

# 내용기반 이미지 검색을 위한 영역별 색상차 분석

## Regional Color Feature Analysis for Content-based Image Retrieval

안재욱, 문성빈 (연세대학교 문헌정보학과)

Jae-wook Ahn, Sung Been Moon  
Dept. of Library and Information Science, Yonsei University

내용기반 이미지 검색에서는 이미지의 하위 영역을 구분하는 방식에 대하여 다양한 접근이 이루어져 왔다. 그중 한 가지가 Stricker와 Dimai가 제안한, 이미지를 다섯 개의 영역으로 나누고 그 가운데 주제 객체가 위치할 것을 가정하여 높은 가중치를 부여하는 방법인데, 본 연구에서는 이와 같은 가정이 타당할 것인가를 S.K. Chang의 PIM(Picture Information Measure) 엔트로피를 계산하여 검증하려 하였다. 실험 결과 이미지의 중앙과 그 외부 영역 사이에는 유의미한 차이가 존재하는 것으로 나타났다. 따라서 Stricker와 Dimai의 방식을 지지할 수 있을 것으로 결론 내릴 수 있다.

### 1. 서론

#### 1.1. 내용기반 이미지 검색

1990년대 들어 컴퓨터와 네트워크 기술의 빠른 발달은 다양한 매체를 이용한 정보의 소장능을 가능하게 하고 이에 대한 이용자들의 정보요구는 끊임없이 증가하고 있다. 이용자들은 지금까지의 텍스트 정보뿐만 아니라 디지털 이미지와 동영상, 디지털 음악과 같은 멀티미디어 정보에 깊은 관심을 갖고 있고, 그러한 경향은 앞으로도 계속될 것이다. 멀티미디어 정보 가운데에서도 특히 이미지는 다른 매체들의 기본이 되는 유형으로서, 이를 쉽고 효과적으로 검색 가능하게 하는 연구가 계속되어 왔다.

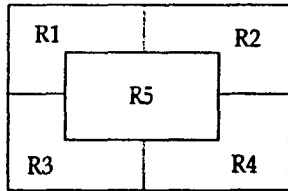
이미지 정보는 학제적이고, 이용자의 주관적인 해석에 의존하는 정도가 강하며, 색인시의 일관성이 매우 떨어진다는 점에서 취급이 쉽지 않다고 알려져

있다(Baxter and Anderson, 1995). 따라서 이미지의 객관적인 특성을 자동 추출하여 자동 색인하는 내용기반 이미지 검색 방식이 큰 관심을 얻고 있다. 내용기반 검색 및 색인에서 주로 이용되는 속성으로는 색상(color), 형태(shape), 무늬(texture) 등이 있는데, 이 중에서도 색상은 정보의 추출과 분석이 비교적 용이하다는 장점을 갖고 있다.

#### 1.2. Stricker와 Dimai의 실험

색상에 의해 이미지의 내용을 색인하는 방식은 크게 전역색상분포(global color distribution)와 지역색상분포(local color distribution)를 이용하는 방식으로 나뉜다. 전자는 이미지 전체의 색상 정보를 색인하며, 후자에서는 이미지 속에 포함된 객체 또는 그 하부영역에 대한 위치 정보를 갖고 있어야 한다

(Stricker and Dimai, 1996). 두 방식은 각각 장단점을 갖게 되고 그 용도 또한 구분되는데, 많은 시스템들은 두 가지를 함께 사용하여 장점들만을 취하고 있기도 하다. Stricker와 Dimai는 이에 대해 최소한의 공간정보만을 색인에 사용해야 한다고 말하고 있다. 그 이유는 이미지를 객체에 의해 구분할 경우 이미지의 회전과 변형으로 인하여 원하는 결과를 이끌어내지 못할 수 있고, 이미지를 다수의 하위 영역으로 구분할 때에는 수작업을 동반해야 할 수가 있으나 그러기에는 처리해야 하는 정보의 양이 너무 많기 때문이다. 또한 전역과 지역 색인을 조합할 경우, 효과적이고 직관적으로 검색 시스템을 구성하기 어려운 정도로 과잉 정보의 문제가 발생할 수 있다는 것이다.



<그림 1> Stricker와 Dimai의 영역 구분

따라서 이들은 전체 이미지의 색상을 단일 히스토그램에 저장하거나, 색상에 의해 이미지에 포함된 객체를 추출하는 것과는 다른 방식을 취했는데, 즉 <그림 1>과 같이 대상 이미지의 영역을 다섯 개로 구분하고 검색 실험을 수행하여 향상된 결과를 보고하고 있다. 이때의 특징은, 첫째 다섯 영역의 경계선 부분에는 퍼지 개념을 도입하였고, 둘째 다섯 영역 중에서도 가운데의 영역에는 바깥 영역보다 높은 가중치를 부여한다는 것이다(Stricker and Dimai, 1996). 이때 가운데 영역에 더 높은 가중치를 부여한 근거는 관찰 결과 중요한 객체들이 가운데 위치하는 경우가 많다는 것인데, 이것은 경험적으로는 타당하다고 할 수 있을지 몰라도 객관적인 근거를 가졌다고 보기는 힘들다. 따라서 본 연구에서는 표본 이미지들을 통계적으로 분석하여 Stricker와 Dimai가 가정했던 것처럼 이미지들의 중앙 영역에 중요한 객체가 위치하는가, 따라서 중앙 영역에 높은 가중치를 줄 수 있는가를 검증하였다.

## 2. 연구의 가설

이 연구에서 관심 영역에 과연 중요한 객체가 존재하는가의 여부를 판단하는 기준이 된 것은 PIM(Picture Information Measure)으로, Shannon의 엔트로피 개념을 이미지 정보에 적용시킬 수 있도록 S.K. Chang이 그레이 레벨 변화의 최소 개념에 기반하여 제안한 것이다. Shannon의 엔트로피가 영상에서의 공간적 구조를 반영할 수 없는데 비해 S.K. Chang의 엔트로피는 이를 반영할 수 있으며, 이를 정량화한 것이 PIM이다. PIM은 다음과 같이 정의된다(김태희, 정동석, 1999).

$$PIM = \sum_{i=0}^{L-1} h(i) - \text{Max}_i(h(i)) \quad (1)$$

L : 양자화 레벨 수, h(i) : 레벨 i에 대한 히스토그램

그리고, PIM을 구하기 위한 관심 블록의 픽셀 수를 고려한 정규화 PIM은 다음과 같이 정의된다.(S.K. Chang, 1989)

$$NPIM = 1 - \frac{\text{Max}_i(h(i))}{\sum_{i=0}^{L-1} h(i)} \quad (2)$$

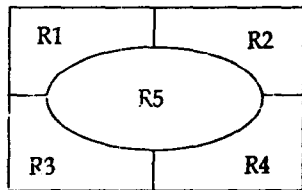
(0 ≤ NPIM ≤ 1)

따라서 PIM이 클수록 이미지에 포함된 픽셀들의 변화는 다양하고, 적을수록 그 변화가 단조롭다고 할 수 있다. 한 가지 색으로 구성된 이미지 내의 영역의 PIM은 0이 되며, 해당 영역 내에 다양한 색의 픽셀이 포함될수록 그 값이 커지는 것이다. 김태희와 정동석(1999)은 이러한 PIM을 이용한 이미지 검색 기법을 소개하면서 대부분의 이미지에 있어서 PIM이 크면 클수록 엔트로피와 색상 변화(그레이스케일 정보가 대상이 되므로 엄밀히 말하면 밝기의 변화)가 커지게 되고, 따라서 그 영역이 많은 정보를 갖고 있다고 말하고 있다. 본 연구에서도 이와 같은 입장을 수용하여 다음과 같이 가설을 설정하였다.

- 1) 가설 : 이미지의 중앙 영역에는 주변 영역보다 중요한 객체가 포함될 확률이 높을 것이다.
- 2) 조작적 가설 : 이미지의 중앙 영역의 PIM값이 주변 영역의 PIM값보다 크다. 따라서 중앙 영역의 PIM과 주변 영역의 PIM의 차이는 0보다 크다.

### 3. 실험

실험에 사용된 이미지들은 인터넷으로부터 수집된 서양 회화 이미지 1,358 장과 미국 NOAA(National Oceanic and Atmospheric Administration)의 사진 컬렉션 중에서 일반 항목 356 장을 선정하였다. 이들을 먼저 명도만을 갖는 그레이스케일 이미지로 변환한 후 영역별로 엔트로피를 구하였고, 각 영역에 포함된 픽셀의 수가 각각 다르므로 표준화된 엔트로피인 NPIM을 사용하여 비교를 가능하게 하였다(식 2). 영역의 구분은 <그림 2>와 같은데, 중앙 영역은 전체 이미지의 1/2 면적을 갖는 타원으로 설정하였으며, 그 외부 영역은 네 개의 영역으로 구분하였다.



<그림 2> 영역 구분



<그림 3> PIM의 예

이 이미지를 그레이스케일로 변환할 때에는 그 양자화 레벨에 차등을 둘 수 있다. 보통 컬러 이미지는 픽셀당 256 단계의 R, G, B로 구성되어 이의 평균을 구하면 256 단계의 그레이스케일로 변환된다. 이 실험에서는 256 레벨에 더하여 64, 32 레벨도 함께 고려하였는데, 그 이유는 이미지를 양자화할 경우 인간의 눈으로 이미지의 주제를 판별할 때에는 크게 중요하다고 할 수 없는 픽셀들의 작은 변

화는 무시할 수 있고, 따라서 영역간 PIM의 차이보다 명확하게 알 수 있기 때문이다. <그림 3>의 이미지에 대하여 256 그레이스케일 레벨일 때의 영역별 엔트로피를 구하여 비교한 것이 <표 1>이다. 비록 이미지의 주제가 되는 앵무새가 본 실험에서 상정한 것과 같은 타원 안에 완벽하게 포함되지는 않지만 상당 부분이 그 안에 포함되어 있고, 따라서 외부 영역에 해당하는 R1, R2, R3, R4의 PIM 및 내영역을 모두 포함하는 R1~R4 영역의 PIM이 중앙에 해당하는 R5의 PIM에 비해 작은 것을 알 수 있다. 이것은 앵무새 부분의 그레이스케일 변화가 배경이 되는 부분의 그레이스케일 변화보다 큰 것이 반영된 결과이다.

<표 1> PIM의 예

영역	R1	R2	R3	R4	R1~R4	R5
NPIM	0.90	0.98	0.98	0.98	0.97	0.99

<표 2> 영역별 엔트로피(NPIM) 평균

대상 이미지	회화 이미지			NOAA 컬렉션			
	그레이스케일 레벨	256	64	32	256	64	32
영역	R1	0.96	0.88	0.80	0.87	0.76	0.67
	R2	0.96	0.86	0.79	0.86	0.75	0.66
	R3	0.95	0.88	0.81	0.82	0.71	0.64
	R4	0.95	0.98	0.81	0.81	0.71	0.64
	R1~R4	0.96	0.91	0.84	0.86	0.77	0.71
R5	0.97	0.92	0.87	0.93	0.86	0.79	

먼저 두 이미지 컬렉션의 영역별 엔트로피를 구하고, 그 평균을 비교해 보면 <표 2>와 같다. 예상했던 것과 같이 주변 영역(R1~R4)의 엔트로피가 중앙 영역(R5)보다 작은 것을 확인할 수 있고 그레이스케일 레벨이 작아질수록, 즉 양자화 정도가 커질수록 엔트로피도 작아지는 것을 알 수 있는데, 이것은 양자화에 의해 작은 색상의 변화는 무시되었기 때문이다. 또한 이미지의 유형에 의해 비교해 보면 회화 이미지보다는 사진 이미지로 구성된 NOAA 컬렉션의 엔트로피 값이 작다는 사실을 알 수 있는데, 이것은 개념적으로 동일한 객체라 할지라도 사

진으로 찍었을 때보다 인간의 손에 의해 그려졌을 경우 불규칙한 면을 더 많이 가질 수 있다는 점을 반영하고 있는 것이다.

<표 3, 4>는 가설에서 상정하였던 중앙 영역과 주변 영역의 차이를 구하여 T-test를 수행한 결과를 정리한 것이다. 평균의 경우와 마찬가지로 차이의 평균도 양자화 정도가 클수록 크게 나타났고, 사진 이미지가 회화 이미지보다 더 큰 값을 나타내고 있다. T-test에 의해 얻어진 t 값은 단측 검증시의 임계치 1.645 ( $\alpha=0.05$ , 자유도 1357, 355) 보다 모두 크게 나타났으므로 이 실험에서 얻어진 결론은 유의하다고 볼 수 있으며, 따라서 이미지의 중앙 영역과 주변 영역 사이에 엔트로피의 차이가 존재한다는 가설은 지지된다고 할 수 있을 것이다.

<표 3> 영역별 엔트로피(NPIM)차이의 평균

대상 이미지	회화 이미지			NOAA 컬렉션			
그레이 스케일 레벨	256	64	32	256	64	32	
영역	R5-R1	0.02	0.04	0.67	0.07	0.10	0.12
	R5-R2	0.02	0.05	0.07	0.07	0.11	0.13
	R5-R3	0.02	0.04	0.06	0.12	0.15	0.16
	R5-R4	0.02	0.04	0.06	0.12	0.15	0.16
	R5-(R1~R4)	0.01	0.02	0.02	0.08	0.09	0.09

<표 4> 영역별 엔트로피(NPIM)차이의 T-test 결과

대상 이미지	회화 이미지			NOAA 컬렉션			
그레이 스케일 레벨	256	64	32	256	64	32	
영역	R5-R1	10.05	16.54	20.43	8.26	10.89	12.41
	R5-R2	11.16	17.57	20.73	12.75	12.40	13.83
	R5-R3	10.01	14.39	17.15	9.97	14.66	14.01
	R5-R4	10.70	14.82	17.80	13.31	14.78	14.58
	R5-(R1~R4)	7.18	7.96	8.68	12.75	11.91	10.48

#### 4. 결론

이 연구에서는 Stricker와 Dimai가 가정했던 '이미지의 중앙 영역에는 중요한 객체가 존재할 것이다'라는 가설을 영역별 엔트로피인 PIM 값을 구하

여 비교해 봄으로써 실제로 검증해 보았다. 실험 결과 이미지의 중앙 영역이 주변 영역에 비해 높은 PIM 값을 갖는 것으로 나타났고, 통계적으로도 유의미한 것으로 검증되었다. 실험 대상이었던 두 이미지 집단 가운데 서양 회화 이미지에서는 차이가 비교적 낮게 나타났다. 이는 회화 이미지가 갖고 있는 속성, 즉 불규칙한 면적이 사진에 비해 상대적으로 많다는 사실에서 기인한 것으로 볼 수 있다.

따라서 Stricker와 Dimai의 가설은 타당하다고 결론 내릴 수 있으며, 이미지 검색 시스템에서도 영역에 따라 다른 가중치를 부여하는 방식이 검색 결과에 긍정적인 영향을 줄 것임을 유추할 수 있다.

물론 이 연구에서 가정했던 바와 같이 이미지 속에서 주제와 관련된 중요한 객체가 항상 높은 엔트로피를 가질 것이라고 무조건 단정할 수는 없다. 그것은 주제적으로 중요한 객체가 주변 배경에 비해 단순한 색상 분포를 가질 수도 있기 때문이다. 따라서 인간에 의한 수작업 색인 결과와 비교하여 주제 객체가 높은 엔트로피를 갖는가를 확인하는 작업이 앞으로의 후속 연구에서 계속될 필요가 있을 것으로 보인다.

#### 5. 참고문헌

1. 김태희, 정동석, "엔트로피와 색채 특징을 이용한 영상 검색 기법", 정보과학회 논문지 제 26권 제 3호, 1999.
2. Graeme Baxter, Douglas Anderson, Image Indexing and Retrieval: Some Problems and Proposed Solutions, *New Library World*, Vol. 96, 1995.
3. Shi-Kuo Chang, Principles of Pictorial Information Systems Design, *Prentice-Hall*, 1989.
4. Markus Stricker, Alexander Dimai, Color Indexing with Weak Spatial Constraints, *SPIE conference*, 1996.