

# 우편물 이미지 획득 시스템 개발

송계관, 임길택, 김두식, 남윤석

한국전자통신연구원 우정기술연구부 자동구분처리연구팀

## Development of Postal Image Acquisition System for Sequence Sorting

Jae-Gwan Song, Kil-Tak Lim, Doo-Sik Kim, Yun-Seok Nam

Electronics and Telecommunications Research Institute  
Automatic Sorting and Processing Research Team

e-mail : jgsong@etri.re.kr

### 요약

우편물의 자동구분은, 우편물을 OVIS(OCR-Video coding Integrated System)에 자동으로 공급하고 우편물의 수취인 주소영역을 카메라를 이용하여 획득한 다음 우편번호를 인식하여 바코드로 변환하여 인쇄하게 되고, 이 우편물은 LSM(Letter Sorting Machine)으로 보내져 BCR(Bar Code Reader)에 의해 인쇄된 바코드를 판독하여 행선지별로 구분하는 과정을 거친다. 주소의 번지 이하 부분은 배달원의 수작업에 의해 최종 배달지점 순서대로 정렬한 다음 배달하게 된다. 이 부분의 작업에 소요되는 시간은 배달원 일일 평균 4 시간에 달하며 원가절감 대상으로 지적되고 있다. 이 부분을 자동화 하여 우편물 처리시간을 단축하고 생산성을 향상하는 방안이 대두되고 있으며, 이를 해결하기 위해 번지 부분까지 OCR 을 이용, 인식하여 우편번호 및 순로 데이터 베이스에서 인식결과에 해당하는 코드를 추출하여 해당 구분 칸으로 우편물을 분류하는 방식을 택하면 집배원이 우편물을 배달하는 순로까지 자동으로 정렬할 수 있게 된다. 본 논문은 수취인 주소영역의 주소부분을 자동판독하기 위한 시스템을 개발한 내용을 다루고자 한다.

### 1. 서론

OCR(Optical Character Reader)이 개발되기 전의 우편물 처리는 우편물에 인쇄된 우편번호를 작업자가 육안으로 식별하여 수작업으로 분류하여 처리해 왔다. 그러나 증가하는 우편물량을 처리하기에는 인적·시간적으로 한계가 있으므로 자동구분기계를 이용한 우편물 자동구분의 필요성이 대두되었으며, 인력 의존도를 최소화하고 효율적인 우편배달 서비스를 제공하기 위하여 OCR 기술을 이용한 자동구분기에 의한 우편물 처리방법을 개발하여 업무에 적용하고 있다 [1,2,3,4]. 소형통상우편물의 자동구분처리에는 OVIS(OCR-Video coding Integrated System)에 우편물을 낱장으로 자동 삽입하고 우편물이 이동하는 동안 우편물에 기재된 우편번호를 OCR 에서 인식한 다음 이 정보를 형광 색의 바코드로 우편물에 인쇄하여 LSM(Letter Sorting Machine)으로 보낸다. LSM 에

서는 우편물에 인쇄된 바코드를 판독하여 자동으로 구분하는 절차를 거치게 된다[5,6]. 각 지역별로 운송된 우편물은 배달구역 단위로 배달원에게 할당되고, 배달원은 순서에 맞게 우편물을 정리하여 배달하게 된다[7,8]. 이 우편물을 정리하는 과정이 전체 우편물을 처리하는 원가 비율의 약 27%를 차지하고 있고 오 구분되는 경우가 있어 우편물을 신속하고 정확하게 처리하는데 저해 요인이 되고 있다[5,6]. 이를 해결하기 위해 본 논문은 우편주소의 번지 부분을 자동으로 인식할 수 있는 시스템을 개발한 내용에 대해 기술하고자 한다. 본 논문의 제 1 장에서는 시스템을 개발한 배경 및 필요성에 대해 기술하였고, 제 2 장에서는 우편물의 순로구분 자동화 동향에 대해 다루었으며, 제 3 장에서는 우편물 영상획득 시스템의 개발 내용에 대해 다루었다. 마지막으로 제 4 장에서는 순로구분을 자동으로 실시하기 위한 향후 연구방향 및 내용에 대해 기술하였다.

2. 관련 연구

2.1 집배국 업무분석

집배국별로 구분된 우편물이 우편집중국으로부터 배달되던 집배국에서는 이 우편물은 동별로 구분하고 이를 다시 집배구별로 구분한 다음, 다시 집배원별로 구분하게 된다. 집배원별로 구분된 우편물은 집배원의 배달순서에 맞게 순서대로 구분하게 된다. 이 과정을 그림 1에 나타내었다.

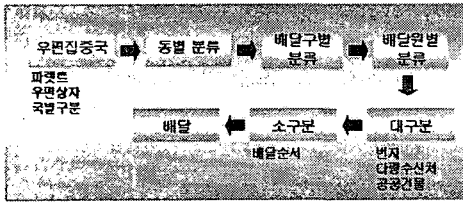
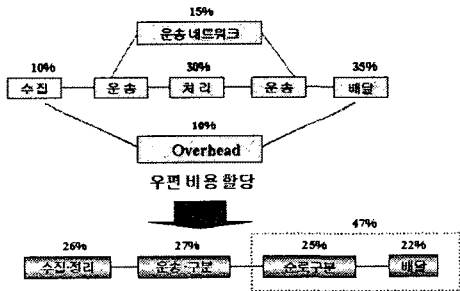


그림 1. 집배국의 업무 처리절차

집배국에 근무하는 집배원 1인당 일일 4시간을 위와 같은 업무를 처리하기 위해 시간을 소요하고 있고, 이것은 우편물 처리원가의 약 27%를 차지하고 있다(그림 2 참조).



Anderson consulting, Siemens

그림 2. 과정별 우편물 처리 원가 분석

현재의 우편물 처리 절차를 살펴보면, 우체통, 우체국 등을 통해 수집된 우편물은 우편집중국으로 운송되고 자동기계를 이용하여 구분한 다음 각 지역별로 운송된다. 지역 우편집중국에서는 관할 구역별 배달 우체국으로 운송하게 되고, 각 우체국에서는 배달원별로 우편물을 할당하여 각 가정이나 사무실 등에 우편물을 배달하게 된다. 우편물을 처리하는 과정별로 원가 비율을 분석해 보면, 우편물을 수집, 정리하는데 26%, 운송 및 구분에 27%, 순로구분 및 배달에 47%의 원가가 소요된다. 이 중에서 가장 많은 부분을 차지하고 있는 순로구분 및 배달부분을 축소하기 위해 순로구분 자동화를 도입할 필요가 있다[11].

2.2 순로구분의 정의

배달원이 우편물을 배달하기 위해 자신의 구역을 순

회하는 경로를 순로라 하고, 이 순로에 따라 우편물을 순서대로 정렬하는 것을 순로구분이라 한다. 이 순로 및 순로구분은 현재의 동 단위로 구분된 우편물을 수작업에 의해 처리하고 있으며, 우편물을 효율적으로 처리하기 위한 매우 중요한 수단이다.

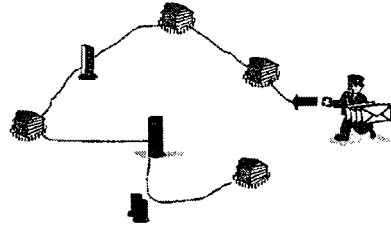


그림 3. 순로 개념도

그림 3은 순로의 개념을 보여주는 그림으로서 우편물 처리의 가장 마지막 처리 단계에 해당한다.

3. 우편물 이미지 획득 시스템

우편물의 이미지를 획득하기 위하여 우편물을 자동으로 이송 시키기 위한 장치를 그림 4에 나타내었다.

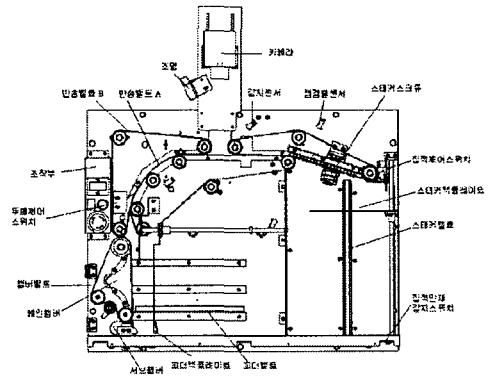


그림 4. 우편물 이송장치 설계도

우편물 이송장치는 우편물을 공급하기 위한 공급대 (Infeed stacker), 우편물을 낱장으로 이송시키기 위한 진공흡입밸브(Vacuum valve), 우편물을 이송시키기 위한 이송벨트(Transportation belt), 우편물의 이미지를 획득하기 위한 카메라(Camera), 우편물을 적재하기 위한 스태커(Stacker)로 구성된다. 진공흡입밸브는 공기의 흡입력을 이용하여 우편물을 한장씩 빨아들여 이송 공급을 방지하며 우편물을 이송벨트에 전달하는 역할을 한다. 이송벨트의 속도는 1.2m/sec의 속도를 가지며, 길이가 200mm인 우편물을 초당 4 통씩 이송할 수 있다. 이송벨트 사이에 설치된 카메라는 라인 스캔(Line scan) 방식의 카메라로서 진행되는 우편물을 라인 단위로 스캔한 다음 이를 합성하여 하나의 이미지를 완성한다.

완성된 이미지는 컴퓨터에 내장된 영상처리 보드에

저장되며 순차적으로 인식기로 보내져 인식을 하게 되고, 이 결과는 각종 통계자료로 활용된다.

그림 5는 개발된 시스템의 우편물 처리 절차를 보여 주며, 그림 6은 획득된 우편 영상의 처리 과정을 보여 주고 있다.

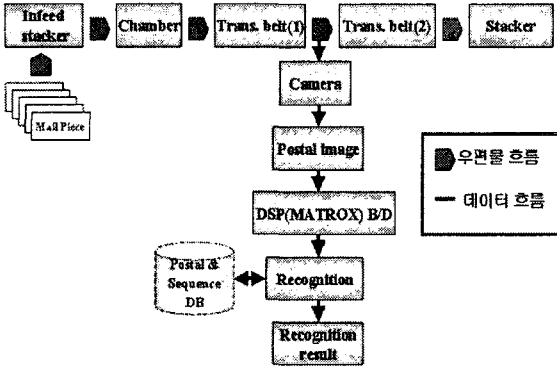


그림 5. 영상획득 시스템에서의 우편물 처리

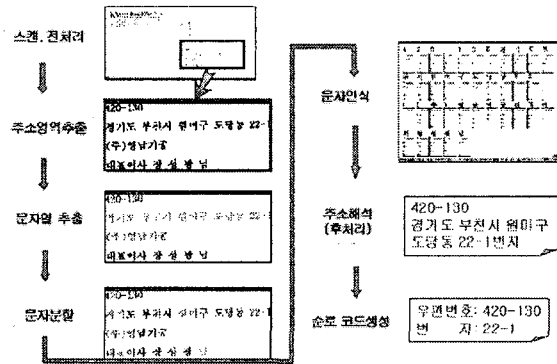


그림 6. 문자인식 절차

그림 7은 한글주소 인식 시스템의 구조도로서, 이미지 획득 부분, 전처리 부분, 문자인식 부분, 주소해석 부분으로 구성된다.

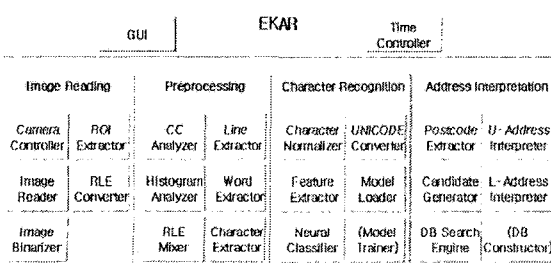


그림 7. 한글 인식 시스템 구조도

### 3.1 이미지 획득 부분

이미지 획득 장치로는 매트록스 사의 라인스캐너를 이용하였다. 초당 4 통의 속도로 이송되는 우편물의 영상을 획득하여, 각 화소의 명도 값이 0 부터 255 사이의 256 단계의 값으로 디지털화 된 그레이(gray) 영상은 PC의 영상처리 보드에 전송된다.

### 3.2 전처리 부분

그레이 영상으로부터 수신자 주소 영역을 포함하는 최소 크기의 영상만을 추출한다. 추출된 주소 영상을 이진 영상으로 변환하는 이진화 과정에서는 Otsu[12]가 제안한 방법 또는 이차원 엔트로피 분석에 의한 방법을 이용하여 이진화를 위한 임계 값을 추정한다.

인쇄체 주소 영상은 각 화소 값을 대상으로 수평 투영을 하여 흑화소가 나타나지 않는 틈(white gap)을 기준으로 문자열을 구분한다. 필기체 주소 영상은 문자열의 기울어짐이나 수평 겹침이 비교적 많이 발생하므로 흑화소에 대한 8 방향 연결요소 분석을 수행한 후, 각 연결요소의 최소외접사각형의 위치와 크기 정보를 이용하여 좌우로 인접한 연결요소를 묶음(grouping)으로서 문자열을 구성한다.

인쇄체 문자열 영상에서 문자 분할 후보 위치를 얻기 위하여 각 화소 값을 대상으로 수직 투영을 하여 흑화소가 나타나지 않는 틈을 기준으로 문자 분할 후보 위치를 얻고, 추정된 평균 문자 폭 정보를 이용하여 문자 병합과 문자 분할을 하면서 문자 분할 경로를 구성한다. 필기체 문자열 영상에서는 연결요소 분석과 획 성분 분석을 병행하여 보다 정교한 문자 분할을 수행한다.

### 3.3 문자 인식 부분

문자 분할에 의하여 추출된 각 문자 영상들을 크기 정규화하고, 특징을 얻는다. 특징 추출 과정에서는 획 윤곽선의 방향 성분을 고려하는 그물망 특징을 추출하고, 이렇게 추출된 특징 정보는 미리 학습되어진 다층 신경망에 입력되어 문자 인식이 수행된다. 하나의 은닉층을 갖는 다층 신경망은 총 8 개가 있으며, 첫번째 다층 신경망에서는 한글의 여섯 유형(종성의 유무 및 모음의 수직/수평/혼합 형태에 따라) 또는 영·숫자를 구분하고, 이 결과에 따라 나머지 7 개의 다층 신경망이 선택적으로 동작한다.

한글을 대상으로 하는 다층 신경망을 학습할 때에는 우리나라의 우편주소 체계 중 시/군/구/동의 이름에 사용되는 글자를 우선 포함하되, 기타 하위 계층의 주소 기입에 자주 나타나는 글자를 보충하여 학습하였다.

문자 인식의 결과로는 하나의 문자 영상에 대하여 총 7 개의 후보를 출력하며, 각 후보에 대한 신뢰값

을 함께 출력하여 주소 해석 시에 이를 고려한다.

### 3.4 주소 해석 부분

주소 해석 과정은 우리나라의 우편 체계를 반영하여 상위부 주소 해석과 하위부 주소 해석으로 구분된다.

상위부 주소 해석에서는 주소 영상의 하단에 위치한 숫자열을 우선 인식하여 우편번호를 미리 추정한 후 이를 이용하여 주소 문자열을 해석하는 방법과, 우편번호 추출을 실패하는 경우에 주소 문자열만을 고려하여 주소의 상위 체계부터 주소 DB의 사전 탐색 방식으로 해석하는 방법으로 나누어진다.

하위부 주소 해석 기능은 우리나라의 주소 DB가 미흡한 관계로 현재는 대전 지역에 대해서만 부분적으로 실험이 이루어졌다.

### 4. 맺음말

우편물의 순로구분을 위해 갖춰야 할 요소는 주소 및 순로 데이터베이스의 구축, 바코드 정보체계 설계, 그리고 순로구분 시설의 확충 등이다.

우정국은 과거의 우편물 수작업 체계에서 자동구분기를 이용한 기계화 집중국을 구축하고 우편물을 보다 신속하고 정확하게 배달하기 위한 노력을 기울이고 있다. 더 나아가 순로구분기를 이용하여 우편물을 배달원의 배달 순서에 맞게 자동화 한다면 익일 배달체계를 실현할 수 있을 것으로 전망된다. 이를 위해 순로구분 자동화의 3대 요소인 바코드의 Format을 4-State로 전환하고 기존의 배달주소 DB와 순로 DB를 구축하며 순로구분기를 보급하는 등 체계를 갖춰야 한다. 또한, 순로구분을 우편집중국이나 배달우체국에서 실시하는 정책적인 결정이 선행되어야 하며, 순로 코드 설계 시 인식되지 못한 우편물에 대한 VCS(Video Coding System)을 고려하여 작업자가 쉽게 타건(Key-in)할 수 있도록 하여야 한다.

### 참고 문헌

1. Donald Clysdale, "Canada Post Corporation's Point of Call Identifier," Advanced Technology Conference, Vol. 2, pp.771-782, December 1992.
2. 박문성, 송재관, 우동진, "우편물 자동처리 촉진을 위한 우편용 우편번호바코드 검증시스템," 한국정보처리학회 논문지 제6권 제4호, pp.968-976, 1999. 4.
3. 유중돈, 권인소 "우편 자동화 기술 - 바코드 인식 기술에 관한 연구," 우정기술 및 서비스 workshop, pp.3-23, 1996. 6.
4. 송재관, "우편용 완성형 한글 4상 바코드 체계 설계," 한국정보처리학회 춘계 학술발표논문집, pp.987-990, 1999. 4.
5. 김태명, "다량우편 이용우편번호에 의한 우편번호 바코드 인쇄 제도," 제4회 우정기술 Workshop, pp.104-106, 1998. 8.
6. 송재관, "우편용 우편번호바코드 검증시스템 개발," 제3회 시스템통합연구회 가을 워크숍, pp.8-12, 1998. 11.
7. 정보통신부, "우편번호에 의한 우편번호 바코드 인쇄 제도," 1999. 5.
8. 한국전자통신연구원, "순로구분 자동처리 시스템," 제5회 우정기술 Workshop, p.68, 2000.10.
9. <http://www.iptp.go.jp/>, 1995.8.
10. <http://www.usps.com/>, 2000. 7.
11. NEC Corp., "신 우편처리 자동화의 흐름," 제5회 우정기술 Workshop, pp.95-106, 2000.10.
12. S. D. Yanowitz and A. M. Bruckstein, "A New Method for Image Segmentation," Proc. of the 9th Int. Conf. on Pattern Recognition, pp. 270-275, 1988.