2$  to  $\pi$  [Sparling et al], using:

$$\alpha = 2\pi + \text{arc cot} \frac{I_d \cos 2\theta}{I_c} \quad \text{for } i_1 > i_2 \ \& \ i_4 < i_6 \quad (3a)$$

$$30 \quad \alpha = \pi + \text{arc cot} \frac{I_d \cos 2\theta}{I_c} \quad \text{for } i_1 < i_2 \quad (3b)$$

$$\alpha = \text{arc cot} \frac{I_d \cos 2\theta}{I_c} \quad \text{else} \quad (3c)$$

A further refinement was introduced to ensure high modulation over the entire data field and hence to reduce noise by merging the ideas given above with the expression given by Quiroga and Gonzalez [1997] to give:

$$35 \quad \alpha = 2\pi + \arctan \frac{I_s \cos 2\theta + I_c \cos 2\theta}{I_d} \quad \text{for } i_1 > i_2 \ \& \ i_4 < i_6 \quad (4a)$$

$$\alpha = \pi + \arctan \frac{I_s \cos 2\theta + I_c \cos 2\theta}{I_d} \quad \text{for } i_1 < i_2 \quad (4b)$$

$$\alpha = \arctan \frac{I_s \cos 2\theta + I_c \cos 2\theta}{I_d} \quad \text{else} \quad (4c)$$

Maps of the isoclinic angle and relative retardation obtained using expressions (1) and (4) are shown in figures 1a and b. The influence of the isoclinic angle on the relative  
40 retardation map is clearly seen at locations where the isoclinic angle is 45 degrees and similarly the influence of the relative retardation on the map of isoclinic angle is clearly seen at the locations of half order and integer fringes. In order to demodulate this data, the map of isoclinic angle is examined and the boundaries of the regions associated with each principal stress are identified using a zero-crossing boundary  
45 detection technique. This technique is similar to edge detection methods in image processing. Briefly, in this technique two matrices, termed ‘positive’ and ‘negative’ are created that are equal in size to the data array being considered and initially populated with zero values. A raster scan of the isoclinic map is performed pixel by pixel with a two by two window formed by the pixels to the right, below and diagonal  
50 to the right of the pixel being considered. The sign of the isoclinic angle at each pixel in the window is examined, if the sign of all the values are the same then no action is taken and the scan moves on the next pixel, otherwise for those pixels in the window with a negative values of isoclinic angle the corresponding zero in the ‘negative’ matrix is changed to one and similarly for the pixels with positive values the  
55 corresponding zero in the ‘positive’ matrix is changed to one. The scan continues so that when it is completed the location of all the zero-crossings is identified by unity values in the ‘positive’ and ‘negative’ matrices. Negative areas are defined as those enclosed by a discontinuity represented by a positive or upward step in the isoclinic angle which will occur at pixel locations defined by the unity values in ‘positive’  
60 matrix, similarly positive areas are enclosed by a negative or downward step in the isoclinic angle defined in the ‘negative’ matrix. The zero-crossing boundaries and corresponding areas for the data in figure 1a are shown in figure 1c with the positive areas outlined in white and the negative ones in black. These areas are subsequently unwrapped commencing from  $A(i)$ , which is best assumed to be the largest, and  
65 employing a quality map, based on the phase derivative variance procedure proposed by Pritt [1996]. The quality of a pixel at location  $(n,m)$  was identified using

$$Q_{\theta}(n,m) = -\frac{1}{w_1^2} \left[ \sqrt{\sum_{i,j \in w_1, w_1} (\Delta\psi_x(i,j) - \overline{\Delta\psi_x}(n,m))^2} + \sqrt{\sum_{i,j \in w_1, w_1} (\Delta\psi_y(i,j) - \overline{\Delta\psi_y}(n,m))^2} \right] \quad (5)$$

where  $\Delta\psi_x(i, j)$  is the difference in the partial derivative with respect to  $x$  of the isoclinic angle obtained from equation (1) for the pixel located at  $(i, j)$  in a window  $w_1 \times w_1$  centred on the pixel  $(n, m)$ . Prior to applying the difference operation, the partial derivative is smoothed in a ‘wrapping’ operation that involves scanning the partial derivative data for spikes of height  $h/2$  and translating the values in the spike by  $h$ , where  $h$  is  $\pi/2$  for the isoclinic data.  $\overline{\Delta\psi_x}(n, m)$  is the mean value of the difference in the partial derivative over the window.  $\Delta\psi_y(i, j)$  and  $\overline{\Delta\psi_y}(n, m)$  are corresponding parameters in the  $y$  direction. It is worthy of comment the  $\Delta\psi_{x,y}(i, j)$  term is effectively a second order partial derivative of the phase map. . The boundary of  $A(i)$  is searched for the highest quality pixel and the area on the other side of the boundary at this point is  $A(i+1)$ . When a discontinuity exists between areas  $A(i)$  and  $A(i+1)$  of magnitude between  $\pi/4$  and  $3\pi/4$  then value of the isoclinic angle for all points within  $A(i+1)$  are translated by  $-\pi/2$ ; similarly when discontinuities between areas  $A(i)$  and  $A(i+1)$  of magnitude between  $-\pi/4$  and  $-3\pi/4$  are identified then all the points with  $A(i+1)$  are translated by  $\pi/2$ . Then, the area  $A(i+1)$  is incorporated into  $A(i)$  and the pixel with the highest quality value on the new boundary is identified and the process repeated until all the areas have been considered. It should be noted that only the discontinuity across the boundary at the pixel with the value is  $Q_{\theta max}$  is considered at each step. The implications of the above process is that discontinuities of height  $\pi$  are allowed to remain in the map which is consistent with the classical definition of the isoclinic angle [Frocht, 1940]; and secondly that zero-crossings occurring because the isoclinic is zero valued are ignored.

Real data, particularly from reflection photoelasticity, tends to contain speckle which causes an enormous number of small areas to be identified by the boundary detection technique and the subsequent analysis of these areas extends the processing time unacceptably. The processing time can be reduced by an order of magnitude if the isoclinic map is filtered prior to applying the boundary detection technique and if the smallest areas are left wrapped. The filtering process employed was a Wiener filter [Press et al, 1992] which is an adaptive filtering technique that employs a small amount of smoothing when large variances and vice versa. The filter was applied to the  $I_s$  and  $I_c$  data prior to their use in equation (1) to generate the isoclinic angle.

100 The speckle generated by the diffuse nature of the reflective layer in a photoelastic coating can cause many very small areas to be identified by the zero-crossing boundary detection technique. However, these small areas were found to not significantly effect the form of the isochromatic map of when a quality-guided unwrapping routine was employed although their influence could be very severe when  
 105 a raster or flood-fill routine is used. Hence it is possible to leave the smallest areas unprocessed and achieve a significant reduction in the processing time for demodulation. Typically, between 5% and 10% of the area in the map of isoclinic angle is left unprocessed. In the program written to implement the new approach the size of the window ( $w_2 \times w_2$ ) for the filter and the percentage ( $\rho$ ) of area left wrapped  
 110 were linearly related as:

$$\rho = \frac{10}{(w_2)_{\max}} w_2 \quad (6)$$

where  $w_{\max}$  is dependent on the size of the data array such that

$$(w_2)_{\max} = \frac{20MN}{256^2} \quad (7)$$

where  $N$  and  $M$  are the number of pixels in the  $x$  and  $y$  direction of the data map and  
 115  $w_2$  can be set by the operator to be  $0 < w_2 \leq (w_2)_{\max}$  but has a default value of  $0.5(w_2)_{\max}$ . The selection of  $w_2$  allows the algorithm to be adjusted based on experienced to optimise its performance with data sets from different experimental environments such as transmission and reflection photoelastic arrangements. Another important feature of the implementation of the above process is that the filtered  
 120 isoclinic data is only used to generate the zero-crossing boundaries and the unfiltered isoclinic data is unwrapped so that any errors introduced by the filtering process are not propagated into the maps of stress direction and magnitude. For the data shown in figure 1c the window size,  $w_2$  was 20 pixels and the demodulation process results in the isoclinic map shown in figure 1d.

125 The demodulated isoclinic angle is substituted into expression (4) to obtain a map of the relative retardation which is also demodulated from the influence of the isoclinic angle (figure 1e). A quality-guided approach [Pritt, 1996] is used to unwrap the relative retardation using a phase derivative variance quality map [Ghigli & Pritt, 1998] which is similar to the one used in the demodulation process described above

130 except that the partial derivatives of the relative retardation are employed in equation  
(5) and the threshold height,  $h$  for the spike removal is  $2\pi$ .

In this case an unwrapping algorithm prepared for the fringe projection technique  
[Heredia & Patterson, 2003] has been used because it was readily available and  
already been optimised for speed of processing. Processing times were found to be  
135 similar to those required by the Wang & Patterson [1995] algorithm. At the end of the  
unwrapping process the results are translated along the fringe order axis so that the  
minimum fringe order is assigned as zero. If a zero order fringe is present in the data  
field, then no calibration of the unwrapped map will be necessary, otherwise the map  
of isochromatic fringe order can be calibrated by providing a known fringe order at  
140 any point in the data field. The association with either the maximum or minimum  
principal stress direction of the seed or initial area  $A(i)$  in the demodulation process  
controls whether the isochromatic map that results from the unwrapping process is  
upright or requires inversion. In the latter case, the isochromatic fringe order  
distribution is reflected about its median value. A typical result is shown in figure 1f.

145

### References

- Frocht, M.M., 1940, *Photoelasticity, vol.1*, John Wiley & Sons Inc., New York, 1940.
- Ghiglia, D.C., Pritt, M.D., 1998, *Two-dimensional phase unwrapping*; Wiley-Interscience, New York .
- 150 Heredia-Ortiz, M., Patterson, E.A., 2003, On the industrial applications of moiré and fringe projection techniques, *Strain*, 39:95-100.
- Patterson, E.A., and Wang, Z.F., 1991. Towards full-field automated photoelastic analysis of complex components. *Strain*, 27:49-56.
- Patterson, E.A., & Taroni, M., 2003, High frequency quantitative photoelasticity applied to jet engine components, *Proc SEM Ann Conf. Experimental & Applied Mechanics*, Charlotte, NC. paper no. 5, also submitted for publication in *Experimental Mechanics*.
- 155
- Press, W.H.; Flannery, B.P.; Teukolsky, S.A.; and Vetterling, W.T., 1992, Optimal (Wiener) Filtering with the FFT. §13.3 in *Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed.* Cambridge, England: Cambridge University Press, pp.539-542.
- 160
- Pritt, M.D., 1996, Phase unwrapping by means of multi-grid techniques for interferometric SAR. *IEEE Trans. Geo. & Remote Sensing*, 34(3):728-738.
- Quiroga, J.A., and Gonzalez-Cano, A., 1997, Phase measuring algorithm from extraction of isochromatics of photoelastic fringe patterns, *Appl. Optics* 36(32):8397-8402.
- 165

Ramesh, K., and Ganapathy, V., 1996. Phase-shifting methodologies in photoelastic analysis – application of Jones calculus. *J. Strain Analysis* 31:423-432.

170 Sparling, S., Gagnon, J., Komorowski, J., *Photoelastic analysis by 6-image phase-stepping with RGB input*. NRC Internal report LTR-St-2154 4, NRC Ottawa

Wang, Z.F., and Patterson, E.A., 1995, Use of phase-stepping with demodulation and fuzzy sets for birefringence measurement. *Optics and Lasers in Engineering*, 22:91-104.

175

**Table 1**– Six-step schemes for digital transmission photoelasticity [Patterson & Wang, 1991]

180

Image No	$\phi$	$\beta$	Light intensity
$i_1$	0	$\pi/4$	$i_1 = i_m + i_v \cos \alpha$
$i_2$	0	$-\pi/4$	$i_2 = i_m - i_v \cos \alpha$
$i_3$	0	0	$i_3 = i_m - i_v \sin \alpha \sin 2\theta$
$i_4$	$\pi/4$	$\pi/4$	$i_4 = i_m + i_v \cos \alpha \sin 2\theta$
$i_5$	$\pi/2$	$\pi/2$	$i_5 = i_m + i_v \sin \alpha \sin 2\theta$
$i_6$	$3\pi/4$	$3\pi/4$	$i_6 = i_m - i_v \cos \alpha \sin 2\theta$

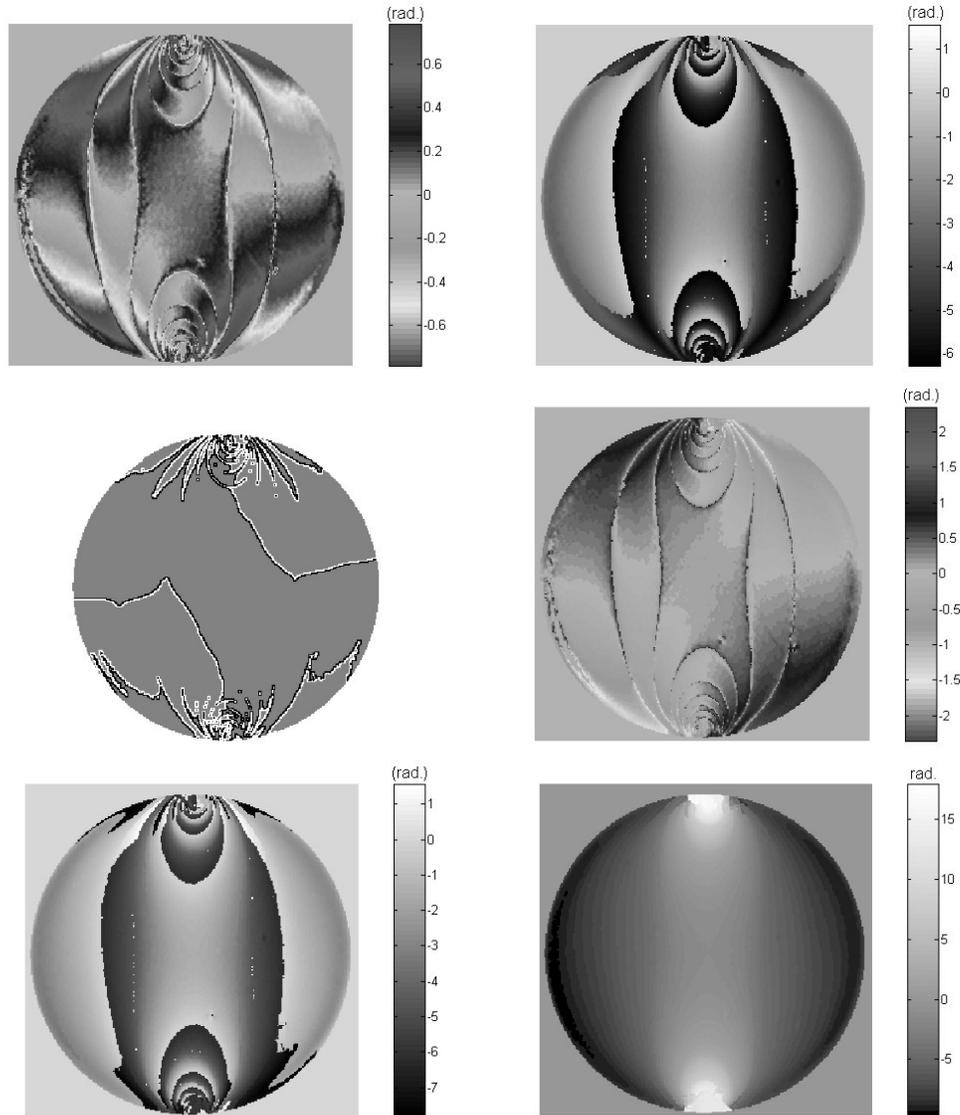


Figure 1: Simulation results from a disc loaded in diametral compression (*from top by rows*):

- 185 (a) Map of the isoclinic angle obtained using expression (1) from six images based on those described in Table 1. The influence of half and integer fringe orders can be seen as discontinuities and the isoclinic angle switches allegiance between the principal stresses across these discontinuities. The map is wrapped with periods of  $\pi/2$ .
- (b) Map of the relative retardation obtained using expression (3) from six images as in (a). The influence of the isoclinic angle can be seen at locations with an isoclinic angle of 45 degrees. The map is wrapped with periods of  $\pi$  with period ends at  $0, \pi, 2\pi\dots$
- 190 (c) Zero-crossing boundaries. Areas within the boundaries are associated with a single principal stress. Positive areas are enclosed by white lines and negative ones by black lines. Positive areas are protrusions surrounded by a step down to their neighbours whilst negative areas are indents surrounded by a step up to their neighbours. The window size,  $w_2$  used to compute this map was 20 pixels.
- 195 (d) Demodulated map of isoclinic angle generated using the zero-crossing boundaries in (c) to unwrap the data in (a). Some noise remains at the locations of half and integer fringe orders. The periodicity of the data is  $\pi$  corresponding to the classical definition of isoclinic angle, and it is everywhere associated with the same principal stress.
- (e) Demodulated map of relative retardation produced using the data in (b) and (d).
- (f) Unwrapped map of relative retardation generated from the data in (e) using the quality guided algorithm.