Simplified Visual Bits – 7/24

Joel Jurik

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[1] and Dr. Rahul Sukthankar[2,3]

[1]University of Central Florida [2]Intel Research Pittsburgh [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].

Figure 1: Examples of SIFT features

Once the SIFT descriptors for all images have been obtained, we randomly generate weights using a uniform distribution from -1000 to +1000 with 128 dimensions (the same as SIFT). We then find the linear projection using these weights and descriptors, and find the best projection and threshold using the GentleBoost algorithm [3]:

Gentle AdaBoost

1. Start with weights $w_i = 1/N$, $i = 1, 2, \ldots, N$, $F(x) = 0$.
2. Repeat for $m = 1, 2, \ldots, M$:
   (a) Fit the regression function $f_m(x)$ by weighted least-squares of y_i to x_i with weights w_i.
   (b) Update $F(x) \leftarrow F(x) + f_m(x)$.
   (c) Update $w_i \leftarrow w_i \exp(-y_i f_m(x_i))$ and renormalize.
3. Output the classifier $\text{sign}[F(x)] = \text{sign}[\sum_{m=1}^{M} f_m(x)]$.

Figure 2: GentleBoost algorithm

Using GentleBoost, we are able to construct a sequence of visual bits that represent an image at a much higher level than SIFT. We take the aggregates of visual bits to train the SVM, which will create a model used for testing.

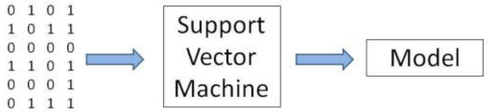


Figure 3: Training using visual bits

Training Parameters

For training the visual bits system, we use the following parameters:

- Weight distribution: Uniform [-1000,+1000]
- Number of weights: 10,000
- Rounds of boosting: 200

We compare our performance, against a baseline object category recognition that uses k-means + SVM with the following parameter:

- Number of clusters: 1000

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.

Figure 4: Examples of images used.

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.

Results

System	Accuracy
Visual Bits	89%
K-means	86%

Figure 5: Results. Visual Bits system shows improvement upon standard clustering approaches

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

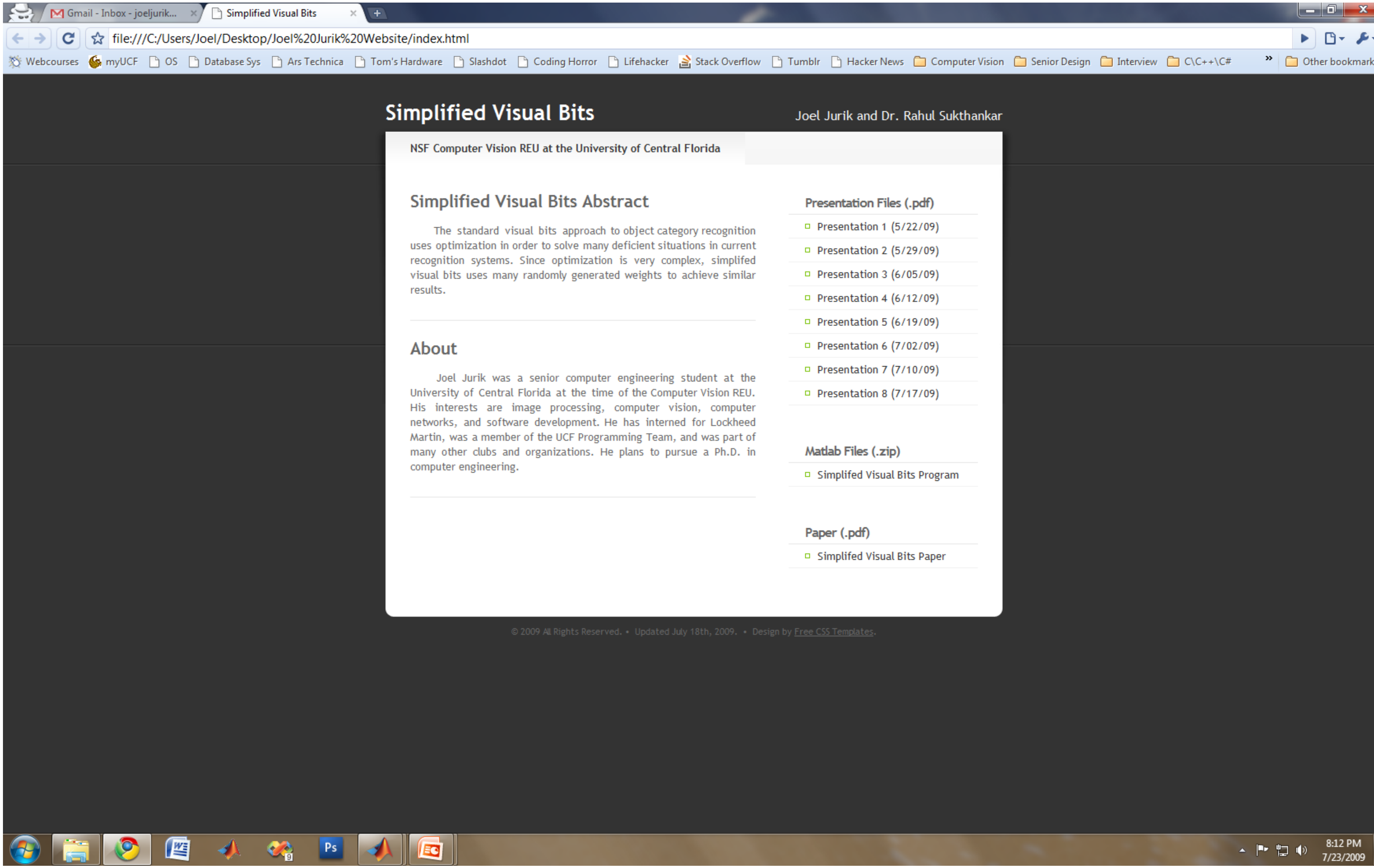
[1] L. Yang, R. Jin, R. Sukthankar, F. Jurie: Unifying Discriminative Visual Codebook Generation with Classifier Training for Object Category Recognition, CVPR 2008

[2] D. G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints, IJCV 2004

[3] J. Friedman, J., Hastie, T., Tibshirani, R.: Additive Logistic Regression: A Statistical View Of Boosting, Annals of Statistics, Vol. 28, 1998.

[4] P. Viola, M. Jones: Robust Real-time Object Recognition, IJCV 2001

Website


Gmail - Inbox - joeljurik...
Simplified Visual Bits
file:///C:/Users/Joel/Desktop/Joel%20Jurik%20Website/index.html
Webcourses
myUCF
OS
Database Sys
Ars Technica
Tom's Hardware
Slashdot
Coding Horror
Lifehacker
Stack Overflow
Tumblr
Hacker News
Computer Vision
Senior Design
Interview
C\C++\C#
Other bookmarks
Simplified Visual Bits
Joel Jurik and Dr. Rahul Sukthankar
NSF Computer Vision REU at the University of Central Florida
Simplified Visual Bits Abstract
The standard visual bits approach to object category recognition uses optimization in order to solve many deficient situations in current recognition systems. Since optimization is very complex, simplifed 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 processing, 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)
Simplifed Visual Bits Program
Paper (.pdf)
Simplifed Visual Bits Paper
© 2009 All Rights Reserved. • Updated July 18th, 2009. • Design by Free CSS Templates.
8:12 PM
7/23/2009

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

System	Accuracy
Visual Bits	89%
K-means	86%

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