

Biológiai jelek Hilbert-térbeli approximációja

Dózsa Tamás

Programtervező Informatikus BSc

Eötvös Loránd Tudományegyetem, Informatikai Kar

Budapest,

2017. június x.

Tartalomjegyzék

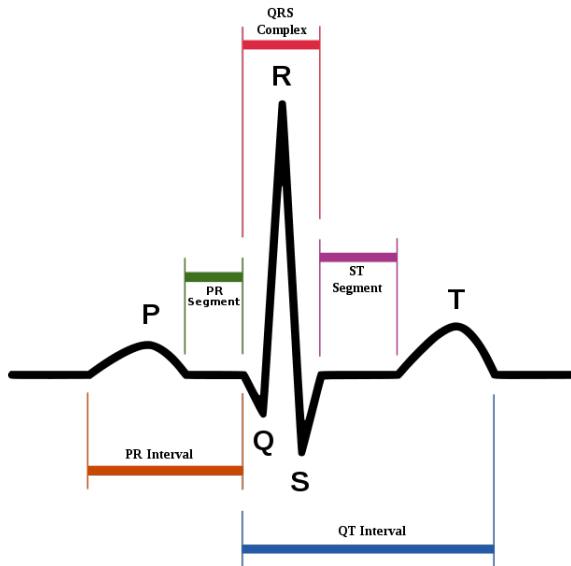
- 1 Bevezetés
- 2 Előzmények
- 3 Kidolgozott módszer
- 4 Implementáció
- 5 Tesztek, eredmények

Motiváció

- Az EKG elemzése nagy jelentőséggel bír az orvostudományban.
- Hosszú mérések, vagy több elvezetés esetén költséges a tárolás.
- A jel zajjal terhelt, ami nehezíti a pontos diagnózis felállítását.
- Hullámszegmensek meghatározása fontos probléma.

Alkalmazások

- **Tömörítés**
- Zajszűrés
- QRS detektálás
- Szívütések osztályozása



Approximáció Hilbert-terekben

Definíciók

- Legyen $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ Hilbert-tér, ahol

$$\langle f, g \rangle := \int_{-\infty}^{\infty} f(t)g(t) \, dt$$

- $\{ \Phi_k : 0 \leq k \leq n \}$ ortonormált rendszer (ONR)
- Ortogonális projekció:

$$S_n f := \sum_{k=0}^n \langle f, \Phi_k \rangle \Phi_k$$

Approximáció Hilbert-terekben

Definíciók

- Az approximáció hibája:

$$\|f - S_n f\|^2 = \|f\|^2 - \sum_{k=0}^n |\langle f, \Phi_k \rangle|^2.$$

Feladat

- EKG esetén másodpercenként több száz adat keletkezik.
- A dolgozat célja a jel **tömörítése**.
- Ezért véges dimenziós altereket használunk (Euklideszi tér).

Hermite-polinomok

Definíció

- Rekurzió:

$$H_n(x) = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x) \quad (n \geq 2)$$

$$H_{-1}(x) = 0, \quad H_0(x) = 1, \quad H_1(x) = 2x$$

- Ortogonalitás:

$$\langle H_n, H_m \rangle_\rho := \int_{-\infty}^{\infty} H_n(t) H_m(t) \rho(t) \, dt$$

$$\langle H_n, H_m \rangle_\rho = 0 \quad (n \neq m)$$

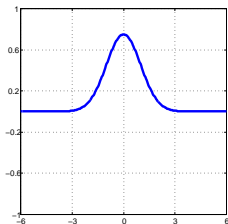
- Súlyfüggvény:

$$\rho(t) := e^{-t^2}$$

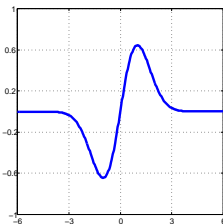
Hermite-függvények

Definíció

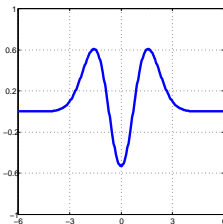
$$\Phi_n(x) := H_n(x)e^{-x^2/2}/\sqrt{\pi^{1/2}2^n n!} \quad (n \in \mathbb{N}, x \in \mathbb{R})$$



(a) $\Phi_0(x)$



(b) $\Phi_1(x)$



(c) $\Phi_2(x)$

Figure: A Hermite-függvényrendszer első három tagja.

Hermite-függvények

Tulajdonságok

- ONR
- Gyorsan tartanak 0-hoz:

$$|\Phi_n(x)| \leq M_n e^{-x^2/4} \leq M_n \quad (x \in \mathbb{R}, n \in \mathbb{N})$$

- Stabil, másodrendű rekurzió:

$$\Phi_n(x) = \sqrt{\frac{2}{n}} x \Phi_{n-1} - \sqrt{\frac{n-1}{n}} \Phi_{n-2}(x) \quad (n \geq 2)$$

$$\Phi_0(x) = e^{-x^2/2} / \pi^{1/4}, \quad \Phi_1(x) = \sqrt{2} x e^{-x^2/2} / \pi^{1/4} \quad (x \in \mathbb{R})$$

Hermite-függvények

Tulajdonságok

- Φ_n deriváltja:

$$\Phi'_n(x) = \sqrt{2n}\Phi_{n-1}(x) - x\Phi_n(x) \quad (n \geq 0, x \in \mathbb{R})$$

- Alakjuk korrelál az EKG-vel.
- Korábbi cikkek is használták ¹ ².

¹R. Jane, S. Olmos, P. Laguna, and P. Caminal, *Adaptive Hermite models for ECG data compression: Performance and evaluation with automatic wave detection*, in Proc. of the Int. Conf. on Comp. in Card., 1993, pp. 389–392.

²A. Sandryhaila, S. Saba, M. Püschel, and J. Kovacevic, *Efficient compression of QRS complexes using Hermite expansion*, IEEE Trans. on Sig. Proc., vol. 60, no. 2, pp. 947–955, 2012.

Rendszerek affin transzformáltja

Ötlet

- Transzlációs, és dilatációs paraméterek bevezetése:

$$\Phi_k^{a,\lambda}(x) := \Phi_k(\lambda x + a) \quad (x, a \in \mathbb{R}, \lambda > 0).$$

- $\left\{ \sqrt{\lambda} \Phi_k^{a,\lambda} : 0 \leq k \leq n \right\}$ is ONR.
- Legjobb közelítés:

$$S_n^{a,\lambda} f := \sum_{k=0}^n \langle f, \Phi_k^{a,\lambda} \rangle \Phi_k^{a,\lambda} \quad (n \in \mathbb{N}, a \in \mathbb{R}, \lambda > 0).$$

- Implementációs szempontok miatt $f^{a,\lambda}$ -t használjuk.

Rendszerek affin transzformáltja

Célfüggvény definíciója

- A Bessel egyenlőtlenség alapján a hibaformula:

$$D_n^2(a, \lambda) := \|f\|^2 - \sum_{k=0}^n |\langle f, \Phi_k^{a, \lambda} \rangle|^2.$$

- Használható továbbá a

$$F_n(a, \lambda) := \sum_{k=0}^n |\langle f, \Phi_k^{a, \lambda} \rangle|^2.$$

Problémák

- Létezik-e a hibafüggvényeknek minimum, maximum helye?
- Nem triviális: az értelmezési tartomány nem kompakt halmaz.

Optimalizációs módszerek

Cél

A legjobb transzlációs, dilatációs paraméterek megtalálása.

Módszerek

- Determinisztikus:
 - Gradiens módszer (parciális deriváltak szükségesek)
 - **Nelder–Mead (NM) szimplex alapú módszer**
- Nem determinisztikus:
 - Particle Swarm Optimization (PSO), raj alapú optimalizáció

Matching Pursuit algoritmus

Cél

- Egy szívütés – több függvényrendszer.
- Mohó algoritmus, minden körben a hibafüggvény minimumát keresi (a reziduum jelre nézve).

Előnyök

- Alkalmas a szegmensek detektálására.
- Adaptív reprezentáció.
- Pontosabb approximáció kevesebb együtthatóval.
- A dolgozatban használt együtthatók száma: 7, 6, 2 .

Matching Pursuit algoritmus

Figure: QRS paramétereinek optimalizációja.

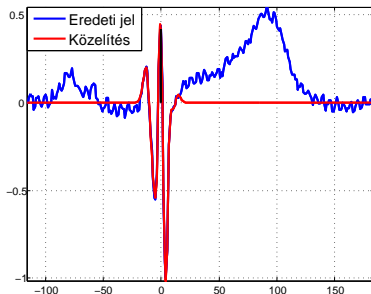
Matching Pursuit algoritmus

Figure: T paramétereinek optimalizációja.

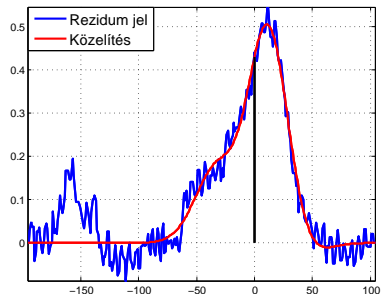
Matching Pursuit algoritmus

Figure: P paramétereinek optimalizációja.

Matching Pursuit algoritmus



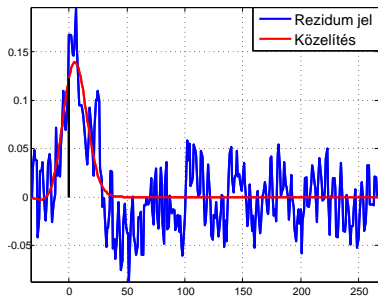
(a) A QRS approximációja.



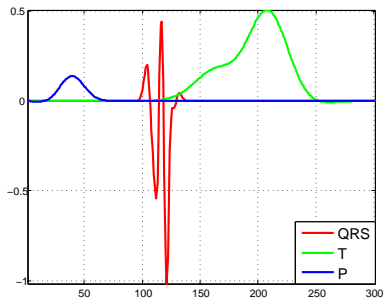
(b) A T hullám approximációja.

Figure: Az MP algoritmus lépései (1-2).

Matching Pursuit algoritmus



(a) A P hullám approximációja.



(b) Szeparált szívütés.

Figure: Az MP algoritmus lépései (3-4).

Megvalósítás

Felépítés

- c++ nyelven megírt jelfeldolgozó rész.
- Webes felhasználói felület.

A tömörítés implementációja

Modulok

- Objektum orientált felépítés.
- Jól definiált interfészek.
- Elkülönített szolgáltatások.
- Újrafelhasználhatóság, könnyen bővíthetőség.

A tömörítés implementációja

Alkalmazott technikák, érdekességek

- Összesen 9 c++ nyelven megírt modul.
- Algoritmusok: Nelder-Mead, MP, Hermite függvény gyökeinek megtalálása.
- c++ 11 adottságainak felhasználása: lambda függvények.

A hatékonyság szem előtt tartása

- Dinamikus memória használata (Hermite rendszerek).
- c++ sajátosságainak kihasználása, adatábrázolás során.
- A Standard Library szolgáltatásainak igénybevétele.

A NelderMead modul

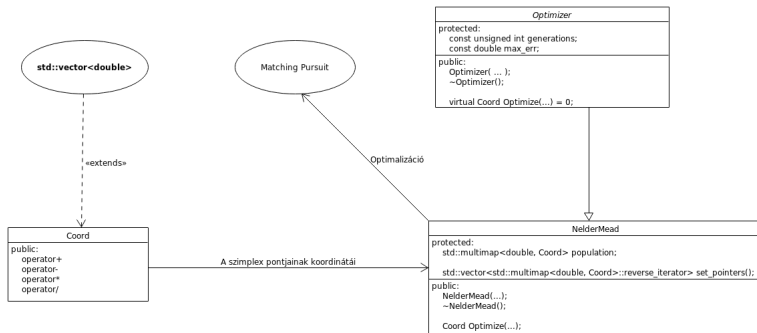


Figure: A NelderMead modul UML diagrammja

A felhasználói felület

Alkalmazott technikák

- Feladatspecifikus nyelvválasztás. Felhasznált programnyelvek: php, javascript, bash.
- Külső technológiák integrálása: GoogleCharts, jQuery.

Kényelem orientált felépítés

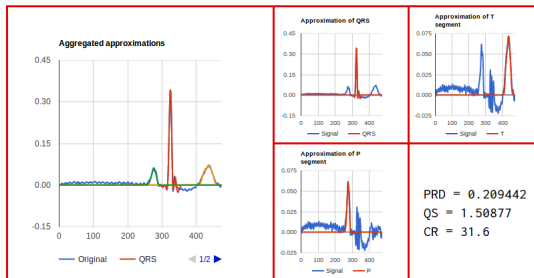
- Átlátható horizontális menüsor.
- Ismertető felület.
- Esztétikus megjelenés. Tervező grafikus által szerkesztett külalak.

A felhasználói felület

EKG JELEK TÖMÖRÍTÉSE
HERMITE FÜGGVÉNYEK SEGÍTSÉGÉVEL

ISMERTETŐ TÖMÖRÍTÉS KITÖMÖRÍTÉS ANIMÁCIÓ

Példa tömörítés:



Dóra Tamás Gábor
tamasi.dora@gmail.com
ELTE IK

Designed by Dóra Tamás
doras.tamas@gmail.com

Figure: Az animation.php

A fejlesztés jellemzése

A rendszer felépítése

- Könnyű telepítés, fordítás.
- Logikus könyvtár felépítés.
- Kód, modul tesztek, webes tartalom elszeparálása.
- Verziókezelő rendszer (git) alkalmazása.

A kód jellemzése

- Azonos kódolási stílus használata.
- Konstans értékek használata. "Varázsszámok" szűrése.
- Beszédes azonosítók használata. (pl. `t_inActionSig_ofs`).

A programcsomag tesztelése

Modul tesztek

- Minden modulhoz külön modulteszt.
- Eredmények, modulteszt kódja: `./src/mt/[modul neve]`

Tesztelés

Értékelés

- Percentage root mean square difference (PRD):

$$\text{PRD} = \frac{\|S_n^{a,\lambda} f - f\|_2}{\|f - \bar{f}\|_2} \times 100.$$

- Compression ration (CR):

$$\text{CR} = \frac{\text{eredeti EKG mérete}}{\text{tömörített EKG mérete}} \times 100.$$

- Quality score (QS):

$$\text{QS} = \frac{\text{CR}}{\text{PRD}}.$$

Tesztelés

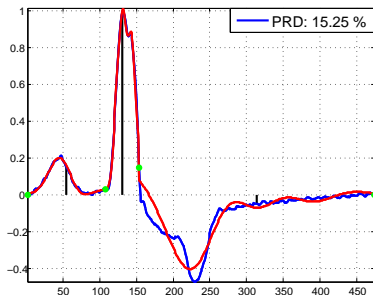
Technikai adatok

- PhysioNet EKG adatbázisból 6 egyenként fél órás rekordon.
- 12830 szívütés, rekordonként 1400-1600 másodperces futásidő.

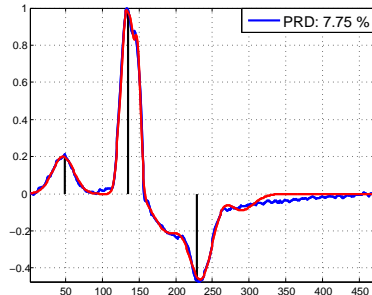
	Hiba (PRD %)				A tömörítés aránya (CR 1 : X)				Quality Score (CR : PRD)			
Rec.	Eredeti	NM	PSO	JPEG2	Eredeti	NM	PSO	JPEG2	Eredeti	NM	PSO	JPEG2
101	11.20	11.10	11.13	11.07	29.71	27.22	27.22	18.94	2.65	2.45	2.47	1.71
117	13.20	11.81	17.66	11.66	36.07	33.06	33.06	23.66	2.73	2.79	1.87	2.02
118	19.83	17.79	16.65	17.72	24.34	22.30	22.30	32.91	1.22	1.25	1.33	1.85
119	14.27	8.76	10.20	8.80	27.89	25.55	25.55	23.85	1.95	2.91	2.51	2.71
201	13.51	12.17	12.17	12.14	28.21	25.35	25.35	13.15	2.08	2.08	2.08	1.08
213	19.92	18.28	17.60	18.29	17.08	15.64	15.64	35.23	0.85	0.85	0.88	1.92
Átlag	14.22	12.55	13.83	12.51	26.83	24.45	24.51	23.86	1.91	2.06	1.85	1.88

Table: A tömörítés összehasonlítása különböző módszerek esetén.

Tesztelés



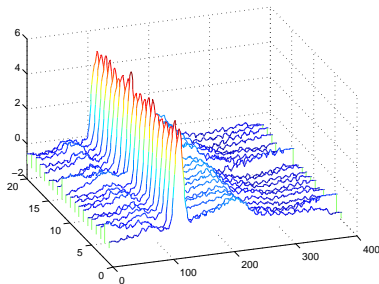
(a) Eredeti módszer.



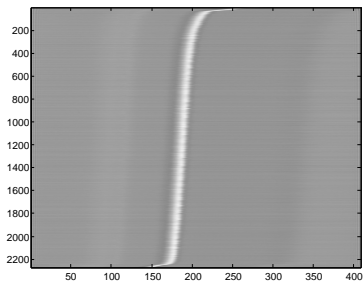
(b) Saját módszer.

Figure: Asszimmetrikus EKG jel közelítése (119-es rekord).

Más algoritmusokkal való összehasonlítás



(a) 180 szomszédos szívütés.



(b) Szívütések 2D képe.

Figure: EKG jelek feldolgozása képtömörítő algoritmusok segítségével.

Összefoglalás

Eredmények

- Új ONR szerkesztése translációval és dilatációval.
- Optimum létezésének belátása.
- Alkalmazás EKG jelek tömörítésére, szegmentálására.
- Tesztelés valós adatsorozatokon.
- Hatékony implementáció elkészítése, webes felhasználói felülettel.
- Összehasonlítás más módszerekkel: fix Hermite, JPEG2000.
- Módszer publikálása nemzetközi folyóiratban [5].

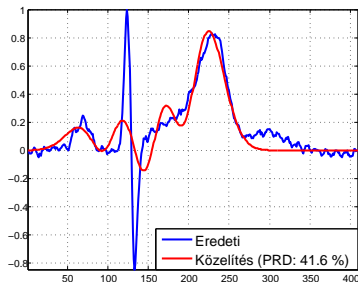
Konklúzó

- A módszer hatékonyan alkalmazható EKG jelek tömörítésére.
- Pontosabb közelítés érhető el asszimetrikus jelek esetén.
- Közel valósidejű feldolgozás lehetséges.

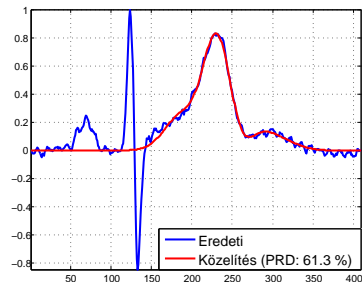
Irodalomjegyzék

-  B. Szokefalvi-Nagy, “Valós függvények és függvénysorok,” in *Polygon Könyvtár*, Szeged, HU, 2002.
-  G. Szegő, “Orthogonal polynomials,” *AMS Colloquium Publications*, New York, USA, 3rd edition, 1967.
-  W. Gautschi, “Orthogonal Polynomials, Computation and Approximation,” in *Numerical Mathematics and Scientific Computation*, Oxford University Press, Oxford, UK, 2004.
-  S. G. Mallat, Z. Zhang, “Matching pursuit in time-frequency dictionary,” in *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
-  T. Dozsa, P. Kovacs, „ECG Signal Compression Using Adaptive Hermite Functions,” *Advances in Intelligent Systems and Computing*, AISC, volume 399

Kitekintés



(a) Rossz eset.



(b) Jó eset.

Figure: A mohó stratégia hátránya.

Nelder–Mead szimplex módszer

Cél

A $D_n^2(a, \lambda)$ hibafüggvény minimumhelyének meghatározása.

Algoritmus

- Kizárólag az $f(x_3) \leq f(x_2) \leq f(x_1)$ értékekre támaszkodik.
- Keressük az $x' = x_1 + \alpha(x - x_1)$ ($\alpha \in \mathbb{R}$) pontot, melyre

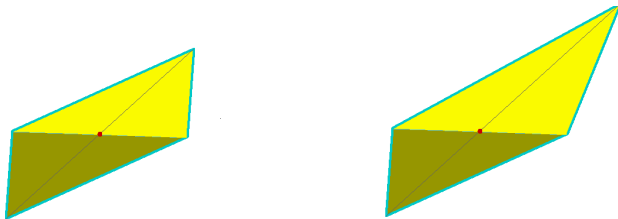
$$f(x') \leq f(x_2) \leq f(x_1),$$

ahol $x = (x_2 + x_3)/2$ a két kisebb érték között húzott szakasz felezőpontja.

Nelder–Mead szimplex módszer

Geometriai transzformációk

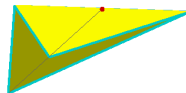
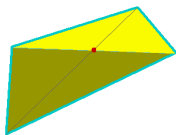
- $\alpha = 2$ esetén pont egy x -re vonatkozó középpontos tükrözés (T_1).
- $\alpha > 2$ esetén a tükrözés és nyújtás (T_2).
- $1 < \alpha < 2$ esetén tükrözés és zsugorítás (T_3).
- $-1 < \alpha < 0$ esetén egyszerű zsugorítás (T_4).



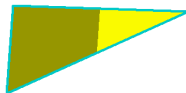
(a) Tükrözés ($T_1 : \alpha = 2$).

(b) T-Nyújtás ($T_2 : \alpha = 2.5$).

Nelder–Mead szimplex módszer



(c) T-Összehúzás ($T_3 : \alpha = 1.5$). (d) Összehúzás ($T_4 : -1 < \alpha < 0$).



(e) Kicsinyítés x_3 -ból (T_5).

Diagnosztikai torzulás jellemzése

Alternatívák

- WDD (2000): orvosdiagnosztikában használt jellemzők közvetlen átvétele ¹.
 - Hátrány: a felhasznált feature-k detektálása nagyon nehéz.
- WWPRD (2006): wavelet együtthatók súlyozott PRD-je ².
 - Előny: könnyen implementálható, csak QRS detektálás szükséges.

¹Y. Zigel, A. Cohen, and A. Katz, *The Weighted Diagnostic Distortion (WDD) Measure for ECG Signal Compression*, IEEE Trans. on Biomed. Eng., vol. 47, no. 11, pp. 1422–1430, 2000.

²A. S. Al-Fahoum, *Quality Assessment of ECG Compression Techniques Using a Wavelet-Based Diagnostic Measure*, IEEE Trans. on Inf. Tech. in Biomed., vol. 10, no. 1, pp. 182–191, 2006.

Diagnosztikai torzulás jellemzése

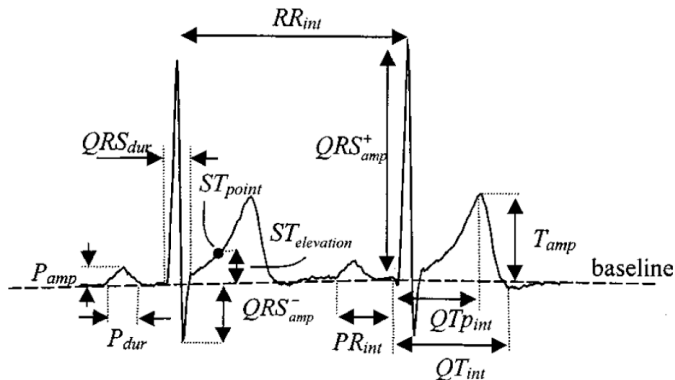


Figure: Egy szívütés diagnosztikai jellemzői (WDD).

Diagnosztikai torzulás jellemzése

Mérték	Minősítés				
	Tökéletes	Nagyon jó	Jó	Nem Rossz	Rossz
PRD	0-4.33	4.33-7.8	7.8-11.59	11.59-22.5	22.5 <
WWPRD_h	0-7.4	7.4-15.45	15.45-25.18	25.18-37.4	37.4 <

Table: Tömörítés diagnosztikai minősége PRD és WWPRD_h alapján.