Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

MIN-Fakultät
Fachbereich Informatik
Arbeitsbereich SAV/BV (KOGS)

Image Processing 1 (IP1)
Bildverarbeitung 1

Lecture 18 – Motion Analysis 2

Winter Semester 2014/15

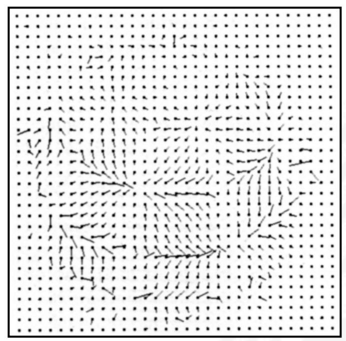
Dr. Benjamin Seppke
Prof. Siegfried Stiehl

Optical Flow and Segmentation

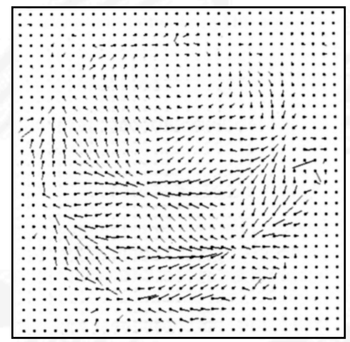
The optical flow smoothness constraint is not valid at occluding boundaries ("silhouettes"). In order to inhibit the constraint, one may try to segment the image based on optical flow discontinuities while performing the iterations.

(From B.K.P. Horn, Robot Vision, 1986)

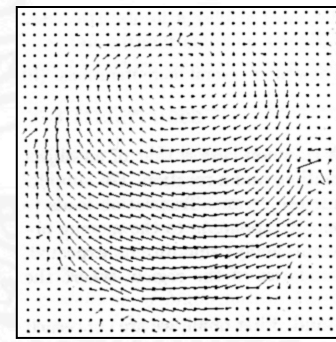
Checkered sphere rotating before randomly textured background



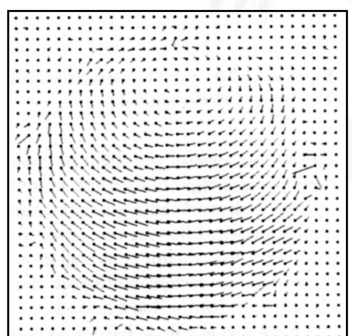
1. iteration



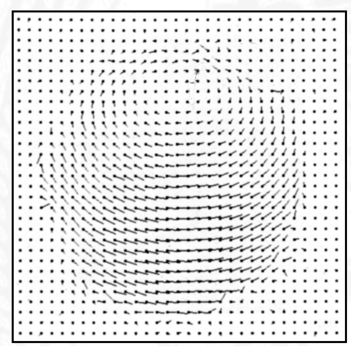
4. iteration



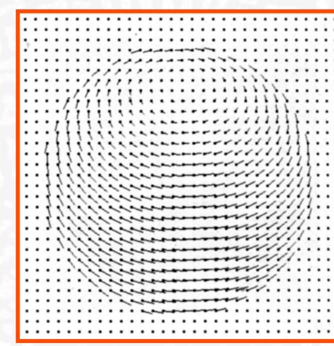
16. iteration



64. iteration



final result

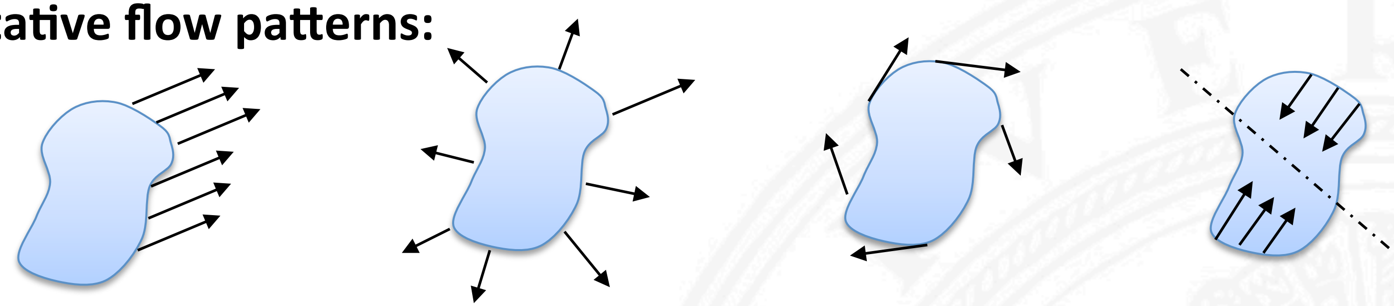


ideal result

Optical Flow Patterns

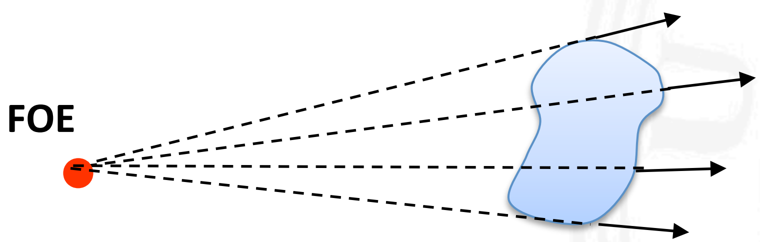
Complex optical flow fields may be segmented into components which show a consistent qualitative pattern.

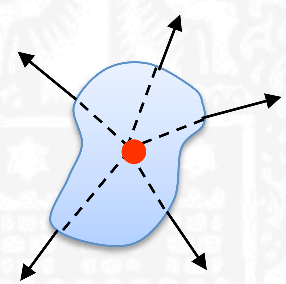
Qualitative flow patterns:



translation at constant distance | **translation in depth** | **rotation at constant distance** | **rotation about axis parallel to image plane**

General translation results in a flow pattern with a focus of expansion (FOE):



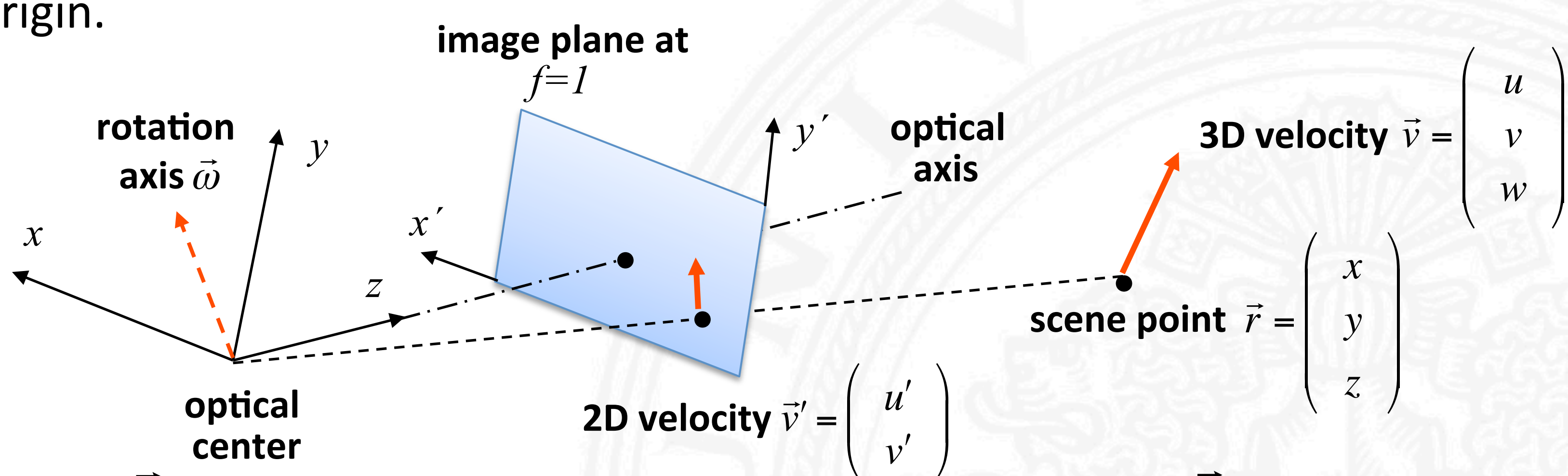


As the direction of motion changes, the FOE changes its location.

Optical Flow and 3D Motion I

In general, optical flow may be caused by an unknown 3D motion of an unkown surface. How do the flow components u', v' depend on the 3D motion parameters?

Assume camera motion in a static scene, optical axis = z-axis, rotation about the origin.



3D velocity $\vec{v}$ of a point $\vec{r}$ is determined by rotational velocity $\vec{\omega}$ and translational velocity $\vec{t}$:

$$\vec{v} = -\vec{t} - \vec{\omega} \times \vec{r}$$

Optical Flow and 3D Motion II

By taking the component form of $\vec{v} = -\vec{t} - \vec{\omega} \times \vec{r}$ with $\vec{t}^T = \begin{pmatrix} t_x & t_y & t_z \end{pmatrix}$, $\vec{\omega}^T = \begin{pmatrix} a & b & c \end{pmatrix}$ and $\vec{r}^T = \begin{pmatrix} x & y & z \end{pmatrix}$ and computing the perspective projection we get

$$u' = \frac{\dot{x}}{z} - \frac{x\dot{z}}{z^2} = \left(-\frac{t_x}{z} - b + cy' \right) - x' \left(-\frac{t_z}{z} - ay' + bx' \right)$$

$$v' = \frac{\dot{y}}{z} - \frac{y\dot{z}}{z^2} = \left(-\frac{t_y}{z} - cx' + a \right) - y' \left(-\frac{t_z}{z} - ay' + bx' \right)$$

Observation of u' and v' at location (x', y') gives 2 equations for 7 unknowns. Note that motion of a point at distance kz with translation $k\underline{t}$ and the same rotation ω will give the same optical flow, k any scale factor.

The translational and rotational parts may be separated:

$$u'_{translation} = -\frac{t_x + x' t_z}{z} \qquad \mathrm{u}'_{rotation} = \mathrm{ax'y'} - \mathrm{b(x'^2} + 1) + cy'$$

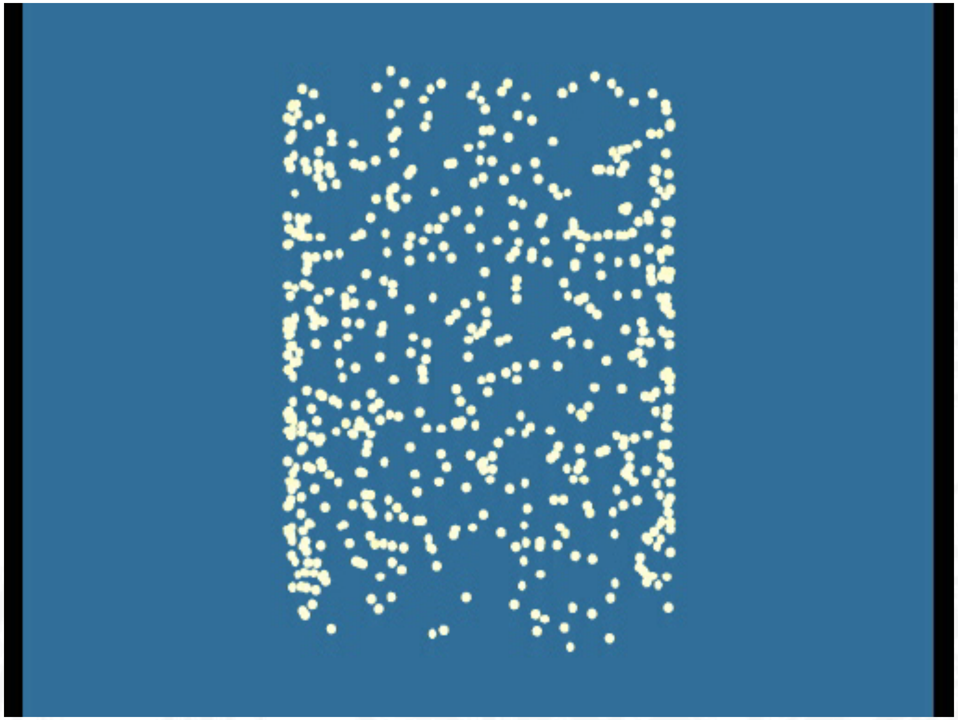
$$v'_{translation} = -\frac{t_y + y' t_z}{z} \qquad \mathrm{v}'_{rotation} = \mathrm{a(y'^2} + 1) - \mathrm{b}x'\mathrm{y}' + cx'$$

For pure translation we have 2 equations for 3 unknows (z fixed arbitrarily).

3D Motion Analysis Based on 2D Point Displacements

2D displacements of points observed on an unknown moving rigid body may provide information about

- the 3D structure of the points
- the 3D motion parameters



Rotating cylinder experiment by S. Ullman (1981)

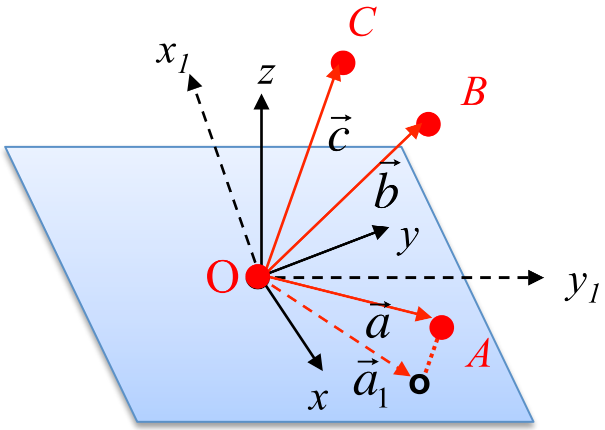
Cases of interest:

- stationary camera, moving object(s)
- moving camera, stationary object(s)
- moving camera, moving object(s)

} **camera motion parameters may be known**

Structure from Motion I

Ullman showed 1979 that the spatial structure of 4 rigidly connected non-coplanar points may be recovered from 3 orthographic projections.

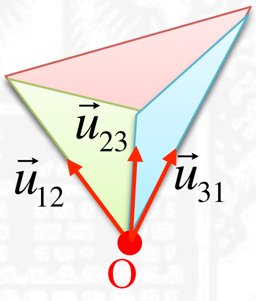


projection plane P_1

O, A, B, C	4 rigid points
$\vec{a}, \vec{b}, \vec{c}$	vectors to A, B, C
Π_1, Π_2, Π_3	projection planes
x_i, y_i	coordinate axes of P_i
$\vec{a}_i, \vec{b}_i, \vec{c}_i$	coordinate pairs of points A, B, C in projection plane Π_i

The problem is to determine the spatial orientations of Π_1, Π_2, Π_3 from the 9 projection coordinate pairs $\vec{a}_i, \vec{b}_i, \vec{c}_i$, $i = 1, 2, 3.$

The 3 projection planes intersect and form a tetrahedron.
$\vec{u}_{12}, \vec{u}_{23}, \vec{u}_{31}$ are unit vectors along the intersections.
The idea is to determine the $\vec{u}_{i,j}$ from the observed coordinates $\vec{a}_i, \vec{b}_i, \vec{c}_i$



Structure from Motion II

The projection coordinates are:

$$a_{1_x} = \vec{a}^T \vec{x}_1 \qquad a_{1_y} = \vec{a}^T \vec{y}_1$$

$$b_{1_x} = \vec{b}^T \vec{x}_1 \qquad b_{1_y} = \vec{b}^T \vec{y}_1$$

$$c_{1_x} = \vec{c}^T \vec{x}_1 \qquad c_{1_y} = \vec{c}^T \vec{y}_1$$

Since each $\vec{u}_{ij}$ lies in both planes Π_i and Π_j, it can be written as

$$\vec{u}_{ij} = \alpha_{ij}\vec{x}_i + \beta\vec{y}_i$$

$$\vec{u}_{ij} = \gamma_{ij}\vec{x}_j + \delta\vec{y}_j$$

$$\alpha_{ij}\vec{x}_i + \beta\vec{y}_i = \gamma_{ij}\vec{x}_j + \delta\vec{y}_j$$

Multiplying with $(\vec{a}_i)^T, (\vec{b}_i)^T, (\vec{c}_i)^T$ we get

$$\alpha_{ij}a_{i_x} + \beta a_{i_y} = \gamma_{ij}a_{j_x} + \delta a_{j_y}$$

$$\alpha_{ij}b_{i_x} + \beta b_{i_y} = \gamma_{ij}b_{j_x} + \delta b_{j_y}$$

$$\alpha_{ij}c_{i_x} + \beta c_{i_y} = \gamma_{ij}c_{j_x} + \delta c_{j_y}$$

Solve for α_{ij}, β_{ij}, γ_{ij}, δ_{ij}. using the constraints $\alpha_{ij}^2 + \beta_{ij}^2 = 1$ and $\gamma_{ij}^2 + \delta_{ij}^2 = 1$

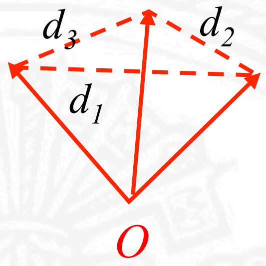
Structure from Motion III

From the coefficients $\alpha_{ij}, \beta_{ij}, \gamma_{ij}, \delta_{ij}$ one can compute the distances between the 3 unit vectors $\vec{u}_{12}, \vec{u}_{23}, \vec{u}_{31}$:

$$d_1 = \left\| \vec{u}_{23} - \vec{u}_{12} \right\| = \left\| (\alpha_{23} - \alpha_{12}) \vec{x}_i + (\beta_{23} - \beta_{12}) \vec{y}_i \right\| = (\alpha_{23} - \alpha_{12})^2 + (\beta_{23} - \beta_{12})^2$$

$$d_2 = (\alpha_{31} - \alpha_{23})^2 + (\beta_{31} - \beta_{23})^2$$

$$d_3 = (\alpha_{12} - \alpha_{31})^2 + (\beta_{12} - \beta_{31})^2$$

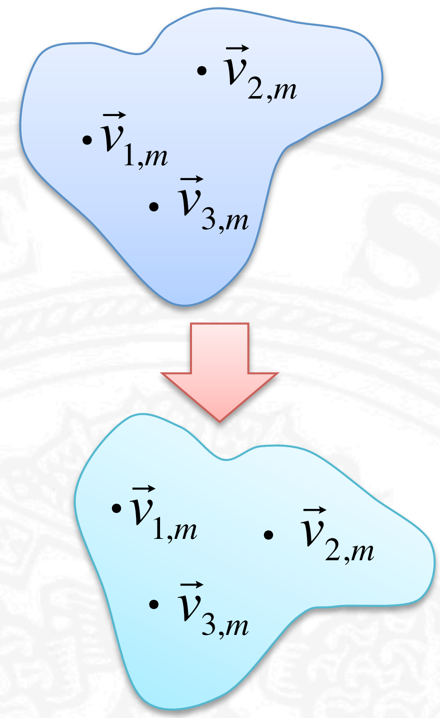


Hence the relative angles of the projection planes are determined.

The spatial positions of A, B, C relative to the projection planes (and to the origin O) can be determined by intersecting the projection rays perpendicular on the projected points $\vec{a}_i, \vec{b}_i, \vec{c}_i$.

Perspective 3D Analysis of Point Displacements

- relative motion of one rigid object and one camera
- observation of P points in M views



For each point $\vec{v}_p$ in 2 consecutive images we have:

$\vec{v}_{p,m+1} = R_m \vec{v}_{p,m} + \vec{t}_m$ motion equation

$\vec{v}_{p,m} = \lambda_{p,m} \vec{v}'_{p,m}$ projection equation

For P points in M images we have

$3MP$	unknown 3D point coordinates $\vec{v}_{p,m}$
$6(M\text{-}1)$	unkown motion parameters R_m and $\vec{t}_m$
MP	unknown projection parameters $\lambda_{p,m}$
$3(M\text{-}1)P$	motion equations
$3MP$	projection equations
1	arbitrary scaling parameter

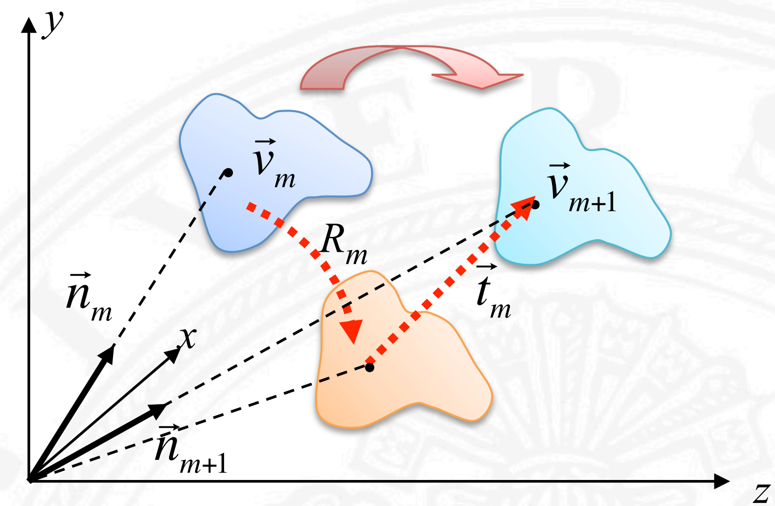
equations ≥ # unknowns → $P \geq 3 + \dfrac{2}{2M-3}$ →

M	P
2	5
3	4
4	4
5	4

Essential Matrix

Geometrical constraints derived from 2 views of a point in motion

- motion between image m and $m+1$ may be decomposed into
  1. rotation R_m about origin of coordinate system (= optical center)
  2. translation $\vec{t}_m$
- observations are given by direction vectors $\vec{n}_m$ and $\vec{n}_{m+1}$ along projection rays.



$R_m\vec{n}_m$, $\vec{t}_m$ and $\vec{n}_{m+1}$ are coplanar: $\left(\vec{t}_m \times R_m\vec{n}_m\right)^T \vec{n}_{m+1} = 0$

After some manipulation: $\vec{n}_m E_m n_{m+1} = 0$ $\quad E$ = essential matrix

with $E_m = \begin{pmatrix} | & | & | \\ \vec{t}_m \times \vec{r}_1 & \vec{t}_m \times \vec{r}_2 & \vec{t}_m \times \vec{r}_3 \\ | & | & | \end{pmatrix}$ and $R_m = \begin{pmatrix} | & | & | \\ \vec{r}_1 & \vec{r}_2 & \vec{r}_3 \\ | & | & | \end{pmatrix}$

Solving for the Essential Matrix

$\left(\vec{n}_m\right)^T E_m \vec{n}_{m+1} = 0$ formally one equation for 9 unknowns e_{ij}

But:

- only 6 degrees of freedom (3 rotation angles, 3 translation components)
- e_{ij} can only be determined up to a scale factor

Basic solution approach:

- observe P points in 2 views, $P >> 8$
- fix e_{11} arbitrarily
- solve an overconstrained system of equations for the other 8 unknown coefficients e_{ij}

E may be decomposed into S and R by Singular Value Decomposition (SVD).

E may be written as $E = S\,R^{-1}$ with R = rot. matrix and $S = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}$

Note: S (and therefore E) has rank 2

Singular Value Decomposition of E

Any $m \times n$ matrix A, $m \geq n$, may be decomposed as $A = U\,D\,V^T$ where

- U has orthonormal columns $m \times n$
- D is non-negative diagonal $n \times n$
- V^T has orthonormal rows $n \times n$

This can be applied to E to give $E = U\,D\,V^T$ with:

$$R = U\,G\,V^T \quad or \quad R = U\,G^T\,V^T$$

$$S = V\,Z\,V^T$$

where $G = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $Z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

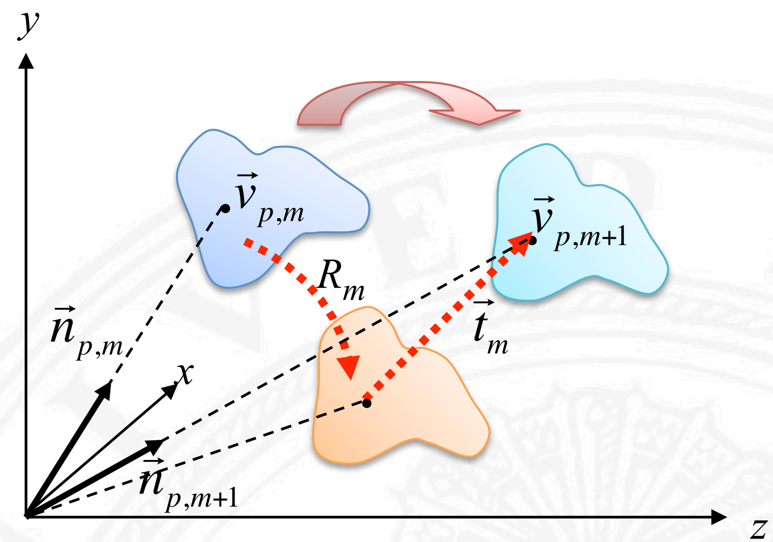
Alternative 3D Motion Constraint

Nagel and Neumann 82

Consider 2 views of 3 points $\vec{v}_{p,m}$, $p = 1 \ldots 3,\ m = 1, 2$

The planes through $R_m \vec{n}_{p,m}$ and $\vec{n}_{p,m+1}$ all intersect in $\vec{t}_m$

→ the normals of the planes are coplanar!



Coplanarity condition for 3 vectors $\vec{a}, \vec{b}, \vec{c}$: $\left(\vec{a} \times \vec{b}\right)^T \vec{c} = 0$ hence:

$$\left(\left(R_m \vec{n}_{1,m} \times \vec{n}_{1,m+1}\right) \times \left(R_m \vec{n}_{2,m} \times \vec{n}_{2,m+1}\right)\right)^T \left(R_m \vec{n}_{3,m} \times \vec{n}_{3,m+1}\right) = 0$$

Nonlinear equation with 3 unknown rotation parameters.

=> Observation of at least 5 points required to solve for the unknowns.

Reminder: Homogeneous Coordinates

- (N+1)-dimensional notation for points in N-dimensional Euclidean space
- allows to express projection and translation as linear operations

Normal coordinates: $\vec{v}^T = \begin{pmatrix} x & y & z \end{pmatrix}$

Homogeneous coordinates: $\vec{v}^T = \begin{pmatrix} wx & wy & wz & w \end{pmatrix}$ $w \neq 0$ **is arbitrary constant**

Rotation and translation in homogeneous coordinates:

$$\vec{v}' = A\vec{v} \quad \text{with} \quad A = \begin{pmatrix} R & \vec{t} \\ \vec{0} & 1 \end{pmatrix}$$

Projection in homogeneous coordinates:

$$\vec{v}' = B\vec{v} \quad \text{with} \quad B = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Divide the first N components by the (N+1)th component to recover normal coordinates

From Homogeneous World Coordinates to Homogeneous Image Coordinates

$\vec{v}^T = \begin{pmatrix} x & y & z \end{pmatrix}$ **scene coordinates**

$\left(\vec{v}_p\right)^T = \begin{pmatrix} x''_p & y''_p \end{pmatrix}$ **image coordinates**

$$\begin{pmatrix} wx''_p \\ wy''_p \\ w \end{pmatrix} = \begin{pmatrix} KR & K\vec{t} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = M \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad \Rightarrow \quad \vec{v}_p = M\vec{v}$$

$$K = \begin{pmatrix} fa & fb & x_{p_0} \\ 0 & fc & y_{p_0} \\ 0 & 0 & 1 \end{pmatrix}$$

intrinsic camera parameters ("camera calibration matrix")

fa **= scaling in** x_P**-axis**
fc **= scaling in** y_P**-axis**
fb **= slant of axes**
x_{P0}, y_{P0} **= "principal point" (optical center in image plane)**

$R, \vec{t}$ **extrinsic camera parameters**

$M = 3 \times 4$ **projective matrix**