ME336 Collaborative Robot Learning Spring 2019 Wednesday, March 22

Lecture 05
Images & Cameras

Song Chaoyang

Assistant Professor

Department of Mechanical and Energy Engineering

songcy@sustech.edu.cn

SUSTech
Southern University
of Science and Technology

A Dynamical System Approach for Softly Catching a Flying Object: Theory and Experiment

By Seyed Sina Mirrazavi Salehian; Mahdi Khoramshahi; Aude Billard (EPFL)

S. S. M. Salehian, M. Khoramshahi and A. Billard, "A Dynamical System Approach for Softly Catching a Flying Object: Theory and Experiment," in IEEE Transactions on Robotics, vol. 32, no. 2, pp. 462-471, April 2016. doi: 10.1109/TRO.2016.2536749

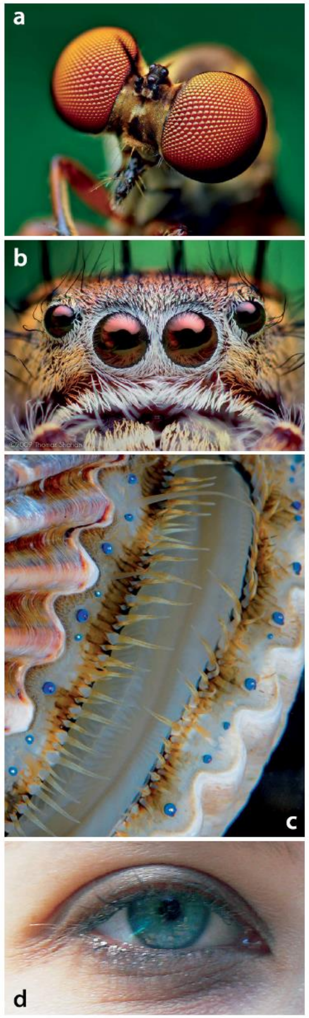
From Animals, to Computers, then Robots

The Concept of Vision

AncoraSIR.com

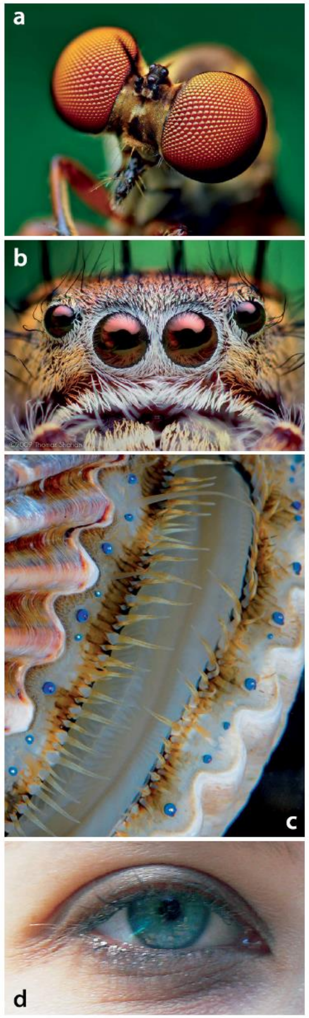
From Animals, to Computers, then Robots

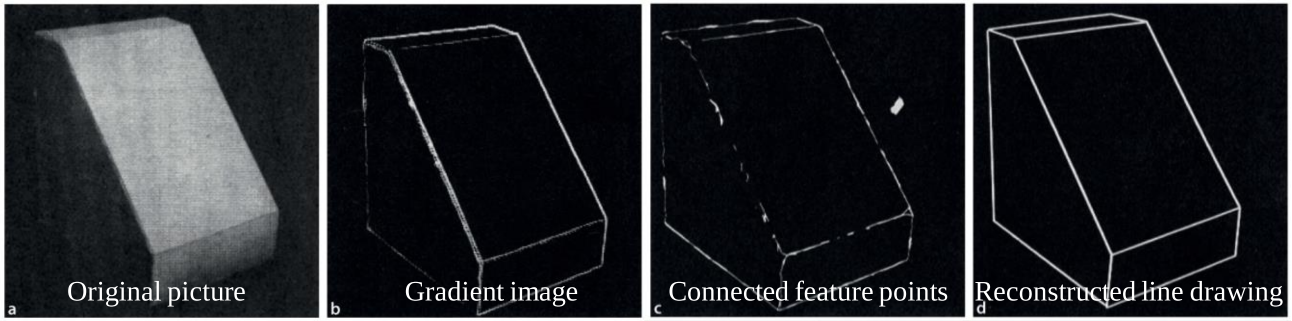
The Concept of Vision



From Animals, to Computers, then Robots

The Concept of Vision

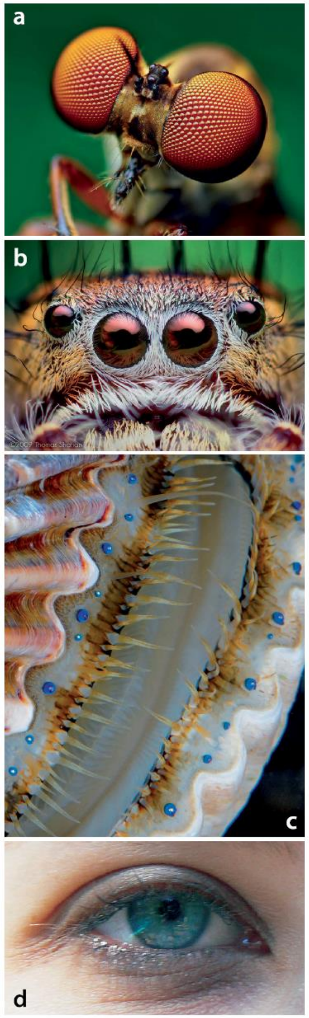


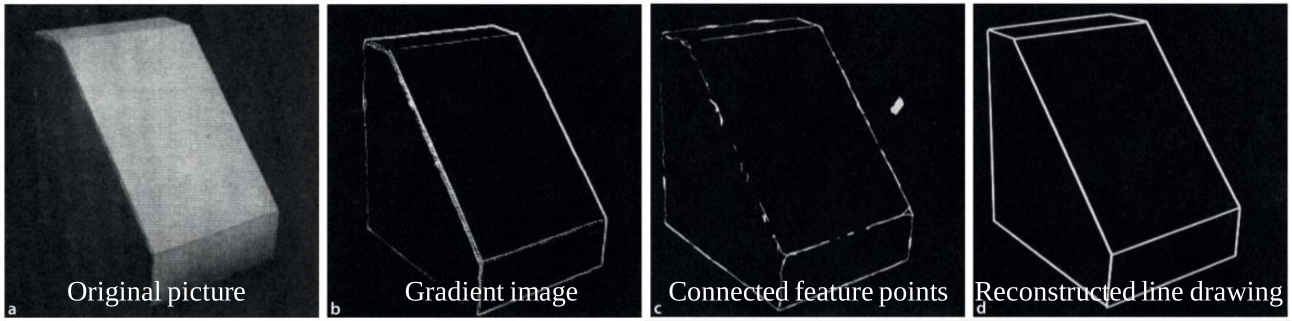


Early results in computer vision for estimating the shape and pose of objects,
from the PhD work of L. G. Roberts at MIT Lincoln Lab in 1963.

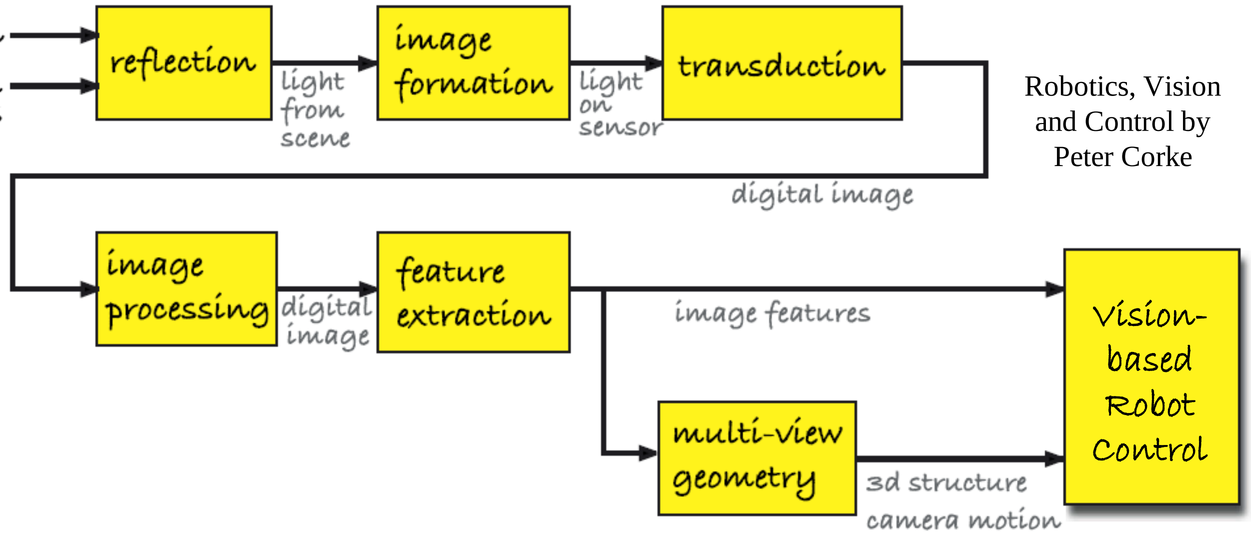
From Animals, to Computers, then Robots

The Concept of Vision





Early results in computer vision for estimating the shape and pose of objects, from the PhD work of L. G. Roberts at MIT Lincoln Lab in 1963.



Robotics, Vision and Control by Peter Corke

What Does Vision Tell Us About the World?

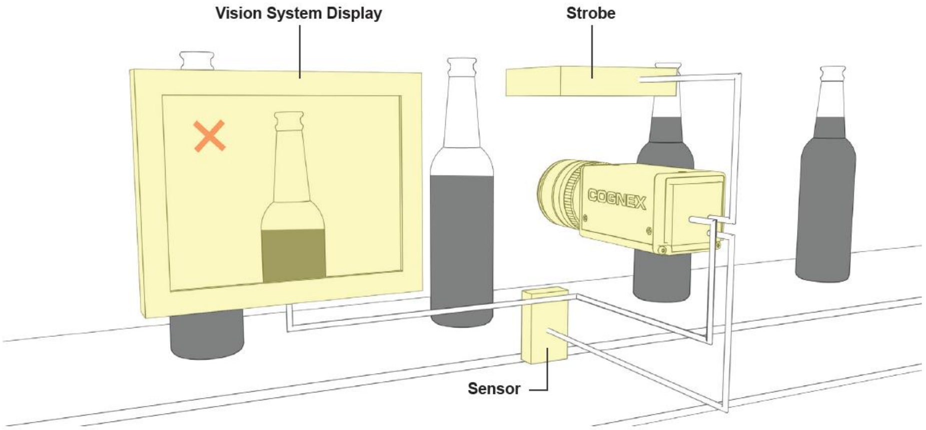
Or the Robot as a Machine

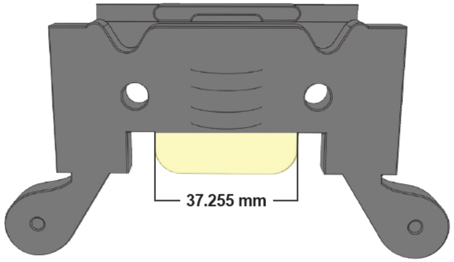
AncoraSIR.com

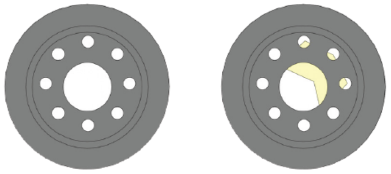
What Does Vision Tell Us About the World?

Or the Robot as a Machine

- Static features
  - Distance
  - Color
  - Shape
  - Texture
  - Environment
  - …







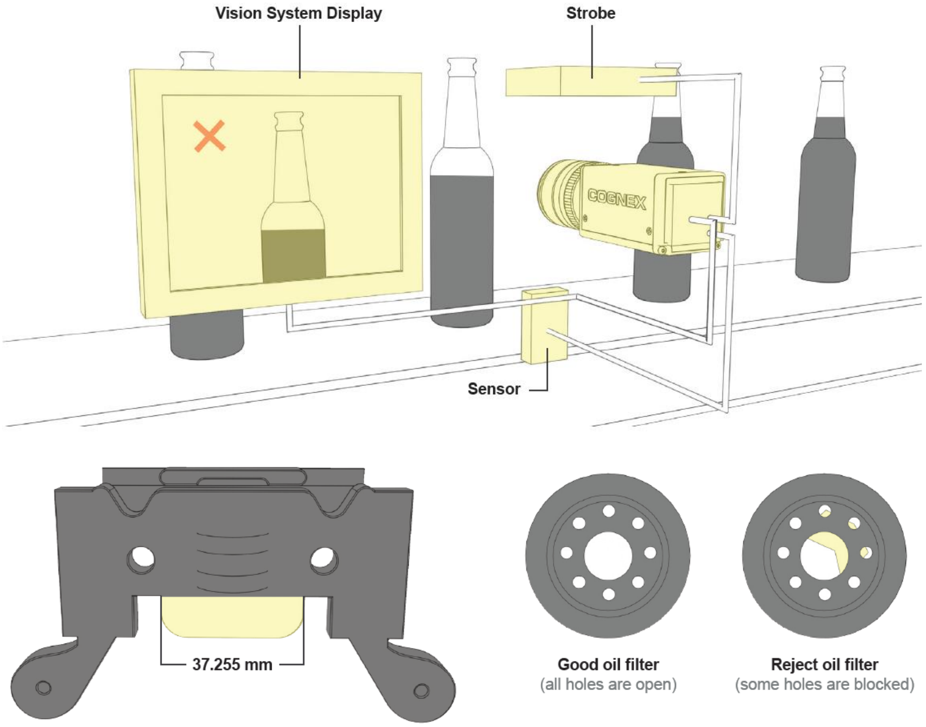
Good oil filter
(all holes are open)

Reject oil filter
(some holes are blocked)

What Does Vision Tell Us About the World?

Or the Robot as a Machine

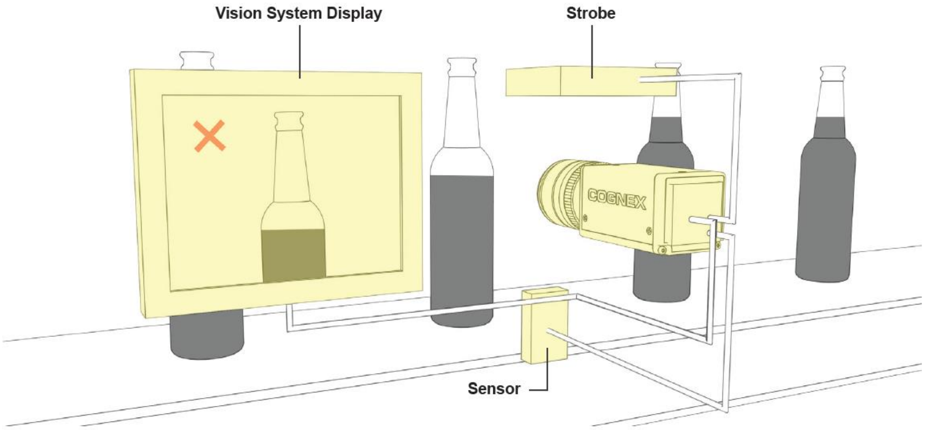
- Static features
  - Distance
  - Color
  - Shape
  - Texture
  - Environment
  - …
- Dynamic motions

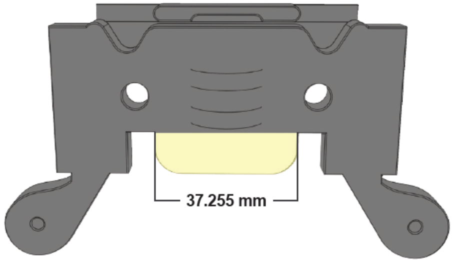


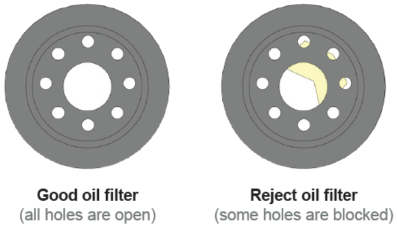
What Does Vision Tell Us About the World?

Or the Robot as a Machine

- Static features
  - Distance
  - Color
  - Shape
  - Texture
  - Environment
  - …
- Dynamic motions
- Understanding of the behavior



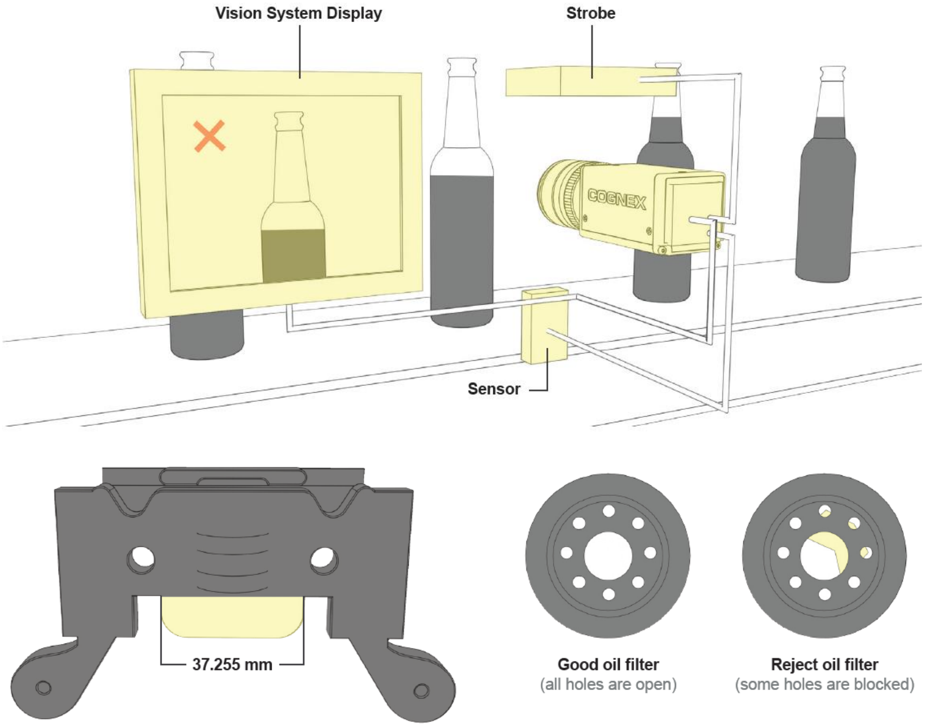




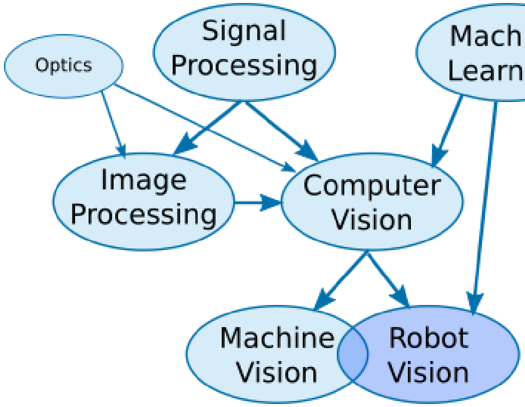
What Does Vision Tell Us About the World?

Or the Robot as a Machine

- Static features
  - Distance
  - Color
  - Shape
  - Texture
  - Environment
  - …
- Dynamic motions
- Understanding of the behavior
- *An important method to interact with the physical world*



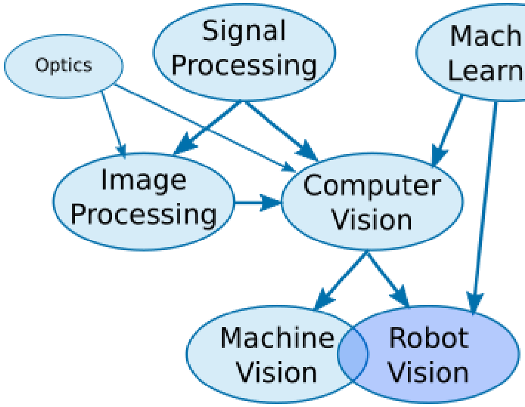
Differentiating Concepts about Vision



AncoraSIR.com

Differentiating Concepts about Vision

- ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

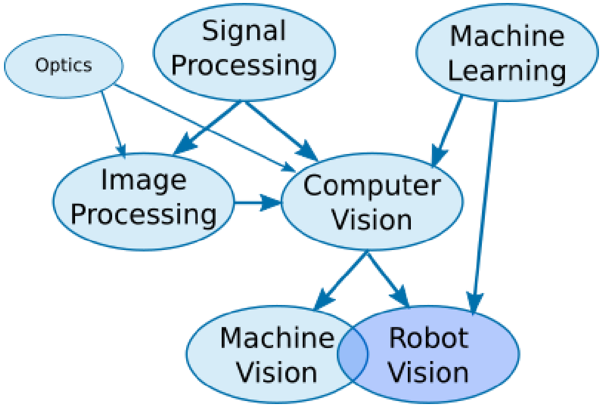


AncoraSIR.com

Differentiating Concepts about Vision

- ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

- ***Image Processing*** techniques are primarily used to improve the quality of an image, convert it into another format (like a histogram) or otherwise change it for further processing.



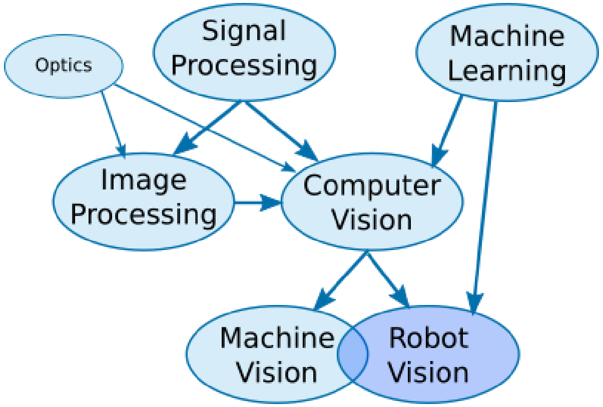
AncoraSIR.com

Differentiating Concepts about Vision

- ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

- ***Computer Vision*** is more about extracting information from images to make sense of them.

- ***Image Processing*** techniques are primarily used to improve the quality of an image, convert it into another format (like a histogram) or otherwise change it for further processing.



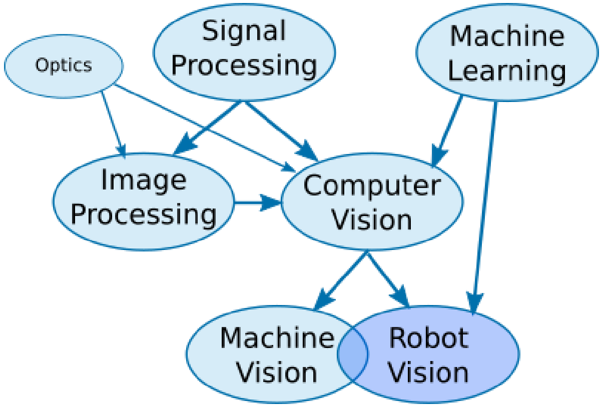
AncoraSIR.com

Differentiating Concepts about Vision

- □ ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

- □ ***Computer Vision*** is more about extracting information from images to make sense of them.

- □ ***Image Processing*** techniques are primarily used to improve the quality of an image, convert it into another format (like a histogram) or otherwise change it for further processing.



- □ ***Machine Vision*** refers to the industrial use of vision for automatic inspection, process control and robot guidance.

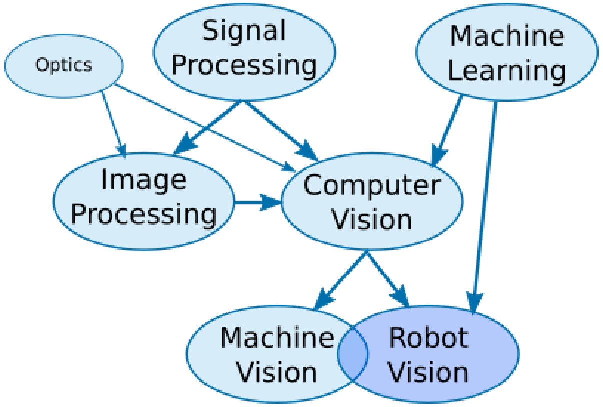
AncoraSIR.com

Differentiating Concepts about Vision

- ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

- ***Computer Vision*** is more about extracting information from images to make sense of them.

- ***Image Processing*** techniques are primarily used to improve the quality of an image, convert it into another format (like a histogram) or otherwise change it for further processing.



- ***Robotic Vision*** involves using a combination of camera hardware and computer algorithms to allow robots to process visual data from the world and execute physical actions.

- ***Machine Vision*** refers to the industrial use of vision for automatic inspection, process control and robot guidance.

AncoraSIR.com

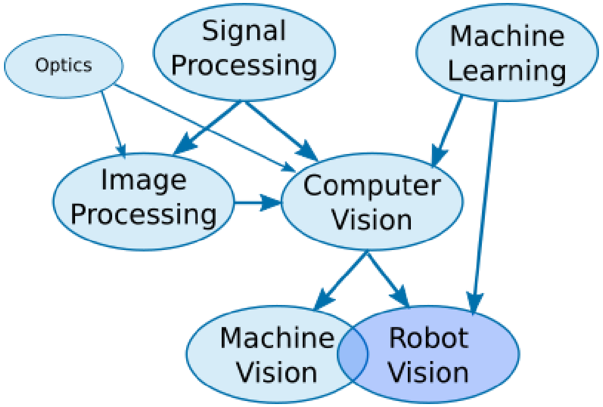
SUSTech Southern University of Science and Technology

Differentiating Concepts about Vision

- □ ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

- □ ***Computer Vision*** is more about extracting information from images to make sense of them.

- □ ***Image Processing*** techniques are primarily used to improve the quality of an image, convert it into another format (like a histogram) or otherwise change it for further processing.



- □ ***Machine Learning*** is focused on recognizing patterns in data.

- □ ***Robotic Vision*** involves using a combination of camera hardware and computer algorithms to allow robots to process visual data from the world and execute physical actions.

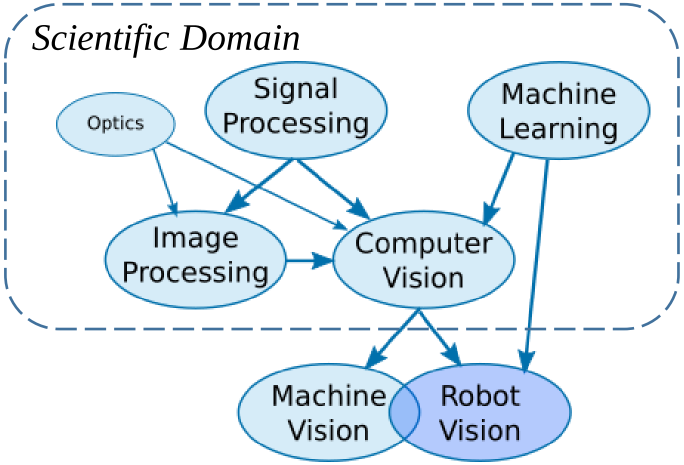
- □ ***Machine Vision*** refers to the industrial use of vision for automatic inspection, process control and robot guidance.

AncoraSIR.com

Differentiating Concepts about Vision

- ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

- ***Computer Vision*** is more about extracting information from images to make sense of them.



- ***Image Processing*** techniques are primarily used to improve the quality of an image, convert it into another format (like a histogram) or otherwise change it for further processing.

- ***Machine Learning*** is focused on recognizing patterns in data.

- ***Robotic Vision*** involves using a combination of camera hardware and computer algorithms to allow robots to process visual data from the world and execute physical actions.

- ***Machine Vision*** refers to the industrial use of vision for automatic inspection, process control and robot guidance.

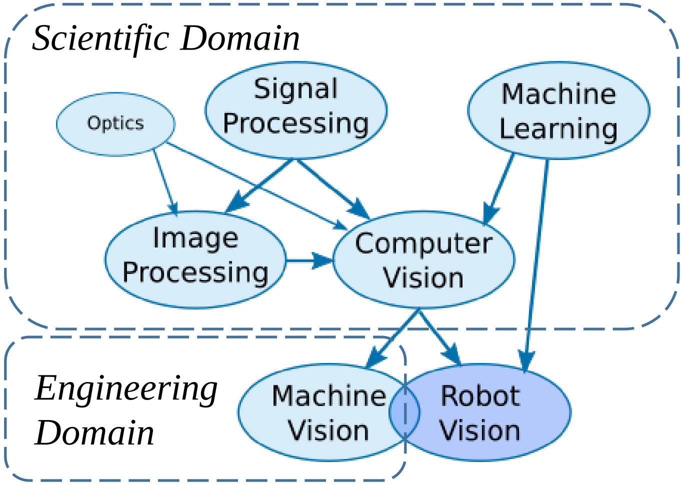
AncoraSIR.com

SUSTech
Southern University of Science and Technology

Differentiating Concepts about Vision

□ ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

□ ***Computer Vision*** is more about extracting information from images to make sense of them.



□ ***Image Processing*** techniques are primarily used to improve the quality of an image, convert it into another format (like a histogram) or otherwise change it for further processing.

□ ***Machine Learning*** is focused on recognizing patterns in data.

□ ***Robotic Vision*** involves using a combination of camera hardware and computer algorithms to allow robots to process visual data from the world and execute physical actions.

□ ***Machine Vision*** refers to the industrial use of vision for automatic inspection, process control and robot guidance.

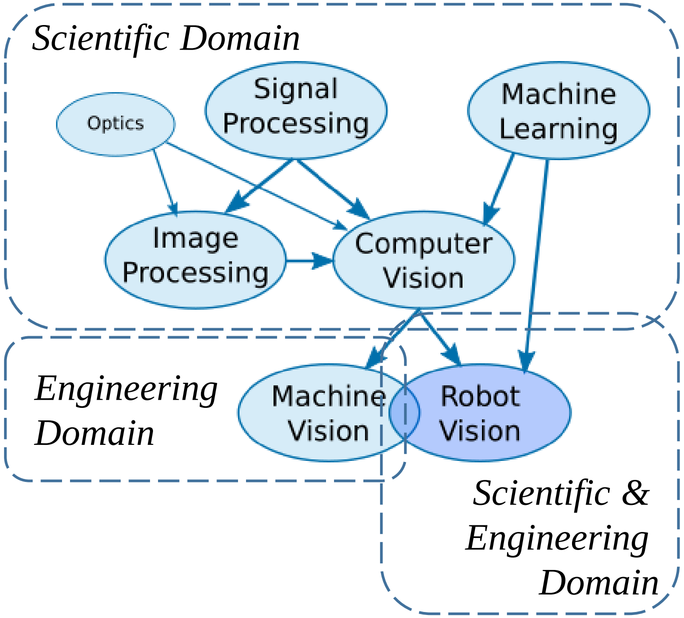
AncoraSIR.com

Differentiating Concepts about Vision

- ***Signal Processing*** involves processing electronic signals to either clean them up, extract information, prepare them to output to a display or prepare them for further processing. Anything can be a signal, more or less.

- ***Computer Vision*** is more about extracting information from images to make sense of them.

- ***Image Processing*** techniques are primarily used to improve the quality of an image, convert it into another format (like a histogram) or otherwise change it for further processing.



- ***Machine Learning*** is focused on recognizing patterns in data.

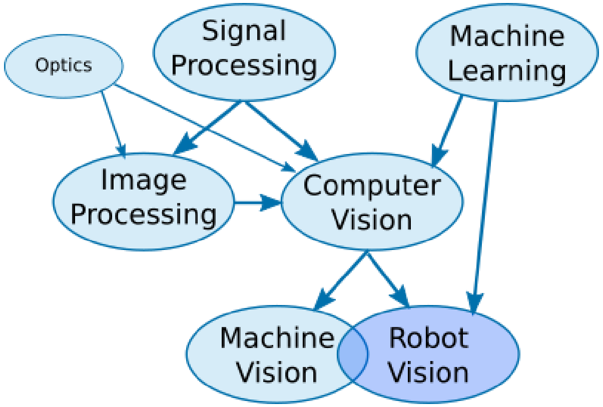
- ***Robotic Vision*** involves using a combination of camera hardware and computer algorithms to allow robots to process visual data from the world and execute physical actions.

- ***Machine Vision*** refers to the industrial use of vision for automatic inspection, process control and robot guidance.

AncoraSIR.com

Differentiating Concepts about Vision

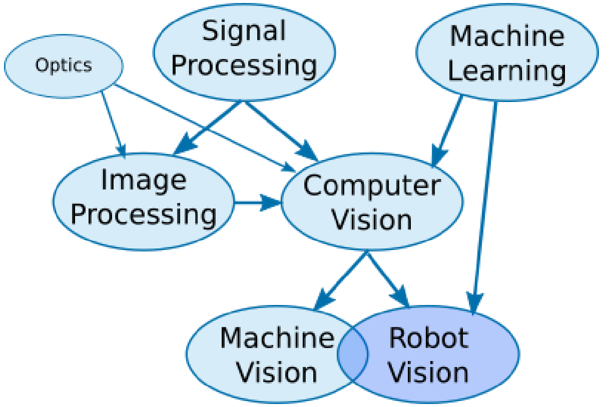
Technique	Input	Output
Signal Processing	Electrical signals	Electrical signals
Image Processing	Images	Images
Computer Vision	Images	Information/features
Pattern Recognition/Machine Learning	Information/features	Information
Machine Vision	Images	Information
Robot Vision	**Images**	**Physical Action**



AncoraSIR.com

Differentiating Concepts about Vision

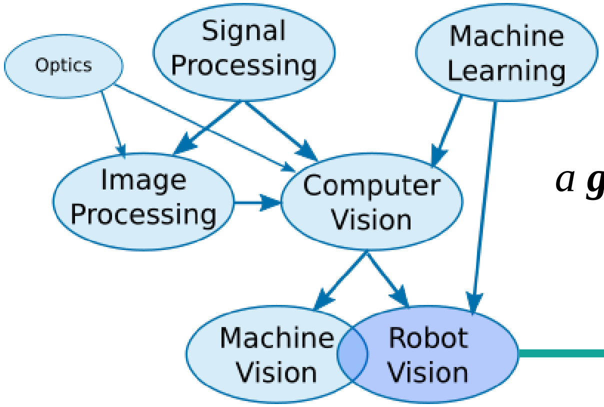
Technique	Input	Output
Signal Processing	Electrical signals	Electrical signals
Image Processing	Images	Images
Computer Vision	Images	Information/features
Pattern Recognition/Machine Learning	Information/features	Information
Machine Vision	Images	Information
Robot Vision	**Images**	**Physical Action**



AncoraSIR.com

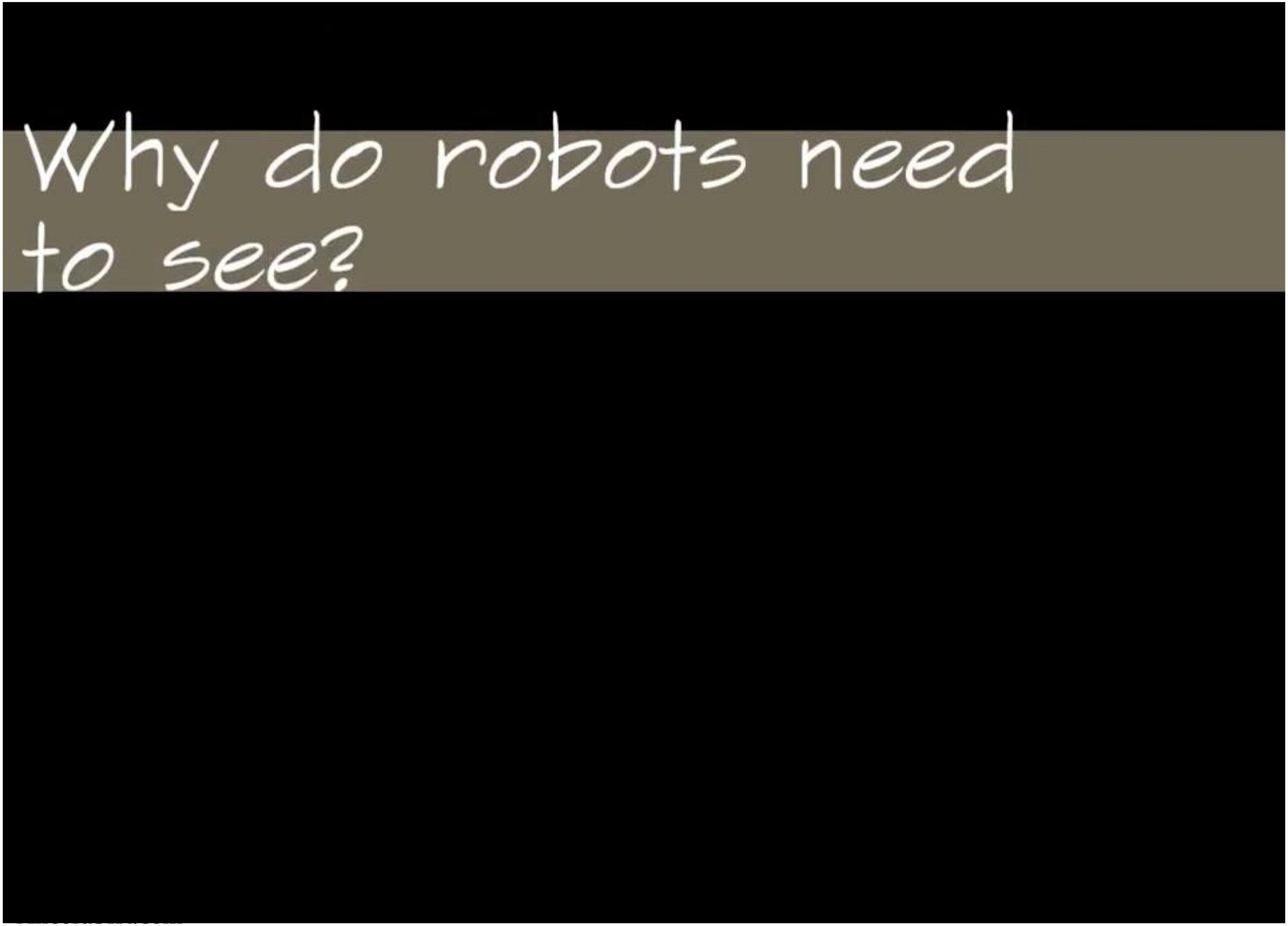
Differentiating Concepts about Vision

Technique	Input	Output
Signal Processing	Electrical signals	Electrical signals
Image Processing	Images	Images
Computer Vision	Images	Information/features
Pattern Recognition/Machine Learning	Information/features	Information
Machine Vision	Images	Information
Robot Vision	**Images**	**Physical Action**



*a **goal** oriented **machine** that can **sense**, **plan** and **act***

AncoraSIR.com



Why do robots need to see?

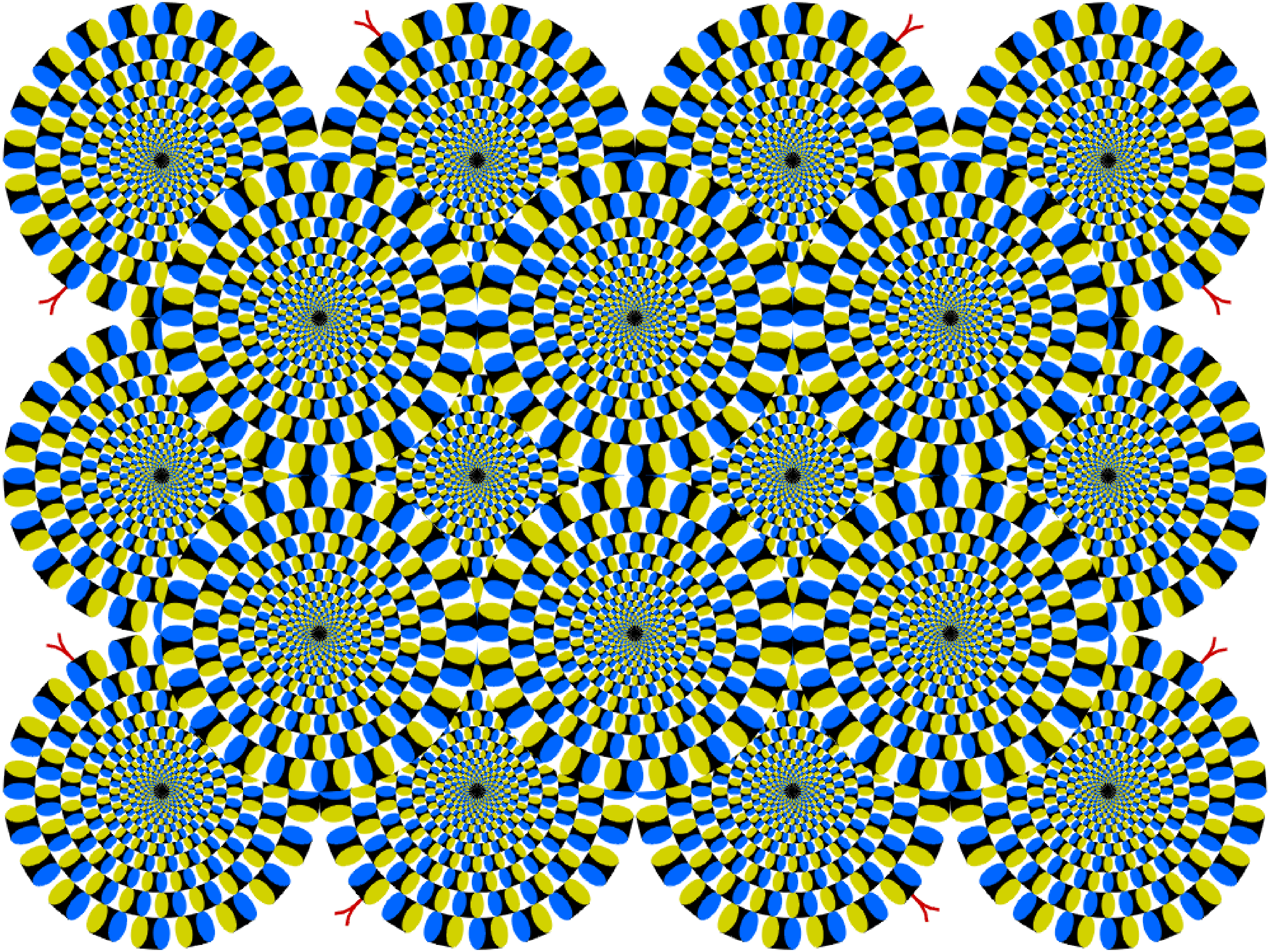


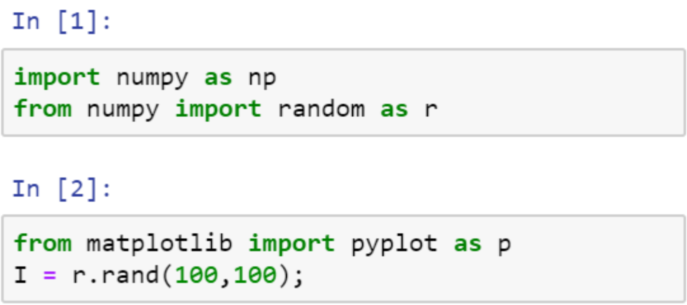
Image Representation

Example of a grayscale [0, 1] image within a planar area of size [*m*, *n*]

AncoraSIR.com

Image Representation

Example of a grayscale [0, 1] image within a planar area of size [*m*, *n*]



In [1]:

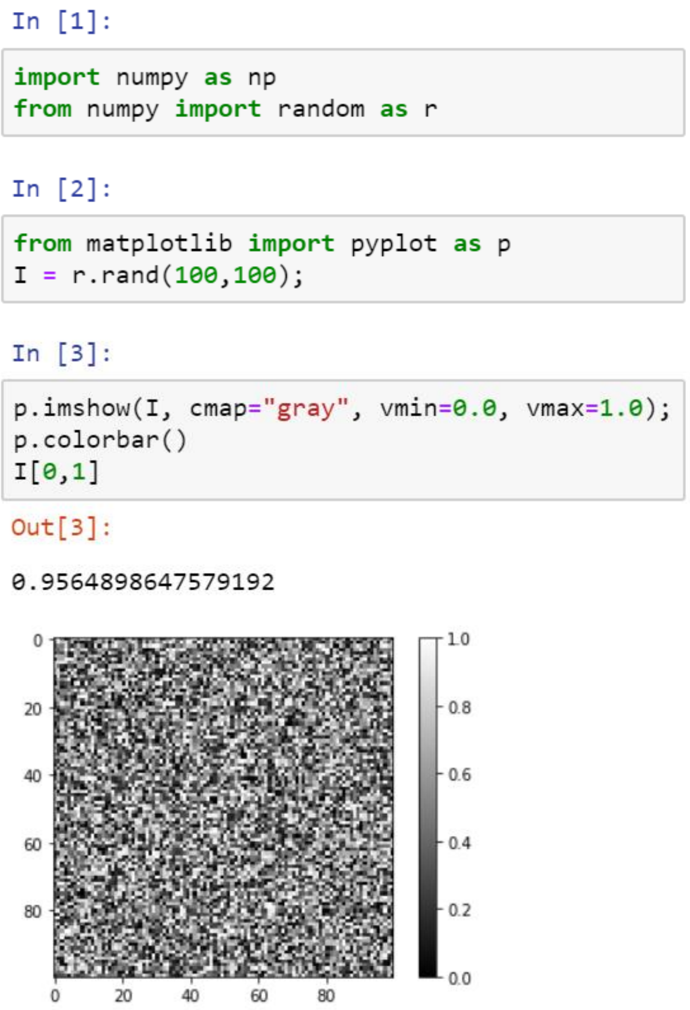
```
import numpy as np
from numpy import random as r
```

In [2]:

```
from matplotlib import pyplot as p
I = r.rand(100,100);
```

Image Representation

Example of a grayscale [0, 1] image within a planar area of size [*m*, *n*]



In [1]:

```
import numpy as np
from numpy import random as r
```

In [2]:

```
from matplotlib import pyplot as p
I = r.rand(100,100);
```

In [3]:

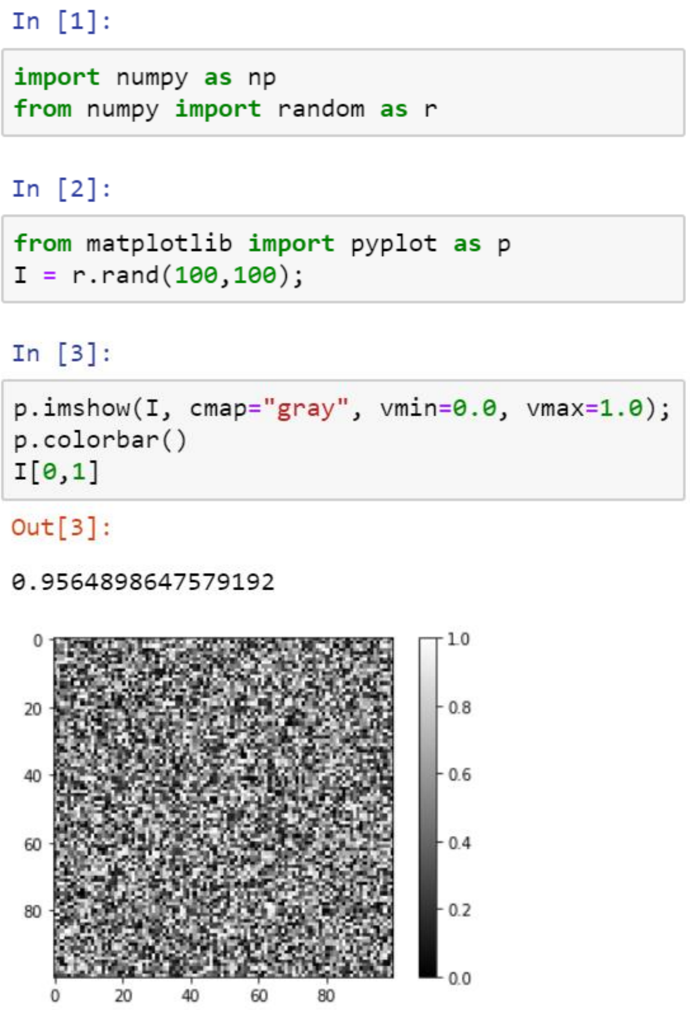
```
p.imshow(I, cmap="gray", vmin=0.0, vmax=1.0);
p.colorbar()
I[0,1]
```

Out[3]:

```
0.9564898647579192
```

Image Representation

Example of a grayscale [0, 1] image within a planar area of size [*m*, *n*]



In [1]:

```
import numpy as np
from numpy import random as r
```

In [2]:

```
from matplotlib import pyplot as p
I = r.rand(100,100);
```

In [3]:

```
p.imshow(I, cmap="gray", vmin=0.0, vmax=1.0);
p.colorbar()
I[0,1]
```

Out[3]:

```
0.9564898647579192
```

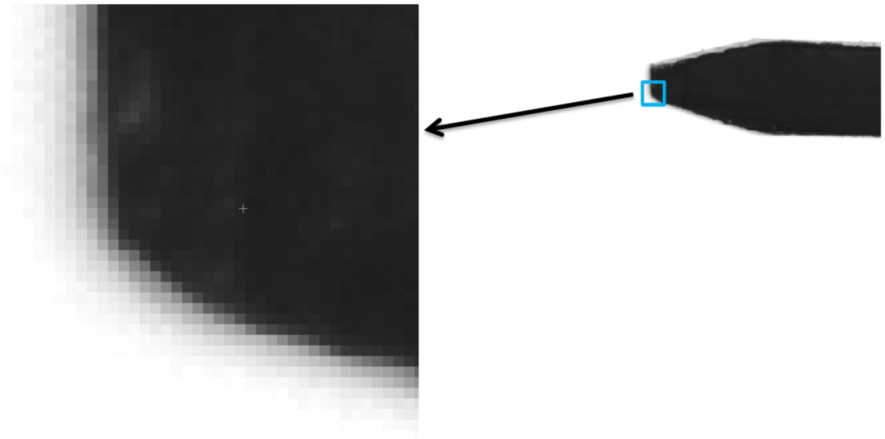
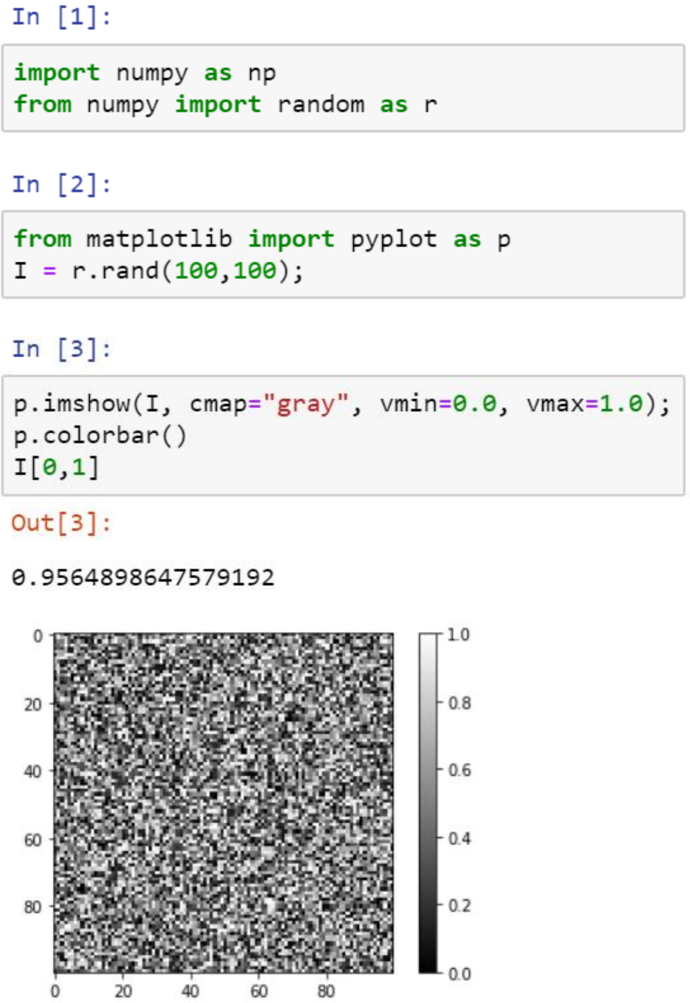
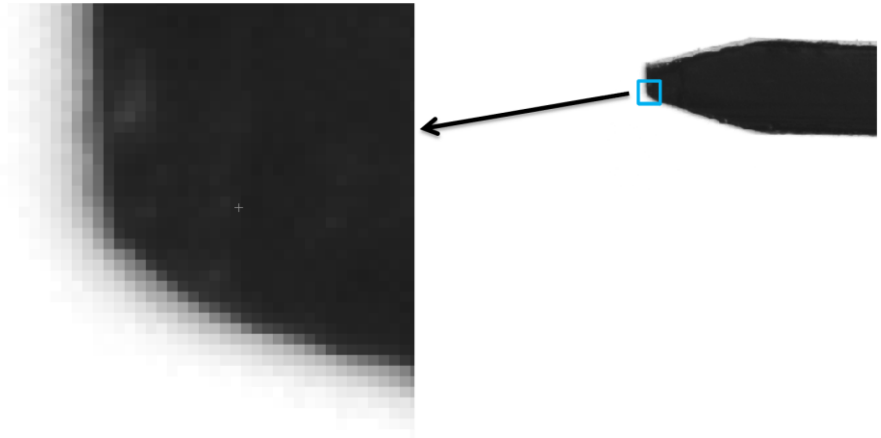


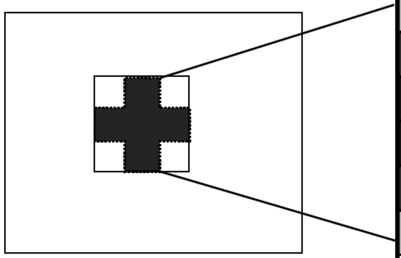
Image Representation

Example of a grayscale [0, 1] image within a planar area of size [*m*, *n*]



```
In [1]:
import numpy as np
from numpy import random as r

In [2]:
from matplotlib import pyplot as p
I = r.rand(100,100);

In [3]:
p.imshow(I, cmap="gray", vmin=0.0, vmax=1.0);
p.colorbar()
I[0,1]

Out[3]:
0.9564898647579192
```

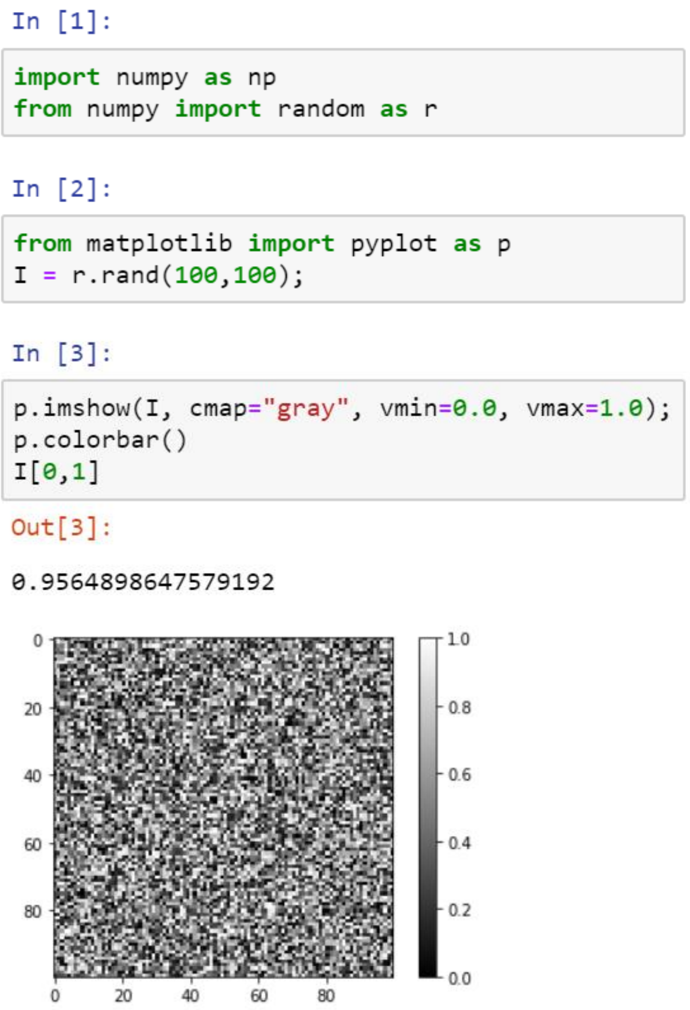




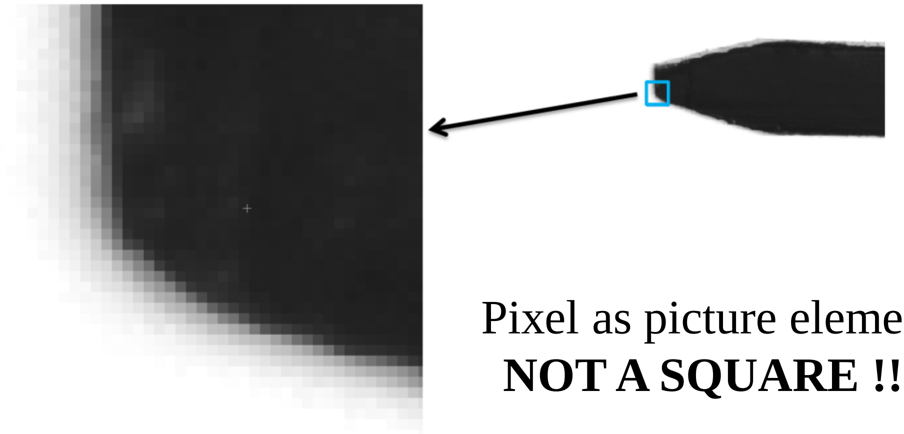
255	255	255	105	51	41	43	49	101	255	255	255
255	255	255	116	62	44	42	57	120	255	255	255
255	255	255	112	68	41	46	58	117	255	255	255
105	110	111	109	60	42	48	61	115	112	114	108
60	68	62	57	42	41	46	41	43	49	42	41
44	42	41	46	46	42	48	44	42	42	46	42
41	46	42	48	44	42	41	41	46	43	49	42
59	54	60	59	41	46	42	46	46	42	48	46
100	120	120	115	51	41	43	49	110	116	118	105
255	255	255	118	62	44	42	57	115	255	255	255
255	255	255	121	68	41	46	58	120	255	255	255
255	255	255	100	60	42	48	61	105	255	255	255

Image Representation

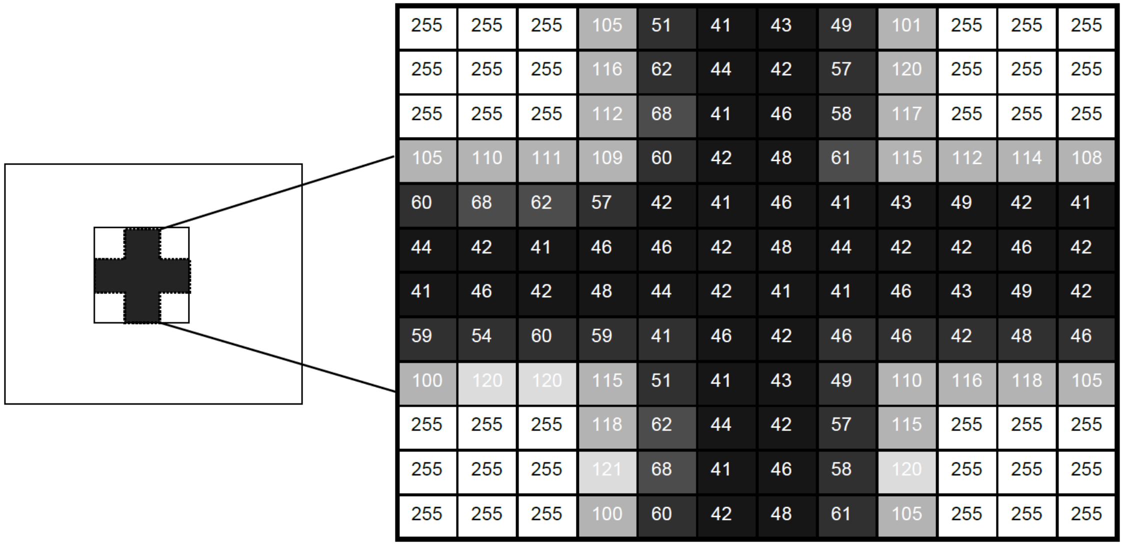
Example of a grayscale [0, 1] image within a planar area of size [*m*, *n*]



```
In [1]:
import numpy as np
from numpy import random as r

In [2]:
from matplotlib import pyplot as p
I = r.rand(100,100);

In [3]:
p.imshow(I, cmap="gray", vmin=0.0, vmax=1.0);
p.colorbar()
I[0,1]

Out[3]:
0.9564898647579192
```



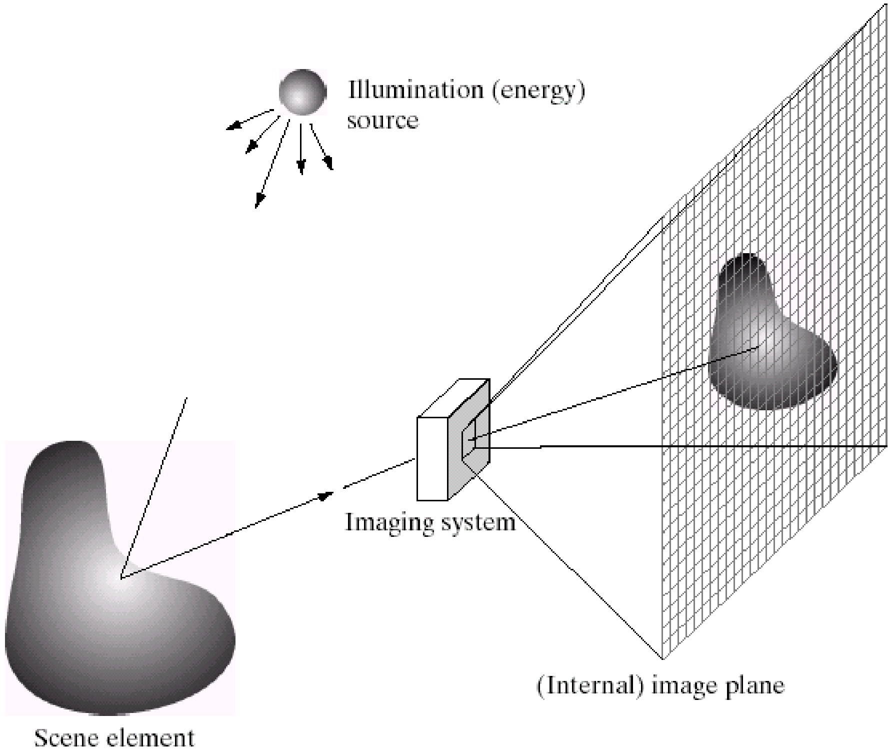
Pixel as picture element
NOT A SQUARE !!!



255	255	255	105	51	41	43	49	101	255	255	255
255	255	255	116	62	44	42	57	120	255	255	255
255	255	255	112	68	41	46	58	117	255	255	255
105	110	111	109	60	42	48	61	115	112	114	108
60	68	62	57	42	41	46	41	43	49	42	41
44	42	41	46	46	42	48	44	42	42	46	42
41	46	42	48	44	42	41	41	46	43	49	42
59	54	60	59	41	46	42	46	46	42	48	46
100	120	120	115	51	41	43	49	110	116	118	105
255	255	255	118	62	44	42	57	115	255	255	255
255	255	255	121	68	41	46	58	120	255	255	255
255	255	255	100	60	42	48	61	105	255	255	255

Formulation of An Image

As a 2D sampling of signal

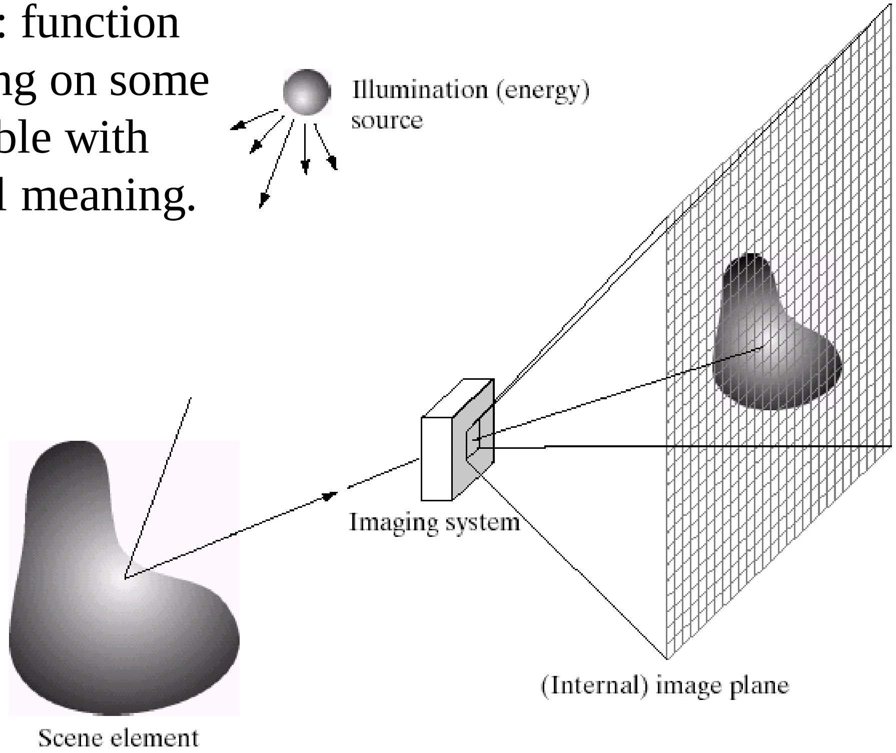


AncoraSIR.com

Formulation of An Image

As a 2D sampling of signal

Signal: function depending on some variable with physical meaning.



AncoraSIR.com

Formulation of An Image

As a 2D sampling of signal

Signal: function depending on some variable with physical meaning.

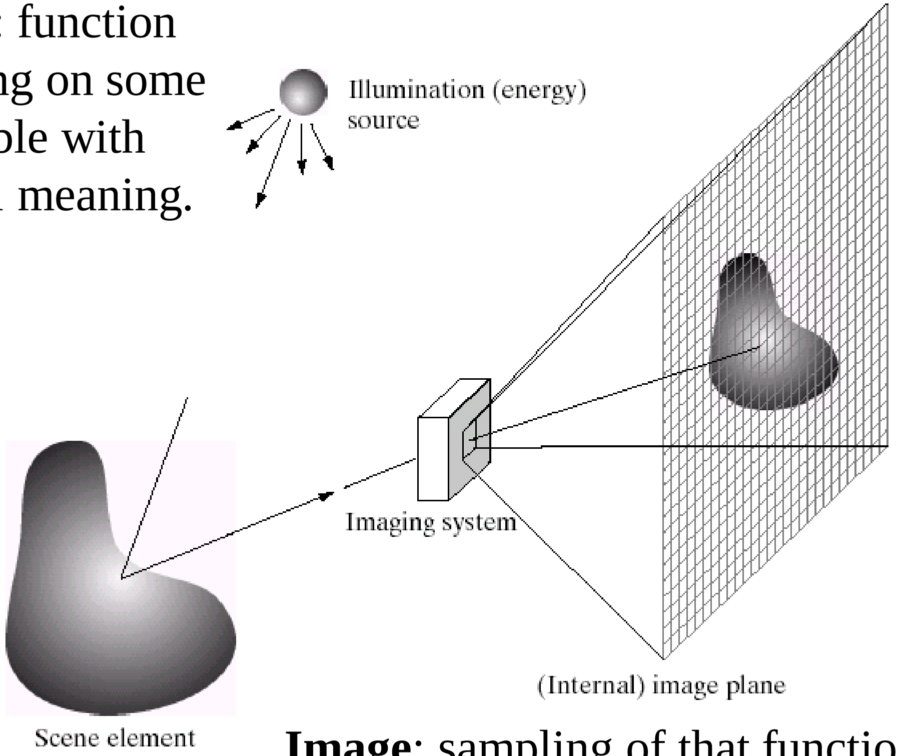


Image: sampling of that function.

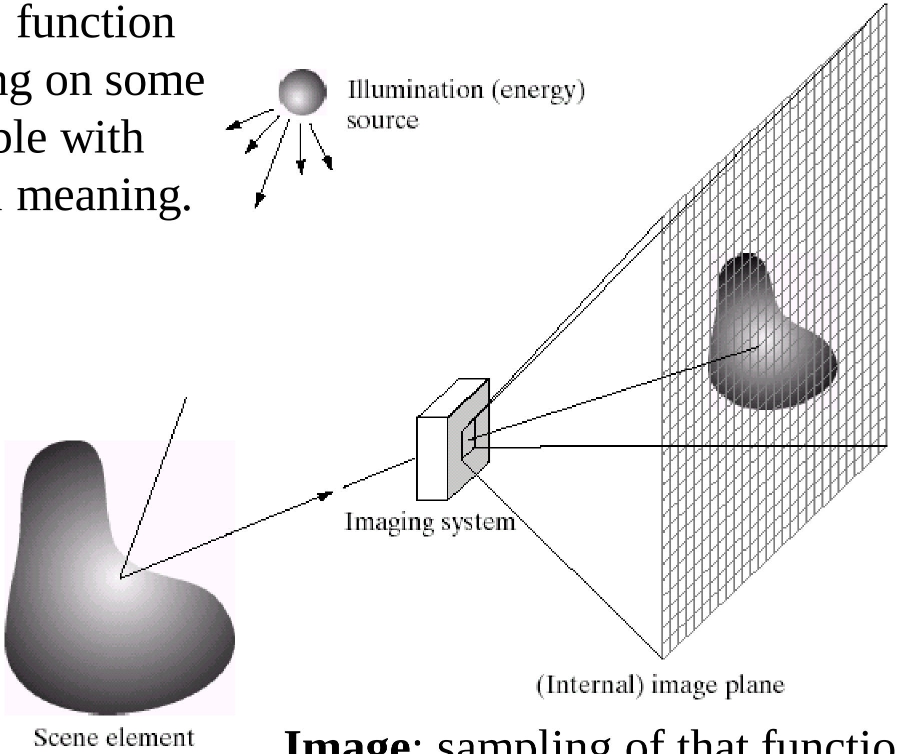
- 2 variables: xy coordinates
- 3 variables: xy + time (video)
- ‘Brightness’ is the value of the function for visible light

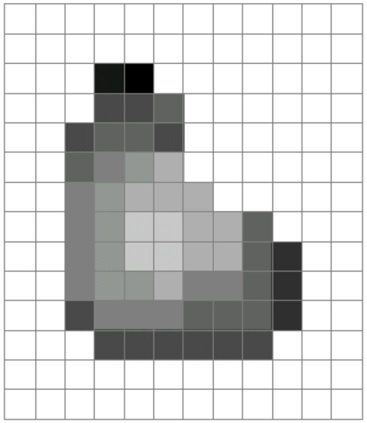
AncoraSIR.com

Formulation of An Image

As a 2D sampling of signal

Signal: function depending on some variable with physical meaning.





Can be other physical values too: *temperature*, *pressure*, *depth* …

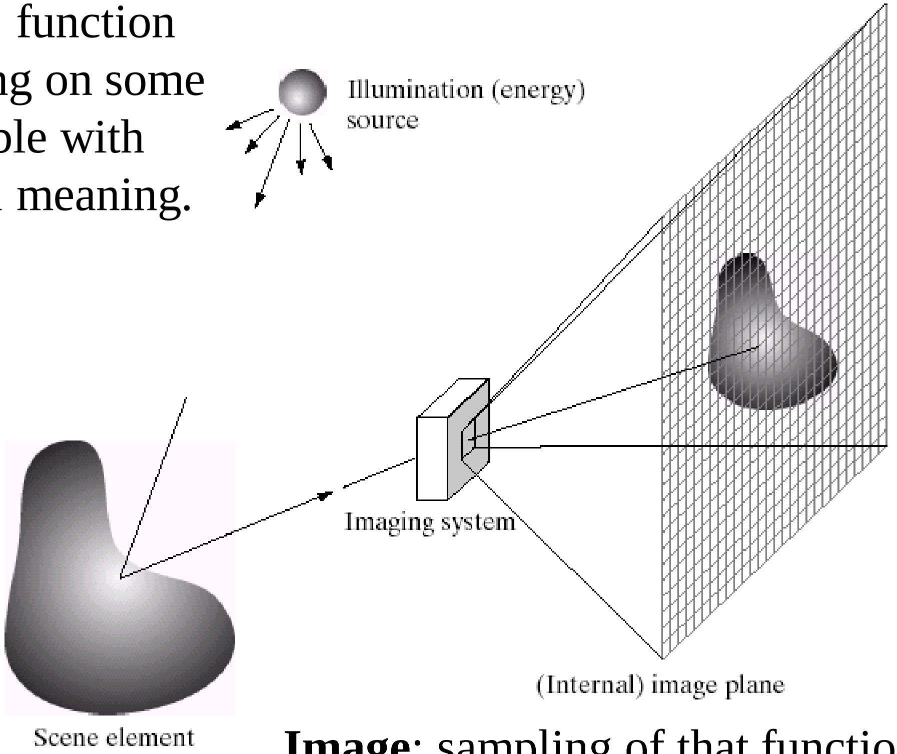
Image: sampling of that function.

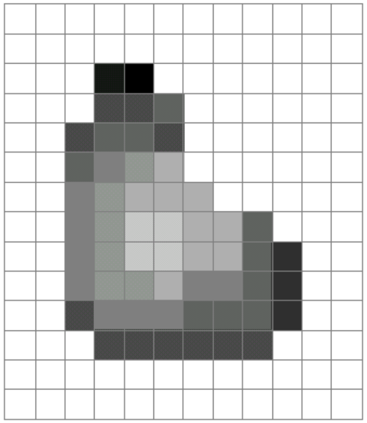
- 2 variables: xy coordinates
- 3 variables: xy + time (video)
- ‘Brightness’ is the value of the function for visible light

Formulation of An Image

As a 2D sampling of signal

Signal: function depending on some variable with physical meaning.

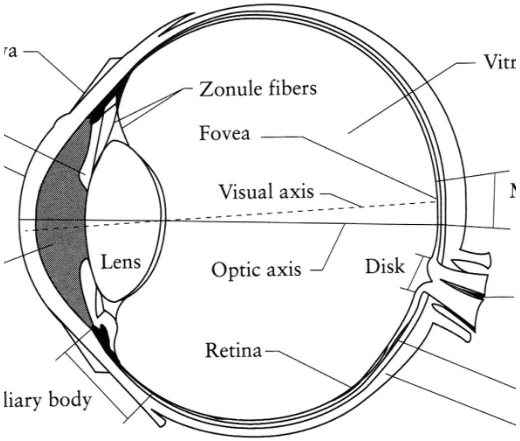




Can be other physical values too: *temperature*, *pressure*, *depth* …

Image: sampling of that function.

- 2 variables: xy coordinates
- 3 variables: xy + time (video)
- ‘Brightness’ is the value of the function for visible light



AncoraSIR.com

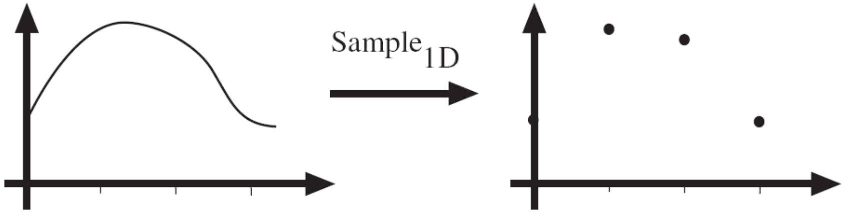
Sampling Physical World Using Images

Physical Understanding of Images

AncoraSIR.com

Sampling Physical World Using Images

Physical Understanding of Images

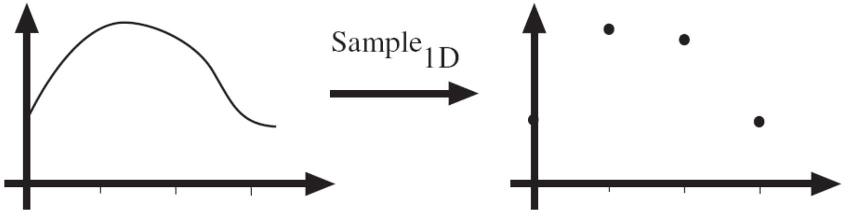


Sampling in 1D takes a function, and returns a vector whose elements are values of that function at the sample points.

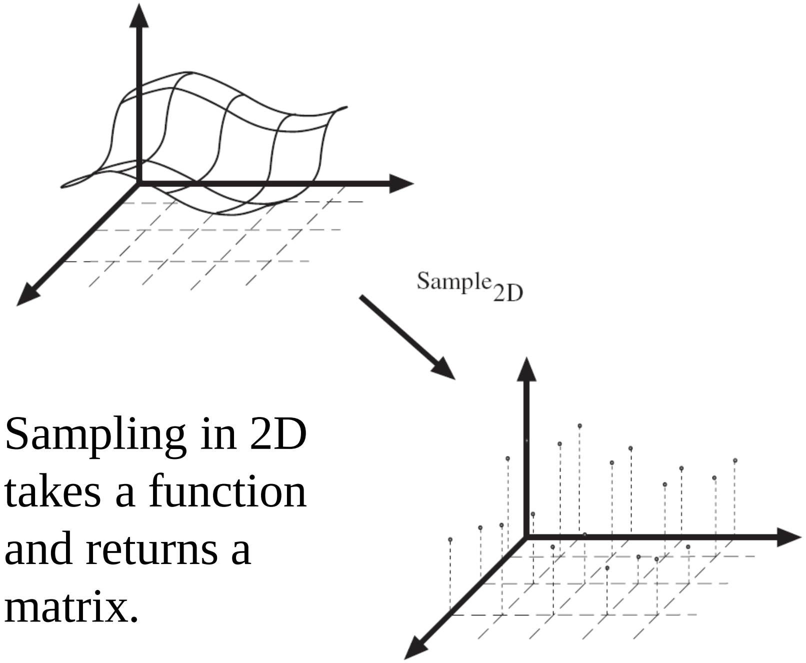
AncoraSIR.com

Sampling Physical World Using Images

Physical Understanding of Images



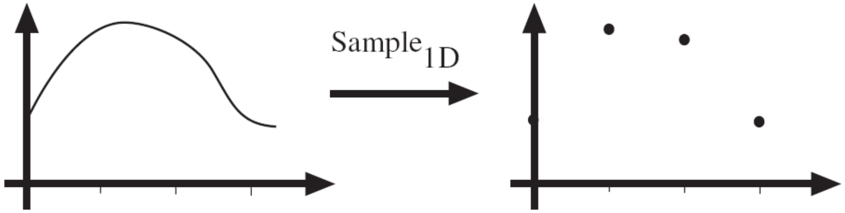
Sampling in 1D takes a function, and returns a vector whose elements are values of that function at the sample points.



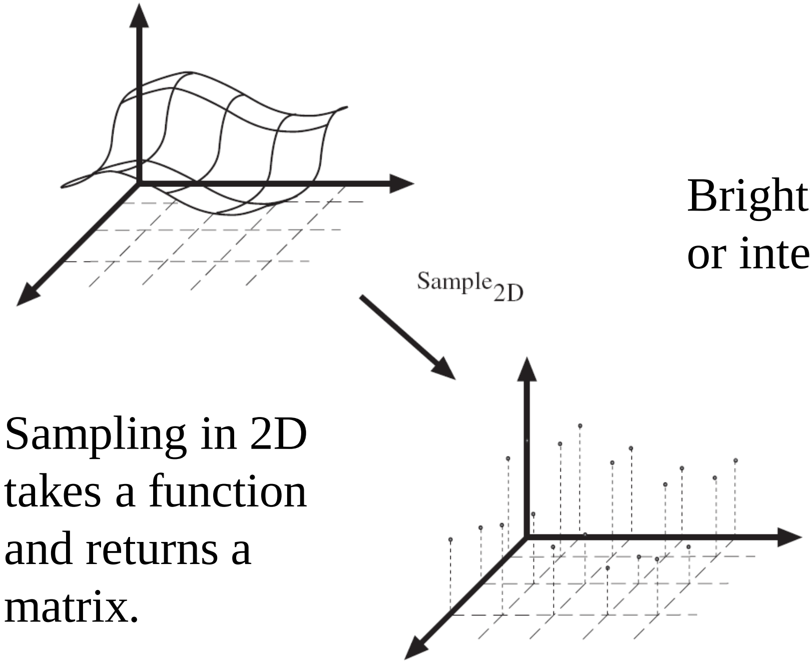
Sampling in 2D takes a function and returns a matrix.

Sampling Physical World Using Images

Physical Understanding of Images

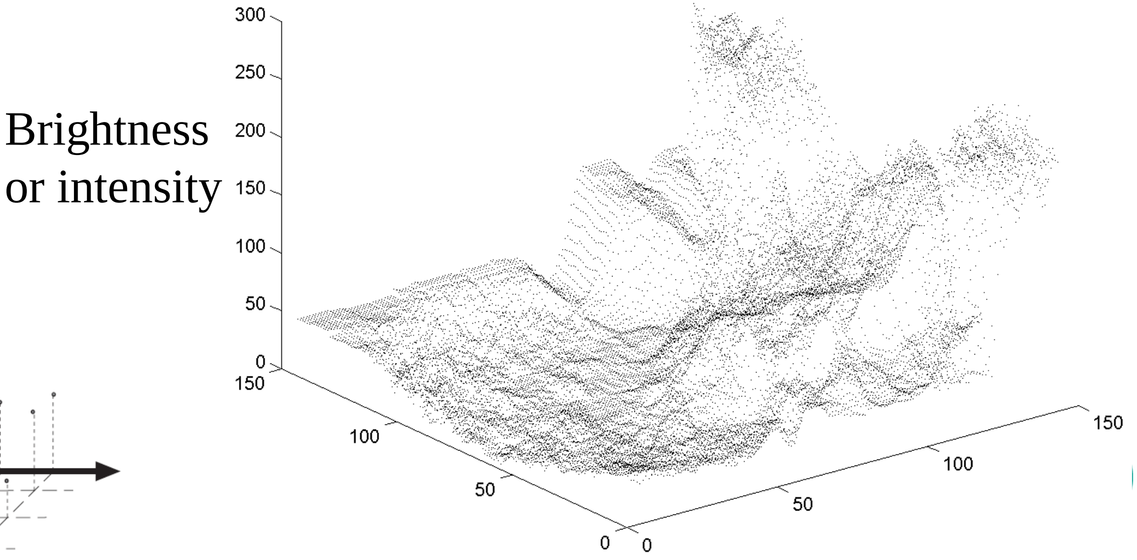


Sampling in 1D takes a function, and returns a vector whose elements are values of that function at the sample points.

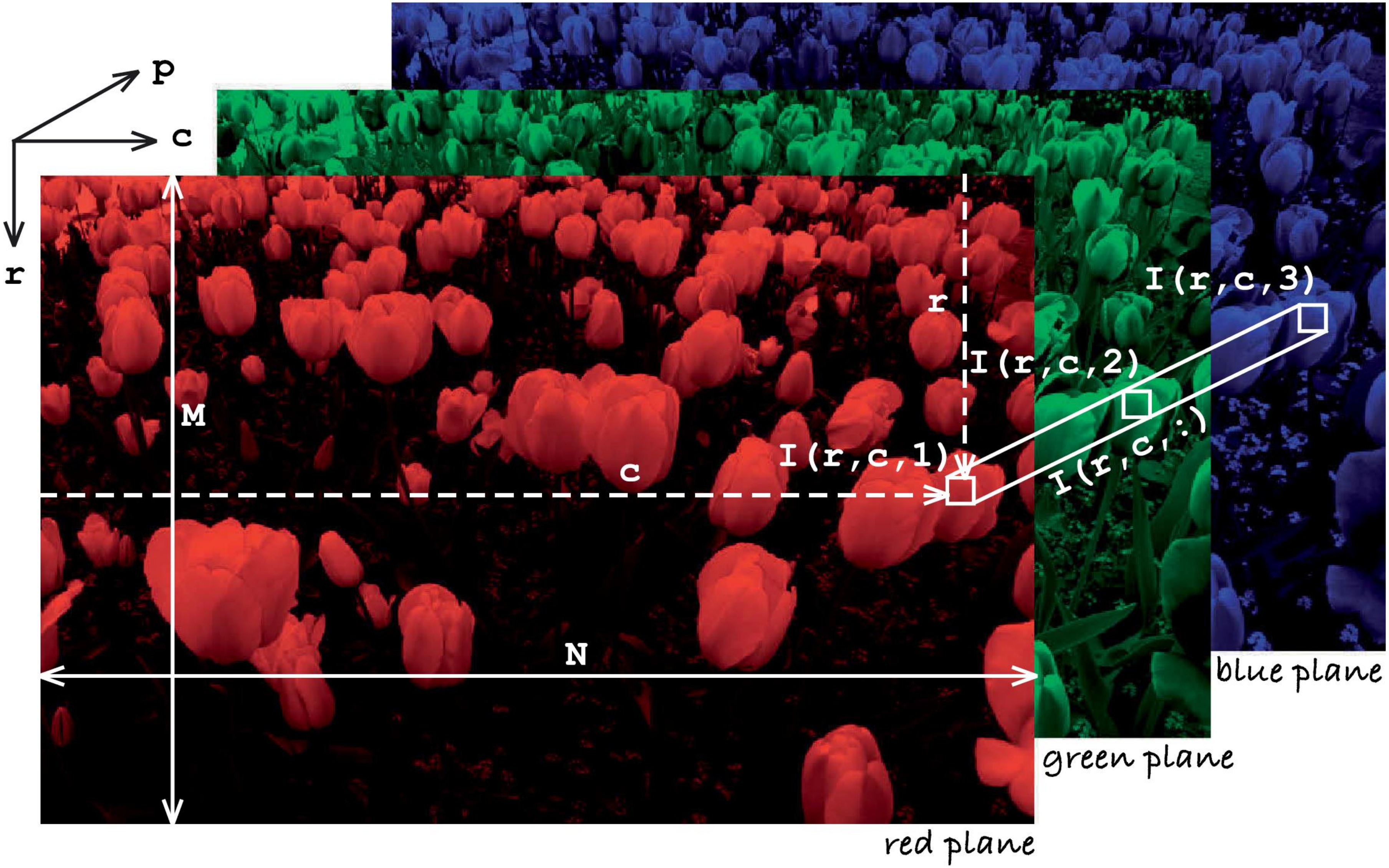


Sampling in 2D takes a function and returns a matrix.

Grayscale Digital Image



An RGB Image



A Robotic Way of Interpreting Images

An important method of sensing the environment

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*

AncoraSIR.com

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed* ***by a robot*** *and people might get hurt*

AncoraSIR.com

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed **by a robot** and people might get hurt*

Machine Vision *actually measures but no action required.*

AncoraSIR.com

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed* ***by a robot*** *and people might get hurt*

Machine Vision *actually measures but no action required.*

Image Size: u, v

AncoraSIR.com

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed* ***by a robot*** *and people might get hurt*

Machine Vision *actually measures but no action required.*

Image Size: *u*, *v*

Color Space: **R**ed, **G**reen, **B**lue
Grayscale: **G**ray

AncoraSIR.com

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed **by a robot** and people might get hurt*

***Machine Vision** actually measures but no action required.*

Image Size: *u*, *v*

Color Space: **R**ed, **G**reen, **B**lue
Grayscale: **G**ray

***[0, 1]** as normalized form, not an integer*
***[0, 255]** as a byte number of range 2^8=256 from 0 to 255,*
all in integer forms

AncoraSIR.com

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed* ***by a robot*** *and people might get hurt*

Machine Vision *actually measures but no action required.*

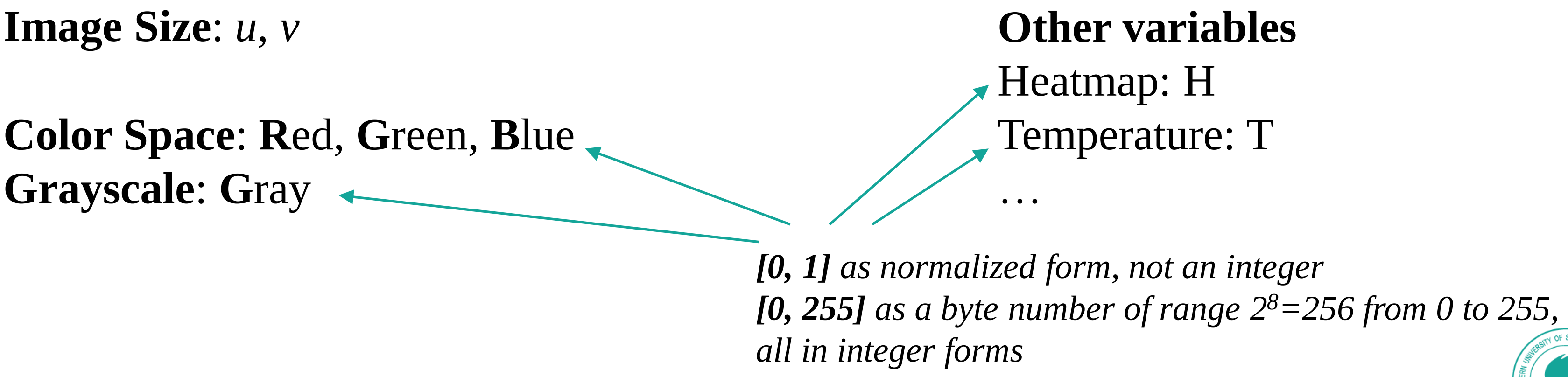


Image Size: *u*, *v*

Color Space: **R**ed, **G**reen, **B**lue
Grayscale: **G**ray

Other variables
Heatmap: H
Temperature: T
…

[0, 1] *as normalized form, not an integer*
[0, 255] *as a byte number of range 2^8=256 from 0 to 255, all in integer forms*

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed* ***by a robot*** *and people might get hurt*

Machine Vision *actually measures but no action required.*

Image Size: *u*, *v*

Color Space: **R**ed, **G**reen, **B**lue

Grayscale: **Gray**

Point Cloud: $x(u, v)$, $y(u, v)$, $z(u, v)$

Other variables

Heatmap: H

Temperature: T

…

[0, 1] *as normalized form, not an integer*

[0, 255] *as a byte number of range* 2^8*=256 from 0 to 255, all in integer forms*

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed* ***by a robot*** *and people might get hurt*

Machine Vision *actually measures but no action required.*

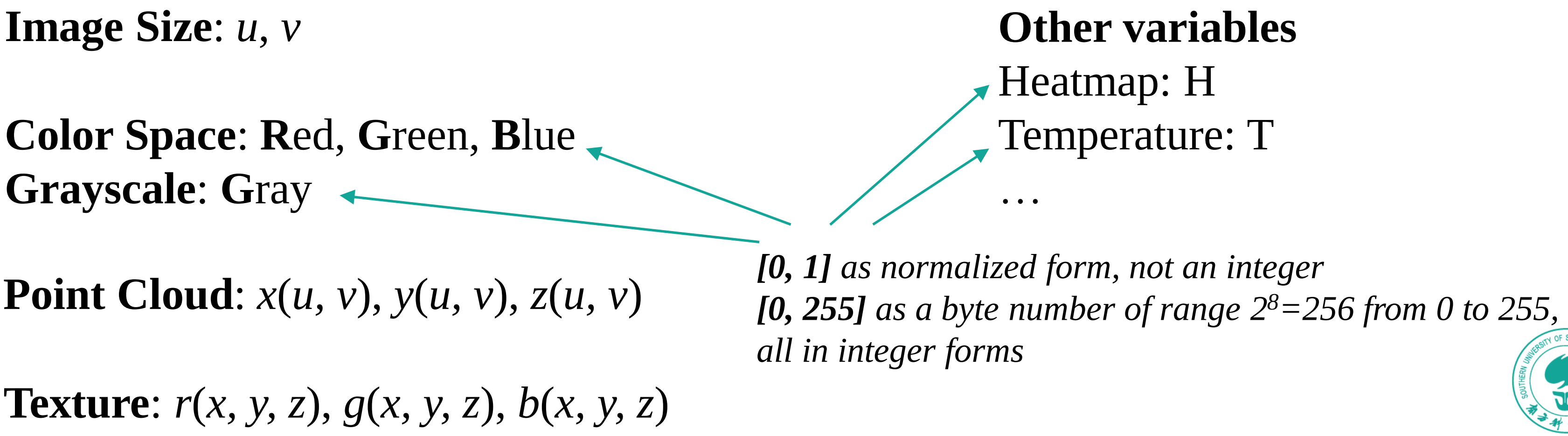


Image Size: *u*, *v*

Color Space: **R**ed, **G**reen, **B**lue
Grayscale: **Gray**

Point Cloud: $x(u, v)$, $y(u, v)$, $z(u, v)$

Texture: $r(x, y, z)$, $g(x, y, z)$, $b(x, y, z)$

Other variables
Heatmap: H
Temperature: T
…

[0, 1] *as normalized form, not an integer*
[0, 255] *as a byte number of range* 2^8*=256 from 0 to 255, all in integer forms*

A Robotic Way of Interpreting Images

An important method of sensing the environment

- Computer Vision
  - Digitization of physical world in multi-dimensional linear algebra
  - *Physical meaning is not a required way of interpretation or usage*
- Robotic Vision
  - Same as Computer Vision, but with a focus on Physical Interpretation
  - *Because actions need to be executed **by a robot** and people might get hurt*

Machine Vision *actually measures but no action required.*

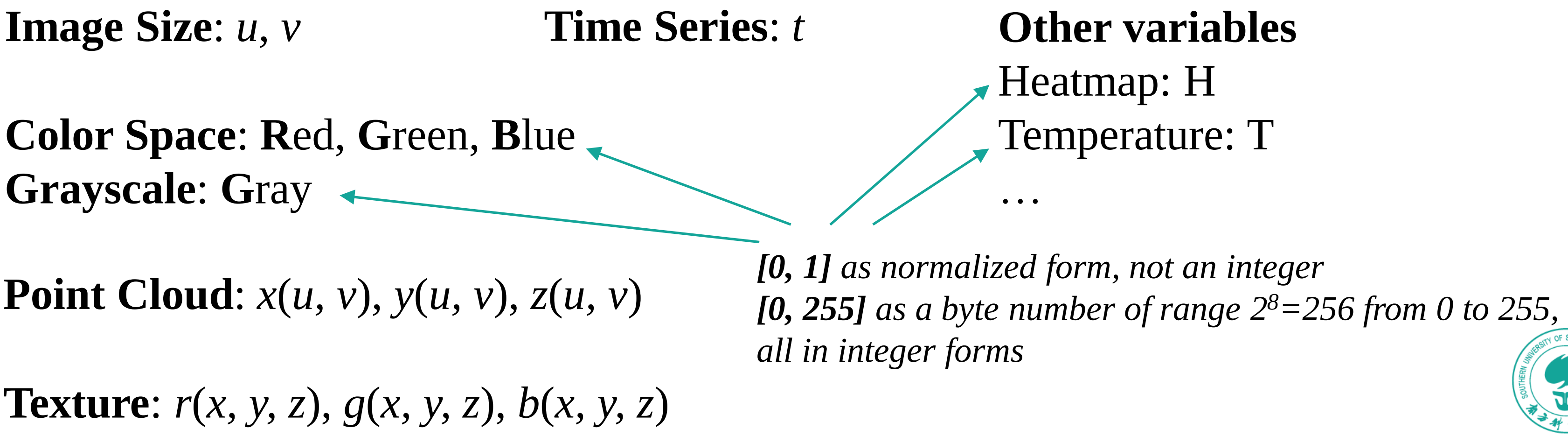


Image Size: *u*, *v*

Time Series: *t*

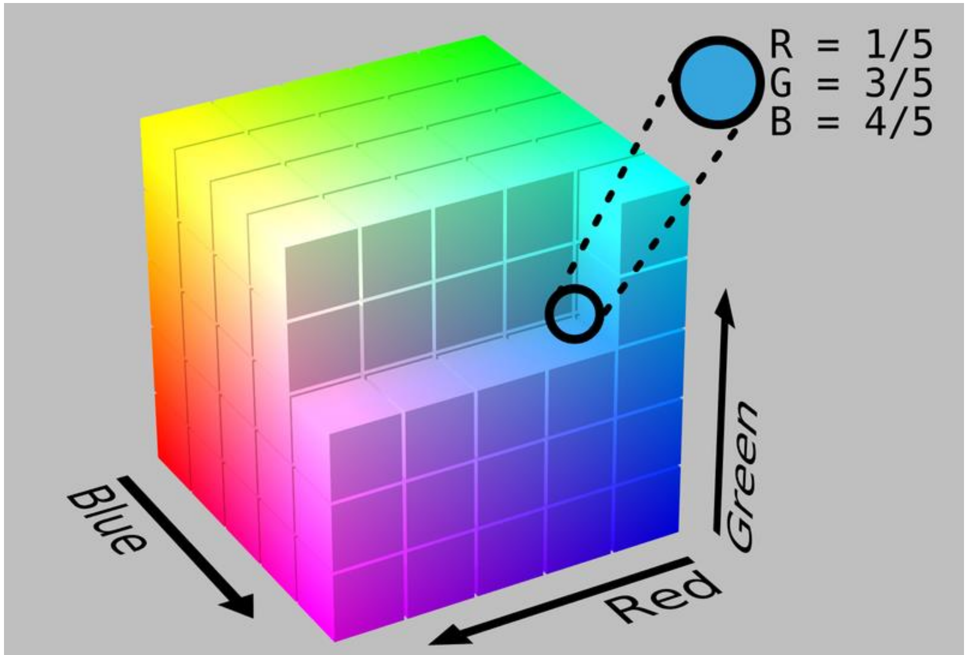
Other variables
Heatmap: H
Temperature: T
…

Color Space: **R**ed, **G**reen, **B**lue
Grayscale: **G**ray

Point Cloud: $x(u, v), y(u, v), z(u, v)$

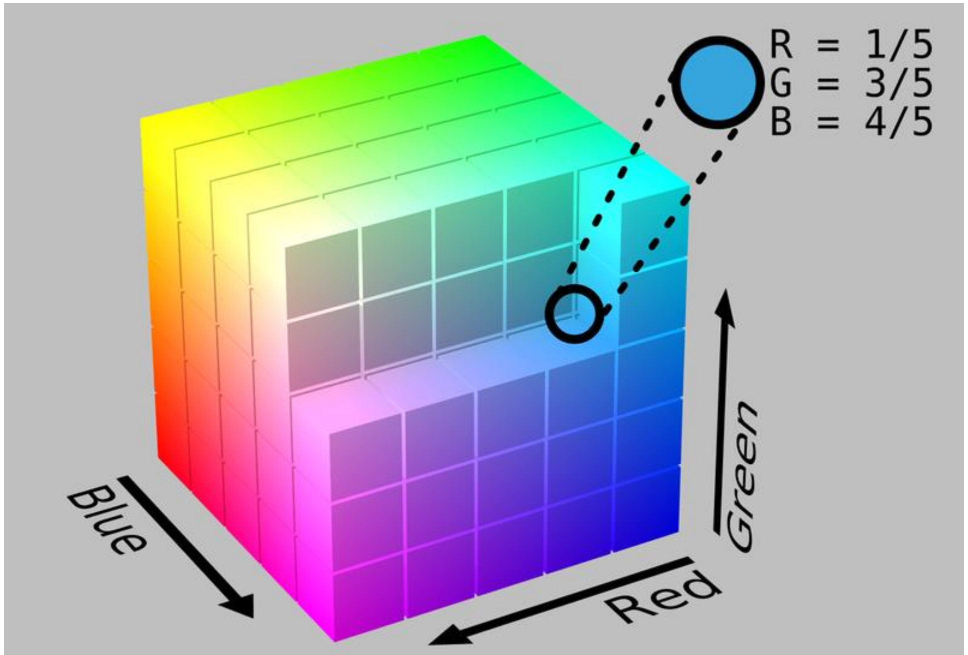
Texture: $r(x, y, z), g(x, y, z), b(x, y, z)$

[0, 1] *as normalized form, not an integer*
[0, 255] *as a byte number of range* $2^8=256$ *from 0 to 255, all in integer forms*



Color Space: **R**ed, **G**reen, **B**lue

AncoraSIR.com



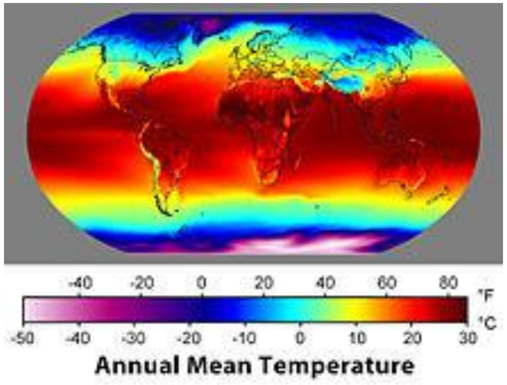
Color Space: **R**ed, **G**reen, **B**lue

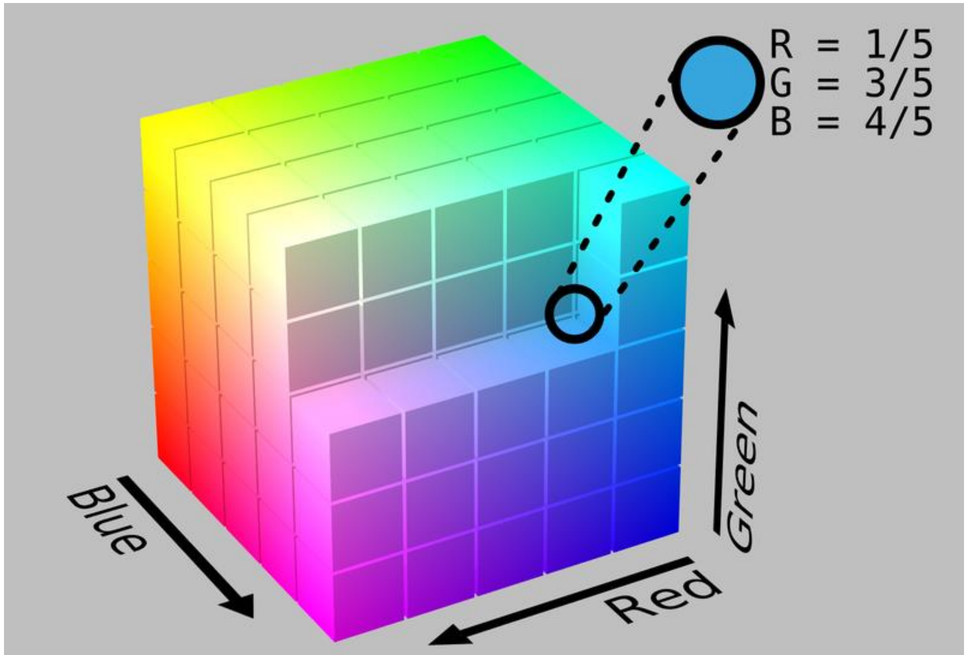
Other variables

Heatmap: H

Temperature: T

…





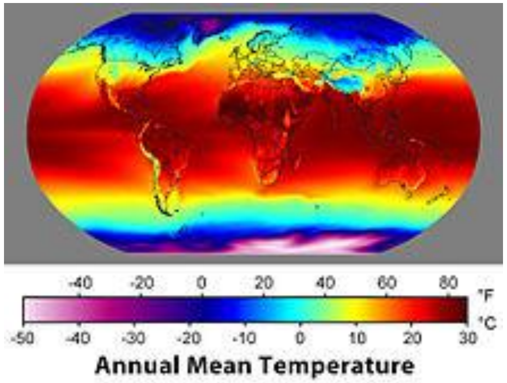
Color Space: **R**ed, **G**reen, **B**lue

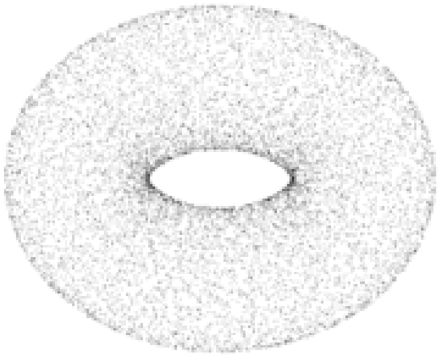
Other variables

Heatmap: H

Temperature: T

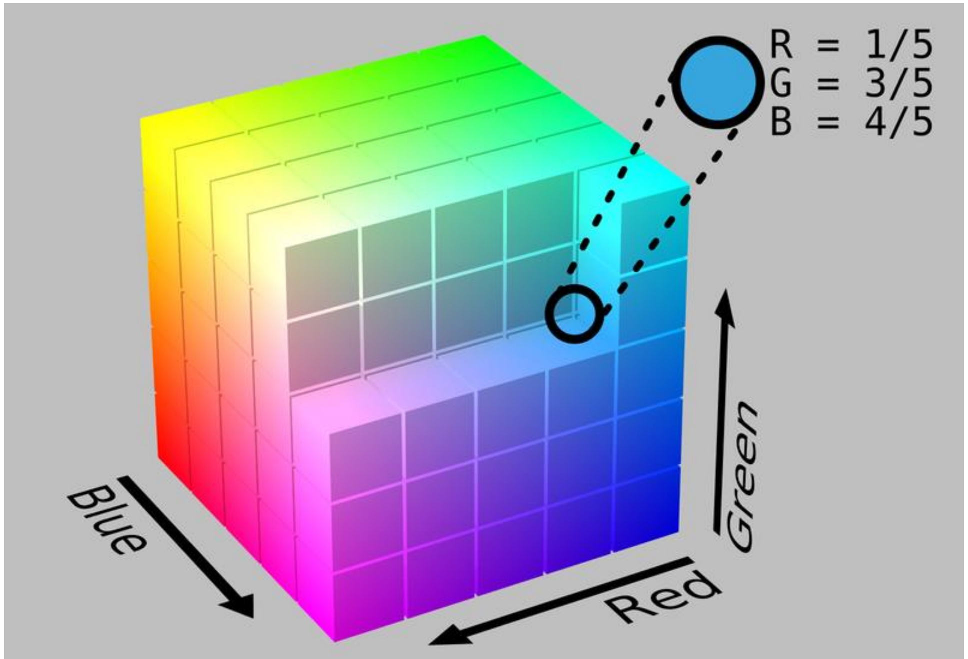
…





Point Cloud: $x(u, v)$, $y(u, v)$, $z(u, v)$

AncoraSIR.com



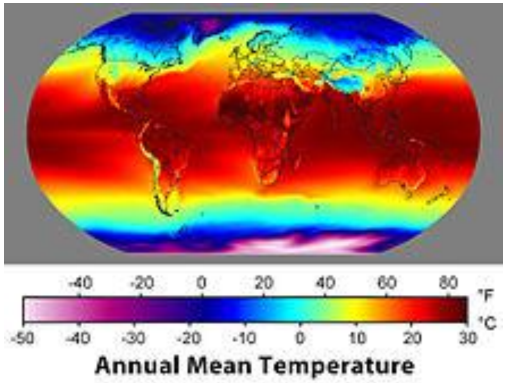
Color Space: **R**ed, **G**reen, **B**lue

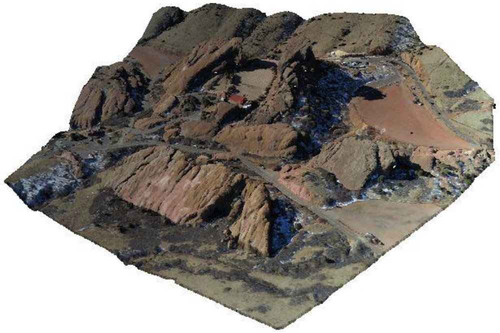
Other variables

Heatmap: H

Temperature: T

…



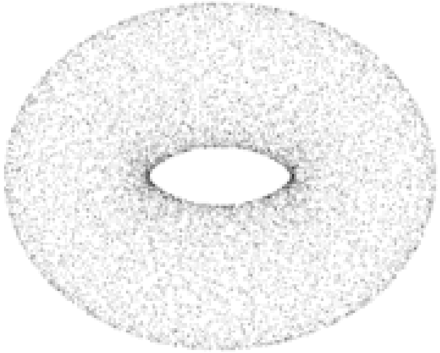


Texture:

$r(x, y, z)$,

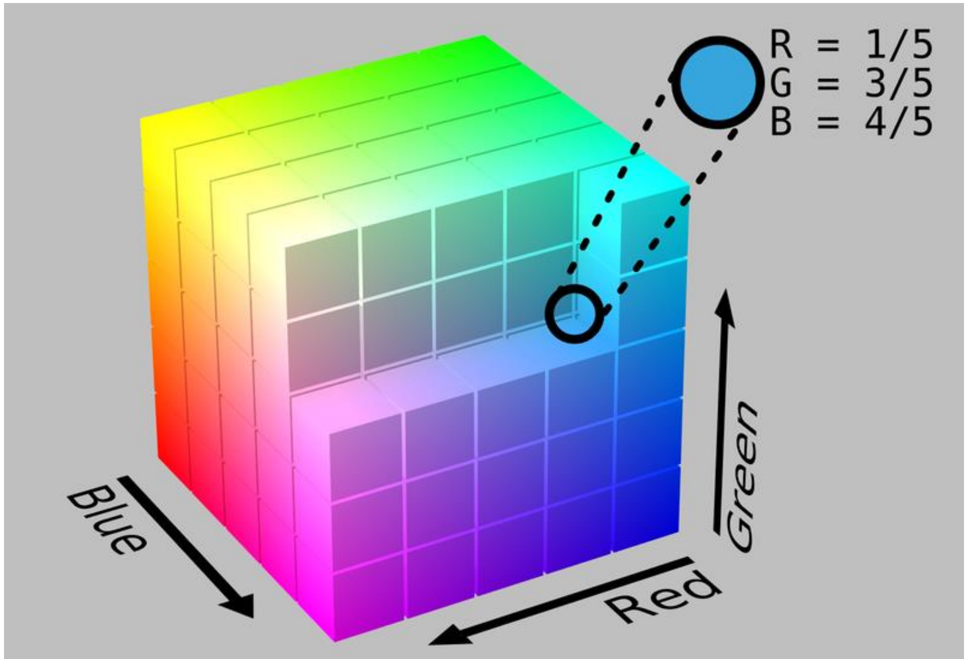
$g(x, y, z)$,

$b(x, y, z)$



Point Cloud: $x(u, v)$, $y(u, v)$, $z(u, v)$

AncoraSIR.com



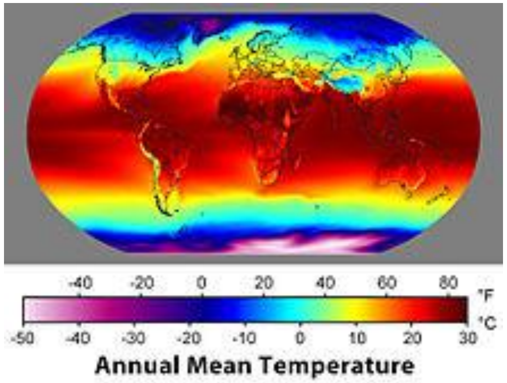
Color Space: **R**ed, **G**reen, **B**lue

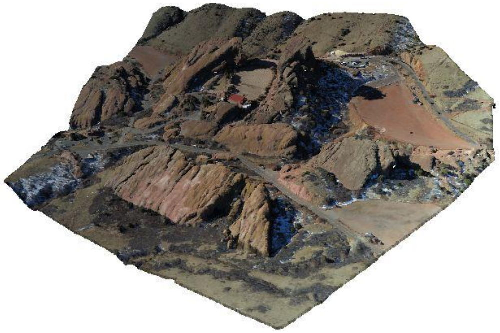
Other variables

Heatmap: H

Temperature: T

…



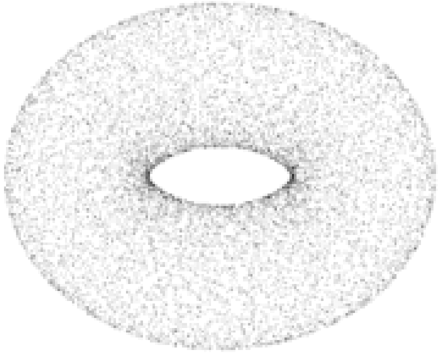


Texture:

r(*x*, *y*, *z*),

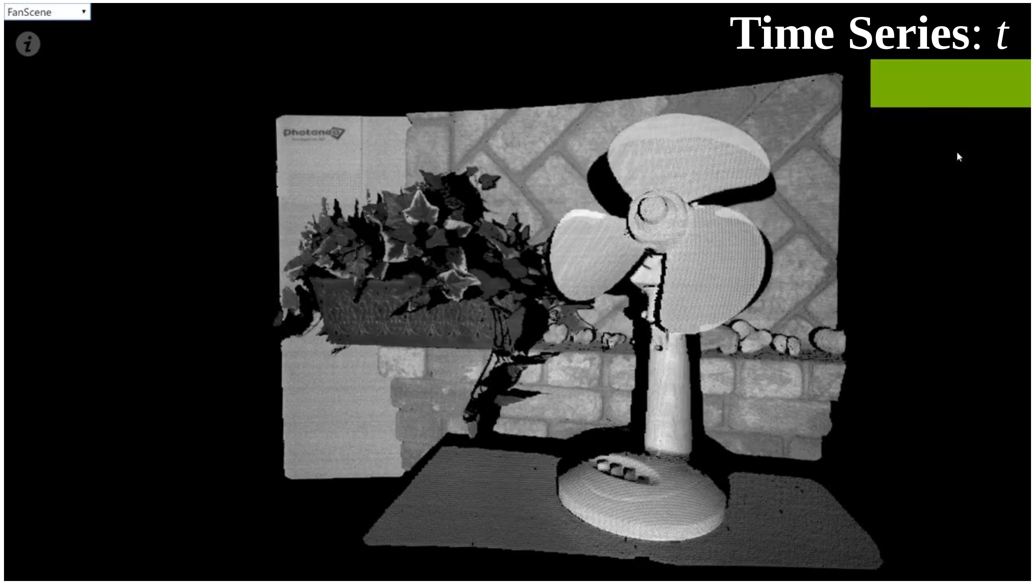
g(*x*, *y*, *z*),

b(*x*, *y*, *z*)



Point Cloud: $x(u, v), y(u, v), z(u, v)$

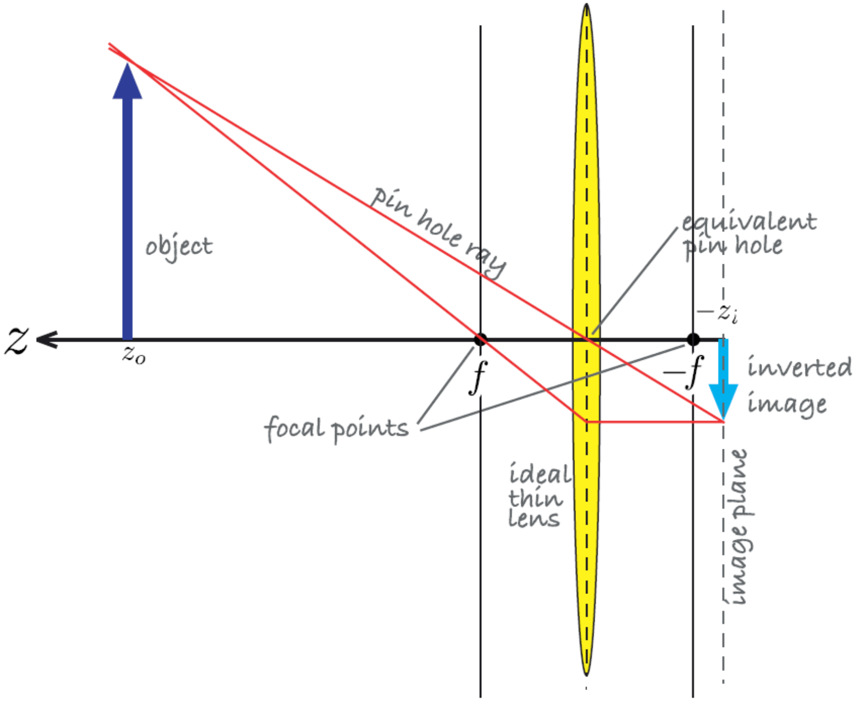
AncoraSIR.com



Time Series: *t*

Perspective Transform

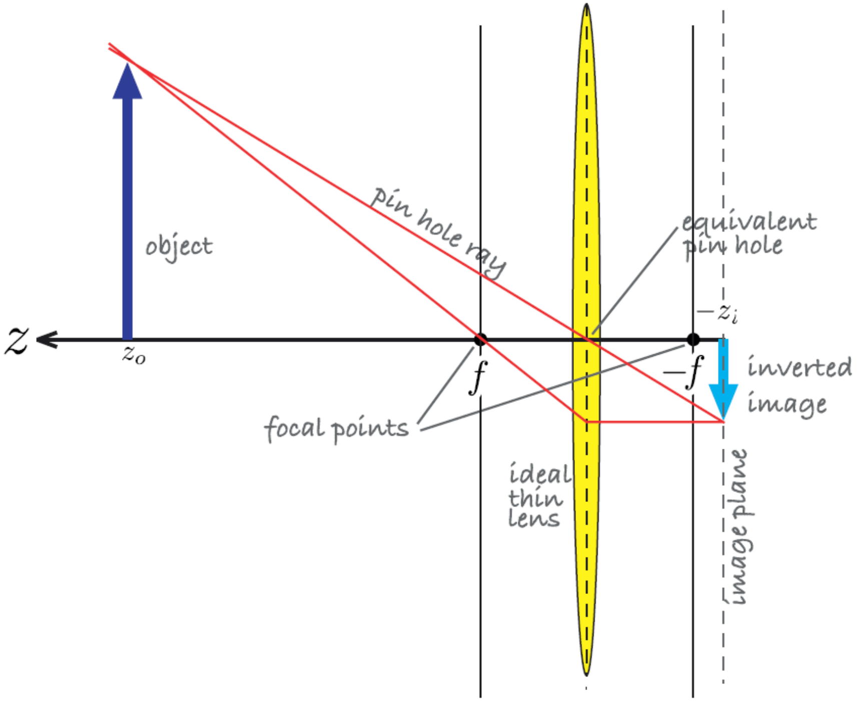
Camera Models



Perspective Transform

Camera Models

Lens Law $\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$

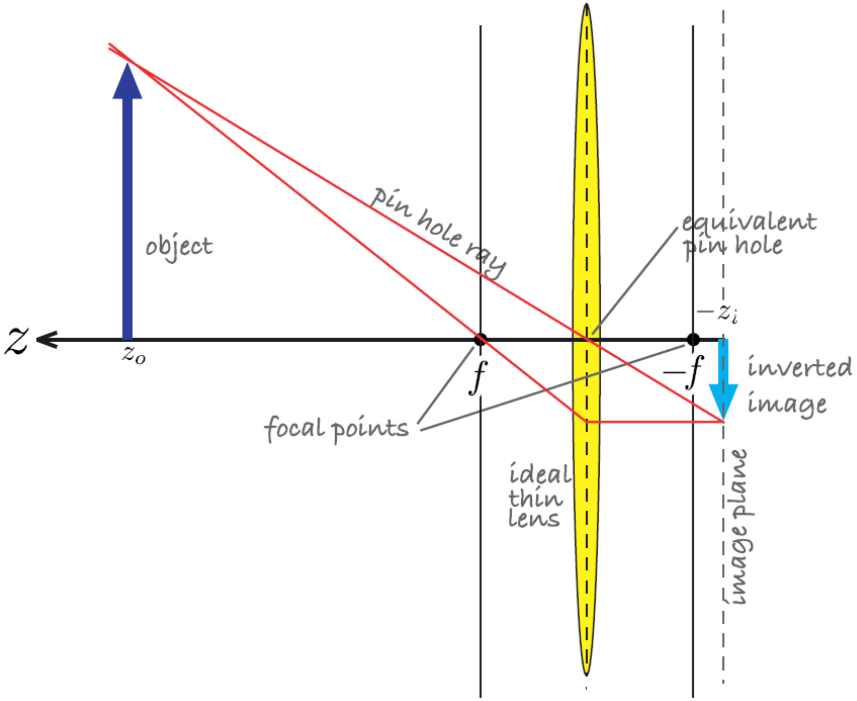


Perspective Transform

Camera Models

Lens Law $\frac{1}{z_{\mathrm{o}}} + \frac{1}{z_{\mathrm{i}}} = \frac{1}{f}$

the distance
to the object



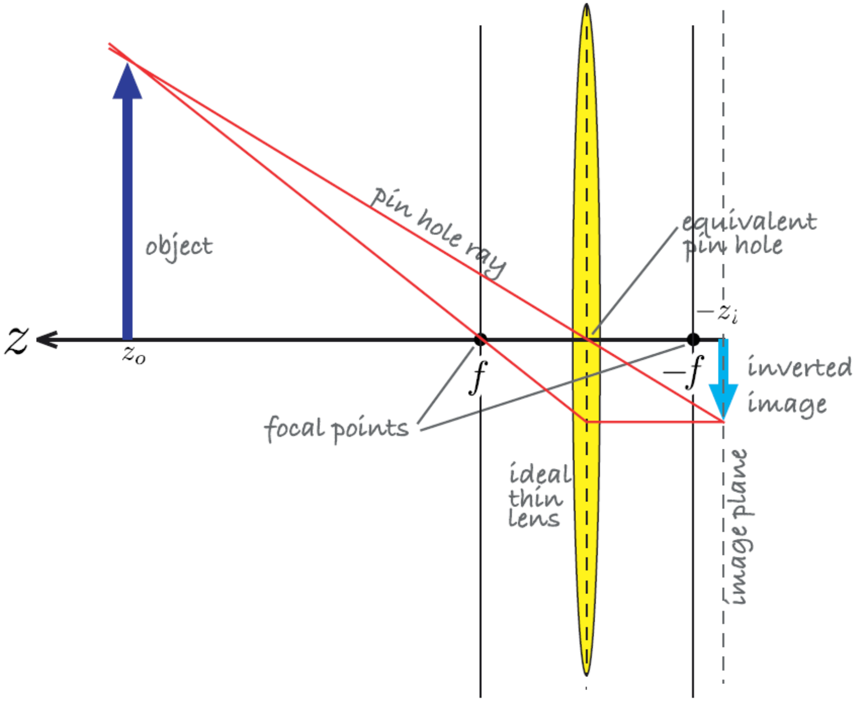
Perspective Transform

Camera Models

Lens Law $\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$

the distance to the object

the distance to the image



Perspective Transform

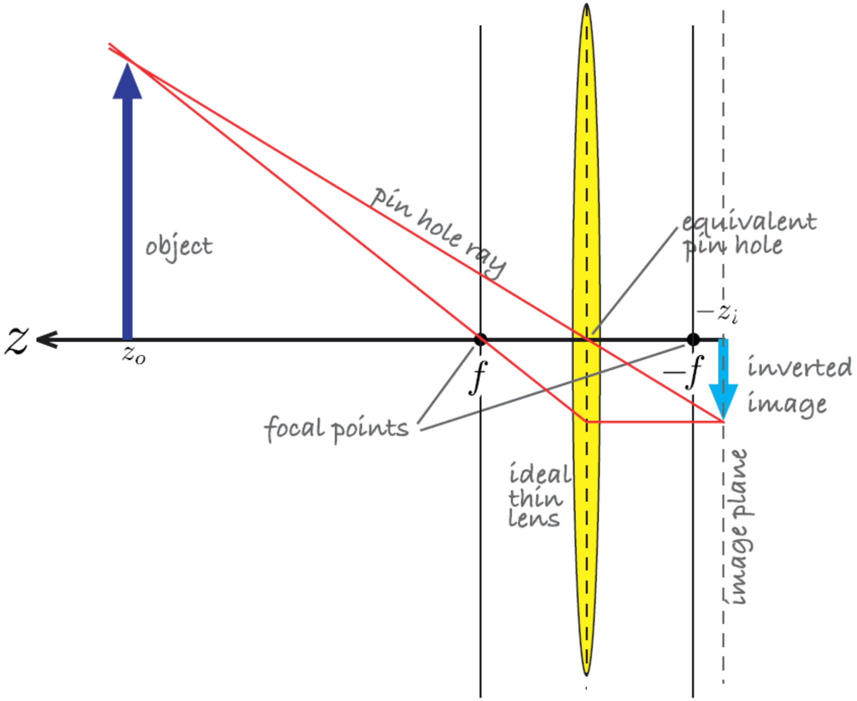
Camera Models

Lens Law $\frac{1}{z_\mathrm{o}} + \frac{1}{z_\mathrm{i}} = \frac{1}{f}$

the focal length of the lens

the distance to the object

the distance to the image



Perspective Transform

Camera Models

Lens Law $$\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$$

the focal length of the lens

the distance to the object

the distance to the image

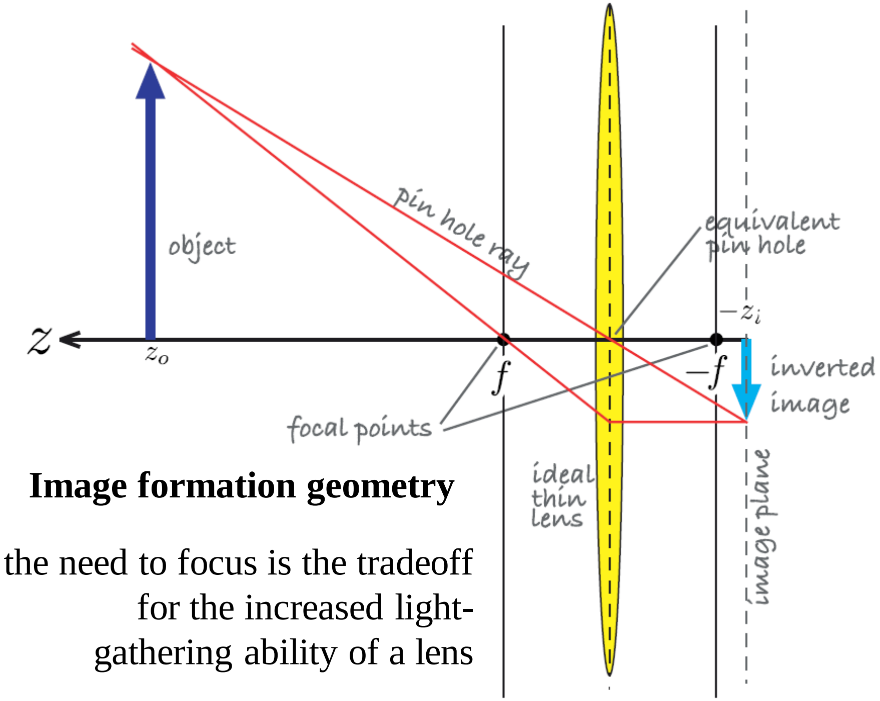


Image formation geometry

the need to focus is the tradeoff for the increased light-gathering ability of a lens

Perspective Transform

Camera Models

Lens Law $\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$

the focal length of the lens

the distance to the object

the distance to the image

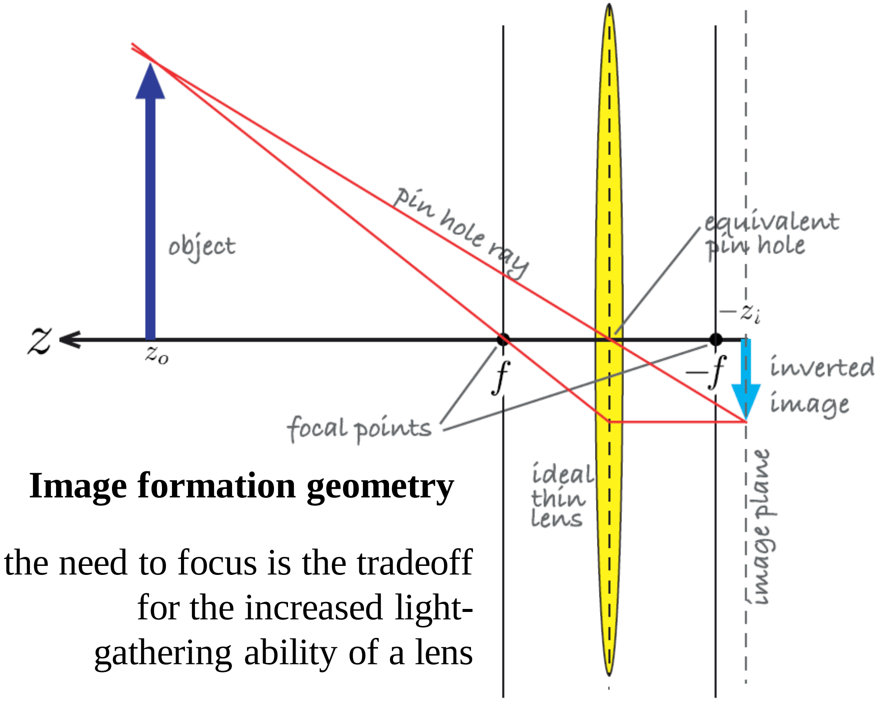
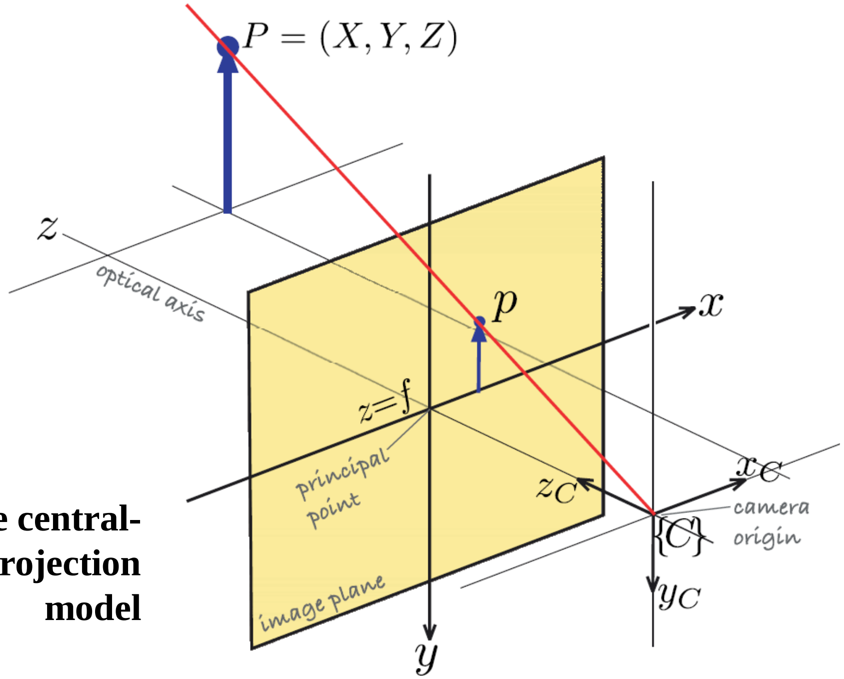


Image formation geometry

the need to focus is the tradeoff for the increased light-gathering ability of a lens



The central-projection model

Perspective Transform

Camera Models

Lens Law $\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$

the focal length of the lens

the distance to the object

the distance to the image

perspective projection $x = f\frac{X}{Z},\ y = f\frac{Y}{Z}$

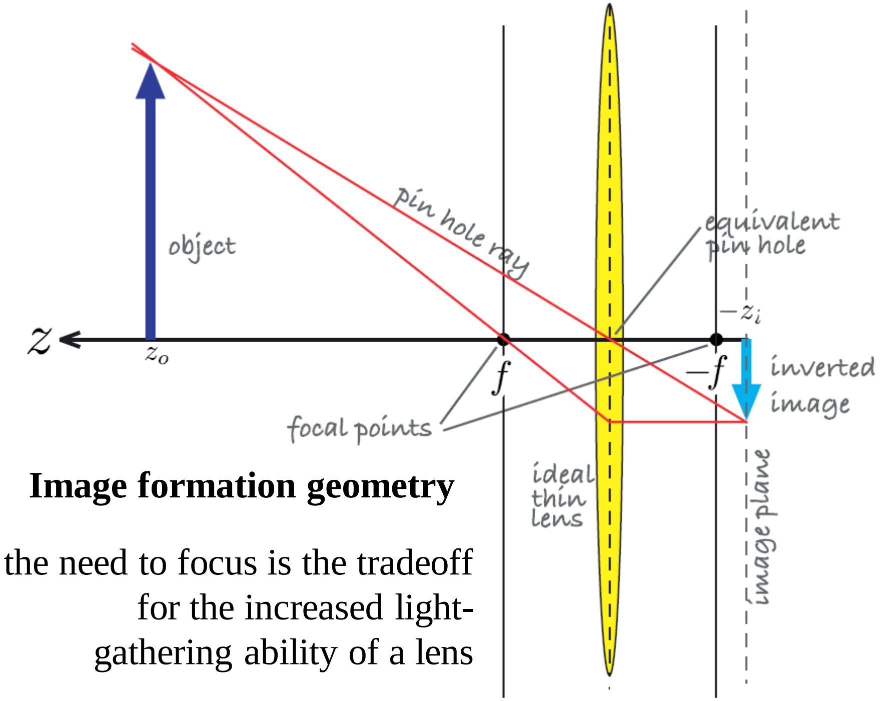
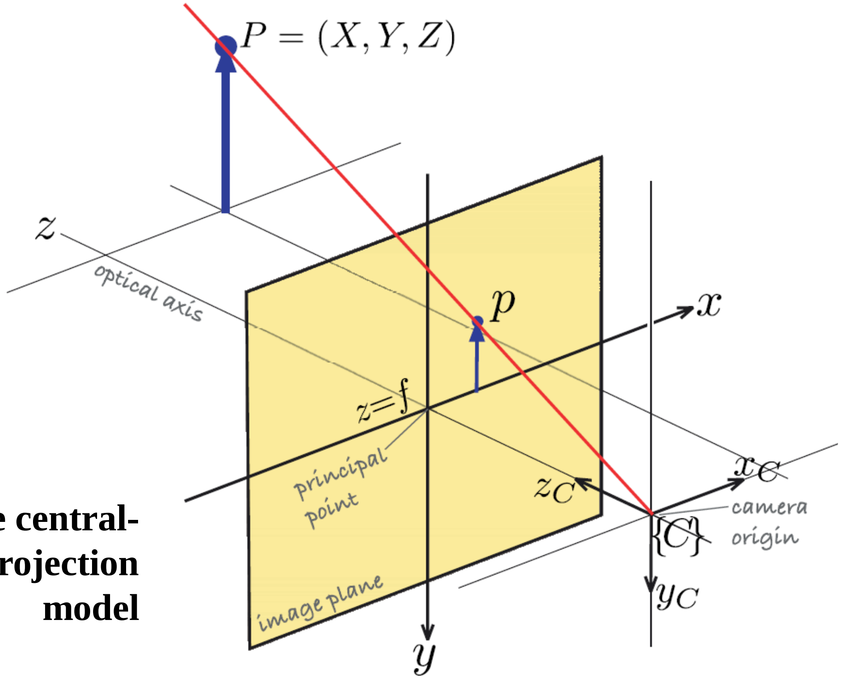


Image formation geometry

the need to focus is the tradeoff for the increased light-gathering ability of a lens



The central-projection model

Perspective Transform

Camera Models

Lens Law $\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$

the focal length of the lens

the distance to the object

the distance to the image

perspective projection $x = f\frac{X}{Z},\ y = f\frac{Y}{Z}$

$\boldsymbol{P} = (X, Y, Z)$

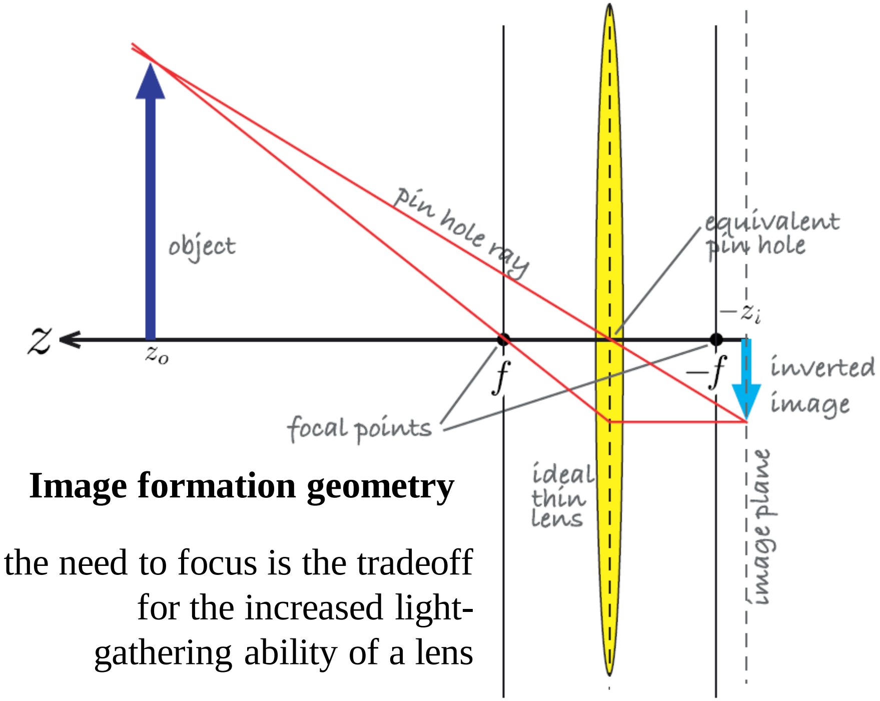
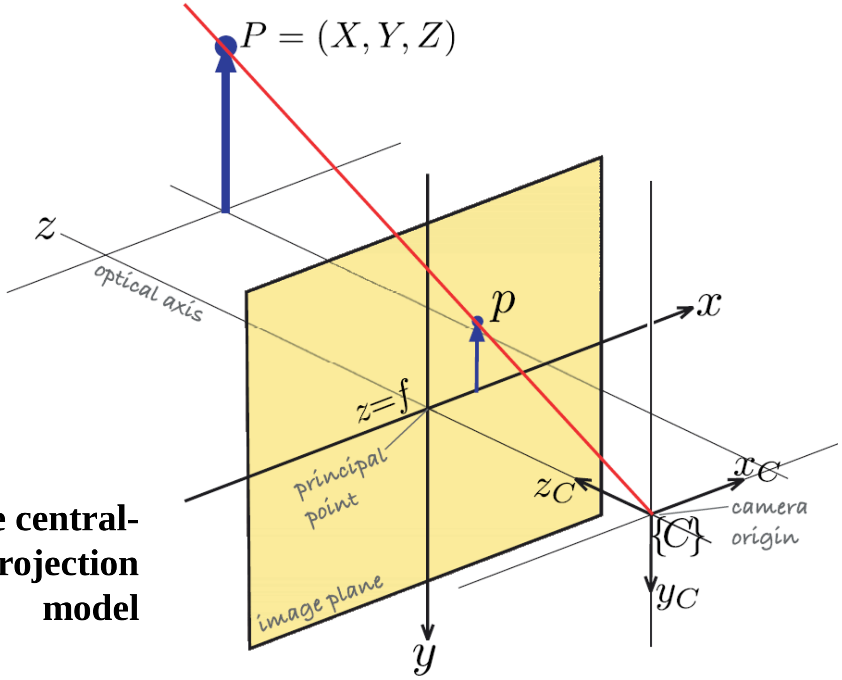


Image formation geometry

the need to focus is the tradeoff for the increased light-gathering ability of a lens



The central-projection model

Perspective Transform

Camera Models

Lens Law $\frac{1}{z_\text{o}} + \frac{1}{z_\text{i}} = \frac{1}{f}$

the focal length of the lens

the distance to the object

the distance to the image

perspective projection $x = f\frac{X}{Z},\ y = f\frac{Y}{Z}$

$$\boldsymbol{P} = (X, Y, Z) \xrightarrow{f} \boldsymbol{p} = (x, y)$$

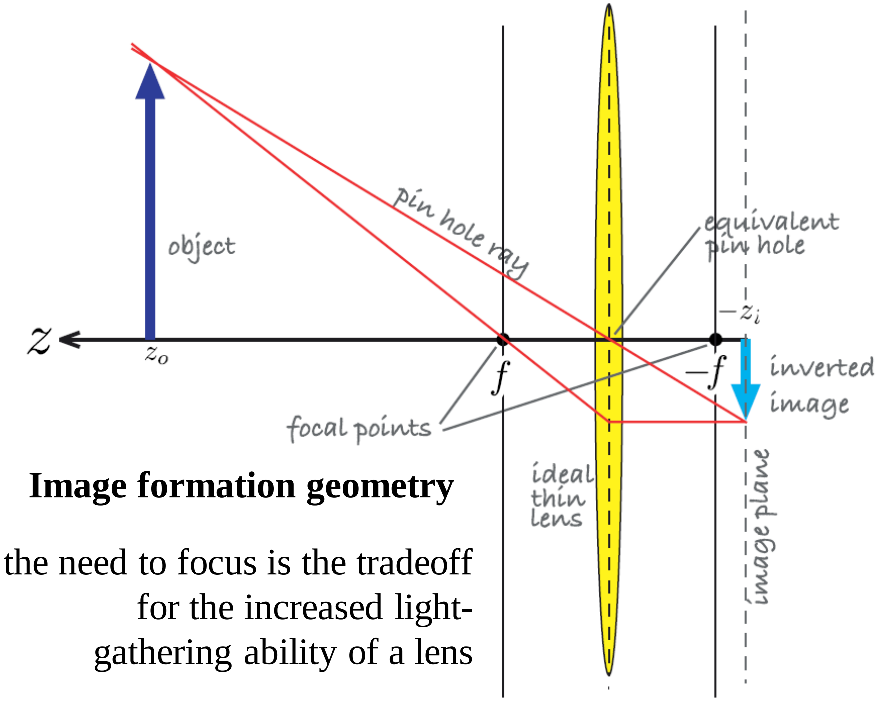
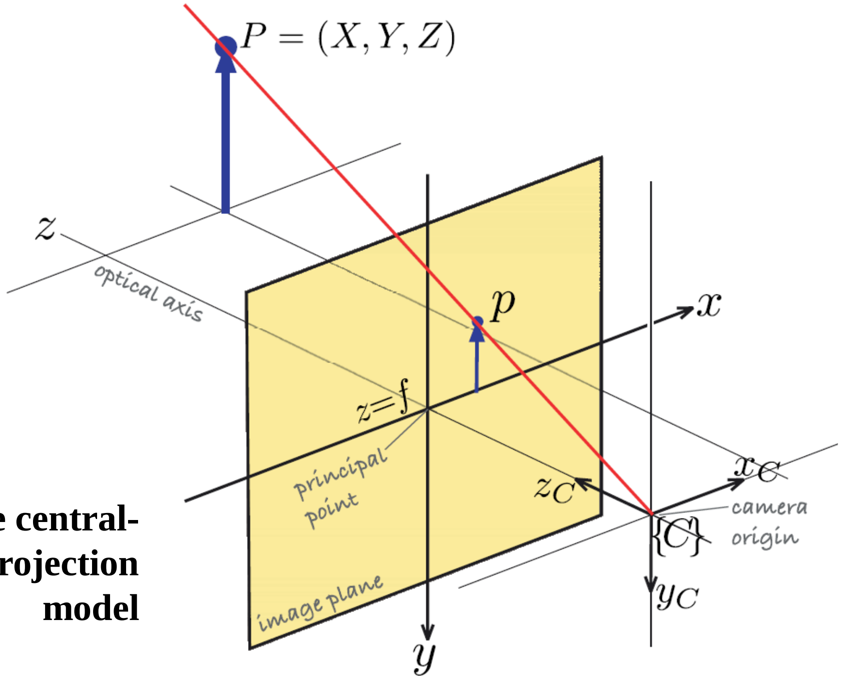


Image formation geometry

the need to focus is the tradeoff for the increased light-gathering ability of a lens



The central-projection model

Perspective Transform

Camera Models

Lens Law $\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$

the focal length of the lens

the distance to the object

the distance to the image

perspective projection $x = f\frac{X}{Z},\ y = f\frac{Y}{Z}$

$\boldsymbol{P} = (X, Y, Z) \xrightarrow{f} \boldsymbol{p} = (x, y)$

$\tilde{\boldsymbol{p}} = (x', y', z')$

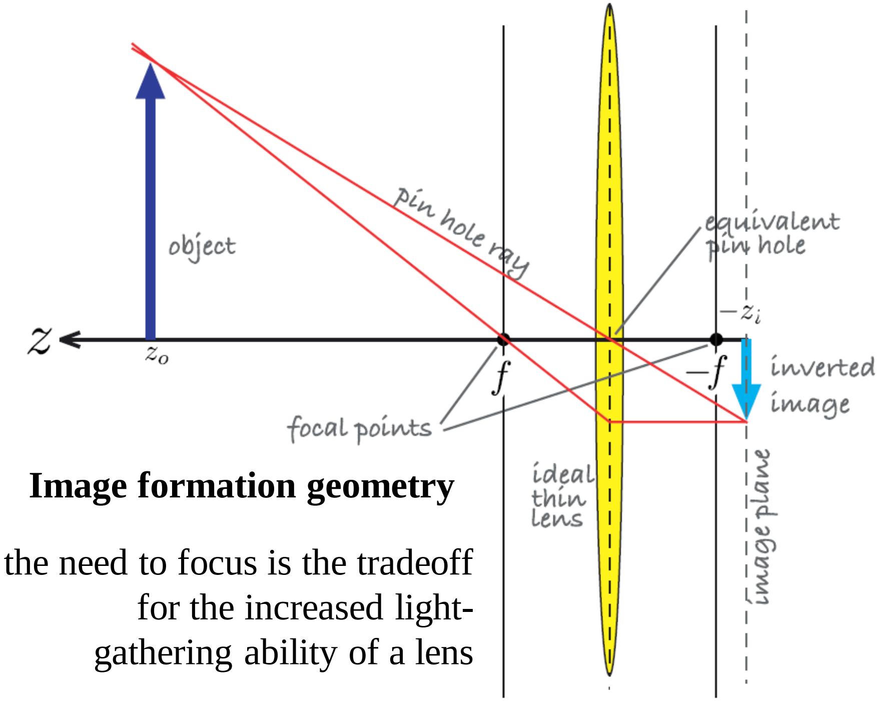
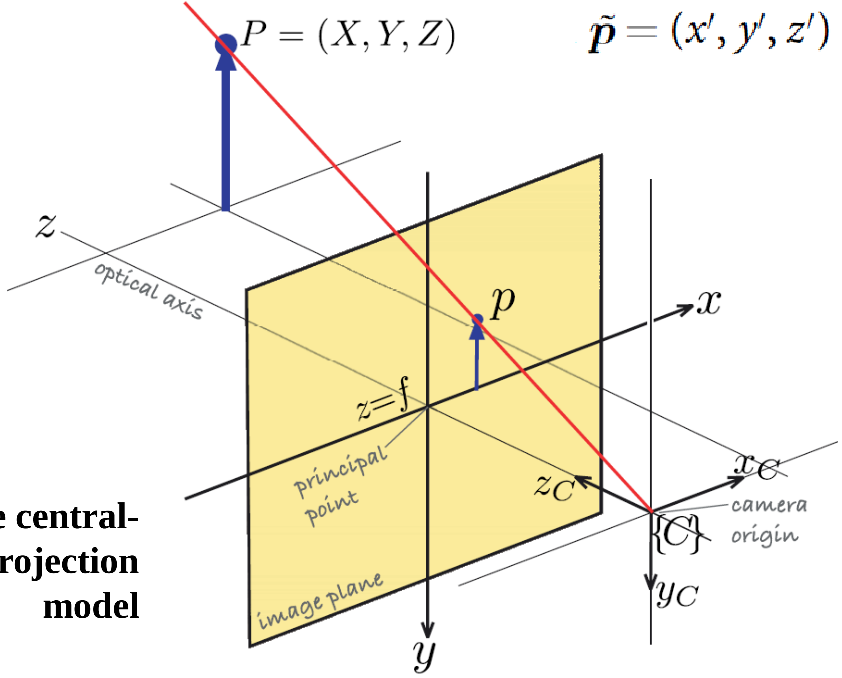


Image formation geometry

the need to focus is the tradeoff for the increased light-gathering ability of a lens



The central-projection model

Perspective Transform

Camera Models

Lens Law $\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$

the focal length of the lens

the distance to the object

the distance to the image

perspective projection $x = f\frac{X}{Z},\ y = f\frac{Y}{Z}$

$\boldsymbol{P} = (X, Y, Z) \xrightarrow{f} \boldsymbol{p} = (x, y)$

homogeneous form

$\tilde{\boldsymbol{p}} = (x', y', z')$

$x' = \frac{fX}{z'},\ y' = \frac{fY}{z'},\ z' = Z$

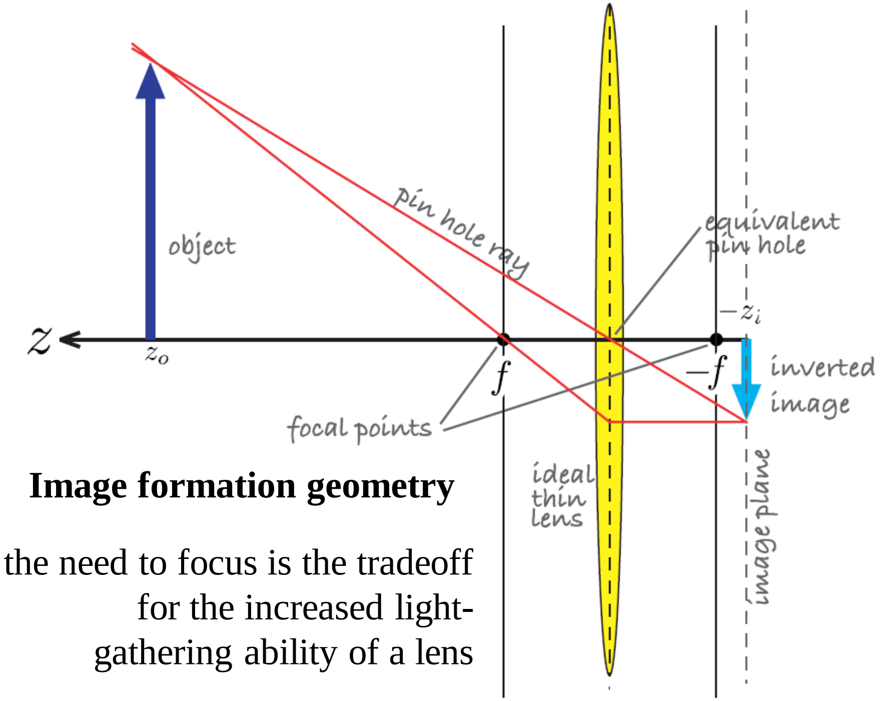
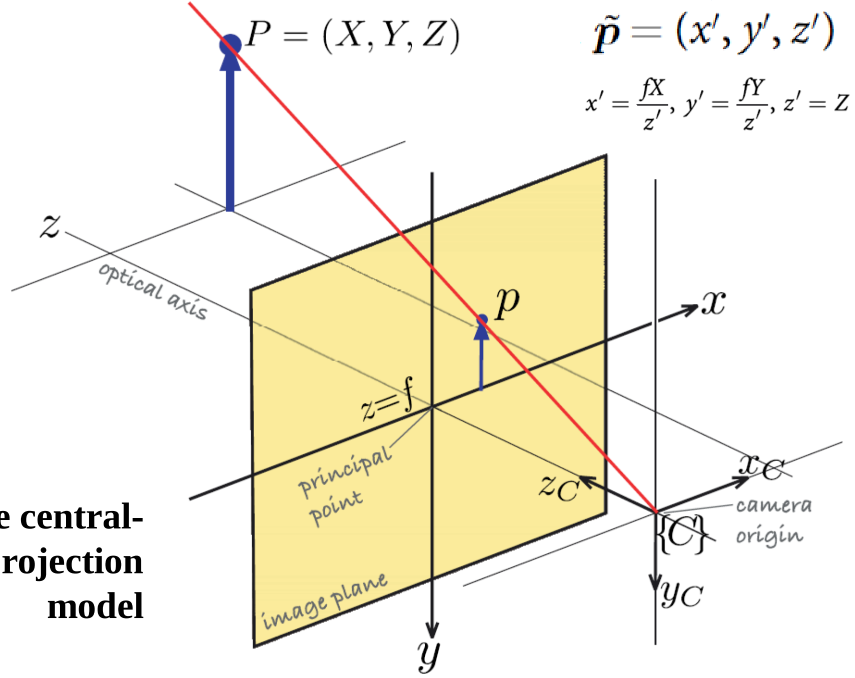


Image formation geometry

the need to focus is the tradeoff for the increased light-gathering ability of a lens



The central-projection model

Characteristics

Perspective Transform

- A mapping from 3D space to the 2D image plane $\mathbb{R}^3 \mapsto \mathbb{R}^2$.

AncoraSIR.com

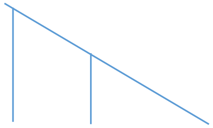
Characteristics

Perspective Transform

- A mapping from 3D space to the 2D image plane

$\mathbb{R}^3 \mapsto \mathbb{R}^2$.

- Straight lines in the world are projected to straight lines on the image plane.



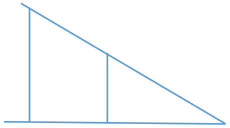
AncoraSIR.com

Characteristics

Perspective Transform

- A mapping from 3D space to the 2D image plane
- Straight lines in the world are projected to straight lines on the image plane.
- Parallel lines in the world are projected to lines that intersect at a vanishing point.
  - The exception are lines in the plane parallel to the image plane which do not converge.

$\mathbb{R}^3 \mapsto \mathbb{R}^2.$



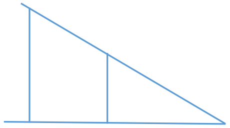
AncoraSIR.com

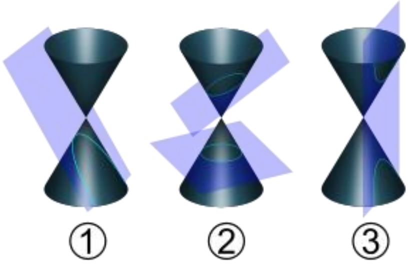
Characteristics

Perspective Transform

- A mapping from 3D space to the 2D image plane
- Straight lines in the world are projected to straight lines on the image plane.
- Parallel lines in the world are projected to lines that intersect at a vanishing point.
  - The exception are lines in the plane parallel to the image plane which do not converge.
- Conics in the world are projected to conics on the image plane.

$\mathbb{R}^3 \mapsto \mathbb{R}^2.$





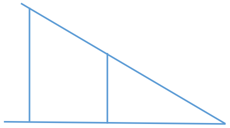
AncoraSIR.com

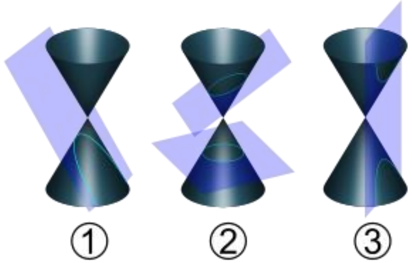
Characteristics

Perspective Transform

- A mapping from 3D space to the 2D image plane
- Straight lines in the world are projected to straight lines on the image plane.
- Parallel lines in the world are projected to lines that intersect at a vanishing point.
    - The exception are lines in the plane parallel to the image plane which do not converge.
- Conics in the world are projected to conics on the image plane.
- **The mapping is not one-to-one and a unique inverse does not exist.**

$$\mathbb{R}^3 \mapsto \mathbb{R}^2.$$





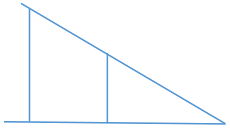
$$\boldsymbol{P} = (X, Y, Z) \xrightarrow{f} \boldsymbol{p} = (x, y)$$

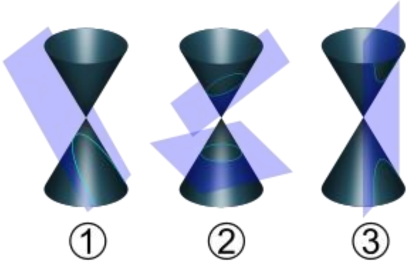
Characteristics

Perspective Transform

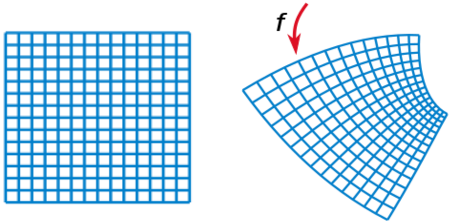
- A mapping from 3D space to the 2D image plane
- Straight lines in the world are projected to straight lines on the image plane.
- Parallel lines in the world are projected to lines that intersect at a vanishing point.
  - The exception are lines in the plane parallel to the image plane which do not converge.
- Conics in the world are projected to conics on the image plane.
- **The mapping is not one-to-one and a unique inverse does not exist.**
- The transformation is not conformal
  - It does not preserve shape since internal angles are not preserved, different from translation, rotation and scaling.

$\mathbb{R}^3 \mapsto \mathbb{R}^2.$





$$\boldsymbol{P} = (X, Y, Z) \xrightarrow{f} \boldsymbol{p} = (x, y)$$



Retinal Image Plane Coordinates

Written in homogeneous form

$\boldsymbol{P} = (X, Y, Z)$

AncoraSIR.com

Retinal Image Plane Coordinates

Written in homogeneous form

$$\begin{tikzcd}
\boldsymbol{P} = (X, Y, Z) \arrow[d, "{x = f\frac{X}{Z},\ y = f\frac{Y}{Z}}"] \\
\boldsymbol{p} = (x, y)
\end{tikzcd}$$

Retinal Image Plane Coordinates

Written in homogeneous form

$$\begin{tikzcd}
\boldsymbol{P} = (X, Y, Z) \arrow[d, "{x = f\frac{X}{Z},\ y = f\frac{Y}{Z}}"'] & \\
\boldsymbol{p} = (x, y) \arrow[r, "\text{homogeneous form}"'] & \tilde{\boldsymbol{p}} = (x', y', z')
\end{tikzcd}$$

$$x = \frac{x'}{z'},\ y = \frac{y'}{z'}$$

AncoraSIR.com

Retinal Image Plane Coordinates

Written in homogeneous form

$$\begin{tikzcd}
\boldsymbol{P}=(X,Y,Z) \arrow[dd, "{x=f\frac{X}{Z},\ y=f\frac{Y}{Z}}"'] \arrow[ddrr, "{x'=\frac{fX}{z'},\ y'=\frac{fY}{z'},\ z'=Z}", "{\tilde{\boldsymbol{p}}=\begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}}"'] & & \\
& & \\
\boldsymbol{p}=(x,y) \arrow[rr, "\text{homogeneous form}"'] & & \tilde{\boldsymbol{p}}=(x',y',z')
\end{tikzcd}$$

$$x=\frac{x'}{z'},\ y=\frac{y'}{z'}$$

Retinal Image Plane Coordinates

Written in homogeneous form

$$\begin{tikzcd}
\mathbf{P} = (X, Y, Z) \arrow[rr, "\text{homogeneous form}"] \arrow[ddd, "{x = f\frac{X}{Z},\ y = f\frac{Y}{Z}}"'] \arrow[dr, dash] & & {}^{C}\tilde{\mathbf{P}} = (X, Y, Z, 1)^{T} \\
& x' = \frac{fX}{z'},\ y' = \frac{fY}{z'},\ z' = Z & \\
& \tilde{\mathbf{p}} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \arrow[dr] & \\
\mathbf{p} = (x, y) \arrow[rr, "\text{homogeneous form}"'] & & \tilde{\mathbf{p}} = (x', y', z')
\end{tikzcd}$$

$$x = \frac{x'}{z'},\ y = \frac{y'}{z'}$$

AncoraSIR.com

Retinal Image Plane Coordinates

Written in homogeneous form

$$\begin{tikzcd}[ampersand replacement=\&]
\boldsymbol{P} = (X, Y, Z) \arrow[r, "\text{homogeneous form}"] \arrow[d, "{x = f\frac{X}{Z},\ y = f\frac{Y}{Z}}"'] \arrow[dr, "{x' = \frac{fX}{z'},\ y' = \frac{fY}{z'},\ z' = Z \quad \tilde{\boldsymbol{p}} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}}"] \& {}^{C}\tilde{\boldsymbol{P}} = (X, Y, Z, 1)^T \arrow[d, "{\text{camera matrix} \quad \tilde{\boldsymbol{p}} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} {}^{C}\tilde{\boldsymbol{P}}}"] \\
\boldsymbol{p} = (x, y) \arrow[r, "\text{homogeneous form}"'] \& \tilde{\boldsymbol{p}} = (x', y', z')
\end{tikzcd}$$

$$x = \frac{x'}{z'},\ y = \frac{y'}{z'}$$

Retinal Image Plane Coordinates

Written in homogeneous form

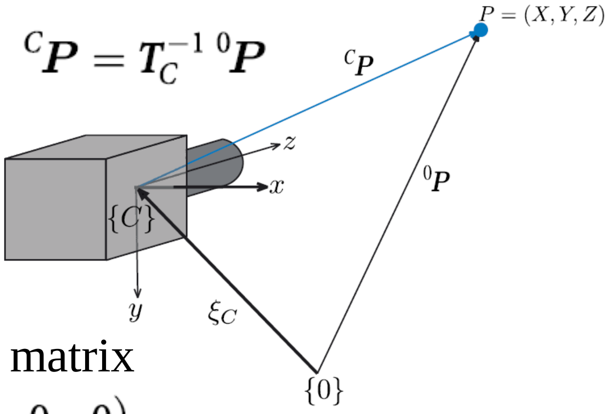
$$\begin{tikzcd}
\boldsymbol{P} = (X, Y, Z) \arrow[rr, "\text{homogeneous form}"] \arrow[dd, "{x = f\frac{X}{Z}, y = f\frac{Y}{Z}}"'] \arrow[dr, dash] & & {}^{C}\tilde{\boldsymbol{P}} = (X, Y, Z, 1)^T \arrow[dd] & \\
& \begin{array}{c} x' = \frac{fX}{z'}, y' = \frac{fY}{z'}, z' = Z \\ \tilde{\boldsymbol{p}} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \end{array} \arrow[dr] & & \begin{array}{c} \text{camera matrix} \\ \tilde{\boldsymbol{p}} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} {}^{C}\tilde{\boldsymbol{P}} \end{array} \arrow[d] \\
\boldsymbol{p} = (x, y) \arrow[rr, "\text{homogeneous form}"', "{x = \frac{x'}{z'}, y = \frac{y'}{z'}}"' near end, shift right=0] & & \tilde{\boldsymbol{p}} = (x', y', z') & \tilde{\boldsymbol{p}} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} {}^{C}\tilde{\boldsymbol{P}}
\end{tikzcd}$$

Retinal Image Plane Coordinates

Written in homogeneous form

$$\begin{tikzcd}
\boldsymbol{P} = (X, Y, Z) \arrow[rr, "\text{homogeneous form}"] \arrow[dd, "x = f\frac{X}{Z},\ y = f\frac{Y}{Z}"'] \arrow[dr] & & {}^{C}\tilde{\boldsymbol{P}} = (X, Y, Z, 1)^T \arrow[dd] \\
& {\begin{array}{c} x' = \frac{fX}{z'},\ y' = \frac{fY}{z'},\ z' = Z \\ \tilde{\boldsymbol{p}} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \end{array}} \arrow[dr] & \\
\boldsymbol{p} = (x, y) \arrow[rr, "\text{homogeneous form}"'] & & \tilde{\boldsymbol{p}} = (x', y', z')
\end{tikzcd}$$

$$x = \frac{x'}{z'},\ y = \frac{y'}{z'}$$



$${}^{C}\boldsymbol{P} = \boldsymbol{T}_C^{-1}\, {}^{0}\boldsymbol{P}$$

camera matrix

$$\tilde{\boldsymbol{p}} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} {}^{C}\tilde{\boldsymbol{P}}$$

$$\tilde{\boldsymbol{p}} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} {}^{C}\tilde{\boldsymbol{P}}$$

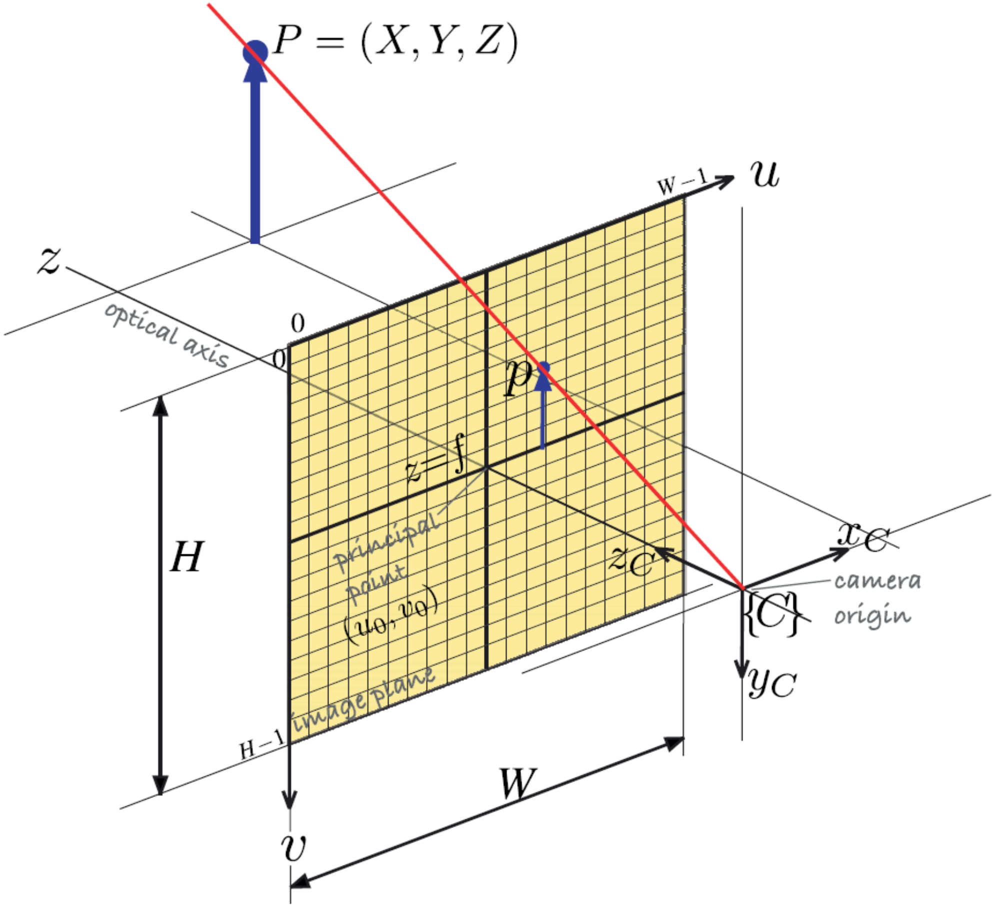
What about the physical coordinates on the actual image?

AncoraSIR.com

Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W**× **H** grid of image pixels
  - The pixel coordinates (*u*, *v*)

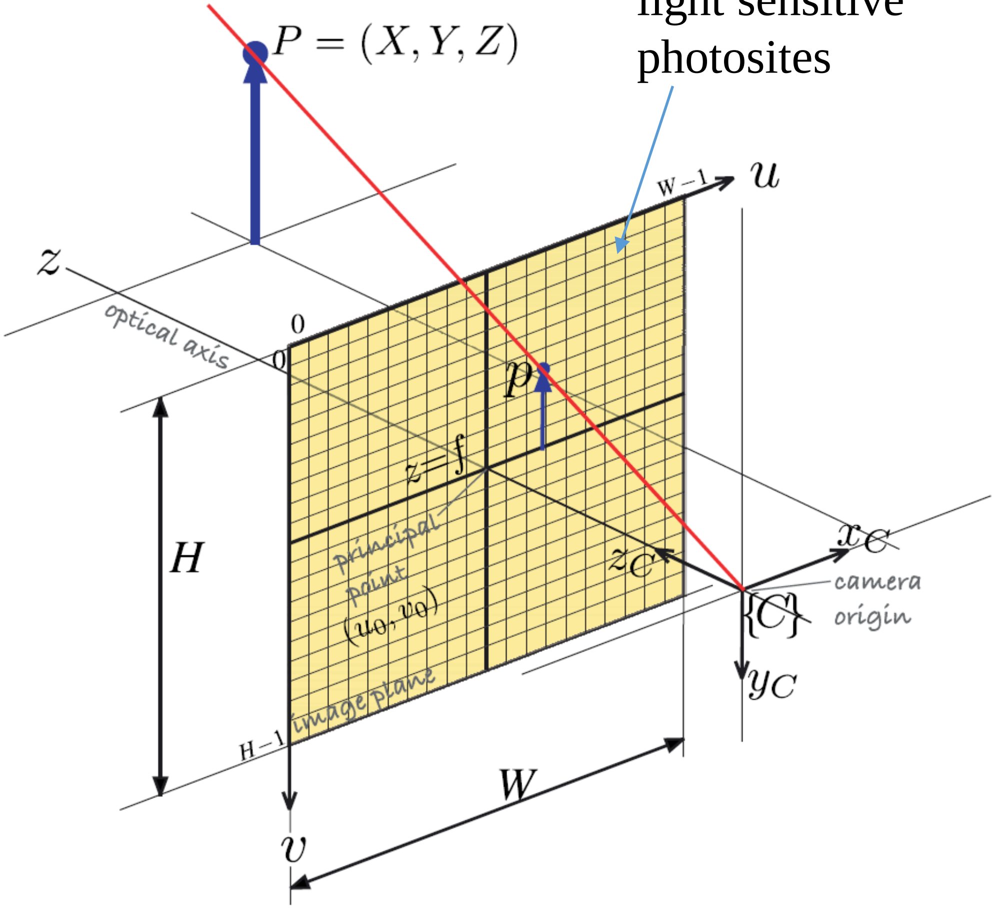


AncoraSIR.com

Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W**× **H** grid of image pixels
  - The pixel coordinates (u, v)



light sensitive photosites

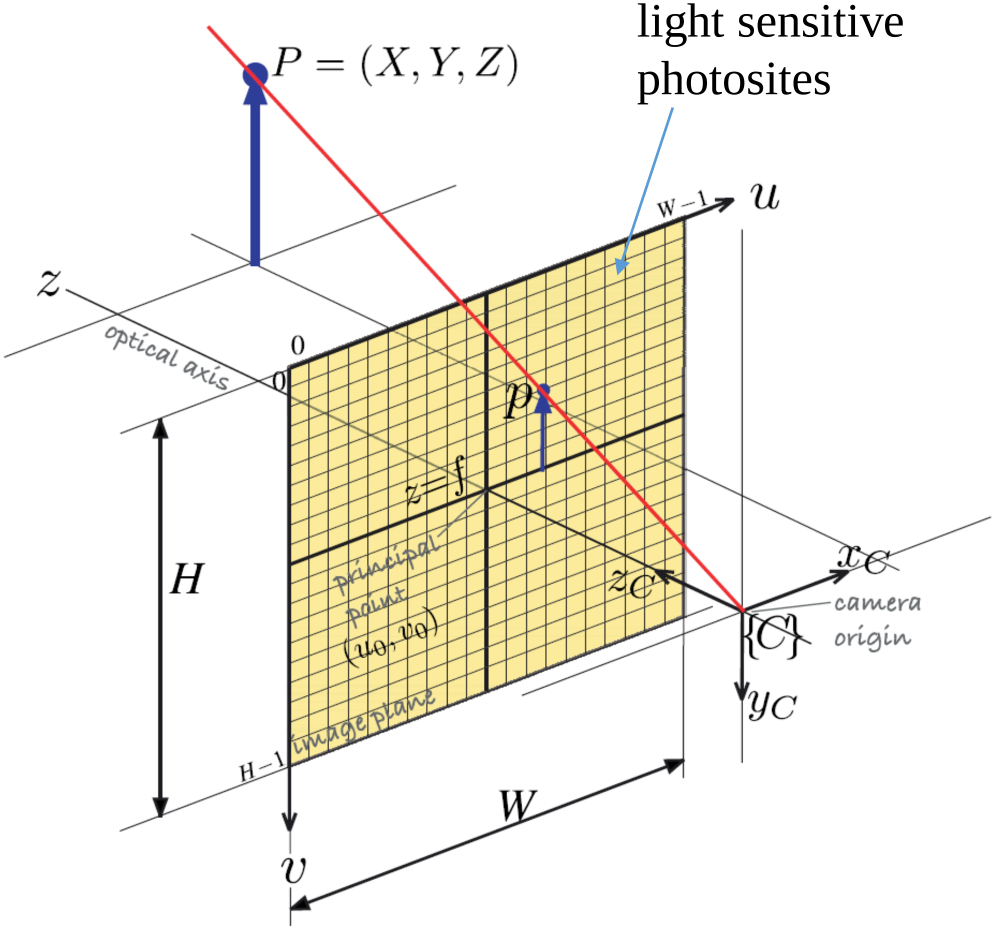
AncoraSIR.com

Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W**× **H** grid of image pixels
  - The pixel coordinates (u, v)

$$u = \frac{x}{\rho_w} + u_0,\ v = \frac{y}{\rho_h} + v_0$$



AncoraSIR.com

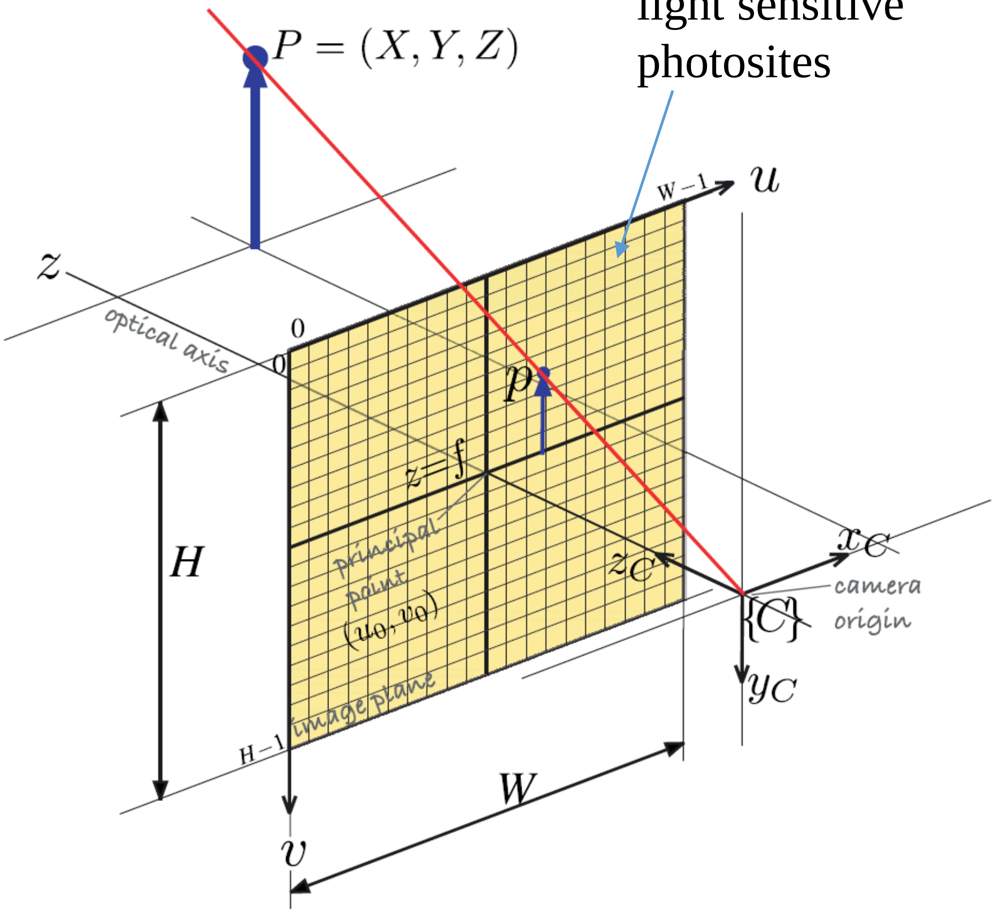
Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W× H** grid of image pixels
    - The pixel coordinates (u, v)

Principal point in pixel coordinate

$$u = \frac{x}{\rho_w} + u_0,\ v = \frac{y}{\rho_h} + v_0$$



light sensitive photosites

Express w.r.t the Camera

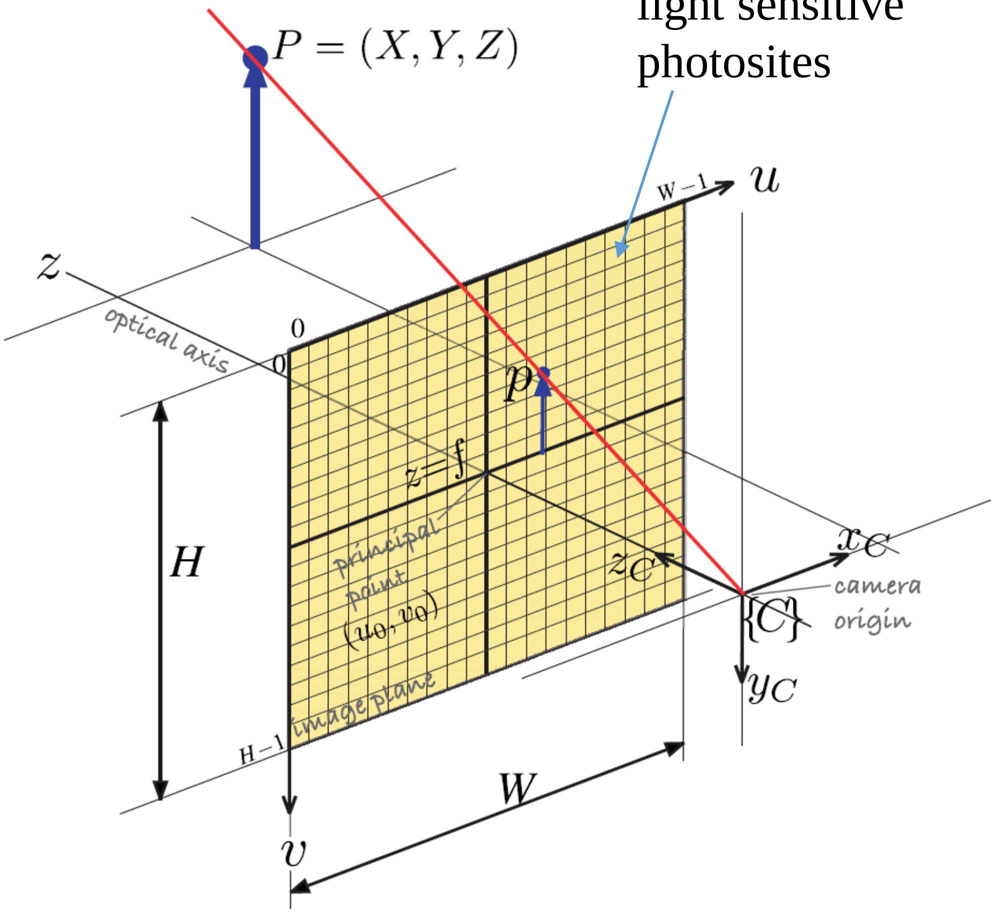
Physical Meanings of Camera Pixels

- A **camera sensor** with a **W**× **H** grid of image pixels
  - The pixel coordinates (*u*, *v*)

Principal point in pixel coordinate

$$u = \frac{x}{\rho_w} + u_0,\ v = \frac{y}{\rho_h} + v_0$$

width and height of each pixel



light sensitive photosites

AncoraSIR.com

Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W**× **H** grid of image pixels
  - The pixel coordinates (u, v)

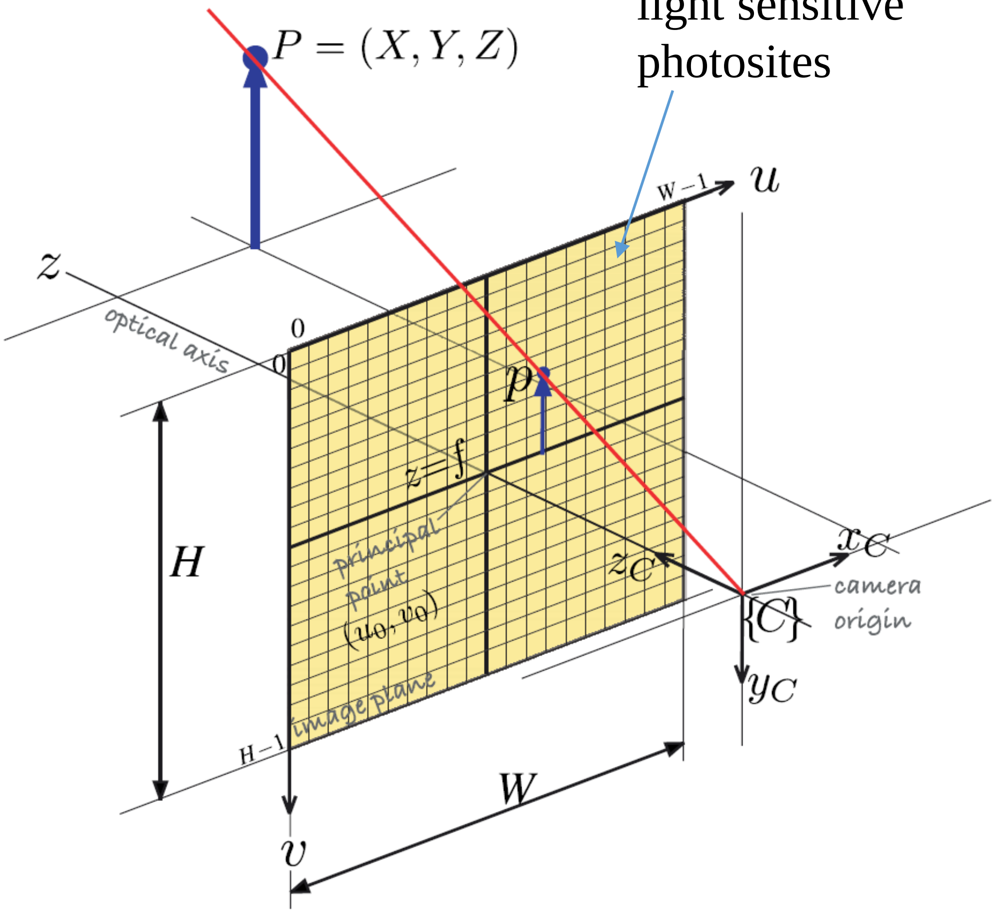
Principal point in pixel coordinate

$$u = \frac{x}{\rho_w} + u_0,\ v = \frac{y}{\rho_h} + v_0$$

width and height of each pixel

$$\tilde{\boldsymbol{p}} = (u', v', w')$$

pixel coordinate



light sensitive photosites

Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W** × **H** grid of image pixels
  - The pixel coordinates (u, v)

Principal point in pixel coordinate

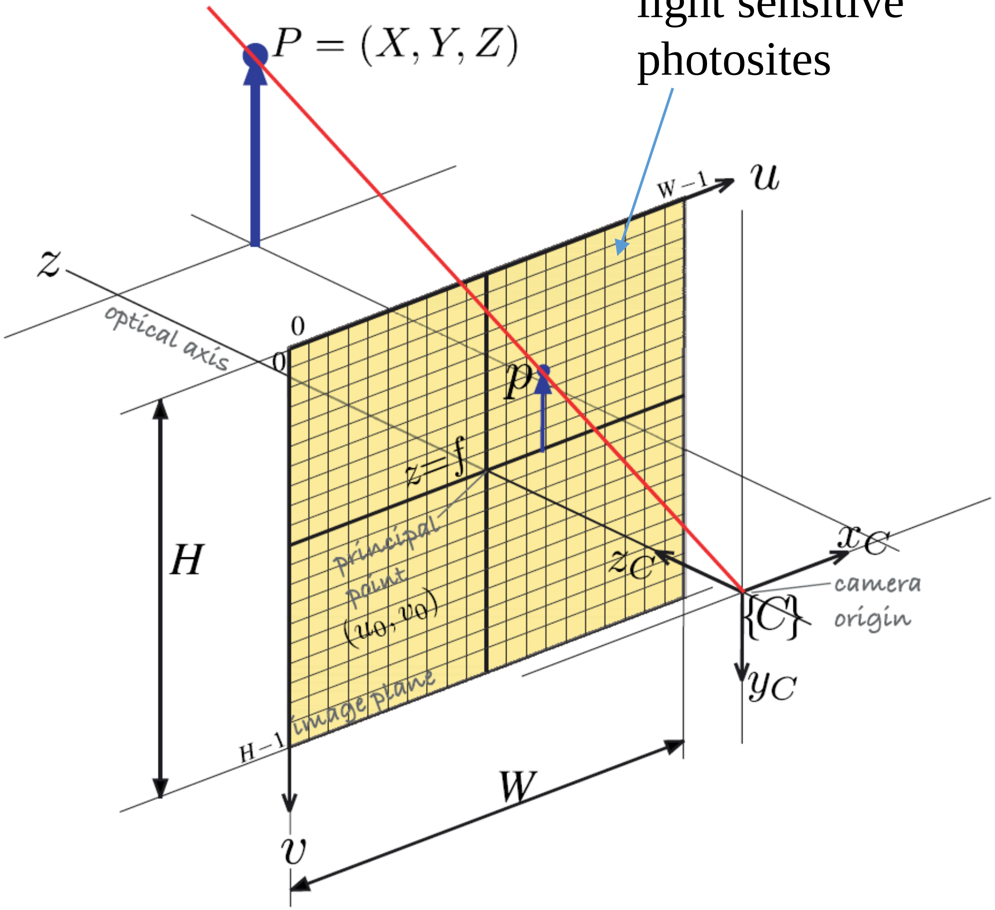
$$u = \frac{x}{\rho_w} + u_0,\ v = \frac{y}{\rho_h} + v_0$$

width and height of each pixel

$$\tilde{p} = (u', v', w')$$

pixel coordinate

$$u = \frac{u'}{w'},\ v = \frac{v'}{w'}$$



light sensitive photosites

AncoraSIR.com

Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W**× **H** grid of image pixels
  - The pixel coordinates (u, v)

Principal point in pixel coordinate

$$u = \frac{x}{\rho_w} + u_0,\ v = \frac{y}{\rho_h} + v_0$$

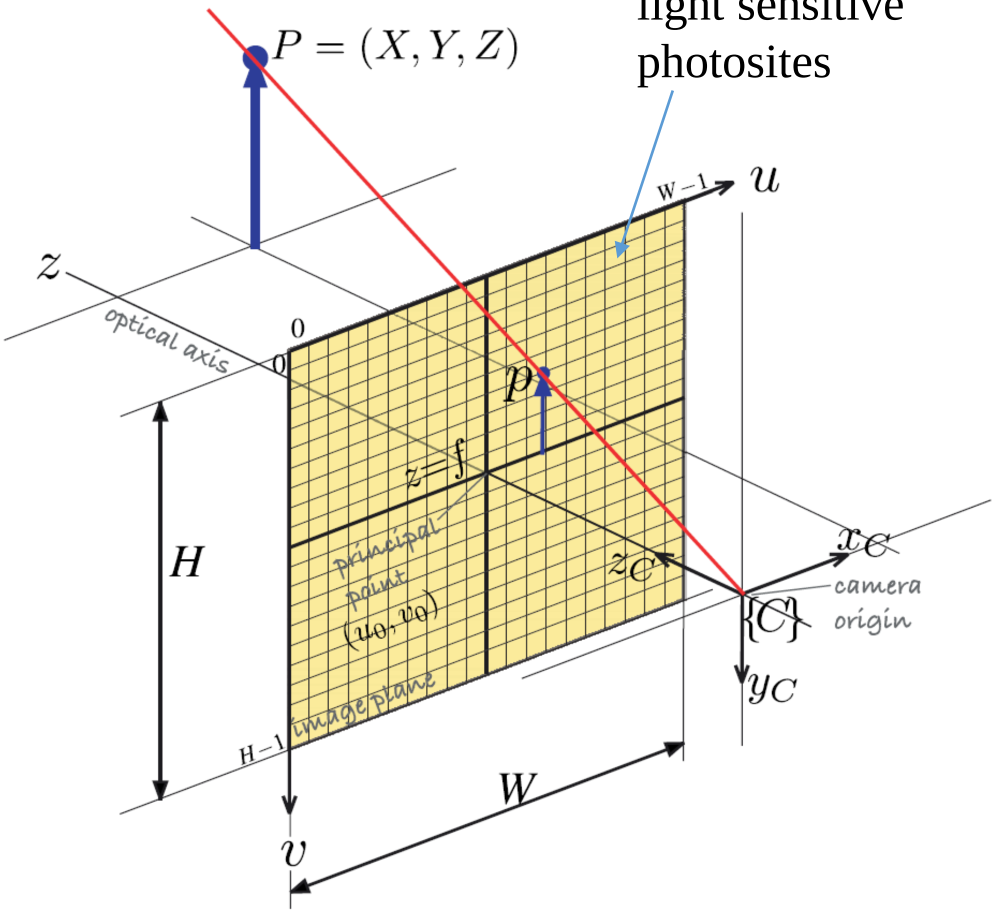
width and height of each pixel

$$\tilde{\boldsymbol{p}} = (u', v', w')$$

pixel coordinate

$$u = \frac{u'}{w'},\ v = \frac{v'}{w'}$$

$$\tilde{\boldsymbol{p}} = \underbrace{\begin{pmatrix} 1/\rho_w & 0 & u_0 \\ 0 & 1/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{K} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} {}^{C}\tilde{\boldsymbol{P}}$$



light sensitive photosites

AncoraSIR.com

Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W**× **H** grid of image pixels
  - The pixel coordinates (*u*, *v*)

Principal point in pixel coordinate

$$u = \frac{x}{\rho_w} + u_0,\ v = \frac{y}{\rho_h} + v_0$$

width and height of each pixel

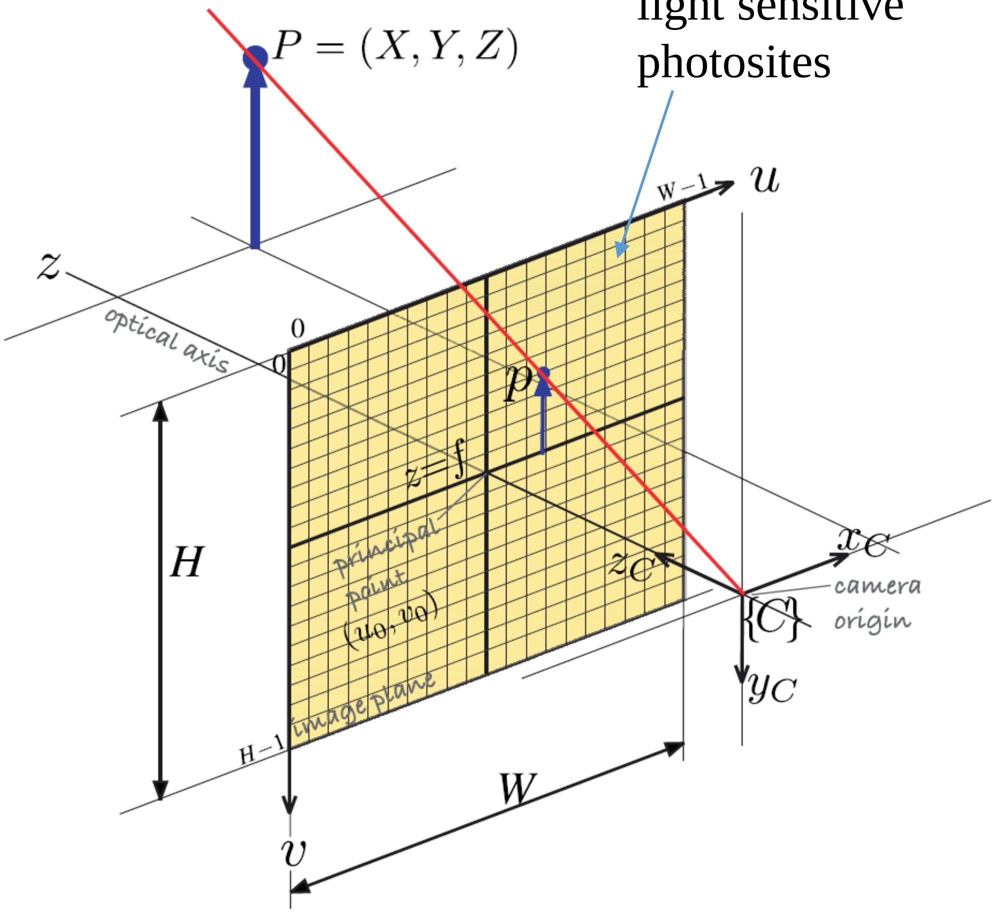
$$\tilde{\boldsymbol{p}} = (u', v', w')$$

pixel coordinate

$$u = \frac{u'}{w'},\ v = \frac{v'}{w'}$$

$$\tilde{\boldsymbol{p}} = \underbrace{\begin{pmatrix} 1/\rho_w & 0 & u_0 \\ 0 & 1/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{K} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} {}^{C}\tilde{\boldsymbol{P}}$$

camera parameter matrix



light sensitive photosites

AncoraSIR.com

Express w.r.t the Camera

Physical Meanings of Camera Pixels

- A **camera sensor** with a **W**× **H** grid of image pixels
  - The pixel coordinates (u, v)

Principal point in pixel coordinate

$$u = \frac{x}{\rho_w} + u_0,\ v = \frac{y}{\rho_h} + v_0$$

width and height of each pixel

$$\tilde{\boldsymbol{p}} = (u', v', w')$$

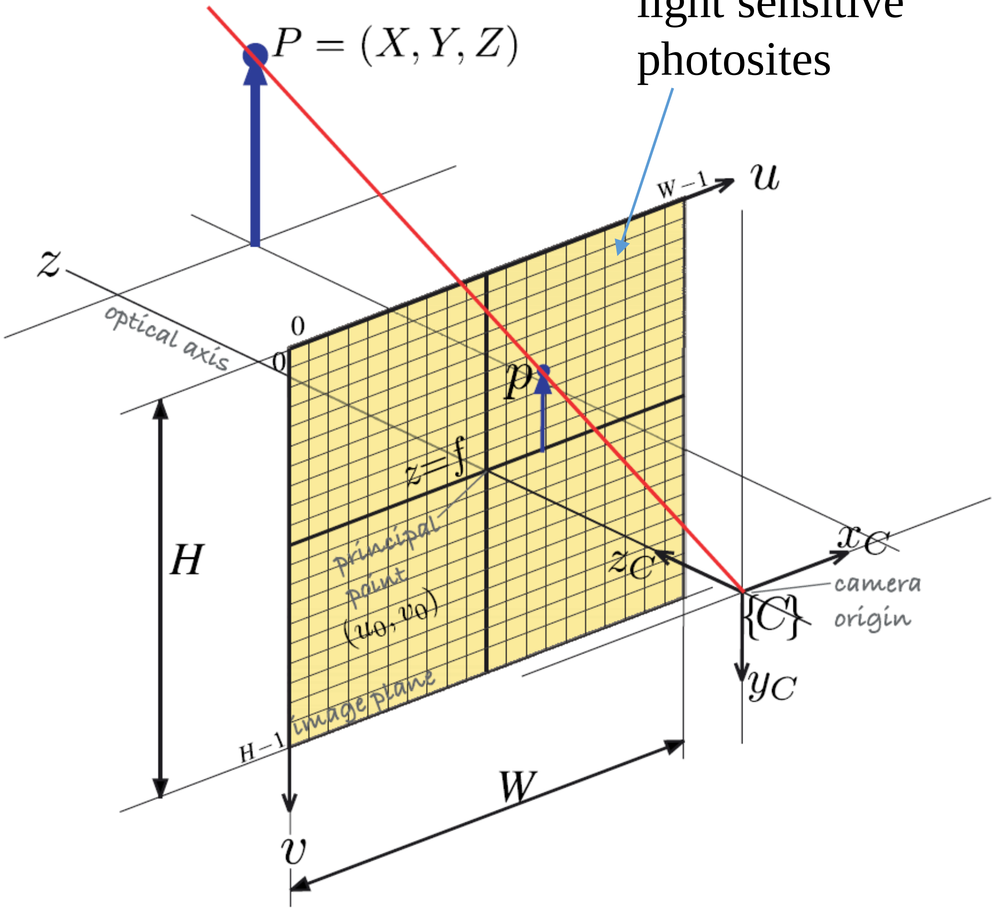
pixel coordinate

$$u = \frac{u'}{w'},\ v = \frac{v'}{w'}$$

camera (projection) matrix

$$\tilde{\boldsymbol{p}} = \underbrace{\begin{pmatrix} 1/\rho_w & 0 & u_0 \\ 0 & 1/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{K} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} {}^{C}\tilde{\boldsymbol{P}}$$

camera parameter matrix



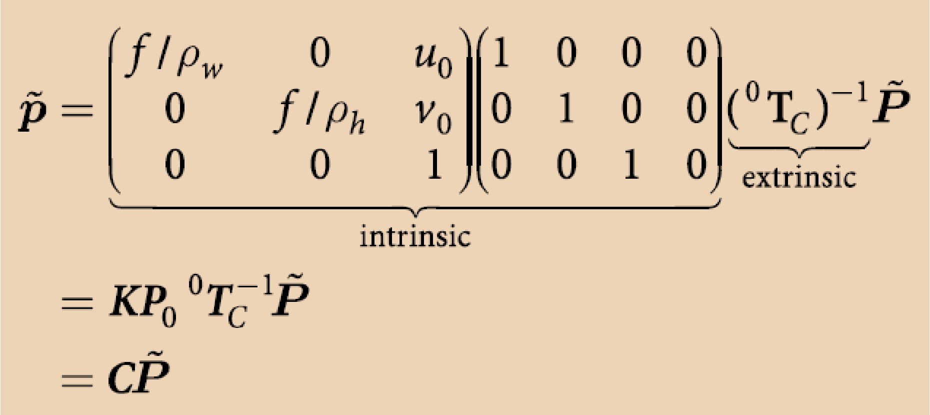
light sensitive photosites

AncoraSIR.com

Camera Projection In General Form

Still, something is missing

- The 3x4 Camera Calibration Matrix
  - Performs scaling, translation and perspective projection



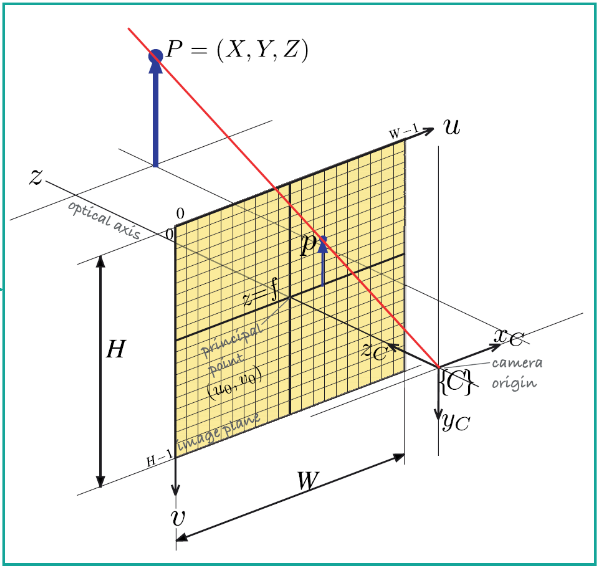
$$
\begin{aligned}
\tilde{\boldsymbol{p}} &= \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{({}^{0}\mathrm{T}_{C})^{-1}}_{\text{extrinsic}} \tilde{\boldsymbol{P}} \\
&= \boldsymbol{K}\boldsymbol{P}_0 \, {}^{0}\boldsymbol{T}_C^{-1} \tilde{\boldsymbol{P}} \\
&= \boldsymbol{C}\tilde{\boldsymbol{P}}
\end{aligned}
$$

Camera Projection In General Form

Still, something is missing

- The 3x4 Camera Calibration Matrix
  - Performs scaling, translation and perspective projection

$$
\begin{aligned}
\tilde{p} &= \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{({}^{0}\mathrm{T}_{C})^{-1}}_{\text{extrinsic}} \tilde{\boldsymbol{P}} \\
&= \boldsymbol{K}\boldsymbol{P}_0 \, {}^{0}\boldsymbol{T}_C^{-1} \tilde{\boldsymbol{P}} \\
&= \boldsymbol{C}\tilde{\boldsymbol{P}}
\end{aligned}
$$



AncoraSIR.com

Camera Projection In General Form

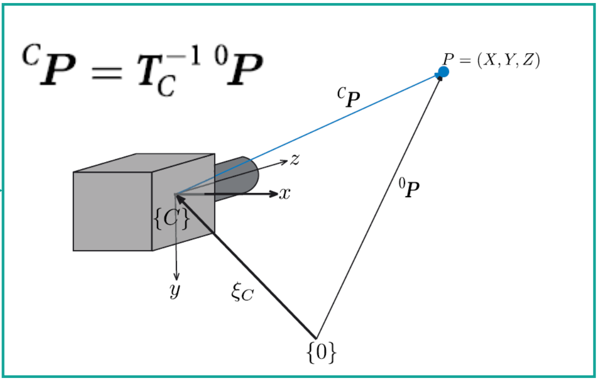
Still, something is missing

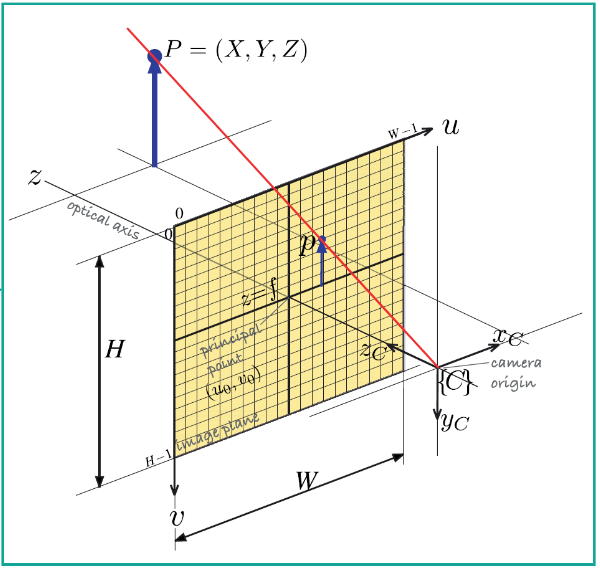
- The 3x4 Camera Calibration Matrix
  - Performs scaling, translation and perspective projection

$$\tilde{p} = \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{({}^{0}\mathrm{T}_C)^{-1}}_{\text{extrinsic}} \tilde{P}$$

$$= KP_0\, {}^{0}T_C^{-1}\tilde{P}$$

$$= C\tilde{P}$$





Camera Projection In General Form

Still, something is missing

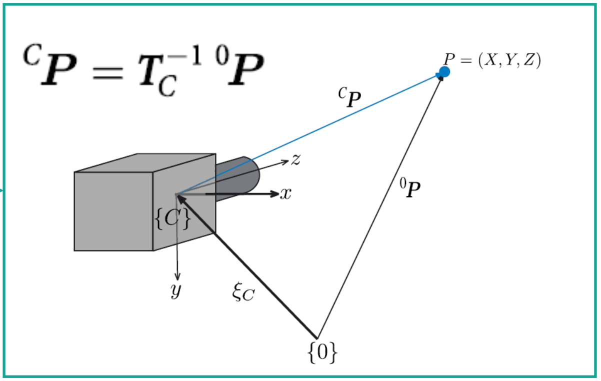
- The 3x4 Camera Calibration Matrix
  - Performs scaling, translation and perspective projection

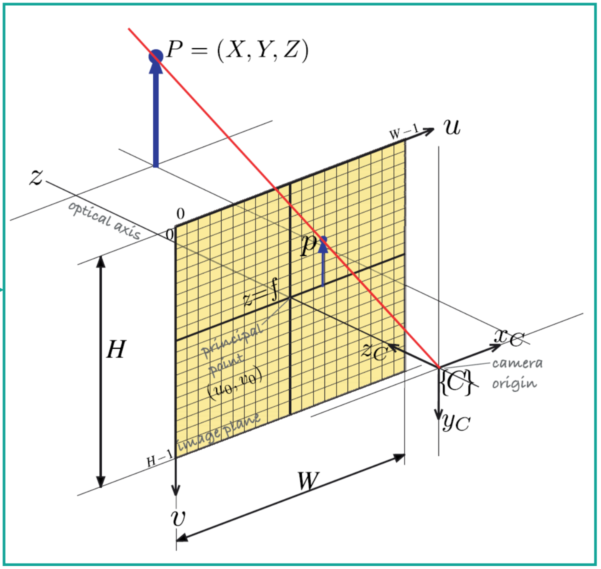
$$\tilde{p} = \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{\left({}^{0}\mathrm{T}_C\right)^{-1}}_{\text{extrinsic}} \tilde{P}$$

$$= KP_0\,{}^{0}T_C^{-1}\tilde{P}$$

$$= C\tilde{P}$$

$$p = \mathcal{P}(P, K, \xi_C)$$





Camera Projection In General Form

Still, something is missing

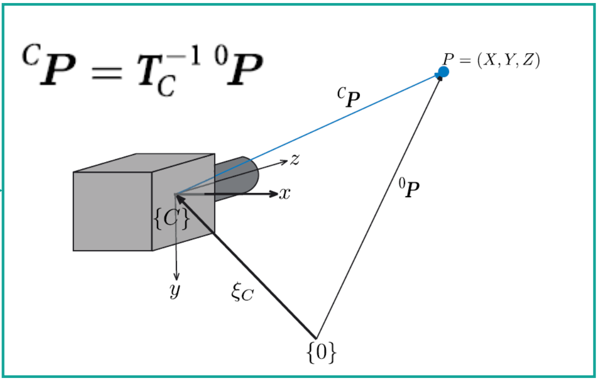
- The 3x4 Camera Calibration Matrix
  - Performs scaling, translation and perspective projection

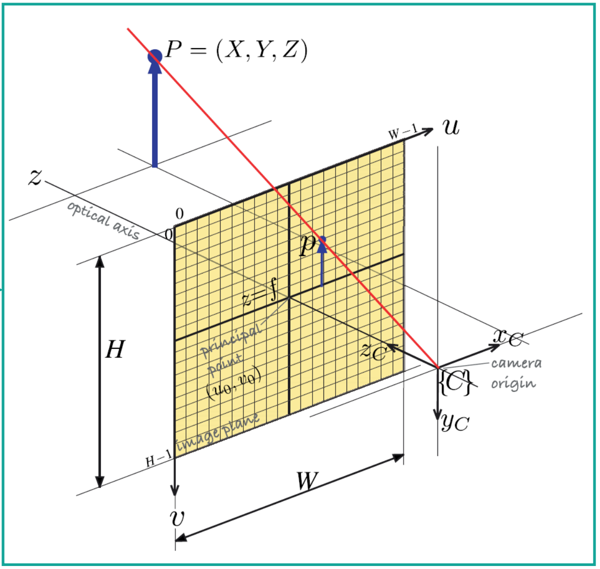
$$\tilde{p} = \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{\left({}^{0}\mathrm{T}_{C}\right)^{-1}}_{\text{extrinsic}} \tilde{P}$$

$$= K P_0\, {}^{0}T_C^{-1} \tilde{P}$$

$$= C\tilde{P}$$

$$p = \mathcal{P}(\boxed{P}, K, \xi_C)$$





Camera Projection In General Form

Still, something is missing

- The 3x4 Camera Calibration Matrix
  - Performs scaling, translation and perspective projection

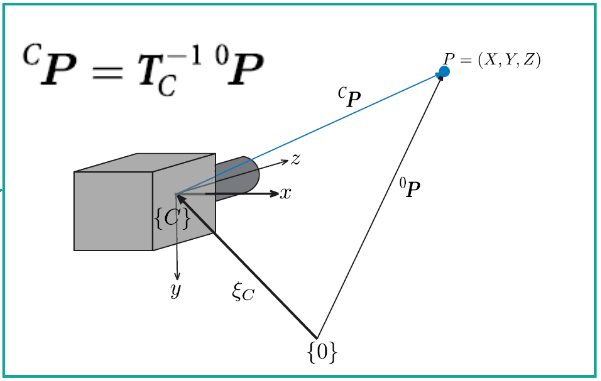
$$\tilde{\boldsymbol{p}} = \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{\left({}^{0}\mathbf{T}_C\right)^{-1}}_{\text{extrinsic}} \tilde{\boldsymbol{P}}$$

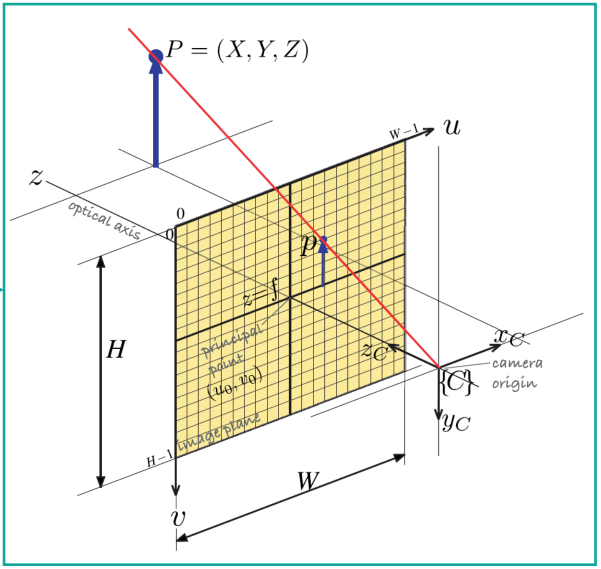
$$= \boldsymbol{K}\boldsymbol{P}_0\, {}^{0}\boldsymbol{T}_C^{-1} \tilde{\boldsymbol{P}}$$

$$= \boldsymbol{C}\tilde{\boldsymbol{P}}$$

$$\boldsymbol{p} = \mathcal{P}(\boxed{\boldsymbol{P}}, \boxed{\boldsymbol{K}}, \xi_C)$$

Camera parameter matrix: **5** parameters


${}^{C}\boldsymbol{P} = \boldsymbol{T}_C^{-1}\, {}^{0}\boldsymbol{P}$
$P = (X, Y, Z)$
${}^{C}P$
${}^{0}P$
z
x
y
$\{C\}$
ξ_C
$\{0\}$



Camera Projection In General Form

Still, something is missing

- The 3x4 Camera Calibration Matrix
    - Performs scaling, translation and perspective projection

$$\tilde{p} = \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{\left({}^{0}\mathrm{T}_C\right)^{-1}}_{\text{extrinsic}} \tilde{\boldsymbol{P}}$$

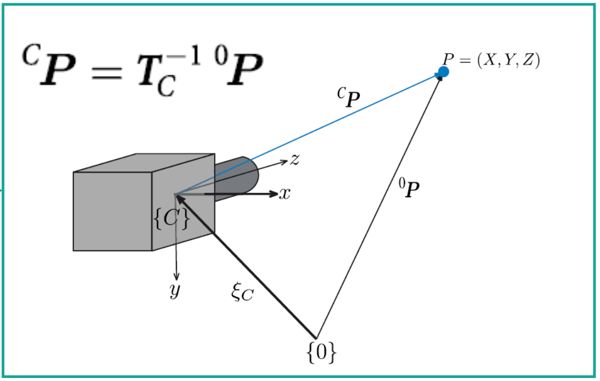
$$= \boldsymbol{K}\boldsymbol{P}_0 \, {}^{0}\boldsymbol{T}_C^{-1} \tilde{\boldsymbol{P}}$$

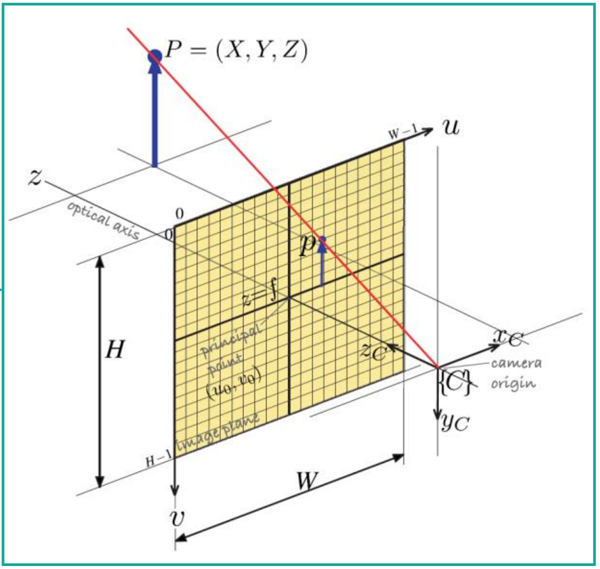
$$= \boldsymbol{C}\tilde{\boldsymbol{P}}$$

$$\boldsymbol{p} = \mathcal{P}(\boxed{\boldsymbol{P}}, \boxed{\boldsymbol{K}}, \boxed{\xi_C})$$

Camera pose: **6** parameters

Camera parameter matrix: **5** parameters


${}^{C}\boldsymbol{P} = \boldsymbol{T}_C^{-1}\,{}^{0}\boldsymbol{P}$
$P = (X, Y, Z)$
${}^{C}P$
z
x
${}^{0}P$
$\{C\}$
y
ξ_C
$\{0\}$



Camera Projection In General Form

Still, something is missing

- The 3x4 Camera Calibration Matrix
    - Performs scaling, translation and perspective projection

$$\tilde{p} = \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{\left({}^0\mathrm{T}_C\right)^{-1}}_{\text{extrinsic}} \tilde{P}$$

$$= KP_0\,{}^0T_C^{-1}\tilde{P}$$

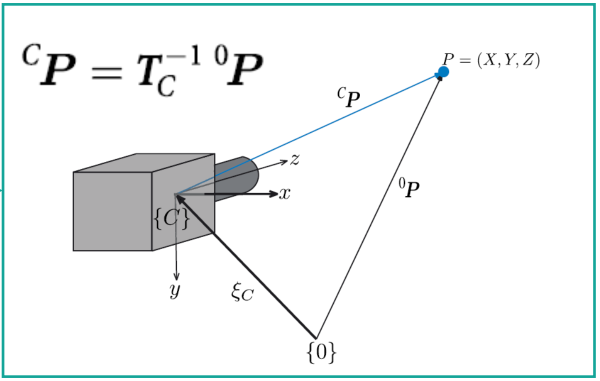
$$= C\tilde{P}$$

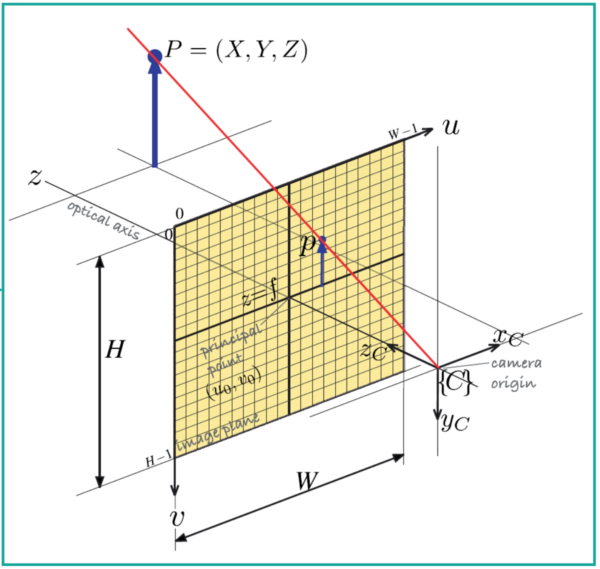
C is 3x4 with **12** elements

$$p = \mathcal{P}(\boxed{P}, \boxed{K}, \boxed{\xi_C})$$

Camera pose: **6** parameters

Camera parameter matrix: **5** parameters





AncoraSIR.com

Camera Projection In General Form

Still, something is missing

- The 3x4 Camera Calibration Matrix
  - Performs scaling, translation and perspective projection

$$\tilde{p} = \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{\left({}^{0}\mathrm{T}_{C}\right)^{-1}}_{\text{extrinsic}} \tilde{P}$$

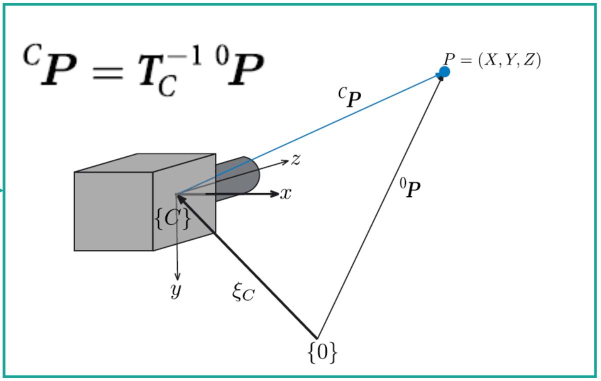
$$= KP_0 \, {}^{0}T_C^{-1} \tilde{P}$$

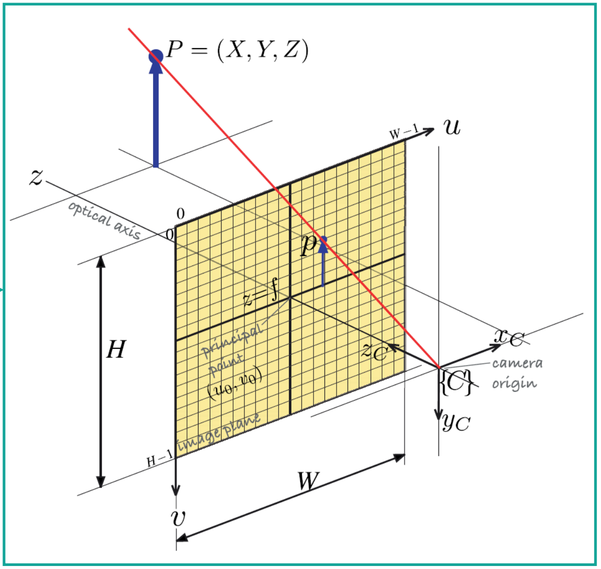
$$= C\tilde{P}$$

$$\tilde{p} = (CH^{-1})(H\tilde{P}) = C'\tilde{P}'$$

C is 3x4 with **12** elements

Unconstrained Overall Scale Factor





$$p = \mathcal{P}(\boxed{P}, \boxed{K}, \boxed{\xi_C})$$

Camera pose: **6** parameters

Camera parameter matrix: **5** parameters

Camera Projection In General Form

Still, something is missing

- The 3x4 Camera Calibration Matrix
  - Performs scaling, translation and perspective projection

$$\tilde{p} = \underbrace{\begin{pmatrix} f/\rho_w & 0 & u_0 \\ 0 & f/\rho_h & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\text{intrinsic}} \underbrace{({}^{0}\mathrm{T}_{C})^{-1}}_{\text{extrinsic}} \tilde{P}$$

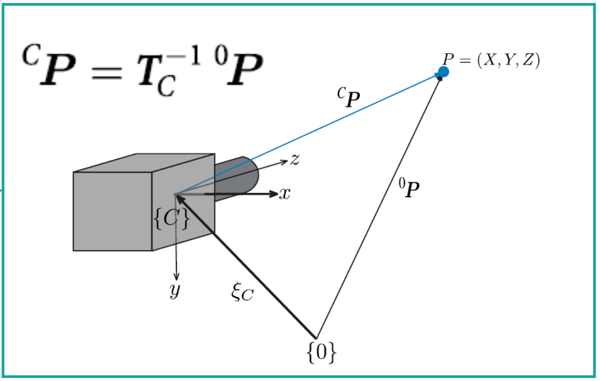
$$= KP_0\,{}^{0}T_C^{-1}\tilde{P}$$

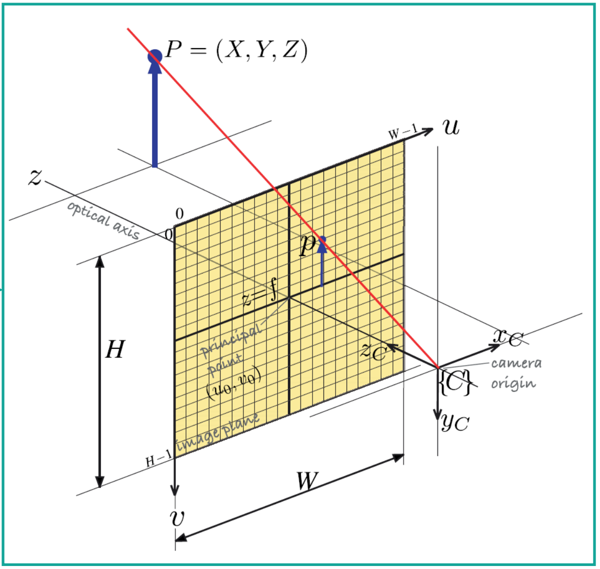
$$= C\tilde{P}$$

$$\tilde{p} = (CH^{-1})(H\tilde{P}) = C'\tilde{P}'$$

C is 3x4 with **12** elements

Unconstrained Overall Scale Factor





$$p = \mathcal{P}(P, K, \xi_C)$$

Camera pose: **6** parameters

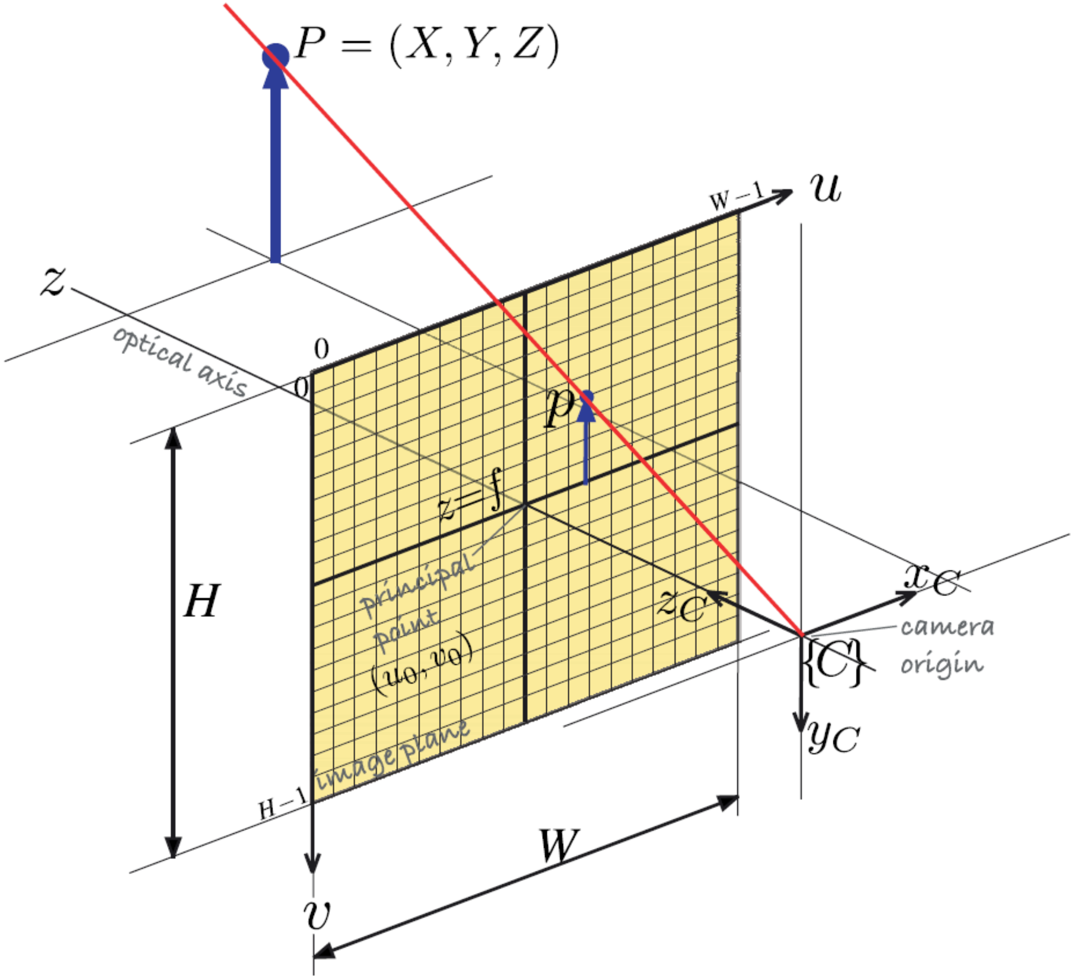
Camera parameter matrix: **5** parameters

It can only be solved if we have information about the camera or the 3D object

AncoraSIR.com

Camera Calibration

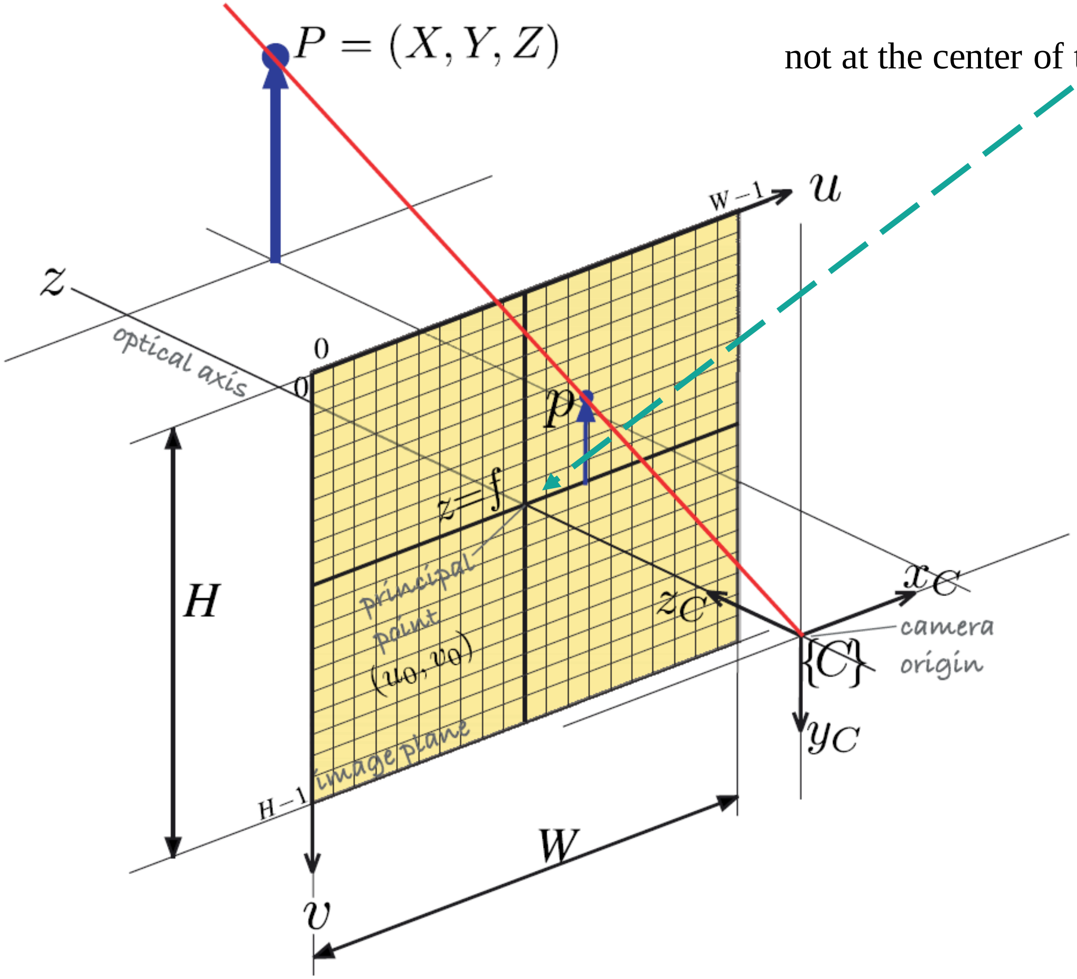
In general, the cameras are not made as modeled



AncoraSIR.com

Camera Calibration

In general, the cameras are not made as modeled

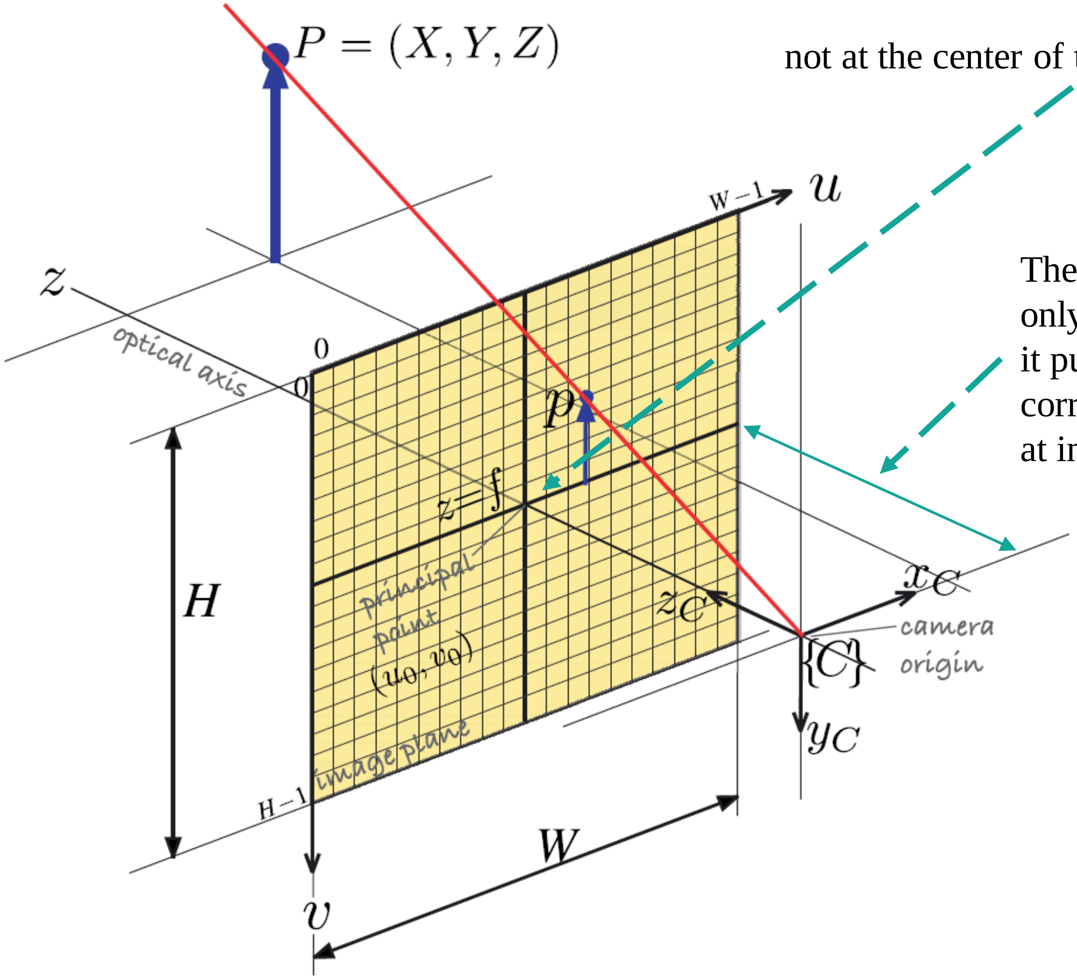


not at the center of the photosite array

AncoraSIR.com

Camera Calibration

In general, the cameras are not made as modeled



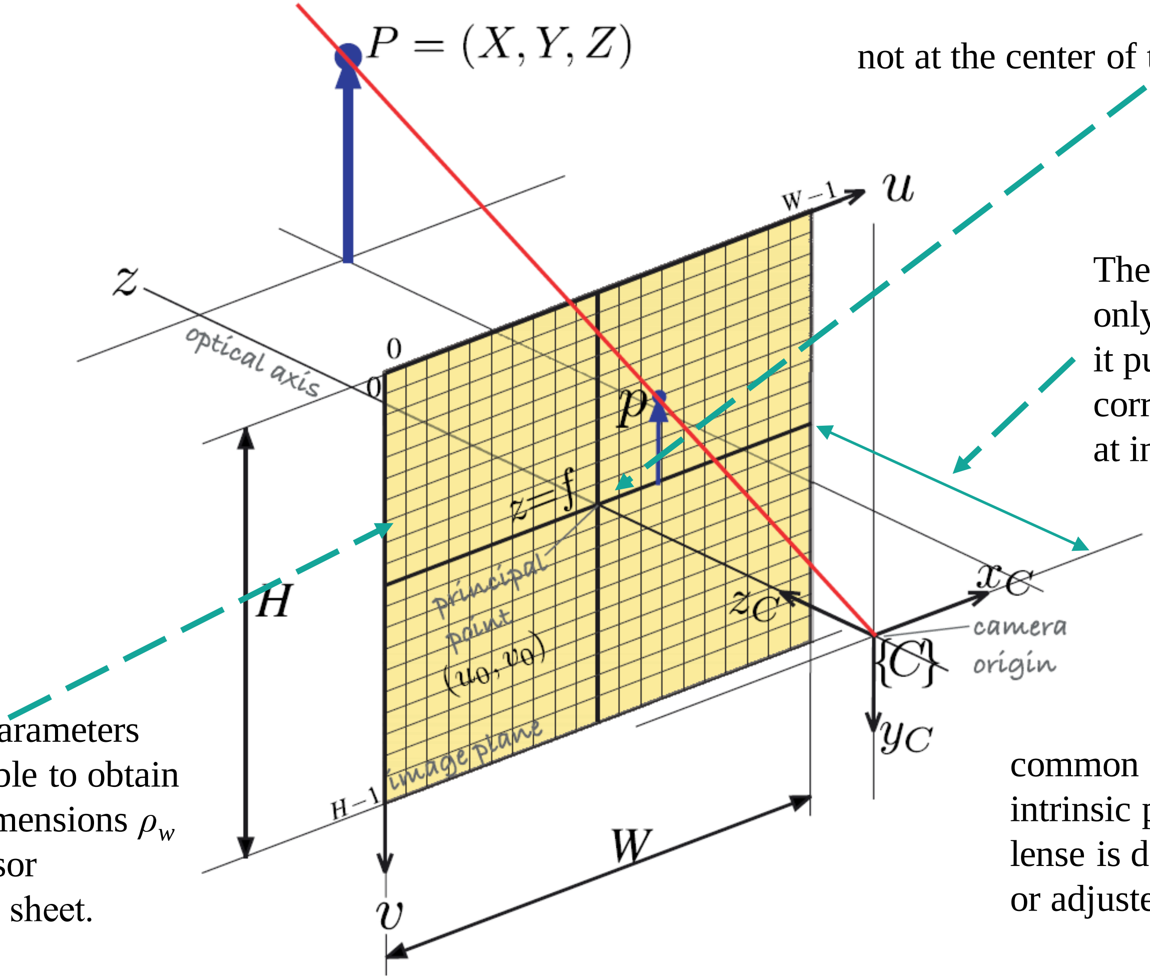
not at the center of the photosite array

The focal length of a lens is only accurate to 4% of what it purports to be, and is only correct if the lens is focused at infinity

AncoraSIR.com

Camera Calibration

In general, the cameras are not made as modeled



not at the center of the photosite array

The focal length of a lens is only accurate to 4% of what it purports to be, and is only correct if the lens is focused at infinity

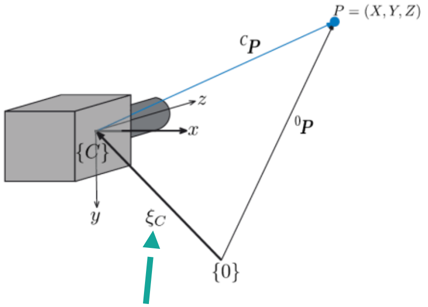
The only intrinsic parameters that it may be possible to obtain are the photosite dimensions ρ_w and ρ_h from the sensor manufacturer's data sheet.

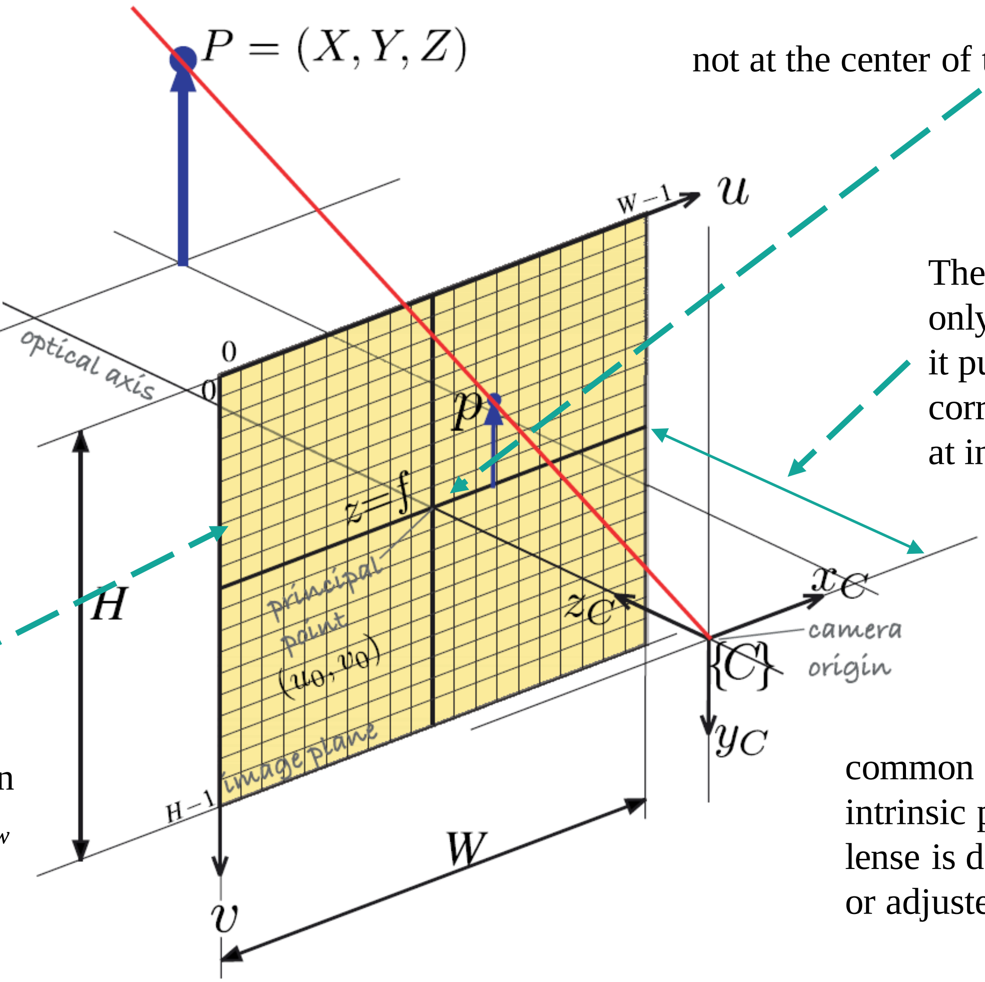
common experience that the intrinsic parameters change if a lense is detached and reattached, or adjusted for focus or aperture.

AncoraSIR.com

Camera Calibration

In general, the cameras are not made as modeled





not at the center of the photosite array

The focal length of a lens is only accurate to 4% of what it purports to be, and is only correct if the lens is focused at infinity

The extrinsic parameters, the camera's pose, raises the question of where exactly is the center point of the camera.

The only intrinsic parameters that it may be possible to obtain are the photosite dimensions ρ_w and ρ_h from the sensor manufacturer's data sheet.

common experience that the intrinsic parameters change if a lense is detached and reattached, or adjusted for focus or aperture.

AncoraSIR.com

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera’s intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

AncoraSIR.com

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera's intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

$$\tilde{p} = C\tilde{P}$$

AncoraSIR.com

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera's intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

$$\tilde{p} = C\tilde{P}$$

$$\tilde{p} = (u, v, 1)$$

$$u = \frac{u'}{w'}, v = \frac{v'}{w'}$$

AncoraSIR.com

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera’s intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

$$\tilde{p} = C\tilde{P} \xrightarrow[u = \frac{u'}{w'},\ v = \frac{v'}{w'}]{\tilde{p} = (u, v, 1)} \begin{array}{l} C_{11}X + C_{12}Y + C_{13}Z + C_{14} - C_{31}uX - C_{32}uY - C_{33}uZ - C_{34}u = 0 \\ C_{21}X + C_{22}Y + C_{23}Z + C_{24} - C_{31}vX - C_{32}vY - C_{33}vZ - C_{34}v = 0 \end{array}$$

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera’s intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

Disregard overall scaling, set to 1

$$\tilde{p} = C\tilde{P}$$

$$\tilde{p} = (u, v, 1)$$

$$u = \frac{u'}{w'}, v = \frac{v'}{w'}$$

$$C_{11}X + C_{12}Y + C_{13}Z + C_{14} - C_{31}uX - C_{32}uY - C_{33}uZ - \boxed{C_{34}}u = 0$$
$$C_{21}X + C_{22}Y + C_{23}Z + C_{24} - C_{31}vX - C_{32}vY - C_{33}vZ - C_{34}v = 0$$

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera’s intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

$$\tilde{p} = C\tilde{P}$$

$$\tilde{p} = (u, v, 1)$$

$$u = \frac{u'}{w'},\ v = \frac{v'}{w'}$$

Disregard overall scaling, set to 1

$$C_{11}X + C_{12}Y + C_{13}Z + C_{14} - C_{31}uX - C_{32}uY - C_{33}uZ - \boxed{C_{34}}u = 0$$

$$C_{21}X + C_{22}Y + C_{23}Z + C_{24} - C_{31}vX - C_{32}vY - C_{33}vZ - C_{34}v = 0$$

Increasing sampling for a solution

$$\begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 \\ & & & & & \vdots & & & & & \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -u_NX_N & -u_NY_N & -u_NZ_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -v_NX_N & -v_NY_N & -v_NZ_N \end{pmatrix} \begin{pmatrix} C_{11} \\ C_{12} \\ \vdots \\ C_{33} \end{pmatrix} = \begin{pmatrix} u_1 \\ v_1 \\ \vdots \\ u_N \\ v_N \end{pmatrix}$$

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera’s intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

Disregard overall scaling, set to 1

$$\tilde{p} = C\tilde{P}$$

$$\tilde{p} = (u, v, 1)$$

$$u = \frac{u'}{w'},\ v = \frac{v'}{w'}$$

$$C_{11}X + C_{12}Y + C_{13}Z + C_{14} - C_{31}uX - C_{32}uY - C_{33}uZ - \boxed{C_{34}}u = 0$$
$$C_{21}X + C_{22}Y + C_{23}Z + C_{24} - C_{31}vX - C_{32}vY - C_{33}vZ - C_{34}v = 0$$

Increasing sampling for a solution

11 unknows to be solved

$$\begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 \\ & & & & & \vdots & & & & & \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -u_NX_N & -u_NY_N & -u_NZ_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -v_NX_N & -v_NY_N & -v_NZ_N \end{pmatrix} \boxed{\begin{pmatrix} C_{11} \\ C_{12} \\ \vdots \\ C_{33} \end{pmatrix}} = \begin{pmatrix} u_1 \\ v_1 \\ \vdots \\ u_N \\ v_N \end{pmatrix}$$

AncoraSIR.com

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera's intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

$$\tilde{p} = C\tilde{P}$$

$$\tilde{p} = (u, v, 1)$$

$$u = \frac{u'}{w'}, \; v = \frac{v'}{w'}$$

Disregard overall scaling, set to 1

$$\begin{aligned} C_{11}X + C_{12}Y + C_{13}Z + C_{14} - C_{31}uX - C_{32}uY - C_{33}uZ - \boxed{C_{34}}u &= 0 \\ C_{21}X + C_{22}Y + C_{23}Z + C_{24} - C_{31}vX - C_{32}vY - C_{33}vZ - C_{34}v &= 0 \end{aligned}$$

Increasing sampling for a solution

11 unknows to be solved

$$\begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 \\ & & & & & \vdots & & & & & \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -u_NX_N & -u_NY_N & -u_NZ_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -v_NX_N & -v_NY_N & -v_NZ_N \end{pmatrix} \boxed{\begin{pmatrix} C_{11} \\ C_{12} \\ \vdots \\ C_{33} \end{pmatrix}} = \begin{pmatrix} u_1 \\ v_1 \\ \vdots \\ u_N \\ v_N \end{pmatrix}$$

$N \geq 6$ for *a* solution, but usually more are used to solve using *least square*

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera’s intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

Disregard overall scaling, set to 1

$$\tilde{p} = C\tilde{P}$$

$$\tilde{p} = (u, v, 1)$$

$$u = \frac{u'}{w'}, v = \frac{v'}{w'}$$

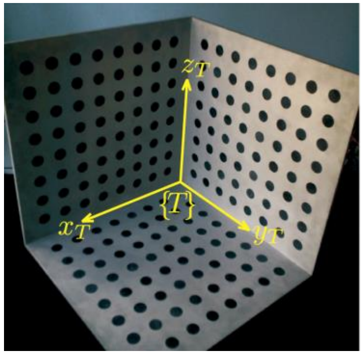
$$C_{11}X + C_{12}Y + C_{13}Z + C_{14} - C_{31}uX - C_{32}uY - C_{33}uZ - \boxed{C_{34}}u = 0$$

$$C_{21}X + C_{22}Y + C_{23}Z + C_{24} - C_{31}vX - C_{32}vY - C_{33}vZ - C_{34}v = 0$$

Increasing sampling for a solution

11 unknows to be solved

(u_i, v_i)



$$\begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 \\ & & & & & \vdots & & & & & \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -u_NX_N & -u_NY_N & -u_NZ_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -v_NX_N & -v_NY_N & -v_NZ_N \end{pmatrix} \boxed{\begin{pmatrix} C_{11} \\ C_{12} \\ \vdots \\ C_{33} \end{pmatrix}} = \begin{pmatrix} u_1 \\ v_1 \\ \vdots \\ u_N \\ v_N \end{pmatrix}$$

$N \geq 6$ for *a* solution, but usually more are used to solve using *least square*

Camera Calibration

Some are done before shipping, some are not, and some are provided with a software to do so

- The process of determining the camera's intrinsic parameters and the extrinsic parameters with respect to the world coordinate system

Disregard overall scaling, set to 1

$$\tilde{p} = (u, v, 1)$$

$$\tilde{p} = C\tilde{P}$$

$$u = \frac{u'}{w'}, v = \frac{v'}{w'}$$

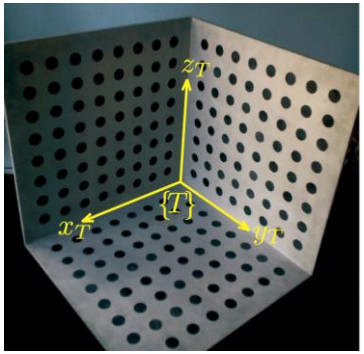
$$(u_i, v_i)$$

$$C_{11}X + C_{12}Y + C_{13}Z + C_{14} - C_{31}uX - C_{32}uY - C_{33}uZ - \boxed{C_{34}}u = 0$$

$$C_{21}X + C_{22}Y + C_{23}Z + C_{24} - C_{31}vX - C_{32}vY - C_{33}vZ - C_{34}v = 0$$

Increasing sampling for a solution

11 unknows to be solved



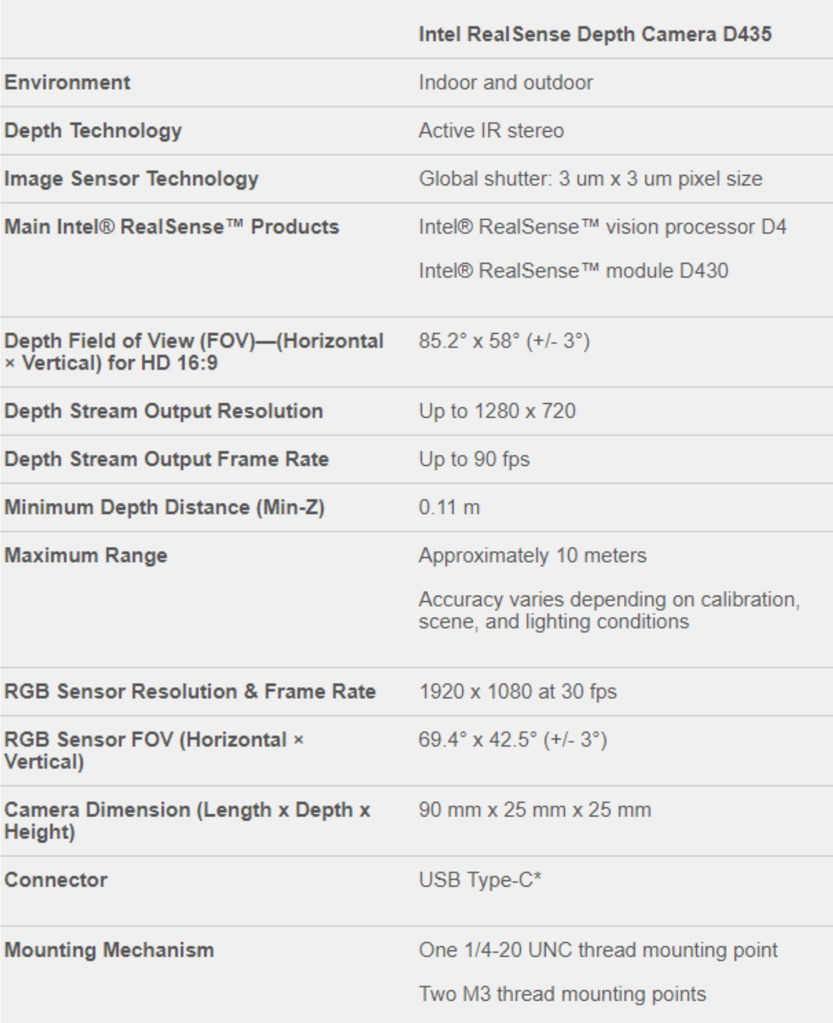
$$\begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 \\ & & & & & \vdots & & & & & \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -u_NX_N & -u_NY_N & -u_NZ_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -v_NX_N & -v_NY_N & -v_NZ_N \end{pmatrix} \boxed{\begin{pmatrix} C_{11} \\ C_{12} \\ \vdots \\ C_{33} \end{pmatrix}} = \begin{pmatrix} u_1 \\ v_1 \\ \vdots \\ u_N \\ v_N \end{pmatrix}$$

$N \geq 6$ for *a* solution, but usually more are used to solve using *least square*

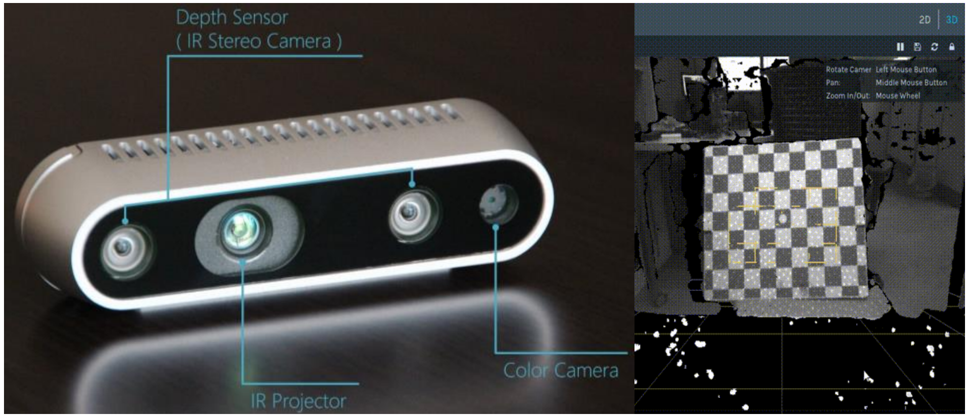
What if the points are coplanar?

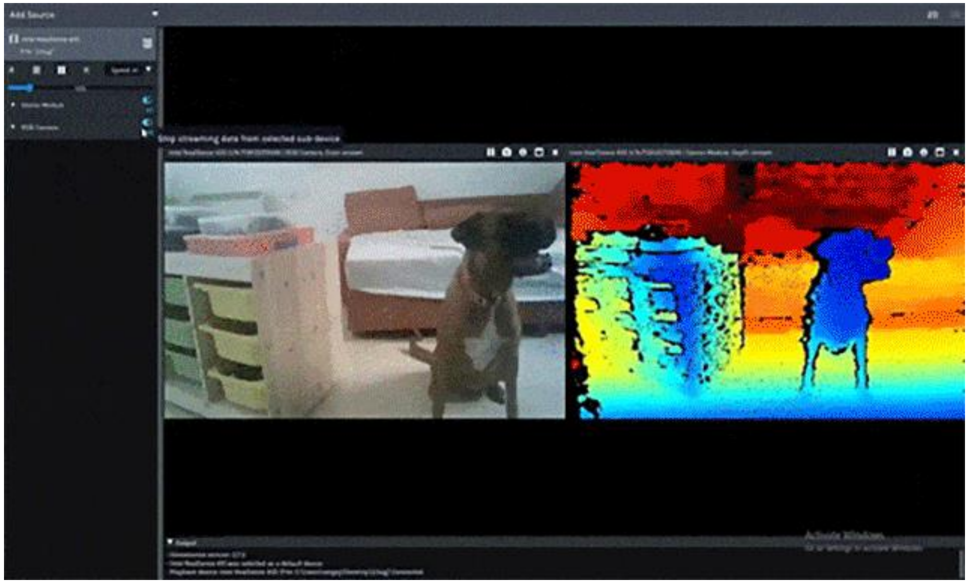
About Intel Realsense D435

Entry level stereo depth sensor with abundant resources at a low cost



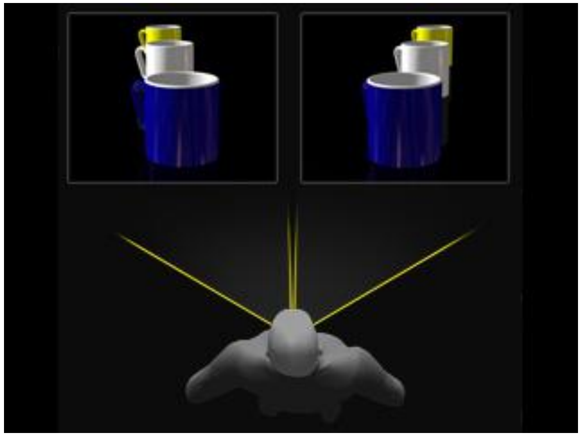
	Intel RealSense Depth Camera D435
Environment	Indoor and outdoor
Depth Technology	Active IR stereo
Image Sensor Technology	Global shutter: 3 um x 3 um pixel size
Main Intel® RealSense™ Products	Intel® RealSense™ vision processor D4 Intel® RealSense™ module D430
Depth Field of View (FOV)—(Horizontal × Vertical) for HD 16:9	85.2° x 58° (+/- 3°)
Depth Stream Output Resolution	Up to 1280 x 720
Depth Stream Output Frame Rate	Up to 90 fps
Minimum Depth Distance (Min-Z)	0.11 m
Maximum Range	Approximately 10 meters Accuracy varies depending on calibration, scene, and lighting conditions
RGB Sensor Resolution & Frame Rate	1920 x 1080 at 30 fps
RGB Sensor FOV (Horizontal × Vertical)	69.4° x 42.5° (+/- 3°)
Camera Dimension (Length x Depth x Height)	90 mm x 25 mm x 25 mm
Connector	USB Type-C*
Mounting Mechanism	One 1/4-20 UNC thread mounting point Two M3 thread mounting points





Stereo Vision

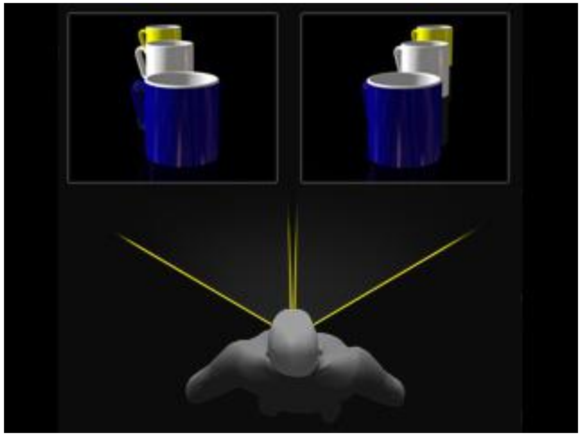
Triangulation Principle

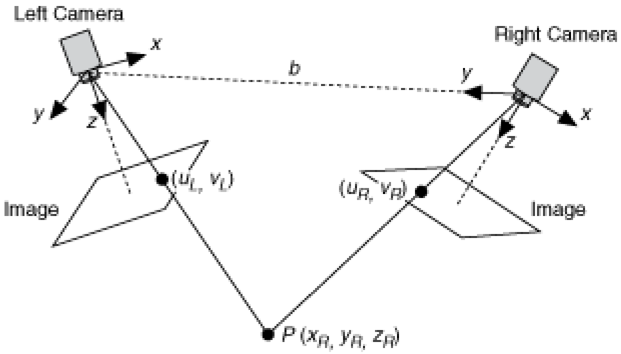


AncoraSIR.com

Stereo Vision

Triangulation Principle

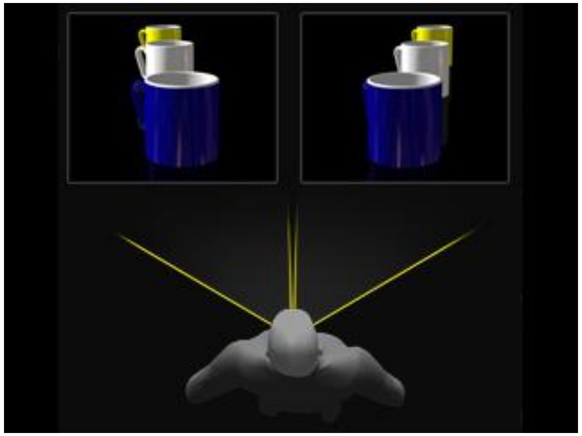


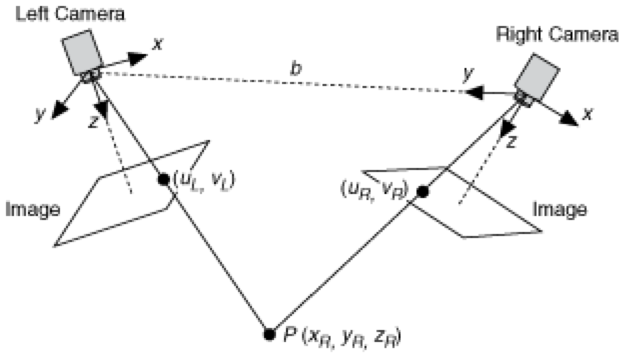


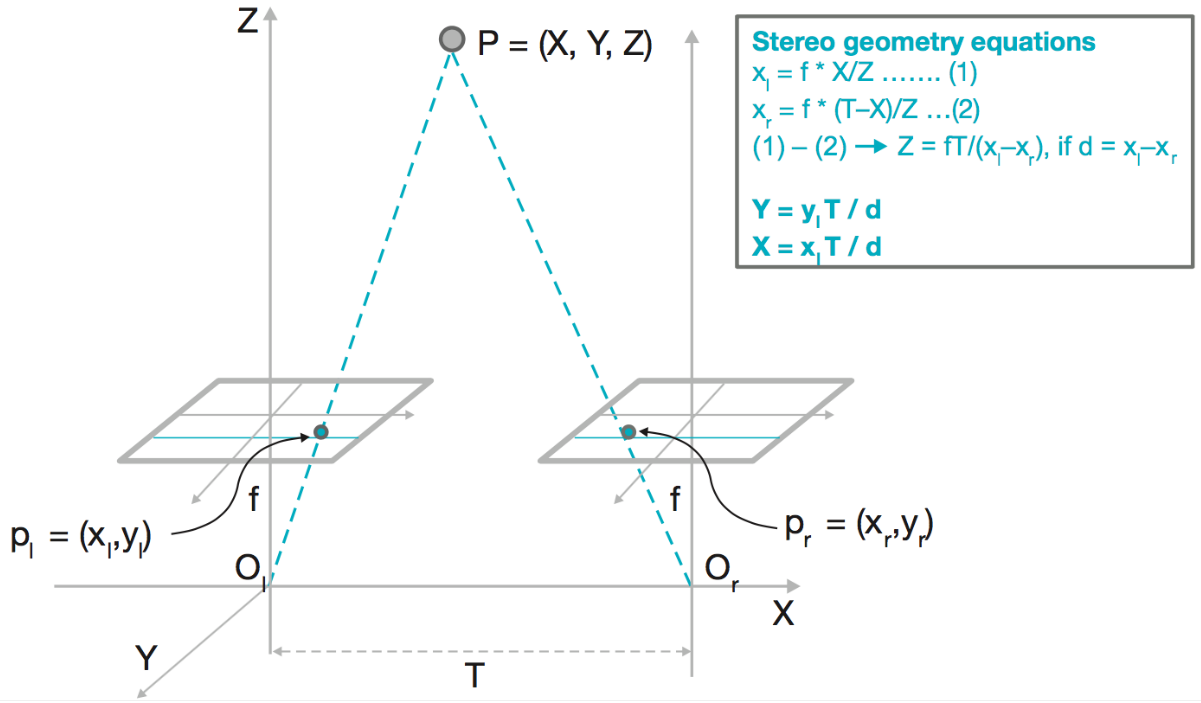
AncoraSIR.com

Stereo Vision

Triangulation Principle







Stereo geometry equations

$x_l = f * X/Z$ .... (1)

$x_r = f * (T–X)/Z$ ...(2)

(1) – (2) → $Z = fT/(x_l–x_r)$, if $d = x_l–x_r$

$Y = y_l T / d$

$X = x_l T / d$

AncoraSIR.com

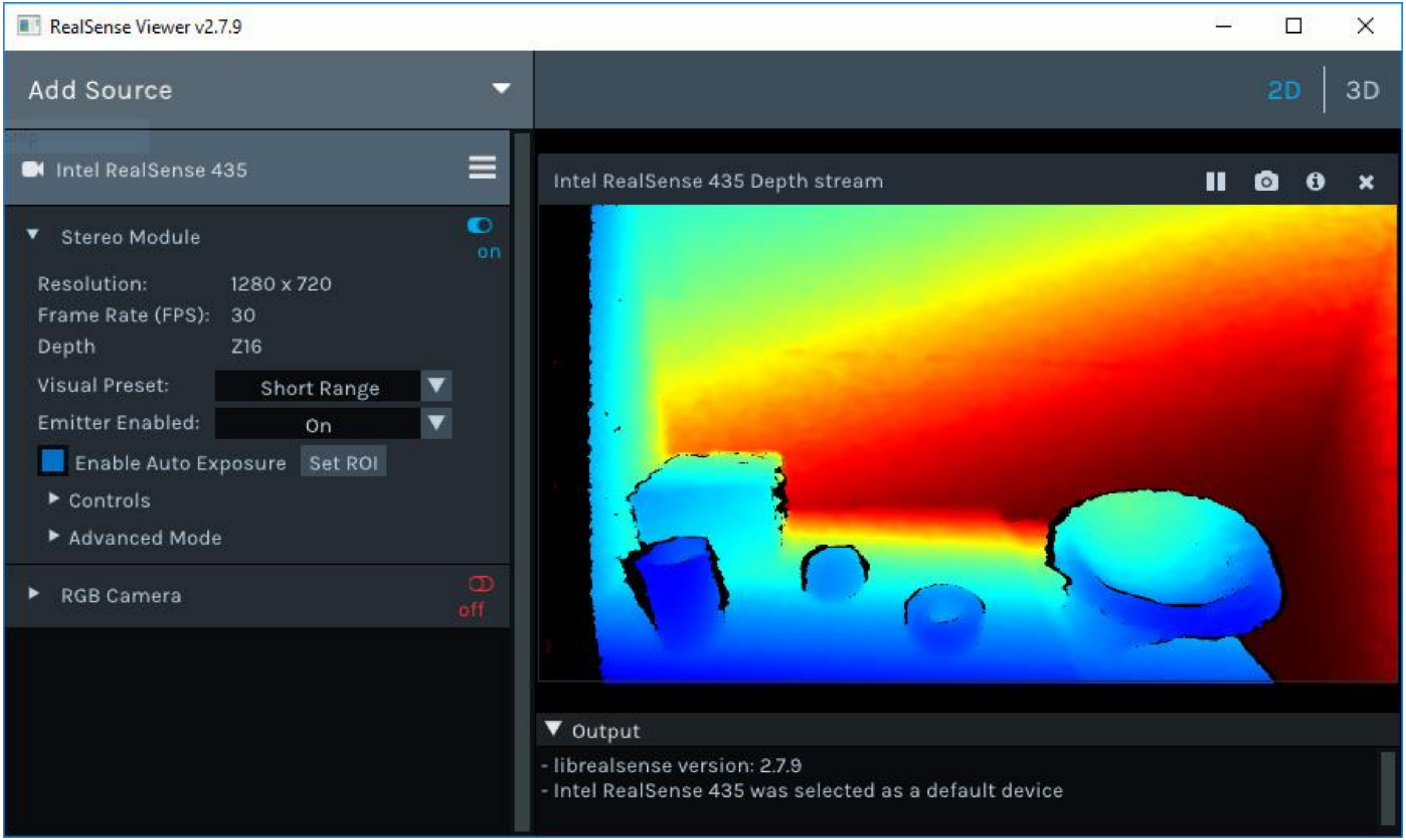
Interfacing Realsense with ROS

Installation

1. Connect the USB Type-C to your host PC.
2. Install the Intel® RealSense™ SDK V2.16(Please use the guidance here: https://realsense.intel.com/sdk-2/#install)
    1. Prep steps as instructed in the website
    2. sudo apt-get install librealsense2=2.16.0-0~realsense0.85
    3. sudo apt-get install librealsense2-utils=2.16.0-0~realsense0.85
    4. sudo apt-get install librealsense2-dev=2.16.0-0~realsense0.85
    5. sudo apt-get install librealsense2-dbg=2.16.0-0~realsense0.85
3. Run the Intel® RealSense™ Viewer (firmware upgrade may be needed if you encounter version error)

AncoraSIR.com

Interfacing Realsense with ROS



Stereo Module on

Resolution: 1280 x 720

Frame Rate (FPS): 30

Depth Z16

Visual Preset: Short Range

Emitter Enabled: On

■ Enable Auto Exposure Set ROI

RGB Camera off

Output

```
- librealsense version: 2.7.9
- Intel RealSense 435 was selected as a default device
```

Interfacing Realsense with ROS

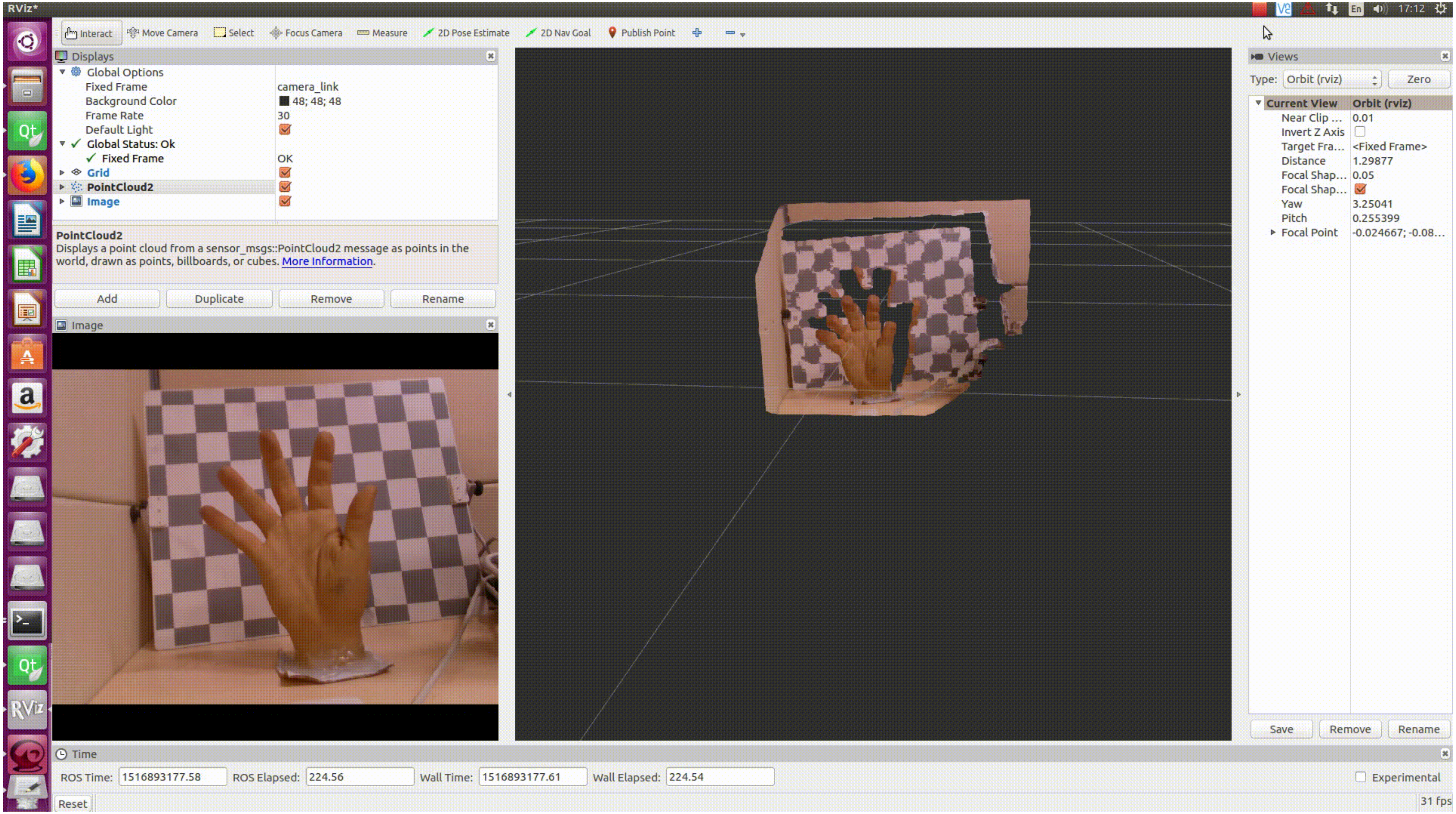
Realsense ROS Package

- Installation
  1. cd ~/catkin_ws/src/
  2. git clone https://github.com/intel-ros/realsense.git
  3. git checkout tag/2.1.0
  4. catkin build & source ~/.bashrc
- Launch: roslaunch realsense2_camera rs_rgbd.launch

AncoraSIR.com

Interfacing Realsense with ROS

Point cloud in Rviz



Displays

Global Options	
Fixed Frame	camera_link
Background Color	■ 48; 48; 48
Frame Rate	30
Default Light	☑
✓ Global Status: Ok	
✓ Fixed Frame	OK
Grid	☑
PointCloud2	☑
Image	☑

PointCloud2
Displays a point cloud from a sensor_msgs::PointCloud2 message as points in the world, drawn as points, billboards, or cubes. More Information.

Image

Views

Type: Orbit (rviz)

Current View	Orbit (rviz)
Near Clip ...	0.01
Invert Z Axis	□
Target Fra...	<Fixed Frame>
Distance	1.29877
Focal Shap...	0.05
Focal Shap...	☑
Yaw	3.25041
Pitch	0.255399
Focal Point	-0.024667; -0.08...

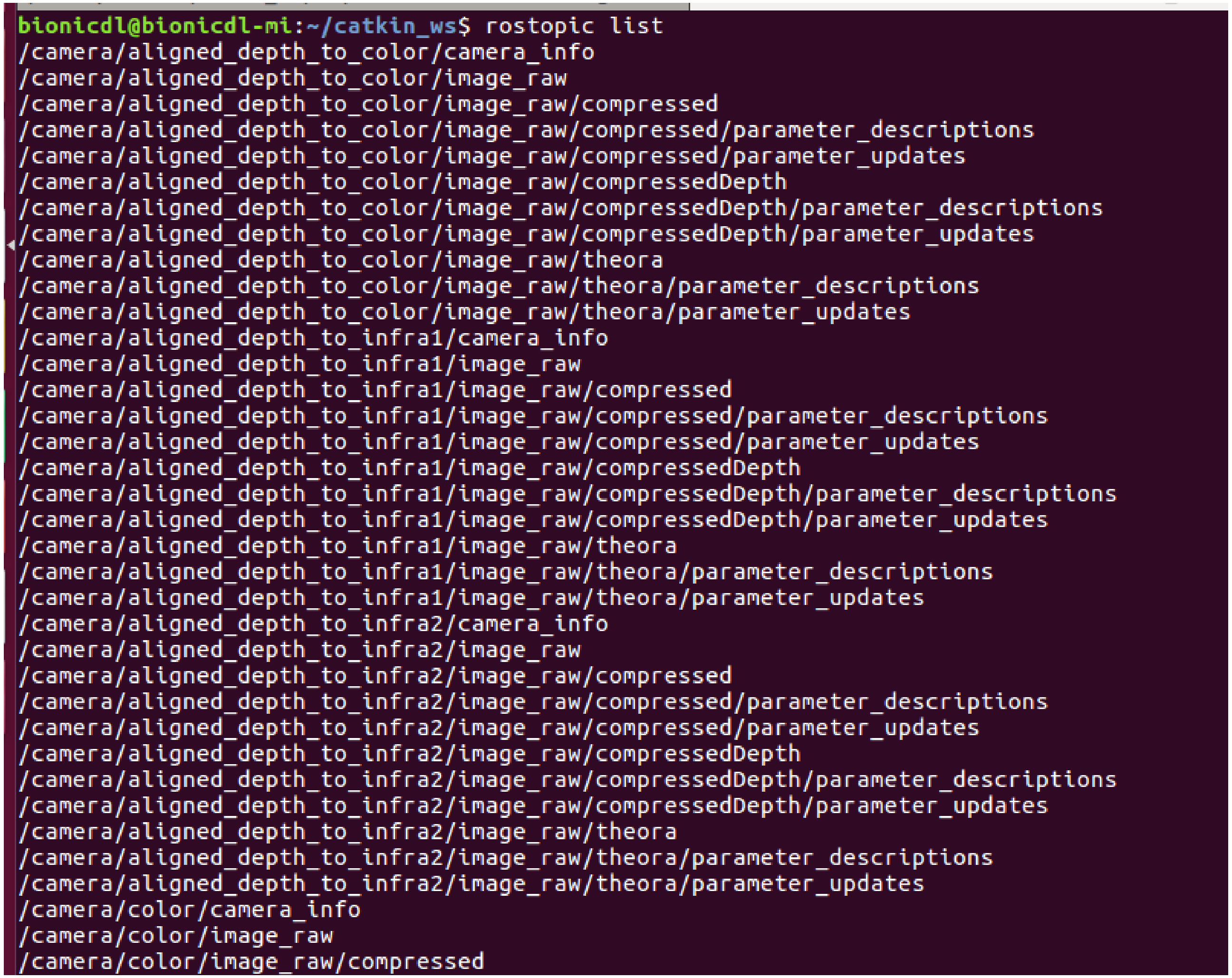
Time

ROS Time: 1516893177.58 ROS Elapsed: 224.56 Wall Time: 1516893177.61 Wall Elapsed: 224.54 □ Experimental

31 fps

Interfacing Realsense with ROS

ROS topics published

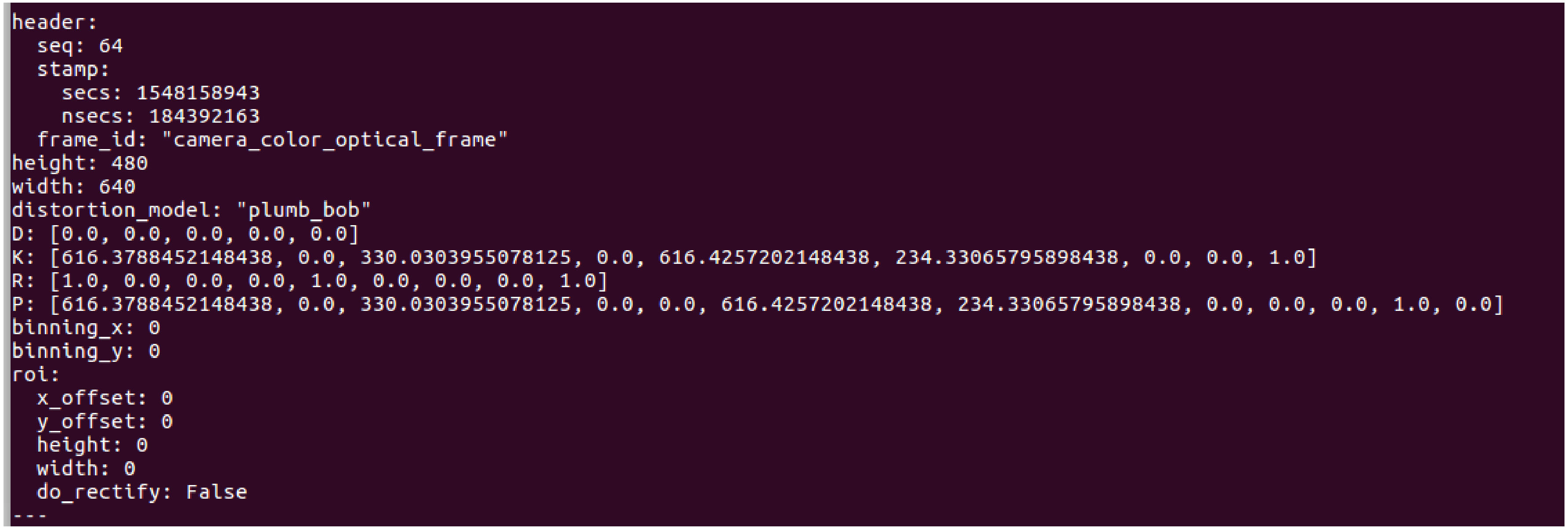


```
bionicdl@bionicdl-mi:~/catkin_ws$ rostopic list
/camera/aligned_depth_to_color/camera_info
/camera/aligned_depth_to_color/image_raw
/camera/aligned_depth_to_color/image_raw/compressed
/camera/aligned_depth_to_color/image_raw/compressed/parameter_descriptions
/camera/aligned_depth_to_color/image_raw/compressed/parameter_updates
/camera/aligned_depth_to_color/image_raw/compressedDepth
/camera/aligned_depth_to_color/image_raw/compressedDepth/parameter_descriptions
/camera/aligned_depth_to_color/image_raw/compressedDepth/parameter_updates
/camera/aligned_depth_to_color/image_raw/theora
/camera/aligned_depth_to_color/image_raw/theora/parameter_descriptions
/camera/aligned_depth_to_color/image_raw/theora/parameter_updates
/camera/aligned_depth_to_infra1/camera_info
/camera/aligned_depth_to_infra1/image_raw
/camera/aligned_depth_to_infra1/image_raw/compressed
/camera/aligned_depth_to_infra1/image_raw/compressed/parameter_descriptions
/camera/aligned_depth_to_infra1/image_raw/compressed/parameter_updates
/camera/aligned_depth_to_infra1/image_raw/compressedDepth
/camera/aligned_depth_to_infra1/image_raw/compressedDepth/parameter_descriptions
/camera/aligned_depth_to_infra1/image_raw/compressedDepth/parameter_updates
/camera/aligned_depth_to_infra1/image_raw/theora
/camera/aligned_depth_to_infra1/image_raw/theora/parameter_descriptions
/camera/aligned_depth_to_infra1/image_raw/theora/parameter_updates
/camera/aligned_depth_to_infra2/camera_info
/camera/aligned_depth_to_infra2/image_raw
/camera/aligned_depth_to_infra2/image_raw/compressed
/camera/aligned_depth_to_infra2/image_raw/compressed/parameter_descriptions
/camera/aligned_depth_to_infra2/image_raw/compressed/parameter_updates
/camera/aligned_depth_to_infra2/image_raw/compressedDepth
/camera/aligned_depth_to_infra2/image_raw/compressedDepth/parameter_descriptions
/camera/aligned_depth_to_infra2/image_raw/compressedDepth/parameter_updates
/camera/aligned_depth_to_infra2/image_raw/theora
/camera/aligned_depth_to_infra2/image_raw/theora/parameter_descriptions
/camera/aligned_depth_to_infra2/image_raw/theora/parameter_updates
/camera/color/camera_info
/camera/color/image_raw
/camera/color/image_raw/compressed
```

Interfacing Realsense with ROS

ROS topics published

- $rostopic ehco /camera/color/camera_info

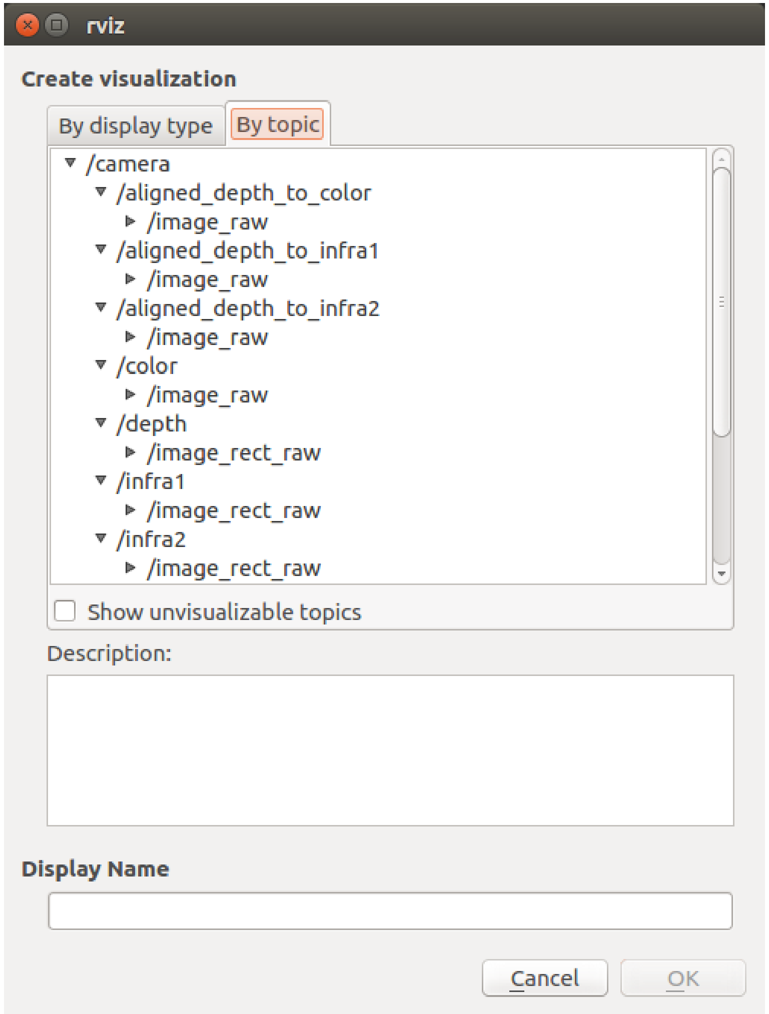


```
header:
  seq: 64
  stamp:
    secs: 1548158943
    nsecs: 184392163
  frame_id: "camera_color_optical_frame"
height: 480
width: 640
distortion_model: "plumb_bob"
D: [0.0, 0.0, 0.0, 0.0, 0.0]
K: [616.3788452148438, 0.0, 330.0303955078125, 0.0, 616.4257202148438, 234.33065795898438, 0.0, 0.0, 1.0]
R: [1.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0, 0.0, 1.0]
P: [616.3788452148438, 0.0, 330.0303955078125, 0.0, 0.0, 616.4257202148438, 234.33065795898438, 0.0, 0.0, 0.0, 1.0, 0.0]
binning_x: 0
binning_y: 0
roi:
  x_offset: 0
  y_offset: 0
  height: 0
  width: 0
  do_rectify: False
---
```

Interfacing Realsense with ROS

subscribe to ROS topic of relevant image

```
# TODO: Create Subscribers

depth_sub = message_filters.Subscriber("/camera/aligned_depth_to_color/image_raw", Image)


rgb_sub = message_filters.Subscriber("/camera/color/image_raw", Image)


ts = message_filters.TimeSynchronizer([depth_sub, rgb_sub], 1)

ts.registerCallback(callback)
```



```
▼ /camera
  ▼ /aligned_depth_to_color
    ▶ /image_raw
  ▼ /aligned_depth_to_infra1
    ▶ /image_raw
  ▼ /aligned_depth_to_infra2
    ▶ /image_raw
  ▼ /color
    ▶ /image_raw
  ▼ /depth
    ▶ /image_rect_raw
  ▼ /infra1
    ▶ /image_rect_raw
  ▼ /infra2
    ▶ /image_rect_raw
```

□ Show unvisualizable topics

Description:

Display Name

AncoraSIR.com

Homework

- Write a perception ros node subscribing to ros topic of relevant image.
- Add a mover function to move the robot arm in the perception ros node.
- Prepare for the next lab session:
  - Install opencv and pcl

Thank you!

Prof. Song Chaoyang

- Dr. Wan Fang (sophie.fwan@hotmail.com)

AncoraSIR.com