

웨이브렛 변환영역의 정보를 신경망 학습 통한 영상검색에 관한 연구

최병도^o 조영 박장한 남궁재찬
광운대학교 컴퓨터공학과

cbd@center.duksung.ac.kr cake0709@hanmail.net (parkch, namjc)@gwu.ac.kr

A Study on Image Search for Neural Network learning to Information of Wavelet Transform region

Byoung-do Choi^o, Young Cho, Chang-han Park, Jae-chan Namkung
Dept. of Computer Engineering, KwangWoon University

요 약

본 논문에서는 웨이브렛 변환 영역의 정보를 신경망 학습을 통하여 영상검색에 관한 연구를 제안하였다. 영상검색은 연구가 이루어지고 있지만, 영상의 특징을 정확하게 표현한다는 것은 현실적으로 어렵기 때문에 영상의 저장 및 검색에 많은 어려움이 있다. 따라서 영상데이터의 효율적인 저장 및 검색을 위해서는 공간 영역보다는 변환 영역에서의 특징추출 방법이 고려되어야 한다.

본 논문에서는 웨이브렛 변환 후 생성되는 저주파 대역의 영상을 일정한 크기로 ($2^n \times 2^n$) 분할한 다음 각 블록의 표준편차를 구하고, 주어진 경계 값을 기준으로 작성된 블록 맵을 유사성의 척도로 이용하여 유사한 영상을 함께 모아 카테고리 분류에 의한 저장을 한다. 또한 질의영상에 대한 블록 맵을 신경망 학습을 통해 해당 카테고리리를 찾아 1:1매칭을 통한 검색을 함으로써 검색 시간을 줄이고, 제안된 시스템 효율을 증대 시킬 수 있었다.

1. 서 론

컴퓨터의 빠른 보급과 인터넷 이용의 급변한 구조 속에서 다양한 멀티미디어 데이터는 빠르고 효율적인 검색 기술을 요구하고 있다. 특히 영상검색 기술에 대하여 많은 연구가 이루어지고 있으나 영상을 저장하거나 검색하는데 많은 어려움이 있다. 이는 멀티미디어 데이터를 중에서 영상은 시각적이며 공간적인 정보를 포함하는 자료로서 멀티미디어 응용에서 가장 많이 사용되고 있다. 또한 다른 멀티미디어 자료와 마찬가지로 대용량 저장과 출력 방식의 다양성, 공간관계 표현, 비정형이라는 특징을 가지고 있기 때문에 영상의 특징을 한마디로 정확하게 표현한다는 것이 현실적으로 어렵기 때문이다[7][8]. 이러한 멀티미디어 데이터의 검색 기법에는 주석 기반 검색 기법과 내용기반 검색 기법이 있다[7]. 최근 멀티미디어 데이터에 컴퓨터 비전이나 이미지 처리 및 컴퓨터 그래픽스와 같은 분야들의 연구결과로 보다 정확한 정보를 추출하는데 도움이 되고 있다. 일반적으로 내용기반 검색에서는 완전 일치가 아닌 어느 정도의 오차를 가질 수 있는 유사 일치 검색을 주로 사용한다. 따라서 검색의 정확도는 질의 데이터와 검색 데이터의 유사율을 구하는 문제로 축약될 수 있다[4].

본 논문의 구성은 다음과 같다. 2장에서는 영상검색 기법과 웨이브렛 변환 및 신경망과 관련된 연구들을 서술하였으며, 3장에서는 본 논문에서 제안하는 시스템의 특징과 웨이브렛 변환 및 신경망이 이용한 알고리즘을 서술하였다. 4장에서는 실험 결과 및 고찰에 대하여 서술하였고, 마지막으로 5장에서는 결론 및 향후 연구방향을 제시한다.

2. 관련연구

2.1 데이터베이스 검색

영상의 내용을 기반으로 하는 검색은 전통적인 데이터베이스 검색과 비교하여 차이점을 보면, 첫째 내용 기반 검색은 유사성에 근거한 검색이다. 둘째, 사용자와 시스템과의 상호작용을 통해 다양한 검색 가능성을 제공한다. 셋째, 시각적인 질의를

필요로 하는 멀티미디어 시나리오와 같은 응용단계에서 사용될 수 있다. 넷째, 사용자는 유사도에 따른 후보영상을 최종 사용자가 선택할 수 있도록 여러 장의 영상들을 아이콘화 하여 표현한다. 영상 자체를 근거로 한 내용기반 검색 시스템은 기존의 텍스트 검색에 부가적인 역할을 해줄 수도 있고 멀티미디어 정보 제공자, 내용 저작자, 최종 사용자와 연계되어 검색을 도와줄 수도 있다. 이러한 영상 데이터베이스 특징 중 유사성 검색은 주어진 표본과의 유사한 정보를 나타내는 유사성 척도에 의해 검색하는 것으로서 유사성 척도는 유사성 검색의 성능과 밀접한 관련이 있으며 문제 영역의 특성에 따라 다양한 척도를 가질 수 있다.

2.2 웨이브렛 변환

컴퓨터 비전에서 이용된 다해상도(multi-resolution) 분석 방법이나 음성과 영상 압축에서 사용되던 부 대역(sub-band) 코딩 기법과 응용 수학에서 사용된 웨이브렛 시리즈 전개 등 많은 기본 기법들이 최근에 들어 웨이브렛 이론의 특수한 응용으로 밝혀졌다. 웨이브렛 해석은 연속 신호와 이산 신호의 경우에 모두 적용될 수 있으며, 다양한 분야에서 그 응용 가능성을 인정받고 있다. 웨이브렛 변환은 비정형(non-stationary) 신호의 분석에 유리한 특징을 가져서 고전적인 단구간 푸리에 변환(STFT : Short Time Fourier Transform)이나 가보 변환(gabor transform)을 대체할 새로운 대안으로 대두되고 있다. 웨이브렛 변환이 고전적인 단구간 푸리에 변환과 구별되는 근본적인 차이점은 단구간 푸리에 변환의 경우 모든 주파수 대역에 대하여 동일한 크기의 필터 윈도우를 사용하는 반면, 웨이브렛 변환은 고주파 대역에서는 폭이 좁은 윈도우, 저주파 대역에서는 폭이 넓은 윈도우를 사용한다는 것이다. 웨이브렛 변환은 입력 신호를 특정 기저 함수의 집합으로 분리하는 과정으로도 이해될 수 있다. 웨이브렛 변환에 사용되는 기저 함수의 집합은 하나의 기본 웨이브렛 기저 함수에 대한 시간 축 방향으로의 확대 및 축소 그리고 평행 이동을 통해 얻어진다[6]. 웨이브렛 변환의 원리는 부대역 부호화를 하는 것이다. 이것은

영상을 필터에 통과시켜 영상의 집합으로 만드는 것을 말한다. 각각의 부 영상은 공간적 주파수의 제한된 구간을 포함하고 있으며, 이들 각각의 영상 집합을 부대역이라고 부른다. 이렇게 변환된 부대역은 원 영상과 비교하였을 때, 대역폭이 줄어들게 되므로, 이들을 다운 샘플 할 수 있다. 부대역 부호화의 구성을 살펴보면 분석단계, 코딩단계, 합성단계로 나눌 수 있다. 분석 단계는 필터링과 다운 샘플링을 통해서 원 영상을 여러 개의 부대역으로 나누는 과정이다. 코딩 단계는 각각의 부대역 특성에 따라 한가지, 또는 여러 가지 코딩 기법을 적용하여 코딩한다. 마지막 합성단계는 업 샘플링과 필터링을 거쳐 부대역을 합침으로써 영상을 복원하는 부분이다[1][2][4][5].

2.3 신경망

신경망 회로에 의한 학습(learning)은 학습 규칙이라 부르는 가중치 조정방식에 의하여 이루어진다. 신경망은 학습에 의하여 입력 패턴을 완성하거나, 입력 패턴을 보류하거나 입력함수를 계산하도록 하여 반복 훈련시키는 방법으로 이루어진다. 학습이 되기 전에 신경망 회로상에서 가중치들은 알맞지 않은 값으로 시작하지만, 학습과정이 끝나면 가중치들이 조정되어 상기한 기능들을 잘 수행하는 일면 가중치의 조정도 더 이상 일어나지 않는다[11]. 단층 신경망은 가장 단순한 구조로서 외부 입력을 받아들이는 입력층X와 처리된 결과를 출력하는 출력층Y로 구성된다. 다층 신경망은 여러 계층으로 구성된 신경망 구조이다. 일반적으로 가장 널리 사용되는 것은 3계층 구조이며, 입력층X(외부 입력을 받아들이는 계층), 출력층Y(처리된 결과가 출력되는 계층), 은닉층Z(입력층과 출력층 사이에 위치하여 외부로 나타나지 않는 계층)이 있다. 3계층 신경망 구조에서는 입력층의 입력에 따라 은닉층의 출력이 나오며, 은닉층의 출력은 다시 출력층에 입력되어 최종 출력이 나오게 된다[13].

3. 제안된 알고리즘 설계

3.1 제안된 시스템의 전체 구성도

본 논문에서 제안한 영상검색 시스템은 웨이브렛 변환 영역에서 저주파 대역의 영상을 블록맵으로 처리하고, 각 블록맵의 가중치 값을 신경망에 학습시켜 카테고리 별로 저장 및 검색을 하는 시스템이다. [그림1]은 제안된 영상검색 시스템의 전체 구성도를 나타낸 것이다.

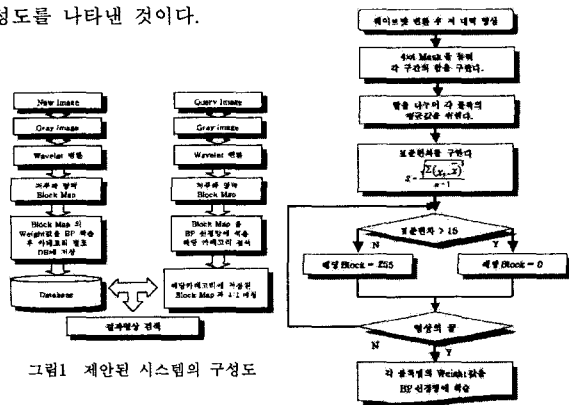


그림1 제안된 시스템의 구성도

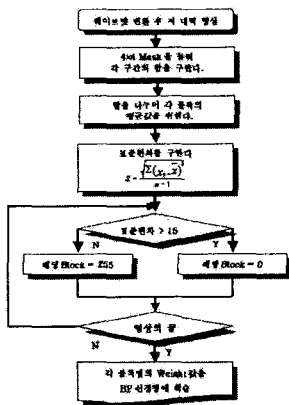


그림3 Block Map의 구성도

3.2 웨이브렛 변환과 블록 맵의 추출

본 논문에서는 웨이브렛 변환후, 저대역 영상은 영상의 의미 있는 정보를 제공할 뿐 아니라 시각적으로 매우 중요한 역할을 한다. 웨이브렛의 변환은 2~5단계가 보편적으로 사용되지만, 본 논문에서는 저대역 영상을 4x4마스크로 블록맵을 만드는 점을 고려하여 1단계 웨이브렛 변환을 하였다. 2차원 영상에서

의 웨이브렛 분해 과정을 보면 각 단계마다 원영상이 4개의 구역으로 나뉜다. [그림2]는 실험영상을 1단계로 웨이브렛 변환한 그림이며, 원영상이 한 단계 웨이브렛 변환될 때, 그림의 가로, 세로가 각각 1/2로 LL(저주파성분), LH(수평성분), HL(수직성분), HH(대각선성분)의 다운 샘플링 되는 것을 알 수 있다 [3]. 웨이브렛 변환후 저대역 영상(128x128)을 (4x4)블록으로 마스크시킨 다음 각 블록에 대한 표준편차를 구했다. 표준편차는 각 블록의 특성을 나타낸 것으로, 표준편차 값 15를 기준으로 블록맵을 생성한 다음 BP(Back Propagation)신경망 학습을 통해 각각 가중치값을 데이터베이스에 저장하게 된다. [그림2]는 Gray영상의 웨이브렛 변환 및 블록맵을 나타낸 것이다.

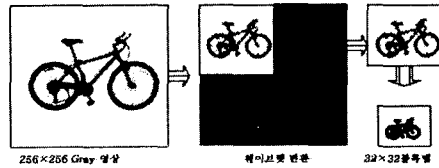


그림2 Gray 영상의 웨이브렛 변환 및 블록 맵

영상을 작은 블록단위로 영역 분류하는 알고리즘은 백터 양자화 프랙탈 압축 등 많은 영상 압축 알고리즘의 전처리 단계로 활용된다. 또한 블록단위 영상분류는 블록의 정확한 영역 분류를 통하여 압축효율을 높이고 탐색 알고리즘의 탐색 범위를 줄여 알고리즘 수행 시간을 단축시키는 데 큰 역할을 한다. 블록단위의 영상분류 방법은 크게 두 가지로 나눌 수 있다. 첫째는 공간영역의 화소 밝기 값을 사용하여 블록의 특성을 분류하는 방법이 있고, 둘째는 주파수 영역의 변환 계수를 사용하여 블록의 특성을 분류하는 방법이 있다[9]. [그림3]은 블록 맵의 구성도를 나타낸 것이다.

본 논문에서는 각 블록의 표준편차 값에 의해 만들어진 블록맵을 BP신경망을 이용한 학습을 통해 해당 카테고리를 찾고, 1:1매칭을 통해 영상을 검색한다. 오류 역전파 알고리즘이라고도 부르는 BP 알고리즘은 순방향 다층 신경망의 학습에 효과적으로 사용할 수 있어서 다양한 분야에 가장 널리 활용되는 학습 알고리즘이다. BP 알고리즘은 출력층의 오차 신호를 이용하여 은닉층과 출력층 간의 연결 강도를 변경하고, 출력층의 오차 신호를 은닉층에 역전파 하여 입력층과 은닉층과의 연결 강도를 변경하는 학습 방법이다. BP 알고리즘을 이용한 신경망의 학습은 크게 3단계로 진행된다. 단계1(학습 패턴을 입력하여 출력을 구함), 단계2(출력과 목표치의 차이(오차)를 구함), 단계3(오차값을 역방향으로 전파시키면서 출력층의 연결 강도 및 은닉층의 연결 강도를 변경)[10].

3.4 질의 영상 검색

본 논문에서 영상검색 방법은 입력된 질의 영상 역시 영상저장과 같은 과정을 통해 Block Map으로 처리한 다음 신경망을 통해 Category별로 저장되어 있는 데이터베이스에서 해당 Category를 찾은 다음 저장된 Block Map과 1:1매칭을 통해서 유사한 영상을 검색하는 시스템이다. [그림5]는 질의영상 구성도를 나타낸 것이다.

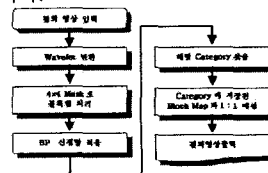


그림4 질의영상검색 구성도

4. 제안된 시스템의 구현 및 비교 평가

[그림5]는 블록 맵된 영상 중에서 각각의 카테고리의 대표패턴을 가지고 BP 알고리즘을 이용하여 학습하는 프로그램이다. 블록 맵된 영상은 32×32크기를 가지므로, 입력노드의 개수는 1024개를 가진다. 질의 영상을 입력하였을 때 블록 맵을 추출하고, 신경망학습을 통해 Data base에 저장되어 있는 해당 카테고리를 찾아 1:1매칭을 통한 영상검색을 한다. [그림6]은 Airplane을 검색하였을 때, Airplane Category의 1번째 Directory에 있음을 보여준다.

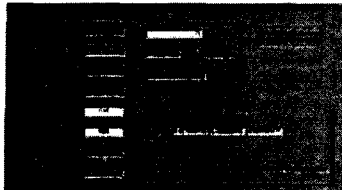


그림5 Back Propagation 학습기

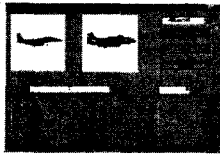


그림6 영상검색 시스템에서 결과를 찾았을 경우

4.3.1 해당 Category 검색을 비교

추출된 블록맵을 신경망에 학습 시켰을 때, 카테고리의 검색을 비교에는 Airplane, Washer, Car 등 3종류의 정형화된 영상을 각각 10개씩 총30개의 영상으로 표준편차 값이 10, 15, 30이상일 때, 카테고리 검색율을 비교한 결과 표준편차 값이 10이나 15이상일 때는 해당 카테고리 검색율이 평균 80%를 보였으나, 표준편차 값이 30이상일 때는 평균 40%의 낮은 검색율을 보였다. [표1]은 신경망학습을 통한 해당 카테고리 검색 확률을 나타낸다.

표1 카테고리 분류를 비교

실험영상	표준편차	표준편차>10	표준편차>15	표준편차>30
Airplane image		100%	100%	80%
Washer image		80%	80%	20%
Car image		60%	60%	20%

4.3.2 검색 시간의 비교

영상 데이터를 신경망을 사용한 패턴매칭 검색방법과 신경망을 사용하지 않는 패턴매칭 검색 시간을 비교하였다. 영상데이터 개수가 90개 이하일 때는 신경망과 패턴 매칭을 사용한 알고리즘과 패턴 매칭만을 사용한 알고리즘 둘 다 비슷한 결과가 나왔으나 300개 이상일 때부터 검색시간의 차이가 나타났다. 1500개 이상의 데이터에서는 배 이상의 차이가 났으며, 신경망을 이용한 카테고리 검색 방법이 좀더 빠르다는 것을 알 수 있었다. 그러나 최악의 경우, 카테고리 검색에 실패를 한 경우에는 패턴매칭만 사용한 경우와 동일한 검색시간을 가진다. [표2]는 검색시간 비교를 나타낸다

표2 검색 시간 비교

데이터 개수	신경망과 패턴매칭 사용시	패턴매칭만 사용시
15	0~1(초)	0~1(초)
90	0~1(초)	0~1(초)
300	1(초)	1~2(초)
1500	1~2(초)	2~6(초)

5. 결론

본 논문에서는 Wavelet변환 영역에서 추출된 블록 맵의 특징 값을 신경망에 학습시켜 해당 카테고리를 찾고1:1패턴 매칭을 통한 검색 시스템을 제안하였다. 제안된 시스템은 영상데이터의 효율적인 저장 및 검색 시간의 단축을 위해 웨이블릿 변환 후 영상 데이터를 저장하였으며, 검색에 사용될 특징추출 방법으로는 변환영역에서 중요 정보가 가장 많은 저주파 대역

의 영상을 일정한 블록으로 나누고 각 블록의 표준편차 값을 기준으로 작성된 블록 맵을 영상 검색의 척도로 이용하였다. 기존의 특징추출 방식(컬러, 질감, 모양)보다 Block Map을 이용함으로써 특징추출 방법을 간략화 하였다. 웨이블릿 변환을 이용하면, 영상 또는 신경망을 이용하지 않은 기존의 검색 방법은 친구간의 패턴 매칭으로 검색 시간이 오래 걸리지만, 제안된 신경망을 이용한 카테고리분류에 의한 저장 및 검색 방법은 해당 카테고리별로 검색함으로써 검색시간을 줄이고, 시스템 효율을 증대시킬 수 있음을 입증하였다. 본 논문에서는 정규화된 영상에 제한을 두었으나, 향후 연구과제로는 좀더 향상된 블록 맵 추출 알고리즘의 제안하여 적용한다면, 카테고리 분류율과 데이터 저장에 있어서도 효율적인 성능 향상을 기대할 수 있다.

참고문헌

[1] A. S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," IEEE Trans. on Image Processing, vol.1, pp. 244-250, Apr. 1992.
 [2] S. G. Mallat, "A theory for multiresolution signal decomposition the wavelet representation," IEEE Trans. Patt. Anal. Mach. Intell., Vol. 11, pp. 674-693, July, 1989
 [3] S. Kim, S. Rhee, J. Jeon, and K. Park, "Interframe Coding Using Two-Stage Variable Block-Size Multiresolution Motion Estimation and Wavelet Decomposition," IEEE Trans. Circuits and Systems for Video Technology, Vol. 8, pp. 399-410, Aug. 1998.
 [4] ISO/IEC CD15444-1, "JPEG2000 Image Coding System," Ver 1.0, Dec. 1999.
 [5] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Addison Wesley, 1992.
 [6] <http://www.intersys21.com/product/idl/wavelet.html>
 [7] 이병구, 김시우, 정진환, "멀티미디어 데이터베이스를 위한 영역지식을 이용한 내용기반 검색", 한국정보과학회 가을 학술 발표논문집 Vol. 24, No.2.
 [8] 정세환, 유현우, 장동식, "VQ를 이용한 영상의 객체 특징 추출과 이를 이용한 내용기반 영상검색", 한국정보과학회 가을 논문집 Vol, 26, No. 2, pp359-361, 1999
 [9] 박동권, "블록단위 영상분할과 프렉탈 영상 부호화", 동국대학교 대학원 석사학위논문, 1997
 [10] 오창석, 뉴로컴퓨터개론, 충북대학교 컴퓨터공학과, pp 20-61, 2000.
 [11] 김희승, 영상인식. 생능출판사, 1998
 [12] 신동규, 신동일, 멀티미디어 데이터베이스, 교보문고, 2000.