Topic Description: Measuring vibrations from video feeds

Presented by the MoD

Background

1. Having the ability to read off vibrations (including heart rate and speech) from a video feed would find applications in communications' enhancement, healthcare, security and seismology. There is previous work in this area, notably MIT's visual microphone (see http://people.csail.mit.edu/mrub/VisualMic/), and an analysis of recovering audio information from laser speckle patterns through imaging (see Z. Zalevsky et al,https://www.osapublishing.org/oe/abstract.cfm?uri=oe-17-24-21566). Although Zalevsky's paper does require laser illumination, it contains a discussion of how images can be converted into spectral patterns.

Topic Description

2. It is currently an open question to what extent audio information can be encoded in video feeds without further active techniques, such as laser or radio frequency illumination. There will be a number of parameters that recovery of audio vibration will depend on. These will include: the frame rate of the camera/camera sensor refresh rate, the number of pixels in the camera sensor, the frequency sensitivity and range of the camera sensor, atmospheric attenuation of light by environment and absorption of sound by the environment.

3. The questions which the ESGI are asked to explore are:
   - What parameters can be used to describe the recovery of audio vibration from video?
   - What would a quantitative model of audio encoding in video look like?
   - What be indicative ranges in which audio could be recovered from video?
   - What are the limits to the fidelity of the audio information that could be recovered?

Data

4. MIT have produced sample video of their visual microphone. It is hoped that video footage can be provided, but a good starting point would be video footage of a mobile phone vibrating while playing a constant sound tone.

Techniques

5. Fourier analysis could be used to extract frequency information from the time evolution of the camera sensor responses (distributed spatially by pixel).
6. Statistical techniques could be used to identify the audio vibration as deviation from the average background (if locally static).