

다량 우편물의 자동 구분처리 향상을 지원하기 위한 고객 바코드 검증시스템 개발

박문성*, 정윤수, 진병운, 이용준
한국전자통신연구원, 물류ID기술연구팀
(mspark, yoonsu, bwjin, yjl)@etri.re.kr

요 약

문자인식 기술에 의해 우편번호와 주소록 인식하고 4-state 바코드를 인쇄하여 배달순서로 자동구분하기 위한 연구가 진행되고 있다. 이 4-state 바코드를 사전에 인쇄하여 접수될 경우에는 바코드의 인쇄규격 및 수록된 정보의 검증이 요구된다. 본 논문에서는 이미지에 의하여 4-state 바코드를 인쇄규격 검증을 위하여 심볼의 두께 및 공간 값이 정확한 값이 획득, 심볼의 훼손상태, 기울기 값, 수록된 정보인 우편번호와 배달순서 코드를 검증할 수 있도록 한 것이다. 시험결과에 의하면 소형통상 우편물 1통에 대하여 인쇄규격 검증을 위한 소요시간을 30~60msec 이내에 가능하게 되었다.

Development of Customer Bar Code Verification System for Improvement of Automatic Sorting Process in the Bulk Mail

Moon-Sung Park*, Yoon-Su Chung, Byeoung-Woon Jin, Young-Joon Lee
Auto ID Technology Team, ETRI

1. 서 론

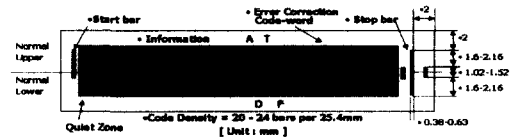
현재, 소형통상 우편물에 대하여 집배원이 배달하는 순서로 자동구분하기 위한 방법이 연구되고 있다. 이러한 절차는 일반 소형통상의 경우에 우편주소록을 인식하고 4-state 바코드에 우편번호와 배달순서 코드를 생성하여 적용하는 방법이 추진되고 있다. 1차 구분은 우편번호 앞 3자리(Optical Character Recognition and Video Coding Integrated System과 Letter Sorting Machine), 2차 구분은 우편번호 뒤 3자리(LSM), 그리고 배달순서 코드에 의한 구분(Delivery Sequence Sorting Machine) 등의 절차로 수행하는 방법이 고려되고 있다. 그러나, 우정 선진국의 경우에도 자동구분 처리 시스템의 OCR에서 우편번호와 주소록 인식하는 비율이 현재 75%(초당 10통 이상일 경우) 정도 수준이다. 이에 따른 오류 우편물(Video Coding 또는 수작업 대상)을 최소화하기 위하여 소형통상 우편물 중에서 가장 많은 비중을 차지하는 다량 우편 이용자들에게(전체 우편물에 60~80%정도 점유) 우편번호, 배달순서 코드 정보를 제공하고 4-state 바코드를 사전에 인쇄토록 하는 방법이 적용되어야 한다. 그리고, 고객에 의하여 인쇄된 바코드의 인쇄규격과 수록된 정보를 정확하게 검증할 수 있게 되면 바코드 판독율을 98%~99%이상으로 유지될 수 있고 오구분의 비율도 최소화될 수 있다. 즉, 고객 바코드가 인쇄된 다량 우편물을 사전에 검증하기 위한 시스템 개발이 필요하다[1,2]. 본 논문에서는 자동구분의 범위를 확대시켜 집배원이 배달하는 순서에 의하여 우편물을 자동으로 구분되는 고객 바코드를 고속으로 인쇄규격을 검증할 수 있도록 하고, 바코드에 수록된 정보의 검증이 가능한 시스템 개발을 목표로 하였다. 제 2장에서는 바코드 시스템의 개요와 바코드의 인쇄규격 검증 및 정보 검증을 위한 요구사항을 검토하고, 제 3장에서는 검증 시스템의 설계, 제 4장에서는 구현된 결과를 보이고, 제 5장에서는 구현 결과를 검토하여 후속 연구사항을 다루었다.

2. 연구 배경

고객에 의하여 우편물상에 인쇄되는 고객 바코드의 인쇄규격을 검토하고 오 구분없이 처리될 수 있도록 바코드 심볼로지의 규격 및 요구사항을 검토하고자 한다.

고객 바코드의 특징은 4개의 높이를 가지며 일정한 심볼의 두께와 공간 값으로 구성된다. 소형 통상 우편물상에 인쇄되는 이 바코드에 대한 인쇄 위치는 우편주소 기재 영역 내에 존재하여야 하며, 우편봉투의 하단 수평라인을 기준으로 최대 5° 이각이 유지되어야 한다. 그리고, 바코드가 존재하는 수평 축상에는 또 다른 심볼 및 문자를 사용하지 않는 것을 원칙으로 하고 있다. 그러나, 바코드 인쇄 규격 검증을 위해서는 몇 도가 조파되었는지 확인할 수 있어야 한다. 바코드에

수록된 정보가 우편주소에 따른 우편번호와 배달순서 코드가 올바르게 적용되었는지 인쇄된 바코드상에 오판독 요인이 존재하여 발생하는 오류와 부분적으로 바코드가 인쇄되지 않거나 훼손으로 인하여 발생한 오류를 정정하기 위하여 적용된 오류정정 코드워드에 대한 검증이 요구된다. 고객 바코드 검증 시스템은 자동구분 처리 시스템에서의 판독방법과 일치하는 방법이 요구된다. 이에 따라 CCD에 의하여 획득된 이미지를 해석하여 인쇄규격 및 정보검증 방법이 요구된다. 구분기에서 사용된 이미지의 해상도는 200dpi(dot per pixel)이므로 검증 시스템 구현을 위한 이미지의 해상도는 200dpi 이상이 되어야 한다[3,5].



$$|수평기울기(\pm 5^\circ)| + |수직기울기(\pm 5^\circ)| < \pm 5^\circ$$



(그림 1) 우편용 4-state 바코드 인쇄규격

인쇄규격에 대한 오차는 (그림 1)에서 보인 오차범위가 최소 0.25mm이므로 이에 대한 규격범위를 고려하여 검증하는 방법이 요구된다. 그리고, 바코드의 정보 검증방법으로는 자동구분 처리 시스템에서 사용중인 우편주소, 우편번호, 배달순서(순로) 코드 등이 일치하는지 검사할 수 있어야 한다. 4-state 바코드에 수록되는 정보는 숫자(2 심볼)와 영문(3 심볼) 또는 완성형 한글(6 심볼)의 표현할 수 있다[1]. (그림 1)과 같이 표현되는 4-state 심볼들은 Tracker(T), Ascender(A), Descender(D), Full Height(H)로 구성된다. 표현 가능한 문자의 수는 하나의 문자정보를 표현하기 위하여 4가지 형태의 심볼의 수로 정의된다. 4-state 바코드 심볼값은 F=0, A=1, D=2, T=3 값으로 설정하고, 심볼 위치값을 4의 계수 값을 적용하였다.

수식 (1)

$$B_n B_{n-1} \dots B_1 = V_n \times 4^{n-1} + V_{n-1} \times 4^{n-2} + \dots + V_1 \times 4^0$$

$$B_n = \text{심볼의 위치}, V_n = \text{심볼의 값}(0,1,2,3)$$

수식(1)은 V_n 심볼값이고, n-1번째 심볼 위치값을 이용하여 심볼로지의 값을 정의한 것이다. 완성형 한글문자(KS 5601, 2,350개)는

49x48의 매트릭스 참조테이블을 생성하여, 행과 열의 값을 수식 (1)에 의하여 계산하고, 열과 행 교차점의 한글문자가 적용되도록 하였다[4,5,6]. 이와 같은 조건들을 바탕으로 고객 바코드 검증 절차에 따른 요구사항은 다음과 같다.

- 1) 검증 대상자에 대한 기초자료 입력 및 우편물 접수
- 2) 우편봉투의 규격(검증자에 의한 검사 결과)의 검증
- 3) 우편주소영역에 바코드 존재 여부 검사(검증자)
- 4) 검증 시스템에 의한 바코드 인쇄규격 검사(심블로지의 기술기, 심블의 두께/높이/간격 값, 심블의 훼손 상태, 오류정정 코드워드의 생성 결과, 수록된 정보의 해석)
- 5) 바코드 규격 검증결과에 따른 보고서 생성 및 검증 결과가 불합격된 이미지 저장(추후 원인분석을 위한 이미지로 활용)
- 6) 수록된 정보 검증은 판독된 결과 중에서 우편번호, 배달순서 코드와 우편주소 정보와 일치 여부 검사되어야 함
- 7) 인쇄규격 및 정보 검증 결과에 따른 보고서 생성 및 데이터베이스에 저장이 필요함
- 8) 고객 바코드 인쇄체도에 시행 방법에 의하여 6개월 단위로 재검사(원인 : 우편번호, 배달순서 코드의 변경 발생)를 받아야 하므로 이전 검증 결과를 참고하여 변경된 사항만 검증하기 위한 참고자료로 활용될 수 있어야 함

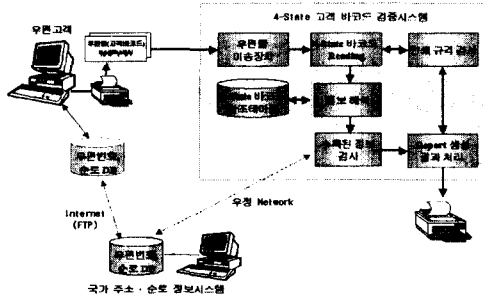
이와 같은 요구사항을 기초로 하여 고객 바코드를 검증할 수 있어야 한다. 이는 우편물의 자동구분 비율 향상 및 운영 비용절감하기 위하여 필수적인 요소이다.

3. 고객 바코드 검증 시스템의 설계

이 장에서는 2장에서 정의한 요구사항을 만족시키기 위하여 고객 바코드의 인쇄규격 및 정보 검증 시스템을 설계한 내용이다. 그리고, 바코드에 수록된 정보를 해석하여 검증하는 과정으로 구분하여 정의하였다.

3.1 고객 바코드 검증시스템 기능 정의

고객 바코드 인쇄규격을 보다 정확하게 검증하기 위해서 바코드 이미지 획득을 위하여 CMOS-CCD(1,024x768)를 사용하고, 우편봉투의 하단 및 우측면을 기준으로 우편주소 기재영역의 범위까지 이미지를 획득하는 방법을 적용하고자 한다. 고객 바코드 검증 시스템은 (그림 2)와 같이 구성하였다. (그림 2)와 같은 고객 바코드 검증시스템 설계된 주요 기능은 다음과 같이 수행될 수 있도록 설계하고자 한다.



(그림 2) 고객 바코드 검증시스템

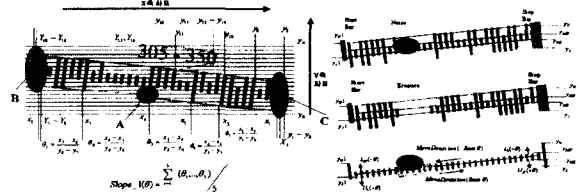
- 1) 우편물을 일정한 간격으로 자동이송 및 제어기능을 포함하는 기능(우편물 이송장치)
- 2) 우편주소 영역의 이미지를 획득하여 메모리에 저장하는 CCD기반 영상획득 기능, 우편 주소영역을 대상으로 함(4-state 바코드 Reading, 라인 스캔 카메라)
- 3) 4-state 바코드가 존재하는 영역을 탐색함과 동시에 심블의 두께, 높이, 기울어진 방향 정보를 생성하는 기능과 이렇게 생성된 정보를 바탕으로 바코드 인쇄규격과 비교하여 차이 값을 생성하는 기능(인쇄규격 검사)
- 4) 획득되어진 심블 값에 의하여 해당 정보를 해석하여 결과(오류정정 코드워드 검증 기능의 포함)를 생성하며 오류인 심블이 발생되면 오류정정 코드워드에 의하여 오류를 정정하는 기능(정보해석), 상기 과정에서는 오류가 검출된 심블의 위치 값이 획득되므로 심블 훼손 정보로 활용
- 5) 심블 값 중에서 우편번호와 배달순서 코드에 의하여 데이터베이스

검색(Local 또는 Remote)하여 우편주소와 일치되는 정보가 사용되었는지 확인하는 기능

- 6) 상기 과정을 반복하여 생성된 검증결과에 따른 합격율을 생성하고 합격 기준에 적합한지 산출하는 기능 이와 같은 과정들에 의하여 생성된 결과에 의하여 고객 바코드 검증결과를 생성하여야 한다. 부가적으로 검증자에 의하여 값도 결과보고서 생성과정에 적용되어 합격율을 생성하는 과정으로 구성하였다.

3.2 인쇄규격 검증을 위한 이미지 탐색 방법

이 절에서는 이미지가 획득되면 4-state 바코드를 탐색하는 방법, 기울기 값의 산출방법, 잡영에 대한 정보 추출 등의 검증절차는 (그림 3)과 같이 정의하였다.



(그림 3) 바코드 영역 검사 및 심블 값 획득

우편영상이 획득되면 그레이 이미지(우편봉투 영상)의 하단부터 검은색 픽셀(pixel)들이 심블의 두께를 갖는지 비교하는 방법을 사용하고, 검출된 픽셀 정보와 다음 수평라인의 정보를 비교하는 과정들에 의하여 바코드일 가능성에 대한 정보(인쇄규격 정보인 심블의 두께, 공간의 범위값)를 기준으로 검사방법으로 설계하였다. 또한, 수평라인을 검사함에 있어 기울어진 방향을 검출하여 수평라인의 검사 대상영역을 최소화 시키는 방법을 포함한다. 바코드가 존재할 가능성에 대한 검사가 완료되면, 시작 심블, 정지 심블, 중앙에 존재하는 트래커 심블 정보에 의하여 중심 축 라인을 생성하여 심블의 값을 해석할 수 있도록 설계하였다. 그리고, 잡영이 존재할 경우와 바코드가 훼손된 값을 산출하기 위하여 심블 두께 및 심블 사이의 간격 값의 분포가 균일한지 확인하고 균일하지 않은 좌표 값에 인접된 심블의 두께 및 간격 값을 산출하여 훼손된 영역을 구별하는 방법을 적용하고자 한다. 검은색 잡영이 바코드 심블내에 존재할 경우에는 심블과 심블사이의 공간에 검은색 픽셀 값이 획득되므로 균일한 심블 두께 및 간격을 유지하는 구간의 정보를 이용하여 훼손된 구간의 정보를 획득하는 방법을 적용하고자 한다.

3.3 정보 검증 방법의 설계

앞 절에서 생성된 결과를 바탕으로 Reed Solomon GF(64) 알고리즘으로 생성한 오류검출 및 정정방법을 수행, 심블 값의 신뢰성을 검사하도록 설계하였다. 오류정정 코드워드는 Reed Solomon 알고리즘의 GF(64) 원시다항식을 적용하여 4개 심블로지(12개 심블)로 생성하였다. 이 방법에 의한 오류정정 범위는 4개 심블로지가 소거되거나 5개의 심블이 다른 값으로 획득되어도 오류가 정정될 수 있다[6,7]. 이 과정에서 오류정정 코드워드가 정확하지, 심블의 정보가 정확하지 검증되므로 별도의 방법으로 오류정정 코드워드에 대한 검증이 요구되지 않는다. 정보의 신뢰성 검증을 위해 획득된 심블 값을 기준으로 문자 정보를 생성하여(우편번호, 배달순서 코드) 자동구분 정보와 일치하는 데이터베이스 검색결과와 일치되는 주소정보가 존재하는지 비교하는 방법으로 구성하였다. 현재 사용되는 우편번호는 집배원별로 할당되어 우편주소 중에서 동 명칭까지 동일하여도 번지 영역에 따라 다른 우편번호를 사용하는 경우가 있으므로 번지 정보를 입력하여 확인하는 과정을 포함한다. 그리고, 수작업으로 검증하는 우편물의 두께 및 우편물의 크기 등에 대한 검사는 검증자에 의하여 검사된 결과를 체크할 수 있도록 하고 이와 같은 검증결과에 의하여 합격/불합격을 결정할 수 있는 결과보고서를 생성되도록 하였다.

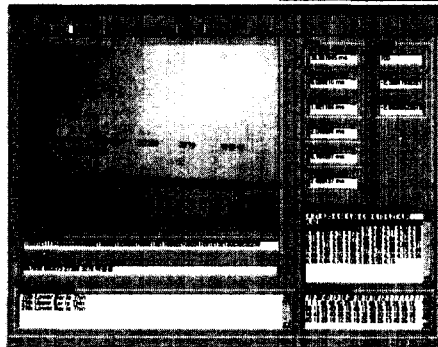
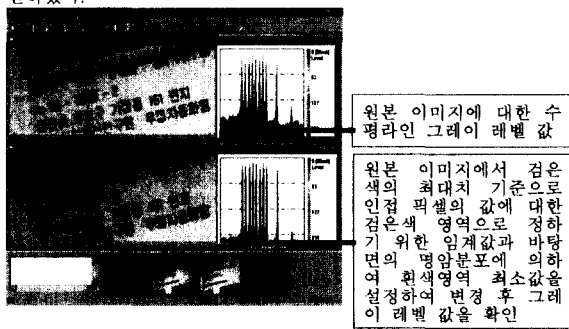
4. 고객 바코드 검증 시스템의 구현

이 장에서는 3장에서 정의된 내용을 기반으로 바코드 규격을 검증하는 시스템의 구현방법과 구현결과에 대하여 다루고자 한다. 우편영상 이미지에서 심블로지 정보의 획득방법을 검토하여 적용하고자 한다. 판독시험을 위하여 표준규격 봉투를 사용하고, 4-state 바코드를 우편주소영역에 인쇄하여 샘플우편물을 생성하였다. 이 샘플우편물의 영상을 획득하여 그레이 이미지 바탕면의 레벨값과 심블값의 크기 분포를 고려한 심블로지 경계값을 생성할 수 있도록 판독 시스템을 구현하였다.

이미지의 크기는 1,024x768로 획득하고 판독대상 영역과 카메라

의 간격을 고정시켜 일정한 이미지가 획득되도록 하였다. 우편봉투의 이미지 중에서 바탕면의 레벨값과 심볼로지 영역의 크기값의 변화분포를 확인하기 위하여 심볼의 두께값을 보다 정확하게 산출되도록 픽셀 값들에 대한 연산 방법으로 탐색대상의 수평라인 선상에서 최대값의 분포에 대하여 검은색 레벨의 값을 결정하기 위한 최소값 기준 생성방법(bipolarize)과 인접 픽셀들의 밝기 정보에 대한 변화량을 기준으로 검은색 또는 흰색 픽셀 값으로 설정하는 방법을 적용하였다. 즉, 바탕면의 밝기 분포와 검은색의 최대값을 기준으로 그레이 레벨값에 대한 임계값으로 변환하면 (그림 4)와 같다.

그레이 이미지의 경계값을 조정하여 시험한 결과에 의하면 60~130값인 경우에 규격검증이 가능한 심볼 두께 및 심볼간 공간값이 정확하게 획득됨을 확인할 수 있었다. 바코드가 검출되기 시작하는 수평라인 상에서 획득된 좌표값과 수평라인을 이동시키면서 트레이커 위치에서 획득된 상태 값을 보인 것이다. 수평 라인선상에 존재하는 좌표값들이 변화되는 분포를 분석하기 위하여 인접된 픽셀간의 차이 값을 기준으로 위치값을 추출하였다. 이 시험결과를 바탕으로 4-state 바코드 심볼로지일 가능성 영역의 좌표값들과 1차 기울기값이 생성된다. 판독기준이 될 수 있는 트레이커들의 위치값을 찾기 위하여 심볼 두께와 심볼사이의 공간값이 균일한 영역과 1차 기울기값을 이용하여 탐색하는 방법을 적용하였다. 이 과정에서 트레이커의 두께, 심볼간 공간값이 정확하게 식별되고, 심볼이 존재하는 위치값들을 획득할 수 있었다. 트레이커를 검사한 결과에 의하여 x, y축의 좌표값을 기준으로 기울어진 각도값을 계산하면 실제 기울기값과 일치됨을 확인하였다.



(그림 4) 심볼의 경계값에 의한 판독시험

이러한 기능들을 추가하여 시험한 결과는 그레이 이미지 상에서의 경계값에 따른 심볼로지일 가능성 좌표값을 추출하고, 1차 기울기 값을 산출을 위하여 36.84msec가 소요되었다. 그리고, 트레이커 영역을 재검사하여 심볼로지 경계값과 2차 기울기값을 계산하는데 0.056msec가 소요되고, 심볼의 상·하위 경계값과 기울기값에 의한 심볼값을 생성하기 위하여 1.82msec정도 소요됨을 확인할 수 있었다.

오류심볼이 발생되었을 경우에 수행되는 Reed Solomon 알고리즘과 심볼로지 정보해석을 위하여 2.23msec가 소요되었다. 심볼의 두께 및 공간값, 훼손 상태 정보를 계산하는데 2.456msec가 소요되었다. 이미지 획득시간을 제외하고 약 43.04msec이내의 우편물 1통에 대한 인쇄규격 검증이 가능하게 되었다. (그림 4)의 결과는 바코드의 인쇄위치, 기울기 등에 관계없이 판독속도가 개선되었음을 보인 것이다. (그림 4)와 (그림 5)에서는 고객 바코드가 인쇄된 우편봉투의 이미지를 획득한 후 시험한 결과이며 바코드가 훼손된 영역 및 고속으로 바코드 심볼로지내에 잠입이 존재할 경우 훼손된 바코드 심볼로

지 위치 및 개수에 대한 정보가 정확하게 획득됨을 확인할 수 있었다. 샘플우편물 200통을 대상으로 시험한 결과에 의하면 30~60msec이내에 바코드 인쇄규격을 검증할 수 있게 되었다. 인쇄규격 검증을 위한 범위는 하나의 픽셀당 0.1mm 단위로 검증이 가능함을 확인할 수 있었다. 이에 따라, 오차 범위가 0.25mm의 범위를 검증하기 위한 조건을 만족하게 되었다.



(그림 5) 고객 바코드 검증 시스템 시험결과

그러나, 심볼들에 대한 인쇄규격을 검증함에 있어 훼손된 심볼이 많은 경우에는 오류정보를 모두 표현하기 어려우므로 심볼에 따른 훼손 상태정보를 간략하게 확인할 수 있도록 두께, 얇다, 높다, 낮다로 표현하는 방법, 심볼의 두께 및 높이 값을 표현하는 방법과 필요에 의하여 상세정보 검색에 의하여 확인할 수 있는 데이터를 생성하였다. 그리고, 인쇄규격 검증 결과에 대한 보고서를 생성할 경우에는 하나의 검증 우편물 단위로 훼손된 심볼의 개수를 저장하여 관리하도록 하고, 규격의 우편물에 대한 이미지 정보를 관리하여 필요시 다시 검증하는 방법을 적용한 것이다.

5. 결론

본 논문에서는 우편봉투에 인쇄된 4-state 바코드를 고속으로 판독하여 인쇄규격을 검증하고, 우편번호 및 배달순서 코드 정보를 검증하는 시스템을 개발한 내용이다. 연구결과에 의하면, 4-state 바코드 심볼로지의 기울어진 각도의 범위값을 정확하게 획득 및 이미지의 회전없이 판독하여 검증할 수 있게 되었으며, 인쇄상태의 불량으로 인한 흰색 또는 검은색 잡영 제거없이 판독하여 훼손된 상태에 대한 정보를 생성할 수 있도록 한 것이다. 또한, 기존의 바코드 판독 시스템은 바코드 심볼로지가 인쇄된 좌·우·상·하 영역에 정해진 규격의 여백(3mm)과 바코드 심볼로지가 존재하는 수평라인 상에 또 다른 문자, 기호, 그림 등의 표기된 경우는 판독할 수 없었으나, 이와 같은 조건에서도 4-state 바코드 심볼로지를 판독할 수 있게 되었다. 심볼로지의 인쇄위치, 크기에 대한 제한 조건없이 판독(휴대용 판독기로 적용 가능)될 수 있게 되었다. 추후 연구사항으로는 우편물 자동구분 처리 시스템에 탑재하고, 개선된 이미지 판독 시스템을 적용할 경우에 우편물 1통을 30~60msec이내에 판독 처리되도록 하기 위한 적용연구와 고객 바코드의 정보 검증방법을 문자인식 기술과 통합하여 자동구분처리와 동시에 고객 바코드를 검증하기 위한 연구가 요구된다. 또한 고객 ID를 활용하여 고객 바코드 인쇄우편물의 서비스 품질관리 및 부가서비스를 제공하기 위한 방안에 대한 연구도 필요하다.

참고 문헌

- [1] Hitoshi Uehara, Yoshiaki Nakamura, Masataka Nakagawa, Terutaka Tanaka, Akira Ohzawa, Ichiro Isawa, Hiroshi Miyoshi, "Research on Barcodes for Mechanized Mail Processing," <http://www.iptp.go.jp/>, July 1995.
- [2] Masataka Nakagawa, Hiroshi Miyoshi, "Barcodes For Use in Mechanized Mail Processing," <http://www.iptp.go.jp/>, June 1995.
- [3] Royal Mail, "Mailsort Customer Barcoding Technical Specification," OXFORD, pp.1-11, April 1995.
- [4] 박문성, "우편용 한글 4-state 바코드," 제4회 우정 workshop, pp.139-152, 1998. 9.
- [5] 박문성, 송재관, 우동진, "우편물 자동처리 추진을 위한 3 out of 5 고객 바코드 검증 시스템," 한국정보과학회 학술발표논문집, 제25권 제2호, pp.496-498, 1998.11.
- [6] 박문성, 김해규, 정희경 "우편용 4-state 바코드 고속 판독방법에 관한 연구," 한국정보처리학회 논문지, 제8-D권 제3호, pp.285-294, 2001. 6.
- [7] 오규택, "Erasure 정정이 가능한 효율적인 Reed-Solomon 복호기용 VLSI," 서울대학교 대학원 전기공학부, 공학석사학위논문, pp.4-36, 1998. 2.