basis. So it is important for human operator to understand lots of information about the physical system status provided and updated by the computer in a short period.

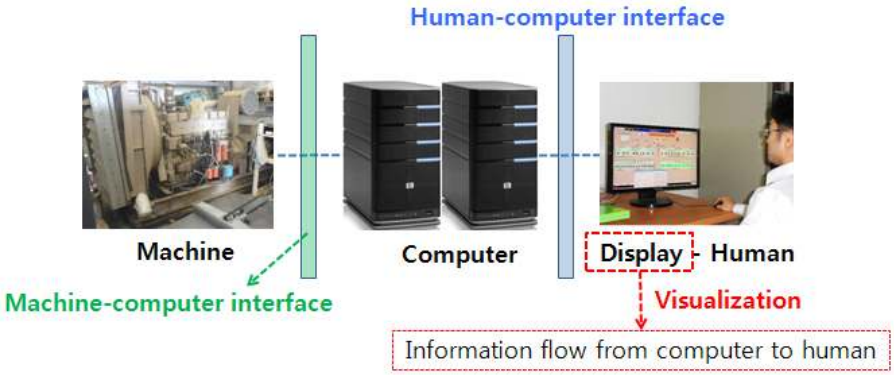

Fig.1 Visualization in Human-Computer Interface

A human has a tendency to feel more comfortable with images than text types on accommodating the information especially when it is complicated and large size. Therefore visualization of information could relieve the human's work load compared to the text mode information processing. This paper proposes a concept design on visualization of data and information in control systems on the aspect of functional and structural methods.

2. Visualization on SCADA system

2.1 SCADA system and HCI

The electric power system is the biggest artificial system on earth and also connected with many other systems because the physical infrastructures of the modern civilization have been built on electricity. The power system is operated by a certain type of control scheme. It started as manual operation of human operators and introduced several automatic processes since local power systems started to be connected with each other and thereby integrated into a bigger system which was difficult to control only by human operator.

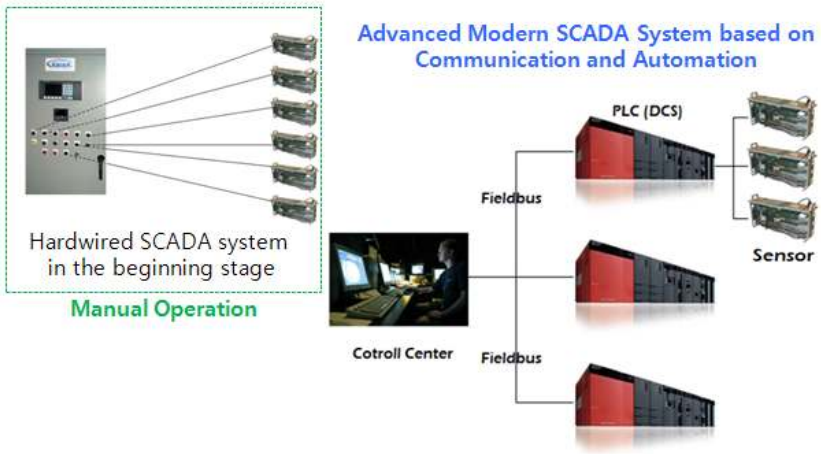

Fig.2 Evolution of SCADA System

As the integration continues with the help of automation and communication technologies the power system grows to a huge system covering the whole country. Therefore it is impossible to operate the system with manual operation. In addition, electricity is not stored, which requires huge size of data and information required to be

processed on real time basis and many components to be controlled interactively in the system. Therefore an automatic control scheme is not an option but an imperative on the current system status. As the information technology advances, the telephone was fully replaced by computer communication. Automation was accelerated with IT and the control system was able to be extended to wide area. It is called SCADA system which stands for supervisory control and data acquisition system. It can be understood as a huge scale of the control system covering wide area to the distance of hundreds of kilometers as shown in Fig.3. In Fig.3, ICCP, DNP, Harris, MODBUS, and TCP/IP are communication protocols which have different data frames and communication procedures. And Fig.4 shows the examples of control centers filled with a huge screen and many displays to provide the visualized information transformed from the mass size of text type data.

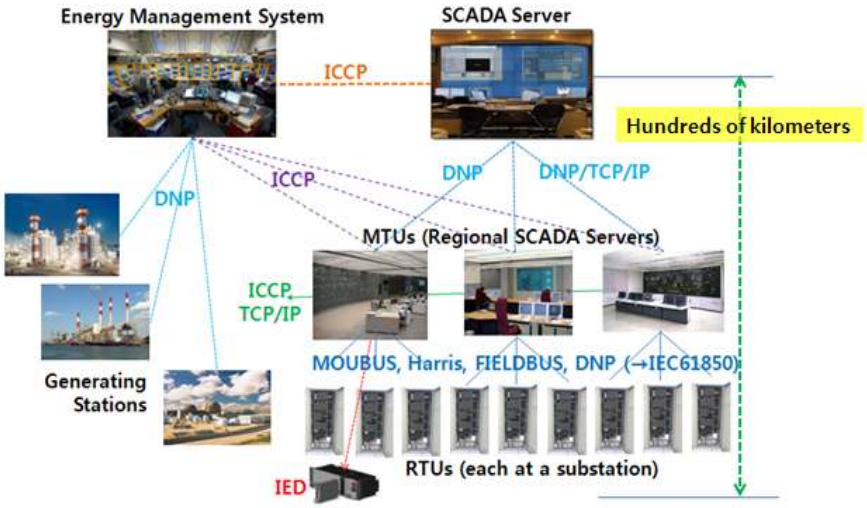

Fig.3 Wide-Area SCADA System

Fig.4 Examples of Monitoring Screen in Control Center

In spite of advanced automation there might happen to be several contingent situations caused by computer programming bugs or unexpected system situation, which still requires human operators to handle the contingencies not covered by the automatic process. In this context, HCI (human-computer interface) is very critical for operators to monitor and control the system. Moreover the mass data in the power system is updated every 2 seconds, which increase the quantity of information more rapidly during a certain time period. There is a

clear limitation on the information quantity which a human can recognize and process in a short time. So there exists a kind of contradiction between two factors, information quantity and human capability. It is expected that visualization takes a role of the interface between two different properties originated from human and machine respectively.

2.2 Data Visualization Process for Producing Information

Data visualization is the graphical representation of information. Bar chart, scatter graphs, and maps are example of simple data visualizations that have been used for decades. Information technology combines the principles of visualization with powerful applications and large data sets to create sophisticated images and animations [1]. Visualization is becoming an essential method to deal with large data sets. One representative characteristic in the SCADA system is to deal with the large size of data retrieved from many numbers of local data acquisition units on remote sites as shown in Fig.3. The SCADA system covers whole country, so the data overwhelm the human capacity on information process, and furthermore the data is updated every 2 seconds which increases the quantity of data exponentially over time horizon. Therefore it is impossible to follow the whole raw data in text forms. Instead it is more efficient to transform the data into the visualized form of information. The visualization does not just mean to make the image form from the raw data but to produce the meaningful information through the process as shown in Fig.5.

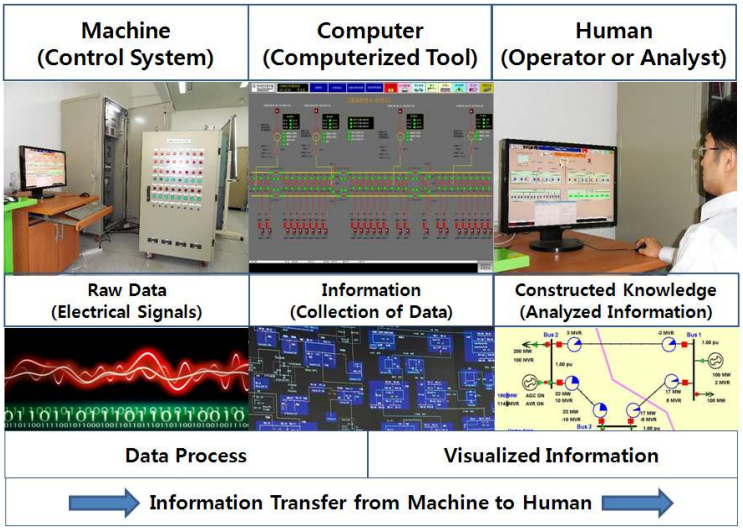

Fig.5 Visualization Process

The terms data and information were used differently in Fig.5. Then what is the difference between data and information? SMALE Consulting Ltd. defined the difference. The term information is defined as useful knowledge derived from the data while the term data is defined as raw, unanalyzed facts, figures and events. This means the information is produced by transforming the data through the process as follows [1]:

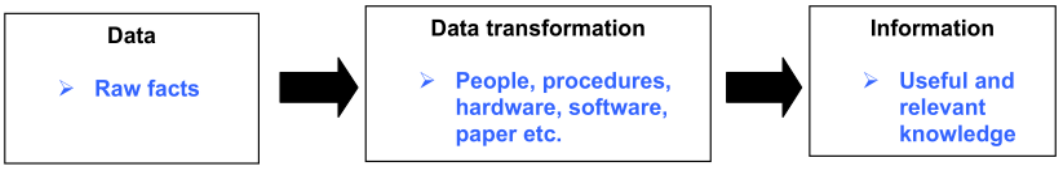

Fig.6 Data Transformation to Information

Considering the communication between RTU and SCADA Server in Fig.3, raw data is retrieved by RTUs from field devices and facilities. Through this process the electronic signal on physical level is transformed into the digital form of data. The data at each RTU is transmitted to the MTU, and then collected and structured in a relational database system at the MTU server in which the raw data is transformed into the form of information at first. The information is transformed into visualized forms for the operators in the MTU. The data and information at MTUs are again transmitted to the SCADA server and used for the overall operation of the entire system.

3. Concept Design for Visualization

3.1 Concept Design for Visualization Method

There are several stages on the flow of processing data and information from field devices to the SCADA server. The data is made to be segmented and related with each other in the MTU server in which the data takes a form of information at first. The data is filtered and organized as required for the regional operation and the transmission to the upper SCADA server. Considering the control system of MTUs there are two aspects of data usage. One is to use the data for the automatic operation process in the control system, and the other is to display the data with various visualized forms for human operators in the MTU. The hierarchical structure of data transformation process is shown in Fig.7 for the inner control process and the visualized display, which is an example of the segmentation of data properties according to functional purposes of the data usage.

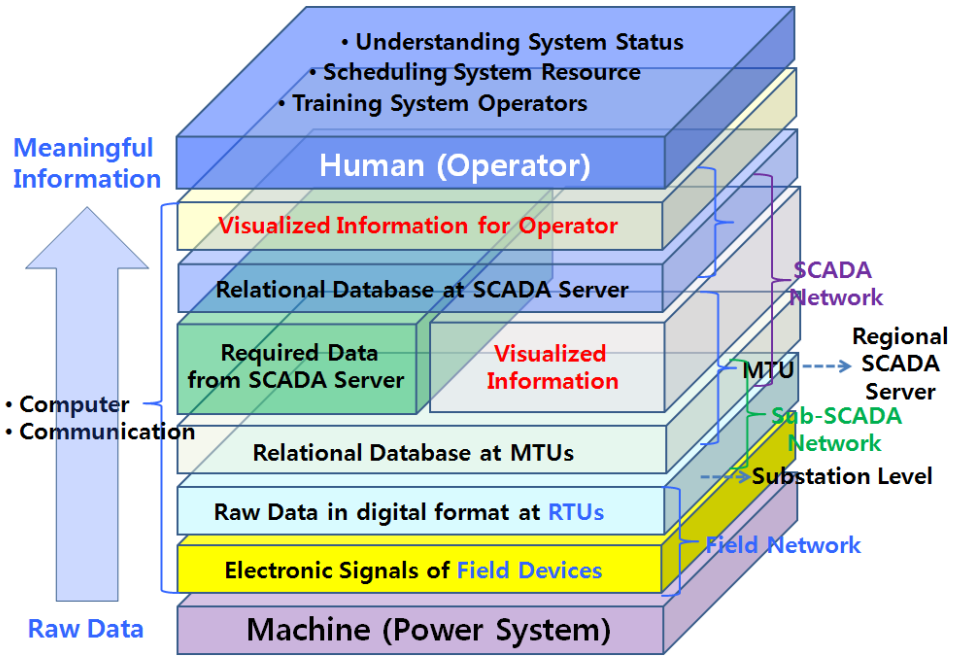

Fig. 7 Concept Design for Hierarchical Data Segmentation

The data could be more segmented by data type layers in that the raw data is transformed into the processed information at each level which defines system environments and control purposes as shown in Fig.7. Segmenting the data increases the flexibility of data usage and transformation by combining and organizing the segmented data sets in many ways. It is considered as information when a functional relationship could be

induced from two variables which are mapped into two segmented datasets as shown in Fig. 8. And this kind of formulation could be applied to multiple datasets more than two with various forms of mathematical relationship.

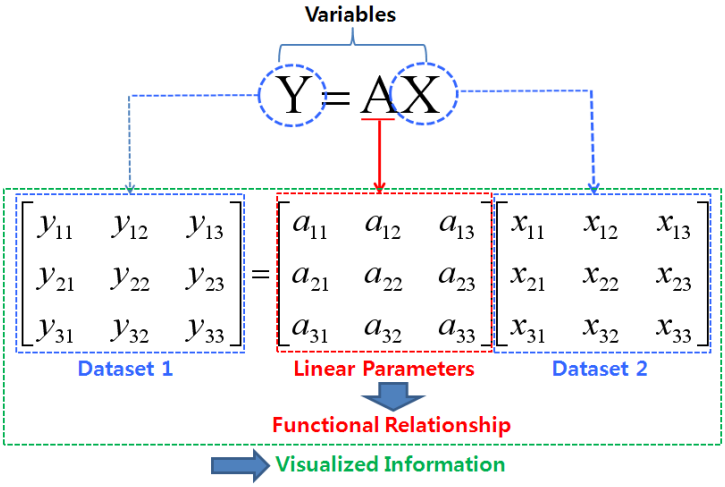

Fig. 8 Mathematical Formulation of Information with Datasets

Information is produced from datasets and higher level information is produced from multiple information sets at the current level. It is a kind of hierarchical structure from low level data to high level information as shown at y-axis in Fig. 9, and the higher level information is more required on human operation level at x-axis than automatic control processes on machine and computer level. The intersection area indicated by the dotted square is considered to be preferred as the visualization area for the human operator.

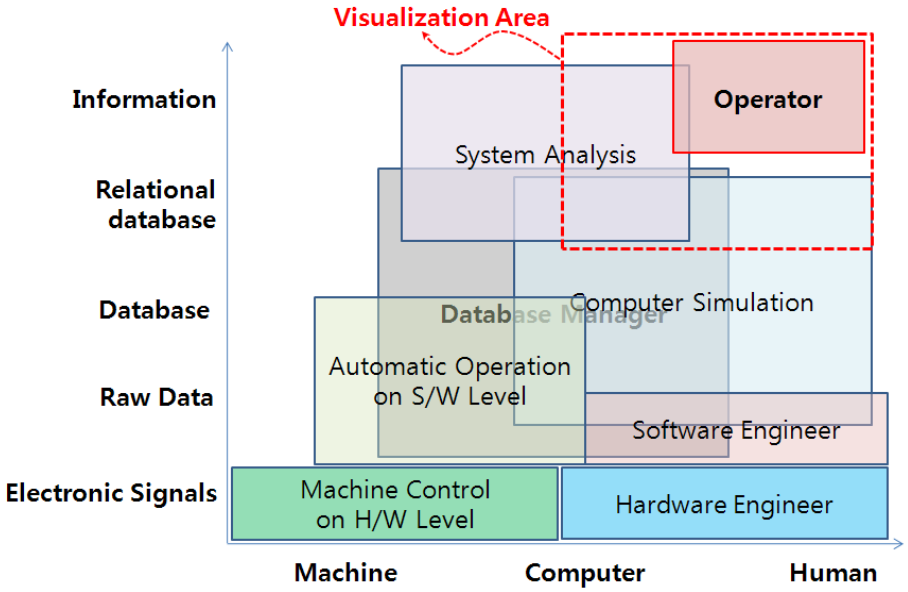

Fig. 9 Concept of Visualization Matrix

However it is possible to change the visualization area dependent on which information is needed and who wants the information as a visualized format. The power system is composed of several layers as shown in 'Multiple Layers of Power Industry[3]' of Fig. 10 and there are many system components and market participants in electricity market at each layer, so there are lots of entities who need various types of information both on quantitative and qualitative aspects. Fig. 10 illustrates the variable visualization area dependent on different layers of the system and information users.

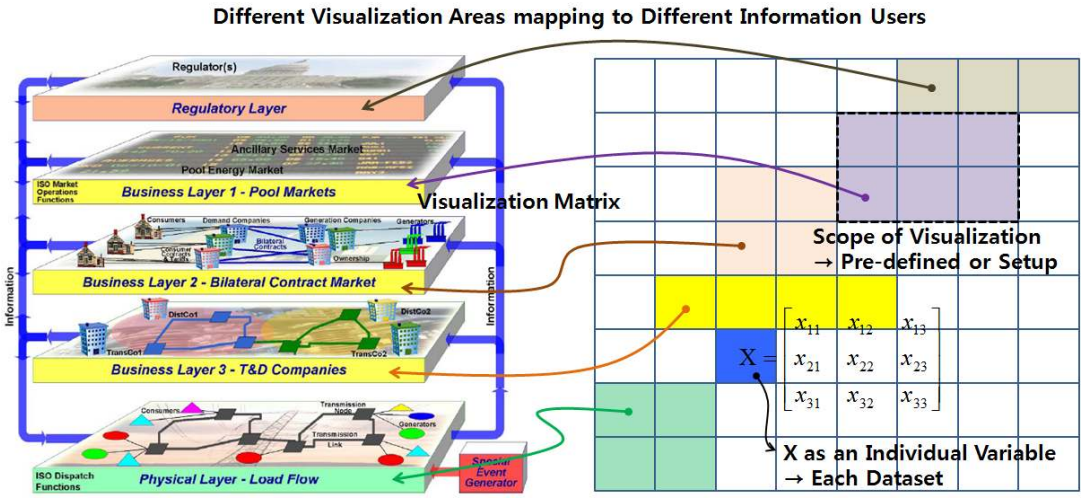

Fig. 10 An Example of Variable Visualization Area

3.2 Concept Design for Visualization Process with Computerized Tools

The data components in visualization matrix are determined the following procedure in Fig. 11. Measured data in the form of analog and digital signals are acquired and then transformed into digital bits by field devices and sensors. The data is transmitted to RTUs and transformed to the raw data format at software level. The data is stored again in the database and the relationship is built based on the function provided by the database program. When it is required to build more complicated functional relationships after data segmentation process in which a least unit of dataset is defined, the functions could be provided with other numerical analysis programs or statistical tools through an API (Application Program Interface) program.

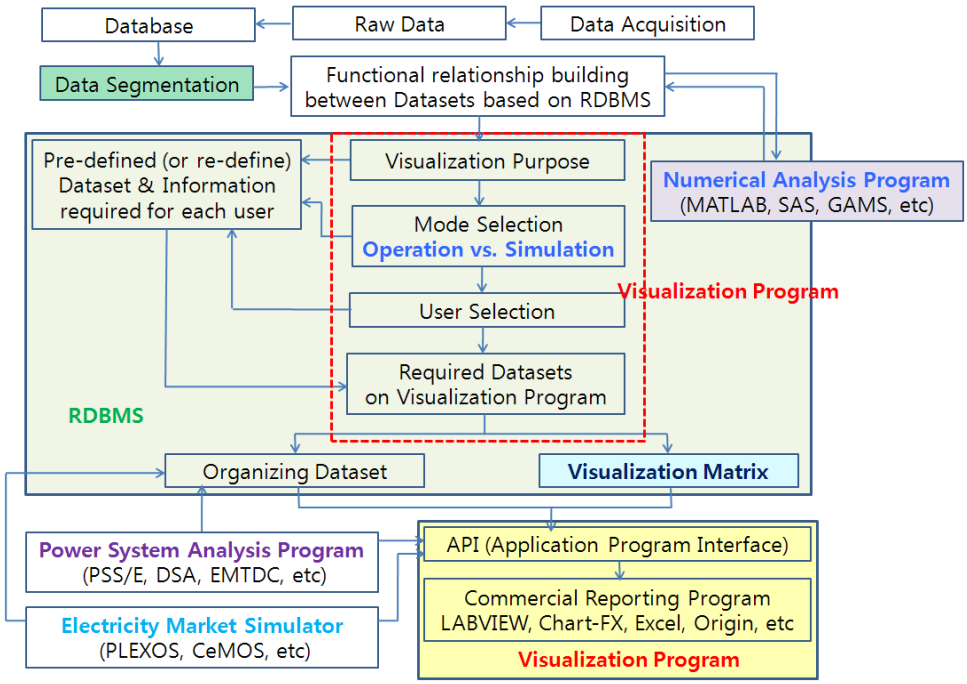

Fig. 11 Procedure for Establishing Visualization Matrix

After the relationship building process finished, several options are recommended to be selected to fix the required dataset for the visualization process such as the purpose, the user, and the mode of the visualization. For

example, in the power industry, the purpose could be understating system dynamics, market simulation, generation dispatch, etc. The user could be system operators, market analyst, system engineer, etc.

On the aspect of software engineering the relation between the data is built based on the RDBMS which stands for relational database management system. ORACLE and MS-ACCESS are commercial database programs widely used in many fields including the power industry. It is available to establish simple relations on RDBMS but impossible to make information based on complicated calculations, which requires another application program. The application program takes the role of making more complicated and functional relationship between datasets, or producing the high level information. The application program should be designed to interact with RDBMS and other visualization tools. The concept design of visualization program is illustrated in Fig. 12 as two aspects of horizontal and vertical conceptual structures.

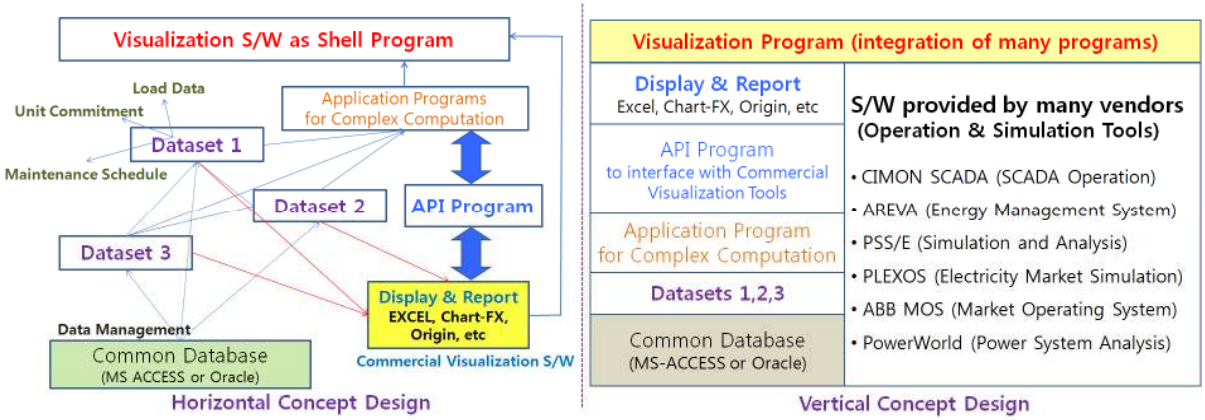

Fig. 12 Horizontal and Vertical Concept Design of Visualization Program

There are many commercial tools provided by vendors with different purposes and usages in the electric power industry. Some of them are the on-line operation tools monitoring and controlling the power system on real time basis while others are off-line simulation tools to analyze the system characteristics with many contingent cases. However the tools are not compatible with each other in spite of the fact that all those tools are needed to exchange the data and share the functions. On this context visualization program is expected to take a role of interfacing different programs with each other as a shell program or an API above all the programs as shown in Fig. 13.

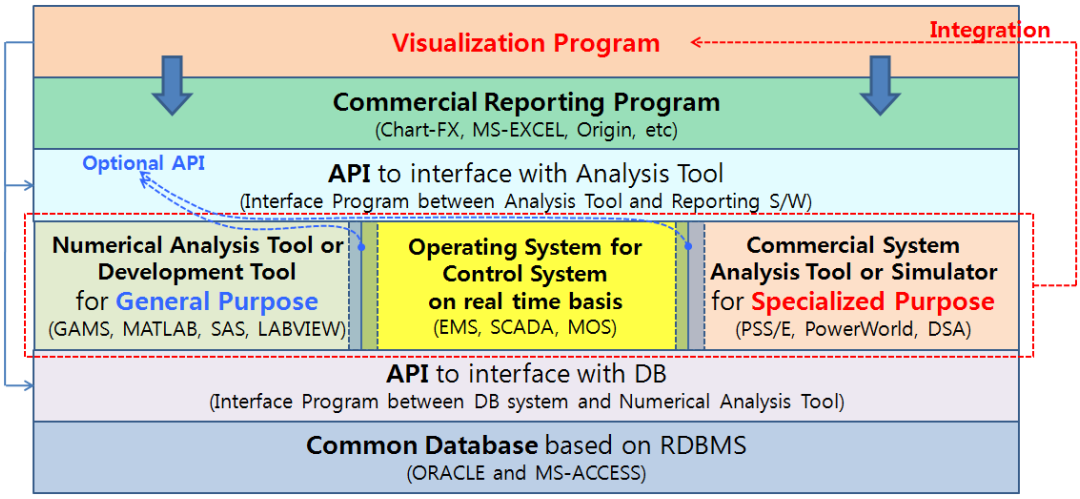

Fig. 13 Visualization Program Hierarchy with API

4. Application of Visualization Matrix to Power Industry

Everything is connected to everything else. It is called the first of ecology which is also applied to information technology. Quoting John Donne, no data is an island, entire of itself, every data is a piece of a data source, a part of a larger business reality. It is the underlying philosophy of the visualization program concept proposed in this paper. It is required to integrate many different functions and information provided by many applications into a business flow. The electric power industry works based on a huge power system with lots of physical facilities and market participants. There are many kinds of computerized tools to control, analyze and simulate the system and market. Although the tools provide various functions and produce lots of information, they are not compatible with each other, which decrease the efficiency of overall understanding at the level of decision making on the both aspects of physical system and business flow. If the tools could share the database and exchange the functions it would enhance the flexibility on the application. They share the relational database at the lowest level as a common input system, and the visualization program at the highest level as a common output system as illustrated in Fig.13.

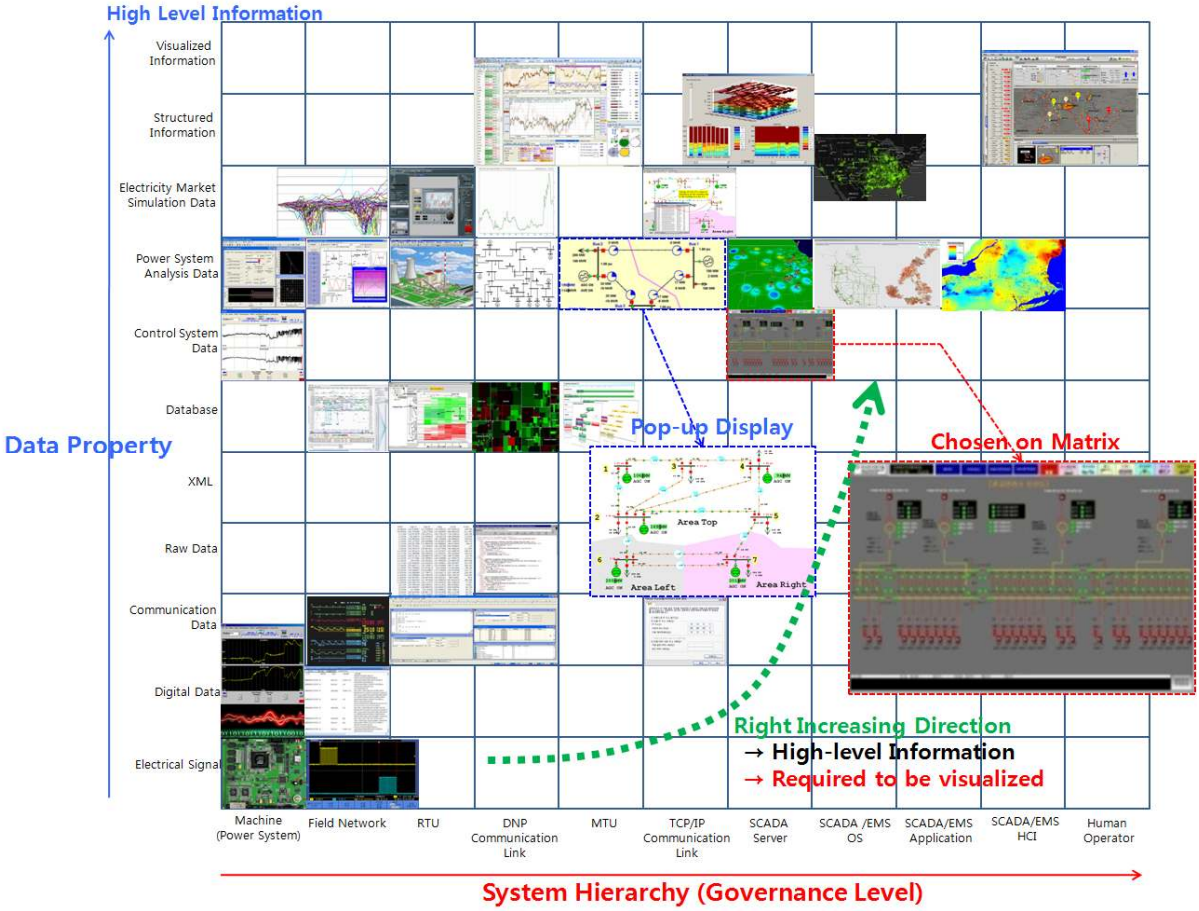

Fig. 14 Application Example of Visualization Matrix in Power Industry

Visualization matrix is a conceptual design of the visualization program as a common output system. When it is available to share the database in the same system it is possible to integrate the visualization process in a reporting program. Otherwise the data is conveyed to the visualization program through the API and then made to be output with a visualized format. In a sense the visualization program is a common API for many different software tools to interface with common database systems. On

visualization matrix each tool could identify its functional position on the entire system and determine what tools to interface with each other, which makes a conceptual outline of visualization program

5. Conclusions

The main goal of data visualization is to communicate information clearly and effectively through graphical means. To convey ideas effectively, both aesthetic form and functionality need to go hand in hand, providing insights into a rather sparse and complex data set by communicating its key-aspects in a more intuitive way [4]. Especially in the electric power system, there are lots of abstract forms of information like reliability, adequacy, security, etc., which are difficult to be described with numerical formats. In this case it is very useful to visualize the information more easy to understand for humans. And the main focus of this paper is to deal with mass size of data by visualization.

5. References

[1] EDUCAUSE Learning Initiative, 7 things you should know about Data Visualization [Online PDF]. Available at <http://net.educause.edu/ir/library/pdf/ELI7030.pdf> [Accessed 20 May 2009]

[2] Boddy, D and Paton, R (1998) *Management: An Introduction*, Prentice Hall

[3] Argonne National Laboratory, Electricity Market Complex Adaptive System (EMCAS) [Online PDF]. Available at <http://www.dis.anl.gov/pubs/61084.pdf> [Accessed 22 May 2009]

[4] OLAP (Online Analytical Processing), Business Analytics and Visualization [Online PDF]. Available at <www.business.mcmaster.ca/courses/k723/note/3.OLAP.ppt >

[5] Klemmer, R.S., Thomsen, M., Phelps-Goodman, E., Lee, R. and Landay, J.A. (2002) Where do web sites come from? Capturing and interacting with design history. In *Proceedings of CHI 2002*, ACM Press, pp 1-8.

[4] Schlosser, A. E. (2003) Experiencing Products in the Virtual World: The Role of Goal and Imager in Influencing Attitudes versus Purchase Intentions, *Journal of Consumer Research*, vol. 30, no. 2, pp 184-198.