

비주얼 브라우저를 이용한 이미지 검색 시스템

이지연 강사(연세대학교)

1. 서론

이미지 정보로의 접근 방식은 크게 두 가지로 나누어 생각할 수 있는데 그 한 가지는 각각의 이미지에 부가된 색인어 혹은 기술어를 탐색어로 이용하는 방법 (text-based access)이고 다른 하나는 이미지의 색채 (color), 이미지 속의 인물이나 물체의 형태 (shape), 혹은 이미지의 구조 (texture)를 이용하여 컴퓨터가 자동 분석한 이미지의 내용을 통하여 탐색을 하는 방법 (content-based access)이다 ([1]; [2]; [3]; [4]; [5]; [6]; [7]). 그러나 이 두 가지 방법 모두 이미지가 전달하는 풍부하고 동시에 주관적인 정보를 검색하는 데 있어서 한계성을 나타내고 있다. 색인어 등의 문자를 이용한 접근 방식은 여러 이용자의 주관적이고 다양한 해석을 수용하는 데 한계가 있으며 내용 분석을 통한 접근 방식은 아직도 검색되는 정보가 일차적 정보 (primary information)에 머무르는 수준이다.

색인어나 내용 분석을 이용한 이미지 검색의 한계성을 고려한 결과, 이 연구에서는 비주얼 브라우저 기법을 채택한 이미지 검색 시스템을 소개하고 검색의 효율성을 측정하였다. 여기에 소개된 비주얼 브라우저 기법은 사진들간에 이미 성립된 연결 경로 (browsing path)를 통하여 이용자가 탐색어를 이용할 필요 없이 원하는 사진을 선택함으로써 이와 비슷한 사진들을 검색하는 방법이다. 브라우저 경로는 이미

사진들에 부가된 색인어나 기술어 (descriptors)를 이용하여 같은 색인어를 공유하고 있는 모든 사진들을 검색하는 것이다. 그러므로, 여기에 소개된 브라우징 기법을 더 엄밀히 말한다면 문자 정보를 기반으로 한 비주얼 브라우징 (text-based visual browsing)이라 하겠다. 이러한 방식은 매뉴얼 혹은 자동으로 이루어진 색인어들이 있을 경우 특별히 많은 추가 노력이 없이 시스템 구축이 가능하다. 반면, 이용자 측면에서는 적절한 색인어의 선택에 따르는 부담을 해결할 수 있으므로 기존의 탐색어를 이용한 방식과는 상당히 다른 탐색 방식을 이용하는 효과를 기대할 수 있다. 이용자들은 자신들이 보는 사진 속의 시각적 정보를 문자로 전환시켜야 하는 과정 없이 직접 시각적 정보에 대응을 할 수 있는 것이다.

이 연구는 위에서 언급된 비주얼 브라우징 기법을 이용하는 이미지 검색 시스템을 구축하여 이용자 탐색 실험을 통하여 그 효율성을 측정하였다. 또한 비주얼 브라우징 기법을 전통적인 키워드 탐색 기법이나, 키워드 기법과 비주얼 브라우징 기법을 혼합한 방법과 비교하였다. 이 연구는 궁극적으로 소개된 비주얼 브라우징 기법이 해석적이고 이차적인 이미지 정보 검색 전반에 이용될 수 있는지를 조사하였다.

2. 연구 배경 및 목적

이 연구는 1990년대부터 활발하게 진행되어 온 이미지 정보에 대한 연구를 바탕으로 이미지 정보의 중요성을 인식하며 출발하였다 ([8]; [9]). 이와 더불어 지난 10여년간의 연구에서 공통적으로 지적된 이미지 색인 작업의 어려움과 문자를 이용한 이미지 기술의 한계를 공감하며 ([10]; [11]; [12]) 이러한 어려움을 완화시킬 수 있는 가능성 중의 하나로 비주얼 브라우징 기법을 제시하였다. 원정보 자체는 시각적인 형태로 존재하는 데 반하여 원정보로의 접근은 이와 다른 문자 형태를 따른다면, 시각정보가 문자정보로 전환되고 해석될 때 원정보가 걸러지고 굴절될 수 있는 가능성을 무시할 수 없는 것이다. 이보다는 이용자가 자신들이 보는 이미지 자체를 이용하여 자신들이 원하는 정보를 선택하게 하고 다음 단계로 나아가게 하는 직접 대응 방식을 이용하는 것이 이용자 측면에서 탐색어 이용 시 생길 수 있는 부담 및 부적절한 탐색어 선택의 문제점을 해결할 수 있는 방안이 될

수 있을 것이다.

이 연구는 우선 현재 널리 이용되고 있는 키워드 탐색 방법이 해석적이고 주관적인 이미지 정보 검색을 얼마나 잘 수행하고 있는가를 알아보기 위한 것이고 다음으로 이 연구의 핵심 부분인 비주얼 브라우징 기법을 소개하고 이 방법이 이미지 검색에 이용될 수 있는 가능성 (feasibility)을 진단하려는 목적 하에 이루어졌다.

3. 연구 방법

비주얼 브라우징 기법의 효율성을 측정하고 키워드 방식이나 혼합 방식과의 비교를 위한 이미지 검색 시스템을 구축하기 위하여 이미지 데이터베이스를 디자인, 구성하였는데 첫 단계는 사진들을 수집하고 색인하는 것이었다. 사진의 종류 선정에 있어서 이차적인 정보 (secondary information) ([13])의 대표적인 예로 감정 정보 (emotional information)를 선택하여 250 여 개의 사진을 선택, 색인하였다. 사진의 선택과 색인 작업은 6 명의 정보학 및 사진학 분야의 전문가와 박사 과정 학생들의 공동 협력 작업에 의해 이루어졌는데 이는 선택과 색인 작업의 객관성을 높이기 위한 것이었다. 이 데이터베이스의 검색 시스템은 세 가지 방법의 탐색 모드를 제공하는

그림 1. 비주얼 브라우징 초기화면

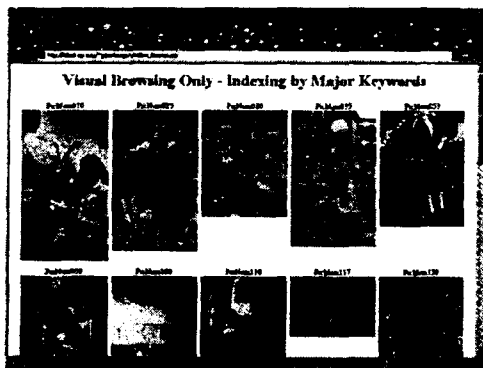


그림 2. 비주얼 브라우징 탐색결과화면

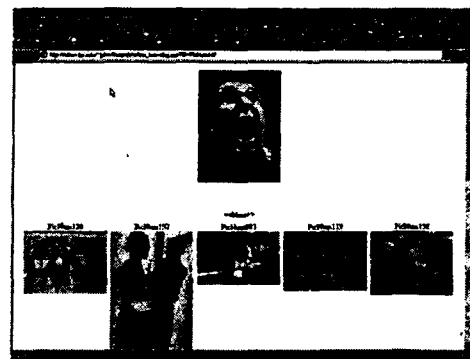


그림 3. 키워드 탐색화면

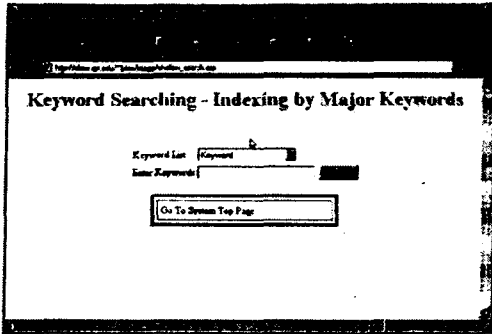
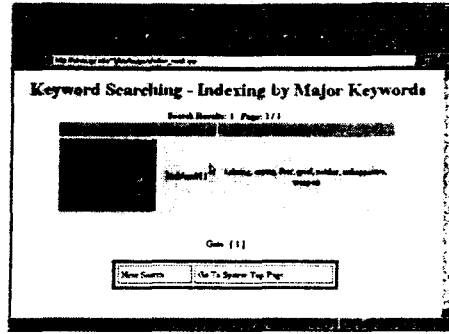


그림 4. 키워드 탐색결과화면



데 위에서 말한 비주얼 브라우징 방식, 키워드 탐색 방식과, 이 두 가지의 혼합 방식이다. 다음 그림들은 비주얼 브라우징 방식과 키워드 방식의 탐색 화면과 탐색 결과 화면이다.

그림 1 의 비주얼 브라우징 초기화면은 25 개의 대표사진으로 구성되어 있다. 각각의 사진은 이 연구에서 다루는 25 개 감정의 각각을 대표하는 사진이다. 25 개 감정은 연구자의 사전 연구에서 연구 대상자들이 주어진 여러 사진을 볼 때 자주 읽혀졌던 감정들을 선택한 것이다. 비주얼 브라우징은 25 개의 대표사진 중 이용자의 탐색 주제에 가장 부합하는 사진을 선택함으로써 시작한다. 이용자는 사진을 직접 선택함으로써 그와 비슷한 사진을 검색하게 되는데 그 예가 그림 2 이다. 선택된 사진은 위에 디스플레이되고 나머지 검색된 사진들이 화면 밑부분에 나열된다. 이용자는 자신이 원하는 정보를 얻을 때까지 계속 브라우징과 사진의 선택을 진행하게 된다.

반면, 키워드 탐색은 이용자가 탐색어를 그림 3 의 탐색어 입력 부분에 직접 입력하거나 그 위의 키워드 리스트를 참조하여 선택함으로써 이루어진다. 이때 탐색어가 복수인 경우 불리언 연산자로 합성할 수 있으며 구 (phrase)인 경우 인용부호 (quotation marks)를 이용하여 구 표시를 하게 된다. 키워드 탐색은 매치되는 키워드에 의하여 이루어지며 그림 4 가 탐색결과를 보여주고 있다. 우선 총 검색된 사진의 수를 표시하고 검색된 사진들을 그 색인어와 함께 보여준다. 이는

이용자로 하여금 citation pearl growing 처럼 탐색의 다음 단계에서 이용될 수 있는 탐색어의 선택을 돕기 위한 것이다.

혼합 방식은 이용자가 비주얼 브라우징과 키워드 탐색 방식을 원하는 대로 선택하고 바꿔가며 사용하는 방법이다.

이렇게 구축된 시스템을 이용하여 25 명을 대상으로 탐색 모드 비교실험을 하였는데 각 실험 대상자는 세 가지의 각기 다른 탐색 방식을 이용하여 정해진 3 가지의 감정을 나타내는 사진들을 검색하도록 하였는데, 예를 들어 키워드 방식을 이용하여 행복을 나타내는 사진을 검색하고, 다음으로 비주얼 브라우징을 이용하여 흥분감의 사진을 검색하고, 마지막으로 혼합 방식을 이용하여 슬픈 감정을 나타내는 사진을 검색하도록 하는 것이었다. 이 때 정해지는 감정과 탐색 방식의 종류 및 순서는 고르게 분포시켰다. 실험 대상자는 실험 시에 자신들의 탐색 전략 및 탐색의 모든 과정에 관련된 생각을 말로 표현하도록 (Thinking-aloud) 하였는데 이는 이용자의 각 탐색 방식에 대한 반응을 포착하기 위함이었다. Think-aloud 에 의해서 수집된 데이터는 내용 분석 (content analysis)과 평균 정확율, 그리고 반복 측정 분산 분석법 (Repeated Measure Analysis of Variance)에 의하여 분석되었다.

4. 연구 결과

세 가지 탐색 방식의 효율성 측정 및 비교는 네 가지 효율성 측정 기준에 의하여 이루어졌는데, 그 기준은 정확율, 검색된 적합 사진 수, 브라우징된 사진 수, 그리고 색인어에 대한 이용자의 동의 정도 (inter-rater reliability)이다. 정확율은 정보 검색의 적합성 판단에 널리 이용되어 온 판단 기준이며 다른 기준들 또한 사전 실험을 통하여 연구 결과에 연관성을 지닌 것으로 파악되어 이 연구의 수치 데이터로 분석에 이용되었다. 다음은 각 평가 기준에 의해 측정된 탐색 방식의 효율성에 대한 보고이다.

4.1. 정확율

실험 대상자들은 탐색을 진행하는 동안 탐색된 사진들의 적합성을 판단하고 그 이유를 설명하였다. 표 1은 실험 참여자들의 적합성 판단에 의한 평균 정확율을 보여 주고 있다.

<표 1> 평균 정확률

	비주얼 브라우징	키워드 탐색	혼합 방식	평균
평균 정확율 (N=25)	46.5	61.9	67.0	58.5

표에서 보는 바와 같이, 혼합 방식이 가장 효율적인 방식으로 나타났으며 그 다음은 키워드 탐색, 그리고 비주얼 브라우징이 가장 덜 효율적인 방식으로 드러났다. 평균 정확율의 차이에 있어서 키워드 탐색과 혼합 방식간에 차이는 이 방법들과 비주얼 브라우징과의 차이에 비해서 크지 않은 것으로 나타났다. 반복 측정 분산 분석법도 이와 비슷한 결과를 보여 주었다.

- 키워드 탐색과 비주얼 브라우징 간에 큰 차이 (a significant difference)가 있는 것으로 분석되었다 (df = 1, f = 8.615, sig. = .007, α = .05).
- 혼합 방식과 비주얼 브라우징 간에도 큰 차이가 있는 것으로 분석되었다 (df = 1, f = 17.740, sig. < .001, α = .05).

25 회 개개의 실험을 분석하면, 18 회 실험 (72%)에서 키워드 탐색이 비주얼 브라우징보다 정확율이 높은 결과를 나타냈으며, 21 회 실험 (84%)에서 혼합 방식이 비주얼 브라우징보다 높은 정확율을 나타내었다. 이 결과는 각각의 두 방식들이 왜 큰 차이가 있는 것으로 분석되었는지를 설명한다고 하겠다.

4.2. 검색된 적합 사진 수

사전 실험을 통하여 연구자에 의하여 관찰된 중요한 점의 하나는 검색된 적합 사진 수 자체가 효율성 측정의 좋은 판단 기준이 될 수 있다는 것이었다. 비주얼 브라우저가 정확율 면에서 다른 방식들에 비해 열등한 것으로 분석되었으나 데이터 분석 중 비주얼 브라우저가 다른 방식들에 비해서 검색된 적합한 사진 수에서는 뒤떨어지지 않는 것으로 나타났기 때문이었다. 다음의 표는 이러한 경향을 보여 주고 있다.

<표 2> 검색된 적합 사진 수

	비주얼 브라우저	키워드 탐색	혼합 방식
검색된 적합 사진 수 (N=25)	625	371	628

분석 결과는 혼합 방식이 가장 우수한 것으로, 그 다음으로 비주얼 브라우저이, 그리고 키워드 방식이 가장 덜 효과적인 방법으로 나타났다. 이 결과는 혼합 방식에 의해 검색된 적합 사진의 대다수 부분이 비주얼 브라우저에 의한 것임을 내포하고 있다. 이러한 결과는 또한 비주얼 브라우저가 키워드 탐색에 비해 많은 수의 적합 사진을 원하는 경우 더 효율적인 방법임을 암시하는 것이기도 하다. 비주얼 브라우저와 키워드 탐색만을 비교한다면, 데이터 분석 결과는 88%의 실험에서 비주얼 브라우저가 키워드 탐색보다 더 많은 적합 사진을 검색했음을 보여 주었다. 이러한 패턴은 키워드 탐색과 비주얼 브라우저 방법간의 반복 측정 분산 분석법에 의한 큰 차이를 나타내어 주고 있다 ($df = 1, f = 11.843, sig. = .002, \alpha = .05$).

4.3. 브라우즈된 사진 수

연구자는 실험을 통해 참여자들이 키워드 탐색보다 비주얼 브라우저 방법을 이용할 때 검색된 사진들을 더 빨리 브라우즈하는 것을 관찰할 수 있었다. 이 브라우즈된

사진 수는 정확율이나 검색된 적합 사진 수와 연관이 있는 것으로 간주되어 여기에 효율성 측정의 기준으로 선택되었다. 표 3 은 실험 참여자들이 검색된 사진들의 적합성을 판단하기 위해 브라우즈한 사진의 수를 보여 주고 있다.

<표 3> 브라우즈된 사진 수

	비주얼 브라우징	키워드 탐색	혼합 방식
브라우즈된 사진 수 (N=25)	1272	615	909

위의 표에서 볼 수 있는 바와 같이, 비주얼 브라우징 방법을 사용할 때 실험 참여자들은 키워드 탐색을 할 때보다 2 배 이상의 사진을 브라우즈한 것으로 나타났다. 대부분의 참여자들 (84%)이 키워드 탐색보다 비주얼 브라우징 방법을 이용할 때 더 많은 사진들을 보면서 적합성을 판단했다. 반복 측정 분산 분석법은 각각 세 가지 방식들 간의 차이점을 발견했는데 그 결과는 다음과 같다.

- 비주얼 브라우징 대 키워드 탐색 ($df = 1, f = 41.236, sig. < .001, \alpha = .05$),
- 키워드 대 혼합 방식 ($df = 1, f = 10.787, sig. = .003, \alpha = .05$),
- 비주얼 브라우징 대 혼합 방식 ($df = 1, f = 25.846, sig. < .001, \alpha = .05$).

이 결과는 비주얼 브라우징이 이용자들로 하여금 더 많은 사진들을 신속하게 보게 함으로써 더 많은 수의 적합한 사진을 검색할 수 있는 기회를 제공하고 있음을 나타내는 것이다.

4.4. 색인어와 이용자 질의어의 일치정도 (Inter-Rater Reliability)

시스템에 의해 제공되는 색인어와 실험 참여자들이 적합성 판단 시 색인어에 대한 동의 정도를 분석하여 이 두 집합 간의 일치 여부를 조사하였다. 분석 방법으로는

실험 참여자에 의하여 부적합한 것으로 판단되었으나 시스템에 의해서 매치되어 검색된 사진 수를 이용하였다.

<표 4> 색인어와 실험 참여자 질의어 사이의 불일치 정도

	색인어에 의해 검색되었으나 실험 참여자에 의하여 브라우저된 사진 중 부적합하다고 판정된 사진 수		
	비주얼 브라우저	키워드 탐색	혼합 방식
합계* (N=25)	317/1272 (25.2%)	221/615 (34.5%)	325/909 (36.0%)

* 분자 부분은 시스템에 의하여 검색되었으나 실험 대상자들에 의하여 부적합한 것으로 판정된 사진 수이고 분모 부분은 브라우저된 총 사진 수이다.

색인어에 대해서 실험 참여자가 동의하지 않은 사진의 수만을 고려한다면, 키워드 탐색이 불일치하는 부분이 가장 적은 것으로 나타났으며 (221), 비주얼 브라우저가 그 다음이고 (317), 혼합 방식이 가장 불일치가 많은 것으로 드러났다 (325). 그러나 이 불일치 정도는 실험 참여자들이 실제로 본 사진의 수와 함께 생각되어야 한다. 그 이유는 더 많은 사진을 볼수록 더 많은 불일치 부분을 발견할 수 있기 때문이다. 따라서 불일치 정도를 이 두 가지 수치의 비율로 나타냈을 때, 비주얼 브라우저가 가장 낮은 불일치 정도를 보여 주었고 (25.2%), 키워드 탐색이 그 다음으로 불일치 정도가 낮았으며 (34.5%), 혼합 방식이 가장 불일치가 높은 것으로 나타났다 (36.0%). 비주얼 브라우저 방법이 색인어와 실험 참여자 판단에 의해 가장 일치도면에서 안정적인 것으로 분석된 이유는 많은 수의 사진이 양쪽 그룹에 의해서 적합한 것으로 판단되었을 뿐만 아니라 많은 다른 사진들이 양쪽 그룹에 의해서 부적합하다고 동의되었기 때문이다. 적합 사진에 대한 일치도 뿐만 아니라 부적합 사진에 대한

일치도 역시 비주얼 브라우징이 일치도 면에서 가장 안정적인 방법으로 나타난 중요한 원인이 되었다.

일반적으로 25%에서 36%의 색인어와 실험 참여자 판단간에 불일치 정도가 발견되었는데 이러한 inter-rater inconsistency 의 문제는 불가피하다고 하겠다. 이것은 왜 많은 이미지 검색 분야의 연구가 색인 작업의 어려움과 색인의 일관성의 중요성을 강조하는 가를 뒷받침하는 것이라 하겠다.

5. 토의 및 결론

연구 결과를 볼 때, 탐색 방법의 효율성은 그 측정방법에 따라 다른 것으로 드러났다. 표 5 에서 각기 다른 효율성 측정 기준에 의한 탐색 방법의 효율성을 요약하였다.

<표 5> 각 효율성 측정 기준에 의한 탐색 방법의 효율성

측정기준	효율성*		
	비주얼 브라우징	키워드 탐색	혼합 방식
1. 정확률	3	2	1
2. 검색된 적합 사진 수	2	3	1
3. 브라우즈된 사진 수	1	3	2
4. 색인어와 이용자 판단 일치도	1	2	3

* 숫자는 어느 방법이 가장 효율적이냐 (1), 다음으로 가장 효율적인가 (2), 혹은 가장 덜 효율적인가 (3)를 나타낸다.

위의 표에서 보여지는 바와 같이, 탐색 방식의 효율성은 이용자의 정보 요구 형태에 따라 달라질 수 있다. 즉 이 연구 결과로 밝혀진 유의한 점 중의 하나는 탐색 방식의 효율성이 처해진 상황에 따라 달리 해석될 수 있다는 것이다 (the nature of

conditional-dependence of effectiveness of different access modes). 이 연구는 문자를 이용한 이미지 검색의 문제점을 인식하며 그 대안책으로 비주얼 브라우징이 대부분의 경우 이러한 문제점을 해결할 수 있을 것이라는 기대에서 출발하였다. 그러나, 실험 데이터를 분석한 결과, 하나의 탐색 방법이 다른 방법보다 항상 우월한 것이 아니라는 것을 밝혀낼 수 있었다. 탐색 방법의 효율성은 보다 복잡한 유형을 나타내고 있으며 그 측정 기준에 따라 달라질 수 있다는 점이다. 이런 효율성이 상황에 따라 다르다는 점은 이용자 요구의 이해가 중요하다는 것을 강조하는 것이다. 예를 들어, 이용자의 정보 요구가 많은 수의 광범위한 사진을 검색하기를 원하고 그에 따르는 시간이나 노력을 들일 수 있는 경우라면 비주얼 브라우징이나 혼합 방식이 적합한 탐색 방법이 될 수 있을 것이다. 그러나, 이용자가 가능하면 적은 시간과 비용을 들이며 탐색의 효율성 (efficiency)을 강조하고자 한다면 키워드 탐색이 적합하다 할 것이다.

또 다른 이 연구 결과의 의의는 기존의 문자를 이용한 이미지 검색의 효과를 넓히면서 동시에 이용자를 위한 탐색의 유연성 (flexibility)과 다양성 (variety)을 제공하는 방법을 살펴 보았다는 점이다. 이 연구에서 제시된, 색인어의 공유에 의한, 비주얼 브라우징 방식은 프로그래밍이나 시스템 구현 면에서 많은 추가 작업을 요하지 않으면서 이용자의 성공적인 검색 면에서는 그 효과가 크다고 하겠다. 실험에 참가한 많은 대상자들이 비주얼 브라우징을 이용할 때의 장점으로 탐색어 선택의 부담에서 자유로움을 언급했으며 시각 정보를 직접 이용하는 것이 이미지에 의하여 전달되는 더 풍부한 정보를 이용하는 데 도움이 된다고 언급하였다.

이 연구는 이미지 검색을 위한 탐색 방법의 하나로 비주얼 브라우징 기법을 제시하였으며 이 방법이 이미지 검색에 이용될 수 있는 긍정적인 가능성 (positive potential)을 보여 주었다. 다음 단계로는 이 비주얼 브라우징 방식이 이미지 검색에 있어서 일반화 (generalizability)될 수 있는가를 연구하는 것이라 하겠다. 다양한 종류의 이미지 데이터와 색인 방식을 이용한 이미지 데이터베이스의 검색에 비주얼 브라우징 방법을 이용함으로써 이 방식의 이용 가능성의 확대를 진단하고자 하는 것이 이 연구의 확장 계획이다.

참고 문헌

- [1] Zheng, M. Metadata Elements for Object Description and Representation: A Case Report from a Digitized Historical Fashion Collection Project. *Journal of the American Society for Information Science*. 50(13): 1193-1208, 1999.
- [2] Rasmussen E. Indexing Images. *Annual Review of Information Science and Technology*, 32: 169-196, 1997.
- [3] Hastings, S. Query Categories in a Study of Intellectual Access to Digitized Art Images. *Proceedings of the Annual Meeting of the American Society for Information Science*, 32: 3-8, 1995.
- [4] Keister, L. User Types and Queries: Impact on Image Access Systems. In *Challenges in Indexing Electronic Text and Images* (Fidel, R. et al., Eds.) Medford NJ: Learned Information, Inc., 1994.
- [5] Enser, P. Query Analysis in a Visual Information Retrieval Context. *Journal of Document and Text Management*, 1(1): 25-52, 1993.
- [6] Idris, F. & Panchanathan, S. Review of Image and Video Indexing Techniques. *Journal of Visual Communication and Image Representation* 8(2): 146-166, 1997.
- [7] Aigrain, P. et al. Content-Based Representation and Retrieval of Visual Media: A State-of-the-Art Review. *Multimedia Tools and Applications* 3(3): 179-202, 1996.
- [8] Lynch, Clifford A. The Technologies of Electronic Imaging. *Journal of the American Society for Information Science* 42(8): 578-585, 1991.
- [9] Besser, Howard. Imaging: Fine Arts. *Journal of the American Society for information Science* 42(8): 589-596, 1991.
- [10] Lunin, Lois F. Image Databases: An Overview. *Proceedings of the 12th National Online Meeting*: 215-219, 1991.

- [11] Rorvig, Gary A. Expanding Archives: Library of Congress Is Not Just Books. *The New York Times*: 11(L), 14(N), September 3, 1990.
- [12] Goodrum, A.A. Image Information Retrieval: An Overview of Current Research. *Journal of Informing Science* 3(2): 63-67, 2000.
- [13] Markey, Karen. Access to Iconographical Research Collections. *Library Trends* 37(2): 154-174, 1988.