TRACKING THE HUMAN BODY VIA A WIRELESS NETWORK OF PYROELECTRIC SENSOR ARRAYS

James Jolly, Joe Bishop and Emilio Nanni
Advisor: Kurt Kosbar
IFT/MST Telemetry Learning Center
Missouri University of Science and Technology
Rolla, Missouri

ABSTRACT

This paper describes the design and construction of a low-cost wireless sensor network (WSN) intended to track a human body walking upright through its physical topology. The network consists of arrays of pyroelectric infrared (PIR) sensors that can detect a moving body up to five meters away within a semicircular field of view. Data is gathered from these arrays and transmitted to a central processor that triangulates the body's position. Important characteristics of both the PIR sensors and the network's asynchronous nature are elaborated upon to illustrate how they affect the interpretation of the data.

INTRODUCTION

Human tracking is an exciting application domain for WSN's, as many existing systems benefit from precise knowledge of an individual's position. For example, many indoor security systems exist that detect the presence of an intruder; however, these systems do little more than note their point of entry into a secure area. Using such systems makes it difficult to infer the path they have taken and their present location in areas without choke-points or corridors. Estimating these both improves the chances of intercepting the intruder and provides insight into what was compromised during the intrusion.

The design of a WSN capable of reliably detecting and estimating the position of a human is challenging because:

- An accurate knowledge of physical deployment of each wireless sensor unit (WSU) is necessary to calculate an accurate position for the human being tracked.
- Sensors on different WSU's will likely be subject to different noise sources.
- Without clock synchronization, sensor observations will be made at different times.
- Observations may not arrive at a station collecting data in the order they are made.

Sensor choice has a large impact on the nature of the tracking task. The tracking of humans with pyroelectric sensors is described in the section on Pyroelectric Sensor Characteristics. These sensors generate a potential based upon changes in incident passive-infrared radiation, and bring about new challenges within the context of the tracking task:

- A pyroelectric sensor detects changes in radiation. As a result, it is not possible to detect very slow-moving and stationary objects, even if they emit considerably more or less infrared radiation than their surroundings.
- It's difficult to gauge the distance to a human body when a change in infrared is detected. Each sensor's output is a function of the speed an object moves in front of it, that object's heat signature, and the object's distance from the sensing elements. It is possible for a waveform generated by a distant fast-moving body to resemble the waveform from a nearby slow-moving body.
- The human body comes in a range of sizes, and may be wearing clothing that interferes with its heat signature.

As a result of the above, a WSN that utilizes pyroelectric sensors to detect a moving human body requires overlapping sensor ranges to accurately triangulate its position. The section on Sensor Array Design describes the arrays of pyroelectric sensors intended for placement, such that their fields of view overlap. The Position Estimation Algorithm section describes how we merged the unique world-views of the arrays to generate an estimate for a moving body within their field of view.

Pyroelectric Sensor Characteristics

Pyroelectric sensors rely on passive infrared radiation emitted by the objects moving in front of them. All humans emit infrared radiation, and unless steps are taken to reduce one's infrared signature, most humans emit infrared radiation of approximately the same wavelength and magnitude.

Plank's Law for black-body radiation relates spectral radiance (I, measured in $\frac{W}{m^3}$) to wavelength for an object not illuminated by other sources of electromagnetic radiation.

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda kT}} - 1}. \tag{1}$$

λ : radiation wavelength (m)
h : Planck constant
c : speed of light (m/s)
k : Boltzmann constant
T : temperature (K)

Spectral radiance peaks between between 5 and 20 μm for a human body at 310 K.

The RE200B, a balanced differential, series-opposed, dual-element pyroelectric sensor was selected to build the sensor arrays. This sensor is manufactured by Nippon Ceramic Co., Ltd., and was selected because of its use in prior human tracking work [1] and affordability compared to other pyroelectric infrared sensors.

Each RE200B detects radiation with wavelength between 7 and 14 μm, a property determined by a glass plate that also protects its two internal sensor elements. These elements are separated such that they witness the same phenomenon from slightly different perspectives and opposed in series such that the voltage across them is zero when they witness the same change in infrared radiation. This balanced differential nature of the sensor greatly reduces interference by background radiation (which is seen by both internal sensor elements and subsequently canceled).

An object passing in front of the sensor generates an output voltage across it that resembles a single period of a sinusoid, the result of two opposing voltage spikes occurring in rapid succession, one associated with each sensor element. The phase angle of the sinusoid can be used to determine the direction of movement across the field of view (FOV) of the sensor.

Unfortunately, the FOV for dual-element sensors is typically large (often as much as 120°). Because objects are more likely to change speed or direction within a large FOV, a Fresnel lens was used to shrink the FOV of each sensor to approximately 15°, making it obvious when a body moves directly in front of it and reducing its susceptibility to changes in temperature in indoor air currents. A Fresnel lens has a much greater surface area than the internal sensor elements, increasing the effective gain of the sensor by increasing the amount of infrared it absorbs.

In indoor testing, a variety of non-human phenomena were captured by these sensors (such as a burst of turbulence produced by an air handler turning on). As a result, there was need to reject sinusoidal pyroelectric sensor outputs that were most likely not produced by human movement. Bounds for the range of frequencies generated by human movement within the 15° FOV of a sensor were estimated using the following formulas:

$$f_{low} = \frac{v_{slow}}{\theta_{sensor}\ d_{long}} \quad f_{high} = \frac{v_{fast}}{\theta_{sensor}\ d_{short}} \tag{2}$$

θ_{sensor} : the FOV of a single sensor
f_{low}, f_{high} : the lowest and highest frequency sinusoids generated by movement
d_{long}, d_{short} : the furthest and shortest distances movement can be reliably detected at
v_{slow}, v_{fast} : the slowest and fastest movement speeds detected by the sensor

Experimentation suggested most human motion occurs at speeds between 0.25 and 2 m/s, and that a sensible effective range for the sensors was between 5 and 1 m. 5 m was picked as an upper bound because it became difficult to discern casually dressed targets moving at low speed at that distance. This was especially true in settings with a great deal of thermal noise.

These speeds yield an approximate minimum and maximum frequency of 0.2 and 7.6 Hz, respectively. An active band-pass filter with a center frequency of approximately 3 Hz and a bandwidth of approximately 10 Hz was selected to both amplify and filter the sensor outputs.

SENSOR ARRAY DESIGN

Figure 1: sheet metal housing

The Crossbow MICA2 Berkeley mote was selected to serve as a platform to gather data from the pyroelectric sensor arrays. Each WSU subsequently consisted of a MICA2, a pyroelectric sensor array, and a custom data acquisition board (DAB) used to filter and amplify the voltages produced by the sensors. Each array was was outfitted with eight pyroelectric sensors, the number of ADC's available on the MICA2.

The WSU array consisted of a sheet metal housing that served as a mounting frame for the Fresnel lenses and pyroelectric sensors as well as electromagnetic shielding for the data acquisition hardware (particularly the wireless hardware itself). The WSU housing held each of the eight sensors at focal length behind its respective lens.

Hoping to as monitor as much area as possible with an acceptable level of tracking accuracy, the centers of each sensor's FOV were separated by 18° such that each WSU monitored a 162° viewshed, with a small amount of dead space in between each sensor in the array.

Unfortunately, sheet metal reflects infrared energy well, and gaps in the housing resulted in infrared "leaks". These gaps were easily sealed with electrical tape, which effectively blocked the unwanted infrared radiation. Dividers were also placed in between each sensor within the housing to prevent infrared from passing through a lens at an oblique angle and reaching an unintended sensor.

POSITION ESTIMATION ALGORITHM

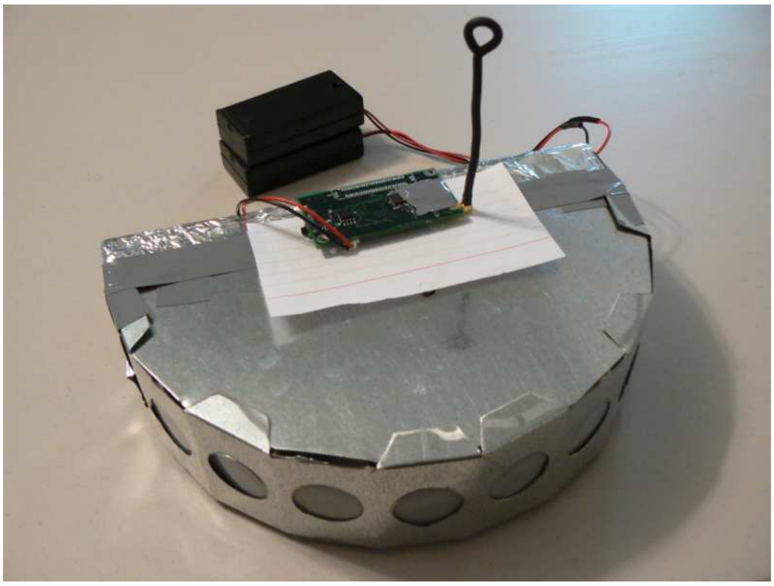

Figure 2: WSU

Each WSU was programmed to continuously sample the amplified and filtered sensor outputs produced by the DAB 15 times a second, transmitting a datagram to a nearby base station after every sample to facilitate real-time tracking. Each datagram contained a unique ID number assigned to its respective WSU and eight sensor output values, all encoded in 32-bit unsigned integers.

The stream of samples from each sensor was digitally filtered upon arrival at the PC. An average was continuously computed over a large window of size w for each sensor's output stream. This resulted in a stream of averages available for a sensor after w values arrived from it at the base station. At that point, the absolute value of the difference between each sample and this average was computed. The resulting stream of values was compared to a sensibly set threshold to determine if a sensor was observing motion. These "detector" values typically lingered near zero and increased when stimulated by human-like motion. Comparing against a small-window size ($w/20$) average of the detector values reduced the occurrence of false-positives by deemphasizing the significance of each individual sample.

Each marked sensor, associated with a particular viewing angle with respect to a particular WSU, then contributed to the continuous calculation of a heading to nearby human movement for that WSU. For each set of eight readings, a heading to the nearest movement was crudely estimated by averaging the angles associated with the sensor outputs that stepped over the threshold discussed above. When a sensor was marked, the routine running on the PC took it into consideration for the heading calculation for 0.5 s.

The PC then used the headings reported by each WSU to pinpoint the location of the movement at any given instant. This was accomplished by finding points of intersection between the headings. Note that each heading, coupled with its associated WSU location, forms a ray. To support all possible physical network topologies, a check for ray intersection was required for each heading-to-heading comparison.

A simple algorithm was then used to estimate the location of movement within the network by using the WSU headings in tandem with their respective positions:

For every pair of nodes in the network, determine if their rays intersect within reasonable range of the network. If a nearby point of intersection exists, note it. Afterward, look at all the points of intersection and average their coordinates. The resulting centroid is an estimate of the humans position.

For n WSU's, this algorithm's worst-case computational complexity is $O(n^2)$, as each WSU must compare itself against every other WSU in an attempt to find points where their rays intersect. In a worst-case scenario, every WSU will calculate a ray that uniquely intersects every other WSU ray. After each comparison the coordinates of the point of intersection can be added to a sum of the coordinate values, adding a constant-time penalty to each of the comparison steps. This sum can then be divided by the number of intersections to find the centroid.

With the arrival of each datagram, the points of intersection between its associated WSU and the other WSU's using the location estimator above were recomputed. This shifted the average of these points, continuously recreating centroids representative the present location of movement. By chaining-together estimates of this location over time, it became possible to record the path a human elected to walk within the network.

TESTING

This body tracking was tested using four WSU's, all placed within range of the base station such that each unit's effective FOV overlapped with each of the other units' FOV's. Overlap regions consisted of two, three, and four overlapping FOV's. Zones with the highest amount of overlap generally performed the highest resolution tracking. This was expected, presumably due to the error made by each unit being averaged-out by the observations of other units with different perspectives. The predicted position was generally within 1 m of the correct position with four units placed 5 m from each other. These units were arranged in a plus-shaped physical network topology such that their viewsheds overlapped in the center of the plus.

In addition to measurement and triangulation challenges associated with the infrared sensors, the message-passing between wireless sensors presented its own unique challenges. Most importantly, each MICA2 did not sample in tandem with the other MICA2's due to the communication overhead involved in running clock synchronization on top of a sensing application. As a result, observations were made at different times by each WSU. As a heuristic, the position estimator used the last message successfully transmitted by each WSU to predict where a human is.

Each MICA2 was programmed to place a sequence number in its messages such that the last message from a WSU could always be identified from the partial order imposed by the sequence numbers. Every position estimate was a function of the values contained in the last messages received from the WSU's.

Despite the information about some rays being older than others, this system worked acceptably when messages arrived from each WSU at about the same intervals, particularly when capturing over 10 samples each second.

FUTURE IMPROVEMENTS

One of the chief barriers to the WSU's accurate tracking a body was thermal noise and electromagnetic interference in the deployment environment not being the same from sensor to sensor. Each sensor has its own unique perspective; some sensors see a lot of noise, while others little. This results in some sensors being capable of detecting movement under the detector threshold, which is set high in the interest of preventing false-positives.

It might be of benefit to statistically model the output of each individual sensor unit in real-time. Then a sensor unit could be marked as having observed movement when a significant change occurs to a signal that its model indicates is highly unlikely. This approach might prevent thermal noise and electromagnetic interference from negatively influencing the position estimation algorithm. It would also make it easier to quantify how significant a change is, and subsequently improve the weighting of angles when calculating a heading for each WSU.

Building such statistical models could be accomplished by software running on the PC, calibrating the network to a particular environment. Quantities of primary interest are signal mean and variance; however, other higher-order statistical moments might be of use. The parameters that govern these models should be learned on a much longer time frame than the expected slowest detectable infrared energy change.

Note that the pyroelectric signals discussed also yield information about the lateral nature of the infrared energy change (i.e. which direction is the body moving in), which was not being taken advantage of in this position estimation method. Recording this information when a threshold is crossed (or a significant deviation from a statistical model is detected) would add an additional one-dimensional vector quantity to each heading generated. Coupled with the perspective of two WSU's not observing something moving in between them, the positions of a human estimated by a network could now be turned into both a point and a two-dimensional vector quantity which describes that point's instantaneous motion in a direction.

A better position estimation algorithm might also reduce the weight of old information when calculating points of intersection. It might also try to calculate an expected future movement profile (likely a simple heading) to help improve future position predictions by using past position estimates.

CONCLUSION

Wireless sensor networks have great potential to perform spatial tracking of human movement using arrays of low-cost pyroelectric sensors. Using such sensors places adds additional complexity to the software needed to merge the unique perspectives of them together. Asynchronous sampling and unreliable transport of sensed data makes this merging problem more challenging.

REFERENCES

[1] Shankar, M., et al., "Human-Tracking Systems Using Pyroelectric Infrared Detectors", Optical Engineering, 45, 106410 (2006)

[2] Armitage, A., Binnie, T. and Kerridge, L., "Measuring Pedestrian Trajectories Using a Pyroelectric Differential Infrared Detector", 83rd Annual Transportation Research Board Annual Meeting, Washington D.C., January, 2004

[3] Ditzel, M. et al., "Multi-Target Data Aggregation and Tracking in Wireless Sensor Networks", Journal of Networks, Vol. 3, No. 1, January 2008