APPLYING DITHERING TO IMPROVE DEPTH MEASUREMENT USING A SENSOR-SHIFTED STEREO CAMERA

Jiandan Chen, Wail Mustafa, Abu Bakr Siddig, Wlodek Kulesza

Blekinge Institute of Technology, SE-371 79 Karlskrona, Sweden (✉ jian.d.chen@bth.se, +46 70 830 1294, wail.mustafa@bth.se, ahsi08@student.bth.se, wlodek.kulesza@bth.se)

Abstract

The sensor-shifted stereo camera provides the mechanism for obtaining 3D information in a wide field of view. This novel kind of stereo requires a simpler matching process in comparison to convergence stereo. In addition to this, the uncertainty of depth estimation of a target point in 3D space is defined by the spatial quantization caused by the digital images. The dithering approach is a way to reduce the depth reconstruction uncertainty through a controlled adjustment of the stereo parameters that shift the spatial quantization levels. In this paper, a mathematical model that relates the stereo setup parameters to the iso-disparities is developed and used for depth estimation. The enhancement of the depth measurement accuracy for this kind of stereo through applying the dithering method is verified by simulation and physical experiment. For the verification, the uncertainty of the depth measurement using dithering is compared with the uncertainty produced by the direct triangulation method. A 49% improvement of the uncertainly in the depth reconstruction is proved.

Keywords: depth reconstruction, dithering, skewed parallel, stereo setup, iso-disparity surfaces.

1. Introduction

Advanced technologies may help people to extend their ability to process visual information. This raises the demand for autonomous systems with high performance sensors. This paper is concerned with a stereo camera and its applications in human activity monitoring. The Intelligent Vision Agent System, IVAS, is a vision and information processing system used in these kinds of applications [1, 2] and [3]. The IVAS gathers data in order to reconstruct 3D information that can be used in health care, security and surveillance applications. The system focuses on an interesting part of the scene by dynamic control of the stereo image. Such a system requires high accuracy in the reconstruction of the 3D information in order to guarantee high performance.

The human activity field is a 3D world, where the location of each point is represented by x, y and z coordinates. However, a camera can only capture a two-dimensional image where each point is represented by x and y coordinates. The stereo system provides the mechanism for acquiring the vital z coordinate [4]. In this case, the z coordinate is referred to as the depth, and the process of acquiring the depth from the stereo system is called depth reconstruction.

The stereo system captures two images in the 3D world. The reconstructed point needs to have a projection in each image, and this can only happen when the point is located in the common field of view (FoV) of the camera pair. To be able to reconstruct 3D information, the system needs to first implement a matching process to find the corresponding points in the two images of the same view.

However, a digital camera quantizes the image plane into an array of pixels that forms the digital image. Because of this, the projection points are approximated and assumed to be located in the centers of the pixels. The difference between the reconstructed depth estimated

from the exact and the discretised projections is referred to as the depth reconstruction uncertainty. The depth reconstruction uncertainty is related to the pixel size of the camera sensor. The selection of an optimal sensor pixel size is discussed in [5].

Regarding the image resolution, the limitation of the pixel size is overcome by combining the information from slightly different low-resolution images of the same scene into a higher-resolution image. This way of enhancing the image resolution is called super-resolution reconstruction [6]. In a similar way, two pairs of images taken by a stereo system with two slightly different setups can be used in combination to reconstruct the depth with enhanced uncertainty. The two different setups can result from a readjustment of one of the stereo parameters. This method is referred to as the dithering approach [7]. The dithering technique is commonly used to overcome the true color issue that exists in colored image printing [8]. The effect of the dithering on the estimation of the sine wave amplitude is studied in [9].

In many cases, the camera has to be rotated to capture a certain view that is initially outside the camera FoV. This rotation causes distortion of the shapes in the captured images. Hence, further processing is required to overcome this distortion. Instead of rotating the whole camera, professional photographers use a technique called a sensor-shifted camera that applies a millimeter shift between the camera lens and the sensor [10]. This shift provides an effect similar to the rotation, so that it captures the wanted view while avoiding the distortion.

Francisco and Bergholm have proposed the use of a sensor-shifted camera in the stereo setup where the sensor has a controlled micro-movement [11]. This kind of stereo system is referred to as the skewed-parallel stereo camera. In that paper, the authors discuss the benefits of using this kind of camera in a stereo system instead of the vergence movement used in the general stereo setup. When compared to the convergence camera stereo setup, the skewed-parallel camera setup requires simpler reconstruction processing. In addition to this, the skewed-parallel camera setup offers a wider common field of view than the parallel camera stereo setup. A similar camera was used by Ben-Ezra et al. to minimize the motion blur in the reconstruction of super-resolution images of a video signal [12].

After an introduction and a summary of the related works, the problem statement and the main contributions are provided in Section 2. In Section 3, the geometric model of the skewed-parallel stereo camera is described followed by the derivation of the dither signal for this stereo system. Section 4 discusses the implementation of the dithering algorithm along with the conducted synthetic and physical validation experiments, and their results. The prototype of the skewed-parallel camera used in the physical experiment is also described in this section. Finally, a conclusion and a recommendation for future work are provided in Section 5.

2. Problem statement and main contributions

The concern of this paper is to improve the depth measurement uncertainty through the use of the dithering algorithm for a specific setup of the skewed-parallel stereo camera. The uncertainty of the depth measurement can also be reduced by decreasing the pixel size but the drawback of this method is that this will also reduce the signal-to-noise ratio. Therefore, we propose the application of the dithering method as a simple and robust way to improve the reconstruction uncertainty. The modeling of the depth reconstruction using the dithering algorithm for this kind of stereo system, and its validation through simulation and physical experiment, are the main problems that need to be resolved before applying the method to real measurement.

This paper contributes to current research by:

− Developing a mathematical model of the depth measurement using the skewed parallel stereo system.

www.czasopisma.pan.pl　　PAN　　www.journals.pan.pl

POLSKA AKADEMIA NAUK

Metrol. Meas. Syst., Vol. XVII (2010), No. 3, pp. 335–348

- Developing and implementing the dithering algorithm based on the depth measurement model.
- Configuring a prototype for the skewed-parallel camera setup.
- Validating the depth reconstruction enhancement of the dithering algorithm for the skewed-parallel camera stereo setup through simulation and physical experiment.

3. Problem analysis and modelling

3.1. The skewed-parallel stereo geometric model

In our approach, we use the pinhole camera model. The setup of the skewed-parallel stereo camera in the x–z plane is shown in Fig. 1 [13]. The center of the coordinates is in the middle of the baseline, B. The baseline is the distance between the optical lenses' centers, o_l and o_r, of the left and the right cameras respectively. The sensors lie on the same horizontal line, and the sensor centers are denoted as c_l and c_r for the left and the right cameras respectively. The shift of the sensor is defined as the horizontal distance between the optical center and the sensor center and it can be different for each camera. The shifts of the sensors of the left and the right cameras are denoted as S_l and S_r respectively. In our notation, the shifts are positive for a movement of the sensors to the right, and negative for an opposite direction. The focal lengths of the two cameras are assumed to be the same and are denoted as f.

The angle between the optical axis and the primary axis, which is defined as the line passing through the sensor center and the center of the lens, is called the convergence angle and it is denoted as α_l and α_r for the left and right cameras respectively. The convergence angle exists as a result of shifting the sensors of the skewed-parallel stereo cameras. This shift has the same effect as rotating the stereo pair since it introduces a fixation point, P_0, and widens the common FoV. Using trigonometry, the convergence angle of the left camera can be derived as:

$$\alpha_l = \tan^{-1}\frac{S_l}{f} \quad \text{for} \quad -\frac{m}{2} \le S_l \le \frac{m}{2}, \tag{1}$$

where m is the length of the sensor plane.

For any point in the space, P, with the depth Z, the projections of the point on the left and right sensors along the x-axis are x_l and x_r respectively (the x coordinates of each sensor cross the sensor middle). Considering the quantization effect, these projections are approximated by the pixel centers, and denoted as x_{Ql} and x_{Qr} for the left and right cameras respectively. In this case, the quantized depth Z_q of the point can be found through:

$$Z_q(n) = \frac{fB}{(n\Delta D + (S_r - S_l))} \quad \text{for} \quad n\Delta D + (S_r - S_l) > 0 \tag{2}$$

with

$$n = x_{Qr} - x_{Ql} = \left\lfloor \frac{x_r}{\Delta D} \right\rceil - \left\lfloor \frac{x_l}{\Delta D} \right\rceil, \tag{3}$$

where ΔD is the length of pixel, n is an integer number representing the target disparity, and the symbol $\lfloor \cdot \rceil$ denotes rounding to the nearest integer.

The depth Z_0 of the fixation point, P_0, being a cross section of the primary axes, can be found by setting n to zero in Eq. (2). The x and z coordinates of the fixation point are:

$$\left(X_{P_0}, Z_{P_0} \right) = \left(\frac{S_r + S_l}{2(S_r - S_l)} B, \frac{fB}{S_r - S_l} \right). \tag{4}$$
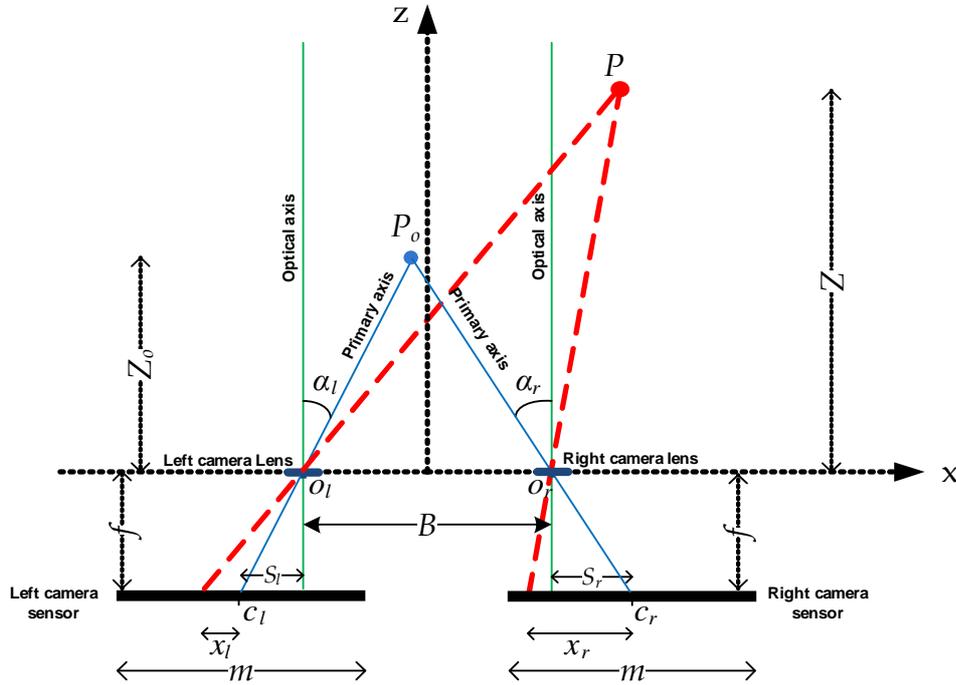

Fig. 1. An *x-z* view of the skewed-parallel stereo camera schematic diagram.

The disparity is the displacement of the corresponding *x*-axis projections of a certain point in space on the left and the right images. From Eq. (2), it can be seen that the depth reconstruction is inversely proportional to the disparity. For each disparity *n*, there exists a corresponding iso-disparity surface that represents the depth of all the points that have the same disparity [7]. The iso-disparity surfaces appear as lines in the *x*–*z* plane as seen in Fig. 2 for the red lines. The interval between the iso-disparity surfaces represents the depth reconstruction quantization uncertainty, which is a nonlinear function of *n*.

3.2. Application of the dither signal

The idea behind the dithering technique is to add noise to the signal prior to the quantization process in order to slightly change the statistical properties of the quantization [2]. The quantizers in this model are the cameras and the quantized signals are the target point projections x_l and x_r for the left and right cameras respectively. Using the dithering technique to reduce the uncertainty of the depth reconstruction for the parallel stereo setup was proposed in [7]. In that paper, authors proved that the uncertainty is reduced by half by applying the two stage discrete binary dither signal. To accomplish this reduction, a dither signal adjusts the stereo setup for a secondary measurement that follows the initial one, and the depth can then be estimated from all these measurements.

In this paper, we use a two-stage discrete binary dither signal for each camera. This means that we make use of four images to calculate the depth of the target. This allows us to estimate the depth with a reduced quantization uncertainty. In the parallel stereo setup, the depth reconstruction uncertainty is halved when the dithering algorithm is applied. The optimal diter signal makes the target projection move from its original position by a distance that is equal to half a pixel size [7].This means that the dither signal shifts the iso-disparity line so that it lies

www.czasopisma.pan.pl   PAN   www.journals.pan.pl
POLSKA AKADEMIA NAUK

Metrol. Meas. Syst., Vol. XVII (2010), No. 3, pp. 335–348

in the middle of the two consecutive iso-disparity lines n_t and n_t+1 between which the target is present. This shift reduces the uncertainty by half.

For the skewed-parallel stereo camera setup, calling (2), the difference between two consecutive iso-disparity lines, ΔZ_t, which represents the depth reconstruction uncertainty, can be found to be:

$$\Delta Z_t = \frac{Bf\,\Delta D}{\left[n_t\Delta D+\left(S_r-S_l\right)\right]\left[\left(n_t+1\right)\Delta D+\left(S_r-S_l\right)\right]},\qquad(5)$$

where t refers to a specific iso-disparity line n_t.

To estimate the dither signal, generated by a shift of one sensor, that affects the iso-disparity line n_t to move exactly into the middle of ΔZ_t, we should first determine the depth in the middle by:

$$Z_t+\frac{\Delta Z_t}{2}=\frac{Bf}{n_t\Delta D+\left(S_r-S_l\right)+\Delta S_t},\qquad(6)$$

where ΔS_t is the shift introduced to one sensor that moves the iso-disparity surfaces to the middle of the two consecutive iso-disparity lines n_t and n_t+1.

Calling (2) and (5), the dither signal, ΔS_t, can be mathematically proven to be:

$$\Delta S_t=-\frac{\left(n_t\Delta D+\left(S_r-S_l\right)\right)\Delta D}{2\left(\left(n_t+1\right)\Delta D+\left(S_r-S_l\right)\right)+\Delta D}.\qquad(7)$$

To verify the dither signal obtained by (7), Matlab 7 [14] and the Epipolar Geometry Toolbox [15] are used. In the verification scenario, a target in the common FoV that belongs to a specific iso-disparity line n_t is chosen. Then, the dither signal is applied to check the movement of the iso-disparity lines. For the simulation, the baseline B is set to 100 mm, the focal lengths f are 25 mm each, the pixel length ΔD is 8.33 μm and the target is assumed to be in the disparity line $n_t = 178$.
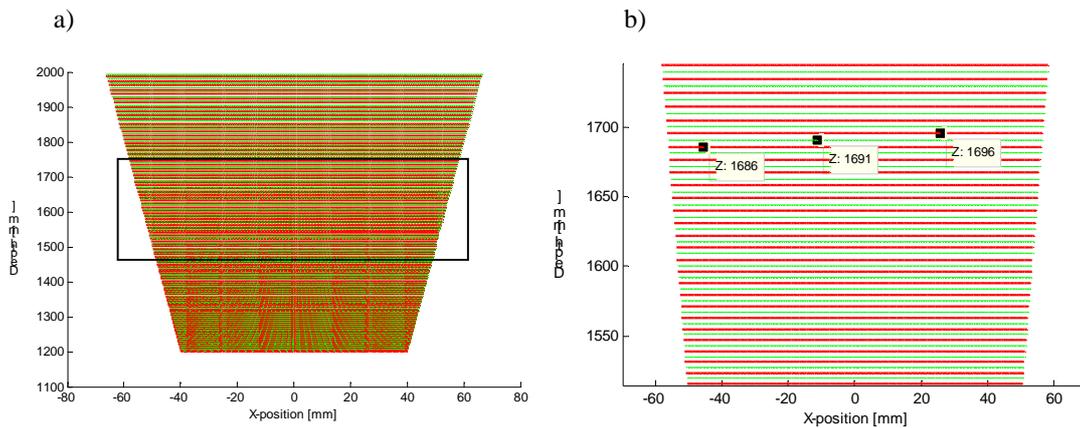

Fig. 2. Iso-disparities obtained by simulation before and after applying the dither signal. Red lines come from the primary and green lines were obtained after the dither signal. b) Zoomed-in area (inside the box) of a).

Fig. 2a shows the iso-disparity lines in 2D. Red lines denote the iso-disparity lines between 1200 mm and 2000 mm of depth for the original setup, while the green lines denote the iso-disparities after introducing the dither signal, ΔS_t, to both cameras. In this case, ΔS_t is found to be −4 μm approximately, which is equal to 0.5 pixels. This shows that the skewed-parallel

stereo camera has the same property as the parallel stereo camera regarding the dither signal [7]. To place the new iso-disparity surfaces in the middle of the old ones, it is required that the projection of the target feature is shifted half a pixel.

Fig. 2b shows a zoomed area around $n_t = 178$. From calculation, before applying the dither signal, the depth for the specified disparity can be determined to 1686 mm and 1696 mm for the next disparity level. After applying the dither signal, the depth for the same disparity $n_t = 178$, is found to be 1691 mm, which falls in the middle between the above-mentioned depths

4. Implementation and validation

4.1. Implementation of the dithering algorithm

From the description of the dithering algorithm in [7], and the description of the estimation of the dither signal for the skewed-parallel stereo camera in Section 3.2, the dithering algorithm can be defined. By applying the dither signal ΔS_t which controls the sensor positions of the left and right cameras, four images are obtained; two images before dithering and two images after dithering. For the disparity calculations, we can combine those four images into six pairs of images. In practice, however, since the dither signal moves the camera sensor a very short distance and the disparity of the two pairs of images taken by the same camera is too small to be useful to extract depth information from this pair. Therefore, only four pairs are considered.

The quantized projections of the target point obtained from the right and left images can be used to create the projection matrices x_{Qr} and x_{Ql} for the right and left cameras respectively as follows:

$$x_{Qr} = \begin{bmatrix} x_{Qr1} & x_{Qr2} \\ -1 & -1 \end{bmatrix} \text{ and } x_{Ql} = \begin{bmatrix} 1 & 1 \\ x_{Ql1} & x_{Ql2} \end{bmatrix}, \tag{8}$$

where x_{Qr1} and x_{Qr2} are the quantized projections of the target on the right image before and after applying the dither signal respectively, while x_{Ql1} and x_{Ql2} are the respective projections on the left image.

From the projection matrices, the dithering matrix, that contains the disparities of the four considered pairs, can be obtained by:

$$\mathbf{d} = x_{Qr}^{\text{T}} . x_{Ql}. \tag{9}$$

The dithering algorithm can be then implemented through the four following steps:
1. Preliminary estimation of the depth from the disparity of the initial pair of images using Eq. (2).
2. Estimation and application of the dither signal, ΔS_t, (7) to shift the sensors of the two cameras.
3. Secondary estimation of the four depths corresponding to the four disparities in $\mathbf{d}$ (9).
4. Calculation of the depth of the target point by averaging the four depths from step 3.

4.2. Synthetic experiment

The synthetic experiment was performed using Matlab 7 and the Epipolar Geometry Toolbox [15]. The simulation environment is a 3D space, with two pinhole skewed parallel cameras. The target is assumed to be 1500 points randomly distributed in a cubic area with the dimension 300 mm × 300 mm × 300 mm. The cubic center is set to be (0, 0, 1600) in *XYZ*

www.czasopisma.pan.pl   PAN   www.journals.pan.pl

POLSKA AKADEMIA NAUK

Metrol. Meas. Syst., Vol. XVII (2010), No. 3, pp. 335–348

coordinates. The target points fall in the common FoV of the stereo camera. Fig. 3 shows a perspective view for this setup in 3D space.
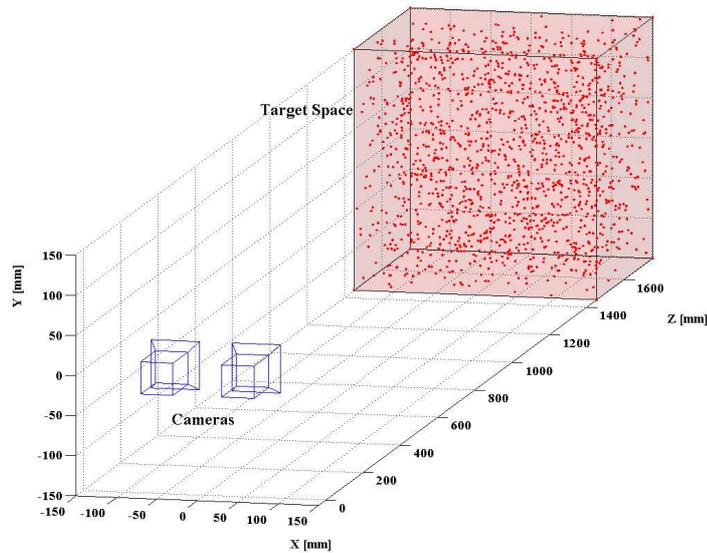

Fig. 3. The simulation setup: The points in the cubic area (red dots) are the targets.

The setup of the stereo camera is: the baseline B is 100 mm, the focal length f is 25 mm, and the pixel size ΔD is 8.33 µm. The simulation scenario is to measure the depth of each point, where the initial shifts of the two cameras are set to zero, using the direct method through Eq. (2), and using the dithering algorithm described in Section 4.1. The results present a comparison between the two methods in order to illustrate how the depth reconstruction uncertainty can be improved by the dithering algorithm for this kind of stereo system.
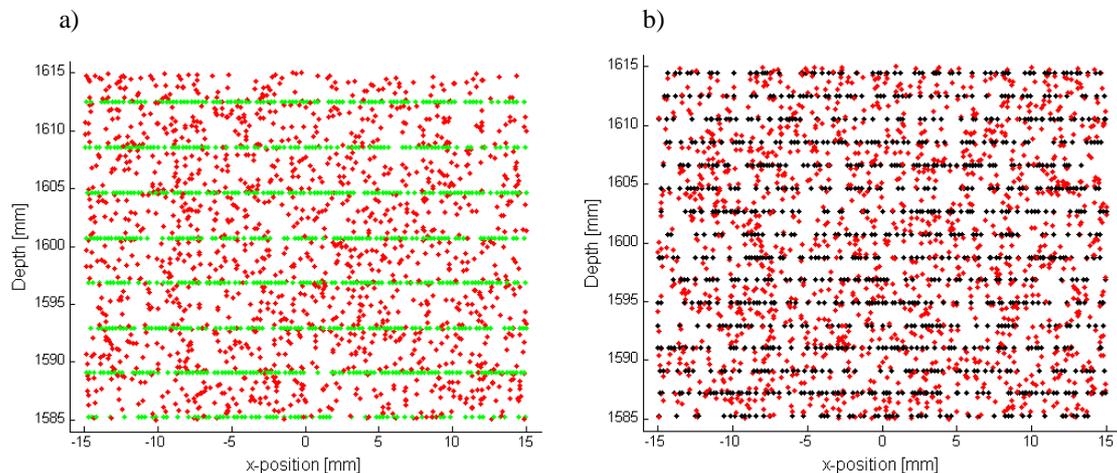

Fig. 4. Top view of the depth reconstruction of the original target points (in red). The green points represent the reconstructed points by the direct method; b) describes the reconstructed points by the dithering algorithm as black points.

Fig. 4 shows the top view of the target points in their original positions and in their reconstructed positions after the depth measurement for the two methods. The figure shows the points within the zoomed range from −15 mm to 15 mm along the *X*-axis, and from 1585 mm to 1615 mm along the *Z*-axis. In Fig. 4a, the red dots represent the original

target points while the green dots represent the depth estimations for these points using the direct reconstruction method. The black dots in Fig. 4b represent the depth estimations for the same target points using the dithering algorithm. The black dots form new iso-disparity surfaces with different intervals that correspond to the reconstruction uncertainty of the method when applying the dithering algorithm.
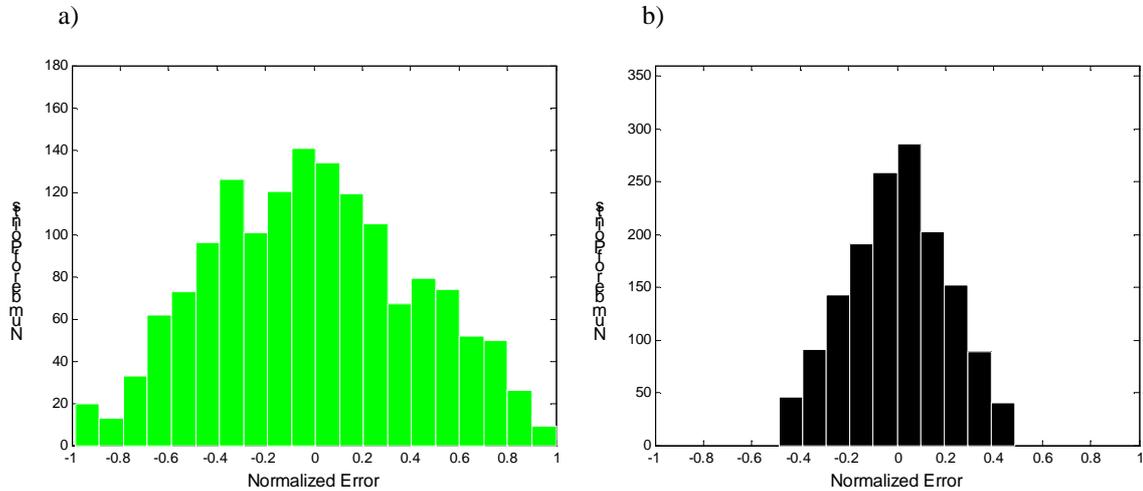
a)   b)

Fig. 5. Histograms of the normalized depth reconstruction error a) for the direct triangular method and b) after applying the dithering algorithm.

4.3. Skewed-parallel Camera Prototype

To physically implement the dithering algorithm for the skewed-parallel camera, it is necessary to make use of a camera with a sensor that is capable of moving in a controlled horizontal movement in relation to the lens. To satisfy this requirement, a prototype can be built by separating the camera body, which contains the sensor, from the lens that is normally attached to that body. Additionally, to control the movement of the camera sensor, the camera body needs to be attached to a micro-movement mechanical device such as an *x*-positioner or a DC motor.

Following the above design considerations, a skewed parallel camera prototype was built to be used for the experimental needs [11]. This prototype was reconfigured for the purpose of this project. The camera prototype contains four main components that are: a camera module, a linear stage controlled by a DC motor, a lens and a metal stand that holds all the components. In Fig. 6, the prototype containing these components along with a target grid pattern that is used for the experiment part of the paper are shown.

The camera module is connected through a cable to a frame grabber card that is installed in a computer where the pictures are stored and processed. This module contains a CCD sensor that represents the shifted sensor in the skewed-parallel camera model. The Sony XC-555P camera module is used in this prototype. The camera module is a color video camera with a 1/2 type sensor [16]. The captured image resolution is 768 (H) x 576 (V) pixels.

In order to provide the horizontal shift of the camera module, the prototype uses a linear stage with a DC motor. The linear stage movement distance is up to 25 mm with a resolution of 0.06 μm/count (motor step) [17]. This resolution is sufficient to provide the required micro-movement capability of the sensor.

To configure this camera prototype as a stereo camera, it has to be placed in two different positions where the distance between the two positions represents the baseline *B* of the stereo camera. The prototype has been attached to a translational stage with an accuracy of 1 mm, allowing an accurate implementation of the position.

The Tamron 23FM25SP lens model is used in this prototype. The lens is a C-mount type with a focal length of 25 mm. The focus of the lens can be adjusted for objects that lie between 0.15 m and ∞ from the front of the lens [18].
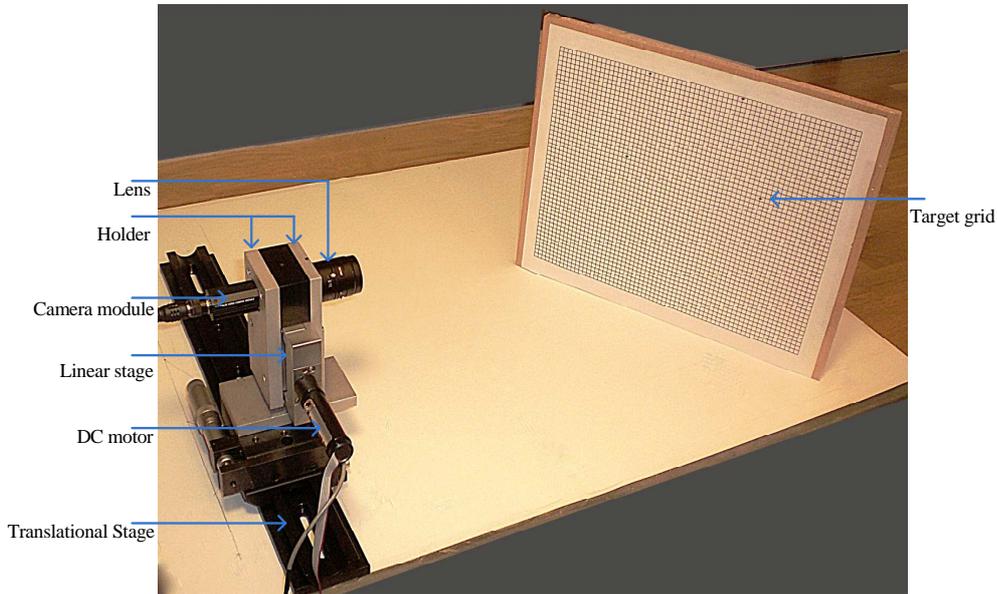

Fig. 6. A side view of the camera prototype showing its components and the experiment setup.

4.4. Physical experiment

To validate that the dithering algorithm enhances the depth reconstruction for the skewed-parallel camera, we designed a physical experiment [19]. The setup of the experiment is shown in Fig. 6 whereas the camera prototype is described in Section 4.3. The targets have been represented as points in a pattern consisting of a grid of lines with a distance of 5 mm between the lines. The pattern has been pasted onto a board that can be easily positioned in front of the stereo camera along the optical axis.

In [7], a way to validate the enhancement of the depth reconstruction is proposed. The idea is to measure the differential depth Z_{AB}, which is the distance between two targets along the optical axis, instead of the absolute depth. The method provides more accurate validation because it avoids measuring the distance between the test target and the center of the lens, something which is difficult to determine.

The reconstruction uncertainty of the differential depth, ΔZ_{AB}, represents a difference between the reference value of Z_{AB} and the reconstructed value by the stereo vision system. The probability distribution function, PDF, of the differential depth quantization uncertainty is described as the convolution of the depth quantization uncertainties at the two target points:

$$p(\Delta Z_{AB}) = p(\Delta Z_A) \otimes p(\Delta Z_B),\qquad(10)$$

where ΔZ_A and ΔZ_B are depth reconstruction quantization uncertainties at the target points A and B respectively and $\otimes$ denotes convolution. The PDF of the depth uncertainty ΔZ is defined as [20]:

$$p(\Delta Z) = \begin{cases} \dfrac{n_t^2 \Delta D}{(Bf)^2}\left(\Delta Z n_t^2 \Delta D + Bf\right), & \Delta Z_m \leq \Delta Z < 0 \\[3mm] -\dfrac{n_t^2 \Delta D}{(Bf)^2}\left(\Delta Z n_t^2 \Delta D - Bf\right), & 0 \leq \Delta Z < \Delta Z_m \\[3mm] 0, & \text{elsewhere} \end{cases} \tag{11}$$

where ΔZ_m is the maximum depth reconstruction uncertainty corresponding the to the interval between iso-disparity surfaces.

According to (10), the range of the differential depth reconstruction quantization uncertainty is the sum of the depth quantization uncertainty ranges of the two corresponding points.
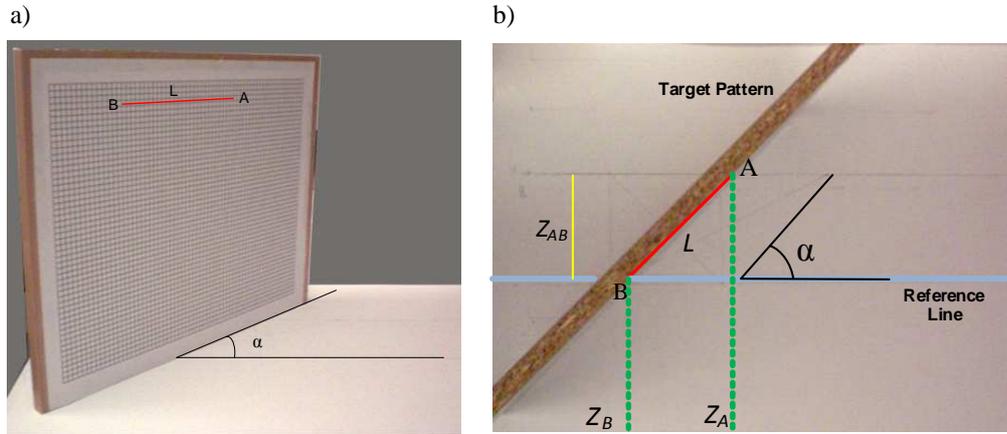

Fig. 7. The target grid positioning: a) front view; b) top view.

To implement the validation method, it is required to define a pair of target points, *A* and *B*, with a reference differential depth. If the target grid pattern is placed parallel to the baseline, *A* and *B* will have a zero differential depth. However, by tilting the pattern with the angle *α*, as shown in Fig. 7a and 7b, a different reference differential depth is obtained. The tilting angle, *α*, is adjusted with respect to a reference line that is parallel to the baseline of the stereo system. If the distance between the two points on the grid, *L*, and the angle *α* are known, the reference differential depth Z_{AB} can be directly obtained by: $Z_{AB} = L \times \sin(\alpha)$.

The experiment was conducted for three different values of the tilting angle α: 0°, 26.6° and 45°. In addition to this, two target pairs were used. The distance between the two target points on the grid *L* was set to 100 mm and 150 mm for the first and the second pair, respectively. The two target pairs were in the common field of view of the stereo system.

The baseline B is equal to 100 mm, and the grid of targets is at a distance of approximately 1600 mm from the baseline. For the camera prototype described in 4.3, the length of the pixel size has been calibrated to 8.33 µm, and the number of motor steps required to shift the camera module half a pixel size is found to be 69 motor steps. The focal length is 25 mm.

Table 1 presents the results of the validation experiment for the two pairs of target points with the three different titling angles. The reference differential depth for each pair at each angle is compared with the reconstructed one by use of both the direct method and the dithering algorithm. The absolute error of the differential depth reconstruction for each method with respect to the reference differential depth is also listed for each case. This error can be estimated as a convolution of quantization errors of the target pair where these errors depend on the distance of each target point from the baseline. The distances cannot be measured exactly but it is possible to determine the minimum distance value to be 1560 mm.

www.czasopisma.pan.pl   PAN   www.journals.pan.pl
POLSKA AKADEMIA NAUK

Metrol. Meas. Syst., Vol. XVII (2010), No. 3, pp. 335–348

From this measure, it is possible to estimate that the quantization error of the target pairs is at least 4.7 mm which corresponds to $n = 253$. Then, the maximum absolute error is at least 9.4 mm for the reconstructed differential depth by the direct method in this setup, and 4.7 mm for the reconstructed differential depth by the dithering method.

From Table 1, it can be noticed that the reconstruction error by the use of the dithering method is about half of that of the direct method. The mean of the absolute reconstruction errors for each method, calculated from the table, is 5.1 mm for the direct method, and 2.6 mm for the dithering algorithm. The improvement in the depth reconstruction accuracy in this experiment is thus 49%.

Table 1. Results of the differential depth reconstruction showing errors by the direct method and the dithering algorithm for the target pairs.

Angle [degree]	Line length [mm]	Reference differential depth distance Z [mm]	The direct reconstruction method		The dithering method	
			Reconstructed Z_d [mm]	Absolute reconstruction error [mm]	Reconstructed Z_m [mm]	Absolute reconstruction error [mm]
0	100.0	0.0	7.4	7.4	3.7	3.7
	150.0	0.0	7.3	7.3	3.7	3.7
26.6	100.0	44.7	47.7	2.7	43.5	1.2
	150.0	67.1	61.0	6.1	63.9	3.2
45	100.0	70.7	68.8	2.0	71.7	1.0
	150.0	106.1	111.2	5.1	109.9	2.8

5. Conclusion

This paper introduces the use of sensor-shifted cameras in the stereo system instead of the conventional cameras and applies the dithering algorithm to improve depth reconstruction. Both the synthetic and the physical experiments verify that applying dithering reduces the depth reconstruction uncertainty by half when compared with the direct method.

It is verified by simulation that the dither signal causes the iso-disparity surfaces of the skewed-parallel stereo camera to shift with respect to their initial positions. The new position is the middle of the iso-disparity intervals before applying the dither signal. This can be interpreted to mean that dithering can be applied to reduce the depth reconstruction uncertainty by half through a multi-stage measurement, *i.e.* the dithering algorithm.

It is also found that the dither signal is equal to half the pixel size of the camera sensor regardless of the first estimation of the depth of the target. Thus, it can be concluded that no primary measurement is needed for the dithering algorithm, which in turn means that it significantly simplifies the use of this kind of camera when determining depth.

In the synthetic experiment, the enhancement was verified using computer simulation. The depth reconstruction of the test target points using the dithering algorithm shows improvement of the uncertainty since the reconstruction of the points forms iso-disparity lines with interval widths reduced by half for the dithering algorithm when compared to the direct method. This improvement can be confirmed by comparing the histograms of the reconstruction error produced by both methods. This comparison shows that the dithering algorithm reduces the span of the error distribution to half the range obtained with the direct method. The depth reconstruction improvement can also be observed in the 48.6% reduction in the standard deviation of the reconstruction error for the simulation targets by the dithering algorithm.

The results of the physical experiment also show reconstruction improvement for all the test targets with an average of 49%, which is close to the theoretical value. For higher experimental accuracy, the differential depth of the target pairs was used in the validation experiment instead of the absolute depth of a single target.

The proposed method can be applied to measure a depth of objects with structural surfaces. Furthermore, the dynamics of the measured objects is limited by the speed of the camera movement. However, it can be extended by using a variable opacity optical attenuation mask directly in front of a camera lens [21].

For further research, the dithering approach can be applied with more steps in order to reduce the uncertainty to less than half.

Acknowledgements

References

[1] W. Kulesza, J. Chen, S. Khatibi: "Arrangement of a Multi Stereo Visual Sensor System for a Human Activities Space". Ed. A. Bhatti, *Stereo Vision*. InTech Education and Publishing, Vienna, 2008, pp. 153–172.

[2] J. Chen, S. Khatibi, W. Kulesza: "Planning of a Multi Stereo Visual Sensor System for a Human Activities Space". *Processing of 2nd International Conference on Computer Vision Theory and Applications*, 2007, pp. 480–485.

[3] J. Chen, S. Khatibi, J. Wirandi, W. Kulesza: "Planning of a Multi Stereo Visual Sensor System for a Human Activities Space – Aspects of Iso-disparity Surface". *Proc. of SPIE on Optics and Photonics in Security and Defence*, vol. 6739, Florence, Italy, Sept., 2007.

[4] R. Hartley, A. Zisserman: *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.

[5] T. Chen, P. Catrysse, A. Gamal, B. Wandell: "How Small Should Pixel Size Be?". *Proc. of SPIE on Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications*, vol. 3965, 2000.

[6] H. Sahabi, A. Basu: "Analysis of Error in Depth Perception with Vergence and Spatially." *Comput. Vis. Image Und.*, vol. 63, no. 3, 1996, pp. 447–461.

[7] J. Chen, S. Khatibi, W. Kulesza: "Depth Reconstruction Uncertainty Analysis and Improvement – the Dithering Approach". *Elsevier Journal of Image and Vision Computing*, vol. 28, no. 9, 2010, pp. 1377–1385.

[8] C. Alasseur, A Constantinides, L. Husson: "Colour Quantisation Through Dithering Techniques". *IEEE International Conference on Image Processing*, vol. 1, 2003, pp. I–469–72.

[9] F. Corrêa Alegria: "Contribution of jitter to the error of amplitude estimation of a sinusoidal signal". *Metrol. Meas. Syst.,* vol. XVI, no. 3, 2009, pp. 465–478.

[10] *Two new perspective control wide-angle lenses*. Canon Technical Hall, *http://www.canon.com/cameramuseum/tech/report/200907/report.html.*

[11] A. Francisco, F. Bergholm: "On the Importance of Being Asymmetric in Stereopsis-or Why We Should Use Skewed Parallel Cameras". *Int. J. Comput. Vis.*, vol. 29, no. 3, 1998, pp. 181–202.

[12] M. Ben-Ezra, A. Zomet, S.K. Nayar: "Video Super-Resolution Using Controlled Subpixel Detector Shifts". *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, 2005, pp. 977–987.

[13] A. Siddig: *Depth Reconstruction Uncertainty Improvement for Skewed Parallel Stereo Pair Cameras Using Dithering Approach*. Master Thesis, Dept. of Signal Processing, Blekinge Inst. of Tech., Karlskrona, Sweden, 2010. (submitted)

[14] *Image Processing Toolbox 7 User's Guide*. The Math Works Inc., 1993, *http://www.mathworks.com/access/helpdesk/help/pdf_doc/images/image_tb.pdf*

[15] G.L. Mariottini, D. Prattichizzo: "EGT for Multiple View Geometry and Visual Servoing: Robotics Vision with Pinhole and Panoramic Cameras". *IEEE Robotics & Automation Magazine*, vol. 12, no. 4, 2005, pp. 26–39.

[16] *CCD Color Video Camera Module*. Sony Corp. ,2002, *http://pro.sony.com/-bbsccms/assets/files/mkt/indauto/Brochures/xc-555_techmanual.pdf*

[17] *Operating Manual MS 38E C-832 DC Motor Controller.* Physik Instrumente (PI) GmbH & Co., 1996, *http://www.physikinstrumente.net/ftpservice/Motor_Controllers/XXX__Oldcontrollers/C-832.DC-MotorController/C832.OperatingManual/MS38E280.pdf*

[18] *2/3 25MM F/1.4 with Lock for MegaPixel Camera*. Tamron Inc., *http://www.tamron.com/cctv/prod/23fm25sp.asp*

[19] W. Mustafa: *Depth Measurement Improvement Using Dithering Method in Sensor-shifted Stereo Cameras*. Master Thesis, Dept. of Signal Processing, Blekinge Inst. of Tech., Karlskrona, Sweden, 2010.

[20] J. Chen: "The Depth Reconstruction Accuracy in a Stereo Vision System". *XLI Intercollegiate Metrology Conference*, Gdansk, Poland, 2009.

[21] H. Farid: *Range Estimation by Optical Differentiation.* Ph.D. Dissertation, University of Pennsylvania, Philadelphia, USA, 1997.