Redundanz entropie erklären falls nötig:

Redundanz >= 0, Minimale Anzahl an Bits pro Symbol = Entropie,

Formel Redundanz

(Entropie-Kompression-Zusammenhang:

Daten niedrige Entropie -> Struktur + Daten sind komprimierbar)

Daten sind rohe Fakten oder Symbole, die an sich keine spezifische Bedeutung haben. Information entsteht durch die Interpretation, Organisation und Strukturierung von Daten, wodurch ein sinnvoller Kontext geschaffen wird

* Informationsgehalt von Daten in Bezug auf die kürzest mögliche algorithmische Beschreibung

Datenanforderungen:

* Bsp. Tabellen (Struktur, wiederkehrende Muster)
* Farbverlaufbilder computergeneriert, Farbwechsel nicht zufällig, sondern nach einem Schema

Random Challenge: 2 Wege

* Kolmogorov Komplexität, Programm schreiben das Sequenz generiert, Programm kleiner als Zahlenfolge selbst.
* Allgemeinen Encoder/ Decoder Schreiben der alle Zufallszahlen komprimieren können muss. Programmgröße egal, da hier alle Zahlenfolgen  
  (Ansatz bewiesen unmöglich)   
  Kolmogorov Komplexität wahrscheinlich unmöglich, aber nicht bewiesen
* Challenge zeigt das es sinnvoll ist sich Gedanken über die Daten zu machen die man komprimieren will

Bilder beschreiben wieso ausgewählt (Benachbarte Pixel Ähnlichkeit / offensichtliche Redundanz, Extrembeispiele)

Kompressionsrate Wertebereich

Äbwägung der Messkriterien:

* Beispiel 1, Echtzeitanwendungen:
* Beispiel 2: Ressourcen im Internet:
* Beispiel 3, Archivierung:

Jeweils unterschiedliche Kriterien wichtig 🡪 nicht nur ein Aspekt anschauen

Algorithmen:

* Huffman Tree (Binärbaum), den separat speichern
* LZ77 über Sliding Window und jeweils Verweise, refferenzen

Kompraten:

* Durchschnittswerte, alle positiv, am besten Filtern + Huffman

RLE:

* Flagge, 151.380 Bytes auf nur 32 Bytes
* Fade 1 und 2 auch komprimierbar + Grund
* Frau, Katze natürliche Bilder schlecht, sogar neg. kompraten

Huffman:

* Nur positive Kompraten, stabil
* Schwäche: viele verschiedene Werte ähnlich häufig

Filter + Huffman:

* Durchschnittl. Am besten
* Katzenbild das von anderen kaum komp. Werden kann hier sogar 25 %, stark auch auf natürlichen Bildern
* Filtern steigert effektivität deutlich, lokale Relationen in Bildern -> weniger verschiedene Werte die unterschiedlich häufig vorkommen

LZ77:

* Schlechter als gedacht, natürliche Bilder überhaupt nicht

Filter + LZ77:

* Testbild Fade 1 deutliche Verbesserung (52 %)
* Die Filterung erzeugt durch Relativwerte wiederkehrende Muster aus einem Farbverlauf
* Nur positiv, natürliche Bilder nur etwas

Deflate:

* Ähnliche Ergebnisse zu LZ77, etwas besser Huffman entfernt aus LZ77 komp. Weitere Redundanz
* Dennoch nicht überzeugend

PNG:

* Überall verbesserung, z.t. deutlich
* Nur Katzenbild nicht

Kompressionszeiten:

* Hardwarabhängig, nur tendenzen wichtig nicht absolute Zeiten, kein perfekte Implementierung
* RLE lineare Laufzeitkomplexität wie theoretisch erwartet
* Huffman sogar schneller, obwohl er O(n log n) hat
* LZ77 deutlich langsamer, Laufzeit ist nicht nur Auflösungs sondern auch Bildinhaltabhängige, Bsp. Bild mit Katze: Laufzeit hangt direkt ¨ damit zusammen, wie weit LZ77 durch das Sliding Window iterieren muss, um einen Treffer zu finden.

Das Katzenbild enthalt deutlich unstrukturiertere Daten als das Bild Fade 2. ¨ Beim Katzenbild wird deutlich langer im Sliding Window ¨ gesucht um einen Treffer zu finden oder es wird kein Treffer gefunden

* Filtern lineare Laufzeit
* Filter + LZ77 etwas schneller, da öfters matches im Sliding Window s
* Kompressionszeit des Deflate- und PNG Algorithmus entspricht weitgehend der Laufzeit des LZ77 Algorithmus

Dekompressionszeiten:

* Revert Filter schneller als Filtern selber O(5n) vs. jetzt O(n); in Praxis sogar eher um Faktor 9 schneller
* RLE Dekompressions Laufzeit stark Datenabhängig (je komprimierter desto schneller im schnitt)
* Huffman etwas langsamer als komp.
* LZ77 Laufzeit stark Datenabhängig (je komprimierter desto schneller im schnitt)
* Dekompressionszeit des Deflate- und des PNG Algorithmus entspricht weitgehend der des LZ77 Algorithmus. LZ77 ist fur die langen Dekompressionszeiten verantwortlich.