Metrische Anpassung der Earth Mover's Distanz zur Ähnlichkeitssuche in Multimedia-Datenbanken

Christian Beecks, Marc Wichterich, Thomas Seidl
Data Management and Exploration, RWTH Aachen University
{beecks, wichterich, seidl}@informatik.rwth-aachen.de

Abstract: Die zunächst im Bereich Computer Vision vorgestellte Earth Mover's Distanz (EMD) hat aufgrund Ihrer hohen Flexibilität über eine Vielzahl von Anwendungen Einzug in die inhaltsbasierte Ähnlichkeitssuche in Multimedia-Datenbanken gehalten. Techniken zur effizienten Bearbeitung von EMD-Anfragen wurden in der jüngeren Vergangenheit vorgestellt und basieren zumeist auf einer mehrstufigen Bearbeitung mittels Indexierung oder Filterung der Datenbank. Eine Nutzung der Flexibilität der EMD zur nachträglichen Anpassung des Ähnlichkeitsmaßes wird jedoch nicht diskutiert. Der vorliegende Beitrag stellt ein mathematisches Rahmenwerk zur Anpassung des EMD-Maßes bei Erhalt der metrischen Eigenschaften vor.

1 Einleitung

Ständig wachsende Datenbanken multimedialen Inhalts erschweren es, einen umfassenden Überblick über die Daten zu erhalten. Auch die Anzahl solcher digitaler Datensammlungen steigt. Als Folge dessen ist es notwendig, große Mengen digitaler Daten übersichtlich und durchsuchbar zu strukturieren.

Der Wunsch nach ähnlichen Inhalten zu suchen wird durch Modelle der Ähnlichkeitssuche erfüllt. Sie unterstützen beispielsweise private Anwender bei der Suche nach ähnlichen Urlaubsbildern oder wissenschaftliche Anwender bei der Auswertung ihrer Arbeiten.

Ein flexibles, auf einem Optimierungsproblem basierendes Distanzmaß, welches an die vorhandenen Daten und an die subjektive Wahrnehmung der Ähnlichkeit angepasst werden kann, ist die sogenannte Earth Mover's Distanz (EMD) [RT01, AWS06b]. Für den effizienten Einsatz der EMD in Multimedia-Datenbanksystemen wurden in den letzten Jahren mehrere Techniken vorgestellt (siehe u. A. [LBS06, AWS06a, AWMS08, WAKS08]). Sie basieren zumeist auf einer mehrstufigen Anfragebearbeitung mittels Indexierung oder Filterung der Datenbank. Eine Nutzung der Flexibilität der EMD zur nachträglichen Anpassung des Ähnlichkeitsmaßes wird jedoch nicht diskutiert. Der vorliegende Beitrag[1] stellt ein mathematisches Rahmenwerk zur Anpassung des EMD-Maßes bei Erhalt der metrischen Eigenschaften vor.

[1]Die dem Beitrag zugrundeliegende Forschung ist Teil des DFG-Projekts "Schnelle inhaltsbasierte Suche in großen Multimedia-Datenbanksystemen mittels der Earth Mover's Distance" (SE 1039/1-2 und SE 1039/1-3).

2 Grundlagen und verwandte Arbeiten

In diesem Abschnitt werden für das weitere Verständnis notwendige Grundlagen und verwandte Arbeiten behandelt.

2.1 Ähnlichkeitssuche

Die Ähnlichkeitssuche ist ein Prozess, bei dem zu einem vorgegebenen Datenobjekt in einer Datenbank ähnliche Datenobjekte gesucht werden. Die Ähnlichkeit zwischen je zwei dieser Datenobjekte ergibt sich aus den Merkmalen der Datenobjekte unter Zuhilfenahme eines geeignet definierten Distanzmaßes.

Das Ergebnis einer Ähnlichkeitssuche in einer Datenbank kann beispielsweise die Menge aller Objekte p der Datenbank sein, für welche die Distanz zu einem Anfrageobjekt q kleiner gleich als eine gegebene Obergrenze ε ist:

$$\{p \in DB | d(q, p) \leq \varepsilon\}$$

Für die Bearbeitung von Suchanfragen existieren effiziente Methoden. Häufig wird hierbei auf Indexstrukturen und Filter zurückgegriffen, welche in eine mehrstufige Anfragebearbeitung integriert werden [FRM94, KSF^{+}98, SK98]. Einen Überblick über den gesamten Themenkomplex der Ähnlichkeitssuche gibt das Buch von Schmitt [Sch05].

2.2 Merkmale und deren Repräsentationen

Die für die inhaltsbasierte Ähnlichkeitssuche relevanten Merkmale eines Multimedia-Objektes können durch Histogramme in einer aggregierten Form erfasst werden. So lassen sich beispielsweise Bilder durch ihre Farben, Formen oder Texturen und das Layout dieser Merkmale beschreiben.

Definition 1 *Sei $\mathcal{M}$ ein Merkmalsraum mit einer paarweise disjunkten Partitionierung $\mathcal{M} = \mathcal{M}_1 \uplus \ldots \uplus \mathcal{M}_n$ und p ein durch $p_1, \ldots, p_k \in \mathcal{M}$ beschriebenes Objekt. Ferner bezeichne $|\mathcal{M}_i^p| = \frac{1}{k} \cdot |\{p_j | p_j \in \mathcal{M}_i, 1 \leq j \leq k\}|$ den Anteil der einer Partition $\mathcal{M}_i$ angehörigen Elemente. Dann ist das Histogramm $h_p \in \mathcal{R}^n$ des Datenobjekts p wie folgt definiert:*

$$h_p = (h_{p,1}, \ldots, h_{p,n}) \textit{ mit } h_{p,i} = |\mathcal{M}_i^p|, 1 \leq i \leq n.$$

Ein Histogramm lässt sich hierbei als ein Punkt auf einer Hyperebene in einem n-dimensionalen Vektorraum auffassen. Die Summe der Histogrammeinträge ist jeweils 1. Im Falle von Farbhistogrammen sind die Merkmale $p_1, \ldots, p_k \in \mathcal{M}$ beispielsweise die Farbwerte von k Pixeln eines Bilds, wobei $\mathcal{M}$ ein dreidimensionaler Farbraum ist (Bsp.: RGB, HSV).

2.3 Distanzmaße und die EMD im Speziellen

Neben den in Abschnitt 2.2 beschriebenen Merkmalen und deren Repräsentationen tragen Distanzmaße wesentlich zur Qualität der Ähnlichkeitssuche bei. Distanzmaße werden eingesetzt, um die Ähnlichkeit zwischen zwei Merkmalsrepräsentationen, im Folgenden Histogramme, zu messen. Ähnliche Repräsentationen der zu vergleichenden Objekte sollen hierbei eine kleine Distanz aufweisen.

Definition 2 *Sei d eine Distanzfunktion und x, y, z Elemente aus einem Raum X. Die Distanzfunktion $d : X \times X \mapsto \mathcal{R}^+$ ist eine Metrik, wenn sie für alle x, y, z die folgenden Eigenschaften erfüllt:*

- *Definitheit:* $d(x, y) = 0$ *gdw.* $x = y$
- *Symmetrie:* $d(x, y) = d(y, x)$
- *Dreiecksungleichung:* $d(x, y) \leq d(x, z) + d(z, y)$.

Metrische Distanzeigenschaften lassen sich für die EMD effizienzgewinnend ausnutzen [LBS06, AWS06a, AWMS08] und ermöglichen den Einsatz von Methoden zur distanzbasierten Indexierung (siehe [Sam06] Kapitel 4.5).

Die Earth Mover's Distanz

Die Earth Mover's Distanz wurde von Rubner und Tomasi [RT01] im Bereich der Computer Vision vorgestellt. Das Anwendungsspektrum der Earth Mover's Distanz ist weit gefächert. Es existieren eine Vielzahl von Publikationen, welche die Earth Mover's Distanz als Ähnlichkeitsmaß für die Suche in Datenbanken einsetzen, z.B. auf Bildern [RT01, JLZZ04], auf Musikstücken [TGV$^+$03] oder auf Vektorfeldern in der Physik [LBH98].

Die Earth Mover's Distanz basiert auf dem Transportproblem der Linearen Optimierung, dessen Lösung durch eine hierfür optimierte Variante des Simplex-Algorithmus [HL90] berechenbar ist. Das Transportproblem beschreibt die Problematik, Produkte von Anbietern auf Verbraucher zu verteilen. Die Kosten, die für den Transport einer Produkteinheit von einem Anbieter zu einem Verbraucher anfallen, sind bekannt. Das Problem dieser Situation besteht darin, eine Verteilung der Produkte zwischen Anbietern und Verbrauchern mit minimalen Transportkosten zu finden. Übertragen auf die EMD, im Kontext von Histogrammen, berechnet diese die minimalen aufsummierten Kosten für die Transformation eines Histogramms in ein anderes Histogramm. Die Kosten für den "Transport" einer Einheit vom Histogrammeintrag $h_{p,i}$ nach $h_{q,j}$ werden hierbei durch eine Kostenmatrix $C = [c_{ij}]$ spezifiziert.

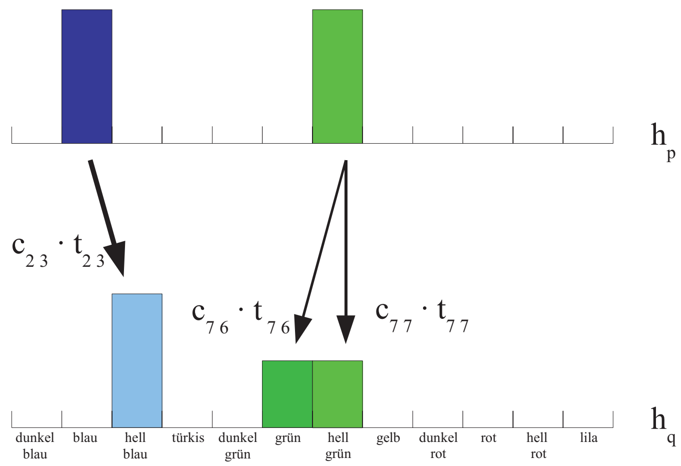

Abbildung 1: Die Earth Mover's Distanz berechnet den optimalen Transport $T = [t_{ij}]$ für die Transformation eines Histogramms h_p zu einem Histogramm h_q bei Kosten $C = [c_{ij}]$. Die Distanz beträgt hier $c_{23}t_{23} + c_{76}t_{76} + c_{77}t_{77}$.

Definition 3 *Seien zwei Histogramme $h_q = (h_{q,1}, \ldots, h_{q,n})$ und $h_p = (h_{p,1}, \ldots, h_{p,n})$ und eine Kostenmatrix $C \in \mathcal{R}^{n\times n}$ gegeben. Die Earth Mover's Distanz $EMD_C(h_q, h_p)$ ist als Optimierungsproblem über alle zulässigen Transportmatrizen $T = [t_{ij}]$ definiert:*

$$EMD_C(h_q, h_p) = \min_{T \in \mathcal{R}^{n\times n}} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} t_{ij}$$

unter den Einschränkungen

$$t_{ij} \geq 0, \quad \forall 1 \leq i, j \leq n, \tag{1}$$

$$\sum_{j=1}^{n} t_{ij} \leq h_{q,i}, \quad \forall 1 \leq i \leq n, \tag{2}$$

$$\sum_{i=1}^{n} t_{ij} \leq h_{p,j}, \quad \forall 1 \leq j \leq n, \tag{3}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} t_{ij} = 1. \tag{4}$$

Die Einschränkungen (1) bis (4) in der Definition 3 der Earth Mover's Distanz entsprechen der intuitiven Vorgehensweise bei der Transformation eines Histogramms in ein zweites Histogramm, wie in Abbildung 1 dargestellt. Die Gleichung (1) stellt sicher, dass alle auftretenden Transporte positiv sind. Die Gleichungen (2) und (3) limitieren die Transporte gemäß den Einträgen der Histogramme. Gleichung (4) formalisiert die Forderung, dass eine vollständige Transformation durchzuführen ist.

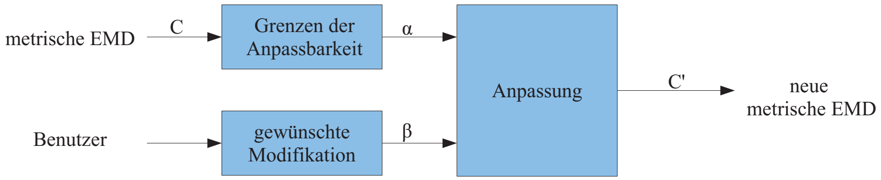

Abbildung 2: Schematischer Ablauf der metrischen Anpassung

3 Ein mathematisches Rahmenwerk zur Adaption der EMD

In diesem Abschnitt werden unsere Ansätze zur Anpassung der Earth Mover's Distanz auf Histogrammen unter Einhaltung der metrischen Eigenschaften beschrieben. Wir entwickeln ein Rahmenwerk welches die Anpassung der Einträge einer Kostenmatrix erlaubt. Abbildung 2 stellt die dazu notwendigen Schritte schematisch dar. Die Grenzen der Anpassbarkeit ergeben sich aus der Kostenmatrix der EMD und werden in Abschnitt 3.1 behandelt. Die eigentliche Anpassung, basierend auf der gewünschten Modifikation und den Grenzen der Anpassbarkeit, wird in Abschnitt 3.2 vorgestellt. Abschließend gibt Abschnitt 3.3 einen beispielhaften Ausblick auf mögliche Techniken zur Bestimmung der gewünschten Modifikation.

3.1 Grenzen der metrischen Anpassung

Die metrischen Eigenschaften des Distanzmaßes EMD_C werden durch die verwendete Kostenmatrix C impliziert. Der nachfolgende Satz zeigt, dass die Kostenmatrix C dazu Eigenschaften erfüllen muss, welche analog zu denen aus Definition 2 sind.

Satz 1 *Sei eine Kostenmatrix $C \in \mathcal{R}^{n \times n}$ gegeben. Die Earth Mover's Distanz EMD_C ist metrisch gdw. folgende Eigenschaften für $1 \leq i, j, k \leq n$ erfüllt sind.*

- *Definitheit: $c_{ij} = 0$ gdw. $i = j$*
- *Symmetrie: $c_{ij} = c_{ji}$*
- *Dreiecksungleichung: $c_{ij} \leq c_{ik} + c_{kj}$*

Beweisskizze. Die Einträge c_{ij} der Kostenmatrix C implizieren ein Distanzmaß zwischen den Dimensionen i und j der zu vergleichenden Histogramme. Dieses erfüllt nach Definition 2 die Eigenschaften einer Metrik. Nach [RT01] ist somit auch EMD_C eine Metrik. Ist eine der drei Eigenschaften für C nicht erfüllt, so lässt sich jeweils ein Beispiel konstruieren, welches zeigt, dass dann auch EMD_C nicht metrisch sein kann. □

Die Adaption der Earth Mover's Distanz erfolgt durch die Modifikation einzelner Einträge der Kostenmatrix. Die Modifikation einzelner Einträge kann durch Änderungsfaktoren realisiert werden, die für jeden Eintrag individuell bestimmbar sind. Eine Multiplikation eines Wertes der Kostenmatrix mit dem entsprechenden Änderungsfaktor bestimmt den resultierenden neuen Wert.

Definition 4 *Sei eine Kostenmatrix $C \in \mathcal{R}^{n\times n}$ gegeben. Weiterhin seien Änderungsfaktoren α_{ij} gegeben. Dann sind die Einträge der neuen Kostenmatrix $C' \in \mathcal{R}^{n\times n}$ wie folgt definiert:*

$$c'_{ij} = c_{ij} \cdot \alpha_{ij}.$$

Die notwendigen Einschränkungen, welche bei der Veränderung der Werte c_{ij} für $i \neq j$ gemäß der vorangegangenen Definition bestehen, ergeben sich aus Satz 1. Hierbei bedarf die Einhaltung der Dreiecksungleichung besonderer Betrachtung, da sie abhängig von den Verhältnissen der Werte untereinander ist. Das aus Satz 1 folgende Lemma zeigt die unteren und oberen Grenzen für die Anpassung eines Eintrags.

Lemma 1 *Sei eine beliebige Kostenmatrix $C \in \mathcal{R}^{n\times n}$ einer metrischen EMD gegeben. Die maximal zulässige Verkleinerung bzw. Vergrößerung α_{ij} des Eintrags $c_{ij}(i \neq j)$ unter Einhaltung der Dreiecksungleichung ist durch die untere Grenze $0 \leq \alpha_{ij}^- \leq 1$ und die obere Grenze $1 \leq \alpha_{ij}^+ \in \mathcal{R}$ beschränkt:*

$$\alpha_{ij}^- = \frac{\max_{k\neq i,j}\{|c_{ik} - c_{jk}|\}}{c_{ij}} \quad \leq \quad \alpha_{ij} \quad \leq \quad \frac{\min_{k\neq i,j}\{c_{ik} + c_{jk}\}}{c_{ij}} = \alpha_{ij}^+.$$

Die Multiplikation des Änderungsfaktors α_{ij} in den zulässigen Grenzen $[\alpha_{ij}^-, \alpha_{ij}^+]$ mit dem Matrixeintrag c_{ij} einer Kostenmatrix C erhält also die Gültigkeit der Dreiecksungleichung. Für den Erhalt der Symmetrie muss jede Änderung an c_{ij} analog auch für c_{ji} durchgeführt werden. Weiterhin bedingt der Erhalt der Definitheit, dass $\alpha_{ij} > 0$ eingehalten wird. Der folgende Satz zeigt, dass eine Betrachtung von Werten $0 < \alpha_{ij} < 1$ jedoch gar nicht notwendig ist, da der Effekt einer jeden solchen Verkleinerung auch per Vergrößerung der übrigen Werte erzielt werden kann.

Satz 2 *Sei eine beliebige Kostenmatrix $C \in \mathcal{R}^{n\times n}$ einer metrischen EMD gegeben. Eine Verkleinerung der Einträge c_{ij} und c_{ji} durch den Änderungsfaktor $\alpha > 0$ mit $\alpha_{ij}^- \leq \alpha \leq 1$ und $\alpha_{ji}^- \leq \alpha \leq 1$ ist äquivalent zu einer Vergrößerung der restlichen Einträge $c_{i'j'}$ durch den Änderungsfaktor $\frac{1}{\alpha} \geq 1$.*

Beweis. Bezeichne $C^- \in \mathcal{R}^{n\times n}$ die Kostenmatrix, die aus der Matrix C entsteht, indem die Einträge c_{ij} und c_{ji} mit dem Faktor α multipliziert werden. Sei weiterhin $C^+ \in \mathcal{R}^{n\times n}$ diejenige Matrix, welche aus der Matrix C durch Multiplikation aller Einträge $c_{i'j'}$ bis auf c_{ij} und c_{ji} mit dem Faktor $\frac{1}{\alpha}$ entsteht. Dann gilt gemäß der Konstruktion $\frac{1}{\alpha} \cdot C^- = C^+$. Die Matrix C^- hält gemäß Lemma 1 die Dreiecksungleichung ein und ist per Konstruktion weiterhin symmetrisch und definit. Die Multiplikation von C^- mit dem Faktor $\alpha > 0$

ändert nichts an den metrischen Eigenschaften der EMD. Somit ist C^+ bis auf Skalierung äquivalent zur Kostenmatrix C^- und folglich auch EMD_{C^+} metrisch. $\square$

Die Vergrößerung mehrerer Einträge einer (ggf. skalierten) Kostenmatrix reicht somit aus, um den Effekt einer Verkleinerung eines einzelnen Eintrags zu erzielen. Jede beliebige metrische Modifikation einer Kostenmatrix der Earth Mover's Distanz ist folglich durch Vergrößerungen einzelner Werte und geeigneter Skalierung der Kostenmatrix möglich. Da bei Ähnlichkeitsanfragen vielfach nur das Verhältnis der Distanzen untereinander von Bedeutung ist, muss die Skalierung häufig nicht durchgeführt werden. In anderen Fällen lässt sich die Skalierung der Kostenmatrix auf die Skalierung der Eingabeparameter der Ähnlichkeitsanfrage zurückführen, da $EMD_{\alpha \cdot C} = \alpha \cdot EMD_C$ gilt.

3.2 Einfluss der gewünschten Anpassung

Bisher wurden die Grenzen der maximal zulässigen Veränderungen beschrieben. Die tatsächlich gewünschten Modifikationen werden im Folgenden auf die zulässigen Vergrößerungsintervalle $[1, \alpha_{ij}^+]$ abgebildet. Von der Bestimmung der gewünschten Modifikation soll hier zunächst abstrahiert werden. Einen Ausblick auf mögliche heuristische Verfahren gibt Abschnitt 3.3. Wir gehen hier davon aus, dass die Modifikationswünsche in Form von Einflussfaktoren $\beta_{ij} \in [0, 1]$ symmetrisch mit $\beta_{ij} = \beta_{ji}$ gegeben sind. Dabei soll sich die Relation untereinander möglichst gut in der Modifikation der Kostenmatrix widerspiegeln. Ein Wert von $\beta_{ij} = 0$ entspricht dem Wunsch, α_{ij} nicht zu vergrößern, während $\beta_{ij} = 1$ dem Wunsch nach möglichst starker Vergrößerung entspricht. Analog entspricht $\beta_{ij} = m \cdot \beta_{i'j'}$ dem Wunsch, c_{ij} im Vergleich zu $c_{i'j'}$ prozentual m mal so stark zu erhöhen.

Zunächst wählen wir vorläufige Vergrößerungen $\widetilde{\alpha}_{ij}$ unterhalb des Minimums α_{min}^+ aller oberen Grenzen α_{ij}^+, so dass den durch β_{ij} spezifizierten Modifikationswünschen entsprochen wird.

Definition 5 *Sei eine Kostenmatrix $C \in \mathcal{R}^{n \times n}$ und Einflussfaktoren β_{ij} gegeben. Die vorläufige Vergrößerung $\widetilde{\alpha}_{ij} \in [1, \alpha_{min}^+]$ ist wie folgt definiert:*

$$\widetilde{\alpha}_{ij} = \beta_{ij} \cdot (\alpha_{min}^+ - 1) + 1.$$

Alle β_{ij} sind kleiner gleich 1, wodurch keines der $\widetilde{\alpha}_{ij}$ größer als α_{min}^+ ist. Da sich aus der Definition der oberen Grenzen ergibt, dass die Vergrößerung eines Wertes die oberen Grenzen der anderen Vergrößerungen niemals verringert, können alle Anpassungen $\widetilde{\alpha}_{ij}$ durchgeführt werden und die Dreiecksungleichung gilt weiterhin. Ein Sonderfall tritt bei $\alpha_{min}^+ = 1$ auf. Sind einzelne Einträge apriori gar nicht vergrößerbar, so kann den Modifikationswünschen nur teilweise entsprochen werden (durch Auslassung der entsprechenden Einträge bei der Bestimmung von α_{min}^+ und bei der Anpassung selber).

Die in Definition 5 vorgestellten vorläufigen Änderungsfaktoren $\widetilde{\alpha}_{ij}$ bewirken die Vergrößerung der Einträge c_{ij} der Kostenmatrix C in einem minimalen, zulässigen Bereich. Hierbei wird jedoch die maximal zulässige Vergrößerung nicht notwendiger Weise erreicht. Durch einen einzelnen zusätzlichen Parameter $\lambda \in [0,1]$ als Einflussstärke lässt sich steuern, inwieweit die maximal zulässige Vergrößerung ausgeschöpft werden soll. Somit lassen sich dann schließlich die Änderungsfaktoren α_{ij} angeben:

Definition 6 *Zu einer Kostenmatrix $C \in \mathcal{R}^{n \times n}$, durch Einflussfaktoren β_{ij} bestimmte vorläufige Vergrößerungsfaktoren $\widetilde{\alpha}_{ij}$ und der Einflussstärke λ sind die Änderungsfaktoren α_{ij} wie folgt definiert:*

$$\alpha_{ij} := \lambda \cdot \rho \cdot (\widetilde{\alpha}_{ij} - 1) + 1.$$

Hierbei ist ρ so zu wählen, dass $\alpha_{ij} \leq \alpha^+_{ij}$ für alle $1 \leq i, j \leq n$ gilt und dass die Gleichheit $\alpha_{ij} = \alpha^+_{ij}$ für mindestens ein Paar i, j bei $\lambda = 1$ gilt.

Nun ist nur noch ρ passend zu bestimmen. Dazu zeigt der nächste Satz eine geeignete Wahl, welche alle geforderten Eigenschaften erfüllt.

Satz 3 *Für α_{ij} gemäß Definition 6 und*

$$\rho = \min_{1 \leq i \neq j \leq n} \left\{ \frac{\alpha^+_{ij} - 1}{\widetilde{\alpha}_{ij} - 1} \right\}$$

als der kleinsten prozentualen Vergrößerung der vorläufigen Änderungsfaktoren $\widetilde{\alpha}_{ij}$ gilt:

$$\exists (i^*, j^*) : \Big(\alpha_{i^* j^*} = \alpha^+_{i^* j^*} \wedge \forall (i', j') \neq (i^*, j^*) : \alpha_{i' j'} \leq \alpha^+_{i' j'} \Big)$$

Beweis. Bezeichne i^*, j^* ein Paar von Indizes, welche ρ gemäß Satz 3 minimieren. Dann gilt für $\lambda = 1$:

$$\begin{aligned} \alpha_{i^* j^*} &= \lambda \cdot \rho \cdot (\widetilde{\alpha}_{i^* j^*} - 1) + 1 \\ &= \frac{\alpha^+_{i^* j^*} - 1}{\widetilde{\alpha}_{i^* j^*} - 1} \cdot (\widetilde{\alpha}_{i^* j^*} - 1) + 1 \\ &= \alpha^+_{i^* j^*} \end{aligned}$$

Sei nun i', j' ein beliebiges Paar ungleich i^*, j^* und λ beliebig aus $[0,1]$. Dann gilt:

$$\begin{aligned} \alpha_{i' j'} &= \lambda \cdot \rho \cdot (\widetilde{\alpha}_{i' j'} - 1) + 1 \\ &\leq 1 \cdot \rho \cdot (\widetilde{\alpha}_{i' j'} - 1) + 1 \quad = \quad \frac{\alpha^+_{i^* j^*} - 1}{\widetilde{\alpha}_{i^* j^*} - 1} \cdot (\widetilde{\alpha}_{i' j'} - 1) + 1 \\ &\leq \frac{\alpha^+_{i' j'} - 1}{\widetilde{\alpha}_{i' j'} - 1} \cdot (\widetilde{\alpha}_{i' j'} - 1) + 1 \quad = \quad \alpha^+_{i' j'} \end{aligned}$$

□

Nachdem schließlich die zentrale Definition 6 die metrische Anpassung der Earth Mover's Distanz über die Wahl der Einflussfaktoren β_{ij} ermöglicht, beschreiben wir nachfolgend kurz einige Ansätze zur Wahl dieser Einflussfaktoren.

3.3 Zur Wahl der Einflussfaktoren

Die metrische Anpassung der Earth Mover's Distanz basiert auf der Modifikation der Kostenmatrix mittels den Einflussfaktoren β_{ij}. Während die Modifikation in dem vorangegangenen Abschnitt theoretisch beschrieben wird, soll nachfolgend die Frage nach Heuristiken zur Bestimmung des Anpassungsbedarfs beispielhaft diskutieren werden.

Eine erste Möglichkeit besteht in einer direkten Bewertung der Ähnlichkeitswahrnehmung im Merkmalsraum durch den Benutzer. Ist der Merkmalsraum z.B. ein Farbraum, so steht jeder Histogrammeintrag für einen Farbbereich. Nun ist es möglich, dem Benutzer Paare von repräsentativen Farben zu präsentieren und nach der Ähnlichkeit auf einer fixen Skala zu fragen. Werden beiden Farben als unähnlicher eingeschätzt als durch die Kostenmatrix abgedeckt, so kann ein entsprechend großer Wert β_{ij} gesetzt werden. Fällt es dem Benutzer schwer die Ähnlichkeit auf einer Skala anzugeben, so kann dies gegebenenfalls durch die Verwendung von mehr als zwei Farben pro Einschätzung vermieden werden. Hierbei muss der Benutzer nur entscheiden, ob eine gegebene Farbe ähnlicher oder unähnlicher zu einer zweiten im Vergleich zu einer dritten ist.

Statt eine Bewertung im, für den Benutzer eventuell abstrakten, Merkmalsraum vorzunehmen, kann dies über Techniken des Relevance Feedbacks auch direkt auf der Ebene der Datenbankobjekte geschehen. Hierbei können zum Beispiel statistische Eigenschaften, der durch den Benutzer als relevant eingeschätzten Objekte, in Form einer Kovarianzmatrix $K \in \mathcal{R}^{n \times n}$ gesammelt werden und Rückschlüsse von k_{ij} auf β_{ij} gezogen werden. Durch eine datenbezogene Anpassung läßt sich das vorgestellte Rahmenwerk ggf. für EMD-spezifisches "Distance Metric Learning" [YJ06] einsetzen.

Die konkrete Auswahl und die Untersuchung benutzerbasierter Heuristiken, welche sich in das hier vorgestellte Rahmenwerk integrieren, wollen wir in weiteren Arbeiten angehen.

4 Zusammenfassung und Ausblick

In unserem Beitrag haben wir ein mathematisches Rahmenwerk für die Anpassung der äußerst flexiblen Earth Mover's Distanz bei Erhalt der metrischen Eigenschaften vorgestellt. Zusammen mit den in der letzten Zeit vorgestellten Beschleunigungstechniken lassen sich nunmehr effiziente und effektive EMD-basierte Systeme zur Ähnlichkeitssuche in großen Datenbanken erstellen, welche an die Anwendungsanforderungen angepasst werden können.

In weiteren Arbeiten planen wir konkrete Strategien zur Wahl der Anpassung mittels des hier vorgestellten Rahmenwerkes zu entwickeln.

Literatur

[AWMS08] I. Assent, M. Wichterich, T. Meisen und T. Seidl. Efficient Similarity Search Using the Earth Mover's Distance for Large Multimedia Databases. In *Proc. IEEE Int. Conf. on Data Engineering (ICDE)*, 2008.

[AWS06a] I. Assent, A. Wenning und T. Seidl. Approximation Techniques for Indexing the Earth Mover's Distance in Multimedia Databases. In *Proc. IEEE Int. Conf. on Data Engineering (ICDE)*, 2006.

[AWS06b] I. Assent, M. Wichterich und T. Seidl. Adaptable Distance Functions for Similarity-based Multimedia Retrieval. *Datenbank-Spektrum*, 6(19):23–31, 2006.

[FRM94] C. Faloutsos, M. Ranganathan und Y. Manolopoulos. Fast Subsequence Matching in Time-Series Databases. In *Proc. Int. Conf. on Management of Data (SIGMOD)*, 1994.

[HL90] F. S. Hillier und G. J. Lieberman. *Introduction to Mathematical Programming.* McGraw-Hill, 1990.

[JLZZ04] F. Jing, M. Li, H. Zhang und B. Zhang. An Efficient and Effective Region-based Image Retrieval Framework. *IEEE TIP*, 13(5):699–709, 2004.

[KSF+98] F. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel und Z. Protopapas. Fast and Effective Retrieval of Medical Tumor Shapes. *IEEE TKDE*, 10(6):889–904, 1998.

[LBH98] Y. Lavin, R. Batra und L. Hesselink. Feature Comparisons of Vector Fields using Earth Mover's Distance. In *Proc. IEEE Conf. on Visualization*, 1998.

[LBS06] V. Ljosa, A. Bhattacharya und A. K. Singh. Indexing Spatially Sensitive Distance Measures Using Multi-resolution Lower Bounds. In *Proc Int. Conf. on Extending Database Technology (EDBT)*, 2006.

[RT01] Y. Rubner und C. Tomasi. *Perceptual Metrics for Image Database Navigation*. Kluwer Academic Publishers, 2001.

[Sam06] H. Samet. *Foundations of Multidimensional and Metric Data Structures*. Morgan Kaufmann, 2006.

[Sch05] I. Schmitt. *Ähnlichkeitssuche in Multimedia-Datenbanken - Retrieval, Suchalgorithmen und Anfragebehandlung*. Oldenbourg Verlag, 2005.

[SK98] T. Seidl und H.-P. Kriegel. Optimal Multi-Step k-Nearest Neighbor Search. In *Proc. Int. Conf. on Management of Data (SIGMOD)*, 1998.

[TGV+03] R. Typke, P. Giannopoulos, R. C. Veltkamp, F. Wiering und R. van Oostrum. Using Transportation Distances for Measuring Melodic Similarity. In *Proc. Int. Conf. on Music Information Retrieval (ISMIR)*, 2003.

[WAKS08] M. Wichterich, I. Assent, P. Kranen und T. Seidl. Efficient EMD-based Similarity Search in Multimedia Databases via Flexible Dimensionality Reduction. In *Proc. Int. Conf. on Management of Data (SIGMOD)*, 2008.

[YJ06] Liu Yang und Rong Jin. Distance Metric Learning: A Comprehensive Survey, 2006. `http://www.cse.msu.edu/~yangliu1/frame_survey_v2.pdf`.