Unique Identification of 50,000+ Virtual Reality Users from Head & Hand Motion Data

Vivek Nair · Wenbo Guo · Justus Mattern · Rui Wang · James F. O’Brien · Louis Rosenberg · Dawn Song

https://doi.org/10.48550/arXiv.2302.08927

Contributors

Vivek C. Nair
UC Berkeley

Wenbo Guo
UC Berkeley

Justus Mattern
RWTH Aachen

Rui Wang
UC Berkeley

James F. O'Brien
UC Berkeley

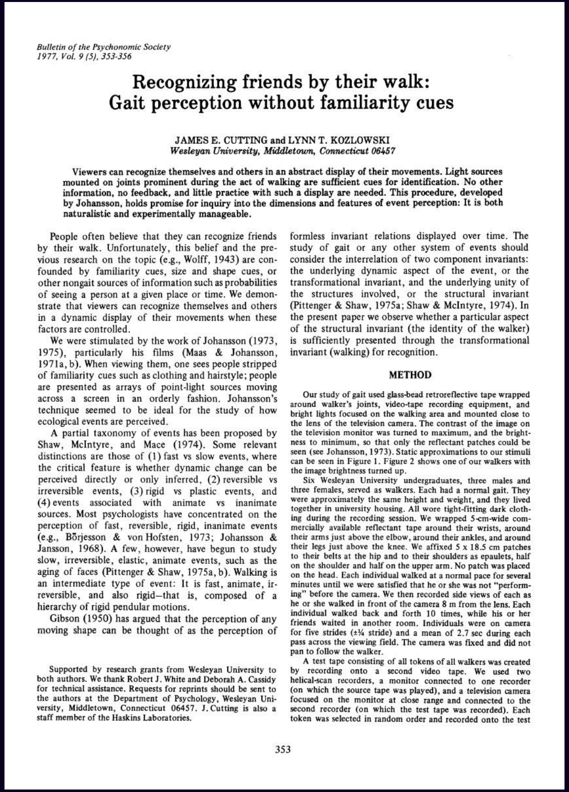
Louis Rosenberg
Unanimous AI

Dawn Song
UC Berkeley

Acknowledgments

* Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of their employer or the supporting entities.

Identity in motion



Bulletin of the Psychonomic Society
1977, Vol. 9 (5), 353-356

Recognizing friends by their walk: Gait perception without familiarity cues

JAMES E. CUTTING and LYNN T. KOZLOWSKI
Wesleyan University, Middletown, Connecticut 06457

Viewers can recognize themselves and others in an abstract display of their movements. Light sources mounted on joints prominent during the act of walking are sufficient cues for identification. No other information, no feedback, and little practice with such a display are needed. This procedure, developed by Johansson, holds promise for inquiry into the dimensions and features of event perception: It is both naturalistic and experimentally manageable.

People often believe that they can recognize friends by their walk. Unfortunately, this belief and the previous research on the topic (e.g., Wolff, 1943) are confounded by familiarity cues, size and shape cues, or other nongait sources of information such as probabilities of seeing a person at a given place or time. We demonstrate that viewers can recognize themselves and others in a dynamic display of their movements when these factors are controlled.

We were stimulated by the work of Johansson (1973, 1975), particularly his films (Maas & Johansson, 1971a, b). When viewing them, one sees people stripped of familiarity cues such as clothing and hairstyle; people are presented as arrays of point-light sources moving across a screen in an orderly fashion. Johansson's technique seemed to be ideal for the study of how ecological events are perceived.

A partial taxonomy of events has been proposed by Shaw, McIntyre, and Mace (1974). Some relevant distinctions are those of (1) fast vs slow events, where the critical feature is whether dynamic change can be perceived directly or only inferred, (2) reversible vs irreversible events, (3) rigid vs plastic events, and (4) events associated with animate vs inanimate sources. Most psychologists have concentrated on the perception of fast, reversible, rigid, inanimate events (e.g., Bőrjesson & von Hofsten, 1973; Johansson & Jansson, 1968). A few, however, have begun to study slow, irreversible, elastic, animate events, such as the aging of faces (Pittenger & Shaw, 1975a, b). Walking is an intermediate type of event: It is fast, animate, irreversible, and also rigid—that is, composed of a hierarchy of rigid pendular motions.

Gibson (1950) has argued that the perception of any moving shape can be thought of as the perception of formless invariant relations displayed over time. The study of gait or any other system of events should consider the interrelation of two component invariants: the underlying dynamic aspect of the event, or the transformational invariant, and the underlying unity of the structures involved, or the structural invariant (Pittenger & Shaw, 1975a; Shaw & McIntyre, 1974). In the present paper we observe whether a particular aspect of the structural invariant (the identity of the walker) is sufficiently presented through the transformational invariant (walking) for recognition.

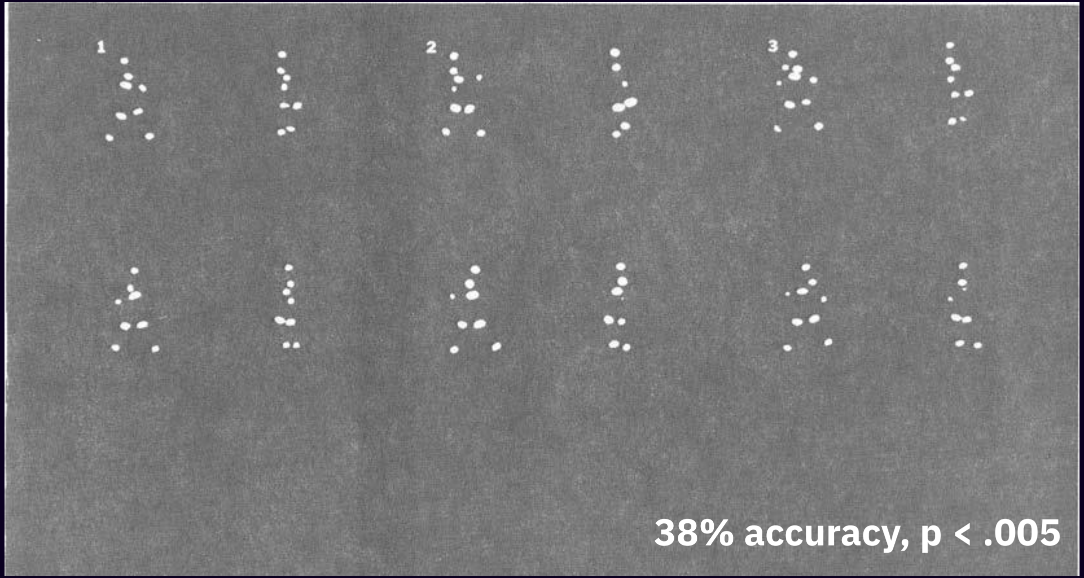
Supported by research grants from Wesleyan University to both authors. We thank Robert J. White and Deborah A. Cassidy for technical assistance. Requests for reprints should be sent to the authors at the Department of Psychology, Wesleyan University, Middletown, Connecticut 06457. J. Cutting is also a staff member of the Haskins Laboratories.

METHOD

Our study of gait used glass-bead retroreflective tape wrapped around walker's joints, video-tape recording equipment, and bright lights focused on the walking area and mounted close to the lens of the television camera. The contrast of the image on the television monitor was turned to maximum, and the brightness to minimum, so that only the reflectant patches could be seen (see Johansson, 1973). Static approximations to our stimuli can be seen in Figure 1. Figure 2 shows one of our walkers with the image brightness turned up.

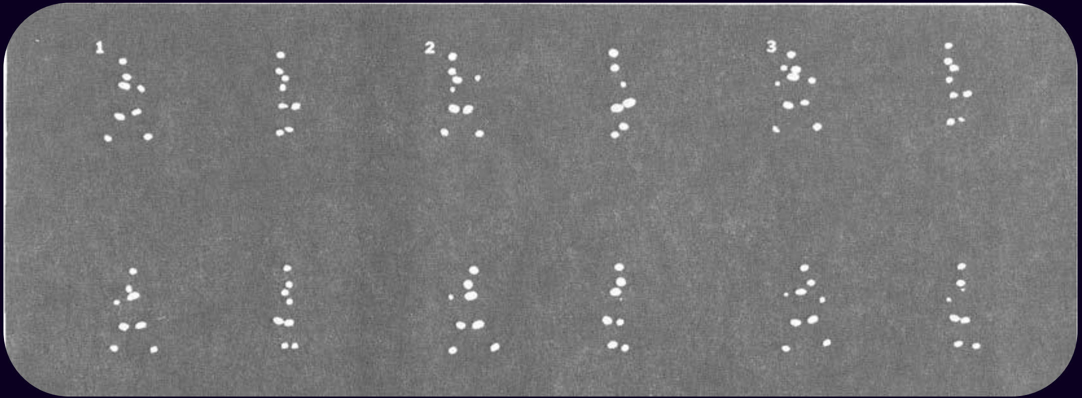
Six Wesleyan University undergraduates, three males and three females, served as walkers. Each had a normal gait. They were approximately the same height and weight, and they lived together in university housing. All wore tight-fitting dark clothing during the recording session. We wrapped 5-cm-wide commercially available reflectant tape around their wrists, around their arms just above the elbow, around their ankles, and around their legs just above the knee. We affixed 5 x 18.5 cm patches to their belts at the hip and to their shoulders as epaulets, half on the shoulder and half on the upper arm. No patch was placed on the head. Each individual walked at a normal pace for several minutes until we were satisfied that he or she was not "performing" before the camera. We then recorded side views of each as he or she walked in front of the camera 8 m from the lens. Each individual walked back and forth 10 times, while his or her friends waited in another room. Individuals were on camera for five strides (±¼ stride) and a mean of 2.7 sec during each pass across the viewing field. The camera was fixed and did not pan to follow the walker.

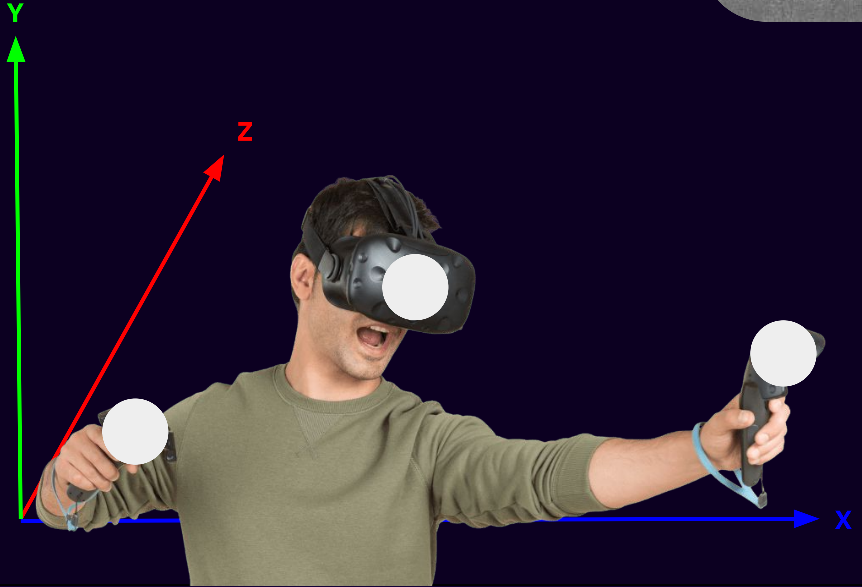
A test tape consisting of all tokens of all walkers was created by recording onto a second video tape. We used two helical-scan recorders, a monitor connected to one recorder (on which the source tape was played), and a television camera focused on the monitor at close range and connected to the second recorder (on which the test tape was recorded). Each token was selected in random order and recorded onto the test

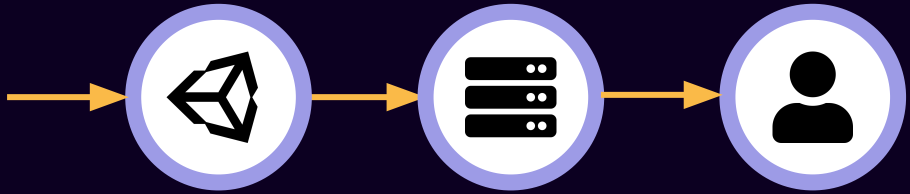


Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9(5), 353–356. https://doi.org/10.3758/BF03337021

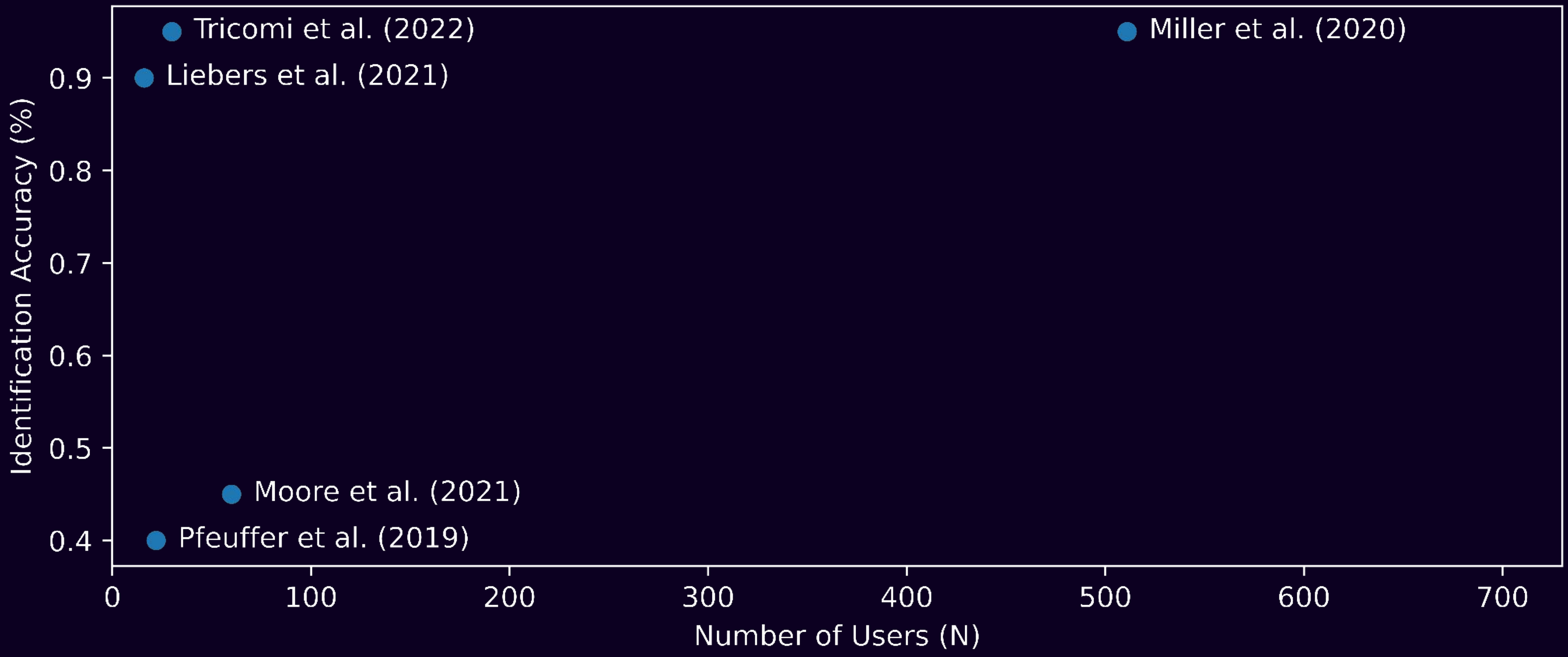
Motion in VR



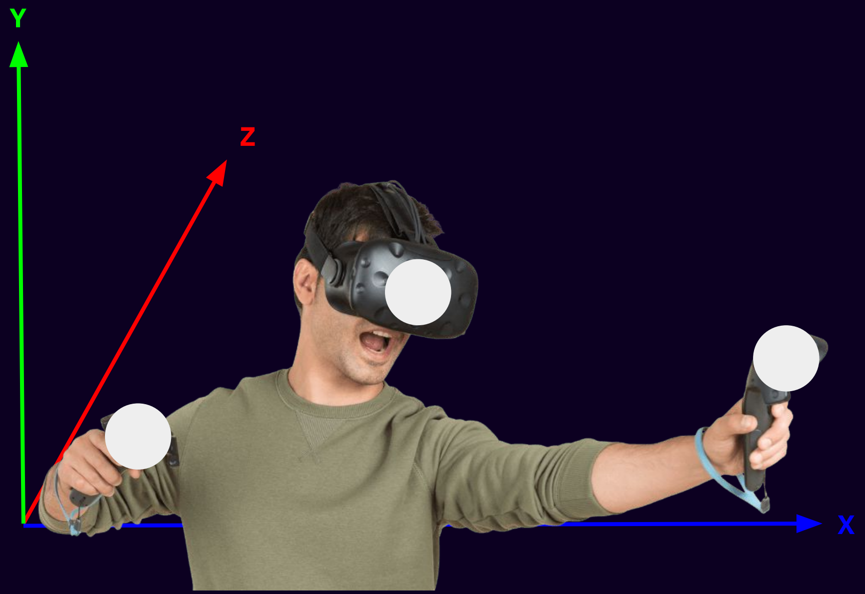


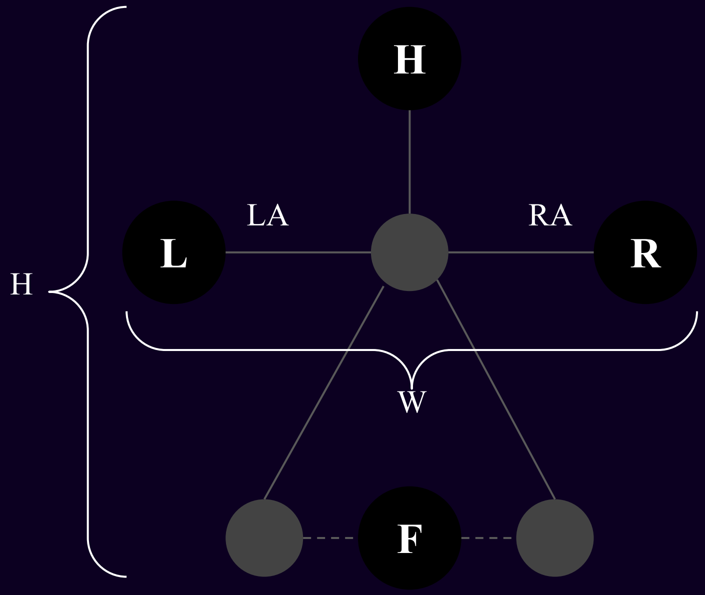


VR identification studies



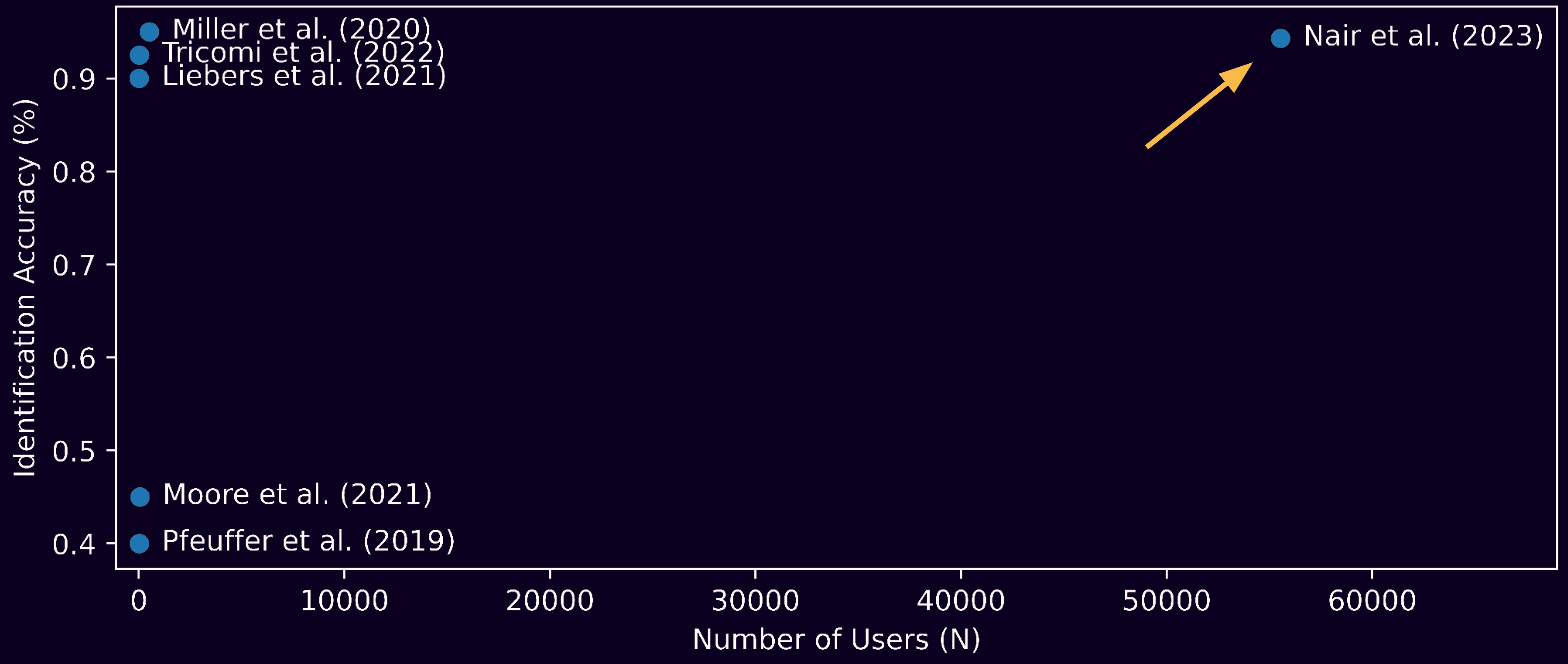
Static identifiers





* Enough to identify ≈4,000 users

VR identification studies



Believer Imagine Dragons Rustic rustic

Ranked 1.44★ Expert

Hide filters and sorting

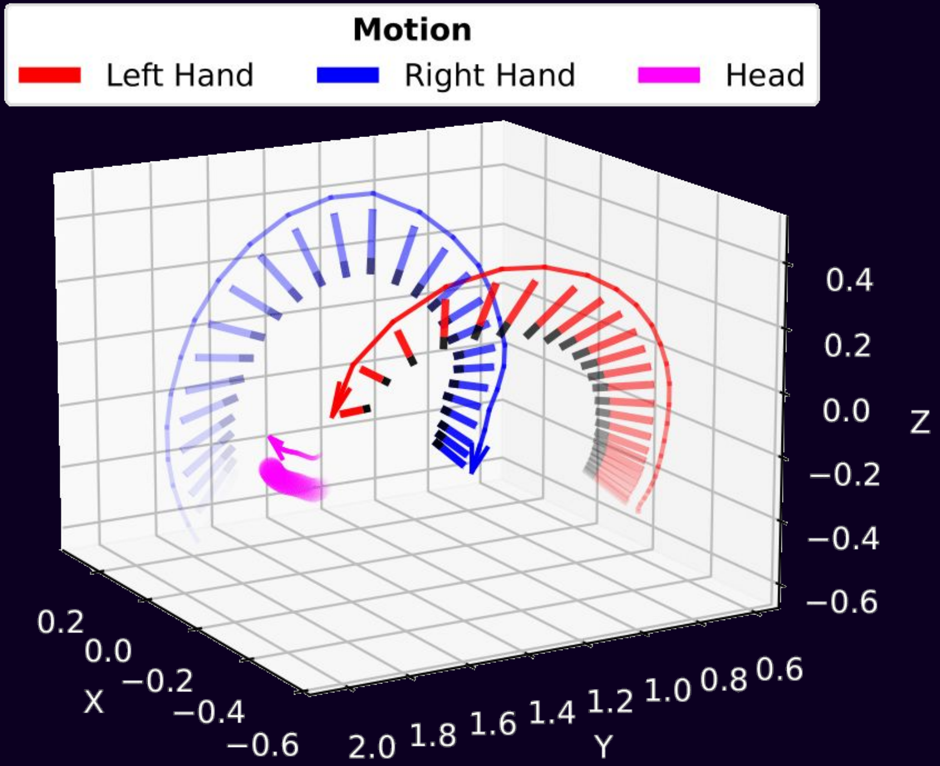
Rank | Acc | Recent | 115 Streak | Pauses | Play count

Rank	Player	Clan	When	pp	Acc	Score
#1	Marsh	LUCK	2 months ago	122.88pp	97.34%	481,912 SF
#2	MochiPower123		4 months ago	120.33pp	97.14%	480,915 DA, SF
#3	Pajima	ART LUCK	5 months ago	118.74pp	97.01%	480,258 DA, SF
#4	CheezitYT		2 months ago	106.03pp	95.13%	470,957 SF, GN
#5	Hunter	MEEP ST	6 months ago	105.34pp	95.64%	473,481 DA, SF
#6	Dark_Jimmy		5 months ago	103.39pp	95.39%	472,275 DA, SF
#7	Lautrick98	SKIL RBKS	3 months ago	102.54pp	95.28%	471,731 SF
#8	Cohenlinklater	BSC	3 months ago	102.04pp	94.59%	468,271 SF, GN
#9	Emerald knight YT	RAGE FROG	3 months ago	101.23pp	94.47%	467,686 SF, GN
#10	OEN | Ani	BSC OEN	last month	101.05pp	95.09%	470,746 SF

1 - 10 / 1470 1 2 3 4 5 6 7 8 9 10 ... 147

Show more details

BeatLeader dataset



USERS
55,541

REPLAYS
2,669,886

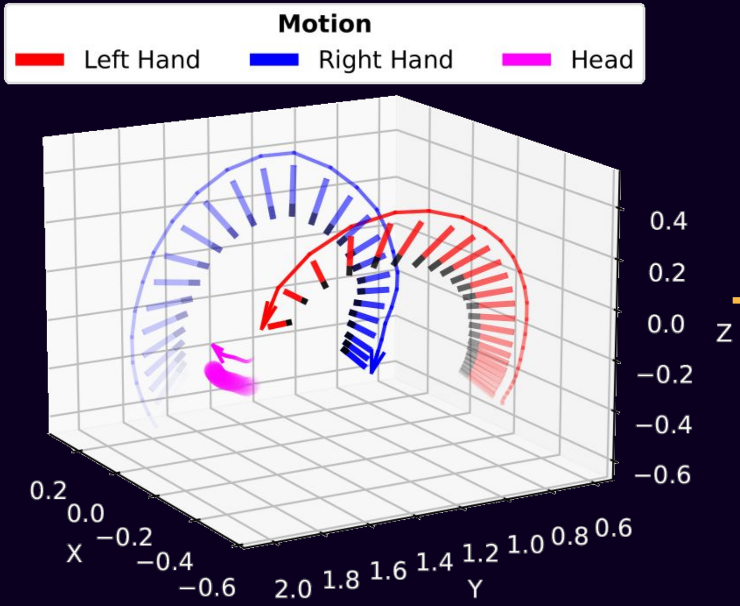
COUNTRIES
40+

SESSIONS
713,013

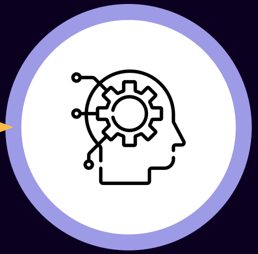
VR DEVICES
20+

DATA
3.96 TB

Motion features



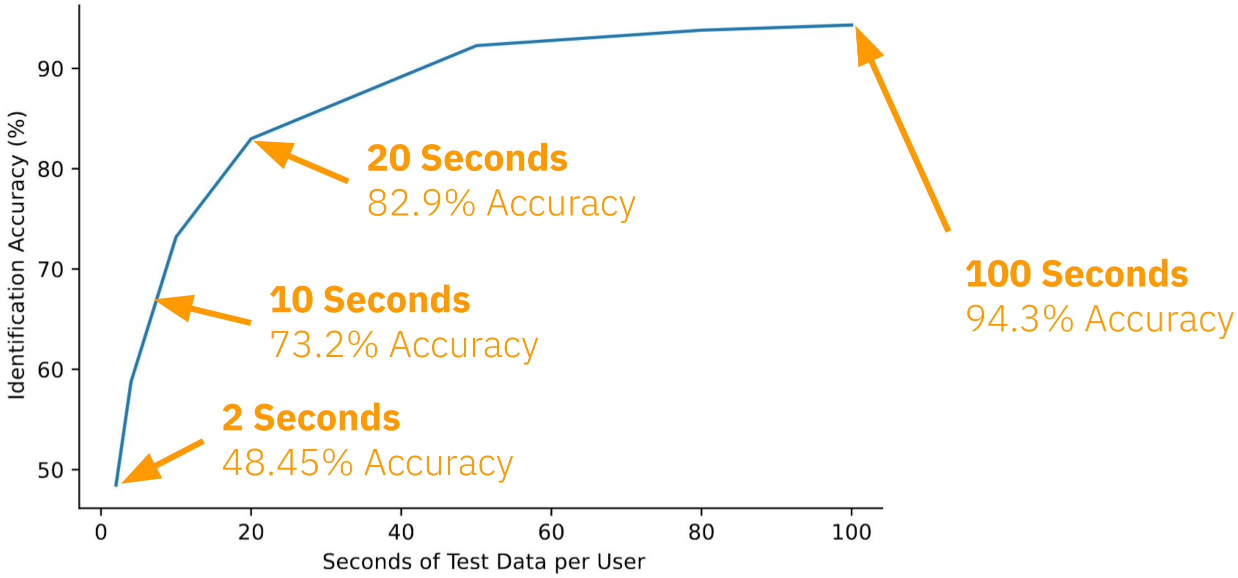
$$\{pos_x, pos_y, pos_z, rot_i, rot_j, rot_k, rot_l\}$$
$$\times$$
$$\{\mathrm{min}, \mathrm{max}, \mathrm{mean}, \mathrm{med}, \mathrm{stdev}\}$$
$$\times$$
$$\{\mathrm{head}, \mathrm{left\_hand}, \mathrm{right\_hand}\}$$
$$=$$
$$[105 \text{ dimensions}]$$



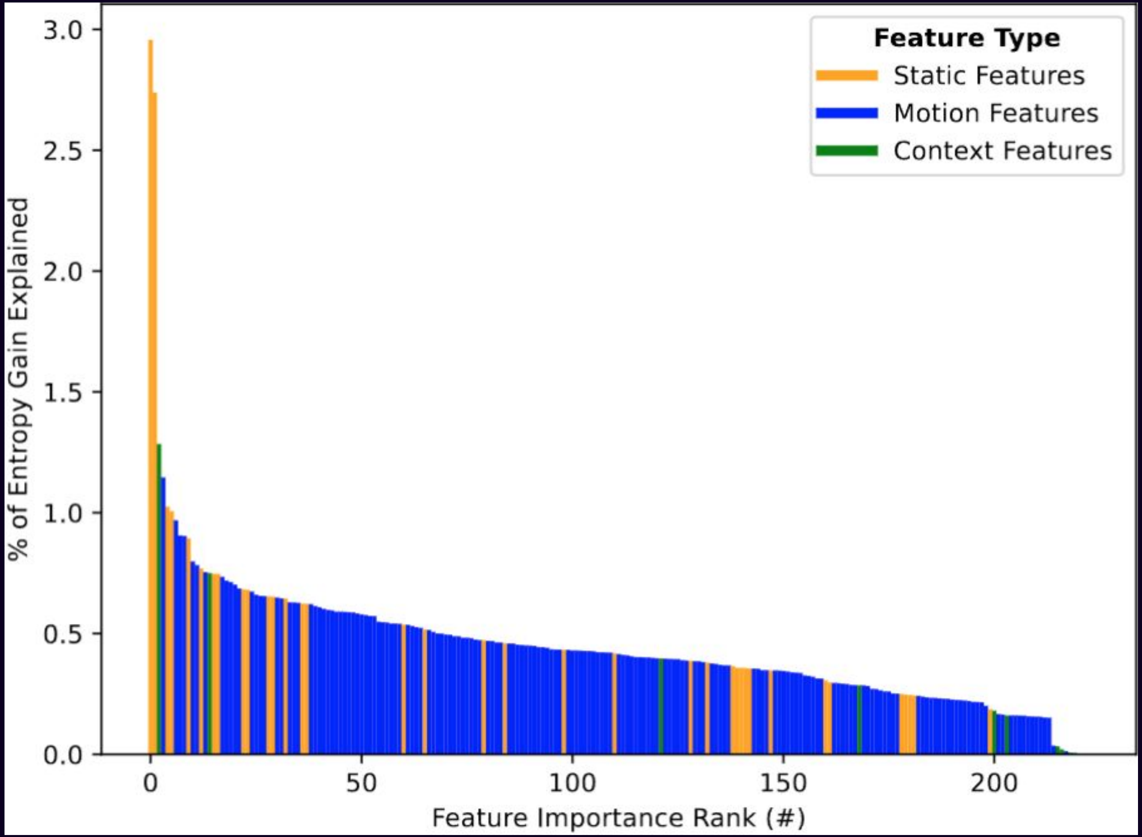
LightGBM

Results

Layer	# of Models	Accuracy (per Model)	Accuracy (per Layer)
Layer 1	10	93.1%	90.2%
Layer 2	10	93.1%	90.2%
Layers 1 & 2	20	93.1%	91.0%
Layer 3	5	84.0%	84.0%
Layers 1, 2, & 3	25	91.3%	94.3%



Model explanations



STATIC FEATURES

22.9% of entropy gain

Enough to identify ≈ 4,000 users

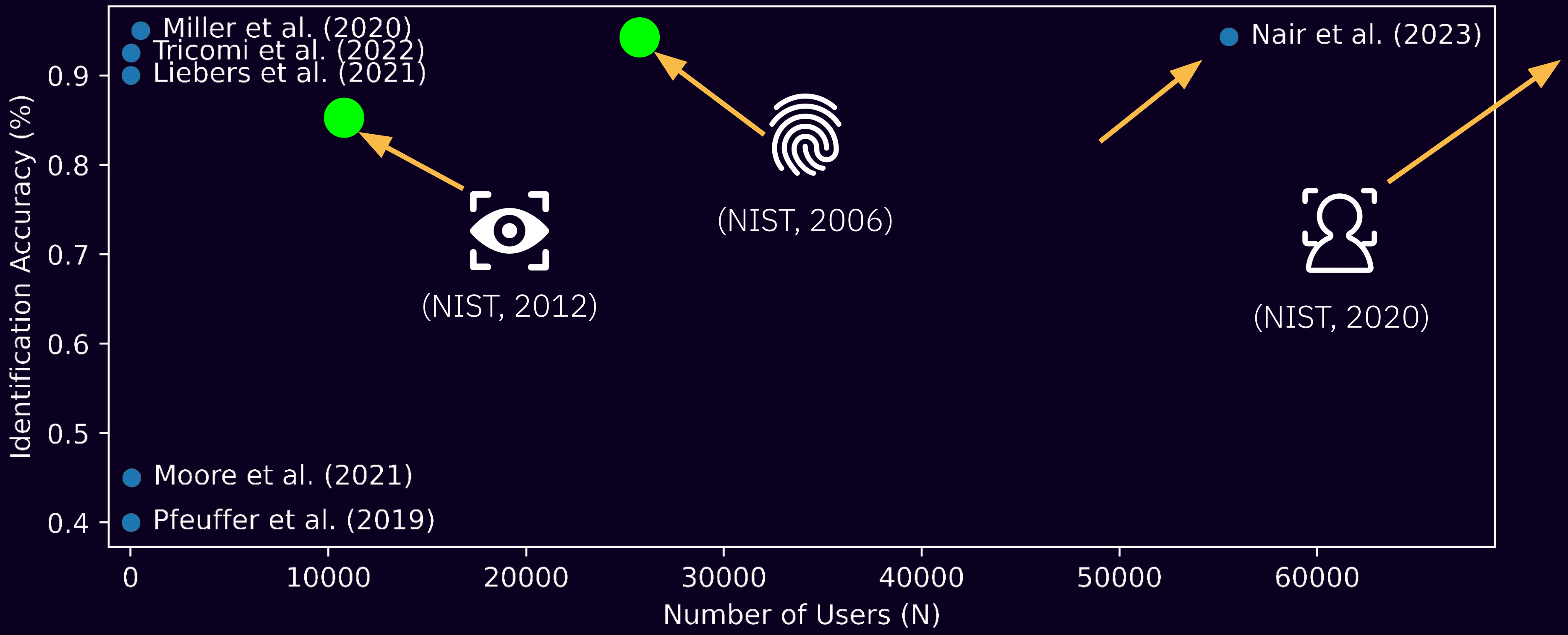
MOTION FEATURES

73.9% of entropy gain

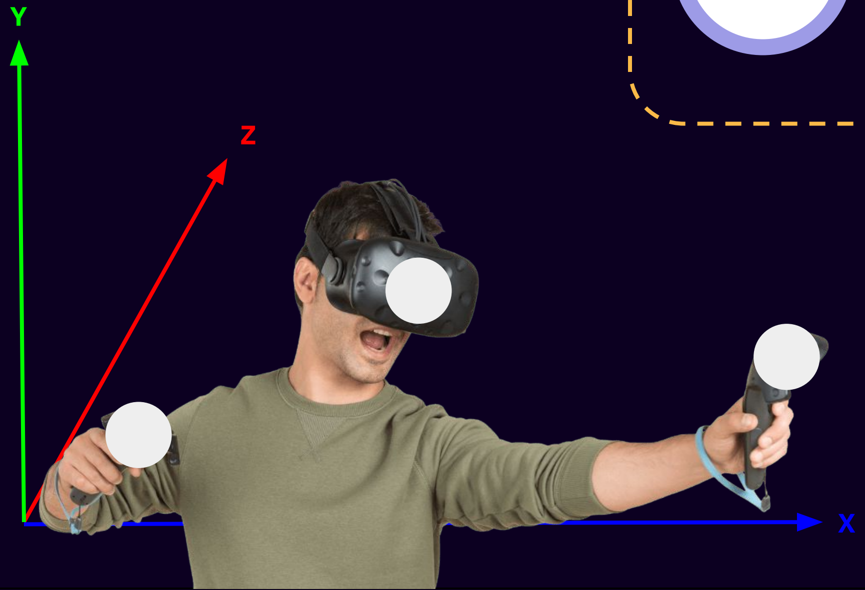
CONTEXT FEATURES

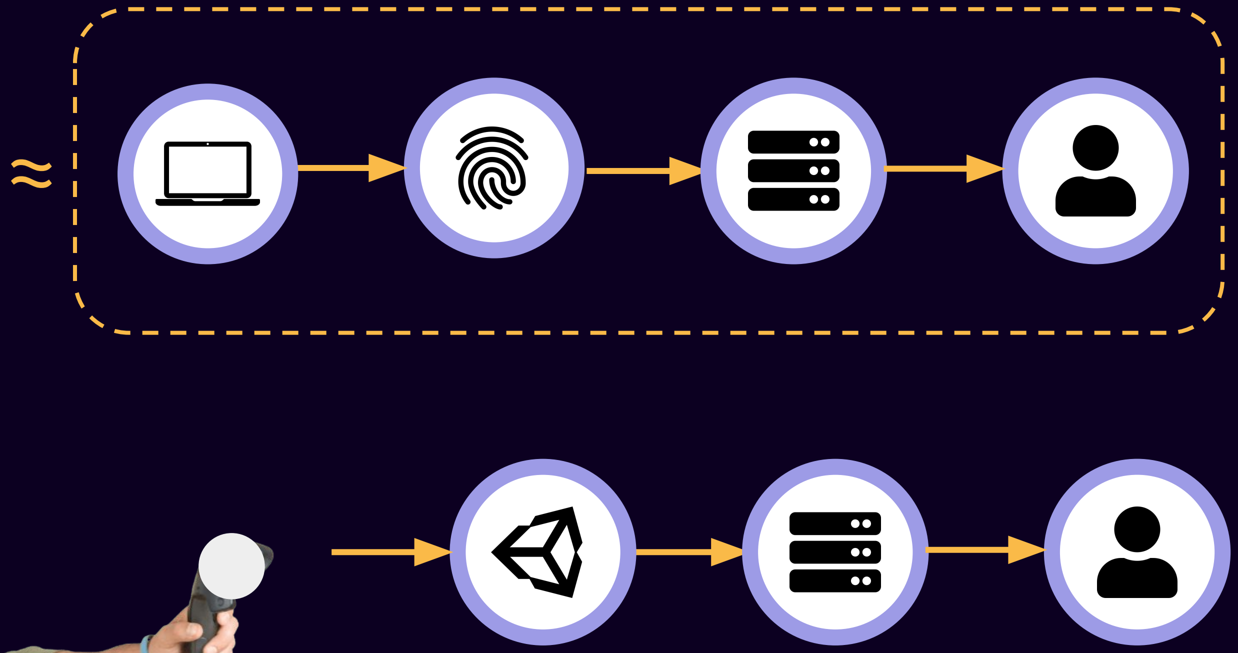
3.2% of entropy gain

Other biometrics



Implications for VR privacy





Thank you!

50,000+ User Identification Study
https://rdi.berkeley.edu/vr-identification/
https://arxiv.org/abs/2302.08927
https://github.com/MetaGuard/Identification

Other RDI Metaverse Research
https://rdi.berkeley.edu/metaverse/