

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Dynamic 3D-Vision

K.-D. Kuhnert , M. Langer, M. Stommel and A. Kolb
*University Siegen
 Germany*

1. Introduction

Measuring and recognising the surfaces of the surrounding world forms a ubiquitous problem in automation and robotics. The knowledge of the environment allows a flexible and autonomous behaviour in different situations. Stereo vision belongs to the most popular techniques for gathering this information because it provides the dense depth information necessary for complex grasping tasks. Compared to laser scanners stereo cameras also have the advantage of higher framerates, so they are widely used for mobile robots. Porta (2005) e.g. uses the Small Vision stereo system (Konolige, 1997) to enhance localisation of a mobile robot: Features from depth maps are used additionally to appearance based intensity features. Other examples include Zhu et al. (2004) or Kang et al. (1995). However, the main problems of stereo vision remain speed and robustness. In order to accelerate the time-consuming registration of the stereo images and avoid specialised hardware, Sun (2002) employs an intelligent subregioning mechanism which reduces the search space of the correspondence analysis. Another approach builds upon the usage of modern SIMD processor instructions as documented by Sunyoto et al. (2004). Kim et al. (2005) on the other hand achieve real-time behaviour by segmenting foreground objects from the background. Depth information is then only updated for moving objects. Of course this approach is problematic for mobile robots. The lack of robustness of stereo analysis for particular scenes mainly arises from depth discontinuities and ambiguous surface texture. Kang et al. (1995) avoid these ambiguities by projecting textured light on the scene, but this is no general solution. Kim et al. (2005) made experiments with an adaptive matching window to increase the accuracy near edges. Zhao and Katupitiya (2006) examined the effect of occlusion and developed a method that detects occlusion areas and adapts a matching window appropriately. To evaluate and compare the robustness of different stereo algorithms, Scharstein et al. (2001) propose a taxonomy for different stereo algorithms and create a testbed including stereo images with groundtruth. Using this testbed we will document the results of the software system for the computation of dense disparity maps presented here. Our stereo system unites some of the speed optimisations mentioned above and hence achieves real-time behaviour. The calibration procedure and some comments on the brightness change constraint will be given. We will also present results for the distance measurements with a PMD camera ("Photonic Mixer Device", Schwarte (2001), Kraft et al. (2004)) which is a technique for measuring the distance of an object by the time of flight of an active infrared illumination. The calibration procedure and the specifics of the measurements will be described, especially for scenes with surfaces almost in parallel to the

optical axis of the system. Finally the results on combining the stereo and PMD technique will be given, discussing the advantages and disadvantages.

2. Stereo

Depth measurement by stereo vision algorithm is done by computing a so called (sparse or dense) *disparity map*. Disparity maps encode the depth for each reference pixel in the stereo images. Using the well-known stereo geometry formula 12 and the disparity map, one can easily calculate the depth value of any given pixel. The calculation of disparity maps leads to the so called *stereo correspondence problem*.

The stereo correspondence problem can be formulated as the problem to efficiently traverse a two-dimensional search space consisting of any intensity value at any position in the stereo images. Algorithms finding corresponding pixel pairs in acceptable time will be presented later.

When one pixel of one stereo image is compared with one pixel of the other stereo image, the degree of correspondence between these pixels is calculated using a certain metric. Two of the most prominent metrics are the *sum of absolute differences (SAD)* and the *sum of squared differences (SSD)*. Both metrics have in common that they calculate the correspondence value over a certain block size. This block size is usually represented by a rectangular region pixel by pixel around the pixels that are to be matched. SAD sums up both rectangular regions and takes the absolute value of the difference of these sums, whereas SSD takes the squared value of the difference of these sums. Both functions can be efficiently computed using modern SIMD (Single Instruction, Multiple Data) processor instruction sets.

2.1 Review

In this part we review various stereo correspondence algorithms for the traversing of the two-dimensional search space mentioned earlier. First these algorithms are briefly described, then concrete implementations of these algorithms are discussed.

All of the reviewed algorithms belong to the class of *block-matching* methods along a *horizontal scanline*. Therefore, the images must be - at least approximately - rectified if the optical axes of the cameras are not adjusted in parallel. Due to the rectification and the epipolar constraint (see Faugeras, 1993), it can be presumed that corresponding pixels in stereo image pairs can be found on the same horizontal lines in both images. Then we define a three-dimensional correspondence candidate matrix (also called *cost matrix*) $C(x,y,d(x,y))$ holding all disparity value candidates d of a given reference pixel at position (x,y) . After building this matrix, one needs to find efficient algorithms for reordering the optimal disparity value for any given reference pixel. Several geometric and object specific constraints reduce the search space and lead to an increase in both speed and quality of the results of the correspondence analysis. A set of the most important constraints is given in the following.

The so called *brightness change constraint* states that if a pixel in one of the stereo images has a corresponding pixel in the other image, the intensity values of these pixels need to be same. Fulfilling this requirement makes a reliable pixel matching even possible, because all of the later discussed algorithms work intensity based. This is actually one of the most important constraints concerning the quality of the correspondence analysis' results.

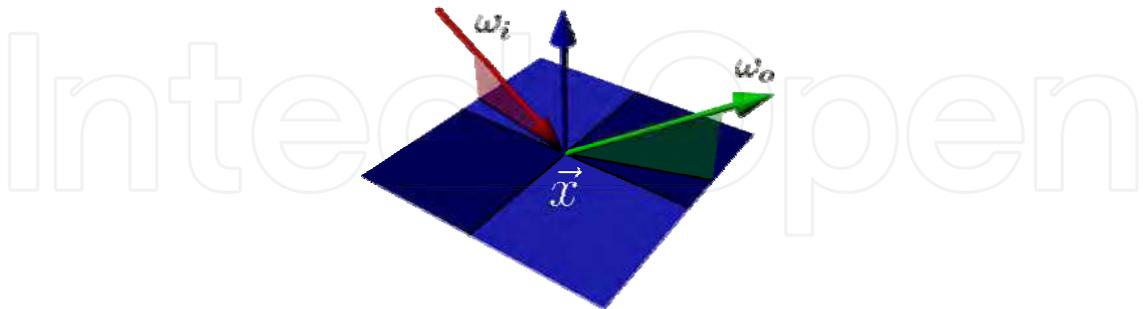


Figure 1. Incoming light ray (marked red) with angle $\omega_i = (\theta_i, \phi_i)$ and outgoing/reflected light ray (marked green) with angle $\omega_o = (\theta_o, \phi_o)$ at point \mathbf{x}

Object surfaces need to be *piecewise smooth*. This prerequisite ensures that corresponding pixel pairs in both images have almost the same disparity values. This enables a further reduction of the search space, because corresponding pixels can be found around a certain offset on a scanline in the analyzed image.

The object surfaces are highly *textured*. The more textured the surfaces are, the more reliably corresponding pixels can be found. If for instance an object surface has only one colour, it is almost impossible to distinguish corresponding pixels from non-corresponding pixels on these surfaces because they look identical.

The *monotony constraint* demands that, if the stereo images are rectified, corresponding pixel pairs are to be searched in the same direction on a scanline. That means fulfilling the monotony constraint, the pixel pairs' occurrence is ordered on the scanlines.

As a last important constraint the objects are supposed to be *equally visible* for both cameras of the stereo-system. That means, objects in a scene are not partially occluded. Partial occlusion leads to different projections of the same object onto the image planes of the stereo-cameras, because the cameras look at the objects from slightly different angles. The consequence of that effect is that the left and the right stereo image possess different, even mutually exclusive information of the object in the scene. Hence errors can be expected when trying to match pixels in one image along a scanline in the other image where the other image actually has no information about these pixels at all.

As being said, the brightness change constraint belongs to the most important constraints for the set of intensity based stereo matching algorithms. The majority of errors occurring in the matching process can be classified as violations of this constraint. The reason why this constraint is so easily violated is mainly due to the angle dependent *reflection properties* of an object as well as the problem of *subpixel edge shift*.

The reflection properties of an object mainly depend on the surface of the object as well as on the light reaching the object surface. These two factors determine the equation of the reflected light intensity

$$L_o(\mathbf{x}, \omega_o) = \int_{\Omega} f(\mathbf{x}, \omega_i \rightarrow \omega_o) L_i(\mathbf{x}, \omega_i) \cos \theta_i d\omega_i, \quad (1)$$

where $\omega_i = (\theta_i, \varphi_i)$ and $\omega_o = (\theta_o, \varphi_o)$ denote the solid angle of incoming and reflected light rays at point \mathbf{x} as shown in Figure 1.

L1 and L2 denote the light intensities of the incoming and outgoing direction of the light rays. The term $f(\mathbf{x}, \omega_i \rightarrow \omega_o)$ represents the so called *bidirectional reflectance distribution function* (BRDF) measuring the physical reflectance behaviour of the surface material. For any light ray with given entry angle Ω hitting the object at point \mathbf{x} , the BRDF yields the quotient of irradiance and emittance of any reflected light ray. The model of the BRDF is based on the concept of so called *micro facets*. Micro facets are microscopically small mirrors randomly aligned and distributed all over the object's surface. The alignment of these mirrors is determined by the probability distribution of eq. 2:

$$f(\mathbf{x}, \omega_i \rightarrow \omega_o) = \frac{1}{\pi} \frac{F \cdot D \cdot G}{\langle \omega_o \circ \mathbf{n} \rangle \langle \omega_i \circ \mathbf{n} \rangle}, \quad (2)$$

F denotes the *fresnel-factor* which models the fraction properties of the object material, D is the *probability distribution function* of the micro facets and G is the *geometry factor* modelling shading between the micro facets. An in-depth look into the complex mathematical deduction of these parameters can be found in related textbooks. To yield the overall light intensity reflected from a point \mathbf{x} , one has to integrate over the entire spatial angle Ω . This complex procedure requires extensive computation time.

The problem of subpixel edge shift results from the inherently limited resolution of cameras. To measure the light intensity of a pixel of recorded scene, the camera has to integrate the light intensity over a certain area of the scene predetermined by the camera resolution. This causes problems on edges in the scene, because edges typically mark an abrupt change in light intensity values. So the integration process averages the light intensities in the given areas over the edges. Due to high angle dependency of the projection of edges (as described earlier) in such a subpixel integration area, the intensity values of corresponding pixels may differ significantly. This problem can be reduced just slightly by using high resolution cameras.

Hence, when performing intensity based stereo correspondence analysis, one always has to consider these inherent problems. There is no general solution to address these problems and the results need to be interpreted respectively. In the following sections some of the most popular approaches of stereo correspondence analysis are presented.

2.1.1 Winner Takes it All (WTA)

One of the simplest algorithms for searching in the matching matrix is the *Winner Takes it All* method. WTA works as follows: For a given reference pixel WTA walks through the cost matrix selecting that pixel which has the lowest difference to the reference pixel. This is a simple minimum search on a given set of numbers. It is also a local method, because the algorithm operates only on *one* vector of the matching matrix for each reference pixel.

The biggest advantage of such a primitive local method is the easy way of its implementation. Furthermore WTA can most easily be optimized. The disadvantages on the other hand are that local methods highly depend on the constraints discussed earlier. There is also the possibility that identical pixels of the cost matrix are assigned to reference pixels more than once.

2.1.2 Global Methods

In contrast to local methods, global methods perceive the stereo correspondence problems as the problem of minimizing a global energy function. The goal is to find a disparity function d which minimizes the global energy function

$$E(d) = E_d(d) + \lambda E_s(d), \quad (3)$$

as it is described in (Scharstein & Szeliski, 2001). The term $E_d(d)$ measures how well the disparity function d matches the stereo pixel pair which is to be evaluated. This is done by eq. 4 which takes all possible corresponding pixel pairs into account. λ is a global constant denoting a weight for $E_s(d)$.

$$E_d(d) = \sum_{(x,y)} C(x, y, d(x, y)) \quad (4)$$

$E_s(d)$ is responsible for introducing the *piecewise smoothness*-constraint.

After stating the global energy function to be optimized, we now show some popular algorithm for its optimization.

2.1.3 Dynamic Programming (DP) with Scanline Optimization (SO)

The optimization of regular functions $E_s(d)$ using naïve approaches is an NP complete problem. The use of a *dynamic programming* approach helps finding the global minimum of mutually independent scanlines in polynomial time.

DP was first introduced by Richard Bellman in 1953. It describes the process of solving complex problems where one has to find the best decisions to solve this problem one after another. The basic idea of DP is to break down the complex problem into smaller subproblems. After solving these subproblems optimally, the original problem can be solved. The subproblems themselves are broken down into smaller subproblems and so forth until the subproblems have a trivial solution. An important step in DP is memorising (also called *memoization*) the solutions of the already solved subproblems. Otherwise one would have to compute the solutions to the same subproblems over and over again as in the simple recursive approach of computing Fibonacci numbers. DP is especially well suited for finding shortest paths in matrices with associated cost values.

In our case the correspondence analysis is treated like an n-dimensional search problem. The matching costs (defined as the SAD or SSD values of the examined pixel pairs) of *every* point of a scanline assemble the search space. Hence this is a global method. In contrast to local methods (like WTA) not just certain pairs of pixels are taken into account, but a two-dimensional matrix of candidate pixels is considered. Such a matrix is constructed of all reference pixels versus their corresponding candidate pixels. To find the best corresponding pixel out of the candidates, one has to find the shortest path with lowest matching costs in the matrix.

We use *x-d-submatrices* of the whole cost matrix $C(x, y, d(x, y))$ as subproblems that are to be optimized. Firstly a certain area in which the path will be searched is preselected. The preselection is done to save computing time and the width of that area may vary. Secondly for every pair of corresponding scanlines in the left and right stereo image, the shortest path through the matrix with least pairwise matching costs is selected.

An extension to the classical dynamic programming approach presented above, is the scanline optimization for which the same prerequisites as discussed above apply. An additional cost-constant is introduced distributing penalties to the candidate pixels, if the values of neighbouring pixels differ too much (like it is the case at object edges). Hence big jumps in disparities are highly penalized disqualifying themselves for the later path search. The decision when to penalize a disparity value is predefined.

Dynamic programming algorithms have the advantage of a higher robustness compared to local methods. This especially holds true if some of the constraints of section 2.1 are not fulfilled. The algorithms exploit the monotony and piecewise-smoothness constraint so that the search space for disparity values can be reduced to pixels located on object surfaces. So it is not possible to select a candidate pixel that is actually located far from the object surface but coincidentally possesses almost an identical intensity value to the reference pixel.

The disadvantage on the other hand is the problem of finding a suitable path in the search space concerning partially occluded pixels. Another drawback is the heavily increased computing time compared to simple local methods.

2.1.4 Simulated Annealing (SA)

Simulated annealing is a heuristically based method in computer science. It is used to solve optimization problems that have a high complexity, which makes going through all combinations to solve the problem computationally infeasible.

SA is inspired by natural annealing processes. For instance, the slow annealing of liquid metal provides enough time for the molecules inside the metal to align them in a way to build a stable crystal structure. That way a low energy state is achieved close to the optimum. Transferred to our optimization problem the temperature corresponds to an acceptance threshold. Below that threshold an intermediate result of the optimization process is still allowed to temporarily worsen.

There are two major algorithms in this field: the *metropolis algorithm* and the *hill climbing algorithm*. The hill climbing algorithm has the ability to leave a local optimum trying to find an even better one. Hence it is the more sophisticated algorithm. A more detailed description of these two algorithms can be found in (Metropolis; Rosenbluth & Teller, 1953). Simulated annealing yields a very high confidence in the stereo correspondence analysis. Alas it has a very bad runtime due to the very high computational complexity. There can be no real-time behaviour expected from algorithms falling into this category.

2.2 Robustness

We compared various stereo correspondence algorithms with regard to the run-time behaviour and the quality of the computed (dense) disparity maps. We concentrated our work around the implementation of Scharstein/Szeliski (Scharstein & Szeliski, 2001) from the Middlebury University. We optimised their original algorithms with a series of steps like smoothing the images first and working on recursively down sampled sub-images to decrease overall computing time as proposed by Sun (2002). To further decrease the computation time, we also used the MMX processor extensions as proposed by Sunyoto et al. (2004). Using sample stereo-images and their corresponding *groundtruth* images from the Middlebury University we were able to calculate the quality of the resulting disparity map counting the incorrect disparity values (Figure 2). We used the following formula to calculate these values:

$$BP = \frac{1}{N} \sum_{(x,y)} (d_C(x,y) - d_T(x,y) > \delta_d) \quad (5)$$

The variable d_C denotes one element of the calculated disparity map, whereas d_T is an element of the groundtruth map at position (x,y) . δ_d is a fault tolerance value. We set it to 1.0.



Figure 2. Left image and groundtruth image of sample stereo image

Matcher	Method		Time [s]	Bad Pixels [%]
Schar/Szel	WTA	SAD	1.04	4.60
		SSD	1.12	4.85
	DP/SO	SAD	4.34	6.46
		SSD	4.35	6.63
	SA	SAD	158.28	4.35
		SSD	147.88	4.54
	building matching-matrix		0.95	N/A
optimized	WTA	SAD	0.25	7.47
Schar/Szel		SSD	0.27	6.14
		DP/SO	SAD	1.09
	SSD		1.05	6.96
	SA	SAD	39.09	5.74
		SSD	37.10	5.14
		building matching-matrix		0.23

Table 1. Performance and quality results of different matching algorithms measured on an AMD Athlon XP 1700+ system with 512MB RAM

The results of the different algorithms for the example pictures shown in Figure 2 are presented in table 1. We used a fixed block size of 16x16 pixels. In the top half of the table one can see the values obtained from the original version of the Scharstein/Szeliski implementation. The lower half shows the values achieved with the optimizations extending this implementation. The second column enlists the different matching algorithms used in conjunction with either SAD or SSD as correspondence measurement. The last two columns

show the run-time behaviour as well as the quality (measured in the rate of falsely computed disparity values) of the implementation computing the disparity maps of the pictures like shown in Figure 2. The pictures are 32 Bit truecolour images with images sizes of 450 x 375 pixels. Figure 3 shows the corresponding disparity map and the difference image of the groundtruth image with the computed stereo image.

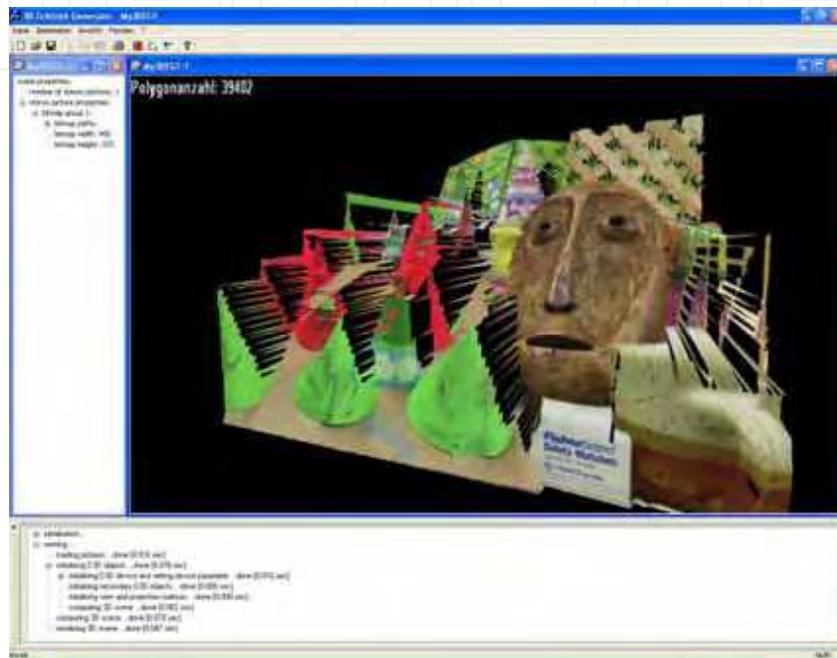


Figure 4. Reconstructed 2.5D-view of the sample scene



Figure 5. A left and a right image of a stereo camera

Table 1 clearly shows that the simple WTA method performs best (in the original *and* optimized version) with regard to run-time behaviour and is only slightly worse than SA with regard to the quality of the computed disparity values. SA on the other hand is about 150 times slower than the simple WTA algorithm and yields to an improvement of just 0.25% - 2.00% regarding disparity quality.

The difference picture of the calculated disparity maps (Figure 3 – right side) shows that errors mainly occur on edges due to partial occlusion. The brightness change constraint in this case is preserved hence leading to no additional errors. It can be expected that under normal conditions the brightness change constraint will be violated and DP/SO may yield better results than the simple WTA, because DP/SO is more robust in this case.

Figure 3. Computed disparity map (using WTA with SSD method) and difference map according to the groundtruth map

We also used other images mainly taken from the (Stereo Vision Research Page, 2007). The results of those pictures resemble the results presented in table 1.

In Figure 4 we present a reconstruction of the original scene using the stereo-images and their corresponding disparity map. This of course cannot be a full 3D-view of the scene because there is only one stereo image pair made from a certain angle. Hence we call this a 2.5D-view of the original scene. The viewer can clearly distinguish which objects were closer to the camera and which were lying in the background.

2.3 Variable Block Size and Interest Operators

Some cameras offer the possibility to automatically control the brightness of the image. Concerning single camera images this is often an improvement. However, for the matching of two images from different cameras this is a clear disadvantage. In a stereo setup the length of the baseline between the cameras causes deviations in the fields of view of the cameras, so only a part of the scene can be seen in both images. The rest of the images is different for the two cameras and influences the brightness control. As a result, also the common image parts have a different brightness. In that case we made good experiences by matching the derivation of the image instead of the pure image intensities.

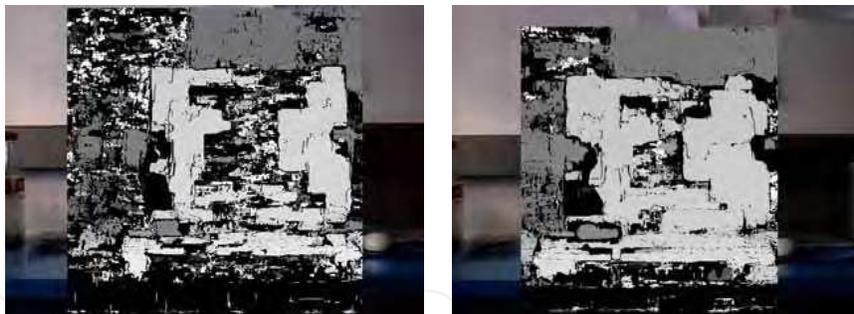


Figure 6. Disparities for a block size of 30x10 pixels and 60x20 pixels

The derivative can also be used as an interest operator, to accelerate matching and in particular to determine the size of the image regions being matched. The following explanations refer to a WTA block-matching algorithm with SSD as similarity measure. The algorithm finds matches by comparing a block from the left image to all positions on the corresponding epipolar line in the right image. Interest operators were introduced by Moravec (1977). Their purpose is to limit image processing operations to the relevant image area and save computational cost. A review on different methods can be found in Bähr and Vögtle (1991). We use the horizontal derivative to find textured image positions in the left image which can be matched robustly to the right image, and to find image positions in the

right image which are candidates for a correspondence. Other positions in the right image are not compared, which heavily reduces the computational effort. Since the robustness of stereo analysis relies on texture, we examined the use of maxima of the derivative as positions for matching. The result is a comparatively sparse disparity map because for a given block in the left image there is only one matching position per edge in the right image. The number of positions depends also on the threshold for maxima detection and often a correct maximum is not recognized because the gradient falls below a given threshold. However, the results are quite robust. To increase the number of correspondences, we replaced the maxima detection by simple thresholds for the gradient magnitude. We use a higher threshold for the left image and a lower threshold for the right image, to make sure we obtain all corresponding positions in the right image despite the image noise. The results are both more dense and more robust.

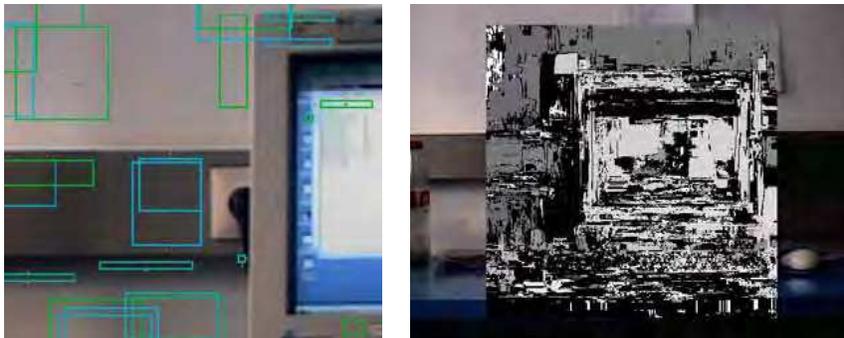


Figure 7. Adaptive block size for the cross method and resulting disparities

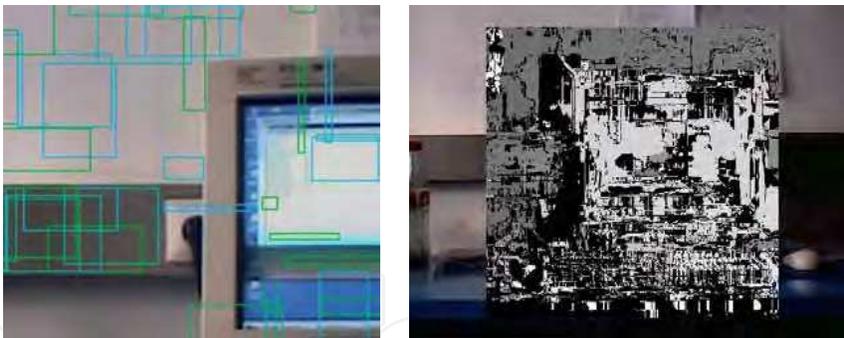


Figure 8. Adaptive block size for the region method and resulting disparities

The stability of the stereo analysis depends also on the chosen block size. Figure 6. shows the disparity maps for the stereo image pair from Figure 5 for different block sizes. Light values indicate a higher disparity, dark areas a lower disparity. Missing values are marked as black. They result from the plausibility test proposed by Faugeras (1993): The matching result in the right image is searched for in the left image. The resulting second match should be the original block in the left image. If this test fails, the disparity value is discarded. For a block size of 30×10 pixels 53 percent of all possible disparity values could be computed. For a block size of 60×20 pixels this value increases to 59 percent. As can be seen from the picture, this is an improvement especially for homogeneous regions. The drawbacks of an

increased block size are the additional computational effort and the lower accuracy, in particular near depth discontinuities.

To overcome these disadvantages, a locally adaptive block size is introduced. In homogeneous regions a higher block size is chosen than in textured regions to make matching more robust if there are few characteristic image features. To reduce the inaccuracy near depth discontinuities, bigger blocks are not allowed to cross sharp edges. For every image position the block size is determined by evaluating the gradient magnitude along the respective image row and column. Starting with an initially small block, the block size is increased horizontally and vertically until the borders of the block reach an edge in the row or column through the centre pixel. If there are no edges the block growth is stopped after a certain maximum size has been reached. The search along a cross was chosen because size determination can be performed very quickly. Figure 7 shows some of the blocks as rectangles. Little dots indicate the edges which stopped the block growth. The right part of Figure 7 shows the resulting disparity map for a maximum block size of 60x60 pixels. In comparison to the disparity map in Figure 6 the results for homogeneous regions have improved. On the other hand, some of the stable but inaccurate disparities near the edges of the monitor are missing now. This is a result of the smaller block size in textured image regions and absolutely correct taking into account the occlusion effects in that area. A drawback of this method is that only edges lying on a cross through the block center have an influence on the block size. The disparity values near the corners of the monitor hence are still inaccurate. To improve that, a second strategy for the size determination is examined. Starting from an initial block size the block is expanded alternating by one row or one column, respectively, if there is no edge in this new row or column. If the block growth is stopped in one dimension, it continues in the other dimension until also there an edge has been reached. The results can be seen in Figure 8. Concerning the corners of the monitor, the disparities are more accurate now, while the remaining values are similar to the results before. For both methods of size determination the gain in density and accuracy of the disparity maps has the disadvantage of a higher computational cost, which is primarily a result of the increased block size for homogeneous regions. The computation of the disparity map by the cross method took three times longer than for a fixed block size of 30x10 pixels. The method, which tested the whole area of a block for edges, was even four times slower. It should also be mentioned that the matching results for homogeneous regions are comparatively unstable even for large blocks. A soft edge at the border of a block caused by a depth change in the scene together with smoothing etc. can dominate the whole structure inside that block and thus influence the resulting disparity value.

2.4 Post Processing: Disparity Histogram and Subpixel Accurate Disparity

Ideally, disparity maps consist of big areas with steadily changing disparities for flat surfaces and abrupt changes for depth discontinuities. In homogeneous regions the results deviate strongly from this ideal. After removing most of these unstable disparities using a suitable interest operator, mid-size areas of homogeneous disparity remain. These areas are surrounded by areas without results. The remaining false matches appear as single disparities deviating much from their neighbourhood. These observations motivate the following assumptions: (a) A correct disparity value belongs to a certain homogeneous surface. It probably appears there multiple times. (b) Errors are rare. (c) Wrong disparities

do not belong to a certain surface and hence take on arbitrary values. The same wrong value appears probably only a few times. That means in reverse: Frequent disparity values are usually right, while rare values are often wrong. For our experiments we computed the histogram of disparity values and used a threshold on the histogram to discard uncommon values. Our experiments lead to good results for thresholds in the range of a few tenth of a percent. With these thresholds sometimes up to 50 percent of the removed values were actually wrong disparities. Of course, the results depend also on the image contents and perhaps a comparison with neighbouring disparity values could lead to further improvements. In general this is a good supporting method if most of the unstable disparities are already filtered out by an interest operator. Then this method discards a high percentage of wrong values at almost no computational cost without removing too many disparity values.

For applications which require a high precision rather than high speed the disparity can be computed with subpixel accuracy. A robust way to determine the subpixel shift between two corresponding block is to minimise the square error

$$e = \sum_x (g(x+s) - f(x))^2 \quad (8)$$

between the pixel intensities g of the block in the left image and the intensities f of the corresponding block right image. The variable x denotes a pixel position inside the block and the variable s denotes a subpixel shift along the epipolar line. The intensity of a block with subpixel shift is linearly interpolated using the image derivative according to the formula

$$g(x+s) = g(x) + sg'(x) . \quad (9)$$

The subpixel shift is then obtained by finding the root of the derivation of the error, i.e. the value of s for which

$$e' = \sum (2sg^2 + 2gg' - 2g'f) = 0 . \quad (10)$$

The subpixel shift results thus to

$$s = \frac{\sum 2g'f - \sum 2gg'}{\sum 2g^2} . \quad (11)$$

In practice, this method leads to smooth subpixel shifts in areas of uniform disparity. Outliers occur only where the disparity value already deviates from the neighbourhood for pixel accuracy. Since at wrong positions the subpixel shift often is greater than half a pixel, the subpixel shift is well suited to indicate wrong disparity values. Since such great values correspond to a neighbouring matching position with a lower subpixel shift, the disparity must already be wrong at pixel level and thus can be discarded.

2.5 Camera Calibration and Accuracy of the Distance Measurements

We use a stereo setup with two *DFK 21F04* cameras by *The Imaging Source* and *Cosmicar/Pentax* lenses to compute the accuracy of the distance measurements. These cameras provide images with a resolution of 640×480 pixels. They are mounted on separate 10mm aluminium plates which can be adjusted in yaw, pitch and roll angle. The baseline

length of the camera setup is 20 cm. The lens aperture is set to 5.6, the focus to infinity. The lens of the right camera has a focal length of 8mm.

To save computation time during operation the stereo images are not rectified in software. Instead we rely on a careful manual adjustment of the camera orientation. By changing the zoom of the lens of the left camera, the image sizes are brought into accordance with pixel accuracy. Because the stereo algorithms we use belong to the category of scanline matching, the roll and pitch angles of the cameras are adjusted in a way that the line correspondence between the left and the right image is maximised. We obtain an error of less than 1/100 pixels for the roll angle and less than 1 pixel for the pitch angle. For reasons of simplicity, stereo systems are often built with parallel optical axes. But for a working distance of 1.5m-4m that was chosen with regard to a later data fusion with the PMD camera, the images had a common field of view too small for stereo analysis. Therefore, the optical axes are directed towards each other, so that both camera images centre an object at a distance of 4m.

Assuming a pinhole camera model, we can compute the distance z of a point in 3D-space by the well-known formula

$$z = \frac{b}{x_r/f - x_l/f}, \quad (12)$$

where f denotes the focal length of both cameras and x_r and x_l denote the horizontal coordinate of the corresponding position in the right and the left camera image. The variable b denotes the length of the baseline. For the proposed stereo setup the formula changes to

$$z = \frac{b}{\tan(\alpha + \tan^{-1}(-x_r u/f_1)) - \tan(\beta + \tan^{-1}(-x_l u/f_2))}. \quad (13)$$

Here, α and β denote the deviation of the yaw angles of the optical axes from a parallel setup. The values f_1 and f_2 are used because the focal length is not necessarily the same for both cameras. The variable u denotes the size of one pixel on the CCD-sensor. It is given by the Sony ICX098BQ data sheet as 5.6 μ m/pixel.

B	0.0024691
α	0.3045662
β	0.3041429
F_1	0.7583599
F_2	0.7580368

Table 2. Camera parameters obtained by the genetic algorithm

To find the parameters b , α , β , f_1 and f_2 , a series of sample images of a flat, highly textured test surface is taken at known distances between 1.6m and 4m. For the stereo analysis we use the "Winner Takes It All" method with the SSD similarity measure based on the implementation by Scharstein et al. (2001). The resulting pixel accurate disparity values ($x_r - x_l$) are averaged over the test surface. Then a genetic algorithm was used to find a good parameter set that minimises the integrated squared error between the distance values obtained by the formula above and the measured distance. A standard deviation of 0.0077 was achieved for the distance values (in meter), when the algorithm was stopped after a sufficient number of cycles. The resulting parameter values are given by table 2. Averaging the disparity values in practice results in the loss of the absolute image position. For every

averaged disparity value the absolute coordinates x_r and x_l can be computed by adding an arbitrary offset to the disparity, keeping in mind that the disparity is valid for a larger image region. Some optimisation methods do not take this into account and thus find parameter sets with a significantly lower standard deviation, but then the distance values are only plausible for the chosen offset.

3. PMD

The advent of the photonic mixer device (PMD) leads to new possibilities in real-time depth measurement. Compared to the use of stereo correspondence analysis, laser scanning or other depth information yielding technology, PMD-cameras have the advantage of recording a scene providing intensity images *and* depth images at once. No further time consuming computation needs to be done. PMD cameras integrate the sensor hardware plus the needed software for gathering the images in one device. Figure 9 shows a PMD camera.



Figure 9. Picture of a PMD camera

3.1 General Operation

The principle of a PMD camera will be briefly described. The camera emits an amplitude modulated light signal, which is reflected back onto the camera sensor array by the surface. The sensor is coupled with the modulation emitter and because of that capable of separating the electrons, which are generated by the reflected photons according to their distance. This process of comparing the optical signal with the electrical reference signal of the emitter is responsible for gathering three-dimensional information of the scene.

The scheme of depth measuring with PMD sensors is shown in Figure 10. The modulation source sends amplitude modulated light (usually at wavelengths of about 800nm) to the object surfaces. The modulation frequency is set to 20MHz, so that a full oscillation of the signal has a length of 15m. This also limits the band of unambiguously yielding depth information to 7.5m (half the oscillation length). Hence beyond 7.5m the camera yields false depth information and should only be used at ranges below that limit.

There are two light penetrable photo gates on the surface of the semiconductor elements, which are set to a voltage equivalent to the light modulation. The incoming photons release electrons in the underlying p-layer of the photo gate. Due to the voltage the semiconductor is set to, a potential gap is induced. This deflects the electrons to one of the reading diodes. Changing the polarity of the voltage the electrons are deflected to the other reading diode. That way we get two capacitors collecting electrons. Unmodulated background light is distributed equally to both capacitors thus eliminating its effect for detection. The modulated light on the other hand together with the push-pull voltage causes correlated

readout signals, which directly correspond to the phase difference. Hence the “mixing” is done by mixing electrons of the incoming signals with the signal of the push-pull voltage. The incoming signal is integrated over its amplitude and the measured result is put into its corresponding pot. The difference of the capacity of both pots is directly related to the phase difference which is linearly related to the distance of the camera to the recorded object.

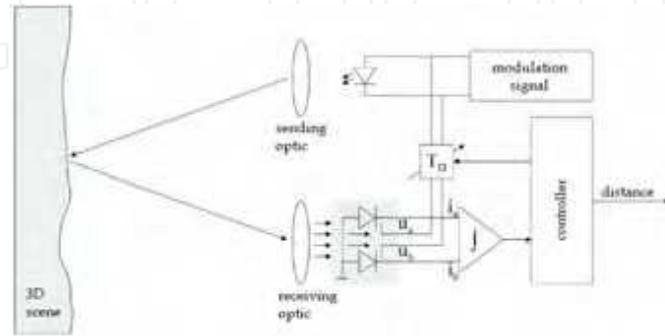


Figure 10. Scheme of a PMD measurement system

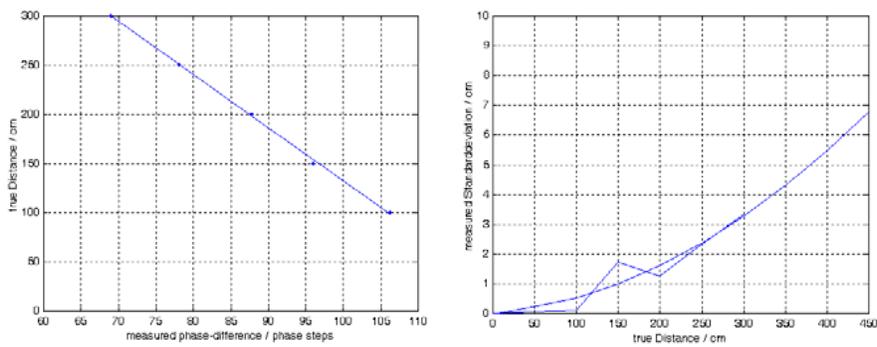


Figure 11. Left: Sample measurements (dots) and linear approximation of the distance function. Right: Standard deviation of the measurements as function of the distance

3.2 Calibration

In our experiments we use a 1K PMD camera with ambient light suppression, which has a resolution of 16x64 pixels. The camera provides a measurement of the phase difference for every pixel and the modulation ratio which is a measure for the signal quality. The phase difference corresponds to the distance between the surface of an object and the camera. For ease of use we transform the distance data to a representation in Cartesian coordinates. To this end the camera geometry is approximated by central projection and the illumination by a point source at the centre of projection. With these approximations we obtain directly distance measurements in polar coordinates: The angles of the coordinates are built between the rays from the centre of projection through the grid cells of the sensor array. The distance is given by the measurements themselves. This representation is then converted to Cartesian coordinates. A side effect of this procedure is that a slight increase of the lateral resolution towards the image borders is visible now. It is caused by the large aperture angle of 70.5 degree of the camera. To increase the accuracy of our approximations, lens distortion was

corrected during the coordinate transformation. Besides that, the coordinate system was shifted by 1.3 pixels horizontally and 2.5 pixels vertically to account for an offset between the optical axis and the centre of the sensor chip.

To determine the relation between the phase difference and the distance, a flat surface with high reflectance was recorded for several distances. The phase difference was averaged over the middle 5 pixels of the array because there our approximations have the smallest error. The distance function $z(\varphi)$ was then approximated by the linear function

$$z(\varphi) = -5.4022\varphi + 672.5143 \quad (14)$$

with a remaining maximum error of about three percent (see Figure 11 left). Beside the distance function also the relation between distance and standard deviation is of major importance because the accuracy of the measurements depends heavily on the amount of light received by the camera. Since the illumination decreases quadratically with the distance, a second order polynomial was fitted to the data. The resulting function

$$\sigma(\varphi) = 2.734 \cdot 10^{-5} \varphi^2 + 2.86723 \cdot 10^{-3} \varphi - 4.229692 \cdot 10^{-2}. \quad (15)$$

can be seen in Figure 11 right. Figure 12 shows a calibrated distance matrix of the test surface used for camera calibration as well as two nearer marker objects.

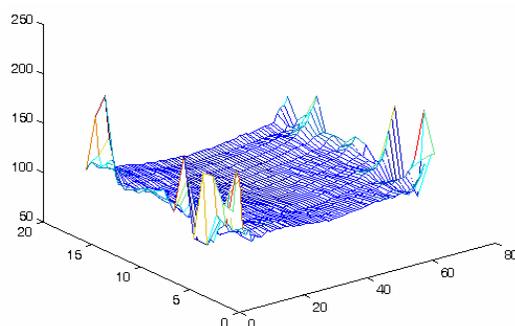


Figure 12. Distance matrix of a flat surface at 3m distance

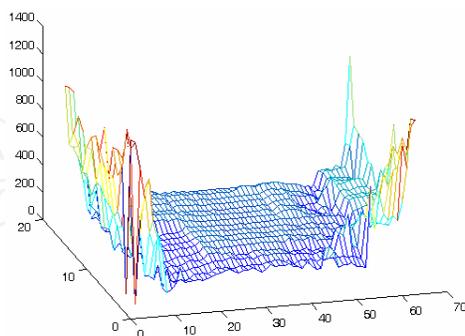


Figure 13. Calibrated distance measurements for a near flat surface (1m distance)

The remaining errors are mainly caused by model inaccuracy for small distances, by too low signal strength and by difficulties with the recorded scene. It turned out that for distances

smaller than 1.5 meters the camera model is not appropriate and the light source should better be modelled by a transmitter with finite area (see Figure 13). The problem of a too low signal occurs mainly at the borders of the sensor array and for very distant objects. But also the reflectance of the objects plays a role and the angle between the light source and the optical axis. Besides that, there is a nonlinear relation between the phase and the signal amplitude due to deviations from the ideal sinus wave shape. To account for these problems, a threshold for the modulation ratio which indicates the signal strength is introduced. Pixels with a modulation ratio below 30 percent are ignored. This procedure ensures that the measurements are in good accordance to function 15.

Beside the reflectance of the recorded objects also their geometry can cause inaccuracies. If the border of an object is mapped to one sensor element, the sensor receives a mixed signal from the object in the front and the background. Then the recorded signal is a linear combination of two sinus waves which are weighted by the reflectance and distance of both objects. It is also possible that the signal is composed of more sinus waves if there are more objects occluding each other. For these sensor elements the true distance cannot be computed. A reasonable assumption then is that the real distance is somewhere between the neighbouring distances. A simple way of handling this is to introduce a minimum and a maximum depth map. The minimum depth map is result of a 3x3 minimum operator on the distance matrix, the maximum map is the result of an analogue 3x3 maximum operator. This is a fast method and accounts for the distance dependency of the error.

3.3 Experimental Results

We present in this section some results from experiments that were conducted to find out more about the real world behaviour of the PMD camera. The goal of the experiments was to gain a calibrating function for any pixel yielding a mapping from a measured value to a standardised value. First we tried to do the calibration process in front of a white wall, but later a board was taken instead of the wall. The reason for that will be addressed later in this section.

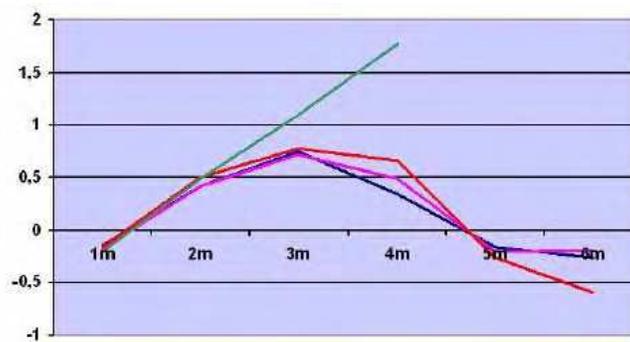


Figure 14. Phase differences of the “wall-scene” and the the “board-scene” (green line)

The camera was placed in front of the wall at distances of 1m to 6m with steps of 1m, which is in the range of 7.5m (see section 4.1). Each time a full measurement was taken by the camera. For all the experiments a tape measure was used instead of the internal alignment function of the camera to measure exactly the distance of the camera to the wall. To

suppress the effect of noise 2500 measurements were taken and every pixel was integrated over a 5x5 area at each distance step.

Up to a distance of 3m the phase difference values show expected linear behaviour. But at distances beyond 3m the values start decreasing again (see Figure 14). This might be due to the fact that the surface of the wall is too smooth so that the light is not reflected in a diffuse manner. Hence the modulated light might not find its way back to the PMD sensor. Further tests showed that this effect is almost independent to the exposure time of the camera.

As an alternative we chose a board with fine but coarse surface resulting in a better diffuse light reflection. With this, the camera showed the expected behaviour. The phase difference values are linearly increasing even beyond the point of 3m (green line). It has to be said that the calibration process could not measure distances of 5m and 6m, because above 4m the board was not big enough to cover the entire image plane of the camera.

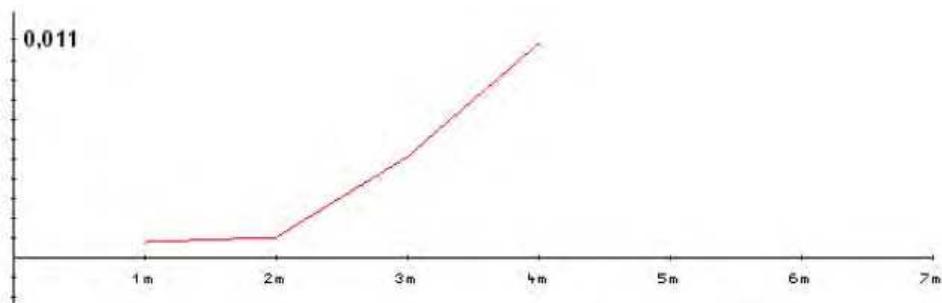


Figure 15. Standard deviations of measurements

We also compared the standard deviations of measurements. The results are presented in Figure 15. This clearly shows the influence of noise at higher distances. The higher noise ratio leads to the increase of the standard deviation.

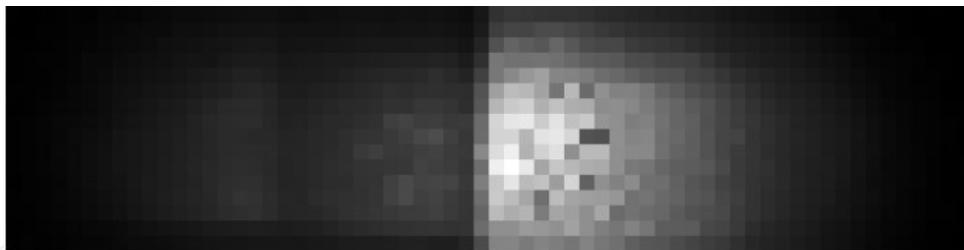


Figure 17. Modulation image of the edge scene

Beside the calibration process we used the camera trying to measure the depth of various scenes. One of the most interesting scenes is the "edge-scene" presented in Figure 16 a) and b). The goal of this experiment is to find out if the camera produces reliable results on object edges. The distance of the wall on the left is 150cm, the distance of the right wall is 80cm. For the scene in Figure 16 a) the depth measurement of the camera is quite accurate with 158cm for the far and 77cm for the near wall differing only 3%-5% from the exact depth values. Moving the camera slightly to the left (as shown in Figure 16 b) in contrast yields unreliable results. The camera looks on the edge at a very beaked angle filling almost a two pixel column. The modulation picture (Figure 17) shows that the degree of modulation

decreases rapidly in the border area of the two walls not allowing any reliable measurements. The reason is the flat angle transporting the light away from the camera instead of reflecting it right back to the sensors. Hence for reliable results one has to make sure the camera is aligned orthogonally towards flat objects. The phase difference picture for Figure 16 a) is presented in fig 18.

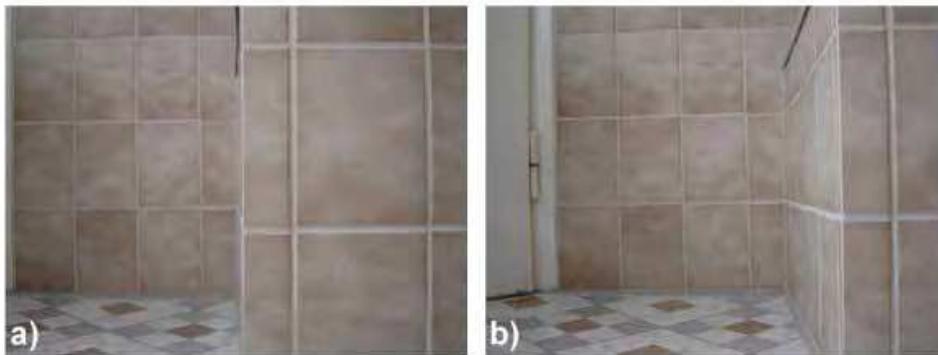


Figure 16. Edge scene - a) camera looking at edge of the wall orthogonally - b) camera looking at edge of the wall at a beaked angle

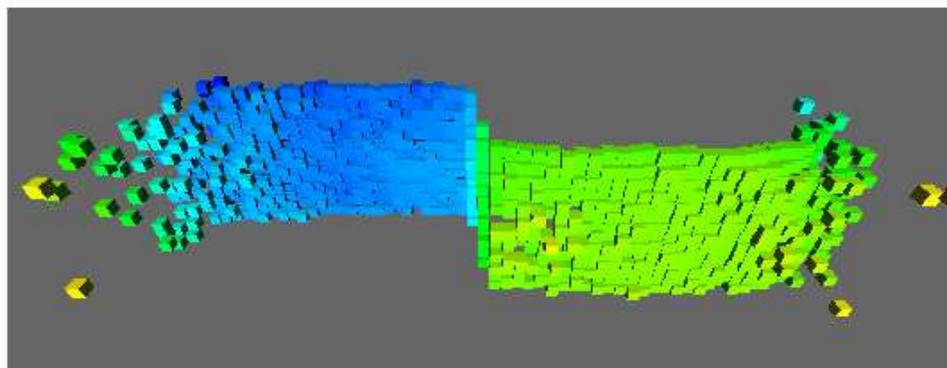


Figure 18. Phase differences for the "wall-scene" using a 3D-view

4. Fusion of PMD and Stereo Data

In comparing the stereo measurements and the results of the PMD camera, the stereo system has the advantage of a high precision with regard to the distance measurements as well as the lateral resolution. Unfortunately, the results depend heavily on the image contents and can be very unstable for ambiguous scenes. The PMD-camera on the other hand provides stable results independently of the surface texture. Here, the coarse resolution of the sensor is unfavourable. Also, the accuracy of the depth measurements is inferior to the results of the stereo camera. Therefore, we made experiments to fuse the results of both techniques in order to obtain depth measurements both robust and precise.

4.1 Fusion Mechanism

Our fusion method is based on the intersection of confidence intervals for the results of both camera types. A schematic of the fusion algorithm is given in Figure 19. The starting point for the computation of the confidence intervals are the depth maps of the camera systems. The depth maps consist of the image coordinates together with the corresponding depth values. Since the camera geometry is known for both systems from the calibration step, the depth values are stored as points in Cartesian coordinates with a metric coordinate system centred at the camera. During camera calibration also the standard deviation of the depth measurements is determined. For the stereo camera the confidence intervals are computed from a depth map by adding, respectively subtracting, twice the standard deviation of the depth values. This corresponds to a 95 percent interval around the mean of a normal distribution. The result is a minimum and a maximum depth map, respectively. For missing depth values, i.e. values recognised as unstable, a large depth interval of 10m is set.

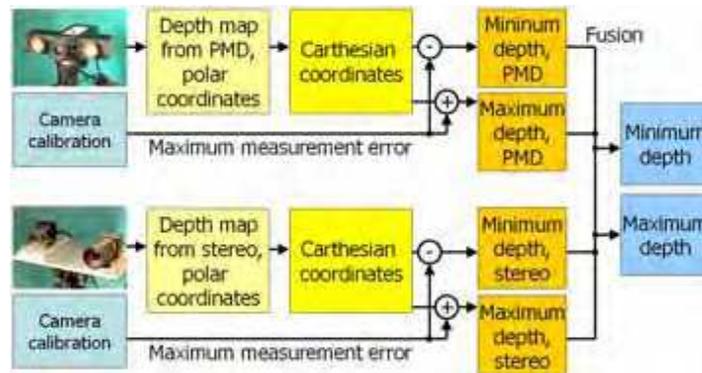


Figure 19. Fusion of the distance data

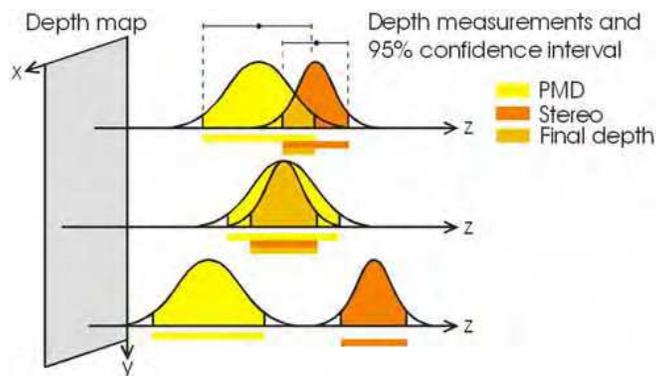


Figure 20. Intersecting the depth intervals of the stereo and the PMD camera

For the PMD-camera the minimum and maximum depth maps are computed by taking the minimum and maximum value of a 3x3 environment around every point in the depth image. This is because for occluding edges a mixed signal of the nearer and farther object surface is received. Then the depth maps are corrected by the measurement error. To compensate for different aperture angles of the camera systems and different lateral

resolutions the depth map of the PMD camera is scaled by appropriate factors in vertical and horizontal direction.

The depth maps now form arrays which give a depth interval for every x - and y -coordinate. Data fusion is then done by intersecting the intervals of the PMD-camera with the intervals of the stereo camera. Figure 20 illustrates three different cases. The first case is that the depth intervals partially overlap. The area marked as 'final depth' is the intersection of the two depth intervals. The second case shown is that one depth interval completely covers the other one. Here, both sensors deliver the same depth value, but with different measurement accuracy. The third case is that there is no overlap. In that case one sensor or both deliver wrong values. Without taking further assumptions, nothing more can be said here about the true depth value, so the final value is marked as missing.

4.2 Experimental Results

Figure 21 shows a scheme of the experimental setup and the recorded scene. The experiments were conducted in a corridor of the university. The only changes to the original scene are the person we asked to stand in the corridor and the low carton placed in the foreground. In particular, we did not facilitate the depth recognition by e.g. hanging up highly textured

posters. The environment can thus be considered as a natural indoor scene (for office buildings). The distance between the cameras and the objects of the scene was 3m to 4m. The stereo camera was placed behind the PMD camera because of the smaller aperture angle of 23 degree compared to the 70.5 degree of the PMD camera. It was also placed 30cm higher to avoid the PMD camera of being visible in the stereo images.

Figure 22 shows a pair of images from the stereo camera. These images are problematic for stereo analysis in many ways. First, there are big homogeneous regions, primarily the white walls and the white column. Scanline stereo or block matching fails in these regions due to the lack of characteristic image features. Secondly, the shirt of the person shows a repetitive pattern causing multiple solutions to the correspondence problem. In the remaining parts of the image stereo analysis is affected by occlusion. This concerns mainly the ceiling and the environment around the person. Figure 23 shows the resulting disparity and depth maps. Since most parts of the images lead to unstable results, a sobel operator was applied to the images to find areas where stable results can be expected. Disparity values in unstable regions are discarded. In the given depth map these areas are marked black. They are replaced by the mentioned 10m interval before fusion with the PMD camera.

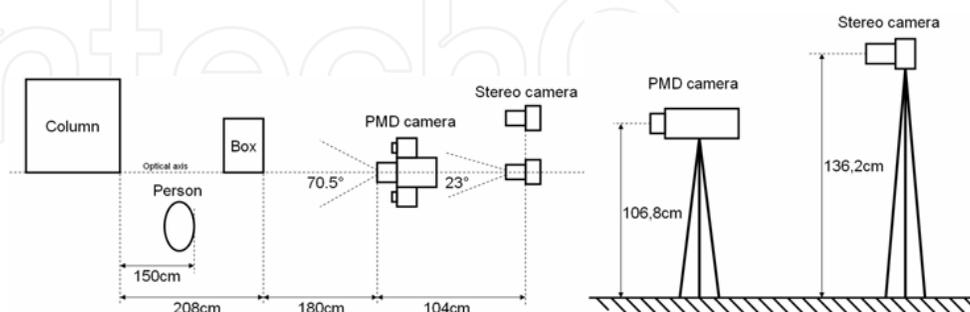


Figure 21. Left: Top view of the experimental setup. Right: Side view of the experimental setup



Figure 22. Left and right image of the stereo camera



Figure 23. Left: Disparity map. Right: Resulting depth map (in image coordinates)

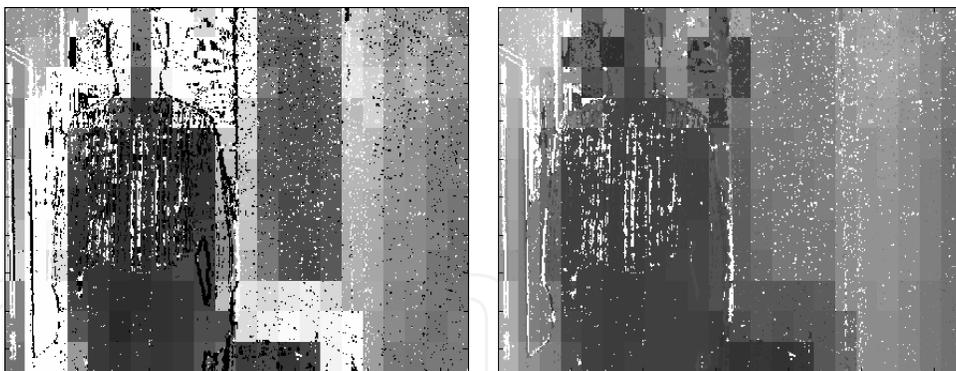


Figure 24. Size of the overlapping depth intervals (left) and average depth (right)

The results of the sensor fusion can be seen in Figure 24. The left picture shows the size of the overlapping depth intervals, the right image shows the mean depth. The coarse block structure results from the low resolution of the PMD camera. Coordinates without overlapping depth intervals are marked as white. The bright blocks in the left figure indicate big depth intervals around the person, at the edges of the column and around the box. This inaccuracy results from depth discontinuities because there the PMD camera receives a

signal both from the far and the near surface. By contrast, the stereo camera is quite accurate for these positions and delivers in general the distance to the nearer object. Although the depth map from the stereo analysis is relatively sparse, the result after sensor fusion is dense. Missing results from stereo are replaced by values from the PMD camera because the stereo system provides sufficient large depth intervals when a result is unknown. The few missing values in the fused depth map occur when both systems report a high accuracy for their results, although one of the depth values or the accuracy itself is wrong. Although the fusion mechanism itself is comparatively straight forward, it seems to preserve the advantages of both 3D sensors while avoiding their disadvantages

6. Conclusion

Common state of the art mechanisms for the measurement of the surrounding environment in real time usually pose a trade off between high speed, robustness and accuracy. With applications for mobile robots in mind, this work focuses on the faster methods stereo analysis and PMD camera. Our research aims at the computation of robust and dense depth maps in real time.

First, the performance of three standard stereo algorithms is examined with regard to two different measurements of similarity. The subsequent optimisation of the standard methods by using modern SIMD instructions and programming techniques like e.g. recursive subdivision leads to an increase of speed by a factor of four. As a result, for the Winner-Takes-It-All algorithm we achieve a computation time of 250ms (plus 230ms for building the cost matrix), which can be considered real time. The accuracy of our stereo setup is determined experimentally and a scene reconstructed from stereo data is shown.

To improve robustness and speed the image derivative is evaluated. In order to deal with poorly structured environments experiments with an adaptive block size are conducted. The resulting disparity maps are more dense but the resulting block sizes for homogeneous regions slow down the correspondence analysis. Hence, this approach is not suitable for real time. A fast post processing step dealing with a disparity histogram is introduced to discard wrong matches. The subpixel disparity is computed as a measure of plausibility.

As a comparatively new technique the PMD camera is used for distance measurement. The PMD camera provides directly the depth information for every pixel without the intensive computation that characterises stereo analysis. The camera was calibrated with an accuracy of 5 percent for distances over 1.5m. For smaller distances a more complex model than a pinhole camera with a point light source is needed. We observe that the measurement error increases quadratically with the distance, which is an effect of the reduced amount of light received from distant surfaces. Other inaccuracies result from the reflectance properties of the recorded surfaces or extreme geometric arrangements of the scene. A big advantage of the PMD camera is that it does not rely on the texture of a surface or the visibility of objects in a second camera like a stereo camera. With an (adjustable) integration time of 80ms per image it is also much faster. On the other hand, the stereo camera has a higher image resolution as well as higher depth accuracy. Especially the behaviour on and near edges is better. We thus made experiments combining both methods and they turn out to compensate the disadvantages of each other very well. As a result, we obtain robust and dense depth information in real time.

7. References

- Bähr, H.-P. & Vögtle, T. (1991). *Digitale Bildverarbeitung. Anwendung in Photogrammetrie, Kartographie und Fernerkundung*, Wichmann, Karlsruhe, ISBN 3-87907-224-8
- Faugeras, O. (1993). *Three-Dimensional Computer Vision – A Geometric Viewpoint*, MIT Press, Cambridge, Mass., ISBN 0-262-06158-9
- Kang, S.; Webb, J.A.; Zitnick, C. & Kanade, T. (1995). A multibaseline stereo system with active illumination and real-time image acquisition, *Proceedings of the Fifth International Conference on Computer Vision (ICCV '95)*, June 1995, pp. 88-93.
- Kim, H.; Min, D.B. & Sohn, K. (2005). Real-Time Stereo Using Foreground Segmentation and Hierarchical Disparity Estimation, in *Ho, Y.S. & Kim, H.J. (Eds.): PCM 2005, Part I, LNCS 3767*, Springer-Verlag, Berlin, Heidelberg, pp. 384-395.
- Konolige, K. (1997). Small Vision System: Hardware and Implementation, *In Proceedings of the 8th International Symposium on Robotics Research*, Japan.
- Kraft, H.; Frey, J.; Moeller, T.; Albrecht, M.; Grothof, M.; Schink, B.; Hess, H. & Buxbaum, B. (2004). 3D-Camera of high 3D-frame rate, depth-resolution and background light elimination based on improved PMD (photonic mixer device)-technologies, OPTO, Nuernberg, May 2004.
- Moravec, H.P. (1977). Towards automatic visual obstacle avoidance. *Proceedings of the Fifth International Joint Conference on Artificial Intelligence, IJCAI-77*.
- Metropolis, N.; Rosenbluth, A.; Rosenbluth N.; Teller A. & Teller E. (1953). Equation of State Calculations by Fast Computing Machines, *Journal of Chemical Physics*, article 21, pages 1087-1159.
- Porta, J.M.; Verbeek, J.J. & Kröse, B.J.A. (2005). Active appearance-based robot localization using stereo vision, *Autonomous Robots* 18(1), pages 59-80.
- Scharstein, D.; Szeliski, R. & Zabih, R. (2001). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, *Proceedings of the IEEE Workshop on Stereo and Multi-Baseline Vision*, Kauai, HI, Dec. 2001.
- Schwarte, R. (2001). *Dynamic 3D-Vision, IEEE Int. Symp. on Electron Devices for Microwave and Opto-electronic Applications, EDMO 2001*, Wien.
- Stereo Vision Research Page* (2007). <http://cat.middlebury.edu/stereo/newdata.html>, Middlebury University, 30.1.2007.
- Sun, Ch. (2002). Fast Stereo Matching using Rectangular Subregioning and 3D Maximum Surface Techniques, *International Journal of Computer Vision*, 47(1/2/3), May 2002
- Sunyoto, H.; van der Mark, H. & Gavrilu, D.M. (2004). A Comparative Study of Fast Dense Stereo Vision Algorithms, University of Parma, Italy - 2004 *IEEE Intelligent Vehicle Symposium*.
- Zhao, J. & Katupitiya, J. (2006a). A Fast Stereo Vision Algorithm with Improved Performance at Object Borders, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'06)*, Beijing, China, October 9-15, 2006.
- Zhao, J. & Katupitiya, J. (2006b). A Dynamic Programming Approach Based Stereo Vision Algorithm Improving Object Border Performance, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'06)*, Beijing, China, October 9-15, 2006.
- Zhu, Z.; Karupiah, D.R.; Riseman, E. & Hanson, A. (2004). Adaptive panoramic stereo vision for human tracking with cooperative mobile robots, *Robotics and Automation Magazine*, Special Issue on Panoramic Robots, 14 (10): pages 69-78.



Vision Systems: Applications

Edited by Goro Obinata and Ashish Dutta

ISBN 978-3-902613-01-1

Hard cover, 608 pages

Publisher I-Tech Education and Publishing

Published online 01, June, 2007

Published in print edition June, 2007

Computer Vision is the most important key in developing autonomous navigation systems for interaction with the environment. It also leads us to marvel at the functioning of our own vision system. In this book we have collected the latest applications of vision research from around the world. It contains both the conventional research areas like mobile robot navigation and map building, and more recent applications such as, micro vision, etc. The first seven chapters contain the newer applications of vision like micro vision, grasping using vision, behavior based perception, inspection of railways and humanitarian demining. The later chapters deal with applications of vision in mobile robot navigation, camera calibration, object detection in vision search, map building, etc.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

K.-D. Kuhnert, M. Langer, M. Stommel and A. Kolb (2007). Dynamic 3D-Vision, Vision Systems: Applications, Goro Obinata and Ashish Dutta (Ed.), ISBN: 978-3-902613-01-1, InTech, Available from:
http://www.intechopen.com/books/vision_systems_applications/dynamic_3d-vision

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2007 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen