A Virtual Assistant to Help Dysphagia Patients Eat Safely at Home

Michael Freed, Brian Burns, Aaron Heller, Daniel Sanchez, *Sharon Beaumont-Bowman
SRI International, Menlo Park California / *Brooklyn College, Brooklyn New York
{first name, last name}@sri.com, sharonb@brooklyn.cuny.edu

Abstract

For millions of people with swallowing disorders, preventing potentially deadly aspiration pneumonia requires following prescribed safe eating strategies. But adherence is poor, and caregivers' ability to encourage adherence is limited by the onerous and socially aversive need to monitoring another's eating. We have developed an early prototype for an intelligent assistant that monitors adherence and provides feedback to the patient, and tested monitoring precision with healthy subjects for one strategy called a "chin tuck." Results indicate that adaptations of current generation machine vision and personal assistant technologies can effectively monitor chin tuck adherence, and suggest feasibility of a more general assistant that encourages adherence to a range of safe eating strategies.

Dysphagia Patients Need Help With Adherence

Dysphagia, or difficulty swallowing, is a widespread and often devastating disorder that affects 10–30% of the elderly population and high percentages of patients with neurological conditions such as stroke (50–75%) and Parkinson's Disease (up to 95%). Dysphagia creates numerous risks; chief among them is aspiration pneumonia—an infection caused by accidental ingestion of bacteria-laden food into the lungs with mortality ranging from 10–70% (DeLegge, 2002).

Clinicians frequently prescribe risk-reducing compensatory strategies such as tucking the chin to the chest before swallowing to protect the airway, and making an "effortful swallow" to clear residual food in the pharynx. These strategies have been shown to significantly improve patient health and well-being (Low et al., 2001). However, few patients can continuously self-monitor eating behavior (Shinn et al., 2013), leaving the task of policing adherence to caregivers. This creates significant relationship strain and caregiver burden. "It is poignant to note that we can assist people with dysphagia by first assisting their caregivers (Cichero and Altman, 2012)."

Our preliminary results indicate the feasibility of shifting the burden of monitoring and encouraging adherence to a device that actively monitors patient-eating behavior and provides appropriate real-time feedback. Dysphagia patients are often advised to eat in front of a mirror to enhance self-awareness. We have constructed a proof-of-concept prototype for a laptop- or tablet-based application that, when positioned in front of the patient, functions as both a mirror and a source of feedback to assist use of safe compensatory strategies. The prototype successfully monitors compliance for a single eating strategy—specifically, whether the patient is compliant with flexing the head downward into a "chin tuck" after spooning food into the mouth. We believe that currently available machine-vision and machine-learning algorithms can be used to monitor adherence with the most commonly prescribed safe eating strategies, and that this monitoring can be done in real time by using widely available consumer electronics hardware. Our ultimate goal is to develop a fully functional, patient-friendly Dysphagia Coach application that increases patients' eating safety and well-being while decreasing dysphagia-related mortality and caregiver burden.

Figure 1. Dysphagia Coach prototype showing detection of user and feedback on a chin tuck maneuver.

Prototype Approach and Initial Results

As illustrated in Figure 1, our prototype provides a simple chin tuck monitoring and user feedback capability. A monitoring session begins when the system detects a face within an estimated threshold distance from, and orientation, towards the camera (Viola and Jones, 2004), indicating that the patient may be ready to eat. It recognizes spoons with and without food, and tracks movement to and from the mouth. When the system detects that the spoon has been inserted into the mouth, it monitors for a large, downward rotation, or "tuck," of the head, followed by an upward rotation into a rest position. If a tuck is detected after spoon insertion and before any subsequent insertion, the system gives the patient positive feedback. Otherwise, it indicates that the tuck was omitted.

Of the requirements explored using our prototype, the need to precisely estimate head angle using sensors and processors on commodity hardware raised the most significant question of technical feasibility. We selected an approach incorporating landmark-based machine vision algorithms (Sagonas et al., 2013) running on a standard, camera-equipped laptop. Landmark algorithms find standard head and neck reference points, such as the tip of the nose, in a video image, and then map them to corresponding points in a reference digital head/neck model. Head pose is then estimated by computing the best 3D rotation fit between points and model (Demonthon and Davis, 1995). Importantly for detecting chin tucks, tilt angle can be treated as approximation of head-torso angle in the sagittal plane. Pilot data (n=5) showed an RMS estimation error of 3.6 degrees. This result was less than the within-subject variability of 5.2 degrees for correctly performed chin tucks, suggesting that the error is not clinically significant (Fig. 2, row 3 – red points are estimated head angles for shown frames; green circles are actual).

We designed the prototype to get feedback from clinicians on the concept, and to gauge the technical feasibility of the approach. Feedback confirmed the potential benefit to patients and provided guidance in prioritizing which safe eating strategies are most important to monitor. Clinicians also indicated strong interest in using the technology to collect patient performance data, both to improve care for individual patients and to advance scientific understanding of dysphagia.

Demonstration

The demonstration includes an interactive working prototype able to monitor eating behavior and provide feedback on chin tuck adherence, as well as videos and other materials that support discussion of additional forms of in-home disability support.

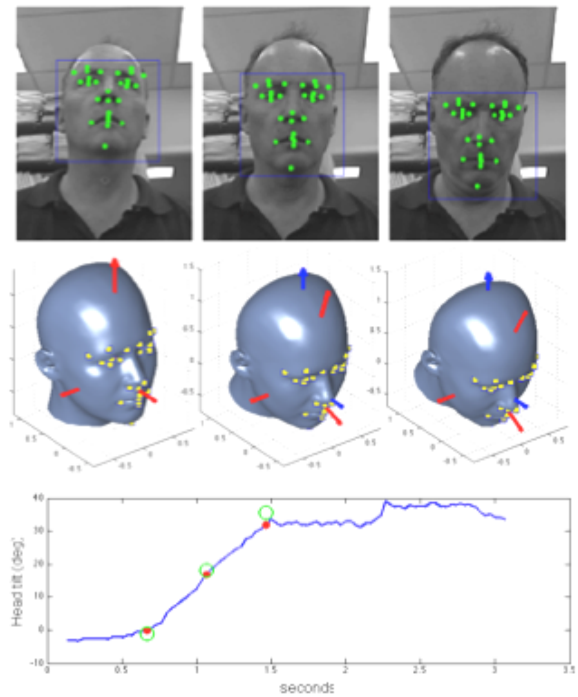

Figure 2. Chin-tuck angle is determined by finding landmark head/neck locations in video, mapping to a digital head model, and then computing the head pose. Results show accuracy to within 4 degrees.

References

Cichero JA, Altman KW. Definition, prevalence and burden of oropharyngeal dysphagia: A serious problem among older adults worldwide and the impact on prognosis and hospital resources. *Nestle Nutr Inst Workshop Ser.* 2012;72:1–11.

DeLegge MH. Aspiration pneumonia: Incidence, mortality, and at-risk populations. *J Parenter Enteral Nutr.* 2002;26(6 Suppl):S19-24.

Dementhon, D. and Davis, L. Model-based object pose in 25 lines of code. *International Journal of Computer Vision*; 15(1):123-141.

Low J, Wyles C, Wilkinson T, Sainsbury R. The effect of compliance on clinical outcomes for patients with dysphagia on videofluoroscopy. *Dysphagia.* 2001;16(2):123–127. doi:10.1007/s004550011002.

Sagonas C, Tzimiropoulos G, Zafeiriou S, Pantic M. 300 Faces in-the-wild challenge: The first facial landmark localization challenge. *Proc International Conference on Computer Vision (ICCVW).* 2013;397(403):2-8. doi:10.1109/ICCVW.2013.59.

Shinn EH, Basen-engquist K, Baum G, et al. Adherence to preventive exercises and self-reported swallowing outcomes in post-radiation head and neck cancer patients. *Head Neck.* 2013;35(12):1707–1712. doi:10.1002/HED.

Viola P, Jones M. Robust real-time object detection. *Int J Comput Vis.* 2004;57(2):137-154.