

영상검색에서의 다중 피쳐 결합을 위한 변형된 보다 카운트 방법

정세윤, 김규현, 전병태, 이재연, 배영래

한국전자통신연구원

Tel: 042-860-5724/ Fax: 042-860-4844

Modified Borda Count Method for Combining Multiple Features of Image Retrieval

Se Yoon Jeong, Kyuheon Kim, Byung Tae Chun, Jae Yeon Lee and Younglae J. Bae.
Electronics and Telecommunications Research Institute.

E-mail: jsy@etri.re.kr

Abstract

In this paper, we propose an image retrieval system using the MBCM(Modified Borda Count method) in CME(Combining Multiple Experts). It combines color-, shape- and texture-based retrieval sub-systems. CME method can complementarily combine results of each retrieval system, which uses different features. There are some problems when the Borda count method in pattern recognition is applied to image retrieval. Thus, we propose a modified Borda count method to solve these problems. In the experiment, our method reduces false positive errors and produces better results than that of each retrieval module that uses only one feature.

I. 서론

내용기반 영상 검색은 멀티미디어 데이터 베이스의 핵심 기술 중에 하나이다. 현재 세계적으로 관련 연구가 활발히 진행되고 있으며, 국제 표준화 작업도 MPEG 위원회의 MPEG-7에서 진행되고 있다. 내용기반 영상 검색은 영상의 내용을 정량적인 값으로 표현해서 색인하고 검색하는 방식이다. 이 정량적인 값을 피쳐라고 한다. 주로 사용되는 피쳐는 컬러(color), 질감(Texture), 모양(Shape)이다. 대부분의 영상검색에 관한 연구들은 이들 중 한 가지 피쳐만을 사용하고 있다. 한가지 피쳐만을 사용하여 만족할 만한 검색 결과를 얻

는 것은 무척 어려운 일이다. 다행히 각 피쳐 검색 결과에는 상호 보완적인 성격이 있다. 예로 컬러 검색결과가 안 좋았는데, 질감 검색에서는 좋은 검색 결과가 나오는 경우, 또는 모양 검색 결과는 안 좋았는데, 컬러 검색결과 좋은 경우이다. 이러한 경우에 각 피쳐의 검색 결과를 종합하여 더 좋은 검색 결과를 얻을 수 있다. 본 논문에서는 이러한 각 피쳐 검색 방법을 효과적으로 결합하여 보다 좋은 검색 결과를 얻을 수 있는 다중 피쳐 검색 방법을 제안한다. 이를 위해서 패턴 인식에서 사용되는 CME 개념에 기반한 다중 분류기 결합 방법을 참고하였다. 패턴 인식에서 사용되는 다중 분류기 결합 방법 중에서도 영상 검색에 적용하기 적합한 방법인 Borda Count 방법을 적용하였다. 하지만, Borda Count 방법 그대로를 영상 검색에 적용하기에는 몇 가지 문제점이 있다. 이 문제점을 해결하기 위해 본 논문에서는 변형된 Borda Count 방법을 제안하였다.

II. BORDA COUNT METHOD

CME 개념에 근거한 다중 분류기 결합 방법의 대표적인 것에는 행위 지식 공간(BKS: Behavior Knowledge Space)을 이용한 방법, 다득표(majority voting) 방법, Borda Count 방법, 논리적 회귀분석을 이용한 가중화된 Borda Count 방법, 선형 확신도 누적(LCA: Linear Confidence Accumulation)방법, 퍼지융합(fuzzy fusion) 방법, 신경망을 이용한 방법,

퍼지 행위지식 공간 방법, 다단계 결합 방법등이 있다[1].

이들 방법 중에 보다 카운트 방법은 영상 검색에 적용 하기가 용이하고 구현하기 쉬워 본 논문에서는 보다 카운트 방법을 사용하여 각 영상 검색기의 결과를 결합하였다.

보다 카운트 방법은 각 분류기가 순서정보를 제공할 때 적합한 방식이다. 영상 검색기에서는 질의영상에 대해 유사한 영상들을 순서적으로 출력하므로 보다 카운트 방법을 적용할 수 있다.

수식 (1)은 일반적이 CME 정의식이다.

$$\left. \begin{aligned} e^1(x) &= C_0^1, \dots, C_{M-1}^1 \\ e^2(x) &= C_0^2, \dots, C_{M-1}^2 \\ &\vdots \\ e^K(x) &= C_0^K, \dots, C_{M-1}^K \end{aligned} \right\} \rightarrow CME(x) = y \quad (1)$$

여기서, $e^k(x)$ 는 입력 패턴 x 에 대한 k 번째 분류기의 출력이고, 수식 (1)의 의미는 CME가 k 개의 출력기의 출력을 종합하여 입력 패턴 x 에 대해 최종적으로 y 클래스라고 판단한다는 의미이다.

일반적인 CME 정의식인 수식(1)의 출력 부분만 각 결합 방법의 출력식으로 바꾸면 각 결합 방법들의 정의식이 된다.

$$\Rightarrow CME(x) = j, \quad \text{if } B_j = \max_{1 \leq i \leq M} \{B_i\} \quad (2)$$

수식 (2)는 보다 카운트 방법의 출력식이다. 보다 카운트 방법은 입력 패턴 x 에 대해 각 클래스의 보다 카운트를 구하고, 보다 카운트가 가장 큰 j 클래스를 결과로 출력하는 방법이다.

$$B_i = \sum_{k=1}^K (w^k \cdot G^k(i)) \quad (3)$$

수식 (3)은 보다 카운트 계산식이다. 각 분류기의 i 클래스의 G 값을 합해 보다 카운트 값을 구한다. 여기서 가중치는 각 분류기에 동일한 값을 사용할 수도 있고, 다른 값을 사용할 수도 있다.

$$G^k(i) = M - r, \text{ if } C_r^k = i \quad (4)$$

수식 (4)에서 M 값은 전체 클래스의 개수이고, C_r^k 의 의미는 i 클래스가 k 번째 분류기의 출력에서 순위가 r 이라는 것을 의미한다.

III. MBCM(Modified Borda Count Method)

보다 카운트 방법을 영상 검색에 그대로 적용하기에는 몇 가지 문제점이 있다. 본 절에서는 이러한 문제점들과 그 해결책에 대해서 설명한다.

1. 결과 출력 문제

패턴 인식에서 사용되던 보다 카운트 방법을 영상 검색에 적용하는데 있어서 첫번째 문제점은 결과 출력 부분이다. 패턴 인식에서는 입력 패턴과 가장 유사한 클래스 하나만 결과로 출력을 하고, 영상 검색에서는 데이터 베이스 내에서 입력질의 영상과 유사한 영상 모두를 결과로 출력한다. 따라서, 보다 카운트 방법의 결과 출력부분을 변경하여야 한다.

수식(5)는 변경된 결과 출력식이다.

$$CME(x) = C_0^{cme}, C_1^{cme}, \dots, C_{M-1}^{cme}, \quad (5)$$

By descending order of B_i

2. M 값

패턴 인식의 보다 카운트 방법에서 M 값의 정의는 전체 클래스의 개수였다. 이 정의를 영상 인식에 적용하기에는 문제점이 있다. 이는 영상 데이터 베이스의 크기가 계속 증가하기 때문이다. 또한, 이 값은 매우 큰 값을 가지므로, 보다 카운트 값은 크게 되고, 따라서, 약간의 차이는 의미가 없게 되어 전체 검색기의 분별력이 떨어지게 된다. 영상 검색에 적용하기 위해서는 새로운 M 값의 정의가 필요하다. 수식 (1)을 보면 M 값의 다른 의미가 각 분류기의 결과 개수를 의미한다는 것을 알 수 있다. 이 의미를 기반으로 영상 검색에 적합한 M 식을 정의하면 수식 (6)와 같다.

$$M = \max\{N(e_1(x)), N(e_2(x)), \dots, N(e_k(x))\} \quad (6)$$

여기서 $N(e_k(x))$ 는 k 번째 검색기의 검색 결과 개수이고, 영상 검색에서 질의 영상에 따라 각 검색기의 검색 결과 개수가 다르므로 이 값들 중에서 최대값을 M 값으로 정하였다.

3. Borda count 계산상의 문제

보다 카운트 방법에서 문제점 중 하나는 가장 열악한 검색기의 결과가 전체 결과에 가장 큰 영향을 미친다는 것이다. 열악한 검색기의 결과에는

대부분 false positive 에러가 포함되어 있다. 이러한, false positive error 의 결과에서의 순위는 검색기 마다 크게 다르다. 이는 열악한 검색기에서는 false positive error 이기 때문에 순위가 낮지만 다른 검색기에서는 정상적인 검색 결과라 순위가 높기 때문이다. 이러한 성질을 고려한다면 결과를 종합할 때 false positive error 의 영향을 줄일 수 있다.

$$B'_i = \prod_{k=1}^K (w^k \cdot G^k(i)) \quad (7)$$

이러한 특성을 고려하여 수식(3)을 수식 (7)로 수정 하였다. 수식(7)에서는 전체 보다 카운트를 합할 때 덧셈 대신에 곱셈을 사용하고 있다. 이는 산술 평균대신 기하 평균을 사용한다는 의미이다.

$$\sqrt{ab} \leq \frac{a+b}{2} \quad (8)$$

산술 기하 평균 부등식에서 등호는 a와 b가 같은 값을 가질 때 성립한다. 즉, 산술 평균값이 같은 경우라도 a와 b가 서로 비슷한 경우의 기하 평균값이 크다는 것이다.

즉, 수식 (3)의 보다 카운트가 같더라도 수식(7)을 이용해 구한 보다 카운트 값은 각 검색기의 결과에서의 순위가 유사한 경우가 큰 값을 갖게 된다. false positive error 는 검색기의 결과에서의 순위 변동이 크므로 상대적으로 작은 값을 갖게 되어 전체 결과에서 미치는 영향이 줄어들게 된다.

이처럼 보다 카운트를 계산할 때 산술평균대신 기하 평균을 사용하면 false positive error 를 줄일 수 있다.

4. 각각의 검색기 결과에서의 빈도 고려

전체 검색 결과 영상 중에서 각 검색기 결과마다 포함된 영상은 질의 영상과 유사할 확률이 매우 높고 상대적으로 한 검색기 결과에만 포함된 영상은 질의 영상과 유사할 확률이 낮다. 즉, 여러 검색기의 결과에 포함된 영상일수록 질의 영상과 유사할 확률이 높다. 이러한 성질을 고려하기 위하여 보다 카운트를 계산할 때 가산점을 더하였다.

$$B'_i = \prod_{k=1}^K (w^k \cdot G^k(i)) + M^{f-1} \quad (9)$$

여기서 f 는 빈도수로 i 번째 영상이 몇 개의 검색

기의 결과에 나타났는지를 의미한다.

즉, 식 (9)에서 가산점은 결과에 포함된 빈도에 의해 결정되는 것을 알 수 있다.

IV. 검색 시스템

그림 1은 MBCM 을 사용한 검색 시스템의 블록 도이다.

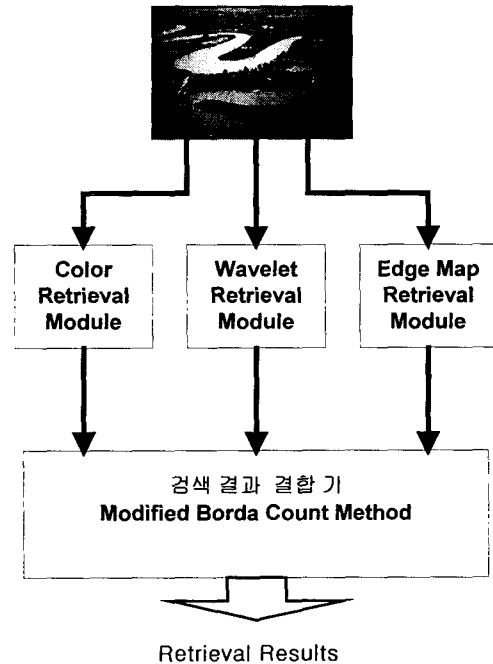


그림 1. Retrieval System using MBCM

전체 검색 시스템은 컬러 검색, 웨이블릿 검색, 에지맵 검색 모듈의 검색 결과를 MBCM 을 사용하여 결합해서 최종 검색 결과를 출력한다.

1. 컬러 검색 모듈

컬러 검색모듈은 컬러 모멘트 피쳐를 사용하여 검색을 한다. 컬러 모멘트는 다음과 같이 구하여진다. 먼저 영상을 엔트로피가 높은 영역과 낮은 영역으로 분리한다. 다음으로 각 영역에 대해 HSV 컬러 공간에서 H, S, V 채널별로 1 차, 2 차, 3 차 모멘트, 즉, 평균, 분산, 스쿼값을 계산한다. 영역이 2 개이므로 전체 피쳐의 개수는 18 개이다. 피쳐 수에 비해 검색 결과가 아주 우수한 검색 방식이다[2,3].

2. 웨이블릿 검색 모듈

웨이블릿은 영상의 질감적인 특징을 잘 표현해준

다. 이 방법은 먼저 영상을 4 layer 웨이블릿 변환을 한다. 그 다음 가장 주파수가 낮은 밴드의 평균과 분산을 구하여 색인하고, 이 밴드와 이웃 3개의 밴드의 웨이블릿 계수를 색인한다. 검색에서는 먼저 평균과 분산값만을 비교하여 검색을 수행한 후 한 번 걸러진 검색결과들의 웨이블릿 계수를 비교하는 2단계 방식의 검색을 수행한다. 이 검색 방식은 검색 속도를 높이기 위해 2단계로 구성된 것이 특징이다[4,5].

3. 에지 맵 검색 모듈

에지 맵은 영상의 모양을 검색하는 방식으로 사용된 3가지 검색기 중에서 가장 간단한 방식이다. 먼저 영상의 에지를 구한다. 에지를 구한 후 영상을 16x16 블록으로 나눈 후 각 블록의 에지 픽셀의 개수를 세어 이 값들을 색인한 후 검색하는 방식이다.

V. 실험 및 고찰

실험에 사용된 데이터 베이스는 MPEG-7의 S3 data set을 사용하였다. S3 data set은 주로 자연 풍경 사진이다. 이 중에서 질의 영상으로 10개를 선택하였고 각 질의 영상은 10개의 유사영상을 가진다.

제안된 MBCM이 각 검색 모듈의 결과를 잘 결합하여 좋은 검색 결과를 출력하였는지를 판단하기 위해 검색 결과에서 상위 10등 내에 false positive error가 몇 개 있는지를 구하였다. 표 1은 이 결과표이다.

Table 1. The number of false positive errors among top ten retrieved results

	color	wavelet	edge map	our method
query 1	6	7	9	7
query 2	5	7	8	4
query 3	9	9	8	8
query 4	4	4	9	6
query 5	6	7	9	6
query 6	6	8	9	4
query 7	5	6	5	4
query 8	2	3	8	2
query 9	5	9	9	7
query 10	7	8	7	4
average	5.5	6.8	8.1	5.2

표 1에서 알 수 있듯이 MBCM의 false positive error 개수가 제일 적다. 즉, MBCM이 검색결과에

서 false positive error를 줄여 더 좋은 검색 결과를 제공한다는 것을 알 수 있다.

VI. 결론

본 논문에서는 MBCM을 이용한 영상 검색 시스템을 제안하였다. 실험결과 제안된 시스템은 서로 다른 피처를 사용하는 각 검색 모듈의 결과를 결합하여 보다 우수한 검색결과를 출력하였다. 검색 결과 결합방법으로 보다 카운트 방법을 사용하였는데, 패턴 인식에서 사용되던 보다 카운트 방법을 영상 검색에 적용하기에는 몇 가지 문제점이 있어 MBCM을 사용하였다.

실험 결과 제안된 MBCM 방법이 검색결과들의 false positive error를 줄여 보다 좋은 검색결과를 출력한다는 것을 알 수 있었다. 현재의 결합 방법은 검색 결과의 순위 정보만을 사용하고 있다. 검색 결과에는 순위 정보 뿐만 아니라 유사도 정보도 있다. 이 정보도 같이 활용한다면 더 좋은 검색 결과를 얻을 수 있을 것이다.

참고문헌

- [1] 김봉근, "다중 분류기 결합을 위한 일반화된 행위지식 공간", 숭실대 박사학위 논문, 1997.
- [2] M. Stricker and M. Orengo, "Similarity of Color Images," Proc. Of SPIE: Storage and Retrieval for Image and Video Databases III, Vol. 2420, pp. 381-392, Shi-Kuo Chang, "Principles of Pictorial Information Systems Design", Prentice-Hall, pp. 61-81, 1989.
- [3] 정세운, 배영래, "Content Based Image Retrieval Using Region Color Distribution Features" 제 10회 영상 처리 이해에 관한 워크샵, pp. 45-50, Jan. 1998.
- [4] James Ze Wang and Sha Xin Wei, "Wavelet-Based image indexing techniques with Partial sketch Retrieval capability," Proceeding of the 4th Forum on Research and Technology Advances in Digital Libraries, pp. 13-24, Washington D.C. May.1997.
- [5] 정세운, 배영래, "Wavelet Transform을 이용한 key-Frame검색", 한국정보과학회 추계학술대회, pp. 509-511, 1999.