Listen to the singing paper!

Recovering sound from inaudible videos

Anders Harbøll Christiansen and Emil Gensby, Bachelor students of Math and Technology
Technical University of Denmark

Our project was inspired by another project at MIT found online (A. Davis et al. , The Visual Microphone: Passive Recovery of Sound from Video, 2014). We recorded high-speed video of a piece of paper exposed to a song. The video itself contained no sound, but from the image data alone we recovered a signal, which when played resembled the original sound.

1. Are you the paranoid type?

Sound is something that we can 'hear'. Our eardrums detect the minute vibrations in air pressure and interpret those as sound. However, those vibrations also affect our surroundings, not just our ears. This means that everyday objects reveal the sound of their surroundings.

Normally we don't notice everyday objects vibrating, however with a high-speed camera it is possible to capture these vibrations. This is even possible through soundproof glass. Our aim, however, is simply to recover the song that we play and recognize it through a song-recognizing app.

2. Optical Flow

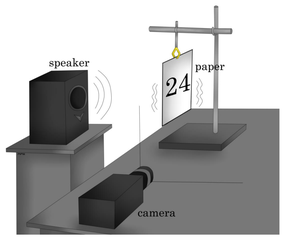
We use the high-speed camera to record the video. This is in a highly controlled environment, as the experiment is quite sensitive. To estimate the displacement of the paper from the video we need solve the aperture problem of optical flow. The aperture problem describes the movement of an object in a series of images, $I(x, y, t)$, by the use of the two directional derivatives and the time-directional derivative, as:

$$I_x V_x + I_y V_y = -I_t$$

where I_x, I_y and I_t are the three derivatives and V_x and V_y are matrices containing the motion of every pixel. In our case however, we can assume that the motion of the image is singular, that is we can replace V_x and V_y by Δx and Δy respectively. The estimated motion gives us a signal that we clean to eliminate noise.

The Power of Love ♫ ♪

Recording setup



Recording

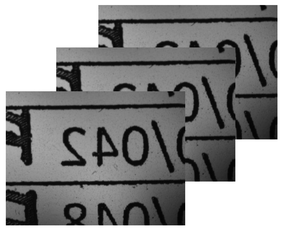
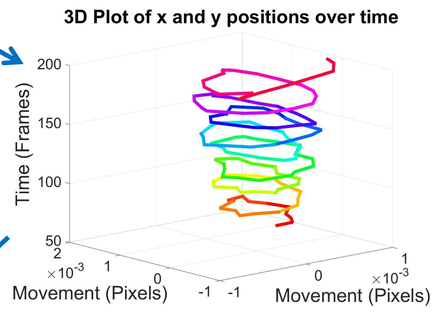


Image data

Movement Estimation



Recovered signal

Noise Reduction

Hear for yourself!

The Power of Love ♫ ♪

3. The Result

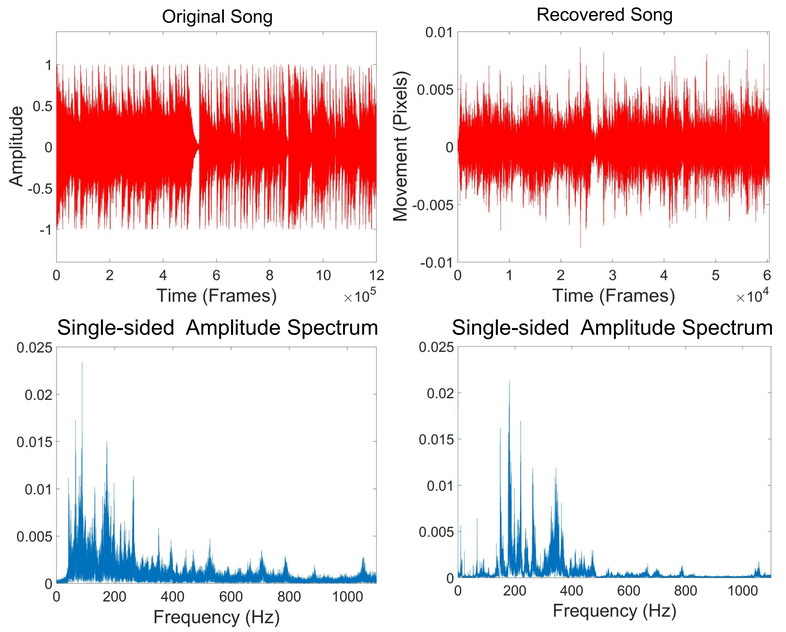
The song Power of Love performed by Huey Lewis & The News was played.

We successfully recovered the song! The recovered song was correctly recognized by the song-recognizing app, Shazam. Qualitatively, it sounds like listening to the song through a wall.

4. Discussion

We compare the original signal and the recovered signal both in the time-domain and the frequency domain, shown below.

There is a loss of some frequencies, mostly higher frequencies, but also some of the lower frequencies. We attribute this to the complex interaction between sound and material.



Comparison between original and recovered signal

Acknowledgements

We would like to thank a number of people without whom this project would not have gotten off the ground. First, we would like to thank A. B. Dahl, V. A. Dahl, C. H. Trinderup and M. J. Emerson for being our supervisors every Friday. We would also like to thank J. Brunskog from DTU Electro for information about sound and the setup, L. T. Nielsen and T. Kiørboe from DTU Aqua for helping us handle the high-speed camera as well as O. Christensen from DTU Compute for his expertise on signal filtering. Finally, thanks to M. M. Holst for providing hand drawn images used in this poster. Thank you to all of you for reading as well!