Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
ISPRS Journal of Photogrammetry & Remote Sensing 60 (2006) 269–283
PHOTOGRAMMETRY & REMOTE SENSING
www.elsevier.com/locate/isprsjprs

Measuring the distance of vegetation from powerlines using stereo vision

Changming Sun [a,*], Ronald Jones [a,1], Hugues Talbot [a,2], Xiaoliang Wu [b], Kevin Cheong [a,3], Richard Beare [a,4], Michael Buckley [a], Mark Berman [a]

[a] *CSIRO Mathematical and Information Sciences, Locked Bag 17, North Ryde, NSW 1670, Australia*
[b] *CSIRO Mathematical and Information Sciences, Private Bag No 5, Wembley WA 6913, Australia*

Received 10 September 2005; received in revised form 17 March 2006; accepted 17 March 2006
Available online 11 May 2006

Abstract

Electricity distribution companies in many countries are required to maintain a regulated clearance space around all powerlines for bushfire mitigation and safety purposes. Vegetation encroachment of high voltage electricity line clearance space is a major problem for electricity distribution utilities. If not properly controlled, vegetation encroachment can lead to bushfire and public safety risks as well as degrading electricity supply reliability. In this paper we describe a prototype airborne system for the automated measurement of the distance of vegetation from powerlines using stereo vision from a stream of stereo images. A fundamental problem with the images from the prototype system is that the powerlines are usually difficult to see, although the power poles are visible. The proposed strategy has been to recover the vegetation surface using stereo vision techniques, identify successive power poles, model the powerlines between successive poles as a catenary, and measure the distance between the vegetation surface and the modelled line. Some suggestions about how to improve the system are also made.

Keywords: stereo matching; powerline inspection; power pole segmentation; vegetation clearance; 3D vegetation surface

1. Introduction

Infrastructure components such as powerlines, telecommunication lines, and oil and gas pipelines, are often run above ground, for very long distances, extending in "corridors". Encroachment of vegetation on the infrastructure components (whether they be powerlines, telecommunication lines or pipelines) is undesirable. Vegetation may damage powerlines or telecommunication lines, or limit access to pipelines. However one of the most disastrous outcomes is bushfire, resulting from contact between powerlines and vegetation. Inspection of infrastructure corridors is therefore required to monitor vegetation.

Current methods for checking clearances involve laborious, ground or airborne based visual inspection of electricity distribution networks to determine which trees must be cleared, together with extensive aerial

* Corresponding author. Fax: +61 2 9325 3200.
E-mail address: changming.sun@csiro.au (C. Sun).
[1] Present address: DSTO Information Sciences Laboratory, Edinburgh SA 5111, Australia.
[2] Present address: Laboratoire Algorithmique et Architecture des Systèmes Informatiques, ESIEE, Paris, France.
[3] Present address: Bureau of Meteorology, Melbourne Vic 3001, Australia.
[4] Present address: Medicine, Nursing and Health Sciences, Monash University, Vic 3800 Australia.

0924-2716/$ - see front matter © 2006 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2006.03.004

audits of the network to ensure effective clearance prior to the bushfire season. These are expensive, time consuming, sometimes dangerous, and subject to observer bias and fatigue or a failure to observe trouble spots at the right time. Further, human beings are not very good at judging perspective from a distance. It may therefore be difficult to tell from a supervised inspection whether vegetation is encroaching on infrastructure components or not. For example, when a powerline is being inspected from an aircraft, the inspector is usually required to observe the line from directly above. In this situation, it is very difficult to judge the distance between the vegetation and the lines.

There is a need for a system and method which enables the automated mapping or inspection of corridor-type infrastructure. Such a system should quantify encroachment by vegetation and detect other faults, thus enabling correct maintenance planning. A numerical method for predicting the performance of a powerline proximity warning device installed on a vehicle equipped with any type of boom has been proposed (Nguyen et al., 1996). Remote sensing techniques have been used for land-cover anomaly detection along pipeline rights-of-way and for transmission corridor encroachment detection (Gauthier et al., 2001; Campanella et al., 1995). The image resolution for currently available remotely sensed images is still low. A photogrammetric solution, called Danger Tree, has been developed using aerial photography (Dall, 1991). It requires a number of parameters about the transmission lines. It also requires aerial photographs on two occasions: once when the trees have no leaves, and once when the tree leaves are at their fullest. Jones and colleagues have developed techniques for aerial inspection of overhead powerlines using video on a helicopter platform (Jones, 2000; Jones and Earp, 2002; Whitworth et al., 2001; Williams et al., 2001; Jones et al., 2003; Golightly and Jones, 2003, 2005). However, they have not been used for measuring the distance between powerlines and nearby vegetation. Laser range finding techniques have sometimes been used for such purposes (Ackermann, 1999; Carter and Shrestha, 1998; Hyde et al., 1996; FLI-MAP, 2005). Systems which use laser scanning techniques tend to be more expensive than systems using video cameras. Even with laser range finding techniques, video images are often needed for visual interpretation purposes. In this project, we study the possibility of developing a cost effective airborne image capture and processing system to automatically measure the clearance of vegetation from powerlines to support bushfire mitigation operations.

The paper is organised as follows. Section 2 gives a brief description of the prototype system and our approach. Section 3 describes our methods for stereo image processing, including camera orientation estimation and stereo image normalisation, dense stereo matching, and 3D calculation. We present our methods for power pole segmentation using the normalised stereo images in Section 4. The digital elevation model (DEM) and orthoimage generation and mosaicking are presented in Section 5. The distance calculation, computational speed and accuracy, and visualisation aspects are presented in Sections 6, 7 and 8 respectively. Finally in Section 9 we make some suggestions about how the system might be improved.

2. The prototype system and our approach

In this prototype system, a fixed-wing aircraft (Cessna 182RG) flies at approximately 80 m above the ground, with forward looking video cameras that capture a continuous series of color images of the powerline infrastructure mounted beneath the tip of each wing. (The flying height for a fixed-wing aircraft is of necessity much higher than that for a helicopter, which is usually between 3 and 10 m above the powerlines.) The angle between the optical axes of the two cameras is about 7°. The cameras are looking down at about 35° (i.e. the angle between the optical axis and the horizon is about 35°). The pixel resolution on the ground is about 10 cm close to the centre of the image. The distance between the two cameras is 10 m. The system carries a differential global positioning system (DGPS) and an inertial navigation system (INS) to provide position and orientation data for the aircraft. The accuracy of the DGPS is better than 0.5 m. Fig. 1 gives a schematic illustration of the system.

A fundamental problem with the images currently available is that the powerlines are usually difficult to see, although power poles are visible. The proposed strategy is to identify successive power poles, and to model the lines between poles as a catenary. In fact, because of variations in line tension with load, temperature, etc., a range of catenary curves will be used, resulting in a catenary "envelope". The aim is then to measure the distance from the envelope to nearby vegetation. Another complication is that successive power poles are in different image "frames", usually about 10 to 14 images apart. Frames between successive power poles need to be "mosaicked" into a single pole-to-pole image in order to measure the distance between

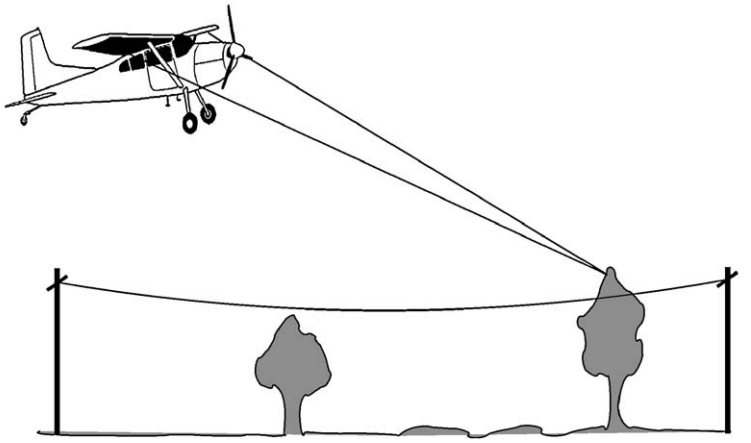

Fig. 1. A fixed-wing aircraft flying above powerlines and collecting streams of stereo images.

the catenary envelope and the trees. Therefore the strategy used in this project is:

(1) Use stereo vision to obtain a 3D surface from video images.
(2) Find the power poles automatically from images and match them in stereo images.
(3) Mosaic the images to create a pole-to-pole image.
(4) Infer the catenary envelope in each pole-to-pole image.
(5) Calculate the distances from the catenary envelope to trees.

The following steps will need to be carried out to produce the relevant information in a pole-to-pole image and they are described in detail later in this paper:

(1) Automatically calculate the relative orientation of the stereo cameras for each pair of images, since wind and vibration effects continually change the camera geometry.
(2) Normalise individual image pairs to generate epipolar stereo images based on the relative camera orientation.
(3) Match individual image pairs to obtain disparity maps.
(4) Identify power poles using image information.
(5) Compute 3D information from matched pairs. This step includes the calculation of the 3D positions of the tops of the power poles and the 3D surface of each 3D scene.
(6) Calculate the catenary envelope of powerlines using the 3D positions of the tops of the poles within a span.
(7) Generate a DEM using the disparity map and camera parameters; generate an orthoimage using the DEM, image and camera information.
(8) Mosaic the DEM and orthoimages into a 3D pole-to-pole image.
(9) Identify tree boundaries in each 3D pole-to-pole image.
(10) Measure appropriate distances from the catenary envelope to tree boundaries.

3. 3D reconstruction using stereo vision

3.1. Input camera parameters and stereo images

There are 6 external parameters recorded that describe the position and orientation of each camera when each image was acquired. These are related to the DGPS and INS measurements. The internal camera parameters are also known through camera calibration. These parameters, identical for each image, are: camera focal length, camera lens distortion, pixel size of the CCD array in the X direction, pixel size of the CCD array in the Y direction, and the location of the principal point. Both the internal and the external parameters will be used in 3D calculations. The color images were initially recorded on SuperVHS tapes, and digitised off-line. These VHS tapes have been used before this study for visual checking of the power poles. The image size is 736×560 pixels. In future systems, digital cameras will be used for image acquisition. For each pole-to-pole span, there are usually 10 to 14 frames. They are oblique views of the scenes. In this project, only lightness information has been used to match image points. The lightness image is obtained by averaging the red, green and blue bands of the original color image. Color

information was used in the power pole segmentation stage.

3.2. Camera orientation and image normalisation

The relative camera parameters do not stay constant due to vibration of the aircraft wings, so it is necessary to estimate the true camera orientation parameters automatically from image features. We used Zhang et al.'s method for robust feature matching (Zhang et al., 1995). Because the point matching step is automatic, mismatches can occur. Therefore it is necessary to identify and remove those mismatches for the estimation of the camera geometries. The parameters related to this geometry include the three rotation angles and the three elements of the translation vector (two elements if normalised). The geometry obtained will be used for stereo image normalisation and for 3D calculation.

After the relative orientation parameters of the cameras have been obtained, stereo image normalisation can be performed. The operation of *normalisation* is meant to ensure a simple epipolar geometry for a stereo pair (Faugeras, 1993, Section 6.3). By *simple geometry* we mean that the epipolar lines are parallel to the image rows, i.e. the normalisation process transforms (usually non-horizontal) epipolar lines into horizontal scanlines. The new image plane is chosen so that it is parallel to the line connecting the two camera centres (Ayache and Hansen, 1988; Ayache and Lustman, 1991; Kang et al., 1994). Normalisation using three images (a stereo pair plus one image in a neighboring pair) can also be carried out (Sun, 2003). In the normalised stereo images, corresponding points or matching points lie on the same horizontal scanlines in the left and right epipolar images. This allows for simpler and more efficient dense stereo matching. Stereo image normalisation will also be useful for power pole segmentation and matching.

3.3. Dense stereo matching

Disparity is the difference in pixel location between matched features in different images. This disparity image together with the camera parameters can be used to calculate the 3D surface. We use the normalised left and right images to generate a disparity map using fast dense stereo matching.

We have developed a fast and reliable stereo matching algorithm which produces a dense disparity map by using fast cross correlation, rectangular subregioning (RSR) and 3D maximum-surface techniques in a coarse-to-fine (pyramid) scheme (Sun, 1997, 2002). Fast correlation is achieved by using the box-filtering technique whose speed is independent of the size of the correlation window and by segmenting the stereo images into rectangular subimages at different levels of the pyramid. By working with rectangular subimages, not only can the speed of the correlation be further increased, the intermediate memory storage requirement can also be reduced. The disparity map for the stereo images is found in the 3D correlation coefficient volume by obtaining the global 3D maximum-surface rather than simply choosing the position that gives the local maximum correlation coefficient

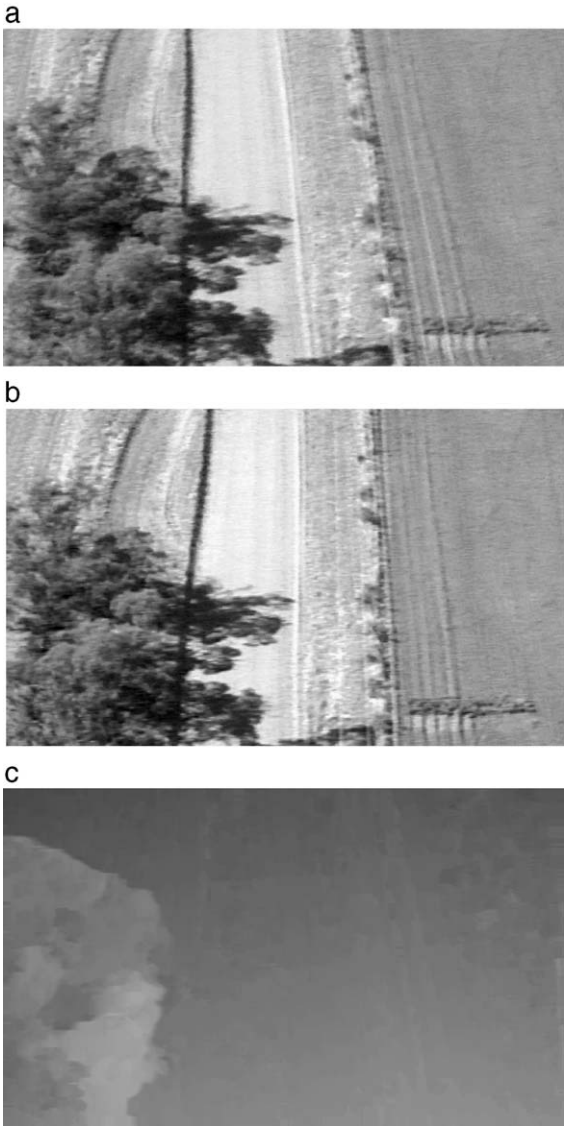

Fig. 2. (a,b) One pair of the normalised stereo images; (c) the disparity image obtained (notice the shadow region of the tree on the ground has the same disparity as the rest of the ground).

value for each pixel. The 3D maximum-surface is obtained using a two-stage dynamic programming (TSDP) technique. Detailed descriptions of the algorithms are given in Sun (1997, 2002). Fig. 2 shows a normalised image and the disparity image obtained by applying our stereo matching technique.

3.4. 3D calculation

After obtaining the corresponding points (as in a disparity map) between the left and right 2D images, the 3D position of the same physical point can be obtained. The coordinates of the 3D point can be obtained by finding the intersection of the two light rays passing through the camera centres and the two image matching points. If we calculate the 3D position for every matched point on the image, a 3D surface of the scene can be obtained. The 3D positions of the tops of the power poles are calculated in the same way. The image positions of the power poles are obtained using the method described in Section 4. The final 3D information is calculated in the global reference system, and therefore can be related to other geographical data such as that found in a geographical information system (GIS).

4. Power pole segmentation

The aim of the power pole segmentation procedure is to find the pole when it appears in an image pair and then to find the position along the pole where the powerline is attached. Note that there is insufficient resolution in the current set of images to locate the pole insulators (which would indicate exactly where the line is attached to the pole). The width of a pole at the top is about 30 cm. Instead, our approach is to find either the intersection point of the cross arm with the pole or, in cases where there is no cross arm attached to the pole, to find an accurate measure of the top of the pole. Input to the segmentation procedure is a pair of left and right normalised color input stereo images as described in the previous section.

4.1. Finding candidates for poles and cross arms

The procedure for segmenting candidates for poles and cross arms begins with a preprocessing step to mark 'background' parts of the image that we know cannot possibly be poles, in particular green vegetation regions. There are two assumptions made about poles during this processing stage: poles are gray (not green or some other color) and poles are light (they are not black and there is adequate lighting on the pole). These two assumptions are the basis for a simple segmentation procedure to find background regions in the image. After transforming the color input image into an hue, lightness and saturation image, the saturation component of this image is then used to identify regions that are not gray and the lightness component is used to find dark regions.

The next step in the segmentation process is to identify candidates for poles in the input image (as illustrated by a subimage in Fig. 3a). The sequence of steps is as follows:

(1) Perform a linear vertical median filter on the lightness component of the hue, lightness and saturation image (Fig. 3b).
(2) Perform a horizontal linear median filter on the vertical median filter result and take the absolute difference between the two images.
(3) Threshold this result and remove those pixels previously identified as background pixels (Fig. 3c).
(4) Perform a vertical linear median filter on this result to connect up broken line elements and remove isolated noise.
(5) Perform a dynamic line opening (see below for a description of this new filter) to remove spurious features on the sides of linear objects.
(6) Remove short lines to produce the final image of pole candidates (Fig. 3d).

As mentioned above, a new filter called the *dynamic line opening* is used to clean up linear features in binary images. The filter treats each object in the binary image independently. The direction θ_i of the major axis of the best fit ellipse (Haralick and Shapiro, 1992, Sect. A.7) to the binary object i is calculated and the maximum length L_i of the object in direction θ_i is determined (using a minimum bounding box in direction θ_i). Then each object is filtered by removing line segments (in direction θ_i) within the object that are shorter than some fraction of L_i; in our procedure, this fraction is one half. The filter tends to remove noise from the side of the object and preserve only the main form of the object.

The process for finding cross arms follows that for finding pole candidates. As the aircraft is flying along the powerlines, the cross arms usually appear approximately horizontally in the image. Therefore, the direction of filters used for cross arm detection are along the horizontal direction. The number of cross arms on a pole varies. It can be 1, 2, 3 or no cross arms. When there is more than one cross arm, the separation between them varies from about 90 to about 210 cm depending on the type of pole. An example result for

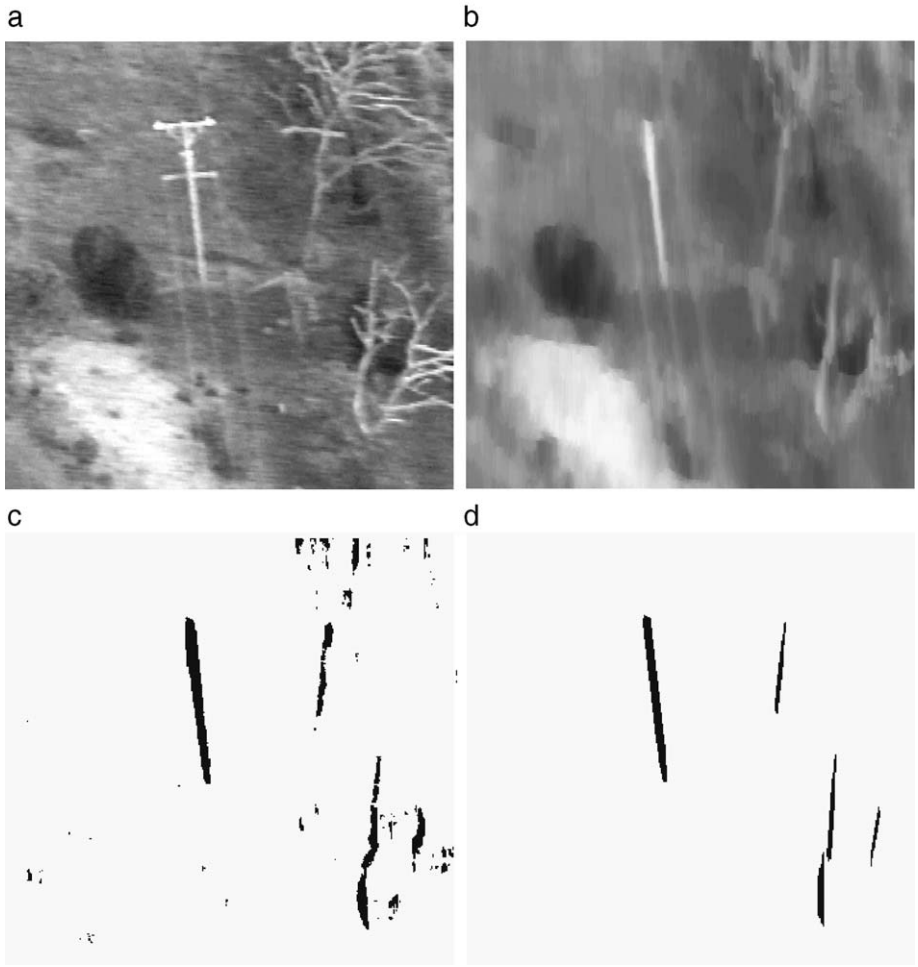

Fig. 3. Finding candidates for poles in the input image. (a) Subset of an input image. (b) Result from the vertical median filter when applied to the lightness component image. (c) Threshold of the absolute difference image between (b) and the result of a horizontal median filter, restricted to regions that are not in the background of the image. (d) Final result after a dynamic line opening and length threshold is applied.

this process is shown in Fig. 4. A few steps for the pole or cross arm detection involve the use of median filtering along linear windows. Other methods such as the Sobel filter or the Hough Transform could also be used for pole or cross arm detection. However, we have found the above procedure to be satisfactory for our purposes.

4.2. Matching candidates for poles and cross arms

After running the background detection, and the pole and cross arm segmentation procedures on both the left and right normalised images, we have images of pole and cross arm candidates for the image pair. Fig. 5 shows an example set of results. The input left and right images are shown in Fig. 5a and b respectively. The corresponding pole candidates found are shown in Fig. 5c and d and the cross arm candidates found are shown in Fig. 5e and f.

4.2.1. Matching pole candidates

The first step in the pole matching process is to label and obtain summary shape statistics (such as area, bounding box and best fit ellipse statistics) for the pole candidates in the left and right images. Any candidates that have a small vertical height or are not lying within the central region of the image are discarded. We say that a left and a right candidate form a valid pair of poles if they satisfy the following constraints:

(1) There is vertical overlap of the two candidates.
(2) The angular separation between the candidates is within a specified angle range. Note that a pair of pole candidates will typically form a ‘V’ shape,

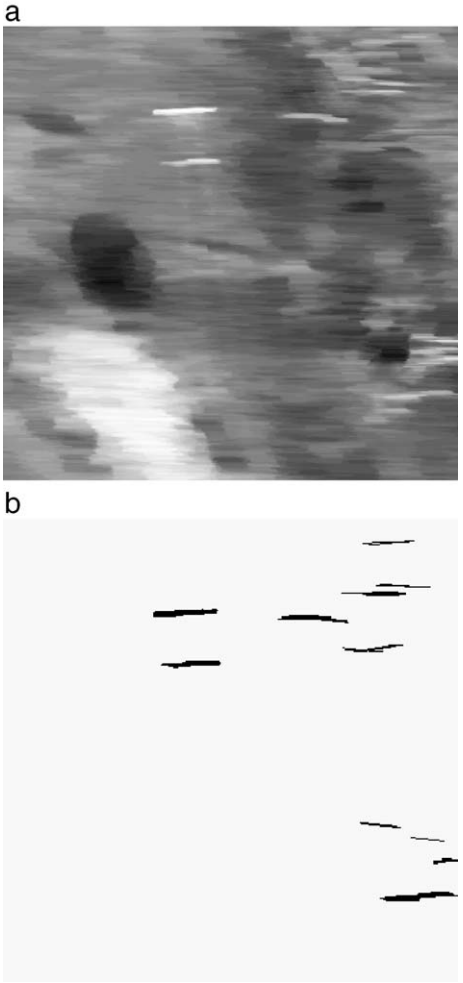

Fig. 4. Finding candidates for cross arms in the input image (as in Fig. 3a). (a) Result from the horizontal median filter when applied to the lightness component image. (b) Final result after cleaning and applying an opening by union.

where the right pole candidate is actually positioned to the left hand side of the left pole candidate.

(3) The horizontal distance between the centroid of the right and the left pole candidates is within a specified range.

4.2.2. Matching cross arm candidates

Once a set of valid pairs of poles has been found, the next stage in the matching process is to find and match any cross arms that intersect the pole candidates. The process comprises the following steps for each valid pair of poles:

(1) Fit a line to the left pole candidate in the valid pair that passes through its centroid at an angle given by the direction of the major axis of the best fit ellipse to the pole candidate.
(2) Find all the cross arm candidates in the left image that intersect this line, between two vertical extremes given by the bottom of the pole candidate and the top of the pole candidate plus a few extra pixels.
(3) Record the vertical position of each intersecting cross arm at the centroid point where it intersects the line.
(4) Repeat the above process for the pole candidates in the right image to find the right intersecting cross arms.
(5) Match a pair of left and right intersecting cross arms by finding the pair that has the minimum vertical separation.
(6) Repeat the above step, excluding cross arms that have already been matched, until either there are no remaining left cross arms or no remaining right cross arms.

The matching cross arms for the example valid pair in Fig. 5 are shown in Fig. 6. Here, the vertical separation marks where to search for intersecting cross arms along the line fitted to the pole candidate; in this case there are two intersecting cross arms for both the left and right candidates.

4.2.3. Choosing the best valid pair of pole candidates

We now have matched valid pairs of pole candidates and have found any matching cross arms for each valid pair. The final stage in the process is to select the best valid pair of pole candidates, which will hopefully correspond to the true power pole in the input images. If we consider two valid pairs of poles A and B, we consider that A is a better solution than B if it satisfies one of the following conditions:

(1) A has the same number of cross arms as B and has a greater vertical height than B.
(2) A has more cross arms than B and has at least half the vertical height of B.
(3) A has fewer cross arms than B and has at least twice the vertical height of B.

By comparing all the valid pairs, the best solution for a valid pair is determined. Fig. 7 shows a collection of results superimposed upon the original normalised input images. Segmented poles are indicated in black and cross bar intersection points are indicated by white discs.

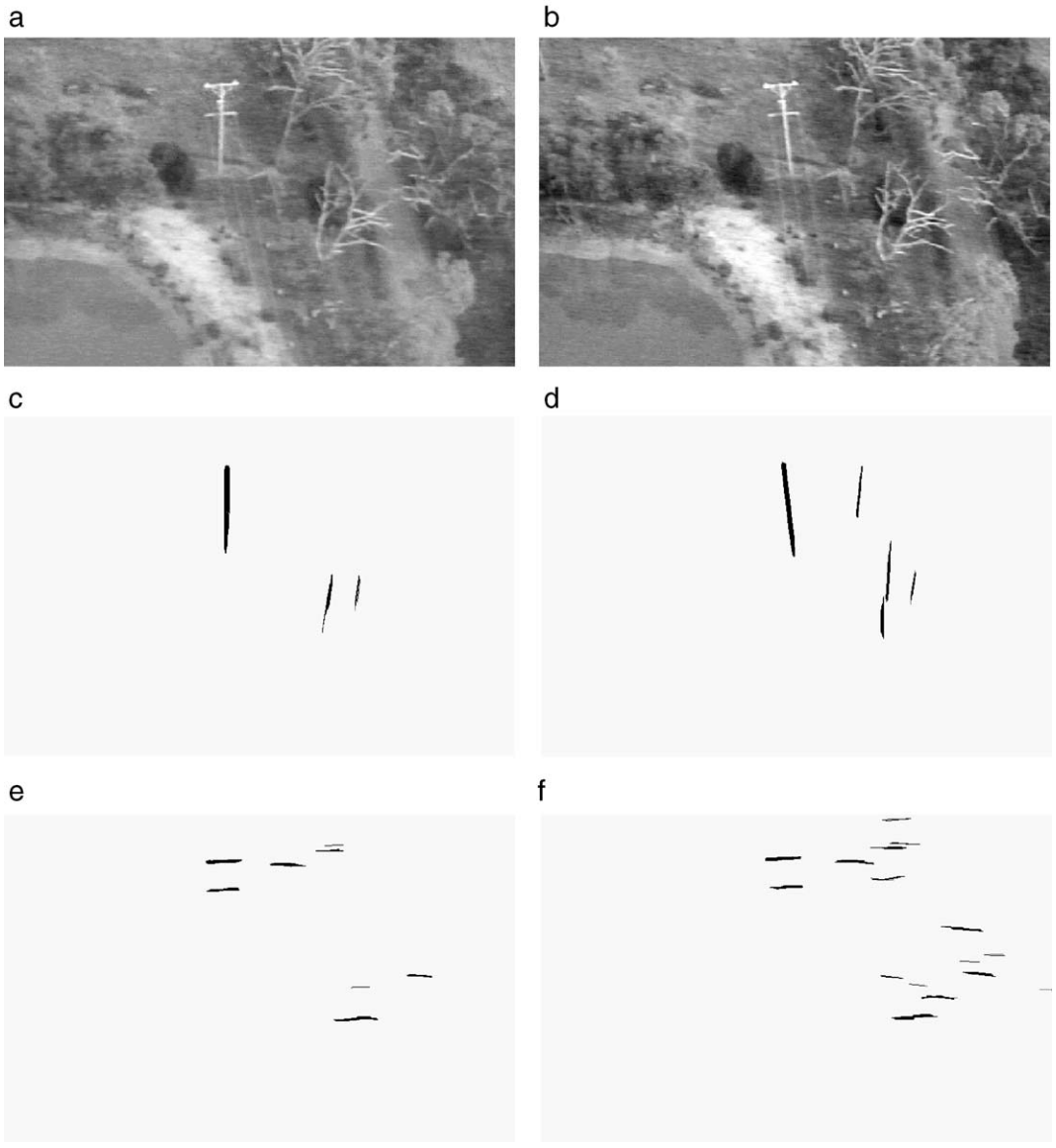

Fig. 5. Results for the pole and cross arm segmentation procedures. (a) Left normalised input image. (b) Right normalised input image. (c) Pole candidates in the left image. (d) Pole candidates in the right image. (e) Cross arm candidates in the left image. (f) Cross arm candidates in the right image.

4.3. Poles with no cross arms

In cases where there is no cross arm attached to the power pole, we need to use the position of the top of the pole as a measure for where the powerline is attached. Therefore, we refine the result given by the pole segmentation procedure using a watershed function (Vincent and Soille, 1991). The watershed function requires as input a marker for the pole, a marker for the background (called the external marker), and also a segmentation function. The first stage in the process is to clean and refine the previous segmentation result for the pole so as to obtain a good pole marker. This is done by the use of the skeleton operation (Serra, 1982) and the rank-max opening operation.

We now turn our attention to finding an external marker for the watershed function. A rank-min closing is used to fill in regions in the external marker that are not linear features such as poles. We also make use of the pole marker found in the previous section to make sure that the external marker does not overlap the pole. These markers are shown in Fig. 8a, where the pole

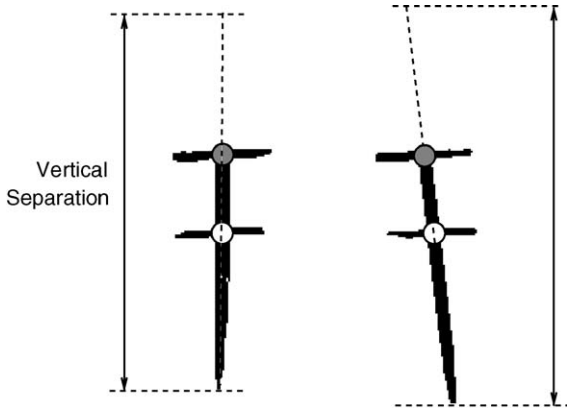

Fig. 6. Matching cross arms to valid pairs of pole candidates.

marker is in white and the external marker in gray. The segmentation function we use is the image gradient that was derived during the process of finding a pole marker. The result of the watershed function is shown in Fig. 8b, where the final segmentation of the pole is shown in white. A close up of the pole is shown in Fig. 8d. For comparison, the previous segmentation result for the pole, obtained at the end of Section 4.2, is shown in Fig. 8c. It is clear from this that the new segmentation result for the pole gives a better estimate of the top of the pole.

It is necessary to extract the coordinates of the top of the pole so that they may be used to find the position along the pole where the powerline is attached. This is done by simply fitting a line to the segmented pole using

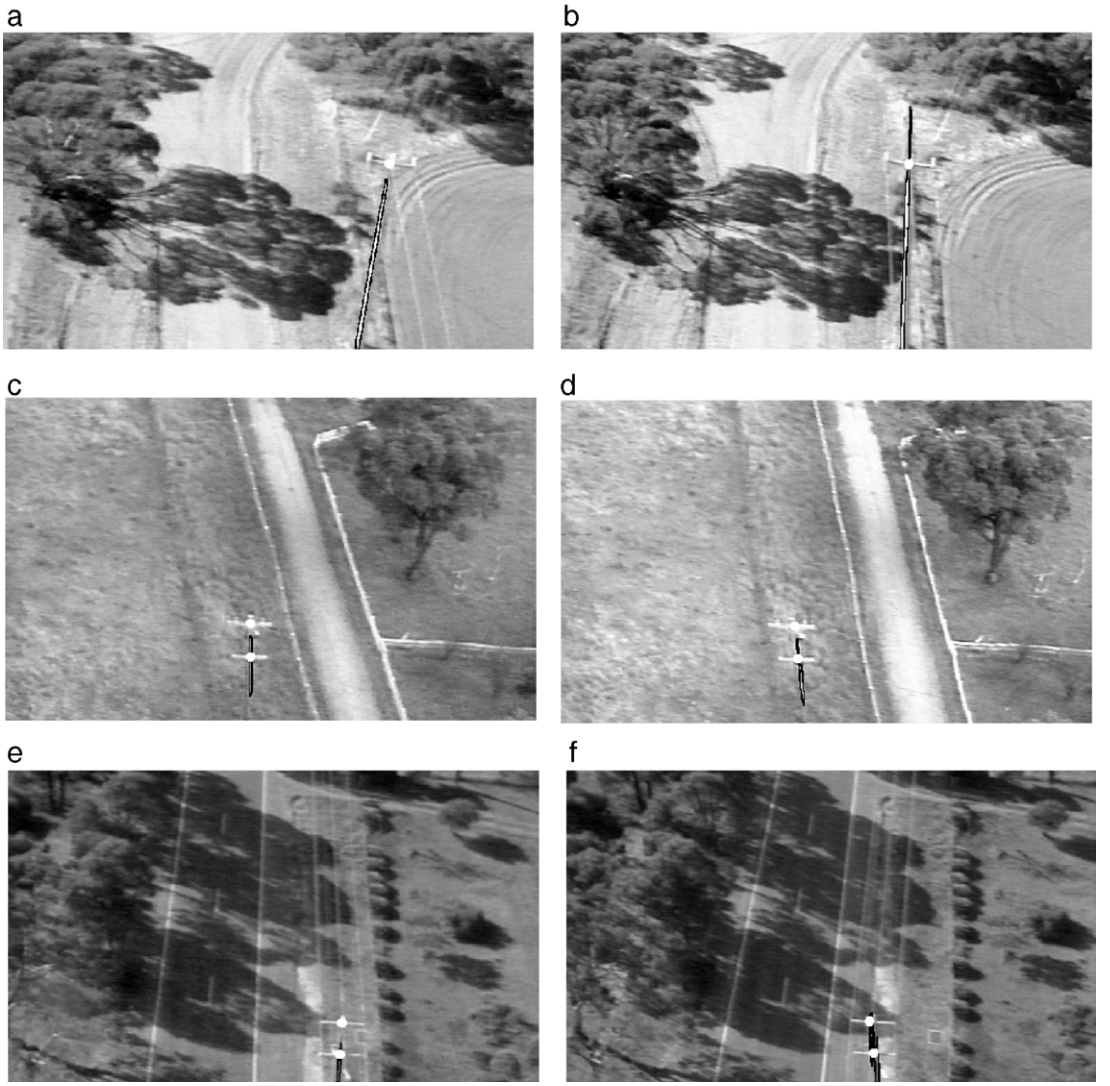

Fig. 7. Various results from the segmentation procedure. Segmented poles are indicated in black and cross bar intersection points are indicated by white discs.

the best-fit ellipse statistics of centroid and orientation. By inserting the top vertical position into the equation for this line, the corresponding top horizontal position is readily extracted.

4.4. Summary results for pole segmentation

We now present a summary of the segmentation results for several representative spans of images that were normalised. Usually a span of images contains 10 to 14 image pairs that runs from one pole to the next. A pole will often appear in about four or five of these image pairs. Our algorithm will usually find the pole in two or three of these pairs.

We present results for a set of eight spans. Table 1 shows the results for these spans, each of which is indicated by a number 1 to 8 at the top of the table. Under each number is a column of results which gives the success rate for the span on an image basis and then an overall success (tick) or fail (cross) classification for the span. For example, for the first span, 5 out of 5 of the image pairs that contained a pole were successfully segmented. By 'successfully segmented' we mean that the pole was found and the position along the pole where the powerline is attached was correctly determined (either the pole/cross arm intersection point or the top of the pole for those poles without cross arms). For this same span, 9 out of 9 of the image pairs that did not

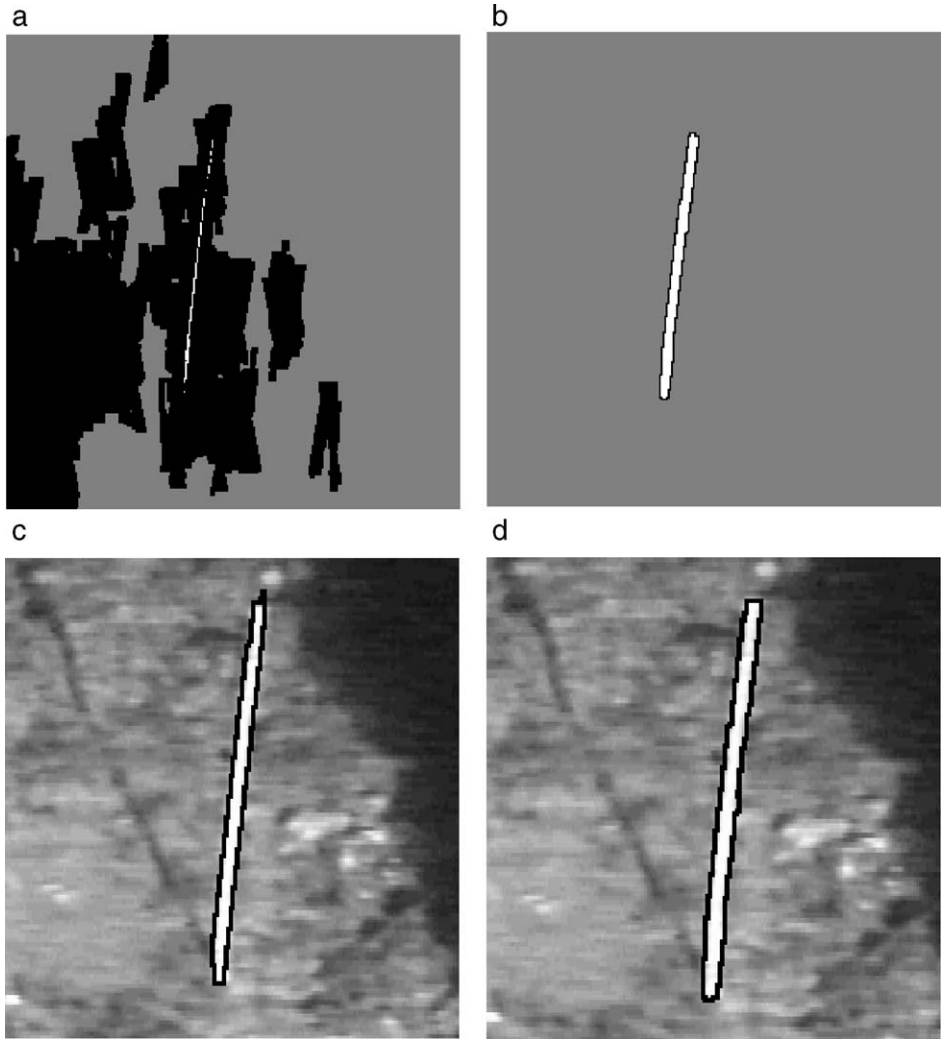

Fig. 8. Applying the watershed function. (a) Combined pole in white and external markers in gray. (b) Result from the watershed function, where the final result for the pole is in white. (c) Original pole candidate superimposed on the image. (d) The improved segmentation result using the watershed process; the top of the segmented pole is now correctly positioned.

Table 1
Summary of results for 8 spans of images

Span	1	2	3	4	5	6	7	8	
Success: pole	5/5	5/5	2/4	4/4	3/6	1/3	6/6	4/4	30/37 (81%)
Non pole	9/9	8/8	8/8	9/9	4/5	11/11	6/6	17/17	72/73 (98%)
Failure: pole			2/4*		3/6*	2/3*			7/37 (19%)
Non pole					1/5				1/73 (2%)
Success	√	√	√	√	×	×	√	√	

Success (tick) or fail (cross). Asterisk denotes failures due to poor image quality. See Section 4.4 for details.

contain poles were correctly segmented. That is, the procedure returned no pole objects for these images and was not fooled into finding an object that was not a pole (for example a branch of a tree or a fence post). Overall then, we have determined that this span was a success and this is indicated by the tick at the bottom of the column.

As any given pole will appear in two or three consecutive images, we only need to segment the pole and correctly determine the position of the powerline in any one of these images. Therefore, it is possible for the segmentation procedure to fail to find a pole in a single image pair but find it in another image pair; this will be recorded as an overall success for the span. We denote failures due to poor image quality in Table 1 with an asterisk (in fact, most of the failures were caused by poor image quality, as discussed further below).

Span 5 in Table 1 shows the only case where there was a failure on an image pair that did not contain a pole. Overall, 6 out of the 8 spans processed were deemed to be successfully segmented. On an image basis, 81% of the image pairs that contained poles were successfully segmented (as indicated on the right hand side of the table) and 98% of the image pairs that did not contain poles were successfully segmented.

5. DEM and orthoimage generation and mosaicking

5.1. DEM and orthoimage generation

Since the powerlines cannot be seen in the majority of the images, their position will be modelled using the knowledge of where the power poles are. This requires consecutive power poles to appear in a single image. Unfortunately, power poles are several images apart. In the first sequence of test images, power poles appear in approximately every 10th to 14th image in the sequence. In some of the more recent test images, particularly the set over a typical rural network, power poles are more unevenly spaced. To create the pole-to-pole images, we need to generate a DEM and orthoimages. In our case the known output image grid will be the Australian Map Grid. The location of each data item is given in easting and northing coordinates which have been interpolated onto a regular grid. This means that information derived from these images can easily be related to other geographical data.

From the 3D calculation stage, each pair of matching points generates a point in 3D space. Within the field of view of the cameras, all the 3D points which have corresponding image matching points have a 3D location measurement. However, the 3D points obtained do not lie on a regular grid. From these irregular 3D points, interpolation was performed in the local region to obtain the 3D points on a regular grid (Bartier and Keller, 1996). The grid spacings or the distances of neighboring points in the X and Y directions need to be given. The spacing of the grid will relate to the size of the DEM.

After the DEM of a local region has been obtained, its corresponding orthorectified image or orthoimage can be generated. Generating an orthoimage needs the DEM, camera parameters (internal and external) and the original image. For each element in the DEM, we know its 3D X, Y, and Z position. Using the known camera parameters, its corresponding image location of a DEM position can be obtained using the collinearity equations (Kraus, 1993, p.278). Therefore we have a corresponding image intensity value related to this DEM position. This obtained image location is unlikely to be at the exact position of a pixel in the input image grid. The image intensity value in the output image is determined by bilinear interpolation. Fig. 9a shows an image of the DEM obtained from Fig. 2a and b. The black color around the borders of the image indicates missing values. The pixel intensity in Fig. 9a relates to the heights of DEM. Fig. 9b shows the obtained orthoimage.

5.2. Mosaicking of DEMs and orthoimages

Once the images have been aligned on a common grid, mosaicking is performed to overlay the images on each other. In the regions where the images overlap, a decision needs to be made about how to determine what

intensity values are transferred to the output image. The options are to place one image on top of (overwrite) the other or to average the intensity values of the two (or more) images. The latter option can produce a visually smoother image by smoothing any overall brightness differences between the images, but if the alignment of the images is not sufficiently accurate, it can also cause spatial blurring of image features in the overlap region. The overwriting option will be used for these examples, since the time lapse between successive images is minimal, and hence illumination differences should be minimal.

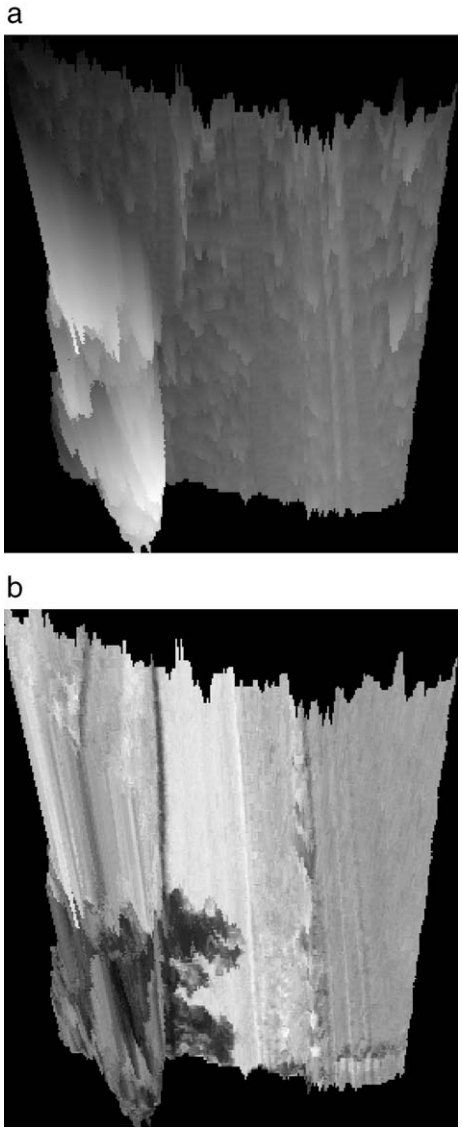

Fig. 9. (a) DEM. The black color around the borders of the image indicates missing values. The pixel intensity relates to the heights of DEM. (b) Orthoimage corresponding to (a).

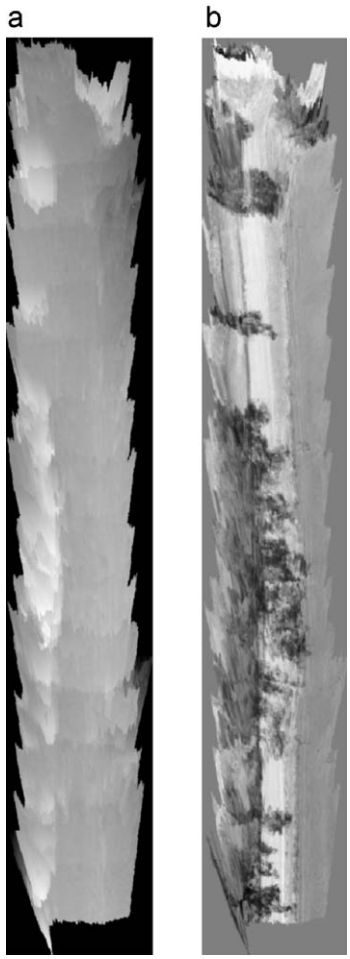

Fig. 10. (a) Mosaicked DEM and the (b) mosaicked orthoimages.

It is also necessary to mosaic several orthoimages into a single large orthoimage. Both DEM and orthoimage information are necessary to mosaic orthoimages geometrically, since the geometric information is provided by the DEM. Fig. 10 shows the mosaicked DEM (from 14 DEMs) and the corresponding mosaicked orthoimage. Due to small errors on the external orientation parameters, the joining for several small blocks of DEMs at the bottom of the mosaicked DEM in Fig. 10 seems unsmooth. This effect can be reduced if a bundle adjustment process is carried out for estimating the external orientation parameters within a span so that the parameters from neighboring pairs are consistent with each other.

6. Measuring the distance of trees from powerlines

This section explains how the distance from the powerline to neighboring trees and vegetation is

calculated. It then describes a method of determining if the vegetation lies within an envelope centred about the powerline. The calculation of distances from the powerline to the underlying surface is performed after the full 3D surface between two power poles has been generated, using stereo matching and mosaicking. This method is illustrated using a span of mosaicked images of landscapes which contains a powerline suspended between two power poles.

The input data is a landscape (3D surface) stored in DEM data format. A pixel in the image has a gray level intensity which corresponds to the height of the landscape at that pixel point. Now, given the 3D location of the tops of two power poles, we can calculate the location of the powerlines modelled by a catenary, the envelope that surrounds the powerlines and decide whether vegetation lies within this envelope.

The minimum safe distance of an object to the powerline is known as a clearance. This measure is given for a specific conductor type and span length of a powerline (the Euclidean distance between two power poles when viewed from the top). Clearances to the powerline nearer the power pole are smaller than clearances along the centre 2/3 span of the powerline. In this paper a clearance distance of 150 cm is used for near spans (1/6 at the beginning and end of the powerline) and 200 cm for centre spans. Each clearance distance represents the radius of a circular envelope in 2D space centred about the powerline when viewed along the powerline, and hence forms a cylindrical envelope along the span of the powerline in 3D space.

A flexible, inelastic powerline suspended between two power poles assumes the shape of a catenary. Its equation is:

$$z = C\left(\cosh\left(\frac{y}{C}\right) - 1\right)$$

where $C = H/W$ is a catenary constant, H is the horizontal component of tension (5000 N (Newton)) and W is the resultant distributed conductor load (0.15 N/m). z denotes the perpendicular height between the powerline and the tangent intersecting the lowest point on the catenary (Y axis).

Fig. 11 shows a sample output of power pole positions, a powerline and catenary envelope, landscape and trees (powerline sagging in this figure has been magnified for illustration purposes). Fig. 11a is a top view of the landscape with the position of the power poles and powerline represented as white pixels. The gray level intensities are proportional to the height of each pixel in the landscape. The catenary envelope is represented by black pixels, and vegetation found within the catenary envelope is represented by white blobs. Fig. 11b is a side view of the landscape, viewed from the right of the first panel. Each pixel point in the landscape image has a height. Again, white pixels represent the powerline modelled by a catenary and white blobs represent vegetation found within the catenary envelope. Fig. 11c is another side view of the landscape, viewed from the right side of the first panel. It is similar to the second panel except that the landscape points plotted along the X axis are bounded by the catenary envelope; gray pixels represent the ground and white blobs represent vegetation within the catenary envelope.

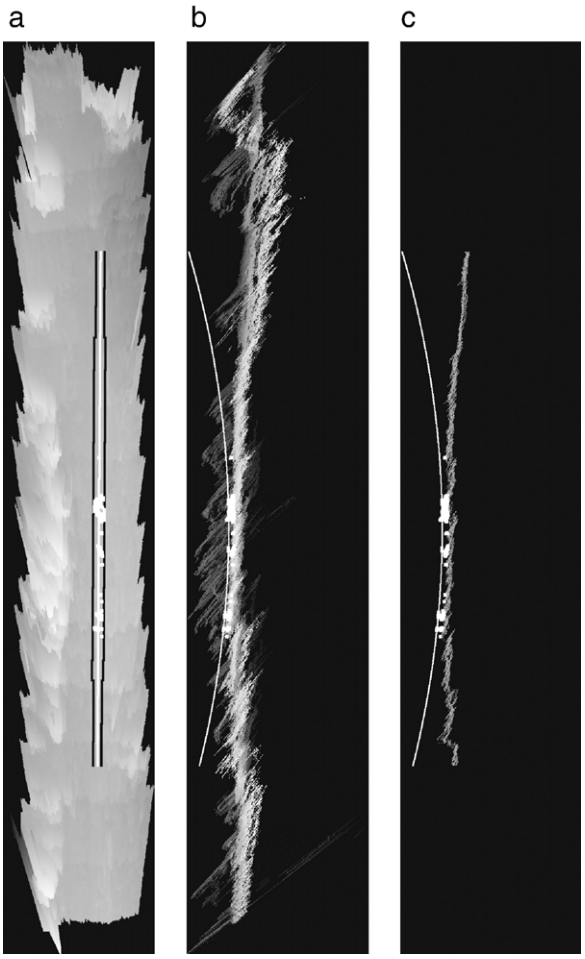

Fig. 11. (a) Top view of a powerline (with envelope) suspended on 2 power poles over a landscape. (b) Side view of a powerline and all objects in the landscape. (c) Side view of a powerline and objects within the catenary envelope (powerline sagging in (b) and (c) has been magnified). Powerline represented as white pixels. The gray level intensities are proportional to the height of each pixel in the landscape. The catenary envelope is represented by black pixels, and vegetation found within the catenary envelope is represented by white blobs.

7. Computational speed and accuracy

In this section we discuss the computational speed and system accuracy issues. Table 2 shows the computational time for each stage of the system on a 2.5 GHz CPU. The pole segmentation step takes the longest time because it is run inside a software environment which does not have efficient I/O procedures. The pole segmentation time can be much reduced if the algorithm is implemented in standalone programs. The timings given for the feature matching, orientation estimation, stereo matching, DEM/orthoimage generation, and pole segmentation stages are for each pair of images. Assuming we have 10 pairs of images within a span, the total processing time for a span will be approximately 170 s. Currently all the processing is carried out off-line.

To estimate the 3D point calculation errors, a total of 18 points have been surveyed at three different locations, 6 at each location. At each location, 4 points are on features on the ground or on building corners, and 2 points are on the top and the bottom of a power pole. The average RMSE error for these 18 points between the surveyed and the calculated X, Y, and Z components was about 23 cm.

8. 3D visualisation

The stereo processing produces a DEM of the terrain. The DEM files are high resolution, dense, height maps (approximately 1000×4000 pixels). The simplest way to display such an object in three dimensions is to break it up into triangles. Unfortunately a simple minded approach produces far too many triangles for typical display hardware to handle in an interactive fashion. Approximations that dramatically reduce the number of triangles while only introducing small errors are essential. A free software package called *terra* has been used to do this (http://www-2.cs.cmu.edu/afs/cs/user/garland/www/scape/terra.html). Terra attempts to minimize the number of triangles and the error in the approximation. It employs a greedy insertion algorithm that searches for the point with the highest error and inserts a new vertex there. A Delaunay triangulation is used to update the mesh in an efficient manner.

Table 2
Computational speed for each stage of the system (on a 2.5 GHz CPU)

Processing step:	Timings:
Feature matching	1.33 s/pair
Orientation estimation	1.05 s/pair
Stereo matching	0.92 s/pair
DEM/orthoimage generation	0.48 s/pair
Pole segmentation	12.79 s/pair
DEM mosaicking	0.54 s/span
Orthoimage mosaicking	0.07 s/span
Distance calculation	0.54 s/span

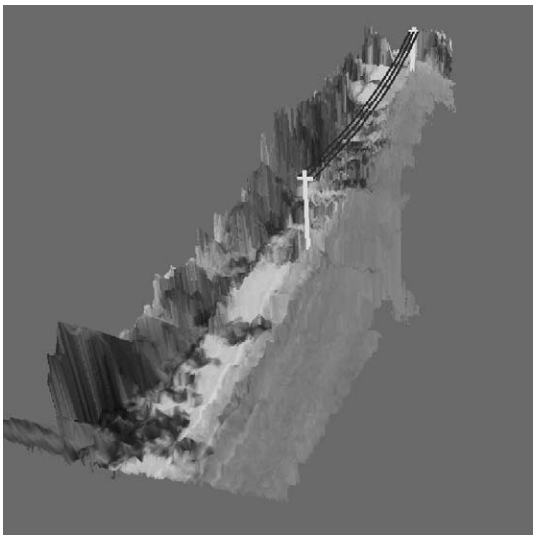

Fig. 12. A perspective view of the mosaicked 3D surface with power poles and powerlines.

Knowing the 3D positions of the tops of the power poles within a span of the images and the 3D terrain or vegetation surface, power poles can be drawn above the surface with the calculated height. Cross arms can also be modelled. From the 3D positions of the tops of the poles, the catenary equations of powerlines can be calculated and modelled. Fig. 12 shows a perspective view of the mosaicked 3D surface with power poles and powerlines on top of them.

9. Discussion and future work

We have presented techniques for automatically measuring the distance of vegetation from powerlines using stereo vision techniques. The proposed strategy has been to identify successive power poles and to model the powerlines between successive poles as a catenary. The vegetation surface is recovered using stereo vision techniques and by mosaicking several individual 3D surfaces from each pair of stereo images into a single 3D map within a successive pair of poles. Then the distance is measured from the catenary to the recovered vegetation surface.

By using high resolution digital cameras, the powerlines are more likely to be seen from images. In this case, the 3D position of powerlines could be determined from just image information without the need to know

parameters such as the load and tension of powerlines. If we need to obtain information such as tree species in addition to the surface of the trees, certain kinds of color or multispectral images will be necessary. Some of the techniques described have the potential to be applied in the mapping and positioning of linear and other features such as roads, railways, pipelines, fibre optic cables, streams and rivers, and, on a large scale, valley systems and shorelines. It could also be used for inventory updating, and geohazard and slope stability assessment of roads and railways.

The outcomes of this project lead us to conclude that automation of tree encroachment detection is technically possible and economically viable. We now have a clear understanding of future applications of automated image capture and analysis in distribution network management.

Acknowledgements

The financial support from Powercor Australia for this work is gratefully acknowledged. We thank the anonymous reviewers for their very constructive comments. We are grateful to Cherylann Biegler, Bob Coulter, Kevin Cryan, and Suzanne Lavery for their help during the course of this project.

References

Ackermann, F., 1999. Airborne laser scanning — present status and future expectations. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2/3), 64–67.

Ayache, N., Hansen, C., 1988. Rectification of images for binocular and trinocular stereovision. Proceedings of International Conference on Pattern Recognition, vol. 1. Ergife Palace Hotel, Rome, Italy, pp. 11–16.

Ayache, N., Lustman, F., 1991. Trinocular stereo vision for robotics. IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (1), 73–85.

Bartier, P.M., Keller, C.P., 1996. Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). Computers & Geosciences 22 (7), 795–799.

Campanella, R., Davis, B., Occhi, L., 1995. Transmission corridor encroachment detection: how remote sensing and gis can help. Earth Observation Magazine 4 (8), 23–25.

Carter, W.E., Shrestha, R.L., 1998. Engineering applications of airborne scanning lasers: reports from the field. Photogrammetric Engineering and Remote Sensing 64 (4), 246–253.

Dall, J.A., 1991. Danger tree surveys for transmission line maintenance. Technical Papers of the ACSM-ASPRS Annual Convention: Photogrammetry and Primary Data Acquisition, vol. 5. American Society for Photogrammetry and Remote Sensing, Baltimore, Maryland, pp. 61–67.

Faugeras, O., 1993. Three-Dimensional Computer Vision: A Geometric Viewpoint. The MIT Press.

FLI-MAP, 2005. FLI-MAP; Corridor Mapping. http://www.flimap.nl/. Accessed March 17, 2006.

Gauthier, R.P., Maloley, M., Fung, K.B., 2001. Land-cover anomaly detection along pipeline rights-of-way. Photogrammetric Engineering and Remote Sensing 67 (12), 1377–1389.

Golightly, I., Jones, D.I., 2003. Corner detection and matching for visual tracking during power line inspection. Image and Vision Computing 21 (9), 827–840.

Golightly, I., Jones, D.I., 2005. Visual control of an unmanned aerial vehicle for power line inspection. The 12th International Conference on Advanced Robotics, Seattle, Washington, USA, pp. 288–295.

Haralick, R., Shapiro, L., 1992. Computer and Robot Vision, vol. I. Addison-Wesley, Readings, Massachusetts.

Hyde, R.T., Wise, M.G., Stokes, R.H., Brasher Jr., E.C., 1996. Aircraft-based topographical data collection and processing system. US Patent, 5,557,397.

Jones, D.I., 2000. Aerial inspection of overhead power lines using video: estimation of image blurring due to vehicle and camera motion. IEE Proceedings. Vision, Image and Signal Processing 147 (2), 157–166.

Jones, D.I., Earp, G.K., 2002. Camera sightline pointing requirements for aerial inspection of overhead power lines. Electric Power Systems Research 57 (2), 73–82.

Jones, D.I., Whitworth, C.C., Duller, A.W.G., 2003. Image processing for the visual location of power line poles. In: Morrow, P.J., Scotney, B.W. (Eds.), Proceedings of Irish Machine Vision and Image Processing Conference. University of Ulster, Coleraine, pp. 177–184.

Kang, S.B., Webb, J.A., Zitnick, C.L., Kanade, T., 1994. An active multibaseline stereo system with real-time image acquisition. Tech. Rep. CMU-CS-94-167, School of Computer Science. Carnegie Mellon University.

Kraus, K., 1993. Photogrammetry, vol. 1. Dümmler, Bonn.

Nguyen, D.H., Paques, J., Laliberté, L., Bourbonniére, R., 1996. Numerical analysis of power line proximity warning device using electrical field measurement. Hazard Prevention 32 (1), 18–25.

Serra, J., 1982. Image Analysis and Mathematical Morphology. Academic Press.

Sun, C., 1997. A fast stereo matching method. Digital Image Computing: Techniques and Applications. Massey University, Auckland, New Zealand, pp. 95–100.

Sun, C., 2002. Fast stereo matching using rectangular subregioning and 3D maximum-surface techniques. International Journal of Computer Vision 47 (1/2/3), 99–117.

Sun, C., 2003. Uncalibrated three-view image rectification. Image and Vision Computing 21 (3), 259–269.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (6), 583–598.

Whitworth, C.C., Duller, A.W.G., Jones, D.I., Earp, G.K., 2001. Aerial video inspection of overhead power lines. The IEE Power Engineering Journal 15 (1), 25–32.

Williams, M., Jones, D.I., Earp, G.K., 2001. Obstacle avoidance during aerial inspection of power lines. Aircraft Engineering and Aerospace Technology 73 (5), 472–479.

Zhang, Z., Deriche, R., Faugeras, O., Luong, Q.-T., 1995. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence 78 (1/2), 87–119.