A Real-Time Human Stress Monitoring System
Using Dynamic Bayesian Network

Wenhui Liao, Weihong Zhang, Zhiwei Zhu and Qiang Ji
{liaow, zhangw9, zhuz, jiq}@rpi.edu
Department of Electrical, Computer and Systems Engineering
Rensselaer Polytechnic Institute, Troy, NY 12180-3590, USA

Abstract
We present a real time non-invasive system that infers user stress level from evidences of different modalities. The evidences include physical appearance (facial expression, eye movements, and head movements) extracted from video via visual sensors, physiological conditions collected from an emotional mouse, behavioral data from user interaction activities with the computer, and performance measures. We provide a Dynamic Bayesian Network (DBN) framework to model the user stress and these evidences. We describe the computer vision techniques we used to extract the visual evidences, the DBN model for modeling stress and the associated factors, and the active sensing strategy to collect the most informative evidences for efficient stress inference. Our experiments show that the inferred user stress level by our system is consistent with that predicted by psychological theories.

1 Introduction

Human stress is a state of tension that is created when a person responds to the demands and pressures that arise from work, family and other external sources, as well as those that are internally generated from self imposed demands, obligations and self-criticism. Due to the adverse effects of stress in our daily life, it is important to monitor such an unhealthy state in a timely manner and treat it properly.

The causes and manifesting features of human stress have been extensively investigated in psychology, computer vision, physiology, behavioral science, ergonomics and human factor engineering. In spite of the findings from diverse disciplines, it is still a notoriously challenging task to develop a practical human stress monitoring system. In this paper, we address two impediments on the road. First, we propose a dynamic Bayesian network (DBN), a plausible framework for reasoning under uncertainty, to judiciously model the causes of stress and its various manifesting evidences. Since human stress is volatile, dynamic, uncertainty-encompassing and user-dependent, it is inadequate to use evidences from one single modality in stress modeling. Some physical symptoms such as fast heart rate and rapid breathing are not unique to stress. In order to infer stress accurately, other physiological or behavioral evidences must be used to complement the inference. This necessitates evidence synthesis from multiple modalities. A DBN is an ideal tool to integrate these evidences. It is capable of addressing uncertainty in stress modeling, integrating the causes of stress, representing probabilistic relations among causes and evidences, modeling the dynamics in stress development and providing efficient inference solutions. Specifically, our non-invasive real-time stress monitoring system integrates the causes of stress and the evidences from four modalities – physical appearance (facial expression, eye movements, and head movements) from visual sensors, physiological features from an emotional mouse, behavioral and performance evidences. From a technical point of view, we also integrate computer vision techniques to extract features from the real-time videos, active sensing techniques to choose optimal evidence subsets, and parameter learning techniques to refine a BN to fit with individual users. Our experiments show that the system performance is improved by both combining individual evidences and exploiting the active sensing technique.

Second, we bring research results from psychology in validating a stress monitoring system. One fundamental difficulty in validating a system is lack of ground-truth stress. Some experiments have shown that even user self-reports are erroneous and unreliable. Fortunately, several psychological theories find that in a task-specific environment, user stress level varies with the workload imposed to the user and the job control of the assigned tasks [8, 4]. One innovation in our system validation is to compare the inferred stress level against job characteristics, instead of the (unavailable) ground truth human stress.
2 Related Work

Human stress has been studied in a number of diverse disciplines. Psychologists define emotions, in particular stress, as valenced (positive or negative) reactions to situations consisting of events, actors, and objects [11]. Computer scientists find that facial expressions have a systematic, coherent, and meaningful structure that can be mapped to affective states [2, 3]. Physiologists demonstrate that high stress level is related to the symptoms of faster heart beat, rapid breathing, increased sweating, cool skin, cold hands and feet, feelings of nausea, tense muscles and alike. Behavioral scientists show that high stress level promotes negative thinking, damages self-confidence, narrows attention, disrupts focus and concentration and makes it difficult to cope with distractions. Ergonomic studies indicate the Inverted-U relationship exists between stress and performance of a task [4, 1].

Researchers have developed inference approaches or pragmatic systems to recognize user stress level. The approaches or systems differ from each other in either the evidence modalities, or inference techniques, or both. Physiological measures (like electromyogram, electrocardiogram, respiration, and skin conductance) are exploited to detect stress in a car driver [5, 13]. One more example using physiological evidences is a skin temperature measuring system, developed for non-contact stress evaluation [9]. In [14], the sympathetic and parasympathetic activities of the heart of a human are used to determine human stress level via wavelet decomposition and fuzzy logic techniques. [15] combines several physiological signals and visual features (eye closure, head movement) to monitor driver drowsiness and stress in a driver simulator. Our system differs from the cited ones in that it uses evidences from multiple modalities and employs a DBN approach. It is the first one that synthesizes evidences from four modalities in user stress modelling. The effectiveness of multiple-modality information fusion is demonstrated by the increasing accuracy of inferred stress with the number of source evidences.

3 System Overview

The experimental environment is shown in Figure 1. In experiments, a user sits in front of a computer screen and responds to the tasks presented in the screen. Three cameras are used to capture the user real-time videos, which will be used to extract visual evidences. Meanwhile an emotional mouse (see the right chart of Figure 1) is used to collect physiological evidences. The hardware specification can be found in [6]. In order to alter user stress level, the environment is designed to be task-specific: the user is required to respond to asynchronously generated two types of tasks: a math task is about an addition/subtraction arithmetic of two two-digit integers, and an audio task is about the alphabetic precedence of two consecutively presented letters. For a math task, the user has to indicate his choice on the correctness of the arithmetic operations; for an audio task, the user has to indicate the alphabetic precedence between the current letter and the previous one.

![Figure 1. The experimental environment](image)

The stress monitoring system is presented in Figure 2. It consists of four main components. First, visual, physiological, behavioral and performance evidences are extracted from their sources. Second, a DBN is customized to individual users with machine learning techniques. Third, active sensing technologies are exploited to select a subset of most informative evidences in revealing stress. Finally, a DBN inference engine produces user stress level, the output of the system. The system then goes to the next step. The components will be detailed in the rest of the paper.

![Figure 2. The conceptual components of the stress monitoring system](image)

3.1 Feature Extraction

We discuss how to extract user evidences from four-modality sources, namely visual, physiological, behavioral and performance evidences.
3.1.1 Visual Features

We extracted nine visual evidences from the real time video. They are Blinking Frequency (BF), Average Eye Closure Speed (AECS), Percentage of Saccadic Eye Movement (PerSac), Gaze Spatial Distribution (GazeDis), Percentage of Large Pupil Dilation (PerLPD), Pupil Ratio Variation (PRV), Head Movement, Mouth Openness and Eyebrow Movement. The video images are captured by the visual sensor shown in Figure 3. The sensor is a suite of three cameras: one wide-angle camera focusing on the face and two narrow-angle cameras focusing on the eyes.

![Figure 3. The visual sensor](image)

The entire extraction procedure is divided into four relatively separate components – eye detection and tracking (extracting BF and AECS), gaze estimation (extracting PerSac, GazeDis, PerLPD and PRV), facial feature tracking (extracting Mouth Openness and Eyebrow Movement) and face pose estimation (extracting Head Movement). Visual feature extraction starts with eye detection and tracking, which serves as the basis for subsequent components. A robust eye detection and tracking approach is developed via the combination of the appearance based mean-shift tracking technique and bright pupil effect under Infrared light sources [6]. Thanks to this combination, the eyes can be tracked under various face orientations and variable lighting conditions. Even though the eyes are completely closed or partially occluded due to oblique face orientations, our eye tracker can still track them accurately. After tracking the eyes successfully, the eyelid movement can be subsequently monitored and the relevant eyelid movement parameters can be computed accurately.

In eye gaze estimation, we employed a computational dynamic head compensation model [18]. The model can automatically update the gaze mapping function to accommodate the 3D head position changes when the head moves. Consequently, the gaze tracking technique allows free head movements in front of the camera but still achieving very high gaze accuracy; meanwhile, the technique reduces the number of gaze calibration procedure to one time for each user. After tracking the eye gaze successfully, the gaze movement can be monitored and the relevant gaze parameters can be computed accurately.

To analyze the facial expressions, twenty-seven facial features around eyes and mouth are selected for tracking [6]. Each facial feature is represented by a set of multi-scale and multi-orientation Gabor wavelets. At each frame, based on the possible region for each facial feature as constrained by the detected pupils, the initial positions of each facial feature can be located via Gabor wavelets matching. In order to achieve a robust and accurate detection, the initial feature positions are then refined by a flexible global shape model based on ASM (Active Shape Model) that constrains the spatial relationships between the detected facial features. To account for face poses, the global face shape model, which is learned under frontal faces, is dynamically deformed via the previously estimated face pose parameters to accommodate the face geometry changes. Thus, the correct global spatial constraints can always be imposed over the facial features so that they can be still tracked robustly under varying face orientations. Moreover, we also introduce a multi-state face shape model in order to handle different facial expressions. Finally, a confidence verification procedure is carried out as a post-processing to handle cases of mis-tracking or self-occlusion. As a result, our technique is robust and insensitive to the variations in lighting, head motion, and facial expression.

Given the tracked facial features, the 3D non-rigid motion of the facial features caused by the facial expression is estimated using a motion extraction method. It will automatically eliminate the 3D head motion from the tracked facial features, therefore, the 3D facial feature motion caused by the facial expression can be extracted under arbitrary face orientations. Then based on the extracted 3D motion of the facial features, a probabilistic framework is utilized to recognize the facial expressions by integrating the Dynamic Bayesian Networks (DBN) with the Facial Action Units (AU) from psychological view. Because of the successful modelling of the spatial and temporal behaviors of the facial expression via the proposed framework, the facial expressions can be recognized robustly and accurately under various face orientations.

In face pose estimation, we developed a technique that automatically estimates the 3D face pose based on the discovered facial features [17]. First, in order to minimize the effect of the facial expressions, our approach only chooses a set of rigid facial features that will not move or move slightly under various facial expressions for the face pose estimation. Second, these rigid facial features are used to build a face shape model, whose 3D information is first initialized from a 3D generic face model. With the use of a frontal face image, the generic 3D face shape model is individualized automatically for each person. Thirdly, based on the personalized 3D face shape model and its correspond-
ing tracked facial features, our approach exploits a robust RANSAC based method to estimate the face pose parameters. Since this method automatically removes the inaccurate facial features, face pose parameters can be always estimated from a set of facial features that are accurately detected in the face image.

3.2. Physiological, Behavioral and Performance Evidence

In addition to visual evidences, we exploit physiological, behavioral and performance evidences in the model since they are also indicators of human stress [4, 12, 7]. To collect physiological evidences, an emotional mouse is especially built from a normal mouse by equipping with physical sensors. The emotional mouse measures heart rate, skin temperature, Galvanic skin response (GSR) and finger pressure. The mouse is designed to be non-intrusive. (A lengthy description of the emotional mouse is deferred to a longer version of the paper.) For behavioral evidences, we monitor user’s interaction activities with the computer. Two evidences are the number of mouse clicks and mouse pressure from fingers in a time interval. For performance data, we collect math error rate, math response time, audio error rate and audio response time which are extracted from a log-file that keeps track of user’s responses to the tasks.

4 DBN Construction and Inferences

4.1 The DBN Model

A Bayesian Network (BN) is a directed acyclic graph (DAG) that represents a joint probability distribution among a set of variables. In a BN, nodes denote variables and the links among nodes denote the conditional dependencies among variables. The dependencies are characterized by a conditional probability table (CPT) for each node. A DBN additionally models the temporal relations of the variables. A DBN is a probabilistic framework for uncertainty knowledge representation, information fusion and reasoning under uncertainty. These capabilities are critical to the success of a stress modelling and inference system.

We construct the DBN in Figure 4 to model human stress. The DBN is split by time steps. At one step, the diagram breaks into two portions. The upper portion, from the top to “stress” node, depicts the elements that can alter human stress. These elements include user’s stress level at the previous step, the workloads, the environmental context that can be simple or complex, specific characters of the user such as her traits that can be extroverted or introverted and importance of the goal that she is pursuing. This portion is called predictive portion. On the other hand, the lower portion of the diagram, from the “stress” node to leaf (evidence) nodes, depicts the observable features that reveal stress. These features include the quantifiable measures on her physical appearance, physiology, behaviors, and performance. This portion is called diagnostic portion. The hybrid structure enables a DBN to combine the predictive factors and observable evidences in stress inference.

In the diagnostic portion, to model correlations among evidences from the same modality, an intermediate node is introduced for each type of evidences. For instance, a “physical” node is introduced to link “stress” and the observable visual features. The intuition is that the user affect influences her physical status; in turn, her physical status influences the observable features such as blinking frequency, AECS and others. For the same reason, three other nodes “physiological”, “behavior” and “performance” are added. These variables, represented as the intermediate nodes, are hidden. They, however, are necessary to model the correlations among evidential variables. Similarly, we introduce several hidden nodes “eye movement”, “gaze”, “head movement”, “facial expressions”, “pupil” to model the correlations among evidential variables in the same group. Such a hierarchical structure successfully integrates multiple-modality evidences.

To capture the dynamics among variables, we consider two types of conditional dependencies for variables at two adjacent time steps. In the first type, a link exists in an influencing factor of user stress, showing how the influencing factor evolves over time. In our design, one link exists from the workload at the previous step to that at the current one. This link depicts how the workloads vary along time. Later on, we will exploit this to study how stress is influenced by the workloads. Apparently, links can be added for the context, traits and goal nodes. In the second type, a link exists from stress at the previous step to stress at the current one. This link depicts how stress at the previous step, together with other influencing factors, produces impacts over stress at the current step. In particular, as the workload, context, user traits and goal remain fixed, the link shows how stress self-develops.

4.2 Model Parameterization and Dynamic Inferences

Model Parameterization determines the conditional dependencies given the model structure of the BN, i.e., the conditional probability table (CPT) for each node. First, we resort to domain expert in specifying the initial CPTs. Then, we use the EM learning algorithm [10] to refine the CPTs based on the data from different subjects.

Dynamic inference estimates the user stress level at each time step \( t \). We first introduce the notations and then define the problem. We shall use the first character of a node to refer to the node, i.e. \( w \) referring to \textit{workload}. In addition,
we subscribe a variable by a step $t$ to refer to the variable at time $t$, i.e., $s_t$ for stress node at time $t$. Under these notations, the DBN model specifies three probabilistic relations specified by the DBN model in Figure 4: the workload transition model $p(w_t|w_{t-1})$ (it is assumed that $c_t$, $t_t$ and $g_t$ are time invariant in the experiments), the stress transition model $p(s_t|s_{t-1}, w_t, c_t, t_t, g_t)$ and the evidence generation model $p(z_t|s_t)$ where $z_t$ is the set of evidences observed at step $t$. The inference at step $t$ is to calculate the probability $p(s_t|z_{1:t})$ where $z_{1:t}$ is the set of all available evidences $z_{1:t} = \{z_k, k = 1,\ldots, t\}$ up to time $t$. In case $t = 0$, $p(s_t|z_{1:t})$ degenerates to the prior $p(s_0)$.

From a Bayesian point of view, the task is to recursively compute $p(s_t|z_{1:t})$ from $p(s_{t-1}|z_{1:t-1})$. Conventionally, the task can be accomplished at two stages: prediction using the predictive portion of the DBN and correction using the diagnostic portion. However, since our model has a temporal link with workload, the probability for workload $w_t$ must be updated over time. By Total Probability Theorem, $p(w_t) = \sum_{w_{t-1}} p(w_t|w_{t-1})p(w_{t-1})$.

Then, in the prediction stage, the prior probability of user stress at step $t$ is calculated by the Chapman-Kolmogorov equation. The prior is denoted by $p(s_t|z_{1:t-1})$.

$$p(s_t|z_{1:t-1}) = \sum_{s_{t-1}, w_t} p(w_t)p(s_{t-1}|z_{1:t-1})p(s_t|s_{t-1}, w_t, c_t, t_t, g_t).$$

The equation exploits the Total Probability Theorem and the fact that $c_t$, $t_t$ and $g_t$ are time invariant.

In the correction stage, the evidence set $z_t$ is used to update the prior $p(s_t|z_{1:t-1})$ by Bayes’ rule:

$$p(s_t|z_{1:t}) = \frac{p(z_t|s_t)p(s_t|z_{1:t-1})}{p(z_t|z_{1:t-1})} = \frac{p(z_t|s_t)p(s_t|z_{1:t-1})}{\sum_s p(z_t|s)p(s|z_{1:t-1})}.$$

In summary, the inference at step $t$ sequentially calculates the probability $p(w_t)$, the prior probability $p(s_t|z_{1:t-1})$ and the posterior probability $p(s_t|z_{1:t})$.

## 5 Active Sensing

In order to minimize the adverse consequence that the human stress may bring about, the active sensing technique can be employed to infer user stress level in an accurate and efficient manner. For accurate inferences, active sensing can choose the most informative evidences and also rule out non-informative or conflicting evidences in revealing human stress. For efficient inferences, active sensing can elect to instantiate fewer evidence variables and thereby make an inference algorithm faster.

The mutual information concept offers a standard solution to choose the evidences. Let $s_t$ be the stress node, and $e$ be a set of evidence nodes $\{e(1),\ldots, e(n)\}$. The mutual information $I(s_t; e)$ is defined as

$$I(s_t; e) = - \sum_{e(1),\ldots, e(n)} p(e(1),\ldots, e(n)) \ln p(e(1),\ldots, e(n)) + \sum_{s_t, e(1),\ldots, e(n)} p(s_t, e(1),\ldots, e(n)) \ln \frac{p(e(1),\ldots, e(n), s_t)}{p(s_t|z_{1:t-1})}$$

where the sum is taken over the possible values of the variables, and $p(s_t|z_{1:t-1})$ is the predicted probability in the dynamic inference. The optimal evidence set $e^*$ is the one that maximizes $I(s_t; e)$. The evidence set $e^*$ is passed to the inference engine, which will construct the posterior probability $p(s_t|z_{1:t})$.
that maximizes $I(s; e)$. The evidences in the set $e^*$ should be the most informative in revealing the user stress. Note that since the evidence selection criterion depends on the predicted probability $p(s|z_{1:t-1})$, it is able to dynamically select evidences at each time step.

We remark that a search over all possible evidence combinations grows exponentially with the number of evidence nodes. Thus, in the experiments, we use a greedy approach to choose a certain number of the available evidences by ranking the quantity $I(s; e)$. The results show that the system performance increases with the number of evidences.

6 Experiments

We have conducted experiments with five subjects of different ages, genders and races. We report our analysis on the relations between stress and individual evidences, the stress recognition results, and the effectiveness of the active sensing technique in improving the recognition performance.

6.1 Methodology

Our validation method rests on the existing psychological theories. The existing results show that occupational stress was affected by two job characteristics: demand and control [8]. Demand refers to the amount of attention and effort required to carry out one’s job. We will interchangeably use demand and workload here. Control primarily refers to the decision-making freedom present in a job. It is predicted and confirmed that a user becomes more stressful when demands were higher or when control was lower [16]. This gives us a referent to study the relations between task characteristics and observed features. On one hand, by deliberatively changing the workloads and control of tasks, the user stress level can be changed. On the other hand, these changes may be embodied in the user physical appearance features, behavioral features and performance measures. Hence, if the user control over tasks remains unchanged and only the workloads undergo changes, in an effective framework the inferred stress levels should be able to reflect these changes. In our confined laboratory environment, it is reasonable to assume that a user’s control over math and audio tasks remain unchanged. Therefore, we can use workload to represent stress in the subsequent discussions.

6.2 Stress versus Individual Evidences

The experimental data shows that the individual features are sensitive to the stress. As stress increases, a participant blinks less frequently, closes the eyes faster, dilates the pupils more often, and focuses the eye gaze on the screen more often and remains there longer. Also, as stress decreases, a participant moves the head and opens the mouth less frequently, clicks the mouse button harder, the heart rate increases, and GSR decreases. We illustrate the relations between stress and three evidences – PerLPD, GazeDis and AECS as an example.

First, we draw the evidences and stress over time steps. The top three charts of Figure 5 present the relations between stress and each of the three evidence variables in a 20-minute period for one subject. For each time step in the x-direction, its value in y-direction is the running average in a 36-second time interval. Note that the values of visual evidences have been normalized into the range $[0.0, 1.0]$. It can be seen that approximately PerLPD and GazeDis are positively correlated to stress, whereas AECS is negatively correlated to stress.

Second, to obtain the quantitative relation between human stress and individual evidences, we conducted correlation analysis. The correlation coefficients along time steps are plotted in the bottom charts of Figure 5. The x-direction denotes the time steps, whereas the y-direction denotes the coefficients. It can be seen that at most time steps the correlation coefficients between stress and PerLPD (GazeDis) are positive, whereas the coefficients between stress and AECS is negative. This is consistent with our preceding analysis.

However, individual evidences are not completely reliable. Let us take the coefficients between stress and PerLPD as an example (the bottom chart in Column (a)). Although it is approximately positively correlated to stress, negative coefficients occur in a time period; in addition, at some time steps, the coefficients fall below 0.5, thereby indicating that the correlation between stress and PerLPD is somehow weak. However, with the DBN model by combining the individual features, the inferred stress level has a very high correlation with the ground-truth stress, as shown in the next section.

6.2.1 Stress Recognition

The experiments prove that our system can successfully monitor human stress. The results for Subject A are shown in the top two charts in Figure 6. In each chart, a solid curve denotes the ground-truth stress levels, while a dashed curve denotes the inferred stress levels. Note that due to our discussion in the validation subsection, for simplicity we use workload as an indicator of the ground-truth stress level. The top-left chart shows inference results with six evidences, whereas the top-right shows those with ten evidences. The evidences are dynamically determined by the information-theoretic criterion. We see that inferred stress curves roughly trace the ground-truth curves. This suggests that the inferred stress results reasonably agree with those
predicted by psychological theories. In addition, we see that at most time steps the inferred stress curve from ten sensors is closer to the solid curve than that from six sensors.

\begin{table}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Subject & A & B & C & D & E \\
\hline
Corr\_Max\_single & 0.68 & 0.62 & 0.76 & 0.56 & 0.67 \\
Corr\_Min\_single & 0.28 & 0.25 & 0.24 & 0.27 & 0.27 \\
Corr\_Combine & 0.89 & 0.85 & 0.92 & 0.79 & 0.86 \\
\hline
\end{tabular}
\caption{Correlation coefficients between inferred stress and ground-truth stress}
\end{table}

Finally, to give results for more than one subjects and show how multiple evidences outperform individual evidences in stress inferences, Table 1 presents the overall correlation coefficients between the ground-truth stress and inferred stress in a 80-minute period for five subjects. For a subject, its overall correlation coefficient is the average of coefficients across all time steps. In the table, Corr\_Max\_single (Corr\_Min\_single) indicates the maximum (minimum) absolute value of the overall correlation coefficient between stress and inferred stress using the individual features; Corr\_Combine indicates the overall correlation coefficient for inference stress using all available features. The magnitudes of the coefficients indicate the degree to which inferred stress level is correlated to ground-truth stress. We see that the overall coefficients using all the evidences are relatively large (their upper bound is 1.0). The table shows that the inferred stress level is quite closely correlated to the ground-truth. This indicates that the Bayesian technique is effective in monitoring the human stress. In addition, the table shows that the overall coefficients using all evidences are larger than those using individual evidences. This indicates that the Bayesian inference technique is effective in integrating multiple evidences from multiple modalities.

6.2.2 Active Sensing

The remaining charts in Figure 6 demonstrate that the active sensing approach outperforms a passive (random selection) approach. The middle two charts show the inference performances for the random selection strategy, where the left uses six evidences and the right uses ten evidences. To quantitatively show how much the active sensing approach outperforms, we conducted statistic correlation analysis and drawn the coefficients in the bottom two charts. In both charts, the solid (dashed) curve denotes the correlation coefficients along time steps between the ground-truth stress and inferred stress with the active (passive) sensing approach. It is observed that in both charts, almost all the time, the solid curve lies above the dashed curve. This implies that the inferred stress is more correlated to the ground-truth stress in the active sensing case. Consequently, the active sensing approach is effective in improving the inference performance.

These correlation coefficients also suggest that for the same sensing strategy, the inference performance can be improved as more optimal evidences are used to recognize the subject’s stress. For instance, in the active sensing case, the overall average value of the correlation coefficients (between ground-truth stress and inferred stress) in bottom-left chart for six evidences is around 0.5, whereas the overall average value in the bottom-right for ten evidences is around 0.8. Hence, to some extent, when more evidences are used, the inferred results are more correlated to the ground-truth stress. Nonetheless, it is difficult to characterize formal conditions to ensure this monotonicity.

7 Conclusions and Directions

We present a real time non-invasive system that infers user stress level from evidences of four modalities. A DBN framework is employed to model the user stress and these
evidences. The experiments show that the inferred user stress level by our system is consistent with that predicted by psychological theories. We also demonstrate the effectiveness of DBN framework in fusing multiple-modality evidences and the active sensing strategy in choosing evidences. Our future work is to extend the current framework to affective state recognition in addition to stress modelling.

References


