

4-방향마스크와 프로젝션을 이용한 손상된 문서에서의 숫자 영상 복원

최선아⁰, 윤미진, 강동구, 김도현, 차의영

부산대학교 전자계산학과
(kiddy, ddaengii, dkkang1, uliminer, eycha}@harmony.cs.pusan.ac.kr

Restoration of corrupted digit image

Using 4-neighborhood mask and projection

Sun-Ah Choi⁰, Mi-Jin Yoon, Dong-Koo Kang, Do-Hyeon Kim, Eui-Young Cha

Dept. of Computer Science, Pusan National University

요 약

본 논문에서는 잘못된 인쇄로 인한 문서상의 잡영이나 문자 훼손이 있는 문서를 복원 하고자 한다. 제안하는 방법은 스캐너로 읽어들이는 문서영상을 잡영 제거를 한 다음 훼손된 숫자 영상에 대해서 프로젝션을 이용하여 숫자 열을 낱알의 숫자로 분할한다. 각각의 숫자에 대해서 크기가 일정하도록 정규화를 시킨 다음, Backpropagation을 이용하여 훼손된 숫자를 학습하였다. 학습시킨 다음 원 영상과 훼손된 영상을 각 픽셀단위로 비교하여 4-방향 마스크를 이용하여 원래의 숫자 영상으로 복원하도록 한다.

1. 서 론

문서나 서적을 복사 시에 원본과는 달리 인쇄문서에 잡영이 섞이거나 제대로 복사가 안된 경우 영상의 훼손으로 글씨를 제대로 알아보지 못하는 경우가 많다. 이 같은 경우, 원영상을 복원하기 위한 방법으로 프로젝션을 이용하여 훼손영역을 검출하는 방법[1]과 모폴로지를 이용한 방법[2], 인식기 이용한 방법 [3]등이 있다. 모폴로지를 이용한 경우, 그 훼손된 영역의 복원이 부정확하다는 단점이 있고, 인식기를 이용한 경우는 처리 시간이 길어진다는 단점을 가지고 있다[1]. 프로젝션을 사용하는 방법은 불필요한 부분 제거시 원 숫자와 함께 제거되어 다시 복원하는 경우 오류가 발생하는 문제를 안고 있다.

본 논문에서는 영상이 훼손된 문서를 300dpi로 스캔으로 읽어들이어 숫자 열을 자르고, 프로젝션을 통해 각각의 숫자로 나눈다.

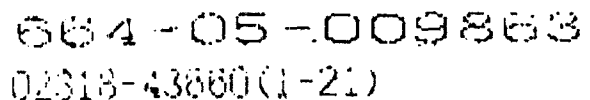
Backpropagation을 이용하여 훼손된 숫자들을 학습을 시켜 인식하게 한 다음, 학습된 결과의 올바른 숫자와 훼손된 영역을 교차하여 원 영상으로 복구하고자 한다. 여기서 사용된 글씨체는 굴림체이고 글씨크기를 정규화시켰다.

본 논문의 2장에서는 잡영 제거 및 숫자 열을 각 하나의 숫자로 나누는 전처리 과정을 거치고, 3장에서는 신경망을 이용, 구체적으로 영상을 복원하여 4장에서는 실험환경과 결과를 분석하고, 5장에서는 결론 및 향후 과제를 말하고자 한다.

2. 문자 분할

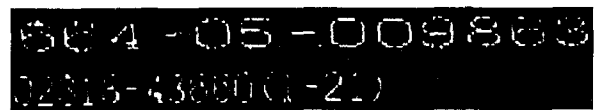
2.1. 입력된 영상의 잡영제거

그림 1은 인쇄된 문서상의 숫자 영상으로 잡영이 많이 있고 인쇄 상태가 좋지 않음을 알 수 있다. 잡영 제거를 위해 간단한 필터를 이용하기로 한다[4]. Size Filter는 잡영을 제거하는데 아주 효과적이고 방법도 간단하기에 자주 사용하는 방법이다. 그림 2는 Size Filter를 이용하여 잡영을 제거한 결과 영상을 보여주고 있다.



664 - 05 - 009863
02318-43880(1-21)

a. 원영상



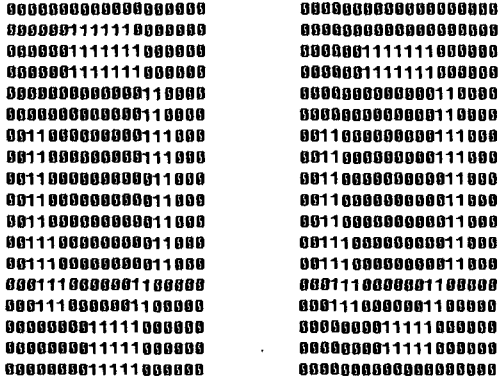
664 - 05 - 009863
02318-43880(1-21)

b. 반전 이미지

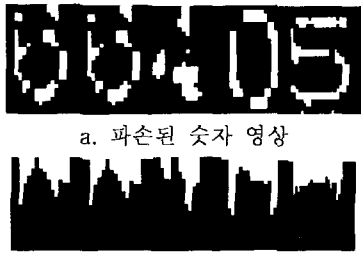
그림 1. 스캔으로 읽어들이는 훼손된 숫자 영상

2.2. 프로젝션을 이용한 숫자열 낱알 분할

필터링을 사용하여 잡영을 다 제거한 후 우리가 복원할 관심 숫자 영역을 수직 프로젝션을 그림 3과 같이 나타낼 수 있다. 그림 3의 결과를 이용하여 히스토그램의 낮은 값을 찾아 이를 기준으로 나누어주는 방법으로 숫자를 분할 할 수 있다[5].



a. 잡영이 섞인 숫자 영상 b. 잡영이 제거된 숫자 영상
그림 2. Size Filter를 이용한 잡영 제거



a. 파손된 숫자 영상 b. 수직 Projection
그림 3. 숫자열의 Projection

3. Backpropagation을 이용한 파손 숫자 복원

3.1. BP 인식 알고리즘

패턴의 개수를 10개로 하여 은닉 층을 2개 두고, 단극성 시그모이드 함수를 사용하였다. Learning Rate를 0.1로 offset Modify Rate를 0.1, Momentum Rate를 0.2로 두고 학습 반복횟수를 1000번 수행하였다. BP 알고리즘은 다음과 같다[6].

- ① 인식하고자 하는 숫자를 입력한다.
- ② 학습과정과 동일한 방법으로 출력 층의 실제 출력 값을 산출한다.
- ③ 출력 값과 가장 유사한 목표 값을 선택하고 그에 해당하는 숫자를 선택한다. 만약 이미 학습과정에서 입력층과 은닉층, 그리고 은닉 층과 은닉층간의 가중치가 수정되었고 비슷한 입력패턴이 들어갈 경우 수정된 가중치에 의해 산출된 출력 값은 그에 해당하는 목표값과 아주 유사하게 된다.

3.2. 숫자 영상 복원

BP를 이용하여 파손된 숫자를 인식한 다음, 그 결과를 이용하여 원영상과 비교를 한다. 우선, 정규화 되지 않은 이미지를 크기 25 × 30으로 정규화 한다. 이 정규화된 입력받은 이미지를 각 픽셀 단위로 파손 영상과 비교하

여 영상이 손상된 부분을 찾아 복원하는 방법을 사용하기로 한다.



a. 파손된 영상 b. 올바른 영상
그림 4. 파손된 영상과 올바른 영상의 비교

위의 그림4는 정규화 시킨 이미지의 원영상과 파손된 영상간의 비교를 각각 수직, 수평 projection을 통해 나타냄을 보여준다. 그림에서 보는 바와 같이 손상된 이미지와 깨끗한 이미지간의 수평, 수직 projection의 차이가 남을 알 수 있다. 그림 4의 (a)의 경우, 손상된 숫자 9라는 이미지를 인식하고 나면 데이터베이스에 저장되어 있는 정규화 된 숫자 9와 비교하여 숫자의 손상된 부분을 프로젝션으로 찾는다. 손상된 영역이라고 예상되는 여러 후보 영역을 두 영상의 픽셀단위로 비교하여 4-방향 마스크를 이용하여 복원한다. 단, 여러 후보 영역 중 숫자 영상 안의 후보영역은 복원 대상에서 제외시킨다. 이미 숫자 영상이기 때문에 복원시킬 필요가 없기 때문이다. 나머지 후보 영역 복원 시, 이웃하는 4방향 중 하나라도 흰 픽셀과 겹쳐지는 부분이 있으면 그 부분을 흰색으로 복원시킨다. 이 과정은 단 한번으로 복원이 안되므로 후보 영역이 사라질 때까지 반복한다. 그림 5는 마스크를 이용한 복원과정을 나타내고 있다. 다음 그림 6은 한 숫자이미지를 예로 결과 영상을 보여주고 있다. 그림 6.a는 원 이미지를 나타내고, 그림 6.b는 스캔으로 읽어들이는 파손된 영상을 반전시킨 이미지이다. 이 두 그림 a 와 b를 각 픽셀단위로 비교하여 파손된 영역을 복원시킨 영상이 그림 6.c. 와 같이 나타내고 있다. 회색조로 색칠된 부분이 그림 b.에서의 누락된 영역을 복원시킨 결과이다.

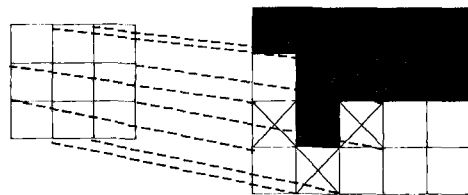


그림 5. 4방향 마스크를 이용한 훼손영역 복원 과정



그림 6. 실험데이터 복원 결과



그림 7. 실제 영상 복원 결과

그림 6는 실험을 하기 위해 임의로 문자를 훼손한 영상으로 복원이 깨끗하게 됨을 볼 수 있다. 그림 7은 실제 잡영이 심한 문서에서 얻은 영상을 복원한 것으로 대체적으로 복원이 잘 되었음을 알 수 있었다.

하지만, 신경 회로망을 이용하여 숫자 인식 시에 파손 데이터를 오인식 하는 경우 그 복원 결과가 잘못 되었음을 알 수 있었다. 그림 8은 숫자 8을 3으로 오인식 하여 복원이 잘못된 예를 보여주고 있다.

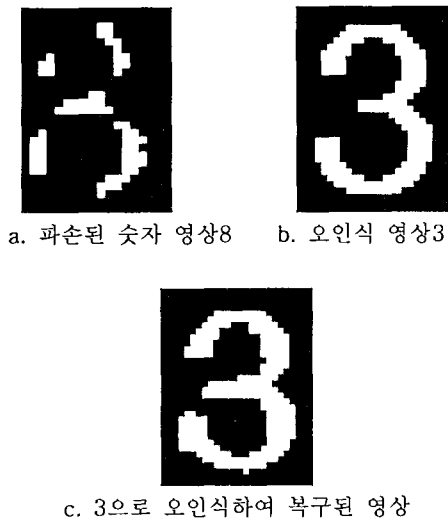


그림 8. 복원 실패 영상

4. 실험환경 및 결과

제안하는 방법을 Pentium III 800MHz PC에서 Visual C++ 6.0을 사용하여 구현하였다. 영상의 크기는 모두 25 × 30의 크기로 정규화하고 글꼴은 굴림체 한 종류로 두었다. 입력 대표 패턴의 갯수를 0 ~ 9 까지 10개로 실험했다. 임의로 만든 실험 데이터에 대해서는 매우 깨끗하게 영상을 모두 복원하였고 잘못된 인쇄로 인해 얻어진 파손된 영상을 실험한 결과 대체적으로 복구가 잘 됨을 알 수 있었다.

5. 결론 및 향후과제

다층 퍼셉트론의 일종인 Backpropagation을 이용하여 미리 파손된 데이터를 인식함으로써 올바르게 숫자 영상을 복구 할 수 있었다. 크기를 하나의 폰트로 정규화 시킴으로 복원 과정을 단순화시킬 수 있었다.

하지만, 신경 회로망을 사용하여 숫자를 인식함으로 그만큼 처리 시간이 길어지는 단점이 있다. 또한 숫자를 오인식 할 경우, 재점검 단계가 없으므로 복원과정에서 잘못된 영상으로 복원될 경우가 있다. 이를 해결하기 위해서 숫자가 잘못 오인식 되지는 않았는지, 그리고 복원 과정에서 잘못 되지 않았는지를 확인하는 재점검 단계가 필요하겠다. 또한 글씨체에 제한을 두지 않고, 기울어진 영상이나 굵게 쓰여진 영상들도 복원 할 수 있도록 유연성이 있어야 하겠다.

6. 참고문헌

- [1] 이창현, 최영우, 김경환, 이일병, "서식 문서의 선과 접촉된 숫자열 복원에 관한 연구", 정보과학회 논문지, 28권 6호, pp. 439-448, 2001.
- [2] D. Guillevic, C. y. Suen, "Cursive Script Recognition: A Fast Reader Scheme", Proc. of 2nd International Conference on Document Analysis and Recognition, pp.311-314, 1993.
- [3] D. Wang, S. N.Srihari, "Analysis of Formation and Artificial Intelligence", Vol. 8, No. 5, pp.1031-1052, 1994.
- [4] Ramesh Jain, Rangachar Kasturi, Brian G Schunck, "Machine Vision", pp. 47-49, 1995.
- [5] J. R. Parker, "Algorithms For Image Proccsing And Computer Vision", pp. 280-286, 1997
- [6] 김대수, "신경망 이론과 응용(1)", pp. 91-120, 1992