DENSE IMAGE MATCHING – DIM

SLIDES BY B. ALSADIK

UNIVERSITY
OF TWENTE.

IMAGE MATCHING

- Point clouds can be generated from images using the so-called **Image Matching Algorithms** (IMA)

- **Image Matching** problem consists of finding a unique mapping between the points belonging to two images of the same scene.
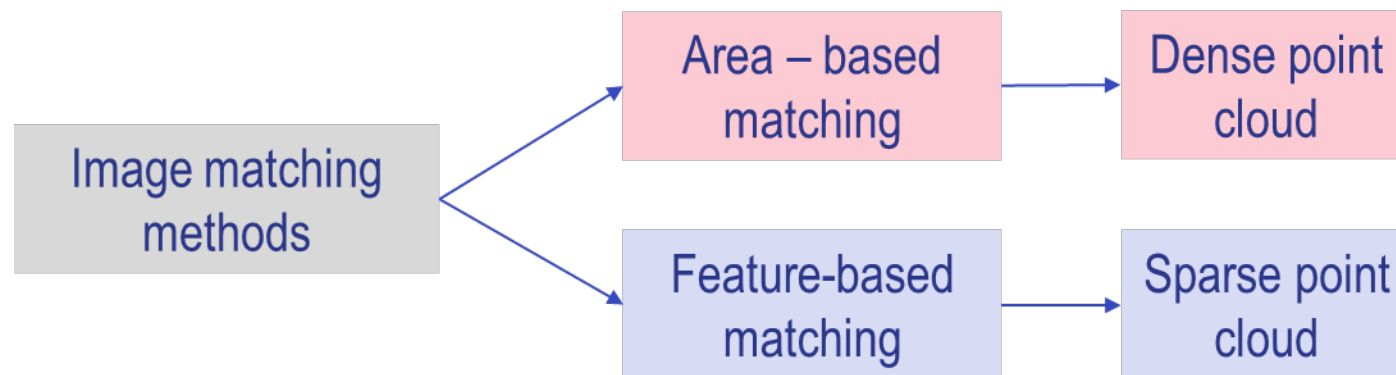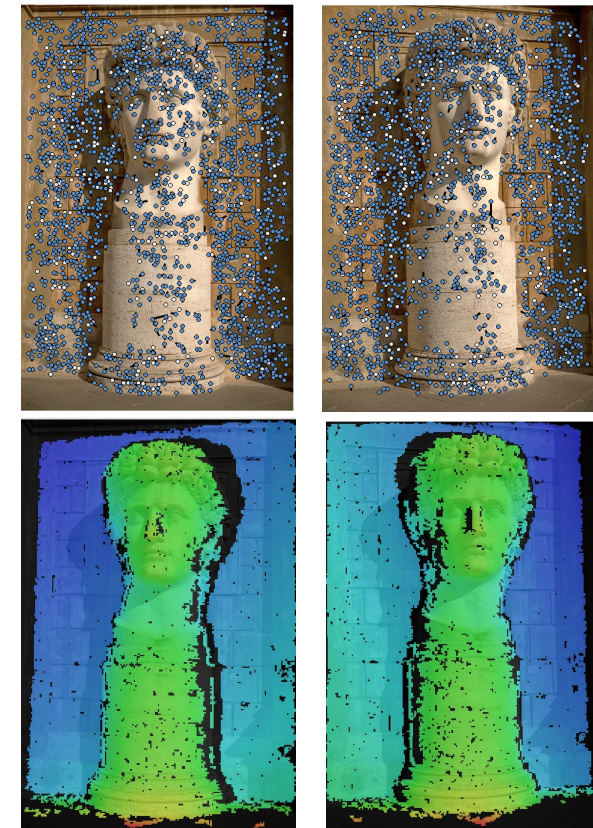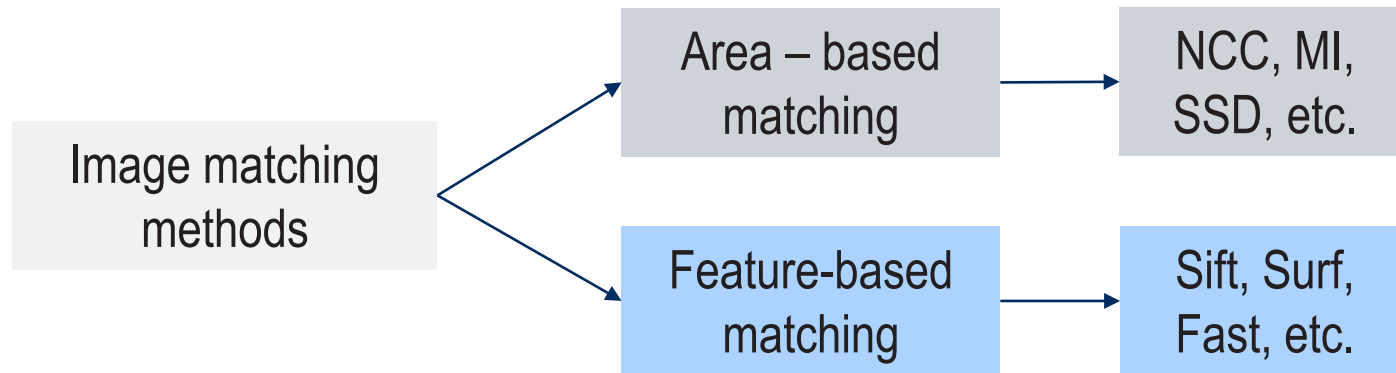
```
Image matching methods ──┬──► Area – based matching ──► Dense point cloud
                         └──► Feature-based matching ──► Sparse point cloud
```

IMAGE MATCHING

Image matching methods

→ Area – based matching → NCC, MI, SSD, etc.

→ Feature-based matching → Sift, Surf, Fast, etc.

Area – based matching
- Computational expensive
- Sensitive to noise, illumination and geometric distortion
- Pixel-wise and accurate

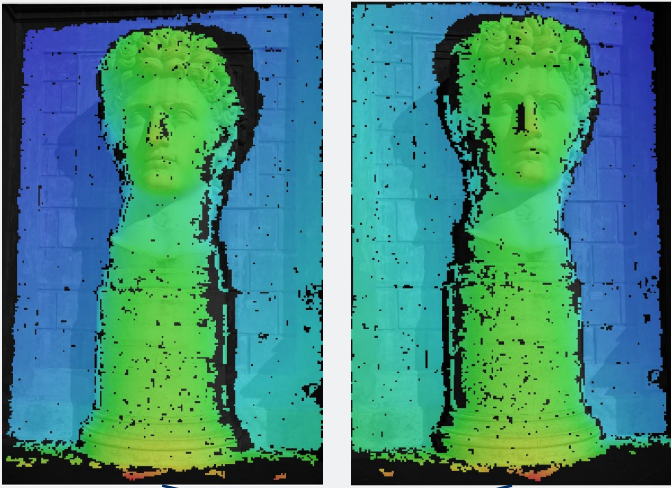Feature – based matching
- Time efficient
- Robust to geometric distortion, illumination and distortion
- Less accurate

Source: B. Alsadik. www.itc.nl

UNIVERSITY
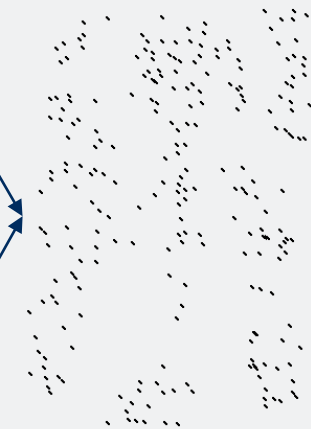OF TWENTE.
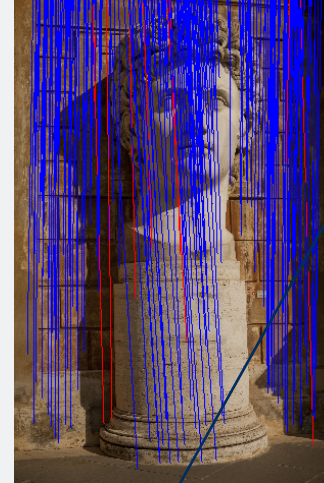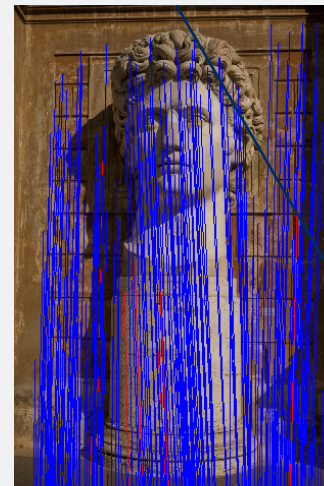
IMAGE MATCHING

Area based matching

Depth map 1   Depth map 2
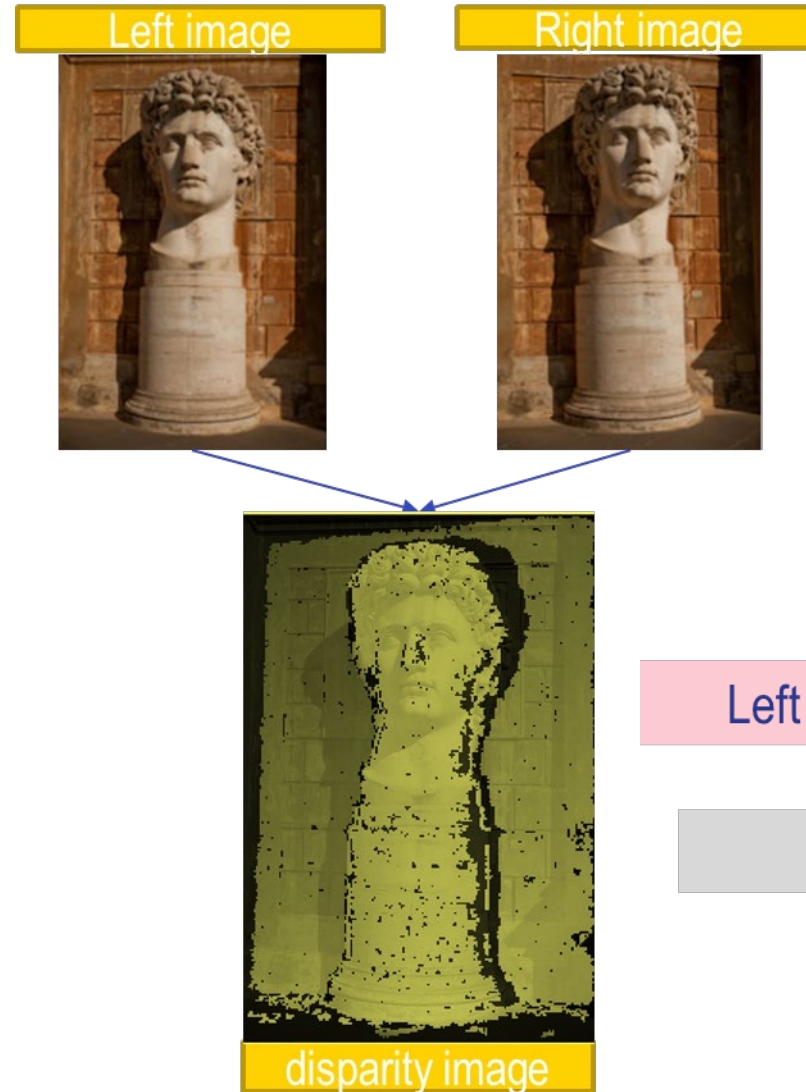

Dense point cloud

Feature based matching

Sparse point cloud

Source: B. Alsadik. www.itc.nl

UNIVERSITY
OF TWENTE.

WHAT IS DISPARITY?

- We need first to understand the term of disparity which should be computed using the DIM. By matching every pixel in the left-hand image with its corresponding pixel in the right-hand image
- Then computing the distance between the pixel values (the disparities) we should end up with a disparity image.

- Disparity images have a peculiar ghost look due to the complete lack of surface color and texture.

Q/ How the derived disparity can lead us to create a point cloud?

Left image   Right image

disparity image

Rectified images

Left image   Right image

Dense Image Matching

Disparity map

UNIVERSITY OF TWENTE.

Source: B. Alsadik. www.itc.nl

WHAT IS DISPARITY?

- in the right-hand image the spot has moved to the left by 45 pixels. If we searched more points, we would see that disparity is greater for points that are closer to the camera (point B≅45) than the distant points (like point A≅ 22).

http://vision.middlebury.edu/stereo/

UNIVERSITY
OF TWENTE.
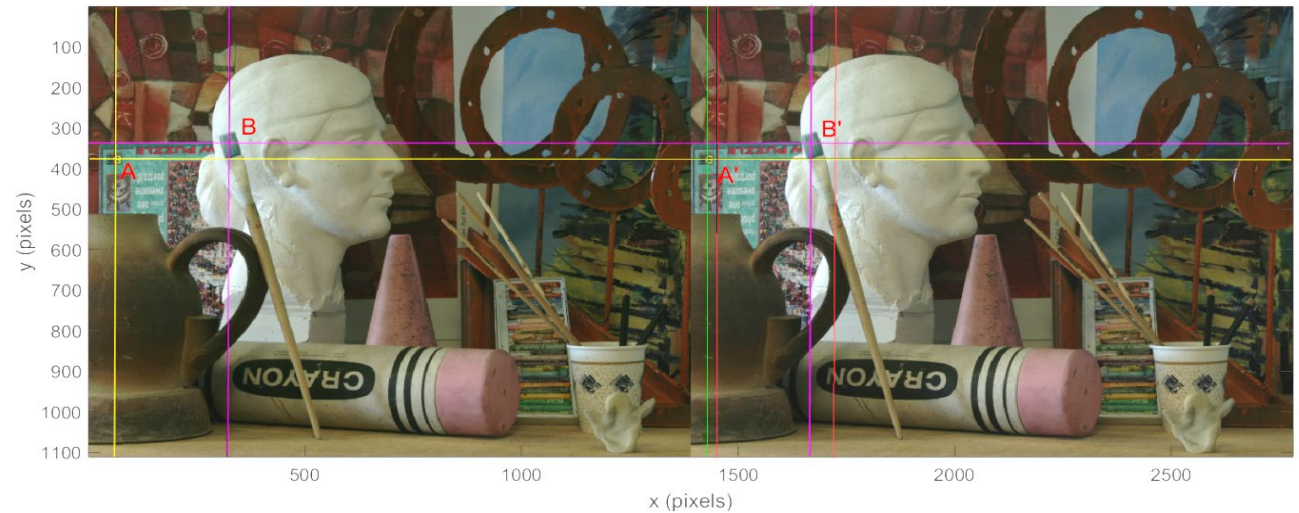
AREA-BASED TECHNIQUE – TEMPLATE MATCHING

- In the area-based matching, instead of estimating the disparity in a pixel-to-pixel relation, we apply window area to area matching.

- We will have a template window and a moving searching window; The goodness of the matching will be computed by comparing the cost in both window

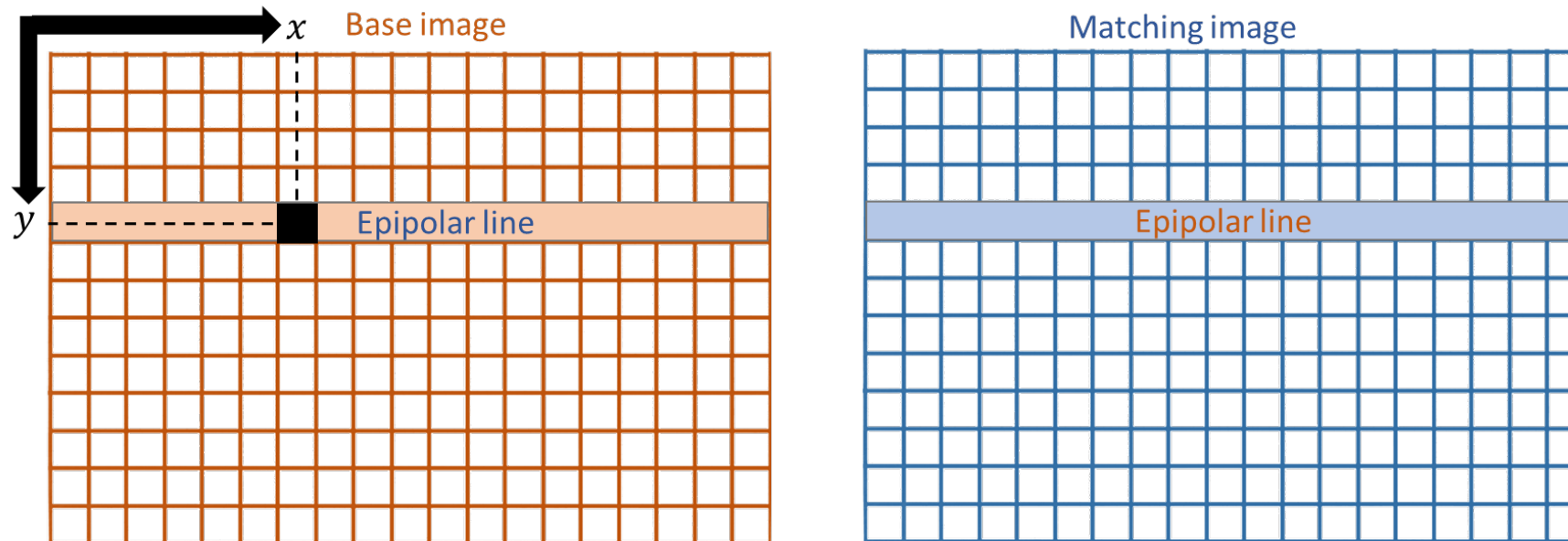- As mentioned, cost can be estimated using NCC, SAD, or MI or else

Template window

Source: B. Alsadik. www.itc.nl

UNIVERSITY
OF TWENTE.

DISPARITY CALCULATION

How to find the optimal disparity of pixel $p(x, y)$ in the left image?

1- image rectification step $\rightarrow$ Y-parallax is removed
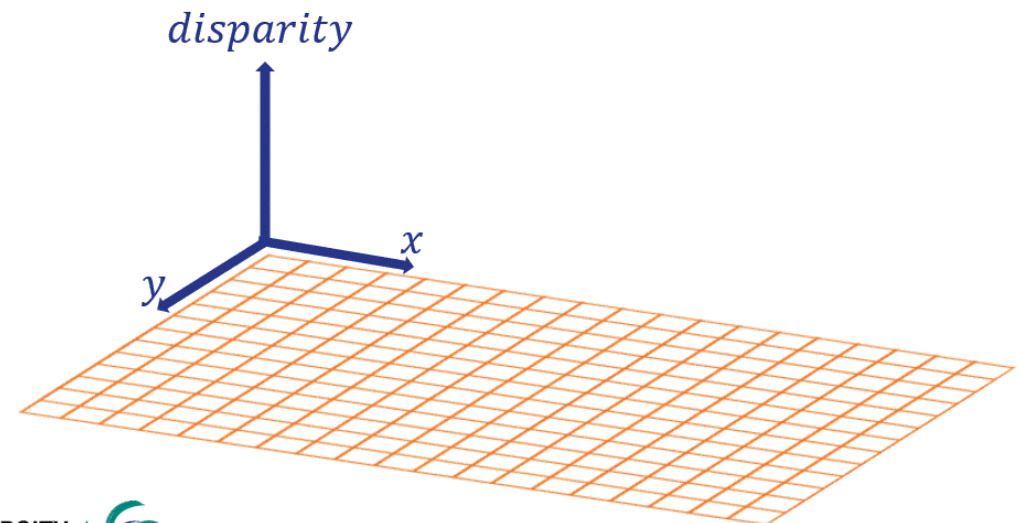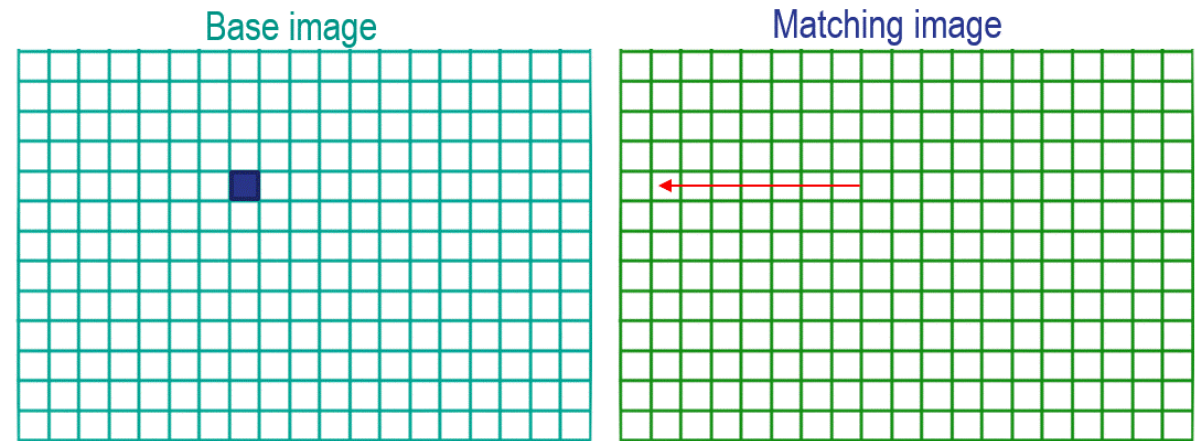Epipolar lines in both stereo images are horizontal and parallel

Source: B. Alsadik. www.itc.nl

UNIVERSITY
OF TWENTE.

DISPARITY CALCULATION

How to find the optimal disparity of pixel $p(x, y)$ in the left image?

2- Search range determination: Determine the range of disparities to consider for the pixel $p(x, y)$.

Base image

Matching image

disparity

x

y

DISPARITY CALCULATION - MATCHING COST CUBE

How to find the optimal disparity of pixel $p(x, y)$ in the left image?

3- Cost computation: For each disparity value within the search range, compute a matching cost between the pixel $p(x, y)$ in the left image and corresponding pixels in the right image along the same epipolar line.

disparity

x

y

matching cost

Min. cost

UNIVERSITY OF TWENTE.
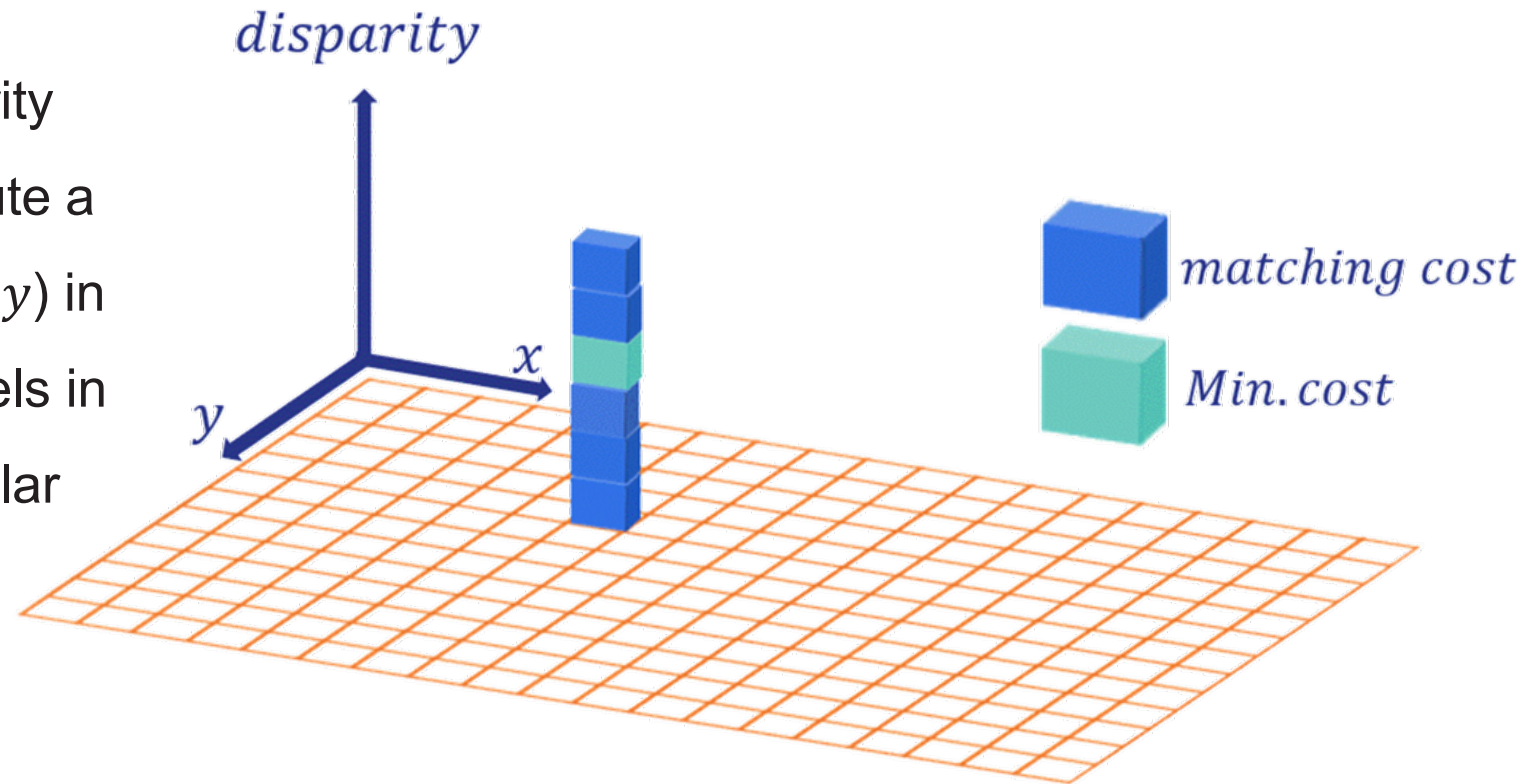
DISPARITY CALCULATION - COST AGGREGATION

How to find the optimal disparity of pixel $p(x, y)$ in the left image?

4- Cost aggregation: Apply a cost aggregation technique to combine the matching costs over a local neighborhood of the pixel $p(x, y)$.

5- Disparity selection: Choose the optimal disparity based on the aggregated costs. In the case of normalized cross-correlation NCC, the disparity with the highest cross-correlation value or the lowest normalized cross-correlation error is typically selected as the optimal disparity.

UNIVERSITY
OF TWENTE.

AREA-BASED TECHNIQUE

- Normalized Cross Correlation → determine the position in a search image with the highest similarity with a reference pattern (template).

$$\gamma(u,v) = \frac{\sum_{x,y}\left[f(x,y)-\bar{f}_{u,v}\right]\left[t(x-u,y-v)-\bar{t}\right]}{\left\{\sum_{x,y}\left[f(x,y)-\bar{f}_{u,v}\right]^2 \sum_{x,y}\left[t(x-u,y-v)-\bar{t}\right]^2\right\}^{0.5}}$$

Where

- f is the image
- $\bar{t}$ is the mean of the template
- $\bar{f}_{u,v}$ is the mean of $f(x,y)$ in the region under the template

Template window

Search window

Source: B. Alsadik. www.itc.nl

UNIVERSITY
OF TWENTE.

NCC– PRINCIPLE IN 1D

- Search for the maximum of correlation
    - — Template profile
    - — Search profile

Correlation coefficient = 0.65

- Way to measure similarity:
    - **NCC**: (Normalized) CC
    - **MSE**: Mean Squared Error
    - **MAD**: Mean Absolute Difference
   - **MI**: Mutual Information

Correlation coefficient = 0.95
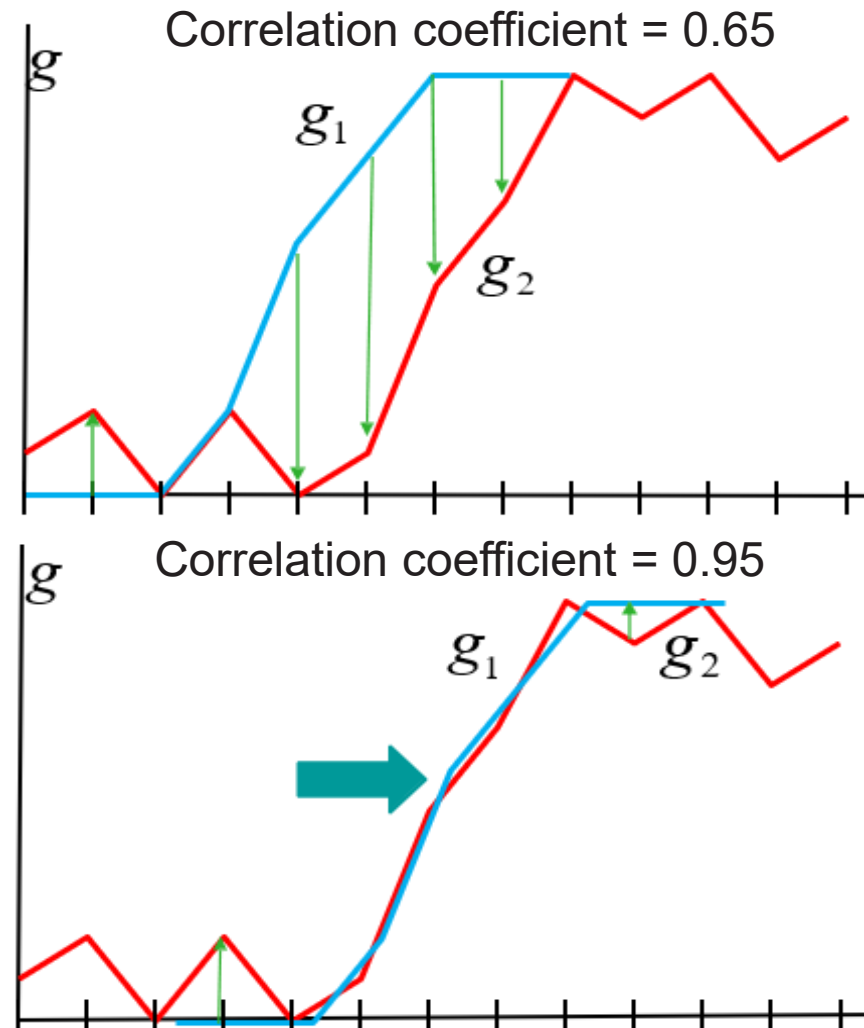
AREA-BASED MATCHING – 2D

- Also called Template Matching

- The *red shaded region* shows the window W that is the set of pixels used to compute the output pixel (shown in *red*)

- Each pixel in the output image is given by

$$\mathbf{O}[u, v] = s(\mathbf{T}, \mathcal{W}), \forall (u, v) \in \mathbf{I}$$

- where $\mathbf{T}$ is the w × w template, the pattern of pixels we are looking for.

- $\mathbf{W}$ is the w × w window centered at (u, v) in the input image.
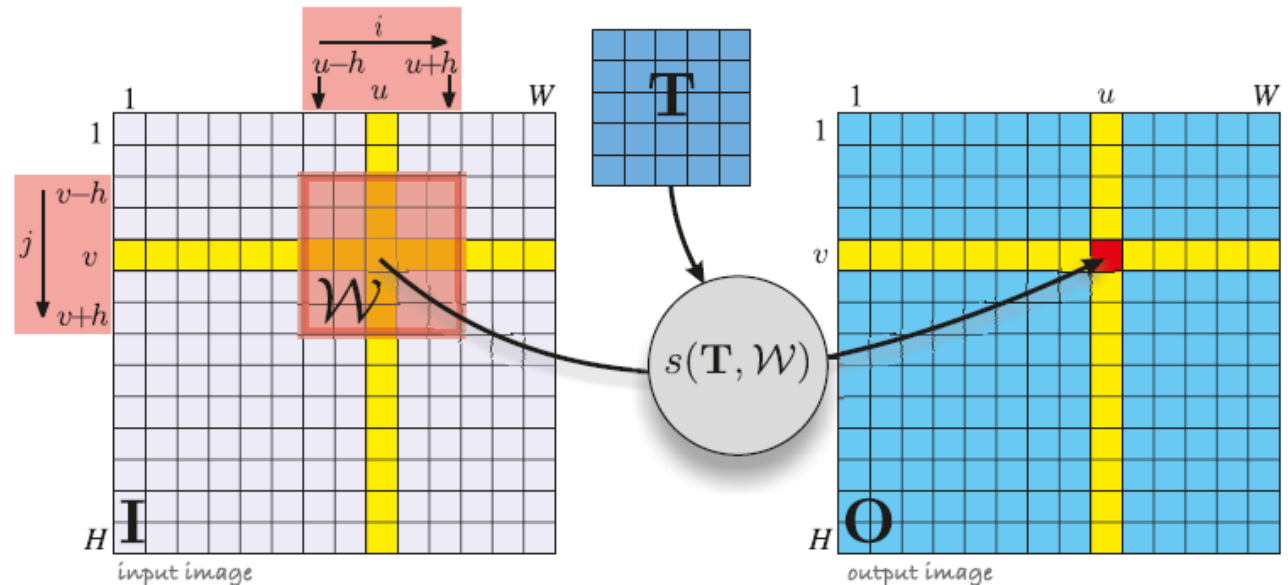
Robotics, Vision and Control, Peter Corke, Springer

AREA-BASED MATCHING

- **Problem:** best match is not necessarily correct match because of noise, repetitive patterns, different illumination, occlusion, in texture less (or -poor) areas the best match is arbitrary

- Different ways to reduce the number of mismatches:

AREA-BASED MATCHING

Add **epipolar constraint** to reduce the search area

Source: cyclomedia technology BV
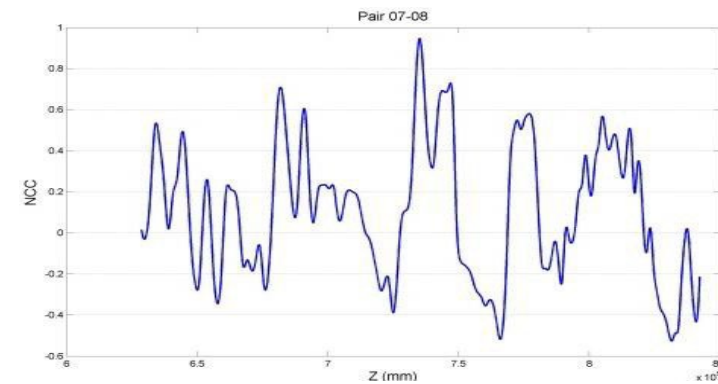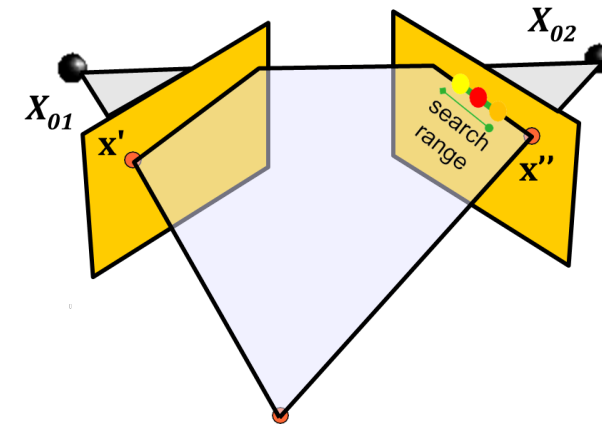
AREA-BASED MATCHING

- Given the relative position between images, they are often resampled in order to have homologous points on the same row.

- Epipolar lines are then horizontal.

- This will make the matching process considerably faster as the search will be in horizontal (1D search).

(a) (b) (c) (d) (e) (f)
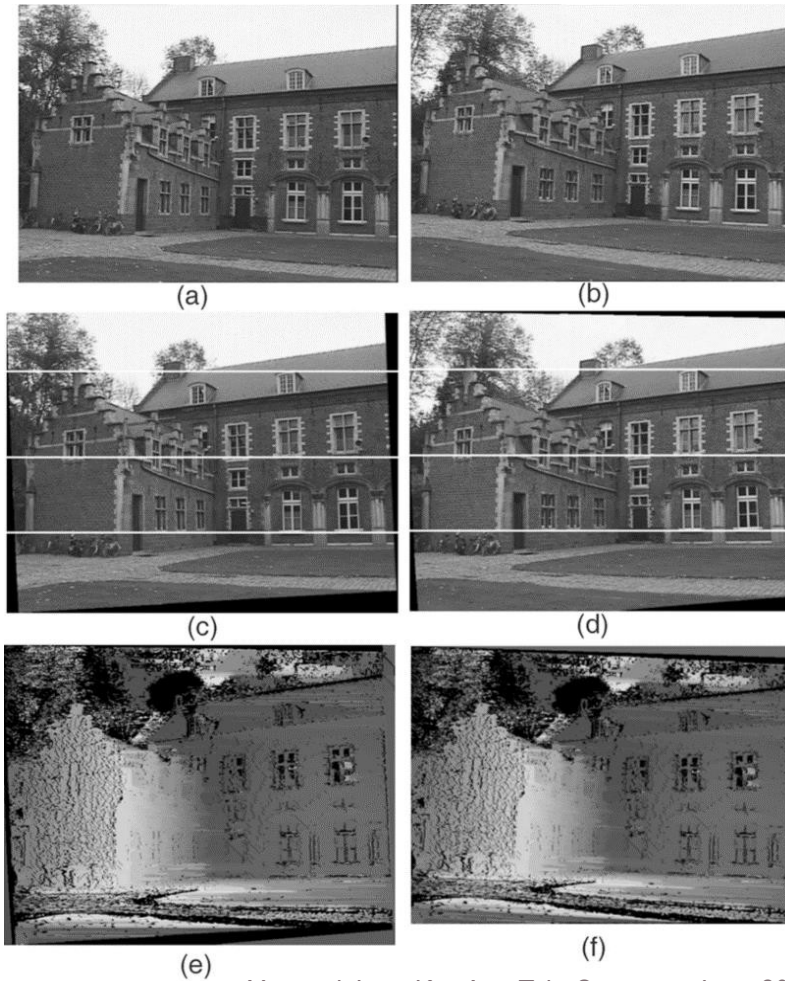
Yunseok Lee, Kar-Ann Toh, Sangyoun Lee, 2008

UNIVERSITY
OF TWENTE.

AREA-BASED MATCHING

- Use of image **scales**

- Refine the 3D information at each step and use it as starting point in the following iteration (higher resolution).

- The advantage of this procedure is to progressively build the point cloud at each step and to use it to reduce the search area in the following steps.

- In other words, we reduce the length of the epipolar line in the search images.

UNIVERSITY
OF TWENTE.
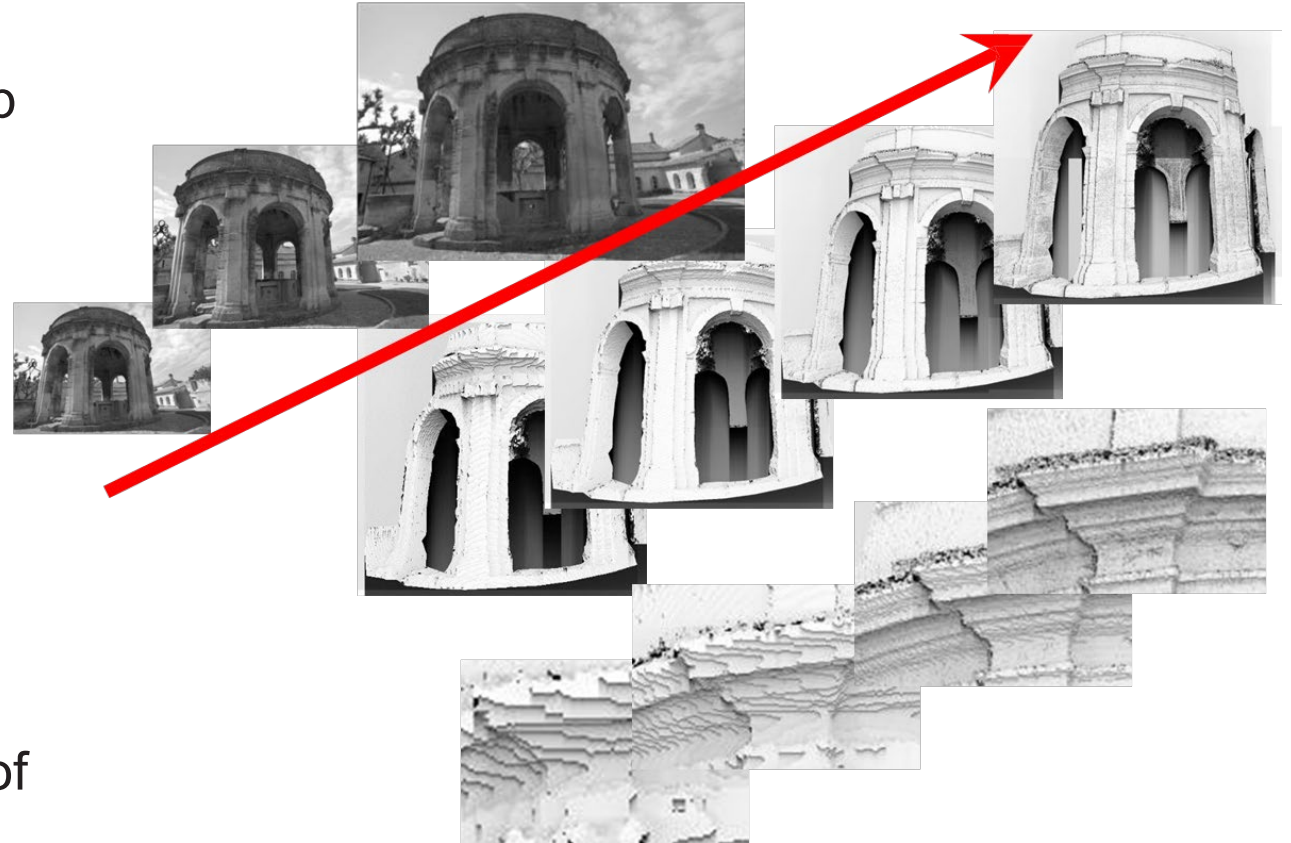
AREA-BASED MATCHING

- Use **more than** 2 images

- More than 2 images are used to match points: many problems on repetitive patterns are solved and a more accurate positioning is achieved.


UNIVERSITY
OF TWENTE.
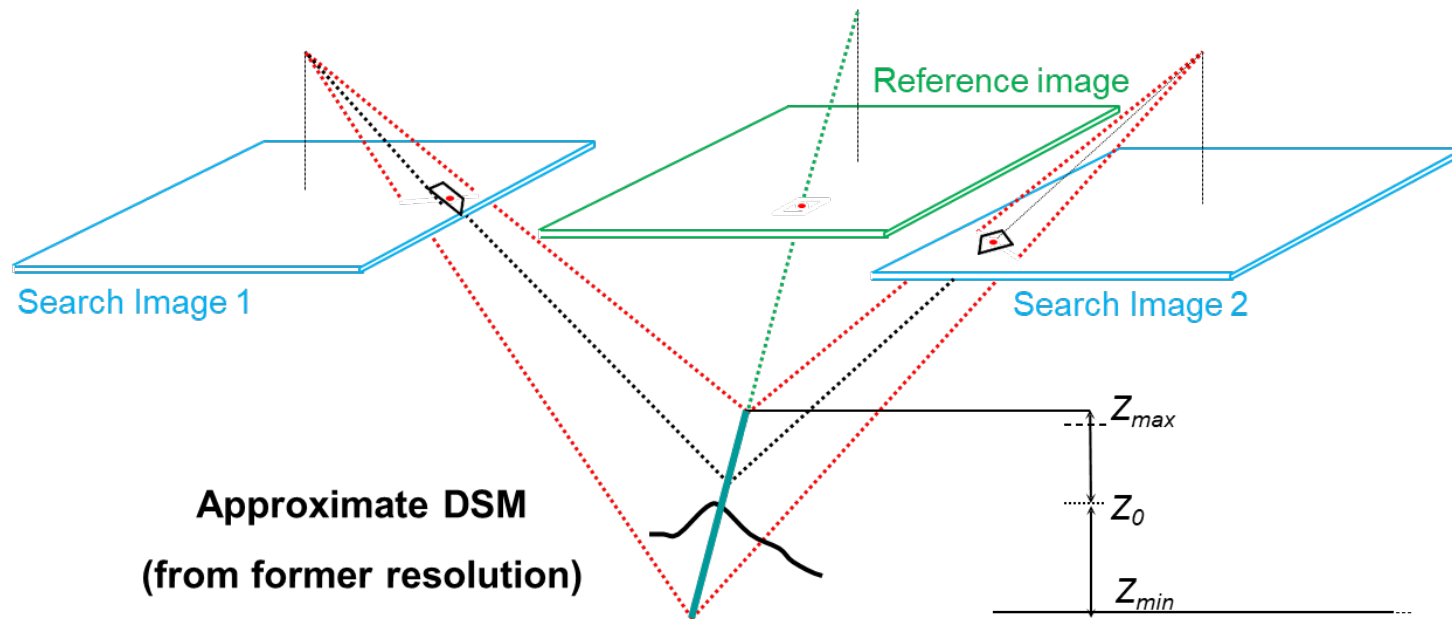
AREA-BASED MATCHING

- Use **more than** 2 images

- The correlations of the stereopairs can be merged in a unique graphs that "summarizes" the information of all the images.

- The correlation peak is more evident on the correct match.
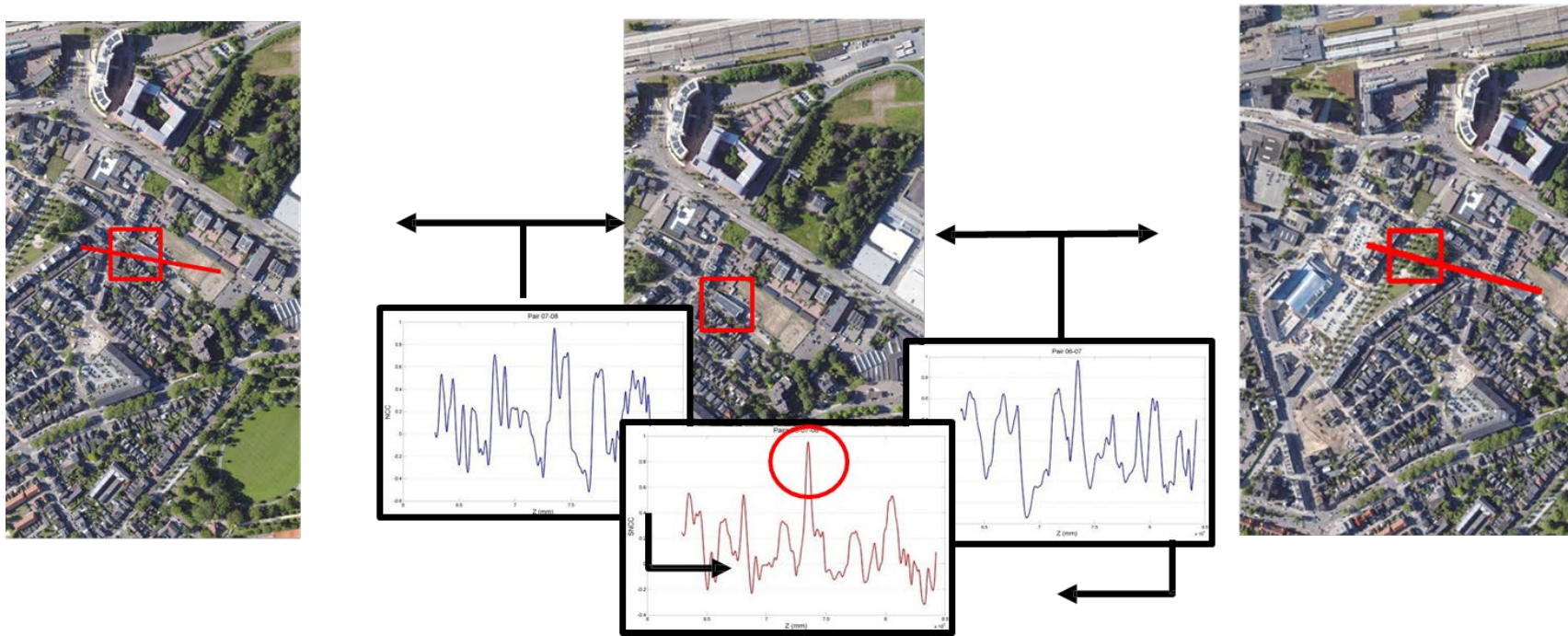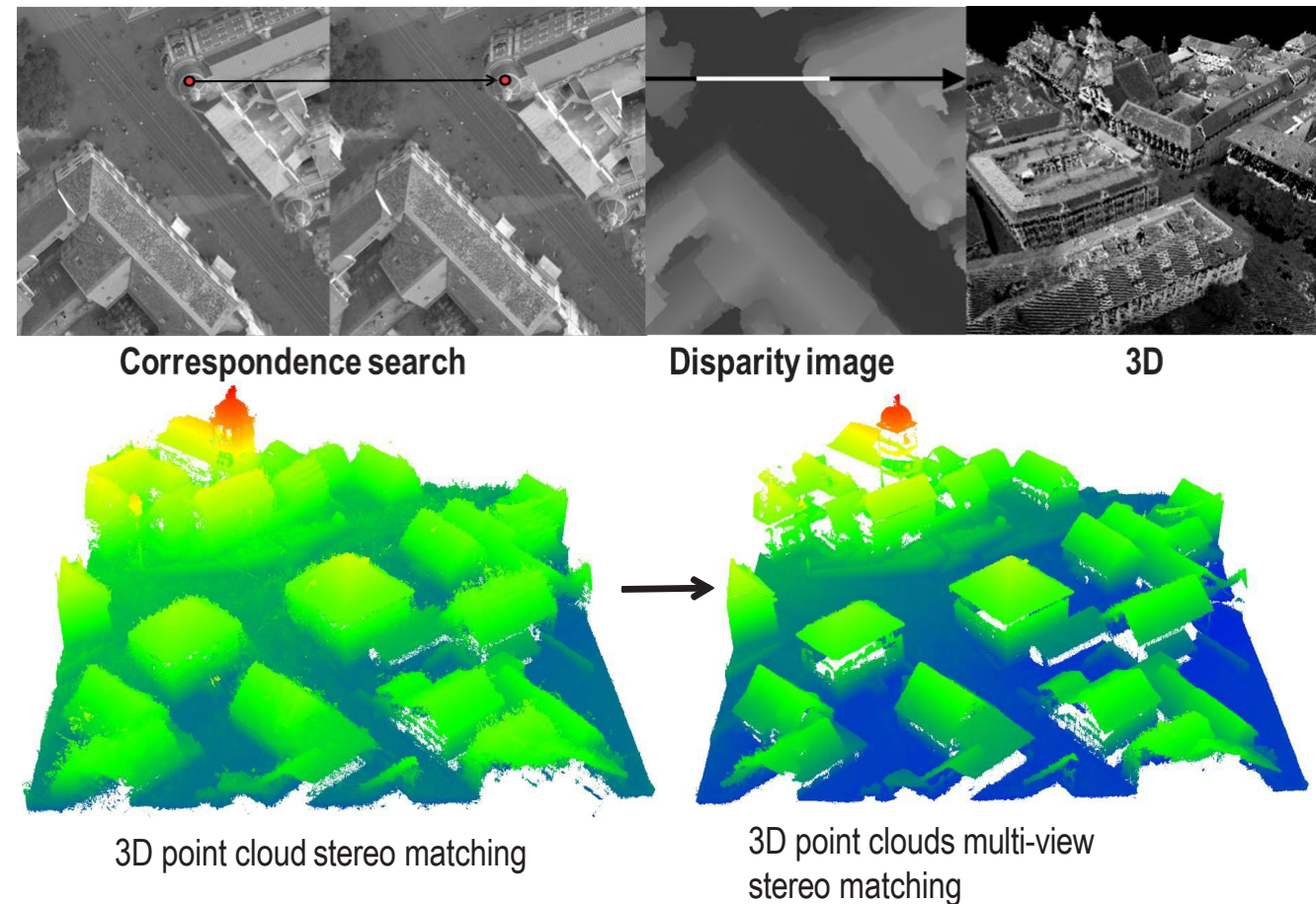
AREA-BASED MATCHING

- Use **more than** 2 images: multi-stereo

- **Multi-image** is often not able to cope with occluded areas: some images do not "see" the point and give a wrong contribution to the match: need for a **visibility map**

- **Multi-stereo** do not have this problem, but the issue is how to merge the different point clouds
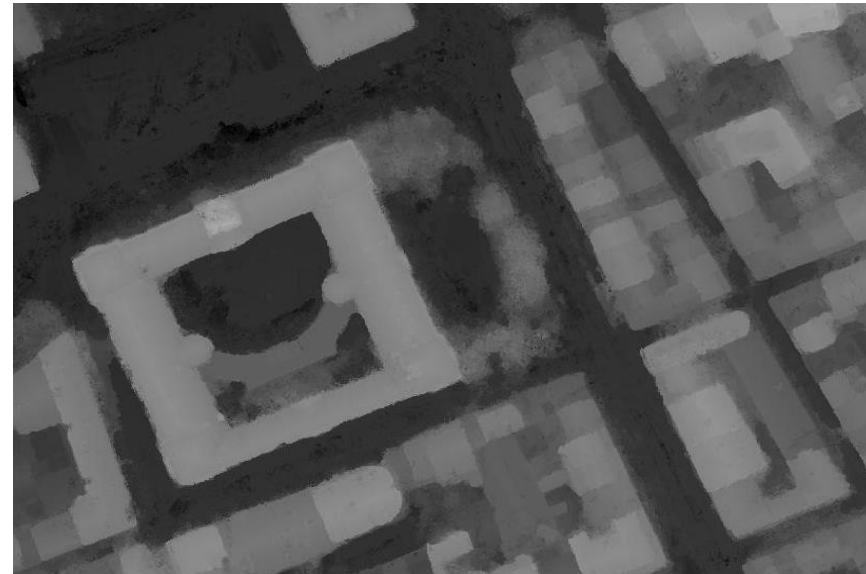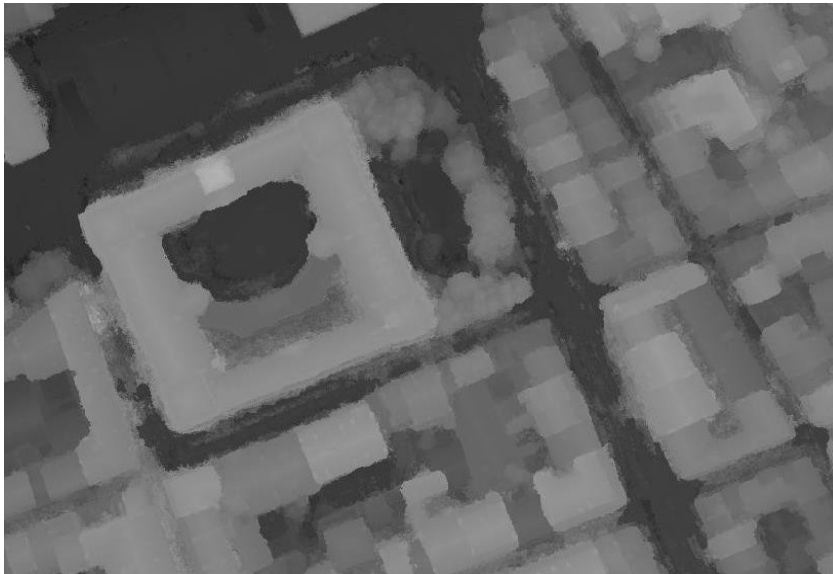
Source : Sure, Konrad Wenzel

UNIVERSITY
OF TWENTE.

AREA-BASED MATCHING

- multi-view stereo offers a more comprehensive and denser point cloud by fusing information from multiple images and viewpoints.

- Multi-view stereo can produce more accurate and detailed reconstructions, but it requires more computational resources


Correspondence search **Disparity image** **3D**

3D point cloud stereo matching

3D point clouds multi-view stereo matching

DVW Stuttgart 2017

UNIVERSITY OF TWENTE.

AREA-BASED MATCHING

Compared to local matching like NCC method, we prefer the global matching methods. Why?

- Robustness to textureless regions:

- Handling occlusions

- Addressing depth discontinuities

- Handling large disparity ranges
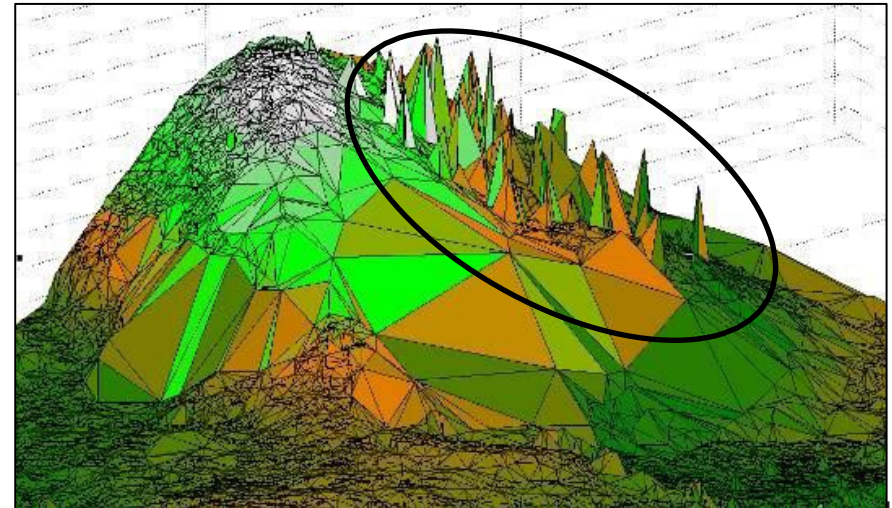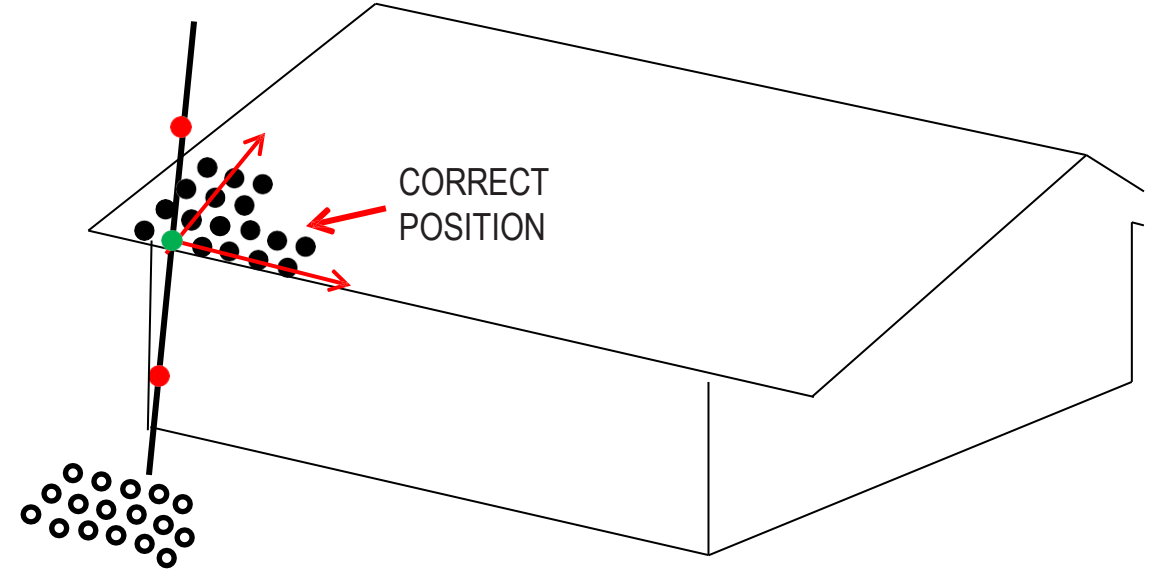
AREA-BASED MATCHING

- **Local algorithms**
  - Window based algorithm: correlation
  - Evaluate one pixel at a time
  - Considering only neighboring pixels
  - Poor at detecting sudden depth variation and occlusions,
  - Blurred edges, loss of small objects

- **Global algorithms**
  - Consider information that is available in the whole image
  - Estimate disparity such that a global energy function is minimized
  - Smaller patches for similarity measure
  - Smooth surfaces are forced by introducing penalty terms for discontinuities

AREA-BASED MATCHING

- **Global:** optimization approach using a global model
→ on the whole image (when possible): 2D optimization

- **Semi-global:** approximation fast numerical solution
→ a discrete solution considering a limited number of directions: 1D optimization

- The results achieved by global or semi-global approaches are very similar.

- Due to their higher speed the semi-global approaches are the ones used in commercial solutions.
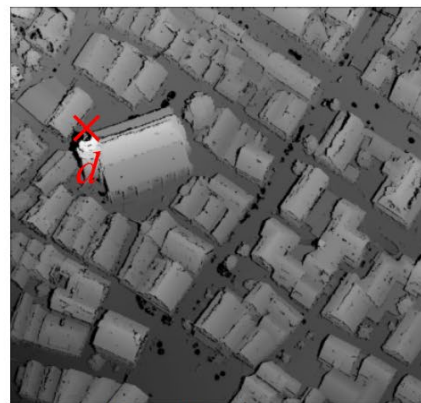
SEMI GLOBAL MATCHING SGM

- The input is a stereo pair: epipolar images are computed in order to reduce the search to a 1D problem.

- For each pixel, the disparity map is computed.

- The disparity image or parallax image stores the correspondence information as disparity, which is correlated to the depth.
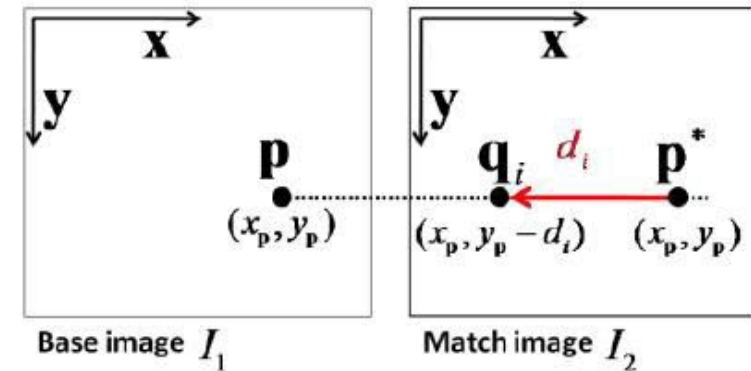

Base image
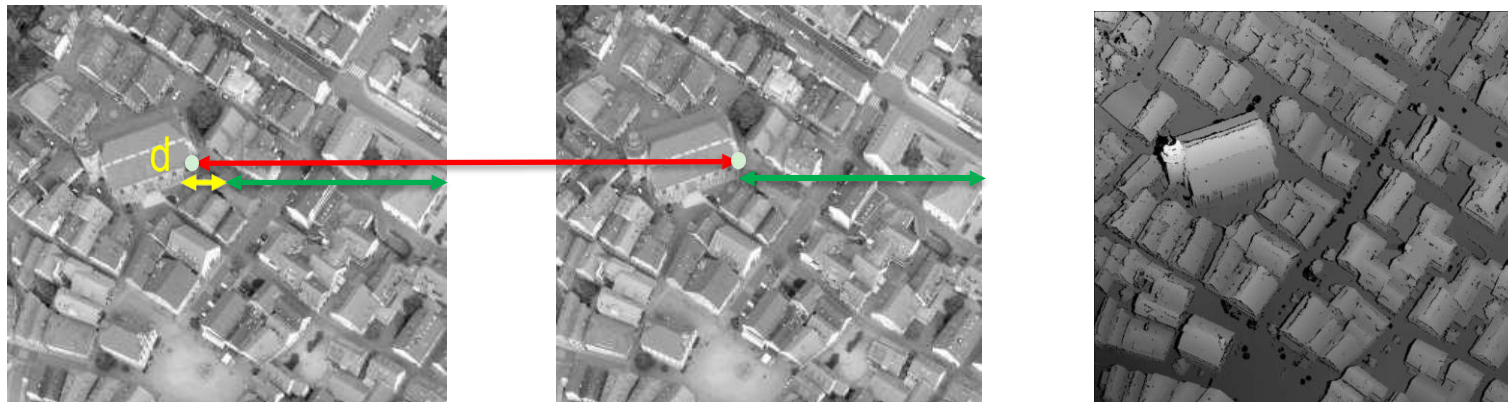
Match image

Parallax image

Base image I_1

Match image I_2

UNIVERSITY
OF TWENTE.

SEMI GLOBAL MATCHING SGM

Continuity, smooth surfaces → Neighboring pixel usually have similar disparities, i.e. similar values

Same pixel order → Pixel along an epipolar line in the left image will have the same order on the corresponding epipolar line on the right image

Limited values for parallaxes → maximum value of parallax allowed between corresponding points in the 2 images

Norbert Haala, University of Stuttgart

UNIVERSITY
OF TWENTE.

SEMI GLOBAL MATCHING SGM

- Definition of a global energy function which needs to be optimized.
Function has the form

$$E(d) = E_{data}(d) + E_{smooth}(d)$$

With:
- d: disparity image

- E_{data}: Term reflecting the total, aggregated costs given disparities d (e.g., cross correlation)

- E_{smooth}: Energy that codes smoothness (e.g., penalties for disparity changes in close neighborhood, preserve discontinuities)

Goal: minimize $E(d)$, i.e. find a disparity image which minimizes the energy function

SEMI GLOBAL MATCHING SGM

- The cost function is computed for all the pixels
- → each pixel can be matched

$$E(d) = E_{data}(d) + E_{smooth}(d)$$

- Ambiguities are avoided by additional continuity constraint

- Add costs (Penalty) for disparity changes of neighboring pixels
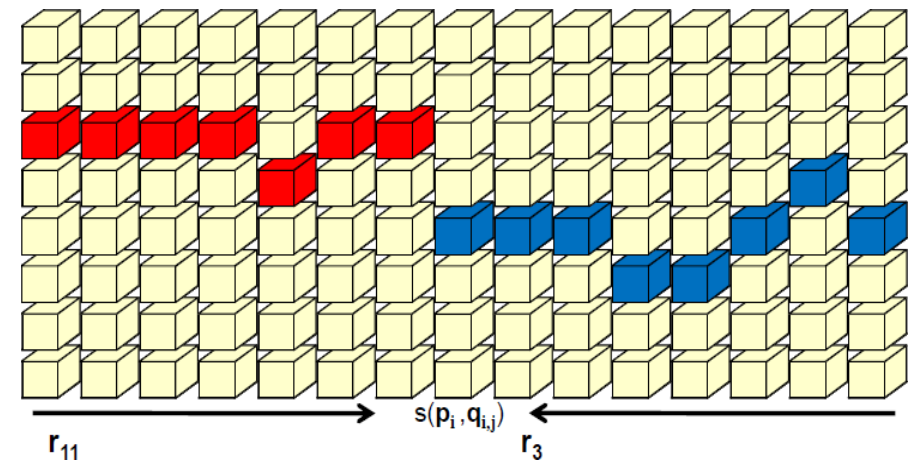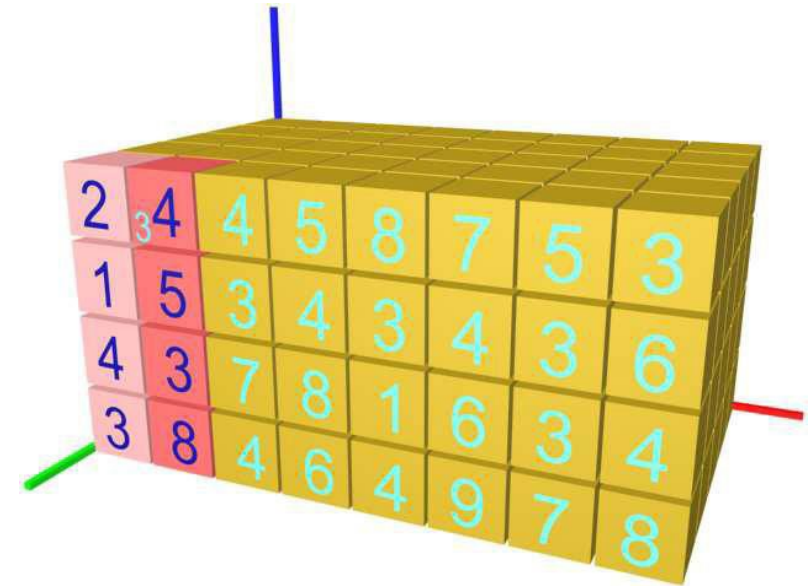
SEMI GLOBAL MATCHING SGM

How to find the optimal disparity of pixel $p(x, y)$ in the left image?

4- Cost aggregation: Apply a cost aggregation technique to combine the matching costs over a local neighborhood of the pixel $p(x, y)$.

UNIVERSITY OF TWENTE.
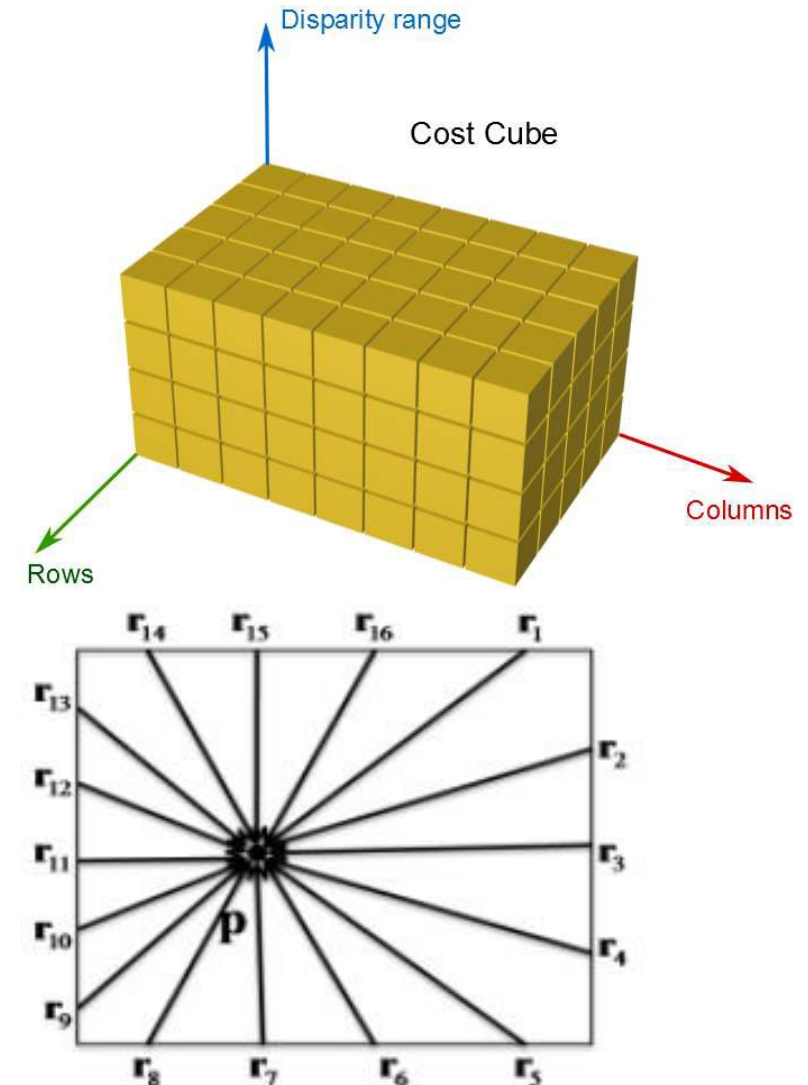
SEMI GLOBAL MATCHING SGM

- Compute the energy function for all the combinations (add smoothing term)
- Choose the best disparity considering the minimum cost and the different directions

$s(p_i, q_{i,j})$

r_{11} r_3

UNIVERSITY
OF TWENTE.

SGM VS TRADITIONAL LOCAL ALGORITHMS
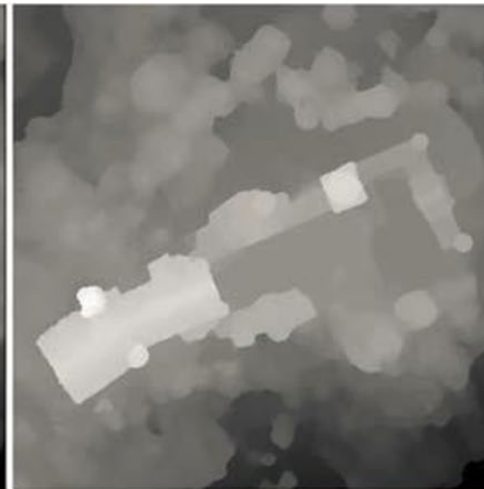
- Higher quality on building borders

- Less mismatches

- Smoother shapes

- Pixel based approaches **1 pixel → 1 point**

image   Cross correlation   Semi-global matching

Source: Hirschmüller, Heiko (2005) - Accurate and efficient stereo
processing by Semi Global Matching and Mutual Information

UNIVERSITY
OF TWENTE.

DEPTH (POINT CLOUD) CALCULATION

- Depth at any given point can be computed if the disparity at that point is known.
  - After disparity being computed for every pixel, depth can be derived for each pixel, As mentioned
  - For the parallel-axis camera geometry, the epipolar lines are parallel and horizontal so conjugate points have the same y-coordinates.
  - When depth (or Z coordinate) is computed, the dense point cloud can be found since the other two coordinates X,Y can be easily computed

- $depth\ (Z) = \dfrac{(baseline * focal\ length)}{disparity}$
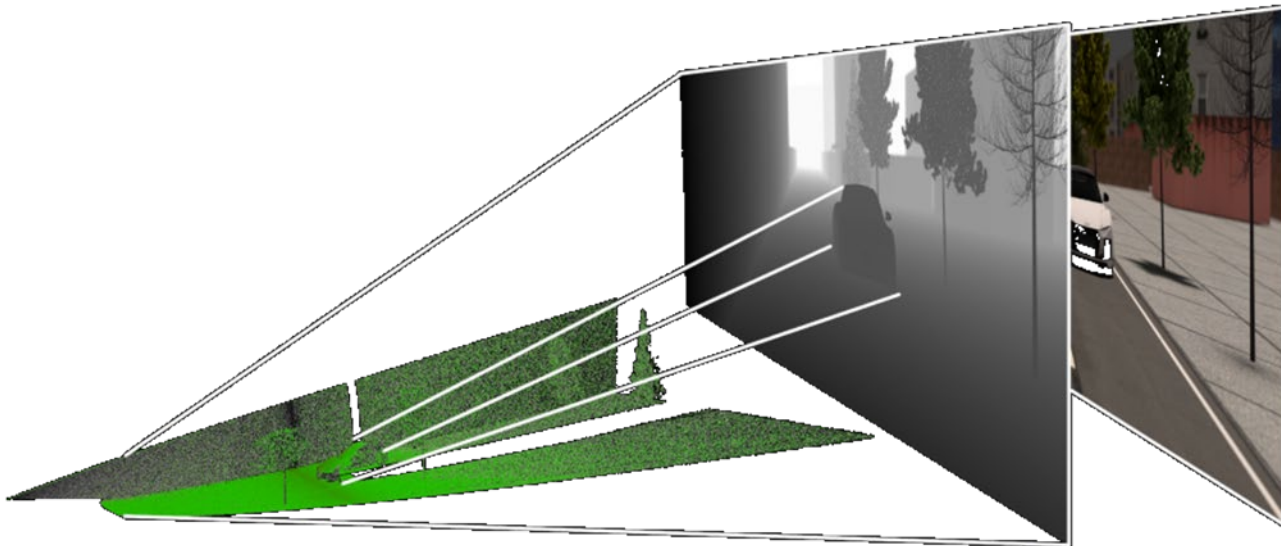
- $X = x_l Z / focal\ length$
- $Y = y_l Z / focal\ length$

Baseline

Disparity

Triangulation

Depth

R,G,B +
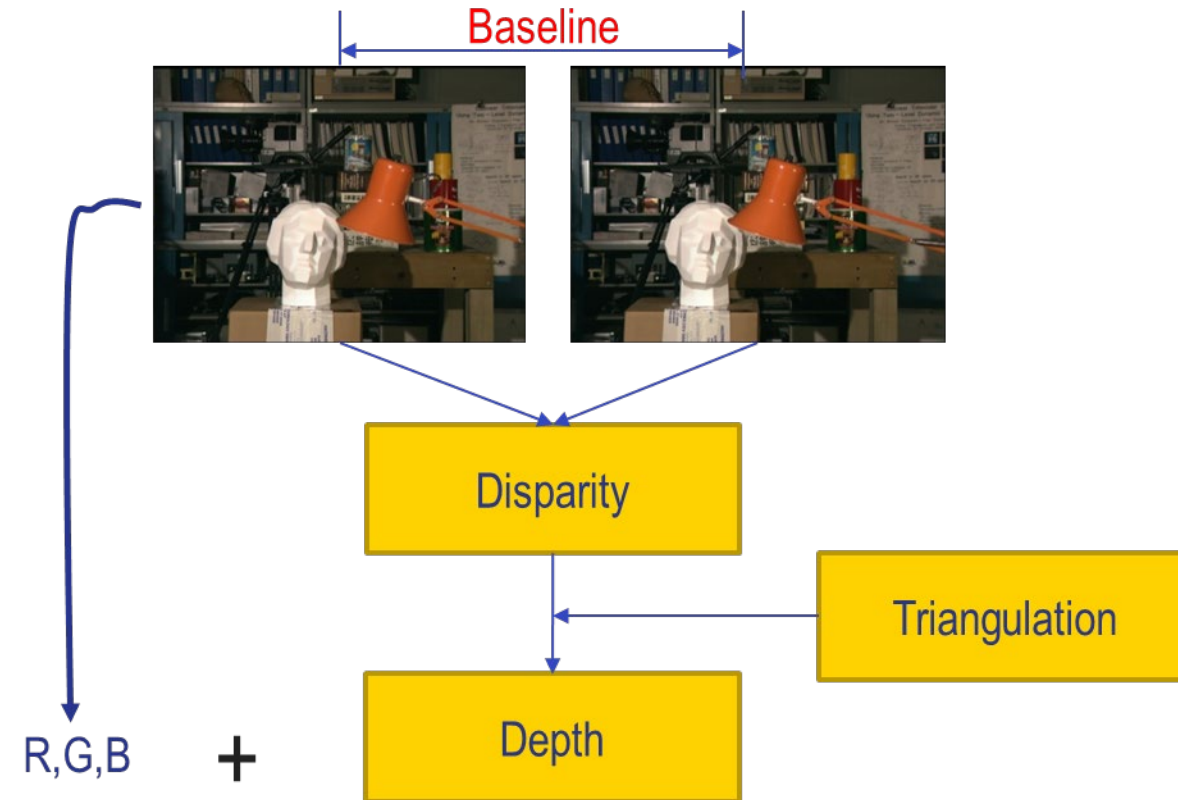
DENSE POINT CLOUD

- Depth images ($R, G, B, depth$) can be converted to X, Y, and Z for every matching pixel.

- Then we can get a dense point cloud.

- If we know the exterior orientation of the stereo images, then we can produce the dense point clouds X, Y, Z in real scale and orientation.
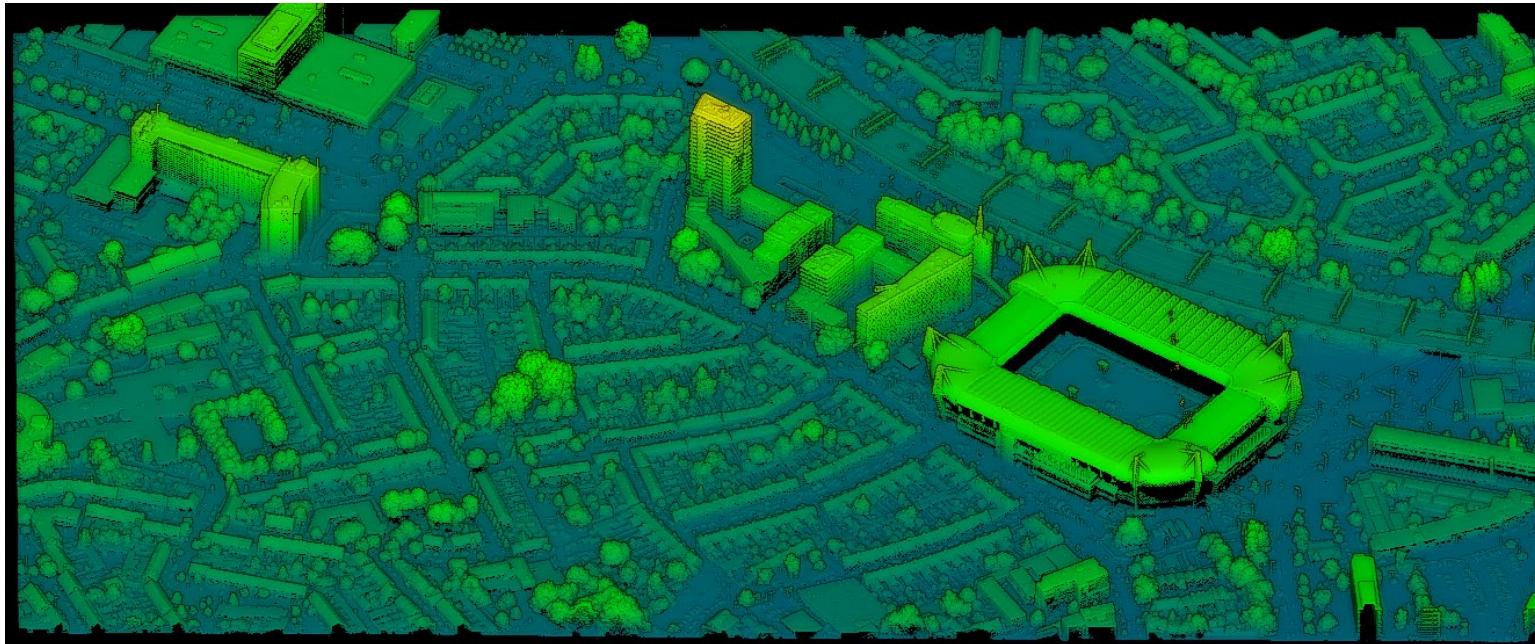

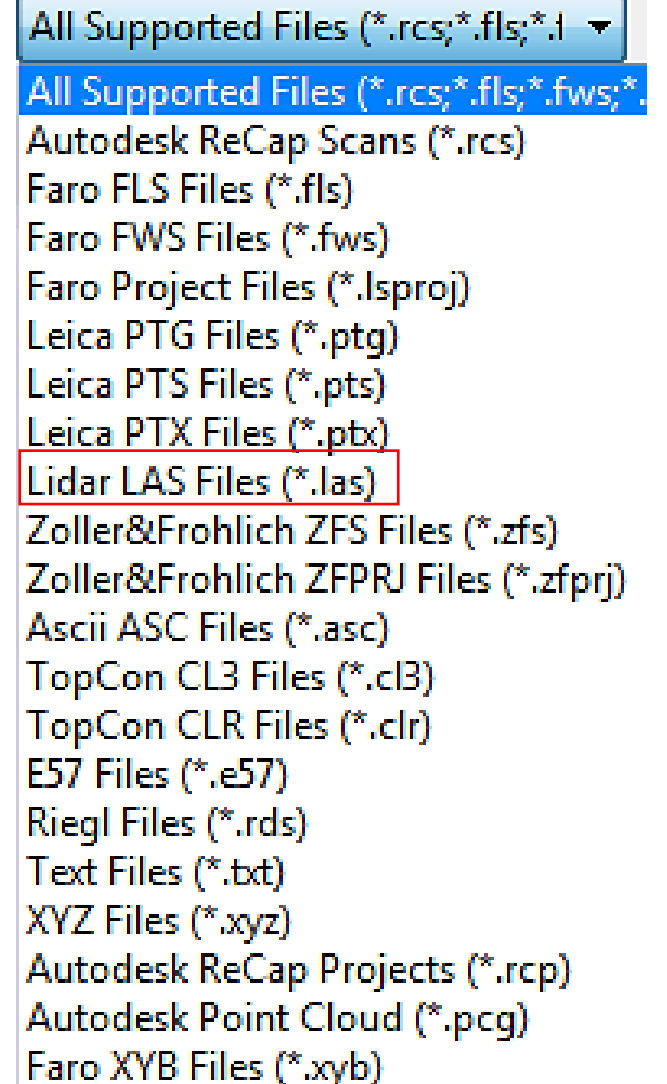Point cloud (green) calculated from depth map (grayscale)

UNIVERSITY
OF TWENTE.

DENSE POINT CLOUD

- A derived point cloud can be defined as a set of points $Pi = (X, Y, Z) \quad T \epsilon R3, i = 1, \ldots n$, fixed in one common coordinate system.

- The point sequence might be ordered but often the points are stored in an unorganized mode.

https://www.cyclomedia.com/en

UNIVERSITY
OF TWENTE.

POINT CLOUD FORMAT

- Hardware producers have created exclusive PC formats.

- The LAS format was developed by the ASPRS to overcome the issues with ASCII for Lidar point clouds.

- Common ASCII formats could be OBJ, XYZ,PTX while binary formats include PCD, LAS, PLY, E57, etc.

All Supported Files (*.rcs;*.fls;*.f ▼

All Supported Files (*.rcs;*.fls;*.fws;*.
Autodesk ReCap Scans (*.rcs)
Faro FLS Files (*.fls)
Faro FWS Files (*.fws)
Faro Project Files (*.lsproj)
Leica PTG Files (*.ptg)
Leica PTS Files (*.pts)
Leica PTX Files (*.ptx)
Lidar LAS Files (*.las)
Zoller&Frohlich ZFS Files (*.zfs)
Zoller&Frohlich ZFPRJ Files (*.zfprj)
Ascii ASC Files (*.asc)
TopCon CL3 Files (*.cl3)
TopCon CLR Files (*.clr)
E57 Files (*.e57)
Riegl Files (*.rds)
Text Files (*.txt)
XYZ Files (*.xyz)
Autodesk ReCap Projects (*.rcp)
Autodesk Point Cloud (*.pcg)
Faro XYB Files (*.xyb)

UNIVERSITY
OF TWENTE.

WHAT HAVE WE LEARNED?

At the end of this lecture, students will be able to:

- Explain the difference between DIM and FIM.

- Describe the disparity and depth.

- Explain the Semi Global Matching technique.

- Describe the template matching principle.

UNIVERSITY
OF TWENTE.

THANK YOU

UNIVERSITY
OF TWENTE.