**W1.** **Feature extraction from structured data.**

**2. Feature extraction from continuous signals (Month 1-14)**

**Ivan Michieli**: Extraction of features in the form of singular spectrum and modified principal components (Month 7-10) – završeno u maju 2015.

Metoda rastava višedimenzionalnih podataka nekog dinamičkog sustava na glavne funkcionalne komponente (principal component analyses - PCA) se direktno nastavlja na metode ugradnje vremenskih serija u n-dimenzionalni fazni prostor. Rastavom na glavne komponente (svojstvene vrijednosti kovarijantne matrice podataka) podaci se dekoreliraju i prikazuju u projekciji u kojoj su dinamičke karakteristike najviše izražene (maksimalna varijanca podataka)! Na taj je način moguće smanjiti dimenzionalnost matrice podataka zanemarujući projekcije s malom ukupnom varijancom ispod nekog zadanog praga! Kod analize vremenske serije nekog signala (jednodimenzionalni set podataka) serija se ugrađuje u pseudo fazni prostor tj. formira se matrica pseudo vektora odabrane dimenzionalnosti i onda se izlučuju glavne komponente takvog sustava. U ovisnosti o načinu ugradnje u fazni prostor, odabirom dimenzionalnosti i različitih parametara ugradnje kao i izborom težinske funkcije moguće je dobiti različite setove modificiranih glavnih komponenti s odabranim istaknutim svojstvima! Može se pokazati da ako se dimenzija pseudo-vektora povećava PCA metoda teži diskretnoj Fourrier-ovoj analizi vremenske serije (Gibson et.al. 1992., Vautard and Ghil 1989)! Kod velikog broja podataka signala važno je ograničiti dimenzionalnost analize na računalno prihvatljivu vrijednost! Dostupni algoritmi za proračun glavnih osi i svojstvenih vrijednosti kvadratnih simetričnih kovarijantnih matrica mogu biti ograničavajući faktor analize! Najbolji pristup ovom problemu je preko bitno efikasnijih algoritama za rastav na singularne vrijednosti (Singular value decomposition – SVD)! U linearnoj algebri SVD je jedna od metoda faktorizacije realne (ili kompleksne) **M** (*m* × *n* ) matrice podataka na način da **M** = **UΣV\***, gdje je **U** realna *m* × *m* unitarna matrica čije kolone su ortonormalni vektori (tzv. lijevi singularni vektori), **Σ** je *m* × *n* kvadratna dijagonalna matrica singularnih vrijednosti i **V\*** je *n* × *n* transponirana unitarna matrica čije kolone su tzv. desni singularni vektori! Uz određenu normalizaciju vremenske serije može se pokazati da su desni singularni vektori jednaki glavnim osima a singularne vrijednosti jednake glavnim komponentama PCA metode! Za primjenu metode na odabrane baze podataka (PhysioBank „gait database“ i „interbeat (RR) interval database”) definirane su procedure u programskom paketu “Mathematica” preko ugrađenih funkcija za SVD analizu. Uspješno su testirani proračuni matrica podataka s ugrađenim pseudo-vektorima dimezije od nekoliko tisuća točaka signala!