Filter Basket Analysis

Decent Espresso

Software to analyze images of espresso filter baskets, and assist in quantifying manufacturing quality.

John Weiss, Ph.D.

2/18/2016

Table of Contents

1. Introduction

Espresso filter baskets have a series of holes punched in the basket bottom. Ideally, these holes should be consistent in size and shape. But baskets vary greatly in terms of manufacturing quality.

This software provides a means of quantifying basket quality. Given an image of the filter basket bottom, the area of each hole is determined, and a histogram of hole areas is generated. This histogram can be plotted for quick visual determination of hole consistency.

2. Compilation

To build the software, you must have a C++ compiler installed, and all search paths set correctly.

- Unzip the software into a folder, and open a terminal window in the *src* directory.

- Using g++ (Linux or Windows), type

  ```
  make
  ```
  or
  ```
  g++ *.cpp -o fba
  ```

- Using Visual C++ (Windows), type

  ```
  nmake –f Makefile.vc
  ```
  or
  ```
  cl *.cpp -o fba
  ```

You may wish to move the resulting executable file (*fba* or *fba.exe*) to a directory on your executable program search path, so that you can run it simply by typing “fba”.

3. Program Usage

The filter basket analysis software is designed to be run from the command line, by issuing the following command:

```
fba basket.png [measurement] [bin size] [threshold] [contrast stretch]
```

This command can be scripted to run a set of images just as easily as one image. For example, to run all the basket images in the current folder:

```
for f in *.png; do fba $f; done      (Linux)
for %f in (*.png) do fba %f          (Windows)
```

4. Program parameters

basket.png
The only required program input is a filter basket image in PNG format. Filter basket images must feature a dark basket on a light background. Holes should be light and in focus. *Correct lighting and exposure is essential to proper functioning of the software.* If possible, crop the image to contain only the dark filter basket on a white surround. Placing the filter basket on a light box, with low ambient illumination, is ideal. Resolutions of at least 2Mpixel are recommended. Higher resolution images give more accurate results, but take longer to process.

measurement
The first command-line option allows conversion from pixels to standard units of measurement (such as millimeters). Hole areas and diameters are measured in pixels. If a physical measurement is taken on the basket to determine the maximum horizontal distance between holes, the program will automatically convert pixels into the same units as the measurement. The default value of 0 does no conversion.

This conversion can also be done with spreadsheet functionality. A sample Excel spreadsheet (*sample.xlsx*) is provided for illustration. This is discussed in detail later in the document.

bin size
The second command-line option sets the histogram bin size. Basket analysis produces a list of hole areas (and diameters). To plot this data as a histogram of hole sizes, the holes are sorted by area, and a binned histogram is generated. The default is to determine bin size automatically, and generate a histogram with 11 bins. You may override this default behavior by entering a different bin size. For example, a size of 5 will count the number of holes with area 0-4, 5-9, 10-14, etc. A value of 0 will invoke default behavior (automatically determine the bin size). The actual bin size is noted in the CSV file.

threshold
The third command-line option sets the binary threshold for segmenting foreground and background pixels. Grayscale pixel intensities range from 0 (black) to 255 (white). By default, the threshold is set to 128 (middle gray). If the lighting and exposure are off, it may be necessary to adjust this threshold for optimal results. Lighter images may benefit from raising the threshold. Darker images may benefit from lowering the threshold. Note that changing the threshold may alter the detected hole areas (higher thresholds may result in smaller hole areas).

contrast stretch
The fourth command-line option turns on automatic contrast stretching. This helps compensate for poor exposure, by ensuring that the darkest areas are mapped to pixel intensity 0, and the lightest areas are mapped to pixel intensity 255. By default, contrast stretching is turned off. Supplying any value for this option (e.g., "true") will turn it on.

Command-line options must be supplied in the correct order. For example, if the basket measurement is 50mm, you wish to set the binary threshold to 160 (instead of 128), retain the automatic histogram bin size, and leave off contrast stretch:

```
fba basket.png 50 0 160
```

5. Program output

For an input image *basket.png*, the program will produce two output files: *basket.csv* and *basket.cc.png*.

CSV file

basket.csv is a CSV (comma separated value) file, suitable for examination in a text editor or for import into a spreadsheet program such as Excel. The CSV file includes the following information:

- row 1: filter basket image name and dimensions
- row 2: binary threshold
- row 3: number of connected components detected (i.e., number of holes)
- row 4: sorted list of hole areas
- rows 6-12: summary statistics about hole areas (range, avg, stddev, etc.)
- rows 14-16: conversion factors
- row 18: histogram bin size
- rows 19-31: binned histogram of hole areas and diameters, suitable for scatter plotting

Connected components file

basket.cc.png is a pseudocolor image of the labeled connected components, which correspond to detected holes in the filter basket. It is primarily for verification purposes. Detected holes are color-coded, and a box is drawn around the area of holes. By comparing this image with the original basket image, you can quickly and easily evaluate hole detection accuracy. If any filter basket holes are not detected, or non-hole areas are labeled, or some hole areas are not correctly marked, then there are problems with the analysis. The best option is typically to take another image of the basket. But if acquiring another image is impractical, you may be able to correct these problems by changing the binary threshold and/or applying a contrast stretch (see below for examples).

6. Conversion formulas

The software determines basket hole area by counting the number of pixels in each hole. Assuming circular holes, we can compute the radius and diameter of the hole:

$$A = \pi r^2 \qquad \rightarrow \qquad r = \sqrt{\frac{A}{\pi}} \qquad \rightarrow \qquad d = 2r$$

These values are computed from the image, and are all in pixels. To convert from pixels to standard units such as millimeters, we need a physical measurement to relate to the image. If you measure the maximum distance between holes, either horizontally or vertically across the basket bottom, the software will do the conversion for you. The conversion factor is simply the ratio of the measurement in standard units (e.g., mm) to the number of pixels across the same distance:

$$f = \frac{units}{pixels}$$

The number of pixels is computed by the software, but you must provide the physical measurement. The units are up to you; millimeters is a good choice, but if you prefer inches, then provide the measurement in inches. To convert linear measurements from pixels to standard units, simply multiply the conversion factor times the diameter (or radius):

$$d \times f$$

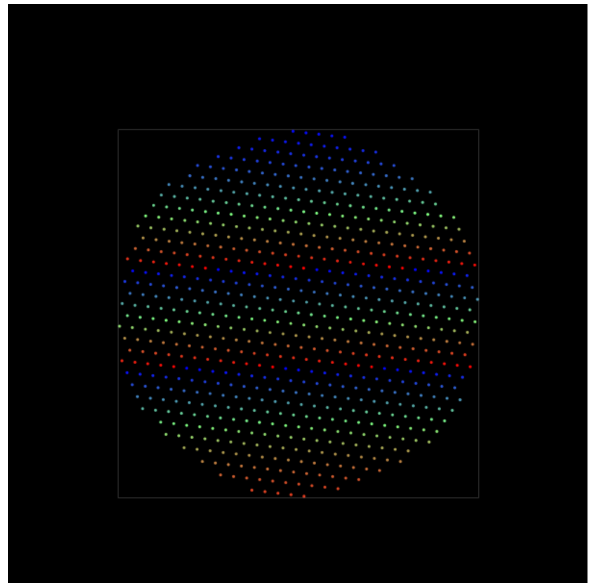
Area measurements are in squared units (mm^2, $inch^2$, etc.), so you must use the square of the conversion factor:

$$d \times f^2 = d \times f \times f$$

The software does this for you, but it can also be done in a spreadsheet. A sample spreadsheet is provided for illustration (*sample.xlsx*). This spreadsheet was created from the CSV output of basket analysis. Rows 15 and 16 give the horizontal and vertical distance in pixels and standard units, followed by the conversion factor. (The software uses the horizontal distance, but either one should work.) Below the standard CSV output, the basket hole area and diameter histograms were duplicated. The computation of area and diameter in standard units (assuming a measurement of 50mm) was performed using the above formulas. The values obtained are exactly the same as those provided by the software (compare rows 20-31 with rows 34-45).

7. Example usage: *fba DE5108-2.png 50*

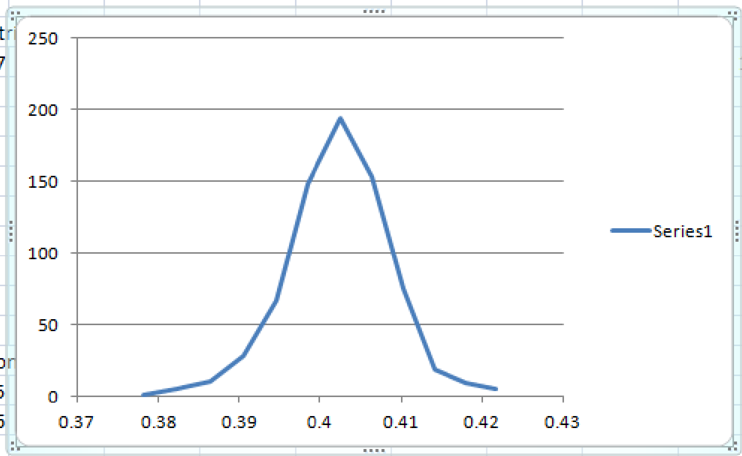
DE5108-2.png



DE5108-2.cc.png

sample.xlsx

	A	B	C	D	E	F	G	H	I	J	K	L
1	filter bask	3201 rows	3201 columns									
2	binary threshold: 128											
3	715 connected components (sorted and tr											
4	182	184	185	187								190
5												
6	Summary Statistics											
7	n	715										
8	min	182										
9	max	225										
10	mean	205.3										
11	stddev	6.4										
12	median	206										
13												
14		pixels	units	conversion								
15	horizontal	2002	50	0.024975								
16	vertical cc	2042	50	0.024486								
17												
18	histogram with bin size 4											
19	area(pixel	n		area(units	n		diam(pixe	n		diam(unit	n	
20	180	1		0.112275	1		15.14	1		0.378092	1	
21	184	5		0.11477	5		15.31	5		0.38227	5	
22	188	11		0.117265	11		15.47	11		0.386402	11	
23	192	28		0.11976	28		15.64	28		0.390492	28	
24	196	67		0.122255	67		15.8	67		0.394538	67	
25	200	148		0.12475	148		15.96	148		0.398544	148	
26	204	194		0.127245	194		16.12	194		0.402509	194	
27	208	153		0.12974	153		16.27	153		0.406436	153	
28	212	75		0.132235	75		16.43	75		0.410326	75	
29	216	19		0.13473	19		16.58	19		0.414179	19	
30	220	9		0.137225	9		16.74	9		0.417996	9	
31	224	5		0.13972	5		16.89	5		0.421779	5	
32												



DE5108.csv (Excel screenshot)

8. Adjusting default parameter values

fba image.png [measurement] [bin_size] [threshold] [contrast_stretch]

The only required parameter is the filter basket image filename (basket.png). Generally you will want to provide a measurement, to convert pixels into standard units. Occasionally you may wish to alter other parameters. For example, if 11 bins in the histogram does not yield a satisfactory plot of hole sizes, you can try manually adjusting the bin size. Analysis problems can often be addressed with a manually specified threshold can sometimes provide a solution. This is illustrated by the provided sample images.

DE5108-0.png:

- good sharp image, defaults work
- uncropped large image (21Mpixel) slows down analysis, better if cropped
  fba.exe DE5108-0.png

DE5108-1.png:

- tightly cropped version of DE5108-0.png
- basket analysis works, but small bright region above right of basket is classed as a hole
- solution: use slightly higher threshold than default of 128
  fba.exe DE5108-1.png 50 0 140

DE5108-2.png:

- ideally cropped version of DE5108-0.png: dark basket, bright surround, bright holes
- basket analysis works using default parameter values
  fba.exe DE5108-2.png

basket01.png:

- blurry low-rez image, but can still be analyzed using default parameter values
  fba.exe basket01.png

basket02.png:

- image with high ambient illumination, can still be analyzed using default parameter values
  fba.exe basket02.png

basket03.png:

- blurry, poorly lit, low-rez image, but can still be analyzed
- problem: bright areas on rim are classed as basket holes
- solution: use higher threshold, e.g. 165 (found by trial and error):
  fba.exe basket03.png 0 0 165

basket04.png:

- good sharp image
- problem: many small highlight areas on basket rim mess up the analysis
- solution: use higher threshold, e.g. 165 (found by trial and error):
  fba.exe basket04.png 0 0 165

basket05.png:

- image is completely unsuitable for analysis, does not have light holes on dark basket
- solution: retake image with lighting from below