

POUŽITÍ METODY ICA PŘI STANOVENÍ VÝZNAMNÝCH ZNAKŮ V SIGNÁLECH KARDIOVASKULÁRNÍHO SYSTÉMU KONÍ

IDENTIFICATION OF SIGNIFICANT FEATURES IN SIGNALS OF EQUINE CARDIOVASCULAR SYSTEM USING ICA

M.Kohut^{a)}, J. Holčík^{a)}, J. Hanák^{b)}, R. Kabeš^{b)}

^{a)} České vysoké učení technické v Praze, Fakulta biomedicínského inženýrství, náměstí Sítná 3105, 272 01, Kladno, Česká republika, kohut@ubmi.cvut.cz

^{b)} Veterinární a farmaceutická univerzita Brno, Klinika chorob koní, Palackého 1/3, 612 42, Brno, Česká republika

Abstrakt Článek popisuje využití metody analýzy nezávislých komponent (ICA) pro řešení problematiky náhlé kardiovaskulární smrti koní při operacích. Úvodní kapitola seznamuje čtenáře s danou problematikou a charakterizuje současný stav. Druhá část je věnována popisu vlastní metody použité k řešení a vstupních dat. Třetí a čtvrtá část demonstrovuje získané výsledky včetně grafů a shrnuje získané poznatky. Bylo zjištěno, že ne všechny složky obsažené ve vstupních datech jsou podstatné pro dostatečně vyjádření sledovaných procesů - nejčastějším případem (6 záznamů z 10) byla situace, kdy pro dostatečně přesný popis stačily pouze dvě nezávislé komponenty.

Summary The paper deals with application of the Independent Component Analysis (ICA) for looking for causes of the sudden cardiovascular death in horses during surgery. The first chapter describes the solved problem and state of art. Next part describes the ICA method and its properties and an input data format. In the third and fourth parts results of the analysis are presented. It was found that not all the three important components contained in the input data must be also part of the output data set. The most frequent case (6 records of 10) is that only two components were sufficient for description of input data.

1. ÚVOD

I přes veškerý pokrok v oblasti veterinární anesteziologie zůstává úmrtnost koní, která je v přímé souvislosti s celkovou anestézií, stále vysoká. Podle nejrozsáhlejší a nejkompaktnější studie s účastí 6255 koní [Johnston GM et al., 1995] je četnost náhlých a neočekávaných úmrtí 0,9% a ve 39% případech těchto úmrtí bylo příčinou selhání kardiovaskulárního systému.

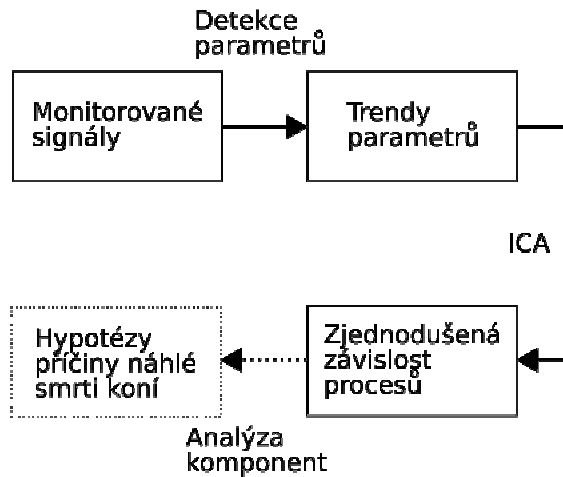
Dosavadní výzkumy na tuzemských (Klinika chorob koní VFUB) i zahraničních pracovištích ukazují, že monitorování běžných vitálních parametrů pomocí elektrokardiografie, pulzní oxymetrie, přímého měření krevního tlaku, analýzy koncentrace plynů v dýchacím systému či stavu acidobazické rovnováhy nemusí vždy odhalit blížící se selhání kardiovaskulárního systému při anestézii. Nicméně klinické zkušenosti ukazují, že existuje souvislost mezi stavem a chovem, resp. tréninkem koní a náchylností k náhlé smrti (paradoxně je zaznamenávána větší úmrtnost u dobře trénovaných koní). Abychom tuto hypotézu ověřili (příp. vyvrátili), je potřeba se mimo jiné zabývat analýzou měřených signálů, které primárně charakterizují dynamiku vývoje hodnot veličin popisujících stav koně při anestézii.

Určitým nedostatkem dosavadních výzkumů v dané oblasti je skutečnost, že až dosud vytvořené modely na našem pracovišti jsou výsledkem analýzy dat, která charakterizují činnost kardiovaskulárního

systému koní pouze v bdělém stavu (ať v klidu nebo po podráždění elektrickým impulsem). Zmíněný přístup nelze aplikovat během anestézie, protože neexistuje možnost cíleně ovlivňovat činnost a reakce kardiovaskulárního systému nezávisle na probíhající operaci. Proto je nutné využít postupů, které umožní kontinuální vyhodnocení průběhu jednotlivých veličin, příp. jejich vazeb. Jelikož reálná data získaná monitorování vitálních funkcí obsahují vyjma užitečných informací i parazitní a redundantní, je možností, jak odstranit méně užitečné složky a zvýraznit, resp. poodhalit ty více informativní a tím zjednodušit závislost dílčích procesů, metoda rozkladu na nezávislé komponenty (ICA).

2. POUŽITÉ METODY

Pro řešení problému máme v současnosti k dispozici signály popisující činnost kardiorespiračního systému 10 koní při anestézii (primárně, EKG, tlakové křivky, respirační křivky, saturace O₂), které byly získány při monitorování životních funkcí koní v průběhu operačních zákroků na Klinice chorob koní VFUB systémem Datex-Ohmeda S/5TM. Dané signály postihují složku řízení činnosti kardiovaskulárního systému, elektro-mechanickou činnost srdce i činnost dýchacího systému a předpokládáme tedy, že obsahují informace, které mohou objasnit příčinu náhlé smrti.



Obr.1. Blokové schéma řešení projektu (Block diagram of the solved task).

Analýza nezávislých komponent (ICA) [1][2] je statistická metoda, která je používána především při výběru příznaků popisujících analyzovaná data, příp. při hledání skrytých zdrojů dat.

Metoda ICA je definována vztahem

$$\mathbf{s} = \mathbf{W}\mathbf{x}, \quad (1)$$

kde $\mathbf{x} = [x_1, x_2, \dots, x_m]^T$ je n -rozměrný vektor zkoumaných dat a $\mathbf{s} = [s_1, s_2, \dots, s_m]^T$ je obraz vektoru \mathbf{x} . Transformační matice $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]^T$ představuje lineární transformaci zkoumaných dat \mathbf{x} na vektor \mathbf{s} . Cílem ICA je tedy nalézt koeficienty transformační matice \mathbf{W} tak, aby vypočtené nezávislé komponenty byly statisticky nezávislé. Tyto nezávislé komponenty představují ony zdroje signálů, které jsou skryté v množině zkoumaných dat \mathbf{x} . Pořadí a polarita nezávislých komponent nemusí odpovídat pořadí a polaritě skutečných zdrojů.

Vzájemná statistická nezávislost je pro náhodné veličiny s_1, s_2, \dots, s_m definována vztahem

$$p(s_1, s_2, \dots, s_m) = p_1(s_1) \cdot p_2(s_2) \cdot \dots \cdot p_m(s_m) \quad (2)$$

kde $p(s_1, s_2, \dots, s_m)$ je sdružená funkce rozložení hustoty pravděpodobnosti a $p_i(s_i)$ je marginální funkce rozložení hustoty pravděpodobnosti náhodné veličiny s_i .

Posouzení významnosti jednotlivých nezávislých komponent provádíme na základě odchylky Δ_k zpětně rekonstruovaných průběhů veličin od původních originálních dat při odstranění posuzované komponenty. Rekonstrukce je prováděna prostřednictvím inverzní matice \mathbf{W}^{-1} . Při použití všech nezávislých komponent pro rekonstrukci bude odchylka nulová a s odstraněním komponent, které jsou významné tato odchylka poroste.

Dále pro kontrolu provádíme také vizuální porovnání rekonstruovaných průběhů s originálními.

Odchylku Δ_k jsme vyjádřili pomocí vztahu

$$\Delta_k = \frac{\sqrt{\sum_{i=1}^m \sum_{j=1}^n (x_{ij} - x_{ij}^*)^2}}{m \cdot n} * 1000, \quad k=1,2,\dots,m. \quad (3)$$

kde x_{ij} jsou prvky matice zkoumaných dat \mathbf{X} , x_{ij}^* jsou prvky rekonstruované matice \mathbf{X}^* , m je počet monitorovaných veličin, n je počet vzorků a k je posuzovaná nezávislá komponenta.

Nalezení nezávislých komponent je prováděno prostřednictvím programu ICALAB, volně dostupného z [3]. Program nabízí pro výpočet více algoritmů, v našem případě byl použit algoritmus SOBI [4]. Tento byl zvolen na základě předchozích experimentů s daty, u kterých byly skryté zdroje předem známy.

3. VÝSLEDKY

Délka monitorovaných signálů odpovídá délce probíhající operace, řádově se jedná o hodiny. Tyto signály obsahují neodstranitelné artefakty a tak pro analýzu byly pro ověření analýzy vybírány minimálně 15 minutové konzistentní intervaly. Signály jsou navzorkovány s maximálním využitím kapacity monitorovacího systému s ohledem na jejich charakter, tzn. EKG signál 300Hz, signál tlakové křivky 100Hz, ostatní signály 25Hz.

Z jednotlivých monitorovaných signálů byly získány trendy 7 parametrů:

- interval mezi hodnotou systolického tlaku a vlnou R, v grafech označení x1;
- diastolický tlak (nejnižší hodnota tlakové křivky předcházející systolickému tlaku), v grafech označení x2;
- srdeční frekvence stanovená z průběhu EKG (na základě intervalu mezi sousedními vlnami R), v grafech označení x3;
- srdeční frekvence stanovená z tlakové křivky (na základě intervalu mezi sousedními systolami), v grafech označení x4;
- střední tlak (střední hodnota tlaku mezi dvěma systolami), v grafech označení x5;
- minimální tlak (nejnižší hodnota tlaku mezi dvěma systolami), v grafech označení x6;
- systolický tlak (nejvyšší hodnota tlaku v jedné srdeční periodě), v grafech označení x7.

Vzhledem ke skutečnosti, že jednotky všech parametrů nejsou ekvivalentní, bylo potřeba průběhy hodnot parametrů před použitím vlastní metody ICA normalizovat.

Zvolili jsme 2 způsoby:

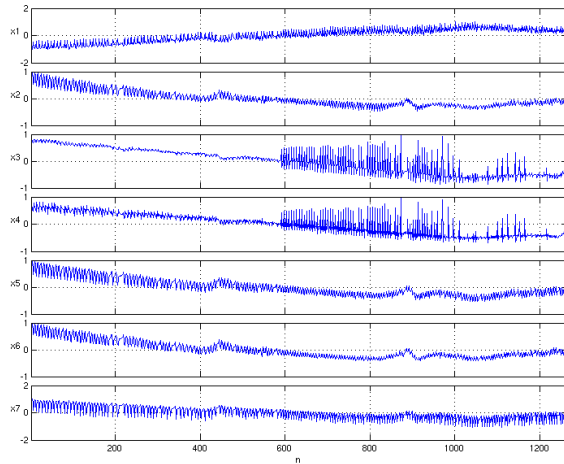
a) normalizace vzhledem ke směrodatné odchylce v jednotlivých průbězích (pro porovnání s výsledky metody analýzy hlavních komponent – PCA [5]);

b) normalizace vzhledem k maximální hodnotě v jednotlivých průbězích.

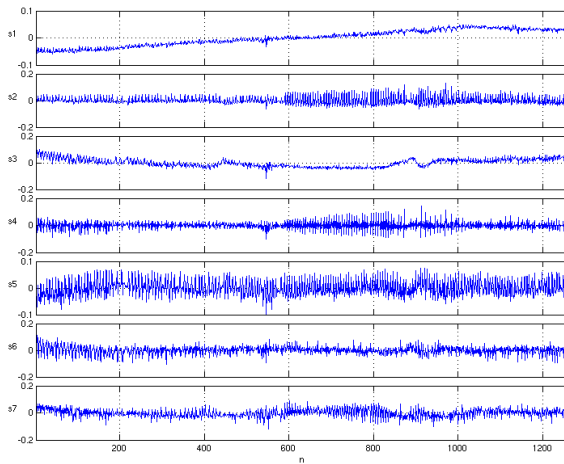
Konečně u všech průběhů byla odstraněna stejnosměrná složka.

Rozložením na nezávislé komponenty jsme zjistili, že způsoby normalizace neovlivňují výsledky z hlediska kvalitativního, ale spíše kvantitativního.

Následující obrázky (obr.2 - 7) demonstrují některé charakteristické případy.



Obr.2. Pacient A – průběhy vstupních parametrických signálů normalizované vůči maximální hodnotě s nulovou stejnosměrnou složkou (Patient A – input signals normalized to maximum value and with zero mean value).

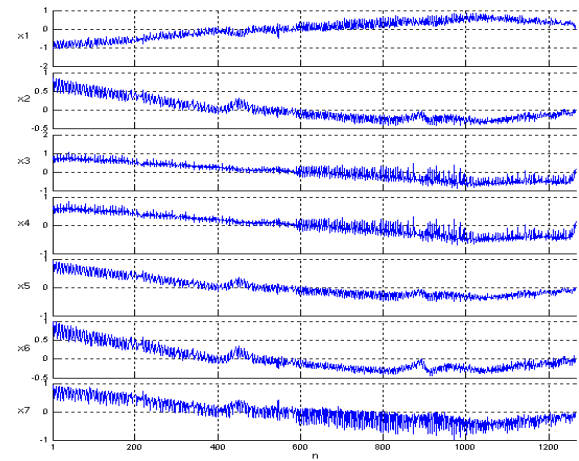


Obr.3. Pacient A – průběhy hodnot vypočtených nezávislých komponent (Patient A – computed independent components).

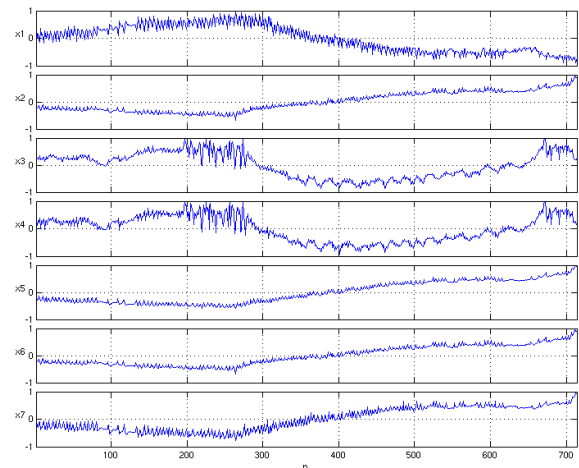
Vypočtením odchylky pro charakteristické případy podle vztahu (3) jsme získali údaje uvedené v tab.1 a tab.2, které kvantifikují významnost jednotlivých nezávislých komponent (obr.3, resp. obr.7) při zpětné rekonstrukci monitorovaných dat.

Tab. 1. Pacient A – odchylky rekonstruovaných průběhů od vstupních dat (Patient A - deviations of the reconstructed signals from input data).

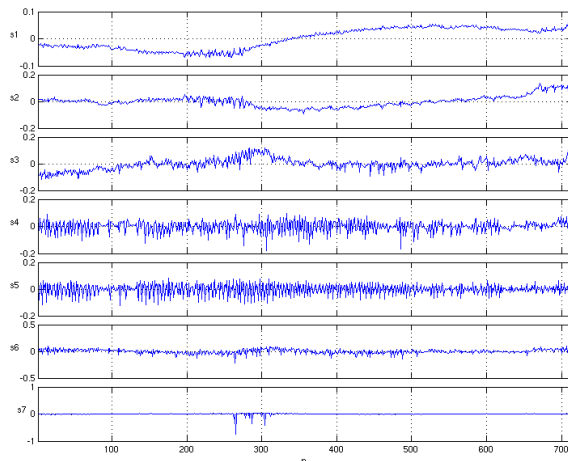
Δ_1	3,6124
Δ_2	1,0891
Δ_3	0,9865
Δ_4	0,8773
Δ_5	0,4356
Δ_6	0,4434
Δ_7	0,4830



Obr.4. Pacient A – rekonstruované průběhy parametrů s využitím prvních třech nezávislých komponent – s_1 , s_2 , s_3 (Patient A – signals reconstructed from three independent components – s_1 , s_2 , s_3).



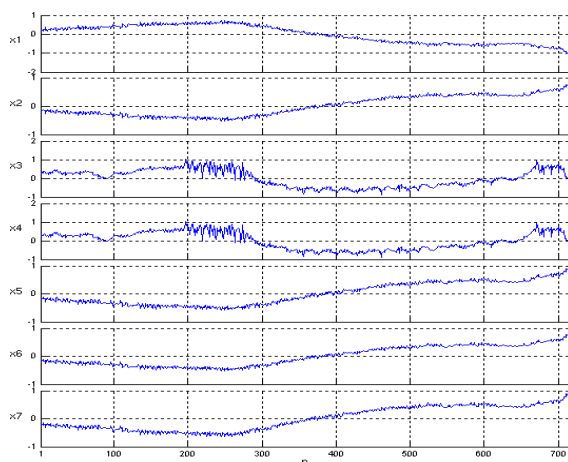
Obr.5. Pacient B – průběhy vstupních parametrických signálů normalizované vůči maximální hodnotě s nulovou stejnosměrnou složkou. (Patient B – input signals normalized to maximum value and with zero mean value)



Obr. 6. Pacient B – průběhy hodnot vypočtených nezávislých komponent (Patient A – computed independent components).

Tab. 2. Pacient B – odchylky rekonstruovaných průběhů od vstupních dat (Patient B - deviations of the reconstructed signals from input data).

Δ_1	5,1260
Δ_2	2,9207
Δ_3	0,6268
Δ_4	0,6035
Δ_5	0,6107
Δ_6	0,4550
Δ_7	0,0604



Obr. 7. Pacient B – rekonstruované průběhy parametrů s využitím prvních dvou nezávislých komponent (s1, s2) (Patient A – reconstructed signals from two independent components – s1, s2).

4. DISKUSE

Získané výsledky lze z obecnějšího hlediska kategorizovat do dvou, resp. tří skupin (třetí však obsahovala pouze jednoho pacienta).

Pacient A je zástupcem kategorie jedinců, u kterých jsou významné tři nezávislé komponenty (skryté zdroje). To vyplývá jednak z tab.1 a také z obr.4, kde rekonstruované průběhy poměrně dobře charakterizují originální data (obr.2).

Konkrétně v uvedeném příkladu je možné si všimnout v průbězích originálních parametrů (obr.2) jistých fenoménů. Jedná se o poměrně lineární pokles, resp. nárůst hodnot s časem u parametrů krevního tlaku (x2, x5, x6, x7) a srdeční frekvence (x3, x4), resp. parametru zpoždění systoly za vlnou R (x1). Tento jev vystihuje (obr.3) první nezávislá komponenta (s1). Jak bylo uvedeno, polarita komponent nemusí vždy odpovídat polaritě originálních signálů. Dalším rysem je výrazné rozkmitání hodnot u parametrů srdeční frekvence, které je vystiženo druhou nezávislou komponentou (s2). Třetím jevem je existence zvlnění ve dvou místech průběhu parametrů krevního tlaku, které je optimálně charakterizováno třetí nezávislou komponentou (s3).

Zmíněné tři skryté zdroje jsou tedy odpovědné za průběhy všech parametrů (srdeční frekvence, krevních tlaků a parametru zpoždění systoly za vlnou R). Lze pozorovat existenci jistých řídicích mechanismů v pozadí těchto tří fyziologických veličin. Čistě z hlediska hodnoty odchylky se mohou jevit významné i další komponenty (u pacienta A se jedná konkrétně o komponentu s4), ty však ovlivňují zejména rozptyl rekonstruovaných parametrů, nikoliv však morfologii průběhu.

U jedinců další kategorie (např. pacient B) se ukazuje významnost pouze dvou komponent, která je vidět z hodnot v tab.2 a z vizuálního porovnání originálních průběhů (obr.5) a rekonstruovaných (obr.7).

Z originálních průběhů (obr.5) lze pozorovat (až na polaritu) obdobný průběh parametrů krevního tlaku a zpoždění systoly za vlnou R. Tento znak vystihuje při zpětné rekonstrukci první nezávislá komponenta (obr.6). Vývoj parametrů srdeční frekvence při pozornějším prohlédnutí charakterizuje především druhá nezávislá komponenta navzdory skutečnosti, že se její průběh jeví podobný komponentě první.

Zde je tedy zaznamenána existence pouze dvou řídicích mechanismů zmíněných fyziologických veličin. Stejně jako u předchozí kategorie také zde další komponenty přispívají pouze k rozptylu rekonstruovaných průběhů.

Třetí kategorie je zastoupena pouze jedním pacientem a je charakteristická pouze jednou významnou nezávislou komponentou, která stojí v pozadí řízení krevního tlaku, srdeční frekvence a zpoždění systoly za vlnou R.

5. ZÁVĚR

ICA je obecně známou a používanou metodou pro analýzu biosignálů, nicméně pro řešení

problematiky náhlé kardiovaskulární smrti koní je podle dostupných informací používána poprvé.

Z analyzovaných 10 pacientů byli první kategorii příslušní 3 pacienti, druhé kategorii 6 pacientů a třetí kategorii 1 pacient.

Dříve užitá metoda PCA ukázala také existenci tří kategorií, nicméně ty nejsou plně ekvivalentní s výsledky metody ICA. Proto bude potřeba konfrontovat výsledky s klinickými záznamy jednotlivých pacientů a případně zavést další metody předzpracování či použít jiných algoritmů rozkladu.

Poděkování

Tento výzkum je podporován projektem GAČR 102/04/0887 „Metody a technické prostředky pro analýzu náhlé kardiovaskulární smrti koní“.

LITERATURA

- [1] Černošek, A. *Metody eliminace artefaktů v EEG záznamech*. Ostrava., Doktorská disertační práce. FEI VŠB - TU Ostrava. (2002).
- [2] Hyvärinen, A., Oja, E. *Independent Component Analysis: Algorithms and Applications*. Neural Networks Research Center, Helsinki University of Technology, (2000)
<http://www.bsp.brain.riken.jp/ICApub/NN00.pdf>
- [3] Cichocki, A., Amari, S., Siwek, K., Tanaka, T. a spol., ICALAB Toolboxes,
<http://www.bsp.brain.riken.jp/ICALAB>.
- [4] Belouchrani, A., Abed-Meraim, K., Cardoso, J.F., Moulines, E. *Second-order blind separation of temporally correlated sources*. Int. Conf. on Digital Signal Processing, Kypr, s. 346-351, (1993).
- [5] Kohut, M., Holčík, J., Hanák, J., Kabeš R. *Identification of Significant Features in Signals of Equine Cardiovascular System Using PCA*. YBERC05, St. Lesná, Slovensko, (2005).