

LIS/GIS의 D/B구축을 위한 수치지도의 위치오차분석에 관한 연구 Position Error Analysis of Digital Map for LIS/GIS Database

조규전* · 이영진** · 홍용현***

Cho, Kyu-Jon · Lee, Young-Jin · Hong, Yong-Hyun

요 旨

LIS/GIS의 D/B구축을 위한 수치지도에서의 위치정보는 지물의 위치판독뿐만 아니라 이를 이용한 LIS/GIS의 구조화나 시설물 관리, 도시관리, 지적관리, 공간분석 등에서 중요한 자료로 이용되고 있다. 본 연구에서는 수치지도에서 평면위치자료의 오차분석을 위하여 지형별 특성에 따라 지형, 밀도, 경사도의 세가지 형태로 분류하고, 지형도의 평면위치를 참값으로 하여 수치지도와의 상대적 오차를 분석하였다. 수치실험결과, X, Y 방향에 관계없이 면구조의 데이터보다 선구조의 데이터에 대하여 과대오차가 더 많이 발생함을 알 수 있었으며, 지형별 분석의 경우에 있어서는 교외지와 산악지에서의 평균오차가 작게 나타났다. 밀도별 분석에 있어서는 느슨한 밀도를 기준으로 중간밀도와 조밀한 밀도에서의 평균오차가 작게 나타나고, 경사도별 분석에 있어서는 평지를 기준으로 급경사지와 완경사지에서의 평균오차가 작게 나타남을 알 수 있었다.

ABSTRACT

In the digital mapping for data base construction, the positional information is used as an important tool in LIS/GIS(Land Information System/Geographic Information System) that is used for a facility management, urban/cadastral management as well as in spatial analysis. In this paper, for an error analysis of X, Y coordinates data on digital map, test area was classified by topography, density, and slope. The coordinates on topographic map were assumed as true values and they were compared with the coordinates on digital map. A result of the numerical test show that a vector data of line type had more gross error than vector data of polygon type. And, SME(standard mean error) of urban or intermountain area had small values compared to that of suburban area in topography analysis. The SME of dense and middle zone had small values compared to that of loose zone in density analysis. In another slope analysis, the SME of steep of gentle slope had small values compared to that of flatland.

1. 서 론

최근 여러 분야에서 사용되고 있는 LIS/GIS(Land Information System/Geographic Information System)는 토지, 지적, 지도, 지형 등과 관련된 다양한 정보를 특성에 따라 공간적 위치에 적합되도록 조직적으로 구축하고 저장, 분석, 처리하여 모든 분야에 적용할 수 있는 종합정보체계를 말한다.^{9,11,12)}

LIS/GIS의 처리과정은 자료입력, 관리, 분석, 결과출

력이라는 4단계로 이루어지고 있으며 정확한 분석과 결과를 위해서는 입력되는 자료의 정확도가 매우 중요한 의미를 갖는다. 특히 입력자료중 지형·지물에 관계된 3차원 자료를 얻기 위해서는 많은 시간과 경비가 소요되며 정확성이 유지되어야 한다. 따라서 보다 경제적이면서도 정확한 수치 데이터를 얻을 수 있는 방법으로 기존 지형도나 주제도로부터 필요한 자료를 추출해 내는 방법을 사용하여야 한다.

최근 해석도화기에 의해 컴퓨터를 이용한 수치지도 제작 방법이 소개되면서 지도 제작의 새로운 국면을 맞이하고 있다. 컴퓨터 그래픽에 의한 CAD/CAM/CAE의 이용이 증가하고 그 처리 기술이 다양해지고 있으며,

* 경기대학교 공과대학 토목공학과 교수

** 경일대학교 공과대학 측지공학과 조교수

*** 경기대학교 대학원 토목공학과

지도를 이용한 많은 분야에서도 컴퓨터를 활용한 지도 제작활동을 모색하고 있다.^{3,9)} 기존의 측량의 좌표계산 등에서 주로 이용하던 컴퓨터가 각종 지도의 내용을 수치화하여 처리하고 가공하므로써 다양한 분석과 출력을 행하는 일련의 수치지도제작 기술을 도입하기에 이르렀다. 종래의 지도 제작 및 이용 방법은 주로 인력에 의한 지도 정보 획득, 제도, 인쇄의 과정을 거침으로써 지도 작성 시간이 길고, 내용 수정이 있을 때 마다 번거로운 작업을 해야만 했다. 그러나 수치지도 제작 방법에 의한 지도제작은 지도 자료의 획득, 가공, 처리가 컴퓨터에 의해 체계적으로 이루어져서 그 정확성이 보장되며 지도성곽의 작성, 관리, 이용등이 일관성있게 처리될 수 있게 되었다.

수치지도의 제작방법은 기존의 지도를 디지털화하거나 스캐닝한 자료를 벡터화하여 편집하는 방법을 사용한다.^{2,3)} 스캐닝 방법에 의한 수치지도는 스캐닝 방법이나 지형·지물의 종류에 따라 그 정확도가 달라지므로 이에 대한 연구와 평가가 매우 중요하다. 수치지도 제작과정에서 사용되는 정보데이터 중에서 평면 위치 자료, 표고자료들은 매우 중요한 것들이다. 특히 평면 위치 자료나 표고 수치 자료는 GIS를 응용한 지하 시설물 관리나 도로, 철도 등의 노선설정, 계획, 그 외 지물의 위치 관측 등에서 중요한 자료로 이용되고 있으나, 이들에 대한 정확도 검증과 오차의 원인 및 그 해결방안에 대한 연구가 매우 미미한 실정이다. 따라서 본 연구는 수치지도의 지형 특성과 지물 특성에 따른 X, Y 평면 위치 오차의 크기와 분포 상황을 고찰하여 그 원인을 규명해보고 그 결과를 앞으로 제작되는 수치지도의 정확도 향상과 LIS/GIS의 데이터베이스 구축과 관리 및 그 응용성과에 기여토록 하는 데 목적이 있다.

2. 오차분석 이론

일반적으로 관측결과에서 발생하는 오차는 과대오차(gross error)와 정오차(systematic error) 그리고 우연오차(random error)의 세 경우로 구분한다.

정오차는 함수식에 의해 표현되어 질 수 있는 일정한 법칙에 따라 발생되므로 그 원인과 상태만 알면 쉽게 제거되어진다. 우연오차는 측정시의 조건이 수시로 변화하기 때문에 그 원인을 찾기 힘들거나 모르게 나타난다. 이러한 우연오차는 발생빈도, 크기 등을 전혀 알

수 없는 무작위성 오차로 정규분포와 같은 확률론적인 방법으로 다루어진다. 과대오차는 관측자의 부주의나 측정 방법의 잘못된 선택으로 발생되어지는 오차로 앞의 두 형태에 비하여 그 크기가 매우 크므로 쉽게 발견되며 오차로 취급하기가 곤란하지만 실제적으로 측정값에 큰 영향을 준다.^{3,4)}

본 논문에서는 정오차의 성분은 지도 제작당시 완전히 소거되어진 것으로 가정하고 우연오차의 성분과 자료 추출시에 발생되어 질 수 있는 과대오차 성분만을 고려하였다.

2.1 오차처리

통계적 측면에서 측정값은 확률분포의 한 표본이며 확률법칙을 따른다고 본다. 그러므로 과대오차를 포함하고 있는 측정값은 표본측정이 만들어진 모분포와 동일한 분포에 속한다고 볼 수 없다. 즉, 과대오차를 제거해야 그 측정값들을 사용할 수 있다는 것이다.

따라서 본 논문에서는 과대오차의 추출과 제거방법으로서 두 가지를 사용하였다.

2.1.1 정규분포에 의한 과대오차 검출^{5,6)}

식 (1)은 정규분포 $N(\mu, \sigma^2)$ 을 따르는 관측값 x 가 평균에서 σ 의 상수배의 범위에 존재하는 확률을 의미한다. 즉, 어떤 관측값 x 가 평균에서 3σ 또는 4σ 의 범위에 존재할 확률이 각각 99.73%와 100%를 보여주고 있으며, 보통 이 범위를 우연오차의 허용한계로 보고 이 한계보다 큰 범위의 오차를 과대오차로 보고 소거한다. 따라서, 특별한 경우에 있어서 $\pm 0.6745\sigma_x$ 는 67.45%의 신뢰구간을 의미하며, $\pm 1.96\sigma_x$ 는 95%, $\pm 2.57\sigma_x$ 는 99%의 신뢰구간, 그리고 $\pm 3.29\sigma_x$ 는 99.9%의 신뢰구간을 나타낸다. 일반적으로 측정값의 편차 $(x_i - \bar{x}) \pm 3.29\sigma_x$ 보다 큰 경우에 이를 과대오차로 보고 소거한다.

$$\begin{aligned} P[\mu - 0.6745 \sigma_x \leq x \leq \mu + 0.6745 \sigma_x] &= 0.500 \\ P[\mu - 1.9600 \sigma_x \leq x \leq \mu + 1.9600 \sigma_x] &= 0.950 \\ P[\mu - 2.5760 \sigma_x \leq x \leq \mu + 2.5760 \sigma_x] &= 0.990 \\ P[\mu - 3.2900 \sigma_x \leq x \leq \mu + 3.2900 \sigma_x] &= 0.999 \end{aligned} \quad (1)$$

2.2.2 Data-Snooping 방법^{10,12,13)}

이 방법은 실제오차에 대한 임의의 분포과정으로부터 유도된다. 관측값의 실제오차 ϵ 가 미지수이기 때문에 잔차 v 와 미지수 X 가 오차의 기대값으로 사용되며 어떠한 귀무가설 아래에서 잔차에 대하여 통계적으로

조정된다.

검정통계량 T는 다음과 같이 나타낸다.

$$T = \frac{V^T P V}{(n-u) \sigma_0^2} = \frac{\sigma^2}{\sigma_0^2} \leq F(1-\alpha, r, INF) \quad (2)$$

여기서, V는 잔차행렬

P는 경중을 행렬

n-u는 잉여관측수

σ_0^2 는 사전분산

σ^2 는 사후분산

α 는 유의 수준을 나타낸다.

Baarda는 일반 잔차대신에 시스템의 기하조건이 포함된 표준화잔차(standardized residual)를 사용하여 다음 식 (3)과 같은 검정값 W_i 를 조사하였다.

$$W_i = \frac{V_i}{\sigma_{v_i}} = \frac{V_i}{\sigma_0 \sqrt{(Q_{VV})_{ii}}} \quad (3)$$

여기서, $(Q_{VV})_{ii}$ 는 잔차의 경중을 행렬의 i번째 대각선 요소이며

$$Q_{VV} = I - A(A^T P A)^{-1} A^T P \quad (4)$$

로 표시된다. 여기서 A는 관측방정식의 행렬이며 I는 단위행렬이다.

검정통계량 W_i 는 기각값 K와 비교하여 $W_i > K$ 이면 관측값 i에 과대오차가 포함되었다는 대립가설을 채택하게 된다. 기각값 K를 Baarda는 4.123, Förstner는 3.29로 제안하였다.

과대오차 V_i 가 W_i 에 미치는 실제영향은

$$\delta_i = \nabla W_i = \frac{\nabla V_i}{\sigma_{v_i}} = \frac{\nabla V_i}{\sigma_0 \sqrt{(Q_{VV})_{ii}}} = \frac{V_i}{\sigma_0} \sqrt{(Q_{VV})_{ii}} \quad (5)$$

가 되며 δ_i 를 비중심모수(non-centrality parameter)라 한다.

귀무가설과 대립가설의 차인 δ_i 의 하한계 δ_0 는

$$\delta_0(\alpha_0, \beta_0) \approx \phi^{-1}\left(1 - \frac{\alpha_0}{2}\right) + \phi^{-1}(\beta_0) \quad (6)$$

로 나타내며 $\alpha_0=0.1\%$, $\beta_0=80\%$ 일 때, $\delta_0=4.13$, $\alpha_0=1\%$, $\beta_0=80\%$ 일 때, $\delta_0=3.42$ 이다. 따라서 식 (5)의 δ_i 를 하한계 δ_0 로 대치하면 관측값에서의 과대오차 하한계 V_i 를 구할 수 있다.

$$\nabla_0 V_i = \sigma_0 \frac{\delta_0}{\sqrt{(Q_{VV})_{ii}}} \quad (7)$$

3. 수치실험

3.1 대상지역의 선정과 점의 선정

본 연구는 선정된 대상지역에 대해서 지형의 특성을 고려하기 위하여 시가지, 교외지, 산악지등 세가지 지형 형태로 분류하였으며 지물 특성을 고려하기 위하여 시가지는 밀집도에 따라 세등급으로 재분류를 하였다.

비교 데이터로 사용되는 X, Y 평면 좌표값을 얻기 위한 대상점의 선정은 건물과 등고선을 중심으로, 건물의 좌표는 모서리점을, 등고선은 다른 선형 데이터와 교차하는 점들을 기준으로 1,056점을 무작위로 선택하였다. 지형도에서의 대상점에 대한 좌표 획득은 디지털 이터를 이용하여 X, Y 평면직각좌표를 취득하고, 동일 지점에 대한 수치지도에서의 좌표값은 AutoCAD를 이용하여 평면직각좌표를 취득하여 두 좌표값의 차를 비교분석하였다. 비교값 산출을 위한 수치지도는 건설교통부 국립지리원에서 스캐닝 방법에 의하여 제작되고, 벡터편집까지 완료된 수치지도를 사용하였다. 표 1, 표

표 1. 도엽별 선정 내용

도엽명	선점기준			합 계
	건 물	등 고 선		
016 도엽	267 점	63 점		388 점
025 도엽	154 점	282 점		436 점
026 도엽	151 점	139 점		290 점
합 계	572 점	484 점		1,056 점

표 2. 분석 방법별 선정 내용

지형별 분석			밀도별 분석			경사별 분석		
시가지	교외지	산악지	조밀	중간	느슨	급경사	완경사	평야지
348 점	136 점	45 점	245점	145 점	151 점	123 점	123점	20 점
총 529 점			총 541 점			총 266 점		

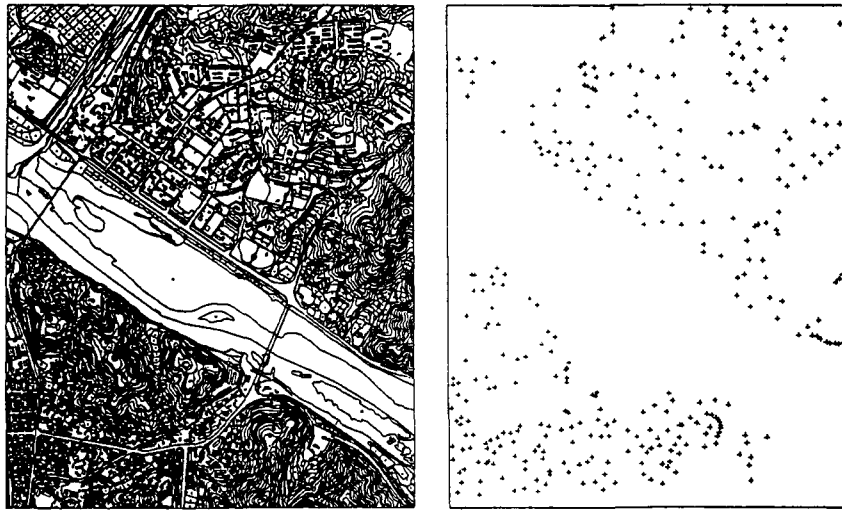


그림 1. 공주 016도엽과 점의 위치

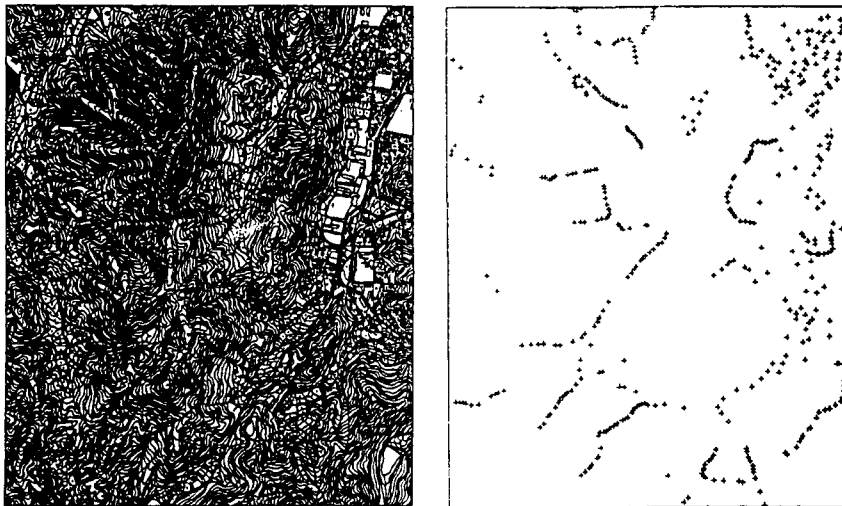


그림 2. 공주 025도엽과 점의 위치

2와 그림 1, 그림 2, 그림 3은 선점된 내용과 대상지역에 대한 지형도를 보여준다.

3.2 과대오차 검출

본 연구에서는 과대오차의 제거방법으로 정규분포에 의한 방법과 Baarda의 방법을 사용하고자 하였으나, 망 조정의 경우와는 달리 알고있는 점들이 많았기 때문에 정규분포에 따른 과대오차를 추출·제거하는 방법을 취하였다. 대상 지역에서 99.9%가 신뢰도인 3.29σ 보다 벗어나는 오차를 과대오차로 보고 통계분석에서 제외시켰다. 전체 1056 점 중에 수치지도와 지형도와의 비

교과정에서 누락되거나 비교불가능한 149개 점을 제외하고 선택된 표본 907개 점에 대하여 분석결과, 이 중 X 방향오차에 대한 과대오차는 전체 비율 0.55%에 상당하는 7개 점에서 발견되었으며, Y 방향오차에 대한 과대오차는 전체 비율 1.54%에 상당하는 14개 점에서 발견되었다. 그 현황은 표 3과 같다.

표 3에서 알 수 있듯이 X, Y 방향에 관계없이 면구조의 데이터보다 선구조의 데이터에서 더 많은 과대오차가 나타남을 알 수 있다.

또한 전체 X, Y 방향오차에 대한 분석과 동일한 방법으로 지형별 분석, 밀도별 분석, 경사별 분석의 방법에

대한 과대오차를 소거하였으며, 그 결과는 표 4와 같다.
과대오차가 발생된 점들을 지형도와 수치지도 상에

서 비교하여 살펴보면, 등고선의 경우에 있어서는 도로
나 간류, 지류, 하천선과 같은 동일한 선데이터와 접하

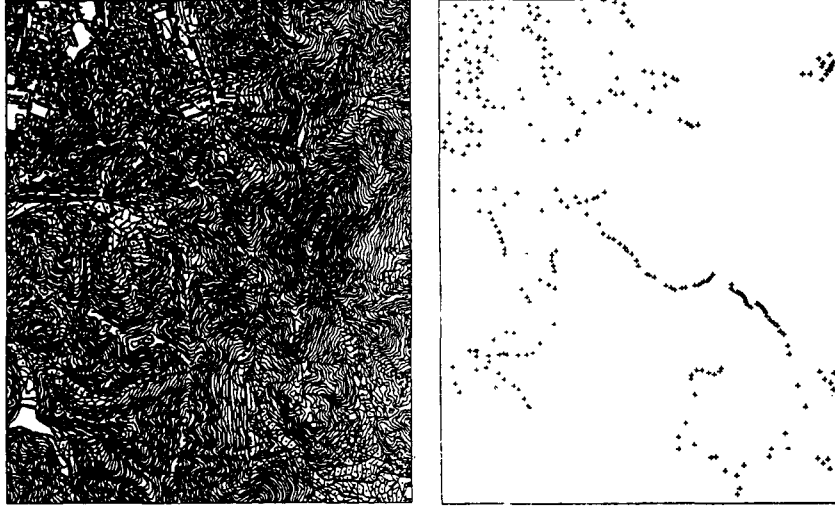


그림 3. 공주 026도엽과 점의 위치

표 3 (a). 전체 X 방향오차에 대한 과대오차 검출

3.29σ	도엽	016 도엽		025 도엽		026 도엽		합 계	비 고
		건 물	등고선	건 물	등고선	건 물	등고선		
3.5770	.	.	3	.	2	.	.	5 점	1차 검색
3.3618	.	.	1	.	.	.	1	2 점	2차 검색
3.3228	3차 검색
합 계			4		2		1	7 점	

표 3 (b). 전체 Y 방향오차에 대한 과대오차 검출

3.29σ	도엽	016 도엽		025 도엽		026 도엽		합 계	비 고	
		건 물	등고선	건 물	등고선	건 물	등고선			
3.2766	.	.	1	.	4	.	2	7 점	1차 검색	
2.6316	.	.	.	2	1	1	.	4 점	2차 검색	
2.5677	.	.	.	1	1	.	.	2 점	3차 검색	
2.5385	1	.	1 점	4차 검색	
2.5344	5차 검색	
합 계			1		3		6	2	2	14 점

표 4. 특성별 과대오차 현황

현 황	특 성	지형별 분석			밀도별 분석			경사도별 분석		평 지
		시가지	교외지	산악지	조밀	중간	느슨	급경사	완경사	
측점수		348 점	136 점	45 점	245 점	145 점	151 점	123 점	123 점	20 점
과대 오차	X	2 점	.	.	1 점	1 점
	Y	3 점	.	.	2 점	1 점	.	7 점	.	.

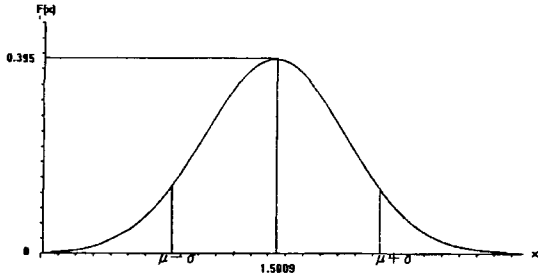


그림 4. X 방향오차에 대한 정규분포도

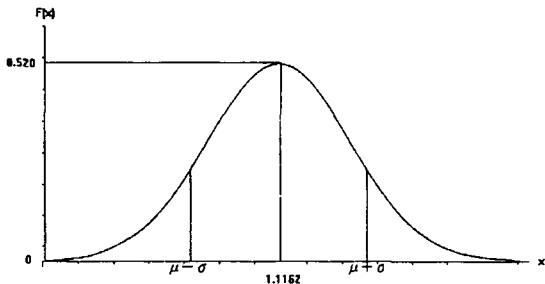


그림 5. Y 방향오차에 대한 정규분포도

거나 급격하게 변하는 부분에서 발견되었으며, 건물의 경우에 있어서는 도로 경계선, 등고선, 지류등과 같은 선데이터와 일치하거나 접하는 부분 또는 선의 두께가 변하는 부분에서 과대오차가 발생하였다. 이러한 과대오차가 발생한 원인으로서는 스캐닝 과정에서 읽어들이진 래스터 데이터를 벡터 데이터로 변환하는 과정에서 완벽하게 벡터데이터로 변환시키지 못하여 발생하는 기계오차, 또는 수동 편집 과정에서 작업자의 손의 흔들림에 의하여 발생하는 과도한 오차로 판단된다.

아래 그림 4와 그림 5는 각각 X 좌표오차와 Y좌표 오차에 대해서 과대오차를 제거한 다음의 표본의 분포 상태를 나타낸 것이다.

4. 분석 및 고찰

본 논문에서는 지형특성에 따른 X, Y 평면위치의 오

차분석을 위하여 지형특성에 따라 시가지, 교외지, 산악지로 구분 하였으며, 시가지의 경우 밀도에 따라 다시 세 등급으로 분류하여 분석하였고, 경사도에 따라 세가지 형태를 분석하였다. 각각의 특성에 따른 분석 결과는 표 5와 같다.

지형별 분석의 경우에 있어서 X 방향과 Y 방향에 대한 평면오차의 크기가 교외지, 산악지, 시가지의 순서로 크게 나타나고 밀도별 분석에 있어서는 느슨한 밀도, 중간 밀도, 조밀한 밀도의 순서로 크게 나타나고 경

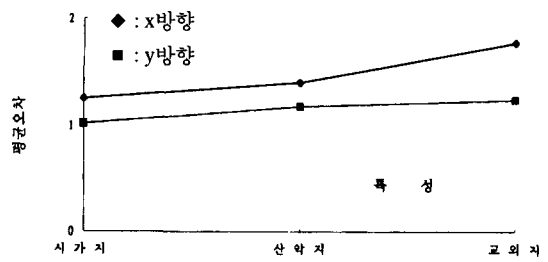


그림 6. X, Y 방향에 대한 지형별 분석의 평균오차

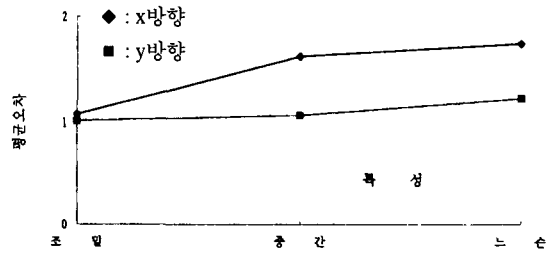


그림 7. X, Y 방향에 대한 밀도별 분석의 평균오차

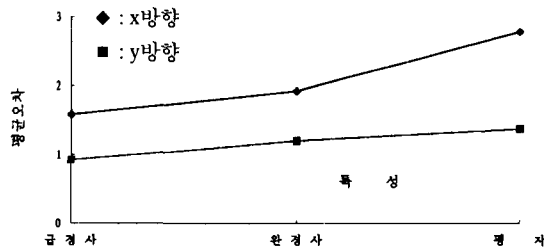


그림 8. X, Y 방향에 대한 경사도별 분석의 평균오차

표 5. 특성별 평균오차 현황

(단위 : m)

특성	지형별 분석			밀도별 분석			경사도별 분석		
	시가지	교외지	산악지	조밀	중간	느슨	급경사	완경사	평지
X 방향 평균오차	1.2527	1.7767	1.3986	1.0649	1.6241	1.7446	1.5651	1.9139	2.7782
Y 방향 평균오차	1.0116	1.2427	1.1765	1.0055	1.0552	1.2250	0.9306	1.2000	1.3641
평균	1.13215	1.5097	1.28755	1.0352	1.33965	1.4848	1.74785	1.55695	2.07115

사도별 분석에 있어서는 평지, 완경사, 급경사의 순서로 평면오차가 크게 나타났다.

교외지에 대한 평균오차가 산악지와 시가지보다 크게 나타나는 이유로는 여러 종류의 데이터, 예를 들어 주택, 등고선, 도로, 기호등 복합적인 데이터 구조에 따른 근접 또는 중첩에 의하여 인식과정에서 영향을 주는 것으로 판단된다. 즉, 일률적인 데이터 구조만을 스캐닝 할 때의 정확도에 비해 떨어지는 것으로 판단되며, 느슨한 밀도와 평지에서의 평균오차의 경우에도 같은 이유인 것으로 판단된다. 그림 6, 그림 7, 그림 8은 X, Y 방향에 대한 특성별 분석 결과를 보여준다.

전체적으로 X방향오차의 평균오차보다 Y방향오차의 평균오차가 더 작게 나타나는데 이는 사용한 스캐너(드럼스캐너)의 회전방향(Y축 방향)과 스캐너의 인식방향(X축 방향)의 속도차에 따른 영향으로 판단된다.

5. 결 론

본 논문은 LIS/GIS의 D/B구축을 위한 수치지도의 오차 분석에 관한 연구로서 각각 특성이 다른 3 도엽(축척 1:5,000)의 수치지도로부터 지형별 특성, 밀도별 특성, 경사도별 특성에 따라 지형도와와의 X, Y 평면오차를 비교·검토하였으며 다음과 같은 결론을 얻었다.

1. 수치지도에서 발생하는 과대오차는 주로 래스터 데이터를 벡터데이터로 변환하는 과정에서 발생하는 것으로 판단된다.

2. X, Y 방향오차에 대한 분석 결과 X, Y의 방향에 관계없이 면구조의 데이터보다 선구조의 데이터에 대한 위치오차가 더 크게 나타남을 알 수 있었다.

3. 특성별 분류에 대한 분석결과로 지형별 분석에서는 교외지의 평균오차가 시가지나 산악지의 경우보다 크게 나타나고 밀도별 분석에서는 느슨한 밀도에서의 평균오차가 가장 크게 나타났으며 경사도별 분석에서는 평야에서의 평균오차가 크게 나타남을 알 수 있었다.

4. 회전통에 부착된 지형도를 인식하는 방법인 드럼

스캐너의 경우에 있어서는 X 방향의 평균오차에 비해 Y 방향의 평균오차가 크게 나오는 것으로 판단된다.

5. 앞으로 LIS/GIS D/B 구축에서는 대축척의 지적도 현황도 등을 수치화 하는데 따르는 투영오차, 전개오차, 기준점오차, 제작오차 등에 대하여 체계적으로 분석하고, 지적도와 지형도의 중첩에 관한 연구가 요망된다.

참고문헌

1. 백은기, 이영진, "등고선도의 자동작성에 관한 실험적 연구," 한국측지학회지, 제 2권, 제 1호, 1984, pp. 49-53.
2. 유복모, "지형공간정보론," 동명사, 1994.
3. 이석찬, "표준 측량학," 선진문화사, 1989.
4. 이석찬, 고영호, 이영진, "돌출오차의 검출과 측지망의 신뢰도," 대한토목학회 논문집, 제 7권, 제 1호, 1987, pp. 1-9.
5. 조규전, "표준측량학," 보성출판사, 1996.
6. Ansoult, M. M., P.J. Soille, and J.A. Loodts, "Mathematical Morphology: A tool for Automated GIS data Acquisition from scanned Thematic Maps," P.E. & R.S., Vol. 56, No. 9, 1990, pp. 1263-1271.
7. Bernhardsen, T. "Geographic Information System," 1992.
8. Chrisman, N. R. "The Error Component in Spatial Data" in "Geographic Information Systems(Vol. 1): principles," Longman, 1991, pp. 165-174.
9. Goodhild, M., R. Haining, and S. Wise, "Intergrating GIS and spatial data analysis: problem and possibilities," int. J. of Geographic Information Systems, Vol. 6, No. 5, 1992, pp. 407-423, 1992.
10. Greenlee, D. D. "Raster and Vector Processing for Scanned Linework," Auto-Carto 8, Maryland, 1987, pp. 639-640.
11. Peng, Z. and K. J. Dueker, "Error and Accuracy in Spatial Data Allocation," GIS/LIS Proceeding 93', Vol. 2., 1993, pp. 592-602.
12. Smith, T. R. and K. K. Park, "Algebraic approach to spatial reasoning," IJGIS, Vol. 6, No. 3, 1992, pp. 177-192.
13. Taylor, D. R. F. "Cartographic Visualization and Spatial data handling," Advanceed in GIS Research, Vol. 1, 1994, pp. 16-28.