Robust Visual Golf Club Tracking

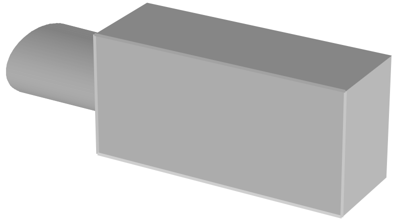
Vincent Lepetit, Nicolas Gehring, Pascal Fua
CVlab - EPFL

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

DARTFISH

Aim

1.5 sec



- Completely automatic video based system.
  - No user intervention.
  - Use of a single video (PAL) camera.
  - No external expensive devices.
  - No specifically instrumented golf clubs or clothing.
- Usable in natural environments
  - Cluttered (fixed) background

Club: thin, specular reflexion, moves very fast (up to 170km/h).

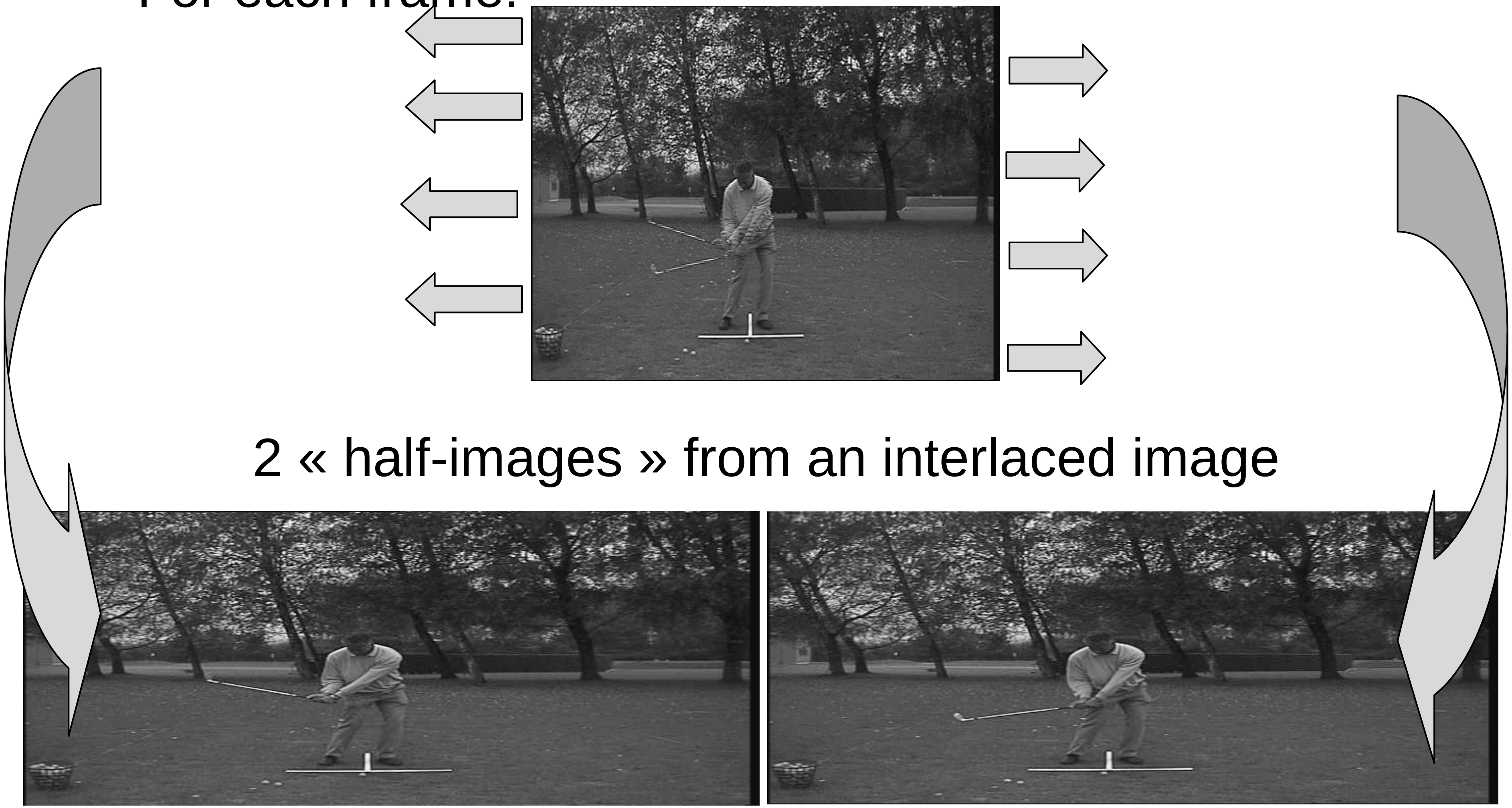
t

$t+\frac{1}{25}\,\mathrm{sec}$

- Club extraction
- Tracking algorithm with
  - local motion model
  - global motion model

Dealing with interlaced images

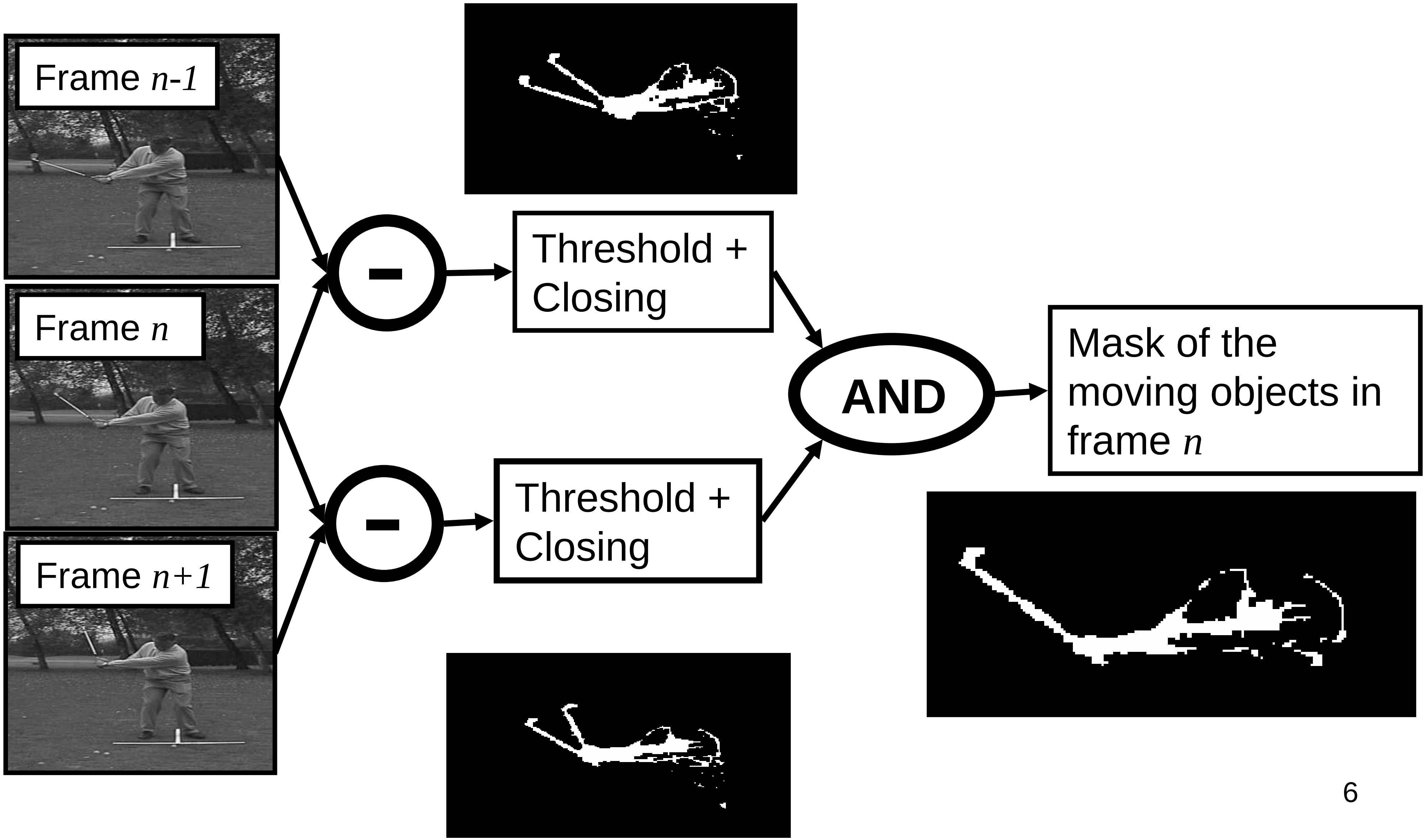
- For each frame:



2 « half-images » from an interlaced image

50 « half-images » per second

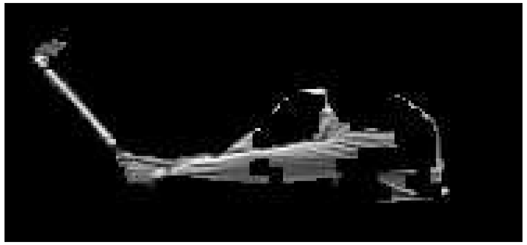
Detection of moving objects



Hypothesis generation

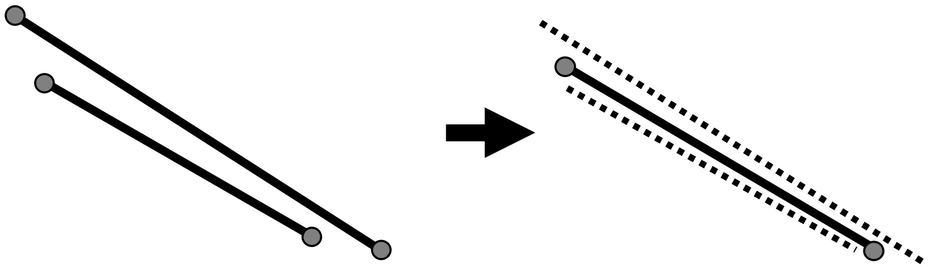
Detection of adjacent parallel segments under the moving-object mask

1. Edges detection

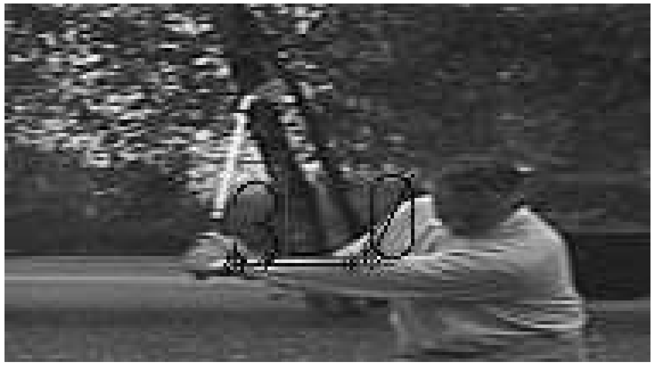

2. Segment detection (contour extraction, polygonal approximation)

3. Parallel segment detection and fusion




Hypothesis generation

Some results of the parallel segments detection



Hypothesis generation

- The resulting segment is only a part of the shaft

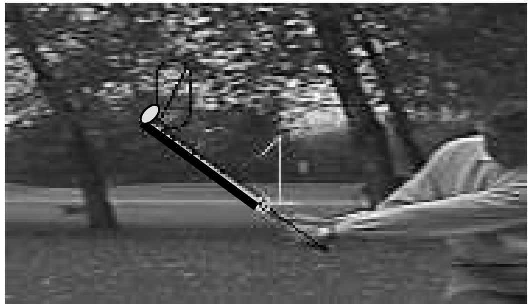
  Search for the shaft end-points

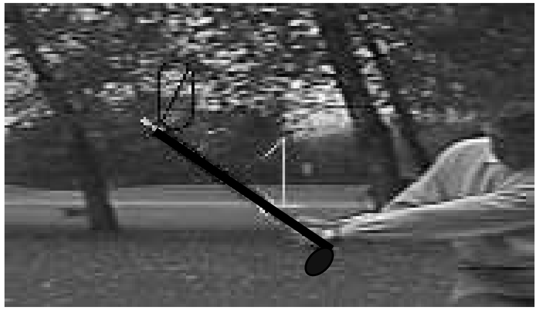
- Looking for the club head in the moving-object mask (as the last white point)

- Looking for the “hands” in the color image

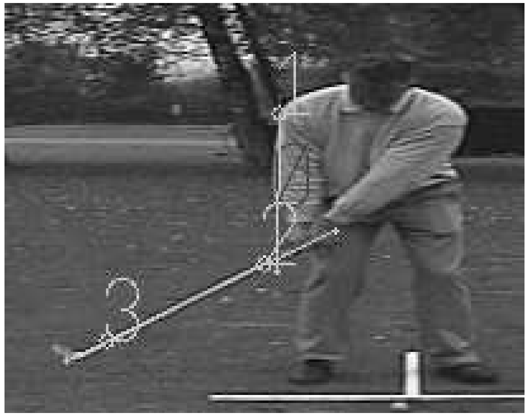
Hypothesis generation

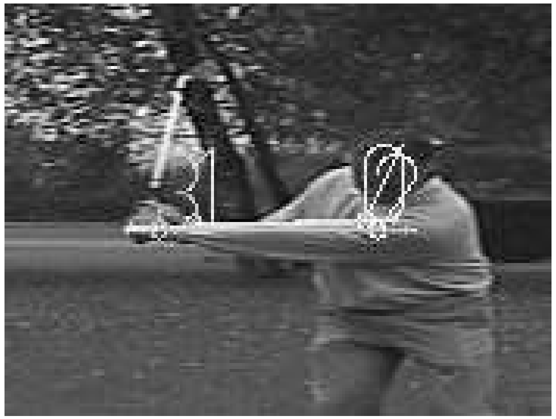
- Results
  - In this example, two hypothesis:





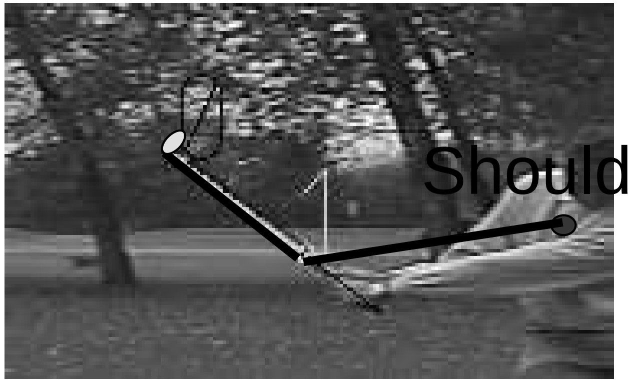
  - Others results:



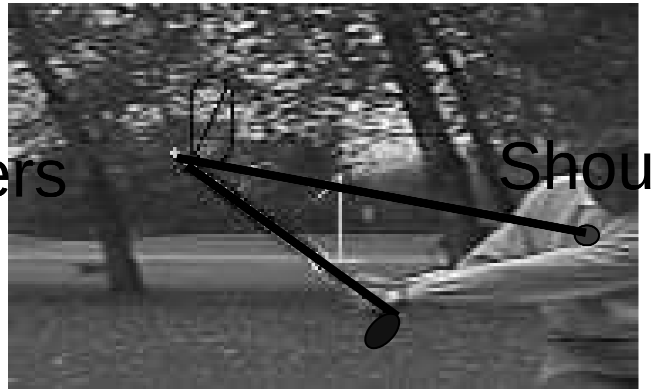


Hypothesis generation

- Heuristics for removing some hypotheses
  - Given a 2D point somewhere between the golfer shoulders, we can remove some physically impossible hypotheses



Possible

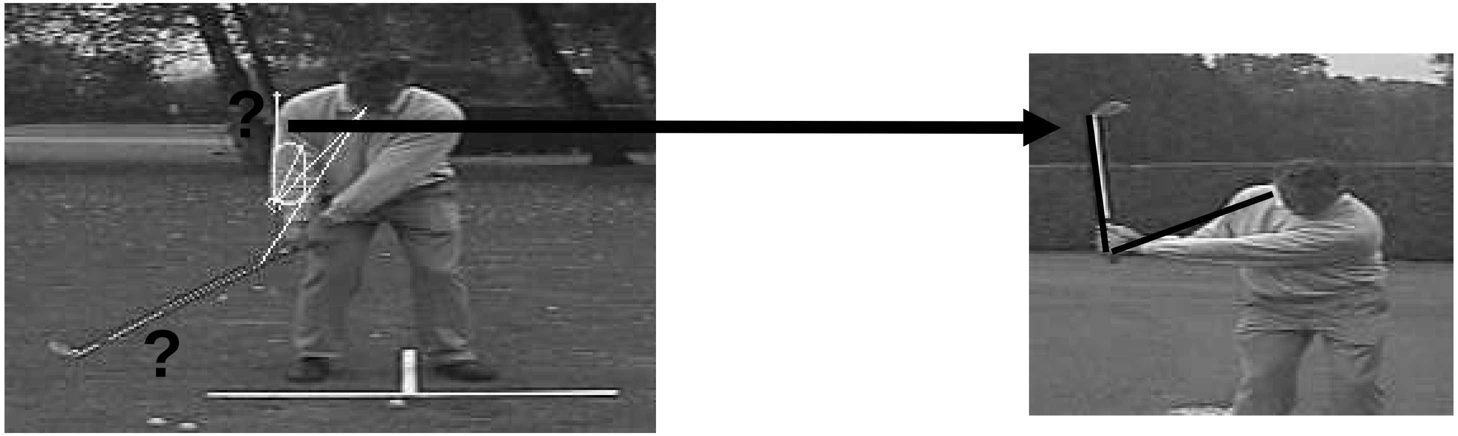


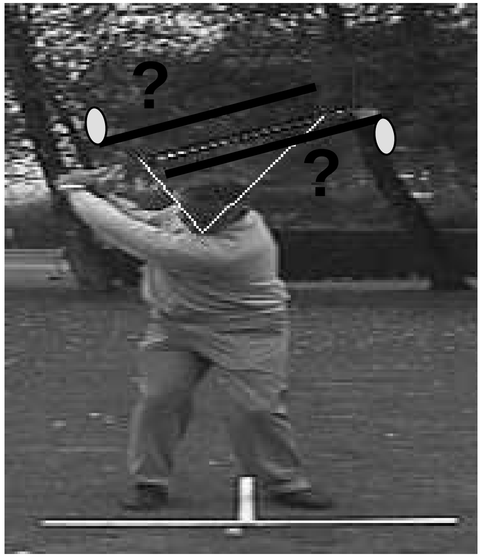
Impossible

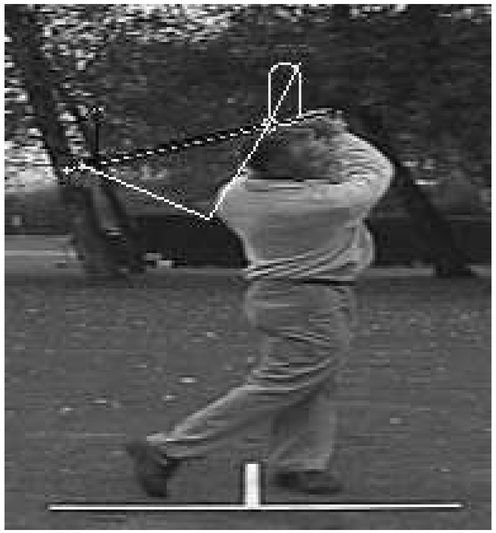
- No accurate position for this point needed
- Can be easily provided by the user

Why we still need a tracking algorithm ?

- Some wrong hypotheses can not be removed:







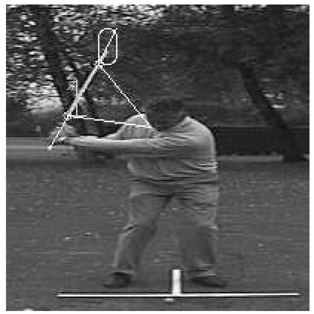
Tracking

- Many visual tracking techniques have been proposed in the computer vision literature.
  - Data Association approaches (MHT)
  - ConDensation

  - Based on recursive motion models: $X_{t+1} = f(X_t)$
  - Difficult to consider a specific motion such as a golf swing.
  - Suffer from a lack of robustness for practical applications when:
    - Frequent mis-detections
    - Large acceleration
    - Abrupt motion changes

New tracking algorithm

- Idea:
  - Take into account previous frames + next frames
  - Consider the detections in these frames to locally estimate the club shaft motion

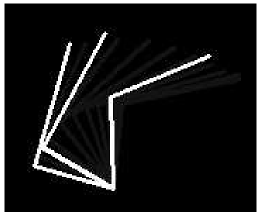
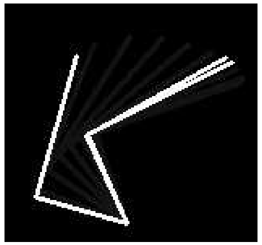
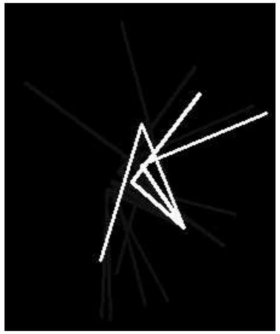
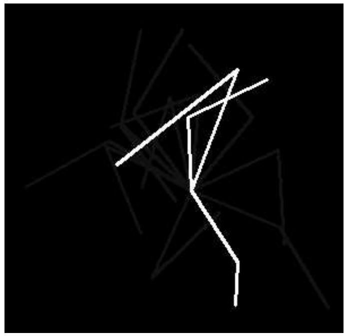


New algorithm

MLESAC applied to tracking:

- Choose randomly 3 frames (in the previous and next frames)
- Choose randomly one detection in these frames
- Compute the shaft motion assuming a locally constant acceleration
- Estimate the shaft position in the previous and next frames

  Several examples:

- Compute the support of the predicted motion *i.e.* the number of frames where there is a detection near the predicted position

- Repeat and keep the shaft motion with the maximum likelihood

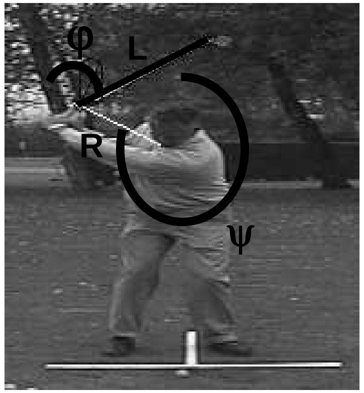
Deals easily with mis-detections and false-alarms

Robust motion estimation

Motion estimation

- **Parametrisation** of the shaft (double-pendulum model):

  s = [Shoulders, L, R, Ψ, φ]



- **Estimation**
  - From the three randomly selected shafts s_i, s_j, s_k,
  - assuming a constant acceleration for all the parameters,
    - we can predict the position, velocity and acceleration of the shaft

    $\mathbf{s}_{0}$ = [Shoulders $_0$, L $_0$, R $_0$, Ψ_0, φ_0] in the current frame:

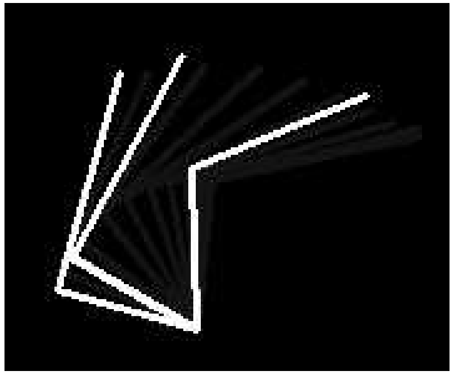
$$A = \begin{bmatrix} 1. & i & \frac{i(i-1)}{2} \\ 1. & j & \frac{j(j-1)}{2} \\ 1. & k & \frac{k(k-1)}{2} \end{bmatrix}$$

$$\begin{bmatrix} \Psi_0 \\ \dot{\Psi}_0 \\ \ddot{\Psi}_0 \end{bmatrix} = A^{-1} \begin{bmatrix} \Psi_i \\ \Psi_j \\ \Psi_k \end{bmatrix}$$

    - we can also predict the position in the other frames:

$$\Psi_n = \Psi_0 + n\dot{\Psi}_0 + \frac{n(n-1)}{2}\ddot{\Psi}_0$$

$$\ldots$$



Maximum Likelihood Estimation

- M_t motion at time t
- $Z_t = \{\mathbf{z}_{t-nB} \ldots \mathbf{z}_{t+nA}\}$ detection sets for frames t-n_B to $t+n_A$

$$M_t = \arg\max_{M} p(Z_t \mid M)\delta(M)$$

- Random sampling: $\tilde{M} = \arg\max_{M_S} p(Z_t \mid M_S)\delta(M_S)$

is an initial estimate of M_t, and refined using all the correct detections

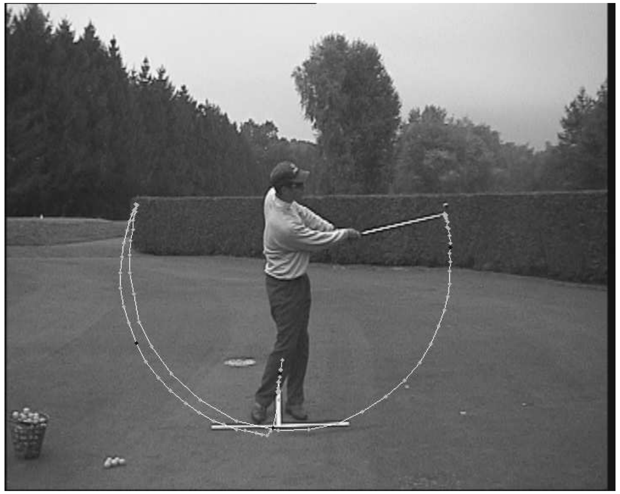
$$p(Z_t \mid M_S) = \prod_{i=n_B}^{+n_A} \underbrace{p(\mathbf{z}_{t+i} \mid y_{t+i,S})}_{\text{Classical observation model}}$$

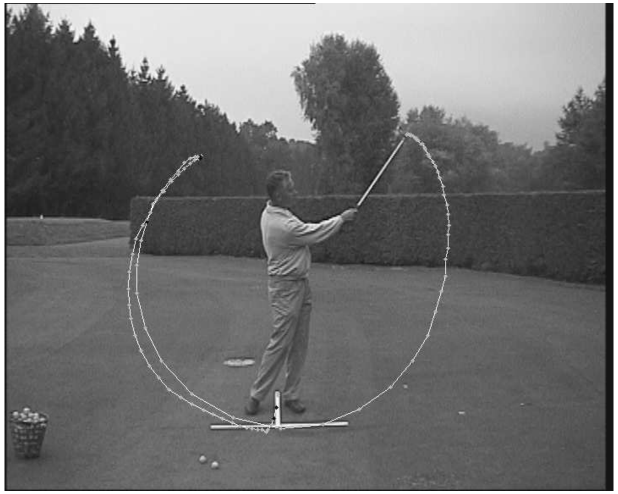
Advantages

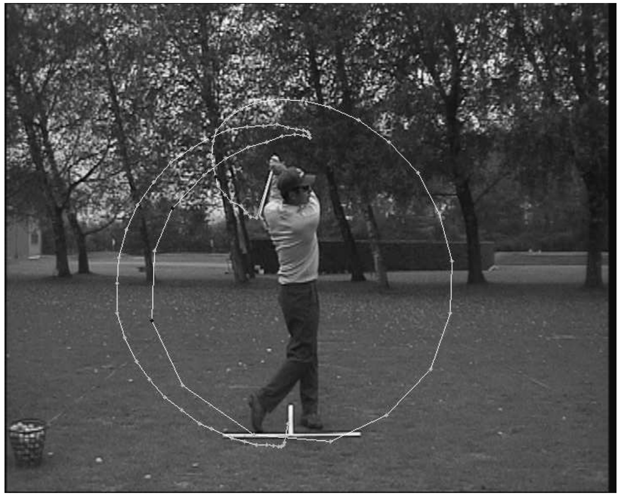
- The shaft position can be estimated when it is not detected, with very good accuracy:

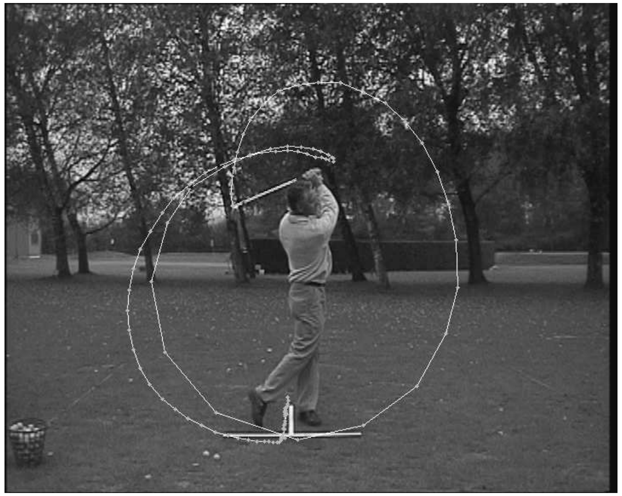
- Using the next frames makes the tracker:
  - More robust
  - More accurate
  - **Almost Automatic!**

Other results
same parameters









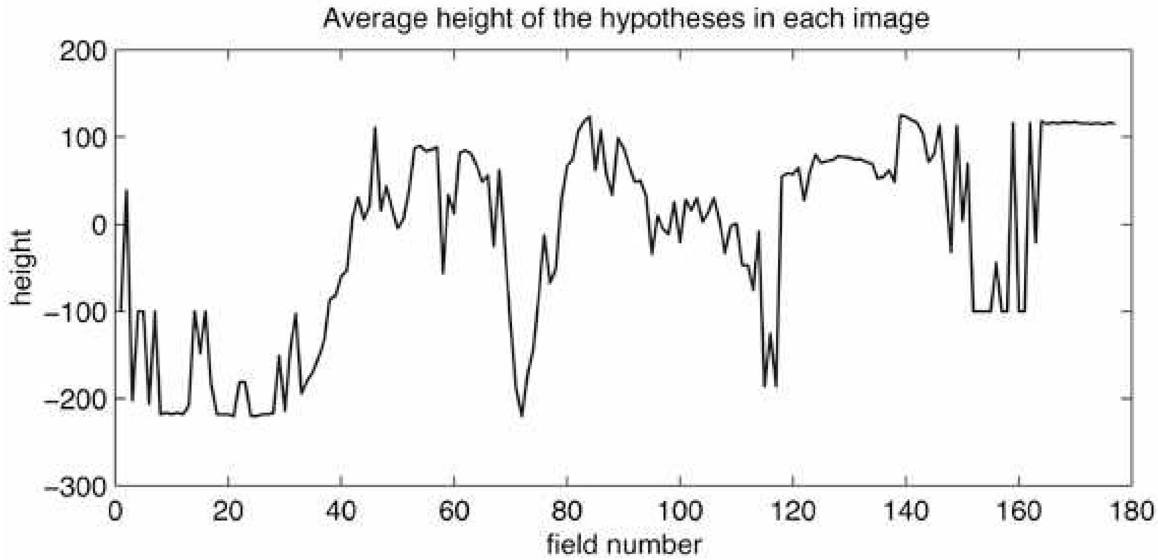
Making the tracker more robust

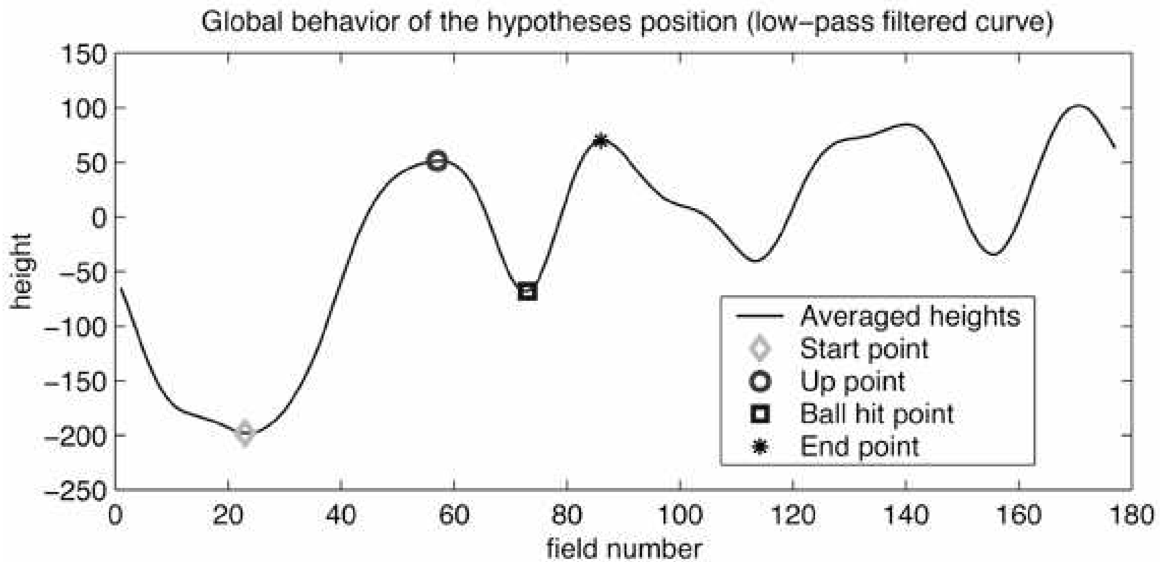
- Using a global motion model

Temporal Segmentation of the Sequence

upswing

downswing





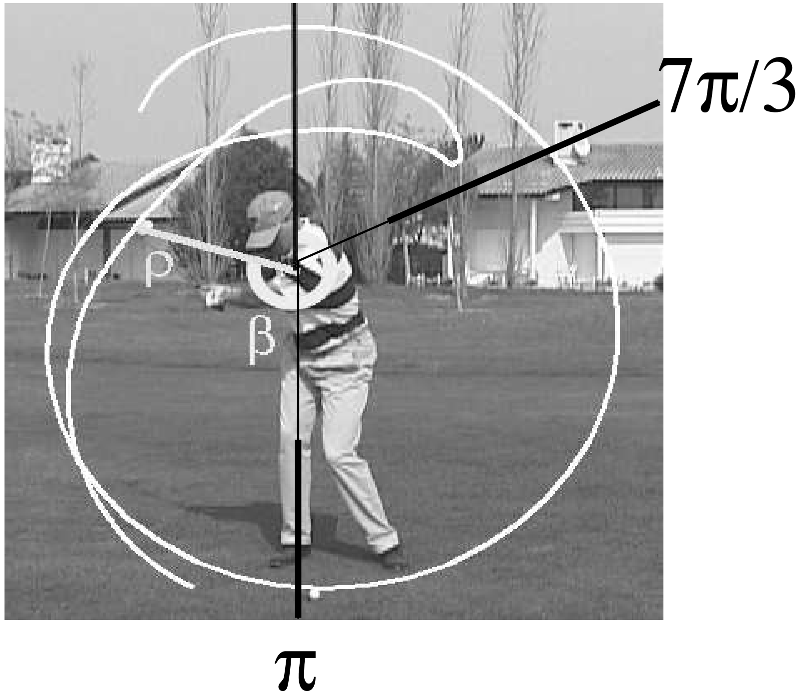
Trajectory Estimation

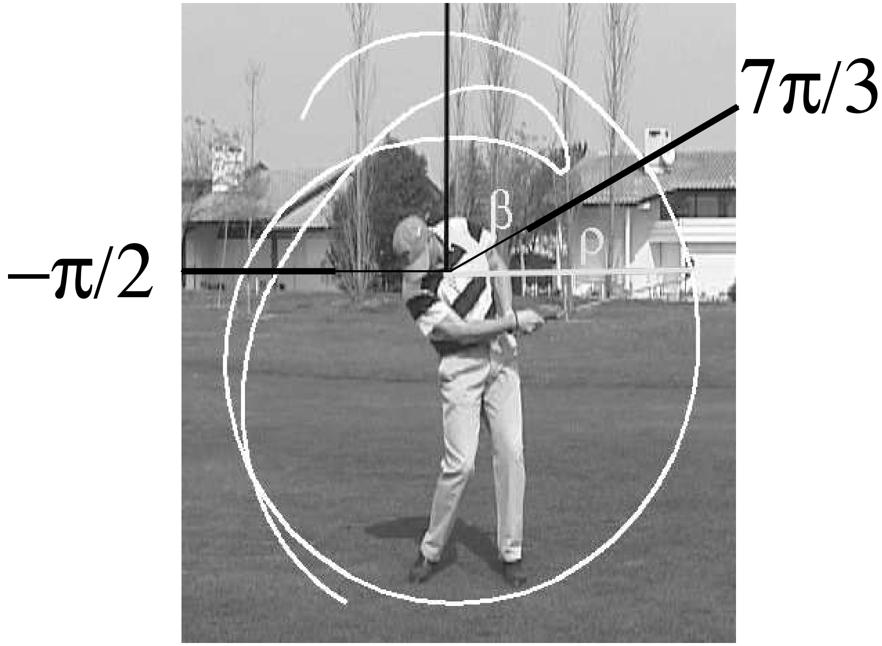
Separated global motions models for

– upswing

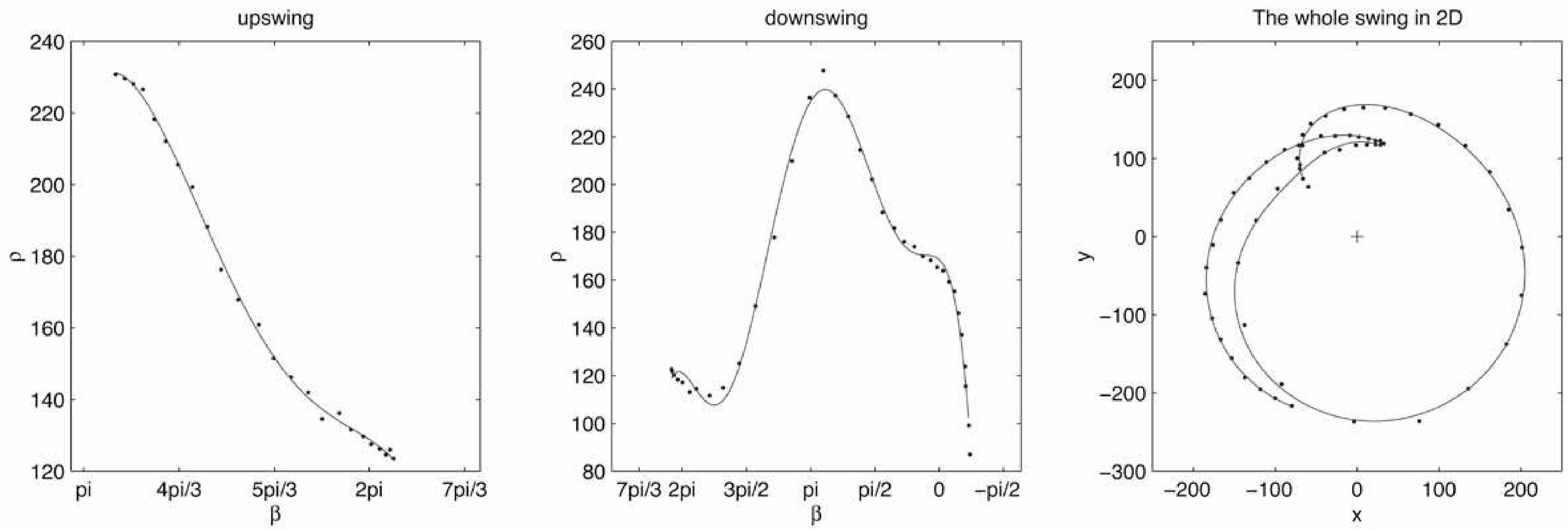
– downswing

expressed in polar coordinates system





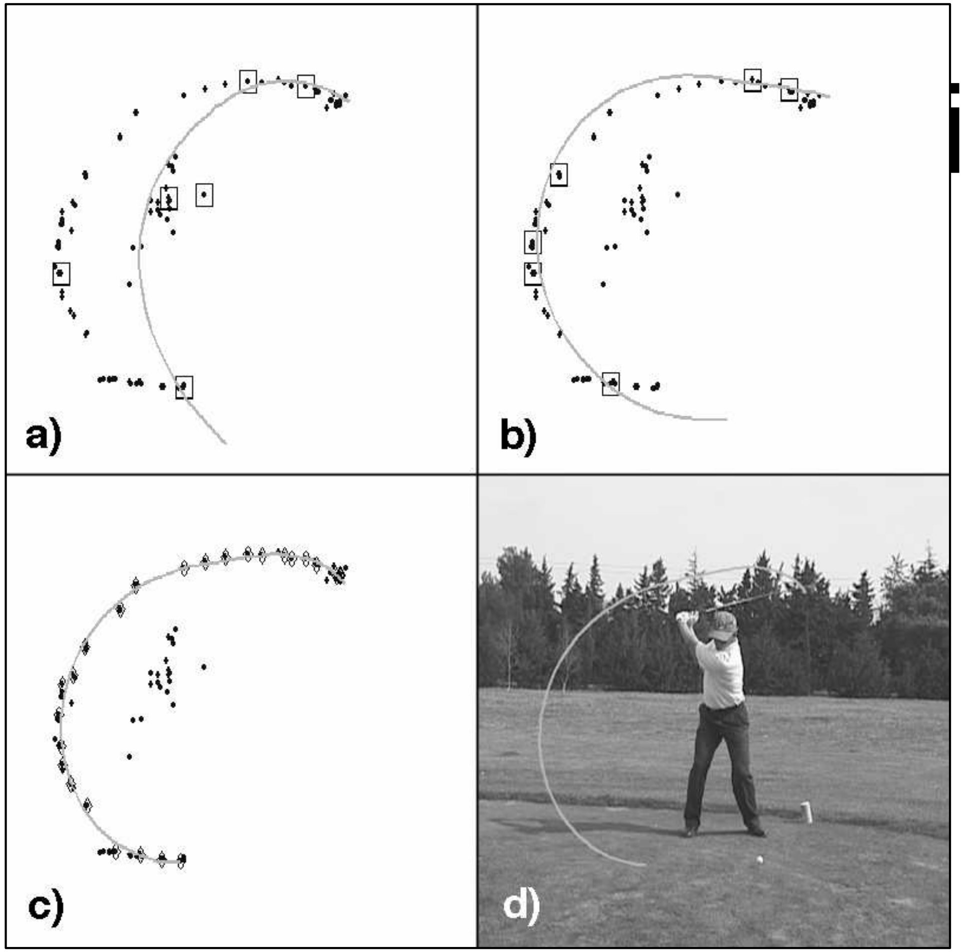
Simple polynomial functions of rather small degrees



Upswing
deg. 4 polynomial

Downswing
deg. 6 polynomial

imation



Same algorithm than before, with a global motion model

•Robust estimation (b)

•Refinement using the complete support (c)

Experimental results

Further Research

- Define a better global motion model (PCA ?)
- Analysis of the motion parameters

- Tracking of the golfer body