SLEEP POSE RECOGNITION IN AN ICU USING MULTLIMODAL DATA AND ENVIRONMENTAL FEEDBACK

UCSB-SBCH
VRL-TMRL
2 July 2015

CARLOS TORRES, JEFFREY C. FRIED,

SCOTT D. HAMMOND, and B.S. MANJUNATH

MOTIVATION

HEALTHCARE PROBLEM!

Intensive Care Unit (ICU)

ICU STATISTICS IN THE U.S.*

ICU#	+65 Year Old	Cost	Stay	Mortality	Year 2020
~5M patients/year	46% (2.3M)		9.3days	10-30% (+7% per day)	

Questions:

1. What are (some) the problems in the ICU?
2. What can we do to help?

** Src: Online - US Department of Health and Human Services. April 2014*

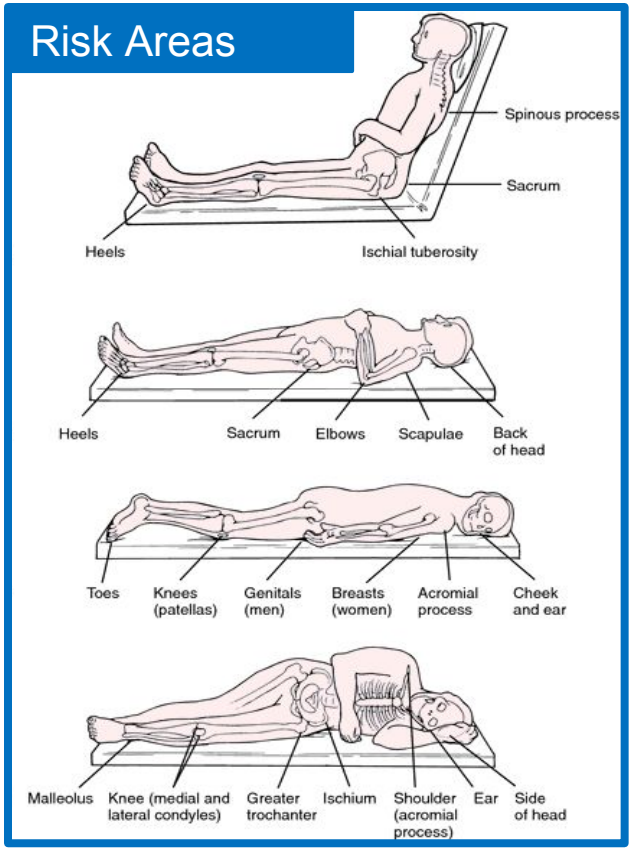
HEALTHCARE – Pose Analysis in the ICU

- Problem 1: Sleep Deprivation* (SD) - sleep hygiene
  - "Bad night" → ICU stay +10%
  - Sleep positions → Quality of sleep
  - Obtrusive measurements + Surveys
  - No prevention

*Sleep in the Critically Ill Patient, Weinhouse and Schwab. Sleep 2006
Factors affecting sleep in the ICU. Bihari et al. JCSM 2012

HEALTHCARE – Pose Analysis in the ICU

- Problem 2: Decubitus Ulcerations** (DUs) – Bed Sores (wounds)
  - 2.5M (80% occur in ICU)
  - Pose (Bony areas)
  - Braden scale (subjective & observational)
  - Rounds & Patient Rotation (2hr, <20%)



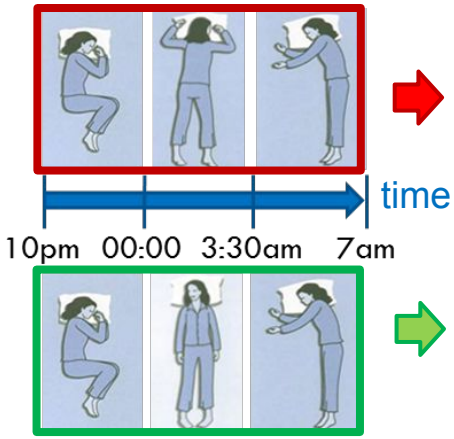
***Preventing Pressure Ulcers in Hospitals. Soban et al. Journal on Quality & Patient Safety 2011*
Online Medical-Dictionary: pressure ulcer, retrieved April 2014

HEALTHCARE – Clinical Impact

Automated Analysis of Sleep Hygiene and DU Incidence/Prevention

- Non-intrusive **automated** data collection and analysis
- Incidence & risk evaluation from **evidence** (measurements vs observations)
- **Individualize** therapies using quantifiable data

SLEEP DISORDER ANALYSIS

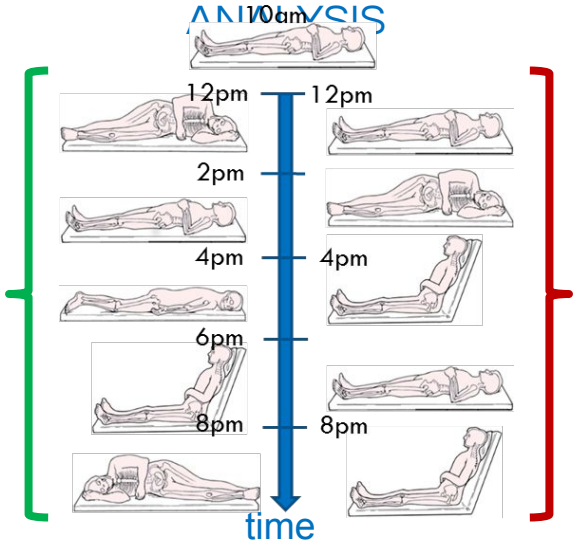


Sequence X
(MD: Bad Sleep)

Sequence Y
(MD: Good Sleep)

DECUBITUS ULCERATION ANALYSIS

Sequence A
DU risk: **0**
(lowest)



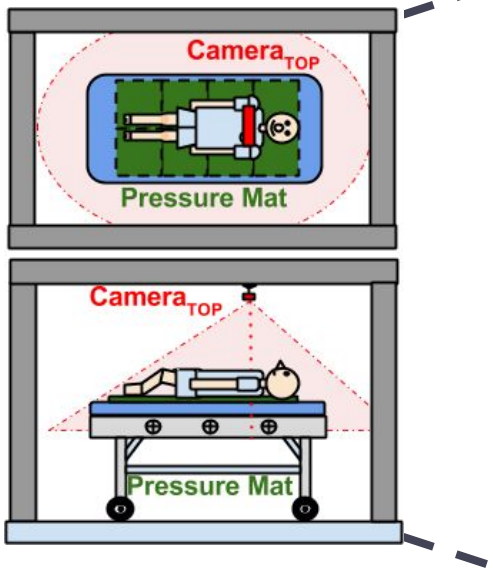
Sequence B
DU risk: **10**
(highest)

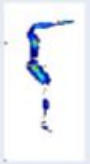
Suitability of Existing Methods

Dry Lab: Mock-Up ICU

Mockup – Data From an Ideal ICU Scenario

PHYSICAL SETUP



NAME	Soldier	Log	Yearner	Fetal	Freefaller	Starfish
POSE						
RGB (R)						
Depth (D)						
4-Pressure						

Shotton et al. Kinect API [1]

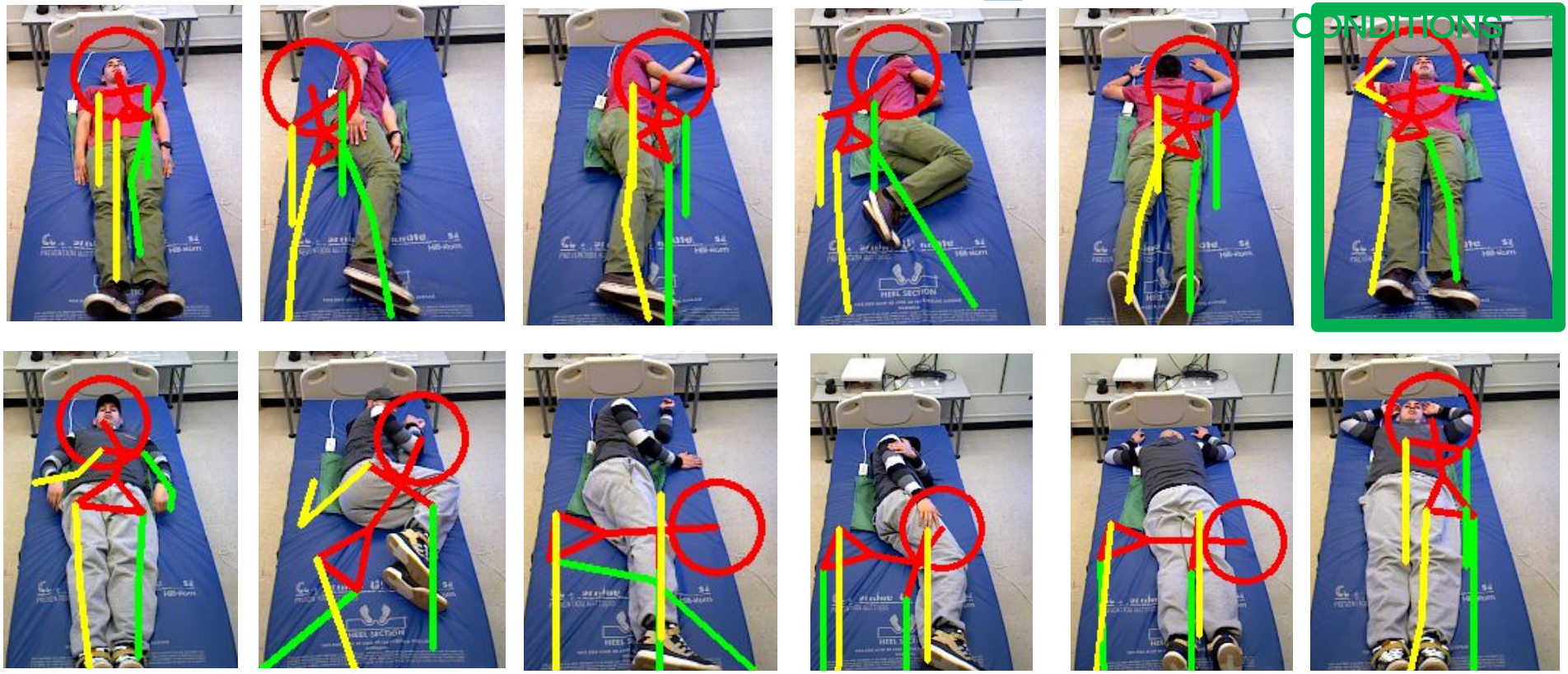
DEPTH ASSUMPTIONS

- Depth Contrast
- Orientation wrt camera

FAILURE CAUSES

- Insufficient Depth Contrast
- Occlusions

BEST CONDITIONS



Yang et al. Flexible Mixture of Parts [2]

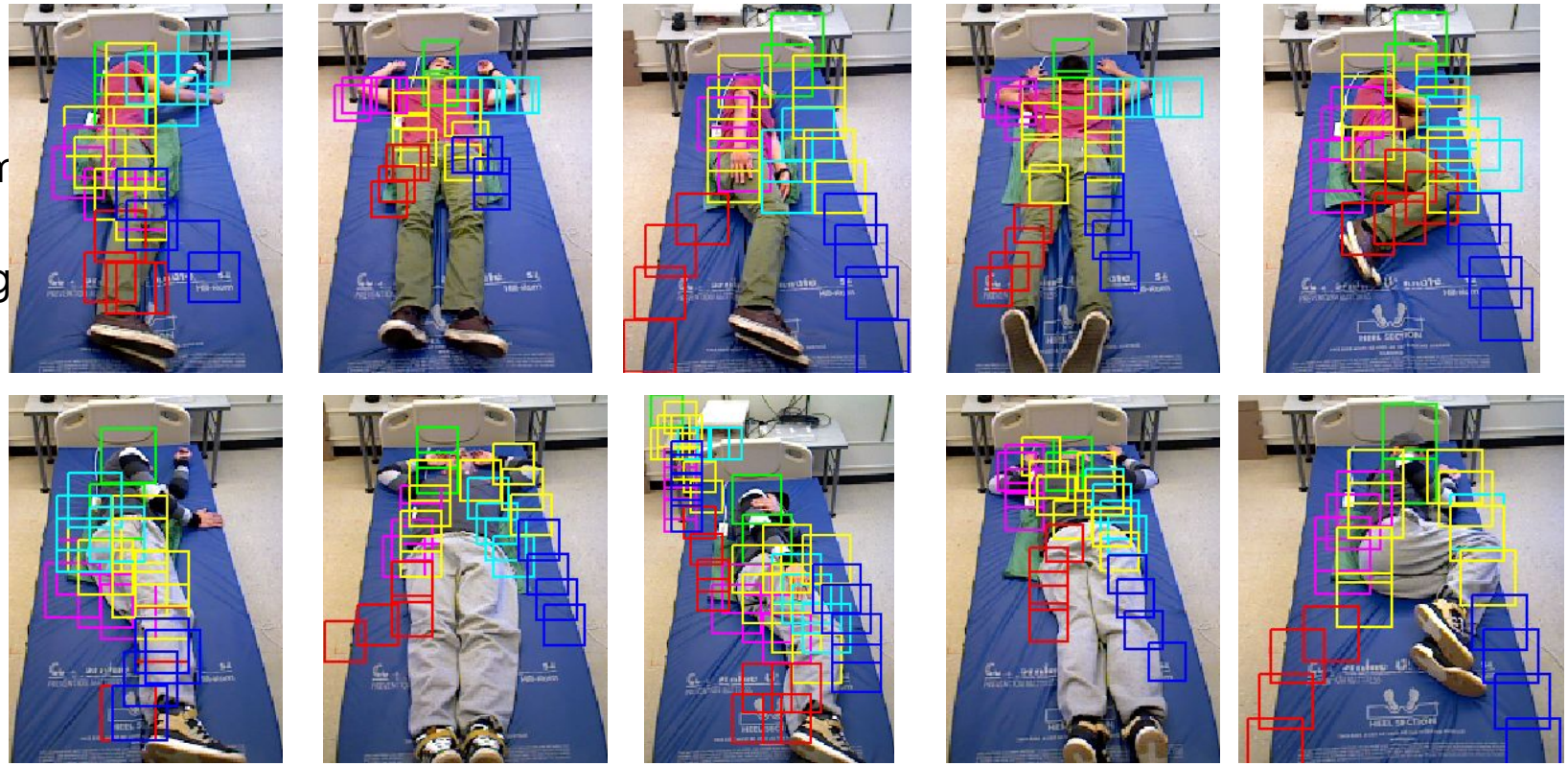
RGB ASSUMPTIONS

- View Point
- "Good" Illumination
- No Occlusions

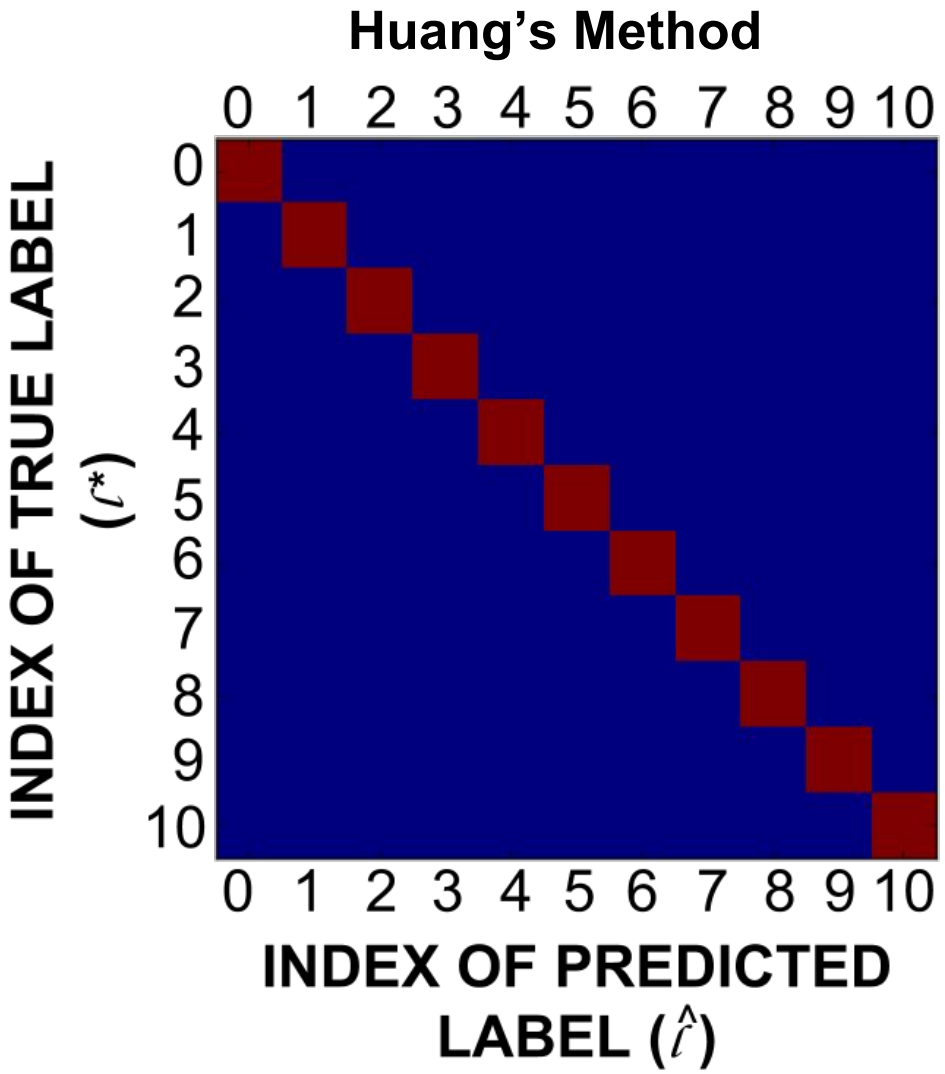
FAILURE CAUSES

- Clutter
- Minor self-occlusions

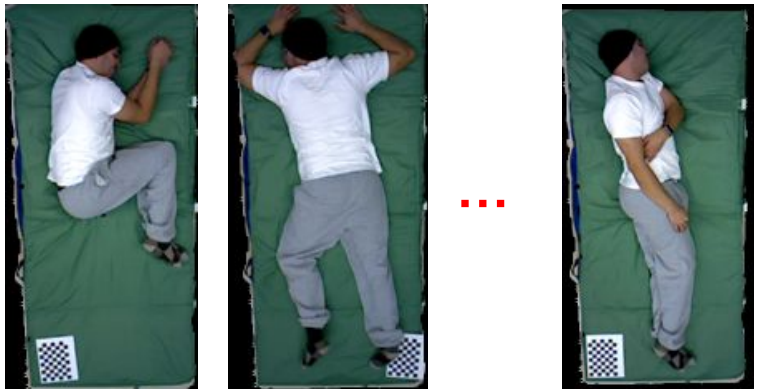
- Head
- Torso
- Right Arm
- Left Arm
- Right Leg
- Left Leg



Huang et al. RGB-Pressure Classification [3]



Ideal ICU Scenes



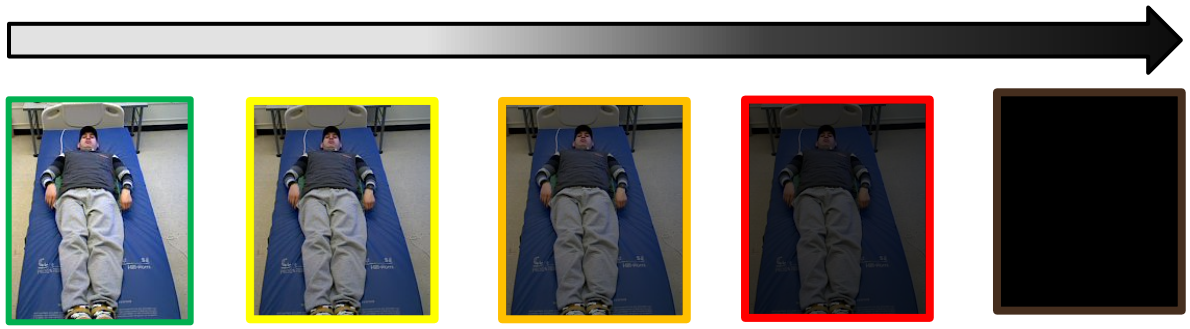
…

PROMISING!

Challenges of ICU Scenarios

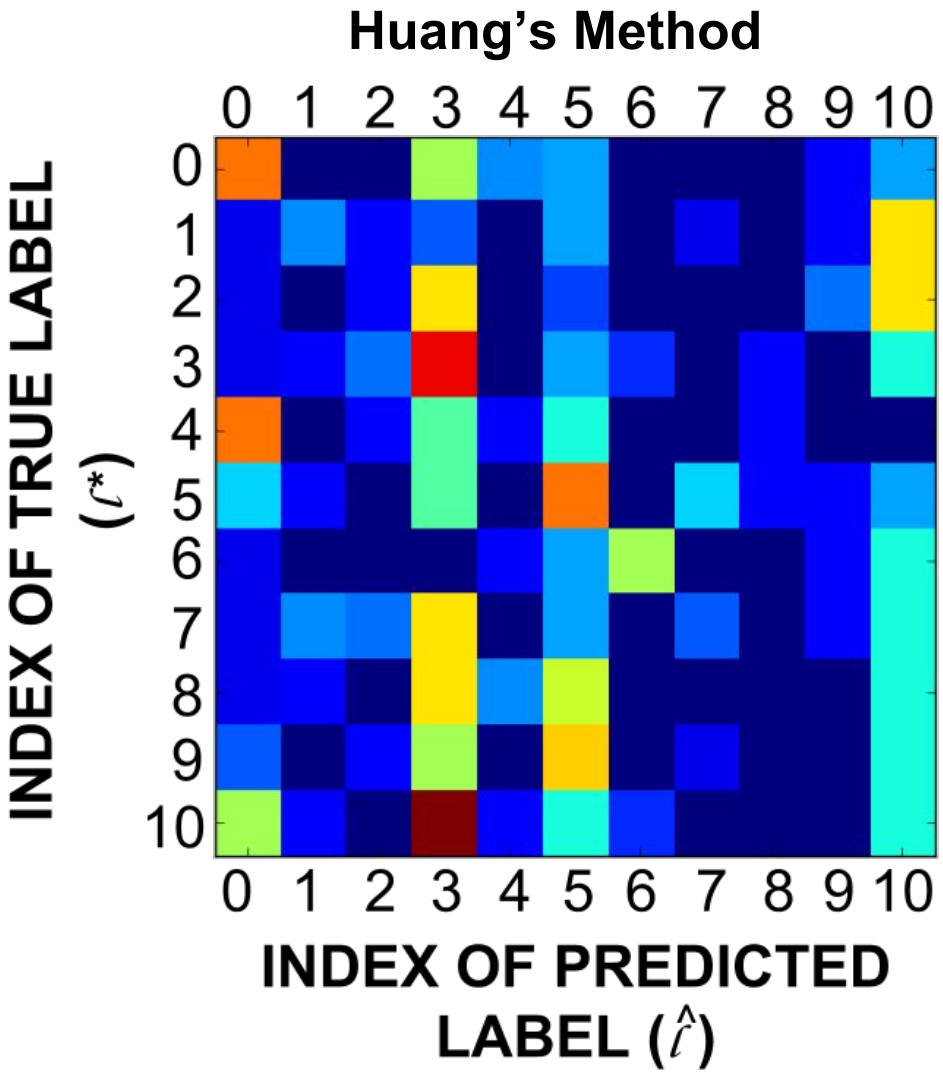
The ICU is a **natural** (unstructured) scenario

Illumination variations



Occlusions (blankets)

Huang et al. RGB-Pressure Classification [3]

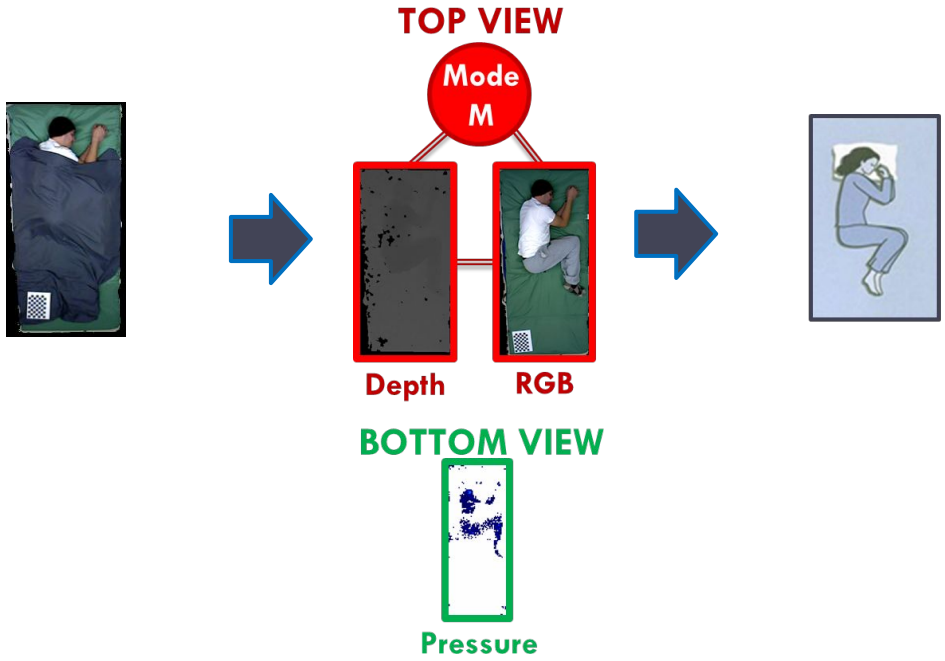


Dim and Occluded
(Pillow + Blanket)

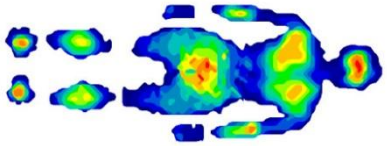
Realistic ICU Scenes

RGB-D-P to Tackle Natural ICU Scenes

- Insights from suitability tests
  - RGB-D-P can be complementary on some scenes
  - Some modalities do better than others on different scenes
- Collect multimodal data and record the environmental conditions



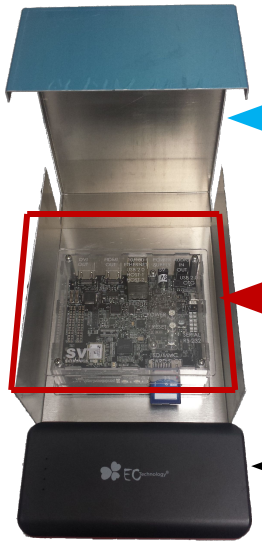
Multimodal Modules + Room Sensors



Tekscan BPMS Pressure System

Bottom View (f_P)

RGB + D Camera



Enclosure

Panda Board

Battery

Illumination Robustness (f_R and f_D)

Environmental Measurements

Modality/Scene Indicators (c)

Data: R-D-P Poses and Scenes

S0 – empty bed

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)										
Depth (D)										
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

Data: R-D-P Poses and Scenes

S0 – empty bed

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)										
Depth (D)										
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

Proposed Approach

Multimodal Data + Environmental Feedback

$$\hat{z}_k = \arg\max_{z \in Z} P(z|X_k^c, c)$$

$\hat{z}_k$, estimated pose label for the given data point

Z = {bed, log R, fetal L, …, soldier U}, pose labels

$X_k^c = \{f_R, \ldots, f_m\}_M$, input data point (set of features) → X_k

k, data point index over data from c

c, scene (e.g., bright + clear)

m, modality indicator over M

Combine to Deal with ICU Scenes

QUER

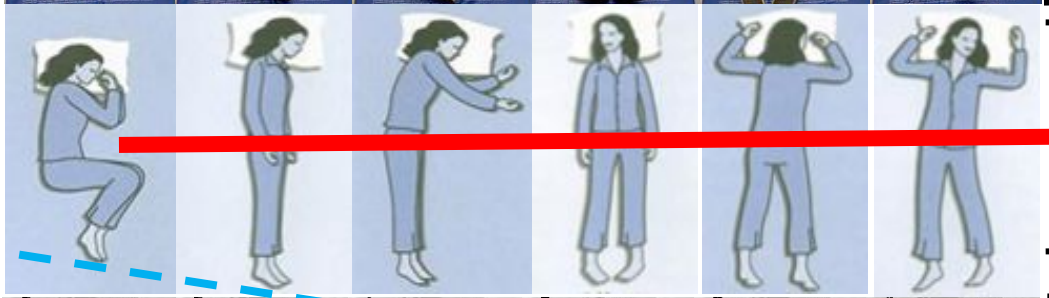
RG B +

Dept h +

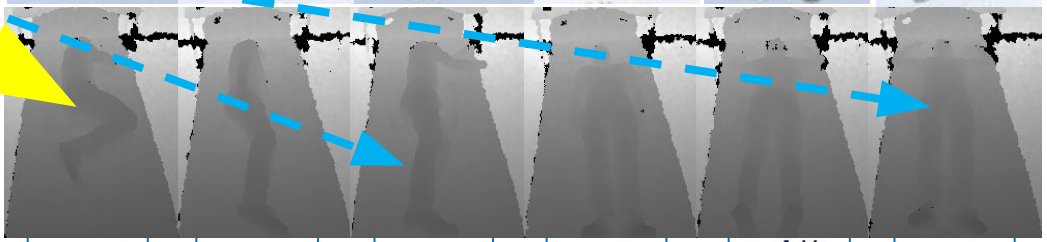
Pressu re

DICTIONARY OF POSES

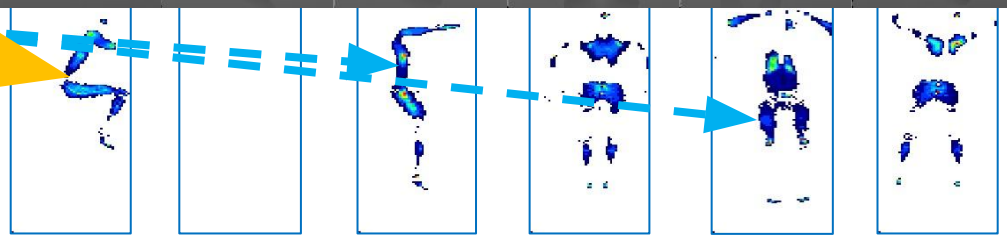
RG B



L



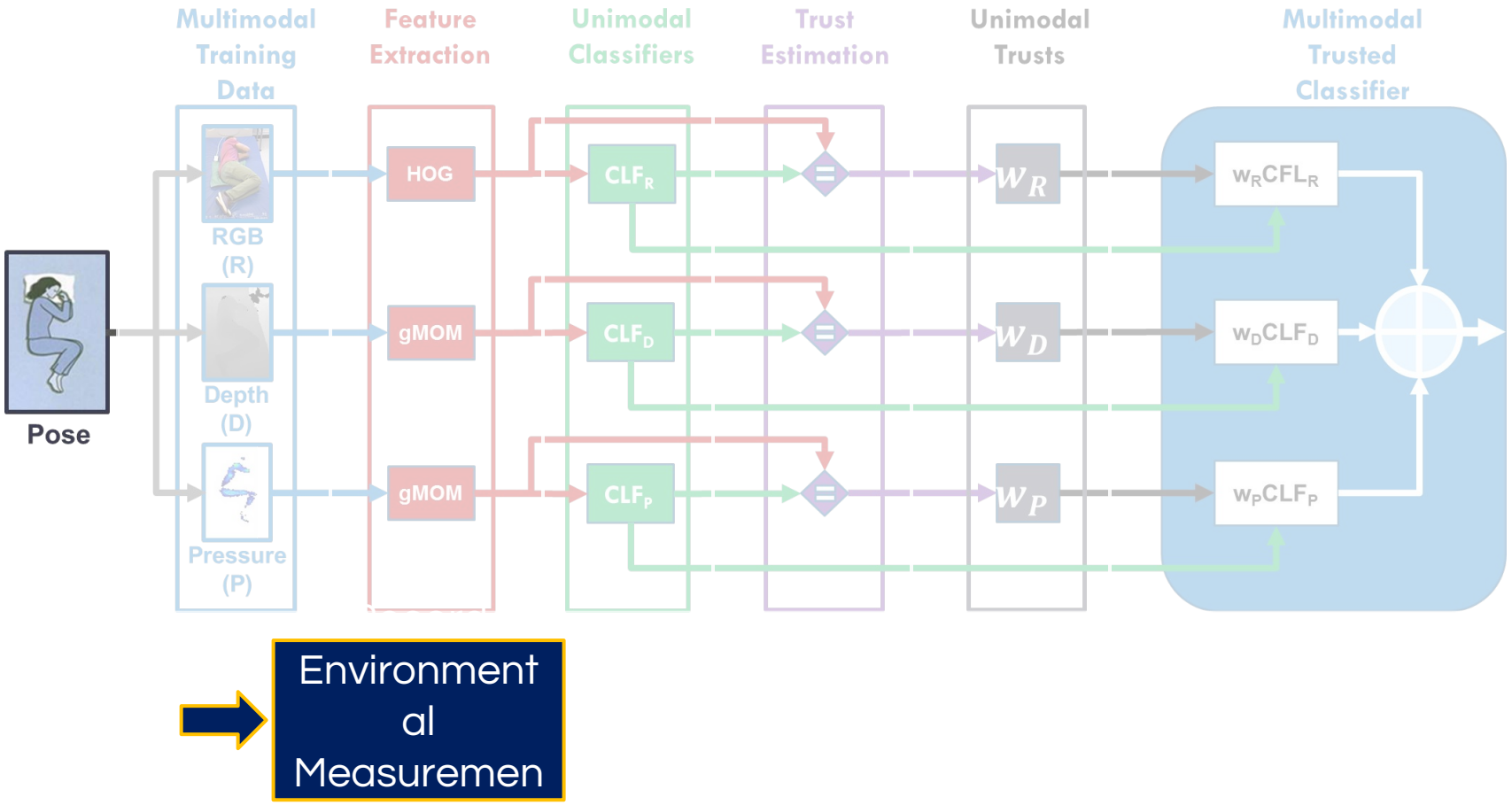
Depth



Pressur e

OUTPUT (DESIRED)

System – Multimodal Trusted Classifier



Features: Gradients(R) + Shape (D+P)

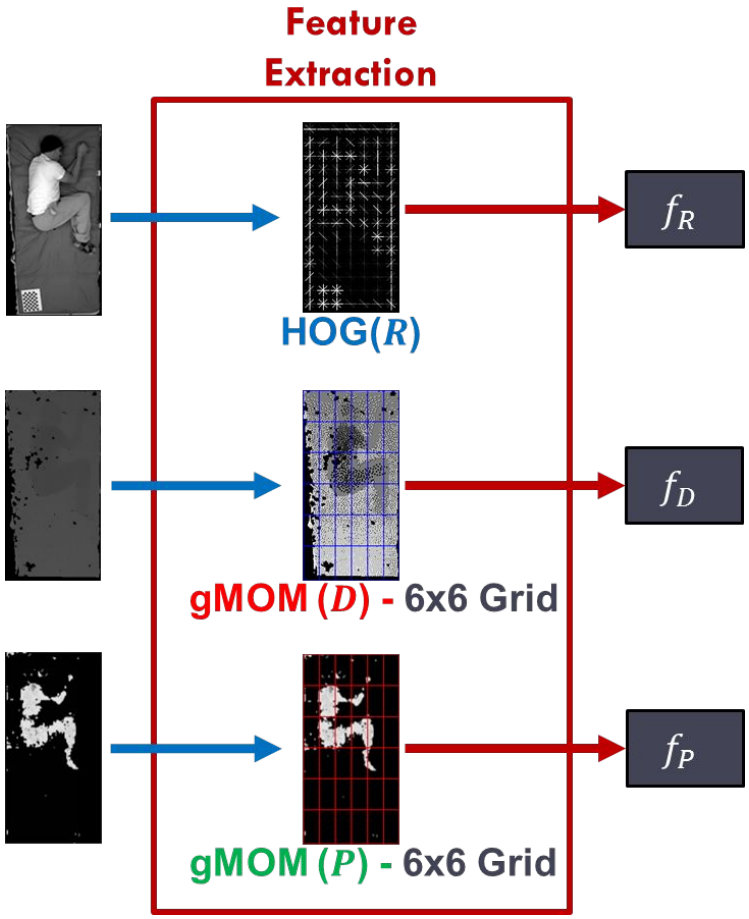
- Histogram of Oriented Gradients (HOG) [4]

  RGB: f_R– 8424 elements

- Moments up to 3rd order (gMOM)[5]

  Depth: f_D– 360 elements

  Pressure: f_P (P) – 360 elements



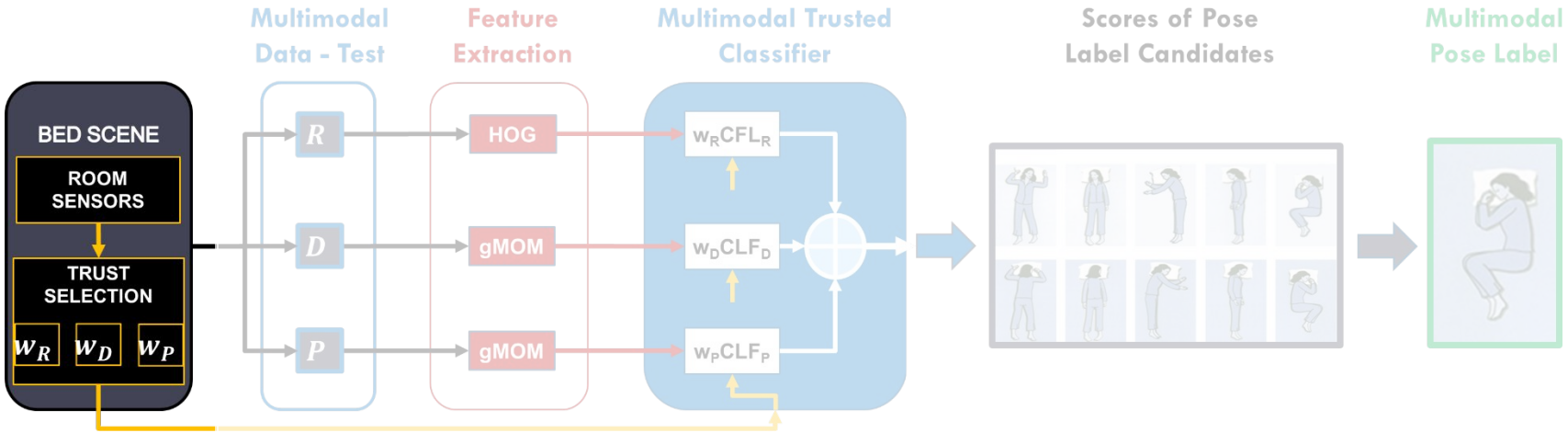
Unimodal Evaluation (Trusts Estimation)

$$\widehat{w}_m = \frac{\sum_{k=1}^{K} \vec{\mathrm{b}}[k,m]}{K} \quad \rightarrow \quad w_m = \frac{\widehat{w}_m}{\sum_m \widehat{w}_m}$$

where

- $\widehat{w}_m$, trust for modality m and scene c (omitted)
- w_m, normalized modality trust (sum is one)
- $\vec{\mathbf{b}}$, performance measure of the classifier for m and c (e.g., mean accuracy)
- k runs over the training samples K from scene c
- m modality indicator (R, D, P)

Multimodal Trusted Classifier-Testing



Environmental Measurements

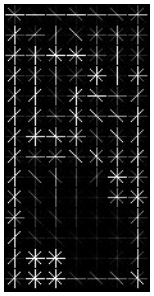
Bright + Clear	0.4	0.3	0.3
Dark + Blanket	0.1	0.3	0.6

Room sensors dictate which trust values to use for the “measured” scene conditions

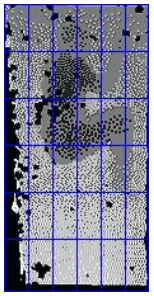
Multimodal Features

Given an $\{R, D, P\}$ data point (feature vector set) indexed by k

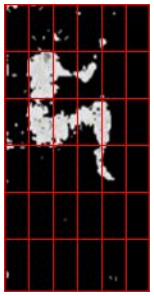
$$X_k = \{f_m\}_M = \{f_R, f_D, f_P\}_k$$
$$= \{\mathrm{HOG}(R), g\mathrm{MOM}(D), g\mathrm{MOM}(P)\}_k$$



HOG(R)



gMOM(D)



gMOM(P)

Find the label $\boldsymbol{z_{\hat{l},k}}$ indexed by $\hat{l}$ using all the features and trusts in M (i.e., f_m and w_m, $m \in M$)

Unimodal Trusted Label Scores

Unimodal trusted scores

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

where

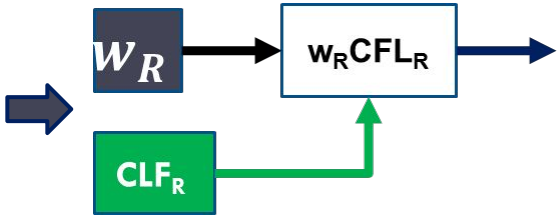
$S^c_{m,k}$, label scores from the trusted unimodal classifier

w^c_m, trust value for f_m under scene c

$\mathrm{CLF}^c_m = \{\hat{s}_{1,k}(f_m), \ldots, \hat{s}_{L,k}(f_m)\}$, vector of scores from classifier m

L , number of labels

c, scene

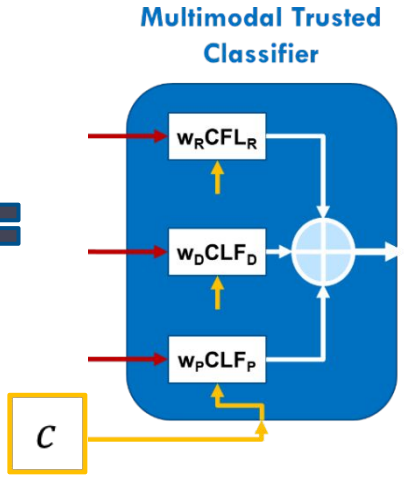


Multimodal Trusted Score Computation

- Multimodal score S_k^c from trusted unimodal scores

$$S_k^c = \sum_{m=1}^{M} S_{m,k}^c = \sum_{m=1}^{M} \left(w_m^c \{\hat{s}_{l,k}(f_m)\}_L^c \right)$$

$$\sum_{m=1}^{M} \left(w_m^c \{\hat{s}_{l,k}(f_m)\}_L^c \right) =$$



where

$\{\hat{s}_{l,k}(f_m)\}_L^c = \{\hat{s}_{1,k}(f_m), \ldots, \hat{s}_{L,k}(f_m)\}^c$, label scores from classifier m

Multimodal Trusted Score Computation

Finally, find the index $\hat{l}$ of the estimated pose label $Z_{\hat{l}}$:

$$\hat{l} = \arg\max_{l \in L}(S_k^c) = \arg\max_{l \in L} \sum_{m=1}^{M} \left(w_m^c \{ \hat{s}_{l,k}(f_m) \}_L^c \right)$$

Scores of Pose Label Candidates

Multimodal Pose Label

$$\arg\max_{l \in L}(S_k^c)$$

$$\hat{l} \rightarrow \hat{z}_k = Z_{\hat{l}}$$

Unimodal Classification Accuracy

SCENE		Support Vector Classifier				Linear Discriminant Analysis			
LIGHT	OCCLUSION	RDP	R	D	P	RDP	R	D	P
BRIGHT	**Overall**	**81**	**79**	**20**	**64**	**52**	**81**	**82**	**94**
	Clear	75	100	23	55	99	100	97	93
	Blanket	60	76	22	56	85	79	78	93
	Blanket + Pillow	86	86	24	66	80	86	78	80
	Pillow	66	86	23	56	85	80	75	80
DARK	**Overall**	**62**	**10**	**21**	**64**	**85**	**9**	**80**	**86**
	Clear	67	12	29	55	95	12	97	94
	Blanket	55	9	22	55	70	5	78	94
	Blanket + Pillow	57	9	24	56	72	5	75	80
	Pillow	56	9	25	56	80	5	75	84

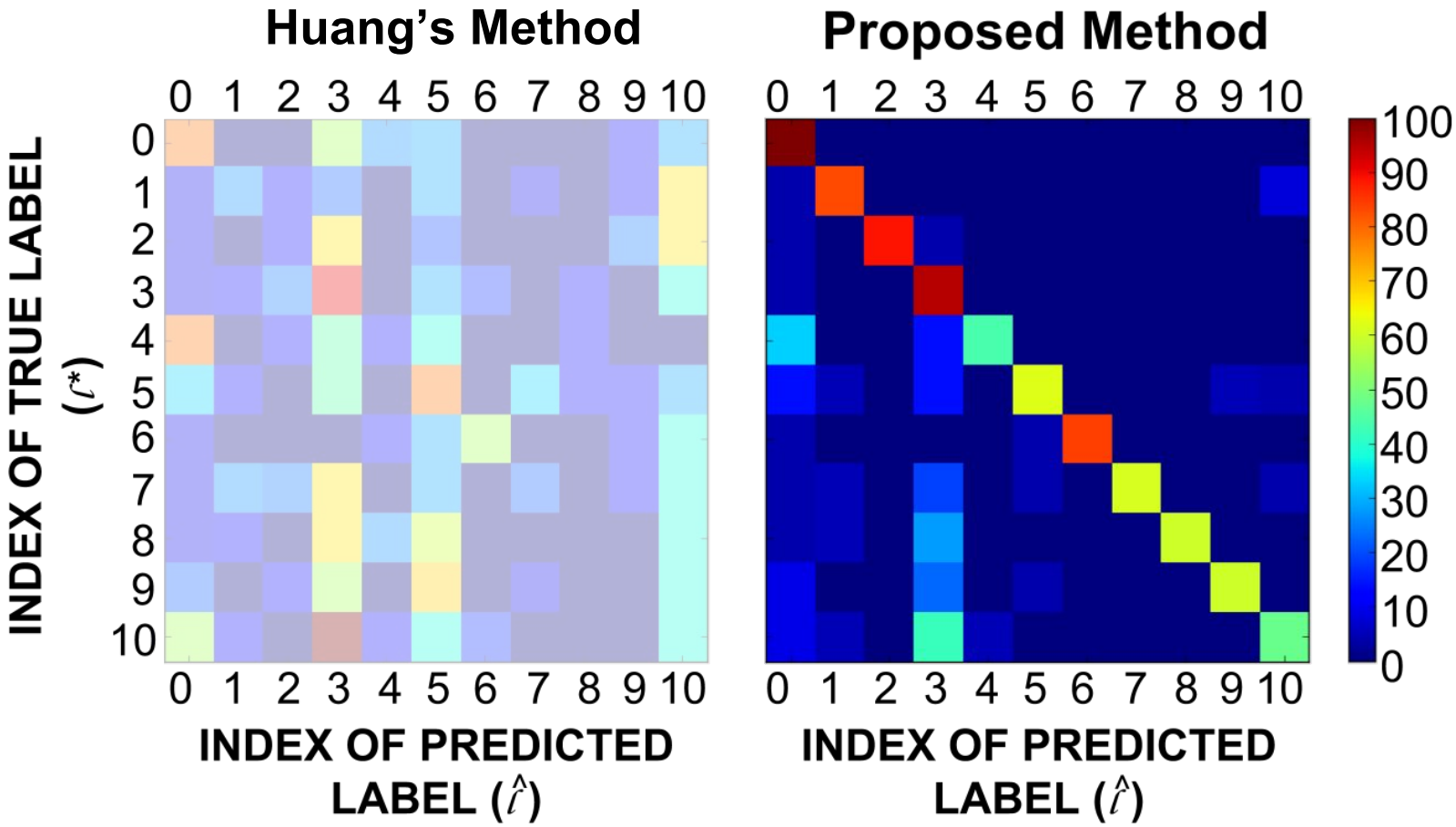
"TAKE AWAY"

NO SINGLE MODALITY NOR CONCATENATION SOLVES THE PROBLEM!

Trusted Multimodal – Classification Accuracy

SCENE		Competing		Proposed	
LIGHT	OCCLUSION	MaVL	Huang [3]	SVC	LDA
BRIGHT	**Overall**	**75**	**73**	**99**	**98**
	Clear	82	100	**100**	**100**
	Blanket	65	8	**86**	**80**
	Blanket + Pillow	54	6	**86**	**84**
	Pillow	80	58	**90**	**90**
DARK	**Overall**	**45**	**N/A**	**55**	**33**
	Clear	17	N/A	**81**	**85**
	Blanket	**20**	N/A	**20**	**19**
	Blanket + **Pillow**	**32**	N/A	18	18
	Pillow	**60**	N/A	24	22

Confusion Matrices – Dim and Occluded



Trust Adjustment w_m^* for Missing Modality n

$$w_m^* = w_m \left(1 + \frac{|w_n - w_m|}{\sum_{m=1}^{M} w_m}\right), 1 \le m \le M$$

SCENE		SVC			LDA		
LIGHT	OCCLUSION	RD \P	RP \D	DP \R	RD \P	RP \D	DP \R
BRIGHT	Overall	89	89	100	88	89	88
	Clear	100	100	100	100	100	100
	Blanket	85	90	95	80	85	92
	Blanket + Pillow	84	84	94	89	84	84
	Pillow	85	88	87	90	85	95
DARK	Overall	26	42	52	38	71	81
	Clear	54	48	73	30	74	76
	Blanket	8	22	35	23	69	77
	Blanket + Pillow	6	27	30	13	53	69
	Pillow	12	37	45	37	65	74

"TAKE AWAY"

PRESSURE IS THE MOST INFORMATIVE MODALITY

… BUT

Conclusion

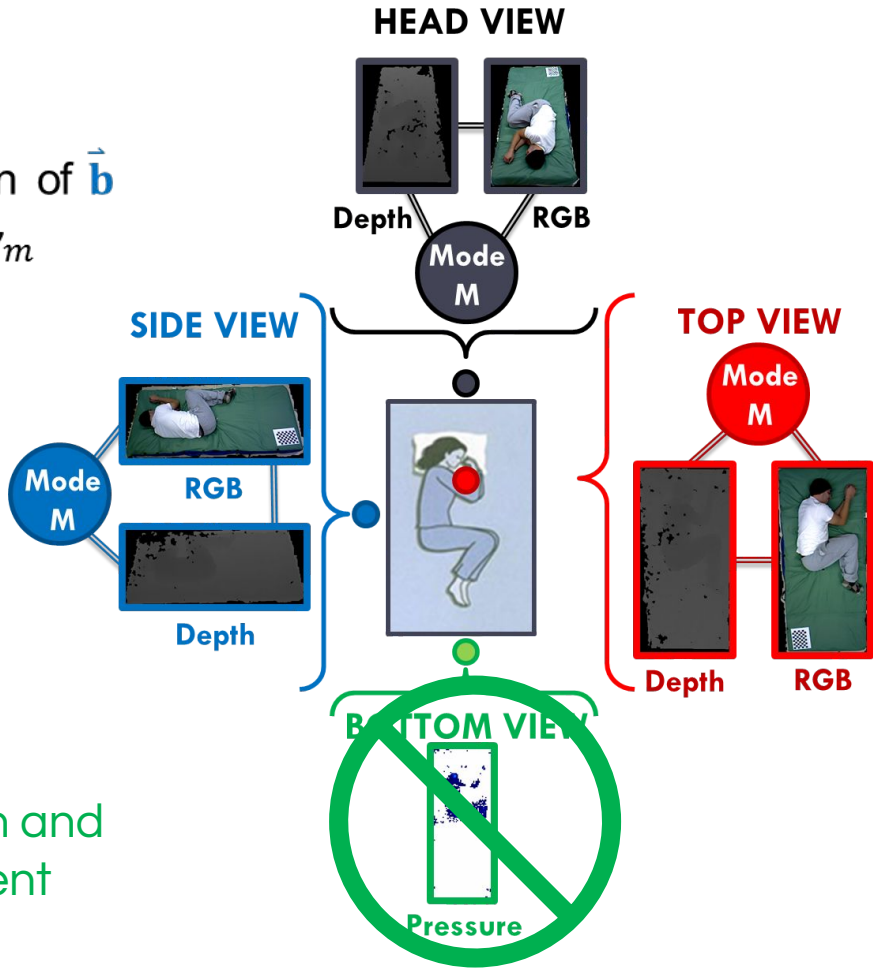
1. Evaluation of **Existing** methods and unimodal approaches
   - Shotton et al. – not suitable
   - Yang et al. – not suitable
   - Huang et al. – suitable for ideal scenarios ONLY!

2. **Trusted Multimodal** + environmental feedback is **promising**
   - RGB, Depth, and Pressure contribution based on environment
   - Improvement over existing methods
   - Robustness to illumination, occlusions, and sensor failures

3. May enable **automated** temporal **analysis** of ICU patients
   - Sleep Hygiene
   - Pressure Ulcers

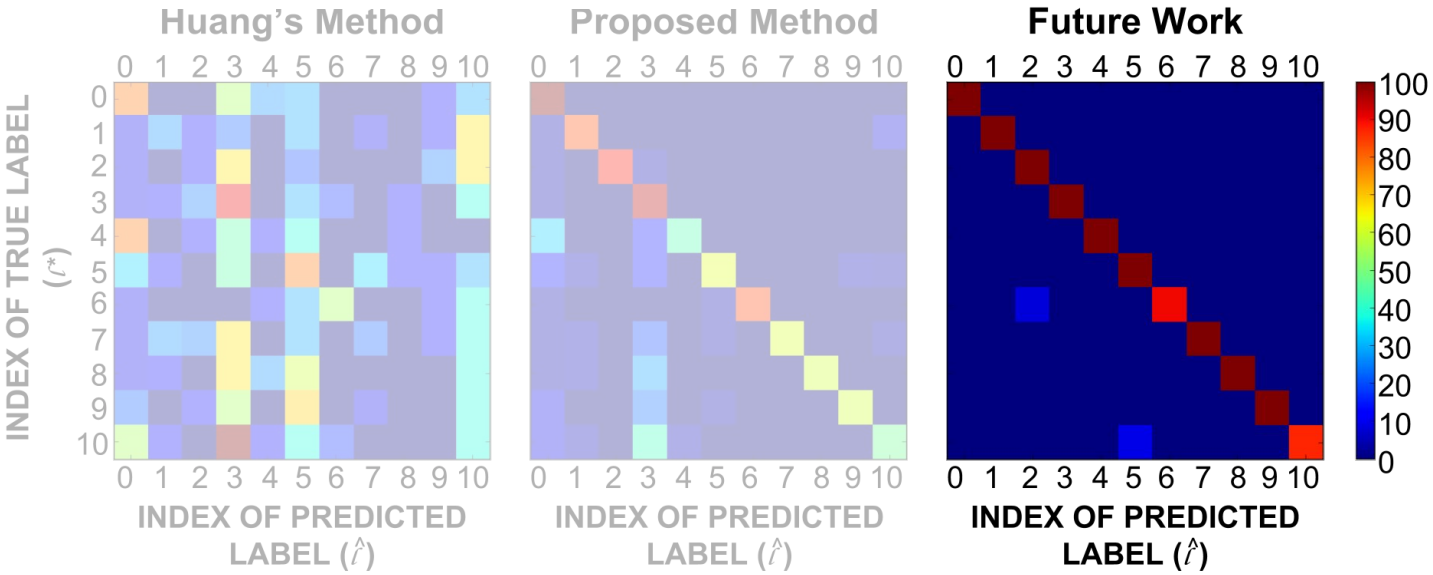
Future: Multimodal Multiview

- Objectives
  - Increase Robustness
  - Avoid Pressure Mat
  - Multimodal Construction of $\vec{\mathbf{b}}$
  - Optimize Solution for w_m



Sanitation and Deployment

Future: Multimodal Multiview (Preliminary)

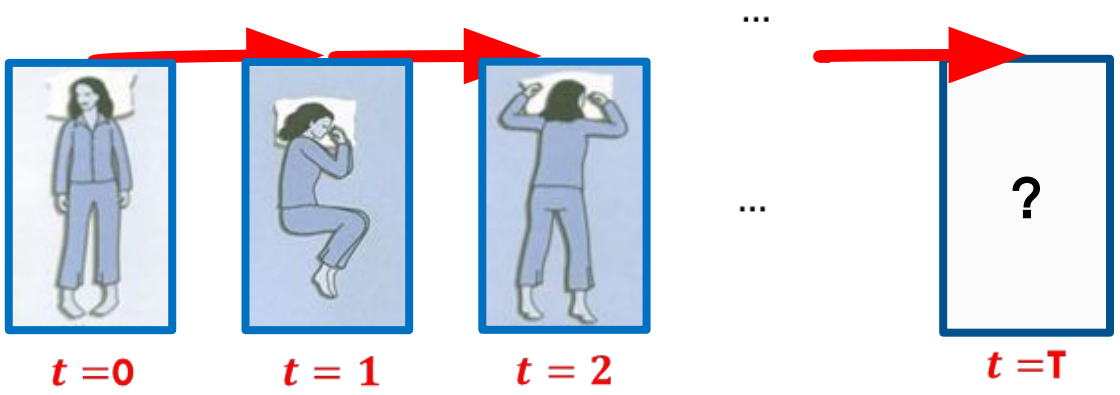


Dim and Occluded

Future: Temporal Multimodal Multiview

Temporal multimodal multiview pose analysis

- Temporal analysis of SD & DU: pose duration & transitions
- Clinical visualization of patient poses (summary & models)



References

1. Jamie Shotton, Ross Girshick, Andrew Fitzgibbon, Toby Sharp, Mat Cook, Mark Finocchio, Richard Moore, Pushmeet Kohli, Antonio Criminisi, Alex Kipman, and Andrew Blake. ”*Efficient Human Pose Estimation from Single Depth Images”.* In Decision Forests for Computer Vision and Medical Image Analysis, Springer, 2013.
2. Y. Yang, D. Ramanan. “*Articulated Human Detection with Flexible Mixtures of Parts”*. In IEEE PAMI 2014.
3. Weimin Huang, Aung Wai, Siang Foo, Jit Biswas, Chi-Chun Hsia, and Koujuch Liou. Multimodal Sleeping Posture Classification. International Conference on Pattern Recognition (ICPR) 2010.
4. Dalal, Navneet, and Bill Triggs. “*Histograms of oriented gradients for human detection”*. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE, 2005.
5. Hu, Ming-Kuei. "*Visual pattern recognition by moment invariants*." *Information Theory, IRE Transactions on* 8.2 (1962): 179-187.

THANK
YOU!
Q & A

E-MAIL: CARLOSTORRES@ECE.UCSB.EDU