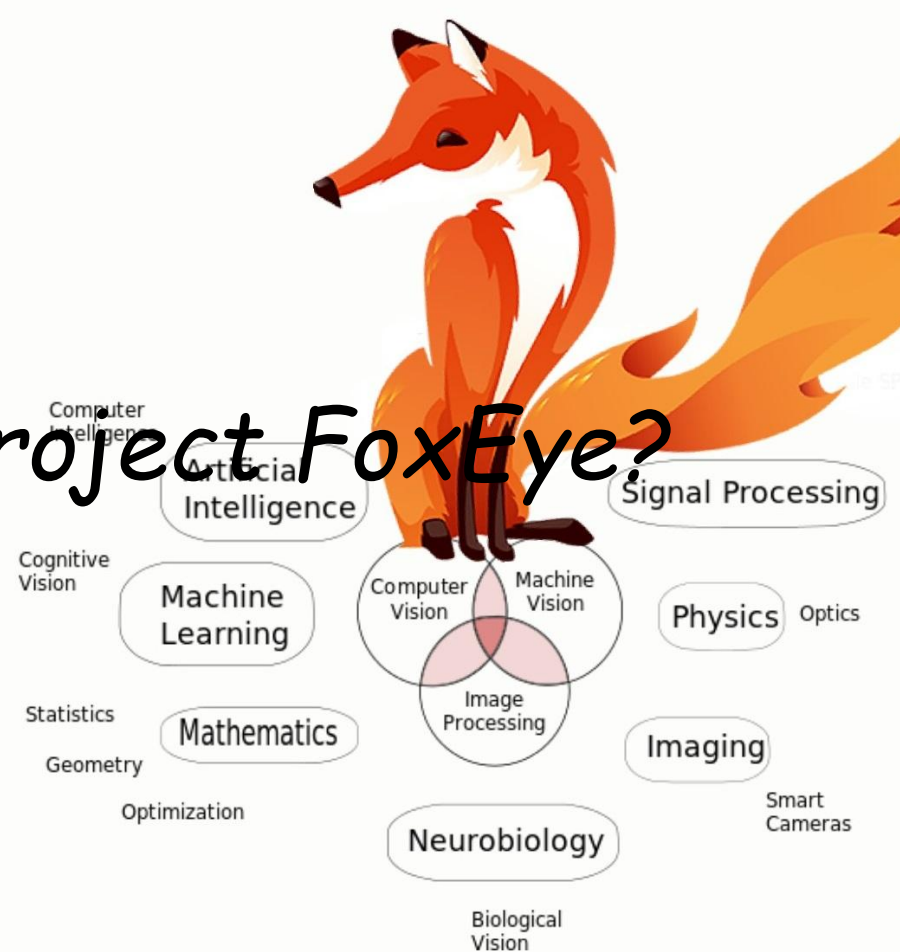


# Project FoxEye

Bring Modern Image Processing and  
Computer Vision Technologies to the Web

Chia-hung Tai(ctai), Tzu-hao Kuo(Kaku)  
TPE Multimedia, Mozilla

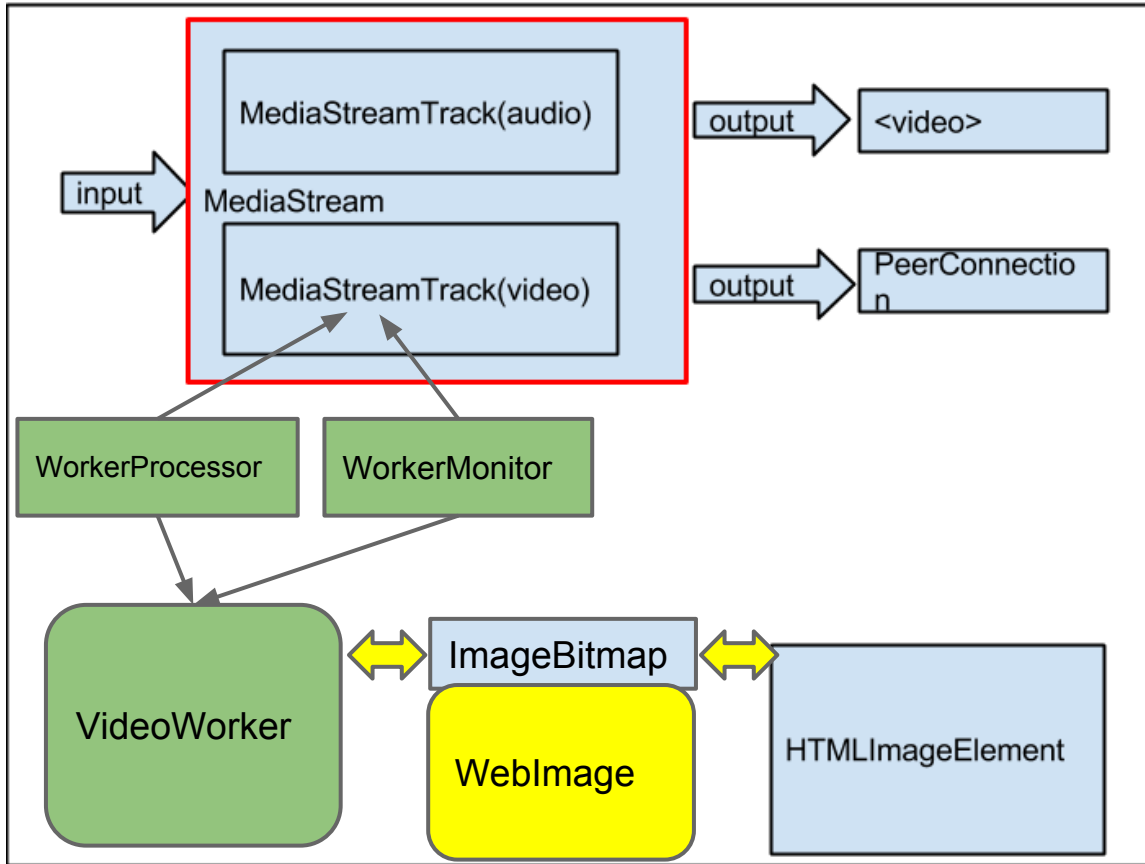
# What is Project FoxEye?



# How it works

- Provide an image processing and computer vision library, WebImage, for some features.
- Associate MediaStreamTrack with Workers
- OfflineMediaContext for rendering as soon as possible

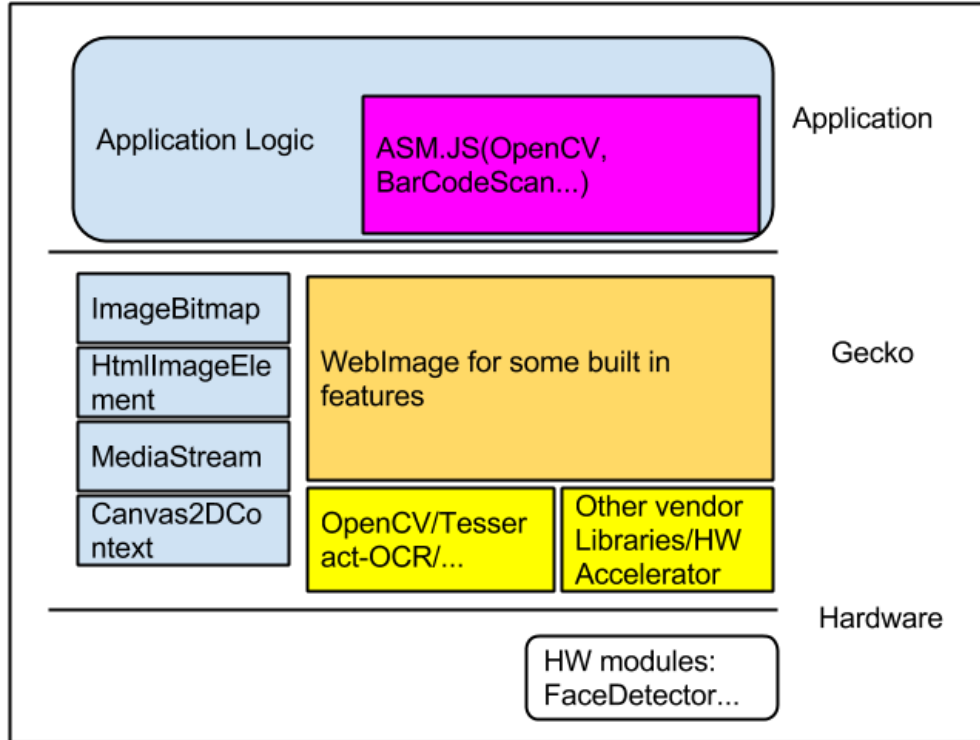
# How it works (Cont'd)



# WebImage

- Reduce the WebImage scope as small as we can-> use asm.js as much as we can.
- Some image processing/computer vision features in asm.js might run poorly in B2G case.
  - Will do the comparison and analysis later.
- For some HW accelerator case...like face detection in B2G
- Can be cross platforms through asm.js

# WebImage(Cont'd)



# API Draft

```
1 partial interface MediaStreamTrack : EventTarget {
2 // WorkerMonitor is for the case of just frames analysis without any modification.
3 // 'parameters' cloned and passed in VideoProcessEvent
4 void addWorkerMonitor(Worker worker, object parameters);
5 void removeWorkerMonitor(Worker worker, object parameters);
6
7 // WorkerProcessor is for the case of modifying frames must be generated in real time.
8 // 'parameters' cloned and passed in VideoProcessEvent
9 MediaStreamTrack addWorkerProcessor(Worker worker, object parameters);
10 void removeWorkerProcessor(Worker worker, object parameters);
11 };
12
13 interface VideoWorkerGlobalScope : DedicatedWorkerGlobalScope {
14 attribute EventHandler onvideoprocess;
15 };
16
17 interface VideoProcessEvent: Event {
18 readonly attribute DOMString trackId;
19 readonly attribute double playbackTime;
20 readonly attribute ImageBitmap inputFrame;
21 attribute ImageBitmap outputFrame;
22 readonly attribute object parameters;
23 };
```

---

# Sample code(Main JS)

Main javascript file:

```
1  var myMediaStream;
2  navigator.getUserMedia({video:true, audio:false}, function(localMediaStream) {
3    myMediaStream = localMediaStream;
4    var videoTracks = myMediaStream.getVideoTracks();
5    var track = videoTracks[0];
6    var myWorker = new Worker("textRecognition.js");
7    track.addWorkerMonitor(myWorker));
8    myWorker.onmessage = function (oEvent) {
9      console.log("Worker recognized: " + oEvent.data);
10   };
11   var elem = document.getElementById('videoelem');
12   elem.mozSrcObject = dest.stream;
13   elem.play();
14 }, null);
```



# Sample code(worker)

---

```
1  var textDetector = WebImage.createTextDetector(img.width, img.height)
2
3  onvideoprocess = function (event) {
4      var img = event.inputFrame;
5      // Do text recognition.
6      // We might use built-in detection function or OpenCV in asm.js
7      var words= textDetector.findText(img);
8      var recognizedText;
9      for (var ix = 0; ix < words.length; ix++) {
10         recognizedText = recognizedText + words[ix] + " ";
11     }
12     postMessage(recognizedText);
13 }
```

---

**Demo**

# Demo

Website:

<http://people.mozilla.org/~tkuo/>

Source code:

- [https://github.com/kakukogou/opencvjs\\_demo\\_segmentation](https://github.com/kakukogou/opencvjs_demo_segmentation)
- [https://github.com/kakukogou/opencvjs\\_demo\\_facedetection](https://github.com/kakukogou/opencvjs_demo_facedetection)

# Demo

HTML5 + JavaScript only...

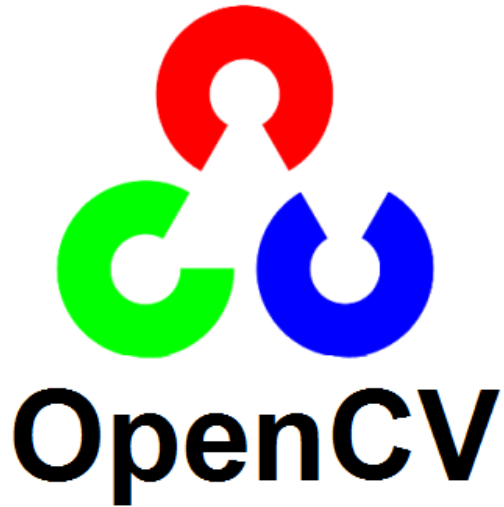
No WebGL

# What have we done...

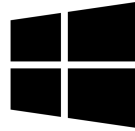
- OpenCV.js
- Gecko + ImageBitmap''

# OpenCV

Open Source Computer Vision

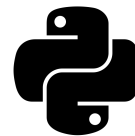


Platforms:



iOS

Languages:



# OpenCV.js

Open Source Computer Vision



**OpenCV**

Platforms:



Languages:

**JS**

# OpenCV.js ... compile



***emscripten***

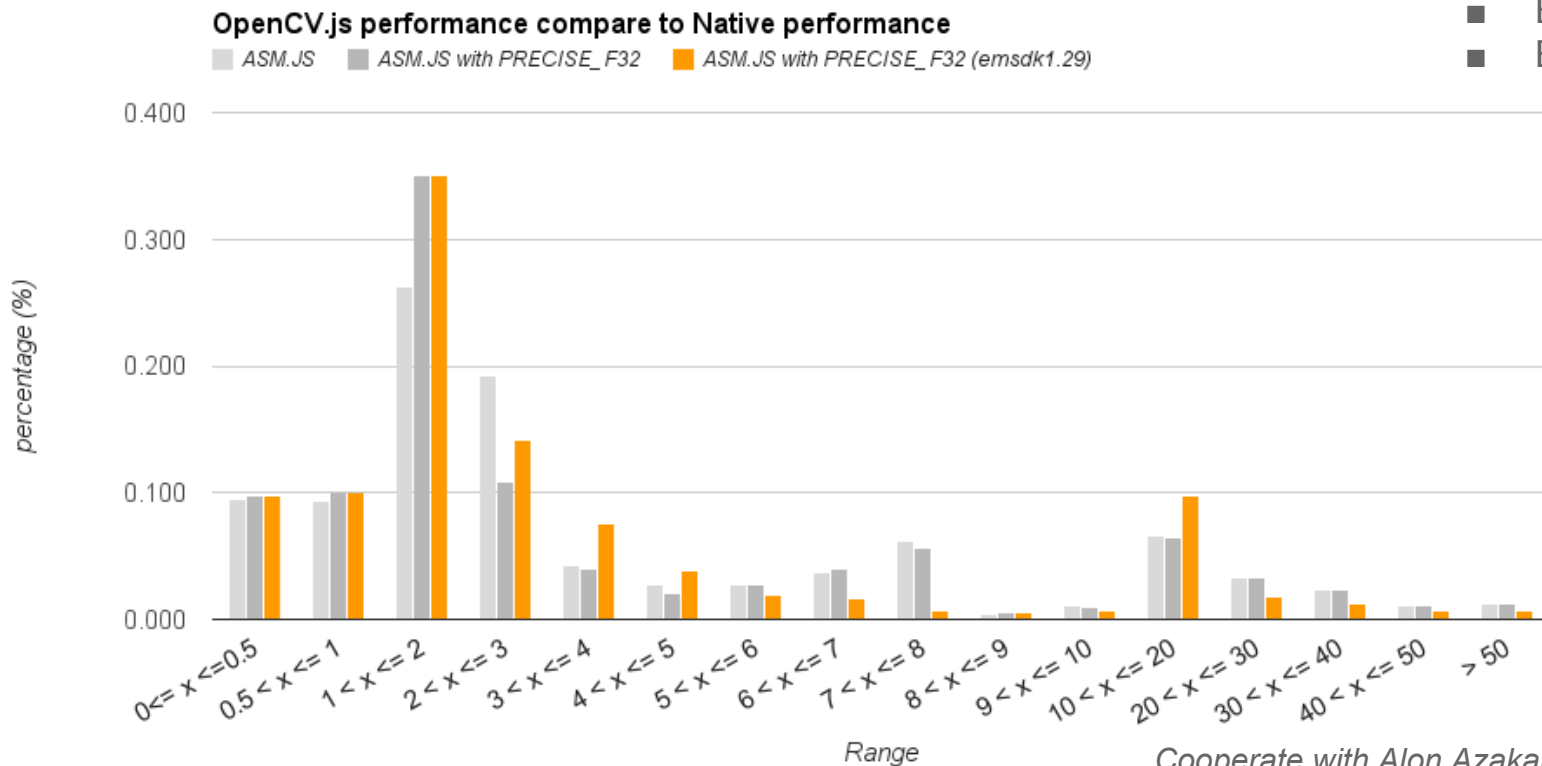
**Guide on github:**

<https://github.com/kakukogou/opencv/tree/opencvjs>



# OpenCV.js ... performance

- Bug 1116674
- Bug 1121860
- Bug 1121877
- Bug 1121908

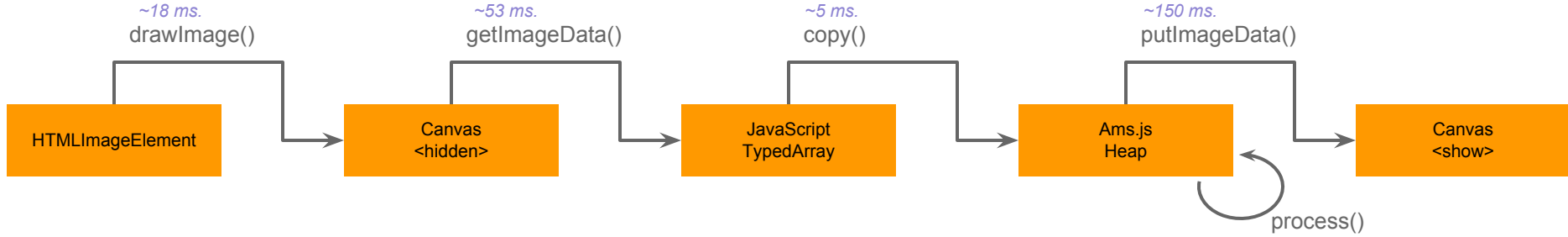


Cooperate with Alon Azakai and Luke Wagner.

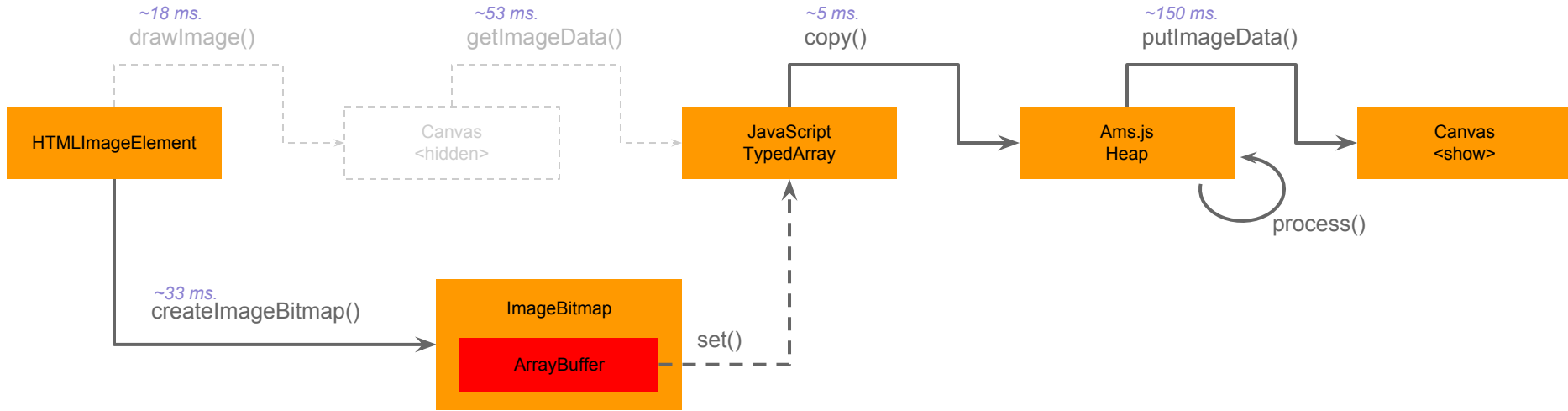
# What have we done...

- OpenCV.js
- Gecko + ImageBitmap''

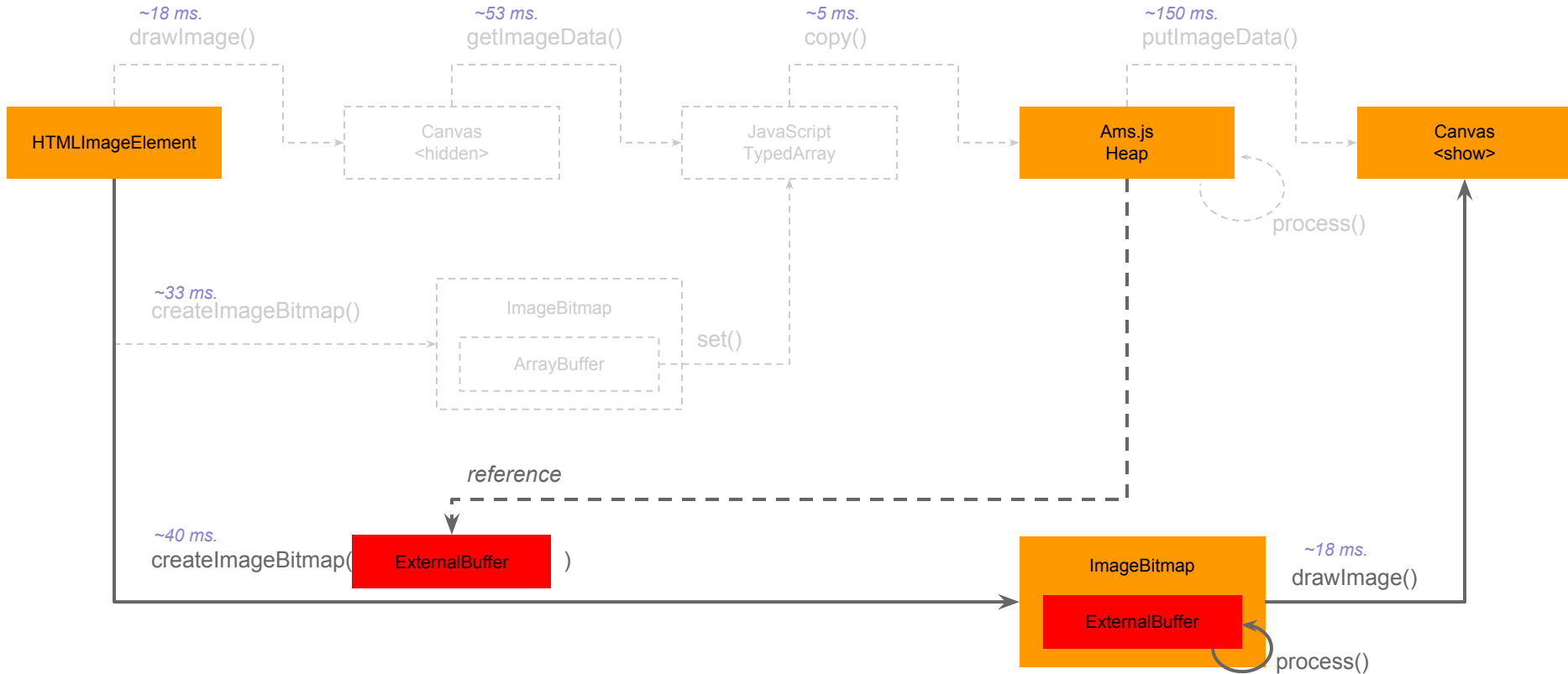
# Gecko



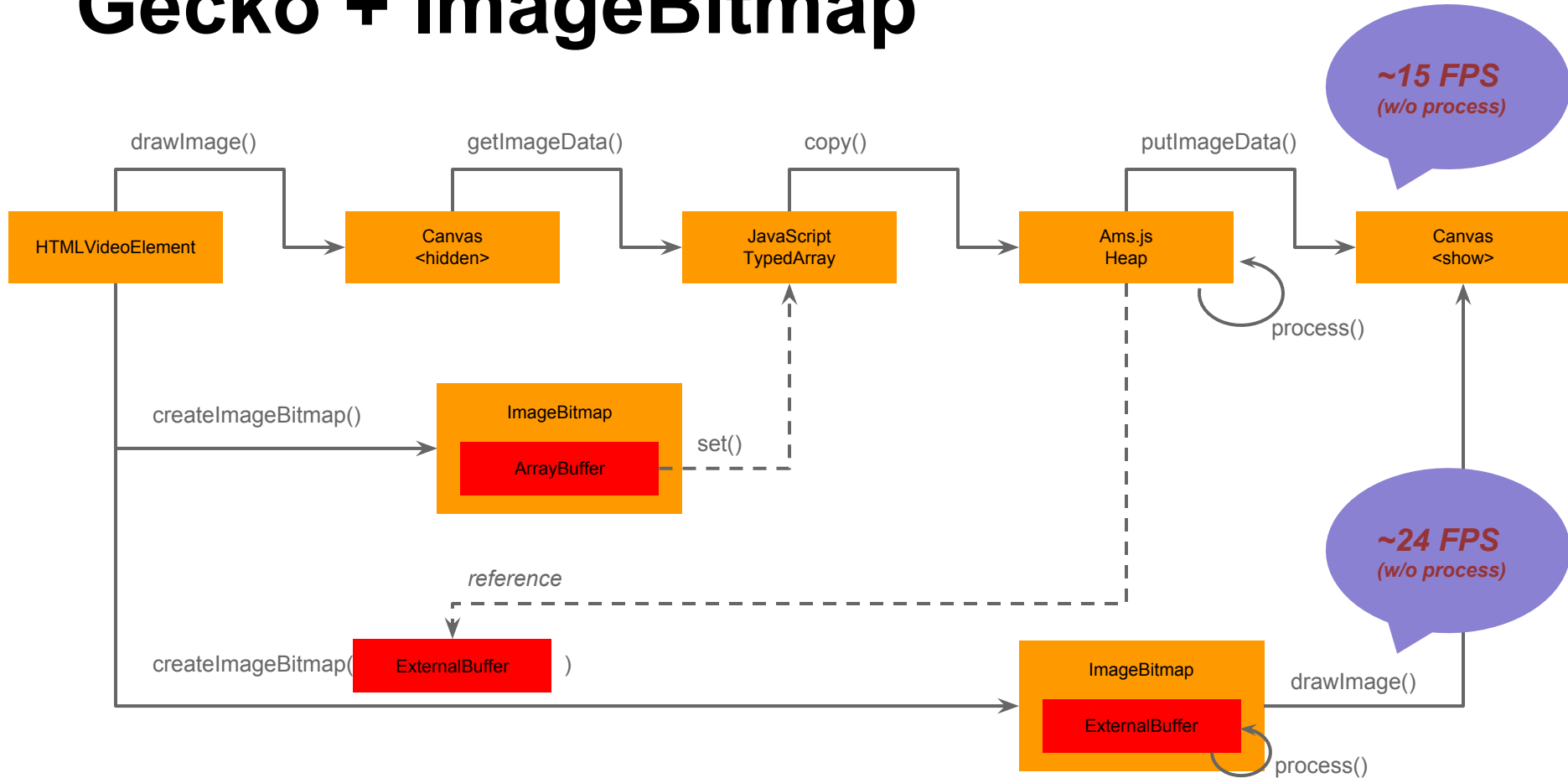
# Gecko + ImageBitmap'



# Gecko + ImageBitmap''



# Gecko + ImageBitmap''



# ImageBitmap''

```
interface ImageBitmap {  
  readonly attribute unsigned long width;  
  readonly attribute unsigned long height;  
  readonly attribute unsigned long channels;  
  readonly attribute unsigned long stride;  
  readonly attribute ImageBitmapFormat format;  
  readonly attribute ArrayBuffer data;  
  readonly attribute unsigned long offset;  
  readonly attribute unsigned long length;  
};
```

```
interface ImageBitmapFactories {  
  Promise<ImageBitmap> createImageBitmap(ImageBitmapSource image);  
  Promise<ImageBitmap> createImageBitmap(ImageBitmapSource image, long sx, long sy, long sw, long sh);  
  Promise<ImageBitmap> createImageBitmap(ImageBitmapSource image,  
                                          ArrayBuffer buffer, unsigned long offset, unsigned long length);  
  Promise<ImageBitmap> createImageBitmap(ImageBitmapSource image, long sx, long sy, long sw, long sh,  
                                          ArrayBuffer buffer, unsigned long offset, unsigned long length);  
};
```

# Drawbacks of Canvas2DContext

- Performance issue
- It is **polling mechanism**, you can't guarantee frame by frame processing
- It uses ImageData (**RGBA only**)
- Canvas2DContext is **not off main thread**
- **No offline processing** (Huge problem in video editor case)






# Current Status(Bug 1100203)

- Bug 1108950: Hooks VideoWorker with MediaStreamTrack(Including VideoProcessEvent)
- Bug 1116674: OpenCV-asm.js on B2G and Desktop Performance
- Bug 1044102: ImageBitmap implementation
- Bug 1108956: OfflineMediaContext support

# Use Cases

- Digital Image Processing(DIP) for camera:
  - Face In, see [Sony Face In](#)
  - Augmented Reality, see [IKEA AR](#)
  - Camera Panorama,
  - Fisheye camera,
  - Comic Effect,
  - Long term, might need Android Camera HAL 3 to control camera
    - Smile Snapshot
    - Gesture Snapshot
    - HDR
    - Video Stabilization
  - Bar code scanner
- Photo and video editing
  - Video Editor, see [WeVideo on Android](#)
  - A faster way for video editing tools.
  - Lots of existing image effects can be used for photo and video editing.
  - <https://www.facebook.com/thanks>

# Use Cases

- Object Recognition in Image(Not only FX OS, but also browser):
  - Shopping Assistant, see [Amazon Firefly](#) 
  - Face Detection/Tracking,
  - Face Recognition,
  - Text Recognition,
  - Text Selection in Image,
    - See <http://projectnaptha.com/> 
  - Text Inpainting,
  - Image Segmentation,
  - Text translation on image, see [Waygo](#) 
- Duo Camera:
  - Nature Interaction(Gesture, Body Motion Tracking)
  - Interactive Foreground Extraction

# Firefox OS V3 Ideation | Content and Context Awareness Web

Team FoxEye,



Let the Web provide **right information** in **right time** and **right place**.

## Where is the idea from?

What if we jump out of box, not only think of a new phone but also a new Web. A strategy to full leverage the treasure we owned, Firefox Browser, is come to our mind.



*Problem Statement:*

Thanks to the internet.

We live in an age of information overload.



How are the symptoms of information overload **after age of smartphone?**  
Getting better?

No! Even worse ...



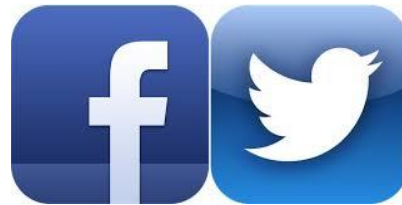
# What is wrong with it?

Too much content!

For example: WWW, FB, Twitter, contacts, gallery...

Thanks to mobile phone.

We get information, anywhere, anytime.





We got a lot of contents. But when we need it...



Mozilla, as a browser and a Web OS provider,  
What can we do?

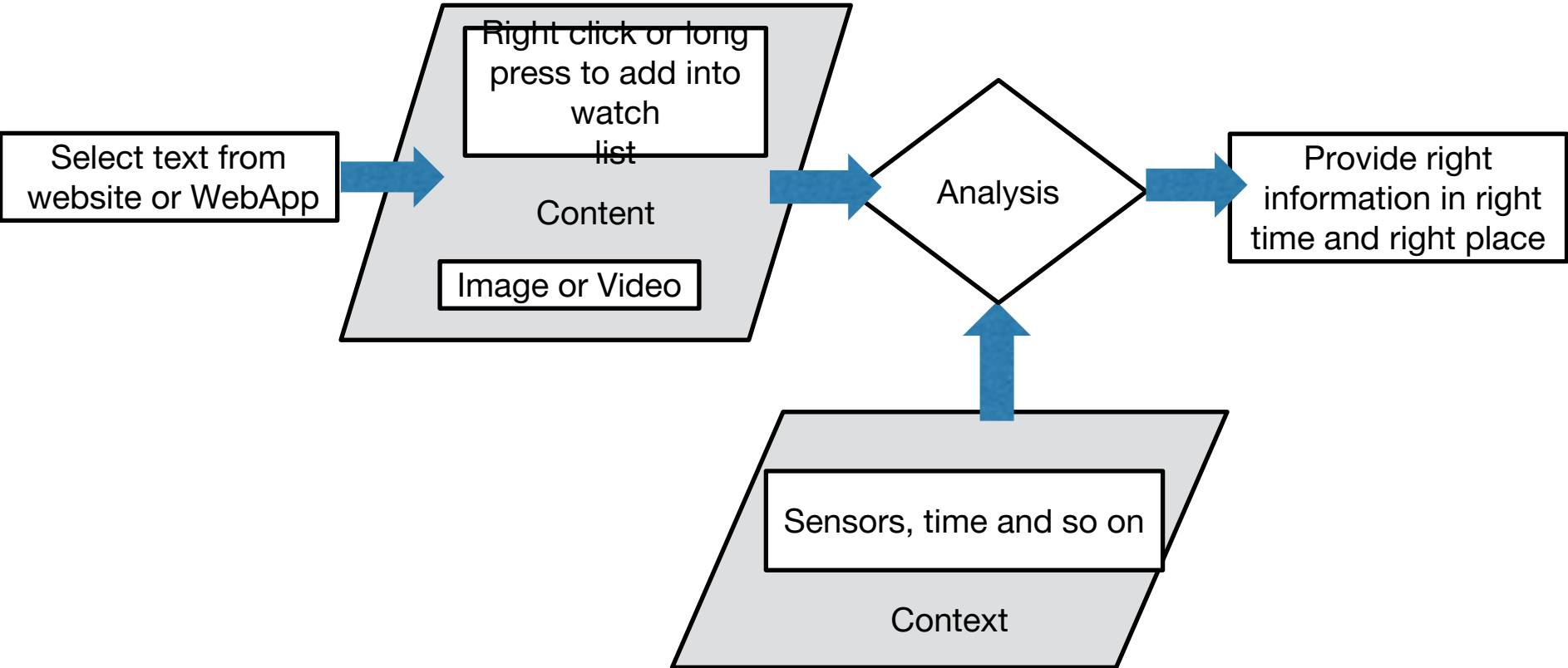


## What is **content and context awareness web?**

*Let the Web knows **what you need, what you want, what you care about** and then provides the **right information in right time and right place.***



# The flow chart



# *Content Awareness Web*

What if the Web knew *Interstellar* is a movie? What kind of new experience can bring to the end user?

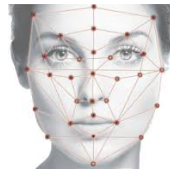
# Concept Detail (Content Awareness Web)

- When you read the post of your friend about the movie, Interstellar, in Facebook, you can add the keyword, Interstellar, into a browser-wide watch list. Once the movie will be on air in your TV channels, you will get a notification to remind you this event from your browser or your Firefox OS devices. You might be able to record the movie in your Firefox OS TV.



## Concept Detail (Content Awareness Web)

- Multimedia content analysis(face recognition): When you hang out and take some pictures with your friends, you can use face recognition to identify your friends and share the pictures with them. So your friend might see the new pictures in their album and you can see the new pictures from your contacts and album app.





# Context Awareness Web

What if the Web know you are in a meeting, or in transportation or walking or sleeping or running? What amazing thing will happen?

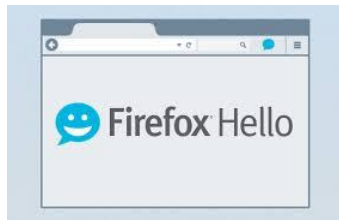






## *Concept Detail (Context Awareness Web)*

- When your calendar shows that you are in a meeting, people who want to call you will get a status update that you are in a meeting. They can choose to send you a SMS, this message could display on your Firefox OS phone or your Firefox browser Hello messenger.
- When you are in transportation or aboard traveling, you might not be available to get a phone call. Your friends can reach you in proper way, for example, through a messenger.



Combine content and context  
awareness web...  
We can provide more...

# Content and Context Awareness Web



- When you walking in the street, you can check your Firefox OS devices to see whether is there a restaurant in your watch list nearby you or not. A restaurant twittered by your friend is just one block away. I added it into my watch list from my Firefox OS devices last month. Decided to have lunch in this restaurant.



# Content and Context Awareness Web (Continued)

- You added LEGO Mindstorms EV3 into your watch list for a while. Today you walk by Toys"R"Us and get a message. What! 20% off on LEGO Mindstorms EV3 today! You walk into the store and buy the toy back home. Because the shop owner use a easy web maker to connect their inventory database, Firefox OS can use near-by network to get the product list and compare to watch list.



amazon  
Your Amazon.com Today's Deals Gift Cards Sell Help

Shop by Department Toys & Games

Search Toys & Games

Toys & Games Best Sellers New Releases Preschool Toys Boy's Toys Girl's Toys Games Outdoor Hobby, Model & Trains Deals Best Toys of the Month Kids' B

LEGO Mindstorms EV3 31313  
by LEGO  
★★★★☆ 254 customer reviews | 99 answered questions

Price: \$359.90 + \$7.49 shipping

Only 1 left in stock.  
Ship from and sold by Ship\_In\_24

- Intelligent EV3 Brick with powerful ARM9 processor, Wi-Fi USB p
- 3 Interactive servo motors, remote control, IRP, and redesigned c
- Control your robot via the infrared sensor system. Innovative softw
- Building Instructions for 5 robots included. Download bonus build
- Control and command your robot with the enclosed remote contro

**Item Under Review**  
While this item is available from other marketplace sellers we may be something wrong with our inventory of th  
We're working to fix the problem as quickly as possible.

49 new from \$349.99 1 collectible from \$333.95

### 3 Questions



*What value (the importance, worth or usefulness of something) does your idea deliver?*  
*Answer:* Give end user what they want in right time and right place. Improve end user satisfaction.

*What technological advantage will your idea deliver and why is this important?*  
*Answer:* This idea fully utilize the advantage that we owned a browser in multiple platforms. The idea breaks the barriers of watch-list from one store only to whole Web. It also leverage the power of cross-platform. So we can give user a new experience, give what they want(watch list, content awareness), in right time and right place(context awareness, cross platform).

*Why would someone invest time or pay money for this idea?*  
*Answer:* If you know what user want, that means more accurate advertisements. And that means a lot of profits behind. You can satisfy not only end-user but also shop owner. That will be win-win situation.

*Thank You*

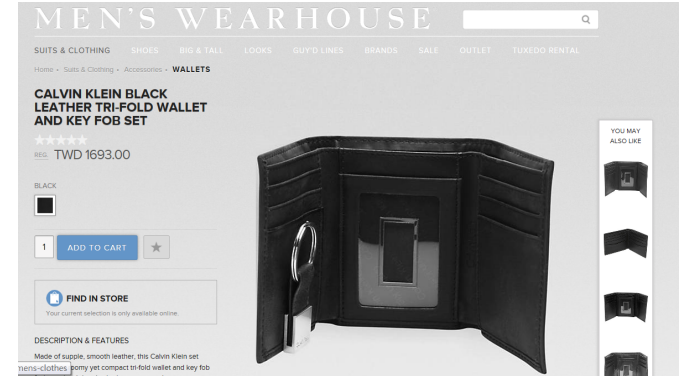


Let the Web provide **right information** in **right time** and **right place**.

**Q&A**

## Concept Detail (Content Awareness Web) continued

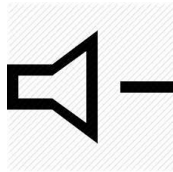
- Text analysis: When you read the post of your friend about the new wallet in Facebook, you can add the keyword, Calvin Klein Black Leather Tri-Fold Wallet, into a web-wide watch list. Once the wallet is on sale in Yahoo shopping, you will get notification in Firefox Browser or Firefox OS devices.





## Concept Detail (Context Awareness Web) continued

- When you go jogging with headphone and someone sends you a message, your music player will lower the volume and read the message through text-to-speech(TTS).
- When you are sleeping, the phone will guess that by ambient light sensor, time and during charging. So the notification will be lower or turn into vibration except phone call and alarm.
- When you are using Firefox Browser in your laptop or desktop, the browser can sync your data from your Firefox OS devices. You can use your browser to manage, edit your pictures and videos. You can run a Photo Story Maker to generate a short video for your trip.

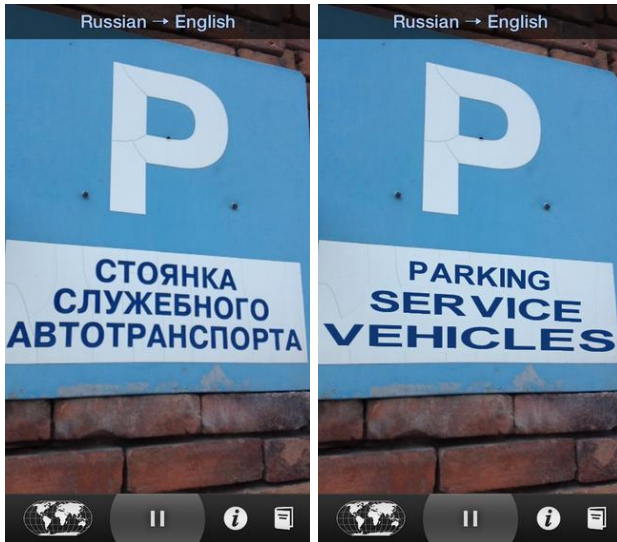


*Unlimited Potentials...*

*Text recognition related use cases*

# Use cases(Word Lens, Waygo)

On-fly offline translator  
with your camera

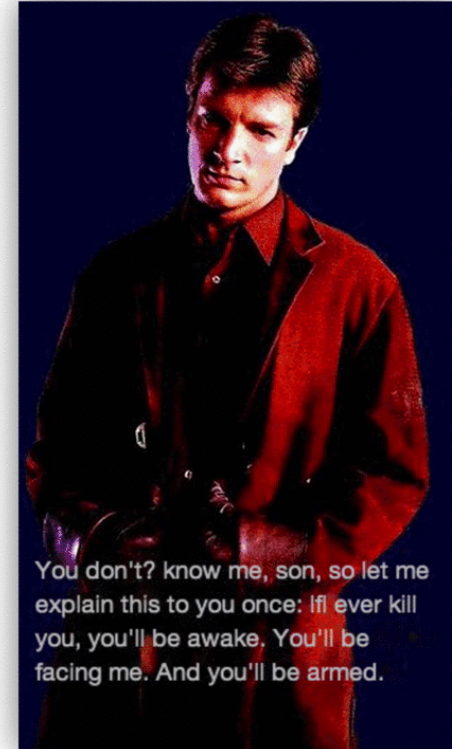
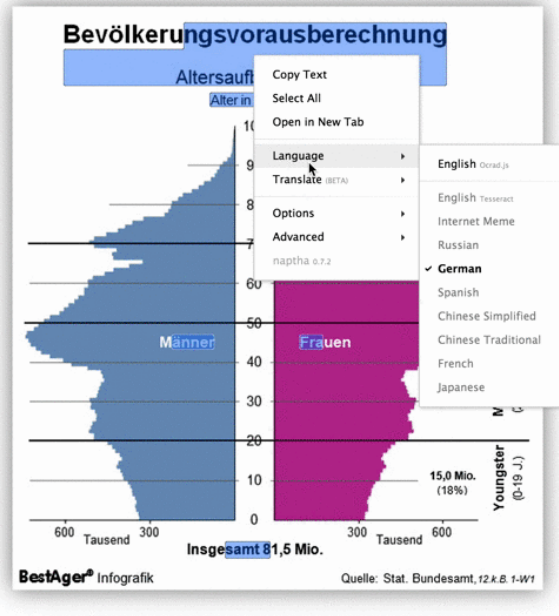
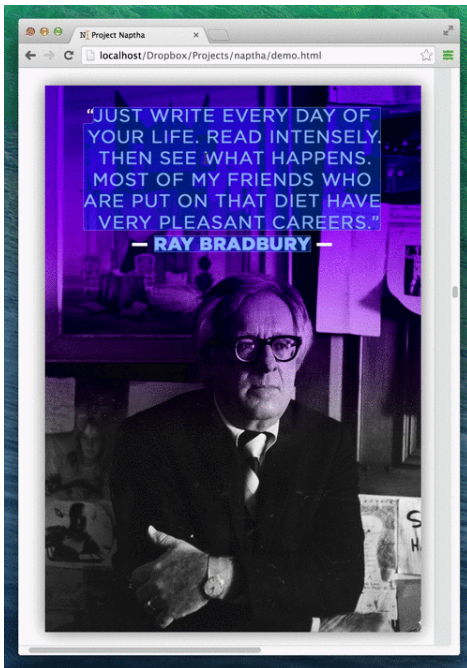


# Use cases(Amazon FireFly)

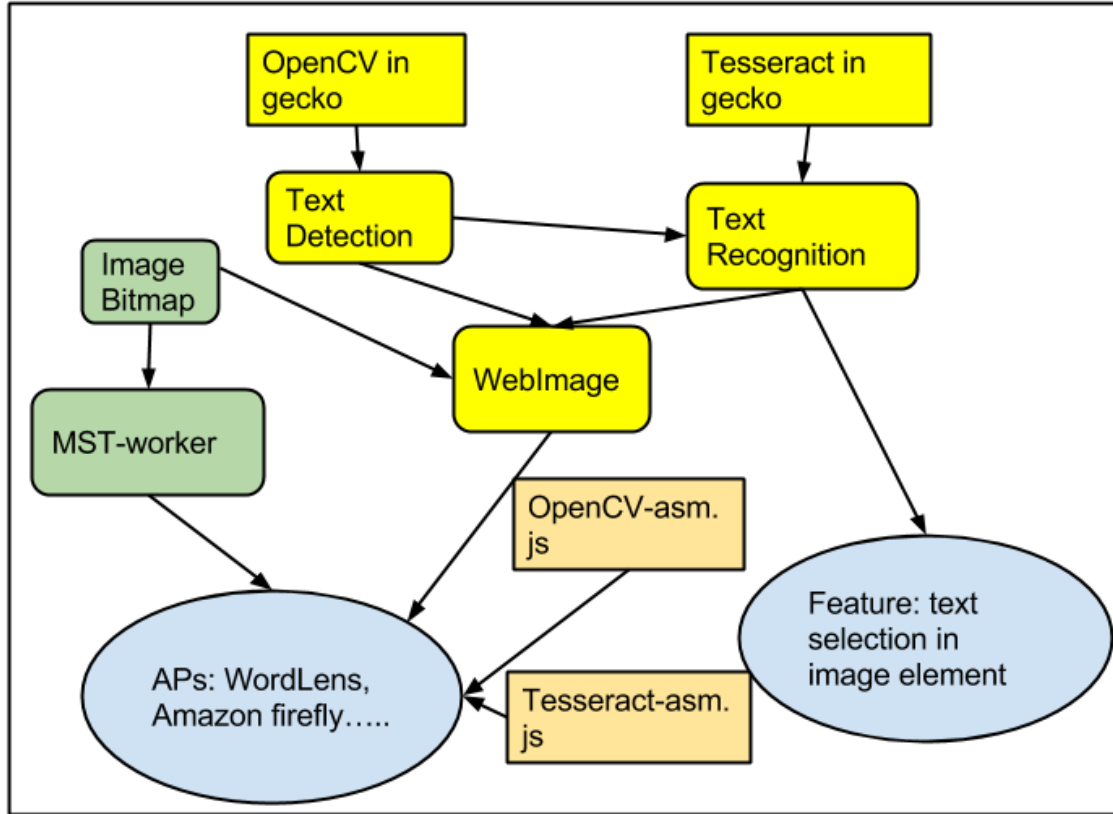


# Use cases(Text selection in Image, browser)

<http://projectnaptha.com/>



# Task dependency(Text recognition)



*Camera effects related use cases*



# Use cases(Augmented Reality)



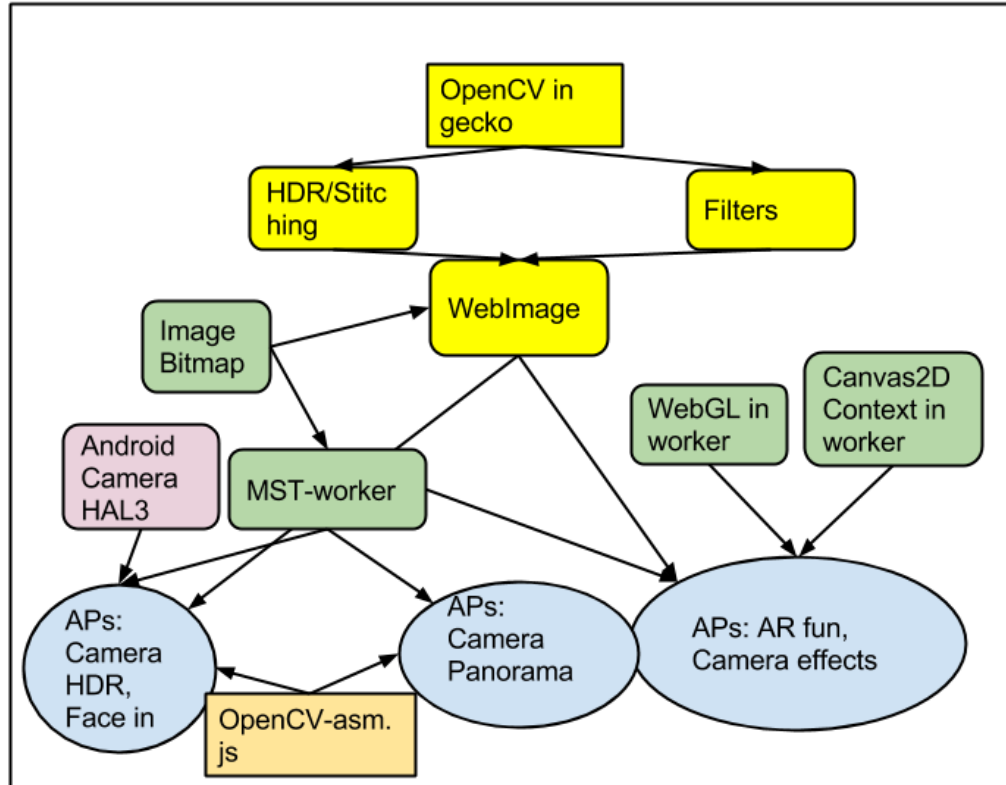
# Use cases(Face in)



# Use cases(Camera Panorama, HDR)



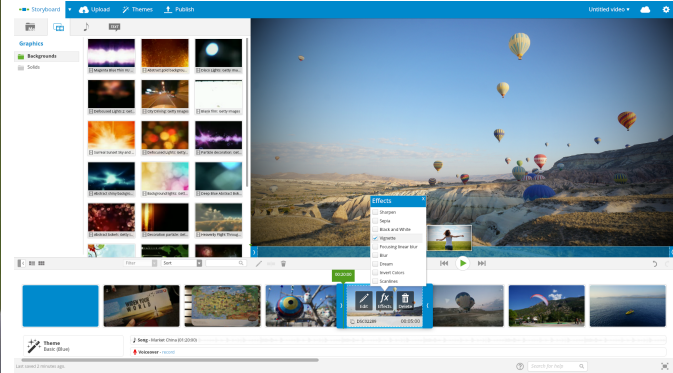
# Task Dependency(Camera Effects)



*Video editor related use cases*

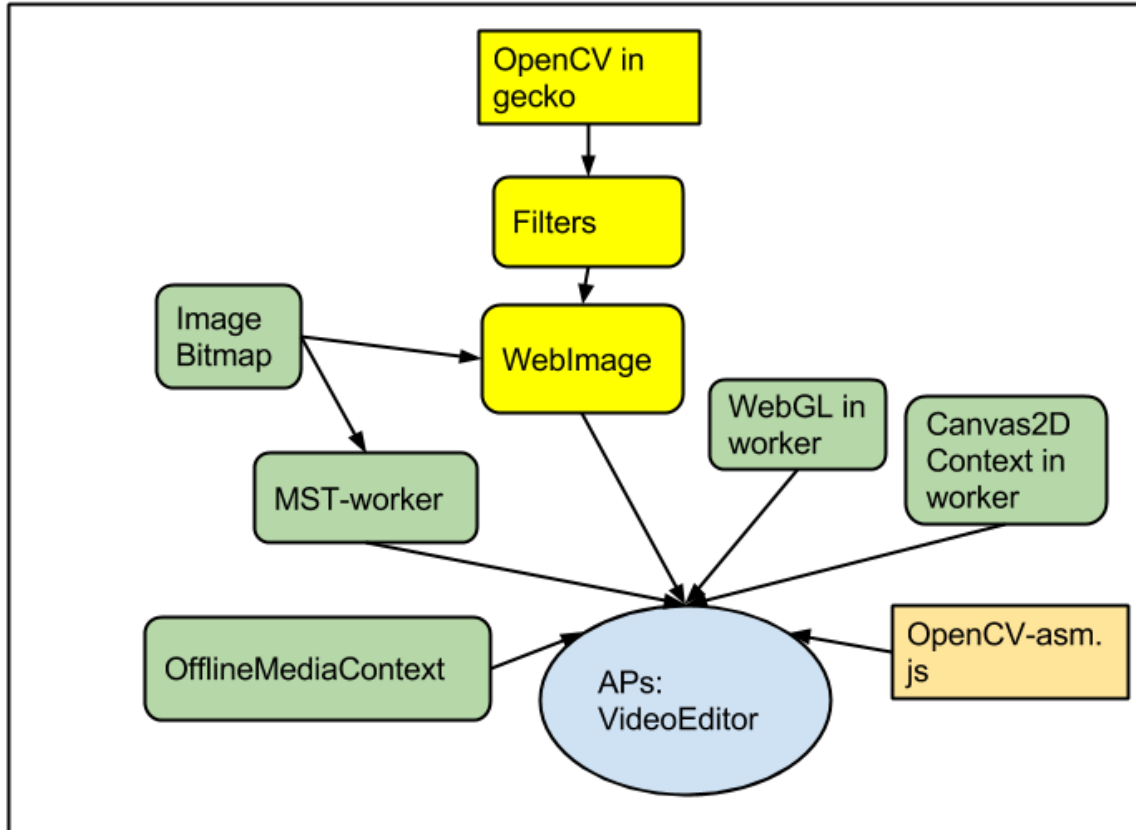
# Use cases(VideoEditor)

WeVideo  
Video Editor  
for Android  
Demo



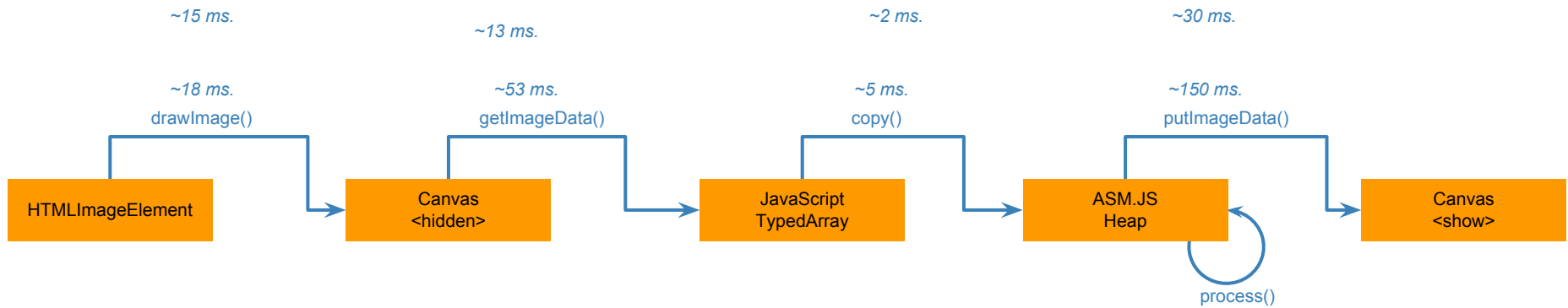


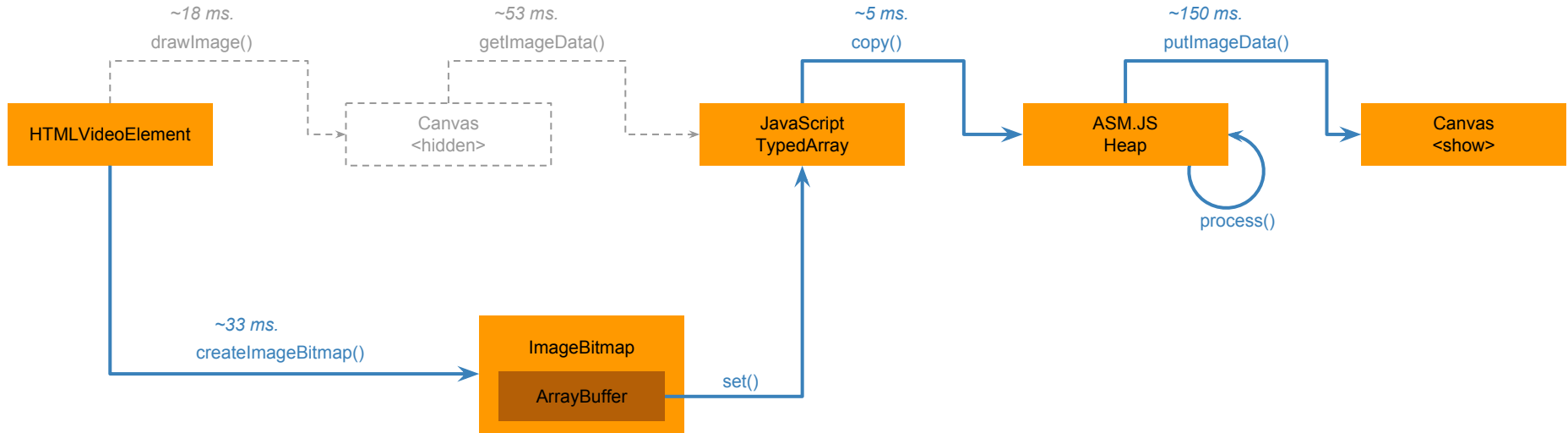
# Task Dependency(Video Editor)



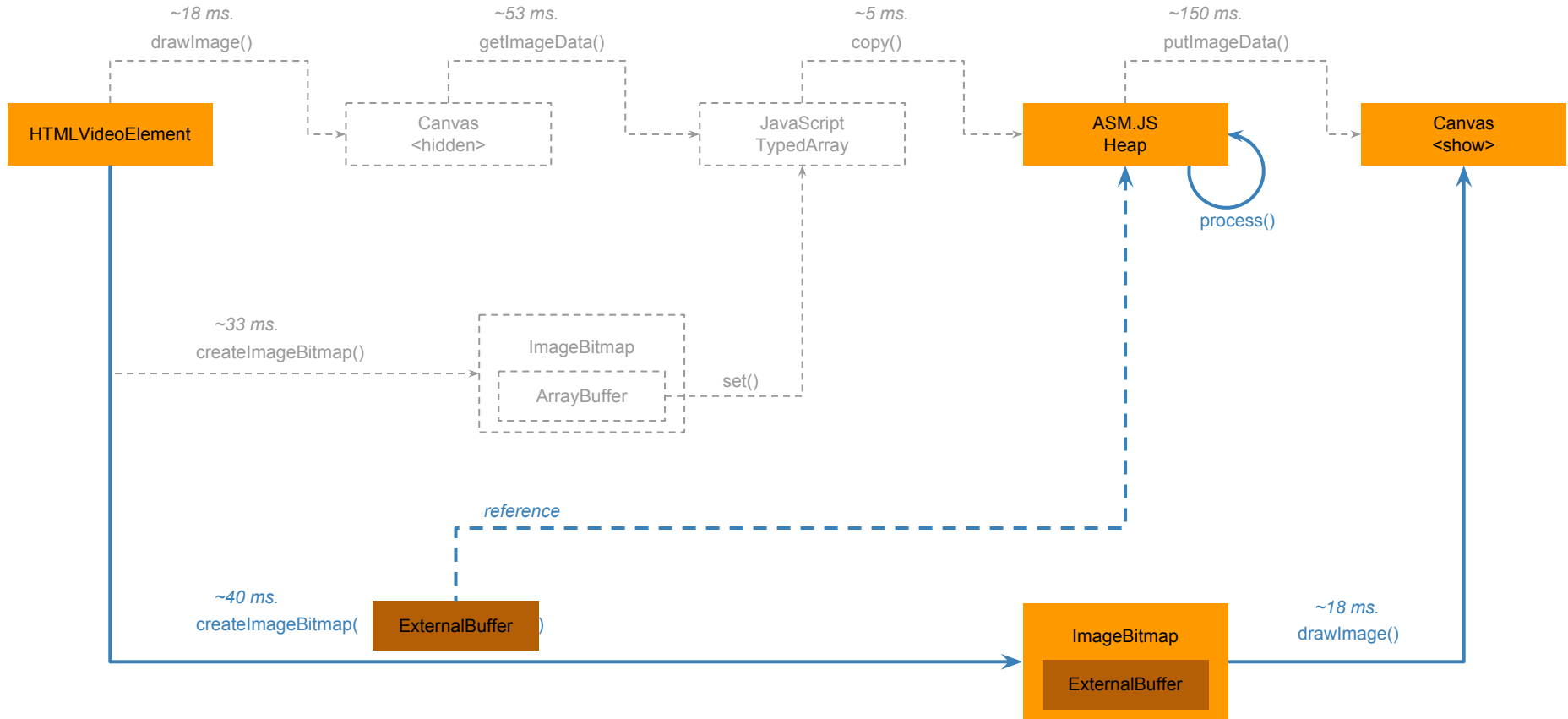


Backup slides





Performance is tested with 640x480, RGBA data under Flame device.



Performance is tested with 640x480, RGBA data under Flame device.

# What have we done...

- OpenCV.js

- enable the powerful library on web platform

- Gecko + FoxEye

- streamline and empower the media processing framework

# What we are going to do next...

## ■ OpenCV.js

- fine-tune performance
- define JavaScript interface

## ■ Gecko + FoxEye

- continue to enrich the framework
- VideoWorker + OfflineMediaContext

*Any possible solution...*

# Comparison: Canvas2DContext

```
// Compute and display the next frame
this.renderFrame = function() {
    // Acquire a video frame from the video element
    this.ctx.drawImage(this.video, 0, 0, this.video.videoWidth,
        this.video.videoHeight,0,0,this.width, this.height);
    var data = this.ctx.getImageData(0, 0, this.width, this.height);
    // Apply image effect
    this.effect.filter(data,this.effect.defaultValues);
    // Render to viewport
    this.viewport.putImageData(data, 0, 0);
return;
};
```



# **Current Web Technology Way (HTMLImageElement)**

# Improvement by FoxEye (HTMLImageElement)