Title

Sleep Pose Recognition in an ICU Using Multimodal Data and Environmental Feedback

ISBN

978-3-319-20903-6

Authors

Torres, Carlos
Hammond, Scott D
Fried, Jeffrey C
et al.

Publication Date

2015

DOI

10.1007/978-3-319-20904-3_6

Peer reviewed

Sleep Pose Recognition in an ICU Using Multimodal Data and Environmental Feedback

Carlos Torres[1(✉)], Scott D. Hammond[1], Jeffrey C. Fried[2], and B.S. Manjunath[1]

[1] Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, USA
{calostorres,manj}@ece.ucsb.edu, shammond@tmrl.ucsb.edu
http://vision.ece.ucsb.edu
[2] Santa Barbara Cottage Hospital, Santa Barbara, USA
jfried@sbch.org

Abstract. Clinical evidence suggests that sleep pose analysis can shed light onto patient recovery rates and responses to therapies. In this work, we introduce a formulation that combines features from multimodal data to classify human sleep poses in an Intensive Care Unit (ICU) environment. As opposed to the current methods that combine data from multiple sensors to generate a single feature, we extract features independently. We then use these features to estimate candidate labels and infer a pose. Our method uses modality trusts – each modality's classification ability – to handle variable scene conditions and to deal with sensor malfunctions. Specifically, we exploit shape and appearance features extracted from three sensor modalities: RGB, depth, and pressure. Classification results indicate that our method achieves 100 % accuracy (outperforming previous techniques by 6 %) in bright and clear (ideal) scenes, 70 % in poorly illuminated scenes, and 90 % in occluded ones.

1 Introduction

The tenets of evidence-based medicine implore clinicians and researchers to collect and process all available data in a specific healthcare setting. New methods for non-disruptive monitoring and analysis of patient sleep poses, patterns, and quality add objective metrics for predicting and evaluating health-related scenarios. There are clear clinical examples where patient poses are correlated to medical conditions. For example, sleep positions affect the symptoms of sleep apnea – where airway obstructions are greatest in supine positions [20]. The symptoms of gastroesophageal reflux disease (GERD) are reduced by laying on the side [11]. Body positioning is important in acute lung injury and prone positioning has been shown to improve outcomes in adult respiratory distress syndrome [5]. Prone and

This project is supported in part by the Institute for Collaborative Biotechnologies (ICB) through grant W911NF-09-0001 from the U.S. Army Research Office. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

supine positions worsen back and spine problems, so lateral positioning is recommended by medical experts [4]. Physicians recommend that pregnant women lay on their sides to improve fetal blood flow [16]. The standard of care for immobile ICU patients is to rotate them every two hours to prevent decubitus ulcers, but this is rarely accomplished or effective [22].

The previous examples show that poses can be manipulated to improve patients' health. Therefore, accurate pose detection and classification is relevant to healthcare. The findings in [2,10,23] correlate body positions to various effects on health and quality of sleep of ICU patients. The authors state that identification of sleep poses in natural scenarios helps to evaluate sleep and to improve diagnosis and treatment of sleep disorders. Current physiological systems use machines that physically connect to the patients, making them disruptive and intrusive. Purely observational systems use images and pressure arrays to estimate poses but have been unable to handle natural scenes – indoor ICU scenes with variable illumination and occlusions such as blankets and pillows.

There are two major approaches for the study of sleep. One approach uses bio-status data to monitor a patient's metabolic state during sleep [12,14,18]. The polysomnogram is the standard equipment used in these studies. Its motion-restricting probes connect to the patient's head, face, and respiratory system, monitoring brain activity, rapid-eye-motion (REM) signals, and levels of oxygen and carbon-dioxide in the blood. The second approach is based on the identification of sleep patterns using non-intrusive equipment and human observers [7,15]. Computer vision methods are used in [13,15,18] but are limited to ideal scenes. In both approaches, the staging needed for observation affects the measurements. In order to overcome these issues, we propose to use three non-invasive, independent sensor modalities: RGB, depth, and pressure. Existing techniques are able to estimate human poses in ideal scenes using these modalities independently, but they fail in challenging ones. In [24] the authors present a generative approach that uses deformable parts model (DPM), commonly used in RGB images. Unfortunately, the DPM method requires images with relatively uniform illumination and with only minor self-occlusions. The discriminative approach from [21] uses depth images and is robust to illumination changes. However, this method requires clean depth segmentation and contrast, and it fails under occlusions. Neither of these methods works in unconstrained ICU scenarios.

Our work is most similar to [9], where standard RGB images and a low-resolution pressure array were used to classify sleep poses from static images. Their method used normalized geometric and load distribution features that depended on a clear view of the scene and the actor. They used interdependent data from RGB and pressure sensors – if one modality failed, no result was produced. Our method uses data from three modalities independently and then combines their estimation results using modality trusts to infer the final pose label. Moreover, our classification method is independent of body type and we use it to improve the unimodal decision of two common classifiers: Linear Discriminant Analysis (LDA) and Support Vector Classifier (SVC). The major contributions of this work are:

1. A new system configuration of complementary sensors to analyze sleep poses in healthcare scenarios. This modular system can be easily adapted to address a number of healthcare tasks and natural indoor scene constraints.
2. A new set of tuned features that capture the shape and appearance of human sleep poses.
3. The formulation of a novel multimodal concept for this application called modality trust, which leverages the ability of each individual modality for reliably representation of the human sleep poses.

2 System Description

The proposed system shown in Fig. 1 uses three sensor modalities: a single Carmine camera, with standard RGB and depth sensors by Primesense, and a high-resolution, pressure-sensing mattress by Tekscan. The Carmine and Tekscan devices are controlled by DuoCore computers, which communicate via TCP-IP and are synchronized using Network Time Protocol (NTP). The sensors monitor the bed and actors in a variety of poses and scenes as described in Sect. 2.1. The scene context (e.g., illumination and occlusions) is captured by the illumination, proximity, and radio-frequency identification (RFID) sensors.

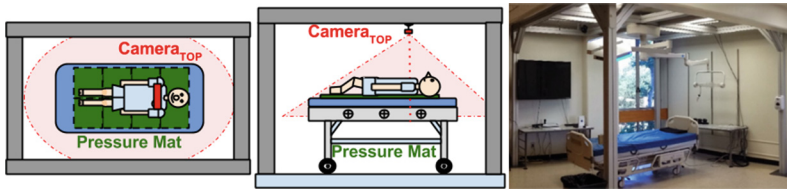

Fig. 1. Top and side profiles (left) of the multimodal system with top camera view (red) and pressure mat (green) and mock-up ICU (right) for data collection and testing (Color figure online).

2.1 Data Collection

Sleep poses are collected from five actors, who were asked to assume each of the ten poses from set $Z = \{$Background, Soldier U, Soldier D, Faller R, Faller L, Log R, Log L, Yearner R, Yearner L, Fetal R, Fetal L$\}$. The set Z has size L and is indexed by l. The letters in the labels U and D stand for facing-Up and facing-Down and L and R stand for laying-on-Left and laying-on-Right. The letter and subscript z_l is used to identify a specific pose label (e.g., $z_0 =$ Background). The scene conditions are simulated using three illumination levels: bright (light sensor within 70–90 % saturation), medium (50–70 %), and dark (below 50 %) and four occlusion types: clear (no occlusion), blanket (covering 90 % of the actor's body), blanket and pillow, and pillow (between actor's upper body and the pressure mat). The illumination intensities are assigned using the percent saturation

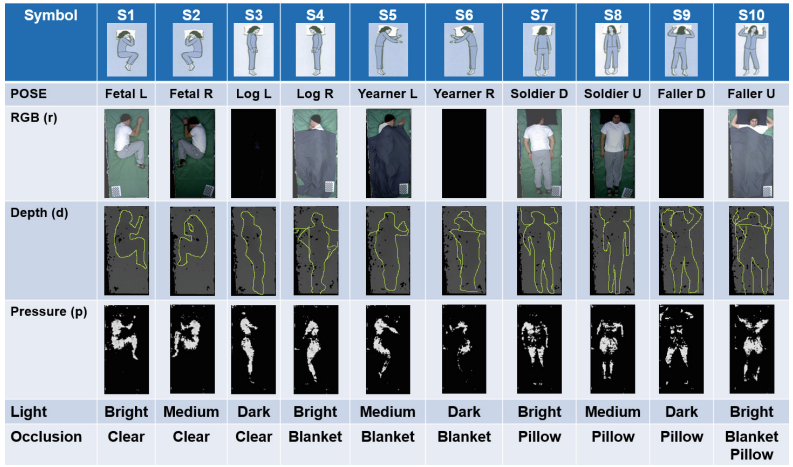

Fig. 2. Sample dictionary of sleep poses showing one actor in various poses and scenes. The top row shows the pose symbol, configuration, and orientation. The second row shows the pose names, where L and R indicates Left or the Right orientation, and U and D indicate facing Up or Down. The third, fourth, and fifth rows are the R, D, P pose representations. The D images on the fourth row are manually delineated to highlight the background and body differences. Finally, the bottom two rows describe the scene.

values and the occlusions are detected using inexpensive RFID and proximity sensors, all by .NET Gadgeteer. The combination of the illumination levels and occlusion types generates a 12-element scene-set $C = \{$(bright, medium, dark) $\times$ (clear, blanket, pillow, blanket + pillow)$\}$. Single illumination and occlusion combination (e.g., bright and clear) is represented using $c \in C$. The dataset is created assuming one scene to be the combination of one actor in one pose and under a single scene condition. From one scene four measurements are collected – three modalities from one camera view (RGB, depth, and a synthetic binary mask) and one pressure image. The data collection included background (bed without actor) images, and images of the actors in each of the 10 poses (11 classes including the background) under each of the 12 scene conditions. The process is repeated ten times for each of the five actors; this generates a dataset of 26,400 images (5 actors $\times$ 10 sessions $\times$ 4 images $\times$ 11 classes $\times$ 12 scenes). The modalities are calibrated using the methods from [6]. Sample data is shown in Fig. 2 and the complete set is available online at http://vision.ece.ucsb/research.

2.2 Feature Extraction

Features are extracted from R, D, P images after subtracting the background, converting them grayscale, and normalizing their pixel intensities.

Histogram of Oriented Gradients (HOG). The proposed formulation uses the HOG feature descriptor [3], extracted from R images, based on its ability to represent human limb structures, which is demonstrated in [24].

Image Geometric Moments (gMOM). Moments [8] extracted from the D and P images are used to describe the shape of the poses and are computed via:

$$\hat{M}_{i,j} = \sum_{x,y} I(x,y) x^j y^i, \tag{1}$$

where moment order is given by $i+j$, $i,j \geq 0$ are the horizontal and vertical orders, and I is the binary intensity value (0 or 1) for pixel at coordinates x, y. Its abilities for pose shape representation were demonstrated in [1,19]. The images are tiled using a six-by-six grid and the raw pixel values ([0,1]) are used to compute up to the third moment from each block.

3 Multimodal Classification with Trust

The modality trust (w_m^c) is defined as the ability of feature vector f_m for pose classification. The vector f_m is extracted from modality m under scene conditions c. The trusts are estimated at training, using all the features in the subset X_{train}^c to compare estimated pose label $\hat{z}_k$ to the ground-truth label z_k^* and to record the matches. The learned trusts are used to infer a final multimodal label.

3.1 Trust Estimation

The set of modality trusts $\{w_1, w_2, \ldots w_M\}^c$ is estimated for modalities in N and condition c. The estimation of the modality trusts is divided into three stages: unimodal training, classifier validation, and trust normalization.

Unimodal Training. In this step unimodal SVC and LDA classifiers CLF_m^c are trained using the features f_m in X_{train}^c. Each of the unimodal classifier outputs a vector of length L of the form $[\hat{s}_{l,k}(f_m)] = [\hat{s}_{1,k}(f_m), \ \ldots, \hat{s}_{L,k}(f_m)]$. Given a datapoint X_k (with M unimodal feature vectors f_m), the $\hat{s}$ elements contain the scores for each of the L labels in Z.

Classifier Validation. At this stage the estimated unimodal labels ${Z_{\hat{l},k}}^{(m)}$ are compared to the ground truth label $Z_{l,k}^*$ from data point X_k. The label matches are stored in the array $\mathbf{b}$ of dimensions $[K, M]$ for all k datapoints ($1 \leq k \leq K$, and $K = |X_{train}|$) using Algorithm 1.

Trust Normalization. Finally, the trusts are estimated with following equation:

$$w_m = \frac{\sum_{k=1}^{K} \mathbf{b}[k,m]}{K}, \tag{2}$$

and normalized so that the sum is one.

Algorithm 1. Unimodal Classifier Validation Vector (**b**)

procedure COMPARE($Z_{\hat{l},k}{}^{(m)}, Z^*_{l,k}$) ▷ Estimated and ground truth labels
  $\mathbf{b} \leftarrow 0, m, k = 0,$ ▷ Initialize array
  for k **do**
    for m **do**
      if $Z_{\hat{l},k}{}^{(m)} = Z^*_{l,k}$ **then**
        $\mathbf{b}[k, m] \leftarrow 1$
      else
        $\mathbf{b}[k, m] \leftarrow 0$
      end if
    end for
  end for
  return b ▷ Vector of size K, M
end procedure

3.2 Multimodal Formulation

The overall description of the system is shown in Fig. 3. It uses the unimodal sensor training data (features) to estimate the trust values. Then the system uses the trusts to refine mulitmodal classification and produce a final pose label for a given test datapoint. First, the system applies a weighted scoring formulation to the unimodal label candidates obtained from the features $\left[f_1, f_2, f_3\right] = \left[\text{HOG}(R),\ \text{gMOM}(D),\ \text{gMOM}(P)\right]$ of datapoint X_k. Finally, the multimodal label $Z_{\hat{l},k}$ is the one with the maximum weighted score as follows:

$$S^c_{m,k} = S^c(X_k[f_m]) = w^c_m \text{CLF}^c_m(f_m), \tag{3}$$

where w^c_m represents the predictive power of feature f_m with scene conditions c, and CLF^c_m is the unimodal classifier score vector $\left[\hat{s}_{1,k}(f_m),\ \ldots, \hat{s}_{L,k}(f_m)\right]$

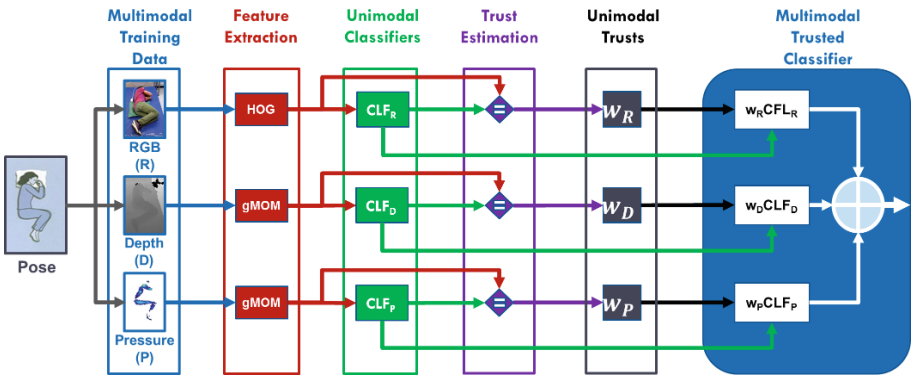

Fig. 3. Diagram of our proposed multimodal sleep pose classification method. The system uses modalities to exploit various scene and sensor properties. Second, features are extracted from the R, D, P pose representations and used to train unimodal classifiers and estimate modality trusts. Finally, trusts are used to refine the output.

(elements are label scores). Thus $S^c_{m,k}$ has L elements representing the unimodal label scores for an input X_k. The multimodal score is computed using:

$$S^c_k = \sum_{m=1}^{M} S^c_{m,k} = \sum_{m=1}^{M} \Big(w^c_m [\hat{s}_{1,k}(f_m), \ldots, \hat{s}_{L,k}(f_m)]^c \Big), \tag{4}$$

The vector S^c_k has L candidate scores values for each k and is computed via:

$$S^c_k = \sum_{m=1}^{M} \Big(w^c_m \{\hat{s}_{l,k}(f_m)\}^c_L \Big). \tag{5}$$

Therefore, given an input vector $X_k = \{f_m\}_M$ from scene c the estimated pose label is $Z_{\hat{l}}$, and the index $\hat{l}$ is computed using the following equations:

$$\hat{l} = \arg\max_{l \in L} \Big(S^c_k \Big), \tag{6}$$

where $\hat{l}$ is the index of the label with the highest trusted score from:

$$\hat{l} = \arg\max_{l \in L} \left(\sum_{m=1}^{M} w^c_m \{\hat{s}_{l,k}(f_m)\}^c_L \right). \tag{7}$$

Missing Modalities. Hardware malfunctions were simulated by omitting information from one modality (set its value to zero), proportionally adjusting the trusts of the remaining ones, and testing the system with the new trust values.

4 Experiments

Experiments are conducted using two classification methods: multi-class linear SVC and LDA from [17]. The experiments use five-fold cross-validation scheme for all reported accuracies. Results indicate that illumination affects R performance, while the performances of D and P remain constant. Recognition using R and D is affected by visual occlusions (blankets) and P is affected by pillows.

Unimodal. Initially, the system is trained/tested with a single concatenated vector (RDP), and the unimodal vectors R, D, P. This assessment provides a performance basis for classification and justifies the need for a multimodal approach. Results indicate that neither the concatenation of all nor the use of a single modality can be used directly to recognize poses across all scenes.

Multimodal. The multimodal experiments show that our system reliably classifies sleep poses in ICU scenarios using modality trust. To the best of our knowledge, there is no other method that considers our range of scenarios. Performance contrast of the system in various scenes is shown in Table 1. The table includes classification accuracies of the trusted multimodal system, a Majority-Vote-Learner (MaVL), and an in-house implementation of the method from [9].

Table 1. Mean multimodal sleep pose classification accuracy of two competing methods and our proposed multimodal trust using SVM (SVC) and LDA classifiers. Our method matches the performance of two competing methods in bright clear scenes and it outperforms them by a range of approximately 30 to 50 %.

Scene		Competing		Proposed	
Illumination	Occlusion	MaVL (RDP)	Huang (RP)	SVC (RDP)	LDA (RDP)
Bright	Clear	80	100	100	100
	Blanket	82	8	85.8	80.4
	Blanket + Pillow	65	6	85.8	83.6
	Pillow	54	58	90	90
Medium	Clear	80	88	100	100
	Blanket	65	7	85.3	80.6
	Blanket + Pillow	57	7	85.3	83.6
	Pillow	78	37	90	90
Dark	Clear	17	–	81.2	85
	Blanket	20	–	20.0	19.2
	Blanket + Pillow	32	–	17.7	18.6
	Pillow	60	–	24.5	22.3

Table 2. Mean classification accuracy with incomplete multimodal information. One modality is removed (\) and the modality trust values of the remaining ones is adjusted for SVC and LDA estimated labels.

Scene		SVC			LDA		
Illumination	Occlusion	$RD \setminus P$	$RP \setminus D$	$DP \setminus R$	$RD \setminus P$	$RP \setminus D$	$DP \setminus R$
Bright	Clear	100	100	100	100	100	100
	Blanket	85	90	95	80	85	92
	Blanket + Pillow	80	85	90	88.6	83.6	83.6
	Pillow	85	88	87	90	85	95
Medium	Clear	100	100	100	100	100	100
	Blanket	70	80	75	68.6	78.6	88.6
	Blanket + Pillow	65	70	71	73.5	81.6	83.6
	Pillow	81	85	87	77.3	82	85
Dark	Clear	54.1	47.7	72.7	29.5	74.1	76.4
	Blanket	7.5	30	35	23.2	68.6	76.8
	Blanket + Pillow	6	27	30	12.8	53	68.9
	Pillow	12	37	45	36.3	65.1	73.7

Missing Modalities. We test the limits of our system by omitting one modality and adjusting the contributions of the remaining modalities. We report SVC and LDA accuracies for all considered scenes in Table 2. Results show that our system performs poorly without pressure information, achieving a classification accuracy of 6 % using SVC and 12.8 % using LDA for dark and occluded scenes.

Confusion Matrices. The confusion matrices of our method and [9] are compared in Fig. 4. The main diagonal on the right shows that our method outperforms the competing method in challenging scenarios.

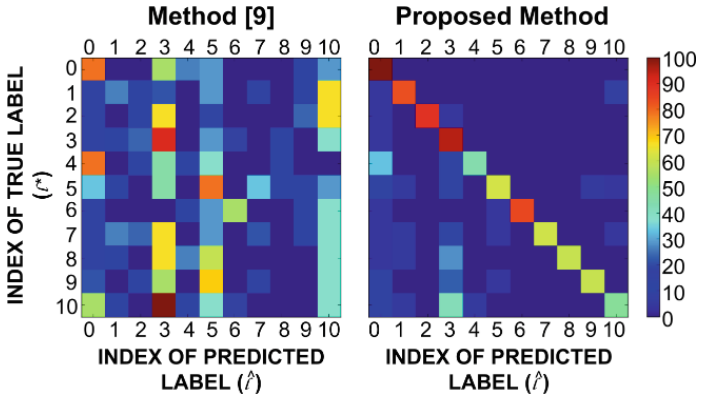

Fig. 4. Confusion matrices of implemented method from [9] with 16 % and our proposed method with 70 % accuracies for dark and occluded scenarios. The confusion matrices show how the indexes of the estimated labels $\hat{l}$ (x-axis) match the actual labels l^* (y-axis). The main diagonal indicates that our method (right) performs better.

5 Discussion

The parameters used for the computation of HOG features are: four orientations, 16×16 pixels per cell, and two-by-two cells per block. The geometric moment parameters were empirically tuned to achieve the highest pose classification accuracy possible. First, moments were extracted from the whole image (one-by-one grid) and their classification performance was tested. The one-by-one grid pose descriptors achieved a mean classification accuracy of 21 % over all scenes and 31 % for the bright, clear one. The grid dimensions were sequentially increased and revealed that a six-by-six grid yielded the highest accuracy without dramatically increasing computation time. Using the six-by-six grid, the system achieved a mean accuracy of 79 % over all scenes and 97 % for the bright and clear scene. Concatenation of features from the whole image and the six-by-six grid did not improve classification. The moment order was tuned alongside the grid dimensions. Shape descriptors were generated by computing up to the third geometric moment from each block; this yielded a ten-element vector per block or 360-element vector per image. Greater order moments increased estimation errors as reported in [8] and did not improve classification. The implementation of [9] to classify poses achieved an accuracy of 100 % in scenarios with bright and medium illumination. The performance increase (the authors reported a 94 % accuracy) is likely due to tunning parameters, higher resolution and complete bed coverage of the Tekscan mat. The $C = 0.5$ parameter for SVC was estimated during training with a validation dataset.

5.1 Conclusion

In this work, we presented a multimodal system to classify sleep poses in natural ICU scenarios. The system handles challenging conditions by relying on measurable variables from environmental sensors. We validated the sensor selection and features experimentally and showed that they provide accurate representations of sleep poses. Quantitative results indicate that the system has a performance increase of 6 % with respect to two existing methods in ideal scenarios and outperforms them significantly in dark and occluded ones. Reliability of the method was tested by sequentially omitting information from one modality and adjusting the remaining modalities via interpolation. With this scheme, the multimodal system achieved a pose classification accuracy of 47 % in challenging scenes.

6 Future Work

The multimodal system achieved high classification accuracies for most conditions; however, some scenarios caused a performance drop (e.g., 70 % in dark scenes) as shown in Table 1 and require further investigation. The system performs reliable classification of sleep poses in natural static ICU scenes. Nevertheless, we are exploring methods that integrate temporal information for the analysis of pose transitions and patterns. Extensions of this work will investigate new methods that are robust to an unconstrained set of body pose configurations, which better represent the poses of bed-ridden patients. Clinical deployment may impede the use of pressure mats due to sanitation requirements, so we are actively devising techniques that do not require this modality. Future work will explore new methods to estimate trust and combine multimodal classifiers (i.e., boosting), avoid the use of pressure mats, and integrate temporal information.

References

1. Ahad, M.A.R., Tan, J.K., Kim, H., Ishikawa, S.: Motion history image: its variants and applications. Mach. Vis. Appl. **23**(2), 255–281 (2012)
2. Bihari, S., McEvoy, R.D., Matheson, E., Kim, S., Woodman, R.J., Bersten, A.D.: Factors affecting sleep quality of patients in intensive care unit. J. Clin. Sleep Med. Official Publ. Am. Acad. Sleep Med. **8**(3), 301 (2012)
3. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of the IEEE Conference Computer Vision and Pattern Recognition (2005)
4. Gordon, S.J., Grimmer, K.A., Trott, P.: Understanding sleep quality and waking cervico-thoracic symptoms. Int. J. Allied Health Sci. Pract. **5**, 1–12 (2007)
5. Guérin, C., Reignier, J., Richard, J.C., Beuret, P., Gacouin, A., Boulain, T., Mercier, E., Badet, M., Mercat, A., Baudin, O., et al.: Prone positioning in severe acute respiratory distress syndrome. New Engl. J. Med. **368**(23), 2159–2168 (2013)
6. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision, 2nd edn. Cambridge University Press (2004). ISBN: 0521540518
7. Hsia, C.C., Liou, K., Aung, A., Foo, V., Huang, W., Biswas, J.: Analysis and comparison of sleeping posture classification methods using pressure sensitive bed system. In: IEEE International Conference on Engineering in Medicine and Biology Society (2009)

8. Hu, M.K.: Visual pattern recognition by moment invariants. IEEE Trans. Inform. Theory **8**(2), 179–187 (1962)
9. Huang, W., Wai, A.A.P., Foo, S.F., Biswas, J., Hsia, C.C., Liou, K.: Multimodal sleeping posture classification. In: IEEE International Conference on Pattern Recognition (2010)
10. Idzikowski, C.: Sleep position gives personality clue. BBC News, 16 September 2003
11. Khoury, R.M., Camacho-Lobato, L., Katz, P.O., Mohiuddin, M.A., Castell, D.O.: Influence of spontaneous sleep positions on nighttime recumbent reflux in patients with gastroesophageal reflux disease. Am. J. Gastroenterol. **94**(8), 2069–2073 (1999)
12. Koprinska, I., Pfurtscheller, G., Flotzinger, D.: Sleep classification in infants by decision tree-based neural networks. Artif. Intell. Med. **8**(4), 387–401 (1996)
13. Kuo, C.H., Yang, F.C., Tsai, M.Y., Ming-Yih, L.: Artificial neural networks based sleep motion recognition using night vision cameras. Biomed. Eng. Appl. Basis Commun. **16**(02), 79–86 (2004)
14. Lewicke, A., Sazonov, E., Corwin, M.J., Neuman, M., Schuckers, S.: Sleep versus wake classification from heart rate variability using computational intelligence: consideration of rejection in classification models. IEEE Trans. Biomed. Eng. **55**(1), 108–118 (2008)
15. Liao, W.H., Yang, C.M.: Video-based activity and movement pattern analysis in overnight sleep studies. In: IEEE International Conference on Pattern Recognition (2008)
16. Morong, S., Hermsen, B., de Vries, N.: Sleep position and pregnancy. In: de Vries, N., et al. (eds.) Positional Therapy in Obstructive Sleep Apnea, pp. 163–173. Springer, New York (2015)
17. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
18. Penzel, T., Conradt, R.: Computer based sleep recording and analysis. Sleep Med. Rev. **4**(2), 131–148 (2000)
19. Ramagiri, S., Kavi, R., Kulathumani, V.: Real-time multi-view human action recognition using a wireless camera network. In: International IEEE Conference on Distributed Smart Cameras (2011)
20. Sahlin, C., Franklin, K.A., Stenlund, H., Lindberg, E.: Sleep in women: normal values for sleep stages and position and the effect of age, obesity, sleep apnea, smoking, alcohol and hypertension. Sleep Med. **10**(9), 1025–1030 (2009)
21. Shotton, J., Girshick, R., Fitzgibbon, A., Sharp, T., Cook, M., Finocchio, M., Moore, R., Kohli, P., Criminisi, A., Kipman, A., et al.: Efficient human pose estimation from single depth images. IEEE Trans. Pattern Anal. Mach. Intell. **35**(12), 2821–2840 (2013)
22. Soban, L., Hempel, S., Ewing, B., Miles, J.N., Rubenstein, L.V.: Preventing pressure ulcers in hospitals. Jt Comm. J. Qual. Patient Saf. **37**(6), 245–252 (2011)
23. Weinhouse, G.L., Schwab, R.J.: Sleep in the critically ill patient. Sleep-New York Then Westchester **29**(5), 707 (2006)
24. Yang, Y., Ramanan, D.: Articulated human detection with flexible mixtures of parts. IEEE Trans. Pattern Anal. Mach. Intell. **35**(12), 2878–2890 (2013)