

# Rectification of SPOT Data for Stereo Image Matching

G. P. Otto, Lecturer,  
Department of Computer Science,  
University College London, Gower Street, London WC1E 6BT  
U. K.  
Commission No: III

## ABSTRACT

Many stereo matching algorithms (e.g. Ohta<sup>6</sup> or Grimson<sup>3</sup>) rely upon the assumption that the input data is in a form in which the corresponding epipolar lines are known, and are parallel to scan-lines. This paper considers the case of data from the SPOT satellite, and:

- (a) shows that a DEM is required before such rectification can be done accurately;
- (b) estimates the magnitude of the errors when only inaccurate disparity (or height) information is available, and shows that an iterative version of one of the "epi-polar constrained" algorithms is feasible;
- (c) and discusses how the parameters required for the rectification can be derived from the SPOT data alone.

These results are also applicable to many other line-scanning sensor systems.

## 1. Introduction

The geometrical aspects of stereo ranging can be divided into two classes: those which are needed to allow some particular matching algorithm to work, or which speed up the matching; and those which are needed to convert the disparities obtained from the matching into a set of positions in some desired reference frame. We will only be considering the first class here.

The motivation for this work is twofold:

- (i) Many stereo algorithms (e.g. Ohta & Kanade,<sup>6</sup> PMF,<sup>7</sup> Grimson,<sup>3</sup> Baker & Binford<sup>1</sup>) assume that the sensor geometry is such that the input data has horizontal, aligned, epipolar lines. (Or that it can be resampled into such a format.) Many others have some constraints on the sensor geometry (e.g. Nishihara<sup>5</sup>). We wish to know whether such algorithms can be adapted for SPOT data; and if so, how.
- (ii) Almost all stereo algorithms will run faster (and often give better quality results) if we make use of the sensor geometry to constrain the searching for matches. The traditional methods of doing this do not work well with SPOT data, since it is a scanning line sensor (which has a rather different geometry, and many more parameters than normal cameras), and is looking at a scene which is flat compared with the sensor-scene range, so that determining the relative orientation parameters from the image data is an ill-constrained problem.

Section 2 shows that SPOT data cannot be accurately rectified into a format where the "scan lines" are all horizontal, and aligned in the two images, unless good estimates of the disparities (or scene heights) are already available. Section 3 estimates the errors involved if we make a "best attempt" at such a rectification. People who are not interested in all the details should just look at section 3.1.3, which gives a couple of examples and briefly discusses the consequences of such errors. Section 4 discusses how the parameters required for the rectification can be derived from the SPOT data alone.

## 2. SPOT data cannot be rectified into an exact "epipolar aligned" † format without a DEM

As we mentioned above, many stereo algorithms assume that the sensor geometry is such that the input data has horizontal, aligned, epipolar lines, or that it can be resampled into such a format. This assumption is quite reasonable in the context of robotic vision, since the cameras can be set up in (more-or-less) whatever fashion is

† We will use the term "epipolar aligned" in this document to denote (stereo) images in which the epipolar lines are all horizontal, and aligned so that corresponding epipolar lines are at the same position in each image.

convenient. It is not unduly restrictive for any stereo system which uses sensors in which all the imaged rays go through a single point (e.g. the perspective point in the case of a conventional camera), since the image data can be resampled suitably once the sensor geometry is known. (See appendix 1.) This section shows why it *is* a problem for scanning line sensors, such as SPOT.

### 2.1. Definition of "aligned epipolar" format

The phrase "epipolar line" (and related terms) are defined in terms of a conventional stereo camera setup, where there are exactly two perspective centres, one for each camera. Since the SPOT geometry is different, we will start by defining exactly what we mean by "aligned epipolar" format, and why it is the appropriate definition for our purposes.

The assumption which is built into the stereo algorithms that we are mentioning is:

**that if we consider any point in one image, then the corresponding point in the other image must lie on the same scan line.**

(The corresponding feature/point may, of course, be occluded.)

In mathematical terms, we have a coordinate system for the two images such that corresponding pairs of points must have the same second† coordinate; i.e. that if a point-feature has coordinates  $(x, y)$  in one image, then it has coordinates  $(x', y)$  in the other.

If, and only if, a pair of images has such a set of coordinates, then we will refer to it as "epipolar alignable".

(N.B. In this document, we will not discuss the resampling of the data from whatever grid it was originally sampled on, to whatever grid of points its value is required at. This is an important issue, but one which is largely independent of what coordinates are being used.)

**Aside:** Those who like to think in geometrical terms may be interested to note that stereo input data can be "epipolar aligned" if, and only if, the 3-D disparity space can be decomposed into a 1-D set of 2-D disparity spaces.

### 2.2. SPOT sensor geometry

This section very briefly summarises the SPOT sensor geometry. For a fuller description, see (e.g.) Rosso<sup>8</sup> and Chevrel et al.<sup>2</sup>

The panchromatic SPOT images are really a composite of a succession of 1-D images, formed by a 6000 element CCD array. These 1-D images are taken at 1.5ms intervals, as the satellite orbits the earth at  $\sim 7$ km/s. The linear array is approximately perpendicular to the satellite's track, and the plane formed by the array & the perspective centre is approximately vertical. (For convenience, we will refer to this plane as the "visible plane" hereafter.) For stereo work, the array is frequently tilted so that it is looking at angles off the vertical of up to 25° or so, but the "visible plane" remains approximately vertical & perpendicular to the satellite track.

### 2.3. Determining which lines in the second image correspond to chosen points in the first

If we are to convert SPOT data into an epipolar aligned format, then we must be able to identify, *for each point in one image*, along which *line* in the second image, the corresponding point must lie. This section shows that this can be done if we know the sensor geometry; no knowledge of the imaged scene is needed to do this.

**Throughout the sections 2 and 3, we assume that we know everything we wish to about the imaging geometry.** e.g. The orbital & attitude parameters are fully known, as well as the sensor focal length etc. etc. (We will discuss a more realistic situation in section 4.)

Consider a point (any point) in one image. Since we know the imaging geometry, we know the line in space along which the corresponding scene feature must lie. (The line through the points at which the perspective centre and the appropriate element in the linear array were, when that line of the image was formed.) We also know the positions of the "visible planes" for each line of the other image; thus we can determine where this line intersects each visible plane. From this, we can determine which sensor element would "see" the feature, if it were at that position on the line. i.e. We know the set of possible corresponding points, and they form a line in the second image.

---

† We lose no real generality by assuming that it is the *second* coordinate. It would, of course, be the same *first* coordinate if we used (line, sample) coordinates.

(The procedure above may not be the best way of doing the calculation in practice, but it does show that it can be done.)

#### 2.4. Attempting to use this information to do "epipolar alignment"

It might appear that this is all we need to do the epipolar alignment. But, what it means is that we know two functions,  $f$  and  $g$ , such that:

- (a) given any point  $\underline{x}$  in the first image, we know that its corresponding point must lie at  $f(\underline{x}, t)$  in the second image for some  $t$ , and
- (b) given any point  $\underline{x}$  in the second image, we know that its corresponding point must lie at  $g(\underline{x}, t)$  in the first image for some  $t$ .

Let's see how far we get trying to use this to define "epipolar aligned" coordinates:

- (i) Let's select a point in the 1st image,  $(x_1, y_1)$  say. The corresponding line in the second image must have the same  $y$  coordinate, so we will *define* the new  $y$  coordinate of each of these points to be  $y_1$ . (We won't worry about what their new  $x$  coordinate should be at this stage.)
- (ii) Take a point on this line in the second image,  $(x_2, y_1)$  say. There must be a line in the first image which corresponds to this point; and they must also have the same  $y$  coordinate. Clearly, we wish to define that the new  $y$  coordinate of all these points is  $y_1$ , but we have to check that this is compatible with whatever else we have defined. So far, we have only fixed coordinates for the originally-selected point in the first image, and this lies on the line (because it is one of the possible corresponding points). So, no problem — we can define the  $y$  coordinate of this line of points in the 1st image to be  $y_1$ .
- (iii) Take another point on the line in the second image,  $(x_3, y_1)$  say. There must be a line in the first image which corresponds to this point; but is it the same line as we obtained in part (ii)? Both of these lines must pass through the original point  $(x_1, y_1)$ , but they might not coincide elsewhere. If they don't, we have a problem, since we will have two different lines of  $y = y_1$ , and we will not be able to define a well-behaved (continuous) coordinate system.†

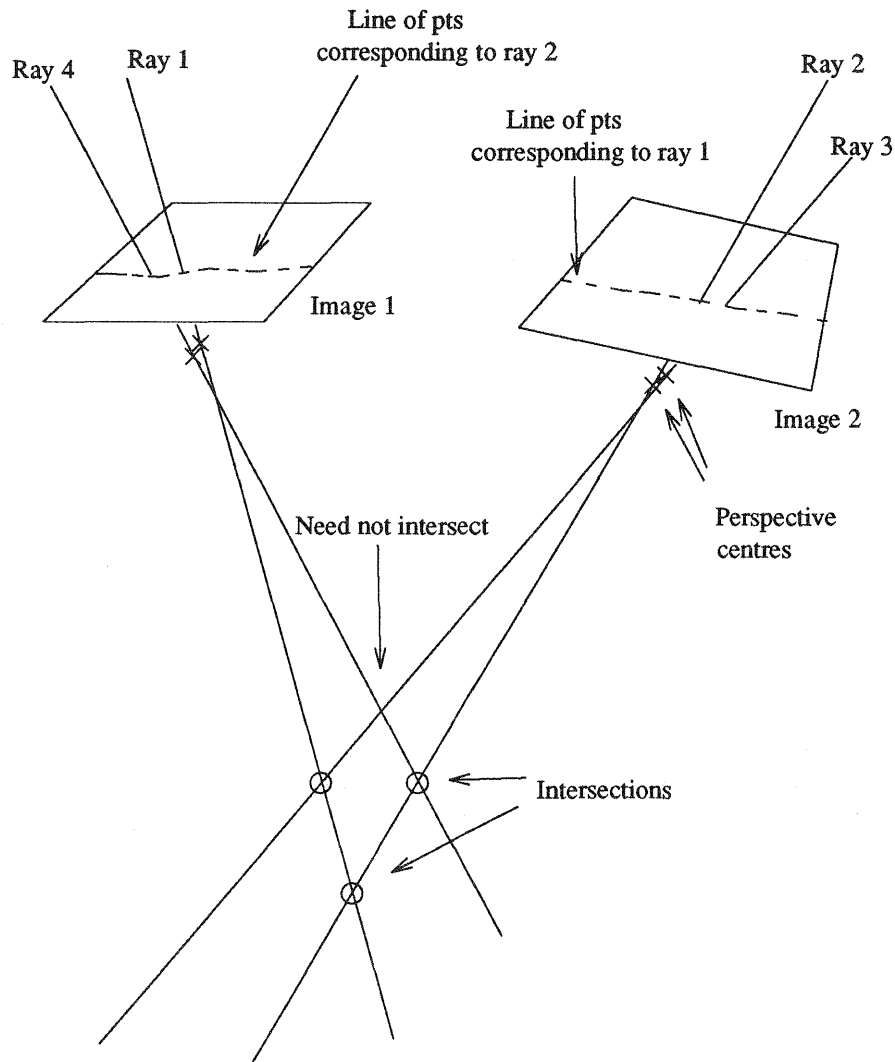
In the next section we will show that, in general, the two lines do not coincide, and that we cannot therefore define a set of "epipolar aligned" coordinates for a SPOT image pair.

#### 2.5. Proof that this won't work for SPOT data (in general)

Consider the diagram below.

---

† We could define one in which both images had the same  $y$  coordinate everywhere, but we would lose the property that nearby image points have coordinates which are nearly equal, and vice versa.

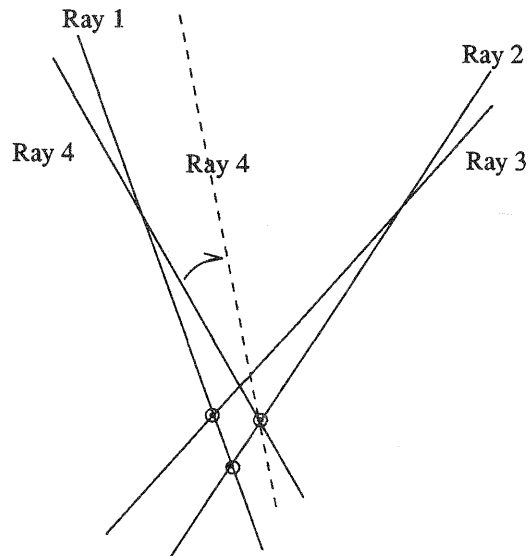


Take any point in image 1, and draw the line through it and the position of the sensors perspective centre at the time that it was imaged. This gives us ray 1. Now, as we described in the previous section, we can consider the various points along ray 1, and find out where they would get imaged in image 2. In this diagram, we have drawn in two such points, and the imaging rays through them (ray 2 and 3). Note that rays 2 and 3 will *not*, in general, go through the same perspective centre — i.e. they will go through two different perspective centres, as drawn on the diagram.

Similarly, if we construct ray 4 to go through a point on ray 2, it will, in general, go through a different perspective centre from ray 1. Now, since rays 2 and 3 give possible matches for ray 1, they must have the same "scan-line" coordinate as ray 1; similarly rays 1 and 4 must have the same "scan-line" coordinate as ray 2. This means that rays 3 and 4 must have the same "scan-line" coordinate.

If rays 3 and 4 intersect, there is no problem; if they don't, then they need to have different "scan-line" coordinates, or else they will get confused with one of the rays which do intersect the other. So, we have to decide whether rays 3 and 4 will intersect or not.

To do this, look at the diagram below.



The solid lines are the rays from the previous diagram, without any of the surrounding detail. The only fixed aspects of these lines is that they must intersect at the 3 circled points, and they must pass through the 4 perspective centres. However, although the satellite tracks will be approximately parallel, they are not exactly so; and in general, the positions of the 4 perspective centres can vary considerably. In particular, the track the satellite follows when taking image 1 is not directly connected to the one it follows when taking image 2, and so (for example) the position of the 4th perspective point can vary somewhat (in all 3 dimensions) relative to the other three. Ray 4 must still pass through the circled intersection (because of the way we constructed those lines), but this still allows it two degrees of freedom. Thus, it can move (say) to the position illustrated by the dotted line. A little consideration will make it clear that most possible positions will not intersect ray 3. In fact, in mathematical terms, "almost all" of the possible positions will not intersect, which means that in practice they will never exactly intersect.

To summarize: rays 3 and 4 must intersect if we are to be able to "epipolar align" the data, but in practice they will never exactly intersect — thus an exact "epipolar alignment" of SPOT data is not possible using just information about the sensor geometry.

N.B. The proof above is not formal, but it could be turned into such a proof if necessary. I have used this form of proof in the belief that it would be more understandable; I apologise to those who would have preferred more algebra.

### 3. Magnitude of the errors involved in "epipolar alignment" of SPOT

The previous section has shown that exact "epipolar alignment" cannot be done using sensor geometry information alone. But the errors might be small enough in practice that we could ignore them, so an estimate of the errors involved in any "approximate epipolar remapping" is required.† That is what this section provides.

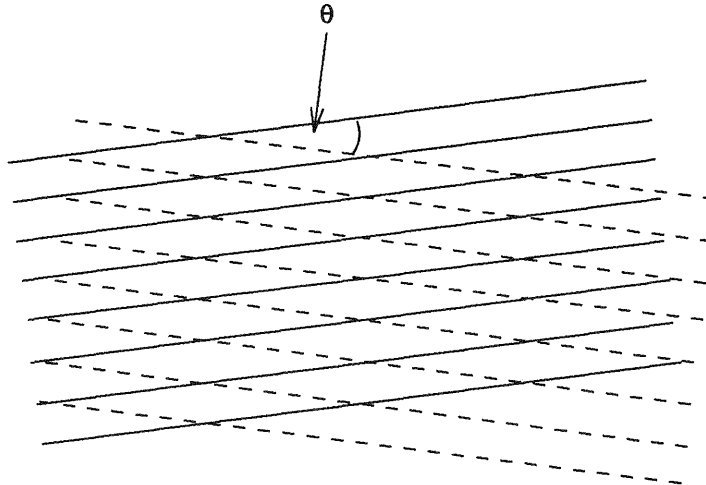
Roughly speaking, the SPOT "visible planes" are vertical (to within  $\sim 1/2^\circ$ ), and are parallel (to within  $\sim 10^{-3}^\circ$ ) within a given image. However, the visible planes of one image are typically at an angle of several degrees to the visible planes of the other (e.g.  $\sim 6^\circ$  in the case of a left and a right view of Aix-en-Provence‡). Let us therefore assume, temporarily, that the visible planes are exactly parallel & vertical within an image.

Let us project the imaging rays vertically downwards onto some reference plane (e.g. a nominal ground plane). (i.e. Take a plan view.) Because the visible planes are assumed to be vertical, all of the imaging rays within each plane will project onto the same line on the reference plane. Furthermore, because the planes (of each image separately) are assumed to be parallel, the projections of each image's visible planes will form a set of

† For example, if the "visible planes" in the two images are exactly parallel, then exact epipolar alignment is possible; perhaps the SPOT "visible planes" are close enough to parallel for the errors to be negligible.

‡ Image scene id's: S1H1860601105433 and S1H1860518102337; angles of incidence  $20^\circ$  from left, &  $26^\circ$  from right.

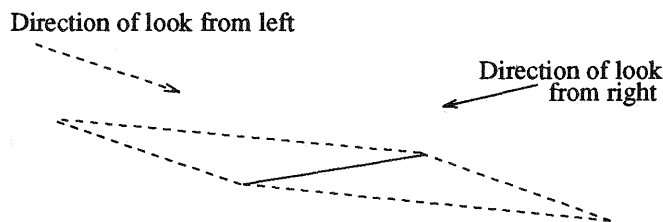
parallel lines on the reference plane. Thus we have a picture something like



where the dotted lines are the projections of the imaging rays from a satellite to the left†, the solid lines are the projections of the imaging rays from a satellite to the right, and  $\theta$  is the angle between the two sets of lines.

Now consider an imaging ray in one image — the left-hand view, say. As it passes from left to right through the scene, it intersects many visible planes of the right-hand view. If we clip this ray at some maximum and minimum height, then we will just have a line segment, and we can calculate how many right-hand visible planes it will intersect. For example, if the angle of incidence of the ray is  $25^\circ$ , and we allow a maximum height range of  $10^4$  metres, then the line segment will have a length of about  $5 \times 10^3$  m in plan; this will intersect about 50 scan-lines (visible planes) in the right-hand view if  $\theta$  is about  $6^\circ$ .

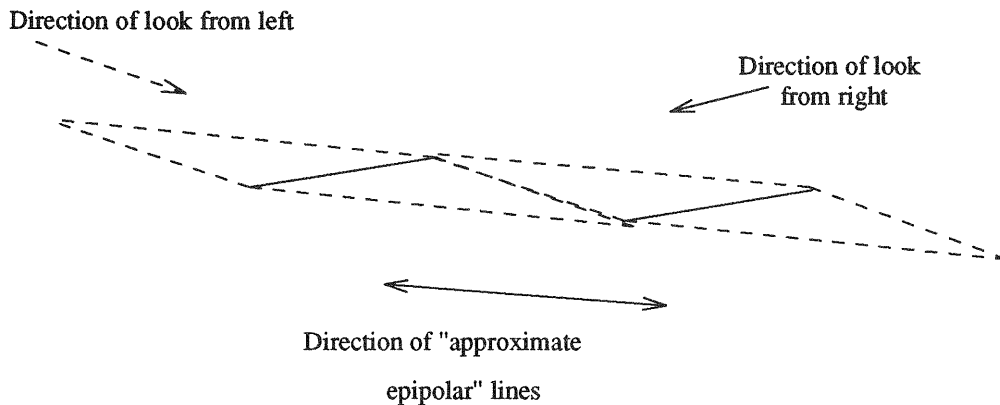
Using this basic idea, we can draw out the region which could correspond to any given pixel in one image. Thus, if we consider a pixel in the right-hand view, it could correspond to any point on the *solid* line in the diagram below.



Now let us consider which points in the left-hand view might correspond to that pixel. Any of the left-hand visible planes which intersect that solid line will contain a ray which intersects the right-hand ray; let us just consider the sections of those rays which lie within the height boundaries we have set. Clearly these intersecting line segments will be parallel to the projections of the visible planes; furthermore, at the left-hand end of the solid line, the intersecting line segment will lie entirely to the left, since that end of the line is at the minimum height boundary. Similarly, at the right-hand end of the solid line, the intersecting line segment will lie entirely to the right, since that end lies at the maximum height boundary. Thus, the area enclosed by the dotted line shows the projection of the intersecting line segments.

The significance of this region is that any point within it could match to one of the pixels in the left-hand image which was a potential correspondence to the original right-hand pixel. This should become clearer if we repeat the process.

† The picture is drawn on the assumption that the scene is in the northern hemisphere.



The diagram above shows the swath on the reference plane which is traced out when we find all the pixels which might match to a pixel which might match to a pixel . . . which might match to our originally selected pixel. In the ideal case, in which we can remap into an epipolar-aligned format, this should be a line. The fact that it isn't, just confirms the proof we gave in the previous section. However, this view gives us something extra: if we draw "approximate epipolar lines" along the centres of such swathes†, then

- (a) we know that any other approximation to an epipolar line must involve at least as much maximum error (unless it has better estimates of the height limits); and
- (b) we know how far any potential match might actually be from this approximate epipolar line.

Two things remain to be done here: deriving the formulae for the approximate epipolar lines and the errors involved in using them, and justifying the approximations. The next two sub-sections will handle these aspects.

### 3.1. Formulae for the error estimates

In the sub-sections below, let

- $\alpha_l$  denote the "SCENE ORIENTATION ANGLE" of the left-hand view
- $\alpha_r$  denote the "SCENE ORIENTATION ANGLE" of the right-hand view
- $i_l$  denote the "ANGLE OF INCIDENCE" of the left-hand view
- $i_r$  denote the "ANGLE OF INCIDENCE" of the right-hand view

These angles can all be read from the SPOT headers — the "SCENE ORIENTATION ANGLE" of a SPOT image is essentially the angle at which the visible planes intersect the ground reference plane, and the "ANGLE OF INCIDENCE" is the angle the imaging rays make with the vertical at the ground plane. (See SPOT-Image<sup>9</sup> for more information.)

Furthermore, let

- $\Delta\alpha$  denote  $\alpha_l - \alpha_r$
- $\Delta ht$  denote the difference between the maximum and minimum possible scene height

#### 3.1.1. Angle of orientation of the "approximate epipolar lines"

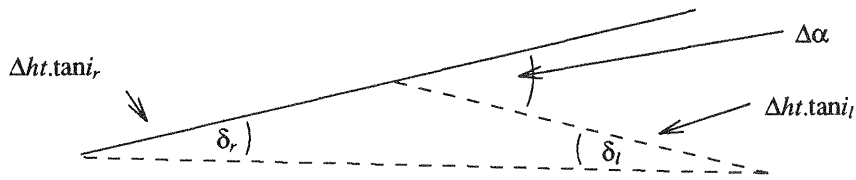
Each segment of a left-hand ray which lies within the height range will project onto a line segment at an angle  $\alpha_l$  to the horizontal, and  $\Delta ht \cdot \tan i_l$  long. Similarly, each segment of a right-hand ray which lies within the height range will project onto a line segment at an angle  $\alpha_r$  to the horizontal, and  $\Delta ht \cdot \tan i_r$  long. If we make a triangle with one of each of these, the hypotenuse is parallel to the approximate epipolar lines. (See the diagram above.) This gives us the diagram below.

† It is possible to resample in such a way — see section 4 if you have doubts.

Direction of look from left



Direction of look from right



Simple trigonometry then gives us that the length of the hypotenuse is

$$\Delta ht \sqrt{\tan^2 i_l + \tan^2 i_r + 2 \tan i_l \tan i_r \cos \Delta \alpha}$$

Then the sin rule gives us that

$$\begin{aligned} \frac{\sin \delta_l}{\tan i_r} &= \frac{\sin(\pi - \Delta \alpha)}{\sqrt{\tan^2 i_l + \tan^2 i_r + 2 \tan i_l \tan i_r \cos \Delta \alpha}} \\ &= \frac{\sin \Delta \alpha}{\sqrt{\tan^2 i_l + \tan^2 i_r + 2 \tan i_l \tan i_r \cos \Delta \alpha}} \end{aligned}$$

Thus

$$\delta_l = \sin^{-1} \left[ \frac{\tan i_r \cdot \sin \Delta \alpha}{\sqrt{\tan^2 i_l + \tan^2 i_r + 2 \tan i_l \tan i_r \cos \Delta \alpha}} \right]$$

Similarly,

$$\delta_r = \sin^{-1} \left[ \frac{\tan i_l \cdot \sin \Delta \alpha}{\sqrt{\tan^2 i_l + \tan^2 i_r + 2 \tan i_l \tan i_r \cos \Delta \alpha}} \right]$$

Readers may be reassured to note that both expressions reduce to

$$\frac{\sin \Delta \alpha}{2}$$

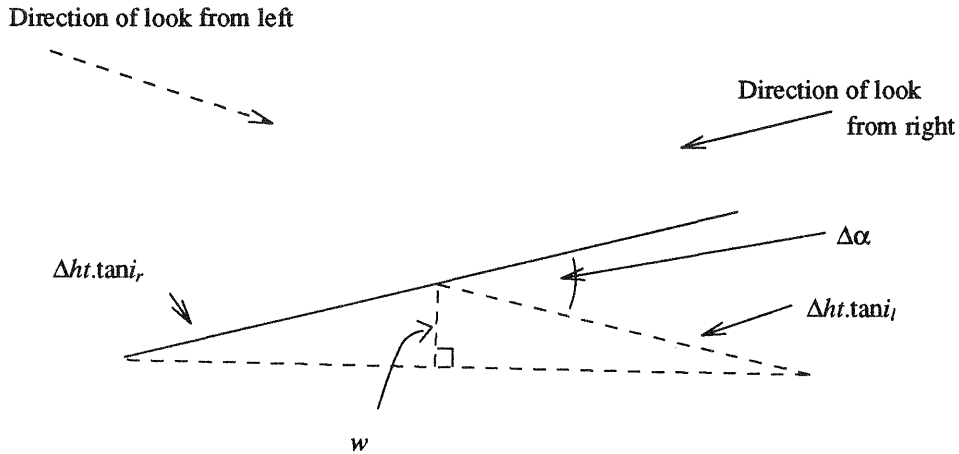
if  $i_l = i_r$ .

Knowing these angles ( $\delta_l, \delta_r$ ) allows us to plot the approximate epipolar lines on the plan view, and to estimate the error of this approximation. In addition, they are the main information we require to do the resampling of each image.

### 3.1.2. The error in the "epipolar-aligned" format

$w$  on the diagram below gives the width of the "approximate epipolar swath", so the maximum error (if we take the middle of the swath as our approximation) is  $\frac{w}{2}$ .





A little geometry gives us that

$$w = \left[ \Delta ht . \tan i_l \right] . \sin \delta_l$$

and thus that the maximum error is

$$\frac{1}{2} \cdot \frac{\Delta ht . \tan i_l . \tan i_r . \sin \Delta \alpha}{\sqrt{\tan^2 i_l + \tan^2 i_r + 2 \tan i_l \tan i_r \cos \Delta \alpha}}$$

which reduces to

$$\frac{1}{2} . \Delta ht . \tan i . \sin \left[ \frac{\Delta \alpha}{2} \right]$$

if  $i_l = i_r$ .

### 3.1.3. Some examples

A few examples will illustrate how significant the error is:

Matching the left-hand SPOT scene of Aix-en-Provence to a near-vertical view<sup>†</sup>, and assuming a maximum height uncertainty of 1000m, we get a maximum error when resampling to "epipolar" of 1.35m. i.e. If we were to resample just once to epipolar, then the true position (in 3-space) of a corresponding pair of image points might be as much as 1.35m away from the "epipolar" line/plane to which we have constrained it. Since each pixel represents a 10m×10m square on the ground, this error should be quite tolerable for almost all purposes. (But note that the error scales linearly with the height uncertainty. If we were looking at part of the Himalayas, the height uncertainty might be 10,000m, so a similar "left→vertical" match would have errors of up to 13.5m.)

Matching the left-hand SPOT scene of Aix-en-Provence to the right-hand scene, and assuming a maximum height uncertainty of 1000m, we get a maximum error when resampling to "epipolar" of 10.2m. This is OK for coarse matching, but is probably unacceptable for Photogrammetric work. The saving grace is that we could use such a remapping to get a coarse estimate, which could then be used to refine the mapping. In two or three steps this would be as accurate as we need.

Incidentally, note that the main reason for the "left→vert" errors being much smaller than the "left→right" errors is *not* the smaller base-to-height ratio. It is because the "vertical" image has an angle of incidence  $\approx 0$ , so that its  $\tan i$  term is approximately zero.

### 3.2. Justification of the approximations

Several approximations have been made in the process of deriving the error estimates above; this section is a brief justification of why these approximations do not significantly affect the result.

<sup>†</sup> The near-vertical scene (scene id: S1H1860512103904) has an angle of incidence of 3.6°; see previous footnote for information about left & right views.

The main approximation is that all non-linear terms (e.g. those due to curvature of earth, and the convergence of rays on the perspective points) have been ignored. However, all these effects only become significant over an extended region of the image, so we can ignore them if we consider the analysis above as applying only to small patches.

The other notable approximation used above was to simply use the absolute orientation data from the SPOT headers, which is known not to be completely accurate. However, we only need order-of-magnitude figures for the errors — the header information is quite accurate enough for that.

#### 4. Resampling to "approximate epipolar" format

The previous sections have assumed that we know everything we wish to about the sensor geometry. Unfortunately, this is unlikely to be true in practice, so this section discusses how the information we need can be extracted from the image data.

The crucial point to note is that we cannot derive a full set of absolute orientation parameters from the data because some parameters produce effectively indistinguishable effects on the image data — for example, deviations in pitch of the satellite are very difficult to distinguish from deviations in the along-track direction, because the scanner is linear.<sup>4</sup> However, if the effects of such parameters are effectively indistinguishable, we do not need to distinguish between them for the purpose of rectifying our data. Thus, the key to having a well-constrained problem is to use the smallest set of parameters which can model all the distortions we need to compensate for, and no more.

This rest of this section gives a few examples of transformations appropriate for rectification for stereo matching — the free parameters in these transformations are the only ones which need to be determined. Since each parameter has a distinguishable effect, the set of parameters can be measured from the image data by measuring the disparities at a small number of points.

##### 4.1. A linear transformation

Section 3.1 calculated the directions of the best approximations to epipolar lines.<sup>†</sup> Given this, our new coordinate system clearly ought to have one of its axes parallel to these lines everywhere. Furthermore, in order to align them, we need to allow for a vertical offset. This gives us a minimal transformation:

$$(x'_l, y'_l) = (x_l, y_l + a_l x_l + b)$$

$$(x'_r, y'_r) = (x_r, y_r + a_r x_r)$$

where  $x, y$  are the original coordinates in each image ( $x$  along a scan-line,  $y$  being scan-line number),  $x', y'$  being the new coordinates in each image,  $a_l, a_r$  being functions of  $\delta_l, \delta_r$  such that the  $y' = \text{const}$  lie are at  $\delta_l, \delta_r$  to the  $y = \text{const}$  lines in each image.<sup>‡</sup>

A minor addition to this is to allow for a minor scale change along the  $y$  axis by modifying the first equation to

$$(x'_l, y'_l) = (x_l, c y_l + a_l x_l + b)$$

where  $c$  is calculated to best align the  $y' = \text{const}$  axes.

Note that these remappings have been chosen so that the parameters can be derived from (a set of) image→image disparities, without any external reference data. Clearly, a rough set of parameters can be calculated using only the SPOT headers; the positional parameters coming from the "predicted Lat, Long" figures, since they are more accurate (and easier to use) than the orbital information given.

##### 4.1.1. Distortion effects ...

The remapping above is basically a shear, so the image is distorted slightly. If this is important (e.g. for correlation algorithms such as PRISM), then a non-trivial  $x \rightarrow x'$  transformation can be introduced, which rotates the  $x = \text{const}$  lines by the same amount as the  $y = \text{const}$  lines (and also does the magnification, if appropriate). This

<sup>†</sup> Under the assumption that the minimum & maximum possible heights were the same everywhere in the scene. The calculations need to be modified somewhat if these heights change over a scene.

<sup>‡</sup> Note that the pixels in SPOT images are typically not quite square. Allowance needs to be made for this (and shear effects due to, for example, the earth's rotation) when converting from distances and angles in the ground plane into pixel coordinates. However, if the transformation parameters are being derived directly from measurements of disparity, such considerations can be ignored.

does not introduce any new parameters, and so we will not discuss it further here.

#### 4.2. Non-linear transformations

The linear remapping above should be quite adequate over small regions of a SPOT image, but non-linear effects are too large to be ignored when whole images are being considered.

Non-linearities arising from the curvature of earth, and the convergence of rays on the perspective points, could be modelled well by a 2nd-order transformation, in the same style as the linear transformation. However, once accurate work is required, some knowledge of the relief needs to be incorporated (see section 3), and this means that no simple polynomial (or rational) transformation will be adequate. (Because relief is so irregular.) In addition, attitude variations of the satellite can cause significant and not-very-regular distortions (up to 1-2 pixels in the along-track direction). Thus, once large regions of the image need to be rectified, some form of patch-by-patch transformation is needed — piece-wise linear is probably most appropriate unless the terrain is particularly flat.

#### 5. Acknowledgements

This work was done as part of a collaborative research project (on "Real-Time 2½D Vision Systems"), funded by the UK's Alvey Directorate. The author would also like to thank Ian Dowman, Dave Gugan† & Peter Muller (all of the Department of Photogrammetry and Surveying at University College London) for helpful discussions.

#### Appendix 1 — Stereo images taken using conventional cameras can be resampled to an "epi-polar aligned" format

In order to obtain "epi-polar aligned" data directly from a pair of conventional cameras (which each have a single perspective centre), the "film planes" need to be coplanar & equidistant from the perspective centres. If the cameras were not actually in such a configuration, then a "virtual film plane" satisfying this criterion can be selected, and the data from each camera reprojected from the true "film plane" onto the virtual film plane. Provided the camera orientation data is known accurately, this is a (logically) straightforward task, even if the true "film plane" isn't actually planar, because of the single perspective centre.

#### References

1. H. H. Baker and T. O. Binford, "Depth from Edge and Intensity Based Stereo," *Proc. 7th International Joint Conference on Artificial Intelligence*, vol. 2, pp. 631-636, Vancouver, 1981.
2. M. Chevrel, M. Courtis, and G. Weill, "The SPOT satellite remote sensing mission," *Photogrammetric Engineering and Remote Sensing*, vol. 47, no. 8, pp. 1163-1171, 1981.
3. W. E. L. Grimson, "Computational Experiments with a Feature Based Stereo Algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, no. 1, pp. 17-34, 1985.
4. D. J. Gugan, "Practical Aspects of Topographic Mapping from SPOT Imagery," *Photogrammetric Record*, vol. 12, no. 69, pp. 349-355, April 1987.
5. H. K. Nishihara, "Practical real-time imaging stereo matcher," *Opt. Eng.*, vol. 23, no. 5, pp. 536-45, 1984.
6. Y. Ohta and T. Kanade, "Stereo by Intra- and Inter- Scanline Search Using Dynamic Programming," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, no. 2, pp. 139-154, 1985.
7. S. Pollard, J. Mayhew, and J. Frisby, "PMF: A stereo correspondence algorithm using a disparity gradient limit," *Perception*, vol. 14, pp. 449-470, 1985.
8. R. Rosso, "The SPOT satellite: assessment of the geometric quality of the images," *International Archives of Photogrammetry*, vol. 22, no. 4, pp. 455-473, 1978.
9. SPOT-Image, *The SPOT Standard CCT Format*, SPOT Image, Toulouse, 1985.

---

† Now at Laser-Scan Laboratories, Cambridge, UK.