Watching the Watchers: Automatically Inferring TV Content From Outdoor Light Effusions

•••

Yi Xu, Jan-Michael Frahm and Fabian Monrose
CCS 2014

Bart Kosciarz

Introduction + Why Should You Care?

Exploit emanations of changes in light to reveal TV content

Can be done from 70+ meters away

Privacy concerns

- Religious beliefs, political views, private things
- U.S. Video Privacy Act of 1998
- 67% of people watch TV during dinner

Related Work

Power usage + power line electromagnetic interference

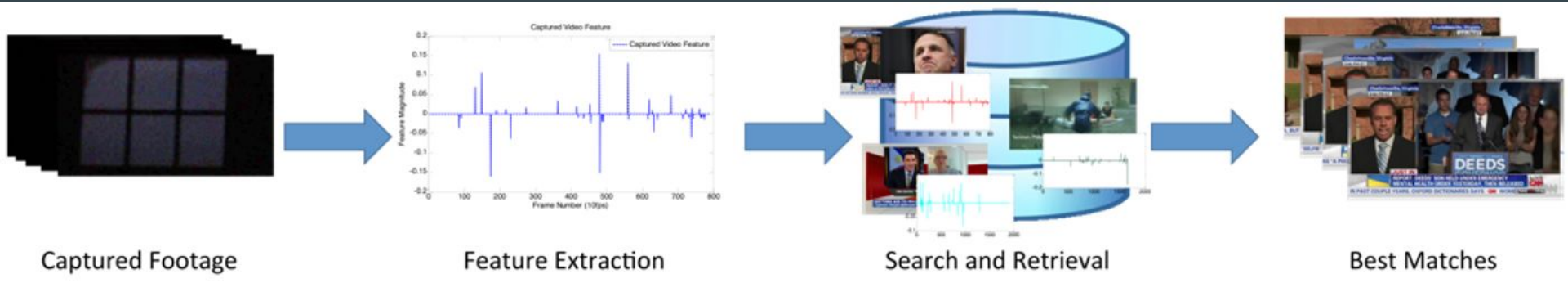
- Depends on TV model / structure of power system

Shiny object reflections

- Recover static image
- Require a view of the screen

Overview

Can we infer content based on brightness changes in a room?



Sugar, Spice, and Everything Nice

What we care about to pull this off

- Quality of captured information (SNR)
- Entropy of observed information
- Length of captured signal
- Size + uniqueness of reference library

Methodology - Feature Extraction

- Compute average pixel brightness for each frame
- Gradient of average brightness signal is what we care about
  - 95% of consecutive frames have the same average intensity
- Feature vector = composition of peaks

Also do this for every video in the database

Methodology - Finding the Best Match

Nearest neighbor search across subsequences

Similarity metric for correlation between two signals

- Assumes the same starting point of both signals
- Computationally hard to exhaustively search
- Takes around 188 seconds to locate a video from 54,000 videos

Methodology - Finding the Best Match

- Sliding window of length 512 over the gradient feature
- Omit all peaks below 30% of the strongest peak's magnitude
- Compute histogram of pairwise distance between peaks

- Index peak features in a K-d tree
- "Found" when best match is stable for 3 iterations
- Search time goes down to 10 seconds

Reference Library

- 10,000 movies
- 24,000 news clips
- 10,000 music videos
- 10,000 TV shows

Over 18,800 hours of video

Extract feature vectors for all of these

Experimental Setup

Record the reflection of TV from a white wall

Distance of 3 meters

Randomly select 62 sequences from the library

Capture with

- Logitech HD Pro Webcam C920
- 60D Canon DSLR

Standard test

Lights off

24 inch screen

Random starting point

Capture Length	60s	90s	120s	180s	240s	270s
Success Rate	39%	49%	54%	70%	85%	94 %

Impact of Room Brightness

Capture 5 videos in 3 different settings

Illumination settings	SNR	Segment Length
Normal brightness level room light off	70	180s
50% brightness level room light off	33	270s
Normal brightness level room light on	15	300s

Impact of Screen Size

Screen Size	SNR	Worst Case Length
24 inch	5	270s
30 inch	48	180s
50 inch	109	180s

Other Factors + Tests

Library Size

- Vary size from 4,000 to 54,000 videos (x 13.5)
- Worst case length from 200s to 240s (x 1.2)

Outdoors

- Attacker positioned on sidewalk
- Observing 3rd floor office window

Outdoors - Results

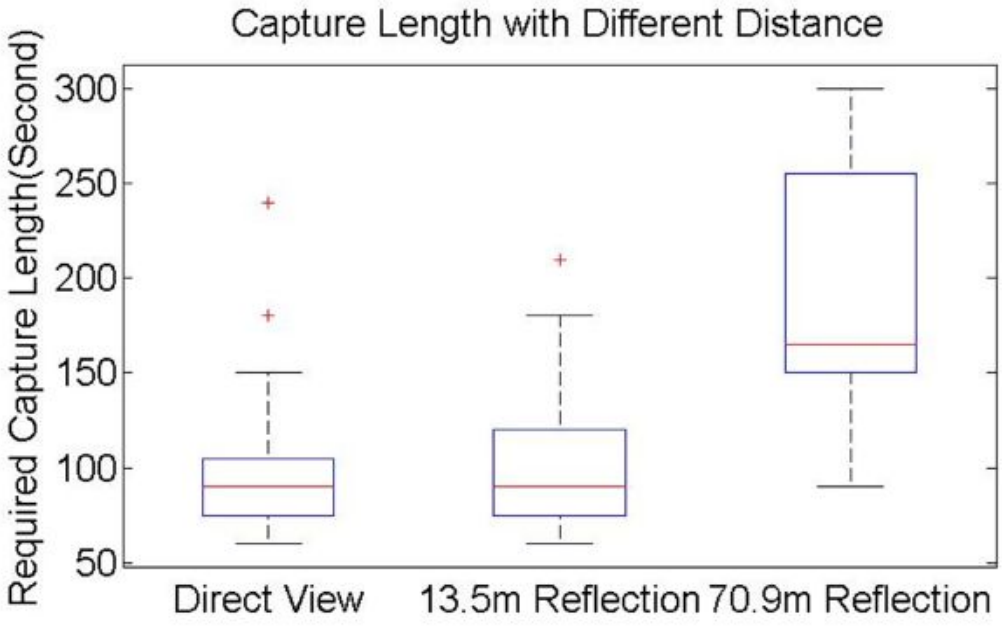
Various distance tests

Average worst case

- 100 seconds at 13.5m
- 190 seconds at 70.9m

Aerial View

Scenario



Mitigations

Curtains

- Vinyl: 3/4 videos after 270 seconds
- Black: 0/4 videos

Lower screen brightness

Flood light

- Blinds camera but doesn’t thwart HDR

Adaptive lighting system

Discussion

What are the key contributions of this paper?

What are the limitations of this approach/Is this attack practical?

How much do people actually care about being targeted by this?