IN-BED PATIENT MOTION AND POSE ANALYSIS USING DEPTH VIDEOS FOR PRESSURE ULCER PREVENTION

Ming-Ching Chang, Ting Yu†, Kun Duan‡ Jiajia Luo, Peter Tu*

GE Global Researh Center
Niskayuna, NY USA

Michael Priebe, Elena Wood, Max Stachura§

Charlie Norwood VA Medical Center
Augusta, GA USA

ABSTRACT

We present a real-time depth based computer vision system for pressure ulcer prevention, in-bed patient care and monitoring. Our system can effectively determine whether or not a mobility-compromised patient has been correctly repositioned at the required frequency. A depth sensor is used to detect and recognize patient movements, motion patterns, and pose positions. If the patient has stayed in an unchanged pose for too long and needs pressure releasing movements, our system can notify caregivers for repositioning or assistance. Privacy concerns are mitigated by removing the RGB components of the video stream from the camera capturing, and only processing depth measurements. We collaborated with clinical practitioners at the Charlie Norwood VA Medical Center for in-field data collection and experimental evaluation. A web portal front-end is developed such that all historical patient movements, pose positions, and repositioning data can be organized to support *telehealth* applications.

***Index Terms*—** motion analysis, pose analysis, depth video, pressure ulcer prevention, patient care.

1. INTRODUCTION

In the United States, pressure ulcers have become a cause of significant morbidity and mortality among hospitalized, institutionalized, and mobility-compromised individuals. Each year over 2.5 million people develop pressure ulcers, resulting in more than 60,000 deaths. The estimated hospital cost associated with patient stays and care with pressure ulcers totals more than \$11 billion annually [1], a significant financial burden to the health care system. To reduce the incidences of pressure ulcers, a variety of means of prevention and treatment have been established, including skin/wound assessment, use of support surfaces and risk assessment [2], periodic patient repositioning [3], nutritional status optimization, skin moisturization, infection assessment and prevention, and use of healing monitoring tools [4].

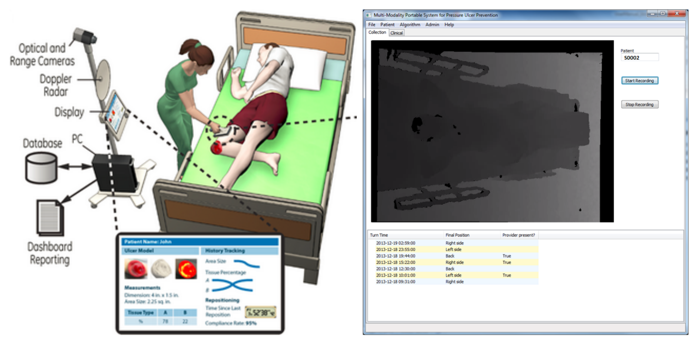


Fig. 1. **System Overview.** We developed a depth video system for pressure ulcer prevention and in-bed patient care. Typical depth view of a mobility-compromised patient is shown on the right. By monitoring the movements and poses of the patient continuously, 24/7, our system determines unchanged poses and can notifying caregivers for repositioning.

Guidelines recommend that a mobility-compromised patient needs to be repositioned for every 2 to 4 hours. However, under current clinical practices, the effectiveness of the repositioning protocol usually becomes diminished due to the lack of resources needed to monitor the occurrences of proper patient repositioning. Both prevention and rehabilitation of pressure ulcers require constant assessment of patient mobility and ulcer progression. This requirement, however, cannot be easily met, especially in remote areas where nursing shortages are common. In this pilot study, we focus on the development of a real-time system that can automatically monitor the movements and pose positions of mobility-compromised patients in 24/7 (see Fig.1). Our system determines if the patient has stayed in an unchanged pose for extended periods of time. In such cases, our system can notify caregivers for repositioning or assistance, so that pose changes or pressure release actions can be performed.

We collaborated with clinical practitioners at Augusta University (AU) and the Charlie Norwood VA Medical Center (CNVAMC) under the support of the Department of Veterans Affairs (VA). In-field data collection and evaluation was

*Ming-Ching Chang was with GE Global Research and is currently with the Computer Engineering Department, Univ. at Albany, SUNY, NY, USA.

†Ting Yu was with GE Global Research and is currently with Nest Labs - An Alphabet Company, Palo Alto, CA, USA.

‡Kun Duan was with GE Global Research and is currently with Snapchat Inc., Los Angeles, CA, USA.

§Elena Wood and Max E. Stachura are also with the Augusta University (AU) Center for Telehealth, Augusta, GA, USA.

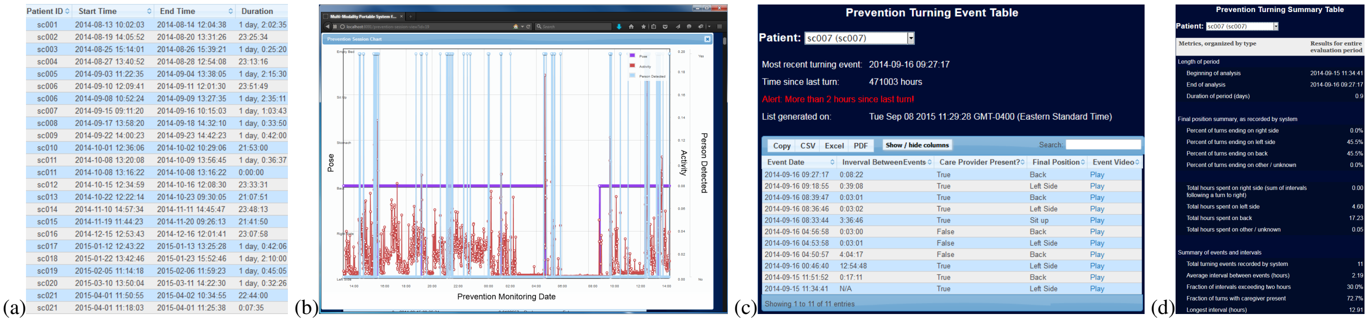


Fig. 2. **Web portal** for pressure ulcer prevention. (a) Patient data sessions. (b) Session graph of estimated patient motion activity (red), poses (purple), and in-bed person detection (cyan) over time. The shown subject has spent > 12 hours lying on the back prior to being turned to the left. (c) Event table for pose changes and repositioning needs. (d) Patient care summary.

performed at the Spinal Cord Injury Unit (SCIU) at the CN-VAMC. Both the practitioner's expertise and inputs as well as the access to the clinical notes regarding the subjects in our experiments were valuable in the development of this system.

The system was developed based on a 3D depth (or range) sensor positioned over the patient bed. We focus on two aspects of measurements in order to analyze and monitor the patient's repositioning needs. (1) **gross body motion estimation**, computed based on the aggregation of motion over the patient bed in a time-window as described in §2, and (2) **in-bed pose analysis** by extracting shape related features from the depth sequence over a rectified coordinate system, in order to recognize the lying or sitting poses as described in §3. Depth cameras enable night monitoring as they are independent of lighting conditions. Also, depth video streams are processed in real-time without requiring storage, thus mitigating privacy concerns. The system can identify instances when the patient has not undergone sufficient pressure releasing actions over prolonged periods of time. As such instances are identified, both the patient and the caregivers will receive real-time notifications. Our repositioning-based pressure ulcer prevention scheme is compact, portable, easy to use by experts and non-experts, and easily deployable for use in hospitals, clinics, and home environments. An on-line web portal front-end was also developed as in Fig.2. All historical patient movements, poses, and repositioning data can be stored and organized for *telehealth* applications.

Related Works: The use of depth cameras is advantageous and popular in human motion analysis, for a survey see [5, 6]. For in-bed patient care, Martinez *et al.* [7, 8, 9] construct a Bed Aligned Map (BAM) by rectifying the bed ROI of the depth frame into 10×10cm cells, and extract average cell height over the bed mattress to estimate bed occupancy, body localization, actions and sleeping positions.

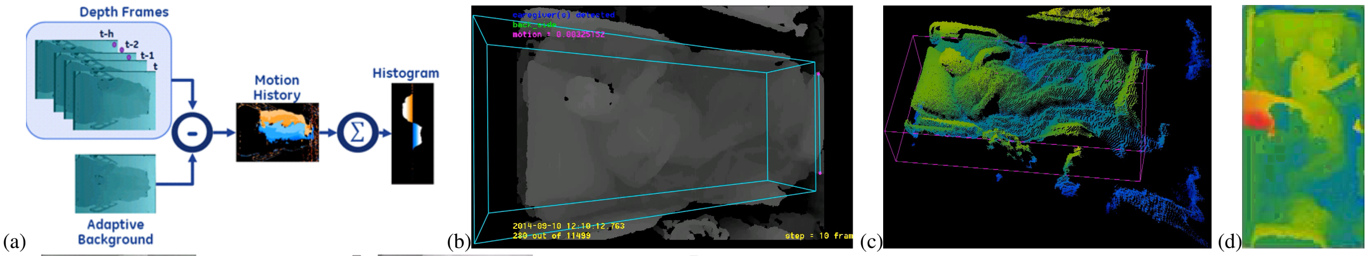
2. MOTION / ACTIVITY MONITORING

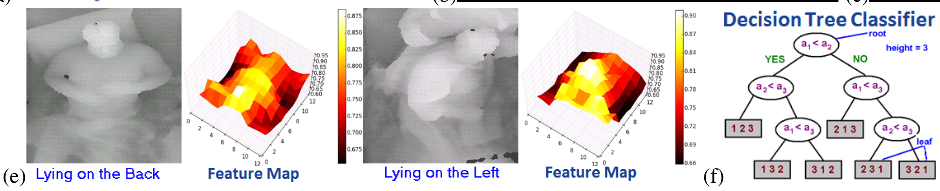
Our goal was to extract motion features from the depth video so as to characterize gross activities of the patient. Lack of activity over extended periods of time might indicate a need for repositioning. However, this simple rule alone can introduce false alerts, since lack of activity is not always caused by inability to move — rather it could be simply due to the absence of the patient, or the patient is sleeping. Our solution is to estimate the gross motion as a means to correlate with the patient's overall mobility, which can be combined with the pose analysis in §3 as a whole to determine the triggering of a repositioning event.

Our motion detection algorithm is based on three steps. *(i)* We adopt a standard **foreground/background** (FGBG) analysis [10] on the depth sequence, where each depth pixel in the view represents the distance from the sensor. Specifically, a per-pixel probabilistic model is built to separate stationary background pixels from pixels with significant changes. Background subtraction is performed followed by morphological filtering. *(ii)* Each depth image can also be regarded as a 3D point cloud, as shown in Fig.3(b,c). We apply a camera to ground-plane **coordinate transformation**, in order to estimate the *height* of each point sample from the ground. In our experiments the patient together with the bed are mostly within the height of $[0.8, 1.8]$ meters. We thus neglect any motion outside this range. *(iii)* To robustly deal with noise arising from small errors and frame variations, we aggregate motion signals over a time window following the **motion history image** (MHI) representation [11]. Specifically, MHI is a static image, where pixel intensity stores the recency of motion in a sequence, so that "motion history" is directly represented in a single image. Temporally distant frames in the past weigh less in the aggregated motion via a decay parameter. The gross *motion magnitude* m is computed on the fly from the bed ROI of the MHI to indicate the patient mobility. Figs. 4, 5, 6 illustrate experimental results.

3. IN-BED PATIENT POSE CLASSIFICATION

The objective of this section is to recognize the patient's lying or sitting pose positions on the bed based on a single depth image. This capability enables our system to monitor the patient pose continuously in real-time, in order to determine the repositioning needs. We focus on the classification of **6** patient pose positions as listed in Fig.3(g): patient lying on the (1) left side, or (2) right side, (3) lying on the back, (4) lying

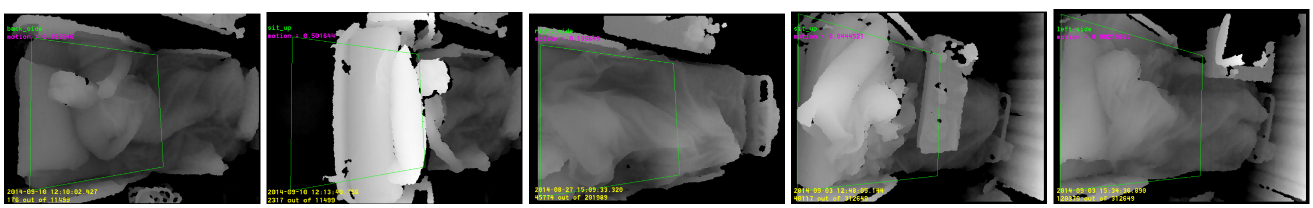




(g)

Lying on the left (LeftSide)
Lying on the right (RightSide)
Lying on the back (BackSide)
Lying on the stomach *i.e.* (FaceDown)
Sitting up (SideUp)
Empty bed (EmptyBed)

Fig. 3. **Patient pose estimation.** (a) The workflow consists of three main steps: **(I) depth view rectification** and adaptive background modeling as shown in (b,c), **(II) feature extraction** from rectified depth map in (d,e), and **(III) decision forest classification** in (f). (g) lists the six in-bed patient pose classes to be recognized.



(a) back side ($m = 0.159$) (b) sit up ($m = 0.501$) (c) right side ($m = 0.139$) (d) sit up ($m = 0.044$) (e) left side ($m = 0.002$)

Fig. 4. Per-frame motion estimation and pose classification examples, where the motion magnitude m is described in §2.

on the stomach, (5) sitting up, and (6) an empty bed. The explicit classification of the "empty bed" state vs. other states can be regarded as a **in-bed person detector**. Our system can thus learn to automatically filter out the time when the patient is removed from the bed (for bowel or other medical services), see Fig.2(b) for an example. Our pose classification workflow consists of three main steps as described below.

(I) Depth view rectification. To remove variations due to camera positioning, we first apply a rectification to convert all the depth pixels within the bed ROI to a 3D world coordinate system. Following the camera to ground-plane coordinate transformation described in §2, we define the world origin as the upper-left corner of the patient bed, the x- and y-axes to align with the short and long sides of bed respectively, the z-axis to point upwards, and the units to be in meters. We keep the points within the height of $h \in [0.8, 1.8]$ meters to obtain the 3D ROI cuboid of the patient bed, as illustrated in Fig.3(b,c). A rectified *height image* $\mathcal{H} = h(i, j)$ can then be created by re-sampling the point cloud as in Fig.3(d).

(II) Feature extraction. We extract two types of features from $\mathcal{H}$ and concatenate them to obtain a ~ 720 dimensional feature vector for pose classification. First, we divide $\mathcal{H}$ into grids (similar to the Bed Aligned Map in [7]) as in Fig.3(e) and extract the **median height** of each grid. To consider scale variability, we chose multiple grid sizes along ground plane x and y directions (of 10×10cm, of 20×10cm, and of 40×40cm). Within each grid, we calculate the median of the height value for all pixels inside the grid. All calculated values are concatenated to form a feature vector. Secondly, we calculate the **pairwise median height difference** between a selected set of grid pairs. Such pairwise features resemble the famous Local Binary Pattern (LBP) [12]. This turns out to provide great discriminative power for pose classification.

(III) Decision forest classification. A *multi-layer decision forest* [13] is trained to recognize the six in-bed lying pose configurations based on the extracted feature vectors. The training of the pose classifier requires annotated samples. To effectively obtain a large set of representative samples, we define "key frames" in the training video to represent a change of patient pose. It is reasonable to assume that the patient pose does not change between consecutive key frames. We manually groundtruthed all key frames in the training videos. Once the pose types between key frames are annotated, we take samples from the frames with known pose labels for training. We found that a few thousand labeled frames are sufficient to train the decision-forest based pose classifier. Fig.4 shows a few examples of the per-frame pose classification results.

Determining repositioning needs: Raw per-frame pose classification as shown in Fig.5(b) appears to be noisy. We apply temporal filtering to remove the spikes and fluctuations as a post-processing step. We further aggregated the classification results within one minute as a final robust estimator. Finally, the lack of pose change(s) over a pre-defined threshold of **4** hours triggers a repositioning event. Fig.6 shows

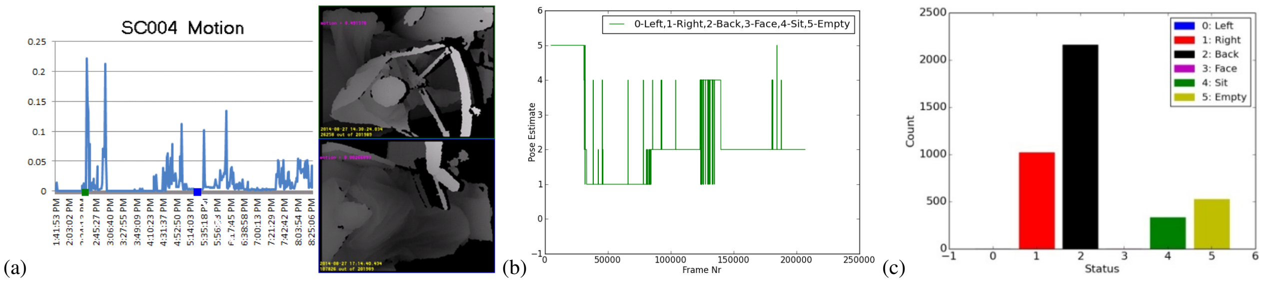


Fig. 5. **Case study** of patient SC004: (a) body motion estimation for about 7 hours of recording, (b) pose classification results (labels *vs.* frame numbers), and (c) pose statistics over time.

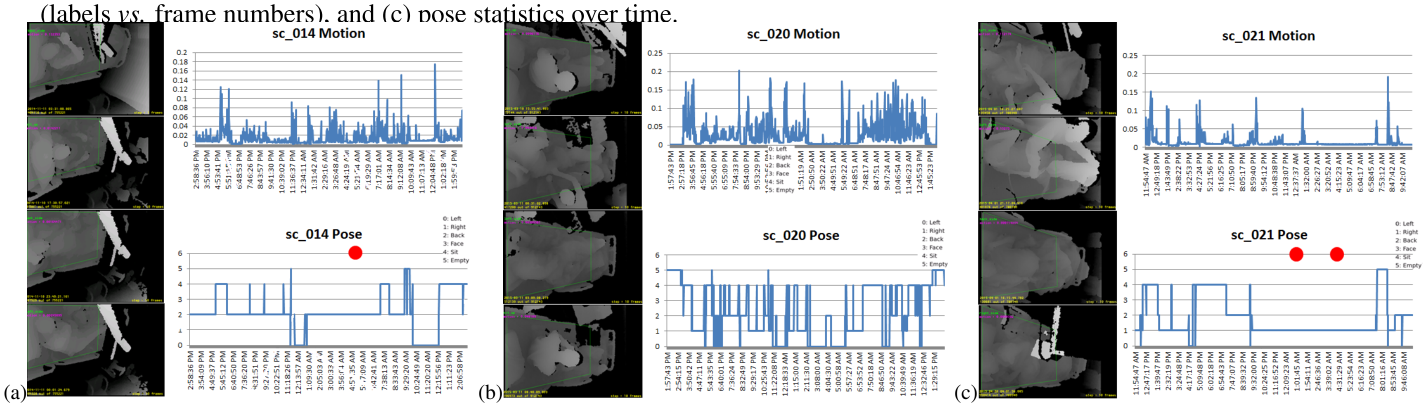


Fig. 6. **Case study** of motion and pose analyses for three subjects. Red dots represent triggered events of repositioning needs.

three such missing repositioning events in red dots.

4. RESULTS AND DISCUSSIONS

We used the Microsoft Kinect [1] to collect depth video from **21** subjects at the CNVAMC site, where each subject is monitored between 24 to 48 hours in compliance with IRB regularizations. The recorded data are divided roughly equally into a training set and a testing set, while subjects in the training set are excluded from the testing set. Our system is mounted on an IV like stand which can be readily positioned next to the patient bed. Fig.2(a) summarizes the collected dataset that is organized on the web portal.

We discuss results on selected subjects and compare our video analytic findings to these clinical notes. Fig.5 shows a case study of patient SC004, who was immobile and had developed pressure ulcers that were not healing well. Observe the classification statistics shown in Fig.5(c) that SC004 never lays on the left side. This was due to wound discomfort.

Fig.6 visualizes both the motion and pose analyses for three subjects. SC014 in Fig.6(a) is seriously immobile and constantly needs repositioning. It turns out that most detected motions are due to patient transit (from bed to bed) using a lifting gantry. In contrast, SC020 in Fig.6(b) is highly mobile and no repositioning is required. SC021 in Fig.6(c) has very little spontaneous motion, and most observed activities are from the caregiver. SC021 has been repositioned well during the daytime, however there were two missing repositionings overnight. Finally the sleep pattern of the subject is recognizable by measuring the amount of relatively inactive motions and unchanged pose positions.

[1]Note that the Kinect skeleton model does not work in our case even if the RGB signal is used, because in-bed patient poses can include heavy occlusions (under blankets) with complex limb configurations. For this reason, alternative forms of analysis must be developed.

5. CONCLUSIONS

We have presented a video analytic system as an aid to pressure ulcer prevention. Our system performs both patient motion monitoring and pose classification from depth video feeds. It can effectively analyze the patient's motion and lying pose positions, and notify caregivers if repositioning is required. In addition to pressure ulcer care, clinics can also make use of our mobility and pose analysis system to assess the patient's overall healing progress and behavioral characteristics such as the sleeping conditions. Our system is easily deployable for use in hospitals, clinics, and home environments. The integration of the video analytic system to a web portal can assist patient data organization and support for telehealth services.

Future works: State-of-art models such as Deep Neural Networks [14] could be applied for pose classification. Also, state transition estimators such as the Hidden Markov Model (HMM) can be applied to improve the repositioning decisions.

Acknowledgments: This work is partially supported by the U.S. Department of Veterans Affairs Innovation Initiative (VAi2) #302 - Multi-modality Portable Systems for Pressure Ulcer Prevention and Care.

6. REFERENCES

[1] Chandan K. Sen, Gayle M. Gordillo, Sashwati Roy, Robert Kirsner, Lynn Lambert, Thomas K. Hunt, Finnn Gottrup, Geoffrey C. Gurtner, and Michael T. Longaker, "Human skin wounds: a major and snowballing threat to public health and the economy," *Wound Repair and Regeneration*, vol. 17, no. 6, pp. 763–771, 2009.

[2] European Pressure Ulcer Advisory Panel and National Pressure Ulcer Advisory Panel, *Pressure ulcer prevention, quick reference guide*, 2009.

[3] M. Reddy, S. S. Gill, and P. A. Rochon, "Preventing pressure ulcers," *The Journal of the American Medical Association*, vol. 296, pp. 974–984, 2006.

[4] Dimitrios I. Kosmopoulos and Fotini L. Tzevelekou, "Automated pressure ulcer lesion diagnosis for telemedicine systems," *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, no. 5, pp. 18–22, 2007.

[5] Lulu Chen, Hong Wei, and James Ferryman, "A suevey of human motion analysis using depth imagery," *Pattern Recognition Letters*, vol. 34, no. 15, pp. 1995–2006, 2013.

[6] Jungong Han, Ling Shao, Dong Zu, and Jamie Shotton, "Enhanced computer vision with Microsoft Kinect sensor: A review," *IEEE transactions on cybernetics*, vol. 43, no. 5, pp. 1318–1334, 2013.

[7] Manuel Martinez, Boris Schauerte, and Rainer Stiefelhagen, ""bam!" depth-based body analysis in critical care," in *International Conference on Computer Analysis of Images and Patterns (CAIP)*, New York, NY, USA, 2013, pp. 465–472, Springer-Verlag New York, Inc.

[8] Manuel Martinez, Lukas Rybok, and Rainer Stiefelhagen, "Action recognition in bed using BAMs for assisted living and elderly care," in *IAPR International Conference on Machine Vision Applications (MVA)*, 2015.

[9] Timo Grimm, Manuel Martinez, Andreas Benz, and Rainer Stiefelhagen, "Sleep position classification from a depth camera using bed aligned maps," in *International Conference on Pattern Recognition (ICPR)*, 2016.

[10] C. Stauffer and W.E.L. Grimson, "Adaptive background mixture models for real-time tracking," 1998, vol. 2, pp. 246–252.

[11] James W. Davis, "Hierarchical motion history images for recognizing human motion," in *IEEE Workshop on Detection and Recognition of Events in Video*, July 2001.

[12] T. Ahonen, A. Hadid, and M. Pietikainen, "Face recognition with local binary patterns," 2004, pp. 469–481.

[13] Antonio Criminisi and Jamie Shotton, *Decision Forests for Computer Vision and Medical Image Analysis*, Springer Publishing Company, Incorporated, 2013.

[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.