sein muß. Für so ein Paar definieren wir

Definition: *a)* Die Kovarianz eines solchen Paars (X,Y) ist

$$\begin{aligned}\operatorname{cov}(X,Y) &= E\big((X-E(X))(Y-E(Y))\big)\\ &= \sum_{i=0}^{m}\sum_{j=0}^{n}\pi_{ij}\big((x_i-E(X)\big)\big((y_j-E(Y)\big)\,.\end{aligned}$$

b) Die Korrelation von (X,Y) ist $\rho(X,Y)=\dfrac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}$.

Wir bezeichnen die beiden Zufallsvariablen X und Y entsprechend der üblichen Definition für Ereignisse als als voneinander unabhängig, falls für alle i,j gilt: $\pi_{ij}=p_iq_j$. Alsdann ist

$$\begin{aligned}\operatorname{cov}(X,Y) &= \sum_{i=0}^{m}\sum_{j=0}^{n}\pi_{ij}\big(x_i-E(X)\big)\big((y_j-E(Y)\big)\\ &= \sum_{i=0}^{m}\sum_{j=0}^{m}\Big[p_i(x_i-E(X))\big)\Big]\Big[q_j\big(y_j-E(Y)\big)\Big]\\ &= \left(\sum_{i=0}^{m}p_i\big(x_i-E(X)\big)\right)\left(\sum_{j=0}^{n}q_j\big((y_j-E(Y)\big)\big)\right)=0\,,\end{aligned}$$

denn

$$\sum_{i=0}^{m}p_i\big((x_i-E(X)\big)=\sum_{i=0}^{m}p_ix_i-\sum_{i=0}^{m}p_iE(X)=\sum_{i=0}^{m}p_ix_i-E(X)$$

verschwindet nach Definition des Erwartungswerts.

Damit haben zwei voneinander unabhängige Zufallsvariablen also Kovarianz und Korrelation null; man sagt auch, sie seien unkorreliert. Bei Bilddaten wird das im allgemeinen nicht der Fall sein; hier wird man im Gegenteil davon ausgehen, daß die Zufallsvariablen zu benachbarten Pixeln sehr stark miteinander korrelieren. Wir können beispielsweise annehmen, daß

$$Y=\rho X+Z$$

ist mit einer von X unabhängigen Zufallsvariablen Z und einer positiven reellen Zahl $\rho<1$. Dann ist

$$E(Y)=E(\rho X+Z)=\rho E(X)+E(Z)\,;$$

falls wir annehmen, daß X und Y denselben Erwartungswert haben, ist daher

$$E(Z)=(1-\rho)E(X)\,,$$

was für ein ρ nahe eins deutlich kleiner ist als $E(Z)$.

In dieser Situation ist

$$\begin{aligned}\operatorname{cov}(X,Y) &= \sum_{i=0}^{m}\sum_{j=0}^{n}\pi_{ij}\big(x_i-E(X)\big)\big((y_j-E(Y)\big)\\ &= \sum_{i=0}^{m}\sum_{j=0}^{n}\pi_{ij}\big(x_i-E(X)\big)\big(\rho x_i+z_j-\rho E(X)-E(Z)\big)\\ &= \sum_{i=0}^{m}\sum_{j=0}^{n}\pi_{ij}\big(x_i-E(X)\big)\Big[\rho\big(x_i-E(X)\big)+\big(z_j-E(Z)\big)\Big]\\ &= \sum_{i=0}^{m}\sum_{j=0}^{n}\pi_{ij}\rho\big(x_i-E(X)\big)^2\\ &\qquad+\sum_{i=0}^{m}\sum_{j=0}^{n}\pi_{ij}\big(x_i-E(X)\big)\big(z_j-E(Z)\big)\\ &= \rho\sum_{i=0}^{m}p_i\big(x_i-E(X)\big)^2+\operatorname{cov}(X,Z)=\rho\sigma_X^2\,,\end{aligned}$$

da X und Z voneinander unabhängige Zufallsvariablen sind.

Wenn wir jetzt noch annehmen, daß $\sigma_X=\sigma_Y$ ist, folgt

$$\rho(X,Y)=\frac{\rho\sigma_X^2}{\sigma_X\sigma_Y}=\rho\,,$$

wir können auf diese Weise also für beliebiges $\rho\in[0,1]$ ein Paar voneinander anhängiger Zufallsvariablen mir Korrelation ρ erzeugen.

Besser noch: Wann immer zwei Zufallsvariablen mit gleicher Standardabweichung Korrelation ρ haben, sind wir immer im obigen Fall, denn definieren wir eine neue Zufallsvariable Z durch $Z = Y - \rho X$, so ist $E(Z) = E(Y) - \rho E(X)$, und das Paar (X, Z) ist unkorreliert, da

$$\begin{aligned}
\operatorname{cov}(X,Z) &= \sum_{i=0}^{m}\sum_{j=0}^{n}\bigl(x_i - E(X)\bigr)\bigl(z_j - E(Z)\bigr)\\
&= \sum_{i=0}^{m}\sum_{j=0}^{n}\bigl(x_i - E(X)\bigr)\Bigl[\bigl(y_j - E(Y)\bigr) - \rho\bigl(x_i - E(X)\bigr)\Bigr]\\
&= \sum_{i=0}^{m}\sum_{j=0}^{n}\bigl(x_i - E(X)\bigr)\bigl(y_j - E(Y)\bigr)\\
&\qquad - \rho\sum_{i=0}^{m}\sum_{j=0}^{n}\bigl(x_i - E(X)\bigr)^2\\
&= \operatorname{cov}(X,Y) - \rho\sigma_X^2 = \rho\sigma_X\sigma_Y - \rho\sigma_X^2 = 0\,.
\end{aligned}$$

f) Modellierung von Bild- und Audiodaten durch Zufallsvariable

Unser Modell für Audio- und Bilddaten ist nun folgendes: Die Daten werden gegeben durch einen Vektor *bzw.* eine Matrix voneinander abhängiger Zufallsvariablen mit gemeinsamem Erwartungswert μ und Standardabweichung σ; zwei benachbarte Zufallsvariablen sollen jeweils Korrelation ρ haben. Diese Zahl ρ bezeichnen wir als die *Autokorrelation* des jeweiligen Signals. Wir schätzen sie anhand eines gegebenen Bilds oder Musikstücks als Korrelation zwischen benachbarten Daten.

Die Länge des Vektors oder der Matrix könnte beispielsweise mit der Länge des Musikstücks oder der Bildgröße übereinstimmen; meist betrachtet man aber des Rechenaufwands wegen nur Teilstücke. In der Bildverarbeitung etwa gehen die gängigen Standards von 8×8-Matrizen aus.

Da der Begriff der Autokorrelation das wohl am schwersten verständliche der hier eingeführten statistischen Konzepte ist, sind die sechs

Testbilder in Richtung fallender Autokorrelation geordnet: Die höchste Autokorrelation hat mit $\rho = 0{,}98$ das Bild „Peppers“, wo die recht homogenen Flächen der Paprikaschoten dafür sorgen, daß sich ein Pixel nur selten von seinen Nachbarn unterscheidet; auch „Lenna“ und „Sailboat“ werden von flächigen Strukturen dominiert. Bei „Stream“ kommen in stärkerem Maße feine Verästelungen von Bäumen und Büschen ins Spiel, so daß die Autokorrelation auf 0,94 absinkt, und bei „Tifffany“ und „Baboon“ schließlich sorgen die vielen Haare für feine Details, die die Autokorrelation auf 0,87 *bzw.* 0,86 herunterdrücken.

Um zu sehen, wie sich die Autokorrelation zur Komprimierung der Daten ausnutzen läßt, betrachten wir der Einfachheit halber zunächst nur eine Folge $X_1, \ldots, X_n$ von Zufallsvariablen; der gemeinsame Erwartungswert sei μ, die gemeinsame Standardabweichung σ, und die Korrelation zwischen X_i und X_{i+1} sei jeweils ρ.

Nach dem oben Gesagten gibt es dann für jede der Zufallsvariablen X_i mit $i < n$ eine davon unabhängige Zufallsvariable Z_i, so daß

$$X_{i+1} = \rho X_i + Z_i$$

ist; entsprechend ist für $i < n - 1$

$$X_{i+2} = \rho X_{i+1} + Z_{i+1} = \rho^2 X_i + \rho Z_i + Z_{i+1}$$

usw.; wenn wir zusätzlich annehmen, daß alle Z_i voneinander unabhängig sind, ist also $\rho(X_i, X_{i+2}) = \rho^2$ und allgemein

$$\rho(X_i, X_j) = \rho^{|i-j|}\,.$$

In der Signalverarbeitung spricht man bei einer solchen Folge von Zufallsvariablen von einem *autoregressive Prozeß;* der hier betrachtete allereinfachste Fall, bei dem alle Korrelationen nur von der Korrelation zwischen zwei benachbarten Zufallsvariablen abhängen, wird als AR(1)-Modell bezeichnet; in der Sprechweise der Wahrscheinlichkeitstheorie handelt es sich hier um spezielle sogenannte MARKOV-Ketten.

Der russische Mathematiker ANDREĬ ANDREEVICH MARKOV (1856–1922) studierte in Sankt Petersburg, wo er später auch Professor wurde. Er beschäftigte sich zunächst hauptsächlich mit Zahlentheorie und Analysis; erst später kommen die wahrscheinlichkeitstheoretischen Arbeiten, für die er heute vor allem bekannt ist. MARKOV-Ketten sind Prozesse ohne Erinnerung, in denen das zukünftige Verhalten nur vom augenblicklichen Zustand abhängt, nicht aber von der Geschichte des Systems. Damit sind sie gerade hier bei Bilddaten nur eine unvollkommene Approximation an die Realität, aber dennoch sehr nützlich.

Falls wir bei einer solchen Folge von Zufallsvariablen die Werte von $X_1, \dots, X_n$ nacheinander übertragen, übertragen wir zuerst den Wert von X_1, dann mit X_2 noch einmal zu $100 \times \rho$ % denselben Wert, mit X_3 dasselbe noch einmal zu $100 \times \rho^2$ %, *usw.*

Eine offensichtliche Alternative hierzu wäre, nur den Wert von X_1 zu übertragen und ansonsten nur die Werte der Z_i. Eine ähnliche Vorgehensweise wird tatsächlich gelegentlich angewandt, allerdings macht man es sich dann noch einfacher und überträgt nur die *Differenzen,* also

$$X_1, X_2 - X_1, \dots, X_n - X_{n-1}\,.$$

Der Nachteil dieses Verfahrens ist, daß sowohl diese Differenzen als auch die Z_i von Zeit zu Zeit sehr groß werden *müssen,* da es in fast jedem Bild oder Musikstück gelegentliche aprupte Veränderungen gibt.

g) Komprimierung durch Dekorrelation

Die Idee hinter allen Komprimierungsverfahren, die auf Transformationen beruhen, ist es, anstelle der Zufallsvariablen X_i geeignete Linearkombinationen

$$Y_i = \sum_{j=1}^{n} \alpha_{ij} X_j$$

zu betrachten, wobei (α_{ij}) eine *invertierbare* $n \times n$-Matrix ist, so daß sich auch umgekehrt die X_i wieder aus den Y_j rekonstruieren lassen. Diese Matrix wird so gewählt, daß die neuen Variablen möglichst unkorreliert sind und daß man die Größe der neuen Variablen möglichst gut abschätzen kann.

Letzterer Aspekt erfordert statistische Betrachtungen, auf die wir hier verzichten wollen; die Dekorrelation der Zufallsvariablen aber führt uns geradewegs zu Eigenvektoren symmetrischer Matrizen:

Wir definieren für eine Folge von Zufallsvariablen deren *Korrelationsmatrix*

$$\mathrm{Kor}(X_1, \dots, X_n) \in \mathbb{R}^{n \times n}$$

dadurch, daß der Eintrag an der Stelle ij dieser Matrizen jeweils die Korrelation $\rho(X_i, X_j)$ sein soll.

Das Ideal, auf das wir hinarbeiten, sind Zufallsvariablen, deren Korrelationsmatrix eine Diagonalmatrix ist, denn dann sind je zwei verschiedene Variablen unkorreliert.

Da die Korrelationsmatrix eine symmetrische Matrix ist, gibt es eine Orthonormalbasis des $\mathbb{R}^n$ aus reellen Eigenvektoren, bezüglich derer sie Diagonalgestalt hat; die Vektoren dieser Orthonormalbasis seien

$$\vec{b}_1 = \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{1n} \end{pmatrix}, \dots, \vec{b}_n = \begin{pmatrix} \alpha_{n1} \\ \vdots \\ \alpha_{nn} \end{pmatrix}.$$

Wir definieren die neuen Zufallsvariablen durch

$$Y_i = \sum_{j=1}^{n} \alpha_{ij} X_j\,;$$

dann ist

$$\begin{aligned}
\rho(Y_i, Y_k) = \vec{v}_{Y_i} \cdot \vec{v}_{Y_k} &= \left(\sum_{j=1}^{n} \alpha_{ij} \vec{v}_{X_j}\right) \cdot \left(\sum_{\ell=1}^{n} \alpha_{k\ell} \vec{v}_{X_\ell}\right) \\
&= \sum_{j=1}^{n} \sum_{\ell=1}^{n} \alpha_{ij}\, \vec{v}_{X_j} \cdot \vec{v}_{X_\ell}\, \alpha_{k\ell} = \sum_{j=1}^{n} \sum_{\ell=1}^{n} \alpha_{ij}\, \rho(X_j, X_\ell)\, \alpha_{k\ell} \\
&= \sum_{\ell=1}^{n} \left(\sum_{j=1}^{n} \alpha_{ij}\, \rho(X_j, X_\ell)\right) \alpha_{k\ell}\,.
\end{aligned}$$

Der Inhalt der großen Klammer ist offensichtlich der Eintrag an der Stelle $i\ell$ der Produktmatrix $A \cdot \mathrm{Kor}(X_1, \dots, X_n)$, wobei $A = (\alpha_{ij})$ die

Matrix der Koeffizienten α_{ij} ist, und die Summation über ℓ macht daraus den Eintrag an der Stelle ik des Produkts mit ${}^tA$. Insgesamt haben wir also gezeigt, daß

$$\mathrm{Kor}(Y_1,\dots,Y_n) = A\cdot\mathrm{Kor}(X_1,\dots,X_n)\cdot{}^tA$$

ist. Nun müssen wir nur noch beachten, daß die Spaltenvektoren der Matrix A als die Vektoren einer Orthonormalbasis des $\mathbb{R}^n$ gewählt waren; der Eintrag an der Stelle ij der Matrix $A\,{}^tA$ ist also das Standardskalarprodukt des i-ten und des j-ten Vektors aus einer Orthonormalbasis und somit null für $i\neq j$ und eins für $i=j$. Daher ist $A\cdot{}^tA=E$, also ${}^tA^{-1}$ und somit auch

$$\mathrm{Kor}(Y_1,\dots,Y_n) = A\cdot\mathrm{Kor}(X_1,\dots,X_n)\cdot A^{-1}\,.$$

Damit ist $\mathrm{Kor}(Y_1,\dots,Y_n)$ eine Diagonalmatrix, denn für jede Matrix $B\in\mathbb{R}^{n\times n}$ ist ABA^{-1} die Matrix B bezüglich der Basis aus den Spaltenvektoren vom A. Diese Basis besteht hier aber aus lauter Eigenvektoren der Korrelationsmatrix, die transformierte Matrix ist also eine Diagonalmatrix.

Unter den Annahmen unseres statistischen Modells können wir also jede Folge von Zufallsvariablen durch eine lineare Transformation in eine Folge unkorrelierter Zufallsvariablen überführen. Diese Transformation bezeichnet man, obwohl sie zuerst von HOTELLING vorgeschlagen wurde, als KARHUNEN-LOÈVE-Transformation.

HAROLD HOTELLING (1895–1973) war ein amerikanischer Statistiker und Ökonom; er lehrte an der Columbia University und der University of North Carolina. In einer 1933 veröffentlichten Arbeit im *Journal of Educational Psychology* schlug er erstmalig diese Transformation vor, die von Statistikern heute in Anlehnung an den Titel seiner Arbeit meist als *Hauptkomponentenanalyse* bezeichnet wird. In Europa erschien die Transformation fast gleichzeitig um 1947 *bzw.* 1948 in wahrscheinlichkeitstheoretischen Arbeiten des Finnen KARI KARHUNEN (*1915) und des Franzosen MICHEL LOÈVE (1907–1979), nach denen sie in der technischen Literatur benannt wird.

Die Matrix A der linearen Transformation hängt nur von ρ ab und kann daher für gängige Werte von ρ vorberechnet werden; die KARHUNEN-LOÈVE-Transformation ist also einfach die Multiplikation mit einer bekannten Matrix.

h) Die diskrete Cosinus-Transformation

Für die Multiplikation zweier $n\times n$-Matrizen benötigt man allerdings n^3 Multiplikationen und noch einmal $n^2(n-1)$ Additionen; der Aufwand steigt mit großem n also sehr stark an. In der Praxis gibt man sich daher mit einen Kompromiß zufrieden und zerlegt eine Folge von Zufallszahlen in kurze Teilsequenzen, die bei eindimensionalen Folgen typischerweise die Länge 8 haben; dies ist beispielsweise der Standard bei Musik-CDs.

Allerdings wird weder bei Musik-CDs noch sonst irgendwo die KARHUNEN-LOÈVE-Transformation wirklich angewandt. Der Grund liegt an der Struktur der Eigenvektoren der Korrelationsmatrix: Betrachten wir etwa als typisches Beispiel den $n=8$; dann haben wir die Matrix

$$\mathrm{Cov}(X_1,\dots,X_8)=\begin{pmatrix}1&\rho&\rho^2&\rho^3&\rho^4&\rho^5&\rho^6&\rho^7\\ \rho&1&\rho&\rho^2&\rho^3&\rho^4&\rho^5&\rho^6\\ \rho^2&\rho&1&\rho&\rho^2&\rho^3&\rho^4&\rho^5\\ \rho^3&\rho^2&\rho&1&\rho&\rho^2&\rho^3&\rho^4\\ \rho^4&\rho^3&\rho^2&\rho&1&\rho&\rho^2&\rho^3\\ \rho^5&\rho^4&\rho^3&\rho^2&\rho&1&\rho&\rho^2\\ \rho^6&\rho^5&\rho^4&\rho^3&\rho^2&\rho&1&\rho\\ \rho^7&\rho^6&\rho^5&\rho^4&\rho^3&\rho^2&\rho&1\end{pmatrix}.$$

Ihre Eigenwerte für $\rho=0{,}95$ können zumindest näherungsweise berechnet werden, und auch die Eigenvektoren lassen sich bestimmen. Diese sollen hier jedoch nicht numerisch angegeben werden: Eine Folge von acht reellen Zahlen ist schließlich im allgemeinen eher unanschaulich. Stattdessen sind in den Abbildungen 59 bis 66 die Eigenvektoren *graphisch* dargestellt, wobei einem Vektor

$$(a_1,\dots,a_8)\in\mathbb{R}^8$$

die acht Striche vom Punkt $(i, 0)$ bis (i, a_i) in der Ebenen entsprechen sollen. Zusätzlich ist in jedes dieser Diagramme noch eine der Kurven

$$y = \cos\left(\frac{(2x-1)(j-1)\pi}{16}\right)$$

für $j = 1, \ldots, 8$ eingezeichnet; wie man sieht, lassen sich die Komponenten der Eigenvektoren sehr gut durch diese Cosinuswerte annähern. Dies gilt nicht nur für den speziellen Wert $\rho = 0{,}95$, sondern für jeden Wert von ρ, der hinreichend nahe bei eins liegt.

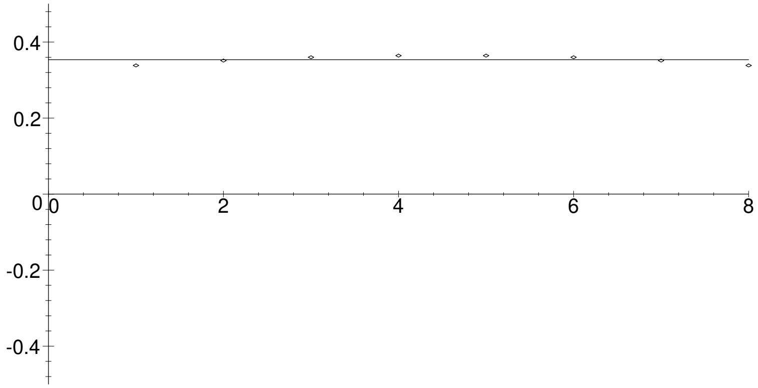

Abb. 59: Der erste Eigenvektor der Korrelationsmatrix

Aus diesem Grund arbeitet man in der Praxis lieber mit den Cosinuswerten; der Basiswechsel hin zur Basis aus den Cosinusvektoren bezeichnet man als *diskrete Cosinustransformation.* Ihr Hauptvorteil gegenüber der KARHUNEN-LOÈVE-Transformation ist, daß sie durch einen schnellen Algorithmus berechnet werden kann, der anstelle des Aufwands n^3 für n Matrixmultiplikationen nur den Aufwand $n^2 \log n$ hat. Für Einzelheiten sei auf die Vorlesung *Numerik I* verwiesen.

Die diskrete Cosinustransformation ist Teil fast aller gängiger Normen zur Bildkomprimierung: Sowohl der JPEG-Standard für Photographien, die Standards MPEG 1 und 2 für digitale (Unterhaltungs-)Videos als

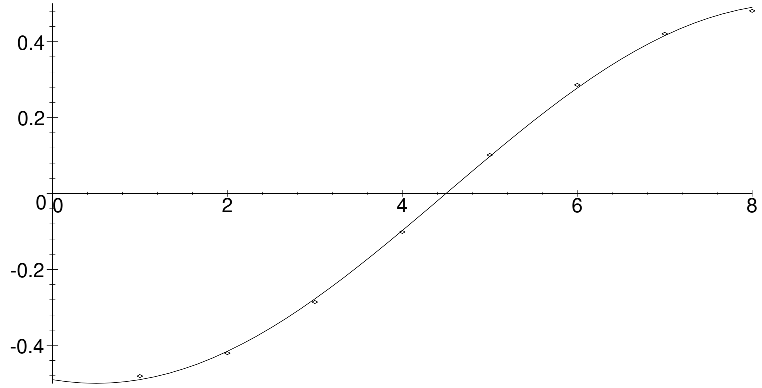

Abb. 60: Der zweite Eigenvektor der Korrelationsmatrix

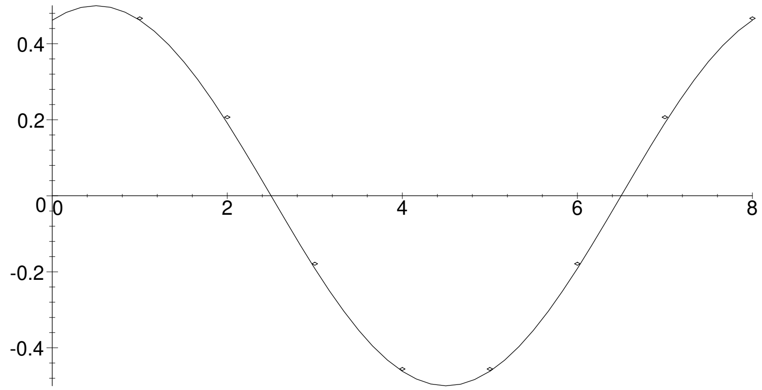

Abb. 61: Der dritte Eigenvektor der Korrelationsmatrix

auch der Standard CCITT H.261 für Videokonferenzen enthalten (neben anderen Bestandteilen) jeweils eine diskrete Cosinustransformation.

Die Transformation allein ist natürlich noch keine Komprimierung: Schließlich haben wir nur einen Vektor in einer anderen Basis hingeschrieben, und die Anzahl der reellen Zahlen, die man zur Beschreibung eines solchen Vektors benötigt, ist unabhängig von der Basis. Der

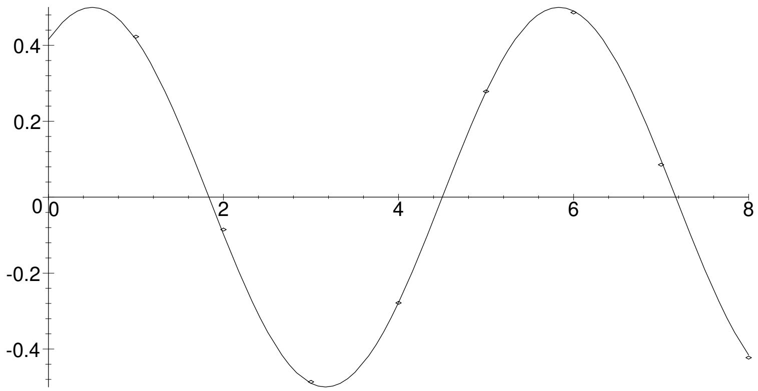

Abb. 62: Der vierte Eigenvektor der Korrelationsmatrix

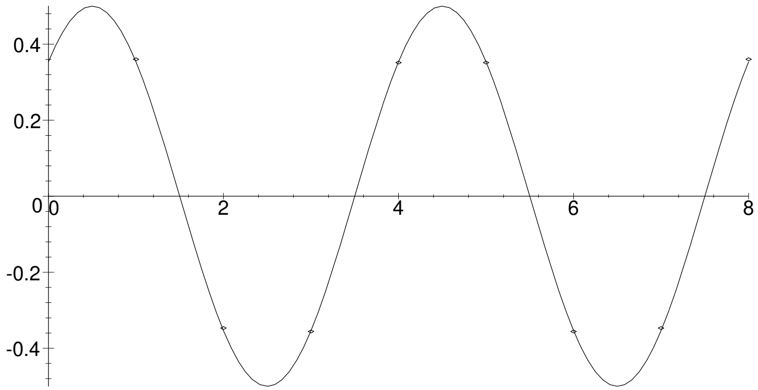

Abb. 63: Der fünfte Eigenvektor der Korrelationsmatrix

wesentliche Vorteil der neuen Basis ist, daß man statistisch recht gute Aussagen über die Größe der Komponenten machen können. Hier wollen wir auf exakte statistische Berechnungen verzichten und stattdessen informell diskutieren, warum dies der Fall sein könnte.

Wie die Abbildungen der Basisvektoren zur KARHUNEN-LOÈVE-Transformationen und die Formeln für die Basisvektoren zur diskreten

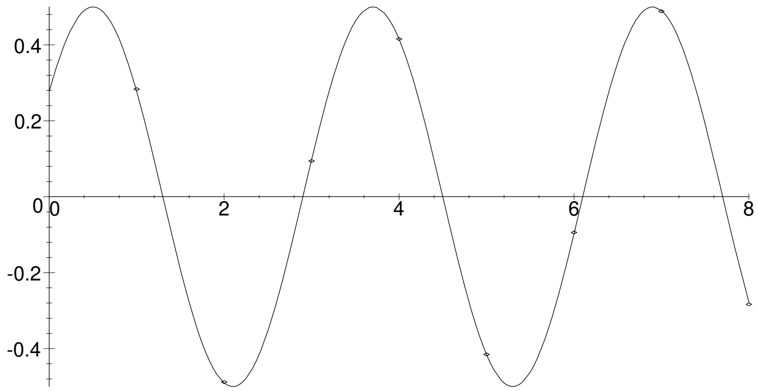

Abb. 64: Der sechste Eigenvektor der Korrelationsmatrix

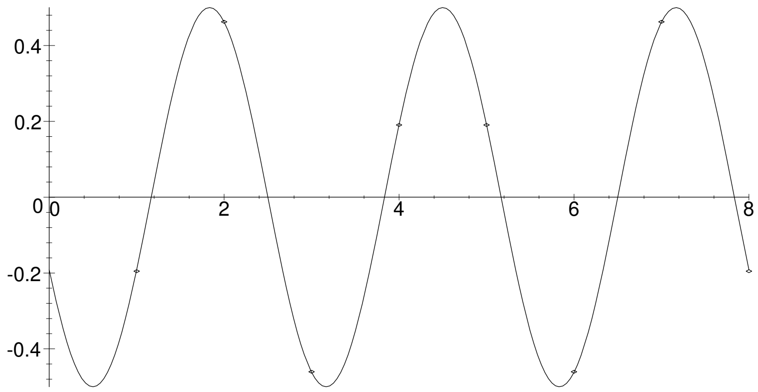

Abb. 65: Der siebte Eigenvektor der Korrelationsmatrix

Cosinustransformation zeigen, werden die Basisvektoren, wenn man sie in der hier angegebenen Reihenfolge betrachtet, immer hochfrequenter. In einem hinreichend fein abgetasteten Bild oder Audiosignal erwarten wir, daß hochfrequente Schwankungen keine große Rolle spielen und somit die entsprechenden Basisvektoren nur kleine Koeffizienten haben oder in vielen Fällen sogar gleich gar nicht auftreten. Dementsprechend

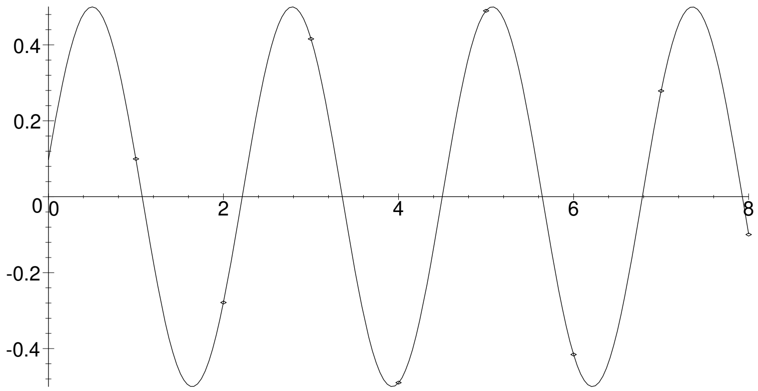

Abb. 66: Der achte Eigenvektor der Korrelationsmatrix

genügt es, für die Übertragung dieser Koeffizienten nur wenige Bits bereitzustellen; bei nur geringen Abstrichen an die Qualität kann man auf gewisse Koeffizienten sogar ganz verzichten.

Ein Kompressionsverfahren wird daher, je nach Anspruch an die Qualität, entweder alle Koeffizienten des Signals in der neuen Basis übertragen und durch eine geeignete Darstellung der Daten dafür sorgen, daß Folgen von Nullen nur wenig Platz benötigen, oder aber es wird nur eine Auswahl der Koeffizienten übertragen und auch für diese jeweils festlegen, wie viele Bit dafür in Anspruch genommen werden. Diese Anzahl wird umso geringer sein, je höher die Frequenz des jeweiligen Basisvektors ist; bei einigen Verfahren wie etwa JPEG können die Anzahlen auch variablel in Abhängigkeit von einer Qualitätszahl gewählt werden.

Zum Schluß sei noch ganz kurz erwähnt, daß die KARHUNEN-LOÈVE-Transformation und damit (mit ganz geringen Abstrichen) auch die diskrete Cosinustransformation zwar die Korrelationsmatrix in optimaler Weise diagonalisieren, daß aber daraus nicht folgt, daß sie auch optimale Kompressionsverfahren liefern: Außer der Kovarianz gibt es noch weitere Quellen für Redundanz eines Bildes.

Ein gewisser Nachteil der Cosinustransformation ist außerdem, daß man

für aprupte Übergänge, wie sie etwa bei Kanten immer wieder einmal auftauchen, die hochfrequenten Basisvektoren braucht, die dann aber nicht nur die Kante selbst beeinflussen, sondern das gesamte Quadrat, auf das die Transformation angewandt wird.

Eine bessere Möglichkeit wäre es daher, wenn man anstelle von Cosinusfunktionen Funktionen verwenden könnte, die sowohl im Zeit- als auch im Frequenzbereich lokalisiert sind. Solche Funktionen gibt es in der Tat, etwa die sogenannten *Wavelets.* Hierbei handelt es sich um schnell abklingende Wellen, und neuere Arbeiten deuten darauf hin, daß diese für gewisse Bildmodelle (die im Gegensatz zum hier betrachteten nicht mit Wahrscheinlichkeiten arbeiten) nicht zu weit vom Optimum entfernt sein sollten. Im Rahmen dieser Vorlesung ist es jedoch zeitlich weder möglich, auf diese Modelle einzugehen, noch ist an eine genauere Behandlung von Wavelets zu denken.

Einen allgemein verständlichen Überblick über Wavelets findet man etwa bei

BARBARA BURKE HUBBARD: Wavelets: Die Mathematik der kleinen Wellen, *Birkhäuser* 1997;

das zitierte Optimalitätsresultat ist beschrieben im Vortrag

STÉPHANE MALLAT: Applied Mathematics meets signal processing

auf dem Internationalen Mathematikerkongress 1998 in Berlin, nachzulesen in Band I der Proceedings, S. 319–338, oder unter http://www.mathematik.uni-bielefeld.de/documenta/xvol-icm/00/Mallat.MAN.html.

Eine für Technische Informatiker gut geeignete fundierte Einführung in diesen Themenkreis ist

STÉPHANE MALLAT: A wavelet tour of signal processing, *Academic Press, 1998.*

E N D E

S C H Ö N E F E R I E N !