Dr. Herbert Süße

Lehrstuhl “Digitale Bildverarbeitung”

Jena, den 18.5.2007

Übungsaufgaben im Fach “Einführung in die Informatik für Physikstudenten” zur Vorbereitung auf die Klausur

9. Gegeben sei das folgende Kodebuch $K = \{0, 01\}$. Ist dies ein I-Kode? Wäre eine kodierte Symbolkette eindeutig dekodierbar? Geben Sie ein Beispiel an für einen Kode mit variabler Kodewortlänge, der kein I-Kode ist.

10. Das fundamentale Kodierungs-Theorem gibt uns eine Schranke für die mittlere Kodewortlänge an, nämlich die Entropie plus 1 Bit. Wie muß eine Verteilung der Quellensymbole aussehen, damit diese obere Schranke erreicht bzw. näherungsweise erreicht wird?

11. Gegeben sei das Grauwertbild G

$$G = \begin{pmatrix} 2 & 1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 3 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 & 1 \end{pmatrix}$$

Berechnen Sie für das Grauwertbild G, welches nur die Grauwerte {0,1,2,3} besitzt, die Entropie, einen binären Huffman-Kode und einen binären Shannon-Fano-Kode. Für beide Kodes sind die mittleren Kodelängen anzugeben und mit der Entropie zu vergleichen.

12. Welche der folgenden Kodes können keine Huffman-Kodes sein, und zwar völlig unabhängig davon, welche Wahrscheinlichkeitsverteilung vorliegt?
a) $\{$0,10.11$\}$
b) $\{$00,01,10,110$\}$
c) $\{$01,10$\}$

13. Gegeben sei die Symbolkette ABRAKADABRA. Berechnen Sie:
a) Das Kodebuch eines Wörterbuchverfahrens, wenn man mit 8 Bit beginnt.
b) Die ganze Zahl, die als Kode für die Symbolkette nach der arithmetischen Kodierung verwendet wird.
c1) Transformieren Sie diese Symbolkette in eine äquivalente nach dem Burrows-Wheeler-Algorithmus.
c2) Wenden Sie darauf die Move-to front Methode an (Folge der Indizes ist ausreichend)

14. Gegeben sei eine unabhängige Informationsquelle S mit n Quellensymbolen a_i und den Wahrscheinlichkeiten p_i. Beweisen Sie für die Entropie H:

$$H(S^3) = 3 \cdot H(S)$$

.

15. Wichtig für Datenkompressionen sind die Statistiken höherer Ordnung. Dies versucht man oft durch eine geeignete Prädiktion zu erreichen. Im Praktikum benutzten Sie zur Prädiktion den identischen Operator, d.h. $g(i,j) = g(i,j-1)$. Überlegen Sie sich was lediglich im Unterschied dazu zu tun wäre, wenn Sie als Prädiktor $g(i,j) = g(i-1,j) \ + \ 0.5(g(i,j-1) - g(i-1,j-1))$ benutzen.

Ist dies ein "günstiger" Prädiktor, bzw. was ist die Idee für diesen Prädiktor?

16. Für die bedingten Entropien ist es wichtig, die sogenannte Bayes-Regel der Wahrscheinlichkeitsrechnung zu kennen, sie stellt den Zusammenhang zwischen Verbundwahrscheinlichkeiten und bedingten Wahrscheinlichkeiten her:

$$P(X/Y) \;=\; \frac{P(X,Y)}{P(Y)}$$

Zeigen Sie, daß

$$H(X,Y) \;=\; H(X) \;+\; H(Y/X) \;=\; H(Y) \;+\; H(X/Y)$$

für die Entropie der beiden Zufallsvariablen X, Y gilt. Benutzen Sie dabei die Bayes-Regel und nehmen Sie an, daß X die Zustände $x_1, x_2, ..., x_n$ und Y die Zustände $y_1, y_2, ..., y_m$ mit den Wahrscheinlichkeiten $P(x_i), i \;=\; 1, ..., n$ bzw. $P(y_j), j \;=\; 1, ..., m$ annehmen können. Unter einer bedingten Entropie verstehen wir dann z.B.

$$H(Y/X) \;=\; \sum_{i=1}^{n} P(x_i)H(Y/x_i)$$

17. Die sogenannte "mutual information" ist formal das gleiche wie die Transinformation, sie wird jedoch oft mit $I(X,Y)$ bezeichnet und ist demzufolge gleich

$$I(X,Y) \;=\; H(X) \;-\; H(X/Y)$$

Zeigen Sie, daß dieser Ausdruck sich durch

$$I(X,Y) \;=\; \sum_{i}\sum_{j} P(x_i, y_j) log_2 \frac{P(x_i, y_j)}{P(x_i)P(y_j)}$$

darstellen läßt (Verwenden Sie wieder die Bayes-Regel). Überlegen Sie, wieso dieser Ausdruck ein Informations-Maß für die Unabhängigkeit zweier Zufallsvariablen darstellt.