WebRTC - NV Use cases

Rijubrata Bhaumik, Intel

Use Cases :

Face Detection
Background Concealment
Eye Contact Correction
Speech To Text
Noise Suppression

https://aizone/~tuukkat/fac

Not secure | aizone

Guest

getUserMedia Face Detection Test

video

metadata


```
{
  "captureTime": 429435,
```

Face Detection - Why ?

Face Detection on Video Conferencing.

Support WebRTC-NV use cases like Funny Hats, etc ..

Client options - OpenCV.js / TensorFlow.js with WASM/GPU backend

Cloud based solutions like Face API from Azure Cognitive Services or Face Detection from Google Cloud's Vision API.

Solution ? Use camera stack (Image Processing Unit) / client side XPU.

Face Detection - What ?

Shape detection API (WICG) does not support streaming. Ergonomics with TransformStream (BreakoutBox) vs exposing on MediaStreamTrack

All native platforms support FaceDetection (% driver support)

Many platforms have BLINK, SMILE, LANDMARKS, EYE detections

Many client AI solutions give gender, age, head-pose, emotion, anti-spoof

https://github.com/w3c/mediacapture-image/issues/289

Face Detection

```
// Check if face detection is supported by the browser.
const supports = navigator.mediaDevices.getSupportedConstraints();
if (supports.faceDetectionMode) {
       // Browser supports camera face detection.
} else {
       throw('Face detection is not supported');
}

// Open camera with face detection enabled and show to user.
const stream = await navigator.mediaDevices.getUserMedia({
       video: {faceDetectionMode: "simple"}
});
const video = document.querySelector("video");
video.srcObject = stream;

// Get face detection results for the latest frame
const videoTracks = stream.getVideoTracks();
const videoTrack = videoTracks[0];
const settings = videoTrack.getSettings();
if (settings.faceDetectionMode && settings.faceDetectionMode != "off") {
       const attributes = await videoTrack.getAttributes();
       for (const face of attributes.detectedFaces) {
       console.log(
       ` Face @ (${face.boundingBox.x}, ${face.boundingBox.y}),` +
       ` size ${face.boundingBox.width}x${face.boundingBox.height}`);
       }
}
```

Eye Contact Correction

Surface Pro X and all Windows 11 compatible devices modulo driver support.

Facetime for new iOS devices have **Attention Correction.**

https://github.com/w3c/mediacapture-image/issues/290

Resolution : on hold, revisit later.

Background Concealment

Every video-conferencing app has the Background Blur (BB) and Background Replacement (BR) feature as a top feature.

Web versions are anyway using some form of AI inference to implement this feature mostly using TFLite's WASM backend and BodyPix

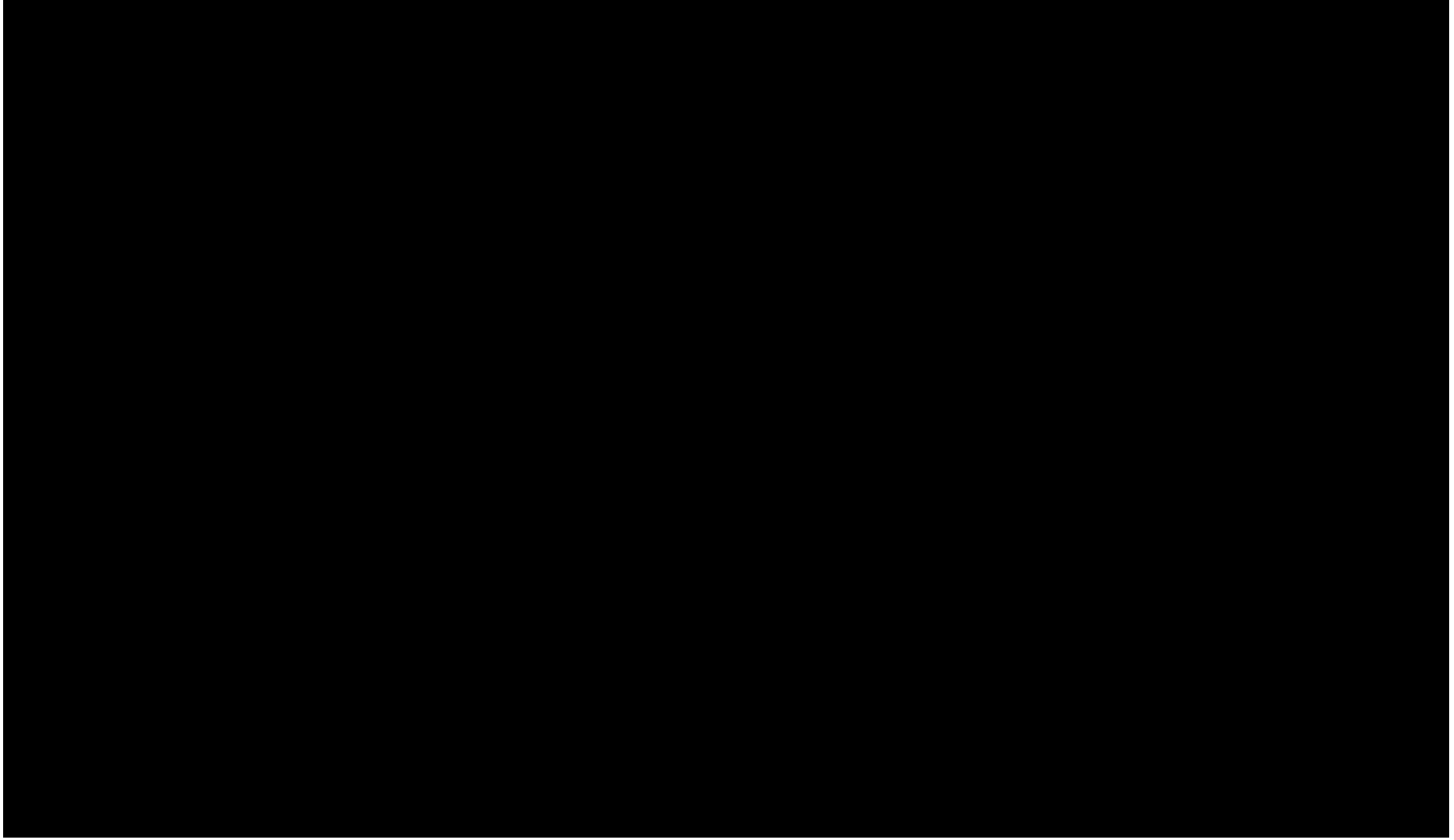
BB and BR should have a combined API?

https://github.com/w3c/mediacapture-image/issues/291

Background Concealment cont’d ..

```
// Open camera and show to user.
const stream = await navigator.mediaDevices.getUserMedia({
    video: true
});
const video = document.querySelector("video");
video.srcObject = stream;

// Try to apply background concealment.
const [videoTrack] = stream.getVideoTracks();
const videoCapabilities = videoTrack.getCapabilities();
if (videoCapabilities.backgroundReplacement) {
    const image = new Image(height, width);
    image.src = "mybackground.jpg";
    await videoTrack.applyConstraints({
        advanced: [{backgroundReplacement: image}]
    });
} else if (videoCapabilities.backgroundBlurLevel) {
      await videoTrack.applyConstraints({
        advanced: [{backgroundBlurLevel: 1}]
      });
} else {
      throw('Background concealment is not supported');
}
```



Speech to Text – Why ?

Web Speech API – Wide number of use cases + Computation on cloud.

Do we want something on MediaStreamTrack, just for WebRTC calls, preferably with local/client-side computation.

SoDA plans ? Google has been building the Live captions feature on Chromium

Noise Suppression -- API?

https://www.w3.org/TR/mediacapture-streams/#def-constraint-noiseSuppression

Boolean to enum { cloud, local} ?

local inference can be on a lowPower/ASIC if such a h/w is present, else on CPU

Save network bandwidth

Privacy – your data does not leave your device

Let’s bridge the native gap

eero, tuukka, zoltan, jianlin, mamatha, wei-fu, priya {intel}
harald, guido, chris, francois, thomas, riley and many more