

# **Epipolar line extraction using feature matching**

**Ovidiu Ghita**

Vision System Laboratory, School of Electronic Engineering,  
Dublin City University, Dublin 9, Ireland  
ghitao@eeng.dcu.ie

**John Mallon**

Vision System Laboratory, School of Electronic Engineering,  
Dublin City University, Dublin 9, Ireland  
john.mallon2@mail.dcu.ie

**Paul F. Whelan**

Vision System Laboratory, School of Electronic Engineering,  
Dublin City University, Dublin 9, Ireland  
whelanp@eeng.dcu.ie

## **Abstract**

Knowledge of the epipolar lines is of utmost importance in stereo vision since they describe the geometrical relationship between the world points and their projections on the imaging sensors. In this paper we propose a two-phase method to solve the epipolar geometry. Firstly, we extract the corners from a stereo pair and their correspondences are determined by imposing a number of matching constraints. The second phase refines the initial matching results by detecting and eliminating the incorrect matching decisions. The algorithm has been developed for a stereo rig which is used in conjunction with a mobile robot and experimental results are reported.

**Keywords:** Stereo vision, epipolar geometry, corner detection, feature matching.

# Epipolar line extraction using feature matching

## Abstract

Knowledge of the epipolar lines is of utmost importance in stereo vision since they describe the geometrical relationship between the world points and their projections on the imaging sensors. In this paper we propose a two-phase method to solve the epipolar geometry. Firstly, we extract the corners from a stereo pair and their correspondences are determined by imposing a number of matching constraints. The second phase refines the initial matching results by detecting and eliminating the incorrect matching decisions. The algorithm has been developed for a stereo rig which is used in conjunction with a mobile robot and experimental results are reported.

**Keywords:** Stereo vision, epipolar geometry, corner detection, feature matching.

## 1 Introduction

Estimating which points in one image correspond to which points in the other image of a stereo pair is the central problem in stereo vision [Okutomi and Kanade, 1993][Dhond and Aggarwal, 1989]. This problem is commonly referred to as stereo correspondence and is usually solved assuming the stereo cameras are fixed and their relative positions are known [Grimson, 1985]. Prior to solving the correspondence problem, the stereo images have to be *rectified* [Faugeras, 1993] in order to align the epipolar lines to the line joining the optical centres of the stereo cameras. Conventionally to avoid image rectification and to simplify camera calibration and stereo matching, the stereo cameras are arranged in a parallel configuration [Woods et al, 1993]. Thus, such a stereo rig generates horizontal epipolar lines and as a consequence the stereo search space is reduced to one dimension along the horizontal axis.

When the stereo vision system is used in conjunction with a vehicle or a mobile robot, due to various terrain conditions the orientation of the epipolar lines may vary considerably from the horizontal axis. As a consequence the complexity of the stereo matching increases significantly. To compensate for this problem, it is necessary to extract a set of *robust* epipolar lines which give an indication regarding the spatial position of the stereo cameras. This information can be either used directly when the stereo matching is carried out along these lines or used in the image rectification process.

Extracting the epipolar lines robustly is a difficult problem since this requires a set of stable points that cover all regions in the stereo images. Traditionally, this problem is addressed by estimating the *fundamental matrix* [Luong and Faugeras, 1996] which describes the projective transformation between the points contained in a image pair. Consequently, the epipolar geometry can be completely determined once the fundamental matrix is computed. One of the main problems associated with this approach is the fact that the fundamental matrix is very sensitive to error in the point location [Gracias and Santos, 1997]. In this paper we tackle this problem by determining the epipolar lines from a set of points of interest such as corners. The use of corners for such purposes proved to be quite adequate because they can be robustly extracted in both images. The epipolar lines are detected by solving the correspondence between the corners in the stereo images, while mismatches are discarded by applying several matching consistency constraints.

The paper is organised as follows. In Section 2 the main ideas related to epipolar geometry are revised. Section 3 describes the algorithm for epipolar line recovery using feature matching

and in Section 4 several experimental results are presented and discussed. Finally, Section 5 concludes this paper.

## 2 Epipolar geometry

Let us consider a scene point  $P(x,y,z)$  that is visible from both stereo cameras. Given its projection on the image plane associated with the left image sensor, its corresponding projection on the right image plane has to be situated on the epipolar line (epipolar constraint). The epipolar line is the intersection between the epipolar plane (defined by the optical centres of the stereo cameras  $C_l$  and  $C_r$  and the scene point under investigation  $P$ ) and the image plane. As can be observed in Figure 1 where the epipolar geometry is illustrated, the epipolar lines define the relative orientation and position of the scene points in the stereo images. The knowledge of the cameras parameters and their spatial orientations allow us to determine more easily the projections on the image planes of the world points. Thus, recovering the epipolar lines is one important step towards solving the stereo correspondence problem.

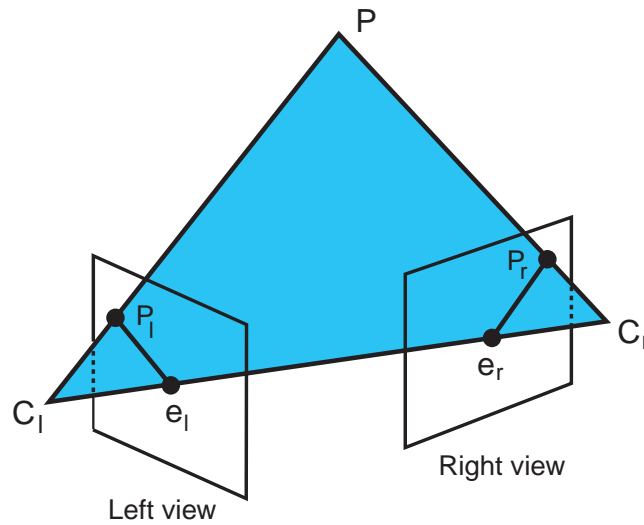


Figure 1. The epipolar geometry. The epipolar lines  $\{P_l, e_l\}$  and  $\{P_r, e_r\}$  are defined by the intersection of the left and right image planes with the epipolar plane  $(C_l, C_r, P)$ . The points  $e_l$  and  $e_r$  are the epipoles and are defined by the intersection of the image planes with the baseline  $\{C_l, C_r\}$ . Note that the epipolar lines are parallel when the stereo cameras are arranged in a parallel configuration.

In projective geometry, the relationship between the co-ordinates of two corresponding points in two stereo images are defined as follows:

$$u_i^T F u_i' = 0 \quad (1)$$

where  $u_i, u_i'$  are the projective co-ordinates of the matched scene points and  $F$  is the *fundamental matrix*. This matrix is 3 x 3 dimensional and in order to determine its elements at least 8 corresponding points between the stereo images are required [Torr and Murray, 1997]. Each corresponding pair of points yields one equation and Equation 1 can be re-written as:

$$a^T f = 0 \quad (2)$$

where  $a$  is a vector containing the co-ordinates of the corresponding points and  $f$  are the elements of  $F$ . Although this approach is common to solve the epipolar geometry, its applicability is restricted to cases when the corresponding points are located precisely. However in real-world applications, very often a stereo rig comprises two cameras arranged in a parallel fashion. Thus, even when the stereo vision system is not fixed, the resulting image distortions are reduced to a planar rotation about the  $z$  axis. Hence, the epipolar lines can be recovered using feature matching.

### 3 Feature matching

In general, the feature matching is a ill-posed problem since for a given point in one stereo image, due to occlusions or missing parts, a corresponding point in the other image may not exist. Furthermore, if the feature points are not associated with prominent textural information or the scene to be analysed reveals repetitive patterns, the matching algorithm may return more than one possible match. Thus, the selection of stable feature points from a stereo pair is a key problem in robust epipolar line recovery. To address this problem two questions have to be answered. The first is what primitives should be employed for matching, while the second question relates to the strategy that has to be employed in order to achieve robust feature matching. The next sections will address these issues in detail.

#### 3.1 Detection of features of interest

The aim of feature (or primitive) extraction is to reduce the large number of pixels contained in an image into a list of primitives that are distinct with respect to their neighbourhood. This has a major impact on the feature matching algorithm since only these lists have to be analysed in order to extract the epipolar lines. In order to be useful, features should be stable enough to be robustly recovered in each stereo image, immune to noise and invariant to geometric transformations. Features may include local primitives such as corners, edges or global primitives when they consist of polygons or image structures [Heipke, 1996]. In general global features are very sparse and difficult to extract and in addition they tend to be more application dependent. As opposed to global features, the local primitives are more general and are very suitable for use when no *a priori* information is available. Therefore, the local features are ideal to be applied to epipolar line recovery. Among local features, corners define the areas in the image where the image intensity changes abruptly, thus they are well distinguished from their neighbouring points. In the next section the SUSAN corner detector [Smith, 1992] will be briefly introduced.

#### 3.2 The SUSAN corner detector

The principle behind the SUSAN corner detector is as follows: it starts by taking a small circular mask (5 x 5 with 3 pixels added at each side) and moves its centre sequentially across every pixel in the image and evaluates the local information at each step. The intensity level of the pixel at the centre of the mask (or nucleus) is evaluated and all surrounding pixels within the mask with similar values in agreement with a threshold value  $t$  are grouped together into the SUSAN area. This is implemented by the relationship illustrated in Equation 3.

$$c(r, r_0) = 100e^{-\left(\frac{I(r) - I(r_0)}{t}\right)^6} \quad (3)$$

where  $r_0$  is the position of the nucleus,  $r$  is the position of any other pixel within the mask,  $I(r)$  is the intensity level at position  $r$ ,  $t$  is the intensity difference threshold and  $c(r_0, r)$  is the output of the comparison. For each pixel in the mask a running total is computed as follows:

$$n(r, r_0) = \sum_r c(r, r_0) \quad (4)$$

Then,  $n$  is compared with a geometric threshold  $g$  which is set at  $\frac{n_{max}}{2}$ , where  $n_{max}$  is the maximum value  $n$  can take ( $n_{max} = 3700$  since the mask contains 37 pixels and 100 is the maximum value of  $c$ ). It is important to note that the intensity difference threshold selects the minimum contrast of the corners which will be detected, while the geometric threshold controls the strength of the corners. In this way by lowering  $g$  only the sharp corners will be detected. The next step involves non-maximum suppression in order to remove false responses. This is carried out by imposing the condition that all the pixels between the centre of gravity of USAN and the centre of the mask must belong to the USAN.

Choosing the right set of threshold parameters is a difficult problem. Ideally, the corners detected should be stable and not grouped in a particular region of the image. For this implementation we selected the intensity difference threshold  $t = 20$  while the geometric threshold  $g$  is set at the default value. This set of parameters typically gives 400 corners (8-bit 512 x 512 images) when the corner detector is applied to indoor scenes. This number is sufficient for epipolar line recovery and assures that corners are detected from all regions in the image.

### 3.3 Corner correspondence

The common approach to match corners in a stereo pair is to take a small region of pixels surrounding the corner to be matched in the first stereo image and compare this with a similar window around each of the potential matching corners in the other image [Smith et al, 1998]. Each comparison yields a score related to the measure of similarity between the corners under examination, the highest matching score representing the best match. The measure of similarity can be evaluated in several ways, but very often for this purpose the *sum of squared differences* (SSD) function [Fusiello et al, 2000] and the *cross correlation function* (CCF) [Smith et al, 1998] are employed.

$$E_{SSD} = \sum_{x=0}^M \sum_{y=1}^N (I_l(x, y) - I_r(x, y))^2 \quad (5)$$

$$E_{CCF} = \frac{\sum_{x=0}^M \sum_{y=0}^N [I_l(x, y) - \bar{I}_l(x, y)][I_r(x, y) - \bar{I}_r(x, y)]}{\sqrt{\sum_{x=0}^M \sum_{y=0}^N (I_l(x, y) - \bar{I}_l(x, y))^2 \sum_{x=0}^M \sum_{y=0}^N (I_r(x, y) - \bar{I}_r(x, y))^2}} \quad (6)$$

where  $E$  is the output of evaluation,  $M, N$  are the window dimensions,  $I_l(x, y)$  and  $I_r(x, y)$  are the stereo intensity images,  $\bar{I}_l(x, y)$  and  $\bar{I}_r(x, y)$  are the average intensity levels of the pixels within the window in the left and right images. Equation 5 illustrates the relationship that implements the SSD function and indicates that the smaller it's resulting value the better the match between the features. Conversely, in the case of CCF (see Equation 6) the best match is represented by the maximum returned value.

### 3.3.1 Algorithm

The corner correspondence algorithm consists of several stages. Initially, for each corner in the right image a list of possible corresponding corners in the left image are selected in agreement with the disparity constraint of the half image width. In other words, the  $x$  co-ordinate of the potential corners in the left image has to be higher than the  $x$  co-ordinate of the corner to be matched in the right image. Also, in order to increase the matching accuracy and reduce ambiguity, a maximum disparity constraint was also imposed. Then, the list of potential corners is ordered in agreement with a similarity measure (SSD or CCF). The corner that returns the highest score is a valid match if it respects the following constraints:

1. It has to respect the confidence constraint which states that the corner's matching score  $e_n$  has to be consistent with respect to a confidence threshold  $\delta$ , which is dependent on the metric used to evaluate the measure of similarity.
2. It has to fulfil the reliability constraint when compared to the next best match:  $e_n - e_{n-1} < \kappa$ , where  $\kappa$  is the reliability threshold.

These matching constraints were employed in order to discard weak matches. Nevertheless the application of these constraints will reduce the number of corners matched in the stereo images, but at the same time the number of incorrect matches are drastically reduced.

### 3.3.2 Refining the matching process. Outlier detection

Although the matching constraints discussed in the previous section significantly reduce the number of outliers, errors still occur when the corners are not distinct or the scene reveals repetitive patterns. However the number of outliers is significantly smaller than the number of inliers, therefore they can be removed by analysing the data set distribution. To achieve this we need to analyse the direction of the epipolar lines derived from mismatched corners since they have a different orientation compared to epipolar lines that are detected correctly. In this regard, the orientation of each epipolar line is detected and the average orientation is calculated. If the difference between the orientation of the epipolar line and the average orientation is higher than an orientation threshold  $\theta$  (for this implementation  $\theta$  was set at  $5^\circ$ ), the epipolar line under investigation is derived from a mismatched corner and consequently is discarded.

## 4 Experiments and results

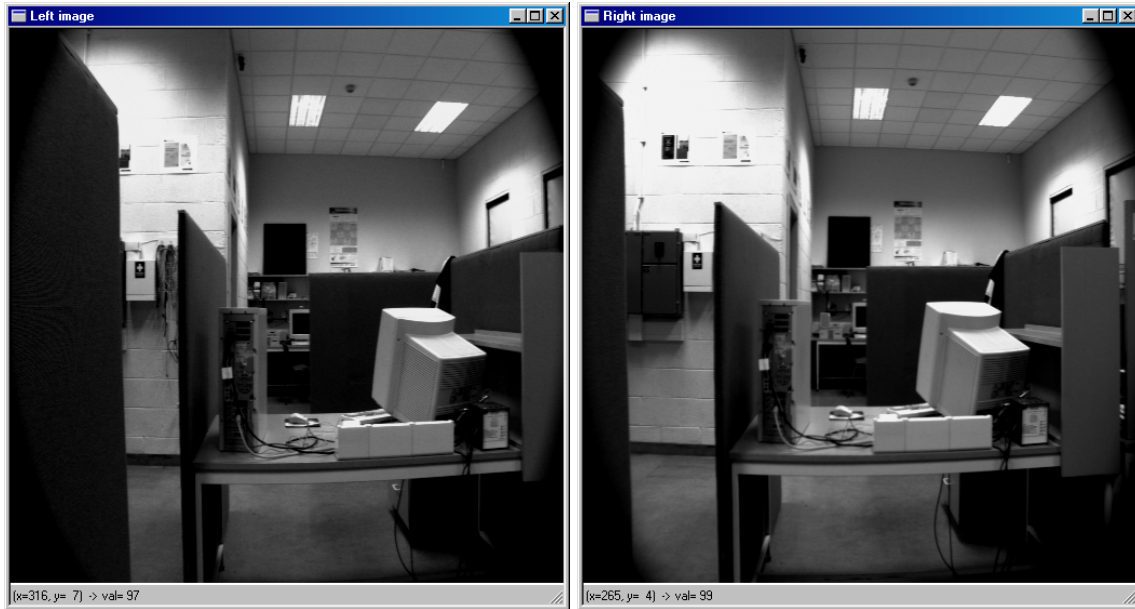
The experiments were conducted on stereo images (8-bit 512 x 512) defined by laboratory scenes. The stereo images were captured using two Sony X77 cameras fitted with 6 mm Computar lenses and digitised by two ITI frame grabbers. The baseline between cameras was set at 20 cm.

A special emphasis was placed on selecting the metric used to evaluate the similarity between the corners in the stereo images. In our experiments we evaluate the performance of the feature matching algorithm using a 5 x 5 window.

Metric	Corners not matched	Corners matched	Corners mismatched
SSD	33.56%	57.32%	9.12%
CCF	32.15%	59.09%	8.76%

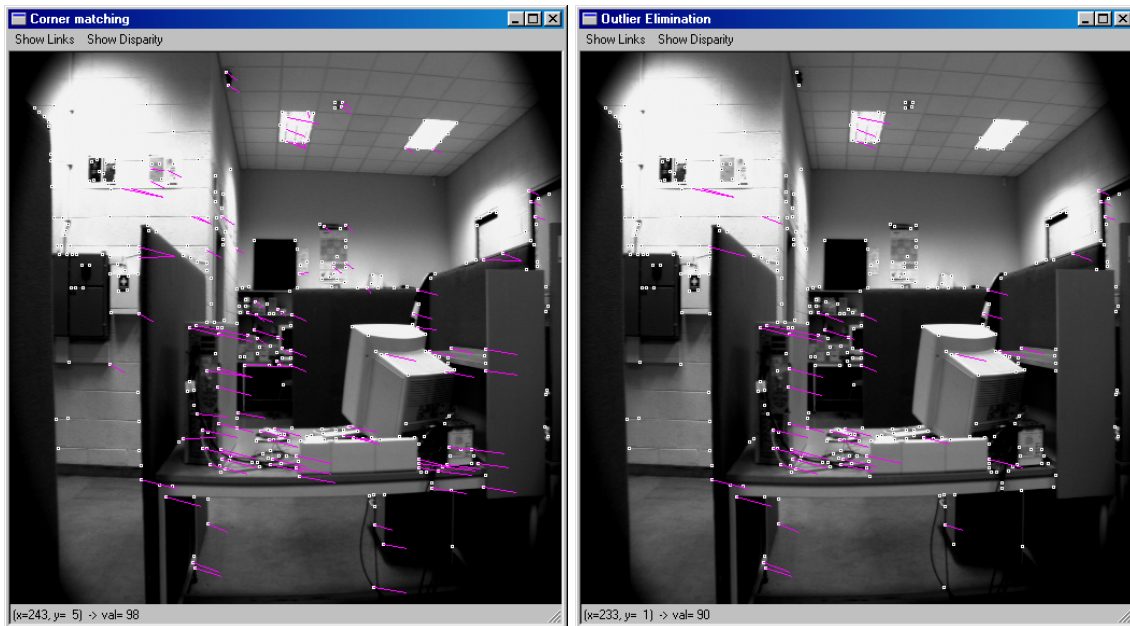
Table 1. Performance of the corner matching algorithm. These results represent the mean of a sample set of 10 different orientation of the stereo rig where the rotational range is within  $-20^\circ$  to  $20^\circ$ .

In both cases the algorithm successfully detects the correspondences between the corners in the stereo views. The CCF performed slightly better than SSD when the stereo images contain regions with high specularity. Quantitative results are depicted in Table 1.



(a)

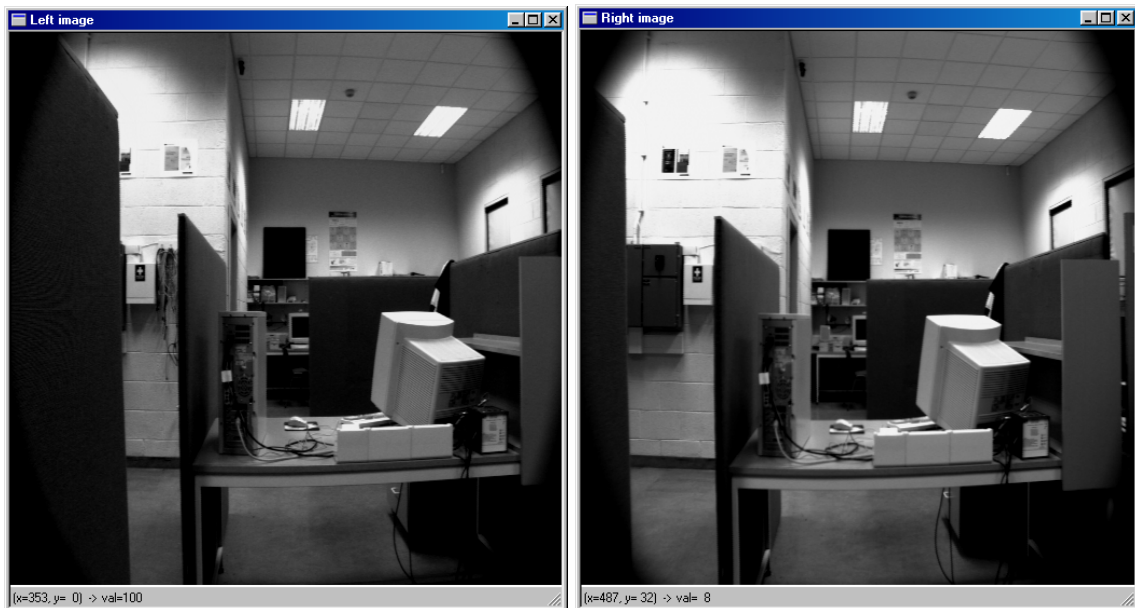
(b)



(c)

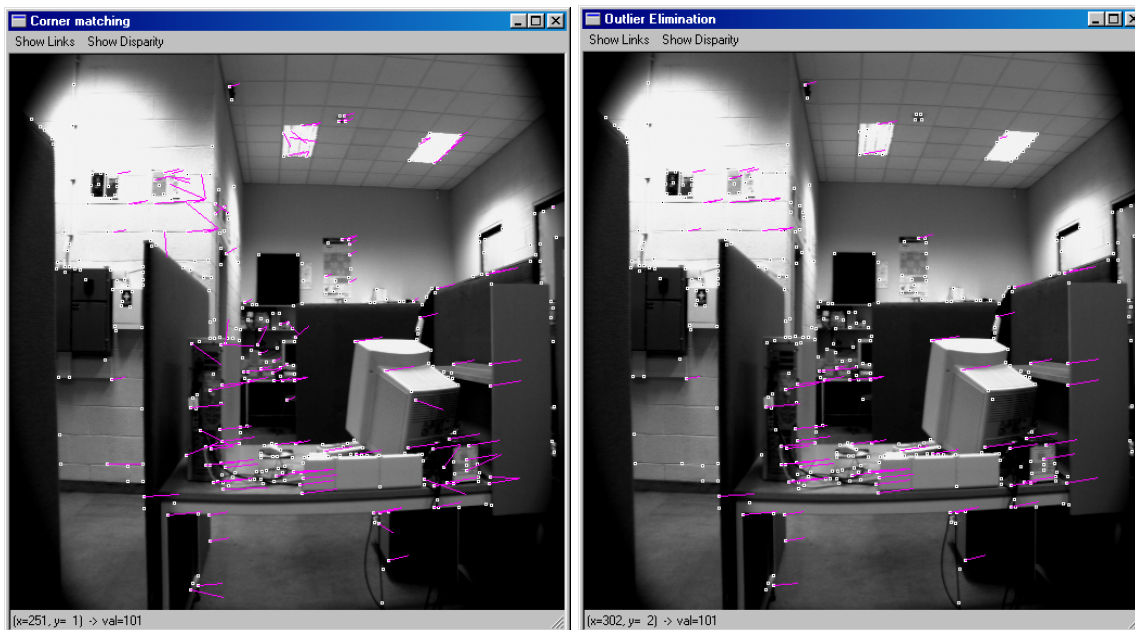
(d)

Figure 2. The epipolar line recovery. (a) The left view. (b) The right view. (c) Corner matching. (d) Outlier detection and elimination. Note that the aperture of the lenses cause circular dark regions on the sides of the images.



(a)

(b)



(c)

(d)

Figure 3. Additional results for a different orientation. (a) The left view. (b) The right view. (c) Corner matching. (d) Outlier detection and elimination.

Figures 2 and 3 illustrate the recovery of the epipolar lines for two orientations of the stereo head. In Figures 2(c) and 3(c) the corners are shown superimposed over the images. The lines indicate the corners matched in both views, while the magnitude represents the disparity. It can be



observed that incorrectly detected epipolar lines derived from mismatched corners are successfully identified and removed (refer to Figure 2(d) and 3(d)). At the same time some correctly matched epipolar lines associated with distant scene points are incorrectly removed due to the warping effect of the 6 mm lenses. This is corroborated with the fact that for distant points in the scene the epipolar lines become shorter and their orientation is detected with less precision. However, this issue has a little impact on the overall results since a large number of epipolar lines are correctly detected.

In the future, we intend to increase the level of initial matches by employing a corner matching strategy that uses a variable window. The warping problems will be tackled using an outlier detection scheme that evaluates the local matching consistency.

## 5 Conclusion

Key to robust epipolar line recovery is the selection of distinct stable features so they are detected consistently in each stereo image. Among other features, corners are very well suited for this purpose since they are well distinguished from their neighbouring pixels and insensitive to geometric transformations. The epipolar lines are recovered by solving the corner correspondence in the stereo images. In order to reduce the matching ambiguity and subsequently the incorrect matches, several constraints such as the maximum disparity and the half image width were imposed. Detecting and removing the outliers is another step towards improving the corner matching consistency. The experimental data indicates that this algorithm returns reliable epipolar line recovery and the present framework is intended to be used in the development of a real-time stereo range sensor.

## 6 References

- [Dhond and Aggarwal, 1989] Dhond U.R. and Aggarwal J.K. (1989). Structure from stereo - a review. *IEEE Transactions on Systems Man and Cybernetics*, 19(6):1489-1509.
- [Faugeras, 1993] Faugeras O.D. (1993). *Three-dimensional computer vision*, MIT Press.
- [Fusiello et al, 2000] Fusiello A., Vito R. and Trucco E. (2000). Symmetric stereo with multiple windowing. *International Journal of Pattern Recognition and Artificial Intelligence*, 14(8):1053-1066.
- [Gracias and Santos, 1997] Gracias N. and Santos J.V. (1997). Robust estimation of the fundamental matrix and stereo correspondences. *5th International Symposium on Intelligent Robotic Systems*, Stockholm, Sweden.
- [Grimson, 1985] Grimson W.E.L. (1985). Computational experiments with a feature based stereo algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(1):17-34.
- [Heipke, 1996] heipke C. (1996) Overview of image matching techniques. *OEEPE - Workshop on the Application of Digital Photogrammetric Workstations*, no. 33:173-189.
- [Luong and Faugeras, 1996] Luong Q.T. and Faugeras O.D. (1996). The fundamental matrix: theory, algorithms and stability analysis. *International Journal of Computer Vision*, 17(1):43-76.
- [Okutomi and Kande, 1993] Okutomi M. and Kanade T. (1993). A multiple-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(4):353-363.
- [Smith et al, 1998] Smith P., Sinclair D., Cippola R. and Wood K. (1998). Effective corner matching. *Proceedings of the British Machine Vision Conference*, Southampton, UK.
- [Smith, 1992] Smith S.M. (1992). *Feature based image sequence understanding*. PhD thesis, Robotics Research Group, Oxford University.

- [Torr and Murray, 1997] Torr P.H.S. and Murray D.W. (1997). The development and comparison of robust methods for estimating the fundamental matrix. *International Journal of Computer Vision*, 24(3):271-300.
- [Woods et al, 1993] Woods A., Docherty T. and Koch R. (1993). Image distortions in stereoscopic video systems. *Proceedings of the SPIE*, vol. 1915, San Jose, CA, USA.