Automated Text Mining
Comparison of Japanese and USA Multi-Robot Research

Robert J. Watts
U.S. Army Tank Automotive Research, Development & Engineering Center (TARDEC)
http://www.tacom.army.mil/tardec

Alan Porter, Ph.D. and Brian Minsk
Search Technology, Inc
http://www.searchtech.com
### Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

**Authors:** Robert Watts; Alan Porter; Brian Minsk

**Performing Organization:**
Search Technology, Inc, 6025 The Corners Parkway, Suite 202, Norcross, GA, 30092

**Sponsoring/Monitoring Agency:**
U.S. Army TARDEC, 6501 East Eleven Mile Rd, Warren, Mi, 48397-5000

**Distribution/Availability Statement:**
Approved for public release; distribution unlimited

**Security Classification:**
- Report: unclassified
- Abstract: unclassified
- This Page: unclassified

**Limitation of Abstract:**
Public Release

**Number of Pages:**
14
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Briefing Outline:

1. What’s Tech OASIS
2. Discuss the Data being Analyzed
   • Field Delimited
   • Multi-Robot Research
3. Processes for Segmenting Data
   • Deductive – Expert Opinion
   • Inductive – PCA based analysis
4. Expectancy Measure
5. Expectancy Measure applied to Segmented data
6. Observations & Interpretations
7. Conclusions & Recommendations

http://www.theVantagePoint.com
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Tech OASIS - A Software System for:
- Knowledge Discovery in Large Text Databases
- Profiling Thousands of Research Abstracts

Technology Scanning
Identifying new technologies and new developments in existing technologies

Technology Profiling
Identify key people and organizations

Technology Mapping and Decomposition
Identify dependencies and relationships

Technology Trending
Establish how a technology has emerged

Technology Forecasting
Project how a technology could evolve
As research progresses in distributed robotic systems, more and more aspects of multirobot systems are being explored. This Special Issue on Advances in Multirobot Systems provides a broad sampling of the research that is currently ongoing in the field of distributed mobile robot systems. To help categorize this research, we have identified seven primary research topics within multirobot systems: biological inspirations, communication, architectures, localization/mapping/exploration, object transport and manipulation, motion coordination, and reconfigurable robots. This editorial examines these research areas and discusses the Special Issue papers in this context. We conclude by identifying several additional open research issues in distributed mobile robotic systems.
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

### IEEE Imposed Categories

**“Guest Editorial, Advances in Multirobot Systems”**

- 354 EI Compendex & INSPEC abstracts
- Expert Perceived Research Categories
  - 324 abstracts grouped

### Deductive Categories
- Expert Field Awareness (e.g., Reconfigurable)

### 7+2 IEEE Imposed Categories

<table>
<thead>
<tr>
<th>IEEE MultiRobot Groups</th>
<th># Records</th>
<th>142</th>
<th>139</th>
<th>114</th>
<th>61</th>
<th>44</th>
<th>42</th>
<th>41</th>
<th>30</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture Allocation Control</td>
<td></td>
<td>142</td>
<td>55</td>
<td>45</td>
<td>27</td>
<td>27</td>
<td>24</td>
<td>14</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Motion Coordination</td>
<td></td>
<td>55</td>
<td>139</td>
<td>38</td>
<td>20</td>
<td>22</td>
<td>17</td>
<td>16</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>Communication</td>
<td></td>
<td>45</td>
<td>38</td>
<td>114</td>
<td>31</td>
<td>25</td>
<td>10</td>
<td>7</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Localization Mapping Exploration</td>
<td></td>
<td>27</td>
<td>20</td>
<td>31</td>
<td>61</td>
<td>9</td>
<td>8</td>
<td>5</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Biological</td>
<td></td>
<td>27</td>
<td>22</td>
<td>25</td>
<td>9</td>
<td>61</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Transport Manipulation Grasping</td>
<td></td>
<td>24</td>
<td>17</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>44</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Robot Learning</td>
<td></td>
<td>14</td>
<td>16</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>42</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Human Interface</td>
<td></td>
<td>9</td>
<td>19</td>
<td>12</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>41</td>
<td>2</td>
</tr>
<tr>
<td>IEEE OTHER</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Reconfigurable</td>
<td></td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Tech OASIS Automated Analyses:
- PCA based factors
- PCA – NO Singular Factor Solution

PCD Analysis Standardizes PCA
- Maximizes Inclusion of Abstracts in Factors, Number of Factors & Number of High Loading Factor Defining Terms
- Minimizes Abstracts in Multiple Factors

Min/Max Analysis - Analogous to Minimizing Entropy & Maximizing Cohesiveness of Factors
Expectancy Measure

Likelihood of item in one field having \( T \) or more abstracts in a specific category of a second field.

- Cumulative Binomial Distribution
- Detailed View group size / file size defines success probability \( p \)
- Field View item frequency \( n \) times \( p \) defines expected frequency
- Cumulative tail calculation based on whether the Detail View item frequency \( T \) > or < than expected

If a list item actually occurs \( T \) times in the records common to the records of a second list item and \( T \) is greater than or equal to the expected value, we get:

\[
p(X \geq T; n, p) = \sum_{r=T}^{n} \binom{n}{r} p^r (1 - p)^{n-r}
\]

Similarly, if \( T \) is less than or equal to the expected value, we get:

\[
p(X \leq T; n, p) = \sum_{r=0}^{T} \binom{n}{r} p^r (1 - p)^{n-r}
\]
Expectancy Measure

Likelihood of item in one field having $T$ or more abstracts in a specific category of a second field.

File Size = 107 abstracts
Motion Coordination => 46 abstracts
Probability $p = 46/107 = .43$

Field View Freq $Y = 30$
Expected Freq in Detail View = 13
Observed Detail View Freq = 19

Expectancy Measure = $1 - \sum_{r=T}^{n} \left( \binom{n}{r} p^r (1 - p)^{n-r} \right)$
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Expectancy Measure
- Anomaly – Expert Input
  - Low < -0.9
    - Protect Competitive Advantage IP
    - Publication lull prior to patent applications
    - Non-active in area
  - High > 0.9
    - Research Focus Area
    - Bias result of National Conference in subject area
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

**Expectancy Measure**
- Anomaly – Expert Input

**PCD** - fewer hi-low Expectancy grps

**Holistic Approach** – Multi-measure Pervasive Findings

**PCD Factor Names change over time (e.g., position to motion control and adaptive control to ...) depicting Tech Maturity

*OTHER* - Non-consensus

<table>
<thead>
<tr>
<th>#</th>
<th>Grp</th>
<th>Exp.</th>
<th>Metric</th>
<th>Cluster Group</th>
<th>Japan Sources</th>
<th>USA Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>11</strong></td>
<td>16</td>
<td>0.892</td>
<td>2.73</td>
<td>1998-99: Human computer interaction</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td>10</td>
<td>0.863</td>
<td>1.61</td>
<td>1998-99: Adaptive control systems</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td><strong>7</strong></td>
<td>12</td>
<td>0.636</td>
<td>1.53</td>
<td>1998-99: Robot learning</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td>11</td>
<td>0.654</td>
<td>1.22</td>
<td>1998-99: Position control</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td><strong>10</strong></td>
<td>21</td>
<td>-0.443</td>
<td>-0.83</td>
<td>1998-99: <strong>OTHER</strong></td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td><strong>21</strong></td>
<td>44</td>
<td>-0.478</td>
<td>-1.00</td>
<td>1998-99: learning (artificial intelligence)</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td><strong>12</strong></td>
<td>26</td>
<td>-0.512</td>
<td>-1.09</td>
<td>1998-99: Intelligent control</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td>16</td>
<td>-0.671</td>
<td>-1.79</td>
<td>1998-99: real-time systems</td>
<td>6</td>
<td>44</td>
</tr>
<tr>
<td><strong>18</strong></td>
<td>43</td>
<td>-0.731</td>
<td>-1.75</td>
<td>1998-99: cooperative systems</td>
<td>5</td>
<td>43</td>
</tr>
<tr>
<td><strong>5</strong></td>
<td>13</td>
<td>0.888</td>
<td>1.39</td>
<td>2000-01: Robot programming</td>
<td>11</td>
<td>24</td>
</tr>
<tr>
<td><strong>8</strong></td>
<td>25</td>
<td>0.847</td>
<td>1.25</td>
<td>2000-01: Algorithms</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td>20</td>
<td>0.765</td>
<td>1.06</td>
<td>2000-01: Computer simulation</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td><strong>8</strong></td>
<td>34</td>
<td>0.538</td>
<td>0.70</td>
<td>2000-01: multi-robot systems</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td><strong>7</strong></td>
<td>30</td>
<td>0.517</td>
<td>0.67</td>
<td>2000-01: multi-agent systems</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>12</td>
<td>0.511</td>
<td>0.68</td>
<td>2001-03: Distributed parameter control systems</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>19</td>
<td>-0.394</td>
<td>-1.87</td>
<td>2001-03: Manipulators</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>24</td>
<td>-0.617</td>
<td>-3.70</td>
<td>2001-03: <strong>OTHER</strong></td>
<td>9</td>
<td>34</td>
</tr>
<tr>
<td><strong>12</strong></td>
<td>27</td>
<td>0.946</td>
<td>1.70</td>
<td>2002-03: multi-robot systems</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>7</td>
<td>0.87</td>
<td>2.03</td>
<td>2002-03: Intelligent robots</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>7</td>
<td>0.87</td>
<td>2.03</td>
<td>2002-03: Manipulators</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td><strong>5</strong></td>
<td>14</td>
<td>0.965</td>
<td>1.00</td>
<td>2002-03: Motion control</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td><strong>9</strong></td>
<td>31</td>
<td>0.481</td>
<td>0.68</td>
<td>2002-03: Collision avoidance</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>8</td>
<td>-0.31</td>
<td>-1.86</td>
<td>2002-03: Collision avoidance</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>9</td>
<td>-0.488</td>
<td>-2.11</td>
<td>2002-03: Robustness (control systems)</td>
<td>10</td>
<td>31</td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>6</td>
<td>-0.507</td>
<td>-3.04</td>
<td>2002-03: sensor fusion</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>11</td>
<td>-0.608</td>
<td>-3.34</td>
<td>2002-03: Control system analysis</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>24</td>
<td>-0.836</td>
<td>-6.01</td>
<td>2002-03: <strong>OTHER</strong></td>
<td>3</td>
<td>14</td>
</tr>
</tbody>
</table>
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Japanese 2002-03 Multi-Robot Research

Cros-Correlation Map
Time-Slice Descriptors Combined
Descriptors Combined C: (Relev.
Links >= 0.400000 shown
> 0.75 0 (0)
0.50 - 0.75 6 (0)
0.25 - 0.50 6 (5)
< 0.25 0 (27)

2002-03: Collision avoidance
2002-03: Motion control
2002-03: Robustness (control systems)
2002-03: Manipulators
2002-03: multi-agent systems
2002-03: Control system analysis
2002-03: Intelligent robots
2002-03: *OTHER*

Affiliations Combined C:Japan
1 Keio University, Yokohama
1 Osaka University, Osaka,
1 Saga University, Saga-shi
1 Nara Institute of Science & Tech
1 Tohoku University, Sendai

2002-03: Collision avoidance
2002-03: Motion control
2002-03: Robustness (control systems)
2002-03: Multi-robot systems
2002-03: *OTHER*

Affiliations Combined C:USA
1 University of Pennsylvania
1 The Robotics Institute, C
1 Jet Propulsion Laboratory

2002-03: Multi-robot systems
2002-03: Motion control
2002-03: Intelligent robots
2002-03: sensor fusion
2002-03: Control system analysis

Affiliations Combined USA A.
1 University of Pennsylvania
1 The Robotics Institute, C
1 Jet Propulsion Laboratory

2002-03: Collision avoidance
2002-03: Multi-robot systems
2002-03: Motion control
2002-03: Sensor fusion
2002-03: Control system analysis

Affiliations Combined USA A.
1 The Robotics Institute, C
1 University of Pennsylvania
1 Jet Propulsion Laboratory

2002-03: Robustness (control systems)
2002-03: Manipulators
2002-03: multi-agent systems
2002-03: Control system analysis
2002-03: Intelligent robots

Affiliations Combined USA A.
1 The Robotics Institute, C
1 University of Pennsylvania
1 Jet Propulsion Laboratory
1 Oak Ridge National Labora
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

**Observations:**

**Expectancy Measure =>** Japanese - less emphasis than expected on biological approaches, reconfigurable robots & architecture allocation control (IEEE) and *OTHER* (PCD)

=> Japanese - more emphasis than expected on human interface & motion coordination (IEEE) and Multi-robot systems & manipulators (PCD)

=> USA sources – less emphasis than expected on human interface & robot learning (IEEE)

=> USA more emphasis than expected on reconfigurable robots & communication (IEEE) and sensor fusion & *OTHER* (PCD)

**Expert Opinion:** Japanese focus more on Industrial Robots and Human Aiding Robots. Must Determine Implications of low *OTHER* expectancy.
Automated Text Mining Comparison of Japanese and USA Multi-Robot Research

Conclusions & Recommendations

• Overview of Tech OASIS & Text-Mining Capabilities
• Analyzed Field Delimited Data on Subject of Multi-Robot Research
• Approaches for Segmentation of the data:
  ➢ **Deductive** (i.e., Expert Perceived) Categories
    ✓ Easier to Use to Generalize Observations over time
    ✓ Field Experts Understand…Acceptance
    ✓ But…Bias to Present Time Period
  ➢ **Inductive** (i.e. PCD Derived) Categories
    ✓ Standardizes Analysis
    ✓ Enables Technology Maturity “Subjective” Assessment
    ✓ but…Biased by high numbers of low frequency sources of tech papers
Expectancy Measure – Ascertain Topical Emphasis Areas & Identifies Unexpected Patterns….as do other measures

Use Holistic Approach…Pervasive Patterns…Include Field Experts

Tech OASIS / VantagePoint Automates Clustering / Categorization of Information to Enable and Improve:

• Cognition of Broad Field of Research
• Elicit Research Questions from noted Anomalies
• Promote Innovation through Expert Involvement

http://www.theVantagePoint.com