

A Similarity Wave View : 이미지 질의 결과의 시각화 방안

진 병 정^{*}, 엄 기 현
동국대학교 대학원 컴퓨터공학과

A Similarity Wave View : An Visualization Mechanism of Image Query Result

Byungjung Jin^{*}, Kyhyun Um
Dept. of Computer Engineering, Graduate school, Dongguk University
(jinbyju, khum)@cakra.dongguk.ac.kr

요약

기존 이미지 데이터베이스 검색 시스템은 질의 결과로 유사도가 높은 이미지들을 모아 제시해 준다. 그런데, 질의 결과로 제시되는 이미지는 전부 같은 크기의 이미지로 표시되고 일차원적으로 단순 나열하기 때문에 결과 이미지의 의미파악이나 질의 처리 과정의 정보를 확인하는 것이 명확하지 않다. 또한 하나의 화면에 나타낼 수 있는 이미지의 수도 매우 제한된다. 따라서 본 논문은 이미지가 표현되는 화면 공간을 5개의 Wave로 나누고, 유사도 순서로 영역 크기를 할당하여 질의 결과 이미지들을 이차원으로 배열하여 제시하고 관리하는 방안을 제안하는 것이다. 그래서 한정된 화면 공간의 활용도를 극대화하고, 질의 결과에 대한 사용자 인지도를 높이며, 질의 결과의 유사도 정보를 유추하거나 분석하기가 용이한 방안이 될 것이다.

1. 서론

사용자 인터페이스 중 직접조작(Direct manipulation) 개념이 성공함에 따라 시각화되고 그래픽 방법이 컴퓨터와의 주요 대화 수단으로 사용되고 있다. 컴퓨터 하드웨어의 발전으로 정보의 시각화와 그래픽 사용자 인터페이스의 역할은 그만큼 확대되고 있다. 정보 시각화는 대용량의 정보를 사용자에게 효율적으로 제시하는 것을 목표로 한다. 정보 시각화는 정보를 화면에 효율적으로 제시하여 사용자에게 정보에 대한 이해를 높이고, 사용자가 원하는 정보를 쉽게 찾도록 하자는 것이다[1].

이미지 데이터베이스 검색 시스템을 구축하기 위해서는 사용자가 이해하고 이용하기 쉬운 사용자 인터페이스 개발 기술이 필요하다. 특히 사용자가 원하는 질의를 생성하고 질의 결과를 이미지의 특성에 맞고 사용자가 이해하기 쉽게 제시하는 기술이 아직 미흡한 실정이다. 즉 기존 이미지 데이터베이스 검색 시스템은 질의 결과에 대해 단순히 순위를 부여해서 일차원으로 나열하고 있어, 질의 결과에 맞는 시각화가 부족하다.

이미지 데이터베이스 검색시스템의 질의 결과는 질의 이미지와 데이터베이스내의 이미지 사이의 유사도 값에 따라 순서화된 집합이다. 따라서 본 논문은 내용 기반 이미지 검색 시스템의 질의 결과 이미지를 시각화 대상으로 하며, 사용자 질의 결과로 나온 이미지를 유사도 순서대로 이미지가 차지하는 공간을 할당해 화면에 보여 줌으로써 화면 공간 활용도와 사용자 인지도를 높이고, 결과에서 이미지의 기본 정보에 따라 결과 집합을 재배치할 수 있는 관리 방안을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구를 통해 기존 정보 시각화의 특징과 문제점을 분석한다. 3장에서는 질의 결과 이미지를 시각화하기 위한 기법을 제안한다. 4장에서는 제안한 방법에 대한 성능을 기존 시스템과 비교 분석하고, 5장에서 결론을 맺는다.

2. 관련연구

기존 정보 시각화 시스템과 그들이 다루는 데이터를 유형별로 분류해 보면 크게 선형(linear), 2차원 맵(2D map), 3차원(3D), 시간(temporal) 속성, 다차원, 트리, 네트워크 등의 데이터를 들 수 있다[1]. 선형 데이터는 텍스트 문서, 프로그램 원시 코드 등으로 순차적인 형태이다. SeeSoft[2]는 수천 라인의 프로그램 원시코드를 한 화면에 보여주도록 화면 구성 기법을 제시하였다. 대용량의 텍스트 문서를 시각화한 Document Lens[3]는 하나의 화면에서 문서의 전체 내용과 문서 한 페이지의 상세한 내용을 동시에 표현하고 있다. 이는 문서를 화면에 표현하기 위해 기하학적 변형을 해야 한다. 2차원 맵 데이터는 지도나 평면도 등의 형태로 지리 정보 시스템이 대표적인 예가 된다. Themescape[4]는 문서의 집합을 지도상에 위치적인 정보로 표시하였다. 3차원 형태(3D)는 고해상도로 인체의 이미지를 향해하려는 Visible Human Explorer[5]가 대표적이다. 3차원 시각화는 방향 상실(disorientation), 겹침(occlusion)의 문제가 발생하며, 이를 해결하기 위한 연구가 계속 진행중인 분야이다.

시간 속성 데이터는 시작 시간과 끝 시간이 존재하고 아 이템들이 겹칠 수 있는 형태이다. Perspective wall[6]이 대표적인 시스템이다. 이는 프로젝트 관리를 시간순서(Time line)에 따라 표현하고 있으며 디스플레이 공간을 수평적으

로 변형하여 처리하고 있다. LifeLines[7]는 환자 기록을 시간순서에 따라 요약한 정보를 화면에 제시한다. 다차원 데이터를 사용하는 시스템들은 데이터베이스 시각화에 관련된 VisDB[8], 스프레드시트와 비슷한 형태의 Table Lens[9]가 있다. 트리 형태의 데이터를 사용하는 시스템들은 2D Space filling 알고리즘에 의해 트리를 사각형으로 표현하는 트리맵[10]과 3차원 Cone and Cam Tree[11], TreeBrowser[12], Hyperbolic Trees Browser[13] 등이 있으며 노드-링크 다이어그램으로 커다란 트리 구조를 표현하고 있다. 네트워크 데이터는 항목들 사이의 관계가 때때로 트리 구조로 쉽게 표현되지 못하는 경우가 있다. 이는 임의의 항목이 여러 항목과 연결되기 때문에 사용자 작업이 복잡하고, 시스템 구현이 어렵다.

지금까지 살펴본 정보 시각화 시스템의 특징은 제한된 화면을 효율적으로 이용하고, 시각화 대상이 되는 데이터를 사용자에게 이해하기 쉽게 보이는 것이 최종 목표라는 면에서는 동일하다.

대표적인 내용기반 이미지 검색 시스템에는 QBIC[14], WebSeek[15], NETRA[16], C-BIRD[17] 등이 있다. 이 시스템들은 이미지의 시각적 특징을 추출하고 인덱스를 구성하는 방법과 질의를 주는 방법이 서로 다르다. 그러나 질의 결과를 제시하는 방법은 공통적으로 유사도 순위로 단순히 선행적으로 나열하고 있다. 하나의 화면에 질의 결과로 제시하는 이미지의 수는 QBIC이 12개, WebSeek이 15개, NETRA가 6개, C-BIRD가 2~48개이다. C-BIRD는 최대 48개 이미지를 표현하지만 이미지의 크기가 작아서 사용자 인지도가 매우 낮다. 이로 인하여 이미지를 확대화면으로 보아야 한다는 단점이 있다. QBIC, WebSeek, C-BIRD는 질의 결과 제시 화면에 있는 어떤 이미지를 선택하면 새로운 화면으로 이미지의 상세 정보를 제시하고 있다. 이는 사용자 작업을 두 개의 화면에서 반복하게 함으로써 작업의 연속성이 부족하고 능률이 저하된다.

3. 질의 결과 이미지의 배치와 관리

본 논문에서 제안하는 Similarity Wave View(SWV)에서 Similarity는 사용자 질의와 결과 이미지를 사이의 유사도를 의미하고, Wave는 결과 이미지의 유사도에 따라 할당된 화면 공간을 뜻한다. SWV에서 이미지를 화면에 배치하는 기준으로 사용자 질의 이미지에 대한 결과 이미지의 유사도 값이 된다. 즉 유사도 값이 높은 이미지를 영역이 큰 Wave 공간에 할당하는 것이다.

3.1 이미지 배치 방법

본 논문에서 제시하는 SWV는 한 화면의 공간을 5개의 Wave로 구성한다. 각 Wave에 디스플레이될 수 있는 이미지의 수를 식으로 표현하면 [식1]과 같다.

$$W_n = 4n - 3 \quad (1 \leq n \leq 5) \dots\dots\dots [식1]$$

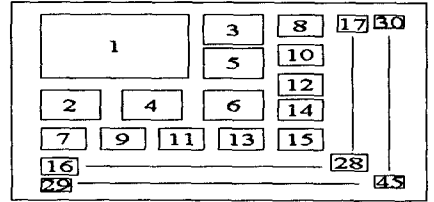
첫째 Wave(W_1) 값은 1이 된다. 질의 이미지와 유사도가 가장 높은 한 개의 이미지가 배치된다. $W_2 = 5$, $W_3 = 9$, $W_4 = 13$, $W_5 = 17$ 이 된다. 결국 하나의 화면에 나타낼 수 있는 최대 이미지 수(M)는 45개이다. 사용자가 인지 가능하다고 생각하는 최소의 이미지 크기를 고려할 때 가장 적당하다. 물론 결과 이미지를 더 많이 제시할 수도 있지만, 본 논문에서는 질의 결과 이미지의 전체 개요를 제공하는 것이

목적은 아니다.

각 Wave의 이미지 크기는 W_1 의 이미지 크기(Sw_1)를 K pixel로 정의하면, 식[2]에 의해 구할 수 있다.

$$\begin{aligned} Sw_1 &= K \\ Sw_2 &= (Sw_1 + Sw_1 / 2) / 3 \\ Sw_3 &= Sw_2 * 3 / 4 \\ Sw_4 &= Sw_3 * 5 / 6 \\ Sw_5 &= Sw_4 * 7 / 8 \dots\dots\dots [식2] \end{aligned}$$

그림[1]은 결과 이미지를 유사도 순위에 따라 화면에 배치하는 순서를 보인 것이다.

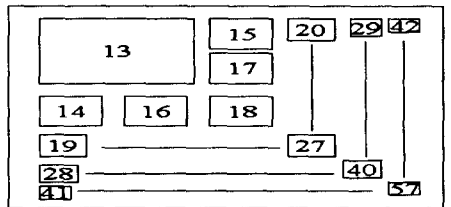


[그림1] 유사도 순위에 따른 이미지 배치

3.2 결과 이미지 관리 방법

사용자 질의에 대한 결과 이미지의 갯수(n)가 화면에 나타낼 수 있는 최대 이미지 갯수(M)보다 크면 W_2 에서 W_5 사이의 이미지를 선택하면 다음 결과 이미지를 볼 수 있다. n과 M이 같다면 그림[1]과 같이 배치된다. n이 M보다 작다면 사용되지 않는 Wave의 이미지 영역은 비어 있게 된다.

그림[1]에서 사용자가 만약 W_2 의 13번째 이미지를 선택했다면 해당 이미지는 W_1 으로 이동하면서 이동 결과는 그림[2]와 같다.



[그림2] 사용자 선택에 따른 이미지 재배치

또한 사용자는 이미지 기본 정보인 색상, 질감, 위치정보, 모자이크로 결과 이미지를 재배치할 수 있도록 했다.

3.3 질의 결과 인터페이스

그림[3]은 본 논문에서 제시한 내용 기반 이미지 검색 질의 결과 인터페이스 화면이다. 화면의 왼쪽의 상단에서부터 오른쪽 하단까지 5개의 Wave를 이루는 것을 볼 수 있다. W_1 은 질의 요구사항과 유사도 결과가 가장 높은 이미지가 위치하게 된다. 4개의 Wave에 있는 이미지 중 하나를 사용자가 선택(click)하면 W_1 로 이동해서 좀더 상세한 화면으로 결과 이미지를 볼 수 있게 된다.



[그림3] 질의 결과 인터페이스 화면

사용자가 최초 질의에 대한 결과 이미지에 대해 색상, 질감, 위치정보, 모자익을 조합한 유사도 값으로 질의 결과를 재배치할 수 있다. 여기서 색상 체크 박스를 선택하고 질의 재배치 버튼을 누르면, 질의 이미지에 대한 결과 이미지를 색상 유사도 순위에 의해 다시 배치하게 된다. 체크 박스의 이미지 기본 정보를 선택하지 않은 상태에서 질의 재배치 버튼을 누르면, 질의 결과는 최초 질의 인터페이스에서 질의한 결과로 재배치된다.

4. 성능 평가

아래의 표[1]은 결과 이미지의 수, 이미지 확대 정보 제시 방법, 화면에 제시하는 이미지의 크기를 기준으로 기존 이미지 데이터베이스 검색 시스템의 질의 결과 제시 인터페이스와 본 논문에서 제안한 SWV를 비교 분석했다.

시스템	QJIC	WebSnek	NETRA	C-BIRD	SWV
결과 이미지 수	12	15	6	2-48	45
이미지 확대 정보	분리된 화면에 제시	분리된 화면에 제시	없음	분리된 화면에 제시	하나의 화면에 제시
화면에 보인 이미지 크기	고정	고정	고정	고정	유사도에 따라 가변

표[1] 성능 비교 분석표

본 논문에서 제시한 방법은 기존의 이미지 데이터베이스 검색 시스템의 질의 결과 인터페이스보다 이미지를 유사도 값에 따라 화면에 보이는 크기를 다르게 하고 하나의 화면에서 작은 이미지를 확대하여 볼 수 있게 함으로써 사용자 인지도를 높였다. 또한 영역의 크기를 조절하여 질의 결과 이미지가 화면에 많이 제시될 수 있도록 하여 화면 공간 활용도를 높였다.

5. 결론

본 논문에서는 내용 기반 이미지 데이터베이스의 질의 결과를 시각화하기 위해 SWV를 제안하였다. SWV는 화면

공간을 5개의 Wave로 나누어 질의 결과 이미지를 유사도 순위에 따라 배치하였으며, 이미지 기본 정보인 색상, 질감, 위치, 모자익 정보를 이용하여 질의 결과를 사용자 의도에 따라 재배치하도록 하였다.

본 논문에서 제시한 SWV는 질의 이미지와 유사도가 높은 결과 이미지를 크게 보이도록 함으로써 사용자 인지도를 높였으며, 사용자가 이미지에 관한 정보를 분석하는 편의성을 제공했다. 또한 유사도가 낮은 결과 이미지는 Wave 영역의 크기를 작게 하여 공간 활용도를 향상시켰다.

6. 참고문헌

- [1] Ben Shneiderman, "Designing the User Interface", Third edition, ADDISON-WESLEY 1998
- [2] Eick, Stephen G., Steffen, Joseph L. and Sumner, Jr. Eric E., SeeSoft: A tool for visualizing line-oriented software statistics, IEEE Transactions Software Engineering, 18, 11(1992) 957-968
- [3] Robertson, G. G. and J. D. Mackinlay (1993). The Document Lens. Proceedings of the ACM Conference on User Interface Software and Technology (UIST '93), Atlanta, Georgia, November, ACM: 101-108.
- [4] Wise, James A. Visualizing the non-visual : Spatial analysis and interaction with information from text documents, Proc. IEEE information visualization '95, 51-58
- [5] <http://www.cs.umd.edu/hcil/visible-human/vhe.shtml> by Chris North, 1996
- [6] Mackinlay, J. D., G. G. Robertson, and S. K. Card. (1991). The perspective wall: Detail and context smoothly integrated. ACM Conference on Human Factors in Computing Systems (CHI '91), ACM: 173-179.
- [7] Brett Milash, Catherine Plaisant, and Anne Rose: "LifeLines visualizing personal histories" Proceedings of the CHI '96 conference companion on Human factors in computing systems: common ground, 1996, Page 392
- [8] Keim, D.A. and Kriegel, H., VisDB: Database exploration using Multidimensional visualization, IEEE Computer Graphics and Applications (September 1994), 40-49
- [9] Rao, R. and S. Card (1994). The Table Lens: Merging graphical and symbolic representations in an interactive focus + context visualization for tabular information, ACM Conference on Human Factors in Software (CHI '94), Boston, Massachusetts, ACM: 111-117.
- [10] Ben Shneiderman, Tree Visualization with Tree-maps: A 2-d Space-filling Approach. ACM Transaction on Graphics (11)1 (January 1992), 92-99.
- [11] Robertson, G. G., J. D. Mackinlay, and S. K. Card. (1991). Cone trees: Animated 3D visualizations of hierarchical information. ACM Conference on Human Factors in Computing Systems (CHI '91), ACM: 189-194.
- [12] Kumar, Browsing hierarchical data with multi-level dynamic queries and pruning, International Journal of Human-Computer Studies, 46, 1(January 1997), 103-124
- [13] Lamping, John : A focus + context technique based on hyperbolic geometry for visualizing large hierarchies, Proc. of CHI 95 Conference: Human Factors in computing Systems, ACM, New York(1995) 401-408
- [14] <http://www.qbic.almaden.ibm.com/~qbic/>
- [15] <http://www.ctr.columbia.edu/webseek/>
- [16] <http://maya.ece.ucsb.edu/Netra/>
- [17] <http://jupiter.cs.sfu.ca/Cbird/CBIRD.cgi>