Cloud Point Labelling in Optical Motion Capture Systems

– Thesis Dissertation –

Juan Luis Jiménez Bascones

Advisor: *Prof. Dr. Manuel Graña Romay*
Computational Intelligence Group
University of the Basque Country (UPV/EHU)

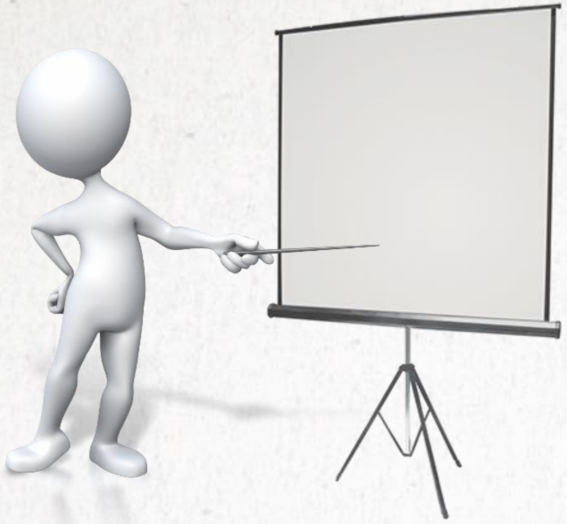
Donostia – San Sebastián – Feb 15, 2019

ABSTRACT

- Provide a broad view of the Mocap State of the Art
  - Applications
  - Technologies
- Optical mocap pipeline
  - Role played by marker labelling
- Main contribution: **brand new labelling algorithm** for Optical Mocap Systems
  - Machine learning approach: use of machine learning techniques vs. rule-based
  - Get rid of predictive information: labelling vs. tracking
- Assessment of the algorithm against a ground truth harvested on purpose

SUMMARY

1. Introduction
   - Motivation
   - State of the Art
   - Overview of an Optical Mocap workflow
2. Contribution
   - Problem statement
   - Correct labelling detection
   - Partial solvers
   - Solver ensemble
3. Results
4. Conclusions



1 INTRODUCTION

WELCOME TO MOCAP

1 INTRODUCTION

Motivation

State of The Art:

- Interest in Mocap
- Mocap technologies
- Labelling techniques

The Optical Mocap Pipeline

WHAT IS MOCAP

MOtion CAPture

- Set of methods and techniques
- Record and analyse the movement
- Fields of application: entertainment, medicine, sports, …

Target movement

✓ Complete human bodies
✓ Specific parts (facial, hands capture)
✓ Machinery, tools, environmental objects
✓ Animals

MOTIVATION

Optical marker labelling as a still open problem

- Professional background
- Problem hard to handle due to the fuzzy nature of both data and rules
- Scarce number of publications dealing with the issue
- Most of them from a kinematic-predictive model approach
- Existing algorithms hard to tune

Something should be done!

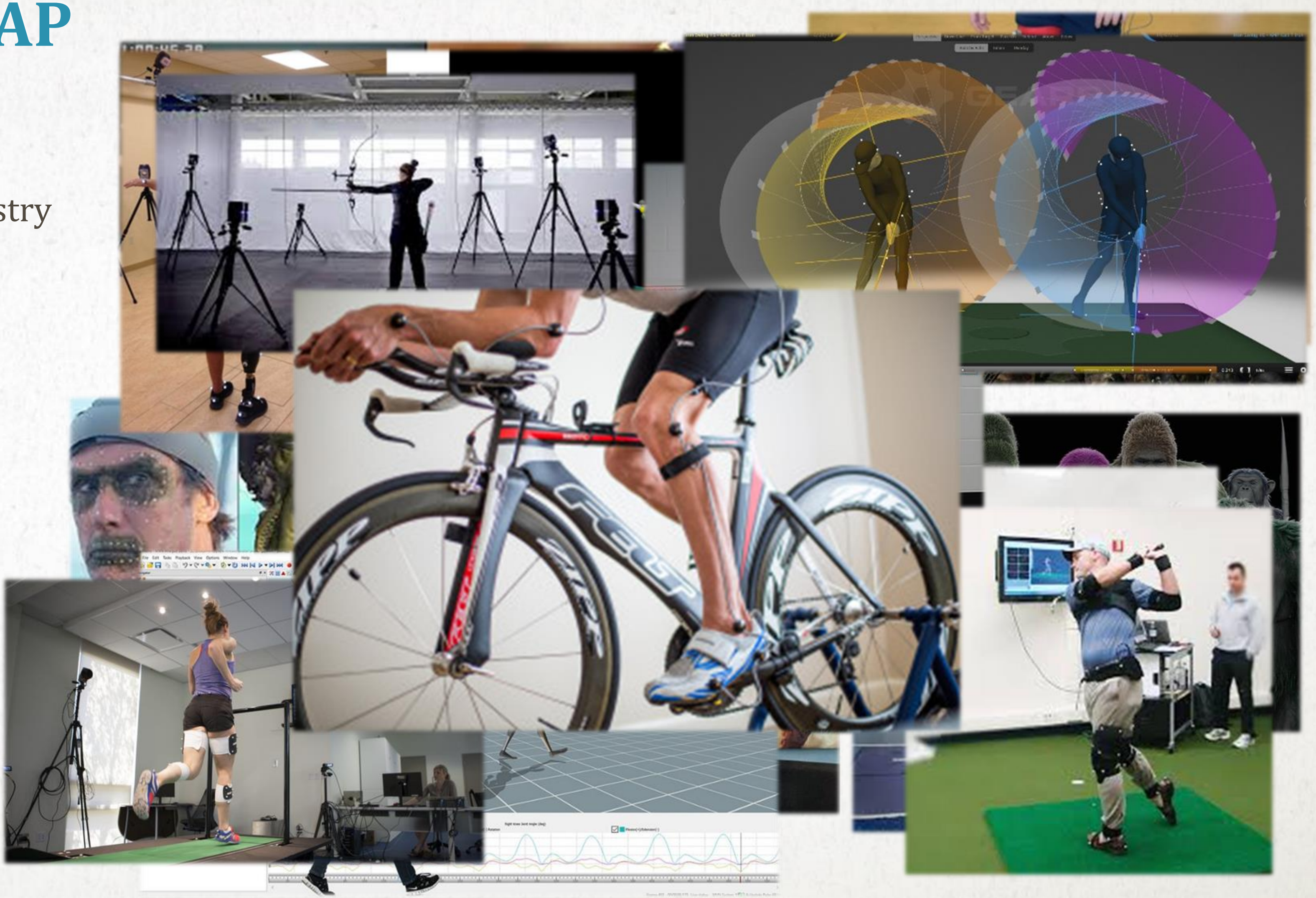
INTEREST IN MOCAP

Applications

- Movies, TV and gaming industry
- Medical analysis
- Ergonomics
- Sport analysis
- Activity recognition
- …

Research in Mocap

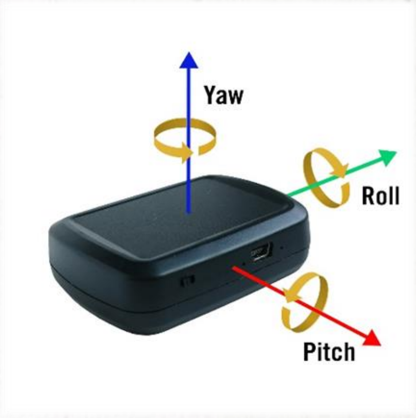
- System assessment
- More affordable solutions
- New detection methods
- …



MOCAP TECHNOLOGIES (I)

Wearables

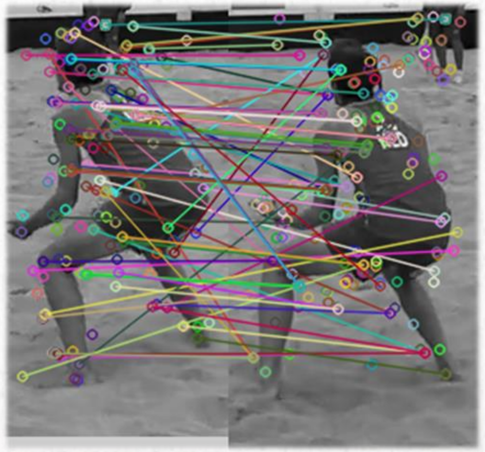
- Electromechanical
- Electromagnetic
- IMUs
  - Gyros, acc, mag;
  - Relative measurement
  - Drifting
- Alternative devices
  - Radio frequency
  - Optical fiber
  - Flexible nanomaterial

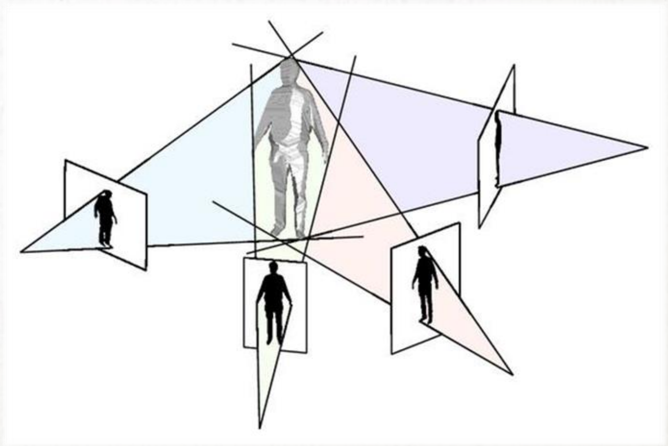


MOCAP TECHNOLOGIES (II)

Markerless Optical Systems

- Range image
- Image-matching
  - SIFT, SURF
- Silhouette extraction

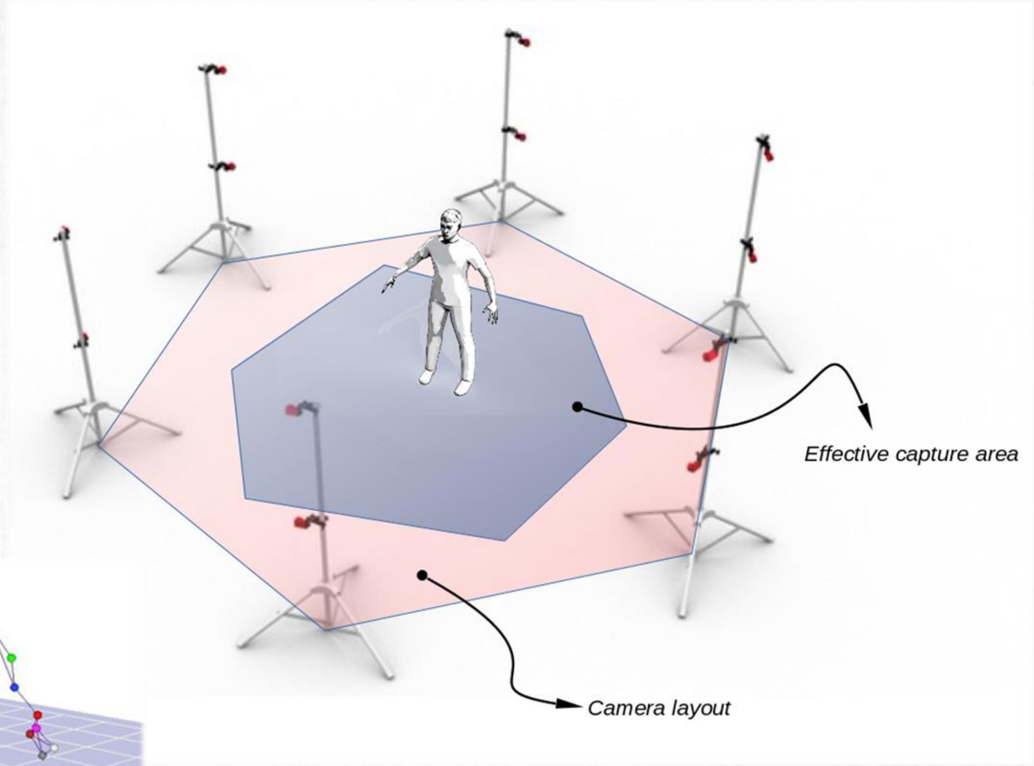


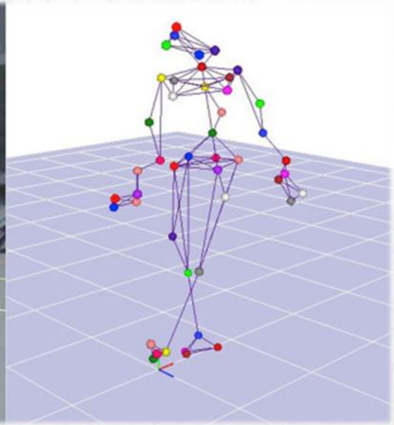


MOCAP TECHNOLOGIES (III)

Marker-based Optical Systems

- Passive markers
- IR Lightning
- Calibrated cameras
- Accurate

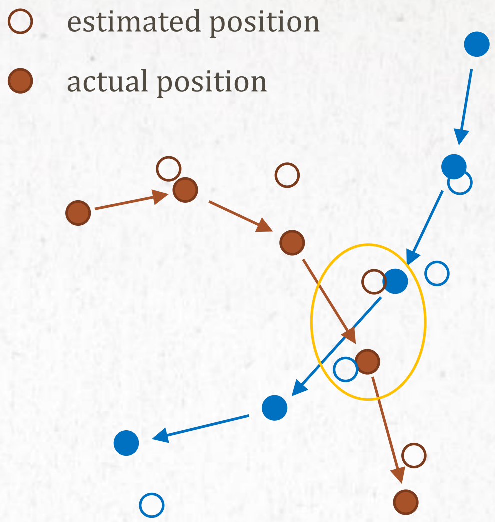


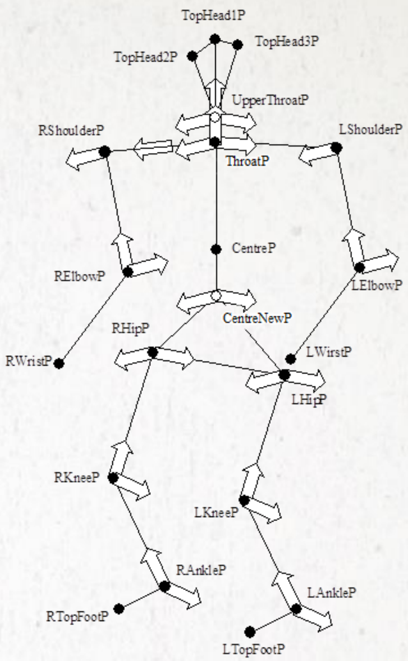


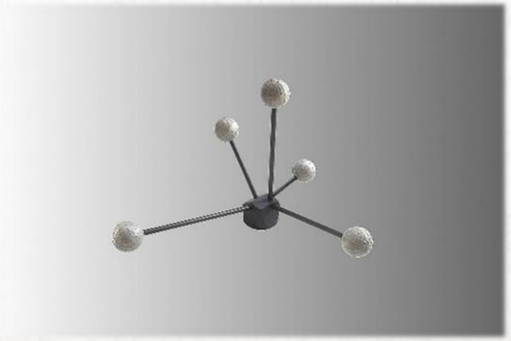
TRACKING TECHNIQUES

Usual approaches

- Trajectory continuity along the time
- Geometric restrictions: pairwise distances, means and standard deviations
- Kinematic constraints: underlaying bone-joint mechanism
  - DOFs as unknowns
  - flesh misleading effect
  - sometimes use true rigid body targets
- Target function and iterative minimisation methods
- Hard to tune: tolerances



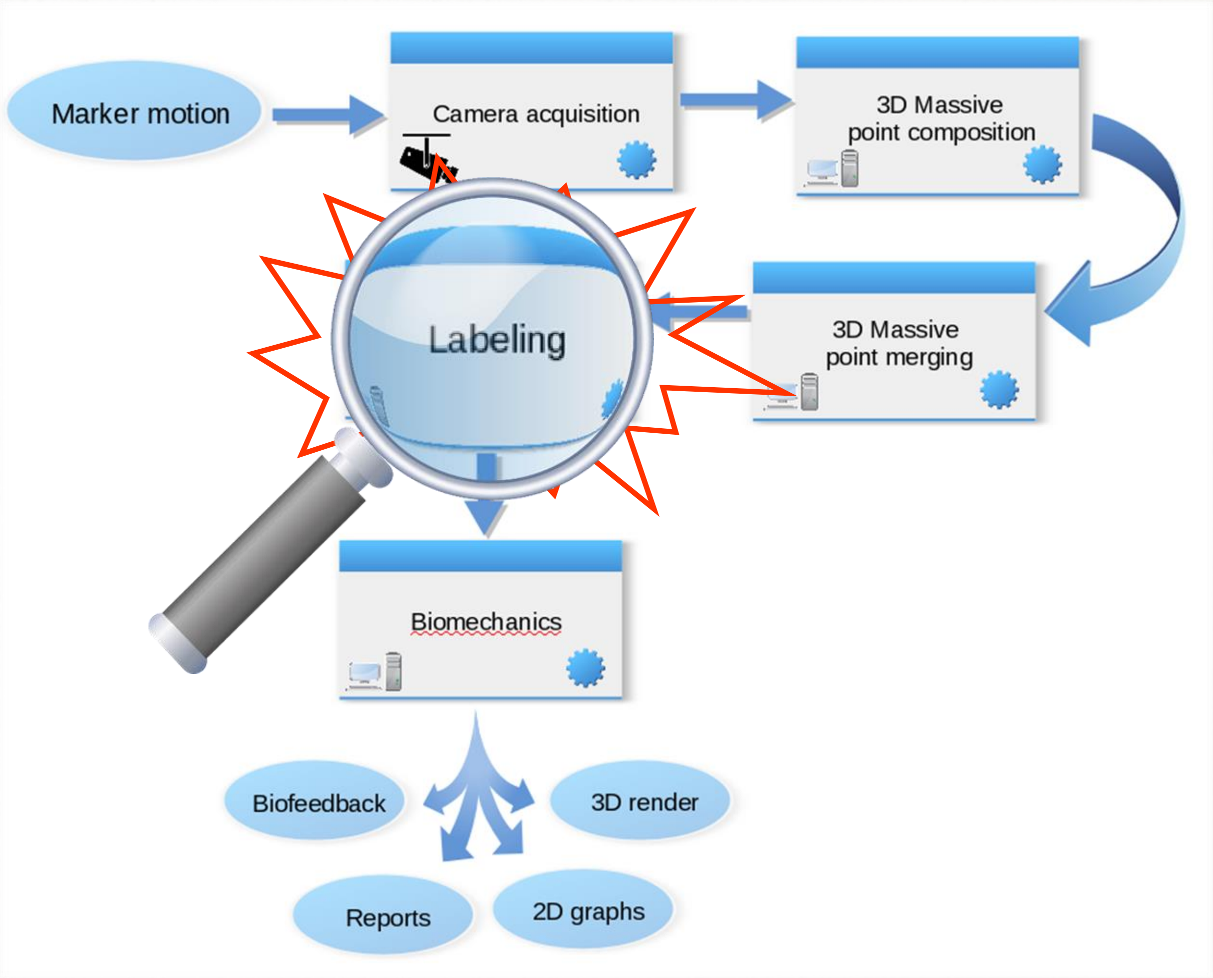




PROCESS STAGES (I)

Optical Mocap workflow

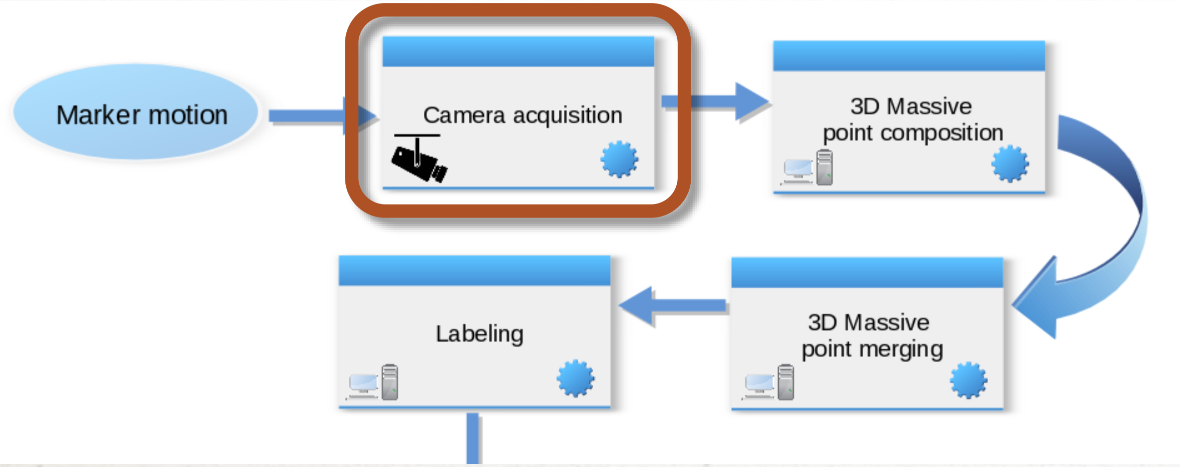
- Known efficient solutions for each stage
- Centred in labelling – not extensive description
- Get to known the problem boundary conditions

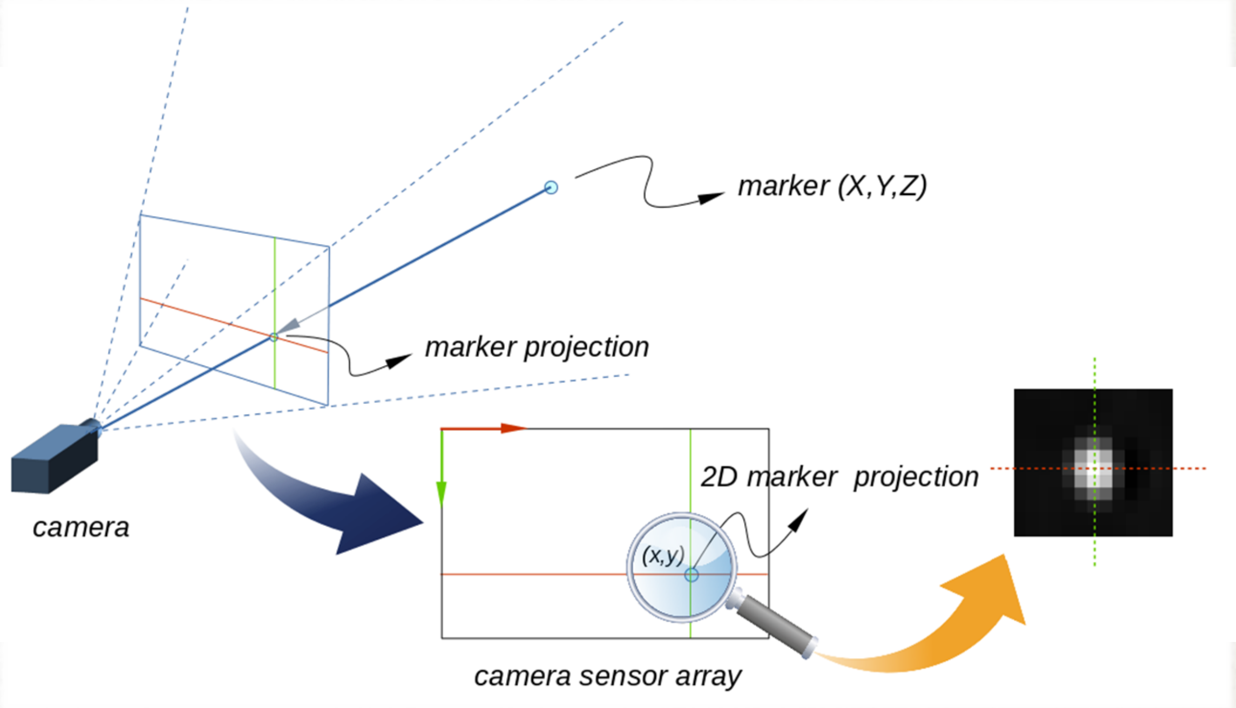


PROCESS STAGES (I)

Marker detection

- IR lightning and reflectivity
- 2D segmentation
- **Anonymous** XY pixel coordinates

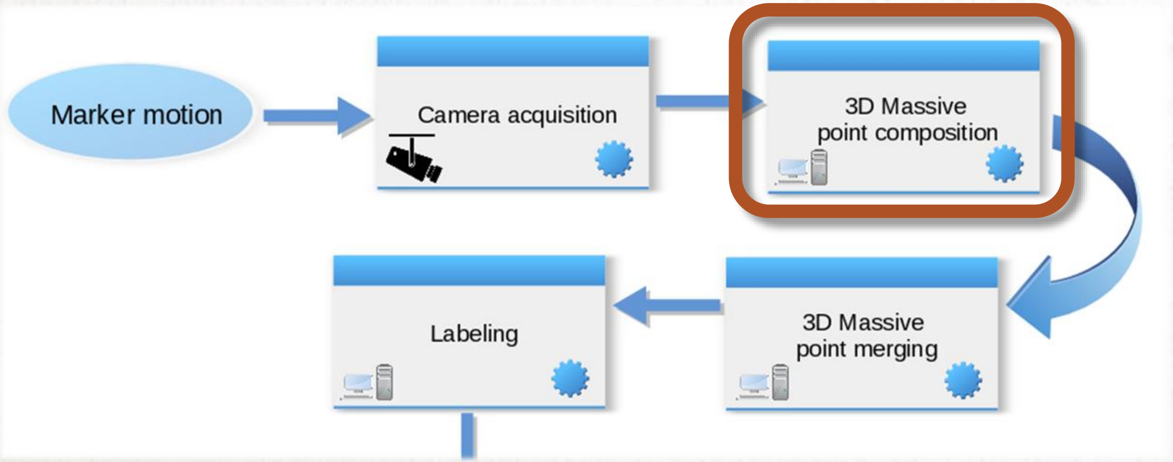


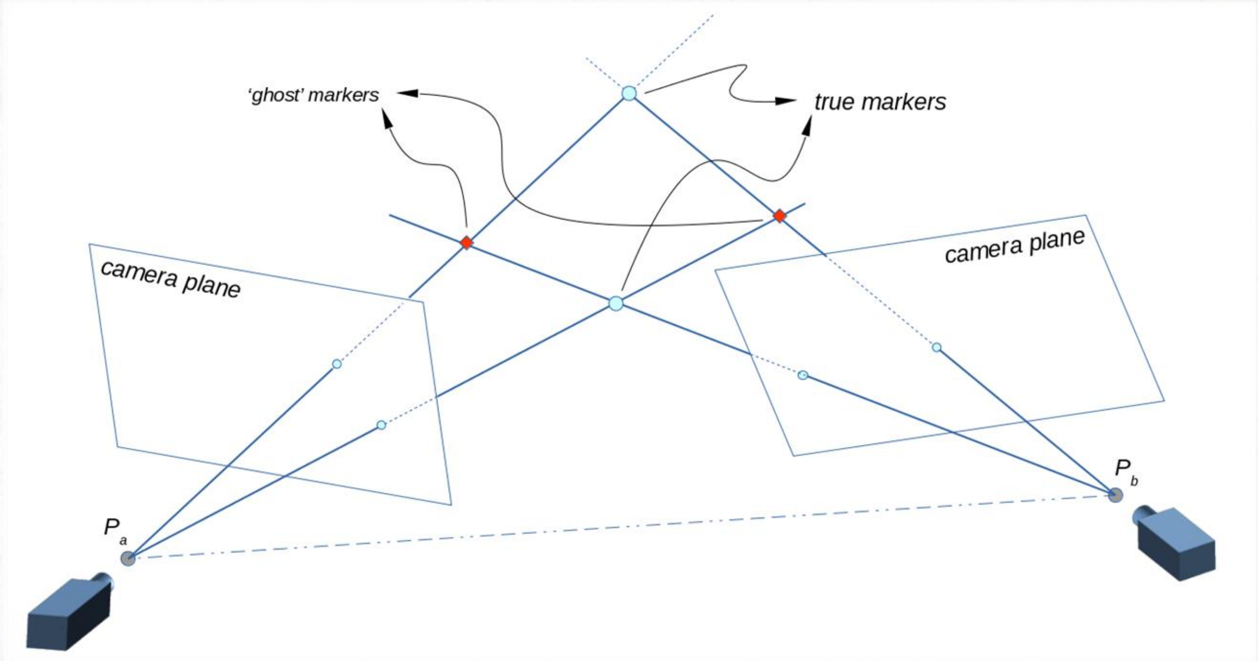


PROCESS STAGES (II)

3D composition (photogrammetry)

- Among all camera and XY combinations
- Anonymous XYZ metric coordinates
- Measurement noise
- Random movement (flesh, clothing, …)
- Ghost markers and **occlusions**

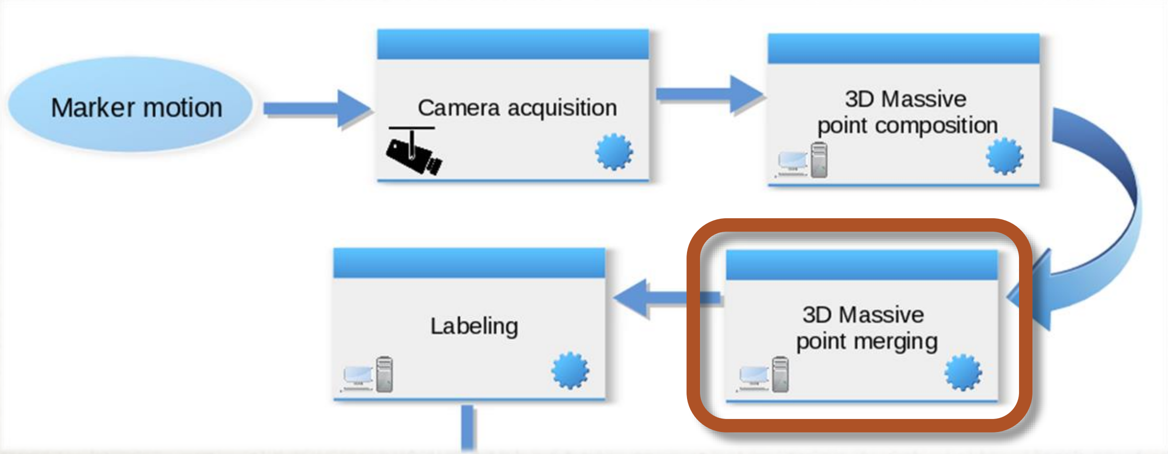


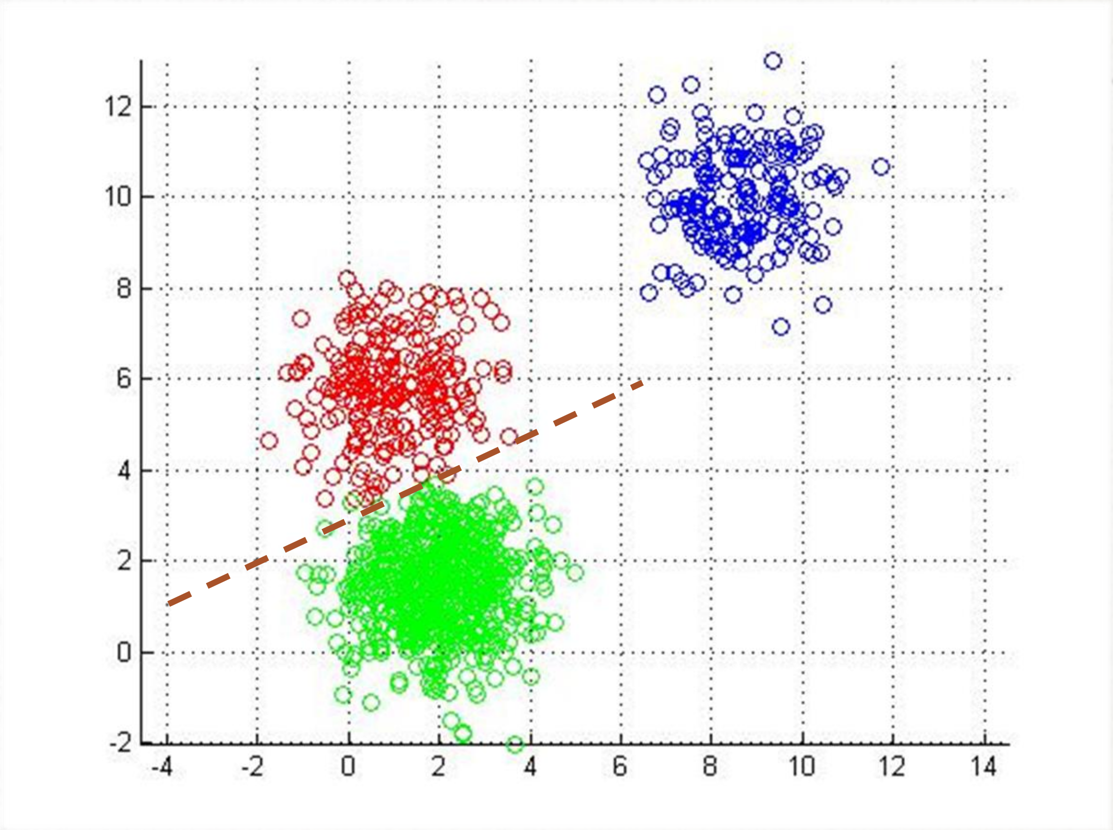


PROCESS STAGES (III)

3D merging

- One point instance per camera pair
- One point per *actual* point
- Remove duplicates
- Make out too near points





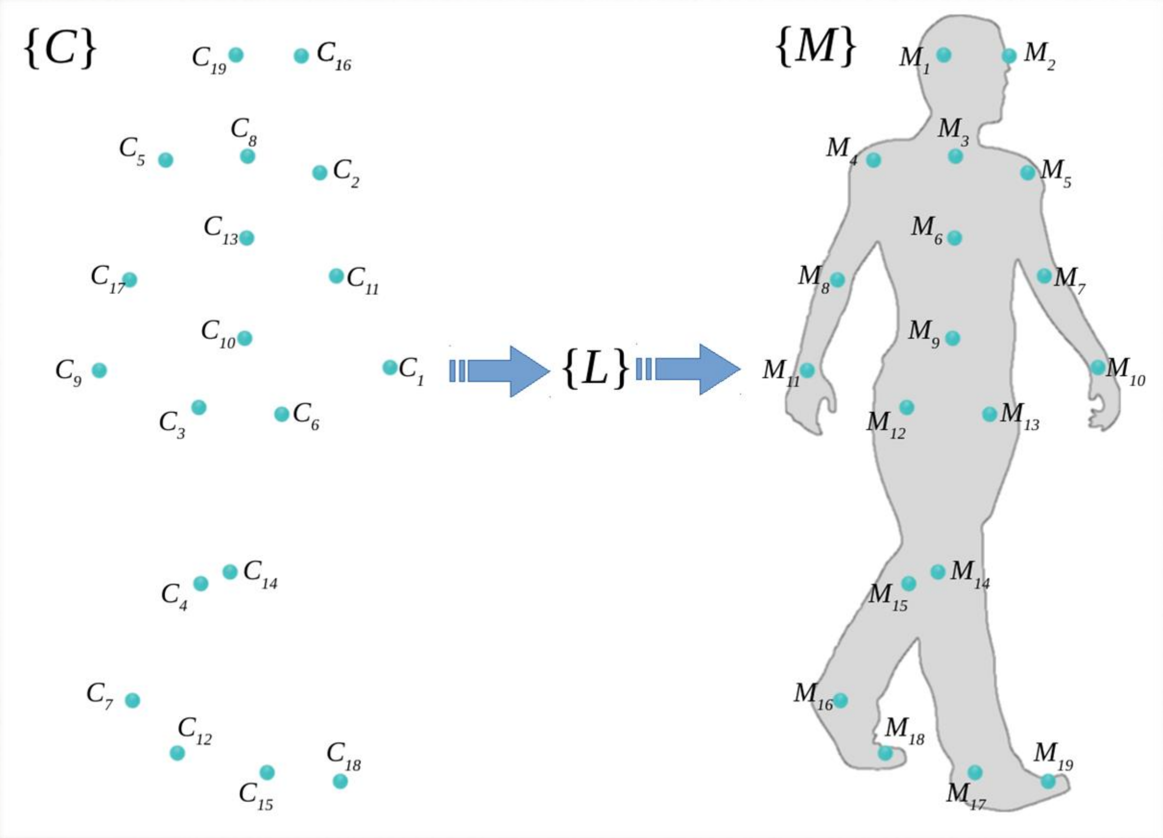
2
PHD THESIS CONTRIBUTION

TAMING THE MARKERS

2
PHD THESIS CONTRIBUTION

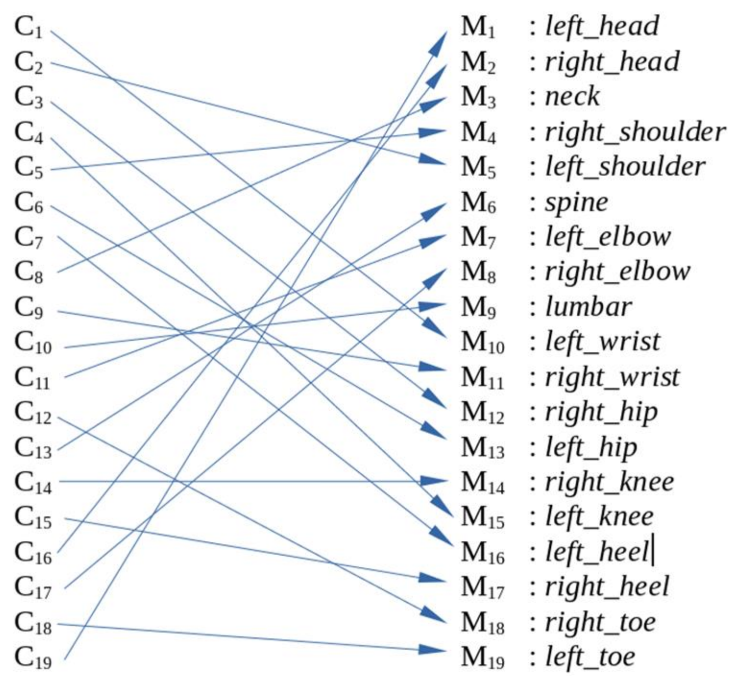
Problem statement
Correct labelling detection
Partial solvers
Solver mining
Solver ensemble

PROBLEM STATEMENT (I)



$L=\{19,16,8,5,2,13,11,17,10,1,9,3,6,14,4,7,15,12,18\}$

$M_i = C_{L(i)}$ $\quad i=1..n$



PROBLEM STATEMENT (II)

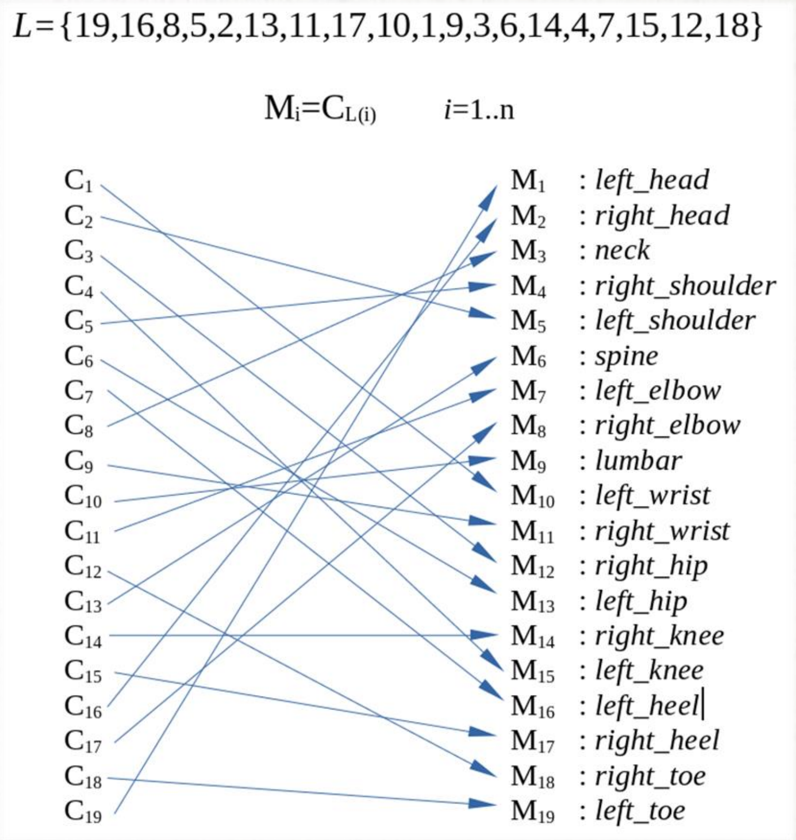
Marker connection

- Input: anonymous candidate list (the cloud)
- Coded as a vector of integers: the **unknown**
- Each integer is either:
  - One candidate index (>0): **labelling assignment**
  - Zero (=0): **occlusion assignment**
- Up to n! different choices for *L* (in a simplified scenario)

$$L_t = \{l_1^t, l_2^t, \ldots, l_n^t\}$$

$$l_i^t \in \{\mathbb{N}, 0\}\ ,\ 0 \leq l_i^t \leq m$$

$$(l_i^t \neq 0) \Rightarrow (l_i^t \neq l_j^t\ \forall j \in \{1, \ldots, n\} - \{i\})$$



PROBLEM STATEMENT (III)

Goal:

- Algorithm to calculate feasible labelling L for each each frame, given a set of candidates C and a marker distribution M:

$$\mathcal{L} = \mathcal{L}(M, C_t)$$

Subject to some requirements:

1. Hit rates in the number of correct candidate assignments above a given target:

$$P_i(l_i) > \tau_i \qquad \begin{matrix} l_i \neq 0 \\ \tau_i \gg 99\% \end{matrix}$$

2. Keep the false occlusion assignment as low as possible
3. Suitable for real time processing (up to 100 frames per second)

THE ROAD MAP

The problem has been addressed in **two** main stages:

1. Initially, we work under the **no occlusion hypothesis**
   - We take on the assumption that not marker is missing in the candidate list
   - Devise a mechanism (*oracle*) to tell between a correct and a wrong labelling L
   - Algorithm (*solver*) to generate all the feasible correct labelling, where no occlusion assignments are possible

2. Then, we undertake the real case: **occlusion happens**
   - *Divide and conquer*: split the marker model into smaller ones (*partial solvers mining*)
   - Apply the algorithm from 1. to generate partial labelling on each subset, when possible
   - Assemble (*solver ensemble*) the partial labelling into a global solution, assuming unidentifiable subsets and thus being its markers assigned as occluded

CORRECT LABELLING DETECTION (I)

The *Oracle*:

- Tell whether a given L is correct or not
- Tackled as a classification problem:

$$\phi(M, C_t, L_t) = \begin{cases} L_t \text{ is correct} & \rightarrow 1 \\ L_t \text{ is no correct} & \rightarrow 0 \end{cases}$$

We define weak classifiers:

- Geometric features over marker sets
- Combinatoric formulation: thousand of instances are available
- Adding a valid range → weak classifier
- n! different choices (simplified scenario)

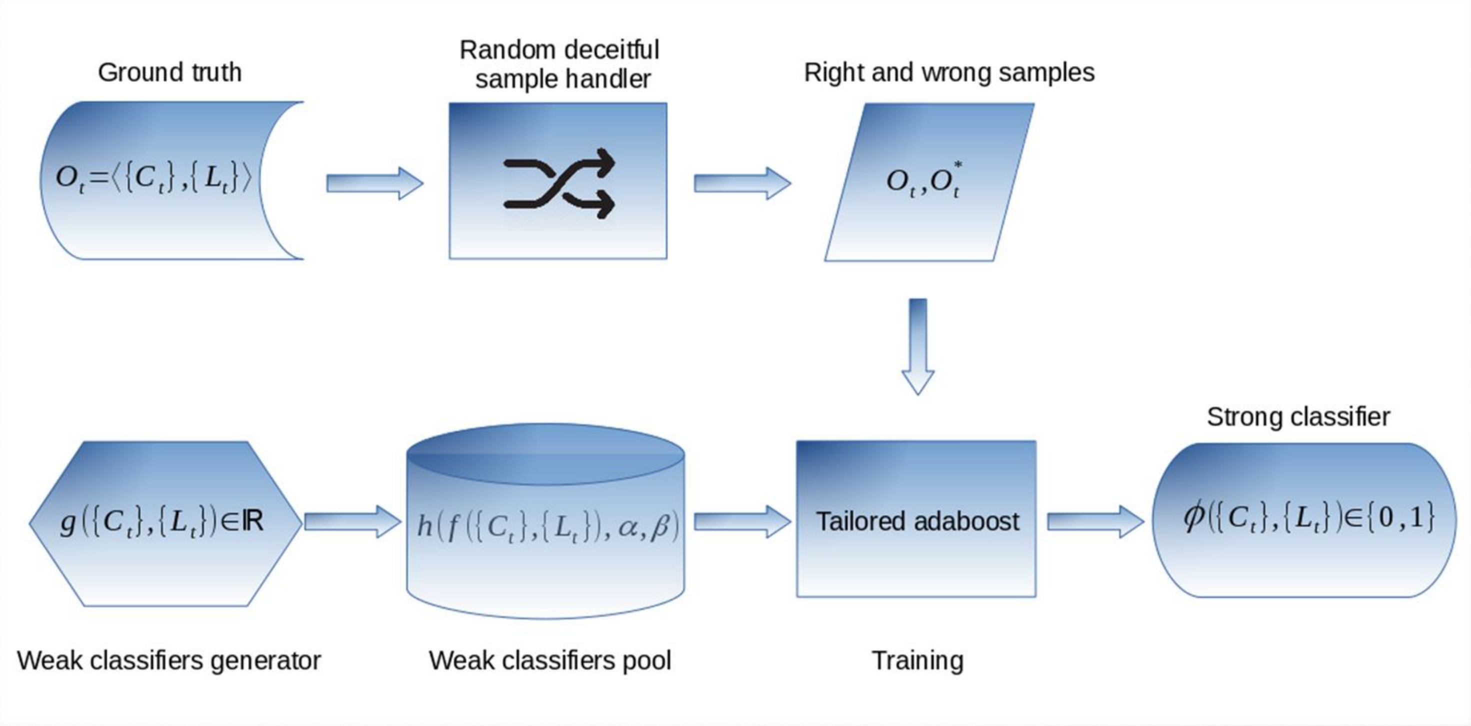
Geometric property	g	# points	points	expression
Angle between consecutive angles	g_1	3	A, B, C	$\arccos\left(\frac{AB \cdot AC}{\lvert AB\rvert \cdot \lvert AC\rvert}\right)$
Distance between points	g_2	2	A, B	$\lvert AB\rvert$
Similarity ratio between segments	g_3	4	A, B, C, D	$2\frac{\lvert AB\rvert - \lvert CD\rvert}{\lvert AB\rvert + \lvert CD\rvert}$
Height difference between two points	g_4	2	A, B	$A_y - B_y$
Distance ratio between consecutive segments	g_5	3	A, B, C	$\frac{\lvert AB\rvert}{\lvert AC\rvert}$
Angle between two segments	g_6	4	A, B, C, D	$\arccos\left(\frac{AB \cdot CD}{\lvert AB\rvert \cdot \lvert CD\rvert}\right)$
Angle between a segment and the vertical	g_7	2	A, B	$\arccos\left(\frac{AB \cdot Y}{\lvert AB\rvert}\right)$
Triangle area	g_8	3	A, B, C	$\frac{1}{2}\lvert AB \times AC\rvert$
Y component of cross vector	g_9	3	A, B, C	$\lvert AB \times AC\rvert \cdot \{0,1,0\}$

CORRECT LABELLING DETECTION (II)

How to build a strong classifier

- Selection by an AdaBoost approach
- Algorithm to automatically gather an ensemble of weak classifiers
- Input data: ground truth and target ratio
- Output: strong classifier

$$\phi(M, C_t, L_t) = \begin{cases} L_t \textit{ is correct } \rightarrow 1 \\ L_t \textit{ is no correct } \rightarrow 0 \end{cases}$$



GENERATING RIGHT LABELLING

The solver

- Form

$$\mathcal{L} = \{L^1, L^2, ...\} = S(C, M, \phi)$$

- Input data: candidate cloud
- Output data: set of feasible labelling

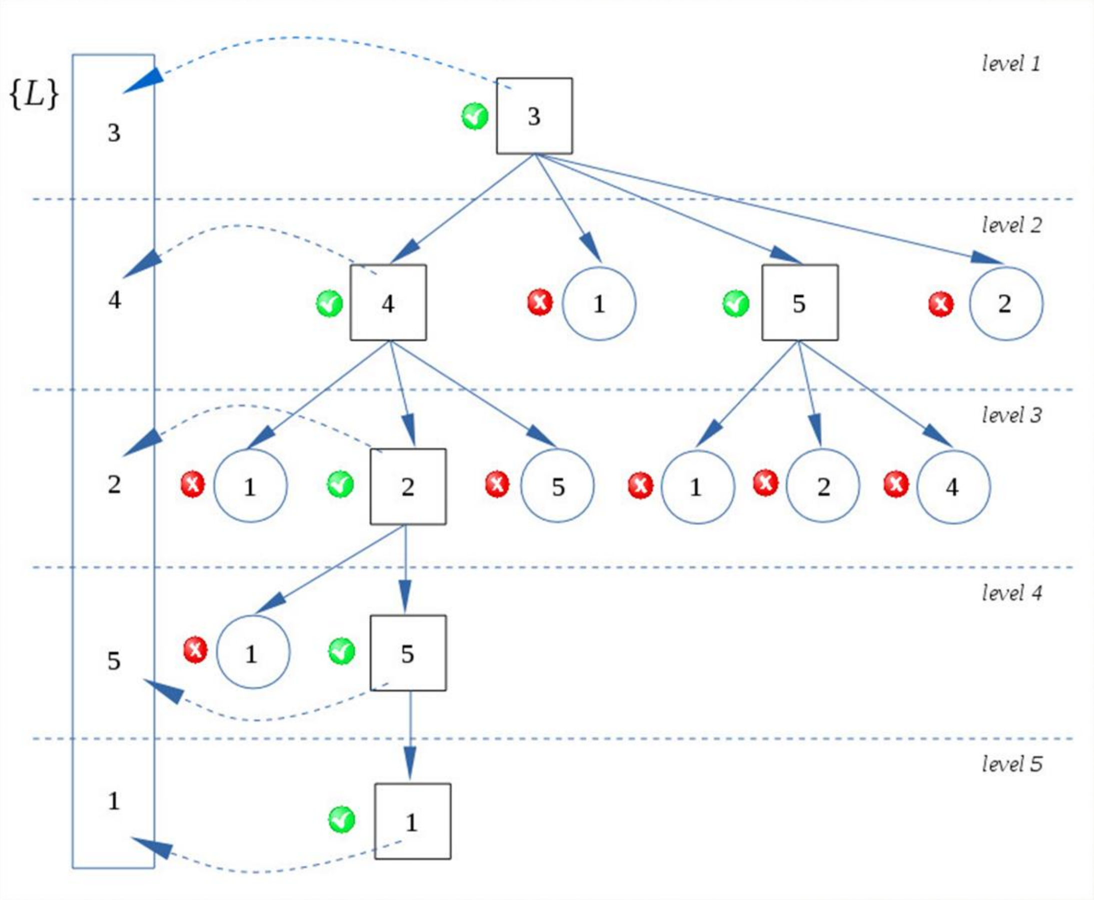
$$\phi(M, C, L^i) = true$$

- **Cannot deal with occlusions!**

$$l_j \in L^i \qquad l_j > 0$$

- Efficient three search algorithm exploiting the trained strong classifier
- Evaluable hit ratio per marker

$$P(S, m_i) = P_i(S) \in [0, 1]$$



PARTIAL SOLVER

Partial solver

- Definition

$$S(C, M_s, \phi_s) \qquad \begin{aligned} L_s &\subseteq L \\ M_s &\subseteq M \end{aligned}$$

- Size: number of marker it work on

Facts:

1. Null hit ratio

$$\text{if } m_i \notin M_s \Rightarrow P_i(S_s, m_i) = NaN$$

2. Increasing hit ratio with solver size

$$m_i \in M_A \subset M_B, |M_B| > |M_A| \Rightarrow P_i(S_B) \geq P_i(S_A)$$

3. Existence of small solvers showing high hit ratios
4. Optimal model design rule

$$\text{if } M_s \equiv M \rightarrow \exists S \setminus P_i(S_s(C, M_s, \phi_s)) = P_i(S(C, M, \phi)) = 1$$

PARTIAL SOLVER ENSEMBLE

Definition:

$$\Omega = \{S_1, S_2, \ldots, S_N\}$$

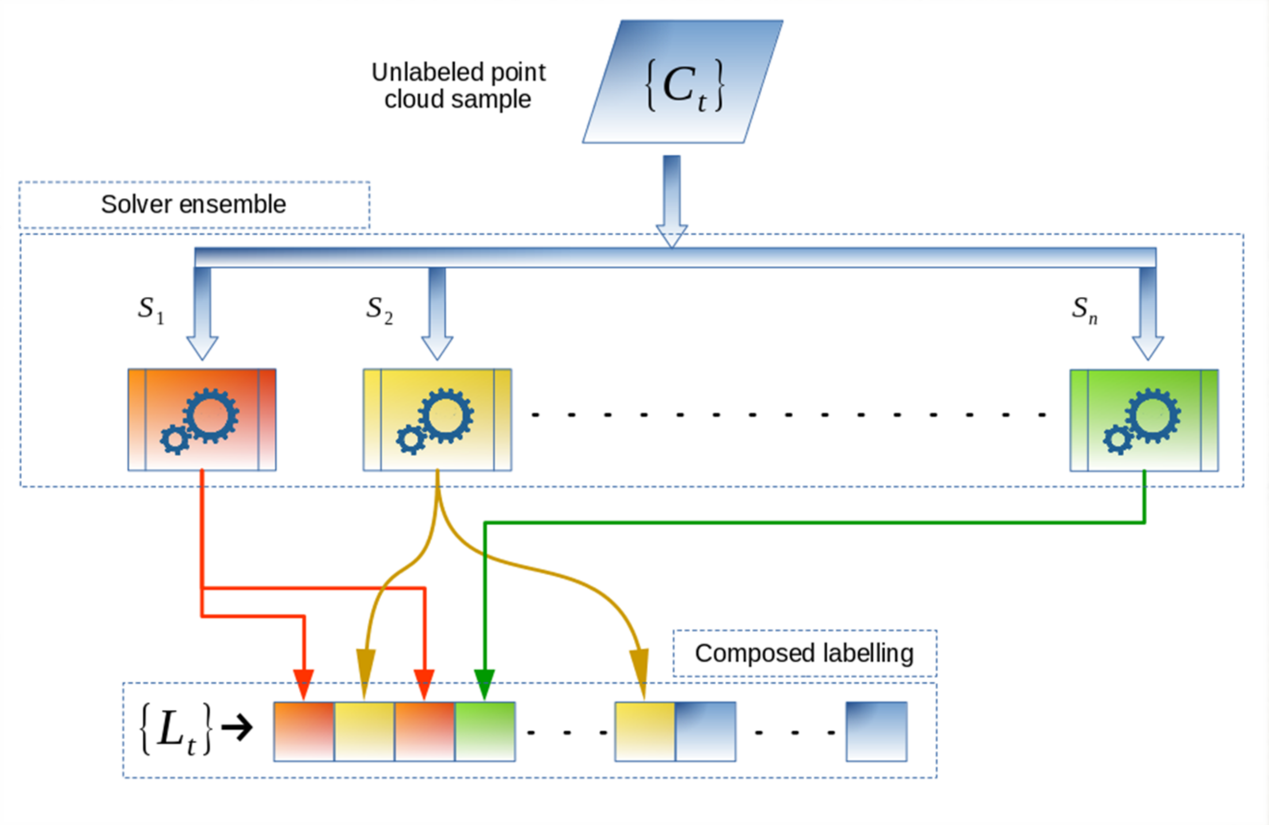
$$M_{s_1} \cup M_{s_2} \cup \ldots \cup M_{s_N} = M$$

Facts:

- **Can deal with occlusions!**
- Size of the members and hit ratio: a trade off
- Combinatoric enumeration: thousand of them available

$$\sum_{i=1}^{n_m} \binom{n_m}{i}, \text{ where } n_m = |M|$$

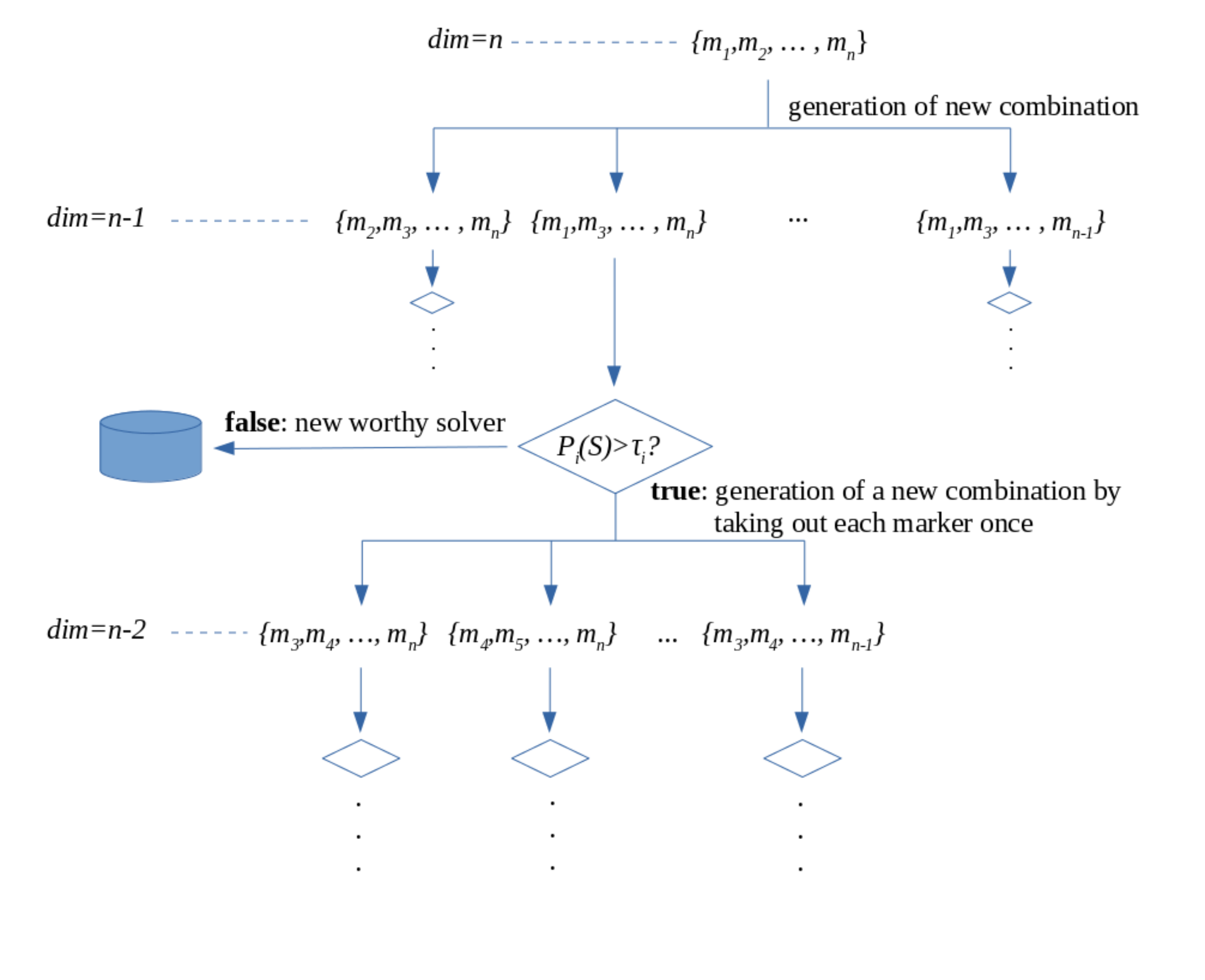
i. e., for n=15 → 32767 different possibilities



PARTIAL SOLVER MINING (I)

Greedy search algorithm

- Bottom-up
- Up-bottom
- Efficient but thorough search
- Target hit ratio as tuning parameter



PARTIAL SOLVER MINING (II)

Genetic search algorithm

- Each solver treated as specimen
- Chromosome: array of Booleans $\{b_i\}$

$$b_i = 1 \; if \; m_i \in M_s, 0 \; otherwise$$

$$L_s = \{l_3, l_5, l_7, l_{10}\} \rightarrow$$

0	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**1**

Genetic operations

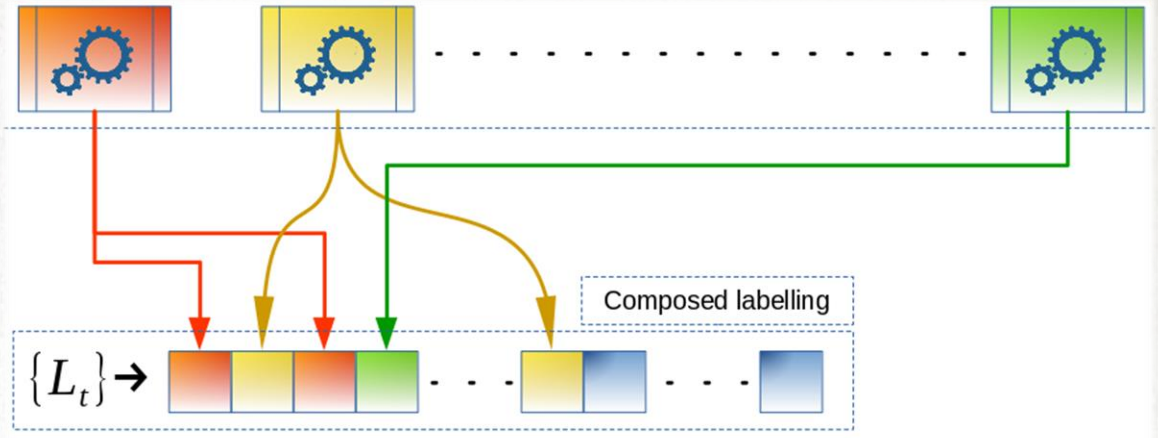
- Crossover, mutation, biased towards small size solvers
- Selection: hit ration as fitness function
- Massive die out after a number of generations

Rather hard to tune

- Multiple simulations
- Usually converges to tantamount valid solutions

SOLVER ENSEMBLE LABELLING

- Each *elite* solver contributes to the complete labelling L
- The consensus is determined by means of a decision matrix
  - ✓ A row per candidate, a column per marker
  - ✓ Each cell holds the suggestion of each solver
  - ✓ Single null entry: assignment
  - ✓ Empty column: null assignment (occlusion)
  - ✓ More than a not-null entry: ambiguity → null assignment
- Target hit ratio ↔ null assignment trade off
  - ↑ hit ratio, ↑ solver size, ↑ occlusion assignments
  - ↓ hit ratio, ↓ solver size, ↓ occlusion assignments



	m_1	m_2	$\cdots$	m_j	$\cdots$	m_{n_m}
c_1	$\varnothing$	$\{S_s^{1,2}\}$	$\cdots$	$\varnothing$	$\cdots$	$\varnothing$
c_2	$\varnothing$	$\varnothing$	$\cdots$	$\varnothing$	$\cdots$	$\varnothing$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
c_i	$\varnothing$	$\varnothing$	$\cdots$	$\{S_s^{i,j}\}$	$\cdots$	$\varnothing$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
c_{n_c}	$\varnothing$	$\varnothing$	$\cdots$	$\varnothing$	$\cdots$	$\{S_s^{n_c,n_m}\}$

3
RESULTS

WHEN THE RUBBER MEETS THE ROAD

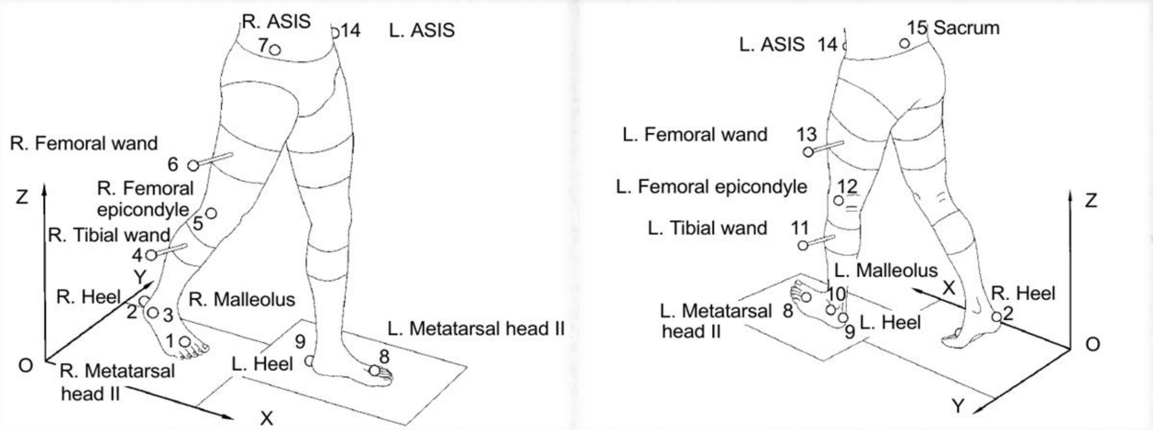
3
RESULTS

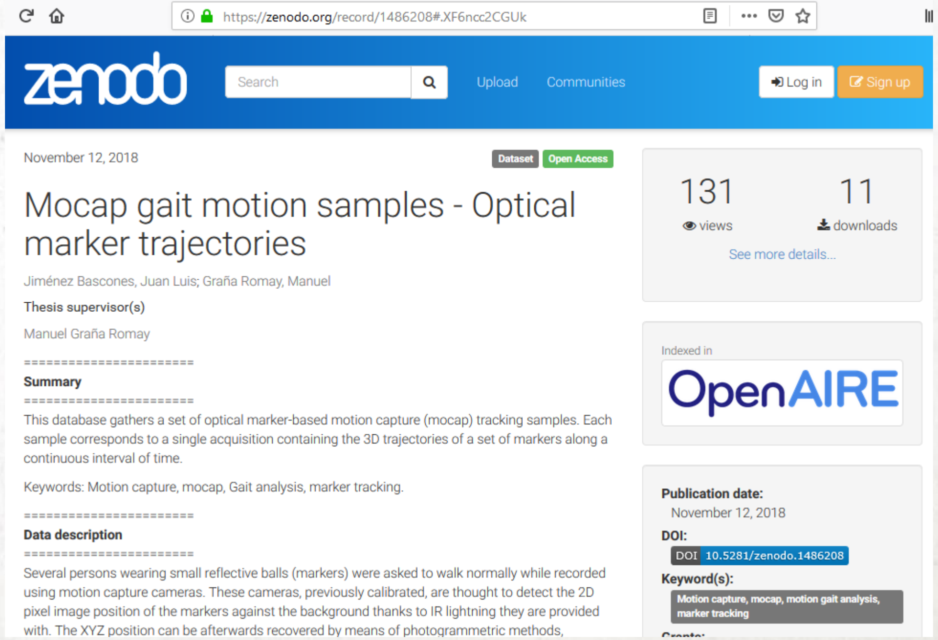
Ground truth
Partial solver performance
Solver ensemble performance

EXPERIMENTAL DATA

Ground truth

- Helen Hayes marker system – 15 markers
- Clinical gait analysis purposes
- Six mocap cameras, native resolution 800x800 pixels
- Sampling frequency: 100Hz
- Database size: 71 recordings between 1.5 and 7 seconds long, summing up more than 20.000 frames
- Manually verified
- Openly available at zenodo.org

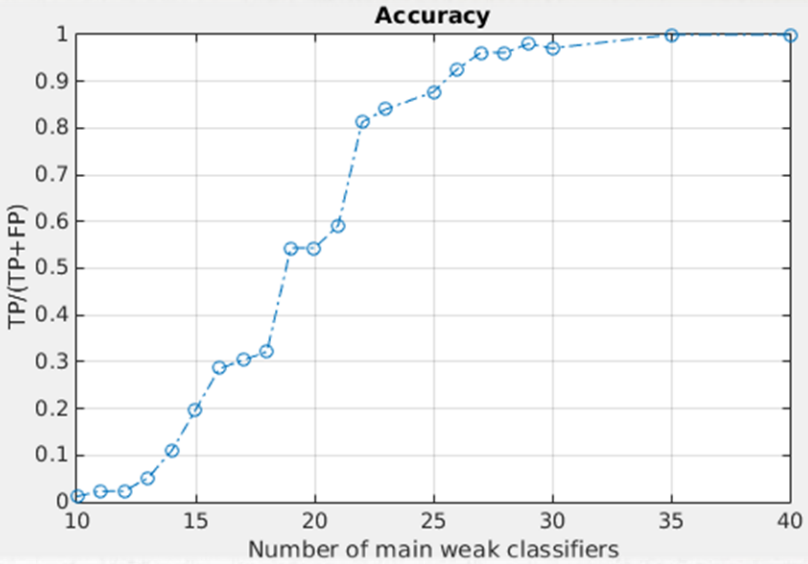


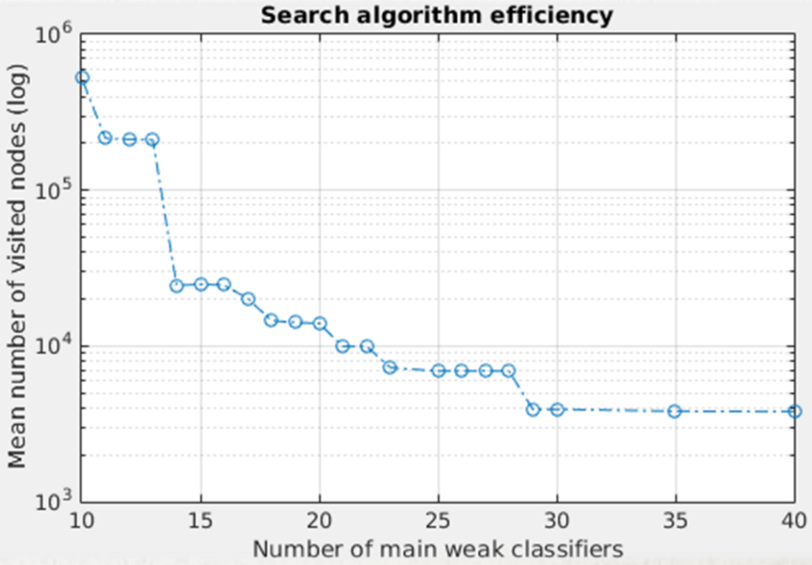


STRONG CLASSIFIER

- Holds the labelling knowledge used by the solver: makes out between a given labelling is right or wrong
- The most relevant are selected by an tailored Adaboost implementation
- Direct effects in the solver performance:
  - ✓ Hit ratio
  - ✓ Processing time
- Do not deal with occlusions

	Weak classifier	Score (%)	Sum score (%)
1	TriangleNormal_Y(R_asis,L_asis,sacrum)	18.82	18.82
2	Dist(R_malleolus,R_heel)	12.91	31.74
3	Dist(L_malleolus,L_heel)	12.84	44.59
4	Dist(R_femoral_epicondyle,R_tibial_band)	11.85	56.45
5	Dist(L_femoral_wand,L_femoral_epicondyle)	11.51	67.96
6	Dist(L_tibial_wand,L_meta_h)	10.87	78.84
7	CoordDiff_Y(R_femoral_wand,R_meta_h)	10.43	89.28
8	TriangleNormal_Y(sacrum,R_meta_h,L_meta_h)	1.90	91.17
9	Dist(R_femoral_wand,R_femoral_epicondyle)	1.84	93.02





Cloud Point Labelling in Optical Mocap
Juan Luis JIMÉNEZ BASCONES
CIG – UPV/EHU

SOLVER ENSAMBLE (I)

- Worthy solvers mining
- Wished hit ratio as tuning parameter or the mining
- Samples taken from the ground truth with occlusion simulation
- Assessment of the solver ensemble in terms of:
  - ✓ False positive assignments (FA)
  - ✓ False occlusion assignments (FO)

Test conditions	
Number of markers	15
Target hit rate	99.99%
Target failure rate	0.01%
Occlusions per frame	4
Number of test frames	16384

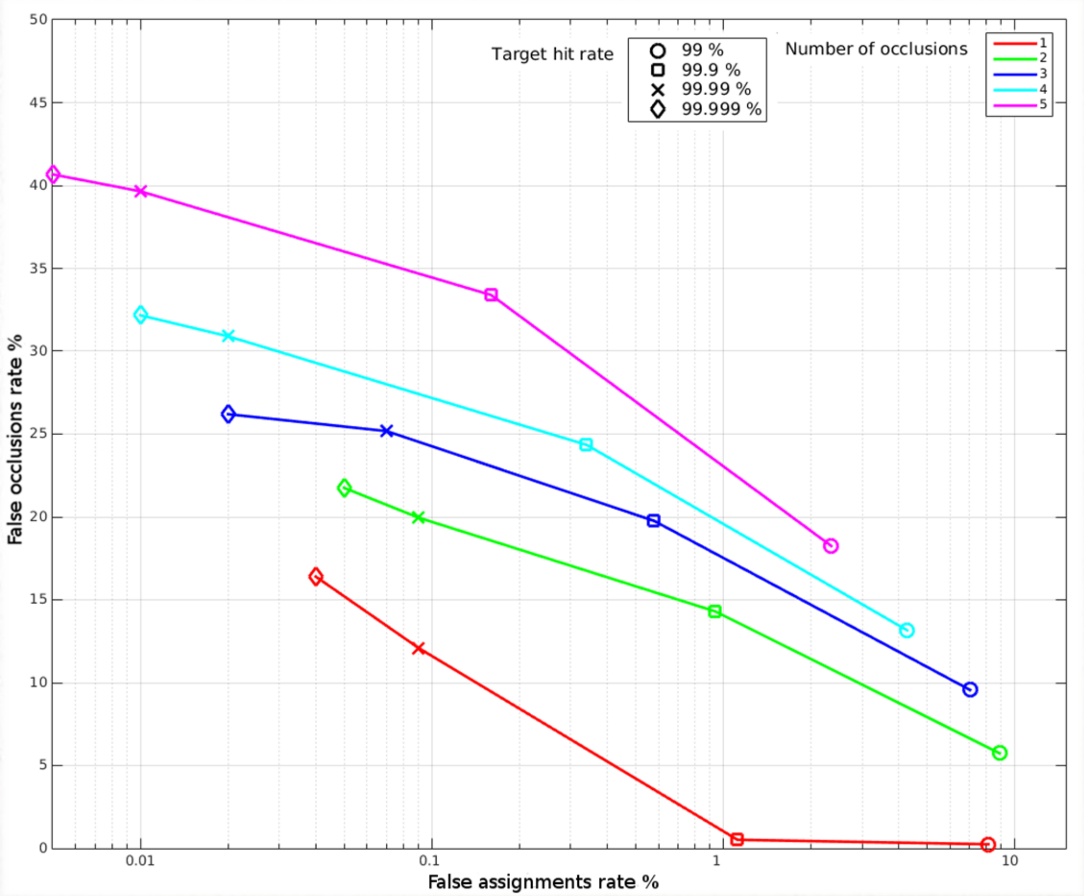
Marker ID	FA #	FA %	FO #	FO %
r_asis	2	0.02%	300	6.38%
l_asis	2	0.02%	254	5.55%
s2	0	0.00%	286	6.11%
r_l_thigh	0	0.00%	3564	45.03%
l_l_thigh	1	0.01%	220	4.85%
r_knee	0	0.00%	1912	30.03%
l_knee	1	0.01%	3532	44.58%
r_calf	1	0.01%	194	4.20%
l_calf	3	0.03%	218	4.71%
r_ankle	3	0.03%	2302	34.02%
l_ankle	1	0.01%	4348	49.30%
r_heel	4	0.05%	3195	42.14%
l_heel	1	0.01%	3579	45.06%
r_toe	4	0.04%	2332	34.67%
l_toe	4	0.05%	3627	45.13%
Average	1.8	0.02%	1991	31.16%

SOLVER ENSAMBLE (II)

- False assignment – False occlusions trade off
- Hit ratio as tuning parameter

False assignments rate					
Target marker hit rate		99.000%	99.900%	99.990%	99.999%
Number of true occlusions per frame	1	8.13%	1.12%	0.09%	0.04%
	2	8.89%	0.94%	0.09%	0.05%
	3	7.04%	0.58%	0.07%	0.02%
	4	4.28%	0.34%	0.02%	0.01%
	5	2.35%	0.16%	0.01%	0.00%

False occlusions rate					
Target marker hit rate		99.000%	99.900%	99.990%	99.999%
Number of true occlusions per frame	1	0.25%	0.52%	12.10%	16.40%
	2	5.74%	14.30%	19.96%	21.76%
	3	9.56%	19.76%	25.18%	26.20%
	4	13.16%	24.35%	30.90%	32.17%
	5	18.25%	33.39%	39.65%	40.67%



DISCUSSION

- Algorithm behaviour
  - Lack of individual discriminating characteristics for identifying the markers separately (XYZ as only information)
  - Sort of community: the ID for each members is supported by the guess on the ID given to its partners by means of weak classifier
  - The solver mining is a sort community finder
- Algorithm benefits
  - ✓ Works on each frame independently
  - ✓ Parallelizable
  - ✓ Scalable

4
CONCLUSIONS

... AND WHAT'S NEXT

4
CONCLUSIONS

Achievements
Limitations
Future work

ACHIEVEMENTS

- Fulfilment of requirements
  - ✓ Reliable against severe and lasting occlusions
  - ✓ Released from previous mistakes – robust recovery (tracking vs labelling)
  - ✓ Efficient
  - ✓ Trainable to specific cases
- Contribution
  - ✓ Original approach, breaking away from traditional *tracking* skeleton-driven algorithms
  - ✓ Approach according to the fuzzy nature of the problem: learning vs. hand-coded rules
  - ✓ Windfalls:
    - substantiation of the awareness of mutual maker ID backup
    - disclosure of hit ratio-false occlusions trade off
    - suitability (*capturability*) of a maker distribution
    - per-marker customizable hit ratio

LIMITATIONS

Associated to the learning process itself

- × Supervised – not that easy to gather a big enough ground truth
- × Sensitive to ground truth corruption – errors hard to catch
- × Sticks to the given samples – wont work with atypical ones
- × Does not take advantage of timed information …
- ✓ … but it could be easily accommodated

$$g_k = g_k\left(M_i, M_j, \ldots, \dot{M}_i, \dot{M}_j, \ldots, \ddot{M}_i, \ddot{M}_j, \ldots\right)$$

$$\dot{M}_i = \left\{\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial y}{\partial t}\right\}$$

FUTURE WORK

New research topics:

- Unsupervised learning
- Further testing with different kind of movements
- Treatment of cases out of the learning set
- Testing of the parallelizable version of the algorithm
- Incorporation of temporal information
- Incorporation of dynamic descriptors

QUESTIONS

... & ANSWERS

QUESTIONS

- ?
- ??
- ???
- ...


Thank you