



Ciencia e Ingeniería Neogranadina

ISSN: 0124-8170

revistaing@unimilitar.edu.co

Universidad Militar Nueva Granada

Colombia

Amézquita Becerra, Germán; Pérez Castillo, José Nelson
SERVICIO GRID PARA LA CLASIFICACIÓN NO SUPERVISADA DE IMÁGENES SATELITALES
UTILIZANDO AUTÓMATAS CELULARES
Ciencia e Ingeniería Neogranadina, vol. 19, núm. 2, diciembre, 2009, pp. 59-76
Universidad Militar Nueva Granada
Bogotá, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=91112316005>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

SERVICIO GRID PARA LA CLASIFICACIÓN NO SUPERVISADA DE IMÁGENES SATELITALES UTILIZANDO AUTÓMATAS CELULARES

GRID SERVICE FOR UNSUPERVISED CLASSIFICATION OF SATELLITE IMAGES USING CELLULAR AUTOMATA

Germán, Amézquita Becerra

*Ingeniero de Sistemas, Docente Universidad Pedagógica y Tecnológica de Colombia.
(Tunja, Boyacá Colombia). Grupo Internacional de Investigación en Informática, Comunicaciones y
Gestión del Conocimiento – GICOGE, Universidad Distrital “Francisco José de Caldas”,
Bogotá Colombia, german.amezquita@gmail.com.*

José Nelson, Pérez Castillo

*Doctor en Informática Universidad de Oviedo. Director Grupo Internacional de Investigación en
Informática, Comunicaciones y Gestión del Conocimiento - GICOGE, Investigador y Docente de
planta Universidad Distrital Francisco José de Caldas. Bogotá Colombia jnperezc@gmail.com.*

Fecha de recepción: 9 de septiembre de 2009

Fecha de aprobación: 20 de octubre de 2009

RESUMEN

Este artículo describe el proceso investigativo por el cual se exploró y puso en práctica la tecnología Grid como elemento promotor del proceso de clasificación no supervisada de imágenes satelitales LANDSAT; se ha implementado un servicio Grid que aplica autómatas celulares como técnica de inteligencia artificial. Las anteriores tareas se han logrado estableciendo una infraestructura Grid y un modelo de implementación que da soporte al servicio Grid. El modelo del servicio visualiza el contenedor de servicios y el contenedor de “portlet”, los cuales se integran para formar tanto el cliente como el servidor. El autómata celular utilizado, está definido en vecindarios de dos dimensiones y se establece la proyección de los niveles digitales de tres de las bandas de la imagen y la ubicación de cada una de ellas sobre el vecindario definido. El algoritmo se basa en reglas de transición que generan modificaciones a los vecindarios hasta obtener las categorías deseadas. Como resultado del proceso de clasificación se genera una nueva imagen en la cual se expresan las categorías en valores de 0 a 255 y se establece falso color para visualizar los resultados obtenidos.

Palabras clave: Servicios Grid, autómatas celulares, clasificación de imágenes satelitales.

ABSTRACT

This paper describes the research process to explore and implement a Grid technology as a promoter of unsupervised, LANDSAT satellite image classification process. In fact, a Grid service has been implemented using cellular automata as an artificial intelligence technique. Those tasks have been achieved by establishing a Grid infrastructure and an implementation model supporting the Grid service. The service model displays the service container and the portlet container, which are integrated to form both client and server. The cellular automaton used is defined by two-dimension neighborhoods, and also a digital level projection of each three bands of the image and location on a defined neighborhood is established. The algorithm is based on transition rules generating changes to neighborhoods to get the desired categories. As a result of the classification process a new image is generated in which categories are expressed in values from 0 to 255 but providing false color to display the results.

Keywords: Grid Services, cellular automata, classification of satellite images.

INTRODUCCIÓN

La computación Grid, corresponde a una de las tendencias para el desarrollo de nuevos sistemas de información, con objetivos definidos en beneficio de los usuarios, por lo tanto se plantea la integración de los sistemas actuales, para lograr compartir los recursos tecnológicos y humanos[1], de tal manera que las entidades participantes aumenten sus capacidades de procesamiento aprovechando los recursos de otras instituciones.

Las empresas durante muchos años han adquirido tecnología que cumple con sus necesidades, sin embargo, los requerimientos tanto de almacenamiento como de procesamiento día a día crecen en las empresas, lo que implica la adquisición de nuevos equipos que den soporte a dichos requerimientos, los cuales muchas veces exceden las capacidades económicas de la empresa [2]; esta razón le ha dado fuerza a proyectos de investigación que se orienten hacia la búsqueda de recursos compartidos con otras empresas y brindar unas mejores prestaciones que suplan las necesidades de recursos, tanto de hardware y software como humano, siendo justamente éste aspecto el que denota la amplia relevancia de las tecnologías Grid.

Las necesidades de conocimiento de las entidades y centros de investigación estiman cada día mayores capacidades de almacenamiento, procesamiento y acceso a dispositivos inalcanzables, tanto por su tecnología como por sus costos, siendo esta una nueva razón que promueve el uso de la Grid.

La aplicabilidad de la computación Grid, puede ser en cualquier área del conocimiento: negocios, gobierno, investigación, entretenimiento, entre otras[3]. Una de las áreas

dónde esta tecnología encuentra mayor aplicabilidad es el procesamiento de imágenes, dado que requiere de altos niveles de prestaciones en cuanto a las capacidades de procesamiento, lo cual hace de la computación Grid, la plataforma ideal para el desarrollo de este trabajo.

Este proyecto de investigación esta dirigido hacia la clasificación no supervisada de imágenes satelitales y utiliza como técnica de inteligencia artificial los autómatas celulares.

Un autómata celular se compone de una cuadrícula uniforme regular y un conjunto de reglas que permiten que este evolucione y tenga un comportamiento a través de periodos discretos para obtener el resultado del procesamiento; justamente el proyecto relacionado por este artículo emplea tales autómatas celulares para clasificar de manera no supervisada imágenes geográficas obtenidas a partir del satélite LANDSAT.

En este documento se define el modelo sobre el cual se implementa el servicio Grid y a partir del mismo se describen los elementos y se establece una arquitectura orientada a la computación Grid, que da el soporte técnico de hardware y software para garantizar el entorno adecuado de lo propuesto preliminarmente.

1. PRELIMINARES

Esta sección describe los conceptos que apoyaron el desarrollo del proyecto de investigación, como la tecnología Grid implementada para el desarrollo del proyecto, el área específica del conocimiento a procesar (imágenes satelitales) y la técnica de programación utilizada para procesar las imágenes satelitales.

1.1. GRID

La tecnología GRID en los últimos tiempos se ha convertido en un tema de interés tanto de las comunidades académicas, como de las instituciones gubernamentales, por ello el Gobierno Colombiano a través la Red Nacional Académica de Tecnología Avanzada RENATA [4], ha establecido los mecanismos para que las universidades y los centros de investigación implementen tecnologías que permita un mejor aprovechamiento de sus recursos tecnológicos y humanos en pro de la investigación.

La tecnología GRID ha evolucionado y su definición por consiguiente ha sufrido los cambios respectivos, en la actualidad se encuentran varias definiciones, como las referenciadas por el autor Frédéric Magoulès, en su libro "Grid Resource Management" [5], quien propone una definición ajustada a los cambios tecnológicos, así: "Grid es una infraestructura de hardware y software que provee acceso transparente, fiable, generalizado y coherente, a recursos distribuidos a gran escala, propietarios y compartidos

por múltiples organizaciones administrativas, con el fin de prestar soporte o apoyo a una gran gama de aplicaciones con el deseo de calidad de servicio.”

La evolución de la Grid se ha establecido en tres generaciones, en donde la primera, se enfatizó en interconectar las grandes supercomputadoras para proveer acceso a los recursos; bajo estos lineamientos se presentaron dos proyectos representativos, uno orientado a dar soporte a aplicaciones específicas y el otro orientado a la ejecución de diferentes aplicaciones de alto desempeño [5]. La segunda generación se enfocó a desarrollar el centro de la tecnología Grid, y se establecieron los “middleware”, que tiene la tarea de hacer transparente el acceso a los recursos, permitiendo realizar entre otras las siguientes tareas: administración de recursos grid, programación de recursos y despliegue de un portal GRID [5]. La tercera generación, representa la convergencia entre la computación GRID y la tecnología de servicios WEB y tiene como idea principal facilitar el ensamblaje de recursos integrados en la grid [5].

1.2. COMPUTACIÓN GRID

Las empresas e instituciones de investigación han comprendido que una de las formas de buscar la solución a los requerimientos de procesamiento y extracción de conocimiento de la información existente, es compartiendo recursos Inter-institucionales, de tal manera que las instituciones se benefician mutuamente [6]. Las empresas cuentan con infraestructuras en las que han integrado servidores, computadores de escritorio, portátiles, dispositivos móviles y dispositivos especiales; dependiendo de las necesidades cada empresa establece los elementos que puede compartir y a los que necesitarían tener acceso.

Teniendo en cuenta que muchas de las necesidades de los centros de investigación implica el uso de la computación y el tratamiento intensivo de datos, los adelantos en la tecnología han establecido varios tipos de computación Grid así: Computación distribuida, computación de alto rendimiento, computación por demanda, computación intensiva de datos y computación colaborativa [5]. Cada uno de estos tipos de computación Grid, apunta a cubrir las estrategias específicas utilizadas por los centros de investigación para sus actividades.

1.3. SERVICIOS GRID

La Infraestructura de Servicios de Grid Abierta (*Open Grid Services Infrastructure OGSi*), define los servicios Grid como la conformación de un servicio WEB más el estado de los recursos, incluyendo los mecanismos para la creación, manejo e intercambio de información entre entidades, facilitando la notificación asincrónica de cambios del estado de los servicios, administración de colecciones de servicios, administración de los ciclos de vida de un servicio y manejo de notificación de fallas [5]. La especificación de la OGSi

establece las interfaces y el comportamiento de los servicios grid a través del lenguaje de descripción de servicios Web (*Web Services Description Language, WSDL*) y del esquema de Lenguaje de Marcado Extensible (*eXtensible Markup Language, XML*).

Los servicios Web se han construido para fines específicos y sobre componentes de tecnologías propietarias, con estrictas normas para permitir la comunicación e interacción entre aplicaciones; mientras que los servicios Grid intentan establecer una estructura que sea heterogénea, de tal manera que en la actualidad se está presentando una convergencia entre las dos tecnologías [5] (Figura 1), es así como Globus Alliance ofrece mecanismos para desarrollar Grids a través de servicios Web.

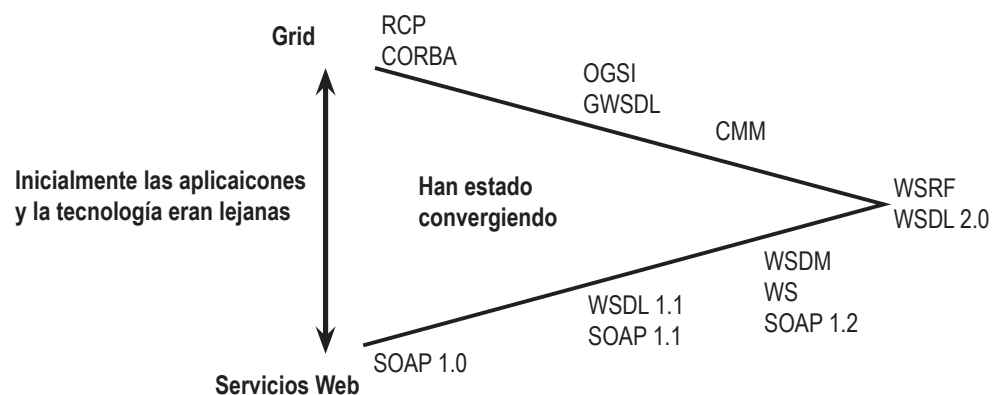


Figura 1. Convergencia entre servicios Web y servicios Grid

Fuente: Frédéric Magoulès [5].

1.4. IMÁGENES SATELITALES

Una imagen satelital es una representación visual de la información capturada por un sensor montado en un satélite artificial. La imagen satelital es la representación de la realidad en forma de una matriz de datos, donde cada celda o píxel es una porción del área observada con un único valor [7, 8].

Las imágenes satelitales almacenan por lo general las variaciones de la intensidad de energía reflejada por la superficie y registrada por el sensor. Estos valores se conocen como **Niveles Digitales (ND)** para cada píxel y normalmente oscilan entre 0 y 255.

Las imágenes satelitales ofrecen información valiosa para innumerables aplicaciones: planificación urbana, vigilancia del medio ambiente, gestión de cultivos, exploración minera, desarrollo de mercados, localización de bienes raíces, entre otros. En forma adicional a la información que se puede observar a través del ojo humano, las imágenes contienen detalles “ocultos” es así que en la información espacial registrada por

las imágenes, se puede observar el contenido mineral de las rocas, la humedad del suelo, la salud de la vegetación, la composición física de los edificios y otros detalles imperceptibles por el ojo humano, esta información se logra por la reflectancia o el denominado contenido espectral de la imagen satelital [9].

Uno de los métodos de mayor relevancia para el procesamiento de imágenes satelitales es la clasificación no supervisada, la que hace parte del núcleo investigativo del presente proyecto y en la que se establecen algoritmos para el análisis sin tener conocimiento previo del contenido de la imagen, por lo tanto sólo requieren de la imagen a clasificar y algunos parámetros que limiten el número de clases que se generarán. Estos algoritmos basan su efecto en la búsqueda de clases con suficiente separabilidad espectral para diferenciar los elementos [10, 11]. Para la clasificación existen varios métodos entre los que se destacan: k-Medias e ISODATA.

El método de clasificación k-Medias, calcula inicialmente el término medio de las clases extraídas de la imagen, luego de forma iterativa los píxeles son insertados en las clases más cercanas utilizando la técnica de la mínima distancia. En cada iteración se calcula la media de la clase y se vuelven a clasificar todos los píxeles hasta alcanzar las clases deseadas [10, 12].

El método de clasificación ISODATA, requiere conocer la cantidad de agrupaciones que se desean utilizar, luego se localiza la zona de mayor reflectancia, los píxeles se asignan al cluster utilizando la técnica de mínima distancia, los clusters se asocian, disgregan o borran en función de la máxima distancia de clases o del número mínimo de píxeles ocupados por la clase, por último se repiten las iteraciones hasta que el cambio entre ellas sea mínimo [10, 12].

1.5. AUTÓMATAS CELULARES

Los autómatas celulares fueron diseñados por Stanislaw Ulam (1909-1984) y John von Neumann (1903-1957) para proporcionar un marco formal a la investigación del comportamiento de sistemas complejos [13]; y se basan en estructuras estáticas, las cuales se rigen por conjuntos de reglas dentro del marco del sistema que se esté analizando. De acuerdo la literatura existente se encuentran diferentes definiciones, así:

Definición descrita por Wolfram [13, 14], quien precisa los autómatas celulares como: “Idealizaciones matemáticas de sistemas físicos en los cuales el paso del tiempo es concreto y la cantidad física toma un conjunto de valores discretos. Un autómata celular toma un conjunto finito de valores discretos y consta de una cuadrícula uniforme regular o “arreglo”, usualmente infinita en extensión, con una variable discreta en cada sitio o “celda”. El estado de un autómata celular es completamente especificado por los valores de las variables de cada sitio. Un autómata celular evoluciona en pasos de tiempo

discretos, el valor de la variable o celda es afectado por los valores de variables o celdas vecinas. El vecindario es típicamente considerado el sitio en si mismo y todos los sitios inmediatamente adyacentes. Las variables de cada sitio son actualizadas simultáneamente (“sincrónicamente”), con base en los valores de las variables en su vecindario en los pasos precedentes, y acordes a un definitivo conjunto de “reglas locales”.

La segunda definición seleccionada representa la visión matemática [15]:

$$A = (S_d, Q, N, \delta),$$

Donde:

A = Representa un autómata celular

S_d = Representa el espacio de las celdas, y d la dimensión

Q = Corresponde al conjunto de estados

N = Representa el vecindario celular

δ = Corresponde a la regla de estado de transición basada en el vecindario

Comparando las definiciones y los objetivos del proyecto, se puede observar que los autómatas celulares son perfectamente aplicables al procesamiento de imágenes, por su conformación en estructuras estáticas y su contenido puede ser procesado aplicando reglas para obtener los resultados deseados.

Las características básicas de los autómatas celulares se centran en el establecimiento de una geometría que corresponde a una división o teselación del espacio del dominio en celdas de la misma figura (cuadrados, polígonos, triángulos), dispuestos en una o varias dimensiones, de tal manera que cada una de las geometrías al ser integradas dentro de un procesamiento, permiten el análisis y simulación del comportamiento de los estados del dominio [14, 16].

Con base en la geometría seleccionada se definen vecindarios (Figuras 2 y 3) y sobre estos se establecen las reglas de comportamiento, las cuales generan las transiciones de cada uno de los elementos pertenecientes al dominio.

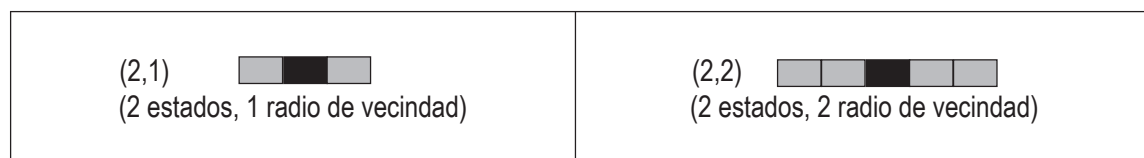


Figura 2. Vecindarios para una dimensión

Fuente: Martínez, G.J . [17]

En la geometría de dos dimensiones, se han establecido cuatro tipos de vecindades (Figura 3).

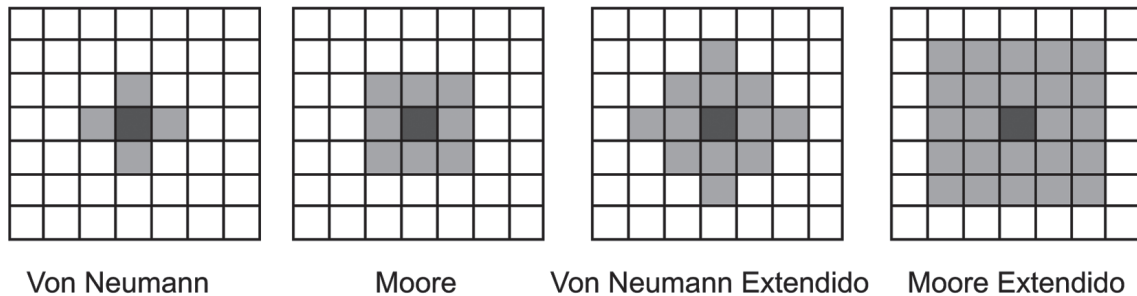


Figura 3. Vecindarios para dos dimensiones

Fuente: Martínez, G.J. [17]

2. IMPLEMENTACIÓN

Esta sección describe los detalles técnicos que se tuvieron en cuenta para implementar la infraestructura Grid, el servicio Grid y el algoritmo utilizado para lograr la clasificación no supervisada de imágenes.

2.1. IMPLEMENTACIÓN DE LA INFRAESTRUCTURA GRID

La infraestructura Grid dispuesta para el proyecto de investigación se compone de elementos que garantizan un laboratorio completo para la conceptualización, aprendizaje y experimentación; en la figura 4 se observan los elementos que conforman la infraestructura Grid utilizada:

- Servidor GT4, posee el contenedor de servicios Grid y está implementado con Globus Toolkit 4.0.1 (Globus es un programa de investigación sobre tecnologías GRID, producido por el grupo Globus Alliance, el cual esta respaldado por varios centros de investigaciones de universidades y compañías del sector privado (IBM, HP, entre otras) [18-20]; servidor sobre el cual se instala el servicio Grid para la clasificación no supervisada de imágenes satelitales.
- Autoridad Certificadora: Genera los certificados para poder establecer los niveles de confianza de los servidores y los usuarios a través de las credenciales emitidas.
- Servidor Myproxy: corresponde al repositorio de credenciales que garantizan niveles de confianza entre entidades, lo que operativamente se dispone para que múltiples servidores puedan recuperarlas con prontitud.

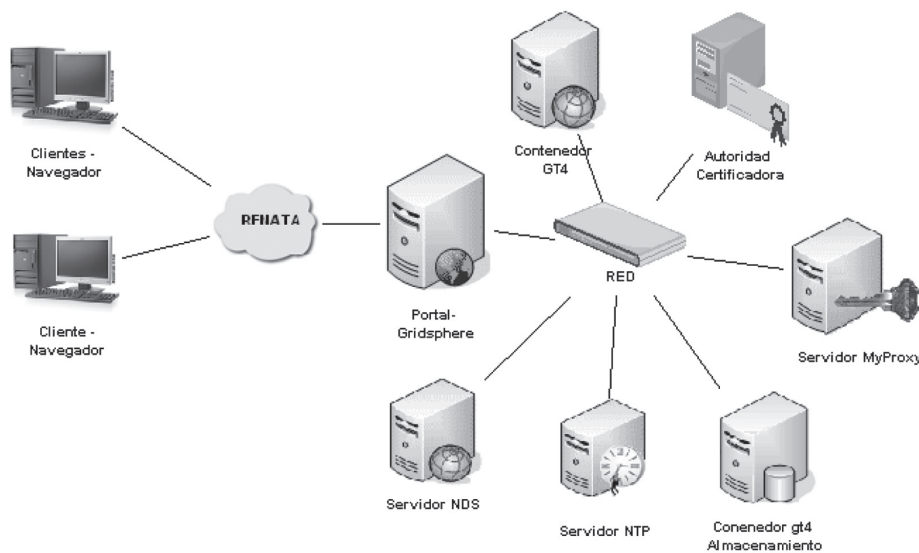


Figura 4. Infraestructura Grid (Universidad Distrital Francisco José de caldas)

- Servidor de almacenamiento, sobre el cual se almacenan todas las imágenes que se desean procesar.
- Servidor de Tiempo: se establece con el servicio de protocolo de tiempo de red (Network Time Protocol, NTP) de Linux que facilitará la sincronización de los servidores integrantes de la Grid.
- Portal: el cual fue implementado con Gridsphere 2.2.10 (desarrollado bajo la licencia open-source) y sobre el cual se implementa gridportlet, sobre este último se implementan los portlets que invocan el servicio Grid.

El proceso de integración de la infraestructura establecida, se logra con el trabajo colaborativo de estudiantes de maestría en Ciencias de la Información y las Comunicaciones pertenecientes al grupo GICOG y algunas asesorías externas a la Universidad Distrital Francisco José de Caldas. Entre los resultados obtenidos, se pueden resaltar:

- Generación de documentación semánticamente validada y aceptada por los integrantes del grupo, lo cual permitió que cada uno de ellos aportara en la construcción y el posterior aprovechamiento de la información para la replicación de la infraestructura en sus lugares de trabajo.
- La infraestructura establecida representa el resultado del trabajo de expertos con conocimiento en informática, sin embargo se hace evidente que después de terminada esta fase, se deben involucrar otros grupos de investigación de todas las áreas del conocimiento, para que dirijan sus proyectos hacia el aprovechamiento intensivo de la infraestructura Grid implementada.

- La infraestructura representa la base mínima necesaria para establecer servicios Grid, pero ha quedado lista para iniciar su proceso de crecimiento e integración con otras instituciones.

2.2. IMPLEMENTACIÓN DEL SERVICIO GRID

Durante el desarrollo del proyecto de investigación se ha implementado la infraestructura Grid y los algoritmos que permiten establecer el servicio Grid para la clasificación no supervisada de imágenes satelitales; la figura 5 ilustra el modelo del servicio Grid desarrollado, en donde se integra el servicio y la interacción con el mismo.

Dentro del modelo se visualizan letras, las cuales van a servir de guía para la siguiente descripción:

(A): Desde el portlet se solicita la ubicación del servicio, el cual es devuelto en un punto de referencia extremo (EPR), que contiene la ubicación del servicio y las especificaciones técnicas, para poder invocarlo.

(B): Igualmente desde el portlet se selecciona la imagen que se va a procesar, la cual está ubicada en el recurso de almacenamiento. Es importante aclarar que la ubicación de las imágenes no está limitada al recurso de almacenamiento definido en la infraestructura Grid, sino a todos los recursos disponibles en la grid y que tengan el servicio de Gridftp debidamente configurado.

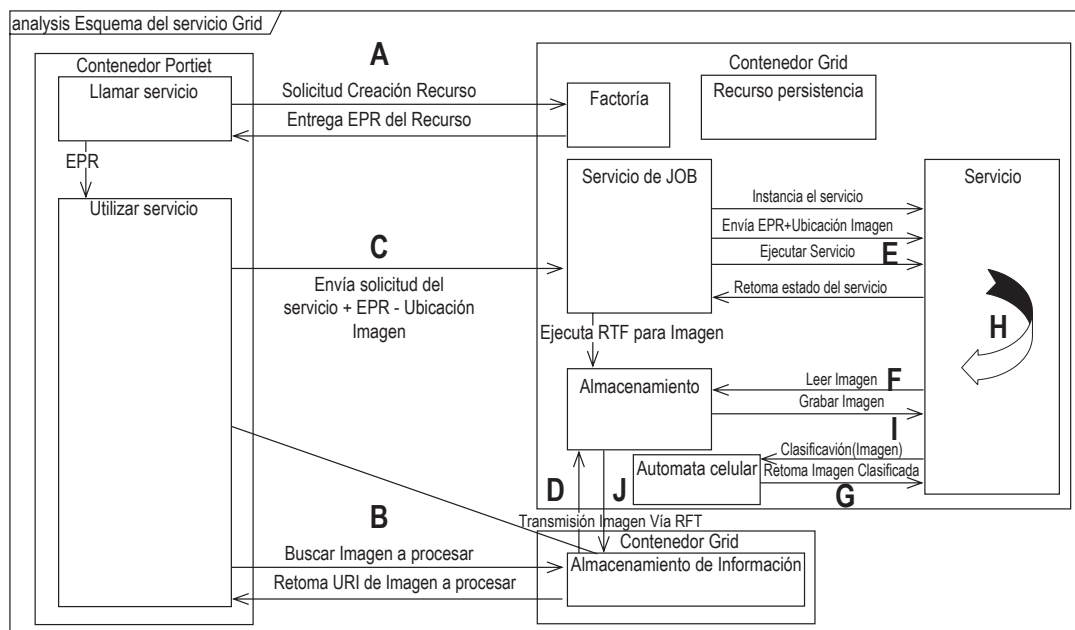


Figura 5. Modelo del servicio Grid

(C): Para la correcta invocación del servicio, se hace necesario definir algunos parámetros como los visualizados en la figura 6.

Figura 6. Vista de captura de parámetros

Se deben seleccionar tres bandas de la imagen, las cuales harán parte del procesamiento, y el número de categorías, el cual corresponde al número máximo de categorías que se generarán tras el procesamiento de la imagen y el número de iteraciones que corresponde a la cantidad de ciclos que se harán para lograr la clasificación. Es importante anotar que entre más iteraciones mejor será el resultado aunque más demorada la obtención del valor final, sin embargo, es de aclarar que el excesivo número de iteraciones puede llevar al caos el resultado final. Finalmente se envía el EPR, Ubicación de la imagen a procesar y los parámetros establecidos en la figura 6, al contenedor Grid como un “Job”, desde este instante se libera la conexión entre el contenedor de portlets y el contenedor Grid. El estado del procesamiento se puede monitorear a través del servicio de monitoreo que posee el contenedor de servicios y el contenedor de portlets.

(D): El servicio de “Jobs” como su primera tarea se encarga de transmitir vía Gridftp la imagen desde el lugar descrito en los parámetros a un lugar temporal dentro del contenedor de servicios Grid.

(E): El servicio de “Jobs” después de haber confirmado la terminación de la transmisión de la imagen, invoca el servicio Grid. Posteriormente se carga la imagen a memoria y se invoca el algoritmo del autómata celular.

(F): El cargue de la imagen se realiza desde la ubicación temporal extrayendo las tres bandas definidas en los parámetros y cargándolas a la memoria para su posterior procesamiento.

(G) Se procede a cargar el algoritmo del autómata celular a memoria, el cual se encarga de realizar el proceso de clasificación no supervisada de imágenes satelitales.

(H): Teniendo la imagen en memoria y los parámetros definidos por el usuario, como son la cantidad de categorías y las iteraciones, el servicio de clasificación ejecuta los algoritmos allí definidos tantas veces como se hayan definido. Al terminar este procesamiento se obtiene una nueva imagen.

(I): La imagen resultante del proceso de clasificación es almacenada en formato “tif” en la ubicación temporal del contenedor de servicios Grid.

(J) Por último el “Job” encargado de realizar la coordinación de las tareas y tras confirmar la terminación del servicio de clasificación de imágenes, procede a transmitir vía gridftp la imagen resultante al recurso de almacenamiento.

2.3. IMPLEMENTACIÓN DEL ALGORITMO

El autómata celular implementado aplica sus reglas sobre los puntos espaciales conformados por los valores digitales de las tres bandas que se hayan seleccionado para el procesamiento (figura 7).

El vecindario está definido según Moore (3x3) [17], y se establece la utilización de tres bandas, cada una de ellas representa uno de los ejes; como se puede apreciar en la figura 7, cada una de las proyecciones de los valores digitales de las tres bandas seleccionadas (banda2, banda3, banda4) representan un vecino.

Se establecen posibles centroides, tomados al dividir cada eje en tamaños de $(\text{valor-Maximo-valorMinimo}) / (N\text{categorias} * 2)$ y obteniendo como resultado cubos más pequeños, de los cuales se toma el centro y se definen como centroides temporales. Por cada una de las iteraciones del proceso de clasificación se redefinen los centróides, eliminando los que menos puntos tengan dentro del espacio de alcance de cada centroide. La unidad de medida para definir la cercanía entre puntos del espacio, está dada por la distancia euclidiana.

Las reglas de transición establecidas permiten que el centro del vecindario cambie su valor, tomando el valor dentro del mismo plano (x,y,z) acercando o igualando el centro al centroide identificado como el más cercano al vecindario.

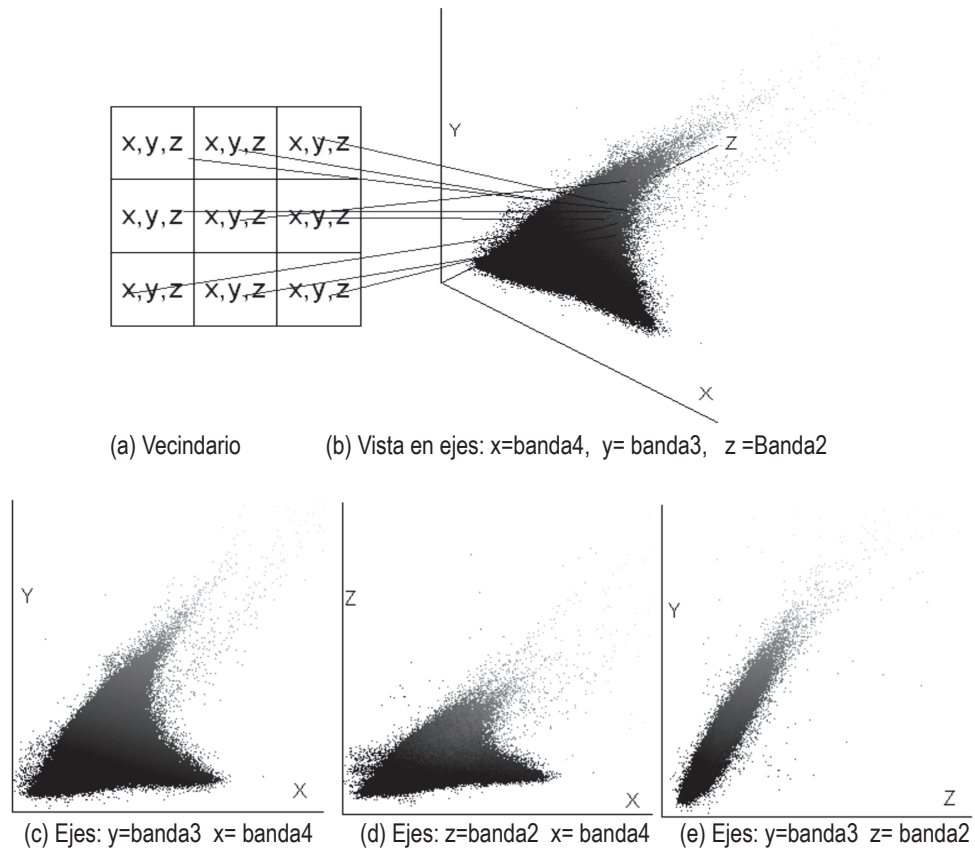


Figura 7. Proyección de valores de píxel de las tres bandas (4,3,2)

Todos los vecinos pertenecen al mismo centroide. (Figura 8)

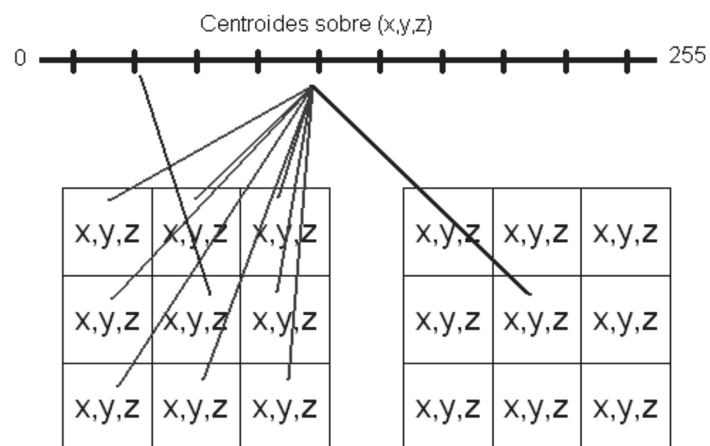


Figura 8. Centro igual al centroide de todos los vecinos.

Acercamiento al centroide más cercano a la mayoría de vecinos. (Figura 9)

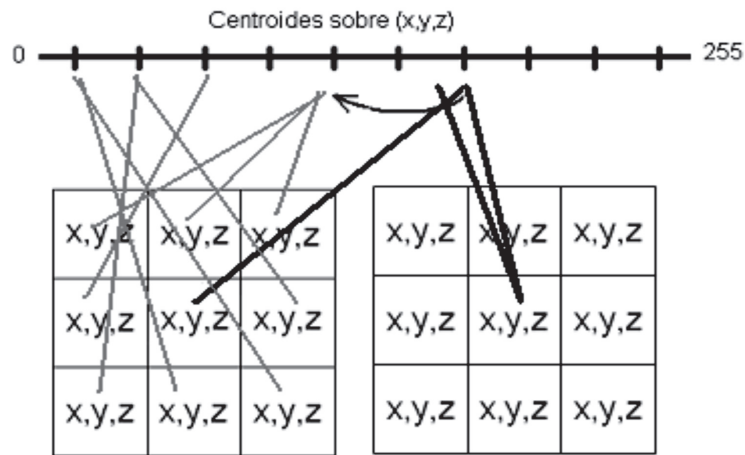


Figura 9. Acercamiento al centroide más representativo de todos los vecinos

3. CASO DE ESTUDIO

La puesta en funcionamiento de lo establecido por el proyecto involucró tres aspectos como fueron: implementación de la infraestructura, utilización de autómatas celulares como técnica de inteligencia artificial y la temática “Clasificación no supervisada de imágenes satelitales”. En donde se hizo evidente que cada una de ellas requería de las otras para lograr los resultados deseados.

3.1. INTEGRACIÓN DE LA INFRAESTRUCTURA Y EL PROCESAMIENTO DE CLASIFICACIÓN DE IMÁGENES

El proyecto razón de este documento, tiene como objetivos el establecimiento de un laboratorio experimental para el despliegue de servicios Grid y la definición e implementación de servicios Grids y en especial la Clasificación no supervisada de imágenes satelitales utilizando autómatas celulares.

Desde el punto de vista de computación de alto desempeño basado en una Grid, el prototipo implementado, representa un proceso de experimentación que da como resultado uno de los nodos que entrará a formar parte de la Grid Nacional e internacional a través de RENATA. Por lo tanto el aprovechamiento de los recursos informáticos inter-institucionales como lo plantea la computación Grid no se ha alcanzado.

Adicionalmente, el usuario de la tecnología Grid espera encontrar con la tecnología Grid tiempos de respuesta “excelentes” cuando se habla de poder de cómputo; este

proyecto no contó con esos tiempos de respuesta, por los pocos recursos disponibles; sin embargo la infraestructura y el servicio desarrollado se encuentra listo para hacer parte de la Grid Nacional e Internacional, y será en ese instante cuando se pueda realmente aprovechar las bondades de la tecnología Grid.

La utilización de la infraestructura Grid, representó la puesta en práctica de varios de los preceptos establecidos por la tecnología Grid: aprovechamiento de recursos, tales como el humano, aportado por lo integrantes del grupo de investigación GICOGE, quienes pertenecen a varias universidades del país, también se establecieron los niveles de confianza necesarios para garantizar la integración de los recursos tecnológicos y la correcta utilización de los mismos.

3.2. INFORMACIÓN UTILIZADA

La aplicación de servicio Grid se realizó sobre la imagen multiespectral LANDSAT de Bogotá en formato "img", con 6 bandas y corresponde a las coordenadas ULX: 983766.5051, ULY: 1019938.46240, LRX: 1008118.5092 y LRY: 992034.4623 (Figura 10).

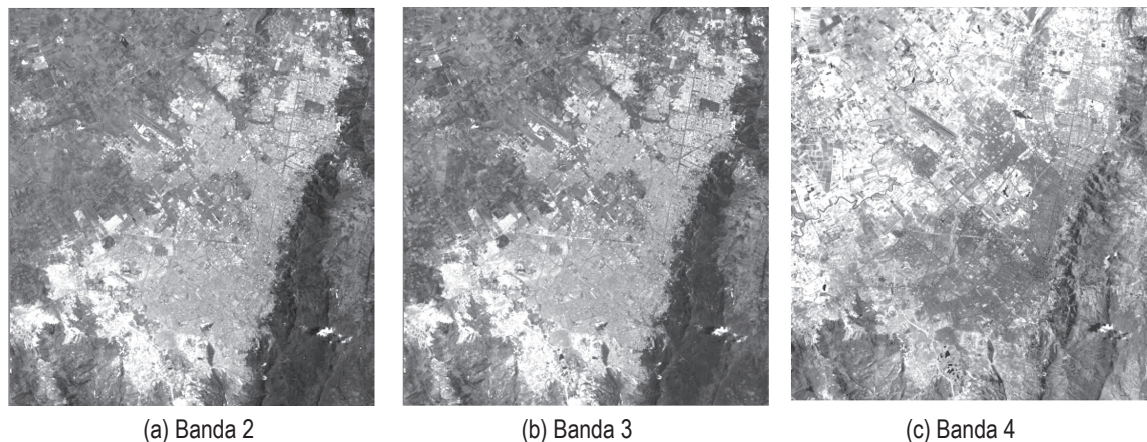


Figura 10. Imagen Landsat de Bogota.

Fuente: LANDSAT

La imagen tiene 762 * 873 píxeles y 8 bits por píxel, con una resolución del píxel de 32 metros, para la investigación se tomaron las bandas 2,3 y 4 y se estableció utilizar como máximo 10 iteraciones y 13 categorías.

3.3. CLASIFICACIÓN NO SUPERVISADA DE LA IMAGEN SATELITAL APLICANDO AUTÓMATAS CELULARES

La figura 11 contiene la imagen resultante de aplicar el algoritmo con autómatas celulares y se encuentra en formato "tif", formato que puede ser visualizado por una gran

cantidad de programas. Los píxeles de la imagen están dispuestos en valores de 0 a 255 en orden consecutivo y con una definición por cada valor de falso color para facilitar su visualización.

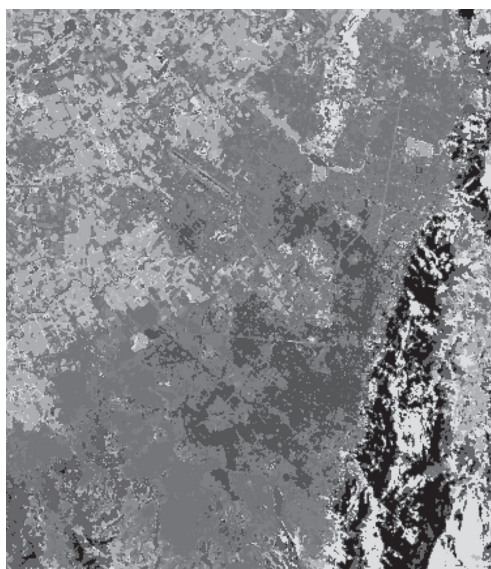


Figura 11. Imagen resultante del proceso de clasificación con 13 categorías y 10 iteraciones

Los colores asignados a la imagen (falso color), se seleccionan de los centroides resultantes del proceso de clasificación; como se aprecia en la figura 7, cada vecino posee los valores digitales de cada una de las bandas seleccionadas para la clasificación, y serán estos valores los que queden asignados a los centroides después del proceso de clasificación y por consiguiente representan el color en términos de RGB.

La utilización del servicio y en especial la asignación de los parámetros (bandas a utilizar, categorías e iteraciones) para la correcta ejecución, depende de las necesidades de los usuarios, y en ese instante serán ellos los que puedan definir si el resultado representa un gran aporte al trabajo que se encuentran desarrollando. Es importante resaltar que dependiendo las necesidades de los investigadores, tanto las técnicas de inteligencia artificial, como las imágenes seleccionadas para el proceso pueden no ser pertinentes o no ser de gran aporte a los intereses de los usuarios.

Teniendo en cuenta proyectos en etapa de desarrollo y pertenecientes al mismo grupo de investigación, se está trabajando en explorar otras técnicas de inteligencia artificial (rough set y redes neuronales) para la clasificación de imágenes satelitales, y según lo que se ha logrado observar y en conversaciones con los investigadores involucrados, se visualiza la necesidad de combinar varias técnicas de inteligencia para hacer una extracción de conocimiento más intensivo y por consiguiente de mayor intereses para los usuarios.

4. CONCLUSIONES Y TRABAJOS FUTUROS

Definitivamente la computación Grid abre nuevos caminos de investigación tanto desde el punto de vista de la integración de recursos (Hardware y Software), la cual es ideal para aquellos investigadores orientados al hardware, los sistemas operativos y las comunicaciones, como desde la aplicabilidad hacia las áreas del conocimiento, hechos que rompen todas aquellas barreras que limitan el alcance de muchos proyectos de investigación.

El desarrollo de este tipo de proyectos implica la conformación de grupos interdisciplinarios, ya que las áreas del conocimiento tanto tecnológicas como temáticas son diversas.

Se hace necesario el establecimiento de mecanismos de cooperación Inter.-institucional para lograr alcanzar mejores resultados, de tal manera que se integren recursos existentes en cada una de las instituciones y se pongan al servicio de las demás, logrando así establecer realmente una infraestructura de cómputo de altas prestaciones.

Aprovechando las capacidades alcanzadas tras la implementación de la arquitectura Grid, se hace evidente la oportunidad de extraer más conocimiento de la información almacenada en las imágenes satelitales, lo cual es factible integrando varias técnicas de inteligencia artificial.

El procesamiento de grandes volúmenes de información requiere de la integración de numerosos recursos de cómputo en clusters y técnicas de procesamiento para buscar mejores resultados, tanto en tiempo de respuesta como en calidad de la información.

REFERENCIAS

- [1] Tianfield, H., *Towards agent based grid resource management*, in *Cluster Computing and the Grid, 2005. CCGrid 2005. IEEE International Symposium on*. 2005.
- [2] Zhichun, X., W.T. Craig, and L. Wing Ning, *Data Processing Workflow Automation in Grid Architecture*, in *Proceedings of the Fifth International Conference on Grid and Cooperative Computing Workshops - Volume 00 %@ 0-7695-2695-0*. 2006, IEEE Computer Society. p. 189-195.
- [3] Fran Berman, G.F., Tony Hey, *The Grid: past, present, future*. Grid Computing – Making the Global Infrastructure a Reality., 2003.
- [4] Conectividad, A.d., RED NACIONAL ACADÉMICA DE TECNOLOGÍA AVANZADA, COLOMBIA.

- [5] Frédéric Magoulès, T.-M.-H.N., Lei Yu, Grid Resource Management - Toward Virtual and Services Compliant Grid Computing. Taylor & Francis Group, LLC, 2009.
- [6] Jacob, B., et al., Introduction to Grid Computing. 2005(IBM International Technical Support Organization).
- [7] ESRI-chile. LANDSAT 7 ETM+. [cited 2008 <http://www.esri-chile.com/biblioteca/landsat7etm.pdf>].
- [8] Fernández, I.A., El satélite Landsat- analisis visual de imagenes obtenidas del sensor ETM+ 2001.
- [9] SRGIS, J.M.L.-. Guía básica sobre Imágenes Satelitales y sus productos. 2006.
- [10] Universidad de Jaen, d.d.l.c., Geodésiza y Fotogrametría, prácticas de teledetección.
- [11] Consortium, O.G., Web Image Classification Service (WICS) Implementation Specification. 2005.
- [12] Visual_Information_Solutions_ITT, Tutotial: Classification Methods.
- [13] Wolfram, S., Cellular Automata and Complexity: Collected Papers.
- [14] Bas Straatman, A.H., Conrad Power, Guy Engelen, Roger White3, The Use of Cellular Automata for Spatial Modelling and Decision Support in Coastal Zones and Estuaria. Project group "Estuaria en Kusten", 2001.
- [15] Ling, C., et al., A Novel Ant Clustering Algorithm Based on Cellular Automata, in IAT '04: Proceedings of the Intelligent Agent Technology, IEEE/WIC/ACM International Conference on (IAT'04). 2004, IEEE Computer Society: Washington, DC, USA %L 1157678. p. 148-154.
- [16] Palash, S., A brief history of cellular automata. 2000, ACM Press. p. 80-107.
- [17] Martínez, G.J., Introducción a la simulación de procesos con autómata celular. Departamento de Posgrado, Escuela Superior de Cómputo Instituto Politécnico Nacional, 2006.
- [18] Sotomayor, B., Globus Toolkit 4 - programming Java Services. ELSEVIER 2006.
- [19] Monroy, J.A.G., Globus Toolkit. E.T.S.I.Telecomunicación- Departamento Ingeniería Sistemas Telemáticos.
- [20] Bart, J., et al., Enabling Applications for Grid Computing with Globus. 2003: IBM Redbooks.