# Practical Course: Vision-based Navigation Summer Term 2015 

Lecture 1: Basics

Dr. Jörg Stückler

## What we will cover today

- Linear algebra, notation
- 3D geometry
- Projective geometry
- Camera intrinsics
- Epipolar geometry
- Robot Operating System (ROS)


## What we will cover today

- Linear algebra, notation
- 3D geometry
- Projective geometry
- Camera intrinsics
- Epipolar geometry
- Robot Operating System (ROS)


## Vectors

- Vector and its coordinates

$$
\mathbf{x}=\left(\begin{array}{c}
x_{1} \\
x_{2} \\
\vdots \\
x_{n}
\end{array}\right) \in \mathbb{R}^{n}
$$

- Vectors represent points in an ndimensional space



## Vector Operations

- Scalar multiplication
- Addition/subtraction
- Length
- Normalized vector
- Dot product
- Cross product



## Vector Operations

- Scalar multiplication
- Addition/subtraction
- Length
- Normalized vector
- Dot product
- Cross product



## Vector Operations

- Scalar multiplication
- Addition/subtraction
- Length
- Normalized vector
- Dot product
- Cross product


$$
\|x\|_{2}=\|x\|=\sqrt{x_{1}^{2}+x_{2}^{2}+\ldots}
$$

## Vector Operations

- Scalar multiplication
- Addition/subtraction
- Length
- Normalized vector
- Dot product
- Cross product



## Vector Operations

- Scalar multiplication
- Addition/subtraction
- Length
- Normalized vector
- Dot product
- Cross product


$$
\mathbf{x} \cdot \mathbf{y}=\|\mathbf{x}\|\|\mathbf{y}\| \cos \theta
$$

$\mathbf{x}, \mathbf{y}$ are orthogonal if $\mathbf{x} \cdot \mathbf{y}=0$
$\mathbf{y}$ is lin. dependent from $\left\{\mathbf{x}_{1}, \mathbf{x}_{2}, \ldots\right\}$ if $\mathbf{y}=\sum_{i} k_{i} \mathbf{x}_{i}$

## Vector Operations

- Scalar multiplication
- Addition/subtraction
- Length
- Normalized vector
- Dot product
- Cross product



## Cross Product

- Definition $\quad \mathbf{x} \times \mathbf{y}=\left(\begin{array}{l}x_{2} y_{3}-x_{3} y_{2} \\ x_{3} y_{1}-x_{1} y_{3} \\ x_{1} y_{2}-x_{2} y_{1}\end{array}\right)$
- Matrix notation for the cross product

$$
[\mathbf{x}]_{\times}=\left(\begin{array}{ccc}
0 & -x_{3} & x_{2} \\
x_{3} & 0 & -x_{1} \\
-x_{2} & x_{1} & 0
\end{array}\right)
$$

- Verify that $\quad \mathbf{x} \times \mathbf{y}=[\mathbf{x}]_{\times} \mathbf{y}$


## Matrices

- Rectangular array of numbers

$$
X=\left(\begin{array}{cccc}
x_{11} & x_{12} & \ldots & x_{1 m} \\
x_{21} & x_{22} & \ldots & x_{2 m} \\
\vdots & & & \\
x_{n 1} & x_{n 2} & \ldots & x_{n m}
\end{array}\right) \in \mathbb{R}^{n \times m} \downarrow \begin{gathered}
\text { rows columns } \\
\downarrow
\end{gathered}
$$

- First index refers to row
- Second index refers to column


## Matrices

- Column vectors of a matrix

- Geometric interpretation: for example, column vectors can form basis of a coordinate system


## Matrices

- Row vectors of a matrix


## Matrices

- Square matrix
- Diagonal matrix
- Upper and lower triangular matrix
- Symmetric matrix
- Skew-symmetric matrix
- (Semi-)positive definite matrix
- Invertible matrix
- Orthonormal matrix
- Matrix rank


## Matrices

- Square matrix
- Diagonal matrix
- Upper and lower triangular matrix
- Symmetric matrix $\quad X=X^{\top}$
- Skew-symmetric matrix
- (Semi-)positive definite matrix
- Invertible matrix

$$
\mathbf{a}^{\top} X \mathbf{a} \geq 0
$$

- Orthonormal matrix
- Matrix rank


## Matrix Operations

- Scalar multiplication
- Addition/subtraction
- Transposition
- Matrix-vector multiplication
- Matrix-matrix multiplication
- Inversion


## Matrix-Vector Multiplication

$$
X \cdot \mathbf{b}=\sum_{k=1}^{n} \mathbf{x}_{* k} \cdot b_{k}
$$



- Geometric interpretation:

A linear combination of the columns of $X$ scaled by the coefficients of $\mathbf{b}$
$\rightarrow$ coordinate transf. from local to global frame

## Matrix-Matrix Multiplication

- Operator
- Definition

$$
\begin{aligned}
& \mathbb{R}^{n \times m} \times \mathbb{R}^{m \times p} \rightarrow \mathbb{R}^{n \times p} \\
& C=A B \\
& \quad=A\left(\begin{array}{llll}
\mathbf{b}_{* 1} & \mathbf{b}_{* 2} & \cdots \mathbf{b}_{* p}
\end{array}\right)
\end{aligned}
$$

- Interpretation: transformation of coordinate systems
- Can be used to concatenate transforms


## Matrix-Matrix Multiplication

- Not commutative (in general) $\quad A B \neq B A$
- Associative

$$
A(B C)=(A B) C
$$

- Transpose

$$
(A B)^{\top}=B^{\top} A^{\top}
$$

## Matrix Inversion

- If $A$ is a square matrix of full rank, then there is a unique matrix $B=A^{-1}$ such that $A B=I$.
- Different ways to compute, e.g., Gauss-Jordan elimination, LU decomposition, ...
- When $A$ is orthonormal, then $A^{-1}=A^{\top}$


## What we will cover today

- Linear algebra, notation
- 3D geometry
- Projective geometry
- Camera intrinsics
- Epipolar geometry
- Robot Operating System (ROS)


## Geometric Primitives in 3D

- 3D point (same as before)
- Augmented vector

$$
\overline{\mathbf{x}}=\left(\begin{array}{c}
x \\
y \\
z \\
1
\end{array}\right) \in \mathbb{R}^{4}
$$

- Homogeneous coordinates

$$
\mathbf{x}=\left(\begin{array}{l}
x \\
y \\
z
\end{array}\right) \in \mathbb{R}^{3}
$$

$$
\tilde{\mathbf{x}}=\left(\begin{array}{c}
\tilde{x} \\
\tilde{y} \\
\tilde{z} \\
\tilde{w}
\end{array}\right) \in \mathbb{P}^{3}
$$

## 3D Transformations

- Translation

$$
\overline{\mathbf{x}}^{\prime}=\underbrace{\left(\begin{array}{cc}
\mathbf{I} & \mathbf{t} \\
\mathbf{0}^{\top} & 1
\end{array}\right)}_{4 \times 4} \overline{\mathbf{x}}
$$

- Euclidean transform (translation + rotation), (also called the Special Euclidean group SE(3))

$$
\overline{\mathbf{x}}^{\prime}=\left(\begin{array}{cc}
\mathbf{R} & \mathbf{t} \\
\mathbf{0}^{\top} & 1
\end{array}\right) \overline{\mathbf{x}}
$$

- Scaled rotation, affine transform, projective transform...


## 3D Transformations

| Transformation | Matrix | \# DoF | Preserves | Icon |
| :--- | :--- | :--- | :--- | :--- |
| translation | $[\boldsymbol{I} \mid \boldsymbol{t}]_{3 \times 4}$ | 3 | orientation |  |
| rigid (Euclidean) | $[\boldsymbol{R} \mid \boldsymbol{t}]_{3 \times 4}$ | 6 | lengths |  |
| similarity | $[s \boldsymbol{R} \mid \boldsymbol{t}]_{3 \times 4}$ | 7 | angles |  |
| affine | $[\boldsymbol{A}]_{3 \times 4}$ | 12 | parallelism |  |
| projective | $[\tilde{\boldsymbol{H}}]_{4 \times 4}$ | 15 | straight lines |  |

## 3D Euclidean Transformtions

- Translation $\mathbf{t}$ has 3 degrees of freedom
- Rotation $R$ has 3 degrees of freedom

$$
X=\left(\begin{array}{ll}
R & \mathbf{t} \\
\mathbf{0} & 1
\end{array}\right)=\left(\begin{array}{cccc}
r_{11} & r_{12} & r_{13} & t_{1} \\
r_{21} & r_{22} & r_{23} & t_{2} \\
r_{31} & r_{32} & r_{33} & t_{3} \\
0 & 0 & 0 & 1
\end{array}\right)
$$

## 3D Rotations

- Rotation matrix $R$ (also called the special orientation group $\mathrm{SO}(3)$ )

Alternative representations

- Euler angles
- Axis/angle
- Unit quaternion


## Rotation Matrix

- Orthonormal $3 \times 3$ matrix

$$
R=\left(\begin{array}{lll}
r_{11} & r_{12} & r_{13} \\
r_{21} & r_{22} & r_{23} \\
r_{31} & r_{32} & r_{33}
\end{array}\right)
$$

- Advantage: Can be easily concatenated and inverted (how?)
- Disadvantage: Over-parameterized (9 parameters instead of 3)


## Euler Angles

- Product of 3 consecutive rotations (e.g., around $X-Y-Z$ axes)
- Roll-pitch-yaw convention is very common in aerial navigation (DIN 9300)



## Roll-Pitch-Yaw Convention

- Yaw $\Psi$, Pitch $\Theta$, Roll $\Phi$ to rotation matrix

$$
\begin{aligned}
R & =R_{Z}(\Psi) R_{Y}(\Theta) R_{X}(\Phi) \\
& =\left(\begin{array}{ccc}
1 & 0 & 0 \\
0 & \cos \Phi & \sin \Phi \\
0 & -\sin \Phi & \cos \Phi
\end{array}\right)\left(\begin{array}{ccc}
\cos \Theta & 0 & -\sin \Theta \\
0 & 1 & 0 \\
\sin \Theta & 0 & \cos \Theta
\end{array}\right)\left(\begin{array}{ccc}
\cos \Psi & \sin \Psi & 0 \\
-\sin \Psi & \cos \Psi & 0 \\
0 & 0 & 1
\end{array}\right) \\
& =\left(\begin{array}{ccc}
\cos \Theta \cos \Psi & \cos \Theta \sin \Psi & -\sin \Theta \\
\sin \Phi \sin \Theta \cos \Psi-\cos \Phi \sin \Psi & \sin \Phi \sin \Theta \sin \Psi+\cos \Phi \cos \Psi & \sin \Phi \cos \Theta \\
\cos \Phi \sin \Theta \cos \Psi+\sin \Phi \sin \Psi & \cos \Phi \sin \Theta \sin \Psi-\sin \Phi \cos \Psi & \cos \Phi \cos \Theta
\end{array}\right)
\end{aligned}
$$

- Rotation matrix to Yaw-Pitch-Roll

$$
\begin{aligned}
\phi & =\operatorname{Atan} 2\left(-r_{31}, \sqrt{r_{11}^{2}+r_{21}^{2}}\right) \\
\psi & =-\operatorname{Atan} 2\left(\frac{r_{21}}{\cos (\phi)}, \frac{r_{11}}{\cos (\phi)}\right) \\
\theta & =\operatorname{Atan} 2\left(\frac{r_{32}}{\cos (\phi)}, \frac{r_{33}}{\cos (\phi)}\right)
\end{aligned}
$$

## Gimbal Lock

- When the axes align, one degree-of-freedom (DOF) is lost...



## Axis/Angle

- Represent rotation by
- rotation axis $\hat{\mathbf{n}}$ and
- rotation angle $\theta$
- 4 parameters $(\hat{\mathbf{n}}, \theta)$
- 3 parameters $\boldsymbol{\omega}=\theta \hat{\mathbf{n}}$
- length is rotation angle
- also called the angular velocity
- minimal but not unique (why?)


## Conversion

- Rodriguez' formula

$$
R(\hat{\mathbf{n}}, \theta)=I+\sin \theta[\hat{\mathbf{n}}]_{\times}+(1-\cos \theta)[\hat{\mathbf{n}}]_{\times}^{2}
$$

- Inverse

$$
\theta=\cos ^{-1}\left(\frac{\operatorname{trace}(R)-1}{2}\right), \hat{\mathbf{n}}=\frac{1}{2 \sin \theta}\left(\begin{array}{l}
r_{32}-r_{23} \\
r_{13}-r_{31} \\
r_{21}-r_{12}
\end{array}\right)
$$

- see: An Invitation to 3D Vision, Y. Ma, S. Soatto, J. Kosecka, S. Sastry, Chapter 2 (available online)


## Unit Quaternions

- Quaternion

$$
\mathbf{q}=\left(q_{x}, q_{y}, q_{z}, q_{w}\right)^{\top} \in \mathbb{R}^{4}
$$

- Unit quaternions have $\|\mathbf{q}\|=1$
- Opposite sign quaternions represent the same rotation
- Otherwise unique $\mathbf{q}=-\mathbf{q}$



## Unit Quaternions

- Advantage: multiplication and inversion operations are efficient
- Quaternion-Quaternion Multiplication

$$
\begin{aligned}
\mathbf{q}_{0} \mathbf{q}_{1} & =\left(\mathbf{v}_{0}, w_{0}\right)\left(\mathbf{v}_{1}, w_{1}\right) \\
& =\left(\mathbf{v}_{0} \times \mathbf{v}_{1}+w_{0} \mathbf{v}_{1}+w_{1} \mathbf{v}_{0}, w_{0} w_{1}-\mathbf{v}_{0} \mathbf{v}_{1}\right)
\end{aligned}
$$

- Inverse (flip sign of vor w)

$$
\begin{aligned}
\mathbf{q}^{-1} & =(\mathbf{v}, w)^{-1} \\
& =(\mathbf{v},-w)
\end{aligned}
$$

## Unit Quaternions

- Quaternion-Vector multiplication (rotate point p with rotation q)

$$
\mathbf{p}^{\prime}=\mathbf{q} \overline{\mathbf{p}} \mathbf{q}^{-1}
$$

with

$$
\overline{\mathbf{p}}=(x, y, z, 0)^{\top}
$$

- Relation to Axis/Angle representation

$$
\mathbf{q}=(\mathbf{v}, w)=\left(\sin \frac{\theta}{2} \hat{\mathbf{n}}, \cos \frac{\theta}{2}\right)
$$

## 3D Orientations

- Note: In general, it is very hard to "read" 3D orientations/rotations, no matter in what representation
- Observation: They are usually easy to visualize and can then be intuitively interpreted
- Advice: Use 3D visualization tools for debugging (RVIZ, libqgIviewer, ...)


## C++ Libraries for Lin. Alg./Geometry

- Many C++ libraries exist for linear algebra and 3D geometry
- Typically conversion necessary
- Examples:
- C arrays, std::vector (no linear alg. functions)
- gsl (gnu scientific library, many functions, plain C)
- boost::array (used by ROS messages)
- Bullet library (3D geometry, used by ROS tf)
- Eigen (both linear algebra and geometry)
- Sophus (Eigen extensions, SE(3)/se(3), see Lecture 2)


## What we will cover today

- Linear algebra, notation
- 3D geometry
- Projective geometry
- Camera intrinsics
- Epipolar geometry
- Robot Operating System (ROS)


## Pin-hole Camera

- Lit scene emits light
- Film/sensor is light sensitive



## 3D to 2D Perspective Projection



## 3D to 2D Perspective Projection

- 3D point $\mathbf{p}$ (in the camera frame)
- 2D point x (on the image plane)
- Pin-hole camera model

$$
\tilde{\mathbf{x}}=\left(\begin{array}{llll}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{array}\right) \tilde{\mathbf{p}}
$$

- Remember, $\tilde{\mathrm{x}}$ is homogeneous, need to normalize

$$
\tilde{\mathrm{x}}=\left(\begin{array}{c}
\tilde{x} \\
\tilde{y} \\
\tilde{z}
\end{array}\right) \quad \Rightarrow \quad \mathrm{x}=\binom{\tilde{x} / \tilde{z}}{\tilde{y} / \tilde{z}}
$$

## Camera Intrinsics

- So far, 2D point is given in meters on image plane
- But: we want 2D point be measured in pixels (as the sensor does)



## Camera Intrinsics

- Need to apply some scaling/offset

$$
\tilde{\mathbf{x}}=\underbrace{\left(\begin{array}{ccc}
f_{x} & s & c_{x} \\
0 & f_{y} & c_{y} \\
0 & 0 & 1
\end{array}\right)}_{\text {intrinsics } K} \underbrace{\left(\begin{array}{cccc}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{array}\right)}_{\text {projection }} \tilde{\mathbf{p}}
$$

- Focal length $f_{x}, f_{y}$
- Camera center $c_{x}, c_{y}$
- Skew $s$


## Camera Extrinsics

- Assume $\tilde{\mathbf{p}}_{w}$ is given in world coordinates
- Transform from world to camera (also called the camera extrinsics)

$$
\tilde{\mathbf{p}}=\left(\begin{array}{cc}
R & \mathbf{t} \\
\mathbf{0}^{\top} & 1
\end{array}\right) \tilde{\mathbf{p}}_{w}
$$

- Full camera matrix

$$
\tilde{\mathbf{x}}=\left(\begin{array}{ccc}
f_{x} & s & c_{x} \\
0 & f_{y} & c_{y} \\
0 & 0 & 1
\end{array}\right)\left(\begin{array}{ll}
R & \mathbf{t}
\end{array}\right) \tilde{\mathbf{p}}_{w}
$$

## Real Camera Lenses

- Lit scene emits light
- Film/sensor is light sensitive
- A lens focuses rays onto the film/sensor



## Real Cameras

- Radial distortion of the image
- Caused by imperfect lenses
- Deviations are most noticeable for rays that pass through the edge of the lens



## Radial Distortion

- Radial distortion of the image
- Caused by imperfect lenses
- Deviations are most noticeable for rays that pass through the edge of the lens
- Typically compensated with a low-order polynomial

$$
\begin{aligned}
& \hat{x}_{c}=x_{c}\left(1+\kappa_{1} r_{c}^{2}+\kappa_{2} r_{c}^{4}\right) \\
& \hat{y}_{c}=y_{c}\left(1+\kappa_{1} r_{c}^{2}+\kappa_{2} r_{c}^{4}\right)
\end{aligned}
$$

## Digital Cameras

- Vignetting
- De-bayering
- Rolling shutter and motion blur
- Compression (JPG)
- Noise



## Epipolar Geometry

- Let's consider two cameras that observe a 3D world point

$R, \mathbf{t}$


## Epipolar Geometry

- The line connecting both camera centers is called the baseline

(line joining both camera centers)


## Epipolar Geometry

- Given the image of a point in one view, what can we say about its position in another?

- A point in one image "generates" a line in another image (called the epipolar line)


## What we will cover today

- Linear algebra, notation
- 3D geometry
- Projective geometry
- Camera intrinsics
- Epipolar geometry
- Robot Operating System (ROS)


## Modern Robot Architectures

- Robots became rather complex systems
- Often, a large set of individual capabilities is needed
- Flexible composition of different capabilities for different tasks


## Best Practices for Robot Architectures

- Modular
- Robust
- De-centralized
- Facilitate software re-use
- Hardware and software abstraction
- Provide introspection
- Data logging and playback
- Easy to learn and to extend


## Robotic Middleware

- Provides infrastructure
- Communication between modules
- Data logging facilities
- Tools for visualization
- Several open-source systems available
- ROS (Robot Operating System),
- Player/Stage,
- CARMEN,
- YARP,
- OROCOS


## Example Architecture for Navigation



## PR2 Software Architecture

- Two 7-DOF arms, grippers, torso, 2-DOF head
- 7 cameras, 2 laser scanners
- Two 8-core CPUs, 3 network switches
- 73 nodes, 328 message topics, 174 services



## Communication Paradigms

- Message-based communication

- Direct (shared) memory access



## Forms of Communication

- Push
- Pull
- Publisher/subscriber
- Publish to blackboard
- Remote procedure calls / service calls
- Preemptive tasks / actions


## Push

- Broadcast
- One-way communication
- Send as the information is generated by the producer P



## Pull

- Data is delivered upon request by the consumer C (e.g., a map of the building)
- Useful if the consumer C controls the process and the data is not required (or available) at high frequency



## Publisher/Subscriber

- The consumer $C$ requests a subscription for the data by the producer $P$ (e.g., a camera or GPS)
- The producer $P$ sends the subscribed data as it is generated to C
- Data generated according to a trigger (e.g., sensor data, computations, other messages, ...)



## Publish to Blackboard

- The producer $P$ sends data to the blackboard (e.g., parameter server)
- A consumer C pull data from the blackboard B
- Only the last instance of data is stored in the blackboard B



## Service Calls

- The client $C$ sends a request to the server $S$
- The server returns the result
- The client waits for the result (synchronous communication)
- Also called: Remote Procedure Call



## Actions (Preemptive Tasks)

- The client requests the execution of an enduring action (e.g., navigate to a goal location)
- The server executes this action and sends continuously status updates
- Task execution may be canceled from both sides (e.g., timeout, new navigation goal,...)


## Robot Operating System (ROS)

- We will use ROS in the lab course
- http://www.ros.org/
- Installation instructions, tutorials, docs



## Concepts in ROS

- Nodes: programs that communicate with each other
- Messages: data structure (e.g., "Image")
- Topics: typed message channels to which nodes can publish/subscribe (e.g., "/camera1/image_color")
- Parameters: stored in a blackboard



## Software Management

- Package: atomic unit of building, contains one or more nodes and/or message definitions
- Build system "catkin" on top of CMake to ease building packages and ROS specific code, and to resolve dependencies between packages
- Repository: contains several packages, typically one repository per institution


## Useful Tools

- catkin_init_workspace
- catkin_create_pkg
- catkin_make
- roscore
- rosnode list/info
- rostopic list/echo
- rosbag record/play
- rosrun


## Tutorials in ROS



## ROS/ Tutorials

## 1. Beginner Level

1. Installing and Configuring Your ROS Environment

This tutorials walks you through installing ROS and setting up your ROS environment on your computer.
2. Navigating the ROS Filesystem

This tutorial introduces ROS filesystem concepts, and covers using the roscd, rosls, and rospack command-line tools.
3. Creating a ROS Package

This tutorial covers using roscreate-pkg to create a new package, and rospack to list package dependencies.
4. Building a ROS Package

This tutorial covers using rosmake to build a package, and rosdep to install system dependencies.
5. Understanding ROS Nodes

This tutorial introduces ROS graph concepts and discusses the use of roscore, rosnode, and rosrun command-line tools.
6. Understanding ROS Topics

This tutorial introduces ROS topics as well as using the rostopic and rxplot command-line tools.
7. Understanding ROS Services and Parameters

This tutorial introduces ROS services, and parameters as well as using the rosservice and rosparam command-line tools.
8. Using rxconsole and roslaunch

This tutorial introduces ROS using rxconsole and rxloggerlevel for debugging and roslaunch for starting

## Questions?

