SIGNAL DECOMPOSITION USING EMD TRANSFORMATION

Zdeněk Mžourek
Bachelor Programme (3), FEEC BUT
E-mail: xmzour01@stud.feec.vutbr.cz

Supervised by: Zdeněk Smékal
E-mail: smekal@feec.vutbr.cz

Abstract: Aim of this paper is to introduce empiric mode decomposition (EMD) as a new adaptive method for decomposing nonlinear and non-stationary signal into intrinsic mode functions (IMFs). This method is capable of decomposing any complicated data set into finite and often small number of IMFs [1] and because of its adaptive nature is highly efficient. With the Hilbert transformation of instrinic mode functions, we can obtain instantaneous frequencies as the functions of time that give sharp indentifications of imbedded structures.

Keywords: empiric mode decomposition, instantaneous frequency, intrinsic mode functions, time-frequency analysis

1 ÚVOD

Hilbert-Huangova transformace (HHT) je nová metoda pro analýzu nelineárních a nestacionárních signálů. Tato metoda se skládá ze dvou částí. První je empirická modální dekompozice (EMD), na kterou se zde hlavně zaměříme, a druhá je vytvoření tzv. Hilbertova spektra (distribuce energie v časově-frekvenční oblasti) za pomoci Hilbertovy transformace (viz [1]). EMD rozkládá analyzovaný signál na vlastní modální funkce (IMF), které jsou založeny na lokálních vlastnostech signálu. To umožnuje fyzikálně smysluplně definovat okamžitý kmitočet, a tedy odstraňuje nutnost použití často „nežádoucích“ harmonických funkcí k reprezentaci nelineálních a nestacionárních signálů.

2 HILBERT-HUANGOVA TRANSFORAMCE

2.1 Okamžitý kmitočet

V klasické Fourierově analýze je kmitočet definován jako funkce sinu nebo kosinu trvající po celou délku naměřeného souboru dat a také s konstantní amplitudou. Pro určení kmitočtu potřebujeme alespoň jednu celou oscilaci, takže s kratším úsekem dat nejsme kmitočet schopni určit. Tímto způsobem definovaný kmitočet nemá pro nestacionární data velký význam, a proto se zavádí tzv. okamžitý kmitočet.

Okamžitý kmitočet musíme zavést jednoznačně a to s využitím Hilbertovy transformace, s jejíž pomocí jsme schopni vytvořit analytickou funkci

$$z(t) = s(t) + \mathrm{j}s_i(t) = a(t)\mathrm{e}^{\mathrm{j}\varphi(t)}, \tag{1}$$

kde $s(t)$ je původní signál, $s_i(t)$ je imaginární funkce získaná pomocí Hilbertovy transformace signálu $s(t)$, $a(t)$ je okamžitá amplituda a $\varphi(t)$ je okamžitá fáze.

Okamžitý kmitočet pak můžeme definovat jako

$$\omega(t) = \frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}. \tag{2}$$

Tato definice představuje funkci jedné proměnné závislé na čase t, a proto v libovolném čase existuje právě jedna hodnota kmitočtu. Kdybychom chtěli získat okamžitý kmitočet z obecného signálu $s(t)$ (tzn. neaplikujeme na něj žádné restriktivní podmínky), tak se může stát, že okamžitý kmitočet $\omega(t)$ bude dosahovat i záporných hodnot. Záporné hodnoty kmitočtu nemají žádný fyzikální význam, a proto jsou k analýze signálu $s(t)$ nepoužitelné.

Aby výsledná okamžitá frekvence $\omega(t)$ měla fyzikální opodstatnění, tak musíme zavést jisté podmínky, které signál $s(t)$ musí splňovat. Záporné hodnoty okamžitého kmitočtu $\omega(t)$ mohou nastat i v případě, že reálná část Fourierovy transformace signálu $s(t)$ má pouze kladné kmitočty, takže podle [1] se doporučuje použít jiný požadavek, a to konkrétně aby signál $s(t)$ byl lokálně symetrický vůči nulové střední hodnotě. To by mělo zaručit, že výsledný okamžitý kmitočet bude mít fyzikální interpretaci (signál bude reprezentovat pouze jednu komponentu, tzv. monokomponentu). Více se lze dočíst v [1].

2.2 Vlastní modální funkce

Vlastní modální funkce (IMF) má obsahovat fyzikálně smysluplný okamžitý kmitočet, a proto musí splňovat následují dvě podmínky:

1. symetrie obálek tvořených lokálními maximy a minimy vůči nulové střední hodnotě (tzn. vůči „časové“ ose),
2. a počet lokálních extrémů musí odpovídat počtu průsečíků s časovou osou.

V ideálním případě by první podmínka měla znít stejně, jak bylo popsáno výše, ale to pro nestacionární data není možné (viz [1]). Z tohoto důvodu se pro určení střední hodnoty použijí obálky tvořené lokálními extrémy (viz obr. 1).

O průběhu IMF můžeme říct, že se v podstatě jedná o amplitudově či kmitočtově (případně obojí) modulovaný signál. Bohužel většina analyzovaných dat neodpovídá definici IMF, a proto potřebujeme způsob, jak tyto data na IMF rozložit.

2.3 Empirická modální dekompozice

Empirická modální dekompozice (EMD) je metoda, která slouží k rozkladu libovolného nelineárního a nestacionárního signálu do konečného počtu IMF. Od svého představení v roce 1998 vzniklo mnoho různých modifikací a optimalizací této metody (viz například [2], [3]). EMD má v dnešní době spoustu využití, a to například odstranění šumu v signálu, separace různých audio zdrojů pomocí prostorové lokalizace, analýza obrazu a spousta dalších.

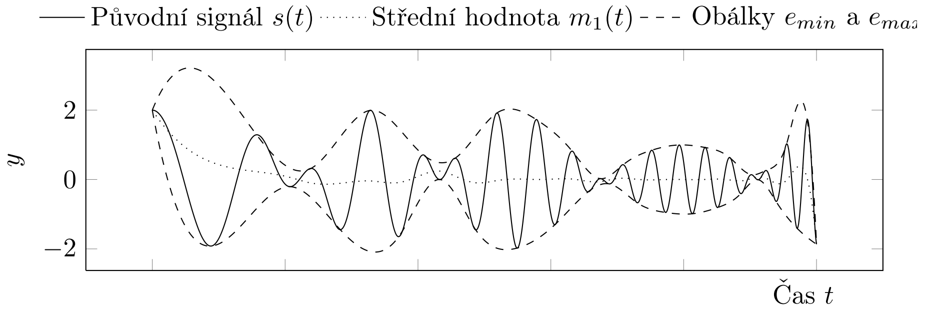
Hlavním neduhem této metody je její definice pomocí algoritmu namísto analytické formulace, která by dovolila teoretickou analýzu, a proto je potřeba k porozumění jejího chování v různých podmínkách podniknout řadu experimentů (viz [2]).

Postup by se dal shrnout v následujících krocích. Nejdříve musíme identifikovat všechny lokální extrémy signálu $s(t)$, poté získáme obálky $e_{min}(t)$ (resp. $e_{max}(t)$) pomocí interpolace lokálních minim (resp. maxim). Vypočítáme střední hodnotu $m(t) = (e_{min}(t) + e_{max}(t))/2$, získáme složku $h(t) = s(t) - m(t)$ a pokračujeme s iterací na složce $h(t)$.

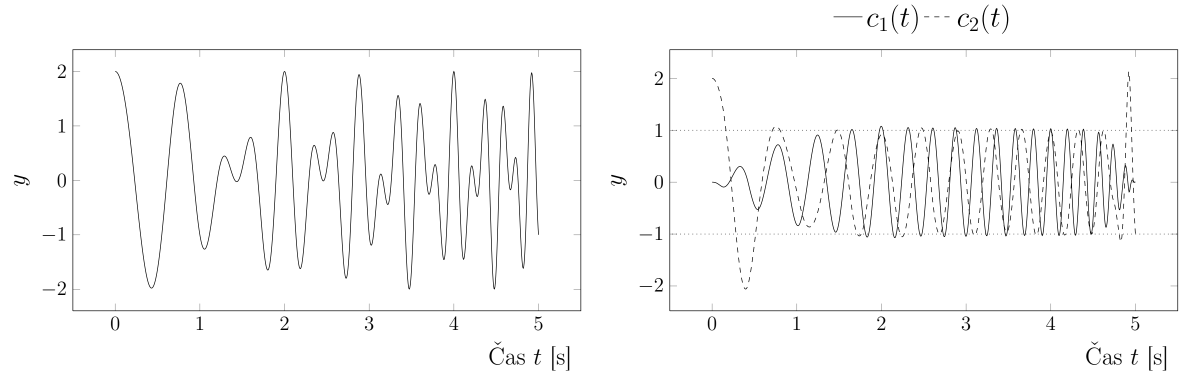
Iterace se zastaví, jakmile složka $h(t)$ bude splňovat podmínky kladené na IMF. Existují různá kritéria podle kterých lze určit kdy iteraci ukončit, aby nedošlo k tzv. přeiterování, při kterém dochází ke ztrátě informací. Například se používá velikost směrodatné odchylky [1] nebo ohodnocení fluktuace střední hodnoty [2].

K interpolaci se používá kubický spline, protože se ukázal jako nejvhodnější [2]. Pomocí úprav jeho definice můžeme dosáhnout lepších výsledků dekompozice [3].

Efektivnost EMD je dobře vidět na rozložení signálu $s(t)$ (obr. 2a), který je tvořen součtem dvou sinusových průběhů s konstatní amplitudou a lineárně vzrůstajícím kmitočtem. Výsledkem dekompozice jsou dvě IMF složky $c_1(t)$ a $c_2(t)$ (obr. 2b), které svým průběhem odpovídají původním sinusovým průběhům. Odchylky na okrajích jsou způsobeny konečnou délkou signálu $s(t)$.



Obrázek 1: Ukázka rozkladu signálu $s(t)$ pomocí EMD.



(a) Analyzovaný signál $s(t)$

(b) Výsledné IMF složky

Obrázek 2: Příklad rozložení $s(t)$ na jednotlivé IMF.

3 ZÁVĚR

EMD je primárně určená pro rozklad nestacionárních a nelineárních signálů. Hilbertovo spektrum v porovnání s Fourierovou nebo vlnkovou transformací dosahuje mnohem lepších výsledků (příklady lze najít v [1]). Pro dosažení optimálních výsledků je potřeba zvolit optimální kritérium pro zastavení iterace, případně upravit samotný algoritmus (viz [2], [3]). EMD má různé neduhy (problémy s identifikací lokálních extrémů a interpolací, přeiterování, a další), ale při správném použití se dá jejich vliv výrazně minimalizovat.

REFERENCE

[1] HUANG, N. E. et al. *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*. *Proceedings of the Royal Society A Mathematical Physical and Engineering Sciences*. 1998, 454, 1971, s. 903–995. ISSN 13645021.

[2] RILLING, G. et al. On empirical mode decomposition and its algorithms. In *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP*, 3, s. 8–11, 2003.

[3] SI-SHENG, L.; ZHANG, T.; FENG, L. *Causes and solutions of overshoot and undershoot and end swing in Hilbert-Huang transform*. *Acta Seismologica Sinica*. 2005, 18, 5, s. 602–610. ISSN 1000-9116.