RESEARCH

Open Access

Design and implementation of a compressive infrared sampling for motion acquisition

Tong Liu[1] and Jun Liu[2*]

Abstract

This article proposes a compressive infrared sampling method in pursuit of the acquisition and processing of human motion simultaneously. The spatial-temporal changes caused by the movements of the human body are intrinsical clues for determining the semantics of motion, while the movements of short-term changes can be considered as a sparse distribution compared with the sensing region. Several pyroelectric infrared (PIR) sensors with pseudo-random-coded Fresnel lenses are introduced to acquire and compress motion information synchronously. The compressive PIR array has the ability to record the changes in the thermal radiation field caused by movements and encode the motion information into low-dimensional sensory outputs directly. Therefore, the problem of recognizing a high-dimensional image sequence is cast as a low-dimensional sequence recognition process. A database involving various kinds of motion played by several people is built. Hausdorff distance-based template matching is employed for motion recognition. Experimental studies are conducted to validate the proposed method.

1 Introduction

How to effectively acquire the human motion information is of key importance for analyzing and interpreting behavior. Wearable sensors and isomorphic vision sensors are the most used sensing methods for motion acquisition. The wearable sensor-based sensing method is able to directly obtain the motion information of specific limbs or joints. It can form the low-dimensional sensor output-based feature representation [1]. However, the wearable sensor belongs to the intrusive sensing method. It needs the observed person to wear or wrap the sensor on the body, so the feeling of comfort will be affected.

The isometric vision sensor-based sensing is the non-intrusive method for motion acquisition. It has shown strong potential application prospects with the developing advantages of the sensor's low cost and networking [2,3]. In order to achieve the recognition and understanding of visual motion, it needs to extract the useful information and feature from the high-dimensional image data. Some representations of the feature, such as the geometric model of the human body, spatial-temporal patterns, appearance, area, contours, and optical flow model,

have been demonstrated to be the effective methods [4]. However, a large number of studies have found that the dimension of the visual image-based feature is high. It will increase the computational complexity of recognition in the information processing and be not suitable for the real-time requirements of systems.

The field of computer vision research has found that the visual image-based data contains considerable redundant information. This makes the data analysis complicated in computation [4,5]. It needs further refinement to remove the redundant parts of the data for forming the low-dimensional feature. Many parsimony models referred as dimension reduction are seeking efficient transformation and recognition based on a low-dimensional representation, such as principal component analysis (PCA) [6], isometric feature mapping (Isomap) [7], and local linear embedding (LLE) [8]. One can remove the redundant information sampled densely by generic sensors and improve the efficiency in the subsequent data analysis procedure [9]. However, there is an imbalance between the data acquisition and data utilization. Great efforts are made to compress the features for moving the course of dimension. This is a wasteful process both for sensing and computational resources. Moreover, the compression algorithms of visual feature heavily rely on software computing platforms, while little work is focused on the

*Correspondence: virginliujun@gmail.com
[2]College of Physics & Electronic Information Engineering, Wenzhou University, Wenzhou 325035, China
Full list of author information is available at the end of the article

physical hardware in pursuit of integrating the compression processing directly.

The newly proposed compressed sensing (CS) theory is able to integrate the sparse signal acquisition and compression in a single process [10-12]. This sensing mechanism avoids the generation and processing of the redundant information. The CS theory supported that the single-pixel camera is able to use the incoherent basis to sample the observed space; then, the scene could be recorded with a low-dimensional form [13,14]. The dimension of measurements is only relevant to the sparsity of the raw signal, and the incoherent basis has the form of random matrix. The random matrix-based measurement operation is the most common compression and can be embedded into the physical hardware. In addition, a random matrix measurement is independent on the sparse data, and the property of data-friendly compression is preferable. The raw image of the scene can be recovered by solving optimization problems.

The *a priori* condition for CS is that the raw signal itself or in some transform domain can be sparsely represented. Some studies have been focused on the parsimony of motion perception in human vision, such as moving light displays (MLDs) [15,16] and spatial-temporal salient points [17-19]. Two major characteristics of the human motion can be found from the above studies. First, the moving and changing information of body motion is the effective clue for recognizing and understanding the human behavior. Second, the short-term changes in human motion can be considered to be sparsely distributed in spatial domain. This raises a simple question: for a given motion signal, whether it is possible to incorporate the direct acquisition and compression into a single sensing process by certain physical coder.

In general, the visible light-based vision sensors do not have the ability of acquiring the motion information directly. However, the pyroelectric infrared (PIR) sensor, due to its inherent capability in motion detection, has the ability of directly acquiring the motion information [20-28]. In addition, the PIR sensor can form the infrared multiplexed sensing with the visual modulation by Fresnel lens. This sensing mechanism supports the implementation of compressed infrared sampling. Although there has been a lot of work exploiting the combination of PIR sensors and Fresnel lens arrays to form the multiplexed infrared sensing, establishing the compressed infrared sampling model for acquiring the motion information is an unsolved problem.

In this article, we exploit the compressive infrared sampling towards the acquisition and processing of human motion. The spatial-temporal changes caused by the movement of the human body are intrinsical clues for determining the semantics of motion, while the movements with short-term changes can be considered as a

sparse distribution compared with the interface region. Several PIR sensors with pseudo-random-coded Fresnel lenses are introduced to acquire and compress motion clues synchronously. The compressive PIR array has the ability to record the changes in the thermal radiation field caused by human movements and encode the motion information into low-dimensional sensory outputs directly. Therefore, the problem of recognizing a high-dimensional image sequence is cast as a low-dimensional sequence recognition process. Hausdorff distance-based template matching is employed for validating the usefulness of the proposed sensing method. In the experimental analysis, a database involving various kinds of motion played by several people is built. The relations between the compressive dimension and correct recognition rate and the compressive dimension and the time consumed by recognition are compared detailedly, and the application of the proposed sensing paradigm to human-computer interaction is addressed.

While the proposed method is based on the following assumption: The motion of body is constrained in a predefined interface space, we assume that each motion is normalized and the same semantical motions posed by different persons should have small spatial-temporal variances. Although our assumption oversimplifies the general motion recognition problem, our purpose in this article is to present the idea of incorporating the PIR sensor and random compression theory into a single hardware measurement process. More solutions for the real challenges can be found in [29].

The rest of this article is organized as follows: In Section 2, we introduce the related infrared sensing model. Section 3 describes the random compression theory. Section 4 gives the design of compressive infrared sampling for motion acquisition. Section 5 presents the experimental setup and gives the Hausdorff distance-based recognition. The summary and conclusions of this article are given in Section 6.

2 Pyroelectric infrared sensor model
2.1 Sensing model
There has recently been considerable interests in PIR sensors for human motion detection and analysis [20-28]. The PIR sensor is made of pyroelectric materials which is sensitive to the thermal radiation with the wavelength between 5 and 14 μm. When the thermal radiation transfers to the sensor and causes temperature changes, the pyroelectric sensor material will produce an equal number of opposite electric charge in its polar. This will produce a weak voltage. Thus, this sensing process has three advantages. First, it is only sensitive to human motion and supports the motion extraction directly in the hardware. Second, its performance is robust to illumination changes and complex background, so troubles in

traditional camera-based vision system can be removed. Third, through the modulation of the sensor's field of view (FOV) by Fresnel lens, the sensor can be achieved on a specific observation area in an optical multiplexing pattern. If we bridge incoherent projections from the interface region to measurement space rationally, the main information can be recorded in a low-dimensional representation.

The human body is able to make the thermal radiation exchanges with the surroundings at room temperature. By studying the pyroelectric materials, Hossain and Rashid gived the simplified equation of pyroelectric current [30]:

$$I(t) = p_s \frac{\mathrm{d}\mathcal{T}(t)}{\mathrm{d}t} = \frac{p_s \eta}{\mathcal{C}} e^{-\frac{\mathcal{G}}{\mathcal{C}}t} U(t) * \frac{\mathrm{d}\phi(t)}{\mathrm{d}t}, \qquad (1)$$

where p_s is the pyroelectric constant and related to the pyroelectric material. $\mathcal{T}(t)$ is the difference of temperature between the sensor and ambient environment and has the form $\mathcal{T}(t) = \frac{\eta}{\mathcal{C}} e^{-\frac{\mathcal{G}}{\mathcal{C}}t} U(t) * \phi(t)$, where $U(t)$ is the unit step function, $*$ is the convolution operator, $\mathcal{C}$ is the heat capacity of the sensor, $\mathcal{G}$ is the total thermal conductivity between the sensor and the environment, η is the rate of absorption of the sensor, and $\phi(t)$ is the received thermal power of the PIR sensor at time t and can be simplified based on the Stefan-Boltzman law as

$$\phi(t) = \frac{A(t)\varepsilon_h k_B A_s (\mathcal{T}_h^4 - \mathcal{T}_c^4)}{d_0^2} + n(t), \qquad (2)$$

where A_s is the surface area of the sensor, d_0 is the distance between the thermal radiation source and sensor, $\mathcal{T}_h$ is the temperature of human (37°C), and $\mathcal{T}_c$ is the ambient temperature in Kelvin. ε_h and k_B are the Stefan-Boltzmann's constant and emissivity factor, respectively, and $n(t)$ is the noise. $A(t)$ is the surface area of the human body that can be observed by the sensor.

Figure 1 presents the transmission of thermal radiation and sensing process of a PIR sensor. We set $H(t) = \frac{p_s \eta}{\mathcal{C}} e^{-\frac{\mathcal{G}}{\mathcal{C}}t} U(t)$ and $s(\mathbf{r}, t) = \frac{\mathrm{d}\phi(t)}{\mathrm{d}t}$, where the $H(t)$ is defined as the step response function and $s(\mathbf{r}, t)$ is the density distribution function of the changing thermal radiation at the space $\mathbf{r}$. By integrating an external resistor and an amplifier, the PIR sensor's output is refined as

$$m(t) = H(t) * \int_{\Omega} v(\mathbf{r}) s(\mathbf{r}, t) \mathrm{d}\mathbf{r}, \qquad (3)$$

where $v(\mathbf{r})$ is the visibility function, which is '1' when $\mathbf{r}$ is visible to the sensor, otherwise is '0'. The visible function $v(\mathbf{r})$ is able to be achieved by the Fresnel lens physically. The Fresnel lens is made of low-cost plastic and has two main functions. First, it has the capability of focusing the changes of thermal radiation onto the sensor; thus,

the sensitivity can be enhanced. Second, according to the requirements of sensing tasks, it can reshape and code the FOVs of the sensor.

Shankar et al. used the black body as simulated radiation sources of the human body and found that the upper and lower cutoff frequencies of the sensor is 0.7 and 2 Hz, respectively, [25]. The step response of the sensor can be approximated as $H(t) \approx C_v \delta(t - \tau), \tau = 1.8$ s, where C_v is the voltage constant, δ is the impulse function, and τ is the delay constant. In order to facilitate the following discussions, we set $H(t) = \delta(t)$. Equation 3 can be simplified as

$$m(t) = \int_{\Omega} v(\mathbf{r}) s(\mathbf{r}, t) \mathrm{d}\mathbf{r}, \qquad (4)$$

where Ω is the coverage area of the sensor.

2.2 Reference structure-based infrared sampling model

In the community of spatial imaging, the geometry reference structure-based tomography is the sensing mode that samples the spatial information selectively in connection with the tasks [31]. Its core idea is to use the reference structure-based coding modulation to build a projection mapping from sampling space to measurement space. In this article, using the combinations of the PIR sensor and Fresnel lens, we build an infrared sampling model, as shown in Figure 2a. Here, the reference structure is achieved by Fresnel lens physically.

Let us first assume that the FOV of a PIR sensor Ω is divided into L non-overlapping sub-cell Ω_i, having the form
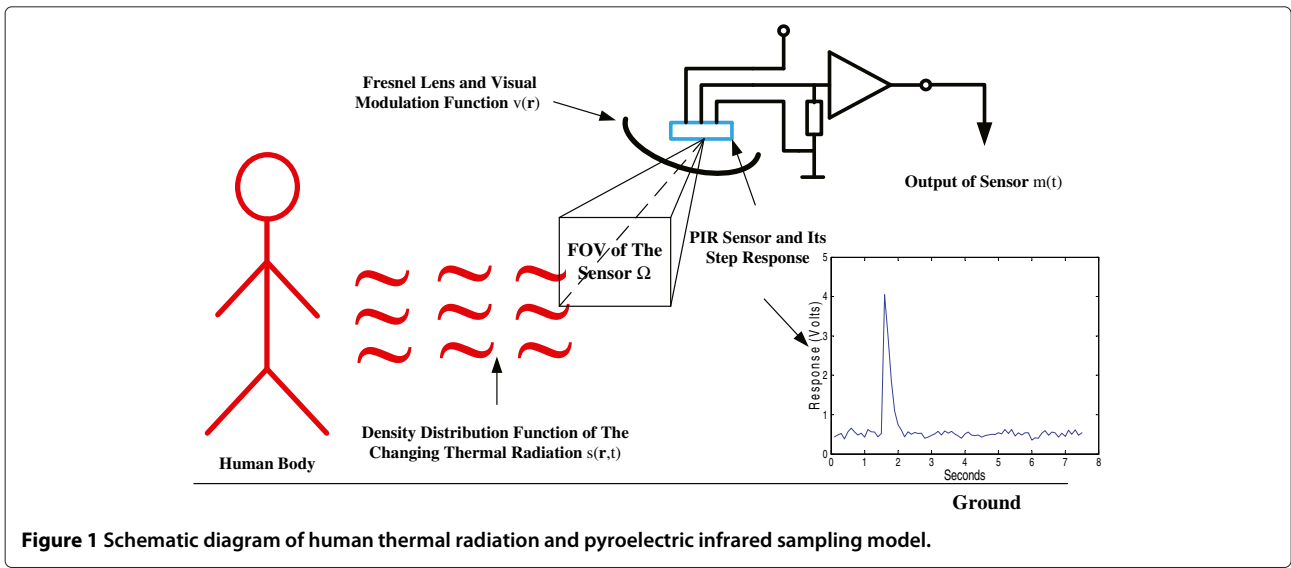
$$\Omega = \bigcup_i \Omega_i, \ \ \Omega_i \bigcap \Omega_j = \emptyset, \ \ i, j = 1, 2, \ldots, L, \qquad (5)$$

where Ω_i is the ith sub-cell of the raw interface space. If the FOVs of the PIR sensor are discrete, we denote v_j as the visible function for the jth sensor; the jth output in the sensor arrays is

$$m_j(t) = \int_{\Omega} v_j(\mathbf{r}) s(\mathbf{r}, t) \mathrm{d}\mathbf{r} = \sum_{i=1}^{L} \int_{\Omega_i} v_j(\mathbf{r}) s(\mathbf{r}, t) \mathrm{d}\mathbf{r} \qquad (6)$$

$$= \sum_{i=1}^{L} v_{ji} \int_{\Omega_i} s(\mathbf{r}, t) \mathrm{d}\mathbf{r} = \sum_{i=1}^{L} v_{ji} s_i(t),$$

where v_{ji} bridges the visibility between the jth sensor and ith sub-cell Ω_i and $s_i(t) = \int_{\Omega_i} s(\mathbf{r}, t) \mathrm{d}\mathbf{r}$ is the integration of the thermal radiation changes in the cell Ω_i at time t. We set $\mathbf{v}_j = \mathrm{row}\,[v_j(\mathbf{r})]$ and $\mathbf{s}(t) = \mathrm{col}\,[s_i(t)]$, respectively, and Equation 6 is rewritten in matrix notation as

$$\mathbf{m}(t) = \mathrm{col}[m_j(t)] = \mathbf{V}\mathbf{s}(t), \qquad (7)$$

Figure 1 Schematic diagram of human thermal radiation and pyroelectric infrared sampling model.

where $\mathbf{V} = [v_{ji}]$ determines the spatial transform of the thermal radiation and is able to be implemented by the Fresnel lens physically.

In general, a PIR sensor commonly consists of single-, dual-, or quad-element detectors. The single-element sensor must add a thermal compensation module to remove the sensitivity to ambient temperature. Quad-element sensors have the inherent advantage that the output is the difference between the voltages obtained from each of the elements of the sensor [25]. The environmental effects can be removed. Figure 2a shows the sampling model of a quad-element PIR sensor, and its output is denoted as

$$m(t) = m_1(t) + m_2(t) - m_3(t) - m_4(t), \qquad (8)$$

where $m_1(t) \ldots m_4(t)$ are the separated output of four elements, respectively. Hence, the visual FOV of the sub-cell can be further divided into four regions by the quad-element PIR sensor, which is denoted as

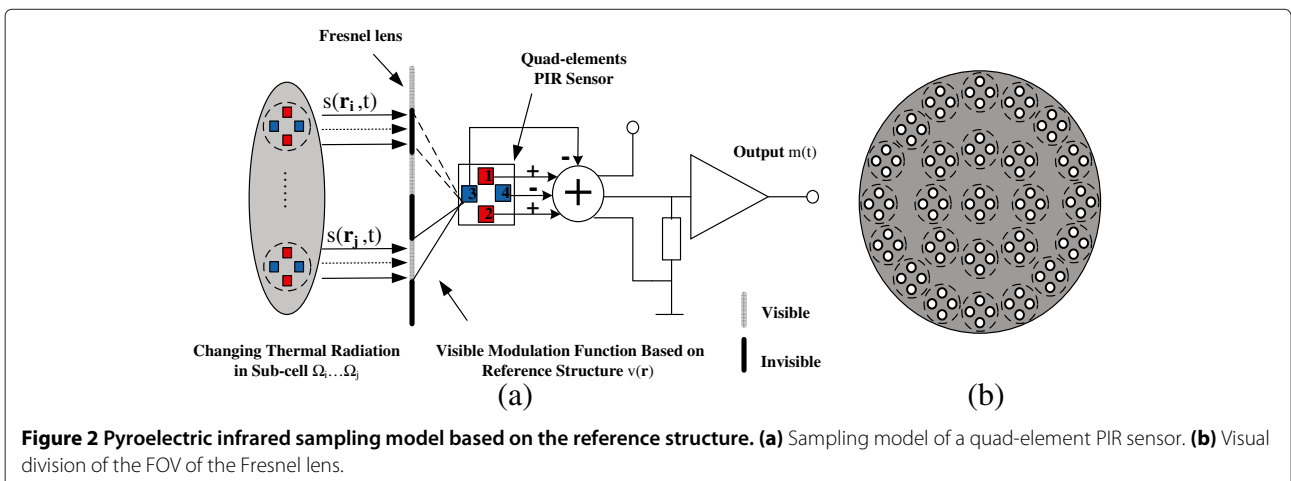$$\Omega_i = \Omega_{i_1} \cup \Omega_{i_2} \cup \Omega_{i_3} \cup \Omega_{i_4}. \qquad (9)$$

The output of the sensor is refined as

$$m(t) = \sum_{i=1}^{L} v_i s_i(t) = \sum_{i=1}^{L} (v_{i_1} s_{i_1}(t) + v_{i_2} s_{i_2}(t) \\ - v_{i_3} s_{i_3}(t) - v_{i_4} s_{i_4}(t)). \qquad (10)$$

Due to the Fresnel lens masks encoding the quad-element PIR sensor integrally visible or invisible for a particular cell, there is $v_i = v_{i_1} = v_{i_2} = v_{i_3} = v_{i_4}$ and

$$m(t) = \sum_{i=1}^{L} v_i (s_{i_1}(t) + s_{i_2}(t) - s_{i_3}(t) - s_{i_4}(t)). \qquad (11)$$

Then, the output of the jth sensor is

Figure 2 Pyroelectric infrared sampling model based on the reference structure. (a) Sampling model of a quad-element PIR sensor. **(b)** Visual division of the FOV of the Fresnel lens.

$$m_j(t) = \sum_{i=1}^{L} v_{ji}(s_{i_1}(t) + s_{i_2}(t) - s_{i_3}(t) - s_{i_4}(t)). \quad (12)$$

Figure 2b shows the Fresnel lens containing 25 non-overlapping cells; thus, each PIR sensor is divided into four sub-cells to form a symmetrical subtraction.

2.3 Sparsity analysis on motion representation

The *a priori* condition for compressive sampling is that the raw signal itself or in some transform domain can be sparsely represented. It is necessary to analyze the motion representation. This is the key to acquire the motion compressively in an efficient way.

Based on the previous PIR sensing model, the sensor will generate approximated impulse response on the changing thermal radiation, while the changing thermal radiation is controlled by the received thermal power $\phi(t)$. If we assume that when the moving subjects keep a fixed distance from the sensor, both body and ambient temperatures are isothermal and the sensor's noise is small, then $\frac{d\phi(t)}{dt}$ is only associated with the visible surface of the body $A(t)$ and can be represented as the moving body parts.

To extract the moving body parts and prove sparsity, we first designed a set of gymnastics to build a motion database. There are 14 kinds of gymnastic motions, including the local movements generated by the arms and legs and the synergistic motions of the upper and lower limbs. Figure 3 gives the sequential images of each kind of motion. All the motions are constrained to perform at a predefined region for keeping the fixed distance from the sensor array. There are five lab members who participated in our experiments, and each member does each motion six times repeatedly. The members are with the most common heights and weights; the range of height is from 160 to 180 cm. Thus, we collected 30 image sequences for each motion; the motions are sampled at 25 frames/s.

In what follows, the sophisticated optical flow method is employed to extract the changing body parts [32]. Examples of three motions are shown in Figure 4a,b,c. We select three frames and three optical flow images in each category of motions for visualization and then compute the intensity of motion flow to represent the changing body parts as shown in the third row of each sub-figure. The large intensity coefficients are represented by light pixels, while small coefficients are represented by dark pixels. It can be observed that most of the coefficients of motion flow are close to zeros. We also compute the average distribution of intensity of motion flow on the designed gymnastic motions and plot the corresponding histogram in Figure 4d. Again, most coefficients are very small to zero, meaning the short-term changing body parts are sparsely distributed. This fact motivates us to set up a compressive infrared sampling for motion acquisition.

3 Random matrix-based compression

In the community of data mining and dimension reduction, the random matrix-based compression or projection has attracted the attentions of a large number of researchers. It has the advantages of low generation complexity, low distance-preserving distortion, and the ability of accelerating the data processing. Given a high-dimensional and sparse data set, such as the thermal radiation space, it is natural to ask whether it could be embedded into a lower dimensional space without suffering great distortion.

Johnson-Lindenstrauss (JL) lemma gives the intuition for designing the infrared sampling towards non-adaptive and stable compressed acquisition method. The original formulation of JL lemma is stated as in [33]: given a parameter $\alpha > 0$ and an integer n_0. If M is a positive integer and $M > O(\alpha^{-2} \log n_0)$, there exists a Lipschitz mapping $\mathbf{V} : \mathbf{R}^N \rightarrow \mathbf{R}^M$ for the set $\mathbf{S} \in \mathbf{R}^N$ which is composed of n_0 points. The mapping is denoted as

$$(1-\alpha) \parallel \mathbf{S}_1 - \mathbf{S}_2 \parallel_2 \leq \parallel \mathbf{V}\mathbf{S}_1 - \mathbf{V}\mathbf{S}_2 \parallel_2 \leq (1+\alpha) \parallel \mathbf{S}_1 - \mathbf{S}_2 \parallel_2 \quad (13)$$

for every $\mathbf{S}_1, \mathbf{S}_2 \in \mathbf{S}$. The JL lemma shows that the set $\mathbf{S}$ in N-dimensional Euclidean space can be mapped on the M-dimensional Euclidean space by the compression matrix $\mathbf{V}$. The JL lemma provides a compression and dimension reduction idea, which if we are able to design the applicable compression projection $\mathbf{V}$, then the data processing calculated in the original high-dimensional space is transformed to a low-dimensional space. Johnson and Lindenstrauss demonstrated inequality (13) and the existence of the $\mathbf{V}$ from the perspective of geometric approximation. However, they did not give the indication of how to design the $\mathbf{V}$ for a specific data set [33].
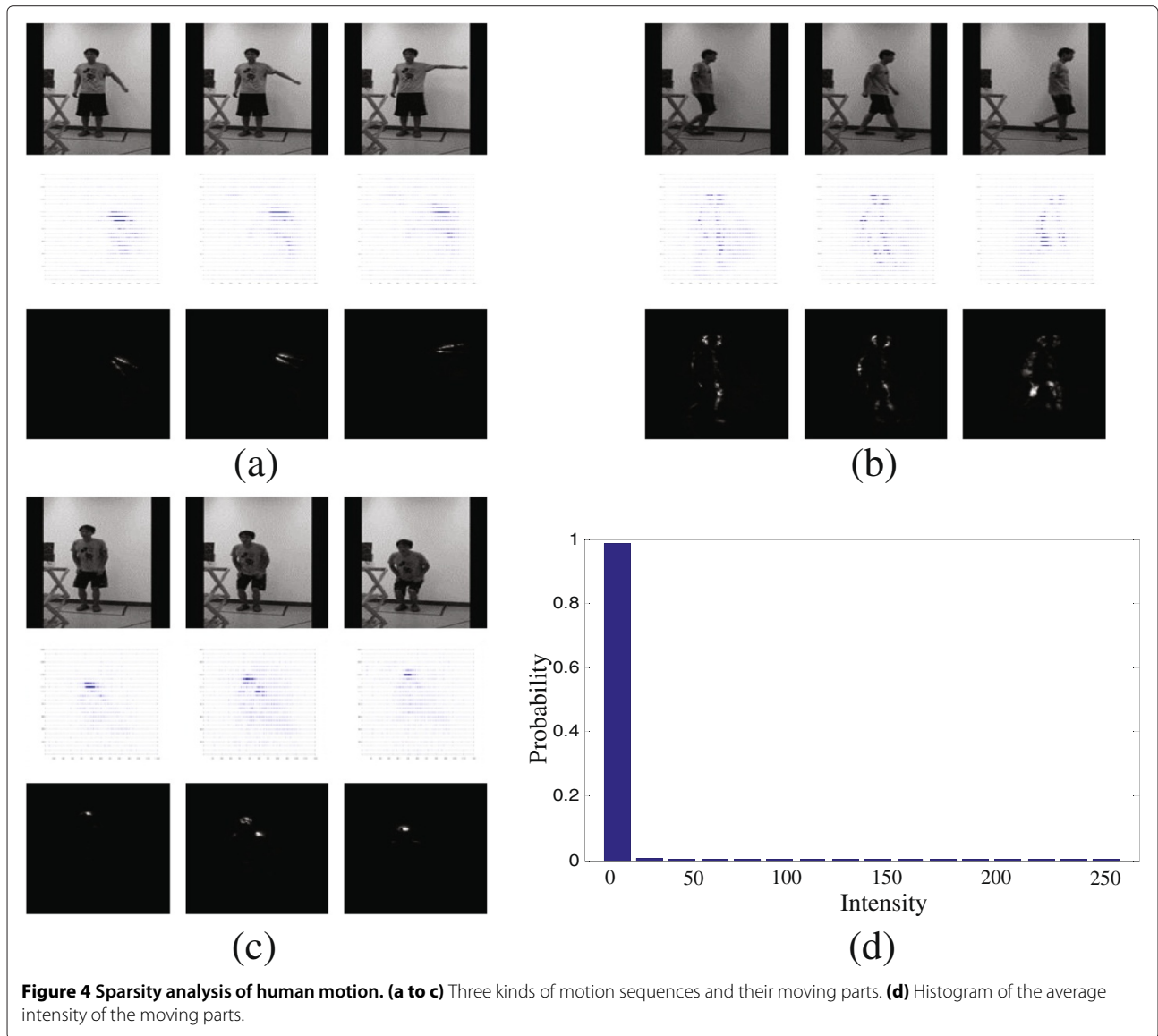
In the subsequent studies, Dasgupta and Gupta provided the proof of JL lemma using the probability theory [34] and pointed out that the entries v_{ij} in the matrix $\mathbf{V}$ were able to be built by the independent Gaussian random variable, meaning as $v_{ij} \sim \mathcal{N}(0, 1)$. When the number of samples satisfies $M \geq 4(\alpha^2/2 - \alpha^3/3)^{-1} \ln n_0$, inequality (13) will hold with high probability. However, the random Gaussian variable v_{ij} contains consecutive floating point numbers; it is difficult to physically integrate or realize in many areas of engineering.

Achlioptas simplified the proof of JL lemma from the perspective of probability theory [35]. The more simple and easily implemented random compression matrix is given. If M is an integer satisfying

$$M \geq \left(\frac{4 + 2\beta}{\alpha^2/2 - \alpha^3/3} \right) \ln(n_0), \quad (14)$$
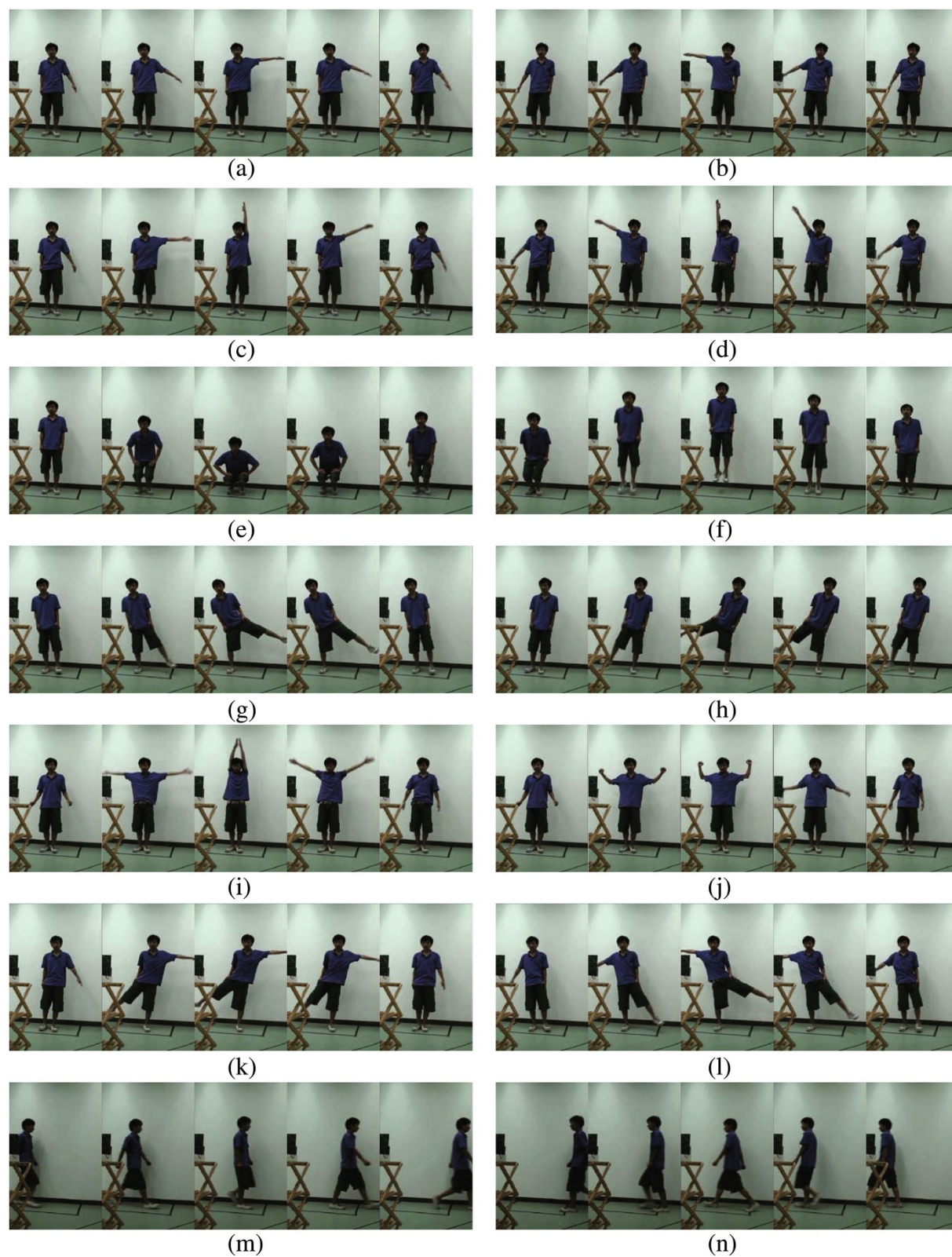
Figure 3 Sequential images of each kind of motion in experiments. (a to n) The 14 kinds of motion in the database.

Figure 4 Sparsity analysis of human motion. (a to c) Three kinds of motion sequences and their moving parts. **(d)** Histogram of the average intensity of the moving parts.

and the projection entry v_{ij} has the form with random Bernoulli distribution

$$v_{ij} := \begin{cases} +1 & \text{with probability } 0.5 \\ -1 & \text{with probability } 0.5, \end{cases} \quad (15)$$

or to meet

$$v_{ij} := \sqrt{3} \begin{cases} 1 & \text{with probability } \frac{1}{6} \\ 0 & \text{with probability } \frac{2}{3} \\ -1 & \text{with probability } \frac{1}{6}, \end{cases} \quad (16)$$

inequality (13) will hold with high probability. The Bernoulli distribution-based compressed projection matrix **V**, due to the simple physical signification, is widely applied to the engineering field.

4 Compressive infrared sampling

4.1 Random matrix-based compressive infrared sampling

Random matrix-based dimension reduction and CS theory provides powerful tools for the design of compressive infrared sampling. According to Achlioptas's statistical results [35], if the v_{ij} is a random variable with the symmetric Bernoulli distribution, then the matrix **V** is able to achieve the dimension reduction and have the approximate distance-preserving property as described in inequality (13). According to Section 2, the random symmetric Bernoulli distribution-based sensing matrix can be achieved using the optical multiplexing. The combination of PIR sensor and the Fresnel lens supports the physical implementation of optical multiplexing. To be specific, for the random matrix stated as in Equation 15, the compressive infrared sampling is achieved by the single-element

PIR sensor and Fresnel lens encoded with random masks; for the random matrix satisfying Equation 16, the compressive infrared sampling is achieved by the random-rotated quad-element PIR sensors and Fresnel lenses encoded with random masks. In this article, we adopt the second physical method for designing the compressive infrared sensing.

If we assume that the body's movement is constrained in a fixed interaction space, so that a specific motion with certain semantics is composed of an infrared sequence of moving body parts. According to sampling theory for the radiation space, the original interface space can be divided into coarser-grained and non-overlapping cells. This division is achieved by the isometric mapping between the interface space and Fresnel lens. The sub-cells on the Fresnel lens have the homologous distribution to the traditional visual sensor pixels. When the body moves into the sensing space, the feature of motion can be represented by the changes of thermal radiation. According to the theory of random compression proposed by Achlioptas [35], we use the random distribution of Equation 13 to modulate each of the FOV of the Fresnel lens. Figure 2b shows the Fresnel lens containing 25 non-overlapping cells; thus, each PIR sensor is divided into four sub-cells to form a symmetrical subtraction. We first select two thirds of the all FOVs of the Fresnel lens on each PIR sensor and mask them. This operation will make the changes of thermal radiation in the sub-cell to be not visible for the PIR sensor. Then, we randomly select a half from all the sensors and then rotate them 90°. The above combined operation enables the sensing matrix of the sensor array to have a pseudo-random property and forms the fashion of compression infrared sampling for motion information:

$$\mathbf{m}(t) = \mathbf{V}\mathbf{s}(t). \tag{17}$$

Figure 5 presents the diagram of the proposed compressed infrared sampling. In this article, we employ 16 PIR sensors to measure the motion information in raw space parallelly, having $\mathbf{m} \in \mathscr{R}^{16}$. The sub-lens on Fresnel lens is further divided into four sub-cells using the quad-element PIR sensors, so the original measured space is divided into $25 \times 4 = 100$ non-overlapping sub-cells and $\mathbf{s} \in \mathscr{R}^{100}$. The measurement matrix $\mathbf{V} \in \mathbf{R}^{16 \times 100}$ compresses the original 100-dimensional states of the thermal radiation changes into the 16-dimensional sensor outputs with a non-adaptive way.

4.2 Statistical analysis of random measurement matrix

Figure 6a presents the pseudo-random measurement matrix used in this article. The white pixels in this figure represent the visible entries '1', the black pixels represent the visible entries '−1', while the gray pixels denote the invisible entries '0'. Given the measurement matrix, it is necessary to verify its effectiveness for inequality (13). However, due to the unknown knowledge of specific set $\mathbf{S}$ and its element number n_0, it is hard to demonstrate that inequality (13) holds with high probability directly.

Kaski proposed the cosine of the angle between two vectors to measure the distortion of similarity when random compression is used [36]. His method gives quantitative assessments on random compression. In this article, we employ his statistical results to assess the effectiveness of pseudo-random measurement matrix. First, assuming two vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ are given, the inner product of two measurement vectors $\mathbf{m}_1$ and $\mathbf{m}_2$ by the random matrix $\mathbf{V}$ can be expressed as follows [36]:

$$\mathbf{m}_1^T \mathbf{m}_2 = \mathbf{s}_1^T \mathbf{V}^T \mathbf{V} \mathbf{s}_2, \tag{18}$$

The matrix $\mathbf{V}^T \mathbf{V}$ can be decomposed as $\mathbf{V}^T \mathbf{V} = \mathbf{I} + \epsilon$, where $\mathbf{I}$ is the identity matrix and the matrix ϵ denotes the entities off the diagonal:

$$\epsilon_{ij} := \begin{cases} \mathbf{v}_i^T \mathbf{v}_j & \text{for } i \neq j, \\ 0 & \text{for } i = j. \end{cases} \tag{19}$$

Then, Equation 18 can be rewritten as

$$\mathbf{m}_1^T \mathbf{m}_2 = \mathbf{s}_1^T \mathbf{s}_2 + \sum_{i \neq j} \epsilon_{ij} s_{1_i} s_{2_j}. \tag{20}$$

The diagonal entities in matrix $\mathbf{V}^T \mathbf{V}$ should be equal to unity since the measurement vector $\mathbf{v}_i$ has the normalized weights with equal probability, while the non-diagonal entries ϵ_{ij} should be equal to zero [36]. However, the vector $\mathbf{v}_i$ and $\mathbf{v}_j$ are not orthogonal in practice, which causes the non-diagonal entries ϵ_{ij} to be small but not to zero. The similarities of the original vectors will generate distortions with the non-zero entries ϵ_{ij}, which can be seen in Equation 20.

If the random measurement matrix is fixed, it is possible to use the statistical properties of entries ϵ_{ij} to analyze the distortions generated by compression. The ideal average of ϵ_{ij} is $E[\epsilon_{ij}] = 0$, with approximate variance $\sigma_\epsilon^2 \approx 1/M = 0.0625$. According to the variance σ_ϵ^2, we can infer that the more measurements and sparser original vectors will generate smaller distortions when random compression is used. The actual average value of ϵ_{ij} based on previous description is −0.0105, and the variance is 0.1518. Figure 6b gives the ideal and actual distribution of ϵ_{ij}. Although the actual variance is larger than the ideal one, the ϵ_{ij} will have the smaller value associated with the sparse vector.

5 Experiments and results
5.1 Experimental setup
Figure 7a presents a PIR sensor module; both the length and width are 4 cm. The PIR sensor unit locates in the

Figure 5 Pseudo-random-coded Fresnel lense-based compressive infrared sampling.
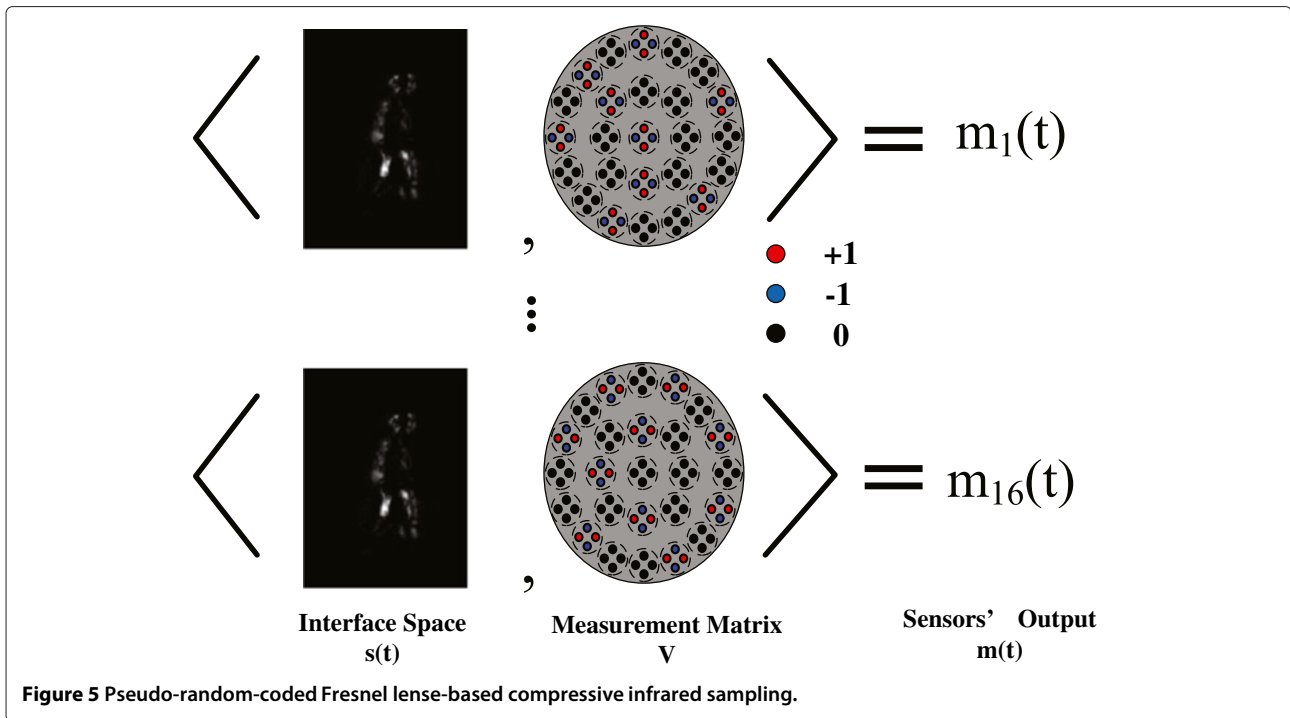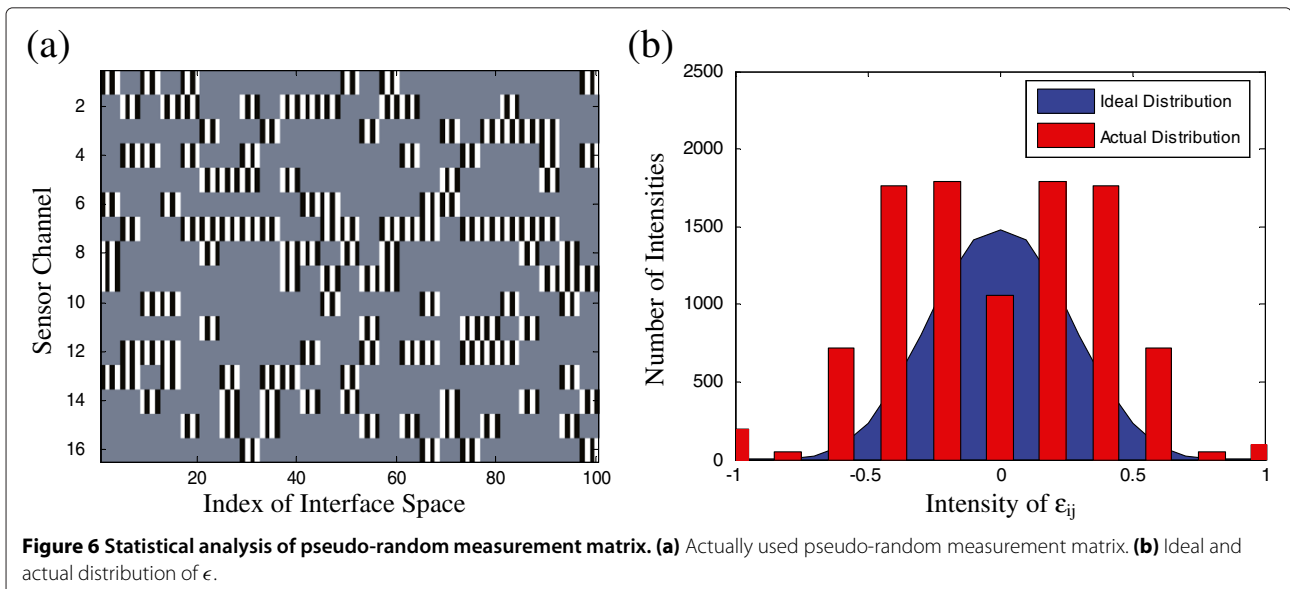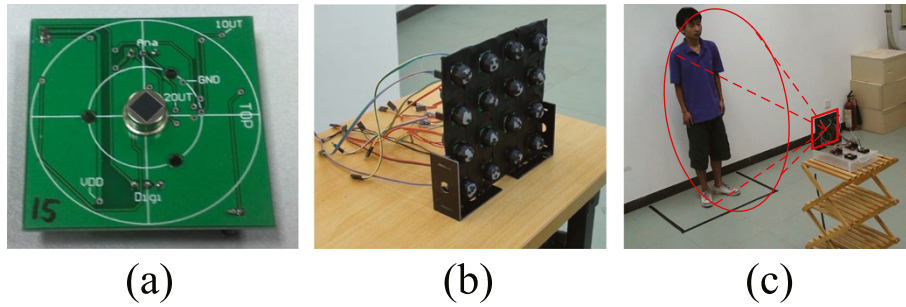
center of the module. Figure 7b shows the prototype of our proposed sensing system for the acquisition of motion information. The sensor array is composed of 16 quad-element PIR sensors. The PIR sensors *D205b* commercially available are employed for sampling the thermal radiation changes [37]. Both the horizontal and vertical range of the sensor's FOV are about 95°. We use the smart system-on-chip (SoC) *CC*2430 to sample the signal with the frequency of 10 Hz. Figure 7c shows the experimental setup for real-time measurement of the body motion. We assume that the movements are restricted inside the virtual box in front of the person. The distance between the sensor unit and the subject is 1.5 m. When the limbs of the body move through the interface region, the corresponding sensors will be activated.

In order to test the validity of the proposed method, we designed a set of gymnastics to build a motion database. There are 14 kinds of gymnastic motions, including the local movements generated by the arms and legs and the synergistic motions of the upper and lower limbs. The

Figure 6 Statistical analysis of pseudo-random measurement matrix. (a) Actually used pseudo-random measurement matrix. **(b)** Ideal and actual distribution of ϵ.

Figure 7 Prototype device of the compressive infrared sampling. (a) A PIR sensor module. **(b)** Hardware prototype. **(c)** Typical experimental scenarios.

ambient temperature is kept at 25°C. Figure 3 gives the sequential images of each kind of motion. Figure 8 shows the typical sensor output vector when a person walks across the interface space.

In the following experiments, we do assume that the gymnastic motion played by each member is required to be constrained in the interaction region. The spatial and temporal differences of the same semantic motions are small. However, the general motion recognition and behavior understanding should be considered the more challenging problems. Although our assumptions simplify the motion recognition, the purpose of this article is to exploit the compressed infrared sampling to directly acquire human motion in a compressive form.

5.2 Measurement sequence segmentation

We use the energy-based detection method to segment the motion signal directly. First, the collected signal is normalized by removing its direct current component. Second, the normalized signal is enframed by splitting the signal into overlapping frames and denoted as

$$\mathbf{m}_s(t) = [m_s(t - d), \dots m_s(t) \dots, m_s(t + d)]^T, \quad (21)$$

where $m_s(t)$ is the sth sensor's frame vector at time t and d is the width of the radius in enframing operation. In our subsequent experiments, the d is 5. Third, the energy signal is obtained by accumulating the squared value in each column of the enframed signal, and a predefined threshold value is used for generating the start-end boundaries. The overall motion segmentation may be determined based on the synthetically enframed signal $\mathbf{M}(t) = [\mathbf{m}_1(t) \dots \mathbf{m}_{16}(t)]^T$.
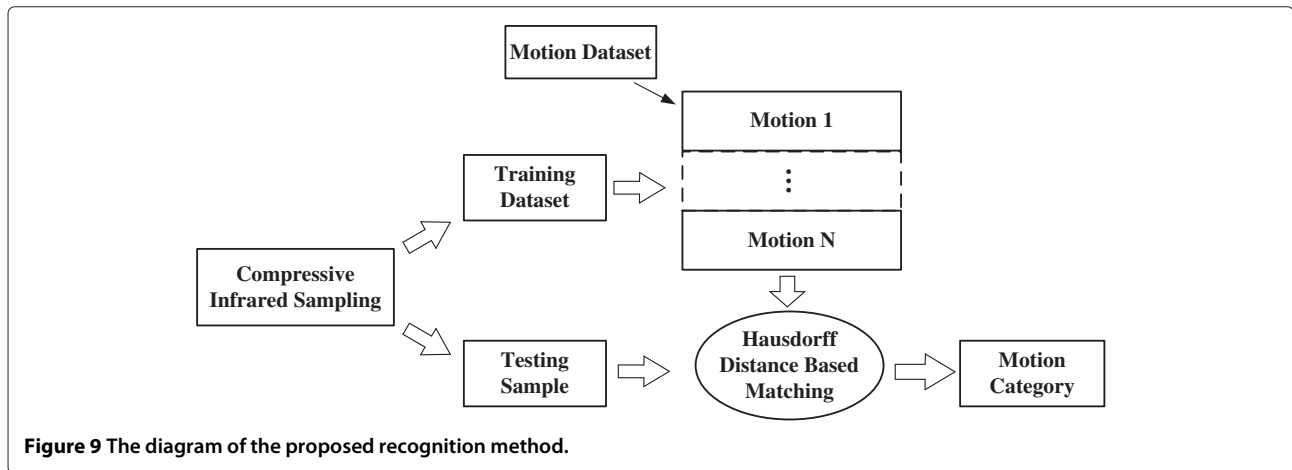
5.3 Hausdorff distance-based recognition method

Hausdorff distance is a lightweight tool to measure the similarity between two different temporal sequences. It is able to overcome the different lengths of time between changes in the sequence and timing offset, and the output also includes the temporal constraints implicitly. Hausdorff distance is widely used in the temporal sequence matching and recognition [38,39]. Figure 9 gives the flow chart of recognition method.

Let us assume that a test sequence $\mathbf{M}_1 = [\mathbf{m}_1(1), \dots, \mathbf{m}_1(t_1), \dots, \mathbf{m}_1(T_1)]$ is given, and the reference sequence is $\mathbf{M}_2 = [\mathbf{m}_2(1), \dots, \mathbf{m}_2(t_2), \dots, \mathbf{m}_2(T_2)]$. T_1 and T_2 are the length of the two signal sequences,
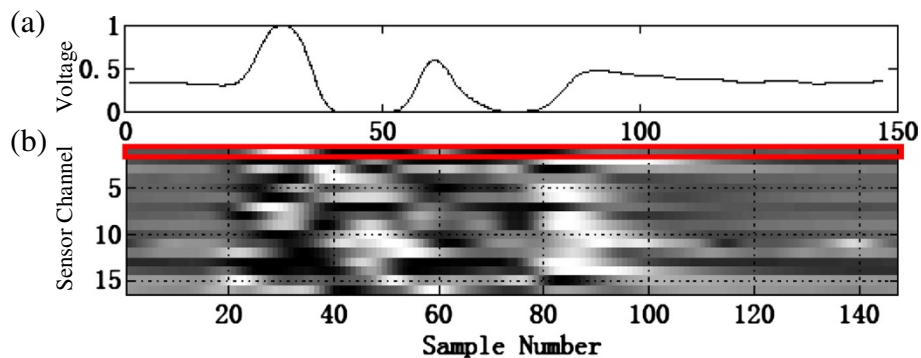
Figure 8 Typical signals collected from the sensor array when a person walks through the interface space. (a) Signal collected from sensor 1. **(b)** Fused sequence by concatenating 16 sensor outputs into a higher dimensional vector.

Figure 9 The diagram of the proposed recognition method.

respectively. The average Hausdorff distance between two sequences is

$$D_h(\mathbf{M}_1, \mathbf{M}_2) = \mathop{\mathbf{mean}}_{1 \le t_1 \le T_1} \left(\min_{1 \le t_2 \le T_2} (\|\hat{\mathbf{M}}_1(t_1) - |\hat{\mathbf{M}}_2(t_2)\|) \right),$$

(22)

where $\hat{\mathbf{M}}_1$ and $\hat{\mathbf{M}}_2$ are the energy normalized sequence. The Hausdorff distance is denoted as

$$D_{\text{hausdorff}}(\mathbf{M}_1, \mathbf{M}_2) = \frac{1}{2} [D_h(\mathbf{M}_1, \mathbf{M}_2) + D_h(\mathbf{M}_2, \mathbf{M}_1)].$$

(23)

If we have pre-stored some reference samples of different motions in the database, then a coming test sequence needs to be matched with each reference sample. The category of the test sequence is determined by the following nearest neighbor rule:

$$w^* = \arg\min_{1 \le w \le W} D_{\text{hausdorff}}(\mathbf{M_1}, \mathbf{M}_w),$$

(24)

where w^* is the category to be recognized and W is the total number of categories in the motion database.
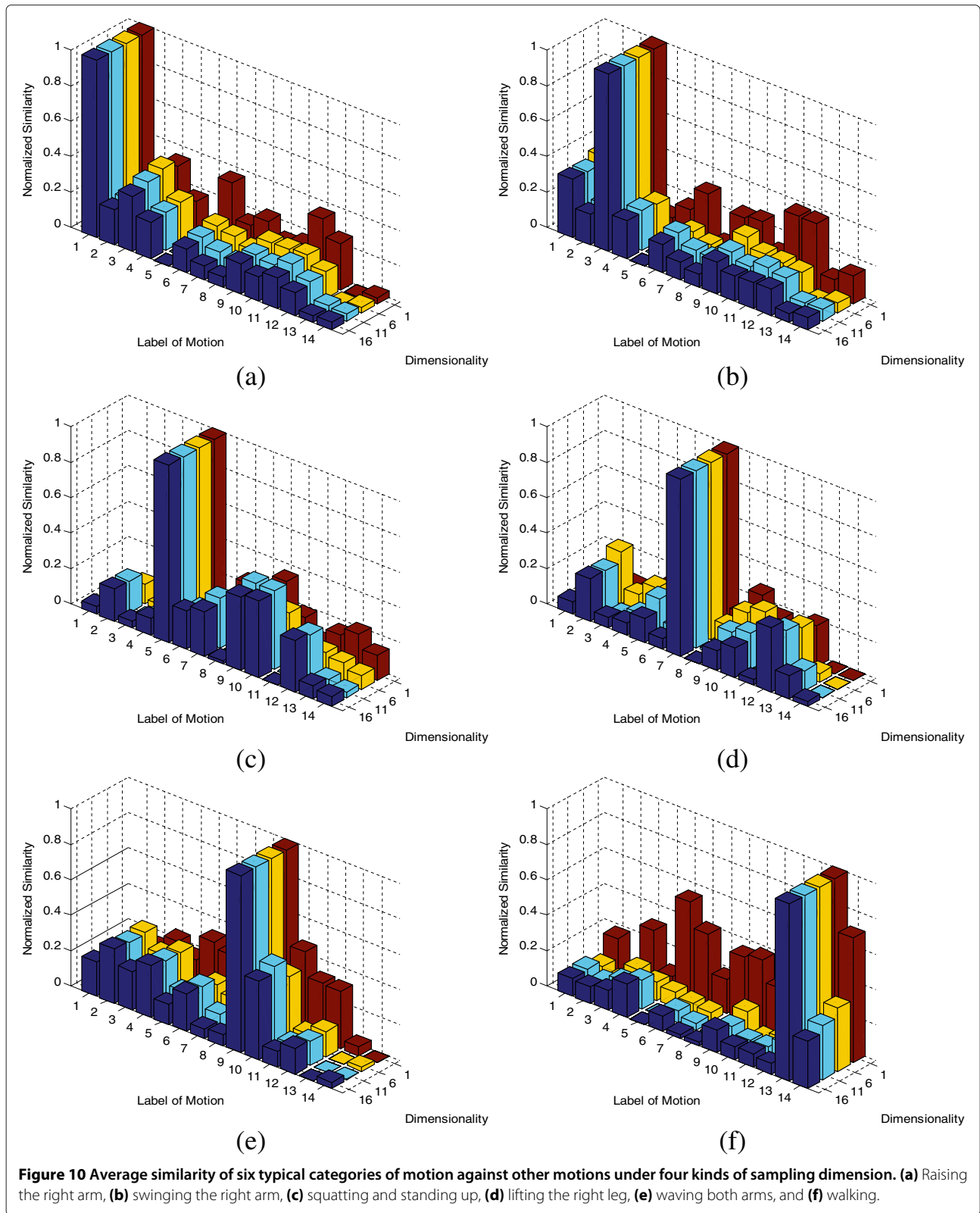
5.4 Experimental results and discussion

The number of measurements or dimensions is an important parameter in the proposed sensing method. By the statistical experiments, we compare the recognition performances in different numbers of dimension. For the collection of the reference set, we randomly select half of the samples from each motion set as the template samples, and the remaining samples are used for test and analysis. The following statistical experimental results are based on 30 cross-validations.

We first statistically give the average similarity of six typical categories of motion against the other motions under different sampling dimensions, as shown in Figure 10. The normalized similarity is calculated based

on the reciprocal of average Hausdorff distance. From the experimental results, it can be seen that each kind of motion samples has the confusing similarity against the other reference samples with low sampling dimension, especially the one-dimensional infrared sampling on synergistic motions of the upper and lower limbs. Thus, the acquired information cannot support to determine the category of the testing motion. When the number of compressed sampling dimension increases to 6 or 11, it is obvious that each kind of the testing motion only has the maximum normalized similarity with the truly recognized category.
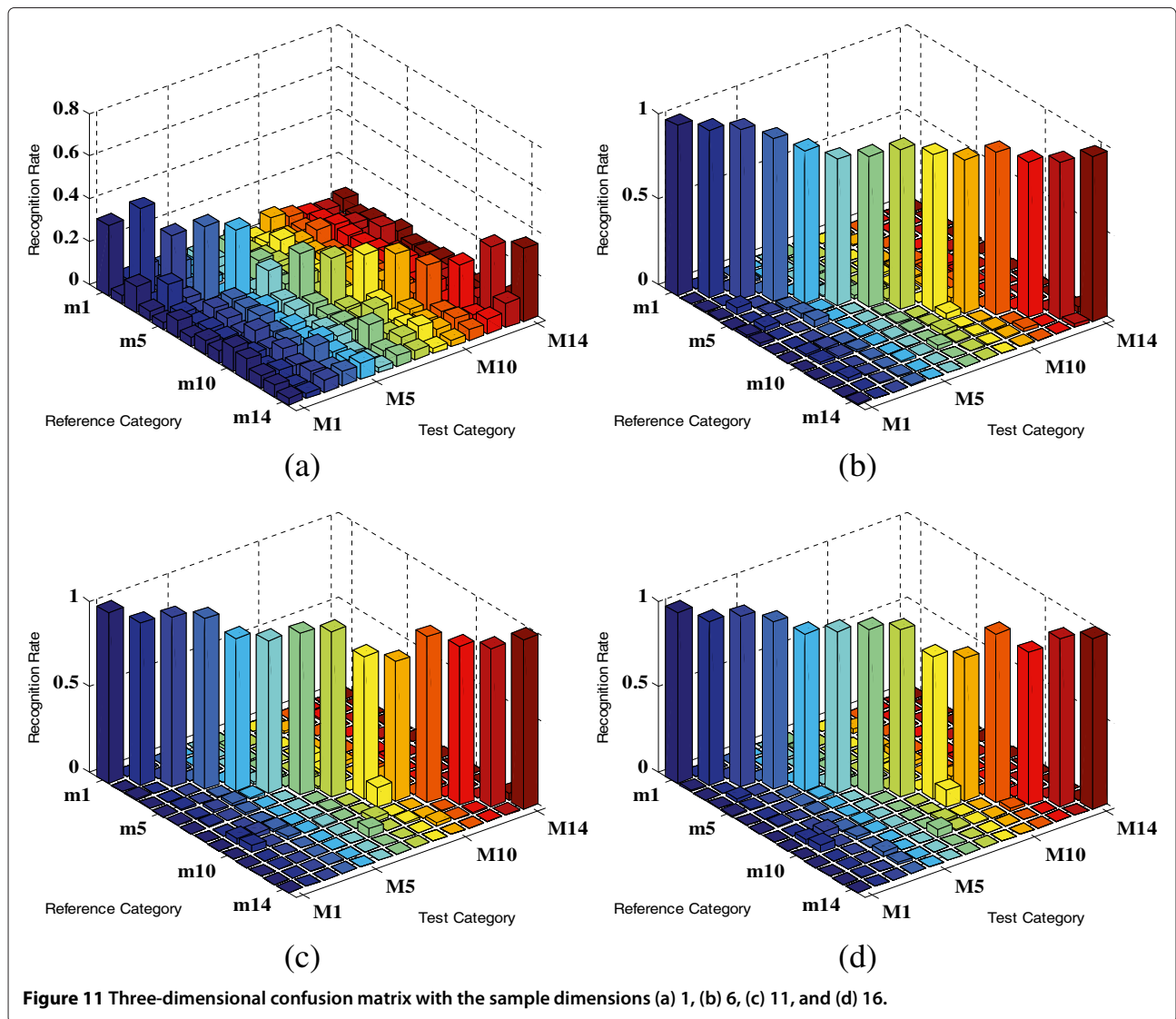
Figure 11 shows the three-dimensional confusion matrix based on the proposed sensing method with four kinds of sampling dimension. The confusion matrix shows the motion reference category (left) versus the test category (right). Each bar (m_i, M_j) in the matrix denotes the percentage of motion M_j being recognized as motion m_i. The percentage of the correctly recognized motion can be obtained by calculating the trace of matrix. The remaining lower bars present the percentage of misclassification. It can be seen when the sample dimension is 1, due to the small sample dimension, the compression sampling signal cannot form the distance-preserving map with the original high-dimensional space. It causes a highly erroneous recognition rate. It is also obvious that when the number of compressed sampling dimension increases to 6 or 11, the proposed method is able to obtain higher correct recognition rate.

Table 1 presents the average correct rate with four types of compressive dimension using the proposed sensing method. When the compressed sampling dimension is 1, the average correct recognition rate is low. The reason for this phenomenon is that if the sampling number is small, the isometric distance between the original high-dimensional state signal of motion and the low-dimensional sampling signal cannot achieve low

Figure 10 Average similarity of six typical categories of motion against other motions under four kinds of sampling dimension. (a) Raising the right arm, **(b)** swinging the right arm, **(c)** squatting and standing up, **(d)** lifting the right leg, **(e)** waving both arms, and **(f)** walking.

distortion with high probability. While the number of compressive dimension reaches to 6, the sensing method will get a better performance.

Table 1 also gives the average recognition time consumed by different compressed sampling dimensions. The relevant algorithms in our experimental studies run on

Figure 11 Three-dimensional confusion matrix with the sample dimensions (a) 1, (b) 6, (c) 11, and (d) 16.

an Intel Pentium 4 2.8-GHz computer by Matlab codes. It can been seen that the larger compressed sampling dimension will consume the longer processing time, and it has approximately linear growing trend based on the Hausdorff distance matching method. It should be noted that the recognition is very fast, since the testing feature is represented by a low-dimensional sequence. Even in the case of 16 dimensions, the average recognition

time will not exceed 36 ms. For some real-time recognition system, designers can make a tradeoff on the number of sampling dimension according to the sensing resources, data processing time, and average recognition rate.

6 Conclusions

A compressive sampling-based PIR sensor array for human motion sensing has been developed and evaluated in this article. The PIR sensors and pseudo-random masked Fresnel lens arrays are used for efficient motion feature transformation. Compressive dimension reduction theory supports that sparse or compressible motion information can still preserve its statistical feature in the measurement space. By modeling the low-dimensional sequential features, we can achieve motion recognition, which is confirmed by experiments.

Table 1 The average correct rates and recognition time for motion detection with four types of dimensionality

	Dimensionality			
	1	6	11	16
Average correct rate (%)	29.47	91.81	94.17	94.83
Average recognition time (ms)	24.66	28.23	31.53	35.05

The proposed sensing method gives rise to two main advantages. First, the sensing module is able to acquire and compress motion information synchronously. Second, the problem of recognizing a high-dimensional signal is transformed into the low-dimensional space, and the computational time of recognition can be saved.

However, there are some limitations in the practical application. First, the sensing method relies on the assumption that the motion is constrained in a predefined interface region, so the distance between the body and sensor array is fixed. Indeed, many motions are often associated with more freedom. Second, different from the classical compressive sensing paradigm, the low-dimensional sensor outputs are used for motion classification without reconstruction. Developing more effective performance analysis method for the sensor system is our future work. Third, other parameters such as the location of the sensor unit and the distance between the sensor unit and the subject to be captured could also be considered to improve the performance of the system. Although the prototype sensor system has the limitations, the sensing method has provided a proof of concept with respect to using a combination of PIR sensor and CS theory for compressive motion acquisition.

Competing interests

Acknowledgements

Author details
[1]Department of Electronic Science, Huizhou University, Guangdong 516001, China. [2]College of Physics & Electronic Information Engineering, Wenzhou University, Wenzhou 325035, China.

References
1. C Yang, Y Hsu, A review of accelerometry-based wearable motion detectors for physical activity monitoring. Sensors. **10**(8), 7772–7788 (2010)
2. A Yilmaz, O Javed, M Shah, Object tracking: a survey. ACM Comput Surv. **38**(4), 1–45 (2006)
3. P Turaga, R Chellappa, V Subrahmanian, O Udrea, Machine recognition of human activities: a survey. IEEE Trans. Circ. Syst. Video Technol. **18**(11), 1473–1488 (2008)
4. R Poppe, A survey on vision-based human action recognition. Image Vis Comput. **28**(6), 976–990 (2010)
5. L Wang, D Suter, Visual learning and recognition of sequential data manifolds with applications to human movement analysis. Comput. Vis. Image Understand. **110**, 153–172 (2008)
6. I Jolliffe, *Principal Component Analysis*. (Springer, Berlin, 1986)
7. J Tenenbaum, V Silva, J Langford, A global geometric framework for nonlinear dimensionality reduction. Science. **290**(5500), 2319–2323 (2000)
8. S Roweis, L Saul, Nonlinear dimensionality reduction by locally linear embedding. Science. **290**(5500), 2323–2326 (2000)
9. I Tošić, P Frossard, Dictionary learning. IEEE Signal Process. Mag. **28**(2), 27–38 (2011)
10. DL Donoho, Compressedsensing Trans. IEEE. Inf. Theory. **52**, 1289–1306 (2006)
11. R Baraniuk, Compressive sensing. IEEE Signal Process. Mag. **24**(4), 118–121 (2007)
12. E Candès, M Wakin, An introduction to compressive sampling. IEEE Signal Process. Mag. **25**(2), 21–30 (2008)
13. M Duarte, M Davenport, D Takhar, J Laska, T Sun, K Kelly, R Baraniuk, Single-pixel imaging via compressive sampling. IEEE Signal Process. Mag. **25**(2), 83–91 (2008)
14. J Romberg, Imaging via compressive sampling. IEEE Signal Process Mag. **25**(2), 14–20 (2008)
15. G Johansson, Visual perception of biological motion and a model for its analysis. Atten. Percept. Psychophys. **14**, 201–211 (1973)
16. G Johansson, Visual motion perception. Sci. Am. **232**, 76–88 (1975)
17. A Oikonomopoulos, I Patras, M Pantic, Spatiotemporal salient points for visual recognition of human actions. IEEE Trans. Syst., Man Cybern. Part B: Cybernetics. **36**, 710–719 (2005)
18. P Dollar, V Rabaud, G Cottrell, S Belongie, Behavior recognition via sparse spatio-temporal features, in *2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005* (IEEE, Piscataway, 2005), pp. 65–72
19. J Niebles, H Wang, L Fei-Fei, Unsupervised learning of human action categories using spatial-temporal words. Int. J. Comput. Vis. **79**(3), 299–318 (2008)
20. R Fuksis, M Greitans, E Hermanis, Motion analysis and remote control system using pyroelectric infrared sensors. Electron. Electrical Eng. **86**(6), 69–72 (2008)
21. A Sixsmith, N Johnson, A smart sensor to detect the falls of the elderly. IEEE Pervasive Comput. **3**(2), 42–47 (2004)
22. J Burchett, M Shankar, AB Hamza, BD Guenther, N Pitsianis, DJ Brady, Lightweight biometric detection system for human classification using pyroelectric infrared detectors. Appl. Optics. **45**(13), 3031–3037 (2006)
23. JS Fang, Q Hao, DJ Brady, BD Guenther, KY Hsu, Real-time human identification using a pyroelectric infrared detector array and hidden Markov models. Opt. Express. **14**(15), 6643–6658 (2006)
24. JS Fang, Q Hao, DJ Brady, BD Guenther, KY Hsu, A pyroelectric infrared biometric system for real-time walker recognition by use of a maximum likelihood principal components estimation (MLPCE) method. Opt. Express. **15**(6), 3271–3284 (2007)
25. M Shankar, JB Burchett, Q Hao, BD Guenther, DJ Brady, Human-tracking systems using pyroelectric infrared detectors. Opt. Eng. **45**, 106401 (2006)
26. Q Hao, D Brady, B Guenther, J Burchett, M Shankar, S Feller, Human tracking with wireless distributed pyroelectric sensors. IEEE Sensors J. **6**(6), 1683–1696 (2006)
27. T Liu, X Guo, G Wang, Elderly-falling detection using distributed direction-sensitive pyroelectric infrared sensor arrays. Multidimensional Syst. Signal Process. **23**, 451–467 (2012)
28. T Liu, J Liu, Feature-specific biometric sensing using ceiling view based pyroelectric infrared sensors. EURASIP J. Adv. Signal Process (2012). doi:10.1186/1687–6180–2012–206
29. S Mitra, T Acharya, Gesture recognition: a survey. IEEE Trans. Syst. Man Cybernet. Part C: Appl. Rev. **37**(3), 311–324 (2007)
30. A Hossain, M Rashid, Pyroelectric detectors and their applications. IEEE Trans. Ind. Appl. **27**(5), 824–829 (1991)
31. D Brady, N Pitsianis, X Sun, Reference structure tomography. J. Opt. Soc. Am. A. **21**(7), 1140–1147 (2004)
32. BD Lucas, T Kanade, An iterative image registration technique with an application to stereo vision, in *The 7th International Joint Conference on Artificial intelligence,* vol. 2 (Morgan Kaufmann, San Francisco, 1981), pp. 674–679
33. W Johnson, J Lindenstrauss, Extensions of Lipschitz mappings into a Hilbert space, in *Conference in Modern Analysis and Probability,* vol. 26 (American Mathematical Society, Providence, 1984), pp. 189–206
34. S Dasgupta, A Gupta, An elementary proof of the Johnson-Lindenstrauss Lemma. Technical report, International Computer Science Institute, Berkeley, 1999

35. D Achlioptas, Database-friendly random projections: Johnson-Lindenstrauss with binary coins. J. Comput. Syst. Sci. **66**, 671–687 (2003)

36. S Kaski, Dimensionality reduction by random mapping: fast similarity computation for clustering, in *The 1998 IEEE World Congress on Computational Intelligence,* vol. 1 (IEEE, Piscataway, 1998), pp. 413–418

37. The PIR Sensor Co. Ltd. http://pirsensor.bloombiz.com. Accessed 7 May 2013

38. DG Sim, OK Kwon, RH Park, Object matching algorithms using robust Hausdorff distance measures. IEEE Trans. Image Process. **8**(3), 425–429 (1999)

39. SH Kim, RH Park, An efficient algorithm for video sequence matching using the modified Hausdorff distance and the directed divergence. IEEE Trans. Circ. Syst. Video Technol. **12**(7), 592–596 (2002)