An ensemble of weak classifiers for pattern recognition in motion capture clouds of points

J.L. Jiménez Bascones[1,2], Manuel Graña[2]

[1]Motion Capture Unit, STT-Systems (San Sebastian), [2]Computational Intelligence Group, University of the Basque Country (UPV/EHU)

May 23, 2017

Abstract

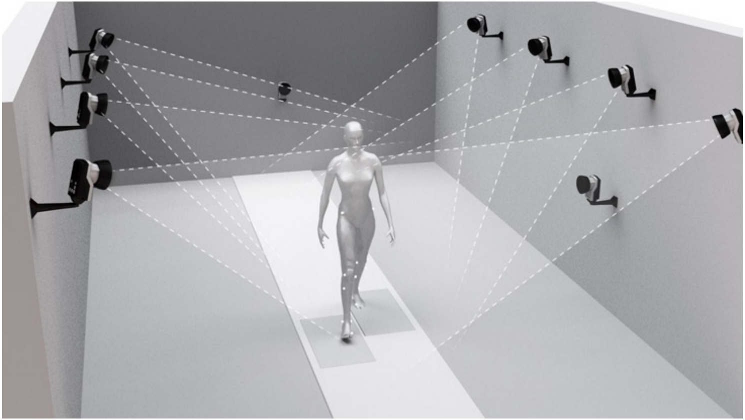
- problem: labeling a cloud of points as a classification problem by an ensemble of weak classifiers.
  - we define a set of geometrical features over small subsets of the cloud of points.
  - we apply an Adaboost like strategy to select a collection of features
- Two problems:
  - Verifying that a labeling is correct (binary classification),
  - generate the labeling of the points in the cloud.
- real life dataset obtained from the measurement of gait motion of persons,
  - ground truth labeling defined manually.
- Results are encouraging on real life data

contents

Contents

- The process of recording the movement of objects (very often human bodies) so that they are digitized into a computer model is known as *motion capture* (MoCap).
- This technology is widely employed in many scientific and industrial fields like entertainment, clinical analysis, rehab and sports.
  - A paradigmatical application is gait analysis.
- Capture systems can be
  - optical (where a set of cameras is used to record the movement) and
  - non-optical (those that use inertial, magnetic or mechanical devices).

- Optical systems capture the movement by means of a set of calibrated and synchronized cameras deployed around the *scene*, recording images at a constant frame rate.
  - Frame by frame, a set of 2D points (passive markers) are extracted from the camera images
  - 3D coordinates re computed by photogrammetric techniques .
  - We do not consider other information (i.e. color codes, surrounding image or fiducial schemes).
- unique identification of the candidates points makes biomechanical calculations possible.



contributions

1. Formulation of the labeling correctness as a classification problem;
2. Proposal of a way of computing geometrical features over the cloud of points which allow to define weak classifiers;
3. An Adaboost approach to build the ensemble classifier from a collection of weak classifiers;
4. Label generator by using the weak classifiers to guide the process;
5. We demonstrate the validity of approach on a large dataset obtained from the real industrial practice

Contents

- keeping the labeling through the time using trajectory estimators,
  - Kalman filter tuned to fit each particular marker behavior.
  - error prone to deal with marker occlusions (points kept out of sight of the cameras) lasting several consecutive frames.
- use of the underlying human skeleton by the identification of the markers belonging to the same body limb.
  - maintaining relative distances
  - The identification of a reappeared maker is backed up by those sharing the same limb.
  - This method may fail in case of massive occlusions where nearly all markers from the same limb have been hidden for too long.
- commercial solutions
  - *Cortex* (developed by Motion Analysis), *Track Manager* (from Qualisys) or *Clima* (by STT Systems).
  - no information about the details of the internal tracking

Contents

Problem statement

- n passive marker over the object whose movement has to be tracked.
  - Each marker has a predefined and constant position over the body
  - and a unique ID ('*right-shoulder*', '*left-knee*',...) is given.
- *model* $\{M\}=\{M_1, M_2, ..., M_n\}$,
- *candidate points* are gathered in $\{C_t\}=\{C_1^t, C_2^t, ..., C_m^t\}$, $t = \{0, 1, 2, ..., T\}$
- When $m \neq n$
  - some real marker is hidden to the cameras (occlusion) or
  - *ghost* points make their appearance on the scene
- challenge : correctly match the elements from M and C using only geometric information.

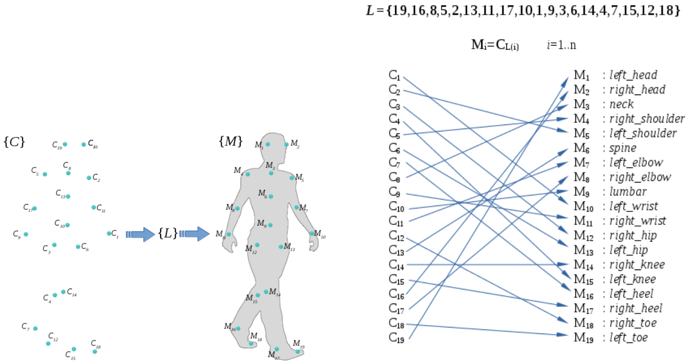


Figure: Example of a humanoid model labeling L.

- labeling $L : C_t \rightarrow M$ coded as the integer vector $L_t = \{l_1^t, l_2^t, ..., l_n^t\}$ where:

$$l_i^t \in \{\mathbb{N}, 0\}, \; 0 \leq l_i^t \leq m \tag{1}$$

$$\left(l_i^t \neq 0\right) \Rightarrow \left(l_i^t \neq l_j^t \; \forall j \in \{1, \ldots, n\} - \{i\}\right) \tag{2}$$

  - $l_i^t > 0$ connects the marker M_i with candidate point $C^t_{l_i^t}$,
  - $l_i^t = 0$ means that marker M_i has no match among the candidate points (i.e it has been occluded).
  - No two elements of L contain the same non-zero mapping since a given candidate cannot be simultaneously assigned to more than one marker.

labeling correctness detection

- Given a marker model and a set of candidate points, the challenge is to decide whether a given labeling L is correct or not as a whole,
  - i.e. if one label is incorrect the whole labeling is incorrect.
- We build a two-class classifier where class 1 is the correct labeling

$$\phi\left(M, C_t, L_t\right) = \begin{cases} L_t \text{ correct} \longrightarrow 1 \\ \text{not correct} \longrightarrow 0 \end{cases}$$

Label generation

- The challenge is to generate the correct labels of the candidate points using the weak classifiers that have been developed for the detection of correct labelings.
- Here the decision is independent for each point,
  - so we might have an incomplete labeling.

Contents

Geometric Features and weak classifiers

- Given the XYZ coordinates of the points, we define geometric function g yielding scalar values.
- Examples of geometric functions are listed in the table 1, each corresponding to a geometric property of the polygon defined by the set of points.

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	$\arccos\left(\frac{AB\cdot Y}{\lvert AB\rvert}\right)$

- The number of features grows combinatorially with the size of the cloud,
    - it is of the order of $\binom{n}{k}$, where
        - n is the number of points in the cloud, and
        - k the number of points considered by the feature.
- Therefore, the possible geometrical functions must be limited, and effective features must be selected from the pool of all potential features.

- Adaboost: each feature defines a weak classifier,
- each feature has a range of natural values $[\alpha, \beta]$ for correct labelings
- weak classifiers:

$$h\left(f_k^S\left(M, L_t, C_t\right), \alpha, \beta\right) = \left\{ \begin{array}{ll} 1 & \text{if} \quad \alpha < f_k^S\left(M, L_t, C_t\right) < \beta \\ 0 & \qquad\qquad \text{otherwise} \end{array} \right. , \tag{3}$$

where

- f_k^S is a feature built with geometric function $g_k()$
- applied to a subset of points $S \subset M$
- selected from the cloud C_t,
- $[\alpha, \beta]$ its natural interval, and
- class 1 denotes correct labeling of the cloud of points.

Building a strong classifier has the following steps:

1. Generate all the possible features.
2. Determine the natural interval of values for each feature,
   1. $\alpha_k^S = \min_C f_k^S(C)$ and $\beta_k^S = \max_C f_k^S(C)$, where all clouds C are correctly labeled.
3. Select the minimal collection of features that ensures a desired accuracy level of the ensemble of weak classifiers.
   1. any collection of weak classifiers will provide very high sensitivity but very low specificity
   2. greedy selection of the weak classifier providing the biggest increase of accuracy by decreasing the number of false positives.

The ensemble of weak classifiers

- $\mathbf{O} = \{O_i\}$ the set of learning observations $O_i = \{C_i, L_i, b_i\}$ corresponding to a common model M.
  - Each observation has a cloud of points C_i and the labeling L_i that maps it into the model.
- b_i encodes the correctness of the mapping,
  - $b_{ij} = 0$ if the label of the j-th cloud point is incorrect, thus the whole labeling is incorrect.
- The training data generation:
  - correct labels available
  - incorrect labeled observations generated by permutation of the labels in a selected correct observation.
  - $\mathbf{O}^* = \{O_i^*\}$ denote the incorrect samples

Ensemble of weak classifiers

- collection of features whose corresponding weak classifier is weighted by its accuracy gain relative to the remaining weak classifiers.
- ensembles output:

$$\phi_J(M, C, L) = \frac{\sum_{j=1}^{J} w_j h_j(f_k^S(C), \alpha_k^S, \beta_k^S)}{\sum_{j=1}^{J} w_j}, \tag{4}$$

where

- j refers to the order of selection of the feature for inclusion in the ensemble,
- J is the size of the ensemble.

Ensemble training

- Adaboost strategy: greedy selection of the weak classifier that maximize the increase in accuracy.
- Initially, all weights are initialized to zero and the set of selected weak classifiers is empty.
- In a loop we feed all classifiers with observations of different error severity
  - If $\phi_J(M, C, L)$ does reject the incorrect sample no further process is done.
  - If not, the weights of unselected weak classifiers that reject it are updated according to the error severity.
- After a number of incorrect observations is processed, add the weak classifier having the greatest weight.
- The whole process eventually ends up when a given threshold on the accuracy of the strong classifier is reached.

Label generation

- Given an ensemble of weak classifiers $\phi_J(M, C, L)$,
  - the number of weak classifiers giving positive outcome
    - measure of how well the vector of integers L links the model points M and the candidate points C.
  - labeling of a cloud of points: L that maximizes the number of weak positive classifications to achieve $\phi_J(M, C, L) = 1$.
- no marker is occluded and no points other than the ones to be labeled are present in the input data.
- the number of possible configurations for L is $n!$.

Label generation

- branch and bound strategy using the following properties:
  - classifier ϕ can be evaluated over a *partial solution* where only a subset of elements of L as meaningful labels.
  - weak classifiers using unassigned labels are ignored;
  - a single weak classifier rejecting a permutation definitively rules it out,
  - not all the elements of ϕ must be computed,
  - a single weak classifier can be computed from a handful of points (usually from 2 to 6) which represents a subset of the vector L;

Contents

Experimental data

- labeled manually
- 70 sequences of non-pathological gait analysis movement involving
  - 14 different people from different ages and body shapes walking at random paces.
- Each sequence was recorded at 100Hz.
  - The average duration of the sequences is three seconds,
  - more than 20.000 point cloud frames.
  - These labeled clouds are correct labeled data
- Point clouds with incorrect labeling (class 0) are generated by random permutations of the labels of correct labeling data.
- The point cloud sequence capture follows the Helen Hayes lower train protocol
- *STT's* proprietary optical motion tracking system, including 6 infrared synchronized cameras of 800x800 pixel resolution

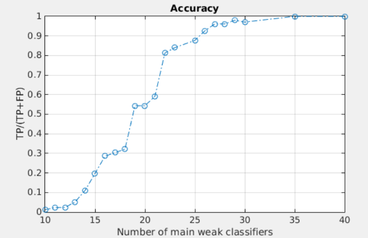
labeling correctness

- geometrical functions g_2, g_4 and g_9 from Table 1
  - 665 weak classifiers.
  - The training algorithm selects 40 weak classifiers. T
- he ensemble classifier achieves an accuracy over 99% after the presentation of more than 10^7 negative samples with diverse error severity.

Table: First selected weak classifiers

	Weak classifier	Score (%)	Cum sc
1	*TriangleNormal_Y(R_asis,L_asis,sacrum)*	*18.82*	*18*
2	*Dist(R_malleolus,R_heel)*	*12.91*	*31*
3	*Dist(L_malleolus,L_heel)*	*12.84*	*44*
4	*Dist(R_femoral_epicondyle,R_tibial_band)*	*11.85*	*56*
5	*Dist(L_femoral_wand,L_femoral_epicondyle)*	*11.51*	*67*
6	*Dist(L_tibial_wand,L_meta_h)*	*10.87*	*78*
7	*CoordDiff_Y(R_femoral_wand,R_meta_h)*	*10.43*	*89*
8	*TriangleNormal_Y(sacrum,R_meta_h,L_meta_h)*	*1.90*	*91*
9	*Dist(R_femoral_wand,R_femoral_epicondyle)*	*1.84*	*93*

Label generation results



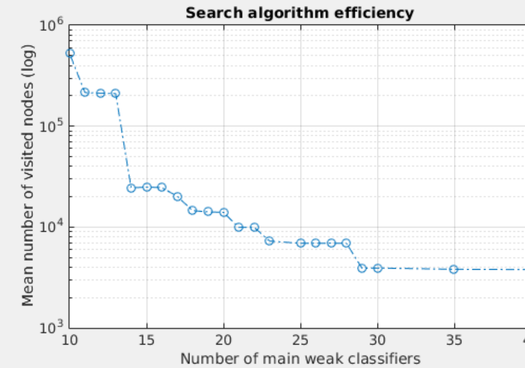


Figure: Accuracy (10-fold cv) and efficiency assessment depending on the number of weak classifiers.

Contents

- We have dealt with the problem of labeling a cloud of points according to a given model using a simple classification approach that does not use any semantic prior structural information, such as anatomical locations or graphs of expected relative positions.
- We define geometrically based features that are evaluated over the cloud of candidate points.
- Corresponding weak classifiers are defined and trained on the available data.
- Adaboost-like selection of the minimal collection of weak classifiers achieving a given accuracy threshold

- The experimental validation is carried out over a dataset obtained from the real industrial experience of gait analysis.
- We achieve encouraging results in both correct labeling detection and label generation tasks.
- labeling of a cloud of points can be carried out in times of the order of $10^{-3}s$, which raises expectation for its use in real time.
- Future work will be addressed to deal with occlusions, and to make the approach simultaneously valid for several models, i.e. the ensemble may be able to detect which model is best fit for the cloud of points.