HOCHSCHULE
RAVENSBURG-WEINGARTEN
UNIVERSITY
OF APPLIED SCIENCES

COMPUTER VISION

COMPULSORY PROJECTS

www.rwu.de
info@rwu.de

GENERAL RULES AND INFORMATION

Compulsory Projects
The reports

In these slides, you will get a collection of all information needed to solve the tasks of the Hands-On projects.

For each task, you will have to **hand in a report via Moodle with a strict deadline**. In this report you will have to answer all questions defined in the task descriptions.

All limitations of your reports (e.g. page-count) will also be published via Moodle.

Compulsory Projects
The reports

A report should always carry all information for the reader to reproduce your algorithms to get the same results as you. At least provide always the following:

- **What?**

  What is your answer or result to the question?

- **How?**

  How did you come to your conclusion? How did you calculate your result?

- **Why?**

  Why did you use your approach? Why did you define the function the way you did it?

Compulsory Projects
The reports

A report should always carry all information for the reader to reproduce your algorithms to get the same results as you. At least provide always the following:

- **What?**
- **How?**
- **Why?**

Describe your own functions in detail to allow a reproduction by the reader.
Here, pseudo code or a flow chart can be helpful to describe longer or more complex functions.

Compulsory Projects
The reports

Follow the general rules for a scientific report. Read the “How to cite” guide provided.

Any kind of plagiarism will result in an automatic failing grade (5.0) for all students involved.

Compulsory Projects
The reports

As a typical What You See Is What You Get (WYSIWYG) editing software (e.g. Microsoft Word, Apple Pages) will reach its limitations for a scientific report, **we will use LaTeX only**.

To generate your report, use one of the many available LaTeX editors (e.g. TeXworks): https://en.wikipedia.org/wiki/Comparison_of_TeX_editors

As you will most likely have to use LaTeX for compiling your Master thesis, this is a good place to start practicing.

Compulsory Projects
The reports

There are many templates available:

https://www.latextemplates.com/

We also have RWU templates:

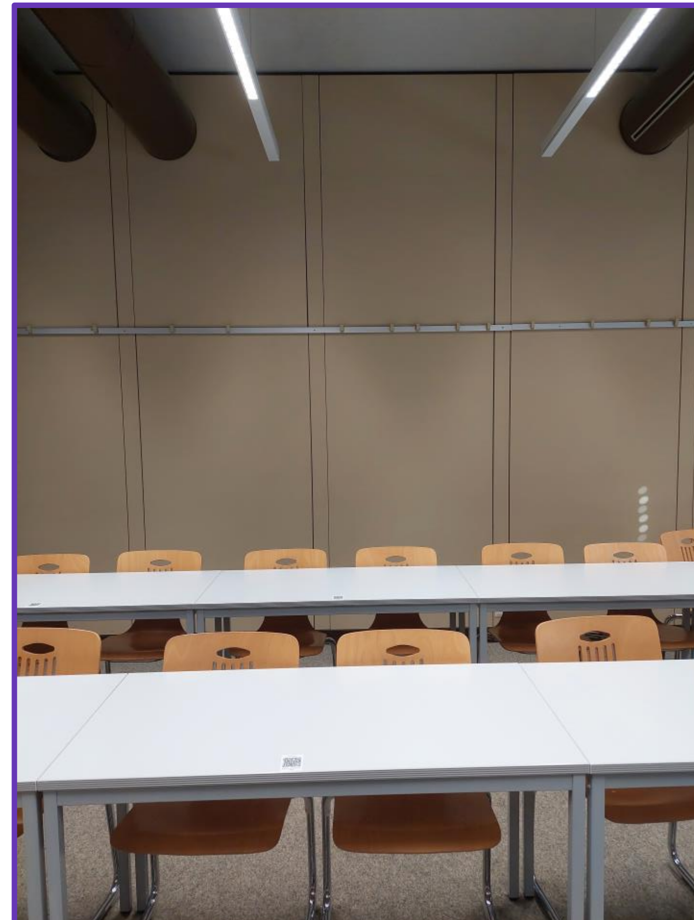
https://fbe-gitlab.hs-weingarten.de/prj-corporate-design/latex-rwustyle

Now, let us talk about what we want to do....

TASK 1:
(SIMPLE) AUGMENTED REALITY

Augmented Reality
Mixing reality with digital information

Take a look at the empty wall of H238.

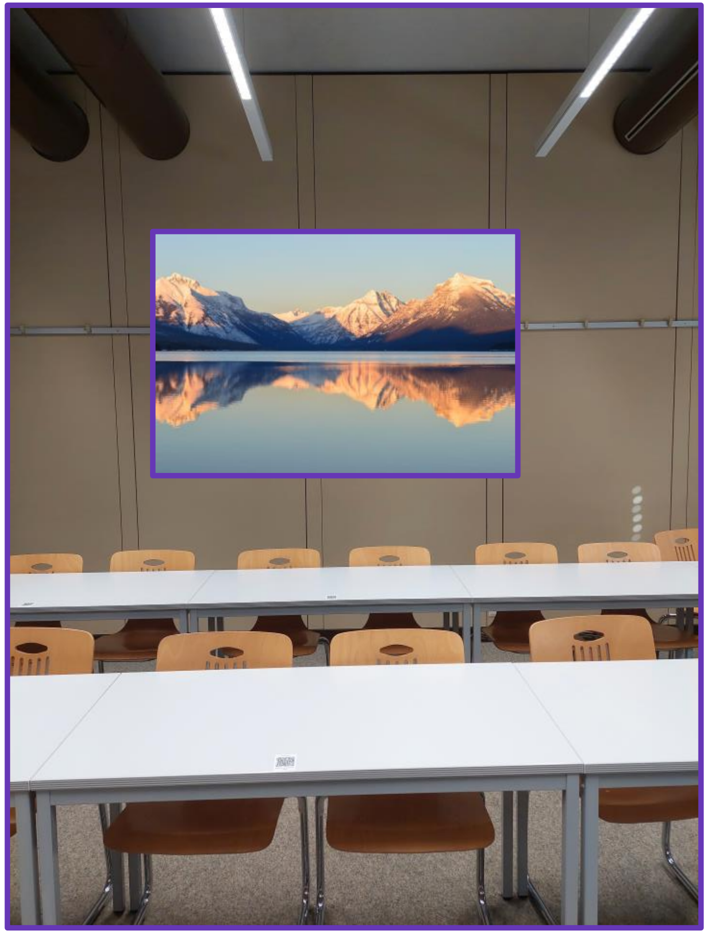


We could make this room nicer with some additional ornamentation....

Augmented Reality
Mixing reality with digital information

A poster would be nice!

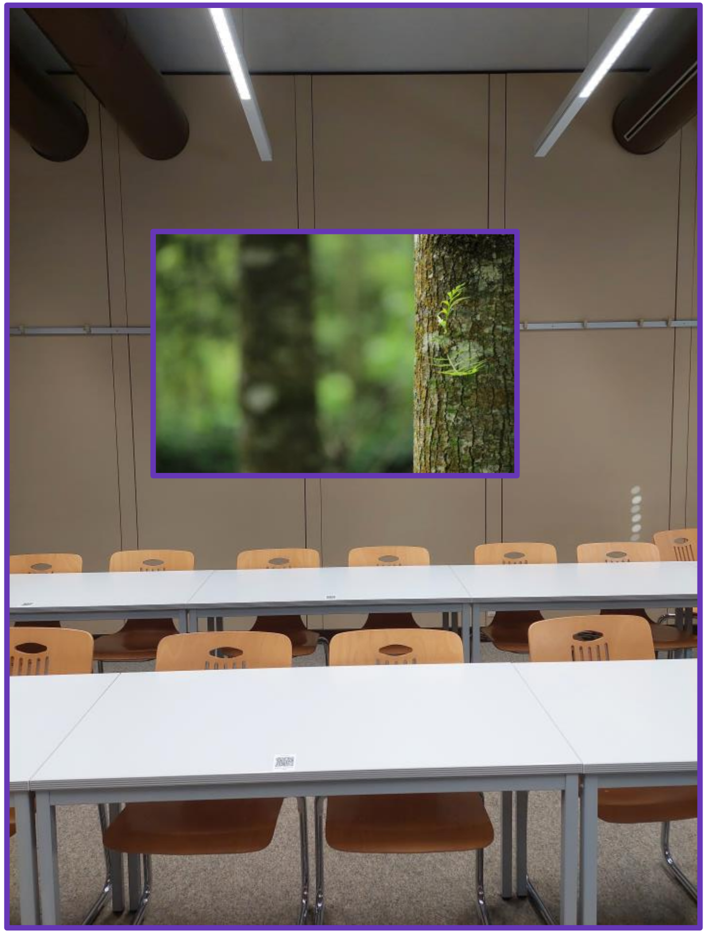
We can past some images onto the image to see how that would look like to find a nice motive….



Augmented Reality
Mixing reality with digital information

A poster would be nice!

We can past some images onto the image to see how that would look like to find a nice motive....

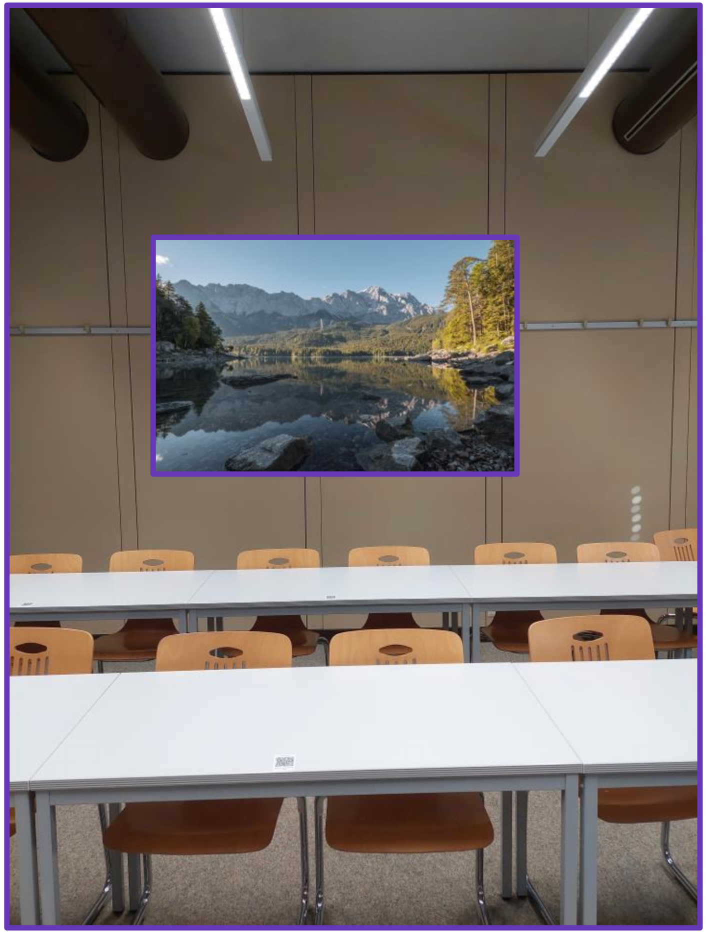


Augmented Reality
Mixing reality with digital information

A poster would be nice!

We can past some images onto the image to see how that would look like to find a nice motive. Okay, this looks nice enough.

But a simple copy paste will not work!

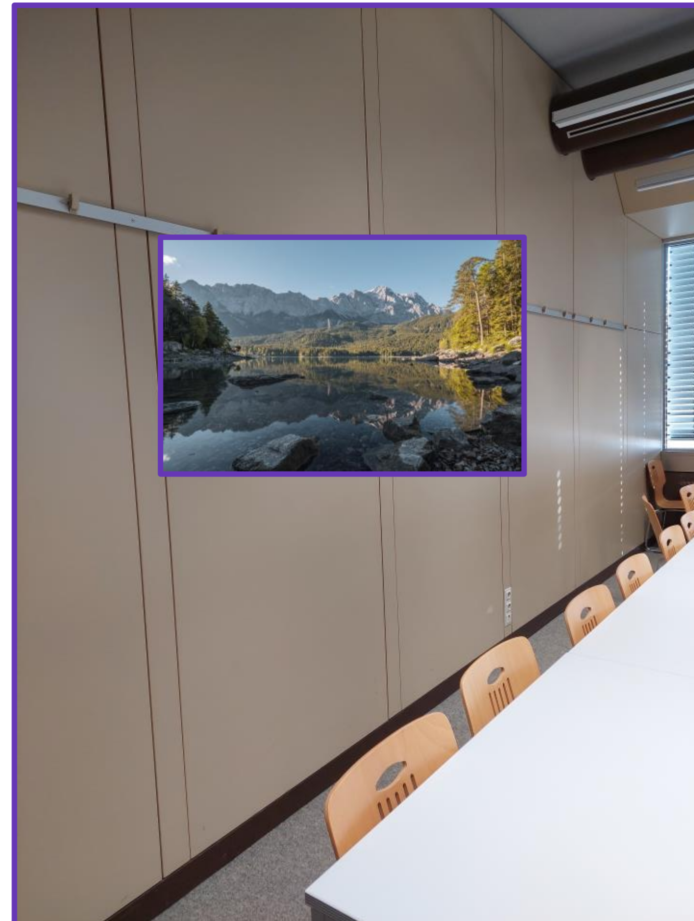


Augmented Reality
Mixing reality with digital information

But a simple copy paste will not work!

We need additional information of the 3D positioning of the wall inside the scene which is projected onto the 2D image.

For this, we can use some defined markers…

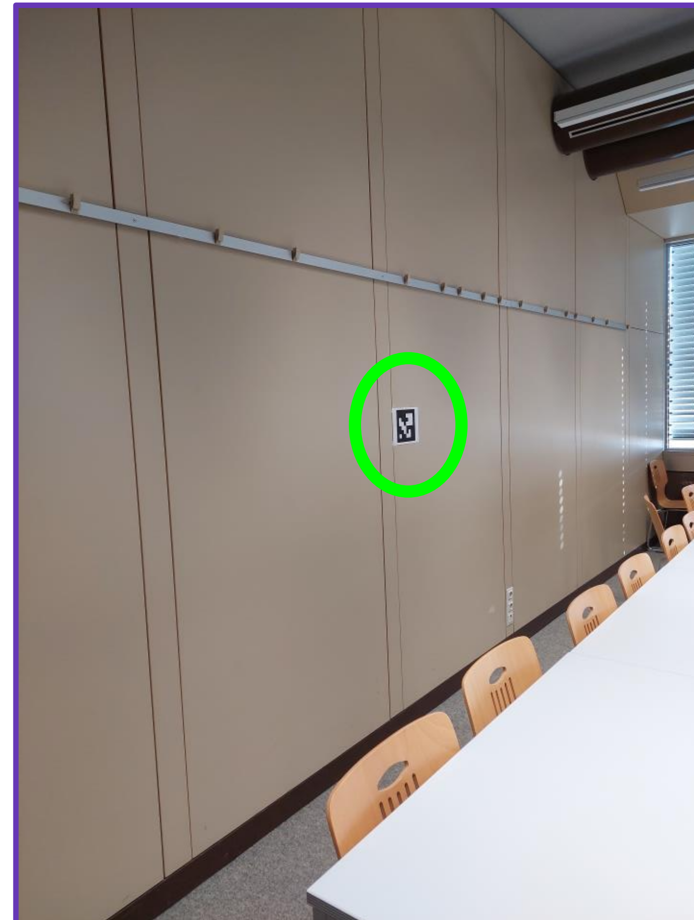


Augmented Reality
Mixing reality with digital information

But a simple copy paste will not work!

We need additional information of the 3D positioning of the wall inside the scene which is projected onto the 2D image.

For this, we can use some defined markers as seen on the right (marker is marked with a green circle).

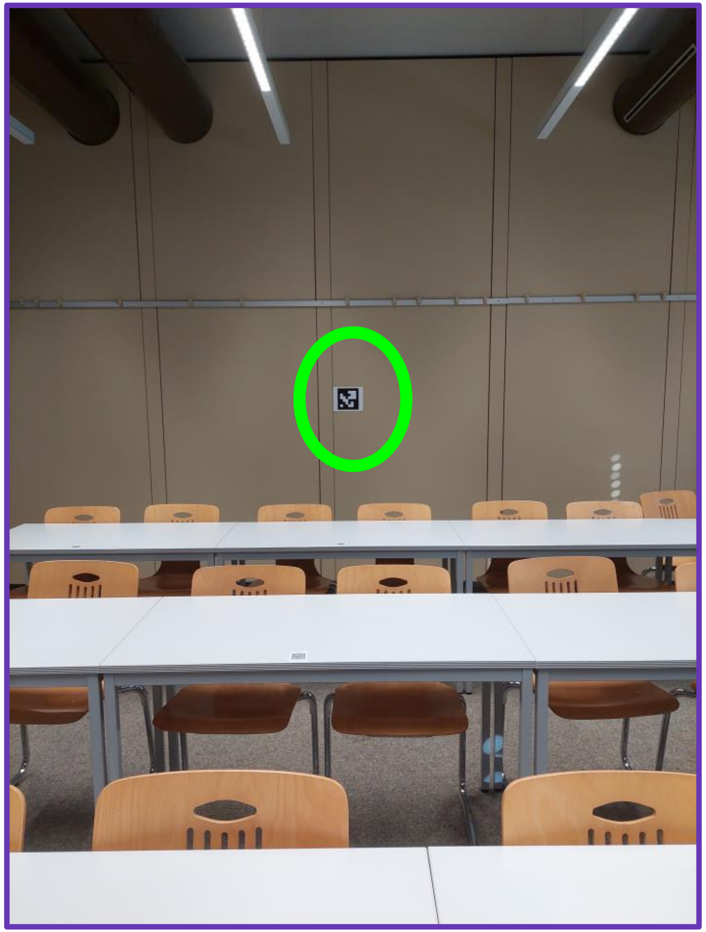


Augmented Reality
Mixing reality with digital information

But a simple copy paste will not work!

We need additional information of the 3D positioning of the wall inside the scene which is projected onto the 2D image.

For this, we can use some defined markers as seen on the right (marker is marked with a green circle).



Augmented Reality
AR marker

These markers should be defined in a way to allow save and stable detection even when printed on simple paper.

Markers like this are typically used in **augmented reality (AR)** applications or calibration processes.

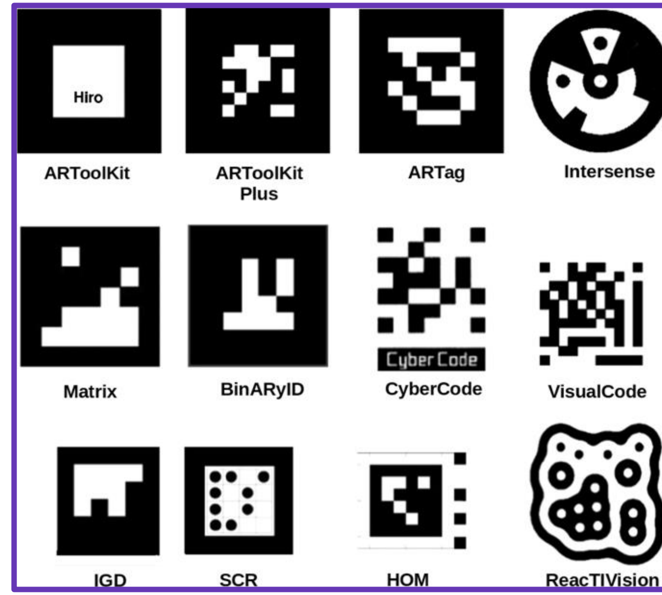


Fig.: Several AR marker.
Source: [GMM]

Augmented Reality
ArUco Marker

One very stable version of these AR markers are the markers developed by the university of Córdoba (UCO).

These markers as seen on the right are therefore called **ArUco markers**. They can be easily detected using OpenCV.

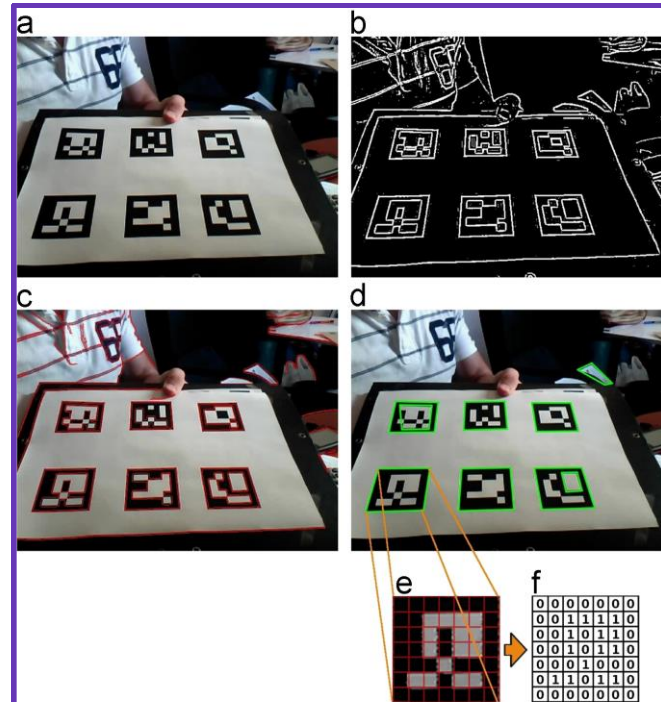
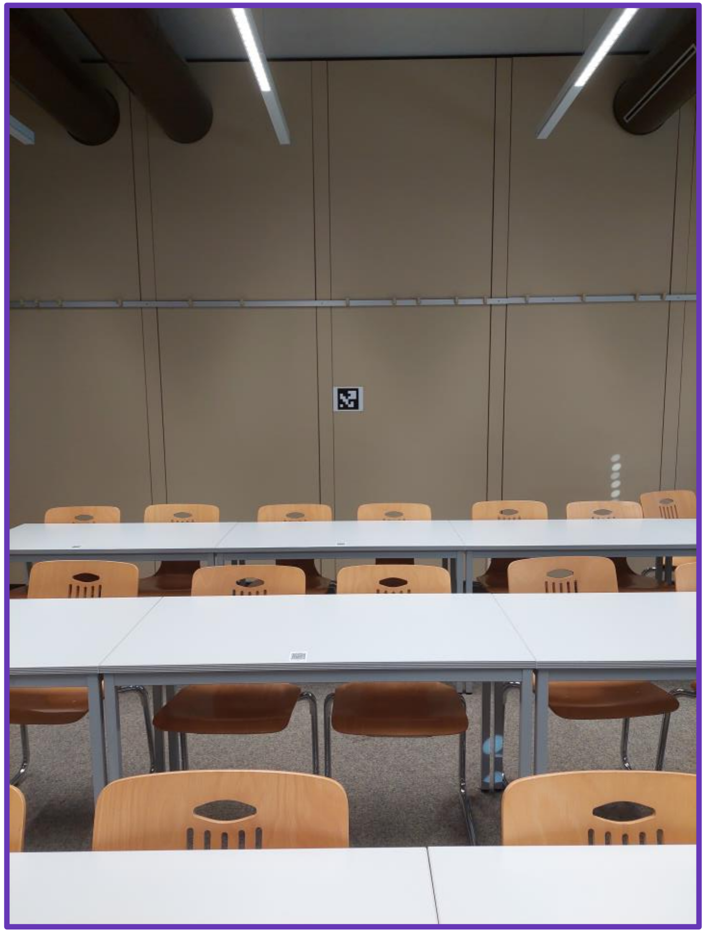


Fig.: Some printed ArUco markers and their step by step detection. Source: [GMM]

Augmented Reality
Mixing reality with digital information

Using the dataset “Room with ArUco Markers” provided on our Moodle course, we can try this approach.

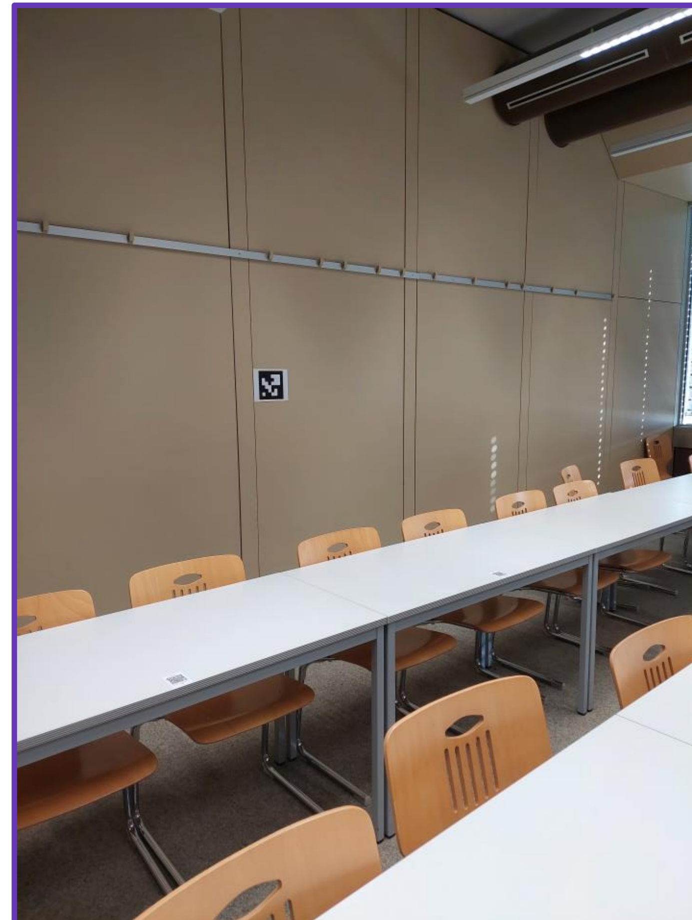
Hopefully we can detect and place on all images the corresponding ArUco markers.



Augmented Reality
Mixing reality with digital information

Using the dataset “Room with ArUco Markers” provided on our Moodle course, we can try this approach.

Hopefully we can detect and place on all images the corresponding ArUco markers.



Augmented Reality
Mixing reality with digital information

Using the dataset “Room with ArUco Markers” provided on our Moodle course, we can try this approach.

Hopefully we can detect and place on all images the corresponding ArUco markers.

Augmented Reality
Mixing reality with digital information

Using the dataset “Room with ArUco Markers” provided on our Moodle course, we can try this approach.

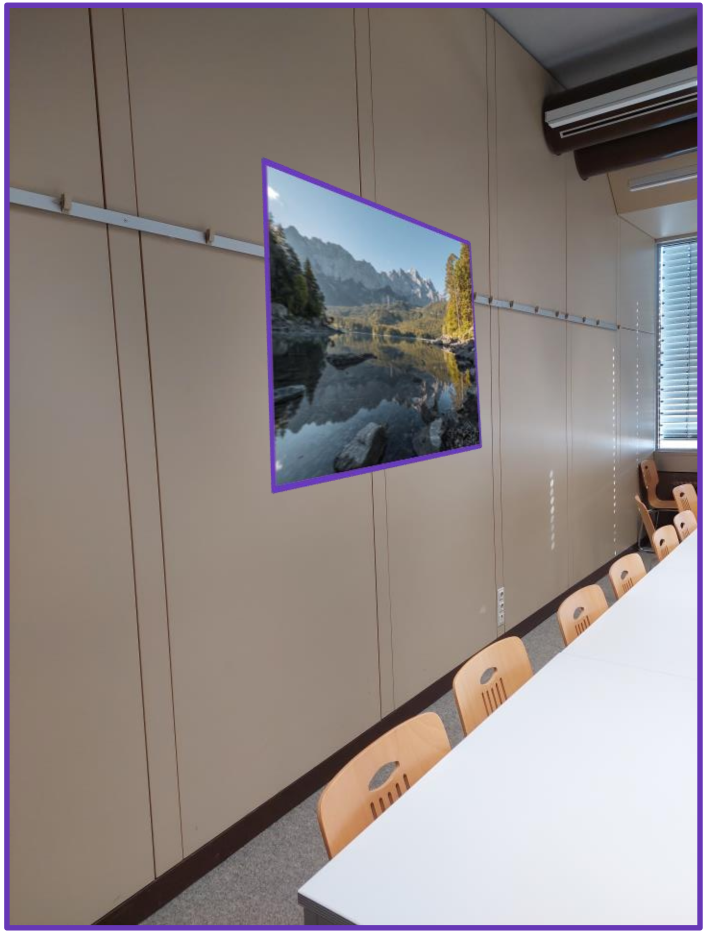
Hopefully we can detect and place on all images the corresponding ArUco markers.

Augmented Reality
Mixing reality with digital information

Using the dataset “Room with ArUco Markers” provided on our Moodle course, we can try this approach.

Hopefully we can detect and place on all images the corresponding ArUco markers.

This will allow us to place the poster onto the projected 3D plane!



Augmented Reality
Mixing reality with digital information

Warning:

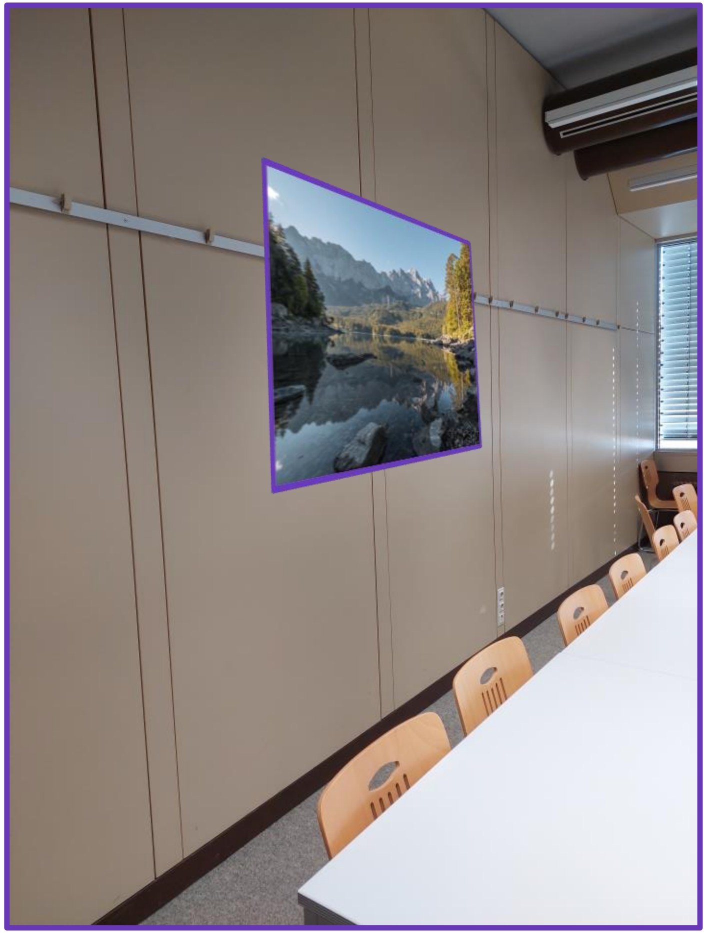
This is only a mock-up what you should achieve using the ArUco markers. The image on the right is done by hand and not correct.

Augmented Reality
Mixing reality with digital information

Your algorithm:

To understand the problem and its solution, read the original paper [GMM] as well as the corresponding OpenCV documentation.

Provide the geometrical/mathematical reasoning how to use the detection to correctly place a poster on the wall.

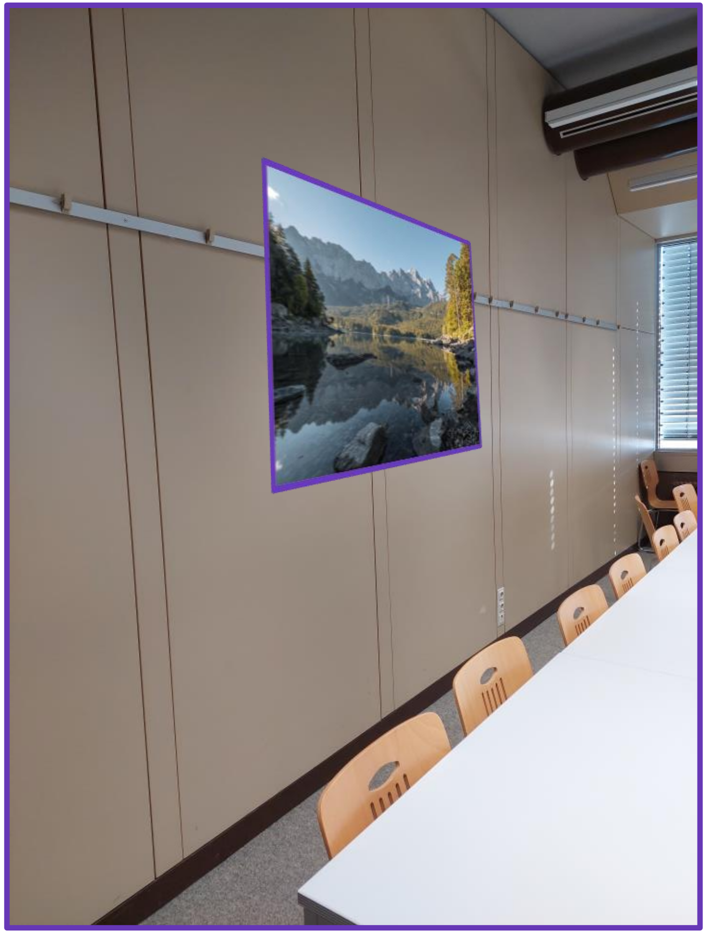


Augmented Reality
Mixing reality with digital information

Your evaluation:

After you read through the material, it will be easy to validate your result using the information provided in the images.

Decide for every image whether or not your result is acceptable. Again, provide all geometrical/mathematical information needed to understand how you evaluated your result.



Task 1: Augmented Reality
What we have to do

Goals of task 1:

Use the dataset provided via moodle in the folder "Room with ArUco Markers".

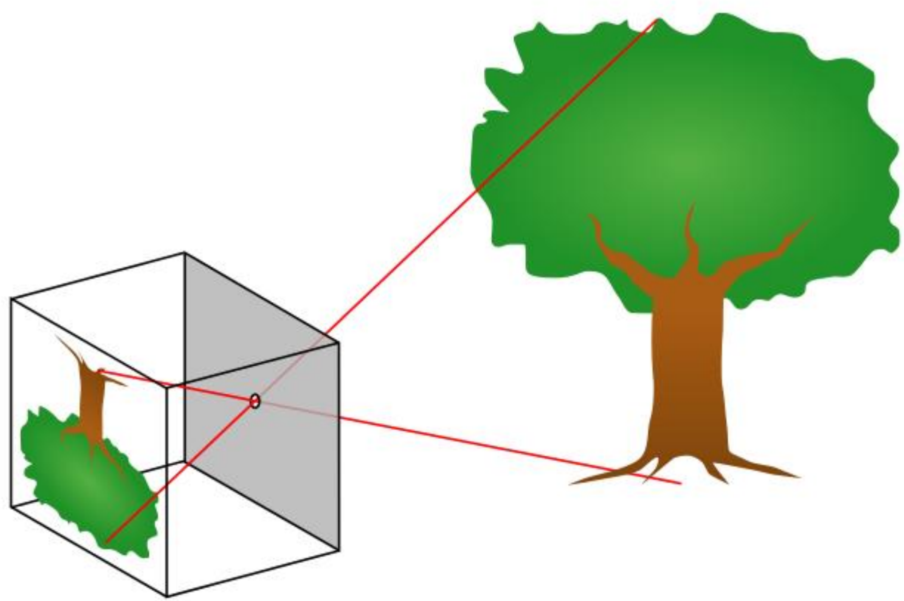
- Try to detect the ArUco markers in every image as described above.
- Provide your **AR images** for every frame in your report. Use the same poster in all images and place them always at the same position.
- Evaluate your result as described above.

Describe in your report how you did solve this problem, provide all information needed to reproduce your algorithms. Include images where they are helpful to describe your steps.

TASK 2:
CAMERA CALIBRATION AND DEPTH ESTIMATION

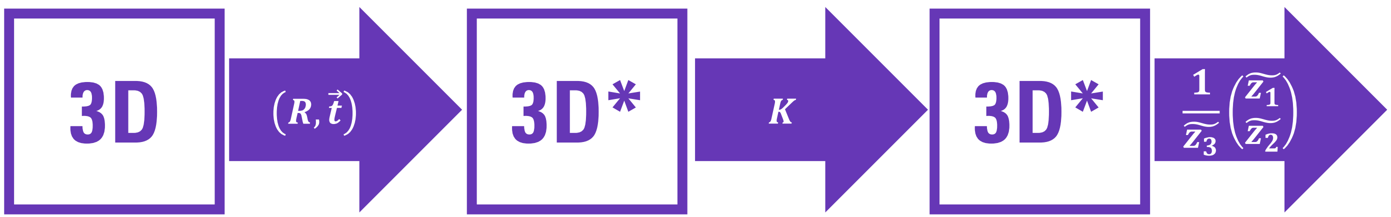
Camera Calibration
The pinhole camera model

In chapter 7, we will find a description how **3D environment** is projected onto the **2D image**.



Camera Calibration
The pinhole camera model

Using the results of chapter 7, we will be able to project the 3D environment onto the 2D image:



World Coordinates

Camera Coordinates

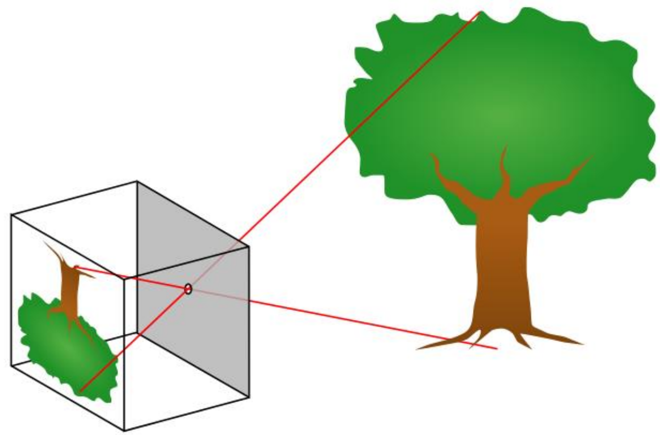
Helper Coordinates $\begin{pmatrix}\widetilde{z_1}\\ \widetilde{z_2}\\ \widetilde{z_3}\end{pmatrix}$

Pixel Coordinates $\begin{pmatrix}z_1\\ z_2\end{pmatrix}$

Camera Calibration
The intrinsic matrix

For this projection, we defined the intrinsic matrix K where

$$K := \begin{pmatrix} f \cdot k_1 & 0 & s_1 \\ 0 & f \cdot k_2 & s_2 \\ 0 & 0 & 1 \end{pmatrix}$$

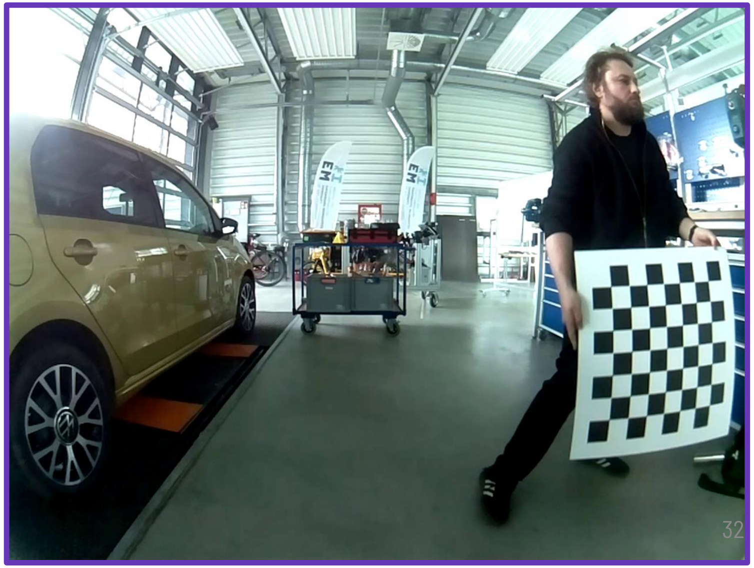


The parameters $\boldsymbol{f}, \boldsymbol{k_1}, \boldsymbol{k_2}, \boldsymbol{s_1}$, and $\boldsymbol{s_2}$ are defined by the structure of the camera sensor itself, they are called the **inner** or **intrinsic parameters**.

Camera Calibration
Estimating the intrinsic matrix

This matrix can be estimated using recordings of a checkerboard as seen on the right

Ideally, these recordings are done in different positions to get an overall good estimation.

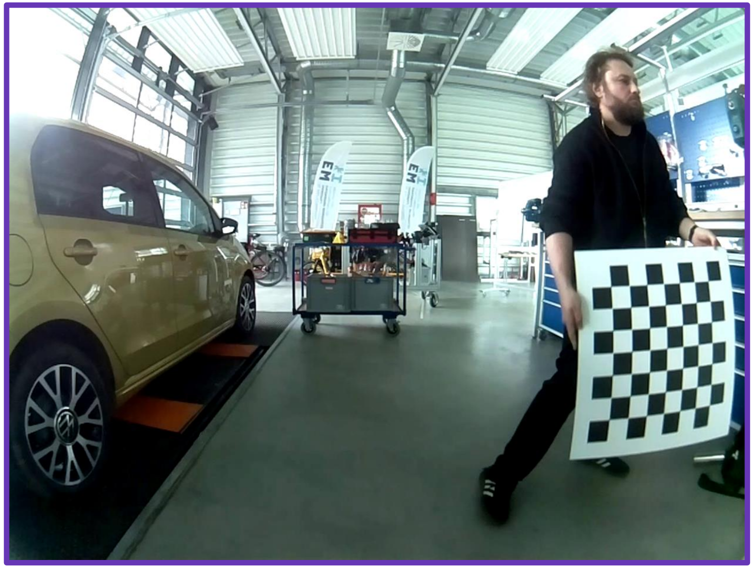


Camera Calibration

One more problem, though

But we realize another problem right away:

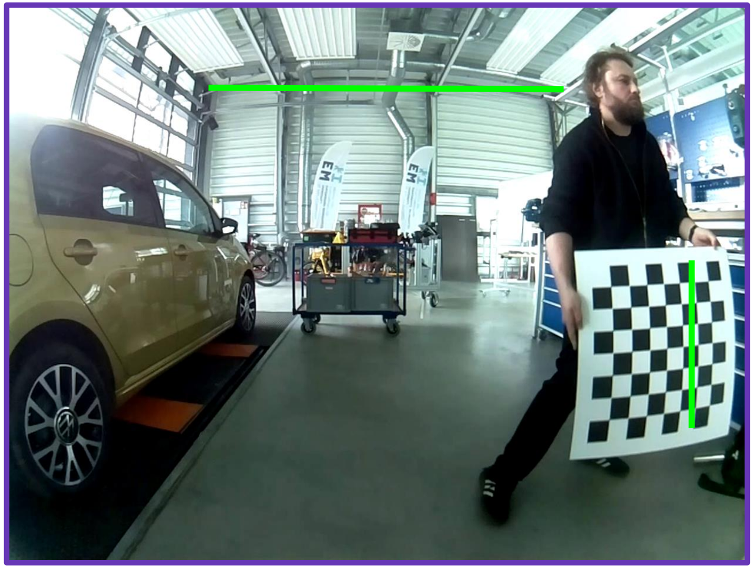
The camera used is not following the pinhole camera mode!



Camera Calibration
One more problem, though

Objects with nice geometries are not represented as expected.

Here, the edges of the checkerboard, the edges of the cupboard, or the edges on the ceiling are not on a line!



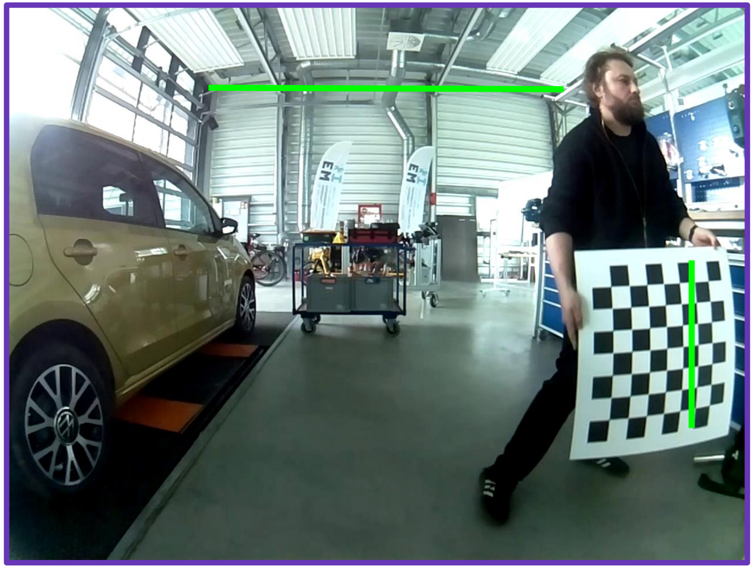

EM
EM


EM

Camera Calibration
Lens distortion

This effect is called lens distortion (covered in chapter 8) and can also be tackled using the recordings with the checkerboard.

In our moodle course, you can find the images to the right in video **“videoWW1_calibration.avi”** the folder **“Camera Calibration and Depth Estimation”**.



Camera Calibration
Lens distortion

After you are done with the undistortion, you will be able to provide the following image (or similar):

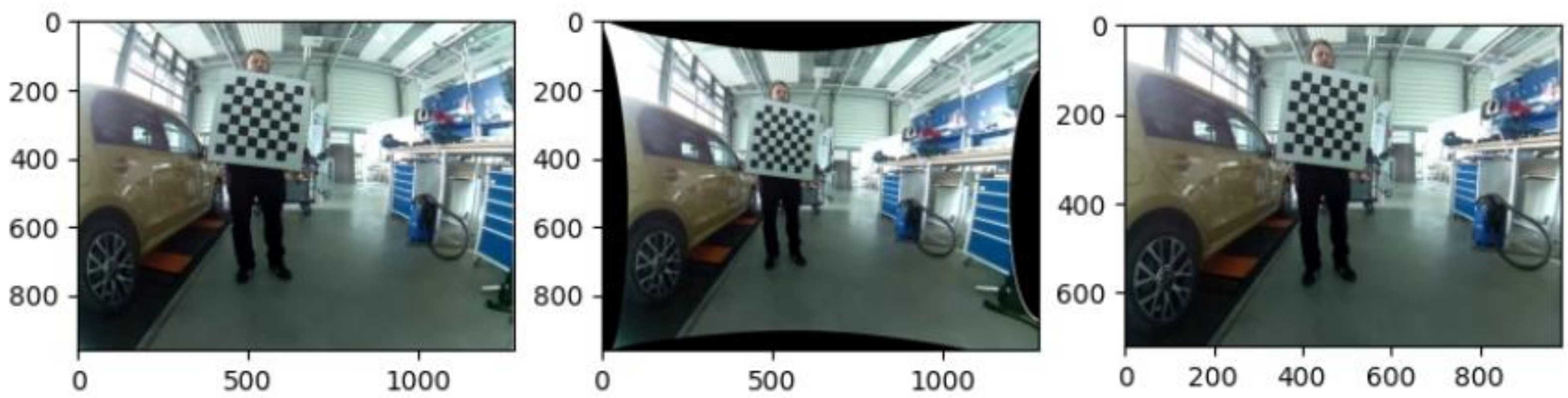


Fig.: From left to right: distorted, undistorted, undistorted + cropped.
Note that on the image on the right almost no distortion is visible.

Task 2: Camera Calibration and Depth Estimation
What we have to do

Goals of task 2 – First part: Undistortion

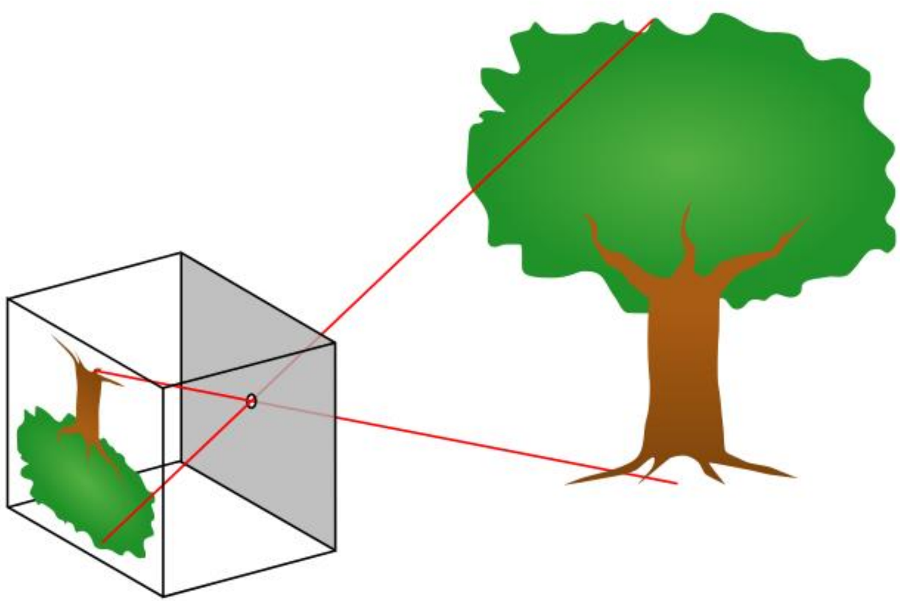
Use the video "**videoWW1_calibration.avi**" provided. Undistort the recordings:

- Provide the distortion coefficients
- Provide the three images (distorted, undistorted, undistorted + cropped) seen above for at least four different images from the video.

No we can proceed to the depth estimation...

Camera Calibration
The pinhole camera model

After the undistortion is done, we can use the pinhole camera model and the results of chapter 7, we will be able to project the 3D environment onto the 2D image:



Camera Calibration

The pinhole camera model

Using the results of chapter 7, we will be able to project the 3D environment onto the 2D image:

$$\mathbf{3D} \xrightarrow{(R,\vec{t})} \mathbf{3D^*} \xrightarrow{K} \mathbf{3D^*} \xrightarrow{\frac{1}{\widetilde{z_3}}\begin{pmatrix}\widetilde{z_1}\\ \widetilde{z_2}\end{pmatrix}} \mathbf{2D}$$

World Coordinates

Camera Coordinates

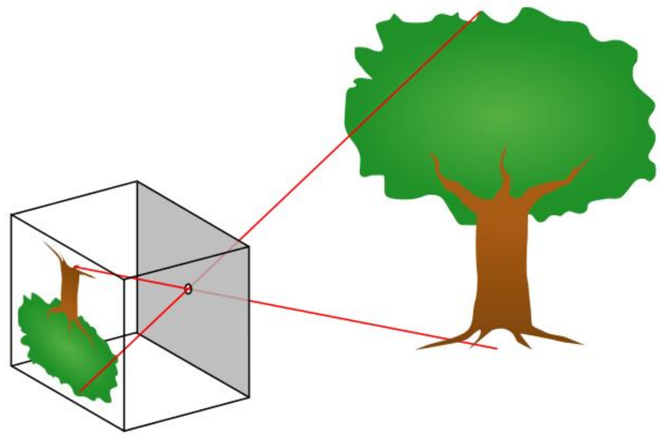
Helper Coordinates $\begin{pmatrix}\widetilde{z_1}\\ \widetilde{z_2}\\ \widetilde{z_3}\end{pmatrix}$

Pixel Coordinates $\begin{pmatrix}z_1\\ z_2\end{pmatrix}$

Camera Calibration
The intrinsic matrix

For this projection, we will define the intrinsic matrix K where

$$K \coloneqq \begin{pmatrix} f \cdot k_1 & 0 & s_1 \\ 0 & f \cdot k_2 & s_2 \\ 0 & 0 & 1 \end{pmatrix}$$

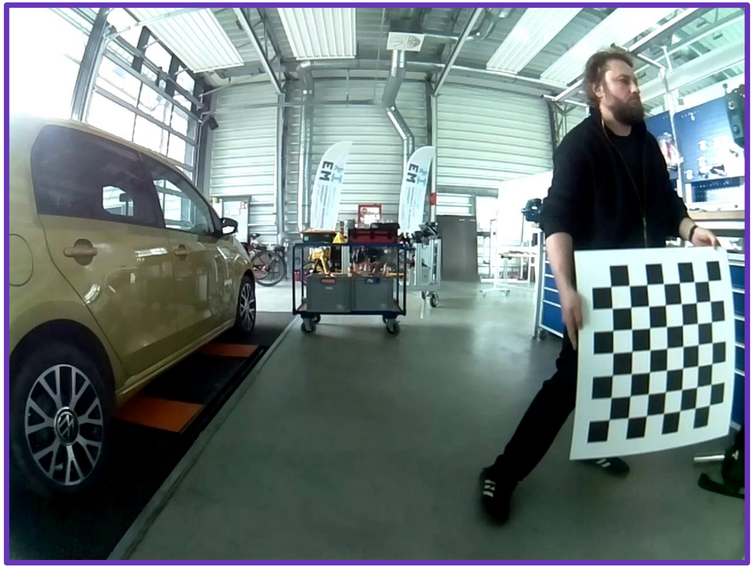


The parameters $\boldsymbol{f}, \boldsymbol{k_1}, \boldsymbol{k_2}, \boldsymbol{s_1}$, and $\boldsymbol{s_2}$ are defined by the structure of the camera sensor itself, they are called the **inner** or **intrinsic parameters**.

Depth Information
Using the intrinsic matrix

Before we can use the pinhole camera model, we have to get rid of the camera distortion. As we now have access to the undistorted images, we can use these recording to calculate the intrinsic matrix!

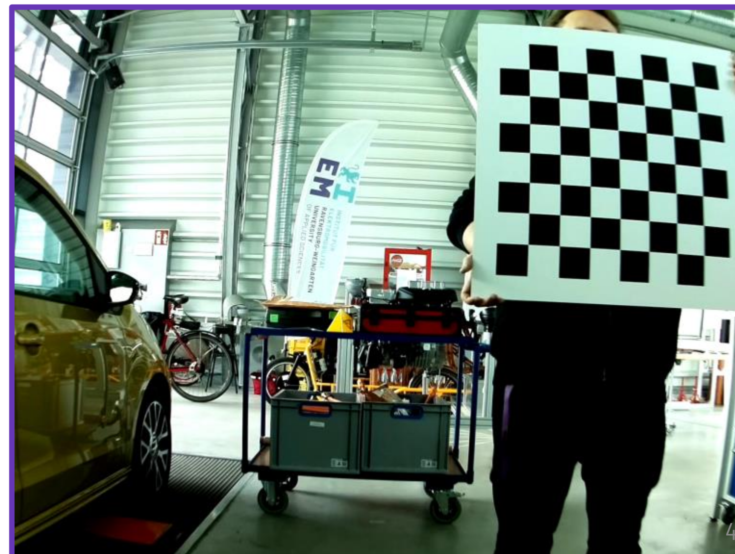
We want to use this information to estimate the distances in a recording!



Depth Information
Using the intrinsic matrix

We want to use this information to estimate the distances in a recording!

To allow easy access to this next step, we do provide additional recordings which are not heavily distorted.



Depth Information
Using the intrinsic matrix

Fig.: How far away from the camera is the trailer of the bicycle in this image?

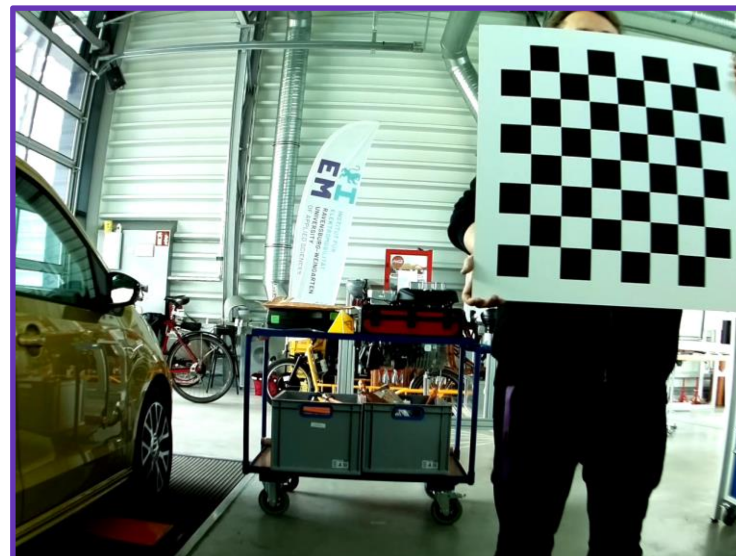
Depth Information
Using the intrinsic matrix

To make sure that we use the correct intrinsic matrix in our estimation, we provide enough images in the beginning of the video to allow a **new camera calibration for each video**.

Depth Information
Using the intrinsic matrix

To simplify our situation, we assume that the world coordinate system is the same as the camera coordinate system.

But how exactly can we use the intrinsic matrix K to estimate the distance?



Depth Information
Using the intrinsic matrix

Using the intrinsic matrix K we get a full line as potential 3D positions for a single pixel $(z_1, z_2)^t$ in the image:

$$L = \left\{ \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \in \mathbb{R}^3; \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = z \cdot R^t K^{-1} \begin{pmatrix} z_1 \\ z_2 \\ 1 \end{pmatrix} - R^t \cdot \vec{t} \text{ and } z \in \mathbb{R} \right\}$$

Depth Information
Using the intrinsic matrix

Using also the simplification of the scenario ($R = E_3$ and $\vec{t} = \vec{0}$), we get the following equation:

$$L = \left\{ \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \in \mathbb{R}^3 ; \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = zK^{-1} \begin{pmatrix} z_1 \\ z_2 \\ 1 \end{pmatrix} \text{and } z \in \mathbb{R} \right\}$$

Depth Information
Using the intrinsic matrix

In addition to the recordings, we now get the true size of the checkerboard:

The full size of the checkerboard is exactly 60 cm by 60 cm. The pattern covers exactly 50.6 cm by 50.6 cm on the board.

We now want to use this information to estimate the distances and sizes of two objects.

Depth Information
Using the intrinsic matrix

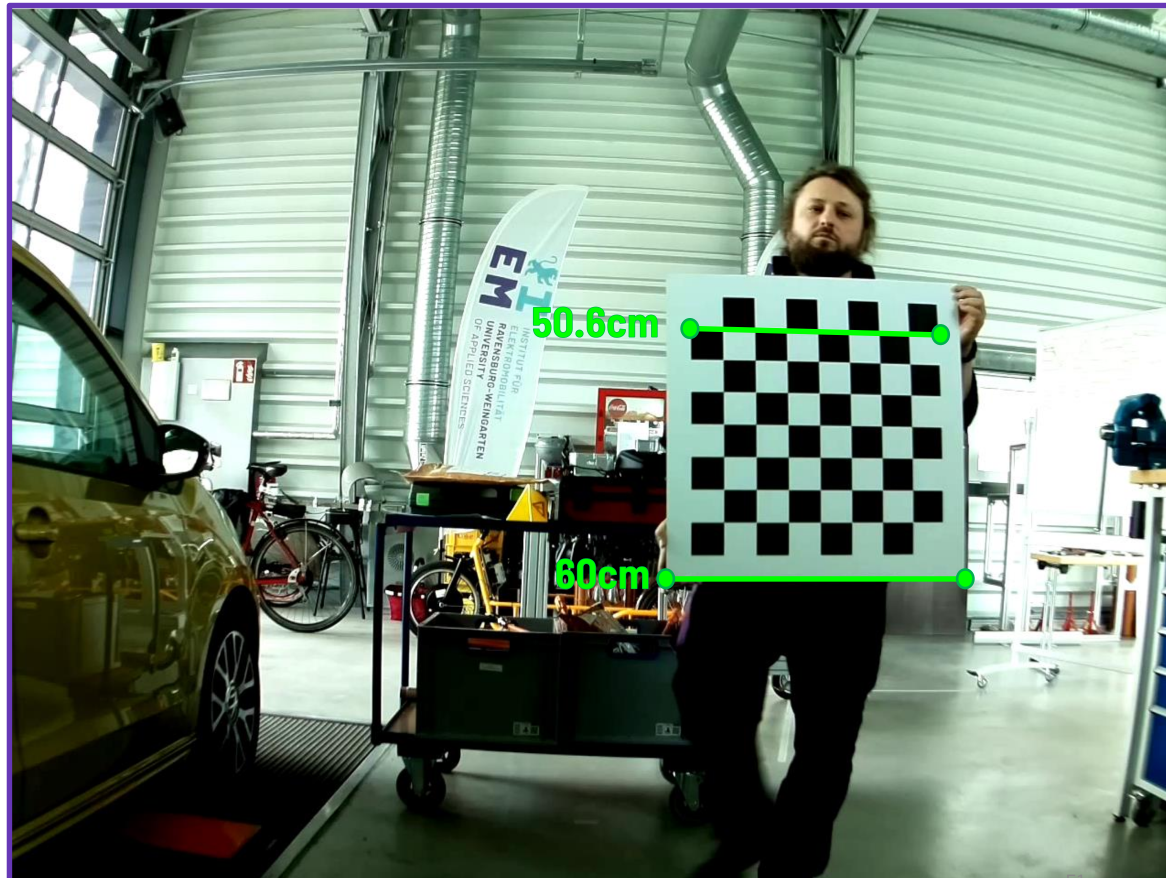


Fig.: Actual size of the board and the pattern.
Board: 60cm x 60cm
Pattern: 50.6cm x 50.6cm

Depth Information
Using the intrinsic matrix

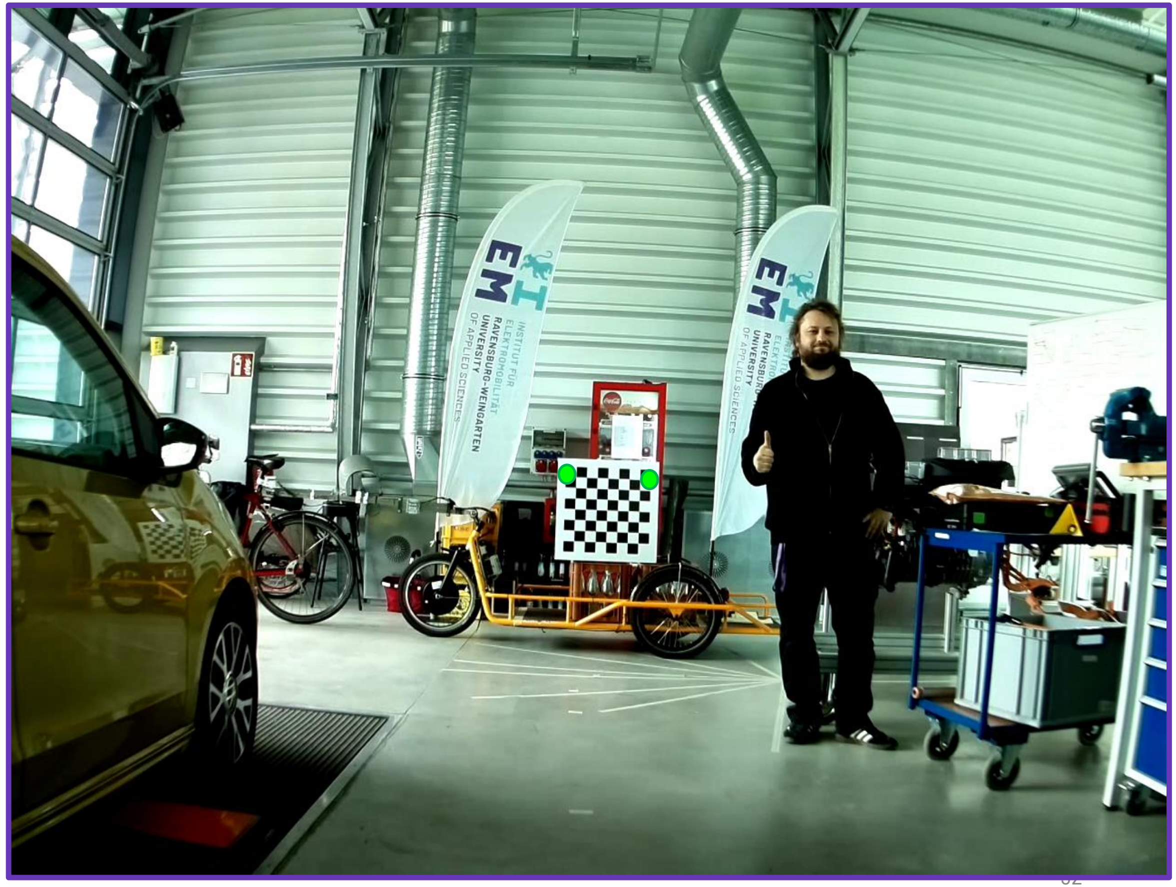


Fig.: We can select two points in any frame. You can do this by manual selection or using the result of a checkerboard detector.

Depth Information
Using the intrinsic matrix

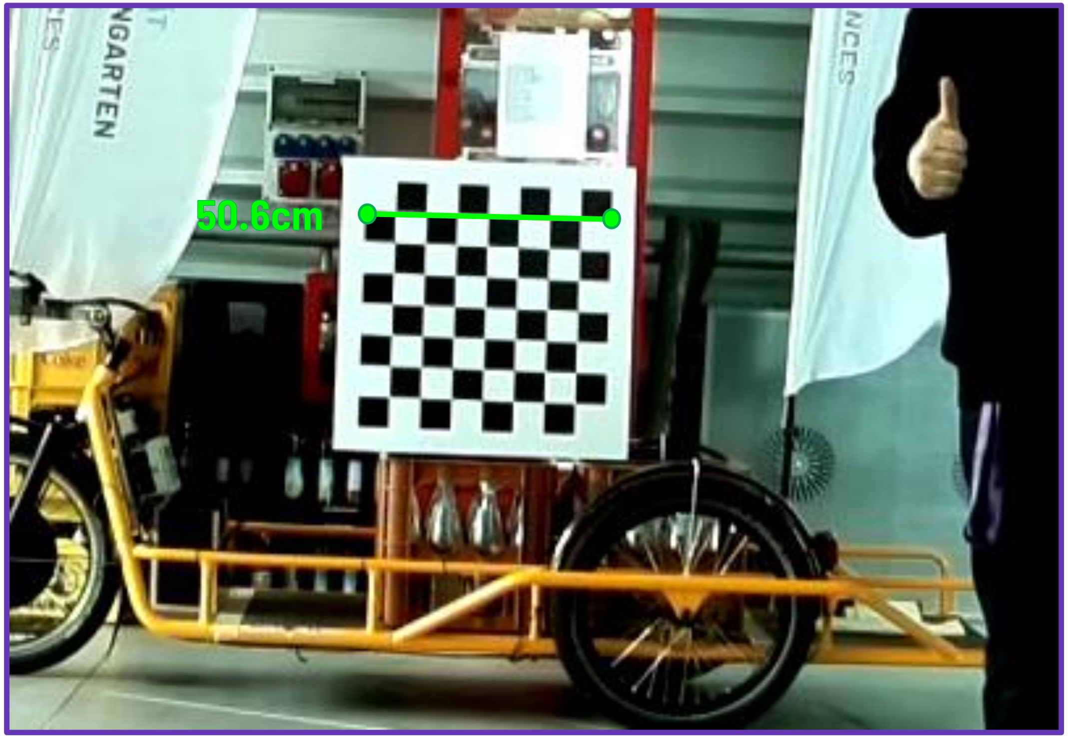
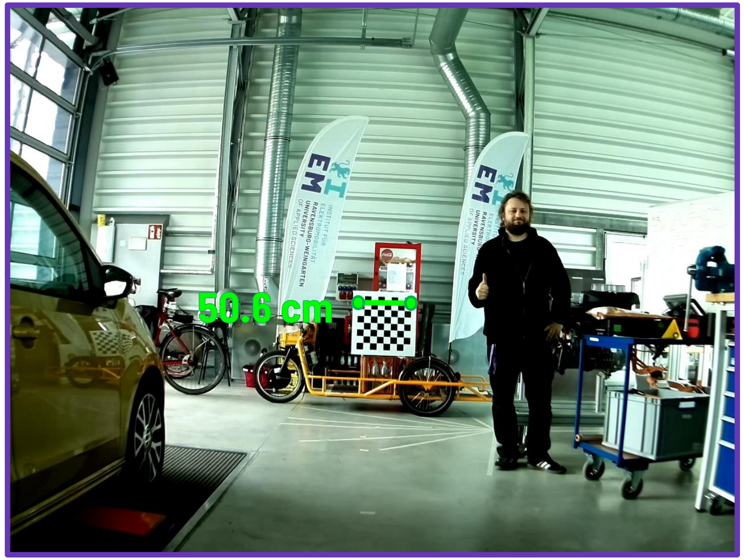


Fig.: The real distance between these points is known: 50.6 cm. Look closely: the actual edge of the board is covered by the metal construct. So we need to use the size of the actual pattern not the board itself.

Depth Information
Using the intrinsic matrix

Both points generate a line of the potential 3D position as seen above.

We assume the checkerboard to be perpendicular to the camera.



Using the additional information that the two points have to be 50.6cm away from each other, we should be able to estimate the distance of the motor block to the camera using some geometry!

Depth Information
Testing your solution

You will realize that the video is provided with three objects in three different distances.

Fig.: Tool cart in unknown distance.

Depth Information
Testing your solution

You will realize that the video is provided with three objects in three different distances.

Fig.: Motor block in known distance of 4.7 m.

Depth Information
Testing your solution

You will realize that the video is provided with three objects in three different distances.

Fig.: Bike trailer in unknown distance.

Depth Information
Testing your solution

Fig.: Marked in green: recording with the motor block in known distance of **4.7m**.

Depth Information
Testing your solution

You can use the recording to test your approach.

In this recording, the checkerboard is placed at the motor block at a distance of 4.7m.

Fig.: Motor block in known distance of 4.7 m.

Depth Information
Testing your solution

In the first part of task 02, you will realize that the undistortion works best if a lot of images are provided.

For the calculation of the intrinsic matrix, this is not necessarily so!

Actually, we got the best results by using only a single frame!

Fig.: Motor block in known distance of 4.7 m.

Depth Information
Testing your solution

Using the full video for the estimation of the camera matrix will be to error prone as some also misdetections of the checker board will be used for the process.

Use the known distance of the motor block to decide, which frames can be used to estimate the camera matrix best!

Fig.: Motor block in known distance of 4.7 m.

Task 2: Camera Calibration and Depth Estimation
What we have to do

Goals of task 2 – First part: Undistortion

Use the video “**videoWW1_calibration.avi**” provided. Undistort the recordings:

- Provide the distortion coefficients
- Provide the three images (distorted, undistorted, undistorted + cropped) seen above for at least four different images from the video.

No we can proceed to the depth estimation...

Task 2: Camera Calibration and Depth Estimation
What we have to do

Goals of task 2 – Second part: Depth estimation

Use the video "**videoHD1.avi**" provided via moodle.

- Provide all intrinsic camera parameters and the frame ID(s) you used.
- Provide the geometrical/mathematical reasoning how you did calculate the distance. **Follow the approach above by using the two lines calculated with the help of the intrinsic matrix. Provide sketches where needed.**
- Provide your distance estimation for all three objects.

Describe in your report how you did solve this problem, provide all information needed to reproduce your algorithms. Include images where they are helpful to describe your steps.

FURTHER READING

Further Reading
Computer Vision

[GMM] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., & Marín-Jiménez, M. J. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, *47*(6), 2280-2292.

VIEL SPASS UND VIEL ERFOLG!

HOCHSCHULE
RAVENSBURG-WEINGARTEN
UNIVERSITY
OF APPLIED SCIENCES

Doggenriedstraße
88250 Weingarten

Postfach 3022
88216 Weingarten

www.rwu.de
info@rwu.de