Pyramid Imaging Inc
945 East 11th Avenue
Tampa, FL 33605
https://pyramidimaging.com
Sales@pyramidimaging.com
(813) 984-0125

Pyramid Imaging

chromasens
Imaging for Professionals

3DPIXA: Limitations of block-matching algorithms on depth reconstruction

Whitepaper

Authors
Sebastian Georgi
Timo Eckhard

Executive Summary:

The aim of this work is to provide a comprehensive overview on the limitations of the block-matching approach in height reconstruction for passive stereoscopy. It describes a variety of different artefacts that appear at certain image features. The occurrence of the artefacts is explained by means of the algorithms underlying mathematical model which is introduced in the first chapter. Images captured with a 3DPIXA exemplify the visual appearance. Additionally, we use simulated stereo images where appropriate to illustrate the difference between the calculated height and the ground truth from simulation.

1. Introduction

The height measurement of a stereoscopic 3D system is based on the fact that features in the images of two cameras with distinct perspective view on the scene are shifted relative to each other. The magnitude of the shift can be used to calculate the distance of the stereo camera to the feature. Based on geometric principles one can show that this shift is along one line only which is called the epipolar line. In case of the Chromasens 3DPIXA product line, the epipolar line is along the x coordinate of the camera images. The 3D stereo measurement therefore relies on a reliable method to find x coordinates of corresponding features in the two camera images. A necessary preprocessing step for this concept is correcting for any geometric distortions in the stereo images, called rectification. All concepts discussed here assume the images are already rectified. Also, one camera has to be arbitrarily chosen as the reference camera relative to which the shift is measured which is the right image for 3DPIXA cameras.

The 3DPIXA uses a method based on block-matching to find these correspondences. The algorithm loops over each pixel of the reference camera image and uses the vicinity of this pixel (a “block” or “window”) as a template. In the other camera image the vicinities of all the pixels that are on the same y-coordinate and at any x-coordinate in a defined search range are compared to this template block. For each pair a value expressing the similarity between the two blocks is calculated with an appropriate mathematical function. Chromasens 3D algorithm uses the normalized cross correlation value between the two blocks to calculate this similarity.

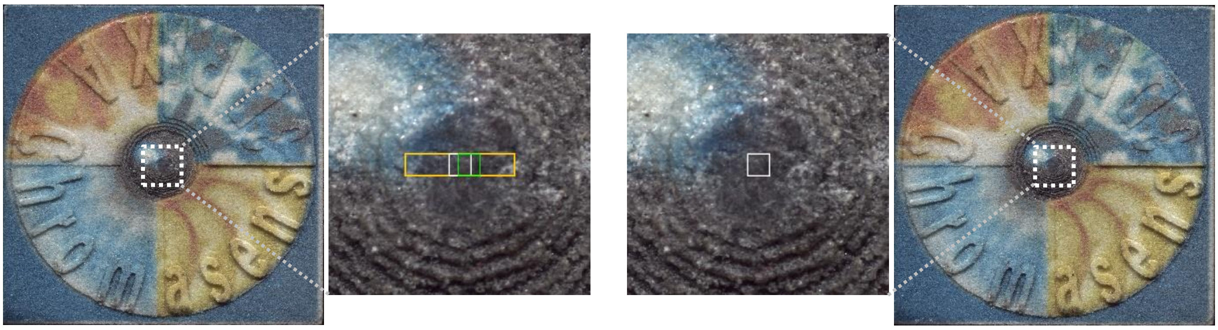

Figure 1: Reference camera image (right) and comparative camera image (left) of a 3DPIXA camera with a detail cutout of the middle section. The template block in the right detail image is marked in white. The same position is marked in white in the left image as well as the search range in yellow and the best match position in green.

Figure 1 visualizes the block-matching concept using images of a Chromasens illustration target, captured with a 10µm resolution 3DPIXA (CP000600-C01-010-0056). The template for one pixel of the image is indicated by a white rectangle in the detail image of the right camera. At the same position in the left image the template block is also depicted but due to the stereo perspective the image content in this block is different. A similarity value between template and comparative block is calculated for all blocks of the left image that have a certain distance to the template position which is marked in yellow in the image. The best match (highest similarity value) to the template in the left image is instead at the position indicated with a green rectangle. The distance between the white and green block is called the disparity value. Through calibration of the camera the distance of the feature in the template to the camera can be calculated from this disparity value. The set of x coordinates that is checked in the comparative image is called the (disparity) search range which can be directly converted to the available height range of this particular camera. Note that Chromasens 3DPIXA cameras are built to have the best overlap between left and right image in the focal plane. Consequently, there is zero disparity between the (rectified) images everywhere in the focal plane and objects even further away from the camera have negative disparities.

Block-matching has proven useful for the height measurement on surfaces with a distinctive texture and only smooth height change. However, the algorithm intrinsically assumes that the object height relative to the camera is constant within the template window size. Furthermore, the similarity measure can only find the best match if the image content actually appears similar in left and right perspective views of the stereo image scene. In practice, there are often areas in images where these strong requirements are not met, resulting in certain effects on the reconstructed height image. The aim of this whitepaper is to describe the cause of different height estimation artifacts that are visible at distinctive image features. To accurately convey the causes of the artifacts it is necessary to first introduce the concept of the cost volume which is the topic of the next chapter.

2. Cost volume

As explained in the introduction the block-matching algorithm is based on a suitable similarity metric which calculates one single value for the comparison of two image blocks of a fixed size. This value is usually expressed as a “cost” which means it is normalized in such a way that it is zero if the two blocks are identical and one when there is no similarity between blocks.

The result of performing block-matching on the entire image pair is the so-called cost volume. It contains one cost value for each x and y position and for each possible x coordinate displacement value between template and comparative bock within the defined search range. The two input images need to have the same size which is then also the size of the cost volume in the x and y dimension. The size in the displacement dimension is the length of the search range which the user can define. The selection of an appropriate search range is crucial for ensuring that the whole height range needed for the application is covered.

For each x and y position there are as many cost values as there are integer displacements in the search range. The displacement value with the lowest associated cost value represents the best possible match which is the disparity that is used to determine the height for this position. A more precise value can be calculated through sub-pixel interpolation between the costs neighboring the minimal value along the displacement dimension.

The reason for height artifacts appearing in block-matching can often be deduced through careful examination of the cost volume. But as the cost volume is a three-dimensional matrix of values only sections of it can be visualized. In the following graphs we generally show the cost volume as a 2D color map with one image coordinate and the displacement as axes and the cost value represented by a color scale. Whenever possible, simulated stereoscopic images were used for which a ground truth disparity (without any matching artefacts) is known.

Figure 2 shows a composition of different views for a simulation of a slightly sloped plane. The top left image shows the simulated image of the right camera image. The texture used for this simulated image is a random pattern with a cutoff at high frequencies which gives very good and unambiguous correlation results. Next to this image are two disparity maps which encode the ground truth disparity value (middle) and the reconstructed disparity value(right) in a false color representation for each x and y position. The disparity maps coincide pixelwise to the right camera image because we set up the algorithm such that the template block for each position is taken from the right image.

The lower left image is a cost volume representation for a section along the plane where zero cost (maximum similarity) is encoded as white color. On the right the calculated disparity value and ground truth disparity are plotted for the same section. From these two plots it is clearly visible how the calculated disparity follows along the minimum of the cost volume.

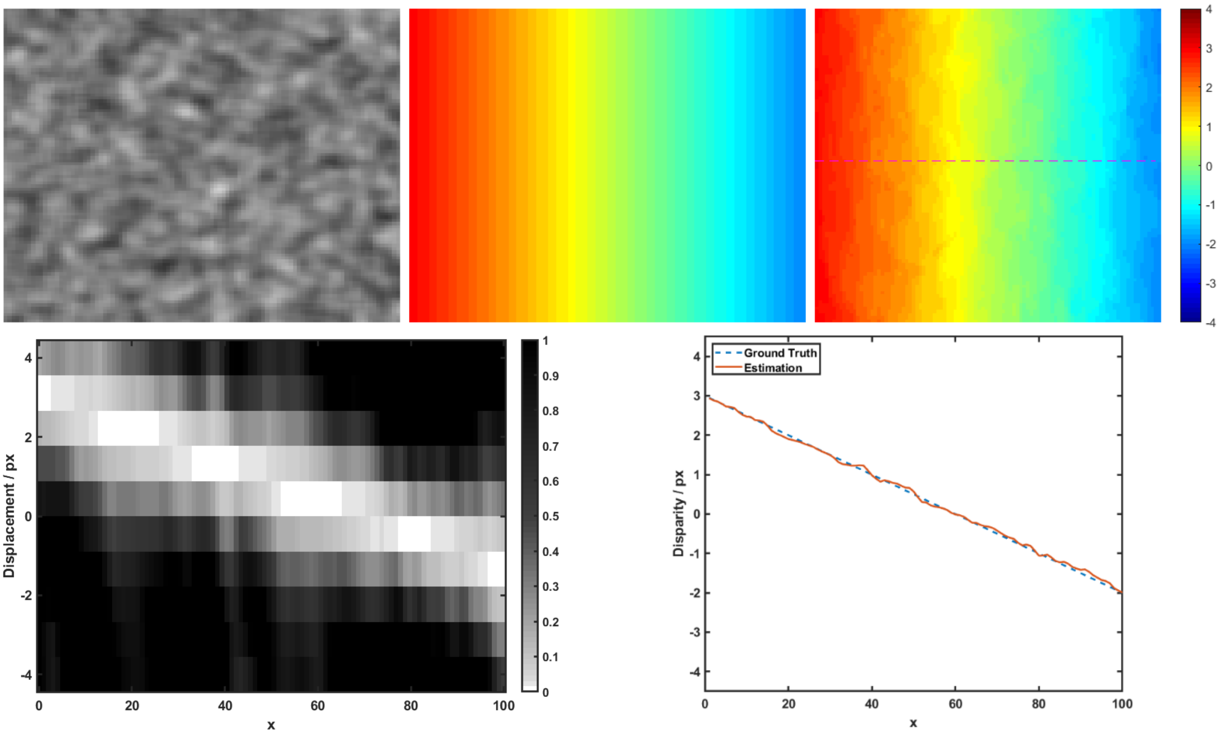

Figure 2: Top left: Simulated right camera image. **Top middle:** Ground truth disparity map. **Top right:** Reconstructed disparity map with magenta dashed line indicating the detailed view in the bottom of this figure. **Bottom left:** Section of cost volume corresponding to the magenta dashed line. **Bottom right:** ground truth and block-matching result.

This slightly sloped plane is an image feature where no artefacts appear due to the block-matching algorithm which is why the estimated height and the ground truth are both nearly identical. Other height features or different kind of textures can lead to the artefacts in the estimated height. The following chapters will give a comprehensive overview off all these artefacts with example scans and explaining the reason for their appearance using the cost volume concept.

3. Search window containing a height discontinuity

The basic assumption of the block-matching approach is that all the features contained in the template block are at the same height. For smooth surfaces this is usually fulfilled to a sufficient degree even if the surface is slightly tilted. But if the window contains a height discontinuity this assumption is no longer true and height artifacts appear. As the epipolar line is along the x direction height discontinuities will behave differently if they are also along the x axis (horizontal) than for any other case. The two cases are therefore described separately.

A height step will always consist of a lower and an upper level that meet at the edge. These levels can differ in the amount of visible texture as well as in the apparent brightness. Each of those two cases leads to different kinds of height artifacts so we show each case separately for both types of edges described before.

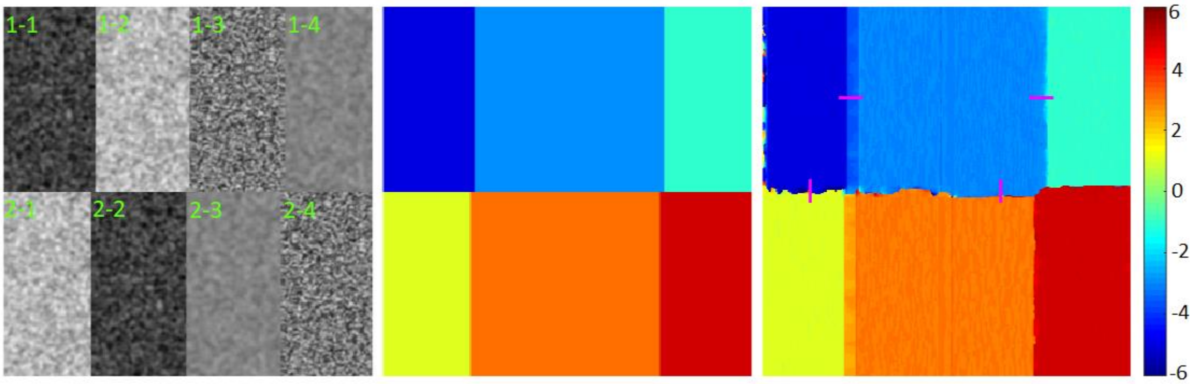

Figure 3: Overview of the height step simulation with fields for different brightness and texture. **Left:** Simulated right camera image. **Middle:** Ground truth disparity map. **Right:** Reconstructed disparity map with the four sections used for the evaluation shown in magenta.

Figure 3 illustrates the simulation we will use to showcase the different effects. On the left half of the image are height steps with different brightness and similar texture and on the right half are fields with different visible texture but similar brightness. In each half we will look at the cost volume of one horizontal and one vertical section to explain why the edges of the different height steps look different in the height image depending on the texture.

3.1 Horizontal edges

The first horizontal height step we will discuss here is the one where the brightness of the two levels is strongly different. These edges will generally appear “ragged” and may contain some extreme outliers. Figure 4 shows an example image of a black carton on top of white paper captured with a 3DPIXA where this behavior is visible.

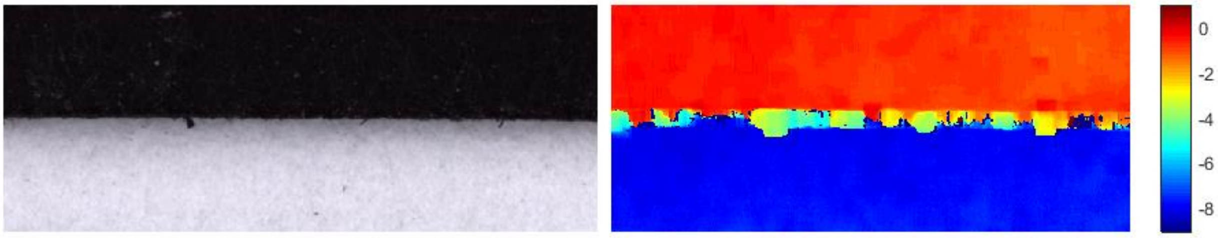

Figure 4 Right camera image and height map for a horizontal edge with brightness difference between the levels captured with a 3DPIXA

The content of a template block at this edge is partly rather dark and quite bright in the other part. This is still largely true when moving the block along the disparity search range. This will cause the block-matching to have a very low cost value for all displacement values as long as this contrast jump (which coincides to the height edge) is contained in the matching window. Due to the uniformly low values the minimum is not well defined and might even take any value that is contained in the search range. This behavior can be seen in the y axis section of the cost volume from the simulation (between ROI 1-1 and ROI 2-1) which is shown in figure 5. Please note that the color bar has been adjusted to a smaller range to show the small cost differences in the edge region.

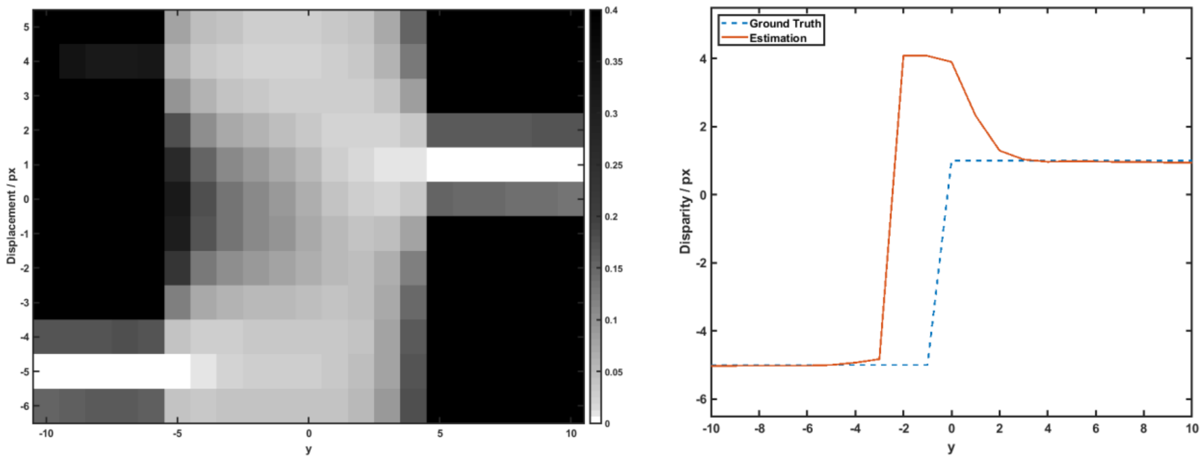

Figure 5: Left: Cost volume section along y axis for a horizontal step with different brightnesses between top and bottom level. The height step is at zero position and the block-matching window size is 11 pixels. **Right:** simulation ground truth and block-matching result.

The second case we want to elucidate here is the one on the right half of the simulated images where one of the two levels contains a much more pronounced texture than the other and the brightness is very similar. A common example for is case would be a height step with the same material at both levels but with one level near the focal plane of the camera and the other level further away (outside the depth of field).

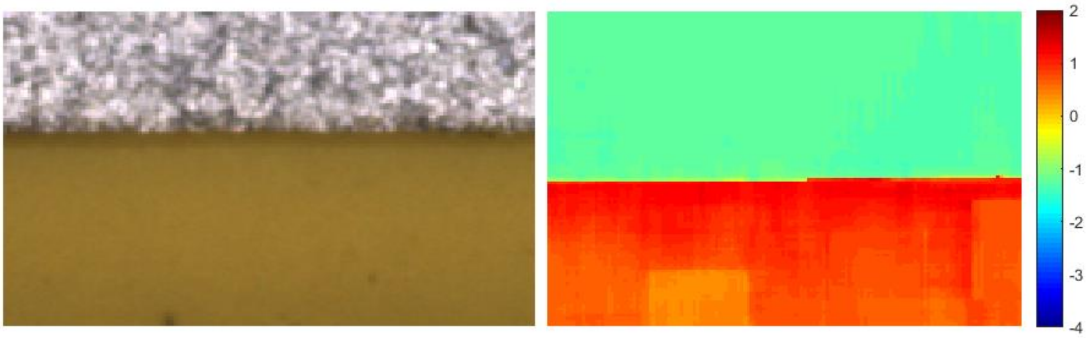

Figure 6 Right camera image and height map for a horizontal edge with a strong difference of visible texture between the levels captured with a 3DPIXA

In figure 6 we show a different example with a more pronounced texture difference example consisting of sandblasted aluminum and an adhesive tape attached to it. One can see that the edge position in the height image is shifted relative to the position of the edge in the camera image.

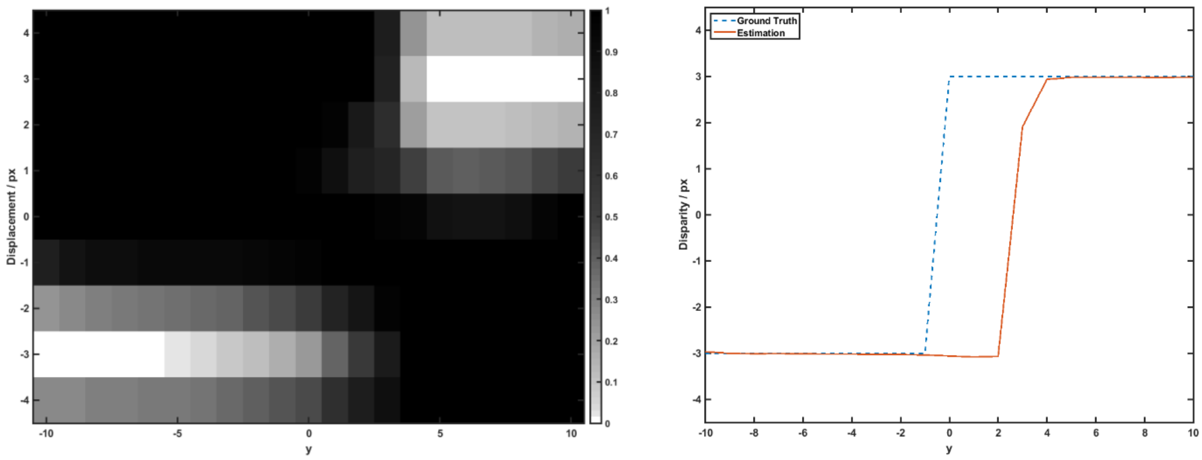

Figure 7: Left: Cost volume section along y axis for a horizontal step with different visible texture between top and bottom level. The height step is at zero position and the block-matching window size is 11 pixels. **Right:** simulation ground truth and block-matching result

Figure 7 shows the cost volume section and the corresponding height estimation for the relevant edge in the simulation (between ROI 1-3 and ROI 2-3). As can be seen in the overview in figure 3 for this section the lower level shows much more texture than the upper level. The level with the increased visible texture will influence the correlation value much more than the lower value. So even if only a few lines of the higher texture are contained within the window the cost already shows a higher similarity for this disparity value. This effectively increases the area where the height of the upper level is estimated beyond the real boundaries (but will not introduce additional outliers). The outline of the edge in the height image will therefore be similar to the real shape of the edge but shifted away from the higher textured level. As this effect is highly deterministic there is the possibility to correct for it using only the contrast information from the camera image and the height map. One additional effect not shown here that complicates this approach is that the effect is swapped if there are strong (local) increases in texture (e.g. a small contamination particle) on the low contrast level. The edge shift in the vicinity of the particle is therefore also in the other direction leading to “indentations” in the edge shape.

In a real-world application there can also be a strong brightness and texture difference at the same time at height edges. In this case the local estimation will be determined by whichever effect is the most dominant in the current window.

3.2 Non-horizontal edges

We mentioned in the beginning of this chapter that horizontal edges differ from all others as they are along the displacement direction (the epipolar line). Another difference between the edge orientations is that non-horizontal edges always have the sidewall between the levels of the height step visible in one stereo image and in the other image it is always occluded. The right camera captures the sidewalls on the right side while the left camera those on the left.

As there is no correspondence for the sidewall in the other stereo image the estimated height will depend on the brightness and texture of the two levels and the characteristics of the sidewall as well. This will lead to artifacts appearing at reasonably big height steps (>> 1 disparity), but the shape and height deviation will depend entirely on the specifics of the sample. At edges where the sidewall is not visible the maximum size of the artifact is generally equal to the window size and it is centered around the edge position. At edges with visible sidewalls the whole area of the sidewall will show height artifacts. The effect of the sidewall occlusion influences all stereo matching algorithms and not only block matching. Therefore we use very small height steps in the next analyses where the effect of the sidewall becomes negligible and the effects of brightness and texture dominate.

In the case of different visible textures, the vertical height steps behave similar to the horizontal case in that the level with the stronger visible texture is enlarged beyond its border. Figure 8 shows the vertical edge from the same sample shown in the respective horizontal case. In the height map the position of the edge is visibly shifted to the right where the tape with low visible texture is. There are also some outliers in the edge region but those are more likely due to the heightened susceptibility to errors of low texture in general.

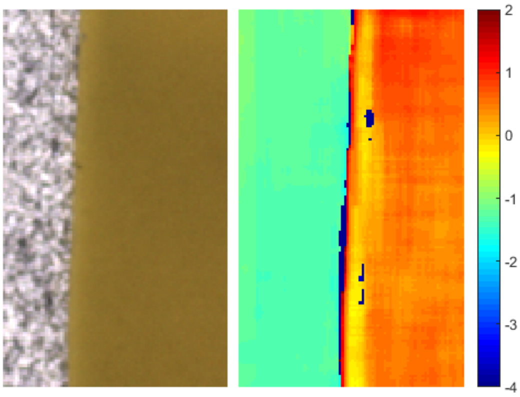

Figure 8 Right camera image and height map for a vertical edge with a strong difference of visible texture between the levels captured with a 3DPIXA

Figure 9 shows the cost volume and height estimation of the respective simulation section (between ROI 1-3 and ROI 1-4). The outline of the edge in the height image is similar to the real shape of the edge but shifted outwards from the higher texture area. So for small height steps and different texture visibility, edges of all directions look very similar.

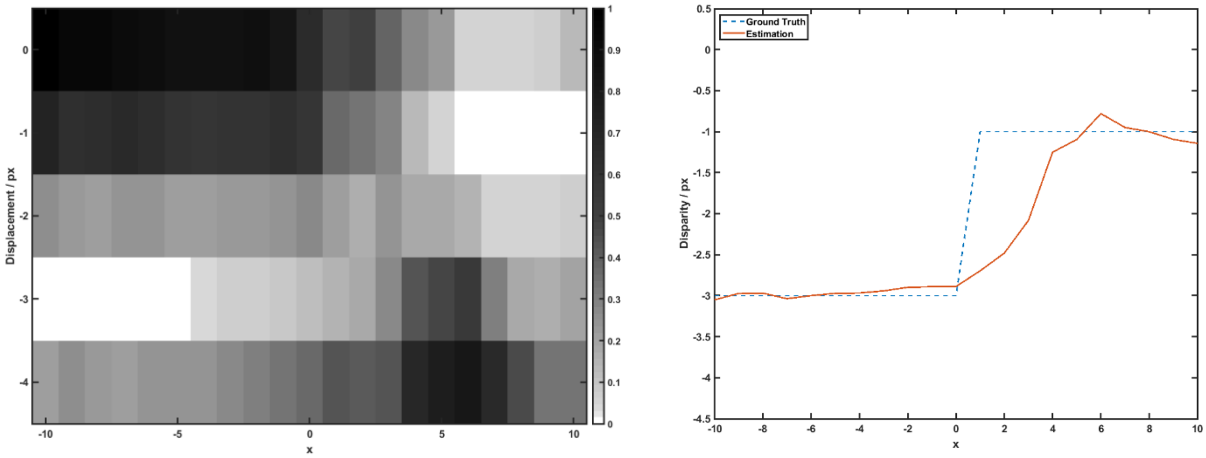

Figure 9: Left: Cost volume section along x axis for a vertical step with different visible texture between top and bottom level. The height step is at zero position and the block-matching window size is 11 pixels. **Right:** simulation ground truth and block-matching result

The remaining case are non-horizontal edges with different brightnesses. Figure 10 shows such an edge captured with a 3DPIXA camera. There is a clearly visible stripe of constant height at the border between the two levels. The height shown for this artefact is neither height from the levels but one in between those.

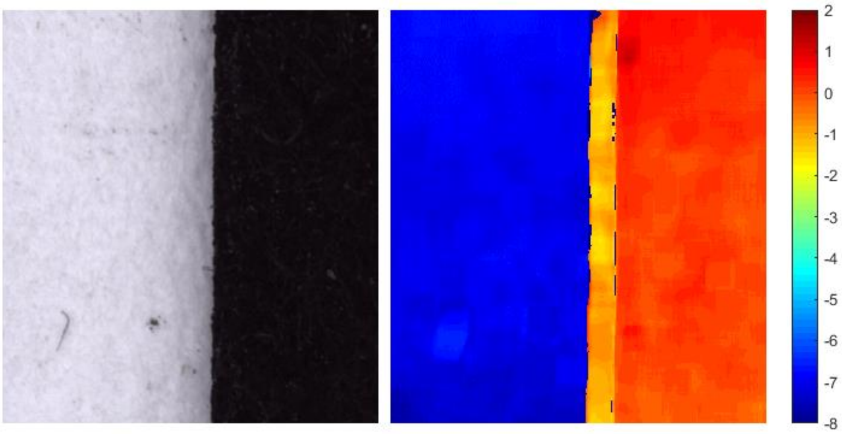

Figure 10: Right camera image and height map for a vertical edge with a strong difference in brightness between the levels captured with a 3DPIXA

Similar to the horizontal case part of any template covering part of the edge will be bright and the other part will be dark. Any block from the left image where the same pixels as in the template will be bright and dark will have a very low cost value nearly independent on any additional texture in the block. So for all templates containing this contrast jump the disparity is locked in to the one with the contrast jump at the same position which is always one that is in between the disparity of the two levels.

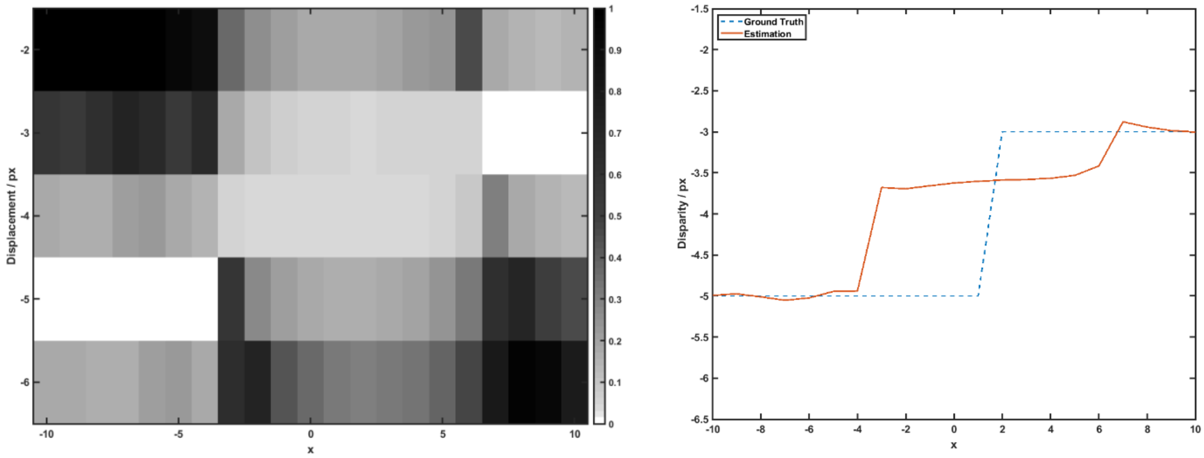

Figure 11:Left: Cost volume section along x axis for a vertical step with different brightness between top and bottom level. The height step is at zero position and the block-matching window size is 11 pixels. **Right:** simulation ground truth and block-matching result

Figure 11 shows the cost volume and height estimation of the respective simulation section (between ROI 1-1 and ROI 1-2). The values of the costs along the displacement direction is nearly constant for all positions where the block overlaps the edge leading to the described effect. This locking effect of a template window containing a pronounced contrast feature is also the reason for the next artefact which shows height artefacts on smooth surfaces.

4. Effects of strong contrasts on smooth surfaces

The height of smooth surfaces without any steps is generally decently reconstructed with the block-matching approach. But at strong contrast features on a flat surface with otherwise homogenous texture artefacts will occur. The size of the contrast feature can be only a few pixels like a dark spot (e.g. dirt stain) on a bright surface or also extended like the border between areas with different brightnesses.

As long as any part of the high-contrast feature is contained within the window the block-matching will calculate the highest similarity at a disparity value where the contrast step is at the same position in both blocks. It is the actual disparity value when the position of the contrast step is in the middle of the template. For a position further away from the middle the correct disparity is already different if the surface is tilted. Figure 12 shows this effect on the example of a tilted plane which has a dark spot at zero position. Besides some edge effects the estimated height basically stays the same in the range of -10 to +10 pixels around the defect which is exactly the size of the window used here.

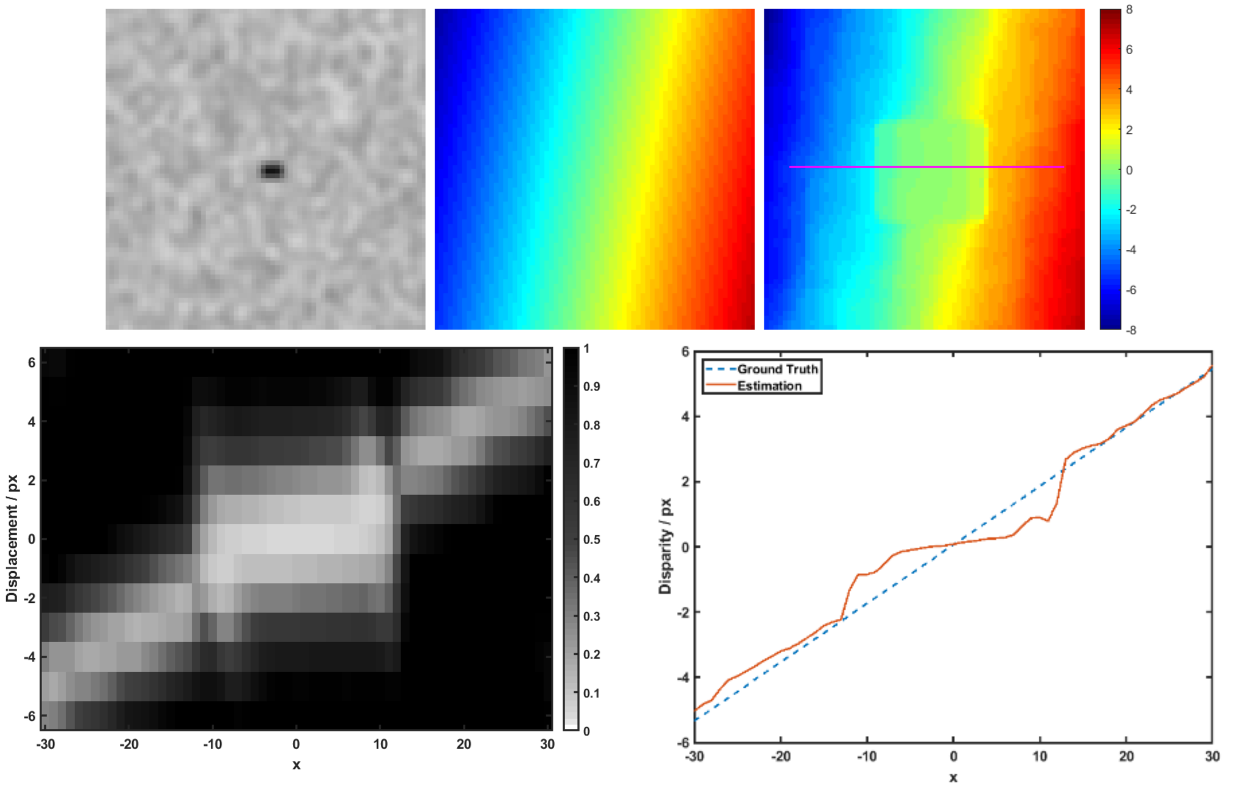

Figure 12:Top left: *Simulated right camera image with contrast defect.* ***Top middle:*** *Ground truth disparity map* ***Top right:*** *Reconstructed disparity map with bottom section shown in magenta.* ***Bottom left:*** *Cost volume section along x.* ***Bottom right:*** *simulation ground truth and block-matching result disparity for the section*

This defect will increase in visibility the higher the slope of the surface, as the height difference at the edges of the window gets bigger. But even on totally flat surfaces the effect can be visible. This is because the estimated height will remain constant as long as the high-contrast feature is in the vicinity of the matching block while the surrounding areas show minuscule height changes due to surface roughness or height estimation variance.

5. Scratches

A scratch is a very narrow defect that has noticeable height changes with steep sidewalls that meet each other without any flat surface at the bottom. A scratch usually has some noticeable length, a defect that is very small in both directions is usually called a pinhole. Pinhole defects basically have the same challenges as scratches for block-matching and as such the concepts discussed here directly transfer to them as well even if they are not explicitly named. Both defects cannot be estimated correctly with block-matching if their size (in any dimension) is smaller than a reasonable window size for block matching (7x7 pixels or more depending on the texture).

This chapter shows depth estimations for a simulated horizontal and vertical scratch and explains the problems that prevent a reasonable height estimation. Figure 13 shows the simulation image and the disparity result for the two different scratches. In this chapter we will also discuss some real-world application influences that cannot be simulated but nevertheless have additional negative effects on the height estimation of scratches.

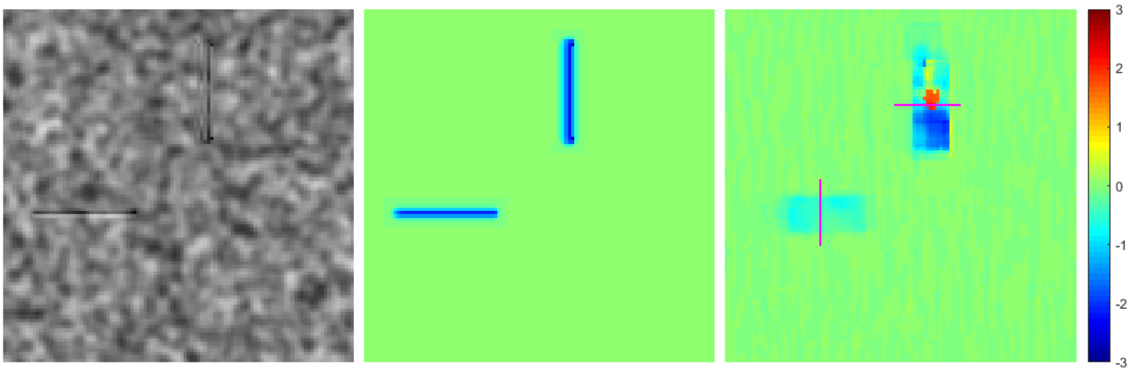

Figure 13: Left: Simulated right camera image with scratch defects. **Middle:** Ground truth false color disparity map. **Right:** Block matching disparity estimation with the analyzed cross sections marked in magenta

The easiest scratch to measure is a horizontal one. In figure 14 we show the cost volume and disparity plot for the simulated scratch. The estimated height here is much shallower than the actual depth of the scratch. In this case the template window containing the horizontal scratch will consists of only a few lines in the block that differ from the surrounding disparity of the pristine surface. As there are only a few lines they also only have a slight influence on the calculated cost values of the whole block (also containing the surrounding) which leads to the shallow depth estimation. Additionally, the size of the scratch gets overestimated due to the enlarging effect of the block matching.

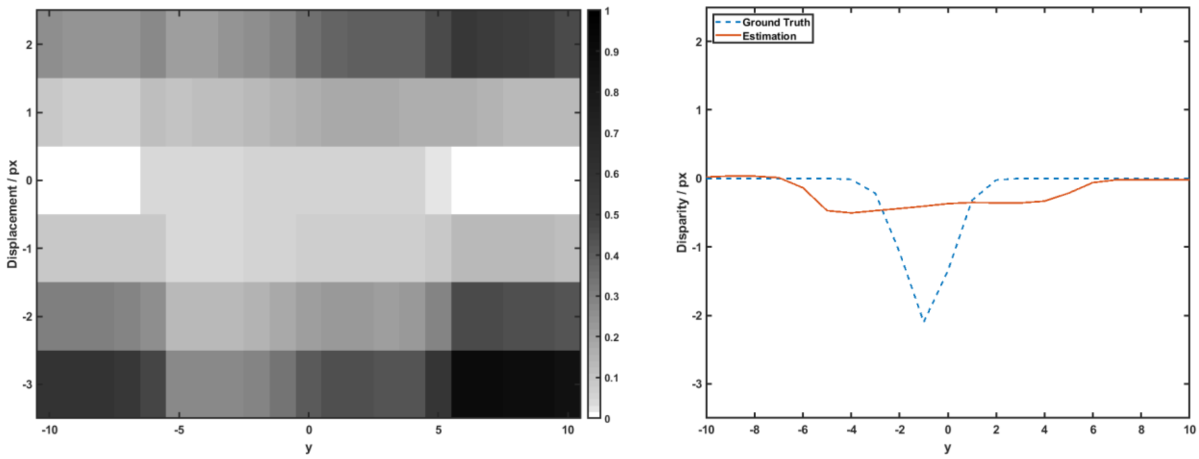

Figure 14: **Left:** Cost volume section along y axis crossing a horizontal scratch with a block-matching window size of 11 pixels. **Right:** simulation ground truth and block-matching result

While the results of the horizontal scratch may still be reasonable a vertical scratch can produce height artifacts which estimate any disparity value as shown in figure 15. These scratches have steep sidewalls that are each only visible to one camera while they are occluded to the other. Analogous to the big height steps in vertical direction discussed in chapter 3.2 the result of the block matching is unpredictable as it strongly depends on the appearance of the side walls. For very thin scratches no height change at all is often estimated while bigger scratches are prone to show extreme outliers.

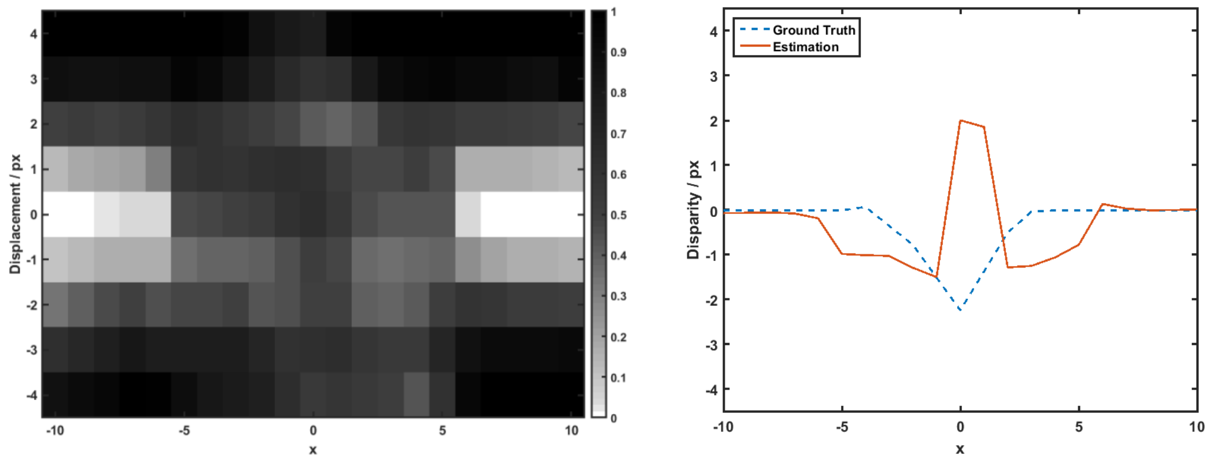

Figure 15: **Left:** Cost volume section along x axis crossing a vertical scratch with a block-matching window size of 11 pixels. **Right:** simulation ground truth and block-matching result

In a real-world application there is the additional complication that the two perspectives will appear with different brightness, depending on the angle-dependent reflection properties of the sample. Steep edges might appear bright in one perspective as they are near to the gloss angle while the other perspective is much darker. This difference will lead to significant outliers in the height estimation. Another common problem is that scratches often have a small elevation at the sides where the material from the trench is moved to during the scratching process. Depending on the texture of these bulges the estimated height of the whole scratch can be higher than the base height.

Due to the described problems we strongly discourage to search for scratch defects in the height image using block matching reconstruction. The identification of scratches is a task that can often be handled more easily on the color images provided by the stereo cameras given a suitable illumination setup.

6. Mismatching due to repeating pattern structures

The basic assumption for the block-matching algorithm is that two windows with image content are the most similar to each other if they show the same region of the sample. Implied in this statement is that the similarity needs to decrease whenever blocks from different sample regions are compared to each other. This assumption is not necessarily true if the samples texture shows a pattern that is repeating along the displacement direction. As no global information is considered by the block-matching algorithm the repetitions can lead to mix-ups which produce outliers in the disparity estimation.

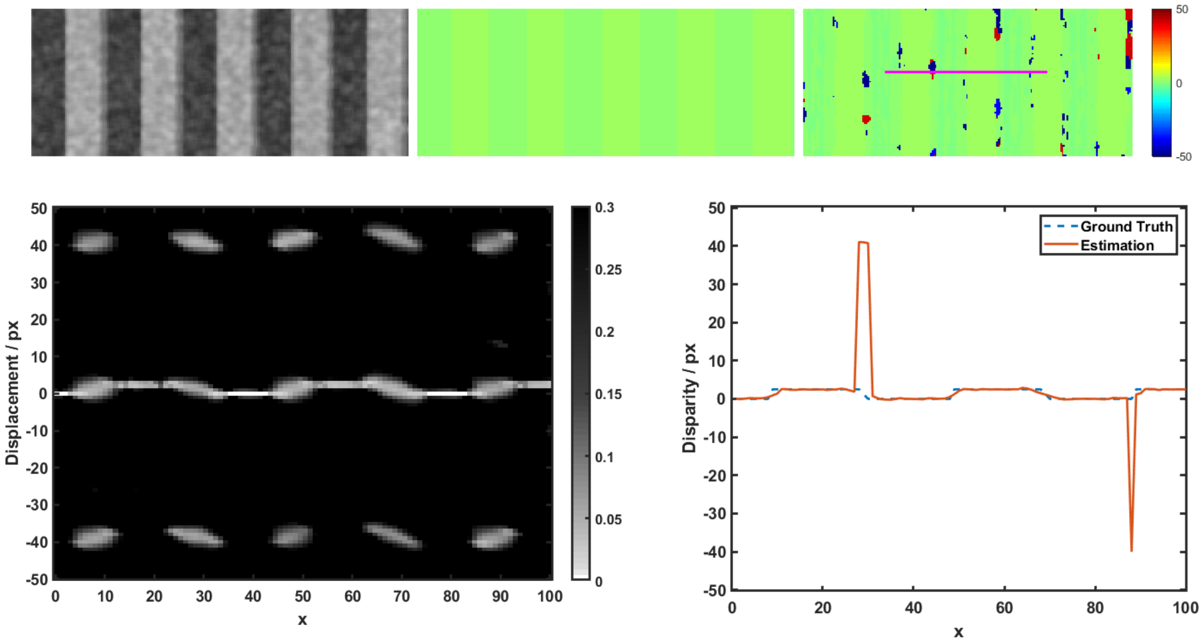

Figure 16:Top left: Simulated right camera image with repeating line pattern **Top middle:** Ground truth disparity map **Top right:** Reconstructed disparity map with bottom section shown in magenta **Bottom left:** Cost volume section along x **Bottom right:** simulation ground truth and block-matching result disparity for the section

Figure 16 shows the cost volume and height estimation for a pattern of vertical lines with a line distance of 40 pixels. The lines are alternating dark and bright and are overlaid with additive noise. At the edges of the lines the strong contrast of the brightness change dominates the cost calculation. As this is the same at each repetition of the pattern the cost volume shows high similarity at -40 and +40 pixels relative to the actual disparity. Due to random fluctuations any one of the edges can have the lowest cost value which leads to outliers in the disparity that are offset to the correct disparity by a multiple of the pattern period. There is an effective way to avoid this height artefact occurring through reducing the disparity search range to a value smaller than the pattern period. But this is not always possible especially if the sample height covers a big range.

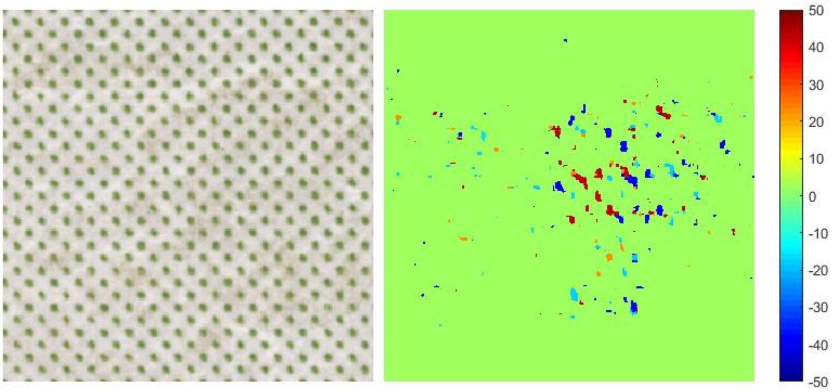

Figure 17: Left: Right camera image of a packaging cardboard detail with a repeating point pattern **Right:** Block matching disparity reconstruction showing repetition artefacts

The pattern chosen here also incorporates actual height changes whenever the pattern changes which in this case are steps that are bigger than the window size. But the general repetition artefact discussed here will also occur if the pattern period is smaller than the window size and also on a flat plane that has a repeating pattern only as its texture. Additionally, the pattern can theoretically be anything that is repeating along the direction of the disparity search range. In real-world applications the most common type of repeating structures are actually vertical lines e.g. several vertical wirebonds next to each other, conductor tracks on PCBS or milling tracks on machine parts. Other repeating structures are often found in the printing industry, for example on half-tone printed packaging cardboards. An example of such a sample and the height map containing the artefacts from the repetition is shown in figure 17.

7. Limitations on surface tilt

It was mentioned in the introduction that the principle of block matching assumes that there is only a flat surface contained within the matching window. Additionally, we mentioned that tilted or rounded surfaces without any height steps are also possible to be matched without artefacts in most cases. But with increasing tilt of the surface relative to the camera normal the height estimation will get gradually worse and above a certain threshold there will be height artefacts appearing as well. In this chapter we look at the two axes (x and y) along which the surface might be tilted separately as well as the influence of the window size on the quality of matching.

A tilt along x (scan-line direction) means that one side of the object is nearer to the camera than the other – resulting in magnification changes of the image system along this dimension. Depending on the tilt angle one camera image is stretched while for the other camera the effect is reversed, and all pixels are squeezed together tighter as demonstrated in figure 18.

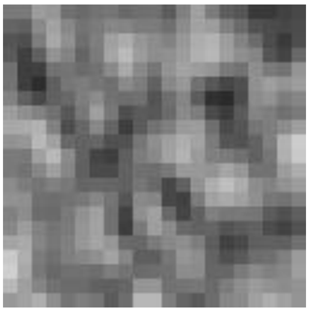

Figure 18: Matching window from left and right camera for a plane tilted by -0.5 disparities per pixel along the x axis

At the correct matching position, the row in the middle of the window is the same in both camera images but they are opposingly distorted at the edges. Naturally the distortion will decrease the calculated similarity between the two blocks. A reduced similarity value means that other random matches can have a lower cost value which will lead to the appearance of outliers which are height values that are completely wrong.

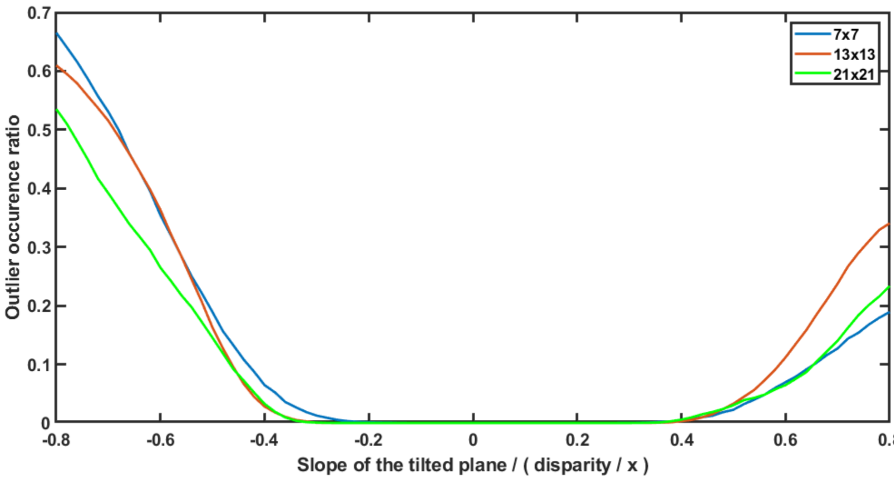

Figure 19: Outlier occurrence over slope along x direction for different matching window sizes.

On flat surfaces with low texture smaller window sizes are much more likely to lead to outliers in the height map as there is less information available. But here the similarity will also decrease with bigger windows as the differences at the edges of the matching window increase with the window size. Figure 19 shows the ration of outliers occurring in the estimated disparity map of simulated image pairs for a plane tilted at different angles along x. There is no noticeable trend difference between the different window sizes which means that the two contrasting effects described before are of similar magnitude effectively cancelling each other.

In addition to the outliers occurring with increasing tilt there is another effect reducing the reliability of disparities estimated for tilted slopes. A block that has a higher contrast feature at one edge will have the lowest cost value shifted to a position that is in between the best match for the middle position of the block and the best match for a pixel at the high contrast edge of the block. These shifts increase the measurement uncertainty on tilted planes, but the estimated height values are still distributed around the best match as they will occur on both edges. This effect is more pronounced with bigger window sizes as the high contrast features can be further away from the best match thereby shifting the minimum further away.

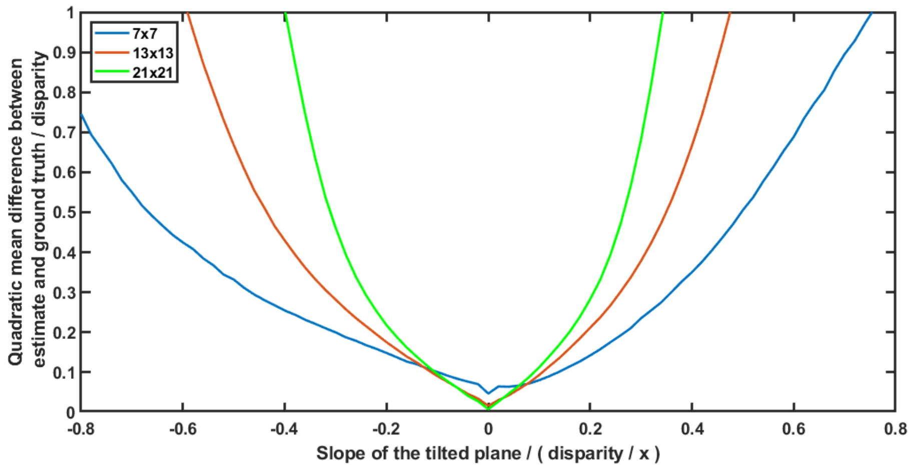

Figure 20: Dependency of the obtainable measurement certainty on the slope of the measured surface along x direction for different matching window sizes. Outliers are filtered out.

Figure 20 shows the plot for the measurement uncertainty which was calculated by the root mean square from the difference between estimate and ground truth for each pixel of the slope with the same simulated images from before. For this measurement the outliers were filtered out. As expected from the considerations above, bigger window sizes yield better results for very low tilt angles but get worse much faster than smaller window sizes. The smaller window sizes are therefore preferred for calculation of image features with steep slopes.

The slopes of the tilted planes against which all of the plots in this chapter are shown is actually the slope when correcting for the parallax effect. We call this perspective the central view. In the right camera image the measured tilt along x $m_{R,X}$ would differ from the central view tilt $m_{C,X}$ by the following relation:

$$m_{R,X} = \frac{m_{C,X}}{1 - {m_{C,X}}/{2}}$$

The same considerations and simulations can also be done for a surface sloped along the y direction which means that the disparity changes between each new captured line. Here the images of the two cameras are vertically sheared relative to each other with the middle row of the matching window being the same in both images at the correct disparity. This is illustrated with windows from a simulated image pair in figure 21.

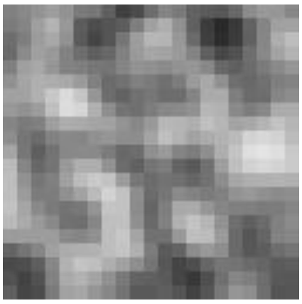

Figure 21: Matching window from left and right camera for a plane tilted by -0.5 disparities per pixel along the y direction

Figure 22 shows the ration of outliers occurring in the estimated disparity map of simulated image pairs for a plane tilted at different angles along y. Contrary to the curves shown for the slope along x this graph has a clearly visible dependency on the window size with bigger windows showing less outliers. But the threshold for any outliers appearing at all is roughly the same at a tilt slope of 0.25.

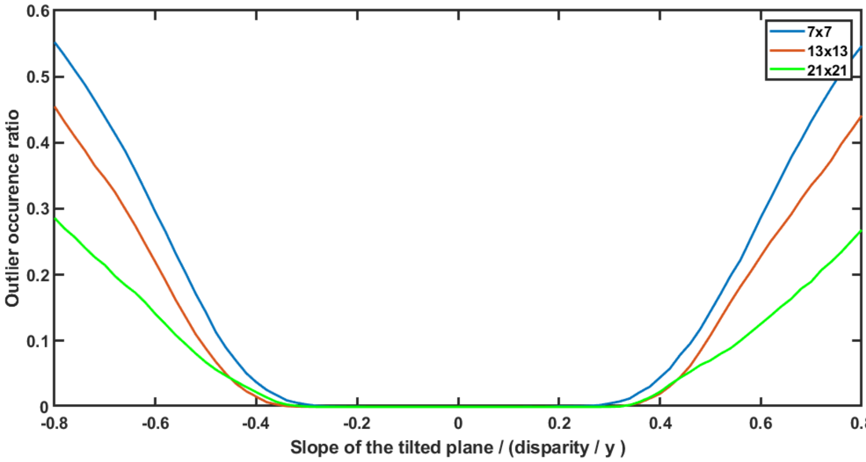

Figure 22: Dependency of the normalized number of outliers occurring in dependence on the slope along y direction of the measured surface for different matching window sizes.

Figure 23 shows the graph for the measurement certainty in dependency on the y slope. It shows similar behavior to the one shown for the x direction in that smaller windows are more precise at steeper slopes.

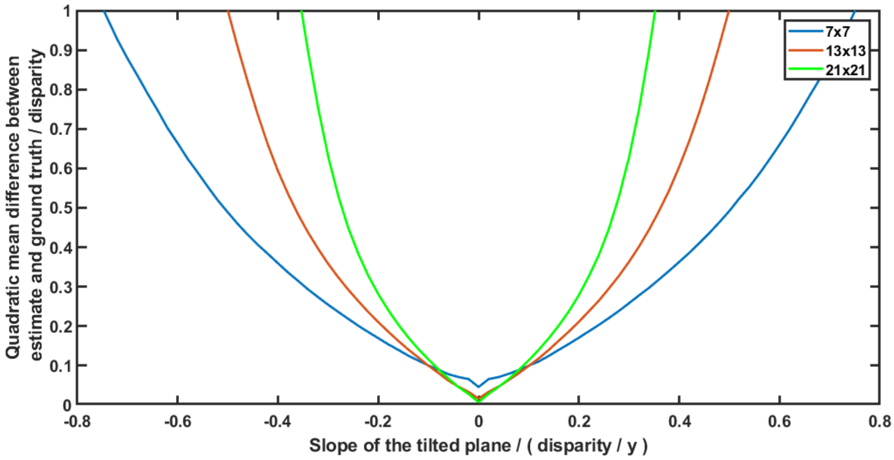

Figure 23: Dependency of the obtainable measurement certainty on the slope of the measured surface along x direction for different matching window sizes. Outliers are filtered out.

A slope tilted in x and y simultaneously will show both the magnification effect and the shearing affect at the same time. The occurrence of both effects at the same time very likely increases the measurement uncertainty even more. But otherwise the general behavior is expected to look qualitatively very similar. The best approach to measuring steep slopes is therefore to correlate with a small window and to apply an appropriate outlier filter afterwards.

The slope for a tilt in y direction does not need to be corrected for the central view perspective as described for the x slope earlier in this chapter. For both slopes their unit (that depends on the disparity) used here cannot be applied directly anyway. Instead the unit needed for a feasibility assessment would be the tilted plane angle relative to a flat plane. Consequently, the slope needs to be converted to an angle. But as the height difference by one disparity pixel is usually not the same as the lateral distance of one pixel the conversion can only be computed when this relation is known. We call this camera specific parameter *p* here which shortens the conversion formula to:

$$\alpha = \tan^{-1}(p\, m_c)$$

This parameter and the conversion for three distinct tilt angles are listed in the following table for three different Chromasens 3DPIXA models:

Camera	***p***	**m_C=0.1** **Generally good results (all sizes)**	**m_C=0.3** **Outliers start to appear (all sizes)**	**m_C=0.5** **Measurement error 10 times bigger (7x7)**
CP000520-D02-005-0035	2.1	*α=12°*	*α=32°*	*α=46°*
CP000470-C01-015-0040	4.9	*α=26°*	*α=56°*	*α=68°*
CP000470-C01-030-0105	7.1	*α=35°*	*α=65°*	*α=74°*

These numbers are meant as a guideline to assess the feasibility of your application based on the shape of your sample.

8. Occlusion and parallax displacement

We explained in chapter 3.2 how a side wall of a non-horizontal height step can negatively influence the height estimation due to the side wall that is only visible in one of the images from the stereo pair. While the side wall is one of the complications at height steps there is also an occlusion effect which hides image features from the bottom plane from one of the cameras as can be seen in figure 24 on the right side. The left subplot shows that the occlusion is also occurring at bridge structures which do not have any side walls.

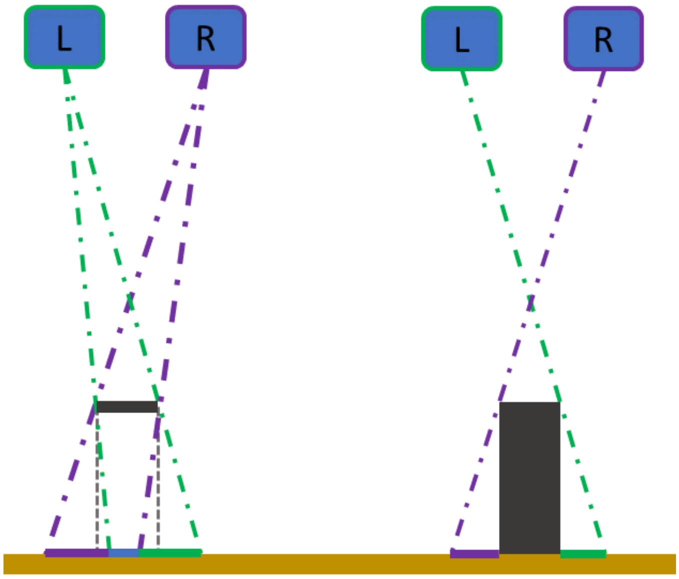

Figure 24: Schematic sketch of the visibility of different image regions around a height feature without sidewalls (left) and with sidewalls (right). The color on the bottom denotes different regions of visibility with the green area not visible in the left image, the violet area not visible in the right image and the blue region not visible in any image.

In the sketch the bottom areas marked in green are only visible to the left camera and as such these areas will not be represented in the height map. The purple areas are visible to the right (template) camera and are therefore included in the height map. But as the correspondent vicinity in the left (comparative) image is occluded the similarity will be low everywhere and the estimated disparity value is random. If there is no sidewall than there is also the blue region which is not visible in any of the two images. If the bridge is narrow enough this region can vanish and the areas directly below it will be visible in both images allowing for both the height of the bridge and the underlying base to be computed. There will still be the area that is only visible to the right image which has no height data, but this area is no longer directly adjacent to the height feature.

If the bridge is smaller than the actual window size used for matching, there is also a parallax effect. The upper level will be surrounded by a different background in the left and right image which can significantly reduce how reliably the matching block for the upper feature can be found. This effect becomes apparent when looking at so called “wirebonds” which are very thin wires connecting electronic components. Wires connecting two components might stretch over an area where the base is much lower and has a changing background as shown in figure 25.

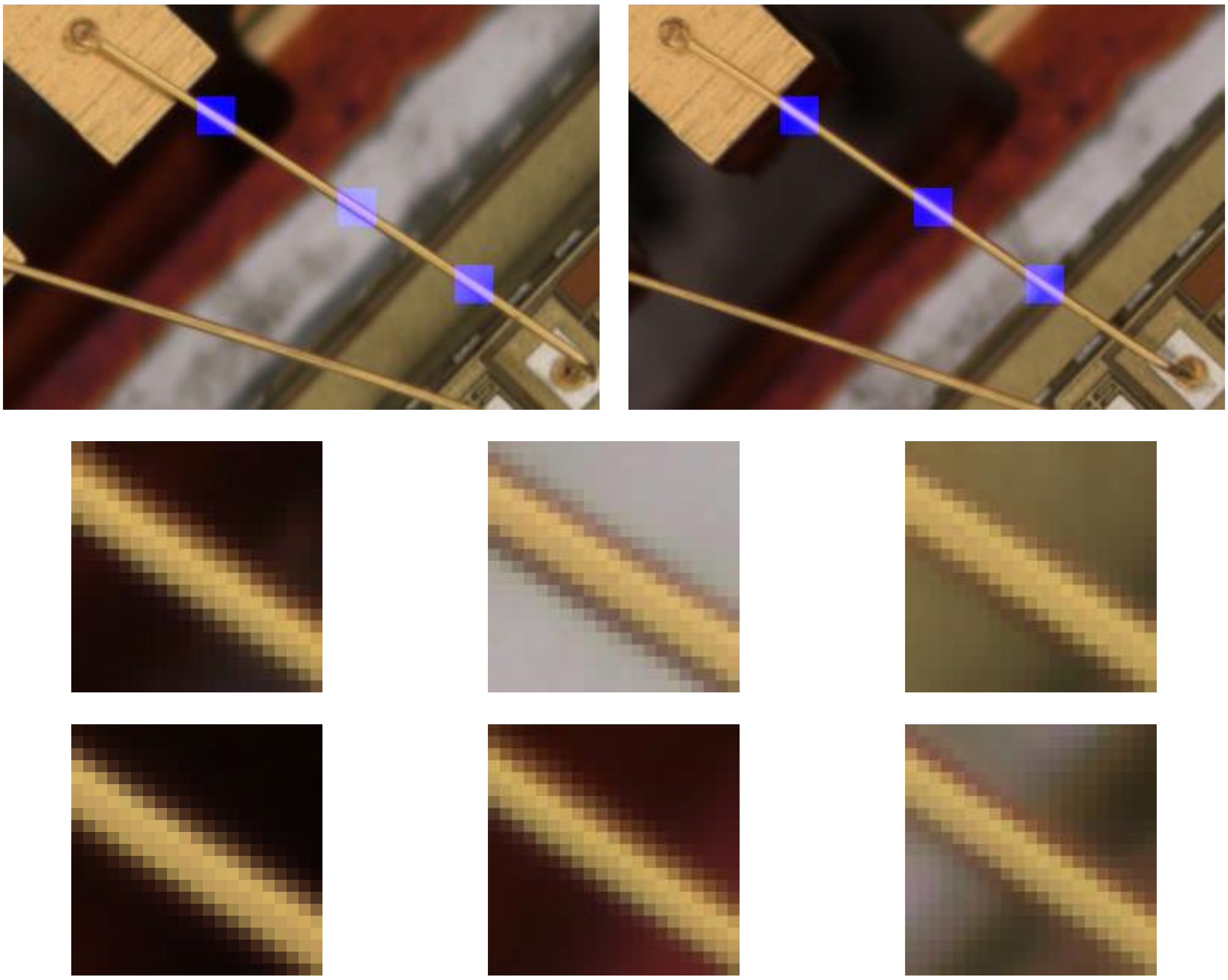

Figure 25: Wirebonding sample with changing backgrounds. The top row shows the overview image with matching windows for the wire at three different epipolar lines marked in blue. Below are details of these windows with the left images in the middle row and in the bottom row the windows from the right image.

The background change behind the second and third windows will significantly increase the matching cost at the disparity value corresponding to the wire height. Depending on the matching algorithm and the surrounding features of the sample this can lead to a match at a completely wrong height or to missing heights on the wire if there is some threshold on the maximum allowable cost value. A smaller window size will decrease the influence of the background on the matching cost and indeed the window size used here for visualization is much bigger than what is sensible for matching given the size of the wire. With a window size smaller than the wire diameter the dependency on the background could in principle be avoided. But as wirebonds have essentially no visible texture the correct height matching depends on matching the contrast at the two edges which necessarily needs to include some background information. The recommended approach to deal with this problem is to apply image processing which homogenizes the background area prior to the stereo matching.

There is one more effect that can occur at height features without sidewalls when the base area is highly reflective. In this case the bottom side of this feature will be visible as a reflection on the base. Due to parallax displacement the reflection is at different positions in the two stereo images. When matched with each other the resulting height will be the base minus the height of the upper feature as is depicted in figure 26.

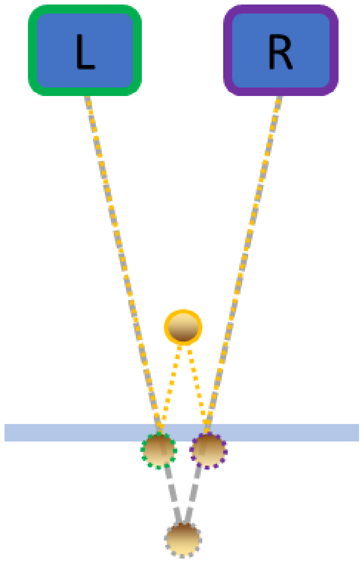

Figure 26: Sketch of a wirebond on top of a reflecting surface. The real wire is depicted with a solid orange border, the reflection seen by the right camera with a dotted violet border and the reflection seen by the left camera with a dotted green border. The assumed height of the reflection after reconstruction is shown with a dotted grey border.

This effect can also occur on wirebond samples where the reflecting surface is the so-called die. One such example is shown in figure 27. The wire is visible in the camera images as a slightly darker area on the die which changes on which direction of the wire it is positioned in the two images. Accordingly, the reconstructed height in this area is much lower than the actual base height of the die.

Figure 27: Wirebond sample with visible reflection of the wire on the die in the image from the left (left) and right (middle) camera and Chromasens 3DAPI height map (right)

This image was captured using a diffuse illumination (Chromasens Corona II Tubelight). As there is light incident from all directions onto the die surface it also emits light in all directions which illuminates the wire even from underneath. Because of that the reflection image is not dark but shows the golden color of the wire which is just very similar to the die color. With a directional lightning like a coaxial illumination the wire reflection would appear dark like a shadow as the underside of the wire is not illuminated from the needed incident angle to reach the angle. The fact that it is not a shadow, but a reflection is an important distinction because the effect will not vanish with the use of a diffuse illumination like a shadow would.

Unlike the other effects described so far, the reflection issue is due to the image content of left and right camera showing different features at the same location. As such this effect is not exclusive to the block-matching algorithm and will occur with every 3D algorithm. This is also the case for the effects that are due to the illumination differences between the two cameras described in the next section.

9. Illumination Effects

As already explained at the beginning of this paper the general idea of stereo matching is based on finding matching features in the two camera images. Any algorithmic way of finding these features relies on the fact that they look very similar in both images. A difference in brightness between the images will generally increase the cost of matching to different degrees depending on the algorithm. The NCC approach is quite robust in this regard as it includes a normalization factor for the brightness. Nevertheless, we strongly advise to setup the cameras in a way that they both show any features with the same intensity in all regions of the images. This means both cameras need to be white point calibrated to the same reference and have a correction for DSNU (Chromasens: black reference), PRNU and vignetting (Chromasens: white reference) active. These steps will ensure that diffuse surfaces will look as similar as possible to each other in the two camera images.

Reflective surfaces will still seem different between the two cameras if the illumination is not a line light of appropriate length, especially if the surface is curved. Figure 28 shows a stereo image pair of a tuss screw head illuminated with a point-like light source. At the top rim there is a gloss point which is at different positions in the left and right image. Due to the different perspectives of the two cameras the brightfield condition is met at different surface tilts which means that the other camera will not see a gloss point if the other is in brightfield condition. Therefore point-like light sources should never be used for stereo imaging with a line-scan camera.

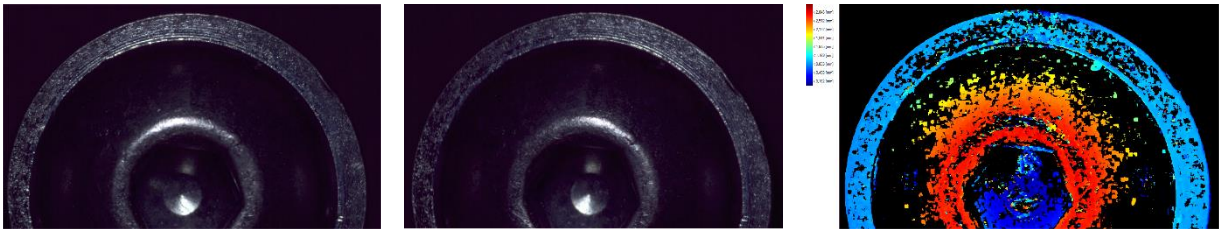

Figure 28: Left and right camera image (left, middle) and calculated height map of a truss screw head illuminated with a point-like light source

Figure 29 shows the same tuss screw head illuminated with a line light source. The gloss point now instead becomes a whole line that stretches along the rim. Due to the extended line light it is possible to meet the brightfield condition for both cameras at the same positions along the rim. Even though they each see different parts of the light the homogeneity of the illumination ensures that the condition is met at the same positions. There is however an edge effect visible at the left and right side of the screw where one camera is in brightfield condition and the other one is not due to the limited length of the illumination.

Figure 29: Left and right camera image (left, middle) and calculated height map of a truss screw head illuminated with a line light source

For most inspection tasks the gloss points of a directed light source are undesired, and a diffuse illumination is needed. Chromasens offers a so-called tube light which is the line-scan equivalent to the dome light used for area cameras. Figure 30 shows the screw head illuminated using a tube light. The tube light illuminates all areas of the screw head homogenously up to a given maximum angle. The angle coverage of this illumination is limited as there is a minimum distance between sample and illumination due to the linear stage needed for line scanning. There is a small left-right offset in the covered angle which is due to edge effects from the viewport slit in the middle of the lightning. This viewport produces the dark line across the middle of the screw which has a slightly different position in the left and right image as well.

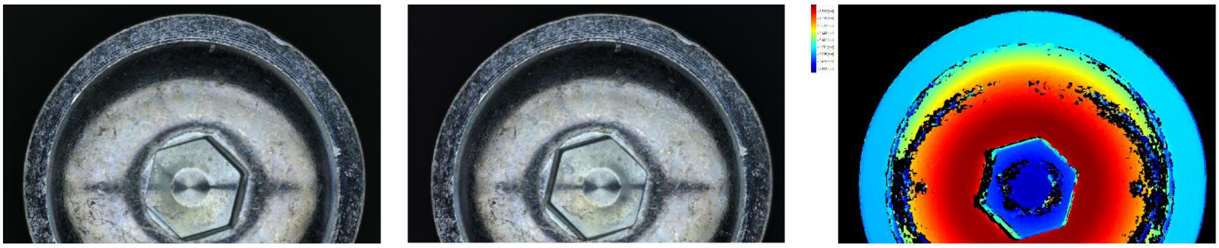

Figure 30: Left and right camera image (left, middle) and calculated height map of a truss screw head illuminated with a tube light source

On a wide assortment of possible applications, the edge effects will not be visible, and the tube light will provide the most similar looking stereo image pairs. But in case these edge effects are occurring at positions that need to be inspected a rotating transport instead of a linear stage might be a possible solution. On a rotating stage adapted to the sample size the surface tilt in scan direction will stay constant which allows for a better adaption of the illumination to the measurement requirements.

10. Summary

Chromasens 3DPIXA cameras can be used to measure 3D features in a fast and precise manner. Depending on the design of the system, many application requirements can be fulfilled. The cameras are delivered with a height reconstruction algorithm based on block-matching that can be accessed by the user via an API interface. This algorithm offers a comparably high computing speed performance and also a high repeatability precision in the reconstructed heights. But at the same time it also suffers from various systematic artefacts.

In this whitepaper we provide a comprehensive overview of all the challenging image contents that typically lead to these artifacts. It is meant as a guide to understanding height artefacts appearing in your application. Once understood it might be possible to lessen the impact of the artefact, either through system design modifications (e. g. scene lighting) or through accounting for the artefact in terms of adequate image post processing. As the size of many artefacts are roughly equal to the size of the matching window, we also encourage you to test your sample with different block sizes and observe the influence this will have on your height map. A few applications like scratch identification or measurement of side-wall slopes are never possible with the default block matching algorithm of the 3DPIXA due to its limitations. However, this does not mean these applications are impossible to be solved. Chromasens can support you in finding tailored solution for these advanced 3DPIXA measurement tasks.

Chromasens GmbH
Max-Stromeyer-Straße 116
78467 Constance
Germany

Phone: +49 (0) 7531 876-0
Email: info@chromasens.de