EPIPOLAR IMAGES

Javier A. Múgica, October 2009

The information needed to generate the epipolar images is scattered between three places in two different papers: in "Elements of Orientation", the section 'Theory of one photograph' for the computation of a faked image; the last part of the section 'Models' in that same paper for the explanation of the epipolar images themselves; the paper "Orthophotographs" for the explanation of the resampling technique; and if we further want to take into account the distortion, as will usually be the case, the section 'Photograph with distortion' of the first paper is needed. For this reason in this paper a complete example is developed.

Example

Let the photographs be a pixel grid of $2400\times1800\,\mathrm{px}$. Let it have a radial distortion given by

$$\mathcal{D}_r = -31.5s - 35.8s^2 + 186s^3 - 92.2s^4, \qquad s = r/1500$$

and no tangential distortion; let the coordinates of the principal point in the fiducial system and the focal length be

$$(x_p^c, y_p^c) = (50.4, -18.5) \qquad \mathrm{f} = 1611.0$$

Finally, let the outer orientation parameters of the two images be

$$\mathrm{O}_1 = (1609.03, 999.57, 39.71) \qquad \mathrm{O}_2 = (1641.20, 1002.99, 40.02)$$

$$\mathrm{M}_1 = \begin{pmatrix} 0.86840 & 0.49583 & 0.00657 \\ -0.49572 & 0.86838 & -0.01351 \\ -0.01240 & 0.00848 & 0.99989 \end{pmatrix} \qquad \mathrm{M}_2 = \begin{pmatrix} 0.98216 & 0.18373 & 0.04014 \\ -0.18235 & 0.87823 & 0.44211 \\ 0.04598 & -0.44154 & 0.89607 \end{pmatrix}$$

It is required to create the epipolar images.

Since the epipolar image is rotated with respect to the original one there cannot be an exact correspondence between pixeles. The epipolar images will be created with the same focal length; this preserves as much as possible the pixel resolution. Since we can choose the value of the parameters (x_p^c, y_p^c), they will be chosen equal to (0,0), i.e., the fiducial system will be centered at the principal point.

The first step is to find the rotation matrix $\mathfrak{M}$ of the new images. We recall from the paper "Elements of Orientation" that

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{m}_1 \\ \mathfrak{m}_2 \\ \mathfrak{m}_3 \end{pmatrix}, \qquad \mathfrak{m}_2 = \frac{\mathbf{s} \wedge \mathfrak{m}_1}{|\mathbf{s} \wedge \mathfrak{m}_1|}, \qquad \mathfrak{m}_3 = \mathfrak{m}_1 \wedge \mathfrak{m}_2$$

where $\mathfrak{m}_1$ is the unitary vector in the direction and sense of $\overrightarrow{\mathrm{O}_1\mathrm{O}_2}$ and the auxiliary vector $\mathbf{s}$ can be chosen freely (so there is an infinity of solutions). Note however that vectors $\mathbf{s}$ differing only in modulus or determining the same plane together with $\mathfrak{m}_1$ will determine the same $\mathfrak{m}_2$, so there is just one degree of freedom in $\mathfrak{m}_2$.

These are common choices for the vector $\mathbf{s}$:

- $\mathbf{m}_3$ of the matrix M_1: The solution that modifies the least the orientation of the first image.
- $\mathbf{m}_3$ of the matriz M_2: The solution that modifies the least the orientation of the second image.
- $(0, 0, 1)$: The solution in which the image axes are closest to the vertical direction

We will choose the last one of these. So:

$$\overrightarrow{\mathrm{O}_1\mathrm{O}_2} = (32.17, 3.42, 0.31), \qquad \mathfrak{m}_1 = (0.99435, 0.10571, 0.00958)$$

$$\mathbf{s} \wedge \mathfrak{m}_1 = (0, 0, 1) \wedge \mathfrak{m}_1 = (-0.10571, 0.99435, 0), \qquad \mathfrak{m}_2 = (-0.10571, 0.99440, 0)$$

$$\mathfrak{m}_3 = \mathfrak{m}_1 \wedge \mathfrak{m}_2 = (-0.00953, -0.00101, 0.99995)$$

$$\mathfrak{M} = \begin{pmatrix} 0.99435 & 0.10571 & 0.00958 \\ -0.10571 & 0.99440 & 0.00000 \\ -0.00953 & -0.00101 & 0.99995 \end{pmatrix}$$

The matrices needed for resampling are the inverses of

$$\mathrm{N}_1 = \mathfrak{M}\mathrm{M}_1^{-1} = \begin{pmatrix} 0.91597 & -0.40125 & -0.00185 \\ 0.40125 & 0.91592 & 0.00974 \\ -0.00221 & -0.00967 & 0.99995 \end{pmatrix}$$

$$\mathrm{N}_2 = \mathfrak{M}\mathrm{M}_2^{-1} = \begin{pmatrix} 0.99642 & -0.08425 & 0.00763 \\ 0.07887 & 0.89259 & -0.44392 \\ 0.03059 & 0.44294 & 0.89603 \end{pmatrix}$$

We will compute a point belonging to the transformed of the first image; the computations for the second image are carried in the same way, using the parameters of that image instead of those of the first one. Let us suppose therefore that we want to find the value of the pixel (-1000,300) of the transformed first image. The coordinates of the transformed (epipolar) image are denoted by $(\mathfrak{x}, \mathfrak{y})$, and $\mathfrak{f}$ its focal length. So we are given the point

$$(\mathfrak{x}^c, \mathfrak{y}^c) = (-1000, 300)$$

Since we have taken the fiducial system of the epipolar image centered at the principal point, i.e., $(\mathfrak{x}_p^c, \mathfrak{y}_p^c) = (0, 0)$, these are also the coordinates in the principal system. So $(\mathfrak{x}, \mathfrak{y}) = (-1000, 300)$.

We will use the formulas given in "Elements of Orientation", but taking into account that we wish to resample, i.e., pass from $(\mathfrak{x}, \mathfrak{y})$ to (x, y), so the formulas have to be reversed. The formulas for resampling are

$$\begin{aligned}
x &= -\mathrm{f}\,\frac{n_{11}\mathfrak{x} + n_{21}\mathfrak{y} - n_{31}\mathfrak{f}}{n_{13}\mathfrak{x} + n_{23}\mathfrak{y} - n_{33}\mathfrak{f}} \\
y &= -\mathrm{f}\,\frac{n_{12}\mathfrak{x} + n_{22}\mathfrak{y} - n_{32}\mathfrak{f}}{n_{13}\mathfrak{x} + n_{23}\mathfrak{y} - n_{33}\mathfrak{f}} \\
z &= -\mathrm{f}
\end{aligned}$$

By substituting the computed and given values the formulas remain

$$\begin{aligned}
x &= -1611.0\,\frac{0.91597\mathfrak{x} + 0.40125\mathfrak{y} + 3.56}{-0.00185\mathfrak{x} + 0.00974\mathfrak{y} - 1610.93} \\
y &= -1611.0\,\frac{-0.40125\mathfrak{x} + 0.91592\mathfrak{y} + 15.57}{-0.00185\mathfrak{x} + 0.00974\mathfrak{y} - 1610.93}
\end{aligned}$$

Applying these formulas to the point $(\mathfrak{x}, \mathfrak{y}) = (-1000, 300)$ we get

$$x = -794.4 \qquad y = 693.7$$

But the transformations are not finished, for these coordinates are the ones in the principal system. The original image has a distortion as well as a fiducial system not centered at the principal point. First the distortion is applied.

$$r = \sqrt{x^2 + y^2} = 1054.7, \qquad s = r/1500 = 0.70311$$

$$\mathcal{D}_r = -31.5s - 35.8s^2 + 186s^3 - 92.2s^4 = 2.3$$

$$\mathcal{D}_x = \mathcal{D}_r(x/r) = 2.3\cdot(-0.75) = -1.7 \qquad \mathcal{D}_y = \mathcal{D}_r(y/r) = 2.3\cdot 0.66 = 1.5$$

So

$$(x', y') = (x, y) + (\mathcal{D}_x, \mathcal{D}_y) = (-796.1, 695.2)$$

And afterwards the offset of the fiducial center.

$$(x^c, y^c) = (x, y) + (x^c_p, y^c_p) = (-796.1, 695.2) + (50.4, -18.5) = (-745.7, 676.7)$$

This is the pixel in the original image:

$$(-1000, 300) \longleftarrow (-745.7, 676.7)$$

So we can assign to the pixel $(-1000, 300)$ of the epipolar image the value of the pixel $(-746, 677)$ of the original image, or perform some kind of interpolation between the values of this and the adjacent pixels, as explained in the paper "Orthophotographs".

Pixel to fiducial

Hitherto we have supposed that the fiducial system is always the measuring system. This was done by definition, therefore the developments are valid for any measuring system whatsoever. Nowadays the measuring system is always the pixel matrix, either because the image is a digital image or because it has been scanned and thus converted into digital, and the coordinates of a point are given but its pixel position within the matrix.

But digital images are actual files that are stored in a given format, and formats were not developed for the needs of fotogrammetry. The coordinate system which they carry is implicitly defined by the way pixels are refered to within them. This is: the first, x, coordinate equals zero for the leftmost pixel and increases to the right; the second, y, coordinate equals zero for the topmost pixel and increases to the bottom. This system has the orientation reversed, so it cannot directly be the fiducial system. A previous transformation form pixel to fiducial coordinates is needed, consisting in the change of the sign of the y coordinate. But there are images in which the pixel size is not the same for both x and y. If for the fiducial system the y pixel size is taken as the unit, in addition to the change in sign of the y values, the x coordinates need to be multiplied by the ratio of sizes x/y.

These transformations are trivial, but there is another restriction which is not as immediate to deal with as the previous ones in resampled images. It is the fact that there cannot be negative pixels. In the epipolar images, for instance, we have chosen to place the origin of the fiducial system at the principal point, which is the natural choice. This cannot possibly be the origin of the pixel system.

The complete path from pixel coordinates of a transformed image to pixel coordinates of the original one is

$$\mathfrak{pixel} \rightleftarrows \mathfrak{fiducial} = \mathfrak{principal} \rightleftarrows \text{principal} \rightleftarrows \text{fiducial} \rightleftarrows \text{pixel}$$

It is also customary to include in the pixel to fiducial transformation a displacement of the origin to the center of the matrix. So this transformation in the original image would be

$$x^c = k(x^p - \mathrm{T}_x)$$
$$y^c = -(y^p - \mathrm{T}_y)$$

where k is the ratio of the x pixel size to the y pixel size. A difference between the intended value of k and the actual one is an affinity distortion, and can be computed together with the other distortions, so that the value of k in this formula can be taken to be the round value, for example 1 or 2. But it may be desired to exclude that affinity component from the distortion function. For example, because it is too big compared to the other distortions. In that case the value of k will be the real one, not the theoretic one, including in it the distortion. For example, 0.9992 or 2.005.

In the epipolar image we have taken a perfect fiducial system, since the parameters defining it can be chosen; thus, its was taken without distortions and its origin at the principal point. For the same reason the value of k for the epipolar image will be taken equal to the theoretic value. If in the real image it was for instance 2.005, for the epipolar image the value 2 would be taken. The affinity distortion of 2.005/2 in the original pixel system is thus removed.

The values of T_x and T_y have to be computed so that no pixel from the transformed image has a negative coordinate.

Care is needed with the precise definition of the pixel system. A pixel is not a point; it is a square (or a rectangle), so it cannot be said that "the pixel" has such and such coordinates, $(100, 0)$ for instance, unless we are supposing that those are the coordinates of the pixel center. But there is another possible choice, which consists in taking round coordinates for the pixel corners. Suppose for example that there are 1000 pixels in the x direction. With the later choice the x coordinates of pixel centers would range from 0.5 to 999.5 and the total image from 0 to 1000. With the former, the pixel centers would range from 0 to 999 and the image from -0.5 to 999.5.

Here we will choose to give integer values to the pixel centers, for this is the practice followed in the paper "Ortophotographs", where the different resampling techniques are explained.

Example: Let the transformation from pixel to fiducial of the original image be

$$x^c = 0.9992(x^p - 1199.5)$$
$$y^c = -(y^p - 899.5)$$

in agreement with the giving of integer coordinates to the pixel centers.

In the epipolar image the translations $\mathfrak{T}_\mathrm{x}$ and $\mathfrak{T}_\mathrm{y}$ are subject to the requirement that no pixel shall have negative coordinates. Therefore we have to find the range of the transformed image. This is done by computing the transformated of the corners of the original image.

The pixel at the top-left corner has coordinates $(x^p, y^p) = (0, 0)$ and fiducial coordinates

$$x^c = 0.9992(0 - 1199.5) = -1198.5$$
$$y^c = -(0 - 899.5) = 899.5$$

The coordinates in the principal system with distortion are $(x', y') = (x^c, y^c) - (50.4, -18.5) = (-1248.9, 918.0)$. Its distortion is found:

$$r = 1550.0, \qquad s = 1.0333, \qquad \mathcal{D}_r = 29.3$$

$$\mathcal{D}_x = \mathcal{D}_r(x/r) = -23.6 \qquad \mathcal{D}_y = \mathcal{D}_r(y/r) = 17.4$$

And so $(x, y) = (-1225.3, 900.6)$.

The value of r used to enter the formula of the distortion has to be the real one, but when removing the distortion, i.e., going from (x', y') to (x, y) as we are doing now, we use the value based on (x', y'). So the obtained coordinates are not the exact ones, but others which are very near the exact ones. In order that this transformation be the inverse of the one from (x, y) to (x', y'), a second iteration should be carried out using the values of r based on the coordinates (x, y) computed, $(-1225.3, 900.6)$ in our example, which are much closer to the exact ones than (x', y'). The second iteration yields:

$$r = 1520.7, \qquad s = 1.0138, \qquad \mathcal{D}_r = 27.7$$
$$\mathcal{D}_x = -22.3 \qquad \mathcal{D}_y = 16.4$$

And so $(x, y) = (-1226.6, 901.6)$.

What follows is the passing from principal coordinates of the original image to principal coordinates of the transformed image:

$$\mathfrak{x} = -\mathfrak{f}\,\frac{n_{11}x + n_{12}y - n_{13}\mathrm{f}}{n_{31}x + n_{32}y - n_{33}\mathrm{f}} = -1476.9$$
$$\mathfrak{y} = -\mathfrak{f}\,\frac{n_{21}x + n_{22}y - n_{23}\mathrm{f}}{n_{31}x + n_{32}y - n_{33}\mathrm{f}} = 316.7$$

If the same process is carried out for the four corners the transformed coordinates are found to be

$$(\mathfrak{x}, \mathfrak{y})_1 = (\mathfrak{x}^c, \mathfrak{y}^c)_1 = (-1476.9, 316.7)$$
$$(\mathfrak{x}, \mathfrak{y})_2 = (\mathfrak{x}^c, \mathfrak{y}^c)_2 = (671.3, 1256.6)$$
$$(\mathfrak{x}, \mathfrak{y})_3 = (\mathfrak{x}^c, \mathfrak{y}^c)_3 = (-779.0, -1310.6)$$
$$(\mathfrak{x}, \mathfrak{y})_4 = (\mathfrak{x}^c, \mathfrak{y}^c)_4 = (1392.5, -358.2)$$

The pixel coordinates cannot be negative. We have

$$\mathfrak{x}^p = 1/\mathfrak{k}\,\mathfrak{x}^c + \mathfrak{T}_x$$
$$\mathfrak{y}^p = -\mathfrak{y}^c + \mathfrak{T}_y$$

$\mathfrak{k}$ is the theoretic round value, which in this example is 1. (If the value of k for the real image were for example 2.005, the value of $\mathfrak{k}$ here would equal 2). So

$$\mathfrak{x}^p = \mathfrak{x}^c + \mathfrak{T}_x$$
$$\mathfrak{y}^p = -\mathfrak{y}^c + \mathfrak{T}_y$$

The most negative value of $\mathfrak{x}^c$ is that of corner 1, and that of $-\mathfrak{y}^c$ is the one of corner 2. Based on these values we can take $\mathfrak{T}_x = 1477$ and $\mathfrak{T}_y = 1257$. With this values the pixel coordinates of the epipolar image are

$$(\mathfrak{x}^p, \mathfrak{y}^p)_1 = (0.1, 940.3)$$
$$(\mathfrak{x}^p, \mathfrak{y}^p)_2 = (2148.3, 0.4)$$
$$(\mathfrak{x}^p, \mathfrak{y}^p)_3 = (698.0, 2567.6)$$
$$(\mathfrak{x}^p, \mathfrak{y}^p)_4 = (2869.5, 1615.2)$$

The range of pixel center coordinates of the epipolar image is therefore (0–2870) × (0–2568), which is an image of 2871 × 2569.

Since the epipolar image is rotated with respect to the original one some areas of the 2871 × 2569 rectangle won't correspond to points of the original image and will thus be blank.

We recall the chain conecting $(\mathfrak{x}^p, \mathfrak{y}^p)$ with (x^p, y^p).

$$\mathfrak{pixel} \rightleftarrows \mathfrak{fiducial} = \mathfrak{principal} \rightleftarrows \text{principal} \rightleftarrows \text{fiducial} \rightleftarrows \text{pixel}$$

The example developed in the previous section ommited the transformations from pixel to fiducial. So now the example will be completed.

Suppose that we are asked to find the value of the pixel $(477, 957)$ of the epipolar image.

1st. $(\mathfrak{x}^p, \mathfrak{y}^p) \to (\mathfrak{x}^c, \mathfrak{y}^c)$:

$$\mathfrak{x}^c = \mathfrak{k}(\mathfrak{x}^p - \mathfrak{T}_x), \qquad \mathfrak{x}^c = 1(\mathfrak{x}^p - 1477) = -1000$$
$$\mathfrak{y}^c = -(\mathfrak{y}^p - \mathfrak{T}_y), \qquad \mathfrak{y}^c = -(\mathfrak{x}^p - 1257) = 300$$

2nd. $(\mathfrak{x}^c, \mathfrak{y}^c) \to (\mathfrak{x}, \mathfrak{y})$:

$$(\mathfrak{x}^c, \mathfrak{y}^c) = (\mathfrak{x}, \mathfrak{y}) = (-1000, 300)$$

3rd. $(\mathfrak{x}, \mathfrak{y}) \to (x, y)$:

$$x = -\mathrm{f}\,\frac{n_{11}\mathfrak{x} + n_{21}\mathfrak{y} - n_{31}\mathfrak{f}}{n_{13}\mathfrak{x} + n_{23}\mathfrak{y} - n_{33}\mathfrak{f}} = -794.4$$

$$y = -\mathrm{f}\,\frac{n_{12}\mathfrak{x} + n_{22}\mathfrak{y} - n_{32}\mathfrak{f}}{n_{13}\mathfrak{x} + n_{23}\mathfrak{y} - n_{33}\mathfrak{f}} = 693.7$$

4^{th}. $(x, y) \to (x^c, y^c)$:

a) $(x, y) \to (x', y')$

$$(x', y') = (x, y) + (\mathcal{D}_x, \mathcal{D}_y) = (-794.4, 693.7) + (-1.7, 1.5) = (-796.1, 695.2)$$

b) $(x', y') \to (x^c, y^c)$

$$(x^c, y^c) = (x', y') + (x^c_p, y^c_p) = (-796.1, 695.2) + (50.4, -18.5) = (-745.7, 676.7)$$

5^{th}. $(x^c, y^c) \to (x^p, y^p)$:

$$\begin{aligned} x^p &= 1/k\, x^c + \mathrm{T}_x, \qquad & x^p &= 1/0.9992\, x^c + 1199.5 = 453.2 \\ y^p &= -y^c + \mathrm{T}_y, \qquad & y^p &= -y^c + 899.5 = 222.8 \end{aligned}$$