

배전설비도면 자동인식입력에 관한 연구

송재주, 김계영, 이봉재, 한철성, 조선구
한국전력공사 전력연구원 정보통신그룹
e-mail:leeboja, kimgy, cshan, josg@kepri.re.kr

A Study on Automatic Recognition of Power Distribution Facility Map

Jae-Ju Song, Gye-Young Kim, Bonng-Jae Yi, Chil-Sung Han,
Seon-Ku Cho
Computer&Communication Group, Korea Electric Power
Research Institute

요약

본 논문에서는 한국전력공사의 배전지리정보시스템을 보다 효과적으로 구축하기 위하여 배전설비도면으로부터 설비정보를 자동취득하여 배전지리정보시스템의 데이터베이스에 입력하는 방법에 관하여 기술한다. 제안된 방법은 충남지사와 인천지사 관내 배전설비도면 약 30매를 대상으로 실험하고 그 결과를 제시한다.

1. 서론

지리정보시스템은 전력, 통신, 가스 등의 국가기반자원의 관리 및 국토 이용의 효율성을 증대하는데 유용한 도구이며, 그 특성상 범국가적인 차원에서 이루어지는 것이 가장 효과적이다. 구미의 선진국에서 이미 1960년대부터 국가지리정보시스템을 구축하기 시작하였지만, 우리나라는 1995년도 NGIS 추진위원회가 설립됨으로써 국가차원의 지리정보시스템 구축 사업이 시작되었다[1]. 국가지리정보시스템 구축을 위하여 정부는 요소기술개발 지원, 표준 제정 및 기본도 입력을 담당하고, 한국전력공사 또는 한국통신공사와 같은 사회기반설비 운용사는 각 운용사가 가지는 설비정보를 데이터베이스화한 후, 국가지리정보시스템과 연계하여 운용되도록 하여야 한다. 특히, 한국전력공사에서는 배전 GIS를 구축하여 국가 GIS와 연계하여 운용하여야 하는데, 이를 위하여 한국전력공사가 가지고 있는 방대한 배전설비정보를 데이터베이스에 입력하는 작업은 필연적이다.

배전 GIS 구축에 필요한 초기자료를 취득하는 방법으로는 다음과 같은 세 가지 방법이 고려될 수 있다. 첫 번째는 스캔된 도면에 있는 객체의 위치

및 관련 정보를 수작업으로 입력하는 방법이며, 두 번째는 벡터화(vectorization) 하는 방법이고, 마지막으로 세 번째는 자동인식에 의한 방법이다. 위의 세 가지 방법 중에서 세 번째 방법인 자동인식에 의한 방법이 가장 효과적인 것으로 고려되고 있다 [2].

본 논문에서는 배전설비도면으로부터 설비정보를 자동취득하여 배전 GIS 데이터베이스에 입력하는 방법에 관하여 기술한다.

본 논문의 구성은 다음과 같다. 제 1장에서는 본 논문의 기본방향에 대하여 기술하였다. 제 2장에서 배전설비도면의 특성 및 인식대상에 관하여, 제 3장에서는 배전설비도면 자동인식 방법에 대하여 각각 설명한다. 제 4장에서는 실험결과 및 성능평가에 관하여 기술한 후, 제 5장에서 결론 및 향후연구에 관하여 논술한다.

2. 배전설비도면의 특징 및 인식대상

본 논문의 목표는 배전설비도면으로부터 설비정보 즉, 설비의 종류, 위치, 연결관계 및 속성정보를 자동취득하여 배전 GIS 데이터베이스에 입력하는 방법을 개발하는 것이다. 이를 위하여 먼저 배전설비

도면의 특성 및 인식대상을 분석하면 다음과 같다. 본 논문에서 인식하고자 배전설비도면의 특성은 표 1와 같다.

표 1. 배전설비도면의 특성

분류	특성
원도재질	polyester film
원도크기	66cm×46cm
원도축척	1:1200(실거리 500m×500m) 1:3000(실거리 1km×1km)
주요구성	도면영역:배전설비도, 지적도, 구획선 도면정보영역:관리구번호, 축척, 지역명
스캐너	명암등급 : 256 gray scale 해상도 : 300DPI
디지털도면	형식 : PCX 크기 : 수 MB ~ 수십 MB 화소수 : 7568×5433(pixels)

표 2. 배전설비도면기호의 분류 및 표현

번호	전 주		전 선	
	기호	이름	기호	이름
1		단주		고압전선
2		A, H주		저압전선
3		삼각주		고압케이블
4		삼각주		저압케이블
5		사각주		인입선류
6		변대주		절분선류
소계	6개		7개	
번호	보조설비		기기류	
	기호	이름	기호	이름
1		수평지선		개폐기 (IS)
2		보호망선		개폐기 (GS)
3		보통지선		리클로저
4		가공지선		색셔날 라이저
5				COS
6				피뢰기
소계	4개		6개	
	합계		23개	

배전설비도면에서 배전설비정보를 표현하기 위해

서는 표 2과 같은 기호와 기호가 가지는 속성을 기술하는 영자와 숫자 등의 문자들로 구성되므로, 본 논문의 자동인식 대상은 표 2의 기호와 영자 및 숫자이다.

3. 배전설비도면 자동인식의 처리흐름

배전설비도면을 자동인식하여 배전 GIS 데이터베이스에 입력하기 위한 처리과정은 그림 1과 같이 사전정보추출, 문자학습, 전처리, 자동인식, 인식결과 편집 및 처리, 배전 GIS DB와 연계 등 6가지 단계로 구성된다.

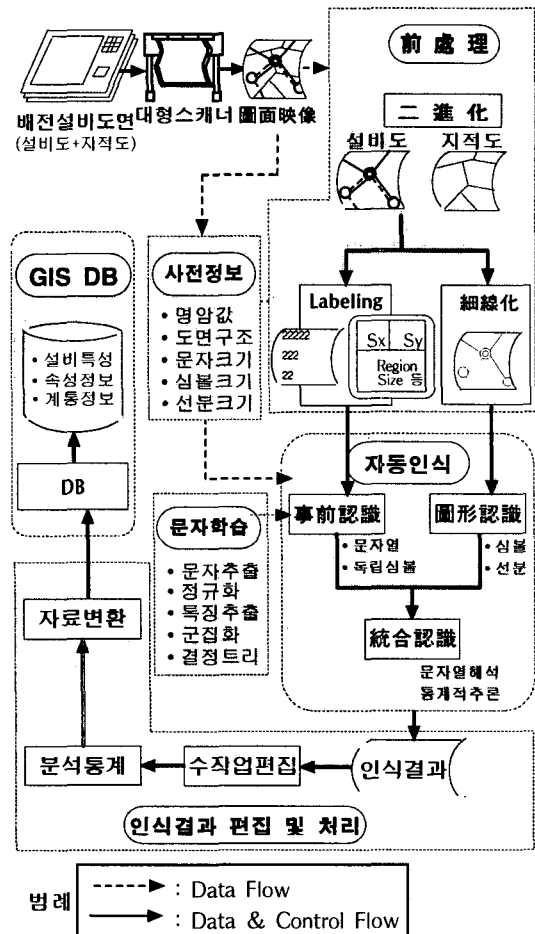


그림 1. 배전설비도면 자동인식의 처리흐름

사전정보추출 단계에서는 배전설비도면을 자동인식하는 과정에서 필요한 정보 즉, 동일한 형태를 가지는 배전설비도면들의 공통적인 정보인 도면구조,

문자크기, 기호크기, 선분간격 등의 정보를 추출한다.

문자학습 단계는 인식대상문자를 컴퓨터에 가르쳐 주는 과정으로 입력영상에서 학습용 문자를 추출하는 문자추출 및 정규화(Normalization), 추출된 문자가 가지는 특성을 산출하는 문자특징추출, 추출된 문자특징 중에서 문자군들 간의 구분을 잘 할 수 있도록 특징요소를 선별하는 FD(Fisher's Discriminant)분석, FD 분석에 의하여 선택된 특징들로 형성되는 특징공간에서 유사한 특성을 가지는 문자군으로 분류하는 군집화, 군집화된 문자들을 트리 형태로 생성하는 결정트리생성 등의 기능으로 구성된다. 사전정보추출 단계와 문자학습 단계는 매도면별로 수행되는 것은 아니고 도면의 기재요소가 미리 등록되지 않은 별도의 특성을 가질 때에 한하여 작업자 판단에 따라 수행한다.

전처리 단계는 입력영상으로부터 인식과정에서 필요한 영상을 생성하는 과정으로 다음과 같이 구성된다. 제1원도인 지적도와 제2원도인 설비도를 분리하여 두 개의 영상을 생성하는 이진화(Binarization), 제2원도에서 연결된 구성요소(Connected Component)별로 고유번호를 부여한 영상을 생성 및 영역정보를 추출하는 레이블링(Labeling), 그리고 선분을 추적할 때 발생하는 모호성을 제거하기 위해 설비도영상에 있는 선분들의 폭을 하나의 화소로 된 영상, 즉 세선영상을 생성하는 세선화(Thinning) 등으로 구성된다.

자동인식 단계는 도면기재요소 즉, 고압선 및 저압선과 같은 선분, 전주와 변대주 등을 나타내는 기호, 그리고 설비속성을 표현하는 문자를 구분하여 인식한 후, 이들간의 연관성을 통해 인식결과를 완성하는 과정으로 사전인식, 도형인식, 그리고 통합인식으로 구성된다. 사전인식에서는 개별문자 및 독립기호를 인식한다. 도형인식은 사전인식에서 제외된 영역을 인식하는 것으로 다음과 같은 절차에 의하여 이루어진다. 먼저, 세선영상을 참조하여 인식처리의 기본 단위원 세그먼트 정보를 추출한 후, 추출된 세그먼트의 해석 가능한 가설을 생성한 다음, 생성된 가설을 검증함으로써 배전설비도면에 있는 기호 및 선분을 인식한다. 통합인식에서는 문자인식을 통하여 인식된 개별문자들을 조합하여 도면기재규칙상의 의미를 가지는 문자열을 추출하며, 가설검증 결과가 모호한 경우에 가장 근접한 결과를 얻기 위한 통계적 추론(Probabilistic Reasoning)을 수

행한다. 상기와 같은 과정을 통하여 인식된 도면정보는 각 설비별 설치구조 및 연결구조에 따라 개별설비별로 위치정보와 함께 저장되며, 인식이 불완전한 정보는 다음 단계인 인식결과 편집과정에서 수동으로 편집 및 수정한다.

인식결과 편집 및 처리 단계에서는 자동인식 단계에서 오인식되거나 미인식된 인식대상을 수작업으로 편집 및 수정하며, 그 결과자료는 배전 GIS 데이터베이스의 모델에 적합한 자료구조로 생성한 후 저장한다.

4. 구현 및 실험결과

이 장에서는 실험결과에 관하여 기술한다. 실험에 사용한 컴퓨터는 IBM-PC/586으로 CPU 속도는 450Mhz이며, 프로그래밍 언어는 Visual C++이다. 영상입력은 Vidar TruScan 500 스캐너를 사용하여 입력하였다. 입력영상의 화소밀도는 300DPI이며 영상크기는 7568× 5433 화소이고 한 화소가 나타낼 수 있는 밝기의 종류는 256이다.

표 3. 주요설비별 인식을 통계

도면 질	처리 시간	설비 종류	설비 수	인식 수	정상 인식	오 인식	미 인식	인식률 (%)
중요	10분 55초	전 주	97	94	93	1	4	96
		고압선	52	62	52	10	-	100
		저압선	28	24	24	-	4	86
		문 자	1228	1074	846	228	69	79
		문자열	368	323	174	149	194	47
보통	16분 10초	전 주	204	205	191	14	13	94
		고압선	94	95	85	10	9	90
		저압선	32	32	29	3	3	91
		문 자	2644	2312	1942	370	703	73
		문자열	712	654	340	314	372	48
나쁨	15분 44초	전 주	177	173	165	8	12	93
		고압선	106	111	98	13	8	92
		저압선	37	32	29	5	8	78
		문 자	2046	1746	1427	319	619	70
		문자열	612	543	314	229	298	51
합계	42분 09초	전 주	478	472	449	23	29	94
		고압선	252	268	235	33	21	93
		저압선	97	88	82	8	15	85
		문 자	5918	5132	4215	917	1391	71
		문자열	1692	1520	828	471	864	50

본 논문에서 제안된 배전설비도면 자동인식입력 방법은 충남지사와 인천지사 관내 배전설비도면 약 30매를 대상으로 성능평가를 수행하였다. 성능평가는 다음과 같이 두 가지 관점에 주안점을 두었다.

첫 번째 관점은 실험대상도면의 질에 따른 성능변화를 관찰하는 것이다. 이를 위하여 실험대상도면을 그 질에 따라 좋음, 보통, 나쁨으로 분류하여 각 그룹의 평균인식율을 산출하였다. 두 번째는 배전설비도면의 구성하는 설비별 성능변화를 관찰하는 것이다. 이를 위하여 전주/변대주, 고압선, 저압선, 수평지선 그리고 문자 및 문자열 각각에 대한 각 그룹의 평균인식율을 산출하였다. 표 2의 인식대상 중에서 상기의 설비들만으로 성능을 평가한 것은 이들을 제외한 다른 설비들은 그 빈도수가 매우 적어 인식율에 대한 통계를 취득하는 것이 어려울 뿐아니라 성능평가의 의미가 없기 때문이다. 표 3에서는 성능평가를 수행한 결과를 보인 것이다.

본 논문에서 구현한 배전설비도면 자동인식 소프트웨어는 그림 3과 같이 구성된다. 즉, 주메뉴, 도구바, 도면영역제어창, 처리결과표시창, 도면표시창, 및 처리상태표시창으로 구성된다.

정보를 자동취득하는 방법에 대하여 기술하고 실험을 통한 성능평가결과를 제시했다.

성능평가결과 주요기호인 전주/변대주, 고압선, 저압선, 수평지선 등의 인식율은 현장에 적용할 만큼 높은 반면 문자 및 문자열 인식율은 저조했다. 이는 도면에서 문자의 기재 방향이 불규칙하기 때문이다. 따라서, 향후에는 문자 및 문자열 인식율을 향상시키기 위하여 기재방향에 독립적인 문자 및 문자열 인식에 관하여 더 많은 연구를 수행하여야 할 것으로 사료된다.

[참 고 문 헌]

- (1) NGIS 총괄분과위원회, 국가지리정보체계(NGIS) 구축 기본계획, 1997. 8.
- (2) J. E. Den Hartog, T. K. Ten Kate, and J. J. Gerbrands, "Knowledge-based interpretation of utility maps," *CVIP*, Vol 63, No. 1, pp. 105-117, 1996.

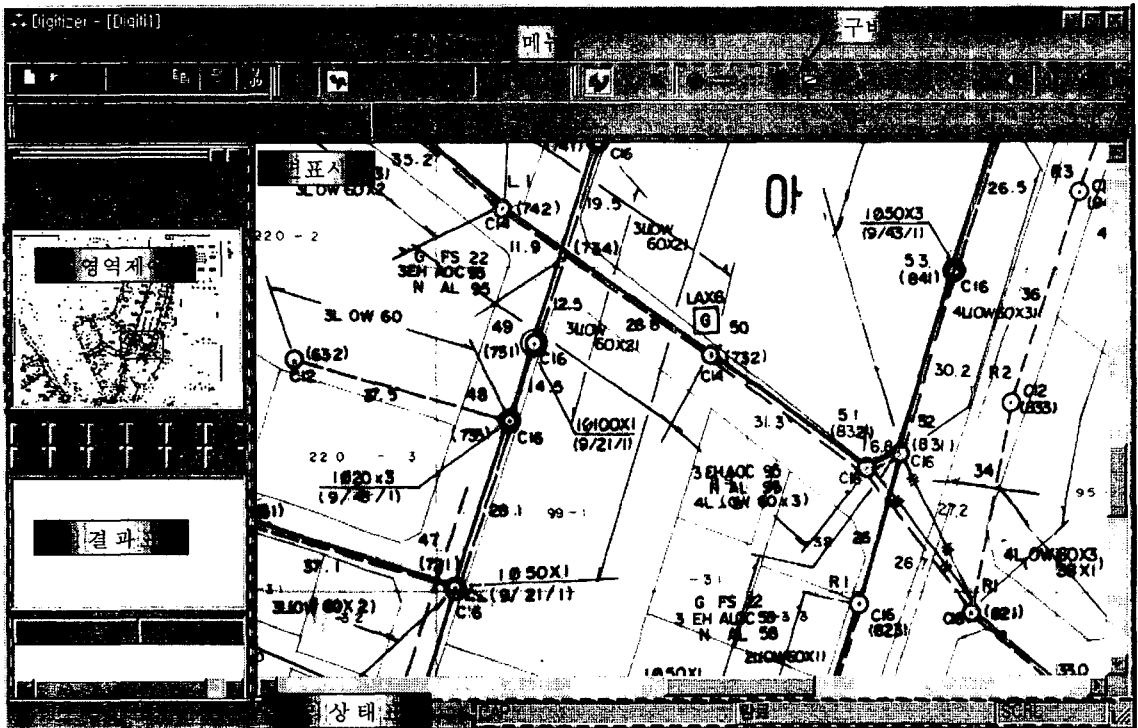


그림 3. GIS용 설비도면 자동인식 소프트웨어

4. 결론 및 향후연구

본 논문에서는 한국전력공사의 배전 GIS를 효과적으로 구축하기 위하여 배전설비도면으로부터 설비