

패턴인식에 의한 고문헌의 진위 판정 시스템

조 동 옥 *

* : 충북과학대학교 정보통신공학과

Author Verification System of Ancient Literatures by Pattern Recognition

Dong Uk Cho *

* : Chungbuk Provincial Univ. of Science & Technology

요 약

본 논문에서는 영상처리 및 패턴 인식 기법을 이용하여 고문헌의 진위판정을 행하는 시스템을 제안하고자 한다. 통상 고문헌은 재질, 서지 형태, 판본과 서문 그리고 내사본과 낙관, 인장, 소장인등에 의해 진위 판정을 행한다. 본 논문에서는 이중 판본에 대해 진위 판정을 행하는 방법을 제안하고자 한다. 이를 위해 조선시대의 37가지 활자본에 대해 영상의 경계 구조 특성 파악에 의한 히스토그램 분석과 각 획의 각도와 길이 그리고 획의 두께등에 의해 진위 판정을 행하고자 한다. 끝으로 실험에 의해 본 논문의 유용성을 입증하고자 한다.

1. 서론

고문헌, 고미술품과 도자기등에 대한 진위판정은 극소수의 전문가들에 의해 행해져왔다. 그러나 이를 문학, 언어학, 통계학, 컴퓨터공학등 다 분야 학제간의 공동 연구로 자동적으로 진위 판정을 행하고자 하는 시스템의 개발에 대해 사회적인 욕구가 증가하고 있다[1]~[3]. 예로서 동경 미술관에서 소장하고 있는 미술 작품 591점중 228점이 위작인 것으로 발표되었다[4]. 따라서 고문헌이나 고미술품의 진위 판정을 자동적으로 행할 수 있는 시스템이 개발되고 이같은 제품이 보편화되면 위작 시비에 대한 종식뿐만 아니라 위작을 행하고자 하는 시도도 미연에 방지할 수 있는 효과도 가져올수 있다. 이를 위해 본 저자는 미술품의 저자 판정을 위한 연구를 수행[5], [6]해 왔으며 미술품에 대한 연구뿐 아니라 고문헌에 대해서도 진위 판정을 행하고자 연구 범위를 확대하고 있다. 현재 고문헌의 진위판정을 전문가에게 의뢰했을시 통상 재질, 서지형태, 활자본, 서문, 내사본, 낙관, 인장, 소장인등을 살펴서 진위 판정을 행한다. 아래(그림1)에 고문헌의 전형적인 예를 나타내었다. 이중 가장 크

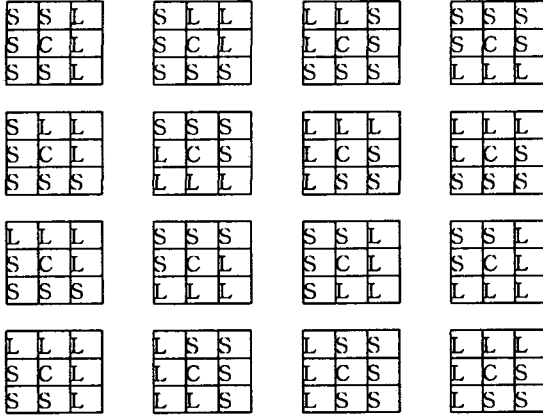
게 위작이 이루어지는 경우가 바로 활자본을 위조해서 원래 연도보다 오래된 문헌으로 조작하는 경우이다. 따라서 본 논문에서는 우선적으로 조선 태종시대의 계미자를 중심으로 활자본을 연구하고 이의 표준 문자 패턴을 생성하여 위조된 활자본 패턴과 유사성을 측정하여 진위 판정을 행하는 방법을 제안하고자 한다.



(그림1) 고문헌의 전형적인 예

2. 활자의 경계 구조 파악에 의한 특징 추출

활자체는 그 경계 형태의 등근 정도나 직선등의처리 결과에 따라 진위 판정에 좋은 벡터가 된다. 따라서 아래 (그림2)와 같은 활자의 경계 형태 분석을 통해 특징 벡터로 사용한다.



(그림2) 경계 구조 형태

$$\Psi_{large}(x) = \frac{x}{255} \text{ ----- (1)}$$

여기서 $x = |C - L|$

$$\Psi_{small}(x) = \frac{-x + 255}{255} \text{ ----- (2)}$$

여기서 $x = |C - S|$

이중 256×256 영상에 대해 16×16의 부영상으로 각 경계 구조의 히스토그램을 형성하고 정합시 효율의 증대를 위해 히스토그램을 4개의 영역으로 나눈다. 나눈 영역에서 4개의 점을 주된 점으로 선정하여 3차식으로 이를 표현한다. 최종적으로 3차식의 계수는 예로서 x^3 의 계수 a 는 하식(3)에 의해 결정된다.

$$a = \frac{D_1}{D} \text{ ----- (3)}$$

여기서

$$D_1 = \begin{vmatrix} y_1 & x_1^2 & x_1 & 1 \\ y_2 & x_2^2 & x_2 & 1 \\ y_3 & x_3^2 & x_3 & 1 \\ y_4 & x_4^2 & x_4 & 1 \end{vmatrix}$$

$$D = \begin{vmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ x_2^3 & x_2^2 & x_2 & 1 \\ x_3^3 & x_3^2 & x_3 & 1 \\ x_4^3 & x_4^2 & x_4 & 1 \end{vmatrix}$$

3. 활자의 획에 대한 특징 추출

활자 획에 대한 특징 벡터로는 획의 시작점과 끝점을 이용한 각, 획의 길이(length)와 획의 두께(width)로 한다.

3.1 획의 각과 길이의 추출

획의 각과 길이를 추출하기 위하여 경계선으로부터 획의 직선 성분을 추출해야 한다. 이를 위해 Hough 변환[7]을 적용한다. 즉, 주어진 직선 L상의 임의의 한점 (X_i, Y_i) 를 지나는 모든 직선들은 아래식으로 정의할 수 있다.

$$X_i \cos \theta + Y_i \sin \theta = R \text{ ----- (4)}$$

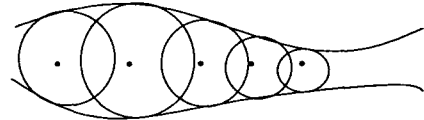
이때 주어진 직선 L상의 모든 점을 지나는 직선은 누적 배열 (θ, R) 에서 한 점에 축적된다는 원리에 의해 직선을 검출하며 이때 획의 각과 길이는 각각 하식 (5), (6)에 의해 구할수 있다.

$$\theta = \tan^{-1} \left(\frac{x_2 - x_1}{y_2 - y_1} \right) \text{ ----- (5)}$$

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \text{ ----- (6)}$$

3.2 획의 두께 검출

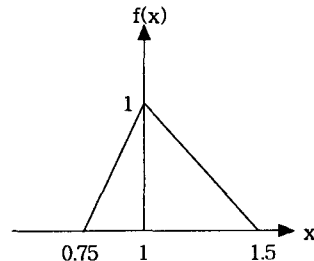
획의 두께는 아래 그림(3)과 같이 획의 내부에 인접하는 최대 원을 통해 획의 두께를 추출한다.



(그림3) 획의 두께 추출

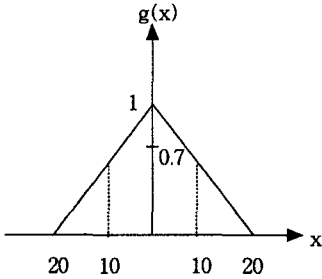
4. 표준 문자 패턴에 대한 유사도 함수의 정의

획의 길이와 두께에 대한 유사도 함수는 아래 (그림4)와 같이 정의한다.



(그림4) 획의 길이와 두께에 대한 유사도 함수

여기서 x 의 값은 표준 문자 패턴의 값과 비교할 문자 패턴과의 값의 비를 나타낸다. 마찬가지로 각에 대한 유사도 함수는 아래 (그림5)와 같이 정의한다.



(그림5) 각에 대한 유사도 함수

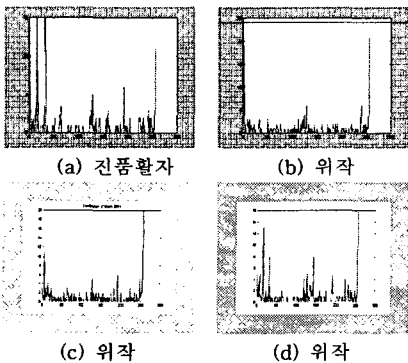
5. 실험 및 고찰

본 논문에서의 실험은 IBM-PC 상에서 행해졌다. 아래 <표1>에 태종 3년 계미자중에 6개의 표본활자와 이에 대한 위작 활자를 표에 나타내었다.

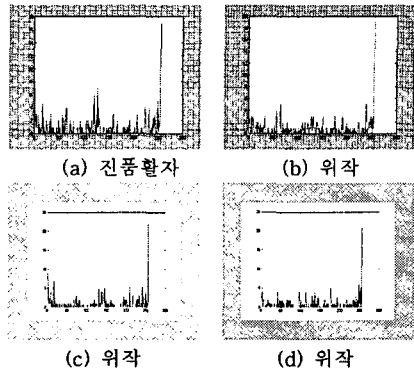
<표1> 태종 3년 계미자에 대한 예

원 활자	위 작 활 자
有	有 有 有
前	前 前 前
位	位 位 位
去	去 去 去
官	官 官 官

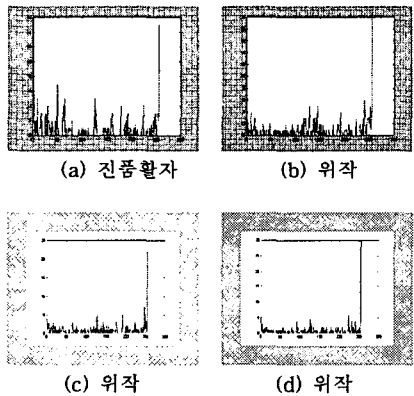
또한 아래 (그림6)~(그림10)에 각각의 활자체에 대한 경계선 특징 형태의 히스토그램을 나타내었다.



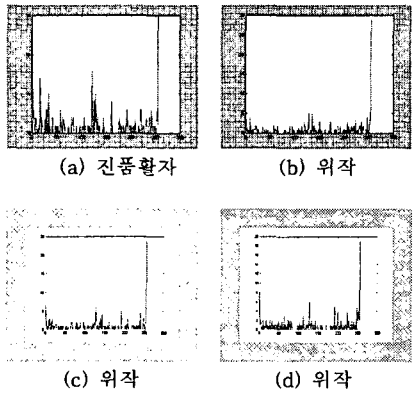
(그림6) 경계 형태의 분석 히스토그램(有)



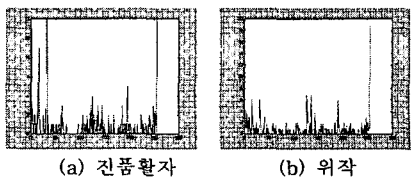
(그림7) 경계 형태의 분석 히스토그램(前)



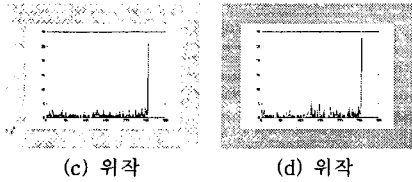
(그림8) 경계 형태의 분석 히스토그램(位)



(그림9) 경계 형태의 분석 히스토그램(去)



(a) 진품활자 (b) 위작



(그림10) 경계 형태의 분석 히스토그램(官)

실험결과에서 알수 있듯이 진품과 위작간의 히스토그램 분석 결과가 확실한 차이가 있음을 알수 있어 경계 형태 분석은 진위 판정에 유용한 벡터가 됨을 알수 있었다. 차후 제안된 알고리즘에 대한 지속적인 실험 수행 및 문제점 파악과 개선등이 행해져야 하리라 여겨진다.

6. 결론

본 논문에서는 고문헌의 진위 판정을 행하기 위한 방법론을 제안하였다. 진품 고문헌에 대해 특징 벡터값을 DB로 저장해 놓는다. 이때 특징벡터로는 경계 형태에 대한 히스토그램 분석결과, 획의 두께, 길이, 각도등과 같은 특징 벡터등을 선정하였다. 실험결과 경계 형태에 대한 분석을 통해 진품과 위작과의 차이를 극명하게 확인할 수 있었으며 차후 다른 특징벡터에 대한 실험수행과 상품화에 대해 지속적인 연구가 수행되어야 하리라 여겨진다.

참고문헌

- [1] R. Sablating et al, Structure Analysis of Paintings Based on Brush Strokes, Anti- Counterfeiting in Art, IS & T/SPIE's 10th Ann. Symp. Electronic Imaging, San Jose, USA, Jan., 1998
- [2] A. Perrig, Albrecht Diirer oder Die Heimlichkeit det deutschen Keteerei, Weinheim, 1987
- [3] A. Perrig, Michelangelo's drawings, New Haven U. a, 1990
- [4] R.D.Lord, Studies in the History of Probability & Statistics VIII, De Morgan and the Statistical Study of Literature Style, Biometrika, 45; 1958
- [5] 조동욱의 1인, "영상처리를 이용한 미술품의 저작자 진위 판단 시스템," 한국통신학회 하 계종합학술대회 논문집, Vol.23, No.2, 2001
- [6] 조동욱, "고미술품 · 고문헌의 진위판정시스템," 한국컴퓨터산업교육학회 논문지, Vol.2, No.5, 2001
- [7] 조동욱의 2인, "정규화에 의한 2차원 물체인 식 알고리즘의 제시," 한국정보처리학회 논문 지, Vol.5, No.8, 1998