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 algorithms use 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 different features from feature space \( F \), a local distance measure is defined: \( c : F \times F \rightarrow \mathbb{R} \geq 0 \). 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 \( T \):

\[
W = \left\{ (i_w(q), j_w(q)) \mid q = 1, \ldots, Q, \max(I, J) \leq Q \leq I + J - 1 \right\}
\]

(1)

with \( i_w(q) \in \{1, \ldots, I\} \) and \( j_w(q) \in \{1, \ldots, J\} \). 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

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

(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]:

\[
\begin{align*}
DTW(C(I \times V), T(J \times V)) &= d_E(C_I^y, T_J^y) + \\
& \min \left\{ DTW(C((i-1) \times V), T(J \times V)), \\
& DTW(C(I \times V), T((J-1) \times V)), \\
& DTW(C(I \times V), T((J-1) \times V)) \right\}
\end{align*}
\]

(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 cellphone, GiSTEQ GPS device and Campaignr software 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 \( \times \) 24 matrices (4 trips for the every route) and these data are different to the testing data which consist of three normalized 30 \( \times \) 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 successful 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 sets matrices these weights corresponded to \( GP, IM \) and \( S \) (GPS, image and signal strength training data). A new matrix

\[
FZ = w_1GP + w_2IM + w_3S
\]

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

\[
w_{1s}GP + w_{2s}IM + w_{3s}SS
\]

matrix (\( GP, IM \) and \( SS \) are \( 30 \times 30 \) 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 an 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
TABLE I
TRIP CLASSIFICATION PERFORMANCE: THE LEARNED WEIGHTS FOR EACH OF THE THREE SOURCES, AND THE ACCURACY (ACC) OF THE CLASSIFIERS USING DIFFERENT SOURCES.

<table>
<thead>
<tr>
<th>Sources</th>
<th>w1</th>
<th>w2</th>
<th>w3</th>
<th>Acc (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>56.66</td>
</tr>
<tr>
<td>IMG</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>66.66</td>
</tr>
<tr>
<td>GPS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>80.00</td>
</tr>
<tr>
<td>GPS,IMG</td>
<td>0.9720</td>
<td>0.0280</td>
<td>-</td>
<td>80.00</td>
</tr>
<tr>
<td>IMG,SS</td>
<td>-</td>
<td>0.8050</td>
<td>0.1950</td>
<td>70.00</td>
</tr>
<tr>
<td>GPS,SS</td>
<td>0.9840</td>
<td>-</td>
<td>0.0160</td>
<td>80.00</td>
</tr>
<tr>
<td>GPS,IMG,SS</td>
<td>0.8310</td>
<td>0.0090</td>
<td>0.1600</td>
<td>83.33</td>
</tr>
</tbody>
</table>

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 (1/15Hz). 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.

V. ACKNOWLEDGMENTS

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

REFERENCES

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.

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.