Head movements in context of speech during stress induction

Giorgos Giannakakis*, Dimitris Manousos*, Panagiotis Simos†*, and Manolis Tsiknakis†*
*Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH), Heraklio, Greece
Email: ggian@ics.forth.gr
†School of Medicine, Division of Psychiatry, Heraklio, Greece
‡Department of Informatics Engineering, Technological Educational Institute of Crete, Heraklio, Crete

Abstract—This paper focuses on the analysis of head movements in the context of speech during stress and neutral conditions. An experimental protocol was followed including tasks of neutral and stressful emotional states induced by different types of stressors. Translational and rotational head movements and velocities were computationally estimated from 2-dimensional facial landmarks in order to assess stress states. In parallel, the effect of speech production on head motility was investigated. The results indicate that stress conditions increase head motility, in both translational or rotational movement features. Besides, there is a clear involvement of speech in the increase of head motility. However, the intensity of head motility can be attributed to the combined effect of stress or arousal and speech and not just the effect of speech production.

Keywords—stress; head movements; head pose; speech; AAM

I. INTRODUCTION

Head movement behaviour is considered to be part of non-verbal communication that is expressed in various aspects of everyday life. They belong to the human’s upper part body gestures serving as signs in the course of human interaction as well as of emotional states in general. There are several head actions such as lowering, raising, tilting, nodding, shaking which have specific meanings and they are recognizable in cross-cultural communication. It is interesting that head gestures are observed also in blind persons that share many, if not all, characteristics of similar behaviour among persons with intact vision [1], [2].

Alongside their communicative importance, head movements may convey affective information. In [3], head motion was modelled in order to recognize behavioural characteristics, among them expression of positive and negative emotions, while in [4] head movements presented significant differentiations between positive and negative affect in infants. In [5], gesture analysis performed using among other features head motion information in order to recognize affective states with increased accuracy. In [6] head nods and shakes were employed among other features in order to discriminate complex emotional situations. Head movements related features have been used in limited studies estimating whether a subject is in a stressed state. It has been reported that head movements during stressful conditions are more frequent [7], more rapid [8] and that, in general, there is greater overall head motion [9], [10].

In addition, there are studies arguing that head movements are modulated by ongoing speech [11]–[13]. In [13], the linguistic functions of head movements as well as their semantics during conversations are presented. Moreover, head shifts were associated with the text structure and prosody [14]. Also, head nodding occurs when a person wants to emphasize a section of his/her speech or a specific word. In [15], head postural shifts (PS) are defined as wide, linear movements which occur mainly before speech initiation.

In the literature, studies addressing the problem of stress detection through head motion features are limited. Moreover, to our knowledge there is no combined analysis of stress and ongoing speech on head movements. This study investigates affective information both the independent as well as the combined effect of stress and overt speech on head movement patterns.

II. METHODS

A. Experimental procedure and population

Data were recorded during the first data acquisition campaign (Semeoticon Reference Dataset SRD’14) within the context of the research project described in [16]. From this dataset, the recordings regarding stress assessment (Semeoticon Reference Dataset for Stress Assessment SRD’14) were used which are described in detail in [10].

The study sample included 23 participants (16 males, 7 females) aged 45.1±10.6 years. Each participant performed the whole experimental procedure (all the 12 different tasks) which means 69 neutral and 184 stressful sessions (69 with speech, 115 without speech) that were used in the analysis.

Table I

<table>
<thead>
<tr>
<th>#</th>
<th>Experimental task</th>
<th>Speech</th>
<th>Duration (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Neutral (NT - 3 tasks)</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Interview (INT)</td>
<td>Y</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Reading Text (RT)</td>
<td>Y</td>
<td>½</td>
</tr>
<tr>
<td>4</td>
<td>Anxious event recall (AER)</td>
<td>N</td>
<td>½</td>
</tr>
<tr>
<td>5</td>
<td>Stressful event recall (SER)</td>
<td>N</td>
<td>½</td>
</tr>
<tr>
<td>6</td>
<td>Images from IAPS database (IAPS)</td>
<td>Y</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Stroop Colour-Word Test (SCW)</td>
<td>Y</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>Adventure video (AV)</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>Psychological pressure video (PPV)</td>
<td>N</td>
<td>2</td>
</tr>
</tbody>
</table>
Three (3) neutral (reference) tasks, 1 relaxed task and 8 stressful tasks (interview (INT), reading Text (RT), stressful event recall (AER and SER), viewing images from IAPS database [17], the Stroop Colour-Word Test (SCWT) [18] and viewing stressful videos (AV and PPV)) took place (a more detailed tasks description and rationale can be read in [10]). These stressors are listed in Table I along with the overall task duration and whether overt speech was involved.

B. Head movements

The head movements can be expressed along 3 translational axes: x axis (horizontal, forward-backward head movement), y axis (lateral, left-right head movement), z-axis (vertical, up-down head movement) and 3 rotational axes: yaw (left-right head turn), pitch (up-down head turn) and roll (turn head towards-away from shoulder). These axes are schematically represented in Figure 1.

![Figure 1. The translational (x,y,z axis) and rotational (yaw, pitch, roll) head movements](image1.png)

The computational estimation of facial landmarks was performed using the inverse compositional image alignment algorithm of the Active Appearance Models (AAM) [19], [20] resulting in 68 total facial landmarks.

C. Translational motion features

For the estimation of the translational head movements, 5 facial landmarks \( p_k, k = 1, 2, ..., 5 \) around the nose were used (from the total 68) whose distribution is shown in Figure 2. These landmarks were selected because they are not affected by motion associated with facial expressions (e.g. mouth deformation, eyebrow raising, frown etc) or speech (chin and cheek movement) and they contain merely translational head motion information.

The relational position of these landmarks in pairs determine the distances \( d_k, k = 1, 2, ..., 5 \) in the y-z plane. The mean head movement \( \overline{d} \) was defined as the mean value of these distances across all \( N \) video frames during each task

\[
\overline{d} = \frac{1}{N} \sum_{i=1}^{N} d = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{5} \| p_k(t) - p_k(t_{ref}) \|
\]

where \( p_k(t_{ref}) \) is the reference landmarks position. The latter was defined as the 100\(^{th}\) frame (2nd sec of recording) where the subjects are in initial conditions.

The mean head movement velocity \( \overline{v} \) was defined as the mean value of distances between successive frames and it can be estimated by the equation

\[
\overline{v} = \frac{\overline{d}}{T} = \frac{1}{N} \sum_{t=2}^{N} \sum_{k=1}^{5} \| p_k(t) - p_k(t-1) \|
\]

where \( p_k, k = 1, 2, ..., 5 \) the landmark points, \( N \) is the number of frames and \( \| \cdot \| \) is the Euclidean norm.

The velocity features were considered more appropriate comparing with position features as they are not subject to the reference point which may affect the estimated features. These measures were also estimated individually in lateral and vertical axes providing the measures \( d_y, d_z, v_y, v_z \) respectively, representing type of motion along these axes. Their relation can be shown in the following equations

\[
d = \sqrt{d_y^2 + d_z^2}, v = \sqrt{v_y^2 + v_z^2}
\]

D. Rotational motion features

Head pose also conveys important information about head motility that can be estimated using computer vision techniques [21]. Head pose features measure angular displacements on rotational axes representing head nods (pitch), turns (yaw) and lateral inclinations (roll). In this study, the head pose was estimated by projecting a 3D face model onto
2D AAM using the Pose from Orthography and Scaling with Iterations (POSIT) algorithm [22]. The rotational axes of these features are represented with the blue, red and green vectors respectively in figure 3.

In addition to angular displacements, angular velocities were calculated as the derivatives of the corresponding angular displacements respectively. It should be noted that the different types of movements may overlap with each other and it is in most cases difficult to decouple them, e.g. pitch rotation - vertical displacement, yaw rotation - lateral displacement.

III. RESULTS

The head movements were estimated, as described in the previous section, using the translational features (total movement, movement y, movement z, total velocity, velocity y, velocity z) and the rotational features (pitch, yaw, roll angles and pitch, yaw, roll angular velocities). Data were grouped by emotional state (2 levels: stress, no stress), vocal channel (2 levels: speech, no speech).

A. Head movements translational features

While investigating head movement translational velocities during experimental tasks, it was observed that the most prominent increases occurred during speech tasks, i.e. during the Interview (INT), the Reading Text (RT) task and the Stroop Colour Word Test task (SCWT).

A two-way ANOVA was performed in order to examine the effects of stress and speech on head movements velocity. In this analysis it was assumed that the RT task would induce lower stress levels in relation to INT, SCWT, an assumption corroborated by the heart rate measures. Stress leads to significant increased head velocity ($F(1,272) = 4.06, p = 0.045 < 0.05$). Also, overt speech lead to a prominent increase in head velocity in relation to silent conditions ($F(1,272) = 176.89, p < 0.01$). However, as expected, there was no interaction effect between speech and stress conditions ($p = 0.77$).

Among the three experimental tasks that involved speech, the greater head motility was noted during the interview task followed by RT and SCWT tasks. Pairwise comparisons revealed that during the interview task, head motility was significantly increased in relation to the reading text task ($p = 0.021 < 0.05$) and SCWT ($p = 0.020 < 0.05$). Corresponding changes were found in the heart rate activity (maximal pairwise increase during Stroop task) which supports the view that stress leads to increased velocity of head motion as compared to a less stressful task, i.e. the text reading task. This can be attributed to an increased cognitive load while searching proper words in answering questions during the interview. On the other hand, no significant differences were observed between the reading text task and the SCWT task despite the assumption that the latter is somewhat less demanding given that subjects were asked to produce a limited set of utterances (names of 3 colors) rather than producing spontaneous, semantically and syntactically complex linguistic output and respond to the experimenter’s questions. In addition, the lateral (y-axis) and vertical (z-axis) movement velocities were investigated. Their behaviour was similar to the overall head movement velocity across experimental tasks in terms of significant findings.

During the stressful tasks, the mean lateral head velocity increased as compared to the vertical one, with the increase being significant during the interview task ($p < 0.01$), the viewing of stressful images from IAPS database, the adventure video and the psychological pressure video ($p < 0.05$). On the other hand, during the reading text and the SCWT tasks the vertical velocity was greater. This difference may be attributed to the direction of the head reading the text and the rhythmicity of word presentation during Stroop test but it should be further examined.

Also, the head movement velocities along the y,z axes were investigated for each subject individually. It was found that lateral and vertical head velocities were highly correlated ($r$ ranging from 0.50 to 0.91) on all but a single neutral task. This implies a kind of coordination of human head during different types of head movements in a wide
range neutral and varying stress levels tasks. It is interesting that the highest correlation \( r = 0.91 \) was observed in the interview task where the subjects were asked to give a brief verbal presentation of themselves.

B. Head pose features

Investigating the head pose features of pitch, yaw, roll and their corresponding velocities, it was observed similar variation in relation to translational features mainly for the velocity features. Distribution of the pitch motion and velocity values are shown in figure 5.

![Figure 5. Pitch angular displacement (left panel) and pitch velocity (right panel) distribution for the experimental tasks.](image)

A two-way ANOVA was performed in order to examine the effects of stress and speech on the pitch (turn head up-down). More negative head pitch and increased head pitch velocity was observed during speech conditions \( F(1,272) = 82.33, p < 0.01 \), in the absence of a significant interaction between speech and stress \( p = 0.78 \). Pairwise comparisons showed that pitch was significantly more negative during interview (INT) \( p = 0.017 < 0.05 \) and reading text (RT) task \( p = 0.010 < 0.05 \). This denotes that during stressful tasks (only those that involve speech), participants raised their head more frequently than in to neutral tasks.

Neither amplitude nor velocity of yaw movements varied significantly as a function of elevated stress (i.e., between stressful and neutral tasks). However, the yaw velocity was significant increased in interview (INT), reading text (RT), viewing images from IAPS and Stroop Colour Word (SCWT) task in relation to neutral tasks. Roll movements increased in interview (INT), reading text (RT) tasks in relation to neutral and the roll velocity in interview (INT), reading text(RT), viewing images from IAPS and Stroop Colour Word Test (SCWT) in relation to neutral tasks.

IV. DISCUSSION

The findings of this study indicate that stress conditions can lead to increased head motility expressed in both head movement and head velocity. This differentiation is typical as compared to neutral or relaxed state. Increased task difficulty, which was also associated with increased arousal, was linked to further increase in head motility (in terms of both amplitude and velocity).

The most prominent and significant head movement increases of any type (translational, rotational) were observed during the interview task, reading text and Stroop-Colour Word Test. Correspondingly, heart rate as an indicator of arousal was also increased during stressors. Importantly, however, significant positive correlations between heart rate and head movement parameters were noted only in tasks associated with significant increases in heart rate.

There is the notion that head movements are associated with the cognitive and language processes involved in speech production. Increased motility is presented during speaking as some studies suggest and is also supported by the analysis of this study. However, results indicate that head movements relate to the intensity of induced stress and specifically with concurrent arousal levels. Stress and speech have both effect on head velocity, but there was no interaction effect between them. It is deduced that head movements and velocities are modulated by the stress load and their variation is not just the effect of speech production. The interpretation of head motion features should be cautious when speech also occurs where its effect should be taken into account.

Regarding pose features, they follow similar task-related variation with translational features. The pitch was significantly reduced only in stress-inducing tasks which also involved overt speech indicating a tendency participants to raise their head in relation to neutral tasks.

An interesting point of methodological interest is the fact that lateral and vertical head movements are correlated. This implies a type of coordination between these directions whose mechanisms should be further investigated. There were specific tasks that the lateral movement was relatively increased (e.g. in interview task) and in some others where the vertical movement was relatively increased (e.g. Stroop Colour Word Test, Reading Text) which can be attributed to the specific task characteristics, i.e. direction of text reading or the rhythmicity of word presentation.

A qualitative observation is that head movements during speech relate to the participants’ overall level of emotional expressiveness, which is expected to correlate strongly with prosodic speech features. This tendency may be more pronounced during the initial parts of an interaction through the production of redundant gestural information in order to limit the possibility of misunderstandings [12]. This is supported also in this study where the interview task presented the most prominent increase in head movement amplitude and velocity.

ACKNOWLEDGMENT

This study was supported by IKY scholarships programme and co-financed by the action entitled Reinforcement of Postdoctoral Researchers, in the framework of the Operational Programme Human Resources Development Program, Education and Life-long Learning of the NSRF 2014-2020.
REFERENCES


