



MrSID: Ein modernes Geodaten-bildformat



MrSID: Ein modernes Geodatenbildformat

In diesem Whitepaper stellt Ihnen Extensis das Konzept der Komprimierung vor sowie die MrSID-Technologie und die Funktionen, mit denen das MrSID-Format zu Ihren Anwendungen und Arbeitsprozessen beitragen kann.

I. Warum Datenkomprimierung notwendig ist

Mit der zunehmenden Nutzung mobiler Geräte und dem Beginn von Cloud-Computing wurde es erstmals extrem einfach, auf Geodaten zuzugreifen und diese zu analysieren. In gleichem Maße ist die Nachfrage nach guten Daten gestiegen. Um dieser Nachfrage gerecht zu werden, sind kommerzielle Satelliten immer leistungsfähiger geworden und unbemannte Luftfahrzeuge werden immer beliebter. Leider hat die starke Vermehrung von Geodatenbildern ihre eigenen Tücken. Organisationen, die so facettenreich sind wie die Anbieter von Web-Content und Bezirksregierungen, stehen nun gleichermaßen vor der Aufgabe, riesige Datenmengen speichern und darauf zugreifen zu müssen.

Da digitale Bilddateien so groß sind, erfordert das Beibehalten von Bildern als unkomprimierte RAW-Daten immense physische Speicherkapazitäten. Außerdem benötigen Sie für den Zugriff auf RAW-Daten Netzwerke mit hoher Bandbreite und großen Arbeitsspeichern. Eine übliche Alternative besteht darin, Bilder in komprimierter Form in Dateiformaten wie JPEG abzuspeichern. Obwohl die JPEG-Versionen der Bilder einen schnelleren Zugriff auf Bildübersichten mit niedrigerer Auflösung ermöglichen, ist ihre Qualität nicht für die Analyse und Verwertung in hoher Auflösung geeignet. Dies führt oft dazu, dass mehrere Versionen jedes Datensatzes mit unterschiedlichen Auflösungen oder Komprimierungsverhältnissen gespeichert werden – eine zum Durchsuchen, eine für die Analyse und so weiter. Das Problem des Speicherns und des Beibehaltens der Daten wird schlimmer, nicht besser.

Wir wissen durch die Zusammenarbeit mit Kunden wie dem United States Geological Survey (USGS) und der National Geospatial-Intelligence Agency (NGA), dass die Arbeitsprozesse heutiger Geodaten regelmäßig Unterstützung für Dateien in einer Größe von Hunderten von Gigabytes erfordern. Aber auch das Speichern der Bilder ohne merklichen Qualitätsverlust, Zugriff auf mehrere Auflösungen oder Schnellübersichten ist ein Muss. Genau wie der effiziente Direktzugriff in die Datei zur Unterstützung beliebiger Szenen-Anfragen und so weiter.

Seit zwei Jahrzehnten ist die patentierte MrSID-Technologie hinsichtlich dieser Anforderungen führend in der GIS-Branche. Unsere Anwendungen wie GeoExpress® und Express Server® werden täglich von Tausenden von Menschen zum Verschlüsseln und zur Übertragung der Bilder benutzt. Und viele Tausende mehr verwenden MrSID-Dateien in Hunderten von Anwendungen einschließlich Esri ArcGIS, ERDAS Imagine und Google Earth.

Mit dem MG4™ -Format bietet Extensis die Komprimierung für Datensätze wie Multispektralbilder und LiDAR-Punktwolken. In diesem Whitepaper stellen wir Ihnen die MrSID-Technologie vor: Welche Funktionen sie bietet, wo sie angewendet werden kann und wie sie funktioniert.

II. Die MrSID-Technologie

Was ist Komprimierung?

Komprimierung bedeutet lediglich, Daten kompakter zu machen, so dass sie weniger Platz einnehmen. Aber der Begriff hat in unserer Branche manchmal negative Assoziationen: Oft führte die Komprimierung von Bildern zu einer Verschlechterung der Datenqualität. Sie hatten entweder eine hohe bildqualität oder sparten Speicherplatz – aber nicht beides.

Nehmen wir ein typisches Luftbild in 8-Bit-Farbe, das 6.000 Pixel breit und 6.000 Pixel hoch ist. Das bedeutet, die RAW-Daten bestehen aus etwa 100 Megabyte (MB) auf der Festplatte: also 36 Millionen Pixel (6.000 × 6.000). Für jedes Pixel werden drei Bytes (drei Farbbänder von jeweils einem Byte) benötigt. Ein üblicher Arbeitsprozess kann Dutzende oder eigentlich sogar Hunderte solcher Bilder haben, die als Kacheln angeordnet sind und so ein komplettes Mosaik bilden. Aber betrachten wir mal nur das eine Bild genauer.

Wenn die Pixel in einem *unkomprimierten* (RAW-) Format wie diesem gespeichert sind, sind die Bilddaten überhaupt nicht komprimiert: die Kapazität an Speicherplatz, die benötigt wird, entspricht der Anzahl der Bytes, die benötigt werden, um die Pixel darzustellen (plus etwas mehr Kapazität für Metadaten wie die Geo-Positionen der Eckpunkte). Geodatenformate wie GeoTIFF speichern Daten in diesem RAW- oder unkomprimierten Format: Jedes

einzelne Pixel wird vollständig dargestellt. Unkomprimierte Formate haben den Vorteil, dass die Daten jedes Pixels genau so dargestellt werden, wie es ursprünglich aufgezeichnet wurde – das nennen wir verlustfreie Bilddaten. Der offensichtliche Nachteil ist natürlich der Speicherplatz, den Sie insgesamt benötigen. Durch den Einsatz verschiedener Algorithmen können wir die Pixeldaten jedoch oft in einer effizienteren Form darstellen.

Gefühlsmäßig kann man sich das so vorstellen: Wenn die Daten aus einer Abfolge von fünf identischen Werten wie **123 123 123 123 123** bestünden, könnten Sie diese stattdessen mit einer Art Kurzschrift wie etwa **123 [5]** verschlüsseln, in der Sie

den Wert nur einmal speichern, aber eine „Wiederholzahl“ hinzufügen. Techniken wie diese liefern Daten, die komprimiert sind. Das bedeutet, dass weniger Speicherplatz benötigt wird, aber die Bilddaten trotzdem verlustfrei sind.

Wenn wir unsere Komprimierungsalgorithmen weiter vorantreiben, können wir Techniken finden, die die Daten zwar nicht genau für jedes Pixel speichern, sondern nur die Annäherung der Daten. Wiederum gefühlsmäßig: eine hochpräzise Zahl wie **3,1415**, die fünf Ziffern erfordert, könnte nur als die einstellige Zahl **3** gespeichert werden, also mit einer 80 %-igen Reduzierung des Ziffernraums (ein Komprimierungsverhältnis von 5:1). Das JPEG-Dateiformat, das häufig für kleine Bilder im Internet verwendet wird, bedient sich dieser Art von verlustbehafteter Komprimierung. Die Strafe für solche verlustbehafteten Komprimierungsverfahren ist, dass Sie normalerweise die Bildqualität verlieren: die Kanten sind dann nicht so scharf, die Farben erscheinen flacher, kleine Artefakte entstehen. Für etwas weniger anspruchsvolle Geodaten-Arbeitsprozesse mag dies akzeptabel sein.

Kurz gesagt haben wir ein Spektrum von drei Arten von Komprimierung und Datenverlust: unkomprimiert und verlustfrei, komprimiert und verlustfrei, und komprimiert und verlustbehaftet. Die unbedarften Techniken, die gerade beschrieben wurden, um diese Arten der Komprimierung zu erreichen funktionieren zwar, würden aber auf keinen Fall hohe Komprimierungsverhältnisse, eine hohe Bildqualität und gute Ergebnisse liefern. Um die Komprimierung in der realen Welt zu verbessern, müssen die Algorithmen für bestimmte Arten von Daten entwickelt werden, wie Geodatenbilder und bestimmte Arten von Arbeitsprozessen.

MrSID-Technologie: Qualität und gute Ergebnisse

Die von der MrSID-Technologie verwendeten Komprimierungstechniken liefern sowohl qualitativ hochwertige Bilder als auch gute Ergebnisse. Dabei werden wir den anspruchsvollen Arbeitsprozessen unserer Branche gerecht.

Bildqualität

Die verlustfreie Komprimierung der MrSID-Technologie ergibt Komprimierungsraten von 2:1 für typische Bilder. Das heißt, Sie brauchen nur die Hälfte des Speicherplatzes und behalten trotzdem Ihre numerisch identischen Originaldaten.

Für weitere Einsparungen beim Speicherplatz kann die verlustbehaftete Komprimierung der MrSID-Technologie typische Verhältnisse von bis zu 20:1 herausholen, und trotzdem noch eine Bildqualität bieten, bei denen die Daten für die meisten Workflows nicht mit bloßem Auge zu unterscheiden sind. Wir nennen diese Art der Komprimierung optisch verlustfrei: Es gibt immer noch einen Datenverlust, aber im Hinblick auf das, wofür das Bild verwendet wird, ist die Größe des Verlustes kaum wahrnehmbar.

Höhere Komprimierungsverhältnisse sind natürlich möglich. Abhängig davon, wie viel Sie an Bildqualität beibehalten müssen und je nach Art Ihrer Originalbilder, können Verhältnisse von 40:1 und darüber hinaus angewendet werden.

Codierungsleistung

Geo-Datensätze können bis zu mehreren Terabytes an Speicherplatz beanspruchen. Manchmal ist der Datensatz nur eine große Datei. Oft aber besteht er aus Hunderten kleinerer Kacheln, die zusammen ein großes Mosaik bilden. So

oder so, mit vielen Anwendungen ist die Menge an Speicher, die für die Verarbeitung und Komprimierung so großer Datenmengen benötigt wird, untragbar. Denn dafür brauchen Sie Hardware auf Serverniveau oder der Endbenutzer wird dazu gezwungen, mit einem „Stückwerk“ an Bildern zu arbeiten – also nur mit einem Teilsatz der Kacheln auf einmal.

Die MrSID-Technologie wurde entwickelt, um dieses Problem zu beheben. Dank der vollen Unterstützung für Eingabe- und Ausgabedateien, die größer als 2 GB sind und der Unterstützung für 64-Bit-Prozessoren, gibt es praktisch keine Begrenzung für die Größe der Bilder mehr, die komprimiert werden können.

Darstellungsleistung

Selbst bei einem Verhältnis von 20:1 wird aus einem 20-TB großen, komprimierten Datensatz immer noch eine extrem große (1 TB) MrSID-Datei. Die Benutzer sind oft ängstlich, wenn Sie mit Dateien dieser Größe arbeiten, da viele GIS-Anwendungen versuchen werden, die ganze Datei einzulesen. Dies verursacht eine übermäßige und gelegentlich fatale Beanspruchung der CPU und des Speichers. Zwei Aspekte der MrSID-Technologie lösen dieses Problem.

Erstens erstellt die von uns verwendete Codierungstechnik mehrere Auflösungen des Bildes innerhalb der generierten MrSID-Datei. Das ist ähnlich dem Trick, „Bildpyramiden“ zu erstellen, aber die MrSID-Technik erstellt keine zusätzlichen Dateien für alle die Pyramidenebenen, noch übernimmt sie die entsprechend freigewordene Kapazität an Speicherplatz. Stattdessen sind die Auflösungsstufen inhärent im Codierungsschema und vollständig in der einzelnen MrSID-Ausgabedatei selbst enthalten. Diese Ebenen werden so skaliert, dass jede Ebene einem Viertel der vorherigen Ebene entspricht: zum Beispiel das Vollbild von 1024 x 1024, dann das auf ein Viertel skalierte Bild von 512 x 512, dann das auf ein Sechzehntel skalierte Bild von 256 x 256 und so weiter bis zu einem „Icon“ oder „Thumbnail“ von üblicherweise 32 x 32. Die Anwendungen können so eingestellt werden, dass sie nur die erforderliche Detailtiefe extrahieren und verarbeiten, ohne das gesamte Bild mit voller Auflösung dekodieren (und vielleicht dann manuell abtasten) zu müssen.

Die MrSID-Technologie ermöglicht es nicht nur, die gewünschte Auflösungsebene anzufordern, sondern sie bietet auch eine selektive Dekompression. Damit können Anwendungen nur die Szene (das geografische Gebiet) aus der Datei anfordern und dekodieren, die von Interesse ist. Einige andere Dateiformate und Komprimierungsverfahren erfordern dagegen, dass das gesamte Bild decodiert wird, selbst wenn nur ein kleiner Teil auf dem Bildschirm angezeigt werden soll.

Ein kurzer Abriss zu MrSID

Der Prototyp für die MrSID-Technologie wurde 1992 im Los Alamos National Laboratory entwickelt. LizardTech entwickelte eine kommerzielle Folgeversion der Technologie, die MrSID Generation 2 (MG2-) Format genannt wurde und führte diese 1998 ein. Die nächste Version, das MrSID Generation 3 (MG3-) Format, das 2002 eingeführt wurde, bot eine bessere Bildqualität und wichtige Funktionen wie die verlustfreie Kodierung. Die neueste Version des MrSID-Dateiformats, die MrSID Generation 4 (MG4) wurde 2009 eingeführt. Das MG4-Format unterstützt multispektrale Bilder und Alpha-Bänder.

Zusätzlich wurde im MG4-Format die Unterstützung für LiDAR-Daten eingeführt. Die Benutzer von MG4 LiDAR-Daten erhalten viele der gleichen wichtigen Funktionen, die in unserer Raster-Kompressionstechnologie vorhanden sind – einschließlich verlustfreier Komprimierung, nahezu verlustfreier Komprimierung und selektiver Decodierung.

LizardTech wurde 2018 in die Extensis-Produktfamilie aufgenommen. Damit ist die MrSID-Technologie nun ein Teil der Suite von Anwendungen und Tools von Extensis, die von der Codierung mit vollem Funktionsumfang bis hin zu Anwendungen für Image-Server für leichte Viewer reichen. Darüber hinaus werden die MrSID-Dateien durch unsere Integrationspartner in Hunderten von GIS-Anwendungen unterstützt. Ob Sie ältere MG2-Dateien in Ihren älteren Anwendungen verwenden, vom Vorteil der verlustfreien Dateien des MG3-Formats in Ihren aktuellen Arbeitsprozessen profitieren oder neuerdings effizient multispektrale Raster- oder LiDAR-Punktwolken-Datensätze kodieren müssen – die MrSID-Komprimierungstechnologien entsprechen Ihren Anforderungen.

Über Extensis

In den 25 Jahren seines Bestehens ist Extensis® ein führender Entwickler von Lösungen, die Unternehmen helfen, den ROI und den Wert ihrer digitalen Assets, Schriften und großen Bilder zu steigern. Die Lösungen von Extensis werden von mehr als 100.000 professionellen Anwendern und 5.000 Unternehmen auf der ganzen Welt eingesetzt und beschleunigen Arbeitsabläufe, damit Kunden ihre Ziele schneller erreichen können.

Extensis wurde 1993 gegründet und hat seinen Hauptsitz in Portland, Oregon sowie Niederlassungen in Seattle, New York, Großbritannien, Frankreich, Deutschland und Australien. Um mehr über Extensis umfassende Lösungen für Digital Asset Management, Schriftarten-Asset-Management, Bild-Asset-Management und Bildkomprimierung zu erfahren besuchen Sie unsere Website <https://www.extensis.com/de-de> oder folgen uns auf Twitter @extensis.
