

주소해석 및 검색을 위한 우편주소파일 설계

(Design of Postal Address File for Address Interpretation and Retrieval)

장태우*, 김호연**, 임길택***

(Tai-Woo Chang, Ho-Yon Kim, Kil-Taek Lim)

요약 우편물을 배달순서대로 정렬하는 순로구분의 자동화를 위해서는 우편물에 표기된 주소를 인식해야 하며, 이 과정에서 표준화된 우편주소 데이터베이스는 필수적이다. 우편주소 데이터베이스는 주소인식뿐만 아니라 우편물 정보처리 등 다양한 측면에서 사용 가능하여 우편물 처리의 자동화와 정보화에 매우 중요하다. 본 연구는 우편주소 데이터베이스에서의 정보체계 표준을 위한 스키마 및 이에 따른 우편주소파일의 설계 방법과 설계된 스키마에 따른 주소해석 과정 및 방법에 관한 것이다. 주소를 표준화된 형태로 변환하는 데에 필요한 정보를 분석하여 이를 우편주소파일 설계과정에 포함시킴으로써 향후 순로구분 자동화를 위한 한글 주소인식 시스템과 고객 바코드 인증을 위한 주소 데이터베이스 매칭 및 주소검색 시스템에 활용할 수 있도록 하기 위해 우편주소파일을 설계하였다.

핵심주제어: 우편주소파일, 주소해석, 스키마 설계

Abstract In order to automate the process of mail sorting by delivery sequence, it is necessary to prepare a postal address database and to interpret written addresses on the mail-pieces with the database and OCR technology. The address database is a critical factor of automation and informatization of postal service since it could be used not only in address recognition but also in various mail processing. In this study, we design the schema of postal address database, design the postal address file based on it and explain the method of address interpretation and retrieval using it. We analyze information requirements for transformation of postal address into the standardized format and consider them in the process of design. The postal address file can be used by address matching or retrieval system as well as by Hangul address recognition system for automation of delivery sequence mail-sorting.

Key Words: postal address file, address interpretation, schema design

1. 서론

현재 우리나라에서는 전체 50억 정도의 우편물들이 전국 22개 우편집중국에서의 발송 및 도착 구분 작업의 자동화를 통해 집배우체국(이하 집배국)까지 전달된다. 집배국에서는 우편물을 배달경

로별로 구분하는 순로구분 작업이 수작업으로 이뤄지고 있으며, 이는 우편물 처리 비용 상승의 주요인 중의 하나로 지적되고 있어 순로구분 작업의 자동화에 대한 필요성이 대두되었다. 이에 따라 순로구분 자동처리 시스템이 개발되었으며, 이 시스템의 효율적 사용을 위해서는 우편번호와 배달점(delivery point) 및 순로 정보를 효율적으로 관리하기 위한 데이터베이스의 구축이 필요하다[6].

또한 우편물 구분 자동화를 위해서 반드시 해결

* 경기대학교 산업공학과 전임강사

** 한국전자통신연구원 우정기술연구센터 선임연구원/팀장

*** 경주대학교 컴퓨터멀티미디어공학부 조교수

해야 할 문제 중 하나가 배달지 주소 정보의 코드화이다. 코드화는 주소 문자를 자동으로 인식하거나 미리 표준화된 바코드로 인쇄된 코드 정보를 읽음으로써 가능하며, 이 중 어떤 방법을 이용하더라도 표준화된 주소 데이터베이스는 필수적이다. 전자에서는 인식된 주소 문자열을 표준 주소 문자열로 바꾸기 위해서, 후자에서는 기존의 주소 데이터베이스를 표준 주소 형식으로 변환하고 검색하기 위해서 표준 주소 데이터베이스를 이용해야 한다. 이처럼 표준화된 주소 데이터베이스는 다양한 변형을 갖고 표기된 주소를 표준화된 형태로 변환시킨 뒤 이를 고유의 코드로 바꾸는 과정에 필요하다.

일반적으로 우편물에 표기된 주소는 표준화된 주소와 다를 수 있다. 주소는 자연어와 같은 자유도를 갖지는 않으나 나름대로 변형과 생략이 가능한 문법을 갖는다. 이러한 정보는 표준화된 주소 데이터베이스에 포함되어 있어야 한다. 주소 표기의 문법은 해석이 가능한 형태로 분석되어 데이터베이스에 내재되어 있어야 한다. 본 연구에서는 표준화된 스키마를 따르는 일련의 인스턴스들을 우편주소파일(Postal Address File: PAF)이라 정의하며, 이에 대한 스키마를 정의하고 그 예를 설명한다. 우편주소파일의 설계 목적은 주소의 표준형 변환에 필요한 정보를 분석하고 이를 포함시킴으로써 향후 순로구분 자동화를 위한 한글 주소인식 시스템과 고객 바코드 인증을 위한 주소 데이터베이스 매칭 및 검색 시스템 등에 활용할 수 있도록 하기 위해서이다. 영국과 호주에서 사용하는 PAF와 목적은 같지만 한국적 상황에 적용하기 위한 스키마와 설계 개념에서 차이를 갖는다[8].

예를 들어 우편주소 자동인식 시스템에서 우편물에 기재된 주소의 내용을 그대로 인식한 결과를 주소인식 결과라고 하는데, 주소 데이터베이스를 참조하여 주소인식 결과를 해석하여 아래와 같이 표준화된 주소로 변환하게 된다.

- 주소인식 결과: 서울 종로구 을지3가 256-1번지 (우)100-193
- 주소해석 결과: 서울특별시 중구 을지로3가 256-1 (우)100-847

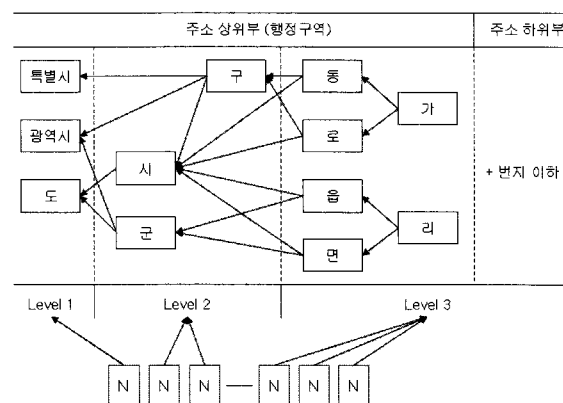
이하 2절에서 우편주소파일의 설계 개념과 8가지 주소 특징에 대한 처리 가능성에 대해 언급한

후, 3절에서 스키마의 설계와 우편주소파일을 이루는 각 테이블들의 관계에 대해 설명하고, 4절에서 우편주소파일을 사용한 주소 해석 및 검색 방법에 대해 간략히 설명한다. 5절에서 우편업무에 활용하기 위한 표준 우편주소파일의 확장에 대해 논하고 6절에서 결론을 맺는다.

2. 주소 체계에 따른 설계 개념

2.1 우리나라 주소 체계와 표현의 다양성

우리나라의 우편주소 체계는 도로(Street) 중심의 서양 주소체계와 달리 시, 도, 구, 군, 동, 읍, 면, 리 등의 기본 행정구역을 세분화하여 포함하고 있으며, 기본 행정구역 내에서 구역을 기준으로 번지를 할당하여 주소를 결정하는 지번 체계이다. 우편번호 역시 구역과 지번그룹, 다량배달처 등으로 세분화하여 구성되어 있다. 이러한 우편주소 체계를 구조화하여 살펴보면 (그림 1)과 같다[1][4].



(그림 1) 한국의 우편주소 체계

(그림 1)에서 특별시, 광역시, 도를 Level 1로, 구, 시, 군을 Level 2, 동, 가, 로, 읍, 면, 리를 Level 3로 표현한다. Level 1은 우편번호 첫째자리에, Level 2는 둘째와 셋째자리에, Level 3은 뒤세 자리에 대응된다[5].

우편물의 순로구분 작업을 자동화하기 위해서는 우편물에 표기된 주소를 자동 인식해야 하는데 우편주소는 <표 1>과 같은 특징들이 있어 주소해석 과정이 필요하며, 다양한 표현 형태를 수용하여 데

<표 1> 우편주소의 다양한 특징과 예

특 징	예 (표준형)	예 (변형)
① 계층적 행정구역 체계	서울특별시 종로구 을지로3가	서울 종로 을지로3가
② 법정지명, 행정지명의 세분화 및 대표 지역 명칭 사용	인후동1가, 인후동2가	인후1동, 인후2동, 인후3동, 인후동
③ 택지개발 또는 공단 지역의 가 지번 형태의 주소 존재	남동공단 37블록 49로트	남동공단 37B49L
④ 사서함의 사용	대전유성우체국사서함 106호	대전 유성 PO Box 106
⑤ 건물, 아파트 등의 표기 형태가 다양	목동7단지, 한빛타운아파트	목동@, 한빛Apt.
⑥ 건물, 아파트 등의 동, 호 수 등의 세부 주소 형태가 다양	710동 1234호	710-1234
⑦ 번지 표기 형태가 다양	256번지 1호, 256-1	256-1번지, 256의 1, 산4-1
⑧ 동일 지점에 대한 유사 명칭 사용	한국과학기술원	과기원

이터베이스로 저장한 후 검색할 필요성 역시 존재한다. 이 외에도 더 이상 존재하지 않는 건물 또는 번지 등을 일반인들이 사용할 경우를 위해 데이터를 저장하는 것을 고려할 필요도 있다. 이를 위해 우편주소파일을 위한 데이터베이스 스키마를 다음과 같은 주요 개념을 가지고 설계한다.

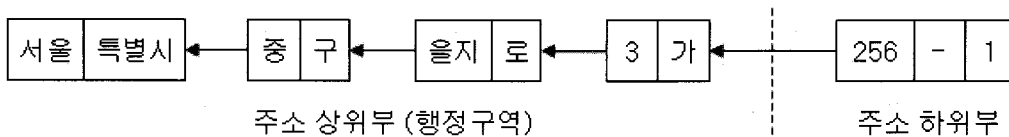
2.2 설계 개념

주소해석을 위해서는 한국의 일반적인 주소체계로부터 주소정보를 정형화시켜 데이터베이스로 표현하는 과정이 우선 필요하다. 앞서 언급한 바와 같이 행정구역과 지번 기준으로 구성되는 우리나라의 주소체계는 하나의 구역이 다시 작은 구역으로 나뉘기도 하고, 동일한 번지를 갖는 구역 내에 여러 개의 업체가 존재하거나 다수의 세대가 거주하기도 한다. 예를 들어 아파트와 아파트 단지, 대형 빌딩은 동일 번지 내에 다수의 배달지점이 존재하는 경우이다. 이러한 국내 상황과 앞 절에서 언급한 특징들을 고려하여 데이터 표현의 중복을 피하면서 구분이 필요한 최종 배달지점까지 코드

가 부여될 수 있도록 우편주소파일이 설계되어야 한다. 따라서 본 연구에서는 다음과 같은 설계의 핵심 개념을 세우고 주소파일을 설계하였다.

1. 계층적 행정 구역의 일관성 있는 표현
2. 번지가 포함된 주 배달점(Primary Delivery Point)과 번지 내에서 세분화되는 배달점의 표현
3. 주소 표기의 변형 흡수를 위한 키워드와 동의어 표현

첫 번째는 <표 1>의 특징 ①, ②에 대해 우리나라 주소의 계층적 구조를 반영하되 그 계층의 깊이(Depth)가 일정하지 않은 것은 보완할 수 있는 표현을 목표로 하였다. 우리나라 주소는 (그림 1)의 행정구역 계층에서 볼 수 있는 바와 같이 순서 및 깊이가 일정하지 않은 구조를 가지고 있다. 이를 일관성 있게 표현하려면 주소의 최하위 행정구역부터 시작하여 최상위 행정구역까지 단계별로 나타내는 형식을 취하면 된다. 여기에서 최하위 행정구역이란 번지 바로 앞의 단위 행정구역을 말한다. 최하위 행정구역 다음에는 번지 이하의 정보가



(그림 2) 주소 문자열의 계층적 연결

오게 될 것이다. (그림 2)는 실제 주소를 예로 들어 행정구역 부분과 번지 부분으로 분리한 후 이를 어떻게 계층적으로 표현하는지 보여주고 있다. 주어진 주소가 '서울특별시 종로구 을지로3가 256-1'일 때 행정구역은 256-1이라는 번지부 앞의 3가에서부터 시작하여 서울특별시까지 연결 리스트와 같이 표현된다.

이와 같이 가장 작은 단위의 행정구역부터 시작하여 상위 행정구역을 따라가는 순환계층 구조를 효과적으로 표현할 수 있도록 데이터베이스의 형식을 정의하면 데이터의 중복을 피하면서 주소를 쉽게 표현할 수 있는 주소파일을 만들 수 있다.

두 번째 설계 개념은 주 배달점과 상세 배달점을 구분하는 것이다. 우리나라의 상황에 맞게 하기 위해 주 배달점은 번지에 상응하는 배달점으로 정의하였고 상세 배달점은 주 배달점만으로는 구분되지 않는 배달점으로 정의하였다. 고유 번지와 단독주택이건, 공동주택이건, 사무실이 모여 있는 빌딩이건 상관없이 그 번지에 해당하는 건물을 모두 통틀어서 주 배달점으로 정의하여 하나의 테이블로 표현하였으며, 그에 속한 상세 배달점들은 주 배달점에 대한 연결 키를 필드로 포함하는 하나의 다른 테이블로 표현하였다. 세부 설계 과정에서 특징 ③, ④도 반영할 수 있도록 하였다.

세 번째 설계 개념은 우편주소파일을 주소해석 혹은 주소 문자열 분석에 이용할 수 있도록 설계하는 것이다. 즉, 주소 문자열에 나타날 수 있는 키워드와 동의어를 따로 정의함으로써 이의 변형이나 생략을 일관성 있는 규칙 적용에 의해 처리할 수 있도록 만드는 것으로 <표 1>의 나머지 특징들(⑤~⑧)을 반영할 수 있다. 키워드는 행정구역 단위나 공통으로 나타나는 건물 등의 단위를 위한 것이며, 동의어는 동일 지점에 대한 유사 명칭 사용 등을 위한 것이다. 이는 생략형으로 표기

되거나 유사 명칭으로 표기되는 것과 같이 표준명칭 및 표준형 행정구역 단위로 표기되어 있지 않은 주소 문자열을 인식하여 최종 배달점 정보로 코드화하기 위해서 필요하다.

이러한 세 가지 주된 개념에 대해 총 5개의 테이블을 구성하며, 각각의 역할은 <표 2>와 같고 각 테이블의 구조와 필드별 의미는 3절에서 설명한다.

3. 우편주소파일 구조

아래와 같이 일반적으로 사용되는 네 가지의 주소들을 예로 PAF로 구성하기 위한 설계 방법을 설명한다.

Addr1 (빌딩): 서울특별시 영등포구 여의도동 17-1 금산빌딩 우편구역

Addr2 (공단): 인천 남동구 남동공단 37블록 49로트 우정기계공업

Addr3 (아파트): 경기도 성남시 분당구 서현동 대림한빛APT 103동 222호

Addr4 (복합건물): 서울 종로구 장사동 173번지 세운상가 가동 2층 바열 217호

3.1 행정구역 (ADMIN_DIST) 테이블

ADMIN_DIST는 (그림 2)에서 본 것과 같이 주소 상위부에 대해 개별 행정구역 단위를 계층적으로 구성하여 표현하기 위한 테이블이며, 각 레코드는 고유 번호(ADMIN_DIST_ID)로써 구분된다.

Addr1과 같은 경우 특별시, 구, 동의 세 가지 단위 행정구역이 포함되어 있다. 이를 표현하기 위해 세 개의 레코드가 필요하게 되는데, 각 레코드

<표 2> 우편주소파일을 구성하는 테이블 설명

테이블 명	설명 및 기능
ADMIN_DIST	계층적인 행정구역(Administrative District)을 표현하기 위한 테이블
PDP	번지 단위로 부여된 주 배달점(Primary Delivery Point)을 표현하기 위한 테이블
DDP	번지만으로 표현되지 않는 상세 배달점(Detailed Delivery Point)을 표현하기 위한 테이블
KEYWORD	주소에서 사용되는 키워드(Keyword)를 정의한 테이블
SYNONYM	주소에서 사용되는 동의어(Synonym) 또는 유사어를 정의한 테이블

ADMIN_DIST_ID	ADMIN_DIST_NAME	ADMIN_DIST_TYPE	HIGHER_ADMIN_DIST_ID	UPPER_POSTCODE	LOWER_POSTCODE
345678901	서울	특	00000000	N/A	N/A
...					
234567890	영등포	구	345678901	150	N/A
...					
123456789	여의도	동	234567890	150	010

(그림 3) 계층적 행정구역 표현의 레코드 뷰

는 '서울+특별시', '영등포+구', '여의도+동'과 같이 행정구역 이름(ADMIN_DIST_NAME)과 유형(ADMIN_DIST_TYPE)으로 분리되어 저장된다. 가능한 ADMIN_DIST_TYPE은 '특(특별시), 광(광역시), 도, 시, 구, 군, 동, 읍, 면, 리, 로, 가'이다. 계층상의 하위 행정구역의 경우 상위 행정구역을 가리키기 위한 필드(HIGHER_ADMIN_DIST_ID)를 가지게 되는데, *Addr1*을 레코드 뷰로 보면 (그림 3)과 같다. 참고로, 서울특별시와 6대 광역시(부산, 대구, 인천, 대전, 광주, 울산), 9개 도(강원, 경기, 경북, 경남, 전북, 전남, 충북, 충남, 제주)는 상위 행정구역이 없으므로 HIGHER_ADMIN_DIST_ID를 00000000으로 설정한다.

(그림 1)에서 설명한 바와 같이 행정구역은 우편번호와 밀접한 관계를 가지는데 Level1과 Level2의 시군구까지는 우편번호 상위 3자리와 일치한다. 이를 위한 필드(UPPER_POSTCODE)를 하나 두고 구체적으로 우편번호가 지정될 수 없는 곳은 N/A로 기록한다. *Addr1*에서 서울특별시의 경우는 상위 3자리 중 첫 번째 수인 1을 제외하면 서울특별시 전체를 대표할 수 있는 코드가 없기 때문에 상위 3자리가 N/A로 기록되었다. 그러나 그 하위 행정구역인 영등포구와 여의도동은 모두 150으로 동일한 값을 가질 수 있다.

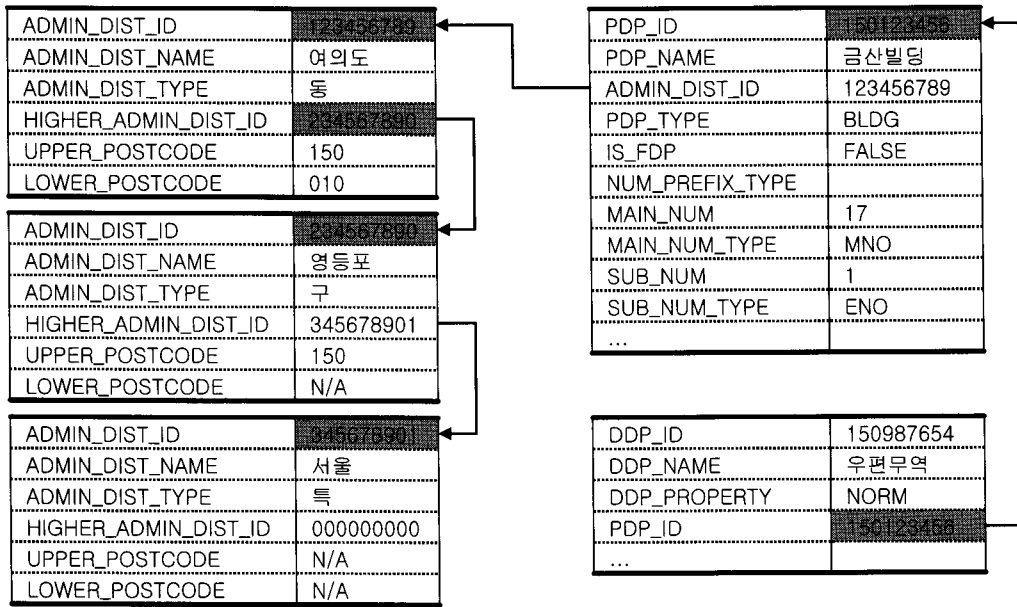
우편번호 6자리 중 하위 3자리 역시 읍면동 단위의 행정구역을 일부 가리키기도 하므로 하나의 필드(LOWER_POSTCODE)를 할당하고 구체적으로 우편번호가 지정될 수 없는 곳은 N/A로 기록된다. *Addr1*에서 영등포구에 소속된 하위 행정구역(동)의 하위 3자리는 서로 동일하지 않기 때문에 특정한 값을 지정해 줄 수가 없으므로 N/A로 지정되었다. 하위 행정구역은 우편번호 하위 3자리가 모두 지정될 수 있다. 만일 하나의 동에 번지에 따라 우편번호가 달리 부여되어 여러 개의 우편번호

가 사용되고 있다면 그 동을 대표하는 우편번호를 기입하도록 한다.

3.2 주 배달점 (PDP) 테이블

PDP는 단위 행정구역 내 번지 중심의 주 배달점을 정의하기 위한 테이블로 소속된 행정구역을 가리키기 위한 필드(ADMIN_DIST_ID)를 가진다. PDP는 번지를 기준으로 정의되므로 아파트, 빌딩, 단독주택, 상가 등과 사서함이 PDP 수준에서 표기될 수 있으며, 이를 대표하는 공식적인 이름은 별도의 필드(PDP_NAME)로 저장하고 일반적으로 사용되는 다른 명칭이 존재할 경우는 동의어(Synonym) 테이블에 저장한다. 주소 문자열에서 '아파트', '빌딩' 등은 변형이 가능하기 때문에 키워드가 따로 정의된 단어는 PDP_NAME에서는 생략해야 하며 '하이츠', '맨션'과 같이 변형이 없어서 키워드를 따로 정의하지 않은 단어는 포함하여 저장한다. 이러한 아파트, 공동주택, 빌딩, 단독주택 등의 배달점 유형과 교육기관과 관련한 유형은 추가 필드(PDP_TYPE)에 정의한다. 또한 아파트, 빌딩과 같은 경우 배달 업무에 있어서는 최종 배달점이 아니므로 상세 배달점을 추가해야 하는데, 최종 배달점 여부는 IS_FDP (Is a Final Delivery Point?) 필드의 값에 의해서 결정한다. IS_FDP의 값이 'N'이면 PDP_ID를 외부 참조키로 가지는 상세 배달점(DDP) 테이블의 레코드들이 존재한다.

번지부의 표현에 대해서도 번지부 앞의 문자열, 주 번지, 부 번지, 구성요소들의 각 문자열을 분리하여 저장하도록 한다. 먼저 일반 번지와 번지부 앞에 나타나는 산 번지, 사서함은 각각 NOR, SAN, POT를 값으로 하는 필드(NUM_PREFIX_TYPE)로 구분한다. 또한 번지부의 번지와 호(부 번지)는 각각 두 가지의 필드



(그림 4) ADMIN_DIST, PDP, DDP 간 연결 관계의 예

(MAIN_NUM, SUB_NUM)로 구분하며, ‘번지’, ‘호’, ‘의’, ‘-’와 같은 구성요소 역시 각각 두 가지의 필드(MAIN_NUM_TYPE, SUB_NUM_TYPE)로 구분한다. 공단 및 개발 지구의 블록, 로트와 같이 지번 개념과 유사한 형태로 사용되는 번지부 다음의 정보를 가리키도록 추가 필드(INFO_NUM1, INFO_NUM2, INFO_NUM3)를 구성하며, 이의 유형에 대한 필드(INFO_NUM1_TYPE, INFO_NUM2_TYPE, INFO_NUM3_TYPE)도 추가한다.

이와 같은 설계에 대해 예를 들면, **Addr1**의 경우 여의도동까지는 행정구역 계층에서 표현되고 ‘17-1번지 금산빌딩’이 주 배달점이 된다. 이 중 ‘금산’은 PDP_NAME에, ‘빌딩’에 해당하는 ‘BLDG’는 PDP_TYPE에 저장된다. 그리고 번지부에서 ‘17’, ‘-’, ‘1’은 각각 MAIN_NUM, MAIN_NUM_TYPE, SUB_NUM에 저장된다. **Addr2**의 경우 ‘37’, ‘블록’, ‘49’, ‘로트’는 각각 INFO_NUM1, INFO_NUM1_TYPE, INFO_NUM2, INFO_NUM2_TYPE에 해당된다. 추가 변형 번지구조를 위해 INFO_NUM3, INFO_NUM3_TYPE 필드를 추가한다.

3.3 상세 배달점 (DDP) 테이블

DDP는 주 배달점이 최종 배달점이 아닌 경우의 상세 배달점들을 가리킬 때 사용된다. 예를 들어서 **Addr1**의 경우 ‘금산빌딩’까지는 주 배달점이 되고 우편무역은 상세 배달점이 되며, **Addr3**과 같은 경우에 ‘대림한빛APT’까지는 주 배달점이 되고 103동 222호가 상세 배달점이 된다. 즉, 하나의 번지에 상응하는 다량배달처가 주 배달점이 되고, 해당 주 배달점 건물에 속한 모든 세대, 상점 또는 부서 등이 상세 배달점이 된다. **Addr1**의 ‘우편무역’과 같이 배달점의 공식적인 명칭은 상호명, 회사명, 부서명 등을 위한 필드(DDP_NAME)에 저장한다. 우편업무를 위해 배달점의 특성을 지정할 필요가 있는데, 이를 위해 우편물 배달이 가능한지, 배달과 관련하여 다른 제약사항이 있지 않은지 확인하기 위한 필드(DDP_PROPERTY)를 추가로 둔다.

상세 배달점이 속한 주 배달점의 ID를 필드(PDP_ID)로 묶으로써 행정구역까지의 전체 주소를 유추할 수 있다. **Addr1**을 예로 들 경우 (그림 4)와 같이 행정구역, 주 배달점, 상세 배달점이 연결 관계를 갖게 된다.

상세 배달점에 추가되는 주소정보는 그 유형과 함께 최대 4개까지 추가할 수 있도록 모두 8개의 필드를 둔다. 복합 건물의 경우 등에 해당되는데, 예를 들어 **Addr4**의 경우는 다음과 같이 각 정보

를 할당해 줄 수 있다. 각 유형에 대한 코드는 부록에 별도로 정의한다.

DTL_INFO1 = “가”
 DTL_INFO2 = “2”
 DTL_INFO3 = “바”
 DTL_INFO4 = “217”
 DTL_INFO1_TYPE = “BDN”
 DTL_INFO2_TYPE = “LVL”
 DTL_INFO3_TYPE = “YOUL”
 DTL_INFO4_TYPE = “HO”

3.4 키워드 (KEYWORD) 테이블

키워드는 주소에 나타나는 문자나 단어 중에서 단순한 지명이 아니면서 특별한 의미를 갖되 규칙에 의해서 생략이 가능하거나 표기 방법이 다양한 것을 의미한다. 대표적인 예로 행정구역 명칭에 나타나는 키워드와 건물명에 나타나는 키워드를 들 수 있으며, **Addr3** 문자열을 예로 들 경우 ‘도, 시, 구, 동, APT, 동, 호’의 키워드를 분리할 수 있다.

이처럼 주소요소를 고유명과 키워드로 나누는 이유는 키워드가 개별적인 의미를 갖기 때문에 키워드별로 변형이나 생략 규칙이 정해지기 때문이다. 예를 들어, ‘성남시’, ‘수원시’, ‘전주시’ 등에 나타난 키워드 ‘시’가 생략되어 ‘성남’, ‘수원’, ‘전주’와 같이 표기되기도 하고 ‘대구광역시’, ‘광주광역시’가 ‘대구시’, ‘광주시’ 혹은 ‘대구’, ‘광주’와 같이 다르게 표기될 수 있는 것은 키워드의 의미가 다

르고 이에 따라 다른 규칙이 적용되기 때문이라고 할 수 있다. 마찬가지로 ‘대림한빛아파트’가 ‘대림한빛Apt’ 혹은 ‘대림한빛@’ 등과 같이 다양하게 표기될 수 있는데, 이를 일반적인 형태로 정의하여 모든 아파트 표기에 적용하기 위해서 키워드가 필요하다고 할 수 있다.

키워드는 각각의 유형을 정하여 필드(KWD_TYPE)를 할당하였고 이들을 행정구역인지 배달점인지 등에 따라 범주화하여 별도의 필드(KWD_CATEGORY)에 저장하도록 하였으며, 사용 예를 들면 <표 3>과 같다. 만약 키워드가 행정구역을 의미한다면 KWD_CATEGORY의 값은 “ADT”가 되고, KWD_TYPE은 ADMIN_DIST_TYPE이 참조하게 된다.

각 키워드의 타입에 대해 예를 들면 <표 4>와 같으며, 네 가지의 각 레코드에 대한 설명은 아래와 같다.

1. 해당 키워드가 주 배달점(PDP)에서 이용되는 것이고, 그 주 배달점의 타입(PDP_TYPE)이 아파트(APTM)인 경우 키워드로 @가 쓰일 수 있으며, 그 사용 빈도(속성)는 높지 않음을 (SLDM) 의미한다.
2. 해당 키워드가 번지의 점두사로 이용되는 것이고, 그 타입이 산 번지(S)인 경우 키워드로 ‘산’이 쓰일 수 있으며, 그 사용 속성은 표준형(STD)임을 의미한다.
3. 해당 키워드가 번지부(DNT) 본번(MNO)의 인자이며 그 인자가 “-”인 경우를 가리킨다. 또한 그 속성이 USAL(USUAL), 즉 통상적으로 많이 쓰임을 나타낸다.

<표 3> KWD_CATEGORY의 유형과 관련 필드 및 사용 예

값	설명	관련 필드	예
ADT	행정구역의 타입을 의미한다.	ADMIN_DIST.ADMIN_DIST_TYPE	시, 구, 동
PDT	해당 레코드의 키워드가 주 배달점에 관한 키워드임을 의미한다.	PDP.PDP_TYPE	(주), 아파트, 빌딩
DNT	해당 키워드가 주 배달점의 번지부나 상세 배달점의 번호부를 나타내는 인자임을 의미한다. 다시 말하면, 주 배달점이나 상세 배달점의 본번, 부번, 블록, 로트, 동, 층, 열, 호 등을 위한 키워드임을 가리킨다.	PDP.MAIN_NUM_TYPE, PDP.SUB_NUM_TYPE, DDP.INFO_NUM#_TYPE, DDP.DTL_INFO#_TYPE	-, 산, 블록, 로트, 번지, 호
NPT	해당 레코드의 키워드가 번지 등의 숫자 앞에 나타나는 문자를 가리킨다.	PDP.NUM_PREFIX_TYPE	산, 사서함

<표 4> 키워드 타입에 대한 예

예		KWD_CATEGORY	KWD_TYPE	KWD_STRING	KWD_PROPERTY
1	대림한빛@ 103-201	PDT	APTM	@	SLDM
2	산 129-1번지	NPT	SAN	산	STD
3	103-201	DNT	MNO	-	USAL
4	103-201	DNT	ENO	Null	USAL

4. 해당 키워드가 번지부(DNT) 부번(ENO)의 인자에 해당하며 통상적으로 생략될 수 있음(Null)을 나타낸다.

주소 표기 시 키워드가 실제로 사용되는 값은 KWD_STRING 필드에 저장되는데, 예를 들어 해당 키워드가 APT인 경우 실제로 사용되는 문자열은 “Apt”, “@”, “A”, “아파트” 등이 될 것이다. 이러한 다양한 형식을 여러 개의 레코드를 이용하여 나열함으로써 표현할 수 있다. 또한 키워드가 어떻게 사용되고 있는지의 속성을 나타내기 위해 별도의 필드(KWD_PROPERTY)를 두고 각 키워드의 사용빈도에 대한 정보를 담음으로써 실제 주소 매칭 과정에서 활용할 수 있도록 하였다. 그 속성들은 STD, USAL, SLDM, UNUS의 네 가지 경우로 표현된다. STD는 Standard의 약자로 규정된 표준형을 의미한다. USAL은 Usual의 약자로, 표준형은 아니지만 사람들에 의해 자주 사용되는 표기인 경우를 말한다. SLDM은 Seldom의 약자로, 간혹 사용되는 경우를 말한다. UNUS는 Unusual의 약자로 SLDM보다 사용빈도가 낮아 별로 사용되지 않는 경우를 말한다. 예를 들어, 인자가 “아파트”인 경우, 표준형은 “아파트” 이므로 STD의 속성을 갖는다고 말할 수 있다. “Apt”는 통상 자주 사용되므로 USAL의 속성을 갖는다. “@”과 “A”는 각각 그 사용빈도상 UNUS와 SLDM을 갖는다고 정의할 수 있다.

키워드가 따로 정의되어 있고 문자열이 KWD_STRING에 선언되어 있으면 테이블에 고유명사를 입력할 때 키워드 부분을 제외하고 입력해야 한다. 예를 들어, 대림한빛아파트를 입력할 때 PDP_NAME은 ‘대림한빛’이 되어야 한다.

3.5 동의어 (SYNONYM)

SYNONYM은 동일 지역이나 기관, 건물을 지칭하는 여러 가지 이름을 정의하기 위해서 사용되는 테이블이다. 예를 들면, “충청남도”를 “충남”이라고 하거나, “한국과학기술원”을 “과학원”, “과기원”, 혹은 “과학기술원”, “KAIST” 등으로 표기하는 경우에 이를 표현하기 위해서 사용한다. 이러한 동의어들을 따로 정의하는 이유는 다양한 형태의 주소 표기들을 모두 처리할 수 있도록 하기 위해서이며, 정의되지 않은 단어는 주소해석 및 검색이 불가능하다.

본 테이블 역시 키워드와 유사하게 동의어의 유형을 정의하는 필드(SYNONYM_TYPE)를 둔다. 행정구역 명칭의 동의어는 ADS 유형으로, 주 배달점, 상세 배달점에 대한 동의어들을 각각 PDS, DPS 유형으로 정의한다.

정의되는 동의어에 대한 실제 표준명칭은 행정구역 또는 주 배달점, 상세 배달점에서 참조하도록 각 테이블을 참조하는 ID 필드(SYNONYM_ID)와 상기 정의한 유형을 통해 얻을 수 있다. 즉 유형별로 사용되는 SYNONYM_ID는 관련 테이블의 해당 ID를 이용하는데, 표준명칭을 지칭하는 모든 ID들은 각각 자기의 고유한 ID를 가지므로 중복을 피할 수 있기 때문에 그대로 사용된다. 하나의 표준명칭에 여러 개의 동의어가 정의될 수 있으므로 이 필드 값은 테이블간의 관계키로 활용된다.

주어진 SYNONYM_ID에 대응하여 실제 사용되는 여러 개의 동의어는 각 레코드의 SYNONYM_STRING 필드에 저장된다. 하나의 표준명칭에 대해 여러 개의 동의어가 존재할 수 있으므로 동의어는 같은 SYNONYM_ID를 갖는 여러 개의 레코드들에 의해 표현될 수 있다. 정의된 동의어의 사용빈도 등과 같은 특성을 나타내는 필드(SYNONYM_PROPERTY)와 함께 예를 들면 <표 5>와 같이 표현된다. 현재 정의된 속성은 부록에 첨부한다.

<표 5> SYNONYM 테이블의 예

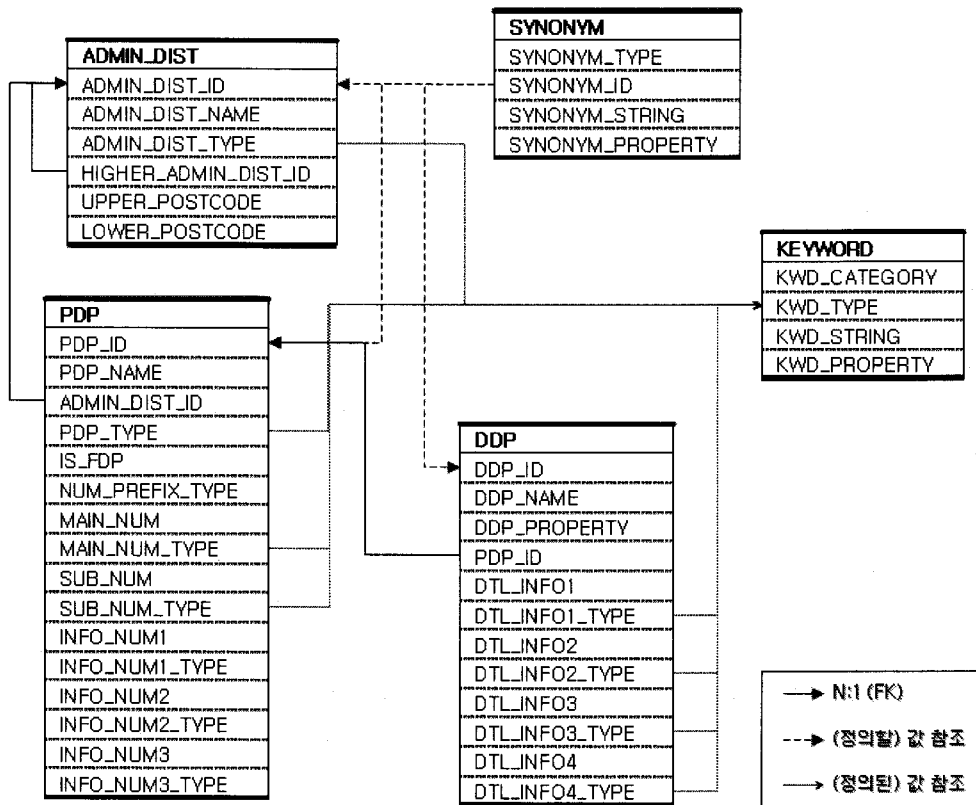
SYNONYM_TYPE	SYNONYM_ID	SYNONYM_STRING	SYNONYM_PROPERTY	표준형
ADS	342523828	충남	USAL	충청남
PDS	436258734	과학기술원	USAL	한국과학기술원
PDS	436258734	과학원	UNUS	
PDS	436258734	과기원	SLDM	
PDS	436258734	KAIST	USAL	
PDS	436258734	카이스트	SLDM	
DPS	236782655	한마음치과	USAL	한마음새로움치과
DPS	236782655	한새치과	SLDM	

3.6 관계도

이상에서 설명한 행정구역(ADMIN_DIST), 주배달점(PDP), 상세 배달점(DDP), 키워드(KEYWORD), 동의어(SYNONYM)의 5개로 구성된 우편주소파일의 테이블들 간의 관계를 나타내면 (그림 5)와 같다.

(그림 5)에서 실선 화살표는 외부 키(foreign key)를 나타내며 개체(테이블)간의 관계차수(cardinality)가 N:1이 됨을 의미한다. HIGHER_ADMIN_DIST_ID는 상위 행정구역의 ADMIN_DIST_ID를 가리키도록 되어 있다.

긴 점선 화살표는 정의할 값을 참조하는 것으로 SYNONYM_ID가 동의어의 유형에 따라 ADMIN_DIST_ID, PDP_ID, DDP_ID의 값들 중



(그림 5) 우편주소파일 테이블간의 관계도

하나를 이용한다는 것을 의미한다. 또한 점선 화살표는 정의된 값을 참조하는 것으로 KEYWORD 테이블에 이미 정의되어 있는 값을 ADMIN_DIST, PDP, DDP 테이블의 각 TYPE에 해당하는 필드에서 참조하는 것을 의미한다. 예를 들어, PDP 테이블의 PDP_TYPE 필드는 키워드 테이블의 KWD_TYPE과 연결되고 그 관계는 일대다의 관계를 갖는다. 즉 주 배달점의 타입은 주 배달점의 인자 값인 키워드와의 관계를 이용해 표현된다.

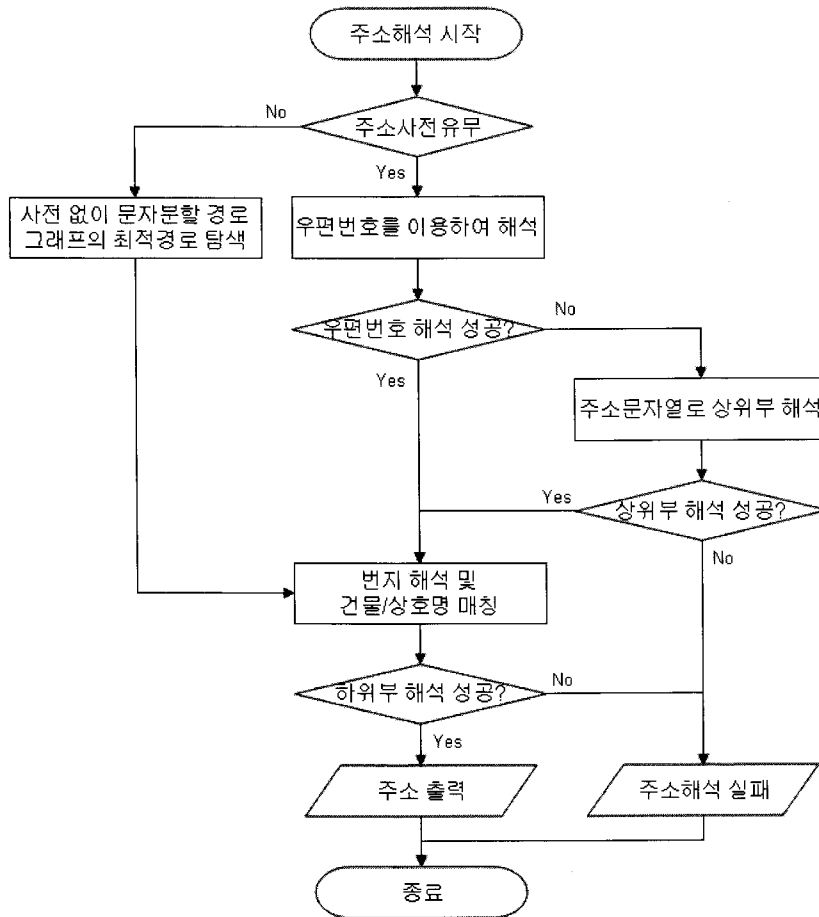
4. PAF를 이용한 주소해석 및 검색

4.1 주소해석

주소해석은 주소가 포함된 문자열을 입력으로

받아 유효한 실제 주소지 정보를 추출하는 것으로, 영상처리 및 문자인식 과정의 결과물들을 이용하여 우편번호와 주소에 대한 문자열을 출력하게 된다. 이때 주소매칭 시의 속도 향상을 위해 PAF 형태의 주소DB로부터 생성한 Tree 구조의 주소(행정구역) 사전을 사용하게 된다. 참고로, 우편번호 사건의 경우는 각 자리 숫자들의 조합을 통한 후보 생성을 위해 Trie 형태로 구성되어 있다. 우편번호 등의 주소인식 과정은 크게 연결 요소 기반 문자열 추출, 문자분할, 문자인식으로, 이중 문자인식 과정은 화소밀도와 방향각도를 이용한 특징 추출, 신경망을 이용한 형식 분류(한글 및 기타 문자 구분 포함), 문자분할 그래프 생성 등으로 구분된다.

주소해석 과정은 (그림 6)과 같다. 먼저 우편번호 인식을 시도하고 그 결과에 따라 상위부 매칭을 시도한다. 우편번호를 이용한 해석을 먼저 시도



(그림 6) 주소해석 과정

하는 이유는 해석의 속도를 빠르게 하기 위해서이다. 실제로 우편번호 인식에 실패하더라도 주소사전을 이용하면 상위부 매칭에 성공할 가능성은 높다. 주소의 상위부, 즉 우편번호에 해당하는 행정구역명까지를 인식한 다음에는 번지와 건물명 등의 주소 하위부를 해석하고 최종 결과를 출력한다 [2].

참고로 우편번호와 상위부 검색 부분은 ADMIN_DIST 테이블의 내용을, 번지 해석 및 건물/상호명 매칭 부분은 PDP 테이블의 것을 이용한다.

4.2 주소검색

전국적인 주소 데이터베이스가 PAF 형식에 따라 구축되어 있는 경우 지능적인 검색이 가능하다. 건물명, 기관명과 같은 하나의 단어를 가지고 검색하거나 계층적인 주소 체계에 따라 검색할 수도 있으며, 연산자 등을 사용해서 검색 결과를 한정시킬 수 있다. 본 연구에서는 앞서 설명한 PAF 형식에 따른 효율적 검색에 대한 예를 보이기 위해 웹 기반의 검색시스템을 구현했으며, 기능 측면에서 간략히 설명한다.

먼저 하나의 단어를 사용할 때에는 행정구역(ADMIN_DIST)과 동의어(SYNONYM) 위주로 검색하게 되며, 동의어 검색의 결과는 최종 배달점까지 출력하게 된다. 연산자(포함, 논리 연산자 등)나 와일드카드(*, ?)를 이용한 부분 검색과 계층적 주소 체계에 따른 문자열 완전 검색에 대한 예는 <표 6>과 같다.

검색시스템은 Tomcat 환경에서 Axis를 이용한

웹서비스로 구현되었는데, 검색 결과는 SOAP(Simple Object Access Protocol) 서버로부터 XML 형태로 받게 된다. 구조화되고 계층적인 XML 형태로 처리함으로써 DB와의 연계가 용이하며 웹서비스를 통해 타 시스템과의 연계도 용이하게 된다. 즉 웹서비스를 사용할 수 있는 정부 및 기업들의 주소검색 및 우편번호검색 서비스에 다양한 검색 방법을 제공해줄 수 있게 된다.

<표 6>에서 '*학원 IN 유성 봉산'을 입력하여 검색한 결과 다운받게 되는 문서의 예는 (그림 7)과 같다. 우편주소파일의 ADMIN_DIST, PDP, DDP 테이블의 해당 부분과 우편번호(POSTCODE)가 XML 문서의 각 ELEMENT로 구성되며, 여러 개의 결과는 각 우편번호, 행정구역, 배달점이 ADDRESS 요소를 구성하여 나타나게 된다. 이러한 정보요소별 계층 구조는 문서에 포함된 DTD에 명시되어 있다.

5. 우편업무 활용을 위한 PAF의 확장

이상과 같이 살펴본 우편주소파일은 주소해석 및 검색 시스템에서의 활용뿐만 아니라 실제 우편 업무에 활용하기 위해 다음과 같은 사항들을 고려하여 확장할 필요가 있다.

- 집배구 및 순로 정보와의 연계 : PDP 및 DDP에 집배구와 순로 정보를 연계하도록 하여 자동 순로구분이 가능하도록 해야 한다. 주소해석을 위한 사전 생성 외에 자동 순로구분기의 작동을 위한 순로파일을 작성해야 하는 것이다.

<표 6> 다양한 검색 방법과 예

방법	입력	결과
하나의 단어를 이용한 기본 검색	유성	대전광역시 유성구 가정동 대전광역시 유성구 갑동 ...
	KAIST	대전광역시 유성구 구성동 373-1번지 한국과학기술원
연산자 및 와일드카드를 이용한 부분 검색	대전 유성 창조*@	대전광역시 유성구 반석동 349번지 창조대군인아파트
	*학원 IN 유성 봉산	대전광역시 유성구 봉산동 181-16번지 한가람바독학원 101호 대전광역시 유성구 봉산동 181-16번지 한가람바독학원 102호 ...
주소 문자열을 이용한 완전 검색	유성 봉산 181-16	대전광역시 유성구 봉산동 181-16번지 한가람바독학원

```

<?xml version="1.0" encoding="EUC-KR"?>
<!DOCTYPE KPOST [
<!ELEMENT KPOST (ADDRESS)*>
<!ELEMENT ADDRESS (POSTCODE, ADMINDIST, PDP, DDP)>
<!ELEMENT POSTCODE (#PCDATA)>
<!ELEMENT ADMINDIST (#PCDATA)>
<!ELEMENT PDP (#PCDATA)>
<!ELEMENT DDP (#PCDATA)>
]>
<KPOST>
  <ADDRESS>
    <POSTCODE>305506</POSTCODE>
    <ADMINDIST>대전광역시 유성구 봉산동</ADMINDIST>
    <PDP>181-16 한가람바독학원</PDP>
    <DDP>101호</DDP>
  </ADDRESS>
  ...
</KPOST>

```

(그림 7) 검색 결과의 구조와 인스턴스를 포함하는 XML 문서

- GIS 정보와의 연계 : PAF 기반의 배달점 관리나 순로 정보의 편집 등에 있어서 번지 주소(PDP)나 상세 주소(DDP)의 건물 정보에 GIS 좌표를 연계하여 관리함으로써 사용자 편의성을 높일 수 있다. 또한 캐나다와 미국 등과 같이 관제, 경로 생성, 시설물 관리 등에도 활용할 수 있다[3]. 이를 응용하여 향후 집배원들이 사용하는 PDA를 통한 PAF 및 순로 정보의 편집도 가능하게 된다.
- 배달점 관련 상세 우편정보 관리 : 우편업무 편의성을 고려하여 배달점에 대한 배달 관련 정보를 별도의 테이블로 추가할 필요가 있다.

이 외에도 PAF 중 행정구역 및 배달점에 대한 이력 정보를 관리해야 데이터의 연속성과 일관성을 유지할 수 있고, 배달점과 연관된 수신인 또는 세대원 정보를 저장하여 미국이나 독일에서와 같이 향후 반송우편물 처리에 주소변경 정보를 활용할 수 있을 것이다[7][9]. 상기 제시된 사항들을 관리하고 활용할 수 있도록 우편주소정보관리시스템 프로토타입을 구현하였다.

를 체계적으로 관리할 수 있도록 본 연구에서는 데이터베이스 스키마와 이에 따른 우편주소파일에 대한 설계 방법을 제시하였다. 계층적 행정구역 표현, 번지 및 번지 내 세부 주소 표현, 주소 변형 표기 표현에 대한 3가지 주된 설계 개념을 중심으로 변형된 표기 예를 보인 8가지 주소 특징에 대한 처리 가능성을 보였으며, 이를 위한 행정구역, 주 배달점, 상세 배달점, 키워드, 동의어의 5가지 테이블을 설명하였다. 또한 PAF를 이용한 주소해석 및 검색 방법에 대해 설명하였고, 순로구분 및 배달정보관리 등의 우편업무에 활용할 수 있도록 PAF를 확장할 때의 고려사항들에 대해서도 언급하였다.

본 연구의 결과는 순로구분 자동화를 위한 한글 주소인식 시스템과 향후 배달점 정보를 포함하는 고객바코드의 인증을 위한 주소 데이터베이스 매칭 및 주소 검색 시스템 등에 활용할 수 있을 것으로 기대된다. 5절에서 언급한 고려사항들을 반영하기 위한 추후 연구와 시스템 구현 및 활용이 필요할 것이다.

6. 결론 및 향후 활용 방향

우편주소의 다양한 특징들을 포함하여 주소정보

참 고 문 헌

[1] 김정환, 이석구, 신미영, 남윤석, “효율적인 순로코드 발생을 위한 고속 한글 주소검색 시스템

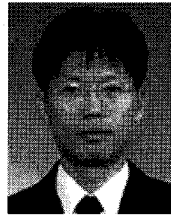
템 개발”, 정보처리학회논문지D, 제8-D권, 제3호, pp.273-284, 2001.

- [2] 김호연, 임길택, 김두식, 남윤석, “서장 우편물 자동처리를 위한 우편영상 인식 시스템”, 정보처리학회논문지B, 제10-B권, 제4호, pp.429-442, 2003.
- [3] 남상우, “우편 업무에 대한 지리정보시스템의 응용”, 우정정보, 23, 1995
- [4] 왕승진, 장태우, 김호연, 남윤석, “한국 우편주소 표기 표준안”, 정보처리학회논문지D, 제11-D권, 제3호, pp.703-708, 2004.
- [5] 우정사업본부, 2005 우편번호부, 정보통신부, 2005.
- [6] 장태우, 정한일, 박찬권, “우편주소 정보화 방안 수립에 관한 연구”, IE Interfaces, Vol. 18, No. 1, pp.104-116, 2005.
- [7] Day, Tom, “Postal Automated Redirection System”, Proceedings of the International Conference on Postal Automation, 2003.
- [8] Royal Mail, PAF Digest Issue 5.0, 2002.
- [9] Stiefelwagen, Christian, “Generation: Automation”, Postal Technology International, 2003.



장 태 우 (Tai-Woo Chang)

- 서울대학교 산업공학과 학사
- 서울대학교 산업공학과 석사
- 서울대학교 산업공학과 박사
- 미국 Pitney Bowes 방문연구원
- 한국전자통신연구원 우정기술연구센터 선임연구원
- 현재 : 경기대학교 산업공학과 전임강사
- 관심분야 : 시스템공학, 정보시스템 통합, 우정물류 등



김 호 연 (Ho-Yon Kim)

- 연세대학교 전산학과 이학사
- 한국과학기술원 전산학과 공학석사
- 한국과학기술원 전산학과 공학박사
- 일본 NHK 放送技術研究所 방문연구원
- 독일 SIEMENS ElectroCom 방문연구원
- 현재: 한국전자통신연구원 우정기술연구센터 선임연구원/팀장
- 관심분야: 패턴인식, 문자인식, HMM, 신경망 등



임 길 택 (Kil-Taek Lim)

- 경북대학교 전자공학과 공학사
- 경북대학교 전자공학과 공학석사
- 경북대학교 전자공학과 공학박사
- 한국전자통신연구원 우정기술연구센터 선임연구원
- 현재: 경주대학교 컴퓨터멀티미디어공학부 조교수
- 관심분야: 패턴인식, 영상처리, 컴퓨터비전, 신경망 등

부 록

PDP_TYPE

배달점	INVD	배달이 불가능한 지역. 토지, 임야, 도로 등 (주소해석에는 사용되지 않음)		
	OLDP	구역이 새로 할당되었거나 분할 등의 이유로 더 이상 유효하지 않음		
	HOUS	이름이 없는 단독주택. 이 경우 PDP_NAME은 세대주가 될 수 있음		
	STOR	일반 상호명 표기에 사용됨 (슈퍼, 미장원, 병원 등)		
	SRVC	서비스		
	APTM	일반적인 아파트		
	APTC	아파트 단지(Apartment complex). 1단지 2단지 등으로 표기될 때		
	APTN	새로운 이름의 아파트로서 아파트로 불릴 수도 있는 주거 단지. 예, 하이츠, 맨션, 빌리지, 마을, 타운 등		
	VILL	빌라		
	HCPX	주택 단지(House complex). 연립, 주택		
	BLDG	~ 빌딩으로 표기되는 건물 (예: 삼성빌딩, LG화재빌딩 등)		
	STRC	일반 건물 (복합건물, 도매상가, 백화점 등)		
	CMPY	일반 회사		
	INCR	주식회사, (주)로 표기되는 회사		
	ORGN	기관		
	YGSO	연구소		
	YGWN	연구원		
POBX	사서함			
교육기관	PRMS	초등학교(Primary school)	MVHS	실업고등학교
	MDLS	중학교(Middle school)	WVHS	여자실업고등학교
	WMDS	여자중학교(women's middle school)	ARHS	예술고등학교
	HGHS	고등학교(high school)	ALHS	체육고등학교
	WMHS	여자고등학교(women's high school)	MRHS	수산업고등학교
	CMHS	상업고등학교(commercial high school)	UNIV	대학교
	CWHS	여자상업고등학교(women's commercial high school)	WUNV	여자대학교
	TEHS	공업고등학교(technical high school)	JCLG	전문대학교
	AGHS	농업고등학교(agricultural high school)	AUNV	예술대학교
	FRHS	외국어고등학교 (foreign high school)	TUNV	공과대학교
	MCHS	종합고등학교	CUNV	체육대학교
	WCHS	여자종합고등학교		

MAIN_NUM_TYPE, SUB_NUM_TYPE, INFO_NUM#_TYPE

MNO	23-4와 같이 dash나 '의'로 표기될 수 있는 위치를 나타내며, 번지의 중간이나 사서함의 중간 등에 쓰임
ENO	23-4와 같이 dash로 분리되는 번지에서 뒷부분을 나타내며, ~번지, ~호를 표기하기 위해 쓰임
EPO	사서함의 끝을 나타내며, 주로 ~호 와 같은 형태로 표기되는 자리에 쓰임
BLK	3블록 1로트와 같이 앞에 나오는 블록(Block)을 가리킨다.
LOT	3블록 1로트와 같이 뒤에 나오는 로트(Lot)를 가리킨다.

DDP_PROPERTY

NORM	배달 가능
INVD	유효하지 않은 배달점
OLD	주소 이전이나 재건축 등의 이유로 유효하지 않음, 배달 불가
FWD	우편물 전달 서비스를 신청한 상태
SVC	기타 특수 서비스를 신청한 상태

DTL_INFO#_TYPE

BDN	행정구역의 동(洞)이 아닌 아파트나 빌딩의 동. 건물 번호(building number)
LVL	건물의 층
YOUL	층 내에서의 열
HO	최종 대상 배달 지점의 실(室) 번호
NULL	미사용

KWD_TYPE

KWD_CATEGORY	KWD_TYPE	의미
ADT	특	특별시, 시
	광	광역시, 시
	도	도
	시	시
	구	구
	군	군
	동	동
	읍	읍
	면	면
	리	리
	로	로
NPT	가	가
	지	지구
	NOR	일반 번지. 아무런 접두어가 없다.
PDT	SAN	산 번지. '산'을 이용
	POT	사서함. '사서함'을 이용
PDT	HOUS, APTM 등	PDP_TYPE으로 사용되는 값 및 의미와 동일
DNT	MNO	"2-4번지" 의 경우 -
	ENO	"23-4번지" 에서 ~번지, ~호
	EPO	~사서함. "사서함 4-5호"에서 중간의 -는 MNO를 이용하여 표기함
	BLK	~블록
	LOT	~로트
	BDN	동
	LVL	층
	YOUL	열
	HO	호
	NULL	미사용

KWD_PROPERTY

STD	해당 키워드의 표준형
USAL	통상적으로 자주 사용되는 경우
UNUS	간혹 사용되는 경우
SLDM	잘 사용되지 않는 경우 (사용되는 것으로는 인정함)

SYNONYM_PROPERTY

USAL	통상적으로 자주 사용되는 경우
UNUS	드물게 사용되는 경우
SLDM	잘 사용되지 않는 경우 (동의어로 인정은 함)
OLDS	변경되었거나 이전에 사용하던 경우