

Clasificación de ECG basada en Características de Escala, Dirección y Ritmo

M Llamedo Soria ^{1,2,3}, JP Martínez Cortés ^{1,3}

¹ Grupo de Tecnologías de las Comunicaciones, Inst. de Investigación en Ingeniería de Aragón, Univ. de Zaragoza, España, {llamedom, jpmart}@unizar.es

² Universidad Tecnológica Nacional, Buenos Aires, Argentina

³ CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)

Resumen

En este trabajo se estudió el desempeño para la clasificación de latidos mediante características basadas en intervalos, ángulos y amplitudes. La mayoría de estas características mayormente fueron calculadas en ambas derivaciones del ECG, y a diferentes escalas de las respectivas transformadas wavelet. Las señales de ECG provienen de la base de datos MIT-BIH Arrhythmia, siguiendo las recomendaciones de etiquetamiento y presentación de resultados propuesta por la AAMI. Las divisiones de datos referentes a entrenamiento, prueba y validación cruzada, realizadas en este trabajo fueron hechas paciente a paciente y no latido a latido como en otros trabajos. En el set de registros de entrenamiento se aplicó un algoritmo de búsqueda flotante de características para obtener el mejor modelo. Dicho modelo, formado por 4 características se evaluó en el set de prueba, obteniendo una exactitud global del 90%; específicamente para la clase de latidos normales se obtuvo una sensibilidad (Se) del 92%, y un valor predictivo positivo (+P) del 85%; para la clase supraventricular, una Se del 88% y un +P del 93%; para la clase ventricular, una Se del 90% y un +P del 92%. El clasificador propuesto basado en características no-Euclídeas se desempeña mejor que otros clasificadores actuales, usando menos características.

1. Introducción

El análisis de la señal electrocardiográfica (ECG) proporciona una técnica no invasiva para el estudio de la actividad del corazón en sus distintas condiciones. Particularmente los algoritmos automáticos de clasificación se focalizan en el análisis del ritmo y la morfología del ECG, y específicamente en las variaciones respecto a la normalidad. Justamente, las variaciones en el ritmo, regularidad, lugar de origen y forma de conducción de los impulsos cardíacos, se denominan arritmias [1]. Mientras que algunas arritmias representan una amenaza inminente (Ej. fibrilación ventricular), existen otras más sutiles que pueden ser una amenaza a largo plazo sin el tratamiento adecuado. Es en estos últimos casos, que registros ECG de larga duración requieren una inspección cuidadosa, donde los algoritmos automáticos de clasificación representan una ayuda significativa en el diagnóstico.

En la última década se han desarrollado algunos algoritmos de clasificación de ECG (Ej. [2]-[5]), pero solo unos pocos tienen metodologías y resultados comparables, a pesar de las recomendaciones de la AAMI

[6] para facilitar la resolución de estos problemas. De dichos métodos, algunos funcionan de manera completamente automática, mientras que otros pueden aprovechar la asistencia de un experto para mejorar su desempeño. La base de datos (BD) utilizada en estos trabajos ha sido la MIT-BIH Arrhythmia [7]. En cuanto a las características utilizadas, los intervalos RR fueron usados por casi todos los grupos. También se utilizaron muestras del complejo QRS diezmado, o transformado mediante polinomios de Hermite o la descomposición wavelet (TW). Otros grupos usaron características que integran la información presente en ambas derivaciones, como el máximo del vectocardiograma (VCG) del complejo QRS (VCG_{MAX}), o el ángulo del VCG *loop* en dicho punto (VCG_{ANGLE}).

Como la metodología para la extracción de características generalmente requiere la delineación previa del ECG, en este trabajo se propone un clasificador que utilice características preferentemente calculadas durante la etapa de delineación. Dichas características deberían ser calculadas preferentemente en la TW y de manera multiderivacional (MD), resultando en un modelo robusto a los ruidos típicamente encontrados en la señal de ECG. Esto concuerda con la metodología propuesta por nuestro grupo para la delineación,

El objetivo de este trabajo es el desarrollo y evaluación de un modelo para la clasificación completamente automática de ECG, bajo la hipótesis de que las características MD y calculadas en la TW deberían proporcionar un mejor desempeño en la clasificación.

2. Metodología

2.1. Origen de las señales de ECG

En este trabajo se utilizó la BD MIT-BIH Arrhythmia para el entrenamiento y evaluación del clasificador. Esta BD consiste en 48 registros de dos derivaciones, con una duración aproximada de 30 minutos muestreados a 360 Hz. Los primeros 23 registros corresponden a la práctica ambulatoria típica, mientras que los 25 restantes fueron seleccionados debido a la presencia de arritmias ventriculares, supraventriculares y nodales. Cabe destacar que las derivaciones no son las mismas a lo largo de la BD, dependiendo de las limitaciones del paciente. Las anotaciones provistas en la BD se usaron tanto para el

entrenamiento, como para la evaluación del clasificador siguiendo las recomendaciones de la AAMI. Se siguió el mismo esquema de división de la BD propuesto en [5] respecto a entrenamiento (DS1) y prueba (DS2), para facilitar la comparación de resultados. Sin embargo, también se estudió entrenar el clasificador en un subset de DS1, dado que en los registros 201 y 207 se presentan latidos con morfologías atípicas, que empeoran la capacidad de generalizar que se pretende lograr durante el entrenamiento. Por otro lado, la clase AAMI Q (latidos inclasificables) fue descartada por su escasa representación en la BD. Finalmente se estudió una variante de etiquetado respecto a la propuesta por AAMI, al unificar la clase ventricular y fusión (de latidos normales y ventriculares) como una única clase ventricular. Esta variante en el etiquetado será referida como AAMI2. El esquema de división de datos se resume en la tabla Tabla 1.

	N	S	V	F	#Rec
DS1	45673	929	3755	412	22
DS1 _M	42502	706	3366	410	20
DS2	44053	1833	3202	388	22
Full MIT-BIH	88175	1635	7121	822	44

DS1 comprende los registros 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230. DS2 comprende los registros 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234. DS1_M es DS1 sin 201 y 207. Los registros 102, 104, 107, 217 fueron excluidos por tener latidos marcapaseados.

Tabla 1. División propuesta para la BD MIT-BIH en set de entrenamiento (DS1) y prueba (DS2). Los registros con latidos marcapaseados fueron excluidos. Las clases de latidos corresponden a N: normales, S: supraventriculares, V: ventriculares y F: fusión.

2.2. Procesamiento de señales

Todos los registros en la MIT-BIH fueron primero procesados de la manera descrita en [5] para quitar los ruidos típicos de ECG. No se ha realizado ninguna normalización de amplitud o energía, debido a que se han estudiado algunas características relacionadas con la amplitud. Como la finalidad de este trabajo es el estudio de un modelo de clasificación, la detección de QRS fue realizada por medio de las anotaciones provistas en la misma BD. Algunas características consideradas están basadas en la transformada wavelet (TW) de la señal de ECG. La TW de una señal continua se define como:

$$W_a x(b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt, a > 0 \quad (1)$$

La TW puede discretizarse usando un esquema diádico en el factor de escala $a = 2^k$ para $k \in \mathbb{Z}^+$ con igual frecuencia de muestreo en cada escala (Algorithme à trous), e implementado como un banco de filtros. Se utilizó como wavelet prototipo $\psi(t)$ una función *spline* cuadrática. Como resultado de esta transformación, la señal original de ECG puede analizarse como una derivada suavizada a diferentes escalas (o bandas de

frecuencia) y desplazamientos de tiempo. En [2] el lector interesado encontrará una discusión más completa respecto a la TW.

2.3. Extracción de características

Siguiendo las conclusiones de trabajos anteriores [2][5] se utilizaron características morfológicas y de intervalos. Como características de intervalos se obtuvo de la serie de intervalos RR las siguientes características: RR[i-1], RR[i], RR[i+1], la mediana de los últimos 10 y 20 intervalos RR como una estimación del ritmo local, y la media de los últimos 10 minutos (RR_{avg}) como una estimación del ritmo global. Como características morfológicas fueron consideradas la anchura del complejo QRS (QRS_w), el máximo del módulo del *loop* de QRS (VCG_{MAX}) y el ángulo del *loop* en dicha posición (VCG_{angle}). Otra característica relacionada a la morfología, es la escala wavelet donde el complejo QRS tiene su máxima proyección. Ya que aquellas señales que evolucionan rápidamente (como un latido normal) tenderán a estar proyectadas en las bajas escalas wavelet (o alta frecuencia). Esta característica se calcula como una suma ponderada, donde

$$A_i = \frac{1}{M} \sum_{m=1}^M |W_i x(t_m)|$$

son la media de las amplitudes absolutas de los picos detectados para las escalas $i=1, 2, \dots, 6$ y $M \leq 2$, siendo M el número de picos detectado por cada escala en los tiempos t_m . Luego se calcula la escala de máxima proyección por cada derivación (QRS_{proj}^{lead}) como:

$$QRS_{proj} = \frac{\sum_{i=1}^6 A_i \cdot i}{\sum_{i=1}^6 A_i}$$

Como las características en nuestros modelos pertenecen a dominios diversos como \mathbb{R} , \mathbb{R}^+ y S^2 (dominio angular o direccional) es necesario transformar o manejar adecuadamente los datos, previo a realizar las tareas de clasificación. En nuestro caso, asumiremos que cada característica está normalmente distribuida y por lo tanto, los valores son válidos en el dominio real. Por este motivo, las características de intervalos o morfológicas definidas en \mathbb{R}^+ deben transformarse a \mathbb{R} , por ejemplo mediante la función logaritmo. De la misma manera, las características angulares necesitan un tratamiento especial que describiremos brevemente, el lector interesado puede encontrar en [8] los detalles metodológicos. Se consideran las características angulares ϑ , como el argumento de un número complejo de módulo unitario $e^{j\vartheta}$, siendo $j^2 = -1$. Luego el valor esperado de esta variable define la dirección media y la varianza angular, contrapartes de la media y varianza tradicionales:

$$E[\exp^{j\vartheta}] = \rho_{\vartheta} e^{j\mu_{\vartheta}^c} \quad (2)$$

$$\mu_{\vartheta}^c = \arg(E[\exp^{j\vartheta}])$$

$$V_{\vartheta}^c = 1 - \|E[\exp^{j\vartheta}]\| = 1 - \rho_{\vartheta}$$

Donde ρ_θ también se conoce como longitud resultante. Luego para un modelo de dimensión F, donde Fw son los índices de las variables angulares, los estimadores de máxima verosimilitud son:

$$(\hat{\mu}_x)_f = \begin{cases} \frac{1}{M} \sum_{m=1}^M x_f(m) & \text{if } f \in Fw \\ \arg \left(\frac{1}{M} \sum_{m=1}^M \exp(jx_f(m)) \right) & \text{else} \end{cases} \quad (3)$$

$$\hat{\Sigma}_x \approx \frac{1}{M-1} \sum_{m=1}^M \mathbf{x}'(m) \mathbf{x}'(m)^T \quad (4)$$

siendo

$$\mathbf{x}'(m) = \begin{cases} x_f(m) - (\hat{\mu}_x)_f & \text{if } f \in Fw \\ (x_f(m) - (\hat{\mu}_x)_f)_{\text{mod } 2\pi} & \text{else} \end{cases}$$

Como se observa en la ecuación 4, $\hat{\Sigma}_x$ puede calcularse mediante la dirección media $(\hat{\mu}_x)_f$ y los datos crudos sin necesidad de otro cálculo adicional.

2.4. Clasificación de ECG

Para esta tarea se utilizaron funciones discriminantes lineales y cuadráticas. Según el criterio de máxima probabilidad a posteriori (MAP), la ecuación general cuadrática resultante para los vectores de características \mathbf{X} , y la i -ésima clase se define:

$$g_i(\mathbf{x}) = -\frac{1}{2} \mathbf{x}^T \Sigma_i^{-1} \mathbf{x} + \mu_i^T \Sigma_i^{-1} \mathbf{x} - \frac{1}{2} \mu_i^T \Sigma_i^{-1} \mu_i - \frac{1}{2} \log(|\Sigma_i|) + \log(P(\omega_i)) \quad (5)$$

donde μ_i , Σ_i y $P(\omega_i)$ son el vector medio, la matriz de covarianza y la probabilidad a priori de la i -ésima clase. Las estimaciones μ_i y Σ_i son calculadas en DS1, considerando probabilidades a priori iguales. La regla de

clasificación asigna un latido representado por \mathbf{X} a la clase i que tenga máximo valor $g_i(\mathbf{x})$.

Cuando las matrices de covarianza Σ_i se consideran iguales para todas las clases, la función discriminante se vuelve lineal en \mathbf{X} . En este caso, la única matriz de covarianza Σ se calcula de la misma manera que en [5]. Por último, la implementación de todas las rutinas de clasificación se realizó sobre el *toolbox* PRtools [9].

2.5. Selección del Modelo y Reducción de Dimensión

Es bien sabido que los modelos de baja dimensionalidad tienden a generalizar mejor a ejemplos no presentados durante el entrenamiento, resultando de esta manera clasificadores más robustos y por tanto realistas. Por este motivo se buscó seleccionar las mejores características para la clasificación por medio de un algoritmo de búsqueda flotante [10]. Por otro lado, siguiendo los resultados de la búsqueda se han explorado otros modelos de manera empírica.

2.6. Evaluación del Desempeño

Todos los modelos fueron evaluados en DS1/DS1_M usando una división en k-partes, donde cada parte será un registro (o paciente) de los 22/20 registros presentes en cada esquema de división. Esto está motivado por el hecho que el clasificador deberá enfrentarse a nuevos pacientes para ser clasificados, en lugar de nuevos latidos de los mismos pacientes. En la referencia [4] se muestra cómo este último enfoque sesga de manera optimista los resultados, debiendo ser evitado.

El desempeño en la clasificación se medirá mediante la exactitud global; luego por cada clase se calculará la sensibilidad (Se) y el valor predictivo positivo (+P) como lo sugiere la AAMI [5] tanto para el set de entrenamiento como para el de prueba. A pesar de que no exista ninguna

Dataset	labeling	size	Normal		Suprav.		Ventr.		Fusion		Total		
			Se	+P	Se	+P	Se	+P	Se	+P	Acc	Se	+P
DS1	AAMI	5	64	76	74	69	31	89	93	54	75	75	67
DS1	AAMI2	7	88	78	80	91	86	87	NA ¹		85	85	85
DS1 _M	AAMI	5	83	87	81	93	79	88	94	74	84	84	86
DS1 _M	AAMI2	4	92	86	86	93	88	88	NA ¹		89	89	89

¹ El etiquetado AAMI2 considera las clases Fusión y Ventricular como una única clase Ventricular.

Tabla 2. Desempeño para cada modelo separando todas las clases presentes en los conjuntos de entrenamiento. Los resultados se expresan en porcentajes.

Dataset	labeling	size	Normal		Suprav.		Ventr.		Fusion		Total			
			Se	+P	Se	+P	Se	+P	Se	+P	Acc	Se	+P	
Este trabajo Chazal et al. [4]	DS1 _M	AAMI	5	72	90	90	94	89	86	91	75	86	86	86
	DS1 _M	AAMI2	4	92	85	88	93	90	92	NA ¹	90	90	90	
	DS1	AAMI	48	87	52	76	39	78	82	89	86	80	80	68

¹ El etiquetado AAMI2 considera las clases Fusión y Ventricular como una única clase Ventricular.

Tabla 3. Evaluación final para los modelos con mejor desempeño, entrenados en DS1 /DS1_M y evaluados en DS2 separando las clases AAMI/2. Todos los resultados están expresados en porcentajes.

Reference	Algorithm				Totals
	f	n	s	v	
F	353	18	1	16	388
N	9053	31918	1884	1331	44186
S	19	43	1648	126	1836
V	268	24	47	2875	3214
Totals	9693	32003	3580	4348	49624

Reference	Algorithm			Totals
	n	s	v	
N	40532	2434	1220	44186
S	126	1622	88	1836
V	319	46	3237	3602
Totals	40977	4102	4545	49624

Tabla 4. Matrices de confusión obtenidas en la evaluación final en DS2, para los modelos presentados en la tabla 3. A la izquierda el modelo usando etiquetado AAMI y a la derecha usando el etiquetado AAMI2.

recomendación de la AAMI para compensar el gran desbalance en la representación de las clases en la BD MIT-BIH, en este trabajo se igualaron la cantidad de latidos por clase antes del cálculo del +P por clase. Esto equivale a modificar las matrices de confusión de la Tabla 4 de manera que todas sus filas sumen 44186. Hasta el momento este problema no se trató en ninguno de los trabajos citados.

Finalmente, el mejor modelo hallado es entrenado en DS1/DS1_M mientras que DS2 se usa exclusivamente para presentar los resultados de desempeño final.

3. Resultados

Los resultados de la selección de modelos se muestran en la tabla 2. El mejor modelo para cada combinación de set de entrenamiento y etiquetado se muestra para comparación. Como también nos interesa evaluar si DS1_M y el etiquetado AAMI2 mejoran el desempeño, seleccionamos 2 modelos para los resultados finales. Los resultados de estos modelos y las matrices de confusión se presentan en las tablas 3 y 4 respectivamente.

Los mejores modelos resaltados en la tabla 3 están compuestos por las características $RR[i]$, RR_{avg} , QRS_{proj}^1 , QRS_{proj}^2 para el etiquetado AAMI2, y para el etiquetado AAMI original se agregó QRS_{angle} al modelo anterior.

4. Discusión y Conclusiones

En primer lugar se evaluaron 4 modelos en DS1/DS1_M con los etiquetamientos AAMI/2, obteniendo mejores resultados los modelos entrenados en DS1_M. De esta manera se comprueba que la presencia de ejemplos de entrenamiento poco frecuentes, como los que se presentan en los registros 201 y 207 degradan la capacidad de generalización, empeorando el desempeño. Al evaluar el desempeño final en DS2, ambos modelos obtuvieron mejores resultados que los presentados en [4], con la ventaja de obtenerse a partir de un modelo de menor tamaño. Los mejores resultados se obtuvieron al usar el etiquetado AAMI2, con un modelo de 4 características ($RR[i]$, RR_{avg} , QRS_{proj}^1 , QRS_{proj}^2) relacionadas con el ritmo cardíaco actual y global, y con la morfología del complejo QRS.

El etiquetado alternativo AAMI2 mejoró la clasificación de las clases resultantes, esto puede deberse a las sutiles diferencias que existen entre las clases V y F en el contexto de las características elegidas. Otra razón podría

ser lo poco representada que está la clase F en la BD MIT-BIH.

Futuros trabajos deberían estudiar la capacidad de generalización de los modelos propuestos a otras BD, como por ejemplo la AHA. Otra mejora posible es la adaptación local al registro bajo estudio, con la ayuda de un experto local.

Agradecimientos

Este trabajo fue financiado por los proyectos TEC-2007-68076-C02-02 de la CICYT y GTC T-30 de la DGA. El CIBER de Bioingeniería, Biomateriales and Nanomedicina es una iniciativa del ISCIII.

Referencias

- [1] Taylor GJ. 150 Practice ECGs: Interpretation and Review. Blackwell Science, 2002. ISBN 0-632-04623-6.
- [2] Martínez J, Almeida R, Olmos S, Rocha A, Laguna P. A wavelet-based ECG delineator: Evaluation on standard databases. IEEE Transactions on Biomedical Engineering 2004; 51:570–581.
- [3] Hu YH, Palreddy S, Tompkins W. A patient-adaptable ecg beat classifier using mixture of experts approach. IEEE Transactions on Biomedical Engineering 1997;44:891–899.
- [4] Lagerholm M, Peterson C, Braccini G, Edenbrandt L, Sörnmo L. Clustering ecg complexes using hermite functions and self-organizing maps. IEEE Transactions on Biomedical Engineering 2000;47:838–848.
- [5] de Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 2004;51:1196–1206.
- [6] Testing and reporting performance results of cardiac rhythm and st-segment measurement algorithms. American National Standard, ANSI/AAMI/ISO EC57, 1998–(R)2008.
- [7] Mark R, Moody G. Mit-bih arrhythmia database 1997. <http://ecg.mit.edu/dbinfo.html>, 1997.
- [8] Bahlmann C. Directional features in online handwriting recognition. Pattern Recognition 2006;39(1):115 – 125. ISSN 0031-3203.
- [9] Duin R, Juszczak P, Paclik P, Pekalska E, deRidder D, Tax D, Verzakov S. Pr-tools, a matlab toolbox for pattern recognition, 2008. URL <http://www.prtools.org>.
- [10] Pudil P, Novovicova J, Kittler J. Floating search methods in feature selection. Pattern Recognition Letters 1994; 15(11):1119–1125.