

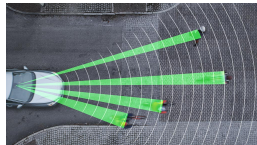
Adaptive Gestenerkennung mit Variationsabschätzung für interaktive Systeme

Maxim Boianetchii Marian Stein

9. Juli 2015

Motivation

- Gestenerkennung in vielen Gebieten gefragt:
 - ▶ Medizin
 - ▶ Automobilindustrie
 - ▶ Unterhaltungsbranche
- Mit aktuellen Methoden nur eingeschränkt möglich
 - ▶ Erkennung teilw. nur nach vollst. Ausführung der Geste
 - ▶ Keine Rückmeldung von Zusatzinformationen über die Geste(z.B. Geschwindigkeit)



Vorgeschlagene Methode

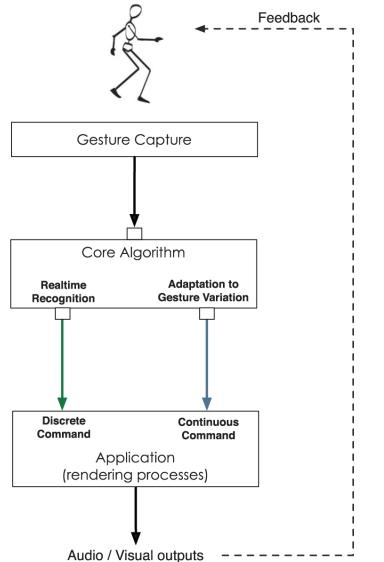
- Verwendung eines Partikelfilters
- frühzeitige Erkennung und Rückmeldung von Variationsinformationen, z. B. Geschwindigkeit, Drehung, etc.
 - ▶ Ermöglicht Anwendungen, bei denen die Benutzer direkt während der Gestenausführung interagieren können

Ähnliche Arbeiten

- Templatebasierte Erkennung (\$1 Recognizer) (Wobbrock et al. [2007]):
 - ▶ Vorverarbeitung der aufgenommenen Daten
 - ▶ Erkennung mithilfe euklidischen Abstands
- Dynamic Time Warping (DTW) (D.M.Gavrila and L.S.Davis [1995], Liu et al. [2009])
 - ▶ Erfassung der gesamten Geste
 - ▶ Anpassung durch Strecken/Stauchen der Vorlage
- Kondensationsalgorithmus (M.Black and A.Jepson [1998])
 - ▶ gleiche Grundidee wie der vorgeschlagene Algorithmus
 - ▶ Anpassung auf Skalierung beschränkt

Interaktionsprinzipien

- neue Interaktionsmöglichkeiten:
 - ▶ Tonmanipulation
 - ▶ Videospiele
 - ▶ etc.
- zwei grundlegende Interaktionen:
 - ▶ Ausführung/Festlegung der Geste
 - ▶ Manipulation der Geste während der Ausführung
- Erkennung der Geste und Abschätzung der Variationen in Echtzeit und kontinuierliche Aktualisierung
- Verwendung eines einzigen Templates pro Geste



zugrundeliegendes Modell

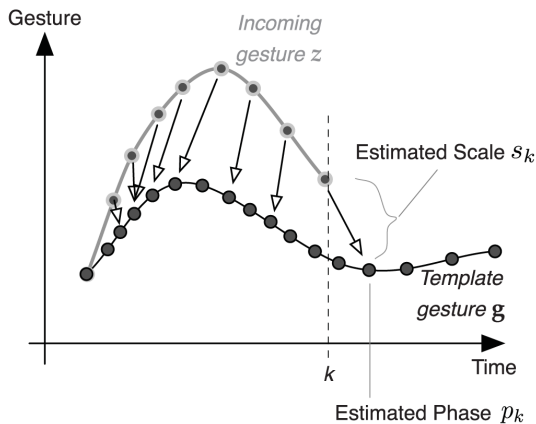
- Geste: Gliedmaßenbewegung, die durch eine Zeitserie einer festen Anzahl Parameter vertreten wird
- Zustandsmodell zum Zeitpunkt k :

$$\begin{cases} x_k = f_{TR}(x_{k-1}, v_{k-1}) \\ z_k = f_{OB}(x_k, w_k; g) \end{cases} \quad (1)$$

- ▶ x_k : Vektor des Systemzustands, mit den Gestenparametern als Elemente
- ▶ f_{TR} Funktion für die Entwicklung des Zustands, abhängig vom vorherigen Zustand und der Abweichungssequenz v_k
- ▶ f_{OB} Funktion, die aus den Messwerten w_k , dem vorherigen Zustand und einer Templategeste g Beobachtungen generiert

zugrundeliegendes Modell

- Die ersten 2 Elemente des Zustandsvektors x_k werden festgelegt als Phase und Geschwindigkeit, weitere Elemente können weitere Parameter enthalten
- f_{TR} ist linear und gaussverteilt
- f_{OB} wird durch eine Student'sche t-Verteilung modelliert

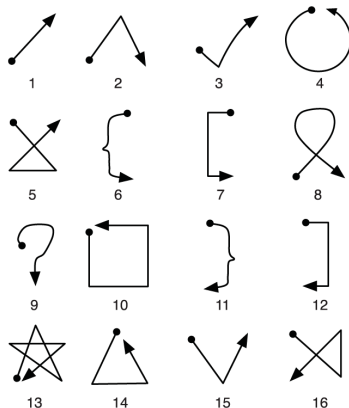


zugrundeliegendes Modell

- Inferenz des Zustandsvektors mithilfe eines Partikelfilters
 - ▶ Zustand wird durch Gewichtung vieler, in diesem Fall gaussverteilter, Partikel abgeschätzt
 - ▶ jedes Partikel repräsentiert einen möglichen Zustand und wird mit seiner Wahrscheinlichkeit gewichtet
 - ▶ Gewichtung der Partikel wird durch f_{OB} beeinflusst
- Der erwartete Gesamtzustand ist dann durch die gewichtete Summe aller Partikel gegeben
- Um zwischen verschiedenen Gesten zu unterscheiden, wird der Zustandsraum um einen Gestenindex erweitert und die Partikel gleichmäßig über alle Gestenindizes verteilt.

Erkennung realer 2D-Gesten

- Gestendatenbank von Wobbrock et al. [2007]:
 - ▶ 2D-Stiftgesten
 - ▶ 16 Gesten * 10 Teilnehmer * 3 Geschwindigkeiten * 10 Versuche = 4800 aufgenommene Gesten
- jeweils ein Datensatz als Template und einer als gemessene Geste
- 4 Versuche:
 - ▶ Erkennung Gesten gleicher Geschwindigkeit
 - ▶ Einfluss der Verteilungsparameter auf die Erkennung
 - ▶ Erkennung bei Beispielen verschiedener Geschwindigkeiten
 - ▶ Früherkennung



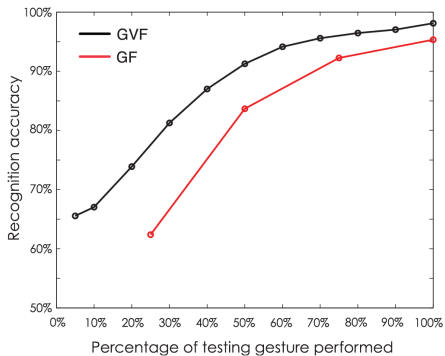
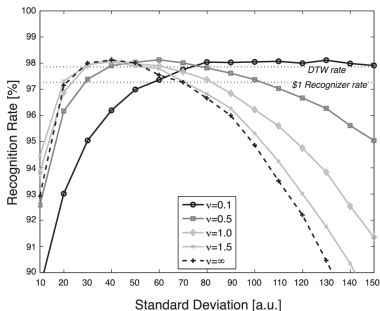
Erkennung realer 2D-Gesten

Testergebnisse

	\$1 recognizer	DTW	GF	GVF
	offline operated after scaling and rotation estimation		online no adaptation of scaling neither rotation	online incremental adaptation of scaling and rotation
Mean	97.27 %	97.86 %	95.78 %	98.11 %
Std	2.38 %	1.76 %	2.06 %	2.35 %

Our model has the following parameterization: $\sigma = 130$, $\nu = 0.1$.

		TRAINING EXAMPLES					
		Slow		Medium		Fast	
		\$1	GVF	\$1	GVF	\$1	GVF
TESTING EXAMPLES	Slow	96.2%	97.3%	94.9%	93.2%	91.6%	85.9%
	Medium	96.1%	95.4%	97.1%	98.6%	94.1%	96.6%
	Fast	92.9%	88.3%	94.6%	97.5%	95.5%	98.2%

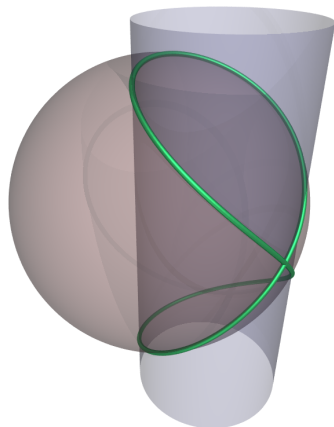


Abschätzung der Varianz anhand synthetischer Daten

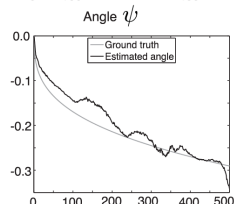
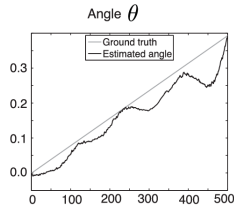
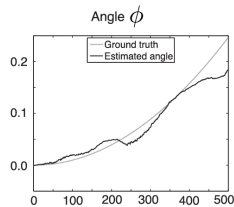
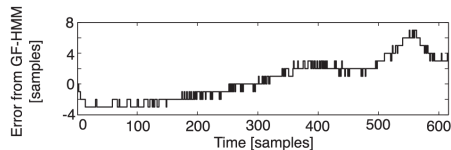
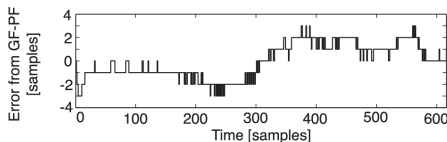
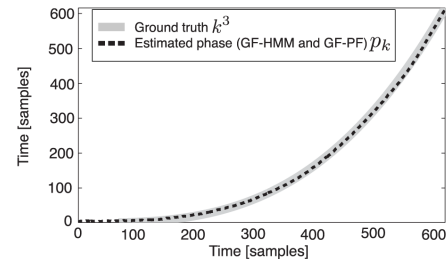
- Test der Varianzabschätzung
- synthetische Daten, da Menschen keine exakte Varianz liefern können
- Gesten in Form einer Viviani-Kurve:

$$C(t) = \begin{cases} x(t) = a(1 + \cos(t)) \\ y(t) = a \sin(t) \\ z(t) = 2a \sin(t/2) \end{cases} \quad (2)$$

- ▶ Vorlage lineare Abtastung der Kurve
- ▶ Testgeste kubische Abtastung + gleichmäßig verteiltes Rauschen

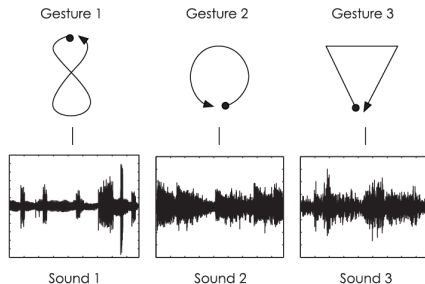


Abschätzung der Varianz anhand synthetischer Daten



Nutzerstudie

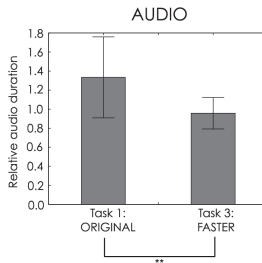
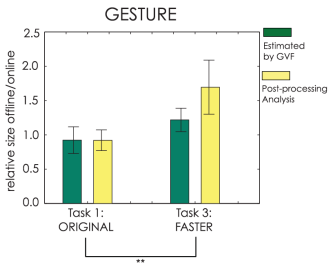
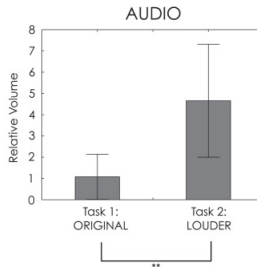
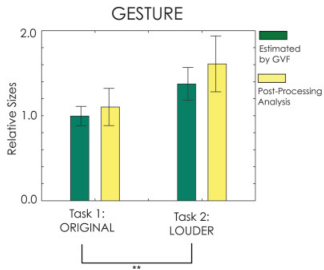
- Untersuchung der Einbindung in Anwendungen
- Gestenausführung startet Wiedergabe eines Tons, Variationen der Geste verändern ihn
 - ▶ Geste gilt als erkannt, wenn Wahrscheinlichkeit $> 50\%$
 - ▶ schneller/langsamer \rightarrow schnellere/langsamere Wiedergabe
 - ▶ Gestengröße \rightarrow Lautstärke
 - ▶ Gestendrehung \rightarrow Hochpassfilter
- Nachberechnung der tatsächlichen Varianzen zum Vergleich mit den geschätzten



- Teilnehmer sollten:
 - ▶ Geste unverändert ausführen
 - ▶ Geste mit einem veränderten Parameter ausführen
 - ▶ ein Parameter während der Ausführung verändern
 - ▶ Geste mit zwei veränderten Parametern ausführen

Nutzerstudie

Ergebnisse mit einem veränderten Parameter

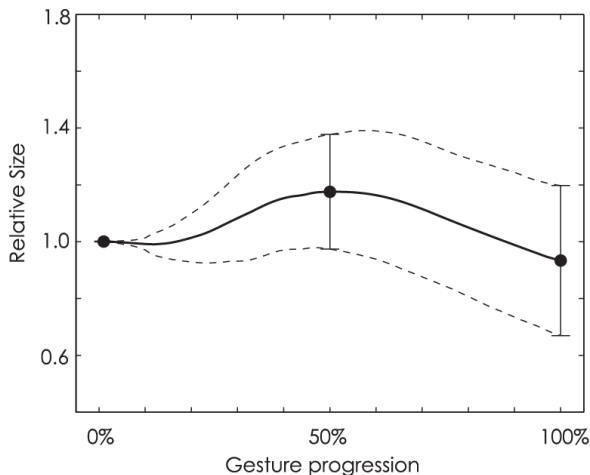


Nutzerstudie

Ergebnisse mit einem kontinuierlich veränderten Parameter

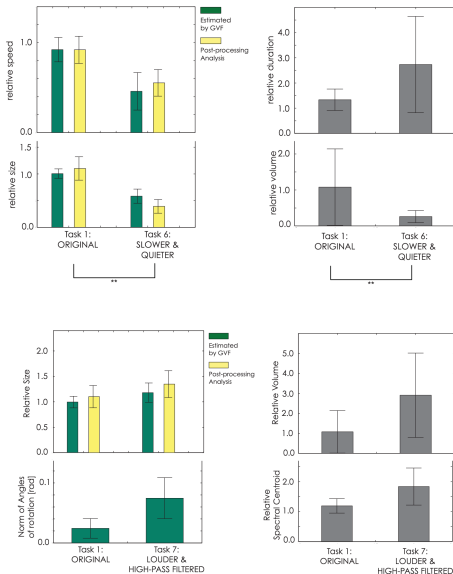
- Teilnehmer sollten Ton zunächst laut spielen, dann leiser werden

Temporal Estimation of the Size by GVF



Nutzerstudie

Ergebnisse mit zwei geänderten Parametern



- Algorithmus zur Erkennung von Gesten, der Variationen in Echtzeit mitteilt.
- im Vergleich durchweg bessere Ergebnisse als andere Algorithmen aus dieser Klasse
- Erkennung starker Variationen der Gesten ohne Beispiele der Variationen
- frühzeitige Erkennung
- somit starke Verbesserung gegenüber anderen Methoden

- D.M.Gavrila and L.S.Davis. Towards 3-d model-based tracking and recognition of human movement: A multi-view approach. *Proceedings of the International Workshop on Automatic Face and Gesture recognition*, 1995.
- J. Liu, L. Zhong, J. Wickramasuriya, and V. Vasudevan. Uwave: Accelerometer-based personalized gesture recognition and its applications. *Pervasive and Mobile Computing* 5, 2009.
- M.Black and A.Jepson. A probabilistic framework formatching tempora ltrajectories:condensation based recognition of gestures and expressions. *Proceedings of the European Conference on Computer Vision (ECCV'98)*, 1998.
- O. Wobbrock, A. D. Wilson, and Y. Li. Gestures without libraries, toolkits or training: A \$1 recognizer for user interface prototypes. *Proceedings of the 20th Annual ACM Synopsium on User Interface Software and Technology*, 2007.