A route identification algorithm for assisted living applications fusing WLAN, GPS and image data

Milan Redžić, Conor Brennan, Noel E O'Connor
CLARITY: Centre for Sensor Web Technologies
Dublin City University, Ireland
Email: milan.redzic,brennanc,oconnorn@eeng.dcu.ie

***Abstract*—Complex environment and objective obstacles are causes that usually require more than one localisation modality. In this work a novel multimodal fusion method for user localisation is presented which combines GPS, images and WLAN signal strength data in order to more accurately find routes one has traversed using all these three modalities at the same time. This method outperforms each method separately and also shows superiority over fusion of other two modalities. Thus it presents a good framework for navigation and ambient assisted-living applications.**

I. Introduction

This paper addresses the automatic identification of often-traversed routes for assisted living applications using three different modalities. Such applications of using large amounts of location data can be of benefit to a variety of users. Some (runners) may wish to know how often they take a particular route whilst jogging. In caring for the elderly, allowing a mobile device to automatically determine whether they have deviated from their normal routine can trigger a notification to their carers. In the life-logging community, route matching can add important structure to the months/years of recorded daily activities. This problem is complicated by a number of factors including the need to track users seamlessly in both indoor and outdoor environments, the need for robustness to slight deviations in the precise path and speed taken along a route. Sometimes exists weak and inaccurate GPS signal due to obstacles, multipath propagation and close buildings that may cause serious errors [6], [1]; in the case of WLAN there is changing and noisy nature of its channel which gives big variation by uncluttered environment [3], [8]. For images affected with great noise, blur and big change of light at the place where they were taken may give very inaccurate results as well[2]. In this work commonly traversed routes are identified with clusters based on sensed data, two of which take the form of wireless signals: GPS and WLAN. The latter is particularly important as it can be used both indoors and outdoors. In addition an efficient image matching algorithm [2] is implemented to process data from images automatically taken along the route. In this work a finite number of routes were identified within the DCU campus. Each route was traversed many times over a period of 6 weeks and data sequences collected automatically on each occasion. Each such traversal of a route is referred to as a *trip* in what follows. Section (II) outlines the use of Multidimensional Dynamic Time Warping (MDTW) and Dynamic Time Warping (DTW) algorithm in order to automatically compare trips corresponding to specific routes based on wireless and image data sensed on each trip. Section (III) outlines the manner in which data was sensed while section (IV) presents results for each modality individually as well as results based on a fusion of the data. Also these results are visually presented in 2D space.

II. DTW and MDTW

In order to find a similarity measure for data collected during different trips the Multidimensional Dynamic Time Warping Algorithm [4], [5] was employed. The classic DTW algorithm uses a local distance measure to determine the similarity between two sequences. These sequences may be discrete signals (time-series) or, more generally, feature sequences sampled at equidistant points in time [9]. In order to compare two different features from feature space F, a local distance measure is defined: $c : F \times F \rightarrow \Re \geq 0$. To measure the similarity between two sequences of data, the first C of length I and the second T of length J, an $I \times J$ distance table D is constructed, where each element of D, $d(i, j)$, represents the local distance between C_i, the i^{th} element of C and T_j, the j^{th} element of T. Warping paths W are then calculated from the distance table, each of which consists of a set of distance table elements that define a mapping and alignment between C and T:

$$W = \left\{ (i_w(q), j_w(q)) \;\middle|\; \begin{array}{l} q = 1, ..., Q, \\ \max(I, J) \leq Q \leq I + J - 1 \end{array} \right\} \quad (1)$$

with $i_w(q) \in \{1, ..., I\}$ and $j_w(q) \in \{1, ..., J\}$. Given $(i_w(q), j_w(q))$ and $(i_w(q-1), j_w(q-1))$, the warping path is restricted by the following conditions [4]: continuity ($i_w(q) - i_w(q-1) \leq 1$ and $j_w(q) - j_w(q-1) \leq 1$), the endpoint ($i_w(1) = j_w(1) = 1$ and $i_w(Q) = I$ and $j_w(Q) = J$) and the monotonicity ($i_w(q-1) \leq i_w(q)$ and $j_w(q-1) \leq j_w(q)$). The similarity between the data sequences can be gauged by identifying the optimal warping path which minimises the overall distance. This minimised distance is given by

$$DTW(C, T) = \min_w \left(\sum_{q=1}^{Q} d(i_w(q), j_w(q)) \right) \quad (2)$$

DTW(C,T) is normalised with the length of the optimal path (compensation due fact that warping paths may depend on the paths' lengths) [4].

Since the data in this paper was multidimensional, we switch to multidimensional sequences $C(I \times V)$ and $T(J \times V)$ (V is the number of variables) and we use d_E, the extended *Euclidean distance* [4] as the local distance measure for two vectors of length V. The DTW distance between two multidimensional sequences $C(I \times V)$ and $T(J \times V)$ can be calculated recursively as [5]:

$$\begin{aligned}&DTW(C(I \times V), T(J \times V)) = d_E(C_I^V, T_J^V) + \\&min\{DTW(C((I-1) \times V), T(J \times V)), \\&DTW(C((I-1) \times V), T((J-1) \times V)), \\&DTW(C(I \times V), T((J-1) \times V))\}\end{aligned} \tag{3}$$

For GPS and WLAN data $DTW(C, T)$ can be thus computed for each pair of trips, and used to populate a distance matrix. In the case of image data the elements of the distance table corresponded to the number of features, matched using the SURF algorithm [2], between every two images (one from each set). A greater weight was put on bi-directional matches since the greater level of confidence ascribed to them (the measure is $d(i,j) = 10B + U_{ij} + U_{ji}$, where B stands for the number of bidirectional matches, U_{ij} the number of unidirectional matches from the i^{th} to the j^{th} image and vice versa). The distance table is then multiplied by -1 so that the optimal path corresponds to the path with most matches [7]. To transform the number of SURF matches between two trips into the distance matrix, a mapping process needed to be defined. It should be monotonically decreasing and produce non-negative values. While there are many such functions, the reciprocal function was used for its simplicity [10].

III. Experimental Set-up

A set of training data was collected simultaneously using a SenseCam, GiSTEQ GPS device and Campaignr software (for collecting signal strengths data) installed on a N95 Nokia cellphone. The data recording was collected at regular time intervals (every 1, 15 and 30 seconds for GPS, SenseCam and Campaignr respectively). Each route was traversed many times over a period of 6 weeks, yielding 30 testing (6 routes $\times$ 5 trips) and 24 training (6 routes $\times$ 4 trips) sets of data. Signal strength information is considered to be 3-dimensional as the same 3 MAC addresses were discernible along each trip. GPS data is deemed to be 2-dimensional. That gives two data-matrices of order $N \times 3$ and $M \times 2$. The MDTW was then applied to each pair of data sequences (for each modality) [4], [5]. In the case of image data the DTW algorithm was applied to every two sets of images taken by the SenseCam.

To find the fusion matrix and the accuracy of the algorithm 1-nearest neighbor approach (1-NN) was employed. There were training and testing types of data that were collected during the experiment process. The sets of training data consist of three normalized 24×24 matrices (4 trips for the every route) and these data are different to the testing data which consist of three normalized 30×30 matrices (5 trips for the every route). For each trip the nearest trip was found and was checked whether that nearest trip belonged to the same route. The number of all succesful nearest trip matches was divided with the total number of trips which gave the accuracy of the algorithm. In order to fuse these modalities together and to calculate the accuracy of such approach, weighting coefficients w_1, w_2 and w_3 were introduced. For the training data set matrices these weights corresponded to GP, IM and S (GPS, image and signal strength training data). A new matrix

$$FZ = w_1 GP + w_2 IM + w_3 S$$

was defined and it was investigated using 1-NN what coefficient values w_{1s}, w_{2s} and w_{3s} would give maximal accuracy for the FZ under the condition of:

$$w_1 + w_2 + w_3 = 1$$

There were several different pairs of values w_{1s}, w_{2s} and w_{3s} which were used for the calculation of the accuracy of the

$$w_{1s} GPS + w_{2s} IMG + w_{3s} SS$$

matrix (GPS,IMG and SS are 30×30 testing set matrices) again using 1-NN. For the only one pair the result gave the maximum and those accuracy and coefficient values were stored.

IV. Analysis of the results

Table I clearly illustrates that GPS is the strongest individual modality. This is further emphasised by the high weight that is placed on this data source by the fusion process. As all our trips were outdoors, this was to be expected.

Figure 4 gives a visual representation of the similarities between trips in different modalities. We used the distance-matrix visualisation algorithm given in [10] to display in $2D$ a representation of the multidimensional trips and their similarities. This algorithm takes the difference between every two trips (distance matrix elements) and makes a chart (trips on the chart are presented as circles) in which the distances between them on the chart match those differences. This iterative algorithm first calculates the target distances between all the trips. Next all the trips were placed randomly on the two-dimensional chart. For every pair of trips the target distance is compared to the current distance and en error term is calculated. Then every trip is moved a small amount closer or further in proportion to the error between the two trips. This procedure is repeated many times until the total amount of error cannot be reduced by moving the trips any more. Examining the GPS results in figure 4(a), it can be seen that the fourth route (trips $16-20$) and the sixth route (trips $26-30$) do not cluster well (red and purple routes shown in fig 3). They traversed environments where the GPS signal was degraded and attenuated (shown as green circles in fig 3), due to tall buildings (the sixth route) and to part of the path going into a tunnel (the fourth route), both of which are known to affect GPS signal quality [6], [1]. The reason why the second and the fourth route failed as the image data were collected randomly during a variety of different conditions (rain/sun, morning/evening/nighttime, obstacles). While SURF features are somewhat robust to changes in lighting [2], large changes cause problems, as shown in figure 1. Consequently

Sources	w_1	w_2	w_3	Acc (%)
SS	-	-	-	56.66
IMG	-	-	-	66.66
GPS	-	-	-	80.00
GPS,IMG	0.9720	0.0280	-	80.00
IMG,SS	-	0.8050	0.1950	70.00
GPS,SS	0.9840	-	0.0160	80.00
GPS,IMG,SS	0.8310	0.0090	0.1600	**83.33**

TABLE I
TRIP CLASSIFICATION PERFORMANCE: THE LEARNED WEIGHTS FOR EACH OF THE THREE SOURCES, AND THE ACCURACY (ACC) OF THE CLASSIFIERS USING DIFFERENT SOURCES.

the results show that only 4 of the 6 routes could be properly clustered using image data alone. For the other routes image-matching performed quite well, considering the low sampling rate it used ($1/15$Hz). The format above the image describes the matching process and is given as $[U_{12}$:$U_{21}]$:$[B]$, where U_{12} is represented with red lines which connect matches on the two images, U_{21} with blue lines and B with the green lines. On its own, WLAN signal strength readings perform worst for trip classification, which was expected due to the many environmental factors that can influence signal strengths outdoors and the fact that only 3 MACs were discernible. Figure 2 shows signal readings for 3 trips from the same route, illustrating the degree of variability inherent in the readings. While each source independently has its own problems, by fusing the three modalities together, we achieved greater accuracy than any individual/combined modality.

Future work will further investigate others weighting methods especially adaptive weighting based on confidence, different fusion methods and the possibility of tracking indoor trips where GPS fails.

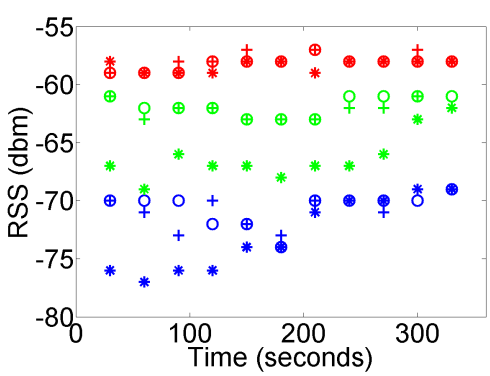


Fig. 2. Signal strengths distribution example: Data from 3 MAC addresses shown in red, green and blue, corresponding to trip 16, 17 and 18 (plotted with circles, crosses and asterisks) respectively in figure 4(b). Note the discrepancy in signal strengths for trip 18 compared to others.

V. ACKNOWLEDGMENTS

This work is supported by Science Foundation Ireland under grant 07/CE/I1147.

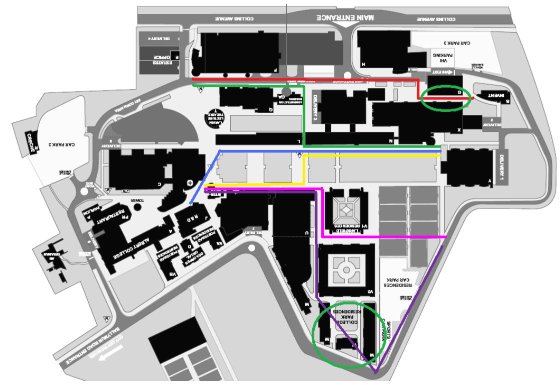
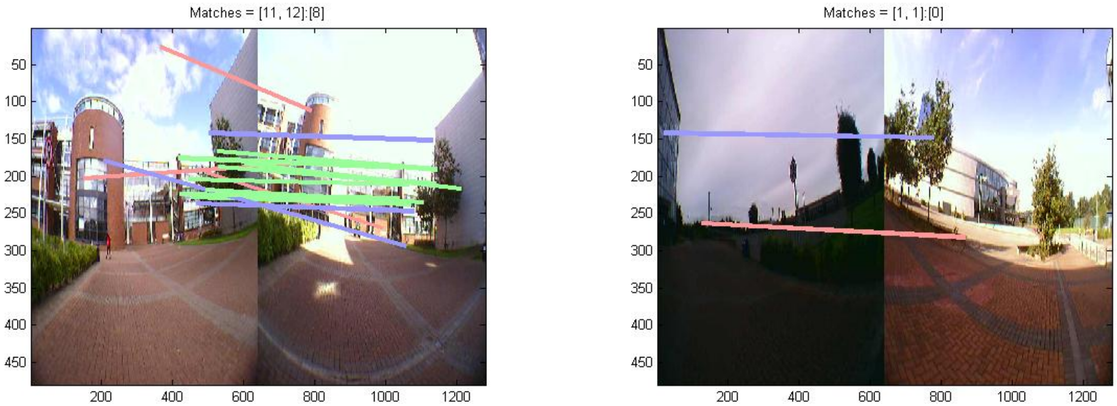


Fig. 3. DCU map with the routes overlaid

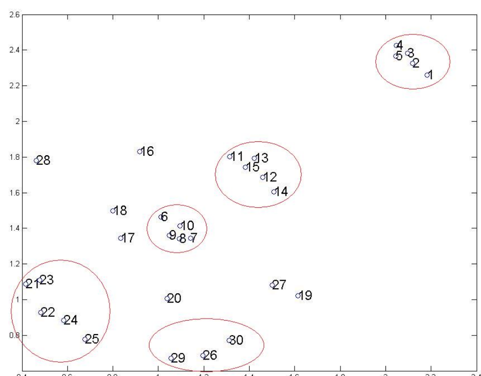
REFERENCES

[1] G. M. et al. Degraded gps signal measurements with a stand-alone high sensitivity receiver. In *Proceedings of National Technical Meeting*. The Institute of Navigation, January 2002.
[2] H. B. et al. Surf: Speeded up robust features. *Computer Vision and Image Understanding (CVIU)*, 110(3):346–359, August 2008.
[3] K. W. et al. A practical evaluation of radio signal strength for ranging-based localization. *SIGMOBILE Mob. Comput. Commun. Rev.*, pages 41–52, January 2007.
[4] M. H. K. et al. Using dynamic time warping for online temporal fusion in multisensor systems. *Information Fusion*, 9(3):370–388, January 2008.
[5] M. V. et al. Indexing multidimensional time-series. *The VLDB Journal*, July 2006.
[6] R. K. et al. Gps signal fading model for urban centres. *IEE Proc.-Microw. Antennas Propag*, August 2003.
[7] R. K. et al. Human motion recognition using isomap and dynamic time warping. *ICCV Workshop on Human Motion*, 2007.
[8] M. Klepal, M. Weyn, W. Najib, I. Bylemans, S. Wibowo, W. Widyawan, and B. Hantono. Ols: opportunistic localization system for smart phones devices. In *MobiHeld '09: Proceedings of the 1st ACM workshop on Networking, systems, and applications for mobile handhelds*, August 2009.
[9] M. Muller. *Information Retrieval for Music and Motion*. Springer, Berlin, Germany, 2007.
[10] T. Segaran. *Programming Collective Intelligence:Building Smart Web 2.0 Applications*. O'Reilly Media, Cambridge, Massachusetts, 2007.

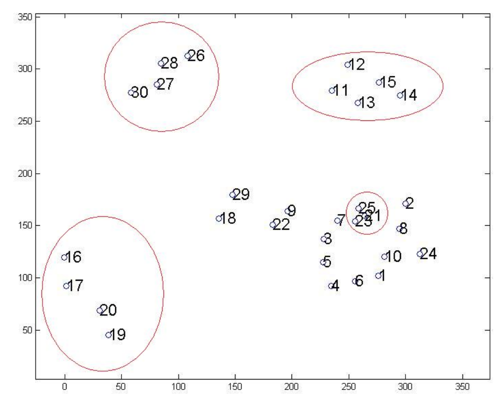


(a) (b)

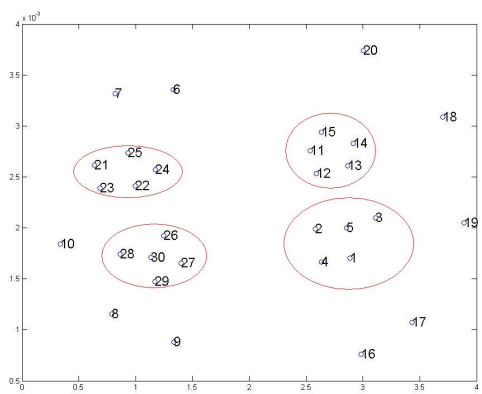
Fig. 1. Image matching examples for trips taken in different light conditions: (a) in similar lighting conditions, many matches are found, (b) matching is more difficult due to lighting changes.



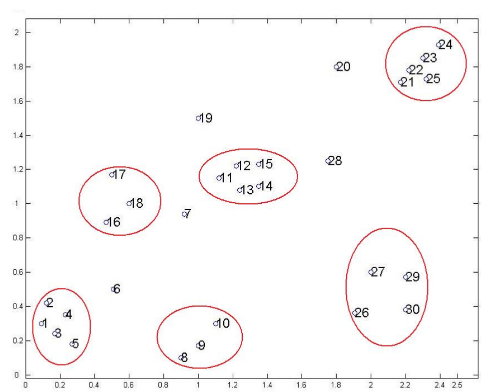
(a) GPS



(b) Wireless signal strengths (WLAN)



(c) Image-based matching



(d) Fusion of all three sources

Fig. 4. Visualisation of trip similarity using different localisation sources: We project the distances between trips into 2-dimensions for visualisation. Circles are drawn to show trips from the same route that tightly cluster together. Route 1 contains trips {1,..,5}, Route 2 contains trips {6,..,10}, etc.