Complexity analysis for the establishment of image correspondences of dense spatial target fields

Hans-Gerd Maas

Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology

ETH - Hoenggerberg, CH - 8093 Zurich

Abstract:

The establishment of stereoscopic correspondences for a large number of targets in a true 3-D application without a continuous surface connecting the targets does often pose difficult problems to automatic or semiautomatic processing systems. If the targets do not show any features which allow for a reliable distinction of candidates, only the geometric criterium of the perpendicular distance of a candidate to the epipolar line can be applied. Depending on the number of targets and the depth extension of object space this may lead to unsolvable ambiguities. As an example for this problem an application of digital photogrammetry to 3-D particle tracking velocimetry can be considered. In this paper two methods will be discussed to reduce the number of ambiguities drastically by employing three or more cameras in special configurations: the method of intersection of epipolar lines and a method with asymmetric arrangement of three cameras known from computer vision. In a detailed analysis of the methods the reduction of the number of expectable ambiguities, which can amount to a reduction factor of up to 100, will be proven and quantified.

1. Introduction

Three-dimensional particle tracking velocimetry (3-D PTV) is one of the most powerful flow measurement techniques. It is based on seeding a flow with small, reflecting, neutrally buoyant particles and recording image sequences of these particles by a stereoscopic camera setup. To achieve a sufficiently high spatial resolution a dense seeding of the flow (1000 - 2000 particles) is usually required. With video technique and methods of digital photogrammetry completely automatic PTV systems can be developed today (Papantoniou/Maas, 1990). Trying to judge the reliability of such a system one has to cope the fact that the high target density causes ambiguities in some steps of the processing of image sequences in order to derive particle trajectories. The data processing from images to trajectories can be divided into the following major processing steps (Maas, 1990):

- Image segmentation / determination of particle image coordinates
- Establishment of correspondences between particle images in different views
- Computation of spatial coordinates
- Tracking
- (Interpolation to regular grid)

Ambiguities may occur as particles partly or totally overlapping each other in one or more views in the segmentation step, as multiple candidates in the epipolar search window in the procedure of the establishment of stereoscopic correspondences and as multiple solutions in tracking. This paper will only address ambiguities in the photogrammetric matching process; ambiguities in image segmentation and in tracking can be estimated following
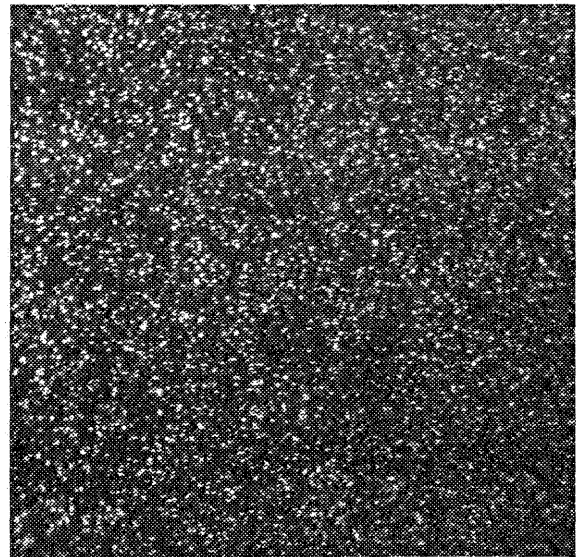
Maas (1992) or Adrian (1991).

Figure 1: Example of particle image (~ 1400 particles)

Figure 1 shows an example of a typical particle image with some 1400 imaged particles. Once the image coordinates of all particles in all images have been determined, correspondences between data of the different images have to be established to be able to calculate the 3-D coordinates. In photogrammetry we employ the epipolar geometry to solve this problem. Knowing the orientation parameters of the cameras from a calibration procedure, proceeding from a point P' in one image an epipolar line in an other image can be defined on which the corresponding point has to be found. In the strict mathematical formulation this line is a straight line, in the more general case with convergent

camera-axes, non-negligible lens distortion and multimedia geometry (object and sensor in media with different refractive indices) the epipolar line will be a slightly bended line. Its length l can be restricted if approximate knowledge about the depth range in object space is available, e.g. the range of the illuminated test section. Adding a certain tolerance width ε to this epipolar line segment (due to data quality) the search area for the corresponding particle image becomes a narrow twodimensional window in image space.

2. Two-camera arrangement

With the large number of imaged particles a problem of ambiguities occurs here, as often two or more particles will be found in the search area. If the particle features like size, shape or color do not allow a reliable distinction of particles, these ambiguities cannot be solved by a system based on only two cameras.

For a quantification of the probability of the occurence of ambiguities a point P centered in object space shall be considered: $X = b_{12}/2$, $Z = (Z_{min} + Z_{max})/2$, $Y = 0$ (Figure 2, consideration in epipolar plane without loss of generality).
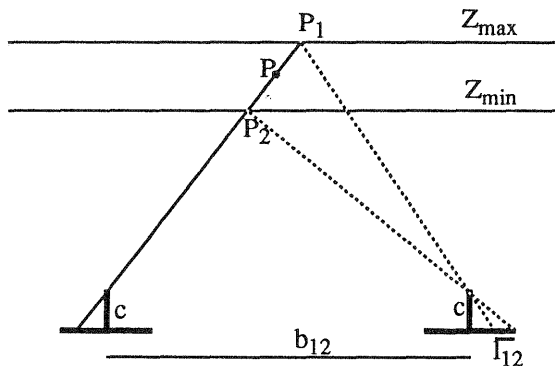

Figure 2: length of epipolar search window

With

$$\frac{X}{Z} = \frac{X_1}{Z_{max}} = \frac{X_2}{Z_{min}} = \frac{x'}{c}$$

$$\Rightarrow \quad X_1 = Z_{max} \cdot \frac{X}{Z}, \quad X_2 = Z_{min} \cdot \frac{X}{Z} \qquad \text{(Eq. 1)}$$

the length of the epipolar search window becomes

$$l_{12} = x''_2 - x''_1 = c \cdot \left(\frac{b_{12} - X_2}{Z_{min}} - \frac{b_{12} - X_1}{Z_{max}} \right)$$

$$= c \cdot \left(\frac{b_{12}}{Z_{min}} - \frac{b_{12}}{Z_{max}} \right)$$

$$= \frac{c \cdot b_{12} \cdot (Z_{max} - Z_{min})}{Z_{min} \cdot Z_{max}}, \qquad \text{(Eq. 2)}$$

and with the average number of ambiguous particles per search window

$$P_{a_{12}} = (n-1) \cdot \frac{2 \cdot l_{12} \cdot \varepsilon \cdot}{F} \qquad \text{(Eq. 3)}$$

one receives the expectable number of ambiguities per stereopair

$$N_a = (n^2 - n) \cdot \frac{2 \cdot c \cdot \varepsilon \cdot b_{12} \cdot (Z_{max} - Z_{min})}{F \cdot Z_{min} \cdot Z_{max}}. \qquad \text{(Eq. 4)}$$

The number of ambiguities grows
- approximately with the square of the number of particles
- linearly with the length of the epipolar line segment
- linearly with the width of the epipolar search window

With realistic suppositions for the number of particles per image and the dimensions of the epipolar search window in a reasonable camera setup the number of ambiguities to be expected becomes that large (see table 1), that a two-camera-system will not allow for a robust solution of the correspondence problem, if the number of targets or the depth range in object space cannot be controlled strictly. Instead algorithms based on three or more cameras rather than two will be discussed in the following, which allow a drastical reduction of the expectable number of ambiguities.

2.1 Intersection of epipolar lines

A consequent solution of the problem is the use of a third camera in a setup as shown in Figure 3 with the aim of reducing the search space from a line plus tolerance to the intersection of lines plus tolerance.
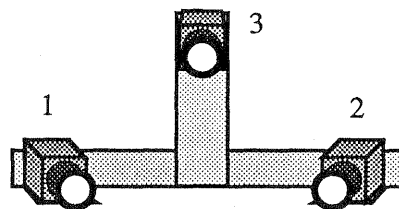

Figure 3: arrangement of three CCD cameras for the method of intersection of epipolar lines
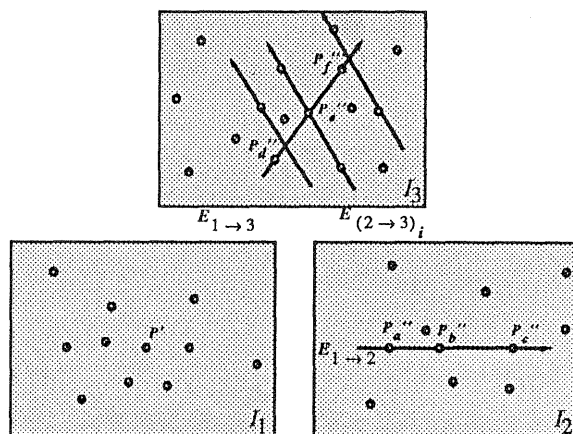
This setup can be exploited as shown in Figure 4:

Figure 4: principle of intersection of epipolar lines

103

Proceeding from a point P' in the image I_1 all epipolar lines $E_{1 \to 2}$ in I_2 and $E_{1 \to 3}$ in I_3 are being derived, on which candidates P_a'', P_b'' and P_c'' resp. P_d''', P_e''' and P_f''' may be assumed to be found. An unambiguous determination of the particle image corresponding to P' can neither be found in I_2 nor in I_3. However if all epipolar lines $E_{(2 \to 3)_i}$ of all candidates P_i in I_2 are being intersected with the epipolar line $E_{1 \to 3}$, there will be a large probability that only one of the intersection points will be close at one of the candidates in I_3 (in Figure 4: P_e'''). This consideration has been implemented via a combinatorics algorithm which tries to find such consistent triplets in the three datasets and rejects points which are members of more than one consistent triplet. Such an unambiguous consistent triplet is a necessary and sufficient condition for the establishment of a correct correspondence. A similar, iterative approach can be found in (Kearney, 1991).

This procedure can reduce the probability of ambiguities drastically, but not totally. The remaining unsolvable, but detectable ambiguities can be seperated into three cases:

1. a 'wrong' candidate Q'' on the epipolar line $E_{1 \to 2}$ has got a corresponding particle image Q''' on $E_{2 \to 3}$, which accidently falls onto the epipolar line $E_{1 \to 3}$:
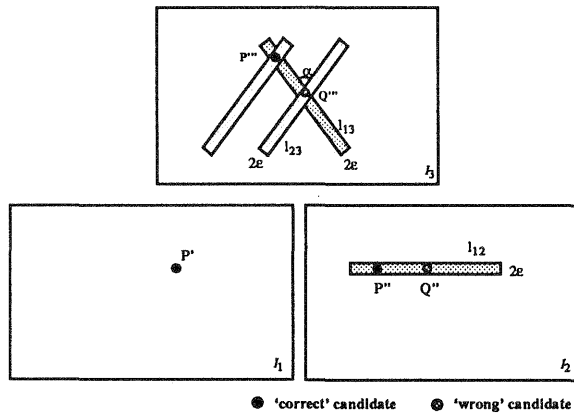

Figure 5: Intersection of epipolar lines - first kind of ambiguity

For a point P centered in object space we get with:

$$l_{12} = c \cdot \frac{b_{12} \cdot (Z_{max} - Z_{min})}{Z_{min} \cdot Z_{max}}$$

$$l_{13} = c \cdot \frac{b_{13} \cdot (Z_{max} - Z_{min})}{Z_{min} \cdot Z_{max}}$$

$$l_{23} = c \cdot \frac{b_{23} \cdot (Z_{max} - Z_{min})}{Z_{min} \cdot Z_{max}} \qquad \text{(Eq. 5)}$$

$$f_{23} = K_{1 \to 3} \cap K_{2 \to 3} = \frac{4 \cdot \varepsilon^2}{\sin \alpha}$$

and with

$$P_{12} = \frac{2 \cdot (n - 1) \cdot c \cdot \varepsilon \cdot b_{12} \cdot (Z_{max} - Z_{min})}{F \cdot Z_{min} \cdot Z_{max}}, \qquad \text{(Eq. 6)}$$

$$P_{23} = \frac{f_{23}}{2 \cdot \varepsilon \cdot l_{23}} = \frac{2 \cdot \varepsilon \cdot Z_{min} \cdot Z_{max}}{c \cdot b_{23} \cdot \sin \alpha \cdot (Z_{max} - Z_{min})} \text{(Eq. 7)}$$

the probability $P_{a(1)}$ of this first kind of ambiguities becomes

$$P_{a(1)} = P_{12} \cdot P_{23} = \frac{4 \cdot (n - 1) \cdot \varepsilon^2 \cdot b_{12}}{F \cdot b_{23} \cdot \sin \alpha}. \qquad \text{(Eq. 8)}$$

2. the epipolar line $E_{2 \to 3}$ of a 'wrong' candidate Q'' on the epipolar line $E_{1 \to 2}$ does also hit the 'correct' candidate P''' on $E_{1 \to 3}$, because a candidate Q'' is placed too close at the 'correct' candidate P''', or because a too short base component b_{13} has been chosen:

● ''correct' candidate   ◎ 'wrong' candidate
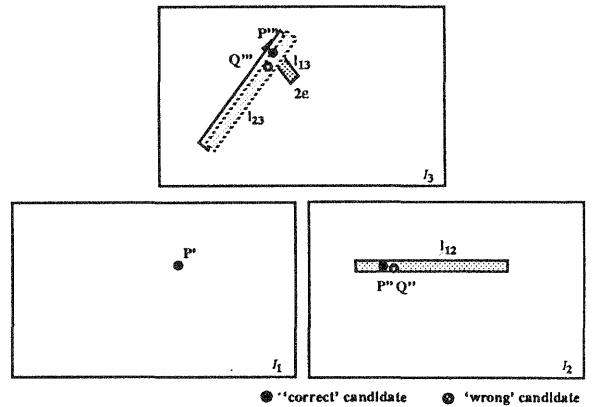
Figure 6: Intersection of epipolar lines - second kind of ambiguity

With (Eq. 5), (Eq. 6) the probability for this second kind of ambiguity is

$$P_{a(2)} = P_{12} \cdot \frac{f_{23}}{2 \cdot \varepsilon \cdot l_{13}}$$

$$= \frac{4 \cdot (n - 1) \cdot \varepsilon^2 \cdot b_{12}}{F \cdot b_{13} \cdot \sin \alpha}. \qquad \text{(Eq. 9)}$$

3. A second candidate R''' is found at the intersection of the epipolar lines $E_{1 \to 3}$ and $E_{2 \to 3}$ of the 'correct' candidate P'/P'' - an event which is often correlated with the occurence of an overlap:

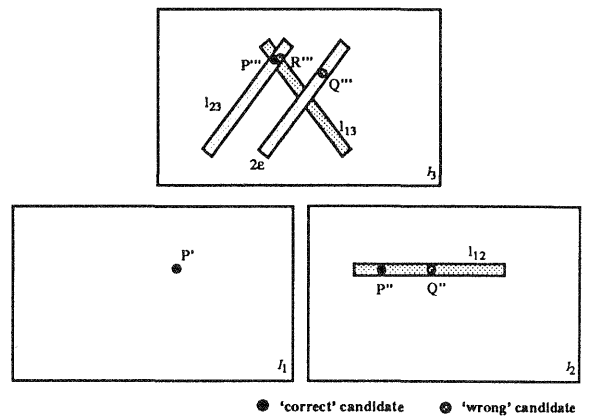● 'correct' candidate   ◎ 'wrong' candidate

Figure 7: Intersection of epipolar lines - third kind of ambiguity

104

The probability for this third kind of ambiguity is

$$P_{a(3)} = \frac{(n-1) \cdot f_{23}}{F} = \frac{4 \cdot (n-1) \cdot \varepsilon^2}{F \cdot \sin \alpha} . \qquad \text{(Eq. 10)}$$

With (Eq. 8), (Eq. 9) and (Eq. 10) the probability of an unsolvable ambiguity in the method of intersection of epipolar lines becomes

$$P_a = P_{a(1)} + P_{a(2)} + P_{a(3)}$$

$$= \left(\frac{4 \cdot (n-1) \cdot \varepsilon^2}{F \cdot \sin \alpha} \cdot \left(1 + \frac{b_{12}}{b_{23}} + \frac{b_{12}}{b_{13}} \right) \right), \qquad \text{(Eq. 11)}$$

and the expectable number of remaining ambiguities becomes

$$N_a = \frac{4 \cdot (n^2 - n) \cdot \varepsilon^2}{F \cdot \sin \alpha} \cdot \left(1 + \frac{b_{12}}{b_{23}} + \frac{b_{12}}{b_{13}} \right). \qquad \text{(Eq. 12)}$$

An optimum (i.e. a minimum number of remaining ambiguities) is achieved with $b_{12} = b_{13} = b_{23}$, which means a configuration of the three projective centers in a equilateral triangle. Other than in a two camera model the number of ambiguities does not depend on the length of the epipolar lines (i.e. on the depth range in object space resp. the baselength) any longer. In total the number of ambiguities is being reduced by at least a factor of 10 (see table 1).

2.2 Collinear arrangement of three cameras

The method of intersection of epipolar lines may be the most evident, but it is not the only way of exploiting a third camera. Using a different algorithm one can also work with three cameras which are arranged in a way that their projective centers are lying on a straight line as shown in Figure 8. In this case possible correspondences between the first and the second image have to be verified by a propagation into the third image.
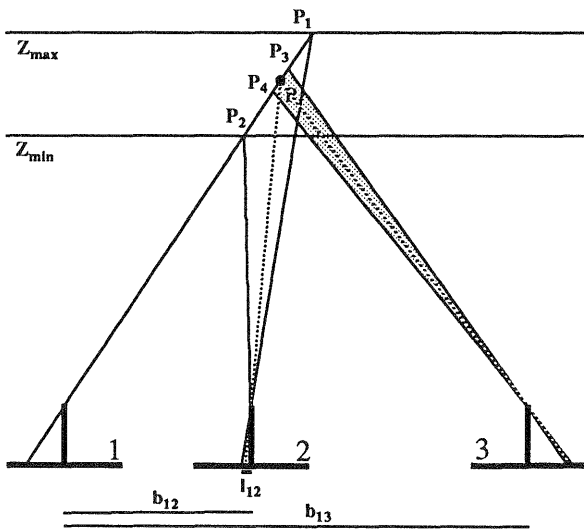

Figure 8: Proceeding with three collinearly arranged cameras

For all possible matches (1-2) a point in object space is being calculated

$$Z = \frac{c \cdot b_{12}}{p_x} \qquad X = Z \cdot \frac{x'}{c} \qquad Y = 0. \qquad \text{(Eq. 13)}$$

Depending on an assumed maximum error ε of the parallax p_x the thus established point(s) will have an error mainly in depth; this leads to a reduced search space Z_3, Z_4 in object space:

$$Z_3 = \frac{c \cdot b_{12}}{p_x - \varepsilon} = \frac{c \cdot b_{12} \cdot Z}{c \cdot b_{12} - (Z \cdot \varepsilon)}$$

$$X_3 = Z_3 \cdot \frac{X}{Z} = Z_3 \cdot \frac{b_{13}}{Z_{max} + Z_{min}}$$

$$Z_4 = \frac{c \cdot b_{12}}{p_x + \varepsilon} = \frac{c \cdot b_{12} \cdot Z}{c \cdot b_{12} + Z \cdot \varepsilon}$$

$$X_4 = Z_4 \cdot \frac{X}{Z} = Z_4 \cdot \frac{b_{13}}{Z_{max} + Z_{min}} \qquad \text{(Eq. 14)}$$

which is being imaged into image 3, where the length l_{123} of the search window becomes

$$l_{123} = x'''_4 - x'''_3$$

$$= c \cdot \left(\frac{b_{13} - X_4}{Z_4} - \frac{b_{13} - X_3}{Z_3} \right)$$

$$= c \cdot \left(\frac{b_{13}}{Z_4} - \frac{b_{13}}{Z_3} \right)$$

$$= c \cdot b_{13} \cdot \left(\frac{(c \cdot b_{12} + Z \cdot \varepsilon)}{c \cdot b_{12} \cdot Z} - \frac{(c \cdot b_{12} - (Z \cdot \varepsilon))}{c \cdot b_{12} \cdot Z} \right)$$

$$= 2 \cdot \varepsilon \cdot \frac{b_{13}}{b_{12}} . \qquad \text{(Eq. 15)}$$

This way one receives a short epipolar search area in image 3 for all the candidates in image 2. If exactly one valid candidate is found in these search spaces the necessary and sufficient condition for a correct correspondence is fulfilled.

A similar proceeding is used by Lotz/Fröschle (1990); they suggest a strongly asymmetric arrangement of cameras as shown in Figure 9 to reduce the probability of occurence of ambiguities.
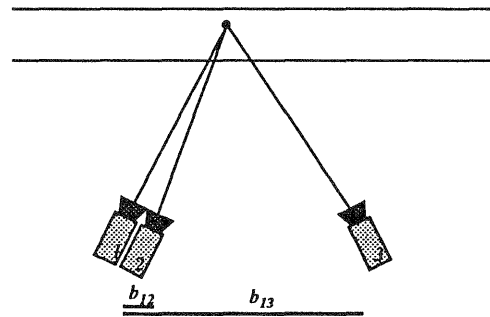

Figure 9: asymmetric camera arrangement (Lotz/Fröschle, 1990)

The short base b_{12} guarantees for a small number of ambiguities in the establishment of correspondences between image 1 and 2, while the long base b_{13} fulfills the requirement of good depth accuracy. As shown later (Eq. 19 - 23), this arrangement can minimize the probability of occurence of ambiguities but does not take into consideration that ambiguities can be solved; thus it does not represent an ideal setup if the total number of unsolvable ambiguities is to be minimized.

Like the method of intersection of epipolar lines this collinear arrangement has some remaining ambiguities, which cannot be solved. Two kinds of ambiguities can be distinguished:
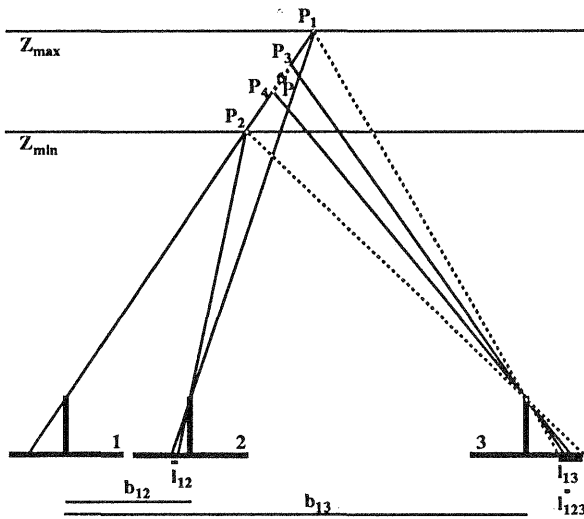

Figure 10: length of epipolar line segments for three-camera-setup

1. A point R''' is accidently imaged in the search area l_{23} of a 'wrong' candidate Q'' on l_{12}.

   With

   $$l_{123} = 2 \cdot \varepsilon \cdot \frac{b_{13}}{b_{12}} \,,$$

   $$l_{23} = c \cdot \frac{(b_{13} - b_{12}) \cdot (Z_{max} - Z_{min})}{Z_{min} \cdot Z_{max}}$$

   one receives

   $$P_{a(1)} = P_{12} \cdot P_{23} = \frac{4 \cdot (n-1) \cdot \varepsilon^2 \cdot b_{13}}{F \cdot (b_{13} - b_{12})} \,. \qquad \text{(Eq. 16)}$$

2. A second point Q''' is detected in the search area l_{23} of the 'correct' candidate:

   $$P_{a(2)} = \frac{2 \cdot (n-1) \cdot \varepsilon \cdot l_{123}}{F}$$

   $$= \frac{4 \cdot (n-1) \cdot \varepsilon^2 \cdot b_{13}}{F \cdot b_{12}} \,. \qquad \text{(Eq. 17)}$$

With (Eq. 17), (Eq. 18) the probability of an unsolvable ambiguity for this camera arrangement becomes

$$P_a = P_{a(1)} + P_{a(2)}$$

$$= \frac{4 \cdot (n-1) \cdot \varepsilon^2 \cdot b_{13}^2}{F \cdot b_{12} \cdot (b_{13} - b_{12})} \,, \qquad \text{(Eq. 18)}$$

and the number of remaining unsolvable ambiguities is

$$N_a = \frac{4 \cdot (n^2 - n) \cdot \varepsilon^2 \cdot b_{13}^2}{F \cdot b_{12} \cdot (b_{13} - b_{12})} \,. \qquad \text{(Eq. 19)}$$

If n, ε and b_{13} are given by the the number of targets, the image quality and the requirements of depth accuracy, the optimum choice of b_{12} can be calculated; for $P_a \rightarrow min$ the derivative $(\partial P_a) / (\partial b_{12})$ has to be zero:

$$\frac{\partial P_a}{\partial b_{12}} \overset{!}{=} 0$$

$$\Rightarrow \quad \frac{4 \cdot (n-1) \cdot \varepsilon^2 \cdot b_{13}}{F} \cdot \left(\frac{1}{(b_{13} - b_{12})^2} - \frac{1}{b_{12}^2} \right) = 0$$

$$\Rightarrow \quad \underline{b_{12} = b_{13}/2} \qquad \text{(Eq. 20)}$$

This shows that the ideal camera arrangement of three collinear cameras is a symmetric arrangement with $b_{12} = b_{23} = b_{13}/2$. Like the method of intersection of epipolar lines the length of the epipolar lines does not have an influence on the number of ambiguities. The efficiency of the method is almost as good as the method of intersection of epipolar lines (see table 1).

2.3 Comparison of the methods

The expectable numbers of remaining ambiguities for the methods discussed above are compiled in table 1 for realistic assumptions for the number of particles (n), the depth range in object space (ΔZ) and the width of the epipolar search area (ε) for a base $b_{13} = 200$ mm and a camera constant $c = 9$ mm:

Table 1: numbers of remaining ambiguities

Number of cameras arrangement	2	3 coll.	3 triang.
parameters			
n = 1000, ε = 10 µm, ΔZ = 40 mm	401	40	35
n = 2000, ε = 10 µm, ΔZ = 40 mm	1605	160	140
n = 1000, ε = 5 µm, ΔZ = 40 mm	201	10	9
n = 1000, ε = 10 µm, ΔZ = 80 mm	802	40	35

With two cameras the expectable numbers of unsolvable (but detectable) ambiguities becomes that large that the method itself becomes questionable. The geometric constraint of a third camera leads to a reduction of the numbers of ambiguities by at least one order of magnitude.

If the number of remaining ambiguities is still considered too large, a further reduction is possible in a straightforward manner by employing a fourth camera and either

106

arranging the projective centers in a square (-> intersection of epipolar lines) or on a line (-> double verification of possible matches). Such an arrangement will lead to a reduction factor of at least 100 and almost press the number of remaining ambiguities against zero. An extension to an arbitrary number of cameras is also possible but will rarely be necessary.
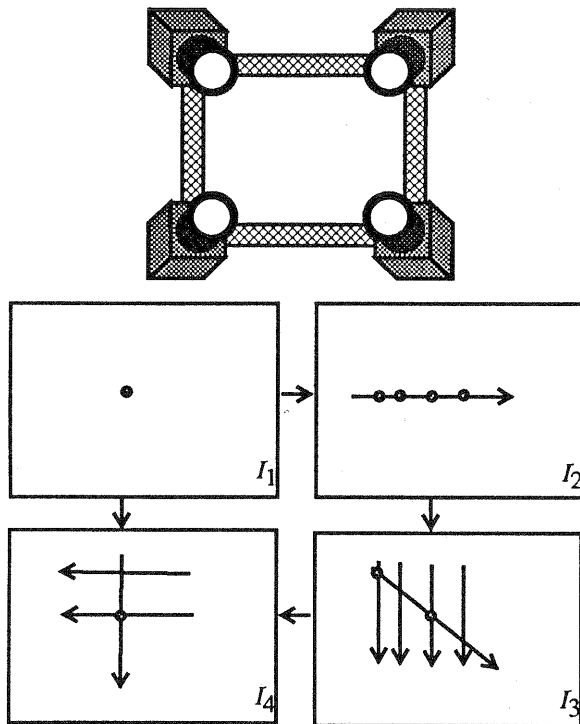

Figure 11: Intersection of epipolar lines in a four-camera arrangement

All the above considerations are only valid for targets randomly distributed in space without a continuous surface. Not randomly distributed targets, e.g. regular dot patterns projected onto a surface to generate an artificial texture (Maas, 1991) may lead to no overlapping targets but much larger numbers of ambiguities, if the pattern is oriented in a way that it is parallel with the epipolar lines in one or more images.

3. Results

To test the method it has been applied to simulated datasets and in several real experiments under various conditions with good success. In the particle tracking velocimetry experiments a maximum of about 1000 instantaneous velocity vectors could be determined with a three camera setup. To achieve a much higher yield seems to be difficult with current CCD-sensor resolution mainly due to image quality and because the number of overlapping particles becomes too large. A two camera system could only give reliable results if the number of particles in the test section and the depth range (i.e. the thickness of the illuminated layer in the water) were strictly controlled. A sample result of particle tracking velocimetry with three cameras is shown in Figure 12

A much higher spatial resolution was achieved when problems with overlapping targets or with ambiguities in tracking could be avoided; in an application of surface measurement with a regular dot pattern projected on a surface of an industrial object which did not show any natural texture (Maas, 1991) it was possible to establish correspondences between more than 5000 discrete points per image of 720 x 574 pixels.
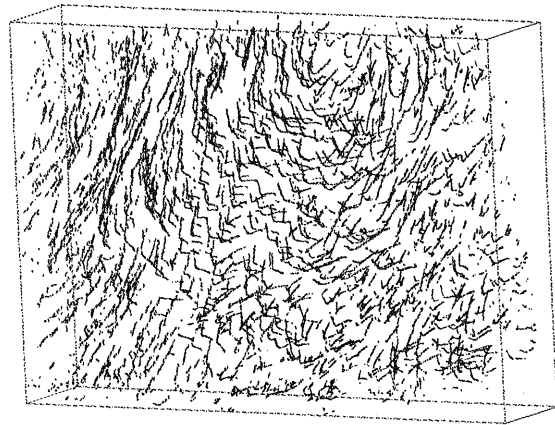

Figure 12: Example results (0.5 sec. flow data in a stirred aquarium)

References:

1. Adrian, R., 1991: Particle-Imaging Techniques for Experimantal Fluid Mechanics. Annual Review of Fluid Mechanics, Vol. 23

2. Kearney, J.K., 1991: Trinocular Correspondence for Particles and Streaks. Dept. of Computer Science, The University of Iowa, Technical Report 91-01

3. Lotz, R., Fröschle, E., 1990: 3D-Vision mittels Stereobildauswertung bei Videobildraten. 12. DAGM-Symposium Mustererkennung, Informatik Fachberichte 254, Springer Verlag

4. Maas, H.-G., 1990: Digital Photogrammetry for Determination of Tracer Particle Coordinates in Turbulent Flow Research. ISPRS Com. V Symposium "Close Range Photogrammetry Meets Machine Vision", 3.-7. September 1990, Zurich, Switzerland, published in SPIE Proceedings Series Vol 1395, Part 1

5. Maas, H.-G., 1991: Automated Surface Reconstruction with Structured Light. Int. Conference on Industrial Vision Metrology, Winnipeg, July 11-12, SPIE Proceedings Series Vol. 1526

6. Maas, H.-G., 1992: Complexity analysis for the determination of dense spatial target fields. 2nd International Workshop on Robust Computer Vision, March 9-12, Bonn, Germany

7. Papantoniou, D., Maas, H.-G., 1990: Recent Advances in 3-D Particle Tracking Velocimetry. Proceedings 5th International Symposium on the Application of Laser Techniques in Fluid Mechanics, Lisbon, July 9-12