Master’s Dissertation 2011

Visual Data Mining of Rechargeable Lithium Ion Battery Deterioration Characteristics

Susumu Mizuguchi
（Student ID Number：81033514）

Supervisor Tetsuro Ogi

March 2012

Graduate School of System Design and Management,
Keio University
Major in System Design and Management

SUMMARY OF MASTER'S DISSERTATION

Student Identification Number	81033514	Name	Susumu Mizuguchi

Title:

Visual Data Mining of Rechargeable Lithium Ion Battery Deterioration Characteristics

Abstract:

With the wide spread of Electric Vehicles and its potential for reusability, it is required to understand how the deterioration of battery can be analyzed through a set of running data. In this research, a visual data mining application is designed and developed in order to conduct interactive and explosive knowledge discovery. The application enables three-dimensional data representation on a set of super high-definition 4K Displays.

Although its data volume is relative large, data, taken from Electric Vehicles over the period of 2 months, can be and should be processed and visualized with faster computation and more detailed data representation. The visual data mining program is thus modified in order to adjust such requirement.

As a result of visual exploration by the visual data mining application created, it was able to extract the data points which may be used to evaluate the deterioration from run data.

Key Word(5 words)

Visual Exploration, Visual Data Mining, High Definition Three-Dimensional Visualization, Lithium Ion Battery, Deterioration Analysis,

Table of Contents

1. Research Background

1.1. An Overview of Rechargeable Lithium Ion Battery

Lithium Ion Battery (Li+B) is a type of battery that employs lithium anode. It is one of the most popular batteries for consumer electronics products. The battery has drawn attention because it is and will be employed to electric vehicles.

In 2008, the U.S. government announced their new policy described as "A Green New Deal" which provoked the development of renewable energy, along with other financial policies. The new policy stated that it is "necessary to create new industry and or market that can expect high growth". The industry or markets indicated here were specifically described as for those that can contribute to sustainable development of both our environment and human being. The industries are not only the ones who were affected by the new policy. To an extent to the policy, Japanese government promised rather an aggressive target if 25 percent reduction of greenhouse gas emission by 2020. In addition to the policy, current situation of global economy raises a concern of energy issues. With emerging markets growing every year, the developing countries are consuming energies which have been in stake for decades. The ways to realize alternate natural energy resources or other renewable energy have been actively discussed, yet these solutions are dependent to weather and other natural conditions, which are uncontrollable. Therefore, it is important to develop the rechargeable battery so that it can keep, distribute, and reuse the energy efficiently. Such an infrastructure of energy utilization is referred as smart grid. Rechargeable lithium ion battery will be the one of most important components in smart grid technology.

Rechargeable battery is, thus, a more important component of which it enables smart grid technology for leveling the fluctuation of electricity demand and supply. Smart grid is a concept of which to connect the energy sources and rechargeable entities under specific protocols defined by voltage, frequency, phase, and harmonic. Smart grid is expected to be realized by Home Energy Management Systems (HEMS). In HEMS network, the electricity is communicated via various network protocols such are ZigBee, Bluetooth, and Z-Wave*, and transcend the information to Smart Meter so that the smart meter can manage the each home equipments in more efficient way.

For electric vehicles, it is also dependent to rechargeable batteries. Majority of electric vehicle today employs rechargeable lithium ion battery(Li+B) for its specific characteristics suitable for use cases of which are superiority in energy density, voltage, charge and discharge energy efficiency, self-discharge rate, and less memory effect. Since lithium-ion batteries comprise a family of battery chemistries that employ various combinations of anode and cathode materials, each combinations has distinct advantage and disadvantage in terms of safety, performance, cost, and other parameters. The most prominent technologies for automotive applications are lithium-nickel-cobalt-aluminum (NCA), lithium-nickel manganese-cobalt (NMC), lithium-manganese spinel

(LMO), lithium titanate (LTO), and lithium-iron phosphate (LFP). The technology that is currently most prevalent in consumer electronics is lithium-cobalt oxide (LCO) and it is considered as unsuitable for electric vehicle due to its safety risks. Each of these automotive batteries would require elaborative monitoring system in which to better balance and control the chemical release of energy, preventing thermal runaway and ensuring reasonably long battery life. Due to availability of data, it is deemed that Li+B signify lithium manganese spinel (LMO).

While the green new deal policy provoked the discussions on environmental concerns such as global warming and the paradigm shift on energy utilization, it is also important to note that the policy indeed was triggered by financial crisis occurred in 2008. The financial crisis was followed by major U.S. automobile players' –referred as Big Three- failures. Obviously the policy not only intended to work on environmental issues, but also to stimulate the economy by investing more than 150 billion dollar over next decade and creating the employment of 5 million people. For Japan, especially in manufacturing industry, of which both automobile and battery industries reside, the employment has already been a major concern with the emergence of new workers from Eastern Asian countries. The work has been transferred to those referred as "Low Cost Region (LCR)". LCR has become a thread to domestic employment. Such a work-migration is apparent in assembly process, where it deemed as less-valuable compared to other processes such are components, intermediate systems, services, and consulting. For instance, in 2010 Nissan Corporation migrated the manufacturing of their MARCH products to developing countries such are Thailand, India, China and Mexico mainly because of cost efficiency. Although Li+B, a material component and an OEM (Original Equipment Manufacturer) product, may follow the same path as MARCH, its majority is currently manufactured in Japan. However, in the long run, as the cost of Li+B has been a hurdle for industry players, some of its manufacturing process will be outsourced to LCR countries in future. One must cultivate in order to survive. Yet, historically speaking, Japan is not really good at this. Companies in Japan are more often praised for its detailed and quality work instead of dynamic market penetration. As Sharp ironically named their new media tablet as "Galapagos", some of Japan-unique technologies have developed in a rather closed way. Li+B is one of few such technologies that would meet the condition to fully utilize the beneficial aspects of Japan's high-standard quality control and uniquely grown technology.

1.2. Visual Data Mining

From data stand point, today, it is possible to buy a hard disk drive that can store all of the world's music at $600. Data is flooding, and it will continue to do so for next few years. IDC Digital Universe Study revealed that, in 2020 there will be 40 more times amount of data used. In 2009, there is about 0.8 zeta bytes used and in 2020, 35 zeta bytes of data are expected to be used. This tremendous change is often referred as "Big Data". Yet when "Big Data" is discussed, there is no defined set of data specifications by measures such as tera bytes and so forth. It is because the definition of such data is different by industry. In general, “Big Data” suggests a set of data with its size beyond the current computational feasibilities or analytical capabilities.

Data mining has been a prominent activity with the emergence of relational database and its development. It is also called “Knowledge Discovery in Database”, and is intended for heuristic knowledge acquisition from the database. However, with growing demand of data volume and complexity, each process of data mining (data preparation, algorithmic transactions, visualization, and perception) must be empowered by innovations.

Visual Data Mining (VDM) is an elaborative data analysis methodology derived from data mining and visualization methodology. While the objective of data mining is to process the large amount of data with machine learning, its process may get more cyclical and repetitive with conventional data visualization methodologies as the number of data sets grow and vary.

Data mining, as it has already been discussed, is aimed for “knowledge discovery”. However, its process can easily become exhaustive due to the amount of iterative processes in visualization. Although there are many data mining applications available today, the applications do not provide rich and interactive visualization interfaces that would lead new finding. Effective data visualization would also contribute to effective communications, thus better understandings of information.

With the objective of data mining is set to knowledge discovery, the role of data reasoning by visualization has become more significant. The advancement of computational technology, such as CPU (Central Processing Unit) and GPU (Graphics Processing Unit) has enabled richer visualization for both consumer and enterprise scale. Furthermore, the faculty possess 4K display environment referred as “CDF (Concurrent Design Facility)”. To utilize such a rich visualization environment as well as generic computing environment is one of motivations toward this research project.

1.3. Project Background

As per the background of research project, I took a part in a research project with Professor Sasaki from Keio University.

The main project is conducted by JARI (Japan Automobile Research Institute). It is a subproject of a national project entitled “Development of Performance Assessment Methodology for In-Car Batteries” under “Next Generation Energy/ Social System Demonstration – The Development of Compound Storage Systemization Technology”. In below, I reprint the excerpt from the project summary.

“ It is certain that Li+B will be employed to rechargeable battery system as a tool to realize the low carbon society. Li+B, on the other hand, has also been employed to in-car battery for Electric Vehicle (EV) and Plug-in Hybrid Electric Vehicle (PHEV), and the countless number of used Li+B will begin to emerge in future. Used batteries may be re-used to other means. Re-use of battery would make it feasible to share the battery costs and accelerate EV/PHEV growth, as well as to promote the system enabled by smart grid technology. Thus it is likely to contribute the early realization of “Next-Generation Energy Social System”.

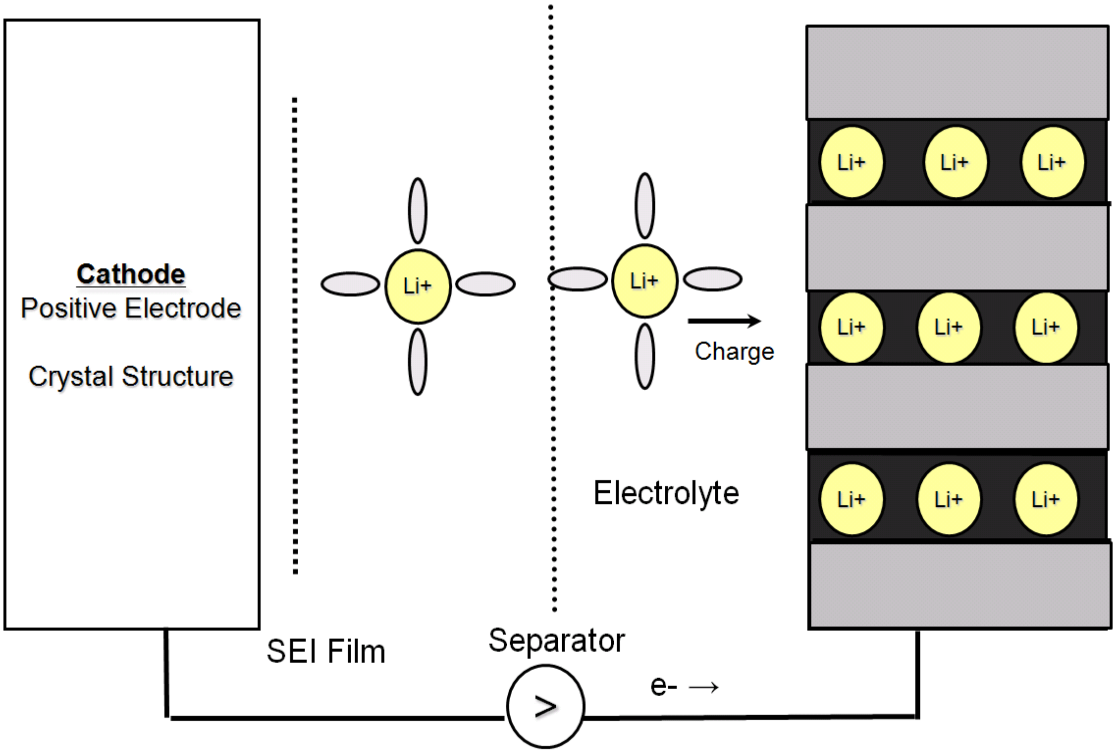
In this particular project, the development of residual performance assessment technology will be delivered and formulate the guideline for in-car battery utilization. A large number of used in-car batteries will begin to emerge and will need re-utilization.

More specifically, the running data and battery data of EV/PHEV from the customers who have purchased an EV or PHEV will be collected. Data also will be backed up from the ones taken from duration testing and regular Chassis Dynamo testing.

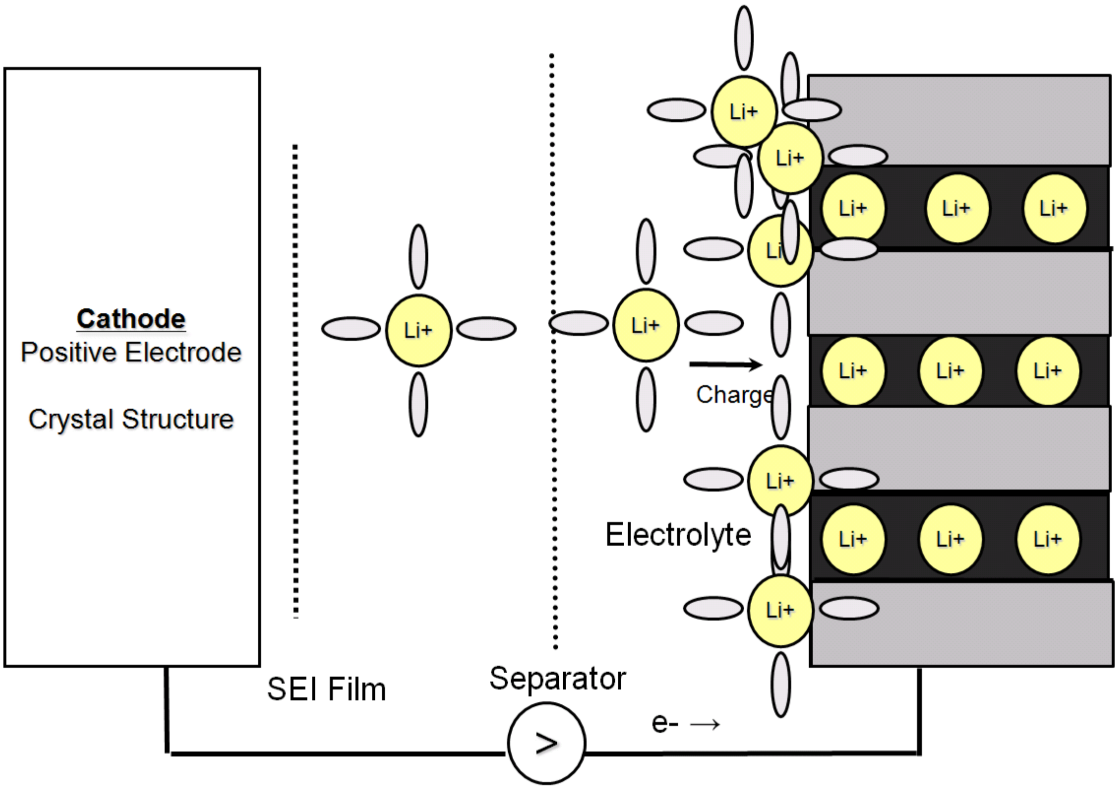
The project will perform an integrating analysis on above mentioned data and information in which to structure a battery model and develop analysis methodology for residual performance assessment. “

The project organization are consisted of 3 parties; R&D of rechargeable batteries, data analysis and modeling team, and steering committee. Institutions such are JARI, NEDO (New Energy and Industrial Technology Development Organization), and Universities (Tokyo, Kyoto, Keio, and Seikei). Professor Sasaki from Keio University belongs to running data analysis working group of data analysis and modeling team.

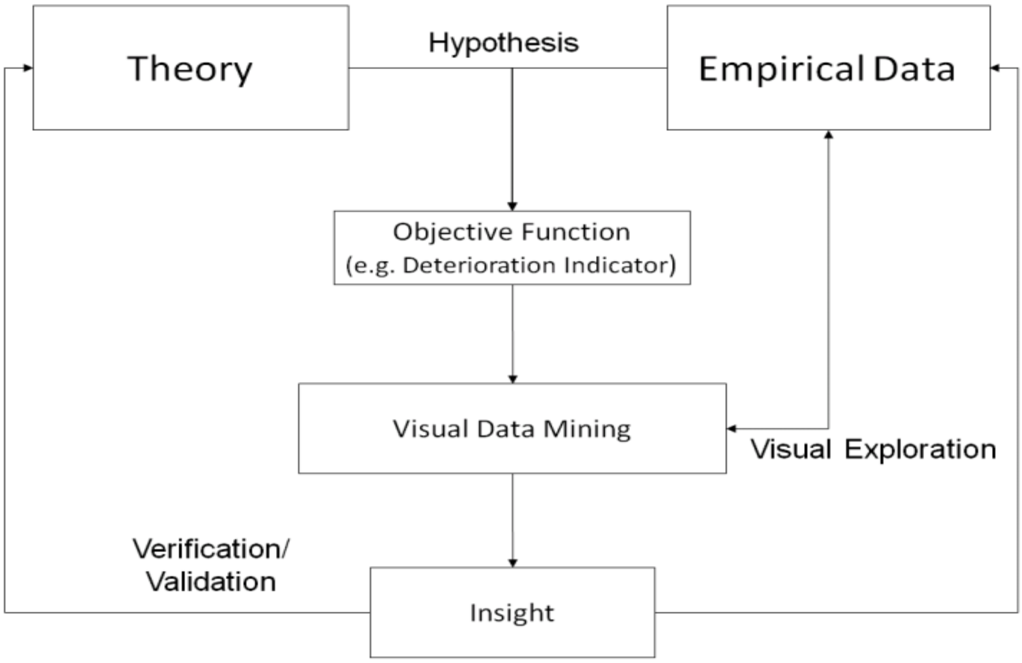
In addition to the above-mentioned data, a series of accelerated deterioration testing will be performed in order to clarify the deterioration mechanism of battery. By analyzing the internal mechanism of such a battery, the deterioration mode and its root cause will be clarified. In order to further understand clearly on battery internal model, particularly how the deterioration would occur over time, simplified model (1), (2) are shown in below. (1) model shows the initial state of battery. Li+ is an ion that is going into anode. Once an ion goes into anode, the battery is charged. However, as the time goes by, there will be the case(s) when ions fail to go inside the anode and remain in electrolyte or stick to the surface of pole. In such cases, as described in (2) model, it can be said that the battery is deteriorated.



(1)Simplified Model of Internal Model at Charge: Initial State



(2)Simplified Model of Internal Model at Charge: Deteriorated State

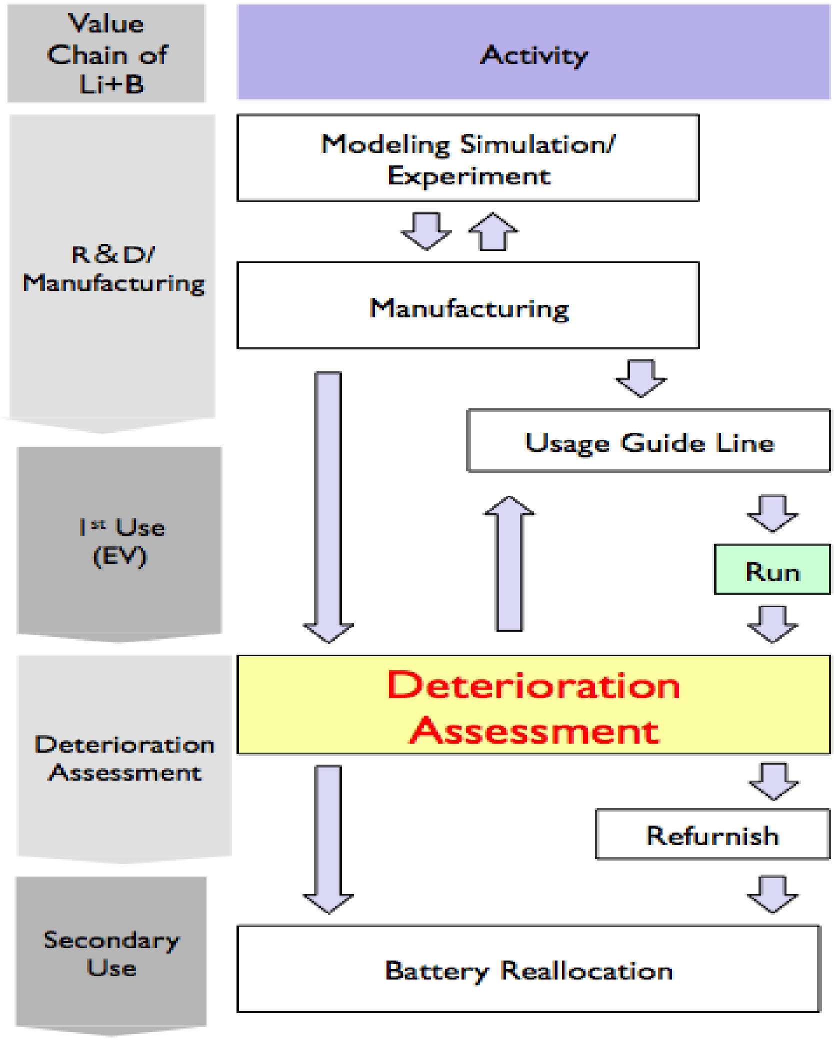


(1). Associated Flow of Battery Theory and VDM.

As described in above (1) associated chart, in this research, it values and emphasizes the blend and integration of theoretical interpretation of battery and insights from empirical data. It is extremely hard to understand the mechanisms by only looking at run data or vice versa, since both portion plays a part of supplement to one another.

As per the extensive topic to the project objective, it is to prevent secondary Li+B market become "Lemon Market" for any particular industry player, for there is a terminology in information economics called "Lemon Market". The term "Lemon Market" describes a state of market when seller knows more about product than buyer and tries to take advantage of it. In the case of rechargeable Li+B, its residual performance assessment must be open and accurate for secondary users so that the trade will be fair and secured. A series of activities is summarized as “Value-Chain of In-Car Li+B” described in below flow chart 1.

In a value chain at flow chart 1, this research is focused to contribute on deterioration assessment as highlighted in yellow. The deterioration assessment is emphasized since it will become one of the most important paths not only for secondary use of Li+B but also to broader means of effective energy utilization.



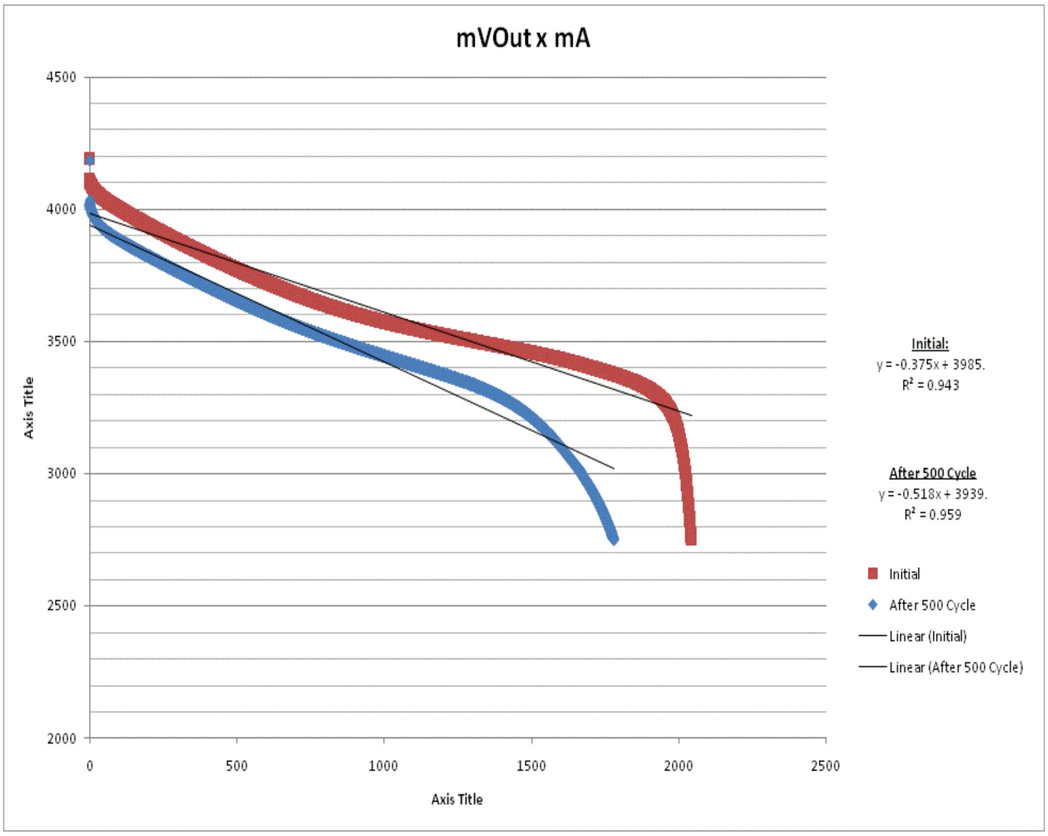
Flow Chart 1. Value Chain of In-Car Li+B

2. Visual Data Mining Applied in Context

2.1. Analysis on Bench Data

As described in previous section, two types of data sets were provided for the analysis; the laboratory experiment data (of which referred as “Bench Data” in this paper) and data from electric vehicles used for specific period (of which referred as “Run Data” in this paper). Bench data is consisted of two types of data; initial battery state and after 500 charge-discharge. Battery cell is 18650 type and it is experimented under 25 ℃ which is constantly maintained at the same level. Upper limit of Voltage is 4.2V (4200mV) and lower limit of Voltage is 2.75V (2750mV).

The figure [F1] shows data plot of mVOut and mA. Initial data shows the trend of $y = -0.375x + 3985$, $R^2 = 0.943$ where R^2 signifies the least square approximation. Data after 500 cycles shows $y = -0.518x + 3939$, $R^2 = 0.959$. Comparing



[F1]

From [F1] chart, the following three things can be observed.

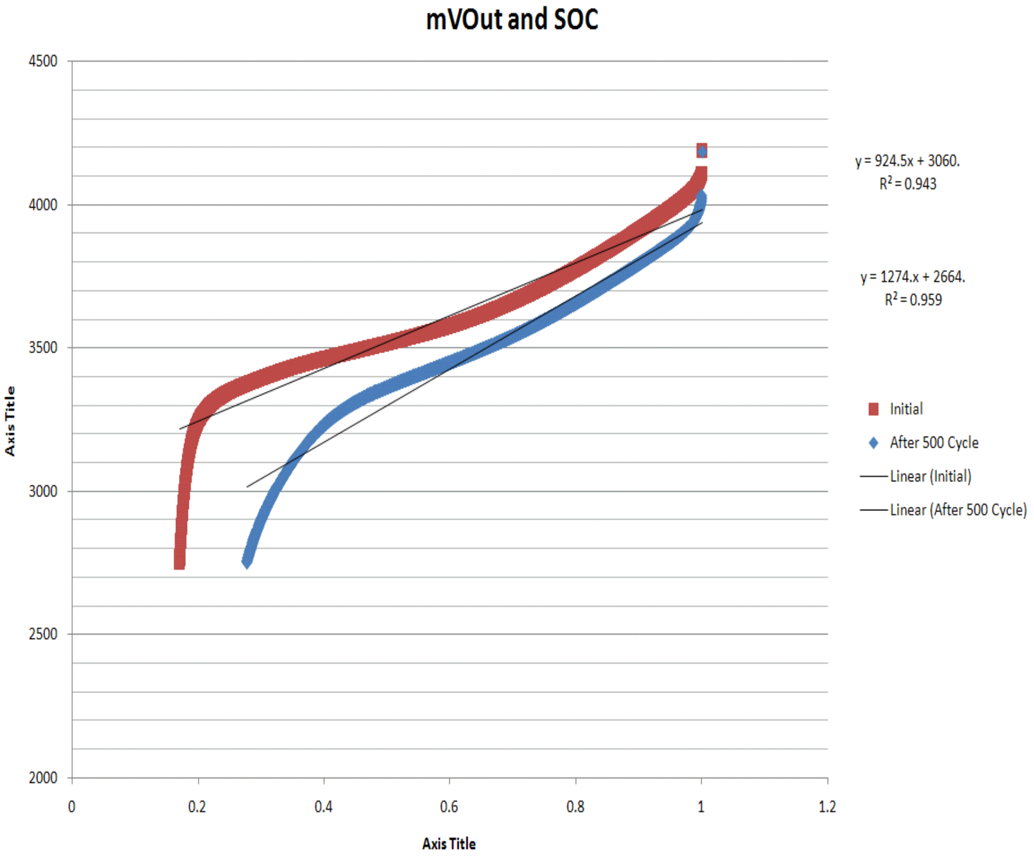
- The slope is greatly differed (from -0.375 to -0.518)
- Y intercept (b) differs by 51mVolt.
- The distribution (R^2) is more constant than other two parameters differ by 0.016.

[F2] plots the data of mVOut on Y axis and SOC (State of Charge) on X axis. SOC is derived from following formula;

$$SOC = 1 - Ib \Sigma n / Watt$$

Watt is derived from mVOut multiplied by mA.

In this chart, initial data shows y = 924.5x + 3060 and $R^2 = 0.943$, while after 500 cycle data shows y = 1274.x + 2664, $R^2 = 0.959$. It also describes the deterioration of battery is found in V Out value at the same SOC point.



[F2]

After the comparison of initial and after 500 cycle data, it can be concluded that 1). Voltage Output (V Out) value becomes smaller after 500 cycle.

2). Slope is bigger after 500 cycles.

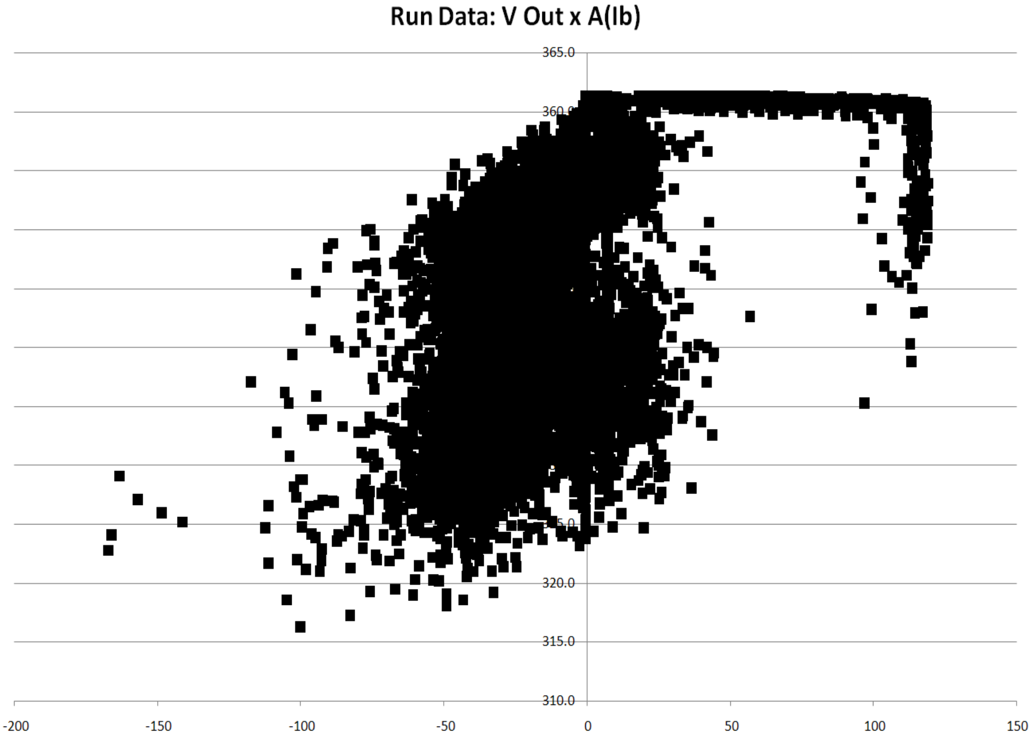
2.2. Analysis on Run Data

Run data is taken from actual running experiments conducted by project team. Running data is consisted of 2.5 million records taken over 3 month (from January to April 2011). The detail of run data is described in below table [T1].

Parameters	Frequency	Unit
Sampling Start Time	At Random	Date and time
Sampling Time	Per Sec	Date and time
Sampling Start Time	At Random	Date and time
Mode	Per Sec	Run, Rapid Charge Regular Charge, Parking
Speed	Per Sec	Km/h
V_Out	Per Sec	Voltage
A	Per Sec	Ampere
ODO Meter	Per Min	Km
AC_Output	Per Min	W
Heater	Per Min	W
Min_Temperature	Per Min	℃
Max_Temperature	Per Min	℃
Service Plug_Temp	Per Min	℃

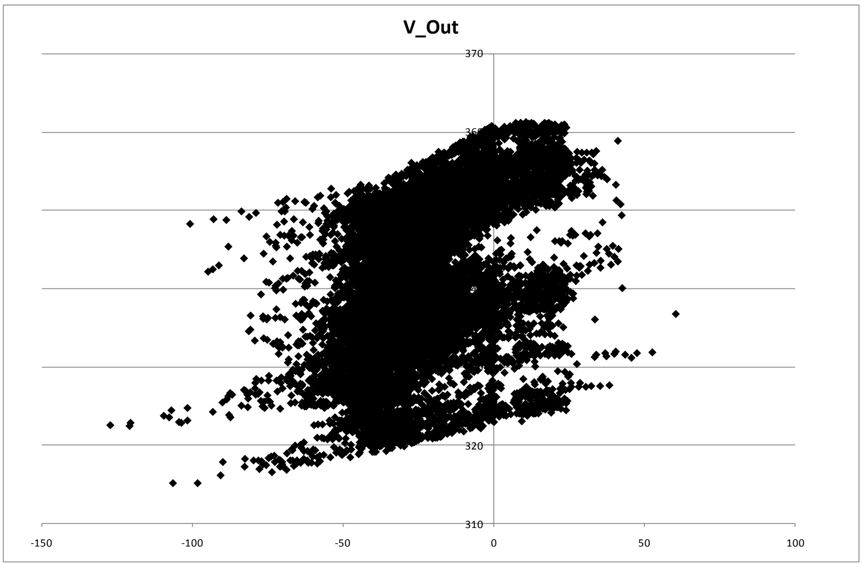
[T1]

An extracted data plot of V Out and A(Ib) is shown in [F3]. From visualization stand point, it is hard to distinguish each single data and how the trend of plot would go from 2 dimensional graphical representations.



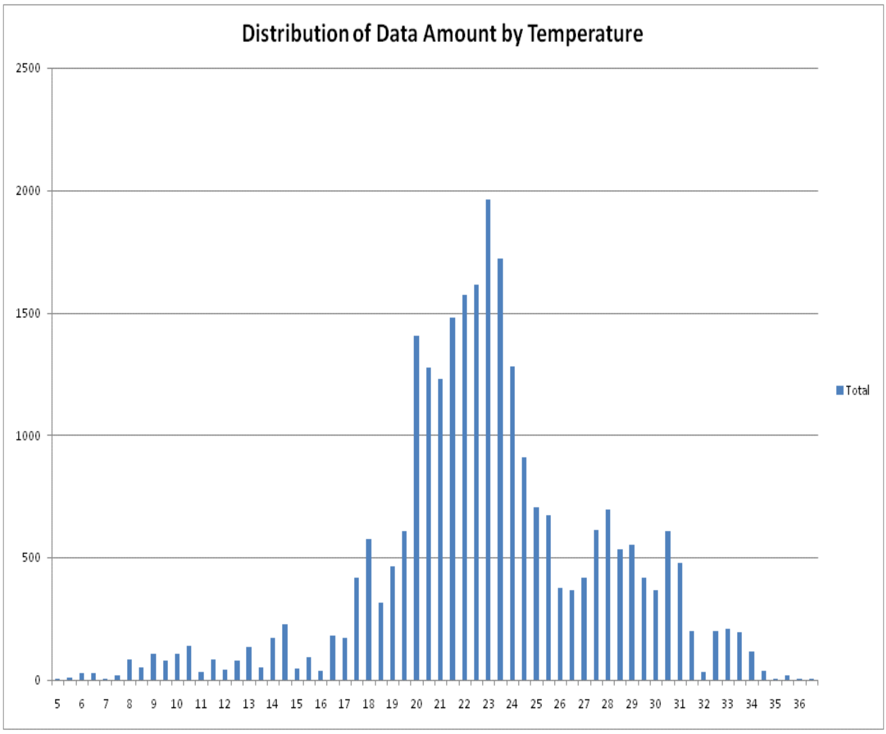
[F3] Run Data VOut and A(Ib) with Regular Charge

The distribution of temperature is described in [F4].Although the temperature has variations roughly from 5℃to 35℃, the largest portion of data resides to circa 25℃, where bench data was kept inside the laboratory experiment. From [F4], all of running state data is extracted. The data visualization is not quite apparent in which to state Ri at specific data range(window).



[F4]

Additionally, the temperatures of which running data have been undertaken, are close to 25℃ for their averages, which would make the comparison available.



[F5]

2.3. Design Specification

In this chapter, it is discussed that design specification and considerations on how the contexts and requirements have come into the shape. As per prerequisite to data visualization, all of the data is stored into one Microsoft Access Database, and the copy link of database is stored in a shared file server so that user can query the data of preference. SQL Studio is used to query the data and put the file into CSV file of which the visualization program refers to.

Design specification of visualization was derived from requirement analysis which was explained in previous section. Based on the requirements, data plotting tool must have following 1)-5) functionalities;

1). To plot the data in three dimension

2). To color the data at certain point of charge/discharge

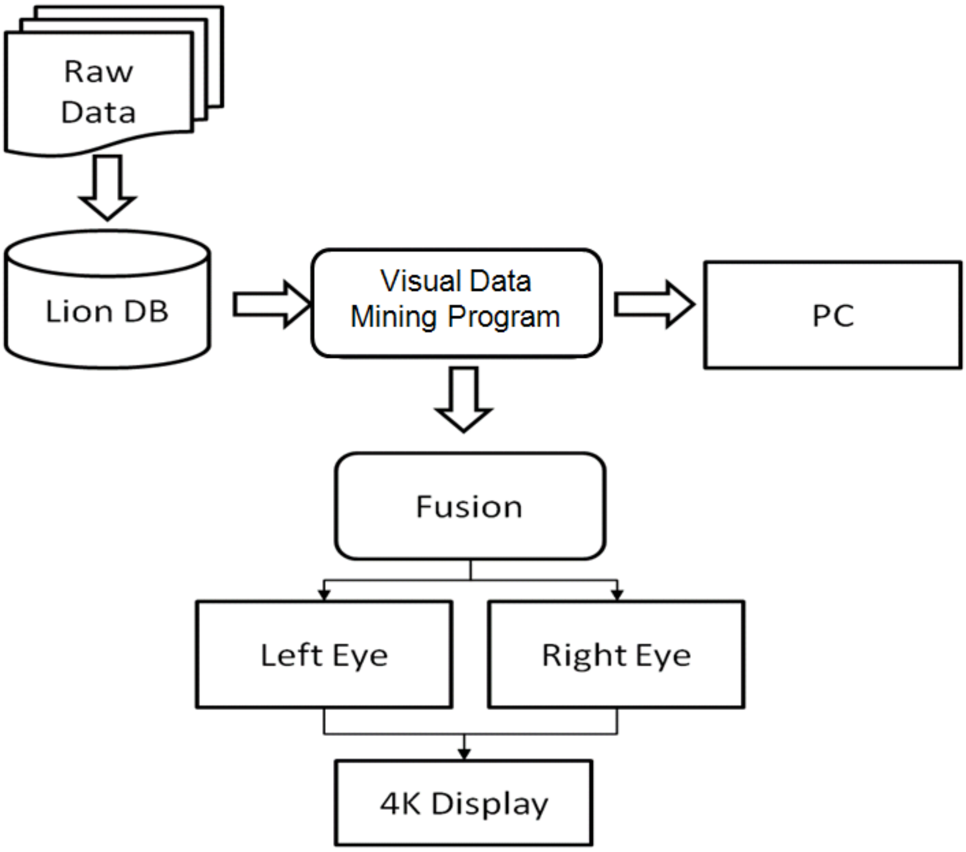
3). To select the data range (expand/shrink)

4). To move the data range

5). To calculate the least squares approximation of selected data range

6). To be able to identify raw data.

The system architecture is described as following exhibit 1. Raw data is formatted in CSV. The data is stored in database, and then the data is queried (usually from one in all data) at request. A set of two 4K projectors display the image to a 180 inch display. In this environment, 1 pixel size equals to almost 1mm square. If a person with 1.0 eyesight sees the screen from 4m of distance, it generates the resolution of which it exceeds cognitive limitation. Each of “Left Eye” image and ”Right Eye” image is calculated by a separate set of computers. “Fusion” is mounted on a separate computer and it extracts Open GL commands from visualization program.



(Exhibit 1)

Control of application is described as followed;

→ key signifies to expand the target range by 1000

← key signified to shrink the target range by 10000

↑ key signifies to move the target range forward by 1000

↓ key signifies to move the target range backward by -1000

c key signifies the draws linear approximation equation of target range and the least square approximation.

q key signifies to move the target range forward by 100

a key signifies to move the target range forward by 10

z key signifies to move the target range forward by 1

u key signifies to move the target range backward by 100

j key signifies to move the target range backward by 10

m key signifies to move the target range backward by 1

w key signifies to expand the target range by 100

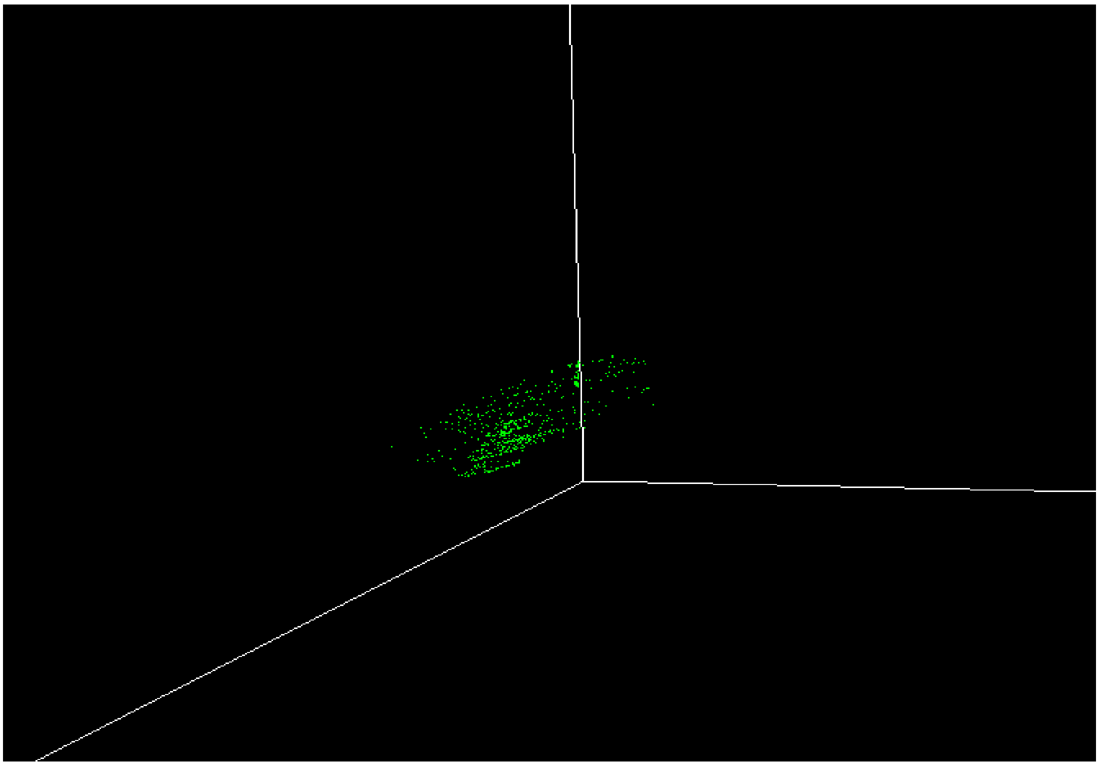
s key signifies to expand the target range by 10

x key signifies to expand the target range by 1

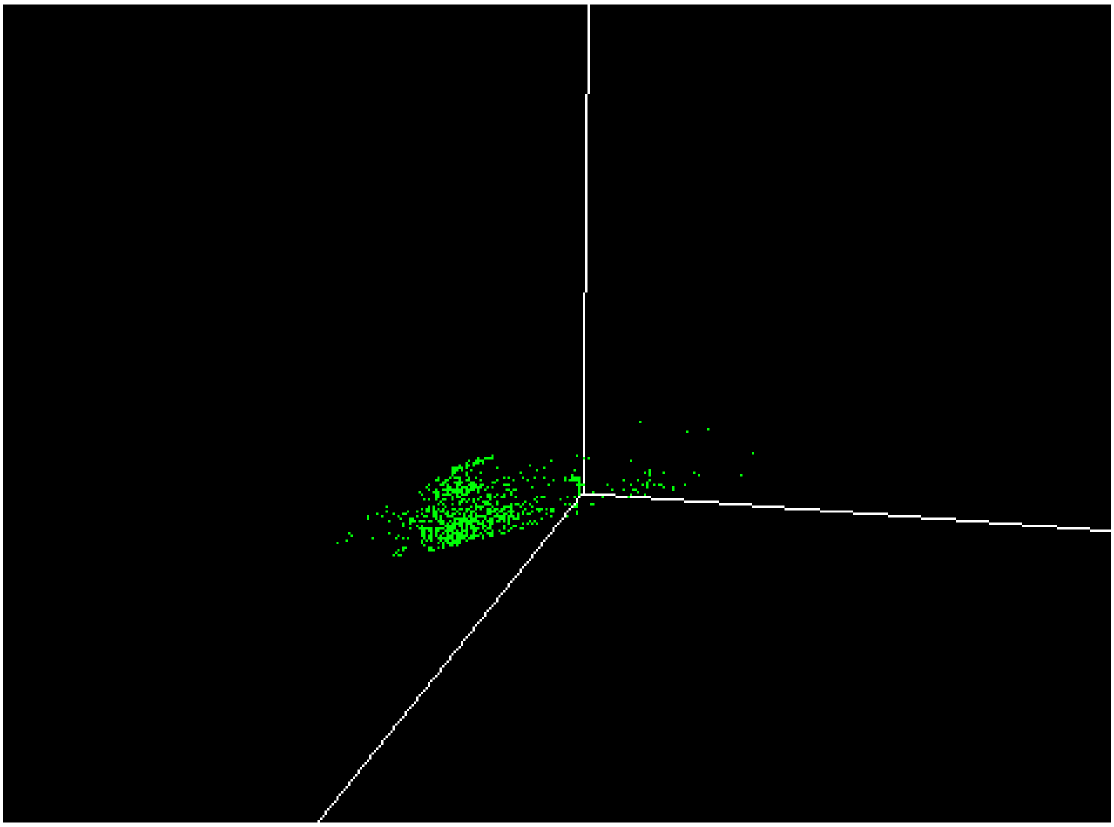
y key signifies to shrink the target range by 100

h key signifies to shrink the target range by 10

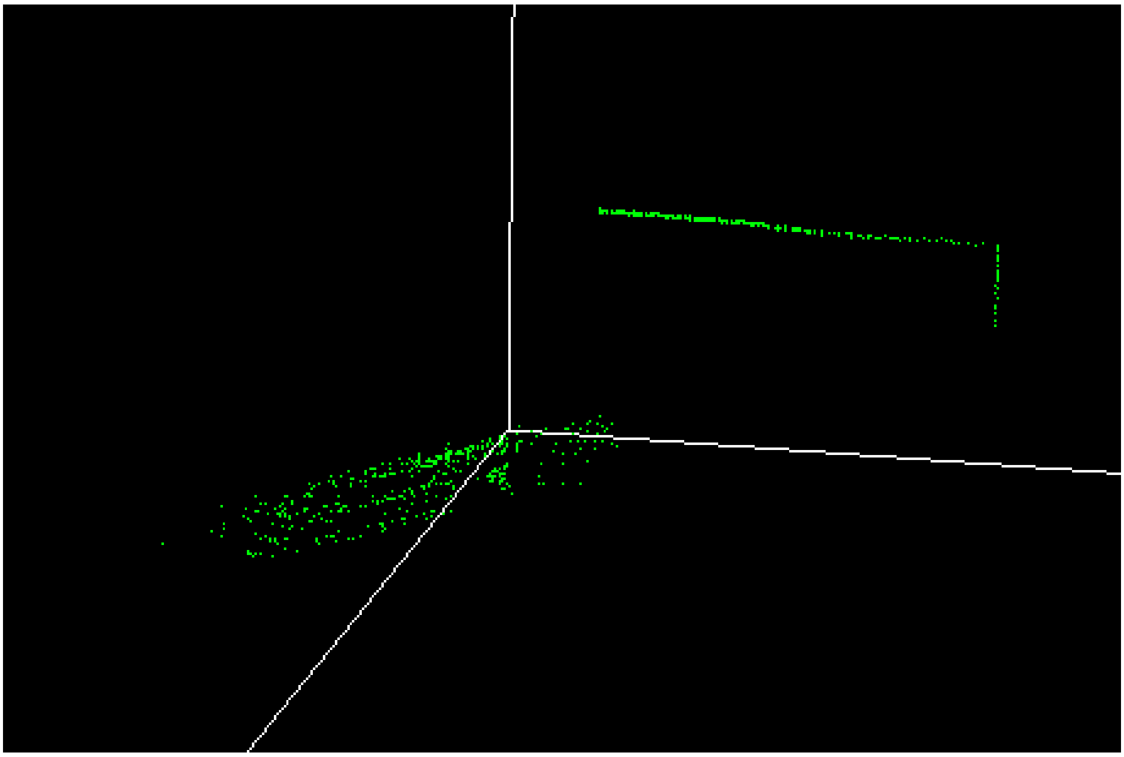
At start-up screen, the image identical to following [F6] must be displayed. The display data range can be adjusted by above-mentioned control keys. In below, a series of movement by pressing “↑” for a cycle (↑ key signifies to move the target range forward by 1000) is performed.



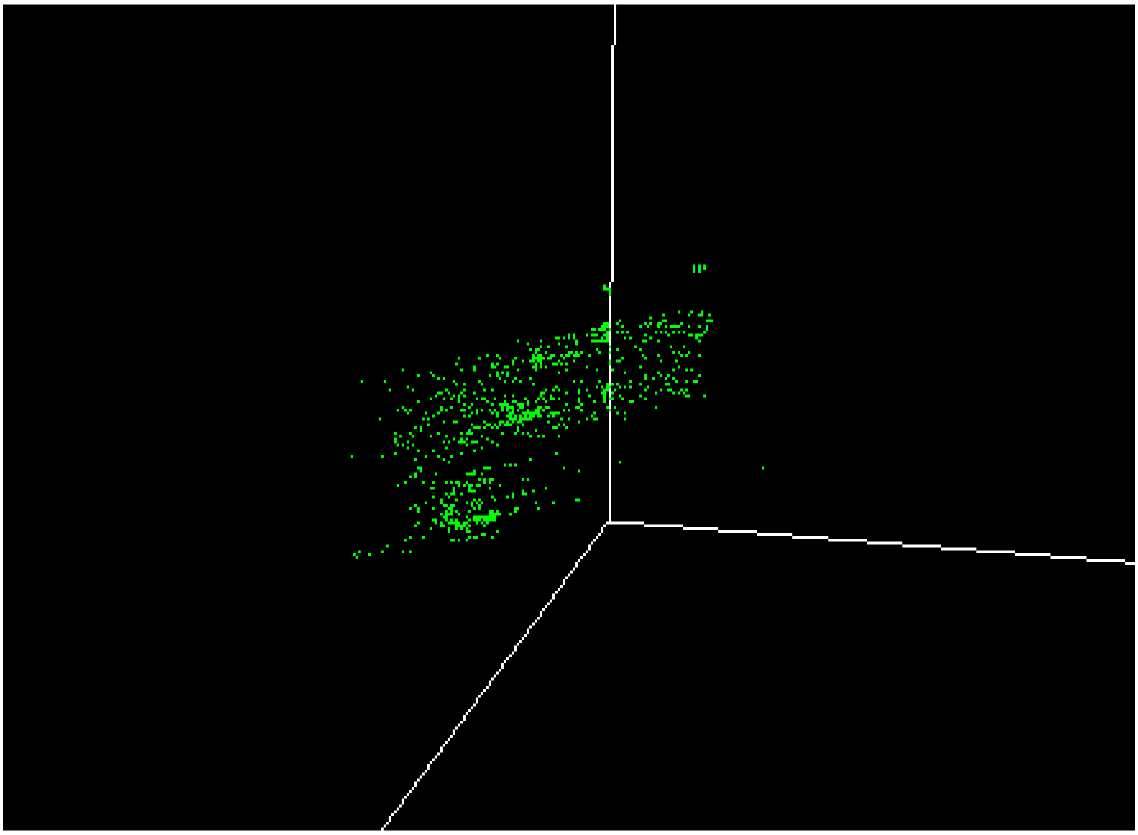
(Output 1)



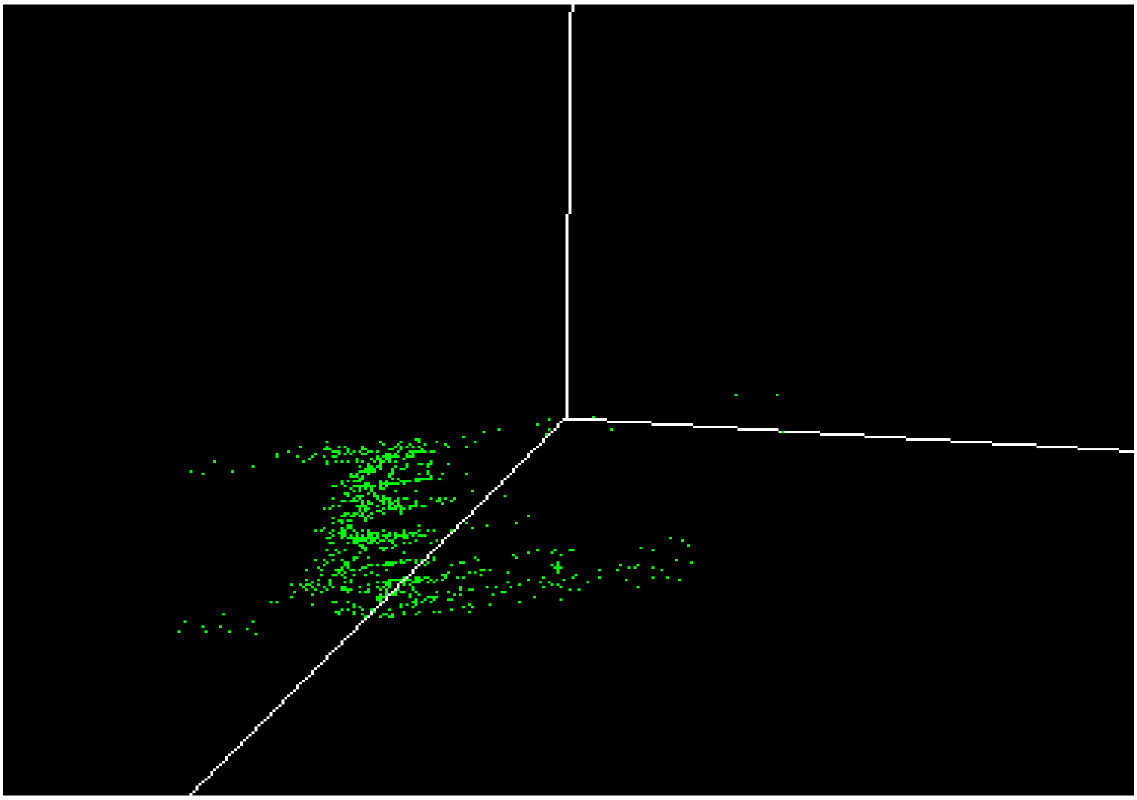
(Output 2)



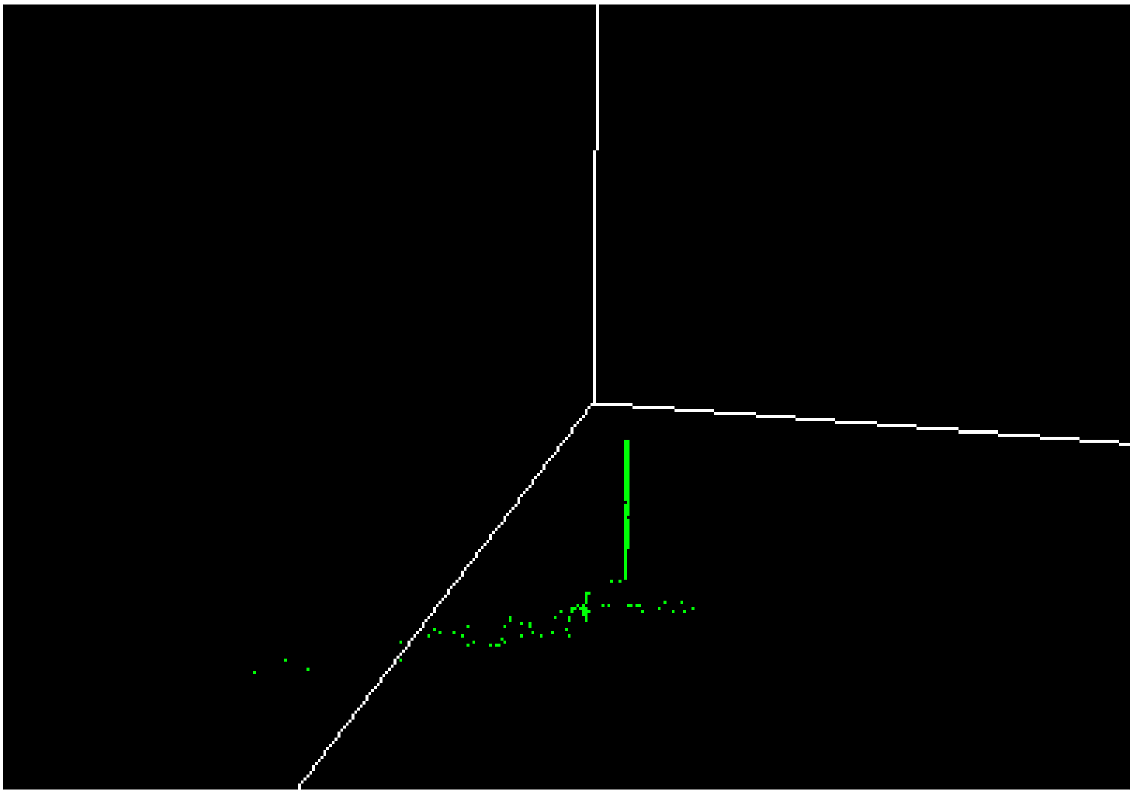
(Output 3)



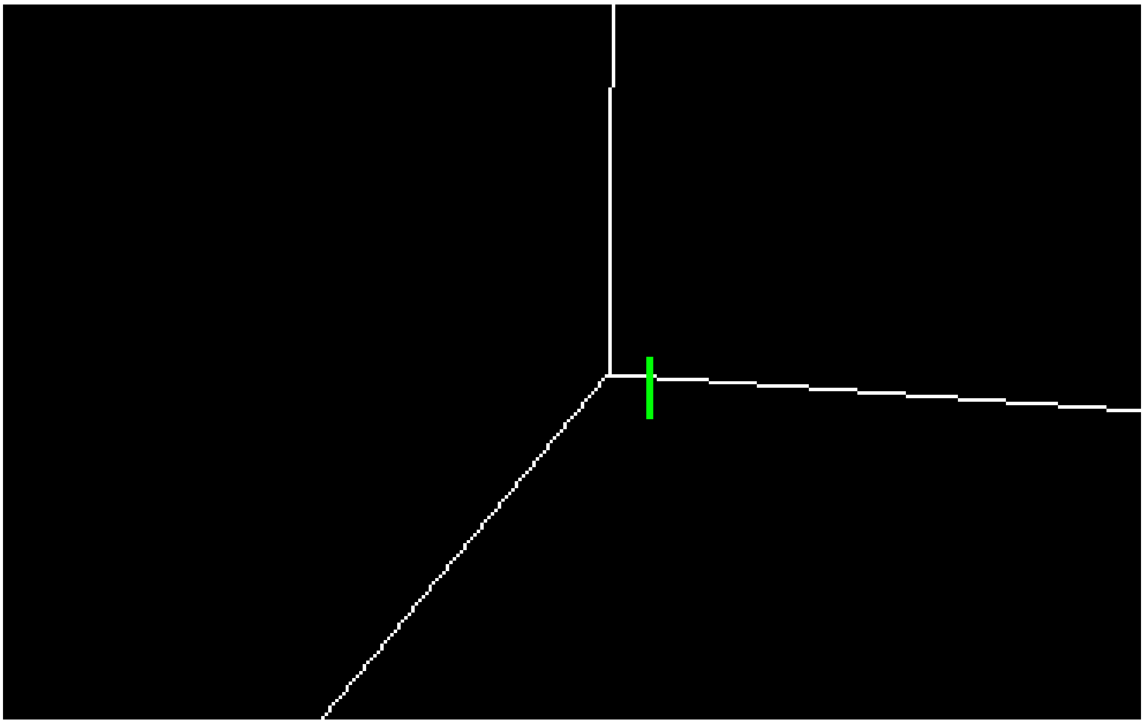
(Output 4)



(Output 5)

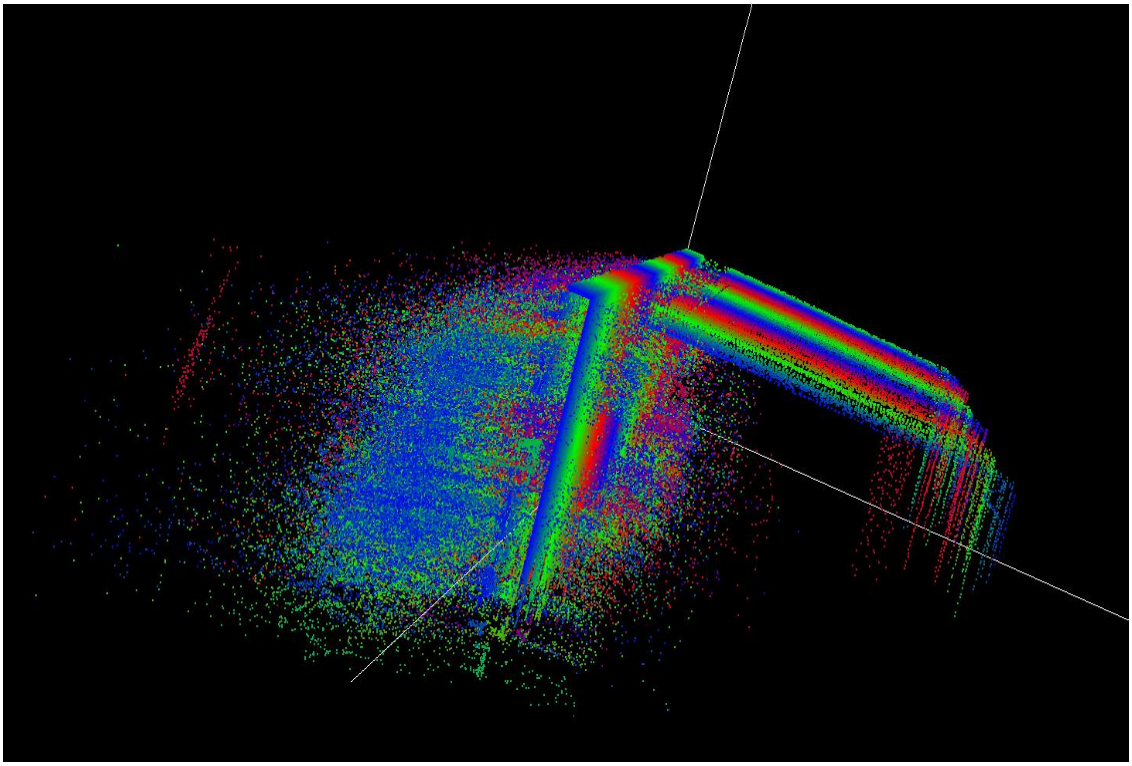


(Output 6)

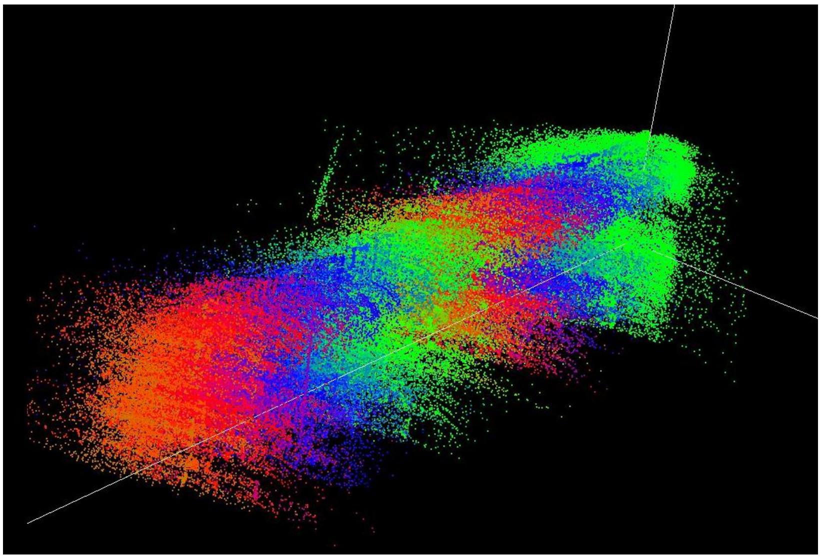


[F6] Moving forward by 1000

In below, [F7] figure shows the data plot of all run data.



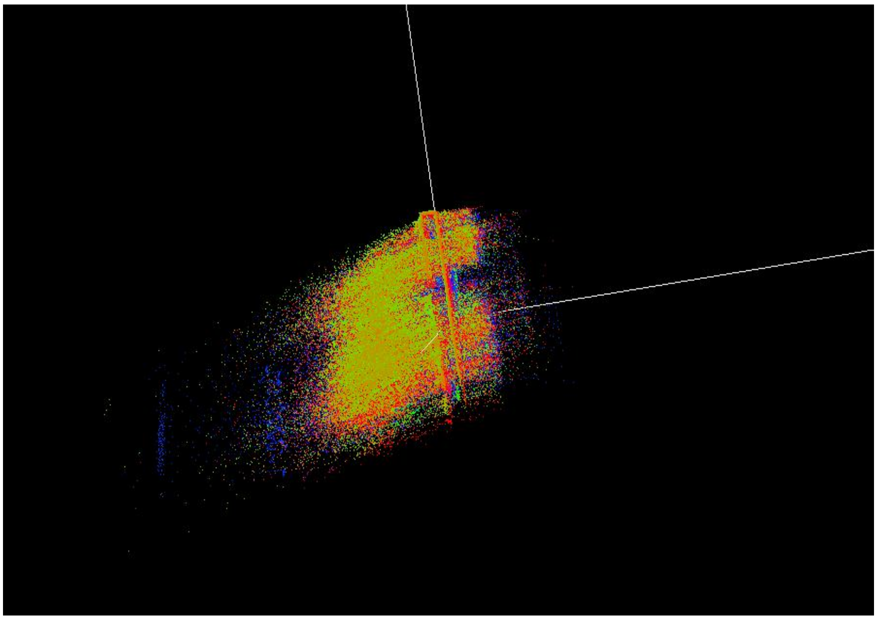
[F7]



[F8]

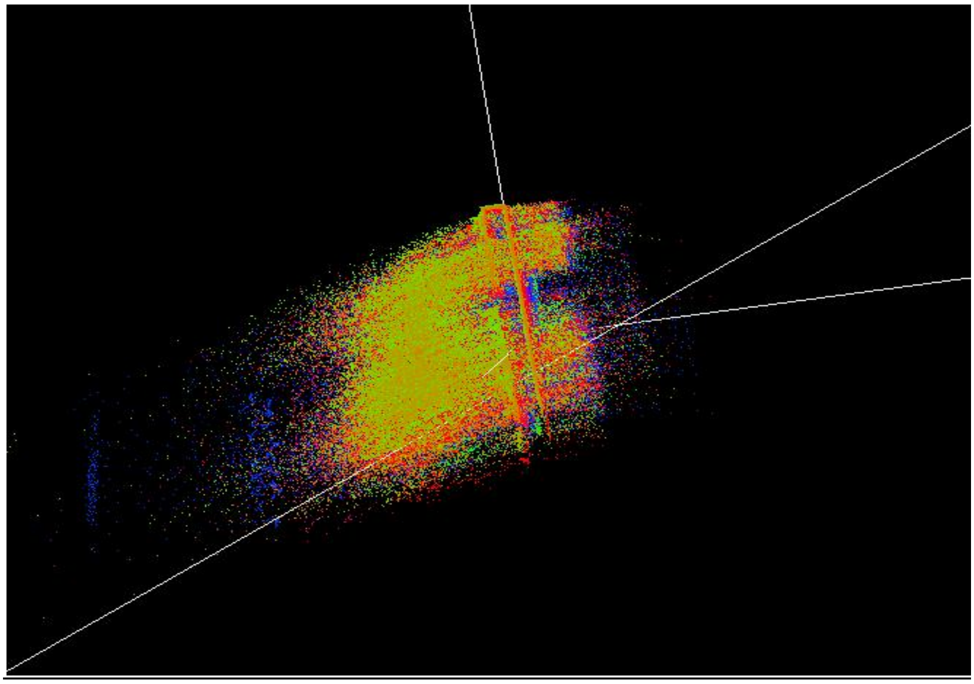
In above, only run data is selected. The color is determined by the V_Out (Voltage value circa 360. Each time the program reads V_Out around 360, the color gradually changes.

Following [F9] is the output of data which include all of running data except rapid charge state.



[F9] All running data except rapid charge

Following [F10] is the output of data which include all of running data except rapid charge and the least square approximation. The least square approximation is Y=ax+b where a=0.444064,b=-0.217158,s=11.858352.



[F10] All running data except rapid charge with the least approximation square

For performance improvements, “displaylist” function of Open GL language is implemented to the system. Displaylist function caches the data representation into memory of computer so that the program does not have to run each time to call 2 million points of data. The performance is improved by double in nominal environment (such as desktop/notebook computer) and 12 times faster in 4K environment. The detail of performance comparison is described in index 1.

Index1: Performance Comparison of Programs

	Original	Displaylist Employed
Nominal	30 Hz	60Hz

4K Display	0.3Hz	6Hz

3. Verification and Validation

3.1. Definition

As far as system engineering is concerned, “V&V (Validation and Verification)” is to testify whether the system is reasonable in objective, as well as it is functionally capable of completing the tasks that were extracted from the context and user requirement. Therefore, validation of system must correspond, and contribute to the objective of research; to know the deterioration mechanism of lithium ion battery. On the other hand, verification of system is intended to clarify whether the system is functioning right.

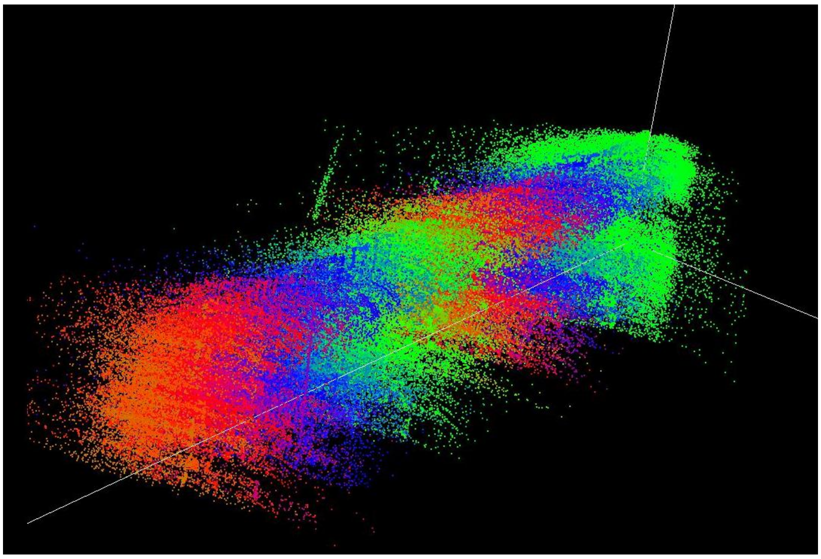
3.2. Verification

The verification of system is done by confirmation of system design which is derived from user requirement. As shown in figures from [F4] to [F6], the target data sets can be arranged at user’s preference. Moreover, the least square approximation is calculated and a linear equation is drawn when the button is pressed at any given target data range.

Control key functionalities are confirmed and shown in below images as [Verification Evidence n : functionality].

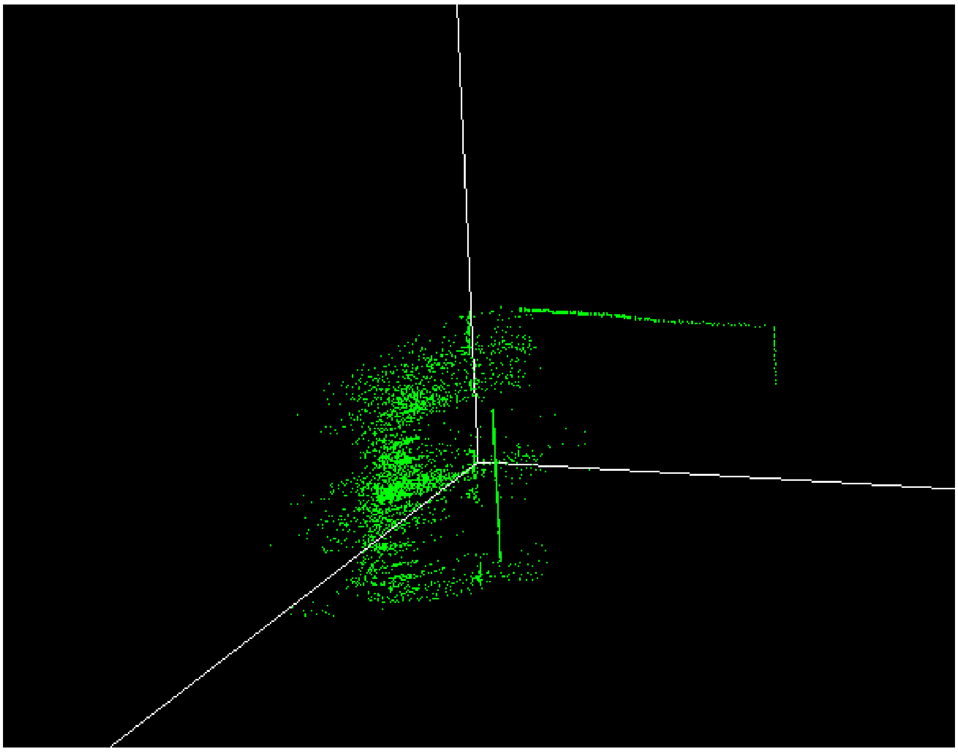
[Verification Evidence 1: 3Dimentional Data Representation]

[Verification Evidence 2: Color the data]

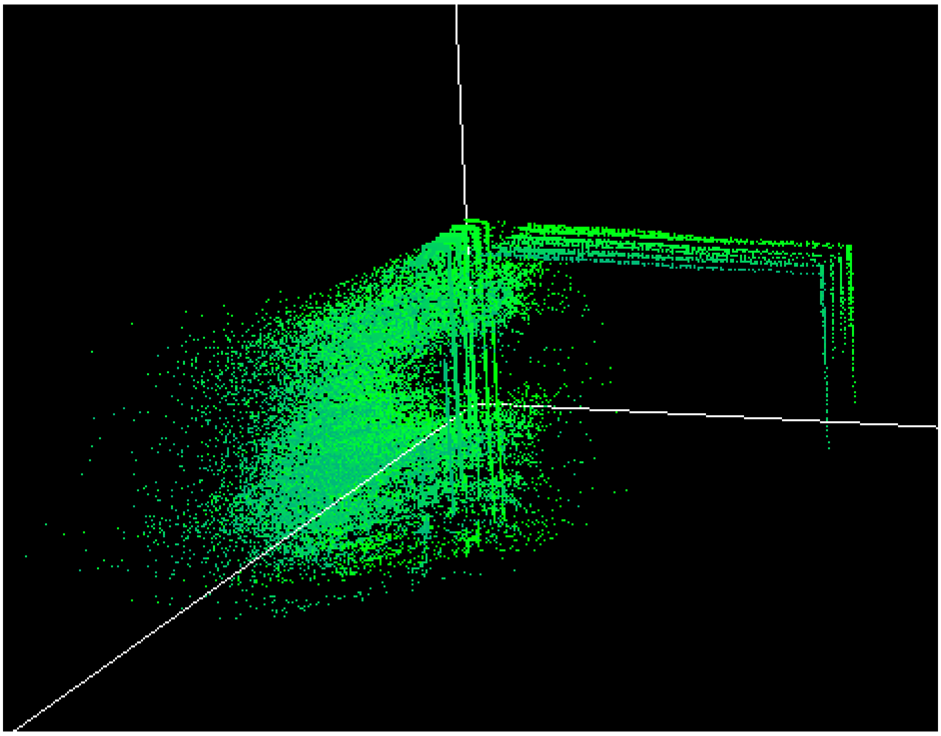


[F11]

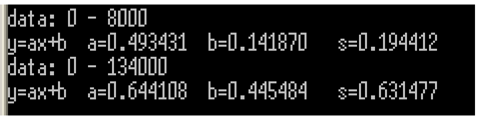
[Verification Evidence 3: Expand Shrink Target Data Window]



[F12]



[F13]

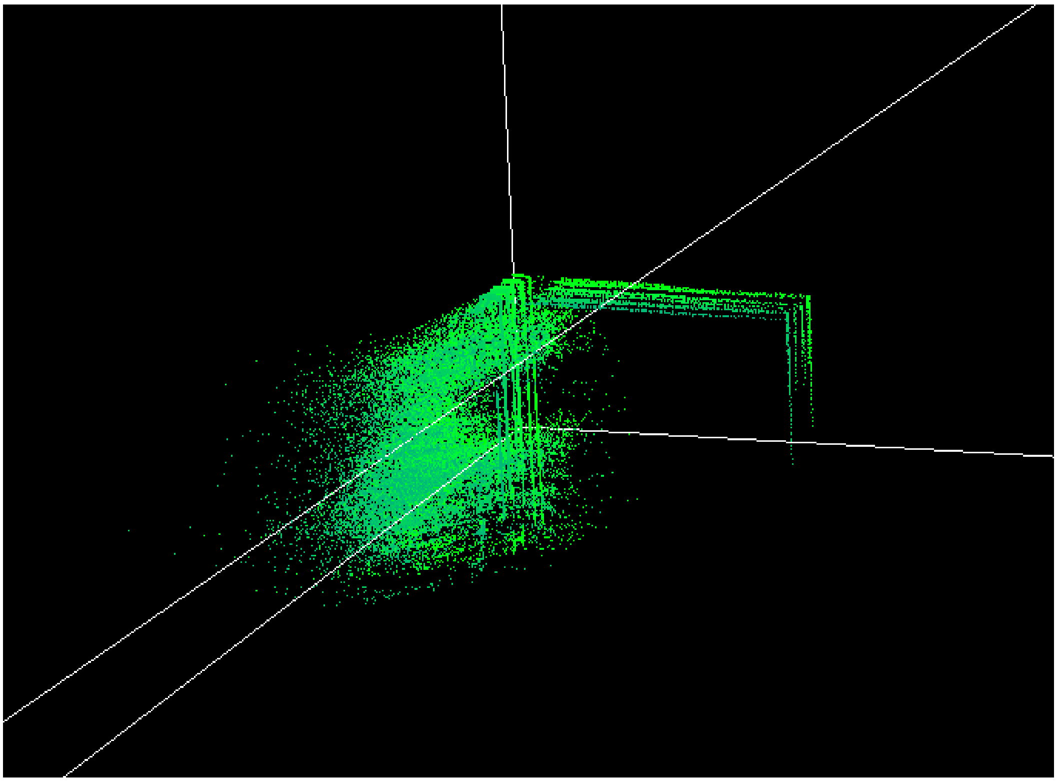


```
data: 0 - 8000
y=ax+b  a=0.493431  b=0.141870   s=0.194412
data: 0 - 134000
y=ax+b  a=0.644108  b=0.445484   s=0.631477
```

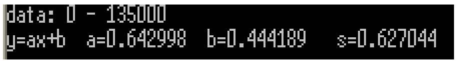
[F14]

[Verification Evidence 4: Calculation of the least square approximation]

[Verification Evidence 5: Data Identification]



[F15]

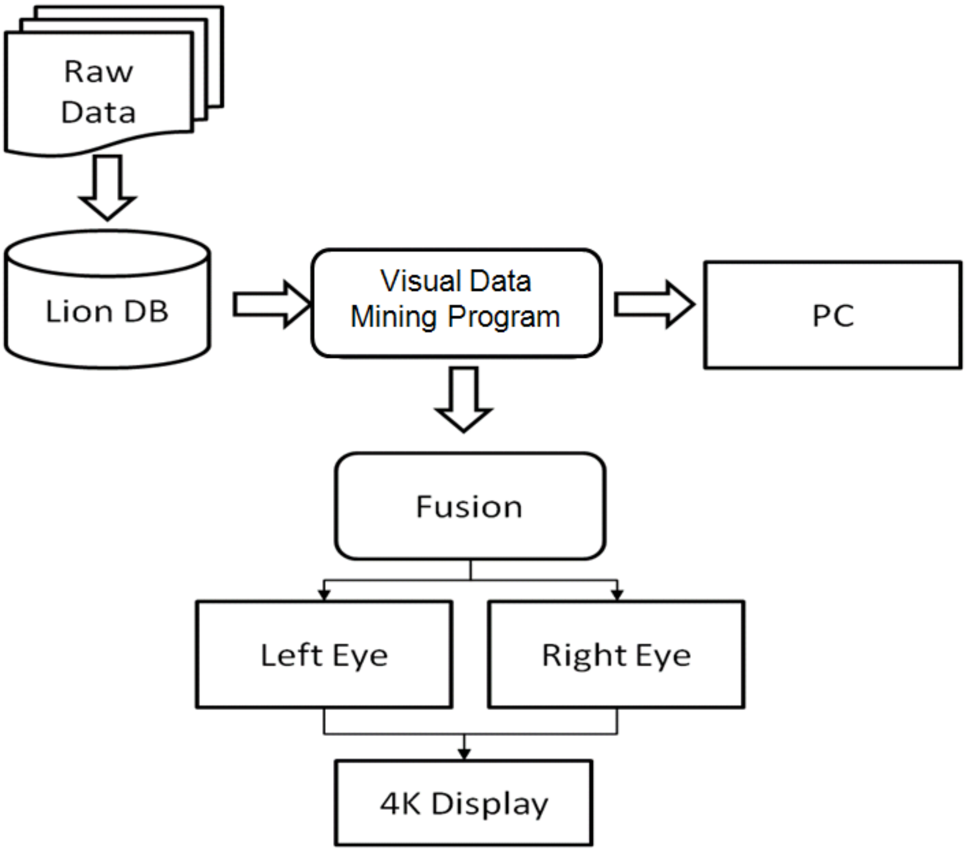


```
data: 0 - 135000
y=ax+b  a=0.642998  b=0.444189   s=0.627044
```

[F16]

3.3. Validation

A systemic methodology for identifying internal resistance from running data may be proposed in following architecture (Exhibit 1). System is capable of high resolution visualization environment (4K) and generic personal computers. Since the window of identify-able Ri is expected to be small (approximately 10 to 100 seconds) and the volume of entire data sets are big (approximately a million seconds), visualization application in high resolution environment is required. Image 1 shows the use of system at 4K display.



(Exhibit 1; recaptured)

Image 1:　Image of VDM Application in Use at 4K Display

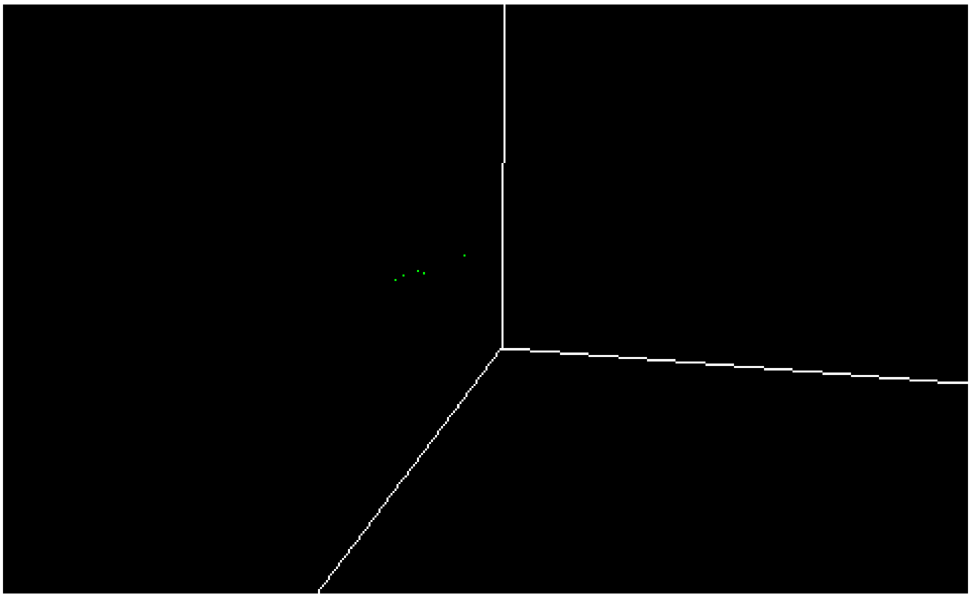
4. Conclusion

From the validation result, it can be confirmed that the time window to find Ri is 30 seconds in running data. Since the objective of VSM application tool is to find out Ri from running data, the optimal time window will contribute to identifying the deterioration of Li+B. From theoretical stand point, Ri is considered as a liner expression of V Out and Ib which can be formulated as V = Ri * Ib +E0.

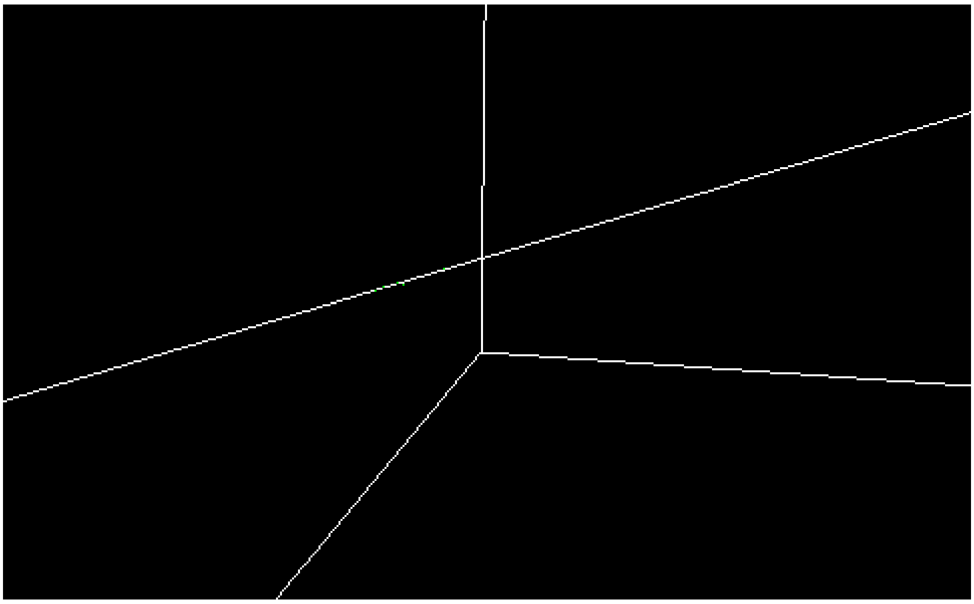
Variations of target data range are testified in order to validate the objectives of visualization tool. From running data, the data sets starting from 717 to

803 and 931 to 975 are selected since it satisfied the conditions in order for Ri to be stable; running state with SOC near 100% and Ri circa 0 (-0.5 ≦Ri≦ 0.5).Using the visualization tool, the relationship of time window and distribution (the least square approximation) is observed. The data sets should satisfy V=Ri * Ib+Eo0 where E0 stands for Potential Electron. The variations of windows are designated to 5, 10, 30. Following [Result] images describe the actual output from application tool.

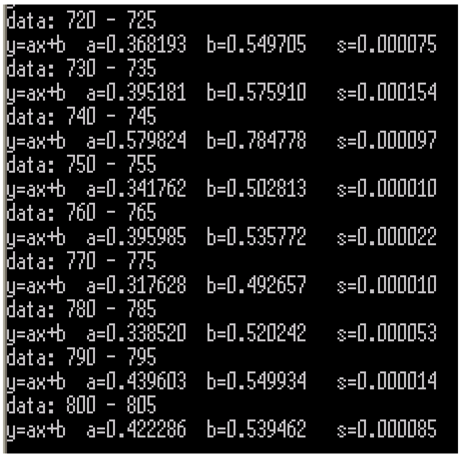
[Result717-803, Window=5]



[F17]



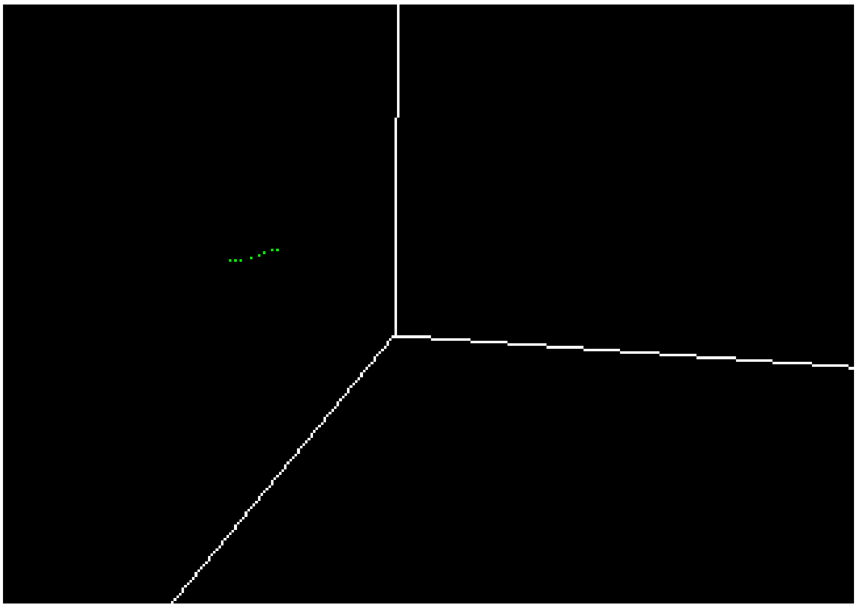
[F18]



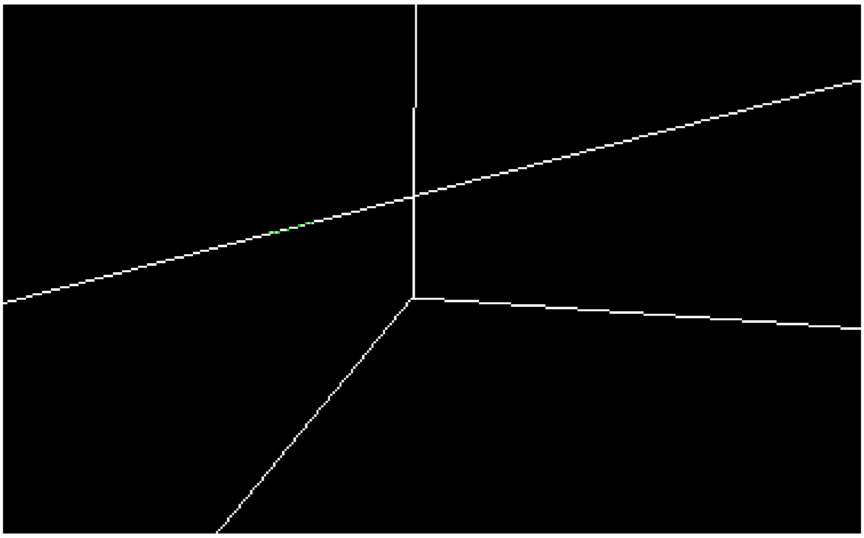
```
data: 720 - 725
y=ax+b  a=0.368193  b=0.549705   s=0.000075
data: 730 - 735
y=ax+b  a=0.395181  b=0.575910   s=0.000154
data: 740 - 745
y=ax+b  a=0.579824  b=0.784778   s=0.000097
data: 750 - 755
y=ax+b  a=0.341762  b=0.502813   s=0.000010
data: 760 - 765
y=ax+b  a=0.395985  b=0.535772   s=0.000022
data: 770 - 775
y=ax+b  a=0.317628  b=0.492657   s=0.000010
data: 780 - 785
y=ax+b  a=0.338520  b=0.520242   s=0.000053
data: 790 - 795
y=ax+b  a=0.439603  b=0.549934   s=0.000014
data: 800 - 805
y=ax+b  a=0.422286  b=0.539462   s=0.000085
```

[F19]

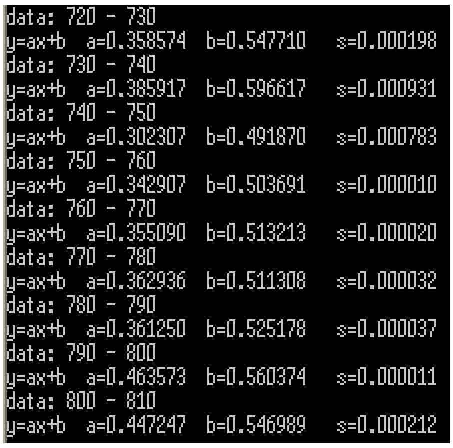
[Result717-803, Window=10]



[F20]



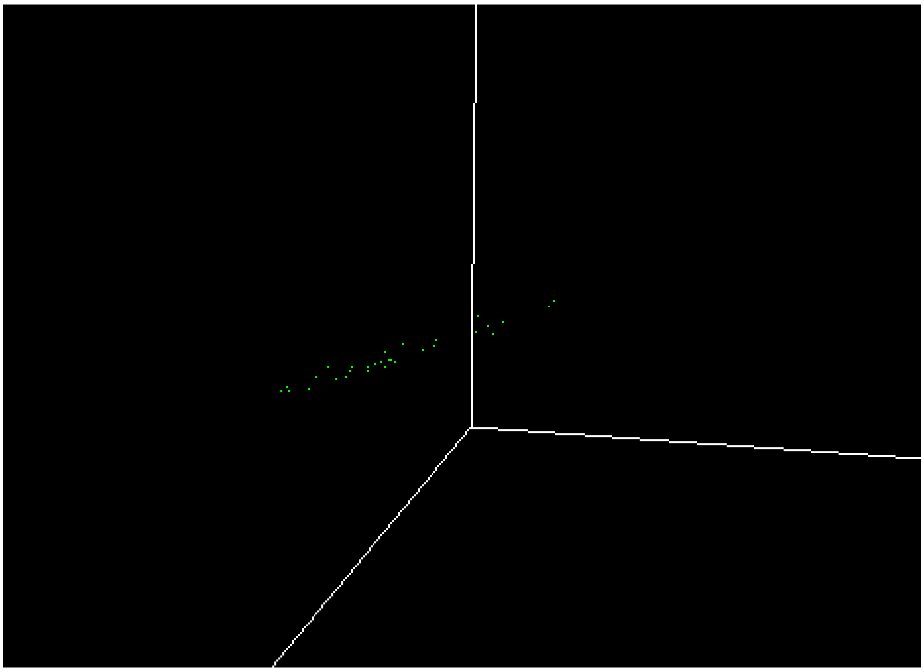
[F21]



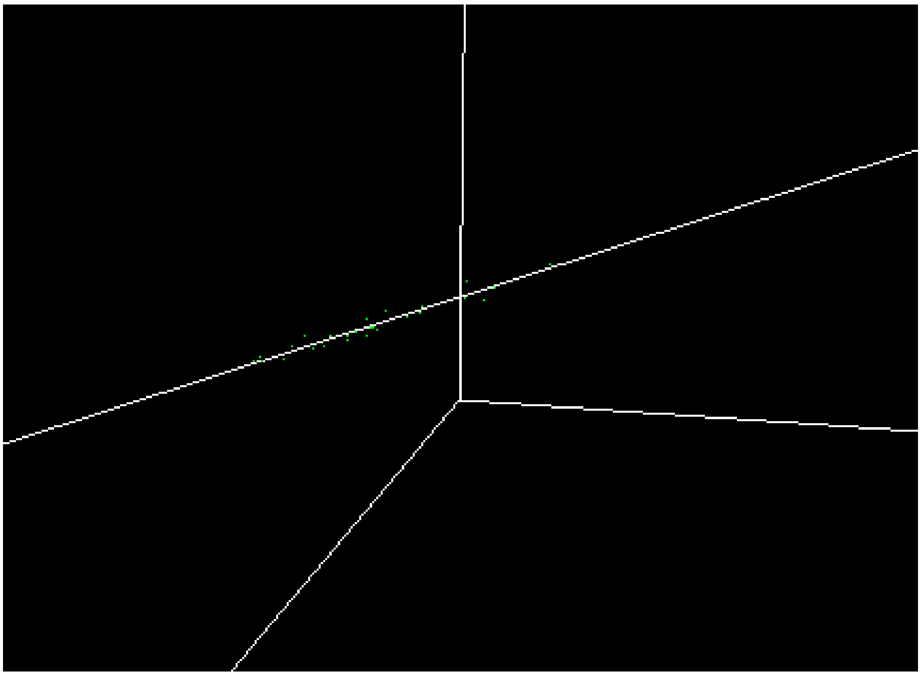
```
data: 720 - 730
y=ax+b  a=0.358574  b=0.547710   s=0.000198
data: 730 - 740
y=ax+b  a=0.385917  b=0.596617   s=0.000931
data: 740 - 750
y=ax+b  a=0.302307  b=0.491870   s=0.000783
data: 750 - 760
y=ax+b  a=0.342907  b=0.503691   s=0.000010
data: 760 - 770
y=ax+b  a=0.355090  b=0.513213   s=0.000020
data: 770 - 780
y=ax+b  a=0.362936  b=0.511308   s=0.000032
data: 780 - 790
y=ax+b  a=0.361250  b=0.525178   s=0.000037
data: 790 - 800
y=ax+b  a=0.463573  b=0.560374   s=0.000011
data: 800 - 810
y=ax+b  a=0.447247  b=0.546989   s=0.000212
```

[F22]

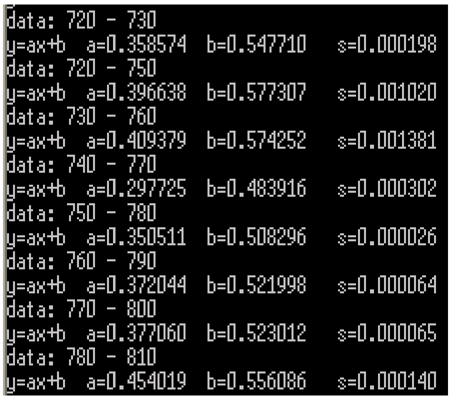
[Result717-803, Window=30]



[F23]



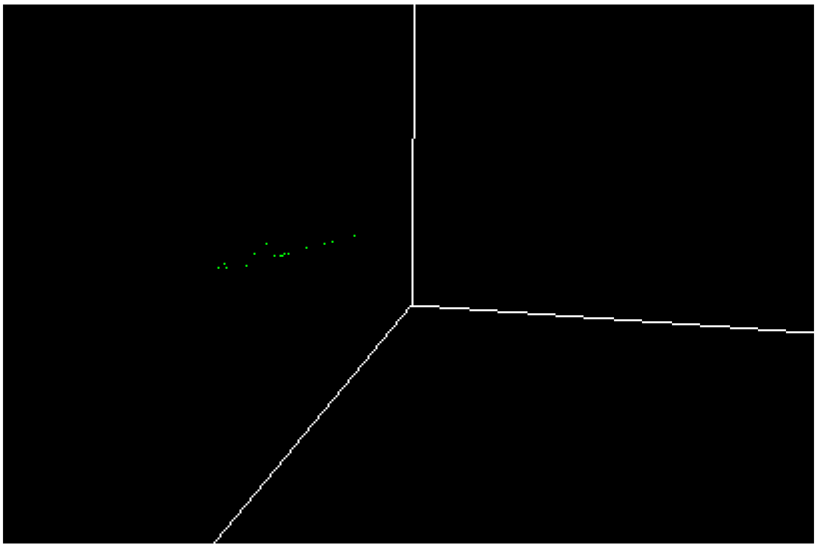
[F24]



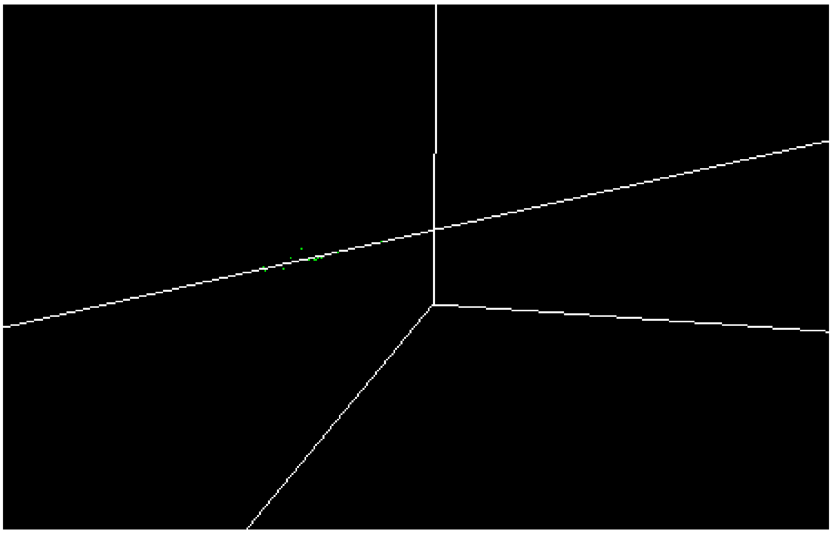
```
data: 720 - 730
y=ax+b  a=0.358574  b=0.547710   s=0.000198
data: 720 - 750
y=ax+b  a=0.396638  b=0.577307   s=0.001020
data: 730 - 760
y=ax+b  a=0.409379  b=0.574252   s=0.001381
data: 740 - 770
y=ax+b  a=0.297725  b=0.483916   s=0.000302
data: 750 - 780
y=ax+b  a=0.350511  b=0.508296   s=0.000026
data: 760 - 790
y=ax+b  a=0.372044  b=0.521998   s=0.000064
data: 770 - 800
y=ax+b  a=0.377060  b=0.523012   s=0.000065
data: 780 - 810
y=ax+b  a=0.454019  b=0.556086   s=0.000140
```

[F25]

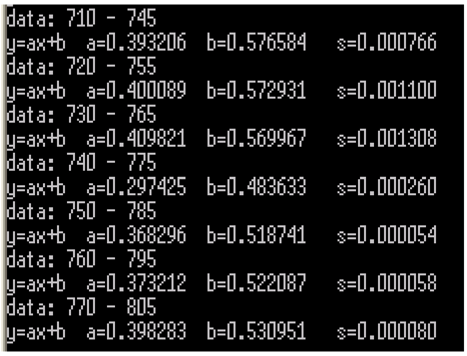
[Result717-803,Window=35]



[F26]



[F27]



```
data: 710 - 745
y=ax+b  a=0.393206  b=0.576584   s=0.000766
data: 720 - 755
y=ax+b  a=0.400089  b=0.572931   s=0.001100
data: 730 - 765
y=ax+b  a=0.409821  b=0.569967   s=0.001308
data: 740 - 775
y=ax+b  a=0.297425  b=0.483633   s=0.000260
data: 750 - 785
y=ax+b  a=0.368296  b=0.518741   s=0.000054
data: 760 - 795
y=ax+b  a=0.373212  b=0.522087   s=0.000058
data: 770 - 805
y=ax+b  a=0.398283  b=0.530951   s=0.000080
```

[F28]

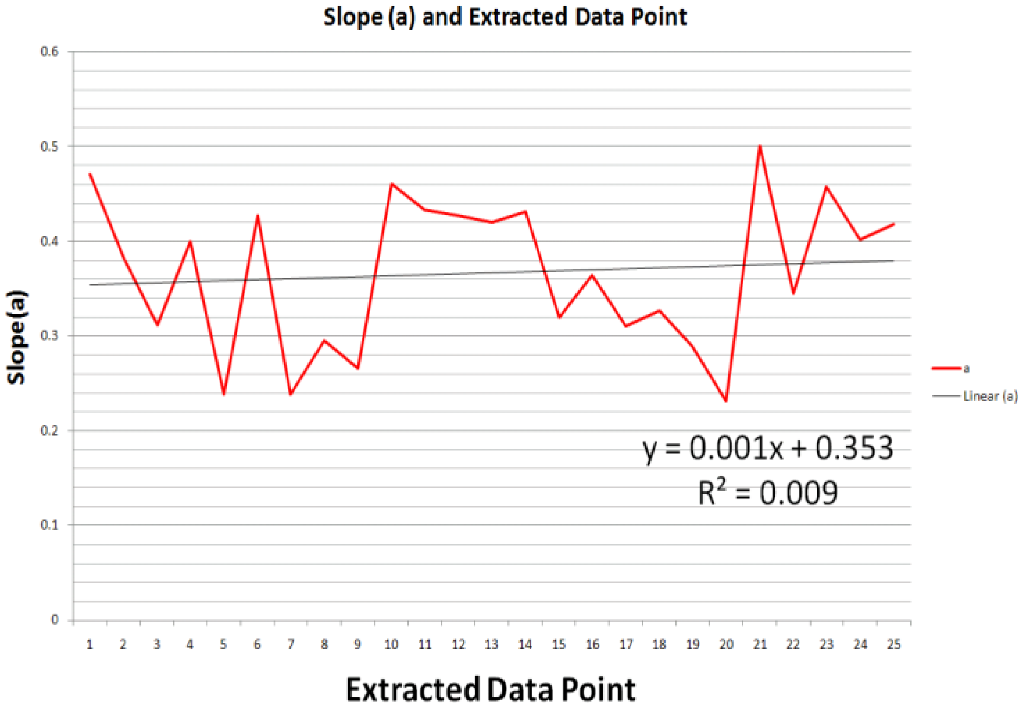
Window	Distribution of S
[Result717-803, Window=5]	4.95222E-05
[Result717-803,Window=10]	0.000355753
[Result717-803,Window=30]	0.000511268
[Result717-803,Window=35]	0.000572342

[T2]

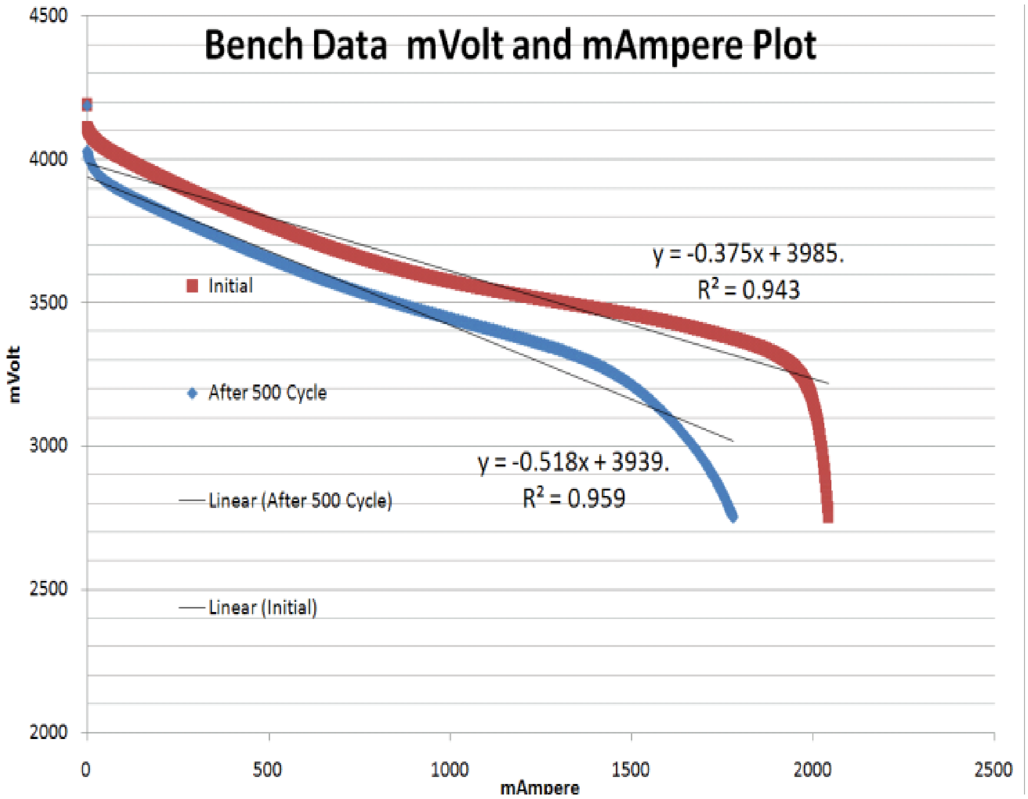
The window at 10 has the least distribution of S.

In future, system would receive real time data and the more accurate deterioration information will be available. Furthermore, as for future work, there are challenges that have to be overcome. In below, the challenges are discussed from two view points; visual data mining and Li+B deterioration analysis.

Having completed the data explorative knowledge finding exercise, as described in previous section, there are more or less 60 data points of which were extracted. From 60 data points, 25 are selected after excluding the time window exceeds 120 seconds. As shown in Figure 10, 1 to 25 data points are plotted chronologically. Slope(a) value fluctuates between 0.25 to 0.50. With its trend, it shows a slight yet gradual growth in value (y=0.001x).



[F29] Extracted Data Point by Visual Data Mining Application



[F30] Bench Data of Initial and After 500 cycle discharge

As per comparison, bench data from laboratory experiment is used. The bench data, as shown in Figure 11, is a comparison of initial battery state and the state after 500 charge-discharge cycle. 500 charge-discharge cycles is equivalent to 2 years of battery usage. The absolute difference in slope (a) between initial and after 500 cycles is 0.375 to 0.518. From the previous chapter, data points extracted by VDM exploration vary from 0.25 to 0.50. Although the two data sets do not make an exact match, it can be said that a series of slope (a) data follows similar trend of bench data characteristics therefore it may be used to assess the deterioration in future.

As for future work, the relation between extracted data window and other parameters such as SOC and temperature is expected to be clarified. Moreover, system performance improvement in order to endure longer period of data sample will also be greatly expected to be worked on future.

From visual data mining view point, automated algorithm of processing the large data will be critical for further enhancement. Given the circumstances, datasets are more static and limited although they are rather huge in volume and in complexity. In future, real time data analysis must be imperative. Thus, more efficient algorithm for data processing and more visually interactive data mining environment will be required so that “finding a needle from haystack” will be possible. Additionally, such an environment must be available not only for exclusive organizations. Its cost-effectiveness and commercial use-case must be studied for extensive opportunities for both business and academia.

From Li+B deterioration analysis view point, its mechanism must continue to be a target for many while its research and development progresses in every aspect of performance, quality, materials, cost-performance, and use-case. A specific issue related this research will be, a clarification on how regeneration and accumulation of usage effect on battery deterioration will be great advancement on enabling more accurate speculation.

5. Reference

- リチウム二次電池(小久見 善八,2008)
- Curbing Global Energy Demand Growth: The Energy Productivity Opportunity (McKinsey Global Institute 2007)
- Batteries for Electric Cars; Challenges, Opportunities, and the Outlook to 2020 (Boston Consulting Group 2010)
- 「クルマ依存」からの脱却 (肌附 安明 2010)
- Big Data: The Next Frontier for Innovation, Competition, and Productivity (McKinsey Global Institute, 2011)
- 技術力で勝る日本がなぜ事業で負けるのか (妹尾堅一郎 2009)
- THE MARKET FOR "LEMONS":QUALITY UNCERTAINTY AND THE MARKET MECHANISM (George A. Akerlof 1970)
- Visual Data Mining: Theory, Techniques and Tools for Visual Analytics (David Hutchinson and other, 2008)
- A Study on Visual Data Mining using Super High Definition Three-Dimensional Environment (So Sato, 2010)

- Finding a Needle in Haystack: Facebook's Photo Storage (Doug Beaver, Sanjeev Kumar, Harry C. Li, Jason Sobel, Peter Vajgel)

6. Acknowledgements

I would like to thank Professor Ogi and Assistant Professor Tateyama for two years of advisory on my master dissertation from their profound knowledge and experiences on visualization and computation. I'd like to thank Professor Sasaki, for technical advices on Lithium Ion Battery as well as being sub-supervisor on my thesis. I'd like to thank Proferssor Teshima, and Professor Toma for being sub-supervisors of my dissertation and for providing objective feedback to my research. I would like to thank my family, friends, co-workers at Hitachi Global Storage Technologies, and all of classmates for your kindness and encouragement.