韓國 OA學會 論文誌
第3卷 第2號, 1998. 6.

98-3-2-2-7

수학적 모폴로지를 이용한 화상인식에 관한 연구

남 태 희*

A Study on the Image Recognition Using Mathematical Morphology

Tae-Hee Nam*

요 약

본 논문의 화상 처리 기법은 최근 다양한 기법을 적용하여 화상 및 문자 인식 방법에 대하여 효율적인 인식 방안으로 많이 제시되고 있지만, 실제로 그 인식 분석 처리 과정이 메우 복잡함으로 인하여 충분히 반영되지 못하고 있는 한계성을 가지고 있다.

따라서 본 논문에서는 모폴로지라는 단순 집합론적 기법, 즉 Erosion과 Dilation 그리고 Opening과 Closing, Structuring Element를 적용하여 화상 인식의 유효성을 검증하고자 한다.

Abstract

The image recognition, with various technics, recently presented an effective recognition scheme both for image and character. The analyzing process, however , is highly complicated so that this does not utilized fully. This paper examines the validity of the image recognition through morphology, which is simple and a theory of sets. The morphology applies erosion and dilation, opening and closing, and structuring element.

---

* 동주대학 무역사무자동화과 조교수
논문접수 : 98.4.22 심사완료 : 98.6.10

Ⅰ. 서론

최근 컴퓨터는 그 기능과 활용면에 있어서 매우 다양해지고 많은 발전을 거듭하고 있다. 이러한 컴퓨터의 발전은 컴퓨터로 하여금 화상 인식 처리 기술 및 멀티미디어 지원을 기본으로하여 화상정보를 보다 빠르고 신속하게 필요한 정보만을 추출하는 것을 필요로하고 있다. 본 논문에서는 화상인식 기술에 있어서 가장 기본적인 단계로 기존 복잡한 알고리즘을 필요로하는 전처리과정에서 수학적인 모폴로지(Morphology)를 이용하여 스캐닝한 화상을 2치로 변환하여 수리형태학을 이용한 집합론적 분리 방법을 적용하고자 한다. 이에 실험으로 제시된 화상은 악보 화상에 기초한 음악 파일 작성을 위한 전처리 단계로서 가사가 있는 보통의 악보에서 가사와 순수 악보 부분을 추출및 분리하는 방법을 제안한다. 제안하는 방법은 모폴로지를 이용한 방법으로 집합론적인 처리만을 이용하는 것이다. 이에 관련한 기존의 연구는 패턴 매칭이나 히스토그램을 이용하여 악상 기호의 특징을 인식하는 정도였고, 중요한 것은 이러한 방법들이 복잡한 알고리즘을 필요로하므로, 본 논문에서는 간편한 방법으로 모폴로지라는 집합론적인 방법을 적용하게 되었다.

Ⅱ. 본 론

1. Morphology

Morphology는 수리형태학으로 주어진 2치화상의 특징 추출을 목적으로하며, 집합론적 조작으로 이루어진 대상 화상의 변형 수법에 관한 일관된 논리체계로 이루어져있다.

1.1 기본연산

모폴로지 처리에 있어서 구조요소(stucturing element)와 기본 연산자인 확장(dilation)과 수축(erosion)은 핵심적인 것이다. 즉 Dilation은 영역에 픽셀들을 첨가하는 것이고, Erosion은 영역들로부터 픽셀들을 없애는 것이다. 그러므로 추가나 삭제는 주어진 꼴의 연산자 마스크인 Stucturing Element에 의해 결정된다. 연산은 구조요소를 사용하며 다음의 규칙을 따르는데, 먼저 Dilation은 적어도 구조요소 중 한 픽셀이라도 원화상에서 영역내에 놓여 있다면 출력화상에서 주목 픽셀을 1로 하며, Erosion은 전체 구조요소가 원화상에서 영역 내에 놓여 있다면 출력 화상에서 주목 픽셀을 1로한다. 이 두가지 연산의 조합으로 Opening과 Closing 연산을 정의할 수 있으며, 그 방법과 용도를 정리하면, Openig은 Erosion한 화상을 Dilation하는 연산으로 화상의 가장자리부분을 완만하게 하거나 노이즈와 같은 작은 영역을 없애는데 사용한다. 그리고 Closing은 Dilation한 화상을 Erosion하는 것으로서 가장자리 부분에 공백을 채운다. 이상의 4가지 기본 연산을 실제 화상에 적용한 예를 다음 그림 1에 나타내었다.

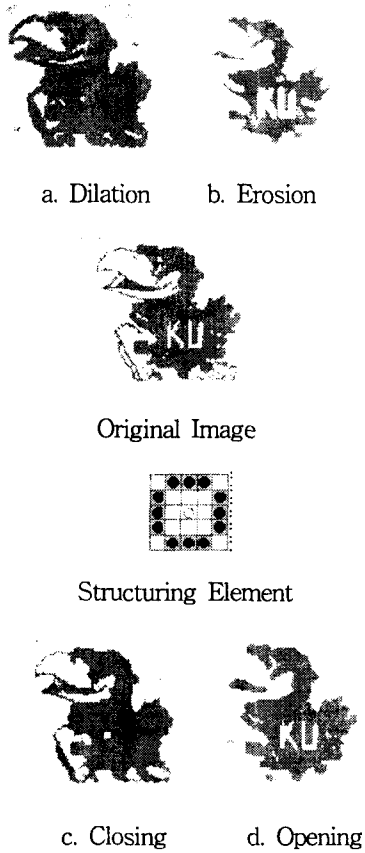


그림 1. 모폴로지 기본 연산
Fig 1. Basic Operation of Morphology

III. 악보 특징 분리

악보는 크게 음표가 있는 오선과 가사부분으로 나눌 수 있는데 여기서는 3 단계를 거쳐 음표, 오선, 가사의 3부분으로 특징을 분리하게 된다. 그 단계는 가사 부분 화상의 획득, 원화상과 차분하여 오선과 음표가 같이 있는 화상을 구하고 마지막 단계로서 그 화상에서 음표와 오선을 분리한다.

1. 가사 및 악보 화상 분리

모폴로지를 이용한 구조요소는 5x5, 3x1 크기로 그림 2와 같다.

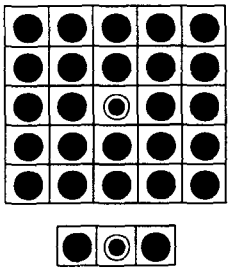


그림 2. 5x5, 3x1 모폴로지 구조요소

Fig 2. Structuring Element of Morphology 5x5 , 3x1

본 structuring element를 악보화상에 적용하여 4회의 Dilation을 취하면 오선과 가사부분의 모든 픽셀이 확장되어 특히 오선 부분이 모두 하나의 영역으로 되며 이 영역은 상대적으로 가사부분이 확장된 영역보다는 크다.

기본 변형에서 Erosion연산을 같은 구조요소로 10회 실시하면 가사부분이 완전히 삭제되고 오선부분은 적어지게 된다. 그리고 다시 10회 Dilation을 취하면 가사부분은 없이 오선부분만 확장된 화상을 얻을 수 있다. 이 화상을 원화상에서 차분하면 약간의 오류가 있는 가사 화상을 얻을 수 있다.

여기서 말하는 오류는 음표의 일부분이나 높은 음자리표의 일부분으로 오선보다 폭이 넓어 가사화상에 같이 남게되나 가사부분의 일렬성을 이용해 제거할 수 있다. 즉 3x1의 구조요소를 이용해 충분히 Dilation한 다음 같은 횟수만큼 Erosion시킨 화상과 먼저 구한 화상인 잡음이 포함된 가사 화상과의 AND 연산을 취함으로써 얻을 수 있다.

이 과정을 다음 식 ①, ②로 나타낼 수 있다.

$$P_{Ch'} = P_{Source} - E^{n}(D^{m} P_{Source}) \quad ..①$$

단 , $m=4$, $n=10$

$$P_{Ch} = P_{Ch'} \cap E^{N}(D^{M}\ P_{Ch'}) \quad ...②$$

단 , $M=30$, $N=30$

D : *dilation*, E : *erosion*

여기서 $P_{Ch'}$는 잡음이 포함된 가사화상, P_{Ch}는 가사만 있는 화상을 뜻하며 P_{Source}는 원화상을 말한다.

다음 단계로 음표와 오선 화상의 획득을 위해서는 식 ①, ②의 과정에서 획득한 가사화상(P_{Ch})을 원화상에서 차분한 화상을 얻을 수 있다. 이 과정을 다음 식 ③으로 나타낼 수 있다.

$$P_{Score} = P_{Source} - P_{Ch} \quad ....③$$

2. 음표와 오선 분리

이 단계에서는 악보화상(P_{Score})에서 음표와 여러가지 음악 기호를 오선에서 분리해내는 것이다. 여기서 사용한 모폴로지 구조요소는 3x1, 1x3 크기의 2가지로 그림 3과 같다.

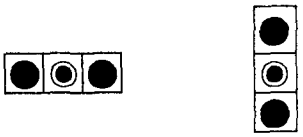


그림 3. 3x1, 1x3 구조요소

Fig 3. Structuring Element 3x1, 1x3

식 ③의 단계에서 획득한 화상, 즉 악보화상 (P_{Score})으로부터 행한다. 오선은 횡축으로만 구성

되어 있으므로 3x1의 구조요소로 충분히 Erosion하면 음자리표, 쉼표 등의 음표부분이 삭제가 되고 오선은 길이가 짧거나 중간이 끊어져 남게된다. 이것을 같은 횟수만큼 Dilation을 취해 원래의 길이만큼 보정해주게 되는데, 본 논문에서는 실험을 통해 그 횟수는 평균 30회로 정하여 실시하였다.

이렇게 하면 오선화상(P_{Five})가 생기고 이 화상을 악보화상(P_{Score})에서 차분하면 손상된 음표화상(P_N')을 얻을 수 있다.

이 화상은 오선이 차분되면서 없어진 부분이 끊어진 것인데 이 간격을 1x3의 구조요소로 2회 Dilation하고 다시 2회 Erosion하여 보정을 하면 약간의 왜곡은 있으나 원화상에 충실한 음표화상(P_N'')이 생긴다.

그리고 좀더 정확한 보정을 위해서는 악보화상인 P_{Score}와 비교연산을 한다. 즉 최종적으로 음표화상인 P_N은 P_N''과 P_{Score}를 AND 연산을 취하고 연속하여 P_{Five}와 AND를 취하여 얻는다.

이상의 분리 과정을 나타낸 식은 아래와 같다.

$$P_{Five} = D^{30}(E^{30}P_{Score}) \quad \cdots\cdots ④$$

$$P_N' = P_{Score} - P_{Five} \quad \cdots\cdots ⑤$$

$$P_N'' = E^2(D^2P_N') \quad \cdots\cdots ⑥$$

$$P_N = (P_N'' \cap P_{Score}) \cap P_{Five} \quad \cdots\cdots ⑦$$

IV. 실험 및 고찰

본 논문의 실험은 586 PC에서 C언어로 구현한 툴을 이용하였고, 실험에 사용한 악보 화상은 600x350의 크기로 300dpi의 해상도에서 스캔하고 툴에 맞는 데이터 형태로 변환하여 실험에 적용하였다. 다음은 대표적인 실험 결과 화상의 예를 보여준다.

그림 4. 원화상
Fig 4. Original Image

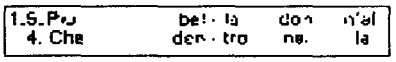


그림 5. 가사 화상
Fig 5. Text Image

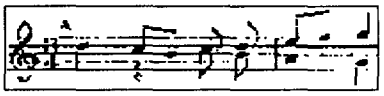


그림 6. 악보 화상
Fig 6. Image of Sheet Music

그림 7. 음표 화상
Fig 7. Image of Music Symbol

그림 8. 오선 화상
Fig 8. Image of Five Line

실험에 대한 표준 화상이 없으므로 100개에 대한 분리 추출의 정확도를 다음 표에 나타내었다.

표 1. 특징별 분리 결과
Table 1. Result of Separation Feature

	가사	음표와 기타기호	오선
에러	23	35	0
정확도	97.4%	95.3%	100%

V. 결 론

본 논문에서 제안된 방법은 화상처리기술의 전처리과정에 해당하는 내용으로 모폴로지라는 단순 집합론적 방법을 이용하여 기존의 복잡한 알고리즘을

적용하여 구현하는 것보다 매우 단순하게 구현할 수 있음을 보인 것으로, 모폴로지의 기본 연산인 Dilation, Erosion 두가지 연산과 5x5, 3x1, 1x3의 구조 요소를 가지고 화상 분리 추출을 수행하였다. 실험에서도 보았듯이 화상의 대상인 악보 전체에 대한 적용이 아직 미흡하나 대상 화상에 대한 적용과 그 분리 추출의 결과에서 알 수 있듯이 상당히 효과적으로 분리 추출됨을 알 수 있다. 그러나 연산 횟수와 순서 획득 등 연산 처리 과정에 있어서 자동적으로 설정하여 실행할 수 있는 작업의 연구가 좀더 필요하며 더 나아가서 분리 추출된 화상중에서 음표 화상의 정확한 보정과 음표가 가지고 있는 박자와 음높이 정보를 획득하는 방법이 좀 더 연구가 진행되어야 할 것이다.

References

[1] Baird, H.S., H. Bunke, K. Yamamoto, "Structured Document Image Analysis," Springer-Verlag, 1992.

[2] Bassmann, H., P.W. Besslich, "Ad Oculos Digital Image Processing," Thomson Publishing, 1995

[3] Castleman, K.R., "Digital Image Processing," Prentice-Hall, 1996.

[4] Crane, R., "Simplified Approach to Image Processing," Prentice-Hall, 1997.

[5] Fisher, Y., "Fractal Image Compression," Springer-Verlag, August 1994.[6] Gonzalez, R.C., R.E. Woods, "Digital Image Processing," Addison-Wesley, June 1992.

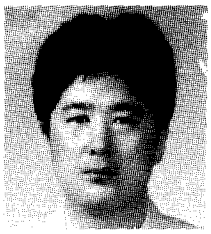
[7] Haralick, R.M., S.R. Sternberg, X. Zhuang, "Image Analysis Using Mathematical Morphology," IEEE Trans. Pattern Anal. Machine Intell., Vol. 9, pp. 532~550, July 1987.

[8] Serra, J., "Image Analysis and Mathematical Morphology" Vol. 2, Academic Press, 1989.

[9] Serra, J., "Image Analysis and Mathematical Morphology" Vol. 1, Academic Press, 1989.

[10] Wornell, G.W., "Signal Processing with Fractals." Prentice-Hall, 1996.

● 저자소개



남 태 희

1989년 : 경성대학교 경영학과 (경영학사)
1992년 : 경성대학교 산업정보학과 (공학석사)
1989년 ~ 1992년 : 우성전산직업전문학교 전산실
1993년 ~ 현재 : 동주대학 무역사무자동화과 조교수
1996년 : 부경대학교 전자공학과 박사과정 수료
관심분야 : 데이터베이스, 화상인식, 정보통신, MIS, GIS