Lightweight biometric detection system for human classification using pyroelectric infrared detectors

John Burchett, Mohan Shankar, A. Ben Hamza, Bob D. Guenther, Nikos Pitsianis, and David J. Brady

We use pyroelectric detectors that are differential in nature to detect motion in humans by their heat emissions. Coded Fresnel lens arrays create boundaries that help to localize humans in space as well as to classify the nature of their motion. We design and implement a low-cost biometric tracking system by using off-the-shelf components. We demonstrate two classification methods by using data gathered from sensor clusters of dual-element pyroelectric detectors with coded Fresnel lens arrays. We propose two algorithms for person identification, a more generalized spectral clustering method and a more rigorous example that uses principal component regression to perform a blind classification. © 2006 Optical Society of America

OCIS codes: 100.5010, 120.0280, 330.5000, 150.6910.

1. Introduction

Biometric systems are widely used in person verification and secure identification. Unique identification is a gateway to many technologies, particularly user-based services that are very local in nature, such as access control to secure locations. Human tracking systems are principally interested in determining the existence and location of humans within regions of open space. While low false-positive biometric methods typically require close proximity and rely on high-resolution data collection accompanied by much computation, tracking applications are often the opposite, designed for distribution over longer ranges and trading pinpoint resolution for reduced sensory and computational requirements.

A tracking sensor and a biometric sensor perform similar measurements, with an output based on the activity of the human within his or her field of view. Gait recognition is an answer to this problem, recognizing targets by their silhouette profile.1 Typical systems often employ high-data-rate imaging devices,2 which when used in the infrared domain are expensive. However, human tracking does not require the high resolution provided by imaging devices3; meaningful tracking is possible with resolution comparable to the body’s cross section. Motion tracking arrays have been implemented to supplement video tracking4 as an alternative to expensive thermal cameras. Distributed infrared sensor networks have begun to meet these needs and are both cheap and easy to construct.5–7

The output of tracking sensors may not uniquely identify a person, but the data are sufficient to distinguish personal characteristics. We demonstrate an algorithm that treats the data gathered from several tracking sensors as a high-dimensional vector and that uses multidimensional scaling techniques to map these vectors to a lower-dimensional space. By performing spectral clustering on the lower-dimensional vectors, we classify the data from multiple walks by different people. Through the mapping process, data corresponding to each person is clustered in the lower-dimensional space. Classification is made possible by the application of this calibrated mapping to a random walk by identifying the cluster to which the mapped vector belongs.

2. Sensor Platform Development

Human bodies radiate heat to their environment as predicted by the Stefan–Boltzmann law.8 The average human frame radiates \( \sim 100 \text{ W/m}^2 \) of power,9 which peaks at about the 9 \( \mu \text{m} \) wavelength, assuming it is an ideal blackbody source. To detect this radiation, we employ the dual-element pyroelectric...
detector PIR325 from Glolab Corporation.\textsuperscript{10} To improve both our collection efficiency and the spatial resolution, we introduce Fresnel lens arrays to create a discontinuous visibility pattern. By segmenting the detectors’ visibility, we cause the detectors to produce a time-varying signal that can be used to both track and classify objects moving in the field of view as shown in Fig. 1. The following platform description is abbreviated; please refer to Refs. 11 and 12 for detailed descriptions of the detector and our platform, respectively. We chose a lens array with 11 lenslet elements, which produces the signal shown in Fig. 1 when a person walks from left to right and then returns.

Each sensor module is assembled by using eight pyroelectric detectors, each with its own lens array. The sensor platform was originally designed for a cheap, distributed collection of low-resolution spatial data regarding the presence of infrared sources, but with proper analysis it can provide a classification of the target while also determining the target’s position.

3. Biometric Classification
All biometric identification systems rely on the construction of a one-to-one mapping between the sensory patterns and the individuals who create those patterns. In traditional systems the complex structure of certain body parts, such as the human iris, is measured optically, analyzed digitally, and then a digital code is computed for each person’s iris. Similarly, other aspects of a person’s behavior, such as a sensor’s data pattern given a prescribed walking pattern, are plausibly unique to that person. Each person acts as a distributed infrared source whose distribution function is determined by, among other things, his or her shape and the infrared emissivity of the person’s skin at every point. Combined with the idiosyncrasies in how individuals carry themselves, even when walking the same prescribed path, their heat will impact a surrounding sensor field in a plausibly unique way. By measuring the sensor response to a person in a prescribed walking pattern, we map this response to a code vector that uniquely identifies the person.

A. Problem Formulation
To classify each person’s walk, we collect a sensor data trace from several detectors possessing a previously demonstrated response pattern. Let $s$ be the number of sensors and $T_j^{(i)}$ be the length of each sensor’s data trace in discrete samples. Denote by $z_j^{(i)}$ a $T_j^{(i)}$-dimensional vector that represents the $j$th sensor measurement (i.e., $z_j^{(i)}$ is a signal of length $T_j^{(i)}$ of the $i$th walk, where $j = 1, \ldots, s$ and $i = 1, \ldots, m$. Note that for each walk $z_j^{(i)}$, we may get a different temporal length $T_j^{(i)}$, and hence we may need to interpolate the signals $z_j^{(i)}$ in time so that the length of each signal $z_j^{(i)}$ is equal to $T = \min(T_j^{(i)})$. Therefore each person’s walk can be represented by a $d$-dimensional vector $x_i = \text{vec}(z_1^{(i)}, \ldots, z_s^{(i)}) \in \mathbb{R}^d$ obtained from $s$ sensor measurements by stacking the $T$-dimensional column vectors $z_j^{(i)}$ underneath one another, where $i = 1, \ldots, m$ and $d = Ts$.

The biometric classification may now be concisely described by the following statement: given a data set of walks, $X = \{x_i\}$, where $1 \leq i \leq m$ and where each walk $x_i$ is a high-dimensional vector. We first employ dimensionality reduction to map these high-dimensional vectors to a lower-dimensional space by using an isomorphic mapping function. The mapping is designed to preserve a geodesic distance metric so that the distance between points in the higher space is reflected approximately in the corresponding distance in the lower space. Next we perform spectral clustering on the resultant vectors; spectral clustering refers to algorithms that cluster points in space by using eigenvectors drawn from the matrix formed by aggregating the vectors themselves. Each vector cluster corresponds to the sensor’s response to a person’s walk; using a graph algorithm, we segment this lower-dimensional space. Discrimination between persons is accomplished by assigning each segment to the person whose walks are most closely mapped to it.

B. Implementation
Our experimental setup uses a sensor platform module, as discussed in Section 2. We gathered data from three different people walking along a prescribed path within the field of view of the module. We used the sensor platform with a radial configuration; the module’s sensors look outward in a radial pattern, with each of the eight detectors oriented 45° from its neighbor. The subjects walk along a straight path tangential to the sensor module at a distance of ~3 m, as shown in Fig. 2. Only three of the eight detectors are oriented toward the subject’s path and generate useful data; these three data streams are collected independently of one another.

Taken as a whole, the three sensor data streams describe the sensor’s response to a person’s walk. We concatenate the three-stream, one-dimensional vec-
tors end to end into a single vector, similar to the use of a single feature vector for facial or iris recognition. Seventeen walks were performed by three people, for a total of 54 data traces. Data analysis of the gathered data streams is performed by using multidimensional scaling (MDS) methods. MDS refers to a collection of mathematical techniques used to embed a high-dimensional set of points described by a dissimilarity matrix into a low-dimensional space such that the geodesic distance between vectors is preserved as discussed in Subsection 3.A.

While there are necessary similarities in the walk data from different persons, each person’s walking style is unique, which is reflected in the subtleties in the timing and amplitude of the signal. However, how this style is reflected in the sensor signal itself is not easily quantified through signal correlation; we quantify the similarities by treating each data sample as a multidimensional vector and by looking at the distance between vectors. If a similar pattern appears in the two vectors, then the vectors will be similarly close in the corresponding dimensions. Using the eigen-decomposition method described in Table 1, we map these vectors to a lower $d$-dimensional representation. By mapping to a lower space, we compact this notion by changing the comparison problem to a graph segmentation problem.

By applying the algorithm steps shown in Table 2 to the mapped vectors, we divide the points into clusters of points in which each cluster is determined to be the walks of one person. Because these walks are calibration data, we know a priori how many clusters to form and the number of data points that should belong to each cluster. We perform a blind segmentation by using eigenvalue analysis, judging the success of our algorithm by how many of each person’s walks are correctly clumped together. The mapping of points to persons results in a segmentation, shown in Fig. 3. With this mapping, only two walks were

![Figure 2. Sensor placement and a resulting data stream generated by a single person.](image-url)
assigned incorrectly, with the remaining walks being clumped correctly into clusters corresponding to each person walking. Treating this segmentation as a calibration step, we prove our algorithm by taking new walk data from each of the three persons and performing the same mapping on the high-dimensional vector resulting from their sensor trace. The mapping of the new walks are boxed in Fig. 3; the mapping function correctly maps each of the unknown walks to the right cluster in the lower-dimensional space.

4. Classification with Principal Component Regression

We use principal component regression to calculate a set of discrimination vectors that will classify the spectrum of a person’s walk into one or more possible states. The state is defined generally, e.g., the person doing the walk, the activity they are performing, whether they are walking quickly or slowly, etc. The weighted linear combination of the measured spectral intensities \( w^T s = \kappa \) projects onto a classified value of \( \kappa = +1 \) for positive instances (belonging to a state) and \( -1 \) for negative instances. Each element of \( w \) reveals the contribution of that spectral region in the decision. Note that different \( w \) vectors may estimate different properties from the same spectral data. This projection method is similar to Woodward’s classification method for determining information about a remote target by using shaped radar pulses.\(^{13}\)

A. Algorithm

We assume that a sufficient dictionary of spectra is available for each state (but not necessarily for each person) to be used for calibration. Given a set of training spectra arranged columnwise as matrix \( A \) of known memberships \( c \), the least-squares (LS) error condition defines a weight vector \( w \) such that

\[
\text{arg min}_{w} \| A^T w - c \|_2. \tag{1}
\]

The LS problem is solved by using the singular-value decomposition (SVD) of the system matrix \( A^T \), where

\[
A = \sum_{k=1}^{n} \sigma_k u_k v_k^T = U \Sigma V^T, \tag{2}
\]

where \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \) are the singular values and orthogonal matrices \( U \) and \( V \), with \( U^T U = I \) and \( V^T V = I \), correspond to left and right singular vectors, respectively.

Since the linear system in Eq. (1) is underconstrained, there are infinite \( w \) that will satisfy it without necessarily being good classifiers for cases not used in training (present in \( A \)). A cross-validation calibration of the training data is used to remedy that. Leave-one-out cross validation determines the rank of the approximation

\[
A_r = \sum_{k=1}^{r} \sigma_k u_k v_k^T = U_r \Sigma_r V_r^T \tag{3}
\]

used for the LS solution by utilizing the maximal rank of the matrix \( A \) that would minimize the worst error.

We would like to note here that the above presen-
tation covers one-dimensional projections to \(-1, 1\) points. However, one can define higher-dimensional projections by defining multiple classifiers \(w\) and target coordinates. In a two-dimensional example, one can define two classifiers \(w_x\) and \(w_y\) such that \(w_x^T w_y = 0\) and can define target coordinates belonging to the vertices of a regular polygon.

B. Testing

We test our discrimination vector by gathering training data streams from several people performing different activities. Three individuals performed multiple walks in different states of activity. A total of 65 walks were performed, combining different subjects and activities as given in Table 3.

We determine an appropriate discrimination vector that maps walks corresponding to a specific state to 1 and walks not corresponding to that state to -1. We demonstrate in Fig. 4 the discrimination vectors for four different states: (1) subject is walking very slowly, (2) subject is taking small steps, (3) subject is swinging arms while walking, and (4) subject is Katie. Each of these states reflect a different classifiable attribute, related to either the subject’s activity or identity. After constructing each vector by using training data, we apply the vector to 65 sets of walk data, computing two separate histograms of the returned values of \(\kappa\). Each plot contains two curves; one plots a histogram of the \(\kappa\) values for walks that match with the prescribed state, and the second one similarly plots the walks that do not match. For example, in the plot labeled “arms swinging,” the left curve describes the \(\kappa\) values for input data in which the subject was not swinging his or her arms, whereas the right curve describes input data in which the subject was swinging his or her arms. The relative heights of the two plots is indicative only of how many walks fell into these two categories; it does not reflect on the accuracy of the method. While ideally the two

<table>
<thead>
<tr>
<th>Table 3. Breakdown of 65 Walks by Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Mohan</td>
</tr>
<tr>
<td>Bob</td>
</tr>
<tr>
<td>Katie</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Fig. 3. Three unknown walks (boxes) are mapped by using a calibrated training set.
curves would be delta functions at $\pm 1$, the separation of the plots is the real indication of accuracy.

We characterize the accuracy of our recognition by using Woodward’s approximation\textsuperscript{13}; Woodward’s $p_{\text{yes}}$ and $p_{\text{no}}$ notations correspond to each plot’s two curves centered at $+1$ and $-1$, respectively. We define a threshold value $\Theta$ similar to Woodward’s $\Theta$, where histogram values above the threshold represent a positive detection. A false-alarm error occurs for values of $p_{\text{no}}$ that are above this threshold, and lost detection occurs when $p_{\text{yes}}$ dips below the threshold. The variables $F$ and $L$ are defined to be the area under the curve of these error regions, respectively. Woodward notes that choosing $\Theta$ can be done to minimize either form of error but that reducing one will increase the other. We choose $\Theta$ to be the point at which $p_{\text{yes}}$ and $p_{\text{no}}$ are equal, as shown by the vertical line splitting the two shaded regions in Fig. 4; assuming that the two distributions are relatively similar, this will reduce our overall error.

The error probabilities can be determined by dividing the two areas $F$ and $L$ by the total area under the two curves, and they are acceptably small as shown by inspection of Fig. 4. The accuracy is also indicated by the separation of the two plots; ideally, the discriminator will map all matching walks to positive $+1$ and all nonmatching walks to $-1$. The separation value given above each plot is defined as the difference between the minimum positive instance $\kappa_+$ and the maximum negative instance $\kappa_-$. When the separation is positive, the discrimination succeeds in clearly separating the positive and negative instances. When the separation is negative, there are instances that cannot be discriminated. The absolute value of the negative separator denotes the size of the region of uncertainty.

5. Conclusions

We have demonstrated a lightweight biometric detection system for human classification by using pyroelectric infrared detectors equipped with Fresnel lens arrays. These detectors are more commonly used for human tracking applications, but through isometric mapping and principal component regression, we are able to extract features that are used to distinguish between different persons and activities. While other biometric mechanisms and devices are clearly superior in accuracy, they are both expensive and can be performed only in close proximity to the device. By using standard medium-range tracking sensors, we extend the operating range of identification systems,

![Fig. 4. (Color online) Discrimination results for four different states.](image)
making possible a spatially distributed biometric classification network.

The authors acknowledge the Army Research Office for supporting this research through grant DAAD19-03-1-0352.

References