

ARTICULATORY SPACE CALIBRATION IN 3D ELECTRO-MAGNETIC ARTICULOGRAPHY

An Ji¹, Michael T. Johnson¹, Jeffrey Berry²

¹Electrical and Computer Engineering, Marquette University, Milwaukee, WI, USA

²Speech Pathology and Audiology, Marquette University, Milwaukee, WI, USA

ABSTRACT

This paper introduces a new method to calibrate data collected using Electro-Magnetic Articulography (EMA) into an appropriate articulatory space. A bite plate record for a target subject is used to define the maxillary occlusal and midsagittal planes, and then a single quaternion rotation is derived to transform the dataset into the new anatomically referenced space. The choice of specific rotation solution is discussed relative to the corresponding anatomical assumptions regarding the original sensor placement and coordinate system. Data were collected using NDI Wave Speech Research System for one pilot subject, and calibration results and consistency throughout the calibration record reviewed. The results show that the rotation solution can accurately and consistently transform all sensors' positions into an articulatory space in which sensor movements and orientations can be easily analyzed. This preliminary study enables the investigation of articulatory kinematics and relationship to acoustics.

Index Terms— Electro-Magnetic Articulography, articulatory space, quaternion representation

1. INTRODUCTION

A number of technologies have been used historically for recording articulator movements. X-ray cinematography [1-2] is effective, but the radiation to the subject's head is a concern. Cine MRI can provide dynamical 3D measurement of the vocal tract but it is somewhat cumbersome and expensive [3-4]. The ultrasound technique is able to capture the surface of the tongue [5-6], but noise, echo artifacts and refractions may affect the results. Electro-Magnetic articulography (EMA) sensing systems has been growing in use in recent years for investigation of articulatory kinematics [7-9]. This system is able to track both position and orientation of the sensors, which can include both 5 DOF (degree of freedom), corresponding to a planar orientation, and 6 DOF options. This enables detailed study of the relationship between acoustic data and articulator movements. It is necessary to calibrate all the raw data from the system's coordinate space into a meaningful anatomical articulatory space [10]. Most datasets available now include the bite plate calibration in their preprocessing stage [11-14]. However, none of them provide a detailed description of this processing sufficient for 3D EMA, or the underlying assumptions on which the calibration is based. In this

paper, a bite plate calibration algorithm based on quaternion rotation method is developed. This approach allows for a single transformation of both the position and orientation information into an articulatory reference space.

The remainder of this paper is organized as follows: Section 2 discusses the desired articulatory space. Section 3 and 4 describe the method in detail, and Section 5 gives results and evaluates the proposed algorithm, with conclusions in Section 6.

2. ARTICULATORY SPACE

In the data set, a 6DOF sensor attached to the center of the forehead defines the local coordinate system. All other sensors are 5DOF, which are less bulky and allow better freedom of movement to articulators. The 6DOF sensor is used as a reference sensor to track the head movement, and the system implements head movement correction internally so that data for all 5DOF sensors is referenced to the 6DOF origin and orientation. The goal of calibration is to define a robust articulatory space and find the optimal translation and rotation to accomplish local to articulatory space coordinate transformation. This target articulatory space is based on each subject's anatomy, as shown in Figure 1. The origin of the new coordinate system is located at the central maxillary incisors [10]. When maxillary and mandibular natural teeth are brought together, a plane of contact can be defined between the occlusal surfaces of the upper and lower teeth. This plane is called the occlusal plane and it represents the x - z plane in our coordinate system.

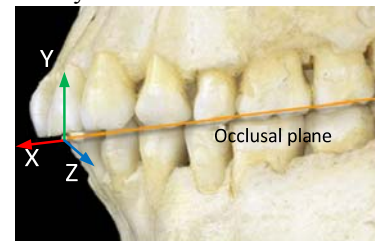


Figure 1: Target articulatory referenced coordinate system. The origin is located at the central maxillary incisors with the x and z axis lying in the occlusal plane (highlighted orange). The y axis is perpendicular to that plane along the midsagittal axis.

The negative x axis follows the midsagittal line of the occlusal plane toward the back of throat. The positive z

axis runs perpendicularly to the x axis on the occlusal plane in the direction of the subject's left. The positive y axis is perpendicular to the occlusal plane in the upward direction. Thus the x - y plane defines the midsagittal plane and the x - z plane defines the maxillary occlusal plane in the new articulatory coordinate system.

A bite-plate calibration record is used for identifying the new coordinate system. One 6DOF sensor is attached at the center of forehead (denoted REF) and two 5DOF sensors are placed on the bite-plate, one at the maxillary central incisors (OS) and one along the midsagittal plane at the bisection between the back molars (MS), as shown in Figure 2 below. To create the bite-plate, two pieces of dental impression wax are softened in warm water and molded onto a tongue depressor. This softened wax is then placed into the subject's mouth and an impression of the bite is taken. Immediately afterward (while the wax is still soft), the experimenters measure the midpoint between the back molars and create an indentation for the placement of the MS sensor. The OS sensor is placed directly in front of the central maxillary incisors. These sensors are pressed into the wax until they are secured and the bite-plate is replaced in the subject's mouth for the bite-plate recording.



Figure 2: Bite-plate with sensor positions

In normal recording of articulatory movements, the bite-plate wax is taken out and 5DOF sensors were placed at the desired recording points. For this experiment a simple configuration was used with two 5DOF sensors placed on the mandibular incisor and second molar. The 6DOF sensor remained attached in the same forehead position throughout. The translation and rotation determined from the bite plate were applied to each data record to map the original coordinate system to the articulatory space system, creating a virtual origin point at the central maxillary incisors and a correctly referenced occlusal and midsagittal plane.

The determination of the translation needed to move the origin of the system is straightforward, since both the REF sensor and the OS sensor are simultaneously present in the bite plate record. However, the rotation needed to create the desired articulatory space is not as straightforward. In addition to the OS and MS sensors, one additional physical assumption is needed to define the midsagittal plane. There are a number of possibilities for this, including for example: 1) assuming that the REF sensor (along with the OS and MS sensors) is placed

precisely in the midsagittal plane, 2) assuming that the REF sensor vertical orientation (y axis) is precisely along the midsagittal plane, or 3) adding an additional sensor to the bite-plate to define the occlusal plane, with the midsagittal plane perpendicular to this.

Each of these assumptions could be a reasonable way to define the midsagittal plane, if care is taken to minimize physical error relative to the necessary sensor placement and orientation. Each assumption also represents a different potential source for calibration error. In this experiment we assumed the first of these configurations, and place the REF sensor physically in the center of the forehead to define the midsagittal plane.

3. PROPOSED METHOD

3.1. Quaternion rotation representation

Since the orientation data for the system are given in quaternion format, the desired rotation will be implemented via a quaternion rotation approach. In computer visualization and animation, the quaternion format is a commonly used method to represent rotation and orientation [15]. A quaternion is a 4-D unit vector $q = [q_0, q_x, q_y, q_z]$ satisfying the following equation:

$$q_0^2 + q_x^2 + q_y^2 + q_z^2 = 1 \quad (1)$$

A quaternion rotation thus lies on the 4-D unit hypersphere. The key application of quaternions to tracking jaw movement in this paper lies in their use to represent rotations. A unit-normalized quaternion can be used to represent a rotation by an angle θ around a unit axis v .

$$q = \left(\cos\left(\frac{\theta}{2}\right), \sin\left(\frac{\theta}{2}\right)v_x, \sin\left(\frac{\theta}{2}\right)v_y, \sin\left(\frac{\theta}{2}\right)v_z \right) \quad (2)$$

where the vector part $\sin(\theta/2)v = [q_x, q_y, q_z]$ defines the rotate axis direction, and the scalar part $\cos(\theta/2) = q_0$ defines the degree of rotation. To rotate a point whose position is represented by the vector \vec{p} by an angle θ along the axis v to a new position p_{final} we apply the quaternion multiply operation

$$p_{final} = QPQ^* \quad (3)$$

where $Q = [\cos(\theta/2), \sin(\theta/2)v]$, $P = [0, \vec{p}]$.

In the NDI EMA system used here, sensor orientations are represented by a quaternion vector, which indicates the amount of rotation a sensor has undergone relative to its established base orientation in the coordinate space.

In the following section, we will derive the unique quaternion rotation needed to transform all position and orientation rotation in local coordinates into the articulatory space coordinates.

3.2. Rotation axis and angle

In the target articulatory space, following translation, the OS represents the origin, with the MS directly on the x axis and the REF sensor at a point in the x - y plane (not

necessarily on the y axis, because the MS-OS-REF triangle may not form a right angle. Since the distance from OS to MS, the distance from OS to REF, and the MS-OS-REF angle can all be directly computed, the exact new coordinate locations for the MS and REF sensors can be easily determined. The resulting necessary calibration rotation is thus the rotation which will rotate the MS-OS-REF triangle onto the new target coordinates. There is a single unique rotation to accomplish this.

The notation is defined as follows: \overline{MS} and \overline{MS}' represent the MS location in local and articulatory coordinates respectively, while \overline{REF} and \overline{REF}' are the corresponding head reference sensor coordinates.

One approach to determining the necessary rotation is to consider the rotation of the MS and REF points separately. There are an infinite number of rotations that will rotate the original \overline{MS} onto \overline{MS}' , and also an infinite number of rotations that will rotate the original \overline{REF} onto \overline{REF}' , and the final solution is the unique intersection of these two sets. Figure 3 illustrates how to describe the set of rotation axes for the MS point. The rotation can be visualized as rotation along the surface of a cone, with the center of the cone as the rotation axis. All possible rotation axes lie in the plane formed by the bisecting vector \overline{BS}_{MS} and the cross-product vector \overline{V}_{MS} .

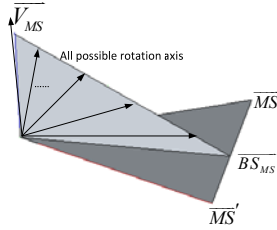


Figure 3: 3D visualization for the rotation axis

By following the same steps we can find another plane that includes all axes which rotate \overline{REF} to \overline{REF}' in the new coordinates. The intersection of these two planes is a unique line which is the single rotation axis we are looking for.

Equations (4), (5), (6), (7), (8) give the calculation method:

$$\overline{V}_{MS} = \langle \overline{MS}, \overline{MS}' \rangle \quad (4)$$

$$\overline{V}_{REF} = \langle \overline{REF}, \overline{REF}' \rangle \quad (5)$$

$$\overline{BS}_{MS} = \frac{\overline{MS} + \overline{MS}'}{2} \quad (6)$$

$$\overline{BS}_{REF} = \frac{\overline{REF} + \overline{REF}'}{2} \quad (7)$$

$$\overline{Axis} = \langle \langle \overline{V}_{MS}, \overline{BS}_{MS} \rangle, \langle \overline{V}_{REF}, \overline{BS}_{REF} \rangle \rangle \quad (8)$$

Where $\langle \rangle$ denotes the cross product operator.

After finding the rotation axis, it is necessary to determine the correct rotation angle. Figure 4 illustrates this computation, based on the visualization of the rotation cone. The following equations show how to calculate the rotation angle θ , and give the resulting quaternion form rotation.

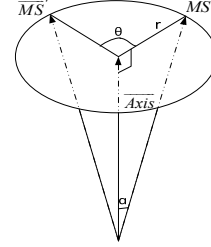


Figure 4: 3D visualization for the rotation angle

Equations (9), (10), (11) calculate the rotation angle and its quaternion form

$$r = |\overline{MS}| \sin \alpha \quad (9)$$

$$\theta = 2 \sin^{-1} \left(\frac{1/2 |\overline{MS} - \overline{MS}'|}{r} \right) \quad (10)$$

$$q = [\cos(\frac{\theta}{2}), \sin(\frac{\theta}{2}) \overline{Axis}_x, \sin(\frac{\theta}{2}) \overline{Axis}_y, \sin(\frac{\theta}{2}) \overline{Axis}_z] \quad (11)$$

This q can applied using equation (3) to any data record to transform into the appropriate articulatory space coordinate system.

4. EXPERIMENTAL METHOD

The NDI Wave Speech Research system [16] was used with a pilot subject to create a bite-plate record, implement the calibration method described above, and then track the positions and orientation of jaw sensors. The EMA system provided head-movement corrected data in the local coordinate space relative to the reference sensor on the forehead. The system samples kinematic data at 400Hz and acoustic data at 44.1kHz.

The rotation calibration was determined from the mean positions of the OS, MS, and REF sensors in the bite plate record, and the equations above were used to perform the resulting rotation on the bite plate record data, as well as on data from a simple jaw movement experiment.

5. RESULTS AND DISCUSSION

Figure 5 and Figure 6 show a trace of the bite plate record data in both the original and transformed coordinate spaces. All units are in millimeters. It can be seen from these figures that the original MS-OS-REF plane is correctly rotated into the x-y plane in the new articulatory space. Table 1 shows the mean values and standard deviations of MS and REF for this data set after calibration, compared to the value determined computationally. The mean value after calibration is the same as the computed value, indicating that the derived

solution correctly transforms the space into the desired coordinate system. The standard deviation gives an indication of physical sensor movement during the calibration process itself. Sources of this error include changes in REF sensor position and/or orientation, changes in OS or MS sensor position, tracking and computation error in the internal head-correction algorithm, and error margins of the EMA data collection system itself. The standard deviation across the bite plate record is in the range of 0.05-0.11 millimeters, suggesting relatively good stability of the transformation to the articulatory-referenced coordinate system. (Note, however, that there was no attempt in this experiment to push the boundaries of these errors, and that in fact the subject was relatively still during the biteplate record.)

Table 1. MS and REF sensor positions in articulatory space

	X	Y	Z
Derived MS value	-34.7538	0	0
MS mean	-34.7538	0	0
MS standard deviation	0.0556	0.0491	0.0657
Derived REF value	-8.6217	77.9419	0
REF mean	-8.6217	77.9419	0
REF standard deviation	0.1140	0.0656	0.0608

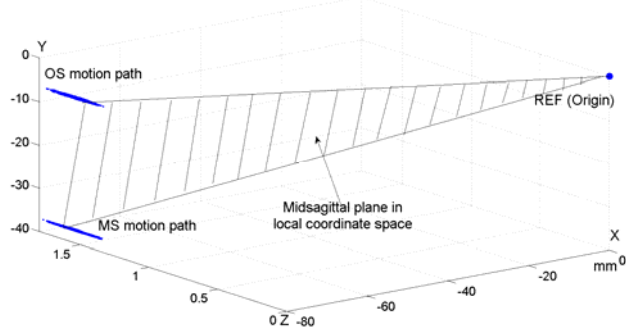


Figure 5: MS, OS and REF position in local coordinates

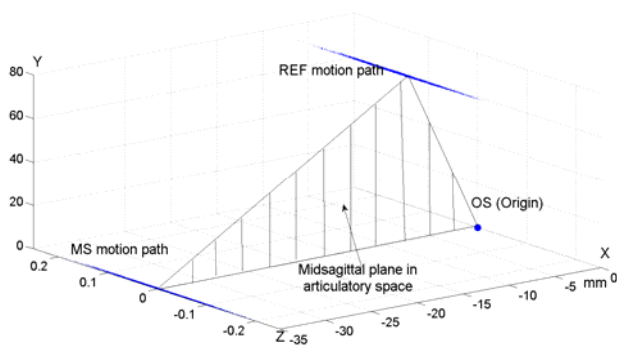


Figure 6: MS, OS and REF position after bite plate calibration

To illustrate the implementation of articulatory kinematics using the new physiologically referenced coordinate space with both position and orientation data, an illustrative example of jaw movement and angle computation from a single sensor. The two-dimensional rigid-body model [15-16] decomposes speech-related jaw

movements into three components: rotation degree around the transverse axis located approximately through the condyles, and horizontal and vertical translation of this axis in the midsagittal plane. These components are easy to calculate when in the articulatory referenced coordinate space. Figure 7 shows the motion path of mandibular sensors and corresponding jaw rotation angle in the midsagittal plane after calibration. One sensor is attached on the mandibular incisor and the other is on the buccal surface of the second molar, with the subject performing a simple repeating range of motion movement on the jaw. Jaw angle can be determined directly from the incisor sensor orientation, as shown in Figure 7.

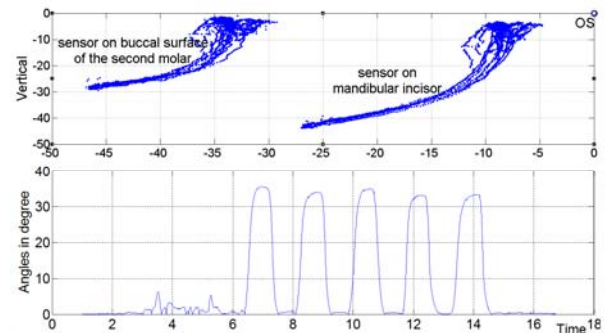


Figure 7: jaw movements in midsagittal plane after calibration

6. CONCLUSION

The work presented here has introduced a calibration approach for transformation of kinematic data into an appropriate and stable articulatory coordinate space, and demonstrated implementation of this approach using a quaternion rotation method. The resulting algorithm is able to transform both position and orientation data into the desired coordinate system. This is a first step enabling investigation of the relationship between articulator kinematics and acoustics in the EMA framework.

7. ACKNOWLEDGEMENT

This research is supported by grant IIS-1141826 from the National Science Foundation.

8. REFERENCES

- [1] Munhall, K. G., Vatikiotis-Bateson, E., And Tohkura, Y., "X-ray film database for speech research," J. Acoust. Soc. Am., pp. 1222-1224, 1998
- [2] Perkell, J. S., "Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study," Research Monograph No. 53, The M.I.T. Press, Cambridge, MA, 1969.
- [3] Masaki, S., Tiede, M. K., Honda, K., Shimada, Y., Fujimoto, I., Nakamura, Y., and Ninomiya, N., "MRI-based speech production study using a synchronized sampling method,"

- Journal of the Acoustical Society of America, col. 20, no. 5, pp. 375-397, 1999.
- [4] Narayanan, S., Nayak, K., Lee, S., Sethy, A., AND Byrd, D., "An approach to real-time magnetic resonance imaging for speech production," *J. Acoust. Soc. Am.*, vol. 115, no.4, pp. 1771-1776, 2004.
- [5] Stone, M.L., Sonies, B.C., Shawker, T. H., Weiss, G., and Nadel, L., "Analysis of real-time ultrasound images of tongue configuration using a grid-digitizing system," *Journal of Phonetics*, vol. 11, no. 3, pp. 207-218, 1983.
- [6] Kaburagi, T. and Honda, M., "An ultrasonic method for monitoring tongue shape and the position of a fixed-point on the tongue surface," *J. Acoust. Soc. Am.*, vol. 95, no. 4, pp. 2268-2270, 1994.
- [7] Schwestka-Polly R, Engelke W, Engelke D., "The importance of electromagnetic articulography in studying tongue motor function in the framework of an orthodontic diagnosis", *Fortschritte der Kieferorthopädie* 53: 3–10, 1992.
- [8] Horn H, Göz G, Bacher M, Müllauer M, Kretschmer I, Axmann-Krcmar D., "Reliability of electromagnetic articulography recording during speaking sequences," *European Journal of Orthodontics* 19: 647–655, 1997.
- [9] Fuchs, S., Perrier, P., and Pompino-Marschall, B., "Speech production and perception: Experimental analyses and models", *ZAS papers in Linguistics*, 40, 2005.
- [10] Westbury, J. R., "The significance and measurement of head position during speech production experiments using the x-ray microbeam system", *J. Acoust. Soc. Am.* Volume 89, Issue 4, pp. 1782-1791, 1991.
- [11] Byrd, D., Browman, C. P., Goldstein, L., and Honorof, D., "Magnetometer and x-ray microbeam comparison," in *Proceedings of the 14th International Congress of Phonetic Sciences*, edited by J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey (American Institute of Physics, New York), pp. 627–630. 1999.
- [12] Krista, R., "The Effect of Palate Morphology on Consonant Articulation in Healthy Speakers", Master thesis, Department of Speech-Language Pathology, University of Toronto, 2011.
- [13] Westbury, J. R., "X-Ray Microbeam Speech Production Database User's Handbook Version 1.0", Jun. 1994.
- [14] Gracco, V. L., Nye, P. W., "Magnetometry in speech articulation research: some misadventures on the road to enlightenment", *Forschungber. Institute Phonet. Sprachl. Kommun. University München* 31: 91-104, 1993.
- [15] John C. Hart, George K. Francis, Louis H. Kauffman, "Visualizing quaternion rotation," *ACM Transactions on Graphics*, v.13 n.3, pp. 256-276, 1994.
- [16] Berry, J., "Accuracy of the NDI Wave Speech Research System," *Journal of Speech, Language, and Hearing Research*, Vol.54 1295-1301, 2011.
- [17] Westbury, J. R. "Mandible and hyoid bone movements during speech," *Journal of Speech and Hearing Research*, 31, pp. 405-416. 1988.
- [18] Edwards J. "Mandibular rotation and translation during speech," Doctoral dissertation, City University of New York, 1985.