

Geometry PIG 데이터 분석을 위한 통합 가시화 기법

김복동¹, 구상욱², 권혁돈³, 정승대⁴, 정순기⁵

경북대학교 컴퓨터공학과^{1,2,3,4,5}

{comthief¹, sokoo², hdkwon³, tinywolf⁴}@vr.knu.ac.kr, skjung@knu.ac.kr⁵

Integrated Visualization Techniques for Analyzing Geometry PIG Data

Bok Dong Kim¹, Sang Ok Koo², Hyok Don Kwon³,
Seong Dae Jung⁴, Soon Ki Jung⁵

Dept. of Computer Engineering, Kyungpook National University^{1,2,3,4,5}

요약

Geometry PIG (Pipeline Inspection Gauge)는 배관 내에 삽입되어 내부를 흐르는 매체에 의해서 추진되는 장치로서 배관의 기하학적 형상을 파악하기 위해 사용된다. Geometry PIG 는 여러 종류의 센서를 지니고 배관 내부를 주행 하면서 탑재된 저장장치에 빠른 샘플링 속도로 데이터를 저장하기 때문에 획득된 많은 양의 데이터를 분석하기 위한 가시화 기법이 필요하다. 본 논문에서는 데이터의 특성을 고려하여 다양한 가시화 기법들의 스키마를 정의하고, 이러한 가시화 기법들을 이용해 geometry PIG 데이터 분석을 위한 통합된 가시화 기법을 제안한다. 통합된 가시화 기법은 각 가시화 기법들을 사용자가 원하는 형태로 배치하며 사용자가 원하는 시점에서 데이터를 파악할 수 있도록 가시화 기법에 따른 동기화와 사용자 인터페이스를 지원한다.

Keyword : large data visualization, visualization technique, data analysis, geometry PIG

1. 서론

컴퓨터를 이용한 데이터의 분석 및 가시화 기술(visualization technique)은 정보처리 기술의 발달로 지형의 공간정보 파악을 위한 지리정보 시스템(GIS; Geographical Information Systems), 기상 예보를 위한 기후 데이터의 분석, 생물학·의학 분야 데이터 표현 등 많은 분야에서 사용되고 있다[1, 2]. Geometry PIG(Pipeline Inspection Gauge)는 배관 내에 삽입되어 내부를 흐르는 매체에 의해서 추진되는 장치로서 배관의 기하학적 형상을 파악하기 위해 사용된다. Geometry PIG 시스템을 이용해 검사할 수 있는 배관 내의 특징정보(feature)들은 weld, bend, dent, ovality 등이 있고, 각 특징정보들은 분석을 통하여 형태, 크기와 위치 등을 알아낼 수 있다[3].

많은 가시화 기법들은 대용량의 데이터를

효율적으로 분석하기 위해 개발 되었다[4]. 적절한 데이터의 가시화 기술의 이용은 사용자가 데이터의 분석과 조사(exploring)를 위해 보다 직관적으로 이해할 수 있게 한다. 그림 1 은 데이터 유형과 각각 다른 유형의 가시화 기법으로 적용한 사례를 보여준다. 본 논문에서는 geometry PIG 데이터 분석 시스템을 위한 통합된 가시화 기법을 제안한다. 효율적인 데이터 분석을 위하여 동기화를 지원하는 View 들과 각 View 적합한 사용자 인터페이스를 정의한다. PIG로부터 획득된 데이터는 하나의 View 를 통해서만 분석하기 어렵기 때문에 다양한 기법들을 통해 가시화 한다. 특히, 정확한 영역을 선택하기 위한 방법과 관심 영역을 선택적으로 확대하여 보기 위하여 fisheye view 가시화 전략으로 Cylinder View 를 적용하였고, Histogram View 를 지원하여 보다

정확한 데이터의 분석이 가능하도록 하였다.

본 논문의 구성은 다음과 같다. 2 장에서는 관련연구들에 대해 알아보고, 3 장에서는 시스템 전체의 구성을 알아보기 위해 geometry PIG 로부터 시스템에 대해 설명한다. 4 장에서는 geometry PIG 데이터 가시화에 적합한 다양한 가시화 기법들 정의하고, 5 장에서는 정의된 가시화 기법들을 기반으로 통합된 가시화 기법에 대해 설명한다. 6 장에서는 통합된 가시화 기법을 이용하여 특징정보의 유형에 따른 분석과정을 분석하고, 마지막 7 장에서 결론 및 향후 과제를 제시 한다.

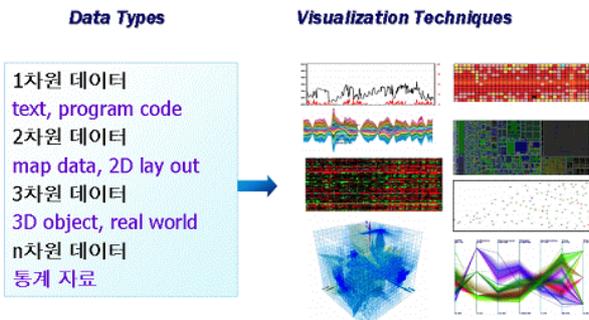


그림 1. 데이터 유형에 따른 가시화 기법들

2. 관련 연구

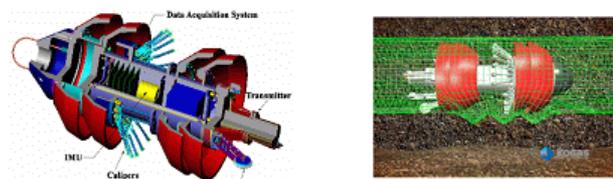
여러 가시화 기법들은 데이터의 유형에 따라 적절한 기법들이 필요하다. 데이터의 유형에는 크게 1 차원 데이터, 2 차원 데이터, 3 차원 데이터, 다차원 데이터, 네트워크 데이터, 트리 형태의 데이터, 계층적(hierarchical) 데이터 등으로 나누어진다[5]. 1 차원 데이터의 경우 텍스트 데이터, 프로그램 소스 코드, 그리고 시간에 따른 데이터 등이 있고 선형적인 특성을 갖는다. 2 차원 데이터의 경우 지형 데이터, 신문과 같은 페이지 배치 형태(layout)와 같이 평면적인 특성을 지닌다. 또한 인체와 같은 3 차원 객체를 표현하기 위한 3 차원 데이터와 통계 데이터베이스나 관계 데이터베이스와 같이 n 차원의 공간상에 표현 가능한 다차원 데이터들은 다양한 분야에서 사용된다[5, 6]. 각 데이터 타입에 따른 적절한 가시화 기법들은 데이터를 분석하고 정보를 추출하기 위하여 사용된다.

많은 양의 데이터를 표현할 경우 스크린

크기의 제한, 표현 색상의 제한 등의 근본적인 제약들로 인해 데이터 표현의 제약이 발생한다. 이러한 문제들을 해결하기 위해 기존 연구들에서 fisheye view 와 같이 view 의 관심 영역을 확대하는 interactive filtering 기법들이 적용되었다[7]. Uwe Rauschenbach[7]는 2 차원 래스터(raster) 영상을 모바일 장치와 같이 작은 화면상에 표현하기 위하여 Rectangular Fisheye View 를 제안하였다. 이와 같이 fisheye view 를 사용함으로써 화면 크기의 제한이 있는 경우에도 부분적인 확대를 통하여 사용자가 원하는 부분을 가시화 할 수 있다.

3. 시스템의 개요

가시화 기법들은 그 분야(domain)에 적합한 유형의 가시화 기법을 사용하기 때문에, 데이터의 특성을 이해하는 것은 매우 중요하다. 이장에서는 데이터의 특성을 이해하기 geometry PIG 의 데이터 획득 과정과 데이터베이스와 분석 시스템과의 관계를 설명한다. 그림 2(a)은 geometry PIG 의 구조를 나타낸다. Geometry PIG 는 finger, odometers, IMU, transmitter, 데이터 획득 장치로 구성되어있다. 데이터 획득 장치는 geometry PIG 가 배관을 주행하는 동안 각 센서들을 통해 측정되는 데이터를 순차적으로 기록한다. 이들 센서 중 finger 는 그림 2(b)에서와 같이 특징정보를 지낼 때 배관의 형상에 따라 다른 값을 획득하게 된다. 이렇게 획득된 데이터는 데이터베이스에 저장되고, 분석 시스템 사용해 특징정보들은 분석된다. 각 센서의 빠른 샘플링 속도에 따라 12 시간 주행 시에는 약 1.82Gbyte 정도의 대용량 데이터가 획득된다.



(a) Geometry PIG의 구조

(b) 배관의 Dent 구간

그림 2. Geometry PIG

분석가는 일반적으로 Line Graph 의 가시화 기법을 가장 많이 사용하여 분석 작업을 진행하며 특징정보의 유형에 따라 Pseudo Color Image, Polar View 등을 사용한다. Geometry PIG 데이터의 경우 weld, bend, tee, dent, valve 등 14 가지 유형의 특징정보를 가지고 있고, 분석가는 데이터 전 구간에 대해 순차적으로 이동하면서 배관내의 문제점과 정보를 분석한다.

4. 가시화 기법들의 정의

이 장에서는 본 논문의 통합된 가시화 기법에 적용되는 Line Graph, Pseudo Color Image, Polar View, 3D Surface, Histogram View 가시화 기법과 Cylinder View 가시화 적용 방법에 대해 설명한다.

4-1 Line Graph

Line Graph 가시화는 많은 응용 프로그램 및 데이터 가시화에서 사용되고 있는 2 차원 가시화 기법 중 하나이다. 2 차원 그래프 표현법은 평면상에서 가로축과 세로축을 갖는다. 본 논문에서는 1 차원으로 획득된 연속적인 데이터를 Line Graph 가시화를 적용하여 가로축을 거리 축, 세로축을 데이터의 값의 크기로 표현한다. 각 finger 들은 하나의 세로축을 사용 함으로서, 유사한 값을 갖는 경우 중첩(overlap)현상이 발생한다. 따라서 finger 들의 구분을 위하여 각 finger 들을 서로 다른 색상으로 구분하고, 사용자는 이들 finger 들의 색상과 라인의 크기, 세로축의 좌표의 시작 값과 끝 값을 임의로 조정할 수 있다. 그림 3 은 geometry PIG 데이터를 Line Graph 가시화로 나타낸 경우이다.

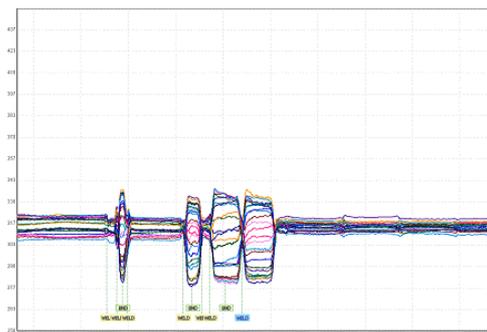


그림 3. Line Graph 가시화

4-2 Pseudo Color Image

Pseudo Color Image 가시화는 데이터를 2 차원 평면상의 이미지에 사상시키는 방법이다. 그림 4 는 세로축을 아래에서부터 1 번 finger 를 부여하여 순서대로 24 번 finger 까지 배치한 경우이다. 사용자는 임의의 Pseudo Color 를 지정하여 색상의 배치를 바꿀 수 있다. 또한 작은 값과 큰 값의 범위를 설정하여 분석에 적합한 색상의 분포를 설정할 수 있다.

그림 4 에서는 Gray 색상을 사상시킨 예이고, 흰색에 가까운 경우는 큰 값을 의미하고, 검은 색에 가까운 경우는 작은 값을 의미한다. 본 논문에서 구현된 Pseudo Color Image 가시화는 원의 형태로 배치된 finger 들을 고려하여, 롤링(rolling) 기능을 지원한다. 예를 들어, 사용자는 세로축 finger 들의 배치를 바꾸어 24 번과 1 번 finger 의 배치를 세로축의 중앙에 위치하도록 배치할 수 있다.

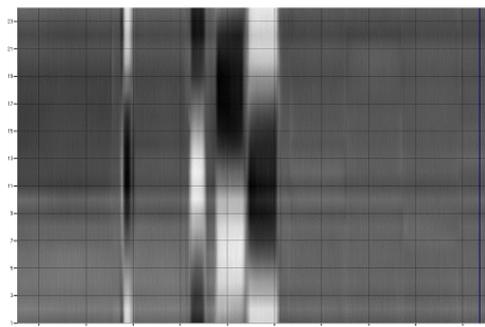


그림 4. Pseudo Color Image 가시화

4-3 3D Surface

스크린의 평면적인 특성 때문에 3 차원 가시화 기법은 값들이 겹쳐서(overplotting) 출력되어 값이 보이지 않거나 겹친 값이 변형된 색상으로 나타나기 때문에, 인터랙션(interaction)과 인지하는데 있어서 직관적이지 못하다는 단점을 지닌다. 본 논문에서 사용된 3D Surface 가시화는 가로축을 마찬가지로 동기화를 위한 거리 축을 사용하며, 세로축을 Pseudo Color Image 가시화와 마찬가지로 finger 의 개수만큼 분할하여 인접한 finger 들을 배치한다. 또한 좌표계의 z 축을 finger 값을 적용시켜 3 차원 평면으로 구성하고

투시투영(perspective projection)을 사용하여 3 차원 좌표계를 표현하였다. 그림 5 는 3D Surface 를 적용하여 가시화 경우이다.

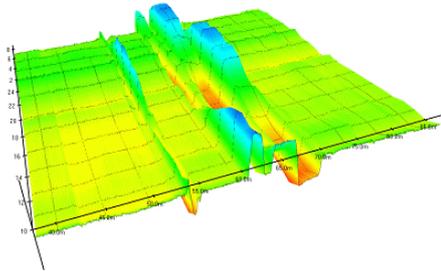


그림 5. 3D Surface 가시화

4-4 Polar View

Polar View 가시화는 배관의 단면을 보는 것과 같은 형태의 가시화 기법으로, 분석 시스템에서는 Line Graph, Pseudo Color Image, 3D Surface 에서 일어나는 사용자의 마우스 인터랙션에 따라 그 위치에서의 finger 값과 주변 finger 값들을 표현한다. Polar View 가시화는 finger 값을 원의 중심으로부터 떨어진 거리로 사상하여 배관의 단면을 보는 것과 유사한 형태의 효과를 지원한다. 그림 5 에서 보이는 중앙의 굵은 선은 현재 마우스의 위치에서 중앙을 기준으로 finger 값만큼 떨어지게 배치하여 finger 값들 전체가 상하좌우로 치우쳐져 있는 경우를 판단할 수 있다.

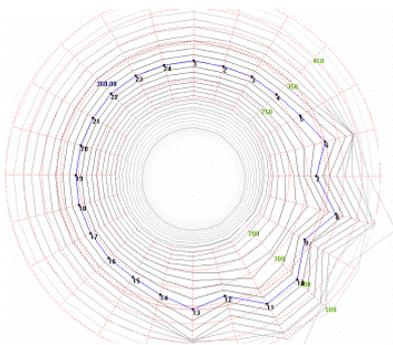


그림 6. Polar View 가시화

4-5 Histogram View

Finger 값은 특징정보를 지날 때 갑자기 변화하는 특성이 있다. 따라서 본 논문에서는 앞뒤 데이터의 차이 값으로 표현하는 Histogram View 를 사용하여 특징정보 분석의 효율성을

높인다. Histogram View 에서는 가로축을 거리 축을 사용하고 세로축을 거리축의 앞 데이터와의 차이 값을 모든 finger 에 대하여 누적 하여 표현 한다.

$$D_k = \sum_{i=0}^N \sqrt{(f_{(k+1,i)} - f_{(k,i)})^2} \quad (식 1)$$

식 1 에서 k 는 화면상에 표현되는 데이터의 개수, N 은 핑거의 개수, i 는 finger 번호를 의미한다. 따라서 i 번째 finger 의 차이 값은 현재(k)위치에서 다음위치($k+1$)의 차이 값의 절대치에 의해서 구해지고, Histogram View 의 생성은 식 1 과 같이 i 번째 finger 의 값을 누적함으로써 구해진다. 그림 7 에서 상단의 그림은 Line Graph 가시화를 나타내고 하단의 그림은 Histogram View 를 적용한 경우를 나타낸다.

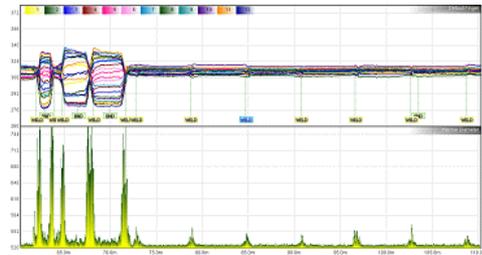


그림 7. Histogram View 가시화

4-6 Cylinder View 의 적용

인터랙션을 지원하는 많은 응용 프로그램 들에서 대용량 데이터의 가시화는 스크린 공간의 제한으로 문제점을 해결하기 위하여 fisheye view 전략들이 제안 되었다[7]. 본 논문에서는 분석 시스템의 가시화 기법 중 가장 많이 사용되는 Line Graph 에 fisheye view 전략을 적용하여 Cylinder View 를 구성하였다. Cylinder View 는 가로 축인 거리 축의 간격을 재배치하여 특정 거리 영역을 확대하는 기법으로 마치 거리 축을 실린더에 사상한 효과를 나타낸다. 그림 8 은 등 간격의 데이터를 재배치하는 과정을 보여준다. 분석가는 데이터를 분석하기 위하여 주요 feature 부분을 화면의 중앙에 위치시키는 경향이 있다. 본 논문에서 사용하는 Cylinder View 는 이런 분석가 행위의 특성을 고려하여 스크린의 중앙을 확대함으로써, 정확한 영역 선택과 시각적 효과를 지원한다.

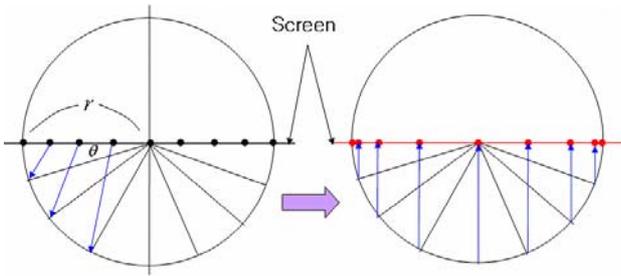


그림 8. 데이터 간격의 재배치 과정

스크린의 크기가 원의 지름과 같을 때, 그림 8에서 r 은 스크린 크기의 반을 의미한다. 이때, 재배치되는 데이터의 위치는 반원을 등 간격으로 나누는 후, 다시 평면으로 사영하여 구한다.

$$x_k = r - r \cos \theta, \quad (\text{식 2})$$

$$\theta = \frac{k}{N} \times 180. \quad (\text{식 3})$$

식 2에서 k 는 데이터의 개수를 의미하고, N 은 평거의 개수를 의미한다. 따라서, 스크린상의 위치 x_k 는 식 3에서 구한 반원을 등 간격으로 나눈 위치에서의 x 위치 값을 빼서 구해낼 수 있다.

5. 가시화 기법의 통합

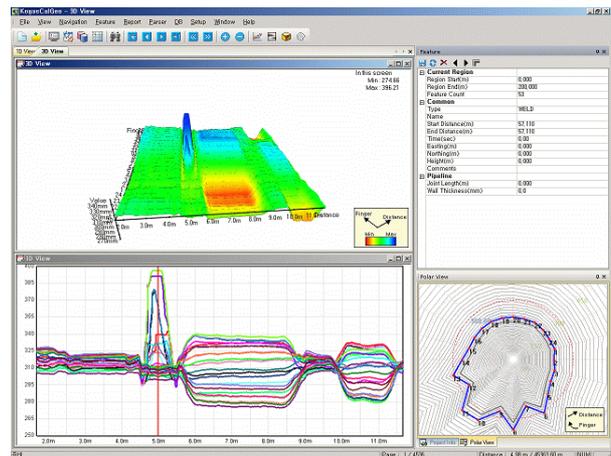
사용자는 통합된 가시화 기법을 이용하여 단일 view가 아닌 여러 view를 사용자가 원하는 형태로 배치하고, 각 view에 따른 적절한 사용자 인터페이스를 사용하여 사용자가 원하는 시점에서 정확한 데이터를 파악할 수 있다. 따라서 통합된 가시화 기법을 적용하기 위하여 각 view들의 특성을 파악하여 view들 간의 연관성을 고려한 동기화와 특징정보를 분석하기 유리한 적응성 있는 view의 배치(layout), 그리고 각 view에서 사용자 요구 사항에 따른 사용자 인터페이스를 지원한다. 가시화 기법들의 동기화를 통해 분석가는 동일 구간에 대해 하나의 특징 정보를 분석하는 과정에서 분석가가 원하는 view를 통해 효율적인 분석이 가능하다.

6. 구현 및 실험 결과

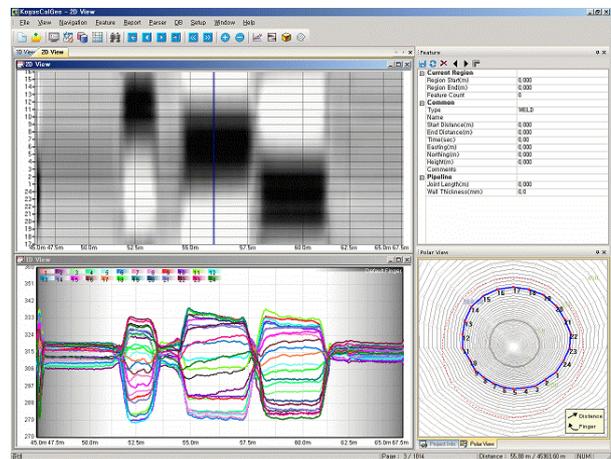
이 장에서는 각 가시화 기법들의 구현 결과와 통합된 가시화 기법의 효율성에 대해 분석한다. 본 논문의 실험에서는 24개의 finger를 가진

geometry PIG 데이터를 가시화 하였다. 분석 시스템의 실험 환경은 펜티엄 4 2.8GHz의 CPU, 512MB의 주 메모리, GeForce 4 MX440의 그래픽 카드를 사용하였다.

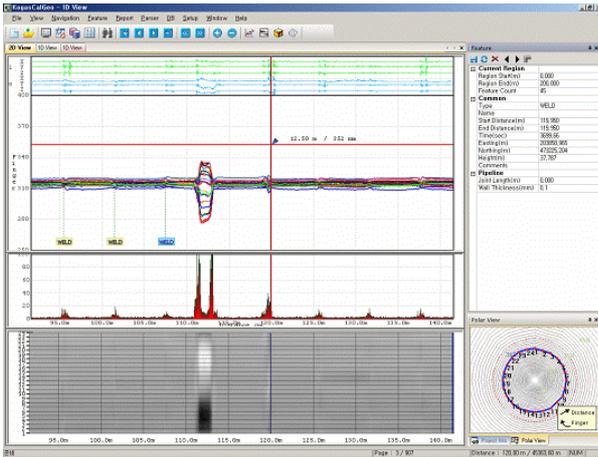
그림 9은 통합된 가시화 기법을 적용하여 다양한 특징정보를 분석하는 과정을 보여준다. 특징정보는 유형에 따라 적합한 형태의 view의 배치와 특징정보 분석모듈은 분석가가 원하는 형태의 데이터 분석이 가능하도록 한다. 그림 21(a)의 경우 배관의 한쪽 구간이 개폐된 형태인 Tee 정보를 3D Surface 가시화와 Polar View를 통하여 사용자가 원하는 정보를 쉽게 얻을 수 있다. 그림 21(b)의 view 배치는 bend 구간의 중심 부분이 어떤 방향으로 치우쳐 있는지를 파악하기 쉽고, 그림 21(c)는 Histogram View와 동기화된 Line Graph가 정확한 특징정보를 입력 가능하다는 것을 보여준다.



(b) Tee 특징정보의 분석



(c) Bend 특징정보의 분석



(d) Weld 특징정보의 분석
그림 9. 통합 가시화 기법

7. 결론 및 향후 과제

본 논문에서는 데이터의 특성을 고려한 다양한 가시화 기법들의 좌표계를 정의하고, 이러한 가시화 기법들을 이용해 geometry PIG 데이터 분석을 위한 통합된 가시화 기법을 구현하였다. 가시화 기법 분석에 필요한 가시화 기법의 좌표계 정의 및 구현을 통해 분석 시스템에 적용함으로써, 분석가는 효율적인 데이터 분석이 가능하였다. 통합된 가시화 기법은 다중 윈도우를 통한 데이터 분석이 가능하고, 사용자 view 배치와 각 view 들의 실시간 동기화를 통해 분석가가 원하는 형태의 데이터 분석이 가능하다. 특히 Histogram View 는 정확한 특징정보 위치의 선택이 가능함을 보였고, 2 차원 가시화 기법들이 적용된 view 들은 Cylinder View 를 이용한 사용자의 요구 사항에 따른 확대 기법을 적용하여 관심영역을 확대하는 효과를 줄 수 있었다.

향후 본 논문의 연구를 기반으로 MFL (Magnetic Flux Leakage)을 비롯한 다른 PIG 시스템의 데이터 분석에 적합한 가시화 기법들에 관한 연구가 필요하다.

감사의 글

본 논문은 한국가스공사 연구개발원의 연구비 지원인 용역 결과물임.

참고 문헌

- [1] Milton H. Shimabukuro, Edilson F. Flores, Maria Cristina F. de Oliveira, and Haim Levkowitz, "Coordinated Views to Assist Exploration of Spatio-Temporal Data: A Case Study," 2nd International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV'04), pages 107-117, 2004.
- [2] Purvi Saraiya, Chris North and Karen Duca, "An Evaluation of Microarray Visualization Tools for Biological Insight," 10th IEEE Symposium on Information Visualization (INFOVIS 2004), pages 1-8, October 10-12, 2004.
- [3] D.K. Kim, S.H. Cho, S.S. Cho, H.R. Yoo, Y.W. Rho and Y.T. Kho, "Development of the Caliper System for a Geometry PIG Based on Magnetic Field Analysis," KSME International Journal, Vol. 17, No. 12: pp. 1835-1843, 2003.
- [4] Theresa-Marie Rhyne, Melanie Tory, Tamara Munzner, Matt Ward, Chris Johnson, David H. Laidlaw, "Information and Scientific Visualization: Separate but Equal or Happy Together at Last," IEEE Visualization 2003 Conference (VIS 2003), pages 611-614, October 19-24, 2003.
- [5] Ben Shneiderman, "The eyes have it: a task by data type taxonomy for information visualizations," Proceedings of the 1996 IEEE Symposium on Visual Languages, pp. 336-343, September 3-6, 1996.
- [6] Susan Havre, Elizabeth Hetzler, Paul Whitney and Lucy Nowell, "ThemeRiver: Visualizing Thematic Changes in Large Document Collection," IEEE Transaction on Visualization and Computer Graphics, Vol. 8, No. 1. pp. 9-20, January-March 2002.
- [7] Carl Gutwin and Chris Fedak, "Interacting with Big Interfaces on Small Screens: a Comparison of Fisheye, Zoom, and Panning Techniques," Proceedings of the 2004 Conference on Graphics Interface, Vol. 62, pp. 145-152, 2004.