

지능형 화상 검색 시스템에서의 사용자 모델을 이용한 사용자 적응

김 용 환[†] · 이 필 규^{††}

요 약

수많은 정보의 홍수 속에 정보 과다는 현대인의 피할 수 없는 문제로 대두되었다. 특히, 인터넷과 컴퓨팅 기술의 발전으로 정보 자원이 급속도로 증가하고 있다. 따라서, 사용자가 원하는 정보를 찾아내는 것은 더욱 어려워졌다. 이러한 정보 검색 문제들을 해결하기 위하여 많은 정보 검색 시스템이 나타나게 되었다. 현재의 정보 검색 시스템들은 문서 검색에서는 사용자의 요구에 맞는 결과를 찾아 주고 있다. 그러나, 화상에 대한 검색 시스템의 연구는 초보 단계이기 때문에 사용자의 요구에 맞는 결과를 출력하지 못하고 있다. 이러한 문제를 해결하기 위해서 본 논문에서는 화상 검색 시스템이 사용자 모델을 이용하여 사용자에게 적용할 수 있는 기능을 부여하기 위하여 지능 사용자 인터페이스에 관한 고찰을 통한 인간-컴퓨터의 상호 작용 모델인 HCOS(Human-Computer Symmetry) 모델을 적용하고 이를 기반으로 화상 검색 시스템에서의 사용자 적응 능력을 갖는 지능 사용자 인터페이스를 제안하였다. 지능 사용자 인터페이스는 정보 검색의 복잡성과 사용자와 시스템간의 의미상의 차이를 감소시켜야 하며 사용자의 질의 성향과 관심을 반영할 수 있도록 학습 기능과 적응 기능을 포함하고 있어야 한다. 이를 위해 본 논문에서는 화상 검색을 위한 사용자 모델에 기계 학습(Machine Learning) 알고리즘인 결정 트리(Decision Tree)와 역전파 신경망(Backpropagation Neural Network)을 사용하였다. 지능 사용자 인터페이스의 화상 검색 실험을 통하여 시스템이 사용자에게 적응하여 검색 효율이 좋아짐을 알 수 있었다.

User Adaptation Using User Model in Intelligent Image Retrieval System

Yong-Hwan Kim[†] · Phill-Kyu Rhee^{††}

ABSTRACT

The information overload with many information resources is an inevitable problem in modern electronic life. It is more difficult to search some information with user's information needs from an uncontrolled flood of many digital information resources, such as the internet which has been rapidly increased. So, many information retrieval systems have been researched and appeared. In text retrieval systems, they have met with user's information needs. While, in image retrieval systems, they have not properly dealt with user's information needs. In this paper, for resolving this problem, we proposed the intelligent user interface for image retrieval. It is based on HCOS(Human-Computer Symmetry) model which is a layered interaction model between a human and computer. Its' methodology is employed to reduce user's information overhead and semantic gap between user and systems. It is implemented with machine learning algorithms, decision tree and backpropagation neural network, for user adaptation capabilities of intelligent image retrieval system(IIRS).

* 본 결과는 정보통신부 정보통신 우수시범학교(대학원) 지원사업에 의해 1999년 인하대학교 교내 연구비 지원으로 수행하였음.

† 정 회 원 : 인하대학교 대학원 전자계산학과

†† 종 신 회 원 : 인하대학교 전자계산공학과 교수

논문접수: 1999년 4월 13일, 심사완료: 1999년 10월 29일

1. 서론

최근에 들어 정보의 홍수라고 불릴 만큼 다량의 정보 속에서 필요한 정보를 찾아내는 것은 매우 중요하다. 즉 현대인에게 정보 과다(information overload) 문제가 점차 쟁점으로 대두되고 있다. 특히, 인터넷과 컴퓨팅 기술의 급속한 발전으로 문자 정보뿐만 아니라 멀티미디어 정보의 급속한 증가에 따라 멀티미디어 정보 검색에 대한 필요성이 증대되고 있다. 이러한 멀티미디어 정보 중에 화상 검색을 위해 많은 연구가 진행되어 Virage[1], MediaRoadMap[7], VisualSeek[13], QBIC[15] 등과 같은 화상 검색 시스템들이 나타났다. 화상 검색 시스템들은 화상을 분석한 특징을 분석한 정보에 따라 사용자가 찾고자하는 화상을 찾아주는 역할을 한다. 그러나, 현재의 화상 검색 시스템들은 사용자의 충분하지 않은 질의 정보만을 가지고 검색을 수행하기 때문에 검색의 결과가 사용자가 원하지 않는 결과를 출력할 수도 있다. 이러한 문제를 해결하기 위해서는 검색을 수행할 때 사용자의 검색 성향과 관심을 파악하여 이를 질의에 반영하여 검색을 수행할 수 있도록 해야 한다[17]. 사용자의 성향과 관심을 파악하기 위한 방안 중에 하나로 사용자 모델을 도입한 지능 사용자 인터페이스 개념을 도입할 수 있다. 정보화 사회로의 진행에 따라서, 일반 사용자의 전문성 내지는 전문화 가능성을 기대할 수 없기 때문에 시스템의 측면에서 사용자들과 시스템 사이의 의미상 차이(semantic gap)가 감소되어야만 한다. 또한, 사용자의 정보 활용도를 향상시킬 수 있어야 한다.

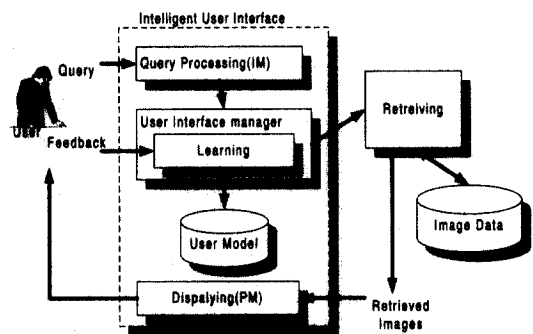
본 논문에서는 화상 검색 시스템을 위한 지능형 사용자 인터페이스를 위한 사람-컴퓨터 상호 작용 모델(Human-Computer Interaction Model)을 적용하고, 이 모델을 기반으로 한 화상 검색에서의 사용자와 시스템 사이의 의미상의 차이를 줄일 수 있는 지능 사용자 인터페이스의 설계 및 구현에 대하여 기술한다.

이에 앞서 언급한 지능 사용자 인터페이스 개념을 도입한 지능형 화상 검색 시스템(IIRS : Intelligent Image Retrieval System)을 설계하는데 있어서 바람직한 전략은 다음과 같다.

- 사용자의 정보검색 부담을 감소시켜 준다.
- 사용자의 의도를 가능하면 적은 횟수의 인터랙션에 파악하여야 한다.

- 대화를 통한 사용자의 의도를 파악하는데 단일 질의를 통해서가 아니라 기존에 축적된 사용자 모델을 기반으로 문맥을 통한 이해한다.
- 사용자의 실수를 포용할 수 있어야 한다.
- 시스템이 사용자의 특성에 적용할 수 있어야 한다.
- 시스템의 성능이 시간이 지남에 따라서 개선되어야 한다.

(그림 1)은 화상 검색 시스템에서의 지능 사용자 인터페이스의 역할을 나타낸 것이다. 지능 사용자 인터페이스에는 질의를 내부 표현으로 바꾸어 주는 질의 처리 기능(Query Processing)을 갖는 IM(Interpretation Module), 사용자 인터페이스 관리자(User Interface manager), 사용자 모델(User Model)과 사용자에게 검색 결과를 출력하는 기능(Displaying)을 갖는 PM(Presentation Module)으로 되어 있다. 사용자 인터페이스 관리자(User Interface Manager)는 사용자 모델의 데이터를 이용하여 사용자 성향과 관심에 적용하는 부분으로 질의에 대한 검색된 결과에 대한 사용자의 피드백을 사용자로부터 받아 학습하는 학습 기능(Learning)으로 이루어져 있다. 사용자 모델은 사용자 질의와 피드백 정보를 저장하여 사용자의 검색 성향과 관심을 파악하는 부분으로 지능 사용자 인터페이스의 중요한 한 부분이다.



(그림 1) 검색 시스템에서의 지능 사용자 인터페이스의 역할

본 논문에서는 지능 사용자 인터페이스의 적용과 학습 능력을 결정 트리(Decision Tree)와 역전파 신경망(Backpropagation Neural Network)이라는 기계 학습(Machine Learning) 알고리즘을 통하여 구현하였다. 이러한 지능 사용자 인터페이스의 설계와 구현은 화상

검색을 원하는 사용자에게 효율적인 검색을 제공하고 찾고자하는 화상을 쉽고 빠르게 찾게 해줄 수 있었다.

2. 지능사용자 인터페이스에 대한 고찰

지능 사용자 인터페이스 기술은 시스템이 사용자의 의미를 파악하는데 중요한 역할을 담당한다. 이 장에서는 지능 사용자 인터페이스의 개념과 사용자 인터페이스 모델인 HCOS 모델(Human-COmputer Symmetry Model)에 관하여 기술한다.

2.1 지능 사용자 인터페이스의 개념

지능 사용자 인터페이스는 여러 방법으로 정의되고 설계될 수가 있다. 지능 사용자 인터페이스의 보다 나은 설계를 위해 사용자 인터페이스 방법론 사이에 구분은 매우 중요하다. Eberts에 의하면, 사용자 인터페이스의 종류는 다음과 같이 네 가지로 나뉘어 질 수 있다[6]: 실험적(Empirical) 방법, 예측 모델(Predictive modeling) 방법, 의인화(Anthropomorphic) 방법, 인지 과학적(Cognitive) 방법이다.

실험적 방법에서는 사용자 인터페이스의 설계자는 일반적으로 다양한 장치나 기술로부터 적당한 선택을 하여야 한다. 올바른 선택을 하였는가의 여부는 실험에 의해서 판단된다. 이 방법의 기본 착상은 설계 단계에 있어서의 대부분의 직관은 가능한 경우를 모두 고려할 수 없으나 실험을 통하여 더 많은 경우를 조사할 수가 있다는 것이다.

예측 모델에 의한 방법의 기본 착상은 프로토타입이 만들어지기 전에 최상의 설계를 예측하자는 것이다. 설계자는 현존하는 모델을 잘 이해하고 어떻게 사용되는지에 대해 잘 이해하고 있어야 한다. 실험적 방법과 비교하여 주된 차이점은 사용자의 행동이 설계 단계에서 반드시 평가되어 져야 한다는 것이다. 이 방법의 전형적인 예가 GEOM 모델이다. GEOM 모델에서는 설계자가 그의 설계를 평가할 수 있도록 하기 위해 상태 전이망(state-transition networking)이 사용된다.

의인화(Anthropomorphic) 방법에서는 인간의 지능이 사용자 인터페이스의 이상적인 모델이라고 간주된다. 따라서, 설계자는 인간과 인간이 대화하는 과정과 방법을 인간과 컴퓨터 사이의 상호 작용의 모델로 적용한다. 이 방법에서 컴퓨터 시스템은 인간의 지능을 모방하려고 노력하며 컴퓨터 시스템은 사람이 다른 사

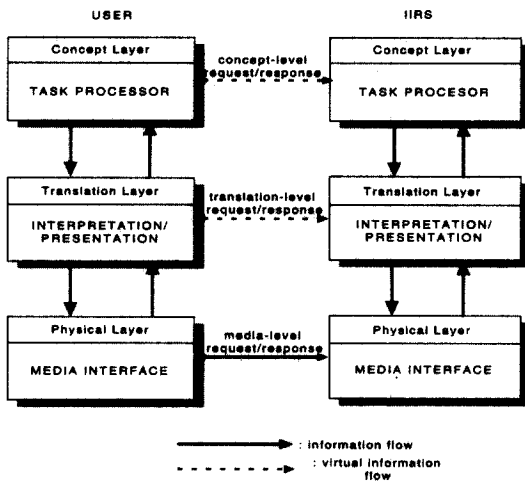
람과 대화하는 것과 유사하게 사용자와 상호 작용을 하도록 하여야 한다. 이 접근 방법에서 지능적 요소들은 자연어 처리, 음성 인식, 패턴 인식, 지식 표현, 사용자 모델 등의 사용자 인터페이스와 연결된다. 앞의 방법들과 구별되는 두 가지 접근 방법이 있는데 이는 인간과 인간의 상호 작용을 모델로 하고, 정보를 취합하는 근원으로 사용하거나, 인간과 컴퓨터의 상호 작용이 불일치 점을 찾기 위해 인간과 인간의 상호 작용과 비교하는 방법이다.

인지과학적 방법에서는 인지과학의 이론을 보다 효율적인 사람과 컴퓨터 정보의 처리를 위해서 사용자 인터페이스에 적용한다. 인지 이론은 사람이 어떻게 정보를 알아내고, 저장하고, 추출하는 방법과 사람이 문제를 해결하고 응답하기 위해서 어떻게 그 지식을 운용하는지를 연구한다. 이 방법에서 세 가지의 가장 중요한 구성 요소는 개념 모델, 정신 모델, 그리고 디스플레이 표현이다. 개념 모델은 컴퓨터 시스템에 대한 설명이다. 정신 모델은 디스플레이 표현과의 상호 작용을 통해서 얻어지는 사용자에 의한 시스템 모델이다. 이 방법의 목표는 개념 모델처럼 정신 모델을 정확하고, 일관성 있고, 완벽하게 해 주는 가장 알맞은 디스플레이 표현을 선택하는 것이다. 인지과학적 방법은 사람을 적용성 있고, 융통성 있고, 문제 해결을 위해 환경과의 상호 작용에 적극적으로 관여하는 것을 기본 전제로 하고 있다. 이 방법은 인간 지능의 모델에 토대를 두고 있으며, 지능 사용자 인터페이스에 필요한 특징을 명시하는 데 사용될 수 있다는 점에서 의인화(anthropomorphic)방법과 유사하다.

2.2. HCOS(Human Computer Symmetry)모델

(그림 2)에서는 2.1절에서 기술한 지능 사용자 인터페이스에 대한 고찰을 기반으로 지능형 화상 검색 시스템(Intelligent Image Retrieval System: IIRS)을 위한 인간-컴퓨터 상호 작용 모델인 HCOS(Human-Computer Symmetry) 모델[17]을 적용하였다. HCOS 모델은 인간의 정보 처리 기능과 컴퓨터 정보 처리 시스템의 기능이 유사 또는 동일한 과정과 방법을 따른다고 가정한다. 따라서, 인간과 컴퓨터 모두 정보 처리 시스템으로 모델링된다. 현존하는 대부분 인간-컴퓨터 대화 모델은, 인간 정보 처리 시스템은 보통 태스크를 개시하고 컴퓨터 시스템은 주로 태스크를 처리한다고 본다. 그러나, HCOS 모델에서는 사람, 컴퓨터 모두 태

스크를 활성화하고 처리할 수 있다고 가정한다. 컴퓨터 시스템은 더 이상 수동적인 객체로서가 아니라 능동적인 정보 처리 객체로 다루어질 수 있다. 또한, 사용자는 더 이상 모호한 객체가 아니라 기능이 분명하게 공학적으로 정의된 객체로 취급된다. HCOS에서, 사용자(인간)와 기계(컴퓨터 정보처리 시스템)는 각각 개념층(concept layer), 번역층(translation layer), 미디어아(physical layer) 세 개의 계층을 가진다.



(그림 2) HCOS 모델

개념층에서는 사용자(USER)는 주로 태스크를 시작하고 태스크 처리 결과를 평가한다. IIRS는 태스크를 처리하며 태스크를 시작한 사용자에게 응답한다. 태스크는 IIRS에서의 작업의 단위이며, 정보 생성(information generation), 정보 수집(information gathering), 정보 편집(information edit), 질의, 응답, 교수 명령 등이 포함된다. 태스크는 IIRS의 적용 도메인에 따라서 보다 구체적으로 정의될 수 있다. 개념 층에서, 사용자는 보통 태스크를 시작하며, IIRS에 의해서 태스크가 처리되도록 요구한다. IIRS는 목적 정보를 생성시킴으로써 사용자의 요구에 응답한다. 사용자는 생성된 정보가 만족스러운지 여부를 결정한다. 만약 사용자가 만족한다면 그 태스크는 종료된다. 그렇지 않으면 사용자는 적절한 수정 명령을 발생함으로써 정보를 정정/재생산해 줄 것을 요구한다. IIRS는 사용자의 요구에 맞게 정보를 생산하여서 사용자에게 응답한다. 사용자와 IIRS는 각각 개념 표현(conceptual representation)

이라 불리는 개념 층의 내부 표현을 가지고 있다.

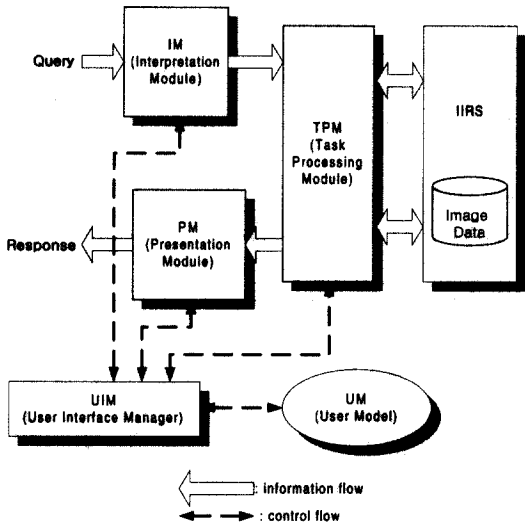
사용자와 IIRS의 번역층은 해석 처리(interpretation processing)와 표현 처리(presentation processing)로 구성되어 있다. 사용자와 IIRS의 표현 처리는 각각 개념 표현을 디스플레이 표현으로 바꾸어 준다. 해석 처리는 디스플레이 표현을 개념 표현으로 바꾸어 준다. IIRS의 모든 구성 요소는 오류를 발생시킬 수 있다. 그러나 모델의 단순화를 위해서 미디어층의 입출력 처리는 오류를 발생시키지 않는다고 가정한다. IIRS의 표현 처리는 모니터에 디스플레이 기술이다. 표현 처리에 있어서 오류가 없다는 가정은 현재 목표로 하는 IIRS에서 단지 간단하고 매우 제한적인 그래픽 기능만이 사용되기 때문에 타당하다. 위의 분석을 바탕으로 다음 3장에서는 IIRS를 위한 지능 사용자 인터페이스의 설계에 대한 구체적인 내용을 기술한다.

3. 지능 사용자 인터페이스의 설계

일반적으로 사용자는 화상 정보 검색의 검색하는 정보의 반복성과 검색하는 성향이 어떠한 패턴이 있고 임의가 아니고 특정 도메인을 주로 조회하는 특성이 있다. 지능 사용자 인터페이스의 요구 사항은 이러한 정보 검색 특성을 기반으로 하여 IIRS를 위한 사용자 인터페이스에 적용성 또는 학습능력을 가지는 확장된 사용자 인터페이스로 규정한다. 이를 위하여 본 논문에서는 주어진 질의에 관련되어 조회되는 응답으로부터 사용자 모델을 구성하고, 사용자와 시스템 사이의 과거의 질의/응답을 근거로 하여 사용자의 질의 성향을 학습하여 IIRS가 다음의 사용자 질의의 수행을 지능적으로 할 수 있도록 한다. (그림 3)은 지능 사용자 인터페이스의 개념 설계에 대한 내용이다.

지능 사용자 인터페이스는 IM, PM, TPM, UM, UIM으로 구성된다. IM(Interpretation Module)은 입력되는 질의를 관리하며 내부 표현인 특징 벡터로 변환한다. 이 질의의 내부 표현은 화상 검색을 위한 기초 자료로 사용되며 동시에 질의의 유형을 학습 또는 판별하기 위하여 UIM의 입력된다. PM(Presentation Module)은 질의에 대한 응답의 디스플레이 관리를 한다. TPM(Task Processing Module)은 내부 표현으로 변환된 질의를 기반으로 사용자에게 출력 정보(응답)를 생성 또는 생성을 위한 화상의 집합과 인터페이스를 담당한다. UM(User Model)은 사용자의 질의 성향에 대한 정

보를 포함하는 데이터베이스 또는 파일 구조를 가지고 있으며, UIM(User Interface Manager)은 주로 UM의 운용과 관련된 결정을 처리를 한다. UIM은 UM을 기반으로 사용자의 질의를 추론하는 질의 성향 학습, 사용자의 만족도 피드백을 이용 UM의 질의 패턴과 화상의 집합과 연계 등을 담당한다.



(그림 3) IUI 개념모델

4. 지능 사용자 인터페이스의 구현

3장의 설계에서 IIRS의 사용자 적응 방법(adaptation method)은 중요하며 사용자와 IIRS 간의 의미상의 차이를 줄일 수 있고 사용자의 질의 패턴을 학습하여 새로운 질의에 대하여 사용자에게 가장 적합한 해결책을 제시할 수 있다. 이러한 적응 방법의 구현은 결정 트리와 역전파 신경망을 이용하여 구현하였다.

4.1 결정 트리

결정 트리는 노드와 에지로 구성된 일반적인 트리이며 각 내부 노드는 학습 데이터가 표현된 한 속성을 가지며, 각 에지는 내부 노드의 속성이 갖는 값을 나타낸다. 일반적으로 데이터가 주어졌을 때 분류를 할 수 있는 결정 트리는 여러 가지가 있을 수 있다. ID3 학습 알고리즘은 많은 학습 데이터로부터 최적의 결정 트리를 구축하여 준다.

학습 데이터는 전체 화상 데이터와 화상 데이터의

속성인 키워드로 이루어진다. ID3 알고리즘은 정보 이론에 따르는 엔트로피 개념을 이용하여 최적의 결정 트리를 구축하여 준다. 정보 이론의 엔트로피는 데이터의 무질서도를 나타내는 정량적인 값이다. 엔트로피는 다음과 같이 정의된다.

$$\text{엔트로피 } H(S) = -\sum_k \left(\frac{|C_k|}{|S|}\right) \log\left(\frac{|C_k|}{|S|}\right) \quad (1)$$

식 (1)에서 S는 입력된 학습 데이터 집합을 의미하며, C_k는 S에 속해 있는 데이터 중에 부류 k인 부분 집합을 나타낸다. |S|와 |C_k|는 각각 집합 S와 C_k의 크기를 나타낸다. 일반적으로 입력되는 화상 데이터들을 구분하기 위한 정보가 없기 때문에 엔트로피가 매우 높다. 그러나, 구분 정보들에 의해 학습이 완료된 상태에서는 각 단말 노드가 하나의 화상으로만 결정되므로 엔트로피는 0이 된다. 따라서 ID3 학습 알고리즘은 엔트로피를 최소화하는 과정을 반복한다.

$$H_{FV[i]}(T) = \sum_{j=1}^n \frac{|T_j|}{|T|} \times H(T_j) \quad (2)$$

구축된 결정 트리에서 각 비단말 노드는 식 (2)의 결과를 최소로 하는 특징 벡터의 한 요소 FV[i]를 나타내며 탐색 과정의 테스트를 위한 속성으로 사용된다. 단말 노드를 구성하는 조건은 2가지가 있다. 먼저 모든 FV[i]에 대하여 식 (2)의 결과 값이 0이 되는 경우로서 입력 데이터 집합의 모든 특징 벡터들이 하나의 데이터를 나타낸다.

구축된 결정 트리는 화상 검색을 위해 사용된다. IM을 통해 생성된 특징 벡터는 결정 트리의 탐색을 통해 하나의 화상과 연계된다. 결정 트리의 탐색 과정은 결정 규칙을 적용해 가는 과정이다.

4.2 역전파 신경망

본 논문에서는 사용자 모델을 역전파 신경망을 이용하여 학습하였다. 역전파 신경망은 직감적으로 이해할 수 있으며 또한 프로그래밍이 비교적 용이하기 때문에 주로 패턴인식, 제어 등에서 많이 사용되는 알고리즘이다.

본 논문에서 사용된 역전파 신경망은 입력층, 은닉층, 출력층의 3개 층으로 구성된다. 각각의 계층의 학습 데이터는 사용자가 검색된 결과의 화상에 대하여 질의와 얼마나 유사한지에 대하여 평가한 피드백 정

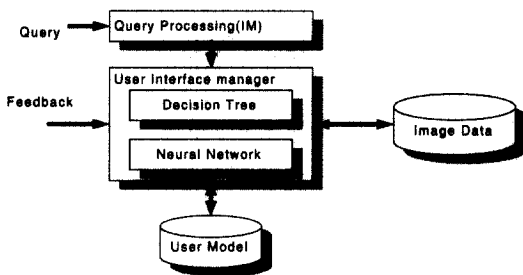
보를 이용하여 학습한다. 이 때 신경망의 오차가 0.1이하로 떨어질 때까지 학습한다.

학습이 완료된 신경망은 각 유니트들간의 연결 강도를 모두 저장하고 있기 때문에 벡터가 입력되고 결정 트리에서 탐색이 불가능하면 원하는 결과에 근접한 값을 신속하게 얻을 수가 있다. 결과값을 얻는 과정은 역전파 신경망 학습 단계의 첫 번째 과정을 적용한다. 즉, 연결 강도와 역치함수를 이용하여 결과를 얻게 된다.

4.3 지능 사용자 인터페이스의 구현

IIRS의 사용자 모델을 이용한 사용자 적용을 할 수 있도록 해주는 부분이 UIM이다. UIM은 앞서 설명한 두 개의 기계 학습 알고리즘을 이용하여 (그림 4)와 같이 구현하였다. 결정 트리와 역전파 신경망은 사용자 모델 관리자 역할 뿐만아니라 그 고유의 기능이 검색 기능을 갖도록 하였다. 따라서 질의에 바로 사용자 모델의 적용 후 검색이 이루어진다. 또한 검색 결과에 대한 사용자 피이드백을 받아 영상 정보와 같이 사용자 모델에 입력되어 사용자의 성향을 학습하게 된다. IIRS의 검색의 대상에 사용된 영상은 자연 환경을 담은 영상으로 65개의 영상이 사용되었으며 각각의 화상에 고정된 키워드 20개마다 가중치가 부여되었다.

이는 학습 알고리즘의 학습 시간을 고려한 데이터의 최대치를 사용한 것이다. 이 영상의 키워드는 결정 트리과 역전파 신경망의 학습 데이터로 사용한다.



(그림 4) 사용자 모델링 관리자의 구현

5. 실험 및 결과

본 논문의 화상 검색의 효율 측정 실험은 인터넷 기반의 CGI 프로그램을 이용하였다. 사용한 화상은 자연 풍경 화상으로 화상의 키워드 등록을 하여 검색을 하였다. 정보 검색 시스템을 평가한다는 것은 검색 시스

템들의 동작과 성능을 검사하는 것이다. 이러한 정보 검색 시스템의 평가는 전체 시스템, 시스템의 부분, 특별한 내부 기능 또는 내부 기능의 구성요소에 대한 평가일 수도 있다. 정보 검색 시스템의 평가에서의 중요한 요소는 질의와 데이터간의 관계를 규정한 “연관성 (Relevance)”이 중요하다. 연관성은 정보 검색 시스템의 효율을 측정하는 데 필요한 요소이다. 그런데, 연관성은 사람의 개념마다 다르기 때문에 같은 질의에 따라 다른 결과가 나올 수 있다. 이러한 사용자마다의 다른 개념은 시스템의 성능 측정 방법이 일정하다해도 실험 결과가 사용자마다 다른 수치의 결과가 나올 수 있다.

본 논문에서의 화상들은 실험을 위해 지능 사용자 인터페이스를 사용한 검색에서 사용하는 질의와 사용된 화상들간의 연관성을 정의하였다. 즉, 어떠한 사용자 질의에 어떠한 화상이 관련이 있는지를 미리 정의하여야 하기 때문에 실험에 사용되는 질의에 사용자가 찾아지기를 희망하는 화상을 연관성을 정의하여 놓았다. 이러한 방법으로 화상 검색을 위한 지능 사용자 인터페이스에 대한 성능을 평가하게 된다. 실험에 사용된 데이터는 자연 풍경을 대상으로 한 화상을 사용하였다.

(그림 5)는 자연 풍경 화상을 사용하여 검색된 결과를 나타낸 것이다. (그림 5(a))는 지능 사용자 인터페이스의 적용하기 전에 사용자 질의에 의해 검색된 화상들이다. 이 검색된 화상 중에 사용자가 세 번째 화상을 선택하여 자신의 만족도를 “만족”하다고 표시하여 그 결과를 시스템이 학습할 수 있도록 하여 준다. 학습한 결과를 다시 같은 질의로 검색을 시도하면 (그림 5(b))와 같이 처음 검색한 결과와 다른 결과를 얻게 된다. 여기에서 살펴보면 사용자가 만족하다고 피이드백 정보를 준 화상이 검색 결과의 순위가 향상된 것을 볼 수가 있는데 이는 사용자의 검색 성향 또는 관심에 적용하였다고 볼 수 있다. (그림 5(c))에서는 네 번째의 화상에 대하여 사용자가 “만족”하다고 표시한 경우에 사용자 질의를 학습한 후 검색의 결과로 (그림 5(d))의 결과가 나타난 것이다. 여기에서도 마찬가지로 (그림 5(c))의 네 번째 결과가 두 번째로 순위의 향상을 가져왔다. 이러한 검색과 피이드백의 반복은 사용자에게 좀 더 자신이 원하는 결과를 빠르게 얻을 수 있게 해준다.

위의 실험의 결과로 지능 사용자 인터페이스의 효율

(그림 5) 지능 사용자 인터페이스의 실험 결과

을 측정하기 위해 위의 연관성을 이용하여 사용자마다 출력된 결과를 조사하여 정량화된 값을 구할 수가 있는데 실험을 각각의 사용자에 대하여 사용자 질의에 의해 검색된 결과를 보고 위의 연관성을 이용하여 다음과 같은 2가지의 검색의 효율을 측정할 수 있는 값을 얻을 수 있다.

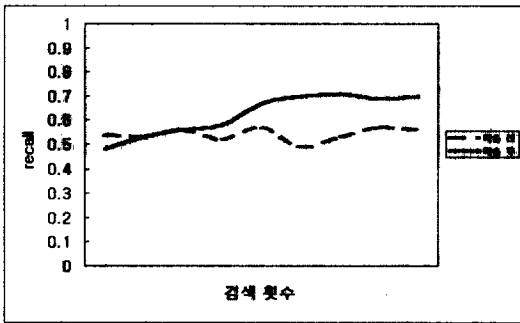
$$\text{재현율(recall)} = \frac{\text{질의와 관련된 검색 화상의 수}}{\text{전체 화상 중 질의와 관련된 화상의 수}}$$

$$\text{정확률(precision)} = \frac{\text{질의와 관련된 검색 화상의 수}}{\text{검색된 전체 화상의 수}}$$

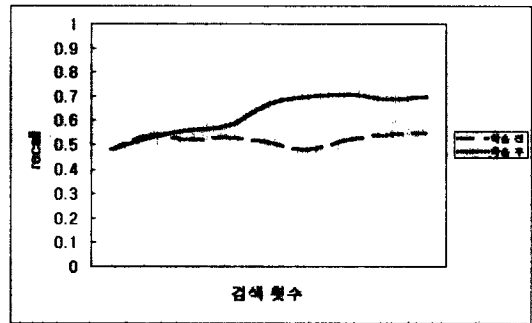
재현율(recall)은 사용자 질의에 의해 검색된 화상이 처음 사용자가 그 질의에 찾아 주기를 희망한 화상의 전체 집합 중에 검색된 화상이 얼마나 되는지를 정량적인 값으로 나타낸 것이고 정확률(precision)은 사용

자 질의에 의해 검색되어진 전체 화상 중에 데이터가 그 질의에 찾아 주기를 희망한 화상이 검색된 화상이 얼마나 되는지를 정량적인 값으로 나타낸 것이다. 이 중에 정확률은 사용자가 얼마나 검색된 결과 중에 찾고자하는 정보를 얼마나 빠른 시간에 찾을 수 있는지를 나타내는 검색 시간의 단축을 의미한다. 따라서, 정확률이 높다는 것은 사용자가 원하는 정보를 짧은 시간에 찾을 수 있음을 의미한다.

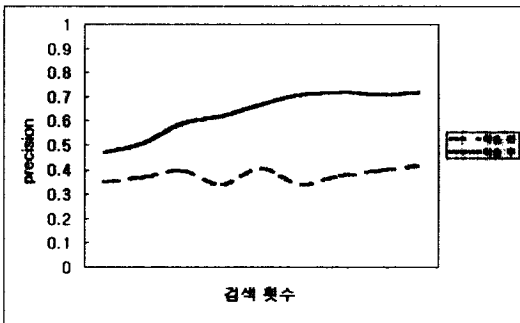
본 논문에서는 사용자마다 지능 사용자 인터페이스를 사용하지 않았을 경우의 검색에서의 재현율과 정확률, 지능사용자 인터페이스를 사용하여 검색하는 경우에 대하여 재현율과 정확률을 계산하였다. 실험에는 자연 풍경을 담은 화상 데이터를 사용하여 2명의 사용자에 대한 실험을 하였다. 이 실험에서 나타난 정확률과 재현율의 수치는 한 번 질의할 때 다음 번의 여러



(a) 재현율(recall)의 비교

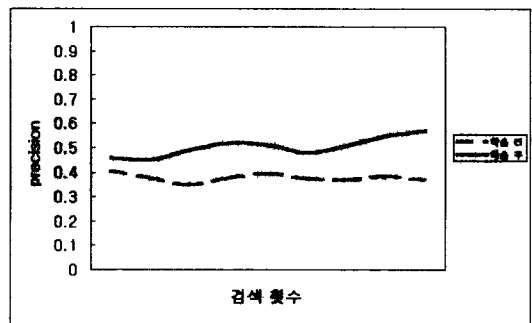


(a) 재현율(recall)의 비교



(b) 정확률(precision)의 비교

(그림 6) 사용자 A의 검색 효율 비교



(b) 정확률(precision)의 비교

(그림 7) 사용자 B의 검색 효율 비교

검색에서의 실험 결과의 평균을 계산한 것이다.

(그림 6)과 (그림 7)은 각각 사용자 A와 사용자 B에 대하여 검색 효율을 지능 사용자 인터페이스를 사용하지 않은 경우와 지능사용자 인터페이스를 사용한 경우의 검색 효율을 나타낸 것이다. 위의 결과들에서 사용자 A는 사용자 질의 학습 전의 검색 효율은 재현율(recall)의 평균은 0.54이고 정확률(precision)은 평균 0.38이 나타났다. 이러한 검색 효율을 사용자 질의 학습 후에는 재현율이 평균 0.71이고 정확률이 평균 0.63으로 나타났다. 이는 재현율, 정확률 모두가 향상된 수치로 나타났다. 그러나, 사용자 질의 학습 후의 검색 효율의 향상은 사용자에 적용함으로써 0.8이상의 수치를 나타내고 있다. (그림 7)의 사용자 B의 검색 결과에서도 재현율은 0.51에서 0.62로 정확률은 0.38에서 0.50으로 나타나 사용자 질의 학습이 개인의 검색 효율을 높일 수 있음을 알 수 있다. 검색 효율의 수치상의 개인적인 차이는 개인마다 화상과 질의의 관련성을 정의하는 것이 다르기 때문이다. 위의 실험 결과와

같이 지능 사용자 인터페이스를 사용하여 검색함으로써 화상 검색 시스템이 각각의 사용자에 적용함에 따라 검색 효율이 증가하였다. 앞의 (그림 5)의 화상의 검색 결과에서와 같이 사용자는 검색 결과에 대하여 만족도를 피드백 정보를 주어 지능 사용자 인터페이스에 학습함으로써 사용자의 검색 성향과 관심을 반영할 수 있었다.

6. 결론 및 향후 연구 과제

화상 검색을 위한 지능 사용자 인터페이스는 인간-컴퓨터 상호 작용 모델인 HCOS(Human - Computer Symmetry) 모델을 적용하여 사용자와 컴퓨터간의 대화를 통하여 시스템이 사용자의 행동에 적응할 수 있도록 설계하였고 실제로 실험 결과에서 나타난 것처럼 자연 화상의 검색에 있어서 사용자의 질의에 사용자의 의도와 연관된 화상을 검색해 낼 수 있었다. 다른 도메인 화상 검색에도 그에 따른 키워드를 정의하여 지능

사용자 인터페이스를 사용한 검색을 지원할 수 있다.

현재의 많은 시스템들이 사용하고 있는 검색 방식에서의 사용자 인터페이스인 사용자가 직접 명시한 사용자의 관심만을 사용자 모델로 학습하여 사용자의 동적인 관심의 변화를 시스템이 적용할 수 없는 것과 비교하면 사용자 인터페이스가 사용자에게 적용할 수 있다는 면에서 본 연구에서 제안한 지능 사용자 인터페이스는 화상 검색을 필요로 하는 사용자에게 원하는 검색 결과를 정확히 찾을 수 있었다.

본 논문에서는 기계 학습 알고리즘인 결정 트리와 역전파 신경망을 이용하여 이러한 사용자에게 적용할 수 있는 지능 사용자 인터페이스를 구현하였다. 이 두 기계 학습 알고리즘은 사용자의 성향과 관심을 파악하는 데는 적당하였다. 그러나, 계속적인 검색으로 인하여 늘어나는 사용자 질의, 즉 학습 데이터가 증가하면 다시 처음부터 학습을 해야하는 문제점을 증가 학습 방법(Incremental learning method)에 대한 구조의 적용을 학습 시간의 단축을 위하여 연구하여야 한다.

참 고 문 헌

- [1] Amarnath Gupta, Visual Information Retrieval : A Virage Perspective(Revision 4), <http://www.virage/wpaper/>.
- [2] Ana B. Benitez, Mandis Beigi, and Shin-Fu Chang, Using Relevance Feedback in Content-based Image Metasearch, IEEE Internet Computing, July · August, 1998.
- [3] A. Tan and C. Teo, Learning User Profiles for Personalized Information Dissemination, Proceedings of 1998 IEEE International Joint Conference on Neural Networks, pp.183-188, May 1998.
- [4] B. Sheth and P. Maes, Evolving Agents for Personalized Information Filtering, Proceedings of the IEEE Conference on Artificial Intelligence for Applications, IEEE Press, 1993.
- [5] E. Bloedorn, I. Mani, and T. Richard MacMillan, Machine Learning of User Profiles : Representational Issues, the 13th International conference on Artificial Intelligence Proceedings (AAAI96), August 1996.
- [6] E. Rich, User Modeling Via Stereotypes, Cognitive Science, Vol.3, pp.329-354. 1979.
- [7] F. Golshani and Y. Park, Content-based image indexing and retrieval in ImageRoadMap, proceedings of SPIE on Multimedia Storage and Archiving System II, pp.194-205, November 1997.
- [8] F. Iris and S. Panchanathan, Review of Image and Video Indexing Techniques, Journal of visual communication and image representation. Vol.8. No.2. pp.146-166, June 1997.
- [9] G. Salton and M. McGill, Introduction to Modern Information Retrieval, New York, McGraw-Hill, 1983.
- [10] H. Ase and S. Kobayashi, Information Retrieval Condition Generation System using Cased-Based Reasoning, PRICAI 90, 1990.
- [11] I. Akoulchina and J. Ganascia, SATELIT-Agent : An Adaptive Interface Based on Learning Interface Agents Technology, The proceedings of the 6th International Conference on User Modeling(UM97), Chia Laguna, Sardinia, June 1997.
- [12] Joep Simons, Using a Semantic User Model to Filter the World Wide Web Proactively, The proceedings of the 6th International Conference on User Modeling(UM97), Chia Laguna, Sardinia, June 1997.
- [13] John. R. smith and Shin-Fu Chang, VisualSeek ; a fully automated Content-based image query system, the Fourth ACM multimedia conference, Boston MA. pp.87-98, November, 1996.
- [14] J. R. Quinlan, Induction of decision tree, Machine Learning 1, pp.81-106, 1986.
- [15] M. Flicker et al. Query by image and video content : The QBIC System. IEEE Computer, 28

(9):23-32, September 1995.

- [16] M. Pazzani, J. Muramatsu and D. Billsus, Syskill and Webert : Identifying interesting web sites, proceedings of the Thirteenth national Conference on Artificial Intelligence and the Eighth Innovate Applications of Artificial Intelligence Conference, pp.54-61, 1996.
- [17] P. Rhee, Y. Kim, B. Shim, "Intelligent Information Retrieval System for the Metadata of Korean Technical Journals," Proceedings of the 2nd International workshop on Information Retrieval with Asian Language, pp.173-184, 1997.
- [18] Paul Yap, et al. Learning Probabilistic User Profiles, AI Magazine, American Association for Artificial Intelligence, Summer 1997.
- [19] R. Kass, Building a User Model Implicitly from a Cooperative Advisory Dialog, UMUI 1:3 pp. 203-258, 1991.
- [20] Rogers Joey, Objected-oriented neural networks in C++, Academic Press, Inc. pp.83-132, 1997.
- [21] Simond Haykin, Neural Networks, Macmillan Publishing Compony, 1994.
- [22] Susan T. Dumais, Combining evidence for effective information filtering, AAAI Spring Symposium on Machine Learning and Information Retrieval Tech Report SS-96-07, AAAI Press, March 1996.



김용환

e-mail:g1991260@inhavision.inha.ac.kr

1997년 인하대학교 전자계산공학과 졸업(학사)

1999년 인하대학교 대학원 전자계산공학과 졸업(공학석사)

1999년~현재 인하대학교 대학원 전자계산공학과 박사과정

1999년~현재 동양공업전문대학 경영정보기술공학부 강사
관심분야 : 멀티미디어 정보 검색, 기계 학습, 인간-컴퓨터 상호 작용



이필규

e-mail:pkrrhee@dragon.inha.ac.kr

1982년 서울대학교 전기공학과 졸업(학사)

1986년 East Texas State University Computer Science 졸업(석사)

1990년 University of SouthWest Louisiana Computer Science 졸업(박사)

1982년~1985년 시스템공학연구소 연구원

1991년~1992년 한국전자통신연구소 선임연구원

1993년~1994년 IBM T.J. Watson Research Center 객원연구원

1992년~1996년 인하대학교 전자계산공학과 조교수

1996년~현재 인하대학교 전자계산공학과 부교수

관심분야 : 영상 처리, 기계 학습, 지능형 사용자 인터페이스, 인간-컴퓨터 상호 작용