



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

ΑΝΟΙΚΤΑ ακαδημαϊκά
μαθήματα ΠΠ

Τεχνητή Νοημοσύνη II

Ενότητα 5: Αντίληψη

Μουστάκας Κωνσταντίνος

Πολυτεχνική Σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας
Υπολογιστών

Σκοποί ενότητας

- Αντίληψη



Περιεχόμενα ενότητας

➤ Αντίληψη



Αντίληψη

Τεχνητή Νοημοσύνη II

Αντίληψη (4)

Σύνοψη

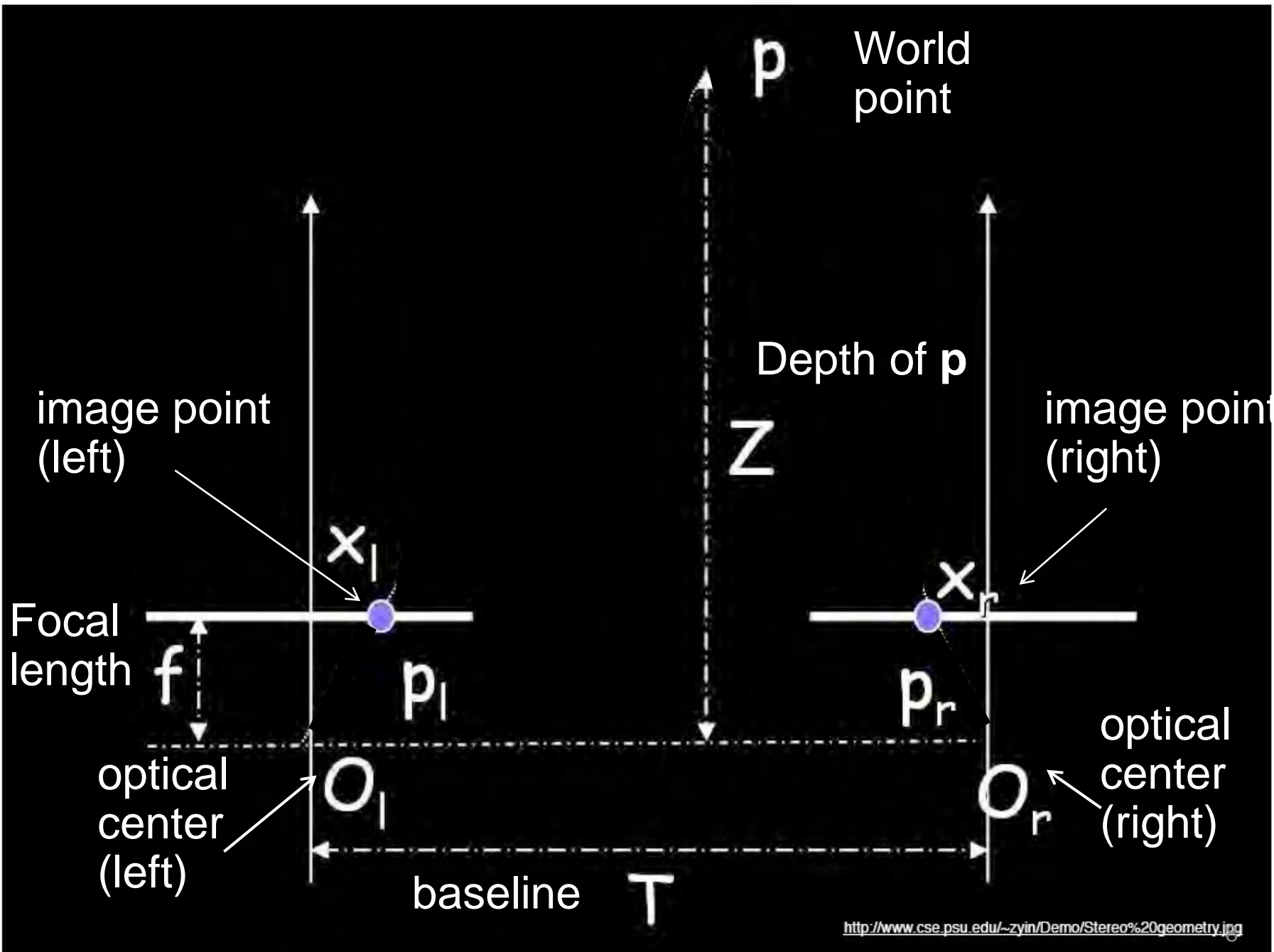
- Διόφθαλμη όραση στον άνθρωπο
- Στερεογράμματα
- Επιπολική γεωμετρία και επιπολικοί περιορισμοί
 - Παράδειγμα με παράλληλους οπτικούς άξονες
 - Γενική περίπτωση με βαθμονομημένες κάμερες



Γεωμετρία στερεοσκοπικού συστήματος

- Υποθέτουμε, παράλληλους οπτικούς άξονες, γνωστές παραμέτρους κάμερας (π.χ. βαθμονομημένες κάμερες):

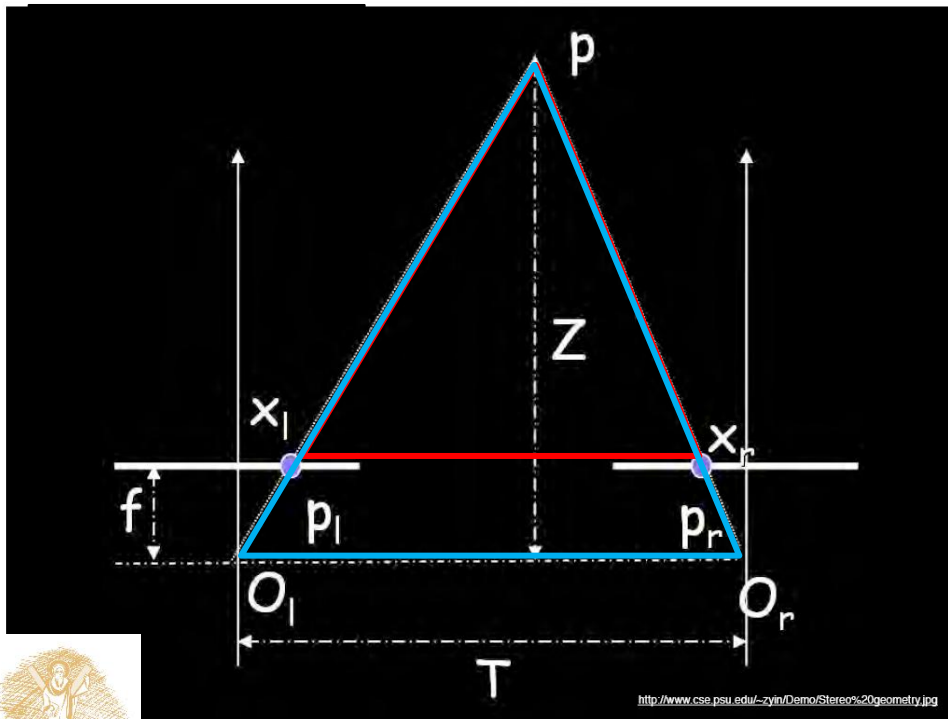




Γεωμετρία στερεοσκοπικού συστήματος

- Υποθέτουμε, παράλληλους οπτικούς άξονες, γνωστές παραμέτρους κάμερας (π.χ. βαθμονομημένες κάμερες):

Όμοια τρίγωνα (p_l, P, p_r) και (O_l, P, O_r) :



$$\frac{T - \Delta x}{Z - f} = \frac{T}{Z}$$

$$Z = f \frac{T}{x_r - x_l}$$

Disparity

Ανομοιότητα

$$x_r - x_l$$

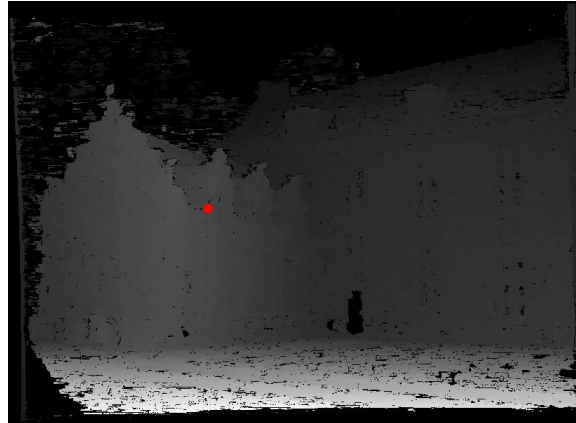


Βάθος βάσει ανομοιότητας

Εικόνα $I(x,y)$



Disparity map $D(x,y)$



Εικόνα $I'(x',y')$

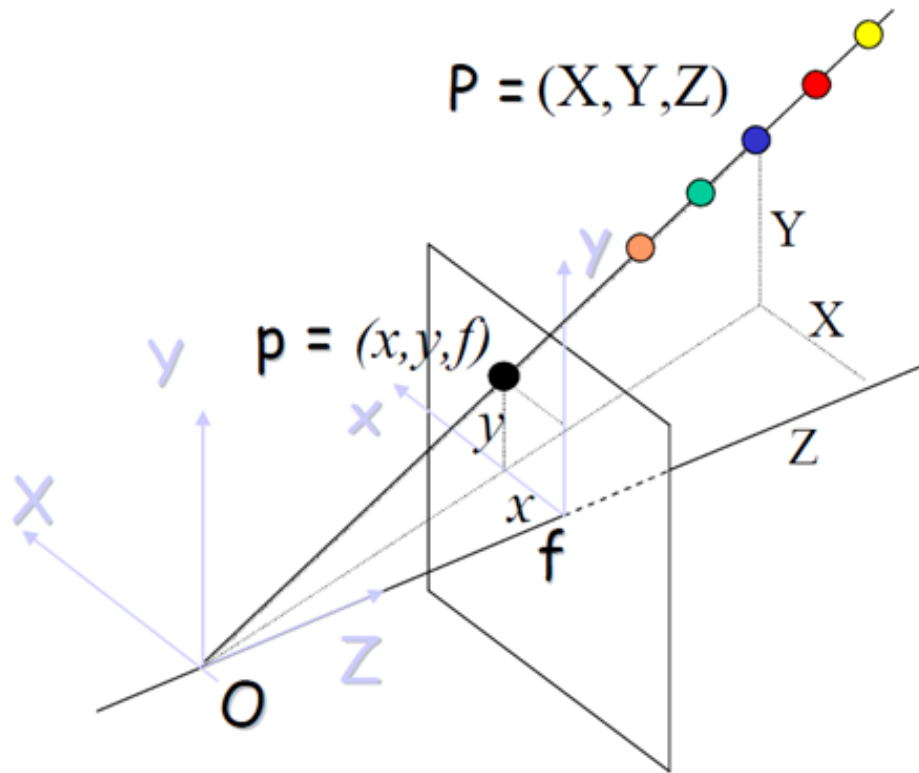


$$(x',y')=(x+D(x,y), y)$$

Οπότε εάν μπορούμε να βρούμε τα **αντίστοιχα σημεία** και στις δύο εικόνες, μπορούμε να υπολογίσουμε σχετικό βάθος



Γιατί στερεοσκοπική όραση;

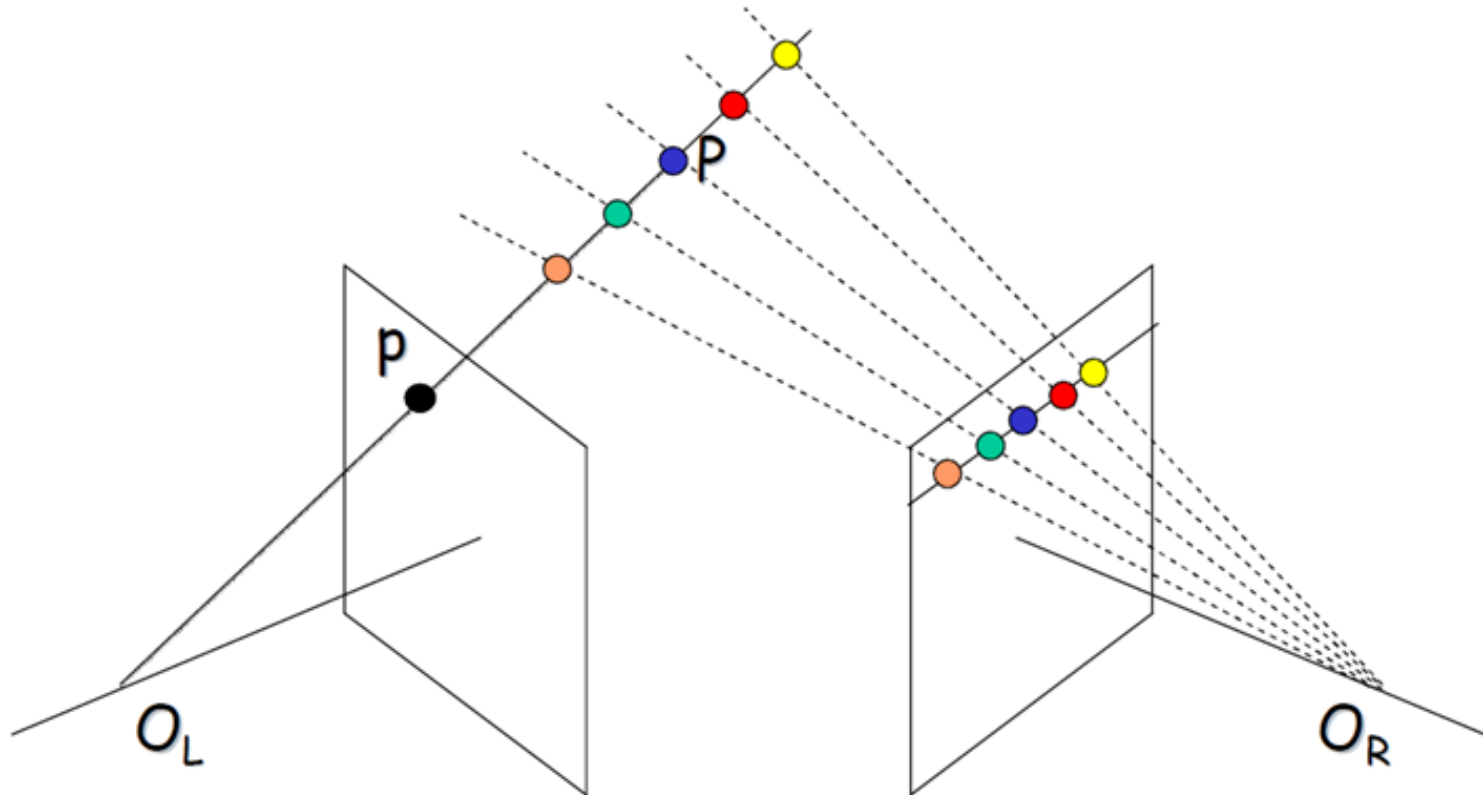


$$x = f \frac{X}{Z} = f \frac{kX}{kZ}$$
$$y = f \frac{Y}{Z} = f \frac{kY}{kZ}$$

Θεμελιώδες πρόβλημα: Όλα τα σημεία στην ακτίνα OP έχουν ίχνος ρ.



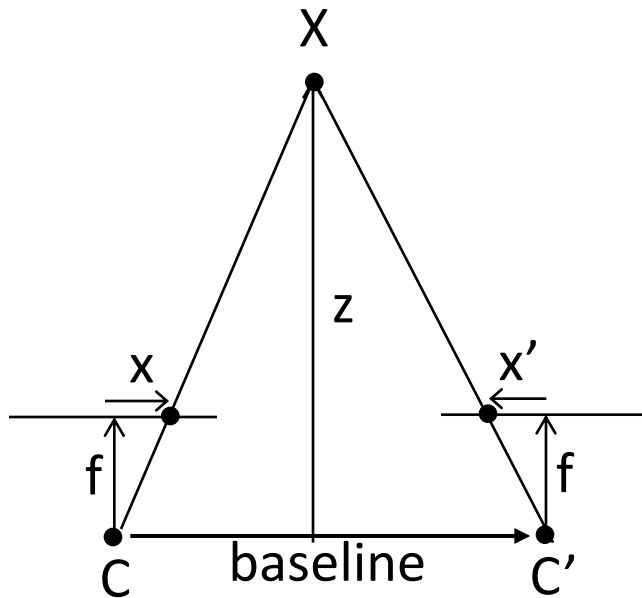
Γιατί στερεοσκοπική όραση;



Μία δεύτερη κάμερα μπορεί να άρει την απροσδιοριστία, επιτρέποντας τον υπολογισμό βάθους μέσω της τριγωνοποίησης



Βάθος βάσει ανομοιότητας



$$(X - X') / f = \text{baseline} / z$$

$$X - X' = (\text{baseline} * f) / z$$

$$z = (\text{baseline} * f) / (X - X')$$



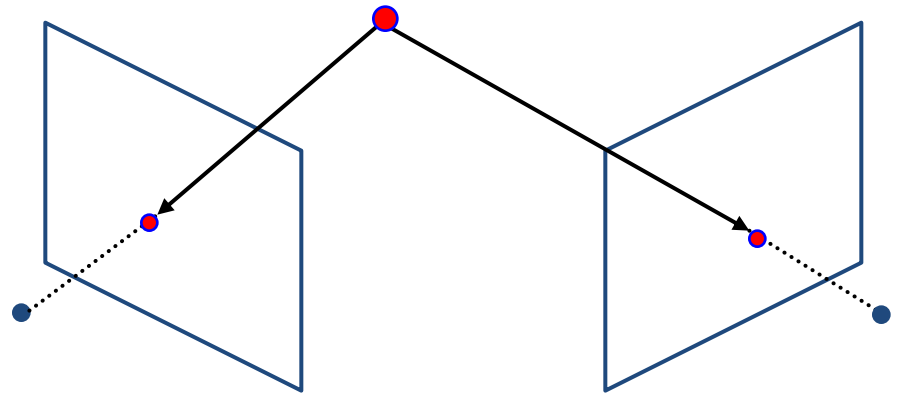
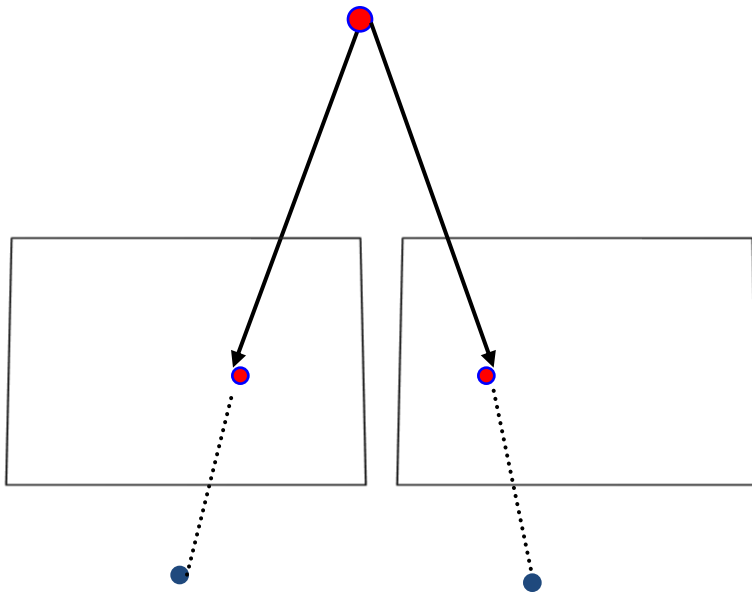
Σύνοψη

- Διόφθαλμη όραση στον άνθρωπο
- Στερεογράμματα
- Επιπολική γεωμετρία και επιπολικοί περιορισμοί
 - Παράδειγμα με παράλληλους οπτικούς άξονες
 - Γενική περίπτωση με βαθμονομημένες κάμερες

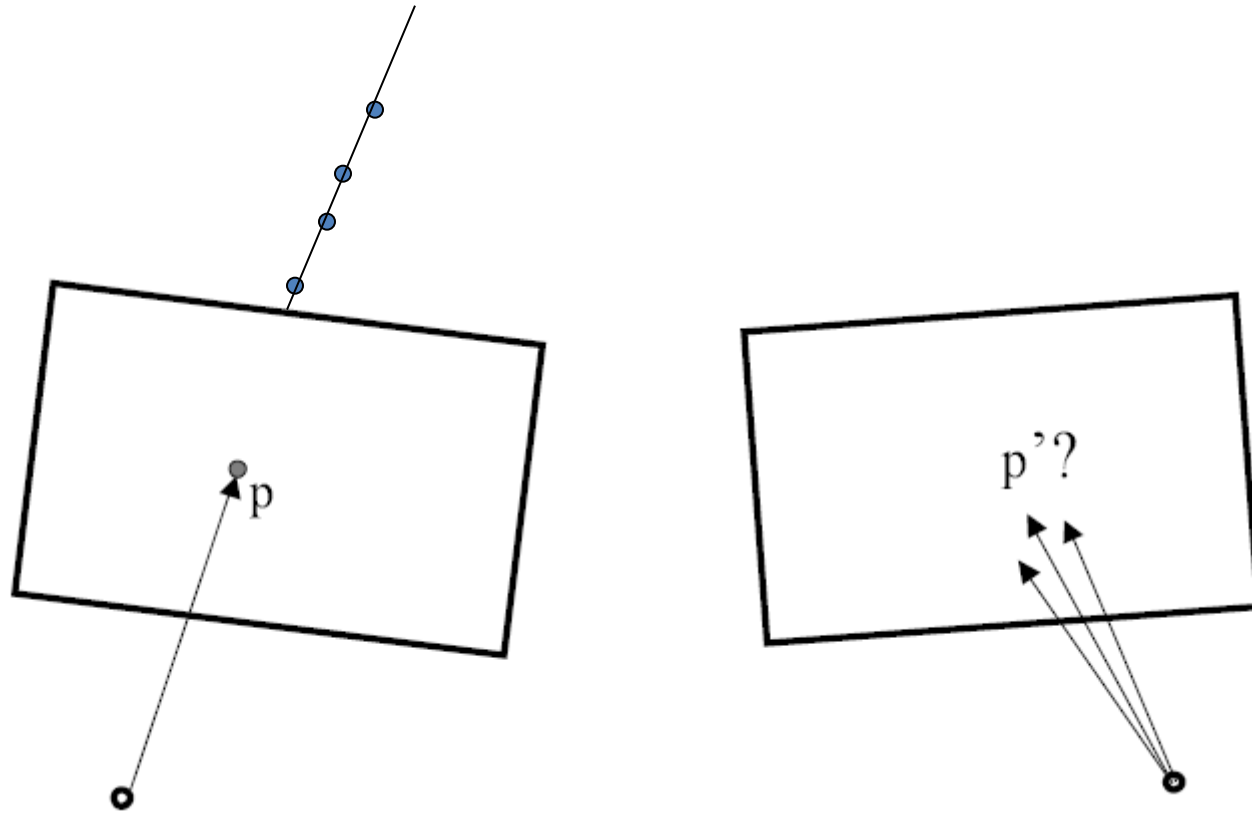


Βαθμονομημένες κάμερες

- Οι δύο κάμερες δεν είναι απαραίτητο να έχουν παράλληλους οπτικούς άξονες



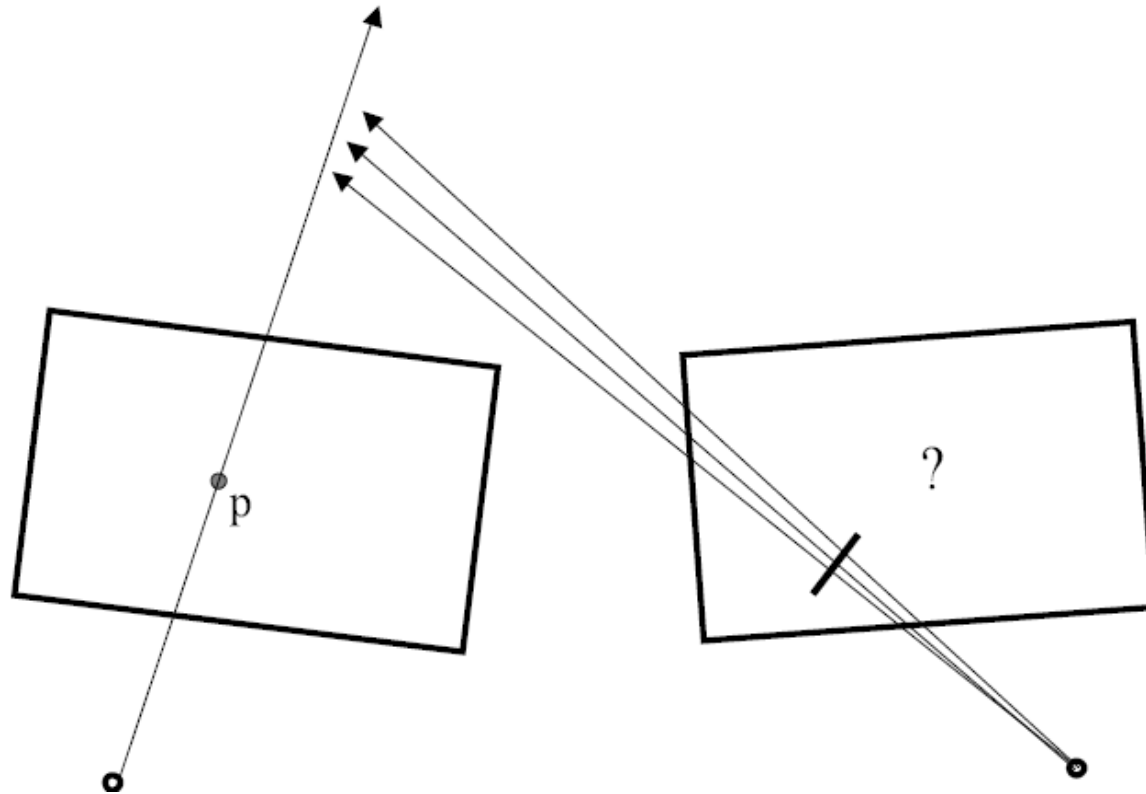
Στερεοσκοπική αντιστοίχιση



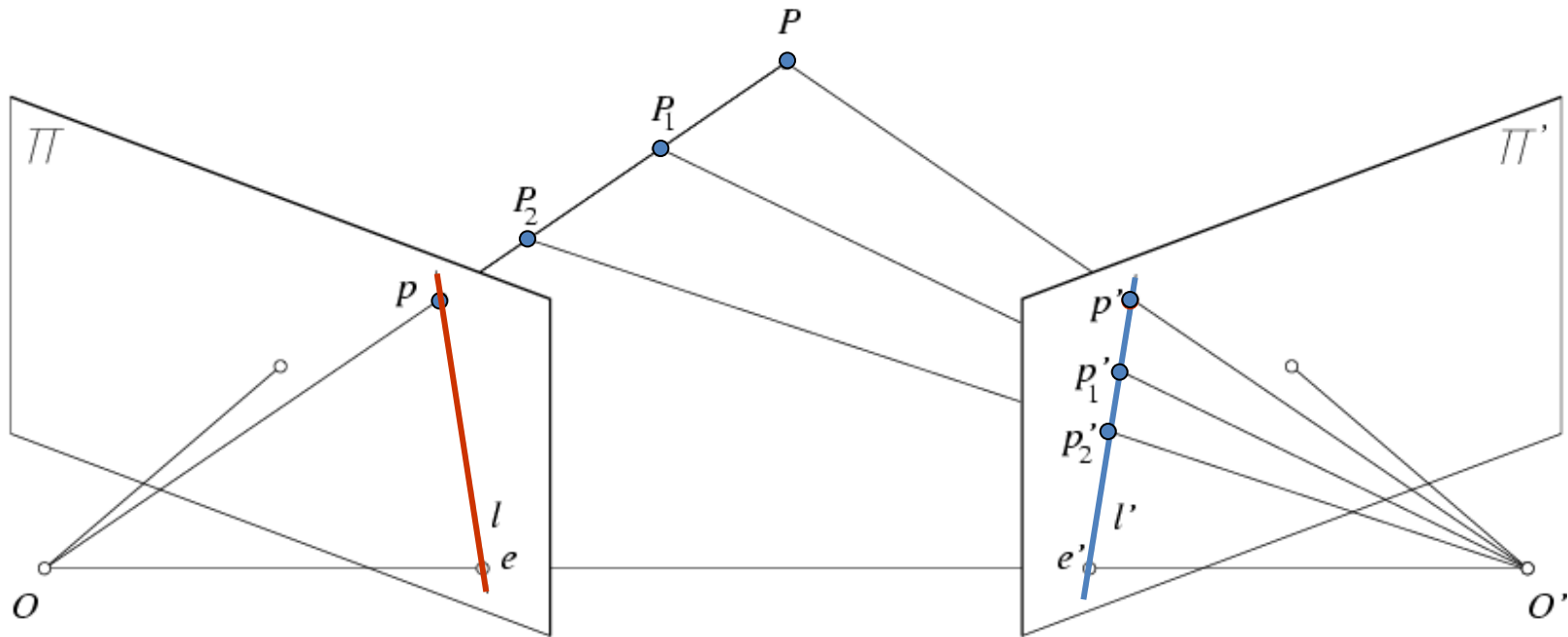
- Δοσμένου του σημείου p στην αριστερή εικόνα, που μπορεί να βρίσκεται το αντίστοιχο σημείο p' ;



Περιορισμοί στερεοσκοπικής αντιστοίχισης



Επιπολικός περιορισμός

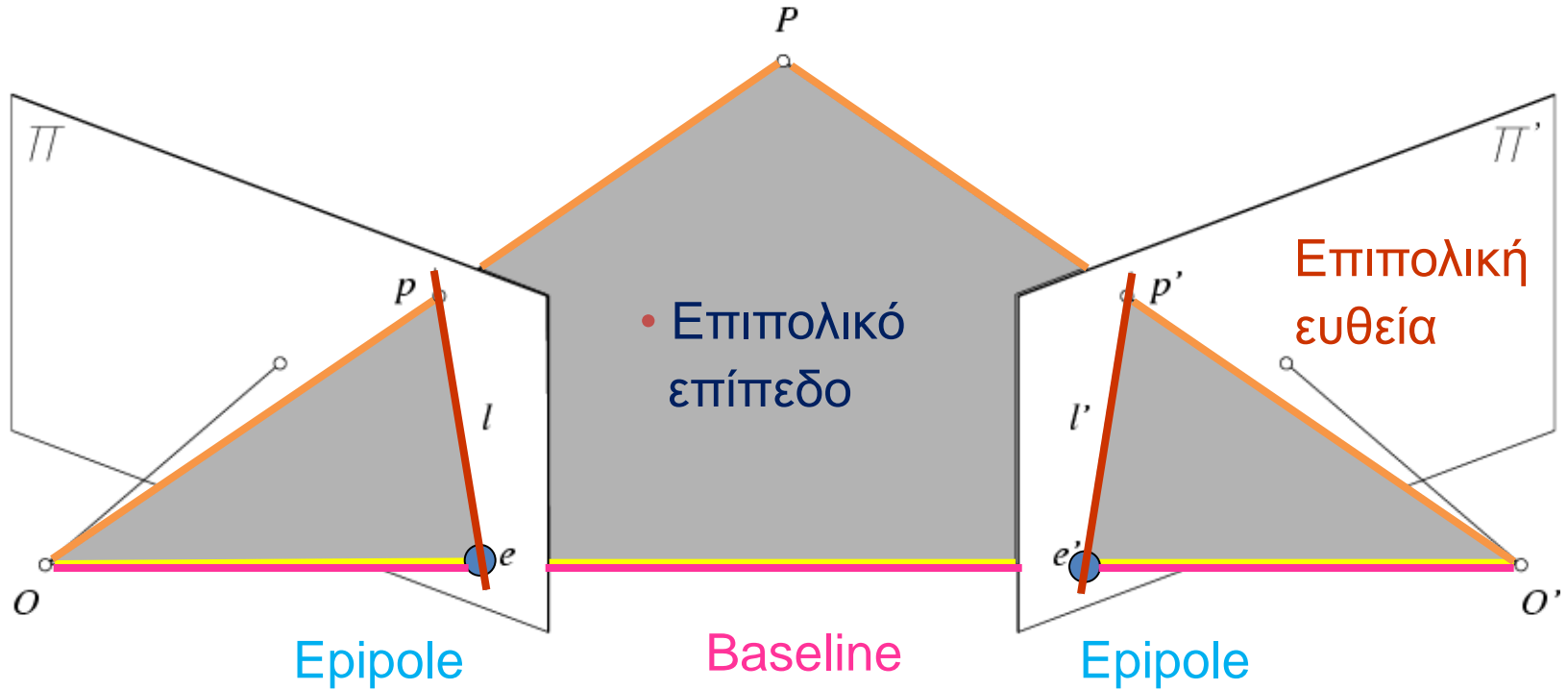


Η γεωμετρία βάσει των δύο όψεων περιορίζει την πιθανή θέση του αντίστοιχου σημείου στη δεύτερη εικόνα, ενός δοσμένου σημείου της αρχικής εικόνας.

- Πρέπει να βρίσκεται πάνω σε μία ευθεία που προκύπτει από την τομή του Π' και του επιπέδου που ορίζεται από το σημείο P και τα οπτικά κέντρα.



Επιπολική γεωμετρία



Επιπολική γεωμετρία: Όροι

- **Baseline:** ευθεία που ενώνει τα κέντρα προβολής
- **Erirole:** σημείο τομής της baseline με το επίπεδο της εικόνας
- **Eriplar plane:** επίπεδο που περιέχει την baseline και το 3Δ σημείο
- **Eriplar line:** τομή του επιπολικού επιπέδου με το επίπεδο της εικόνας

- Όλες οι επιπολικές ευθείες τέμνονται στο επιπολικό σημείο
- Ένα επιπολικό επίπεδο τέμνει τα επίπεδα της αριστερής και της δεξιάς εικόνας στις επιπολικές ευθείες



Γιατί είναι σημαντικός ο επιπολικός περιορισμός;

Επιπολικός περιορισμός



Είναι χρήσιμος γιατί ανάγει το πρόβλημα αντιστοίχισης σε πρόβλημα μονοδιάστατης αναζήτησης πάνω στην επιπολική ευθεία.

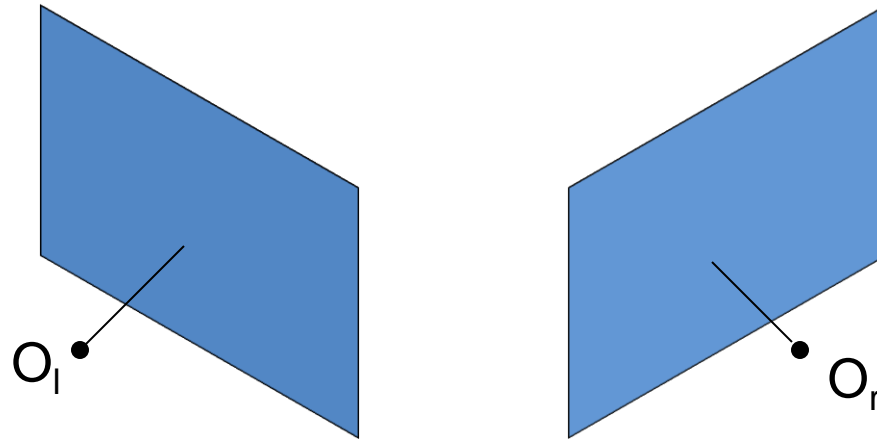


Παράδειγμα

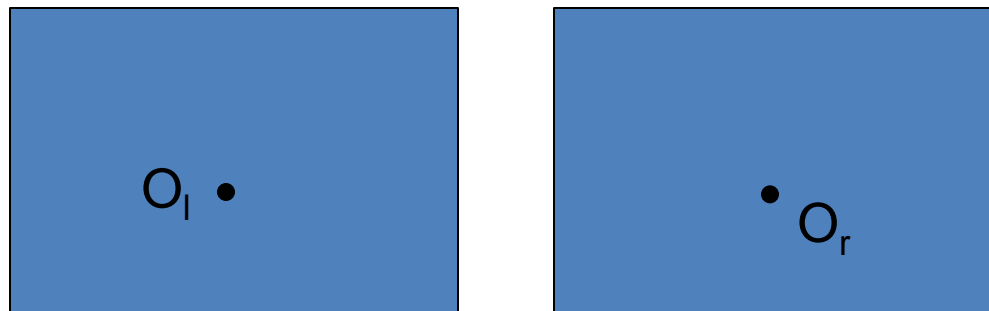


Πως θα ήταν οι επιπολικές ευθείες;

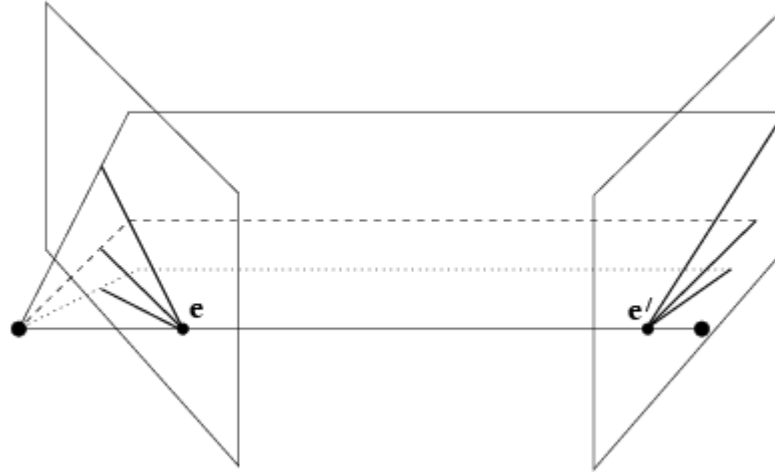
1.



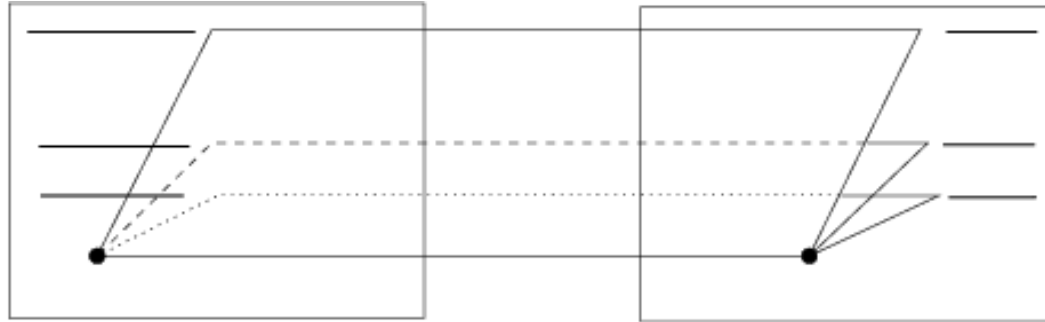
2.



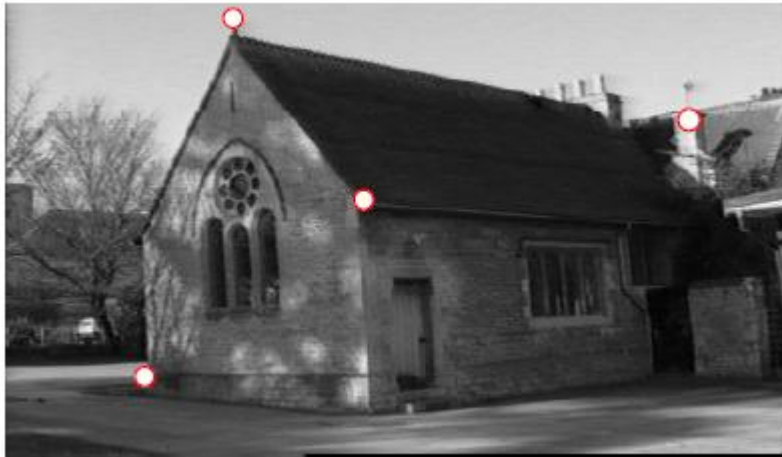
Παράδειγμα: Συγκλίνουσες κάμερες



Παράδειγμα: Παράλληλες κάμερες



Που είναι τα
επιπολικά
σημεία;

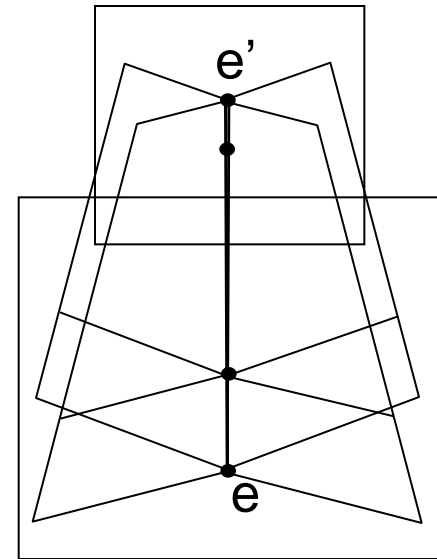
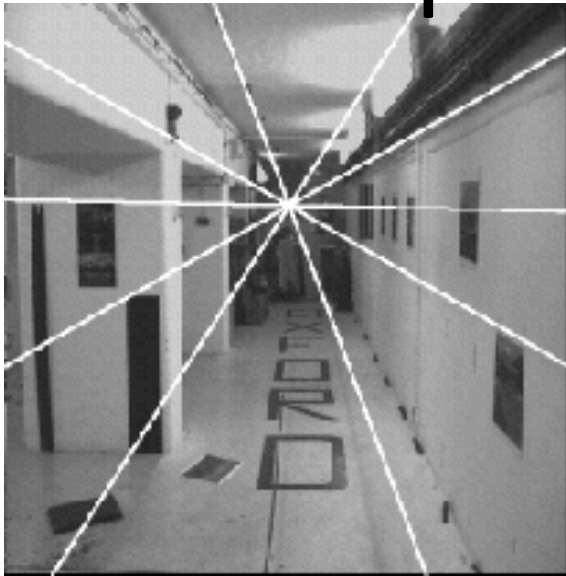


Παράδειγμα: Κίνηση

Πως θα ήταν οι επιπολικές ευθείες εάν η κάμερα κινούνταν προς τα μπροστά;



Παράδειγμα: Κίνηση

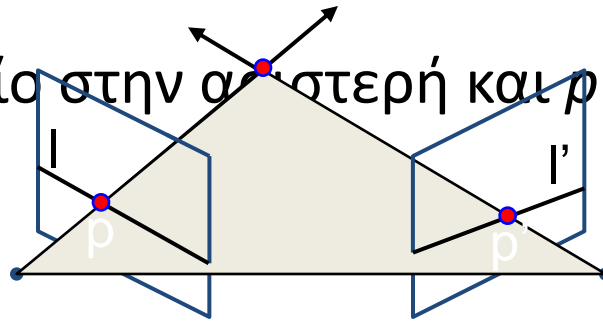


Τα επιπολικά σημεία έχουν ίδιες συντεταγμένες και στις δύο εικόνες. Τα σημεία κινούνται κατά μήκος ημιευθειών που ξεκινούν από το e



Θεμελιώδης πίνακας (Fundamental matrix)

- Έστω p ένα σημείο στην αριστερή και p' ένα στη δεξιά εικόνα



- Επιπολική σχέση
 - Το p προβάλεται στην επιπολική ευθεία l'
 - Το p' προβάλεται στην επιπολική ευθεία l
- Η επιπολική προβολή περιγράφεται από έναν πίνακα F 3×3

$$l' = Fp$$

- Προκύπτει $l \equiv p'F$

$$p'Fp = 0$$

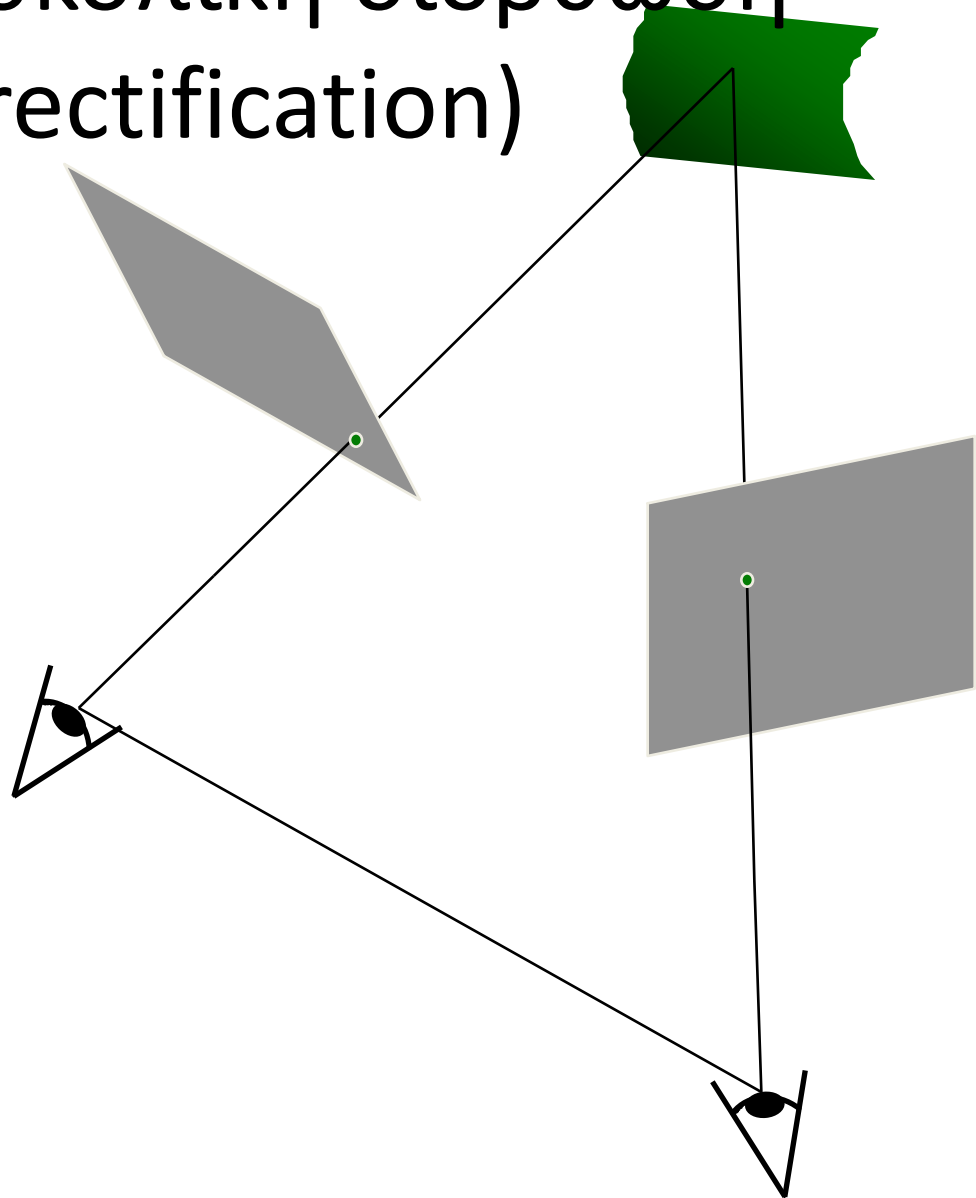


Θεμελιώδης πίνακας

- Ο πίνακας F λέγεται
 - “Essential Matrix”
 - όταν είναι γνωστές οι παράμετροι της κάμερας
 - “Fundamental Matrix”
 - γενικά (μη βαθμονομημένες κάμερες)
- Μπορούμε να λύσουμε ως προς F βάσει ζευγών σημείων
 - Κάθε ζεύγος (p, p') παρέχει μία γραμμική εξίσωση για τα στοιχεία του F
$$p' F p = 0$$
 - 8 σημεία είναι αρκετά για τον υπολογισμό του F (8-point algorithm)

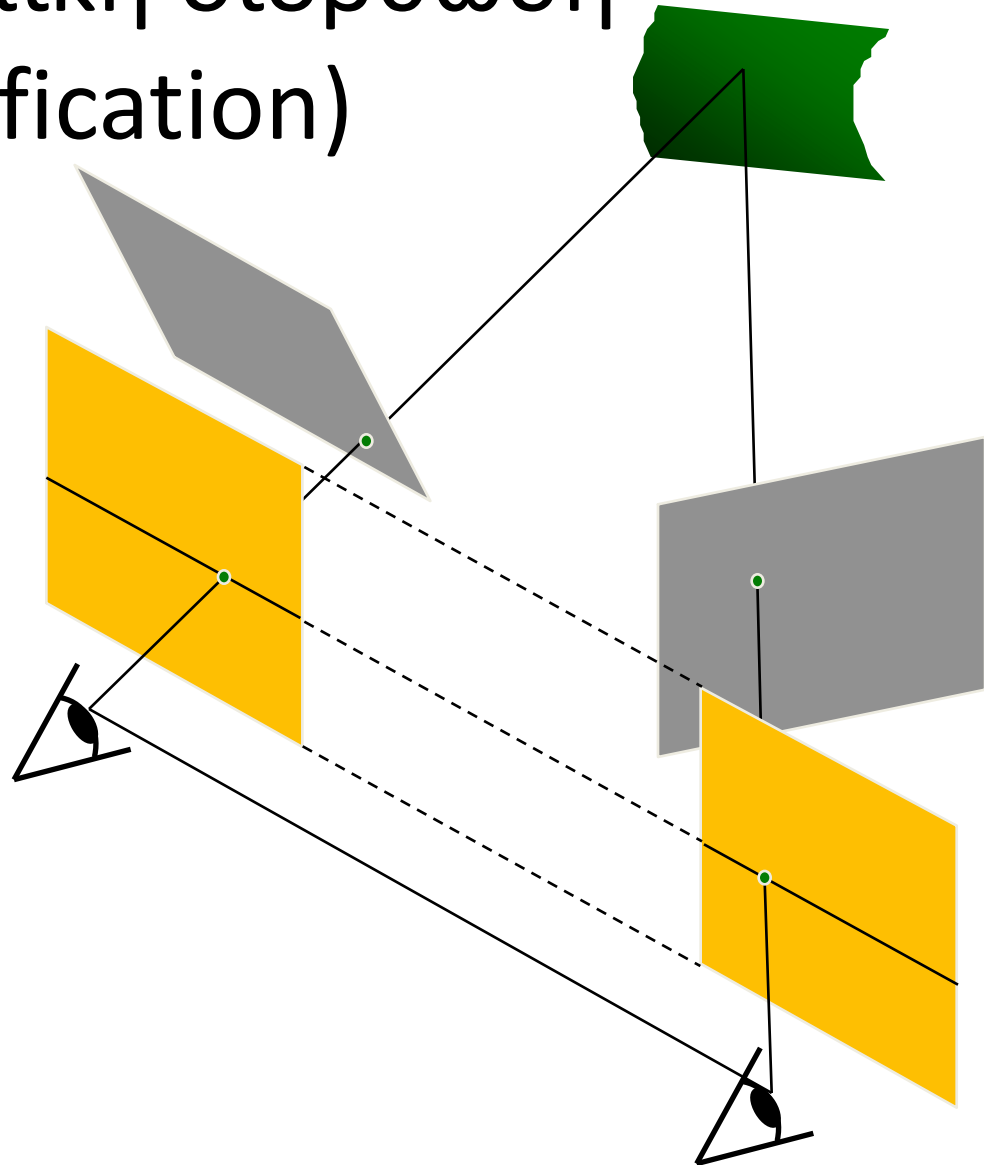


Στερεοσκοπική διόρθωση (rectification)



Στερεοσκοπική διόρθωση

- Επαναπροβολή των εικόνων σε ένα κοινό επίπεδο παράλληλο στην ευθεία μεταξύ των κέντρων προβολής (rectification)
- Η κίνηση των εικονοστοιχείων γίνεται οριζόντια μετά από αυτό το μετασχηματισμό
- Δύο ομογραφίες (3×3 μετασχηματισμός), μία για κάθε εικόνα



Παράδειγμα

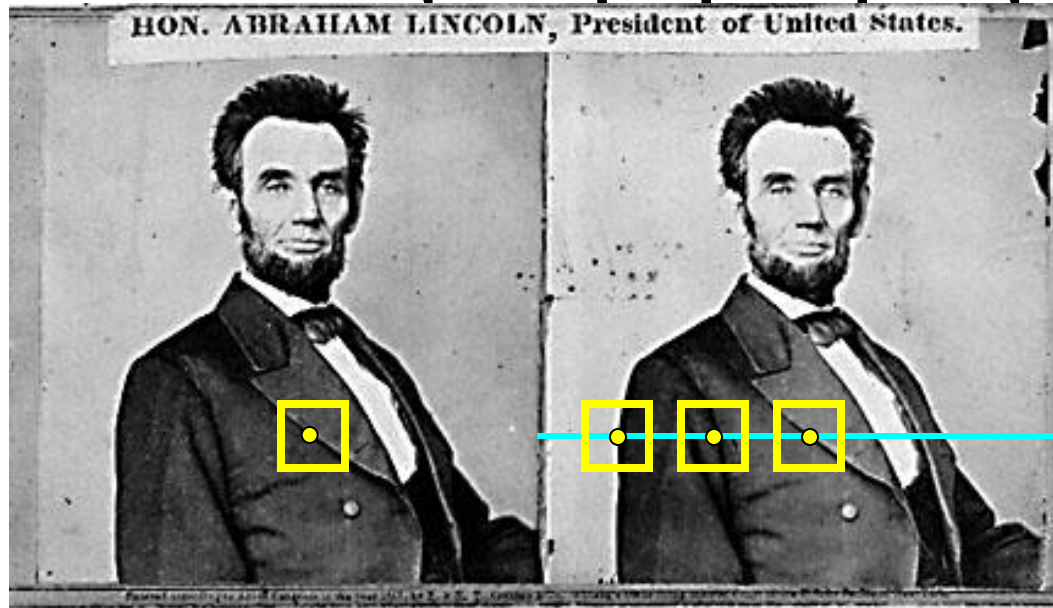


Το πρόβλημα αντιστοίχισης

- Η επιπολική γεωμετρία περιορίζει την αναζήτηση, αλλά ακόμα και έτσι το πρόβλημα παραμένει πολύπλοκο.



Βασικός αλγόριθμος



- Εάν είναι απαραίτητο διόρθωση των δύο εικόνων ώστε να μετασχηματιστούν οι επιπολικές ευθείες σε οριζόντιες
- Για κάθε εικονοστοιχείο x της πρώτης εικόνας
 - Εύρεση της αντίστοιχης οριζόντιας επιπολικής ευθείας στη δεξιά εικόνα
 - Εξέταση όλων των εικονοστοιχείων της ευθείας και επιλογή του βέλτιστου x'
 - Υπολογισμός ανομοιότητας $x-x'$ και ορισμός βάθους ως $\text{βάθος}(x) = fB/(x-x')$



Αναζήτηση αντιστοίχισης



- Συνέλιξη ενός παραθύρου κατά μήκος της δεξιάς επιπολικής ευθείας και σύγκριση του περιεχομένου του παραθύρου αυτού με το παράθυρο αναφοράς στην αριστερή εικόνα
- Κόστος: SSD ή κανονικοποιημένη συσχέτιση

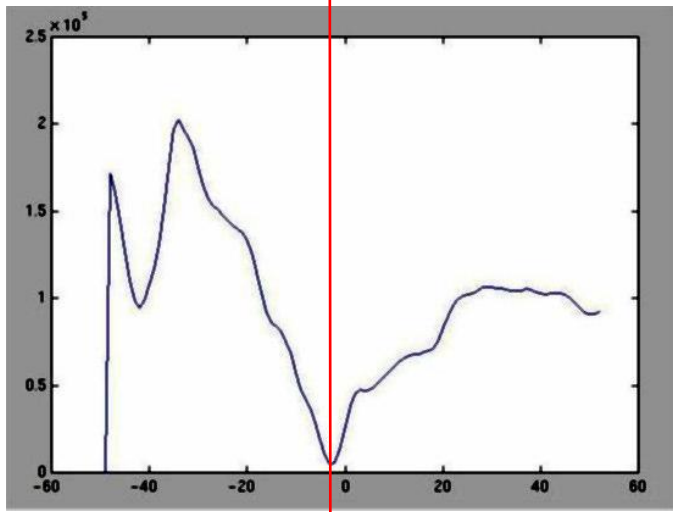


Αναζήτηση αντιστοίχισης

Αριστερή

Δεξιά

ευθεία



SSD

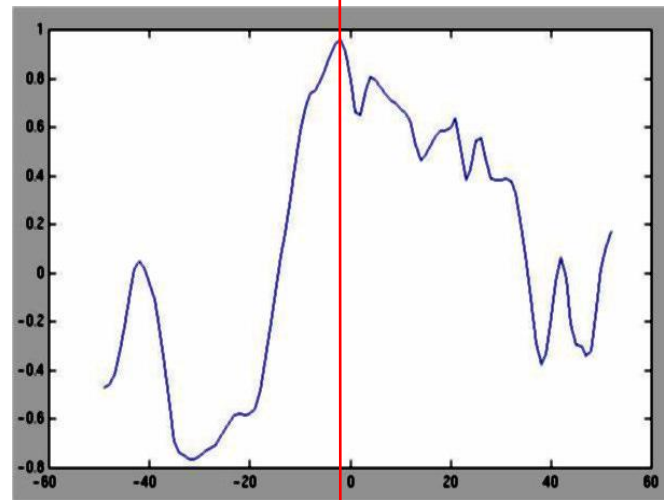


Αναζήτηση αντιστοίχισης

Αριστερή

Δεξιά

ευθεία



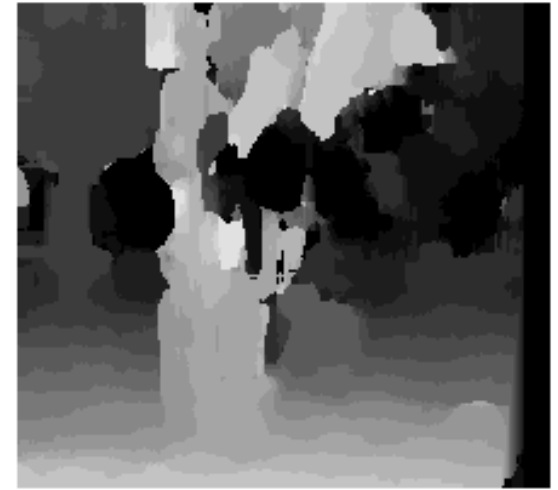
Norm. corr



Επιρροή του μεγέθους του



$W = 3$

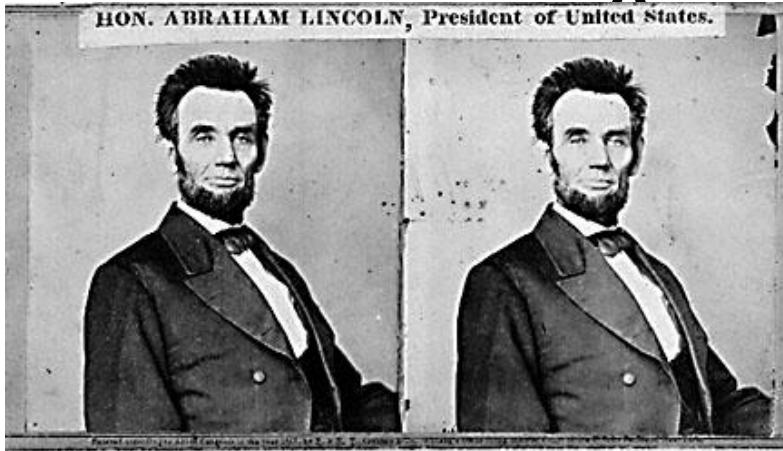


$W = 20$

- Μικρό παράθυρο
 - + Λεπτομέρεια
 - Θόρυβος
- Μεγάλο παράθυρο
 - + Ομαλοί χάρτες ανομοιότητας
 - Λεπτομέρεια



Αποτυχία αντιστοίχισης



Χαμηλόσυχνες περιοχές



Αποκλεισμός, επανάληψη



Μη-Lambertian επιφάνειες, αντανάκλασεις

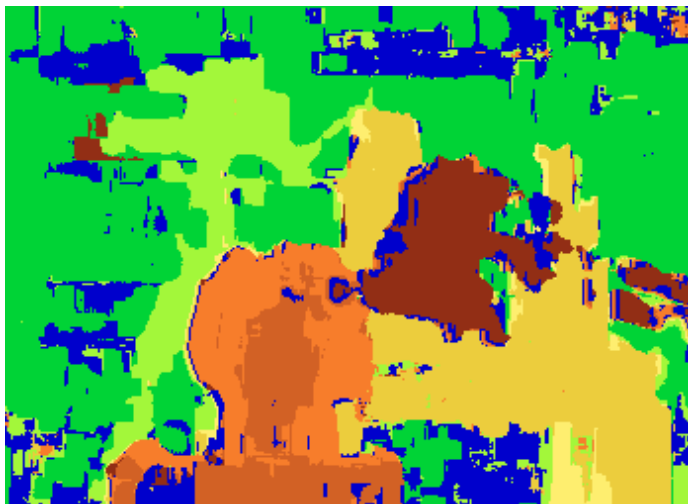


Αποτελέσματα αντιστοίχισης με παράθυρα

Δεδομένα



Αντιστοίχιση βάσει παραθύρου



Ground truth



Βελτίωση;

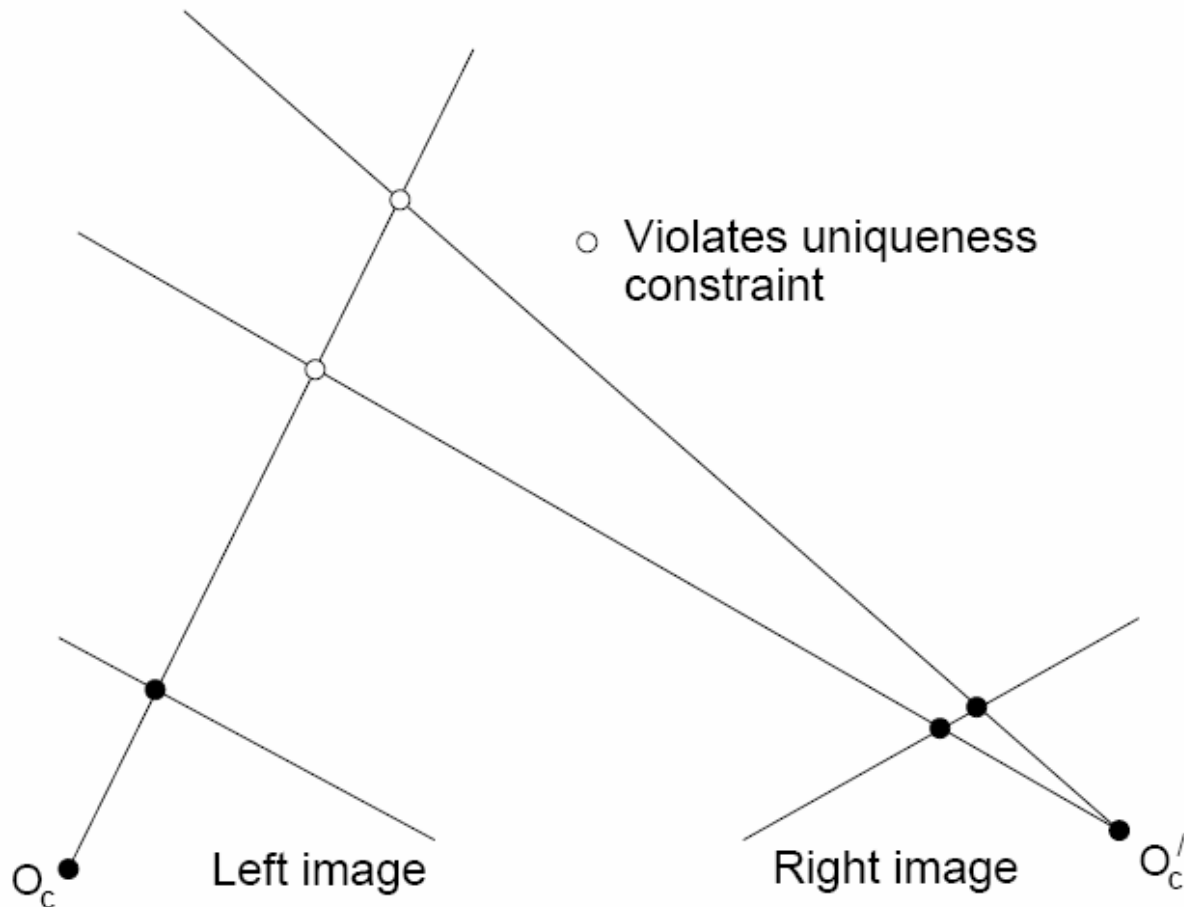
- Μέχρι στιγμής: οι αντιστοιχίσεις είναι ανεξάρτητες για κάθε σημείο
- Τι περιορισμούς ή γνώση (priors) μπορούμε να προσθέσουμε;



Στέρreo περιορισμοί/priors

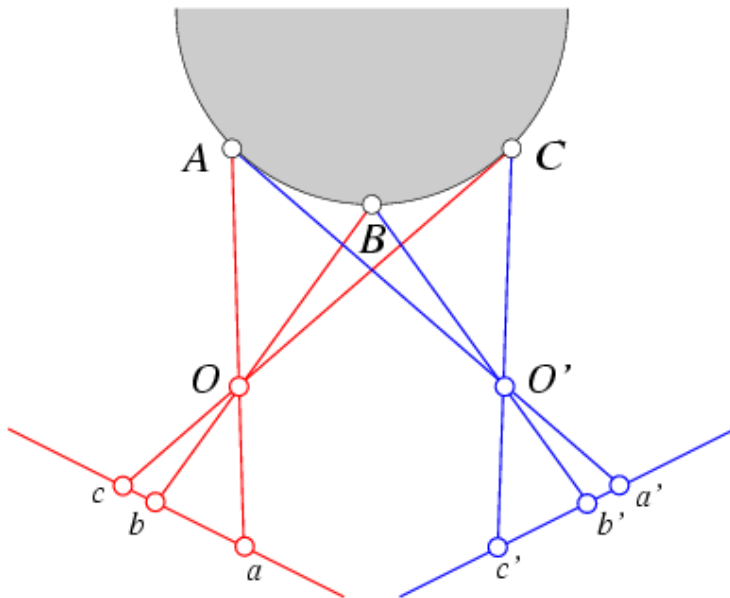
- Μοναδικότητα

- Για κάθε σημείο σε μία εικόνα, πρέπει να υπάρχει το πολύ ένα σημείο αντιστοίχισης στην άλλη εικόνα



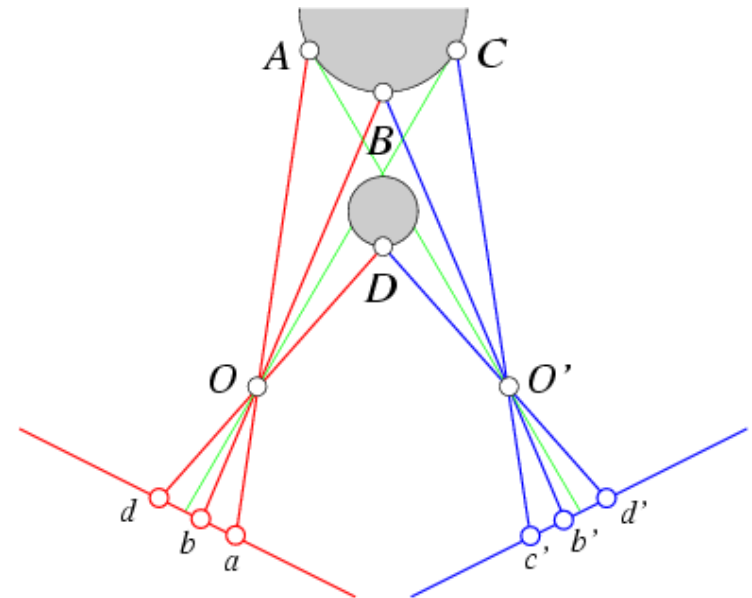
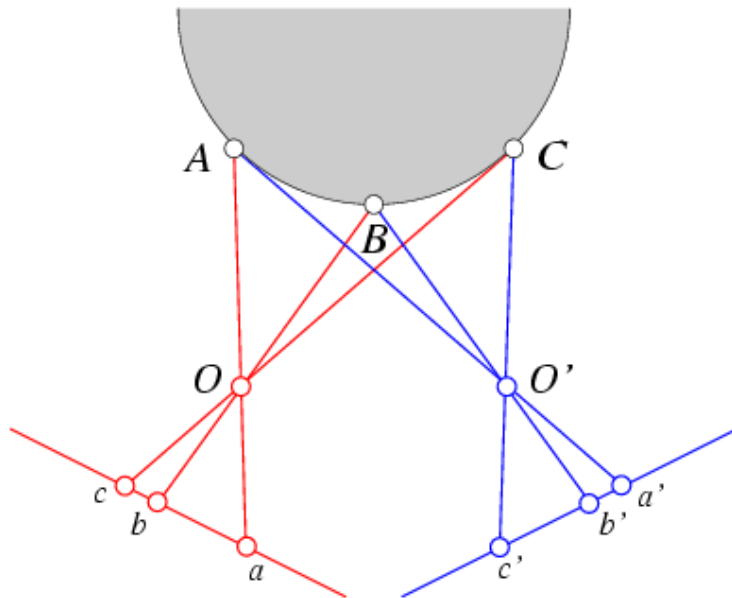
Στέρεο περιορισμοί/priors

- Μοναδικότητα
 - Για κάθε σημείο σε μία εικόνα, πρέπει να υπάρχει το πολύ ένα σημείο αντιστοίχισης στην άλλη εικόνα
- Διάταξη
 - Η διάταξη των αντίστοιχων σημείων πρέπει να είναι ίδια και στις δύο όψεις



Στέρχο περιορισμοί/priors

- Μοναδικότητα
 - Για κάθε σημείο σε μία εικόνα, πρέπει να υπάρχει το πολύ ένα σημείο αντιστοίχισης στην άλλη εικόνα
- Διάταξη
 - Η διάταξη των αντίστοιχων σημείων πρέπει να είναι ίδια και στις δύο όψεις



Η διάταξη δεν ισχύει πάντα



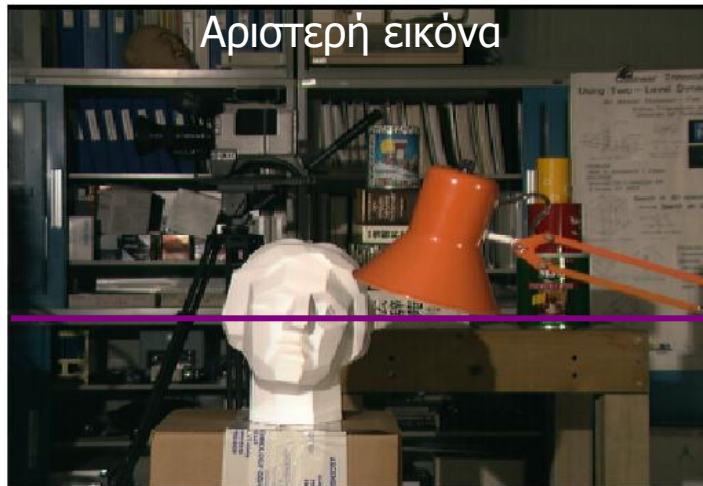
Στέρreo περιορισμοί/priors

- Μοναδικότητα
 - Για κάθε σημείο σε μία εικόνα, πρέπει να υπάρχει το πολύ ένα σημείο αντιστοίχισης στην άλλη εικόνα
- Διάταξη
 - Η διάταξη των αντίστοιχων σημείων πρέπει να είναι ίδια και στις δύο όψεις
- Ομαλότητα
 - Αναμένουμε οι τιμές ανομοιότητας να μεταβάλλονται ομαλά (στο μεγαλύτερο μέρος)



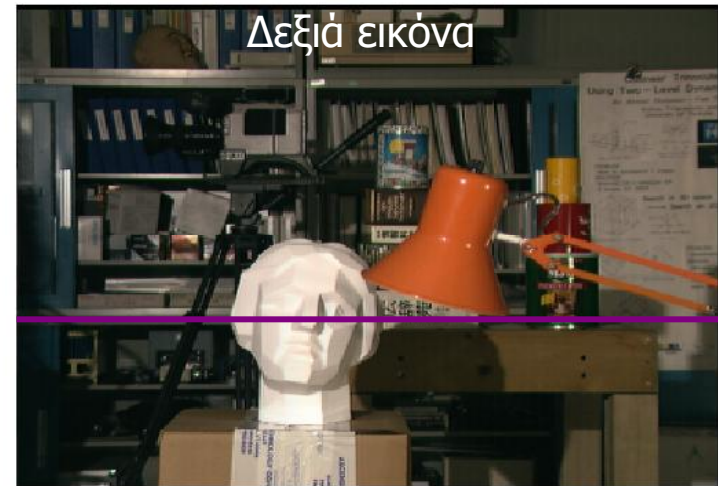
Scanline stereo

- Προσπαθούμε να αντιστοιχίσουμε



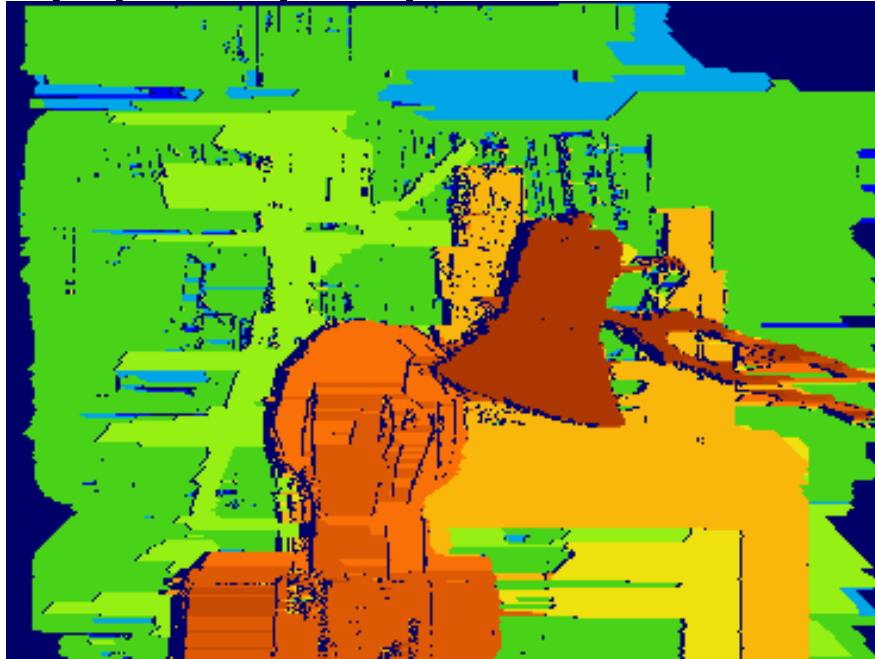
'C

CI

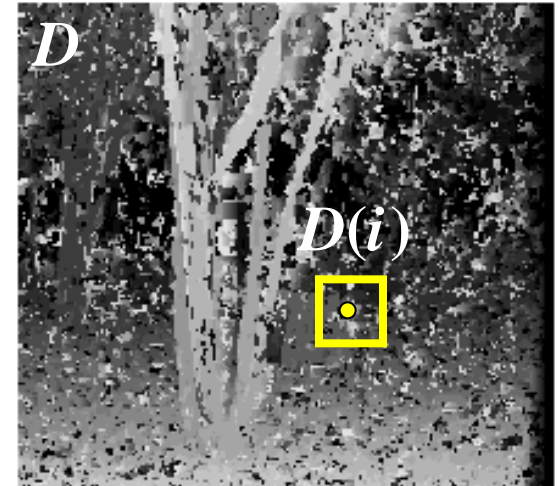
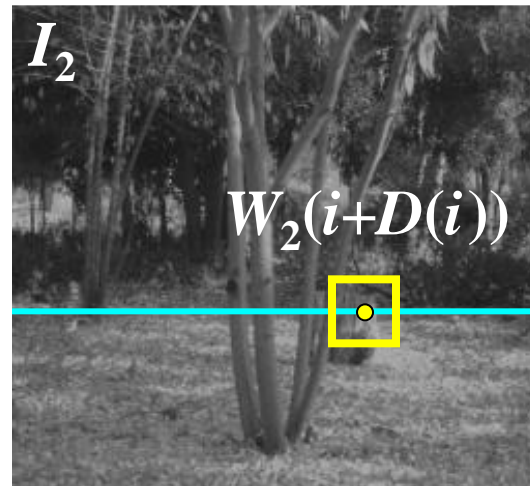
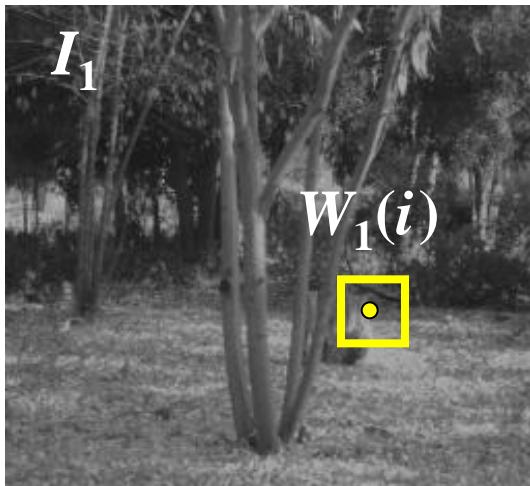


Συνεκτικό stereo σε ένα 2D πλέγμα

- Η λογική του Scanline stereo δημιουργεί ραβδόμορφο σφάλμα



Αντιστοίχιση βάσει ενεργειακής ελαχιστοποίησης

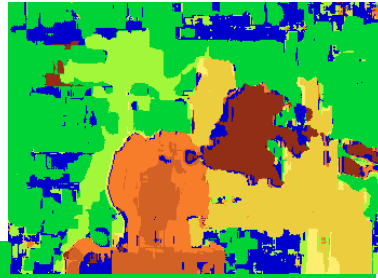


$$E(D) = \underbrace{\sum_i (W_1(i) - W_2(i + D(i)))^2}_{\text{δεδομένα}} + \lambda \underbrace{\sum_{\text{neighbors } i, j} \rho(D(i) - D(j))}_{\text{όρος ομαλότητας}}$$



Πολλοί περιορισμοί μπορούν να εκφραστούν ενεργειακά

Πριν



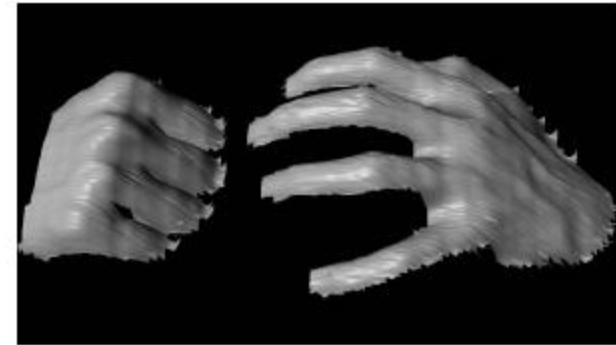
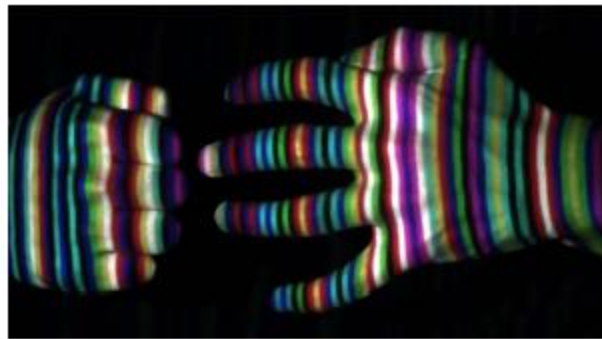
Μετά



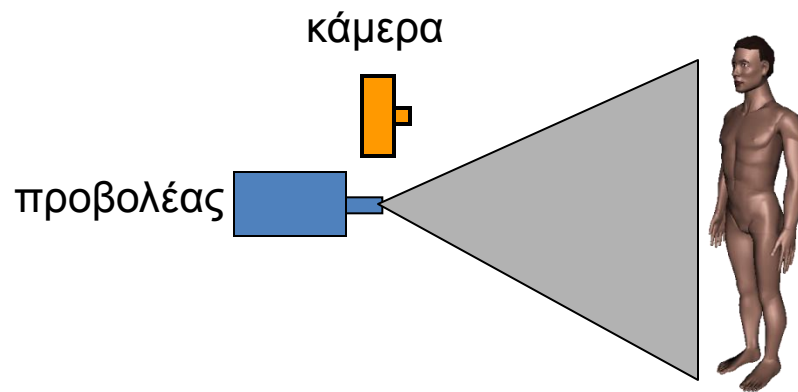
Ground truth



Ενεργό στέρεο με δομημένο φως



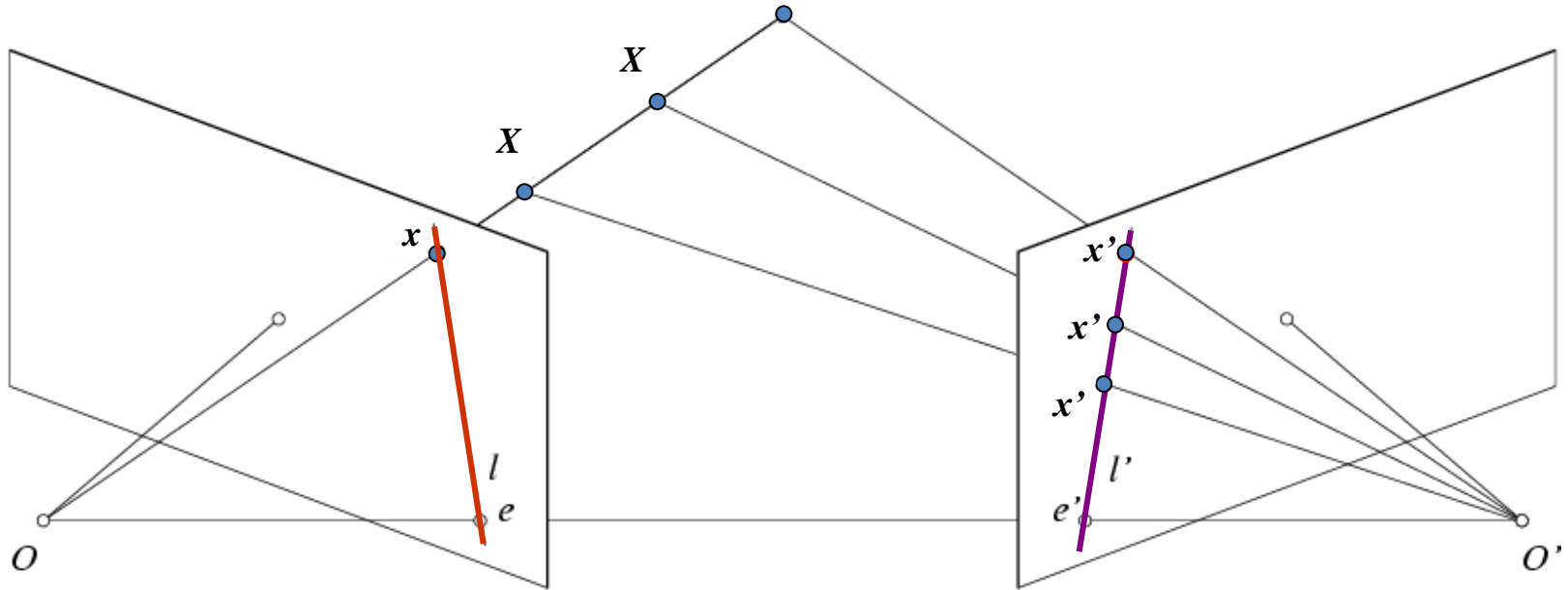
- Προβολή/εκπομπή δομημένου φωτός στο αντικείμενο
 - Απλοποιεί το πρόβλημα της αντιστοίχισης
 - Επιτρέπει να χρησιμοποιήσουμε μία μόνο κάμερα



Kinect: Δομημένο υπέρυθρο φως



Σύνοψη: Ιδέα: Επιπλῶλικός περιορισμός



Το αντίστοιχο σημείο του x πρέπει να βρίσκεται στην αντίστοιχη ευθεία l'

Το αντίστοιχο σημείο του x' πρέπει να βρίσκεται στην αντίστοιχη ευθεία l



Σύνοψη

- Επιπολική γεωμετρία
 - Τα επιπολικά σημεία (eripoles) είναι η τομή της baseline με τα επίπεδα της εικόνας
 - Το αντίστοιχο σημείο στη δεύτερη εικόνα βρίσκεται επί της ευθείας που περνάει από το επιπολικό σημείο
 - Ο θεμελιώδης πίνακας αντιστοιχίζει ένα σημείο σε μία εικόνα σε μία ευθεία σε μία άλλη (την επιπολική ευθεία)
 - Μπορούμε να λύσουμε ως προς F δοσμένων ζευγών σημείων (π.χ., χαρακτηριστικά σημεία)
- Εκτίμηση βάθους
 - Εκτίμηση ανομοιότητας βρίσκοντας τα αντίστοιχα σημεία κατά μήκος των οριζόντιων ευθειών
 - Το βάθος είναι αντιστρόφως ανάλογο της ανομοιότητας



Activity recognition

Activity recognition

Aims to recognize the actions and goals of one or more agents from a series of observations on the agents' actions and the environmental conditions.



Activity recognition

Activity recognition approaches:

In real life activity recognition systems typically follow a hierarchical approach.

- Lower levels: Background-foreground segmentation, tracking and object detection
- Mid-level: Action recognition modules
- High-level: Reasoning engines, which encode the activity semantics based on the lower level action-primitives.



Activity recognition

Two main approaches:

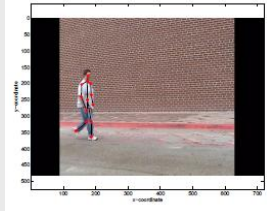
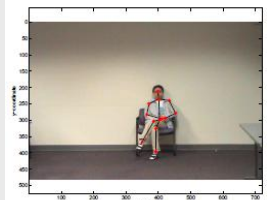
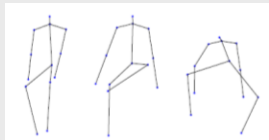
- methods based on various sensors placed on the subject to extract meaningful features
- methods based on video analysis to detect human activity (*one of the most promising applications of computer vision*)



Activity recognition

Approaches of Video based action recognition

Model based



(From Sheikh et al, 2005)

g. Joint angles

- Joint positions

Direct mapping of image cues

Global/holistic



(From Bobick & Davis, 2001)

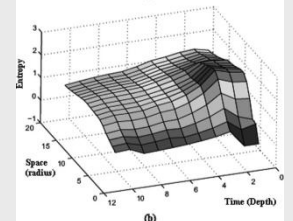
e.g.

- Silhouettes
- Edge images
- Optical flow

Patch-based



(a)



(b)

e.g.

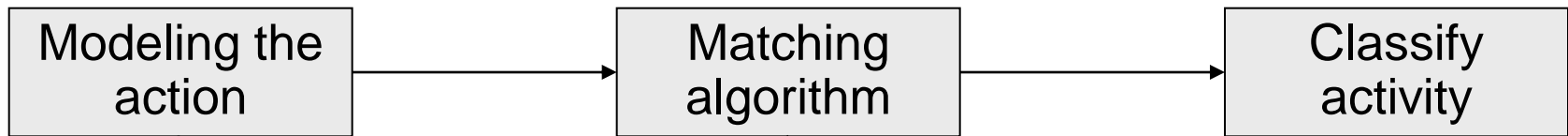
- Spatio-temporal silent points

(Oikonomou et al., 2006)

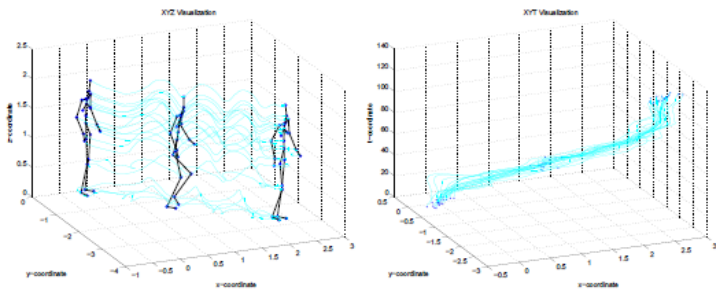


Model-based approach

Activities: (a) walking, (b) sitting, (c) standing, (d) running



Representation of an action in 4-space



Action in XYZ space

Action in XYT space

Matching algorithm

For each action A_i , $i \in 1, 2, \dots, N$ do

1. Normalisation. Compute a similarity transform, transforming the mean of the point to the origin and making the average distance of the points from the origin equal to $\sqrt{2}$ (Separately for each action instance)
2. Compute Subspace Angle between W and W_t :
 - Compute Orthogonal Bases: Uses SVD to reliably compute orthonormal bases of W' and W_t , \tilde{W}' and \tilde{W}_t
 - Compute Projection: Using the iterative procedure described in Bjork and Golup (1973) for $j \in 1, \dots, p$

$$W'_{i+1} = W'_i - WW^T W'_i$$

- Find Angle: Compute $\theta = \arcsin \min(1, \|W'_p\|_2)$

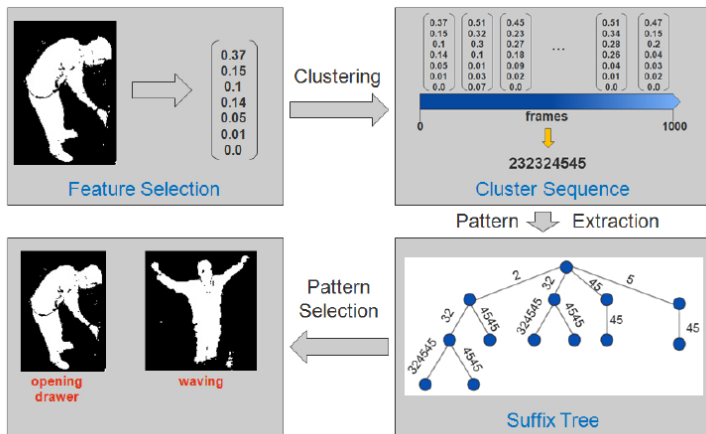
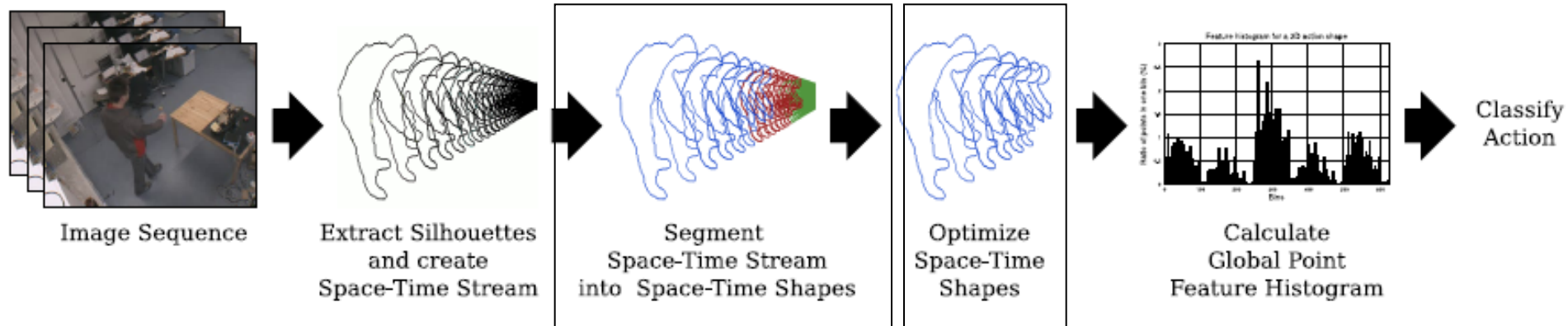
Select $i^* = \arg \max_{i \in \{1, \dots, N\}} \cos(\theta_1)$.

Where W' corresponds to the projection of an action instance, and matrices W_1, W_2, \dots, W_N each modeling the N different actions.

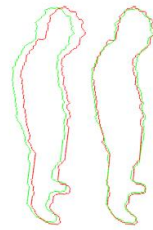
(Sheikh et al., 2005)



Global/holistic approach: Activity recognition in office



Alignment



Minimisation



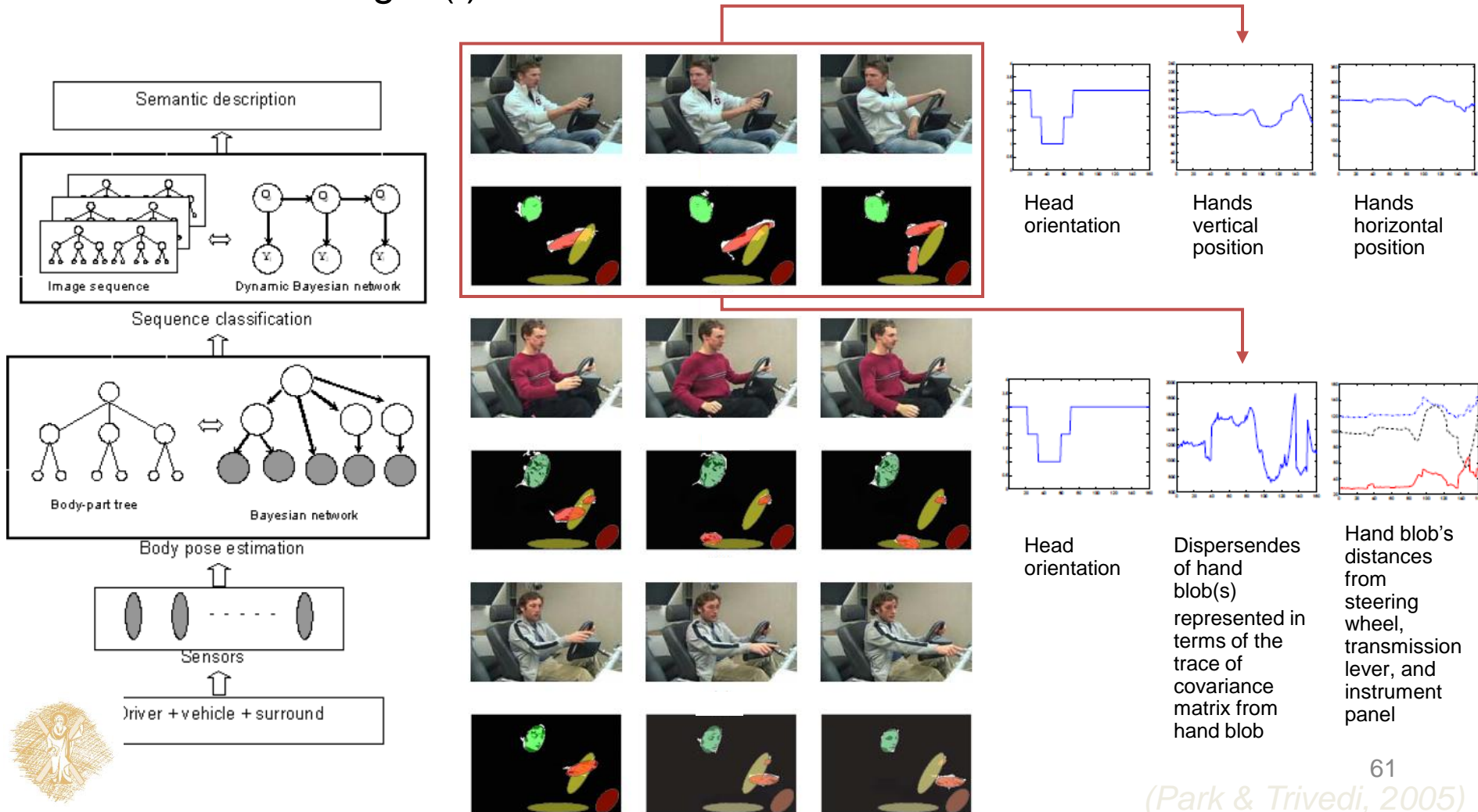
1. For each point $p_i \in P$ belonging to frame f_t , search for its closest corresponding point $q_i \in Q$ in frame f_{t+1} ;
2. Compute the Frobenious norm in distance space of f_t and f_{t+1} as:

$$\|f_t, f_{t+1}\|_F = \sqrt{\sum |q_i - p_i|^2}$$

3. Select f_{t+1} as being different (i.e. unique) than f_t , if $\|f_t, f_{t+1}\|_F \leq d_{thresh}$

Activity recognition in a vehicle

Activities: (a) drive-forward, (b) backup, (c) shift-gear, (d) turn-left, (e) turn-right (f) touch-radio



Motion Images

- **Motion Energy Image (MEI)**

- Describes the motion energy for a given view of action.
- Binary image.

$$E_T(x, y, t) = \bigcup_{i=0}^{\tau-1} D(x, y, t - i)$$

- **Motion History Image (MHI)**

- Intensity image.

$$H_T(x, y, t) = \begin{cases} \tau & , \text{if } D(x, y, t) = 1 \\ \max(0, H(x, y, t - 1) - 1) & , \text{otherwise} \end{cases}$$



Motion Images – Activity Recognition



Motion Energy
Image



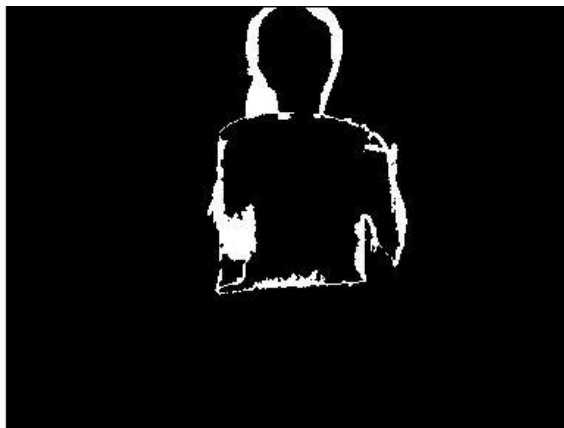
Motion History
Image



MEI Templates of Six Micro-Activities



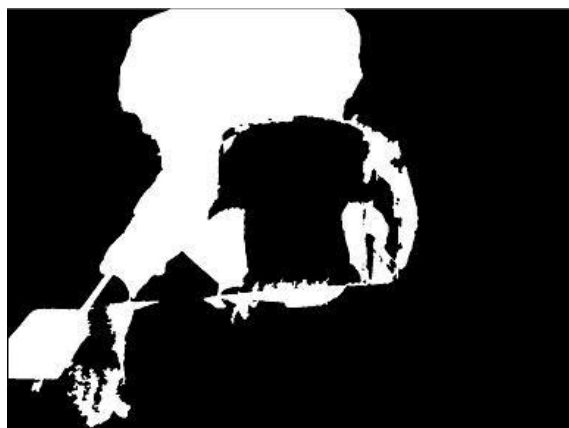
Pick up or put back



Phone Conversation



Activate Alarm



Talking to the microphone



Raising Hands



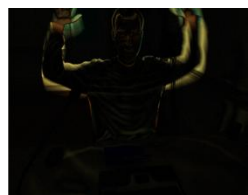
Drink from Glass



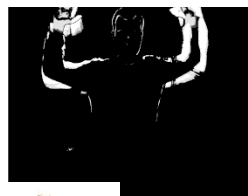
Flow Chart



Grab frames from office activities.



Get difference from sequential frames.



Convert difference images to grayscale.



Extraction of the Motion Energy Image.

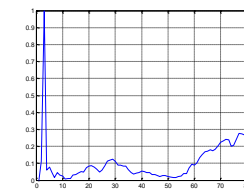
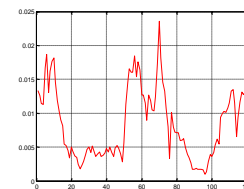
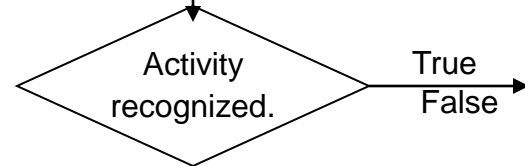
Extraction of the Motion History Image.



Prototype MHI/MEI

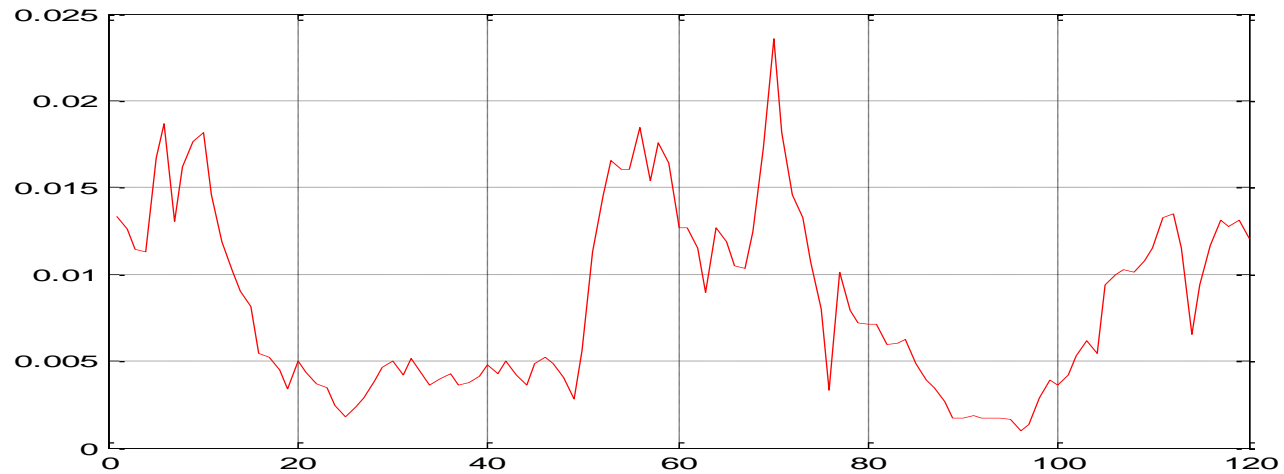
RIT Extraction + CIT Extraction

Comparison of euclidian distance & Correlation Factor



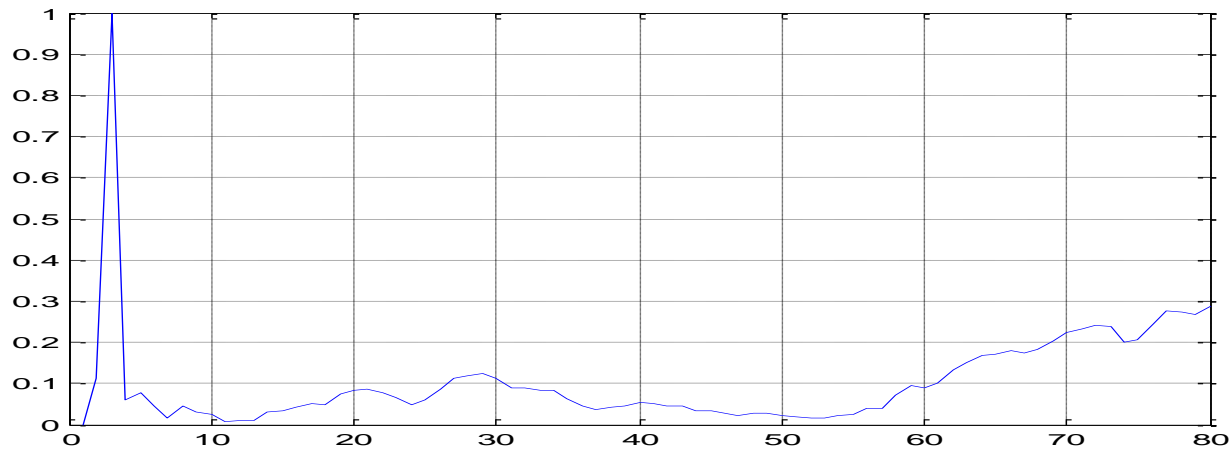
RIT Extraction

1. Face detection
2. Centre of Face
3. RIT Extraction



CIT Extraction

1. Face detection
2. Centre of Face
3. CIT Extraction



Metrics

- Euclidian Distance

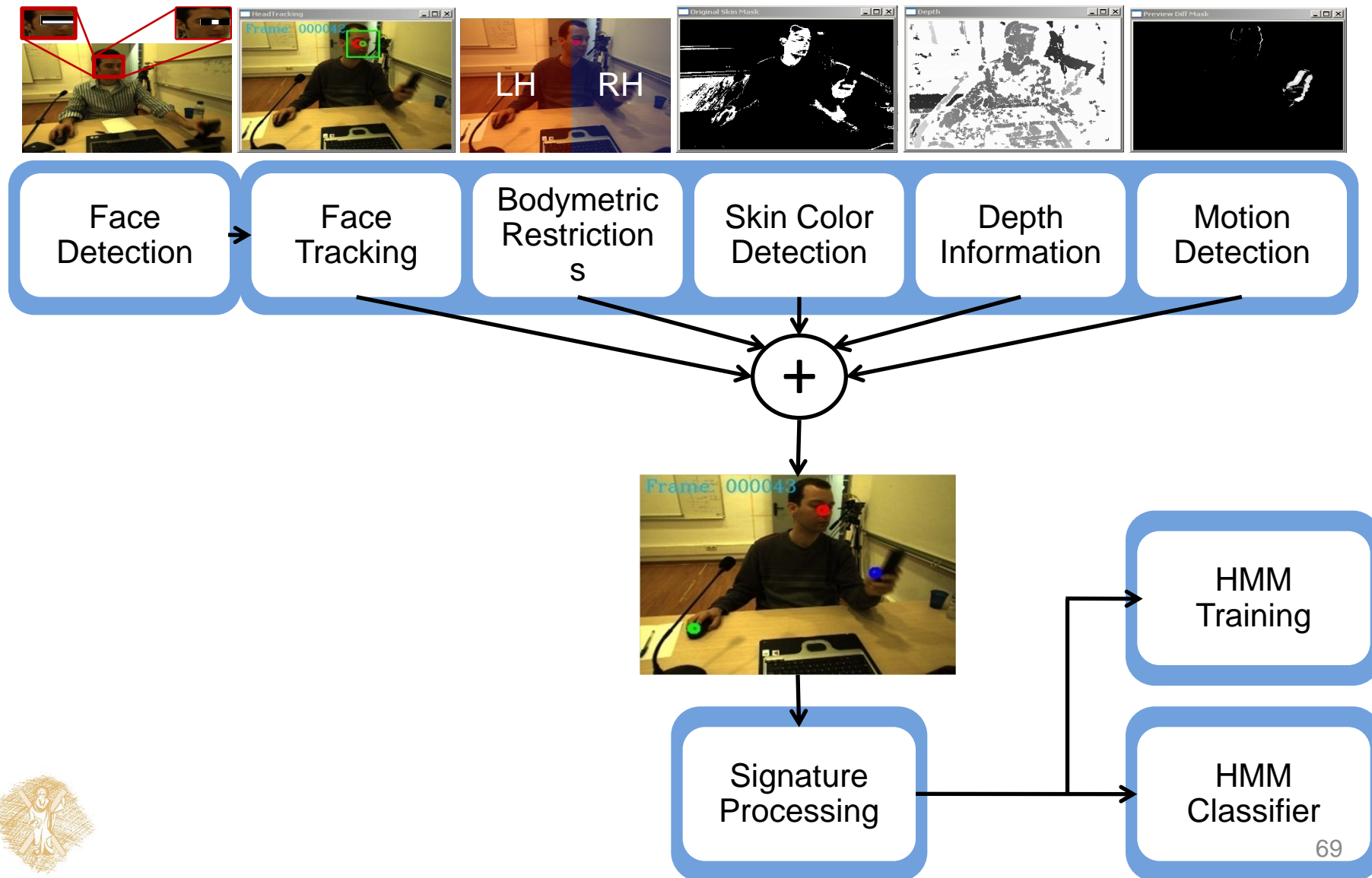
$$D_E = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$\text{corr}(x, y) = \rho_{x,y} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{E((x - \mu_x)(y - \mu_y))}{\sigma_x \sigma_y}$$

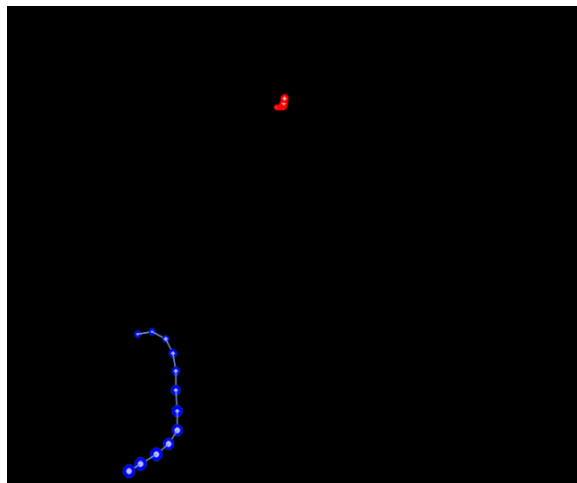
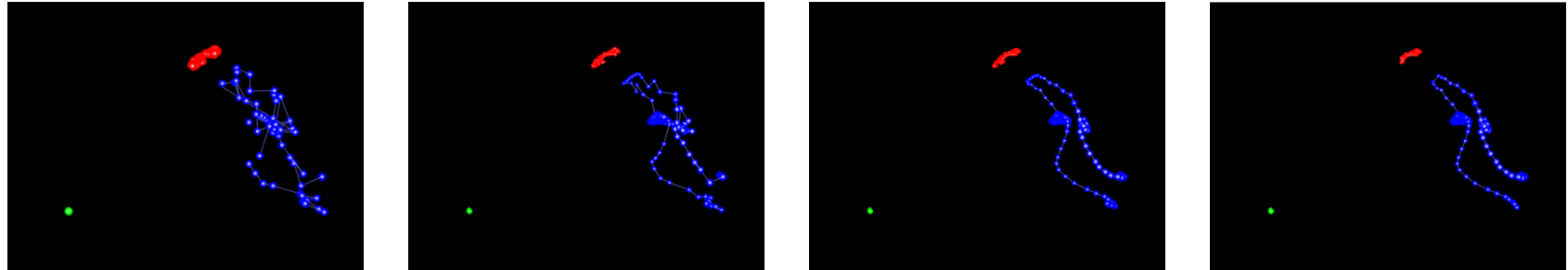
- Correlation Factor - Coefficients



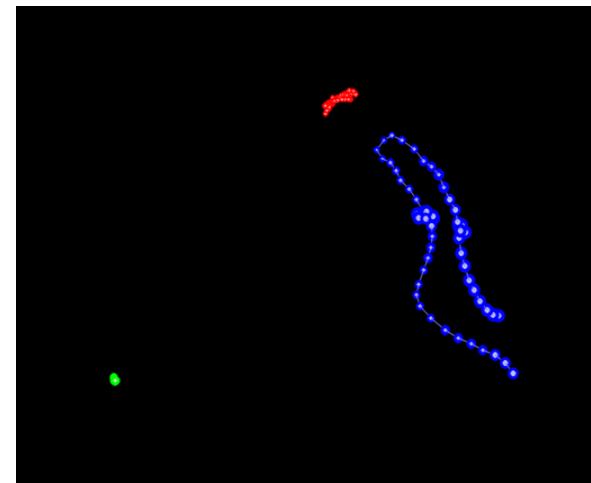
System Overview



Signature Extraction



Activity: Office Panel

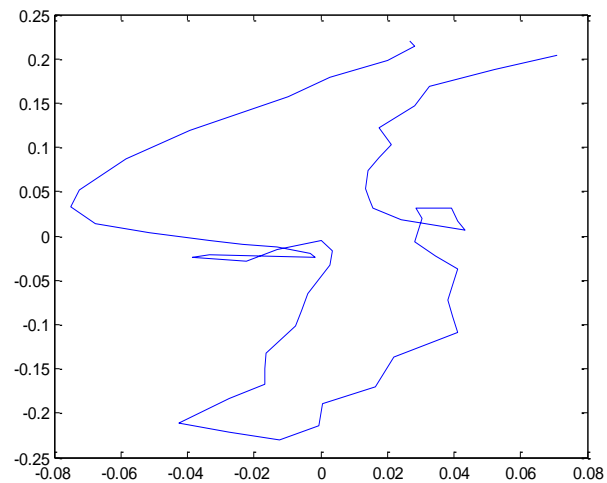
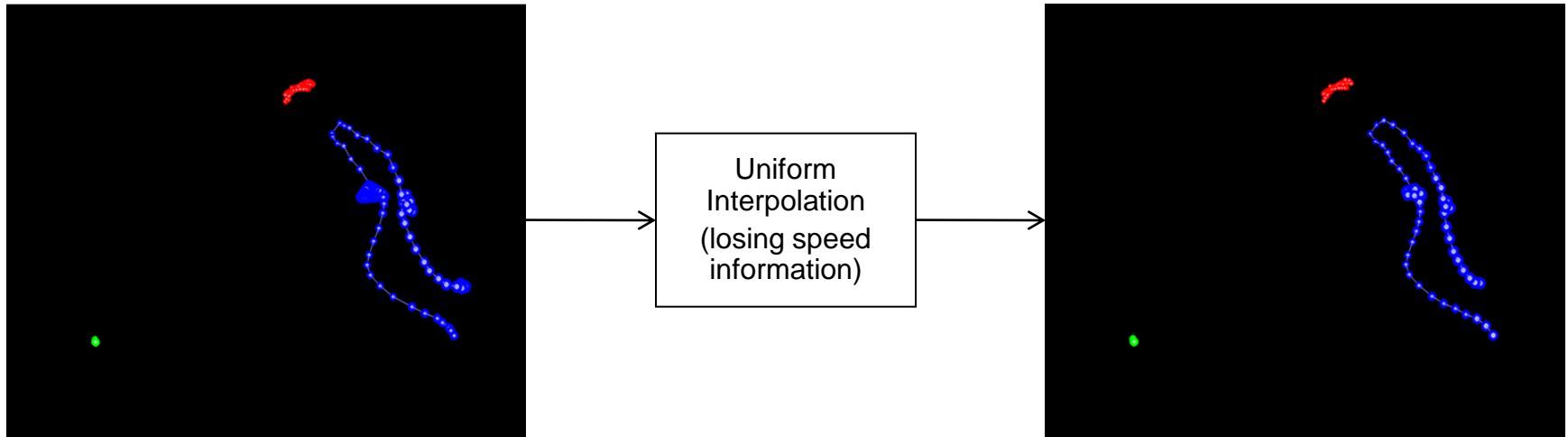


Activity: Phone Conversation



Trajectory's Features for authentication

Trajectory Features Extraction (curvature, torsion)

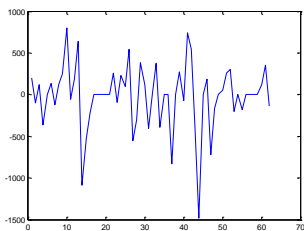


PCA

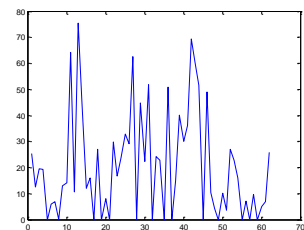


Trajectory's Features for authentication

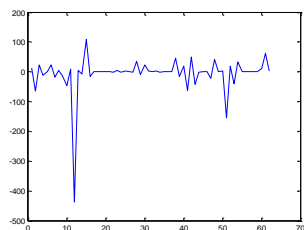
Trajectory Features Extraction (curvature, torsion)



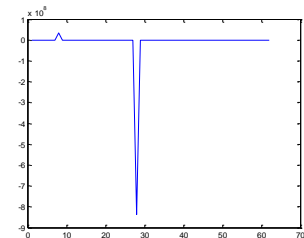
$$\kappa^*(P_i) = 4 \frac{\sqrt{\hat{s}(s-a)(s-b)(s-c)}}{abc}$$



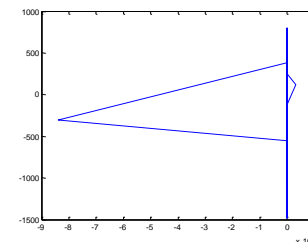
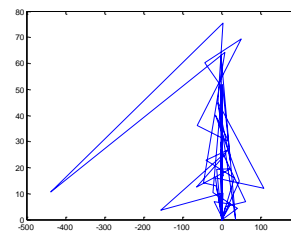
$$\kappa_s^*(P_i) = 3 \frac{\kappa^*(P_{i+1}) - \kappa^*(P_{i-1})}{2a + 2b + d + g}$$



$$\tau^*(P_i) = \frac{1}{2} \left(6 \frac{H^+}{def \cdot \kappa^*(P_i)} + 6 \frac{H^-}{gmm \cdot \kappa^*(P_i)} \right)$$



$$\tau_s^*(P_i) = 4 \frac{\tau^*(P_{i+1}) - \tau^*(P_{i-1}) + r(\tau^*(P_i)\kappa_s^*(P_i)/6\kappa^*(P_i))}{2a + 2b + 2d + h + g}$$



Motion symmetry perception:

a) Curvature sub-signature

b) Torsion sub-signature

,whereas

$$\frac{1}{3!} \begin{vmatrix} x_i & y_i & z_i & 1 \\ x_{i-1} & y_{i-1} & z_{i-1} & 1 \\ x_{i+1} & y_{i+1} & z_{i+1} & 1 \\ x_{i+2} & y_{i+2} & z_{i+2} & 1 \end{vmatrix} = V_{abcdef} = \frac{\Delta_{abc} \cdot H^+}{3}$$

$$r = 2a + 2b - 2d - 3h + g$$

$$s = (a + b + c) / 2$$



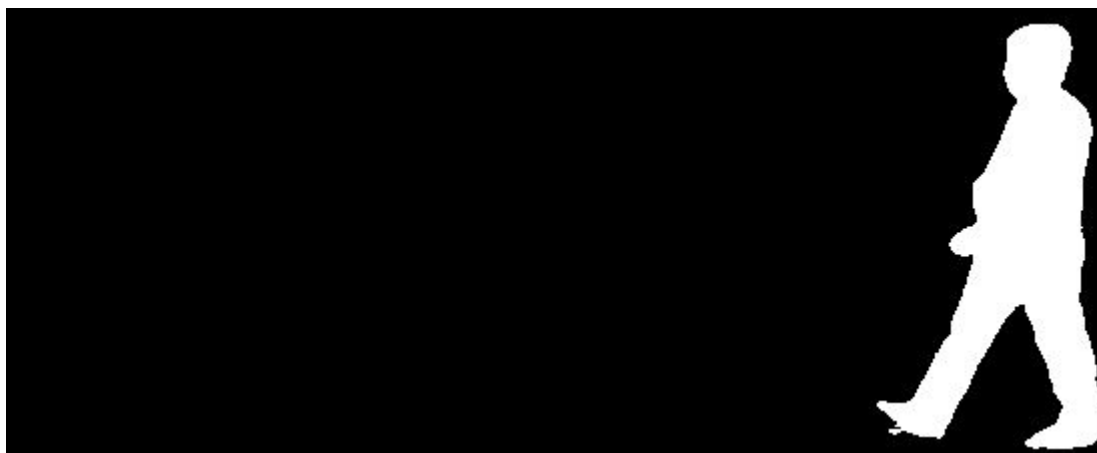
Gait...

- ❖ is the personal, idiosyncratic way in which people walk or move on foot
- ❖ has received significant attention as a biometric, in the last 10-15 years

(Nixon & Carter, 2006; Rahati et al., 2008; Boulgouris et al., 2005)

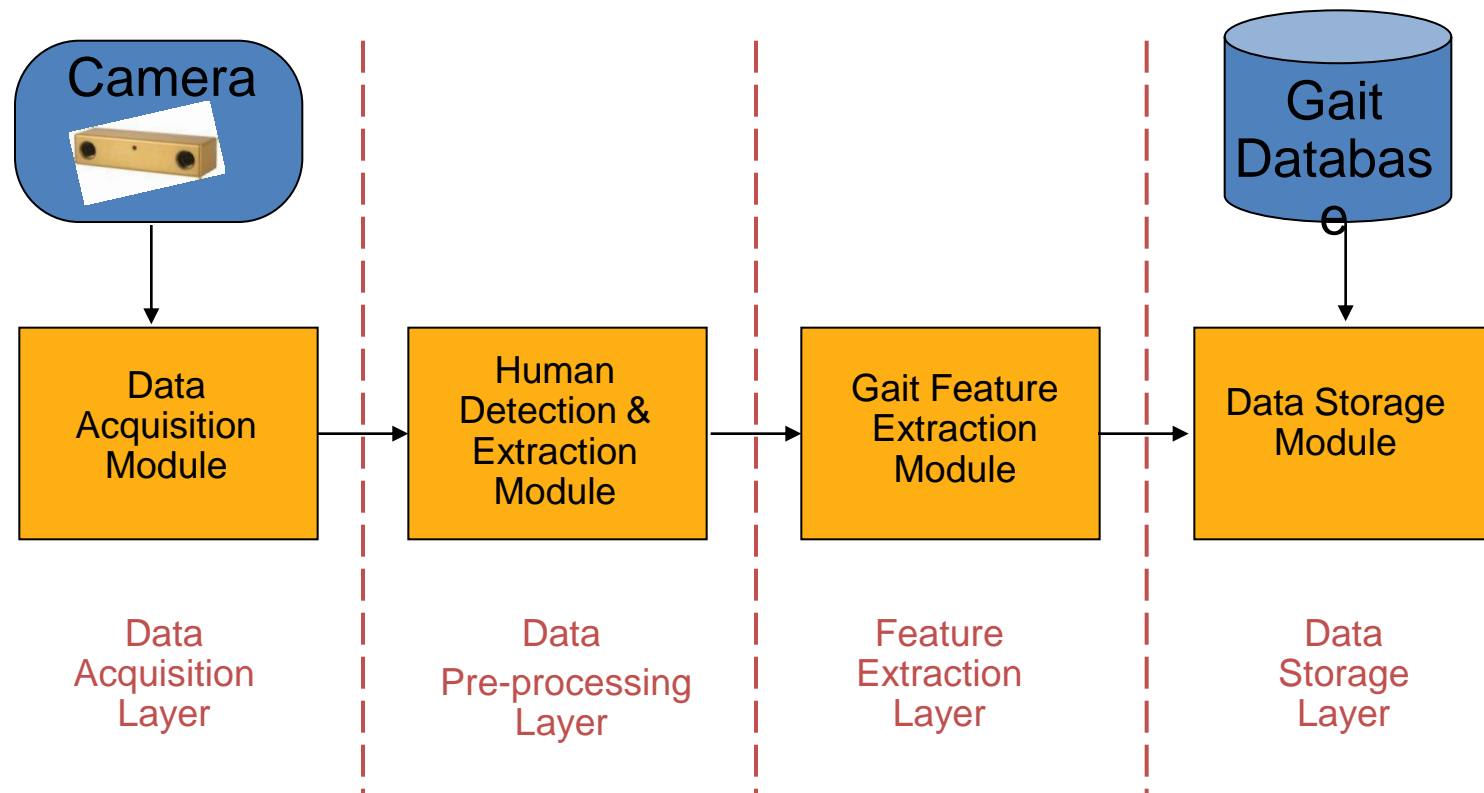
- ❖ has demonstrated high recognition rates

(Zhang et al., 2007; Bouchrika & Nixon, 2007)



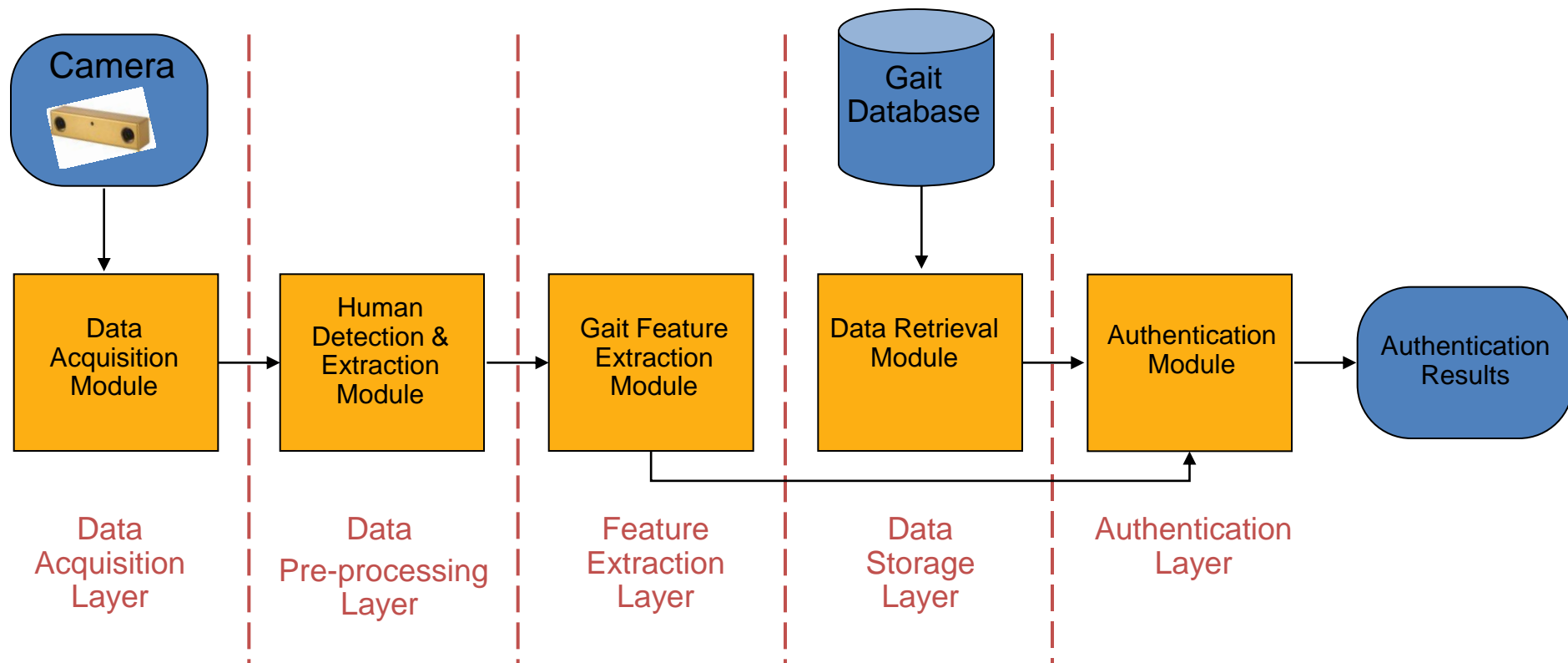
General Gait Authentication System

– Enrolment procedure



General Gait Authentication System

– Authentication procedure



Approaches to the gait recognition

Featured based (model free):

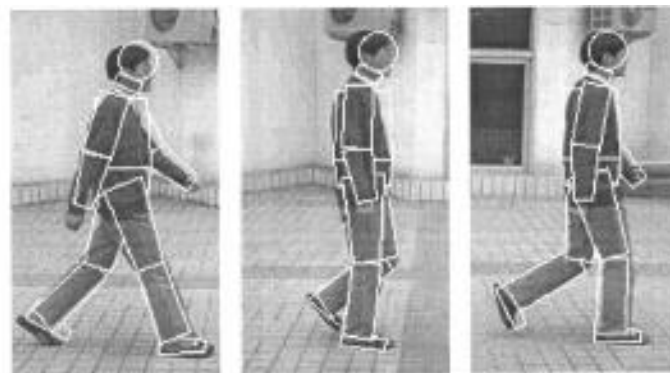
Focus on the spatiotemporal information contained in the silhouette images (low level measurements).



(From Li et al. 2004)

Model based:

Serve as prior knowledge to predict motion parameters, to interpret human dynamics, or to constrain the estimation of low-level image measurements.



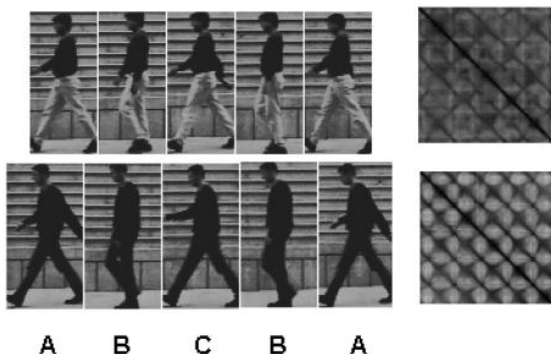
(From Wang et al. 2004)



Features for gait recognition

Featured-based method (1/2)

- Width of the outer contour of the silhouette (*Kale et al. 2004*)
- Entire binary silhouette (*Kale et al. 2004*)
- Boundary vector variations from the centre of silhouette (*Wang et al. 2003*)
- Statistical moments from body parts ellipses (*Lee et al. 2002*)
- Height, stride length and cadence (*BenAbdaker et al. 2002*)
- Height, distance between head and pelvis, max distance between pelvis and feet, and the distance between feet (*Johnson & Bobick 2001*)
- Self similarity plot (*BenAbdelkader et al., 2001*)



$$S(t_1, t_2) = \sum_{(x,y) \in B_{t_1}} |O_{t_1}(x,y) - O_{t_2}(x,y)|,$$

Where $1 \leq t_1, t_2 \leq N$, B_t is the bounding box of the person in frame t_1 , and $O_{t_1}, O_{t_2}, \dots, O_{t_N}$ are the scaled image templates of the person.



Features for gait recognition

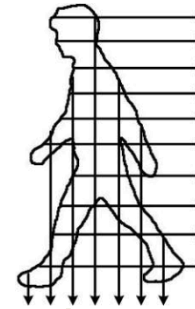
Featured-based method (2/2)

- Horizontal and vertical projections

$$P_h(x) = \sum_{y=1}^{N_c} S(x, y), \quad x = 1, \dots, N_r$$

$$P_v(x) = \sum_{r=1}^{N_r} S(x, y), \quad y = 1, \dots, N_c$$

$$S(x, y) = \begin{cases} 1, & \text{if } (x, y) \text{ is a foreground pixel} \\ 0, & \text{otherwise} \end{cases}$$



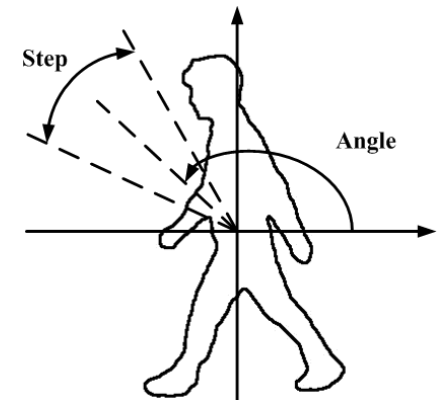
(Liu et al. 2002)

- Angular transform

$$A(\theta) = \frac{1}{N_\theta} \sum_{(x,y) \in \Phi_\theta} S(x, y) \sqrt{(x-x_c)^2 + (y-y_c)^2}$$

where θ is an angle, Φ_θ is the set of the pixels in the circular sector $\left(\theta - \frac{\Delta\theta}{2}, \theta + \frac{\Delta\theta}{2}\right)$ and N_θ is the

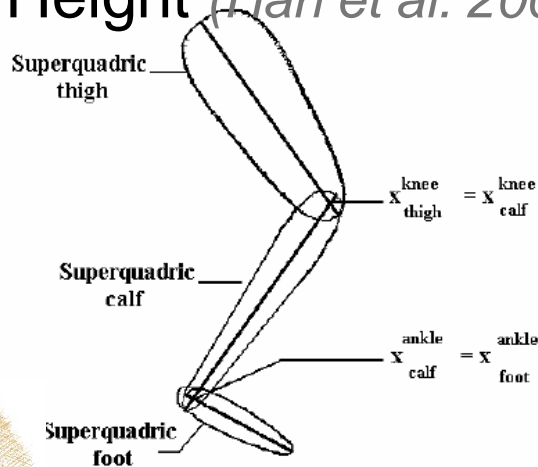
cardinality of Φ_θ (Boulgouris et al. 2004).



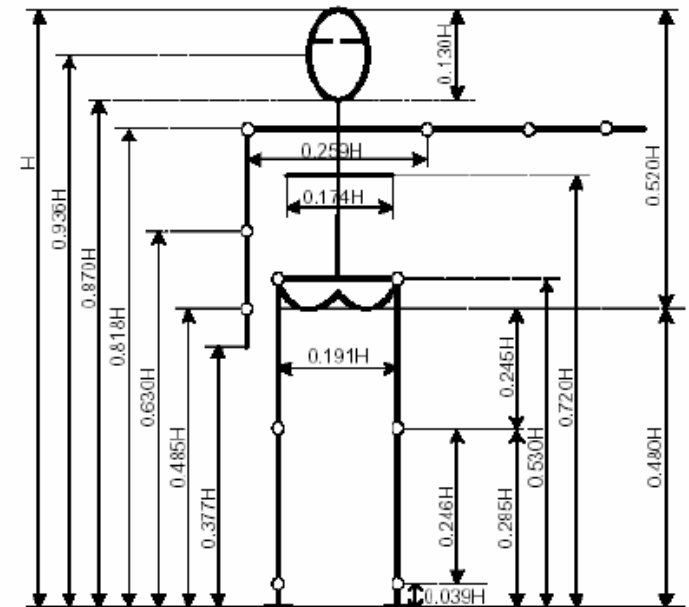
Features for gait recognition

Model-based method (1/2)

- Absolute joint positions (*Zhang et al. 2004*)
- Limb angles (*Zhang et al. 2004, Goffredo et al. 2008, Lu et al. 2006*)
- Lengths of various body parts (*Lu et al. 2006*)
- Widths/thickness of body parts (
- Height (*Han et al. 2006*)



(From Desseree & Legrand, 2005)



Human parts proportions (Winter 2004)

Features for gait recognition

Model-based method (2/2)

- Area of a body component
- Vector distance between the gravity centres of a body component and the whole body

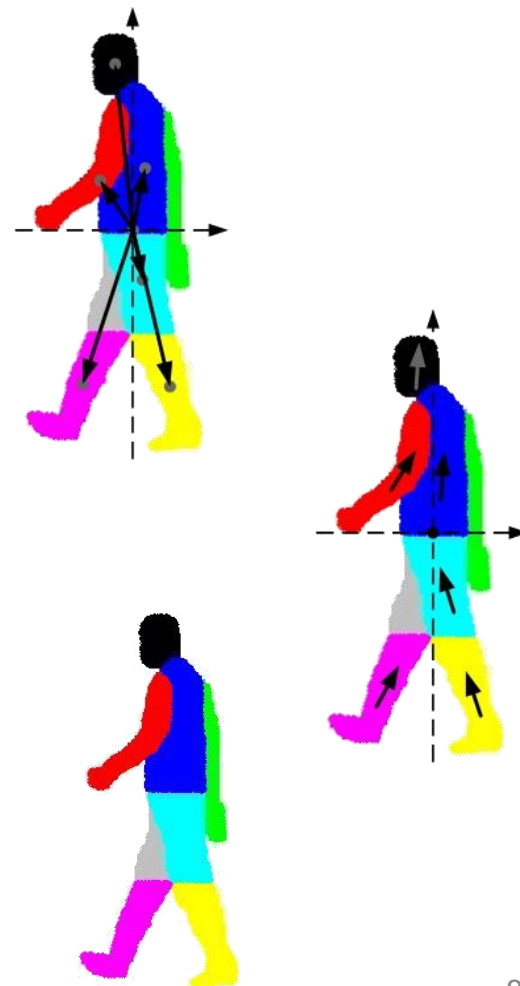
(Huang & Boulgouris, 2009)

- Orientation of a body component

(Huang & Boulgouris, 2009)

- Similarity based on body components

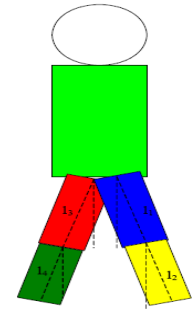
(Boulgouris & Chi, 2007)



Creating a human model

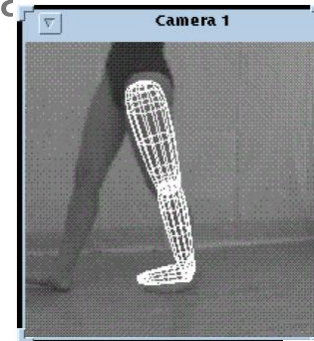
Whole body models

- 14 rigid bodies (*Wang et al. 2004*)
- 5 rigid bodies (*Zhang et al. 2007; Cheng-Chang et al. 2007*)

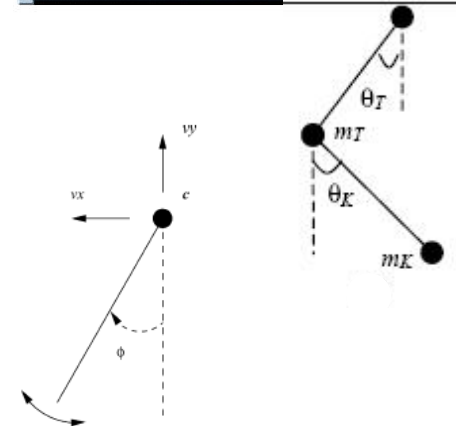


Leg model models

- 3 rigid bodies (thigh, shank, foot)
(*Desseree et al. 2005*)



- 2 rigid bodies (thigh and shank)
(*Yam et al. 2002*)



rigid body (the thigh)

(*Cunado et al. 2003*)

Creating a human model

Constrains:

- Gait is symmetrical and periodical (healthy population)
 - ➔ adequate to define the model for only one leg
 - ➔ the same model can describe either leg
- Dependency of the neighbouring joints: *(Yam et al. 2004)*
 - shoulder-elbow, thigh-knee, knee-ankle
 - lower limb is driven by the upper limb
- *(Wang et al. 2004)*
• minimum and maximum values of the joint angles

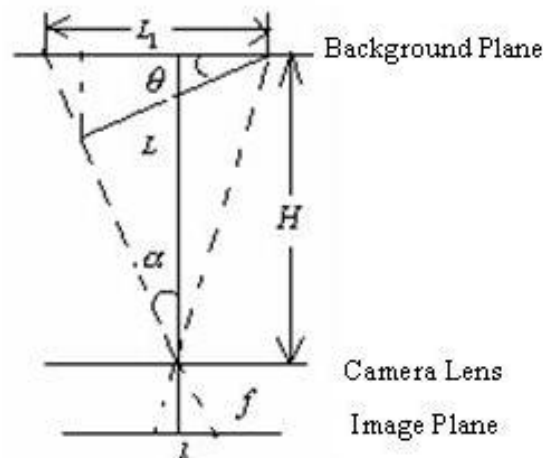


Problems in gait recognition

Overcoming the dependency to the gait direction

- Based on pinhole imaging model and weak-perspective projection, the gait direction is obtained and thus the projection function:

$$L = \frac{lH}{f(\cos \theta + \sin \theta \tan \alpha)}$$



Recognition rate: 50-70%

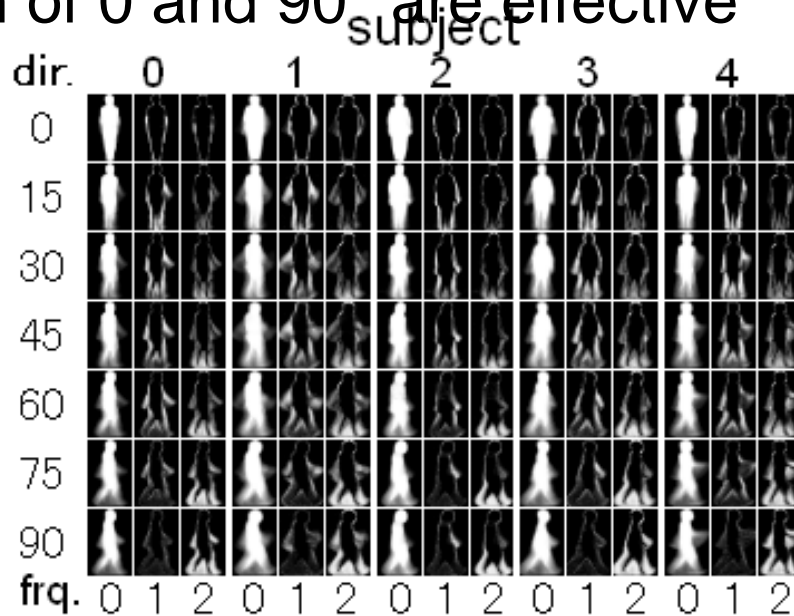
(Han et al. 2006)



Problems in gait recognition

Proven references to be used for gait invariant analysis

- Simple Reference: 45 deg and 135 deg angle views achieve better results
- Two references: Combinations of orthogonal references such as combination of 0 and 90° are effective



(Makihara et al, 2006)

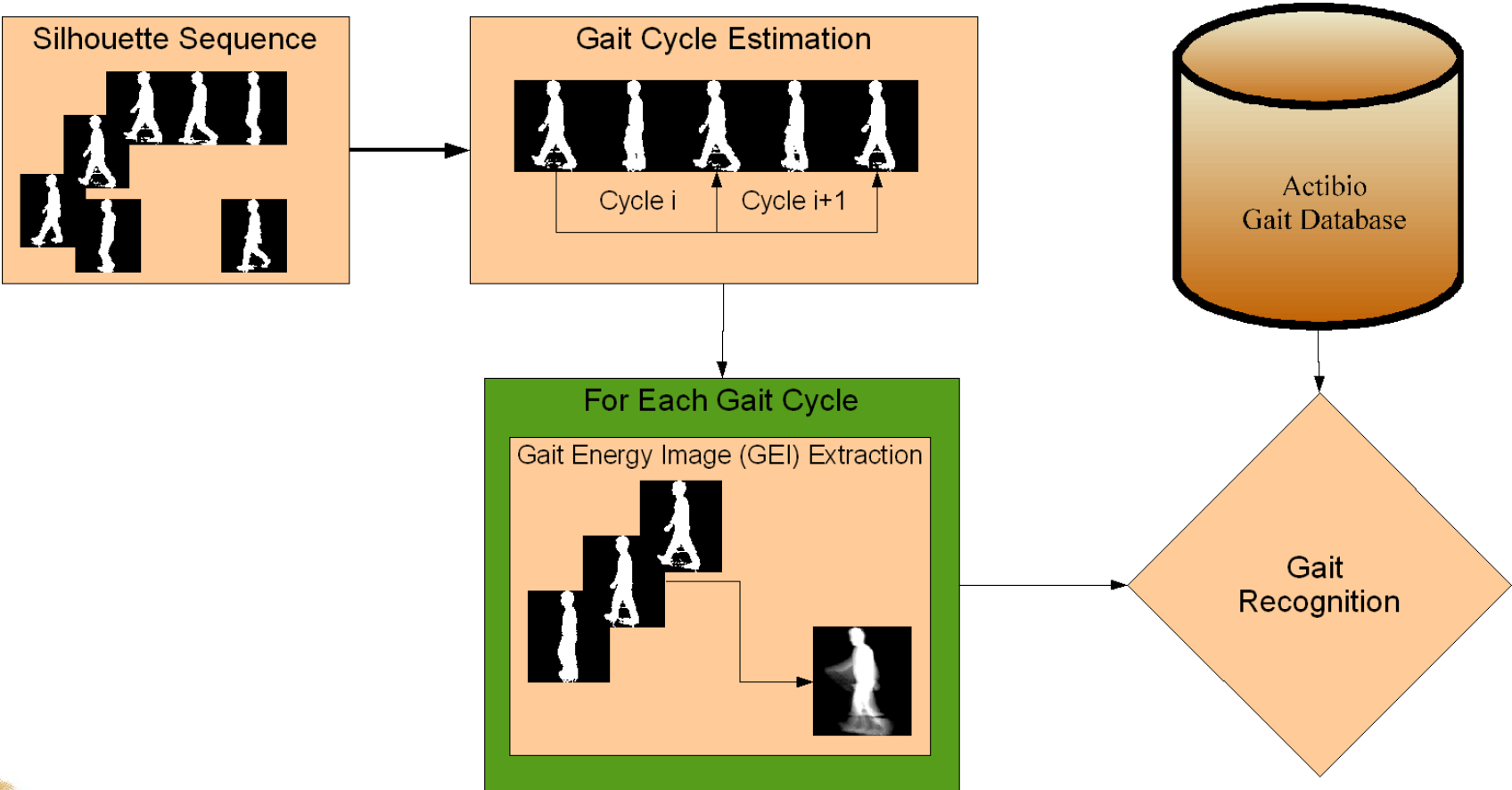


Functionalities

- Live capturing frames / Reading frames from database
- Human Extraction based on original frames
- Live Enrolment and Authentication based on live capturing and Human Extraction
- Offline Enrolment and Authentication on the database



Gait Recognition Algorithm Block Diagram



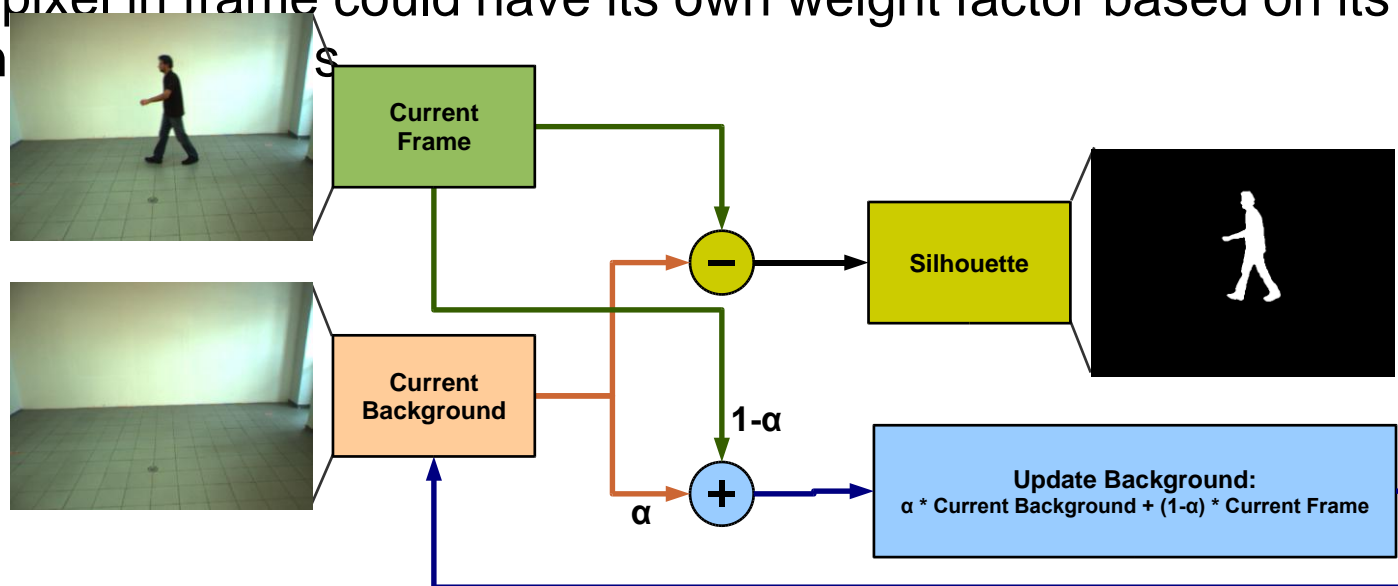
Dynamic background update and silhouette extraction

- Background image is updated every time a new frame arrives using an interpolation rule:

$$\alpha * \text{Current Background} + (1-\alpha) * \text{Current Frame}$$

where α is an adaptive weight factor that changes in every frame.

- Depending on chosen segmentation algorithm, we could have: the same α for all frame pixels, based in global illumination changes or every pixel in frame could have its own weight factor based on its local illumination



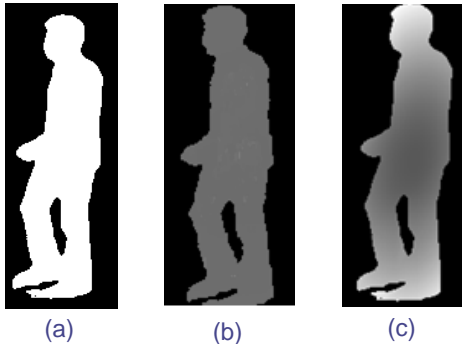
Pre-Processing Stage

Pre-processing of foreground silhouettes
(Noise removal)



Usage of Depth

Silhouette Representation

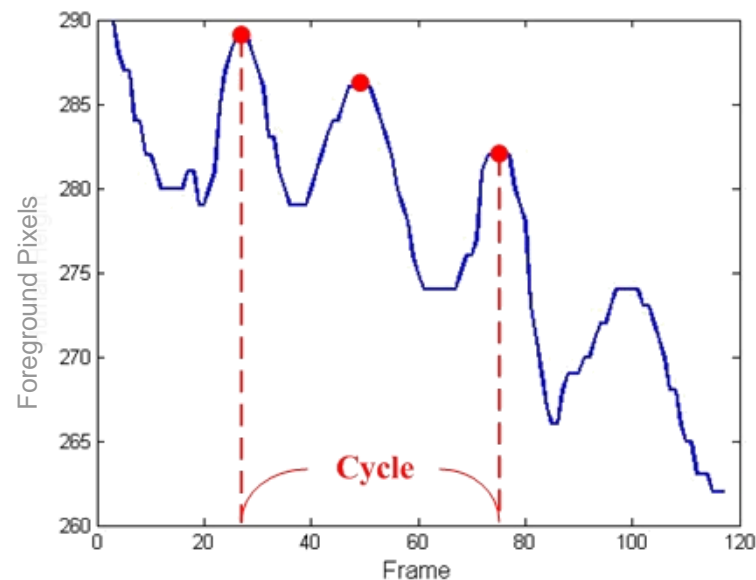
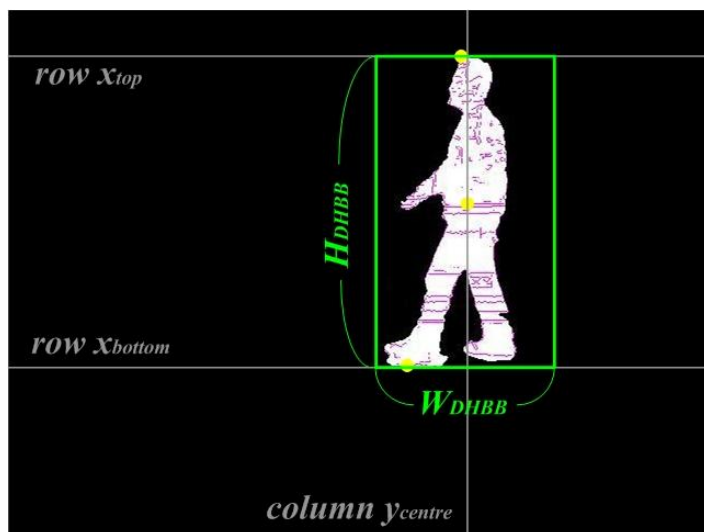


Binary Silhouette (a),
3D Radial Distributed Silhouette (b), and
3D Geodesic Distributed Silhouette (c)



Gait Cycle Estimation: Algorithm I

- Human Height (calculated in the Human Extraction stage).
- Number of foreground pixels in the lower half of the human silhouettes.



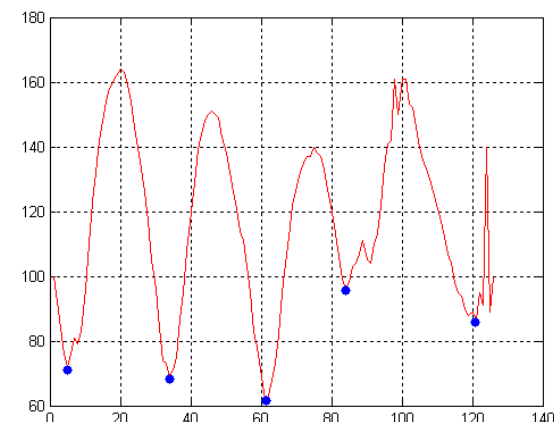
Period : Number of frames in one cycle (NF_c) is calculated by observing the peaks.



Gait Cycle Estimation: Algorithm II



- Period is estimated by calculating the bounding box's width in each frame and getting the local minima for the entire sequence
- Local Minima that are too close (<5 Frames) are discarded.



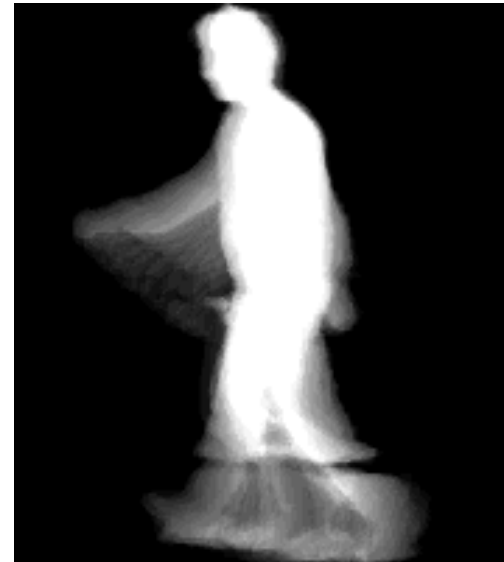
$$C_i = \text{LocalMin}(\text{BoundingBoxWidth}(n) | n \in 1, \text{numberOfFrames})$$



Utilization of Gait Energy Images (GEI)



Gait frame

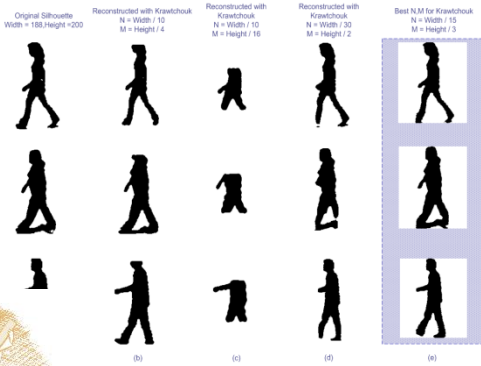
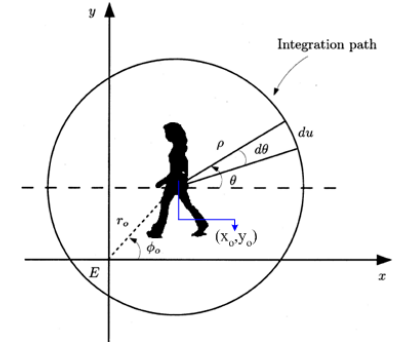
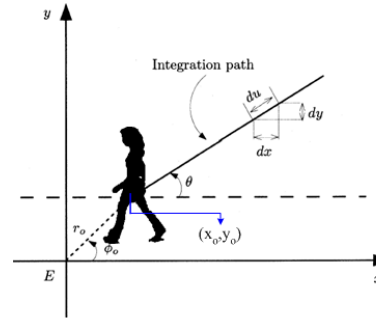
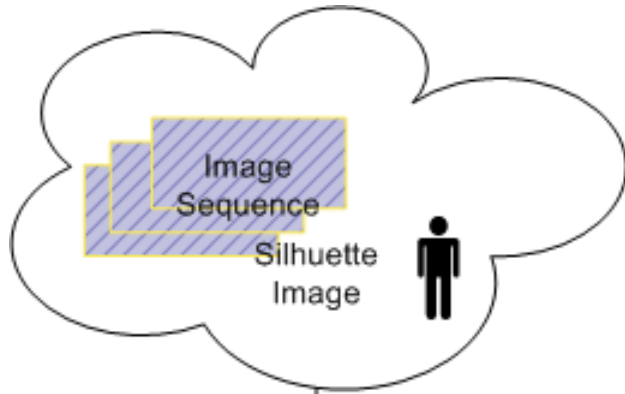


Gait Energy Image

$$GEI = \frac{1}{CycleLength} \cdot \sum_{i=CycleStart}^{CycleEnd} gaitFrame(i)$$



Signature Extraction Stage

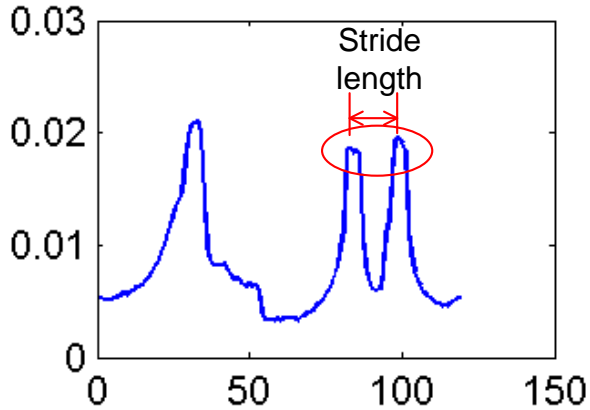


$$Q_{nm} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \bar{K}_n(x; p_1, N-1) \bar{K}_m(y; p_2, M-1) f(x, y)$$

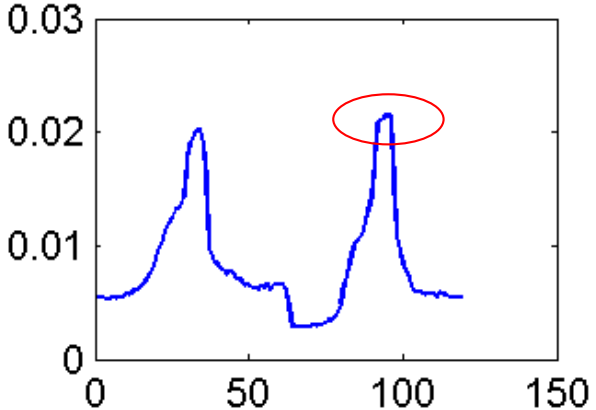


Mapping of gait characteristics into gait transforms

Double support position



Mid-stance position

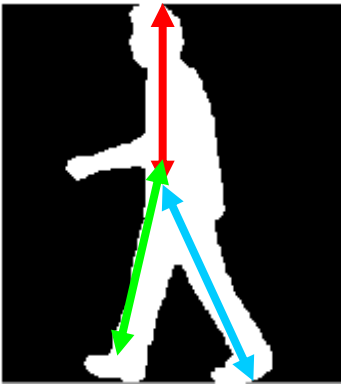


Length of stride mapping in the RIT Transform

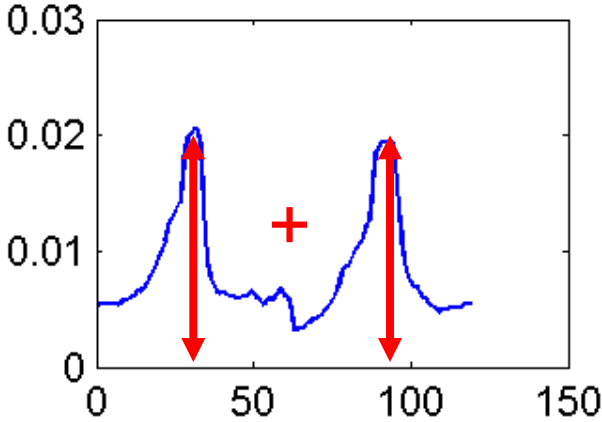
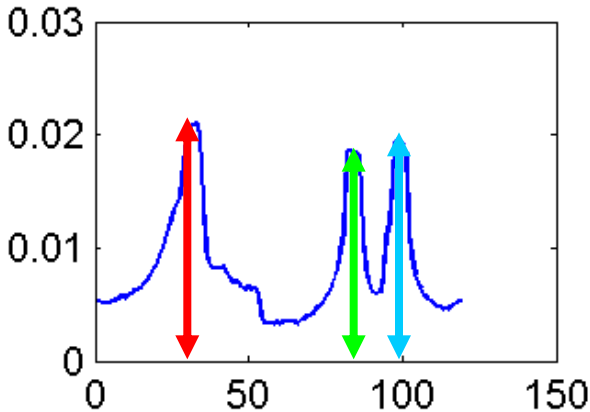
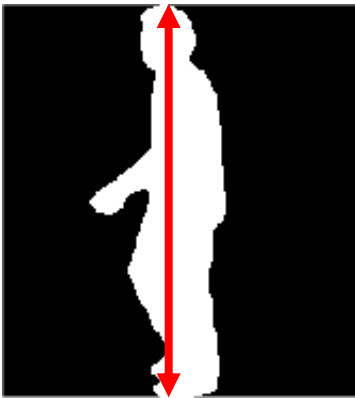


Mapping of gait characteristics into gait transforms

Distance between:
➤ Pelvis and feet
➤ Pelvis and head



Height estimation

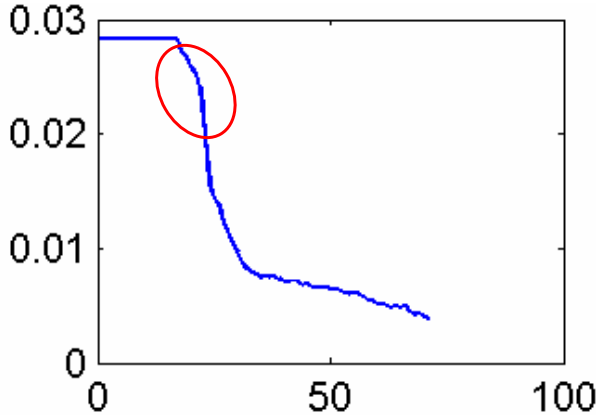
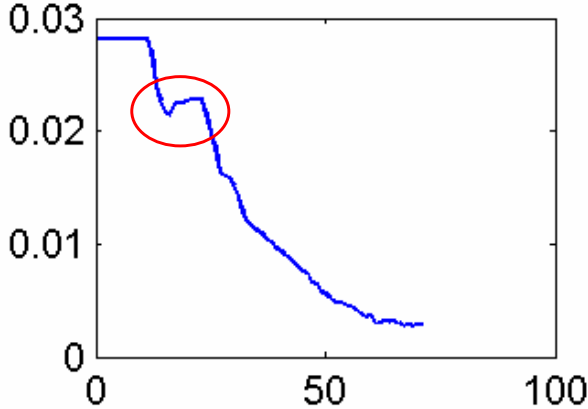
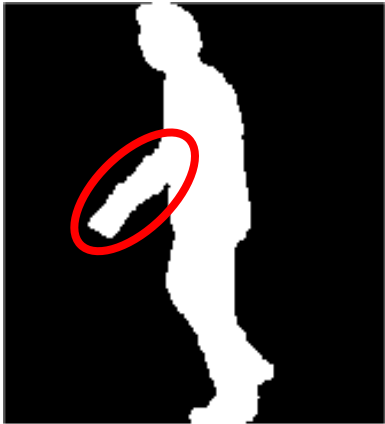


Gait dynamics parameters of the RIT Transform



Mapping of gait characteristics into gait transforms

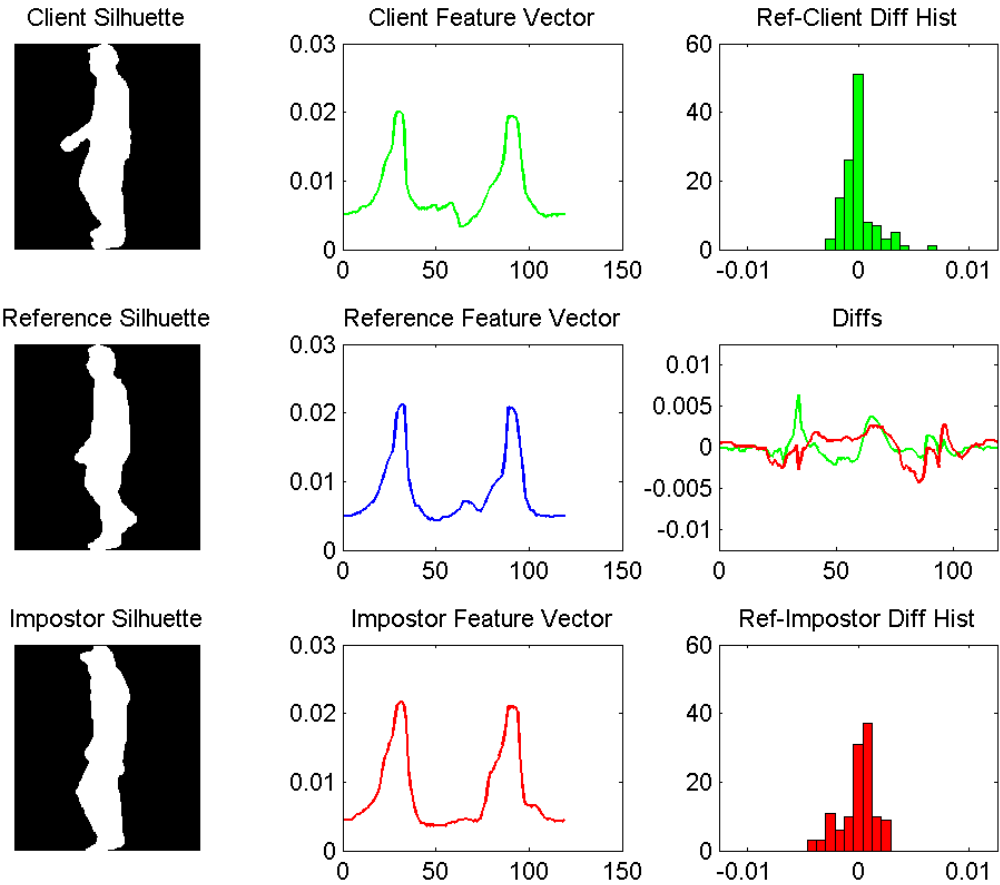
Hand movement



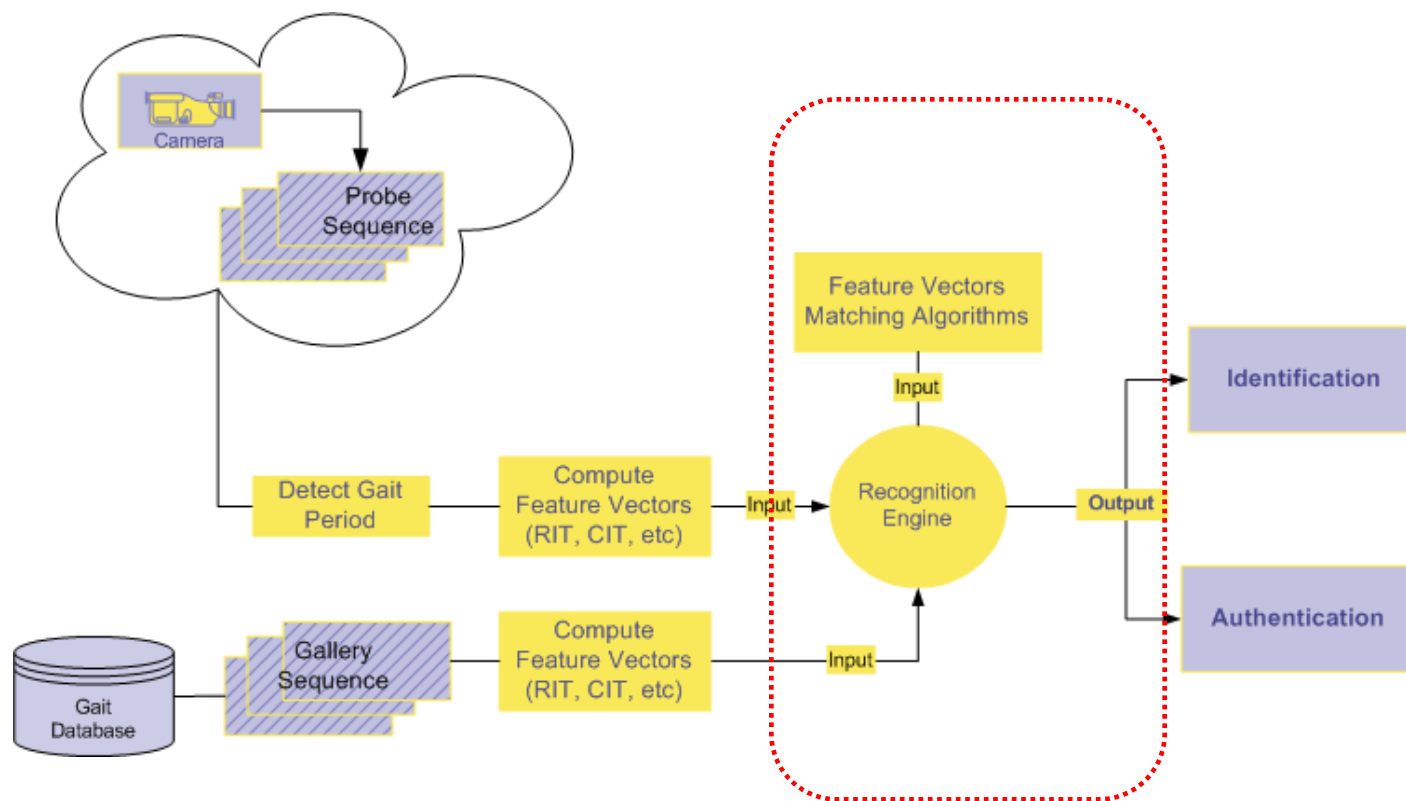
Hand movement detection using the CIT Transform



Mapping of gait characteristics into gait transforms



Template Matching using temporal correlation



For each gait cycle compute the distance between the stored (gallery) and the claimed (probe) gait signature

$$D = \min_l \sum_{i=1}^{Np} \sqrt{\sum_{n=0}^{FVecSize} (\text{ProbFeatVec}_i(n) - \text{GalleryFeatVec}_{i+l}(n))^2}$$



Τέλος Ενότητας

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



ΣΗΜΕΙΩΜΑΤΑ

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση **1.0**.

Έχουν προηγηθεί οι κάτωθι εκδόσεις:

- Έκδοση **1.0**. διαθέσιμη [εδώ](#).



Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών, Φακωτάκης Νίκος, Μουστάκας Κωνσταντίνος, Πέππας Παύλος, Σγάρμπας Κυριάκος. «Τεχνητή Νοημοσύνη II, Αντίληψη». Έκδοση: 1.0. Πάτρα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση:

https://eclass.upatras.gr/modules/course_metadata/opencourses.php?fc=15



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.



Σημείωμα Χρήσης Έργων Τρίτων

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Εικόνες/Σχήματα/Διαγράμματα/Φωτογραφίες

- Η διάλεξη είναι εμπνευσμένη και έχει χρησιμοποιηθεί εκπαιδευτικό και πολυμεσικό υλικό από τις διαλέξεις του James Hays του «Brown University, Computer Science Department» και του αντίστοιχου μαθήματος «CS 143 Introduction to Computer Vision» (<http://cs.brown.edu/courses/cs143/>).

