

Denoising Human-Motion Trajectories Captured with Ultra-Wideband Real-time Location System

Rok Piltaver, Božidara Cvetković, and Boštjan Kaluža
 Department of Intelligent Systems, Jožef Stefan Institute
 Jamova cesta 39, 1000 Ljubljana, Slovenia
 E-mail: {rok.piltaver, boza.cvetkovic, bostjan.kaluza}@ijs.si

Keywords: motion capture data, human motion analysis, filtering, real-time locating system

Received: September 21, 2015

Real-time locating system (RTLS) based on UWB radio technology can be used to track people performing every-day activities. However, the quality of obtained data is relatively low and, therefore it is difficult to perform a reliable advanced analysis of human motion based on it. The paper analyses the noise of RTLS measurements and suggests filtering methods that reduce the impact of the noise on the accuracy of activity recognition. The methods are based on the statistical properties of the noise and human anatomy and motion limitations. First, a rule based method for inserting missing measurement values is suggested and compared with simple insertion of the last known value. Second, an adaptive low-pass filter that reduces impulsive noise is suggested and compared with median filter. Third, a filter that ensures human motion constraints are met is suggested. In addition, an implementation of Kalman filter that can be used to estimate the missing values, estimate the velocity of movement from recorded locations, and for smooth the signal is described. The advantages and limitations of the suggested filtering approach are demonstrated on synthetic and real data. Finally, influence of each phase of the suggested filtering chain on the accuracy of activity recognition is analysed.

Povzetek: Z UWB radijskim sistemom za lociranje v realnem času je mogoče spremljati gibanje človeka pri vsakdanjih opravilih. Kvaliteta tako zajetih podatkov je relativno slaba, kar otežuje natančne analize gibanja. Prispevek analizira šum tako zajetih podatkov in predlaga postopek za zmanjšanje vpliva šuma na točnost prepoznavanje aktivnosti. Metode so osnovane na statističnih lastnostih šuma in omejitvah, ki izhajajo iz anatomije ter fiziologije človeškega telesa. Za vstavljanje manjkajočih vrednosti predlagamo postopek na osnovi pravil in ga primerjamo z vstavljenjem zadnje znane vrednosti. Za odpravi impulznega šuma predlagamo prilagodljiv nizkopasovni filter in ga primerjamo z mediano filtrom. Zadnji v zaporedju je filter, ki zagotovi, da filtrirani podatki ustrezajo omejitvam gibanja človeka. Opisan je tudi Kalmanov filter, ki vstavi manjkajoče vrednosti, oceni hitrost gibanja in odpravi splošen šum. Sistem je ovrednoten na podlagi vpliva vsakega predlaganega filtra na točnost prepoznavanja aktivnosti, prednosti in omejitve filtrov pa so prikazane na sintetičnih in realnih podatkih.

1 Introduction

There is a significant amount of research in human activity recognition, since it is important in many domains such as ambient assisted living, security, sports, and recognition of health problems. The goal of human activity-recognition algorithms [7, 13] is to build a model that maps a sequence of sensor readings (and some additionally computed features) to an activity label, such as walking, sitting, cycling, etc. Such algorithms require that there are no missing values and that the level of noise is low. An important sensor technology that provides sensor measurements, which are useful for human activity recognition, relies on real-time location systems (RTLS) that output 3-dimensional coordinates of tags attached to the human body. High-fidelity optical RTLS such as Vicon [24] and SMART [2] provide accurate measurements (± 2 mm) but often suffer from tag

mislabelling due to problems with tracking when tag occlusion happens. Furthermore, they require a line-of-sight between the tag(s) attached to the human body and several cameras. They are appropriate for use in controlled environment (laboratory or animation studio), but fail in real-world applications as they are expensive, difficult to install, and have limitations such as line-of-sight requirement and confined operational area (e.g. 2×3 m). More affordable system rely on radio technology, which makes them less privacy-invasive and cheaper, but less accurate. Systems based on ultra-wideband technology (UWB) such as Ubisense [22] achieve ± 15 cm accuracy in ideal setting, which makes human activity recognition challenging. The main problem addressed in this paper is how to denoise human-motion trajectories captured with UWB RTLS in order to improve human activity recognition. Methods described in the paper can also be applied in other applica-

tions based on UWB RTLS as well.

Denoising human-motion data captured with UWB RTLS raises several challenges. [13, 16]. First, motion-capture data may contain a certain percentage of missing values due to packet loss or temporal sensor malfunction. Second, sensor noise and environment disturbances cause a percentage of motion-capture data to have a high error –so called outliers – and unstable measurements, which corrupt the reconstruction of human-body posture. The noisy data that violates physical body constraints as well as spatial-temporal motion constraints, which in turn causes additional problems for robust human activity recognition. Finally, some essential features used in activity-recognition process that are computed from noisy measurements, such as velocity and acceleration, have an integral error term, which accumulates error over time.

This paper proposes an efficient approach for denoising human-motion trajectories that filters corrupted motion data and enforces spatial-temporal constraints of human body, which enables a more accurate computation of features used by activity-recognition model. The key idea is to apply a series of filters that address the above-mentioned challenges: (i) inserting missing values, (ii) filtering values with high error, (iii) enforcing spatial-temporal constraints of human body, (iv) smoothing the noise, and (v) estimating derived features such as velocity.

The next section presents related work and compares it with the methods suggested in this paper. Section 3 introduces the real-time locating system used in the experiments, describes how it was used to track human motion and gives detailed analysis of the sensor noise. Section 4 gives an overview of the suggested sequence of filters and explains each of them: filters for dealing with missing values are given in Section 4.1, outlier filters are discussed in Section 4.2, filter that enforces spatial-temporal body constraints is proposed in Section 4.3, and filter for smoothing and estimating velocities is given in Section 4.4. Section 5 evaluates the proposed filters. Evaluation on synthetic data is given in Section 5.1, on real data in Section 5.2 and on human activity recognition in Section 5.3. Two applications using the proposed filters are shortly described in Section 6. The paper conclusion is in Section 7.

2 Related Work

Related work from the field of signal processing provides numerous signal denoising methods. This section provides a quick overview of methods for filtering extreme values, enforcing constraints and estimating values from noisy signals.

Qui et al. [19] reviewed and evaluated various impulsive noise filtering techniques for aircraft engine sensor data. Kernel smoothing and local regression method performed best on slowly changing signals such as ramp signal with white noise and outliers. Cascaded recursive median filter performed best on the step change signals with standard de-

viation of the Gaussian noise lower than half of the change in the signal value.

Verma et al. [23] reached similar conclusions and confirmed that median filter successfully removes outliers while preserving signal features in jet engine gas path measurements.

Sul et al. [21] presented a Kalman filter framework that handles the following problems related to motion capture sensor noise: satisfaction of physical constraints inherent to human body, user-specified kinematic constraints, and noise reduction. The constraints are added to the Kalman filter as an error function that needs to be minimised. The filter also guarantees seamless motion transition between concatenated motion segments and can be used for motion generation.

Musić et al. [14] presented an Extended Kalman filter for filtering sit-to-stand-motion using low cost inertial sensors. They define dynamic human body model and fused it with sensor measurement into an Extended Kalman filter. This approach successfully estimates and filters angles between body segments, angular velocities, angular accelerations, and joint moments.

This paper focuses on measurements captured with UWB RTLS, which are known to have low accuracy. As analysed in Section 3.3, the measurements contain different types of noise, which requires combination of multiple filters. The main contribution of the paper is the complete analysis of a comprehensive set of filters that enable effective sensor readings denoising. This work is based on findings presented in our previous papers [4, 5, 6, 15].

3 Location Trajectories of UWB System

This section first introduces the ultra-wide band (UWB) real-time location system (RTLS) used in the experiments. Second, placement of RTLS tags on human body is described according to the needs of activity-recognition – the domain used for evaluation of the proposed denoising approach. Finally, a detailed analysis of UWB RTLS measurement noise is given as it defines the denoising algorithm requirements and points out the essential evaluation tests.

3.1 Ubisense Location System

A commercially available localization system Ubisense [22] was selected as the sensing component. It allows locating by tracking a set of small tags ($40 \times 40 \times 16$ mm, 25 g), which are attached to a person's body, within an area of up to 30×30 m. A sampling frequency of around 9 Hz can be achieved with no more than four tags simultaneously. Each tag communicates using ultra-wideband radio signal [25] with four to six stationary sensors, for example, mounted on the wall. To calculate the x , y and z position of a tag, both the time

difference of arrival and the angle of arrival of radio signal are used. Location accuracy of about ± 15 cm in each of the three axes can be achieved across approximately 95% of the readings in a typical open environment [22]. However, in a real-life scenarios the absolute measurement error is higher than 100 cm in 1% of measurements, which represents a significant challenge for preprocessing and filtering.

3.2 Tag placement

The effect of body tag placement on classification accuracy was studied in [12], where it turned out that in general more tags enable more accurate classification. However, given large enough noise, even an increased number of tags does not necessarily improve the results. For example, the accuracy of the activity-recognition is comparable when using eight or four tags. Nevertheless, using only one or two tags significantly impacts the classification accuracy. Based on these results and the fact that the Ubisense sampling rate for four tags is limited to 9 Hz, the tags were positioned at the following locations on the body: chest, waist, left and right ankles.

3.3 Noise Analysis

In order to successfully denoise the RTLS measurements, the analysis of noise was conducted first. The Ubisense RTLS was installed and calibrated in a 7×4 m room used for measuring the noise of the RTLS readings. The tags were placed on the following positions (black circles in Figure 1): left ankle, right ankle, waist, and chest. To analyse the noise, static measurements were collected on a grid with 0.5 m density while the person wearing the tags was faced in one of the four directions (north, east, south, west). Over 150 measurements (lasting at least 15 s) were taken at each grid location. As a result, approximately 100,000 measurements were collected at the known locations and orientations. The data was analysed with statistical methods as described below.

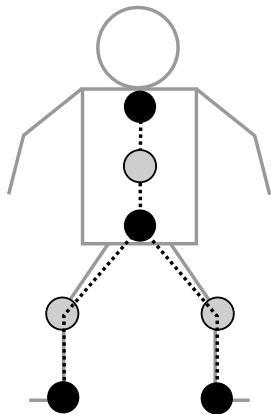


Figure 1: Positions of the RTLS tags on the body.

Figure 2 shows the observed measurement error. The

black dots represent the distance from the true position in 3D while the grey crosses represent the 2D projections. Figure 2 illustrates that the noise level makes applications such as activity recognition challenging.

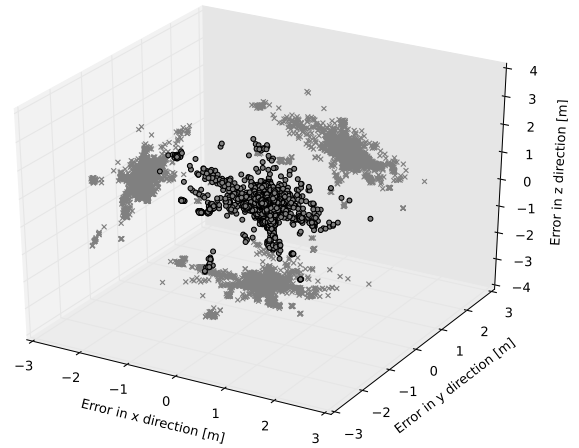


Figure 2: RTLS measurement error in 3D (black circles) with 2D projections (grey crosses).

The noise was further analysed by RTLS error histograms shown in Figure 3 for each of the three directions as well as the combined absolute error. Figure 3 show that the error is the highest in z direction (up-down), and the smallest in x direction. The Shapiro-Wilk test for normality [20] was performed for all three directions and all four tags confirming that the measurement noise is not normally distributed. This was also confirmed by Q-Q plots analysis.

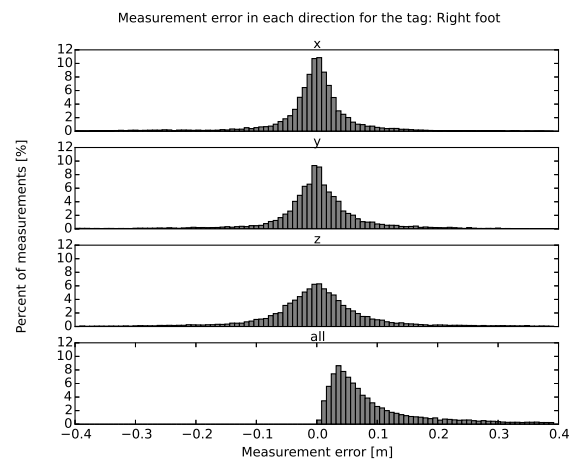


Figure 3: Histograms of RTLS measurement error in all three directions (x, y, z) and the total absolute error.

The standard deviation of measurement error in each direction is between 10.6 and 17.6 cm. Measurement error in single directions is below 10.4 to 29.9 cm (depending on the tag position and direction of error) for 95% of the measurements. The average absolute error is between 9.8 and 14.4 cm depending on the tag placement. The median of the absolute error is between 3.2 and 6.9 cm. The abso-

lute measurement error is below 22.3 to 53.7 cm (depending on tag position) for 95% of the measurements. Only one percent of measurements has absolute error higher than 1.389 m. The tags that are placed higher, for instance, chest, have lower noise level as illustrated by Figure 4.

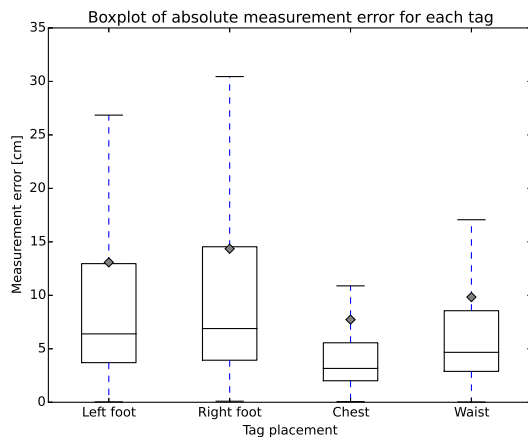


Figure 4: Boxplots (without outliers) of RTLS absolute measurement error for each tag placement. Diamonds represent mean values.

The maximal Spearman rank-order correlation coefficient between error in two directions for any of the four tags is 0.135. This shows that the error in various directions is not correlated. The maximal observed error in single direction is 3.97 m which is more than half of the longest side of the room in which the measurement noise was analysed. Additional noise analysis using auto-correlation confirmed that the noise is random, that is, there is no external process that influences the measurement error.

Up to 0.84% of the RTLS measurements are missing. However, in most cases (0.57%) only one consecutive value is missing, while more than 3 consecutive values (corresponding to 1/3 s or more) between two consecutive measurements, are missing in 0.07% of measurements.

4 Denoising Human Motion

This section propose an efficient approach for denoising human-motion trajectories that filters corrupted motion data and enforces the human-body spatio-temporal constraints thus enabling more accurate feature computation.

The key idea is to apply a series of filters as shown in Figure 5 that mitigate the identified measurement errors. First, the missing values are inserted by either inserting the last known value or by using rule-based insertion. Second, the raw RTLS signal of each tag is filtered to remove the impulse noise, using either a median filter or an adaptive low-pass filter. Third, measurements are corrected by a constraint propagation procedure in order to satisfy the following constraints: human anatomic constraints enforcing expected distances between joints and human motion constraints enforcing acceleration and velocity limits. Finally,

a Kalman filter is applied in order to smooth the signal and obtain an estimation of velocity, which is an important feature for activity recognition.

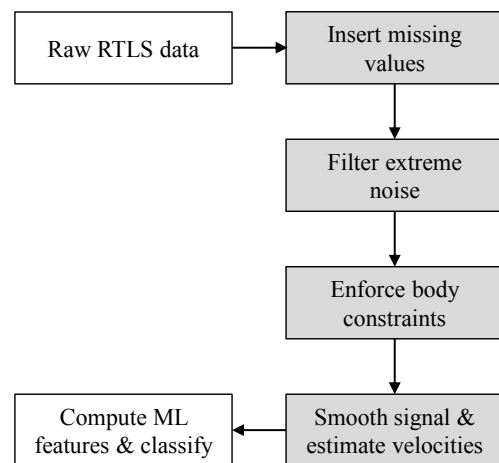


Figure 5: A series of filters for denoising the raw UWB RTLS measurement – preprocessing step for activity recognition.

4.1 Dealing with Missing Values

In most applications, the algorithms for RTLS signal analysis can be simplified if no values are missing and a constant sampling rate is used. However, data from UWB RTLS contains missing values due to packet loss, delay during transmission, sensor failure, or corrupted packets. Therefore, the first step in RTLS denoising is to insert the missing values. This paper compares two approaches described below. Other methods, such as Kalman filter (see Section 4.4) or retrospective interpolation, could be used as well.

4.1.1 Insert the Previous Value

A simple method for dealing with the missing values replaces the missing value x_t at time t with the last known value x_{t-1} according to Equation 1.

$$x'_t = x_{t-1} \quad (1)$$

This approach is simple to implement, has a constant time complexity, and can be executed online. The error of the locations inserted using this approach is reasonable if only one or a few consecutive measurements are missing and sampling frequency is high. Nevertheless, if the tag with missing value is moving, applying this method results in a signal with discontinuous derivation – velocity suddenly changes to zero, which is not desirable.

4.1.2 Rule-based Insertion

Rule-based insertion uses height of the person (*height*) and values of the non-missing tag measurements to estimate the values of the missing tag measurements. The locations of

the tags are denoted as follows: c for chest, w for waist, aR for right ankle, and aL for left ankle tag.

This approach uses a simple rule to first identify the activity of the user and then estimate the positions of the missing tags based on the activity, which is treated as the context.

The identification of the activity is done according to the height of the chest tag. If the tag is below 0.65 m the assumed activity is lying, otherwise the activity is upright. If the identified activity is lying, the values of the missing tags are estimated using one of the rules presented in Algorithm 2.

Algorithm 1: Rule-based insertion if the identified activity is lying.

```

1 if  $w$  and  $aR$  and  $aL$  are missing then
2    $w(x, y, z) = (c_x - height/3, c_y, c_z)$   $aR(x, y, z) =$ 
    $(c_x - height, c_y - 0.2, c_z)$   $aL(x, y, z) =$ 
    $(c_x - height, c_y + 0.2, c_z)$ 
3 else if  $w$  and  $aR$  are missing then
4    $w(x, y, z) = (1/2(c_x + aL_x), c_y, c_z)$   $aR(x, y, z) =$ 
    $(aL_x, 2 * c_y - aL_y, aL_z)$ 
5 else if  $w$  and  $aL$  are missing then
6    $w(x, y, z) = (c_x, 1/2(c_y + aR_y), c_z)$   $aL(x, y, z) =$ 
    $(aR_x, 2 * c_y - aR_y, aR_z)$ 
7 else if  $w$  is missing then
8    $w(x, y, z) = (c_x, c_y, 2/3c_z)$ 
9 else if  $aR$  is missing then
10   $aR(x, y, z) = (aL_x, 2 * w_y - aL_y, aL_z)$ 
11 else if  $aL$  is missing then
12   $aL(x, y, z) = (aR_x, 2 * w_y - aR_y, aR_z)$ 

```

If the identified activity is upright, the values of the missing tags are estimated using one of the rules presented in Algorithm 3.

Algorithm 2: Rule-based insertion if the identified activity is upright.

```

1 if  $w$  and  $aR$  and  $aL$  are missing then
2    $w(x, y, z) = (c_x, c_y, 2/3c_z)$ 
3    $aR(x, y, z) = (c_x - height, c_y - 0.2, 0)$ 
    $aL(x, y, z) = (c_x - height, c_y + 0.2, 0)$ 
4 else if  $w$  and  $aR$  are missing then
5    $w(x, y, z) = (c_x, c_y, 2/3(c_z + aL_z))$   $aR(x, y, z) =$ 
    $(2 * w_x - aL_x, 2 * w_y - aL_y, aL_z)$ 
6 else if  $w$  and  $aL$  are missing then
7    $w(x, y, z) = (c_x, c_y, 2/3(c_z + aR_z))$   $aL(x, y, z) =$ 
    $(2 * w_x - aR_x, 2 * w_y - aR_y, aR_z)$ 
8 else if  $w$  is missing then
9    $w(x, y, z) = (c_x, c_y, 2/3c_z)$ 
10 else if  $aR$  is missing then
11   $aR(x, y, z) = (2 * w_x - aL_x, 2 * w_y - aL_y, aL_z)$ 
12 else if  $aL$  is missing then
13   $aL(x, y, z) = (2 * w_x - aR_x, 2 * w_y - aR_y, aR_z)$ 

```

This approach is constrained by the mandatory availability of the chest tag location upon which the activity is identified.

4.2 Dealing with Outliers

The second filter used in the suggested denoising approach deals with the impulse noise (outliers). As explained in Section 3.3, a few percent of the RTLS measurements are outliers, which should be filtered before other data processing is executed. This paper compares two approaches for outlier filtering explained in the following subsections: the median filter and an adaptive low-pass filter.

4.2.1 Median Filter

Median filter is a non-linear filter that can suppress impulsive, isolated noise without blurring sharp changes in the signal [26]. The filter uses a window of sequential samples with odd length $w = 2n + 1$. At each time step t the filter returns the median of the elements in the window:

$$x'_t = \text{median}(x_{t-n}, \dots, x_t, \dots, x_{t+n}) \quad (2)$$

The only parameter of the median filter is the window length w , which introduces a delay of length $\lfloor w/2 \rfloor$. Too long window may smooth the signal too much, while too short window does not remove the high density noise. A common approach is to choose a window length that preserves the desired signal features and attenuates the impulse noise well.

The majority of computational time for the median filter is spent on calculating the median value of each window, hence an efficient median calculation is crucial for the filter run-time. While a naive approach sorts samples in each window, a histogram-based algorithms implemented with binary search trees are more efficient.

In the case of RTLS data, the median filter is applied at each tag, separately for each dimension. The filter is able to remove isolated spikes in the signal, while parts of the signal with high oscillation remain unsuppressed. However as demonstrated in Section 3.3, the Ubisense RTLS data does not contain long periods with many outliers, which makes the median filter suitable for dealing with outliers.

4.2.2 Adaptive Low-Pass Filter

Another method that filters outliers is the low-pass filter, also termed high-cut filter. It passes signals with a frequency lower than a certain cut-off frequency and attenuates signals with frequencies higher than the cut-off frequency. It provides a smoother form of a signal by removing the short-term fluctuations (outliers) and preserving the longer-term trend. The output x'_t of a discrete low-pass filter is a weighted sum of the input x_t and the preceding output x'_{t-1} for a given constant smoothing factor $0 \leq \alpha \leq 1$ that defines the cut-off frequency:

$$x'_t = \alpha x_t + (1 - \alpha)x'_{t-1} \quad (3)$$

The main idea of the adaptive low-pass filter is to set the smoothing factor α dynamically. If the tag is stationary the cut-off frequency should be lower compared to the cut-off frequency used when the tag is moving. The key challenge is how to detect whether or not the tag is moving.

This is done using movement detection algorithm described in [15]. The algorithm computes a set of attributes from time windows of RTLS data and uses them as the input to the movement detection classifier trained using a machine-learning algorithm. The attributes of the classifier are: average velocity, standard deviation of velocity, average difference between two consecutive velocities, approximate length of travelled path, standard deviation of velocity direction, and average change of direction within the time window. The accuracy of the classifier is above 96%. The miss-classification happens mainly in time windows that include a transition from a stationary state to motion and vice versa. The classifier achieves even higher accuracy during long periods without transitions: when the tag is stationary or when it moves.

The advantage of the adaptive low-pass filter is that the smoothing is dynamically adjusted. Therefore, sequences of stationary locations are smoothed more while the features of the signal during motion are preserved.

4.3 Spatial-Temporal Body Constraints

After the missing values are inserted and the outliers are filtered, more advanced filtering methods are applied. So far each tag was considered as a separated measurement. In reality, however, the tags are attached to a human body, which implies a set of constraints regarding relative tag positions and motion dynamics. In activity-recognition process as well as in other applications using RTLS data, it is usually expected that a set of measurements resembles human body proportions as well as spatial-temporal patterns in natural human motion. Therefore, we developed a filter based on iterative constraint relaxation that: (i) projects measured values in a valid domain; (ii) applies human body proportion constraints to the measured positions; and (iii) constraints spatial-temporal motion patterns.

4.3.1 Mapping Measurements to a Valid Domain

In the first step, an assumption about a valid domain of measurements is made. For example, it is expected that all the measurements are within a cuboid bounded with two extreme points \mathbf{p}_A and \mathbf{p}_B (assuming the coordinate system is aligned with the room) as shown in Figure 6. To keep the measurement \mathbf{p}_t within the expected bounds, it has to be translated to an edge (in case it is not already within the cuboid) as shown in Figure 6. The update step is:

$$\mathbf{p}'_t = \min(\max(\mathbf{p}_t, \mathbf{p}_A), \mathbf{p}_B). \quad (4)$$

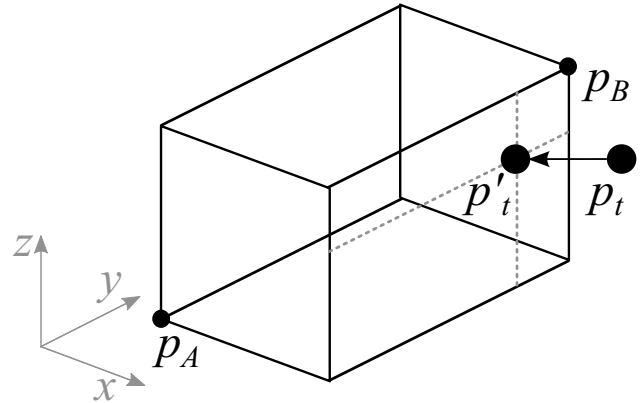


Figure 6: All the measurements are bounded with a cuboid.

4.3.2 Body Constraints

The human body is modelled using rigid-body components, which assume that there is no deformation. Rigid-body components are connected to each other with joints and form an articulated rigid body that approximates the human body as shown with dotted lines in Figure 1. The distance between any two connected joints is constant regardless of external forces exerted on it. Note that at this point on joint constraints are posed.

In our case, the four RTLS tags provide the positions of the joints (ankles, waist and chest), but do not allow the reconstruction of the skeleton displayed in Figure 1 since locations of knees and abdomen are missing. They are reconstructed as follows. Suppose there are two points A and C with known position and a joint B , which interconnects A and C , with an unknown position. Since the distances $r_A = d(A, B)$ (between A and B) and $r_C = d(C, B)$ are known, the point B then lies at the intersection of the two spheres centred at A with radius r_A and at C with radius r_C .

In general, there are three cases to consider: (i) $r_A + r_B = d$, that is, the intersection is a single point; (ii) $r_A + r_B > d$, that is, there is no solution; and (iii) $r_A + r_B < d$, that is, the intersection lies on a circle. In the second case, the point B is positioned on the line between points A and C so that the distances between points is in the same proportion as the lengths of r_A and r_B . In the third case, a new coordinate system is used to calculate the position of the point B . In the new coordinate system the first sphere is centred at the origin and the second sphere is centred at a point on the positive x-axis, at distance d from the origin, as shown in Figure 7. Subtracting the sphere equations gives a set of points representing a circular intersection of the spheres:

$$x = \frac{d^2 - r_C^2 + r_A^2}{2d} \quad (5)$$

$$y^2 + z^2 = r_A^2 - \left(\frac{d^2 - r_C^2 + r_A^2}{2d}\right)^2 \quad (6)$$

The exact position of the point B is not important, hence

an arbitrary point on the circle is picked and transformed to the original coordinate system. As explained below, the distance between joints is enforced with Equations (8) and (9).

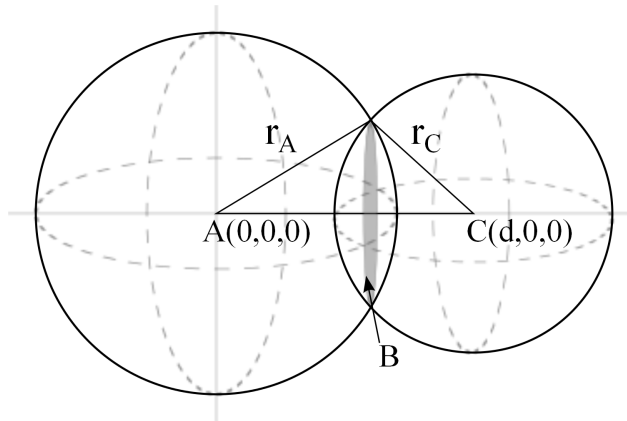


Figure 7: The result of sphere-sphere intersection is a circle.

After all the joint positions are known, constraints between the connected pairs of points can be introduced. For example, suppose the true distance between joints A and B is r_A , that is

$$\|p_A - p_B\| = r_A. \tag{7}$$

If measurements p_A and p_B violate the constraint given by the Equation (7), the position of both points is adjusted. Each point is translated along the line connecting the points for half of the error defined as the difference between the measured and the true distance as shown in Figure 8. The update is:

$$p'_A = p_A + \frac{\|p_B - p_A\| - r_A}{2\|p_B - p_A\|} (p_B - p_A) \tag{8}$$

$$p'_B = p_B - \frac{\|p_B - p_A\| - r_A}{2\|p_B - p_A\|} (p_B - p_A) \tag{9}$$

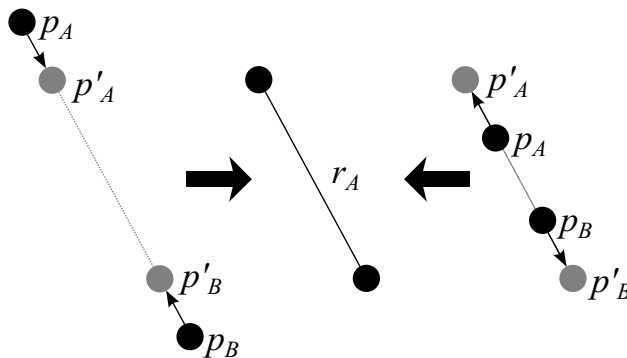


Figure 8: Move the points p_A and p_B to match the constraint given by Equation (7).

4.3.3 Spatial-Temporal Motion Patterns

In addition to the constraints introduced by human body proportions, physical motion constraints of limbs (such as velocity and acceleration) are also considered. Suppose that $a [m/s^2]$ is the greatest possible acceleration of a body joint. This implies that it can travel at most $l_{max} = (v_{t-1} + a\Delta t/2)\Delta t [m]$ in time interval Δt , where $1/\Delta t$ is the sampling frequency. Hence the next position of the joint p_t is limited with a sphere with radius l_{max} , that is

$$\|p_t - p_{t-1}\| \leq l_{max} = (v_{t-1} + a\Delta t/2)\Delta t \tag{10}$$

In the case the new position p_t is outside the sphere, the position is translated onto the edge of the sphere in the direction of the previous position as shown in Figure 9. The update step is:

$$p'_t = p_t + \frac{l_{max}(p_t - p_{t-1})}{\|p_t - p_{t-1}\|} \tag{11}$$

In order to speed up the computations, the sphere can be approximated with a cube with edge length l_{max} . In this case, the new position is limited using Equation 4 where:

$$p_A = p_{t-1} - (l_{max}, l_{max}, l_{max}) \tag{12}$$

$$p_B = p_{t-1} + (l_{max}, l_{max}, l_{max}) \tag{13}$$

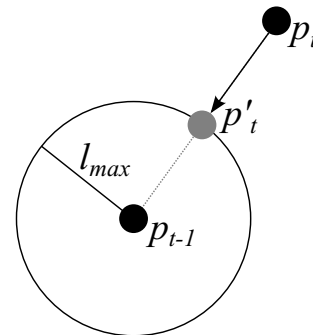


Figure 9: Constrain the maximal distance according to Equation (11).

4.3.4 Iterative Constraint Relaxation

Finally, the three types of constraints are put together. Consider $C = \{\Phi_i\}$ as a set of constraints, where $\Phi(p)$ applies the update step on point p first using Equation (11) and then using Equation (4); that is, $p' \leftarrow \Phi(p)$, while $\Phi(p_A, p_B)$ applies the update step on both points A and B using Equations (8) and (9); that is, $p'_A, p'_B \leftarrow \Phi(p_A, p_B)$. If a constrained between points A and B is not present, then $\Phi(p_A, p_B)$ does not alter the corresponding points. The algorithm 4 takes the set of constraints C and set of points P as an input and iteratively updates the values until the convergence threshold τ_c or maximal number of iterations k is reached [1].

Algorithm 3: Iterative constraint relaxation.

Data: set of constraints \mathbf{C} , set of points \mathbf{P} , maximal number of iterations k , convergence threshold τ_c

Result: filtered set of points \mathbf{P}

```

1 while  $k > 0$  and  $\Delta > \tau_c$  do
2    $\Delta = 0$ ;
3   for  $\mathbf{p} \in \mathbf{P}$  do
4     for  $\mathbf{q} \in \mathbf{P}$  do
5        $\mathbf{p}', \mathbf{q}' \leftarrow \Phi(\mathbf{p}, \mathbf{q})$ ;
6        $\Delta \leftarrow \Delta + |\mathbf{q} - \mathbf{q}'|$ ;
7        $\mathbf{q} \leftarrow \mathbf{q}'$ ;
8      $\mathbf{p}' \leftarrow \Phi(\mathbf{p})$ ;
9      $\Delta \leftarrow \Delta + |\mathbf{p} - \mathbf{p}'|$ ;
10     $\mathbf{p} \leftarrow \mathbf{p}'$ ;
11   $k \leftarrow k - 1$ ;
```

$$\begin{bmatrix} p_{x,t+1} \\ p_{y,t+1} \\ p_{z,t+1} \\ v_{x,t+1} \\ v_{y,t+1} \\ v_{z,t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \Delta_t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta_t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta_t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{x,t} \\ p_{y,t} \\ p_{z,t} \\ v_{x,t} \\ v_{y,t} \\ v_{z,t} \end{bmatrix} + \mathbf{w}_t \quad (15)$$

5 Results

The proposed denoising method was tested on synthetic data, real data, and in an activity-recognition application. The evaluation is based on the insights obtained by the noise analysis and shows that the proposed method successfully deals with missing values, outliers and general noise while preserving the original signal features.

5.1 Evaluation on Synthetic Data

The proposed denoising method was first evaluated on synthetic data. The locations of the four tags (ankles, waist and chest) were simulated as follows. The time-series starts with a period of normal positions of the four tags during standing – simulating a period with no noise. In the second phase, the signal is corrupted by moving the chest tag into outlier position – simulating a long series of outliers. Finally the position of the chest tag returns to normal.

Figure 10 shows x , y , and z coordinates of two tags. The blue line represents the synthetic signal, the green line represents the output of the median filter, the red line represents the output of the Spatio-temporal body constraints filter, and the pink line represents the output of the Kalman filter, that is, the final denoising result.

Figure 10 shows that the output of median filter follows the noisy signal due to a long period of outliers. Short periods of outlier values are successfully filtered by the median filter as explained in the next section. On the other hand, the filter that enforces the human-body proportion and motion constraints follows the noisy signal conservatively. This significantly reduces the error of the chest tag, however it introduces a relatively small error of the waist tag while the ankle tags are not affected. The Kalman filter smooths the transition. This example illustrates the benefits of the Spatio-temporal body constraints filter on the noisy data that can not be filtered appropriately using median filter alone.

5.2 Evaluation on Real Data

An example of filter response is shown in Figure 11, which shows x (top) and y (bottom) coordinates for a tag attached to the waist for $T = 600$ time steps (approximately 67 s). The vertical axis corresponds to the position of the tag in meters. The blue line represents the original location measurements, the green line represents the median filter

4.4 Smoothing and Velocity Estimation

The final step of RTLS denoising smooths the signal and estimates additional quantities such as velocity, which are used as attributes in activity-recognition process. For this task, Kalman filter [3] is used. It is a recursive linear filter for determining the best estimation of the system's state. The main assumption of the Kalman filter is that the underlying system is a linear dynamical system and that all measurement errors have a multivariate Gaussian distribution. In our case the system is a single RTLS tag moving in 3D space.

The Kalman filter performs the following three tasks: smooths the UWB measurements, estimates the velocities of tags, and predicts the missing measurements. In our case, the Kalman filter state is a six dimensional vector \mathbf{x}_t that includes positions and velocities in each of the three dimensions at time t , $\mathbf{x}_t = [p_{x,t}, p_{y,t}, p_{z,t}, v_{x,t}, v_{y,t}, v_{z,t}]^T$. The next state is estimated from the previous state using the following equation:

$$\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t, \quad (14)$$

where matrix \mathbf{F} encodes the linear dynamical system, \mathbf{B} is a control matrix and \mathbf{w}_t is noise covariance matrix. In our case, the Kalman update is simplified to Equation 15. The next state is calculated as a sum of the previous position and a product between the previous velocity and the sampling interval Δt for each direction separately. The velocities remain constant. The measurement noise covariance matrix is set based on UWB RTLS specification, while the process noise covariance matrix is fine-tuned experimentally.

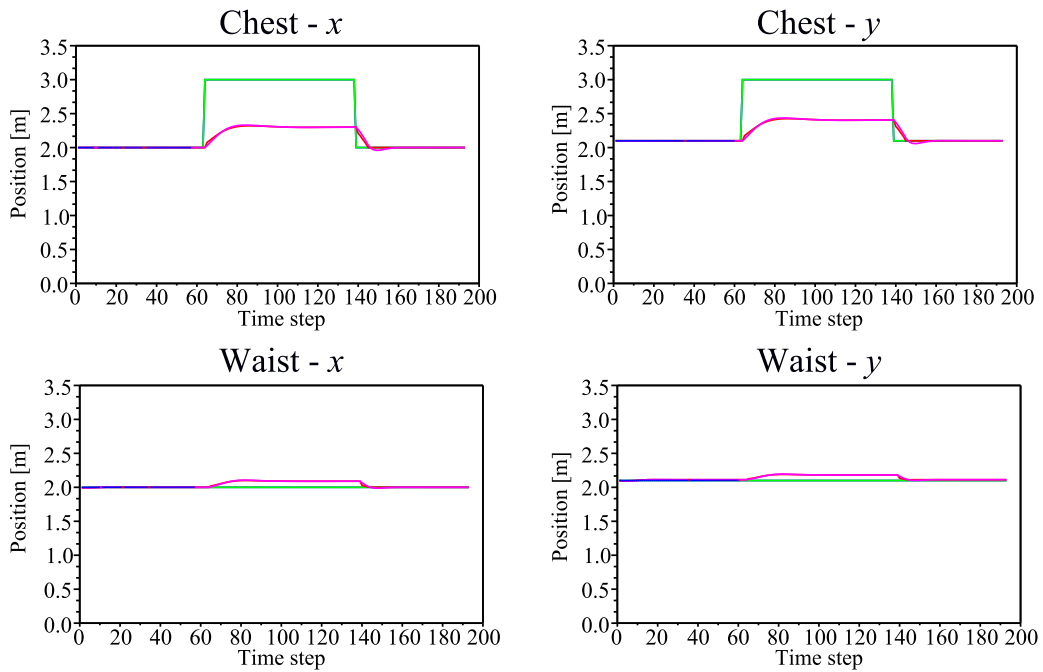


Figure 10: Response of filters on synthetic data with a long period of outliers: measurements of chest tag have a high error for extended period of time.

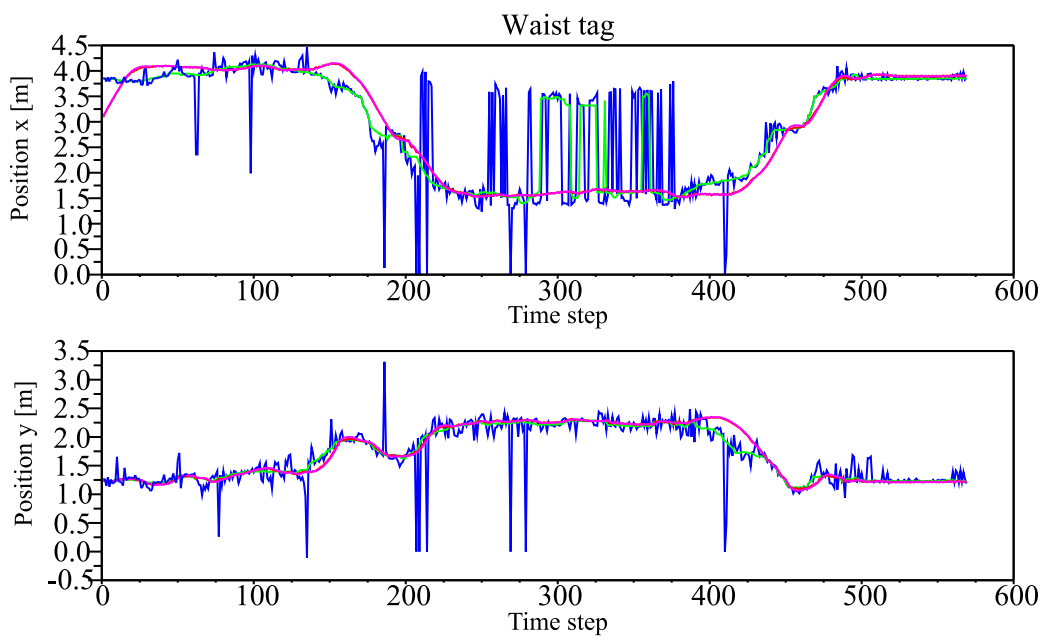


Figure 11: Filtered coordinates x (top) and y (bottom) of a tag attached to the waist.

output, and the red line represents the output of Spatio-temporal body constraints filter followed by the Kalman filter.

Figure 11 illustrates that the median filter successfully filters sparse outliers while it fails when it encounters a long sequence of outliers. In addition it shows that Spatio-temporal body constraints filter successfully correct such errors using the information about positions of the other tags.

5.3 Impact on Human Activity Recognition

In order to quantitatively evaluate the proposed denoising method, combinations of suggested filters were used as the preprocessing step that filtered the raw RTLS data for activity-recognition [6, 9, 10]. The effect on the activity-recognition classification accuracy is analysed in order to evaluate the effect of each suggested filter. The test dataset includes 32,000 to 55,000 instances for each of the ten persons, amounting to over 400,000 instances in total. Leave-one-person-out method was performed to estimate the classification accuracy using each subset of filters (used for raw-data preprocessing). The results are shown in Table 1. One-tailed paired t-test was performed to calculate the statistical significance of differences in classification accuracy (shown in Figure 12) between using various subsets of denoising filters in pre-processing step.

ID	p	r	r+m	r+l	r+m +b	r+m +b+K
A	62.5	65.6	65.4	62.8	76.1	76.8
B	61.2	61.5	60.6	58.9	72.6	73.0
C	55.2	58.0	62.6	59.7	73.7	74.4
D	65.3	67.8	68.7	65.6	74.8	76.6
E	60.4	64.4	68.4	63.4	67.7	68.1
F	64.2	64.1	62.0	61.1	74.2	74.5
G	59.0	59.0	59.2	58.2	71.9	72.2
H	56.5	59.8	65.8	61.5	69.2	72.2
I	61.9	62.1	64.6	63.3	76.3	76.5
J	63.5	64.0	64.0	63.0	75.4	75.4
\bar{x}	61.0	62.6	64.1	61.7	73.2	74.0
σ	3.1	3.0	3.0	2.2	2.7	2.6

Table 1: Accuracy of activity recognition (in %) using subsets of filters for preprocessing.

First, the two methods for inserting the missing values are compared. Classification accuracy using rule-based insertion ($\bar{x} = 62.6\%$, $\sigma = 3.0$) is significantly higher ($p = 0.005$) compared to inserting the last known value ($\bar{x} = 61.0\%$, $\sigma = 3.1$).

Second, the two methods for filtering impulse noise are compared. They are applied after the rule-based insertion of missing values. Classification accuracy using median filter ($\bar{x} = 64.1\%$, $\sigma = 3.0$) is higher compared to using the adaptive low-pass filter ($\bar{x} = 61.7\%$, $\sigma = 2.2$). However fine-tuning the parameters of the adaptive low-pass filter could improve its influence on the classification

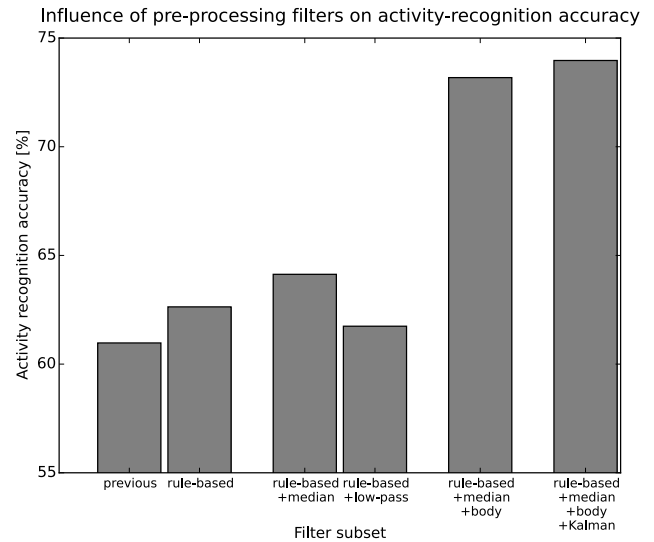


Figure 12: Comparison of activity-recognition accuracy using various subsets of filters for raw-data pre-processing.

accuracy. Furthermore, the observed difference in classification accuracy means that the adaptive low-pass filter is worse than the median filter in this application, however it does not mean that this is so in general. We argue that the adaptive low-pass filter is preferred to median filter for stationary RTLS tag positions since the accuracy of detecting stationary sequences is high ($>96\%$). Adding median filter after rule based insertion of missing values significantly improves classification accuracy ($p = 0.054$) compared to only inserting the missing values.

Third, adding the body filter after the rule-based insertion and median filter significantly ($p \approx 0$) increases classification accuracy to $\bar{x} = 73.2\%$ ($\sigma = 2.7$). Finally, adding Kalman filter at the end of the filter chain significantly ($p = 0.013$) increases classification accuracy to $\bar{x} = 74.0\%$ ($\sigma = 2.6$).

6 Applications

The proposed set of filters was successfully used in two applications. The first is an intelligent system for surveillance of personnel and equipment movement that triggers an alarm when unusual or forbidden activities are detected. The second one is a care system that detects abnormal events (such as falls) or unexpected behaviour that may be related to a health problem in elderly people. Each of the two application and the importance of denoising for robust operation of the application is briefly described below.

The first application is Commander's Right Hand [16, 17, 18]. It is an intelligent system for surveillance of movement of personnel and equipment in high security indoor environment. It uses a RTLS system to monitor movement of personnel and important equipment. It learns a model of the usual behaviour and compares it with the currently observed movement in order to detect abnormalities.

The goal is to alarm the commander about unusual and forbidden activities and enable centralized overview of monitored environment and analysis of past events. Comparison of motion detected by the RTLS and the motion detected by intelligent video surveillance detects motion that is not monitored by the RTLS, which triggers an alarm. Furthermore, the expert system module enables a simple way of specifying prohibited behavior in terms of forbidden motion patterns.

The system uses the filters for inserting the missing values, filtering the outliers, detecting motion, identifying basic activities (lying, sitting, and standing), smoothing the motion trajectories and estimating movement velocities. Only one RTLS tag per person is used in the system, therefore the Spatio-temporal body constraints filter can not be applied.

The second application is the Confidence system [6, 8, 10, 11]. It collects RTLS information about movement of an elderly that lives home alone and wears four tags positioned as described in this paper. The short term motion analysis detects unusual events such as a fall or an accident. The system triggers a rapid actuation of the health services, which decreases the negative consequences of the accident (worsening of injuries, psychological impact of being alone and injured, etc.). The long term motion analysis detects deviations in motion patterns and elderly habits which often correspond to changes in persons health. For instance, when the person's health status is worsening, there is usually less activity. It reflects in longer periods of lying and sitting, less walking and slower speed of walking and general movement. In addition, frequency of visits to the toilet often increases. When such deviation from normal behaviour is detected, the system notifies the caretaker in order to check on the elderly.

The Confidence system [6] uses the denoising method described in this paper as a preprocessing step for motion analysis described above. The denoising provided by the proposed filters significantly improves performance of activity recognition and modelling of the usual behaviour as well as simplifies motion analysis software.

7 Conclusion

The paper analyses the noise of commercially available real-time location system (RTLS) based on ultra-wide band radio technology. The results of the analysis are used to design a series of efficient denoising filters integrated into a denoising method consisting of four steps. The effect of suggested filters on the accuracy of activity recognition (based on RTLS data) is analysed to evaluate the filters. The first step inserts the missing values. A rule-based insertion method is suggested and shown to enable significantly higher accuracy compared to insertion of the last known value. The second step filters the measurements with high error – so called outliers. A low-pass filter with dynamically adapted parameters based on the motion detection is

suggested and compared with median filter. It is empirically shown that the suggested motion detection algorithm achieves accuracy of 96%, which enables adaptive filtering. The third step enforces spatial-temporal constraints of human-body proportions and motion, which additionally reduces the noise by exploiting information about location of the other tags attached to the same person. The filter significantly improves the accuracy of activity-recognition. The fourth and final step smooths the signal and estimates motion velocities, which are used as attributes for activity recognition. If the filter parameters are set correctly, the sequence of filters successfully attenuates RTLS noise while preserving the features of the observed motion. This simplifies the software for intelligent motion analysis and improves its accuracy. Furthermore, the advantages and limitations of suggested filters and their interaction are illustrated on synthetic and real data.

Acknowledgments

The experiments were made within the EU FP7 project Confidence and national project Commander's Right Hand. The authors are very thankful to both project teams, in particular to Mitja Luštrek, Matjaž Gams, Domen Marinčič, Erik Dovgan, Violeta Mirchevska, Blaž Strle, and other colleagues from the Department of Intelligent Systems at the Jožef Stefan Institute, as well as the anonymous volunteers that made experimental recordings possible and Robert Jakomin who made the RTLS noise recordings.

References

- [1] R. Barrett, M. Berry, T. F. Chan, et al. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd edition*. Philadelphia, 1994.
- [2] BTS Bioengineering Corp. *BTS SMART-DX*. <http://www.btsbioengineering.com/products/kinematics/bts-smart-dx/>, 12-03-2015.
- [3] R. E. Kalman. *A New Approach to Linear Filtering and Prediction Problems*. Transaction of the ASME – Journal of Basic Engineering 35, pp. 35–45, 1960.
- [4] B. Kaluža and E. Dovgan. *Glajenje trajektorij gibanja človeškega telesa zajetih z radijsko tehnologijo*. Proceedings of the 13th International Multiconference Information Society – IS 2009, vol. A, pp. 97–100, 2009.
- [5] B. Kaluža. *A Unified Framework for Detection of Suspicious and Anomalous Behavior from Spatio-Temporal Traces*. Informatica (Slovenia) 38(2), 2014.
- [6] B. Kaluža, B. Cvetković, E. Dovgan, H. Gjoreski, M. Gams, M. Luštrek. *Multiagent Care System to Support Independent Living*. International Journal on Artificial Intelligence Tools 23(1), 2013.

- [7] B. Kaluža, M. Gams. *Analysis of Daily-Living Dynamics*. Journal of Ambient Intelligence and Smart Environments 4(5), pp. 403–413, 2012.
- [8] B. Kaluža, V. Mirchevska, E. Dovgan, M. Luštrek, M. Gams. *An Agent-based Approach to Care in Independent Living*. Lecture Notes in Computer Science 6439, pp. 177–186, AmI 2010, 2010.
- [9] B. Kaluža, M. Gams. *An Approach to Analysis of Daily Living Dynamics*. Proceedings of the WCECS 2010, vol. 1, pp. 485–490, 2010.
- [10] B. Kaluža, M. Luštrek, E. Dovgan, M. Gams. *Context-Aware MAS to Support Elderly People*. AAMAS 2012, Valencia, Spain, June 2012.
- [11] M. Luštrek, B. Kaluža, B. Cvetković, E. Dovgan, H. Gjoreski, V. Mirchevska, M. Gams. *Confidence: Ubiquitous Care System to Support Independent Living*. ECAI 2012.
- [12] M. Luštrek, B. Kaluža, E. Dovgan, B. Pogorelc, and M. Gams. *Behavior Analysis Based on Coordinates of Body Tags*. Lecture Notes in Computer Science 5859, pp. 14–23, AmI 2009, 2009.
- [13] M. Luštrek, B. Kaluža. *Fall detection and activity recognition with machine learning*. Informatica 33(2) pp. 205–212, 2009.
- [14] J. Musić, R. Kamnik, M. Munih. *Model based inertial sensing of human body motion kinematics in sit-to-stand movement*. Simulation Modelling Practice and Theory 16, pp. 933–944, 2008.
- [15] R. Piltaver. *Strojno učenje pri načrtovanju algoritmov za razpoznavanje tipov gibanja*. Proceedings of 11th International multiconference Information Society – IS 2008, pp. 37–40, 2008.
- [16] R. Piltaver, E. Dovgan, M. Gams. *An intelligent indoor surveillance system*. Informatica 35(3) pp. 383–390, 2011.
- [17] R. Piltaver, B. Pogorelc, M. Gams. *Ambient intelligence for indoor surveillance* Proceedings of International Joint Conference on Ambient Intelligence – AMI 2011, pp. 5–8, 2011.
- [18] R. Piltaver, M. Gams. *Expert system as a part of intelligent surveillance system* Proceedings of 18th International Electrotechnical and Computer Science Conference – ERK 2009, vol. B, pp. 191–194, 2009.
- [19] H. Qiu, N. Eklund, N. Iyer, X. Hu. *Evaluation of Filtering Techniques for Aircraft Engine Condition Monitoring and Diagnostics*. Proceedings of the International Conference on Prognostics and Health Management, pp. 1–8, 2008.
- [20] S. S. Shapiro, M. B. Wilk. *An analysis of variance test for normality (complete samples)*. Biometrika 52 (3–4), pp. 591–611, 1965.
- [21] C. W. Sul, S. K. Jung, K. Wohn. *Synthesis of Human Motion using Kalman Filter*. Proceedings of the International Workshop on Modelling and Motion Capture Techniques For Virtual Environments 1537, pp. 100–112, 1998.
- [22] Ubisense GmbH. *In-building Location Systems*. <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4449097>, 12-03-2015.
- [23] R. Verma, R. Ganguli. *Denoising Jet Engine Gas Path Measurements Using Nonlinear Filters*. IEEE/ASME Transactions on Mechatronics 10(4), pp. 461–464, 2005.
- [24] Vicon Motion Systems Ltd. *Vicon Bonita*. <http://www.vicon.com/System/Bonita>, 12-03-2015.
- [25] Wikipedia. *Ultra-wideband*. <http://en.wikipedia.org/wiki/Ultra-wideband>, 12-03-2015.
- [26] L. Yin, R. Yang, M. Gabbouj, Y. Neuvo. *Weighted median filters: a tutorial*. IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing 43(3), pp. 157–192, 1996.