

Computational deglutition: Signal and image processing methods to understand swallowing and associated disorders

Ervin Sejdić, Georgia A. Malandraki and James L. Coyle

1. Introduction

Swallowing is a sensorimotor activity by which food, liquids, and saliva pass from the oral cavity to the stomach. It is considered one of the most complex sensorimotor functions due to the high level of coordination needed to accomplish the swallowing task over a very short period of one to two seconds, and the multiple subsystems it involves. Dysphagia (swallowing difficulties) refers to any swallowing disorder, commonly caused by a variety of neurological conditions (e.g., stroke, cerebral palsy, Parkinson's disease), head and neck cancer and its treatment, genetic syndromes, and iatrogenic conditions or trauma. The signs and symptoms of dysphagia range from anterior loss of food while eating, difficulty chewing, subjective difficulty in swallowing food or liquids, to choking or coughing before, during, or after eating due to impaired clearance of swallowed material from the throat into the digestive system. When not effectively treated, dysphagia can cause malnutrition, dehydration, failure of the immune system, psycho-social degradation, and in general, a decreased quality of life.

The major medical consequence of dysphagia is aspiration of food and liquids into the airway, often leading to airway obstruction, pneumonia, and an increased risk of mortality. Dysphagia affects approximately 9 million adults per year in the US [1] and is especially prevalent among the elderly. Characteristically, 50 to 75% of stroke patients and 60% to 70% of patients who undergo radiation therapy for head and neck cancer have dysphagia. In addition, over 60,000 people yearly die from complications associated with swallowing dysfunction. Complications of dysphagia drastically increase healthcare costs. Overall, together with the costs incurred by hospitals, costs of dysphagia in the healthcare system exceed one billion dollars per year.

In the past 30-40 years we have gained an increased understanding of this potentially devastating condition and have made remarkable improvements in the management of dysphagia. Given recent advances in signal and image processing algorithms, we strongly feel that the signal/image processing community is poised to make further fundamental contributions to the understanding of swallowing and swallowing difficulties and improve patient outcomes. There is a widespread need for signal and image processing algorithms that can help clinicians in the management of dysphagia. Therefore, we propose the establishment of a new signal/image processing subfield called ***computational deglutition***. This newly established translational subfield will be a collaboration between clinicians and the signal/image processing community aimed at the development of clinically relevant algorithms that will aid clinicians during the assessment and treatment of swallowing disorders.

2. Swallowing function and the swallowing mechanism

Oropharyngeal swallowing is not simply the act of propelling food and liquids toward the digestive system. It is an intricately timed, short-duration, centrally programmed patterned response designed to deliver nutrients, fluids, and medications to the digestive system while at the same time preventing aspiration of swallowed material into the airway. Within a few seconds, up to two dozen kinematic and valvular events, performed by more than 30 pairs of muscles, occur to simultaneously enable the upper aerodigestive tract to alternate between its respiratory and digestive functions.

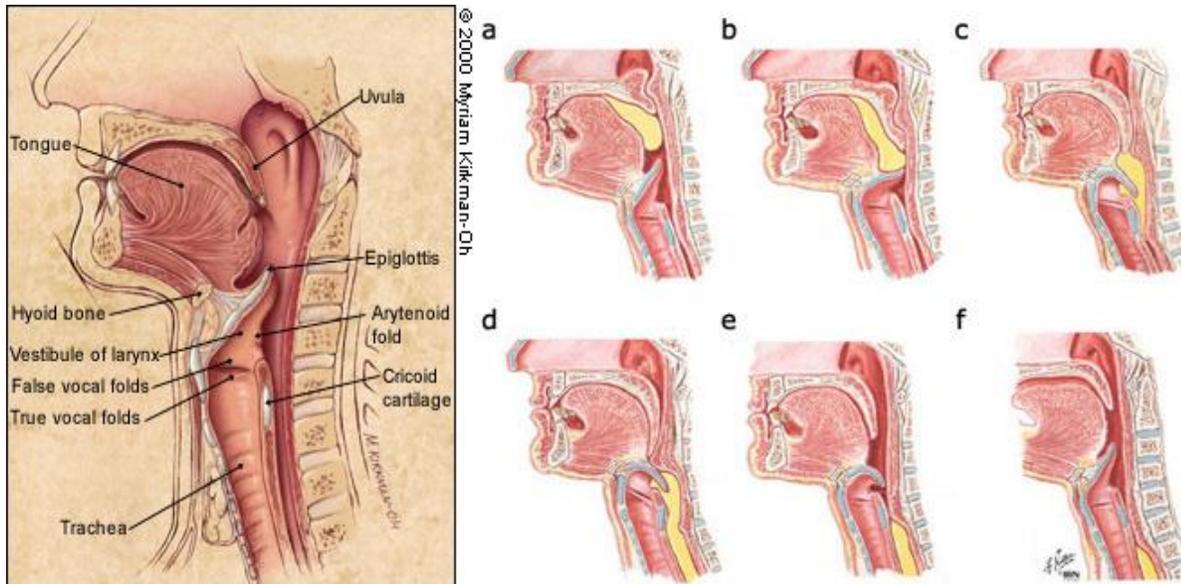


Figure 1 – LEFT: Anatomic landmarks in the sagittal view (Source: J. B. Palmer, J. C. Drennan, and M. Baba, "Evaluation and treatment of swallowing impairments," *American Family Physician*, vol. 61, no. 8, pp. 2453–2462, Apr. 2000.) RIGHT: The swallowing process: (a) the swallow initiation; (b) bolus is propelled by tongue and UES opening anticipating bolus arrival; (c) bolus enters the pharynx associated with epiglottal downward tilt, hyolaryngeal excursion, and UES opening; (d) bolus passes through the pharynx; (e) bolus passes the UES, and the oropharyngeal swallow is completed; (f) the entire bolus is on the esophagus (Source: J. A. Robbins, A. D. Bridges, and A. Taylor, "Oral, pharyngeal and esophageal motor function in aging," *GI Motility Online*, 2006.)

Whether swallowing reflexively during sleep or consciously enjoying a meal, swallowing delivers saliva and ingested nutrients through the pharynx, which is a single tube shared by both the respiratory (airway) and digestive systems, while valving gas (breathing) or food (swallowing) flow between them (see Figure 1 for relevant anatomical landmarks). During oral preparation, sensory receptors in the oral mucosa receive and carry sensory information through afferent pathways to the brainstem and the brain. During this stage, liquids are contained by oral valves, (lips, tongue and soft palate), and solid foods are mechanically reduced by mastication into a relatively cohesive bolus while saliva is mixed with the bolus. When the bolus is considered adequately prepared for transfer to the pharynx it is propelled posteriorly while a cascading sequence of events that direct flow away from the airway and toward the esophagus begins. These events include pharyngeal and laryngeal kinematic events that mediate the opening and closing of respiratory and digestive valves and reconfigure the oropharyngeal cavity, closing the airway, and opening the inlet to the esophagus. Hyolaryngeal excursion, a kinematic pattern

analogous to a series of pulleys between the mandible and skull base on one end and the hyolaryngeal complex on the other, leads to this alternating valving by displacing the larynx anteriorly and superiorly out of the path of the oncoming bolus and closing its inlet valve, while simultaneously contributing to distension of the upper esophageal sphincter (UES).

Laryngeal displacement and airway closure are accompanied by inversion of the epiglottis, the cartilaginous valve at the laryngeal inlet, and closure of the internal larynx, further ensuring protection of the airway. Because the posterior wall of the larynx is shared as the anterior wall of the UES, this upward and forward displacement delivers concurrent traction forces to the UES. This traction is a necessary factor contributing to opening of the digestive valve while progressive pharyngeal pressures continue to propel the bolus into the esophagus. Please refer to Figure 1 for a description of the swallowing process.

3. A brief introduction to dysphagia and common etiologies

With more than 700,000 new cases reported every year in the US, neural damage or impairment (e.g., stroke) serves as the most common cause of dysphagia. Usually, swallowing disorders post stroke are related to disruption of the sensorimotor functions mediated by the cranial nerves, which directly control the structures of the mouth and throat [2]. Naturally, if some of the 30 pairs of oral, pharyngeal and laryngeal muscles are not receiving proper neural inputs, the patient will not be able to have a fully functional swallow and may be unable to adequately transfer a bolus past the larynx and into the esophagus.

Neurodegenerative conditions, such as Huntington's or Parkinson's disease, also often result in swallowing disorders. In these cases, dysphagia is typically manifesting as oral and pharyngeal dis-coordination, rigidity, and/or reduced sensation in the oropharynx, which all can result in mistiming of airway closure and upper esophageal sphincter opening. In addition to neurogenic etiologies, there are several anatomically-related causes of dysphagia as well. Conditions that result in an inflamed and swollen esophagus, such as eosinophilic esophagitis or gastroesophageal reflux, can make it difficult for the patient to transfer a bolus through the esophagus. This can often lead to the feeling of food becoming ``stuck" in the throat. Various abnormal benign or malignant growths, such as tumors, swollen lymph nodes, or esophageal webs, can obstruct the path of a bolus as well, leading to similar feelings of food obstruction and risks of aspiration.

Finally, direct damage to the muscles and structures of the throat can also result in swallowing difficulties. Surgical procedures or radiation therapy typically used to manage head and neck cancer can cause disrupted propulsion and airway protection, and other sources of physical trauma can similarly lead to dysphagia by altering anatomy and sensorimotor integrity.

4. Swallowing and dysphagia assessment

Swallowing assessment can be sorted into two categories: screening and diagnostic testing. Screening tests are relatively simple pass-fail procedures performed by anyone trained in its administration and identifies patients with a high likelihood of having dysphagia and needing further testing. Like screens for breast cancer or heart disease, dysphagia screening provides no

diagnostic information regarding the physiologic nature of the disorder, nor provide information to guide treatment. Screens typically include simple water swallow challenges in which the patient fails the screen if they cough or produce other overt signs of aspiration of the swallowed water [3]. If these signs are absent it is assumed that dysphagia is absent and no intervention or further testing is performed.

Conversely, diagnostic testing identifies the physiologic nature of the disorder and informs the examiner, typically a qualified speech-language pathologist (SLP), about treatment options to mitigate the dysphagia and its adverse effects [4]. After a failed screen, a clinical/bedside evaluation is performed, without the use of instrumentation. It involves a detailed examination of oropharyngeal and laryngeal sensorimotor function, assessment of cognitive status, and observations of the patient swallowing a variety of textures and volumes of foods/liquids. The examiner synthesizes the results and determines whether the cause of dysphagia can be determined and remediated, and in some cases clinical evaluations are adequate to achieve these goals. However, since pharyngeal disorders and impaired airway protection are not observable without imaging technology, the clinical evaluation fails to detect asymptomatic impairments such as silent aspiration, or any pharyngeal events that occur beyond the intraoral view of the examiner. In such cases instrumental testing is performed.

Instrumental diagnostic tests characterize the physiologic nature of the dysphagia and identify potential interventions to mitigate its adverse effects, by elucidating the exact mechanisms of dysphagia along with its underlying causes. However, it is more complex, costly, invasive and time-consuming than a clinical examination, and requires expert clinicians and imaging instrumentation such as fluoroscopic or fiberoptic equipment. Overall, the need for more highly trained personnel increases as the diagnostic process flows from screening to clinical and then instrumental testing respectively.

Non-image based information or signals about some components of swallow function can be collected with noninvasive methods such as surface electromyography and cervical auscultation. Surface electromyography (sEMG) involves placing electrodes on the patient's anterior neck and recording the electrical activity of the underlying muscles during a swallow [5]. The theory is that if the nerves or muscles involved in swallowing are impacted, the signal will change in a clinically significant way when compared to a recording from a healthy patient or a healthy/normal swallow. Surface electromyography can only indirectly describe a swallow since it is limited to monitoring regional muscle activation and does not allow for isolated muscles or other regions to be assessed. As a result, this technique remains mostly experimental and complementary to other diagnostic methods and/or is used as a treatment biofeedback tool.

Cervical auscultation (CA) is another popular screening method in which a clinician listens to the throat with a stethoscope while the patient performs a swallow. The examiner then makes inferences regarding swallow integrity. The theory behind this, much like the sEMG procedure, is that the sounds recorded from a patient with dysphagia will be significantly different than that recorded from a healthy individual. However, stethoscopes and the human auditory system are incapable of transmitting or perceiving the entire spectrum of signals produced during swallowing, thus interpretation of these signals is imprecise and incomplete. Though attractive in its simplicity,

CA is unable to identify specific physiologic events or abnormalities. Currently, high resolution sensors (i.e., piezoelectric sensors, microphones and accelerometers) are under investigation in order to advance cervical auscultation by recording the entire spectrum of displacement, acoustic and vibratory signals emanating from the throat during swallowing.

The most widely accepted imaging method of assessing dysphagia is the videofluoroscopic (VFS) diagnostic examination (Figure 2, upper row) [6, 7]. During this test, the patient is asked to swallow small amounts of food or liquid mixed with a contrast agent, typically barium sulphate.

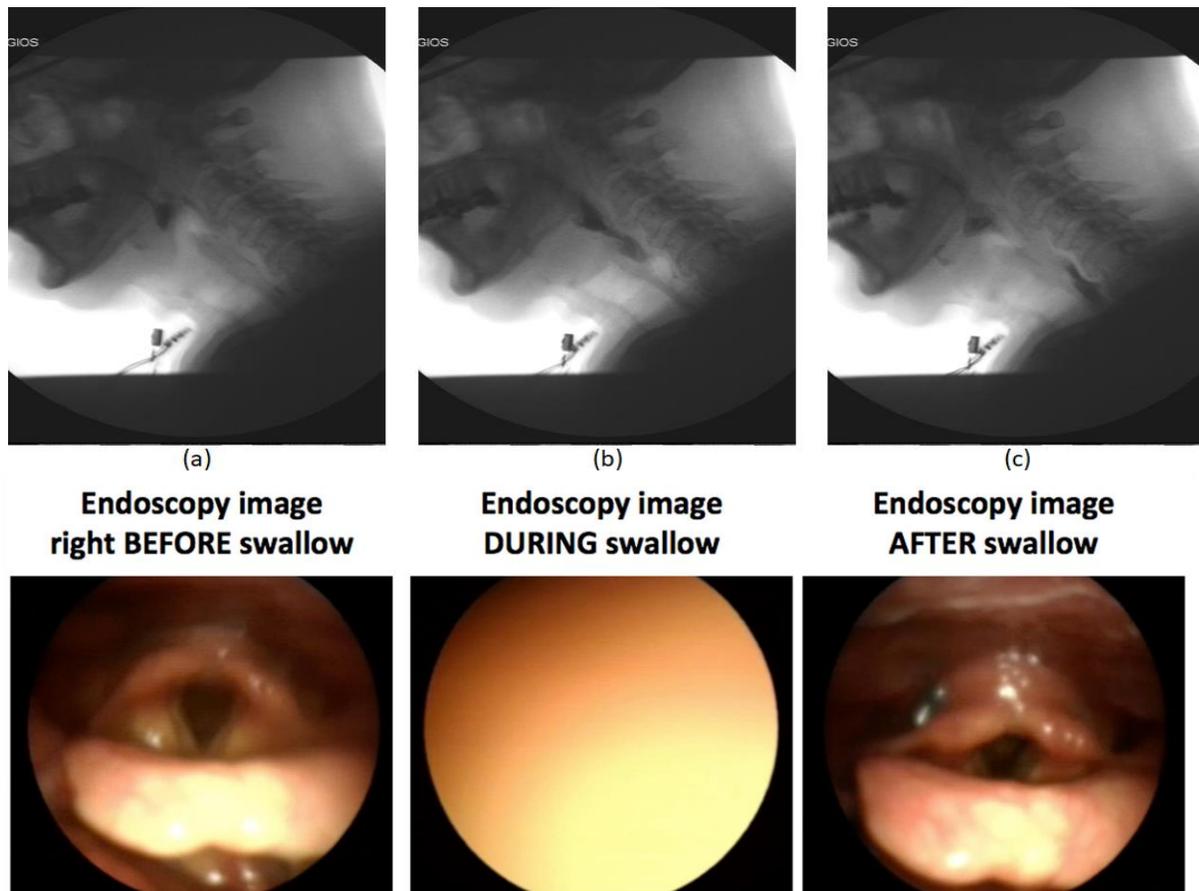


Figure 2 – UP: A patient passing a bolus through the oropharyngeal area as seen in videofluoroscopy. DOWN: In this endoscopic video-sequence, a patient is passing a bolus through the oropharyngeal area. Although structures can be easily viewed, the material swallowed is not as easily visible.

The x-ray equipment is aligned to produce a sagittal view of the oropharynx, pharynx, and upper esophagus containing all of the major swallowing structures, allowing an imaging clinician (i.e., radiologist) and a swallowing specialist (i.e., SLP) to observe and analyze the physiologic events that produce bolus movement in real time, determine which aspects of the swallow are not functioning properly, assess the timing and severity of impaired airway protection, and then deploy trial interventions. All these factors are necessary to form a comprehensive assessment of a patient’s swallow and have led to the widespread adoption of VFS as a diagnostic test.

Fiberoptic endoscopic evaluation of swallowing (FEES) is also used to assess swallowing disorders (Figure 2, lower row). Rather than using an x-ray imaging machine, a small fiberoptic

camera attached to a flexible endoscope is directed into the oropharynx and beyond through the nares while the clinician observes events occurring before and after the swallow. The advantages of this method is that the examiner can directly observe the patient's anatomy without x-ray, and examine much finer details as well as the color of surrounding tissues and symmetry of laryngeal function, both of which can provide important diagnostic information while the patient swallows regular foods (not barium). Because FEES tests are performed without radiation-risks, they can last a longer period of time than VFS, to assess issues like fatigue. However, this method has two key drawbacks relative to VFS. The first is that only a small range of oropharyngeal anatomy is visible at one time due to a limited field of view. Second, FEES techniques cannot view the swallowing mechanism before, during and after the swallow, leading to imaging blindness during the pharyngeal swallow as the pharynx collapses over the camera lens. As a result, FEES cannot meaningfully assess the actions of the pharynx or larynx during a swallow and is blind to oral and esophageal structure and function, further limiting the information provided.

Advanced neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), magnetoencephalography (MEG), and electroencephalography (EEG) are also instrumental methods that provide significant insights into brain activity during swallowing. These methods are mostly used experimentally at this time but are critical to understand and improve computational deglutition. The following sections discuss signal and image processing approaches and challenges relating to the aforementioned instrumental swallowing assessments.

5. Signal processing approaches and challenges

There are many signal processing challenges in computational deglutition. We will focus here on two prevalent cases that rely on physiological signals occurring during swallowing. In the first part, we will rely on signals acquired from the neck (e.g., electromyographic, acoustic), while in the second part, we will focus on electroencephalography (EEG) signals (and concurrently acquired deglutition signals from the neck) during swallowing.

A typical data acquisition and processing setup of deglutition signals acquired from the neck is shown in Figure 3. The first step is a choice of a sensor: accelerometers and microphones are typically used in most contributions [8]. Most recent contributions have showed that a combination of multiple sensors may be the most beneficial to obtain a comprehensive non-invasive assessment of events occurring during swallowing. However, the major issue here is that there is no consensus on sensors to be used for data acquisition, and sensors of varying frequency and bandwidth have been used in different contributions. In general, it is recommended to utilize sensors that have flat frequency response from 0 Hz to 3 kHz, and to sample these signals at 4 kHz. This is a sufficiently high sampling frequency given that most of the frequency content of these signals is below 500-600 Hz. Lastly, when utilizing multiple sensors to acquire deglutition signals, it is strongly recommended that these signals are time-synchronized in hardware via a data-acquisition board/card, that is, the same data-acquisition system is used to synchronously acquire multiple signals. Otherwise, important swallowing events that are very short in duration (shorter than 100 ms) may be misaligned and difficult to compare across different modalities.

The choice of a sensor impacts most of the subsequent signal processing steps. Raw EMG and acoustic signals during swallowing provide potentially valuable information but in a practically useless form, as raw signals cannot be quantitatively compared between subjects or across sessions. Therefore, pre- and post-processing steps are essential. The next typical task is pre-processing of deglutition signals. Here, we typically employ various filtering techniques and/or denoising to remove background/electrical noise, but also to annul the effects of a data acquisition apparatus (i.e., whitening). Filters are developed based on sensors and amplifiers used for data acquisition. Denoising is typically achieved via wavelet denoising. Upon completion of filtering and denoising operations or other required steps such as normalization, segmentation of swallowing recordings into multiple region of interests, typically individual swallows, is carried out. Several different algorithms have been proposed in the literature, mostly relying on some form of machine learning. Nevertheless, the exact steps of the pre-processing tasks differ significantly in contributions published, and this often poses a challenge when trying to re-create previously obtained results. Hence, the entire field would benefit from a systematic approach to pre-processing of physiological recordings obtained from the neck during swallowing.

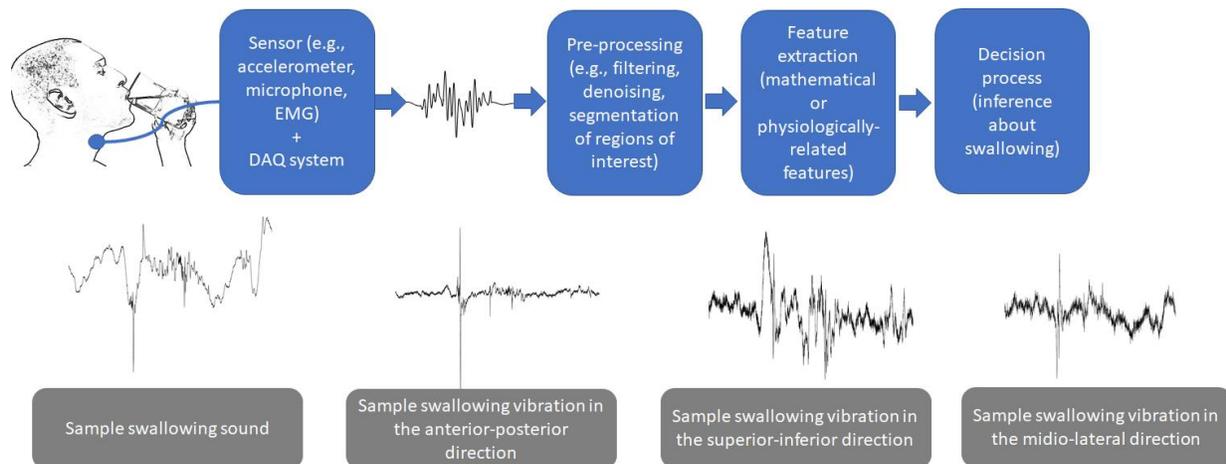


Figure 3 – An overview of a typical setup to acquire and process deglutition signals from the neck. Sample swallowing sound and swallowing vibrations signals in the three anatomical directions are shown as well.

The third step involves feature extraction from deglutition recordings. Most contributions relied on extracting mathematical features in time, frequency, or time-frequency domains. While this approach was warranted in initial contributions, as there was a lack of knowledge about basic properties of deglutition signals, we strongly believe that the field should move towards the extraction of physiologically-relevant features, that is, signal features that can be related to physiological events occurring during swallowing. Hence, we need to acquire simultaneous deglutition recordings during videofluoroscopy or endoscopy imaging to enable us to relate these signals and features to actual swallowing physiologic events. Newer methods based on deep learning may be useful for feature extraction, as these new methods can automatically extract features that maximize class differentiation.

The last step is typically a decision-making process during which we infer about the integrity of the swallowing function or swallowing tasks that were carried out during an

experimental procedure. In many instances, this decision process relies on a statistical analysis of features extracted in the previous step. In recent years, we have witnessed the development of various machine learning algorithms that can aid the decision process. These machine learning approaches mostly relied on differentiating swallowing safety/efficiency states. Nevertheless, there is a wide-open field for the development of machine learning algorithms that can not only infer about the state of the swallowing function, but even infer about swallowed food or drinks. Our recent review paper [8] showed that most machine learning methods have already been used from traditional Bayesian methods to neural networks.

Similar processing steps are taken when inferring about the brain activity via EEG during swallowing [9]. After acquiring EEG signals, one would typically start with pre-processing steps that include low-pass filtering with a cutoff frequency up to 128 Hz, a notch filter at 50/60 Hz and an artifact removal step involving the independent component analysis or other blind source separation algorithms. The pre-processing part can also include a segmentation step, where one identifies regions of interest (i.e., EEG activity during swallowing) for further analysis. The segmentation process can be aided by auxiliary signals such as cervical auscultation recordings (denoted by the blue box). However, to use cervical auscultation recordings during the EEG segmentation process, EEG recordings need to be synchronized with these cervical auscultation recordings, and this is typically achieved via a hardware system, such as a data-acquisition card.

The next step diverges depending on the analysis employed. On one side, researchers can utilize a feature-based analysis, where we attempt to extract features that we may think are relevant for the decision process. These features are mathematical features that are extracted from EEG recordings and often have no physiological meaning. The second approach relies on a network-based analysis where researchers rely on graph theory to establish brain networks. Here, these networks during swallowing can be established in two different ways: (a) during a swallowing process that would include multiple single swallows; or (b) on a swallow-by-swallow basis. The first approach is suitable when one desires to understand a global swallowing network, while the second approach is more suitable for understanding time-dependent changes in swallowing networks, which may be particularly of interest, when clinicians are attempting to understand the effects of various treatments on swallowing safety and efficacy. The network-based approach is also very interesting to the signal processing community, as it opens many interesting problems for the field of graph signal processing. In our own research, we use the vertex-frequency analysis (see a recent lecture note in this magazine [10]) to understand swallow-by-swallow changes in brain networks [11]. However, other graph signal processing approaches are anticipated to be suitable as well.

The last step involves various machine learning techniques, from traditional Bayes classifiers and support vector machines to the newest algorithms such as deep belief networks [12]. While various accuracies have been reported, we feel that most of those results are not generalizable for clinical use. In many cases, these contributions are proposed by signal processing practitioners with little or no understanding of clinical needs. Hence, such contributions are technologically elegant, but are of a small clinical value. Therefore, our signal processing

community needs to work more closely with clinicians to propose clinically relevant technological solutions.

6. Image processing approaches and challenges

Computational deglutition introduces several image processing challenges as well, that are associated either with videofluoroscopy/endoscopy or dynamic MR imaging and neuroimaging (e.g., fMRI) during swallowing. In this section, we briefly review some of these open challenges.

Image processing approaches have been historically constrained to human judgment. There is no dispute regarding the accuracy of human judgments of swallow kinematics, airway protection, residue patterns and swallow efficiency. However, too few clinicians receive advanced training in the performance of these judgments, and even then, reliability ratings can be variable. Efforts to standardize clinical decision-making have succeeded in increasing access of validated decision-making algorithms to clinicians. A penetration-aspiration scale was developed in 1996 to describe the extent of airway compromise during disordered swallowing on a swallow-by-swallow basis and possesses high reliability among trained judges [13]. Residue rating scales have also been developed that use relatively convenient anatomic landmarks with which to make judgments. The Modified Barium Swallow Impairment Profile (MBSImp) was recently developed to characterize seventeen components of oropharyngeal swallowing on a swallow-by-swallow basis and exhibits acceptable inter-rater reliability after training [14]. Other judgments that characterize motor integrity, such as the displacement of the hyoid bone during swallowing, are commonly made in clinical imaging studies and inferences regarding the summative motor functions producing airway closure and UES opening are made. However, these judgments are largely subjective and variable, unfortunately because the evidence indicating the range of typical displacement requires computerized analysis to characterize normal from abnormal.

Efforts to automate certain judgments from videofluoroscopic data are expanding. Currently, residue ratings are possible using a combined human-computer interface that exploits geometric relationships to quantify post-swallow pharyngeal residue (e.g., [15]). Likewise, hyoid bone tracking methods, in which machine learning is combined with expert human judgment to measure and quantify the completeness of hyoid displacement, are under investigation in many labs including our own.

There is a widespread need for algorithms that can aid clinicians in the analysis of videofluoroscopy/endoscopy images. Currently, such images (Figure 2, upper row) are analyzed manually on a frame-by-frame basis. As can be expected, such an approach is time consuming and prone to errors due to fatigue or expertise of a clinician conducting this analysis. The field currently needs algorithms to: (a) segment individual physiological landmarks (e.g., cervical vertebrae) or any transient objects (e.g., a bolus) from other objects present in images, (b) identify the beginning and end of swallows; and (c) identify swallowing safety and efficacy.

In recent years, several researchers, including our team, have also used MRI techniques to investigate swallowing function and neural activity during swallowing. These techniques offer substantial benefits in image quality but also come with their own challenges. Dynamic MRI of swallowing allows for better visualization of soft tissues than videofluoroscopy, and can even

provide insights on muscle integrity. Another advantage of dynamic MRI is that it does not require the use of ionizing radiation and regular food can be evaluated during swallows. Imaging speed used to be superior with videofluoroscopy (30 frames per second), but recently dynamic MR imaging can also achieve serial imaging rates of up to 26 frames per second or more providing increased temporal resolution. This is particularly critical for swallowing events as most are completed in less than a second. Data acquisition challenges that remain include the need to swallow in a supine position (most facilities do not have an upright magnet), and magnetic susceptibility differences that occur at interfaces between air and tissue, which are plentiful in the oropharynx. These artifacts can be successfully addressed either by using multiple shot acquisition sequences, or by using susceptibility corrective reconstruction algorithms during post-processing [16].

Unlike scales and tools designed for videofluoroscopic analysis (e.g., PA Scale, MBSImP), no similar standardized or validated tools exist to enable clinicians to complete respective measurements of swallowing events using MRI. To initiate MR image analysis, accurate registration and segmentation of swallow events and/or anatomical structures is a critical first step, as the number of volumes and slices acquired is large and the amount of anatomical displacements during swallows abundant. Recently, algorithms have been developed that allow semi-automatic segmentations of MRI volumes of the tongue and hyolaryngeal structures and enable faster calculations of displacement/deformation events. These approaches typically include identification of anatomical landmarks by experimenters and calculation of their movement and shape changes during swallowing using advanced statistical methods. Despite their promise, such techniques continue to be validated, and require substantial training and time to be completed. Therefore, at this time their clinical use is significantly limited.

Another popular MRI method that has been used in swallowing research includes task-related functional MRI (task fMRI), which allows us to noninvasively examine brain activations during swallows and has provided important insights on the neurophysiology of human swallowing. Image processing of fMRI data involves sophisticated pre- and post-processing steps as well. After acquiring fMRI images, pre-processing steps would typically include brain extraction, removal of first volumes that correspond to the stabilization period of the magnetic signal, despiking, slice-timing correction, motion correction, spatial smoothing, and bandpass filtering (see Figure 4). Multistage registration and normalization is also performed to register the data on standard anatomical atlases. Currently these steps are automated or semi-automated and performed via a pipeline of commands or GUI systems provided in well-developed fMRI analysis programs such as the Analysis of Functional NeuroImages or FMRIB Software Library.

To compute task onset timings that are used in the post-processing analysis, tasks performed during the fMRI scans are often cued by visual or audio stimuli. The subject must perform the task in strict compliance with the stimulus. Secondary-monitoring devices, including surface electrodes or pneumographic belts placed around the neck, are needed to ensure the subjects' swallows comply with the stimuli [16]. During post-processing, the task onsets are convolved with the canonical hemodynamic response function for use in GLM models to analyze each subject's activation during the scan. Contrasts between GLM model parameters of interest

are then used to compare activations between different tasks (when more than one tasks are examined). Multiple comparisons corrections are further necessary because the large number of brain voxels significantly increases the false positives for any given statistical threshold. Whole brain and region-of-interest (ROI) or seed-based analyses are both widely used. For an example of a single subject whole brain analysis of swallowing brain activations, see Figure 4 (D).

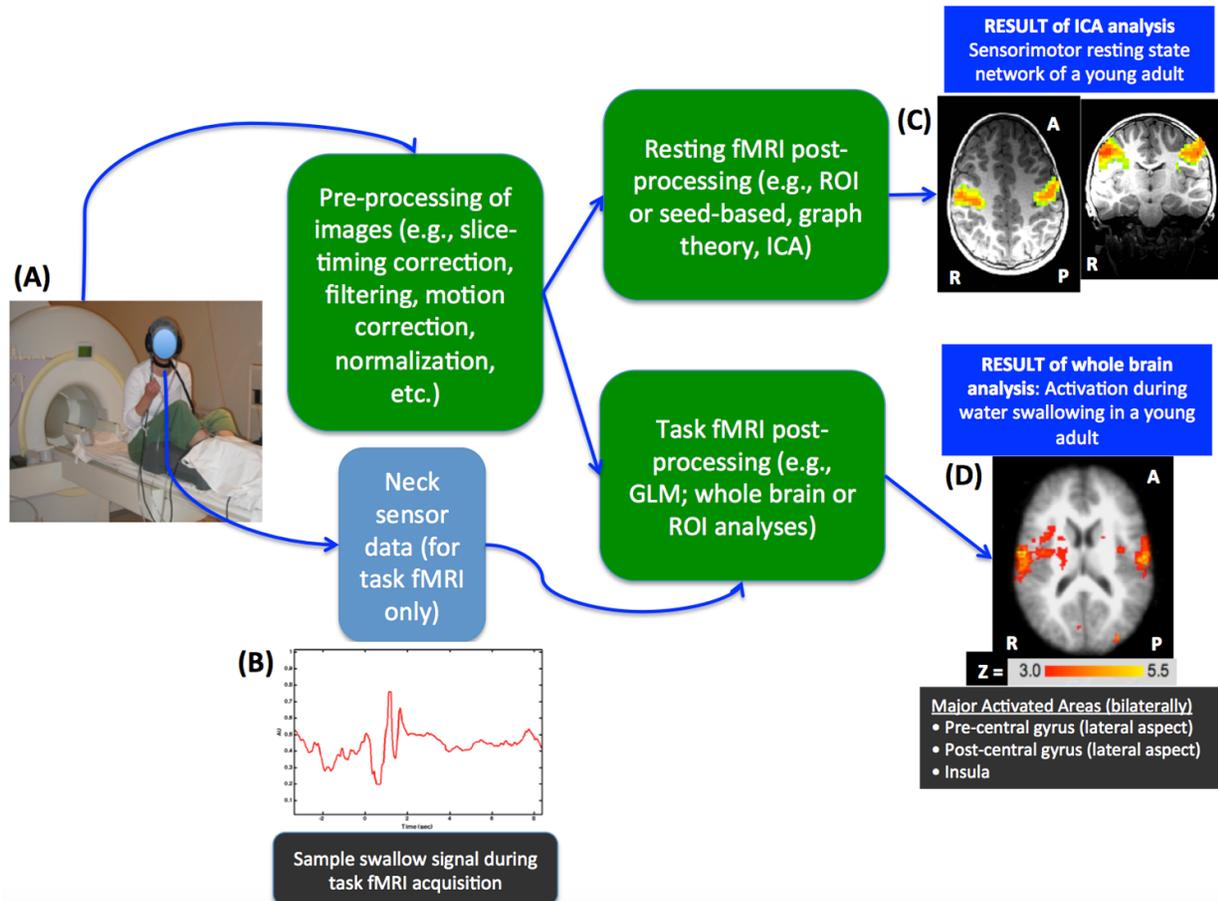


Figure 4 – Setup to acquire and process fMRI images during resting state and during swallowing. (A) Subject shown wearing respiratory bellows around neck over the thyroid cartilage (to capture swallow signal during swallows) before experiment initiation. (B) Time course of the output of the bellows for a water-swallowing trial (red), for a single subject. (C) Results of ICA analysis of a resting-state fMRI scan of a young adult male showing the symmetrically activated sensorimotor network at rest. (D) Results of whole brain task fMRI analysis of a young adult male showing areas of significant activation during water swallowing. The neurological images are shown in radiological convention (the right hemisphere is shown on the left). A = Anterior, P = Posterior, R = Right Hemisphere.

It is important to highlight that to improve signal interpretation accuracy, task-related fMRI results should be interpreted relative to another comparison condition (e.g., rest). Further, during swallowing-specific experiments, motion related artifacts are very common as swallowing includes movements of the neck/throat during scanning, and need to be carefully examined, eliminated, or post-processed [17].

An alternative to task-related fMRI paradigms is the use of resting-state functional connectivity MRI (resting-state-fcMRI). Resting state fcMRI allows us to investigate the functional connections of brain areas at rest and correlate that information with behavioral measures obtained outside the scanner. It is based on the fact that areas of the brain that are functionally

related (even if they are far apart) show low frequency fluctuations of the BOLD (Blood Oxygenation Level-Dependent) signal that have the exact same temporal patterns. As such, resting-state fMRI has helped us identify several resting-state networks in the brain that are altered or even absent in individuals with diseases or in older age compared to healthy young adults. The advantage of resting-state paradigms for studying populations with dysphagia is that patients are not required to swallow in the magnet (in the supine position), a task that is frequently challenging for patients with dysphagia. Pre-processing steps are almost identical to the task based pre-processing analysis with the addition of nuisance factors regression (CSF and white matter regressors) to further improve data quality. For post-processing of resting-state scans, popular methodologies include advanced mathematical models, such as graph theory, independent component analyses techniques, and clustering algorithms. The contribution of resting-state fMRI to our understanding of the neural control of swallowing can only be indirect and remains experimental at this time but holds a lot of promise.

A promising new imaging technology to comprehensively image swallowing physiology and neurophysiology and alleviate some of the challenges of task fMRI was examined by one of our authors and her collaborators [16]. This technology, known as SimulScan, allows the simultaneous dynamic imaging of the oropharyngeal area and functional imaging of the brain during swallowing and provides the ability, for the first time, to directly and simultaneously evaluate both central (brain) and peripheral (oropharyngeal) physiological signals during swallowing. This technique has been used successfully to image natural, uncued, spontaneous swallows and brain activation associated with these swallows in healthy young adults [16] but requires further validation.

Although sophisticated algorithms are now available for the analysis of fMRI images, extensive training and expertise with this methodology are necessary and costs remain prohibitive for clinical use. Therefore, its direct clinical application at this time is questionable, though its contribution to our understanding of the neural control of swallowing and the neuroplastic adaptations needed for functional swallowing is substantial and will continue to increase.

For dynamic MRI, specifically, which in time may be able to replace videofluoroscopy, significant more work from both the image processing and clinical communities in synchrony is needed. For this method as well, the field currently needs algorithms and models to improve segmentation and automated analysis of events and help predict swallowing pathologies and ultimately treatment outcomes.

7. Future directions in computational deglutition

To foster the development of computational deglutition as a field, we encourage researchers to share datasets with other researchers. Specifically, we invite and encourage the community to produce clinical protocols and consent forms that include a clause about publicly sharing de-identified datasets to foster the growth of computational deglutition as a field. Furthermore, we anticipate that such publicly available datasets will also result in faster standardization of instrumentation and development of algorithms that can improve healthcare and patient outcomes.

Over the years, we have often witnessed the signal/image processing community developing computationally or mathematically elegant solutions with a limited practical usability. Computational researchers interested in this new field should work closely with clinicians to ensure that new developments are addressing clinically relevant problems. We, as the community, should also strive to ensure that our new algorithms are applicable across different patients and patient groups, rather than in a limited number of patients. Similarly, the clinical community should work closely with computational researchers to understand how to acquire data in a systematic view to ensure that the collected data is useful for further algorithmic developments.

Authors

Ervin Sejdić (esejdic@ieee.org) is an associate professor at the University of Pittsburgh, Pennsylvania. His research interests include biomedical signal processing, swallowing, and gait. He received the Presidential Early Career Award for Scientists and Engineers in 2016 and the National Science Foundation CAREER Award in 2017. He is a Senior Member of the IEEE.

Georgia A. Malandraki (malandraki@purdue.edu) is an associate professor at Purdue University and a board-certified specialist in swallowing disorders. She received the Early Career Contributions in Research Award by the American-Speech-Language-Hearing Association in 2011. Her research focuses on neuroimaging, neurorehabilitation of swallowing, and telehealth.

James L. Coyle (jcoyle@pitt.edu) is a professor at the University of Pittsburgh and a board-certified specialist in swallowing disorders. He is a Fellow of the American Speech-Language and Hearing Association and received the University of Pittsburgh Chancellor's Distinguished Teaching Award in 2016.

Acknowledgements

Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute Of Child Health & Human Development of the National Institutes of Health under Award Number R01HD092239. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- [1] N. Bhattacharyya, "The prevalence of dysphagia among adults in the United States," *Otolaryngology-Head and Neck Surgery*, vol. 151, no. 5, pp. 765–769, 2014.
- [2] J. Logemann, "The evaluation and treatment of swallowing disorders," *Otolaryngology and Head and Neck Surgery*, vol. 6, no. 1, pp. 395–400, 1998.
- [3] R. Martino, F. Silver, R. Teasell, M. Bayley, G. Nicholson, D. Streiner, and N. Diamant, "The Toronto bedside swallowing screening test (TOR-BSST): Development and validation of a dysphagia screening tool for patients with stroke," *Stroke*, vol. 40, no. 2, pp. 555–561, February 2009.
- [4] J. L. Coyle, "The clinical evaluation: A necessary tool for the dysphagia sleuth," *Perspectives on Swallowing and Swallowing Disorders (Dysphagia)*, vol. 24, no. 1, Feb. 2015.

- [5] R. Ding, C. Larson, J. Logemann, and A. Rademaker, "Surface electromyographic and electroglottographic studies in normal subjects under two swallow conditions: Normal and during the Mendelsohn maneuver," *Dysphagia*, vol. 17, no. 1, pp. 1–12, January 2002.
- [6] J. L. Coyle and J. Robbins, "Assessment and behavioral management of oropharyngeal dysphagia," *Otolaryngology and Head and Neck Surgery*, vol. 5, no. 1, pp. 147–152, 1997.
- [7] J. L. Coyle, *Videofluoroscopy: A Multidisciplinary Team Approach*. San Diego, CA, USA: Plural Publishing, 2012, ch. Biomechanical Analysis, pp. 107–122.
- [8] J. M. Dudik, J. L. Coyle, and E. Sejdic, "Dysphagia screening: Contributions of cervical auscultation signals and modern signal-processing techniques," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 4, pp. 465–477, Aug 2015.
- [9] I. Jestrovic, J. L. Coyle, and E. Sejdic, "Decoding human swallowing via electroencephalography: a state-of-the-art review," *Journal of Neural Engineering*, vol. 12, no. 5, pp. 051001–1–15, Oct. 2015.
- [10] L. Stankovic, M. Dakovic, and E. Sejdic, "Vertex-frequency analysis: A way to localize graph spectral components," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 176–182, 2017.
- [11] I. Jestrovic, J. L. Coyle, and E. Sejdic, "Differences in brain networks during consecutive swallows detected using an optimized vertex–frequency algorithm," *Neuroscience*, vol. 344, pp. 113 – 123, 2017.
- [12] F. Movahedi, J. L. Coyle, and E. Sejdic, "Deep belief networks for electroencephalography: A review of recent contributions and future outlooks," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 3, pp. 642–652, May 2018.
- [13] J. C. Rosenbek, J. A. Robbins, E. B. Roecker, J. L. Coyle, and J. L. Wood, "A penetration-aspiration scale," *Dysphagia*, vol. 11, no. 2, pp. 93–98, 1996.
- [14] B. Martin-Harris, M. B. Brodsky, Y. Michel, D. O. Castell, M. Schleicher, J. Sandidge, R. Maxwell, and J. Blair, "MBS measurement tool for swallow impairment - MBSImp: establishing a standard," *Dysphagia*, vol. 23, no. 4, pp. 392–405, 2008.
- [15] W. G. Pearson, S. M. Molfenter, Z. M. Smith, and C. M. Steele, "Image-based measurement of post-swallow residue: The normalized residue ratio scale," *Dysphagia*, vol. 28, no. 2, pp. 167–177, 2013.
- [16] T. L. Paine, C. A. Conway, G. A. Malandraki, and B. P. Sutton, "Simultaneous dynamic and functional MRI scanning (SimulScan) of natural swallows," *Magnetic Resonance in Medicine*, vol. 65, no. 5, pp. 1247–1252, 2011.
- [17] G. A. Malandraki, S. Johnson, and J. Robbins, "Functional MRI of swallowing: from neurophysiology to neuroplasticity," *Head and Neck*, vol. 33, no. S1, pp. S14–S20, Oct. 2011.