Biometric Evaluation on the Cloud: A Case Study with HumanID Gait Challenge

Ravi Panchumarthy, Ravi Subramanian, and Sudeep Sarkar
Computer Vision and Pattern Recognition Group
Dept. of Computer Science & Engineering,
University Of South Florida, Tampa, FL-33620
Biometric Algorithm Evaluation

• Issues with Datasets
  – Very large (~TB)
  – Difficult to create and share

• Issues with current practice of reporting results
  – Compute intensive
  – Cannot enforce consistency of protocols
  – Isolated and disorganized
  – Cannot perform deep comparison other than based on aggregate measures (e.g. detection rate, false alarm)
## Existing Major Gait Databases

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th># of Subjects</th>
<th>Variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gait Challenge</td>
<td>1.2 TB</td>
<td>122</td>
<td>2 viewpoints, surface, shoe, carrying condition, time</td>
</tr>
<tr>
<td>Soton Large Database</td>
<td>350GB</td>
<td>115</td>
<td>View</td>
</tr>
<tr>
<td>CASIA</td>
<td>10GB</td>
<td>153</td>
<td>11 viewpoints, clothing, carrying condition</td>
</tr>
<tr>
<td>OU-ISIR</td>
<td>529MB (silhouettes only)</td>
<td>4007</td>
<td>25 views, 32 clothing, 9 speeds, gait fluctuations</td>
</tr>
<tr>
<td>TUM-GAID Database</td>
<td>50GB</td>
<td>305</td>
<td>Backpack, shoes</td>
</tr>
</tbody>
</table>
The HumanID Gait Challenge Problem

• Data set of gait video
  – 122 subjects (1870 sequences)
  – Exercise 5 covariates
    • view, shoe, surface, carrying condition, time
  – 1.2 TB of data

• 12 Challenge experiments of increasing difficulty

• Baseline recognition algorithm to measure progress

• Joint effort of USF, NIST, and ND.

• Distributed to more than 50 groups and counting.
# Gallery and Probes

<table>
<thead>
<tr>
<th>Shoe</th>
<th>Concrete</th>
<th>Grass</th>
<th>Concrete</th>
<th>Grass</th>
</tr>
</thead>
</table>

- Collected over four days, May 20-21, 2001 and Nov 15-16, 2001
- 33 subjects common between the May and Nov
## USF Human ID Experiments

<table>
<thead>
<tr>
<th>Comment</th>
<th>Surface</th>
<th>Shoe</th>
<th>Camera</th>
<th>Carry</th>
<th>Time</th>
<th># sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gallery</td>
<td>Grass</td>
<td>A</td>
<td>R</td>
<td>No</td>
<td>May+Nov</td>
<td>122</td>
</tr>
<tr>
<td><strong>Probe A</strong></td>
<td>Grass</td>
<td>A</td>
<td>L</td>
<td>No</td>
<td>May+Nov</td>
<td>122</td>
</tr>
<tr>
<td><strong>Probe B</strong></td>
<td>Grass</td>
<td>B</td>
<td>R</td>
<td>No</td>
<td>May+Nov</td>
<td>54</td>
</tr>
<tr>
<td>Probe C</td>
<td>Grass</td>
<td>B</td>
<td>L</td>
<td>No</td>
<td>May+Nov</td>
<td>54</td>
</tr>
<tr>
<td><strong>Probe D</strong></td>
<td>Concrete</td>
<td>A</td>
<td>R</td>
<td>No</td>
<td>May+Nov</td>
<td>121</td>
</tr>
<tr>
<td><strong>Probe E</strong></td>
<td>Concrete</td>
<td>B</td>
<td>R</td>
<td>No</td>
<td>May+Nov</td>
<td>60</td>
</tr>
<tr>
<td><strong>Probe F</strong></td>
<td>Concrete</td>
<td>A</td>
<td>L</td>
<td>No</td>
<td>May+Nov</td>
<td>121</td>
</tr>
<tr>
<td><strong>Probe G</strong></td>
<td>Concrete</td>
<td>B</td>
<td>L</td>
<td>No</td>
<td>May+Nov</td>
<td>60</td>
</tr>
<tr>
<td><strong>Probe H</strong></td>
<td>Grass</td>
<td>A</td>
<td>R</td>
<td>Yes</td>
<td>May+Nov</td>
<td>120</td>
</tr>
<tr>
<td><strong>Probe I</strong></td>
<td>Grass</td>
<td>B</td>
<td>R</td>
<td>Yes</td>
<td>May+Nov</td>
<td>60</td>
</tr>
<tr>
<td><strong>Probe J</strong></td>
<td>Grass</td>
<td>A</td>
<td>L</td>
<td>Yes</td>
<td>May+Nov</td>
<td>120</td>
</tr>
<tr>
<td><strong>Probe K</strong></td>
<td>Grass</td>
<td>A/B</td>
<td>R</td>
<td>No</td>
<td>Nov</td>
<td>33</td>
</tr>
<tr>
<td><strong>Probe L</strong></td>
<td>Concrete</td>
<td>A/B</td>
<td>R</td>
<td>No</td>
<td>Nov</td>
<td>33</td>
</tr>
</tbody>
</table>
Year to Date performance on USF HumanID Dataset

Identification rate (Gallery size: 122)

Current and Proposed Solutions

- **Algorithm Developers cost (Time + Hardware)**
  - $\$$
  - $\$$
  - $\$$
  - $\$$

- **Performance Evaluator’s cost (Time + Hardware)**
  - $\$$
  - $\$$
  - $\$$
  - $\$$

**Current Solutions:***

- **In-house—self-defined test**
- **In-house—existing benchmark (FERET, HumanID, Gait)**

**Proposed Solutions:***

- **Independent—weakly supervised (FRGC competitions)**
- **Independent—supervised (FRVT-2002)**
- **Independent—strongly supervised: Using Cloud Computing**
- **Independent—strongly supervised: Using Grid Computing (FVC)**
Solution: Evaluation on the Cloud

- Eliminates the need to acquire dataset
- Easy to enforce consistency of protocols
- Easy access to archive of detailed results from many algorithms
GAIT Cloud Sequence Diagram
Login
Upload Code
Algorithm Evaluation
GAIT Cloud – Algorithm Submission

Upload Code

Specifications for Uploading:
1. Must be a .zip file
2. zip file MUST contain 2 folders named ‘src’ and ‘MAKE’
3. ‘src’ folder must contain all the source files
4. MAKE file must be written such that:

   - gaitBadUploadTest.zip 0.3MB Failed
   - Compilation Error. Please Try again.

Compilation Output:

c++ -O -o ./bin/gait_fast_similarity_only GaitFastSimilarityOnly.cpp image.h -lm
GaitFastSimilarityOnly.cpp:190:2: error: invalid preprocessing directive \#define
GaitFastSimilarityOnly.cpp:192: error: ISO C++ forbids declaration of ‘UCHAR’ with no type
GaitFastSimilarityOnly.cpp:194: error: expected ‘;’ before ‘”’ token
GaitFastSimilarityOnly.cpp: In function ‘double HammingSimilarityImage(Image, Image)’:
GaitFastSimilarityOnly.cpp:421: error: ‘UCHAR’ was not declared in this scope
GaitFastSimilarityOnly.cpp:421: error: expected ‘;’ before ‘Intersection’
GaitFastSimilarityOnly.cpp:432: error: ‘Intersection’ was not declared in this scope
GaitFastSimilarityOnly.cpp:432: error: ‘struct Image’ has no member named ‘Pixel’
GaitFastSimilarityOnly.cpp:432: error: ‘struct Image’ has no member named ‘Pixel’

http://marathon.csee.usf.edu/GaitBaseline/gaitcloud
GAIT Cloud – Trigger Experiments

http://marathon.csee.usf.edu/GaitBaseline/gaitcloud
GAIT Cloud: View & Download Results

http://marathon.csee.usf.edu/GaitBaseline/gaitcloud
Execution Time Comparison

For 1 job (12 experiments) using gait baseline algorithm.

Machine Configurations
EC2 Rental Cost Comparison

For 1 job (12 experiments) using gait baseline algorithm.

Machine Configurations

Cost incurred to rent EC2 (USD)

- EC2.m1.large (4): $4.08
- EC2.m2.xlarge (6.5): $5
- EC2.m1.xlarge (9): $5.44
- EC2.m2.2xlarge (13): $6.00
- EC2.c1.xlarge (20): $4.08
- EC2.m2.4xlarge (26): $10.00
Execution Time Vs Cost : Trend Graph

For 1 job (12 experiments) using gait baseline algorithm.

Cost incurred to rent EC2 (USD)

Execution Time in Hours

- 26 ECU's
- 20 ECU's
- 13 ECU's
- 8 ECU's
- 6.5 ECU's
- 4 ECU's
Avg CPU Utilization Comparison

For 1 job (12 experiments) using gait baseline algorithm.
Future Work

• Distributed Worker Architecture
  – MapReduce or All Pairs abstraction on the cloud.

• Meta Analysis of algorithm performance allowing for partial view of error and success conditions.

• Expand the implementation to allow for training of the algorithm.
Thank You

Dr. Sudeep Sarkar
and Ravi Panchumarthy, PhD Candidate,
Dept. of Computer Science and Engineering
University Of South Florida, Tampa, FL
sarkar@mail.usf.edu
ravi1@mail.usf.edu

Gait Cloud: http://marathon.csee.usf.edu/GaitBaseline/gaitcloud
Gait Challenge: http://gaitchallenge.org