

# Stereoanalyse af digitale billeder

Søren I. Olsen

## Oversigt

1. Datamatsyn og stereoanalyse
2. Stereobilleder og stereogeometri
3. Stereosyn hos mennesket
4. Korrespondanceproblemet

Problemer og simplificerende antagelser

5. Algoritmiske metoder

## Datamatsyn

I datamatsyn analyseres et eller flere (gråtone/farve) billeder med det formål at ekstrahere information om scenens rumlige form, bevægelse etc.

Form fra gråtoner	1 billede
Form fra tekstur	1 billede
Form fra stereo	2-n billeder
Form fra lyslægning	3-n billeder
Form fra bevægelse	2-n billeder
Form fra fokus	n billeder
Form fra zoom	2-n billeder

## Stereoanalyse

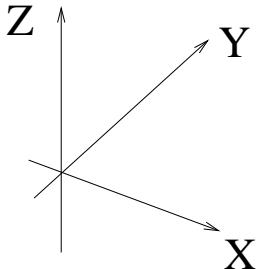
Analyse af 2 billeder af en scene optaget fra forskellig position.

Hvis kamera'enes absolutte position og orientering kendes kan scenen tredimensionale form og position (dybde) bestemmes ud fra korresponderende punkter i de to billeder.

1. Kamerakalibrering
2. Korrespondanceproblemet
3. 3D rekonstruktion af scene

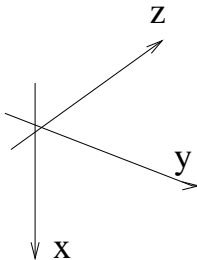
# Kamerakalibrering

## Verdenskoordinatsystem



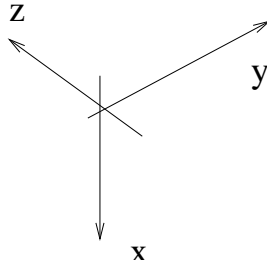
$F_L$

3D Rotation,  
Translation

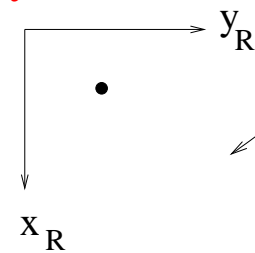
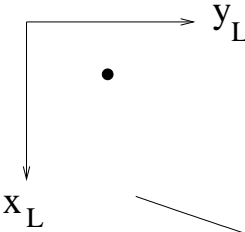


Perspektivisk projektion  
Aspekt ratio, brændvidde  
linse forvrængning etc.

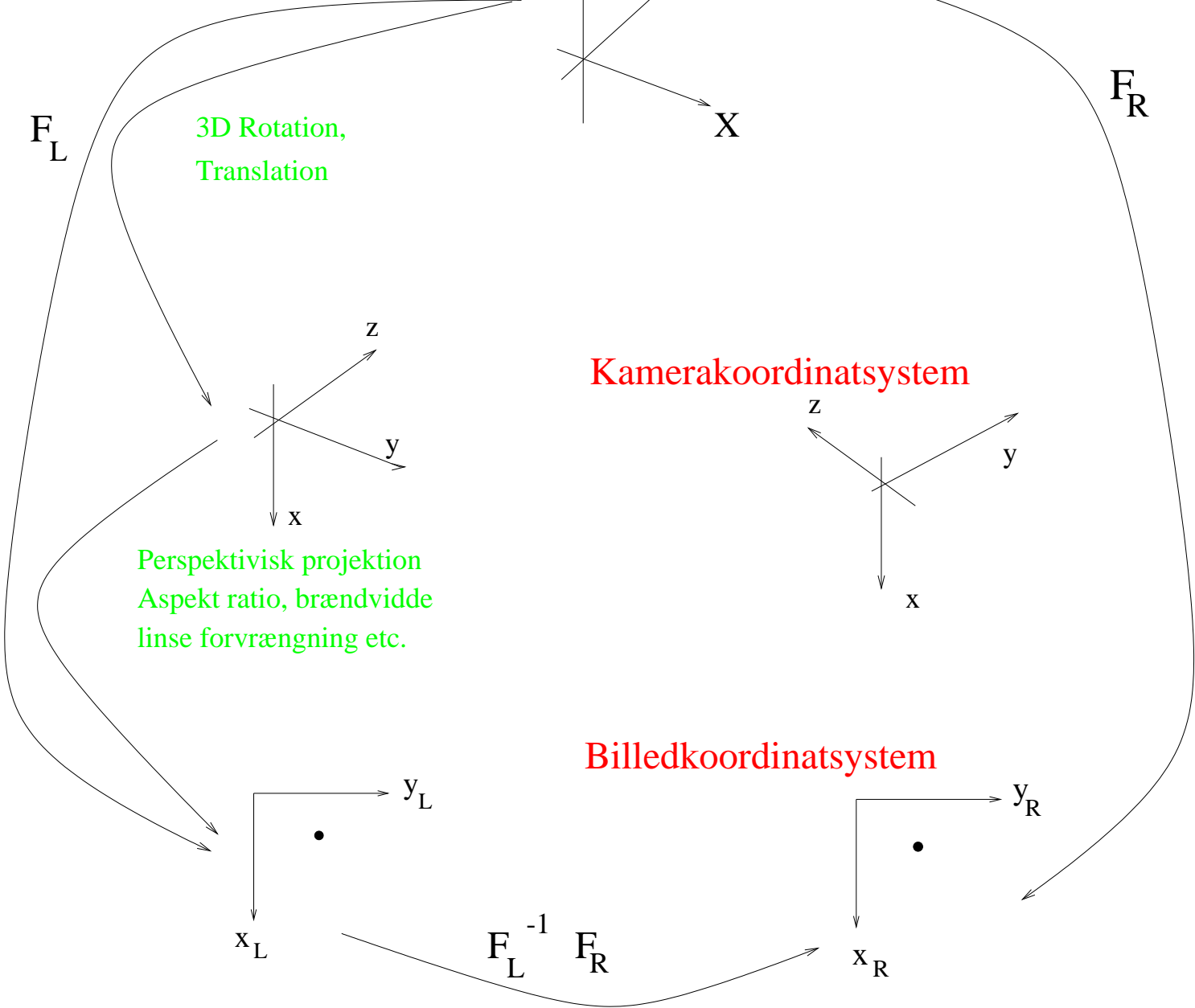
## Kamerakoordinatsystem



## Billedkoordinatsystem



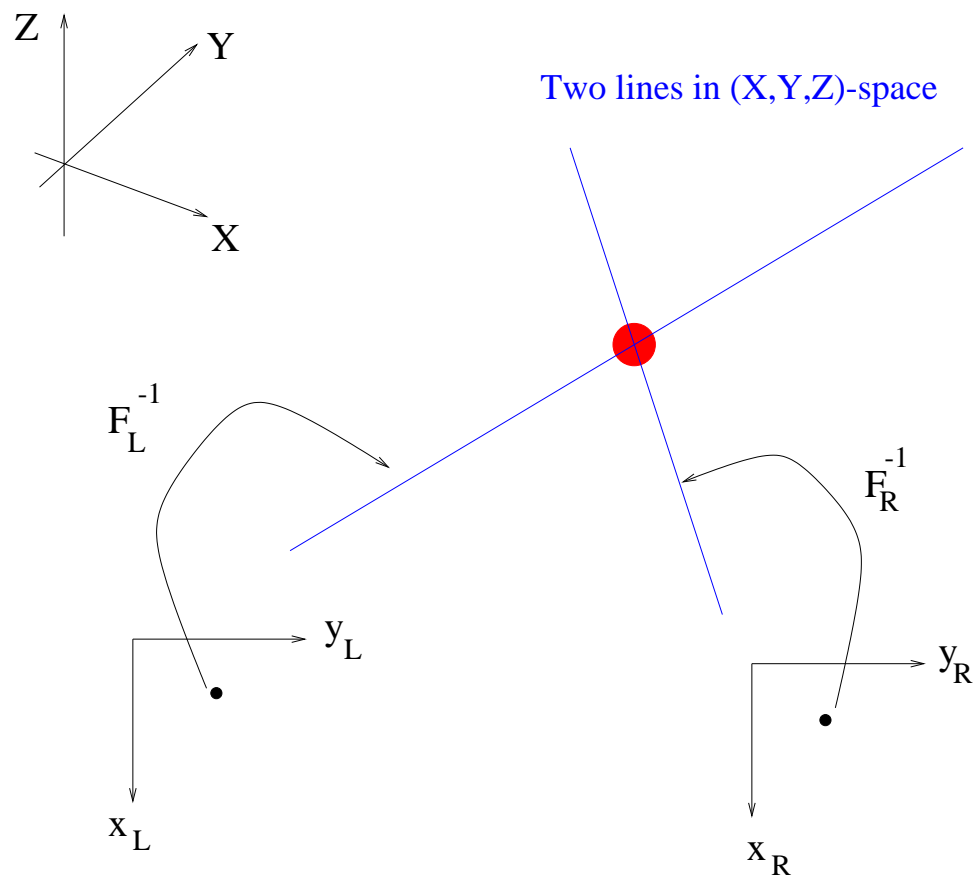
$F_L^{-1}$   $F_R$



## Korrespondanceproblemet

Givet billedpunkt  $(x_L, y_L)$  i det ene kamera, find billedpunkt  $(x_R, y_R)$  i det andet kamera således at de to billedpunkter afbilder samme fysiske scenepunkt.

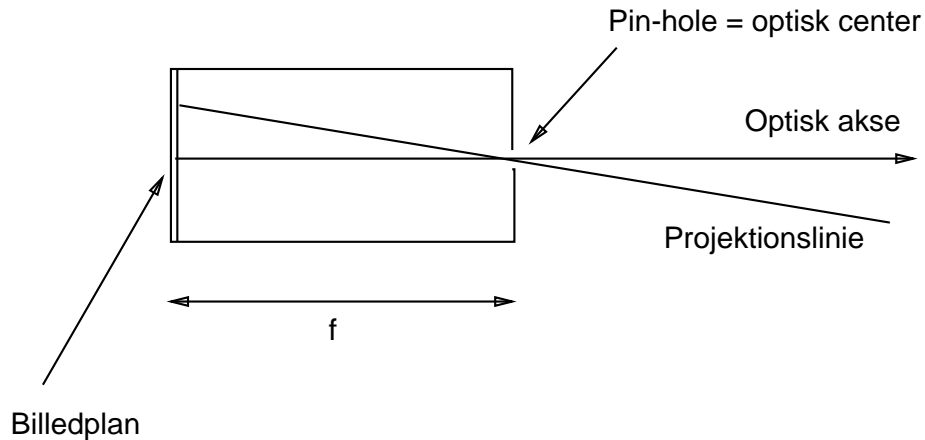
## 3D rekonstruktion



## Anvendelser

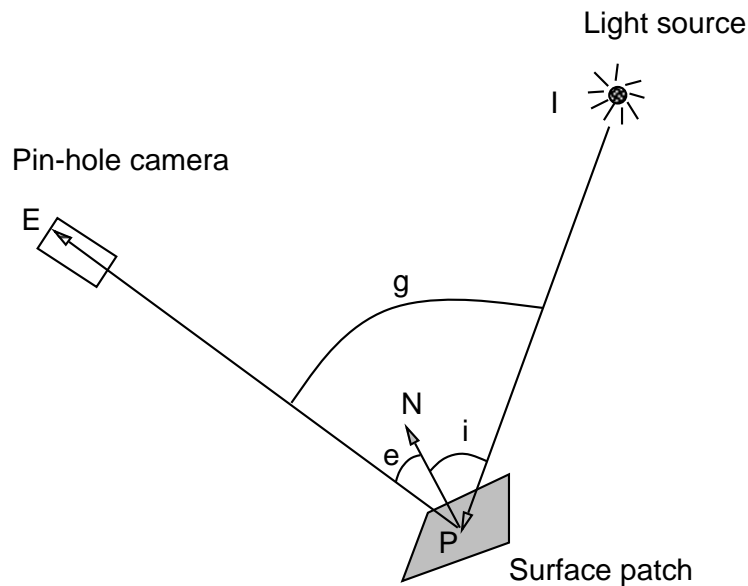
- Astronomi: Planetobservationer
- Topografisk kortlægning ud fra fly eller satellitbilleder, f.eks. i utilgængelige egne eller ved større anlægsarbejder.
- Navigation (f.eks. af selvkørende robotter og biler)
- Opmåling af truede bygningsværker og monumenter
- Underholdning

# Billeddannelse



Billeddannelsen er perspektivisk:  $(x, y) = \frac{f}{Z} (X, Y)$

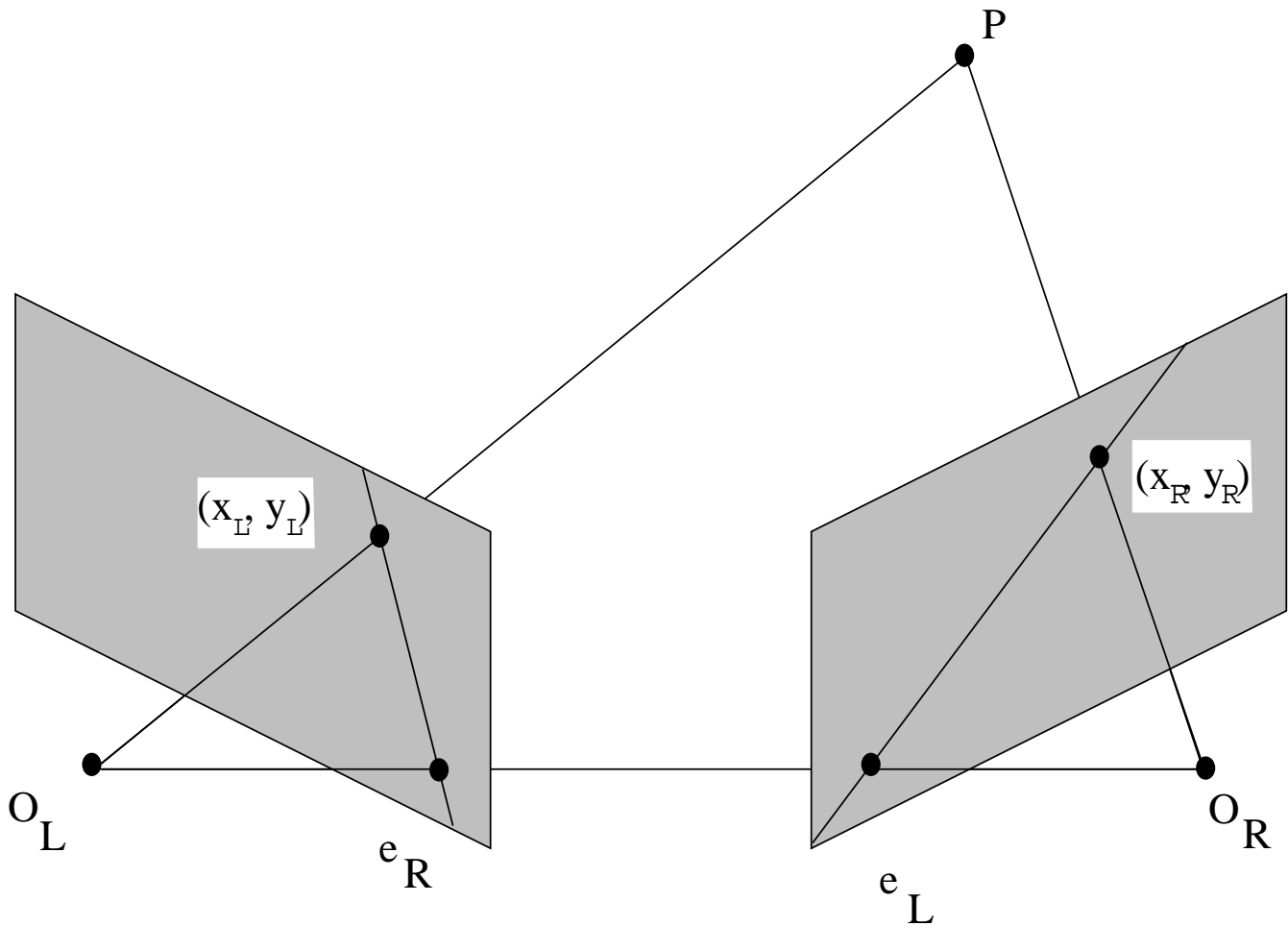
hvor  $(x, y)$  er billedkoordinater og  $(X, Y, Z)$  er verdenskoordinater.



1. Punktet **P** reflekterer forskellig mængde af lys mod de to kamera.
2. Form og størrelse af en projiceret fladeflis vil være forskellig på de to billedplaner.

## Stereogeometri - Epipolarlinier

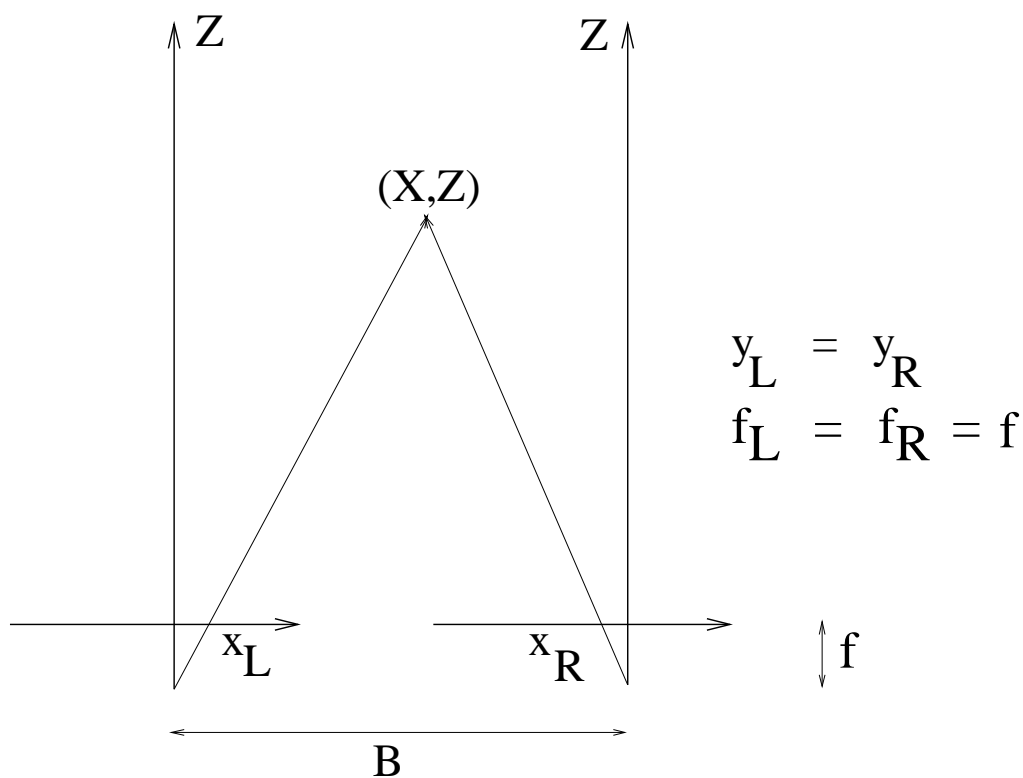
Uanset hvorledes de to kamera er positioneret og orienteret er korrespondanceproblemet 1-dimensionelt, dvs: Givet  $(x_L, y_L)$  vil  $(x_R, y_R)$  ligge på en ret linie i  $R$ -billedplanet.



Stereogeometri med konjugerede epipolarlinier



## Det simpleste tilfælde - Parallele optiske akser



**Disparitet:**  $(d_x, d_y)$

$$d_x = x_R - x_L = f \frac{X - B}{Z} - f \frac{X}{Z} = -f \frac{B}{Z}$$

I dette tilfælde bliver rekonstruktionen:

$$(X, Y, Z) = -\frac{B}{d_x} (x_L, y_L, f)$$

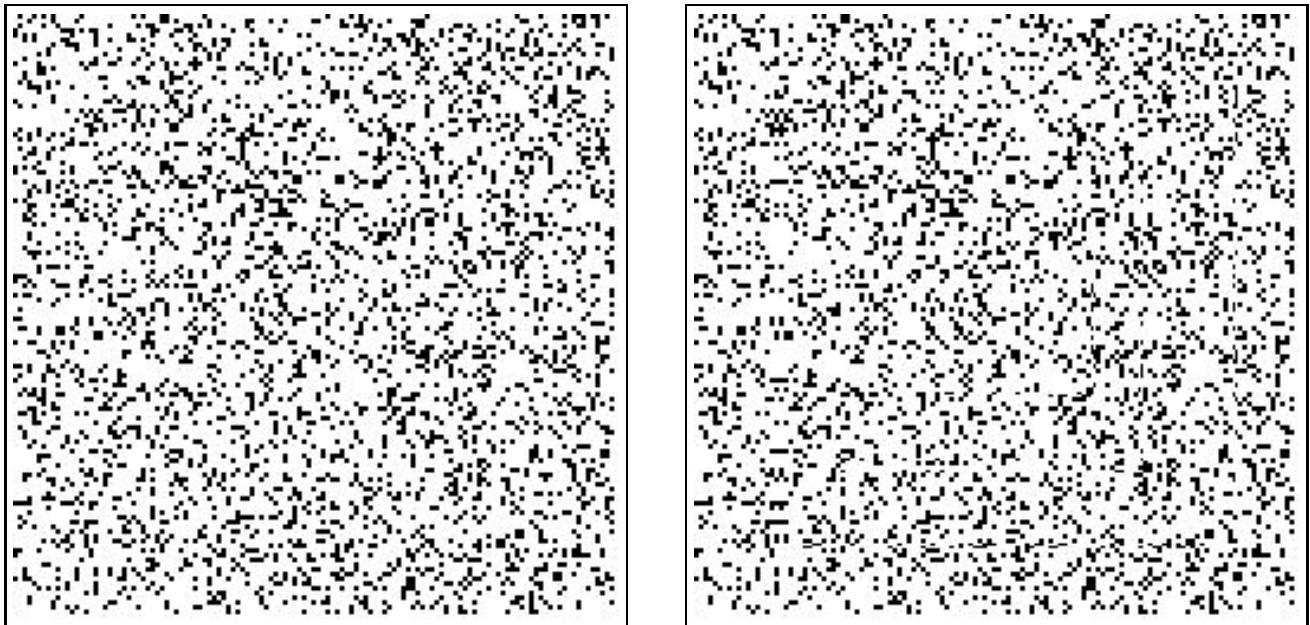
## Epipolarlinieopretning

Givet kamerakalibrering kan de to billeder oprettes således at  $y_L = y_R$

## Stereosyn hos mennesket

Mennesker er gode til stereosyn: Prøv at træde en nål med det ene øje lukket. Både psykologisk og neurofysiologisk er stereosynet veludforsket.

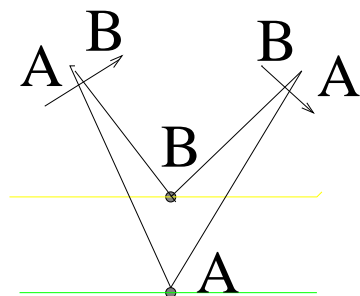
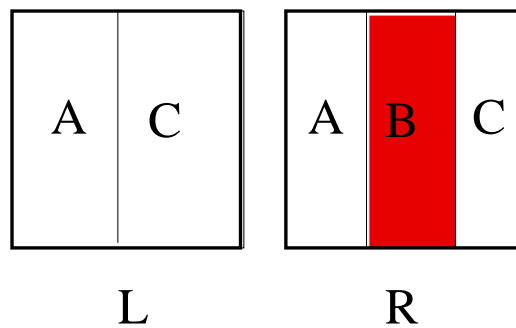
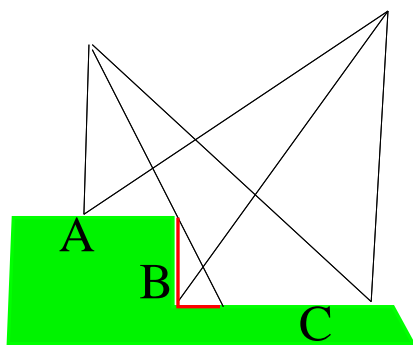
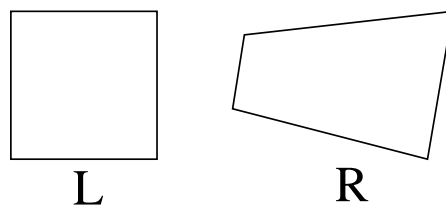
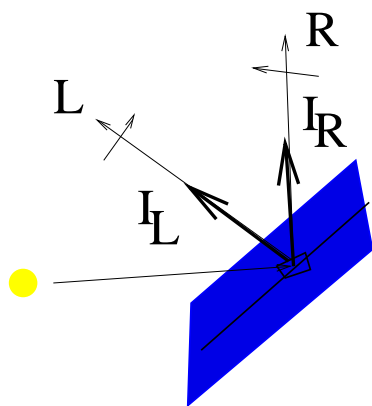
- Farver er ikke betydende for evnen til stereosyn
- Eksistensen af intensitetskontraster (f.eks. fra tekstur) er meget betydende for evnen til stereosyn. En stor del af V1 (det første område af visual cortex) er dedikeret til detektion af intensitetskanter.
- Kendte strukturer er ikke betydende for stereosynet.



Figur 1: Random dot stereogram

## Korrespondanceproblemet: Problemer

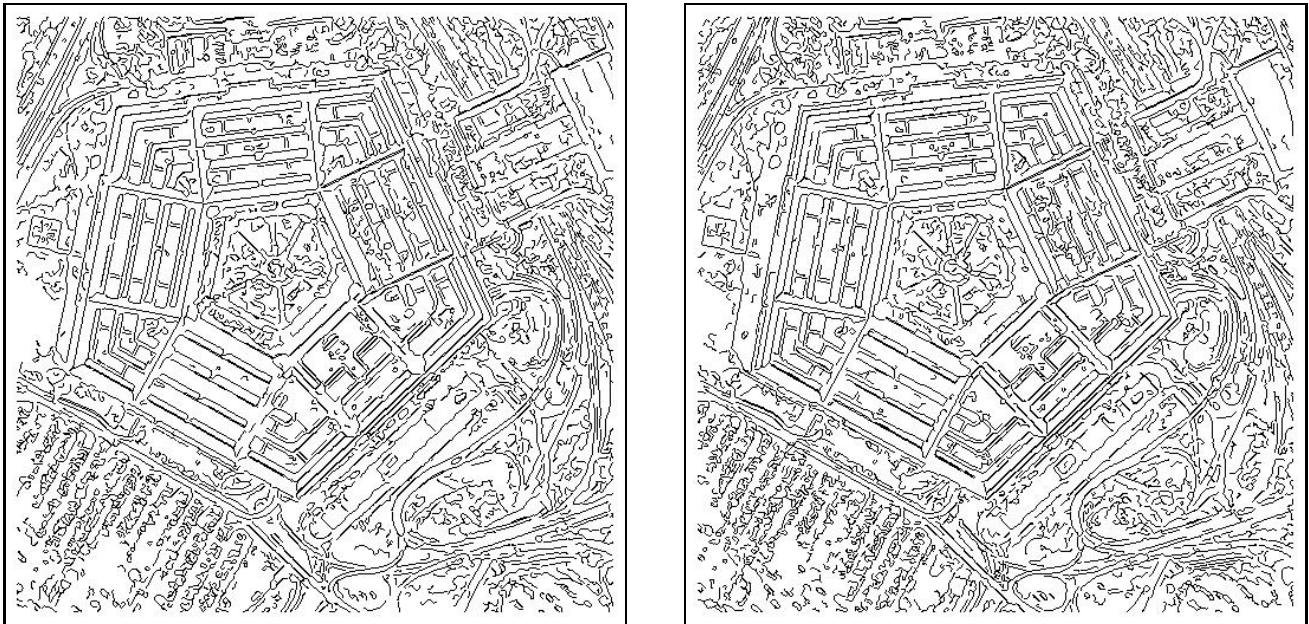
- Intensiteterne i korresponderende punkter er ikke ens
- Korresponderende områder har ikke samme form
- Visse områder kan være usynlige i det ene billede
- Ved transparens kan der opstå rækkefølgeombytning



## Simplificerende antagelser

- Antag scene består af solide objekter (ingen transperens)
- Bestem punkter [features] i billederne, sådan at disse er stabile mht. forskelle kameraposition. Bestem kun dispariteter i featurepunkter.

Kantpunkter, dvs. punkter hvor intensitetsgradientens længde (kontrasten) er lokalt maksimal, er gode (stabile) featurepunkter.

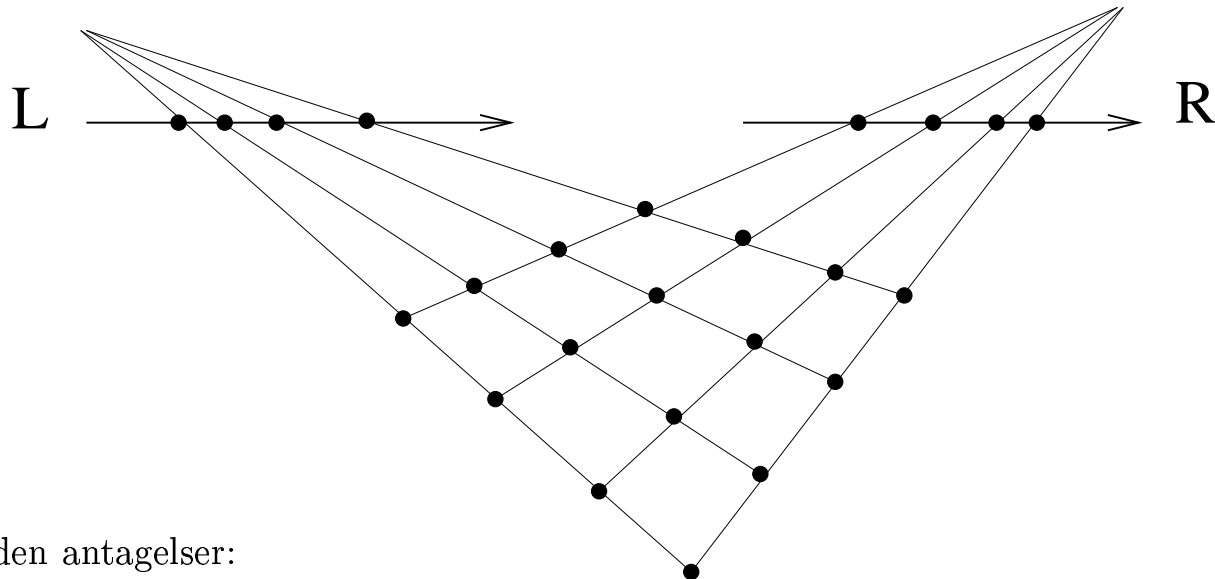


Figur 2: Kanter detekteret i et stereobilledpar

## Korrespondanceproblemet: Kombinatorisk størrelsesorden

Antag  $n$  punkter i venstre billede og  $n$  punkter i højre billede.

Lad  $N(n)$  være antal løsningsmuligheder.



Uden antagelser:

$$N(n) = 2^{n^2}. N(4) = 16536.$$

Antag svag entydighed, dvs. hvert  $x_L$  kan korrespondere med højst et punkt  $x_R$ .

$$N(n) = \sum_{i=0}^n \frac{(n!)^2}{(n-i)!(i!)^2}. N(4) = 204.$$

Antag rækkefølgeordning, dvs:  $x_L^1 < x_L^2 \Rightarrow x_R^1 < x_R^2$ , og svag entydighed.

$$N(n) = \frac{(2n)!}{(n!)^2}. N(4) = 70.$$

Antag stærk entydighed, dvs. hvert  $x_L$  korresponderer med netop et  $x_R$ .

$$N(n) = n!. N(4) = 12.$$

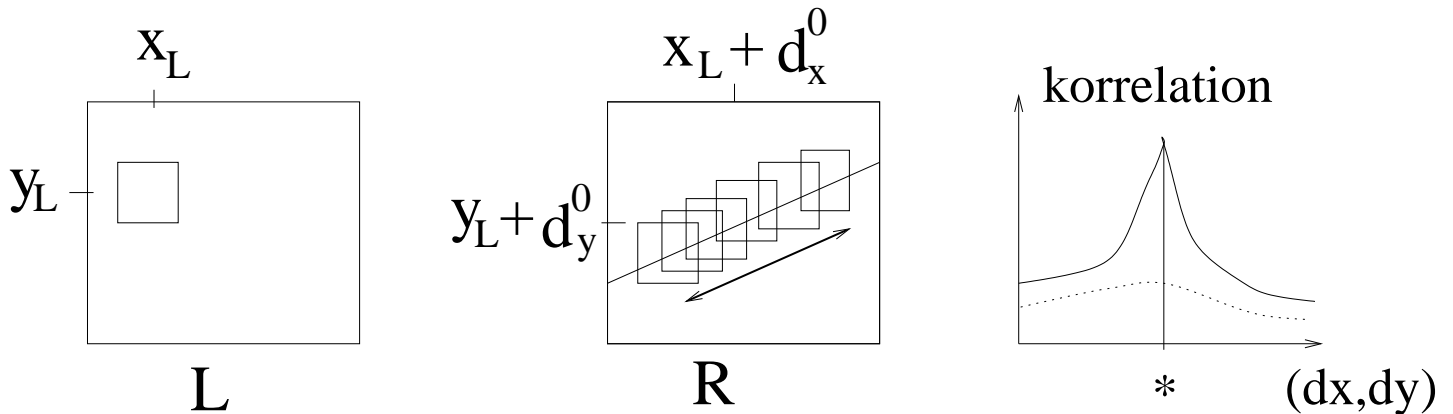
Antag rækkefølgeordning og stærk entydighed.

$$N(n) = 1.$$

## En areal baseret metode

Givet disparitetsestimater  $(d_x^0, d_y^0)$  i punktet  $(x, y)$ .

Find  $(d_x, d_y)$  så korrelationskoefficienten mellem intensiteterne i vinduer centreret i  $(x, y)$  i **L**-billedet og  $(x + d_x, y + d_y)$  i **R**-billedet er maksimal.



## Egenskaber

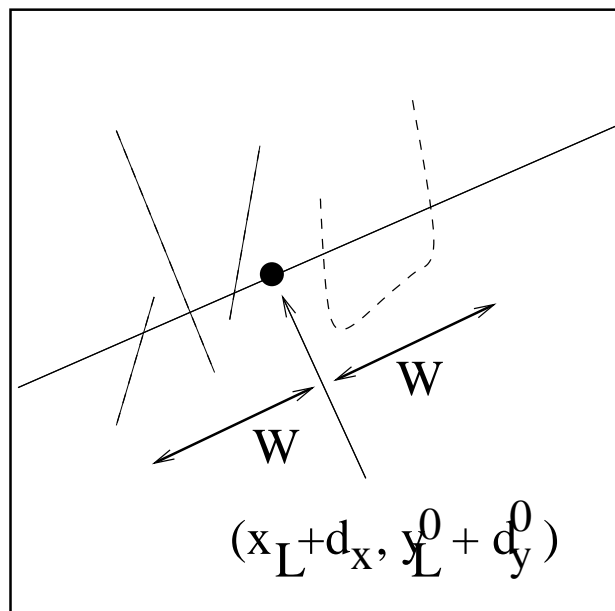
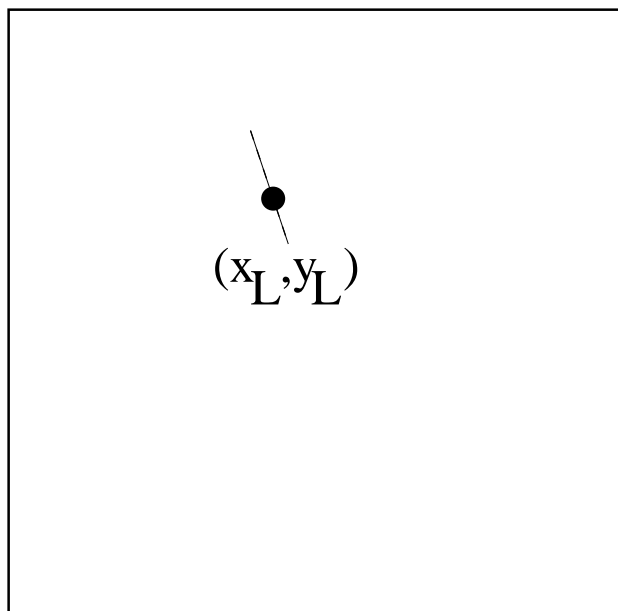
- Beregningstung
- Potentielt nøjagtig, men ikke robust
- Kan modificeres så der tages højde for forvrængning af vinduesformen.

## En feature baseret metode

Knyt featurevektor  $\vec{F} = (f_1, f_2, \dots, f_n)$  til alle kantpunkter  $(x, y)$ .

Simplest: 1 feature = kantorientering

Givet estimat  $(d_x^0, d_y^0)$  af disparitet i **L**-kantpunktet  $(x_L, y_L)$ : Find mulige kandidater til det korresponderende **R**-punkt  $(x_R, y_R)$  som de kantpunkter på **R**-epipolarlinien til  $(x_L, y_L)$  gennem  $(x_L + d_x^0, y_L + d_y^0)$  inden for søgeområde **W**.



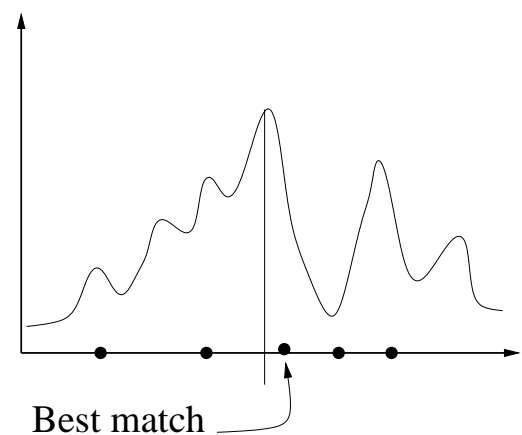
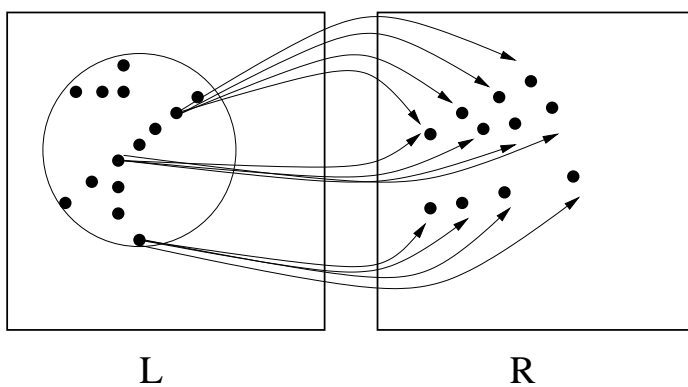
Vælg  $(x_R, y_R)$  så kandidat med orientering mest lig orienteringen i  $(x_L, y_L)$ .  
Accepter kun hvis forskel i orientering er lille (f.eks. 15-20 grader).

## Variationer

- Accepter kun korrespondencer hvis  $L \rightarrow R$ -match stemmer med  $R \rightarrow L$ -match, dvs. hvis  $L(R(\vec{x}_L)) = \vec{x}_L$ .
- Generelt godhedsmål for match  $(x_L, y_L) \rightarrow (x_R, y_R)$

$$J = fkt(\|(d_x - d_x^0, d_y - d_y^0)\|^2) + \sum_{i=1}^n \alpha_i \|f_i^L - f_i^R\|^2$$

- **Alternativ:** Vælg først samtlige kandidater, der opfylder et mindstemål af lighed mellem features, og udskyd endeligt valg til senere passage, hvor mere global information er til rådighed.
- **Alternativ:** Bestem samtlige mulige dispariteter i alle punkter i en omegn om  $(x_L, y_L)$ . Beregn en udglattet fordeling af disparitetsfordelingen. Vælg match, der stemmer bedst med maksimumspunkt i fordeling.

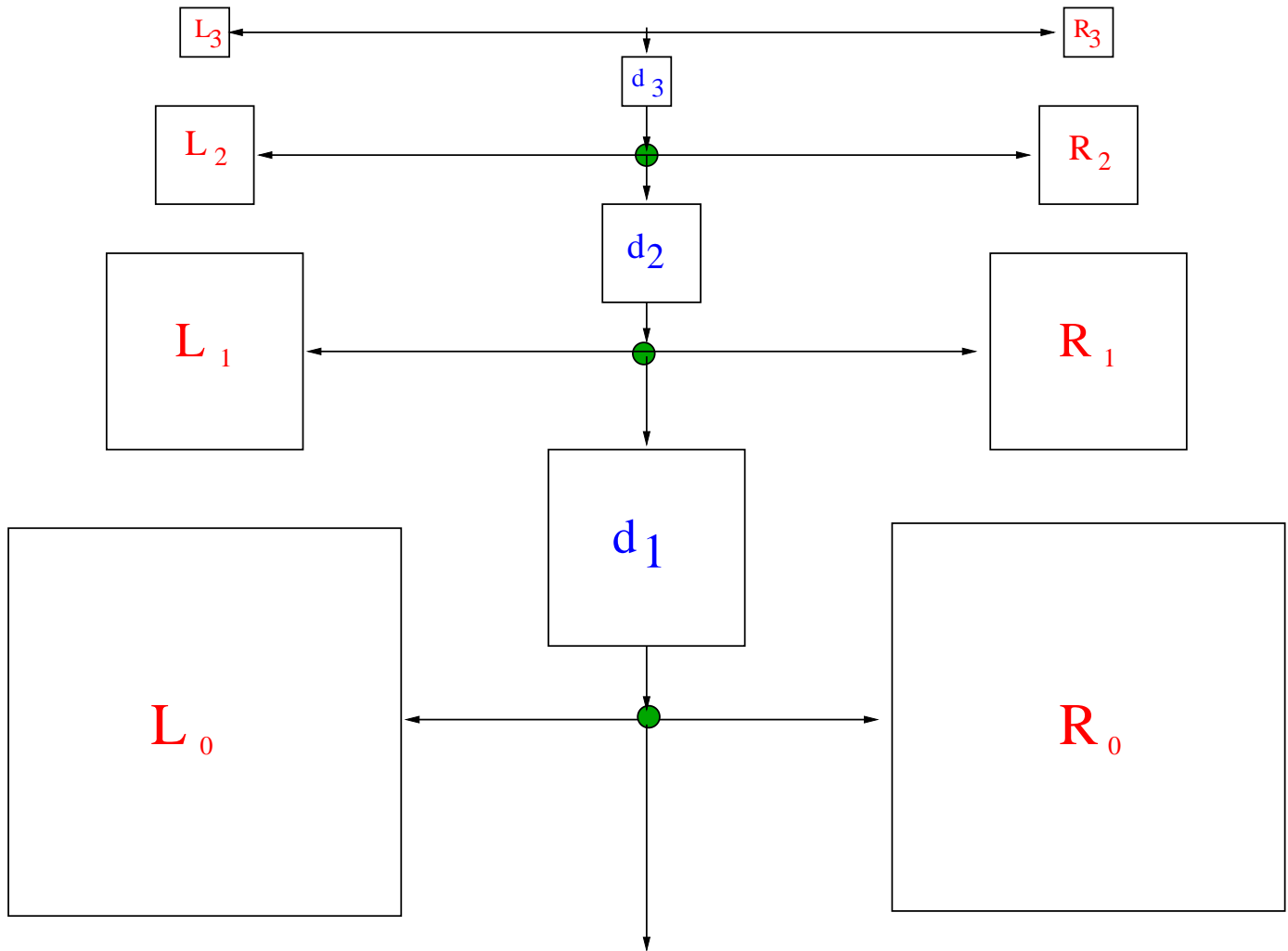




## Grov til fin analyse

I praksis kan dispariteten være stor (fx. 100 pixels) og have stor variation (fx. fra -50 pixel til 50 pixel). For at reducere søgeområdets størrelse er et estimat af dispariteten nødvendigt.

### Metode: Trinvis nedfotografering



Ved en nedfotograferingsgrad på  $1/2$  bliver lagerforbruget:

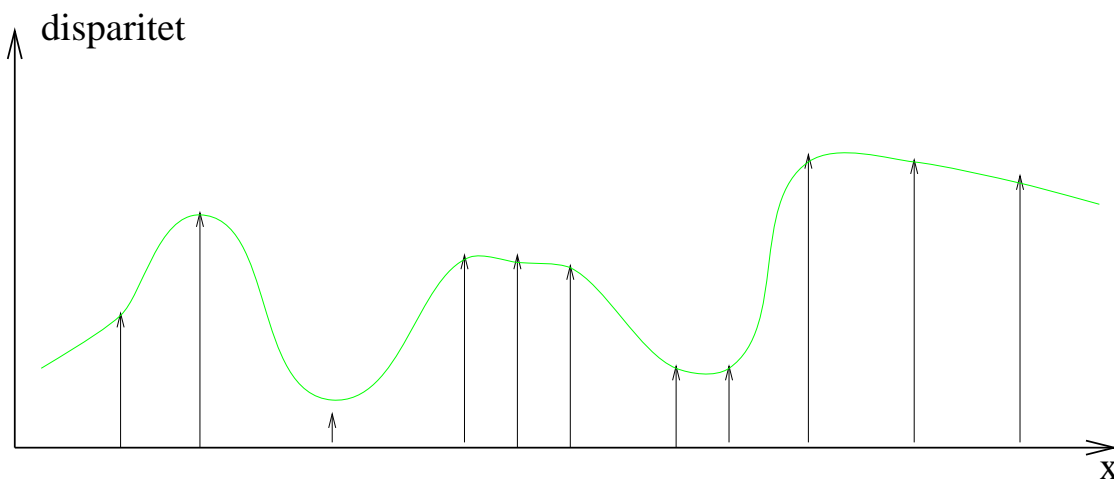
$$1 + \frac{1}{4} + \frac{1}{16} + \dots < \frac{4}{3}$$

## Tæt disparitetsflade

Ved featurebaseret stereoanalyse bestemmes disparitet kun i featurepunkter. Bestemmelse af disparitet overalt kan foretages ved:

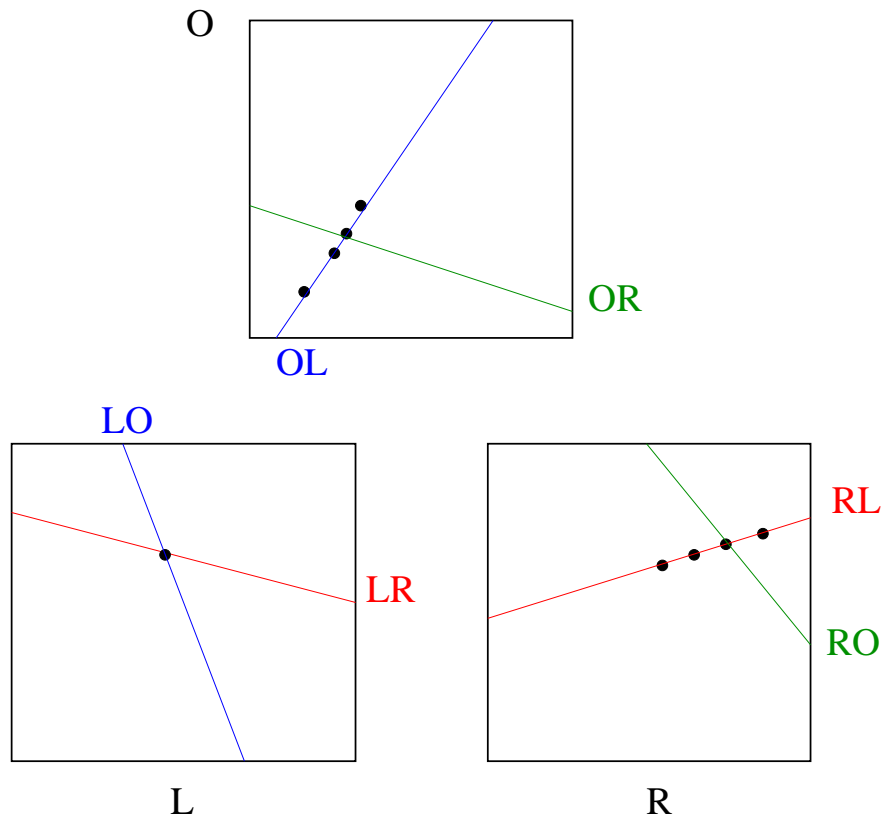
- Anvendelse af intensitetsbaseret (korrelations-) algoritme hvor søgeområde antages lille og bestemt ud fra dispariteten i de nærmeste featurepunkter.
- Ved interpolation, ved at tilpasse en glat flade  $d(x, y)$  til de fundne dispariteter  $c(x, y)$ . Funktionen  $d$  kan f.eks. bestemmes som den funktion, der minimerer  $\Theta(d)$ :

$$\Theta(d(x, y)) = \sum_M |d(x, y) - c(x, y)|^2 + \iint |d_x^2(x, y) + d_y^2(x, y)| dx dy$$



## Fler-kamera stereo

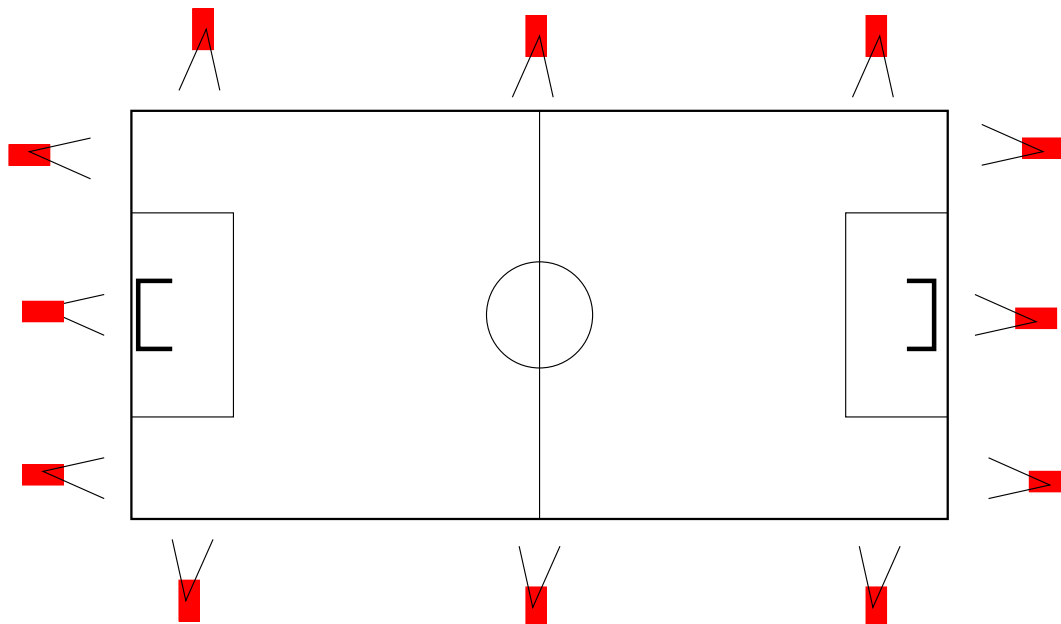
Ved brug af 3 kamera-stereo lægger epipolarliniegeometrien stærke restriktioner på korrespondanceproblemet.



Ved brug af mange kamera'er øges robustheden.

## Var der straffe ?

Placere 10-20 helt stabile (kalibrerede) kamera'er omkring fodboldbane, således at alle positioner dækket af 3-6 kamera'er.



Løbende sporing og 3D-rekonstruktion af de 24 bevægelige objekter.

Var der straffe ?

Var bolden inde ?