Progress

- Tried calling C++ code from Matlab
- Improved previous object recognition system using K-means with SIFT+SVM to compare results
- Tested
- Worked on website
  - www.freecsstemplates.org
- Worked on poster
- Learning more about LaTeX
  - MiKTeX distribution + TeXnicCenter IDE
Learning Visual Bits with Direct Feature Selection
Joel Jurik\textsuperscript{1} and Dr. Rahul Sukthankar\textsuperscript{2,3}
\textsuperscript{1}University of Central Florida \textsuperscript{2}Intel Research Pittsburgh \textsuperscript{3}Robotics Institute, Carnegie Mellon

Abstract
Popular object category recognition systems currently use an approach that represents images by a bag of visual words. However, these systems can be improved in two ways using the framework of visual bits [1]. First, instead of representing each image feature by a single visual word, each feature is represented by a sequence of visual bits. Second, instead of separating the processes of codebook generation and classifier training, we unify them into a single framework.

We propose a new way to learn visual bits using direct feature selection to avoid the complicated optimization framework from [1]. Our results confirm that visual bits outperform the bag of words model on object category recognition.

Explanation
For image representation, we use SIFT [2].

For training the visual bits system, we use the following parameters:
\begin{itemize}
  \item Weight distribution: Uniform [-1000, +1000]
  \item Number of weights: 10,000
  \item Rounds of boosting: 200
\end{itemize}

We compare our performance, against a baseline object category recognition that uses k-means + SVM with the following parameter:
\begin{itemize}
  \item Number of clusters: 1000
\end{itemize}

Each system uses the same 200 training images, in which 100 are positive, and 100 are negative. The positive images are of airplanes, and the negative images can be rhinos, cars, elephants, faces, or minarets. SIFT is used for image representation, and SVM is used during training.

Testing Parameters
For testing, we use a separate set of images from the training set that consists of 100 images, where 50 are positive and 50 are negative. The task at hand is to distinguish between an airplane and non-airplane images.

<table>
<thead>
<tr>
<th>System</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Bits</td>
<td>89%</td>
</tr>
<tr>
<td>K-means</td>
<td>86%</td>
</tr>
</tbody>
</table>

As you can see, the visual bits system outperforms the object recognition system using k-means clustering with the task of distinguishing between airplane and non-airplane images. Object category recognition using multiple categories using visual bits is currently being implemented.

References
Simplified Visual Bits Abstract

The standard visual bits approach to object category recognition uses optimization in order to solve many difficult situations in current recognition systems. Since optimization is very complex, simplified visual bits uses many randomly generated weights to achieve similar results.

About

Joel Jurik was a senior computer engineering student at the University of Central Florida at the time of the Computer Vision REU. His interests are image procssing, computer vision, computer networks, and software development. He has interned for Lockheed Martin, was a member of the UCF Programming Team, and was part of many other clubs and organizations. He plans to pursue a Ph.D. in computer engineering.

Presentation Files (.pdf)
- Presentation 1 (5/22/09)
- Presentation 2 (5/29/09)
- Presentation 3 (6/05/09)
- Presentation 4 (6/12/09)
- Presentation 5 (6/19/09)
- Presentation 6 (7/02/09)
- Presentation 7 (7/10/09)
- Presentation 8 (7/17/09)

Matlab Files (.zip)
- Simplified Visual Bits Program

Paper (.pdf)
- Simplified Visual Bits Paper
Training

- Parameters for both systems:
  - 200 images (100 airplane, 20 rhino, 20 elephant, 20 minaret, 20 faces, 20 cars)
  - Image representation using SIFT
  - SVM during training

- Parameters for visual bits
  - Number of weights: 10,000
  - Weight distribution: Uniform [-1000,+1000]
  - Number of rounds of boosting: 200

- Parameter for K-means
  - 1000 cluster centers
Testing

- 100 images (50 airplane, 10 rhino, 10 elephant, 10 minaret, 10 faces, 10 cars)
- The task is to distinguish between an airplane and non-airplane image
## Results

<table>
<thead>
<tr>
<th>System</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Bits</td>
<td>89%</td>
</tr>
<tr>
<td>K-means</td>
<td>86%</td>
</tr>
</tbody>
</table>
Plan

- Change Visual Bits and K-means systems to recognize multiple categories
- Change to a complete, standard dataset (like PASCAL VOC 2006)
- Start work on paper