Disclaimer: The data set was retrieved from the web by the Yahoo search engine, and is provided AS-IS, without any warranty. Images may be included which are subject to copyright. (If you are the owner of a picture and want it removed, please contact us.) You may choose to use it, at your own risk, without cost for research purposes; it may NOT be downloaded for commercial use. We expect revisions of the dataset/documentation in the future.

1 Dataset version 1.0

Release date: 07/14/05.

2 WWW Access

All HTML files in the directories described below are accessible through the main page at http://visionpc.cs.uiuc.edu/isd/. The dataset directory is at http://visionpc.cs.uiuc.edu/isd/dataset.

3 Structure of the dataset directory

The data directory contains 3 subdirectories; one for each query term (BASS, CRANE, SQUASH). For each query term, a text file is provided:

```
queryterm/annotations.txt
```

- **First column:** Number of the annotated image¹
- **Second column:** Colon (separation character)
- **Third column:** Annotated sense number (cf. table 1 below)

The file `queryterm/dataset.html` permits browsing the dataset.

¹Note that numbering is not necessarily consecutive because annotation took place after some of the automatic filtering of images/HTML documents (cf. filtering criteria in paper).
In addition, for each query term, there are three subdirectories:

1. **queryterm/images**: Annotated images in their true sizes
2. **queryterm/pages**: HTML documents corresponding to the annotated images
3. **queryterm/info**: Text files with the following information about the downloaded and annotated images and their corresponding web pages:
   - the URL of the web page
   - the URL of the image target
   - the URL of the image thumbnail
   - the page TITLE.

The file numbers correspond to the first column in `queryterm/annotations.txt`.

<table>
<thead>
<tr>
<th>Word (#Annot. images)</th>
<th>QueryTerms</th>
<th>Senses</th>
<th>Coverage</th>
<th>Examples of visual annotation cues</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BASS</strong> (2881)</td>
<td>5: bass, bass guitar, bass instrument, bass fishing, sea bass</td>
<td>1. fish 2. <strong>musical instrument</strong> 3. related: fish 4. related: <strong>musical instrument</strong> 5. unrelated 6. people</td>
<td>35% 28% 10% 12% 7%</td>
<td>any fish, people holding catch any bass-looking instrument, playing fishing (gear, boats, farms), food, rel. charts/maps speakers, accessories, works miscellaneous (above senses not applicable) faces, crowd (above senses not applicable)</td>
</tr>
<tr>
<td><strong>CRANE</strong> (2650)</td>
<td>5: crane, construction cranes, whooping crane, sandhill crane, origami cranes</td>
<td>1. <strong>machine</strong> 2. <strong>bird</strong> 3. <strong>origami</strong> 4. related: <strong>machine</strong> 5. related: <strong>bird</strong> 6. related: <strong>origami</strong> 7. people 8. unrelated 9. <strong>karate</strong></td>
<td>21% 26% 4% 11% 11% 1% 7% 18% 1%</td>
<td>machine crane, incl. panoramas crane bird or chick origami bird other machinery, construction, motor, steering, seat egg, other birds, wildlife, insects, hunting, rel. maps/charts origami shapes (stars, pigs), paper folding faces, crowd (above senses not applicable) miscellaneous (above senses not applicable) martial arts</td>
</tr>
<tr>
<td><strong>SQUASH</strong> (1948)</td>
<td>10: squash+: rules, butternut, vegetable, grow, game of, spaghetti, winter, types of, summer</td>
<td>1. <strong>vegetable</strong> 2. <strong>sport</strong> 3. related:vegetable 4. related: <strong>sport</strong> 5. people 6. unrelated</td>
<td>24% 13% 31% 6% 10% 16%</td>
<td>squash vegetable people playing, court, equipment agriculture, food, plant, flower, insect, vegetables other sports, sports complex faces, crowd (above senses not applicable) miscellaneous (above senses not applicable)</td>
</tr>
</tbody>
</table>

Table 1: Web images for three ambiguous query terms were annotated manually out of context (without considering the web page document). For each term, the number of annotated images, the query retrieval terms, the senses, their distribution, and rough sample annotation guidelines are provided, with core senses marked in bold face. Because image retrieval engines restrict hits to 1000 images, query expansion was conducted by adding narrowing query terms from askjeeves.com to increase corpus size. We selected terms relevant to core senses, i.e. the main discrimination phenomenon.

### 4 Structure of the runs directory

This directory contains median results files (from 5 runs with different initializations). There are six files for all three query keyword (BASS, CRANE, SQUASH), each one run for:

- **All+Meta** considered all senses and the meta senses. The meta senses merged related senses with their respective core sense (see below). All annotated images were considered.
- **Core** contains just core senses and only images annotated with core senses were considered.
5 Annotation overview

Senses, sense distributions, and annotation criteria are summarized in Table 1. The images were annotated with one of several labels by one of the authors out of context (without considering the web site and its text). For annotation purposes, images were numbered and displayed on a web page in thumbnail size. In case the thumbnail was not sufficient for disambiguation, the image linked at its true size to the thumbnail was inspected. The true-size view depended on the size of the original picture and showed the image and its name. However, the annotator tried to resist name influence, and make judgments based just on the image. For each query, 2 to 4 core word senses (e.g. squash vegetable and squash sport for SQUASH) were distinguished from inspecting the data. However, because “context” was restricted to the image content, and there was no guarantee that the image actually depicts the query term, additional annotator senses were introduced. Thus, for most core senses, a RELATED label was included, accounting for meanings that seemed related to core meaning but lacked a core sense object in the image. In addition, for each query term, a PEOPLE label was included because such images are common due to the nature of how people take pictures (e.g. portraits of persons or group pictures of crowds, when core or related senses did not apply), as was an UNRELATED label for irrelevant images which did not fit other labels or were undeterminable.

Contact information: {loeff, ebbaalm}@uiuc.edu