

Extracción de Características Cualitativas del Llanto de Bebé y su Clasificación para la Identificación de Patologías Utilizando Modelos Neuro-Difusos

C. A. Reyes-García^{1*}, A. A. Torres-García², M. A. Ruiz-Díaz³

¹ Ciencias y Tecnologías Biomédicas, Instituto Nacional de Astrofísica, Óptica y Electrónica, Tonantzintla, Pue., México

² Coordinación de Ciencias Computacionales, Instituto Nacional de Astrofísica, Óptica y Electrónica, Tonantzintla, Pue. México

³ Universidad Politécnica de Tlaxcala, Zacatelco, Tlaxcala, Mexico

* kargaxxi@inaoep.mx

¹kargaxxi@inaoep.mx, ²kcobaim@hotmail.com, ³mariaantoniaruiz@uptlax.edu.mx

Resumen— El artículo presenta la implementación de un sistema diseñado para extraer las características cualitativas acústicas más representativas de la onda del llanto de bebé y posteriormente las utiliza para su clasificación con el fin de identificar patologías. Para la clasificación se utilizan varios modelos híbridos neuro-difusos que permiten asegurar la precisión de la clasificación. El sistema presentado, una vez validado, permitirá hacer diagnósticos no invasivos en infantes desde etapas tempranas. Se presenta una descripción de la interfaz utilizada para la extracción de características, de los modelos utilizados para clasificación, las bases de datos utilizadas y algunos resultados importantes obtenidos.

Palabras clave— análisis del llanto infantil, identificación automática de características cualitativas, clasificación de llanto, diagnóstico no invasivo.

I. INTRODUCCIÓN

El llanto de los recién nacidos es una expresión funcional de las necesidades biológicas básicas, las condiciones emocionales o psicológicas como el hambre, el frío, el dolor, los calambres e incluso la alegría [1]. Requiere un esfuerzo coordinado de varias regiones cerebrales, principalmente tronco encefálico y sistema límbico, y está relacionado con la respiración y el mecanismo pulmonar. Sus características reflejan el desarrollo y posiblemente la integridad del sistema nervioso central. Por lo tanto, el análisis del llanto infantil es una herramienta complementaria no invasiva adecuada para evaluar el estado físico de los lactantes particularmente importante en el caso de los recién nacidos prematuros. Específicamente, la distinción entre un llanto regular y uno con anomalías es de interés clínico.

Siendo económico y sin contacto, el estudio del llanto del bebé recién nacido ha tenido un crecimiento sobresaliente en las últimas décadas. Varios estudios se refieren tanto al análisis auditivo subjetivo de la voz y el habla como al análisis acústico automático en adultos. Sin embargo, en lo que respecta al llanto del recién nacido, existen pocos métodos automáticos, algunos basados en enfoques clásicos como la transformada de Fourier y el análisis de

autocorrelación [1-5] y otros en técnicas paramétricas [6-7]. Dichos métodos permiten estimar las principales características acústicas, como la frecuencia de vibración de las cuerdas vocales, las frecuencias de resonancia del tracto vocal, la duración del llanto, etc. En los últimos años, varios autores proponen métodos de clasificación para una amplia gama de patologías. Reyes et al. [8-10] han investigado recién nacidos normales, sordos y asfixiantes a través de redes neuronales, selección genética y lógica difusa, Poel et al [11] presentan resultados sobre la clasificación del llanto en recién nacidos trastorno normal y relacionado con la hipoxia que usa redes neuronales de función de base radial con un rendimiento general de clasificación del 85%.

El llanto del recién nacido refleja el desarrollo y posiblemente la integridad del sistema nervioso central, por lo que su análisis es un medio atractivo no invasivo para evaluar el estado físico de los bebés desde etapas muy tempranas de vida. En el análisis del llanto infantil, es de gran importancia identificar las características cualitativas, ya que proporcionan información relevante que permite identificar variaciones o semejanzas entre llantos normales y patológicos, así como diferenciar entre diferentes patologías. Generalmente el análisis de las características cualitativas se realiza de forma manual, por medio de la percepción visual (inspeccionando espectrogramas) y auditiva (escuchando las grabaciones de llanto) de los médicos expertos, quienes, de acuerdo a lo que ven y escuchan, pueden emitir un diagnóstico. En este trabajo se presenta un método que permite la detección automática de unidades de llanto y se presenta un modelo llamado “dodecagrama” que permite identificar automáticamente el tipo de melodía de las unidades de llanto, y finalmente con los valores de la frecuencia fundamental, se identifican automáticamente características cualitativas distintivas como los *shifts*, *glides* y concentraciones de ruido de las unidades de llanto.

II. METODOLOGÍA

El proceso de extracción de características cualitativas inicia con la detección de unidades de llanto. Esta se lleva a cabo con la finalidad de utilizarlas para un análisis posterior, tal es el caso de [12], en el cual se analiza la duración media

de las señales de llanto, la media de la frecuencia fundamental de los llantos así como su forma melódica.

En las grabaciones se encuentran sonidos que no son de utilidad para el análisis del llanto infantil, tal es el caso, de los sonidos producidos por el ambiente y los sonidos inspiratorios producidos por los infantes antes de emitir una unidad de llanto, al cual los médicos llaman inspiración, que en algunos casos, no es útil para análisis posteriores y por lo tanto estos sonidos son tomados en cuenta. Otro punto importante y que se debe considerar es la variedad de ambientes y dispositivos en los que se adquieren las grabaciones así como la intensidad y el tipo de llanto del infante. En las grabaciones pueden presentarse llantos con *pitch* alto, con *pitch* bajo, llantos graves, llantos agudos, etc. y la variación de la intensidad, ya que el infante puede reducir o aumentar la intensidad de su llanto en la misma grabación.

Entonces, la detección de llanto está definida como un procedimiento mediante el cual los sonidos de llanto son seleccionados y separados de la grabación, la detección de llanto se ha llevado a cabo manualmente en varios trabajos como [1], [13].

El método de detección de unidades de llanto es parte de una interfaz implementada en MATLAB. Basados en pruebas experimentales se eliminaron unidades de llanto menores a 200 ms por ser sonidos de muy corta duración que no aportan información útil para un análisis posterior. Ya que en estas unidades es muy difícil determinar otras características de interés para los médicos, como la forma melódica. Se definió también un umbral de energía ($U(e)$) aplicado a las señales similar al propuesto en [14] y que en nuestros experimentos se obtiene de la siguiente manera:

$$U(e) = \frac{E_n}{4} \quad (1)$$

donde E_n es la energía de la señal a corto tiempo.

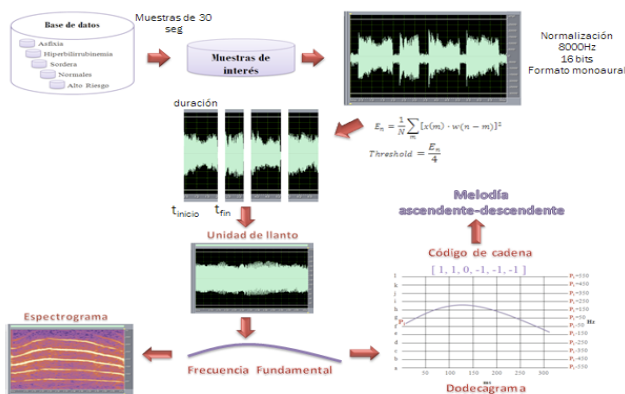


Fig. 1. Esquema general del proceso de extracción y procesamiento de características cualitativas

En la Fig.1 se muestra paso a paso el funcionamiento del método propuesto iniciando con la detección de unidades de llanto. Posteriormente se hace la identificación del tipo de melodía, para la que se desarrolló una modificación del método presentado en [15] al cual hemos llamado método del dodecagrama, y el cual forma parte de la interfaz de extracción de características. En la Fig. 1 se muestra el dodecagrama en el cual cada unidad de llanto se posiciona al centro de las líneas f y g. El valor de las líneas de las líneas está determinado por el valor de la frecuencia fundamental de la primera ventana, el siguiente paso es codificar la unidad de llanto mediante las siguientes reglas: 1 si el valor de la frecuencia fundamental pasa a un renglón superior, 0 si el valor de la frecuencia fundamental se mantiene en el mismo renglón y -1 si el valor de la frecuencia fundamental pasa a un renglón inferior. Dónde el número 1 corresponde a un incremento de la frecuencia fundamental, el 0 sin cambios, y el -1 a un decremento de la frecuencia fundamental. Por lo cual, podemos ver que la unidad de llanto que se muestra en la Fig. 1 tiene una forma melódica de tipo: descendente-ascendente.

Para identificar los *shifts* se miden las diferencias de las frecuencias fundamentales a lo largo de la señal, si la diferencia sobrepasa los 100 Hz se considera *shift* (puede haber más de uno en una unidad de llanto).

De la misma forma para identificar los *glides* se miden las diferencias de las frecuencias fundamentales a lo largo de la señal, si la diferencia sobrepasa los 600 Hz en un espacio de tiempo muy corto se considera *glide* (puede haber más de uno en una unidad de llanto). Un *vibrato* se define como una serie de ondas con al menos cuatro movimientos de ascenso y descenso en la frecuencia fundamental

Métodos de Clasificación:

Para los experimentos de clasificación se implementaron y adaptaron cuatro clasificadores neuro-difusos adaptativos, que son; Clasificador neuro -difuso adaptativo con modificadores lingüísticos (f NFC_LH), clasificador neuro-difuso con modificadores lingüísticos y características seleccionadas (NFC_LH_FS), clasificador neuro-difuso de gradiente conjugado (NFC) y clasificador neuro-difuso de gradiente conjugado acelerado (NFC_acelerado). Los sistemas SCGNFC y SSCGNFC son optimizados mediante algoritmos de gradiente conjugado escalado. En estos dos sistemas, el algoritmo k-means se utiliza para inicializar las reglas difusas. También, la función de pertenencia Gaussiana solamente se utiliza para descripciones de conjuntos difusos. Los otros dos sistemas se basan en modificadores lingüísticos (LH) sintonizados por gradiente conjugado escalado.

Resultados experimentales

El sistema se probó con dos conjuntos de muestras, uno obtenido de la base de datos BabyChillanto del Instituto Nacional de Astrofísica Óptica y Electrónica con muestras de llanto de infantes mexicanos, de la cual se utilizó para prueba una base de llantos de *hiperbilirrubinémicos* que se comparó contra otra base de llantos de *hipotiroidicos*, ambas colectadas en México, la primera en el Instituto Nacional de Rehabilitación y la segunda por el Dr. Mario Mandujano de la UAM Xochimilco. Además, se probó el sistema en otro ambiente para diferenciar entre llanto de bebé prematuro y a término registrados en el Hospital Universitario de Lieja, en Bélgica para evaluar la versatilidad del sistema. Los resultados de clasificación mostrados reflejan las posibilidades de aplicación del sistema completo.

Resultados de Experimentos Seleccionados

Los resultados muestran que las características cualitativas son útiles también para apoyar diagnósticos médicos perceptivos, ya que es posible mostrar en el espectrograma el punto en donde se presenta cada una de ellas.

Tabla 1. Resultados de clasificación para cada clasificador para las clases de *HiperBilirrubinemia vs Hipotiroidismo*, obtenidos en diferentes iteraciones junto con la exactitud promedio y la desviación estándar.

NFC	f NFC_LH	NFC acelerado	NFC_LH_FS
100.00	100.00	100.00	100.00
50.00	50.00	50.00	100.00
100.00	50.00	100.00	50.00
100.00	100.00	100.00	100.00
100.00	100.00	0.00	50.00
100.00	50.00	100.00	50.00
50.00	50.00	50.00	50.00
100.00	66.67	33.33	100.00
Exactitud promedio			
73.33	60.00	60.00	68.33
Desviación standard			
37.02	32.58	40.21	27.72

Tabla 2. Resultados de clasificación para cada clasificador para las clases de *Prematuros vs A Término*, obtenidos en diferentes iteraciones junto con la exactitud promedio y la desviación estándar.

NFC	f NFC_LH	NFC acelerado	FC_LH_FS
100.00	100.00	66.67	100.00
100.00	100.00	100.00	83.33
100.00	100.00	100.00	85.71
83.33	83.33	83.33	83.33
83.33	83.33	83.33	50.00
85.71	42.86	71.43	71.43
100.00	83.33	100.00	66.67
85.71	71.43	100.00	85.71
Exactitud promedio:			
84.52	81.90	84.52	72.14
Desviación estándar			
17.96	17.43	15.64	20.65

IV. DISCUSIÓN

Con el método desarrollado en la base de llantos *Hiperbilirrubinémicos vs Hipotiroidicos* se detectaron 2115 unidades de llanto de 24 grabaciones. Y en la de *Prematuros Vs A Término* se detectaron 2983 unidades de llanto de un total de 65 grabaciones. En cuanto a la precisión en el reconocimiento se puede ver que en el caso de *HiperBilirrubinemia vs Hipotiroidismo* el reconocedor NFC fue el que obtuvo mejores resultados con un 73.33 % y en el caso de *Prematuros vs A Término* dos clasificadores, NFC y NFC acelerado, obtuvieron la precisión más alta que es de 84.52 %. Aunque los resultados parecen no ser muy elevados de cualquier manera son muy alentadores, ya que fueron obtenidos utilizando solo las características cualitativas. Todos los clasificadores son entrenados a través de una serie de iteraciones siguiendo un esquema del tipo *10-fold Cross Validation*. En las Tablas 1 y 2 se muestran los resultados parciales que cada clasificador obtiene al final de cada iteración (*fold*). Si bien se han obtenido mejores resultados utilizando características cuantitativas, el estudio presentado es relevante ya que la extracción de características cualitativas es importante ya que permiten hacer una descripción muy cercana a la utilizada por los médicos especialistas al hacer un estudio perceptivo de las ondas de llanto infantil a través del cual se pueda realizar un diagnóstico.

V. CONCLUSIÓN

Perceptivamente existe una gran similitud entre las ondas de llanto de bebés padeciendo hiperbilirrubinemia y los que padecen hipotiroidismos, por lo que la obtención de una precisión de reconocimiento de 73.33 % utilizando solo las características cualitativas refleja un alto potencial para su posterior aplicación en el desarrollo de sistemas de diagnóstico no invasivo. De igual manera, en el caso de la clasificación de llantos prematuros y a término [17] en la que se obtuvo 84.52 % permite entrever una mayor posibilidad de aplicación de las características cualitativas en otros dominios distintos al diagnóstico. En este caso la mayor precisión se puede explicar por la diferencia existente entre los dos tipos de llanto ocasionada por la mayor madurez del aparato fonatorio de los bebés nacidos a término. El sistema aun requiere de más estudios y experimentar con diversas modificaciones, como la utilización de una combinación de características cualitativas y características cuantitativas. También se pretende aplicar métodos evolutivos de selección para utilizar solo las características que conduzcan a la obtención de los mejores resultados. Y también es conveniente probar

con la aplicación de diversos tipos de clasificadores diferentes a los utilizados en este estudio, como los desarrollados en [18]. Con los resultados obtenidos es posible asegurar que el uso de las características cualitativas para clasificación con diversos fines, abre una gran ventana de oportunidades de aplicación limitadas solo por la imaginación de los desarrolladores.

RECONOCIMIENTOS

Los autores agradecen al proyecto MX14MO06 “Técnicas de análisis y clasificación de voz y expresiones faciales: aplicación a las enfermedades neurológicas en recién nacidos y adultos” del programa ejecutivo de cooperación científica y tecnológica México-Italia financiado por AMEXID de la SRE y el Ministerio de Asuntos Exteriores de Italia.

BIBLIOGRAFÍA

- [1]. Michelsson K, Michelsson O. Phonation in the newborn, infant cry. *Int J Pediatr Otorhinolaryngol* 1999; vol 49 (1), pp S297-S30.
- [2]. Fuller BF. Acoustic discrimination of three types of infant cries. *Nurs Res* 1991; 40(3):156-160.
- [3]. Laufer MZ, Horii Y. Fundamental frequency characteristics of infant non-distress vocalization during the first twenty-four weeks. *J Child Lang* 1977; 4(02):171-184.
- [4]. Reggiannini B, Sheinkopf SJ, Silverman HF, Li X, Lester BM. A Flexible Analysis Tool for the Quantitative Acoustic Assessment of Infant Cry. *J Speech Lang Hear Res* 2013; 56(5):1416-1428.
- [5]. Sirviö P, Michelsson K. Sound-spectrographic cry analysis of normal and abnormal newborn infants. *Folia Phoniatica et Logopaedica* 1976, 28(3):161-173.
- [6]. Hevia-Montiel N, Molino-Minero-Re E, Carrillo-Bermejo AJ. Tortuosidad discreta como medida morfométrica en tumores cerebrales. *Rev Mex Ing Biom.* 38(1):188-98. 2017 DOI: 10.17488/RMIB.38.1.13
- [7]. A. Fort, C. Manfredi, “Acoustic analysis of newborn infant cry signals”, *Med. Eng. Phys.*, vol. 20, pp. 432-442, 1998.
- [8]. Kheddache Y, Tadj C. Resonance Frequencies Behavior in Pathologic Cries of Newborns. *Journal of Voice* 2014. <http://dx.doi.org/10.1016/j.jvoice.2014.04.007>
- [9]. Reyes-Galaviz, O. F., Tirado, E. A., & Reyes-García, C. A. (2004). Classification of infant crying to identify pathologies in recently born babies with ANFIS (pp. 408-415). in *Lecture Notes in Computer Science* 3118, Springer Berlin Heidelberg. *Computers Helping People with Special Needs*, edited by Klaus Miesenberger et. al., Springer, Berlin, 2004, pp 408-415.
- [10]. Cano, O.S.D., Escobedo, D.L., Suaste, I., Ekkel, T., Reyes García, C.A.: A Combined Classifier of Cry Units with New Acoustic Attributes. In: Martínez-Trinidad, J.F., Carrasco Ochoa, J.A., Kittler, J. (eds.) *CIARP 2006*. LNCS, vol. 4225, pp. 416-425. Springer, Heidelberg (2006).
- [11]. Amaro-Camargo, E., & Reyes-García, C. A. (2007). Applying statistical vectors of acoustic characteristics for the automatic classification of infant cry. In *Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues* (pp. 1078-1085). Springer Berlin Heidelberg.
- [12]. Poel M., Ekkel T. Analyzing infant cries using a committee of neural networks in order to detect hypoxia related disorder 2006 *International Journal on Artificial Intelligence Tools* 15 3 397-410.
- [13]. K. Michelsson, K Eklund, P. Leepänen, H. Lyytinen, “Cry characteristics of 172 Healthy 1- to 7-Day-Old Infants”, *International Journal of Phoniatics Speech Therapy and Communication Pathology*, vol. 54, 2002
- [14]. K. Wermke, W. Mende, C. Manfredi, and P. Brusciaglioni, “Developmental aspects of infant's cry melody and formants”, *Med Eng Phys*, vol. 24(7-8), pp. 501-514, 2002.
- [15]. I Saratxaga, I Luengo, E Navas, I Hernández, J Sánchez, I Sainz, “Detección de PITCH en condiciones adversas”, IV Jornada en Tecnología del Habla, Universidad del País Vasco – Euskal Herriko Unibertsitatea pp. 13-18, 2006.
- [16]. Benyó Z. Várallyay G Jr., Illényi A. Melody analysis in the newborn infant cries. pages 11-14, December 2009.
- [17]. Orlandi, S., Garcia, C. A. R., Bandini, A., Donzelli, G., & Manfredi, C. (2016). Application of pattern recognition techniques to the classification of full-term and preterm infant cry. *Journal of Voice*, 30(6), 656-663.
- [18]. A. Rosales-Pérez, C. A. Reyes-García, J. A. Gonzalez, O. F. Reyes-Galaviz, H. J. Escalante, and S. Orlandi, “Classifying infant cry patterns by the Genetic Selection of a Fuzzy Model,” *Biomed. Signal Process. Control*, vol. 17, pp. 38-46, 2015.