



CHALMERS

Master of Science in Intelligent Systems Design
Department of Applied Information Technology

Report No. 2010:XXXX
ISSN: 1651-4749XXX

EFFICIENT FEATURES FOR VISUAL OBJECTS CLASSIFICATION

Azadeh Ghafarinia

Department of Applied Information Technology
Chalmers University of Technology



CHALMERS UNIVERSITY

Chalmers University of Technology
Göteborg, Sweden 2010

Keywords: Image Classification, Feature Extraction, Distance measurement, Battery Sorting

Acknowledgment

Thanks to Claes Strannegård who guided me in this research.

Thanks to Josef Fourier for his brilliant equations.

Thanks to dear Dorreh Dormishian who helped me in writing of this report.

Index of Figures:

Figure 1: Manual battery sorting.....	1
Figure 2: Human Eye Structure (4)	3
Figure 3: An image of a square plotted as 3-d surface and its FFT.....	7
Figure 4: A battery image and a 3d visualization of its grayscale counterpart fft.....	8
Figure 5: A band of FFT represented in previous picture and a section of this band	8
Figure 6: Sample battery images from dataset	9
Figure 7: Simmilar trend of the distance having a training set of 11 (Blue), and 3 (Red)	9
Figure 8: Nearest Neighbor Classifier: (Up) the algorithm; (Down) the pattern represented by the vector $X = (X1, X2)$ is identified with that class associated with the closest stored point: $X1, X2, X3$, (i.e., $C2$). All points in the shaded area are closer to $X2$ than other stored points.	11
Figure 9: a) a typical signal for the current problem, b) a similar signal but from a different type, c) a similar signal from the same type, d) a different signal from a different type.....	12
Figure 10: Performance rate for feature (RGB HISTOgram).....	14
Figure 11: Performance rate for feature (FFT Section)	15
Figure 12: Performance rate for feature (RGB Illumination invariant and fft section)	16
Figure 13: A batery image in three sections.....	16
Figure 14: Performance rate for feature (Three pieces color invariant histogram).....	17
Figure 15: Performance rate for feature (Three pieces color invariant histogram and FFT Slice)	18
Figure 16: Performance rate for feature (Three pieces color invariant hist. Reduced colors)	19
Figure 17: Performance rate for feature (Three pieces color invariant hist. Reduced colors with FFT Slice)...	20
Figure 18: Performance rate for feature (Color feature and fourier transform band)	20
Figure 19: Performance rate for feature (fourier transform band)	21
Figure 20: Diiferent faces of the same battery, showing how uneven a label can be	23
Figure 21: Application which was developed for this study.....	24

CONTENTS

1	INTRODUCTION	1
1.1	BACKGROUND.....	1
1.2	PROBLEM DESCRIPTION	2
1.3	HUMAN VISION SYSTEM	2
1.4	PATTERN RECOGNITION.....	3
1.4.1	<i>Feature extraction</i>	4
1.5	VISUAL FEATURES.....	4
1.5.1	<i>Color Space</i>	4
1.5.2	<i>Color Histogram</i>	5
1.6	FOURIER TRANSFORM	5
1.6.1	<i>Theorems</i>	7
1.6.2	<i>Equations</i>	6
1.6.3	<i>FFTW (Fastest Fourier Transform in the west)</i>	6
2	EXPERIMENTATION	8
2.1	DATA SET	8
2.2	CLASSIFICATION METHOD	10
2.2.1	<i>Distance Function</i>	11
2.3	ANALYZED FEATURES AND RESULTS	12
2.3.1	<i>RGB Color Histogram</i>	13
2.3.2	<i>Fourier Transform Section</i>	14
2.3.3	<i>RGB Illumination Invariant and Fourier Transform</i>	15
2.3.4	<i>Three Pieces of Histogram</i>	16
2.3.5	<i>Three Pieces of Histogram and reduced color percision</i>	18
2.3.6	<i>Color feature and Fourier Transform Band</i>	20
2.3.7	<i>Fourier transform band</i>	20
2.3.8	<i>Unbalanced Training Set</i>	21
2.4	TIMING.....	23
3	DISCUSSION.....	24
4	CONCLUSION	24
5	FUTURE WORK	25

1 INTRODUCTION

Efficient features for visual objects classification is the subject of this study which tries to introduce a practical method for classification of batteries based on their company and battery type. This has been an ongoing research work for the last three years. An overview of the problem and the current status of the research have been provided in this section. Next the framework of the experimentation and the set of tests are elaborated.

1.1 BACKGROUND

Recycling of batteries which requests optical objects classification is the major concern of an ongoing project at a Chalmers Innovation company named Optisort. Because of the diverse chemicals in use and dissimilar recycling process for different battery types, there must be some sort of classification. The current method of sorting the batteries is to pass them on a conveyor belt and get them recognized and sorted by human workers. The job can be considered as hazardous for oxidized batteries and their poisonous vapors (1). Figure 1 show how the job is being done at a recycling station (2).



FIGURE 1: MANUAL BATTERY SORTING

Many different methods and technologies are in use for automatic battery sorting in different countries. In The Netherlands the sorting process uses a sensor which is made by Philips. The sensor hires the magnetic field examiners and is reported to have 99% accuracy. It takes two seconds to inspect a battery and can process one battery at a time which is a slow rate. There is ongoing research work to improve this system. A new sorting line was developed, based on three process steps. "At the first step, mixtures of batteries pass through different sieves where the button batteries are separated. The remaining batteries pass through a magnetic separator. The non-magnetic batteries are the zinc-carbon batteries with casings of paperboard. The magnetic ones pass then through another sieve system, where the

batteries are separated by shape (prismatic and cylindrical) and by size (AAA, AA, C and D). Finally, on a third step, the batteries pass through different sensors (TRI-MAG) that are able to separate different batteries according to their dimensions, mass and electromagnetic properties. An UV sensor separates batteries with a label indicating the presence of mercury” (3). In Germany a visual pattern recognition system was created for battery sorting and is called batteriessortieranlage (BSA). This system is able to process the information on the label, e.g. manufacturer, materials, etc. This system is also 99% precise and has an advantage over the previous system which is its processing rate. The rate is 24 batteries/s. “After a pre-selection from battery sets of the same size and shape, the batteries are sent individually to the detection system. By comparing with a schematic example, the batteries can be placed automatically in the right collection box” (3).

There is another system in Germany (SORBAREC) examines the X-ray images for the batteries to sort them. “This method is able to sort zinc-carbon, manganese-alkaline, nickel-cadmium, nickel-metal-hydride, lithium and mercury. In this process, after hand and size sorting, the batteries are separated from a stock silo via different conveyor belts and fed to the X-ray sensor. The radiography unit consists of an X-ray tube and a sensor installed in a radiation protection cabin. The electrochemical battery type is identified in real time. The batteries fall off the conveyor belt and are pushed out of their trajectory by compressed air blasts from the side or from above. In this fashion several fractions can be reliably separated. Sorting speeds of up to 12 batteries/s are achieved with battery intervals of approximately 7 mm. The analysis is performed by computer, which likewise identifies the battery types based on the gray levels of the X-ray image” (3).

1.2 PROBLEM DESCRIPTION

Consider a system with an input and an output which, the input is a single image which is taken from one side of a battery. The image is in 24bit BMP format and has a size of 120 x 480. The output must be the exact battery class, i.e. the brand and the subtype of battery. For example an output of this system can be Energizer_Metalic, which denotes a certain battery class. There are 34 battery classes for this experimentation but the solution might be generalized to a wider set of batteries.

1.3 HUMAN VISION SYSTEM

To achieve a good interpretation of spatial data, in order to survive or defend, human vision system has evolved over a long period of time. Though the functional mechanisms of human vision have not been completely revealed, there are exact models of specific parts of this system (4).

In a typical case of human vision action, the visual information is sensed through the eye. As shown in a cross section of human-eye in figure-2, different parts of eye perform in different roles till the information is collected to pass to optical nerves. To be able to change the focus on specific items, ciliary muscles are holding the lens as a means of movement. The amount of light passed through the lens is controlled by iris, or pupil, same as an aperture on a camera. A transparent coating called cornea (sclera) provides the protection for these faint parts. The blood vessels in choroid supply nutrition, besides choroids’ opaqueness prevent extra light being passed inside the eye. After passing through the lens,

the light falls over retina and by fixing the eye position and angle, and changing the shape of the lens to find the best focus style, the image is formed on the focal point (fovea). Fovea contains the bulk of sensors despite the blind spot which is the starting of optic nerve and holds no sensors. There are around 100 million sensors located over the retina. Photochemical transmissions caused by these sensors due to light change, causes nerve impulses and therefore a signal is formed and transmitted by the eye to the brain (4).

There are two types of vision sensors in retina, the rods sensitive to black and white and the cones sensitive to color. There are three types of cones; α (to sense the blue light), β (to sense the green light), and γ (to sense the red light). A summation of these three sensor responses covers the whole of the visual spectrum. In low levels of light images are sensed with the rods. For more details concerning visual perception one may refer to Cornsweet (1970) (4).

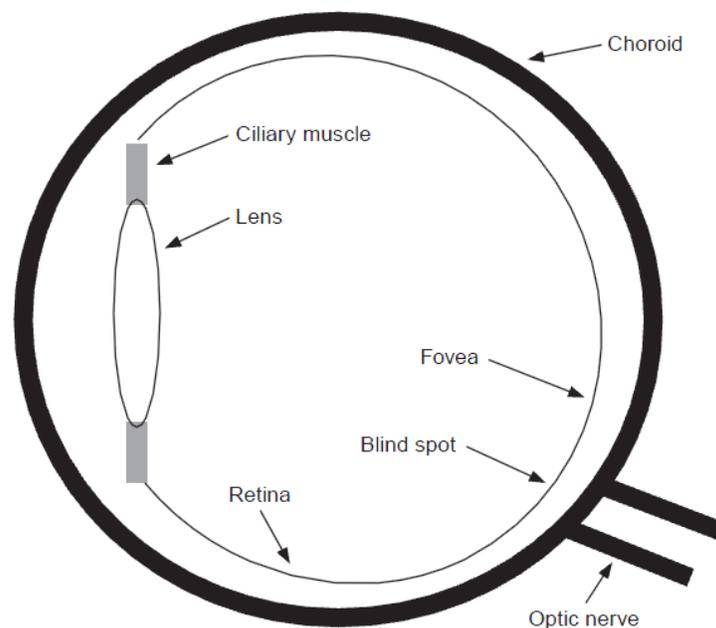


FIGURE 2: HUMAN EYE STRUCTURE (4)

The response of cones and the rods is altered by a logarithmic function and multiplied by a weighting factor to add a control over the contribution of separate sensors. The adjustment of weighting factors yields to have particular filtering properties (4).

1.4 PATTERN RECOGNITION

A well expressing definition for pattern is “Pattern can be defined as a quantitative or structural description of an object or some other entity of interest” (5). Then a pattern class can be defined as “a set of patterns that share some properties in common” (5). Having a sample of input data, the process

of classifying it as a member of other existing categories is called *pattern recognition*. The item to be recognized is either concrete or abstract. Concrete items have a visual or spatial representation, i.e. road maps, speech waveform, image, etc. Abstract items are conceptual and are not represented with concrete data, i.e. the similarity of two works of a musician (5). The image classification fits in concrete items pattern recognition.

Pattern recognition can be performed as a supervised operation, unsupervised operation, or with a neural network model. Three basic steps in the pattern recognition process are data acquisition, preprocessing, and decision making (5).

Any kind of information which identifies an entity is called a *feature* and the act of spotting the features is called *feature extraction*. In the next section feature extraction is elaborated for image processing as the field of interest for this study.

1.4.1 FEATURE EXTRACTION

A pattern or feature can be represented with a set of parallel measurements (6). An image has a considerable amount of visual data. Due to the computational limitations, there must be a limited subset of data which enables the observer to analyze and compare. It is also important to decide which key information represents the image in the best way. A set of features can be useful in an application and useless in others; therefore it is important to identify those features which suit a specific application. A feature can be represented with a vector denoting its numerical values:

$$F = \{f_1, f_2, \dots, f_n\},$$

n is the number of numerical elements in the feature F

1.5 VISUAL FEATURES

A computer image is a two dimensional matrix of pixels. Each pixel value is a representative of its brightness and color. The matrix with M pixels in height and N pixels in width, and having d-bits to describe the brightness values has a size of M*N*d bits. Having d bits for representation of a pixel value allows having a range of 0 to $2^d - 1$ values.

An image can be classified based on high-level or low-level features. Those features which are corresponding to the image information, e.g. color statistics, and are not related to the shape and content, are called low-level features (4). On the other hand, those features which correspond to the content of image, e.g. a company logo, are called high-level features. In the case of battery image classification, if there was an access to the whole label of a battery, high-level features could be defined and used (1). To get familiar with some low-level features a few basic terms and definitions are concisely presented.

1.5.1 COLOR SPACE

The pattern, of which the colors are distributed over an image, is a good cue to distinguish the class of image. Relying on just color might not be the sufficing distinctive attribute of an image from another, but it still has advantages in classification process. (6)

There are different encoding methods or color models for a color histogram. A color model describes how a pixel color is represented and how it relates to the whole image attributes. A set of all possible colors with a color model which forms a color space is called *gamut*.

A well-known color space is RGB, which the color is encoded to its red, green and blue components. RGB model is vulnerable to illumination; as the illumination is increased, the three components magnitudes are increased accordingly. To get around this problem Berens et al. (6) suggests a brightness-independent mapping of RGB values as below:

$$r = \frac{R}{R + G + B}$$

$$g = \frac{G}{R + G + B}$$

$$b = \frac{B}{R + G + B}$$

Based on the three equations:

$$b = 1 - r - g$$

According on above equations, a pixel can be represented with a 2-tuple (r, g), therefore less space is required to represent a pixel. There are many color spaces each representing an image with different numerical values for example based on hue, saturation, etc. but in this work the focus has been put on brightness-independent RGB as represented above.

1.5.2 COLOR HISTOGRAM

The intensity histogram shows how individual brightness levels are occupied in an image; the image contrast is measured by the range of brightness levels. The histogram plots the number of pixels with a particular brightness level. Having a color value of d bits per pixel, which allows having 2^d different numerical brightness representations, will led to a histogram vector representation as:

$$H = \{C_1, C_2, \dots, C_n\}, \quad n = 2^d$$

which C_i is the number of pixels with intensity value of i

1.6 FOURIER TRANSFORM

The *Fourier transform*, like many other linear transforms provides a means of solving the problems in the linear space. It alters the problem from its original domain to a new form which is solvable. It has an

important role in many branches of science. Fourier transform has had many applications since its representation by Josef Fourier in the late 18th century (7).

Any optical, electrical or acoustical wave and its spectrum, both have valuable measurable information. A Fourier transform is a representation of a signal or waveform in the frequency domain or its spectrum, i.e. the signal is mapped to its frequency components. The frequency measure is the rate of repetition in time. The number of repetitions within a second is called one Hertz.

1.6.1 EQUATIONS

Let $F(s)$ be the Fourier transform of a continuous function $f(x)$, it can be gained from the following equation:

$$F(s) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi xs} dx$$

For a discrete function over N values the *discrete Fourier transform* equation is:

$$F(u) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)e^{-j\left(\frac{2\pi}{N}\right)xu}$$

The equation for a two dimensional discrete function is:

$$F(u, v) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)e^{-j\left(\frac{2\pi}{N}\right)(xu+vy)}$$

1.6.2 FFTW (FASTEST FOURIER TRANSFORM IN THE WEST)

There is an algorithm which employs the rearrangement of Fourier transform components, in a way that speeds up the calculation time, and is called Fast Fourier Transform (FFT). There are implementations and libraries for FFT in a diverse set of programming languages and applications. FFTW (Fastest Fourier Transform in the West) is a computer subroutine library in C language for computing the discrete Fourier transforms of one or multi dimensional data. The worst case time order of this library is $O(n \log(n))$ (8). This library has been used for the set of experiments in this work.

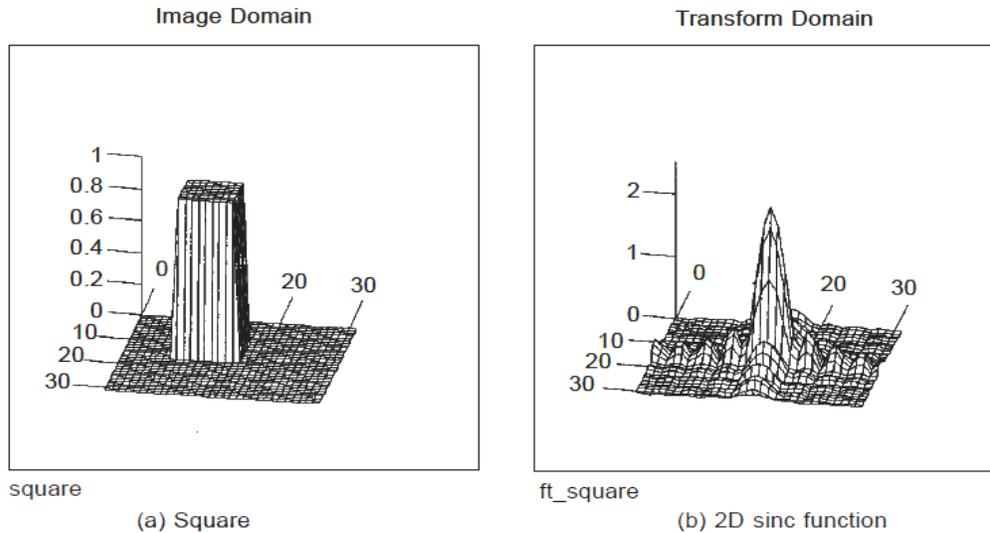


FIGURE 3: AN IMAGE OF A SQUARE PLOTTED AS 3-D SURFACE AND ITS FFT

1.6.3 THEOREMS

There are some basic theorems regarding Fourier transform which some relevant ones are briefly noted in this paragraph. The position of visual objects or features in an image has not any impact on its frequency domain representation. According to *shift theorem* the Fourier transform is *shift invariant*, which means if all the components within an image are shifted by a certain amount the Fourier transform does not change and will remain as it was. The *similarity theorem* denotes that if a signal is compressed in the time scale, it will be expanded in the frequency domain. *Addition theorem* is that the sum of two functions has a Fourier transform equal to sum of their two Fourier transforms (7). “In application, this suggests that we can separate images by looking at their frequency domain components.” (4). The rotation theorem which is of high importance in the domain of this study, is that, when an image is rotated by D degrees in spatial domain, its Fourier transform phase part changes by D degree which means the absolute value does not change.

To be able to map the frequency domain to the spatial domain there are a few hints. The high frequencies are found where the brightness changes rapidly. “Globally, the lower frequencies carry more information whereas locally the higher frequencies contain more information so the corruption of high frequency information is of less importance.” (4)

The following two figures show a battery image and its Fourier transform in three different views. Next to the battery image is the 3Dimensional representation of its Fourier transform in Matlab, and the next figure shows two different partial views of the same transform. These three representations of the Fourier transform of this battery are acquired using FFTW library in a self-developed application by author and Matlab has been used to visualize the data. The two last representations have been used in proposed methods in this work.

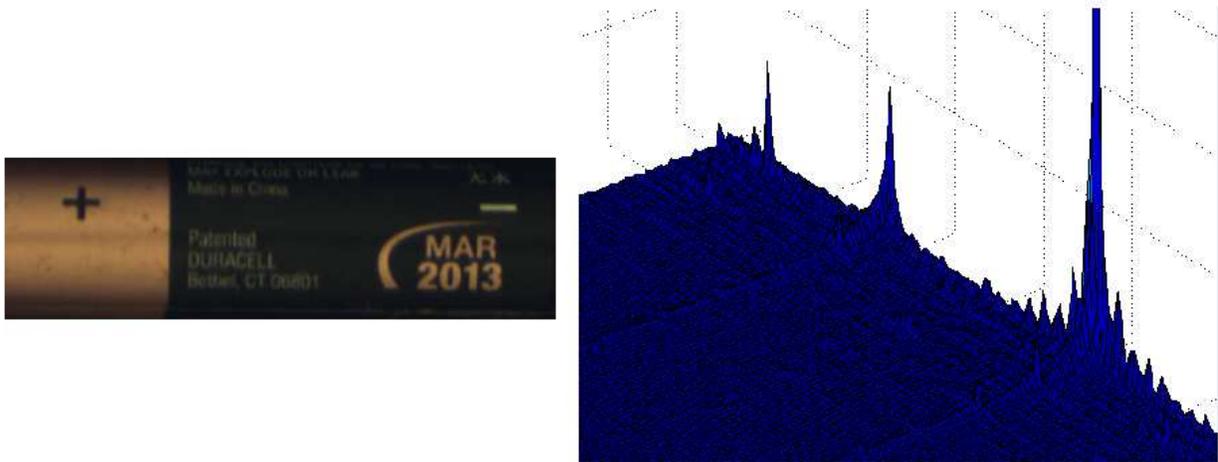


FIGURE 4: A BATTERY IMAGE AND A 3D VISUALIZATION OF ITS GRAYSCALE COUNTERPART FFT

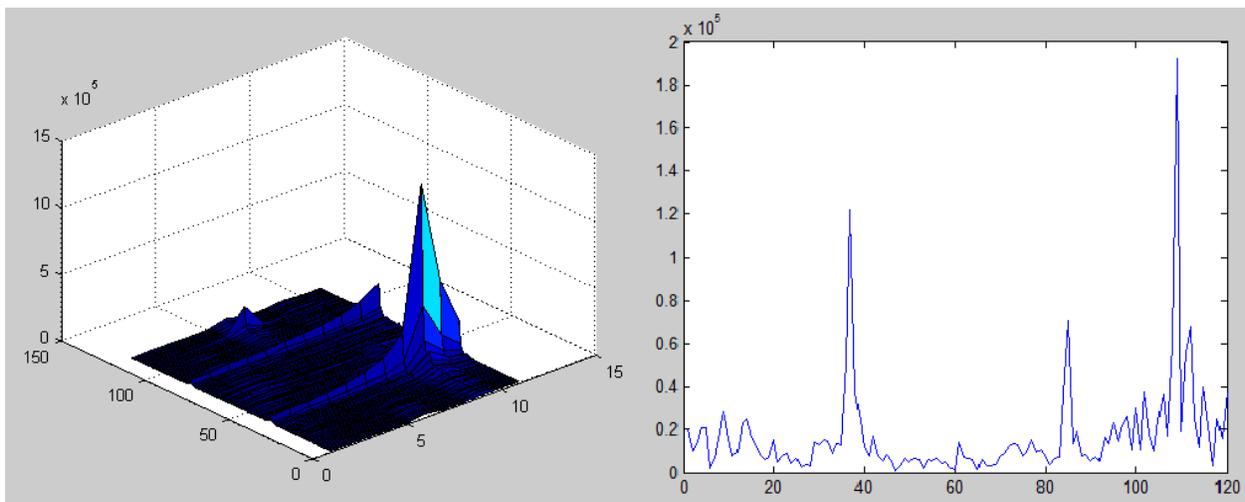


FIGURE 5: A BAND OF FFT REPRESENTED IN PREVIOUS PICTURE AND A SECTION OF THIS BAND

2 EXPERIMENTATION

A set of tests have been performed with focus on two main features, the color and the frequency components. The details about these tests have been elaborated in the next sections.

2.1 DATA SET

The input data was provided by the Optisort Company as a set of more than 7000 images in 34 folders. Each folder contained around 200 images of one battery type; therefore there were 34 different battery types. Each image was in size of 120x408 pixels and encoded in 24-bit bmp format. Pictures were taken from a battery with small alternation of angle around its axis. An input to the system is just one image

which is taken from one arbitrary face of a battery. Some sample images of this dataset are as in bellow figure.

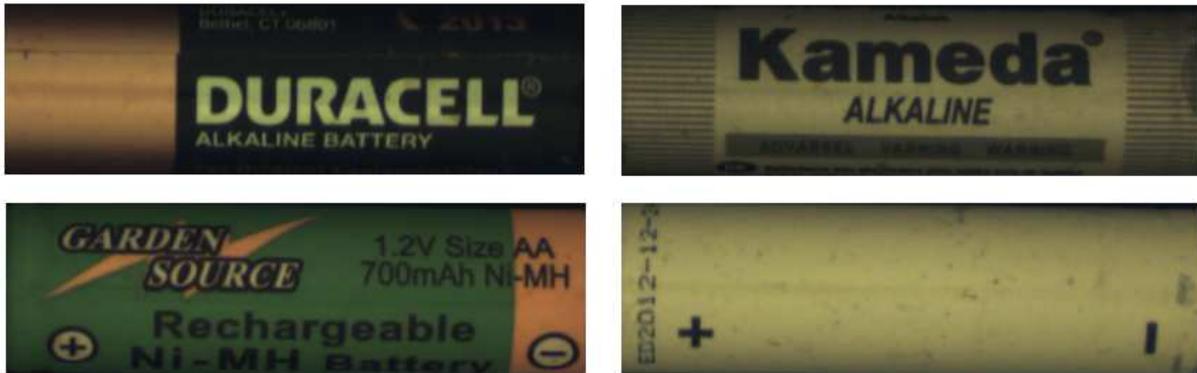


FIGURE 6: SAMPLE BATTERY IMAGES FROM DATASET

The training set is formed with three least overlapping images for each battery type, i.e. there are 34×3 images in the training set. The selection of 3 images rather than any other number of images for training was due to a test on a sample of about 1000 images. Firstly 11 images were selected as a training set for each battery class and the result of a color distance test showed that, having only three images is almost as effective as having 11 images. Below chart shows that the average RGB color distance of each of 250 images against the training set has the same trend for both 11 images in the training set, and 3 images in the training set. Both 11 and 3 image selection for the set is based on a maximum coverage of the battery surface.

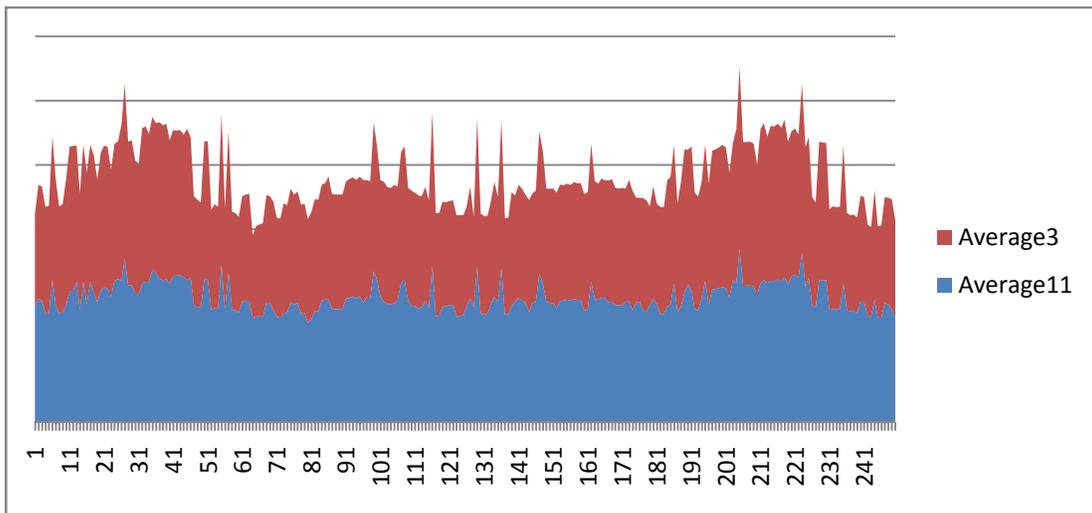
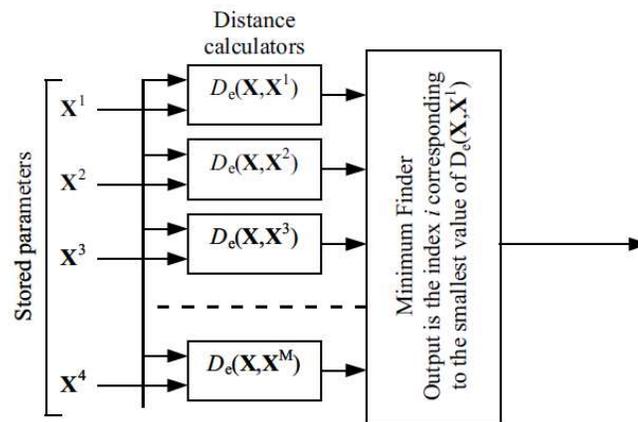


FIGURE 7: SIMILAR TREND OF THE DISTANCE HAVING A TRAINING SET OF 11 (BLUE), AND 3 (RED)

2.2 CLASSIFICATION METHOD

The method chosen for classification was KNN. The intention was to start with KNN and then replace it with SVN or Neural Network in a next phase, but due to its simplicity and the good results which were achieved, it remained the main classification method of this work. In another thesis work on the same issue a wider range of classification methods including the two latter mentioned ones have been applied (1).

KNN (k-nearest neighbors' algorithm) is a classification method which identifies the pattern based on its nearest neighbors, i.e. the closest training data in the feature space. KNN is one of the most straightforward algorithms in machine-learning. $K=1$ means that the object is assigned to its closest instance of training set. To evaluate the input against the training data and find its nearest instance, there must be a feature which allows numerical comparison. A graphical view of KNN is represented in below figure (6).



Input vector:
 $\mathbf{X} = (X_1, X_2, X_3, \dots, X_N)$

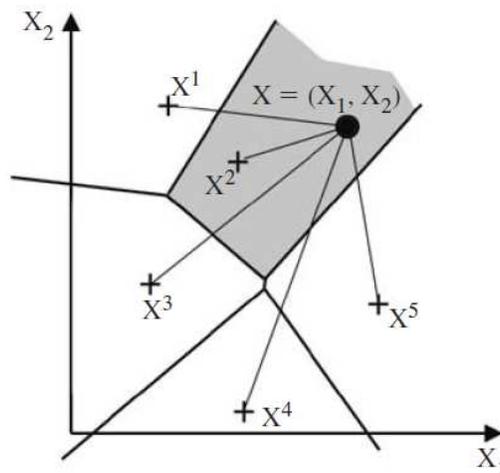


FIGURE 8: NEAREST NEIGHBOR CLASSIFIER: (UP) THE ALGORITHM; (DOWN) THE PATTERN REPRESENTED BY THE VECTOR $X = (X1, X2)$ IS IDENTIFIED WITH THAT CLASS ASSOCIATED WITH THE CLOSEST STORED POINT: $X1, X2, X3$, (I.E., $C2$). ALL POINTS IN THE SHADED AREA ARE CLOSER TO $X2$ THAN OTHER STORED POINTS.

2.2.1 DISTANCE FUNCTION

Next step was to pick a proper distance function. Considering a feature in the training set as a reference vector, V_r and the same feature information for an input image as V_i , there is a diverse range of comparison methods between two vectors to find the similarity or distance of the two.

$$V_r = \{V_{r1}, V_{r2}, \dots, V_{rN}\}$$

$$V_i = \{V_{i1}, V_{i2}, \dots, V_{iN}\}$$

The Euclidean distance which is the ordinary numerical distance of the two points in an N dimensional space is calculated as:

$$D(V_r, V_i) = \sqrt{(V_{r1} - V_{i1})^2 + (V_{r2} - V_{i2})^2 + \dots + (V_{rN} - V_{iN})^2}$$

The *Jaccard similarity coefficient* is another distance function which is defined as the number of intersection members of two sets divided by the number of union members of them. *Tanimo coefficient* which extends the Jaccard coefficient is a measure of similarity between two vectors by finding the angle between them:

$$T(V_r, V_i) = \frac{V_r \cdot V_i}{\|V_r\| \|V_i\|}$$

The product moment correlation is another measure of similarity between two vectors:

$$r_{V_r V_i} = \left(\sum_j \frac{(V_{rj} - \bar{V}_r)}{\sigma_r} \cdot \frac{(V_{ij} - \bar{V}_i)}{\sigma_i} \cdot \frac{1}{N} \right)$$

\bar{V} is the average of vector and Sigma is the *standard deviation*. \bar{V} and Sigma can be achieved with the following equations:

$$\bar{V} = \frac{\sum_{j=1}^N V_j}{N} \quad \text{and} \quad \sigma = \sqrt{\frac{\sum_{j=1}^N (V_j - \bar{V})^2}{N}}$$

All of these three mentioned distance measures have been employed in this work, but a fourth distance function which is an arbitrary measurement turned to be the best one for this experimentation:

$$D = \sum_j |V_{rj} - V_{ij}|$$

This arbitrary distance metric is sensitive to the range of each signal, while Euclidean distance is fitting for finding the distance of two points in an N-Dimensional space, the Jaccard and Correlation coefficients are sensitive to the trend, i.e. if two signals are very far in the range but have the same rise

and fall, they will get a good correlation similarity. Since in this work, the numeric range of color values has a distinctive role, the Euclidean approach gave the best result among all. The next figure tries to demonstrate a typical vector (signal) for this problem and other possible vectors for comparison.

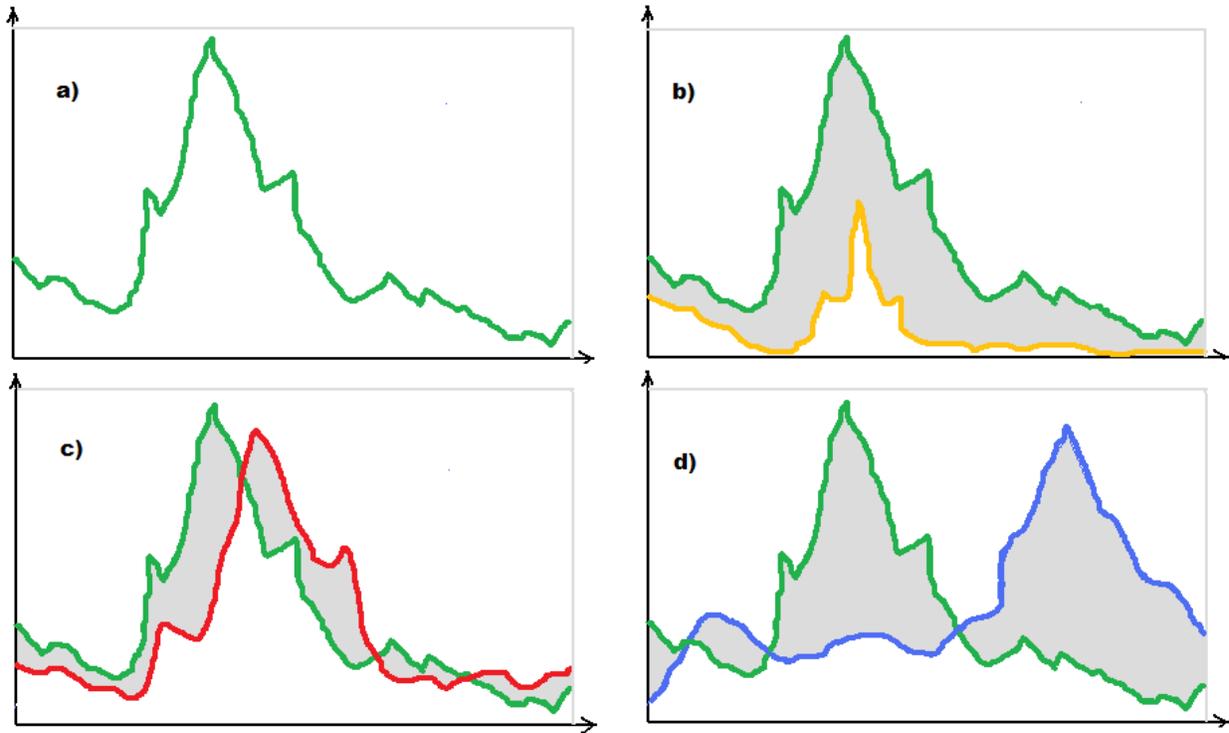


FIGURE 9: A) A TYPICAL SIGNAL FOR THE CURRENT PROBLEM, B) A SIMILAR SIGNAL BUT FROM A DIFFERENT TYPE, C) A SIMILAR SIGNAL FROM THE SAME TYPE, D) A DIFFERENT SIGNAL FROM A DIFFERENT TYPE

The tested features are mostly RGB dependent numerical values and Fourier transformations. The first part of experiments has been based on RGB values. “The distance between two images I1 and I2 is defined as the distance between their respective histograms:

$$|I1 - I2| \equiv |H1 - H2|$$

The smaller the distance is, the closer the similarity of the images.” (6)

Later on, the experiments were re-executed with different distance functions.

2.3 ANALYZED FEATURES AND RESULTS

In these set of experimentation, the focus was on the usefulness of Fourier Transforms as a feature in image classification. Yet different combinations of color features, with or without Fourier Transform have been tested, and finally the Fourier Transform in a grayscale colors-space has been examined.

2.3.1 RGB COLOR HISTOGRAM

The first set of experiments was based on RGB color space and the Euclidean distance of the input to each of representatives of the base battery types. Each pixel RGB values were acquired to form a histogram to shape a histogram vector and the vector was the subject of comparison. The steps are:

- 1- The input image is read and the color matrix is extracted.
- 2- Each battery type is represented with three images which have the minimum overlap.
- 3- The histogram of color matrix of input image is compared to the three histograms of each battery class based on a simple distance function.
- 4- The nearest neighbor is recognized as the battery type of the input image.

The results of running this test are as the following table. The number of false and right verifications, total number of tested images and the performance which is the number of right verifications divided by the total number of tested images are presented. The same structure will be held till the end of this report:

RGB Color Histogram	FALSE	TRUE	Grand Total	Performance
Count of Result	1216	5584	6800	82%

To get deeper in the results, a chart is provided which shows each of 34 battery classes and the rate of right verifications per battery class. The average performance is brought in the end of this chart:

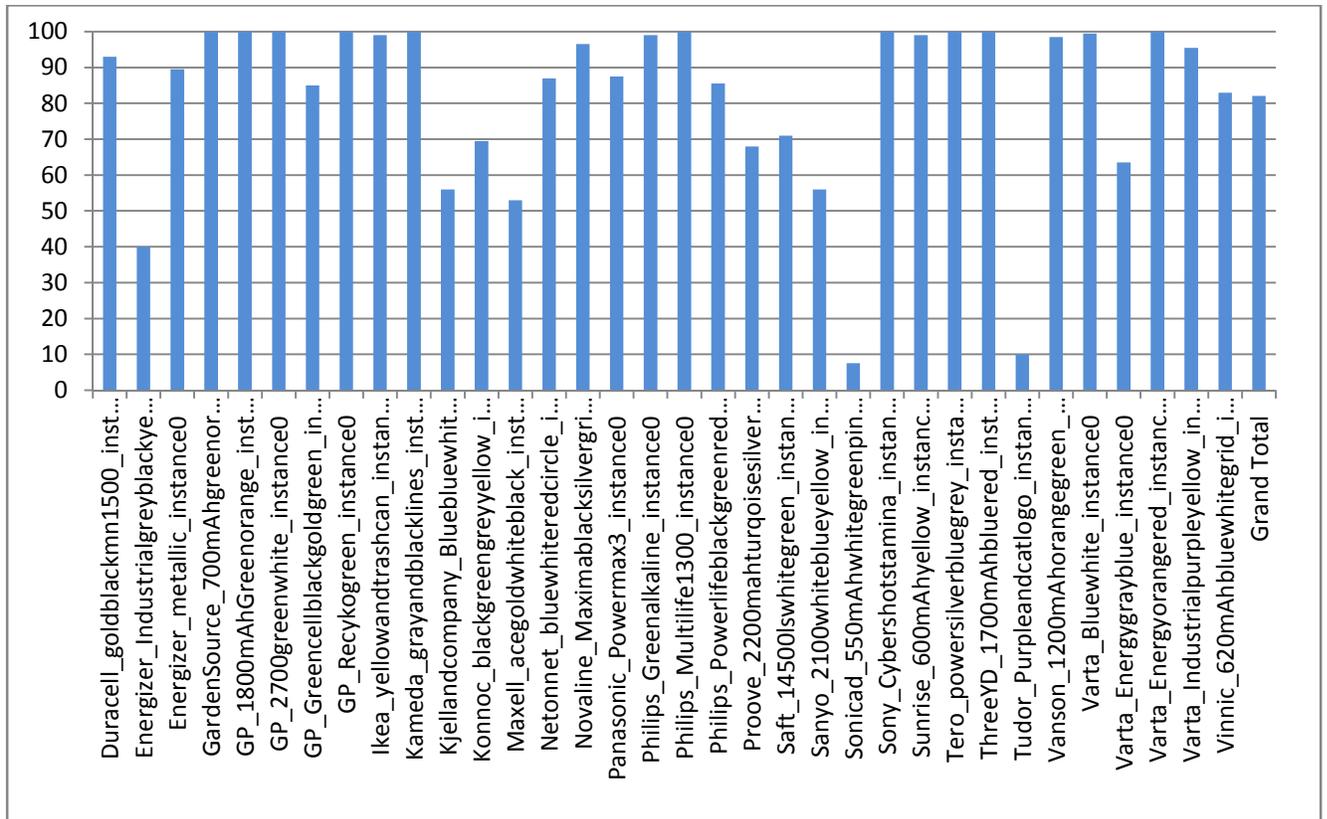


FIGURE 10: PERFORMANCE RATE FOR FEATURE (RGB HISTOGRAM)

2.3.2 FOURIER TRANSFORM SECTION

To analyze the effectiveness of Fourier transform as a distinctive feature, sample images were selected in the dataset and their Fourier transform was carefully observed. As stated in previous sections, and observed in the transforms, most important and distinguishing data in a Fourier transform of an image is at the low-frequency components. So a set of first 30 elements of lowest frequency elements were picked to form a vector which was used in this part of the experiment. Since the Fourier transform results in complex numbers, the ABS of each number was picked as an equivalent measure. The ABS of a complex number is defined as:

$$C = (C_r, C_i), \quad ABS(C) = |\sqrt{C_r^2 + C_i^2}|$$

Due to the rotation theorem which is briefly implied in section 0, this ABS value is not sensitive to the rotation of an image, so the batteries can appear in any direction as the input image and no extra load in training set is required.

The results turned to be:

FFTW Section	FALSE	TRUE	Grand Total	Performance
--------------	-------	------	-------------	-------------

Count of Result	5078	1722	6800	25%
------------------------	-------------	-------------	-------------	------------

The overview of results for different battery classes:

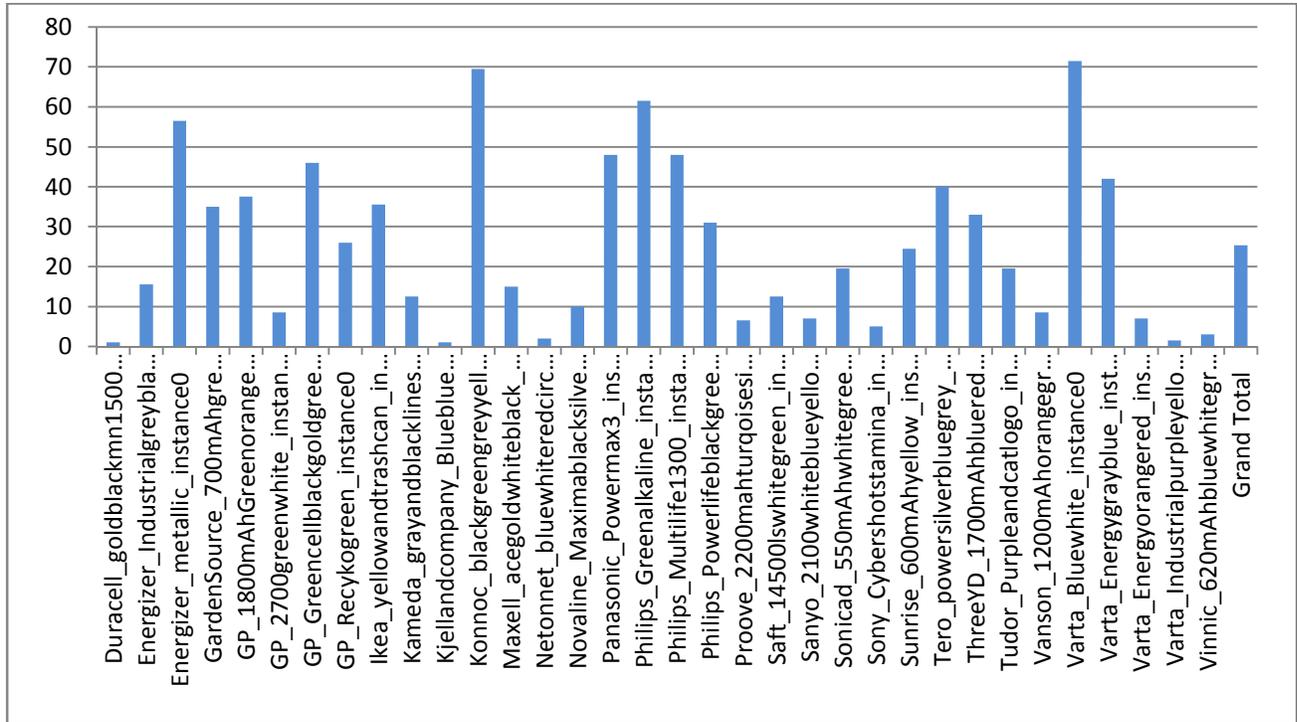


FIGURE 11: PERFORMANCE RATE FOR FEATURE (FFT SECTION)

2.3.3 RGB ILLUMINATION INVARIANT AND FOURIER TRANSFORM

This part of the experiment was designed based on equations presented in 1.5.1 and rely on an illumination-invariant color model. In this test, the color values of each pixel are replaced with the new color value and a histogram has been built based on the new values. Then top 5 nearest battery classes are picked and with a Fourier transform section vector the closest battery has been picked. The results are:

RGB Illumination Invariant FFT	FALSE	TRUE	Grand Total	Throughput
Count of Result	462	6338	6800	93%

The overview of results for different battery classes:

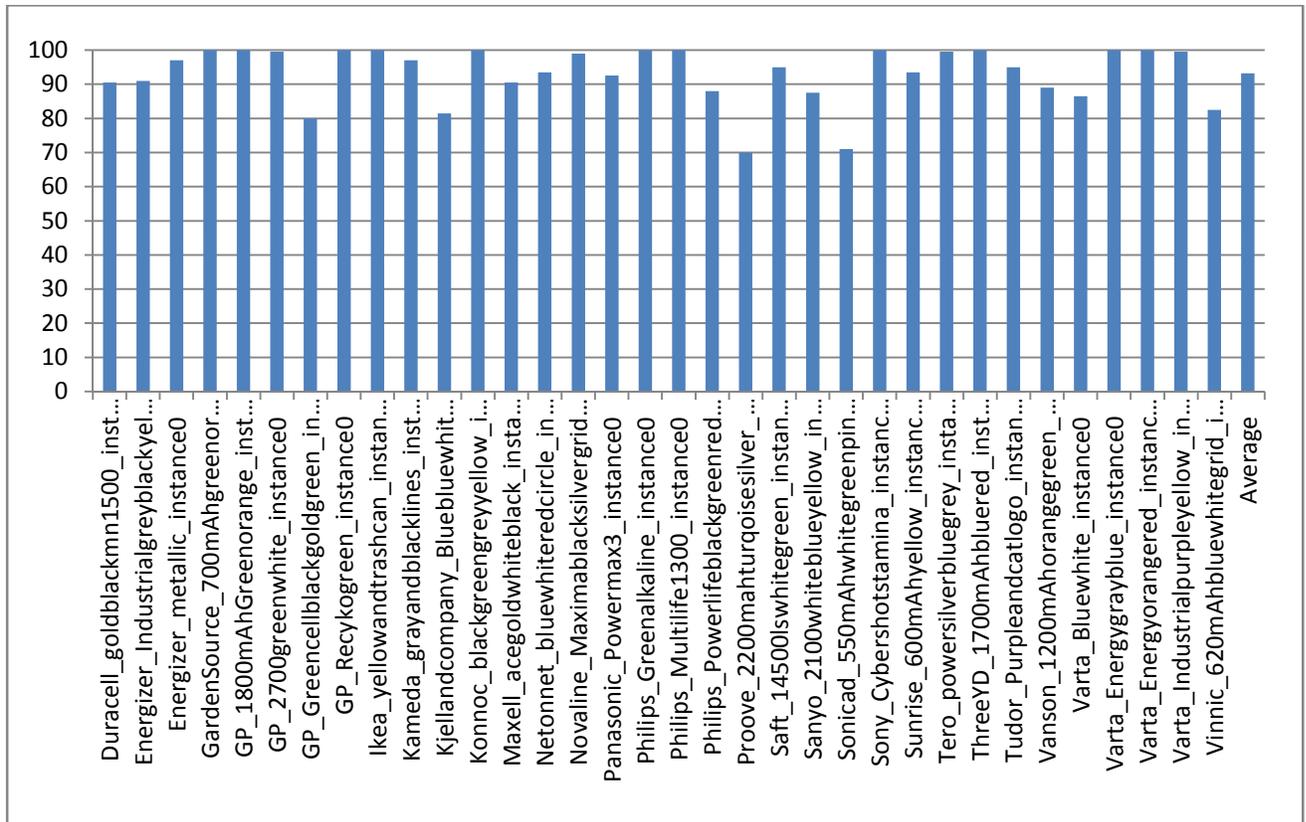


FIGURE 12: PERFORMANCE RATE FOR FEATURE (RGB ILLUMINATION INVARIANT AND FFT SECTION)

2.3.4 THREE PIECES OF HISTOGRAM

Another idea was to divide the battery image to three parts and extract three illumination invariant histograms and form a feature vector by simply attaching these three vectors. Many types of batteries have a label design as below:

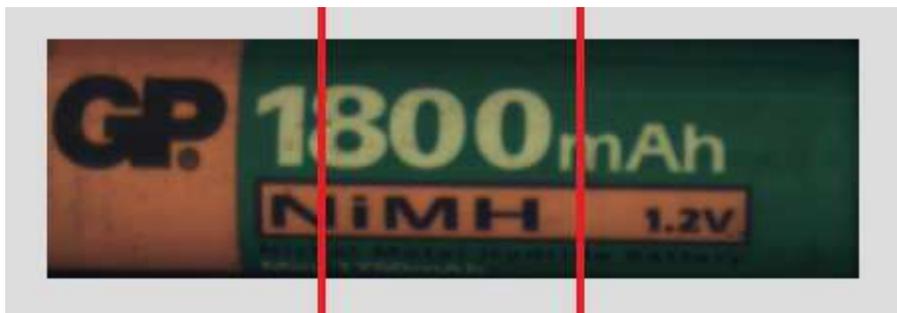


FIGURE 13: A BATTERY IMAGE IN THREE SECTIONS

By dividing the battery surface more information regarding its color structure could be achieved. The test was run and this result was achieved which is 3% better than the RGB feature test:

3 Pieces Color Invariant Histogram	FALSE	TRUE	Grand Total	Performance
Count of Result	996	5804	6800	85%

The overview of results for different battery classes:

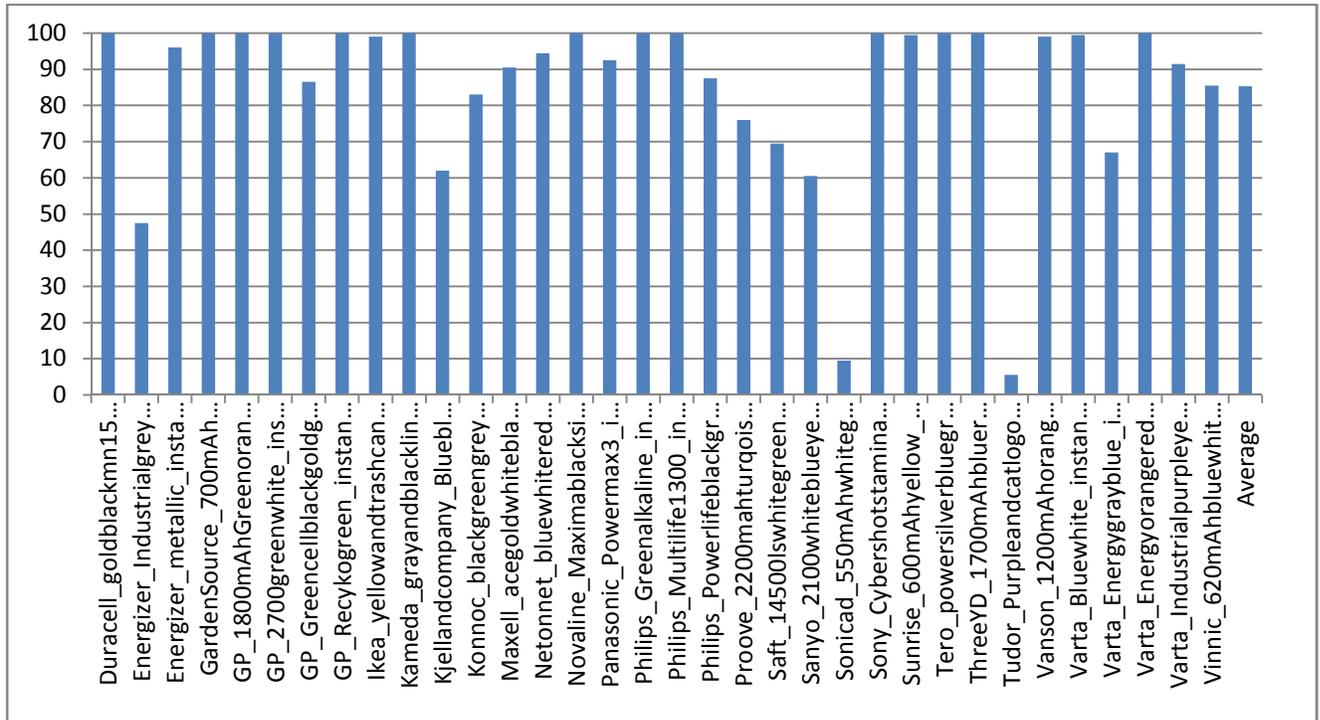


FIGURE 14: PERFORMANCE RATE FOR FEATURE (THREE PIECES COLOR INVARIANT HISTOGRAM)

With Fourier slice on first 5 results:

3 Pieces Histogram & FFT Slice	TRUE	FALSE	Grand Total	Throughput
Count of Result	6398	402	6800	94%

The overview of results for different battery classes:

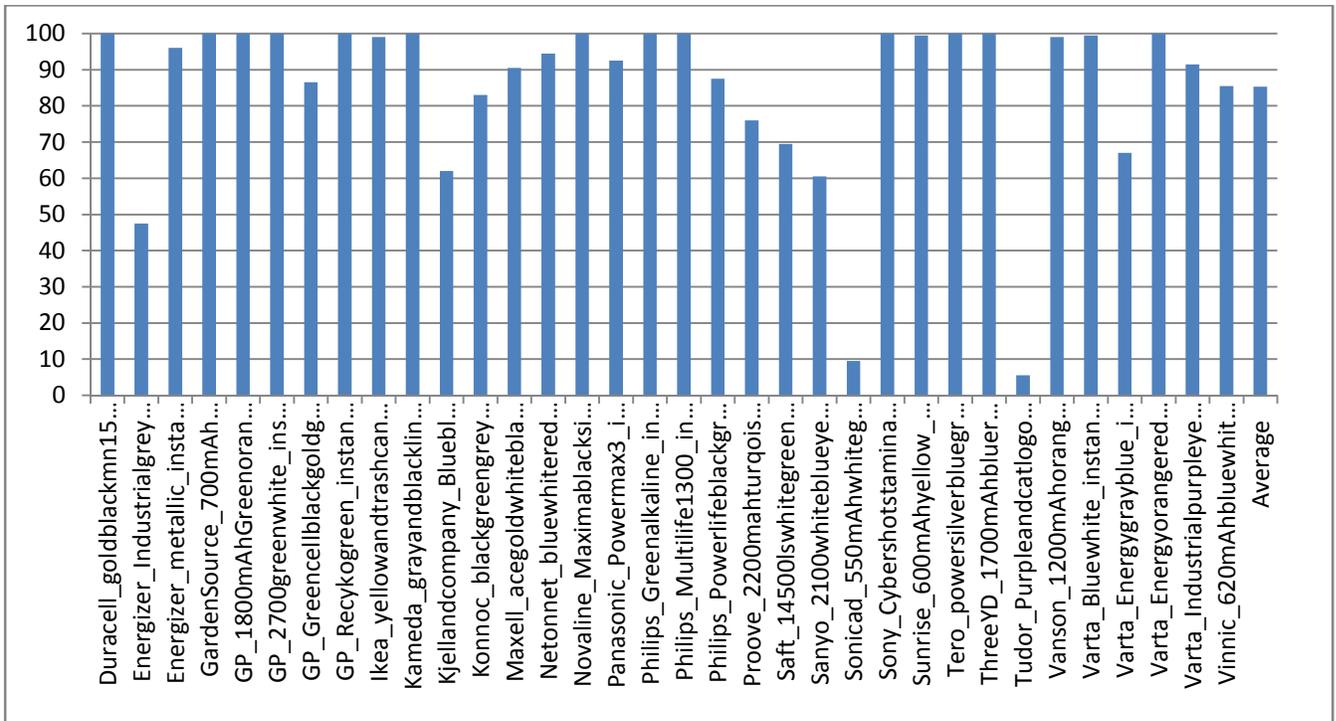


FIGURE 15: PERFORMANCE RATE FOR FEATURE (THREE PIECES COLOR INVARIANT HISTOGRAM AND FFT SLICE)

2.3.5 THREE PIECES OF HISTOGRAM AND REDUCED COLOR PERCISION

The number of illumination invariant r and g has been reduced to half, i.e. 50 color values were picked to represent each pixel red value and 50 to represent green value (instead of 100 values for each color). The result was:

Three Pieces of Hist. Reduced Colors	TRUE	FALSE	Grand Total	Throughput
Count of Result	5823	977	6800	86%

The overview of results for different battery classes:

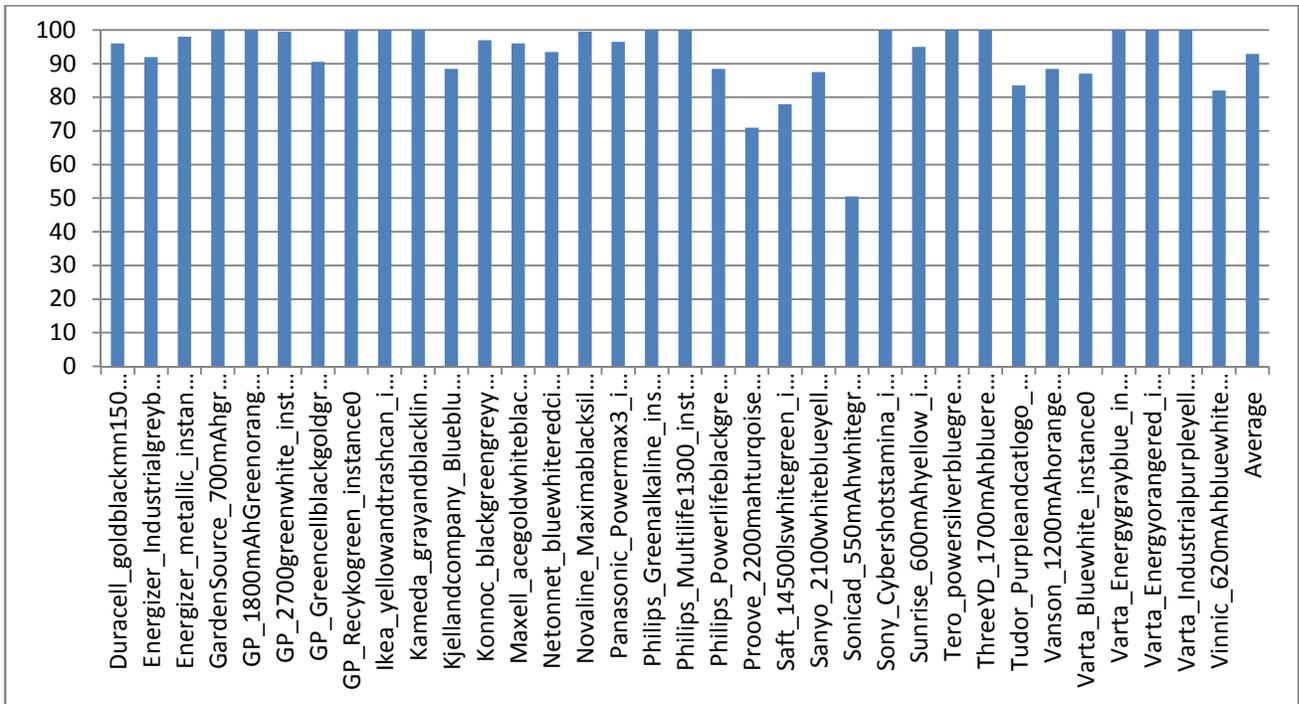


FIGURE 16: PERFORMANCE RATE FOR FEATURE (THREE PIECES COLOR INVARIANT HIST. REDUCED COLORS)

Adding a comparison of the Fourier transform slice to top 5 results in the previous experiment, the results changed as bellow:

3 pieces Color Hist. FFT Section	TRUE	FALSE	Grand Total	Throughput
Count of Result	6317	483	6800	93%

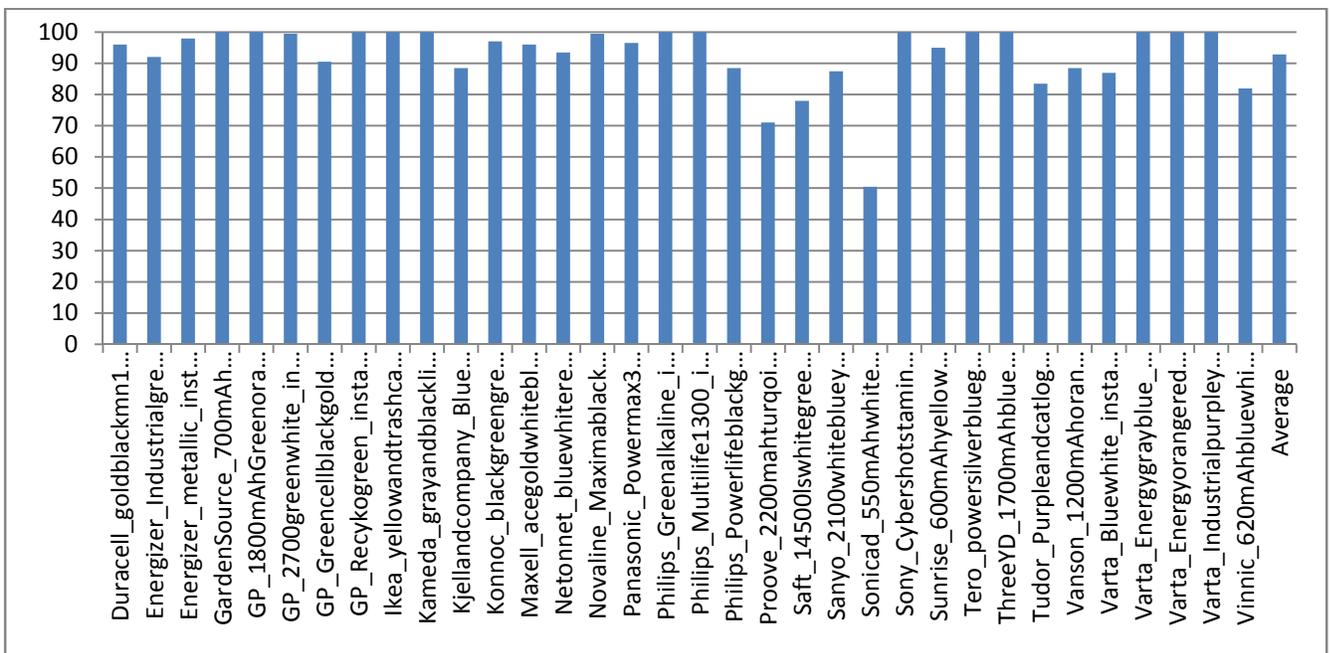


FIGURE 17: PERFORMANCE RATE FOR FEATURE (THREE PIECES COLOR INVARIANT HIST. REDUCED COLORS WITH FFT SLICE)

2.3.6 COLOR INVARIANT FEATURE AND FOURIER TRANSFORM BAND

In the previous experiment, instead of picking a section of a 2D FFTW result, a band which seems to have the most identifying data, and will be mentioned in the rest of this text as *the effective band*, was selected and the data of this band was represented as a 1D vector. So this vector was used to classify the image with the nearest instance in the top 5 results of the last color feature which was applied in previous step. The results were improved as bellow:

Color Invariant and Fourier Transform Band	TRUE	FALSE	Grand Total	Performance
Count of Result	6640	160	6800	98%

The overview of results for different battery classes:

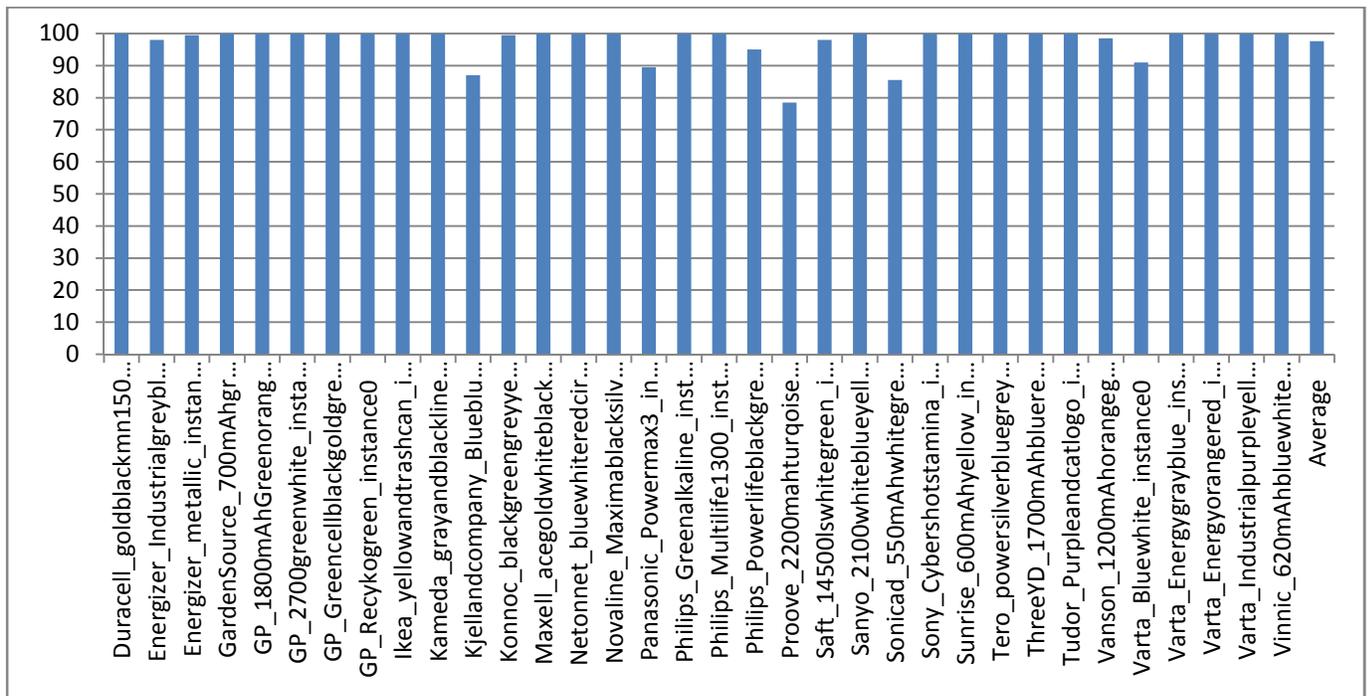


FIGURE 18: PERFORMANCE RATE FOR FEATURE (COLOR FEATURE AND FOURIER TRANSFORM BAND)

2.3.7 FOURIER TRANSFORM BAND

In this part of the experiment, the idea was to focus only on Fourier transform to measure the performance of the *effective band* as a single means of classification and not combined with any color feature. Since the Fourier transform here was extracted based on a grayscale image, the color information is almost ignored.

FFT Band	FALSE	TRUE	Grand Total	Throughput
Count of Result	327	6473	6800	95%

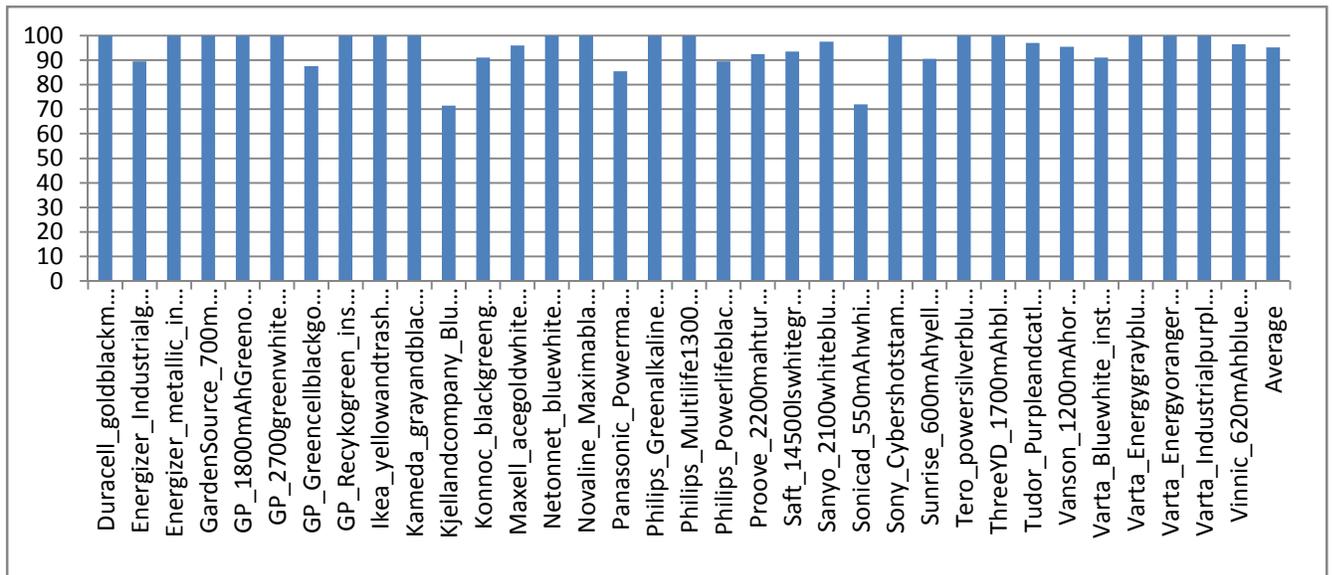


FIGURE 19: PERFORMANCE RATE FOR FEATURE (FOURIER TRANSFORM BAND)

All the test up to this point were based on a training set of 3 images per battery class, but in the next step a new design for the training set is proposed and its effect is shown on improvement on results.

2.3.8 UNBALANCED TRAINING SET

From the results obtained in last section, it was noticed that the performance of the method for 18 battery types is 100% but for others it varies between 70 to 98%. From the log files produced for each battery class, it was obvious that the wrong results are not randomly scattering in between of the right ones, but they are a trail of consecutive images. Based on this evidence it can be concluded that the training set for some batteries does not cover the whole range of battery surface, while for others it is good enough with just three images in the training set. Therefore the structure of test application was changed to support a different number of training images for each battery class. Sample improvements are provided as bellow:

For battery class Kjellandcompany_Bluewhiteyellow, in the last experiment with 3 images as the training data, the performance was:

Battery Class against 3 images	True results	Grand Total	Performance
Kjellandcompany_Bluebluewhiteyellow_instance0	143	200	71.5 %

Increasing the number of training images improved the results, and with 5 training images, the result was:

Battery Class against 5 images	True results	Grand Total	Performance
Kjellandcompany_Bluebluwhiteyellow_instance0	194	200	97 %

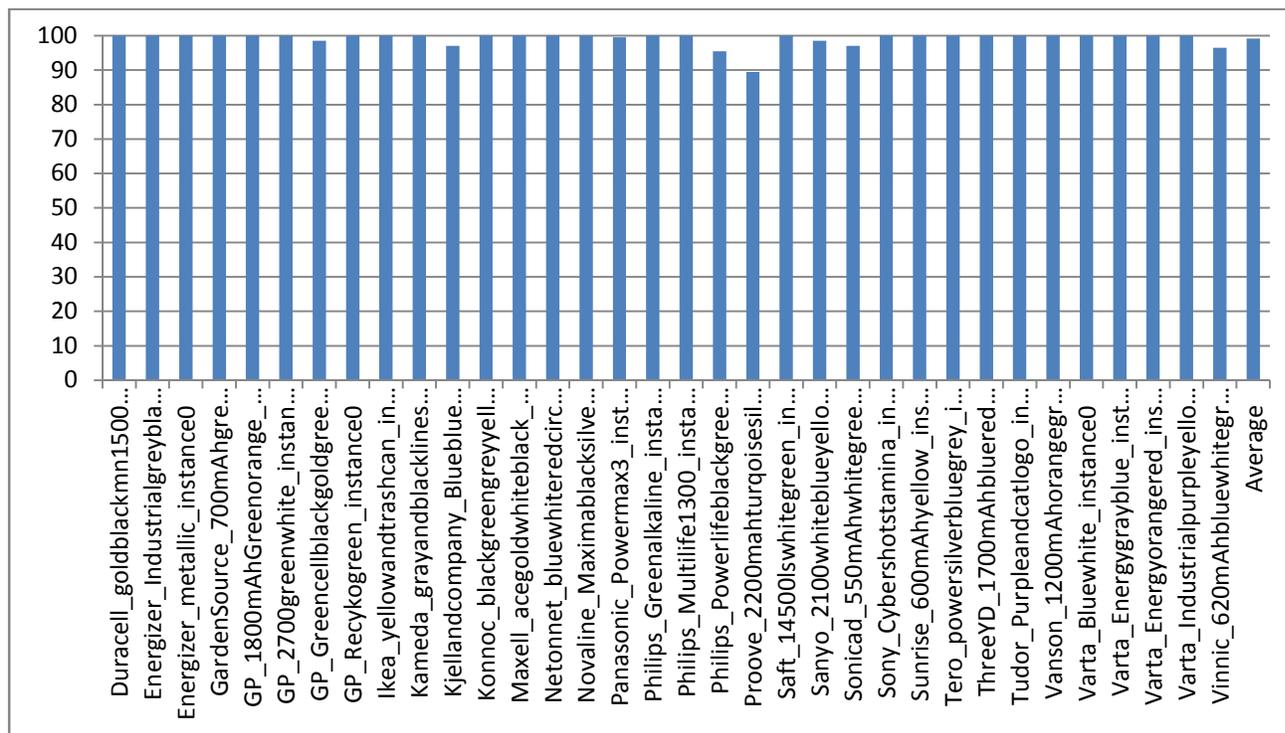
It is needless to say that the selection of training images is a very sensitive step in order to achieve the best performance with the minimum number of training images.

The whole test was repeated with an arbitrary number of training images for each battery type. The number of training images per battery class was:

Number of training images	Number of battery classes
3	19
4	13
5	1
6	1

The result of this test was:

Unbalanced Training & FFT Band	FALSE	TRUE	Grand Total	Performance
Count of Results	56	6744	6800	99%



The most number of training images are used for the type Sonicad_550mAhwhitegreenpink. Having a look at the best tried combination of training images for this battery reveals the reason of such a difference with other battery classes:



FIGURE 20: DIFFERENT FACES OF THE SAME BATTERY, SHOWING HOW UNEVEN A LABEL CAN BE

It can be justified that because of the design of this battery label, more images are required to cover a whole set of images.

2.4 TIMING

Feasibility of an approach is not enough for an industrial usage of it; it must also generate the response within a reasonable amount of time. There are many factors affecting the throughput of a software application, e.g. the hardware configuration, operating system and how tuned they are for a certain usage. A configuration which performs well in an RDBMS application might not be qualified for a computational or graphical application.

These set of experiments were performed on a personal laptop with the configuration of:

CPU: Intel® Core™2 Duo CPU T5750 2.00 GHz

RAM: 4 GB

System Type: 32 bit OS

OS: Microsoft Windows Vista

The programming language was Microsoft C#.NET, though some preliminary tests were performed in Matlab. The author does not claim to be a professional programmer and is not good in memory management in C#.NET; therefore any result achieved may be a subject of improvement. The GUI of developed application is presented in (Figure 21: Application which was developed for this study).

The best timing achieved in this work was with the last presented method, effective FFT band, which has the score of 225 ms per input image. For the combinational method of applying FFT on the color feature, it was 927 ms per input image and for the standalone color features, the response time was around 700 ms per input.

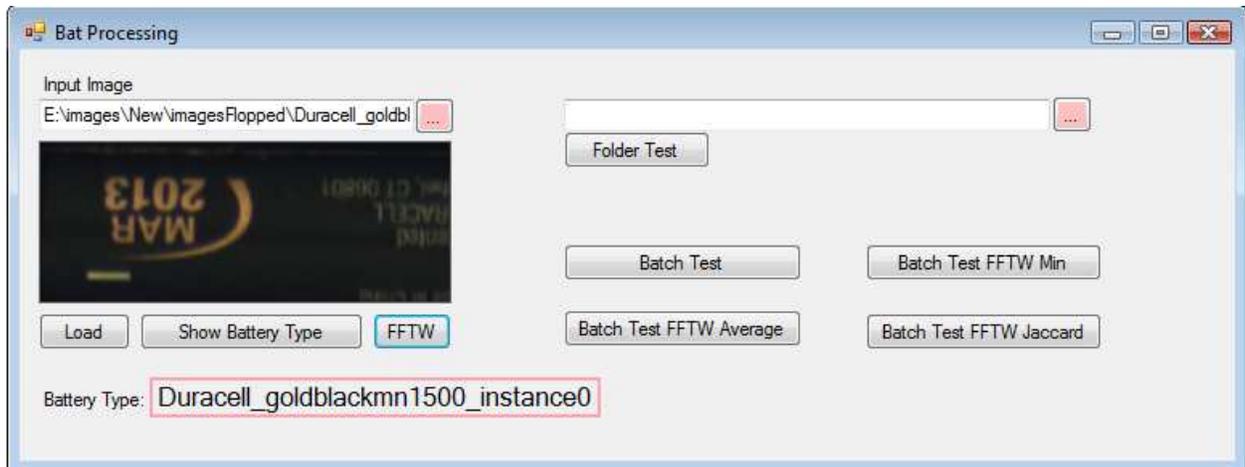


FIGURE 21: APPLICATION WHICH WAS DEVELOPED FOR THIS STUDY

3 DISCUSSION

Any of proposed and tested methods in this work has two outputs per an input image, one is the identified battery class and the other is the numeric distance of the input to this class. The input image is always recognized as one of the 34 batteries in the training dataset. Knowing that there are more battery classes in the world, makes this feature to be undesirable, however this work can be extended to verify such a case.

Another validation problem for the results of this study can be the limited image dataset. All the images which belong to the same class seem to be taken from a very few batteries, though the proposed method is not vulnerable to rotation, and is not so sensitive to illumination. The input dataset lacks the corrupted and noisy images as well.

4 CONCLUSION

The objective of this study was to examine and introduce useful features which are able to successfully classify an input battery image against a training dataset. A range of methods were examined and the results were analyzed. The output of each phase was used to improve the features for a next round of experimentation. Different results were achieved for the proposed features.

The best result achieved during this study was 99% and was performed by the method, named here as effective FFT band. These tests showed that FFT is a good-quality and reliable feature in this dataset and

for some battery classes, it is 100% accurate. To increase the performance over a battery class, having an increased set of training images proved to be fitting. Because of the rotation-invariant feature of FFT on ABS value over an image, it is well suited for detecting flipped or rotated battery images.

5 FUTURE WORK

The results achieved in this work are tempting enough to become a base for further research works in the field. The very next development over this method could be to try to verify if the result is accurate enough or if the battery must be sent to unrecognized bag.

These experiments are performed on a dataset which no special preprocessing was performed on them. Finding a few preprocessing methods which well suits as a combination of Fourier transform band approach which is introduced in this work could be a useful study.

Making a semi-standard datasets to be used for such a work is another time consuming job which helps to have a better validation over any proposed method. Datasets can be arranged having similar and dissimilar battery types to be more accurate.

Reference

1. **Pascale, Giuseppe and Savarino, Stefano.** *Optical Battery Sorting with Computational Intelligence Techniques.* Chalmers University of technology. Gothenburg : s.n., 2008.
2. **Appelkvist, Olof, et al.** *Battery Sorting Prototype.* Chalmers University of Technology. Gothenburg : s.n., 2009.
3. *Recycling of batteries: a review of current processes and technologies.* **Bernardes, A. M., Espinosa, D. C. R. and Tenório, J. A. S.** 1-2, s.l. : Elsevier B.V., 2004, Journal of Power Sources, Vol. 130.
4. **Nixon, Mark S. and Aguado, Alberto S.** *Feature Extraction and Image Processing.* s.l. : Newnes, 2002. ISBN 0 7506 5078 8.
5. **Bow, Sing-Tze.** *Pattern Recognition and Image Preprocessing.* 2. New York : Marcel Dekker, 2002. ISBN: 0-8247-0659-5.
6. *Image indexing using compressed colour.* **Berens, J., Finlayson, G.D. and Qiu, G.** 4, 2000, Vision, Image and Signal Processing, IEE Proceedings, Vol. 147.
7. **Bracewell, Ronald N.** *The Fourier Transform and Its Applications.* s.l. : McGraw-Hill Higher Education, 2000. ISBN 0-07-303938-1.
8. **Frigo, Matteo; Johnson, Steven G.** . FFTW. [Online] MIT. <http://www.fftw.org/>.
9. **Stanković, Radomir S., Moraga, Claudio and Astola, Jaakko.** *Fourier Analysis on Finite Groups with Applications in Signal Processing and System Design.* s.l. : Wiley-IEEE Press, 2005. ISBN-I 3 978-0-471 - 69463-2.