

Zur Standardisierung von Geo-Rasterdiensten – Interoperable Mehrwertdienste für große Rasterarchive

Zusammenfassung

Immer mehr werden Rasterdaten integraler Bestandteil von Geodiensten. Längst ist die Hardware leistungsfähig genug, um auch Objekte in der Größenordnung von Tera- bis Petabytes im Online-Zugriff zu halten. Im nächsten Schritt entstehen daraus Web-basierte Dienste, für die sich damit die Frage nach standardisiertem, interoperablem Zugriff stellt.

Im Open GeoSpatial Consortium (OGC) werden seit längerem u.a. Standards für Georaster-Dienste entwickelt. Kürzlich wurde die Version 1.1 des Web Coverage Service (WCS) freigegeben, und der Web Coverage Processing Service (WCPS) als dynamischer Mehrwertdienst auf Rasterdaten ist in Vorbereitung. WCS und WCPS sollen in diesem Beitrag vorgestellt werden.

Der Autor ist Mitglied der OGC WCS Revision Working Group, Chair der WCPS Working Group und Co-Chair der Coverages Working Group.

Abstract

Increasingly raster data are becoming integral component of geo services. By far current hardware is powerful enough to allow online access even to objects of Terabyte to Petabyte size. As a next step, Web-based services emerge, and this immediately raises the question of standardised, interoperable access. In the Open GeoSpatial Consortium (OGC) geo raster services are being developed as part of their geo service standards. Recently version 1.1 of the Web Coverage Service (WCS) has been released, and the Web Coverage Processing Service (WCPS) as a dynamic value-added raster service is under preparation. Both WCS and WCPS are presented in this contribution. The author is member of the OGC WCS Revision Working Group, Chair of the WCPS Working Group, and Co-Chair of the Coverages Working Group.

OGC Web Map Service (WMS) und Virtual Globes zeigen Wege zumindest für die visuelle Darstellung von Rasterobjekten mit x/y- und x/y/t-Dimensionen auf.

Zusätzlich stellen sich aber zunehmend auch Integrationsanforderungen: Klimaforscher koppeln atmosphärische und ozeanographische Modelle über die Daten. Belüftungssimulationen in Gebäuden führen zur Integration externer Klimadaten mit CAD-Daten. Geschäftsdaten werden mit Raumbezug versehen und dazu mit Geodaten verknüpft. Diese willkürlich herausgegriffenen Beispiele zeigen, dass künftig der Informationsintegration eine wichtige Rolle zukommen wird. Langfristig ist zu erwarten, dass die zunehmende Durchdringung mit Geodiensten dazu führt, dass Geodaten sowohl mit technisch-wissenschaftlichen als auch administrativ-betriebswirtschaftlichen Daten gekoppelt werden müssen.

Die Formulierung von hinreichend abstrakten und allgemeinen, aber doch verständlichen Rasterdienste-Definitionen ist daher eine der großen Herausforderungen: Wie sehen Dienste auf Rasterdaten sinnvollerweise aus, wenn sie Geo-Raster und auch Rasterdaten aus angrenzenden Gebieten unterstützen sollen? Wie betten sich Rasterdienste in die allgemeine Orchestrierung von Geo-Diensten ein? Wie kann eine einfache, effiziente und flexible Ankopplung von client-seitigen Werkzeugen gestaltet werden? Auf der Implementierungsebene stellt sich die Performance-Frage, wenn Terabyte- bis Petabyte-große Objekte interaktiv abgerufen werden.

Das Open GeoSpatial Consortium (OGC, www.opengeospatial.org) arbeitet seit Jahren an Standards auch für Geo-Rasterdienste. In der modularen OGC-Familie von Standards sind dies insbesondere der Web Coverage Service (WCS) sowie der Standard-Vorschlag Web Coverage Processing Service (WCPS). Der Web Map Service (WMS) fällt teilweise ebenfalls darunter. Neuere

AUTOR

Peter Baumann

Rasterdienste

Noch im Jahr 2002 wurde dem Autor dieses Beitrags in einer Vermessungsbehörde vorgehalten, „unsere Orthophotos müssen wir nicht nach außen anbieten, die können sowieso nur Experten verstehen“. Ansichten und Zeiten ändern sich, und spätestens mit GoogleEarth hat sich allgemein die Erkenntnis durchgesetzt, dass auch Rasterdaten in der Orchestrierung ganzheitlicher Geo-Dienste einen wichtigen Platz einnehmen.

Dabei bilden Orthophoto-Karten aus Luft- und Satellitenbildern lediglich

die Spitze des Daten-Eisbergs; ebenso zu berücksichtigen sind 3D-Zeitreihen (x/y/t), 3D-Geophysik/Explorations-Daten (x/y/z), 4D-Klimamodelle (x/y/z/t) und viele weitere. Eine zunehmende Zahl von benachbarten Disziplinen geht ähnliche Wege wie die Geographie, sie konsolidieren ihre Datenbestände und stellt sie per Web für die Expertengemeinde und die Allgemeinheit zur Verfügung: Geophysik, Geochemie, Ozeanographie, Klimamodellierung sowie insgesamt die Erdsystem-Forschung sind hier exemplarisch zu nennen. Datenportale bieten heute üblicherweise komfortable katalogbasierte Suche, diese führt jedoch oft bei den Basisdaten lediglich zu einem einfachen Datei-Download. Mehr Komfort und Flexibilität ist hier gefragt, und Dienste wie

Dienste mit Raster-Komponenten sind der Web Processing Service (WPS) und die Standard-Familie Sensor Web Enabling (SWE).

Im vorliegenden Beitrag wird aus der Arbeit des OGC zum aktuellen Stand der Rasterdienste WCS und WCPS berichtet. Neben den konzeptuellen Modellen der Standards wird eine exemplarische Architektur vorgestellt. Weiterhin werden WCS und WCPS in den Kontext der anderen OGC-Standards gestellt. Eine Zusammenfassung des aktuellen Stands und Ausblicke auf künftige Herausforderungen beschließen den Beitrag.

WCS

Der Begriff Coverage ist in der OGC Abstract Specification (OGC, 2000) und ISO 19123 (ISO, 2003) sehr allgemein gefasst als ein "space-varying phenomenon"; derzeit betrachtet sind allerdings nur reguläre Gitternetze, also Rasterdaten. Im weiteren sei daher „Rasterobjekt“ synonym zu Coverage gebraucht; „Zelle“ bezeichne die wertetragenden Gitterpunkte.

Ein WCS (OGC, 2006b) erlaubt den Zugriff auf Rasterobjekte mit raum-zeitlicher Semantik. Konzeptuell ist eine Coverage definiert durch die Domäne (domain) und den Wertevorrat (range). Die Domäne besteht aus den raum-zeitlichen Achsen: x und y sind obligatorisch, z (Höhe/Tiefe) und t (Zeit) jeweils optional. Entlang jeder Achse ist die Domäne durch Unter- und Obergrenze definiert, so dass sich als Gesamtausdehnung ein achsenparalleler Würfel ergibt.

Der Wertevorrat definiert programmiertechnisch gesehen den Pixel- oder Voxeltyp der Zellen. Eine Zelle kann einen atomaren Wert tragen oder einen zusammengesetzten Wert, deren Felder (fields) über lokal eindeutige Namen angesprochen werden. Alternativ kann jedes Feld selbst wieder individuell zusammengesetzt sein, und zwar als multi-dimensionales Array mit benannten Achsen. Abzuwarten bleibt, ob diese doch recht komplexe Array-Substruktur praktischen Zuspruch findet; vermutlich wird in aller Regel die einfache Feldliste genügen.

Schließlich kennt ein Rasterobjekt noch eine nichtleere Menge von Koordinatensystemen, in denen Resultate zurückgeliefert werden können. Für die ggf.

notwendige Umprojektion muss pro Feld die einzusetzende Interpolationsmethode nearest [neighbour], linear, quadratic, cubic, none angegeben werden. Pro Rasterobjekt lässt sich optional eine Menge von Nullwerten definieren, also Zellenwerte, die „unbekannt“ bedeuten. Mit dem optionalen Parameter Nullwiderstand (null resistance) lässt sich festlegen, wie sich Nullwerte bei Interpolation auswirken; mögliche Einstellungen sind full, half, none und other.

Nach intensiven Diskussionen wurde auf Wunsch eines Herstellers in der Coverage-Beschreibung zusätzlich eine hierarchische Gruppierung von Rasterobjekten eingeführt, welche es erlaubt, Rasterobjekte in Hierarchien beliebiger Tiefe zu gruppieren. Zwischen den Elter/Kind-Knoten ist für die Coverage-Konstituenten BoundingBox, Liste der unterstützten Koordinatensysteme und unterstützte Ausgabeformate eine Vererbungsdefinition, welche jedoch teilweise übergangen werden kann. Ansonsten beinhaltet die Hierarchie keine weitere Semantik und ist auch nicht für den Raster-Zugriff zu verwenden. Diese Struktur wurde nach intensiven Diskussionen auf Wunsch eines Herstellers in der Coverage-Beschreibung, allerdings nicht für den Rasterzugriff erlaubt, da die Semantik weder wohldefiniert noch konsistent ist. Als einfaches Beispiel sei die Vererbung der BoundingBox genannt – sie wird „vererbt wenn nicht anders spezifiziert“; damit kann auf der Ebene einer übergeordneten „abstrakten“ Coverage keinerlei Aussage mehr getroffen werden, ob ihre untergeordneten Elemente innerhalb der BoundingBox des Elternelements liegen oder nicht. Von einer Verwendung dieses Konstrukts kann daher nur abgeraten werden.

Zusammengefasst besteht ein Rasterobjekt i.w. aus den Rasterdaten selbst, einem Namen, einer Menge von erlaubten Koordinatensystemen, Nullwerten und Interpolationsmethoden. Viele Angaben sind optional und mit praktisch sinnvollen Voreinstellungen belegt, um die Handhabung zu erleichtern.

Auf dieser Struktur stellt der WCS ein festes Set von Operationen zur Verfügung: raum-zeitliche Ausschnittsbildung (domain subsetting), Werteextraktion (range subsetting), Skalierung, Umprojektion in ein anderes Koordinaten-Re-

ferenzsystem und schließlich Codierung in einem wählbaren Datenformat (data format encoding).

Die Dienstspezifikation folgt der allgemeinen OGC-Dienststruktur. Zu Beginn einer Sitzung fordert der Client mit einem GetCapabilities-Request eine Übersicht über die verfügbaren Rasterobjekte an. Zu jedem der angebotenen Rasterobjekte lassen sich Details mit einem DescribeCoverage-Request erhalten. Auf Basis dieser Informationen lässt sich schließlich ein GetCoverage-Request absetzen, um Rasterobjekte oder Ausschnitte davon abzurufen. Wie üblich lassen sich Anfragen als http GET-Request (in key-value-pair (KVP) Notation) und POST Request (in XML) formulieren. Die zugehörigen XML Schema Definitionen basieren auf GML (Geography Markup Language) 3.1.1 und OWS Common Implementation Specification 1.0.0. Im folgenden exemplarischen GetCoverage-Request wird der einfacheren Lesbarkeit halber KVP-Notation verwendet. Die Anfrage adressiert ein Rasterobjekt AVHRRcube, um daraus einen x/y-Ausschnitt gemäss der blauen BBOX-Angabe zu bilden; entlang der Zeitachse wird die jüngste Zeitscheibe selektiert. Aus den Feldkomponenten wird radiance und temperature extrahiert. Als Rückgabeformat der Daten, skaliert auf die Größe 246x300, wird HDF-EOS (Hierarchical Data Format - Earth Observing System) gewählt, ein gängiges Datenaustauschformat in der Fernerkundung, für das ein WCS Application Profile definiert ist. Eventuelle Fehlermeldungen sollen in XML zurückgeliefert werden:

`http://myServer/wcsServlet?`

`VERSION=1.1.0 & SERVICE=WCS &
REQUEST=GetCoverage &`

`COVERAGE=LandSatCube &`

`RANGESUBSET=radiance;temperature`

`SRS=EPSG:31464 &`

`BBOX=4636000.0,5717000.0,4687000.0,5768000.0 & TIME=max &`

`WIDTH=246 & HEIGHT=300 & DEPTH=1 &`

`FORMAT=HDF-EOS &`

`EXCEPTIONS=application/vnd.ogc.se_xml`

Das Ergebnis einer WCS-Anfrage wird entweder direkt zurückgegeben oder im Server für weiteren Abruf gespeichert; der Request-Parameter store gibt dem Client die Kontrolle darüber. ►

Eine Unklarheit in der aktuellen WCS-Version 1.1 ist die Frage der Dimension eines Raster-Ergebnisses: Ist das Resultatobjekt der obigen Anfrage dreidimensional mit Dicke 1 oder zweidimensional? Wichtig ist dies beispielsweise bei x/t-Schnitten: man möchte nicht ein 3D-Objekt mit y-Dicke 1, sondern ein 2D-Objekt erhalten, welches in einem 2D-Format wie TIFF oder JPEG im Client präsentiert werden kann.

Anwendungsbeispiel

Im Rahmen des EU-Projekts ESTEDI (www.estedi.org) wurde vom Deutschen Fernerkundungs-Datenzentrum (DFD) der DLR auf Basis des RasterServers rasdaman (vgl. Abschnitt „Referenzimplementierung“) ein WCS-Dienst aufgebaut. Ungefähr 10.000 AVHRR-Bilder wurden zu einer blattschnittfreien Zeitreihe von Europa und Mittelmeer, also einem 3D x/y/t-Würfel zusammengefügt (Abb. 1). Die Anfragen werden vom rasdaman-System mit seiner Raster-Anfragesprache rasql abgearbeitet, das komplette Request-Ergebnis die Bilddaten selbst liegen in einer Oracle-Datenbank. Die raum-zeitliche Ausschnittsbildung wird von der graphischen Schnittstelle in eine WCS-Anfrage verschlüsselt und zum Server übertragen. Dort wird aus den WCS-Parametern eine rasql-Anfrage generiert und diese von rasdaman ausgeführt. Das fertige Ergebnis wird schließlich an den Client zurückgegeben. So lassen sich 1D-Zeitreihen (z.B. „Temperaturkurve über Moskau für ein Jahr“), 2D-Schnitte zu gegebenen Zeitpunkten (z.B. „Meerestemperatur um Griechenland am 20. Juni 2005“) sowie 3D-Teilwürfel (z.B. „Oberflächentemperatur für die letzten fünf Jahre zwischen Italien, Sizilien, Korsika und Sardinien“) extrahieren.

Ein Problem entstand beim Testbetrieb des DFD-Dienstes aus dem einfachen Zugriff: per Mausklick kann der Benutzer seine gewünschten Ausschnitte beliebig konfigurieren – muss allerdings dann das Anfrageergebnis nach wie vor über ein bandbreitenbegrenztes Netz empfangen. Gerade die einfache Handhabung machte bei manchen Anwendern vergessen, dass Gigabyte-große Anfrageergebnisse nach wie vor einige Transportzeit benötigen. Daher wurde zuerst eine Warnung an die Benutzer auf die Suchseite eingefügt, und im späteren Verlauf eine Mengenbe-

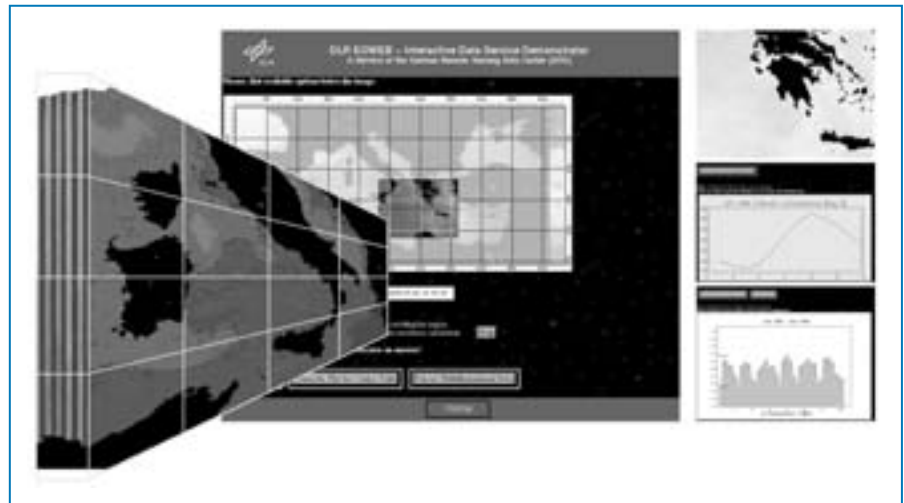


Abb. 1: WCS auf 3D-AVHRR-Zeitreihen: GUI (mitte), 3D-Ergebniswürfel (links), 2D-Zeitscheibe (oben rechts), 1D-Zeitreihen (mitte und unten rechts). Daten mit frdl. Genehmigung des DFD, www.dfd.dlr.de

schränkung eingeführt. Systemseitig kann die einstellbare transparente Transportkompression von rasdaman etwas Abhilfe schaffen, dies hängt aber naturgemäß sehr stark vom Kompressionsverhalten des Datenmaterials ab.

Bei vertikalen Schnitten war die Bearbeitungszeit anfangs unbefriedigend. Wie sich herausstellte, war als Ablagestruktur in rasdaman ein Stapel von ein-pixel-hohen Scheiben gewählt worden; dies spiegelte die Struktur der Eingabedaten wider, also 2D-x/y-Bilder. Bei dieser „blätterteig-artigen“ Struktur muss der Server bei einem vertikalen Durchstich für jedes einzelne Pixel eine ganze Scheibe (entsprechend der Größe einer AVHRR-TIFF-Datei) lesen, was natürlich hochineffizient ist. Nach Konsultation mit dem Hersteller nutzte die DLR einen speziellen Tuningparameter von rasdaman, um die Kachelungsstruktur an das Retrievalverhalten dergestalt anzupassen, dass vom Server beim Einfügen der TIFF-Bilder automatisch eine Partitionierung in kleinere Würfel erfolgte (die Gitterlinien in Abb. 1 symbolisieren dies). Mit dieser Struktur lassen sich nun Schnitte und Durchstiche entlang allen drei Dimensionen gleich effizient bearbeiten.

WCPS

Mit WCS ist eine gute Basis für die wichtigsten Funktionen auf Rasterobjekten gelegt. Mehrfach wurde allerdings an die Arbeitsgruppe herangetragen, auch diese oder jene weitere Verarbeitungsfunktion

in den Standard aufzunehmen. Ein typisches Beispiel bildet der NDVI (Normalized Difference Vegetation Index), zur Bestimmung der Vegetation aus nahem Infrarot- (nir) und Rot-Kanal (red). Er ist mathematisch wohldefiniert als

$$NDVI = (nir - red) / (nir + red)$$

Der Ausdruck wird pixelweise ausgewertet und liefert ein Bild in der Originalauflösung mit Pixelwerten zwischen -1 und +1; hohe Werte signalisieren Vegetation in dem jeweiligen Pixelgebiet.

Ein weiteres Argument für mehr Funktionalität liefert eine Beobachtung, welche von Kerstin Kleese am CCLRC/UK berichtet wird (Kleese, 2000). Wie Auswertungen zeigen, laden die Nutzer von Klimamodell-Daten um einen Faktor 10 höhere Datenvolumina von den FTP-Servern der dortigen Rechenzentren als sie eigentlich benötigen. Der Grund: Das unflexible Angebot von Dateien in exakt der Form, wie sie vom Großrechner erzeugt wurden, zwingt sie zum Herunterladen umfassender Datensets mit anschließender Aufbereitung auf Clientseite. Offensichtlich ließe sich bis zu dem genannten Faktor 10 an Datentransfer einsparen, könnten die Nutzer als Anfrage ihren exakten Bedarf formulieren. Ähnliche Erfahrungen existieren am Deutschen Fernerkundungs-Datenzentrum (DFD) der DLR mit dem Abruf von originalen und prozessierten Fernerkundungsdaten.

Wiewohl also nutzerspezifische Funktionsangebote Vorteile versprechen, blieb die Revision Working Group jedoch be-

wusst bei der Basisfunktionalität, um die WCS-Spezifikation nicht zu überfrachten. Statt dessen wurde mit dem Web Coverage Processing Service (WCPS) ein weiterer Dienststandard für Mehrwert-Funktionalität vorgeschlagen (OGC, 2006a). An Stelle einer zwangsläufig immer unvollständigen, schwer in ein Gesamtkonzept zu integrierenden Liste von Einzelfunktionen bietet WCPS eine Coverage-Anfragesprache; beliebig komplexe Coverage-verarbeitende Ausdrücke lassen sich dynamisch zu einer Anfrage zusammensetzen, welche vom Server ausgewertet wird. Die Coverage-Antwort genügt der Definition der WCS-GetCoverage-Response, ist also voll kompatibel.

WCPS erweitert dazu WCS um den zusätzlichen Request ProcessCoverage. Dem Server wird ein XML-codierter Ausdruck in der WCPS-Raster-Anfragesprache übermittelt, dessen Ergebnis entweder ein oder mehrere Rasterobjekte oder aber skalare Daten (z.B. aggregierte Daten oder Metadaten) zurückliefert. Zur besseren Lesbarkeit ist neben der XML-Syntax auch eine „Abstract Syntax“ definiert, welche sich an XQuery (W3C, 2007) orientiert.

Das folgende Beispiel implementiert den NDVI als WCPS-Request für drei serverseitige Rasterobjekte ModisA, ModisB und ModisC; das Ergebnis wird in HDF-EOS zurückgeliefert:

```
for c in ( ModisA, ModisB, ModisC )
return
  encode( ( c.nir - c.red ) / ( c.nir + c.red ),
    "HDF-EOS" )
```

Die for-Klausel listet die zu inspizieren-den Rasterobjekte und bindet die Laufvariable c daran. Der Raster-Ausdruck in der return-Klausel generiert für jeden Wert von c ein Rasterobjekt, in dem der Rasterausdruck im ersten Parameter der encode-Anweisung ausgewertet wird und das resultierende Rasterobjekt in das Zielformat HDF-EOS, welches im zweiten Parameter angegeben ist, codiert wird. Der Rasterausdruck greift auf die nir- und red-Komponenten der Modis-Szenen zu und bildet pixelweise den NDVI. Komponentenzugriff und arithmetische Operationen bilden eine wichtige Klasse der WCPS-Operationen; tatsächlich sind sämtliche aus Programmiersprachen bekannten arithmetischen, logischen, trigonometrischen und logarithmischen Operationen verfügbar.

Eine weitere wichtige Klasse von Operationen stellt die Ausschnittsbildung dar. Hier sind insbesondere Trimmen (trim) und Schneiden (section) zu nennen. Beim Trimmen wird an Hand einer Ober- und Untergrenze entlang einer oder mehrerer Dimensionen („bounding box“) ein Teilbereich eines Rasterobjekts ausgeschnitten. Beim Schneiden wird entlang einer Koordinate ein Schnitt durch das Objekt gelegt, so dass sich die Dimension reduziert. Das folgende Beispiel kombiniert beide Operationen und extrahiert aus der x/y/t-AVHRR-Zeitreihe aus Abb. 1 eine Temperaturkurve am Punkt x0/y0 und liefert sie als komma-separierte Werte (der Wildcard-Operator „*.“ bedeutet „alles von der Untergrenze bis zur Obergrenze des Objekts“ in der angegebenen Dimension):

```
for a in ( AVHRRCube )
return
  encode( a[ x0, y0, *.* ], "CSV" )
```

Koordinaten lassen sich als Pixelkoordinaten oder in einem definierten räumlichen oder zeitlichen Koordinatenwertes (jeweils unter Angabe des verwendeten Koordinatensystems) angeben, etwa „Sun May 20 00:04:00 CEST 2007“ für eine Zeitkoordinate.

Aggregationsoperationen erlauben es, summarische Werte aus Rasterobjekten abzuleiten. Neben den gängigen Aggregaten wie Summe, Mittelwert, Maximum und Minimum erlaubt die WCPS-Sprache die Definition eigener Aggregate.

In einem weiteren Anwendungsszenario interessiert den Anwender lediglich der prozentuale Anteil von Vegetation in

dem Ausschnitt gegeben durch die Eckpunkte (2001,4001) und (3000,5000), bei einem vorgegebenen Schwellwert für den Vegetationsindex von 0.6 (Abb. 2). Die folgende Anfrage bestimmt die Anzahl der Vegetations-Pixel, normiert sie auf die Größe des betrachteten Gebiets und liefert als Ergebnis eine einzige skalare Größe, eine Häufigkeit zwischen 0 und 1, zurück:

```
for c in ( ModisA, ModisB, ModisC )
return
  count_cells(
    ( ( c.nir - c.red ) / ( c.nir + c.red ) > 0.6 )
    [2001:3000,4001:5000] )
    / ( 1000 * 1000 )
```

Der vorherige NDVI-Ausdruck wird hier zuerst gegen den Schwellwert 0.6 verglichen, was in einer booleschen Matrix resultiert (true = „Vegetation“, false = „keine Vegetation“). Anschließend wird der verlangte Ausschnitt gebildet. Der Einfachheit halber wurden im Beispiel Pixelkoordinaten gewählt; es können jedoch, je nach Art der Dimensionsachse, ebenso Geokoordinaten, zeitliche Koordinaten oder abstrakte Intervallangaben sein; ein Beispiel für den letzteren Fall wäre Elevation, Druck oder Temperatur, wenn diese als eigene Dimensionen eines Datenwürfels angelegt sind.

Der count_cells()-Operator zählt die true-Pixel, welche im letzten Schritt auf die Gesamtgröße des 1000x1000-Ausschnitts normiert wird. Am Ende wird die resultierende Gleitpunktzahl zum Client zurücktransferiert.

Sofort fällt auf, dass die geschilderte Abarbeitungsreihenfolge hochgradig ►



Abb. 2: MODIS-Szene, links: Falschfarben-Komposit (infrarot/rot/grün); rechts: Vegetationsindex mit Schwellwert 0.6 (weiß=true, schwarz=false) Bilder generiert vom WCPS der Jacobs University, Daten: NASA

ineffizient ist: sicherlich wird man zuerst die Ausschnittsbildung vornehmen, um dann die Pixelinhalte zu inspizieren. Andererseits ist die obige Formulierung die kürzeste und übersichtlichste, sollte also dem Anwender nicht verwehrt werden. Optimierungstechniken, wie sie seit langem in Datenbanksystemen eingesetzt werden, können hier helfen. Im vorliegenden Fall kann der Server an Hand von algebraisch formulierten Ersetzungsregeln die effizienteste äquivalente Anfrage bestimmen und die Ausschnittsbildung vor der NDVI-Berechnung stellen. Die algebraische Optimierung von Raster-Anfragen wurde von Ritsch untersucht (Ritsch, 1999) und im rasdaman-System durch ca. 150 Regeln implementiert.

Die obigen Beispiele sollen den Mechanismus der WCPS-Anfragesprache vermitteln; die Spezifikation selbst enthält eine Anzahl weiterer Operationen, welche u.a. die wichtigsten Signal- und Bildverarbeitungsoperationen subsumieren, Reprojektion erlauben sowie die Extraktion von Beschreibungsinformation (z.B. in welchen Projektionssystemen ein Rasterobjekt abgefragt werden kann). Insgesamt reicht die Ausdrucksmächtigkeit der WCPS-Sprache von einfachen Bildoperationen (wie Addition und Überlagerung) über statistische Aggregationsoperationen bis zu Filterkernen und Konvolutionsooperatoren. Dabei bleibt die Sprache sicher in der Auswertung – jede mögliche Anfrage terminiert nachweisbar nach endlich vielen Schritten, so dass der Server nicht unbegrenzt lange blockiert werden kann.

Die Semantik der Sprache ist for-

mal definiert und somit sowohl zwischen Menschen als auch zwischen Maschinen ohne Informationsverlust kommunizierbar. Insbesondere kann automatische Lastverteilung bzw. „service chaining“ realisiert werden; dies ist z.B. mit den Konzepten des allgemeineren Web Processing Service (WPS) nicht möglich, wie bereits in einem früheren Artikel thematisiert (Stollberg et al, 2007).

Abschließend sei nochmals betont, dass die WCPS-Sprache nicht als Endbenutzerschnittstelle angelegt ist. Vielmehr ist sie als eine programmiertechnisch einfach handzuhabende Schnittstelle gedacht, mit der graphisch-interaktive Clients komplexe Anfragen in XML codieren und an den Server übermitteln können. Es obliegt den Entwicklern solcher Clients, die Funktionalität (oder eine problemspezifische Teilmenge davon) dem Anwender in komfortabler Weise anzubieten, ohne letzteren der Syntax auszusetzen.

Referenzimplementierung

Die Entwicklung des WCPS basiert auf der Erfahrung mit Raster-Anfragesprachen aus der Entwicklung und dem Einsatz von rasdaman.

Das rasdaman-System (raster data manager) ist Ergebnis einer Serie von EU-Projekten. Diese Middleware-Technologie erlaubt Manipulation und Abfrage beliebig großer multi-dimensionaler Rasterobjekte, welche vom System in relationalen Datenbanken verwaltet werden (Baumann, 1994). Die rasdaman-Anfragesprache rasql erweitert SQL92 um Rasteroperationen mit formal definierter Semantik (Baumann, 1997). Der Server

optimiert Anfragen und wertet sie gegen die gespeicherten multidimensionalen Rasterobjekte aus, welche partitioniert in einer relationalen Datenbank in Blobs (Binary Large Objects) abgelegt sind.

Untersuchungen an der Jacobs University ergaben, dass eine Kachelgröße von einigen Megabyte besonders geeignet ist, wenn für den Server eine übliche PC-basierte Hardware-Architektur zugrunde gelegt wird (Rathnam, 2006). Statt eines Terabyte-großen Blobs verwaltet die Datenbank also typischerweise Millionen von relativ kleinen „Blöcken“ im Megabyte-Bereich. Derzeit finden vergleichende Blob-Benchmarks auf den marktführenden und Open-Source-Datenbanksystemen statt.

Weitere Merkmale von rasdaman sind flexible multidimensionale Kachelung, transparente Tertiärspeicheranbindung, Hardware/Software-Parallelisierung. Rasdaman befindet sich im kommerziellen Einsatz, u.a. beim französischen Nationalen Geographischen Institut in Paris.

Für die Referenzimplementierung wurde rasdaman benutzt, weil die meisten WCPS-Konzepte bereits implementiert zur Verfügung stehen, wenn auch als Datenbank-Anfragesprache anstatt eines Web Services. Als relationales DBMS wird in der Installation von Jacobs University PostgreSQL verwendet – unter anderem um zu zeigen, dass Open-Source-Datenbanksysteme eine geeignete Basis für solche Dienste bilden. Prinzipiell könnte aber jedes relationale DBMS eingesetzt werden. Abb. 3 zeigt den resultierenden Service-Stack.

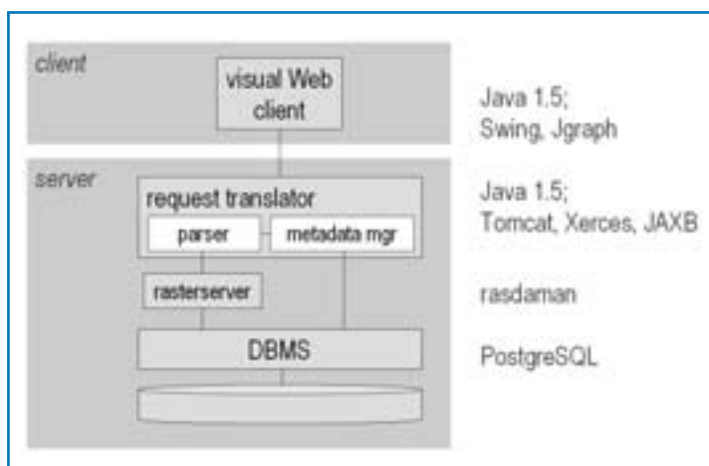


Abb. 3: Architektur der WCPS-Referenzimplementierung mit verwendeten Komponenten.

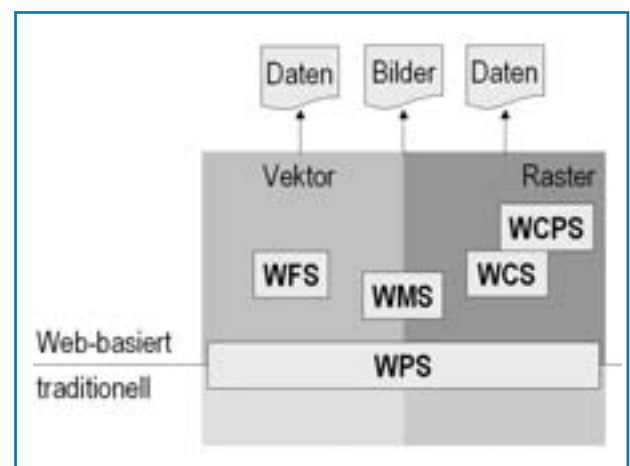


Abb. 4: Einordnung raster-relevanter OGC-Standards.

```

<ProcessDescription processVersion="2" storeSupported="true" statusSupported="false">
  <ows:Identifier>Buffer</ows:Identifier>
  <ows:Title>Create a buffer around a polygon</ows:Title>
  <ows:Abstract>Create a buffer around a single polygon. Accepts the polygon as GML and provides GML output for the buffered feature.</ows:Abstract>
  <ows:Metadata xlink:title="spatial" />
  <ows:Metadata xlink:title="geometry" />
  <ows:Metadata xlink:title="buffer" />
  <ows:Metadata xlink:title="GML" />
  <DataInputs>
    <ows:DataInput>

```

Abb. 5: Die Semantik einer WPS-Anfrage ist formuliert in Title und Abstract und nur für Menschen erkennbar.

Weitere relevante Standards

Die Familie modularer OGC-Standards (i.w.: Service Implementation Specifications und Profiles) basiert auf dem gemeinsamen Architekturmodell von OWS Common. Abb. 4 zeigt eine mögliche Einordnung von OGC-Spezifikationen aus Raster-Sicht.

Der Web Map Service (WMS) Vektor- und Rasterdaten kann ein WMS-Dienst überlagern und erzeugt letztendlich ein einziges Bild, welches z.B. in einem Web-Browser darstellbar ist. Diese konzeptuelle Einfachheit und unmittelbare Web-Tauglichkeit machen den WMS zum unangefochtenen (kostenfreien) „Verkaufsschlager“ der OGC. Aus Raster-sicht ist die Funktionalität zwischen WCS und WMS nicht unähnlich, allerdings haben beide unterschiedliche Zielsetzung: während WMS den Schwerpunkt auf unmittelbar darstellbare Karten-Bilder legt, soll WCS es insbesondere ermöglichen, Daten in ihrer Ursprungsform abzurufen und ggf. clientseitig weiter zu verarbeiten. Beide Standards haben demgemäß ihre speziellen Einsatzschwerpunkte, werden aber auch oft zusammen (und mit einem WFS) eingesetzt.

Der Web Processing Service (WPS) erscheint besonders geeignet, um Legacy-Anwendungen zu „webifizieren“ (OGC, 2006c). Darüberhinaus ist das zugrundeliegende SOAP-Prinzip, Routinen auf entfernten Rechnern aufzurufen, sehr allgemein und für praktisch sämtliche Gebiete einsetzbar (SOAP = Simple Object Access Protocol). Nachteilig ist allerdings, dass lediglich die Funktionsnamen und Parametertypen (also die Syntax des Aufrufs) spezifiziert werden. Was eine WPS-Funktion wirklich bewirkt, steht lediglich in Titel und Kurzfassung – also nicht maschinenlesbar (Abb. 5). Damit lässt sich beispielsweise keine automatische Verteilung einer Anfrage auf mehrere Rechnerknoten („service chain-

ning“) bewirken. WCPS auf der anderen Seite bietet durch seine formal spezifizierte Anfragesprache die Möglichkeit, per Programm die Anfrage zu analysieren und ohne Wissen und Zutun des Clients zu verteilen. Nebenbei bemerkt lässt sich WCPS in natürlicher Weise als WPS-Dienst formulieren. Insofern ergänzen sich also WPS und WCPS.

Ein wichtiger Punkt ist die volle Kompatibilität zwischen WCS und WCPS. Ein Software-Anbieter kann ein (einfacher zu implementierendes) WCS-Produkt anbieten und vielleicht später einen (anspruchsvolleren) WCPS zusätzlich entwickeln, und beide Dienste auf demselben Datenvorrat parallel und kompatibel betreiben.

In jüngerer Zeit entsteht, getragen von großem weltweitem Interesse, das Sensor Web Enablement (SWE, vgl. www.opengeospatial.org/projects/groups/sensorweb). SWE umfasst die Teilstandards Observations and Measurements (O&M), SensorML, TransducerML, Sensor Observation Service (SOS), Sensor Alert Service (SAS), Sensor Processing Service (SPS) und Web Notification Services (WNS). Der Begriff „Sensor“ wird dabei sehr allgemein gesehen und kann u.a. bildgebende Instrumente an Bord von Satelliten beinhalten. Es wird daher wichtig sein, diese neuen Standards mit den Basisstandards wie WFS und WCS zu harmonisieren.

Fazit und Ausblick

Rasterdienste bilden einen immer wichtigere Facette im Spektrum der Geodienste. So bezeichnet Xavier Lopez, Director Spatial bei Oracle, „2-D and 3-D image services the next great wave in geo databases“. Dienste wie WorldWind und GoogleEarth bilden erst den Anfang einer umfassenden vertikalen Integration von Rasterdaten in das Geo-Dienstespektrum, bei gleichzeitiger horizon-

taler Ausdehnung in die verschiedensten Anwendungen. Standards werden daher immer wichtiger, um Interoperabilität der entstehenden Dienste zu sichern. Dem OGC, in Zusammenarbeit mit ISO, OASIS, W3C und weiteren relevanten Organisationen, kommt dabei eine wichtige integrierende Rolle zu.

Ange-sichts der Komplexität der Materie ist der Ansatz, modulare Teilstandards auf Basis einer gemeinsamen konzeptuellen Architektur vorzusehen, jedenfalls zielführend. Dass sich in der Praxis manche der OGC-Standards über-lappen, ist nur natürlich: ähnliche Fragestellungen treten in unterschiedlichsten Ausprägungen und Anwendergruppen auf – Rasterobjekte beispielsweise in Kartographie, Fernerkundung, Sensorik und OLAP (Online Analytical Processing) bei der Analyse von Geschäftsdaten. Die Harmonisierung von Standards bleibt daher eine der großen Herausforderung innerhalb des OGC.

WCS als Basisdienst für Rasterzugriff wurde in der aktuellen Version 1.1 nach erfolgreicher elektronischer Abstimmung Ende Dezember 2006 freigegeben. Die WCS-Arbeitsgruppe hat für Version 1.2 bereits einige Punkte auf der Agenda, unter anderem:

- „transactional service“ oder „harvesting service“: ähnlich zum WFS-T soll ein WCS-T spezifizieren, wie Daten von einem anderen Rechner abgeholt und ins Dienste-Portfolio eingebaut werden können.
- „abstrakte“ Dimensionen und beliebige Schnitte: Rasterobjekte sollen neben raum-zeitlichen Achsen auch „abstrakte“ Achsen besitzen können, etwa Verkaufszahlen oder Spektralbereiche. Schnitte sollen nicht notwendig x und y gleichzeitig besitzen, beispielsweise sollen x/t –Schnitte zulässig sein. Diese Vorschläge resultieren aus den WCPS-Arbeiten.
- Ein bedeutendes Thema wird sicherlich die OGC-interne Abstimmung der verwandten Standards – insbesondere SWE ist hier zu nennen – mit WCS bilden.

WCPS hat seit 2006 den Status eines „Best Practice“ Paper (OGC, 2006a), eine WCPS-Arbeitsgruppe wurde im April 2007 beschlossen; die Verabschiedung als Standard wird für 2007 bzw. 2008 angestrebt. ▶

Wichtig erscheint nun eine umfassende, möglichst breit angelegte Evaluierung der WCS- und WPCS-Spezifikationen durch Implementierungen und Anwendungsstudien. Für WCS laufen bereits seit einiger Zeit parallel zu den Implementierungen in Industrie und Wissenschaft auch praktische Erprobungen; zwei Projekte seien namentlich erwähnt:

- Im OGC-Netzwerk GALEON (Geo-interface to Atmosphere, Land, Earth, Ocean netCDF) wird WCS auf seine Tauglichkeit für 4D/5D-Klimadaten untersucht wird. WPCS wird in der derzeit anlaufenden 2. Phase eingebracht und evaluiert.
- EarthLook ist ein vom Land Bremen gefördertes Vorhaben, um an Hand eines Demonstrators auf maritimen, geophysikalischen und Fernerkundungsdaten die Brauchbarkeit von WCS und WPCS zu evaluieren. Neben den Projektpartnern OHB System AG und Jacobs University ist eine Reihe von Institutionen assoziiert,

insbesondere das Alfred-Wege-ner-Institut Bremerhaven.

Sowohl für WCS als auch für WPCS sind weitere praktische Evaluierungen von hohem Interesse. Wissenschaftliche Kooperationen hierzu sind daher hochwillkommen, um domänenübergreifend nutzbare, interoperable und effizient implementierbare Raster-Standards zu erreichen.

Danksagung

Teile der vorgestellten Arbeiten sind im Projekt EarthLook entstanden, welches von der Bremer Innovations-Agentur (BIA) gefördert wird. Den Gutachtern dankt der Autor für ihre konstruktiven Verbesserungsvorschläge. ■

AUTOR

Peter Baumann
Professor of Computer Science
Jacobs University Bremen
Campus Ring 12, 28759 Bremen
p.baumann@jacobs-university.de

Literatur:

Baumann, P. (1994): On the Management of Multidimensional Discrete Data. VLDB Journal 4(3)1994, Special Issue on Spatial Database Systems, pp. 401-444

Baumann, P. (1997): A Database Array Algebra for Spatio-Temporal Data and Beyond. The Fourth International Workshop on Next Generation Information Technologies and Systems (NGITS '99), July 5-7, 1999, Zikhron Yaakov, Israel, Lecture Notes on Computer Science 1649, Springer Verlag, pp. 76 - 93

Kleese, K. et al (2000): Intelligent Support for High I/O Requirements of Leading Edge Scientific Codes on High-End Computing Systems - The ESTEDI Project. Sixth European SGI/Cray MPP Workshop, 7-8 September 2000, Manchester, UK

ISO (2003): Geographic information – Schema for coverage geometry and functions. ISO DIS 19123, TC 211/SC/WG 2

OGC (2000): The OpenGIS Abstract Specification: Topic 6: The Coverage Type and its

Subtypes. OGC document 00-106

OGC (2006a): Web Coverage Processing Service Implementation Specification. OGC document 06-035r1, 2006

OGC (2006b): Web Coverage Service Implementation Specification. OGC document 06-083r8, 2006

OGC (2006c): Web Processing Service Implementation Specification. OGC document WPS 05-007r4

Rathnam, R. (2006): Partitioning as a Raster Database Tuning Parameter. Jacobs University Bremen, Web Information Systems seminar report

Stollberg, B. et al (2007): Geoprozessierung in Geodateninfrastrukturen - Aufgaben für die nächste Generation. GIS 4: 24-29

W3C (2007): XQuery 1.0: An XML Query Language. W3C Recommendation, 23 January 2007, www.w3.org/TR/xquery/



Der
**Landkreis
 Cuxhaven**

sucht für den
GIS-Service zum
 15.09.2007
 eine/einen

Diplom-Ingenieur/in (FH)

mit dem Schwerpunkt Geographische Informationssysteme

Es handelt sich um eine Teilzeitstelle (19,5 Wochenstunden), die zunächst befristet bis zum 31.03.2008 zu besetzen ist.

Aufgaben:

- Erstellung von Kartendiensten für Intra-/Internet,
- Geodatenmanagement,
- Realisierung von GIS-Projekten sowie
- Unterstützung der Anwender der GIS-Arbeitsplätze.

Anforderungen:

- Dipl.-Ingenieur/in (FH) der Fachrichtungen Wasserwirtschaft, Landschaftsplanung, Regionalplanung, Geoinformatik oder vergleichbarer Studiengänge
- Fundierte Kenntnisse im Bereich Geographische Informationssysteme mit ArcView 9.x bzw. ArcViewGIS 3.x, ArcIMS 9.x und/oder ArcSDE 9.x
- Kenntnisse in den Bereichen Internet (HTML, XML, JavaScript), Datenbanken (MS SQL, MySQL o.ä.), MS-Office-Software, CAD (AutoCAD o.ä.) und Bildbearbeitung
- Einsatz- und Leistungsbereitschaft
- Organisationsfähigkeit, Innovationsbereitschaft und Sorgfalt
- Eigenverantwortliche und selbständige Wahrnehmung des Aufgabenbereichs im Rahmen von Zielvorgaben

Vergütung erfolgt nach Entgeltgruppe 10 TVöD.

Für weitere Auskünfte stehen Herr Fokuhl, GIS-Service, 04721/66 23 67 oder Herr Kleinschmidt, Personalservice, 04721/66 22 40 zur Verfügung.

Bewerbungen mit vollständigen Bewerbungsunterlagen richten Sie bis zum **20.08.2007** an den

Landkreis Cuxhaven,
Interner Servicebetrieb, 27470 Cuxhaven