様　式　C－１９、F－１９－１、Z－１９　（共通）

# 科学研究費助成事業　　研究成果報告書

平成 ３０ 年　 ６ 月 ２０ 日現在

機関番号：６２６１５
研究種目：挑戦的萌芽研究
研究期間：2015～2017
課題番号：１５Ｋ１２０７２

研究課題名（和文）Signal Processing for Non-intrusive Sleep Monitoring

研究課題名（英文）Signal Processing for Non-intrusive Sleep Monitoring

研究代表者
ＣＨＥＵＮＧ　ＧＥＮＥ（CHEUNG, GENE）

国立情報学研究所・コンテンツ科学研究系・准教授

研究者番号：４０５７７４６７

交付決定額（研究期間全体）：（直接経費）　　2,800,000円

研究成果の概要（和文）：研究目的は、ビデオとオーディオの録音と分析によって患者の睡眠の質を非侵襲的に監視することです。日本の高齢者の大部分に影響を与える無呼吸検出に焦点を当てています。記録されたビデオは、Microsoft Kinectカメラで撮影された一連の奥行き画像で、アクティブなテクノロジを使用しているため、画像は周囲の照明条件の影響を受けません。ビデオは、ノイズがあり、ビット深度が低く、グラフ信号復元技術を使用してノイズ除去およびビット深度強化が必要です。教師付き学習のために、ビデオおよびオーディオからの特徴を抽出して、分類器を構築する。設計された分類器は、様々な無呼吸の種類を高い精度で検出することができる。

研究成果の概要（英文）：The goal of this research is to monitor the sleep quality of a patient non-intrusively via video and audio recording and analysis. In particular, we focus on apnea detection, a common and serious sleep condition that affects a large percentage of the Japanese older population. The recorded video is a sequence of depth images captured by a Microsoft Kinect camera, which utilizes active sensing technologies, so that captured images are not affected by ambient lighting conditions. The captured video is noise-corrupted and of low bit-depth, and requires denoising and bit-depth enhancement, performed using graph-signal restoration techniques. Features from video and audio are then extracted for supervised learning to construct a classifier. The designed classifier can then detect different apnea types with high accuracy, and is robust to the patient's sleep pose. The prototype has been deployed in an Australian sleep clinic and has demonstrated its effectiveness.

研究分野：信号処理

キーワード：画像処理　グラフ信号処理

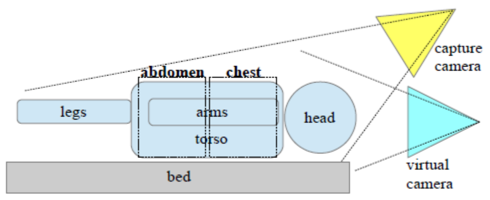


**Fig. 1**. Side view of sleep patient. Torso is divided into two cross sections, each modeled by an ellipse.

## １．研究開始当初の背景

It is well understood that quantity and quality of sleep could significantly affect work productivity. In particular, obstructive sleep apnoea, characterized by repetitive obstruction in the upper airway during sleep, is common in the general population and can have significant negative effect on a person's sleep quality, and hence quality of life and cognitive functions. The condition is diagnosed via attended (in-laboratory) or unattended (ambulatory) diagnostic sleep studies. We address the problem of identifying the obstructive respiratory events in this research.

Existing in-laboratory monitoring devices are cumbersome to use, expensive, and intrusive with multiple body straps and tubes that affect a patient's sleep quality during monitoring. On the other hand, less intrusive sleep monitoring units such as vibration-sensing wristbands (e.g., Fitbit1 and Jawbone UP2) mostly record sleep time, i.e., the quantity rather than the quality of sleep, and are not equipped to detect respiratory events of different kinds during the night.

## ２．研究の目的

Motivated by the shortcomings of in-laboratory monitoring devices and consumer-level sleep monitoring units, our goal is to accurately but non-intrusively detect respiratory events as manually scored by a scientific officer based on data collected by system Alice6 LDxS. Towards this goal, we propose a completely contact-less sleep monitoring system based on depth video and audio processing, suitable for home use. Not relying on the lighting condition of a dark sleeping room, we use a Microsoft (MS) Kinect sensor projecting infrared light patterns to capture depth images of the sleep patient.

## ３．研究の方法

We overview our proposed sleep monitoring system that employs an MS Kinect sensor to capture depth video and audio of a sleep patient. A potential usage of our system is as follows. When a patient stays overnight in a sleep clinic for initial testing, in addition to in hospital system's sensors, we deploy also a Kinect sensor to capture depth video and audio for respiratory

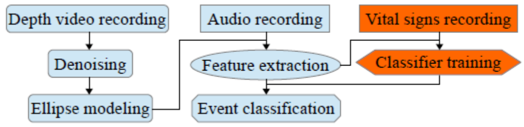


**Fig. 2**. System overview. 'Vital Signs Recording' is for ground truth; Orange: initial training components; Blue: regular usage components.

event classifier training. In subsequent nights at the patient's home, our proposed system that replicates the same Kinect sensor setup is activated to collect depth video and audio data non-intrusively for respiratory event classification. Without the body-attached sensors, this would mean a significant improvement in sleep comfort for the patient when at home.

Specifically, we employ a first-generation MS Kinect depth camera for depth video and audio processing and respiratory event classification. As shown in Fig. 1, the camera is set up at a higher elevation above and away from the head of the patient lying down. This camera location gives an unobstructed view of the patient's torso for depth video capture and analysis. The Kinect camera captures depth images of resolution 640x480 pixels with 11-bit pixel precision at 30 frames per second. The camera can also simultaneously capture audio at 16kHz, 16-bit sample precision with a PCM S16 LE audio codec. Note that though Kinect camera has a 4-microphone array resulting in a 4-channel audio, we use only the first channel for recording.

The main contributions of our research are listed as follows:

1) To facilitate transmission of the acquired data from the patient's home to a sleep clinic for storage and analysis, we propose an alternating-frame video recording scheme, so that different 8 of the 11 bits in captured depth images are extracted at different instants for efficient encoding using H.264 Advanced Video Coding (AVC) video codec. At the decoder, the uncoded 3 bits in each frame can be recovered accurately via block-based search.

2) We perform temporal denoising using a motion vector smoothness prior expressed in the graph-signal domain, so that unwanted flickering can be removed without blurring sharp edges in the depth images.

3) Given the denoised depth video, we track the patient's chest and abdominal movements over time based on our proposed dual-ellipse model.

4) We extract ellipse model features via a wavelet packet transform (WPT), which are combined with audio features extracted via nonnegative matrix factorization (NMF) for a Support Vector Machine (SVM) and neural network (NN)

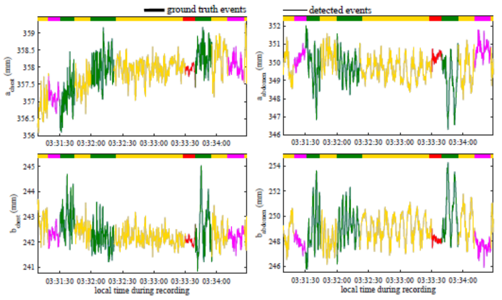


**Fig. 3**. Successfully detected events based on WPT depth video features showing chest_a; chest_b; abdomen_a and abdomen_b during the sideway sleep period. Red: central apnea; Magenta: obstructive and mixed apnea; Yellow: hypopnea; Green: other events.

classifiers to detect respiratory events.

４．研究成果

Using respiratory events scored by a scientific officer (who is blind to our study) based on data collected by system Alice6 LDxS as ground truth, and captured depth video and audio of patients collected at Concord Private Hospital, Australia, we conducted extensive experiments to test our system. First, we show that our depth video compression scheme outperforms a competitor that records only the eight most significant bits in peak signal-to-noise ratio (PSNR). Second, we show that graph-based temporal denoising scheme reduces the flickering effect without over-smoothing. Third, we show that our system can deduce respiratory events of different kinds as scored manually by a scientific officer based on data collected by system Alice6 LdxS with high accuracy.

５．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計　２　件）

(1) Cheng Yang, Gene Cheung, Vladimir Stankovic, “**Estimating Heart Rate and Rhythm via 3D Motion Tracking in Depth Video**,” *IEEE Transactions on Multimedia*, vol. 19, no.7, pp.1625-1636, July 2017. (査読あり)

(2) Cheng Yang, Gene Cheung, Vladimir Stankovic, Kevin Chan, Nobutaka Ono, “**Sleep Apnea Detection via Depth Video & Audio Feature Learning**,” *IEEE Transactions on Multimedia*, vol. 19, no.4, pp.822-835, April 2017. (査読あり)

〔学会発表〕（計　３　件）

(1) Cheng Yang, Gene Cheung, Vladimir Stankovic, “**Estimating Heart Rate via Depth Video Motion Tracking**,” *IEEE International Conference on Multimedia and Expo*, Torino, Italy, June, 2015. (査読あり)

(2) Cheng Yang, Yu Mao, Gene Cheung, Vladimir Stankovic, Kevin Chan, “**Graph-based Depth Video Denoising and Event Detection for Sleep Monitoring**,” *IEEE International Workshop on Multimedia Signal Processing*, Jakarta, Indonesia, September, 2014. (査読あり)

(3) Cheng Yang, Gene Cheung, Kevin Chan, Vladimir Stankovic, “**Sleep Monitoring via Depth Video Recording & Analysis**,” *5th IEEE International Workshop on Hot Topics in 3D* (Hot3D), Chengdu, China, July, 2014. (査読あり)

〔図書〕（計　０　件）

〔産業財産権〕

○出願状況（計　０　件）

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況（計　０　件）

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等
http://research.nii.ac.jp/~cheung/

６．研究組織

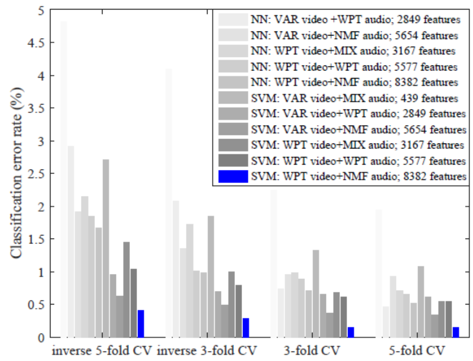


**Fig. 4**. Error rates of classification based on depth video+audio features.

(1)研究代表者
チョン　ジーン (CHEUNG, Gene)
国立情報学研究所、コンテンツ科学研究系、
准教授
　研究者番号：40577467

(2)研究分担者
小野 順貴 (ONO, Nobutaka)
首都大学東京、システムデザイン学部 情報
科学科
教授
　研究者番号： 80334259

(3)連携研究者
　　　　　　　　(　　　)

　研究者番号：

(4)研究協力者
　　　　　　　　(　　　)