
The OpenCV User Guide

Release 2.3

August 15, 2011

CONTENTS

1	Operations with images	1
1.1	Input/Output	1
1.2	Basic operations with images	1
2	Features2d	5
2.1	Detectors	5
2.2	Descriptors	5
2.3	Matching keypoints	5
3	HighGUI	9
3.1	Using Kinect sensor	9

OPERATIONS WITH IMAGES

1.1 Input/Output

Images

Load an image from a file:

```
Mat img = imread(filename)
```

If you read a jpg file, a 3 channel image is created by default. If you need a grayscale image, use:

```
Mat img = imread(filename, 0);
```

Save an image to a file:

```
Mat img = imwrite(filename);
```

XML/YAML

TBD

1.2 Basic operations with images

Accessing pixel intensity values

In order to get pixel intensity value, you have to know the type of an image and the number of channels. Here is an example for a single channel grey scale image (type 8UC1) and pixel coordinates x and y:

```
Scalar intensity = img.at<uchar>(x, y);
```

`intensity.val[0]` contains a value from 0 to 255. Now let us consider a 3 channel image with BGR color ordering (the default format returned by `imread`):

```
Vec3b intensity = img.at<Vec3b>(x, y);  
uchar blue = intensity.val[0];  
uchar green = intensity.val[1];  
uchar red = intensity.val[2];
```

You can use the same method for floating-point images (for example, you can get such an image by running Sobel on a 3 channel image):

```
Vec3f intensity = img.at<Vec3f>(x, y);  
float blue = intensity.val[0];  
float green = intensity.val[1];  
float red = intensity.val[2];
```

The same method can be used to change pixel intensities:

```
img.at<uchar>(x, y) = 128;
```

There are functions in OpenCV, especially from calib3d module, such as `projectPoints`, that take an array of 2D or 3D points in the form of `Mat`. Matrix should contain exactly one column, each row corresponds to a point, matrix type should be 32FC2 or 32FC3 correspondingly. Such a matrix can be easily constructed from `std::vector`:

```
vector<Point2f> points;  
//... fill the array  
Mat pointsMat = Mat(points);
```

One can access a point in this matrix using the same method `texttt{Mat::at}`:

```
Point2f point = pointsMat.at<Point2f>(i, 0);
```

Memory management and reference counting

`Mat` is a structure that keeps matrix/image characteristics (rows and columns number, data type etc) and a pointer to data. So nothing prevents us from having several instances of `Mat` corresponding to the same data. A `Mat` keeps a reference count that tells if data has to be deallocated when a particular instance of `Mat` is destroyed. Here is an example of creating two matrices without copying data:

```
std::vector<Point3f> points;  
// .. fill the array  
Mat pointsMat = Mat(points).reshape(1);
```

As a result we get a 32FC1 matrix with 3 columns instead of 32FC3 matrix with 1 column. `pointsMat` uses data from `points` and will not deallocate the memory when destroyed. In this particular instance, however, developer has to make sure that lifetime of `points` is longer than of `pointsMat`. If we need to copy the data, this is done using, for example, `Mat::copyTo` or `Mat::clone`:

```
Mat img = imread("image.jpg");  
Mat img1 = img.clone();
```

To the contrary with C API where an output image had to be created by developer, an empty output `Mat` can be supplied to each function. Each implementation calls `Mat::create` for a destination matrix. This method allocates data for a matrix if it is empty. If it is not empty and has the correct size and type, the method does nothing. If, however, size or type are different from input arguments, the data is deallocated (and lost) and a new data is allocated. For example:

```
Mat img = imread("image.jpg");  
Mat sobelx;  
Sobel(img, sobelx, CV_32F, 1, 0);
```

Primitive operations

There is a number of convenient operators defined on a matrix. For example, here is how we can make a black image from an existing greyscale image `img`:

```
img = Scalar(0);
```

Selecting a region of interest:

```
Rect r(10, 10, 100, 100);
Mat smallImg = img(r);
```

A conversion from `texttt{Mat}` to C API data structures:

```
Mat img = imread("image.jpg");
IplImage img1 = img;
CvMat m = img;
```

Note that there is no data copying here.

Conversion from color to grey scale:

```
Mat img = imread("image.jpg"); // loading a 8UC3 image
Mat grey;
cvtColor(img, grey, CV_BGR2GRAY);
```

Change image type from 8UC1 to 32FC1:

```
src.convertTo(dst, CV_32F);
```

Visualizing images

It is very useful to see intermediate results of your algorithm during development process. OpenCV provides a convenient way of visualizing images. A 8U image can be shown using:

```
Mat img = imread("image.jpg");

namedWindow("image", CV_WINDOW_AUTOSIZE);
imshow("image", img);
waitKey();
```

A call to `waitKey()` starts a message passing cycle that waits for a key stroke in the "image" window. A 32F image needs to be converted to 8U type. For example:

```
Mat img = imread("image.jpg");
Mat grey;
cvtColor(img, grey, CV_BGR2GREY);

Mat sobelx;
Sobel(grey, sobelx, CV_32F, 1, 0);

double minVal, maxVal;
minMaxLoc(sobelx, &minVal, &maxVal); //find minimum and maximum intensities
Mat draw;
sobelx.convertTo(draw, CV_8U, 255.0/(maxVal - minVal), -minVal);

namedWindow("image", CV_WINDOW_AUTOSIZE);
imshow("image", draw);
waitKey();
```


FEATURES2D

2.1 Detectors

2.2 Descriptors

2.3 Matching keypoints

The code

We will start with a short sample `opencv/samples/cpp/matcher_simple.cpp`:

```
Mat img1 = imread(argv[1], CV_LOAD_IMAGE_GRAYSCALE);
Mat img2 = imread(argv[2], CV_LOAD_IMAGE_GRAYSCALE);
if(img1.empty() || img2.empty())
{
    printf("Can't read one of the images\n");
    return -1;
}

// detecting keypoints
SurfFeatureDetector detector(400);
vector<KeyPoint> keypoints1, keypoints2;
detector.detect(img1, keypoints1);
detector.detect(img2, keypoints2);

// computing descriptors
SurfDescriptorExtractor extractor;
Mat descriptors1, descriptors2;
extractor.compute(img1, keypoints1, descriptors1);
extractor.compute(img2, keypoints2, descriptors2);

// matching descriptors
BruteForceMatcher<L2<float>> matcher;
vector<DMatch> matches;
matcher.match(descriptors1, descriptors2, matches);

// drawing the results
namedWindow("matches", 1);
Mat img_matches;
drawMatches(img1, keypoints1, img2, keypoints2, matches, img_matches);
```

```
imshow("matches", img_matches);  
waitKey(0);
```

The code explained

Let us break the code down.

```
Mat img1 = imread(argv[1], CV_LOAD_IMAGE_GRAYSCALE);  
Mat img2 = imread(argv[2], CV_LOAD_IMAGE_GRAYSCALE);  
if(img1.empty() || img2.empty())  
{  
    printf("Can't read one of the images\n");  
    return -1;  
}
```

We load two images and check if they are loaded correctly.:

```
// detecting keypoints  
FastFeatureDetector detector(15);  
vector<KeyPoint> keypoints1, keypoints2;  
detector.detect(img1, keypoints1);  
detector.detect(img2, keypoints2);
```

First, we create an instance of a keypoint detector. All detectors inherit the abstract `FeatureDetector` interface, but the constructors are algorithm-dependent. The first argument to each detector usually controls the balance between the amount of keypoints and their stability. The range of values is different for different detectors (For instance, *FAST* threshold has the meaning of pixel intensity difference and usually varies in the region $[0,40]$. *SURF* threshold is applied to a Hessian of an image and usually takes on values larger than *100*), so use defaults in case of doubt.

```
// computing descriptors  
SurfDescriptorExtractor extractor;  
Mat descriptors1, descriptors2;  
extractor.compute(img1, keypoints1, descriptors1);  
extractor.compute(img2, keypoints2, descriptors2);
```

We create an instance of descriptor extractor. The most of OpenCV descriptors inherit `DescriptorExtractor` abstract interface. Then we compute descriptors for each of the keypoints. The output `Mat` of the `DescriptorExtractor::compute` method contains a descriptor in a row i for each i -th keypoint. Note that the method can modify the keypoints vector by removing the keypoints such that a descriptor for them is not defined (usually these are the keypoints near image border). The method makes sure that the output keypoints and descriptors are consistent with each other (so that the number of keypoints is equal to the descriptors row count).

```
// matching descriptors  
BruteForceMatcher<L2<float>> matcher;  
vector<DMatch> matches;  
matcher.match(descriptors1, descriptors2, matches);
```

Now that we have descriptors for both images, we can match them. First, we create a matcher that for each descriptor from image 2 does exhaustive search for the nearest descriptor in image 1 using Euclidean metric. Manhattan distance is also implemented as well as a Hamming distance for Brief descriptor. The output vector `matches` contains pairs of corresponding points indices.

```
// drawing the results  
namedWindow("matches", 1);  
Mat img_matches;  
drawMatches(img1, keypoints1, img2, keypoints2, matches, img_matches);
```

```
imshow("matches", img_matches);  
waitKey(0);
```

The final part of the sample is about visualizing the matching results.

HIGHGUI

3.1 Using Kinect sensor

Kinect sensor is supported through VideoCapture class. Depth map, RGB image and some other formats of Kinect output can be retrieved by using familiar interface of VideoCapture.

In order to use Kinect with OpenCV you should do the following preliminary steps:

1. Install OpenNI library (from here [url{http://www.openni.org/downloadfiles}](http://www.openni.org/downloadfiles)) and PrimeSensor Module for OpenNI (from here <https://github.com/avin2/SensorKinect>). The installation should be done to default folders listed in the instructions of these products, e.g.:

```
OpenNI:
  Linux & MacOSX:
    Libs into: /usr/lib
    Includes into: /usr/include/ni
  Windows:
    Libs into: c:/Program Files/OpenNI/Lib
    Includes into: c:/Program Files/OpenNI/Include
PrimeSensor Module:
  Linux & MacOSX:
    Bins into: /usr/bin
  Windows:
    Bins into: c:/Program Files/Prime Sense/Sensor/Bin
```

If one or both products were installed to the other folders, the user should change corresponding CMake variables OPENNI_LIB_DIR, OPENNI_INCLUDE_DIR or/and OPENNI_PRIME_SENSOR_MODULE_BIN_DIR.

2. Configure OpenCV with OpenNI support by setting `WITH_OPENNI` flag in CMake. If OpenNI is found in install folders OpenCV will be built with OpenNI library (see a status `OpenNI` in CMake log) whereas PrimeSensor Modules can not be found (see a status `OpenNI PrimeSensor Modules` in CMake log). Without PrimeSensor module OpenCV will be successfully compiled with OpenNI library, but VideoCapture object will not grab data from Kinect sensor.
3. Build OpenCV.

VideoCapture can retrieve the following Kinect data:

1. **data given from depth generator:**

- `OPENNI_DEPTH_MAP` - depth values in mm (`CV_16UC1`)
- `OPENNI_POINT_CLOUD_MAP` - XYZ in meters (`CV_32FC3`)
- `OPENNI_DISPARITY_MAP` - disparity in pixels (`CV_8UC1`)
- `OPENNI_DISPARITY_MAP_32F` - disparity in pixels (`CV_32FC1`)

- `OPENNI_VALID_DEPTH_MASK` - mask of valid pixels (not occluded, not shaded etc.) (`CV_8UC1`)

2. data given from RGB image generator:

- `OPENNI_BGR_IMAGE` - color image (`CV_8UC3`)
- `OPENNI_GRAY_IMAGE` - gray image (`CV_8UC1`)

In order to get depth map from Kinect use `VideoCapture::operator >>`, e. g.

```
VideoCapture capture( CV_CAP_OPENNI );
for(;;)
{
    Mat depthMap;
    capture >> depthMap;

    if( waitKey( 30 ) >= 0 )
        break;
}
```

For getting several Kinect maps use `VideoCapture::grab` and `VideoCapture::retrieve`, e.g.

```
VideoCapture capture(0); // or CV_CAP_OPENNI
for(;;)
{
    Mat depthMap;
    Mat rgbImage;

    capture.grab();

    capture.retrieve( depthMap, OPENNI_DEPTH_MAP );
    capture.retrieve( bgrImage, OPENNI_BGR_IMAGE );

    if( waitKey( 30 ) >= 0 )
        break;
}
```

For setting and getting some property of Kinect data generators use `VideoCapture::set` and `VideoCapture::get` methods respectively, e.g.

```
VideoCapture capture( CV_CAP_OPENNI );
capture.set( CV_CAP_OPENNI_IMAGE_GENERATOR_OUTPUT_MODE, CV_CAP_OPENNI_VGA_30HZ );
cout << "FPS " << capture.get( CV_CAP_OPENNI_IMAGE_GENERATOR+CV_CAP_PROP_FPS ) << endl;
```

Since two types of Kinect's data generators are supported (image generator and depth generator), there are two flags that should be used to set/get property of the needed generator:

- `CV_CAP_OPENNI_IMAGE_GENERATOR` – A flag for access to the image generator properties.
- `CV_CAP_OPENNI_DEPTH_GENERATOR` – A flag for access to the depth generator properties. This flag value is assumed by default if neither of the two possible values of the property is not set.

Flags specifying the needed generator type must be used in combination with particular generator property. The following properties of cameras available through OpenNI interfaces are supported:

- For image generator:
 - `CV_CAP_PROP_OPENNI_OUTPUT_MODE` – Two output modes are supported: `CV_CAP_OPENNI_VGA_30HZ` used by default (image generator returns images in VGA resolution with 30 FPS) or `CV_CAP_OPENNI_SXGA_15HZ` (image generator returns images in SXGA resolution with 15 FPS); depth generator's maps are always in VGA resolution.
- For depth generator:

- `CV_CAP_PROP_OPENNI_REGISTRATION` – Flag that synchronizes the remapping depth map to image map by changing depth generator's view point (if the flag is "on") or sets this view point to its normal one (if the flag is "off").

Next properties are available for getting only:

- `CV_CAP_PROP_OPENNI_FRAME_MAX_DEPTH` – A maximum supported depth of Kinect in mm.
 - `CV_CAP_PROP_OPENNI_BASELINE` – Baseline value in mm.
 - `CV_CAP_PROP_OPENNI_FOCAL_LENGTH` – A focal length in pixels.
 - `CV_CAP_PROP_FRAME_WIDTH` – Frame width in pixels.
 - `CV_CAP_PROP_FRAME_HEIGHT` – Frame height in pixels.
 - `CV_CAP_PROP_FPS` – Frame rate in FPS.
- Some typical flags combinations “generator type + property” are defined as single flags:
 - `CV_CAP_OPENNI_IMAGE_GENERATOR_OUTPUT_MODE` = `CV_CAP_OPENNI_IMAGE_GENERATOR` + `CV_CAP_PROP_OPENNI_OUTPUT_MODE`
 - `CV_CAP_OPENNI_DEPTH_GENERATOR_BASELINE` = `CV_CAP_OPENNI_DEPTH_GENERATOR` + `CV_CAP_PROP_OPENNI_BASELINE`
 - `CV_CAP_OPENNI_DEPTH_GENERATOR_FOCAL_LENGTH` = `CV_CAP_OPENNI_DEPTH_GENERATOR` + `CV_CAP_PROP_OPENNI_FOCAL_LENGTH`
 - `CV_CAP_OPENNI_DEPTH_GENERATOR_REGISTRATION_ON` = `CV_CAP_OPENNI_DEPTH_GENERATOR` + `CV_CAP_PROP_OPENNI_REGISTRATION_ON`

For more information please refer to a Kinect example of usage `kinect_maps.cpp` in `opencv/samples/cpp` folder.