

Suivi 3D monoculaire appliqué à la détection de chutes

Caroline Rougier, Jean Meunier,
Alain St-Arnaud, Jacqueline Rousseau

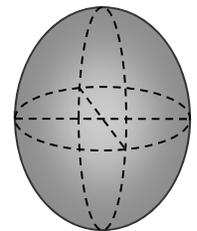
ACFAS 2011 - Sherbroke

Colloque "Vidéo numérique: Traitement, analyse et reconnaissance"

13 mai 2011

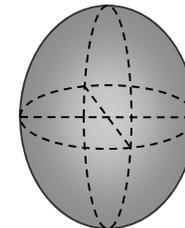
Trajectoire 3D de la tête

- Utile pour les applications de vidéosurveillance
 - Pour localiser une personne dans une pièce
 - Pour la détection de chutes
 - Une chute peut être détectée à partir de vitesses 3D (Etude avec 3 caméras + marqueurs sur la personne (*Wu, 2000*))
- Objectifs
 - Suivi et localisation 3D de la tête en temps réel
 - Avec une seule caméra calibrée, mais aucun marqueurs
 - Tête représentée par une ellipsoïde 3D
 - Tests
 - Précision 3D testée avec la bibliothèque HumanEva (vidéos + données MoCap)
 - Application à la détection de chutes

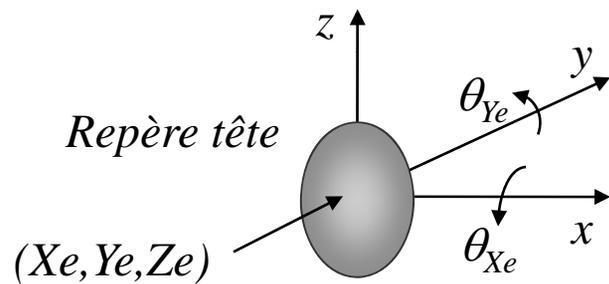


Suivi 3D de la tête

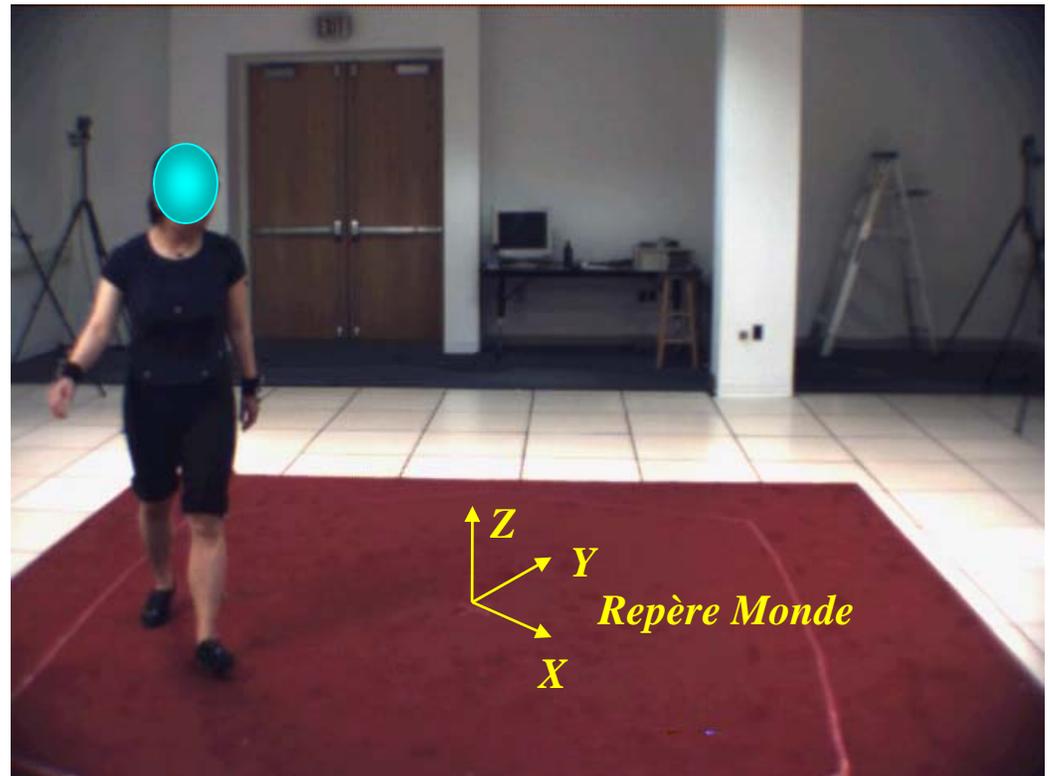
- Pourquoi la tête ?
 - Quasiment toujours visible dans l'image
 - Beaucoup de mouvement lors d'une chute
 - Forme caractéristique ellipsoïdale
- Principe de la méthode:
 - Tête représentée par une ellipsoïde 3D
 - Suivi \Rightarrow filtre à particules
 - Trajectoires avec changements brusques de direction



Représentation 3D de la tête



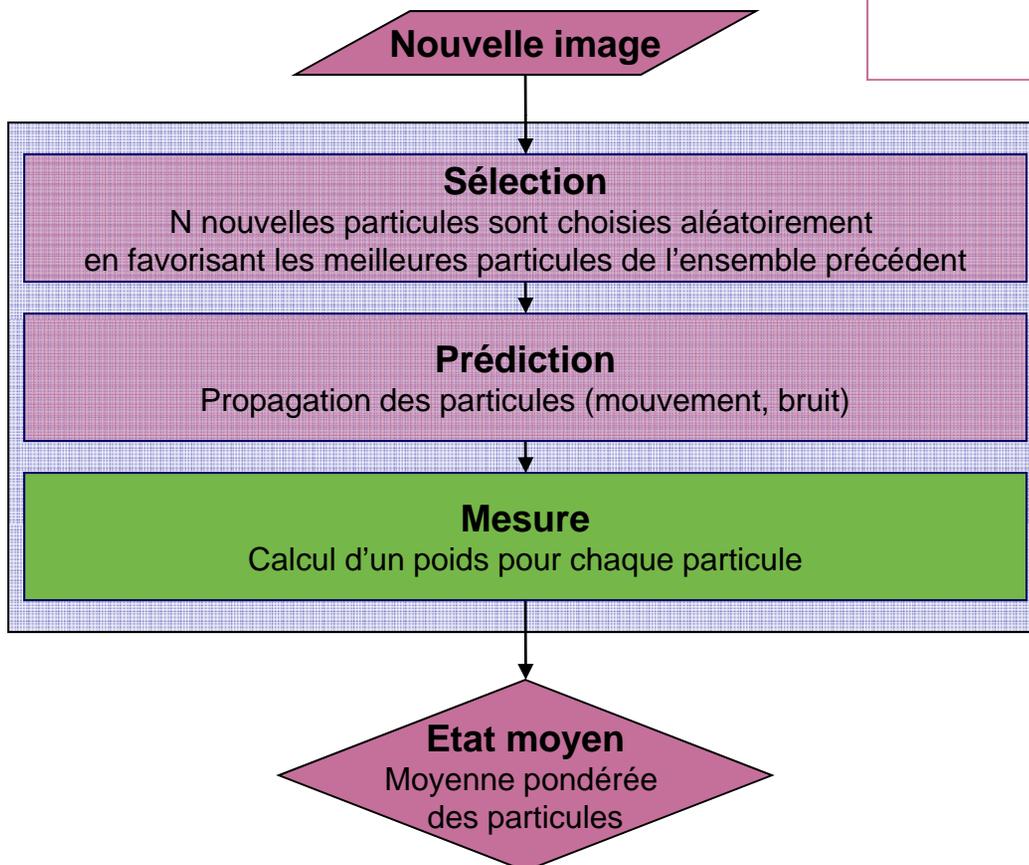
$$[X_e, Y_e, Z_e, \theta_{X_e}, \theta_{Y_e}]$$



Filtre à particules

Chaque particule est une ellipsoïde 3D:

$$s_t^n = [X_e, Y_e, Z_e, \theta_{Xe}, \theta_{Ye}]_t^n$$

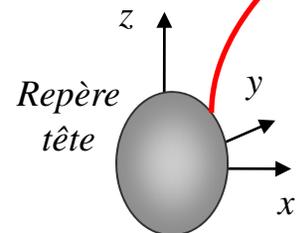
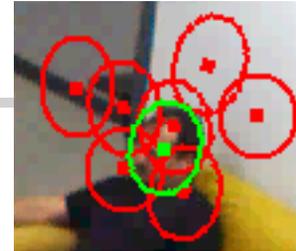


$$s_t^n = A_l s_t'^n + B_l w_t^n$$

⇒ Projection de l'ellipsoïde 3D dans le plan image = ellipse

Mesure

- Mesure:
 - Calcul d'un poids pour chaque particule
 - Projection d'une ellipsoïde 3D dans le plan image = ellipse
- Poids de l'ellipse basé sur:
 - Un coefficient objet
 - Un coefficient couleur
 - Un coefficient corps



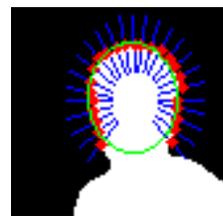
Projection dans
le plan image



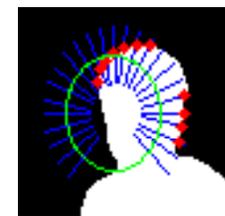
Mesure: coefficient objet



$$C_F = \frac{1}{N_e} \sum_{n=1}^{N_e} \frac{D_e(n) - d_e(n)}{D_e(n)}, \quad C_F \in [0 \dots 1]$$

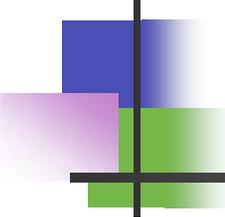


$$C_F = 0.90$$



$$C_F = 0.20$$

⇒ Utile pour la précision 3D



Mesure: coefficient couleur

- Histogramme couleur 3D normalisé (8x8x8 bins)
- Zone rectangulaire incluse dans l'ellipse
- Histogramme de référence mis à jour

$$C_C = \sum_{i=1}^{N_b} \min (H (i) , H_{ref} (i)) , \quad C_C \in [0 \dots 1]$$

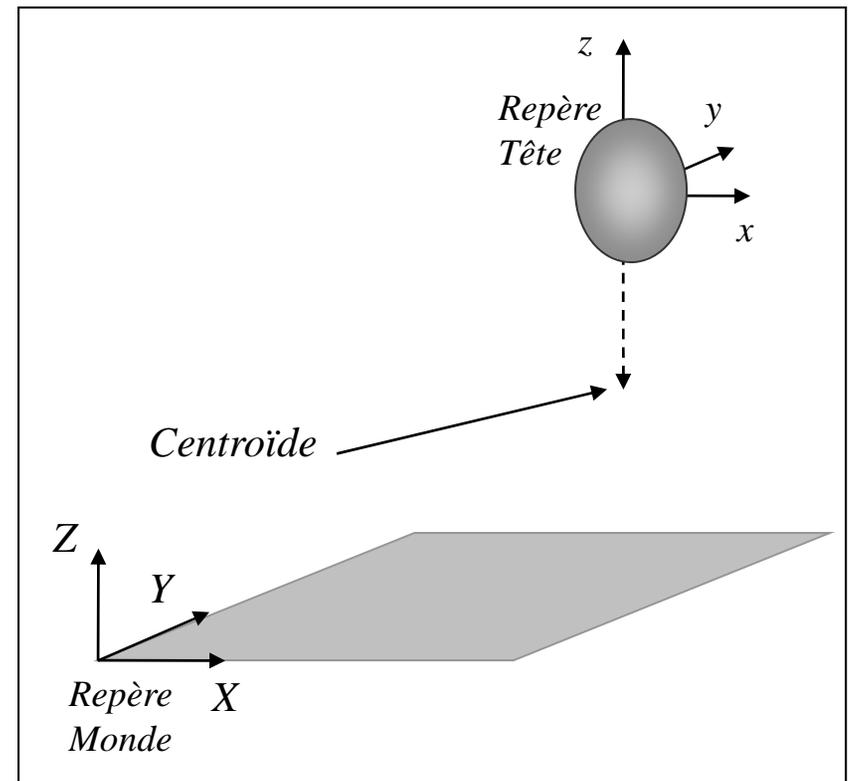
⇒ **Pour ne pas perdre le suivi lors de mouvements importants (ex: chute)**

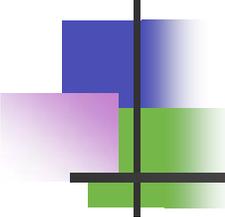
Mesure: coefficient corps

- Projection du point 3D (centroïde personne) doit être proche du centroïde 2D de la silhouette
⇒ distance d_b

$$C_B = \frac{D_b - d_b}{D_b}, \quad C_B \in [0 \dots 1]$$

⇒ Pour éviter des rotations anormales de l'ellipsoïde autour de son centre





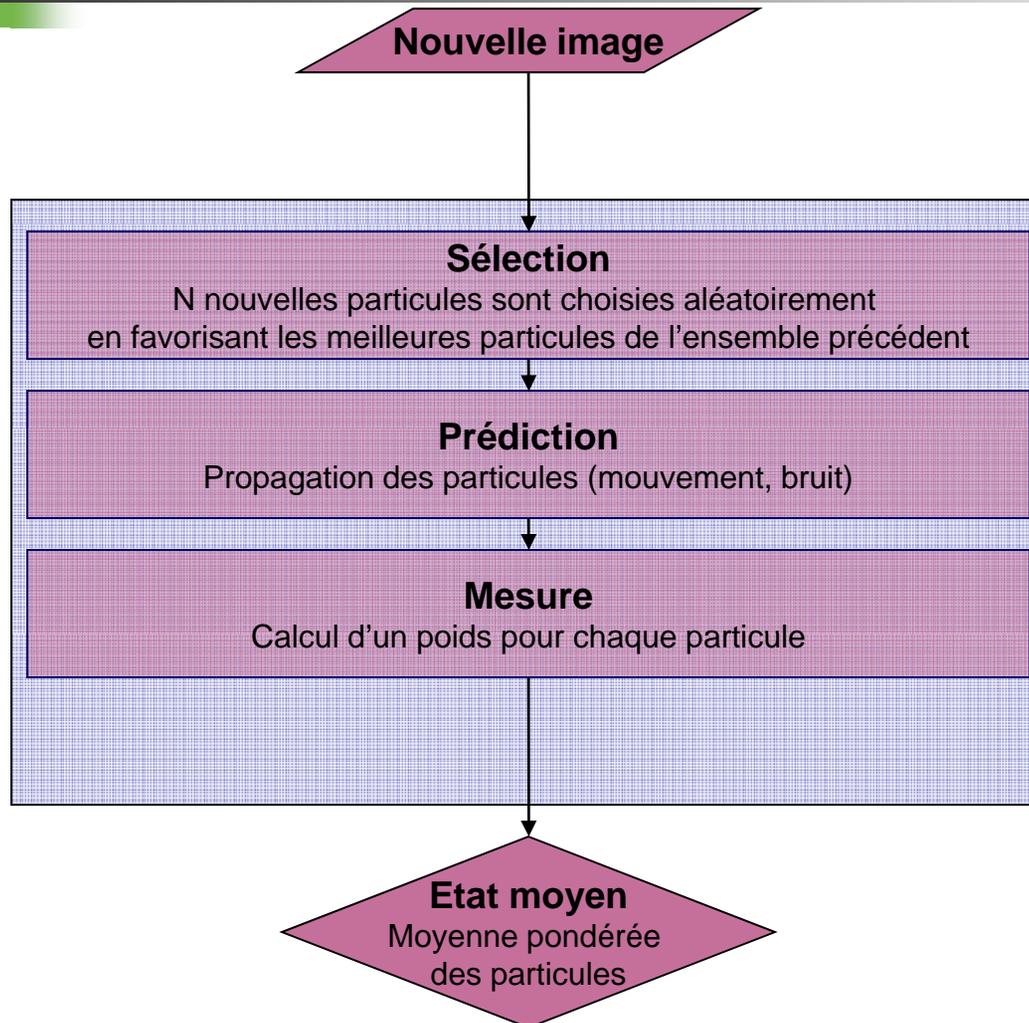
Mesure: poids final

- Amplification des poids ($\sigma=0.15$)

$$C_{final} = \frac{1}{\sqrt{2\pi\sigma}} \exp^{(C_F C_C C_B)/2\sigma^2}$$

⇒ Pour donner plus de poids aux meilleurs particules

Suivi - Filtre à particules original

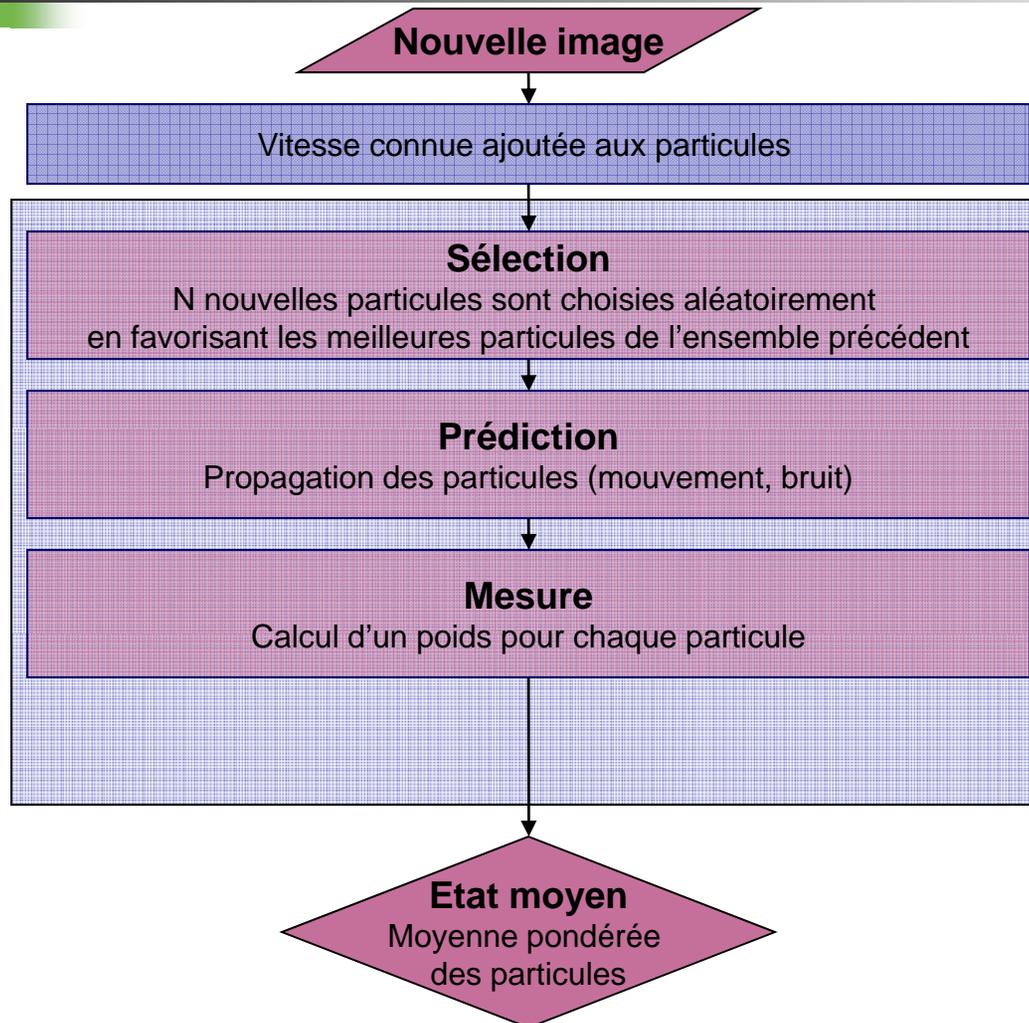


$$s_t^n = [X_e, Y_e, Z_e, \theta_{X_e}, \theta_{Y_e}]_t^n$$

$$s_t^n = A_l s_t'^n + B_l w_t^n$$

$$C_{final} = \frac{1}{\sqrt{2\pi\sigma}} \exp^{(C_F C_C C_B)/2\sigma^2}$$

Suivi - Filtre à particules amélioré



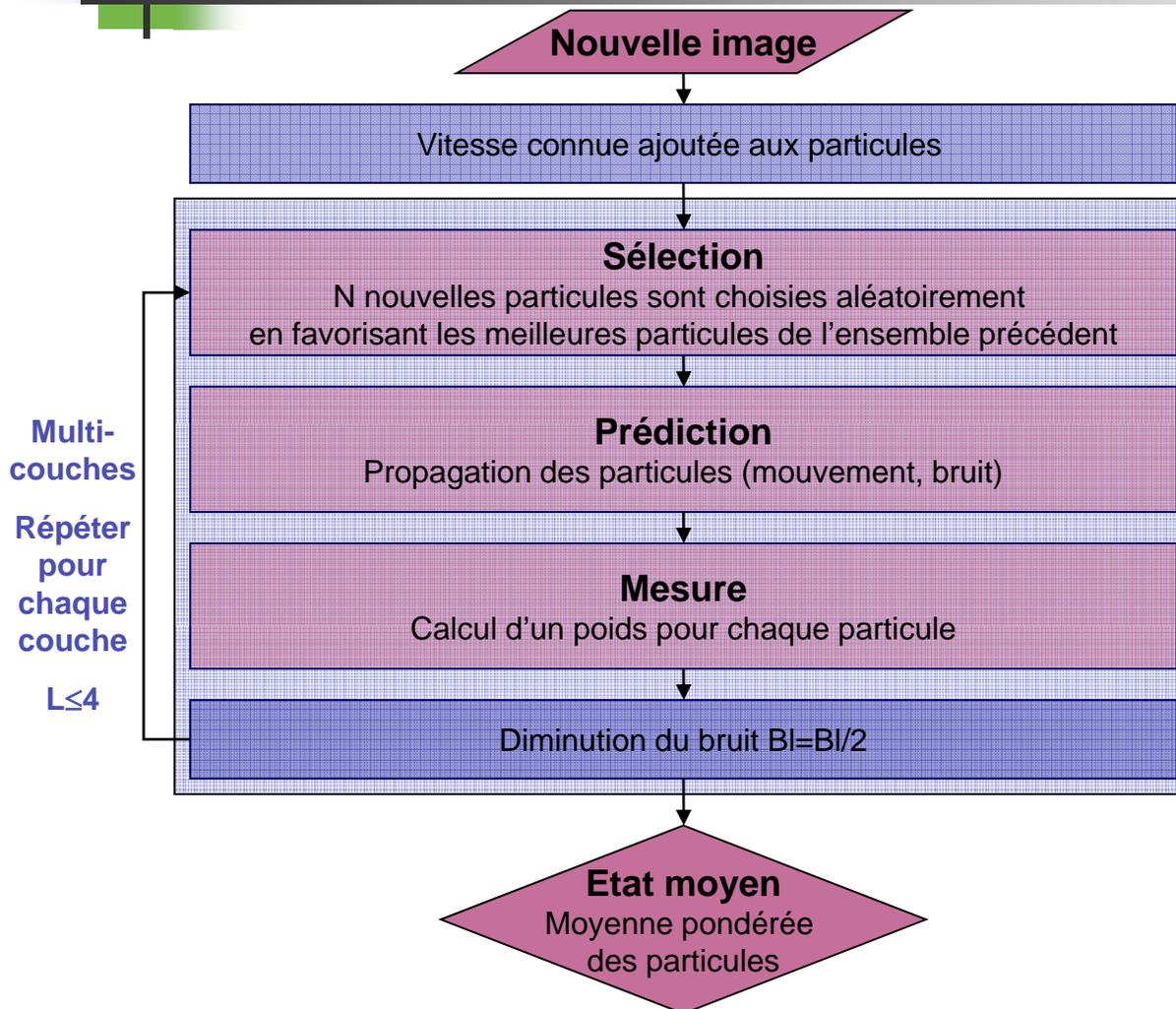
⇒ Ajout de mouvement

$$s_t^n = [X_e, Y_e, Z_e, \theta_{X_e}, \theta_{Y_e}]_t^n$$

$$s_t^n = A_l s_t'^n + B_l w_t^n$$

$$C_{final} = \frac{1}{\sqrt{2\pi\sigma}} \exp^{(C_F C_C C_B)/2\sigma^2}$$

Suivi - Filtre à particules amélioré



⇒ Ajout de mouvement

$$s_t^n = [X_e, Y_e, Z_e, \theta_{X_e}, \theta_{Y_e}]_t^n$$

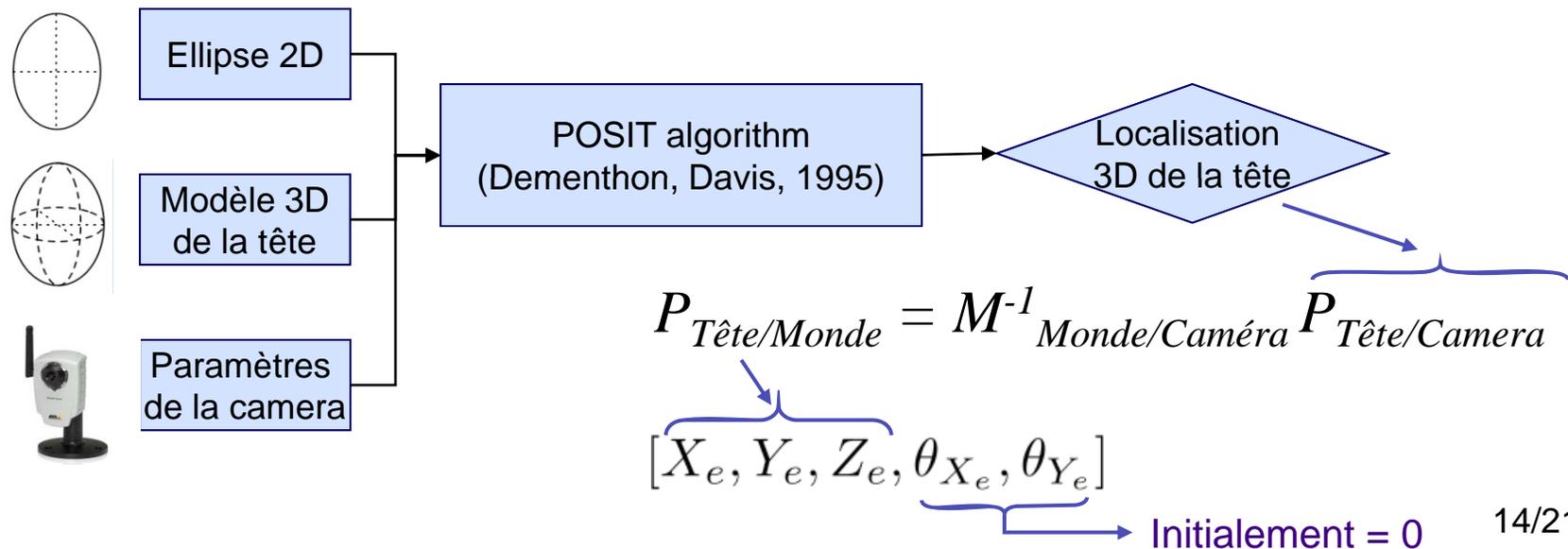
$$s_t^n = A_l s_t'^n + B_l w_t^n$$

$$C_{final} = \frac{1}{\sqrt{2\pi\sigma}} \exp^{(C_F C_C C_B)/2\sigma^2}$$

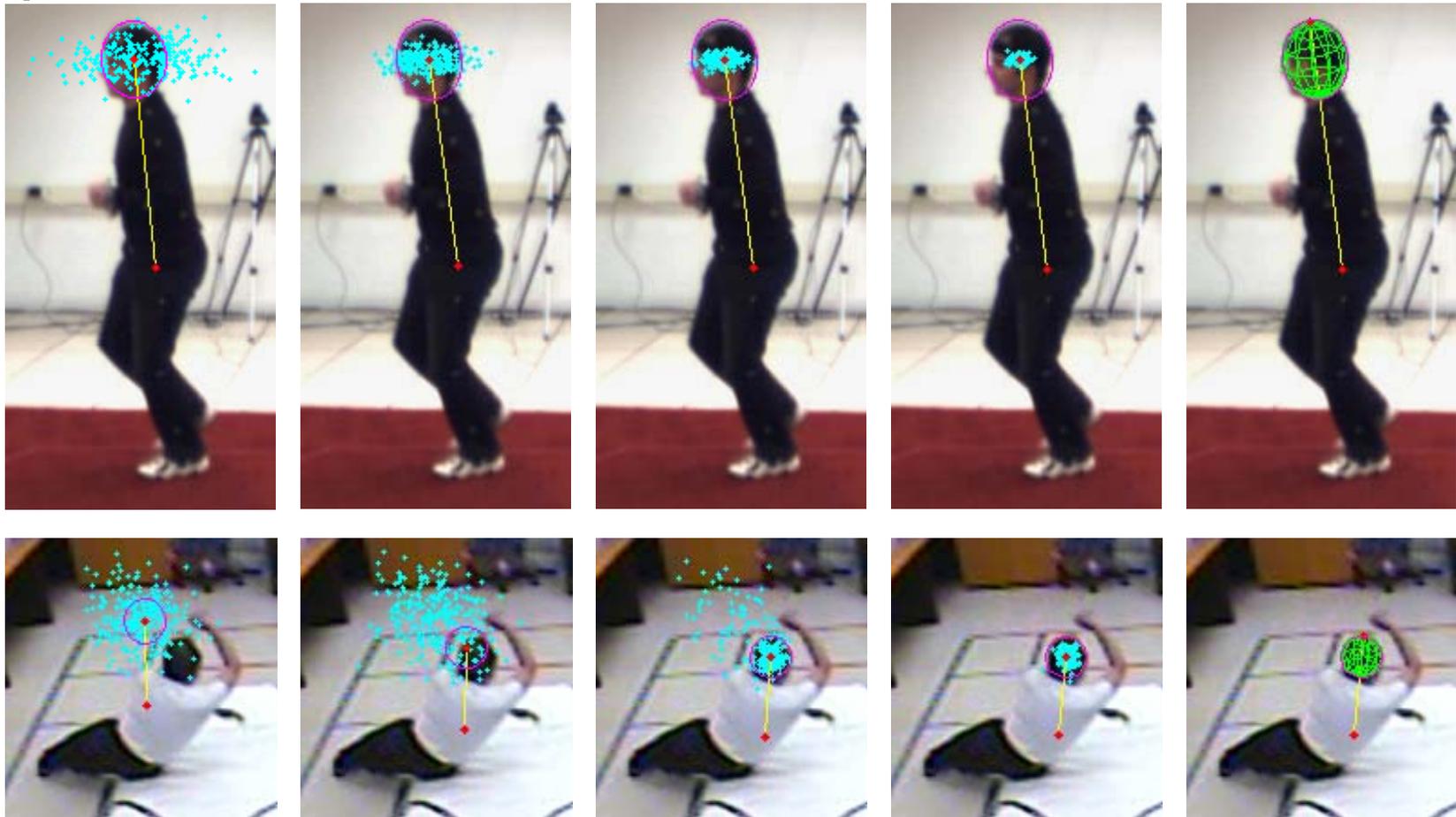
⇒ **Bruit en direction
du mouvement**

Localisation 3D initiale

- Proportions de la tête
 - Une image où la personne est debout
 - Hauteur de la personne connue H
 - Ratio de l'ellipse=1.2 pour une tête
- Localisation 3D initiale

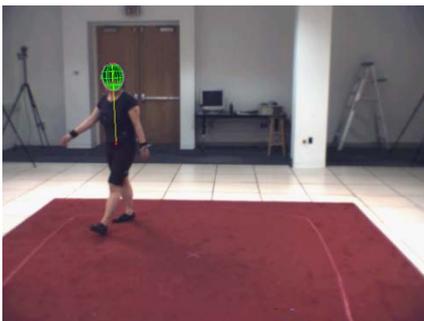
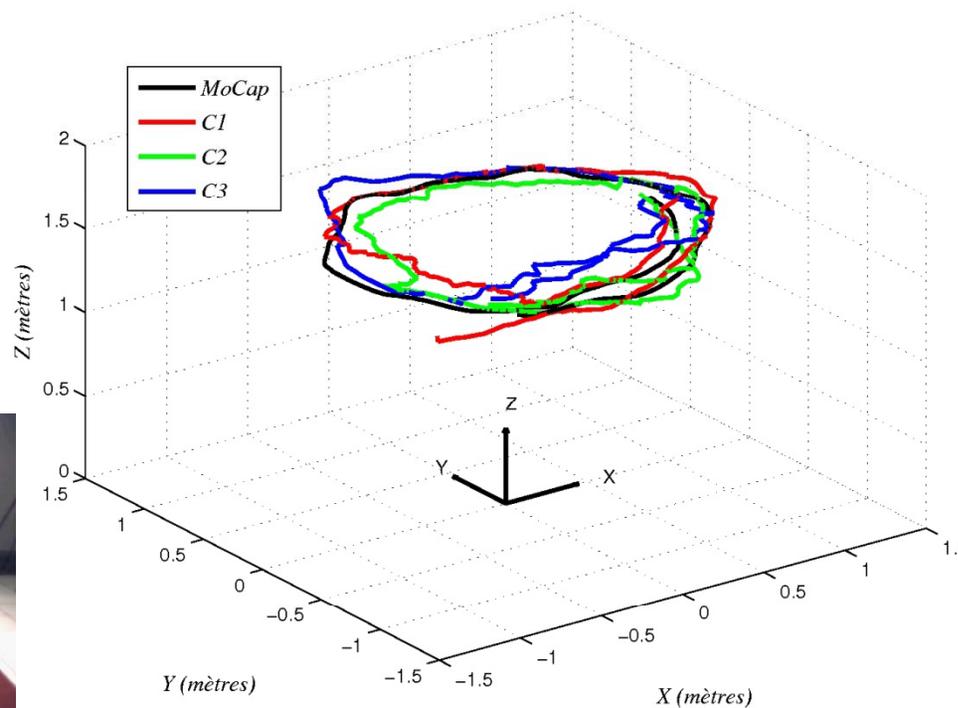


Exemples de suivi



Évaluation de la précision 3D

- Bibliothèque HumanEva
- Données MoCap (Motion Capture)
 - Vraie localisation 3D



Caméra 1

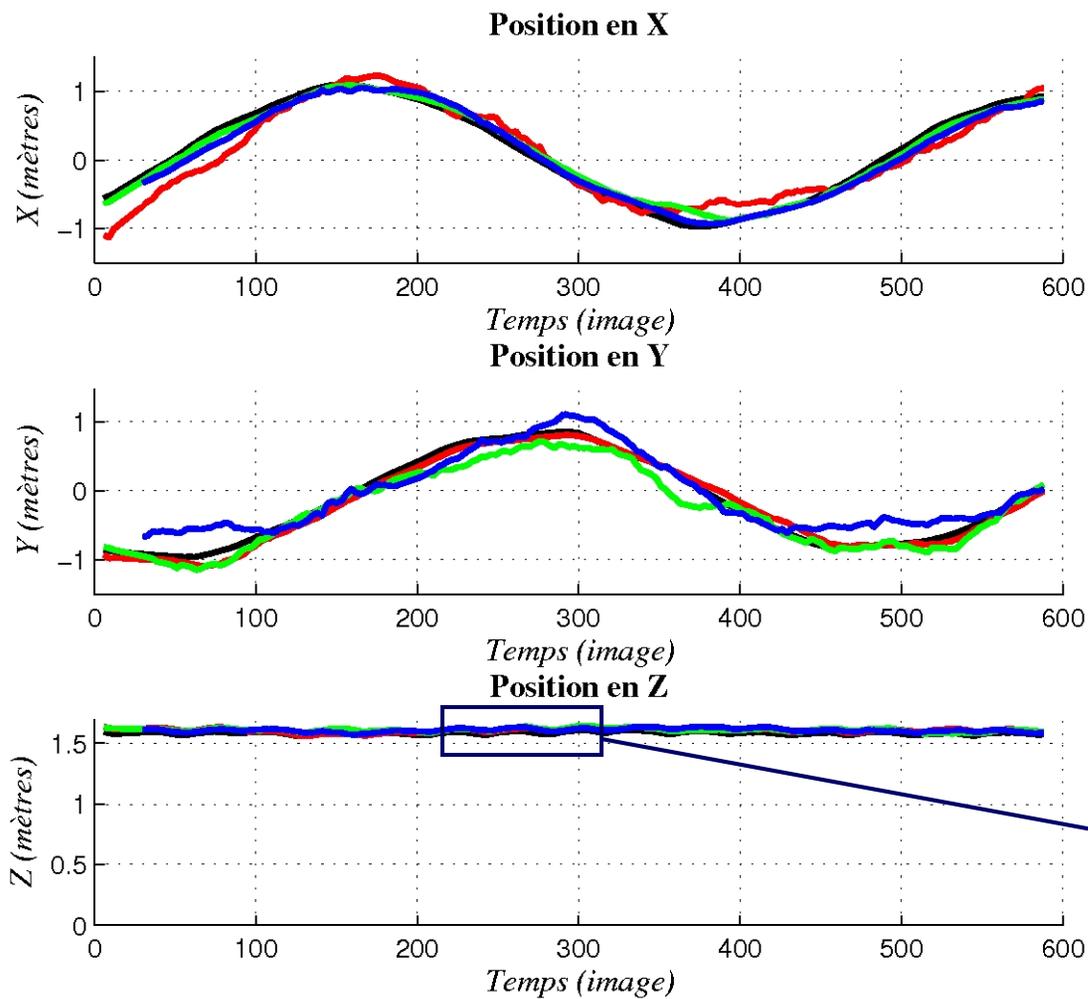


Caméra 2

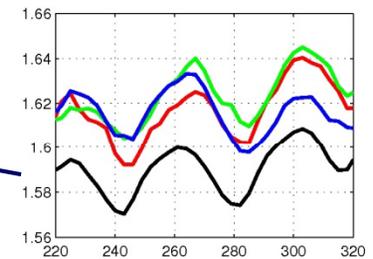


Caméra 3

Évaluation de la précision 3D



Erreur de l'ordre
de 25cm à 5m (5%)



Détection de chutes - exemples

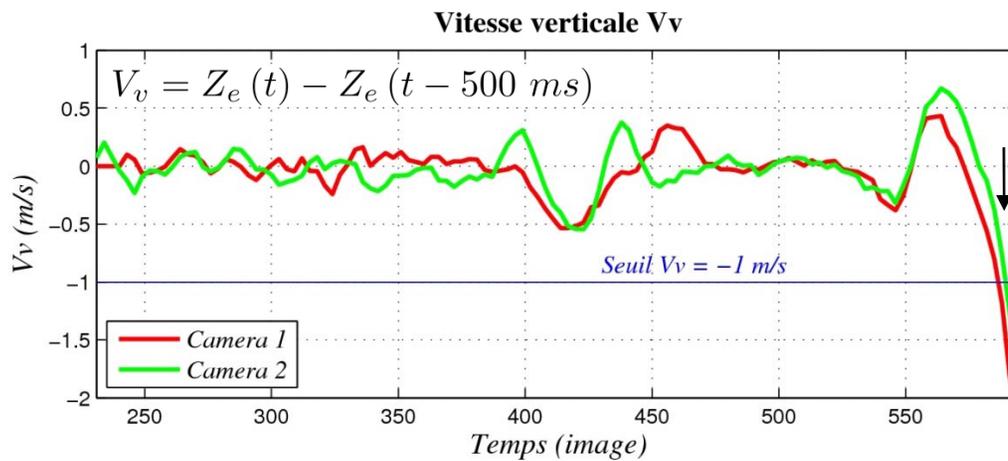
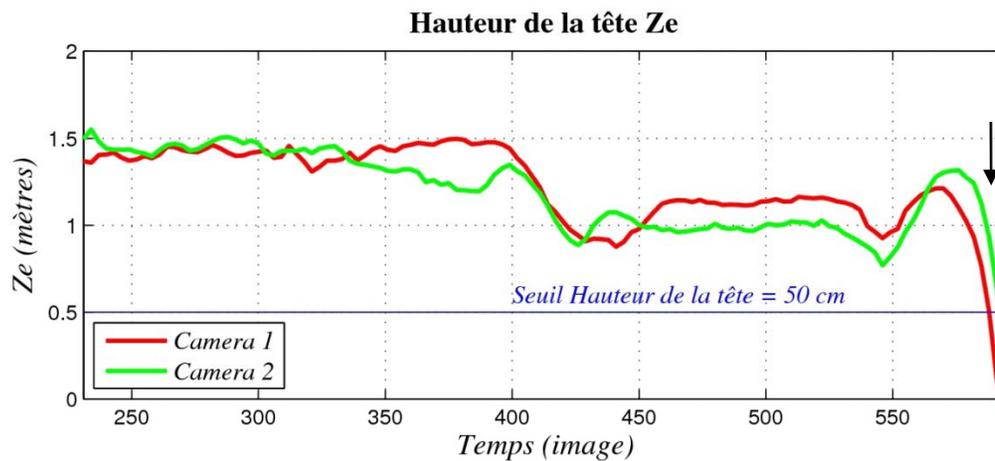
■ Caméra 1



■ Caméra 2



Détection de chutes - exemples



Caméra 1



Caméra 2



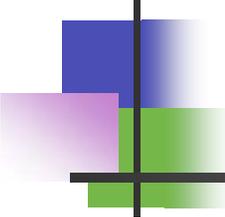
Image 321



Image 423

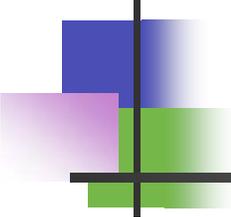


Image 585



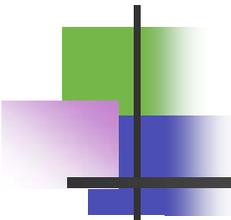
Détection de chutes - Résultats

- Test sur 10 chutes, 2 caméras
- Vitesse verticale de la tête V_v
 - Chute si $V_v < -1\text{m/s}$
 - 100% de sensibilité
 - Pas de fausses détections
- Hauteur de la tête Z_e
 - Chute si $Z_e < 50\text{cm}$
 - 1 seule chute non détectée dans une caméra (problème de segmentation)
 - Pas de fausses détections



Conclusion

- Il est possible de récupérer la trajectoire 3D de la tête avec une seule caméra calibrée
 - Erreur de l'ordre de 25cm à 5m (5%)
⇒ suffisant pour de l'analyse de trajectoires
- Avantages de la méthode
 - Initialisé automatiquement
 - Fonctionne même quand une partie du corps est occultée
 - Occultation avec un objet, entrée/sortie de la scène
 - Filtre à particules multi-couches:
 - Pour plus de précision en un temps raisonnable
- Inconvénients
 - Nécessite que la tête soit visible dans l'image et bien définie

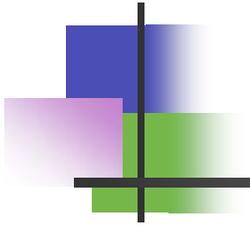


Merci!

Caroline Rougier, Jean Meunier,
Alain St-Arnaud, Jacqueline Rousseau

rougierc@iro.umontreal.ca





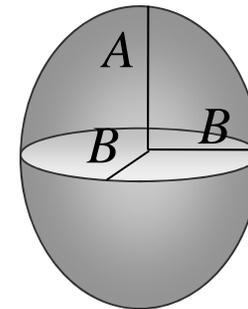
Modèle 3D ellipsoïdal de la tête

- Equation de l'ellipsoïde

$$\frac{x^2}{B^2} + \frac{y^2}{B^2} + \frac{z^2}{A^2} = 1$$

- Forme matricielle

$$Q_H = \begin{pmatrix} \frac{1}{B^2} & 0 & 0 & 0 \\ 0 & \frac{1}{B^2} & 0 & 0 \\ 0 & 0 & \frac{1}{A^2} & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$



Projection du modèle 3D

- Dans le repère caméra, l'ellipsoïde est représenté par une quadrique Q_C :

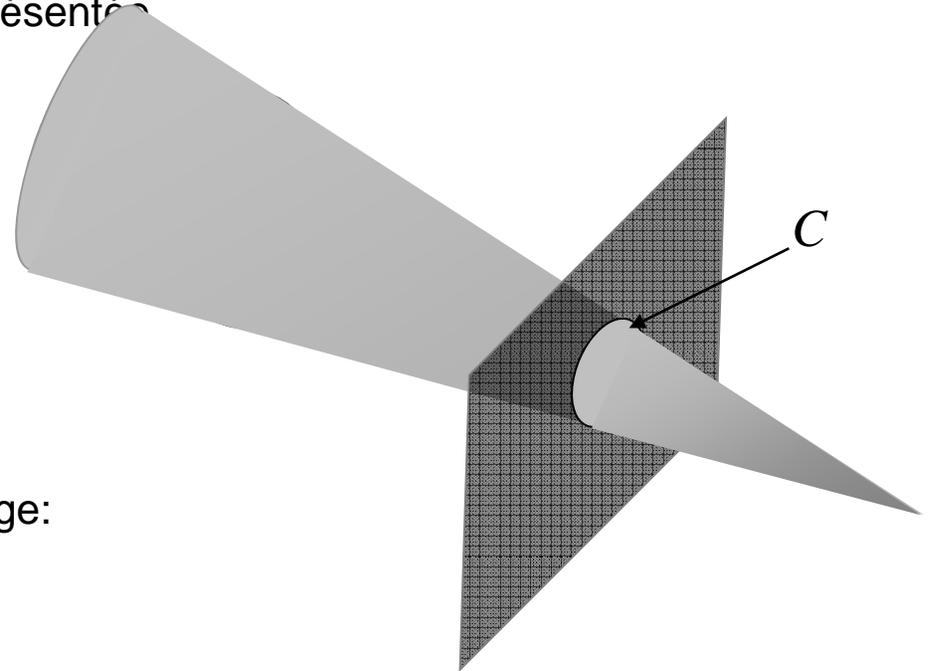
$$[x, y, z, 1]^T Q_C [x, y, z, 1] = 0$$

- Obtenue avec la matrice de projection P :

$$Q_C = P^{-1T} Q_H P^{-1}$$

- Qui donne une conique C dans le plan image:

$$C = Q_{C_{44}} Q_{C_{1:3,1:3}} - Q_{C_{1:3,4}} Q_{C_{1:3,4}}^T$$



■⇒ Paramètres de l'ellipse

Projection du modèle 3D

- Matrice de projection

$$P = K M_{ext} M_{Tête/Monde}$$

- Paramètres internes

$$K = \begin{pmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

- Paramètres externes

Repère monde par rapport au repère caméra

$$M_{ext} = \begin{pmatrix} R & T \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- $M_{Tête/Monde}$

- Rotation/translation de l'ellipsoïde par rapport au repère monde

