



WHITE PAPER

The promise of zero-latency hand tracking

ABSTRACT

The shift from touchscreen to touchless interfaces and from real world entities to multiple mixed and virtual ones creates a new significance for our hands. The need for touchless interaction creates an opportunity to leverage hand movements and gestures. The worlds of mixed and virtual realities require a replacement for traditional input modalities, new ones that allow us to manipulate virtual objects naturally.

The challenge is how to harness the hand tracking opportunity on existing computational devices with low power consumption and zero latency.

This white paper aims to help you uncover the promise of hand tracking, to figure out if this sensor technology presents you with an opportunity to grow your business* and to show how we at ManoMotion have solved the zero-latency challenge.

*PwC "Seeing is believing" [report](#) claims that VR/AR represents a 1,5 trillion-dollar opportunity for the global economy 2030!

IS THIS WHITE PAPER FOR YOU?

Hand tracking empowers devices and applications with the capability to identify hands, interpret human hand gestures intuitively, and even enable hand interactions with virtual objects. This white paper aims to help you identify if and what growth opportunities this technology may present to your business.

We think you are probably responsible for innovation in some way, you might work in product development, R&D, or perhaps you're the CEO/CTO of a tech start-up. You might be exploring AR as a new way to deliver solutions. You may work with IoT, mobile, or VR. Whatever your position, industry, or product, when you've reached the end of this paper, we hope you will have gained an understanding of:

- Hand tracking today and what it is capable of.
- The coming era of the touchless interface.
- The opportunities hand tracking presents in hand safety applications, AR and VR, as well as just about any scenario that involves a person's hands.
- The technical challenges.
- How the ManoMotion framework helps you move from idea to development.

Primarily, we want you to be able to answer this question:

Is hand tracking an opportunity for your business to beat competition and deliver beyond your customers' expectations?

WHAT IS HAND TRACKING?

Hand tracking is a sensor technology that enables an application to leverage hand gestures to control a device or an object without having to touch it. It provides a means to remotely control physical or virtual objects through precision understanding of hand movements.

At the outset, hand tracking relied on sensor-equipped gloves to generate the necessary data to interpret hand movement. Today's gloveless hand tracking uses a camera sensor that takes images of a person's hands, leveraging computer vision* to interpret movement and deliver an accurate set of data points in real time.

*Computer vision is the use of machine learning and AI to interpret images captured by a camera sensor to extract objects and meaning.

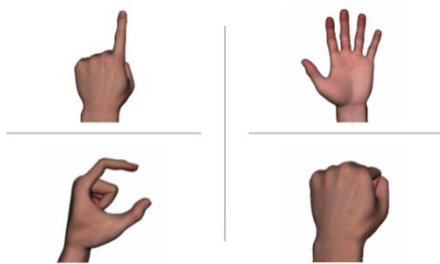
THE ERA OF THE TOUCHLESS INTERFACE

The way we interface with technology — human-computer interaction — has too evolved over time, continually reducing the effort it takes to get machines and devices to do things.

The recent developments, like of touchscreens, gesture-based controls, and voice activation have truly succeeded in taking the grunt work out of device interaction. And perhaps what is most interesting about these newer input modalities is that they tap into our senses, leveraging our innate forms of interaction. Generations are growing up to expect nothing less.

Today, we are on the brink of the next generation: the era of the touchless interface.

Hand tracking can cut costs because it replaces hardware (input devices), with a person's hands. Also, our hands are a natural way for us to communicate and get stuff done. And hands are of particular importance in the world of augmented, mixed, and virtual reality (AR, MR, and VR) because there is no room for traditional input modalities like keyboards and mouse in these environments.



But there is also the opportunity to equip both consumer devices or industrial equipment with sensors and enabling them to leverage hand tracking to detect hands, as a safety feature. Or provide gesture control to give input and control the machine.

Grab, pinch and point will be the future of machine interactions!

The future is here

Strolling relevant tradeshow, you will realize that the future is here. Over the last few years, there has been a multitude of applications demonstrated in public. Just to mention a few:

- clicking on virtual buttons on a keypad by pointing with a finger
- changing channel on a TV by waving left or right with your hand
- opening/closing windows in a car or changing volume on the radio
- navigating information displays/maps on shopping malls
- controlling mobile phones or computers with gestures

THE HAND TRACKING OPPORTUNITY

Hand tracking creates new business opportunities in many areas, here are a few examples of these:

The shift to touchless

The coronavirus pandemic has pushed many touchscreen terminal vendors to reconsider their designs, evolving them toward touchless, gesture-controlled solutions. For a rapid fix, hardware solutions offered by [Ultraleap](#), for example, are a straightforward initial step to transform existing designs into gesture-controlled interfaces.

In the future, it's likely that this type of dedicated hardware solution will be integrated into the devices, and in some scenarios, like mobile, it will be possible to adopt a software-only approach, by leveraging existing device sensors.

Safety, IoT, and sensors everywhere

The IoT-concept is enabled by massive and distributed processing power, sensor proliferation, and ubiquitous connectivity. These combined capabilities create opportunities to design new solutions in consumer and industrial settings that leverage hand tracking.



For example, our joint development with industry leader Altendorf, leverages hand tracking technology to deliver the world's first safety-classified solution for industry-grade saws.

Altendorf is an example of insightful and ambitious leadership looking to solve hands-on problems in a traditional industry with cutting-edge technology. In the solution, Altendorf developed a new safety feature that uses hand tracking to detect when an operator's hands are too close to the cutter, automatically shutting the machine in the event of an imminent accident. The solution provides Altendorf with a unique selling point, helping them to position and differentiate. See the [video](#)!

The shift to virtual

VR

The fine motor skills that our hands are capable of are hard to emulate with buttons and handheld controllers, and so the coming generation of VR applications will leverage hand tracking to enable people to 'touch' and manipulate virtual objects in a way that feels natural — providing developers with a broad range of new application areas to explore.

AR

AR is getting a firm grip in the industry and enterprise market. Top areas of application include service-related tasks, skills transfer, and manufacturing processes.

A common denominator for AR applications is a foundation in the physical world, with a need to leverage opportunities provided by the digital world — such as training on virtual objects and defining business processes without the physical constraints of the real world.

Hand movements play such a pivotal role in the way people communicate and interact with their environment that the ability to track them plays a central part in creating the bridge between the physical world and its digital twin.

Mixed reality

As the headset market evolves from basic mobile clip-ons to advanced mixed reality headsets, understanding and formulating hand tracking becomes more complex. Basic hand detection and control functions can be managed on mobile platforms. Advanced gesture control and hand-object interaction requires greater computing capabilities. The mixed reality headset Varjo XR-3, for example, addresses this opportunity using video see-through.

Mobile AR

With hand tracking solutions available for deployment in smartphone AR apps, Mobile AR may not offer the same level of power as tethered headset-based VR/AR solutions, for example, but it does provide a business opportunity to develop at significantly lower cost for an established and ubiquitous platform.

According to Gartner — who have graduated AR from the Trough of Disillusionment on the Hype Cycle to a mature technology — in 2020, 100 million consumers were projected to shop using mobile AR platforms (Android and iOS).

Tech giants like Google, Apple, and Facebook are explicitly committing to this market opportunity, which is clear through the investment in APIs and development kits like ARKit, MediaPipe, and ARCore that enable developers to explore new type of interactive experiences — including hand tracking.

THE TECHNICAL CHALLENGES

Zero latency

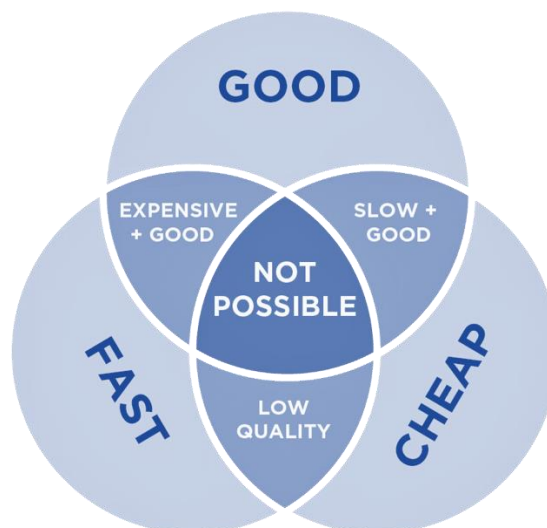
Latency (noun) — *the time delay between cause and effect.*

In hand tracking, latency is the time between a hand movement in reality and the perceived movement in the virtualized environment or application. For hand tracking, latency is primarily a function of computational complexity.

To feel natural, hand tracking latency needs to be less than 15ms. This is what we mean by zero (perceived) latency.

Like most sensor-based technologies, building a reasonably good solution is not that challenging. The challenge lies in being able to deliver a good outcome in real time, with low computational load. It's a classic balance of good-cheap-fast.

- **Good** — the algorithm needs to be good enough to accurately represent hands and their movement
- **Cheap** — it's possible to run the algorithm on available hardware without draining the battery
- **Fast** — it's possible to run the algorithm on available hardware with zero latency



Computational complexity

Conventional hand tracking relies on advanced image and signal processing to translate the highly granular movements of a person's hands in 3D space into a data stream that can be interpreted by an application.

Most applications require hand tracking at millimeter precision with 27 degrees of freedom. This level of precision together with tracking of multiple hands, object identification, and hand-object occlusion recognition creates such a huge state space that it requires the power of a supercomputer to solve in real time.

And so, it seems that meeting the demand to recreate real-world experiences in a mixed reality environment in real time and at the required level of precision is unsolvable with conventional methods.

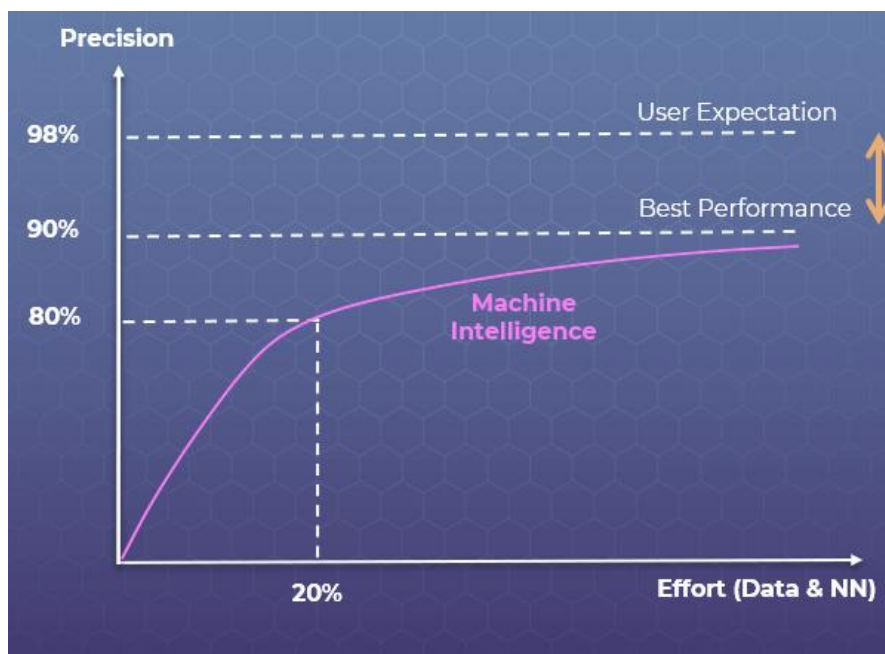


Figure 1: User expectation gap

SOLVING THE TECHNICAL CHALLENGE

ManoMotion’s patented framework RealHands™ solves the technical challenge.

To overcome the latency/computational complexity challenge, ManoMotion has redefined the problem. Instead of trying to solve a massive compute problem, we have adopted an AI machine-learning approach. We have created a data warehouse that stores a representation of all the variations of hands, hand movements, gestures, and skeletal information — including how hands relate to other objects and how they interact with objects.

Our framework leverages the power of computer vision to deliver an AI-based gesture search-engine. It transforms the conventional, computationally-complex, tracking approach into a pattern-matching problem that can be solved accurately and fast. Our approach improves probability of delivering zero latency.

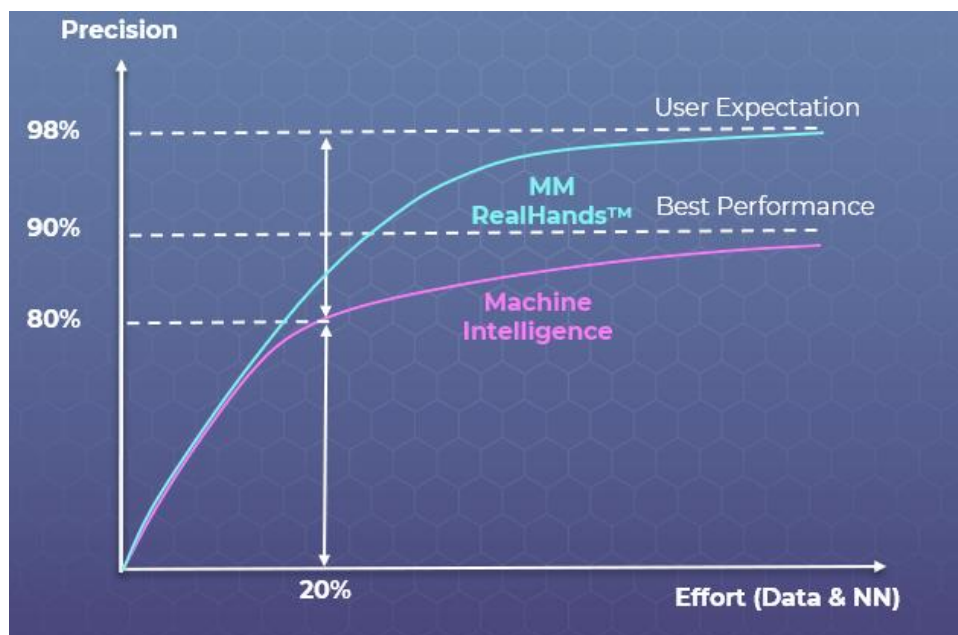


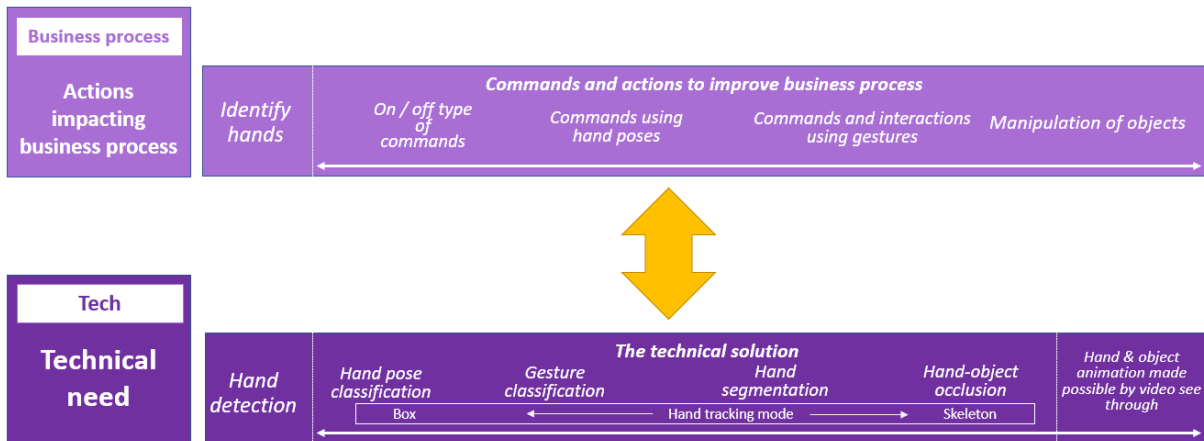
Figure 2: Closing the user expectation gap

Taming the computational hunger of complex AI/ML is a must. RealHands™ makes it possible to reduce computational complexity from 10^{20} to 10^6 . That is crazy!

Shahrouz, CTO and founder of ManoMotion

ARCHITECTING A SOLUTION FRAMEWORK

The ManoMotion Hand Tracking Solution Framework™



To ease the transition from solving the business use case to defining the technical solution, we have created the Manomotion Hand Tracking Solution Framework. Our aim is to connect the purpose of the solution with how we help you to technically solve the challenge.

Complexity of the user experience

- **Hand detection** — detect presence or absence of a user's hand(s) and whether it's a left or right hand
- **Hand position** — locating a user's hand position
- **Defined gestures** — detecting pre-defined hand poses and movements
- **Natural gestures** — detecting complex movements and poses in real-time
- **Object manipulation** — processing user's hand interacting with virtual and/or physical object

Complexity of the environment

Hand recognition — environment characteristics that impact the complexity of hand recognition include lighting conditions, hand occlusion (partial or complete), hand movements, hand position upon recognition, and hand movements during the recognition phase.

Gesture complexity — gestures range from **simple hand movements**, such as point-to-select, to **advanced**, such as turning the handle of a ratchet. Advanced **gestures are often executed in a personalized way**, which adds to the complexity.

Box or full-skeleton mode — user interaction is possible in both modes. Simple point to select, for example, can be achieved in box mode, whereas typing on a virtual keyboard requires full-skeleton tracking that consumes massive computational resources.

Advanced interaction — the ability to manipulate physical or virtual objects without the need for glove-based sensors is a massive differentiator for gloveless hand tracking. New class of HMDs are offering video see-through, creating an opportunity to blend real world and virtual world interactions, including hand tracking.

SOLVING THE BUSINESS CHALLENGE

Our ambition with this paper is to trigger ideas of how the shift from touchscreen to touchless interfaces and from real world entities to multiple mixed and virtual ones creates a new significance for our hands.

And for your business!

Ask yourself the following three questions:

- Have you identified a business opportunity that requires hand tracking at some degree of immersivity — hand recognition, gesture control, or object manipulation — that solves a previously unsolvable problem or provides meaningful user experience?
- Do you see a value in using a software solution that is hardware and sensor agnostic?
- Do you want to embed the solution in a proprietary system or on Android or iOS platform?

If you have answered yes to these three questions, you should consider working with ManoMotion and our patented gesture search-engine solution RealHands™, because we believe we can help you tilt the performance triangle to your advantage, providing zero latency in a computationally efficient way.

WHAT NOW?

How can we help you explore if hand tracking represents an opportunity for your business to beat competition and deliver beyond your customers' expectations?

The first thing we like to do is set up a call, where you tell us about your requirements and the scope of your project.

If we feel we can help you to solve your problem, we'll recommend a one-day exploratory workshop, which we scope in collaboration with you. Or we can send you a remote camera exploration kit, which basically enables you to connect a webcam to your application and start experimenting with hand tracking.

If you decide to go with us, we will send you a project proposal to help get you up and running with our software development kit. And if you want, we'll be there throughout your journey to commercialization, as much or as little as you need.



JOIN US!

Manomotion is on a journey, developing technology enabling the shift from touchscreen to touchless interfaces and from real world entities to multiple mixed and virtual ones, even enabling us to manipulate virtual objects naturally. Let us explore this opportunity together!

www.manomotion.com

sales@manomotion.com

ManoMotion is a diverse team of developers. We have explored machine learning and AI in the context of hand tracking and gesture analysis more than 10 years. Focus is to bring hand tracking functionality to devices using standard and built-in camera solutions. Aim is to always use minimal computing resources and minimal power. Our SDK solution are based on the Android and iOS ecosystems. Short term focus is on mobile solutions. But we are also engaged with HMD vendors, and we see new type of wearables emerging using hand tracking and gesture control as an integrated part of the solution. AR holds fantastic promises for the future – but there is no reason why you should not start building hand tracking and gesture control solutions today.