

Gerenciamento de metadados de grandes volumes de dados de sensoriamento remoto

2017/24086-2 - Projeto de Pesquisa - Regular - FAPESP

Bolsa disponível

15 meses, para início em março de 2019, no INPE em São José dos Campos, SP. Bolsa TT-4 (<http://www.fapesp.br/3162>), para graduado, especialista em Tecnologia de Informação (TI), com dois anos de experiência após a graduação ou título de mestrado na área de TI, sem vínculo empregatício, com dedicação de 40 horas semanais às atividades de apoio ao projeto de pesquisa.

Resumo do projeto

Um dos serviços disponibilizados gratuitamente pelo INPE para a comunidade científica é o catálogo de imagens de sensoriamento remoto (também chamadas de imagens de satélite), disponível em <http://www.dgi.inpe.br/catalogo/>. Este catálogo contém atualmente um volume de aproximadamente 120 TB de dados (imagens em formato GeoTIFF), e está em constante crescimento, em virtude dos satélites capturarem novas imagens diariamente. Por exemplo, o satélite CBERS-2B, enquanto operacional, trabalhava com uma taxa de geração de dados de 120 megapixels por minuto. Selecionando a região de São Paulo, capital, uma consulta no catálogo de imagens do INPE apresenta um total de 640 imagens da mesma região, obtidas desde meados de 1980. Considerando que cada imagem possui aproximadamente 7000 linhas por 7000 colunas, e ocupa em disco (em média) um total de 50 MB, isto representa um total de 32 GB de dados com potencial para mapeamento da mesma região, desde 1980 até os dias atuais. Para exemplificar, isto equivale a 8 mil músicas em formato MP3 (cada uma com 3.5 MB em média) ou ainda a 40 mil documentos de texto de tamanho médio 800 KB. No entanto, este volume de dados nem sempre é aproveitado de maneira integral. Em geral, os pesquisadores realizam buscas nos catálogos de imagens de satélite para fazer o download de cenas de interesse. Os critérios utilizados para a busca, na maior parte dos casos, são baseados em localização da imagem e tipo de sensor.

Nenhum catálogo de imagens de satélite inclui parâmetros de busca mais sofisticados, como nos mecanismos de busca de imagens da internet (exemplo <http://images.google.com/>), que incluem o conteúdo presente nas imagens. No caso de sensoriamento remoto, um exemplo de buscas seria: encontre imagens de satélite que contenham áreas de vegetação, rios ou lagos, e poucas nuvens. Ou ainda: encontre imagens de satélite da Amazônia com áreas de desmatamento maiores que áreas de floresta. Assim, este projeto visa produzir uma metodologia de geração de metadados para imagens de satélite baseada em conteúdo, por meio de operações básicas de processamento de imagens, aliadas a técnicas de mineração de dados e gerenciamento de grandes volumes de dados.

Objetivo

Dados de imagens de satélite em geral são armazenados em conjunto com metadados para auxiliar na sua descrição. Estes metadados geralmente se referem ao tipo de satélite que capturou a imagem, aos limites geográficos mapeados na imagem, ao período em que a imagem foi obtida, além de informações sobre os sensores. No entanto, ainda não existe um conjunto de metadados que represente o conteúdo da imagem. Assim, o objetivo principal desta proposta é o desenvolvimento de um mecanismo de reconhecimento de alvos e características particulares em imagens de satélite e o armazenamento efetivo destes resultados como metadados de um catálogo de imagens. Este mecanismo pode ser implementado como um sistema de computação dividido em módulos (reconhecimento e armazenamento). Este sistema precisa ser robusto para trabalhar em uma base já existente de imagens, além de processar os metadados para novas imagens que venham a ser adicionadas no catálogo.

Resultados esperados

A política de distribuição de imagens, produtos e informações no INPE de maneira gratuita sempre foi motivo de orgulho desta instituição. Assim, as ferramentas de busca geradas por meio deste projeto, quando estiverem funcionando de maneira operacional, serão disponibilizadas por meio de serviços web, disponíveis gratuitamente para usuários cadastrados. Os resultados preliminares serão documentados, além de relatórios e atas de reuniões, em artigos para congressos relacionados. Os resultados finais, explicando as técnicas e soluções criadas, serão enviados para revista internacional, avaliada por pares.

Candidatos

Para mais informações, enviar currículo Lattes atualizado para thales.korting@inpe.br, assunto do e-mail [2017/24086-2].