

Documentation

High performance vessel analysis tools: 3D vessel reconstruction and analysis tool using whole-slide images.

Google Summer of Code 2015

Sayan Maity

Introduction

Extensive research has been performed to monitor the structural changes and spatial relationship of any diseased organ using whole slide histopathology images; resulted in a framework to study 3D whole-slide microscopy image datasets and reconstruct 3D vessel structure to analyze serial liver slices. The proposed standalone framework performs all the sequence of modules of, image registration, segmentation, vessel cross-section association, interpolation, and volumetric rendering. The simulation based research for the research study has been developed in Matlab®. To make this cutting-edge novel idea applicable to large dataset acquired from the patients biopsy result on a daily basis, in this project under Google Summer of Code 2015, an efficient and scalable version of HPC implementation is done using C++ programming language leveraging the powerful and rich open source 3rd party libraries such as openCV, Boost, Matlab api, Armadillo, Boost, Alglib, Cbc, VTK etc.

In this documentation, we will be explaining the pipeline framework components implemented for sequential processing: Registered Images=> Segmentation=> Vessel association => 3D Rendering. The HPC version of the two-step sequential image-based registration component is already been developed before Google Summer of Code 2015. The remaining steps of the pipeline framework have been implemented listed below:

Image Segmentation:

Two subsections comprised of the entire segmentation section:

- 1) Identify vessel tube probability map & produce initial level set function.
- 2) Segment vessel regions with level set method for the image sequence.

Two stage Vessel association:

Two subsections comprised of the entire vessel association section:

- 1) Local Bislade between the successive frames: 1-2;3-4;5-6.
- 2) Global association between the Local Bislade pairs; using the pair generated in the previous step: (1-2)<=> (3-4)<=> (5-6).

3D Rendering:

Two subsections comprised of the entire 3D rendering section:

- 1) Using the associated vessel and their boundaries create the interpolated boundaries using spline interpolation technique, and save the resulted boundaries as binary series of images (*.tif) .
- 2) Use the interpolated boundaries (stacks of series of binary images) to build tetrahedral mesh representation for 3D visualization.

The remainder of this project documentation is organized as follows: we will describe the implementation of all the three tasks: Image Segmentation, Two stage Vessel association & 3D Rendering. For each of them the implemented function definitions will be explained in details including: function definition in C++; technical functionality it will be performing, the data-type of the inputs & outputs of each function. In addition, the third party libraries used in the implementation of a task in which context will be listed in each task. At the end, a script map (flowchart) along with the data-type in the pipeline framework exchanged & output of the sections will be portrayed. An appendix section is added with the issues faced at the time of implementation.

Image Segmentation

II.I. Function Definition

To describe the function definition I have used tree like structure using “tab” to represent the dependencies of functions in color coded.

Name of function: File name=> “segmentation.cpp”

Inputs=>Registered images (“output_image.XX.bmp”)

Functionality=> Set the values for color de-convolution, stain i.e. Hematoxylin, DAB, and red_marker. *Just to note the 3 channels in a RGB image follow BGR order in OpenCV where as in Matlab the order is RGB.*

Output=> The generated initial level set function is saved in yml. Can be read to Matlab and convert to Mat file.

Name of function=> Ocv_ColorDeconvolution(vector< vector<double> > stains, cv::Mat img, Size s)

Inputs=> The stain vales defined by the calling functions; The Registered image ("output_image.XX.bmp") passed by the calling function; & the size of the image
Functionality=> Deconvolves multiple stains from RGB image 'img', given stain matrix 'stains'.

Output=> intensity - an intensity image in C++ double datatype. Each channel is a stain intensity. Channels are ordered same as columns of 'stain' following BGR order.

Name of function=>vec_transpose

Input=> vector<vector<double>>

Functionality => Transpose the 2 dimensional C++ STL vector.

Output=> vector<vector<double>>

Name of function=> ocv_im2vec

Input=> A openCV matrix (cv::Mat)

Functionality => Converts a RGB color image to 3 x MN matrix (Where M,N=> Size of the RGB color image)

Output=> 3 x MN dimensional C++ STL vector: vector<vector<int>>

Name of function=> vec_deconvolution_normalize

Input=> 3 x MN dimensional C++ STL vector: vector<vector<int>>

Functionality => Normalize raw color values according to Rufriok and Johnston's color deconvolution scheme.

Output=> 3 x MN dimensional C++ STL vector: vector<vector<double>>

Name of function=> vec_deconvolution_denormalize

Input=> 3 x MN dimensional C++ STL vector: vector<vector<int>>

Functionality => De-normalize raw color values according to Rufriok and Johnston's color deconvolution scheme.

Output=> 3 x MN dimensional C++ STL vector: vector<vector<double>>

Name of function=> reshpe_back(vector< vector<double> > , Size s)

Input=> 3 x MN dimensional C++ STL vector: vector<vector<int>> & value of M & N explicitly

Functionality => Reshape back 3 x MN dimensional C++ STL vector to MXN openCV matrix (cv::Mat)

Output=> MXN openCV matrix (cv::Mat)

Name of function=> ocv_multiScaleFilter2D (cv::Mat , Size s)

Inputs=> Single channel cv::Mat; the corresponding DAB stain channel & the size of the image

Functionality=> Perform filtering operation in multiple scale & angle. Store the filtered results generated of different scales & angles.

Output=> cv::Mat object containing the filtered result @ different scales & angles.

Name of function=> ocv_Hessian2D (cv::Mat, double sigmas)

Input=> Single channel cv::Mat; the corresponding DAB stain channel; & C++ double precision value 'sigmas' for the kernel generation.

Functionality => Generate kernel coordinates to Build the gaussian 2nd derivatives filters.

Output=> 3-channel cv::Mat containing the filtering result in xx, xy & yy directionality.

Name of function=> ocv_Hessian2D (cv::Mat, double sigmas)

Input=> Single channel cv::Mat; the corresponding DAB stain channel; & C++ double precision value 'sigmas' for the kernel generation.

Functionality => Generate kernel coordinates to Build the gaussian 2nd derivatives filters.

Output=> 3-channel cv::Mat containing the filtering result in xx, xy & yy directionality.

Name of function: File name=> "LSbatch.cpp"

Inputs=>Registered images ("output_image.XX.bmp"); the produced initial level set & color convoluted intensity images after filtering

Functionality=> Update the input level set function in iteration & Segment vessel regions with level set method for the image sequence.

Output=> The segmented vessel contour saved in yml. Can be read to Matlab and convert to Mat file.

Name of function=> matread(const char *file, std::vector<double>)

Inputs=>name of the matlab *.mat file variable to read & the vector to save it's content

Functionality=> Read a matlab *.mat file in C++ & store the data in double precision in a C++ STL vector.

Output=> C++ STL vector.

Name of function=> updatef(cv::Mat u,cv::Mat smoothImg,cv::Mat Pp,cv::Mat g, cv::Mat K,double epsilon, double sigma)

Inputs=>Multiple inputs set by the user,

Functionality=> Update f1, f2 (need to update level set function)

Output=> std::vector<cv::Mat > ; Containing cv::Mat object inside C++ STL vector.

Name of function=> Heaviside(cv::Mat u, double epsilon)

Inputs=>cv::Mat object the & the constant value.

Functionality=> Compute heaviside function.

Output=> cv::Mat

Name of function=> NeumannBoundCond(cv::Mat u)

Inputs=> cv::Mat object

Functionality=> Make a function satisfy Neumann boundary condition

Output=> cv::Mat object

Name of function=> div_norm(cv::Mat u)

Inputs=> cv::Mat object

Functionality=> compute curvature for u with central difference scheme.

Output=> cv::Mat object

Name of function=> Dirac(cv::Mat u, double epsilon)

Inputs=> cv::Mat object

Functionality=> Compute Dirac delta (The delta function is a generalized function that can be defined as the limit of a class of delta sequences. The delta function is sometimes called "Dirac's delta function" or the "impulse symbol")

Output=> cv::Mat object.

Name of function=> distReg_p2(cv::Mat u)

Inputs=> cv::Mat object

Functionality=> First compute first order derivative of the double-well potential & then compute the distance regularization term with the double-well potential.

Output=> Index of the associated vessel object.

II.II. External Libraries Used

Lirbaries Used	Functions served
Matlab api ("mat.h" header file)	To read vessel segmentation information from *.mat files.
OpenCV	To preprocess the segmentation binary images: image dilation; find close contours in the preprocessed binary images, generating kernel for image filtering, image filtering operation.
Boost	For accessing constant values such as pi, epsilon (compiler dependent)

Two stage Vessel association

III.I. Function Definition

Name of function: File name=> “vessel_assosiation.cpp”

Inputs=>Registered images (“output_image.XX.tif”); & the segmented vessel contour from the Segmentation operation.

Functionality=> Generate bi_slide vessel componets by performing a two level vessel objects association computation between the successive slides.

Output=> The segmented vessel contour saved in yml. Can be read to Matlab and convert to Mat file.

Name of function=> read_and_dialate(const char *imagenname, const char *seg_file1,cv::Mat)

Inputs=>Name of the image to read, name of the matlab *.mat file variable to read & the cv::Mat to save it’s content

Functionality=> Read a matlab *.mat file in C++ & store the data in double precision in a C++ STL vector and convert to cv::Mat object and apply image dialation.

Output=> cv::Mat object.

Name of function=> matread(const char *file, std::vector<double>)

Inputs=>name of the matlab *.mat file variable to read & the vector to save it’s content

Functionality=> Read a matlab *.mat file in C++ & store the data in double precision in a C++ STL vector.

Output=> C++ STL vector.

Name of function=> get_boundary_centroid(cv::Mat dialated)

Inputs=> The cv:: Mat component generated after reading & performing dilation operation.

Functionality=> Compute the boundary and centroid of the vessel object detected for the given image.

Output=> vector<vector<Point>> blobs_sorted (the vessel objects sorted by area), vector<double> bolb_area_sorted (area),vector<Point> bolb_center1_sorted (center), vector<Vec4i> hierarchy

Name of function=> get_FSDs(vector<double> X,vector<double> Y)

Inputs=> X & Y coordinates of the boundary of the vessel objects in C++ STL vector.

Functionality=> Calculate Fourier shape descriptors for given boundary.

Output=> vector<double> Fsd (Generated Fourier feature descriptor in a C++ STL vector)

Name of function=> calculate_likelihood(int i,vector<vector <double>> Fsd_slide_1,vector<Point> bolbcenter_1, vector <double> bolbarea_1,vector<vector <double>> Fsd_slide_2,vector<Point> bolbcenter_2, vector <double> bolbarea_2,int problem_id,vector<vector<Point>> blobs_2,vector<double> pre_vec_j,vector<vector<double>> vec3,)

Inputs=>Vessel objects Fourier shape descriptors; area, center.

Functionality=> Calculate the likelihood for each possible association with one-to-one, one-to-two and one-to-none and none-to-one cases.

Output=> vector<double> norm_p_likelihood, vector<vector<int>> C_coeff (Generated Fourier feature descriptor in a C++ STL vector)

Name of function=> get_likelihood_value(vector<double> feature1,Point centroid1,double area1,vector<double> feature2,Point centroid2,double area2,vector<double> cur_vec,vector<double> pre_vec_j,int type, int problem_id)

Inputs=> Vessel objects Fourier shape descriptors; area, center & corresponding problem ID of one-to-one, one-to-two and one-to-none and none-to-one cases.

Functionality=> To compute the value of likelihood by taking into account appearance similarity, spatial distance and trajectory smoothness (for MAP)..

Output=> double "like_values" (Likelihood values in C++ double precision).

Name of function=> get_matching_result(vector<double> vsol, vector<vector<int>> C_coeff, vector<vector <double>> Fsd_slide_1, vector<vector <double>> Fsd_slide_2, vector<Point> bolbcenter_1, vector<Point> bolbcenter_2)

Inputs=> Vessel objects Fourier shape descriptors; area, center & the integer programming optimized matched result.

Functionality=> To get the ids of matched objects in adjacent frame pairs.

Output=> vector<int> parent, vector<vector<int>> t_id (Index of the matched vessel object in a C++ STL container)

III.II. External Libraries Used

Lirbaries Used	Functions served
Matlab api ("mat.h" header file)	To read vessel segmentation information from *.mat files.
OpenCV	To preprocess the segmentation binary images: image dilation; find close contours in the preprocessed binary images etc.
Armadillo	To incorporate numerical functions such as interpolation to find equally spaced points on each closed contour.
Boost	For polygon operation such as union & intersection two compute one-to-two vessel object association.
Cbc	For mixed binary integer programming

3D Rendering

IV.I. Function Definition

Name of function: File name=> "vessel_assosiation.cpp"

Inputs=> After generating bi_slide vessel componets by performing a two level vessel objects association computation between the successive slides the associated id's of the objects been used.

Functionality=> Using the associated vessel and their boundaries create the interpolated boundaries using spline interpolation technique, and save the resulted boundaries as binary series of images (*.tif).

Output=> Series of *.tif 2d binary images.

Name of function: File name=> "vtk_ImageRendering.cpp"

Inputs=> Series of *.tif 2d binary images

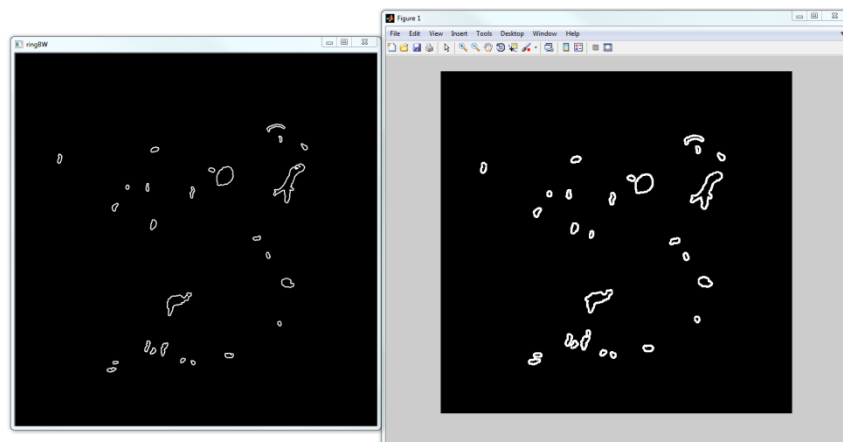
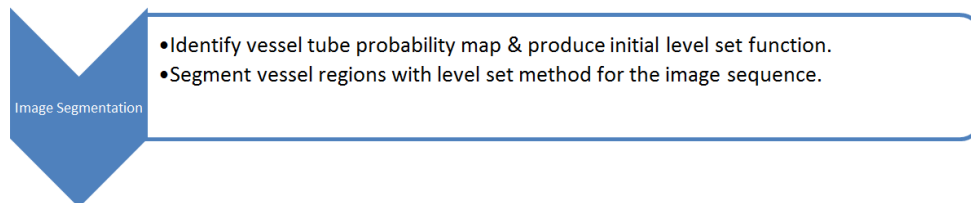
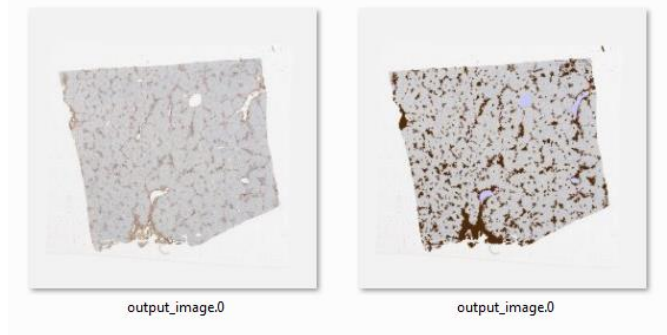
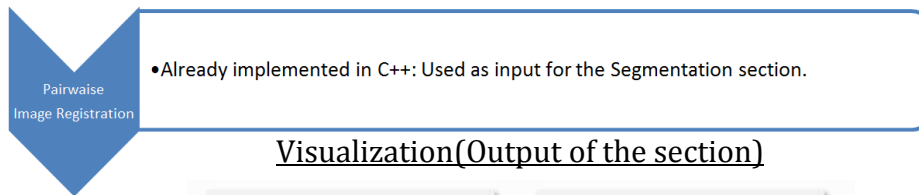
Functionality=> Use the interpolated boundaries (stacks of series of binary images) to build tetrahedral mesh representation for 3D visualization.

Output=> User interaction window tetrahedral mesh representation for 3D visualization.

IV.II. External Libraries Used

Lirbaries Used	Functions served
Alglib	Spline interpolation.
VTK	3D visualization of the vessel structure.

Implementation Flowchart



Left Output of C++ compared with Matlab right



Vessel object
association

- Local Bislide between the successive frames, for example: 1-2;3-4;5-6.
- Global association between the Local Bislide pairs; using the pair generated in the previous step: (1-2)<=> (3-4)<=> (5-6).

Visualization(Output of the section)

Parent Vessel Object ID	Associated Vessel Object ID in Successive frame
...	...
...	...

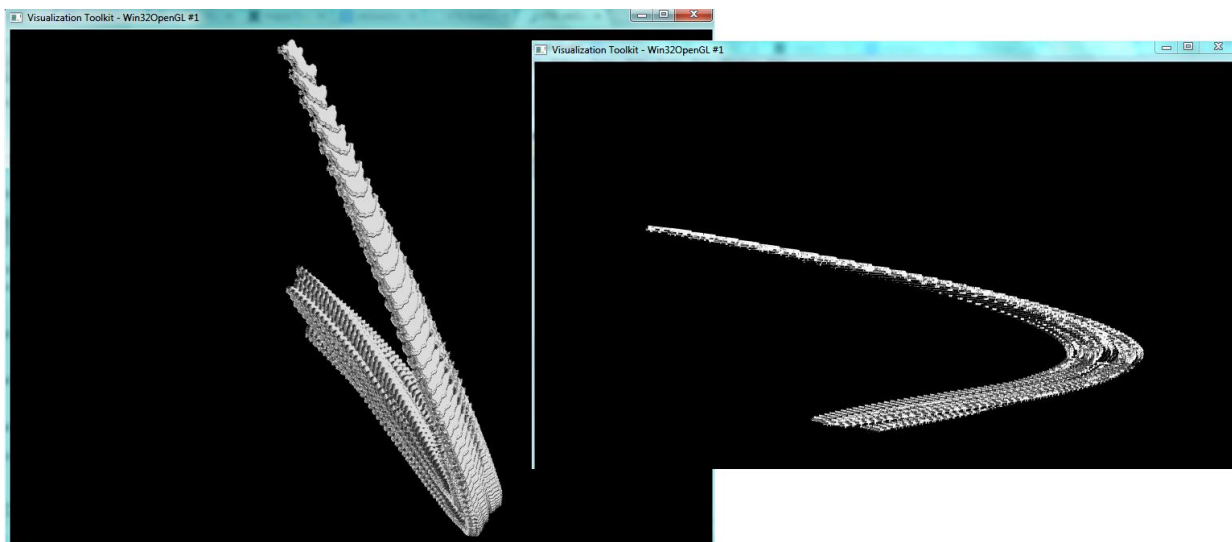
Abstract (toy) table structure to show the output of this section



3D Volume
rendering

- Using the associated vessel and their boundaries create the interpolated boundaries using spline interpolation technique, and save the resulted boundaries as binary series of images (*.tif) .
- Use the interpolated boundaries (stacks of series of binary images) to build tetrahedral mesh representation for 3D visualization.

Visualization(Output of the section)



User Interactive 3D vessel structure (3D window)

Appendix

While Implementing the C++ framework there were few discrepancies faced those will be listed below.

Vessel object detected from binary-segmented image in C++ & Matlab comparison with screenshots

The result screen shot of Variable explorer in Matlab and STL container in C++ (Visual Studio) to show the behavior of "bwboundaries()" in Matlab & "findContours()" function in OpenCV both used to detect closed contours in binary images.

In the "original_matlab\data_seg" directory for the segmented file "output_image.0.bmp" I am attaching the marked screenshot below:

Matlab "bwboundaries()" total 411 closed contours vessel object and 13 largest vessel objects (in areas):

The screenshot displays the MATLAB environment with the following components:

- Variables - sorted:** A table showing the first 13 largest vessel objects. The first row is highlighted in red.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	38746	7115	6501	5683	5.2725e+03	5.2365e+03	4592	3779	3.6235e+03	3.3125e+03	3304	3.0225e+03	2.8235e+03
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													

- Workspace:** Shows the variable `output_image.0` as a `1x411 double` array. The file path `C:\Users\Sayan\Dropbox\original_matlab\data_seg\output_image.0.bmp.mat` is visible.
- Command History:** Shows the execution of the `test_generate_3dvessel` function, including the `length(boundaries)` command.

OpenCV "findContours()" total 526 closed contours vessel object and 13 largest vessel objects (in areas): [Compare the above and below screenshot You can see the areas are matched & did one-one plot also to verify in details]

The screenshot shows a C++ IDE with the following components:

- Source File:** `interpolation_test.cpp`. The code defines a vector of pairs `vp` and sorts it by the second element (index) using `sort(vp.begin(), vp.end(), sort_using_greater_than)`. The sorted vector is then iterated over in a loop.
- Debugger:** A window titled `C:\Users\Sayan\Dropbox\GSoC_2015\Implementation\Read_data\cvtest\Debug\cvtest.exe` shows the state of the program. It displays the `vp` vector and the `vp[i].second` value for the current iteration. The `vp` vector is shown as a table of pairs.
- Call Stack:** A window titled `Call Stack` shows the sequence of function calls. The top call is `cvtest.exe\main(int argc, wchar_t** argv) Line 328`, followed by `cvtest.exe_tmainCRTStartup() Line 552 + 0x19 bytes`, and `kernel.exe\mainCRTStartup() Line 371`.

The `vp` vector data shown in the debugger is as follows:

Index	Value (Pair)
0	(38746.000000000000, 246)
1	(7115.000000000000, 84)
2	(6501.000000000000, 524)
3	(5683.000000000000, 475)
4	(5272.500000000000, 396)
5	(5236.500000000000, 221)
6	(4592.000000000000, 395)
7	(3779.000000000000, 226)
8	(3623.500000000000, 18)
9	(3312.500000000000, 450)
10	(3304.000000000000, 46)
11	(3022.500000000000, 141)
12	(2823.500000000000, 208)

Matlab "bwboundaries()" 9 smallest vessel objects (in areas):

The screenshot displays the MATLAB R2015a environment with three main panels:

- Variables - sorted:** A table showing the contents of the 'sorted' variable. The first row is highlighted with a red border. The data is as follows:

	403	404	405	406	407	408	409	410	411	412	413	414	415
sorted <1x411 double>	77.5000	77	77	76	75.5000	75.5000	72	70	69				
- Command Window:** Displays the message "All warnings have the state 'off'." followed by the prompt "K>".
- Workspace:** Shows the variable 'sorted' with its class 'double' and size '1x411'. The Command History panel below it shows the execution of the 'sorted' command and the 'help sortrows' command.

OpenCV "findContours()" 9 smallest vessel objects compared with Matlab "bwboundaries()" (in areas): [Compare the above and below screenshot You can see the areas are matched and the number of vessel objects are closely matched if we put threshold of area '70']

```

313 //**bolb center by Moments
314
315 bolb_center1.push_back(Point(int(moments(blobs[i]).m10/moments(blobs[i]).m00),int(moments(blobs[i]).m01/moments(blobs[i]).m00)));
316
317 }
318 // This is a vector of {value,index} pairs
319 vector<pair<double,size_t> > vp;
320 vp.reserve(bolb_area.size());
321 for (size_t i = 0 ; i != bolb_area.size() ; i++) {
322     vp.push_back(make_pair(bolb_area[i], i));
323 }
324 // Sorting will put lower values ahead of larger ones,
325 // resolving ties using the original index
326 sort(vp.begin(), vp.end(),sort_using_greater_than);
327 //sort(bolb_center1.begin(), bolb_center1.end(),sort_using_greater_than);
328 vector<Point> bolb_center1;
329 vector<Point> bolb_center2;
330 for (int i = 0 ; i != vp.size() ; i++) {
331     bolb_center1.push_back(Point(vp[i].second, vp[i].first));
332     bolb_center2.push_back(Point(vp[i].second, vp[i].first));
333     if (vp[i].first < 70) break;
334 }
335 }
336 vector<vector<Point>> bolbs_sort;
337 for (int i = 0 ; i != bolb_center1.size() ; i++) {
338     bolbs_sort.push_back(vector<Point>());
339     for (int j = 0 ; j != bolb_center2.size() ; j++) {
340         bolbs_sort[i].push_back(bolb_center2[j]);
341     }
342 }

```

Name	Value	Type
std::vector<std::pair<double, unsigned int>>	(-2.5301711256524607e-098, 1237016955311111083)	std::Vec
std::vector<std::pair<double, unsigned int>>	(77.500000000000000, 0)	std::Vec
bolb_center_sorted	[0]	std::vector
vp	[526]((38746.000000000000, 246),(7115.000000000000, 84),(6501.000000000000, 524),(5683.000000000000, 475),(5272.500000000000, 396),(5236.500000000000, 221),(4592.000000000000, 395),(3779.000000000000, 226),(3623.500000000000, 226))	std::vector

OpenCV "findContours()" detecting vessel object upto area "2.0" where we see before Matlab "bwboundaries()" detecting vessel object upto area "69.0"

```

313 //**bolb center by Moments
314
315 bolb_center1.push_back(Point(int(moments(blobs[i]).m10/moments(blobs[i]).m00),int(moments(blobs[i]).m01/moments(blobs[i]).m00)));
316
317 }
318 // This is a vector of {value,index} pairs
319 vector<pair<double,size_t> > vp;
320 vp.reserve(bolb_area.size());
321 for (size_t i = 0 ; i != bolb_area.size() ; i++) {
322     vp.push_back(make_pair(bolb_area[i], i));
323 }
324 // Sorting will put lower values ahead of larger ones,
325 // resolving ties using the original index
326 sort(vp.begin(), vp.end(),sort_using_greater_than);
327 //sort(bolb_center1.begin(), bolb_center1.end(),sort_using_greater_than);
328 vector<Point> bolb_center1;
329 vector<Point> bolb_center2;
330 for (int i = 0 ; i != vp.size() ; i++) {
331     bolb_center1.push_back(Point(vp[i].second, vp[i].first));
332     bolb_center2.push_back(Point(vp[i].second, vp[i].first));
333     if (vp[i].first < 2.0) break;
334 }
335 }
336 vector<vector<Point>> bolbs_sort;
337 for (int i = 0 ; i != bolb_center1.size() ; i++) {
338     bolbs_sort.push_back(vector<Point>());
339     for (int j = 0 ; j != bolb_center2.size() ; j++) {
340         bolbs_sort[i].push_back(bolb_center2[j]);
341     }
342 }

```

Name	Value	Type
std::vector<std::pair<double, unsigned int>>	(-2.5301711256524607e-098, 1237016955311111083)	std::Vec
std::vector<std::pair<double, unsigned int>>	(77.500000000000000, 0)	std::Vec
bolb_center_sorted	[0]	std::vector
vp	[526]((38746.000000000000, 246),(7115.000000000000, 84),(6501.000000000000, 524),(5683.000000000000, 475),(5272.500000000000, 396),(5236.500000000000, 221),(4592.000000000000, 395),(3779.000000000000, 226),(3623.500000000000, 226))	std::vector

Data type Precision Mismatch in C++ & Matlab comparison with screenshots

While working on the Segmentation section implementation in C++ I am mostly using Open CV library. I faced some issues regarding recreating the same result in C++ compare to Matlab due to precision mismatch of double values in two different compiler.

I did a thorough research and find out there are some existing issue that haven't been solved regarding Matlab & C++(OpenCV) Precision mismatch.

<http://stackoverflow.com/questions/11151609/precision-differences-in-matlab-and-c>

<http://stackoverflow.com/questions/7622012/matlab-and-c-differ-with-cos-function>

I found out:

In **Matlab** 'double' datatype set the CPU floating point precision to 80 bits.

Where-as in **C++**:

- 1) Type float, 32 bits long, has a precision of 7 digits.
- 2) Type double, 64 bits long, has a 15 digits precision.
- 3) Type long double is nominally 80 bits, though a given compiler/OS pairing may store it as 12-16 bytes for alignment purposes.

So to compare result in C++ with Matlab we are supposed to use Type long double in C++.

Here is the **bottle-neck**; **OpenCV** library has its own datatype and it can only handle up to precision of 64 bit :

- 1) CV_32F: "float"=> 32 bits long
- 2) CV_64F: "double"=> 64 bits long
- 3) **OpenCV didn't support long double.**

Based on the facts above when I'm trying to perform filtering operation (In 'multiScaleFilter2D.m' Matlab function) when ever the pixel values are hitting very small exponential (Please See attached Figure 1&2: **OpenCV can only go up-to e-16 compare to Matlab e-31**) any operation (such as squaring, square root & exponential) on those resulting in a huge variation in the result in C++ & OpenCV.

Thus I tried to match the resulted "Ifiltered" & "angles" variables in the 'multiScaleFilter2D.m' Matlab function in the corresponding C++ version:

Outcome 2) "Ifiltered" been going through squaring & exponential operation thus there is some substantial variation between the Matlab function & the corresponding C++ version :

OpenCV: 1673210 (Not too different)

Figure 1:

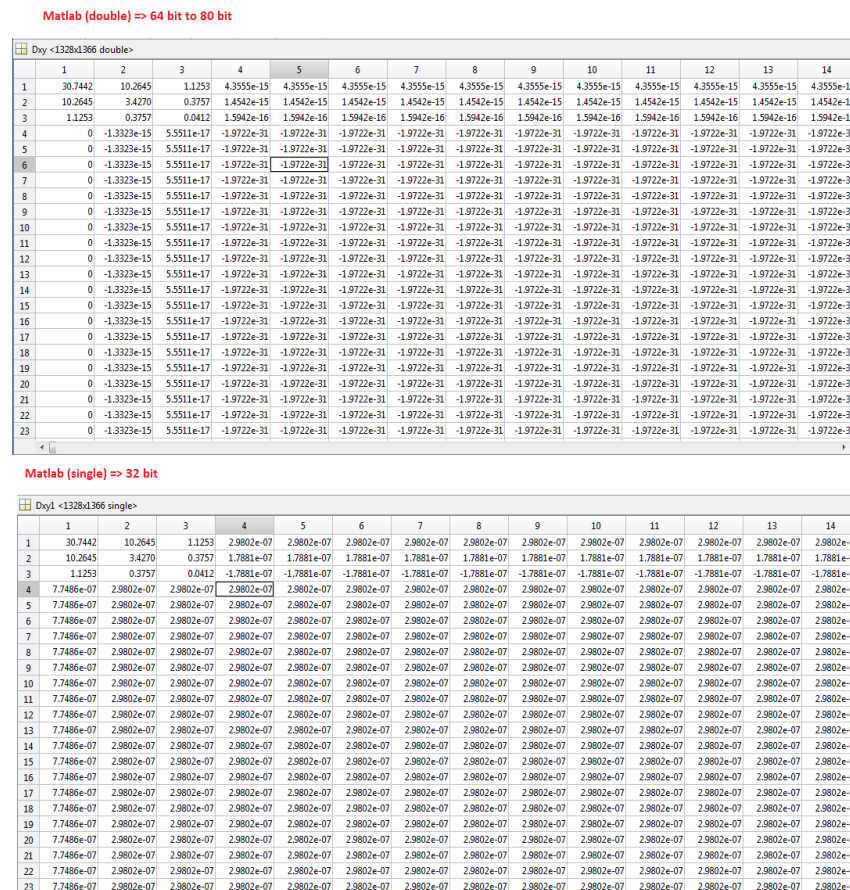
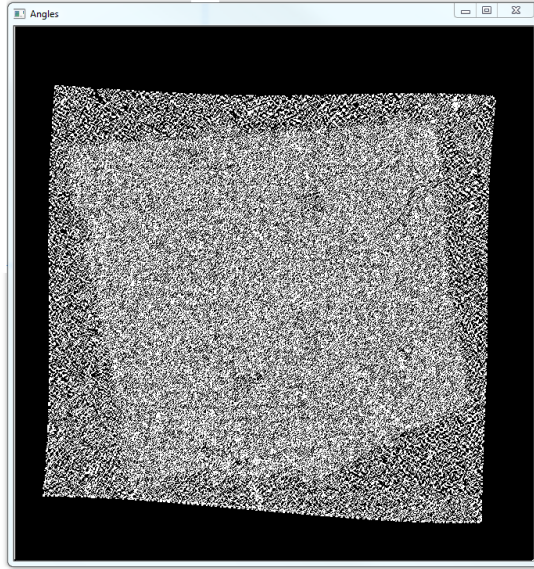


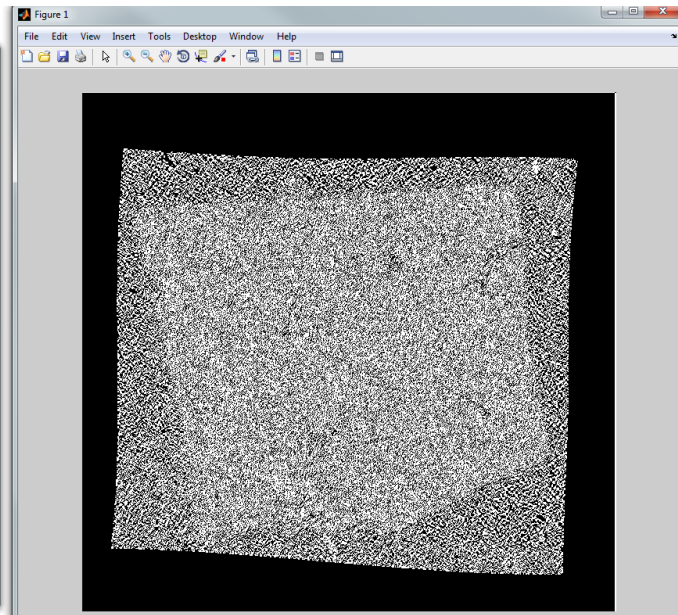
Figure 2:

[illegible][illegible]

Figure 3:



Open CV



Matlab

Figure 4:

