

Videosonifikation am Beispiel von Verkehrsflussdaten

Thomas Hermann, Matthias Milczynski
AG Neuroinformatik, Technische Fakultät, Universität Bielefeld
Bielefeld, Germany
<mailto:{thermannjmmilczyn}@techfak.uni-bielefeld.de>

Zusammenfassung

Sonifikation verwandelt Daten in Klänge, um über das Hören Muster und Strukturen in den Daten zu entdecken. Gerade das Gehör ist hochgradig geeignet, Rhythmen und selbst subtile spektrale Eigenschaften wahrzunehmen, die dem Auge beim Betrachten der Daten leicht entgehen. Generell lassen sich Daten aller Art für Sonifikations-Displays verwenden. In diesem Text stellen wir ein Verfahren zur Verwandlung von *Videodaten*, also Bilddatenströmen in Klangströme vor, welches erlaubt, große Videodatenmengen als hochkomprimiertes Klanggefüge zu erleben. Als Datenquelle für die vorliegenden Sonifikationen wurden Webcam-Aufnahmen des Straßenverkehrsflusses an der A28 im Prättigau verwendet. Wir beschreiben das Verfahren, präsentieren mehrere Sonifikationen und diskutieren Anwendungsfelder von Videosonifikation jenseits des hier zur Illustration ausgewählten Beispiels.

Einführung

Sonifikation transformiert Daten in Klänge. Daten werden im derzeitigen Informationszeitalter mehr und mehr zu einem Rohstoff und es gilt, sie zu “Wissen” zu verdichten, welches handlungsrelevant ist und das Verständnis über die Ursprungsdomäne erhöht. Dieser Veredlungsprozess ist Gegenstand des “Datamining” und verwendet das ganze Repertoire mathematisch-statistischer Verfahren bis hin zum sog. maschinellen Lernen. Voraussetzung zum Einsatz derartiger Analyseverfahren ist allerdings ein Entdecken der Strukturierung – man kann nur finden was man sucht: Wer mit einem Massenspektrometer ein Buch analysiert, wird wohl Papier und Tinte verstehen, aber am Inhalt des Buches – an der Bedeutung des Texts – vorbeischaun.

Für diesen Entdeckungsprozess von Mustern ist der Mensch mit seinen Sinnesfähigkeiten immer noch ein von technischen Systemen unübertroffenes Vorbild. Nicht nur, dass er verschiedene Wahrnehmungs-Sinne mitbringt, er besitzt ein Gehirn, welches Sinn aus den Mustern machen kann, den Suchprozess steuert und ferner die Fähigkeit besitzt, vom Konkreten zu abstrahieren. In dieser “Explorativen Datenanalyse” standen in den letzten Jahrzehnten vornehmlich visuelle Methoden im Fokus, welche abstrakte Daten in grafische Darstellungen transformieren und sie so unserem Wahrnehmungskanal mit der höchsten Bandbreite – dem Auge – für die Analyse zugänglich machen.

Seit etwa 12 Jahren zielt nun die Sonifikation darauf, auch unsere exzellenten Hörfähigkeiten zu erschließen. Diese sog. auditiven Displays bilden oft eine gute Ergänzung zu visuellen Darstellungen, da wir im Hören auf andere Merkmale sensitiv sind als im visuellen (z.B. auf Rhythmus). Sie bieten sich auch an, wenn die visuelle Wahrnehmung nicht verfügbar ist (z.B. für blinde Menschen). Vorteilhaft ist auch, dass wir Klänge im Hintergrund hören können: So wird

man beim Autofahren prompt merken, wenn sich das Motorgeräusch ändert, auch ohne dass man andauernd Aufmerksamkeit hierauf richten muss. Ein trainierter Hörer, wie ein Kfz-Mechaniker, kann das Hören sogar als feines Diagnoseinstrument einsetzen. Das Stethoskop des Mediziners ist ein weiteres Beispiel für diagnostischen Einsatz dieser Art.

Wir gehen anhand eines Beispiels auf akustische Entdeckungsreise und hören in eigentlich besonders 'anschauliche' Daten, nämlich Bilddaten.

Sonifikation von Videodaten

Warum sollte man Videodatenströme sonifizieren? Nun, wie bereits dargestellt, sind die Ohren in der Lage, ganz andere Informationen zu extrahieren als das Auge. Auch wenn das Bild vielfältige Informationen preisgibt, verstecken sich Muster in wahrnehmungsunzugänglichen Nischen. Rhythmische Phänomene mit Periodendauern von einigen Minuten wird man im Video auch visuell sicher bemerken, jedoch wird man kaum ohne Stoppuhr feststellen können, ob der Rhythmus Änderungen unterliegt. Das Gehör nimmt derartige rhythmische Strukturen dagegen weitaus differenzierter auf, vorausgesetzt, sie treten in passender zeitlicher Organisation auf. Gerade an dieser Front erlaubt die Sonifikation weitreichende Manipulationen - so lassen sich z.B. Datenausschnitt und Sonifikationszeit frei wählen, und damit können rhythmische Phänomene auf ganz unterschiedlichen Zeitskalen (z.B. Wochenrhythmus, Tagesrhythmus, Ampelphasen, etc..) in den für das Ohr günstigen Analysebereich verschoben werden. Bei zeitlich geordneten Daten wie Börsenkurse, Klimadaten, oder den hier sonifizierten Verkehrsdaten, liegt eine direkte Nutzung der Zeit als Sonifikationszeit nah, diese Verbindung ist jedoch nicht zwingend. In sog. Sonifikationsmodellen wird zum Beispiel die Zeit bewusst frei gelassen für die Interaktion des Nutzers mit einem aus den Daten kreierte klangfähigen Substrat.

Betrachten wir nun einige Anwendungsmöglichkeiten von Videodatensonifikation:

Ersatz visueller Information: Blinde Menschen können mittels auditiven Displays ein akustisches Substitut für den Sehsinn bekommen ("seeing with sounds"). Ist eine solche Kamera z.B. in der Brille versteckt, verändert sich das Klangbild systematisch mit der betrachteten Szene. Mit etwas Training können Blinde diese Informationen zur Orientierung mitnutzen.

Prozessüberwachung: Eine Vielzahl von Videoströmen kann meist nicht gleichzeitig betrachtet werden - man denke z.B. an 20 Überwachungsmonitore einer Intensivstation. Die Sonifikationen können aber mühelos überlagert werden, wobei markante Klangveränderungen die Aufmerksamkeit auf relevante Änderungen lenken können.

Muster-Entdeckung: In Videosonifikationen lassen sich Muster entdecken, die visuell schwerer zugänglich sind. Im Bereich der Sportwissenschaft haben wir z.B. ein Verfahren entwickelt, Videoaufnahmen einer Handballtrainingsituation zu verklänglichen. Hierbei werden die Positionen von sechs Angriffsspielern und Verteidigern aus dem Bild extrahiert und die Abweichung der Verteidiger von einem Soll-Verhalten entsprechend sonifiziert. Dies erlaubt dem Trainer über das Hören gleichzeitig eventuelles Fehlverhalten mehrerer Spieler zu entdecken, während in der visuellen Spielbeobachtung der Fokus doch durch die Blickrichtung eingeschränkt ist.

Navigationsunterstützung: Die Sonifikation von Videodaten kann helfen, die Navigation in langen Videoströmen zu erleichtern: bei einem schnellen interaktiven Durchscannen kann der Klang auf interessante Momente im Video hinweisen, und so das Suchen (Auf-/Wiederfinden) bestimmter Situationen beschleunigen. Als Datenmaterial für die hier vorgestellten Sonifikationen verwenden wir das in Abstimmung mit dem **Haus am Gern** ausgewählten und von Rudolf Steiner mit einer Webcam aufgezeichnete Videodaten an der Straße A28 im Prättigau. Die Idee besteht darin, den Verkehrsstrom als 'Atemfluß' des 'Organismus Stadt' zu verstehen und ein paar touristische Atemzüge dieser Art hörbar zu machen.

Abbildung 1 zeigt drei typische Einzelbilder der Videos. Das Auge erkennt mühelos eine Straße, Fahrzeuge in verschiedenen Fahrtrichtungen sowie das Wetter. Diese Erkennung bedeutungstragender Objekte ist für uns Menschen ganz selbstverständlich und mühelose. Man muss sich aber vergegenwärtigen, dass sich die Bilder für den Computer lediglich als eine Ansammlung von Helligkeitswerten darstellen, und das Sonifikationsverfahren zunächst nichts von oder über Objekte weis. Die Sonifi-



Abbildung 1: Einige typische Bilder des verwendeten Videomaterials. Die weißen Marker zeigen die Positionen der Vektorquantisierungs-Prototypen.

kation hat daher die Herausforderung, von den reinen Bilddaten ausgehend Strukturen zu offenbaren. Wir verzichten bewusst auf eine Objekterkennung und Spezialisierung, damit das Verfahren generisch auf Videodaten aller Art arbeitet.

Die direkte Verwandlung der Ausgangs-Daten in Klänge scheitert an ihrer schieren Masse: Jedes Einzelbild besteht schon aus ca. 1000000 Zahlen (Helligkeitsinformation zu jedem Bildpunkt und Farbkanal), und die gleichzeitige Verklänglichung in einer gleichhohen Anzahl von Klangelementen ist zum einen technisch nicht realisierbar, andererseits aber auch klanglich nicht sinnvoll. Wichtiger erster Schritt ist also eine Verdichtung auf eine weitaus geringere Anzahl informativer Elemente. Hierfür konzentrieren wir uns auf Änderungen zwischen aufeinanderfolgenden Bildern, und konzentrieren diese weiter durch Vektorquantisierung.

Sonifikation mittels Vektorquantisierung

Wir versuchen eine möglichst allgemeinverständliche Beschreibung unseres Verfahrens zur Datenaufbereitung und umgehen den Dickicht mathematischer Details. Ein qualitatives Verstehen ist jedoch sehr förderlich, um den entstehenden Klangstrom in Bezug auf die Videodaten zu interpretieren. Vektorquantisierung bezeichnet ein Verfahren, eine meist feste Anzahl von Prototypen so zu adaptieren, dass sie eine Datenverteilung in einem Merkmalsraum

nach gewissen Kriterien approximiert. Als Datenquelle verwenden wir das Differenzbild aufeinanderfolgender Videobilder, da unser Verfahren auf zeitliche Änderungen sensitiv sein soll. Als Stimuli verwenden wir diejenigen Bildpunkte, deren Helligkeitsänderung eine bestimmte Schwelle übersteigt. Anhand dieser Stimuli werden nun für jeden Zeitschritt die Prototypen der Vektorquantisierung adaptiert. Anschaulich zieht das Verfahren die Prototypen in Richtung von Änderungen im Bild. Fährt also ein Auto durchs Bild, folgt mit großer Wahrscheinlichkeit ein Prototyp diesem bewegten Objekt. Dass es sich hierbei um ein Objekt, oder gar um ein Auto handelt, kann das Verfahren natürlich nicht wissen. Durch dieses Verfahren reduziert sich die Datenmenge erheblich, nämlich auf die Positionen der Prototypen sowie einiger Merkmale an diesen Positionen. Wir bestimmen z.B. die Farbinformation am Ort der Prototypen, oder die Zahl der Stimuli, die einen Prototypen während der Adaption aktiviert. In Ergänzung dieser "lokalen" Informationsmerkmale bestimmen wir noch eine Reihe globaler Bildmerkmale, beispielsweise die Gesamthelligkeit und die Helligkeit in rot/grün/blau-Kanälen.

Auf diese Weise erhalten wir nun Informationsträger, die im folgenden auf akustische Elemente der Sonifikation abgebildet werden.

Die Gestaltung des Klanggefüges der Sonifikation ist eine eher künstlerisch kreative Arbeit, und zentral für den Sonifikationsprozess. Meist wird man in Überarbeitungs-Zyklen den Verbindungsalgorithmus zum Klang verändern. Das Zusammenkommen von Intuition, Klanggefühl, und vor allem Struktur-Wissen aus dem Datamining macht diesen Schritt zu einem hochgradig interdisziplinären Element. Wir stellen auf der CD mit einer sog. Multiple-Stream Sonifikation eine Realisation vor und konzentrieren uns in den unterschiedlichen Tracks auf die Variation des Kompressionsfaktors als einer der zentralen Determinanten zur Exploration rhythmischer Struktur. Daher klingen alle Sonifikationen im Prinzip ähnlich - sie könnten jedoch auch völlig anders klingen - wie Motorheulen, Vogelgesang, oder Bienenschwärme, je nach Design.

In unseren Sonifikationen werden vornehmlich kontinuierliche Klang-Elemente (sog. stationäre Oszillatoren) verwendet, deren Kontrollparameter durch die vorberechneten Merkmale angesteuert werden. Generell gibt es zwei Sorten von Information, globale und lokale. Die globale Bildhelligkeit und Farbverteilung wird durch die harmonische Untergrund-Klangtextur kommuniziert – je ein Ton steht für Rot/Grün/Blau-Anteil, die Pulsrate und Lautstärke variieren mit der Intensität. Dieser Klangtextur werden Klangereignisse für die lokalisierten Prototypen überlagert. Wir verwenden hier 15 Prototypen und erzeugen für jeden Prototyp einen Klang. Die horizontale Position im Bild wird durch die Stereoposition im Klangbild, die vertikale Position im Bild durch die Frequenz (oder: Tonhöhe) definiert. Die Amplitude (oder Lautstärke) hängt dagegen von der Größe der lokalen Bild-Änderung ab. Somit klingt ein sich von rechts oben nach links unten bewegendes Prototyp wie ein tonhöhe-abfallender Puls, der sich im Klangraum von rechts nach links verschiebt. Plötzlich werden also aufeinanderfolgende systematische Variationen zu Klangkonturen mit einer eigenen akustischen Gestalt. Im Detail verstecken sich noch weitere Modifikationen, wie z.B. Frequenzfilter, die je nach Farbkomposition unterschiedliche Klangscharfe liefern.

Wir belassen es nun bei diesem noch recht einfachen Verfahren, um in den folgenden Klangbeispielen Höreindrücke zu diskutieren.

Klangbeispiele

Für die CD haben wir 6 Tracks komponiert, die unterschiedliche Aspekte der Videodaten-sonifikationen beleuchten (eigentlich: betonen). Track 1 und 2 stellen je 30 min Verkehrsstrom mit geringfügiger zeitlicher Kompression in 5 min dar. Der zeitliche Kompressionsfaktor ist 6. Der Datenteil der CD enthält ein Video, dem die Sonifikation unterlegt ist. Hier sieht man den Verkehrsstrom und überlagert die aus der Vektorquantisierung berechneten Prototypen. Man hört klar die Bewegungsrichtung der Fahrzeuge heraus. Track 1 ist ein Ausschnitt des Nachtverkehrs - die globale Helligkeit ist nahezu verschwindend. Anders im Track 2, wo die harmonische Struktur klar hörbar wird.

Im Klangbeispiel 3 erhöhen wir die Kompression auf 20, um 2 Stunden Verkehrsfluss in 10 min zu sonifizieren. Das Interessante ist, dass hier die Morgendämmerung gewählt wurde, so dass die Veränderung der Bildhelligkeit im Zeitraffer gehört werden kann. Interessant ist ferner, wie sich die Verkehrsdichte bei Tagesanbruch verändert.

Für noch stärkere zeitliche Kompressionen verdichten sich die Prototypenklänge derartig, dass eine fast schon chaotische Textur entsteht. Mit dem hier gewählten Verfahren wird damit leider die Fahrtrichtung unhörbar. Dies zeigt, dass auch das Design auditiver Displays ein dynamischer Prozess ist, und praktisch würde man hier nun das Verfahren selbst modifizieren, wenn man spezifisches Interesse an dieser Information hat. Klangbeispiel 4 und 5 bilden den Gesamtzeitraum des Videomaterials von Freitag Abend bis Sonntagnachmittag ab. Mit zunehmender Kompression verwandelt sich der Tag-Nacht-Rhythmus zu einem Klang-Rhythmus.

Track 6 stellt schließlich 'zum Ausklang' in einen längeren Zeitraum von 15 min. einen ganzen Tag dar. Die Länge ist nicht untypisch für musikalische Stücke, und stellt somit ein allmähliches Struktur-Veränderungshören in den Vordergrund. Es erlaubt zudem, sich langsam in die Sonifikation einzuhören.

Diskussion und Ausblick

Wir hoffen, dass dieses Sonifikationsbeispiel die Grundideen der Sonifikation, die vielfältigen Anwendungsgebiete und die prinzipielle Arbeitsweise illustriert hat. Videodaten wurden gewählt, da sie doch zunächst sehr anschaulich erscheinen, jedoch auch rhythmische Strukturierungen enthalten, in denen die Sonifikation komplementäre Aspekte herausarbeiten kann. Diese Informationen ließen sich freilich auch visuell darstellen, jedoch ist es schwierig, gleichzeitig das Video und die Visualisierung zu betrachten, während die Verteilung auf unterschiedliche Modalitäten die ideale Kombinierbarkeit beider illustriert.

Unsere Welt ist von Veränderung bestimmt. Änderungen sind oft rhythmisch, und das Gehör ein ideales Organ, um Rhythmen wahrzunehmen, und sogar zeitliche Änderungen von Rhythmen zu erkennen. Voraussetzung ist natürlich, dass die zeitliche Struktur an unsere Wahrnehmungsbereich angepasst wird: Gerade durch die Veränderung der zeitlichen Kompression ergeben sich ganz unterschiedliche Strukturen. In unseren Sonifikationen hört man den Tag-Nacht-Rhythmus - über das Hören würde man leicht die saisonale Änderung der Tag/Nacht-Dauern wahrnehmen. Sicher auch die höhere Aktivität durch Feiertage,

Wochenenden, Ferienanfang, etc. Auf Autobahnen wäre interessant, ob sich Verdichtungswellen oder sogar Vorboten für das Erscheinen von Staus akustisch entdecken lassen.

In zwei Projekten untersuchen wir Videosonifikation mit ganz praktischen Zielen: in einer Zusammenarbeit mit der Verhaltensforschung der Universität Bielefeld (Prof. Trillmich, Anke Rehling) sonifizieren wir Videos von Meerschweinchen, um so die Analyse der Veränderung von Saugverhalten Neugeborener und anderer sozialer Aspekte über einen Zeitraum von 4 Wochen nach der Geburt zu unterstützen. Im Vergleich zur traditionellen Analyse (studentenintensive Videobeobachtung) erwarten wir eine Beschleunigung, sowie die Möglichkeit, zuvor unentdeckte Eigenschaften zu finden.

Eine weitere Projekt-Richtung zielt auf einen Echtzeit-Einsatz von Sonifikation. Über eine Webcam aufgenommene Bilder werden unmittelbar in Klänge verwandelt. Blinde Menschen erhalten somit die Möglichkeit interaktiv ihre Umgebung abzutasten (eine Taschenlampe für Blinde). Ein erster Prototyp existiert bereits, jedoch steht eine Klangoptimierung und intensive Tests mit blinden Nutzern noch aus.

Während sich allein aus dem Bereich der Videosonifikation vielversprechende Anwendungen ergeben, ist die allgemeine Aussicht für den Nutzen von Sonifikation in der Wissenschaft ungleich größer. Neben der Entwicklung von Techniken zur Unterstützung von Datenanalyse beliebiger Daten, verfolgen wir spezialisierte Anwendungen im Bereich EEG-Datenanalyse, Sonifikation psychotherapeutischer Gesprächsprotokolle, Wettersonifikation, Diagnose komplexer Netzwerke, Robotersonifikation, Börsendatenonifikation.

Sonifikation ermöglicht, dass wir die exzellenten und hochentwickelten Mustererkennungsfähigkeiten des Gehörs zur Untersuchung abstrakter und komplexer Daten einsetzen können. Wir sind zuversichtlich, dass Sonifikation in Zukunft eine in verschiedenen Bereichen größere Rolle spielen wird. Man wird hören.

Dankworte

Unser besonderer Dank gilt dem Haus am Gern, Rudolf Steiner und Barbara Meyer Cesta für ihr Engagement, ihr Interesse an Sonifikation, und ihren Beitrag zur Verwirklichung: das verwendete Videodatenmaterial wurde von ihnen extra für die Produktion aufgezeichnet. Till Bovermann danken wir für die Unterstützung und viele Anregungen, sowie einen Vorschlag zum Design des CD-Labels. Ferner bedanken wir uns bei Florian Dombois für die Betrachtung aus künstlerischer Perspektive.