

CAMBIOS EN LA CLASIFICACIÓN DE LA VOZ FEMENINA DESPUÉS DEL TRATAMIENTO DE LA PATOLOGÍA VOCAL.

R. Fernández-Baíllo¹, P. Gómez¹, C. Ramírez² J. I. Godino¹B. Scola²

¹Laboratorio de Comunicación Oral (GIPASI) Universidad Politécnica de Madrid, Campus de Montegancedo, s/n 28660 Boadilla del Monte, Madrid. e-mail: roberto@junipera.datsi.fi.upm.es

²Servicio ORL, Hospital Provincial Universitario Gregorio Marañón, C/ Doctor Esquerdo

Resumen

En general en los estudios de patología es importante tener en cuenta factores relacionados con el género. En voz esto se hace más decisivo al existir claras diferencias acústicas y biomecánicas que vienen marcadas por el género. Algunas de estas diferencias son observables en el estudio acústico, pero la relación voz -género se hace más notable al analizar la fuente glótica. La fuente glótica puede ser obtenida mediante métodos de filtrado inverso en los cuales se elimina toda influencia del tracto vocal en la señal de voz resultante. Un posterior análisis del perfil del espectro de densidad de energía y del perfil de la onda glótica nos permite establecer una serie de puntos singulares que son cruciales para la discriminación del género. Igualmente se pueden extraer de la fuente glótica una serie de parámetros acústicos y biomecánicos que permiten clasificar la voz en función del grado patología. Para el siguiente trabajo se muestra los resultados obtenidos en la clasificación de patología para una voz de mujer, en condición pre y post tratamiento, al compararlo con un grupo de voces no-patológicas de su mismo género.

1. Introducción.

La dinámica de los pliegues vocales está determinada por su estructura anatómico-histológica [1], de tal forma que diferencias o cambios en esta estructura ocasionan cambios dinámicos. Es sabido que existen diferencias de género que afectan a las características de los pliegues

vocales, siendo los pliegues vocales en la mujer de una menor masa y longitud en relación al hombre. Esto hace que existan diferencias en la dinámica de los pliegues vocales en la mujer y en el hombre y que el resultado final de la vibración (fuente glótica) también lo sea. Así pues, la fuente glótica incluye bastantes rasgos relativos a la dinámica glótica. La obtención de la fuente glótica se realiza por un proceso de filtrado inverso, en el cual se elimina toda influencia del tracto vocal en la señal de voz. La fuente glótica resultante ofrece tres posibilidades de estudio:

1. *Perfil de la Onda Glótica y del Correlato de Onda Mucosa (MWC)*. (Este es un análisis temporal el cual permite identificar los puntos característicos de apertura y cierre del ciclo de fonación [2]. Este tipo de estudio es muy útil para la identificación de la patología vocal. También permite discriminar, relativamente bien, la voz no-patológica por género. Pero su utilidad para valorar el comportamiento de una voz con respecto a su grupo de género requiere una explicación amplia, la cual va más allá del propósito de este trabajo, ya que incluye factores relacionados con la técnica y el patrón fonatorio)
2. *Singularidades del Espectro de Densidad de Energía del Correlato de Onda Mucosa (SPC-MWC)*.
3. *Parámetros Acústicos y Biomecánicos*.

La metodología de este trabajo se basará en las singularidades del espectro y los posibles parámetros extraídos. El propósito es valorar en la clasificación de un sujeto en condición pre y post-cirugía en relación a su grupo de género.

2. Metodología

2.1. Estimación del género basada en el SPD-MWC.

La voz humana se puede definir como una onda compleja periódica resultante de la transmisión a un medio aéreo de una onda de presión originada por la vibración de los pliegues vocales (*Fuente Glótica*,) y que es sometida a un proceso de transformación en su paso por los órganos supraglóticos (*Filtrado*) hasta su salida por los labios (*Radiación*).

Mediante métodos de filtrado inverso podemos estimar la fuente glótica, eliminando toda influencia del trato vocal [3] [4] [5]. La señal resultante del proceso de filtrado está directamente relacionada con la dinámica de

los pliegues vocales. Posteriormente la Fuente Glótica obtenida puede ser descompuesta 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 [6] 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 (MWC), el cual puede ser estudiado desde una perspectiva temporal o espectral.

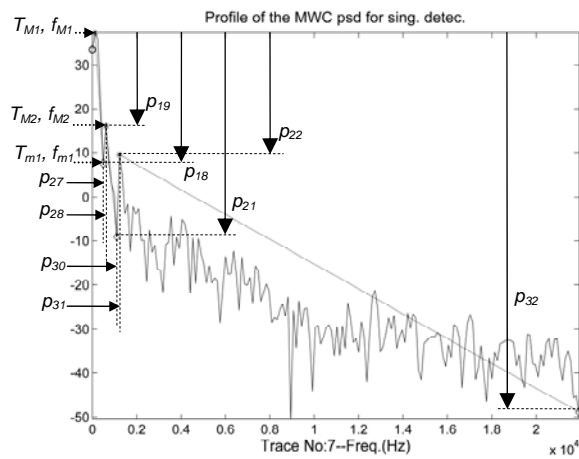


Figura 1. Espectro del densidad de energía del correlato de la onda mucosa mostrando la típica distribución de picos y valles.

La metodología usada para la discriminación del género se basará en el estudio del perfil del espectro de densidad de energía del correlato de onda mucosa (Ver figura 1). Se propone una parametrización basada en la identificación de singularidades de este espectro. De cada punto singular, pico o valle, tomará una posición y amplitud relativa en el espectro en relación al pico mayor (Ver Figura 1). La distribución y tipología de los picos-valles está directamente relacionada con las masas y las tensiones de los pliegues vocales [7] [8]. Para la discriminación del género se utilizaron 100 voces no-patológicas (50 masculinas y 50 femeninas), obtenidas de la base de datos registrada en el proyecto MAPACI [9]. El

objetivo es conseguir una distribución que permita la discriminación de las voces en función de su género con independencia del pitch.

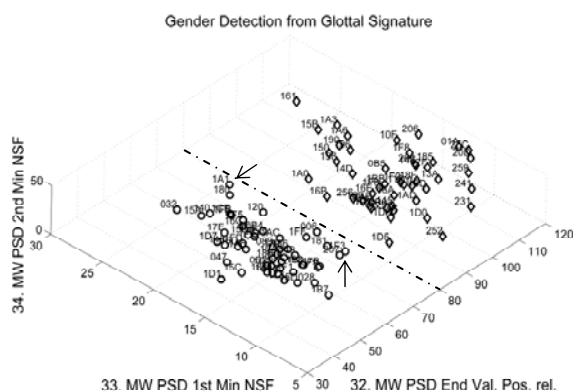


Figura 2. Resultados obtenidos de la clasificación de 100 voces no-patológicas. Se aprecia como los dos grupos de género (femeninas (○) y masculinas (◇)) son correctamente discriminados. Tan solo se ha fallado en la clasificación de dos sujetos (flechas) masculinos que se aproximaron al grupo femenino. (#1A1 y #1F3).

2.2. Análisis de los cambios de clasificación en un grupo de género.

Una vez clasificada la voz por género, se utilizarán 24 voces del grupo femenino, cuya condición es de no-patológicas, como grupo de control. Este grupo se confrontará con de 24 voces patológicas femeninas (incluyendo diversas patologías, todas ellas de tipo orgánico). Posteriormente se realizará un estudio de la dispersión del grupo de voces femeninas (no-patológicas y patológicas) basado en parámetros acústicos y biomecánicos.

De la Fuente glótica se pueden extraer parámetros clásicos $\{p_1...p_{14}\}$ como el Jitter, Shimmer y HNR. Pero también una serie de parámetros biomecánicos $\{p_{35}...p_{46}\}$ referidos a las masas y tensiones de los pliegues vocales. Estos parámetros se obtienen a partir del estudio del perfil del espectro de densidad del correlato de la onda mucosa. En total para el estudio se utilizaron un total de 46 parámetros (Ver figura 1) relacionados con cada ciclo de fonación de cada registro de voz considerado [10]. Posteriormente mediante técnicas de PCA se localizaron los 16

parámetros que resultaron más relevantes para la clasificación de la patología vocal para un grupo de género dado.

2.3. Caso de Estudio.

Se utilizaron registros de voz correspondientes a una mujer de 34 años, no fumadora y actriz de teatro. La cual refería una disfonía crónica, sintomatología vocal en forma de picor y limitación durante el habla y el canto. La exploración médica concluyó la presencia de un pólipo en el pliegue vocal izquierdo. Fue sometida a tratamiento quirúrgico y dos meses después se tomaron de nuevo registros de voz. Los registros de voz pre y post-tratamiento se añadieron al grupo de voces femeninas considerándolos como si fueran dos hablantes distintos. El registro pre-tratamiento (#0E8) se incluyó con las 24 voces patológicas femeninas y el post-tratamiento con las 24 voces normales femeninas (#2DC).

3. Resultados y Discusión

3.1. Resultados de la discriminación de género basada en el *SPD-MWC*.

En un primer estudio se comprobó que la clasificación por género es sensible al perfil espectral [11]. En la Figura 2 se muestran los resultados de dispersión en los que se ve como los dos grupos son perfectamente discriminados. Posteriormente se desarrolló un análisis estadístico para determinar que parámetros eran más decisivos para la clasificación ciega del género. Los parámetros, del total de los extraídos, más sensibles al género fueron p32, p28, p30 y p19, en orden de mayor a menor sensibilidad. Todos estos parámetros hacen referencia a la posición relativa del complejo “Picos-valles” en el perfil espectral. Este estudio deja de manifiesto que existen diferencias espectrales entre los grupos de género. Dado que el perfil espectral está relacionado con las masas y las tensiones de los pliegues vocales, se puede decir que hay diferencias estructurales y dinámicas en la voz en función del género. Características propias de género que deben ser consideradas al afrontar un estudio de discriminación de la patología vocal. Es decir para la detección y clasificación de un sujeto como patológico hay que compararlo, al menos, con su grupo de género.

3.2. Resultados de la clasificación de la voz femenina después del tratamiento.

La clasificación de la patología vocal para un grupo de género basada en el análisis de componentes principales (PCA) del grupo de parámetros extraídos permite diferenciar tres grupos principales. El grupo de voces no-patológica (▼) las cuales están muy agrupadas, el grupo de voces patológicas (◇) la cuales tienen una mayor dispersión y un grupo de voces que están en un estadio intermedio (○) entre la no-patología y la patología (Ver Figura 3). La razón de que el grupo de voces patológicas tenga una mayor dispersión viene dada porque se incluyeron patologías diversas y cada patología tiene una clínica vocal propia.

Como consecuencia del tratamiento el paciente se produce un claro cambio en la clasificación vocal. El cual antes de la intervención quirúrgica está claramente incluido en el grupo de voces patológicas, pero posteriormente queda perfectamente localizado en el grupo de voces normales.

4. Conclusiones.

Del presente trabajo se concluye que la fuente glótica incluye información a partir de la cual se pueden inferir datos referentes a la estructura y dinámica de los pliegues vocales. Igualmente, de la fuente glótica se puede conseguir el correlato de onda mucosa cuyo estudio mediante un análisis espectral permite lograr un perfil en el cual se identifican una serie de puntos singulares que son relevantes para la discriminación del género. En los estudio de la patología de la voz hay que considerar el género del hablante. Ya que se han encontrado diferencias de perfil espectral ligadas al género. Esto quiere decir que existen diferencias de masas y tensiones, y por tanto estas características deben ser valoradas de forma importante ya que la patología vocal de carácter orgánico ocasiona cambios en la masa y tensión del pliegue vocal.

Agradecimientos

Este trabajo se ha desarrollado gracias al proyecto TIC2003-08756, TEC2006-12887-C02-00 del plan nacional I+D+i, del Ministerio de Educación y Ciencia, CCG06-UPM/TIC-0028 del Plan Regional de Investigación Científica e Investigación Tecnológica de la Comunidad de

Madrid y por el proyecto HESPERIA del Programa CENIT, Centro para el Desarrollo Tecnológico Industrial, Ministerio de Industria, España.

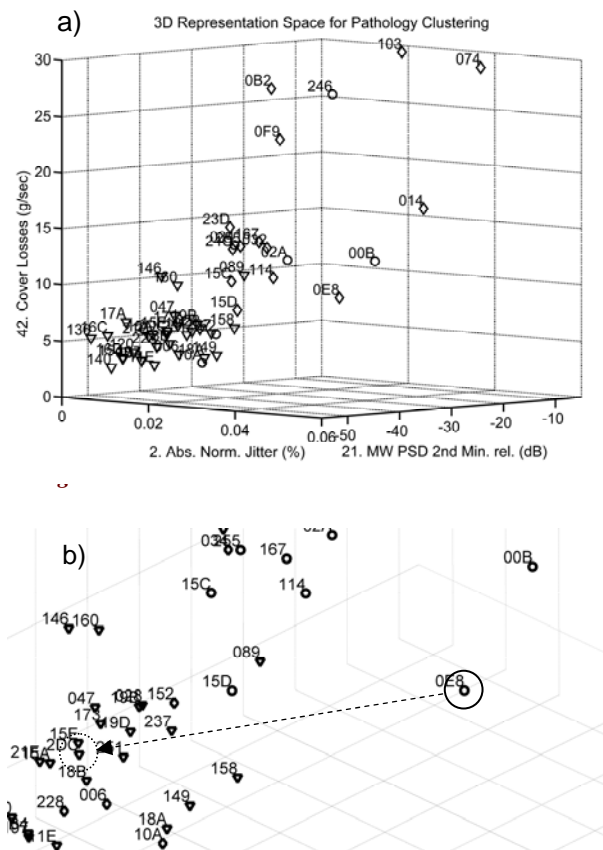


Figura 3. Distribución de os registros de voces patológicas vs. normales para el grupo femenino. El clustering está mostrando los resultados en base a tres parámetros del subconjunto de los doce más relevantes. a): Representación de global del clustering. b): Imagen ampliada para observar el cambio en la clasificación del caso estudio (#0E8 / #2DC).

Bibliografía.

- [1] Hiarino M, Kusita S, Kiyokawa K and cols. Basic and clinical investigation. *Otologia (Fukuska)*, 1975; 21:231-242.
- [2] Fernández-Baillo R, Gómez Vilda P. Clinical Voice Description Based on the Mucosal Wave Correlate. Proc. Of IALP'07. Copenhagen, August 5-9, 2007.
- [3] Sokorin VN, Leonov AS, Thrushkin AV. Estimation of stability and accuracy of inverse problem solution of the vocal tract. *Speech Communication*. Vol. 30, No 1, January 2000, pp 55-74.
- [4] Hayking S. Adaptive filter theory. Prentice-Hall, Englewood Cliffs, NJ. 1996.
- [5] Alku, P., "Parameterisation Methods of the Glottal Flow Estimated by Inverse Filtering", Proc. of VOQUAL'03, Geneva, August 27-29 (2003), pp. 81-87.
- [6] Titze IR. Summary Statement. *Workshop on Acoustic Voice Analysis*. National Center of Voice and Speech. 1994.
- [7] Gómez, P., Rodellar, V., Álvarez, A., Lázaro, J. C., Murphy, K., Díaz, F., Fernández-Baillo, R., "Biometrical Speaker Description based on Vocal Cord Parameterization", Proc. of the ICASSP'06, Vol. 2, 2006, pp. 1036-1039.
- [8] Gómez, P., Fernández-Baillo, R., Rodellar, V. Mazaira, L. M., Martínez, R., Álvarez, A., Godino, J. I., "Biometry of Voice based on the Glottal-Source Spectral Profile", Proc. of the SAFE07: Workshop on Signal Processing Applications for public Security and Forensics, Washington DC, April 12-13, 2007, pp. 120-128.
- [9] <http://www.mapaci.com>
- [10] Gómez, P., Fernández-Baillo., et al., Evaluation of voice pathology based on the estimation of vocal fold biomechanical parameters, *Journal of Voice*. Vol. 21, No. 4, 2007, pp. 450-476.