Model-based Identification of Motion Sensor Placement for Tracking Retraction and Elongation of the Tongue

Yikun K. Wang · Martyn P. Nash · Andrew J. Pullan · Jules A. Kieser · Oliver Röhrle

Received: date / Accepted: date

Abstract Electromagnetic articulography (EMA) is designed to track facial and tongue movements. In practice, the EMA sensors for tracking the movement of the tongue’s surface are placed heuristically. No recommendation exists. Within this paper, a model-based approach providing a mathematical analysis and a computational-based recommendation for the placement of sensors, which is based on the tongue’s envelope of movement, is proposed. For this purpose, an anatomically-detailed Finite Element (FE) model of the tongue has been employed to determine the envelope of motion for retraction and elongation using a forward simulation. Two optimality criteria have been proposed to identify a set of optimal sensor locations based on the pre-computed envelope of motion. The first one is based on the assumption that locations exhibiting large displacements contain the most information regarding the tongue’s movement and are less susceptible to measurement errors. The second one selects sensors exhibiting each the largest displacements in the anterior-posterior, superior-inferior, medial-lateral, and overall direction.

The quality of the two optimality criteria is analysed based on their ability to deduce from the respective sensor locations the corresponding muscle activation parameters of the relevant muscle fibre groups during retraction and elongation by solving the corresponding inverse problem. For this purpose, a statistical analysis has been carried out, in which sensor locations for two different modes of deformation have been subjected to typical measurement errors. Then, for tongue retraction and elongation, the expectation value, the standard deviation, the averaged bias and the averaged coefficient of variation have been computed based on 41 different error-afflicted sensor locations. The results show, that the first optimality criteria is superior to the second one and that the averaged bias and averaged coefficient of variation decreases when the number of sensors is increased from 2, 4 to 6 deployable sensors.

Keywords Tongue Modelling, Sensor Placement Optimisation, Motion Tracking, Retraction, Elongation, Finite Element Modelling, Soft Tissue Modelling, Parameter Estimation, Inverse Problem

1 Introduction

The tongue is a muscle that is composed of several (partially) interlaced muscle fibre groups. The ability to selectively and independently activate the different muscle fibre groups provide the basis for its complex movements. Several clinical studies provide evidence that the dysfunction of single muscle fibre groups can be linked to diseases like dysphagia (Fujiu et al, 1995), dental malocclusion (Cheng et al, 2002), or sleep apnoea (Oliven et al, 2007; Proffit et al, 2007). A better understanding of the underlying activation principles of muscle fibre groups during specific tasks would provide
a significant step forward in analysing and treating diseases associated with unusual tongue movements. From a basic research point of view, a better understanding of the tongue’s function can be achieved either by (i) investigating muscle fibre activity using electromyographic (EMG) recordings or (ii) tracking the tongue’s surface during movement.

A common technique of measuring muscle activity is based on surface EMG recordings (e.g., Baer et al, 1988). However, the complex arrangement of the muscle fibre groups makes it virtually impossible to deduce the activation of particular muscle fibres/muscle fibre groups from such recordings. While needle EMG recordings have previously been performed to investigate the activity of single muscle fibre groups within the tongue (e.g., Sauerland and Harper, 1976), it is again the complex and interlaced muscle fibre arrangement that poses significant challenges. Space-confined (needle) arrays would need to be designed in order to be able to investigate the co-activation of different muscle fibre groups.

The motion of the tongue’s surface can be tracked using image-based or marker-based methodologies. Image-based approaches appeal to X-ray imaging techniques (e.g., Nozaki et al, 2007; Saitoh et al, 2007), standard, tagged, or diffusion magnetic resonance imaging (e.g., Stone et al, 2004; Gilbert et al, 2007), and ultrasound imaging techniques (e.g., Peng et al, 2000). While most of these imaging techniques provide three-dimensional structural information on the tongue, they typically fail to produce images of acceptable spatial and temporal resolution to track the tongue’s movement during specific tasks. If the tracking of internal structures or particular function is secondary, or if one can deduce from the tongue’s surface deformations the internal state (inverse problem), one can also appeal to marker-based tracking methods. Such methods are particularly popular in analysing the kinetics of movement. (The reader is referred to Moeslund and Granum (2001) and Moeslund et al (2006) for an extensive survey on different motion tracking systems.) In movement biomechanics, the most common motion capture systems are based on optically tracking surface markers, which, however, are of limited use for tracking the tongue’s surface as the visibility of surface markers on a tongue is obstructed by the bony walls of the oral cavity. Less common motion tracking systems are based on optically tracking surface markers, which, however, are of limited use for tracking the tongue’s surface as the visibility of surface markers on a tongue is obstructed by the bony walls of the oral cavity. Therefore, electro-magnetic articulographs have been extensively used in the past to track the motion of the tongue and other facial features (Jiang et al, 2002; Steele and Van Lieshout, 2009; Goozee et al, 2000; Katz et al, 1999; Steele and van Lieshout, 2004).

In practice, sensors for tracking the movement of the tongue are heuristically placed on the tongue’s surface. No recommendation exists. Within this paper, we propose a model-based approach to provide a mathematical analysis and a computational-based recommendation for the placement of sensors based on the tongue’s envelope of movement. To determine the tongue’s envelope of movement during specific tasks, principles of forward dynamics (prescribing the level of activity for single muscle fibre groups) have been applied to a generic, anatomically-detailed computational model of the tongue (Wang, 2008). Based on the calculated movements, we selected 2, 4, or 6 locations containing the most information to accurately capture the movement of the tongue. The reduction of the dimensionality of the sensor locations is necessary due to physical constraints of the recording method, e.g. only a small number of recording channels are typically available.

2 Methods

The aim of this study is to provide a model-based recommendation for sensor locations on the tongue during specific movements. Due to limitations on the number of recording channels of a motion capture system (here, a standard AG500), we limit ourselves to a maximum of 6 sensors that can be placed on the accessible parts of the tongue surface.

2.1 Computational tongue model

The computational model of the tongue is generic and based on the female Visible Human data set (Spitzer et al, 1996). The geometrical representation of the tongue is based on a 120-node, 64-element tri-cubic Hermite finite element (FE) mesh that contains the description of 11 different muscle fibre groups (i.e., the intrinsic muscle fibre groups such as the transversus (t), verticalis (v), superior longitudinalis (sl), inferior longitudinalis (il), the extrinsic muscle fibre groups such as the genioglossus anterior (ga), genioglossus posterior (gp), styloglossus (s), hyoglossus (h), and external muscle fibre groups such as geniohyoid (g), digastric (d), and mylohyoid (m)) and serves as reference configuration for the continuum mechanical model.

The mechanical behaviour of the tongue is described by the governing equations of finite elasticity, which
are discretised using tri-cubic Hermite FE basis functions. The fibre direction of each muscle fibre group, \( i \), is denoted at each Gauss point (64 Gauss points per element) by the three-dimensional vector \( \mathbf{a}_i \). For each muscle fibre group, a volume fraction \( d_i \) is introduced to account for the fact that multiple muscle fibre groups can be present at a Gauss point, which represent within the theory of continuum mechanics a homogenised representative volume. We assume that no interactions between the different muscle fibre groups exist. The tongue model used within this paper (cf. Fig. 1) has been created similarly to the one in Wang (2008). Further, the tongue is assumed to be incompressible. Within this paper, the incompressibility constraint is enforced through a Lagrange multiplier \( p \), which can be interpreted as hydrostatic pressure (cf. Spencer (2004); Nash and Hunter (2000)).

In general, the tongue tissue is regarded as a multi-fibre reinforced, incompressible hyper-elastic material (cf. Gerard et al, 2005). The constitutive description of the tongue is given by the 2nd Piola-Kirchhoff stress tensor, \( \mathbf{S}_{\text{tissue}} \), which is an additive composition of the tongue’s ground matrix, \( \mathbf{S}_{\text{matrix}} \), and the stress contribution of each muscle fibre group:

\[
\mathbf{S}_{\text{tissue}} = \mathbf{S}_{\text{matrix}} + 2 \sum_{i \in F} \frac{d_i}{N} \mathbf{g}_i \left( \alpha_i, \lambda_i, \sigma_i^{\text{max}} \right) \mathbf{a}_i \otimes \mathbf{a}_i,
\]

where \( F = \{ t, v, s, \ldots \} \) is the set containing all indices of the muscle fibre groups, \( \mathbf{a}_i \) the fibre direction, \( \lambda_i^2 = \mathbf{a}_i \cdot \mathbf{C}_s \mathbf{a}_i \) is the squared fibre stretch (\( \mathbf{C}_s \) is the right Cauchy-Green solid deformation tensor), \( \alpha_i \in [0, 1] \) an activation parameter for muscle fibre group \( i \) and \( \sigma_i^{\text{max}} \) the maximal contractile stress that can be generated by fibres of muscle group \( i \). The fibre stretch, the maximal contractile stress and the activation parameter of a muscle fibre group are related to each other by the following relationship:

\[
\sigma_i(\alpha_i, \lambda_i, \sigma_i^{\text{max}}) = \sigma_i^{\text{max}} \left[ f_{i}^{\text{passive}}(\lambda_i) + \alpha_i f_{i}^{\text{active}}(\lambda_i) \right],
\]

where \( f_{i}^{\text{passive}} \) and \( f_{i}^{\text{active}} \) are the typical normalised length-force relationships for the passive and active behaviour of skeletal muscle tissue, respectively (cf., for example, Blemker et al (2005), Röhle and Pullan (2007), or Röhle (2010)). In the absence of experimental tests to determine the passive material response (of the tongue’s ground matrix, it is modelled as a Mooney-Rivlin material with coefficients \( c_1 = 0.375 \text{kPa} \) and \( c_2 = 0.175 \text{kPa} \). Further, \( \sigma_i^{\text{max}} \) is chosen to be 150kPa. The parameter \( c_1 \) is hereby associated with the first invariant, \( I_1 \) and \( c_2 \) with the second invariant. Furthermore, Section 4 provides a detailed discussion on the choice and derivation of the Mooney-Rivlin material parameters.

The nodes, at which the boundary conditions are applied, are depicted in Fig. 1(a) as (blue) cubes, (yellow) spheres, and (green) diamonds. The (blue) cubes represent the tongue’s attachment to the jaw and are therefore spatially fixed (zero displacement). The attachment of the tongue to the hyoid bone (yellow spheres) is modelled by fixing all derivatives ensur-
ing that the shape of the hyoid bone is maintained while displacements are allowed. The attachment of the palatoglossus muscle to the lateral sides of the tongue dorsum forms the connection between the tongue and its surrounding tissues. This connection has been modelled by fixing the displacements of the nodes representing the attachment area of the palatoglossus (green diamonds), whilst not constraining their derivatives.

The continuum mechanical model is solved using the software package CMISS\(^1\) (Continuum Mechanics, Image analysis, Signal processing and System Identification), which is developed at the University of Auckland, NZ and has been extensively used for many biomechanical simulations.

2.2 Optimisation of sensor placement

The main aim of this section is to describe the methodology to determine an optimal set of sensor locations on the tongue’s surface during retraction and elongation. Retraction and elongation of the tongue have been chosen within this context, since the muscle fibre groups involved in these movements are known. Gilbert et al (2007) and Napadow et al (1999a,b) report that retraction is caused through the contraction of the transversus and styloglossus muscle fibre groups, while elongation is caused by contractions of the transversus and verticalis muscle fibre groups. The tongue’s envelope of motion can then be determined by the above-described, generic, continuum-mechanical tongue model through solving the resulting movement for a representative sample of different combinations of activation levels. Then, an optimal set of sensor locations can be chosen based on the computed tongue’s envelope of motion using an optimisation criterion reflecting key movement characteristics. The following steps outline the methodology in detail.

2.2.1 Computing the tongue’s envelope of motion

For each type of movement, e.g. retraction and elongation, the tongue’s envelope of motion is determined by utilising the generic continuum-mechanical model outlined in Section 2.1 and principles of forward dynamics, i.e., by choosing a specific set of activation parameters, \(\alpha\), in Eq. 2. Therefore, the envelope of motion for the case of retraction is computed by the forward problem through calculating the resulting mode of deformation given different levels of activation for the transversus muscle fibre group, \(\alpha_t\), and the styloglossus muscle fibre group, \(\alpha_s\). Within this context, the respective levels of activation are multiples of \(\Delta \alpha = 0.1\) and range between 0 (inactive) and 1 (fully activated). The activation levels for the remaining muscle fibre groups are set to 0 (inactive). The set of all computed modes of deformation form the envelope of movement. The same procedure is followed to determine the envelope of motion for elongation. In this case, the levels of activation for all muscle fibre groups remain 0 except for the transversus muscle fibre group, \(\alpha_t\), and the verticalis muscle fibre group, \(\alpha_v\), which are again chosen to be multiples of \(\Delta \alpha = 0.1\) and to range between 0 and 1.

2.2.2 Potential sensor locations

Sensors can only be placed on the external surface of the tongue. Within this work, a total of 48 potential sensor locations have been evenly distributed over half of the tongue’s surface. The number of potential sensor locations correspond to four nodes per surface element face and are generated with respect to the undeformed state (reference configuration). Further, each potential sensor location is identified with a unique numeric index. The set of all sensor locations is denoted by \(\mathcal{S}\).

As a consequence of the general theory of continuum mechanics, each material point \(P\) in the reference configuration (denoted by \(\mathbf{X}\)) has its unique position over time/level of activation (denoted by \(\mathbf{x}(\alpha)\)) being determined by the placement function \(\chi\), where

\[
\mathbf{x}(\alpha) = \chi(P, \alpha), \quad \mathbf{X} = \chi(P, \alpha_0)
\]  

and \(\alpha_0 \equiv 0\). Hence, the displacement at a material point \(P\) is introduced by the difference between the actual position \(\mathbf{x}(\alpha)\) and its initial position \(\mathbf{X}\),

\[
\mathbf{u}(P, \alpha) = \chi(P, \alpha) - \chi(P, \alpha_0).
\]

Hence, \(\chi\) defines for all \(\alpha\) the motion and the displacement of each sensor in \(\mathcal{S}\).

2.2.3 Optimality criteria for selecting sensor locations

Two different optimality criteria for selecting sensor locations on the surface of the tongue have been considered within this work. The first one is based on the assumption that locations exhibiting large displacements contain the most information regarding the tongue’s movement and are less susceptible to measurement errors. The second optimality criterion is also based on movement (maximal displacements) of a potential sensor location, however, its objective function aims to distinguish between movements in anterior-posterior, superior-inferior and medial-lateral direction.

\(^1\) http://www.cmiss.org
The set of all potential sensor combinations consisting of \( n \) different sensors \( \alpha_i \in S \) can be described by

\[
\text{SL}(n) = \left\{ (\alpha_1, \ldots, \alpha_n) \mid \alpha_i \in S, \alpha_i \neq \alpha_j \text{ and } i \neq j \right\}.
\]

The cardinality of \( \text{SL}(n) \),

\[
\text{card} \left( \text{SL}(n) \right) = \binom{N}{n} = \frac{N!}{n!(N-n)!},
\]

denotes the number of possible sensor combinations that could be placed on half of the tongue’s surface \((N = 48)\).

**Optimality Criterion I:** Movement sensitivity of a set of sensor locations can be measured by the Fisher Information matrix. The concept of using the Fischer Information matrix to determine optimal sets of sensor locations is not new. It has previously been applied to many engineering-type applications, e.g., in civil engineering for structural monitoring (Papadimitriou et al., 2000), in environmental engineering for air quality monitoring (Ucinski, 2000), or in aerospace engineering for on-orbit identification and correlation of large space structures (Kanner, 1991).

For a \( n \)-variate multivariate normal distribution with constant covariance matrix, \( \Sigma \), the Fischer Information matrix, \( M(\alpha, \beta) \) with \( \alpha \in [0, 1]^{11} \) \((jab\in [0, 1] \times [0, 1] \times \ldots \times [0, 1])\) and \( \beta \in \text{SL}(n) \), is defined by

\[
M(\alpha, \beta) = \left( \frac{\partial \delta^2(\alpha)}{\partial \alpha} \right)^T \Sigma^{-1} \left( \frac{\partial \delta^2(\alpha)}{\partial \alpha} \right),
\]

where

\[
\delta^2(\alpha) = \left[ u(\alpha_1, \alpha), \ldots, u(\alpha_n, \alpha) \right]^T
\]

and \( \Sigma^{-1} = (1/\sigma^2) I \). The covariance matrix is assumed to be diagonal since the error components in all directions are assumed to be independent of each other. In computing the Fischer Information matrix, the derivative with respect to the level of activation, \( \alpha \), in Eq. 7 is computed numerically using a Newton’s difference quotient with \( \Delta \alpha = 0.1 \). Needless to say that the differentiation is only performed for the non-zero activation parameters, e.g., \( \alpha_i \) and \( \alpha_s \), for investigating retraction.

To maximise the information contained within the envelope of movement over a pre-described parameter space (here the level of activation for the respective muscle fibre groups), the optimal set of sensor locations is found solving the following optimisation problem:

\[
\beta^* = \arg \max_{\beta \in \text{SL}(n)} \left[ \det \left( \sum_{\alpha_i=0}^{1/\Delta \alpha} \sum_{\alpha_j=0}^{1/\Delta \alpha} M(\Delta \alpha, \alpha, \beta) \right) \right],
\]

where \( \alpha_i \) and \( \alpha_j \) correspond to the respective levels of activation of the muscle fibre groups involved in retraction or elongation.

**Optimality Criterion II:** The second optimality criterion aims to distinguish between movements in anterior-posterior, superior-inferior and medial-lateral direction. Therefore, the first three markers shall be chosen based on the maximal displacements in anterior-posterior (AP), superior-inferior (SI) and medial-lateral (ML) direction respectively:

\[
\alpha_1 = \arg \max_{\alpha \in S} \sum_{\alpha_i=0}^{1/\Delta \alpha} \sum_{\alpha_j=0}^{1/\Delta \alpha} \| u_{\text{AP}}(\alpha, \Delta \alpha, \alpha) \|_2,
\]

\[
\alpha_2 = \arg \max_{\alpha \in S \setminus \{\alpha_1\}} \sum_{\alpha_i=0}^{1/\Delta \alpha} \sum_{\alpha_j=0}^{1/\Delta \alpha} \| u_{\text{SI}}(\alpha, \Delta \alpha, \alpha) \|_2,
\]

\[
\alpha_3 = \arg \max_{\alpha \in S \setminus \{\alpha_1, \alpha_2\}} \sum_{\alpha_i=0}^{1/\Delta \alpha} \sum_{\alpha_j=0}^{1/\Delta \alpha} \| u_{\text{ML}}(\alpha, \Delta \alpha, \alpha) \|_2.
\]

For selecting \( n = 4 \) sensors, sensor location \( \alpha_4 \) is chosen such that

\[
\alpha_4 = \arg \max_{\alpha \in S \setminus \{\alpha_1, \alpha_2, \alpha_3\}} \sum_{\alpha_i=0}^{1/\Delta \alpha} \sum_{\alpha_j=0}^{1/\Delta \alpha} \| u(\alpha, \Delta \alpha, \alpha) \|_2.
\]

For \( n > 4 \), one restricts the set of potential sensor combinations, \( \text{SL}(n) \), in such a way that each element of the restricted set contains all of the 4 previously selected sensors. The restricted set will then be used within the above-described selection procedure (Optimality Criterion I) to determine the remaining sensors. In case \( n < 4 \), a selection of the criterion given by Eqs. 10 - 13 is necessary. However, within this work, Optimality Criterion II is only used for selecting a set of 4 sensors.

### 2.3 Quality measures for selected sensor locations

Determining the quality of a chosen set of sensors/markers is challenging. Further, the ultimate goal of motion tracking is not only to track a few sensors, but to predict the movement of any point within the tracked body, here the entire tongue. To do so, one needs to estimate the movement of the entire body based on the movement of a few markers. This can be achieved by determining the activation levels of the muscle fibre groups causing the particular movement, i.e., by solving the inverse problem. The solution of the inverse problem, here computing the activation levels of the respective muscle fibre groups, and the continuum-mechanical model can then be used to determine the motion of the entire tongue (forward problem).
Given a set of sensor locations $\beta \in \text{SL}(n)$ and the (measured) locations of these sensors for a given deformation, denoted by $\tilde{\beta} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n)$, the inverse problem is given by

$$\tilde{\alpha} = \arg \min_{\alpha \in \mathcal{A}} \sum_{i=1}^n \| \tilde{\alpha}_i - \chi(\tilde{\alpha}_i, \alpha) \|_2,$$

(14)

with $\mathcal{A} = [0, 1]^{11}$. Numerically, this inverse problem is solved by the Trust-Region-Reflective algorithm as implemented in MATLAB’s (V7.8, R2009a) optimisation toolbox. The Trust-Region Reflective algorithm is based on approximating the objective function with its Taylor series truncated after the $2^{\text{nd}}$-order derivative term (Hessian). A Newton-like algorithm is then applied to obtain the optimal solution minimising the function value of the truncated Taylor series. For more details about trust-region algorithms, the reader is referred to Moré and Sorensen (1983) and Yuan (2000) and references therein. The algorithm is stopped, if the objective function values for two consecutive iterates, i.e., $\alpha_k$ and $\alpha_{k+1}$, differ by less than $10^{-5}$. To compute for a specific $\alpha_k$, the deformation of the tongue, here $\chi(\tilde{\alpha}_i, \alpha_k)$ for $i = 1, \ldots, n$, CMISS is directly linked to MATLAB. Note, the optimisation problem as given in Eq. 14, is not an 11-dimensional optimisation problem. In case of retraction and elongation, all $\alpha_i$ except for the activation levels of 2 muscle fibre groups are 0. This restriction is taken into account when solving Eq. 14.

Further, to measure the robustness of a particular set of sensor locations, i.e. for a $\beta \in \text{SL}(n)$, the respective sensor locations, $\tilde{\alpha}_i = \chi(\tilde{\alpha}_i, \alpha)$ with $\tilde{\alpha}_i \in \beta$ and a prescribed level of activation $\alpha$, are subjected to a motion-tracking-system-specific measurement error, i.e.,

$$\tilde{\alpha}_i = \tilde{\alpha}_i + \epsilon_i^k,$$

(15)

where $\epsilon_i^k = [e_i^k, e_i^2, e_i^3]^T$ and $\epsilon_i^k \sim N(\mu, \sigma^2)$ with $\mu = 0$ and $\sigma^2 = 0.49$. The choice of the standard deviation is based on previously reported error values for the Articulograph AG500 (cf., Yunusova et al, 2009). As Yunusova et al (2009) do not link their error analysis with any particular probability distribution, a Gaussian distribution is assumed within this work. Next, the inverse problem is solved for sufficiently many $\beta^k = (\tilde{\alpha}_1^k, \ldots, \tilde{\alpha}_n^k)$ such that a statistically relevant expectation value and standard deviation for the estimates of the level of activation (the solution of the inverse problem) can be computed.

The robustness of a set of sensor locations can then be investigated by comparing, for a $\beta \in \text{SL}(n)$, the computed expectation value with the original activation level, $\alpha$, that has been chosen to determine the “exact” sensor locations, $\tilde{\alpha}_i$, in the deformed configuration. Clearly, the better the expectation value matches the originally chosen activation level, the better the set of sensor locations. A further measure for comparing different sets of sensor locations is the standard deviation. The smaller the standard deviation of a particular set of sensor locations, the more robust, and hence the less error-prone, the set of sensor locations.

Hence, the quality of a selected set of sensor locations is evaluated in this work based on the robustness of the inverse method to determine the activation levels of the muscle fibre groups.

3 Numerical Results

All numerical results are based on the assumption that the level of activation is constant for the entire muscle fibre group. Fig. 2 depicts a specific mode of deformation for tongue retraction, e.g. by choosing $\alpha_t = 1.0$ and $\alpha_s = 1.0$ (cf. Fig. 2(a)), and for elongation, e.g. by choosing $\alpha_t = 1.0$ and $\alpha_s = 1.0$ (cf. Fig. 2(b)). Further, a symmetric arrangement of the muscle fibre groups within the tongue is assumed. The symmetry assumption leads to symmetric contractions. For the cases investigated within this context, i.e., retraction and elongation, this is a realistic setting. The symmetry assumption also provides justification to restrict the analysis (as discussed in Section 2.3) to one half of the tongue. The results of this analysis is presented for the case of retraction in Sections 3.1 and for the case of elongation in 3.2.

3.1 Sensor optimisation for retraction (half tongue)

First, retraction is considered. For each $\beta \in \text{SL}(n)$, the objective function for Optimality Criterion I,

$$\mathcal{M}(\beta) = \det \left( \sum_{\alpha_i = 0}^{1/\Delta \alpha} \sum_{\alpha_j = 0}^{1/\Delta \alpha} \mathcal{M}(\Delta \alpha \cdot \alpha, \beta) \right),$$

(16)

is evaluated for $\Delta \alpha = 0.1$. Further, each $\beta \in \text{SL}(n)$ will be associated with an index $l$, such that $\mathcal{M}(\beta_1) < \cdots < \mathcal{M}(\beta_4) < \cdots < \mathcal{M}(\beta_{\text{max}})$. Based on Eq. 6, the number of feasible sensor arrangements (largest index $l$) for $n=2$, 4, and 6 sensors per arrangement totals to 1128, 194580 and 12271512 feasible sensor arrangements, respectively.

The results for $\mathcal{M}(\beta)$ with $\beta \in \text{SL}(4)$ are depicted in Fig. 3, which plots index $l$ versus $\mathcal{M}(\beta_i)$. Within Fig. 3, three different choices of sensor (green squares) arrangements are plotted. Arrangement A (Case A), depicts the optimal set of sensors if Optimality Criterion
Fig. 2 Modes of deformation for retraction with $\alpha_t = 1.0$ and $\alpha_s = 1.0$ (a) and elongation with $\alpha_t = 1.0$ and $\alpha_s = 1.0$ (b). Figures 2(a) and 2(b) additionally depict the principal strains as mirror cones at selected spatial points. The colour spectrum for the strains is depicted in Fig. 2(c). A qualitative comparison of the regional deformation of the tongue is broadly consistent with previously published observations (Napadow et al, 1999a,b).

Fig. 3 Ranking of feasible sensor arrangements, $\beta \in \text{SL}(4)$, for retraction based on the respective function values computed by Eq. 16. The locations for the best (A) and worst (C) choices based on Optimality Criterion I, as well as the index for the best choice based on Optimality Criterion II (B) are indicated. For A, B, and C, the respective sensor locations on the tongue surface are depicted.

$I$ is employed. The optimal set of sensors is denoted by $\beta^\text{opt}_I$. Sensor arrangement B (Case B) represents the optimal choice, $\beta^\text{opt}_I$, for Optimality Criterion II. Based on Optimality Criterion I, marker arrangement B relates to the 4555-th rank (index 190025 out of 194580). In addition to the best sensor locations for Optimality Criterion I and II, the worst choice for a set of sensor locations for capturing retraction, which is denoted by $\beta^\text{worst}_I$, is defined by sensor arrangement C (Case C). To substantiate the claim that the proposed sensor arrangement in Case A is indeed a better choice than the choices in Case B or C, a sensitivity analysis, as described in Section 2.3, was carried out. Within this subsection, one particular state during tongue movement of the tongue was synthetically chosen by computing the resulting deformations of the tongue using $\alpha_t = 0.75$ and $\alpha_s = 0.25$. This particular choice of activation levels is associated with a (generic) mode of deformation, which is likely to occur during retraction of
the tongue (this assumption is based on the fact that the only active muscle fibre groups during retraction are the transversus and the styloglossus muscle fibre group). This simulation (the forward problem) was used to determine the exact sensor locations for the different sensor arrangements, i.e., Case A, B, and C. Next, to investigate the practicability of the proposed sensor optimisation criteria for solving the inverse problem, the sensor locations were subjected to machine-inherent noise, i.e., $e_i^k = [e_i^1, e_i^2, e_i^3]^T$ with $e_i^k \sim N(\mu, \sigma^2)$, $\mu = 0$, and $\sigma^2 = 0.49$ (see also Section 2.3). For the respective sensor arrangements of Case A and B respectively containing 2 and 6 sensors, a total of 41 different perturbations $e_i^k = [e_i^1, e_i^2, e_i^3]^T$, $k = 1, \ldots, 41$, were computed. Similarly, 41 different perturbations were applied to the 4 sensor locations of Case A, B, and C. The aim of perturbing the sensor locations was to test the ability of estimating the actual level of activity for $\alpha_t$ and $\alpha_s$ based on sensor locations that are prone to errors. The results of the estimated activation levels $\alpha_t$ and $\alpha_s$ are presented as scatter plots for 2 sensors in Figure 4(a), for 4 sensors in Figure 4(b), and for 6 sensors in Figure 4(c).

In the following, a statistical analysis of the scatter plots in Figure 4 is presented. Table 1 lists the expectation values and the respective standard deviations of the above described sensitivity analysis.

From Table 1, one observes that the standard deviation of the estimated activation levels decreases for each Case A, B, and C, when the number of deployable sensors is increased from 2 sensors to 4 sensors and ultimately to 6 sensors, except for Case C, when the number of deployable sensors is increased from 4 to 6. This increase is, however, marginal (from 0.2991 in case of 4 sensors to 0.3051 in case of 6 sensors). Further, the difference between the expectation values of inverse problem, $E(\hat{\alpha}_t)$ and $E(\hat{\alpha}_s)$ respectively, and the associated levels of activation of the forward problem ($\alpha_t$ and $\alpha_s$) decreases with an increase in the number of sensors. This is in general true for both variables representing the level of activation, $\alpha_t$ and $\alpha_s$, except for the expectation value of the estimator $\hat{\alpha}_t$ in Case A, if the number of sensors is increased from 4 to 6 sensors. For this case, the difference between the expectation value of the estimator, $\alpha_t$, and the level of activation of the forward problem, $\alpha_t$, chosen for the forward problem, $\alpha_t = 0.75$, increases from 0.0032 for 4 sensors to 0.0088 with 6 sensors. However, the standard deviation of the estimated activation levels significantly decreases from 0.1244 to 0.0761.

Since the estimations for the activation levels $\alpha_t$ and $\alpha_s$ are obtained within one inverse problem, the expectation values of the estimators and the respective standard deviations should not be considered independently. Therefore, an additional overall measure that combines the estimations for $\alpha_t$ and $\alpha_s$ is introduced. This is achieved by first computing the scaled bias of an estimator $\hat{\theta}_i$:

$$B_i(\hat{\theta}_i) = s_i \frac{E(\hat{\theta}_i) - \theta_i}{\hat{\theta}_i}, \quad (17)$$

where $\theta_i$ is either $\alpha_t$ or $\alpha_s$ of the forward problem, i.e., $\alpha_t = 0.75$ and $\alpha_s = 0.25$ respectively, $E(\hat{\theta}_i)$ the expectation values reported in Table 1, and $s_i$ a scaling that depends on the volume fraction of the respective muscle fibre group. Here, the scaling factor for the bias of the transversus muscle fibre group, $s_t$, and the styloglossus muscle fibre group, $s_s$, are defined by

$$s_t = \frac{\int_\Omega d_t(x) \, dx}{\int_\Omega \sum_j \in F d_j(x) \, dx}, \quad (18)$$

and

$$s_s = \frac{\int_\Omega d_s(x) \, dx}{\int_\Omega \sum_j \in F d_j(x) \, dx}, \quad (19)$$

respectively. In Eqs. 18 and 19, $\Omega$ represents the volume of the entire tongue and $d_j$ is the volume fraction of muscle fibre group $j$. The integrals in Eqs. 18 and 19 are approximated by summing up the respective volume fractions at all Gauss points. Note that for each $x \in \Omega$, the sum of all volume fractions is 1. Hence, the respective scaling parameters are computed by:

$$s_t \approx \frac{\sum \# \, of \, GP \, d_t(x)}{\sum \# \, of \, GP} = 0.1772 \quad (20)$$

and

$$s_s \approx \frac{\sum \# \, of \, GP \, d_s(x)}{\sum \# \, of \, GP} = 0.1063. \quad (21)$$

In a similar way as for the averaged bias, one can introduce an overall measure for the standard deviation. This will be achieved by an averaged coefficient of variation. The coefficient of variation, $c_v(\hat{\theta}_i)$ for an estimator $\hat{\theta}_i$ is a normalised measure of dispersion of a probability distribution, i.e.,

$$c_v(\hat{\theta}_i) = \frac{\sigma(\hat{\theta}_i)}{E(\hat{\theta}_i)}. \quad (22)$$

Then, similar to the averaged bias introduced above, the variation of coefficients for the two different estimators is combined to one overall measure by taking into account the scaling parameters of the respective muscle fibre groups.
Fig. 4 Scatter plots visualizing the estimated levels of activation of the transversus (α_t) and styloglossus (α_s) muscle fibre groups for one deformation mode during tongue retraction. The green triangles depict the (α_t, α_s)-tuple obtained by the solving the respective inverse problems under the assumption that the sensor locations are chosen like in Case A, the yellow diamonds represent the results for Case B, and the blue squares plot the obtained activations levels for Case C. The red circle indicates the chosen level of activation for the forward problem.

Hence, the average bias and the average coefficient of variation can be defined for the case of retraction in the following way:

\[ B^p = \frac{1}{s_t + s_s} \left[ \sum_{i \in \{t,s\}} B_i(\hat{\theta}_i) \right] \]  \hspace{1cm} (23)

and

\[ \epsilon^p_v = \frac{1}{s_t + s_s} \left[ \sum_{i \in \{t,s\}} s_i \sigma(\hat{\theta}_i) \right] \left/ \left( E(\hat{\theta}_i) \right) \right. \]  \hspace{1cm} (24)

Table 2 reports, based on the data reported in Table 1, the average bias, \( B^p \), and the average coefficient of variation, \( \epsilon^p_v \), as defined in Eqs. 23 and 24. The average bias and coefficient of variation constitutes one overall measure for the ability to estimate the material parameters, \( \alpha_t \) and \( \alpha_s \), if sensor locations are subject to machine-inherent measurement errors. The bias decreases for each Case if the number of deployable sensors is increased with the exception for Case C, in which the bias for 4 sensors (0.1416) is smaller than the bias for 6 sensors (0.3367). Further, if the number of sensors are kept fixed, then the bias and the coefficient of variation increases from Case A to Case C except for the coefficient of variation for 4 deployable sensors in Case A (0.2693) and Case B (0.2666). However, for this case, the bias of Case A (0.0259) is (much) smaller than the bias of Case B (0.0887). Overall, Table 2 supports the results of Figure 3 that Optimal Criterion I is, at least for the presented case, superior to Optimal Criterion II.
### Table 1: Sensitivity results for tongue retraction

Reported are the expectation value and the standard deviation estimating the activation levels $\alpha_t$ and $\alpha_v$ obtained through inverse estimation using perturbed sensor locations. The respective sensor locations were perturbed (see Section 2.3) based on the results of the forward problem with $\alpha_t = 0.75$ and $\alpha_v = 0.25$. The results for a sample size of 41 is depicted for Case A, B, and C.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha_t$</th>
<th></th>
<th>$\alpha_v$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case A</td>
<td>Case B</td>
<td>Case C</td>
<td></td>
</tr>
<tr>
<td>2 sensors</td>
<td>0.7702 ± 0.1542</td>
<td>0.7547 ± 0.1120</td>
<td>0.6992 ± 0.2889</td>
<td></td>
</tr>
<tr>
<td>4 sensors</td>
<td>0.7532 ± 0.1244</td>
<td>0.7547 ± 0.1120</td>
<td>0.8056 ± 0.1694</td>
<td></td>
</tr>
<tr>
<td>6 sensors</td>
<td>0.7412 ± 0.0761</td>
<td>0.7547 ± 0.1120</td>
<td>0.7928 ± 0.1349</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: Average bias and average coefficient of variation calculated based on the expectation values and respective standard deviations reported in Table 1.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B^p$</td>
</tr>
<tr>
<td>2 sensors</td>
<td>0.1142</td>
</tr>
<tr>
<td>4 sensors</td>
<td>0.0259</td>
</tr>
<tr>
<td>6 sensors</td>
<td>0.0190</td>
</tr>
</tbody>
</table>

### Fig. 5: Sensor placements for tracking tongue retraction using 2, 4, and 6 sensors based on Optimality Criterion I and Optimality Criterion II. The green cubes represent the locations for Case A (optimal), the golden cubes show the locations of Case B (heuristic), while the blue cubes depict the sensor locations of Case C (worst). For 4 sensors, one sensor location coincides in Case A and B. This sensor is highlighted in Fig. 5(b) in red.

For tongue retraction, the optimal and worst sensor locations for $n = 2$, 4 and 6 deployable sensors as well as the 4 optimal sensor locations based on Optimality Criterion II are depicted in Figure 5.

#### 3.2 Sensor optimisation for elongation (half tongue)

The applicability and practicability of the proposed optimality criteria is further investigated by performing a similar analysis as presented in Section 3.1 for a particular mode of deformation during tongue elongation. The primary muscle fibre groups responsible for the elongation of the tongue are the transversus and verticalis muscle fibre groups. In this subsection, the mode of deformation of interest has been computed by choosing activation levels of $\alpha_t = \alpha_v = 0.5$. Then, based on the deformation obtained by solving the forward problem, we compute for each case, i.e., Case A and C for 2 and 6 sensors and Case A, B, and C for the 4 sensors, and based on the machine-inherent measurement errors the perturbed sensor locations. Again, a total of 41 perturbed sensor locations for each case and for each sensor arrangement were considered in the following analysis. For each perturbed sensor location an inverse problem was solved to estimate the respective levels of activation. The search space of the inverse problem hereby was restricted to $0 \leq \alpha_t \leq 1$ and $0 \leq \alpha_v \leq 1$. All other activation level parameters were considered to be 0.

Like for tongue retraction, the best and worst sensor arrangements for each of Optimality Criterion I and Optimality Criterion II were determined. The best and worst sensor arrangements based on Optimality Criterion I are denoted for 2, 4, and 6 sensors by Case A and Case C respectively. The best sensor arrangement based on 4 deployable sensors and Optimality Criterion II is denoted by Case B. The sensor arrangements...
for all cases are obtained by evaluating a similar objective function as proposed for the case of retraction. The only difference between the current objective function and the objective function defined in Eq. 16 are the non-zero activation levels (here: \( \alpha_t \) and \( \alpha_v \) instead of \( \alpha_t \) and \( \alpha_s \)). To extract the desired sensor locations, the respective sensor arrangement indices are ordered by the magnitude of the evaluated objectives function values. Again, the larger the objective function value, the more suitable is the chosen sensor location arrangement to capture the mode of the tongue during elongation.

Figure 6 presents for \( n = 4 \) sensors the objective function values and the locations of the sensors for Case A, B, and C.

Figure 7 depicts the solutions of the inverse problem for all cases. The green triangles depict the \((\alpha_t, \alpha_v)\)-tuple obtained by solving the respective inverse problems for the best sensor placements based on Optimality Criterion I, the yellow diamonds are the estimates based on the best sensor placements obtained by Optimality Criterion II, and blue squares refer to the estimates based on the worst sensor placements obtained by Optimality Criterion I. The red circle indicates the chosen activation levels \((\alpha_t, \alpha_v) = (0.5, 0.5)\), to determine a mode of deformation during tongue elongation.

The closer the estimates of the inverse problem are clustered around \((\alpha_t, \alpha_v) = (0.5, 0.5)\), the less subjective are the estimated activation levels to measurement errors present due to machine-inherent recording errors. To compare between Case A, B and C, a statistical analysis was carried out. Table 3 presents the expectation values of the estimators, \( E(\hat{\alpha}_t) \) and \( E(\hat{\alpha}_v) \) and their standard deviations based on the results for the scatter plots in Figure 7.

From Table 3, one observes that the standard deviations are, for a fixed number of deployable sensors, smaller in Case A than in Case C. Only the standard deviation decreases in the case of \( \alpha_t \) and \( \alpha_v \) of 4 deployed sensors from Case B to Case C. However, for \( \alpha_v \) and 4 deployed sensors, the standard deviation increases substantially from Case B to Case C. By statistically analysing each level of activation independently, i.e., \( \alpha_t \) and \( \alpha_v \), no clear behaviour with respect to the expectation values can be deduced from Table 7. However, one can observe that the expectation values vary less for the estimator of \( \alpha_t \) than for \( \alpha_v \). The range for \( E(\hat{\alpha}_t) \) is between 0.4889 and 0.5403, while the range for \( E(\hat{\alpha}_v) \) spreads from 0.4792 to 0.6022.

Instead of analysing the expectation values and the respective standard deviations of the different activation levels independently, a joint statistical measure is defined in the following. Like in the case for tongue retraction, the average bias and the coefficient of variation is computed by

\[
\bar{B}_c = \frac{1}{s_t + s_v} \left[ \sum_{i = (t,v)} B_i(\hat{\theta}_i) \right]
\]

and

\[
\bar{c}_c = \frac{1}{s_t + s_v} \left[ \sum_{i = (t,v)} s_i \sigma(\hat{\theta}_i) \right]
\]

where \( \theta_i \) is either \( \alpha_t \) or \( \alpha_v \) and the scaling factor for the transversus muscle fibre group, \( s_t \), is 0.1772 and for the verticalis muscle fibre group, \( s_v \), is 0.2087. The results of the average bias and the coefficient of variation for tongue elongation are given in Table 4.

Based on Table 4, one can deduce that the average bias and coefficient of variation increases from Case A to Case C, if the number of deployable sensors remains constant. This underpins the hypothesis that Case A exhibits suitable optimality criteria for placing sensors, if the aim is to perform an inverse estimation to calculate the muscle fibre activation levels based on noise-afflicted sensor measurements. For Case C, one can also observe that the average bias and coefficient of variation is reduced if the number of deployable sensors is increased. The same is true for Case A, if one increases the number of deployable sensors from 2 to 4. Further increasing the number of sensors increases the average bias from 0.0137 for 4 sensors to 0.0423 for 6 sensors. However, the average coefficient of variation reduces from 0.2105 (4 sensors) to 0.1828 (6 sensors).

For tongue elongation, the optimal and worst sensor locations for \( n = 2, 4 \) and 6 deployable sensors as well as the 4 optimal sensor locations based on Optimality Criterion II are depicted in Figure 8.

### 3.3 Numerical behaviour of the objective function

Computing for a specific set of sensor locations the respective activation levels is based on solving the optimisation problem given in Eq. 14. For the Trust-Region-Reflective algorithm, like for most other optimisation algorithms, a convex objective function with a global minimum are desired characteristics. In the following, a visual characterisation of the objective function is presented. To do so, the objective function, which is minimised with respect to the level of activation (cf. Eq. 14), has been evaluated for the admissible parameter space subject to different scenarios, e.g. for retraction and elongation and different sensor arrangements. Figure 9 depicts the contour plots for retraction while choosing the above-computed sensor locations of CASE
Fig. 6 Ranking of feasible sensor arrangements, $\beta_i \in \text{SL}(4)$, for elongation based on the respective objective function values. Further, the locations for the best (A) and worst (C) choices based on Optimality Criterion I, as well as the index for the best choice based on Optimality Criterion II (B) is included. For A, B, and C, the respective sensor locations on the tongue surface are depicted.

Table 3 Sensitivity results for tongue elongation: Reported are the expectation value and the standard deviation estimating the activation levels $\alpha_t$ and $\alpha_v$ obtained through an inverse estimation using perturbed sensor locations. The respective sensor locations were perturbed (see Section 2.3) based on the results of the forward problem with $\alpha_t = \alpha_v = 0.5$. The results for a sample size of 41 is depicted for Case A, B, and C.

Table 4 Average bias and average coefficient of variation calculated based on the expectation values and standard deviations reported in Table 3.

A (Figure 9(a)) and CASE C (Figure 9(b)) using 6 sensors. Figure 10 presents the respective contour plots for tracking elongation with 4 sensors. CASE A is depicted in Figure 10(a) and CASE C in Figure 10(b). The sensors were not subjected to any noise for all cases.

In Figures 9(b) and 10(b), one observes that the smallest contour line encompasses a much larger area than the respective areas in Figures 9(a) and 10(a). Therefore, perturbing sensors for bad sensor placements (adding measurement noise) has a greater effect than for the best sensor placement as the objective function represents the sum of the $L_2$-norms of the differences between the measured sensor locations and the computed sensor locations. Hence, Figures 9 and 10 provide a visual insights on why non-optimal sensor placements.
Fig. 7 Scatter plots visualising the estimated levels of activation of the transversus ($\alpha_t$) and verticalis ($\alpha_v$) muscle fibre groups for one deformation mode during tongue elongation. The green triangles depict the ($\alpha_t, \alpha_v$)-tuple obtained by solving the respective inverse problems under the assumption that the sensor locations are chosen like in Case A, the yellow diamonds represent the results for Case B, and the blue squares plot the obtained activations levels for Case C. The red circle indicates the chosen level of activation for the forward problem.

Fig. 8 Sensor placements for tracking tongue elongation using 2, 4, and 6 sensors based on Optimality Criterion I and Optimality Criterion II. The green cubes represent the locations for Case A, the golden cubes show the locations of Case B, while the blue cubes depict the sensor locations of Case C. For 4 sensors, one sensor location coincides in Case A and B. This sensor is highlighted in Fig. 5(b) in red.
Fig. 9 Contour plot of the objective function given by Eq. 14. Figure 9(a) depicts the objective function in the case of tracking retraction with 6 sensor and the best sensor placement (CASE A) while Figure 9(b) depicts the objective function for worst sensor placement (CASE C).

Fig. 10 Contour plot of the objective function given by Eq. 14. Figure 10(a) depicts the objective function in the case of tracking elongation with 6 sensor and the best sensor placement (CASE A) while Figure 10(b) depicts the objective function for worst sensor placement (CASE C).
exhibit more variations in predicting the respective activation levels.

4 Discussion

Motion tracking systems are commonly used to study the dynamics (and indirectly the biomechanical behaviour) of an object. The challenge of all motion tracking systems is to find suitable marker placements that accurately capture the fundamentals of the tracked motion. In gait analysis, for example, the challenge is to deduce from a (small) set of markers the relative motion of the bones within the musculoskeletal systems. Skin artefacts (e.g., Karlsson and Tranberg, 1999) and wobbling masses (e.g., Schmitt and Günther, 2011) can play a crucial role in tracking gait and have the potential to introduce significant errors to the analysis. In the case of the tongue, the first challenge is to obtain reasonable marker movement as, for the most part, the tongue is hidden within the oral cavity. This can be achieved using electromagnetic articulographs. However, to obtain further insights into the biomechanical behaviour of the tongue, one has to study its mechanical behaviour in addition to the motion. In the case of the tongue, this is a challenging task as the tongue is a complex, extremely flexible and dexterous mechanical object consisting of many interlaced muscle fibre groups. This study provides, for the first time, a model-based identification method to sensibly place sensors on the accessible part of the tongue surface to track different movements, i.e., tongue retraction and elongation. This is achieved based on a pre-computed envelope of motion for different movements. Such a methodology may also provide a suitable replacement for experimental studies that aim to determine optimal sensor locations for gait analysis.

In addition, Magnetic Resonance Imaging (MRI) techniques could be employed to track the tongue Stone et al (2001); Hiiemae and Palmer (2003); Honda et al (2010). In contrast to motion tracking systems, which are limited to track accessible parts of the tongue, MRI has the advantage to obtain information on the tongue’s structure. For example, Diffusion Tensor MRI can provide interesting insights into the internal structures of the tongue, e.g., the tongue’s fibre distribution. Other MRI scanning sequences can provide ways of investigating strain (e.g., Zhong et al, 2008, for skeletal muscles) or tagged MRI allows one to analyse the muscle deformation during motion (Napadow et al, 1999a,b, 2002). However, all MRI scanning methodologies have the same drawback that the acquisition of the images is time-consuming and that tongue movement, e.g. swallowing, happens quickly. This poses a significant challenge to obtain reasonable good image quality as the subject would need to hold a desired pose for a long period of time or needs to be able to perform the process of swallowing extremely slow or in a very repeatable fashion.

One drawback of the presented approach is that the tongue has only been considered as an isolated structure/organ. The interaction of the tongue with the surrounding structures of the oral cavity have not been considered in computing the envelope of motion using the forward continuum-mechanical model. It is clear that interactions of the tongue with the palate and teeth have an influence on the tongue’s mode of deformation. However, since interactions with surrounding structures only occur in cases of higher activation levels, e.g., in the case that the tip of the tongue has been elongated far out of the mouth, the drawbacks of not considering the interactions are limited. Further, taking into account the contact between the tongue and other structures such as the palate and the teeth is not possible without investing additional significant amounts of computational cost. While the forward problem, i.e., computing the envelope of motion, is still feasible, it would have exceeded our computational resources for estimating the level of activation using the inverse problem. However, if contact with surrounding structures does only restrict displacements at surface areas other than the proposed optimal sensor locations, the results and its conclusions, in particular the optimal choice of sensor locations, still hold as the optimality criteria is displacement based. However, should contact have a significant influence on the the overall modes of deformation, then one would have to re-compute the respective envelope of motion and re-apply the proposed methodology to determine the optimal sensor locations for the respective movements. In summary, while the determined sensor locations herein might be subjective to contact, the methodology is not.

The resulting modes of deformations are also directly linked to the chosen material parameters within the continuum-mechanical model. As with all biological tissue, finding appropriate constitutive relations and material parameters is an extremely challenging task. The anatomical structure and the complex muscle fibre distribution within the tongue add further challenges. The material parameters chosen within this study are based on literature data. Cheng et al (2011) report for the tongue an overall shear modulus of 2.67 ± 0.29 kPa (mean ± standard deviation). As the shear modulus for a Mooney-Rivlin material is twice the sum of the two material parameters, i.e., $2(c_1+c_2)$, choosing $c_1 = 0.375$ kPa and $c_2 = 0.175$ kPa for the ground matrix assumes

\begin{align*}
\alpha &= 0.175, \\
\beta &= 0.375, \\
\kappa &= 2.67, \\
\lambda &= 0.29.
\end{align*}
that the ground matrix can contribute up to about 41% to the overall shear. As far as the maximum active fibre stress is considered, Titze (1994) report for the laryngeal muscles an maximum active fibre stress of 100 kPa. De Groot and Van Leeuwen (2004) reports for the maximum active fibre stress in a chameleon tongue of 200 kPa. Based on the assumption that the maximum active stress of the human tongue is somewhere between the laryngeal muscle and the chameleon tongue muscle, we assume a maximum active fibre stress of 150 kPa.

The mechanical model is not only the basis for determining the set of sensor locations based on the envelope of movement. The mechanical model is also used to compute, based on a particular set of activation levels the sensor locations within the actual configuration. Further, the ability to perform the inverse estimation depends on the robustness of the mechanical model. The robustness of the inverse estimation procedure is further linked to the choice of the objective function. Within this work, the objective function of the inverse estimation is solely based on minimising the Euclidean distance between error-prone sensor locations and the "exact" sensor locations obtained from the forward dynamics simulations (using specific activation levels). Given the setting of large deformations, an inverse simulation that is based on minimising only the Euclidean distances between the sensors as objective function might result in non-unique activation levels. Enhancing the objective function with other measures, e.g., the strain, would greatly increase the robustness of the inverse procedure. However, the motion capture setup only provides the actual positions of the markers with respect to a global coordinate system and therefore, limits for now, the objective function to the Euclidean distance. An additional objective function that needs to be investigated within future research could take into account inter-marker distances. However, since we restricted our inverse estimation to the case that only two muscle fibre activation levels are estimated, the proposed objective functions seems to be adequate. If it becomes necessary for the cases of more complex tongue movements to estimate more than two muscle activation parameters, other objective functions might need to be considered.

The self-imposed restriction to two muscle activation parameters while subjecting the sensor locations to a Gaussian-distributed measurement error might result in not being able to find the optimal activation parameters within the search space. For example, if the error-prone sensor locations are not within the range of the envelope of motion obtained through the parameter space (i.e., allowing to vary the respective activation parameters between 0 and 1), then no optimal solution can be found. In such cases an optimal solution might be attainable, if the parameter search space is extended to include more muscle fibre groups or if the range of activation parameters is increased beyond the interval [0,1]. The latter case is, from a physiological modelling point of view, unreasonable. Increasing the search space to other muscle fibre groups might be feasible, however, not desirable as well, as retraction or elongation are assumed to be driven by only two muscle fibre groups. As a result of not finding an optimal solution, the inverse estimation procedure obtains activation levels at the fringe of the parameter space. This, however, is not a drawback of the methodology presented within this work. The overall aim has to be to find sensor locations that are capable of obtaining the actual activation levels even if the sensor recordings are subjected to measurement errors. Needless to say that good estimations are only obtained, if a sequence of estimations based on different measurement errors is computed.

Further, the inverse estimation procedure can be sensitive to the initial starting values. There exists the possibility that the objective function exhibits a plateau or areas with very flat gradients. For such a case, the inverse problem introduces an inherent variability. However, the presented results for the expectation values and the standard deviations show that this variability can be reduced, if suitable sensor locations are chosen. For example, the estimated activation levels of Case C are spread almost over the entire admissible parameter space, while the activation levels for Case A are contained within a much smaller region (see Figures 4 and 7). Further, one can deduce from Section 3 that the optimal sensor locations chosen by means of Optimality Criterion I has a better expectation value/average bias than those based on Optimality Criterion II or the worst sensors locations. Further, this study demonstrates that increasing the number of deployed sensors improves the estimated levels of activation. Deviations from this claim, e.g., the increased average bias in Case A, if 6 sensors are deployed instead of 4, is due to the added noise. An additional test by choosing 12 additional test samples (hence, a total of 53 test samples) reduced the expectation value for 6 sensors and $\alpha_t$ already significantly. The mean and the standard deviation reduced for the 6 sensor case for $\alpha_t$ and $\alpha_e$ to $0.5227 \pm 0.1300$ and $0.5047 \pm 0.0568$ respectively. Moreover, it is necessary to introduce the average bias and coefficient of variation for statistically analysing the obtained activation levels.

In multi-parameter optimisation, one has often the ability to balance the accuracy of a specific optimisation parameter through appropriate scaling. This, however, is not possible within the context of the proposed
optimisation function as both parameters (level of activation) are inherently coupled through the mechanical problem. However, since the mechanical model contains an intrinsic scaling through the volume fraction and the level of activation, additional care needs to be taken when interpreting the results of the inverse problem. Muscle fibre groups with small volume fractions might not significantly contribute to the overall mode of deformation and hence, the expectation value and the standard deviation of an estimator might be more error-afflicted than the estimator for the other muscle fibre group. However, if one assumes that a small volume fraction of a particular muscle fibre group has only a small effect on the overall mode of deformation, one does not worry as long as one is only interested in the overall mode of deformation. This, however, can become a more delicate issue if one is interested in the absolute value of the activation level. In this case, one should be aware that the optimisation procedure might favour a larger muscle fibre group over a smaller one.

Due to the complex muscle fibre architecture and its associated mechanical behaviour, only two specific motions (tongue retraction and elongation) have been considered. Further, to reduce the computational time, symmetry assumptions have been exploited, i.e., the symmetric arrangement of the muscle fibre groups within the tongue leads to symmetric contractions and to symmetric deformations. Further, the presented simulations are based on the assumption that the level of activation is constant throughout the entire muscle fibre group. If one only considers symmetric movements such as tongue elongation or retraction, this is a realistic assumption and therefore also justifies the restriction of the analysis to one half of the tongue.

The number of sensors for each half of the tongue has been limited to a maximum of 6. This due to the physical limitations in recording channels of the articlogroup. A typical setup of an AG500 consists of 12 recording channels. However, placing all 12 sensors on the tongue is not advisable. To obtain the "pure" tongue movement, one also has to consider the movement of the mandible. To obtain the mandible movement with six degrees of freedom, one has to use at least three markers to track the movement of the head, e.g., fix them at the maxillary teeth, and three markers on the mandible, e.g., on the mandibular teeth (compare, for example, Röhrle et al., 2009; Saini et al., 2009). The initial distribution of potential sensor locations (48 per half tongue) have been evenly distributed over the accessible tongue surface. The total number of 48 potential sensor locations also decreases the probability that two neighbouring sensors do overlap after additional noise, which resembles machine-inherent error, has been applied within this study. Such an overlap would provide additionally difficulties in obtaining the levels of activation by means of the inverse problem.

Determining sensor locations for the full tongue would likely result in a different optimal sensor arrangement than just mirroring the optimal sensor locations found for the half tongue. For example, in case that one sensor is on the line of symmetry, it would make in the case of the full tongue no sense to add another sensor at the same location as previously determined. Nevertheless, the same approach as presented herein can be utilised to determine sensor locations on the full tongue. From a computational point of view, the number of deployable sensor locations needs to be reduced. Choosing a total of 96 deployable (twice the deployable sensors on half of the tongue) would result in too many combinations of sensor arrangements. One possibility to circumvent this would be to rule out particular sensor arrangements a priori.

From a practical point of view, it is clear that sensors cannot be placed exactly as computed herein. Recommendations from this work are of qualitative nature. The authors are well aware that there exists a great deal of variability between different subjects in the tongue’s shape and muscle fibre distribution. Nevertheless, this work provides clear evidence for a recommendation of sensor placements if one aims to track tongue retraction or elongation. This claim is substantiated by Figures 3 and 6. There exists several sensible choices of sensor locations, but also many mediocre and bad choices. This also leads to the claim that one should not choose the sensor locations arbitrarily or heuristically, e.g., like for Case B. An heuristic criterion like Optimality Criterion II, which aims to take into account the maximal variation in displacement for each coordinate displacement and the magnitude of the displacement, exhibits a worse average bias and coefficient of variation, if compared to Case A. The overall method is well suited to be also adopted to other cases, in which one would like to determine an optimal set of sensor locations to capture the key characteristics of mechanical deformation, e.g., in gait analysis. Finally, this study has implications for using intra-oral pressure readings from the dorsum of the tongue to reconstruct intra-individual tongue movement patterns during normal and dysphagic swallowing (e.g., Kieser et al., 2011).

Acknowledgements

The author O.R. would like to thank the German Research Foundation (DFG) for financial support of the
project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart and the FP7-PEOPLE-2009-IRSES ”MuscleUp” (#246994). Y.W. would like to acknowledge support from the Foundation for Research in Science and Technology under contract NERF UOAX0406.

References


