#340

Presented at the Winter Conference on Computer Vision Applications (WACV), Lake Placid, NY, March 7-9, 2016

EYE-CU: SLEEP POSE RECOGNITION USING MULTIMODAL MULTIVIEW DATA

Carlos Torres†, Victor Fragoso‡, Scott D. Hammond†, Jeffrey C. Fried*, and B. S. Manjunath†

†University of California Santa Barbara, ‡West Virginia University, and *Santa Barbara Cottage Hospital

web: torrescarlos.com

Motivation

Manual analysis of body poses of bed-ridden patients requires staff to continuously track and record patient poses. Two limitations in the dissemination of pose-related therapies are scarce human resources and unreliable automated systems. This work addresses these issues by introducing a new method and a new system for robust automated classification of sleep poses in an ICU environment. The new method, coupled-constrained Least-Squares (cc-LS), uses multimodal and multiview (MM) data and finds the modality trust values that minimizes the difference between expected and estimated labels

Healthcare Statistics In the ICU

- 5 million people per year are admitted to the ICU.
- 46% are over the age of 65.
- Annual national ICU cost is $130 billion and rising $5 billion per year.
- Average duration of stay in the ICU is 9.3 days.
- Mortality rate is 10-30% and increases by 7% per day.

- Year 2020 estimates
  - ICU elderly population will increase to 69%.
  - Caregiver workforce will shrink by 35%.

Effects of Poses on Health

Sleep Deprivation

- Bad night increases ICU stay by 10%.
- Sleep position(s) → quality of sleep
- Measuring techniques:
  - Intrusive polysomnogram
  - Non-objective surveys

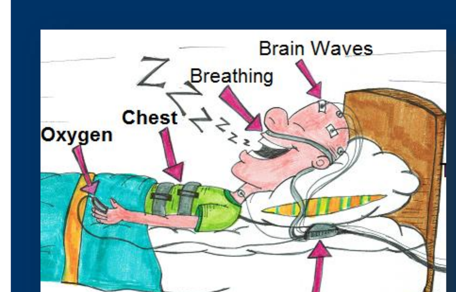

Figure 1. Sleep analysis using polysomnogram electrodes connected to patients head, torso, and finger.

Decubitus Ulcer

- 2.5M cases (80% in ICU)
- Pose and bony areas
- Measuring techniques:
  - Braden scale (subjective)
  - Rounds & pose rotations

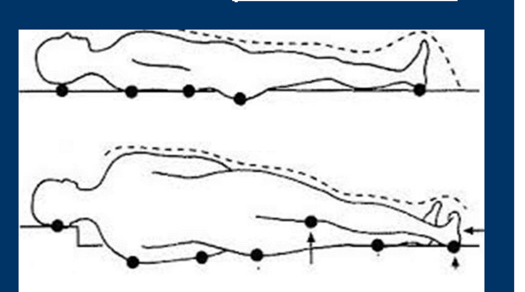

Figure 2. Common decubitus ulcers (bed sores) incidence areas.

Pose Analysis: Current Applications

Sleep Disorder

Decubitus Ulcerations

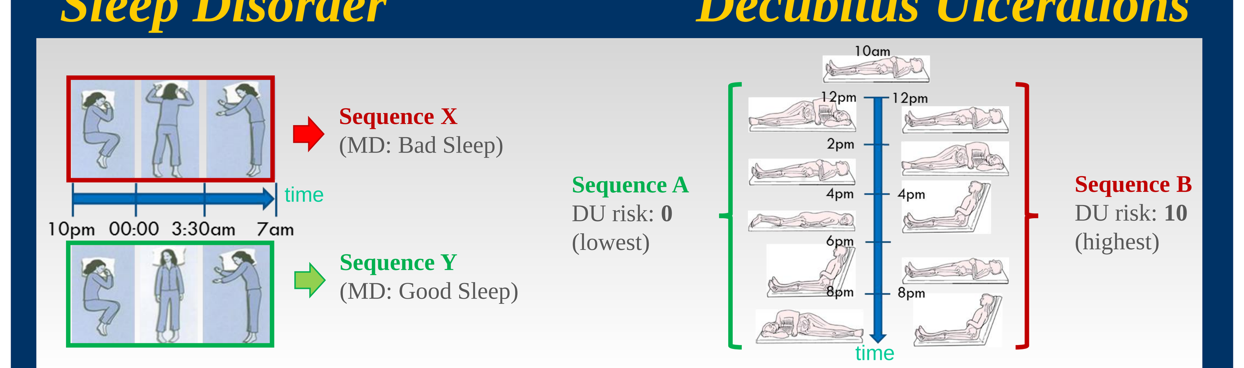

Figure 3. Clinical applications of the Eye-CU system and the cc-LS method for patient sleep pose analysis.

Popular Techniques

Existing techniques include deformable parts models, poses from single-depth images, interdependent pressure and RGB images, and multimodal data. These techniques failed in natural scenarios.

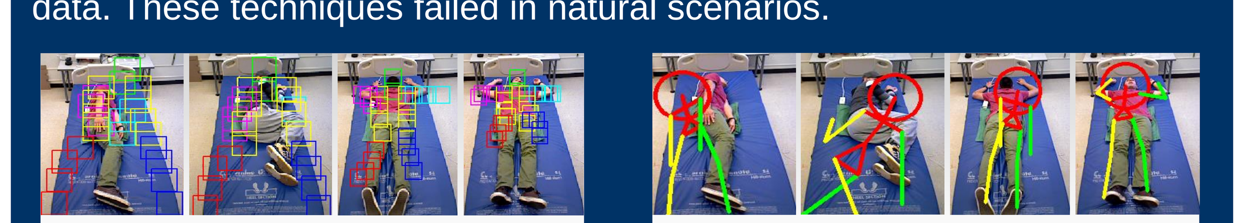

Figure 4. Left: sleep pose performance using deformable part models. Right: sleep pose performance using pose recognition from single depth image (i.e., Kinect API).

The Eye-CU System and the cc-LS Method

Eye-CU System Setup

The Eye-CU system was designed and tested in a mock-up ICU room.

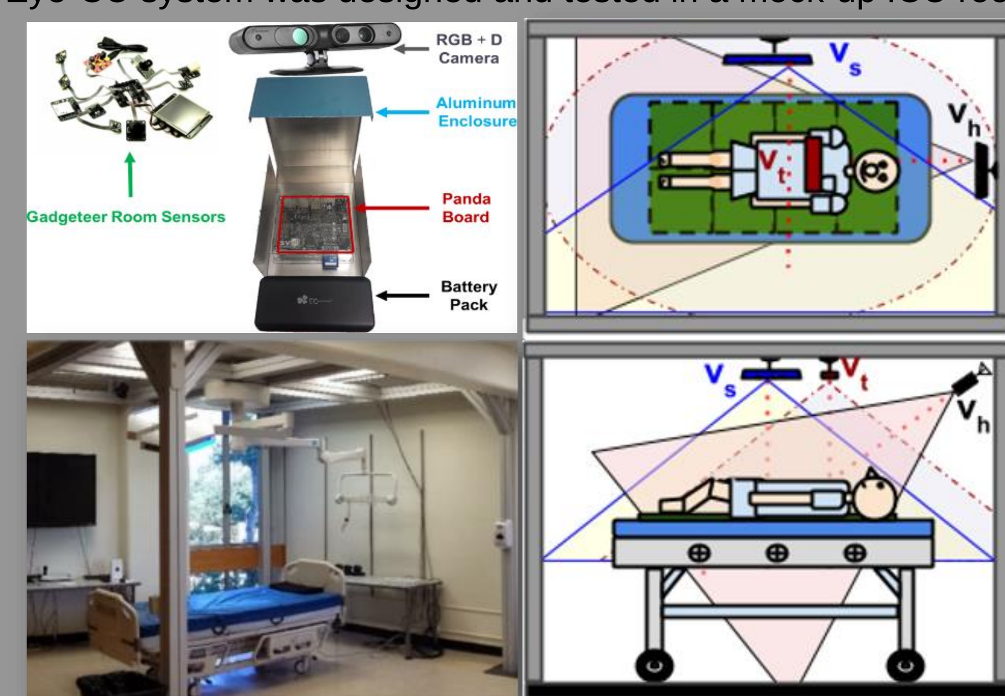

Figure 5. Top left: Eye-CU system node and its components. Top right: perspective top view of the Eye-CU node location w.r.t to the patient and the ICU bed. Bottom left: mock-up ICU room where the system is installed and the sleep pose data is collected. Bottom right: perspective side view of the nodes and the patient.

Feature Vector and Objective

The k-th datapoint extracted from scene c has the following structure:

$$\chi_k^c = \{f_m\}_M^c = \{f_R, f_D, f_P\}_M^c$$

$$= \{\mathrm{HOG}(R), g\mathrm{MOM}(D), gMOM(P)\}_M^c$$

The objective is to find the label index $\hat{l}$, which maximizes the trusted multimodal score:

$$\hat{l} = \arg\max_{l\in L}(S_k^c) = \arg\max_{l\in L} \sum_{m=1}^{M}\left(w_{\mathcal{N}_m}^c\{\mathrm{CLF}(f_{\mathcal{N}_m})\}_L^c\right)$$

$$\mathbf{CLF} = \mathbb{P}(z_l|\chi_k, M = m)$$

$w_{\mathcal{N}_m}^c$ is the trust or weight value for modality m and scene c

$\mathbf{w} = [w_R, w_D, w_P]^{\mathbf{T}}$, c is omitted and $\mathcal{N}_m \to w_m$ ($m = 1$: $w_{\mathcal{N}_m} = w_R$).

Multimodal Multiview Dataset

The dataset will be available online at url: vision.ucsb.edu

Symbol	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
POSE	Fetal L	Fetal R	Log L	Log R	Yearner L	Yearner R	Soldier D	Soldier U	Faller D	Faller U
RGB (R): Views: t s h										
Depth (D) Views: t s h										
Pressure (P)										
Light	Bright	Medium	Dark	Bright	Medium	Dark	Bright	Medium	Dark	Bright
Occlusion	Clear	Clear	Clear	Blanket	Blanket	Blanket	Pillow	Pillow	Pillow	Blanket Pillow

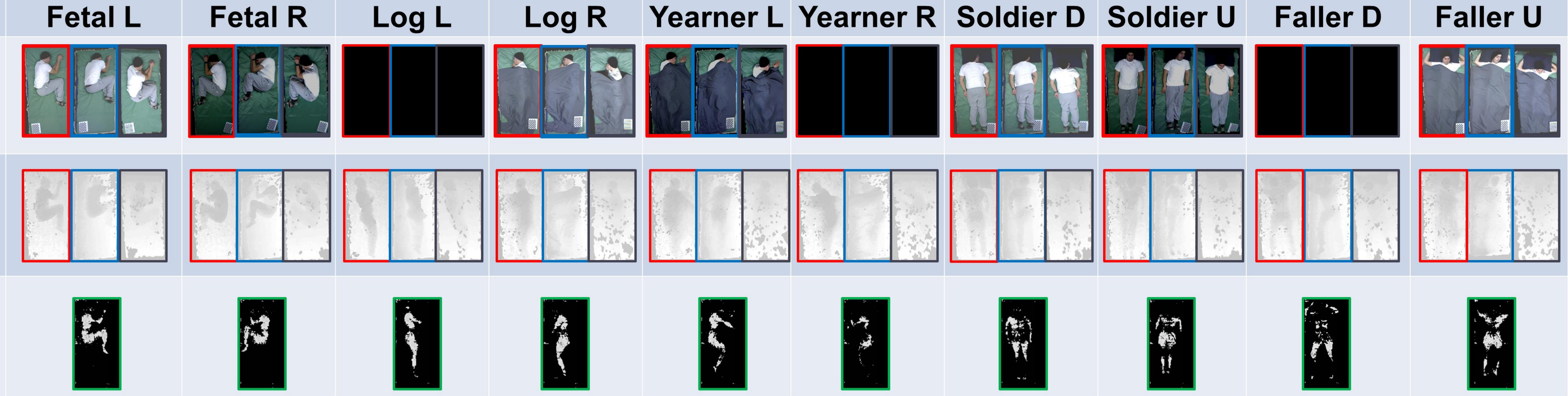

Figure 6. Partial dictionary of multimodal multiview pose data for one actor in 10 poses and 12 scene conditions observed from three views and three modalities

cc-LS Formulation: Computing w

The contribution of each modality and modality-view ($\mathbf{w}$) in the Eye-CU system is estimated by solving multiple optimization problems at once.

$$\underset{\mathbf{w}}{\text{minimize}} \frac{1}{2}\sum_{k=1}^{K}\sum_{l=1}^{L}\left(\sum_{m=1}^{M} s_{k,l,m} w_m - b_{k,l}\right)^2 \qquad (**)$$

$$\text{subject to} \quad \mathbf{1}^T\mathbf{w} = 1$$

$$0 \le w_m \le 1, m = 1, .., M$$

Multimodal Construction

The cc-Ls argument (**) becomes: $\frac{1}{2}\|\mathbf{A}\mathbf{w} - \mathbf{b}\|_2^2$

$$\mathbf{A} = \left[S_{k=1}^T, \dots, S_{k=K}^T\right]_{UM}^T, U = KL,$$

with $S_k(l, m) = s_{k,l,m} = \mathbb{P}(Z_l|\chi_k, M = m)\mathbb{P}(M = m)$

$$\mathrm{b}_m = \left[b_{k=1}^T, \dots, b_{k=K}^T\right]^T,$$

where

$$b_{k,l} = \mathbb{P}(Z = z_l|\chi = \chi_k, Oracle) = \begin{cases} 1, \text{if } \hat{l} = l^* \text{ for } \chi_k \\ 0, \text{otherwise} \end{cases}$$

Finally, $\mathbf{b} = \frac{\sum_{\forall m} \mathrm{b}_m}{M}$

Results

Eye-CU Configurations

Contrast with Competing

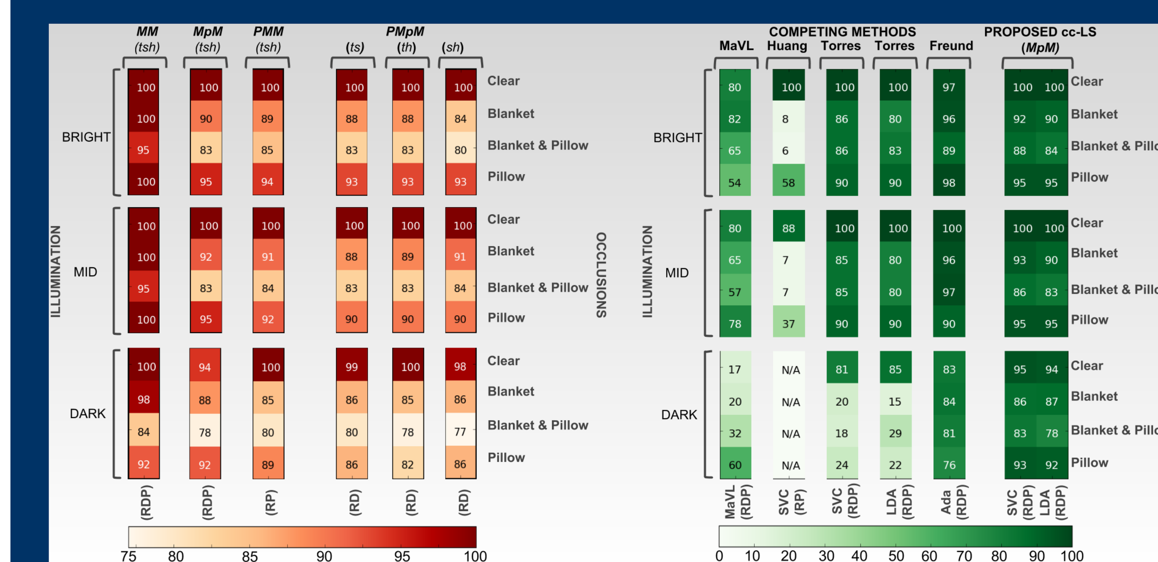

Table 1. Left: Classification performance of the various Eye-CU configurations (based on modalities and views available). Table 2. Right: Classification performance of the MpM cc-LS system configuration and comparable competing methods.

Confusion Matrices: Bright and Clear

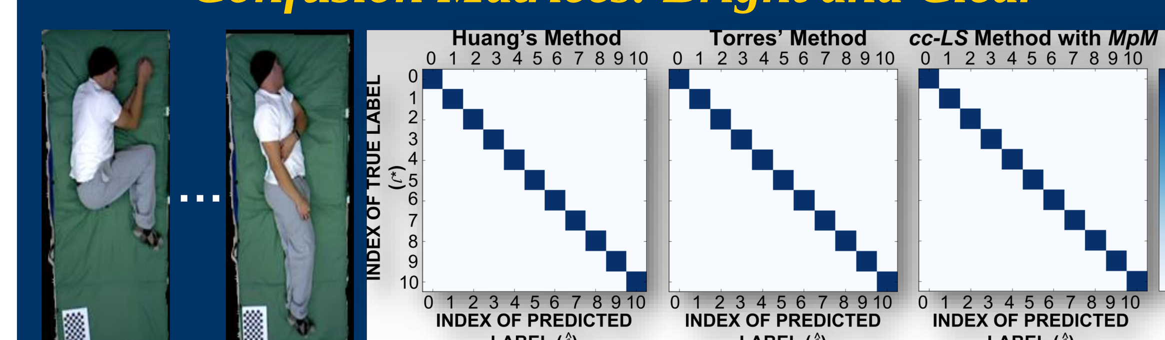

Figure 7. Confusion matrices of the Huang, Torres, and cc-LS with MpM pose classification methods under bight and clear (ideal) scene conditions.

Confusion Matrices: Dim and Occluded

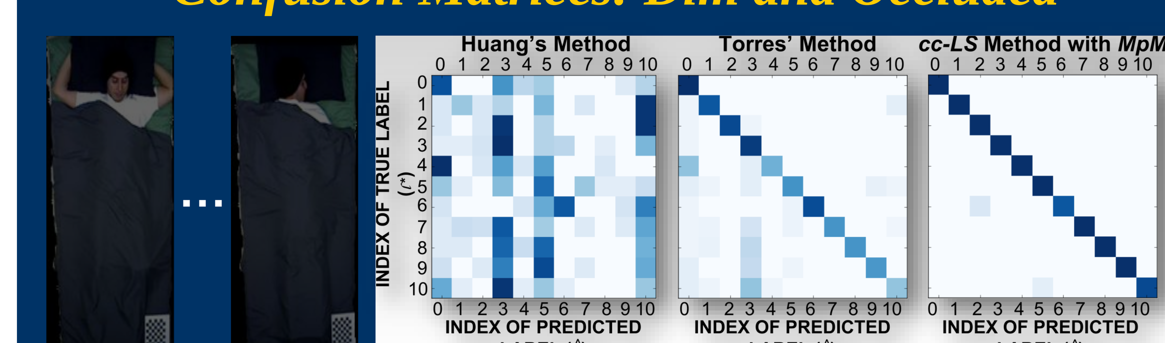

Figure 8. Confusion matrices of the Huang, Torres, and cc-LS with MpM pose classification methods under dim and occluded scene conditions.

Future Work

- Integration of temporal information and pose representation with deep features and artificial neural net architectures.
- Quantify and typify pose sequences (duration and transition) and will investigate removing the constraints from the set of poses.

Main References

1. Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Jrn'l of computer and sys.* Sci., 1997.
2. Huang et al. Multimodal Sleeping posture classification. *In Proc of the IEEE Int'l Conf. on Pattern Recognition (ICPR)*, 2010.
3. Torres et al. Multimodal pose recognition in an ICU using multimodal data and environmental feedback. *In Proc. Of Springer Int'l Conf. on Computer Vision Sys. (ICVS), 2015.*

Acknowledgements

This project was supported in part by the Institute for Collaborative Biotechnologies (ICB) through grant W911NF-09-0001 from the U.S. Army Research Office; and by the U.S. Office of Naval Research (ONR) through grant N00014-12-1-0503. The content of the information does not necessarily reflect the position or the policy of the Government and no official endorsement should be inferred