

Multimedia Retrieval im WS 2011/2012

4. Feature-Transformationsverfahren

4.1 Diskrete Fourier-Transformation

Prof. Dr.-Ing. Marcin Grzegorzek
Juniorprofessur für Mustererkennung
Institut für Bildinformatik im Department ETI
Fakultät IV der Universität Siegen

7. und 8. November 2011



Inhalte und Termine

1. Einführung

1.1 Grundlegende Begriffe

1.2 Suche in einem MMDBS

1.3 MMDBMS-Anwendungen

11.10.2011

2. Prinzipien des Information Retrieval

2.1 Einführung

2.2 Information-Retrieval-Modelle

2.3 Relevance Feedback

2.4 Bewertung von Retrieval-Systemen

17.10.2011

2.5 Nutzerprofile

3. Prinzipien des Multimedia Retrieval

3.1 Besonderheiten der Verwaltung und des Retrievals

3.2 Ablauf des Multimedia-Information-Retrievals

3.3 Daten eines Multimedia-Retrieval-Systems

24.10.2011

3.4 Feature

3.5 Eignung verschiedener Retrieval-Modelle

3.6 Multimedia-Ähnlichkeitsmodell

25.10.2011

4. Feature-Transformationsverfahren

4.1 Diskrete Fourier-Transformation

08.11.2011

4.2 Diskrete Wavelet-Transformation

4.3 Karhunen-Loeve-Transformation

4.4 Latent Semantic Indexing und Singulärwertzerlegung

5. Distanzfunktionen

- 5.1 Eigenschaften und Klassifikation
- 5.2 Distanzfunktionen auf Punkten
- 5.3 Distanzfunktionen auf Binärdaten
- 5.4 Distanzfunktionen auf Sequenzen
- 5.5 Distanzfunktionen auf allgemeinen Mengen

6. Ähnlichkeitsmaße

- 6.1 Einführung
- 6.2 Distanz versus Ähnlichkeit
- 6.3 Grenzen von Ähnlichkeitsmaßen
- 6.4 Konkrete Ähnlichkeitsmaße
- 6.5 Aggregation von Ähnlichkeitswerten
- 6.6 Umwandlung von Distanzen in Ähnlichkeitswerte und Normierung
- 6.7 Partielle Ähnlichkeit

7. Effiziente Algorithmen und Datenstrukturen

7.1 Hochdimensionale Indexstrukturen

7.2 Algorithmen zur Aggregation von Ähnlichkeitswerten

8. Anfragebehandlung

8.1 Einführung

8.2 Konzepte der Anfragebehandlung

8.3 Datenbankmodell

8.4 Sprachen

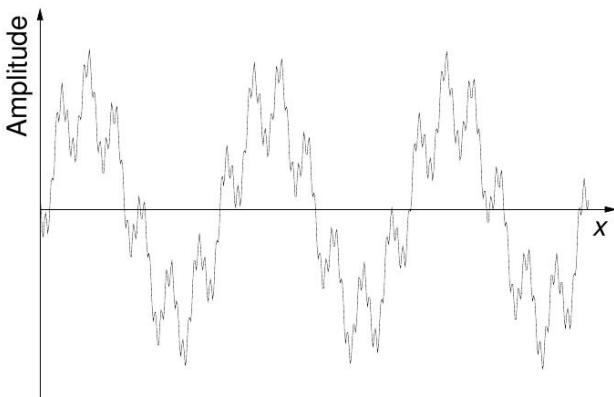
9. Zusammenfassung

DFT - Allgemeines

- ▶ benannt nach franz. Mathematiker Jean Baptiste Joseph Fourier
- ▶ periodische Funktion als Summe von Sinus- und Kosinusfunktionen darstellbar
- ▶ Darstellung im Ortsbereich versus Darstellung im Frequenzbereich
- ▶ äquivalente Darstellungen (ineinander verlustfrei überführbar)
- ▶ Darstellung im Frequenzbereich
 - ▶ niedrige Frequenzen für groben Funktionsverlauf
 - ▶ hohe Frequenzen für Detailinformationen (etwa abrupte Funktionswertänderungen)

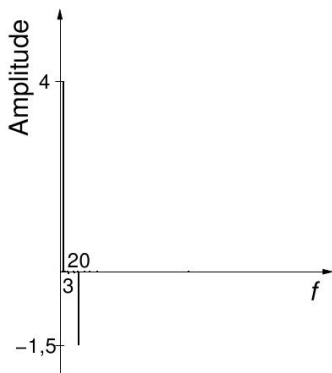
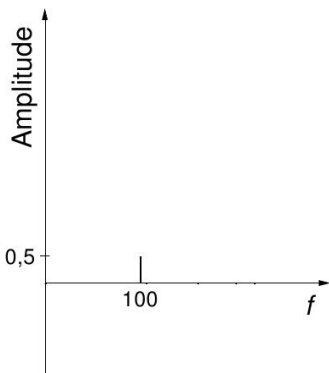
DFT - Erstes Beispiel

gegeben: $f(x) = 4 \sin 3x - 3/2 \sin 20x + 1/2 \cos 100x$



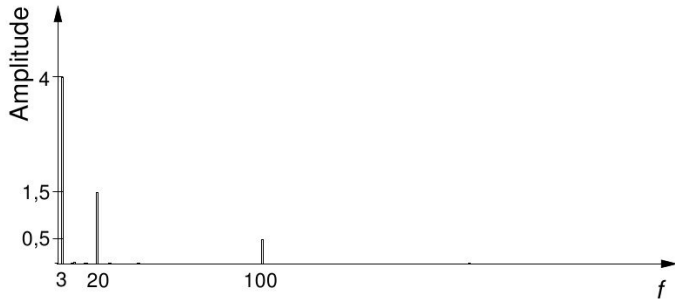
DFT - Erstes Beispiel

Anteil Kosinus- und Sinusschwingungen:



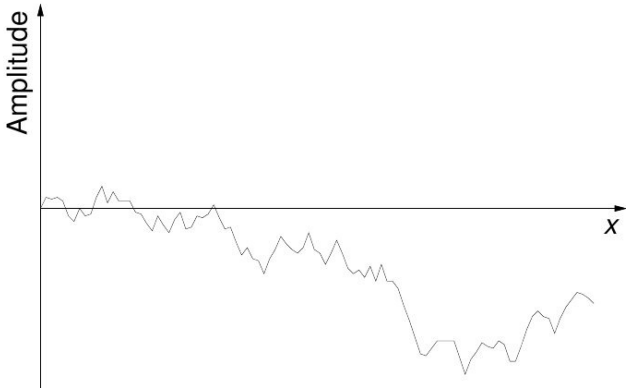
DFT - Erstes Beispiel

Frequenzspektrum:



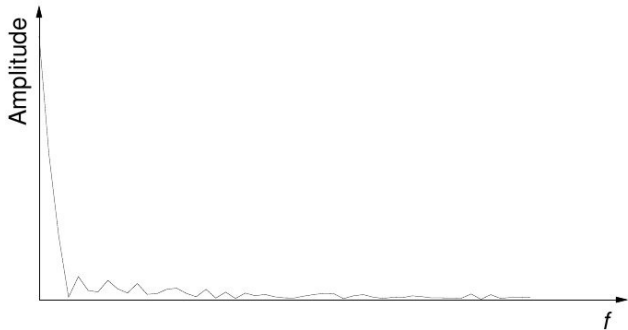
DFT - Zweites Beispiel

zufällig erzeugte Funktion:



DFT - Zweites Beispiel

Frequenzspektrum und Kompaktheit:



DFT - Gedankenexperiment

Darstellung im Orts- und Frequenzbereich:

- ▶ Lautsprecher als Tonquelle erzeugt konstantes Geräusch
→ Orts- bzw. Zeitbereich
- ▶ Eine bestimmte Klaviersaite beginnt zu schwingen, wenn ihre Frequenz im Geräusch enthalten ist. Wenn man die Stärke aller Saitenschwingungen aufzeichnet, erhält man die Darstellung des Signals im Frequenzbereich.
→ Frequenzbereich
- ▶ Rekonstruktion des Geräusches durch Anregen der entsprechenden Klaviersaiten

DFT - Anwendungen

Feature-Normalisierung

- ▶ Unterdrückung von Störeinflüssen innerhalb der Medienobjekte
- ▶ Problem: Trennung Nutz- und Stördaten
- ▶ Trennung manchmal im Frequenzbereich möglich

Feature-Erkennung

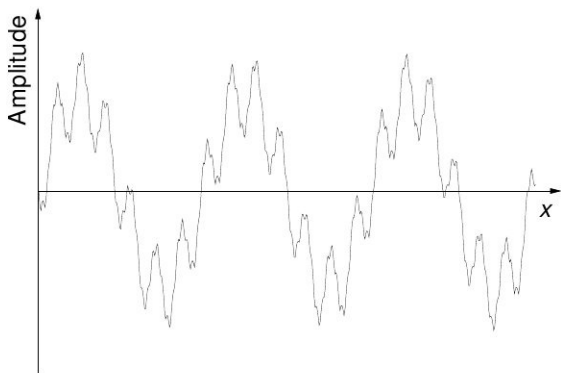
- ▶ Korrespondenz zwischen zu extrahierenden Eigenschaften und Frequenzen möglich

Feature-Aufbereitung

- ▶ Kompaktheit im Frequenzbereich → Minimalität
- ▶ Orthogonalität der Fourier-Koeffizienten

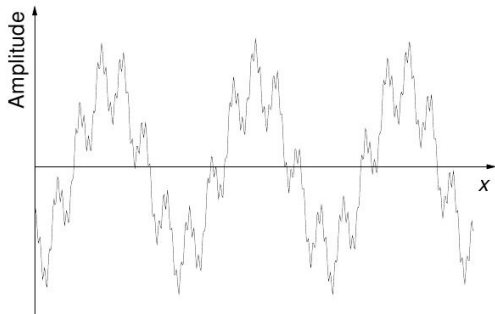
DFT - Beispiel zur Feature-Normalisierung

Beispiel: Entfernung von Frequenz $f=100$



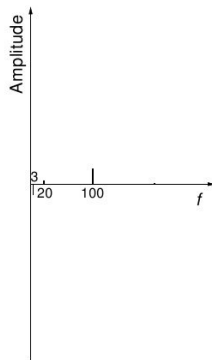
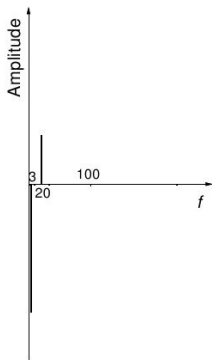
DFT - Feature-Erkennung, Translationsinvarianz

Funktion $f(x) = 4 \sin 3x - 3/2 \sin 20x + 1/2 \cos 100x$ um $\pi/3$ verschoben:



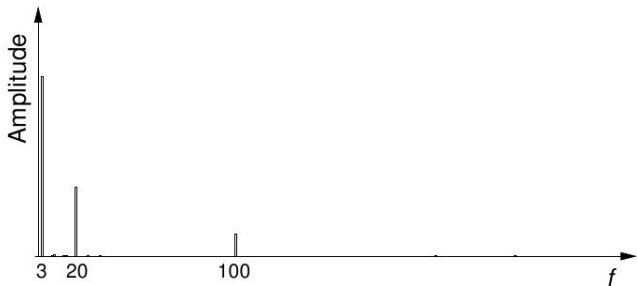
DFT - Feature-Erkennung, Translationsinvarianz

Anteil Kosinus- und Sinusschwingungen:



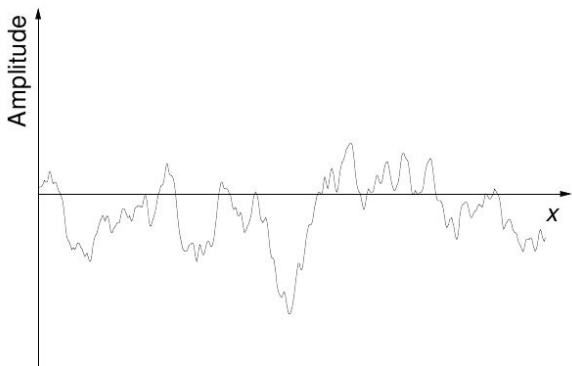
DFT - Feature-Erkennung mit Translationsinvarianz

Frequenzspektrum:



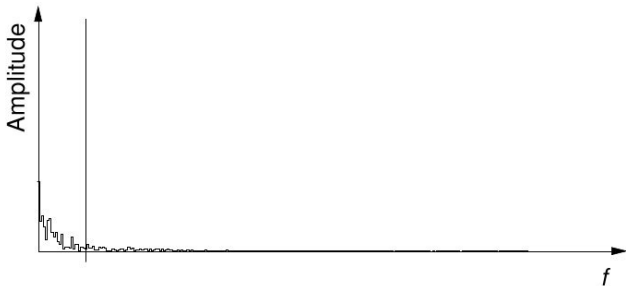
DFT - Feature-Aufbereitung, Minimalität

Ausgangsfunktion:



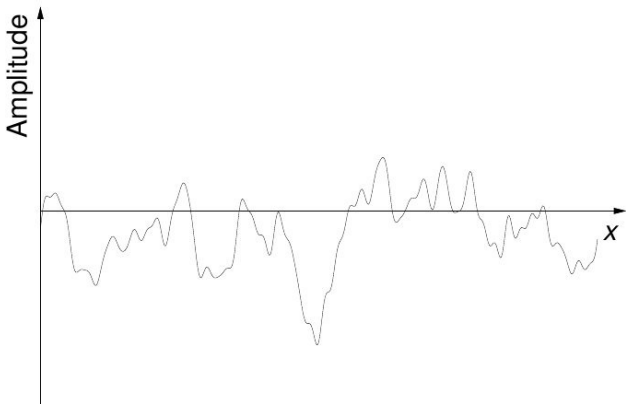
DFT - Feature-Aufbereitung, Minimalität

Frequenzspektrum und Abschneiden hoher Frequenzen:



DFT - Feature-Aufbereitung, Minimalität

approximierte, d.h. minimierte Funktion:



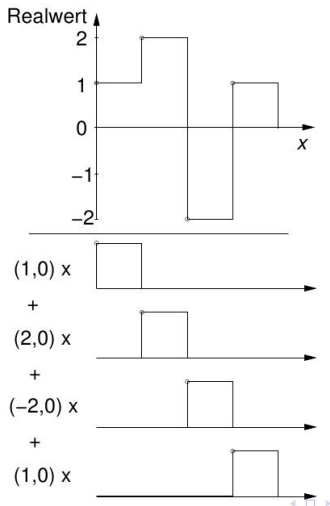
DFT - Berechnung

Erinnerung lineare Algebra und komplexe Zahlen

- diskrete Funktion $f(x)$ als Vektor des **komplexen Vektorraums**:
 $f(x) \in D_n^{\mathbb{C}}$
(Vektorraum $D_n^{\mathbb{C}}$ hat n kanonische Basisvektoren)
- Konjugation einer komplexen Zahl: $\overline{(x, y)} = (x, -y)$
- inneres Produkt für $f, g \in D_n^{\mathbb{C}}$: $\langle f, g \rangle = \sum_{x=0}^{n-1} f(x) * \overline{g(x)}$
- Entwicklungsformel für Orthonormalbasis (v_0, \dots, v_{n-1}) :

$$\forall v \in D_n^{\mathbb{C}} : v = \sum_{i=0}^{n-1} \langle v, v_i \rangle v_i$$

DFT - Beispiel im Vektorraum $D_4^{\mathbb{C}}$



DFT - Fourier-Basis

$$\begin{aligned} e_j(x) &= \frac{1}{\sqrt{n}} e^{\frac{i2\pi jx}{n}} \\ &= \frac{1}{\sqrt{n}} \cos \frac{2\pi jx}{n} + \frac{1}{\sqrt{n}} i \sin \frac{2\pi jx}{n} \end{aligned}$$

mit $j = 0, \dots, n-1$, $x = 0, \dots, n-1$ und $i = \sqrt{-1}$

Orthonormalität:

$$\|e_j(x)\| = \sqrt{\langle e_j(x), e_j(x) \rangle} = 1 \quad \text{für } j = 0, 1, \dots, n-1$$

$$\langle e_j(x), e_k(x) \rangle = 0 \quad \text{für } j, k = 0, 1, \dots, n-1 \text{ und } j \neq k$$

DFT - Fourier-Koeffizienten

- aufgrund Orthonormalität der Fourier-Basis Berechnung der Fourier-Koeffizienten mittels **innerem Produkt**:

$$(re_j, im_j) = \langle f_n(x), e_j(x) \rangle$$

—→ Transformation als einfache Multiplikation mit DFT-Matrix möglich

- Transformation entspricht Rotation im komplexen, hochdimensionalen Raum
- Ergebnis: komplexe Fourier-Koeffizienten
 - Realteil für Kosinusamplituden
 - Imaginärteil für Sinusamplituden

DFT - Transformationsformel

$$\begin{aligned}F_n(j) &= \langle f_n(x), e_j(x) \rangle \\&= \sum_{x=0}^{n-1} f(x) \cdot \overline{e_j(x)} \\&= \sum_{x=0}^{n-1} f(x) \cdot \overline{\frac{1}{\sqrt{n}} e^{\frac{i2\pi jx}{n}}} \\&= \frac{1}{\sqrt{n}} \sum_{x=0}^{n-1} f(x) \cdot e^{-\frac{i2\pi jx}{n}} \\&= \left(\frac{1}{\sqrt{n}} \sum_{x=0}^{n-1} f(x) \cos \frac{2\pi jx}{n}, -\frac{1}{\sqrt{n}} \sum_{x=0}^{n-1} f(x) \sin \frac{2\pi jx}{n} \right)\end{aligned}$$

DFT - Rücktransformation DFT^{-1}

$$\begin{aligned}f_n(x) &= \sum_{j=0}^{n-1} F_n(j) \cdot e_j(x) \\&= \sum_{j=0}^{n-1} F_n(j) \cdot \frac{1}{\sqrt{n}} e^{\frac{i2\pi jx}{n}} \\&= \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} F_n(j) \cdot e^{\frac{i2\pi jx}{n}} \\&= \left(\frac{1}{\sqrt{n}} \sum_{x=0}^{n-1} F(x) \cos \frac{2\pi jx}{n}, \frac{1}{\sqrt{n}} \sum_{x=0}^{n-1} F(x) \sin \frac{2\pi jx}{n} \right)\end{aligned}$$

DFT - Polarkoordinaten

komplexe Zahl (x, y) als Polarkoordinaten mit Länge und Winkel

- Winkel (Phase): $\tan \gamma = \frac{y}{x}$
- Länge: $l = \sqrt{x^2 + y^2}$
- Winkel drückt Verschiebung aus (Sinus versus Kosinus)
- Frequenzspektrum berücksichtigt nur Länge

DFT - Eigenschaften

- **Parseval-Theorem:** $\langle f_n^1(x), f_n^2(x) \rangle = \langle F_n^1(x), F_n^2(x) \rangle$
→ euklidische Distanzen sind im Orts- und Frequenzbereich gleich
- Translation im Ortsbereich ändert ausschließlich Phasenwinkel
- **Symmetrie** der Fourier-Koeffizienten: Werte sind spiegelsymmetrisch
→ n reelle Zahlen reichen zur Darstellung von $F_n(x)$
- **Nyquist-Theorem (Abtasttheorem):** zur Abbildung bestimmter Frequenz mind. doppelt so viele Abtastwerte erforderlich → Symmetrie