



Unterrichtsmaterialien zum Thema

Datenkompression

JAHRGANGSSTUFE 11–13

Material für Lehrkräfte

Projektinformation

Diese Unterrichtsmaterialien sind im Rahmen des Projektes „Columbus Eye – Live-Bilder von der ISS im Schulunterricht“ entstanden. Das Projekt Columbus Eye wird von der Raumfahrt-Agentur des Deutschen Zentrums für Luft- und Raumfahrt e.V. mit Mitteln des Bundesministeriums für Wirtschaft und Energie aufgrund eines Beschlusses des Deutschen Bundestages unter dem Förderkennzeichen 50 JR 1703 gefördert.

Das übergeordnete Projektziel besteht in der Erarbeitung eines umfassenden Angebots an digitalen

Lernmaterialien für den Einsatz im Schulunterricht. Dieses Angebot umfasst interaktive Lerntools und Arbeitsblätter, die über ein Lernportal zur Verfügung gestellt werden.

Für dieses Lehrmaterial, die dazugehörige App und Schülermaterial gilt: © Columbus Eye (CC BY-NC-ND 2.0 DE)

<http://www.columbuseye.rub.de>


 RUHR
UNIVERSITÄT
BOCHUM

 UNIVERSITÄT
BONN

 Bundesministerium
für Wirtschaft
und Klimaschutz


Übersicht

Jahrgangstufe

11 12 13

Niveau

● ● ● ○ ○

Zeitbedarf

90 Minuten

Autoren

 Nico Kraft,
Darius Happe,
Connor Vormstein,
Claudia Lindner

Ziele

Die Schülerinnen und Schüler (SuS) sollen...

- die Übertragung eines Satellitenbildes zur Erde beschreiben,
- verschiedene Kompressionsverfahren erklären,
- die Eignung der verschiedenen Kompressionsverfahren vor dem Hintergrund des Anwendungsbereichs beurteilen und Stellung nehmen,
- den Median-Cut als rekursive Möglichkeit zur Farbreduktion implementieren
- und mit einem iterativen Verfahren vergleichen.

Themen

Verlustfreie und verlustbehaftete Bildkompression

Iterative und rekursive Algorithmen

Satellitenkommunikation

Satellitenbilder

Medien & Material

Arbeitsblatt „Datenkompression“

Lehrkräftematerial „Datenkompression“

Erde-Marker

App „Columbus Eye“ – Part „Datenkompression“

Python-Skript (SuS und LuL)

 DOWNLOAD ON THE
Apple Store

 GET IT ON
Google Play


Didaktische Anmerkungen

Kompetenzen

Die nachfolgend aufgeführten Kompetenzbereiche orientieren sich am Kernlehrplan (KLP) des Landes Baden-Württemberg, da dieser die Kompetenzen gerade im Themenbereich der Datenkompression ausführlich darlegt. Im nächsten Abschnitt werden die inhaltlichen Bezüge zu den KLPs anderer Bundesländer tabellarisch hergestellt.

In dieser Unterrichtseinheit schulen die SuS verschiedene Teil-Kompetenzen (implementieren, dokumentieren, kommentieren, kooperieren, darstellen, analysieren und reflektieren), indem sie...

- vorhandenen Code und Dokumentation nachvollziehen und erklären (2.3 Kommunizieren und Kooperieren 2, 4)
- einen eigenen rekursiven Algorithmus für die Kompression von Bildern implementieren und dokumentieren (Median Cut Algorithmus) (2.2 Modellieren und Implementieren 9, 10; 2.3 Kommunizieren und Kooperieren 3)
- eine iterative Implementierung desselben Algorithmus als solchen klassifizieren, analysieren und die Unterschiede zu ihrer eigenen rekursiven Implementierung beschreiben (2.2 Modellieren und Implementieren 13; 2.4 Analysieren und Bewerten 1, 3)
- in Kleingruppenarbeit (ggf. unter Nutzung geeigneter digitaler Werkzeuge zum kollaborativen Arbeiten) verschiedene Kompressionsverfahren analysieren und ihre Ergebnisse präsentieren (2.2 Modellieren und Implementieren 13; 2.3 Kommunizieren und Kooperieren 5, 6)

Lehrplanbezug

Eine Aufführung ausgewählter Bundesländer:

Bundesland	Jahrgangsstufe	Inhalts-/Lernfeld
Baden-Württemberg	11-12	Daten und Codierung [3,5,7] Algorithmen [1,2,4,8,10,13,16] rekursive Algorithmen
Berlin/Brandenburg	11-12	Softwareentwicklung (Algorithmen)
Hessen	12-13	Objektorientierte Modellierung (iterative und rekursive Algorithmen) Wahlthema Computergrafik (Komprimierung)
Nordrhein-Westfalen	12-13	Algorithmen (iterative und rekursive Algorithmen darstellen und Rekursive implementieren)
Thüringen	Oberstufe	Lernbereich Angewandte Informatik (verlustfreie und –behaftete Datenkompression beschreiben / Verfahren erläutern) Lernbereich Praktische Informatik (iterative und rekursive Algorithmen analysieren und vergleichen)
Schleswig-Holstein	Sek I -> mittlerer Schulabschluss	Daten und Informationen (Klassifizieren verlustfreie und-behaftete Kompression und beschreiben deren Merkmale und Anwendungsbe-reiche)
		Aufgrund des hohen Anspruchsniveaus für die Sek I kann die Implementierung ausgelassen werden.

Voraussetzungen

Es muss eine Möglichkeit bestehen, ein Python-Skript zu öffnen und zu bearbeiten.

Etwa die Hälfte der SuS sollte die App auf ihrem Gerät verfügbar haben. Die Lerneinheiten in der App werden automatisch von der Hochschulcloud sciebo heruntergeladen. Bei Problemen kann [hier](https://sciebo.de/de/hilfe/sciebo-news.html) (<https://sciebo.de/de/hilfe/sciebo-news.html>) nachgeschaut werden.

Fachlich müssen die SuS in der Lage sein, rekursive und iterative Algorithmen zu klassifizieren.

Vorbereitung

Lassen Sie die SuS die App „Columbus Eye“ einige Tage vor der geplanten Stunde herunterladen. Hierzu kann der Link verschickt, der QR-Code ausgeteilt, oder beim Google Play Store in die Suchleiste einfach „Columbus Eye“ eingegeben werden. Der eigentliche Download sollte, um niemandes Datenvolumen zu belasten, von den SuS im heimischen WLAN durchgeführt werden, sofern es kein (zuverlässiges) Schul-W-Lan gibt.

Sobald die App heruntergeladen ist, müssen noch die Daten für den Part „Datenkompression“ hinzugeladen werden.

Hinweis: Möglicherweise funktioniert die App nicht auf allen Smartphones, was mit deren Betriebssystemen und -versionen zusammenhängt. Dies stellt jedoch kein Problem dar. Solange jede Kleingruppe in der späteren Bearbeitung des Arbeitsblattes über ein Smartphone mit funktionierender „Columbus Eye“-App verfügt, können die Aufgaben problemlos durchgeführt werden.

Der Erde-Marker muss nicht für jede/n Lernende/n ausgedruckt werden, da mehrere Gruppen den gleichen Marker nutzen können. Daher muss dieses AB gesondert vom SuS-Material ausgedruckt werden.

Stundenverlaufsplan

Zeit	Phase	Unterrichtsgeschehen	Methodisch-didaktischer Kommentar	Sozialform	Medien
20min	Einstieg	Interaktiver Umgang bei der Datenübertragung einer DESIS-Szene zur Erde unter Herausstellung der Schwierigkeiten und benötigten Dauer der Kommunikation.	Hohe SuS-Aktivität (Erkunden eines realen Beispiels) Nach der Bearbeitung des ersten Aufgabenteils kann auch eine kurze Besprechung / Reflexionsphase erfolgen. Zum Aufgabenteil 1d) kann zudem zuvor auch eine Schätzfrage gestellt werden, ehe der genaue Wert berechnet wird, um so ein niedrigschwelliges Angebot der Beteiligung zu schaffen. Die Ergebnisse werden stichpunktartig notiert. Mit der App zu experimentieren ist ausdrücklich erwünscht.	PA / GA	App / AB Nr. 1
	Leitfrage	Die Fragestellung der Stunde wird herausgearbeitet: z.B. Wie kann die Datenübertragung von der ISS zur Erde verbessert werden?	Die Fragestellung ergibt sich aus der Diskussion in Aufgabe 1 e) und sollte gut sichtbar angeschrieben werden.	UG	Tafel / Beamer
2 min	Gruppeneinteilung		Einteilung in Gruppen á 3-4 SuS, später präsentiert je eine Gruppe pro Kompressionsverfahren; die anderen Gruppen ergänzen.		
15 min	Erarbeitung 1	Zur Beantwortung der Fragestellung erarbeiten die SuS in Kleingruppen drei verschiedene Kompressionsmodi unter Verwendung des ABs und der App. Zudem beurteilen sie die Eignung ihres jeweiligen Kompressionsverfahren für das Satellitenbild des Átnas.	Die Ergebnisse werden für alle in Aufgabenteil c) bei der Präsentation festgehalten. Es können auch digitale Werkzeuge zur kollaborativen Arbeit genutzt werden (MKR).	PA / GA	App / AB Nr. 2 a) – b)
10 min	Zwischensicherung	Im Plenum werden die Ergebnisse aus den Kleingruppen vorgestellt und diskutiert.	Die Präsentationsform ist gerade in Bezug auf die Ausstattung der Schule frei wählbar. Am Ende sollten alle Ergebnisse schriftlich gesichert sein, z.B. auf einem Plakat oder einer Präsentationsfolie.	UG	AB Nr. 2 c) / Tafel / Beamer / Plakat
20 min	Erarbeitung 2	Die SuS implementieren anhand einer Anleitung den Median Cut als Möglichkeit der Kompression durch die Farbreduktion in dem beigefügten Python-Skript.	Die SuS sollen bei Fragen auch Recherchieren und sich so ihre benötigten Informationen finden (-> MKR).	PA / GA	Python-Skript / App / AB Nr. 3)
10 min	Erarbeitung 3	Die SuS beschreiben einen gegebenen iterativen Algorithmus zu diesem Verfahren und klassifizieren ihn als einen solchen.		PA / GA	AB Nr. 4)

Zeit	Phase	Unterrichtsgeschehen	Methodisch-didaktischer Kommentar	Sozialform	Medien
5 min	Sicherung / Transfer	Die SuS nehmen Stellung zu der Aussage: „Es gibt kein perfektes Kompressionsverfahren.“	Beantwortung der Leitfrage unter spezieller Betrachtung des Stundenthemas der Datenkompression und der verschiedenen Anwendungsbereiche, welche auch über das Anwendungsbeispiel der Stunde hinaus gehen kann. Die SuS sollten zunächst selbst Überlegungen anstellen, diese dann kurz in Form einer Murnelphase mit ihren Nachbarn/Kleingruppen austauschen, um so bessere Ergebnisse bei der kurzen Plenumsdiskussion zu erhalten	PA GA UG	App / AB Nr. / 5

Inhaltliche Vertiefung

Zur vertieften Auseinandersetzung mit den Lerninhalten bieten sich zudem folgende Lernvideos von Esero Germany an (<https://esero.de/materialien/lernfilme>):

- Die spektrale Auflösung

Kurzbeschreibung: Spektrale Satellitensensoren sehen die Erde grau. Dies liegt daran, dass sie jeden Wellenlängenbereichen getrennt voneinander in so genannten Spektralkanälen aufnehmen. Die spektrale Auflösung, also die Anzahl der Spektralkanäle, ist eine wichtige Eigenschaft von Satellitensensoren. Doch welche spektrale Auflösung sollte ein Sensor haben und warum können nicht alle Bereiche des elektromagnetischen Spektrums gemessen werden?

Direktlink: <https://www.youtube.com/watch?v=H4orF007AN0&t=2s>

- Die Welt in Infrarot

Kurzbeschreibung: Den meisten Menschen ist Infrarot aus dem Alltag ein Begriff. Wärmebildkameras, Infrarotlampen, aber auch Fernbedingungen oder die Körperscanner am Flughafen verwenden Infrarot. Wir Menschen können Infrarot nicht sehen, nehmen es aber teilweise als Wärme wahr. Also was genau ist Infrarot und wie wird es in der Erdbeobachtung eingesetzt?

Direktlink: <https://www.youtube.com/watch?v=k4m-JBZkMXw>

- Elektromagnetische Spektrum – Einführung

Kurzbeschreibung: Satelliten können Merkmale unserer Umwelt aufzeichnen, die uns Menschen verborgen bleiben. Doch was genau sehen sie denn, was wir nicht sehen?

Direktlink: <https://www.youtube.com/watch?v=p0wWJX0Zp60>

Lösungen

1. Aufgabenstellung

- a) Übertragung dauert sehr lange, da die ISS nur in etwa 10% der Zeit eine Verbindung herstellen kann.
- b) Weg der Übertragung: Es können nun sowohl direkte Verbindungen zwischen ISS und Bodenstation als auch indirekte Verbindungen über die EDRS Satelliten aufgebaut werden.
- c) Die drei Bodenstationen befinden sich alle in Europa, da das E(uropean)DRS ein Projekt der ESA (European Space Agency) ist und es in diesem Fall keine Kooperation mit anderen Raumfahrtagenturen gibt. Dementsprechend sind die ersten Satelliten, wie in der App zu sehen, geostationär über Europa platziert worden.

Um eine redundante Verbindung sicher zu stellen sind es drei Empfangsstationen. Falls an einer oder zwei der Bodenstationen Wartungsarbeiten stattfinden und die Empfangsstationen ausfallen, gibt es mit hoher Wahrscheinlichkeit zumindest eine aktive Station.

- d) $2,5 \text{ GB} = 2.500 \text{ MB}$ Datenpaket
bei einer Übertragungsrate von $1,8 \text{ Gbit/s} = 1.800 \text{ Mbit/s} = 225 \text{ MB/s}$.
 $2.500 \text{ MB} / 225 \text{ MB/s} = 11,11 \text{ s}$ ununterbrochen Verbindung

Weiterführende Hintergrundinformationen/Berechnungen:

Ohne Satellitennetz wird von einer Verbindungszeit von ~10% ausgegangen. Das entspricht einer Übertragungsdauer von 111 Sekunden, also 1 Minute und 51 Sekunden, für o. g. Datenpaket.

Mit Satellitennetz (EDRS) wird von einer Verbindungszeit von ~70% ausgegangen, was einer Übertragungsdauer von 16,66 Sekunden für o. g. Datenpaket entspricht.

- e) individuelle Lösung!
→ mögliche Ansätze oder Ideen sind: mehr Satelliten; bessere / mehr Technik zum Senden, Übertragen und Empfangen; Daten verringern / komprimieren

2. Aufgabenstellung

- a) Erläuterung des jeweiligen Kompressionsverfahrens:

- i. Farbreduktion:

Bei diesem Verfahren wird der Farbraum des Bildes reduziert, um so den Speicherverbrauch stark zu reduzieren. Dabei werden gleichfalls die Bildqualität stark reduziert sowie die Farbwerte verändert und können nicht wiederhergestellt werden. Folglich handelt es sich um ein verlustbehaftetes Kompressionsverfahren.

- ii. Redundanz-/Ähnlichkeitssuche:

Bei der Redundanzreduktion werden Redundanzen und Häufigkeiten genutzt, um Symbole einzusparen. Es handelt es sich um ein verlustfreies Kompressionsverfahren, d.h. es kommt zu keinen Veränderungen der Farbwerte. Diese Methode kann beispielsweise vorher mit einer Ähnlichkeitssuche kombiniert werden. Dieses Verfahren ist ein verlustbehaftetes Kompressionsverfahren, bei dem – je nach verwendetem Grenzwert – nur wenige Informationen verloren gehen.

Hinweis: Tatsächlich ist die Ähnlichkeitssuche zwar ein verlustbehaftetes Verfahren, jedoch sind die Auswirkungen der Kompression im Bild kaum mit dem bloßen Auge zu erkennen. Das Verfahren ist so effektiv, dass erst mit der Nutzung der Lupenfunktion innerhalb der App und bei Betrachtung des 64x64 Pixel Ausschnittes eine Änderung sichtbar wird.

iii. Farbunterabtastung:

Nachdem das Bild in einen Farbraum konvertiert wurde, bei dem die Helligkeit separat abgespeichert wird, findet eine Kompression auf den Farb-Kanälen (Cb und Cr) statt. Der Helligkeitskanal bleibt unangetastet, weshalb die Strukturen eines Bildes weiterhin gut erkennbar bleiben. Die Farb-Kanäle werden reduziert, indem benachbarte Pixel zusammengefasst werden oder wegfallen. Daher handelt es sich um ein verlust-behaftetes Kompressionsverfahren.

- b) Anwendungsbereiche für Hyperspektralbilder gibt es zahlreiche, die von der sehr genauen Unterscheidung von Farben – die Menschen sehen oder auch nicht sehen können – abhängig sind. Dazu gehören bspw. die Unterscheidung von Pflanzen und die Ermittlung des Gesundheitszustandes, Mineraliensuche oder die Identifikation von Gasen in Rauchwolken. Diese Untersuchungen lassen sich auch alle am vorliegenden Bild durchführen. Durch die vielen Farben können sogar Anteile von Oberflächenarten innerhalb jedes Pixels berechnet werden („Spektrale Entmischung“). Daher ist es sehr wichtig, dass diese Farben gut erhalten bleiben. Eine weitere Mischung von Mischpixeln ist dabei unproblematisch. Beim Weglassen, Hinzufügen oder Ändern von Farbwerten allerdings, wie zum Beispiel bei der Farbreduktion, wird der Informationsgehalt des Bildes bis zur Nutzlosigkeit reduziert.

Eignung des jeweiligen Kompressionsverfahrens für das Satellitenbild vom Ätna:

Farbreduktion	Redundanzsuche	Ähnlichkeitssuche	Farbunterabtastung
ungeeignet, da großflächig Informationen verloren gehen	sehr geeignet, da keine optische Veränderung	geeignet, da bei geringem Grenzwert nur wenige Informationen verloren gehen	geeignet je nach Anwendung: Strukturen bleiben erhalten, aber bei Farben können Blöcke gleicher Farbwerte entstehen (Rasterisierungsartefakte)

c) Ergebnispräsentation

3. Aufgabenstellung

Lösung im Programm MedianCutSample_LuL.py:

- Numpy Array
- Zeile 46 bis 57
- Zeile 65, 66
- Zeile 39 bis 44
- Der Code sollte das Bild nun entsprechend komprimiert ausgeben. Eventuell ist die Kompression zu gering, um sichtbar zu sein. Dann können die Eingabewerte angepasst werden. Sollte bei jemandem der Code gar nicht funktionieren, so kann dieser mit der LuL-Version abgeglichen werden.

4. Aufgabenstellung

a) Es gibt mehrere Elemente, in denen sich der selbstgeschriebene Code (rekursiv) vom vorliegenden Code (iterativ) unterscheidet (Im MedianCutSample_LuL.py sind die Zeilenangaben +70 zu nehmen):

- Die Arrays 'oldBuckets' und 'newBuckets' sind neu. Diese beiden Arrays simulieren den Input und Output jedes Iterationsschrittes. 'oldBucket' ist immer der *Input* der aktuellen Iteration und 'newBucket' ist der *Output* der aktuellen Iteration. Dementsprechend wird am Anfang jedes Iterationsschrittes der Output der letzten Iteration in den Input der aktuellen Iteration gespeichert (Zeile: 13).
- Es existiert eine for-Schleife, die eine Zählvariable deklariert; 'currentDepth'. Dieser sogenannte 'Counter' spiegelt immer die aktuelle Iteration wider. In Datenstrukturen fungiert er häufig als Index, da so die Werte der aktuellen Iteration gespeichert werden können. Die for-Schleife zählt von 0 bis 'depth'-1.
- In diesem Code wird der Counter (currentDepth) benutzt, um eine spezifische Anzahl an Datenstrukturen zu erzeugen (Zeile: 14 + 34).

Hinweis zu Zeile 14: Dieser Befehl stellt die mathematische Formel $2^{\text{depth}-1}$ dar.

Hinweis zu Zeile 37: Dieser Befehl erzeugt currentNumberOfBuckets-mal ein mit [0] gefülltes Array innerhalb der 'newBucket' Datenstruktur.

- Die Zeilen 38-46 erzeugen für jeden „alten“ Bucket einen Trennindex, an dem der alte Bucket in zwei „neue“ Buckets getrennt wird. Dabei muss beachtet werden, dass sowohl der „bis-Wert“ für das Array exklusiv ist, als auch Buckets mit einer ungeraden Anzahl an Einträgen existieren können.

Beispiel: 2 alte Buckets mit 9 und 10 Einträgen sollen in der nächsten Generation in 4 neue Buckets geschrieben werden. Die 4 neuen Buckets haben danach die Größen 5,4,5,5.

Hinweis zu Zeile 43: Der Code nimmt nur die Länge des Buckets, der zu trennen ist, plus 1 und dann geteilt durch 2. Dies ist dann die Mitte, an der die Buckets geteilt werden.

- Normalerweise wird bei einem iterativen Code der Counter hoch gezählt (siehe for-Schleife in Zeile 11) und bei einem rekursiven Code herunter gezählt, bzw. die Tiefe des Rekursionsschrittes (siehe im MedianCutSample_LuL.py Zeile 65 + 66: split_into_buckets(..., ..., depth -1)).

Dementsprechend ist die Stoppbedingung im rekursiven Code bei 'depth == 0' (LuL Zeile: 42) und beim iterativen Code der for-Schleife bei 'range(depth)', was eine exklusive Liste erzeugt und einer Stoppbedingung von 'depth' - 1 entspricht (SuS Zeile 11). 'depth' ist hier der durch den Skriptaufruf übergebene Wert für die Anzahl an Farben. Zum Abschluss wird das Numpy Array gespeichert.

b) Unterschied: Beim implementierten Algorithmus handelt es sich um eine Rekursion, während es sich bei dem gegebenen Code um einen iterativen Algorithmus handelt.

5. Aufgabenstellung

Individuelle Lösung!

Diskussionspunkte:

- verlustfreie und -behaftete Kompressionsverfahren eignen sich für unterschiedliche Anwendungen (vgl. Nr. 2)
- Wo finden verlustbehaftete Kompressionsverfahren Anwendung?
 - jpg
 - Social-Media wegen höher Anzahl an verschickten Bildern
 - Kleine Datenpakete auf Webseiten führen zu einem schnelleren Aufbau der Inhalte

mögliche Argumente:

- Verschiedene Anwendungszusammenhänge bedeuten unterschiedliche Anforderungen an die Qualität des Bildes und somit den Bedarf an unterschiedlichen Kompressionsverfahren.
- Generell ist die Unterscheidung von verlustfreien und -behafteten Kompressionsverfahren zu beachten:
 - Verlustfreie Verfahren, z.B. zum Erkennen von Details im Bild → Redundanzsuche.
 - Durch verlustfreie Kompression können Dateien weiterhin sehr groß sein.
 - Für geringen Speicherverbrauch und geringe Anforderungen an die Bildqualität (Social-Media) eignen sich verlustbehaftete Verfahren wie die Ähnlichkeitssuche.
- Ein bestimmtes Kompressionsverfahren funktioniert nicht immer gleich gut für alle Bilder.
Bspw. anhand der Redundanzsuche: Bild mit...
 - ... vielen zusammenhängenden Flächen mit nur wenigen Details → große Einsparung an Speicherplatz.
 - ... heterogenen Flächen mit vielen Details → nur geringe Ersparnis möglich.