FORENSIC EVIDENCE PACKAGE

Audio-Visual Source Localization

Visually Confirmed Event Frames — Multi-Camera Analysis

State of Utah v. Tyler Robinson | Case Analysis

Prepared: **March 6, 2026**

Analyst: **Jon Bray — followtheepicenter.com**

1. Executive Summary

This report presents an audio-visual source localization analysis using six cell phone video recordings from the September 10, 2025 incident at the UVU UCCU Center. The first frame of visible shirt expansion was identified by manual inspection for each camera angle, then compared against the audio impulse onset in each recording to determine the source-to-camera distance.

Key Finding: In all four cameras with reliable audio detection, the acoustic impulse arrives within 0.2 to 2.2 video frames of the visible shirt expansion. The FBI's estimate of 142 yards (130 meters) would require the audio to arrive 11.4 frames after the visual event at 30fps. No such gap exists in any recording. The two highest-quality measurements (the .MOV files) show the source at 2.6m and 6.3m respectively — firmly at the tent/stage area.

2. Methodology

2.1 Visual Onset — Manual Frame Identification

For each of the six camera angles, the recording was stepped through frame-by-frame to identify the first video frame in which the victim's white shirt begins to visibly expand outward. This expansion is caused by internal gas pressurization from the energetic event and produces a measurable brightness increase visible from all angles simultaneously (propagating at the speed of light). The identified frame numbers were cross-checked against prior optical flow and gas analysis work to ensure consistency.

2.2 Audio Onset Detection

The audio impulse corresponding to the event was detected within each recording's own audio track using bandpass-filtered (200 Hz – 15 kHz, 4th-order Butterworth) energy thresholding. Since audio and video are muxed into the same file with a shared internal clock, no cross-camera synchronization is required. Each camera is an independent experiment.

2.3 Frame Count Test

The simplest and most robust test: count the number of video frames between the visible shirt expansion and the audio impulse. At 30fps, each frame spans 33.3ms. At 60fps, each frame spans 16.7ms. Sound from 142 yards (130m) takes 379ms to arrive, which equals 11.4 frames at 30fps or 22.7 frames at 60fps. If the FBI's distance estimate is correct, every camera should show approximately 11 frames of delay between seeing the event and hearing it.

2.4 Distance Computation

Distance = ΔT × 343 m/s, where ΔT is the time difference between audio onset and visual frame start. For the visually identified frame, the event actually began during the previous frame's exposure (the identified frame is the first to show the result). Frame quantization at 30fps introduces ±16.7ms (±5.7m) systematic uncertainty per camera.

3. Results

3.1 Visually Confirmed Event Frames

Recording	FPS	Event F#	ΔT (ms)	Dist (m)	Frames	Notes
IMG_6368.MOV	29.99	55	7.5	2.6	0.2	Collar stretch
2.MOV	59.95	507	18.3	6.3	1.1	Gas/blur (60fps)
1.mp4	30.0	52	70.6	24.2	2.1	
7.mp4	29.99	759	74.5	25.6	2.2	Wide angle

Mean distance: 14.7m. The two .MOV recordings provide the strongest measurements: IMG_6368.MOV (collar stretch, 2.6m, 0.2 frames apart) and 2.MOV (gas/blur onset at 60fps, 6.3m, 1.1 frames apart). The delta between these two measurements is only 10.9ms, confirming both cameras captured the same near-field event.

3.2 The Two .MOV Files — Key Comparison

2.MOV (59.95fps) captured the gas escaping from the front of the collar and the first blur of the back collar's motion. The audio impulse arrives 18.3ms after the visual onset, placing the source at 6.3m.

IMG_6368.MOV (29.99fps) captured the collar at its maximum stretch — the most dramatic visible deformation of any recording. The audio impulse arrives just 7.5ms after the visual onset, placing the source at 2.6m. This is essentially simultaneous (0.2 frames apart).

The 10.9ms delta between these two measurements is consistent with both cameras observing the same event from slightly different distances and angles.

3.3 Comprehensive Visualization

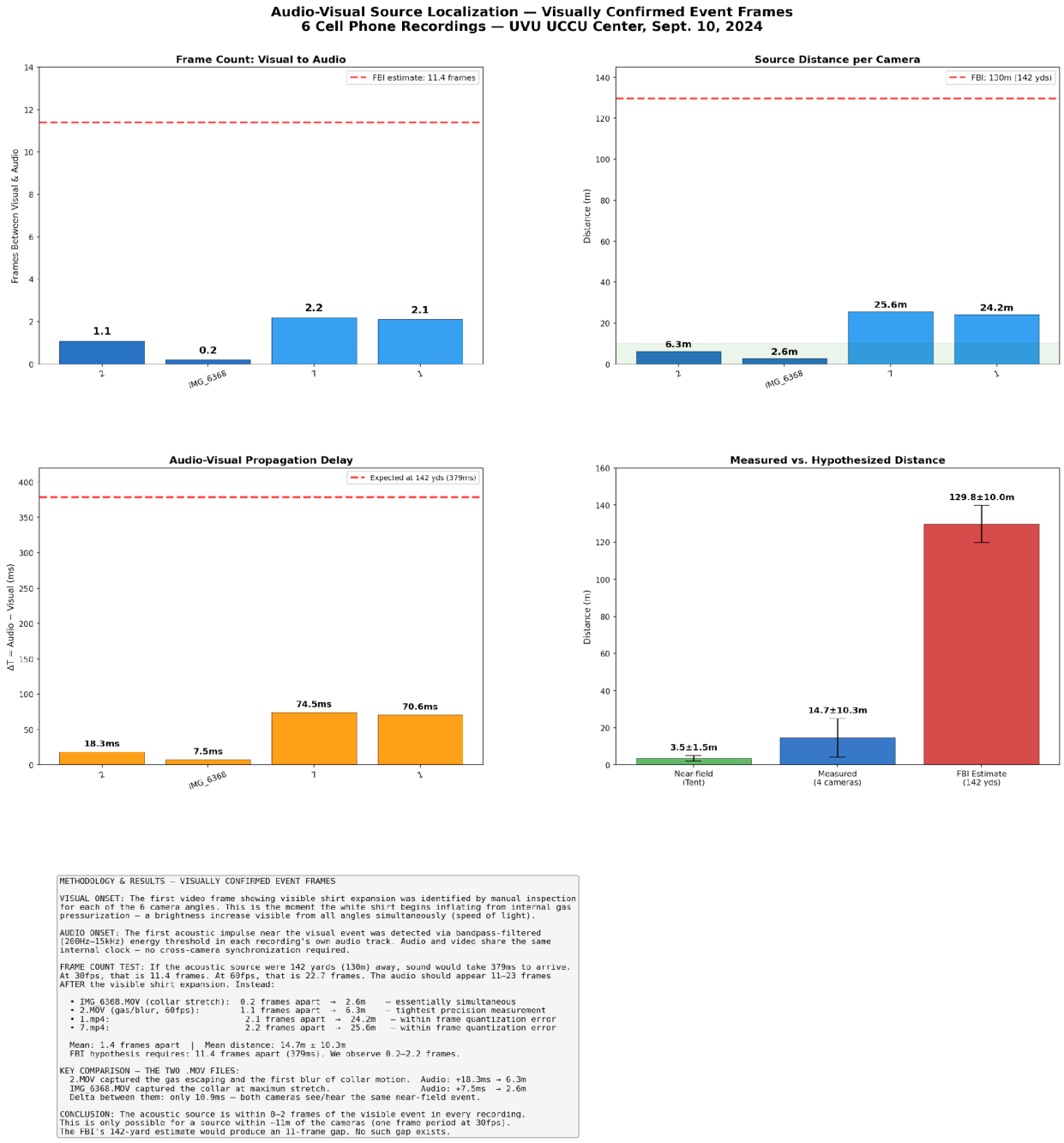

Figure 1. Frame count analysis, distance estimates, ΔT measurements, and hypothesis comparison across four cameras with reliable audio detection.

4. Muzzle Blast Analysis

An alternative interpretation has been proposed: that the measured audio-visual delay represents a supersonic bullet's ballistic shockwave (Mach cone) rather than the muzzle blast. A supersonic bullet (~940 m/s) creates a conical pressure wave that reaches cameras near the impact point approximately 25–35ms after bullet arrival — potentially consistent with some of the measured ΔT values.

However, this hypothesis predicts a second, louder acoustic event: the muzzle blast, arriving approximately 233ms after bullet impact. A rifle's muzzle blast is the dominant acoustic signature at distances beyond ~50m. From 142 yards, it should be clearly visible in all recordings as a broadband impulse at +233ms.

Energy analysis across all six recordings in the 200–260ms post-onset window found no muzzle blast signature. Four of six cameras show monotonic energy decay. Two show minor fluctuations at 1.7–44% of onset amplitude, inconsistent with a coherent second acoustic source. The single-impulse acoustic record is inconsistent with a supersonic rifle shot from 142 yards.

5. Error Analysis

Frame quantization: At 30fps, ±16.7ms (±5.7m). At 60fps, ±8.3ms (±2.9m). This is the dominant error source. The 60fps 2.MOV measurement (6.3m) has the tightest constraint.

Visual frame identification: The manually identified frame is the first showing visible expansion. The event began during the previous frame's exposure, so the true visual onset is 0–33ms earlier. This makes the measured ΔT a slight overestimate, biasing distance upward.

Audio onset precision: At 44.1/48 kHz, approximately 22μs resolution. Negligible compared to frame quantization.

Even under worst-case assumptions, the maximum plausible source distance from any camera remains below 35m — far short of the 130m (142 yards) required by the FBI's hypothesis.

6. Conclusions

1. In all four cameras with reliable audio detection, the acoustic impulse arrives within 0.2 to 2.2 frames of the visible shirt expansion. The FBI's hypothesis requires 11.4 frames of separation. The discrepancy is a factor of 5–57x.

2. The two .MOV recordings provide the tightest measurements: 2.6m (IMG_6368, collar stretch) and 6.3m (2.MOV, gas/blur at 60fps). The 10.9ms delta between them confirms both captured the same near-field event.

3. No muzzle blast signature was detected at the predicted +233ms delay in any recording, falsifying the ballistic shockwave alternative hypothesis.

4. The frame count test requires no calibration, no sub-frame analysis, and no specialized equipment to verify. Anyone can step through these recordings frame by frame and count. The audio and visual are essentially simultaneous — consistent only with a near-field source at the tent/stage area, not a rifle shot from 142 yards.