



ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΠΑΤΡΩΝ  
UNIVERSITY OF PATRAS

ΑΝΟΙΚΤΑ ακαδημαϊκά  
μαθήματα ΠΠ

## Ειδικά Θέματα Υπολογιστικής Όρασης & Γραφικής

Εμμανουήλ Ζ. Ψαράκης & Αθανάσιος Τσακαλίδης  
Πολυτεχνική Σχολή  
Τμήμα Μηχανικών Η/Υ & Πληροφορικής



## Υπολογιστική Όραση

### © Στερεοσκοπική Αντιστοίχιση

Εμμανουήλ Ζ. Ψαράκης

Πολυτεχνική Σχολή

Τμήμα Μηχανικών Η/Υ & Πληροφορικής

# Υπολογιστική Όραση

## *Επισκόπηση Μαθήματος*

### *Στερεοσκοπική Αντιστοίχιση*

- Ανασταλτικοί Παράγοντες
- Βασικές Υποθέσεις
- Μοντέλο κάμερας
- Προβολική Γεωμετρία
- Τεχνικές Στερεοσκοπικής Αντιστοίχισης

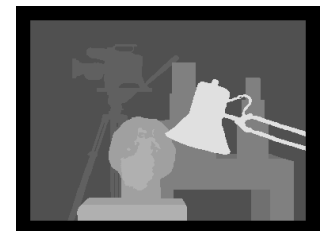
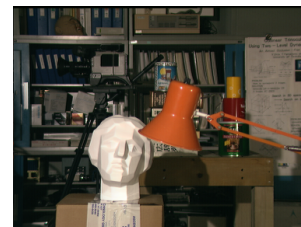
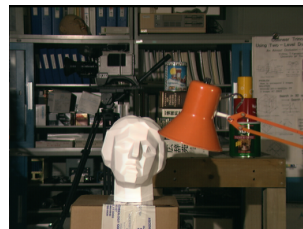


# Υπολογιστική Όραση

## Στερεοσκοπική Αντιστοίχιση

Γενικό πρόβλημα: Αντιστοίχισε το προφίλ αναφοράς  $I(x,y)$  με το προφίλ εισόδου  $J(x,y)$

Πρόβλημα	$I(x,y)$	$J(x,y)$	Σκοπός
Στερεοσκοπική Αντιστοίχιση (Stereo Correspondence)	Αριστερή εικόνα $L(x,y)$	Δεξιά εικόνα $R(x,y)$	Χάρτης Ανομοιότητας



# Υπολογιστική Όραση

## Βασικό ερώτημα (Γενική Περίπτωση)

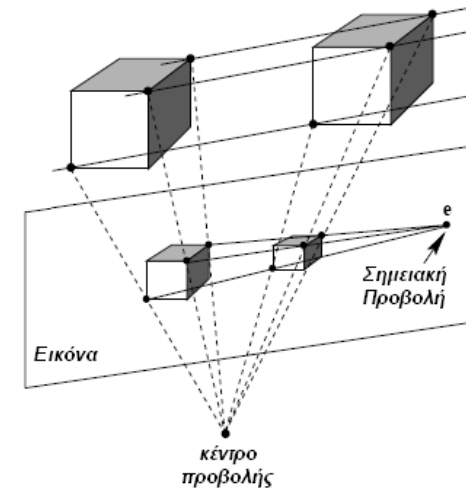
- Δοθέντων δύο εικόνων,
  - ποια είναι τα *αντίστοιχα σημεία* τους;
    - Αντίστοιχα σημεία: προβολές του ίδιου σημείου της σκηνής στις εικόνες
  - ποιος είναι ο μετασχηματισμός, που εφαρμοζόμενος στη μία εικόνα, παρέχει την άλλη;
- Η γεωμετρία του χώρου και ο προσανατολισμός του(ων) αισθητήρα(ων) όρασης **δεν** είναι γνωστά
- **Μόνη** πηγή πληροφορίας: η ένταση φωτεινότητας των εικόνων



# Υπολογιστική Όραση

## Ανασταλτικοί παράγοντες

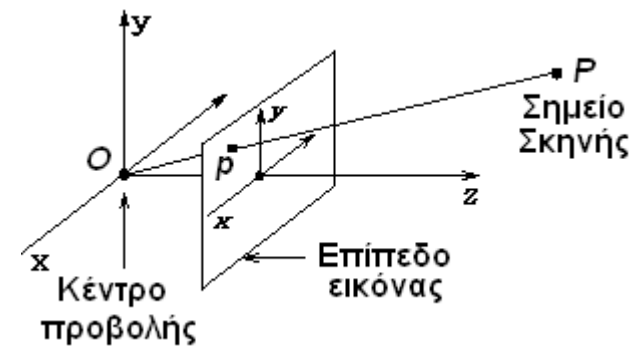
- Ψηφιακή Εικόνα
  - Θόρυβος καταγραφής
  - 1 εικονοστοιχείο αντιστοιχεί σε πολλά σημεία της σκηνής
- Προοπτική Προβολή
  - Προβολή 3D σε 2D (απώλεια πληροφορίας)
- Κίνηση Κάμερας/Σκηνής
  - Παραμόρφωση αντικειμένων
- Μη ομοιόμορφος (φυσικός) φωτισμός
  - -Μη λαμπερτιανές επιφάνειες
- Ασυνέχειες Βάθους
  - Ημι-αποκλεισμένες περιοχές
- Παραμόρφωση φακού
  - Ευθείες μετατρέπονται σε καμπύλες



# Υπολογιστική Όραση

## Ανασταλτικοί παράγοντες

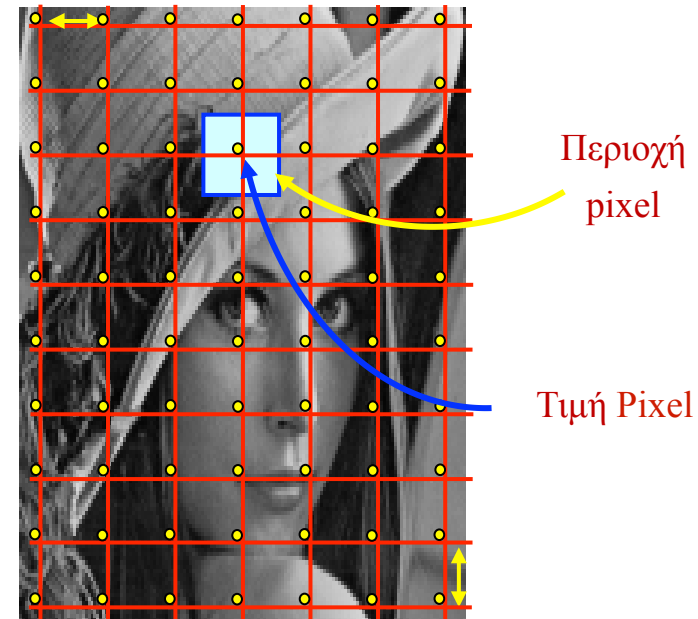
- Ψηφιακή Εικόνα
  - Θόρυβος καταγραφής
  - 1 εικονοστοιχείο αντιστοιχεί σε πολλά σημεία της σκηνής
- Προοπτική Προβολή
  - Προβολή 3D σε 2D (απώλεια πληροφορίας)
- Κίνηση Κάμερας/Σκηνής
  - Παραμόρφωση αντικειμένων
- Μη ομοιόμορφος (φυσικός) φωτισμός
  - -Μη λαμπερτιανές επιφάνειες
- Ασυνέχειες Βάθους
  - Ημι-αποκλεισμένες περιοχές
- Παραμόρφωση φακού
  - Ευθείες μετατρέπονται σε καμπύλες



# Υπολογιστική Όραση

## Ανασταλτικοί παράγοντες

- Ψηφιακή Εικόνα
  - Θόρυβος καταγραφής
  - 1 εικονοστοιχείο αντιστοιχεί σε πολλά σημεία της σκηνής
- Προοπτική Προβολή
  - Προβολή 3D σε 2D (απώλεια πληροφορίας)
- Κίνηση Κάμερας/Σκηνής
  - Παραμόρφωση αντικειμένων
- Μη ομοιόμορφος (φυσικός) φωτισμός
  - -Μη λαμπερτιανές επιφάνειες
- Ασυνέχειες Βάθους
  - Ημι-αποκλεισμένες περιοχές
- Παραμόρφωση φακού
  - Ευθείες μετατρέπονται σε καμπύλες

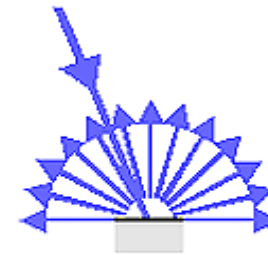




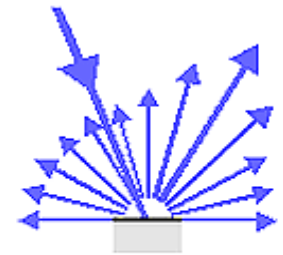
# Υπολογιστική Όραση

## Ανασταλτικοί παράγοντες

- Ψηφιακή Εικόνα
  - Θόρυβος καταγραφής
  - 1 εικονοστοιχείο αντιστοιχεί σε πολλά σημεία της σκηνής
- Προοπτική Προβολή
  - Προβολή 3D σε 2D (απώλεια πληροφορίας)
- Κίνηση Κάμερας/Σκηνής
  - Παραμόρφωση αντικειμένων
- Μη ομοιόμορφος (φυσικός) φωτισμός
  - -Μη λαμπερτιανές επιφάνειες
- Ασυνέχειες Βάθους
  - Ημι-αποκλεισμένες περιοχές
- Παραμόρφωση φακού
  - Ευθείες μετατρέπονται σε καμπύλες



*Λαμπερτιανή  
επιφάνεια*



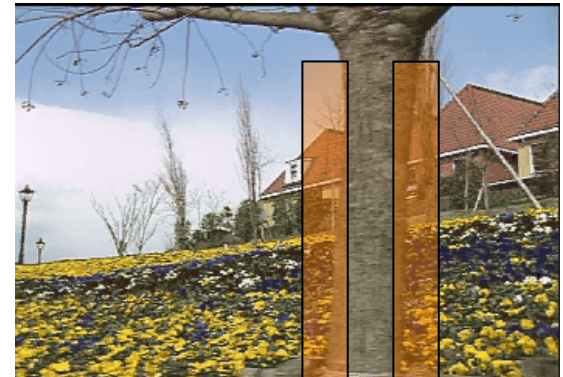
*Μη λαμπερτιανή  
επιφάνεια*



# Υπολογιστική Όραση

## *Ανασταλτικοί παράγοντες*

- Ψηφιακή Εικόνα
  - Θόρυβος καταγραφής
  - 1 εικονοστοιχείο αντιστοιχεί σε πολλά σημεία της σκηνής
- Προοπτική Προβολή
  - Προβολή 3D σε 2D (απώλεια πληροφορίας)
- Κίνηση Κάμερας/Σκηνής
  - Παραμόρφωση αντικειμένων
- Μη ομοιόμορφος (φυσικός) φωτισμός
  - -Μη λαμπερτιανές επιφάνειες
- Ασυνέχειες Βάθους
  - Ημι-αποκλεισμένες περιοχές
- Παραμόρφωση φακού
  - Ευθείες μετατρέπονται σε καμπύλες



# Υπολογιστική Όραση

## *Ανασταλτικοί παράγοντες*

- Ψηφιακή Εικόνα
  - Θόρυβος καταγραφής
  - 1 εικονοστοιχείο αντιστοιχεί σε πολλά σημεία της σκηνής
- Προοπτική Προβολή
  - Προβολή 3D σε 2D (απώλεια πληροφορίας)
- Κίνηση Κάμερας/Σκηνής
  - Παραμόρφωση αντικειμένων
- Μη ομοιόμορφος (φυσικός) φωτισμός
  - -Μη λαμπερτιανές επιφάνειες
- Ασυνέχειες Βάθους
  - Ημι-αποκλεισμένες περιοχές
- Παραμόρφωση φακού
  - Ευθείες μετατρέπονται σε καμπύλες



# Υπολογιστική Όραση

## Βασική Υπόθεση

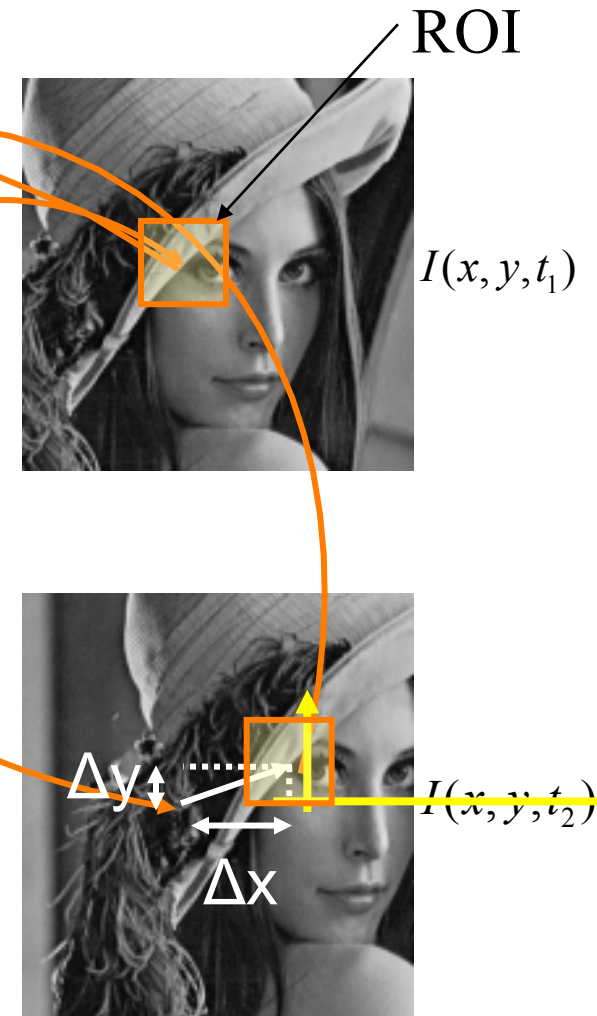
- Σταθερή ένταση φωτεινότητας (Brightness Constancy Assumption) [Horn and Schunk '81]
  - Ένα σημείο της σκηνής απεικονίζεται με την ίδια ένταση φωτεινότητας σε όλες τις διαθέσιμες εικόνες

$$I(x_0, y_0, t_1) = I(x_0 + \Delta x, y_0 + \Delta y, t_2)$$

- Αδυναμία ισχύος σε πρακτικές εφαρμογές
- Καλή προσέγγιση αν
  - $t_2 - t_1 \rightarrow 0$
  - $\Delta x \rightarrow 0$
  - $\Delta y \rightarrow 0$
  - Video με μεγάλο fps

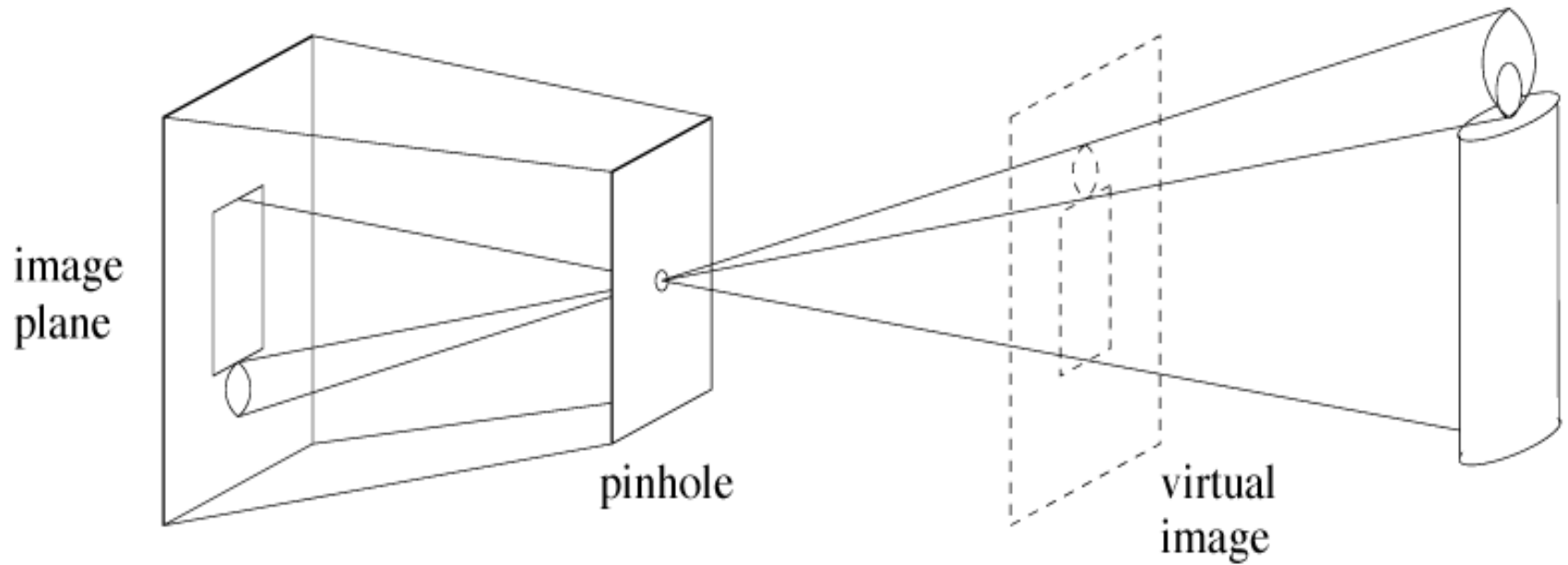
$$I(x, y, t_1) \approx I(x + \Delta x, y + \Delta y, t_2)$$

$$\forall (x, y) \in ROI$$



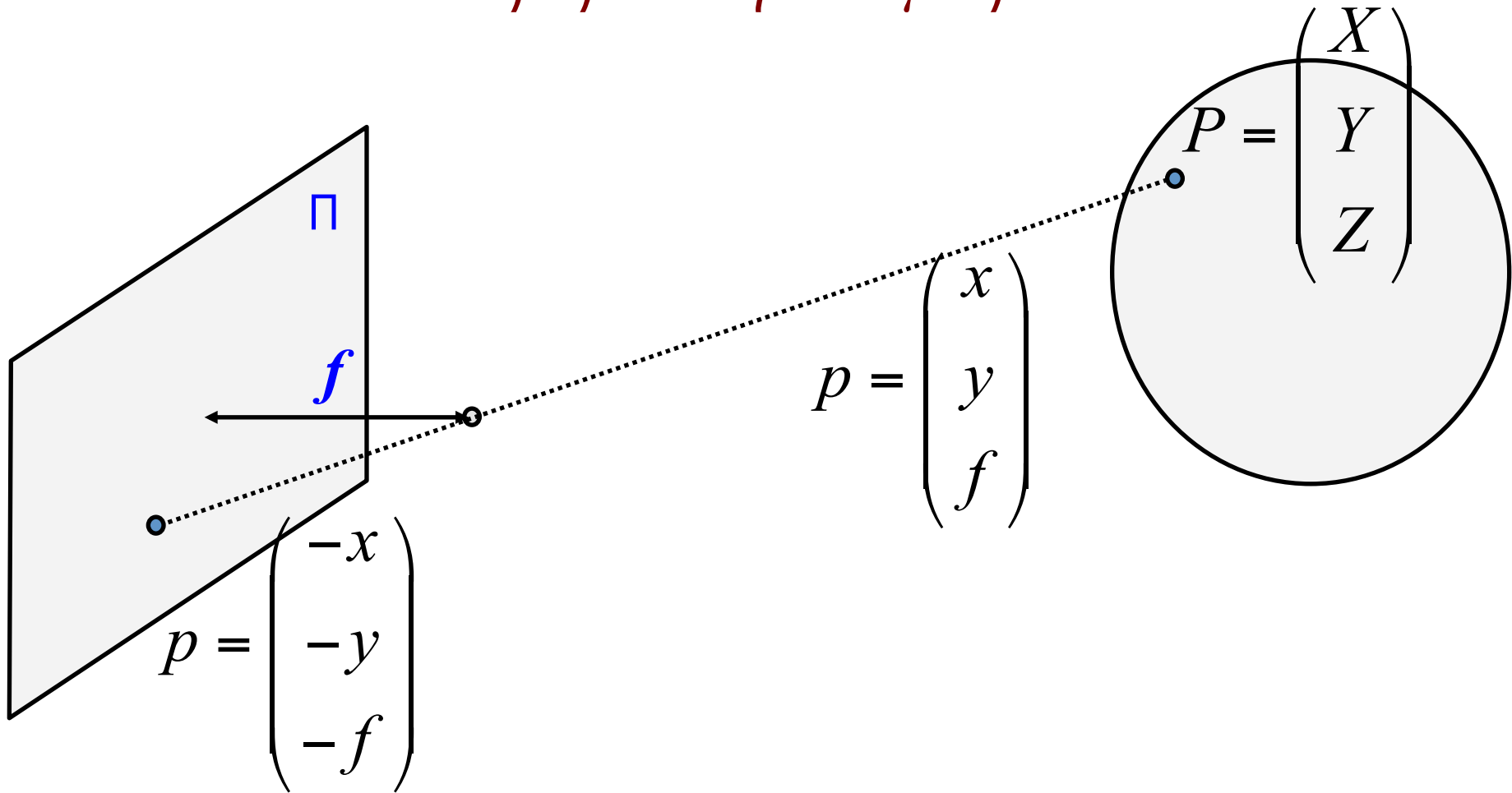
# Υπολογιστική Όραση

## *Μοντέλο Κάμερας*



# Υπολογιστική Όραση

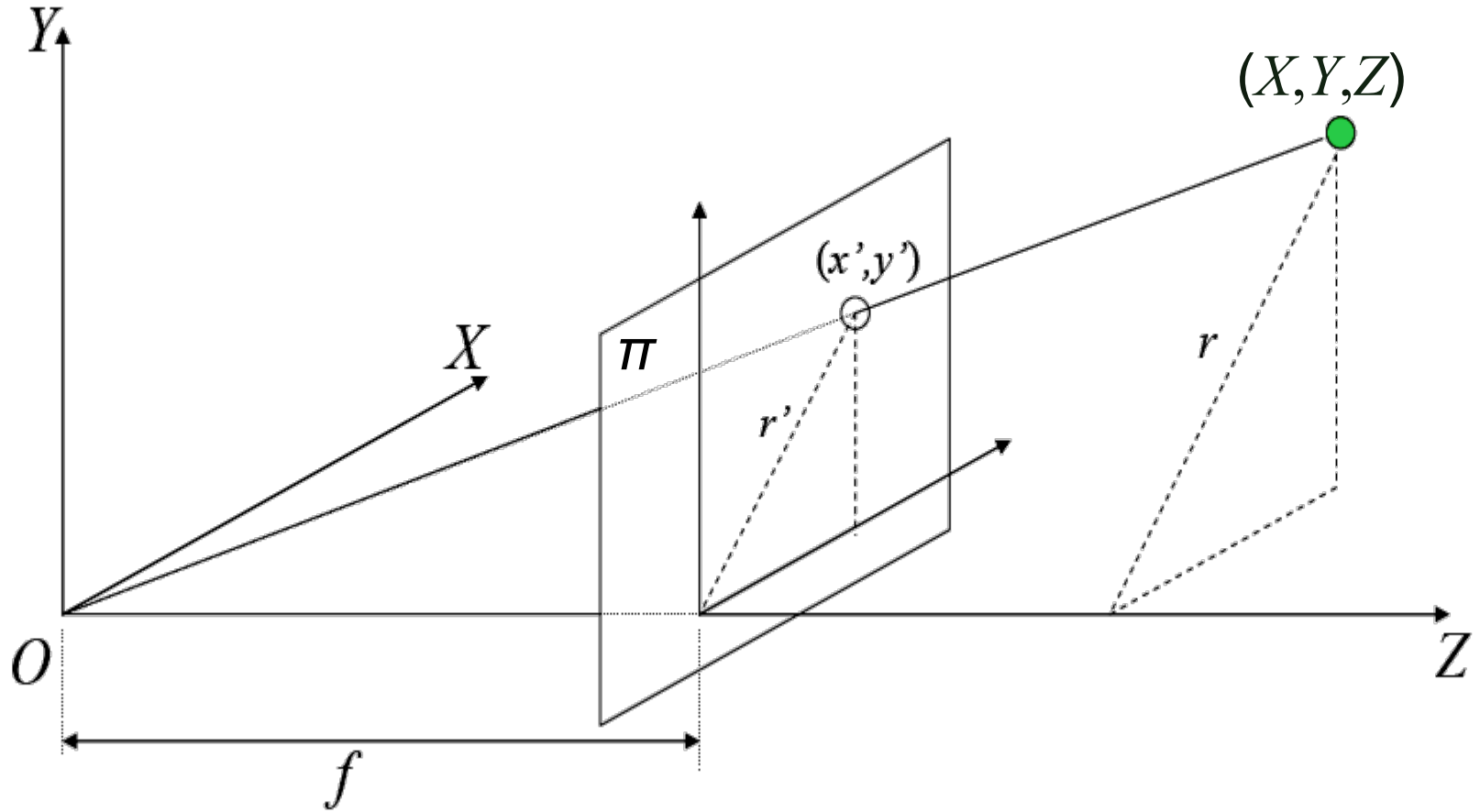
## Προβολική Γεωμετρία



# Υπολογιστική Όραση

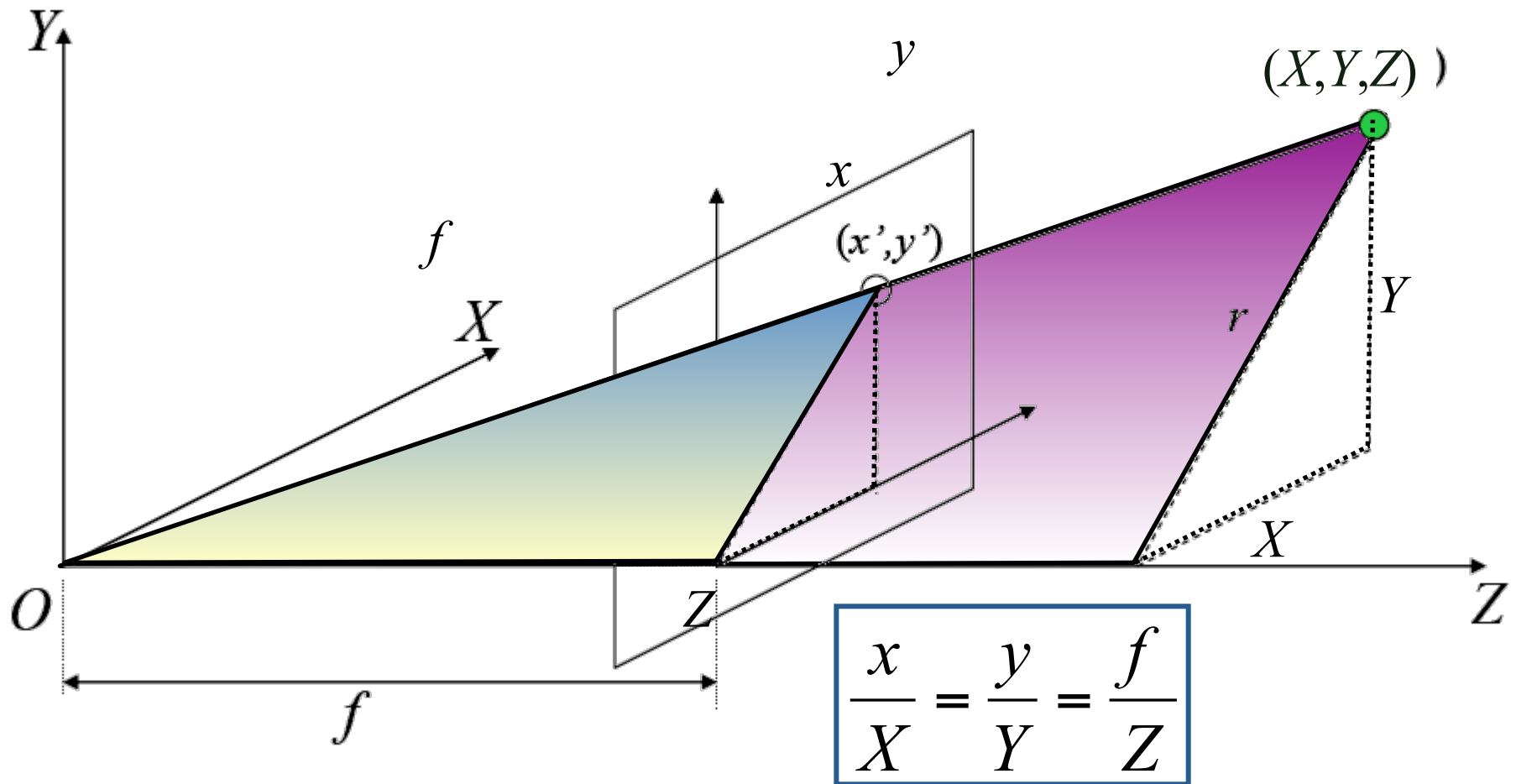
## Συμβολισμοί

- $O$  – Εστιακό Κέντρο
- $\pi$  – Επίπεδο Εικόνας
- $Z$  – Οπτικός Άξονας
- $f$  – Εστιακή Απόσταση



# Υπολογιστική Όραση

## Προβολή

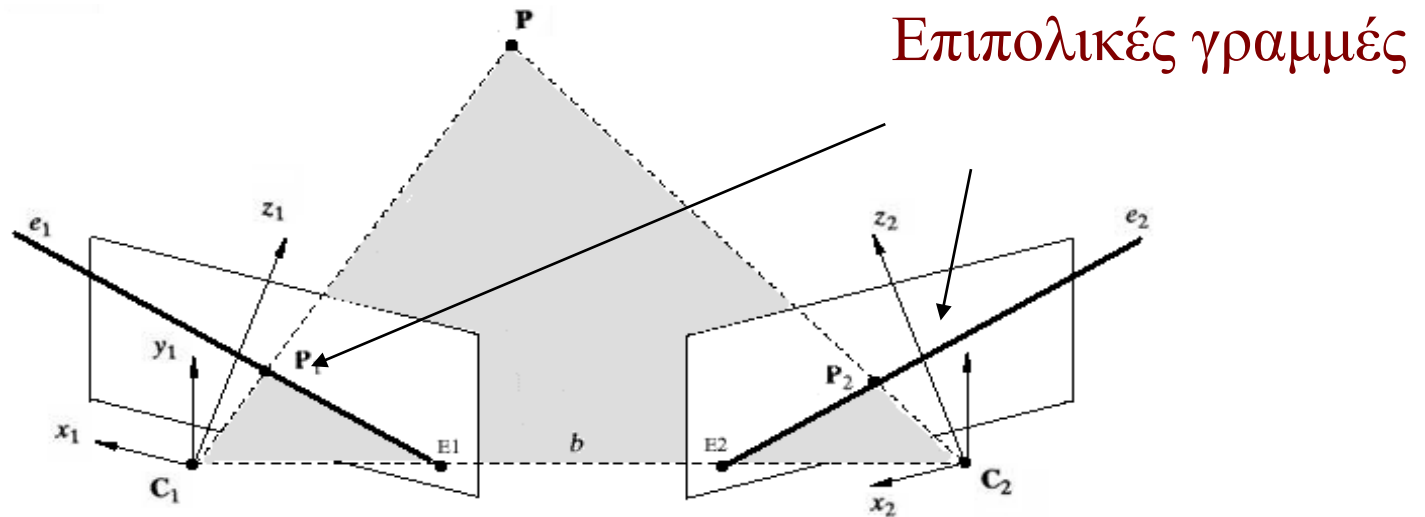




# Υπολογιστική Όραση

## Συστήματα Δύο Αισθητήρων

- Τυχαίος Προσανατολισμός Αισθητήρων

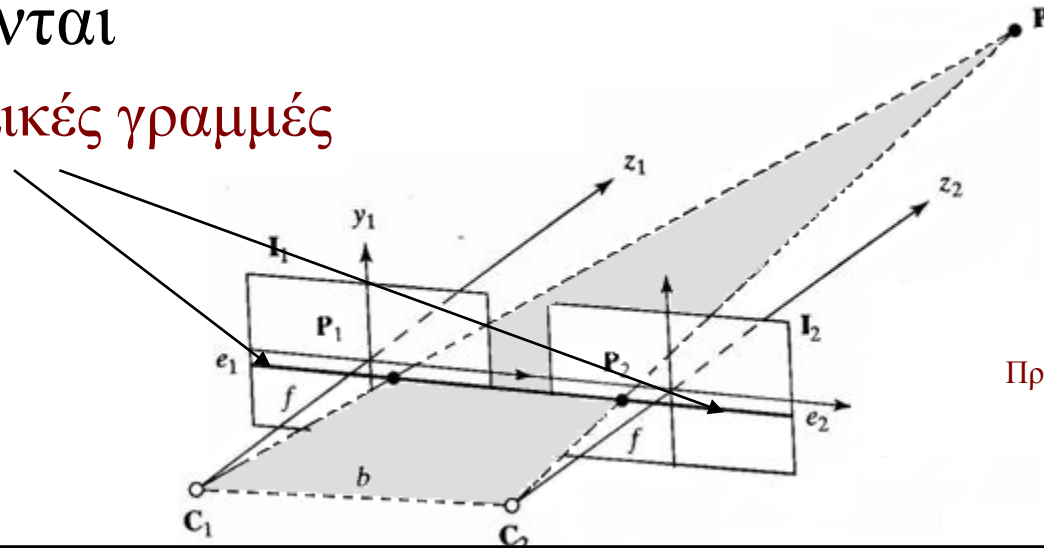


# Υπολογιστική Όραση

## Κανονικό Στερεοσκοπικό Σύστημα

- Παράλληλοι οπτικοί άξονες
- Οι οριζόντιοι άξονες ( $x$ ) των δύο συστημάτων ταυτίζονται

Επιπολικές γραμμές



Κανονικός  
Προσανατολισμός

*Οι επιπολικές γραμμές ταυτίζονται με τις γραμμές των εικόνων*



# Υπολογιστική Όραση

## Κανονική Διάταξη Αισθητήρων

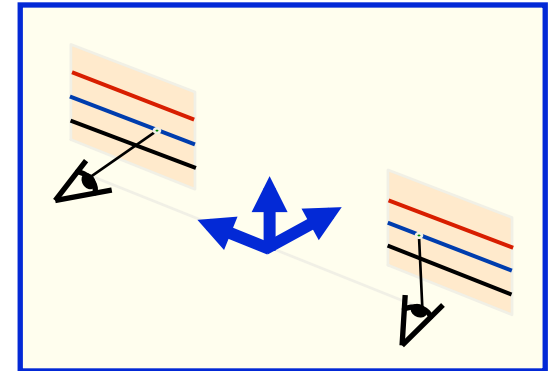
- Η Αρχή του Σ.Σ. στο μέσο του Ε.Τ. που ενώνει τα οπτικά κέντρα

$$x_l = \frac{f(X + b/2)}{Z}, \quad y_l = \frac{fY}{Z}$$

$$x_r = \frac{f(X - b/2)}{Z}, \quad y_r = \frac{fY}{Z}$$

$$x_l - x_r = \frac{fb}{Z}$$

$$X = \frac{b(x_l + x_r)}{2(x_l - x_r)}, \quad Y = \frac{b(y_l + y_r)}{2(x_l - x_r)}, \quad Z = \frac{fb}{x_l - x_r}$$



# Υπολογιστική Όραση

## Ανομοιότητα (*Disparity*)

$$Z = \frac{fb}{x_l - x_r}$$

- Η διαφορά  $d = x_l - x_r$  ονομάζεται *ανομοιότητα*
- Το  $d$  είναι αντιστρόφως ανάλογο του  $Z$
- Το  $d$  είναι ανάλογο του  $b$



# Υπολογιστική Όραση

## Στερεοσκοπική Αντιστοίχιση

- Υπολογισμός *ανομοιότητας* ως προς την εικόνα αναφοράς (π.χ. αριστερή)
  - *Ανομοιότητα*: η απόσταση σε εικονοστοιχεία συζυγών ζευγών όταν τοποθετήσουμε τη μία εικόνα πάνω από την άλλη

$$I_{left}(x, y) = I_{right}(x - dx, y - dy)$$

- Αναζήτηση συζυγών ζευγών (αντιστοιχών σημείων) κατά μήκος των *επιπολικών γραμμών*
- Επιλογή κανονικού συστήματος

$$dy = 0 \Rightarrow \begin{cases} I_{left}(x, y) = I_{right}(x - dx, y) \\ dx > 0 \end{cases}$$



# Υπολογιστική Όραση

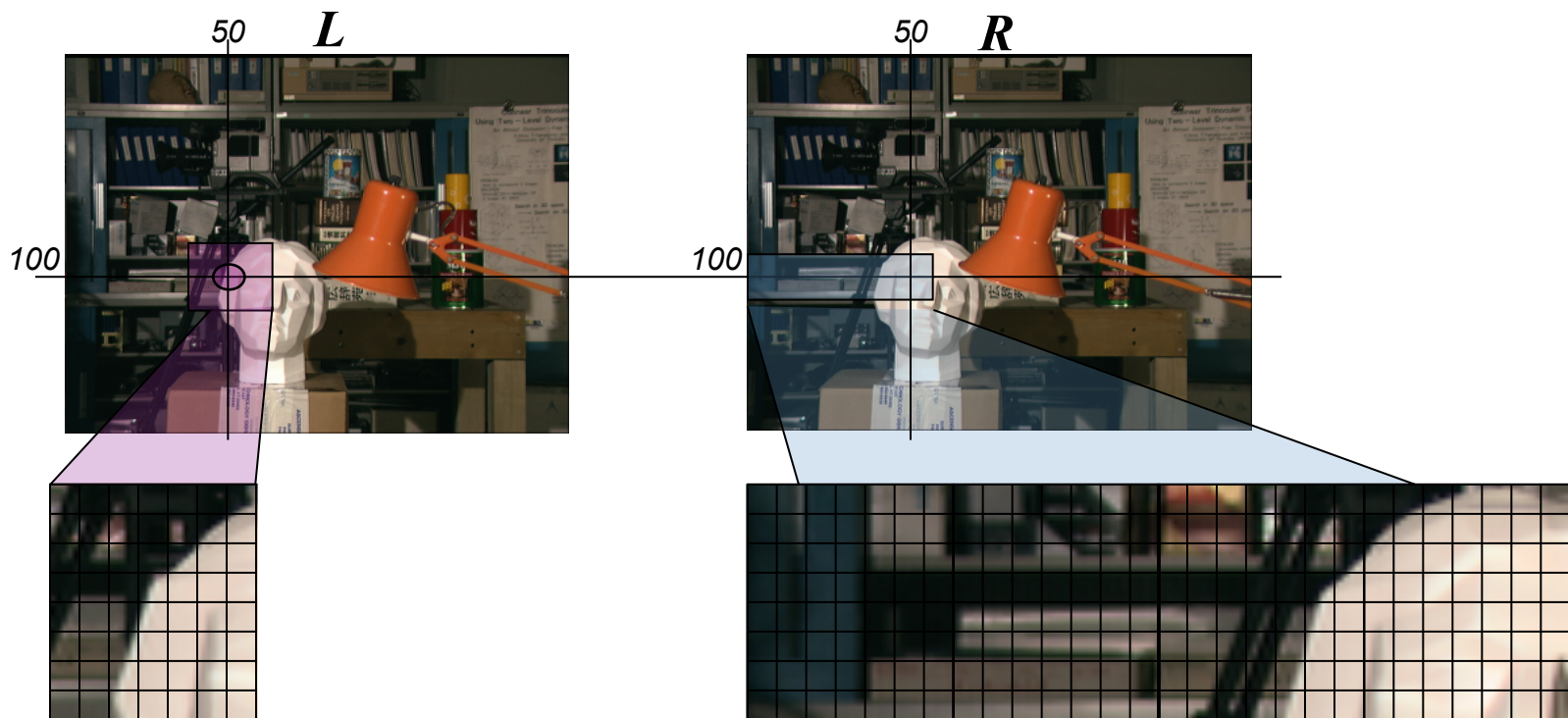
## Μέθοδοι Στερεοσκοπικής Αντιστοίχισης

- Τοπικές μέθοδοι (pixel-wise)
  - Απαραίτητη χρήση παραθύρου (window-based)
  - Επιλογή αντίστοιχου σημείου από πολλά υποψήφια (winner takes all)
- Ημι-ολικές μέθοδοι
  - Δυναμικός προγραμματισμός (row by row)
  - Αναζήτηση βέλτιστου μονοπατιού στο επίπεδο
- Ολικές μέθοδοι
  - Αναζήτηση βέλτιστης επιφάνειας στο χώρο ανομοιότητας (disparity space image)

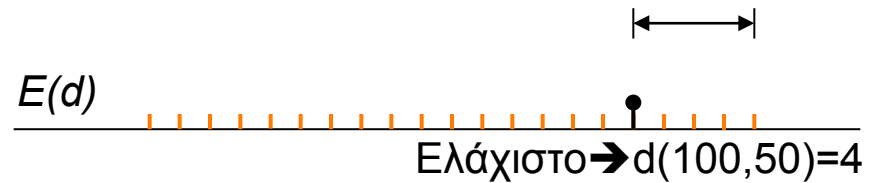


# Υπολογιστική Όραση

## Τοπικές μέθοδοι



$$d(100, 50) = 4 \rightarrow L(100, 50) = R(96, 50)$$



# Υπολογιστική Όραση

## Περιορισμοί και Υποθέσεις

- *Περιορισμοί*
  - Μοναδικότητα: κάθε σημείο της αριστερής εικόνας έχει μοναδικό αντίστοιχο στη δεξιά
- *Υποθέσεις*
  - Σειρά προβολής: η σειρά εμφάνισης δύο σημείων στην αριστερή και δεξιά εικόνα δεν αλλάζει.
  - Η ανομοιότητα σε γειτονικά σημεία δεν μπορεί να ποικίλει έντονα

Η υιοθέτηση  
περιορισμών  
και υποθέσεων



μειώνει το χώρο αναζήτησης  
αντίστοιχων σημείων



μπορεί να προκαλέσει  
διάδοση σφαλμάτων

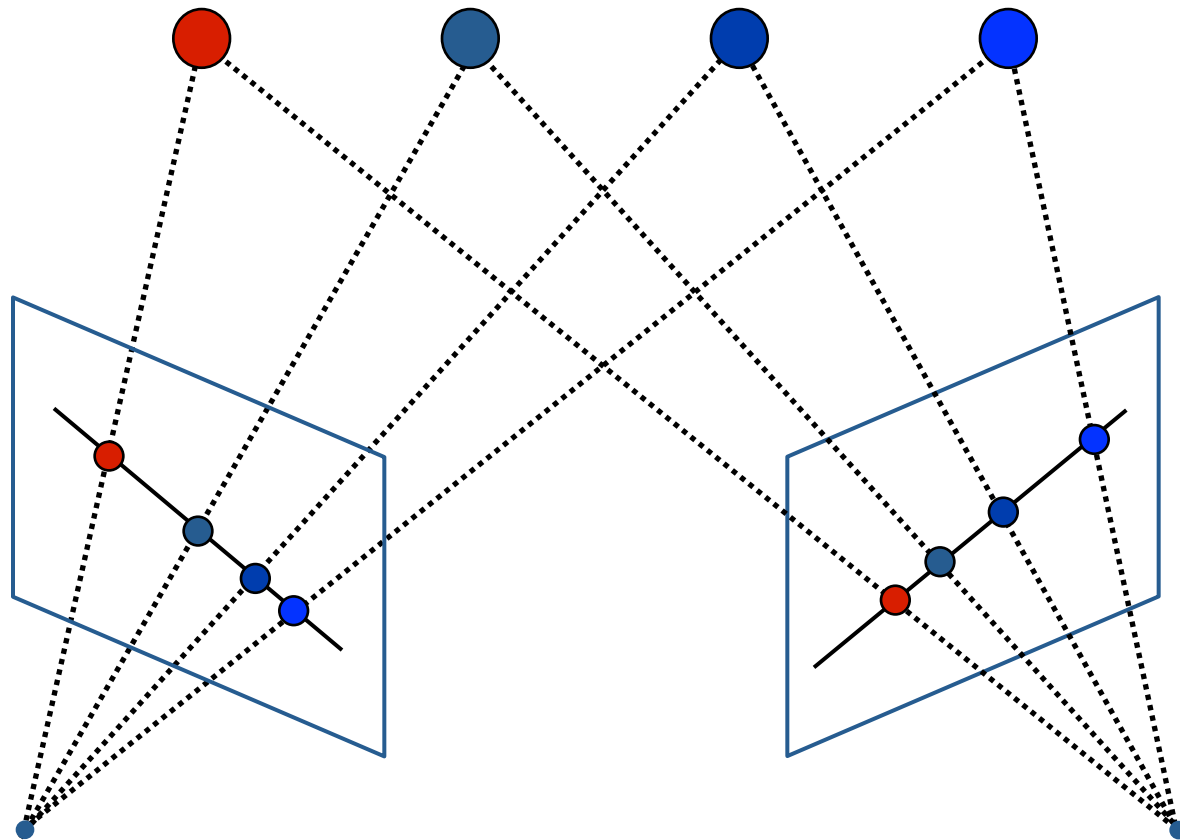




# Υπολογιστική Όραση

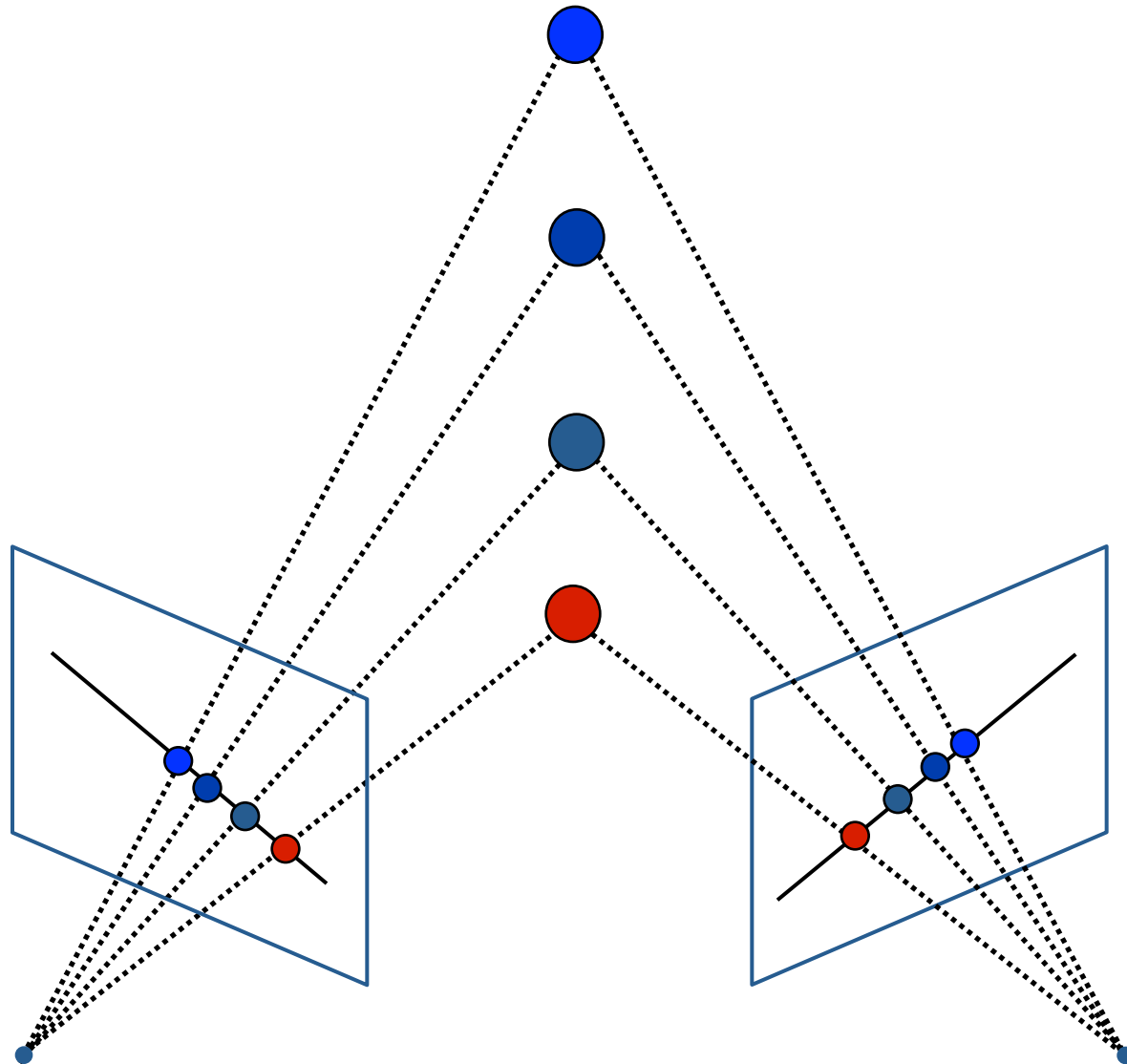
## Περιορισμοί και Υποθέσεις

Υπόθεση: *Σειράς προβολής*



# Υπολογιστική Όραση

## Υπόθεση Σειράς Προβολής



# Υπολογιστική Όραση

## Προβλήματα-Ανασταλτικοί Παράγοντες

Περιοχές μη έντονης υφής

Αριστερή Εικόνα

Δεξιά Εικόνα

Ασυνέχειες Βάθους

Αριστερή Εικόνα

Δεξιά Εικόνα

Φωτομετρικές Παραμορφώσεις

Αριστερή Εικόνα

Δεξιά Εικόνα

Περιοδικότητες

Αριστερή Εικόνα

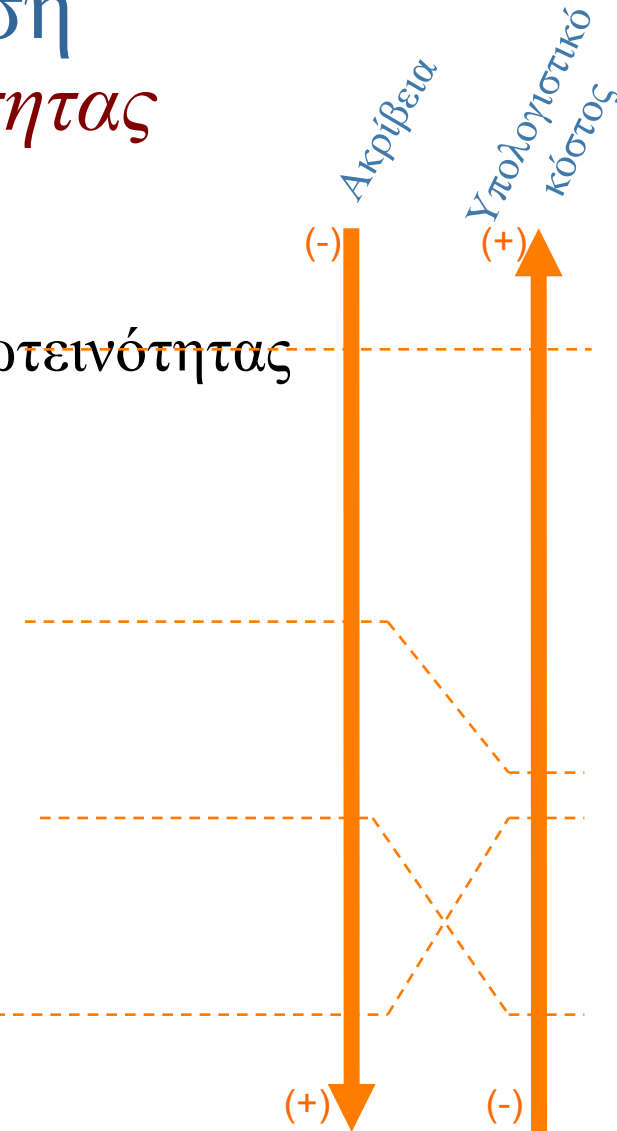
Δεξιά Εικόνα



# Υπολογιστική Όραση

## Ακρίβεια χάρτη ανομοιοτήτας

- Παρεμβολή στη συνάρτηση έντασης φωτεινότητας
  - Δημιουργία εικόνων υψηλότερης ανάλυσης
- Αντιστοίχιση στο πεδίο συχνοτήτων
  - Χρήση πληροφορίας φάσης
- Παρεμβολή στη συνάρτηση κόστους
  - Πολυωνυμική παρεμβολή
- Διαφορική αντιστοίχιση
  - Χρήση πληροφορίας παραγώγου/κλίσης



# Υπολογιστική Όραση

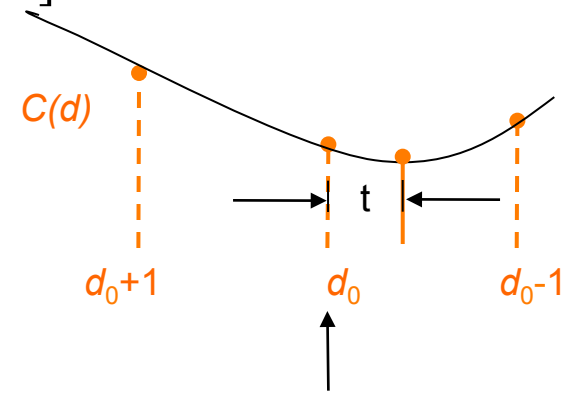
## Παρεμβολή και Διαφορική Αντιστοίχιση

- Παρεμβολή της συνάρτησης κόστους [Anandan '89]

- π.χ. παρεμβολή 2<sup>ου</sup> βαθμού

$$t = \frac{C(d_0 + 1) - C(d_0 - 1)}{4C(d_0) - 2C(d_0 + 1) - 2C(d_0 - 1)}$$

- Βέλτιστη ανομοιότητα:  $d_0 + t$

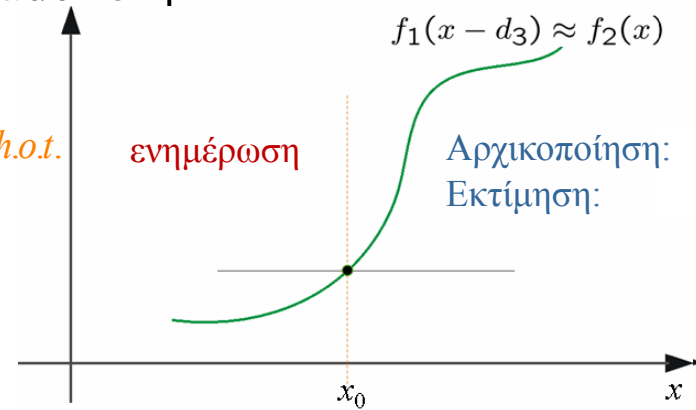


- Διαφορική αντιστοίχιση [Lucas-Kanade '81]

- Χρήση Taylor expansion

$$I_L(x, y) = I_R(x + d + \Delta d, y) \approx I_R(x + d, y) + \frac{\partial I_R(x + d)}{\partial x} \Delta d + h.o.t.$$

- επαναληπτική διαδικασία



# Υπολογιστική Όραση

## Τροποποιημένος συντελεστής συσχέτισης-ECC

- Συντελεστής συσχέτισης

$$\rho_{n,m}(d) = \hat{\mathbf{w}}_L^t(n,m) \hat{\mathbf{w}}_R(n,m-d)$$

- Ακρίβεια εικονοστοιχείου

$$\max_{0 \leq d \leq D} \rho_{n,m}(d)$$

- Πυρήνας παρεμβολής

$$w_R(n,m+\tau) \approx w_R(n,m) + \tau(w_R(n,m) - w_R(n,m-1))$$

- Ενσωμάτωση πυρήνα στο συντελεστή συσχέτισης

$$\rho_{n,m}(d,\tau) = \hat{\mathbf{w}}_L^t(n,m) \hat{\mathbf{w}}_R(n,m-d+\tau)$$



Πρόβλημα Βελτιστοποίησης

$$\max_{0 \leq d \leq D} \max_{\tau} \rho_{n,m}(d,\tau)$$

- Ακρίβεια μικρότερη του εικονοστοιχείου

- Ανεξαρτησία από γραμμικές φωτομετρικές παραμορφώσεις

$$\mathbf{w}(n,m) = \begin{bmatrix} W_{n-3,m-3} \\ W_{n-2,m-3} \\ \vdots \\ W_{n+3,m+3} \end{bmatrix}$$

$$\hat{\mathbf{w}}(n,m) = \frac{\mathbf{w}(n,m) - \bar{\mathbf{w}}(n,m)}{\|\mathbf{w}(n,m) - \bar{\mathbf{w}}(n,m)\|_2}$$



# Υπολογιστική Όραση

## Βελτιστοποίηση ECC

- Συνάρτηση κόστους:

$$\rho_d(\tau) = \frac{\rho_d + \tau(\rho_d - \lambda\rho_{d-1})}{\sqrt{(1 + \lambda^2 - 2\lambda r)\tau^2 + 2(1 - \lambda r)\tau + 1}}$$

$$\text{όπου } \lambda = \frac{\|\mathbf{w}_R(n, m-d-1) - \bar{w}_R(n, m-d-1)\|_2}{\|\mathbf{w}_R(n, m-d) - \bar{w}_R(n, m-d)\|_2}, \quad r = \hat{\mathbf{w}}_R^t(n, m-d)\hat{\mathbf{w}}_R(n, m-d-1)$$

- Δοθέντος  $d_0$ ,

$$\max_{\tau} \rho_{d_0}(\tau)$$

- Κλειστού τύπου λύση οδηγεί σε μηδαμινή αύξηση της πολυπλοκότητας



# Υπολογιστική Όραση

## Υπολογισμός Βέλτιστης Λύσης

Θεώρημα:

Η συνάρτηση  $\rho_{d_0}(\tau)$  παρουσιάζει μοναδικό ακρότατο στη θέση

$$\tau^o = \frac{\rho_{d_0-1} - r\rho_{d_0}}{\lambda(r\rho_{d_0-1} - \rho_{d_0}) + r\rho_{d_0} - \rho_{d_0-1}}.$$

Το ακρότατο αυτό αντιστοιχεί σε ολικό μέγιστο, αν και μόνο αν ο παρανομαστής του  $\tau^o$  είναι αρνητικός.

Στην περίπτωση αυτή, η μέγιστη τιμή της συνάρτησης είναι:

$$\rho_{d_0}(\tau^o) = \sqrt{\frac{\rho_{d_0}^2 + \rho_{d_0-1}^2 - 2r\rho_{d_0}\rho_{d_0-1}}{(1-r^2)}}$$





# Υπολογιστική Όραση

## Αποτελέσματα Προσομοίωσης

Root Mean Square Error				
	Form I		Form II	
shift	NCC	ENCC	NCC	ENCC
0.0613	0.0818	0.0017	0.1145	0.0053
0.1111	0.0800	0.0028	0.1116	0.0088
0.3333	0.0581	0.0064	0.0832	0.0170
0.5000	0.0324	0.0099	0.0590	0.0182
0.8122	0.0758	0.0046	0.1135	0.0122

### Τεχνητές εικόνες

#### Form I

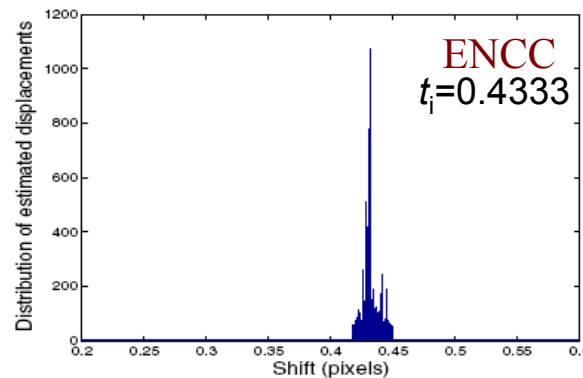
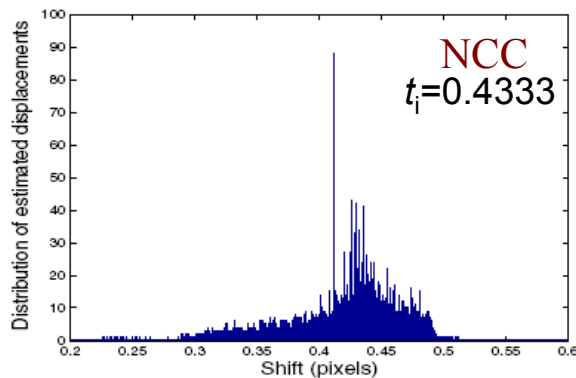
$$R(i, j) = 120\text{sinc}(k_x(i - 50, 1))\text{sinc}(k_y(j - 50, 1))$$

$$L(i, j) = R(i, j - t_j)$$

#### Form II

$$R(i, j) = \frac{1}{2} + \frac{1}{4} \left( \cos\left(\frac{\pi i^2}{P}\right) + \cos\left(\frac{\pi j^2}{P}\right) \right)$$

$$L(i, j) = R(i, j - t_j)$$



Ολική  
Μετατόπιση  
Εικόνων



# Υπολογιστική Όραση

## Στερεοσκοπικές Εικόνες

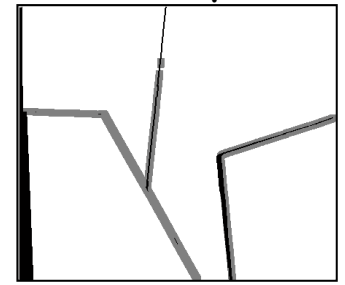
Αριστερή εικόνα

Δεξιά εικόνα

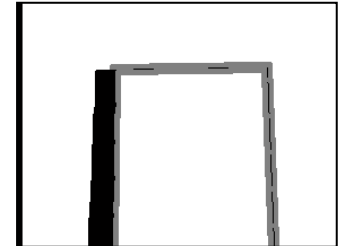
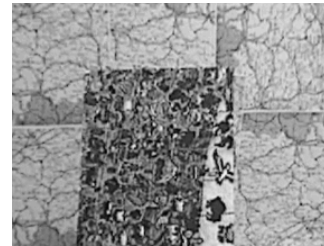
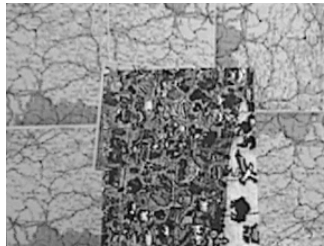
Χάρτης  
Ανομοιότητας

Ασυνέχειες  
Αποκλεισμοί

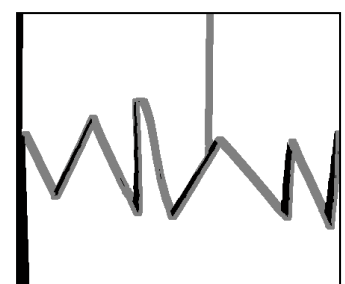
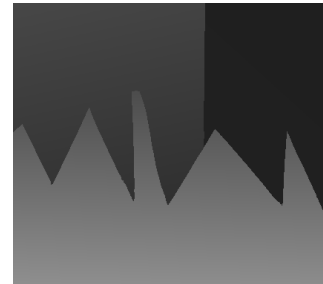
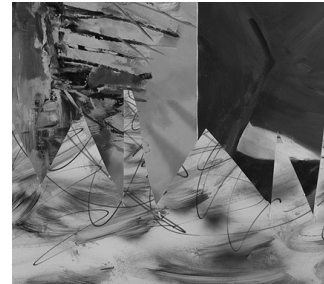
Venus



Map



Sawtooth



# Υπολογιστική Όραση

## Αποτελέσματα Προσομοίωσης

- Τύπος μέτρησης σφάλματος

$$B_R = \frac{1}{N_R} \sum_{(x,y) \in R} |d_C(c,y) - d_G(x,y)| > \delta$$

- Περιοχή ενδιαφέροντος:  $R = \bar{D} \setminus \bar{O}$

$d_C(x,y)$ : εκτίμηση ανομοιότητας

$d_G(x,y)$ : πραγματική ανομοιότητα

$N_R$ : μέγεθος περιοχής ενδιαφέροντος

$D$ : περιοχή ασυνεχειών

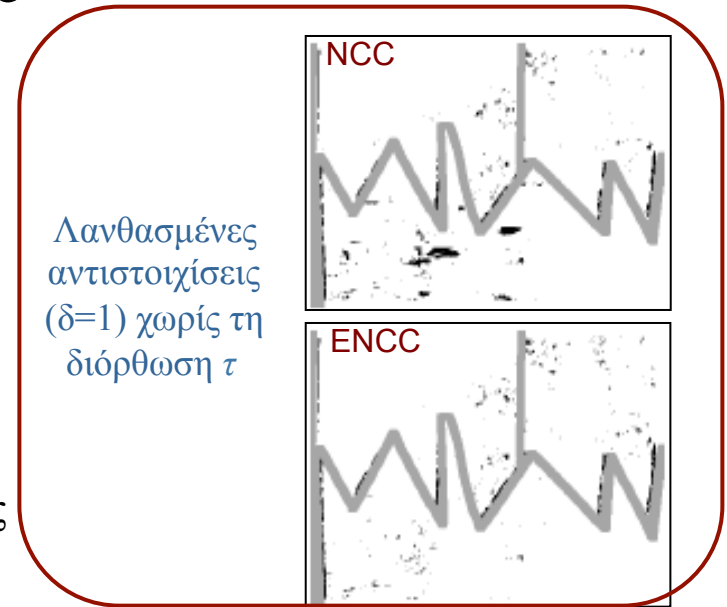
$O$ : περιοχή αποκλεισμών

% $B_R$	Map		Sawtooth		Venus	
	ENCC	NCC	ENCC	NCC	ENCC	NCC
0.25	20.18	24.82	29.69	28.58	15.44	18.13
0.50	3.23	3.42	9.40	9.74	5.72	6.54
0.75	1.20	1.35	4.54	4.96	4.32	4.54
1.00	0.82	1.01	2.72	3.07	3.75	4.02

Αρχικές  
Εικόνες

% $B_R$	Map		Sawtooth		Venus	
	ENCC	NCC	ENCC	NCC	ENCC	NCC
0.25	18.34	23.71	27.95	27.46	12.80	16.32
0.50	2.71	3.08	7.97	8.56	3.91	5.05
0.75	1.08	1.22	3.70	4.26	2.75	3.19
1.00	0.74	0.93	1.99	2.49	2.39	2.89

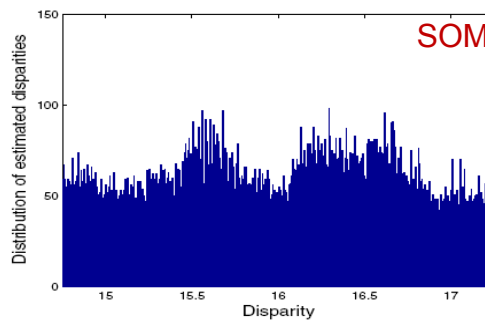
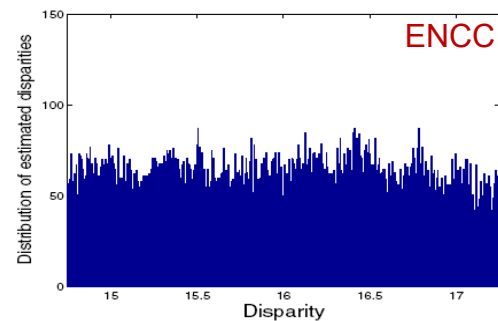
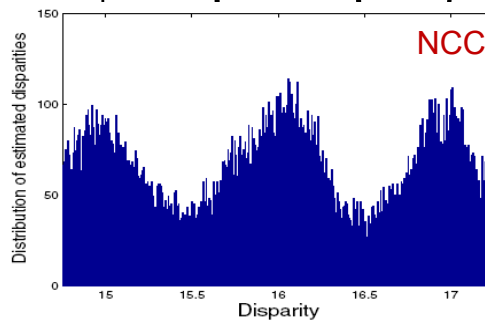
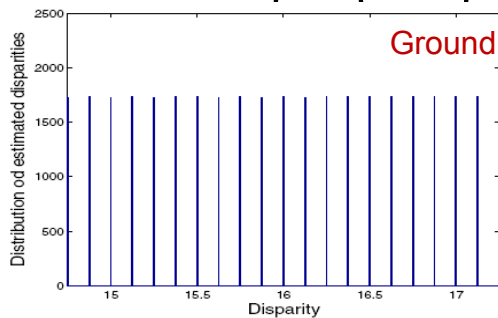
Φωτομετρικά  
παραμορφωμένες  
εικόνες



# Υπολογιστική Όραση

## Αποτελέσματα Προσομοίωσης

- Pixel locking effect (Shimizu-Okutomi '01)
  - Η τάση της κατανομής εκτιμήσεων να δημιουργεί λοβούς γύρω από ακέραιες τιμές
  - Μερική ακύρωση του φαινομένου με πρωθύστερη δράση



Κατανομή εκτιμήσεων ανομοιότητας στην περιοχή [15-,17+] για την εικόνα Sawtooth

SOM: Shimizu-Okutomi Modification



# Υπολογιστική Όραση

## Συμπεράσματα

- Εύρωστη σε φωτομετρικές παραμορφώσεις
- Παροχή ανομοιοτήτων με ακρίβεια μικρότερη του εικονοστοιχείου
- Μικρό υπολογιστικό κόστος
  - Κλειστού τύπου λύση
- Απαλλαγή από το *pixel locking effect*
- Χρήση της βέλτιστης λύσης ως ανιχνευτή προβληματικών σημείων (ημι-αποκλεισμένες περιοχές)



# Υπολογιστική Όραση

## Παραμετρικές Τεχνικές

- Area-based (direct) τεχνικές
  - Αντιστοίχιση βασισμένη στην ένταση φωτεινότητας όλων των εικονοστοιχείων της ROI
    - Απευθείας αναζήτηση παραμετρικού μοντέλου
- Featured-based τεχνικές
  - Αντιστοίχιση βασισμένη σε επιλεγμένα χαρακτηριστικά (γωνίες, ακμές) της ROI
    - Χρήση τελεστή αναγνώρισης χαρακτηριστικών
    - Αντιστοίχιση κοινών χαρακτηριστικών
    - Χρήση παραμετρικού μοντέλου για τη συνολική αντιστοίχιση δοθείσης της αντιστοίχισης χαρακτηριστικών



# Υπολογιστική Όραση

## *Area-based* παραμετρικές τεχνικές

- Ορισμός παραμετρικού μοντέλου
  - Βάσει της φύσης και των απαιτήσεων του προβλήματος
- Ορισμός συνάρτησης κόστους
- Βελτιστοποίηση συνάρτησης κόστους
  - Υπολογισμός των παραμέτρων που βελτιστοποιούν τη συνάρτηση κόστους

### Παράδειγμα

$$W(\mathbf{x}; \mathbf{p}) = \begin{bmatrix} p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

$$\mathbf{x} = [x, y]^t, \mathbf{p} = [p_1, p_2, \dots, p_6]^t$$

$$E(\mathbf{p}) = \sum_{\mathbf{x} \in \text{ROI}} |I_1(\mathbf{x}) - I_2(W(\mathbf{x}; \mathbf{p}))|^2$$

$$\min_{\mathbf{p}} E(\mathbf{p})$$



# Υπολογιστική Όραση

## Τεχνικές Βελτιστοποίησης

- Μέθοδοι πλήρους αναζήτησης (full search)
  - Αναλυτική αναζήτηση των  $N$  παραμέτρων στον  $N$ -D χώρο
  - (-) Υψηλό υπολογιστικό κόστος
  - (-) Πεπερασμένη ακρίβεια
  - (+) Αντιστάθμιση μεγάλων μετατοπίσεων
- Μέθοδοι βασισμένες στην κλίση της έντασης των εικόνων (gradient-based)
  - (+) Μεγαλύτερη ακρίβεια (θεωρητικά ίση με το  $\epsilon$  της μηχανής)
  - (+) Μικρό υπολογιστικό κόστος
  - (+) Δυνατότητα χρήσης επαναληπτικού σχήματος
    - εγκλωβισμός
  - (-) Αδυναμία διαχείρισης μεγάλων μετατοπίσεων
    - Χρήση πυραμιδικού σχήματος
- Υβριδικές μέθοδοι





# Υπολογιστική Όραση

## Γενικό πρόβλημα Ευθυγράμμισης εικόνων

- Ορισμός παραμετρικού μοντέλου  $W(\mathbf{x}; \mathbf{p})$ 
  - $\mathbf{x}=[x,y]^t$ ,  $\mathbf{p}=[p_1,p_2,\dots,p_n]^t$
- Ορισμός μέτρου ομοιότητας μεταξύ εικόνας αναφοράς  $I_R$  (reference image) και γεωμετρικά παραμορφωμένης εικόνας  $I_W$  (warped image)
  - Αντιστάθμιση *φωτομετρικών παραμορφώσεων*



# Υπολογιστική Όραση

## Αντιστάθμιση φωτομετρικών παραμορφώσεων (*contrast-brightness*)

*Lucas – Kanade '81*

$$E_{LK}(a_1, a_2, \mathbf{p}) = \sum_{\mathbf{x} \in ROI} \left( a_1 I_r(\mathbf{x}) + a_2 - I_w(T(\mathbf{x}; \mathbf{p})) \right)^2$$

$$\min_{a_1, a_2, \mathbf{p}} E_{LK}(a_1, a_2, \mathbf{p}) \quad \underline{\text{Επαναληπτικός αλγόριθμος}}$$

*Fuh – Maragos '91*

$$E_{FM}(a_1, a_2, \mathbf{p}) = \sum_{\mathbf{x} \in ROI} \left( I_r(\mathbf{x}) - a_1 I_w(T(\mathbf{x}; \mathbf{p})) - a_2 \right)^2$$

$$\min_{a_1, a_2, \mathbf{p}} E_{FM}(a_1, a_2, \mathbf{p}) \quad \underline{\text{Αναλυτική Αναζήτηση}}$$

*ECC '08*

$$\rho(\mathbf{p}) = \hat{\mathbf{i}}_r^t(\mathbf{x}) \frac{\bar{\mathbf{i}}_w(T(\mathbf{x}; \mathbf{p}))}{\|\bar{\mathbf{i}}_w(T(\mathbf{x}; \mathbf{p}))\|}$$

$$\max_{\mathbf{p}} \rho(\mathbf{p}) \quad \underline{\text{Επαναληπτικός αλγόριθμος}}$$



# Υπολογιστική Όραση

## Σχέση μεταξύ αλγορίθμων

- Ελαχιστοποίηση ως προς τις φωτομετρικές παραμορφώσεις (separable variables)

– *LK*: 
$$\min_{a_1, a_2, \mathbf{p}} E_{LK}(a_1, a_2, \mathbf{p}) \equiv \min_{\mathbf{p}} \left\| \hat{\mathbf{i}}_w(\mathbf{p}) \right\|^2 (1 - \rho(\mathbf{p})^2)$$

– *FM*: 
$$\min_{a_1, a_2, \mathbf{p}} E_{FM}(a_1, a_2, \mathbf{p}) \equiv \min_{\mathbf{p}} \left\| \hat{\mathbf{i}}_r \right\|^2 (1 - \rho(\mathbf{p})^2)$$

Κανένα από τα δύο προβλήματα  
δεν είναι ισοδύναμο με το

$$\max_{\mathbf{p}} \rho(\mathbf{p})$$

Μόνη περίπτωση ισοδυναμίας:

$$\min_{a_1 \geq 0, a_2, \mathbf{p}} E_{FM}(a_1, a_2, \mathbf{p}) \equiv \max_{\mathbf{p}} \rho(\mathbf{p})$$



# Υπολογιστική Όραση

## Αλγόριθμος ECC – Βασική ιδέα

- Κανόνας ενημέρωσης:  $\mathbf{p} = \mathbf{p}_0 + \Delta\mathbf{p}$

- Προσέγγιση  $\mathbf{i}_w(\mathbf{p}) \approx \mathbf{i}_w(\mathbf{p}_0) + \mathbf{G}(\mathbf{p}_0)\Delta\mathbf{p}$

$$\Rightarrow \rho(\mathbf{p}); \rho(\Delta\mathbf{p}; \mathbf{p}_0) = \frac{\hat{\mathbf{i}}_r^t \bar{\mathbf{i}}_w + \hat{\mathbf{i}}_r^t \bar{\mathbf{G}} \Delta\mathbf{p}}{\sqrt{\|\bar{\mathbf{i}}_w\|^2 + 2\bar{\mathbf{i}}_w^t \bar{\mathbf{G}} \Delta\mathbf{p} + \Delta\mathbf{p}^t \bar{\mathbf{G}}^t \bar{\mathbf{G}} \Delta\mathbf{p}}}$$

- $\bar{\mathbf{G}}$  η Ιακωβιανή μήτρα του  $\bar{\mathbf{i}}_w$  ως προς τις παραμέτρους
- Ακολουθία υποδεέστερων μη γραμμικών προβλημάτων βελτιστοποίησης

$$\max_{\Delta\mathbf{p}_j} \rho(\Delta\mathbf{p}_j; \mathbf{p}_{j-1})$$



# Υπολογιστική Όραση

## Υπολογισμός βέλτιστης λύσης

### ΘΕΩΡΗΜΑ

Η συνάρτηση  $\rho(\Delta \mathbf{p}_j; \mathbf{p}_{j-1})$  μεγιστοποιείται για

$$\Delta \mathbf{p}^o = (\bar{\mathbf{G}}^t \bar{\mathbf{G}})^{-1} \bar{\mathbf{G}}^t (\lambda \hat{\mathbf{i}}_r - \bar{\mathbf{i}}_w)$$

• Αν  $\hat{\mathbf{i}}_r^t (\mathbf{I}_N - \mathbf{P}_G) \bar{\mathbf{i}}_w > 0$  τότε το  $\Delta \mathbf{p}^o$  είναι ολικό μέγιστο για

$$\lambda = \frac{\|\bar{\mathbf{i}}_w\|^2 - \bar{\mathbf{i}}_w^t \mathbf{P}_G \bar{\mathbf{i}}_w}{\hat{\mathbf{i}}_r^t \bar{\mathbf{i}}_w - \hat{\mathbf{i}}_r^t \mathbf{P}_G \bar{\mathbf{i}}_w}$$

όπου  $\mathbf{P}_G = \bar{\mathbf{G}} (\bar{\mathbf{G}}^t \bar{\mathbf{G}})^{-1} \bar{\mathbf{G}}^t$

• Αν  $\hat{\mathbf{i}}_r^t (\mathbf{I}_N - \mathbf{P}_G) \bar{\mathbf{i}}_w \leq 0$  τότε το  $\Delta \mathbf{p}^o$  είναι το άκρο ενός διαστήματος και το  $\lambda$  μπορεί να πάρει οποιαδήποτε τιμή ικανοποιεί τους περιορισμούς

$$\rho(\Delta \mathbf{p}_j^o; \mathbf{p}_{j-1}) > \rho(\mathbf{0}; \mathbf{p}_{j-1})$$

$$\rho(\Delta \mathbf{p}_j^o; \mathbf{p}_{j-1}) \geq 0$$



# Υπολογιστική Όραση

## Υπολογισμός βέλτιστης λύσης

### ΛΗΜΜΑ

Μια ικανή συνθήκη για να ισχύουν οι περιορισμοί είναι:

$$\lambda \geq \max \{ \lambda_1, \lambda_2 \}$$

όπου

$$\lambda_1 = \sqrt{\frac{\bar{\mathbf{i}}_w^t \mathbf{P}_G \bar{\mathbf{i}}_w}{\hat{\mathbf{i}}_r^t \mathbf{P}_G \bar{\mathbf{i}}_r}}, \quad \lambda_2 = \frac{\hat{\mathbf{i}}_r^t \mathbf{P}_G \bar{\mathbf{i}}_w - \hat{\mathbf{i}}_r^t \bar{\mathbf{i}}_w}{\hat{\mathbf{i}}_r^t \mathbf{P}_G \bar{\mathbf{i}}_r}$$



# Υπολογιστική Όραση

## *Βήματα Αλγορίθμου FA-ECC*

- Αρχικοποίηση  $\mathbf{p}_0$ 
  - $j=1$
- Επαναληπτική διαδικασία
  1. Υπολόγισε την εικόνα  $I_w(W(\mathbf{x};\mathbf{p}_{j-1}))$
  2. Υπολόγισε την Ιακωβιανή μήτρα  $G(\mathbf{p}_{j-1})$
  3. Υπολόγισε τη βέλτιστη λύση  $\Delta\mathbf{p}_j$  σύμφωνα με το θεώρημα και το λήμμα
  4. Ενημέρωσε τις παραμέτρους  $\mathbf{p}_j = \mathbf{p}_{j-1} + \Delta\mathbf{p}_j$ 
    - Αν  $\|\Delta\mathbf{p}_j\| > \epsilon$ , τότε  $j++$  και πήγαινε στο 1. Διαφορετικά σταμάτα.



# Υπολογιστική Όραση

## Αντίστροφο πρόβλημα – Σύνθεση μετασχηματισμών

- Αντίστροφο πρόβλημα [Hager-Belhumeur '98]
  - Υπολόγισε πως πρέπει να μετασχηματίσεις την  $I_R$  για να αντιστοιχιστεί με την  $I_W$
  - Εφάρμοσε τον **αντίστροφο** μετασχηματισμό στην  $I_W$

Η Hessian μήτρα της βέλτιστης λύσης γίνεται ανεξάρτητη των παραμέτρων

- Σύνθεση μετασχηματισμών [Shum-Szeliski '00]
  - Κανόνας ενημέρωσης
    - $W(\mathbf{x}; \mathbf{p}_j) = W(\mathbf{x}; \mathbf{p}_{j-1}) \circ W(\mathbf{x}; \Delta \mathbf{p}_j)$

Η Ιακωβιανή του μετασχηματισμού γίνεται ανεξάρτητη των παραμέτρων





# Υπολογιστική Όραση

## *Βήματα Αλγορίθμου IC-ECC*

- Αρχικοποίηση  $\mathbf{p}_0$ 
  - $j=1$
  - Υπολόγισε την Ιακωβιανή μήτρα  $G_r(\mathbf{p}_{j-1})$  και τον αντίστροφο  $(G_r^T G_r)^{-1}$
- Επαναληπτική διαδικασία
  1. Υπολόγισε την εικόνα  $I_w(W(\mathbf{x}; \mathbf{p}_{j-1}))$
  2. Υπολόγισε τη βέλτιστη λύση  $\Delta \mathbf{p}_j$  σύμφωνα με το θεώρημα και το λήμμα
  3. Ενημέρωσε το μοντέλο  $W(\mathbf{x}; \mathbf{p}_j) = W(\mathbf{x}; \mathbf{p}_{j-1}) \circ W(\mathbf{x}; \Delta \mathbf{p}_j)^{-1}$ 
    - Αν  $\|\Delta \mathbf{p}_j\| > \epsilon$ , τότε  $j++$  και πήγαινε στο 1. Διαφορετικά σταμάτα.



# Υπολογιστική Όραση

## Επαναληπτικοί Αλγόριθμοι-Σύγκριση

	Πολυπλοκότητα ( $N$ : αριθμός παραμέτρων $K$ : αριθμός εικ/χίων)	Δυνατότητα Εφαρμογής	Ευαισθησία στο θόρυβο
Lucas-Kanade '81 (Forwards Additive LK)	$O(KN^2)$	Οποιοδήποτε μοντέλο	Μικρή
Haager-Belhumeur '98 (Inverse Additive LK)	$O(KN)$	Γραμμικό 2D	Μεγάλη
Shum-Szeliski '00 (Forwards – Compositional LK)	$O(KN^2)$	Ημι-ομάδα	Μικρή
Baker-Matthews '04 (Inverse Compositional LK)	$O(KN)$	Ομάδα	Μεγάλη
FA-ECC (2007)	$O(KN^2)$	Οποιοδήποτε μοντέλο	Μικρή
IC-ECC (2008)	$O(KN)$	Ομάδα	Μεγάλη

