Audio Camera: Real Time Audio Imaging and Analysis

Adam E. O'Donovan (adamod@gmail.com) Ramani Duraiswami (ramani@umiacs.umd.edu)
Perceptual Interfaces and Reality Lab, University of Maryland

What Is The Audio Camera:

- The audio camera is the audible frequency range equivalent of a video camera.
- Generates an acoustic intensity image of the environment that is overlaid upon a real video stream.
- Utilizes GPUs to achieve real-time operation at 30fps.

Hardware:

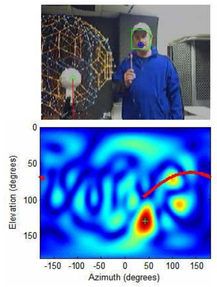
- 64 microphones embedded in the surface of a sound-hard 6 inch sphere.
- Sampling rate of 44100 samples per second per channel transmitted over USB 2.0 to host CPU.
- 640x480 digital video camera embedded in the sphere.
- Video streams at 15fps via USB 2.0.
- Acoustic images generated on an NVIDIA GPU.

Challenges:

- Unlike light, a single point source of sound influences all sensors in the array.
- Generation of a single audio pixel requires a weighted sum of all channels.
- Very high computational load

Vision Guided Beamforming:

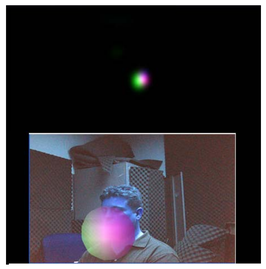
- Visual framework for audio analysis allows us to exploit many existing computer vision algorithms.
- Multiple acoustic/video cameras can be calibrated using standard 8 point vision algorithms.
- We have shown that the audio camera is a central projection camera for acoustic waves which implies an epipolar geometry between video and audio cameras.



- In the figure to the left the bottom panel shows the audio image collected by the audio camera and the top image shows a video stream collected by a separate video camera.
- A face detector processed the video stream and returned likely locations of human speakers in the scene.
- The epipolar geometry then implies a 1-D search space in the acoustic image (shown as the red curved line).
- Efficient search along this line then reveals the active speaker in the acoustic image despite the loud secondary sources.

Acoustic Image Overlay:

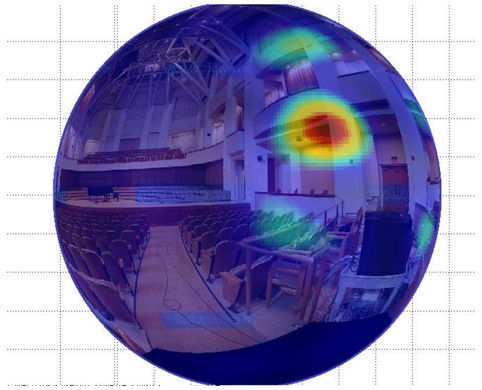
- When the audio and video cameras are collocated the acoustic image can be warped onto the video image.
- Immediately reveals sound emitting objects in the environment.



- The figure to the left shows an active speaker illuminated by the audio camera.
- (Top) acoustic image corresponding to speech
- (Bottom) video stream with acoustic image overlaid.

Concert Hall Acoustics:

- Acoustic engineers are very interested in which objects in a concert hall contribute to which reflection.
- The audio camera can immediately reveal where strong reflections originate.
- Will allow acoustic engineers to identify where to place acoustic damping foam to control and manipulate concert hall reverberation



- In this experiment a loud speaker was placed on the stage underneath the piano at the Dekelbaum concert hall at the University of Maryland.
- A 10ms chirp was played from the stage and the response recorder with the audio camera.
- The image above shows the first reflection of this pulse overlaid on a spherical panorama of the concert hall.

References:

[1] Adam O' Donovan, Ramani Duraiswami, Jan Neumman, Microphone Arrays as Generalized Cameras for Integrated Audio Visual Processing, accepted as an oral presentation, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, 2007

[2] Adam O'Donovan, Ramani Duraiswami, Nail A. Gumerov, Real Time Capture of Audio Images and Their Use with Video Proceedings 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics

[3] Adam O'Donovan, Ramani Duraiswami and Dmitry Zotkin. Imaging Concert Hall Acoustics Using Visual and Audio Cameras. ICASSP 2008.

UMIACS
University of Maryland Institute for Advanced Computer Studies

Audio Camera Website:
http://www.umiacs.umd.edu/~odonovan/Audio_Camera