Tracko

Ad-hoc Mobile 3D Tracking Using Bluetooth Low Energy and Inaudible Signals for Cross-Device Interaction

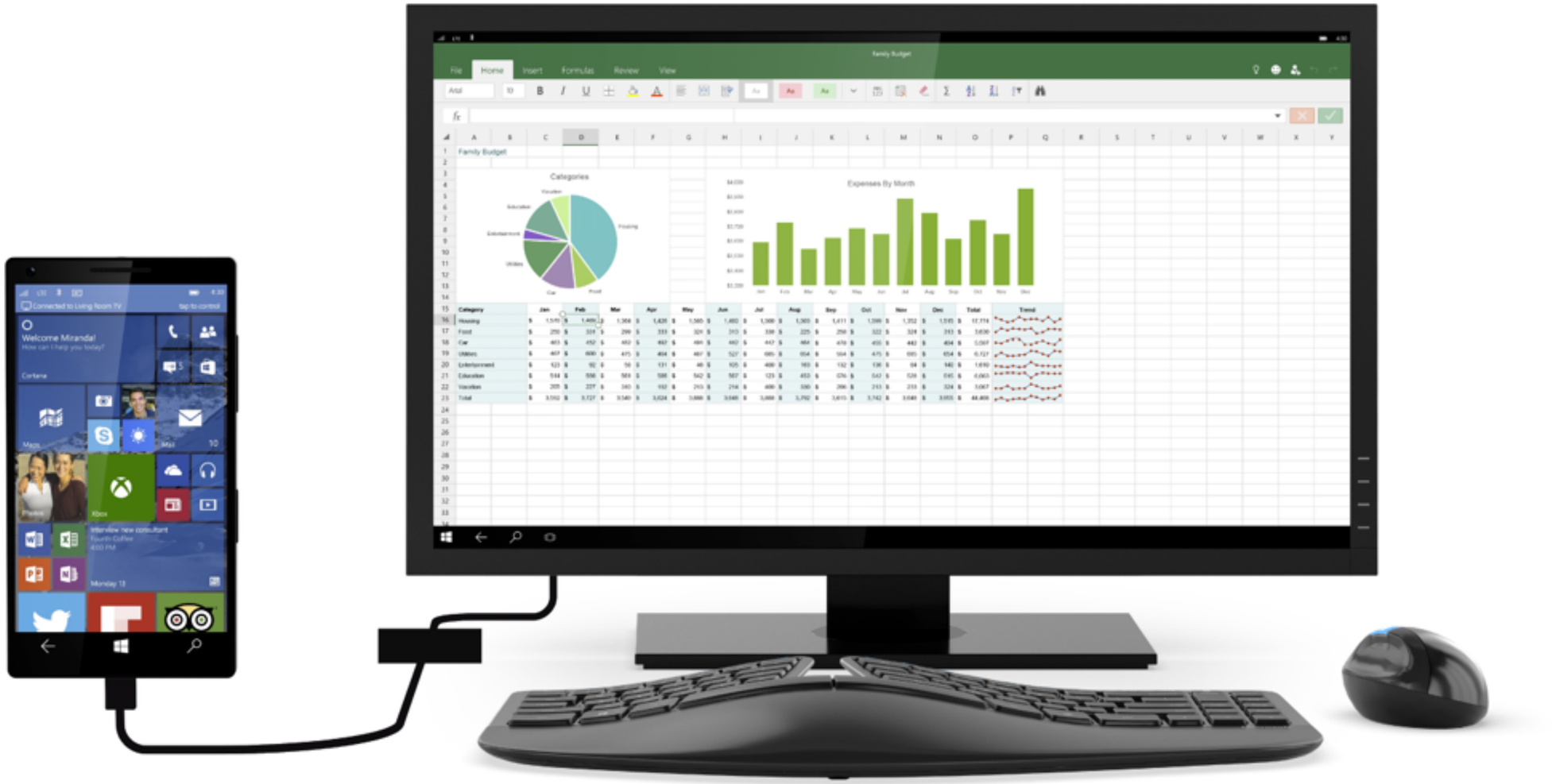
Haojian Jin[1]
Christian Holz[2,3]
Kasper Hornbæk[3]

[1]Yahoo Labs
[2]Microsoft Research
[3]University of Copenhagen

Tracko solves an increasing
need for **mobile infrastructure**

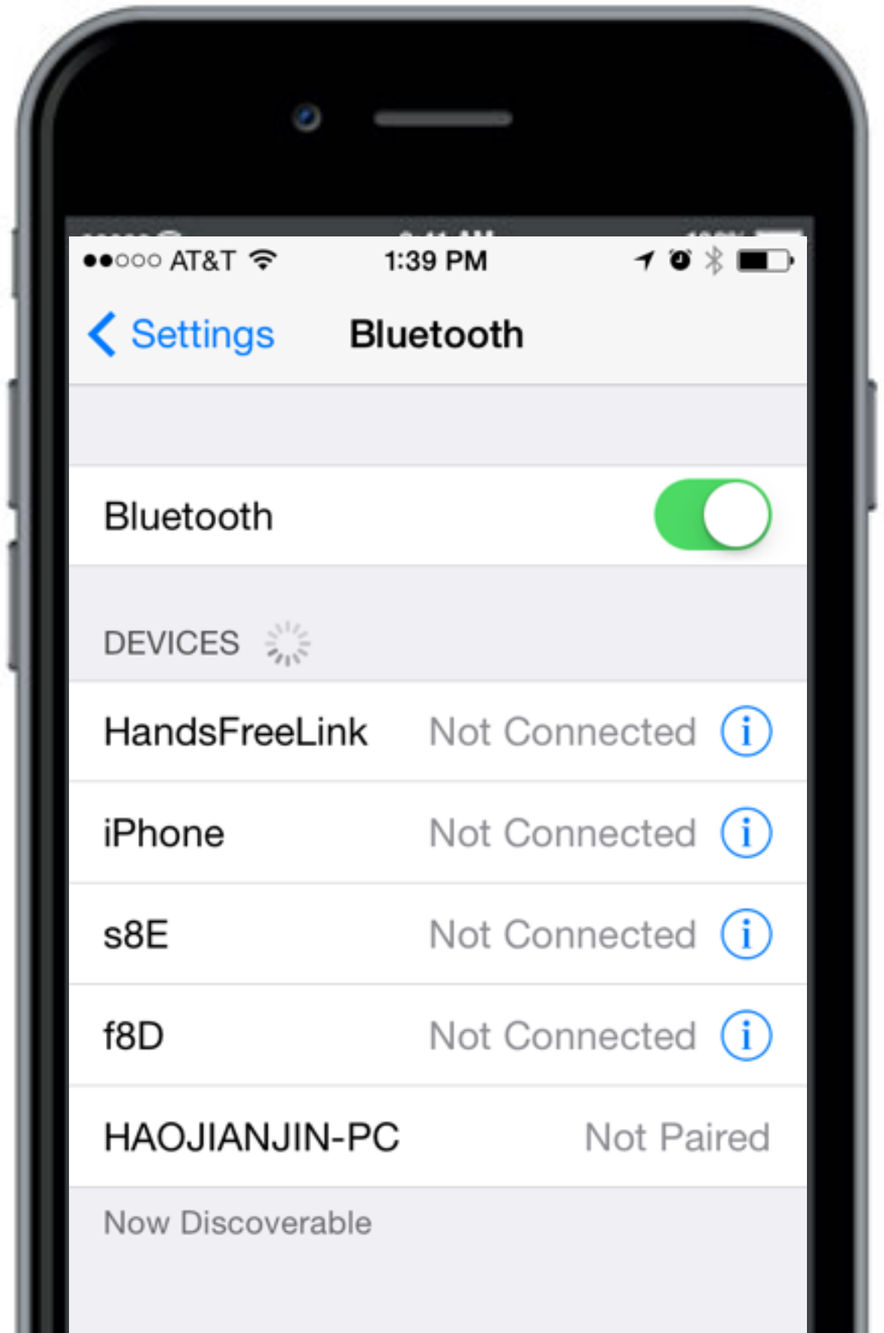
Apple Continuity

Samsung Flow



Microsoft Continuum

how do these devices **see each other?**



detect only the **presence** of surrounding devices

but to track their **actual locations…**


Optitrack

Tracko

ad-hoc **mobile** device to device 3D tracking system

average 3D tracking accuracy of 11.7 cm within 1m

Tracko fuses 3 signal types on commodity devices

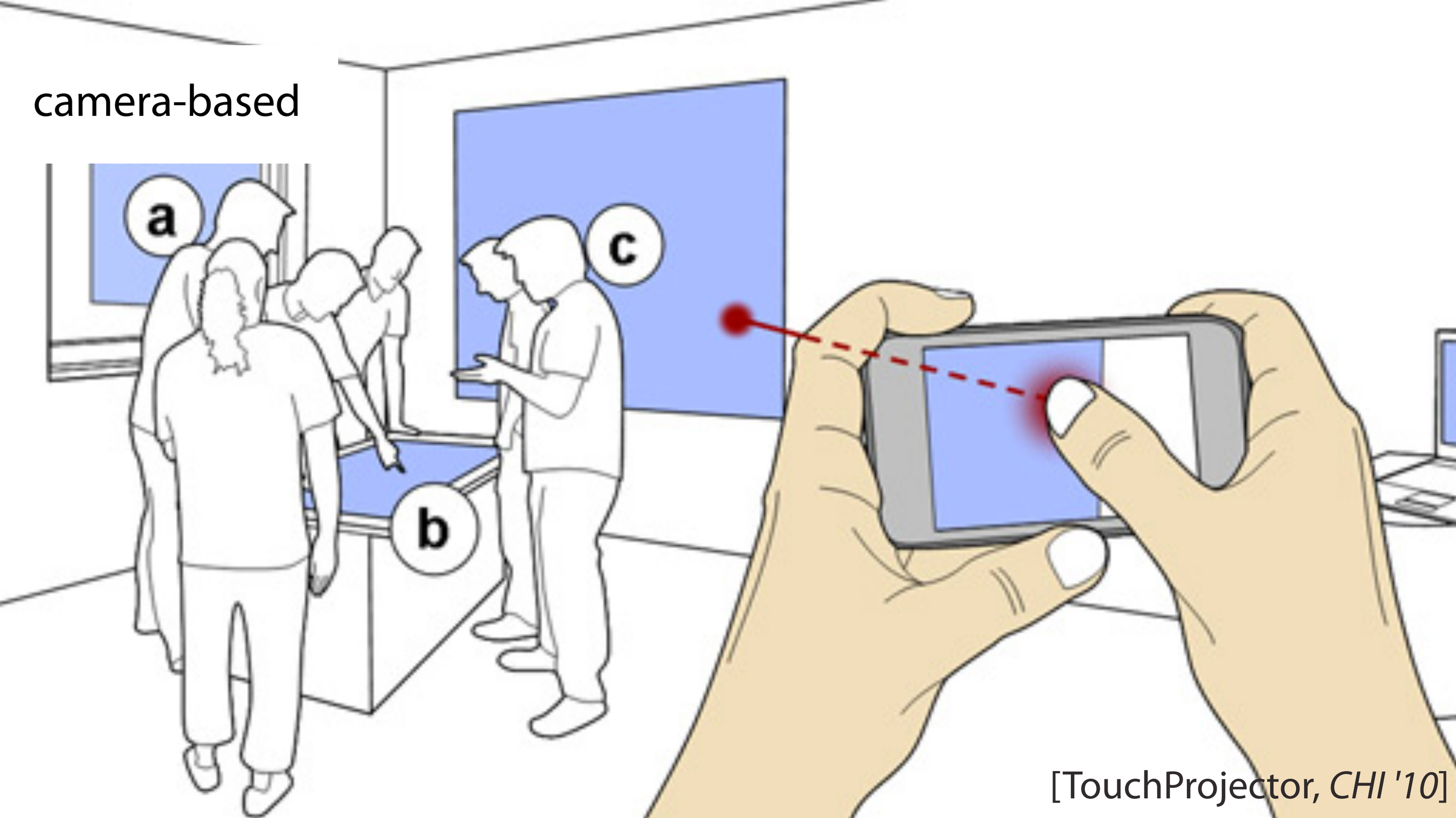
1. bluetooth low energy
2. inaudible stereo signals
3. inertial motion

Tracko: mobile 3D tracking

related work

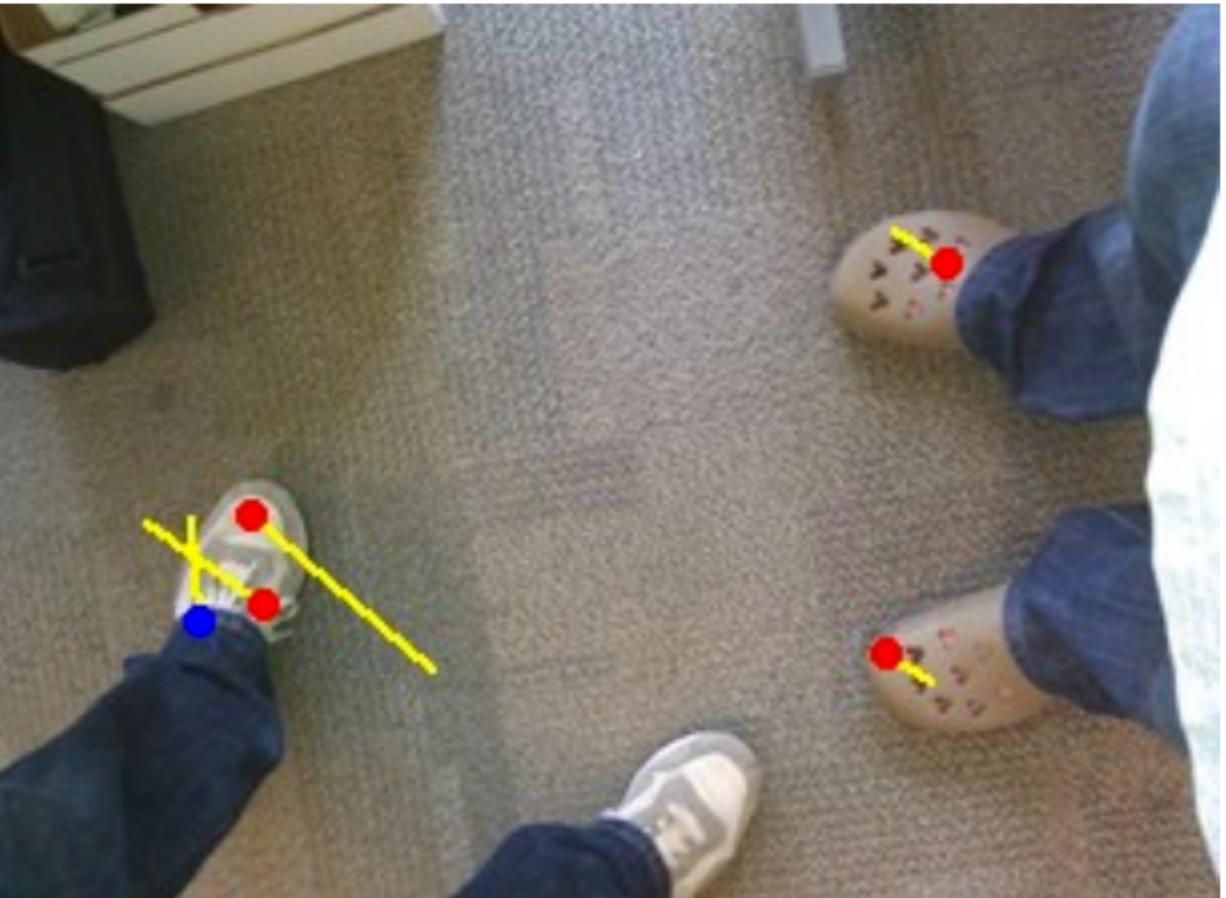
tracking systems using video, radio, audio

camera-based



[TouchProjector, *CHI '10*]

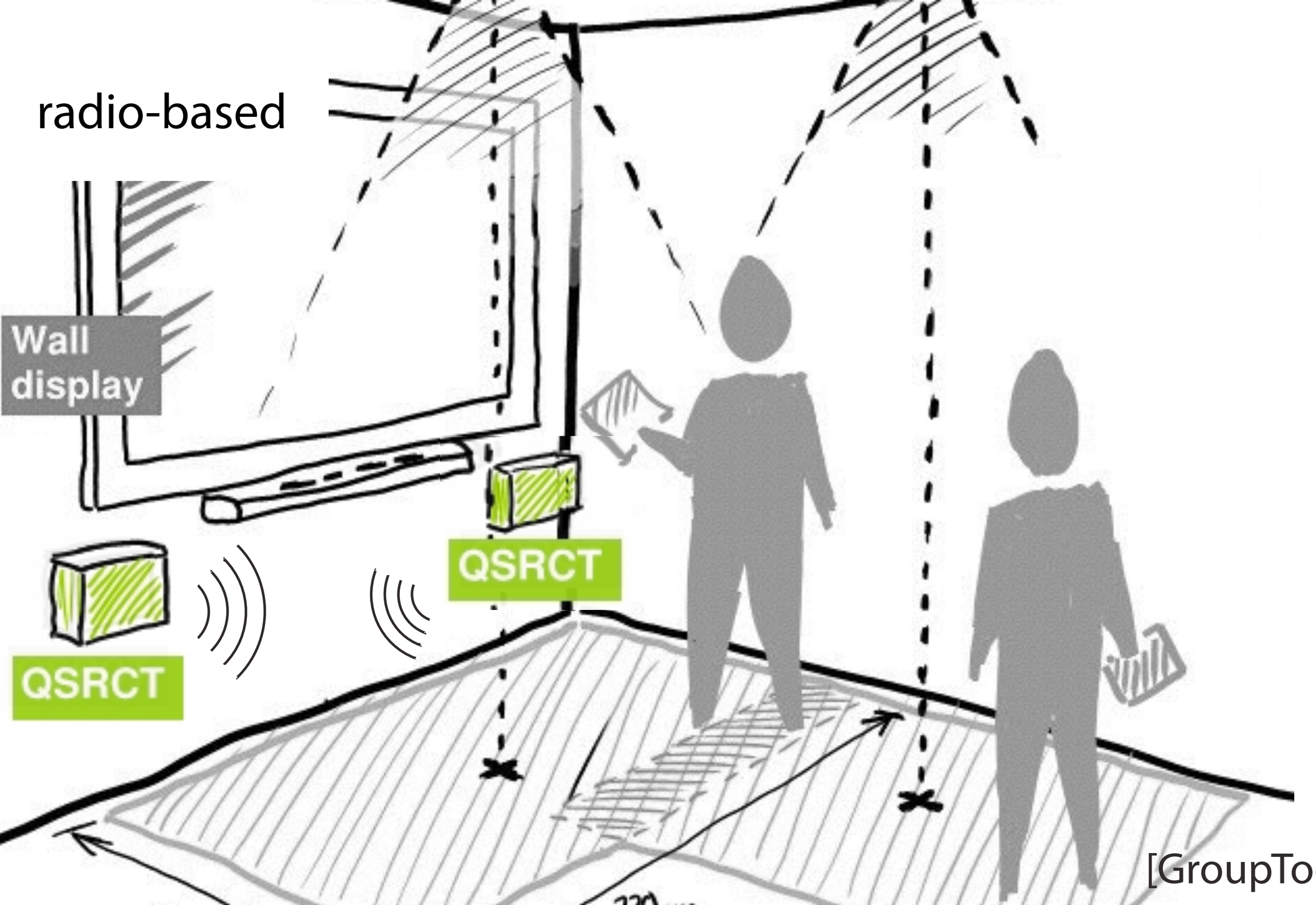
camera-based



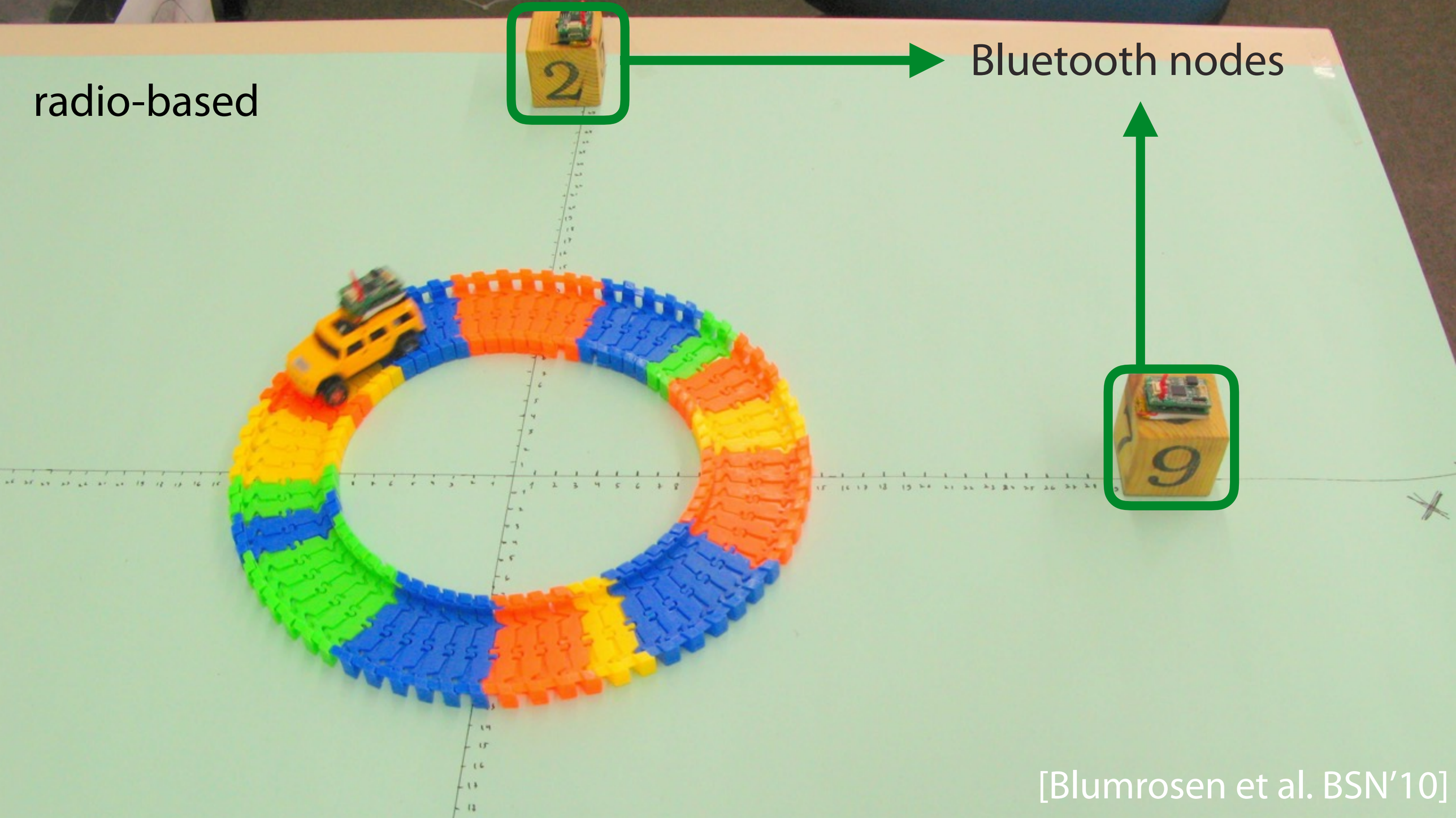
[Orienteer, *CHI '12*]

radio-based

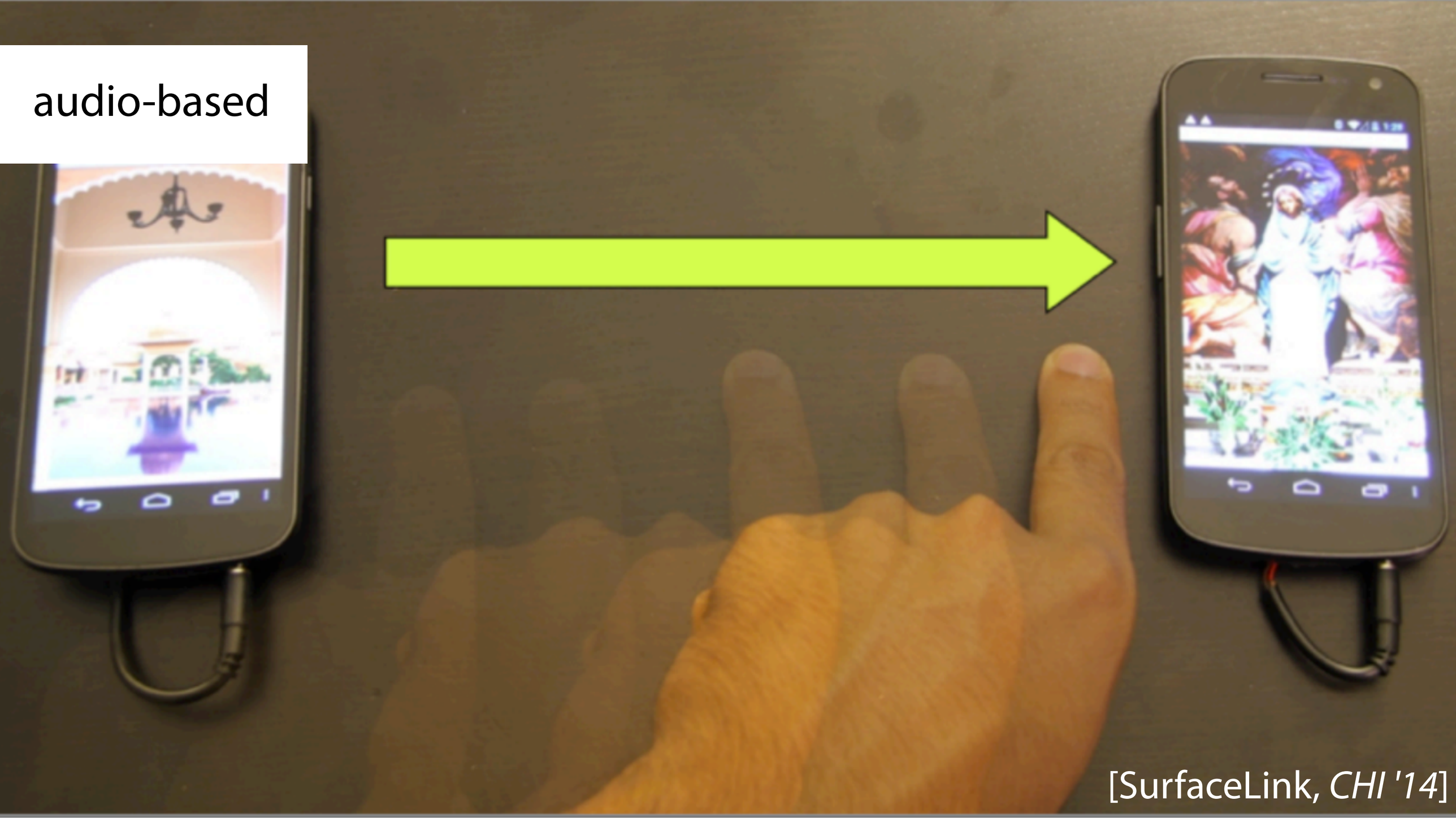
radio-based



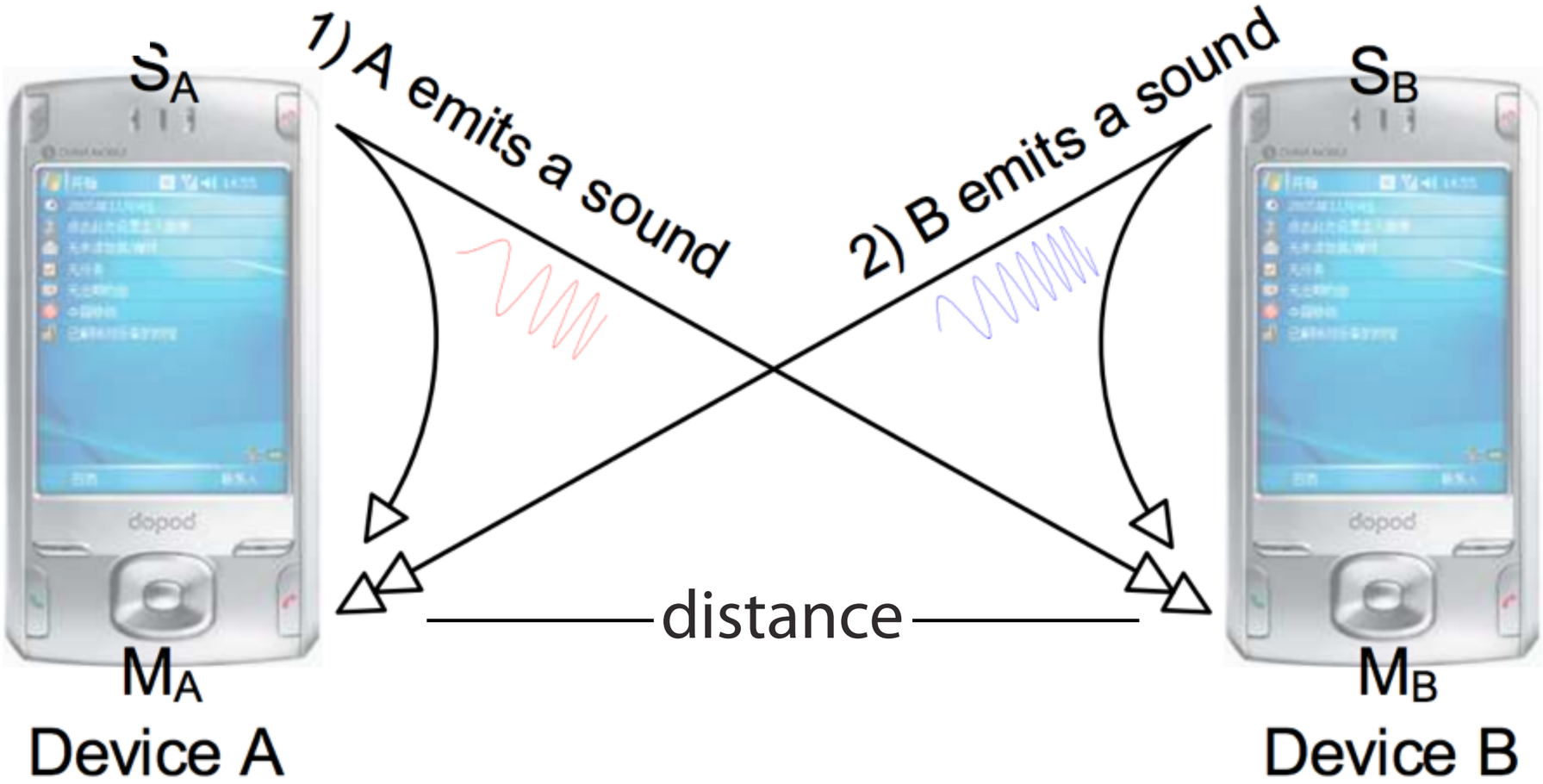
[GroupTogether, *UIST '12*]


radio-based
Bluetooth nodes
2
9
[Blumrosen et al. BSN'10]

audio-based


audio-based
[SurfaceLink, *CHI '14*]

audio-based



[BeepBeep, *Sensys '07*]

audio-based

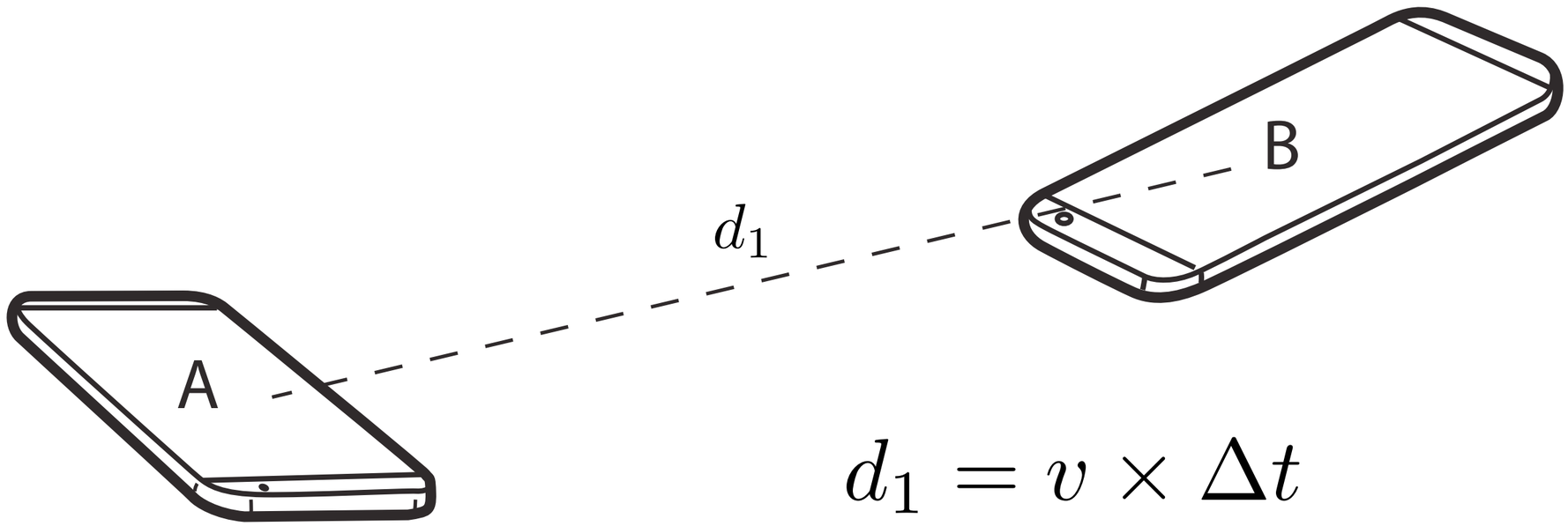
"On the **Feasibility** of Real-Time Phone-to-Phone 3D Localization"

geometric model for **static** devices

[Qiu et al., *Sensys '11*]

background

estimating distances
from exchanged audio signals

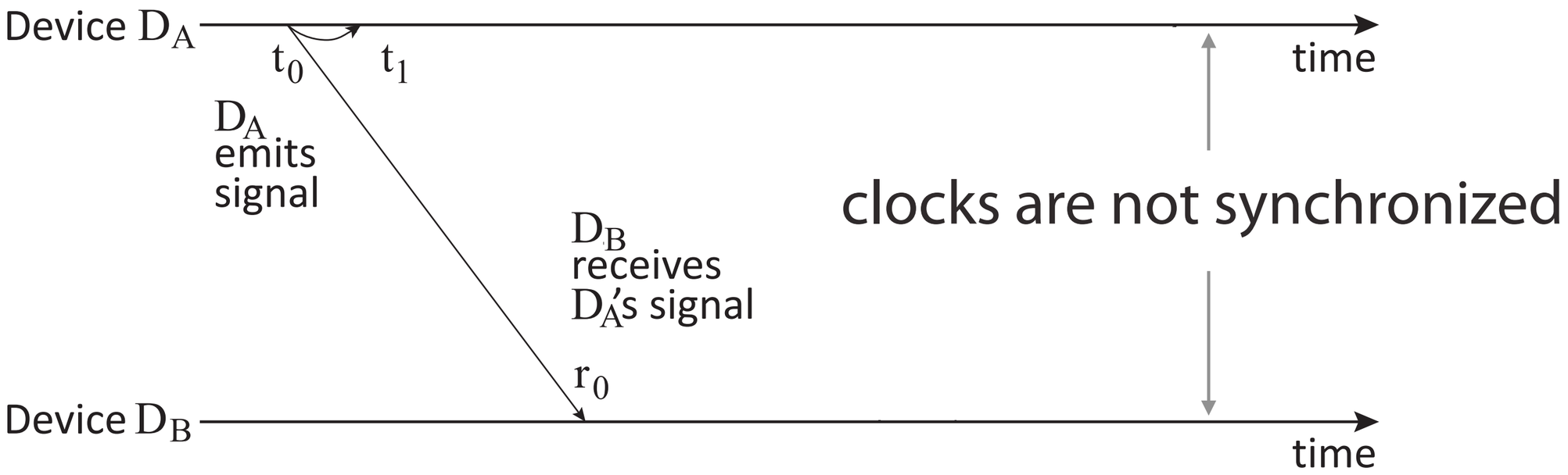


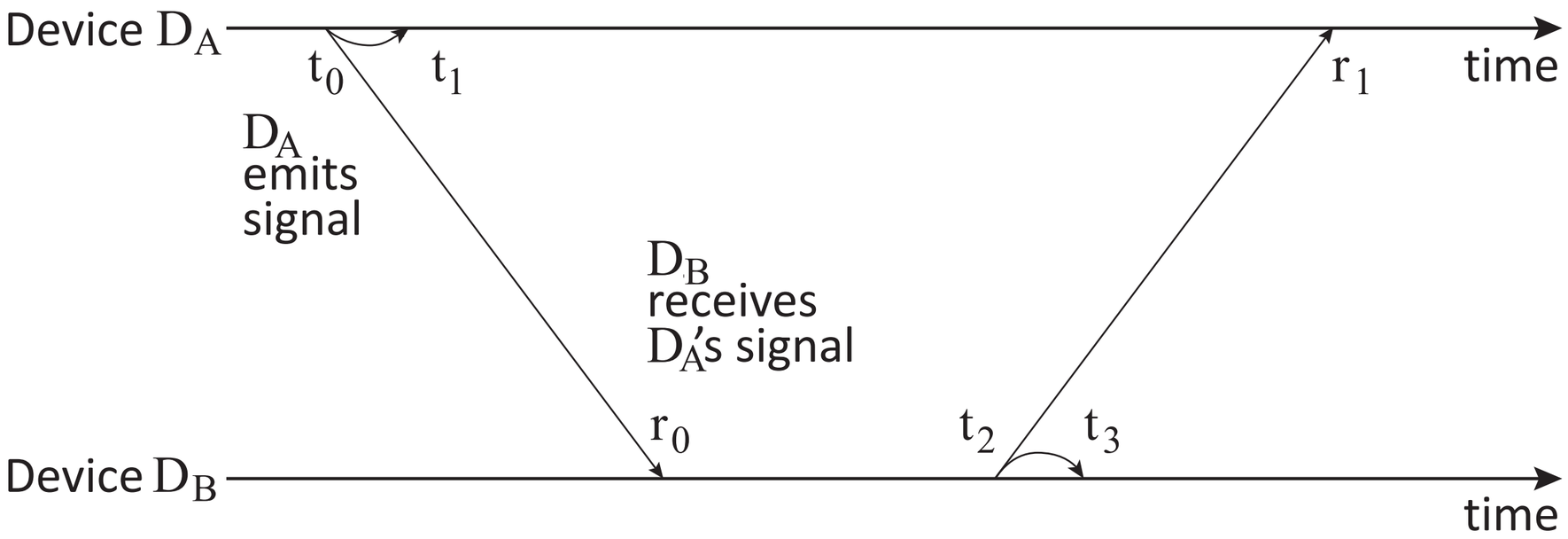
$$d_1 = v \times \Delta t$$

v is the **constant** speed of sound.

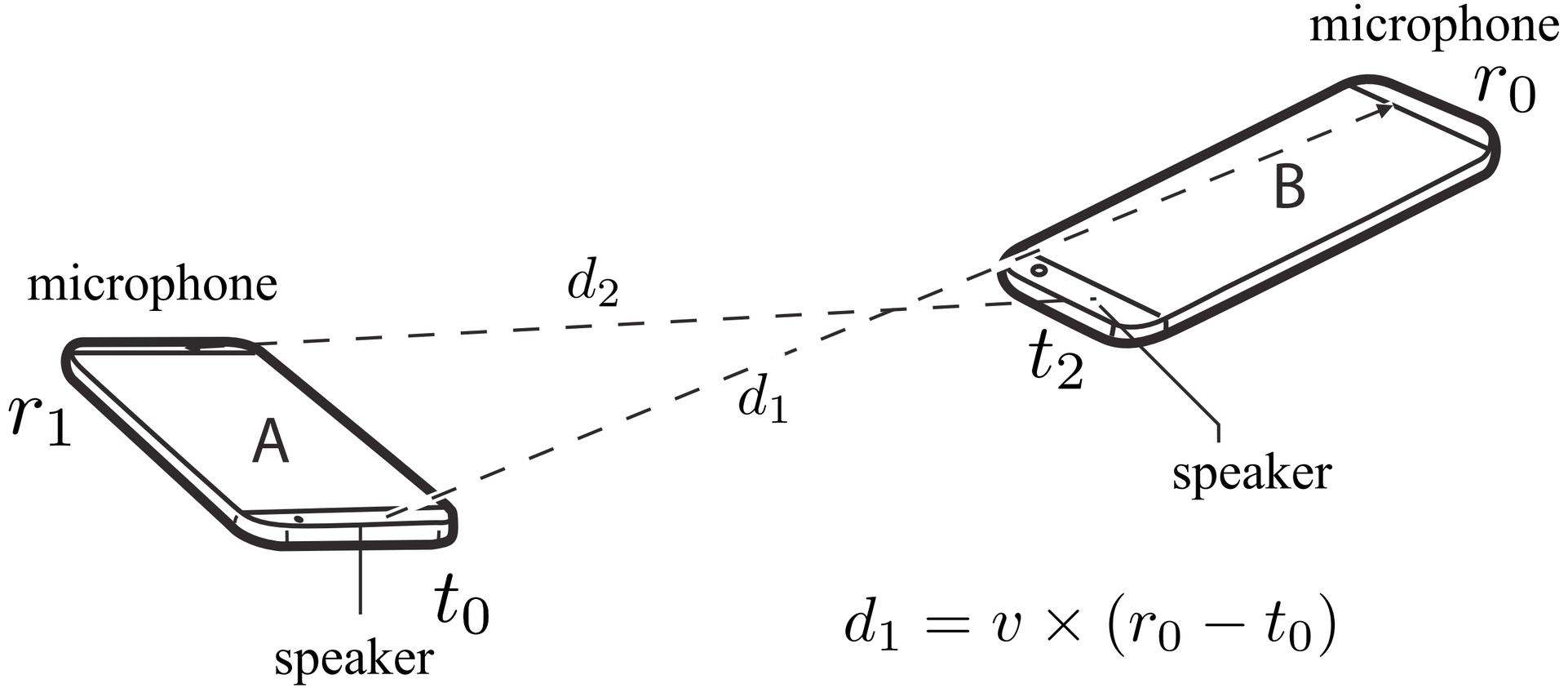
we can calculate the distance d_1

Time of Arrival


Device D_A
t_0
t_1
time
D_A emits signal
clocks are not synchronized
D_B receives D_A's signal
r_0
Device D_B
time



the trick: capture round-trip



$$d_1 = v \times (r_0 - t_0)$$

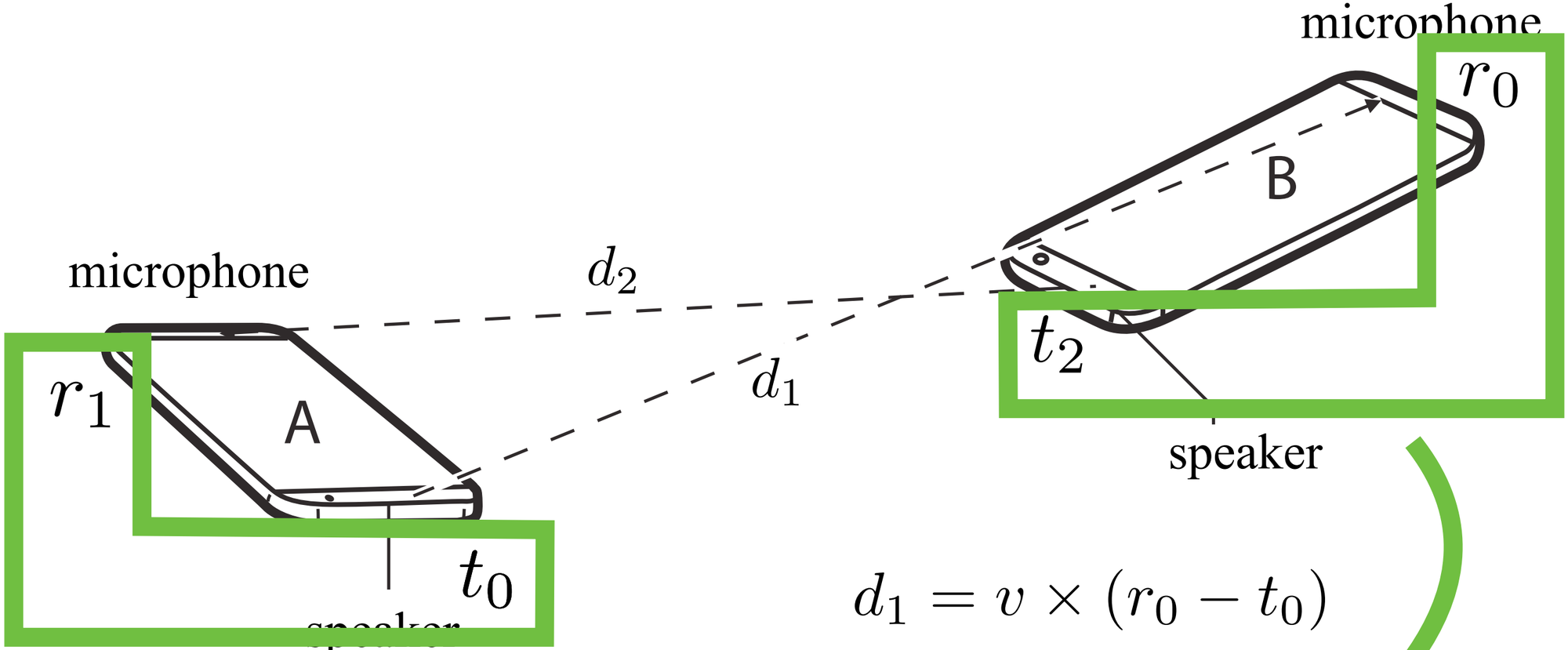
$$d_2 = v \times (r1 - t2)$$

$$d_1 + d_2 = v \times (r_0 - t_0) + v \times (r_1 - t_2)$$

$$= v \times (r_0 - t_2) + v \times (r_1 - t_0)$$

round trip time

[BeepBeep Sensys '07]
[Ganeriwal et al. Sensys'03]



$$d_1 = v \times (r_0 - t_0)$$

$$d_2 = v \times (r1 - t2)$$

$$d_1 + d_2 = v \times (r_0 - t_0) + v \times (r_1 - t_2)$$

$$= v \times (r_0 - t_2) + v \times (r_1 - t_0)$$

round trip time

[BeepBeep Sensys '07]

[Ganeriwal et al. Sensys'03]

accurate round-trip distances

requires no cross-device synchronization

result

Tracko

mobile device to mobile device 3D tracking system

3 signal types

1 Bluetooth low energy (BLE)

2 inaudible acoustic signals

3 inertial sensors (IMU)

1 Bluetooth low energy

the **good:**

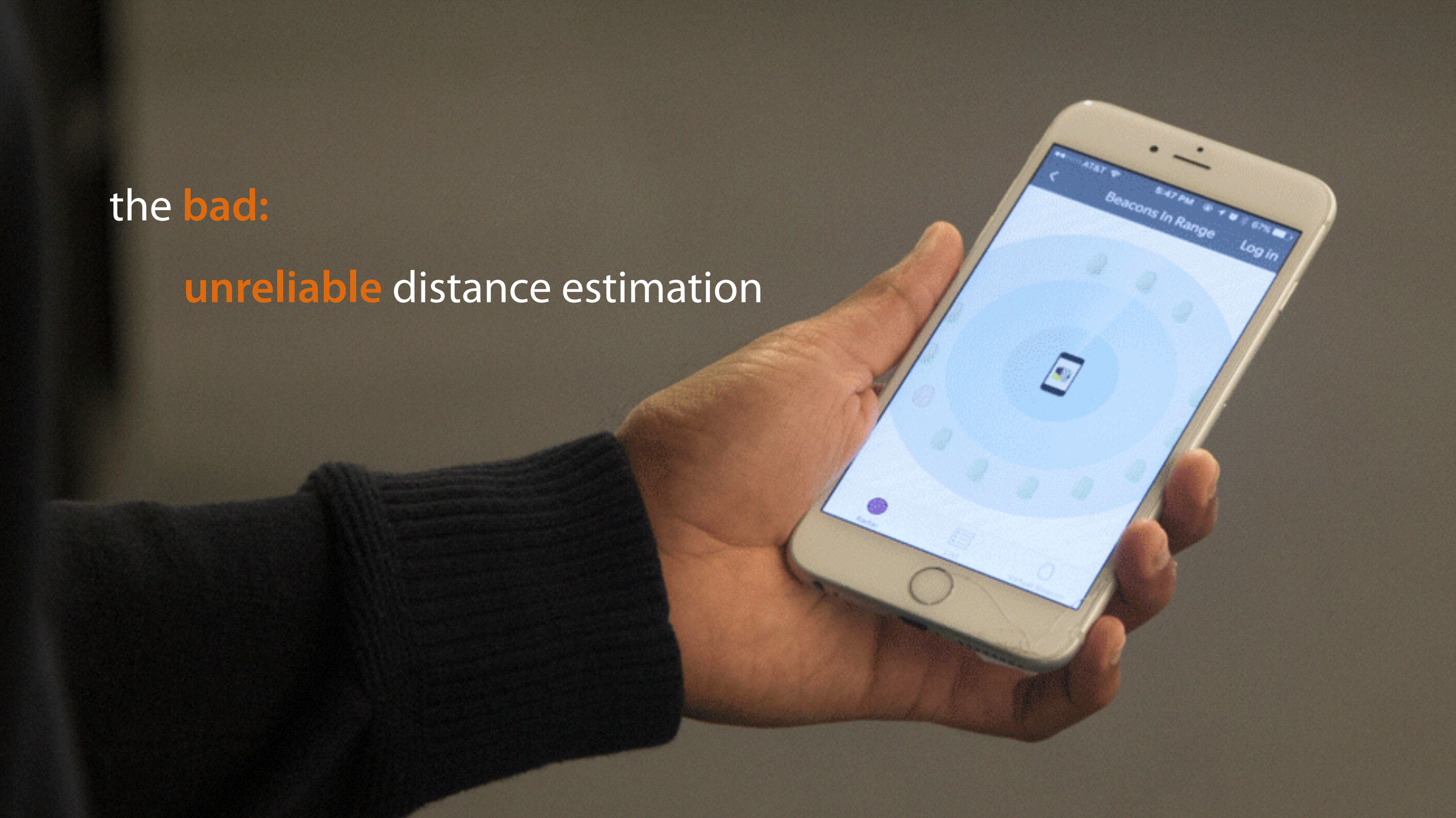
robust signal, good as **fallback**

Tracko detects the **presence** of surrounding devices

Bluetooth low energy

the **bad:**

unreliable distance estimation



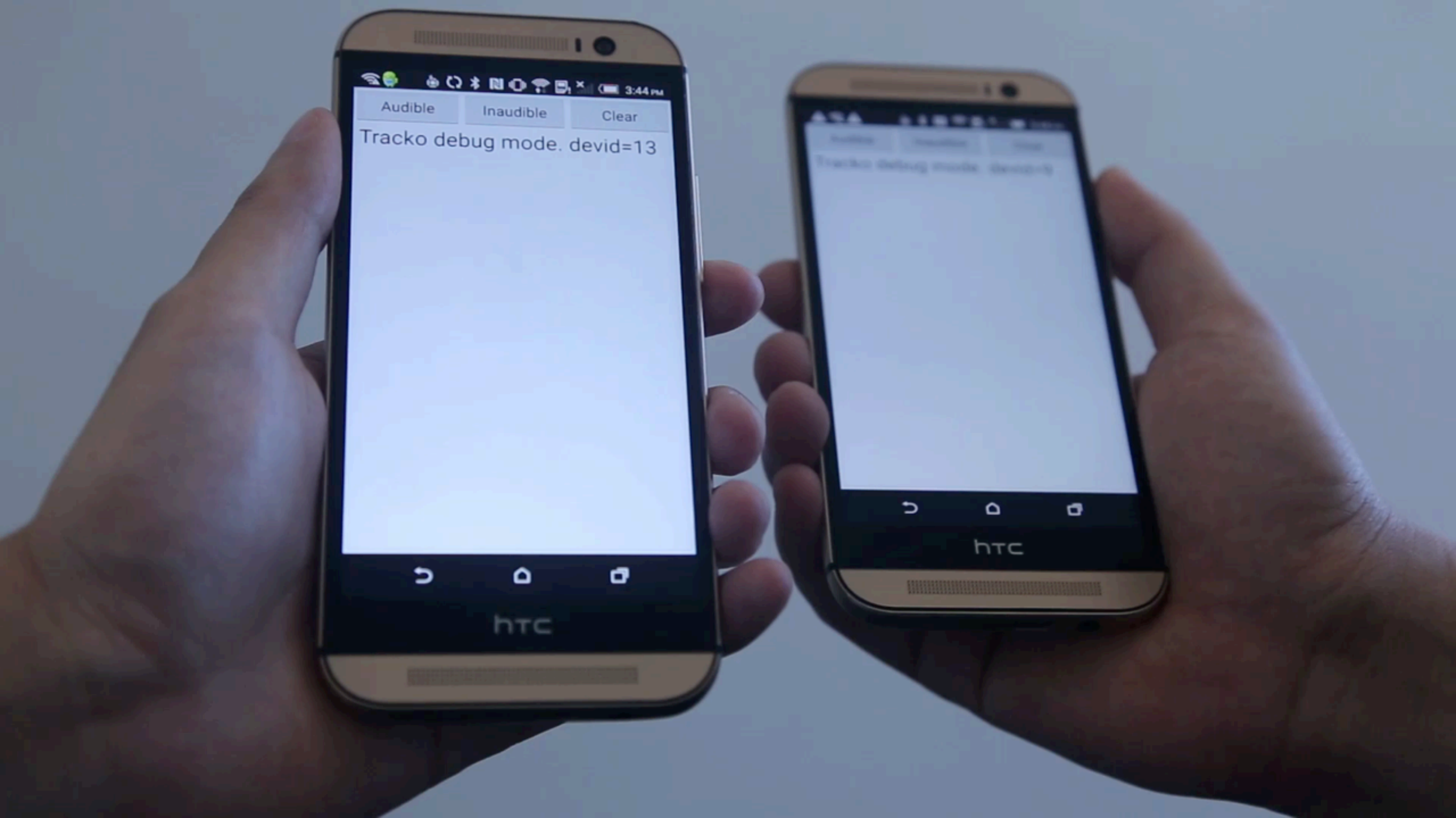
but Tracko already calculates
centimeter-level distances between devices

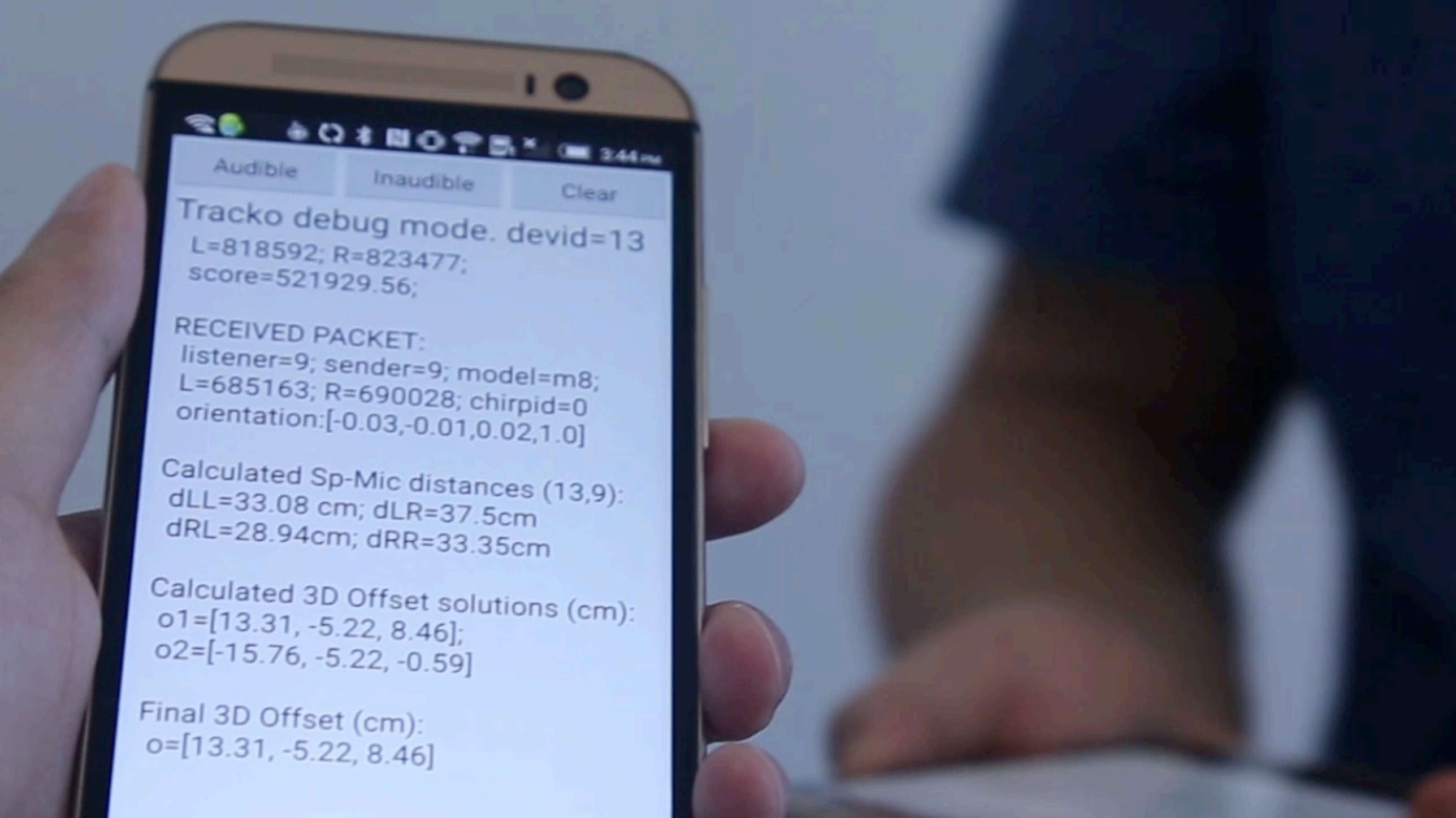
on-the-fly calibration

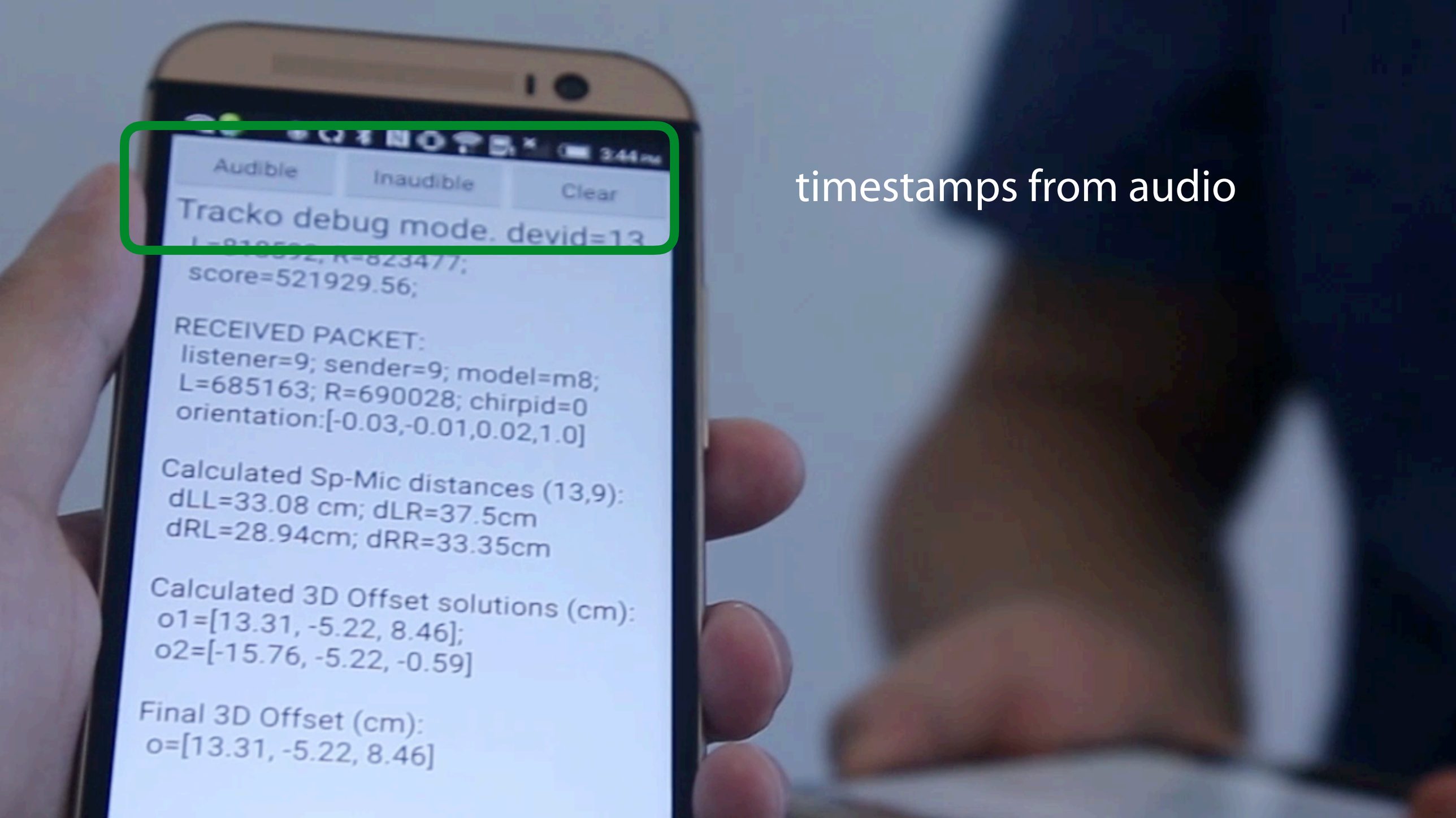
using the distances Tracko obtains through audio

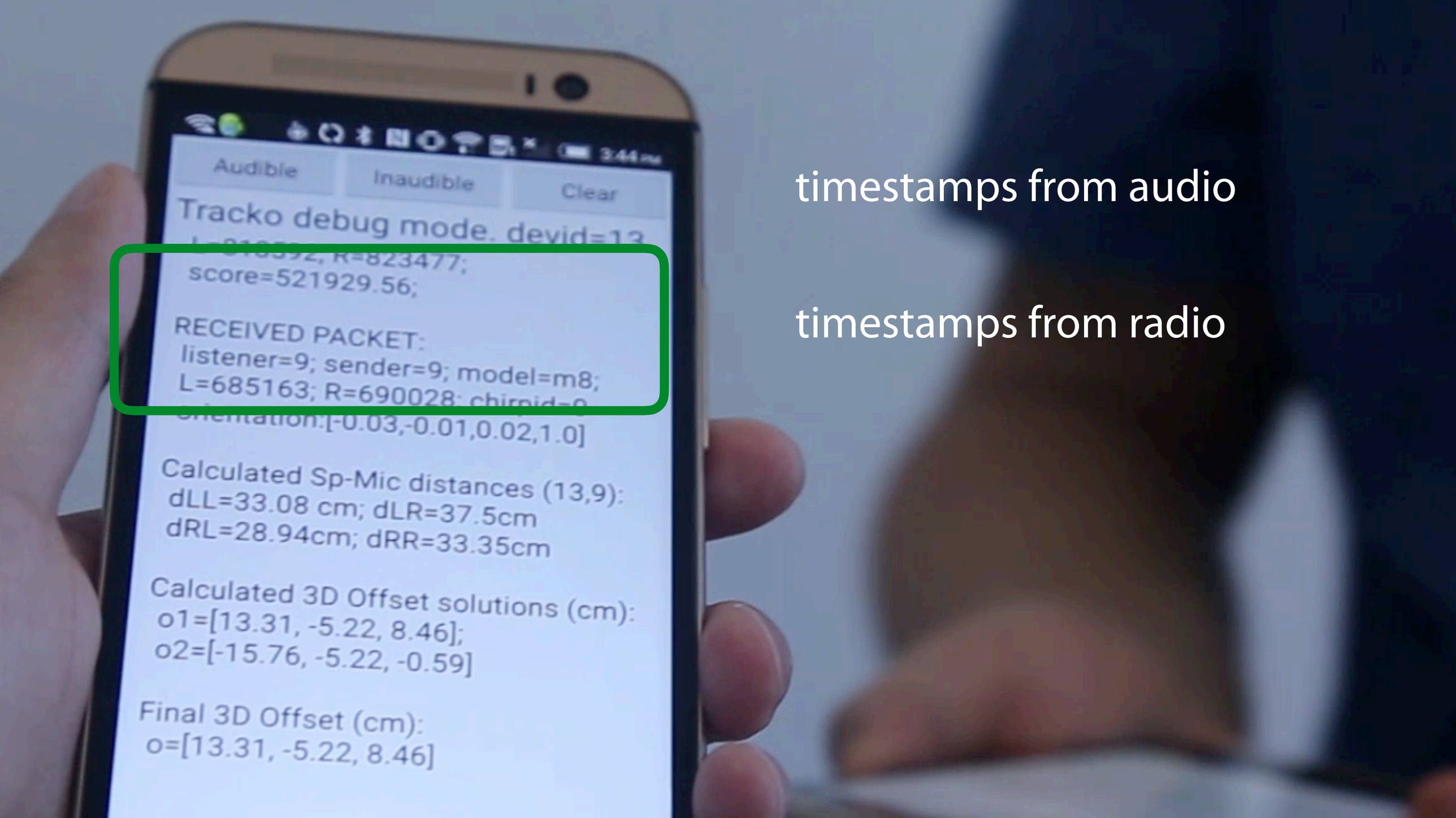
Tracko substantially increases the
accuracy of BLE-based distance estimation

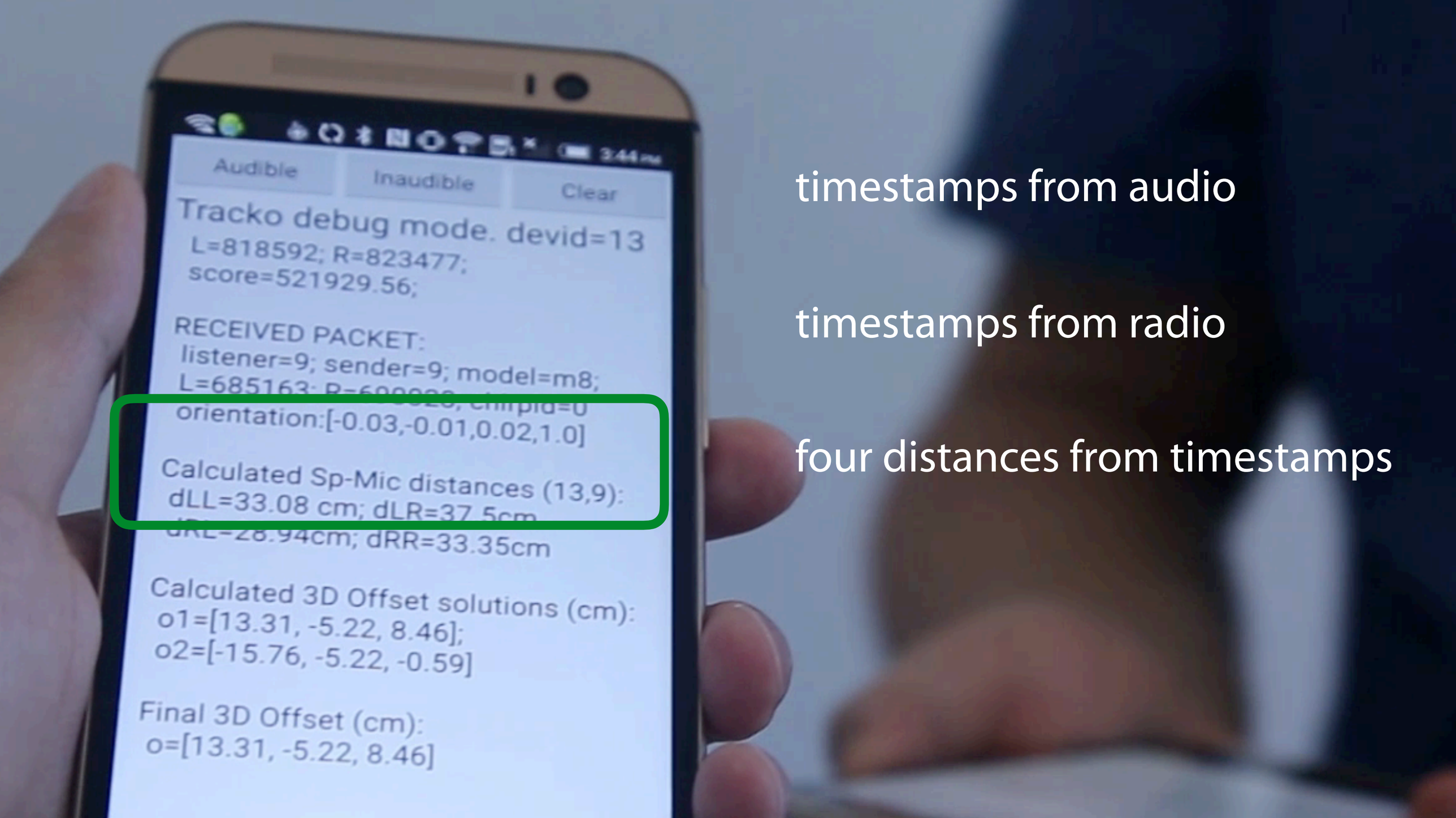
2 Stereo Sounds

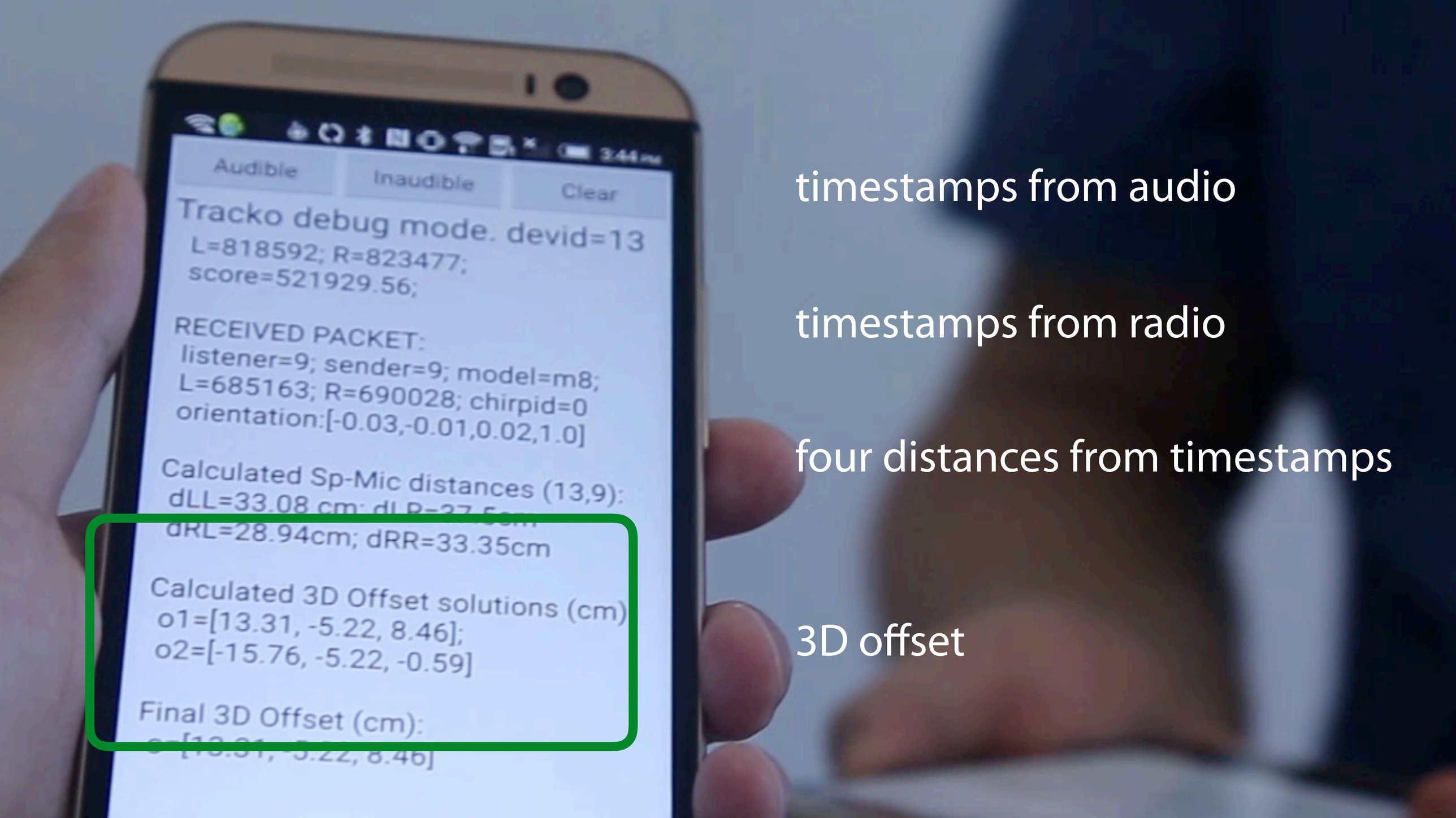

3:44 PM
Audible
Inaudible
Clear
Tracko debug mode. devid=13
htc
htc


Audible
Inaudible
Clear
Tracko debug mode. devid=13
L=818592; R=823477;
score=521929.56;
RECEIVED PACKET:
listener=9; sender=9; model=m8;
L=685163; R=690028; chirpid=0
orientation:[-0.03,-0.01,0.02,1.0]
Calculated Sp-Mic distances (13,9):
dLL=33.08 cm; dLR=37.5cm
dRL=28.94cm; dRR=33.35cm
Calculated 3D Offset solutions (cm):
o1=[13.31, -5.22, 8.46];
o2=[-15.76, -5.22, -0.59]
Final 3D Offset (cm):
o=[13.31, -5.22, 8.46]

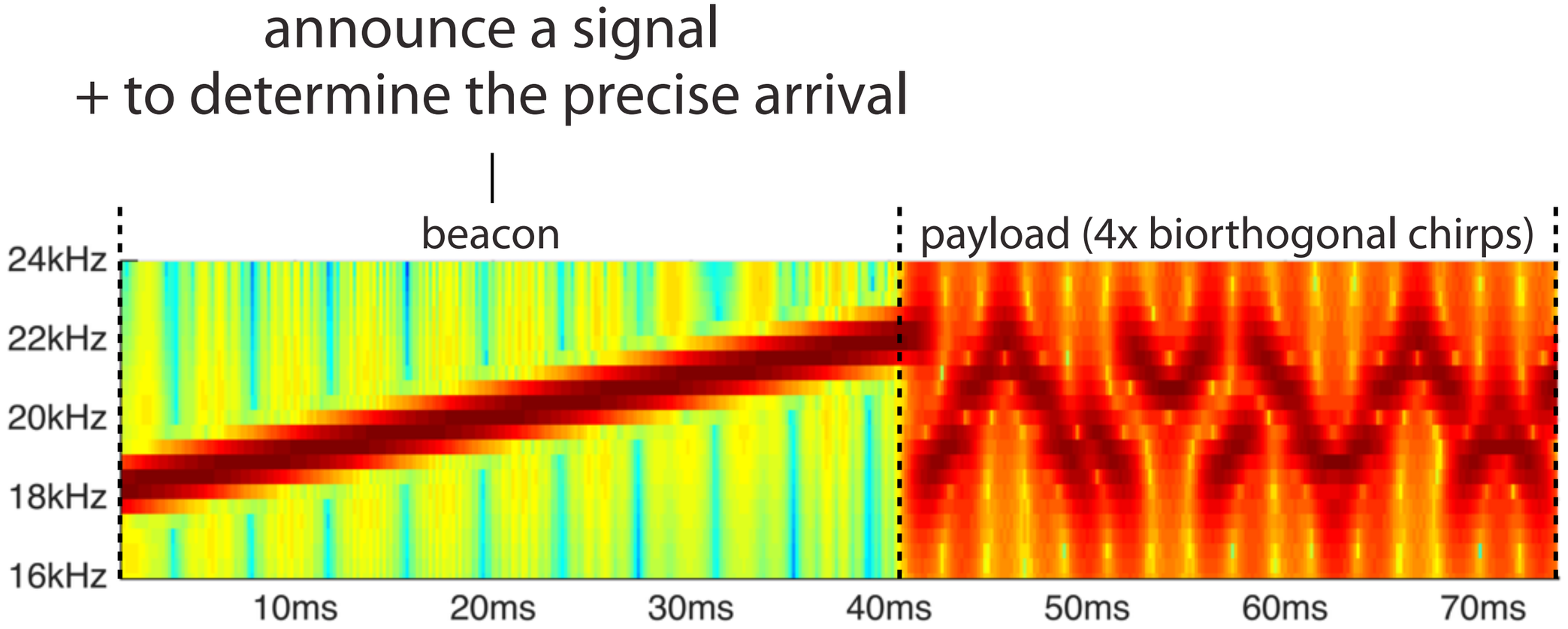

Audible
Inaudible
Clear
Tracko debug mode. devid=13
score=521929.56;
RECEIVED PACKET:
listener=9; sender=9; model=m8;
L=685163; R=690028; chirpid=0
orientation:[-0.03,-0.01,0.02,1.0]
Calculated Sp-Mic distances (13,9):
dLL=33.08 cm; dLR=37.5cm
dRL=28.94cm; dRR=33.35cm
Calculated 3D Offset solutions (cm):
o1=[13.31, -5.22, 8.46];
o2=[-15.76, -5.22, -0.59]
Final 3D Offset (cm):
o=[13.31, -5.22, 8.46]
timestamps from audio


Audible
Inaudible
Clear
Tracko debug mode. devid=13
score=521929.56;
RECEIVED PACKET:
listener=9; sender=9; model=m8;
L=685163; R=690028;
Calculated Sp-Mic distances (13,9):
dLL=33.08 cm; dLR=37.5cm
dRL=28.94cm; dRR=33.35cm
Calculated 3D Offset solutions (cm):
o1=[13.31, -5.22, 8.46];
o2=[-15.76, -5.22, -0.59]
Final 3D Offset (cm):
o=[13.31, -5.22, 8.46]
timestamps from audio
timestamps from radio


Audible
Inaudible
Clear
Tracko debug mode. devid=13
L=818592; R=823477;
score=521929.56;
RECEIVED PACKET:
listener=9; sender=9; model=m8;
orientation:[-0.03,-0.01,0.02,1.0]
Calculated Sp-Mic distances (13,9):
dLL=33.08 cm; dLR=37.5cm
dRR=33.35cm
Calculated 3D Offset solutions (cm):
o1=[13.31, -5.22, 8.46];
o2=[-15.76, -5.22, -0.59]
Final 3D Offset (cm):
o=[13.31, -5.22, 8.46]
timestamps from audio
timestamps from radio
four distances from timestamps

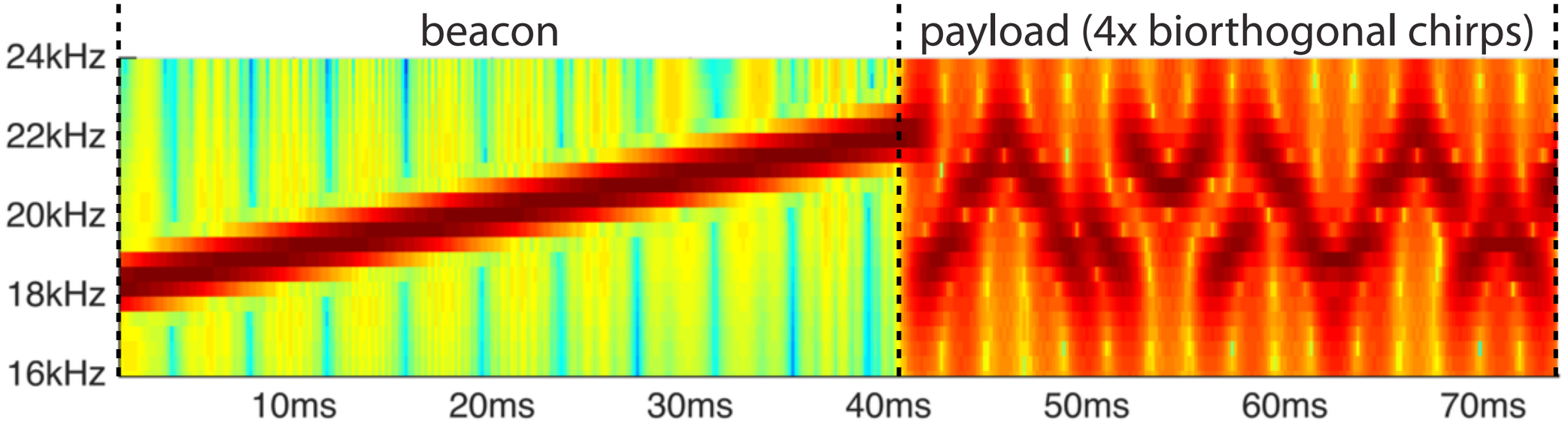

Audible
Inaudible
Clear
Tracko debug mode. devid=13
L=818592; R=823477;
score=521929.56;
RECEIVED PACKET:
listener=9; sender=9; model=m8;
L=685163; R=690028; chirpid=0
orientation:[-0.03,-0.01,0.02,1.0]
Calculated Sp-Mic distances (13,9):
dRL=28.94cm; dRR=33.35cm
Calculated 3D Offset solutions (cm)
o1=[13.31, -5.22, 8.46];
o2=[-15.76, -5.22, -0.59]
Final 3D Offset (cm):
timestamps from audio
timestamps from radio
four distances from timestamps
3D offset

Acoustic signal design



Tracko’s inaudible signals

bi-orthogonal chirps encode the sender of the signal **+ checksum**



Tracko’s inaudible signals

Detecting signals accurately

microphone buffer (48kHz)

n-2	n-1	n	n+1	n+2	.....	m-2	m-1	m	m+1	m+2

$$\text{one frame offset} = \frac{1}{48kHz} \times 340m/s = 0.5cm$$

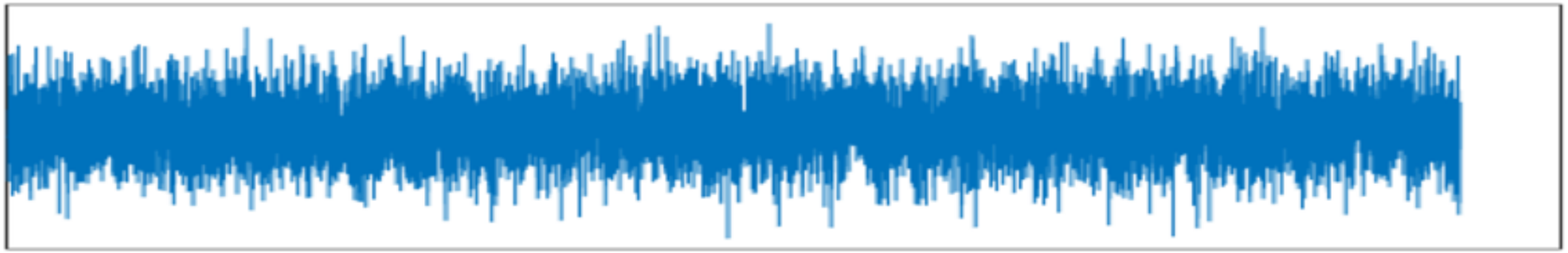
timing granularity

this means that a detection error of
1 frame = 0.5 cm **distance error**

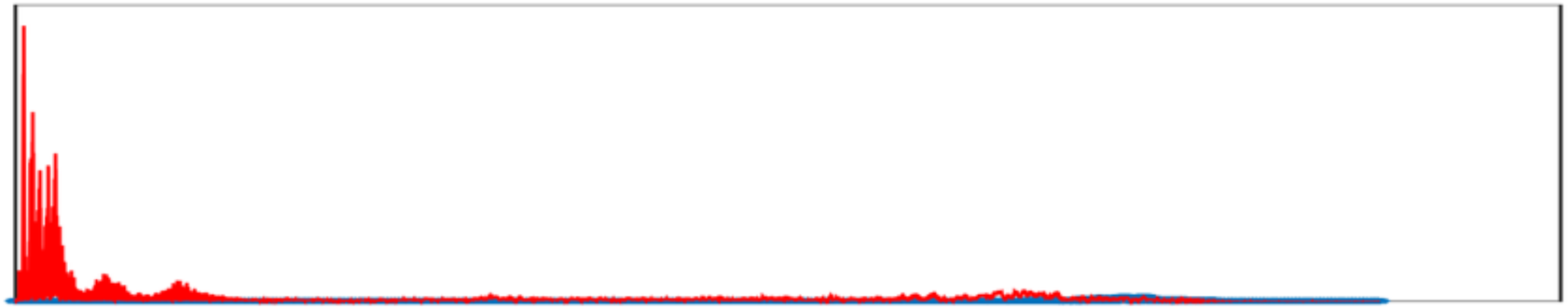
this error **propagates** to Tracko's 3D estimation

accurate detection

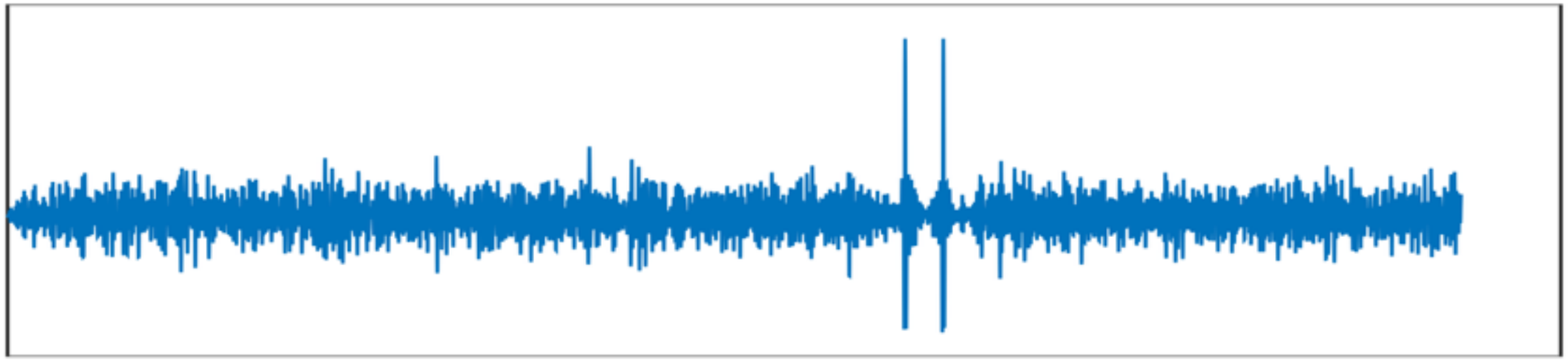
raw signal



high pass

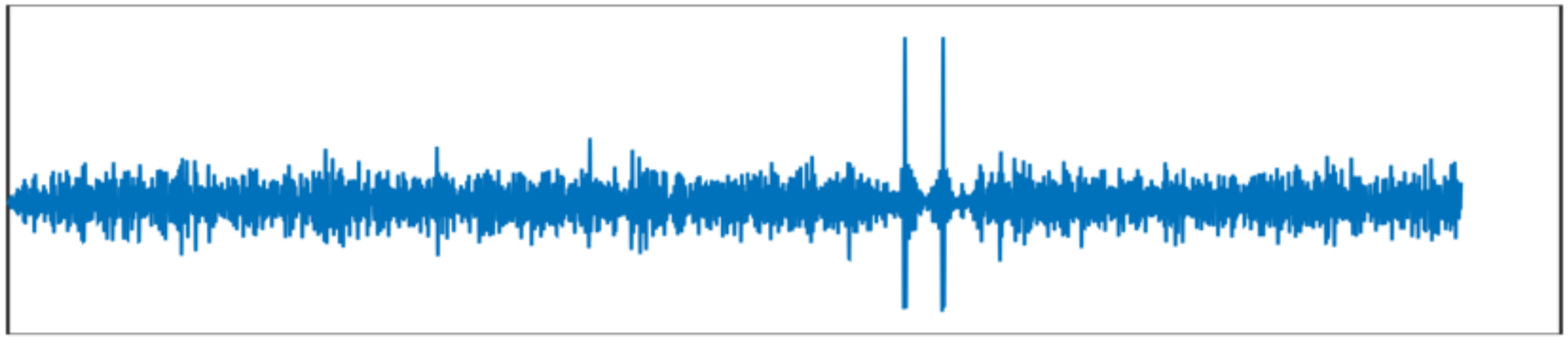


matched filter

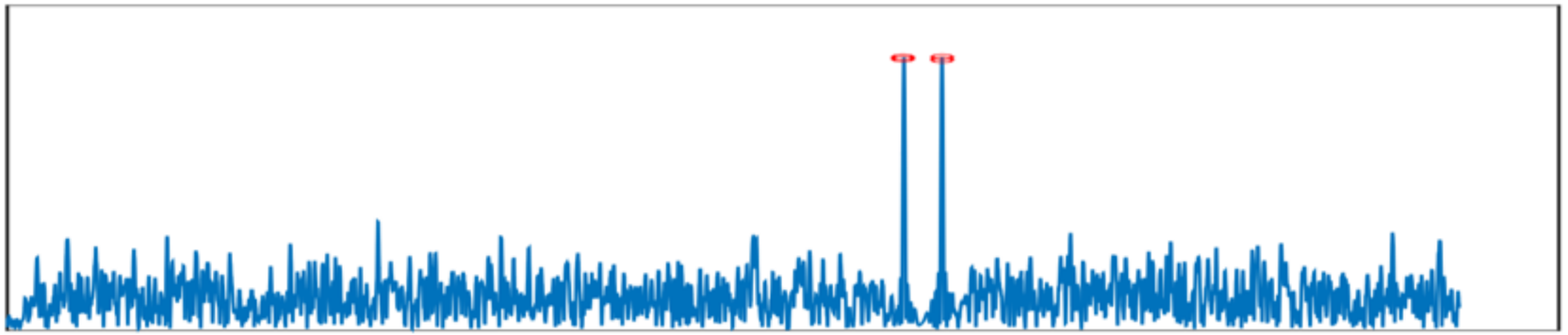


accurate detection

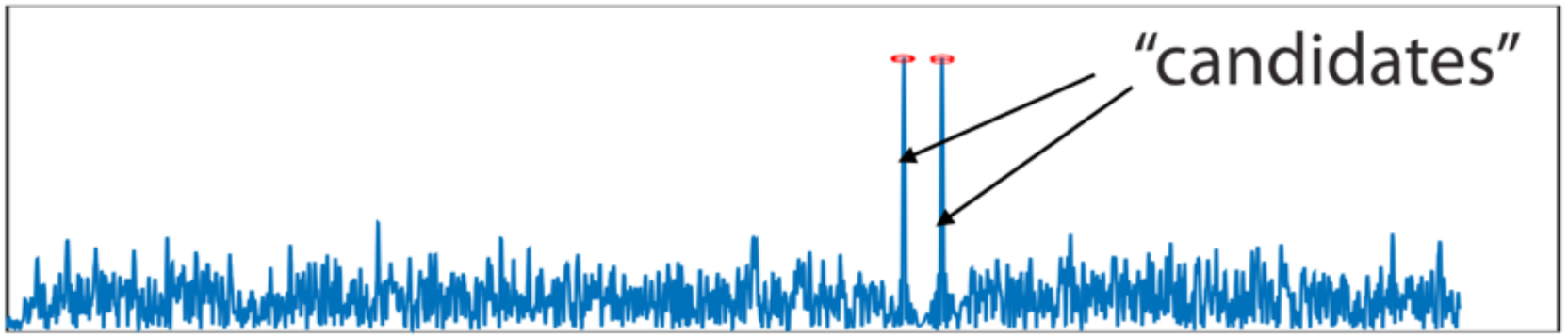
matched filter



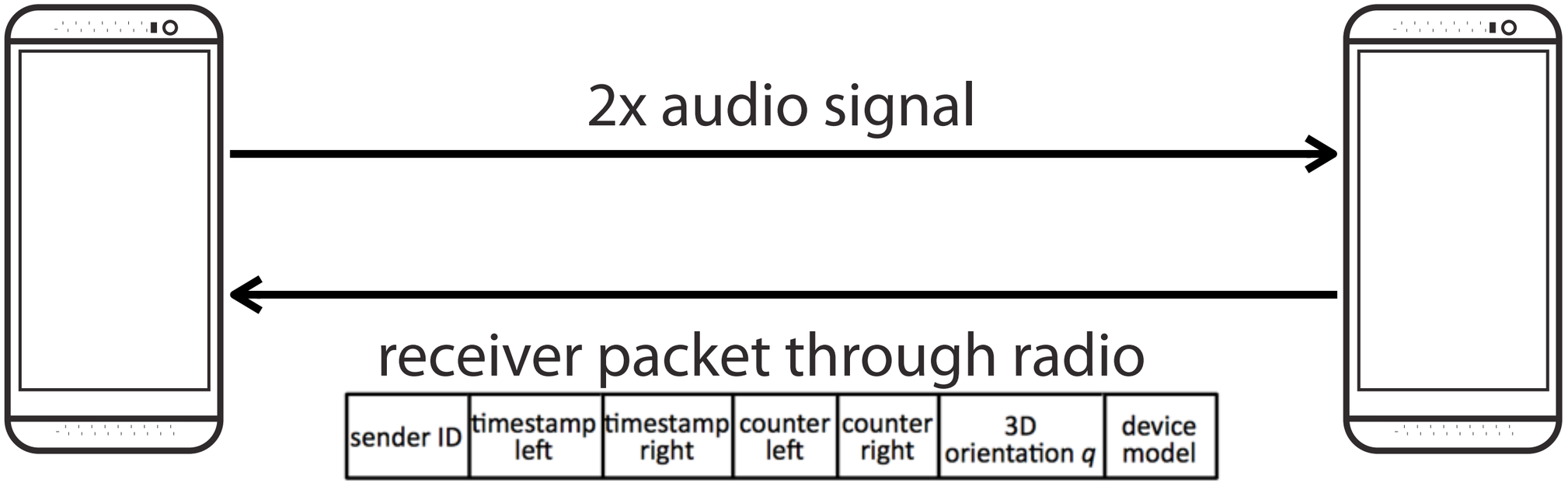
hilbert transform



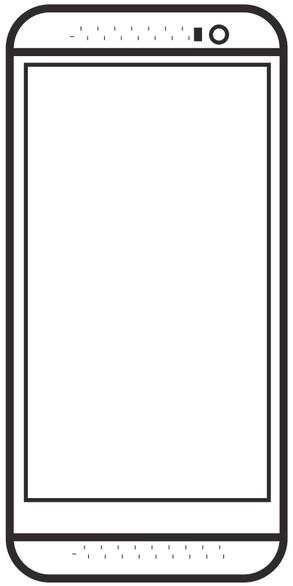
peak selection



calculating distances



receiver packets



each device now has four timestamps:

two local timestamps

two remote timestamps

receiver packets

four timestamps => compute four round-trip distances

left speaker on A => mic on B + **left** speaker on B => mic on A

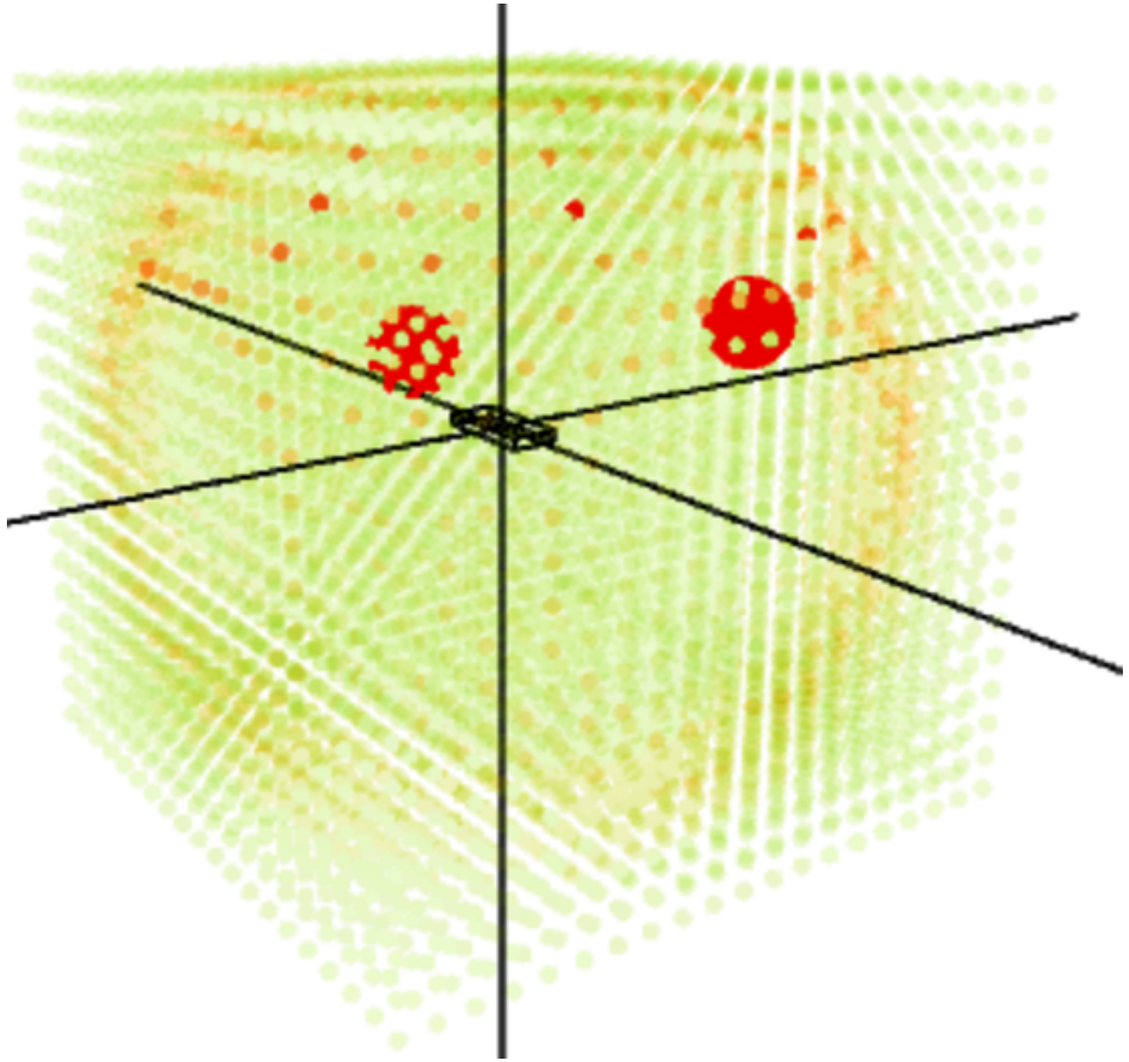
right speaker on A => mic on B + **left** speaker on B => mic on A

left speaker on A => mic on B + **right** speaker on B => mic on A

right speaker on A => mic on B + **right** speaker on B => mic on A

four distances

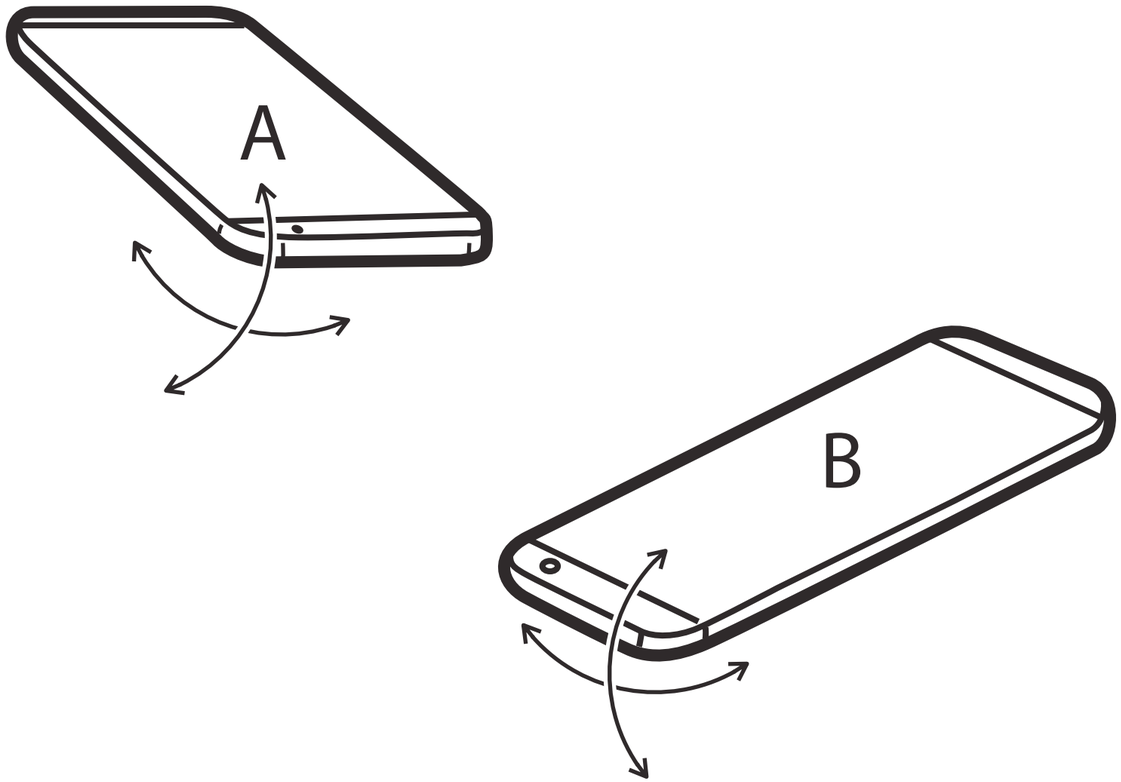
distances -> 3D positions



residual model

=> **generic** solution

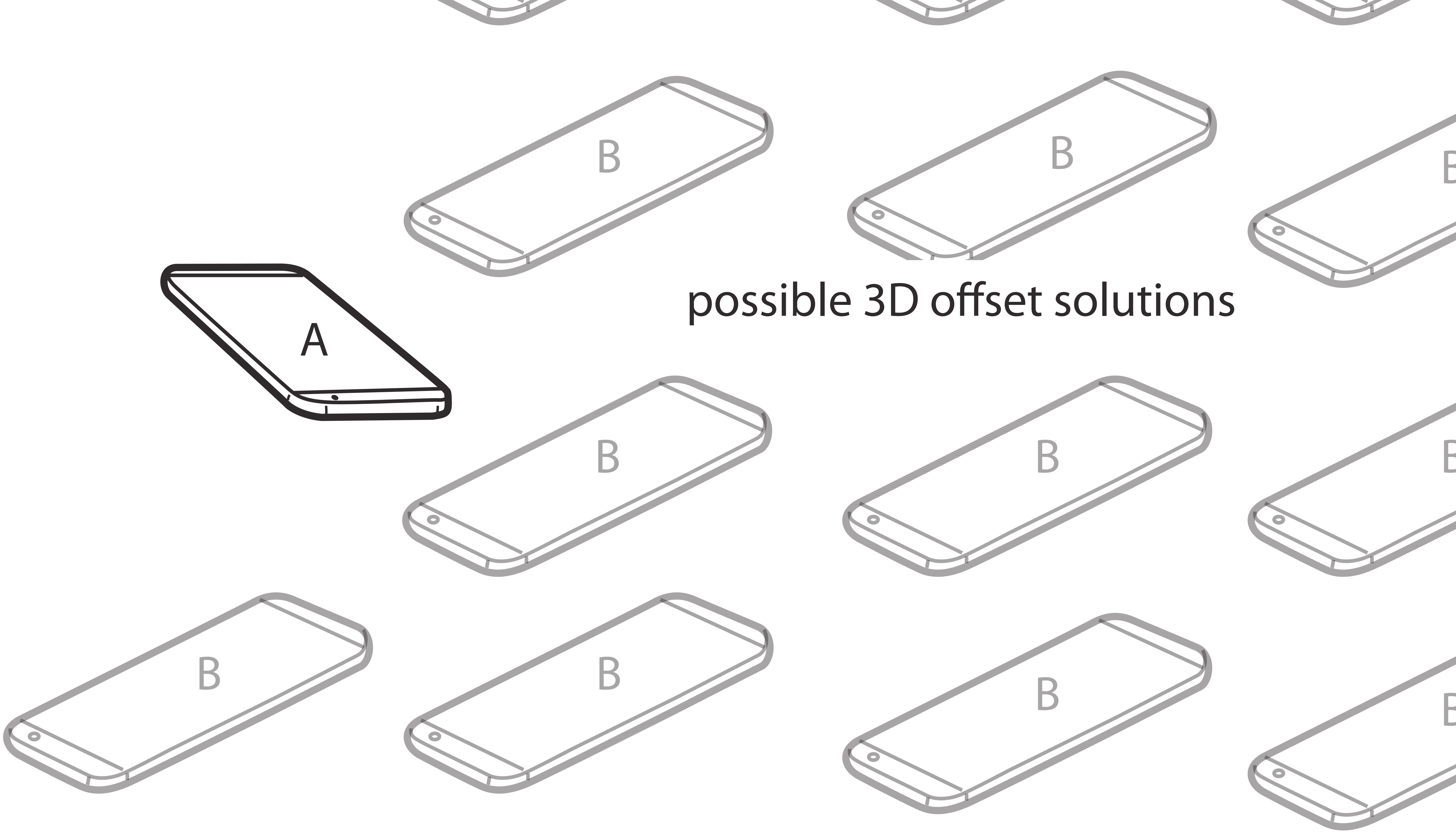
Tracko’s 3D offset calculation



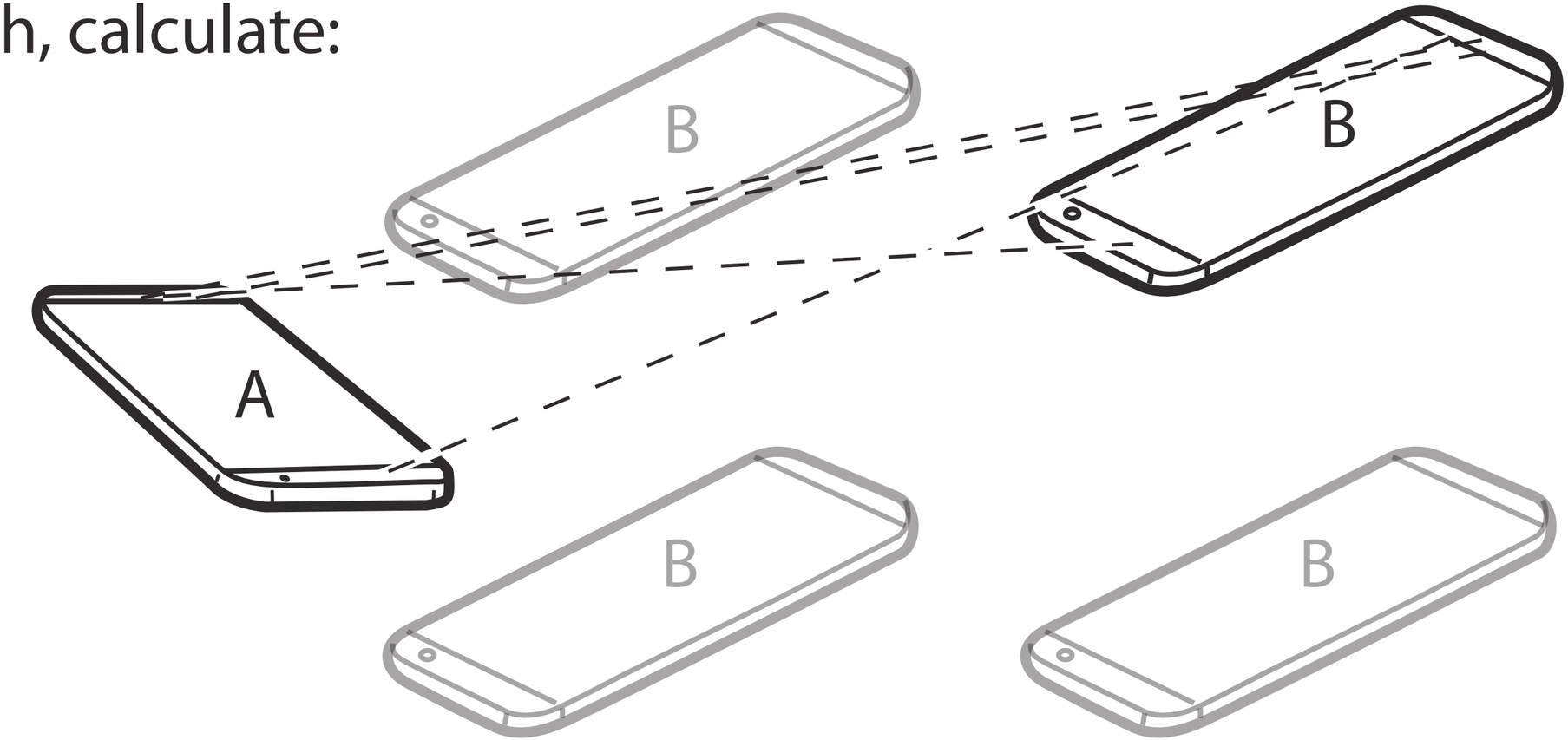
Tracko knows both phones' 3D orientations

accelerometer, gyroscope, magnetometer

3D orientations from the IMU

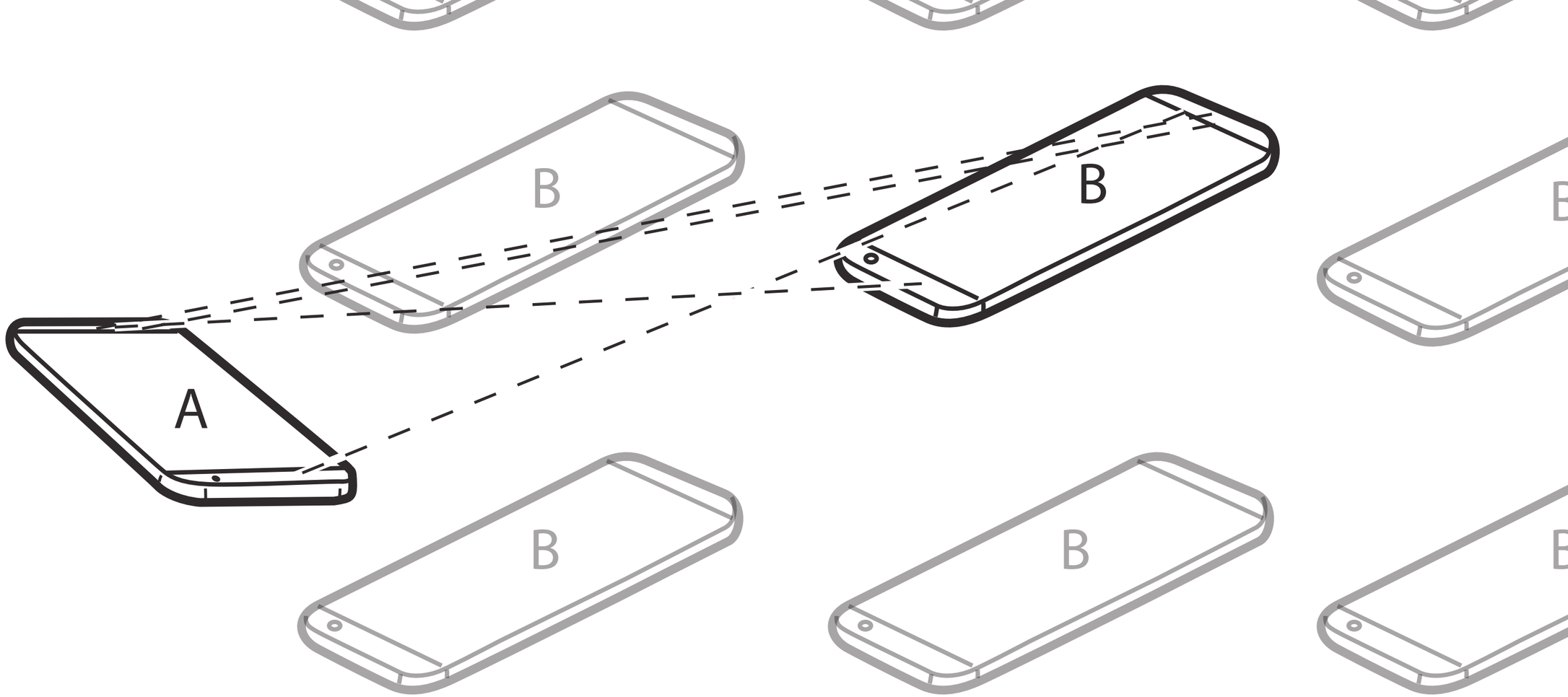

A
B
B
B
B
B
B
B
possible 3D offset solutions

for each, calculate:



$$d^*_{LL} = spA_L mic_B + spB_L mic_A$$
$$d^*_{LR} = spA_L mic_B + spB_R mic_A$$
$$d^*_{RL} = spA_R mic_B + spB_L mic_A$$
$$d^*_{RR} = spA_R mic_B + spB_R mic_A$$

estimated distances



$$SE = \sum (d^*_{XY} - d_{XY})^2, where X, Y \in \{L, R\}$$

gradient descent determines one
3D offset in a **discrete grid**

3 Inertial Sensors

1. gyro transforms Tracko's **local** coordinate system

2. accelerometer detects small movements

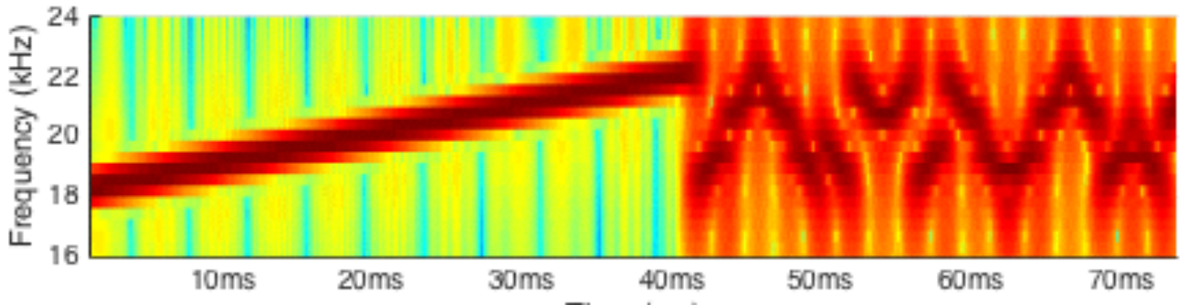
3. upon a new observation, **temporary** adjustments reset

temporary dead reckoning

evaluation

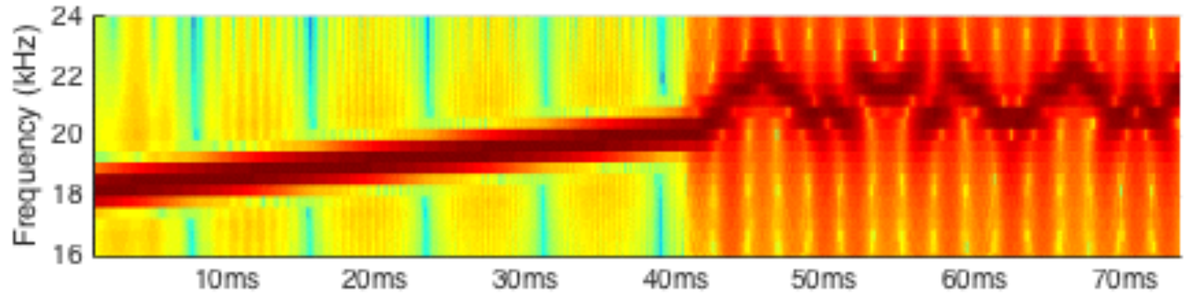
1) audio-based distances and 3D offsets
2) Bluetooth-based distance estimation

frequency ranges



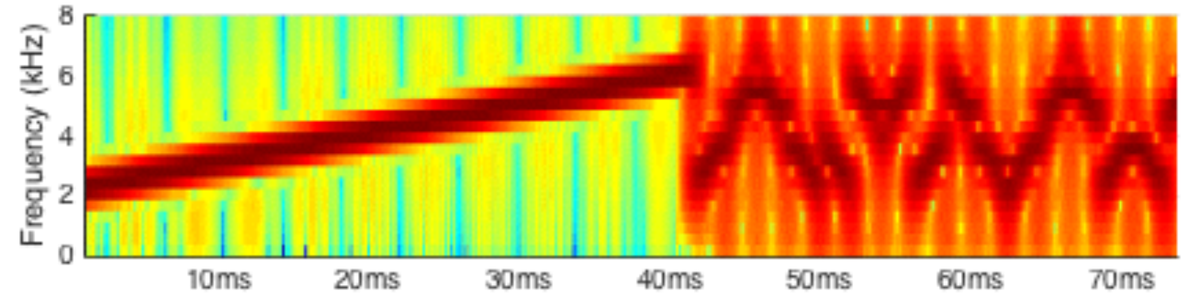
joint inaudible

18k-22k



disjoint inaudible

18k-22k, 20k-22k

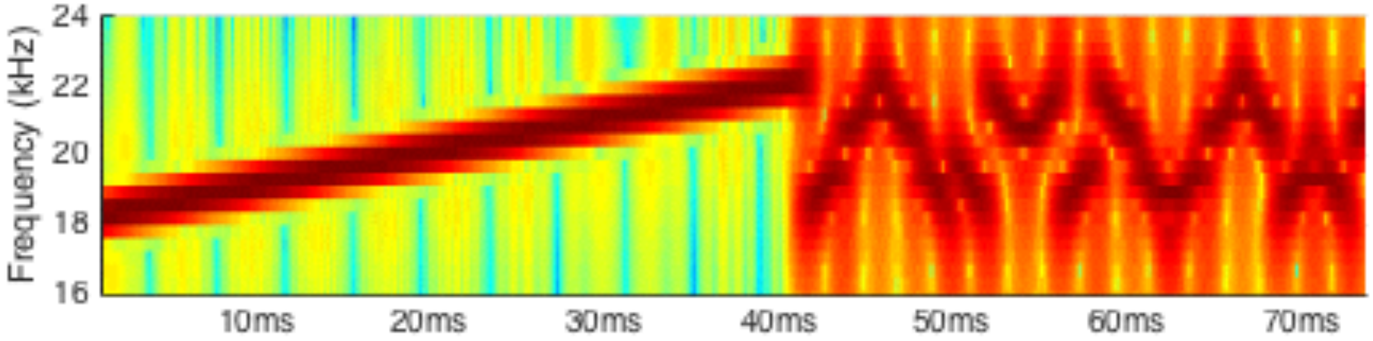


joint audible

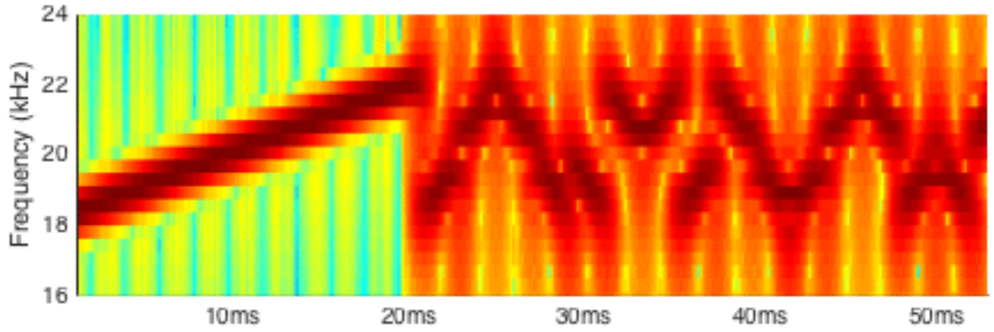
2k-6k

beacon lengths

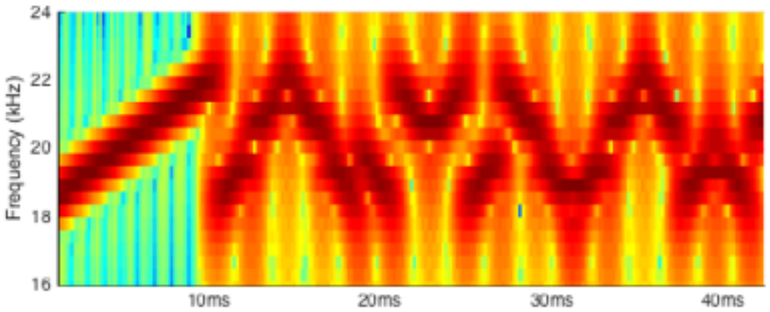
2000 frames



1000 frames

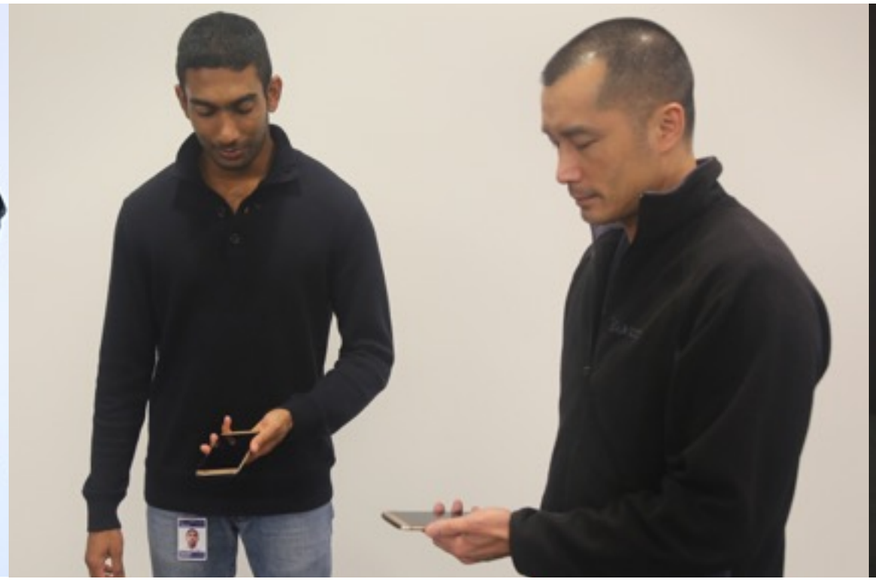


500 frames



distance ranges

close
< 0.5 m



medium
0.5 m – 1 m

far
> 1m

ground truth: Optitrack

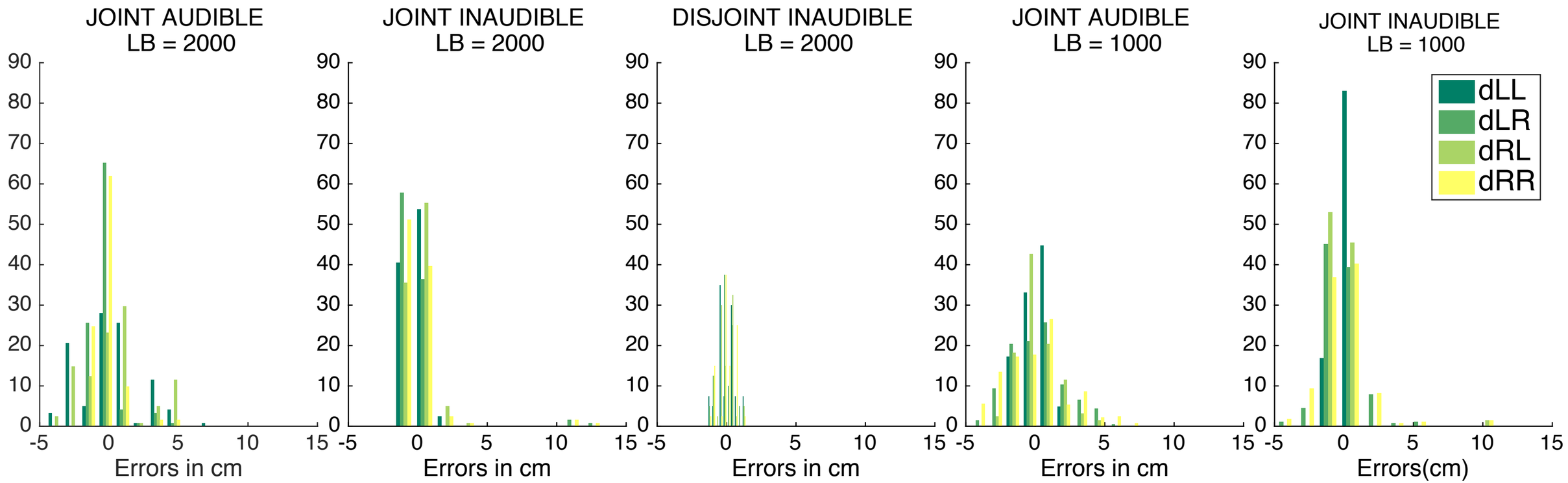
11 cameras
sub-mm accuracy

dependent variables

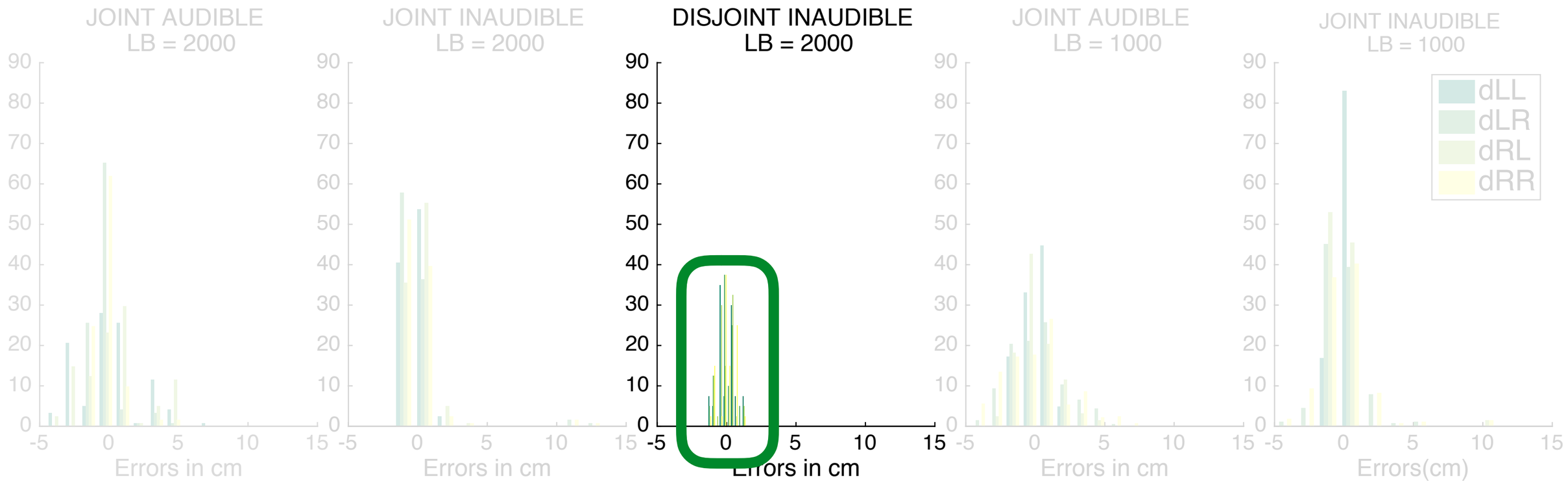
1. audio recognition accuracy => paper
2. distance accuracy
3. 3D offset accuracy

results

distance accuracy

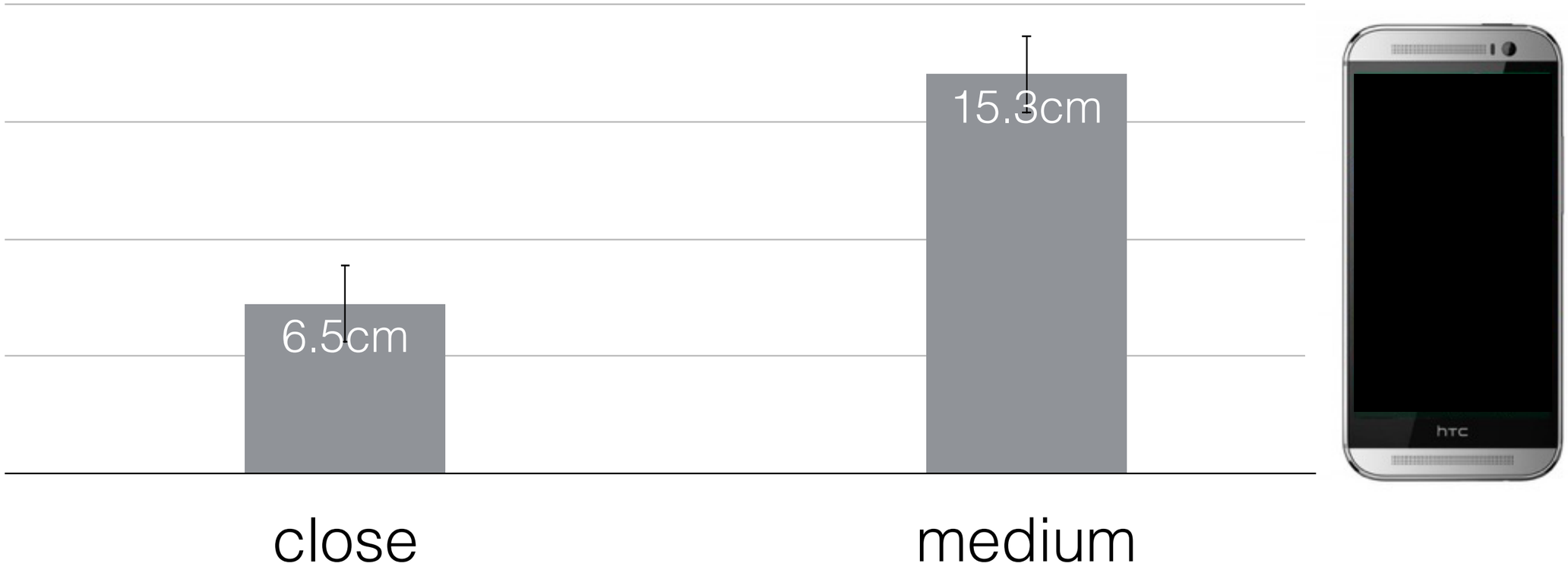


distance accuracy



average distance error = 2 cm

Tracko's 3D accuracy



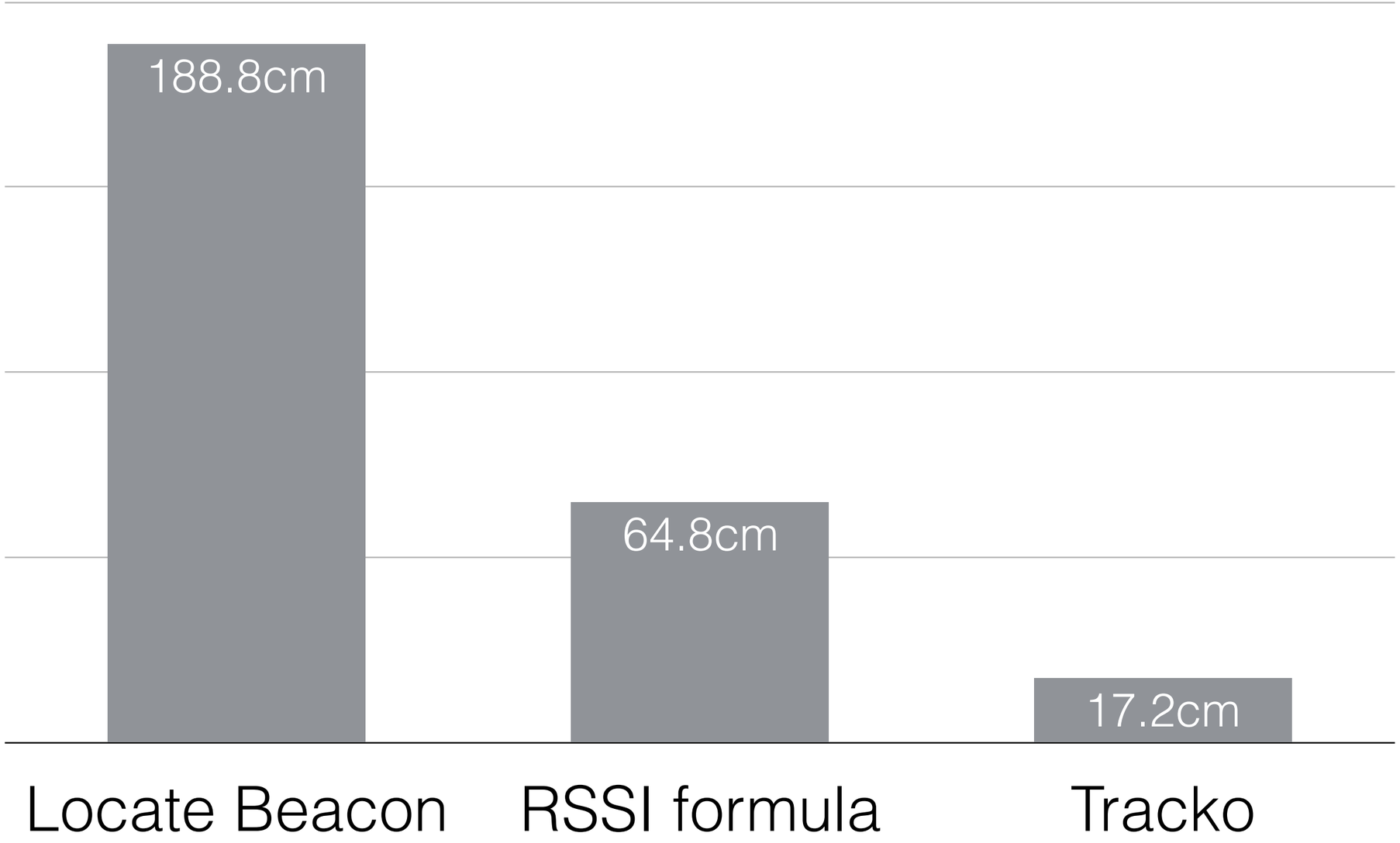
evaluation

1) audio-based distances and 3D offsets

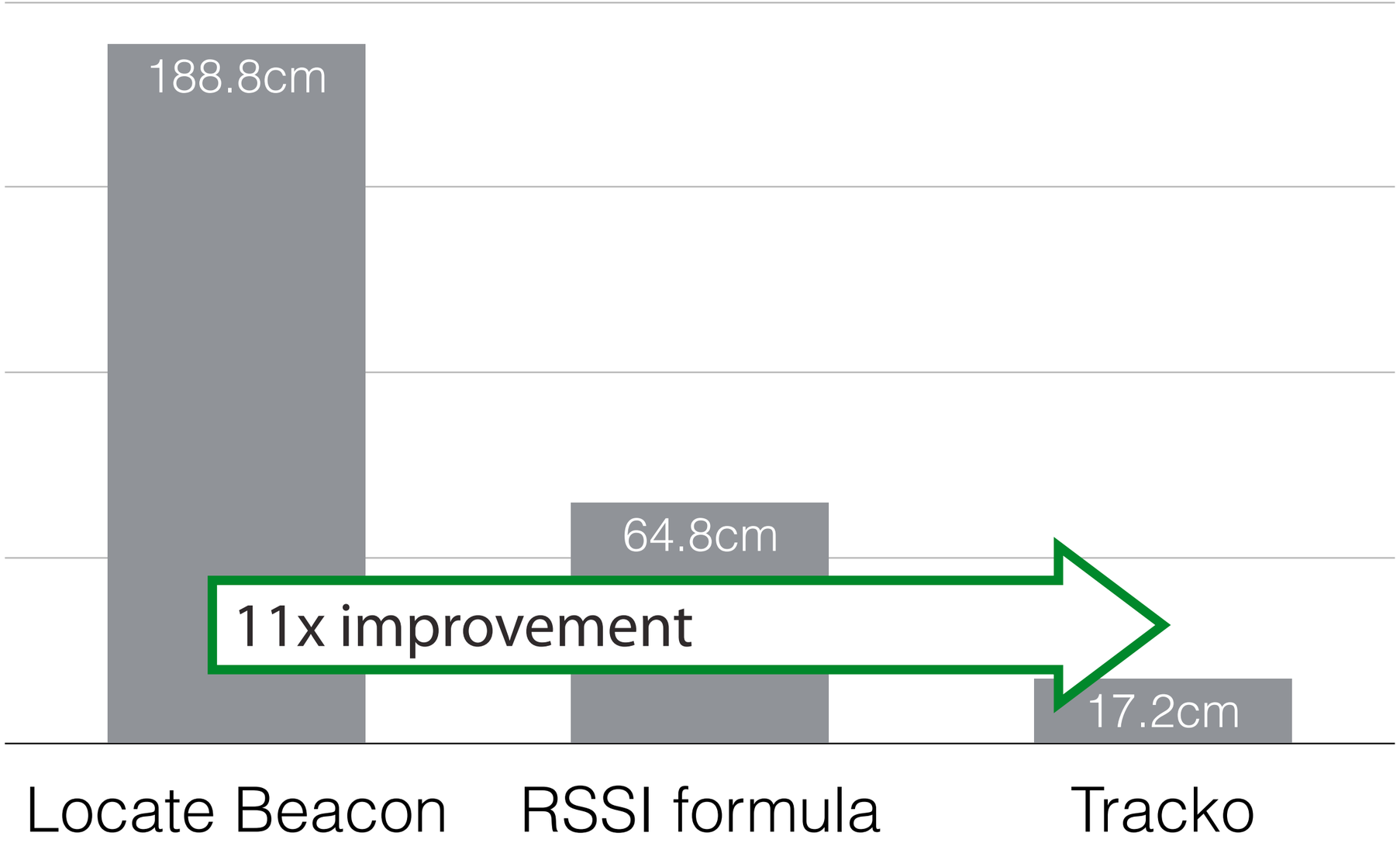
2) Bluetooth-based distance estimation


design

BLE performance



BLE performance



conclusion

mobile device-to-device 3D tracking

works ad-hoc: no synchronization or calibration required

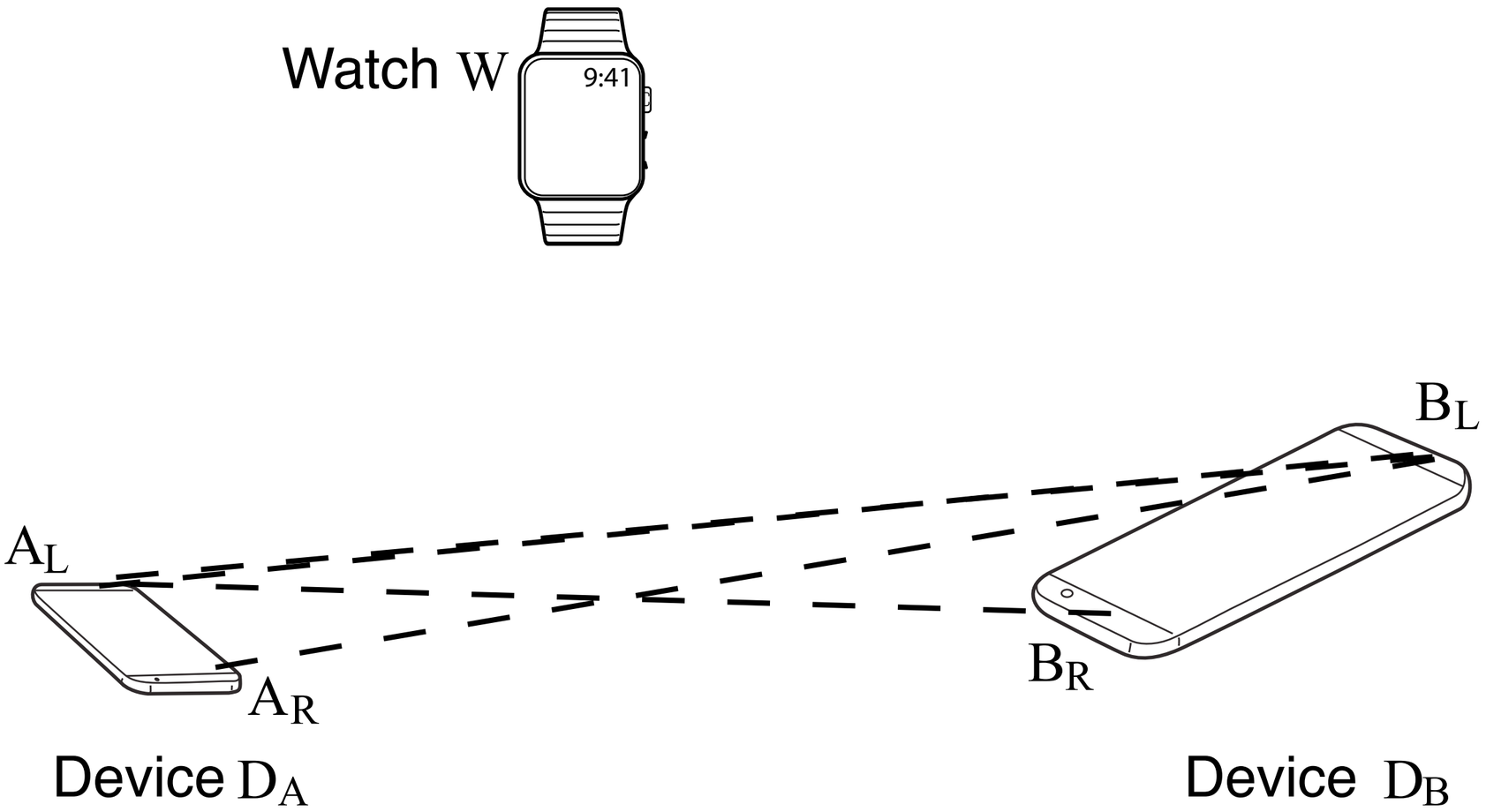
3D tracking accuracy suitable for cross-device interactions

runs on out-of-the-box commodity devices

Tracko

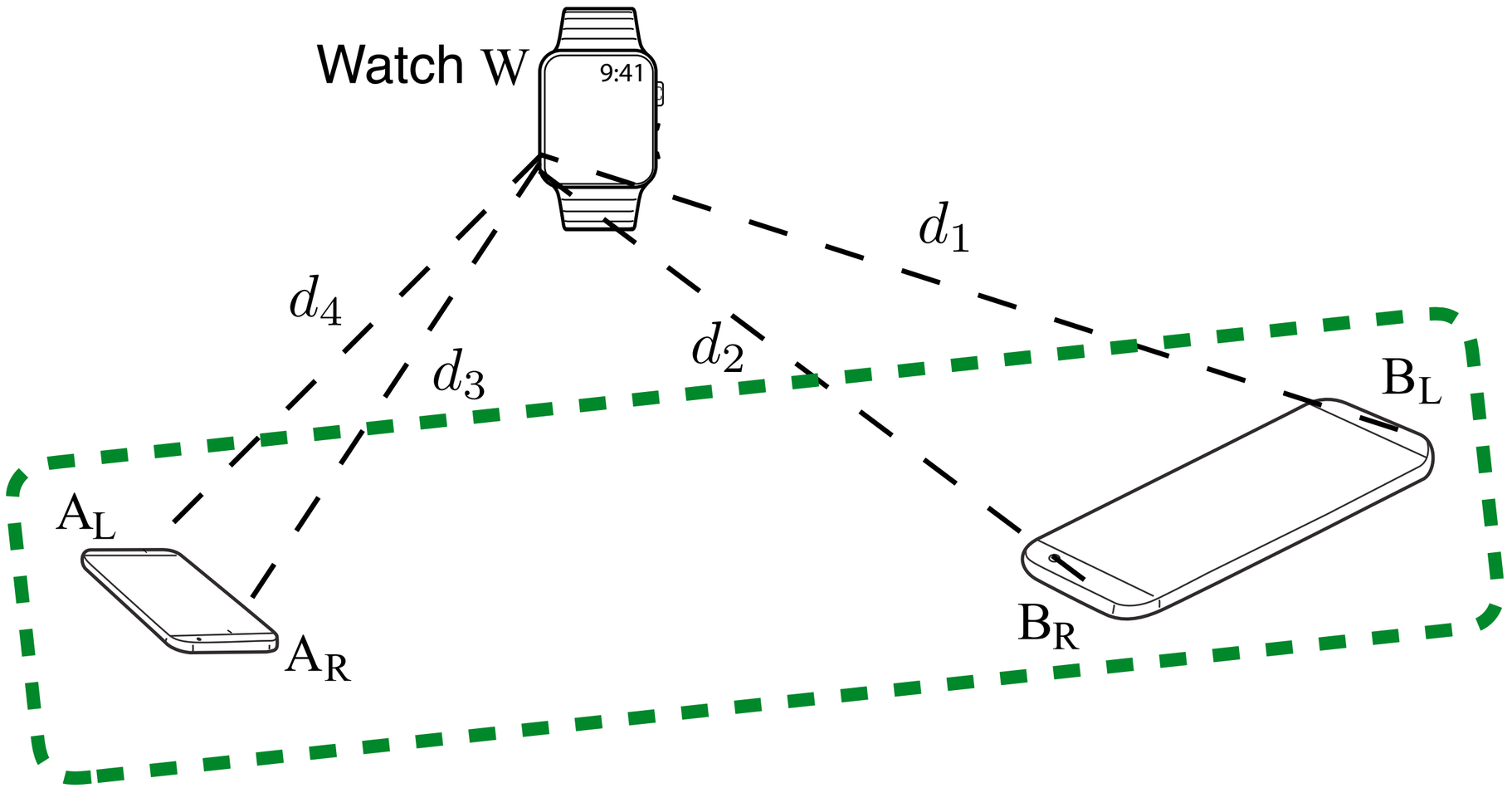

8.8 GOOGLE NEXUS 6P
9 GOOGLE NEXUS 6
SONY XPERIA Z5 PREMIUM
8.3 HTC ONE M9
8.8 HTC ONE (M8)
9.2 SONY XPERIA Z2
8.5 APPLE IPAD MINI
9 SAMSUNG GALAXY TAB PRO 8.4
7.5 BLACKBERRY Z30
9.3 GOOGLE NEXUS 9
8.5 SONY XPERIA Z3V
8 SAMSUNG GALAXY TAB 3 8-INCH
8.5 T-MOBILE G2X
8 ASUS TRANSFORMER BOOK T100TA
LENOVO TAB S8-50
SAMSUNG GALAXY TAB ACTIVE
8 SAMSUNG GALAXY NOTEPRO 12.2
8 APPLE IPAD MINI 3
9.1 HTC ONE
8.5 SONY XPERIA Z3
9 SONY XPERIA Z3 COMPACT
9.1 MOTOROLA MOTO X STYLE
APPLE IPAD PRO
9 APPLE IPAD AIR 2
8.5 KYOCERA BRIGADIER
9.2 APPLE IPAD AIR
SAMSUNG GALAXY TAB PRO 10.1
9 SONY XPERIA TABLET Z
HTC ONE E9+
9 SAMSUNG ATIV Q
8.6 ONEPLUS ONE
HUAWEI MATE S
9 SAMSUNG GALAXY TAB S 8.4
SONY XPERIA Z5 COMPACT
7 SONY XPERIA Z3+
8.5 HTC DESIRE 816
ZTE ZMAX 2
6.5 SONY XPERIA T3
HTC DESIRE 610
HTC ONE M8S
VERTU SIGNATURE TOUCH (2015)
8 SAMSUNG GALAXY TAB 2 (7.0)
9 SONY XPERIA Z2 TABLET
8.5 LG OPTIMUS 2X
9 LG ENV TOUCH
8 HTC BUTTERFLY 2
SONY XPERIA Z2A
LG G PAD X 8.3
9 SAMSUNG GALAXY TAB S 10.5
8 HTC DESIRE 820
8.5 MOTOROLA MOTO G (2014)
6 SAMSUNG GALAXY TAB 3 7.0
MOTOROLA MOTO G LTE (2014)
8.5 ALCATEL ONETOUCH IDOL 3 (5.5")
9 GOOGLE NEXUS 10
6.5 SAMSUNG GALAXY TAB 3 10.1-INCH
LENOVO YOGA TAB 3 PRO
HUAWEI UNION
7 SONY ERICSSON LIVE WITH WALKMAN
SKYGUS V2
8.5 SAMSUNG GALAXY NOTE 10.1 (2014 EDITION)
8 HTC DESIRE EYE
GOOGLE PIXEL C
8.3 HTC ONE MAX
8.7 HTC ONE (E8)
9.5 GOOGLE NEXUS 7 (2013)
ACER PREDATOR 8
8.5 APPLE IPAD MINI 2
APPLE IPAD MINI 4
VIVO XPLAY 3S
ALCATEL ONETOUCH IDOL 3C
HUAWEI G8
7 HTC DESIRE 600
8 ALCATEL ONETOUCH IDOL 3 (4.7")
6.8 ASUS ZENPAD S 8.0
8 LG VOYAGER
LAVA XOLO X900
LENOVO A7-30 A3300
9 GOOGLE NEXUS 7
SONY XPERIA Z4V
HTC ONE ME
9 SAMSUNG GALAXY NOTE 8.0
6.5 AMAZON FIRE PHONE
9.1 MOTOROLA MOTO X PURE EDITION (2015)
8 NOKIA LUMIA 2520
VERIZON WIRELESS ELLIPSIS 7
8.5 MICROSOFT SURFACE RT
7.5 LG OPTIMUS BLACK
7.5 ACER ICONIA TAB A500
SAMSUNG CONVOY 3


…only one microphone



even if the device has single audio unit

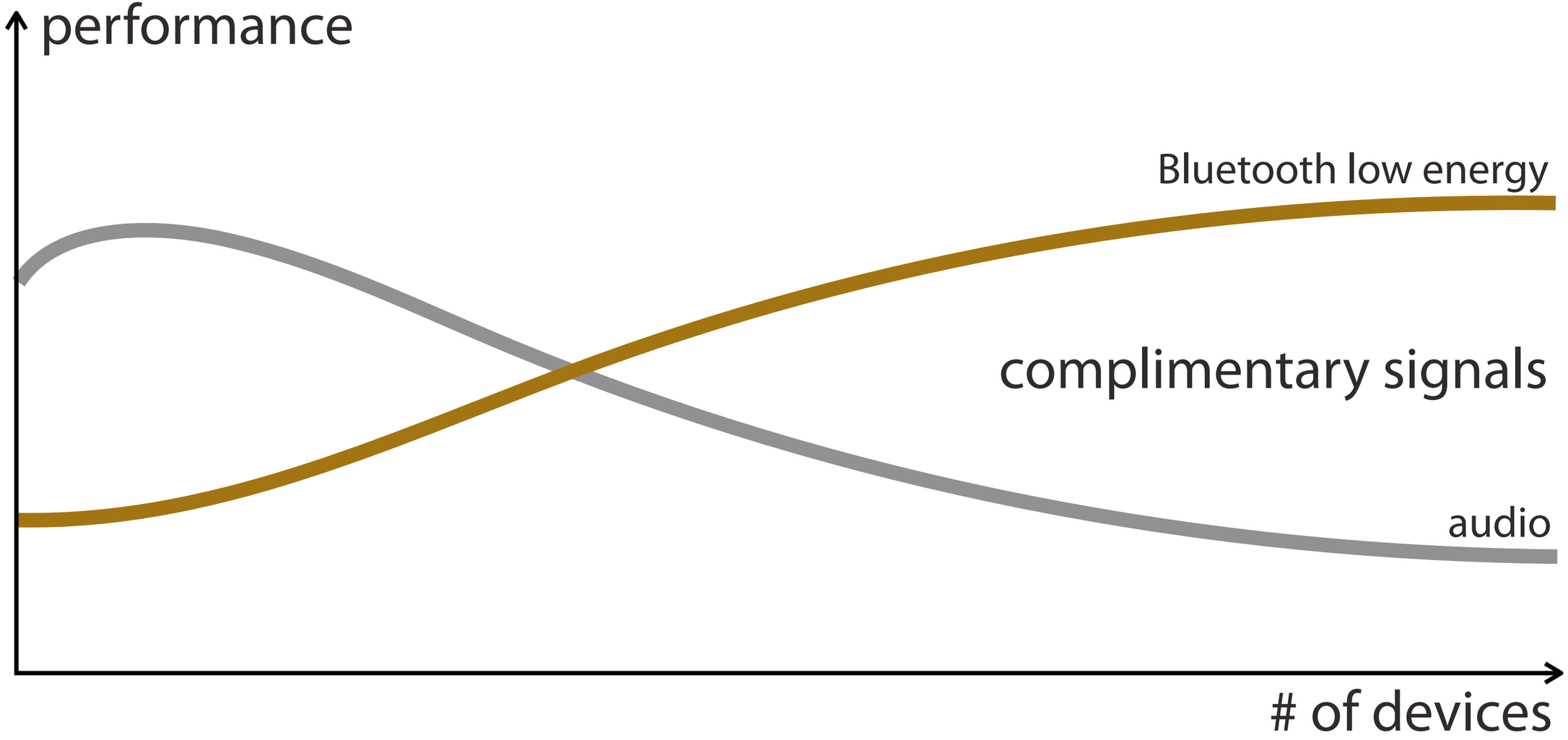
=> paper



one logical device with **four speakers**


built-in stereo
speakers x2
built-in stereo
speakers x2
09:41

Tracko scales with number of devices



Tracko scales with number of devices

and can even integrate into
BLE-based **indoor tracking systems**

Tracko

Ad-hoc Mobile 3D Tracking Using Bluetooth Low Energy and Inaudible Signals for Cross-Device Interaction

Haojian Jin — Yahoo Labs
Christian Holz — Microsoft Research
Kasper Hornbæk — University of Copenhagen