REMOTE PHOTOPLETHYSMOGRAPHY USING NONLINEAR MODE DECOMPOSITION

Halil Demirezen, Çiğdem Eroğlu Erdem
Department of Computer Engineering, Marmara University
halil@demirezen.net, cigdem.erdem@marmara.edu.tr

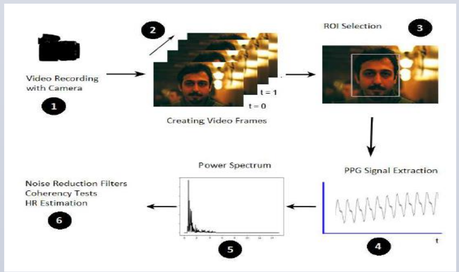
Goal: Improve HeartRate detection success by applying a recently proposed signal decomposition method (NMD) on PPG signal which is extracted from human face region.

Problem:

- Due to being weak in power, PPG signals are affected by noise.
- The measured PPG signal includes not only blood volume pulse but also respiratory wave, illimunation change, rigid and non-rigid head motion. Thus, it needs to be seperated.

Proposed Solution:

- We have applied Nonlinear Mode Decomposition (NMD) method to decompose the PPG signal into nonlinear modes to seperate blood volume pulse.
- PureDL dataset has been used.
- The face regions, thus, ROI (Region-of-Interest) have been extraced by Viola-Jones face detector.



NONLINEAR MODE DECOMPOSITION

- A recently proposed signal decomposition method, which seperates a signal into its physically meaningful oscillations, while removing noise.

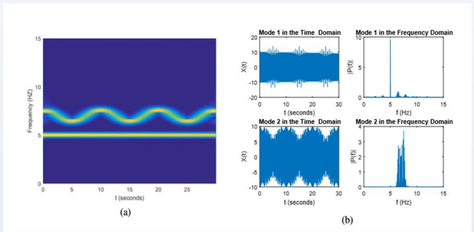
$$s(t) = \sum_i c_i(t) + \eta(t)$$

$$c(t) = A(t)v(\phi(t)) = A(t)\sum_h a_h \cos(h(\phi(t) + \psi_h)$$

- Steps:
  - Extract the fundamental harmonic of an NM from the TFR representation
  - Find all possible harmonics of the fundemental harmonic
  - Select the true harmonics of the same NM
  - Build the nonlinear mode from fundamental and true harmonics, substract it from the signal and repeat
  - Stop when the residual signal meets noise criteria
- Features
  - NMD is extremely noise-robust
  - The parameters of the algorithm are adaptively chosen
  - Returns physically meaningful even though non-sinusodial waveform.
- Example Signal:

$s(t) = (10 - 0.03t)cos\phi_1(t) + 10cos\phi_2(t) + 0.5\eta(t),$

$$\phi_1(t) = 10\pi t,$$
$$\phi_2(t) = 14\pi t + 2\pi \sin(\frac{2\pi t}{10})$$



(a) TFR Representation of signal s(t),
(b) b) recovered modes of s(t)

ALGORITHM

- The Viola-Jones face detector extracts the location and the size of face region
- The ROI is constructed as the %80 in width and %100 in height of the Viola-Jones face box.
- The R, G, B color components are normalized with respect to intensity:

$$r(t) = \frac{R(t)}{R(t)+G(t)+B(t)}, g(t) = \frac{G(t)}{R(t)+G(t)+B(t)},$$
$$b(t) = \frac{B(t)}{R(t)+G(t)+B(t)}$$

- 30 second signal is normalized $\tilde{G} = \frac{g(t)-\mu_g}{\sigma_g}$
- In NMD method, only the G signal is used, while in ICA method, all three signals are used.
- The mode is selected based on the criteria:

$$i_* = argmax_i[\max(|C_i(f)|) - median(|C_i(f)|)]$$

where Ci is the power spectrum of mode *i*

- The frequency with the highest power in the selected mode is the HR frequency estimate
- $HR_{bpm} = HR_f * 60$
- $|HR_{bpm}(t) - HR_{bpm}(t-1)| < 12$, the frequency with next highest frequency is selected.

DATASET

PureDL dataset contains face videos of 8 subjects under 6 different experimental setups:

- **Steady:** the subjects look at the camera from a frontal view and avoid any head motion.
- **Talking:** the subjects are requested to talk without any additional head motion.
- **Slow translation:** the subjects move their heads parallel to the camera plane with a speed of about 7% of the head height in pixels.
- **Fast translation:** the subjects move their heads parallel to the camera plane, with the speed doubled.
- **Slow rotation:** the subjects look at the targets sequentially, which are placed with 35 cm intervals around the camera.
- **Fast rotation:** the subjects look at the targets sequentially, which are placed with 70 cm intervals around the camera.

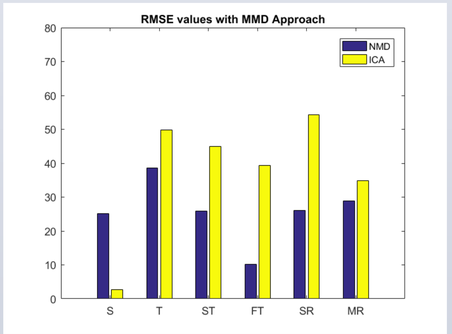
- Each video is 800x600 in resolution and 30 fps and 60 seconds in average.

RESULTS

SETUP	NMD-HR		ICA [3]	
	$\|\mu\|$	σ	$\|\mu\|$	σ
Steady	9.59	23.27	2.71	10.20
Talking	17.41	34.41	40.10	34.05
Slow Translation	5.71	25.22	25.85	36.66
Fast Translation	2.34	9.90	26.76	32.38
Slow Rotation	5.50	25.58	36.83	39.95
Medium Rotation	11.55	26.46	12.10	32.26

Experimental results of the proposed NMD-HR method and the ICA based method. The absolute mean error and the standard deviation of the estimation errors are given in beats per minute.

- The «Talking» setup has the hight average mean error 17.41, 40.10 respectively.
- The NMD signal decomposition method performs better than ICA method in most of the cases.



Experimental results of the proposed NMD-HR method and the ICA based method in terms of RMSE.