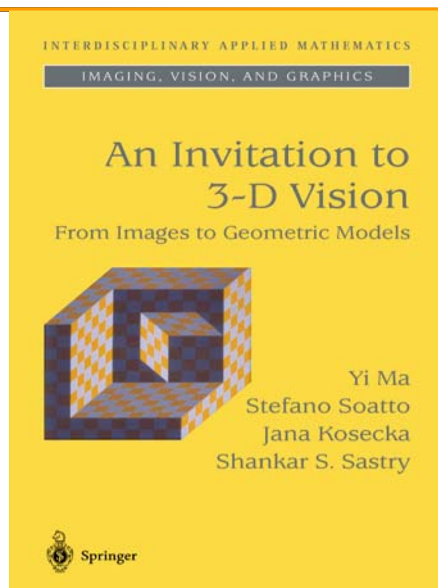
Course 23: Multiple-View Geometry For Image-Based Modeling

Jana Kosecka (CS, GMU)
Yi Ma (ECE, UIUC)
Stefano Soatto (CS, UCLA)
Rene Vidal (Berkeley, John Hopkins)

PRIMARY REFERENCE



Lecture A:
Overview and Introduction

Stefano Soatto

Computer Science Department, ULCA
http://www.cs.ucla.edu/~soatto

COURSE LECTURES

A. Overview and introduction (Soatto)
B. Preliminaries: geometry & image formation (Soatto)
C. Image primitives & correspondence (Soatto)
D. Two-view geometry (Kosecka)
E. Uncalibrated geometry and stratification (Ma)
F. Multiview recon. from points and lines (Vidal)
G. Reconstruction from scene knowledge (Ma)
H. Step-by-step building of 3-D model (Kosecka)
I. Multiple motion estimation & segmentation (Vidal)

COURSE LOGICAL FLOW

IMAGING and VISION: From 3D to 2D and then back again

GEOMETRY FOR TWO VIEWS

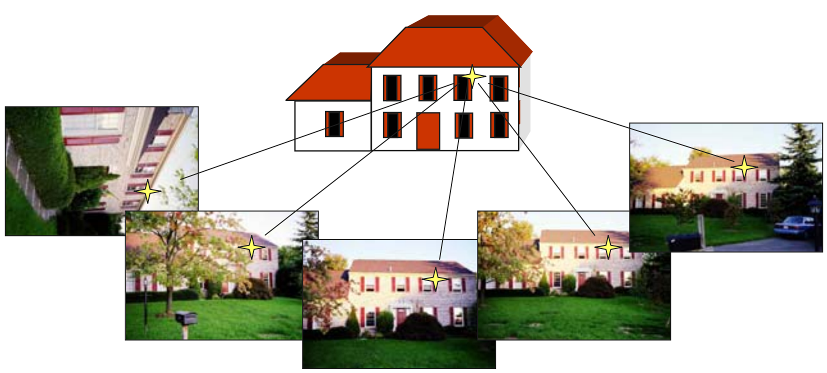
GEOMETRY FOR MULTIPLE VIEWS

MULTIVIEW GEOMETRY WITH SYMMETRY

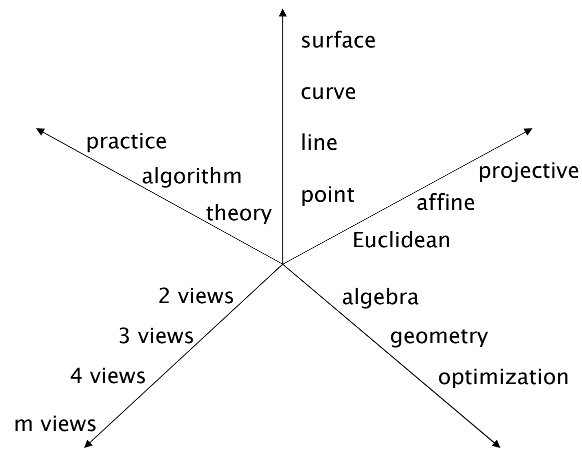
APPLICATIONS: System Implementation, motion segmentation.

Reconstruction from images – **The Fundamental Problem**

Input: Corresponding "features" in multiple perspective images.
Output: Camera pose, calibration, scene structure representation.



Fundamental Problem: An Anatomy of Cases



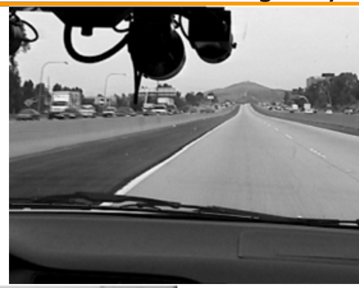
VISION AND GEOMETRY

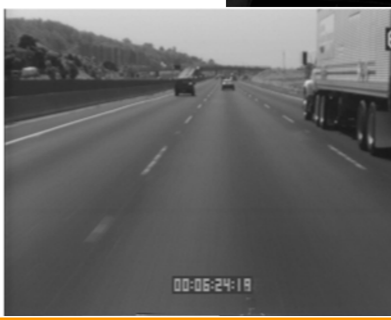
"The rise of projective geometry made such an overwhelming impression on the geometers of the first half of the nineteenth century that they tried to fit all geometric considerations into the projective scheme. ... The dictatorial regime of the projective idea in geometry was first successfully broken by the German astronomer and geometer Mobius, but the classical document of the democratic platform in geometry establishing the group of transformations as the ruling principle in any kind of geometry and yielding equal rights to independent consideration to each and any such group, is F. Klein's Erlangen program."

-- Herman Weyl, *Classic Groups, 1952*

Synonyms: Group = Symmetry

APPLICATIONS – Autonomous Highway Vehicles





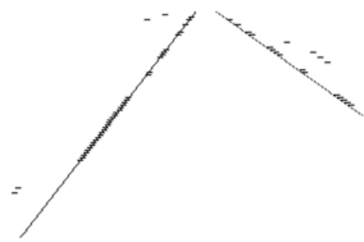
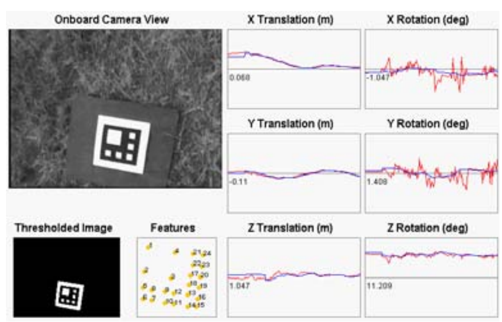


Image courtesy of California PATH

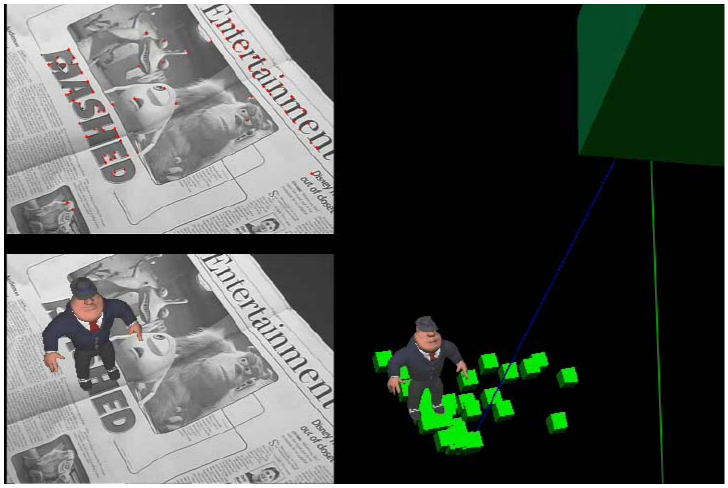
APPLICATIONS – Unmanned Aerial Vehicles (UAVs)



Rate: 10Hz; Accuracy: 5cm, 4°

Courtesy of Berkeley Robotics Lab

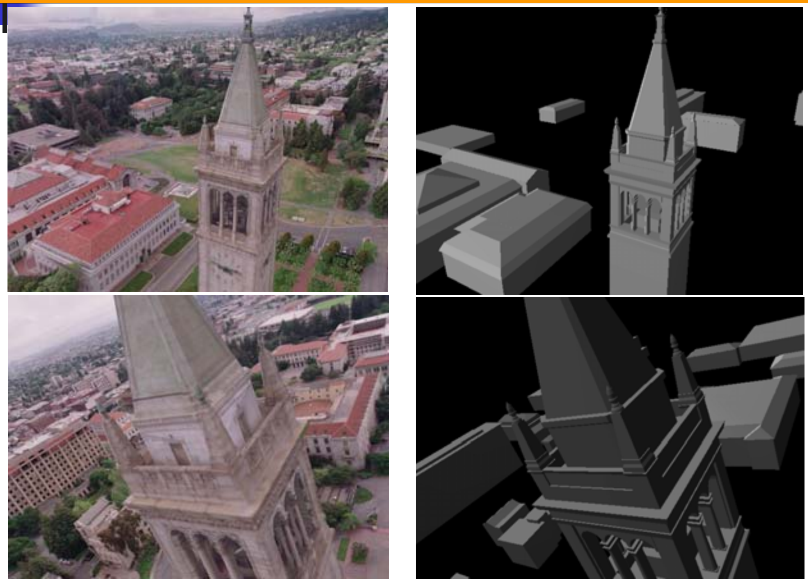
APPLICATIONS – Real-Time Virtual Object Insertion



APPLICATIONS – Real-Time Sports Coverage

First-down line and virtual advertising

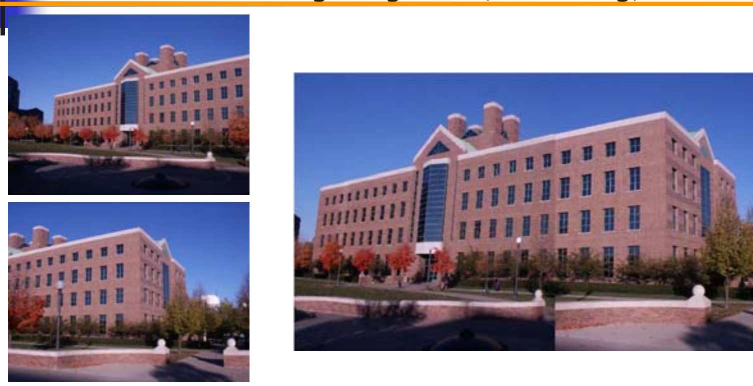
APPLICATIONS – Image Based Modeling and Rendering



August 8, 2004 SIGGRAPH'04, Los Angeles

Image courtesy of Paul Debevec

APPLICATIONS – Image Alignment, Mosaicing, and Morphing

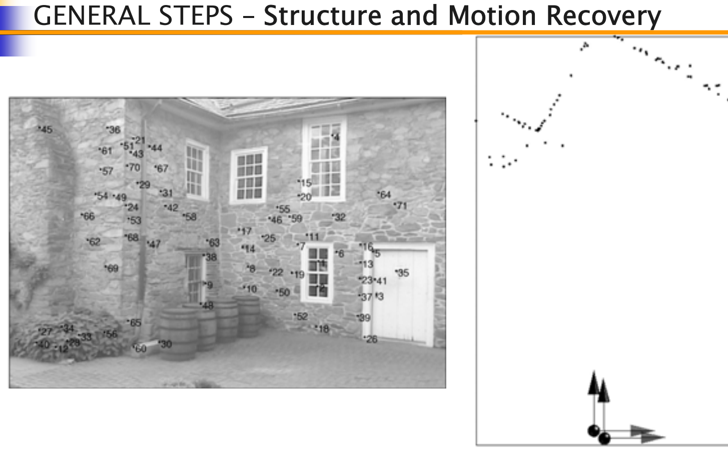


August 8, 2004 SIGGRAPH'04, Los Angeles

GENERAL STEPS – **Feature Selection and Correspondence**

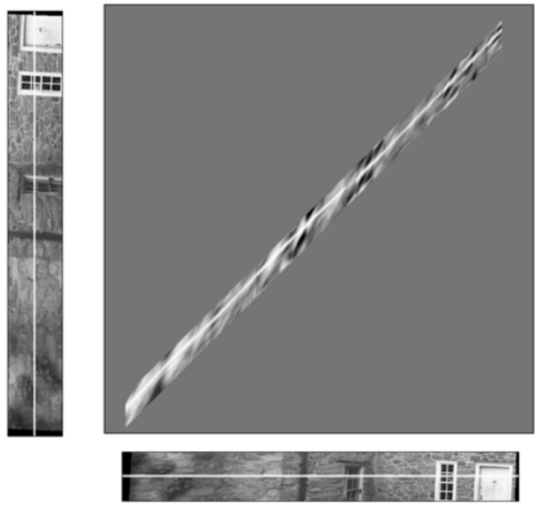
1. Small baselines versus large baselines
2. Point features versus line features

GENERAL STEPS – **Structure and Motion Recovery**

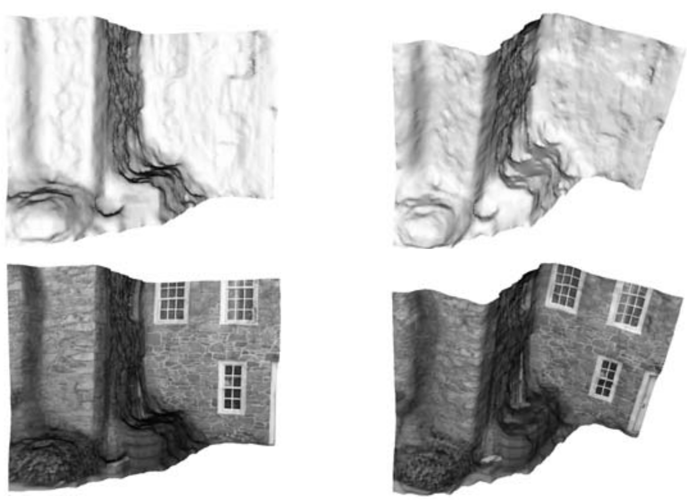


1. Two views versus multiple views
2. Discrete versus continuous motion
3. General versus planar scene
4. Calibrated versus uncalibrated camera
5. One motion versus multiple motions

GENERAL STEPS – Image Stratification and Dense Matching



GENERAL STEPS – 3-D Surface Model and Rendering



1. Point clouds versus surfaces (level sets)
2. Random shapes versus regular structures

Lecture B: Rigid-Body Motion and Imaging Geometry

Stefano Soatto

Computer Science Department, ULCA
http://www.cs.ucla.edu/~soatto

OUTLINE

3-D EUCLIDEAN SPACE & RIGID-BODY MOTION
- Coordinates and coordinate frames
- Rigid-body motion and homogeneous coordinates

GEOMETRIC MODELS OF IMAGE FORMATION
- Pinhole camera model

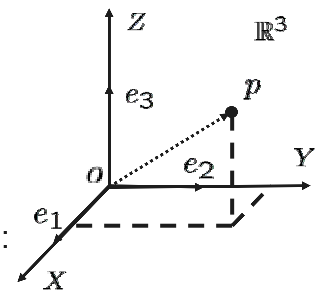
CAMERA INTRINSIC PARAMETERS
- From metric to pixel coordinates

SUMMARY OF NOTATION

3-D EUCLIDEAN SPACE – Cartesian Coordinate Frame

Standard base vectors:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$



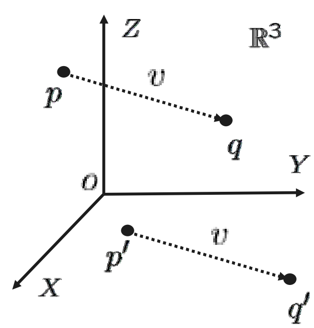
Coordinates of a point p in space:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{R}^3$$

3-D EUCLIDEAN SPACE – Vectors

A "free" vector is defined by a pair of points (p, q):

$$\boldsymbol{X}_p = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \in \mathbb{R}^3, \ \boldsymbol{X}_q = \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} \in \mathbb{R}^3,$$



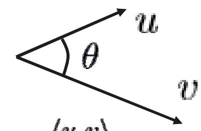
Coordinates of the vector v :

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} X_2 - X_1 \\ Y_2 - Y_1 \\ Z_2 - Z_1 \end{bmatrix} \in \mathbb{R}^3$$

3-D EUCLIDEAN SPACE - Inner Product and Cross Product

Inner product between two vectors:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$



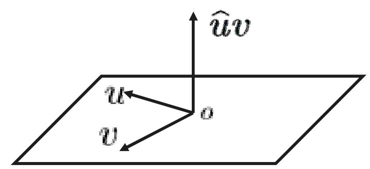
$$\langle u, v \rangle \doteq u^T v = u_1 v_1 + u_2 v_2 + u_3 v_3 \qquad \cos(\theta) = \frac{\langle u,v \rangle}{\|u\|\|v\|}$$

$$\|u\| \doteq \sqrt{u^T u} = \sqrt{u_1^2 + u_2^2 + u_3^3}$$

Cross product between two vectors:

$$u \times v \doteq \widehat{u} v, \quad u, v \in \mathbb{R}^3$$

$$\widehat{u} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$$

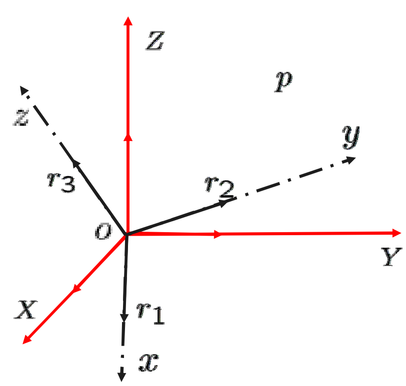


RIGID-BODY MOTION - Rotation

Rotation matrix:

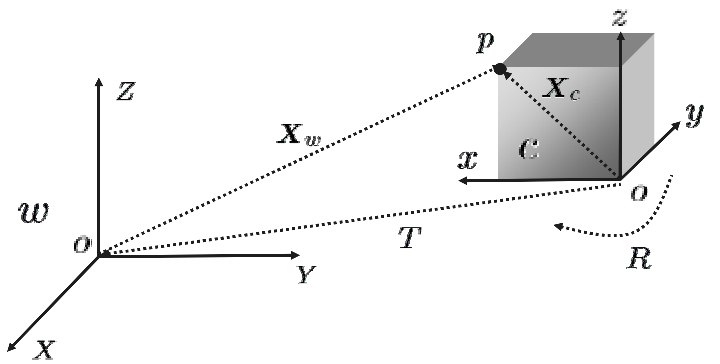
$$R \doteq [r_1, r_2, r_3] \in \mathbb{R}^{3 \times 3}$$

$$R^T R = I, \ \det(R) = +1$$



Coordinates are related by: $\boldsymbol{X}_c = R \boldsymbol{X}_w$

RIGID-BODY MOTION – Rotation and Translation



Coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$

Velocities are related by: $\dot{\boldsymbol{X}}_c = \hat{\omega}\boldsymbol{X}_c + v.$

RIGID-BODY MOTION – Homogeneous Coordinates

3-D coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$

Homogeneous coordinates:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \rightarrow \quad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \in \mathbb{R}^4,$$

Homogeneous coordinates/velocities are related by:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \qquad \begin{bmatrix} \dot{X}_c \\ \dot{Y}_c \\ \dot{Z}_c \\ 1 \end{bmatrix} = \begin{bmatrix} \hat{\omega} & v \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

IMAGE FORMATION – **Perspective Imaging**

"The Scholar of Athens," Raphael, 1518

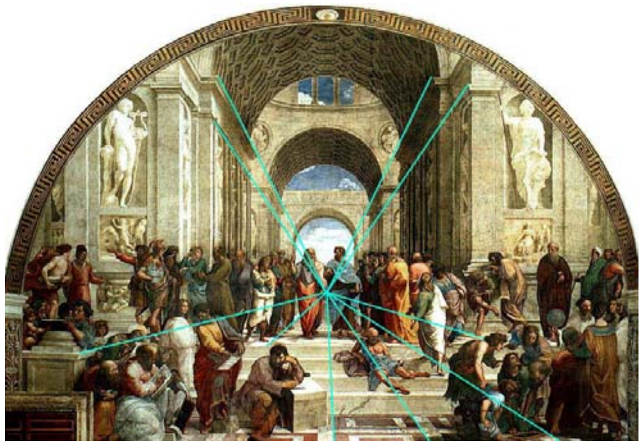
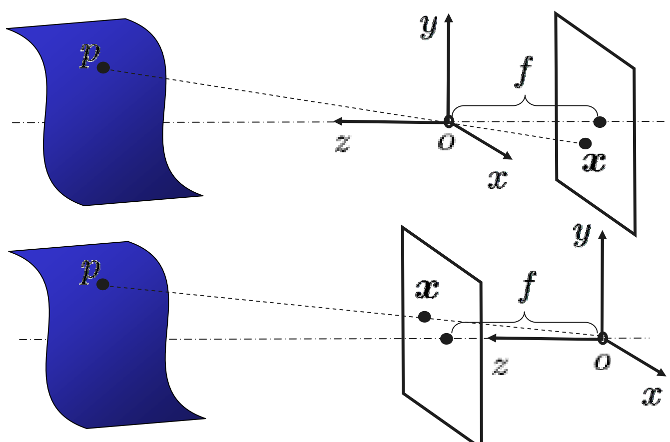


Image courtesy of C. Taylor

IMAGE FORMATION – **Pinhole Camera Model**

Pinhole

Frontal pinhole



$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \to \boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z}\begin{bmatrix} X \\ Y \end{bmatrix}$$

IMAGE FORMATION – **Pinhole Camera Model**

2-D coordinates $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$

Homogeneous coordinates

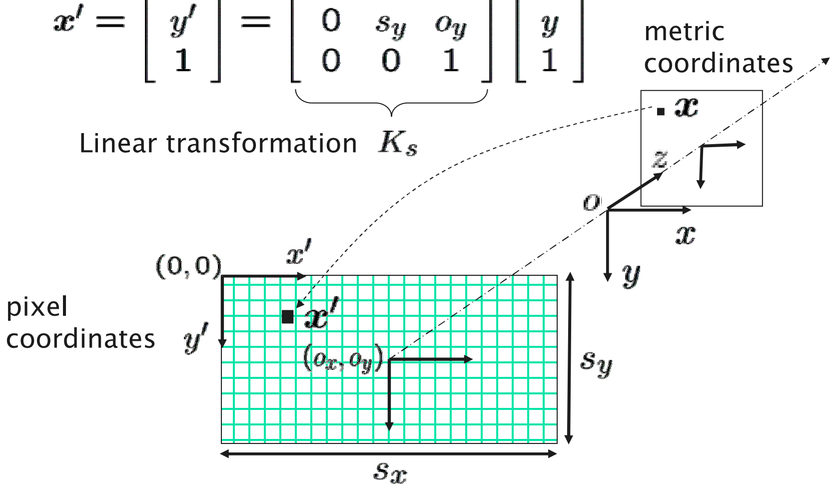
$$\boldsymbol{x} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}, \quad \boldsymbol{X} \rightarrow \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{K_f} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

CAMERA PARAMETERS – **Pixel Coordinates**

$$\boldsymbol{x}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K_s} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Linear transformation K_s



CAMERA PARAMETERS – Calibration Matrix and Camera Model

Pinhole camera | Pixel coordinates

$$\lambda \boldsymbol{x} = K_f \Pi_0 \boldsymbol{X} \qquad \boldsymbol{x}' = K_s \boldsymbol{x}$$

$$\lambda \boldsymbol{x}' = K_s K_f \Pi_0 \mathbf{X} = \underbrace{\begin{bmatrix} f s_x & f s_\theta & o_x \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

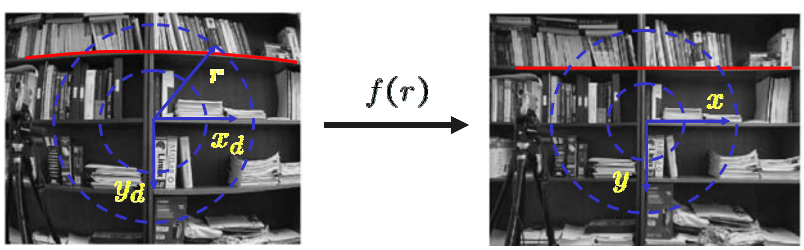
Calibration matrix (intrinsic parameters) $\underbrace{K = K_s K_f \qquad \Pi_0}$

Projection matrix $\Pi = [K, 0] \in \mathbb{R}^{3 \times 4}$

Camera model $\lambda \boldsymbol{x}' = K \Pi_0 \boldsymbol{X} = \Pi \boldsymbol{X}$

CAMERA PARAMETERS – Radial Distortion

Nonlinear transformation along the radial direction



$$\begin{aligned} \boldsymbol{x} &= c + f(r)(\boldsymbol{x}_d - c), \quad r = \|\boldsymbol{x}_d - c\| \\ f(r) &= 1 + a_1 r + a_2 r^2 + a_3 r^3 + a_4 r^4 + \cdots \end{aligned}$$

Distortion correction: make lines straight

IMAGE FORMATION – **Image of a Point**

Homogeneous coordinates of a 3-D point p

$$\boldsymbol{X} = [X, Y, Z, W]^T \in \mathbb{R}^4, \quad (W = 1)$$

Homogeneous coordinates of its 2-D image

$$\boldsymbol{x} = [x, y, z]^T \in \mathbb{R}^3, \quad (z = 1)$$

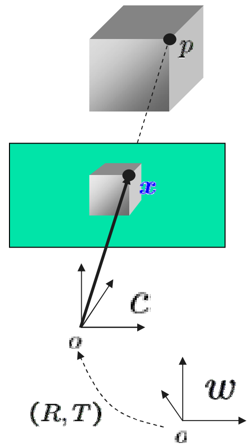
Projection of a 3-D point to an image plane

$$\lambda \boldsymbol{x} = \Pi \boldsymbol{X}$$

$$\lambda \in \mathbb{R}, \ \Pi = [R, T] \in \mathbb{R}^{3\times 4}$$

$$\lambda \boldsymbol{x}' = \Pi \boldsymbol{X}$$

$$\lambda \in \mathbb{R}, \ \Pi = [KR, KT] \in \mathbb{R}^{3\times 4}$$



August 8, 2004 SIGGRAPH'04, Los Angeles

IMAGE FORMATION – **Image of a Line**

Homogeneous representation of a 3-D line L

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{bmatrix} + \mu \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \end{bmatrix}, \quad \mu \in \mathbb{R}$$

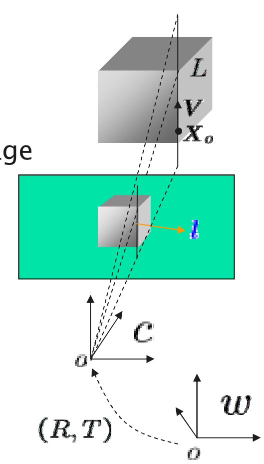
Homogeneous representation of its 2-D image

$$\boldsymbol{l} = [a, b, c]^T \in \mathbb{R}^3$$

Projection of a 3-D line to an image plane

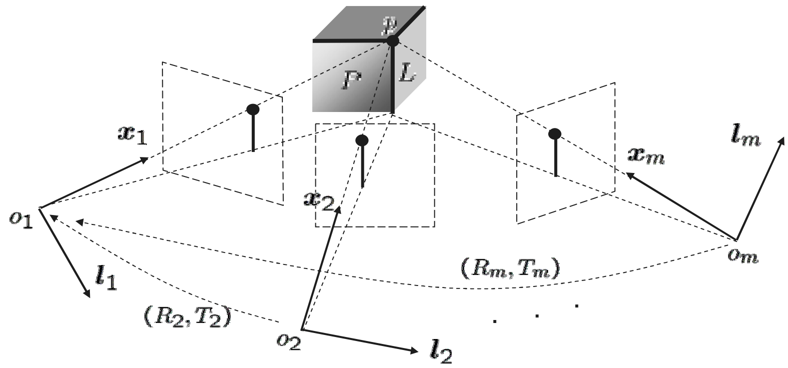
$$\boldsymbol{l}^T \boldsymbol{x} = \boldsymbol{l}^T \Pi \boldsymbol{X} = 0$$

$$\Pi = [KR, KT] \in \mathbb{R}^{3\times 4}$$



August 8, 2004 SIGGRAPH'04, Los Angeles

SUMMARY OF NOTATION – Multiple Images



1. Images are all "incident" at the corresponding features in space;
2. Features in space have many types of incidence relationships;
3. Features in space have many types of metric relationships.