



## Action Recognition for Simple and Complex Actions using Time of Flight Cameras

Serban Oprisescu, Constantin Burlacu, Vasile Buzoloiu<sup>1</sup>, Mihai Ivanovici<sup>2</sup>

<sup>1</sup>Image Processing and Analysis Laboratory, University "Politehnica" of Bucuresti, Romania

<sup>2</sup>Electronics and Computers, Transilvania University, Brasov, Romania

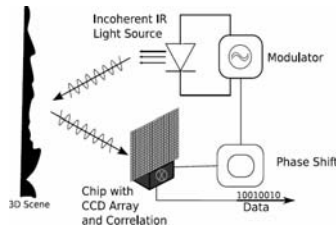
Email: {soprisescu, cburlacu, [buzoloiu@alpha.imag.pub.ro](mailto:buzoloiu@alpha.imag.pub.ro), [M.Ivanovici@gmail.com](mailto:M.Ivanovici@gmail.com)}

## Introduction



- Time-of-Flight (ToF) cameras:
  - active systems
  - record for each pixel both *intensity* and *distance* to the camera;
  - may be successfully used for
    - face detection
    - video surveillance
    - respiration monitoring etc.

## The ToF principle



- It works like a ... *BAT*
- the distance
  - proportional to the phase shift between the *direct* and the *reflected* waves

## Advantages / disadvantages

- Range of "sight" – depending on the integration time
  - "Myopic" behavior for small integration time intervals
  - "Hypermetropic" behavior for long integration time intervals
- Correlate the *intensity* and *distance* images for action recognition
- Segmentation may be easier on distance images (simple thresholding)

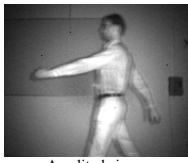
## Goals of the research

- Action Recognition using ToF cameras
  - *Distance* image processing
  - Movement representation by tracking body key points
  - Function characterization for discrimination, using features (e.g. total variation)

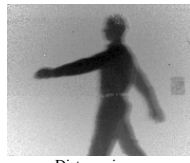
## Action Recognition – Our Approach

- Segmentation of the ToF *distance* image for silhouette extraction
- Skeletonization and Key Points extraction
- Tracking of the key points and computing their trajectories
- Feature extraction
- Action recognition based on computed features

### Silhouette extraction -> Skeleton -> Key Points



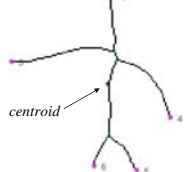
Amplitude image



Distance image



Segmented image



Skeleton with Key Points

Mihai Ivanovici IPCV'09 Las Vegas, Nevada, USA 7

### Task: discriminate between 6 actions



WALK



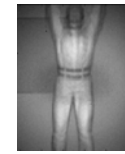
CARRY



RUN



BENT



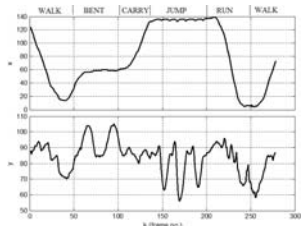
JUMP



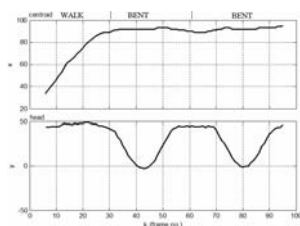
BOX

Mihai Ivanovici IPCV'09 Las Vegas, Nevada, USA 8

### Key Points Trajectories



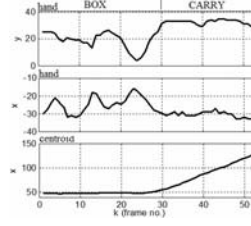
Centroid trajectories on x/y for multiple actions (film\_sr\_MR6)



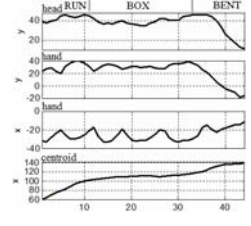
Walk followed by bent, centroid vs. head trajectories

Mihai Ivanovici IPCV'09 Las Vegas, Nevada, USA 9

### Key Points Trajectories



Box followed by carry, centroid vs. hand trajectories



Run, box, bent - centroid vs. hand & head trajectories

Mihai Ivanovici IPCV'09 Las Vegas, Nevada, USA 10

### Experimental results

- Show time! "walk" and "box"



Mihai Ivanovici



IPCV'09 Las Vegas, Nevada, USA

11

### Discriminating the trajectories

- How to discriminate between different movements?
  - Analyzing the complexity of the "signals"
    - Function characterization
    - Fractal analysis (fractal dimension)

Mihai Ivanovici

IPCV'09 Las Vegas, Nevada, USA

12

## Trajectories -> Features

1) The variation of the function:

$$\Delta f = f_{\max} - f_{\min}$$

2) The total variation  $v$  of the function:

$$V = \sum_{k=2}^N |f(k) - f(k-1)|$$

## Trajectories -> Features (II)

3) The real mean speed, computed as the mean of instantaneous speeds for each frame:

$$S_r = \frac{1}{N} \sum_{k=2}^N (f(k) - f(k-1)) / t$$

4) The absolute mean speed, computed as the mean of the absolute value of each instantaneous speed:

$$S_a = \frac{1}{N} \sum_{k=2}^N \left| \frac{f(k) - f(k-1)}{t} \right|$$

## Conditions defining the 6 actions

Action	Conditions			
Walk	$V_x > \text{thr}$	$V_{Hx} \gg \Delta_{Hx}$	$v_x \text{ avr}$	any y small
Carry	$V_x > \text{thr}$	$V_{Hx} \cong 0$	$v_x \text{ avr}$	
Run	$V_x > \text{thr}$	$V_{Hx} \gg \Delta_{Hx}$	$v_x \text{ big}$	
Box	$V_x \gg \Delta_x$	$V_{Hx} \gg \Delta_{Hx}$	$V_{vHx} \text{ big}$	
Jump	$V_x > \text{thr}$	$V_{Hy} > \text{thr}$	$v_y \text{ big}$	any x small
Bent	$V_x > \text{thr}$	$V_{Hy} > \text{thr}$	$v_y \text{ avr}$	

$V_x$  and  $V_y$  are the total variations of  $x(k)$  and  $y(k)$  respectively;  $V_{Hx}$ ,  $V_{Hy}$  are total variations of one key point on hand coordinates;  $v_x$ ,  $v_y$  are the speeds on corresponding coordinates;  $V_{vHx}$  is the total variation of the speed of hand key point on x;  $\Delta_x$  and  $\Delta_{Hx}$  are defined in (1).

## Conclusions

- Using ToF camera
  - Distance image processing
  - Straightforward and more efficient segmentation
- Trajectory characterization
  - Based on function features (e.g. total variation of a function)