A Simple Sequential Pose Recognition Model for Sleep Apnea

Ching-Wei Wang and Andrew Hunter

***Abstract*— Many existing approaches in computer vision to pose estimation make simplifications of the measurement problem, either using silhouettes or assuming knowledge of appearance or color. However, recognizing the pose of a person who is persistently under cover remains challenging. We present a real time monocular-video approach for markerless pose estimation of human body under cover without manual initialization. In order to deal with heavy occlusion, we propose a model that reinforces both feature space and model parameters by adjacent parameters and a novel search framework that aggregates detections over time to produce a more reliable hypothesis. In addition, we have introduced a novel head model, which has the combined effect of improving performance and increasing efficiency. Furthermore, we have proposed a novel representation to estimate upper leg posture using latent features. In evaluation, we demonstrate the techniques to estimate the covered body pose with various postures and obscuration levels in two environmental settings.**

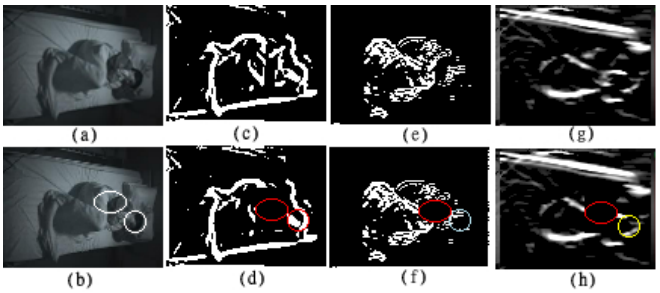


Fig. 1. (a)raw image (b)system output (c)image observation I_t^h by prewitt edge detector(d)reinforced feature space II_t^h with its hypothesis g_t (e)motion data m_t (f)reinforced motion data mm_t with its hypothesis r_t (g)auxiliary image observation J_t (h)reinforced auxiliary image observation JJ_t^h with its hypothesis q_t.

I. Introduction

Estimating human body posture is important for automatic recognition of human activities. There has been considerable work in pose recognition in recent years. Posture of subjects with well-represented appearance or silhouette can be estimated reliably in some systems. However, recognizing the pose of a person who is persistently under cover remains challenging.

In application to diagnosis of sleep activities and syndromes, the goal of our work is to estimate human poses in conditions with persistent heavy occlusion. We do not constrain the forms of the pose nor the level of occlusion where the object may persistently be partially covered, near fully covered, or uncovered, and the occlusion status may change to one another. In addition, we do not require the subject to be uncovered when he/she first appears in the scene nor do we require manual initialization.

The principal sources of difficulty in performing this task include: (a) change in appearance of the subject according to the occlusion level (b) appearance data such as skin color, head-shoulder contour, body outline and ridges of the legs being inaccessible (c) motion data being partial, irregular and obscured by the cover. Many existing approaches to pose estimation make simplifications of the measurement problem, either using motion data (e.g. [1], [4], [5], [17]) to extract silhouettes, or assuming knowledge of appearance or color (e.g. [3], [10], [13], [15]), and the subjects tend to wear close-fitting clothing (or even to be unclothed [3]) in order to extract such information more easily. These methods are too restrictive and not applicable to the problem we encounter in our field of study. Although there is some published research investigating the monitoring of partially occluded humans [6], [14], [19], [23], the methods examined do not deal with pose estimation of consistently and almost wholly occluded subjects. Wang et al. [20], [21] introduced measurement models to detect the upper body consistently under cover. Nevertheless, the method fails with unconstrained poses and when tracking movement in our experiments. In addition, the method estimates only the upper body.

We propose a real time monocular-video approach for markerless pose estimation of human body under cover without manual initialization. The outline of the method is (1) to detect head and torso by an undirected head-torso model (section 3) using the search method (section 5.3); (2) to recognize upper leg posture based on the current torso hypothesis by a novel upper leg model (section 4); (3) to validate the detected human configuration using temporal coherence (section 5.2); and after validation succeeds, (4) to track head and torso using the previous hypothesis, the linking hypotheses and the reinforced features by linking hypotheses (section 2). In evaluation (section 7), the experimental results show that the proposed model is promising to estimate the obscured body pose with various postures and obscuration levels. Furthermore, we compare the proposed method with [20] and show that the performance of the proposed method is largely improved. Regarding the computing performance, without code optimization, it takes 0.1 second to process a frame with a P4 2.4GHz CPU.

Manuscript received July 1, 2008; revised August 15, 2008

C.-W Wang is with the Centre for Visual Surveillance and Machine Perception Research, University of Lincoln, Lincoln, United Kingdom cweiwang@lincoln.ac.uk

A. Hunter is with the Centre for Visual Surveillance and Machine Perception Research, University of Lincoln, Lincoln, United Kingdom ahunter@lincoln.ac.uk

II. Reinforcement by linking parameters

Denoting X as the model parameters and Z as the image observations, we propose a model for obscured subjects by both adding auxiliary image observations $\{z_i^k\}$ for each model parameter x_i and adding relationships $\theta(x_j, \{z_i^k\})$ between the adjacent model parameters and the image observations (Details are presented in the next section). We not only reinforce the model parameter by the adjacent model parameter but also reinforce the obscured feature space in order to generate an accurate model parameter. As the features are weakly represented in our problem domain, the reinforcement by linking hypothesis applied to the obscured feature space gains more advantage than by mere selection of hypothesis. The joint posterior distribution of the proposed model is

$$P(X|Z) \propto \prod \lambda(x_i, x_j) \prod_i (\prod_k (\mathcal{L}(\{z_i^k\}|x_i) - \theta(x_j, \{z_i^k\}))) \tag{1}$$

To estimate the model parameter x_i at time t, we reinforce the image observations $\{z_i^k\}$ for x_i using the known adjacent model parameter x_j from time t-1. We zero the confidence weights of portion of features within the region $\mathcal{A}$ derived from x_j. The image observations can be in different formats, e.g. binary edge maps, motion maps, and gradient edge maps as in Figure 1 and edge box maps as in Figure 2. Implementation details are given in the next two sections.

A. Head Tracker: multiple hypotheses and reinforced features

Denoting the hypotheses of the head and torso at time t as h_t and u_t, the main image observation for the head (prewitt edge maps) as I_t^h, the auxiliary image observation for the head (horizontal oriented edge maps) as J_t^h and the motion data as m_t (see Figure 1), we obtain the head hypothesis:

$$h_t^* = \arg\max_{h_t} P(h_t|h_{t-1}, u_{t-1}, I_t^h, J_t^h, m_t) \tag{2}$$

To estimate h_t^*, we firstly produce reinforced observations II_t^h, JJ_t^h and mm_t using the adjacent model parameter at time t-1, u_{t-1}. We zero the confidence weights of features within the region $\mathcal{A}$ derived from u_{t-1}. For edge maps (II_t^h and JJ_t^h), $\mathcal{A}$ is defined as a vertically expanded area of u_{t-1} for reducing noise. Next, we sample instances within the area with high probability at the previous frame. We then use a edge clustering model(section 3.1) to search the area over II_t^h and mm_t, producing two probability maps. We select two hypotheses g_t and r_t with the highest confidence weight from the two maps respectively. Importantly, we argue that appearance features like II_t^h weight much more than motion information like mm_t because motion may be caused by the cover surface movement or hand movement. Hence, the hypothesis r_t derived from motion cannot be relied to define the state but can assist improving the hypothesis and activating inspecting different evidences. We measure the distance between g_t and r_t to confirm the precision of g_t. If $|g_t - r_t| < \alpha$, where α is the tolerable distance, we define $h_t^* = g_t$; otherwise, we produce an auxiliary hypothesis q_t using JJ_t and h_{t-1} with the appearance model and define the state h_t^* as the average of q_t, g_t and r_t.

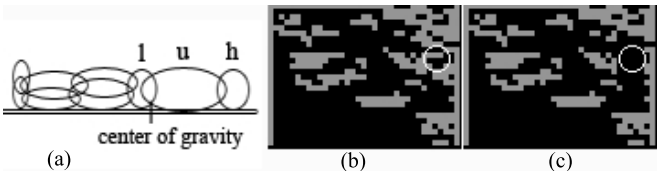


Fig. 2. (a)Biomechanics (b)edge box map with h_t^* (c)reinforced edge box map

In our method, the initial set of hypotheses h_1 and u_1 is proposed from an undirected head-torso model, which is presented in section 3.

B. Torso Tracker with Kinesiology

Motion event is a mixture of target's movement and occluding objects' movements. Motion detected in the target will be used to update the hypothesis; motion by occluding objects will not. With regard to the torso, occluding objects' movements include arm movement and cover surface movement where the subject may pull or remove the cover. To identify real torso motion from motion data, we first examine kinesiology. Considering the earth's gravitational force and biomechanics of human movement [2] in Fig 2 (a), movements of u must be accompanied with movements of h but can happen with or without l movements because the center of the gravity of the body infers that the head cannot be the fixed point to support movements of u and l. That is, denoting a head movement event as C_1, a torso movement as C_2, and a hip movement as C_3, we infer that $P(C_2) = P(C_1 \cap C_2 \cap C_3) + P(C_1 \cap C_2 \cap \sim C_3) = P(C_1 \cap C_2)$.

We estimate new torso hypothesis based on motion, a latent image observation I_t^u, previous torso hypothesis u_{t-1}, the previous adjacent model parameter h_{t-1} and the current adjacent model parameter h_t^*:

$$u_t^* = \arg\max_{u_t} P(u_t|u_{t-1}, h_{t-1}, h_t^*, m_t, I_t^u) \tag{3}$$

When motion detected within the region of the previous torso hypothesis u_{t-1} is over β percentage, it suggests a potential timing to update the hypothesis. To confirm a torso activity occurred, we check if motion occurs within the region of h_{t-1} over γ percent. If true, we then adjust the torso hypothesis based on the two types of image observations. Firstly, we compute the intersection of the motion data and a vertically expanded area $\mathcal{A}$ derived from u_{t-1} using ζ. Denoting the intersection as $\mathcal{D}$, we generate a temporary torso hypothesis using the center of D: $u_t^1 = \overline{D}$. Secondly, we produce a latent image observation, edge box maps (see section 3.2), as in Fig 2 (b) and reinforce the feature by zeroing the confidence weights of features within a region $\mathcal{B}$, where $\mathcal{B}$ is a vertically enlarged area of h_t^*. Thirdly, we input the reinforced features and the current head hypothesis h_t^* to an obscured torso measurement model [20] and generate another torso hypothesis u_t^2. Then, we compare the distance

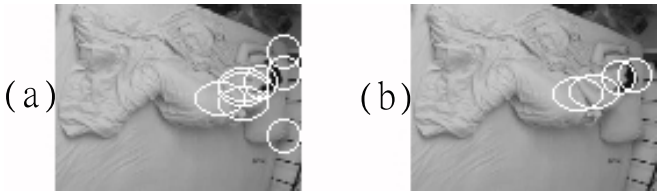


Fig. 3. (a)four hypotheses by boosting model $\{H_i\}$ (b)two hypotheses by auxiliary head model $\{h_k\}$

between u_t^1 and u_t^2 and define u_t^* as follows.(ϱ is the tolerable distance.)

$$u_t^* = \{ \begin{array}{cc} \overline{u_t^1 u_t^2} & , \quad if(|u_t^2 - u_t^1| > \varrho) \\ u_t^2 & , \quad otherwise \end{array} \tag{4}$$

III. Undirected Head-Torso Model

A. Head Detector

A hierarchical boosting model [21] is utilized to propose initial head locations. In this paper, we introduce a novel edge clustering head model both to reduce falsely detected hypotheses and to further refine the hypothesis. The model uses modified binary prewitt edge value features, which are different from the features used by the hierarchical boosting model, to improve detection and estimation by inspecting different evidences.

Given a prewitt edge image J and N estimated heads $\{H_i\}$ by the boosting model, where i = 1 to N, we apply a binary filter to the expanded area of $\{H_i\}$, producing N interested areas $\{\mathcal{Q}_i\}$ and binary features $J_{x,y}^1$. We build a scoring mechanism to count the valid points at each location, obtaining M scores within each area $\mathcal{Q}_i$. The score index is $S(j)_i = \sum J_{x,y}^1$, where $(x, y) \in \mathcal{Q}_i$ and j = 1 to M. Next, we replace head hypothesis H_i with the new hypothesis H_i^1, which obtains the highest score among the area $\mathcal{Q}_i$.

$$H_i^1 = \arg\max_j S(j)_i \tag{5}$$

Afterwards, we disqualify hypotheses with a flatness criteria ω to eliminate smooth areas. If $max_j S(j)_i < \omega$, $H_i^1 \rightarrow$ invalid. In the end, we output a list of validated head hypotheses $\{h_k\}$, where k = 1 to K and $K \leq N$.

B. Torso Detector

Given a raw image, we first apply a horizontal oriented edge detector and secondly transform the edge image into an edge box map **B** [21] as seen in Fig. 5(b), where every data point $B(i, j)$ on the edge box map is either 1 or 0. Given the edge box map **B** and the head hypotheses $\{h_k\}$, we search for a relatively smooth region with reasonable distance and angle from each head hypothesis, i.e. an area near to the head with the lowest interior edge box count, obtaining K pairs of head and torso hypotheses $\{(h_k, u_k)\}$. Given a torso hypothesis u_k, a flatness index $\lambda(u_k)$ is created as follows.

$$\lambda(u_k) = \sum_{(i,j) \in u_k} B(i, j) \tag{6}$$

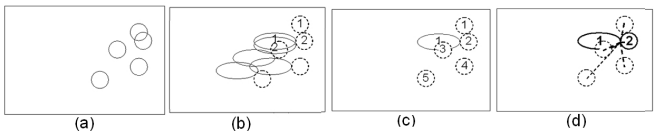


Fig. 4. (a)Head hypotheses $\{h_k\}$ by head detectors (b)Head to torso search: $\{(h_k, u_k)\}$ (c)Compare $\{u_k\}$ and choose the strongest one as u_t^*. (d)Torso to head search: output hypothesis (h_t^*, u_t^*).

Next, we compare the K torso hypotheses $\{u_k\}$ and select the strongest torso candidate u_t^* as follows.

$$\kappa = \min_k \lambda(u_k) \tag{7}$$

if $\kappa > 1$,

$$\{u_l\} = \arg\min_k \lambda(u_k) \tag{8}$$

$$u_t^* = \overline{\{u_l\}} \tag{9}$$

otherwise,

$$u_t^* = \arg\min_k \lambda(u_k) \tag{10}$$

C. Torso to Head Backward Selector

Obtaining the strongest torso candidate u_t^*, a torso-to-head backward selector is developed to choose the strongest head hypothesis. We compute the distance between the joint location o of the torso u_t^* to the head and the centers $\{c_k\}$ of the head hypotheses $\{h_k\}$. We then choose the head hypothesis h_t^* with the closest distance to o as the radius r_k of the head h_k as follows.

$$p = \arg\min_k(||o - c_k| - r_k|) \tag{11}$$

$$h_t^* = h_p \tag{12}$$

Fig 4 illustrates an example of the undirected head and torso model, showing that the backward voting function chooses a head particle in a related reasonable location, and hence h_3 will not be selected. Also, the resulting head-torso pair (h_2, u_1) may not be identical to the original head-to-torso pairs, i.e. (h_1, u_1) or (h_2, u_2).

IV. Upper-legs Pose Estimation

Existing approaches for locating legs use cues like silhouettes, ridges, color blobs, parallel edges, or cone / rectangle shape edge pixels and are under assumption of these features are well represented. However, such an assumption is not applicable for our problem domain. Here, we introduce a novel representation of legs for obscured upper-legs pose recognition. The representation model contains 12 features $\{e_k\}$ to represent the sum of edge boxes in individual subparts $\{L_k\}$ in a given edge box map, where k = 1 to 12.

To test the novel representation, we manually collect 26 images to produce training dataset for constructing 10 different pose templates. Furthermore, instead of collecting images for all poses, we collect images for template T_3, T_5, T_8 and T_9 and use mirror projection theory to make training data for symmetric templates T_4, T_6, T_7 and T_{10}. In training,

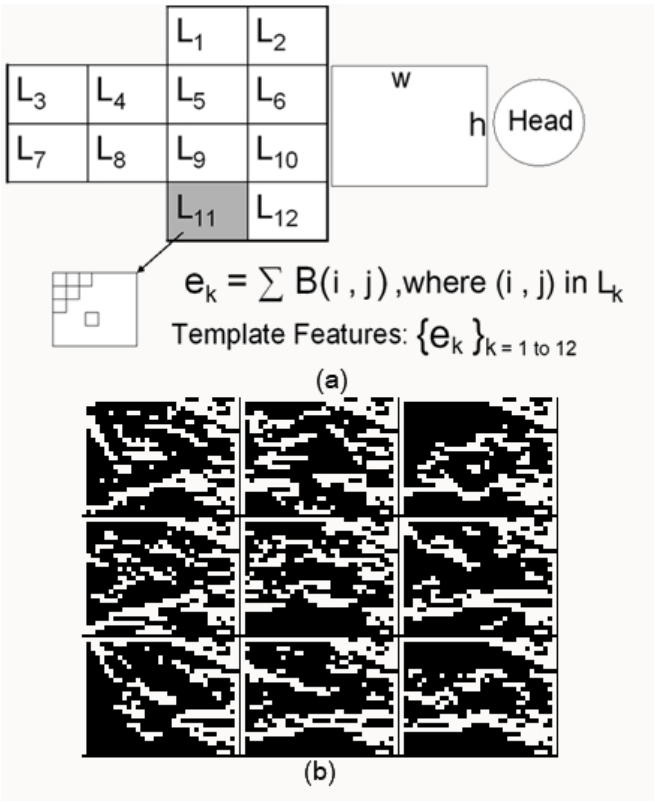


Fig. 5. (a)Representation of upper-legs pose model (b)Some edge box maps.

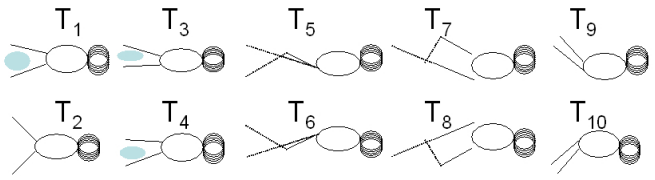


Fig. 6. 10 upper-legs pose templates.

we use a variant of boosting [22] as the learning method for generating a number of classifiers, and the classifiers are built into a binary tree structure. In testing, we apply the model to a new dataset, which includes a number of poses and movements. Interestingly, we discover that although the positions of abstract edge boxes in subareas are disused by such representation where existing approaches tend to maintain local location information, the proposed model is capable to estimate the legs poses in the experimental results.

V. Search Method

As the particle filter has been popularly adopted for articulated human body analysis [3], [8], [12], initially we tested this algorithm as the search framework, but we found that the method is not suitable for obscured subjects (see next section). Hence, we develop a search-before-detect framework, which overcomes the problems encountered when using particle filter.

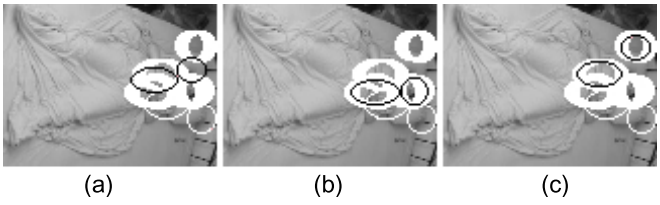


Fig. 7. Problems of particle filter(single scan based detection): (a)Mean (b)Mode t_i(c)Mode t_j.

A. Analysis of Particle Filter

Particle filters can be used in classical tracking or in Track Before Detect (TBD) [3], in which a subject is considered detected when the target likelihood by the tracking process exceeds a threshold. Considering the sampling aspect, TBD integrates the information over time, and detection is based on power/energy that has been integrated over time by multiple scan based sampling. However, in the detection aspect the criterion is still single scan based. Although the samples for defining current system state are derived from the previous samples, in declaring detection, the computation depends on testing whether the likelihood of the weighted mean or the mode over current sample states exceeds a threshold, which is based on the current spatial information. In our problem domain, the subject is consistently and heavily occluded, and we observe that spatial features are unstable and inclined to be inadequate and distorted by strong noise. Thus, we argue that the search method cannot be single scan based either in sampling or in detection. Figure 7 illustrates problems of single scan based detection. With strong noise, the mean position is erroneous as in Fig 7(a), and the mode position becomes unreliable as in Fig 7(c). In addition, particle filter utilizes motion to process stochastic diffusion. However, in our problem domain, motion tends to be partial, noisy and infrequent, and thus we cannot rely on motion for pattern construction.

B. Temporal Coherence

Spatio-temporal approaches have been shown to be advantageous to overcome self-occlusion and image noise in recent research [9], [11], [16]. These methods exploit temporal coherency on feature points such as motion and the silhouette. In contrast, we exploit the property of temporal coherence on system states rather than features, constructing temporally coherent patterns. The distinction between applying temporal coherence to features and system states is that using temporal coherence in feature development risks transition errors from observed data to estimated states, whereas applying temporal coherence property directly on system states avoids such risks.

On each time-step, t, the algorithm evaluates multiple hypotheses for the head position, $\{h_i\}$, and torso position, $\{u_j\}$, and a single hypothesis for the upper legs position, g_t. The strongest hypothesis $H_t = (h_t, u_t, g_t)$ is identified (as described in section 3 and 4). If hypothesis H_t yields a reasonably consistent position over a sufficiently long time

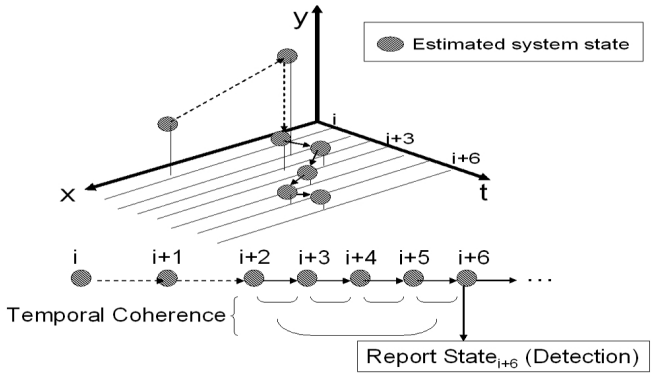


Fig. 8. Temporal coherence on patterns.

period, then detection with H_t is declared. We use a threshold for head displacement, ρ, and a minimum stable period, τ.

Let k be a count of consecutive stable iterations; initialize k=0. On each iteration, t, if $|h_t - h_{t-1}| < \rho$, increment k by 1; otherwise, set k=0. Declare detection using H_t when $k > \tau$. An illustration is given in Figure 8. ($\rho = 0.3, \tau = 3$ are used in our experiments.)

In this work, we use h_t for a temporally coherent relationship to test the performance of applying coherence of states, and intend to explore more complicated models in the future work, integrating the states of all parts.

C. Search Method: Greedy Search + Jumping

As the availability of spatial features available can be variable and the true hypothesis can be temporarily hidden, we conduct the search in an independent greedy manner every frame both to collect as much information as possible and to avoid filtering out the true hypothesis due to sampling. In order to balance the computational cost, we construct a jumping mechanism, which forces the greedy search to skip neighbor rows below a detected point. The jumping function cooperates with an auxiliary computationally efficient measurement model (section 3.1) to find an optimal position in the local area. Given an $M \times N$ region of interest and the current position (x, y), the next search position (a, b) can be formulated as follows. (k =3 is used based on our experiments).

$$(a,b) = \{ \begin{array}{cc} (x+1, y+k) & if(x,y) \in \{h_i\} \\ (x+1, y) & if(x,y) \notin \{h_i\} \wedge (x+1) < M \\ (0, y+1) & otherwise \end{array} \quad (13)$$

Compared with post-processing hypotheses after search using the weighted mean, the mode or clustering [23], the jumping function avoids a merging process and saves effort on both searching and post-processing.

VI. Experimental Results

We evaluate the accuracy of our pose estimation algorithm on three video sequences, containing 2653 frames. The video was acquired with the frame rate 15fps and a resolution of 320*240, using a SONY infrared camcorder (DCR-HC-30E). Fig 9 (a) shows system outputs on a fully covered subject sleeping on different side of the torso and presenting a number of postures; Fig 9 (b) contains different levels of occlusion, including near-completely occluded and the half-uncovered (Please note that more test results can be found in the accompanied video "Seq1.wmv" and "Seq2.wmv"). The same parameters were used in testing, which demonstrate the obscured human pose estimator's ability to follow a wide range of human poses and obscuration. For a quantitative evaluation, we systematically sample frame per 0.3 second from the video clips, obtaining 586 frames. The sub-selected frames are then manually marked to produce a reference standard, and the output of the system is compared to the reference standard. We evaluated our obscured human pose estimation model by calculating how often individual parts are correctly localized, for which we define a precision rate P as the percentage of the overlapping area between the reference standard and estimated output and a recognition rate R(P) as the percentage of the frames to have precision rate greater than given P value. The recognition rates of the head, torso and upper legs pose with precision rate P = 0.7 for the proposed method and the referenced method [20] are presented in Table 1.

TABLE I
Recognition Rates

R(0.7)	Head	Torso	Upper legs
Proposed Method	0.98	0.99	0.74
Referenced Method [20]	0.5	0.92	N/A

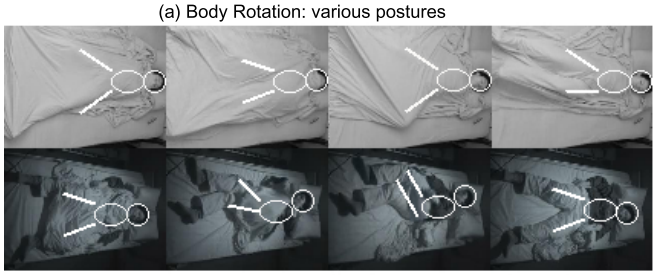
The experimental results show that our algorithm is able to locate the body parts and estimate their pose well even though they are heavily occluded and appear in different postures. Compared to the reference work, the performance is largely improved. Regarding the computing performance, without code optimization, it takes 0.1 second to process a frame with a P4 2.4GHz CPU.

VII. Conclusion

We have presented an efficient monocular-video approach for markerless pose estimation from a consistently fully or partially covered human without manual initialization. In order to deal with heavy occlusion, we have further proposed a model that reinforces both feature space and model parameters by adjacent parameters and a novel search framework that aggregates detections over time to produce a more reliable hypothesis. In addition, we have introduced a novel head model, which has the combined effect of improving performance and increasing efficiency. Furthermore, we have proposed a novel representation to estimate upper leg posture using latent features.

We will continuously develop and enhance methods to locate the rest of the body parts and to recognize human activities during sleep. Amounts of experiments will be conducted using patients and volunteers in the Sleep Lab at the Lincoln County Hospital to determine performance

(a) Body Rotation: various postures



(b) Near fully covered / Partially covered / Uncovered

Fig. 9. System Outputs

on a large scale and to generalize across the variations in human sizes, shapes, and behavior. Furthermore, we will develop a system to diagnose movements characteristic of sleep disturbances.

VIII. ACKNOWLEDGMENTS

This research is jointly supported by United Lincolnshire Hospitals NHS Trust (ULH collaborative Research Grant) and University of Lincoln for the PhD scholarship of C.-W, Wang. The authors gratefully acknowledge the support of Dr. Neil Gravill and Dr. Simon Matusiewicz for their valuable comments in obstructive sleep apnoea, Dr. Chris Hacking for his generous help in clinical engineering and the reviewers for their comments on the paper.

REFERENCES

[1] Agarwal, A., , Triggs, B.: 3D Human Pose from Silhouettes by Relevance Vector Regression. Proceedings of Conference on Computer Vision and Pattern Recognition **2** (2004) 882–888

[2] Carr, G.: Mechanics of sport: a practitioner's guide. Human Kinetics (1997).

[3] Deutscher, J., ,Reid, I.: Articulated Body Motion Capture by Stochastic Search. International Journal of Computer Vision **2** (2005) 185–205

[4] Elgammal, A., , Lee, C.S.: Inferring 3D body pose from silhouettes using activity manifold learning. Proceedings of the Conference on Computer Vision and Pattern Recognition **2** (2004) 681–688

[5] Guo, F., , Qian,G.: Learning and Inference of 3D Human Poses from Gaussian Mixture Modeled Silhouettes. Proceedings of the international Conference on Pattern Recognition **2** (2006) 43–47

[6] Hoey, J.: Tracking using Flocks of Features, with Application to Assisted Handwashing. Proceedings of British Machine Vision Conference **1** (2006) 367–376

[7] Hua, G., ,Yang, M.-H., Wu, Y.: Learning to Estimate Human Pose with Data Driven Belief Propagation. Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition **2** (2005) 747–754

[8] Jenkins, O.C., , Gonzlez, G., , Loper, M.: Tracking human motion and actions for interactive robots. Proceeding of Human-Robot interaction **25** (2007) 365–372

[9] Lan, X., ,Huttenlocher, D.: A unified spatio-temporal articulated model for tracking. Proceedings of Computer Vision and Pattern Recognition **1** (2004) 722–729

[10] Lee, M.W., ,Cohen, I.: A Model-Based Approach for Estimating Human 3D Poses in Static Images. IEEE Transations on Pattern Analysis and Machine Intelligence **6** (2006) 905–916

[11] Li, B., , Meng, Q., , Holstein, H.: Articulated motion reconstruction from feature points. Pattern Recognition **41** (2008) 418–431

[12] McKenna, S.J., , Nait-Charif, H.: Tracking human motion using auxiliary particle filters and iterated likelihood weighting. International Journal of Computer Vision **25** (2007) 852–862

[13] Ramanan, D., ,Forsyth, D.A., ,Zisserman, A.: Tracking People by Learning their Appearance. IEEE Transactions on Pattern Analysis and Machine Intelligence **29** (2007) 65–81

[14] Ramanan, D., ,Forsyth, D.A.: Finding and tracking people from the bottom up. Proceedings of Computer Vision and Pattern Recognition **2** (2003) 467–474

[15] Ren, X., ,Berg, A.C., ,Malik, J.: Recovering Human Body Configurations Using Pairwise Constraints between Parts. Proceedings of the international Conference on Computer Vision **1** (2003) 824–831

[16] Sigal, L., , Bhatia, S., Roth, S., Black, M., , Isard, M.: Tracking loose-limbed people. Proceedings of the Computer Vision and Pattern Recognition **1** (2004) 421–428

[17] Sminchisescu, C., , Kanaujia, A., Li, Z., Metaxas, D.: Discriminative Density Propagation for 3D Human Motion Estimation. Proceedings of the Computer Vision and Pattern Recognition **1** (2005) 390–397

[18] Stone, L.D., , Corwin, T.L., Barlow, C.A.: Bayesian Multiple Target Tracking. Artech House (1999)

[19] Tobias, J., , Geert, C., , Rik, F., , Van, G.L.: Analysis of Human Locomotion based on Partial Measurements. Proceedings of IEEE Workshop on Motion and Video Computing **2** (2005) 248–253

[20] Wang, C.W., ,Hunter, A.: A Novel Approach to Detect the Obscured Upper Body in application to Diagnosis of Obstructive Sleep Apnea. IAENG International Journal of Computer Science **35** (2008) 110–118

[21] Wang, C.W., , Ahmed, A., ,Hunter, A.: Locating the Upper Body of Covered Humans in application to Diagnosis of Obstructive Sleep Apnea. Proceedings of World Congress on Engineering **2** (2007) 662–667

[22] Wang, C.W.: New Ensemble Machine Learning Method for Classification and Prediction on Gene Expression Data. Proceedings of IEEE EMBS Annual International Conference (2006) 3478–3481

[23] Wu, B., , Nevatia, R.: Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors. International Journal of Computer Vision **2** (2007) 247–266

[24] Wu, Y., ,Yu, T.: A Field Model for Human Detection and Tracking. IEEE Transations on Pattern Analysis and Machine Intelligence **5** (2006) 753–765