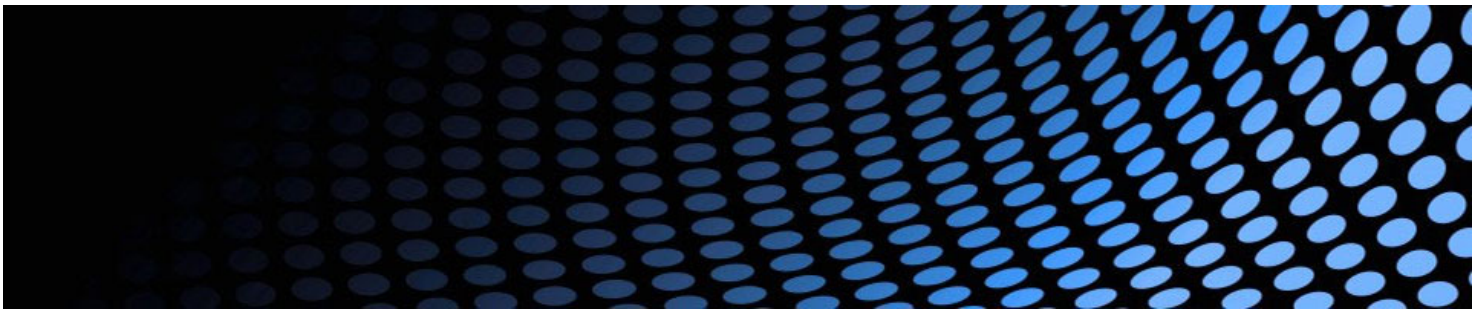

Aug 25

A simple object classifier with Bag-of-Words using OpenCV 2.3 [w/ code]

Categories:

code, opencv, programming, Recommended, Software, video, vision

by Roy

Just wanted to share of some code I've been writing.
So I wanted to create a food classifier, for a cool project down in the Media Lab called FoodCam. It's basically a camera that people put free food under, and they can send an email alert to the entire building to come eat (by pushing a huge button marked "Dinner Bell"). Really a cool thing.

OK let's get down to business.

I followed a very simple technique described in this paper. I know, you say, "A Paper? Really? I'm not gonna read that technical boring stuff, give the bottom line! man.. geez." Well, you are right, except that this paper IS the bottom line, it's dead simple. It's almost a tutorial. It is also referenced by the OpenCV documentation.

Edit (6/5/2014): Another great read for selecting the best color-space and invariant features is this paper by van de Sande et al.

The method is simple:
- Extract features of choice from training set that contains all classes.
- Create a vocabulary of features by clustering the features (kNN, etc). Let's say 1000 features long.
- Train your classifiers (SVMs, Naive-Bayes, boosting, etc) on training set again (preferably a different one), this time check the features in the image for their closest clusters in the vocabulary. Create a histogram of responses for each image to words in the vocabulary, it will be a 1000-entries long vector. Create a sample-label dataset for the training.
- When you get an image you havn't seen - run the classifier and it should, god willing, give you the right class.

Turns out, those crafty guys in WillowGarage have done pretty much all the heavy lifting, so it's up for us to pick the fruit of their hard work. OpenCV 2.3 comes packed with a set of classes, whose names start with BOW for Bag Of Words, that help a lot with implementing this method.

Starting with the first step:

```
Mat training_descriptors(1,extractor->descriptorSize(),extractor->descriptorType());

SurfFeatureDetector detector(400);
vector keypoints;

// computing descriptors
Ptr extractor(
   new OpponentColorDescriptorExtractor(
      Ptr(new SurfDescriptorExtractor())
   )
);

while(..loop a directory? a file?..) {
   Mat img = imread(filepath);
   detector.detect(img, keypoints);
   extractor->compute(img, keypoints, descriptors);
   training_descriptors.push_back(descriptors);
}
```

Simple!
Let's go create a vocabulary then. Luckily, OpenCV has taken care of that, and provide a simple API:

```
BOWKMeansTrainer bowtrainer(1000); //num clusters
bowtrainer.add(training_descriptors);
Mat vocabulary = bowtrainer.cluster();
```

Boom. Vocabulary.
Now, let's train us some SVM classifiers!
We're gonna train a 2-class SVM, in a 1-vs-all kind of way. Meaning we train an SVM that can say "yes" or "no" when choosing between one class and the rest of the classes, hence 1-vs-all.
But first, we need to scour the training set for our histograms (the responses to the vocabulary, remember?):

```
vector<KeyPoint> keypoints;
Mat response_hist;
Mat img;
string filepath;
map<string,Mat> classes_training_data;

Ptr<FeatureDetector > detector(new SurfFeatureDetector());
Ptr<DescriptorMatcher > matcher(new BruteForceMatcher<L2<float> >());
Ptr<DescriptorExtractor > extractor(new OpponentColorDescriptorExtractor(Ptr<DescriptorExtractor>(ne
Ptr<BOWImgDescriptorExtractor> bowide(new BOWImgDescriptorExtractor(extractor,matcher));
bowide->setVocabulary(vocabulary);

#pragma omp parallel for schedule(dynamic,3)
for(..loop a directory?..) {
   img = imread(filepath);
   detector->detect(img,keypoints);
   bowide.compute(img, keypoints, response_hist);

   #pragma omp critical
   {
      if(classes_training_data.count(class_) == 0) { //not yet created...
         classes_training_data[class_].create(0,response_hist.cols,response_hist.type());
         classes_names.push_back(class_);
      }
      classes_training_data[class_].push_back(response_hist);
```

```
    }
    total_samples++;
}
```

Now, two things:
First notice I'm keeping the training data for each class separately, this is because we will need this for later creating the 1-vs-all samples-labels matrices.
Second, I use OpenMP multi(-threading)processing to make the calculation parallel, and hence faster, on multi-core machines (like the one I used). Time is sliced by a whole lot. OpenMP is a gem, use it more. Just a couple of #pragma directives and you're multi-threading.

Alright, data gotten, let's get training:

```
#pragma omp parallel for schedule(dynamic)
for (int i=0;i<classes_names.size();i++) {
   string class_ = classes_names[i];
   cout << omp_get_thread_num() << " training class: " << class_ << ".." << endl;

   Mat samples(0,response_cols,response_type);
   Mat labels(0,1,CV_32FC1);

   //copy class samples and label
   cout << "adding " << classes_training_data[class_].rows << " positive" << endl;
   samples.push_back(classes_training_data[class_]);
   Mat class_label = Mat::ones(classes_training_data[class_].rows, 1, CV_32FC1);
   labels.push_back(class_label);

   //copy rest samples and label
   for (map<string,Mat>::iterator it1 = classes_training_data.begin(); it1 != classes_training_data.
      string not_class_ = (*it1).first;
      if(not_class_.compare(class_)==0) continue; //skip class itself
      samples.push_back(classes_training_data[not_class_]);
      class_label = Mat::zeros(classes_training_data[not_class_].rows, 1, CV_32FC1);
      labels.push_back(class_label);
   }

   cout << "Train.." << endl;
   Mat samples_32f; samples.convertTo(samples_32f, CV_32F);
   if(samples.rows == 0) continue; //phantom class?!
   CvSVM classifier;
   classifier.train(samples_32f,labels);

   //do something with the classifier, like saving it to file
}
```

Again, I parallelize, although the process is not too slow.
Note how I build the samples and the labels, where each time I put in the positive samples and mark the labels '1', and then I put the rest of the samples and label them '0'.

Moving on to .... testing the classifiers!
Nothing seems to me like more fun than creating a confusion matrix! Not really, but let's see how it's done:

```
map<string,map<string,int> > confusion_matrix; // confusionMatrix[classA][classB] = number_of_times_
map<string,CvSVM> classes_classifiers; //This we created earlier

vector<string> files; //load up with images
vector<string> classes; //load up with the respective classes

for(..loop over a directory?..) {
   Mat img = imread(files[i]),resposne_hist;

   vector<KeyPoint> keypoints;
   detector->detect(img,keypoints);
   bowide->compute(img, keypoints, response_hist);

   float minf = FLT_MAX; string minclass;
   for (map<string,CvSVM>::iterator it = classes_classifiers.begin(); it != classes_classifiers.end(
      float res = (*it).second.predict(response_hist,true);
      if (res < minf) {
         minf = res;
         minclass = (*it).first;
      }
   }
   confusion_matrix[minclass][classes[i]]++;
```

```
23 }
```

When you take a look in my files, you will find a much complicated way of doing this. But this is the core idea - look in the image for the response histogram to the vocabulary of features (rather, feature-cluster-ceneters), run it by all the classifiers and take the one with the best score. Simple.
Consider making this parallel as well. No reason for it to be serial.

That's about covers it.

Code

Lately I'm pushing stuff in Github.com using git rather than SVN on googlecode. Donno why, it's just like that.
Get the whole thing at:
https://github.com/royshil/FoodcamClassifier

Follow the build instructions, they're a breeze, and then follow the runnning instructions. It's basically a series of command-line programs you run to get through each step, and in the end you have like a "predictor" service that takes an image and produces a prediction.

Edit (6/5/2014): The dataset can be downloaded from: http://www.media.mit.edu/~roys/shared/foodcamimages.zip

OK guys, have fun classifying stuff!
Roy.

Tags: classification, object, opencv, svm, vision

hew kim • a year ago
HI Roy
Good article, very appreciated for your sharing.
I notice you had provide the train.txt and test.txt file. I wonder how you prepared both file?
• Reply • Share ›

combfish • 2 years ago
maybe I found a bug in make_test_background_image.cpp:

```
if (!accum.data) {
accum.create(img.size(), CV_64FC3);
}

if (img64.size() == accum.size()) {
accum += img64;
}
```

use the accum without Initialization, and get the background.png wrong !

so I changed it :

```
if (!accum.data) {
accum=img64.clone();
continue;
}

cout<<" accum "<<accum.at<cv::vec3d>(0,0)[0]<< endl;
if (img64.size() == accum.size()) {
```

accum += img64;
}

• Reply • Share ›

Jumabek • 2 years ago

Actually it was simple I just changed "j4" part to directory folder ;)

• Reply • Share ›

Jumabek • 2 years ago

Hello, great job bro. Thank you for sharing and comprehensive explanation.
But the problem is I am not in that level to understand this project.
Nevertheless I really have to make it work for my assignment.
When I tried following commands in the command line , i got an error.
Can you help me to make it run. Thanks again!

I run this in the command line

D:\labworks\object recognition\FoodcamClassifier>cmake -D CMAKE_CXX_FLAGS=-fopen
mp . ; make -j4
CMake Error: The source directory "D:/labworks/object recognition/FoodcamClassif
ier/-j4" does not exist.
Specify --help for usage, or press the help button on the CMake GUI.

• Reply • Share ›

lunamystry • 2 years ago

Hi Roy,

Wonderful post, thank you so much!
I am wondering though. I am looking at the code on github: https://github.com/royshil/...

Just the main file, the other files scare me (I have about two week worth of knowledge on CV and all this.)

On line 105 where you are doing the assigning of the histogram to classes and 1-vs-many training, you compute descriptors for each of the train images. You don't seem to calculate new keypoints for the image you are computing descriptors for though. Is this how it is? won't the keypoints variable contain data from the last image you computed descriptors for when you were making the bag of keypoints?

• Reply • Share ›

Chelsea • 2 years ago

Hi,

Would this work on OpenCV 2.4.9? I tried running it but it failed to compile. I was wondering if I needed to have OpenCV 2.3 instead since files may have been changed around?

The error I got was:
make[2]: *** [CMakeFiles/foodcam-predict.dir/predict_common.cpp.o] Error 1
make[1]: *** [CMakeFiles/foodcam-predict.dir/all] Error 2
make: *** [all] Error 2

Thanks.

• Reply • Share ›

Roy ➚ Chelsea • 2 years ago

@Chelsea
It probably needs changed to the API, as it must have changed.
You could give it a try yourself and submit your code, I'll review and merge if it works...

• Reply • Share ›

ismail • 3 years ago

Hey Roy,

I wanted to ask what if we only have two classes? We only train one One Vs All model.

Also and more importantly, when you compare to get the least response what if one of them is negative?What does that mean?

Thank you

Reply • Share ›

Mara • 3 years ago

Hello, Roy! In the training of SVM part, I'm getting an error saying CvSVM is private and train member cannot be accessed. I changed map classes_classifiers to map<string,unique_ptr> classes_classifiers, then classes_classifiers[class_].train(samples_32f,labels) to classes_classifiers[class_]->train(samples_32f, labels), but I'm still getting an error.

I did this

classes_classifiers[class_]->train(samples_32f, labels);
cout << "adding2 " << classes_training_data[class_].rows << " positive" << endl;

to see where it stops working, and it seems that it doesn&#039t reach to adding2.

This part of the code

cout << " Training class: " << class_ << ".." << endl;

also doesn&#039t display class_.

What could be wrong? :(Thank you!

Reply • Share ›

Rajind ➚ Mara • 2 years ago

@Mara did you find any solution to this ? I am facing the same problem.

Reply • Share ›

SN • 3 years ago

I just wanna really thank you for this article, I loved this approach.

Reply • Share ›

Boris • 3 years ago

The same problem with me. The link with dataset is broken.

Reply • Share ›

Amartinez • 3 years ago

Hi guys,

I am trying to run the code of the example and I noticed that in the repository there is not any dataset. Reading the documentation, I got the link where there is the dataset: http://fay.media.mit.edu/fo.... I tried several times but the link seems to be broken. Could anyone here provide me the dataset?.

Thanks in advance.

Reply • Share ›

Kastor • 3 years ago

@Krystal

If u look at the SVM only sample,
u can see that its just a mark. I think its made to distinguish
different runs.

Here is the setting in svm only , main

string file_postfix = "with_colors"

Reply • Share ›

Krystal • 3 years ago

Hey Roy, thanks for you tutorial and code,

I have got the vocabulary and working on the SVM classifier now, and i m not sure about the input argv[2] postfix in the train_bowv.cpp, is it

I have got the vocabulary and working on the SVM classifier now, and I'm not sure about the input argv[2] postfix in the train_bowv.cpp, is it the address ofthe output file?

Thanks

Krystal

• Reply • Share ›

Kastor • 3 years ago

Hi Roy,

1st of all thanks alot for this great article !
Some questions remain.

I see you gain your positive examples by cutting the desired part out of the image,
that means you have to use the sliding window, correct ?

But ist that a good Idea ? I mean u need a good choice for the size of that window.
What is your strategy on that ?

Thanks
Kastor

• Reply • Share ›

Rajni Kant • 3 years ago

Is there any java implementation of this code.

• Reply • Share ›

Adam N • 4 years ago

Hey Roy, awesome article + code! Was just wondering what size training sets you used for building the dictionary and training the SVMs?
Many thanks Adam :)

• Reply • Share ›

Polk • 4 years ago

Hi,

It's a great works and tutorial.
I use Xcode 4.6.6 in OS X mountain lion, OpenCV 2.4.3
opencv works fine in xcode with compiler LLVM GCC 4.2 for OpenMP

I have a problem that all the test images has zero maches and if the set of trainig is limit (example 10 images) all the test images goes to first image.

I use this (line a):
Mat training_descriptors;
Your example use (line b):
Mat training_descriptors(1,extractor->descriptorSize(),extractor->descriptorType());

if I use the second (line b) the process stop and has a problem when reach the line
Mat vocabulary = bowtrainer.cluster();
but if I use the first (line a) it works but the results are extranges.

Thanks for the help.

• Reply • Share ›

Bhaskar • 4 years ago

Hi,

Thanks for the great tutorial. I didnt use OpenMP but I think my results (for something completely diff) are good. I knew the theory and everything but its implementation was a bit out there. Thanks again for all the help.

• Reply • Share ›

Timo • 4 years ago

Hello

Following instructions from main.cpp ends up with empty histograms. Could it be that it is necessary to add the line "detector.detect(img, keypoints)" before each "bowide.compute(img, keypoints, response_hist)" in the loop so that the corresponding keypoints of the current query image are passed to the compute function? Doing so results in a more realistic histogram representation.

Reply • Share ›

Rish • 4 years ago

Hello;
I am working on a low resolution images and I find it diffcult to obtain best matched features using SURF and there a alot of possibilities that Mis-match occurs. Do you think I can use technique of using high quality data set and low quality as well. Please help me out!!! Thank you

Reply • Share ›

Pavel • 4 years ago

Hi, thanks for the great tutorial.

I'd really appreciate if you could suggest me a way to solve this problem http://answers.opencv.org/q...

Many thanks,
Pavel

Reply • Share ›

John • 4 years ago

Hi,

Is there a way to convert it into java android application? I want to recognize leafs instead of pizzas etc.

Thanks,
John

Reply • Share ›

paul • 4 years ago

Hi, nice tutorial!
I have one problem, when I execute "cmake -D CMAKE_CXX_FLAGS=-fopenmp . " I get the following error:

Could not find a configuration file for package "OpenCV" that is compatible with requested version "2.3".

The following configuration files were considered but not accepted:

/opencv/cmake/OpenCVConfig.cmake, version: unknown

Do you know why it doesn't detect the version of opencv?

Reply • Share ›

Giseli • 4 years ago

Thanks for the reply, Roy. I'm tweaking to see the best number of images. Another question... I would like to draw the ROIs of one or more detected objects in the image. In predict_common.cpp I see that you draw circles with the points, so I have a "region". But is possible to get ROIs?

Reply • Share ›

Giseli • 4 years ago

Hi Roy, great post! I was trying to adapt this to detect only if a logo exist on the image. In the train dir is sufficient to have 10 images and in the test dir also some 15 images? Because I can't get the SVM classifier. What is the minimum number of images it needs to work nice?

Reply • Share ›

Roy ➚ Giseli • 4 years ago

@Giseli
I believe there is no minimum number of images, rather a minimum number of extracted features.
The code works to get a 1000-feature vocabulary (with k-means to obtain it), so obviously it needs more than 1000 extracted features.
But you are able to tweak the number of features in the vocabulary. For a small dataset with few categories even 20 features vocabulary can be fine... you really must experiment with the size of the vocabulary to see how it effects the recognition rates.

• Reply • Share ›

Yosafat • 4 years ago

Thanks a lot for your help. It helps me a lot to finish my project. ^_^

• Reply • Share ›

Amr • 5 years ago

Hello thanks a lot for the valuable help :)
i have one question though, i am using C# to make a program like yours and to use opencv i am using the emgu wrapper to wrap opencv in c#. the problem is i cant find the bag of words functions in emgu?? what should i do?

• Reply • Share ›

zernike • 5 years ago

Thank you for this code, it's very interessant,
I ask you, if that I can use just main.cpp to test this code, I can extract features and create vocabulary features, training and testing, it's all in main.cpp??
- what i need training.txt?I not need this file, what you think?
- in your database, I not find directory of TRAIN and TEST.

best anf thank you

• Reply • Share ›

LOL • 5 years ago

Note, the BoW class is finally working as of svn version r8551 or OpenCV 2.4.1 .

There are countless problems prior to this release, and I was surprised this site owner managed to get their code to even work.

Cheers,
J

• Reply • Share ›

jasmine • 5 years ago

back again
I've been trying to build the sln since the time I had commented here !
I asked many people and no one helped me and i tried to run the cmake on Linux but I'm not familiar with it that much
would you please uploaded for me after running the cmake on the files?
I really need to see the execution of your project and until now I can't even run it ! :(
thank u so much

• Reply • Share ›

jasmine • 5 years ago

I swear to god I did what you said but seems the cmake has problems itself . so I made a new project with those classes and headers, and I'm getting 100 unbelievable errors like
" 89 IntelliSense: variable "CV_INLINE" is not a type name c:\opencv2\core\core_c.h 182 1
"
in spit of I gave the VS 2010 all the paths of the include folders and the libs of the opencv !!
it's driving me crazy :(

• Reply • Share ›

Tony • 5 years ago

hi....
really it's a nice work
i have some question :$$
1- your files : every .cpp & .h have a main method ... so every couple give subproject ???
2- you said that i have to run Cmake to have .sln ... but it's not working :((
it gives alot of errors :(((
it ask for the file that contains the source ... i give it the directory of the food file....! that right ??

• Reply • Share ›

Roy ➚ Tony • 5 years ago

@Tony
The project is built using CMake, which will create a subproject for every executable.
Please use CMake.

@jasmine
To compile and use the project you must have a working version of OpenCV 2.3+.
Here are instructions for using OpenCV in applications: http://opencv.willowgarage.com/wiki/Getting_started

The problems you encounter are not problems in my code but in your environment. Sort that out and you are guaranteed to be able to run the programs.

Reply • Share ›

Roy • 5 years ago
@jasmine
I am happy to help you get it to compile, but you must be ready to do the work yourself.
I assume you are running windows. Try to run the CMake GUI program and direct the "Source directory" to the project directory.
Try to "Configure" and you will get specific errors about problems CMake encountered. Or "Generate" will be available right away and you can build the project using MSVS.

Reply • Share ›

jasmine • 5 years ago
well thank you so much , I just wanted to see your program , seems you can't upload it for me .

Reply • Share ›

jasmine • 5 years ago
Cmake is not working , that's all

Reply • Share ›

Roy • 5 years ago
@jasmine
I can't understand the problem from this error...
If you have CMake installed, go to the directory via command line and run "cmake ." and see where that takes you.

Reply • Share ›

jasmine • 5 years ago
I tried to make a new project but I'm getting 1454 error !
can you give me a hand ?

Reply • Share ›

jasmine • 5 years ago
it's not woking at all!
CMake Error: Unable to open cache file for save. C:/Program Files/CMake 2.8/CMakeCache.txt
CMake Error: The source directory "C:/Users/Safaa/Documents/Downloads/Compressed/royshil-FoodcamClassifier-4ba20bb" does not appear to contain CMakeLists.txt.
Specify --help for usage, or press the help button on the CMake GUI.
CMake Error: : System Error: No such file or directory

Reply • Share ›

jasmine • 5 years ago
Cmake didn't work ! would you please upload your project with the sln file :$, I'll be so thankful I wanna see the execution of this great application , we're working in the same zone right now .
my regards

Reply • Share ›

Roy ➔ jasmine • 5 years ago
@jasmine
CMake, in the long run, will be much much better than using the sln!
I suggest you give it another try
What are the errors you encounter?

Reply • Share ›

Lizards • 5 years ago

The dataset in foodcamimages/TRAIN etc. are missing from the foodcamimages.zip.

I am not sure if an unknown password for foodcamextractor.py to recover the image from the IMAP gmail account is really research friendly.

Perhaps it would be nice if somebody shared the image dataset in another structured repository similar to LabelMe.

Reply • Share ›

Girl Rock • 5 years ago

Hi, why I don't get the proper confusion matrix as you do? weird.. this is the output.. I'm using ur files train.txt and test.txt..
49 -> 49:0
50:0
51:0
52:0
53:0
54:0
55:0
56:0
57:0
97:0
cookies:33
fruit_veggie:31
indian:22
italian:45
mexican:3
misc:63
pizza:43

see more

Reply • Share ›

hew kim ➚ Girl Rock • a year ago

Girl Rock
HI I wonder do you have any idea how i should create the train.txt and test.txt by myself? anyway i still can not fully run the project provided by ROY..possible you can give me a guide?

Reply • Share ›

Girl Rock • 5 years ago

What's the point of doing ./make-test-background and produce background.png. Isn't to get more accurate Keypoints? I don't think it is much affected even not utilizing background.png

Reply • Share ›

zwfsgu • 5 years ago

in build_vocabolary.cpp
Rect clipping_rect = Rect(0,120,640,480-120);
img = img(clipping_rect);

the img size often less then clipping_rect

Reply • Share ›

Roy ➚ zwfsgu • 5 years ago

@zwfsgu
In that case what I usually do is "clip the clip": (a nice feature in OpenCV 2.3+ for Rect structs)

Rect clipping_rect = Rect(...);
Rect img_rect = Rect(0,0,img.cols,img.rows);
clipping_rect = clipping_rect & img_rect
img = img(clipping_rect); // <---- now this will never fail. fingers crossed.

Reply • Share ›

Support Our Work

Need help with your project?
Donate and let's talk about it!

Source code

- MoreThanTechnical Code Repository

Categories

Categories Select Category

Archives

- February 2017
- December 2016
- October 2016
- April 2016
- March 2016
- February 2016
- January 2016
- August 2015
- May 2015
- April 2015
- March 2015
- February 2015
- January 2015
- December 2014
- October 2014
- September 2014
- July 2014
- June 2014
- April 2014
- March 2014
- January 2014
- December 2013
- November 2013
- October 2013
- September 2013
- August 2013
- June 2013
- May 2013
- April 2013
- March 2013
- January 2013
- December 2012
- November 2012
- October 2012
- August 2012
- June 2012
- April 2012
- February 2012

- January 2012
- December 2011
- November 2011
- October 2011
- August 2011
- June 2011
- March 2011
- February 2011
- January 2011
- December 2010
- November 2010
- October 2010
- July 2010
- June 2010
- May 2010
- March 2010
- February 2010
- January 2010
- December 2009
- November 2009
- October 2009
- September 2009
- August 2009
- July 2009
- June 2009
- May 2009
- April 2009
- March 2009
- February 2009
- January 2009
- November 2008
- October 2008
- July 2008
- June 2008
- April 2008

Copyright

- Return to top