

브이월드 데스크톱을 위한 대용량 공간정보 데이터 지원 방안 연구

A Study on the Improvement of Large-Volume Scalable Spatial Data for VWorld Desktop

강지훈* · 김현덕** · 김정옥***

Kang, Ji-Hun · Kim, Hyeon-Deok · Kim, Jung-Ok

초 록

최근 IT기술의 발달로 데이터의 양이 급증함에 따라 대용량 데이터를 다루는 빅 데이터 시대가 도래하게 되었다. 다양하고 방대한 빅데이터를 활용하기 위해서는 대용량 데이터 처리 기술을 필요로 하는데, 이러한 흐름에 따라 공간정보 분야에서도 대용량의 공간정보 데이터 서비스 기술이 요구되고 있다. 따라서 본 연구에서는 국외의 대표적인 공간정보 데이터 서비스 기술을 살펴본 후, 국내의 대표적인 지도 소프트웨어인 브이월드 데스크톱에 적용 가능한 KML 포맷의 대용량 데이터 처리 기술을 개발하였다. 개발된 분할 방법과 도구를 검증하기 위해 대용량의 KML 데이터를 이용하여 테스트를 진행하였고, 인덱스 파일과 분할 파일들이 생성되어 브이월드 데스크톱에서 가시화 되었다.

● **주요어** : KML, 공간정보, 대용량 데이터, 분할 방법

ABSTRACT

Recently, as the amount of data increases rapidly, the development of IT technology entered the 'Big Data' era, dealing with large-volume of data at once. In the spatial field, a spatial data service technology is required to use that various and big amount of data. In this study, firstly, we explained the technology of typical spatial information data services abroad, and then we have developed large KML data processing techniques those can be

* 한국국토정보공사 공간정보연구원 국토정보연구실 연구원 (전자우편: kangdaejang@lx.or.kr), 주저자

** 한국국토정보공사 공간정보연구원 국토정보연구실 연구원 (전자우편: hdkim@lx.or.kr), 교신저자

*** 서울대학교 공학연구소 urban informatics 연구센터 부센터장 (전자우편: geostar1@snu.ac.kr), 공동저자

applied as KML format to VWorld desktop. The test was conducted using a large KML data in order to verify the development KML partitioned methods and tools. As a result, the index file and the divided files are produced and it was visible in VWorld desktop.

• **KEYWORDS** : Keyhole Markup Language(KML), Geospatial Information, Large-Volume Data, Partition Method

1. 서론

최근 모바일 혁명, IoT(Internet of Things) 기술 등의 발달과 SNS의 확산으로 데이터의 양이 급증함에 따라 데이터의 생산, 유통, 소비 체계에 큰 변화를 주게 되었다(송태민, 2013). 이는 빅데이터 시대의 도래를 의미하는데, 국가정보화 전략위원회(2011)에서는 빅데이터를 ‘대용량 데이터를 활용·분석하여 가치있는 정보를 추출하고, 생성된 지식을 바탕으로 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술’로 정의하고 있다. 세계 각국에서는 빅데이터가 미래 국가 경쟁력에 큰 영향을 미칠 것으로 예측하여 국가 안전, 재난재해, 질병, 위기 등에 선제적으로 대응할 수 있는 빅데이터 기술 개발에 노력을 기울이고 있다(송태민 등, 2015).

빅데이터는 3V(Volume, Variety, Velocity)의 특성을 가지고 있고, 기본적인 텍스트 데이터에서부터 수치, 영상, 이미지, 센서, 위치, 공간데이터 등과 같이 다양한 유형의 데이터를 포함할 수 있다(김민수 등, 2013). 따라서 이와 같은 다양하고 방대한 데이터를 활용하기 위해 대용량 데이터 처리 기술이 요구되고 있다.

한편, IT기술의 발달은 공간정보 기반의 서비스 제공이 증대되는 데에도 큰 영향을 미쳤다. 그리고 공공데이터 개방 추세는 대중의 공간정보

활용을 더욱 촉진하게 되었다. 이러한 흐름에 따라 UN, USGS 등의 기관에서는 자신들의 콘텐츠를 공간정보와 융합하여 KML 형태로 제공하고 있다. KML(Keyhole Markup Language)은 XML 기반의 마크업 언어로 공간정보를 저장, 표현, 공유하기 위한 대표적인 파일 형식이다(김정옥 등, 2014). KML은 그 범용성을 인정받아 공간정보 분야의 국제 표준화 기구인 OGC(Open Geospatial Consortium)에서 표준 포맷으로 채택되었다. 따라서 KML을 지원하는 응용프로그램이 매우 다양한데, 구글어스, 플리커 등의 구글 제품 외에 ArcGIS, QGIS 등에서도 지원하고 있다(강지훈 등, 2014). 국내에서는 공간정보 플랫폼 역할을 위해 국가에서 구축한 브이월드가 대표적인 KML 저작도구라 할 수 있다.

브이월드는 국가의 공간정보를 통합, 가공하여 수요자가 원하는 정보를 쉽게 검색, 제공하고

〈표 1〉 구글 지도의 KML 지원 기준

항목	기준
정성적 가져올 수 있는 파일의 최대 크기 (원본 KML, 원본 GeoRSS 또는 압축된 KMZ)	3 MB
압축되지 않은 KML 파일의 최대 크기	10 MB
네트워크 링크의 최대 수	10 개
문서 전체 지형지물의 최대 수	1,000 개

(<https://developers.google.com/kml/documentation/mapsSupport>)

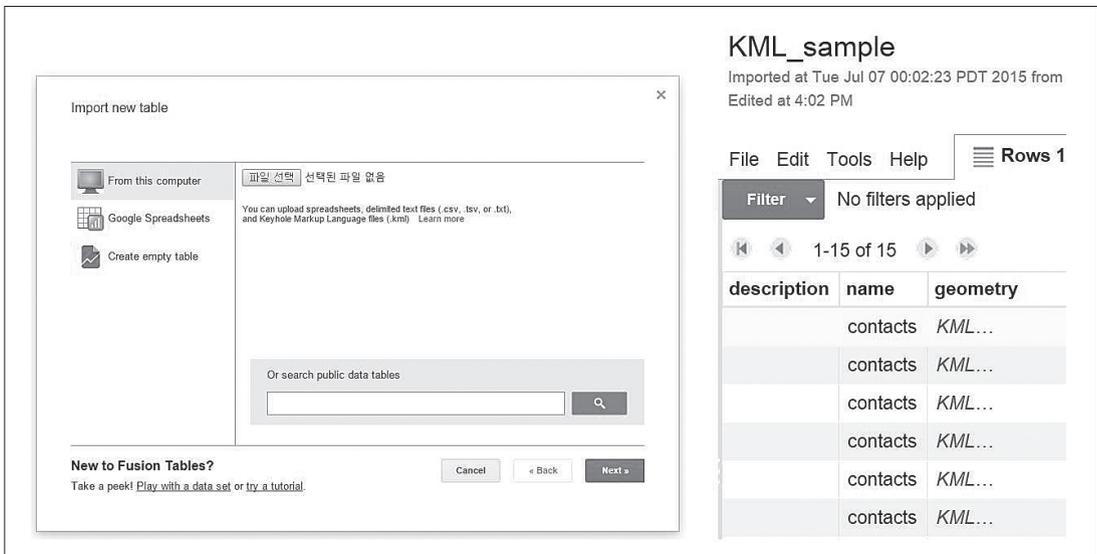
활용할 수 있도록 개방된 통합플랫폼이다. 브이월드 회원가입자 수는 2012년 6월 약 5천 명에서 2014년 6월 약 2만8천 명으로 5배가량 증가하였고, 전체 접속자수(오픈API 포함)는 2013년 상반기 월 평균 약 20만 명에서 2014년 상반기 월 평균 약 254만 명으로 10배 이상 증가하여 최근 사용자가 급증한 것을 알 수 있다(이인수 등, 2014). 그러나 브이월드 만족도 분석 연구에 따르면 몇 가지 문제점이 지적되고 있는데, 그 중 3D 데스크톱의 문제점으로는 KML, KMZ 불러오기의 오류 및 속도저하가 약 40%로 가장 높았다. 이는 3D 데스크톱 서비스가 사용자에게 고품질 데이터를 제공하는데, 컴퓨터 사양에 따라 속도저하 및 불러오기 오류가 다수 발생하는 것으로 인한 문제점으로 판단된다(황정래 등, 2014). 빅데이터 활용의 중요성이 대두되고 대용량

데이터 처리 기술의 수요가 급증함에 따라 브이월드에서도 대용량 데이터를 서비스 할 수 있는 기술 개발이 시급하다. 따라서 본 연구에서는 브이월드 데스크톱을 위한 대용량 KML 데이터 처리 기술을 개발하기 위해 대표적인 대용량 공간정보 데이터 서비스 기술들을 검토하고, 브이월드 데스크톱에 적용가능한 대용량 KML 데이터 처리 방법을 개발하여 제안하고자 한다.

2. 대용량 공간정보 데이터 서비스 기술

2.1. 퓨전 테이블

전 세계를 대상으로 하는 대표적인 웹 기반의 지도 서비스인 구글 지도¹⁾는 KML 데이터에 대해 가시화 가능한 파일 크기와 도형의 수를 <표 1>과 같이 제한하고 있다. 제시한 기준을 초과하는 도형의 수를 가진 데이터를 구글 지도에 호출하게 되



[그림 1] 퓨전 테이블의 사용자 환경

1) <https://www.google.co.kr/maps>

면 도형이 하나도 표시되지 않는다. 이와 같이 기준 이상의 KML 데이터를 구글 지도를 통해 가시화하기 위해서는 구글의 또 다른 서비스인 Fusion Table(이하 퓨전 테이블)을 이용할 수 있다.

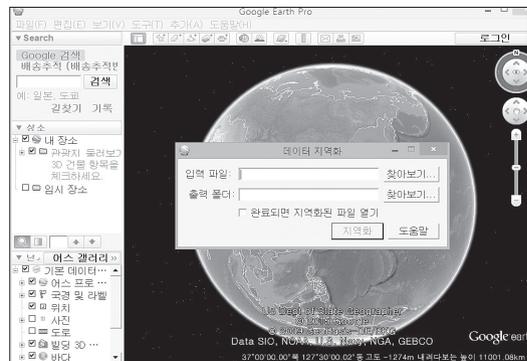
퓨전 테이블은 데이터 관리를 위해 구글에서 제공하는 웹 서비스이다(WIKIPEDIA Google Fusion Table, 2015). 기본적으로 스프레드시트 형태의 데이터 테이블을 저장, 공유, 가시화할 수 있고 사용자 환경은 [그림 1]과 같다. 특별히 가시화를 위해 구글 지도와 연동 기능을 제공하고 이 기능을 사용하면 기존 구글 지도가 지원하는 크기 이상의 KML 데이터를 가시화 할 수 있다. 퓨전 테이블은 point, line, polygon 형태의 KML 데이터를 지원한다. KML 포맷만을 지원하기 때문에, KMZ 포맷의 경우 압축을 풀고 업로드 해야 한다. KML 데이터를 업로드 하면 구글 지도와 연계를 위해 정보들이 geometry 열에 자동으로 채워진다. KML 데이터를 구성하고 있는 하나의 도형은 하나의 <Placemark> 태그로 구성되어 있고, 퓨전 테이블에서는 하나의 행을 생성한다.

KML 데이터는 보통 퓨전 테이블에서 3개의 열(name, description, geometry)을 가진 테이블 데이터로 변환 된다. 하나의 테이블은 10만 개의 행을 생성할 수 있는데, 이는 10만 개의 도형 요소를 가지고 있음을 의미한다. 이는 앞서 언급한 구글 지도의 도형 요소 지원기준 보다 매우 높은 기준이다. 또한 업로드 가능한 KML 데이터 크기는 압축되지 않은 KML 데이터 기준으로

100MB 까지 가능하다. 하나의 셀이 가질 수 있는 문자는 100만개로 제한하고 있고, 언급한 기준들을 넘는 KML 데이터는 업로드 할 수 없다.

2.2. 구글 어스 지역화 도구

사용자를 위한 지구본 형태의 대표적인 가시화 도구인 Google Earth(이하 구글 어스)²⁾는 가장 안정적이고 빠르게 KML 데이터를 표현할 수 있다. 구글 어스에서는 보유한 다양한 콘텐츠를 가시화 하거나 데이터를 저작하기 위한 기본적인 기능을 제공하고 있고, 구글 어스 프로는 구글 어스에 몇 가지 고급 기능을 더하여 서비스 하고 있다. 구글 어스 프로의 도구 메뉴에 Regionator(이하 지역화)를 선택하면 [그림 2]와 같은 사용자 환경을 가진 지역화 기능을 사용할 수 있다. 지역화 기능은 하나의 KML 데이터를 변환해주는 기능으로 하나의 부모 지역 아래 4개의 자식 지역 구조로 되어있으며 사용 방법은 다음과 같다(GOOGLE Regionator, 2007). 지역화 할 KML 파일을 선택하고 지역화 된 파일을 저장



[그림 2] 구글 어스 지역화 도구의 사용자 환경

2) <http://www.google.com/intl/ko/earth/>

할 컴퓨터의 폴더를 찾은 다음 확인을 클릭하여 데이터에 대한 지역화를 진행하고 완성된 파일을 저장한다. 지역화를 구현하기 위해 <Network-Link> 태그와 <Region> 태그를 사용해 기본 데이터를 가공한다. <NetworkLink> 태그를 통해 하나의 파일을 피라미드 구조로 하위 파일을 생성 후 연결하고 <Region> 태그를 이용해 가시화 영역을 제공한다. <Region> 태그는 성능 저하 없이 큰 데이터를 구글 어스에 로딩하기 위한 핵심 기능을 구현하는 태그로 사용자의 시야에 들어오고 화면의 특정 부분을 차지하는 경우에만 메모리에 로드하고 표시하게 된다. 지역화 기능은 구글 코드를 통해 Python 라이브러리로 제공하고 있다. 지역화와 관련된 코어 클래스는 3개 (Regionator 클래스, RegionHandler 클래스, Region 클래스)이나, 실제 사용을 위해서는 구글 어스와 관련되어 연결되어 있는 다양한 클래스를 해석해야하기 때문에 구현이 쉽지 않다.

2.3. 구글의 대용량 서비스 기술과 브이월드 데스크톱

대용량 KML 데이터를 사용하기 위해서는 단순히 오류 없이 읽어 들이는 것 뿐만 아니라 멈춤 현상 없이 작동해야 하지만, 단순히 지도 엔진의 성능만으로는 이를 해결하기 쉽지 않다. 하드웨어 성능에 따라 차이는 있지만, 도형 요소가 절대적으로 많아지게 되면 일반 데스크톱 성능으로는 멈춤 현상 없이 작동하기 어려워지게 된다. 따라서 앞서 살펴본 바와 같이 구글의 지도 플랫폼은 엔진의 성능 개선과 별도로 대용량 KML 데이터의 사용성을 개선하기 위한 별도의 서비스 또는 기능

을 가지고 있다. 하지만 국내에서 가장 많이 사용되는 브이월드는 이와 유사한 대용량 KML 데이터를 지원하는 기능이 아직 개발되지 않았다.

대용량 KML을 지원하는 구글의 서비스는 구글에서만 이용 가능하여 브이월드에는 이용할 수 없다. 퓨전테이블의 경우 웹 기반의 어플리케이션으로 구글 맵스 이외의 지도 플랫폼과는 호환되지 않기 때문에 브이월드 웹 지도 서비스와 연동하기 어려움이 있다. 구글 어스 지역화 도구의 경우 어플리케이션별 지원 명세의 차이로 인해 완벽한 호환이 이루어지지 않는다. <Network-Link>, <Region>, <LoD> 태그 등의 기능을 이용해 링크 형태로 대용량 KML 데이터를 작은 용량의 KML 데이터로 나누어 연동하는데, 구글 어스에 특화된 구조이므로 브이월드 데스크톱에서 바로 사용하기는 어렵다. 이와 같은 검토사항을 기반으로 사용자 KML 데이터의 활용을 지원하는 브이월드 데스크톱에 특화된 대용량 KML 데이터 지원 도구를 설계하였다.

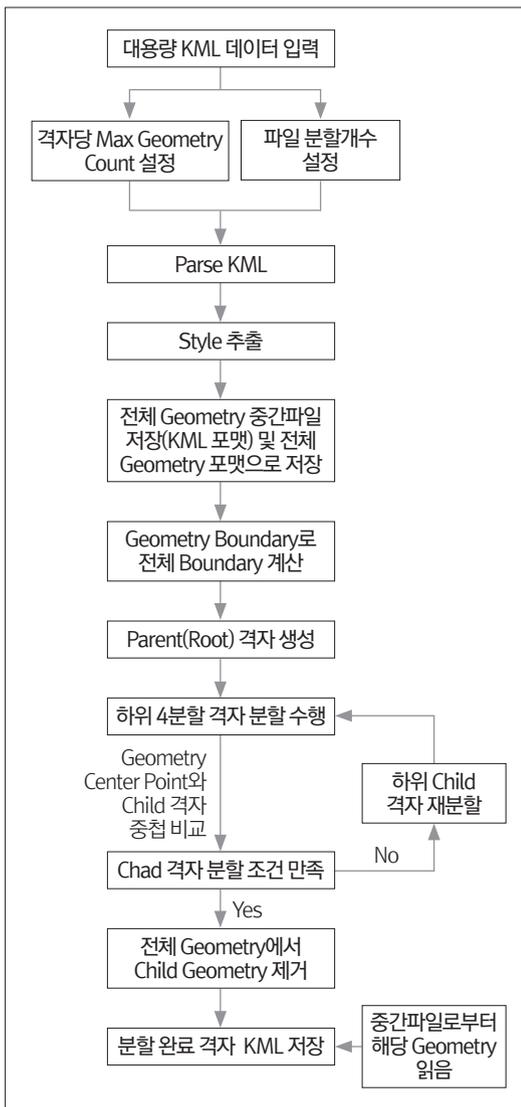
3. 브이월드를 위한 대용량 KML 지원 기술 개발

3.1. 브이월드 데스크톱을 위한 분할 방법

본 연구에서 언급하고 있는 대용량 KML 데이터는 큰 용량으로 인해 브이월드 데스크톱에서 사용이 어렵거나 사용성이 저하되는 파일을 의미한다. 이와 같은 대용량 KML 데이터를 브이월드 데스크톱에서 보다 나은 환경에서 사용가능하도록 위치를 기반으로 하는 분할 서비스를 개발했다. 위치를 고려하기 위해서는 KML 데이터에 포함된 도형요소를 특정 기준으로 증첩여부를 판

단해 분할해야 한다. 본 연구에서는 쿼드트리 방식을 사용하여 전체 도형의 영역을 분할하고자 하는 조건에 부합할 때까지 회귀적으로 분할하였다. 그리고 분할된 데이터가 여전히 대용량인 것을 방지하기 위해 분할된 KML 데이터의 최대 도형 개수와 최대 파일 사이즈를 분할을 위한 초기조건으로 입력하도록 했다.

대용량 KML 데이터는 다수의 도형을 가지고 있기 때문에 이와 같은 중첩여부 판단을 위한 연산에 많은 양의 메모리를 소모하게 된다. 따라서 이는 브이월드 데스크톱 다운현상의 원인이 될 수 있다. 따라서 이것을 개선하기 위해 연산에 불필요한 정보들은 중간 파일을 생성하여 컴퓨터 파일 시스템에 저장해 두고, 연산의 필수 정보만을 메모리에 저장하고 연산하여 사용하였다. 또한, 분할된 데이터가 브이월드 데스크톱에서 오류 없이 열릴 수 있도록 불필요한 태그 사용을 지양했다. 따라서 분할된 파일을 생성하는 과정에서 공간정보 표현을 위해 필수적으로 필요한 <Placemark> 태그의 좌표관련 하위 요소를 제외한 추가적인 태그는 삭제했다. 이와 같은 사항을 고려한 브이월드 데스크톱을 위한 대용량 KML 데이터의 분할 순서는 [그림 3]과 같다.



[그림 3] 분할방법 순서도

3.2. 대용량 데이터 입력과 중간파일 생성

대용량 KML 데이터 분할을 위해서는 우선 원하는 대용량 KML 데이터를 선택하고 분할을 위한 초기 조건을 입력한다. 조건은 크게 2가지로 최대 도형개수나 최대 파일용량을 입력할 수 있다. 최대 도형 개수는 분할된 한 영역에 위치할 수 있는 도형개수의 임계값이고 최대 파일용량은 분할된 파일용량의 임계값이다. 임계값은 자유롭게 설정할 수 있지만 입력 데이터의 도형개수나 파일용량보다 작은 값을 입력해야 분할된 데이터의 대용량화를 막고 분할을 통해 가시화 성능을 개선할 수 있다.

격자의 분할과 중첩 연산을 수행하기에 앞서 점유 메모리를 최소화하기 위해 대용량 KML 데

이터의 전체 도형요소에 대한 지오메트리(Geometry) 정보들만을 따로 추출해 [그림 4]와 같은 구조의 중간파일을 생성해 저장한다. 그리고 지오메트리의 중심점, 중간파일에서 해당 지오메트리가 저장된 위치, 파일 용량만을 사용하여 단순 도형 객체를 만들어 메모리에 저장한 후 분할 과정에 사용한다. 중심점 정보는 향후 중첩 연산의 기준으로 사용하고 중간파일에서의 위치는 분할 완료 격자의 도형정보를 KML 파일로 저장하는데 사용하며 용량정보는 초기 조건의 최대 파일 용량과 비교하기 위해 사용한다.

중간파일
Geomtry(kml) 1
Geomtry(kml) 2
Geomtry(kml) 3
Geomtry(kml) 4
Geomtry(kml) 5
Geomtry(kml) 6
Geomtry(kml) 7
Geomtry(kml) 8
Geomtry(kml) 9
Geomtry(kml) 10
Geomtry(kml) 11
Geomtry(kml) 12
...

[그림 4] 중간파일 구조

3.3. 초기 분할격자의 정의와 중첩 연산

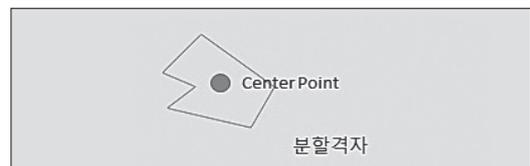
도형단순화 작업과 중간파일 생성작업을 위해 대용량 KML 데이터 전체를 해석하면서 전체 도형요소에 대한 최소 외접 사각형(Minimum Bounding Rectangle)을 생성하고, 이를 루트(root) 격자라 칭한다. 최소 외접 사각형은 X-Y

좌표계 상의 도형 정보에 대한 최소 x, 최대 x, 최소 y, 최대 y 값을 가진 정보이다. 그리고 루트 격자를 4개의 동일한 넓이를 가진 사각형으로 분할하여 [그림 5]와 같은 초기 분할 격자를 생성한다. 이와 같은 초기 분할은 대용량 KML 데이터의 분할을 위한 시작 조건으로 분할하고자 하는 모든 데이터에 적용되고, 분할된 격자는 자식(child) 격자라 칭한다.



[그림 5] 초기 분할격자의 정의

자식 격자와 단순 도형 객체의 중심점 정보를 비교하여 중첩 여부를 검사 한다. 중첩 조건은 [그림 6]과 같고, 이와 같은 조건을 만족하는 도형을 해당 분할격자 안에 있는 것으로 판단한다.



[그림 6] 중첩 판단



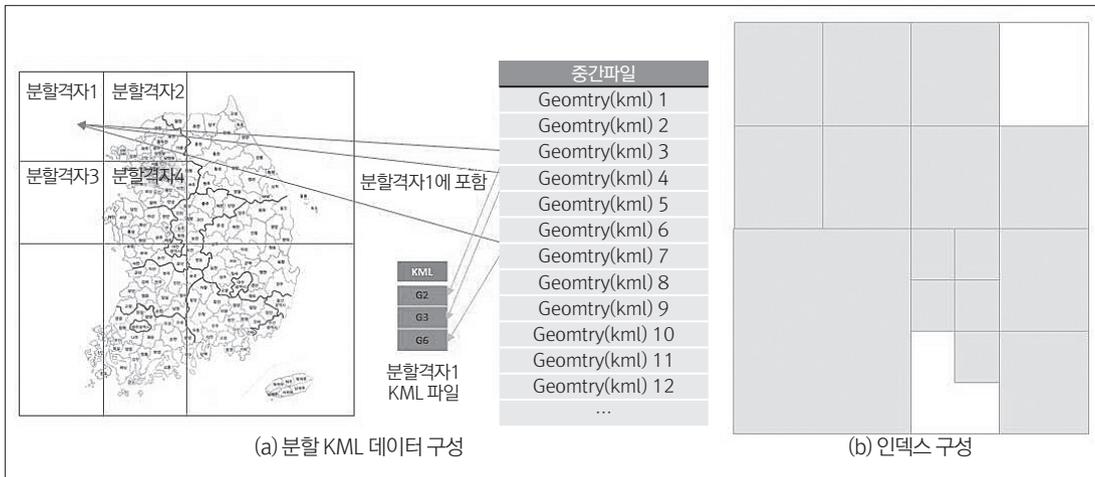
[그림 7] 재분할을 위한 분할격자 생성

기본적으로 단순화된 도형정보를 이용해 전체 도형요소에 대한 중첩 검사를 수행하고, 4개의 자식격자로 분류된 도형의 개수나 파일 용량을 초기 조건과 비교하여 작을 경우 해당 격자에 속한 단순 도형정보는 임시로 저장하고 분할을 완료한다. 하지만 초기 조건보다 큰 값을 가지고 있을

경우 [그림 7]과 같이 4개의 자식격자를 생성하고 재분할을 수행한다. 재분할을 위한 중첩 연산과 초기 조건과의 비교는 처음 분할 프로세스와 동일하다. 예를 들어 초기 조건으로 최대 도형의 개수를 1500개로 설정하고 초기 분할 결과로 [그림 5]의 분할격자1에 2000개의 도형이 분류될 경우, 해당 격자에 대해서는 재분할을 수행하게 된다. 이와 같은 재분할은 모든 자식 격자들이 초기 조건에 만족할 때까지 반복적으로 수행한다.

3.4. 분할된 KML 파일 생성

모든 분할 격자의 도형개수나 파일용량이 기준 값 이하가 되면, 각 분할 격자에 속한 도형 정보들로 구성된 KML 파일들과 인덱스 파일을 생성한다. 격자에 속한 단순도형 객체의 중간파일 위치 정보를 이용해 중간파일에서 해당 도형의 지오메트리 정보를 읽어와 [그림 8(a)]와 같이 분할 KML 파일에 저장한다. 인덱스 파일은 전체 분할 격자 구성현황을 기준으로 분할파일이 생성된 격자에 색상을 부여해 [그림 8(b)]와 같이 생성한다.



[그림 8] 분할 KML 파일과 인덱스 파일 생성

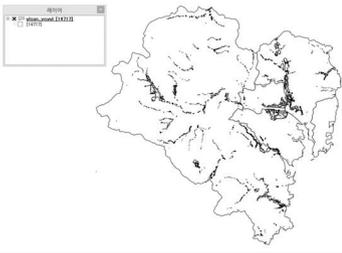
4. 대용량 KML 데이터 분할도구

4.1. 분할도구 평가 환경

연구에서 제안한 브이월드 데스크톱을 위한 대용량 KML 분할 방법의 효과를 검증하기 위하여 브이월드에서 열리지 않는 대용량의 KML 데이터를 이용해 분할 테스트를 수행하였다. 테스트를 위해 앞서 제안한 분할방법을 적용한 분할도구를 개발했다. 그리고 분할도구를 통해 분할

된 데이터를 브이월드 데스크톱에서 사용가능함을 확인했다. 테스트에 사용한 대용량 KML 데이터의 세부정보는 <표 2>와 같다.

<표 2> 테스트 파일의 세부 정보

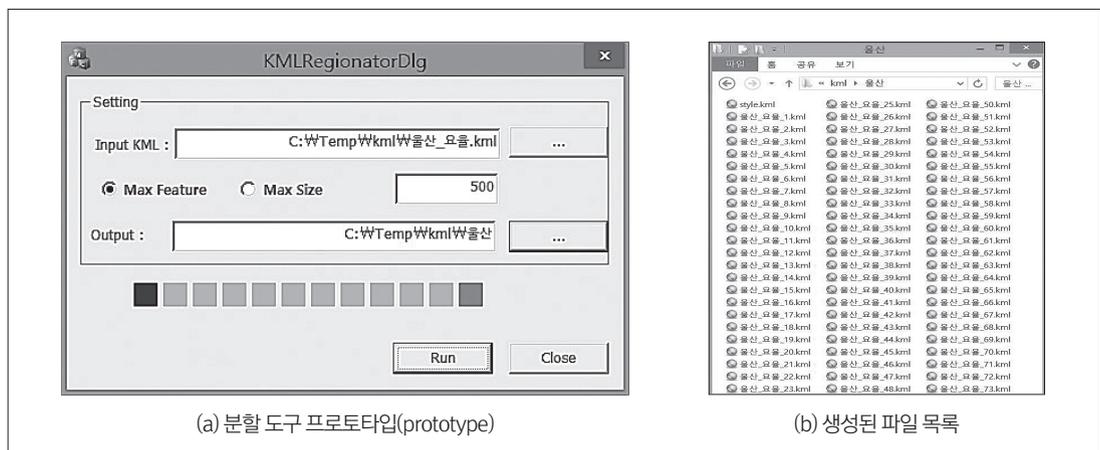
항목	내용
지역적범위	울산광역시
파일용량	21.3MB (압축되지 않은 KML)
지오메트리 형태	폴리곤
도형 개수	14,717
데이터 미리보기	

4.2. 분할 테스트

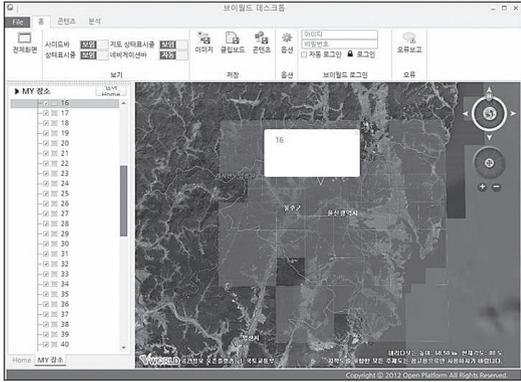
분할 테스트를 위해 [그림 9(a)]와 같이 프로토타입을 개발해 사용했다. 앞서 설명한 테스트 데이터를 선택하고 최대 도형개수 500개를 초기조건으로 설정한 후, 분할 파일을 저장할 폴더를 선택했다. 테스트 파일을 분할하는데 약 3초 정도 소요되었고 분할 결과 총 73개의 분할 파일과 인덱스 파일이 [그림 9(b)]와 같이 생성됐다.

파일명은 입력한 파일 이름 뒤에 숫자 1부터 순차적으로 증가시키며 붙여준다. 그리고 해당 파일들이 어느 격자를 위치시키는지 알기 위해 인덱스 파일의 격자도형 이름을 동일하게 부여했다. 따라서 분할을 완료 후 분할된 KML 파일을 열기 전 인덱스 KML 파일을 우선적으로 확인해야 한다.

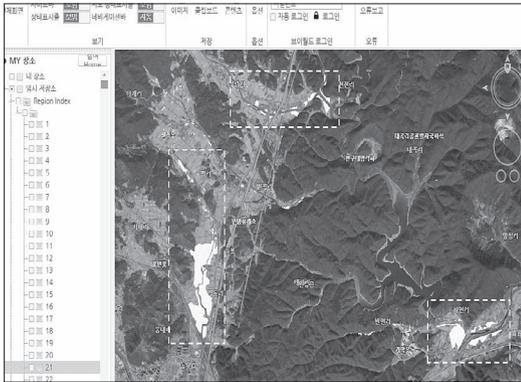
브이월드에서 인덱스 파일을 열면 [그림 10]과



[그림 9] 분할 도구와 분할된 파일 목록



[그림 10] 분할 KML 인덱스 파일



[그림 11] 분할된 파일(16번 격자) 가시화

같이 분할 파일이 있는 격자영역을 표현한다. 그리고 원하는 지역을 선택하면 좌측 격자리스트에서 해당 격자파일의 숫자가 선택된다. 현재 브이월드에서 안정적으로 구현하지 못하지만 글상자를 통해서도 파일의 숫자를 확인 가능하다. 이렇게 확인된 분할파일을 통해 기존에 대용량 KML 데이터에서 확인할 수 없던 정보를 확인할 수 있다.

예를 들어 원하는 격자지역이 데이터가 16번 분할파일인 것을 확인하고 해당영역의 영역 주변 파일들과 함께 열어보면 [그림 11]과 같이 점선 부분처럼 가시화되는 것을 확인할 수 있다.

5. 결론

최근 IT기술의 발달로 데이터의 양이 급증함에 따라 데이터의 생산, 유통, 소비 체계에 큰 변화가 오게 되었고, 대용량 데이터를 활용·분석하여 가치있는 정보를 추출하는 빅데이터 처리 기술의 수요가 증가하고 있다. 이러한 흐름에 따라 공간정보 분야에서도 대용량의 공간정보 데이터 서비스 기술이 요구되고 있다.

본 연구는 브이월드 데스크톱에서 지원하고 있는 KML 포맷에 대한 대용량 데이터 처리 기술을 개발하기 위해 국외의 대표적인 공간정보 데이터 서비스 기술을 살펴보았다. 국내·외를 망라해 가장 많은 사용자가 이용하는 구글 지도와 연계해 사용 가능한 퓨전 테이블은 구글 지도에서 직접 서비스 가능한 KML 데이터 용량 이상의 파일 크기와 도형의 수를 지원한다. 또한 사용자들에게 익숙한 엑셀형태의 사용자 환경을 지원하고 보유한 KML 파일을 업로드하면 자동으로 변환해 하나의 Placemark를 하나의 행에 넣어주기 때문에 누구나 쉽게 이용 가능하다. 하지만 구글 지도와 연계된 서비스이기 때문에 다른 지도 서비스에서 사용할 수 없다. 전문화된 사용자를 위한 구글 어스 프로의 지역화 도구는 대용량 KML 데이터를 고유 태그를 이용해 다수의 하위 KML 데이터로 분할해 성능을 개선한다. 하지만 태그를 사용하는 방식이 구글 어스에 최적화된 구조로 브이월드 데스크톱과 같은 다른 공간정보 데이터 가시화 소프트웨어의 경우 제대로 표현 못할 가능성이 많다.

검토된 기술들을 기반으로 브이월드 데스크톱에 사용 가능한 대용량 KML 데이터의 분할

방법을 개발하였다. 개발한 분할 도구에 대용량 KML 데이터와 분할 조건인 도형 개수 및 파일 크기를 입력하면, 전체 도형에 대한 최소 외접 사각형을 결정하고 사분할 하게 된다. 그리고 메모리 사용을 최소화 하기 위해 도형의 정보들을 중간파일을 생성해 따로 저장하고 분할을 위한 연산은 단순화된 도형정보만을 이용한다. 모든 분할격자가 초기 분할조건 이하가 되면 각 분할격자에 대한 KML 파일과 인덱스 파일을 생성하며 분할을 마친다. 제안한 대용량 KML 데이터 분할도구의 효과를 검증하기 위하여 브이월드에서 열리지 않는 대용량의 KML 데이터로 테스트를 진행했다. 도형 개수를 기준으로 분할하였고, 테스트 결과 인덱스 KML 파일과 분할된 KML 파일들이 생성되었다. 인덱스 KML 파일을 통해 분할 파일이 있는 격자영역을 확인할 수 있었고, 분할파일을 같이 열어봄으로써 분할된 지역이 가시화 되는 것을 확인할 수 있었다.

제안된 분할방안은 용량으로 인해 브이월드에서 사용할 수 없는 데이터를 사용 가능하도록 만드는 것을 목표로 하였다. 그리고 시범적으로 제작한 분할도구를 통해 분할을 통한 대용량 KML 데이터 활용 가능성을 확인 하였다. 이와 같은 KML 분할 서비스가 브이월드 데스크톱의 기능으로 추가된다면, 보다 다양한 사용자 데이터들이 활용될 수 있을 것으로 기대된다.

감사의 글

본 연구는 국토교통부 도시건축개발사업 연구비지원(13도시건축A02)에 의해 수행되었습니다.

【참고문헌】

- 강지훈·김정옥·이준석(2014), 「KML 호환성 개선을 위한 명세서 검증 방안」, 한국측량학회지, 제32권 제4호, pp. 353-361.
- 국가정보화전략위원회(2011), 「빅데이터를 활용한 스마트정부 구현」, p.2.
- 김민수·장인성·이충호·신성웅(2013), 「공간 빅데이터 관련 국내외 동향 분석」.
- 김정옥·강지훈(2014), 「오픈레이어 기반 사용자 친화적 KML 검증도구 설계」, 지적, 제44권 제1호, pp. 165-177.
- 송태민(2013), 「보전복지 빅 데이터 효율적 활용방안」, 한국컴퓨터정보학회지, 제21권 제1호, pp. 45-52.
- 송태민·송주영(2015), 「빅데이터 연구 한 권으로 끝내기」, 한나래, pp.16-18.
- 이인수·이준석(2014), 「공간정보 포털 동향 분석에 관한 연구」, 지적, 제44권 제2호, pp. 125-138.
- 황정래·강진아(2014), 「브이월드 서비스 개선을 위한 사용자 만족도 조사 및 분석에 관한 연구」, 한국공간정보학회지, 제22권 제6호, pp. 23-32.
- Google Project Hosting(2007), Regionator, Google, 미국(<http://code.google.com/p/regionator/wiki/RegionatorAlgorithm>).
- Wikipedia(2015), Google Fusion Tables, Wikimedia Foundation, 미국(http://en.wikipedia.org/wiki/Google_Fusion_Tables).