

우편물 자동구분 처리를 위한 4-state 바코드 정보수록 밀도향상 방법

박문성*, 송재관*, 황재각*, 남윤석*

*전자통신연구원 우정자동화팀

e-mail : mspark@etri.re.kr

Improvement Method of Information Density of 4-state Bar Code for Automatic Processing of Mail Items

Moon-Sung Park*, Jae-Gwan Song*, Jae-Gak Hwang*, Yun-Seok Nam*

*ETRI, Postal Automation Technology Development Team

요 약

우편물 자동구분 처리의 범위를 확장시키기 위하여 집배원의 배달순서로 구분하기 위한 방법이 요구된다. 이는 우편 자동구분처리를 위하여 우편봉투에 인쇄된 주소 및 우편번호를 인식하고 이에 대응하는 자동구분정보를 우편물상에 바코드 인쇄하고 판독하여 구분하게 된다. 또한, 다량우편 이용자들이 의하여 우편물 상에 고객 바코드가 사전에 인쇄된 경우에는 바코드를 판독하여 구분한다. 본 논문에서는 우편물 자동구분 처리를 위하여 적용될 4-state 바코드에 정보수록 밀도를 향상시키기 위한 방법을 적용한 것이다. 시험 결과에 의하면, 4-state 바코드의 정보수록밀도는 기존의 16.67% 보다 약 27.79% 향상시킬 수 있었다.

1. 서론

우편물의 자동구분처리 대상인 우편물은 우편물의 크기 및 서비스 내용에 따라 구분될 수 있다. 일반 소형통상 우편물(시간 당 35,000 통)과 대형 통상우편물을 자동으로 구분할 수 있다.

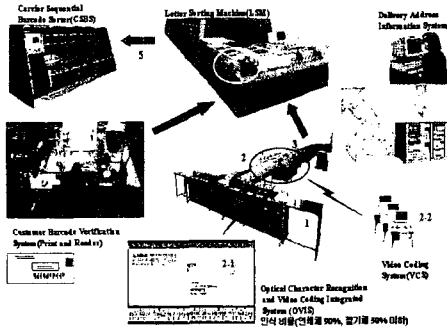
소포인 경우에는 크기 및 다양한 부피가 존재하므로 컨베이어 벨트시스템에 의하여 소포 우편물에 부착된 바코드를 판독하거나, 운영자에 의하여 직접발송 및 도착구분을 위한 정보를 입력하여 구분하는 방법을 병행하여 사용하고 있다. 현재, 우편물의 총 물량을 기준으로 대부분을 차지하고 있는 소형 통상우편물을 집배원이 배달하는 순서로 자동구분 되도록 하기 위한 연구가 진행 중이다. 이와 관련하여, 우편번호를 포함하여 배달순서정보를 적용하기 위한 4-state 바코드시스템을 개발하고 있다. 또한, 4-state 바코드는 다량우편 이용자들이 의하여 우편물에 바코드를 사전에 인쇄하여 우편주소 인식 및 바코드 인쇄과정을 이용하지 않고 구분기에 의하여 자동 구분하는 방법에도 적용될 예정이다. 이 바코드에 정보수록밀도를 높여 다양한 정보를 수록할 수 있도록 하고, 우편물류 처리를 효과적으로 수행하기 위하여 부가적으로 요구될 수 있는 정보가 포함될 수 있

도록 하기 위한 노력이 필요하다[1,2,6]. 이에 따라, 본 논문에서는 우편물 자동구분을 위하여 개발중인 4-state 바코드에 보다 많은 정보를 표현하기 위한 방법을 설계하여 적용한 것이다. 제 2 장에서 우편물 자동구분 처리에 대한 개요, 제 3 장에서는 4-state 바코드의 정보수록밀도 향상 방법 설계 및 바코드 정보수록밀도 적용 이전과 간단히 비교하고, 제 4 장에서는 연구결과와 추후 연구사항에 대하여 다루고자 한다.

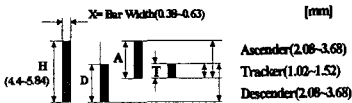
2. 우편물 자동구분 처리의 개요

우편물 자동구분은 우편물량이 증가에 따라 인력 투입이 증가되는 것을 최소화하고, 우편서비스 품질을 향상시키기 위한 목적을 포함하고 있다. 우편물이 수집되어 우편물 처리센터에 접수되면, 우편물들을 주소 기재면을 기준으로 배열한다. 이 우편물을 자동 구분기에 투입하면, 이송되는 우편물의 이미지를 capture 하여 한글주소인식(개발 중)과 우편번호 인식을 실시한다. 이 결과에 의하여 우편번호와 배달순서 정보를 형광색 4-state 바코드로 인쇄한다. 이 우편물은 overhead 컨베이어 벨트를 통하여 구분기로 이송시킨다. 바코드 판독기에 의하여 판독한 후 우편번호 3 자리

에 의하여 1차 구분(발송구분)한다. 그리고, 2차 구분은 뒤 우편번호 3자리로 하게 되며, 1차 구분 결과가 동일지역의 우편물은 구분기에 채투입하여 2차 구분을 실시하고, 타 지역의 우편물은 해당 우편물 처리센터로 운송한 후 구분하는 방법을 적용한다. 또한, 우편물 접수과정에서 고객이 우편물에 4-state 바코드(검은색)를 인쇄하여 접수한 경우에는 검증절차를 통하여 확인한 후, 한글주소 및 우편번호 인식과정 없이 구분기에 직접 투입하여 처리할 수 있도록 할 예정이다(그림 1).



(그림 1) 우편물 자동처리 절차 개념도



(그림 2) 우편용 4-state 바코드 인쇄규격

이 4-state 바코드는 4가지 높이로 구성된 바들을 일정한 간격으로 표현하고, 각 바의 위치에 따라 가중치를 부여하며, 1-5개 바 단위로 하나의 문자를 표현할 수 있다[1, 4, 5].

(그림 2)와 같이 표현되는 4-state 바들은 Tracker(T) 바, Ascender(A) 바, Descender(D) 바, Full Height(H) 바로 구성된다. 4가지 형태의 바들의 어떻게 조합하는가에 의하여 다양한 정보를 수록할 수 있게 된다[1, 5, 8].

현재, 국내 우편물류의 자동화를 위하여 최소한의 공간에 보다 많은 정보를 수록될 수 있도록 설계를 완료한 상태이다. 우정 정보의 유형을 분석하여 숫자(4-state 3 바), 영문자 (4-state 3 바), 조합형 한글(4-state 3 바 x (2-3자소)), 완성형 한글(4-state 6 바)을 포함한 모든 문자정보가 표현될 수 있도록 설계되어 있다. 4-state 바코드 심볼로직 값은 바의 형태에 따라 부여된 F=0, A=1, D=2, T=3의 값과 바의 위치 값을 4의 계수 값에 적용하는 방법으로 구성된다.

[수식 1]

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

[수식 1]을 이용하여 다양한 문자를 표현할 수 있다. 예를 들면, 완성형 한글문자(KS 5601, 2,350개)를 적용하기 위하여 49x48의 매트릭스 참조테이블을 생성하여, 행과 열의 값은 [수식 1]을 적용하여 구하고, 열과 행의 교차점의 문자 정보가 적용되도록 하였다[1, 2, 7, 8].

3. 정보수룩밀도 개선방법

4-state 바코드는 우편번호 6 자리, 배달순서정보(6 자리)로 가정하고, Reed Solomon 알고리즘 GF(64, 57)방법에 의하여 오류정정 코드워드를 생성하게 되면, 12 개 바를 사

용하게 된다. 따라서, 3 개 바에 의하여 하나의 숫자 정보를 표현하도록 하면, 총 48 개의 바가 필요하며, 시작 및 정지 바를 표시하기 위하여 2-4 개의 바를 사용하면 최대 52 개의 바를 사용에 정보를 수록할 수 있다. 이와 같은 결과와 우편물 상에 바코드를 인쇄할 수 있는 공간과 바코드 판독 가능한 영역을 기준으로 총 바의 개수를 산출하면, 82 개 바를 적용할 수 있다[6, 7, 8]. 고객 바코드의 경우에는 우편고객의 우편물 자동구분처리 방법은 물론, 고객관리 및 우편물 처리결과에 대한 정보의 활용 측면이 고려되어야 한다. 이에 따라, 본 절에서는 숫자, 영문, 한글 문자 등의 활용 등을 고려하여 정보수룩밀도 향상을 위한 방법을 설계한 것이다.

3.1 숫자 정보수룩밀도 향상방법

4-state 바코드에 숫자 정보를 표현하기 위하여 필요한 바의 개수를 최소화하기 위한 방법을 다루고자 한다. 숫자 정보는 우편번호, 배달순서정보, 고객식별번호, 우편물 ID 등 다양한 정보가 적용될 수 있다. 이에 따라, 바의 수를 적게 사용하고, 많은 숫자정보를 바코드에 표현하기 위하여 다음과 같은 절차에 의하여 정보수룩밀도를 향상시킬 수 있도록 하였다.

[수식 2]

$$Num_{(a_3)} = \text{int} \left(\frac{Num_{(a_0)}}{64^2} \right)$$

$$Num_{(a_2)} = \text{int} \left(\frac{64^2 \left(\frac{Num_{(a_0)}}{64^2} - Num_{(a_3)} \right)}{64^2} \right) = \text{int} \left(\frac{Num_{(a_0)}}{64^2} - 64Num_{(a_3)} \right)$$

$$Num_{(a_1)} = \text{int} \left(\frac{64^2 \left(\frac{Num_{(a_0)}}{64^2} - Num_{(a_2)} - Num_{(a_3)} \right)}{64^2} \right) = \text{int} \left(\frac{Num_{(a_0)}}{64} - 64Num_{(a_2)} \right)$$

$$Num_{(a_0)} = \text{int} \left(\frac{64^2 \left(\frac{Num_{(a_0)}}{64^2} - Num_{(a_1)} - Num_{(a_2)} - Num_{(a_3)} \right)}{64^2} \right) = \text{int} (Num_{(a_2)} - 64Num_{(a_1)})$$

(1) 숫자 정보를 십진수로 치환

예. 우편번호 123 - 456 = 123456

(2) 십진수로 치환된 값을 64³으로 나눈다.

(3) (2)항의 값의 정수 값을 압축데이터 값 1로 설정한다.

(4) (2)항의 값에서 (3)항의 차를 구한 후, 64²으로 나눈다.

(5) (2)와 (3)항의 방법과 동일한 방법으로 반복하여 64⁰까지 계산한 후 생성되는 값들을 적용한다.

<표 1> 숫자 정보수룩(64 modulo)

	표현정보	64(3)	64(2)	64(1)
데이터(2자리)	99	0.0242	1.5469	35
4-state 값		Not	1	35
바 개수	4		1	3
데이터(3자리)	999	0.2439	15.609	39
4-state 값		Not	15	39
바 개수	5		2	3
데이터(4자리)	9999	2.4412	28.234	15
4-state 값		2	28	15
바 개수	7	1	3	3
데이터(5자리)	99999	24.414	26.484	31
4-state 값		24	26	31
바 개수	9	3	3	3

이러한 방법을 적용하여 숫자를 압축한 경우에 첫번째 자리는 하나의 바 즉, 0~4 의 값 범위이고, 나머지 값들은 4-state 3 바를 적용할 수 있다. 이에 따라, 숫자 6 자리를 압축하면, 총 10 개의 바에 의하여 표현할 수 있다. 영국 방식인 4-state 4 바를 사용할 경우에는 24 개의 바가 요구

되며, 4-state 3 바로 표현할 경우에는 18 개의 바가 요구 된다.

3.2 영문자 정보수축밀도 향상방법

영문자 정보(대문자, 소문자)를 표현함에 있어 4-state 4, 3 바 단위로 생성할 수 있다. 그러나, 영문자의 경우에 중요 주소 정보명칭의 약자로 사용되므로 대문자를 주로 사용하게 된다. 또한, 영문 대문자는 26 개의 바로 표현되며, 영문 대소문자를 모두 적용한 바코드 참조테이블(64 개)을 이용하여 압축정보를 표현하기 위한 방법을 설계하고자 한다. 영문자의 경우에 대문자만을 적용할 경우 27 에 의한 배수와 64 modulo 방법을 혼합하여 적용하면 정보 압축을 용이하게 할 수 있다. 또한, 영문자와 숫자정보를 혼합하여 사용할 경우에는 37 진수에 의하여 10 진수 값을 구한 후에 64 modulo 방법을 적용하면 된다.

[수식 4]

$$Num_{(a3)} = \text{int}(\frac{V_{Rest}}{64^3})$$

$$Num_{(a2)} = \text{int}(\frac{64^3(\frac{V_{4n,Rest}}{64^3} - Num_{(a3)})}{64^2} = \text{int}(\frac{V_{4n,Rest}}{64^2} - 64Num_{(a3)})$$

$$Num_{(a1)} = \text{int}(\frac{64^2(\frac{Num_{(a3)}}{64^2} - Num_{(a2)})}{64} = \text{int}(\frac{Num_{(a3)}}{64} - 64Num_{(a2)})$$

$$Num_{(a0)} = \text{int}(\frac{64(\frac{Num_{(a2)}}{64} - Num_{(a1)})}{64^0} = \text{int}(Num_{(a2)} - 64Num_{(a1)})$$

- 1) $V_n = b_n 27^n + \dots + b_3 27^3 + b_2 27^2 + b_1 27^1 + b_0 27^0$ 의 수식에 의하여 십진수 값으로 치환
- 2) V_n 의 값에서 앞의 N, , , 5, 4, 3, 2, 1 자리 단위로 구분하여 $64^3, , 64^0$ 로 나누어, 4 자리로 구분할 경우에는 다음과 같은 수식에 의하여 각 바코드 심볼로지의 값을 구한다. V_{4n} 은 V_n 에서 앞의 4 자리를 선택한 경우이다.
- 3) 나머지 자릿수로 2)항과 같은 방법을 적용하여 산출할 수 있다.

이 수식에서 V_{Rest} 는 V_n 에서의 나머지 자릿수를 표시한 것이다. 예를 들어, 5 개의 영문 대문자의 경우에 다음과 같은 방법에 의하여 바코드에 표현하도록 하였다. 영문자 정보를 십진수로 치환하고, 영문 대문자의 값이 21b4, 22 b3, 13 b2, 12 b1, 26 b0 의 값으로 가정하여 산출하면, 11, 9, 27, 1, 50 의 값이 되고, 4-state 2 바($V_4=11$), 2 바($V_3=9$), 3 바($V_2=27$), 2 바($V_1=1$), 3 바($V_0=50$) 바의 순으로 계산된 값을 표현하면 총 12 개의 바로 5 개의 영문 대문자를 표현할 수 있다.

3.3 한글 정보수축밀도 향상방법

KS 5601 완성형 한글문자는 2,350 개의 문자로 구성되는데 이러한 문자를 4-state 바코드로 표현하려면, 49 x 49 개의 매트릭스 테이블로 구성하고, 행과 열의 값을 0-48 로 정의하여 사용하면 된다. 이에 따라, 4-state 6 바에 의하여 한글 한문자를 표현할 수 있도록 하고, 이 방법에 행과 열의 값 50 modulo 방법을 적용하여 산출하고, 값을 특정 자릿수로 구분하여 숫자 압축방법을 적용하면 보다 많은 정보를 수록할 수 있을 것으로 기대된다.

- (1) $V_n = b_n 50^n + \dots + b_3 50^3 + b_2 50^2 + b_1 50^1 + b_0 50^0$ 의 수식에 의하여 십진수 값으로 치환
- (2) V_n 의 값에서 앞의 N, , , 5, 4, 3, 2 자리 단위로 구

분하여 $256^3, , 256^0$ 로 나누어, 4 자리로 구분할 경우에는 다음과 같은 수식에 의하여 각 바코드 심볼로지의 값을 구한다. V_{4n} 은 V_n 에서 앞의 4 자리를 선택한 경우이다.

- (3) 나머지 자릿수로 2)항과 같은 방법을 적용하여 산출할 수 있다. 이 수식에서 V_{Rest} 는 V_n 에서 나머지 자릿수를 표시한 것이다.

[수식 5]

$$Num_{(a3)} = \text{int}(\frac{V_{4n}}{256^3})$$

$$Num_{(a2)} = \text{int}(\frac{256^3(\frac{V_{4n}}{256^3} - Num_{(a3)})}{256^2} = \text{int}(\frac{V_{4n}}{256^2} - 256Num_{(a3)})$$

$$Num_{(a1)} = \text{int}(\frac{256^2(\frac{Num_{(a3)}}{256^2} - Num_{(a2)})}{256} = \text{int}(\frac{Num_{(a3)}}{256} - 256Num_{(a2)})$$

$$Num_{(a0)} = \text{int}(\frac{256(\frac{Num_{(a2)}}{256} - Num_{(a1)})}{256^0} = \text{int}(Num_{(a2)} - 256Num_{(a1)})$$

[수식 6]

$$Num_{(a3)} = \text{int}(\frac{V_{Rest}}{64^3})$$

$$Num_{(a2)} = \text{int}(\frac{64^3(\frac{V_{Rest}}{64^3} - Num_{(a3)})}{64^2} = \text{int}(\frac{V_{Rest}}{64^2} - 64Num_{(a3)})$$

$$Num_{(a1)} = \text{int}(\frac{64^2(\frac{Num_{(a3)}}{64^2} - Num_{(a2)})}{64} = \text{int}(\frac{Num_{(a3)}}{64} - 64Num_{(a2)})$$

$$Num_{(a0)} = \text{int}(\frac{64(\frac{Num_{(a2)}}{64} - Num_{(a1)})}{64^0} = \text{int}(Num_{(a2)} - 64Num_{(a1)})$$

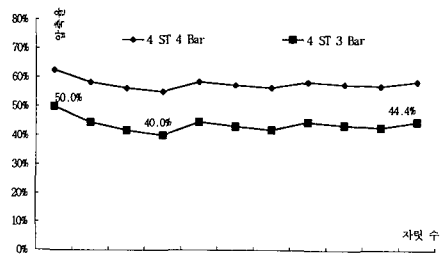
이와 같은 방법을 적용하여 한글문자의 정보수축 밀도를 높일 수 있다. 2 개의 한글 문자를 표현하기 위해서는 4-state 8, 6 바 기준으로 한글 한문자를 표현할 경우에 16-12 개 바가 필요하나 위와 같은 방법을 적용하면, 11 개 바에 의하여 표현될 수 있다.

3.4 비교 분석

앞 절에서는 숫자, 영문, 한글 정보를 바코드로 표현할 경우를 보인 것이다. 정보수축밀도 향상 방법 이전과 적용 후에는 압축이전의 바 수와 압축 후 바의 수의 비율 값을 구하면 된다.

$$CompRate(\%) = \frac{b_{o_{num}} - b_{num}}{b_{o_{num}}} \times 100$$

이와 같이 비율 값을 구하는 수식에 의하여 4-state 4, 3 바의 기준으로 정보 밀도의 압축비율을 계산한 후 비교한 결과는 (그림 3)과 같다.



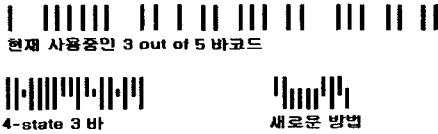
(그림 3) 숫자 정보 압축율(64 modulo)

<표 2> 4-state 3 바에 의한 우편번호 표현

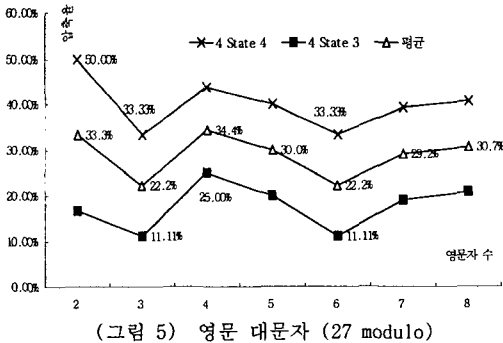
4-state 3 바	우편번호 (305-734)					
바의 값 B2	0	0	0	0	0	0
바의 값 B1	0	0	1	1	0	1
바의 값 B0	3	0	1	3	3	0

<표 3> 정보수축밀도향상 방법에 의한 우편번호 표현

우편번호(305-734)	자릿수			
	3	2	1	0
64 Modulo	1,166	10,642	41,09	6
압축된 10진수 값	1	10	41	6
바의 값 B2	X	0	2	0
바의 값 B1	X	2	2	1
바의 값 B0	1	2	1	2



(그림 4) 우편물 자동구분 처리용 바코드에 우편번호(305-734) 적용(예)



(그림 5) 영문 대문자 (27 modulo)

<표 4> 한글 문자(50 modulo)

표현 정보	조합형	완성형
한글 4 문자	44.44%	16.67%
한글 3 문자	37.04%	5.56%
한글 2 문자	38.89%	8.33%
평균 압축율	40.12%	10.19%

4-state 4 바는 영국의 RM4 BPO, 4-state 3 바는 캐나다, 일본 등에서 적용하여 사용 중인 4-state 3 바와 정보수축 밀도의 향상정도를 보인 것이다. (그림 4)는 현재 사용중인 바코드와 캐나다, 일본(압축방법 적용 안함) 등에서 적용된 바의 개수를 고려하여 비교한 것이다. 영국 및 캐나다는 숫자와 영문자를 혼합하여 사용하는 우편번호를 사용하는 특성이 있고, 캐나다의 경우에는 숫자 및 영문을 혼합하고, 이를 압축하여 적용하고 있다. 숫자 3 자리와 영문자 3 자리를 약 16.67%의 정보압축 방법으로 바코드 표현하여 적용하고 있다. 본 논문에서 제시한 방법을 4-state 3 바 단위로 숫자정보를 표현한 경우에 비교하면, 약 36 바를 20 개 바로 표현할 수 있으므로 16 개 바의 여유공간이 더 확보되고

(기존의 방법보다 압축 비율이 27.79%), 최소한 3 개의 숫자 또는 영문자 정보를 더 표현할 수 있게 되었다. 영문자와 한글문자의 경우에는 압축하지 않은 경우보다 22.3%와 10.19% 정도 압축되는 효과가 있다. 이러한 방법이 적용할 경우에는 적은 단위면적에 많은 정보를 수록할 수 있게 된다. 이에 따라, 자동구분을 위하여 인쇄되는 바의 총 수는 52 개 바에서 16 개 바를 더 인쇄하지 않아도 되므로 인쇄비용도 많이 축소될 수 있을 것으로 기대된다.

4. 맺음말

본 논문에서는 4-state 바코드에 정보수축밀도의 향상시키기 위한 방법을 적용한 것이다. 우편번호만 인쇄한다고 가정하면, 바코드의 인쇄량은 기존보다 약 우편물 2.5 통 단위로 1 통을 더 인쇄할 수 있는 효과가 발생된다. 자동구분하여 배달되는 우편물량이 30 억통 이상이므로 약 12 억통 이상의 우편물에 바코드를 추가로 인쇄할 수 있게 된다. 만일, 형광 바코드 인쇄비용이 1 원이라 가정하면, 30 억통의 우편물을 21.4 억원의 비용으로 바코드를 인쇄할 수 있게 되어 연간 8.6 억원이 절약되는 효과가 발생할 수 있다. 이 우편물량 중에서 다량우편이용자를 통하여 고객 바코드 인쇄물량이 40% 정도가 발생되고 있으며, 이는 우편고객이 우편물을 생성하기 위한 인쇄비용 축소의 요인으로 작용하게 된다. 또한, 인쇄용 프린터의 유지보수 주기를 연장시킬 수 있게 된다. 추후 연구사항으로는 고객 바코드를 활용하여 우편서비스 품질향상을 위한 반송우편물 자동처리, 배달 확인, 우편요금별납자자동처리 등이 가능한 바코드 정보체계의 수립 등에 관한 연구가 필요하다. 또한, 우편물류 처리뿐만 아니라 다양한 물류 정보처리를 위한 바코드로서 활용할 수 있도록 저가의 판독시스템 개발이 요구된다.

참고문헌

- [1] Donald Clysdale, "Canada Post Corporation's Point of Call Identifier," Advanced Technology Conference, Vol.2, pp.771-782, December 1992.
- [2] 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.
- [3] K.B Redersen, Hans Gartner, Walter S. Rosenbaum, "New Applications and Technology Trend in Postal Service," Advanced Technology Conference, pp.823-836. Vol.2, December 1992.
- [4] Masataka Nakagawa, Hiroshi Miyoshi, "Barcodes For Use in Mechanized Mail Processing," <http://www.iptp.go.jp/>, June 1995.
- [5] Royal Mail, "Mailsort Customer Barcoding Technical Specification," OXFORD, pp.1-11. April 1995.
- [6] 박문성, "우편물 자동처리 촉진을 위한 한글 4상 4바코드에 관한 연구," 한국정보처리학회 제2회 시스템통합연구회, pp.92-96. 1998.7.
- [7] 박문성, "우편용 한글 4-state 바코드," 제4회 우정 workshop, pp.139-152. 1998. 9.
- [8] 박문성, 송재관, 우동진 "우편물 자동처리를 위한 한글 4-state 바코드 시스템," 한국정보처리학회 논문지 제7권 제1호, pp.146-155. 2000. 1.