Alex Zhukov VideoGorillas

12M frames per second
How I got into a Netflix movie

Based on a true story :)

PETE J. VALLAS
ADAM WAGNER
BRANDON WATERS
MARCO WIERSCH
ASHLEY WOOD
ALEKSANDAR ZAHARIJEVIC

JOEL A. VANDERAA
JOHN WILLIAM WALTER
SHARON WATT
JACK WIGHTMAN
JASON A. YOUNKMAN

THIS PICTURE MADE UNDER THE JURISDICTION OF

AFFILIATED WITH A.F.L.-C.I.O.-C.L.C.

VIDEOGORILLAS

SAG·AFTRA.

technicolor

FOTOKEM

deluxe


Netflix
https://www.netflix.com/watch/80085566?trackId=13630398&tctx=1%2C1%2Ca693892d-a288-4fd2-a23b-41d5e79e73a2-10517612%2C13e069d0-c92f-45c0-98df-c4688ca1ac66_66551808X6XX1542205127316%2C13e069d0-c92f-45c0-98df-c4688ca1ac66_ROOT
Archival Consultants
PHILIP HALLMAN, UNIVERSITY OF MICHIGAN
PEGGY DOW, UNIVERSITY OF MICHIGAN
JOHN WILLIAM VICKERS, INDIANA UNIVERSITY
SERGE TOUBIANA, LA CINÉMATHÈQUE FRANÇAISE
COSTA-GAVRAS, LA CINÉMATHÈQUE FRANÇAISE
HERVÉ PICHARD, LA CINÉMATHÈQUE FRANÇAISE
KEVIN CLOUD BRECHNER
FRED CHANDLER
Avids Provided by
TECHNICOLOR
Avid Engineer
RWAANA BARNES
Additional Avid Provided by
MOVIOLA
Film Equipment Provided by
CHRISTY'S
Technical Advisors
JASON BRAHMS, VIDEO GORILLAS
ALEX ZHUKOV, VIDEO GORILLAS
Audio Transfers
POST HASTE DIGITAL
Main Titles Designed by
yU+co
End Crawl by
BEN SCHOEN, SCARLET LETTERS
JOSH SCHOEN, SCARLET LETTERS
Film Prints by
FOTOKEM
Film Color Timer
SAUL ESCOBEDO
1:35
The Other Side of the Wind
Property Manager
RAFAEL RECINOS

If brute force doesn’t work you aren’t using enough

Where it all began

Background

Employee #1 at Viewdle face recognition in video startup out of Kiev, Ukraine
Acquired by Google/Motorola

Got to LA to pitch face recognition to movie studios.
Met every single studio.
They never needed face recognition.

Face recognition is cool
can you compare two video files?

Revisions · diff.txt

https://gist.github.com/zhuker/34ac0defea7b67b2e3b9a83e843fb73e/revisions

Revisions

Unified | Split

zhuker **revised** this gist 24 seconds ago.

1 changed file with **12 additions** and **7 deletions.**

19 diff.txt

View file

```
@@ -1,3 +1,9 @@
```

	+ This is an important
	+ notice! It should
	+ therefore be located at
	+ the beginning of this
	+ document!
	+
This part of the	This part of the
document has stayed the	document has stayed the
same from version to	same from version to

```
@@ -8,17 +14,16 @@ would not be helping to
```

compress the size of the	compress the size of the
changes.	changes.
- This paragraph contains	
- text that is outdated.	
- It will be deleted in the	
- near future.	
-	
It is important to spell	It is important to spell
- check this dokument. On	+ check this document. On
the other hand, a	the other hand, a
misspelled word isn't	misspelled word isn't
the end of the world.	the end of the world.
Nothing in the rest of	Nothing in the rest of
this paragraph needs to	this paragraph needs to
be changed. Things can	be changed. Things can
- be added after it.	+ be added after it.
	+
	+ This paragraph contains
	+ important new additions
	+ to this document.

zhuker **created** this gist a minute ago.

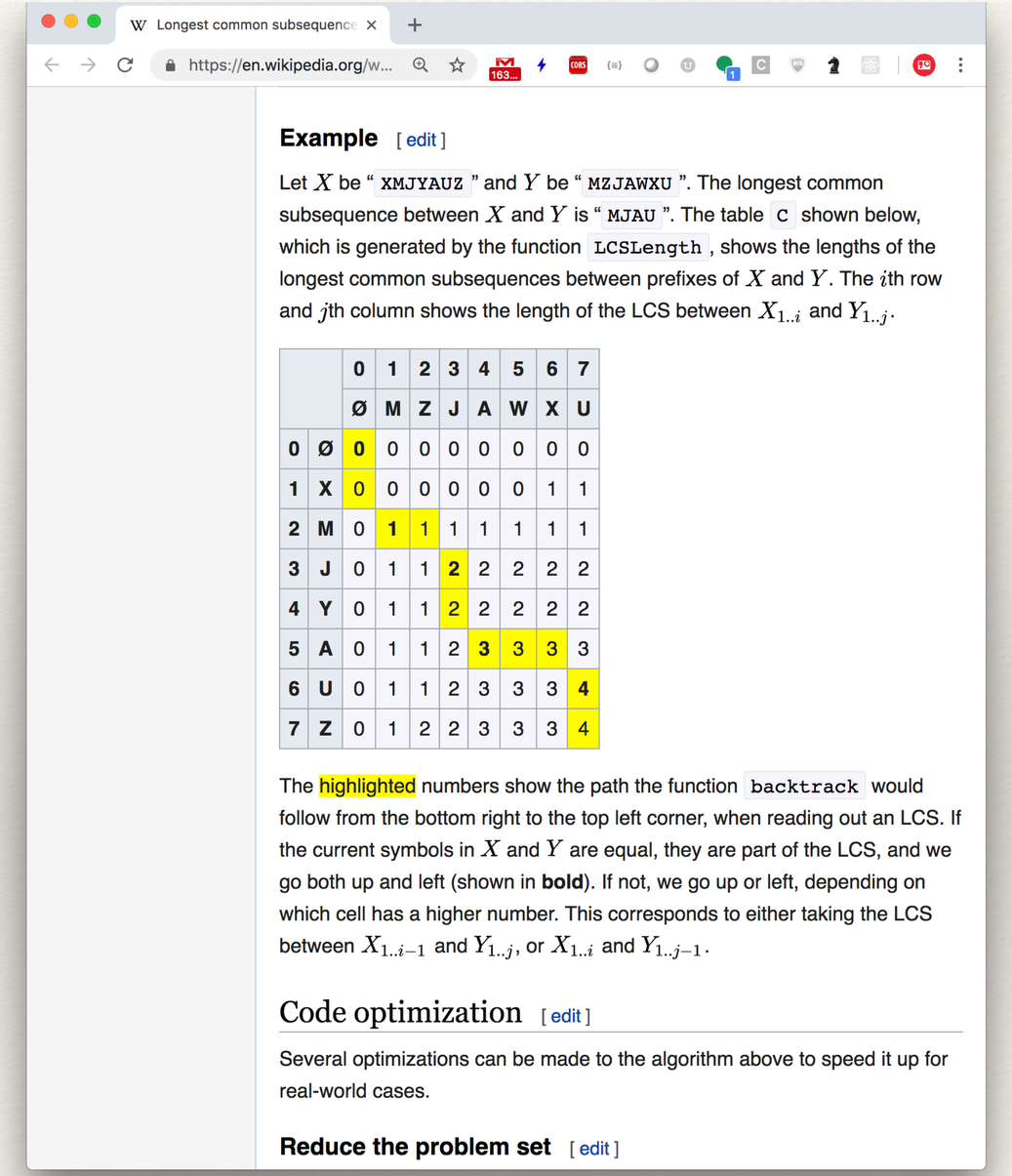
24 diff.txt

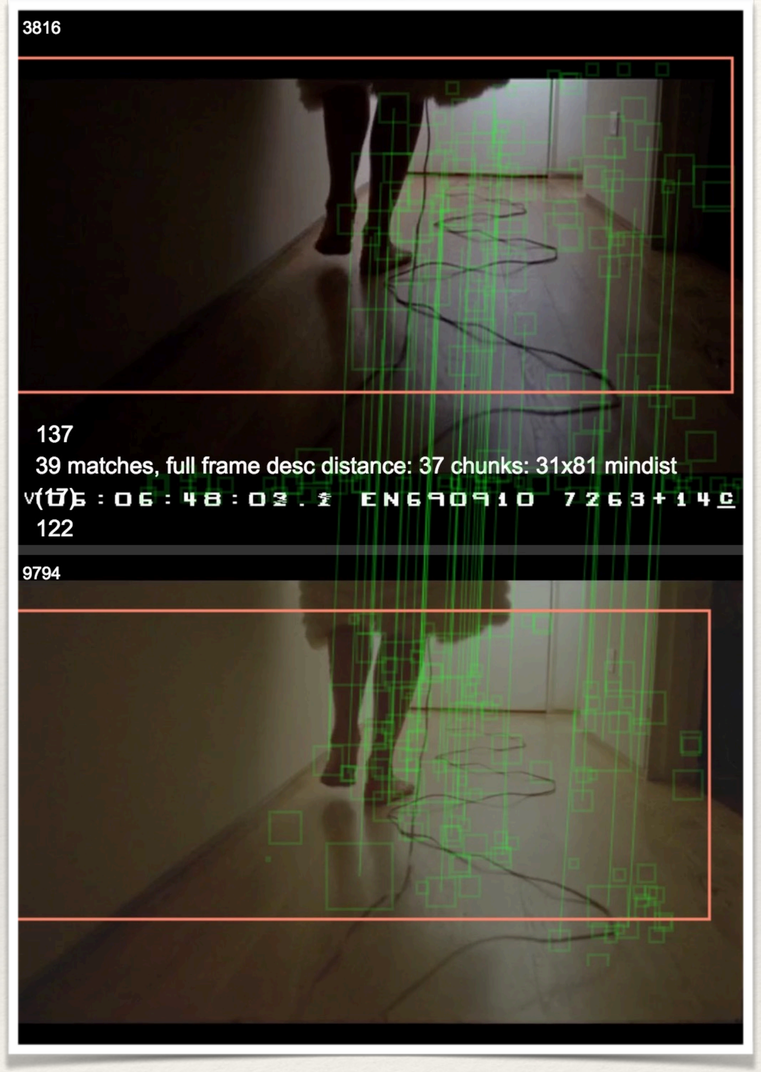
View file

Diff

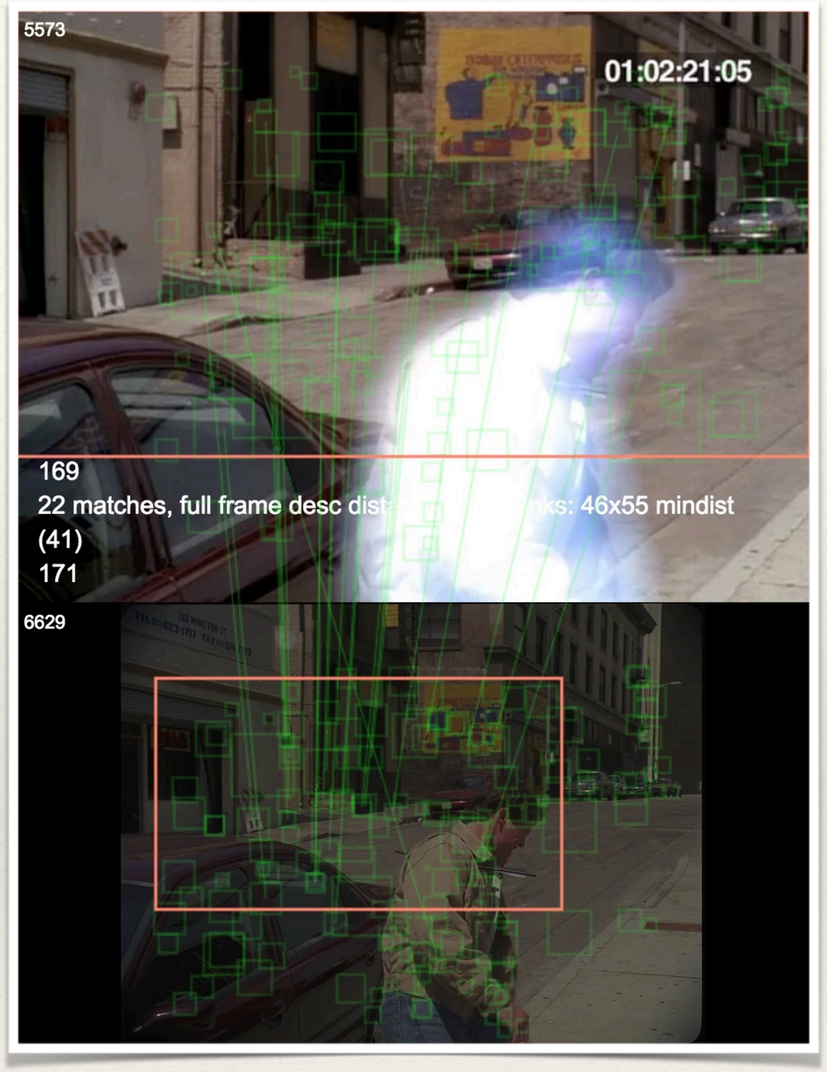
Longest Common Subsequence

https://en.wikipedia.org/wiki/Longest_common_subsequence_problem



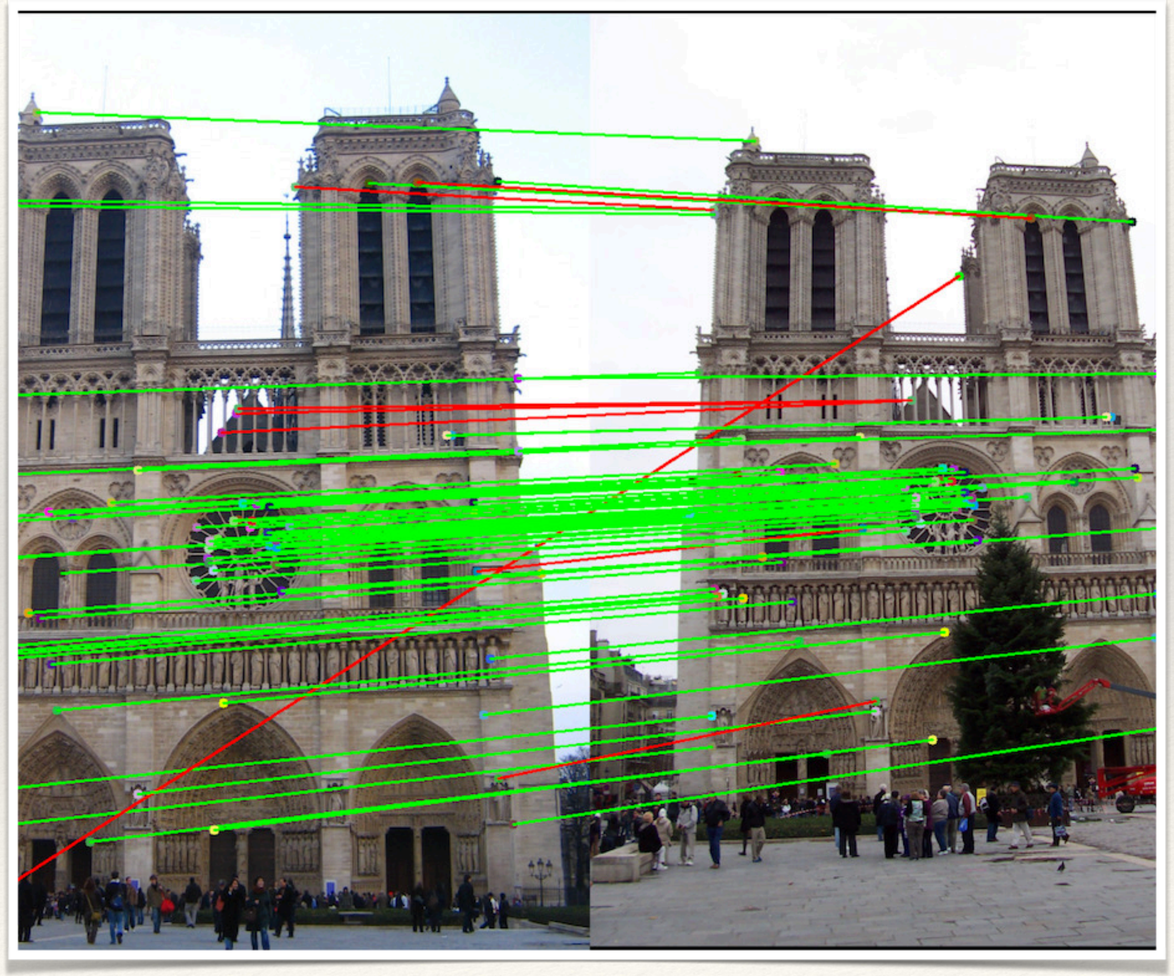

3816
137
39 matches, full frame desc distance: 37 chunks: 31x81 mindist
(17)
122
9794


3294
01:00:46:06
9666


5573
01:02:21:05
169
(41)
171
6629

Old school

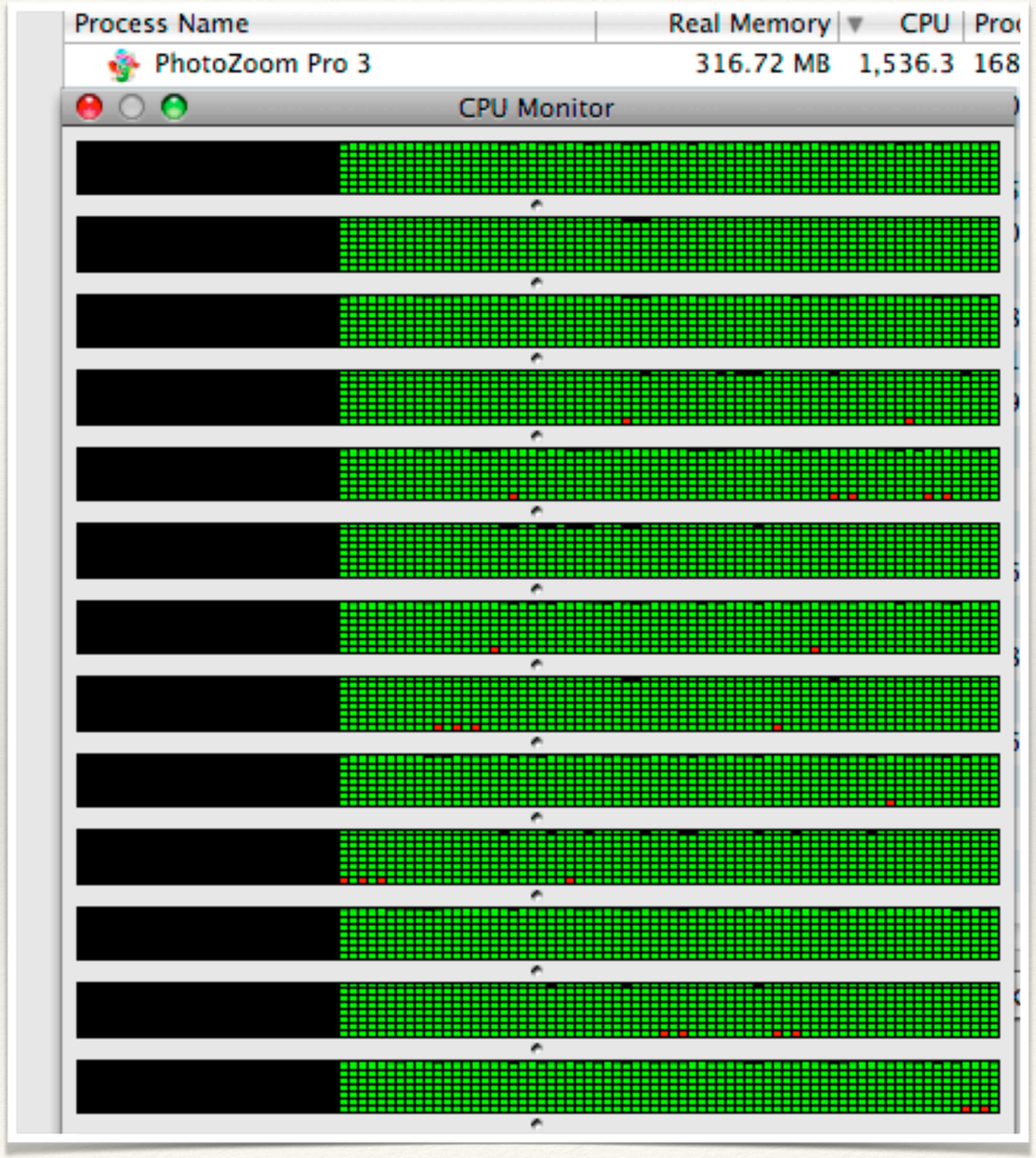
- Why reinvent the wheel?
- Interest point detect/describe
- Match
- Works! 84-92% accuracy
- Good enough
- Awesome!



But there's a catch

Performance

Movie is 200K frames
200K x 200K frame compares = 40B frame compares
64float descriptor at 8TFLOPs = 155 DAYS to compare with bruteforce :)
24GB just to store the descriptor, not talking about any kd-trees etc

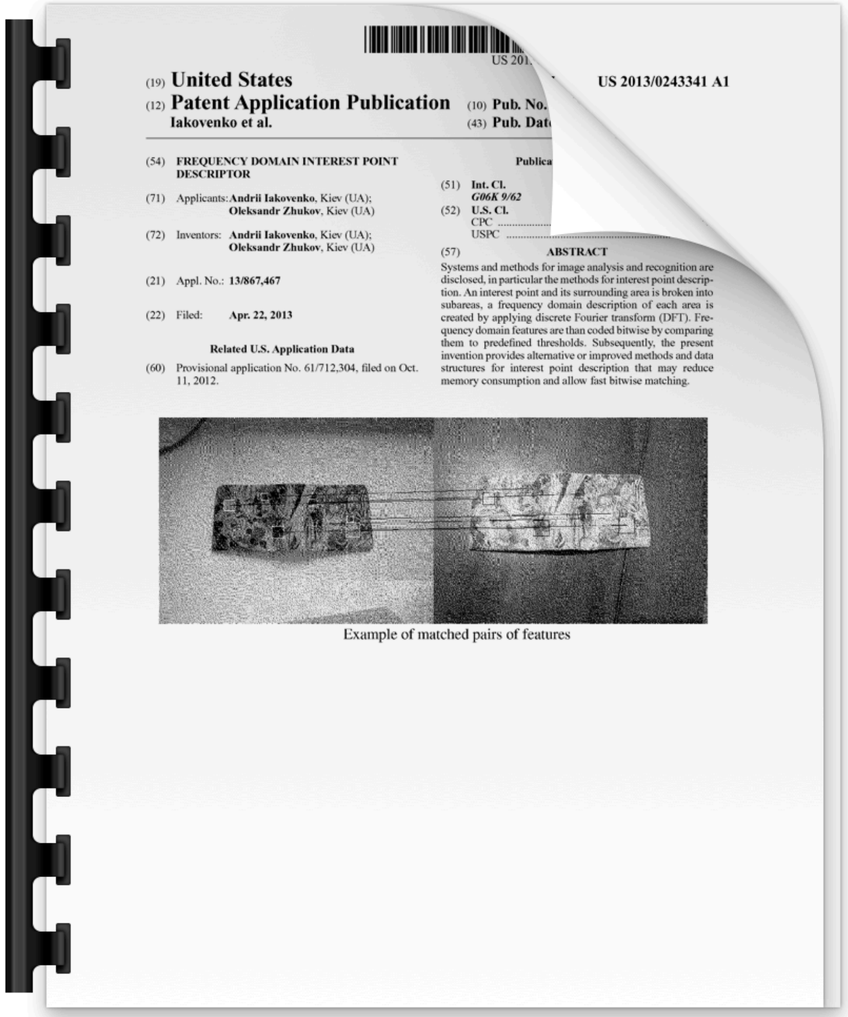


Comes to the rescue

Frequency Domain Descriptor

- Invented here at VideoGorillas
- DCT around interest point
- Frequencies as 160bit vector
- Hamming distance = SUPER fast
- we even bothered to patent it US20130243341A1

```
int hamming_distance(uint64 x, uint64 long y) {
    return popcount(x ^ y);
}
```



(19) **United States**
(12) **Patent Application Publication**
Iakovenko et al.

US 2013/0243341 A1

(54) **FREQUENCY DOMAIN INTEREST POINT DESCRIPTOR**

(71) Applicants: **Andrii Iakovenko**, Kiev (UA); **Oleksandr Zhukov**, Kiev (UA)

(72) Inventors: **Andrii Iakovenko**, Kiev (UA); **Oleksandr Zhukov**, Kiev (UA)

(21) Appl. No.: **13/867,467**

(22) Filed: **Apr. 22, 2013**

Related U.S. Application Data

(60) Provisional application No. 61/712,304, filed on Oct. 11, 2012.

(51) **Int. Cl.**
G06K 9/62

(52) **U.S. Cl.**
CPC .....
USPC .....

(57) **ABSTRACT**

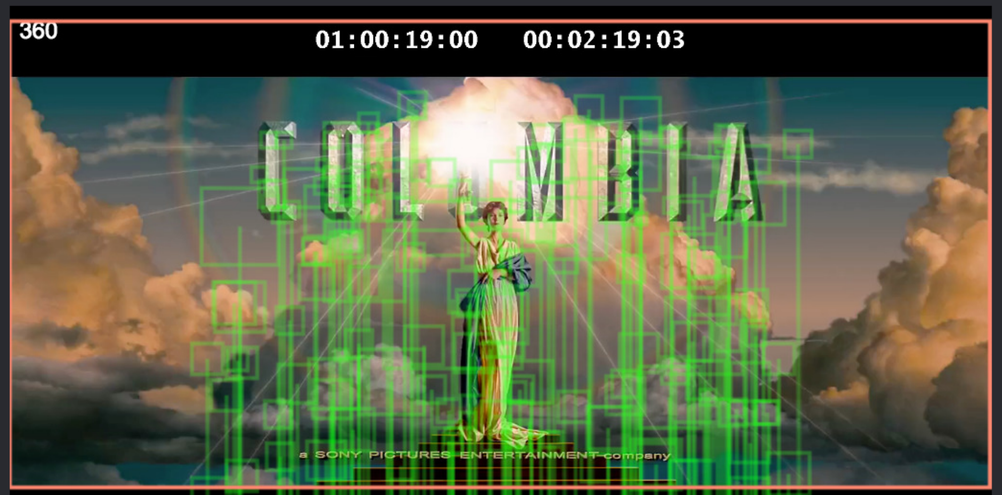
Systems and methods for image analysis and recognition are disclosed, in particular the methods for interest point description. An interest point and its surrounding area is broken into subareas, a frequency domain description of each area is created by applying discrete Fourier transform (DFT). Frequency domain features are than coded bitwise by comparing them to predefined thresholds. Subsequently, the present invention provides alternative or improved methods and data structures for interest point description that may reduce memory consumption and allow fast bitwise matching.

Example of matched pairs of features

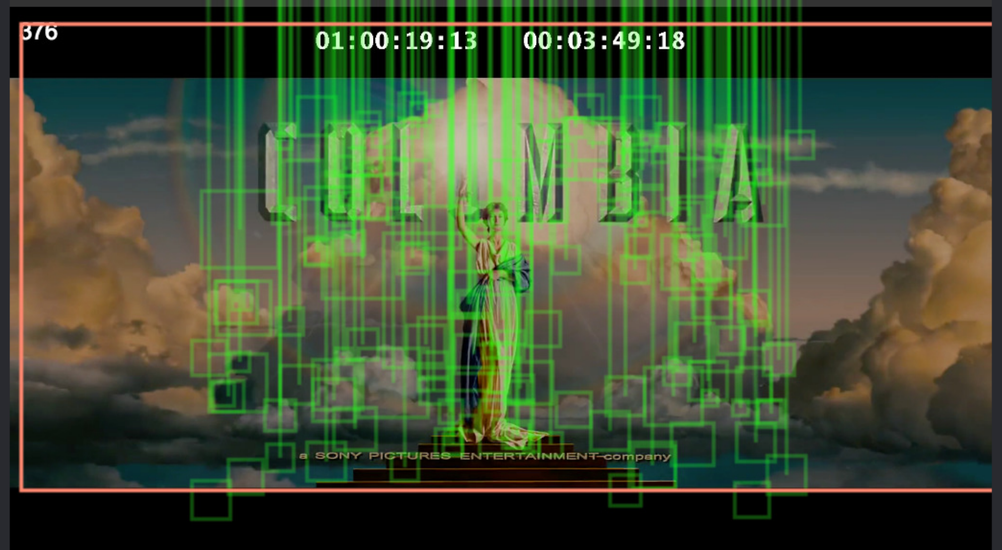
US20130243341A1.pdf

theatrical 3000 frames

directors
3000
frames

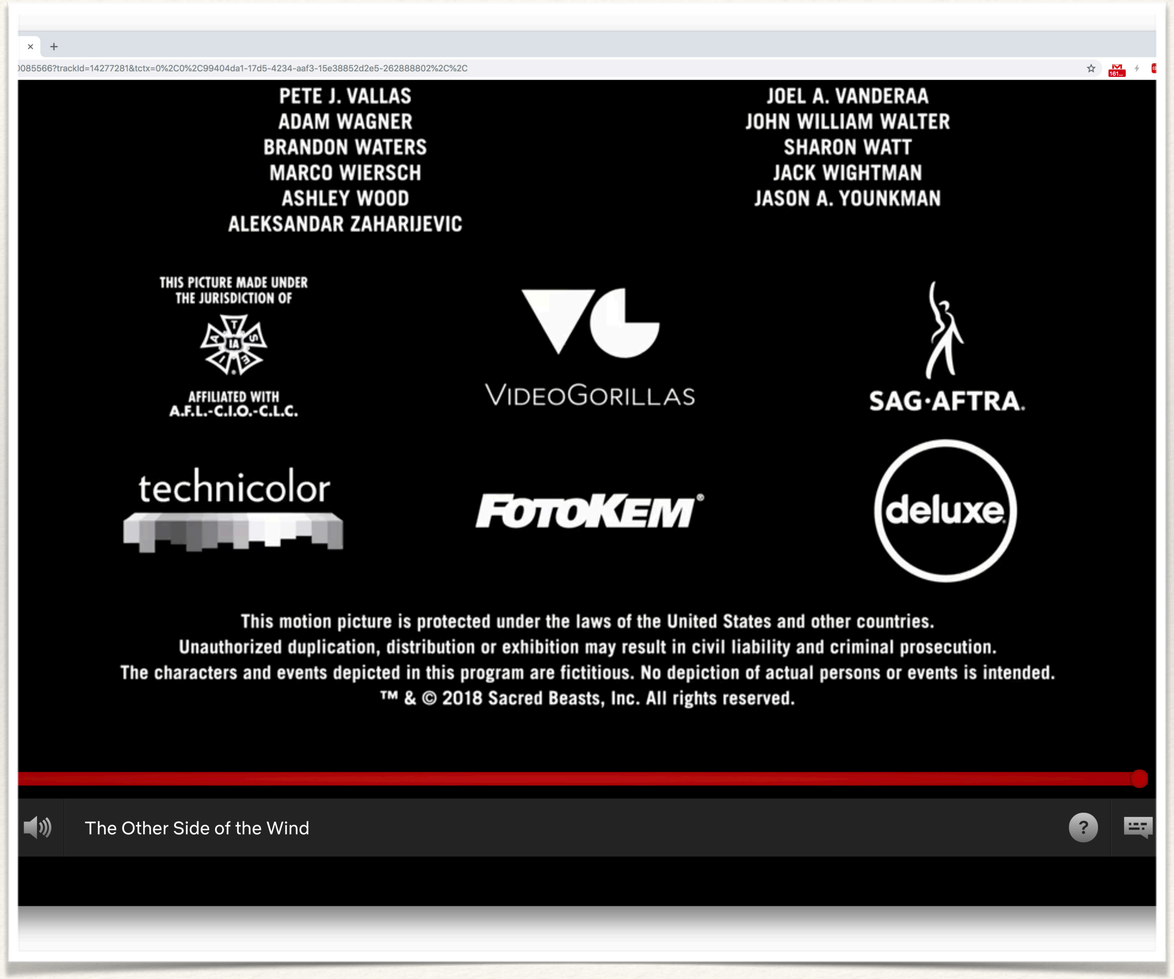


181
113 matches, full frame desc distance: 15 chunks: 3x3 mindist (0)
181



Netflix - Edit Decision Reverse Engineering

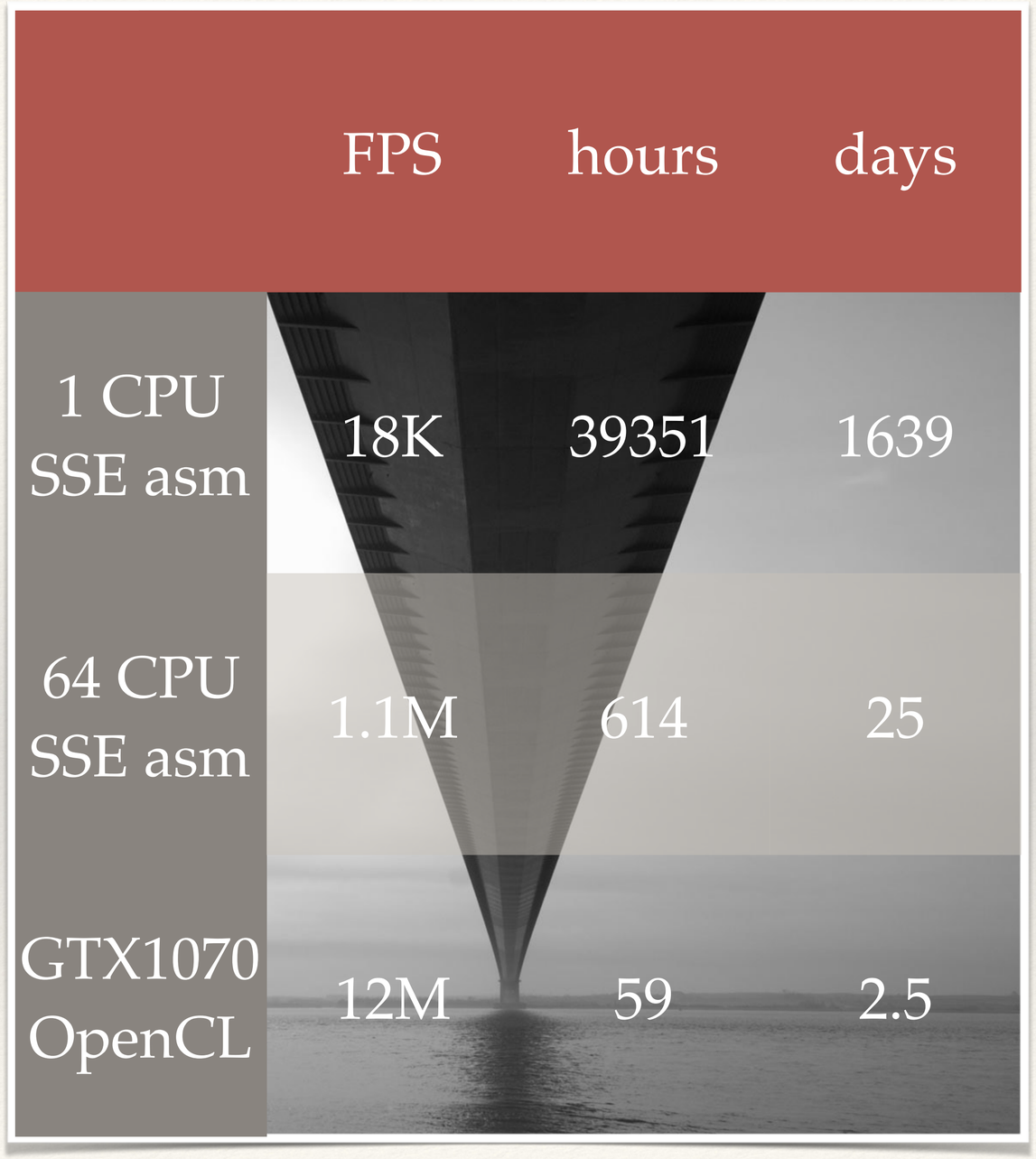
- 300K frames edit from Orson
- 8.5M frames of 4K scans
- No Edit Decisions (EDL)
- We have a binary and a bunch of lines of code, decompile binary extract source code reverse engineer Makefile
- 2.5T (trillion) frames to compare
- Run diff on 13K files
- Manually 5% in 9 months



CPU vs GPU

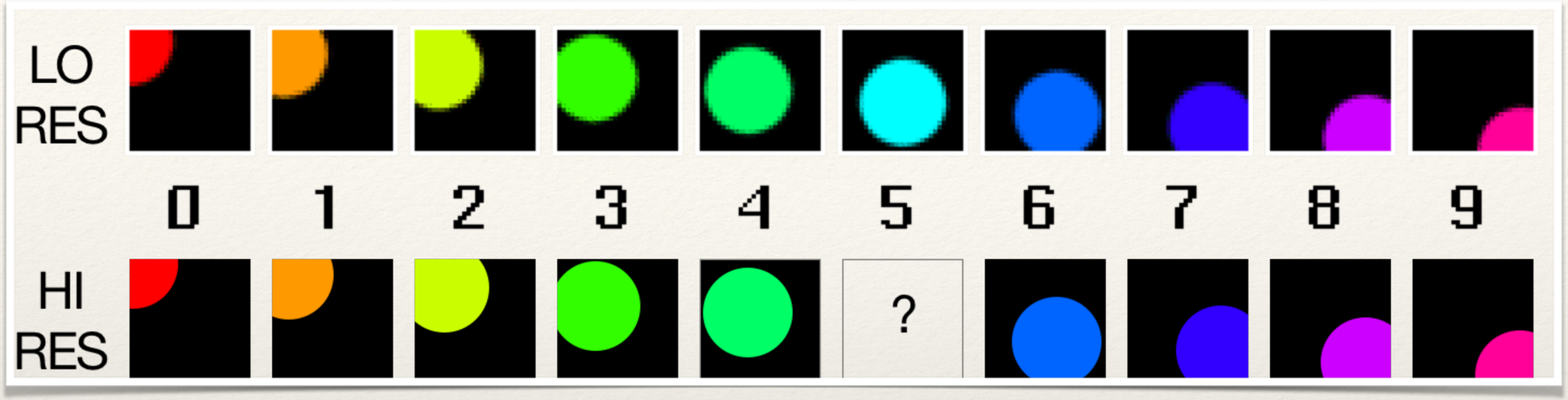
Performance

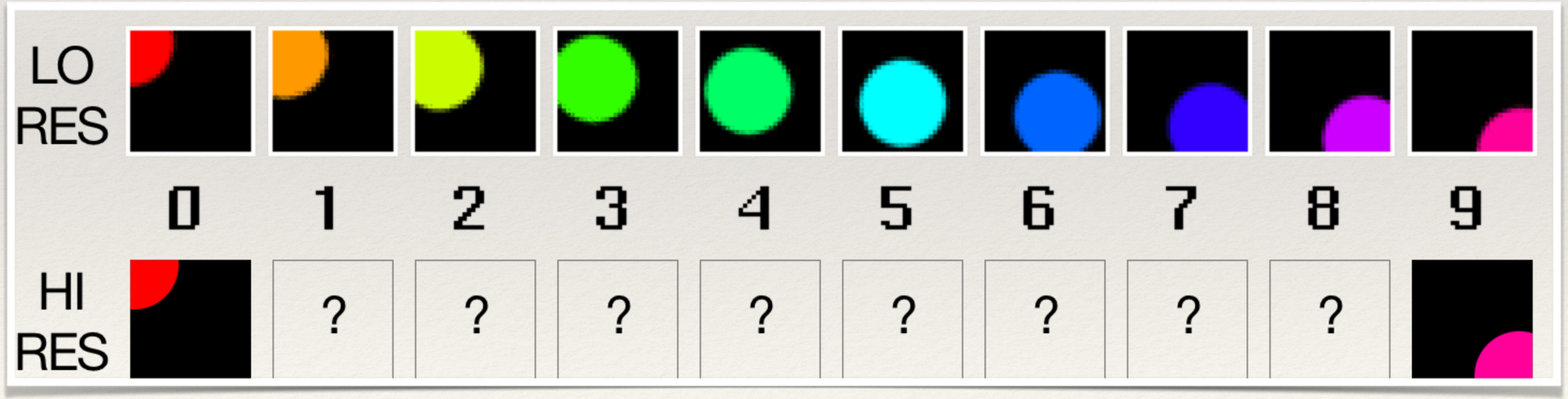
Rent 25 64 core servers for 24 hours
OR
Rewrite in OpenCL



What's next?

- Super resolution
- step1 infer 10% of a movie from 90%
- step2 infer 90% of a movie from 10%
- step3 super res 100%


LO RES
0 1 2 3 4 5 6 7 8 9
HI RES
?


LO RES
0 1 2 3 4 5 6 7 8 9
HI RES
? ? ? ? ? ? ? ?

Alex Zhukov

VideoGorillas