Integration of smartphones and webcam for the measure of spatio-temporal gait parameters

V. Barone, E. Maranesi and S. Fioretti

Abstract— **A very low cost prototype has been made for the spatial and temporal analysis of human movement using an integrated system of last generation smartphones and a high-definition webcam, controlled by a laptop. The system can be used to analyze mainly planar motions in non-structured environments. In this paper, the accelerometer signal as captured by the 3D sensor embedded in one smartphone, and the position of colored markers derived by the webcam frames, are used for the computation of spatial-temporal parameters of gait. Accuracy of results is compared with that obtainable by a gold-standard instrumentation. The system is characterized by a very low cost and by a very high level of automation. It has been thought to be used by non-expert users in ambulatory settings.**

I. INTRODUCTION

3-D accelerometers and webcams are assuming an increasing interest for the ambulatory assessment of human motion, particularly for the extraction of spatial-temporal parameters of gait [1-5], but also for a simplified analysis of motor tasks deemed clinically significant for the functional evaluation of motor disabilities. Literature reports many attempts to use accelerometers, often embedded in more complex measurement units, for the analysis of gait. They are able to give accurate estimate of the temporal gait parameters like, f.i. stride or step duration and walking cadence, but the estimate of the spatial parameters of gait are affected by a rather high incertitude [1]. The problem can be solved by the use of expensive commercial systems based on inertial measurement units. The use of webcams, on the contrary, has reduced timing accuracy, due to the low frame rate that characterizes this low cost instrumentation, but can give satisfactory accuracy in the estimate of spatial gait parameters. If a markerless approach is adopted, an acceptable level of accuracy can be reached at the expense of a high computational effort and making particular care at the measurement environment (f.i. lighting conditions, use of suitable clothing, or background characteristics) [4,5]. Much better results can be obtained by the use of webcams and passive markers. Of course, the gold standard for the kinematic analysis of movement is constituted by optoelectronic stereo-photogrammetric systems [6]. These are very accurate both in time and space but are very expensive and require structured environments (i.e. a laboratory setting and specialized personnel). The approach followed in this paper relies in the integration of both accelerometer and photogrammetric data in a unique system, very cheap and suitable to be used in non-structured environments like ambulatory settings by non-expert users like General Practitioners. In particular the accelerometer data are captured by the inertial measurement units embedded in all last generation smartphones. The use of one webcam and colored markers allows the computation of marker position on a plane: in the present paper the walk path plane is considered. The prototype described in this paper, though able to manage up to three smartphones, has been applied for the measure of spatio-temporal parameters of gait using one smartphone, placed on the trunk at the pelvis level, and a webcam with 6 colored markers. Results have been validated by comparison with those obtained by a 3D optoelectronic stereo-photogrammetric system.

II. MATERIALS AND METHODS

The system is constituted by a laptop that controls up to 3 Android smartphones (Samsung S4), a webcam (Logitech HD Pro C920) and an optional wi-fi access point. For last generation smartphones this latter is not strictly necessary, as any smartphone can act as a router and as a sensor as well. The system architecture is shown in Fig 1.

A suitable *app* collects all the inertial data from the internal sensors (3-axis accelerometer, gyroscope K330 ST Microelectronics, and YAS532 magnetic sensor Yamaha Corporation) and sends them, in wi-fi modality, to the laptop. The data transmission protocol is UDP.

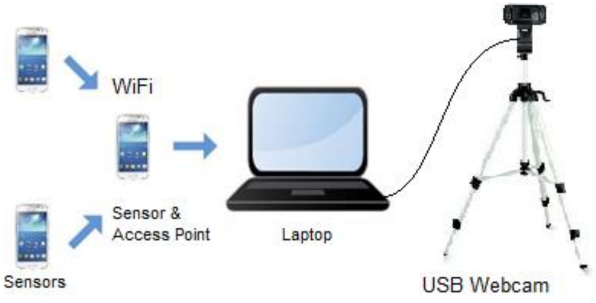

Figure 1: System architecture

V. Barone is PhD student with the Department of Information Engineering - Università Politecnica delle Marche, Ancona, IT-60131 Italy (phone: +39 3206382304; fax: +39 0712204224; e-mail: v.barone@univpm.it).

E. Maranesi is Postdoctoral Researcher with the Department of Information Engineering - Università Politecnica delle Marche, Ancona, IT-60131 Italy (e-mail: e.maranesi@univpm.it).

S. Fioretti is Associate Professor with the Department of Information Engineering - Università Politecnica delle Marche, Ancona, IT-60131 Italy (e-mail: s.fioretti@univpm.it).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

Acquisition

Remote

Record directory

Processing & Analyzing

```
S:299 abs:1.005  R pk:2.2
S:352 abs:1.580  L ts:RL 575 pk:3.5
S:427 abs:2.378  R ts:LR 798 pk:3.9
S:498 abs:3.147  L ts:RL 769 pk:4.1
S:566 abs:3.889  R ts:LR 742 pk:2.6
S:641 abs:4.689  L ts:RL 800 pk:3.7
S:717 abs:5.503  R ts:LR 814 pk:3.4
S:809 abs:6.471  L ts:RL 968 pk:1.8
```

Num Steps: 8

FPS: 29,95

Frame: 89,1312

Time: 3,147

Sample: 498

Measure

- ○ Test
- ○ Left near point: x=174 y=812
- ○ Right near point: x=1808 y=824
- ○ Left far point: x=411 y=522
- ○ Right far point: x=1590 y=516

Left near to Right near: 160

Left near to Left far: 60

Left far to Right far: 160

Right near to Right far: 60

Point

- ● Color1
- ○ Color2
- ○ Color3
- ○ Top: 550
- ○ Bottom: 750
- ○ Edit

Marker

- ○ 1 HR
- ● 2 HL
- ○ 3
- ○ 4
- ○ 5 TP
- ○ 6
- ○ 7
- ○ 8

Active ☑

Max color err: 10

Link marker: 0

Good: 0

Fail: 0

Frame start: 36

Frame end: 152

```
Side:L  x: 157.7 y: 41.1
Side:R  x: 102.6 y: 26.2
Side:L  x: 55.9 y: 45.5
Side:R  x: 12.4 y: 27.2
Synchronism done, first step from photogrammetric system is 1
```

Channel to be calibrated: 0

		Acc X	Acc Y	Acc Z
●	Chan 0	-0.040	0.059	9.831
○	Chan 1	0	0	0
○	Chan 2	0	0	0

		Pitch	Roll
●	Chan 0	-0.388	0.467
○	Chan 1	0	0
○	Chan 2	0	0

```
Open video directory on c:\imu_tmp
Sample loaded: C:\mio\universita\imu\log\tarature4\1\Acc_Orient.csv
chan: 0 imei: 2334278 Correction matrix natural acceleration loaded
0.954651 -0.129724 -0.028168
0.011202 1.012809 -0.002027
0.031983 0.034977 0.984142
chan: 0 imei: 2334278 Correction vector orientation loaded
```

Figure 2: Some windows of software user interface

The acquisition frequency is about 50 Hz; the sampling frequency is variable because it depends on the real-time characteristics of Android operating system. In order to know exactly the time instant of the sensor data acquisition, the *app* associates to each data-packet the absolute system time before sending it to the laptop. The *app* needs to know only two parameters: the IP address of the laptop and the identification number of the smartphone (N=0,1,2). The data acquired from each smartphone are shown graphically on the laptop in real time. In this work, we used only one Samsung S4 smartphone positioned on the dorsal side of the trunk at the pelvis level and held by bandages firmly strapped around the pelvis.

In all tests, a male healthy subject (52 year old, 1.78 m height, and 112 kg mass) walked on regular shoes wearing his usual clothing.

The characteristics of the webcam are: capture rate 30 fps and maximum resolution 1920x1080 pixels. The camera is positioned orthogonally to the motion of the subject pointing downwards at an angle of approximately 40 degrees from the horizontal plane. The measurement field was 3 m wide. The laptop allows to show in real time the images taken by the camera. At the end of the acquisition period, the system automatically splits the video into frames, synchronizes the first frame with the starting time instant of the inertial sensor data acquisition, and reads on the video file the number of video frames per second (Fig 2, L label) in order to maintain synchronization of the video with the acceleration tracks. Passive reflective colored markers are employed. They are placed on the heels (Figure 2, C and D labels). The aim is to track these markers and to derive the gait spatial parameters, i.e. right and left step length and width. Of course, it is not possible to acquire 3D data by one camera, but it is possible to compute the position of a point in a plane. By applying a passive reflective marker on the lateral side of the heel, and determining the instant in which the marker is at the minimum distance from the floor, the system automatically measures on the corresponding video frame the coordinates, in pixels, of the center of the marker relative to the frame reference system. To convert the pixel coordinates into corresponding spatial coordinates on the horizontal walk path plane, at least four markers lying on this plane are necessary (Fig.2, B label): all these markers must be visible from the camera and, to minimize error, should be positioned at the greatest possible distance. The spatial position of these four markers has to be precisely known. Moreover, the four markers have to belong to the same plane of the heel marker during the foot contact with the floor. It is then possible to compute a homographic transformation matrix that allows to derive from the pixel coordinates of a point the position of that point in the walk path plane. The spatial gait parameters are computed taking into account the coordinates of right and left heel markers at the instant in which each marker is at the minimum distance from the floor.

A. Accelerometer calibration

For the static calibration of the acceleration sensor, the system guides the operator through three measures:

- smartphone lying on its left side: let's indicate the ideal acceleration vector with $\mathbf{a}_1 = \begin{bmatrix} g & 0 & 0 \end{bmatrix}^T$ and the actual measure with $\mathbf{a}_1^r = \begin{bmatrix} a_{1x}^r & a_{1y}^r & a_{1z}^r \end{bmatrix}^T$
- smartphone held vertical: the ideal acceleration vector should be $\mathbf{a}_2 = \begin{bmatrix} 0 & g & 0 \end{bmatrix}^T$ while the actual measure is $\mathbf{a}_2^r = \begin{bmatrix} a_{2x}^r & a_{2y}^r & a_{2z}^r \end{bmatrix}^T$
- smartphone held horizontal with the screen upwards: the ideal acceleration vector should be $\mathbf{a}_3 = \begin{bmatrix} 0 & 0 & g \end{bmatrix}^T$ while the actual measure is $\mathbf{a}_3^r = \begin{bmatrix} a_{3x}^r & a_{3y}^r & a_{3z}^r \end{bmatrix}^T$

These steps are facilitated by a dedicated support equipped with a bubble level.

On the basis of the data acquired in these steps the system creates a calibration matrix **T** (see Fig.2, I label), based on the assumption that in each position the corresponding acceleration axis should sense an acceleration equal to *g* (the gravity one). The **T** matrix is given by

$$\mathbf{T} = g\begin{bmatrix} a_{1x}^{r} & a_{1y}^{r} & a_{1z}^{r} \\ a_{2x}^{r} & a_{2y}^{r} & a_{2z}^{r} \\ a_{3x}^{r} & a_{3y}^{r} & a_{3z}^{r} \end{bmatrix}^{-1}$$

Indicating with $\mathbf{a}_m$ the generic acceleration measure, the corrected one is computed as: $\mathbf{a}_{corr} = \mathbf{T}\mathbf{a}_m$

Although the calibration procedure is very fast, it is not always necessary; in fact during the calibration phase the system associates the calibration matrix to the unique IMEI code of each smartphone, so that, after the first calibration, the system is able to automatically recognize the sensor and the associated matrix.

It is possible to verify the correctness of the measure in the calibration window (Fig 2, J label)

B. *Heel Strike determination*

The time instant relative to the heel strike event is determined looking at the vertical component of acceleration. A double-threshold algorithm is used to determine the highest peaks present in the accelerometer trajectory that are considered due to this gait event (Fig 2, A label) [7]. The lower threshold level is computed on the basis of the noise superimposed on the accelerometer signal in static condition; the highest threshold level is determined in an adaptive way starting from a value of 50% of the maximum acceleration value, in order to avoid missing steps.

Right and left steps are determined making reference to the lateral component of acceleration.

C. *Webcam calibration*

The correspondence between the image plane coordinates $\mathbf{p}_i=(x_i,y_i)$ and the walking plane coordinates of *i*-th point $\mathbf{P}_i=(X_i,Y_i)$, as seen by the webcam, is a projective homographic transformation that, in homogeneous coordinates, can be modeled as:

$$\begin{bmatrix} \lambda x_i \\ \lambda y_i \\ \lambda \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix}$$

Dividing the first and second equations by the third one, it follows:

$$x_i = \frac{p_{11}X_i + p_{12}Y_i + p_{13}}{p_{31}X_i + p_{32}Y_i + 1}$$

$$y_i = \frac{p_{21}X_i + p_{22}Y_i + p_{23}}{p_{31}X_i + p_{32}Y_i + 1}$$

Multiplying the two members of the above equations by the same denominator and writing the two resulting equations in matrix form, one obtains:

$$\mathbf{A}_i\mathbf{v} = \mathbf{b}_i$$

where:

$$\mathbf{A}_i = \begin{bmatrix} X_i & Y_i & 1 & 0 & 0 & 0 & -x_iX_i & -x_iY_i \\ 0 & 0 & 0 & X_i & Y_i & 1 & -y_iX_i & -y_iY_i \end{bmatrix}$$

$$\mathbf{b}_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{31} & p_{32} & p_{21} & p_{22} & p_{23} \end{bmatrix}^T$$

The eight parameters of the projective transformation can be estimated in a least squared method if one disposes of at least four markers of known coordinates in the walk path plane. The four markers are placed at the vertices of a quadrangle as shown in Fig 2 (B label); Fig.2 (E label) shows the pixel coordinates of the four vertices, while Fig.2 (F-label) shows the lengths of the quadrangle sides. The estimate of the unknown vector **v** is given by:

$$\mathbf{v} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{B}$$

where:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_n \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{bmatrix}, \text{ and } n \geq 4$$

Once estimated the eight projective parameters, the coordinates of a generic point $\mathbf{P}_i$ in the walking plane can be computed solving the following equation system:

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} p_{11} - p_{22}x_i & p_{12} - p_{23}x_i \\ p_{31} - p_{22}y_i & p_{32} - p_{23}y_i \end{bmatrix}^{-1} \begin{bmatrix} x_i - p_{13} \\ y_i - p_{21} \end{bmatrix}$$

D. *Marker tracking*

The detection algorithm searches, for each marker, and within each frame, all pixels that belong to the set of colors identified by an appropriate color sampling previously done by the user [8, 9]; i.e. from three color samples, chosen for each marker, in three different frames, and taking into account a user defined tolerance for the maximum color error (Fig. 2, G label), the system determines a range for the RGB values within which the marker has to be found. A second algorithm refines the search, computing the average value of the pixel coordinates of the points identified above, and through an iterative process discards the points with greater distance from the center. The procedure runs in a completely automatic manner.

III. RESULTS

The spatio-temporal parameters (step length, width and duration) obtained by the system (Fig.2, H and K labels) have been validated with respect to those obtained by an optoelectronic stereo-photogrammetric system (6-camera Elite BTS, 100 Hz). The markers used for both systems were the same. Tables I-III show specific results relative to three steps of one walking trial performed by a healthy subject.

TABLE I. COMPARISON OF STEP-LENGTHS MEASURED BY THE SYSTEM AND BY THE GOLD-STANDARD SYSTEM IN A THREE STEP GAIT TRIAL

Step Length (SL) (cm)	**Comparison with the gold-standard**		
	System	***Stereometric system***	***Error***
Right SL	55.8	55.3	0.5
Left SL	47.4	46.6	0.8
Right SL	43.4	43.6	-0.2
RMS Error (cm)			0.56

TABLE II. COMPARISON OF STEP-WIDTHS MEASURED BY THE SYSTEM AND BY THE GOLD-STANDARD SYSTEM IN A THREE STEP GAIT TRIAL

Step Width (SW) (cm)	**Comparison with the gold-standard**		
	System	***Stereometric system***	***Error***
Right SW	14.9	14.9	0
Left SW	20.0	19.3	0.7
Right SW	18.5	18.3	0.2
RMS Error (cm)			0.42

TABLE III. COMPARISON OF STEP-DURATIONS MEASURED BY THE SYSTEM AND BY THE GOLD-STANDARD SYSTEM IN A THREE STEP GAIT TRIAL

Step Duration (SD) (ms)	**Comparison with the gold-standard**		
	System	***Stereometric system***	***Error***
Right SD	798	820	22
Left SD	769	750	-19
Right SD	742	730	-12
RMS Error (ms)			18.1

In ten walking trials the RMS error for the right and left step length was 0.58 and 0.40 (cm) respectively, while the RMS error for the step duration was 19.5 ms. The procedure was repeated on 5 subjects (aged 25-55 years): the average errors for the right and left step length were 0.62 and 0.51 cm respectively, whilst for the right and left step widths were 0.71 and 0.64 cm respectively.

IV. DISCUSSION AND CONCLUSIONS

This paper describes the realization of a prototype system based on smartphones and a webcam for the estimation of spatio-temporal parameters of gait. The system is characterized by a very low cost and by a very high level of automation. It has been thought to be used by non-expert users. The inertial unit embedded in the Android smartphone is exclusively used for the timing of gait events because accelerometer data are very accurate to this purpose; step length and width are determined by the photogrammetric procedures associated to the use of the webcam. A very high accuracy in the estimate of these spatial gait parameters has been reached, comparable with that obtainable by the use of sophisticated and highly expensive systems like the optoelectronic stereo-photogrammetric instrumentation. It is worth noting that the declared average percentage error for the estimate of the mean step length by the use of an inertial measurement units placed on the pelvis [1] is about 25%, while in our case it is about 1%. The timing error resulted to be of the order of 20 ms, as expected, due to the sampling frequency of the accelerometer. The present results refer to a system that utilizes only two markers placed on the heels; a forthcoming development of the system will take into account markers placed on the fifth metatarsal heads so that it will be possible to estimate the walking phases in a complete manner. Though this prototype system has been validated on a limited number of healthy subjects each one tested on a limited number of steps, it is being actually tested for treadmill walking in order to monitor the variability of spatio-temporal gait parameters during rehabilitation exercises.

ACKNOWLEDGMENT

The authors wish to thank Dr. M.C. Brambilla and Eng. F. Verdini for their useful suggestions.

REFERENCES

[1] W. Zijlstra, A.L. Hof, “Assessment of spatio-temporal gait parameters from trunk accelerations during human walking”. *Gait & Posture,* vol 18, pp. 1–10, October 2003

[2] M. Henriksen, H. Lunda, R. Moe-Nilssen, H. Bliddal, B. Danneskiod-Samsøe, “Test–retest reliability of trunk accelerometric gait analysis”. *Gait & Posture*, pp. 288–297, June 2004

[3] E. Martin, “Novel method for stride length estimation with body area network accelerometers”. *Biomedical Wireless Technologies-Networks, and Sensing Systems (BioWireleSS)*, pp. 79-82, January 2011

[4] E. Surer, A. Cereatti, E. Grosso, U. Della Croce, “A markerless estimation of the ankle–foot complex 2D kinematics during stance”. *Gait & Posture*, pp. 532-537, February 2011

[5] R. Poppe, “Vision-based human motion analysis: An overview”. *Computer Vision and Image Understanding,* vol. 108, pp.4–18, October 2007

[6] L. Chiari, U. Della Croce, A. Leardini, A. Cappozzo , “A Human movement analysis using stereophotogrammetry Part 2: Instrumental errors”. *Gait and Posture,* vol. 21, 2005, pp. 197–211

[7] K. Tran, T. Le ; Tien Dinh, “A high-accuracy step counting algorithm for iPhones using accelerometer”. *Signal Processing and Information Technology (ISSPIT)*, pp. 213–217, December 2012

[8] I.E. Sampe, N. Amar Vijai, R.M. Tati Latifah, T.Apriantono, “A Study on the effects of lightning and marker color variation to marker detection and tracking accuracy in gait analysis system”. *International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME*E), pp.1-5, November 2009

[9] R.C. Gonzales, R.E. Woods, *Digital Image Processing. Ed. 3.* Prentice Hall, 2007