Negative Latency in the Tactile Internet as Enabler for Global Metaverse Immersion

OPEN CALL ARTICLE

Negative Latency in the Tactile Internet as Enabler for Global Metaverse Immersion

Jonas Schulz, Clemens Dubslaff, Patrick Seeling, Shu-Chen Li, Stefanie Speidel, and Frank H. P. Fitzek

Abstract

The future Metaverse has been proposed as a solution to various social problems, requiring new concepts. This evolution of the Metaverse will be geared toward solving real-world tasks by combining humans and machines in virtual and real spaces. This, in turn, necessitates the inclusion of additional human senses, including the tactile modality, to enable fully immersive human-machine and human-machine-human cooperation. The resulting challenges of communicating ultralow latency information below a 1 ms threshold have guided the Tactile Internet's approaches to decrease various additive delay components. However, physically determined propagation delays geographically limit the future Metaverse when considering all human senses. We propose the concept of negative latency to address this limitation for the next evolution of the Metaverse that enables seamless and global user interactions. Our negative latency concept provides guarantees on predicted user actions to reduce perceived latency with high confidence. We evaluate a proof-ofconcept setup for grabbing tasks with a smart sensor glove and show that negative latencies of several hundred ms are attainable. Taking advantage of this and other advanced technologies in the Tactile Internet, we envision a truly global, immersive, and interactive Metaverse 2.0 that supports meaningful interactions between humans and machines.

Introduction and Motivation

The current view of the Metaverse, as proposed by Meta (formerly Facebook), focuses on creating a social network around the general concept of the Metaverse: democratise access to a customisable, interconnected, social, and highly immersive virtual environment with high-end human-machine interfaces (HMIs) to create a virtual space where people can interact with each other and with digital systems in ways that were previously not possible. This concept of the Metaverse has been the public focus of attention since the early 2020s, e.g., in the context of gaming or virtual collaborative spaces.

Virtual worlds have been broadly deployed since the early 2000s. Second Life, as one of the most prominent examples, is often considered one of the first Metaverse realisations. Early implementations of virtual worlds typically provided simplified interfaces for users to enjoy themselves and explore digital spaces in a social context. These environments can be regarded as refinements of the 3D virtual worlds that were popularised in the video game domain since the early 1980s. Over the years, both categories benefited from continuous improvements of graphics and end system capabilities as well as increasingly capable HMIs.

These advances notwithstanding, the current focus of the Metaverse on the virtual domain neglects how humans interact with virtual **and** physical environments. The next evolutionary step of the Metaverse can be regarded as the addition of physical-space interactions that are driven by virtual-space interactions. In this article, we discuss the enablers and requirements for transferring interactions from the virtual world back to the real world, such as computing, communication, robotics, peripheral devices, and the principles of multisensory human perception and action. Due to space constraints here, we refer the interested reader to [1] for a comprehensive overview of the domain. The amalgamation of human and machine interactions for real-world problem solving can unlock new solutions to some of the most pressing challenges humanity faces today, such as pandemics, aging, medical care, or workforce shortages due to a lack of skills in certain disciplines.

This article contributes an overview of how *negative latencies* in the Tactile Internet will enable the evolution of the Metaverse beyond avatar-based digital social interactions and expand its fully immersive reach to physical spaces on a global scale:

1. We discuss the increased requirements from this multisensory evolution of the Metaverse and the resulting demands on the network

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

Jonas Schulz is with the Deutsche Telekom Chair of Communication Networks, Technical University of Dresden, 01187 Dresden, Germany; Clemens Dubslaff is with the Formal System Analysis Group, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands; Patrick Seeling is with the Department of Computer Science, Central Michigan University, Mount Pleasant, MI 48859 USA; Shu-Chen Li is with the Chair of Lifespan and Developmental Neursoscience, Technical University of Dresden, 01217 Dresden, Germany, and also with the Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI), 01062 Dresden, Germany; Stefanie Speidel is with the National Center for Tumor Diseases Dresden (NCT/UCC), 01307 Dresden, Germany, also with the German Cancer Research Center (DKFZ), 01307 Dresden, Germany, also with the Faculty of Medicine, University Hospital Carl Gustav Carus, 01307 Dresden, Germany, also with the Technical University of Dresden, 01187 Dresden, Germany, also with the Helmholtz-Zentrum Dresden-Rossendorf (HZDR), 01328 Dresden, Germany, and also with the Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI), 01062 Dresden, Germany; Frank H. P. Fitzek (corresponding author) is with the Deutsche Telekom Chair of Communication Networks, Technical University of Dresden, 01187 Dresden, Germany, and also with the Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI), 01062 Dresden, Germany.

Digital Object Identifier: 10.1109/MNET.2024.3375522
Date of Current Version: 16 September 2024
Date of Publication: 11 March 2024

0890-8044/24©2024IEEE

that are predominantly fulfilled with Tactile Internet approaches to reduce service latency.
2. We provide a high-level overview of the origins of latency in the Tactile Internet and how networks are at their physical limits: Current latency levels that have been successfully reduced through many innovations in the Tactile Internet remain additive. Propagation delays limit the geographical reach for a truly immersive experience even below the common metropolitan area network (MAN) ranges of about 50 km.
3. We propose, discuss, and evaluate the concept of *negative latency* to solve the latency challenge and enable a global reach of the future Metaverse. We corroborate our concepts of multimodal sensing and negative latencies with a prototypical implementation and an experiment employing smart gloves.

Enabling the Metaverse Evolution

In this section, we discuss the prerequisites for the next evolutionary step for the Metaverse that will enable revolutionary applications and services in the cyber-physical domain.

Multisensory Requirements

The current Metaverse's focus is on avatar-based interactions in virtual spaces. This approach relies on the visual and auditory senses of the human body for content presentation (actuation) and dedicated controllers for navigation and viewpoint finding for visual display (sensing). Typically, actions in such settings can only be performed through dedicated controller devices geared toward a user's virtual space immersion, not interaction with remote real-world spaces.

Enabling applications and services that connect the virtual space back to the physical realm requires that all human sense representations can be carried over communication networks. Additionally, full immersion also requires haptic perception to make the next generation of the Metaverse perceived by people as accessible and natural. For example, the transmission of touch-based interactions can enable a more intuitive way of sharing skills than the use of visual-audio displays alone [1]. This, in turn, also requires that sensor and actuator technologies become generally available to sense human motion and provide the corresponding haptic stimuli.

Technologies that enable the perception of human motion and touch through wearable sensors and actuators are emerging in commercially available implementations, primarily in the form of smart gloves [2]. Haptic codecs play a crucial role and their development is an active research domain [3], including efforts to consider the potential impacts of aging on haptic perception [4].

Transmission of human senses (such as touch) in a truly perceivable, immersive, and plausible fashion is subject to ultralow latency constraints. The haptic sense (tactile/kinesthetic) has higher temporal precision requirements (approximately 1 ms) than auditory and vision (approx. 3 ms and 15 ms, respectively) in the relevant sensory regions (see, e.g., the review in [1]). As the most stringent latency requirement for human perception, action, and cognition stems from the tactile domain, the notion of the *Tactile Internet* has emerged. The term refers to to the need to support this latency requirement over communication networks. Several new approaches, such as network softwarisation and edge computing, contribute to achieving the goals of the Tactile Internet.

Furthermore, the neurochemical systems of the brain (e.g., dopamine) that are important to regulate perceptual processing decline in the course of aging (see, e.g., [4] for recent reviews). Current Metaverse implementations are challenged to support the required constraints imposed by the psychophysical and neurocognitive processes underlying human perception [5]. The result of the integration of tactile perception for **all** potential users in the Metaverse's next generation is an upper latency requirement. This requirement becomes a latency challenge due to the physical and other limitations of current communication networks, discussed in the following.

The Latency Challenge

We illustrate the overall latency in current computing (input/output, computation, networking, storage) systems with a focus on the individual components required for a realisation of the Metaverse in Fig. 1. In a common scenario where a human interacts with a remote system (virtual or real) and requires sensory feedback, the overall latency is composed of different delay-contributing components [6]: peripherals (sensing and actuating), wireless and fixed network links (communicating), and any additional in-network message and information processing (computing). (We note that this sequence is for a virtual-space interaction, and full remote control in a distant physical space would incorporate additional components for machine operations.)

The initial source of latency is the sensing performed by peripherals (e.g., smart gloves, headsets, or microphones), as sensing itself requires time. (We note that actuation can be regarded in a similar way.) The resulting delay can only be reduced through dedicated hardware

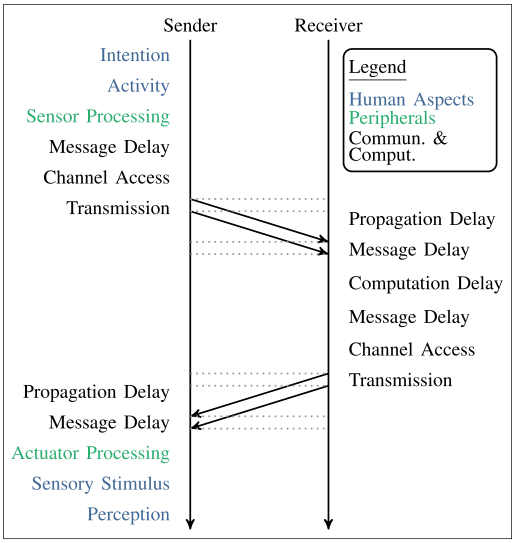

FIGURE 1. Diagram of aspects in human, computation, and networking contributing to latency in the Tactile Internet.

solutions that support ultra-low latencies. Until recently, such solutions have only been of greater interest in the electronic gaming community.

Second, we consider the time needed to send/receive/process individual bits of a message. We consider multiple parts of this delay: channel access delay, transmission delay, and message delay. In shared medium communication systems, such as wireless communication systems, accessing a shared channel is time-consuming (the actual time needed depends on the actual type of access control mechanism employed), as the channel access is divided between multiple users. Increasing the data rate, for example, can decrease this latency component [7]. Additionally, small delays could be added for some initial message pre- or post-processing as message delay.

Limited by physics, the speed of light restricts the time necessary to convey information over a certain distance as propagation delay. In-network computing with software-defined networking (SDN) and network function virtualisation (NFV) enables mobile edge clouds that virtually reduce the distance [8]. The required computation itself introduces an additional delay requiring new concepts. Service function chaining [6] or different forms of computing, such as analog computing, could be employed to reduce computation latency.

The Metaverse and its evolutionary enhancements are based on human interactions and actions within a cyber-physical environment. This human-in-the-loop principle is supported by the Tactile Internet. For example, providing tactile feedback given the constraints of human perceptual systems and their underlying brain mechanisms requires latencies with an upper bound of around 1 ms. The detailed discussion in [6] specifies the delay components with practical numerical boundaries and derives a maximum geographical distance of 25 km between sensing and actuation (see Fig. 2). Moving beyond a radius of about 25 km thus requires new techniques.

The only approach to reduce a physics-limited latency is to derive a new method with a latency component that is less than zero, a *negative latency*. Figure 2 illustrates the idea behind this concept: A negative latency component can increase the geographic reach of the future Metaverse by releasing the propagation delay constraints in the overall system. *Consequentially, a sufficiently large negative latency can extend the geographic reach of the future Metaverse beyond the current MAN boundaries to a global scale.*

The idea of using predictive methods to reduce latency has been discussed in prior works. The survey conducted in [9] highlights the need for intelligent artificial tactile engines that anticipate human actions. Anticipating human actions and intentions opens up the opportunity for the proactive transmission of haptic feedback (e.g., for teleoperation scenarios). The concept of haptic communication and the role of predictive models is also surveyed in [10]. Intelligent engines can analyse, e.g., kinesthetic, user data, and anticipate future network traffic based on recognised hand gestures and actions as provided in [11]. In extension to these prior works, we provide an implementation that showcases the possibilities and challenges of anticipatory computing with guarantees in the future Metaverse.

The Concept of Negative Latency

The term *negative latency* was first introduced by Google in 2019 during an advertisement campaign for the former cloudbased gaming service *Stadia*. In the context of *Stadia*, negative latency refers to the concept of filling a buffer between the server and the player's device with predicted frames of game display for various game continuations. The main purpose of Google's concept of negative latency is to enable cloudbased games to run at faster frame rates, thus providing a smoother gaming experience. Hence, this service can be regarded as an instance of anticipatory computing in which server-side predictions on future frames are made, but without *confidence*.

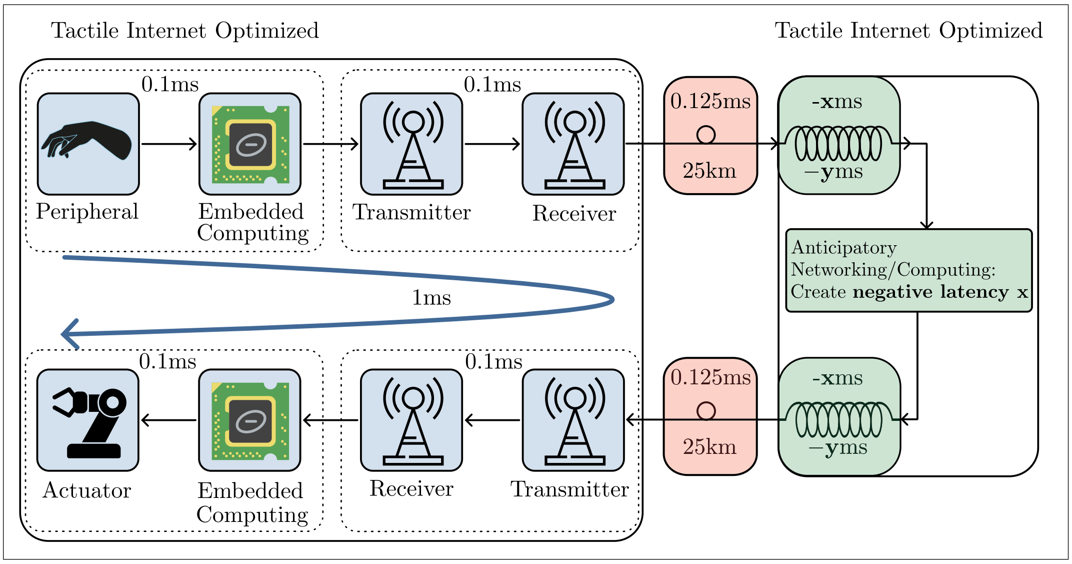

FIGURE 2. End-to-End Latency budget for a sensor-to-actuator loop that meets one millisecond round-trip latency requirement of 5G, recreated and extended from [6].

We argue that this notion of negative latency based on full enumeration is not sufficient in the context of the next generation of the Metaverse.

The future Metaverse must support multimodal interactions in the physical domain, e.g., by controlling robots. This requires reliable low latencies and, therefore, *guarantees* on all the techniques involved to avoid human physical harm. This need for guaranteed low latencies applies to other Tactile Internet applications as well, such as remote surgery, where human lives might depend on functioning systems and their low latency communication. Furthermore, low latencies are not only a key enabler for improving user experiences, but also for ensuring the correct functioning of the experience generation itself.

In general, it is a challenging task to provide guarantees for predictions. Returning to the *Stadia* example, for undercutting 800 ms latency and guaranteeing a 30 fps gaming experience, 24 frames must be computed in advance. Consider a foundational example with only three possible actions that the user can perform, e.g., "left", "straight", and "right". The result would be that more than $2.8 \cdot 10^{11}$ frame sequences are possible, which can hardly be computed.

We propose to define the concept of negative latency as a form of *anticipatory computing* on actions and skills where predictions involved are ensured with sufficiently *high confidence*. Thus, negative latency can be regarded as the addition of quantifiable guarantees to prediction and preemptive computing to undercut latencies.

Quantifying Guarantees on Latency

Combined efforts from different disciplines (see Figure 1 for examples of domains) are required to arrive at guarantees to reduce latency and eventually arrive at negative latency:

1. Neuroscientific studies on perception, cognitive control, and physiological properties contribute to guarantees for human-imposed latency,
2. physical and technological limits such as the speed of light or network performance impose system limits on latency,
3. confidence measures on behaviors and predictions contribute limits to anticipatory computing solution spaces, and
4. models that capture the overall system through formal quantitative analysis contribute guarantees.

The interplays of these factors are important as well, e.g., guarantees on motor behaviours (1) could be used to set boundaries on machine learning predictors (3) and thus improve prediction quality. Combinations of these techniques should result in a confidence measure for predictors that yields, given sensor data, a confidence value between 0 and 1 and an anticipated time of actual execution for each predicted action or skill. Assuming predictors that provide a monotonic increase in confidence over time, we can formalise "high confidence" prediction as exceeding a given confidence threshold, e.g., following industry standard classification of failure probabilities. *Negative latency* subsequently constitutes the time towards an anticipated execution of the action that is predicted with sufficiently high confidence.

k-Negative Latency

High confidence in predictions as a requirement for negative latency reduces the action space that has to be considered but still leaves a significantly large amount of potential outcomes, i.e., multiple actions, possible. One can consider similar body movements required to perform different actions or skills as illustrative example. As long as the number k of actions predictable with high confidence is sufficiently low, an anticipatory computation of consequences on the server and client sides is still feasible and could ensure negative latency by fast local reactions.

We therefore define the confidence in the k actions as the sum of the k most confident actions. The *k-negative latency* of a system is subsequently the maximal time before an action is actually executed, and the k-action confidence is above the fixed confidence threshold. Intuitively, the system can prepare for either of the k most confident actions to be executed in k-negative latency ahead of time.

Negative Latency in Anticipatory Human-Machine Interaction

Peripherals, such as sensors and actuators, represent one of the fundamental building blocks of the evolution of the Metaverse and are required to enable an immersive humanmachine-human or human-machine interaction. As we conceptualised in the previous section, negative latency provides a potential solution to the problem of latency in communication systems, especially when considering long-distance communications. In this section, we highlight how negative latency can be achieved in a foundational example for human-machine interactions.

Human Grasping Actvities

Humans plan their actions. The intention to act, such as picking up an object with the dexterous hand, involves the brain mechanisms that underlie cognitive control and motor planning. Subsequently, these mechanisms guide the implementation of the intended action and convert motor plans into appropriate muscle activities for arm and hand movements. Active inference theories were developed in cognitive neuroscience for predicting human motor planning and goaldirected control, see, e.g., [12], [13]. These approaches can be leveraged for predicting intended actions by humans. Similar predictive inference frameworks [14] and deep neural networks have also been applied in the domain of robotic control (see, e.g., [15]). Together, models and algorithms for anticipating human action and robotic control would allow communication systems to prepare ahead of time, which ultimately could be used to achieve negative latency. We schematically depict the steps of grasping an object in Fig. 3. As the action of grasping an object originates from an intention, muscle activity begins following cognitive planning, and finally, the hand follows the trajectory towards the grasp of an object.

The anticipation of human intention or action before actual muscle activity would require a brain-computer interface, which could be subject to ethical conflicts. Based on the outlined prior research, we postulate that human action can be anticipated with high confidence by interpreting the conducted motion that is sensed via wearable devices and subsequently processed, e.g., with methods from statistics and machine learning. We demonstrate this concept in the remainder of this section with an experimental real-world determination of *k*-negative latency. This represents only an initial foray into determining the capabilities of predictive machine learning algorithms for human action anticipation for a globally connected multisensory Metaverse.

Determination of *k*-Negative Latency

We employ an experiment that focuses on hand movements to showcase how *k*-negative latency (see the section "k-Negative Latency") can be provided. We selected this example as it is one of the most common human actions performed and also prevalent in the domain of human-machine interaction. The *k*-negative latency in our presented approach is derived by anticipating at most *k* possible human hand motion actions with high confidence before the actual human action takes place. We hypothesise that grasping motions and objects excite different patterns of hand and finger motions and that these patterns can subsequently be used to infer the object to be grasped. In turn, a communication system can react before the actual grasp happens. The resulting challenge is to distinguish the motion patterns that occur as soon as possible to provide anticipation of the object to be grasped. We re-emphasise that anticipation inherently comes with confidence (or vice versa with an uncertainty) that needs to be quantified to be considered negative latency. This need especially applies to the Tactile Internet underpinning the Metaverse where high confidences are required to turn anticipation into negative latency and to prevent potential human harm in the cyberphysical space.

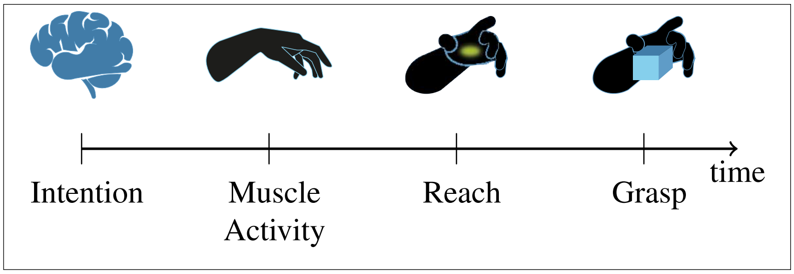

FIGURE 3. Schematic of steps in a human grasping motion.

FIGURE 4. Photo of the experimental setup (light sensor not depicted) during a grasping trial of the espresso cup. The bar plots at the top of the screen represent the classifier's confidence level of the classification decision depicted at the bottom left corner. The finger angle data of the last ten timestamps provided by the smart glove that is passed to the classifier is depicted at the bottom right corner.

Experimental Setup

In our setup, details of the human grasping motion are sensed using a smart glove that measures the orientation of the fingers using inertial sensors at a rate of approximately 20 ms. The smart glove we employ is equipped with six inertial sensors that each produce a four-dimensional number, known as quaternion. A quaternion represents the rotational displacement of the inertial sensor relative to a fixed-earth frame. We measure the elapsed time between the model anticipating the object correctly and the event the object is being lifted via a light sensor that is sampled at the same rate as the smart glove. The response time of the light sensor and the compute time for the machine learning classifier decision require less than 10 ms in total. We consider these values within a reasonable range, especially considering the attainable negative latency (see the section "Results" for details). For the experimental configuration, we use four different objects of daily living (glass, cup, espresso cup, and plate). We depict our overall experimental setup in Fig. 4. In this proof-of-concept demonstration, the kinematic data of hand grasping motions were collected from 3 human subjects, i,e., the first author and two associates.[1]

Machine learning classifiers are trained to detect and interpret patterns based on past observations. For classification purposes, the output of a classifier resembles a vector of possible classification decision and their determined level of likelihood. In our study, we employ a MobileNet machine learning classifier tailored for mobile applications and lowlatency classification tasks. To prevent overfitting and enable real-time deployment on experimental hardware, we reduce the classifier's trainable parameters. Training the model involves recording 50 motion sequences of grasps for four objects by three individuals, splitting the data into 80/20 train and test sets, and augmenting it for fairness through randomization of item placements. We subsequently employ commonly applied methods in the machine learning domain (such as the stochastic gradient descent) to train the classifier to distinguish between the grasping motions of the four different objects based on the data recorded in the train set. We validate the performance of the classifier based on the validation data and fine-tune the parameters of the training procedure until no further improvement in the classification performance is observed.

We further refine the classifier calibration after the training phase by adjusting the classifier's output as follows. The confidence of the classifier is

[1] Exempt of ethical approval in the category benign behavioral assessment with adults, 45 CFR 46104 (d)(3).

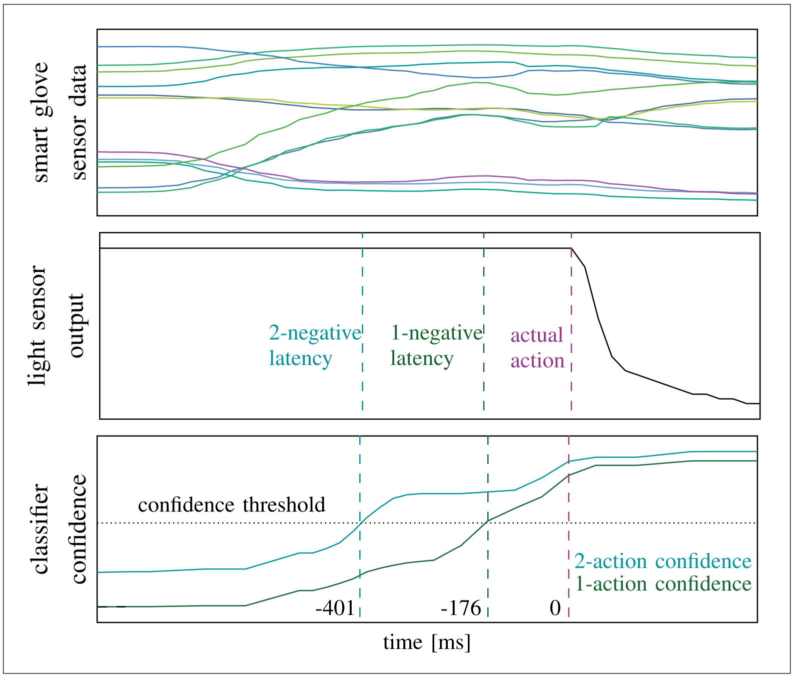

FIGURE 5. Time series data obtained from experimental setup as a user is grasping a glass. Vertical lines correspond to moments in time the confidence threshold is exceeded for 2 or 1 action and the glass is grasped (2-negative latency, 1negative latency, and action finalisation, respectively).

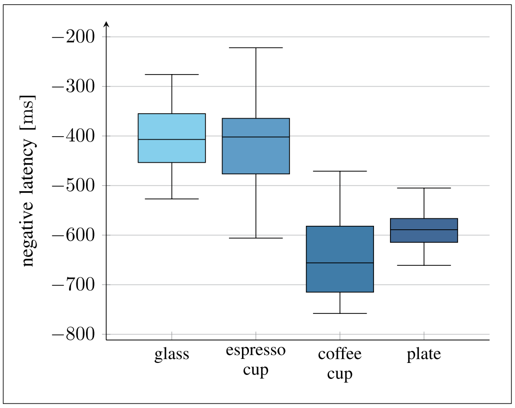

FIGURE 6. Distribution of measured negative latency values of 30 trials per object.

initially determined by calculating the entropy as a measure of uncertainty of the resulting classifier output vector for each individual classification, i.e., actionable objects in our experimental setup. We subsequently normalise this entropy value by the maximally obtainable entropy given the number of actions in our experiment. To this end, we obtain a measure u of uncertainty, ranging from 0 (no uncertainty) to 1 (full uncertainty). We finally define a measure of confidence c for each action as the complement of its uncertainty u, i.e., $c = 1 - u$.

A 1-action confidence corresponds to an individual predicted action's confidence measure. Determination of a k-action confidence measure is based on the accumulation of the confidences of the k most likely predicted actions. A classification decision is accepted if the associated confidence surpasses a threshold. This threshold is empirically determined and chosen to minimise the risk of misprediction and early recognition to achieve a high negative latency budget. In our experiments, we fixed a confidence threshold of 0.5. We highlight that establishing the confidence threshold based on entropy is distinct from setting a threshold on the likelihood estimate produced by the classifier. In the context of a binary classification issue, an entropy-based confidence threshold of 0.5 means that the likelihood estimates would be assigned to the two individual classes in a 0.89 to 0.11 manner.

Results

We illustrate the outcomes for a single experiment trial of grasping a glass in Fig. 5 as an example. The top plot shows data gathered during a grasp from the smart glove, whereby each of the six inertial sensors built into the smart glove produces a four-valued orientation estimate as quaternion. (We note that we are focused on a qualitative illustration of those values here to provide additional context.) The second plot depicts the light sensor measurement that changes upon the moment the smart glove comes in physical contact with the object. We set the time of this initial sensor change to zero. The third plot in Fig. 5 illustrates the model's determined confidence level as part of its classification decision with respect to the most (glass, 1-negative latency) and the two most (glass and espresso cup, 2-negative latency) likely objects to be grasped.

We initially observe that the confidence of the model gradually increases as the grasping motion progresses. This observation can be explained that with the grasping motion advancing, more distinct hand shapes towards the final grasp are exhibited as corroborated by the angular information changes illustrated in the top plot. Fig. 5 also illustrates the 1- and 2-negative latencies determined for our predefined confidence threshold of 0.5. With the time set to zero for the light sensor's indication of the actual grasp, the model's 50% classification confidence for two objects yields a 2-negative latency of –401 ms and for one object a 1-negative latency of –176 ms. Differently put, our approach shows that with confidence requirements on predictions above a quality threshold, anticipation can turn into negative latency that enables a local system to act a significant amount of time before the actual event. Note that in the example trial run, the system has twice the negative latency time to prepare for either glass or espresso cup grasping, before grasping the glass could be predicted with confidence above the quality threshold. We additionally note that higher confidence levels would reduce the negative latency generated.

In evaluating the experiment's performance across various grasping tasks, we conducted 10 trials for each object and individual, presenting the resulting 1-negative latency ahead of actual event times in Fig. 6. Notably, our approach exhibits a discernible range for the negative latency achieved. Specifically, grabbing the coffee cup or plate tends to yield higher negative latency values compared to reaching for a glass or espresso cup. This indicates more distinct movement patterns for these two groups of objects. The larger spread of negative latencies observed with the espresso cup can be attributed to variations in grasp patterns between individuals. Despite limitations in the number of objects and trials, these preliminary findings suggest that anticipatory systems can achieve negative latencies of several hundred milliseconds with high confidence.

Considering the need for *negative latency* in the Tactile Internet to enable the evolution of the Metaverse on a global scale, one could briefly return to the need to provide communication beyond a 25 km radius for a human-in-the-loop system. The worst and best cases observed for the trials illustrated in Fig. 6 are −222 ms for grasping the espresso cup and −758 ms for grasping the coffee cup, respectively. In a bestcase scenario, this would allow tactile information to travel more than five times around the world. More importantly, however, even in the worst case the information could travel more than once around the world – more than enough to enable the seamless immersion and interaction requirements of the Metaverse's next generation on a global scale.

Conclusion and Future Work

The vision of the Metaverse's evolution we presented in this article has tremendous potential to enable applications that pave the way to advance society. With the Tactile Internet as an enabler of immersive human-machine-human or human-machine communication, we propose negative latency as a solution to the latency challenge that limits the future Metaverse's possibilities. A preliminary study investigating the potential for negative latency demonstrates that negative latency of several hundred milliseconds is possible and would provide the necessary geographic and multisensory reach of the next generation Metaverse.

We foresee significant further research needs to evolve the Metaverse. As anticipatory systems rely on statistical methods and artificial intelligence, investigating different frameworks for accuracy and suitability of deployment on Tactile Internet hardware is of significant importance. Moreover, the interplay between anticipation time horizons and attainable confidence levels is crucial to ensure prediction confidence and, subsequently, negative latency across the plethora of application scenarios of the next-generation Metaverse. Methods capable of providing a reliable estimate of their anticipatory confidence will therefore have a tremendous impact. These required advances for the next-generation Metaverse can barely be investigated by communication engineers alone, but have to further take into account physiologic, neuroscientific, physical, and model-based computing designs with quantitative analyses.

Acknowledgment

This work was supported in part by the German Research Foundation (DFG) through the Germany's Excellence Strategy–EXC 2050/1–Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI), Technische Universität Dresden, under Project 390696704; in part by the Federal Ministry of Education and Research of Germany (BMBF) through the programme of "Souverän. Digital. Vernetzt." Joint Project 6G-life, under Project 16KISK001K; and in part by BMBF and German Academic Exchange Service (DAAD), School of Embedded and Composite AI (SECAI), under Project 57616814.

References

[1] F. H. P. Fitzek et al., *Tactile Internet With Human-in-the-Loop*. New York, NY, USA: Academic, 2021. [Online]. Available: https://ceti.one/book
[2] M. Caeiro-Rodrı́guez et al., "A systematic review of commercial smart gloves: Current status and applications," *Sensors*, vol. 21, no. 8, p. 2667, Apr. 2021.
[3] E. Muschter et al., "Perceptual quality assessment of compressed vibrotactile signals through comparative judgment," *IEEE Trans. Haptics*, vol. 14, no. 2, pp. 291–296, Apr. 2021.
[4] S.-C. Li et al., "Human perception across the lifespan as the basis of a tactile Internet for everyone," in *Proc. 27th Eur. Wireless Conf.*, Sep. 2022, pp. 1–6.
[5] S. Senk et al., "Healing hands: The tactile Internet in future tele-healthcare," *Sensors*, vol. 22, no. 4, p. 1404, 2022.
[6] Z. Xiang et al., "Reducing latency in virtual machines: Enabling tactile Internet for human-machine co-working," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1098–1116, May 2019.
[7] P. Schwenteck et al., "6G perspective of mobile network operators, manufacturers, and verticals," *IEEE Netw. Lett.*, vol. 5, no. 3, pp. 169–172, Sep. 2023.
[8] F. Fitzek, F. Granelli, and P. Seeling, *Computing in Communication Networks: From Theory to Practice*. New York, NY, USA: Academic, 2020.
[9] N. Promwongsa et al., "A comprehensive survey of the tactile Internet: State-of-the-art and research directions," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 472–523, 1st Quart., 2020.
[10] X. Shen et al., "Toward immersive communications in 6G," *Frontiers Comput. Sci.*, vol. 4, Jan. 2023, Art. no. 1068478.
[11] K. Antonakoglou et al., "Toward haptic communications over the 5G tactile Internet," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3034–3059, 4th Quart., 2018.
[12] R. A. Adams, S. Shipp, and K. J. Friston, "Predictions not commands: Active inference in the motor system," *Brain Struct. Function*, vol. 218, no. 3, pp. 611–643, May 2013.
[13] G. Pezzulo, F. Rigoli, and K. J. Friston, "Hierarchical active inference: A theory of motivated control," *Trends Cognit. Sci.*, vol. 22, no. 4, pp. 294–306, Apr. 2018.
[14] L. Pio-Lopez et al., "Active inference and robot control: A case study," *J. Roy. Soc. Interface*, vol. 13, no. 122, Sep. 2016, Art. no. 20160616.
[15] A. Ciria et al., "Predictive processing in cognitive robotics: A review," *Neural Comput.*, vol. 33, no. 5, pp. 1402–1432, Apr. 2021.

Biographies

Jonas Schulz (jonas.schulz2@tu-dresden.de) received the Diploma (Dipl.-Ing.) degree in electrical engineering from Technische Universität Dresden, Germany, in 2022. He is currently pursuing the Ph.D. degree with the Deutsche Telekom Chair of Communication Networks, Technische Universität Dresden.

Clemens Dubslaff (c.dubslaff@tue.nl) received the M.Sc. degree in computational logic from UN Lisbon, Portugal, and the Ph.D. from Technische Universität Dresden, Germany. He is currently an Assistant Professor with the Formal System Analysis Group, Eindhoven University of Technology, The Netherlands.

Patrick Seeling (Senior Member, IEEE) (patrick.seeling@cmich.edu) received the Diploma (Dipl.-Ing.) degree in industrial engineering and management from Technische Universität Berlin, Germany, in 2002, and the Ph.D. degree in electrical engineering from Arizona State University, USA, in 2005. He is currently a Professor with the Department of Computer Science, Central Michigan University, USA.

Shu-Chen Li (shu-chen.li@tu-dresden.de) received the M.Sc. and Ph.D. degrees in cognitive psychology from the University of Oklahoma, USA, in 1991 and 1994, respectively. She is currently a Professor of lifespan developmental neuroscience with Technische Universität Dresden, Germany, and the Principle Investigator with CeTI and 6G-life.

Stefanie Speidel (Senior Member, IEEE) (stefanie.speidel@nct-dresden.de) received the Diploma degree in computer science from Universität Karlsruhe in 2005 and the Ph.D. (Dr.-Ing.) degree from the Karlsruhe Institute of Technology in 2009. She is currently a Professor with Technische Universität Dresden, the Director of the National Center for Tumor Diseases (NCT/UCC) Dresden, Germany, and the Head of the Division Translational Surgical Oncology (NCT/UCC).

Frank H. P. Fitzek (Senior Member, IEEE) (frank.fitzek@tu-dresden.de) received the Diploma (Dipl.-Ing.) degree from RWTH Aachen, Germany, in 1997, and the Ph.D. (Dr.-Ing.) degree from Technical University Berlin, Germany, in 2002. He is currently a Professor with Technische Universität Dresden, Germany, coordinating the 5G Lab Germany, 6G-life, and CeTI.