Computer Networks 138 (2018) 192–200

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

HealCam: Energy-efficient and privacy-preserving human vital cycles monitoring on camera-enabled smart devices

Qing Yang[a], Yiran Shen[a,*], Fengyuan Yang[a], Jianpei Zhang[a], Wanli Xue[b], Hongkai Wen[c]

[a] Harbin Engineering University, Building 21A, No.145 of Nantong Street, Nangang District, Harbin, China
[b] School of Computer Science and Technology, University of New South Wales, China
[c] Department of Computer Science, Warwick University, China

ARTICLE INFO

Article history:
Received 15 July 2017
Revised 15 January 2018
Accepted 28 March 2018
Available online 5 April 2018

Keywords:
Compressive sensing
Smart camera
Privacy preservation

ABSTRACT

In this paper, we present HealCam, an energy-efficient and privacy-preserving human vital signs (e.g., respiration cycles) monitoring system on camera-enabled smart devices. HealCam incorporates the related theories of compressive sensing in its system to reduce the sampling rate while preserving data privacy. HealCam saves significant cost on video processing via low-rate and non-uniform random sampling. It also provides a privacy-preserved data collection and enquiry service via lightweight compressive encryption and decryption scheme. According to our evaluations on real datasets, HealCam achieves high accuracy on respiration cycles reconstruction with extremely low average frame rate, i.e., 1 FPS, via non-uniform random sampling compared with traditional uniform sampling strategy. Then we implement HealCam on smartphones to evaluate its resource consumption. The results show that, HealCam is 23.6 times more energy efficient and 26.7 times faster than the original approach on video processing. Its data encryption component is 172 times faster while consumes only 1.01% energy compared with the corresponding state-of-the-art.

1. Introduction

Human vital signs, e.g., respiration and heartbeat cycles, are important references for medical diagnoses and healthcare applications. These vital cycles can be monitored by the dedicated equipments in hospital, e.g., an ElectroCardioGram (ECG) [1] for heartbeat cycles monitoring. However, these equipments are expensive, cumbersome and require special expertise to use; therefore are not appropriate for usage in collecting daily health statistics of patients or the elderly in nursing-house, aged-care facilities or even at home. As the vision of smart healthcare, the future of medical applications should leverage the pervasive smart devices under different environments to provide intelligent, seamless and non-intrusive healthcare services. Therefore, the traditional dedicated and cumbersome equipments are not appropriate.

Extracting the cycles of respiration and heartbeat from continues video recordings of human faces has been extensively studied in the literature [2] and the current state-of-the-art is based on detecting the subtle head motions [3] using the embedded cameras on high performance PCs. The video-based approaches are quite promising as they do not require any special expertise to use and the high-resolution on-board cameras are now pervasive on different embedded smart devices including smartphones, Augmented Reality/Virtual Reality (AR/VR) glasses, tablets and etc. (see Fig. 1). The flexibility and user experience of the video-based approaches will be significantly improved if implemented on the camera-enabled smart devices. However, video processing is known to be computationally prohibitive for resource-constrained smart devices and the computational issues are ignored in previous research.

A straightforward solution for reducing the video computation cost on smart devices is offloading the computationally intensive tasks to the high performance computing facilities, such as local servers or remote cloud computing services [4]. However, offloading the video recordings of individuals' faces is quite privacy-intrusive. Moreover, the extracted vital cycles themselves are significantly private as it can be used to infer individuals' health status. Individuals' health related data has substantially commercial values and it is highly likely that it will be targeted by hackers in the future in the same way that credit card information is now.

In this paper, we focus on monitoring the respiration cycles and addressing the above issues, we propose a novel energy-efficient and privacy-preserving human respiration cycles monitoring system, HealCam, on camera-enabled smart devices by incorporating

* Corresponding author.
E-mail addresses: yangqing@hrbeu.edu.cn (Q. Yang), shenyiran@hrbeu.edu.cn (Y. Shen), yangfengyuan@hrbeu.edu.cn (F. Yang), zhangjianpei@hrbeu.edu.cn (J. Zhang), xuewanli.lee@gmail.com (W. Xue), hongkai.wen@dcs.warwick.ac.uk (H. Wen).

https://doi.org/10.1016/j.comnet.2018.03.033
1389-1286/

Fig. 1. Examples of camera-enabled smart devices.

relevant theories of compressive sensing. The contributions of this paper are as follows,

- In this paper, we propose HealCam for video-based human respiration cycles monitoring on camera-enabled smart devices. To the best of our knowledge, this is the first relevant system running on resource-constrained devices.
- As the minimal frame rate of uniform approach is bounded by the theory of Nyquist frequency and noises, to reduce the sampling rate, we propose a non-uniform random sampling strategy inspired by compressive sensing, instead of traditional uniform sampling to reduce the local video computation cost on the smart devices. According to the evaluation on the real datasets, it achieves significantly higher reconstruction accuracy at extremely low sampling rate compared with the uniform sampling strategy.
- We incorporate a lightweight compressive encryption and decryption algorithm in HealCam, based on compressive sensing, for privacy-preserved data collection and enquiry with smart devices.
- At last, we implement HealCam on smartphones to evaluate its resource consumptions. The results show that HealCam significantly reduces the energy consumption on both local video computing and data encryption.

This paper is organized as followed: Section 2 gives introduction about the work related to monitoring human vital cycles and corresponding encryption methods. Section 3 mentions the detail of HealCam's system architecture. In Section 4 results of experiments conducted on real world dataset are given. Finally Section 5 concludes the paper.

2. Related work

Daily human vital cycles monitoring via smart devices has become one of the research topics with the introduction of Smart Health Projects [5]. Studies [6] showed that the accurate daily health data enabled us to understand how our bodies were functioning and could be helpful in detecting serious illnesses in early stage. Nanotechnology and wireless communications are the most common methods used in smart health applications. In [7], a smart shirt was designed to measure ECG and acceleration signals continuously and transmitted the data to a device connected to an ad hoc network. ZigBee and mobile phones were used to measure ECG and blood glucose in [8]. In another study [9], a sensor integrated into clothing was designed to measure biochemical changes in sweat. The study showed that the sensor had the potential to record real-time variations in sweat during exercise.

Another research mainstream is to process the video recordings of human faces to detect the subtle changes caused by heartbeat or breath [10]. The existing related research can be vastly divided into two categories: color-change based [11,12] and head-motion based approaches[3]. The problem of color-change based approaches are they suffer from the noises from environment lighting changes. However, the existing efforts focused on improving the accuracy, they are not applicable on resource-constrained smart devices due to overwhelming computation from video processing.

With the pervasive availability of high-speed network and high performance cloud services, the most recent research on smart devices is seeking the assistance from the cloud for computation and storage. However, the privacy issues arise when uploading users' data to the cloud. A number of encryption schemes [13] have been proposed to guarantee the data security. Advanced Encryption Standard(AES) is an algorithm proposed in [14], which converts the data files from plaintext format into an incomprehensible format that is called cipher text which could not be read by humans to prevent the unauthorized users from gaining access to the data files. Prabhakar and Joseph [15] proposed a scheme using Advanced Encryption Standard with key length 256 bits(AES-256) to provide complete security for the data during all the stages. These methods could provide ample protection, however, their computation consumptions are intensive. Many other researches have been raised to encrypt the query processing in cloud computation. CryptDB [16] is an online encryption system that allows executing structured query language(SQL) queries over encrypted data using a collection of SQL-aware encryption schemes. Talos [17] is introduced by Shafagh et. al as a system that encrypts query processing for IoT data stored in cloud with encryption keys held by users. However, both of the researches are not focus on energy consumption and efficiency as they are designed for online applications and cloud environment respectively.

Compressive sensing is widely adopted in the area of data collection, which utilizes a sensing matrix to randomly sample and efficiently reconstruct data. Inspired by the properties of compressive sensing, instead of encrypting the raw data directly, in this paper, we propose a new compressive sensing based encryption scheme to preserve the privacy during query processing, which is both secured and lightweight.

3. System architecture

The system architecture of HealCam can be vastly divided into two major stages. The first is the energy-efficient respiration cycles extraction and encryption; the second is privacy-preserving historical records enquiry. Specifically, in the first stage, the respiration cycles are extracted from tracking the head motions of the target in a continuous video stream on smartphones. Non-uniform random sampling is applied to reduce the number of frames for further processing to significantly save the video processing cost. Then to boost the security of the collected data, random Gaussian perturbation is adopted to encrypt the original data and the encrypted data is transferred to and stored in the remote cloud or cloudlet for future enquiry. In the second stage, the user's client on smartphone computes and sends the recovery matrix to the cloud for enquiring historical records. The cloud accesses to the encrypted historical record as demand and computes the sparse representation vector by solving the computationally intensive ℓ_1 optimization. Then the sparse representation vector is sent back to user's client and the original respiration cycles are reconstructed from simple multiplication of representation vector and the learned sparse dictionary.

In this section, we first introduce respiration circles related data extraction and encryption performed in smartphone, which includes targets' feature points tracking, random sampling and encryption. Then communication between smartphone and cloud/server is mentioned in 3.2 that contains efficient historical data reconstruction and final respiration rate computation.

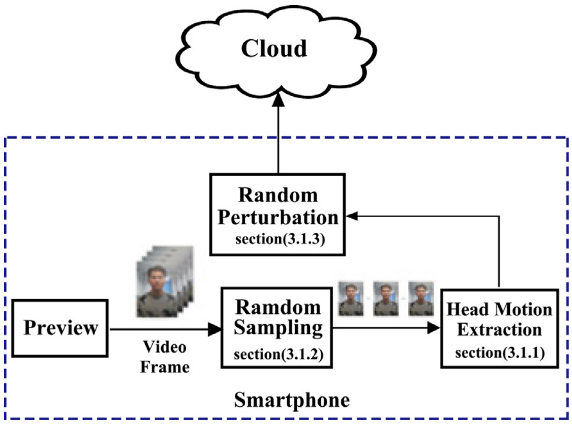


Fig. 2. Respiration cycles extraction and encryption on smartphones.

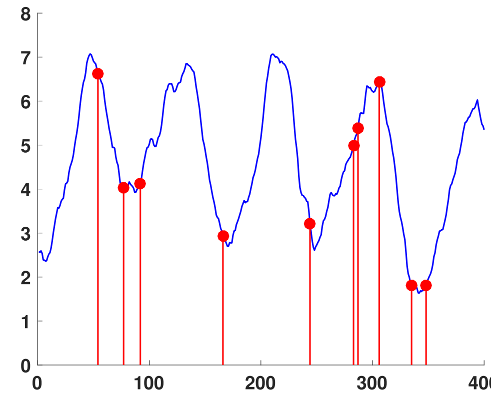


Fig. 3. Feature centroid points tracking and random sampling. (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)

3.1. Respiration cycles extraction and encryption

Fig. 2 presents the steps of extracting and encrypting the data related to respiration cycles. There is medical evidence showing that human's breath and heartbeats can result in subtle movements of their heads [1]. In HealCam, the data related to respiration cycles is extracted from the head motions of the target from a video stream perceived from Live Preview (without saving the video to the internal memory) on smartphones based on the approach proposed by Balakrishnan et al. [3]. To track head motions, instead of capturing the whole face area in the video stream, HealCam extracts feature points of the target and tracks the trajectory of their centroid. Then the centroid motion data is stored as the extraction of the respiration cycles related data. However compared with the existing work, our approach is significantly more energy and computation efficient with non-uniform random sampling inspired by compressive sensing [18].

3.1.1. Features extraction and tracking

Before starting the system, a smartphone should be fixed by a phone holder to avoid irrelevant motions of the hardware. Then the Preview functionality is triggered to obtain a video stream of the target's face and the preview frequency is set as 30 FPS (Frames Per Second). To extract the head motions from of the target, we follow similar approach in [3]: we first apply the Viola and Jones face detector [19,20]to find the region of the face so that to locate the tracking area [21]. The area around mouth and eyes is removed to eliminate the irrelevant motions when the target is blinking or speaking. The location of face in the first face image is set as the reference. Then we adopt Lucas Kanade tracker available in OpenCV [19] to track the relative positions of the feature points in different frames to the reference. The centroids of the feature points in each frame are computed and their displacement trajectory in vertical direction are recorded as the representation of head motions. A demonstration of the trajectory in full frame rate (30 FPS) is shown in Fig. 3 and the corresponding data vector $x = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$ in full frame rate is regarded as the original data vector of the respiration cycles which can be used to infer the respiration rate of the target.

3.1.2. Random sampling

It is known that video processing is computationally intensive and will drain the battery of the smartphone quickly. To save video processing energy, uniformly sampling the video frames is often applied and only the remaining frames will be processed. However, the minimal frame rate of uniform approach is bounded by the theory of Nyquist frequency and noises. We propose non-uniform random sampling inspired by the emerging theory of compressive sensing [18] to further reduce the frame rate. As shown in Fig. 3, different from the uniform approach, non-uniform random sampling picks a small subset of the video frames and stores their corresponding centroids(red points in the figure) in a non-uniform and random sequence which can be expressed as,

$$y = \Phi x \tag{1}$$

where $y \in \mathbb{R}^m$ consists of the relative head motions of the m chosen frames to the reference frame and $m \ll n$. The random selection matrix $\Phi \in \mathbb{R}^{m\times n}$ is a sparse matrix with 1s or 0s as its elements and it is generated locally on smartphones. It is sparse because most of its elements are 0s and only one entry in each row is non-zero (i.e., 1). This sparse matrix is widely used in efficient data collection based on compressive sensing [22–24]. It is worth noting that the features extraction and tracking are processed on the sampled frames directly (i.e., after random sampling) therefore only y will be obtained locally. The original data vector x will be only recovered in enquiry stage which will be detailed in Section 3.2.1.

In uniform sampling approach, an acceptable approximation of the original respiration cycles can be achieved using linear or spline interpolation as the evaluations in Section 4.2. However, in random sampling approach, the original respiration cycles cannot be recovered without the knowledge of the sparse matrix Φ which is stored on user's smartphone. Therefore, random sampling is able to both reduce computational cost and protect the data privacy.

3.1.3. Random perturbation

Though random sampling protects the original data, it is still vulnerable to sophisticated attacks with sufficient observations when fixed sparse matrix is used [25]. Random perturbation is a standard encryption method in the area of data security, whose common methods are adding random noise and matrix to raw data. HealCam exploits the random matrix method to encrypt the data. To further boost the security, we adopt the random perturbation approach proposed in [26] to encrypt the extracted head motions data vector from random sampling,

$$z_i = p_i y_i \tag{2}$$

where $p_i \in \mathbb{R}^{m\times m}$ is a perturbation matrix generated from a *seed* which is simple integer selected and stored locally on smartphones. z_i is the cypher text of y_i in the segment *i*. The *seed* can be used to regenerate the same perturbation matrix in enquiry stage. In the real implementation, data is processed for each 30-s segment therefore the perturbation matrix changes in each 30 s. With the random perturbation, the security of HealCam is improved; the cypher texts can be totally different even the sampled data vectors

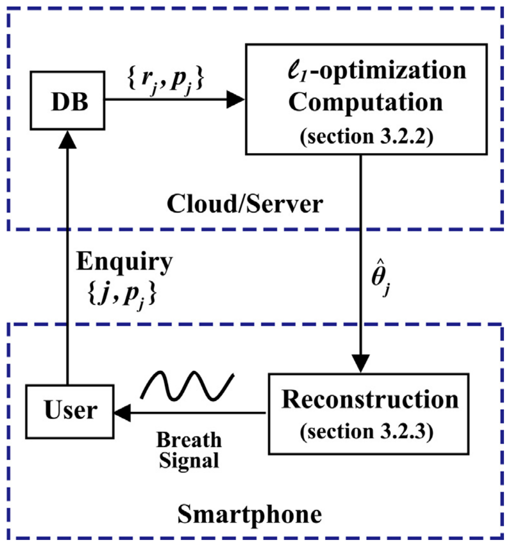


Fig. 4. Historical record enquiry and reconstruction.

are the same. Meanwhile, the perturbation matrix is a square matrix and is invertible which indicates the accuracy can be exactly preserved.

At last, the cypher texts $\{z_1, z_2, z_3, \ldots, z_i, \ldots, \}$ are transferred to cloud and stored for the future enquiry. So far, the privacy of the original respiration cycles has been well protected. The cloud cannot recover or infer the original data without the sparse matrix and the perturbation matrix. Meanwhile, the protection is quite lightweight, it only involves simple matrix multiplication.

3.2. Privacy-preserved historical record enquiry

As only cypher texts are stored in the cloud, data reconstruction needs to be done when users send a enquiry for historical records. We propose a new user enquiry scheme for HealCam in Fig. 4. To achieve energy-efficiency, we split the data reconstruction into two steps: the cloud is responsible for computing sparse representation vector via computationally intensive ℓ_1 optimization; user's smartphone reconstructs the original respiration cycles locally by computing the lightweight matrices multiplication. HealCam facilitates the user-specific dictionary learned from their own training data (collected before using the system) on a trusted third part (e.g., their own P.C.) to protect user's privacy. The detailed description of this stage is as below.

3.2.1. User enquiry

When a mobile user is curious about his historical record, such as the jth data segment, he first regenerates the perturbation matrix p_j with the corresponding seed. Then the recovery matrix $r_j \in \mathbb{R}^{m\times n}$ is formulated as,

$$r_j = p_j \Phi D \tag{3}$$

where $D \in \mathbb{R}^{n\times n}$ is the user-specific sparse dictionary learned from his own data collected before setting up the whole system. Then the recovery matrix r_j and the corresponding segment index j are sent to the cloud to trigger the enquiry about the historical record stored on the cloud.

3.2.2. ℓ_1 Optimization for sparse representation

When the cloud receives the recovery matrix r_j and the index j, it first finds the cypher text z_j according to the index and computes the sparse representation vector of the cypher text by solving the computationally intensive ℓ_1 optimization problem,

$$\hat{\theta}_j = \arg\min ||\theta_j||_1 \qquad s.t. \quad z_j = r_j\theta_j \tag{4}$$

where $\hat{\theta}_j \in \mathbb{R}^n$ is called a sparse representation vector.

The sparse representation vector is then returned to the mobile user for final reconstruction. As the dictionary D is user-specific, the cloud cannot reconstruct or approximate the original respiration cycles without the knowledge of the dictionary.

3.2.3. Reconstruction

The final reconstruction is completed on user's smartphone by simple matrix multiplication,

$$\hat{x}_j = D\hat{\theta}_j \tag{5}$$

where the dictionary D is user-specific and obtained from dictionary learning.

3.2.4. Peak detection and respiration rate computation

To determine the respiration rate in the enquired time segment, a butterworth bandpass filter is firstly applied to remove the irrelevant noises, and the bandwidth we set in HealCam is 0.1 Hz–0.5 Hz, which can cover the scope of respiration rate for adults.

The respiration signals are regular for most of the users and there are some harmonic peaks existing in the signals. To detect the peaks related to respiration, we perform peak detection on the filtered signal. A sliding window is used in peak detection and we label each sample in the signal as a peak if it is the largest value in a window centered at the sample, the frame number series $T = \{t_1, t_2, \ldots, t_i, \ldots\}$of the detected peaks is also generated by this step. Then the mean value of the duration between each peaks Δt is formulated as,

$$\Delta t = mean(diff(T)) \tag{6}$$

where *diff*(T) is a series containing the differentials between each element in T. Apparently, the mean time interval between peaks is $\frac{\Delta t}{f_{reconstruction}}$, and we can derive the respiration rate of the user by

$$RespirationRate = \frac{60 f_{reconstruction}}{\Delta t} \tag{7}$$

where $f_{reconstruction}$ is the frequency(30 Hz in this paper) of the reconstructed signal.

Learning Sparse Dictionary. Given a set of N data vectors $\{x_1, x_2, \ldots, x_N\} \in \mathbb{R}^n$, sparse dictionary learning aims to find a dictionary $D \in \mathcal{R}^{n\times n}$ in which the data vectors can be sparsely represented [27]. Before setting up the whole system, we collected 20 min video recording from each user and divided the video into 30-s segments; each segment corresponds to one data vector for dictionary learning. The data vectors are extracted from the video segments in full frame rate (i.e., 30 FPS). The user-specific sparse dictionaries are learned for each user and they are different from each other. Dictionary learning is computationally intensive but it is once-off and can be computed offline. There are different algorithms for learning sparse dictionary. In this paper, we choose SPAMS proposed in [27] which achieved higher recognition accuracy for GPS trajectory compression in [28] compared with the state-of-the-art methods.

4. Dataset evaluations

The aim of this section is to evaluate the performance on reconstruction accuracy of our proposed sampling approach, i.e., nonuniform Random Sampling using Dictionary Learning (termed as RSDL) on reconstruction. We first determine whether the respiration cycles can be sparsely represented with the learned dictionary. We then evaluate the reconstruction accuracy of RSDL and compare it with the other two methods, RSDCT and uniform Interpolation. RSDCT applies random sampling but uses standard DCT

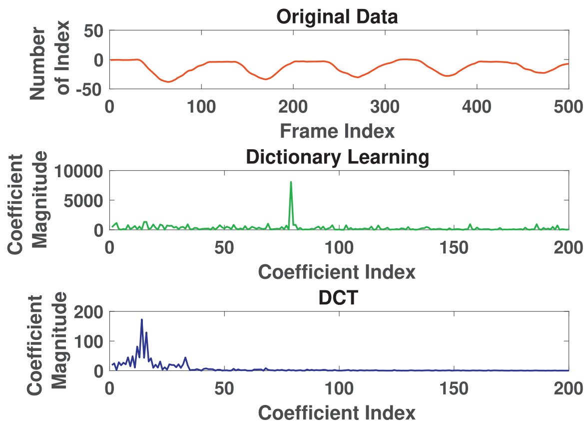


Fig. 5. Sparse coefficients in different domains.

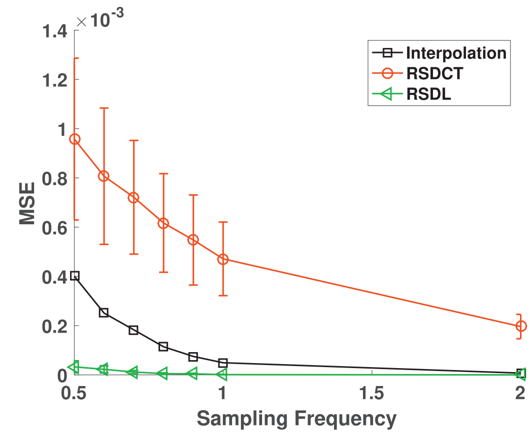


Fig. 6. Reconstruction accuracy of RSDL, RSDCT and Interpolation approaches.

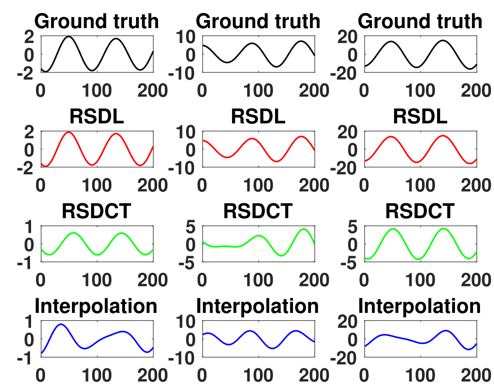


Fig. 7. Samples of reconstruction results.

basis for reconstruction. Uniform interpolation is also included as benchmark.

We recruit 32 subjects (16 males and 16 females)in good health condition by providing some incentive to collect two datasets for evaluation. The data collection consists of two sessions (one dataset for each session). In both sessions, each subject is seated regularly and breath normally. And a 20 min video recording in full frame rate (30 FPS) is collected for everyone. The first dataset (from the first session) is used for dictionary learning while the second dataset (from the second session) is used for accuracy evaluation.

4.1. Sparsity evaluation

The video recordings are first divided into 30-s segments. We then extract a vector of trajectory of the head motions for each segment, i.e., the original respiration cycles in full frame rate. We use the extracted data vectors from the first dataset for training the dictionary and then compute the sparse representations of the data vectors from the second dataset. We find that the learned dictionary has a better performance than the standard DCT basis on sparse representation, i.e., its dominant coefficients are more concentrated. We present an example of the sparse representation coefficients in Fig. 5. The figure on the top is an example of the original data vector. The corresponding sparse coefficients in learned dictionary and DCT domain are shown in the middle and the bottom figures respectively. The results show that sparse representation coefficients are more concentrated with the learned dictionary which indicates better reconstruction accuracy.

4.2. Reconstruction accuracy

In this section, we compare the reconstruction accuracy of three different methods, i.e., RSDL, RSDCT and spline interpolation. The first two methods utilize random sampling while the spline interpolation works on the uniform sampling. We gradually change the average sampling frequency from 0.5 Hz to 2 Hz and compute the corresponding Mean Squared Errors (MSEs) between the original and reconstructed data vectors. The results in Fig. 6 show that our proposed method RSDL achieves significantly lower MSE compared with the other two methods and the performance gain diminishes when the sampling frequency is over 1 Hz. Considering the computational issue, we set the average sampling rate as 1 Hz in our system implementation. We conclude this section by showing a number of samples of reconstructed data vectors from the three different approaches when sampling rate is 1 Hz. In Fig. 7, we present the original and reconstructed respiration cycles from three different subjects in different columns. It can be seen that RSDL achieves very close approximation to the ground truth and it outperforms the other two approaches significantly. RSDL reconstruct randomly sampled data, while uniform sampled data is interpolated to reconstruct. The results of the three examples intuitively reflect that with the same sampling rate, randomly sampled data performs better than the data collected in uniform way.

At last, it is worth noting that spline interpolation and RSDCT cannot preserve the data privacy: it is obvious that uniform sampling does not involve any protection scheme; RSDCT utilizes standard DCT basis which is not user-specific, the cloud is able to reconstruct the original data vector after obtaining the sparse coefficients by using standard DCT basis.

4.3. Respiration rate accuracy

The main application of HealCam is to monitor vital signs of users, in this section, we evaluate the respiration rate accuracy by comparing the respiration rate generated from the 30 Hz video to that from HealCam.There is nor difference comparing results generated from the 30 Hz video to that counted by the subjects when recording video, that we consider respiration rate generated from 30 Hz video as the ground truth data. A sliding window is used to detect peaks related to respiration, and in our experiment, we set 45 as the length of the window. We show the number of peaks detected by ground truth and HealCam as well as the respiration rate generated from the detected peaks in Table 1. Error values are donated by the percentage of the differential value between HealCam and ground truth. The respiration rate results of the whole 32 subjects we recruit shown in Table 1 show that HealCam can achieve significantly high accuracy in monitoring respiration rate, the accu-

Table 1
Respiration rate and numbers of peaks detected from 30Hz video and by HealCam.

	Respiration rate (per minute)		Number of peaks	
Sub.	Ground truth	HealCam (% error)	Ground truth	HealCam (% error)
1	23.8	23.8(0)	12	12(0)
2	27.0	26.8(0.7)	14	14(0)
3	21.2	21.1(0.4)	10	10(0)
4	20.5	20.5(0)	10	10(0)
5	22.8	22.8 (0)	12	12(0)
6	22.2	22.2(0)	12	12 (0)
7	20.7	20.7(0)	10	10(0)
8	21.5	21.5(0)	11	11(0)
9	19.9	19.9(0)	10	10 (0)
10	18.2	18.1(0.5)	9	9(0)
11	26.4	26.4(0)	14	14 (0)
12	30.0	30.0(0)	15	15(0)
13	20.9	23.3(11.5)	10	11(10.0)
14	28.9	28.9 (0)	14	14(0)
15	23.3	23.3(0)	11	11 (0)
16	26.4	24.0(9.1)	12	11(8.3)
17	21.3	21.3(0)	11	11(0)
18	23.5	21.4 (8.9)	12	11 (8.3)
19	23.4	23.4(0)	11	11(0)
20	21.1	21.1(0)	11	11 (0)
21	19.4	19.4(0)	9	9 (0)
22	18.8	18.8(0)	10	10 (0)
23	23.1	23.1(0)	12	12(0)
24	21.7	21.7(0)	11	11(0)
25	23.7	23.7(0)	12	12(0)
26	25.1	25.1 (0)	13	13(0)
27	23.5	23.5 (0)	12	12(0)
28	24.2	24.2(0)	13	13(0)
29	30.4	30.4(0)	16	16(0)
30	27.1	27.1(0)	14	14(0)
31	22.0	22.0(0)	10	10 (0)
32	24.2	21.5(11.2)	10	9 (10.0)

racy of the monitored respiration rate of HealCam is 98.7%, which represents the high practical value of HealCam.

4.4. System evaluation

To evaluate the system cost of HealCam, we implement HealCam on smartphones and high performance server. We also implement the original head motion detection approach proposed in [3] and a recent proposed encryption method, Talos [17], as benchmarks to demonstrate the improvement of HealCam on energy and computational efficiency. Talos is proposed as a efficient version extended from CryptDB [16] however it still consumes over a hundred times more energy and computation than compressive sensing based encryption according to our evaluation.

In our implementation, the smartphone used for implementing the user client is Samsung Galaxy Note 4 which runs Android OS V4.4. It has 2.7GHz processor and 3GB Ram onboard and its battery capacity is 3200 mAh. Then we use a high performance PC to host the server side of HealCam. It is an Apple Mac Pro-running macOS Sierra. It has a 3GHz 8-core Intel Xeon E5 cpu, 32GB 1866 MHz Ram and 3GB AMD Firepro GPU. Wi-Fi is used for data transmission between the smartphones and the server via Wi-Fi router.

We first evaluate the improvement of energy efficiency on video processing. As the server is connected to power source and has relatively high computational resources, we only concern the resource consumptions of user client on smartphones. We use the Trepn Power Profiler [29] (Fig. 8) to measure the energy consumption of the user client and compute the average and standard deviation of the results from 30 independent trials. We set the average frame rate as 1FPS, i.e., we sample 30 frames for each 30-s segment. We divide the energy consumption into three different components (transmission consumption is negligible as our system only uploads little data, i.e., 30 data points for each 30-s segment): Preview, Face Detection and Motion Extraction. As the results shown in Fig. 9(a), HealCam saves significant energy on video processing including face detection, feature extraction and tracking. It reduces the energy consumption in face detection and motion extraction by 27.6 times and 30.9 times respectively and overall it achieves 23.6 times better performance on energy efficiency compared with the original head motion extraction method with full frame rate (30 FPS).

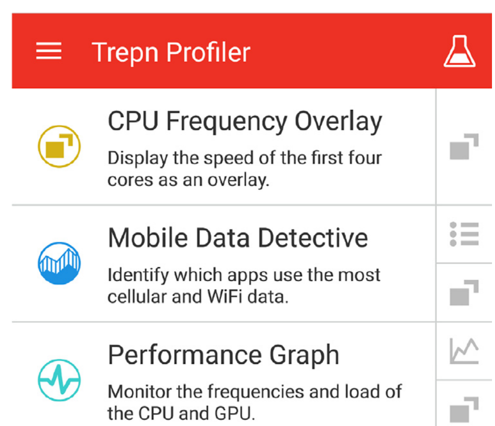
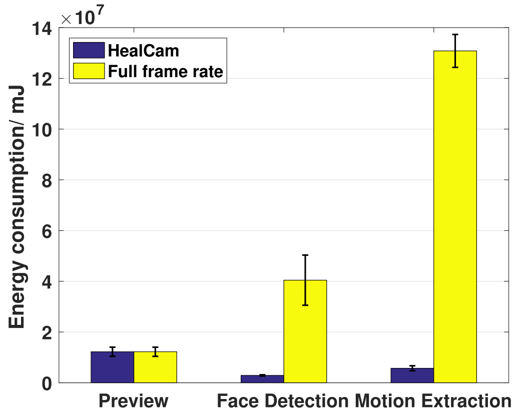
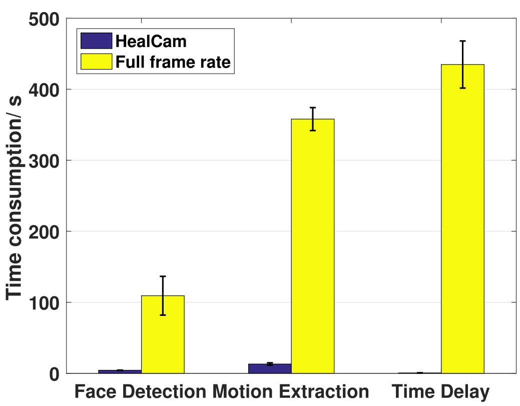


Fig. 8. Trepn Power Pofiler.

We then evaluate the improvement of computational efficiency in video processing. The computation time is obtained from the console of Android studio. Fig. 9(b) demonstrates the computation time of the face detection and motion extraction for each 30-s segment. Again, the average and standard deviation are obtained from 30 independent trials. We also evaluate the time delay of the



(a) Energy Consumption



(b) Processing Time

Fig. 9. Resouce consumption on smartphone.

system on smartphones. The system is implemented in multiple threads and the video frames are processed in sequential manner to reduce the system delay, i.e., face detection and motion extraction will be triggered once a new sampled frame is taken. From the results we can find HealCam is 24.9 times and 27.2 times faster than original approach in face detection and motion extraction respectively and the overall improvement is 26.67 times. The system delay of HealCam on smartphones is within 0.58 s which indicates no cumulative delay. However the original approach keeps processing over 434 s after the 30 s preview. As the system runs on the 30-s segments, the original approach brings overwhelming cumulative system delay for continuous monitoring scenario and is not applicable on smart devices.

As discussed before, our proposed encryption and decryption method is lightweight. It only involves very simple matrix multiplication. We compare it with the latest encryption method Talos in energy and computation efficiency.The encryption part of Talos is implemented in smartphones, and we conduct experiments with our dataset on both Talos and HealCam to encrypt data and measure energy cost. The results show that the compressive sensing based encryption method is 172 times faster and 99 times more energy-efficient than Talos.

5. Conclusion

In this paper, we propose HealCam, an energy-efficient and privacy-preserved respiration cycles monitoring system on camera-enabled smart devices. The overall system design is based on the related theories of compressive sensing. First we apply non-uniform random sampling instead of the traditional uniform sampling to reduce the video processing cost while achieving significantly better reconstruction accuracy. Then by facilitating lightweight compressive sensing based encryption and decryption strategy, the mobile users are able to encrypt and decrypt their data of respiration cycles in highly efficient manner and meanwhile protect their data privacy against the attacks from the cloud.

Acknowledgment

This work is partially supported by National Natural Science Foundation of China under Grant 61702133, Natural Science Foundation of Heilongjiang Province under grant QC2017069, the Fundamental Research Funds for the Central Universities Grant HEUCFJ160601, the China Postdoctoral Science Foundation No.166875 and Heilongjiang Postdoctoral Fund No. LBH-Z16042.

References

[1] C. Takano, Y. Ohta, Heart rate measurement based on a time-lapse image, Med.Eng. Phys. 29 (8) (2007) 853–857.
[2] M.-Z. Poh, D.J. McDuff, R.W. Picard, Advancements in noncontact, multiparameter physiological measurements using a webcam, IEEE Trans. Biomed. Eng. 58 (1) (2011) 7–11.
[3] G. Balakrishnan, F. Durand, J. Guttag, Detecting pulse from head motions in video, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3430–3437.
[4] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, et al., A view of cloud computing, Commun. ACM 53 (4) (2010) 50–58.
[5] M. Macr, D. Mirarchi, C. Pagliaro, P. Russo, P. Vizza, Smart health project 2.0:integration and expectations of smart health topics, ACM Sigbioinform. Record 5 (1) (2015) 1–3.
[6] S. Ahmari, A. Ameri, S. Padmanabhan, K. Parameshwaran, K. Ragheb, M. Mozumdar, Biomesensi: a wearable multi-sensing platform for bio-medical applications, 2015, pp. 372–373.
[7] P. Kumar, Communication in personal healthcare : Issues using 802.15.4 for personal healthcare (2008).
[8] S.W. Lee, Y.J. Kim, G.S. Lee, B.O. Cho, A remote behavioral monitoring system for elders living alone, in: International Conference on Control, Automation and Systems, 2007, pp. 2725–2730.
[9] H. Ren, Q. Max, Bioeffects control in wireless biomedical sensor networks, in: Sensor Ad Hoc Communications Networks, Secon 06 IEEE Communications Society on, 2006, pp. 896–904.
[10] A. Prathosh, P. Praveena, L.K. Mestha, S. Bharadwaj, Estimation of respiratory pattern from video using selective ensemble aggregation, IEEE Trans. Signal Process. 65 (11) (2017) 2902–2916.
[11] S. Suzuki, T. Matsui, S. Gotoh, Y. Mori, B. Takase, M. Ishihara, Development of non-contact monitoring system of heart rate variability (hrv)-an approach of remote sensing for ubiquitous technology, in: International Conference on Ergonomics and Health Aspects of Work with Computers, Springer, 2009, pp. 195–203.
[12] M.Z. Poh, D.J. Mcduff, R.W. Picard, Non-contact, automated cardiac pulse measurements using video imaging and blind source separation., Opt. Express 18 (10) (2010) 10762–10774.
[13] D. Yin, Y. Shen, C. Liu, Attribute couplet attacks and privacy preservation in social networks, IEEE Access 5 (2017) 25295–25305.
[14] A. Tripathi, Data access and integrity with authentication in hybrid cloud (2013).
[15] D.S. Manoj Prabhakar Darsi K.Suresh Joseph, A new approach for providing the data security and secure data transfer in cloud computing, Int. J. Comput. Trends Technol. 4 (5) (2013).
[16] R.A. Popa, C.M.S. Redfield, N. Zeldovich, H. Balakrishnan, Cryptdb: protecting confidentiality with encrypted query processing, 2011, pp. 85–100.
[17] H. Shafagh, A. Hithnawi, A. Dröscher, S. Duquennoy, W. Hu, Talos: Encrypted query processing for the internet of things, in: Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems, ACM, 2015, pp. 197–210.
[18] R.G. Baraniuk, Compressive sensing [lecture notes], IEEE Signal Process Mag 24 (4) (2007) 118–121.
[19] G. Bradski, et al., The opencv library, Doctor Dobbs J. 25 (11) (2000) 120–126.
[20] W. Xu, Y. Shen, N. Bergmann, W. Hu, Sensor-assisted multi-view face recognition system on smart glass, IEEE Trans. Mob. Comput. 17 (1) (2018) 197–210.
[21] Y. Shen, M. Yang, B. Wei, C.T. Chou, W. Hu, Learn to recognise: exploring priors

of sparse face recognition on smartphones, IEEE Trans. Mob. Comput. 16 (6) (2017) 1705–1717.
[22] Y. Shen, W. Hu, R. Rana, C.T. Chou, Nonuniform compressive sensing for heterogeneous wireless sensor networks, IEEE Sens J 13 (6) (2013) 2120–2128.
[23] X. Wu, M. Liu, In-situ soil moisture sensing: measurement scheduling and estimation using compressive sensing, in: Proceedings of the 11th International Conference on Information Processing in Sensor Networks, ACM, 2012, pp. 1–12.
[24] Y. Wang, Z. Yang, J. Zhang, F. Li, H. Wen, Y. Shen, Cs2-collector: a new approach for data collection in wireless sensor networks based on two-dimensional compressive sensing, Sensors 16 (8) (2016) 1318.
[25] Z. Yang, W. Yan, Y. Xiang, On the security of compressed sensing-based signal cryptosystem, IEEE Trans Emerg Top Comput 3 (3) (2015) 363–371.
[26] H. Kargupta, S. Datta, Q. Wang, K. Sivakumar, On the privacy preserving properties of random data perturbation techniques, in: Data Mining, 2003. ICDM 2003. Third IEEE International Conference on, IEEE, 2003, pp. 99–106.
[27] J. Mairal, F. Bach, J. Ponce, G. Sapiro, Online dictionary learning for sparse coding, in: Proceedings of the 26th annual international conference on machine learning, ACM, 2009, pp. 689–696.
[28] R. Rana, M. Yang, T. Wark, C.T. Chou, W. Hu, Simpletrack: Adaptive trajectory compression with deterministic projection matrix for mobile sensor networks, IEEE Sens. J. 15 (1) (2015) 365–373.
[29] https://developer.qualcomm.com/software/trepn-power-profiler.

Qing Yang is a PhD student in the College of Computer Science and Technology, Harbin Engineering University(HEU). She received the bachelor degree of software engineering from HEU. She is a visiting PhD student in the Commonwealth Scientific and Industrial Organization(CSIRO) in Australia. Her main research interests are wearable/ mobile computing and privacy preserving in mobile sensor networks.

Yiran Shen received the PhD degree in computer science and engineering from the University of New South Wales. He is an associate professor in the College of Computer Science and Technology, Harbin Engineering University (HEU). He was SMART Scholar at Singapore-MIT Alliance for Research and Technology before he joined HEU. He publishes regularly at top-tier conferences and journals. His current research interests are wearable/ mobile computing, wireless sensor networks and applications of compressive sensing. He is a member of the IEEE.

Fengyuan Yang is a Graduate student in the College of Computer Science and Technology, Harbin Engineering University (HEU). His current research interests are wearable/ mobile computing.

Jianpei Zhang is the Professor at College of Computer Science and Technology, Harbin Engineering University. He is also the leader of Software and Social Computing Research Group. He received his PhD from Harbin Engineering University. His research interests include data mining, database systems and software engineering.

Wanli Xue is a PhD candidate in Department of Computer Science and Engineer of the University of New South Wales. He is also a member of Networks Research Group of Data61 CSIRO, which used to be NICTA. He received the Master degree from the University of Tasmania. His current research interests as well as PhD topic are about Cyber-Physical Systems and private IoT.

Hongkai Wen is an Assistant Professor in Department of Computer Science, University of Warwick. Before that he obtained his D.Phil at the University of Oxford, and became a post-doctoral researcher in a joint project between Oxford Computer Science and Robotics Institute. Broadly speaking, his research belongs to the area of Cyber-Physical Systems, which use networked smart devices to sense and interactive with the physical world.