

DESCRIPCIÓN BIOMÉTRICA DE LA VOZ A PARTIR DE LA ESTRUCTURA BIOMECÁNICA DE LA CUERDA VOCAL

P. Gómez, R. Fernández, A. Álvarez, L. M. Mazaira, V. Rodellar, J. I. Godino

Grupo de Informática Aplicada al Procesado de Señal e Imagen, Universidad Politécnica de Madrid, Campus de Montegancedo, s/n 28660 Boadilla del Monte, Madrid

e-mail: pedro@pino.datsi.fi.upm.es

RESUMEN

La parametrización del comportamiento dinámico de la cuerda vocal puede ser útil en diferentes campos del procesado de la voz, como en síntesis, detección de patologías del aparato fonador, o en identificación y verificación del locutor. En el presente trabajo trata de la determinación de los parámetros biomecánicos de la cuerda vocal en forma no invasiva, generando estimaciones de las masas, elasticidades y pérdidas asociadas al modelo de cuerda vocal compuesto por el cuerpo y la cubierta elástica de la misma (lamina propia). Se estimaciones de los valores efectivos de dichos parámetros generados para una base de datos de individuos normofónicos. Estos resultados muestran una discrepancia de un orden de magnitud respecto a los valores clásicamente utilizados en modelado directo de la cuerda vocal. Ésta se explicaría mediante el examen del patrón de vibración longitudinal de la cuerda. La parametrización presentada produce descriptores biométricos aplicables en logopedia, rehabilitación y autenticación.

1. INTRODUCCIÓN¹

La parametrización de la onda glótica resulta de gran utilidad en múltiples campos, como en detección e identificación de la patología del aparato fonador [9][3], en el modelado de la cuerda vocal [1], en la síntesis de la voz [11], en la descripción biométrica e identificación del locutor [12], etc. La parametrización directa de la señal de voz íntegra contiene información propia de la biometría del locutor (estructura y dimensiones del tracto vocal, biomecánica de las cuerdas vocales, etc.), pero está sujeta a variaciones fuertemente condicionadas por el proceso articulatorio. Esto significa que rasgos propios de la biometría y de la fonética acústica aparecerán intermezclados. En el escenario propio del reconocimiento de voz es conveniente independizar los

mismos (reconocimiento independiente del locutor). En el contexto de caracterización biométrica es posible que también los rasgos fonéticos aporten información sobre la personalidad del locutor, pero un tratamiento individualizado de ambos niveles podría generar resultados más precisos si los mismos se integran de forma adecuada, siguiendo las tendencias actuales en identificación y verificación del locutor [2]. Parece procedente realizar una caracterización de la biometría del sistema de producción de onda glótica donde se haya separado la influencia del tracto vocal, en la creencia de que la parametrización por separado de los rasgos biométricos de tracto vocal y de pulso glótico pueden describir mejor las características de un locutor dado. Por dicha razón se ha desarrollado una metodología para determinar la biomecánica de la cuerda vocal a partir de la onda glótica obtenida por eliminación de la influencia del tracto vocal. En trabajos previos [8][4] se ha demostrado que la biomecánica de las cuerdas vocales (masas, resortes y pérdidas) es directamente responsable de la distribución espectral de potencia de los correlatos glóticos dinámicos que pueden derivarse de la traza de voz, como es la excitación glótica (*fente glótica*). Bajo ciertas condiciones es posible obtener estimaciones directas de estos parámetros suponiendo que las distribuciones espectrales de los correlatos glóticos están directamente relacionadas con ciertas funciones de la dinámica de la cuerda vocal [6]. El desbalance entre cuerdas, y sus desviaciones con respecto a las estimaciones obtenidas a partir de una muestra de la población general de locutores puede estar directamente relacionado con la patología de la glotis [7]. En el presente trabajo se proporcionan descripciones y estimaciones de los parámetros biomecánicos de las estructuras de cuerpo-cubierta de la cuerda vocal. Se aportan resultados obtenidos a partir de una base de datos que incluye individuos adultos y adolescentes de ambos sexos. Las estimaciones generadas se comparan con los valores utilizados en modelado clásico de la cuerda vocal [14]. Las discrepancias entre los valores teóricos y los obtenidos por este método de estimación ponen en evidencia la necesidad de revisar la forma en que la

¹Este trabajo ha sido subvencionado por los proyectos TIC2002-2273, TIC2003-08756 y TEC2006-12887-C02-00 del Plan Nacional de I+D+I, Ministerio de Educación y Ciencia.

cuerda vocal es considerada en el modelado clásico de la misma.

2. MODELADO DE LA CUERDA VOCAL

El presente estudio se basa en el conocido modelo de 3 masas de Story y Titze [14] que se muestra en la Figura 1.a, donde el cuerpo de la cuerda (formado por los músculos responsables de tensar la misma) aparece representado por las masas M_{bl} y M_{br} (para las cuerdas izquierda y derecha, respectivamente), y por los resortes K_{bl} y K_{br} que describen el comportamiento elástico de los tejidos del cuerpo de la cuerda, ligando la misma a las paredes del cartílago tiro-aritenoide.

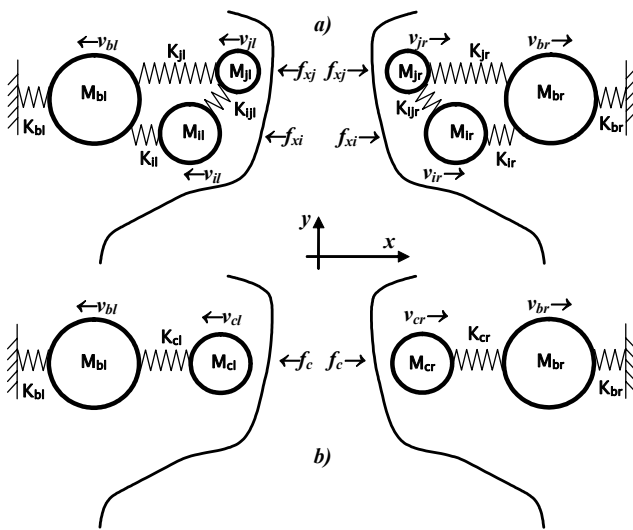


Figura 1. a) Modelo de 3 masas de las cuerdas vocales, donde se muestran los parámetros biomecánicos y las variables dinámicas involucradas en el mismo. b) Modelo de 2 masas derivado del anterior donde las masas de la cubierta se han considerado rígidamente ligadas.

La cubierta de la cuerda (compuesta por una estructura epitelial en capas) se representa por un par de masas por cuerda dado por $M_{il,r}$ y $M_{jl,r}$ ligadas a las masas $M_{bl,r}$ del cuerpo de la cuerda por medio de resortes $K_{il,r}$ y $K_{jl,r}$, interacopladas por los resortes $K_{ijl,r}$. Las masas y resortes son parámetros discretos del modelo asociados a estructuras de tipo puntual. En una simplificación dinámica solo se permite que el sistema se mueva a lo largo del eje x , siendo f_{xij} las fuerzas que actúan sobre cada masa (externas, debido a la presión translótica), siendo $v_{il,r}$, $v_{jl,r}$ y $v_{bl,r}$ las velocidades lineales de las respectivas masas. Si el estudio fuese orientado a poner de manifiesto el papel jugado por la cubierta de la cuerda, en relación con el fenómeno de vibración de la misma conocido como *onda mucosa*, los resortes de acoplamiento entre masas de la cubierta presentarían valores acotados de rigidez, admitiendo que ambas masas pudiesen vibrar en forma relativamente libre. Pero como el presente estudio desea poner de relieve el papel relativo jugado entre las estructuras del cuerpo y de la cubierta de la cuerda, se supondrá que los resortes que acoplan las masas de

cubierta deben verse como ligaduras fuertes (barras de rigidez ilimitada), de modo que la estructura completa puede reducirse al modelo más simple de 2 masas que se muestra en la Figura 1.b, donde

$$\begin{aligned} M_{cl,r} &= M_{il,r} + M_{jl,r} \\ K_{cl,r} &= K_{il,r} + K_{jl,r} \end{aligned} \quad (1)$$

De este modo el estudio de la dinámica de la cuerda vocal puede dividirse en dos, uno dedicado al cuerpo de la cuerda, y otro a la cubierta de la misma. La metodología de procesamiento de la señal de voz utilizada en el presente estudio se basa en la obtención de la fuente glótica, una vez eliminado el tracto vocal de la traza de voz por filtrado inverso, según [1], y en la separación de la onda glótica en dos componentes: una señal de duración de un ciclo de fonación completo (intervalo entre dos cierres glóticos consecutivos, incluyendo la fase de cierre y la de apertura) denominada por Titze [16] como la *onda acústica promedio*, y una componente que conserva los contenidos de alta frecuencia, que se denomina *componente dinámica de cubierta* o también *correlato de onda mucosa*.

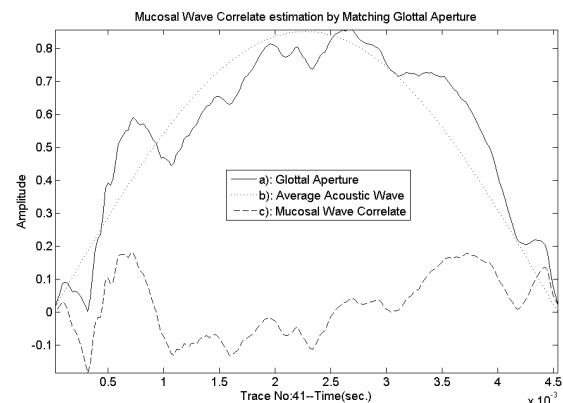


Figura 2. Separación de la fuente glótica (línea continua) en la onda acústica promedio (línea de puntos) y la *componente dinámica de cubierta* (línea de trazos).

La *fente glótica* puede relacionarse con la *apertura glótica*, considerada como la distancia promedio entre las masas de cubierta (consideradas puntuales) durante un ciclo de fonación, y la *componente dinámica de cubierta* asociada con el movimiento relativo de las masas de la cubierta $M_{cl,r}$ con respecto a las masas del cuerpo $M_{bl,r}$ (Figura 1.b). Para llevar a cabo dicha separación se han ensayado diferentes algoritmos, entre ellos el filtrado paso bajo [5]. En la Figura 2 se muestran ambas componentes de la fuente glótica después de su separación para un ciclo de fonación completo.

3. ESTIMACIÓN DE LOS PARÁMETROS DE LA CUERDA VOCAL

En el proceso de estimación de los parámetros del cuerpo y la cubierta de la cuerda vocal se utiliza el equivalente electrodinámico representado por el sistema dado en la Figura 3

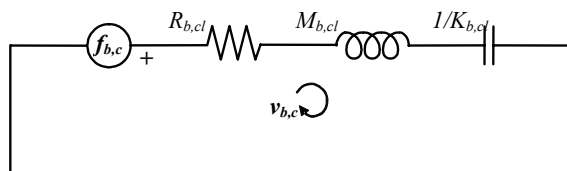


Figura 3. Equivalentes electrodinámicos del cuerpo (b) y la cubierta (c).

Se supondrá que la densidad espectral de potencia de los componentes de la *fente glótica* se hallan relacionados (salvo por un factor de escala) con el módulo al cuadrado de la admitancia de entrada de cada equivalente electromecánico.

3.1. Estimación de los parámetros del cuerpo de la cuerda

La estimación de los parámetros del cuerpo de la cuerda a partir de la *onda acústica promedio* no resulta complicada, toda vez que su densidad espectral de potencia es suave y predecible, ya que la mayoría de la variabilidad presente en la *fente glótica* reside en la *componente dinámica de cubierta*. El proceso de estimación se resume en los siguientes pasos

$$M_{bl,r} = \frac{\omega_2}{\omega_2^2 - \omega_r^2} \left[\frac{T_r - T_2}{T_r T_2} \right]^{1/2} \quad (2)$$

siendo ω_r la frecuencia de resonancia característica del sistema electromecánico equivalente, dada por

$$\omega_r^2 = \frac{K_{bl,r}}{M_{bl,r}} \quad (3)$$

donde el módulo al cuadrado de la admitancia de entrada viene dado por

$$T(\omega) = \frac{1}{\left[R_{bl,r}^2 + \omega^2 M_{bl,r}^2 \right]^2} \quad (4)$$

siendo ω la desviación de frecuencia relativa al punto de resonancia, dada por

$$\omega = \frac{\omega^2 - \omega_r^2}{\omega} \quad (5)$$

reflejando T_r y T_2 los valores de la admitancia dada en (4) para el punto de resonancia y para su frecuencia doble

$$T_r = T(\omega = \omega_r) = \frac{1}{R_{bl,r}^2} \quad (6)$$

$$T_2 = T(\omega = 2\omega_r)$$

El procedimiento de estimación debe detectar el valor del *pitch*, que se utiliza para estimar ω_r . La determinación de T_r y T_2 se lleva a cabo sobre la densidad espectral de potencia de la *onda acústica promedio*. Ello conduce a la determinación de las pérdidas a partir de (6), y de éstas a las masas a partir de (2) y de las tensiones por medio de (3). Como avance a los resultados que se presentan en la sección 4, en la Figura 5 (plantillas 35-37) se pueden ver los resultados

de la estimación de los parámetros biomecánicos para una traza de voz de 0,2 sg. de duración, correspondiente a una fonación de la vocal /a/ por el sujeto 198 (ver la Tabla 1) que se ofrecen junto con su desviación estadística para dar fe de la fiabilidad de las estimaciones.

3.2. Estimación iterativa de los parámetros de la cubierta de la cuerda

La estimación de los parámetros de la *componente dinámica de cubierta* se lleva a cabo en dos partes. Primero se realiza una estimación de tipo grueso siguiendo los pasos (2)-(6) como en el caso de la *onda acústica promedio* anteriormente referido. Como la densidad espectral de potencia de la *componente dinámica de cubierta* es más irregular, y ofrece mayor variabilidad, las estimaciones de los pares $[\omega_r, T_r]$ y $[\omega_n, T_n]$ resultan menos robustas, por lo que se lleva a cabo un refinamiento iterativo a posteriori. Éste se basa en la minimización de la siguiente función de coste

$$L = \int_0^{\omega_s} [T_u(\omega) - T_i(\omega)]^2 d\omega \quad (7)$$

donde ω_s es la *frecuencia de Nyquist*, T_u es la densidad espectral de potencia de la *componente dinámica de cubierta*, y T_i es la estimación iterativa de (4). Esta función muestra un mínimo único, que puede obtenerse, bien por medio del establecimiento de una expresión funcional (lo que no es sencillo dado el carácter no lineal de la misma) o por medio de refinamientos paso a paso de la estimación utilizando la técnica de descenso por gradiente, dada por

$$M_{bl,r}^i = M_{bl,r}^{i-1} \left[1 - 4 \int_0^{\omega_s} (T_u - T_i) T_i^2 \omega^2 d\omega \right] \quad (8)$$

donde i es el índice de iteración, repitiéndose ésta hasta que (7) se reduzca a un valor tan pequeño como se desee. Los resultados de dicha iteración para un ciclo de fonación correspondiente a la misma muestra de voz estudiada anteriormente se muestran en la Figura 4.

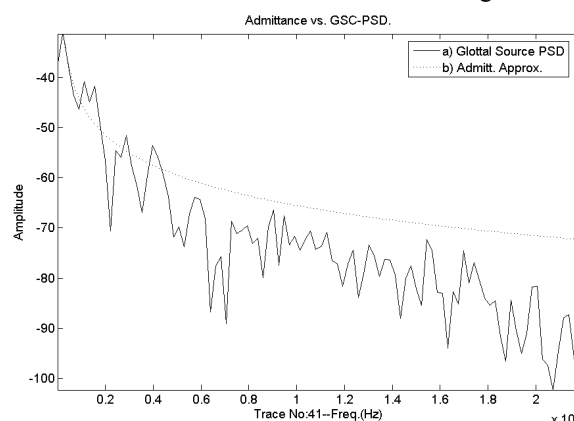


Figura 4. Ajuste de la admitancia del modelo equivalente de 1 masa (línea de puntos) sobre la densidad espectral de potencia (en dB) de la componente dinámica de cubierta (línea continua) tras la última iteración para la estimación de la masa.

Los resultados de estimar los parámetros biomecánicos de la cubierta para cada ciclo de fonación contenido en la traza de voz referida se muestran en la Figura 5 (plantillas 41-43).

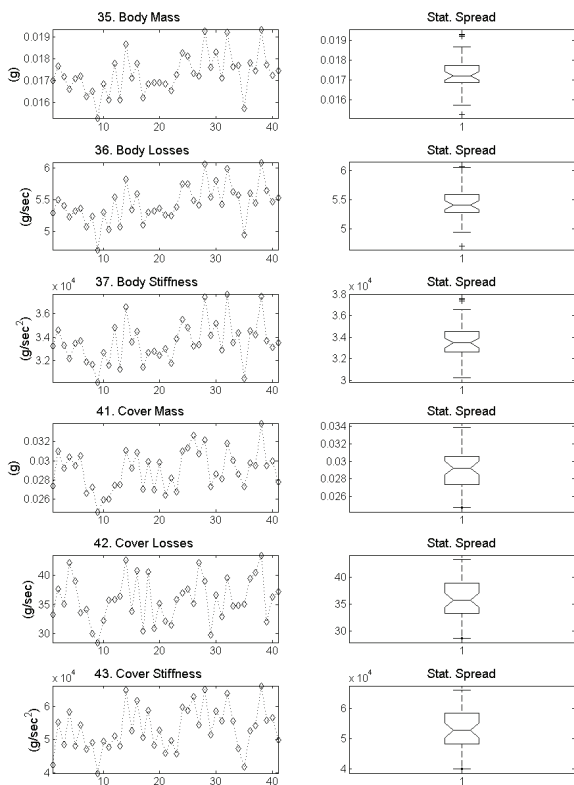


Figura 5. Columna izquierda: Estimaciones de los parámetros de masa, pérdidas y tensión para el cuerpo (35-37) y la cubierta (41-43) de la cuerda del sujeto 198 en cada ciclo de fonación (eje x). Columna derecha: dispersión estadística de cada parámetro.

El refinamiento de la estimación por iteraciones sucesivas puede suponer variaciones de entre un 20% y un 50% del valor medio de la estimación total.

4. RESULTADOS Y DISCUSIÓN

Para estudiar la dispersión de los parámetros biomecánicos típicos en una población con voz sana, se seleccionaron 100 sujetos normofónicos, de los que se obtuvieron diferentes grabaciones. La mitad de los sujetos grabados eran adultos, con edades comprendidas entre 18 y 50 años, y la otra mitad, adolescentes de entre 12 y 18 años. Los locutores se hallaban igualmente distribuidos por sexos. Se realizó la determinación de su condición de normofonía a partir de muestras electroglotográficas y videoendoscópicas, así como por una valoración subjetiva de su índice GRABS [10]. Los resultados que se presentan corresponden a un subconjunto elegido al azar correspondiente a 24 sujetos, de los cuales 16 son adultos, y 8 son adolescentes (la distribución por sexos es paritaria en ambos casos). Las características de cada sujeto se muestran en la Tabla 1.

TfP	Char	M _b	R _b	K _b	M _c	R _c	K _c
191	YF	0.00209	0.626	3.78x10 ³	0.00220	2.24	3.24x10 ³
192	AM	0.00803	1.160	3.33x10 ³	0.00610	4.26	2.27x10 ³
193	YF	0.00521	1.370	7.57x10 ³	0.00477	6.50	7.80x10 ³
194	YM	0.01040	2.230	9.53x10 ³	0.00797	8.36	8.54x10 ³
195	YM	0.00563	1.380	6.73x10 ³	0.00367	4.42	4.69x10 ³
196	AM	0.00463	0.807	2.80x10 ³	0.00290	2.49	2.10x10 ³
197	YF	0.00396	1.270	8.01x10 ³	0.00560	7.00	1.50x10 ⁴
198	YM	0.01730	5.440	3.36x10 ⁴	0.02900	35.9	5.33x10 ⁴
199	AM	0.00584	0.855	2.52x10 ³	0.00445	3.38	3.92x10 ³
19A	YF	0.00451	1.180	6.74x10 ³	0.00406	6.28	1.16x10 ⁴
19B	AM	0.01180	2.100	7.40x10 ³	0.01060	8.61	9.59x10 ³
19C	AM	0.00950	1.410	4.10x10 ³	0.00695	4.65	2.52x10 ³
19D	YM	0.00910	2.990	1.91x10 ⁴	0.00811	10.6	1.49x10 ⁴
19E	AM	0.00542	0.852	2.80x10 ³	0.00381	2.77	1.70x10 ³
19F	AM	0.00674	1.180	4.11x10 ³	0.00526	4.81	4.52x10 ³
1A0	AM	0.00543	0.839	2.49x10 ³	0.00362	2.39	1.25x10 ³
1A5	AF	0.00729	1.950	1.06x10 ⁴	0.00793	7.68	8.73x10 ³
1AA	AF	0.00499	1.380	8.01x10 ³	0.00400	6.11	1.11x10 ⁴
1AB	AF	0.00523	1.440	7.88x10 ³	0.00465	5.61	7.29x10 ³
1AE	AF	0.00513	1.580	9.82x10 ³	0.00766	9.84	1.46x10 ⁴
1B1	AF	0.00406	1.180	6.73x10 ³	0.00591	7.19	8.05x10 ³
1B9	AF	0.00420	1.230	7.18x10 ³	0.00396	5.30	7.21x10 ³
1BF	AF	0.00436	1.190	6.40x10 ³	0.00390	3.96	4.58x10 ³
1C6	AF	0.00419	1.140	6.19x10 ³	0.00227	2.94	2.73x10 ³

Tabla 1. Parámetros biomecánicos para el cuerpo y la cubierta obtenidos a partir de un conjunto de 24 sujetos normofónicos. Características del locutor: AM – adulto masculino, AF – adulto femenino, YM – adolescente masculino, YF – adolescente femenino.

Las grabaciones incluyeron tres realizaciones de la vocal /a/ de unos 3 sg. de duración cada una, de las cuales se seleccionaron fragmentos de 0,2 seg. de duración para el análisis. Estos fragmentos se procesaron para obtener la onda acústica promedio y la componente dinámica de cubierta, a partir de las cuales se estimaron los parámetros biomecánicos de cuerpo y cubierta para cada ciclo de fonación. Las estimaciones para cada parámetro biomecánico correspondientes al sujeto 198 se muestran en secuencia en la columna izquierda de la Figura 5, evaluándose sus promedios y desviaciones típicas, que se adjuntan en las plantillas de la columna derecha de la misma figura. Los promedios resumidos para los 24 sujetos de la base de datos se dan en unidades cgs en las columnas 3-5 (cuerpo) y 6-8 (cubierta) de la Tabla 1. Puede verse a partir de las estimaciones de dichos parámetros biomecánicos que las diferencias fisiológicas en la masa dinámica del cuerpo no son relevantes en sí mismas, pues si bien las masas tienden a ser algo superiores en el caso masculino que en el femenino, y en adultos que en adolescentes, el máximo estimado no corresponde a un individuo adulto masculino (como cabría suponer) sino a un adolescente masculino (sujeto 198). La distribución resulta ser mucho más relevante en el caso de las tensiones de cuerpo, donde en general, la voz adolescente y femenina tiende a ser más tensa que la adulta y masculina. Respecto a las masas de cubierta, se puede decir que se hallan casi siempre por debajo de las de cuerpo, salvo en el caso del sujeto 198, donde las masas de cuerpo y cubierta son de 17,3 mg. y 29 mg. respectivamente. Esto mismo se observa para los sujetos 191, 197, 1A5, 1AE,

y *IB1*, aunque las diferencias no son tan acusadas. Por lo que se refiere a las tensiones de cubierta, se puede ver que en muchos casos pueden llegar a superar las de cuerpo (por ejemplo en los sujetos *193*, *197*, *198*, *199*, *19A*, *19B*, *19F*, *1AA*, *1AE*, *1B1* y *1B9*). El parámetro de elasticidad es el responsable del ajuste de la frecuencia de resonancia, ya que es ajustable mediante el cambio de posición gestual del proceso tiro-aritenoide, y por tanto depende de la altura tonal de la fonación. En todos los casos estudiados el factor de pérdidas asociado a la cubierta es mayor que el observado para el cuerpo de la cuerda. Ello podría deberse a que la mayor parte de la turbulencia y el calentamiento por choque de cuerdas se generan sobre y alrededor de la cubierta.

4.1. Unicidad de las estimaciones

Uno de los problemas subyacentes en este tipo de metodología es el de la unicidad de las estimaciones de masa y elasticidad, ya que el conjunto $\{M, K\}$ acepta infinitos pares de valores mientras cumplan que la relación $K/M = \omega_r^2$ arroje un mismo valor. El factor de pérdidas, admite una referenciación más estricta, ya que viene dado por el ancho de banda relativo de las funciones de densidad espectral de potencia, medido como la altura relativa del pico de resonancia respecto a su anchura. Por otro lado, se pueden formular modelos de pérdidas basados en el proceso de aportación energética debido a la presión transglótica y a la disipación debida al choque y a la turbulencia generada en los períodos próximos al cierre glótico para cada ciclo de fonación, que predicen con bastante exactitud el comportamiento promedio de las pérdidas a lo largo de varios ciclos de fonación vecinos. En base a este hecho, se puede normalizar la densidad espectral de potencia en cada ciclo con referencia a la frecuencia de resonancia y la duración del intervalo de fonación, que permite obtener estimaciones bastante realistas del factor de pérdidas, siendo éste el punto de referencia para determinar el factor de escala de cada ciclo según (6).

4.2. Consistencia de las estimaciones

La afirmación de la validez de las estimaciones no es simple ya que obtener estimaciones directas de los valores de los parámetros asociados a las cuerdas vocales de un sujeto dado *in vivo* es poco menos que imposible. Estimaciones directas de la dinámica de las cuerdas en vibración, utilizando videoendoscopia [15] son posibles para cierto tipo de sonidos (vocales abiertas, sobre todo), a partir de las cuales se pueden establecer procedimientos de estimación similares al mostrado, aunque con los mismos problemas de autenticación, a los que habría que añadir el carácter semi-invasivo de la técnica endoscópica. Pueden diseñarse otros métodos basados en videoendoscopia para estimar la masa de las cuerdas a partir de imágenes de las mismas durante la fonación, aunque en este caso se tropieza con el problema de que la masa estática

(cuerda en reposo) puede diferir sustancialmente de la masa dinámica de la misma (parámetro derivado de las estimaciones dinámicas entre fuerzas y aceleraciones), pues ésta última es solo una fracción de la masa estática, ya que no se puede afirmar que toda la estructura tisular de la cuerda se halle igualmente involucrada en la vibración (en otras palabras, las cuerdas vocales en vibración no son asimilables a las cuerdas libres pinzadas por sus extremos, como el caso de las de una guitarra, pues se hallan conectadas lateralmente en toda su longitud sobre paredes semirrígidas de tipo cartilaginoso por medio de tejido conjuntivo que las une al proceso tiro-aritenoide, lo cual limita grandemente la contribución de parte del cuerpo de la cuerda a la vibración). Esto se muestra con claridad para una cuerda idealizada de sección constante tipo columna, como la que se muestra en la Figura 6.

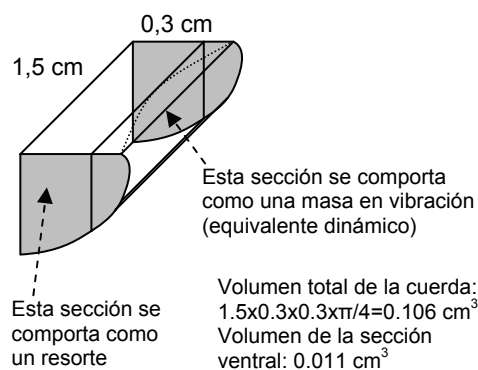


Figura 6. Diferencia entre masa estática y dinámica de la cuerda vocal.

En efecto, a partir de imágenes obtenidas por videoendoscopia estroboscópica se puede apreciar cómo una cuerda sana de las características de la expuesta vibra de forma desigual en toda su longitud, con máxima elongación hacia su zona ventral central, mientras que los extremos apenas vibran. La elongación en la zona central puede ser de alrededor de 1 mm , mientras que la zona correspondiente al cuerpo que no participa en la vibración, sino que sufre compresiones y expansiones puede extenderse en aproximadamente 2 mm de los 3 mm típicos de dimensión transversal para la cuerda. Si se realizan estimaciones de la masa total de una cuerda de estas características se obtienen valores de alrededor de 100 mg . (ver Figura 6), mientras que la masa efectiva de la sección vibrante sería de unos 10 mg . Dependiendo del uso de la cuerda realizado por cada locutor en su gesto fonador propio, el porcentaje de cuerda implicado podría moverse entre un 10% y un 30% del total de la masa estática de la cuerda, lo que incluiría cuerpo y cubierta. Estas estimaciones pueden compararse contra los valores utilizados por los investigadores en modelado directo de la cuerda vocal. Por ejemplo, Story usa en su modelado masas de $0,058 \text{ g}$ (cuerpo), $0,082 \text{ g}$ (cubierta supraglótica) y $0,094 \text{ g}$ (cubierta infraglótica) y parámetros de rigidez de 112.250 dinas/cm , 157.140 dinas/cm y 203.660

dinas/cm respectivamente para un modelo de 3 masas [13], con cuerdas similares a las de la Figura 6 (1,5 cm de largo por 0,3 cm de perfil). Teniendo en cuenta estas consideraciones puede verse que las estimaciones dadas en la Tabla 1 resultan ser más consistentes que los valores utilizados en simulación. Así, la masa total estática utilizada por Story en una sola cuerda sería de 0,234 g (añadiendo la de cuerpo y las dos de cubierta), lo que sobrepasa en más de dos veces la masa que le correspondería de acuerdo con las dimensiones de la cuerda. Al sobredimensionar la masa en este tipo de modelado, habrá que sobredimensionar la tensión que sufre la cuerda, con lo que se utilizan valores que difícilmente serían soportables por el tejido fisiológico sin sufrir un deterioro notorio. Si se asume que los valores de las masas dinámicas son solo una parte reducida de la masa estática, las estimaciones de las tensiones adoptan valores más acordes con la realidad, compatibles con una mayor resistencia al deterioro. En cualquier caso, la tensión de la cuerda es una de las causas de la degeneración de la misma y de la aparición de patología, por lo que su estimación directa resulta de la mayor importancia. El seguimiento de dicho valor en diferentes situaciones gestuales (como en canto, por ejemplo) permitirá a foniatras y logopedas realizar una mejor labor diagnóstica, terapéutica y rehabilitadora.

5. CONCLUSIONES

El trabajo presentado describe una posible metodología para resolver un problema inverso de interés en el tratamiento de la voz: estimar los parámetros biomecánicos de la cuerda vocal a partir de la distribución espectral de potencia de la onda glótica obtenida por filtrado inverso de la voz. Los resultados para un conjunto de locutores de muestra con comportamiento normofónico indican que los rasgos fisiológicos del locutor juegan un papel menos relevante en la distribución de las estimaciones de las masas de las cuerdas que en las de sus tensiones. Por otro lado se han estimado valores para las masas un orden de magnitud por debajo de los utilizados clásicamente en el modelado directo de las cuerdas vocales. Una revisión cuidadosa de la estructura de una cuerda vocal típica y de la implicación de la misma en la vibración apoyaría resultados más en consonancia con los estimados por el procedimiento propuesto que con los utilizados en modelado clásico. Las desviaciones de los parámetros biomecánicos en modo intra-locutor pueden servir como indicadores de esfuerzo vocal abusivo, y en modo inter-locutor pueden utilizarse como detectores de patología y como elementos de caracterización biométrica.

6. BIBLIOGRAFÍA

[1] Akande, O. O. and Murphy, P. J., "Estimation of the vocal tract transfer function with application to glottal wave analysis", *Speech Communication*, Vol. 46, No. 1, May 2005, pp. 1-13.

[2] García, D., Aguilar, J., González, J. and Ortega, J., "Support Vector Machine Fusion of Idiolectal and Acoustic Speaker Information in Spanish Conversational Speech", *Proc. of the ICASSP'03*, Vol. II, pp. 229-232, 2003.

[3] Godino, J. I., Gomez, P., Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. *IEEE Trans Biomed Eng.* Vol. 51, 2004, pp. 380-384.

[4] Gómez, P., Fernández, R., Nieto, A., Díaz, F., Fernández, F. J., Rodellar, V., Álvarez, A., Martínez, R., "Evaluation of Voice Pathology Based on the Estimation of Vocal Fold Biomechanical Parameters", *Journal of Voice*, March 17, 2006.

[5] Gómez, P., Godino, J. I., Álvarez, A., Martínez, R., Nieto, V., Rodellar, V., "Evidence of Glottal Source Spectral Features found in Vocal Fold Dynamics", *Proc. of the ICASSP'05*, 2005, pp. 441-444.

[6] Gómez, P., Godino, J. I., Díaz, F., Álvarez, A., Martínez, R., Rodellar, V., "Biomechanical Parameter Fingerprint in the Mucosal Wave Power Spectral Density", *Proc. of the ICSLP'04*, 2004, pp. 842-845.

[7] Gómez, P., Lázaro, C., Fernández, R., Nieto, A., Godino, J. I., Martínez, R., Díaz, F., Álvarez, A., Murphy, K., Nieto, V., Rodellar, V., Fernández, F. J., "Using biomechanical parameter estimates in voice pathology detection", *Proc. of the 4th Int. Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications MAVEVA'05*, Florence, Italy, October 29-31, pp. 2005.

[8] Gómez, P., Martínez, R., Díaz, F., Lázaro, C., Álvarez, A., Rodellar, V., Nieto, V., "Estimation of vocal cord biomechanical parameters by non-linear inverse filtering of voice", *Proc. of the 3rd Int. Conf. on Non-Linear Speech Processing NOLISP'05*, Barcelona, Spain, April 19-22 2005, pp. 174-183.

[9] Hadjitodorov S, Mitev P. A computer system for acoustic analysis of pathological voices and laryngeal diseases screening. *Med Eng Phys.* 2002;24: 419-429.

[10] Hirano, M., Hibi, S., Yoshida, T., Hirade, Y., Kasuya, H., and Kikuchi, Y., "Acoustic analysis of pathological voice. Some results of clinical application," *Acta Otolaryngologica*, vol. 105, no. 5-6, pp. 432-438, 1988.

[11] McLaughlin, S., "Nonlinear Speech Synthesis", *Proc. of the EUSIPCO'02*, Vol III, 2002, pp. 211-218.

[12] Plumpe, M. D., Quatieri, T. F., Reynolds, D. A., "Modeling of the Glottal Flow Derivative Waveform with Application to Speaker Identification", *IEEE Trans. on Speech and Audio Proc.*, Vol. 7, No. 5, 1999, pp. 569-586.

[13] Story, B. H., "An overview of the physiology, physics and modeling of the sound source for vowels", *Acoust. Sci. & Tech.*, Vol. 23, No. 4, 2002, pp. 195-206.

[14] Story, B. H., and Titze, I. R., "Voice simulation with a bodycover model of the vocal folds", *J. Acoust. Soc. Am.*, Vol. 97 (1995), pp. 1249-1260.

[15] Švec, J. C., Horáček, J., Šram, F., Veselý, J., "Resonance properties of the vocal folds: *In vivo* laryngoscopic investigation of the externally excited laryngeal vibrations", *J. Acoust. Soc. Am.*, Vol. 108(4), pp. 1397-1407, 2000.

[16] Titze, IR. "Summary Statement", *Workshop on Acoustic Voice Analysis*, National Center for Voice and Speech, 1994.