Warum ist die Verwendung von Ähnlichkeitsmaßen generell problematisch? Problematik beim Messen von Ähnlichkeit mittels Distanzmaß.

Distanz und Ähnlichkeit: Distanzfunktionen

Einleitung

- Distanzfunktionen vergleichen die Merkmale zweier Medienobjekte
- Invarianz – drückt aus welche Merkmale zum Vergleich nicht herangezogen werden sollten

Definition

- Binäre Funktion mit folgenden Eigenschaften
    - Selbstidentität, Positivität, Symmetrie, Dreiecksungleichung

Distanzfunktionen

- Einfache Distanzfunktion, Euklidische Distanzfunktion, Minkowski Distanzfunktion, m-Einheitskreise, gewichtete Minkowski-Distanz, Einheitskreise, Quadratische Distanz, Mahalanobis Distanzfunktion, Quadratische Pseudodistanz, Einheitskreis, Bottleneck Distanz

Ähnlichkeitsmaße

- Objekte werden als ähnlich wahrgenommen, wenn sie bei Menschen zu ähnlichen Reizen (Stimuli) führen
- Keine allgemein akzeptierte Definition von Ähnlichkeit
- Ähnlichkeitsmodelle in Mathematik, Statistik, Bildverarbeitung und Mustererkennung
- Ähnlichkeitsmaß: Funktion, die einem Paar von Objekten eine reelle Zahl aus [0,1] zuordnet
- Wert 1 korrespondiert mit maximaler Ähnlichkeit

Distanz vs. Ähnlichkeit

- Viele Ansätze verwenden Distanzfunktionen auf Featurewerten
- Distanzwerte werden auf [0,1] abgebildet
- Distanzeigenschaften für Ähnlichkeitsempfinden zu restriktiv (Untersuchungen in der Psychologie)
- Bedeutet nicht automatisch, dass Distanzfunktionen für Ähnlichkeitsmaße ungeeignet sind.
- Nur nicht grundsätzlich für alle Anwendungen geeignet

Probleme

- Selbstidentität: gilt nicht grundsätzlich
- Positivität: von Tversky als allgemeine Bedigung für menschliches Ähnlichkeitsempfinden widerlegt
- Symmetrie: Rollentausch macht Unterschied
- Dreiecksungleichung: Unterschiede zwischen 2 Objekten werden zu hoch bewertet wenn kein drittes für den Vergleich vorliegt

Dreiecksungleichung

- Unähnlichkeit zwischen A und B wird stärker eingeschätzt als Summe der Unähnlichkeiten zu C

Ähnlichkeitsabstand

- Unähnlichkeitsmaß
- Mindesteigenschaften:
    - Dominanz
    - Konsistenz
    - Transitivität
- Eigenschaften sind allgemeiner als Distanzeigenschaften (z.B. Symmetrie nicht gefordert)
- Bei Anwendung einer monoton wachsenden Funktion auf Werte eines Abstandmaßes bleiben Eigenschaften erhalten

Grenzen von Ähnlichkeitsmaßen

- Weltwissen spielt Rolle bei menschlicher Wahrnehmung
- Ebenen der Inhaltsverarbeitung
    - Syntaktisch (ohne Bedeutung der Objekte)
    - Semantisch (Ähnlichkeitsvergleich)
    - Pragmatisch (Interpretation, thematische Kategorien)

Ähnlichkeitsmaße

- Viele verschiedene Funktionen und Maße vorgeschlagen, Viele Kombinationen
- Leider keine allgemein anerkannte Kombination
- Viele Alternativen zur Auswahl

Geben Sie eine kurze verbale Beschreibung der Kernaussage des Repräsentationssatzes im Feature Kontrast Modell von Tversky. Beschreibe das tversky ähnlichkeitsmaß?

Positivität: von Tversky als allgemeine Bedingung für menschliches Ähnlichkeitsempfinden widerlegt
Feature-Kontrast-Modell

- Matching
- Monotonie
- Unabhängigkeit

Mindesteigenschaften [Tversky und Gati]:

- Dominanz
- Konsistenz
- Transitivität

Für ein Ähnlichkeitsmaß s(a,b) zwischen Objekten a, b auf Grundlage der korrespondierenden Eigenschaftsmengen A und B gelten [Tversky]:

- Matching

  $$s(o_1, o_2) = f(A \cap B, A \setminus B, B \setminus A)$$

- Monotonie
- $$s(a,b) \geq s(a,c) \text{ gdw.}$$
  $$A \cap C \subseteq A \cap B, \quad A \setminus B \subseteq A \setminus C, \quad B \setminus A \subseteq C \setminus A$$
- Unabhängigkeit:
- Übereinstimmung — f(X,Y,Z) sei ein Ähnlichkeitsmaß mit $X = A \cap B$, $Y = A - B$, und Z = B – A. Wir schreiben V ~ W, wenn X, Y und Z existieren, für die eine oder mehrere der Bedingungen gelten:

  $$\begin{aligned} f(V,Y,Z) &= f(W,Y,Z) \\ f(X,V,Z) &= f(X,W,Z) \\ f(X,Y,V) &= f(X,Y,W) \end{aligned}$$

  Für die Unabhängigkeit muss gelten:

  $$s(a,b) \geq s(a',b') \iff s(c,d) \geq s(c',d')$$

Parametrisierbare Funktion, Repräsentationssatz:
Angenommen, s sei ein Ähnlichkeitsmaß, für welches Matching, Monotonie und Unabhängigkeit erfüllt sind. Dann existiert eine Ähnlichkeitsfunktion S, eine nichtnegative Funktion f sowie zwei Konstanten Alpha, Beta >= 0, so dass für alle Objekte a; b; c; d

$$S(a,b) \geq S(c,d) \iff s(a,b) \geq s(c,d)$$

und

$$S(a,b) = f(A \cap B) - \alpha f(A \setminus B) - \beta f(B \setminus A)$$

gelten. Dieser Satz besagt, dass jede Ähnlichkeitsordnung, welche Matching, Monotonie und Unabhängigkeit erfüllt, durch eine Linearkombination der Funktionswerte über der Menge der Gemeinsamkeiten (A nB) und den beiden Mengen der Unterschiede (A\B;B\A) nachgebildet werden kann.
Insbesondere lässt sich das Feature-Kontrast-Modell gut verwenden, um eine gewünschte Asymmetrie nachzubilden. Im vorigen Abschnitt wurde diskutiert, dass die Ähnlichkeit eines Objektes Alpha mit relativ gering hervorstehenden Eigenschaften zu einem Objekt b mit relativ stark hervorstehenden Eigenschaften größer ist als umgekehrt. Im Feature-Konstrast-Modell geht man davon aus, dass die Stärke der Eigenschaften durch die Funktion f ausgedrückt werden kann:

Erklären Sie kurz das Boolesche IR Modell und diskutieren Sie Nachteile.

boolesches Modell: Dokumente werden als Mengen von Indextermen repräsentiert. Suche über Terme in Termmengen; boolescher Junktoren

- Konzepte der Mengentheorie und der booleschen Algebra
- klare Semantik, sehr einfaches Modell
- Gewicht eines Terms bezogen auf ein Text-Dokument binär
- Jedes Dokument wird intern durch die Menge von Indextermen mit Gewicht „1“ repräsentiert
- in Anfrage werden Terme angegeben, die durch boolesche Junktoren, also durch »and«, »or« und »not«, kombiniert werden
- Innerhalb der Vergleichsfunktion werden die durch die Anfrage spezifizierten Anfrageterme in den jeweiligen Dokumenten auf Enthaltensein getestet.
- Ergebnis eines Termtests ist ein boolescher Wert
- boolesche Junktoren zur Kombination einsetzbar

Nachteile:

- **exaktes Modell:** aufgrund binärer Gewichte keine Ähnlichkeitssuche
- **Größe des Ergebnisses:** oft zuviele Dokumente oder keine
- **boolesche Junktoren:** Schwierigkeiten vieler Anwender

Milderung:

- **exaktes Modell:** Umwandlung von Konjunktionen in Disjunktionen; Stufen der Relevanz
- **Größe des Ergebnisses:** einige Systeme haben zweistufiges Suchverfahren (faceted query)
    - Anfrage formuliert und verfeinert, ohne jedoch das Ergebnis anzuzeigen.
    - vollständiges Ergebnis
- **boolesche Junktoren:** all und any statt and und or

Vektorraummodell IR Modell

- sehr weit verbreitetes Retrieval-Modell
- Dokumente werden als Vektoren eines Vektorraums aufgefasst
- Überführung des Retrieval-Problems in das Gebiet der Linearen Algebra
- kann überall dort eingesetzt werden, wo Medienobjekte durch eine feste Anzahl numerischer Merkmalswerte dargestellt werden können und sich Ähnlichkeit auf dieser Basis berechnet läßt
- unterstützt im Gegensatz zum booleschen Modell das Konzept der Ähnlichkeit
- Die Ähnlichkeit wird zwischen zwei Vektoren berechnet
- Anfrage wird durch einen Vektor repräsentiert
- viele Möglichkeiten zur Berechnung der Ähnlichkeit (z. B. Cosinusmaß)
- auch Ähnlichkeit über Distanzfunktionen

Relevance Feedback: Rocchio Verfahren

Browsing — sequentielle Suche, meist wenig sinnvoll, schlechte Anfrageformulierung, vage Vorstellung über das Suchergebnis, **Pseudorelevanz**: automatische Bewertung

Bewertung von Dokumenten: unterschiedliche Bewertungen eines Dokuments bezgl. Verschiedener Dokumenteigenschaften, erhöhter Bewertungsaufwand Dokumentenbewertung kann auslösen:

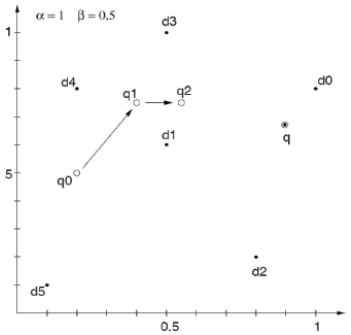
- Anfragemodikation
- Modifikation von Nutzerprofilen
- Modifikation der Dokumentbeschreibungen
- Modifikation des Suchalgorithmus
- Modifikation von Anfragetermgewichten

Verfahren von Rocchio:

- eigentlich entwickelt für Textdokumente, aber auf andere Medientypen übertragbar
- Modifikation von Termgewichten des Anfragevektors im Vektorraummodell
- Termgewichte relevanter Dokumente werden verstärkt und die Termgewichte irrelevanter Dokumente abgeschwächt
- Verschiebung des Anfragepunktes innerhalb des Vektorraums in Richtung der relevanten Dokumente
- Menge »Dr« enthält alle bezüglich der Anfrage »qalt« als relevant markierten Dokumente und
- »Di« alle diesbezüglich irrelevanten Dokumente
- Modifikation des Anfragevektors:

$$q_{neu} = q_{alt} + \frac{\alpha}{|D_r|} \sum_{d_r \in D_r} d_r - \frac{\beta}{|D_i|} \sum_{d_i \in D_i} d_i$$

- Alpha und Beta sind Koeffizienten und gewichten den Einfluss der relevanten und irrelevanten Dokumente



Was ist Spektrogramm?

einfache Darstellungen haben Grenzen:

- Zeit-Domäne zeigt Frequenz-Anteile eines Signals nicht
- Frequenz-Domäne zeigt nicht, wann Frequenzen auftreten

Die kombinierte Darstellung ist das Spektrogramm, Ein Spektrogramm ist die Darstellung des zeitlichen Verlaufes des Spektrums eines Signals. Spektrogramm die Zusammensetzung eines Signals (zum Beispiel eines Klangs oder gesprochener Sprache) aus einzelnen Frequenzen im zeitlichen Verlauf dar.

- x-Achse: Zeit, y-Achse: Frequenzanteile
- Schwärzung (Farbe) eines Punkts: Energie der Frequenz zu dieser Zeit

Vorgehensweise:

- Eingangssignal wird blockweise verarbeitet
- überlappende Segmente des Signals werden verwendet
- sinusoide Fensterfunktionen in Abb. deuten Signalausschnitte an, auf die sich die Analyse in einem Schritt „konzentriert“
- punktweise Multiplikation von Signalblock mit Fensterfunktion

- resultierendes Signal wird Fouriertransformiert
- Aneinanderreihung der Spektralvektoren liefert Zeit-Frequenzdarstellung des Signals
- Etwas präziser ist ein Spektrogramm eine Darstellung der Frequenzverteilung im Phasenraum mit Hilfe der gefensterten Fouriertransformation. Dies ergibt eine komplexwertige Funktion $f(t,\omega)$, die vom Zeitpunkt t und der Frequenz ω abhängt. Das Spektrogramm ist eine Darstellung, die die Werte $|f(t,\omega)|$ in einem Zeit-Frequenz-Diagramm (zum Beispiel farbcodiert) aufträgt.
- Die Interpretation ist dabei die, dass $|f(t,\omega)|$ den Frequenzanteil des Signals zum Zeitpunkt t angibt.

Analysen:

- z.B. Regelmäßigkeit des Auftretens von Frequenzen, Musik vs. Geräusch, mannliche oder weibliche sprache, Musik: Arten von Musik, Umgebungsgeräusche zB. Tierlaute
- häufig statistische Interpretation als nicht normalisierte Dichtefunktion (pdf) über die Frequenz
- erlaubt Berechnung statistischer Parameter, z.B. Lagemaße, Streuung

Erklären Sie kurz die Terminal-Architektur von MPEG-7.

Terminal ist das Teil, das die kodierten Repräsentationen des Multimedia-Contents benutzt. Dabei kann es sich um eine eigenständige Applikation oder einen Teil eines Applikationssystems handeln. Die Architektur besteht aus:

- **Application**
- **Compression Layer:** hier werden die Access-Units geparst und die content description wird rekonstruiert - eine Konvertierung ins Textformat im Rahmen der Decodierung ist nicht Voraussetzung. Verarbeitung kann entweder im Textformat oder in einem proprietären Binärformat erfolgen.
- **Delivery Layer:** stellt Mechanismen für die Synchronisation, Framing und Multiplexing von MPEG-7 Inhalten bereit:
    - MPEG-7 inhalte können unabhängig oder mit dem Content selbst übertragen werden
    - Nicht alle MPEG-7 Streams müssen downstream-fähig sein.
    - Stellt elementary Streams für den Compression Layer zur Verfügung.
    - Elementary streams bestehen aus aufeinanderfolgenden einzeln zugreifbaren Datenpaketen → Access Units (= Kleinste Dateneinheit, der Zeitinformationen zugewiesen werden können)
    - Elementary Streams beinhalten folgende Informationen
        - • Schema information - Struktur der MPEG-7 Description
        - • Descriptions information - komplette oder teilweise Beschreibung des Contents
- **Transmission / Storage Medium:** unterste Ebene, liefert gemultiplexte Streams an den Delivery Layer (DL)

Erklären Sie kurz die Konzepte Rights Data Dictionary (RDD) und Rights Expression Language (REL). In welchem Kontext treten sie auf und wie spielen sie zusammen?

Die RDD und REL treten im Kontext von MPEG-21 auf. REL benutzt das Rights Data Dictionary (RDD) als Vokabular. Die Rights Expression Language und ist ein XML-Schema (daher maschinenlesbar!) zur Definition von verschiedenen Benutzerrechten für ein Digital Item (ist ein strukturiertes, digitales Objekt mit einer standardisierten Repräsentation, Identifikation und Metadateninnerhalb des MPEG-21 Standards; sie sind fundamentale Verarbeitungseinheiten (Transaktion, Verteilung) innerhalb des Frameworks; werden durch Digital Item Declaration Language (DIDL) beschrieben (XML!). Der Schutz von Rechten insbesonders des Urheberrechts ist ein wesentlicher Bestandteil von MPEG-21. Die REL soll daher digitale Inhalte schützen, beabsichtigt die Spezifikation von Kontrolle und der Benutzung von digitalen Inhalten und soll Bedingungen für den Austausch von sensiblen oder privaten digitalen Inhalten unterstützen.
REL kennt folgende Rechte:

- Abspielen
- Ansehen
- Ausdrucken
- Editieren

Außerdem kennt sie verschiedene Conditions:

- Gültigkeit innherhalb bestimmter Zeiträume
- Bezahlarten wie pay-per-use oder flat-fee
- Nutzungsanzahlen wie einmalige oder unbegrenzte Nutzung
- Gebietsgültigkeiten wie Europa oder Amerika

Das Rights Data Dictionary (RDD) definiert verschiedene Bedingungen aus dem Bereich digitaler Rechteverwaltung. Die hier eingetragenen Begriffe sollen eindeutig sein und sind mit einer Auslegung versehen, damit es keine Probleme bei der Implementierung gibt.

Erklären Sie kurz die Ziele von MPEG-21?

- „to define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by different communities." (Vision der Arbeitsgruppe des MPEG-21 Standards)
- Einheitliches Framework, das alle Aspekte der Erstellung, Verteilung und Konsumierung von Multimediainhalten regelt.
- Schaffung einer system- und plattformunabhängigen Umgebung für den Austausch und Gebrauch von Medieninhalten.
- Soll die Integration von multimedialen Inhalten verschiedener Formate (Standards) ermöglichen.
- Vereinheitlichung des Multimediaverkehrs: Multimedia Content Provider mit verschiedenen Geschäftsmodellen sollen auf eine einheitliche Infrastruktur von der Erstellung bis zur Konsumierung von Multimedialen Inhalten zurückgreifen können.
- Effiziente und automatisierte Interoperabilität zwischen verschiedenen Multimedia-Dienstanbietern.
- Standardisierung sämtlicher Komponenten
- Soll Benutzern, Herstellern von Multimediainhalten und -diensten Homogenität, Transparenz und Kompatibilität bieten.
- Integration bereits bestehender jedoch komplexer Technologien zu einem Gesamtsystem (z.B.: Digital Rights Management (DRM))