Maximum a Posteriori (MAP) Estimation of the Route Taken Through a Street Graph

Stuart B. Heinrich
heinrich@lumatic.com

February 28, 2012

Abstract

Estimating the route taken through a street graph from a sequence of consecutively geotagged coordinates is a problem with several possible applications, such as accurately matching panoramas to the streets that they correspond to for street view presentation. In this paper, we show how several different statistical priors can be combined into a Bayesian model that can be efficiently optimized using Belief Propagation (BP).

1 Introduction

In building a system to display street-view level imagery to be associated with turn-by-turn routing instructions from a set of geotagged panoramas, it was initially assumed that selecting the best panorama to display any particular view was a decision that could be made based purely on the geotagged coordinates of the panorama.

Although largely true, it was found that this assumption can sometimes be problematic. In particular, it is ambiguous from the geotagged coordinates which road a panorama was taken from, if the panorama is located at the spatial intersection of two or more roads. This can be problematic when these roads are at different layers, such as commonly occurs when a road crosses underneath a highway, or in the more extreme case when there is a tunnel that lies directly underneath a surface level road.

Other ambiguous cases may arise from discrepancies in the geographic localization of the panoramas vs. the road database. For example, if there is a highway that runs closely parallel to a surface level road, then a slight discrepancy between the GPS coordinates measured from the panoramas and the location of the two roads as recorded in the database might cause the panoramas to be closer to the wrong parallel road. Similarly, if the panoramas were acquired by driving down a side street off a main road, if the side street is shifted by a small amount, then those panoramas might be closer to the main street than to the side street.

As a result of these aforementioned ambiguities, our initial attempts to generate consecutive street level views corresponding to a route through the street database that were based purely on geotagged coordinates would sometimes correct a panorama that was clearly taken from the wrong street.

It may seem that the problem might have been solved by the collection of additional metadata about the panoramas, such as street name. However, street names can also be ambiguous. Thus, unless the panoramas are acquired and geotagged with unique street identifiers from the same street database that is later used for routing purposes, the problem persists.

For our particular problem, the only information available beyond geotagged coordinates of each panorama was an index. Although the panorama indexing scheme was unknown to us, we observed that contiguous subsequences of indices always corresponded to a sequence of panoramas in the order that they were acquired. Therefore, our overall approach was to match these subsequences of coordinates to the streetmap database by exploiting our knowledge of street connectivity. To this end, we have developed a novel maximum *a posteriori* (MAP) approach to optimally assigning panoramas to streets.

Generically, our approach can be used to match any sequence of geographic coordinates to a street map database by exploiting the connectivity of streets. This could be useful for other related problems, such as interpreting a driver's course through a street map, or matching a route that was computed from one street map database into another street map database.

2 Methods

2.1 Problem Description

Let the subsequence of n ordered coordinates to be matched be represented by $r = (\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n)$, where $\mathbf{x} \in \mathbb{R}^2$. Although these coordinates are originally represented as latitude-longitude pairs, we first transform all coordinates into a local

Euclidean coordinate frame using the latest World Geodetic System (WGS84).

The street map is represented by a graph $G = (V, E)$ with vertices V and edges E. An edge $e = (a, b)$ is a directed connection from vertex a to vertex b. For each edge $e \in E$ there is an associated way w_e that is an ordered list of n_e coordinates $\mathbf{x}^e \in \mathbb{R}^2$, $w_e = (\mathbf{x}^e_1, \mathbf{x}^e_2, \dots \mathbf{x}^e_{n_e})$. Two ways w_e and $w_{e'}$ are *connected* if there exists edges $e = (a, b)$ and $e' = (b, c)$.

We refer to an edge e as the *generator* of a coordinate $\mathbf{x}$ if this coordinate was geotagged while actually traversing the associated way w_e. For a particular coordinate $\mathbf{x}$, let the true generator be denoted by $\overline{e}_{\mathbf{x}}$. Thus, the full set of generators is represented by $\overline{\Theta} = (\overline{e}_{\mathbf{x}_1}, \overline{e}_{\mathbf{x}_2}, \dots \overline{e}_{\mathbf{x}_n})$.

Our assumptions are as follows:

1. A street that is closer to a coordinate has a higher likelihood of being the matching street for that coordinate.
2. A street that has a similar tangent vector at its closest point to the coordinate to the tangent vector of the sequence at the coordinate has a higher likelihood of being the matching street for that coordinate.
3. It is highly likely that two consecutive coordinates will be on the same street.
4. It is highly unlikely that two consecutive coordinates will match streets that are not directly connected in the graph structure, as this is a non-viable path through the graph structure. However, we must admit for this possibility, because a vehicle may have traveled off road, or the graph structure may be imperfect.
5. It is not unlikely that two consecutive coordinates will match streets that are directly connected in the graph structure.

2.2 Statistical Model

The first two assumptions can be independently treated as prior probability models. Based on the first assumption only, if we assume that geotagging measurement error is normally distributed with standard deviation σ, then the probability (or likelihood) of measuring a coordinate $\mathbf{x}$ while traversing edge e is given by

$$P_1(\mathbf{x}|e) = \frac{1}{2\pi\sigma^2} e^{-d(\mathbf{x}, w_e(\mathbf{x}))^2/(2\sigma^2)}, \qquad (1)$$

where $w_e(\mathbf{x})$ is the interpolated coordinate on way w_e that is closest to $\mathbf{x}$.

Based on the second assumption only, if we assume that the angle between the tangent vector of successive geotagging measurements and the tangent vector of the way at the closest point to the measurement is normally distributed with standard deviation σ_t, then the likelihood of measuring a coordinate $\mathbf{x}$ while traversing edge e is given by

$$P_2(\mathbf{x}|e) = \frac{1}{2\pi\sigma^2} e^{-(T_r(\mathbf{x}) \cdot T_{w_e}(w_e(\mathbf{x})))/(2\sigma_t^2)}, \qquad (2)$$

where $T_r(\mathbf{x})$ is the tangent vector of sequence r computed at the point $\mathbf{x}$ on r.

The latter three assumptions can be encoded via a piecewise probability function that describes the probability of occurrence for two successive street edges,

$$P_3(e, e') = \begin{cases} 1 & \text{e = e'} \\ 1-\epsilon & w_e \text{ and } w_{e'} \text{ are connected. ,} \\ \epsilon & \text{otherwise.} \end{cases} \qquad (3)$$

for some small value of $\epsilon << 1$. Taking into account the prior probabilities from (1) and (2) as well as the likelihood from (9), the combined posterior probability of the complete sequence of measurements is given by

$$P(r|\overline{\Theta}) = \prod_{i=1}^{n} P_1(\mathbf{x}_i|e_{\mathbf{x}_i}) P_2(\mathbf{x}_i|e_{\mathbf{x}_i}) \prod_{i=2}^{n} P_3(e_{\mathbf{x}_i}, e_{\mathbf{x}_{i-1}}). \quad (4)$$

The log-posterior probability is

$$\log P(r|\overline{\Theta}) = \sum_{i=1}^{n} (\log P_1(\mathbf{x}_i|e_{\mathbf{x}_i}) + \log P_2(\mathbf{x}_i|e_{\mathbf{x}_i})) \qquad (5)$$

$$+ \sum_{i=2}^{n} \log P_3(e_{\mathbf{x}_i}, e_{\mathbf{x}_{i-1}}) \qquad (6)$$

$$= -\sum_{i=1}^{n} \left(\frac{1}{2\sigma^2} d(\mathbf{x}, w_e(\mathbf{x}))^2 + \frac{1}{2\sigma_t^2} \left(T_r(\mathbf{x}) \cdot T_{w_e}(w_e(\mathbf{x})) \right) \right) \qquad (7)$$

$$- \sum_{i=2}^{n} D(e_{\mathbf{x}_i}, e_{\mathbf{x}_{i-1}}) + c \qquad (8)$$

for some constant value c, where

$$C(e, e') = \begin{cases} 0 & \text{e = e'} \\ \tau_1 & w_e \text{ and } w_{e'} \text{ are connected. ,} \\ \tau_2 & \text{otherwise.} \end{cases} \qquad (9)$$

and $\tau_1 = -\log(1-\epsilon)$ is some 'large' value and $\tau_2 = -\log(1-\epsilon)$ is some 'small' value. Thus, the maximum *a posteriori* (MAP) assignment is given by

$$\hat{\Theta}_{MAP} = \underset{\Theta}{\operatorname{argmin}} \left[\sum_{i=1}^{n} \left(\frac{1}{2\sigma^2} d(\mathbf{x}, w_e(\mathbf{x}))^2 \right. \right. \tag{10}$$

$$\left. \left. + \frac{1}{2\sigma_t^2} \left(T_r(\mathbf{x}) \cdot T_{w_e}(w_e(\mathbf{x})) \right) \right) + \sum_{i=2}^{n} C(e_{\mathbf{x}_i}, e_{\mathbf{x}_{i-1}}) \right]. \tag{11}$$

2.3 Belief Propagation

The minimization of (11) can be accomplished optimally via belief propagation (BP). This is an iterative method of message passing, wherein each iteration, each node sends a "message" to each of its neighboring nodes informing that neighbor of the estimated cost of adopting each possible label. Each iteration of BP takes $O(nk^2)$ time, where n is the length of the subsequence r and k is the number of possible labels (i.e., nearby streets) [1]. BP has been proven to converge to the optimal solution on acyclic networks such as this.

Specifically, if we let $\mathcal{P}$ be the set of all $\mathbf{x}$ nodes, and f_p be the *label* (i.e., edge) assigned to each $p \in \mathcal{P}$, and $\mathcal{N}(p)$ be the set of all nodes adjacent to $\mathbf{x}$ (i.e., $\mathbf{x}_{i-1}$ and $\mathbf{x}_{i+1}$), and let $C_p(f_p)$ be the negative log prior probability of p having label f_p according to (1) and (2), and let $C(f_p, f_q)$ be the negative log likelihood of consecutive labels f_p and f_q according to (9), then (11) may be written equivalently as

$$E(f) = \sum_{p \in \mathcal{P}} C_p(f_p) + \sum_{p \in \mathcal{P}, q \in \mathcal{N}(p)} C(f_p, f_q), \tag{12}$$

where $E(f)$ to be minimized is just the log-posterior probability (minus the constant) (8).

Let $m^t_{p \to q}(f_q)$ be the message sent from p to q with label f_q in iteration t. Then the update rule is

$$m^t_{p \to q}(f_q) = \min_{f_p} \left(C(f_p, f_q) + C_p(f_p) + \sum_{s \in \mathcal{N}(p)} m^{t-1}_{s \to p}(f_p) \right). \tag{13}$$

Finally, after sufficient iterations, each node q is assigned the label that minimizes the errors in the final iteration T,

$$f'_q = \underset{f_q}{\operatorname{argmin}} \left(C_p(f_p) + \sum_{p \in \mathcal{N}(q)} m^T_{p \to q}(f_q) \right). \tag{14}$$

2.4 Preliminary Results

The described method has been empirically observed to perform quite well on our data set. Some example results are displayed visually by showing each consecutive segment of a subsequence that has been assigned to the same way in a consistent color (Figure 1). Thus, it can be seen that panoramas are properly matched to streets, and when the route followed doubles back on itself, the subsequences are (correctly) classified separately. We demonstrate also the case of a rather large discrepancy where the street map has an incorrect placement of a side street, yet panoramas are still correctly assigned to that street (Figure 2).

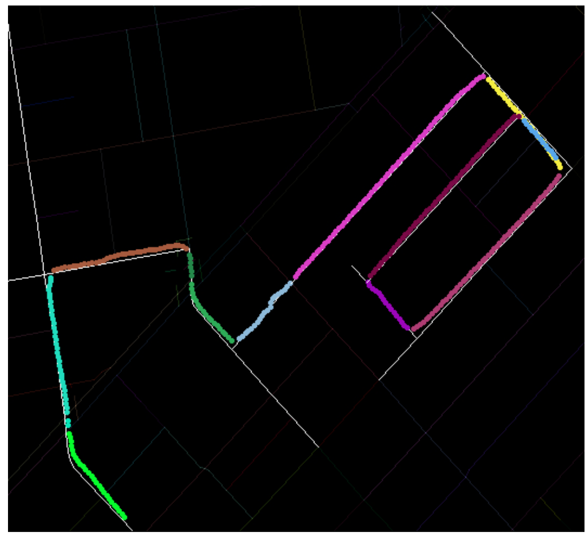

Figure 1: Example of subsequence with randomly colorized subsequences of consecutive panoramas that matched to the same graph edges (streets).

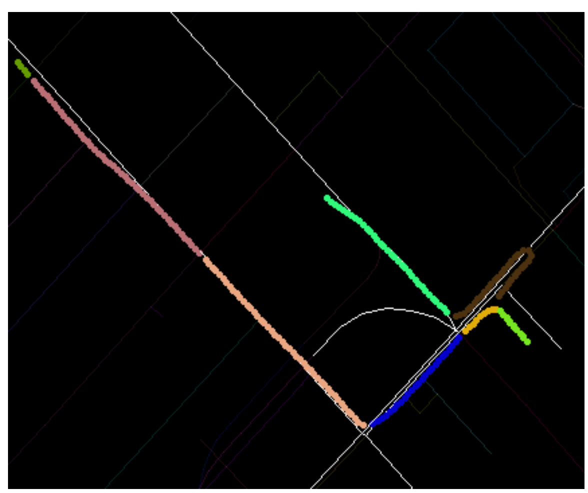

Figure 2: Example of subsequence with randomly colorized subsequences of consecutive panoramas that matched to the same graph edges (streets). Note the correct classification of panoramas to the side street despite, in this case, the large discrepancy in position of the side street.

References

[1] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient belief propagation for early vision. *Int. Journal of Computer Vision*, 70(1), 2006.