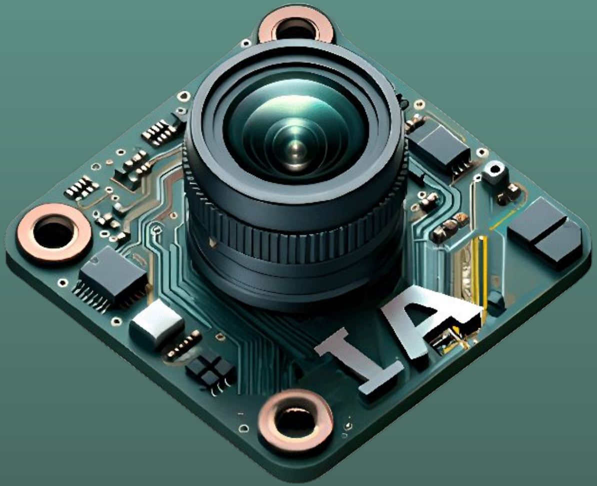
Universidad Industrial de Santander



Semillero de Investigación
“Hands - on”
Computer Vision

Hands-on Computer Vision

SESIÓN 7: SEGMENTACIÓN

Abril 29, 2024 | 4 pm - 6 pm
CENTIC Sala 1-2

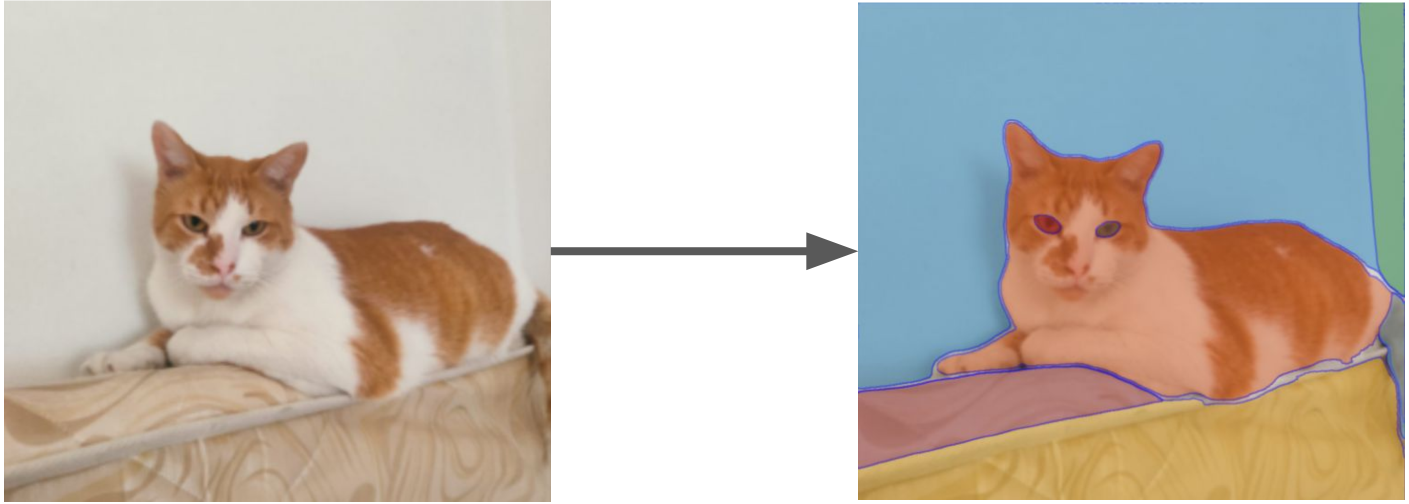
Contenidos del dia

1. ¿que es la segmentación?
2. Aplicaciones
3. Material segmentation
4. Tecnicas
5. tecnicas

1. ¿Que es la segmentación?

Segmentación

El objetivo de la segmentación de imágenes es **clasificar los píxeles** en **regiones destacadas** de la imagen.



Segmentación

Input

segmented

1: Person
2: Purse
3: Plants/Grass
4: Sidewalk
5: Building/Structures

3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	1	1	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	1	1	1	1	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	1	1	3	3	3	5	5	5	5	5	5	5
5	5	3	3	3	3	1	1	3	3	5	5	5	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	4	4	5	5	5
4	4	4	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	1	4	4	4	4	4	4	4

Semantic Labels

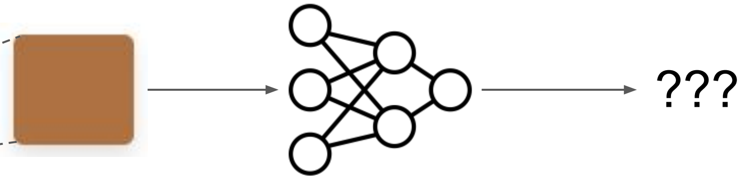
Segmentación

0: Background/Unknown
1: Person
2: Purse
3: Plants/Grass
4: Sidewalk
5: Building/Structures

Segmentación

Cómo clasificar un pixel?

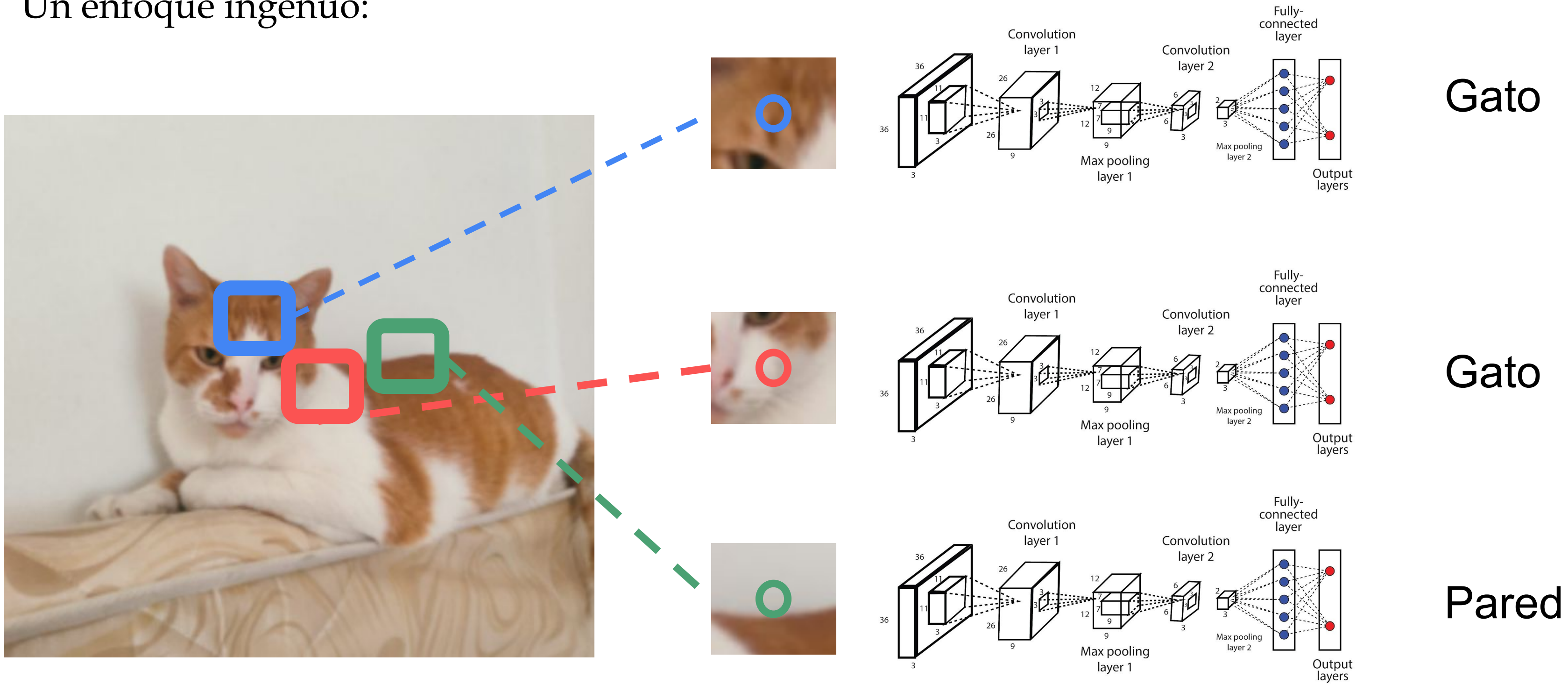
imagen



Imposible clasificar sin contexto!

Segmentación

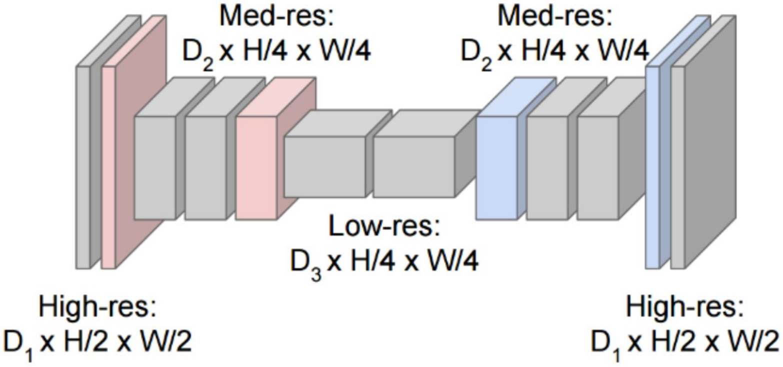
Un enfoque ingenuo:



Segmentación

Entrada

HxWx3



Salida

HxWxC

Segmentación

Volvamos al mismo ejemplo

Input

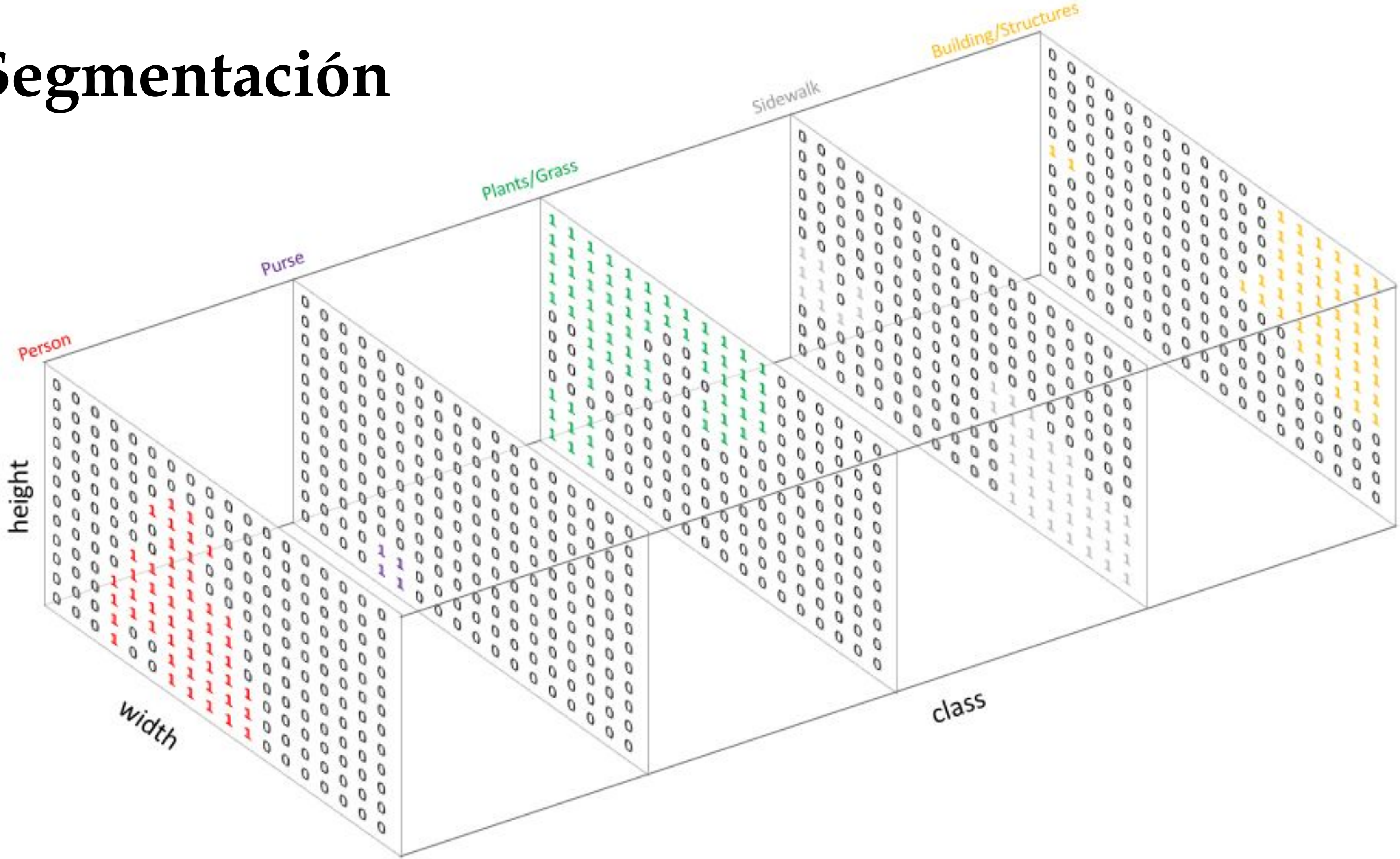
segmented

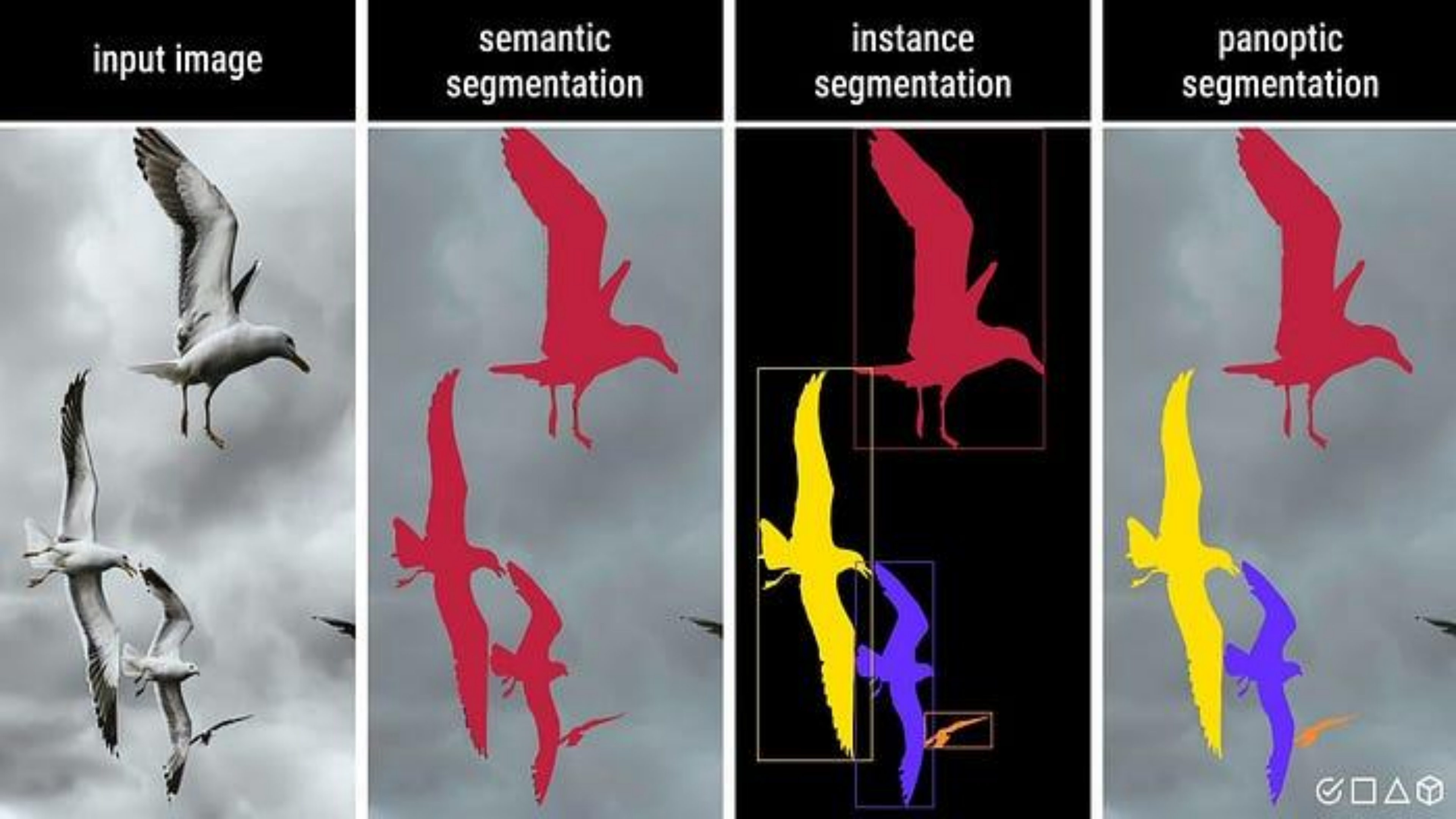
1: Person
2: Purse
3: Plants/Grass
4: Sidewalk
5: Building/Structures

3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	1	1	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	1	1	1	1	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	1	1	3	3	3	5	5	5	5	5	5	5
5	5	3	3	3	3	1	1	3	3	5	5	5	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	4	4	5	5	5
4	4	4	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	1	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	1	4	4	4	4	4	4	4

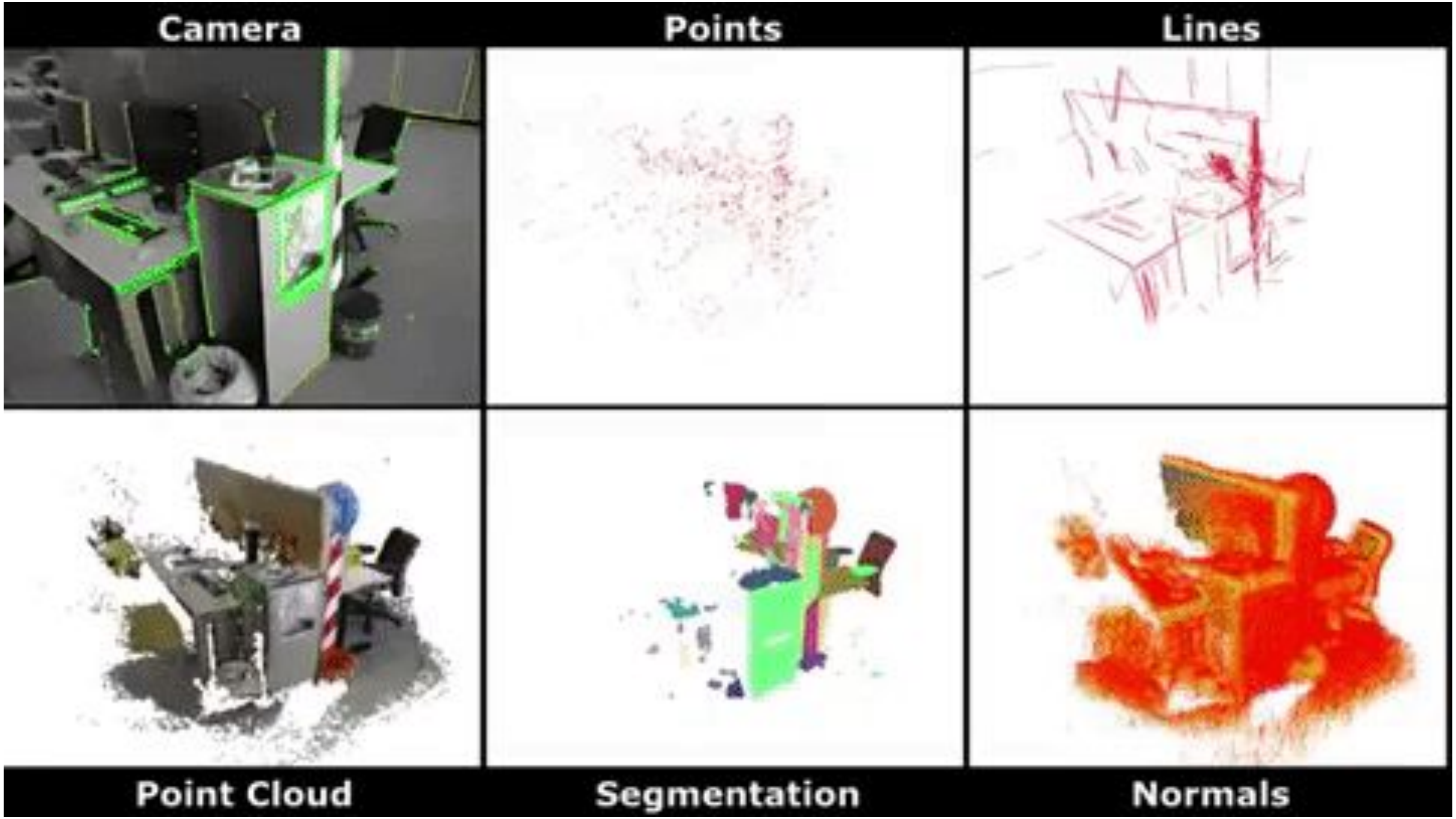
Semantic Labels

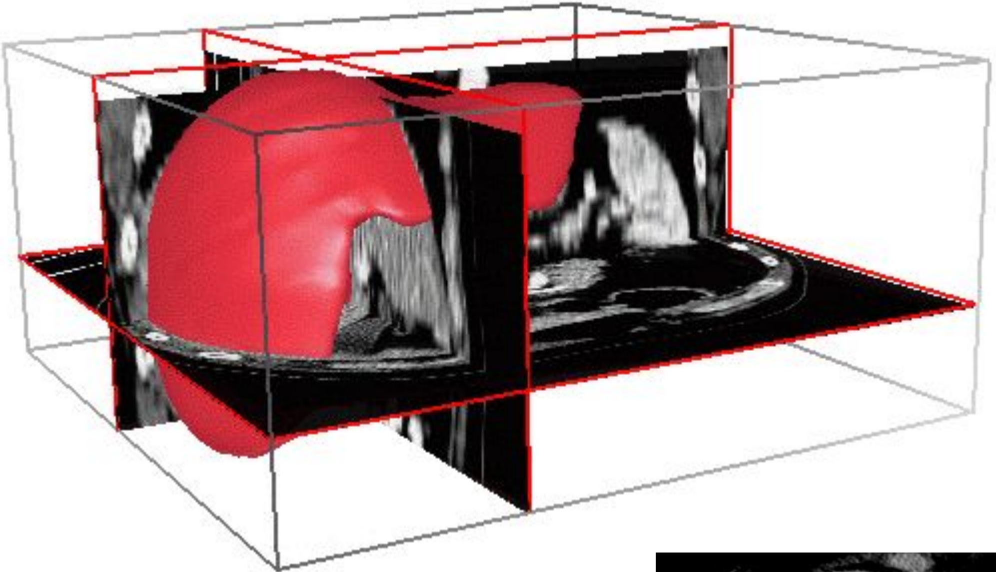
Segmentación

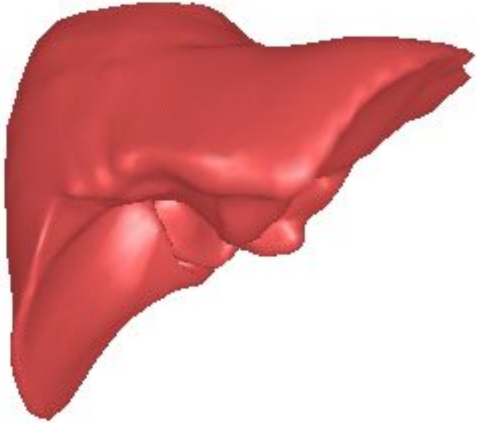


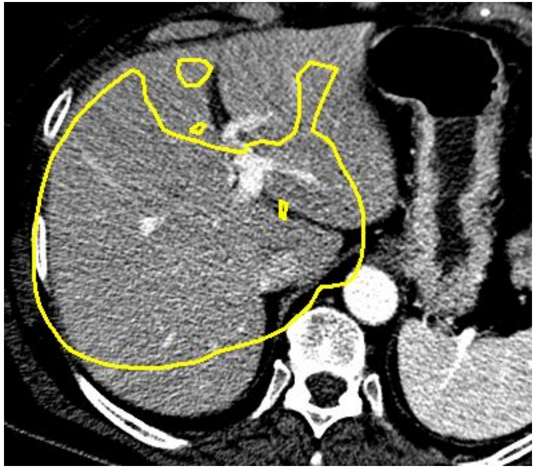

input image
semantic segmentation
instance segmentation
panoptic segmentation

2. Aplicaciones


Camera
Points
Lines
Point Cloud
Segmentation
Normals







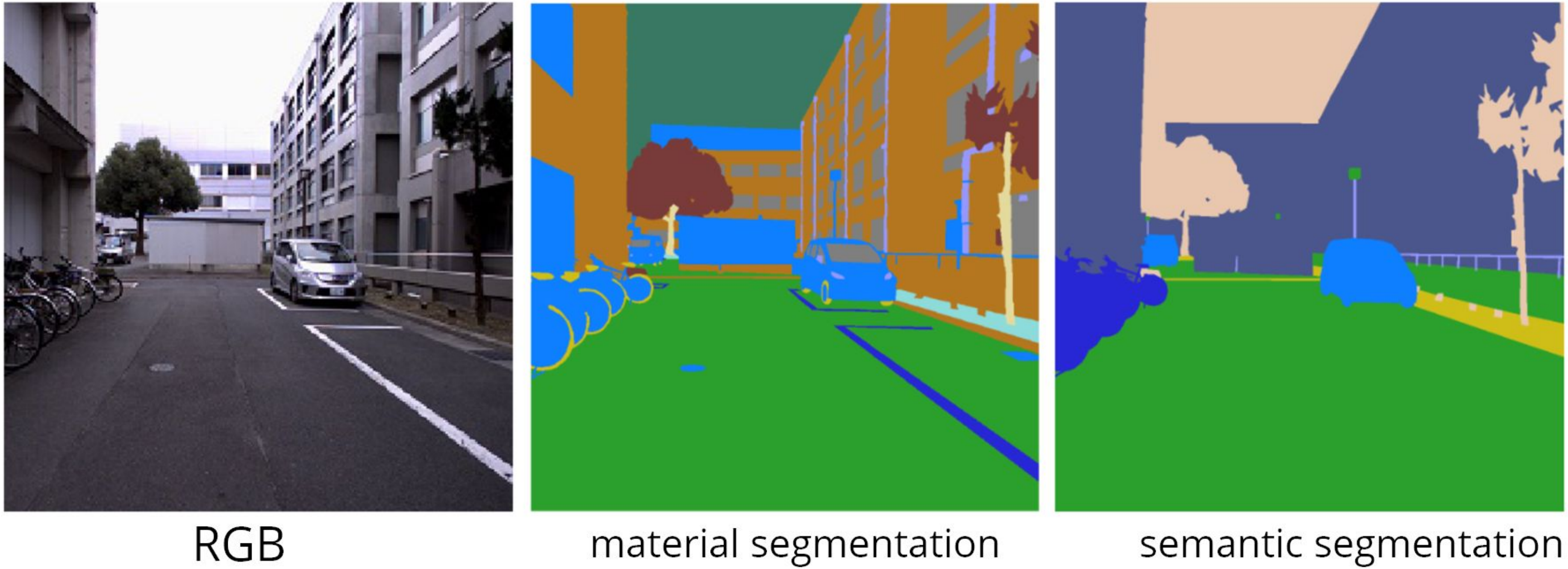

Latency: 1739.4 ms
FPS: 0.5
chair
chair
person
person
chair
chair
person
chair
chair
person
chair
chair
tv
tv
laptop


DeepLab V3 xception_cityscapes_trainfine (GTX980M) INPUT_SIZE=1539
Prediction time: 403ms (2.5 fps) AVG: 356ms (2.8 fps)
WX 86343

3. Segmentación de materiales

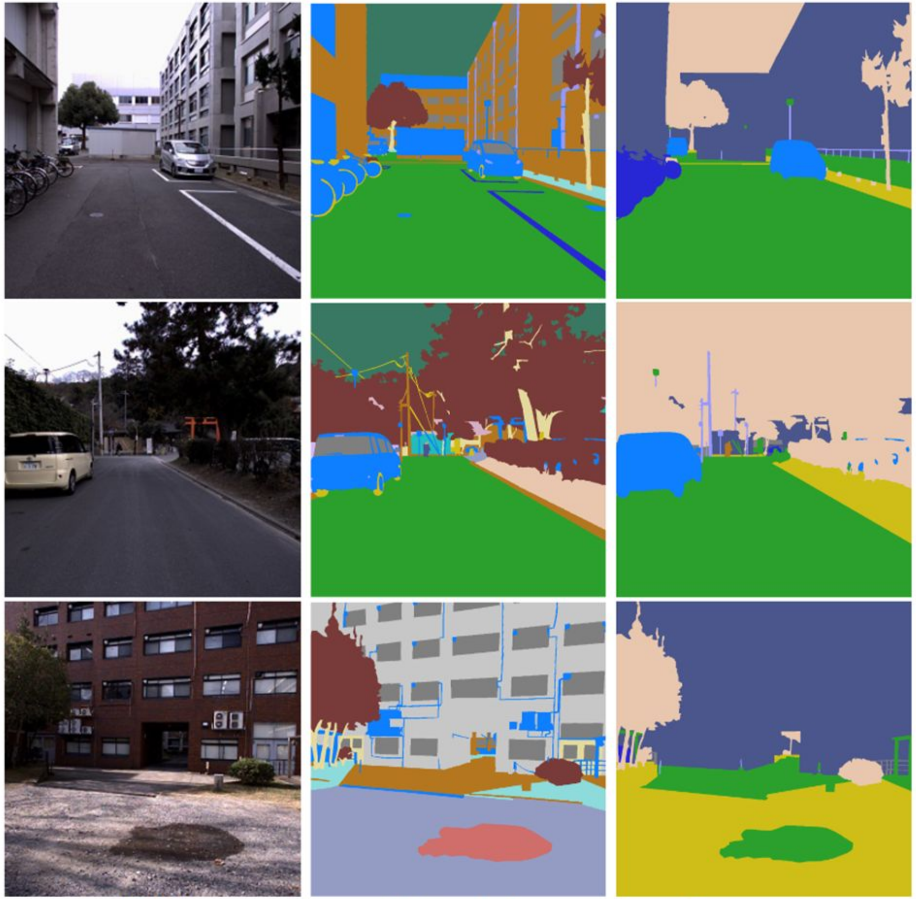

Foliage
Concrete
Plastic
Paint
Metal
Rubber
Asphalt

Semantic segmentation vs Material segmentation



Más información con materiales!

Semantic segmentation vs Material segmentation



RGB Materials Semantics

Segmentación de Materiales

Materiales como el plástico presentan una amplia gama de apariencias según el objeto y la escena

Segmentación de Materiales

Estos patrones de tablero de ajedrez están hechos de **diferentes materiales**

4. Técnicas

BRDF

La reflectancia de muchos materiales varía en función de las direcciones de la luz entrante y saliente, y se describe con la BRDF

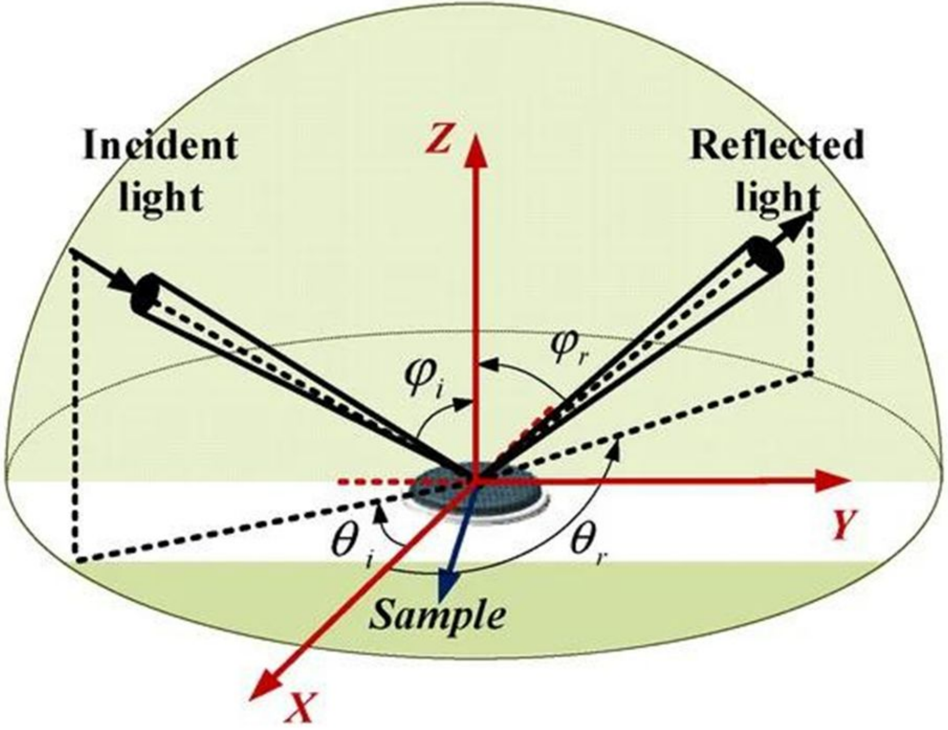
$$f(\theta_r, \varphi_r, \theta_i, \varphi_i) = \frac{L(\theta_r, \varphi_r)}{E(\theta_i, \varphi_i)}$$

$$E(\theta_i, \varphi_i) = \text{Irradiance due to source in direction}(\theta_i, \varphi_i)$$

$$L(\theta_r, \varphi_r) = \text{Radiation of surface in direction}(\theta_r, \varphi_r)$$

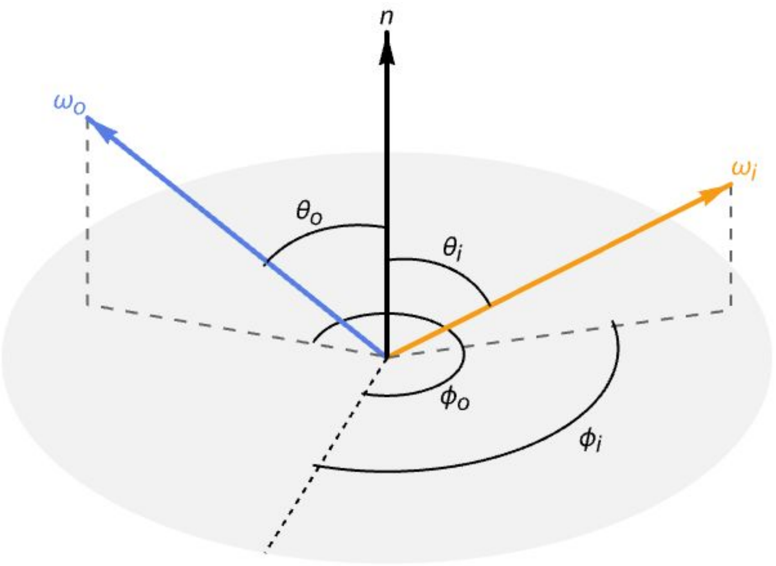
BRDF

$$f(\theta_r, \varphi_r, \theta_i, \varphi_i) = \frac{L(\theta_r, \varphi_r)}{E(\theta_i, \varphi_i)}$$

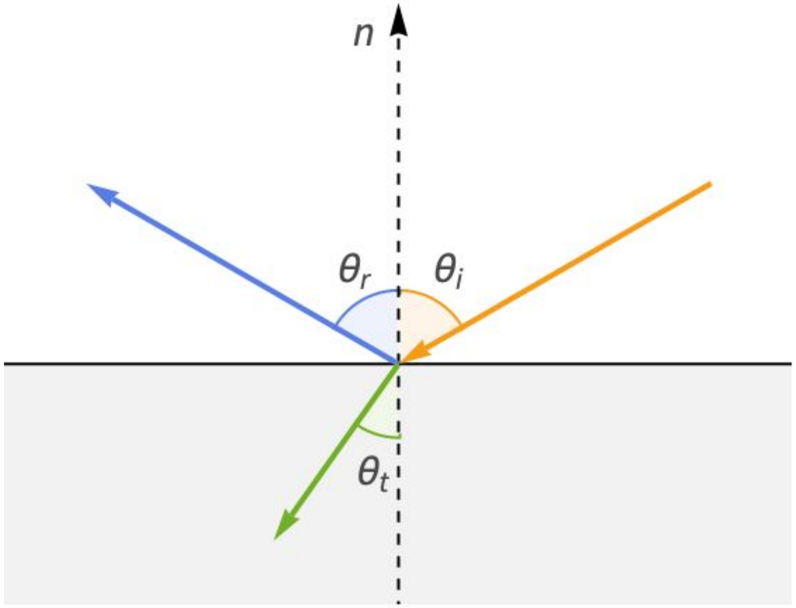


BRDF

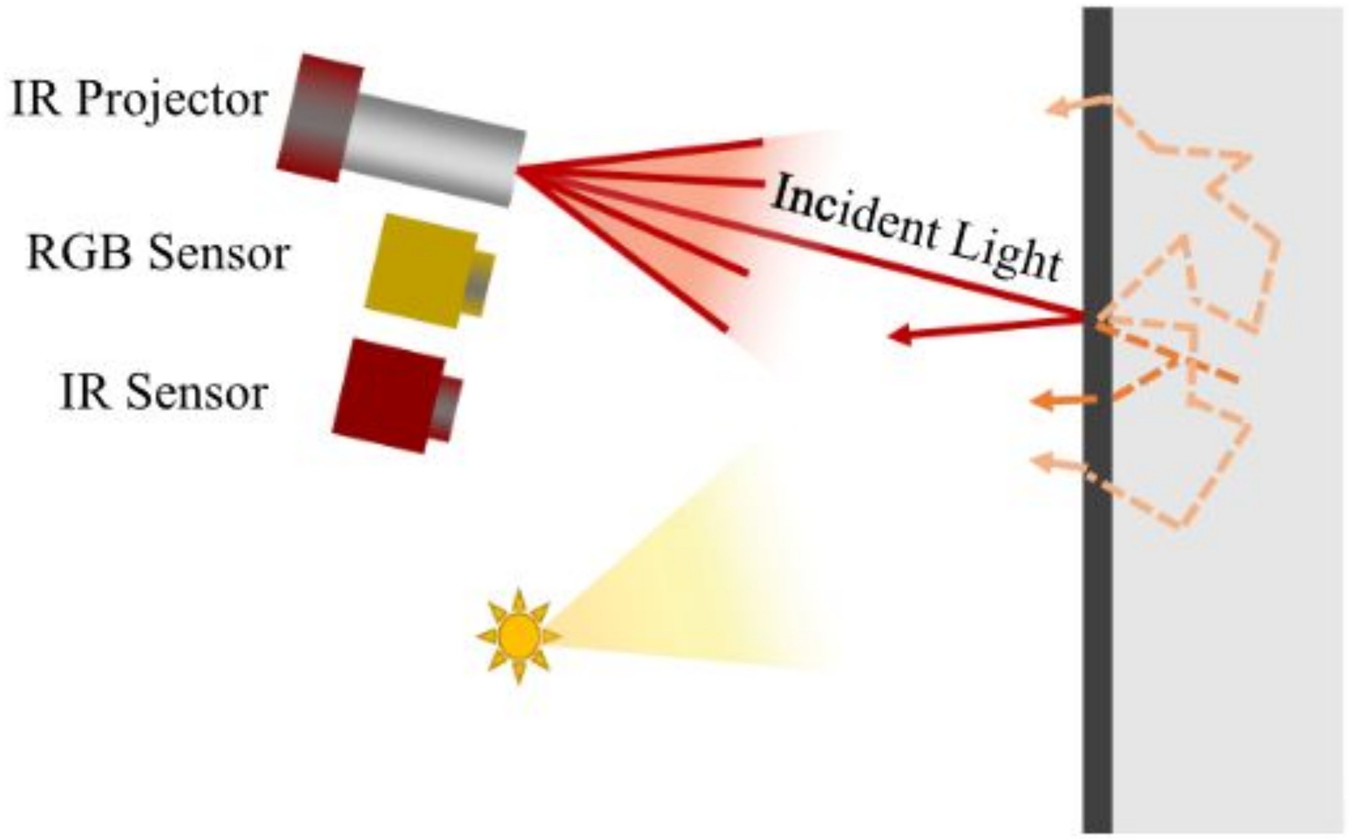
Analisis 3d



Analisis 2d



BRDF



BRDF

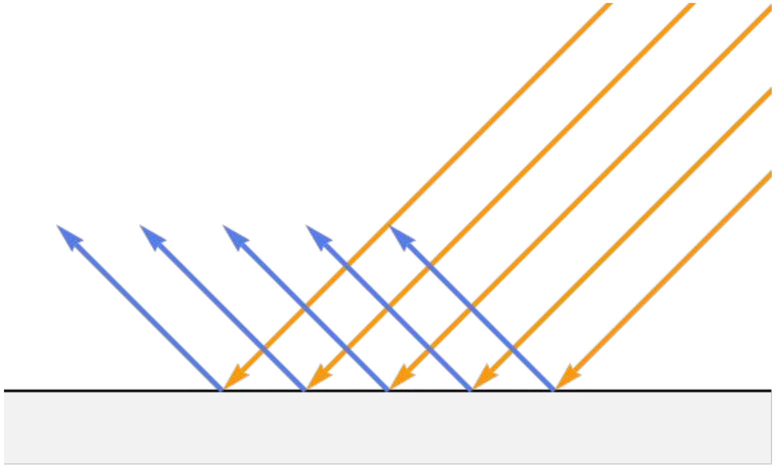
Reflexión de cuerpo

Reflexión superficial

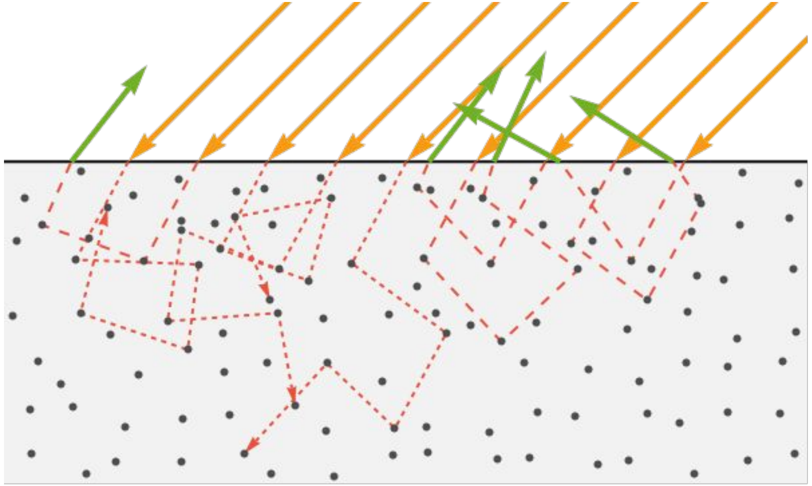
Reflexión hibrida

BRDF

Reflexion especular

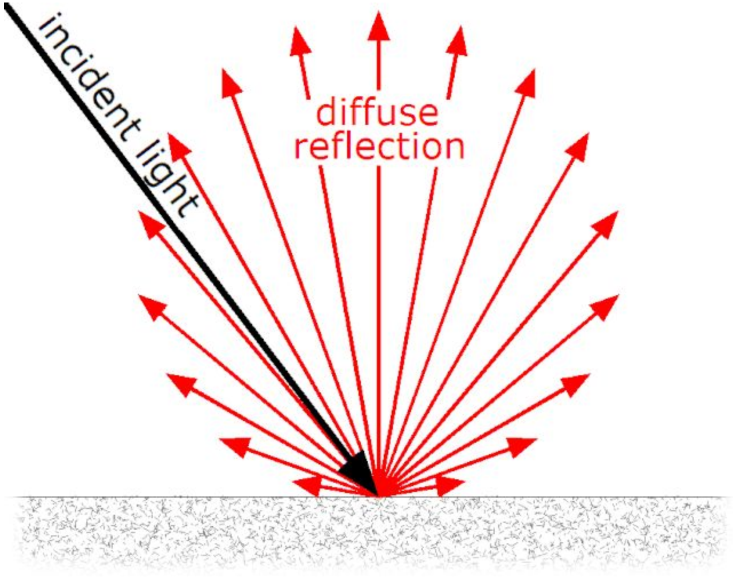


Reflexion difusa



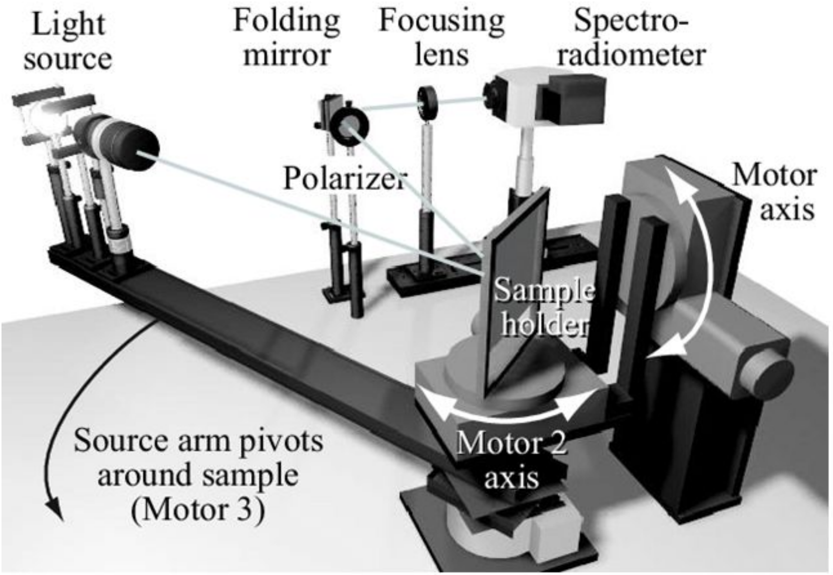
BRDF

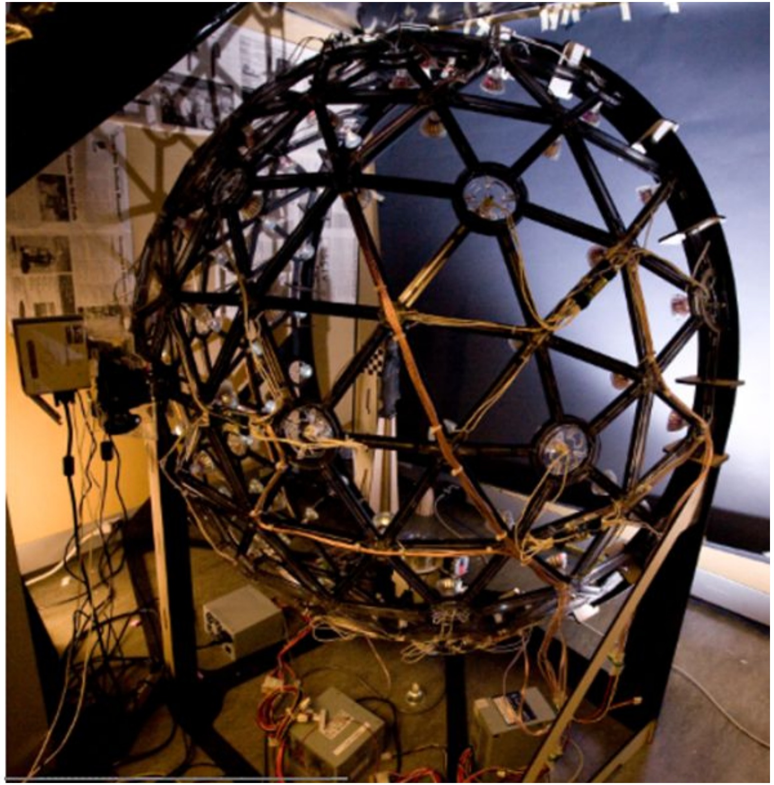
Una aproximación de la reflexión difusa: **superficies lambertianas**



BRDF

Se necesita un gonioreflectómetro para medir la BRDF

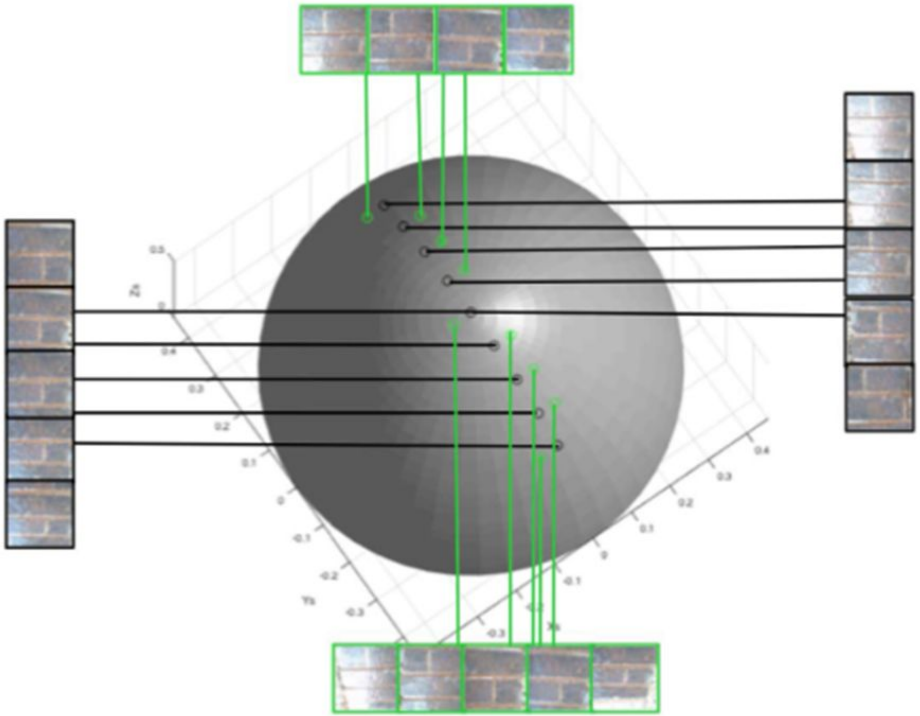


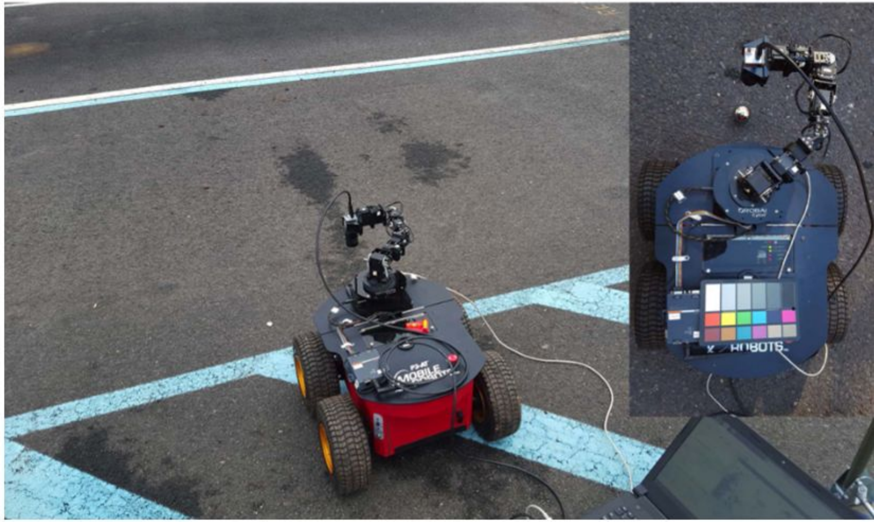


BRDF

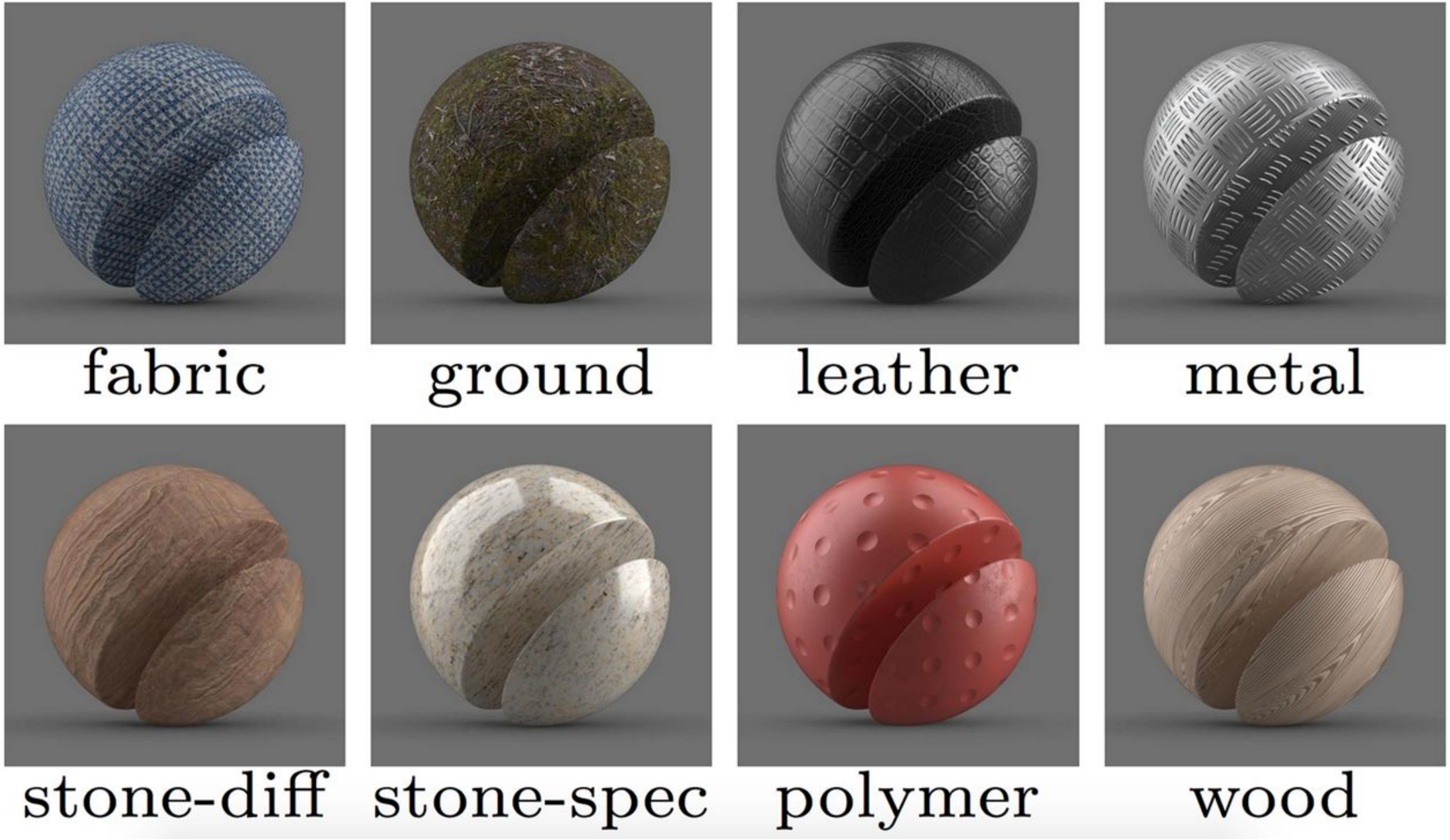
Materials Change with Time

BRDF

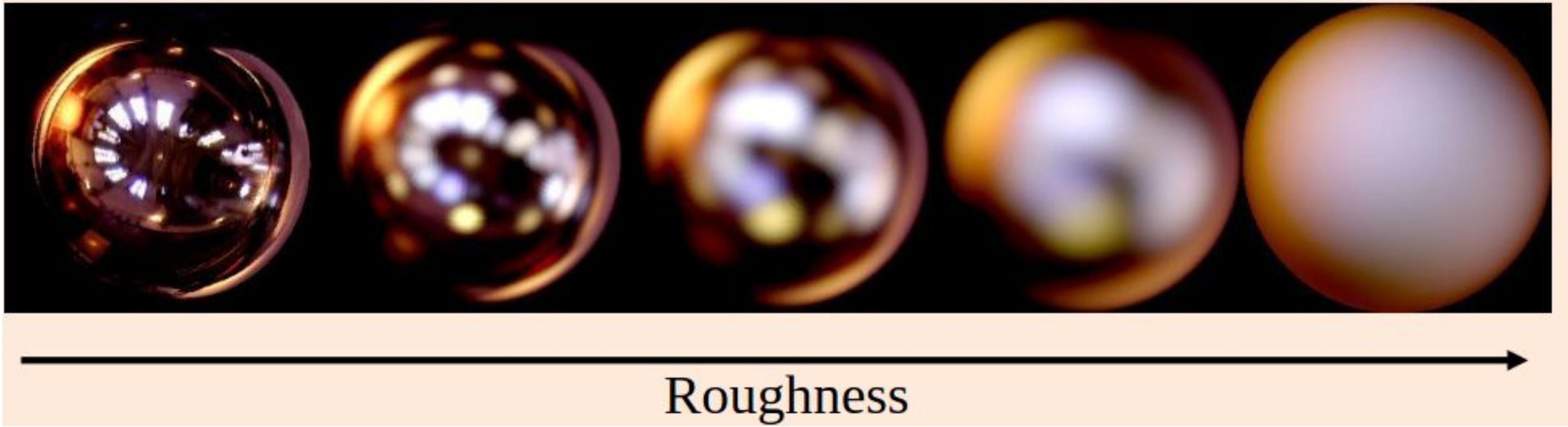




BRDF


fabric
ground
leather
metal
stone-diff
stone-spec
polymer
wood

BRDF



BTF

es una función de seis dimensiones que caracteriza el aspecto de la textura de la superficie de un material en distintas condiciones de iluminación y espacio.

$$f(x, y, \theta_r, \varphi_r, \theta_i, \varphi_i)$$

BTF

es una función de seis dimensiones que caracteriza el aspecto de la textura de la superficie de un material en distintas condiciones de iluminación y espacio.

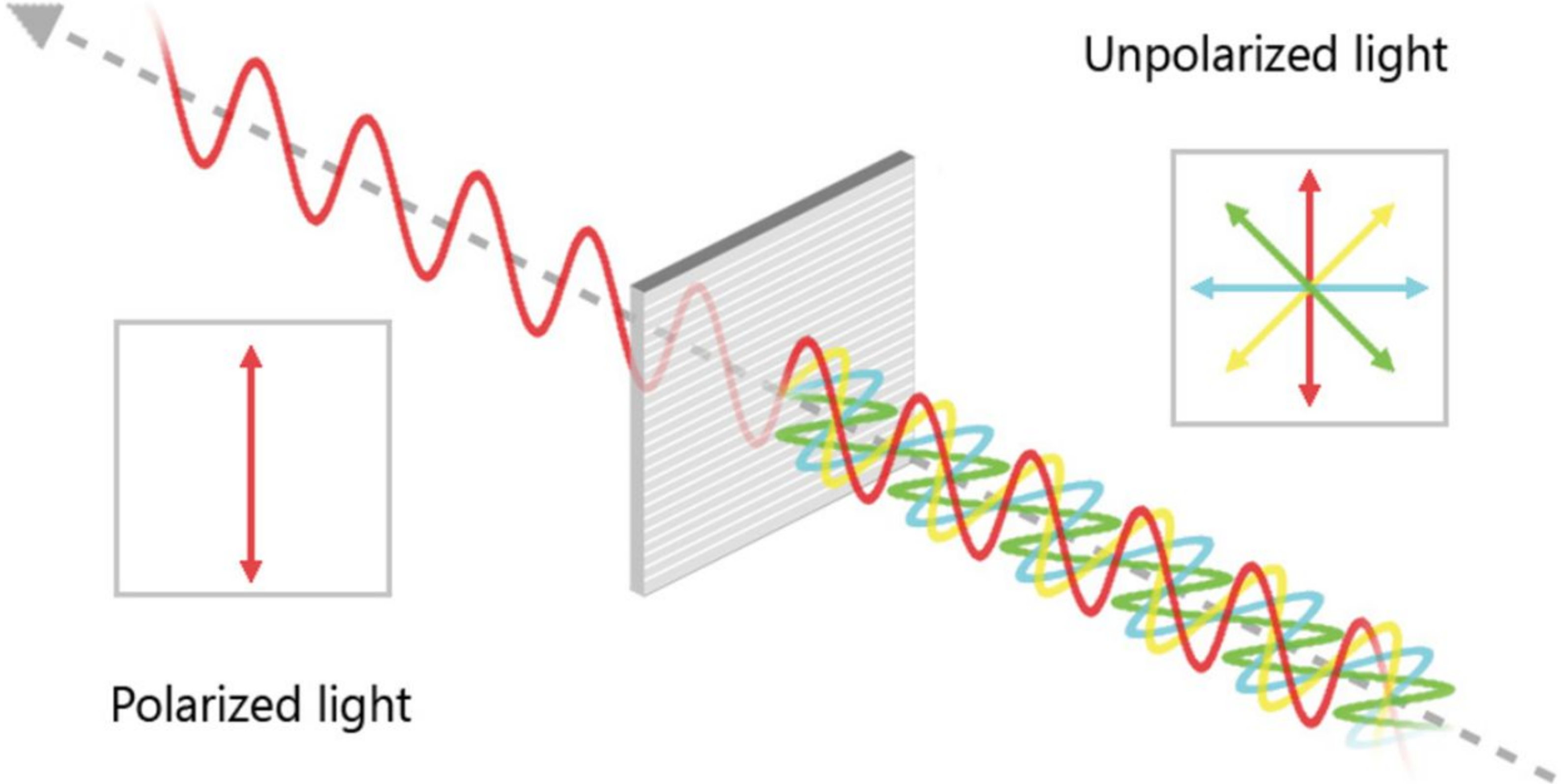
$$f(x, y, \theta_r, \varphi_r, \theta_i, \varphi_i)$$

cada píxel tiene el valor de una BRDF

$$T_{(x,y)}(\theta_r, \varphi_r, \theta_i, \varphi_i)$$

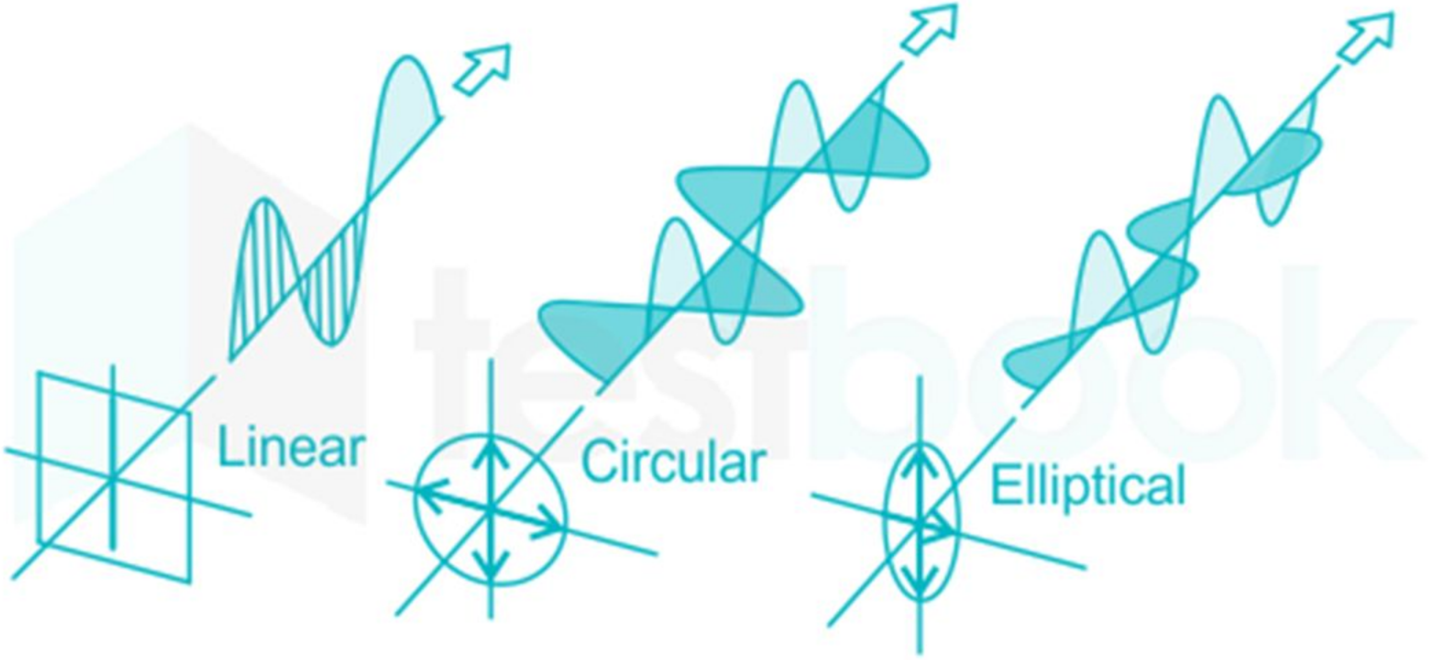
BTF

Polarizacion



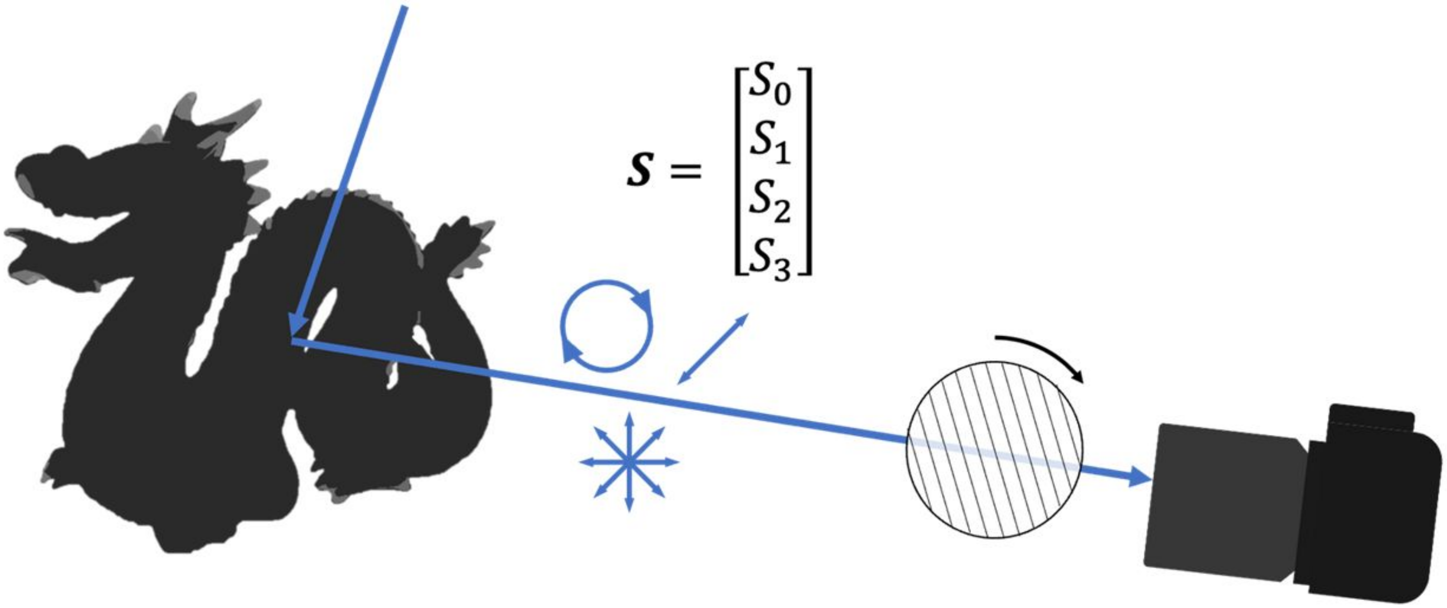
Polarizacion

Para distinguir entre superficies dieléctricas y metálicas basadas en la teoría de reflexión de Fresnel.



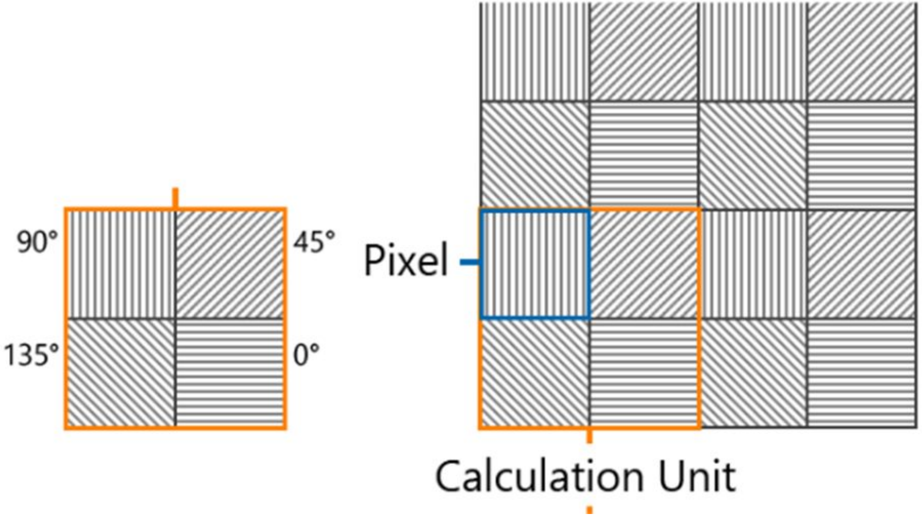
Polarizacion

Para distinguir entre superficies dieléctricas y metálicas basadas en la teoría de reflexión de Fresnel.



Polarizacion

Parametros de stokes



$$S_0 = I = P_{0^\circ} + P_{90^\circ}$$
$$S_1 = Q = P_{0^\circ} - P_{90^\circ}$$
$$S_2 = U = P_{45^\circ} - P_{135^\circ}$$

Polarizacion

Degree of Linear Polarization (DoLp)

The degree of polarization, Π, indicates the ratio of the intensity of the polarized to the intensity of the unpolarized part of the light

$$\Pi = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}$$

Angle of Linear Polarization (AoLp)

The direction of the maximum polarization

$$\Theta = \arctan \frac{S_2}{S_1}$$

Polarizacion

Degree of Linear Polarization (DoLp)

The degree of polarization, Π, indicates the ratio of the intensity of the polarized to the intensity of the unpolarized part of the light

$$\Pi = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}$$

Angle of Linear Polarization (AoLp)

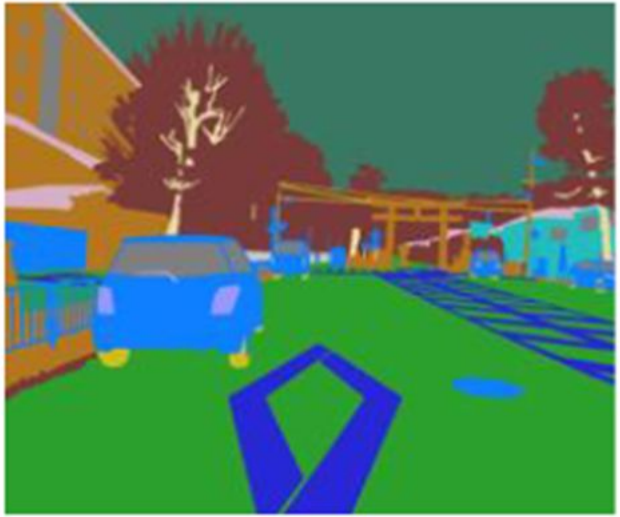
The direction of the maximum polarization

$$\Theta = \arctan \frac{S_2}{S_1}$$

Polarizacion

Angle of Polarization

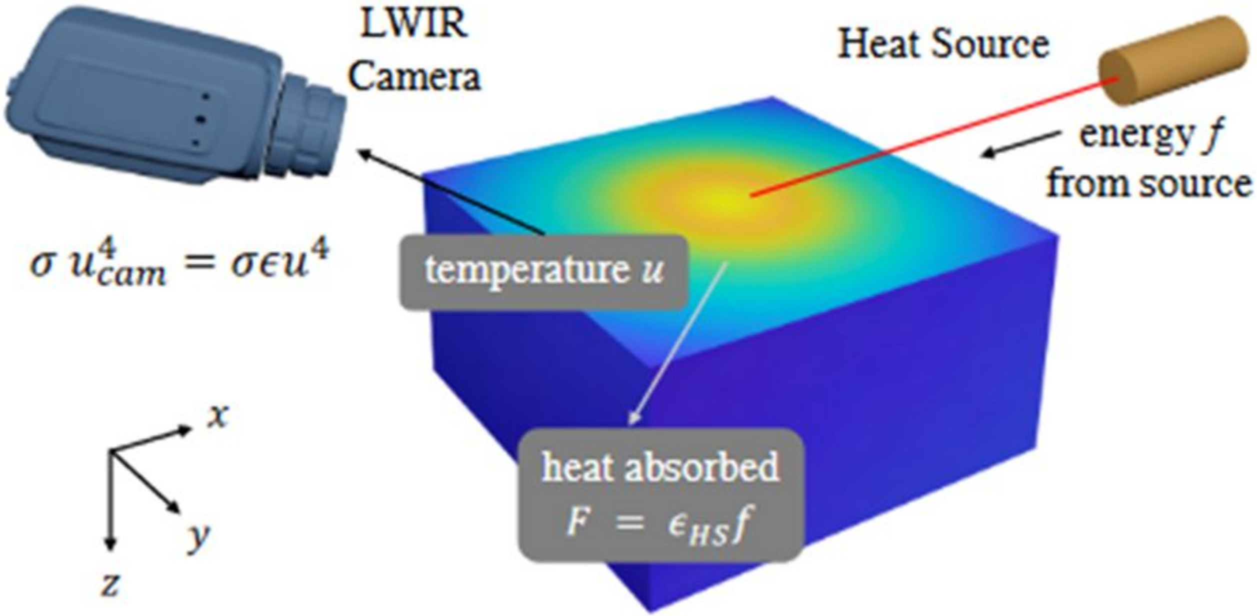
Degree of Polarization



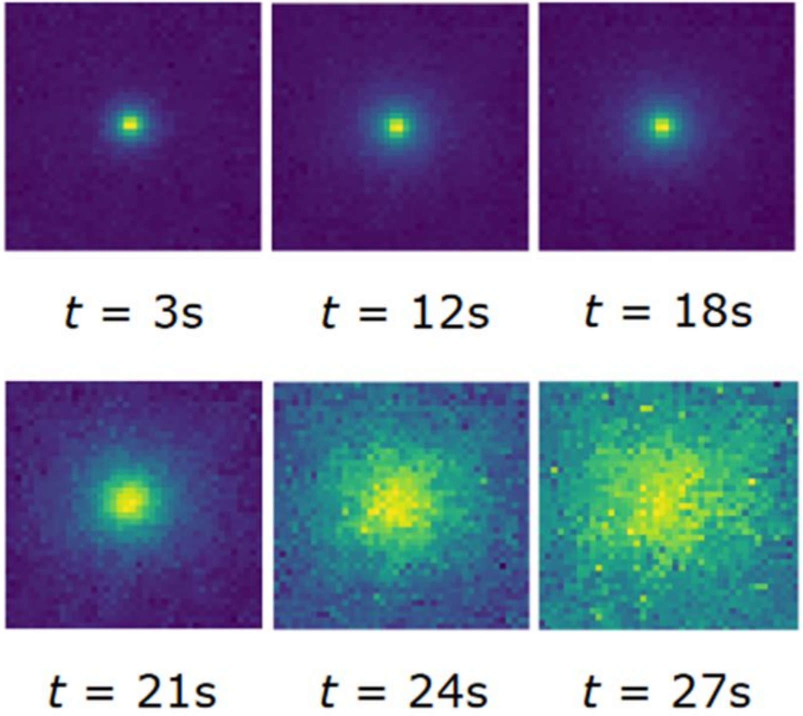
Material Segmentation Annotation

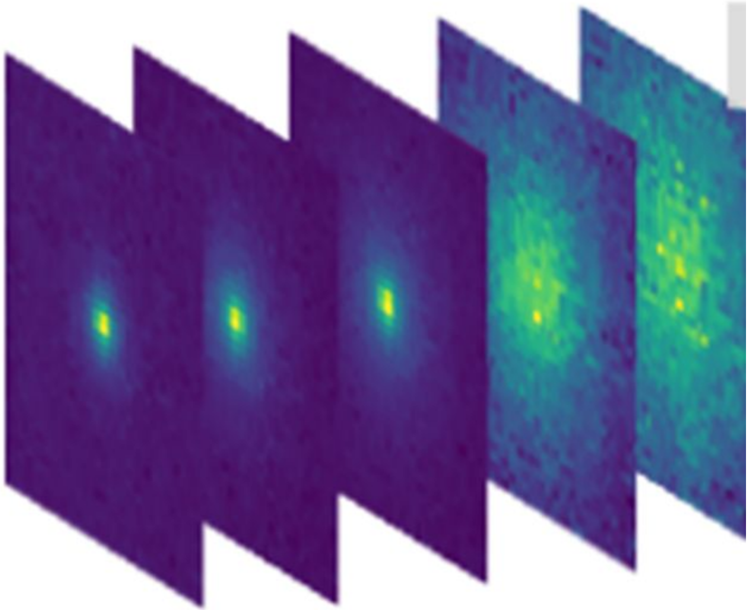
Thermal

Los materiales presentan Funciones de Dispersión Térmica (TSF) únicas cuando se someten a un calentamiento controlado, lo que permite su clasificación

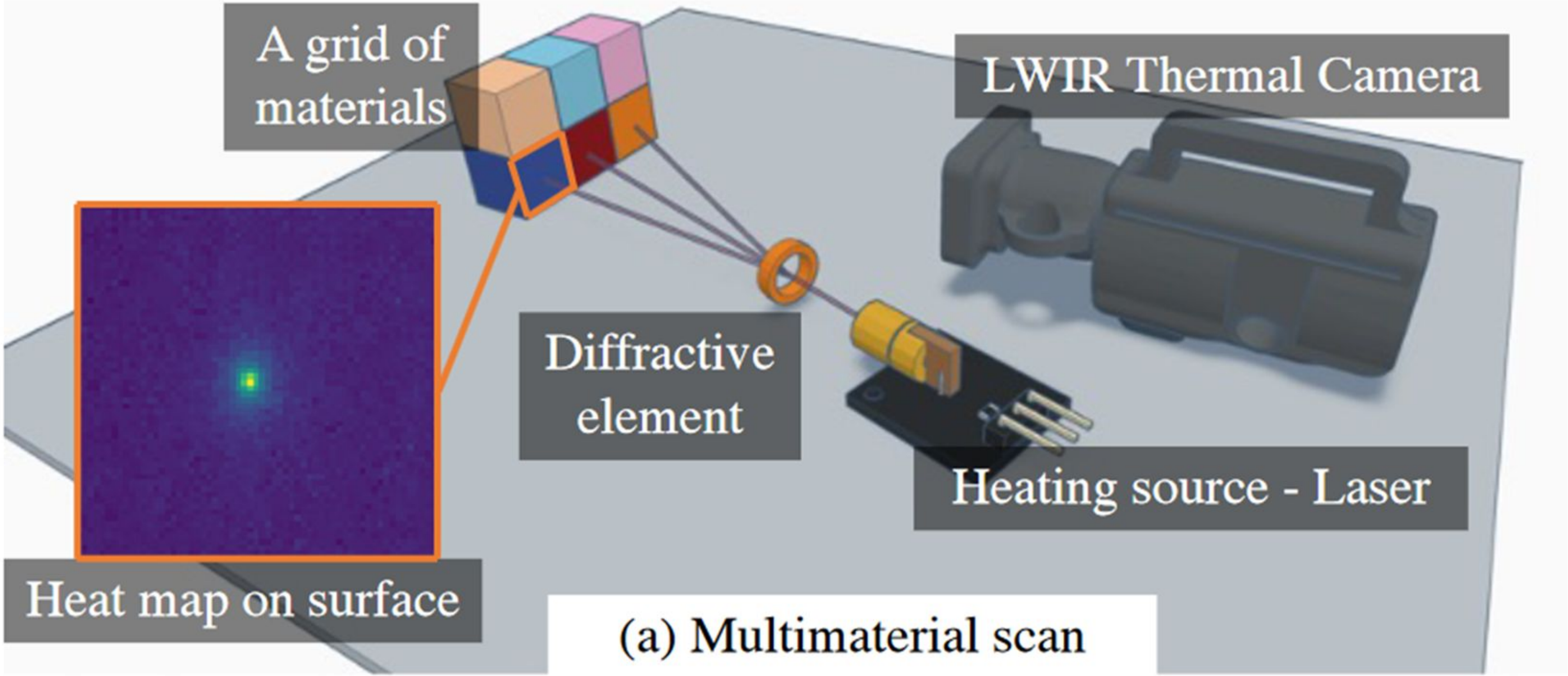


Thermal





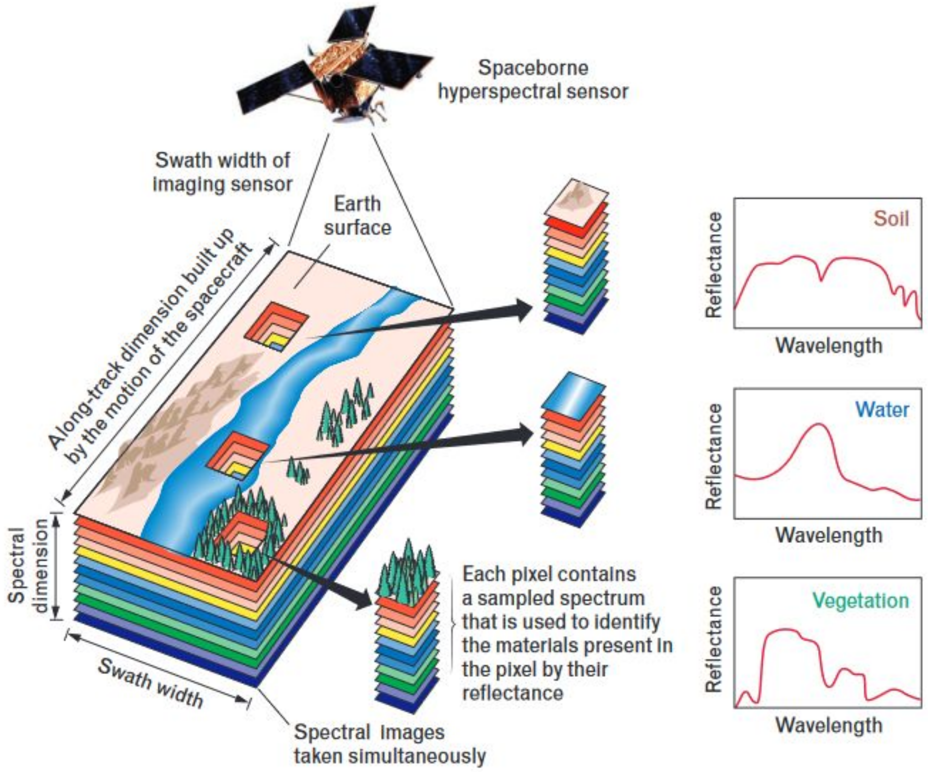
Thermal



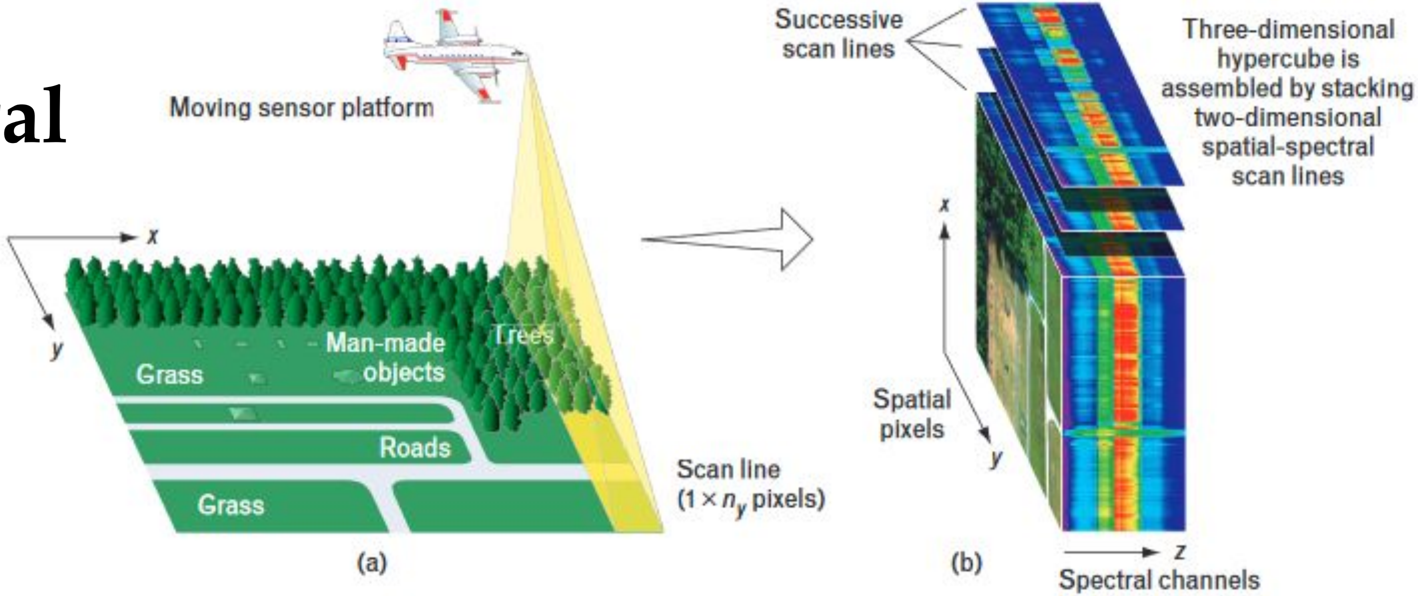
(a) Multimaterial scan

Spectral

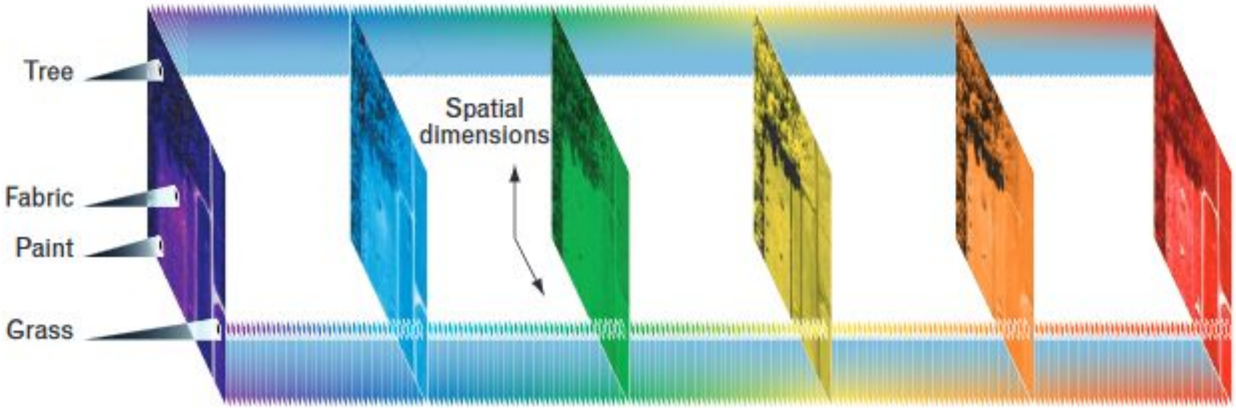
el perfil espectral de las ondas electromagnéticas reflejadas es propio de diversos materiales



Spectral

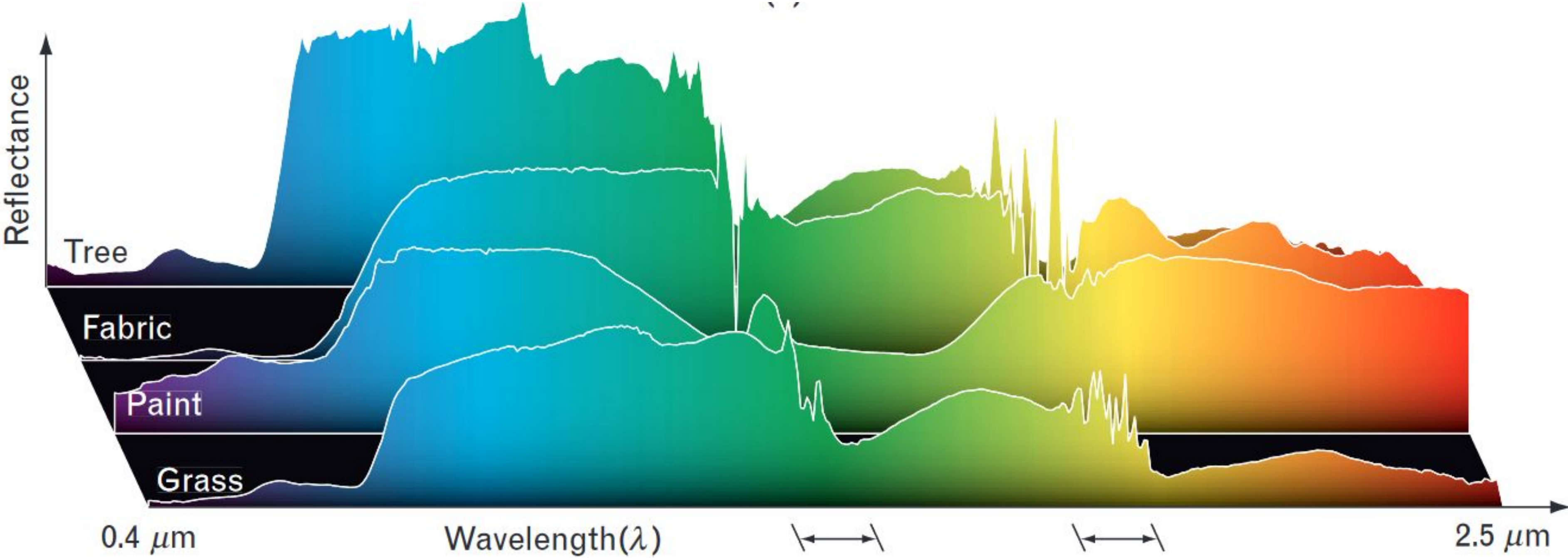


Moving sensor platform
x
y
Grass
Man-made
objects
Trees
Roads
Grass
Scan line
(1 × n_y pixels)
(a)
Successive
scan lines
Three-dimensional
hypercube is
assembled by stacking
two-dimensional
spatial-spectral
scan lines
x
Spatial
pixels
y
(b)
z
Spectral channels

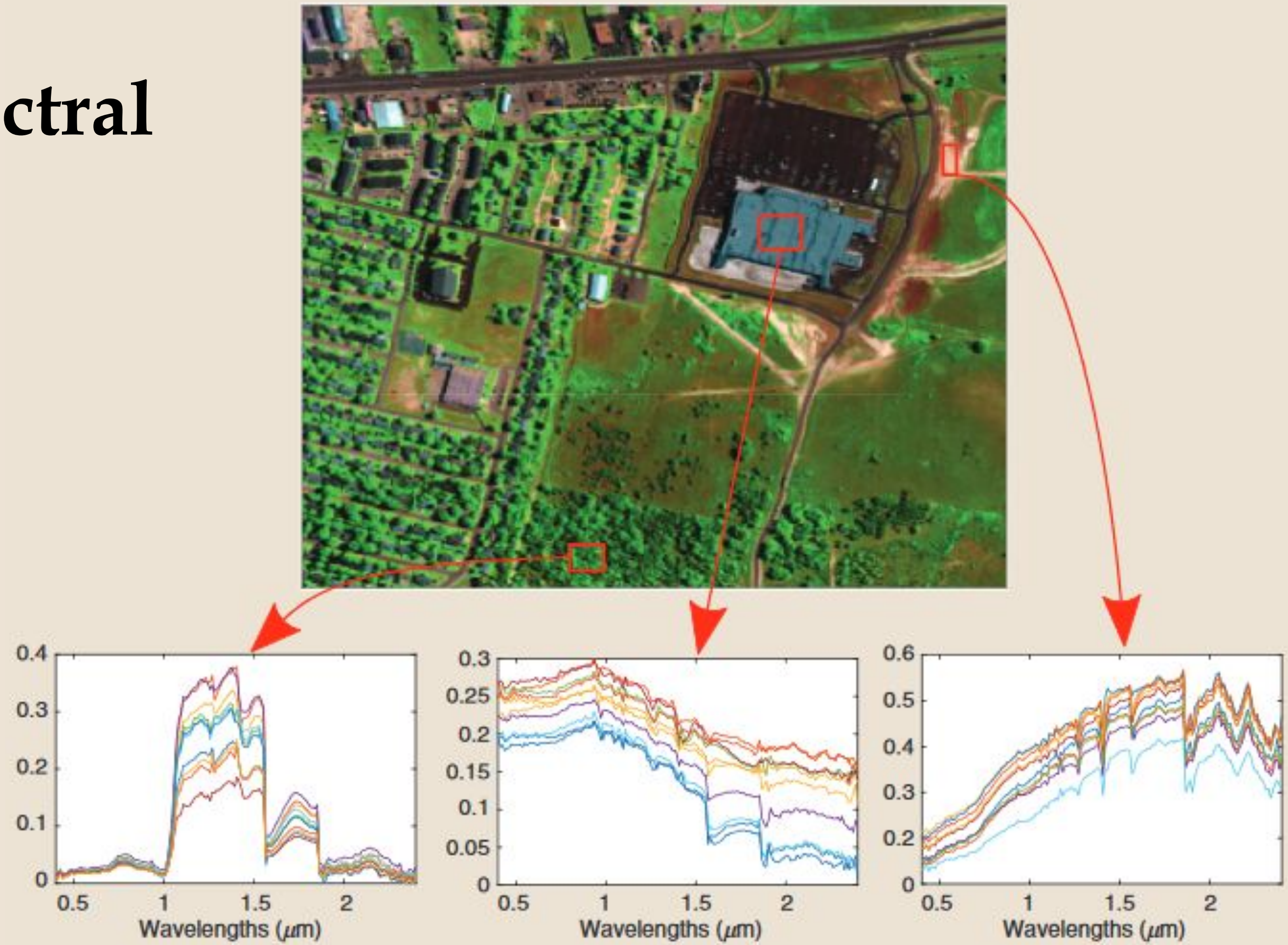


Tree
Fabric
Paint
Grass
Spatial
dimensions

Spectral

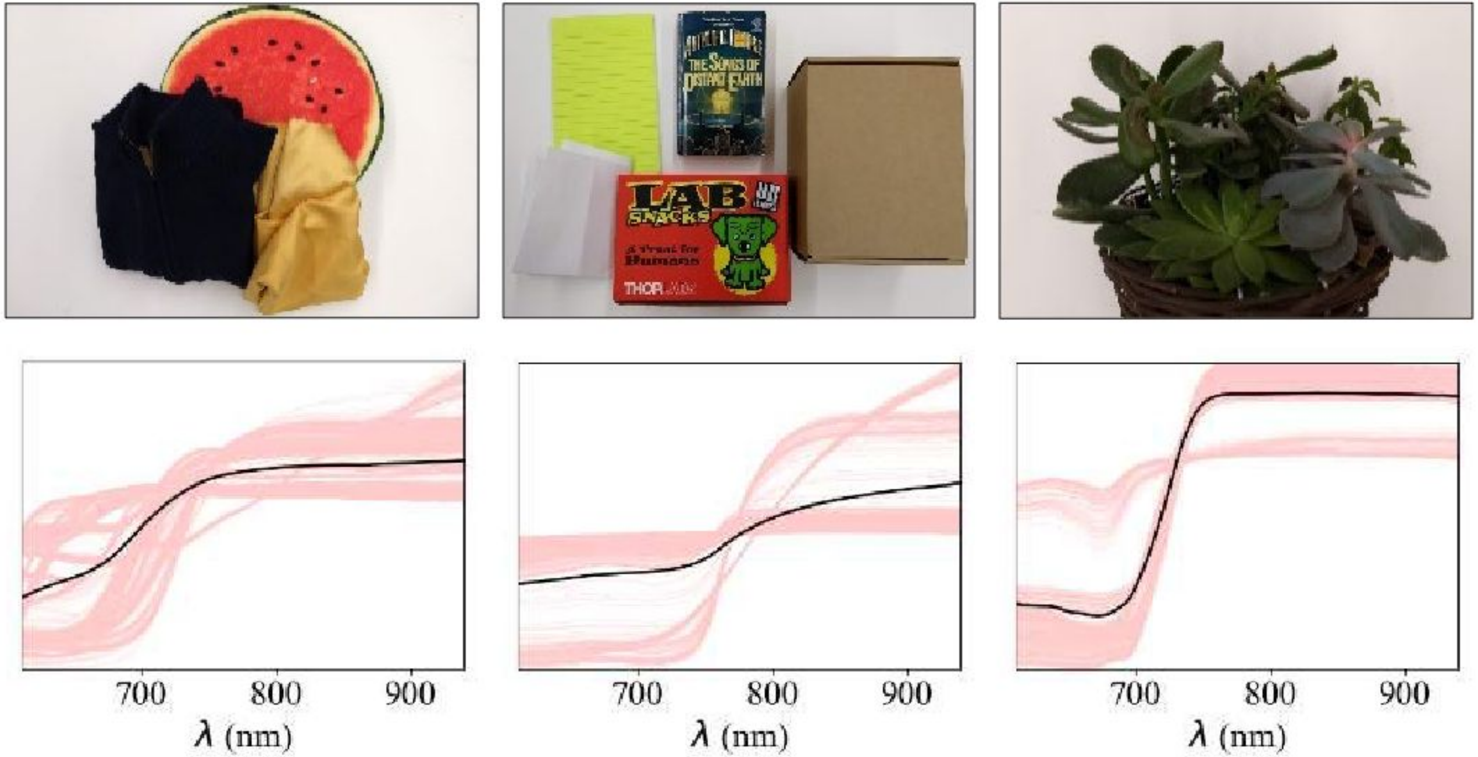

Reflectance
Tree
Fabric
Paint
Grass
0.4 μm
Wavelength(λ)
2.5 μm

Spectral

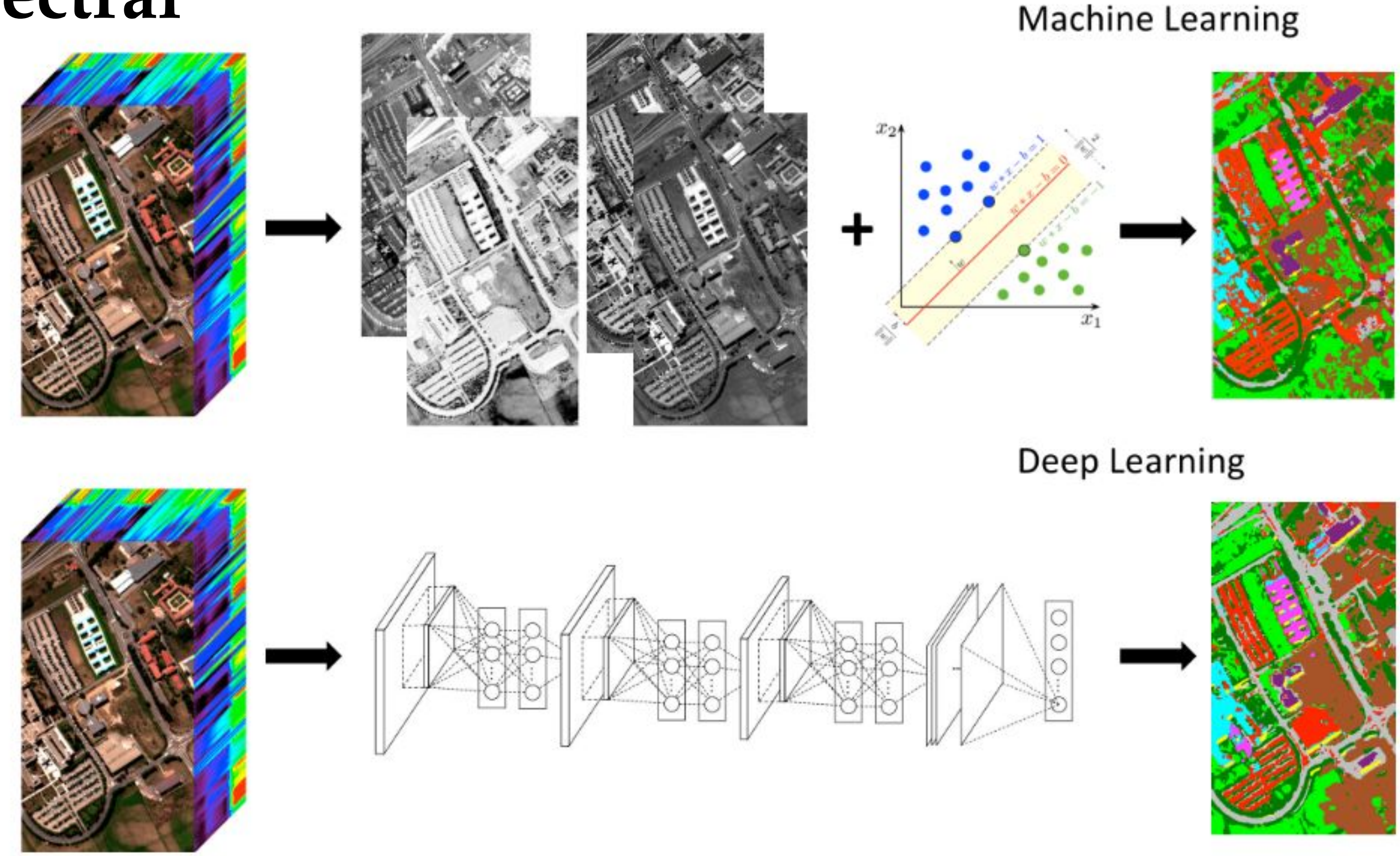


(a) (b) (c)

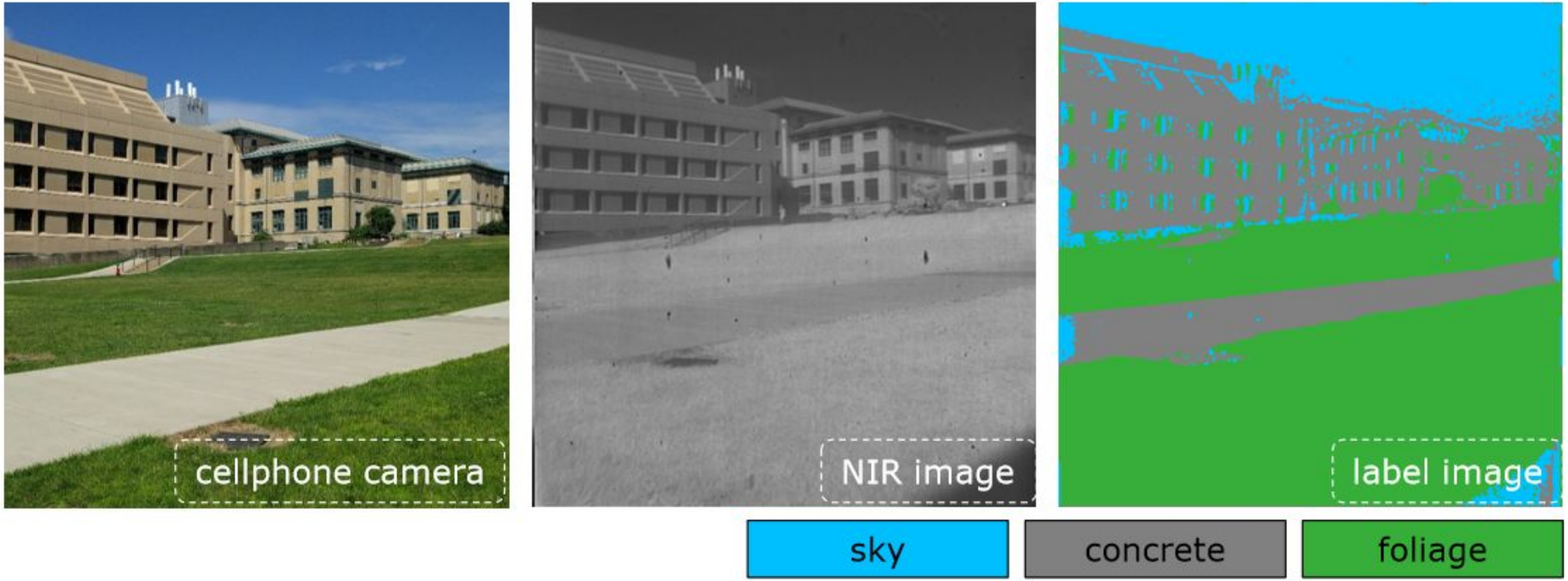
Spectral



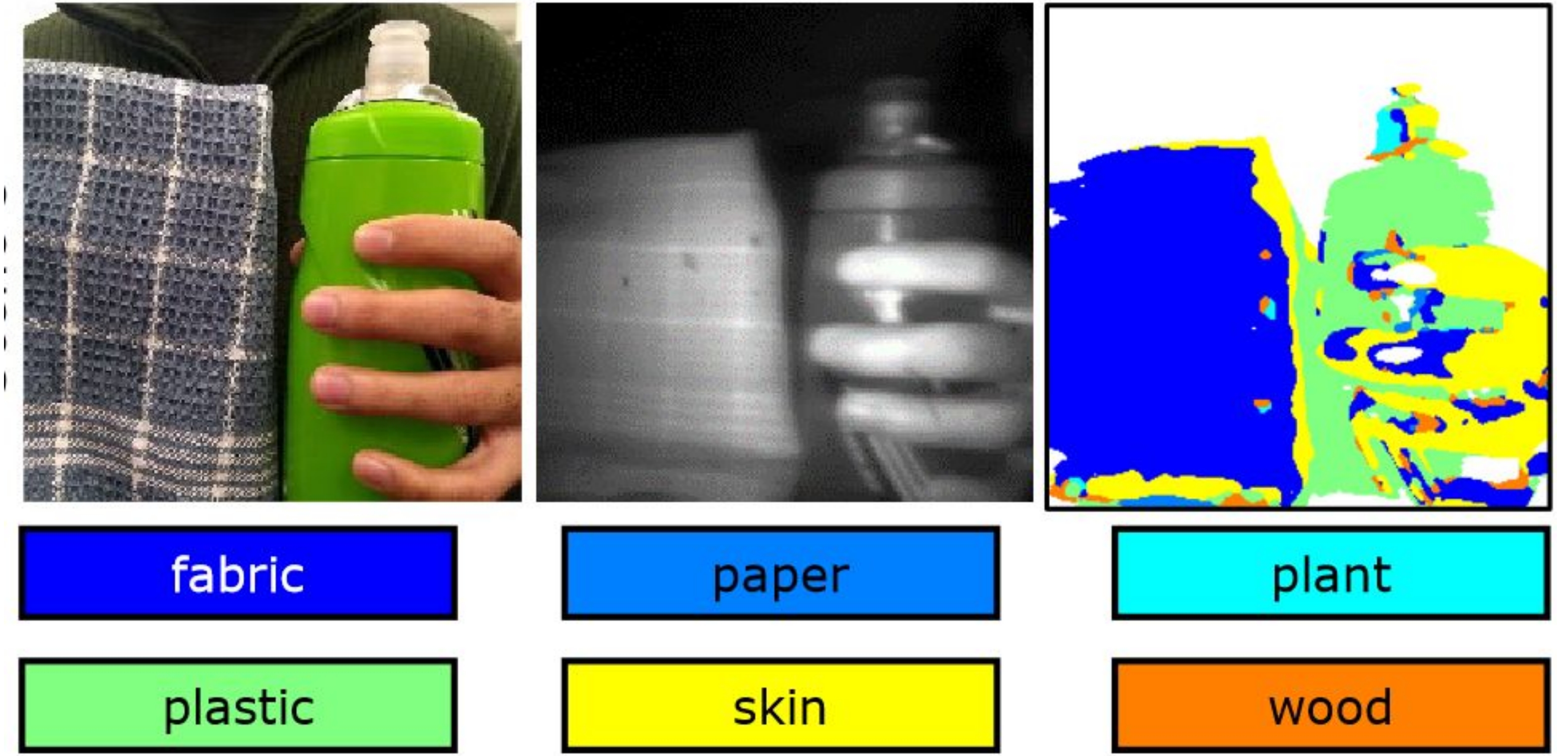
Spectral



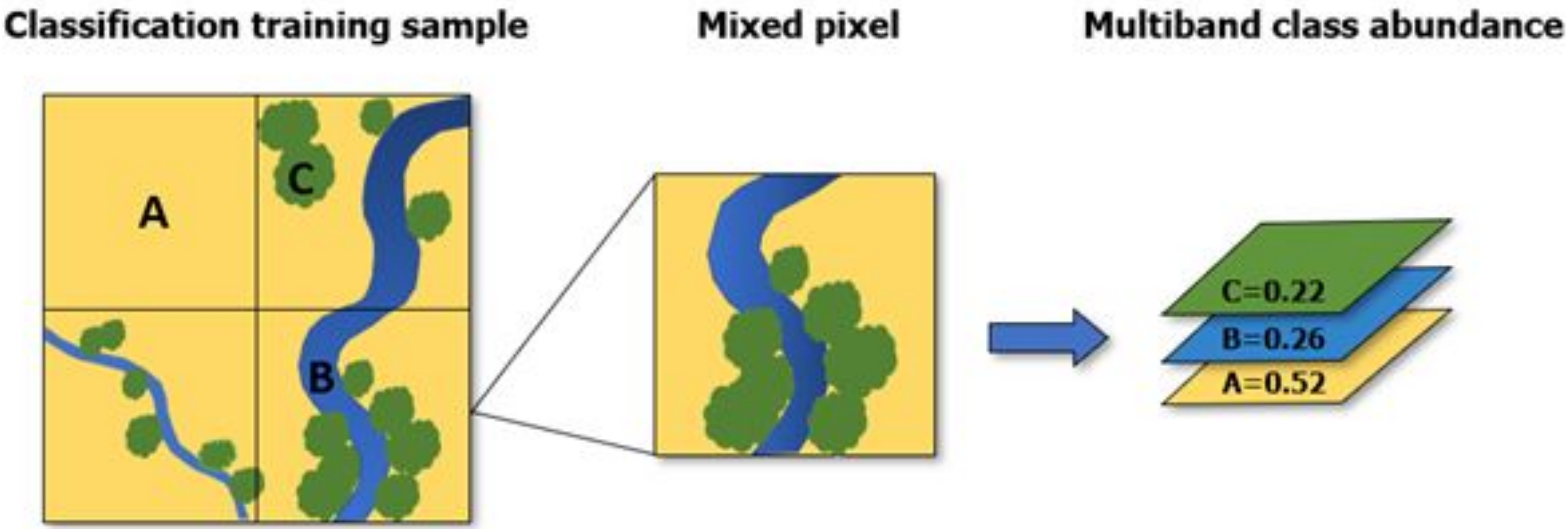
Spectral


cellphone camera
NIR image
label image
sky
concrete
foliage

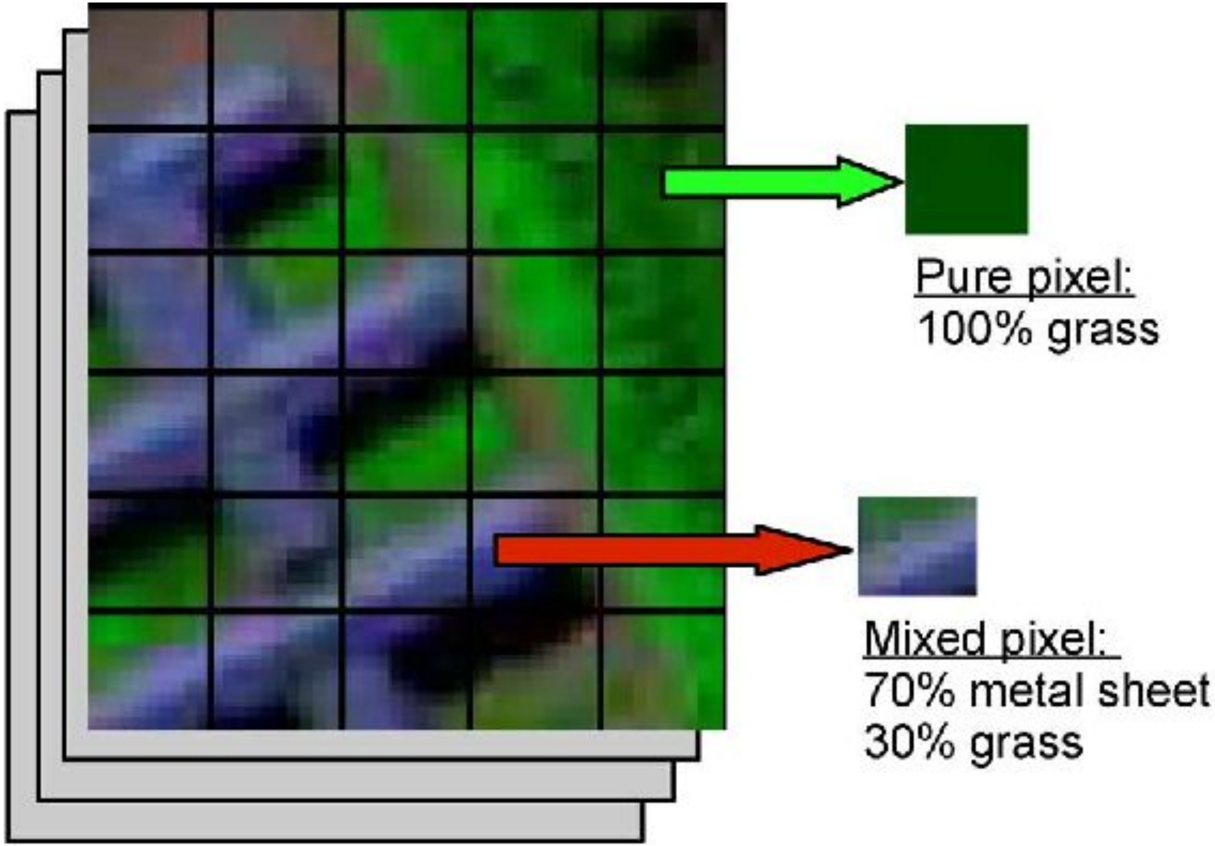
Spectral



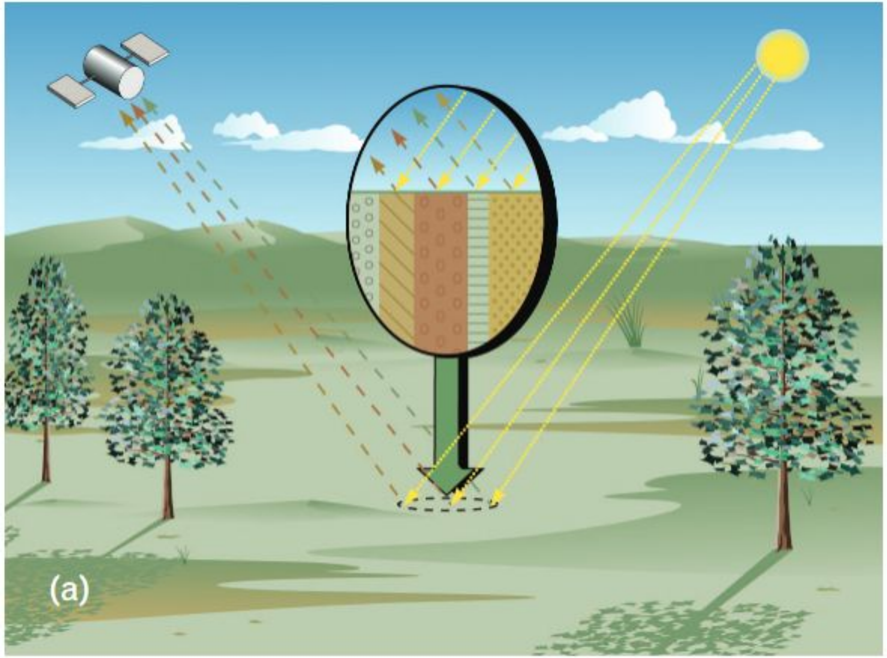
Unmixing


Classification training sample
Mixed pixel
Multiband class abundance
A
C
B
C=0.22
B=0.26
A=0.52

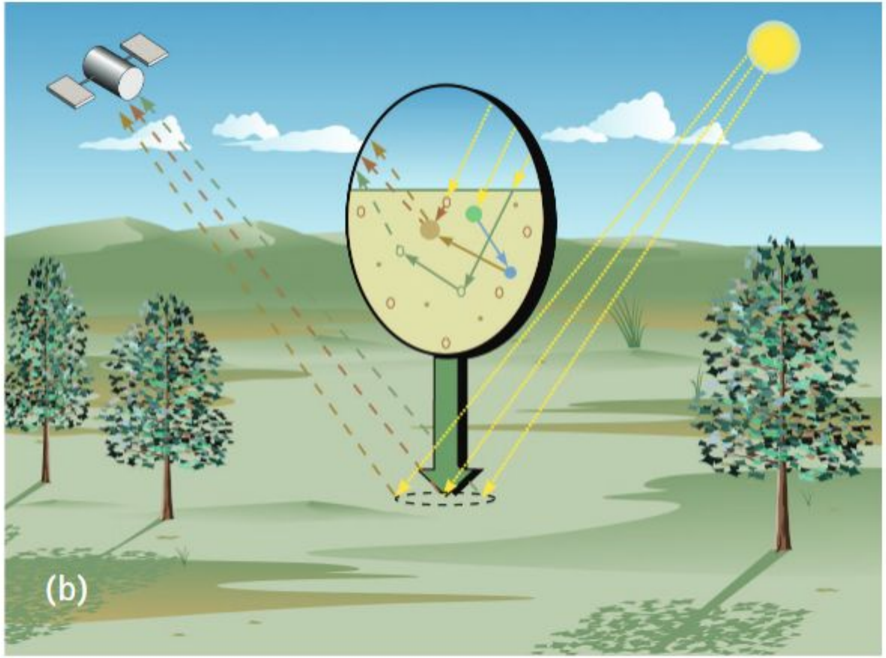
Unmixing



Unmixing

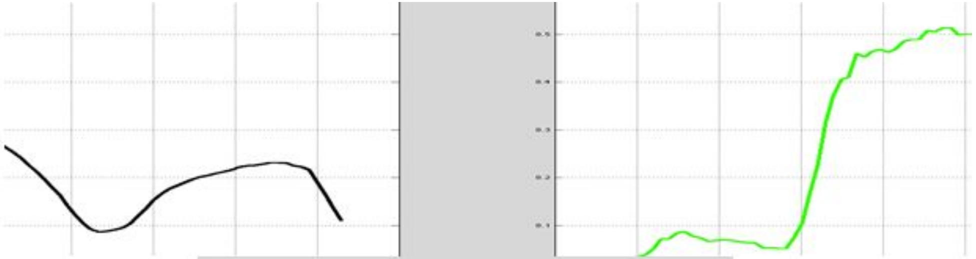


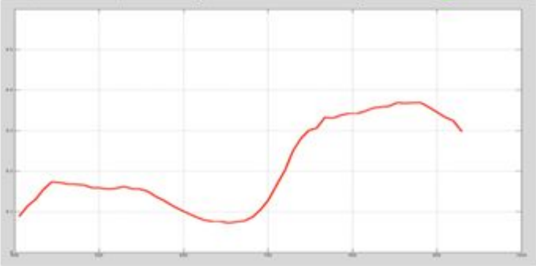
Linear unmixing



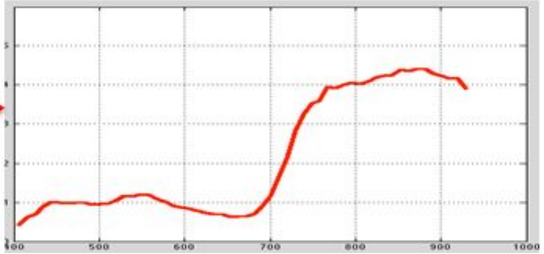
Non-Linear unmixing

Linear Unmixing





0.5Rock+0.5Grass



0.25Rock+0.75Grass

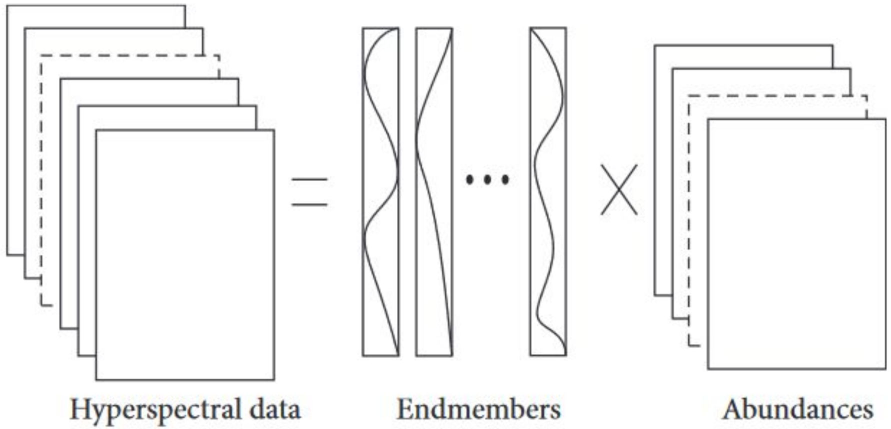
Linear Unmixing

$$\mathbf{Y} = \mathbf{E}\mathbf{A} + \mathbf{N}$$

Subject to

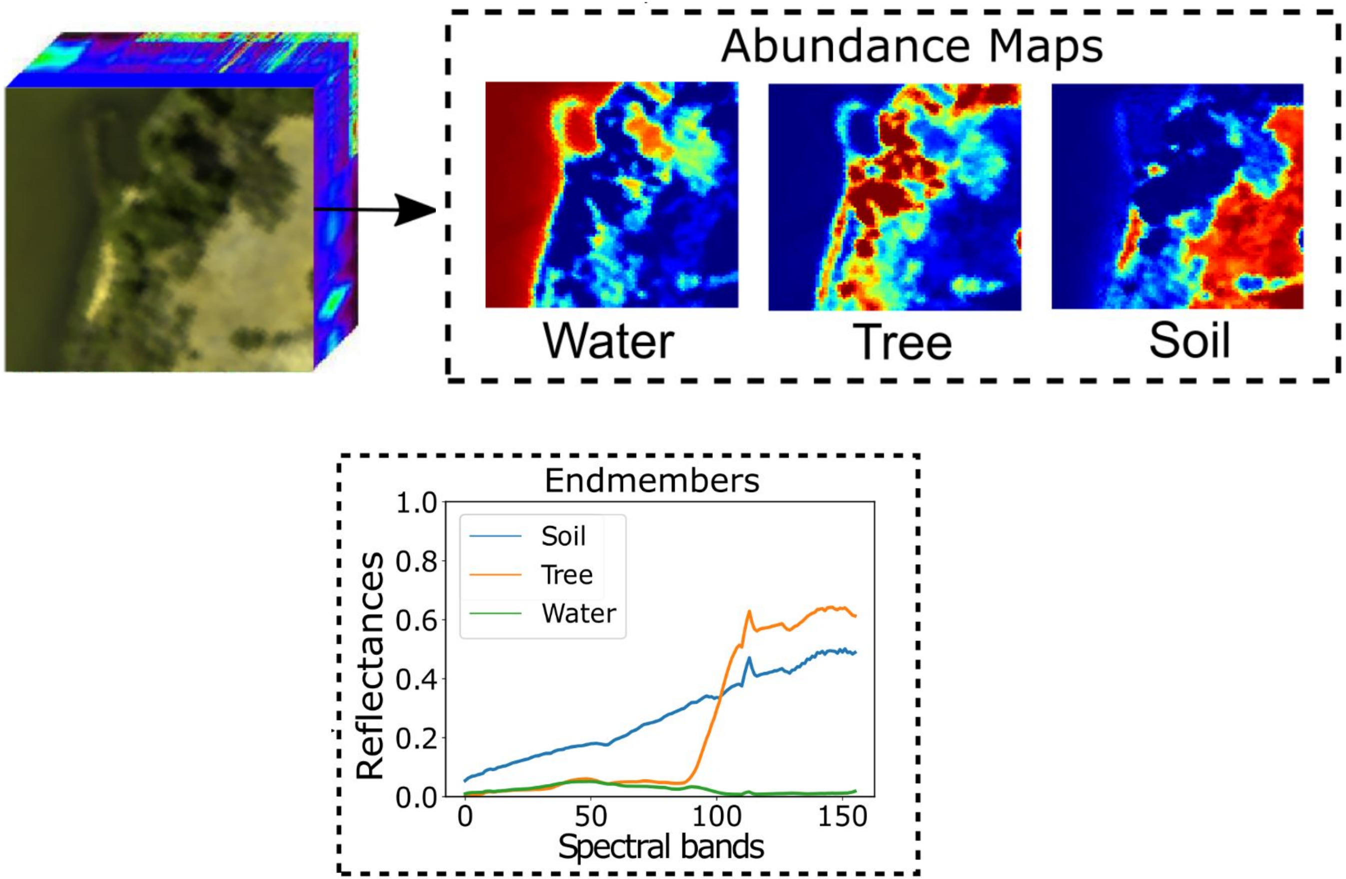
$$\mathbf{A} \geq 0$$

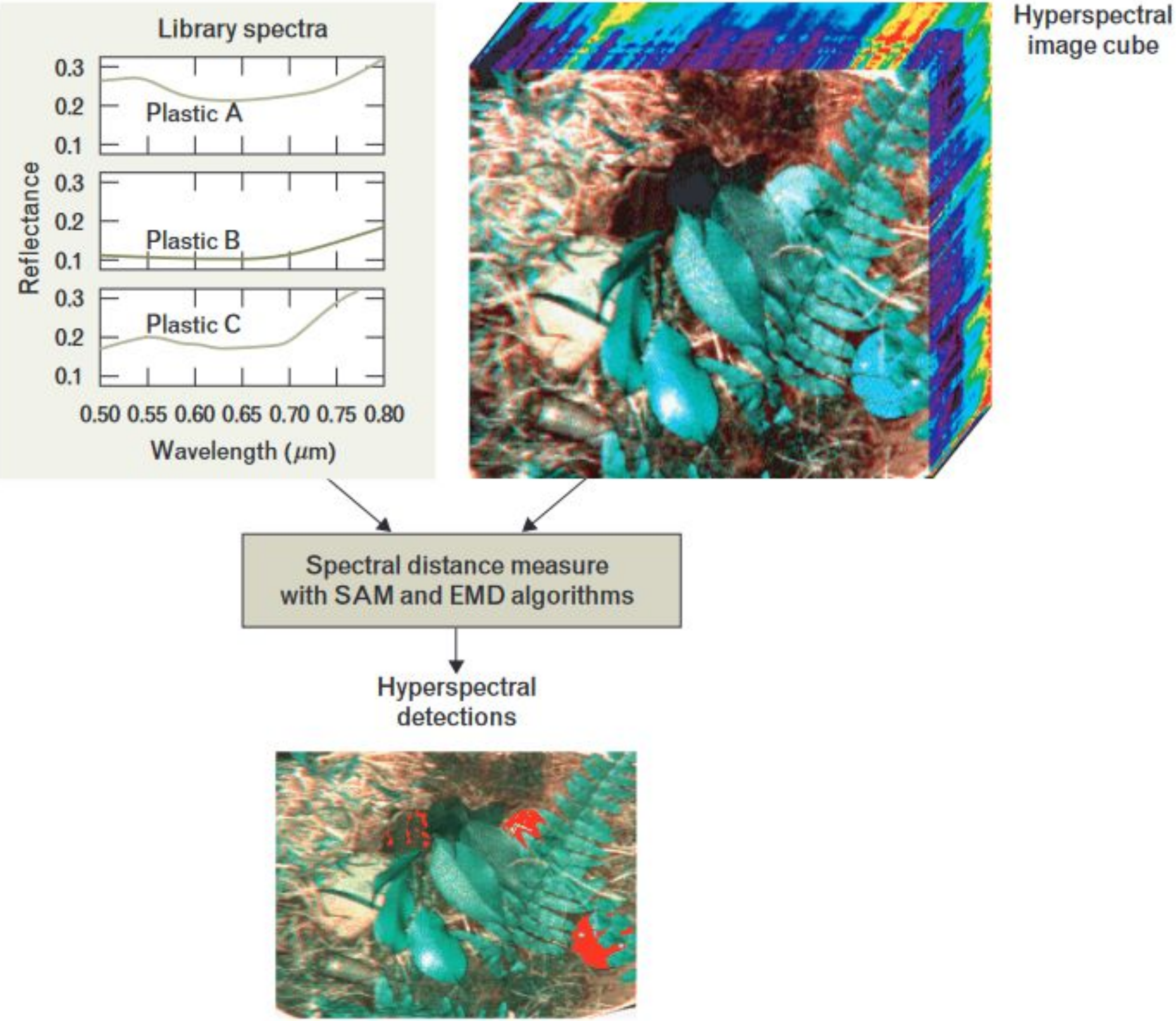
$$\mathbf{1}_R^T \mathbf{A} = \mathbf{1}_n^T$$

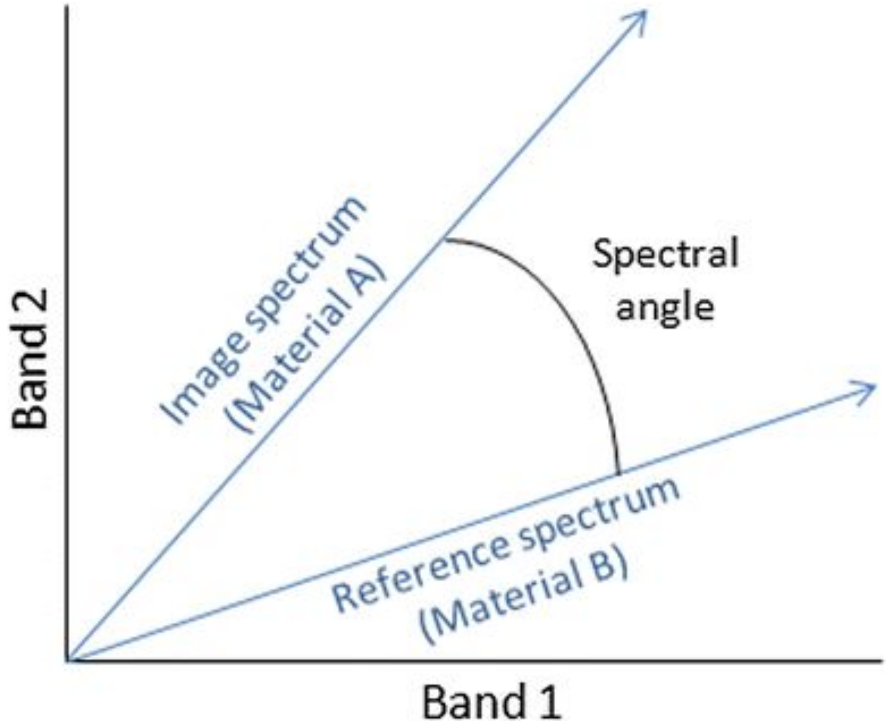


$\mathbf{Y} \in \mathbb{R}^{B \times n}$, where $n = H \cdot W$ $\quad$ $\mathbf{E} \in \mathbb{R}^{B \times R}$ $\quad$ $\mathbf{A} \in \mathbb{R}^{R \times n}$

Linear Unmixing

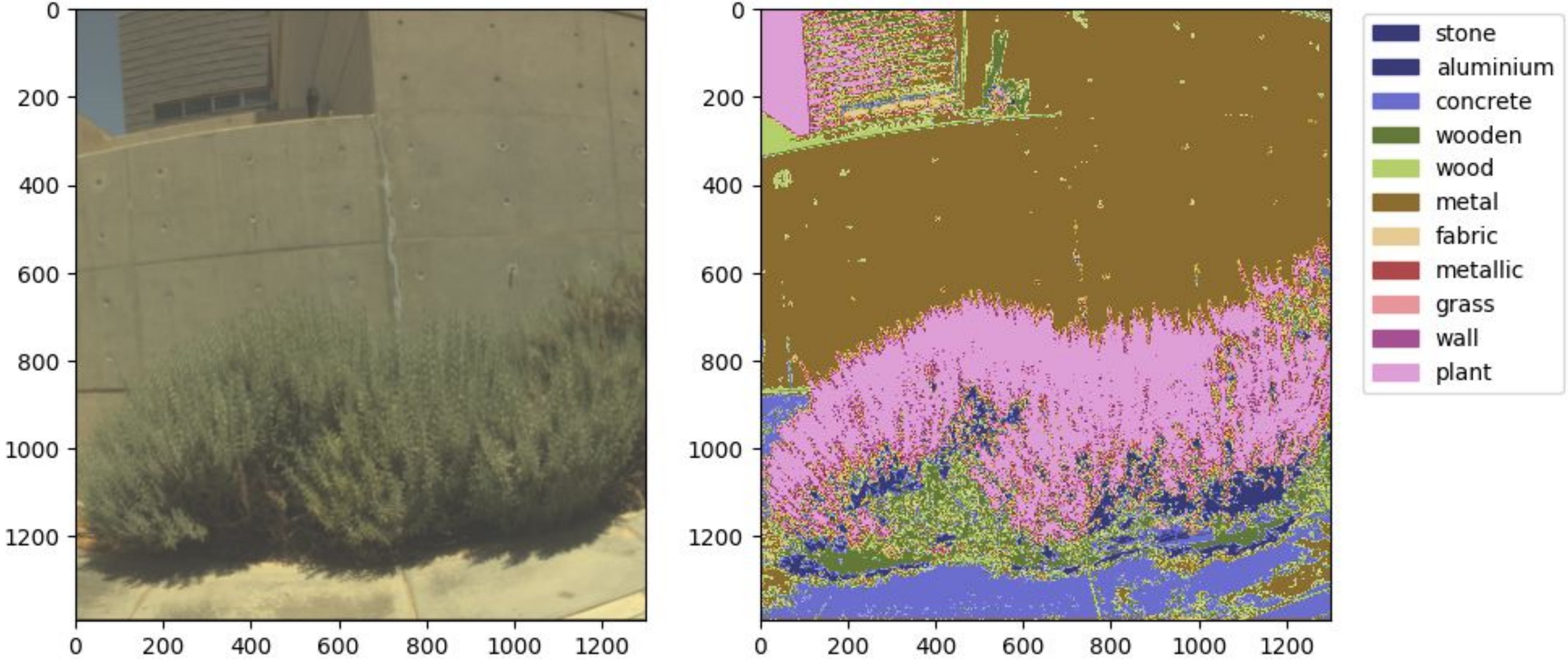



Library spectra
Reflectance
0.3
0.2
0.1
Plastic A
Plastic B
Plastic C
0.50 0.55 0.60 0.65 0.70 0.75 0.80
Wavelength (μm)
Hyperspectral image cube
Spectral distance measure with SAM and EMD algorithms
Hyperspectral detections

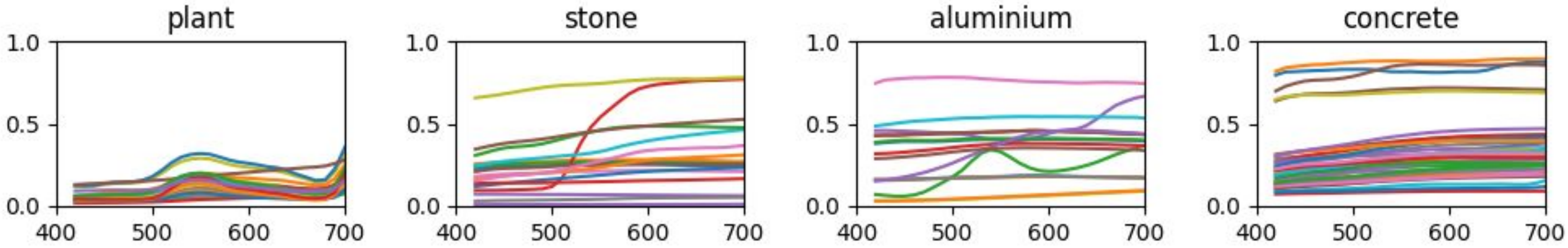


A mayor ángulo menor similitud

A menor ángulo mayor similitud



0
200
400
600
800
1000
1200
0
200
400
600
800
1000
1200
stone
aluminium
concrete
wooden
wood
metal
fabric
metallic
grass
wall
plant


plant
stone
aluminium
concrete
1.0
0.5
0.0
400
500
600
700

Beyond Appearances: Material Segmentation with Embedded Spectral Information from RGB-D imagery

Fabian Perez
Universidad Industrial de Santander
nelson2200183@correo.uis.edu.co

Hoover Rueda-Chacón
Universidad Industrial de Santander
hfarueda@uis.edu.co

Abstract

In the realm of computer vision, material segmentation of natural scenes represents a challenge, driven by the complex and diverse appearances of materials. Traditional approaches often rely on RGB images, which can be deceptive given the variability in appearances due to different lighting conditions. Other methods, that employ polarization or spectral imagery, offer a more reliable material differentiation but their cost and accessibility restrict their everyday usage. In this work, we propose a deep learning framework that bridges the gap between high-fidelity material segmentation and the practical constraints of data acquisition. Our approach leverages a training strategy that employs a paired RGBD-spectral data to incorporate spectral information directly within the neural network. This encoding process is facilitated by a Spectral Feature Mapper (SFM) layer, a novel module that embeds unique spectral characteristics into the network, thus enabling the network to infer materials from standard RGB-D images. Once trained, the model allows to conduct material segmentation on widely available devices without the need for direct spectral data input. In addition, we generate the 3D point cloud from the RGB-D image pair, to provide a richer spatial context for scene understanding. Through simulations using available datasets, and real experiments conducted with an iPad Pro, our method demonstrates superior performance in material segmentation compared to other methods. Code is available at: https://github.com/Factral/Spectral-material-segmentation

1. Introduction

Image segmentation consists on classifying pixels into multiple homogeneous regions, with each region exhibiting

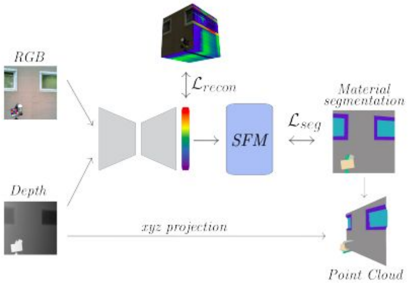
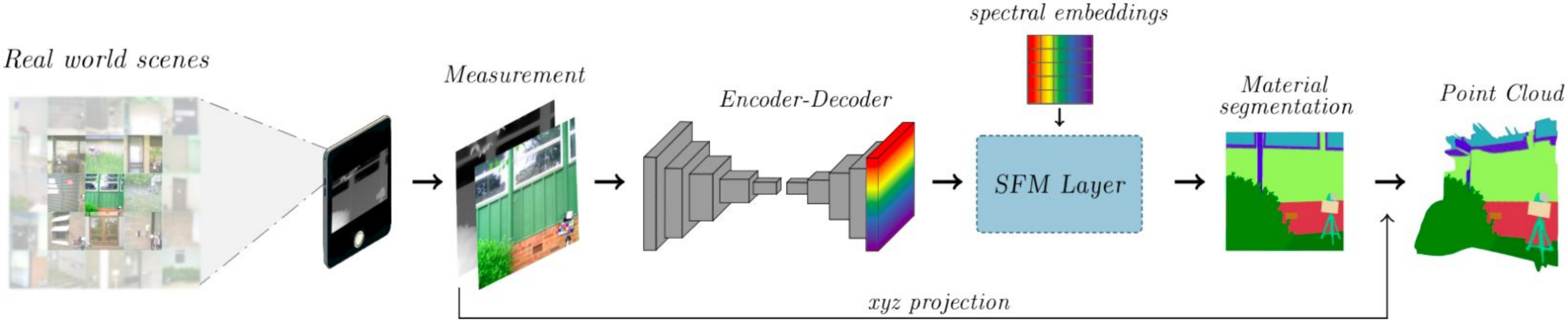


Figure 1. Overview of the training process and resulting output using an RGB-D input. The system employs $\mathcal{L}_{recon}$ for reconstruction of the spectral cube and $\mathcal{L}_{seg}$ for optimizing the segmentation task. A point cloud is then created using an xyz projection.

as it provides the foundational blocks for various applications such as haptics [9], robotic navigation [31], acoustic simulation [2], and grasping [5]. Unlike semantic segmentation, material segmentation is more challenging since spectral reflectance signatures of objects are preferred over color information, for high reliability. However, acquiring spectral information is not an easy nor cheap task, and its mainstream usage is still restricted to laboratories or remote sensing platforms.

In contrast, RGB images are ubiquitous, and color sensors are within the reach of our hand. Nonetheless, extracting material from just 3 channels is challenging, if not impossible, and unreliable. With the advent of deep learning, some alternatives were proposed to do material segmentation from RGB images. Examples include UPerNet



Real world scenes
Measurement
Encoder-Decoder
spectral embeddings
SFM Layer
Material segmentation
Point Cloud
xyz projection

Hands-on:
Segmentation