



MrSID : un format moderne d'image géospatiale



MrSID : Un Format Moderne D'image Géospatiale

Dans ce livre blanc, Extensis présente les rudiments de la compression, de la technologie MrSID et des fonctionnalités pouvant être utilisées dans vos applications et processus grâce au format MrSID.

I. La nécessité de compresser les données

Avec l'accroissement du nombre d'appareils mobiles et l'émergence du cloud, il n'a jamais été aussi facile d'accéder à des données géospatiales et de les analyser. De même, la demande de données *de qualité* n'a jamais été aussi élevée. Pour répondre à cette demande, les satellites commerciaux deviennent de plus en plus puissants et les véhicules aériens sans pilote envahissent le ciel. Toutefois, la prolifération des images géospatiales apporte avec elle son lot de problèmes. Diverses organisations, notamment des fournisseurs de contenus web et des collectivités locales, cherchent aujourd'hui des solutions pour stocker des quantités considérables de données et y accéder.

Les fichiers d'images numériques occupent beaucoup d'espace. Sans compression, les ressources physiques de stockage requises pour conserver l'imagerie brute sont immenses. De même, des réseaux à bande passante élevée et des stations de travail dotées de capacités de mémoire importantes sont nécessaires. La solution adoptée consiste souvent à stocker les images dans un format compressé, notamment en utilisant des formats de fichiers tels que JPEG. Les versions JPEG des images peuvent permettre d'accéder plus rapidement à des images de moindre résolution, mais la qualité n'est pas assez bonne pour réaliser des analyses et les exploiter. De ce fait, plusieurs versions de chaque jeu de données sont généralement conservées, dans des résolutions ou des taux de compression différents. Ainsi, on pourra tantôt utiliser une version pour comparer des images, tantôt une autre pour effectuer des analyses. Les problèmes de stockage et de gestion ne font que s'aggraver au fil du temps.

Nous travaillons avec des clients comme l'Institut d'études géologiques des États-Unis (USGS) et la NGA (National Geospatial-Intelligence Agency ; agence américaine du renseignement géospatial). Souvent, les processus géospatiaux supposent la prise en charge de fichiers représentant des centaines de gigaoctets, le stockage de l'imagerie sans altération de la qualité, l'utilisation de résolutions variées ou l'affichage rapide de différentes vues, et l'accès à des parties spécifiques d'un fichier, entre autres.

Depuis 20 ans, la technologie brevetée MrSID apporte au secteur des SIG (systèmes d'information géographique) des solutions modernes aux difficultés rencontrées. Des milliers de personnes se servent chaque jour de nos applications, telles que GeoExpress® et Express Server®, pour coder et fournir des images. Plusieurs milliers de personnes utilisent aussi des fichiers MrSID dans des centaines d'applications, dont Esri ArcGIS, ERDAS Imagine et Google Earth.

Avec le format MG4™, Extensis peut compresser des jeux de données tels que des images multispectrales et des nuages de points LiDAR. Dans ce livre blanc, nous vous présentons la technologie MrSID : les fonctionnalités qu'elle offre, ses champs d'application et son mode de fonctionnement.

II. Technologie MrSID

Qu'est-ce que la compression ?

La compression désigne simplement le fait de compacter des données afin qu'elles occupent moins d'espace. Toutefois, ce terme prend parfois une connotation négative dans notre secteur d'activité, car la compression de l'imagerie a longtemps entraîné une perte de qualité. Il fallait souvent se résigner à choisir entre résolution élevée et gain d'espace ; concilier les deux était inimaginable.

Prenons l'exemple d'une image aérienne en couleur 8 bits, de 6 000 pixels sur 6 000 pixels. Les données brutes de cette image occuperaient environ 100 mégaoctets (Mo) d'espace : 36 millions de pixels (6 000 x 6 000), sachant que chaque pixel requiert trois octets (trois bandes de couleur d'un octet chacun). Vous pouvez avoir besoin de dizaines, voire de centaines d'images de ce type, arrangées sous forme de dalles pour former une mosaïque complète. Dans ce cas de figure, toutefois, nous ne prenons en considération qu'une seule image.

Lorsque les pixels sont stockés de façon non compressée (données brutes), l'espace disque nécessaire est égal au nombre d'octets correspondant aux pixels de l'image (plus quelques octets pour les métadonnées telles que

les positions géospatiales de certains repères). Les formats de fichiers géospatiaux comme GeoTIFF stockent les données de manière brute, c'est-à-dire sans compression, en incluant tous les pixels. Les formats non compressés présentent l'avantage d'être sans perte, c'est-à-dire que les données enregistrées correspondent exactement aux données d'origine. La grande quantité d'espace disque nécessaire pour conserver les fichiers représente néanmoins un inconvénient majeur.

Toutefois, en nous aidant de divers algorithmes, nous sommes souvent capables de restituer les données des pixels plus efficacement. Supposons que les données se composent d'une séquence de cinq valeurs identiques : **123 123 123 123 123**. Vous pouvez coder cette séquence de manière abrégée : **123[5]**. Vous stockez ainsi la valeur une seule fois et vous y ajoutez un nombre de répétitions. Des techniques de ce type produisent des données compressées, qui occupent moins d'espace disque. En outre, la qualité de l'image n'est pas altérée.

À l'aide des algorithmes de compression, nous pouvons mettre au point des techniques qui évitent de devoir stocker les données correspondant à chaque pixel. Un nombre très précis comme **3,1415**, qui comporte cinq chiffres, peut être stocké en utilisant seulement le chiffre **3**, ce qui représente une réduction de 80 % de l'espace occupé par les chiffres (soit un taux de compression de 5:1). Le format de fichier JPEG, qui est souvent utilisé pour des petites images sur des sites web, est un format de compression avec pertes. Ce type de compression avec pertes s'effectue au détriment de la qualité de l'image. Les contours risquent d'être moins bien définis que sur l'image d'origine, les couleurs peuvent être moins vives et l'image peut subir de petites altérations. Dans certains cas, ce format pourra faire l'affaire si le niveau d'exigence est minime.

En résumé, nous pouvons distinguer trois catégories : des données non compressées et sans perte ; des données compressées et sans perte, et des données compressées, avec pertes. Seules des techniques avancées peuvent offrir un taux de compression élevé, associé à de hauts niveaux de performance et de qualité d'images. Cela passe par la conception d'algorithmes destinés spécifiquement à certains types de données, comme l'imagerie géospatiale, et à certains types de processus de travail.

Technologie MrSID : qualité et performances

La technologie MrSID repose sur des techniques de compression à la fois capables d'offrir une imagerie haute résolution et des niveaux de performance élevés, ainsi que de répondre aux exigences spécifiques de notre secteur d'activité.

Qualité de l'image

Pour l'imagerie standard, la technologie MrSID de compression sans perte produit un taux de compression de 2:1. Cela signifie que vous avez besoin de deux fois moins d'espace de stockage, mais que vous conservez des données identiques aux données d'origine sur le plan numérique.

Les taux de compression avec pertes de la technologie MrSID peuvent aller jusqu'à 20:1, ce qui permet d'économiser encore plus d'espace de stockage tout en offrant une qualité d'image telle qu'aucun changement n'est visible à l'œil nu dans la plupart des processus. Nous qualifions cela de compression sans *perte visible à l'œil nu*.

Cela signifie qu'une perte de données existe, mais qu'elle est imperceptible dans le cadre de l'usage qui est fait de l'image. Des taux de compression encore plus élevés sont évidemment possibles. Selon la qualité d'image dont vous avez besoin et le type d'image d'origine, des taux de 40:1 ou plus peuvent être appliqués.

Performances de codage

Les jeux de données géospatiales peuvent occuper plusieurs téraoctets d'espace disque. Un jeu de données se compose parfois d'un seul fichier volumineux. Toutefois, la plupart du temps, il comprend des centaines de petites dalles qui forment ensemble une grande mosaïque. Quoi qu'il en soit, pour traiter et compresser ces immenses jeux de données, l'utilisateur doit posséder un matériel de catégorie serveur et des capacités de mémoire faramineuses, ou doit travailler de façon fragmentée, en n'affichant qu'un sous-ensemble de dalles à la fois.

La technologie MrSID permet de surmonter ces difficultés. En effet, les fichiers de plus de 2 Go et les processeurs 64 bits sont pris en charge. Ainsi, la taille de l'imagerie pouvant être compressée n'a quasiment plus aucune limite.

Performances d'affichage

Même en appliquant un taux de compression de 20:1, un jeu de données de 20 To compressé produit un fichier MrSID très lourd (1 To). De manière générale, les utilisateurs craignent de travailler sur des fichiers aussi volumineux parce que les applications SIG tentent souvent d'ingérer le fichier entier, ce qui accapare de façon excessive les capacités de mémoire et l'unité centrale, produisant parfois des erreurs irrécupérables.

Deux aspects de la technologie MrSID permettent de les rassurer. Tout d'abord, la technique de codage que nous utilisons crée des résolutions multiples de l'image dans le fichier MrSID généré. Les effets sont similaires à ceux que produisent les créations de pyramides d'images. Toutefois, la technique MrSID n'occasionne pas de fichiers supplémentaires à chaque niveau de la pyramide. Les niveaux de résolution sont inhérents au modèle de codage et intégralement contenus dans un seul fichier de sortie MrSID. Ces niveaux sont échelonnés de sorte que chaque niveau représente un quart du niveau précédent : par exemple, une image complète 1024 x 1024, puis une image à l'échelle d'un quart (512 x 512), puis 1/16e (256 x 256) et ainsi de suite jusqu'à une icône ou une vignette de 32 x 32. Certaines applications sont capables de n'extraire et de ne traiter que le niveau de détail requis, sans devoir décoder (et parfois sous-échantillonner manuellement) l'image entière en résolution maximale.

Outre le choix du niveau de résolution, la technologie MrSID offre une décompression sélective, qui permet aux applications de ne décoder qu'une partie (zone géographique) du fichier. À l'inverse, certains formats de fichiers ou modes de compression requièrent le décodage de l'image entière, même si la zone que vous souhaitez afficher sur l'écran ne représente qu'une petite partie de l'image.

Bref historique de la technologie MrSID

Le prototype de la technologie MrSID a été développé dans les laboratoires américains de recherche de Los Alamos en 1992. En 1998, LizardTech a créé et lancé une version commerciale de la technologie, connue sous le nom de format MrSID Generation 2 (MG2). La version suivante, MrSID Generation 3 (MG3), a fait son apparition en 2002. Ce format a apporté une meilleure qualité d'image et des fonctionnalités clés, comme le codage sans perte. La version la plus récente du format, MrSID Generation 4 (MG4), existe depuis 2009. Le format MG4 prend en charge l'imagerie multispectrale et les bandes alpha.

De plus, le format MG4 prend en charge les données LiDAR. Les utilisateurs des données LiDAR MG4 bénéficient d'un grand nombre des mêmes fonctionnalités clés de notre technologie de compression raster, y compris la compression sans perte, la compression sans perte visible à l'œil nu et le décodage sélectif.

Les produits LizardTech ont été intégrés à la gamme de produits Extensis en 2018 et la technologie MrSID fait désormais partie de l'ensemble d'outils et d'applications d'Extensis, qui comprend des applications complètes de codage, des serveurs d'images et des visionneuses légères. Par ailleurs, grâce aux intégrations de nos partenaires, des centaines d'applications SIG prennent en charge les fichiers MrSID. Que vous utilisiez des fichiers MG2 anciens dans vos applications héritées, des fichiers MG3 sans perte dans vos processus actuels ou que vous souhaitiez coder efficacement des nuages de points LiDAR ou votre imagerie multispectrale, les technologies de compression MrSID vous aident à mener à bien vos projets.

À propos d'Extensis

Depuis 25 ans, Extensis® est l'un des principaux éditeurs de solutions permettant aux professionnels créatifs d'accroître le retour sur investissement ainsi que la valeur de leurs ressources numériques, polices et images. Utilisées par plus de 100 000 professionnels et 5 000 entreprises du monde entier, les solutions d'Extensis accélèrent les flux de travail pour permettre aux clients d'atteindre leurs objectifs plus rapidement.

Fondé en 1993, Extensis est basé à Portland, dans l'Oregon, et dispose également de bureaux à Seattle, à New York, au Royaume-Uni, en France, en Allemagne et en Australie. Pour en savoir plus sur la suite complète de solutions Extensis pour la gestion des ressources numériques, la gestion des ressources de polices, la gestion des ressources d'images et la compression d'images, rendez-vous sur le site <https://www.extensis.com/fr-fr/> ou suivez Extensis sur Twitter @extensis
