MV-Sports

A Motion and Vision Sensor Integration-Based Sports Analysis System

Cheng Zhang[1], Fan Yang[2], Gang Li[1], Qiang Zhai[2], Yi Jiang[2], Dong Xuan[1]
zhang.7804@osu.edu

[1]Dept. of Computer Science and Engineering, The Ohio State University
[2]DeepCode Robotics Co., Ltd.

Sports Analytics

Professional clubs: tactic and coaching

Live sports broadcasting: statistics

Common playfield: training

Sports robot: decision making

Sports Analytics

Professional clubs: tactic and coaching

Live sports broadcasting: statistics

Need Sports Analysis System

Common playfield: training

Sports robot: decision making

The Anatomy of the Sports Analysis System

Human

- Motor learning & Training
- Performance analysis
- Modelling & Simulation
- Game & Competition
- Multimedia & Visualization
- Instruction & Education …

Data

- Data acquisition
- Data processing
- Database system
- Data exchange
- Information & documentation
- Data management …

Infrastructure

- IT networks
- Communication
- Physical space
- Management service
- Distributed system
- ……

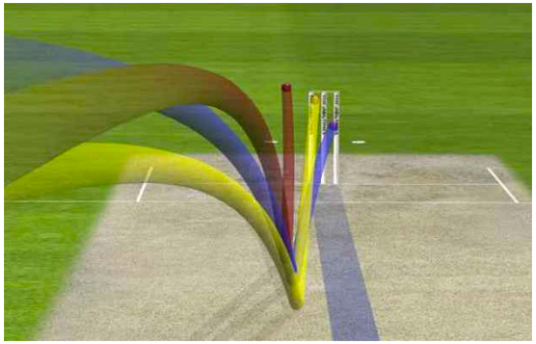
Human, data, infrastructure connection!

Today and the Future

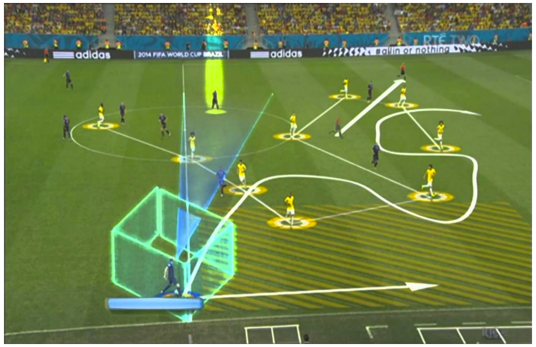
- Today things are controlled by human
  - Video records, offline bio-data analytics, etc.

- The future: accurate, real-time, and autonomous

Human-Centric:
Activity Analysis



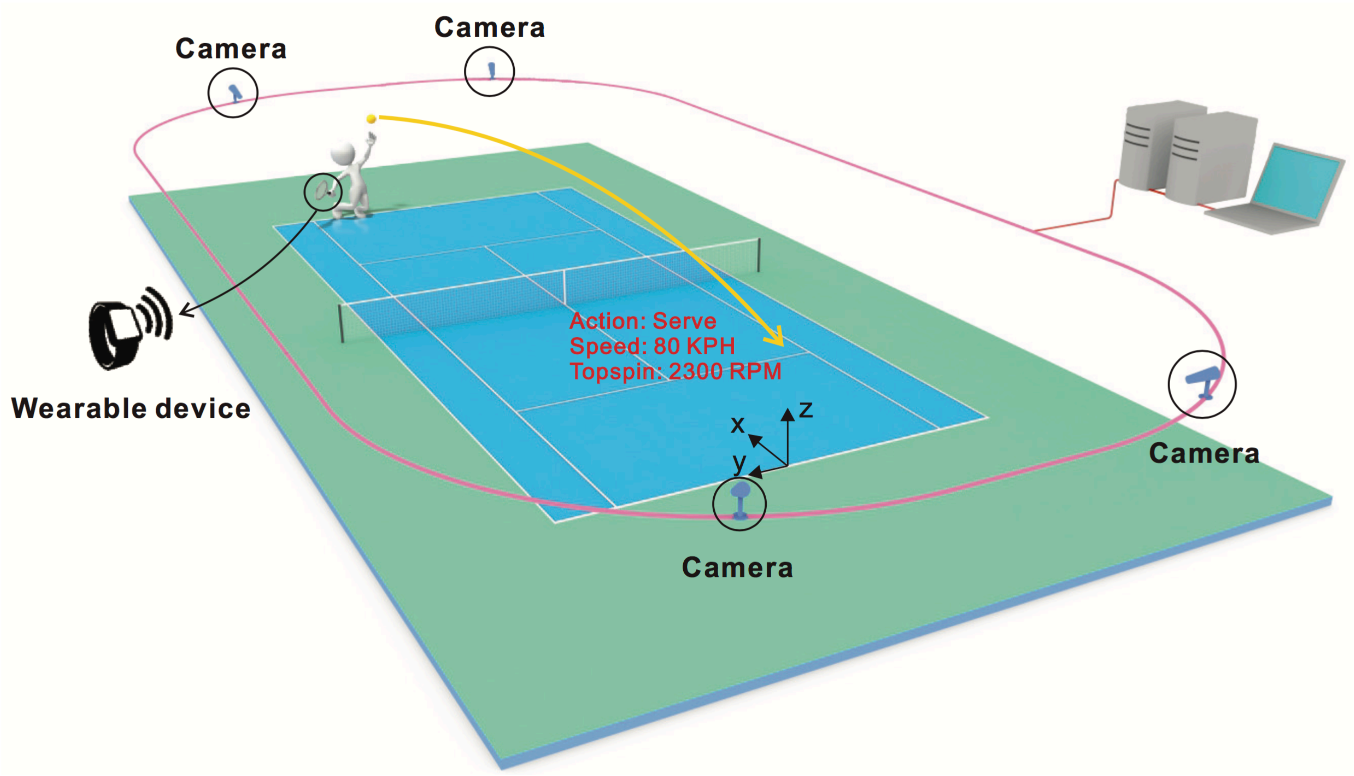
Fast Moving Object:
Ball Tracking



Overall Arbitrary:
Scene Understanding

Our Focus: Tennis as a Case Study
MV-Sports: System Setup

- **Task 1:** Player action recognition
- **Task 2:** High speed moving object tracking and prediction



Three Requirements

- Accurate and fine-grained analysis
- Real-Time
- Low-Cost and non-intrusive

This Paper Asks:

Can we build an affordable and easy-to-deploy system
(cheap sensors and COTS devices)
to achieve accurate and real-time sports analysis?

Existing Solutions Not Suitable

- **Vision sensor-based**
  - Hawk-Eye
    - Expensive ($500k) and not real-time
- **Wireless sensor-based**
  - Wearable device (Zepp, Sony, etc)
    - Not robust and simpleness
  - Embedded wireless sensor (*iBall, NSDI'17*)
    - Too intrusive for broad adoption

Hawk-Eye: High-end cameras

Zepp: Wearable IMU

iBall: IMU+UWB

Core Opportunities

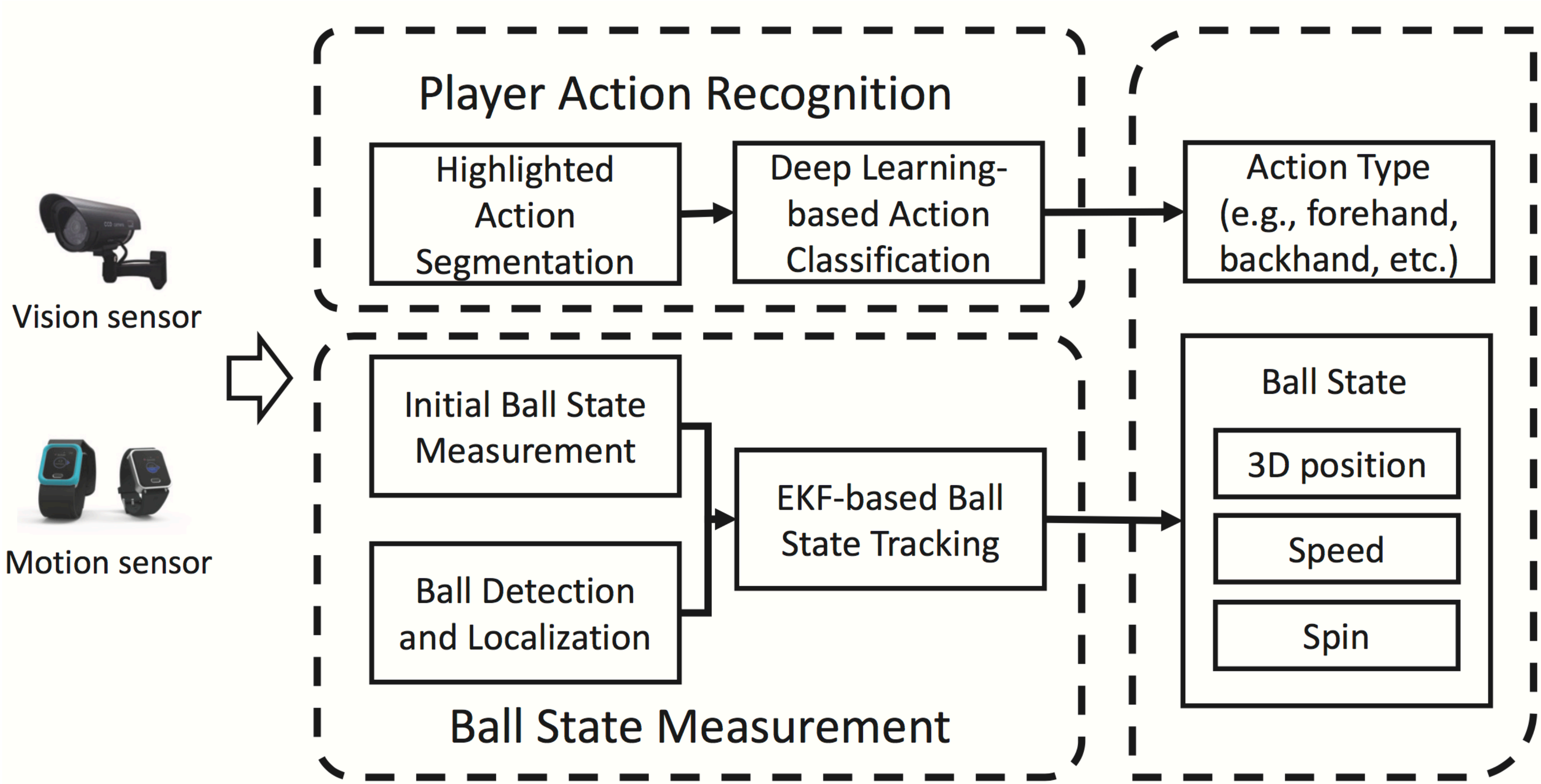
- Observations
  - A majority of existing systems for sports analytics are built with cameras OR wearable devices
  - The motion data (M) is more lightweight and digital quantized
  - The visual data (V) is more precise and intuitionistic
- Key ideas
  - Breaking cost barrier and bridge sensor gap
  - The complementary characteristics of the motion and vision sensors

Core Opportunities

MV-Sports: Sports Analytics based on Motion and Vision Signal Sensing

- Observations
  - A majority of existing systems for sports analytics are built with cameras or wearable devices.
  - The motion data (M) is more lightweight and digital quantized
  - The visual data (V) is more precise and intuitionistic
- Key ideas
  - Breaking cost barrier and bridge sensor gap
  - The complementary characteristics of the motion and vision sensors

MV-Sports: Overview



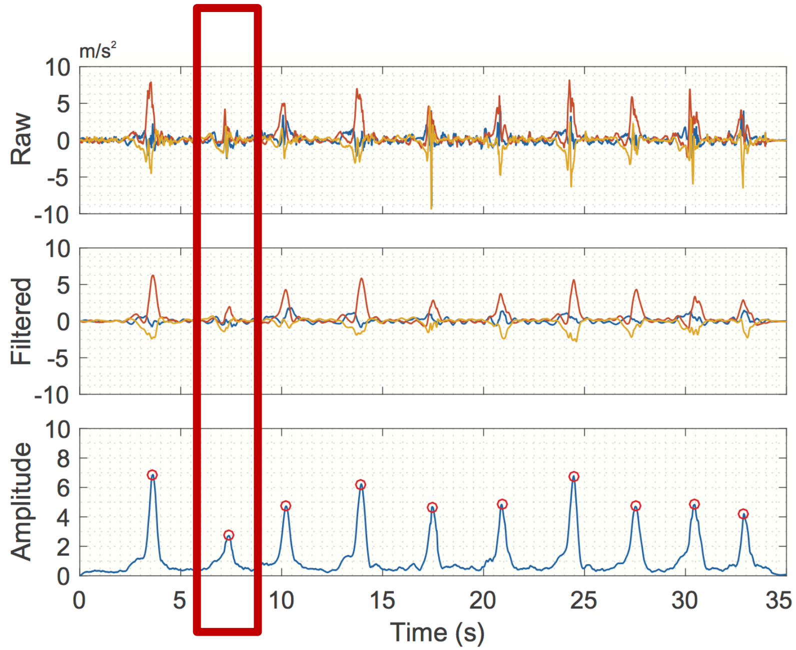
Task 1: Player Action Recognition

- **Goal**
  - To identify five typical tennis shots, i.e., serve, forehand topspin, forehand slice, backhand topspin, and backhand slice
- **Challenges**

  - Conventional video highlight approaches are too heavy!
  - M and V data alone cannot classify action type well!

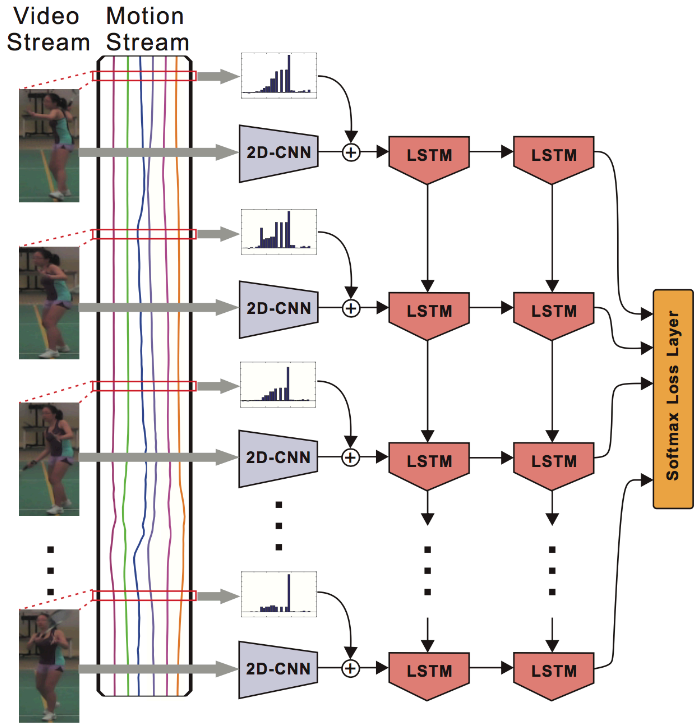
Solutions: Action Segmentation

- How to efficiently highlight key actions during exercise without pain?
  - Leverage motion signals to quickly highlight the meaningful content for action detection, avoiding redundant processing
  - Use an adaptive thresholding strategy in motion data streams

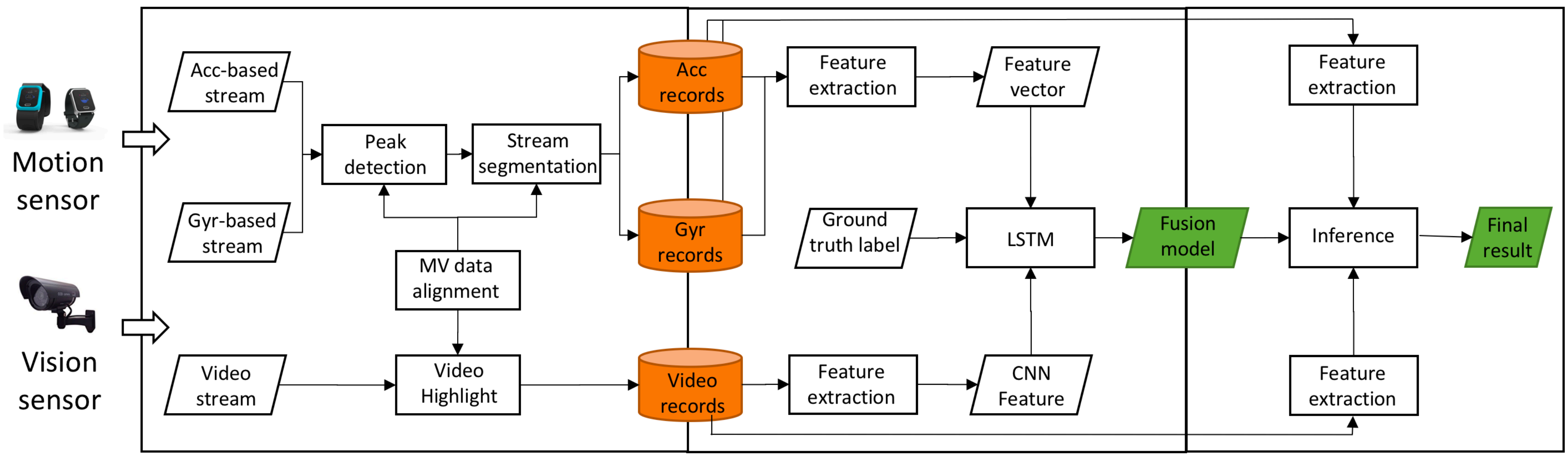


Solutions: LSTM-base MV Integration

- How to fuse and integrate the advantages of motion and vision sensing signals?
  - Integrate the highlighted M and V data via a uniform neural network to train an end-to-end model for classification



Player Action Recognition: Pipeline



Raw Sensor Data Stream

Offline Learning

Online Inference

Task 2: Ball State Measurement

- **Goal**
  - Measure all state includes tennis ball's position, speed, and spin, and it changes when the ball is flying or bouncing

- **Challenges**

  - The ball motion is affected by many forces such as gravity and air drag
  - The ball state is variable and we need to continuously track it during the ball's flight

Two Preliminaries

- **Ball motion physics-based**
  - Initial state when the ball is hit by a racket should be given
  - Analyze the physical process and calculate ball state at each timestamp

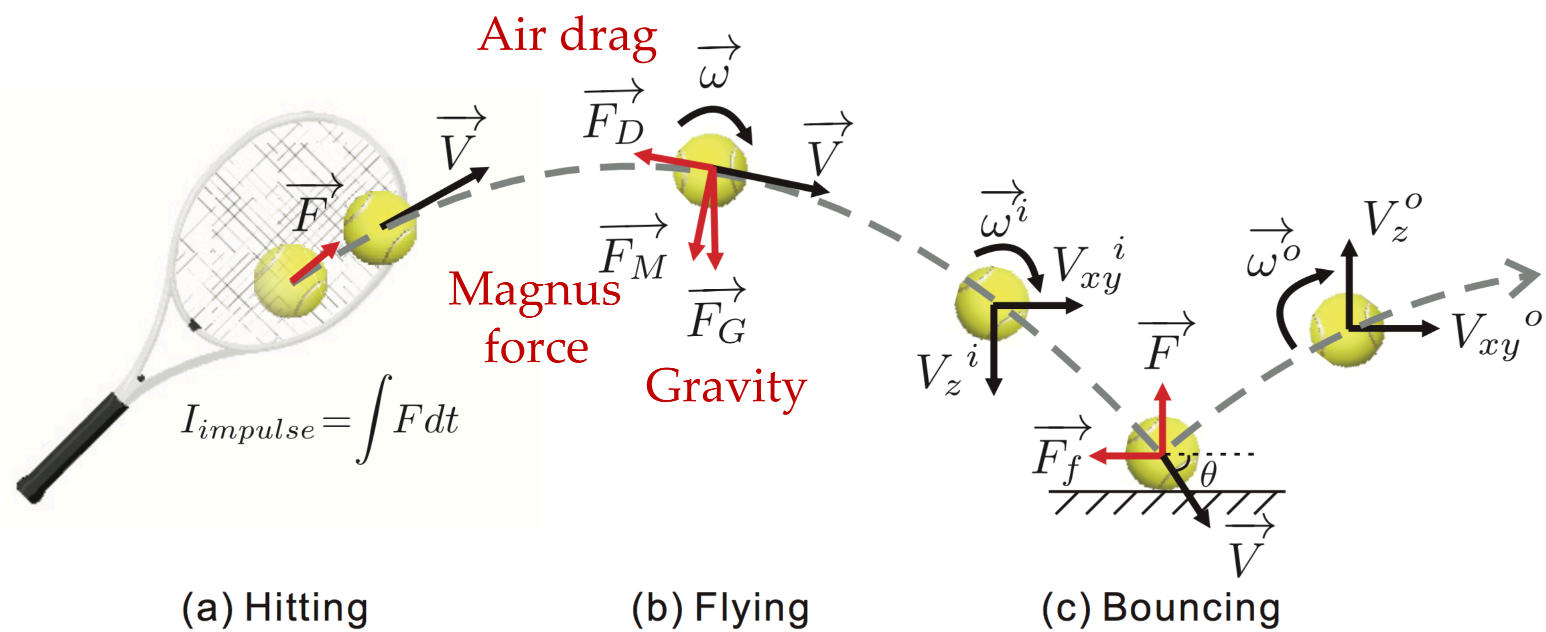
  - Heavily affected by the accuracy of the initial state as all following states are calculated from it
- **Purely vision-based**
  - Continuously track the ball's 3D location during its flight

  - Very sensitive to environmental factors such as lighting conditions

Our Solution

- Two steps based on the integration of motion and visual sensing by formulating into an extended kalman filter (EKF) problem
  - Step1: Determine the roughly accurate ball state at the initial timestamp based on motion sensing of the player's racket
  - Step2: Adopt both ball state tracking methods mentioned above and combine their results optimally to produce accurate and robust tracking result

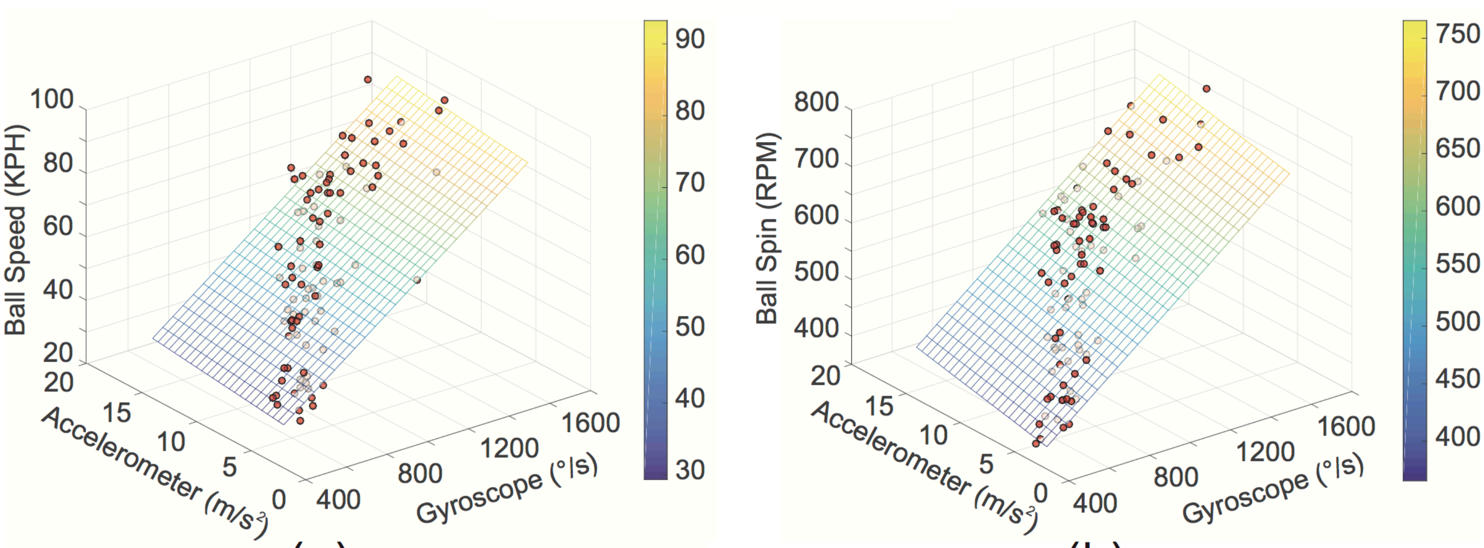
Ball Physical Process



Initial Ball State Measurement: Hitting

- Initial ball state is determined by the impulse of the racket when it hits the ball
- When the player waves the racket and hits the ball, motion sensors capture the force of the hit, which indicates the initial ball state

$$I_{impulse} = \int F dt$$



(a) (b)

Extended Kalman Filter (EKF)

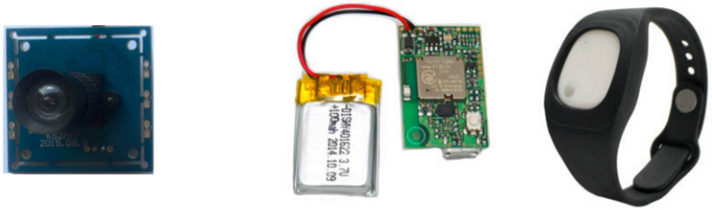
- EKF is the nonlinear version of the Kalman Filter which linearizes about an estimate of the current observed and predicted positions
  - Observation: visual-positioning-based position
  - Prediction: ball-motion-physics-based position

Algorithm 2 EKF-based Tracking at Timestamp t

1: $\overrightarrow{S^{t|t-1}} \leftarrow \overrightarrow{f}(\overrightarrow{S^{t-1}})$ based on Eq. (4)
2: $Sgm^{t|t-1} \leftarrow GSgm^{t-1}G^T + R$
3: $K \leftarrow Sgm^{t|t-1}H^T(HSgm^{t|t-1}H^T + Q)^{-1}$
4: $\overrightarrow{S^t} \leftarrow \overrightarrow{S^{t|t-1}} + K(\overrightarrow{O^t} - H\overrightarrow{S^{t|t-1}})$
5: $Sgm^t \leftarrow (I - KH)Sgm^{t|t-1}$
6: **return** $\overrightarrow{S^t}, Sgm^t$

System Implementation

- Motion part
  - Wearable motion tracking device purchased from mbientlab
  - Accelerometer and gyroscope data
- Vision part
  - Two cheap USB cameras with 1280 x 720 resolution and 60 FPS
- Back-end
  - PC with NVIDIA Tesla M40 GPU

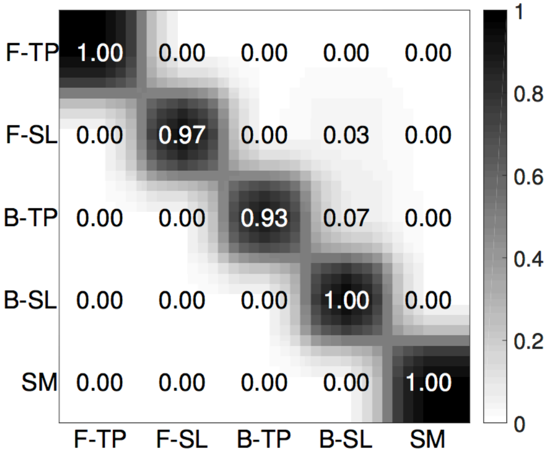


(a) Camera sensor (b) Wearable motion sensor

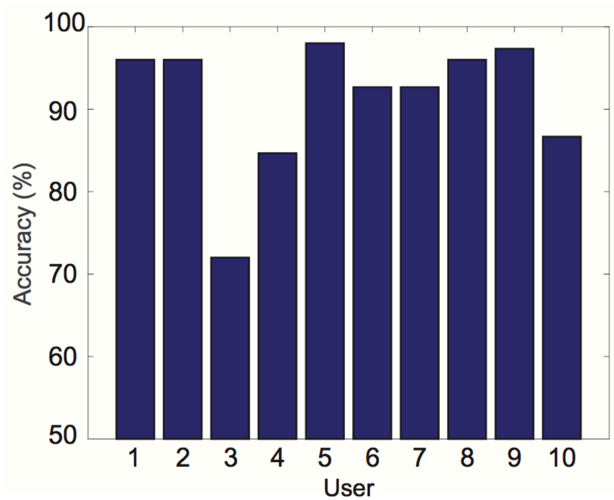
- Ground truth
  - Bushnell-101921 velocity speed gun to measure ball speed
  - High-end cameras with 240 FPS to measure ball spin

Evaluation on Player Action Recognition

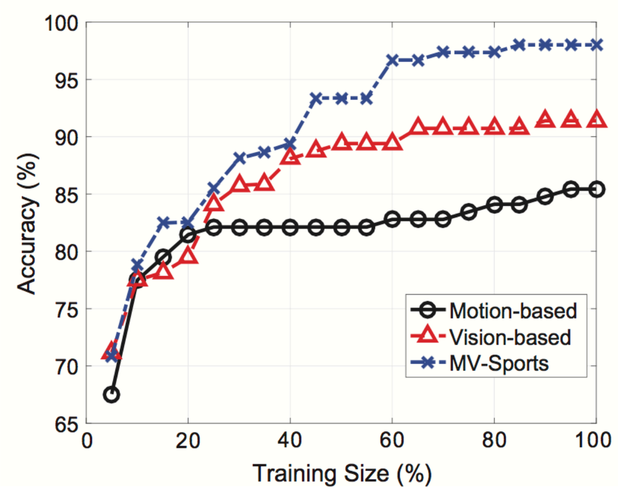
- 10 participants, 5 typical actions in tennis, 50 samples for each action
- Results from 2,500 video clips and motion samples achieve average classification accuracy of 98%, which outperforms conventional pure motion-based or pure vision-based methods



Overall confusion matrix



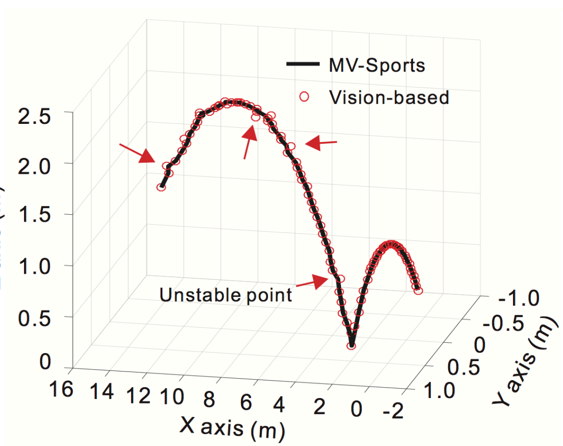
Accuracy of unseen users



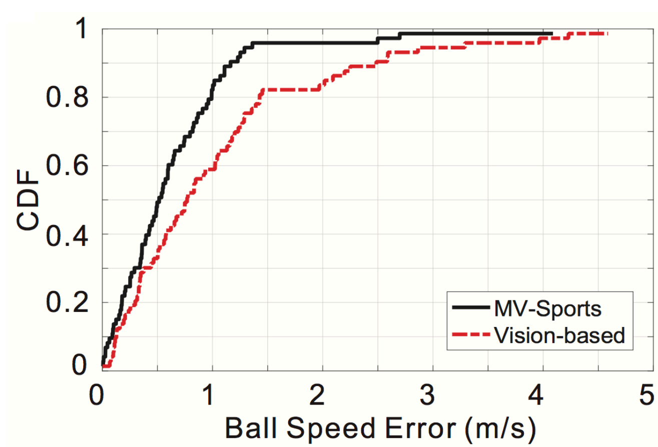
Accuracy of different training data size

Evaluation on Ball State Measurement

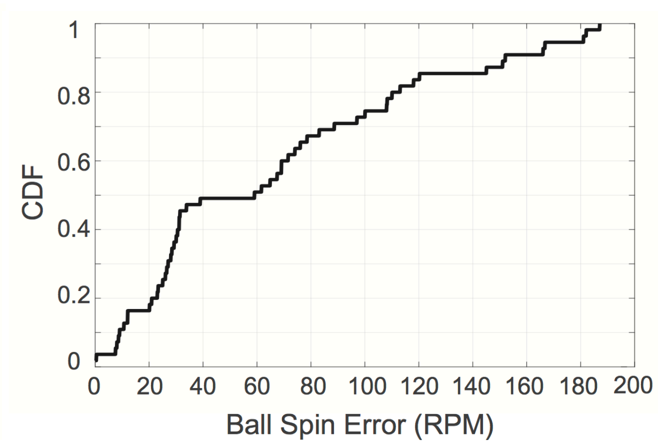
- 100 ball trajectories with different initial speeds and spins
- Compare MV-Sports with pure vision-based approach
- The error is defined as the average difference between the evaluated approach and the ground truth



Ball position tracking



CDF of ball speed error



CDF of ball spin error

Average Processing Time

Procedure	MV-Sports (ms/frame)
Highlight Segmentation	0.2
Player Localization	5.0
Feature Extraction	16.2
Online Action Classification	1.2
Sum	22.6

TABLE I
AVERAGE PROCESSING TIME OF EACH FRAME IN DIFFERENT STAGES OF ACTION RECOGNITION

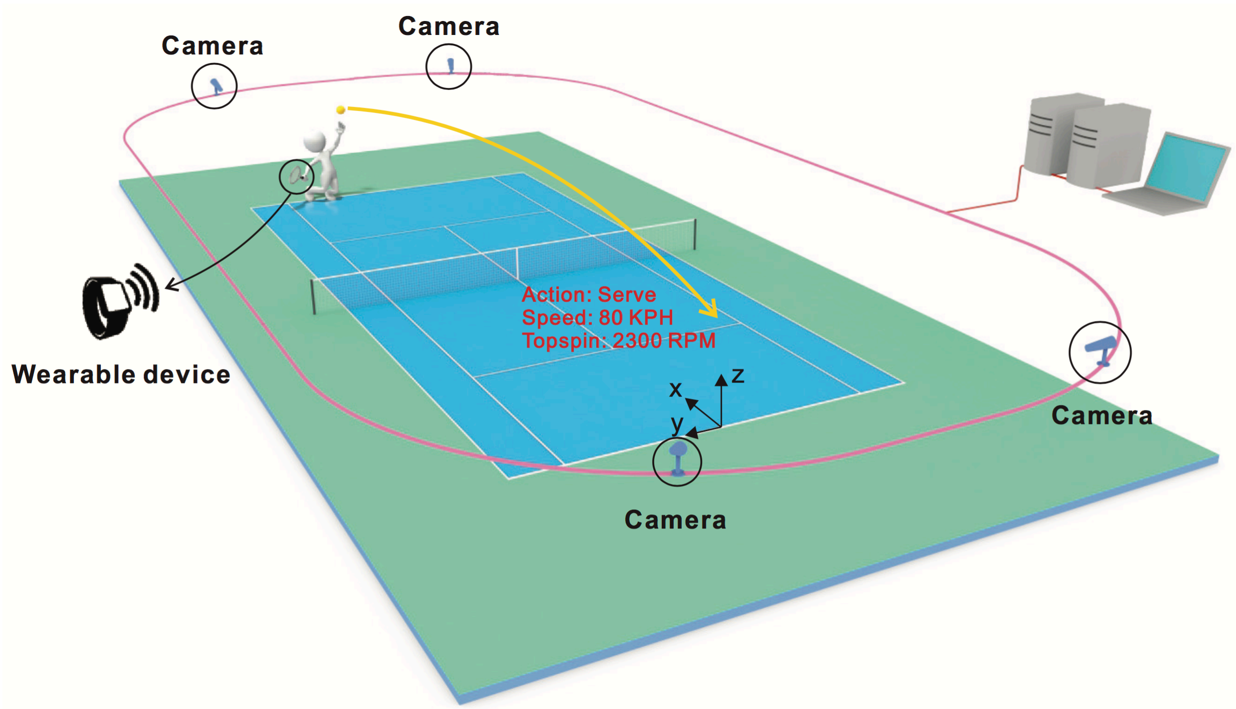
Procedure	MV-Sports (ms/frame)	Vision-based (ms/frame)
Processing Time	10.3	10

TABLE II
AVERAGE PROCESSING TIME OF EACH FRAME OF BALL STATE MEASUREMENT

Summary

- We design MV-Sports that leverages MV data to achieve satisfactory sports analysis driven by data and algorithms
- We implement MV-Sports in a tennis court with COTS devices with high accuracy and real-time performance

The fusion of design, strategy, technology and data to improve Smart Sports Analysis is important in our INFOCOM community!

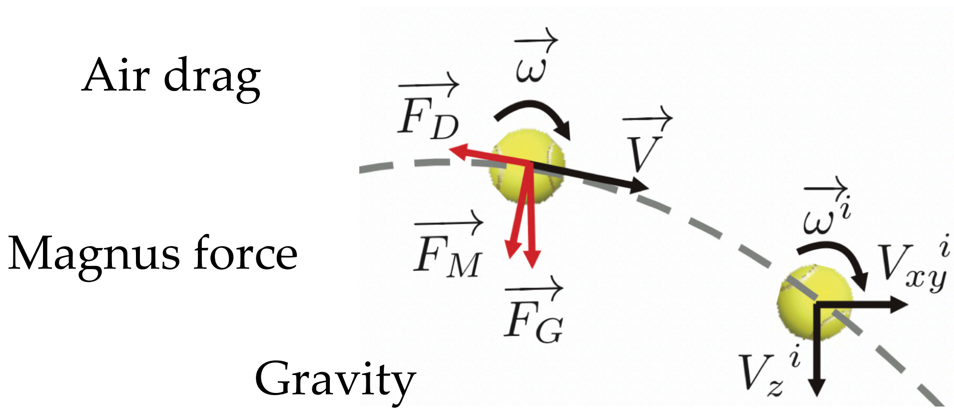


Thanks!
Questions?

Back up slides

Ball Motion Physics: Flying

- When the ball is flying, it is affected by three forces: Gravity, air drag, magnus force

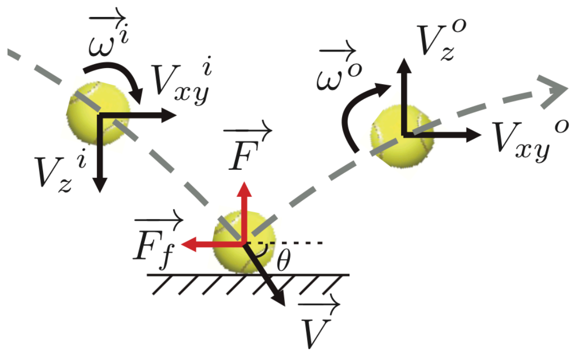


$$
\begin{cases}
\left\| \overrightarrow{{F_G}^t} \right\| = mg \\
\overrightarrow{{F_D}^t} = -\frac{1}{2}\rho C_D A \left\| \overrightarrow{V^t} \right\| \overrightarrow{V^t} \\
\overrightarrow{{F_M}^t} = \frac{1}{2\pi}\rho C_L D^3 \overrightarrow{\omega^t} \times \overrightarrow{V^t}
\end{cases}
\Rightarrow
\overrightarrow{f}(\cdot) =
\begin{cases}
\overrightarrow{P^t} = \overrightarrow{P^{t-1}} + \overrightarrow{V^{t-1}}\Delta t + \frac{1}{2}\frac{\overrightarrow{{F_1}^{t-1}}}{m}\Delta t^2 \\
\overrightarrow{V^t} = \overrightarrow{V^{t-1}} + \frac{\overrightarrow{{F_1}^{t-1}}}{m}\Delta t \\
\overrightarrow{\omega^t} = \overrightarrow{\omega^{t-1}}
\end{cases}
$$

Ball Motion Physics: Bouncing

- The friction between and ball and the ground as well as the elastic deformation are main reasons for the ball state change
  - Vertically

$$V_z{}^o = C_b V_z{}^i$$



  - Horizontally

Sliding friction
$$\begin{cases} V_{xy}{}^o = V_{xy}{}^i - C_f V_z{}^i(1 + C_b) \\ u^o = \frac{5(1+C_b)C_f}{2V_z{}^i} + u^i, \end{cases}$$

Rolling friction
$$\begin{cases} V_{xy}{}^o = \frac{5V_{xy}{}^i + 2u^i}{7} C_s \\ u^o = \frac{5V_{xy}{}^i + 2u^i}{7} C_t, \end{cases}$$

$$\Rightarrow \begin{cases} V_x{}^o = \frac{V_x{}^i V_{xy}{}^o}{V_{xy}{}^i} \\ V_y{}^o = \frac{V_y{}^i V_{xy}{}^o}{V_{xy}{}^i} \\ \omega^o = \frac{u^o}{R}, \end{cases}$$