



VirginiaTech

Ted and Karyn **Hume Center** for National Security and Technology &
SEEC Center: Synergistic Environments for Experimental Computing

Building a General Search Engine for Unstructured Data

Harold Trease

Senior Research Scientist

ICTAS HUME & SEEC Centers & Computer Science Department

Virginia Tech



STEM
Education



Signals
Intelligence



Electronic
Warfare



Resilient
Systems



Data to
Decision

Why Build a Search Engine?

- Lots of data that isn't currently searchable because there is no metadata associated with it
 - image, (noisy) audio, video, communications, IoT
- Multi-INT Big Data -- The majority of unstructured data is “not” text
 - 88% Image/Video, 10% Audio/Communications, 2% Text
 - Video consumes ~75% of Internet bandwidth today (~85 - 90% 2018, Cisco)
- Help analysts/users find the information they want (actually, we provide pointers to information):
 - Google search for any type of digital data
- Multi-Sensor data fusion
- Making use of the structure of data (i.e., galaxy plots)
- Search for people, places, and things including their interactions (i.e., graphs)

Search Engine Technology

- Algorithms and Data Management:
 - Ingest: Unstructured data, compressed. Parallel I/O
 - Signature generation: Transform unstructured data into signatures
 - Signature comparison: Pairwise comparison of “unknown” signature with “known” signatures (database)
 - Connecting signatures into graphs
 - Results and Metrics: Not designed to be 100% accurate, but uses high probabilities
 - User Interfaces and user interaction
- Optimization:
 - Algorithms
 - Hardware and Code Performance
 - Usability
- Hybrid, Heterogeneous Parallel Computing
 - MPI, Hadoop/MapReduce
 - OpenMP, Cuda, OpenCL

Searching for Objects in Image/Video

Two typical frames from cell phone videos



⇐ Search for these objects, called “search criteria”.
Wu, Hao, Kaixi, purple frisbee, building

Search for Wu: http://198.82.148.84/datafission/results/synergy_wu/html_search/

Search Analyze Results:

Source:

/lustre/www-data/mediaData/processedData/Synergy/Wu_01.JPG/Wu_01.JPG



Targets:

- /lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4
- /lustre/www-data/mediaData/processedData/Synergy/20140815_161444.mp4/20140815_161444.mp4

Search Criteria
Searchable Corpus
Search Results:

- Full frame results (ranked)
- Chipout results (ranked)
- Results with metadata
- File, time, frame #

Table of top search frame results: 50

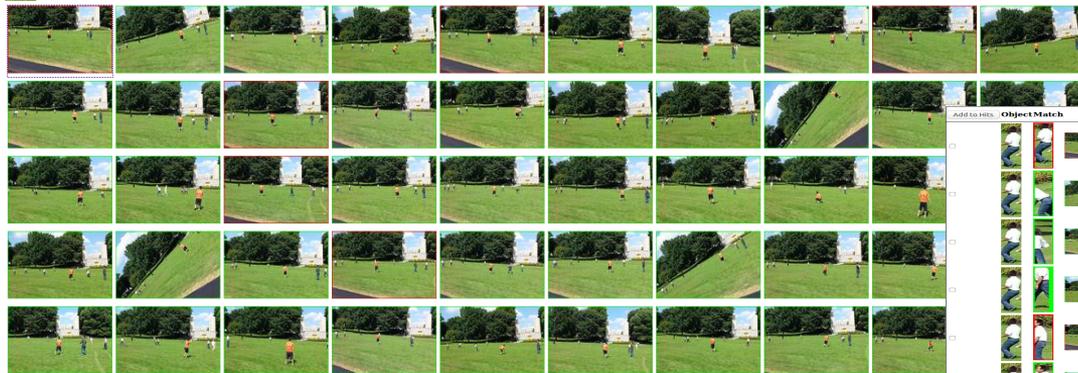
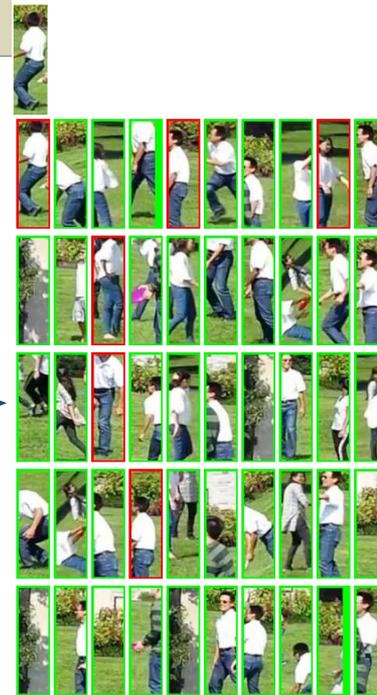


Table of top search OOI results: 50



Add to hits	Object Match	Frame	Segment	Rank	Time
<input type="checkbox"/>			0001:#22.00#0.00#32.44#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161444.mp4/20140815_161444.mp4#toc;file/frame-000006051.gpm	
<input type="checkbox"/>			0002:#96.52#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4#toc;file/frame-000005926.gpm	
<input type="checkbox"/>			0003:#130.85#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4#toc;file/frame-000005986.gpm	
<input type="checkbox"/>			0004:#115.19#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4#toc;file/frame-000005421.gpm	
<input type="checkbox"/>			0005:#20.31#0.00#32.44#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161444.mp4/20140815_161444.mp4#toc;file/frame-000006001.gpm	
<input type="checkbox"/>			0006:#48.52#48.52#50.60#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4#toc;file/frame-000004441.gpm	

Search for a building/place: http://198.82.148.84/datafission/results/synergy_building/html_search/

Source:

/lustre/www-data/mediaData/processedData/Synergy/Building_01.JPG/Building_01.JPG



Targets:

-  /lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4
-  /lustre/www-data/mediaData/processedData/Synergy/20140815_161444.mp4/20140815_161444.mp4

Search Criteria
Searchable Corpus
Search Results:

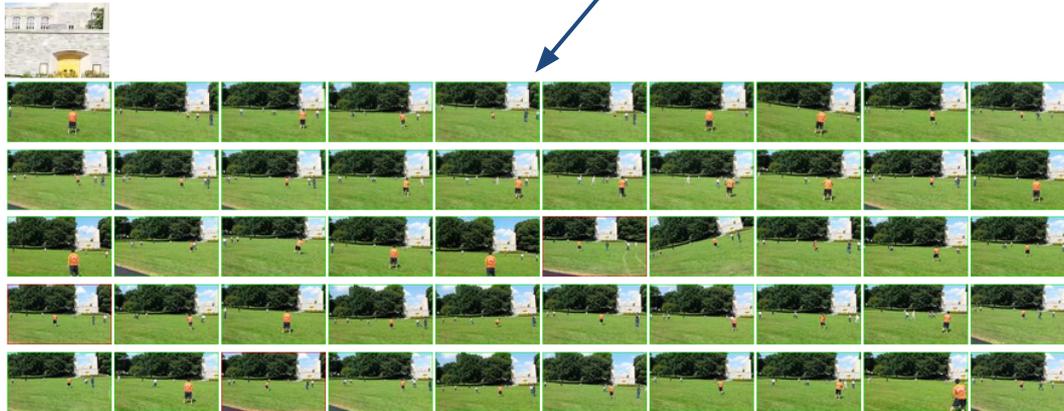
- Full frame results (ranked)
- Chipout results (ranked)
- Results with metadata
- File, time, frame #



Table of top search OOI results: 50



Table of top search frame results: 50



Matches	Object	Match	Frame	Segment	Rank	Time
0					0001 #10.0#0.00#12.80#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000000316.ppm
0					0002 #79.83#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000002371.ppm
0					0003 #42.63#18.83#46.69#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000001266.ppm
0					0004 #40.03#48.52#59.60#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000001436.ppm
0					0005 #91.90#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000002731.ppm
0					0006 #80.04#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000002491.ppm
0					0007 #17.37#12.83#46.69#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000000516.ppm
0					0008 #93.63#0.00#12.80#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000000286.ppm
0					0009 #55.09#48.52#59.60#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/_huc_file/frame.000001636.ppm

Search Analyze Results:

Source:

/lustre/www-data/mediaData/processedData/Synergy/Purple_frisbee.JPG/Purple_frisbee.JPG



Targets:

- /lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4
- /lustre/www-data/mediaData/processedData/Synergy/20140815_161444.mp4/20140815_161444.mp4



Search Criteria
 Searchable Corpus
 Search Results:

- Full frame results (ranked)
- Chipout results (ranked)
- Results with metadata
- File, time, frame #

Table of top search OOI results: 50

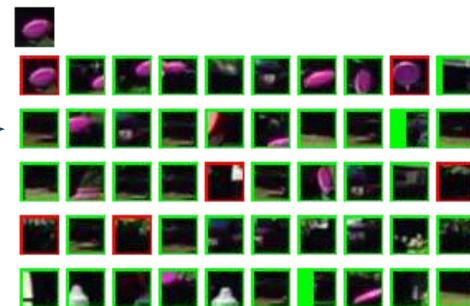
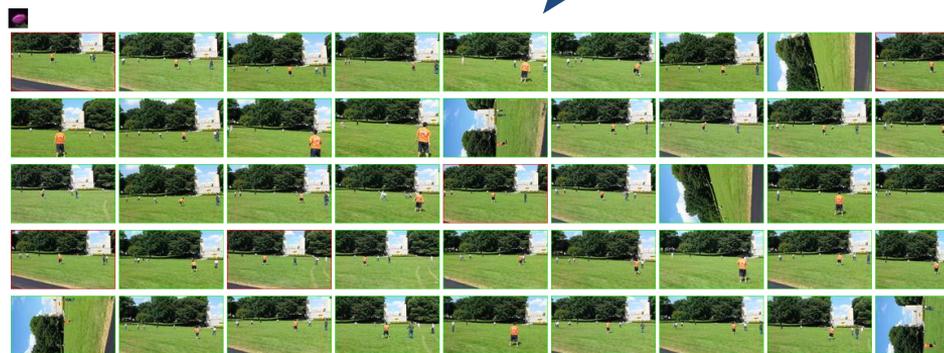


Table of top search frame results: 50



Add to Hits	Object Match	Frame	Segment	Rank	Time
<input type="checkbox"/>				0011:#22.00#00#32.44#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161444.mp4/20140815_161444.mp4/toc_files/frame.00000963.ppm
<input type="checkbox"/>				0002:#136.91#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/toc_files/frame.000004066.ppm
<input type="checkbox"/>				0003:#117.21#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/toc_files/frame.000003481.ppm
<input type="checkbox"/>				0004:#84.38#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/toc_files/frame.000002506.ppm
<input type="checkbox"/>				0005:#34.55#12.83#48.49#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/toc_files/frame.000001026.ppm
<input type="checkbox"/>				0006:#61.65#59.63#72.73#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/toc_files/frame.000001831.ppm
<input type="checkbox"/>				0007:#116.20#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/toc_files/frame.000003451.ppm
<input type="checkbox"/>				0008:#104.08#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/toc_files/frame.000003091.ppm
<input type="checkbox"/>				0009:#30.45#00#32.44#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161444.mp4/20140815_161444.mp4/toc_files/frame.00000961.ppm
<input type="checkbox"/>				0010:#78.82#72.76#148.93#	/lustre/www-data/mediaData/processedData/Synergy/20140815_161203.mp4/20140815_161203.mp4/toc_files/frame.000002341.ppm

Google Image .vs Bing Image vs. VT Search Engine

Search Query



Google Image Search

Google

Web **Images** News Shopping Maps More Search tools

About 31 results (0.97 seconds)

 Image size: 132 x 218
No other sizes of this image found.

Best guess for this image: [abu musab al zarqawi](#)

Abu Musab al-Zarqawi - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Abu_Musab_al-Zarqawi
pronunciation (help·info) **Abū Muṣ'ab az-Zarqāwī**, **Abu Musab** from Zarqā; October 30, 1966 – June 7, 2006), born Ahmad Fadeel al-Nazal al-Khalayleh ...

The Short, Violent Life of Abu Musab al-Zarqawi - The Atlantic
www.theatlantic.com/magazine/archive/2006/07/the...of.../304983/
Jul 1, 2006 - He would later rename himself **Abu Musab al-Zarqawi**. ... Both Huthaifa Azzam and al-Zarqawi would eventually leave Afghanistan to pursue ...

Visually similar images Report images



Bing Image Search

WEB **IMAGES** VIDEOS MAPS NEWS MORE



We couldn't find any matches for this image.

Search tips:

- Please use a valid image file. The image can be a .gif, .jpg, .bmp or .png file type.
- The image must be 10MB or smaller.

Table of top search frame results: 50



VT Search



Google Image Search vs. VT Search Engine

Search Query



Google Image Search

Google **Als_nose.JPG** describe image here

Web **Images** News Shopping Maps More Search tools

Image size: 54 x 80
No other sizes of this image found.

Tip: Try entering a descriptive word in the search box.

Visually similar images [Report images](#)



Bing Image Search

WEB **IMAGES** VIDEOS MAPS NEWS MORE

bing   Image Match

We couldn't find any matches for this image.

Search tips:

- Please use a valid image file. The image can be a .gif, .jpg, .bmp or .png file type.
- The image must be 10MB or smaller.

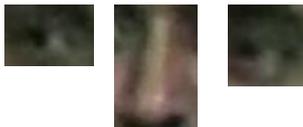
VT Search

Table of top search frame results: 50



Compound Search (i.e., more than one search criteria)

Search
Query



Google Image Search: (Google doesn't support multiple image search criteria)

The screenshot shows a Google search for 'Als_right_eye.JPG'. The search bar contains the filename and a search icon. Below the search bar, the 'Images' tab is selected. A single image result is shown with a size of 44 x 30 pixels. A tip suggests entering a descriptive word, and a section for 'Visually similar images' displays a grid of related image thumbnails.

(compound) VT Search

Search Analyze Results:

Source:

/lustre/www-data/mediaData/processedData/Al/Als_left_eye.JPG/Als_left_eye.JPG

/lustre/www-data/mediaData/processedData/Al/Als_nose.JPG/Als_nose.JPG

/lustre/www-data/mediaData/processedData/Al/Als_right_eye.JPG/Als_right_eye.JPG



Targets:

/lustre/www-data/mediaData/processedData/Al/AlZahari.mp4/AlZahari.mp4

Table of top search frame results: 50





Rank 1 & 2

Search for
RB #22
“and”
the football



Rank 3 & 4



Rank 9 & 11



Rank 15 & 16

3 COOLs



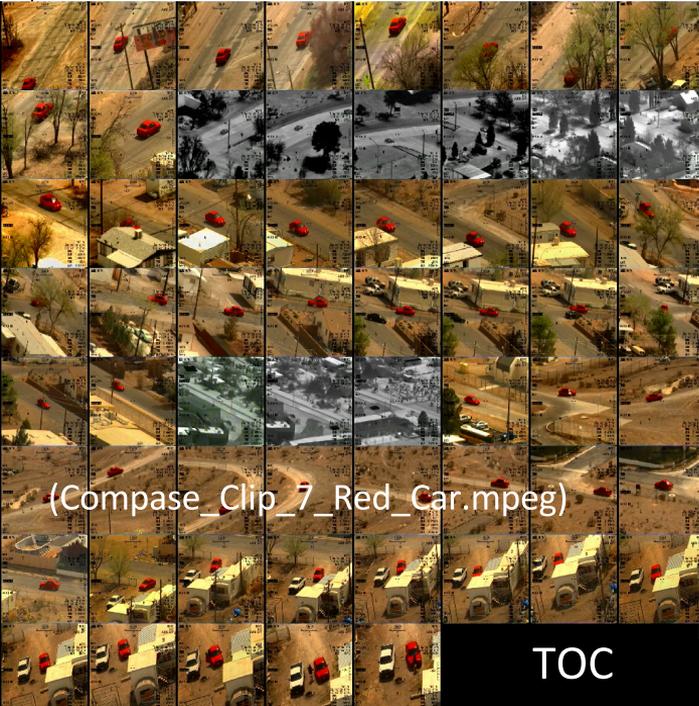
("and" operator)

Pattern-of-Life: Have three vehicles ever been in the same scene at the same time?



("and" operator)

Search for 3 cars using Boolean "and" /"or" operators



(Compass_Clip_7_Red_Car.mpeg)

TOC

Source: /lustre/www-data/video11/sun/White_Dodge_Pickup_Parked.JPG

Target: /lustre/www-data/video11/sun/./slave_video_default/pc.001.000017/f

Rank: 1	SAD: 000.84	source: /lustre/www/White_Dodge_Pick/White_Dodge_Pick/White_Dodge_Pick
Rank: 2	SAD: 000.76	source: /lustre/www/White_Dodge_Pick/White_Dodge_Pick/White_Dodge_Pick
Rank: 3	SAD: 000.51	source: /lustre/www/White_Dodge_Pick/White_Dodge_Pick/White_Dodge_Pick
Rank: 4	SAD: 000.84	source: /lustre/www-data/video11/summaries/Demo_Set_5
Rank: 5	SAD: 001.04	source: /lustre/www/White_Dodge_Pick/White_Dodge_Pick/White_Dodge_Pick
Rank: 6	SAD: 000.64	source: /lustre/www/White_Dodge_Pick/White_Dodge_Pick/White_Dodge_Pick
Rank: 7	SAD: 001.04	source: /lustre/www/White_Dodge_Pick/White_Dodge_Pick/White_Dodge_Pick
Rank: 8	SAD: 001.11	source: /lustre/www/White_Dodge_Pick/White_Dodge_Pick/White_Dodge_Pick
Rank: 9	SAD: 001.11	source: /lustre/www/White_Dodge_Pick/White_Dodge_Pick/White_Dodge_Pick



("or" operator)



("or" operator)

State Farm Logo:



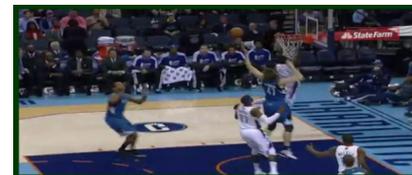
Search for the StateFarm logo in a set of 2012 and 2013 basketball highlight videos

Search Results:

Table of top search frame results: 50



Table of top search OOI results: 50



Data ==> Signatures ==> Database ==> Search Engine

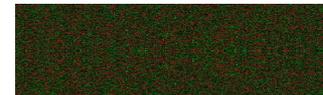
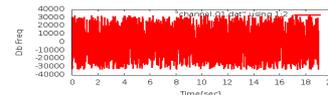
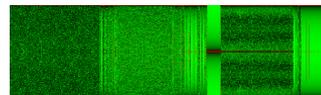
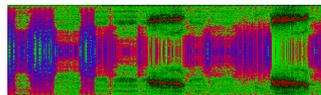
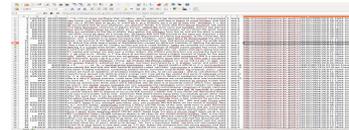
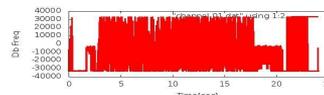
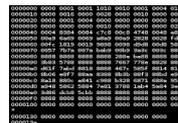
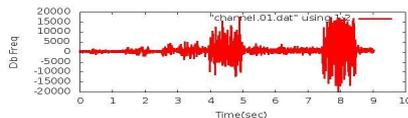
Image: Web, mobile devices

Video: UAVs, Web/IP cameras, Youtube

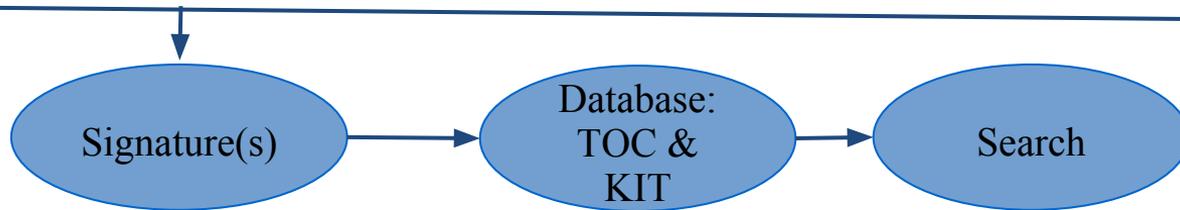
Audio: acoustic sensors, mobile devices

Binary Data: sensors, harddrives, cell phones

Text Documents: logs, chat/email



Basic Abstraction: Generalized pattern search.
Everything is represented as a signal/image.



Signatures: Quantitative Information Representation

What do signatures look like?

Statistical Data Analysis



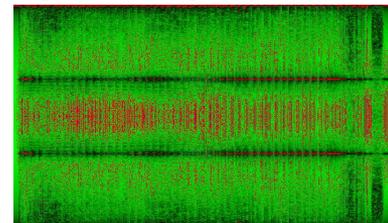
Shannon Entropy [P]log(P)



Spatial Frequencies (DoL)



Spectral Frequencies (FFT)



80-D Signature Feature Vector [4 transforms * 4 spectral * 5 statistical moments]

20-D Native				20-D Entropy				20-D Spatial Frequency				20-D Spectral Frequency			
Grey	Red	Green	Blue												
Five Statistical Moments															

Signature Dope Vector: 151 80 V:20#E:20#S:20#F:20# 66.45 57.47 0.65 2.43 4.02 91.99 91.18 0.69 1.98 2.59 55.02 51.40 1.02 3.72 9.03 53.25 50.39 1.20 4.27 11.36 36.75 61.83 3.07 11.04 38.90 59.85 99.02 1.37 2.93 5.45 40.77 77.17 2.37 6.77 18.72 40.97 80.35 2.27 6.21 16.47 17.53 31.36 3.02 12.97 58.86 18.58 33.76 3.04 13.26 61.06 17.10 31.50 3.05 13.15 60.22 17.81 30.44 2.92 12.50 57.12 40.80 32.62 0.05 1.21 0.14 107.12 111.46 0.13 1.10 0.35 16.22 25.14 2.02 8.35 40.13 0.02 1.77 63.62 4052.87 258199.4 2

Signatures: Why these components? Necessary/Sufficient?

Statistical Data Analysis



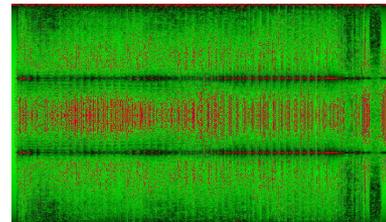
Shannon Entropy [P]log(P)]



Spatial Frequencies (DoL)



Spectral Frequencies (FFT)



80-D Signature Feature Vector [4 transforms * 4 spectral * 5 statistical moments]

20-D Native

20-D Entropy

20-D Spatial
Frequency

20-D Spectral
Frequency

Grey

Red

Green

Blue

Grey

Red

Green

Blue

Grey

Red

Green

Blue

Grey

Red

Green

Blue

Five
Statistical
Moments

Statistics: If two signatures are going to match, then the statistical moments "necessarily" have to be similar.

Entropy: Entropy of man-made vs. natural objects is a discriminator. Natural objects tend to have higher entropy. Man made objects tend to have lower entropy because of (unnatural) uniformity.

Spatial frequencies: Edges, curvature, corners are important discriminators for the human vision system.

Spectral frequency: Used as a discriminator, even though the audible frequency distribution may not mean anything to a human or DSP algorithm.

Signatures: Computational Kernels

Statistical Data Analysis



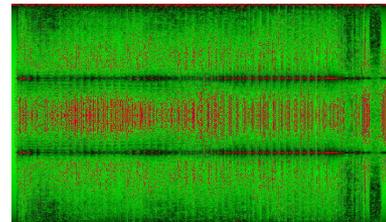
Shannon Entropy $[P()]\log(P())$



Spatial Frequencies (DoL)



Spectral Frequencies (FFT)



80-D Signature Feature Vector [4 transforms * 4 spectral * 5 statistical moments]

20-D Native

20-D Entropy

20-D Spatial
Frequency

20-D Spectral
Frequency

Grey

Red

Green

Blue

Grey

Red

Green

Blue

Grey

Red

Green

Blue

Grey

Red

Green

Blue

Five
Statistical
Moments

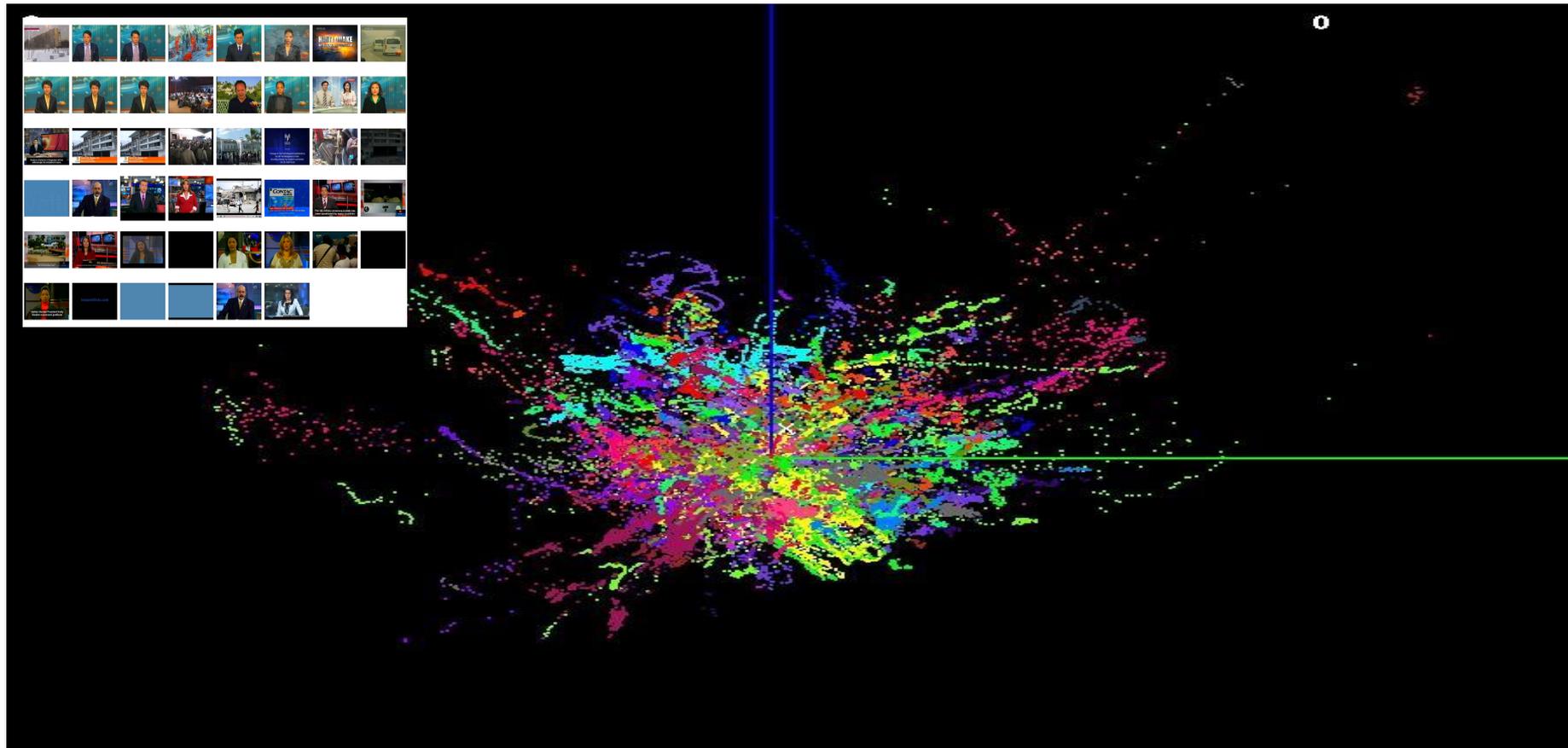
Kernel 1: Mean
Kernel 2: Generating function
for variance, skew, kurtosis,
hyper-skew

Kernel 3: Histogram and
Histogram to $P()$ normalization
Kernel 4: $P()]\log(P())$

Kernel 5: Difference-of-
Laplacians (DoL) edge
detection filter. 5-Point / 9-
Point Stencil

Kernel 6: 1-D FFT

Youtube Videos: Media Space vs. Search Space



Exploiting the Structured of Data

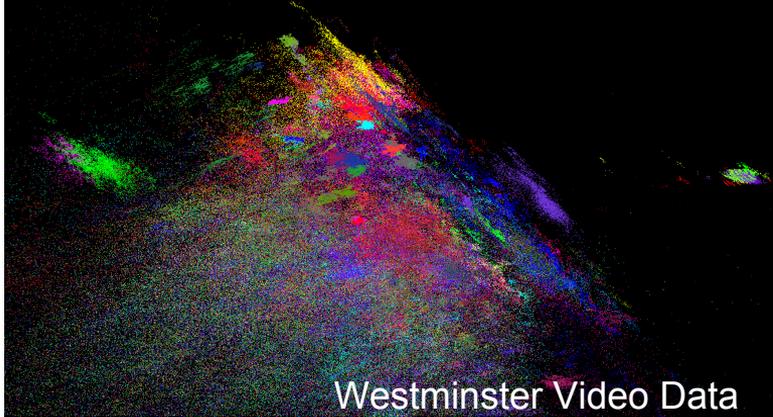
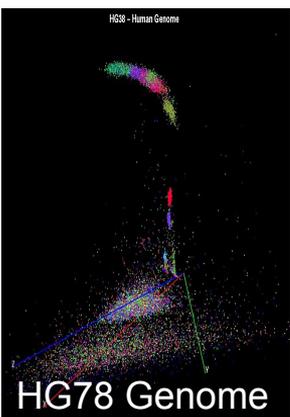
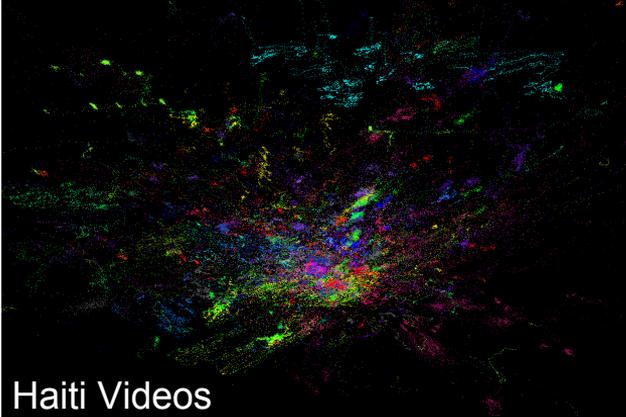
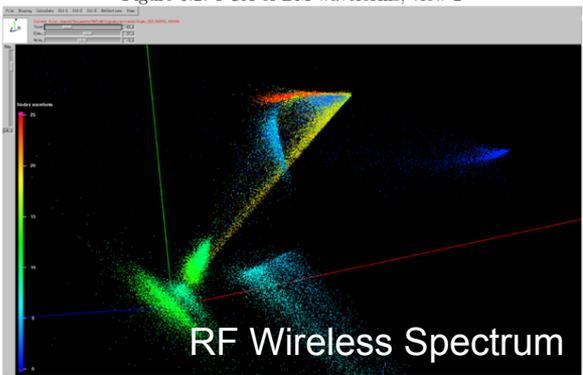
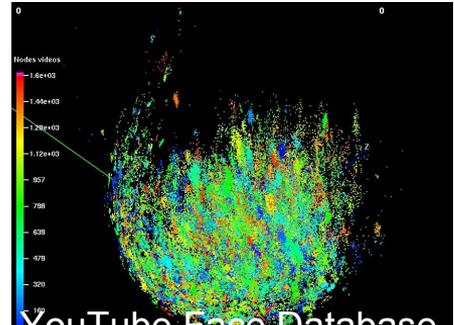
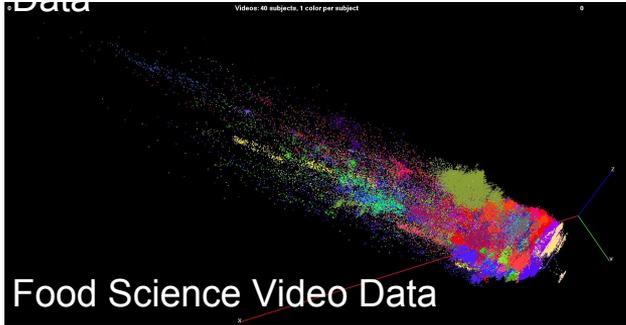
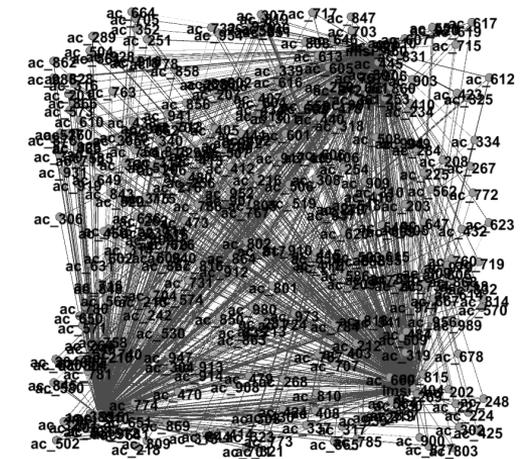
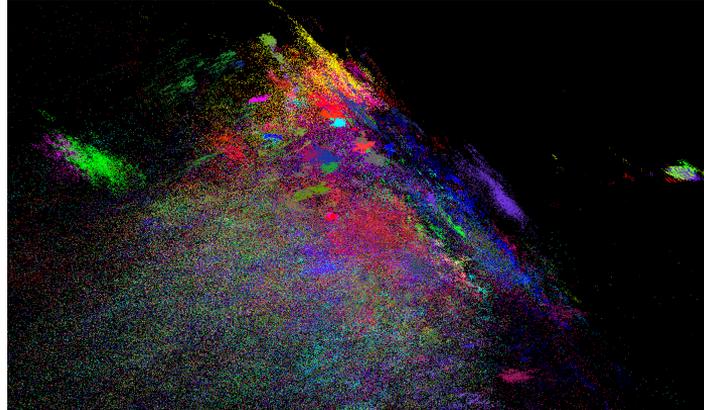
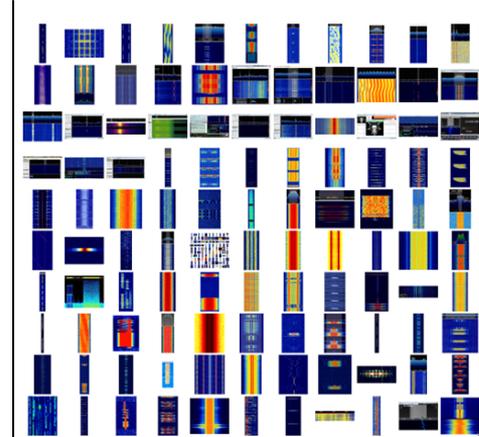


Figure 5.2: PCA of BM waveforms, view 2



Visualization of Data: classification, clustering, graph analysis, summarization/search results



Source: Analyze Results:

Source:

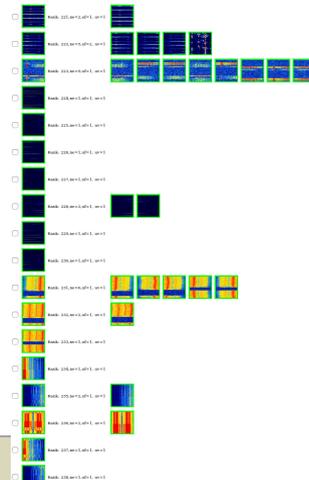
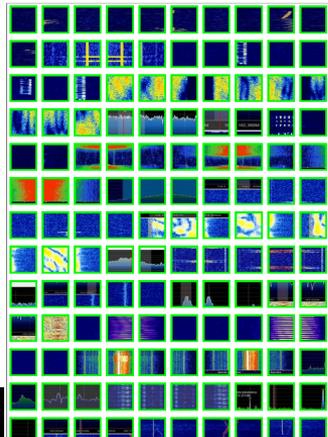
Country: www.databasesearch.com/identifiers/2019img.jpg

Keywords:

Table of top search Name results: 50

Table of top search ORG results: 50

Matrix of Keywords: Click on the following images to extract GOIs: #keywords=911/#entities=1712



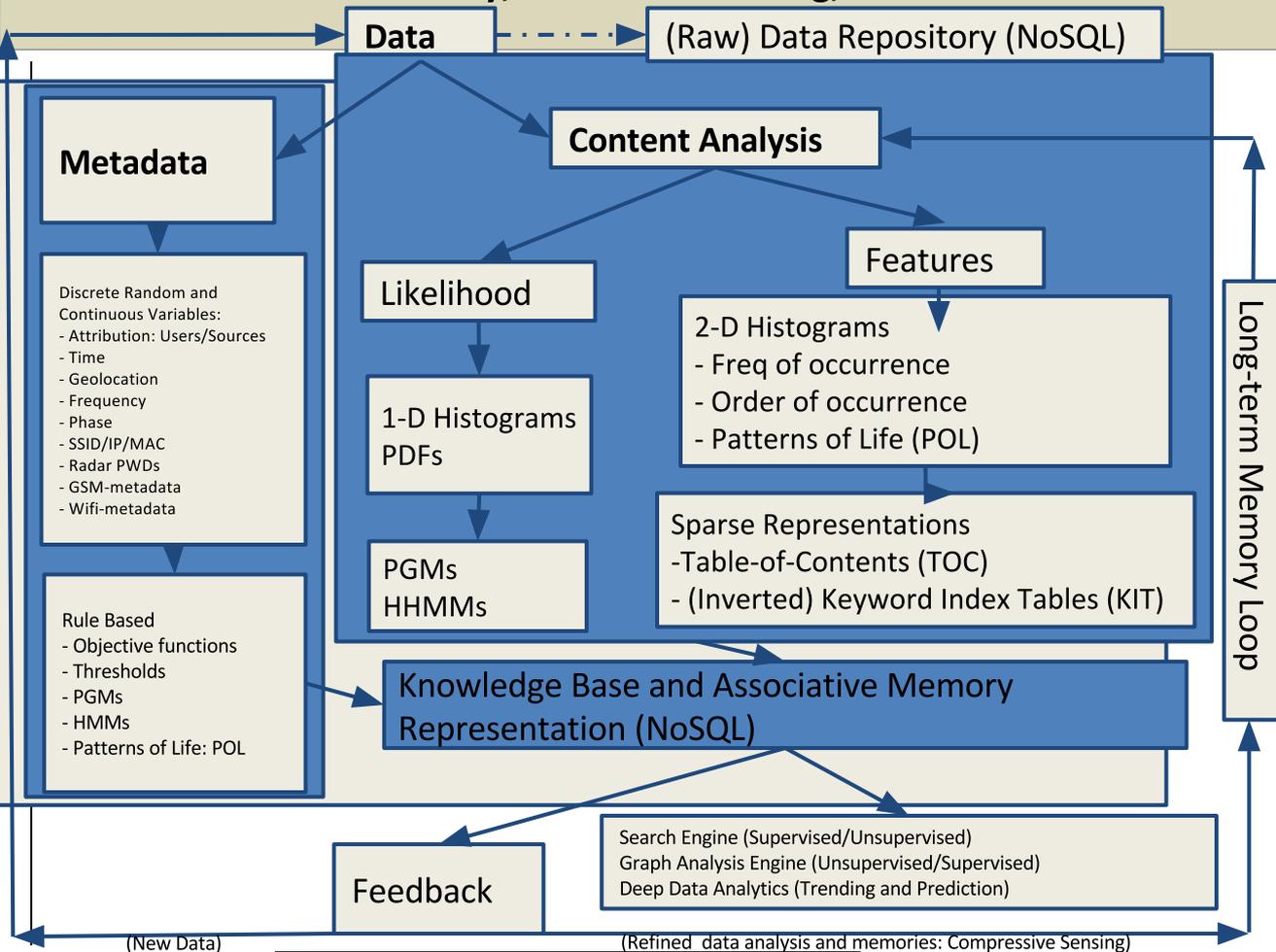
IMSI Blacklist: 10 310
 IMSI Blacklist: 10 404
 IMSI Blacklist: 10 460

Area Code Blacklist: 10 540
 Area Code Blacklist: 10 402
 Area Code Blacklist: 10 509

AC dict entry: 402 31890 402 P Lincoln_and_Omaha_Nebraska
 AC dict entry: 509 31830 509 P Pullman_Spokane_and_Walla_Walla_Washington
 AC dict entry: 540 31020 540 E Fredericksburg_Roanoke_and_Winchester_Virginia

IMSI dict entry: 310 44810 310 390 us **United_States_1** Yorkville_Telephone_Cooperative
 IMSI dict entry: 404 44380 404 30 in **India_91** Usha_Martin_Telecom
 IMSI dict entry: 460 44650 460 01 cn **China_86** China_Unicom

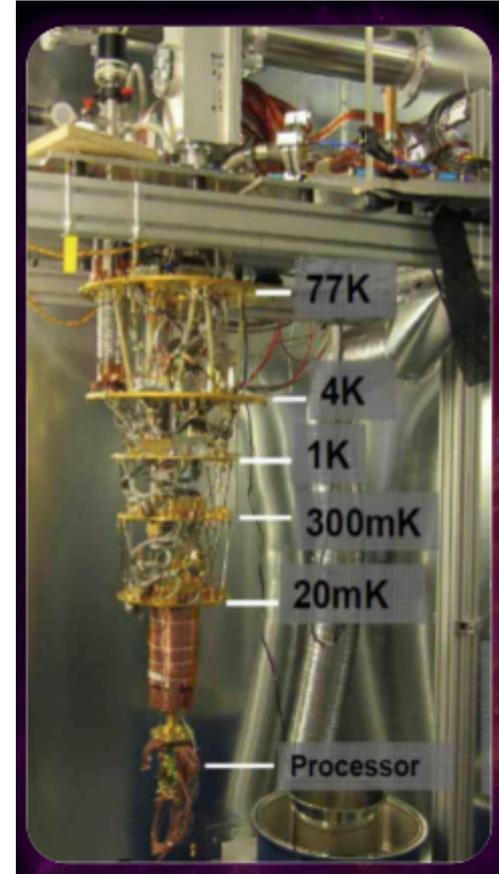
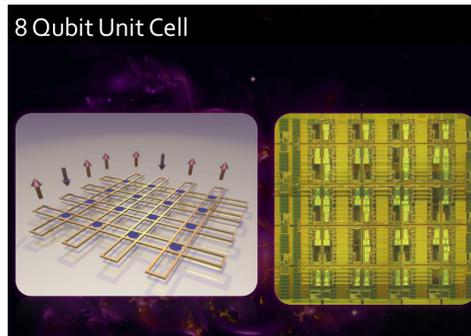
Associative Memory, Machine Learning, Feedback



- Sources: Getting data**
 - Sensor networks
 - Prototypical field data
 - Simulated data (more control over SNR, format, volume)
- Database: Storing data**
 - Streaming vs. archive
 - Ingest, codecs, multi-INT
 - Ingest and indexing/summarization: image, video, audio, sensor data, binary data, text (Transformation into N-dimensional signatures)
 - Ingest and DSPs: GNU Radio
- Knowledge Base: Storing things**
 - Hierarchical associative memories: Remembering things
 - Sparse Representations
 - PDFs (e.g., histograms)
 - NoSQL, SQL, Triple/quad stores
 - Hierarchical representation: subdivision and aggregation of entities
 - Types of metadata: Time, location, mode, modality, frequency of occurrence, order of occurrence
 - Modal metadata: Image, Video, Audio, WiFi, GSM, Radar, etc.
- Search: Finding things.**
 - Supervised/Unsupervised search and graph analysis
 - Content vs. metadata
- Interaction Graphs: Interrelating things**
 - Links, associations, connections, relationships, networks
 - Relating people, places, and things to events, activities
 - Patterns-of-Life
 - Context and Situational Awareness:
 - Metadata (time, (geo)location, attribution)
- Conflict Resolution: Resolving conflicting information**
 - Lack of conflict resolution leads to false positives
- Trending and Prediction:**
 - Big Data vs. Models vs. Deep data analytics
- Belief Propagation: Inferring things**
 - Graph Analysis:
 - What/who is around (centrality)
 - What/who points to me (pagerank)
 - PGMs, Bayesian networks
- Testing and evaluation: V&V&UQ**
 - Real data (Validation)
 - Synthetic data (Validation)
 - Ideal data (Verification)
 - Uncertainty quantification
 - How good is the information
 - Accuracy of the solutions

D-Wave Adiabatic Quantum Computer

- Computes all results at once, reads out one
- Solutions in about 100msec
- Crudely similar to simulated annealing
- Express problem as a QUBO (Quadratic Unconstrained Binary Optimization)
- Transformed into a QMI (Quantum Machine Instruction)
- Results read out as a binary bit vector
- Application areas:
 - Shor's Algorithm (exponential speedups)
 - Prime number factorization
 - Grover's Algorithm (SQRT(N) speedup)
 - Unstructured search
- Application areas:
 - Traveling salesman problem
 - Knapsack problem
 - Machine learning



D-Wave ToQ source code example

Simple program (sample.toq)

- # --- sample.toq --- Regression test
- #
- bool: @a, @b
- bool: @c, @d
- #~~~
- assert : And(Or(@a,@b), Or(@c,@d))
- assert: @c != @d
- assert: (@a+@d >= (@b+@c))
- #~~~
- end:

```
4 Results (400 cases)
(443 occurrences)
=====
# Variable
# Name
1) @a 1 0 1 1
2) @b 0 1 0 1
3) @c 0 0 1 0
4) @d 1 1 0 1
-----
Occurrences' 327 69 42 5
' Enter "toq -QM" for an explanation
of why (Sum(Occurrences) != cases)
```

sample2.toq

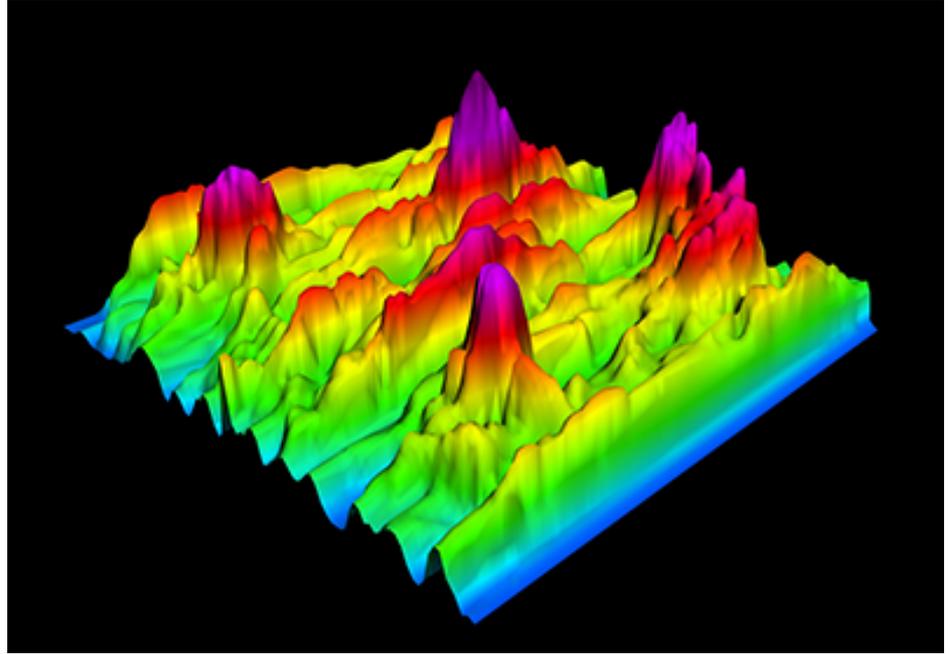
```
# --- sample2.toq --- Regression test
#
bool: @a, @b
bool: @c, @d
real: x, y, z
#~~~
x = pi
y = sqrt(x+2)
z = x-y
assert: And(Or(@a,@b),Or(@c,@d))
if: ((z*x)/3 > 3.11)
    assert: @c != @d
endif:
assert: (@a+@d >= (@b+@c))
#~~~
printvars:
end:
```

sample2.toq (results)

```
4 Results (400 cases)
(471 occurrences)
=====
# Variable
# Name
1) @a 1 1 0 1
2) @b 0 0 1 0
3) @c 0 1 0 1
4) @d 1 0 1 1
-----
Occurrences' 340 89 41 1
' Enter "toq -QM" for an explanation
of why (Sum(Occurrences) != cases)
```

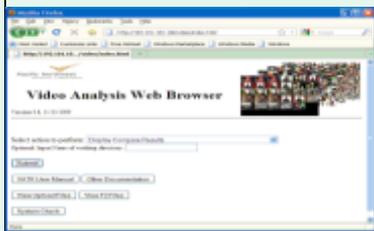
Unstructured Search on the D-Wave Adiabatic Quantum Computer

- Relax a surface onto the surface of “unknown” signatures minus “database” signatures



Virginia Tech Search Engine (VTSE) Platform Architecture

Web Browser



HTTP GUI
Client Interface

Mobile Browser/Apps



HTTP GUI
Client Interface

Rich Client

REST Web Client
Interface

Client/Server Interoperability

M2M Client-Server Interfaces:

- Build using LAMP
- HTTP GUI
- RESTful C++/Curl, Python, Java, and PHP
- API Client-Server Interface for PHP and C++
- Inherits the security domain from client
- Deployable to sensitive environments

Data-Services:

- Data from clients
- Results from database
- Bulk data loads

Server

Python API Server:
- Index&Summarize
- Compare
- Find/Search
- Tracking
- Cleanup

Video Server:
- Algorithm
- TOC, Index, Search
- Trending/Deep
 Analytics/Prediction
- Parallel Computing
- Data Management

(Signatures, TOC, KIT)

DataBase Management:

- MySQL (metadata)
- MySQL (results)
- HyperTable/Hadoop/HDFS