The Visual Microphone: Development and Limitations

Ruairi L. Fraser and Matthew Edgar*
2134394f@student.gla.ac.uk

ABSTRACT

Recent developments in affordable camera technology have brought widespread improvements to specifications including resolution, frame rate and shutter speed. Using a modern high-speed video camera, in conjunction with image processing algorithms, audio information may now be recovered by examining minute changes in intensity on an objects surface. However, the process is limited by object distance, source audio volume, and scene illumination (active or passive).

PROTOTYPE

The prototype system detailed in Fig 1. contains a Mikrotron MC1362 camera fitted with a 80-400mm telephoto lens, a ~671nm laser, a red lens filter to be fitted when lasers in use, and a white LED torch. The camera operates at a maximum frame rate of 500Hz @ 1280x1024pixels, with the possibility to run at over 100kHz with decreased region of interest. The laser has a variable output power of up to 300mW. A beam expander and diffuser was constructed at the output of the laser in order to illuminate a suitable area. A Bluetooth speaker, with maximum output of 98.2dB, was used to provide wireless audio in the vicinity of the object over a range of distances.

FIGURE 1

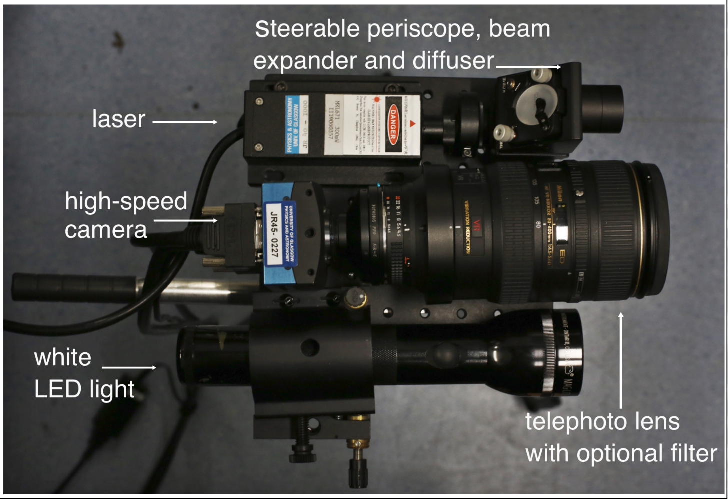


FIGURE 2

RANGE LIMITS

The visual microphone faces significant challenge when recovering under increasing distance. Over greater ranges, the target object will fill less pixels in the high-speed footage, making smaller displacements (RMS ~0.5μm) more difficult to detect. Additionally any noise due to the vibration of the camera itself will be amplified with respect to the target. Illustrated in Fig. 2, the recovery of an identical linear chirp signal is compared at a range of 5, 10 and 15m, under purely ambient lighting conditions. After highpass filtering, the signals are all clearly present, though are seen to decrease in intensity.

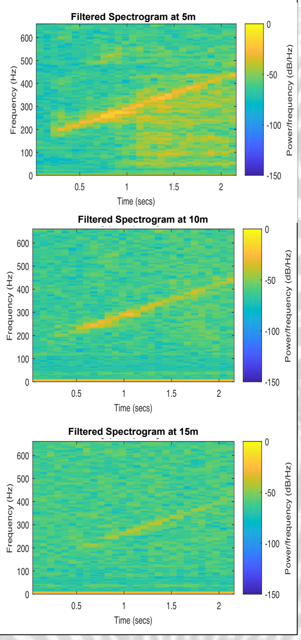
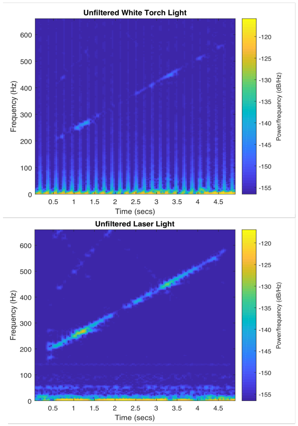


FIGURE 3



LASER IMPLEMENTATION

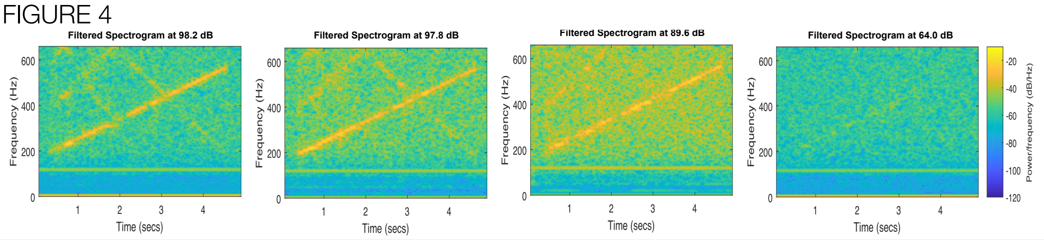
Development of the visual microphone led to the implementation of active laser lighting, in conjunction with the band pass filtering of the camera input, in an attempt to examine only the intensity changes of incident laser light, in turn decreasing the effects of noise such as the flickering of background lights such as fluorescent tubes. In reference to an identical chirp signal Fig 3. compares the performance of the system under white LED light from a DC source, to that with laser illumination and filtered camera input. The improvement is apparent, with clearer recovery and decreased presence of noise like anomalies, before the data has been put through any denoising or filtering algorithms.

SOURCE VOLUME

As recovery of the audio is taken from the physical movements excited by the input audio, the ease of recovery is strongly linked to the volume of the input in the vicinity of the object. Fig 4. shows the deterioration of recovery under laser light, as volume is adjusted in our signal generator, as measured by a decibel meter. The information recovered is limited, despite the highpass filtering and active laser illumination.

All results presented were recovered from footage of a Tunnock's tea cake wrapper. This was selected as the target object most conducive to clear audio recovery. Chirp signals were used as a visually comparable data set.

FIGURE 4



Relevant Publications:
1. Davis A, Rubinstein M, Wadhwa N, Mysore G J, Durand F and Freeman W T 2014
2. Ahn J, Kim Y J and Kim D 2016 Pattern Recognition (ICPR), 2016 23rd International Conference on (IEEE) pp 3927–3932
3. The visual microphone: Passive recovery of sound from video http://people.csail.mit.edu/mrub/VisualMic/ accessed: 2018- 03-08
4. Simoncelli E P and Freeman W T 1995 Image Processing, 1995. Proceedings., International Conference on vol 3 (IEEE) pp 444–447

•SUPA, University of Glasgow, Optics Group, School of Physics and Astronomy, Glasgow, UK, G12 8QQ
Background Image- Richard Maxted,20/10/15, *Horton Stephens Introducing…Richard Maxted*, viewed 18/3/18 https://horton-stephens.com/site/assets/files/1901/maxted_folio26.jpg/