Reviews of 2151 - *"HT-Auth: Unobtrusive VR Headset Authentication via Subtle Head Tremors"*

Reviewer 2 (1AE)

(Review of original submission) 1AE's recommendation to Editors

Major revisions (enumerated in a subsequent field) are required for this to be publishable

(Review of original submission) Impact

Significant impact

(Review of original submission) 1AE's meta-review

The paper introduces HT-Auth, a novel authentication system leveraging subtle head tremors captured by inertial sensors in VR headsets. All reviewers are positive about this paper. The concept is innovative, addressing an important challenge in unobtrusive user authentication for VR environments, and provides a technically sound system design along with detailed evaluations. While the authors present a well-explored first work on this topic, we decided on major revision as the paper requires substantial revisions before publication.
The main concerns lie in the evaluation. The accuracy metrics need reconsideration and additional experiments are required to address consecutive authorization attempts (R2). The localization and noise reduction components of the pipeline lacks sufficient analysis (R4) and specific explanations (R2). The hyper-parameters used in the pipeline are not explained and their impact on performance has not been studied. The current evaluation setting is also limited, not including more realistic VR scenarios, such as dynamic full body motion, diverse VR device characteristics, and the duration of VR sessions. Additionally, the absence of benchmarks against other VR authentication methods makes it challenging to assess the advantages of the proposed system.
Direct comparisons with existing methods are necessary to provide a clearer understanding of its benefits.

The writing and presentation also require significant updates. First, the paper should better frame the proposed authentication technique. The introduction describes the

system as unobtrusive and based on “natural” movements, yet the implementation involves deliberate head tremor actions during authentication. This discrepancy should be clarified by revising both the introduction and discussion sections to avoid misleading readers (R3, R4). Moreover, the paper should include a more comprehensive threat model and address sophisticated attack scenarios (R1, R4). The privacy implications of collecting, storing, and processing head tremor data must also be discussed. Additionally, the paper should explore user perceptions of biometric data security and propose mechanisms to protect sensitive information (R4).

(Review of original submission) 1AE's meta-review: Major / minor revisions

- Improve evaluations: reanalyze the accuracy metrics (R2), explanation of the hyper-parameters and their effect on the performance (R1, R3), additional evaluation of the localization and noise reduction portion of the pipeline (R4), comparison to existing authentication methods (R4).

- Address concerns on the diversity of the evaluation settings with additional experiments or in-depth discussion, and avoid overclaiming contributions (R4)

- Clarify user participation requirements: Clearly state that the head tremors must be actively performed by the user rather than being entirely natural and passive. Update framing and introduction to align expectations. (R3, R4)

- Attack scenarios: Consider more sophisticated attack scenarios such as injection attacks (R3, R4), elaborate on tamper-proof hardware and injection points in the security analysis (R1)

- Improve writing: address unclear sections (R1), provide more discussions on the practical usage scenarios (R4), provide discussions on the privacy implications (R4)

Reviewer 1 (2AE)

(Review of original submission) Contribution to IMWUT

XR (extended reality = VR ... AR) headsets hold big hopes to become "the next big (interactiv wearable/ubiquitous) thing" after the Smtartphone (2007). User input to XR worlds requires intensive further research since mid-air gestures and controllers, both prominent today, are inadequate for many tasks. Authentication - which can be viewed as a special category of input - must receive special attention.

In view of this situation, the idea to use a yet-unexplored biometric input for headset-based authentication represents a very interesting topic for IMWUT.

(Review of original submission) Impact

Medium impact

Ethics

Yes, the paper is in full compliance with IMWUT's policy on Ethics

(Review of original submission) Review

The "contribution" section of this review highlighted the theme and significance of the paper. more concretely, the authors propose involuntary *head tremor* as a unique personal "trait" o be exploited for implicit authentication, meaning that the authentication could run in the background (similar to gait or typing-characteristics based auth.), neither restricting it to a "login" process at the strta of a session nor bothering he user with the interruptions of re-authentication, e.g., after a pause in using an App.

The paper is well written, the authors make a substantial effort to present arguments and information in an understandable / illustrative way. Obviously, the challenge is to convert head movement (sensor) data into a robust representation: this representation must ensure a high likelihood that the comparison of the true user's actual and initially-registered samples match with high likelihood ("accept") while any other person's actual samples don't match ("reject") with high likelihood.

The processing steps, measures and underlying ideas for arriving at such a suitable representation make the core scientific novelty and originality of the paper, the corresponding experimentation (user study) contributes to that scientific "value". Already in section 2.5, the authors show good mastery of typical approaches used as part of biometric authentication; the authors argue well why the wide-spread methods discussed are not suitable for their problem, which also shows that the problem at hand is far from trivial.

In assessment, it seems clear to this reviewer that the presented approach represents a valuable contribution, opens the door for investigation of a yet-unexplored biometric authentication scheme, and presents a quite good first approach to it.

A number of issues call for improvement. Some can be easily resolved by improving the writing. Remaining issues that can't be healed in the context of a minor revision represent drawbacks of a first approach to a novel implicit-authentication scheme; this reviewer finds it totally acceptable that an initial approach is not yet perfect; the open issues should be better highlighted in the writing however.

For the list of issues to be adressed in the revision (either by clarification or by

admission as drawbacks), see the recommendations below

In summary, the recommendation is "accept with minor changes".

(Review of original submission) Recommendation(s) to 1AE

Acceptable with minor (or no) changes

(Review of original submission) Major / minor revisions (recommendation to 1AE)

suboptimal introductory explanation Line 32: the one-dimensional binary categorization of biometrics-based authentication schemes (behavioral vs. physiological) may be somewhat wide spread but is not helpful for non-expert readers as an introduction to the general authorization challenge addressed in the paper. A two-dimensional distinction regarding obtrusiveness (explicit / implicit i.e. foreground / background) on one hand and hardware cost (maybe with three values: zero for hardware common for the devices consiered - here: IMUs in XR headsets, low for techniques like fingerprints where cheap hardware existis meanwhile, high for hardware that is still expensive like EEG/EMG etc.) would be more helpful.

unclear writing or error: Lines 224-229: Something seems to be messed up here and in the next section, or the text is no clear. Signals with heavy PSD are discarded, but for those kept, the interesting signal (0-15Hz) and interfering signals with 0-20Hz will still pose a big problem - which seems to be the reason for applying MODWT. It is neither clear from the present paragraph why denoising can't be applied for heavy-PSD frames nor is it made clear that the remaining frames may contain the same kind of noise (human-motion artifacts) ... at least that's the impression of this reviewer from looking at Fig. 8 and the corresponding text, see below

unclear writing Line 230: this first sentence speaks of micro-movements as the remaining noise, but then, MODWT is shown to remove human motion, see Fig. 8. This confusion blends with the missing clearification addressed above.

improtant issue / missing details Line 230ff (3.2.2): It is hard to understand (if at all) how the process actually works. A little more information on the parameters of G_i would be helpful (mother wavelet? Other parameters?) It should become cleaer how, in the process, one "spectrogram" (head tremor) can identified as signal and all others as noise.

marginal issue Line 349ff (section 3.4.2): The theme adressed here provokes questions about the need to cater for variations in short-term (day-to-day) muscle "behavior" and long-term issues (evolution of muscles). This issue is nicely addressed much later in the paper. A short forward reference in brackets would help the interested reader

typo Line 370: "to" missing between "basically" and "utilize"

related issue Lines 444-445: the chosen parameter settings don't help to understand the details about the MODWT process requested above Readers cannot be expected to read the documentation of the pywt package (nor a backgroud paper on MODWT; basic undertanding of wavelets can be assumed though)

better explanation needed Line 502: BAC is meant for balancing, but is that the right approach? Fingerprint authentication is now common and can hence serve as an analogy for the issue I am trying to highlight here: FRR is much less dramatic than FAR, especially if 1-2 immediately-following attempts succeed with high likelihood and/or if a (usually more tedious) fallback authentication technique is available. Therefore, the use of BAC should be put in perspective.

missing study results Line 502ff: Continuing with the issue raised above, FRR/FAR should be compared w.r.t. to this higher danger of FAR. Ideally, consecutive attempts (within a few seconds) should be measured and discussed: measuring and discussing multi-attempt authenticatio nmay not be possible in the timeframe of a minor-revision cycle, but maybe at least a few informal experiments could be conducted and the issue should be clrearly identified. This is even more important since the proposed approach lends itself to background authentication, which consists "naturally" of regular repetitions of the authentication process.

same issue Line 634: This section would be a good place to elaborate further on the issue of consecutive authorization attempts (with the true user or a different / attacking one wearing the headset)

Unclear issue Line 652ff (5.3.2): This paragraph is unclear, it should be described more precisely. The issue is also related to the entire issue of how tamper-proof hardare. Where would an attacker have to intercept the data into a corresponsponding authentication system and how difficult would that be (would the injection point be easily accessible? How difficult would it possibly be to construct more temper-resistent hardware?).

Related issue Line 652ff (5.3.2): related to the issue above - and maybe worth discussing much earlier in the paper to avoid misunderstandings - is the question if one can willingly generate head tremor? If so, how - and would that lead to stronger "pseudo-tremor" signals than the unwillingly effectuated ones (with higher likelihood of leading to FA than 5.3.1 and 5.3.2)?

(Review of original submission) Confidence

Very confident - I am knowledgeable in the area

Reviewer 3 (Reviewer)

(Review of original submission) Contribution to IMWUT

This paper addresses the challenge of secure and efficient user authentication in Virtual Reality (VR) environments, where traditional methods like passwords or hand gestures are inconvenient and prone to security risks. The authors propose HT-Auth, a novel authentication system leveraging subtle head tremors, which are natural and unique to individuals, captured via inertial sensors in VR headsets. By deriving biometric data from these tremors, the method offers an unobtrusive and reliable authentication solution. Experimental results demonstrate high accuracy with minimal registration samples.

(Review of original submission) Impact

Medium impact

Ethics

Yes, the paper is in full compliance with IMWUT's policy on Ethics

(Review of original submission) Review

Strengths:

The insights of the biomechanics underlying head tremor production are appreciable.
This paper considers user heterogeneity and employs transfer learning to enhance personalization.
The user study validates the practicality and usability of the proposed method.
The open-source code contributes to the community and promotes innovation.

Concerns:

In HT-Auth, Head tremors are captured by zero-permission IMU sensors. An attacker can very easily acquire these sensory inputs. Based on this knowledge, they may launch even more powerful attacks, e.g., injection attacks.
The basic authentication process from the user's perspective requires further clarification. For instance, it should be explained whether the head tremors are user-initiated actions or if the system employs a challenge signal to trigger the authentication process.
How are the hyperparameter values in the design determined? Are they selected through empirical experimentation, or is there a specific rationale behind choosing

these values?

(Review of original submission) Recommendation(s) to 1AE

Major revisions (enumerated in a subsequent field) are required for this to be publishable

(Review of original submission) Major / minor revisions (recommendation to 1AE)

The length of this paper is acceptable.

(Review of original submission) Confidence

Highly confident - I consider myself an expert in the area

Reviewer 4 (Reviewer)

(Review of original submission) Contribution to IMWUT

This work introduces innovative approach to authenticate users on VR devices, using built-in IMUs to capture unique muscular characteristics of head tremors as biometric signatures. The contribution particularly addresses the growing needs for unobtrusive and practical authentication methods in VR environments. This can not only be used in traditional HMDs but also applied to various types of glasses or head-worn devices that's likely to have simple IMU sensors.

(Review of original submission) Impact

Significant impact

Ethics

Yes, the paper is in full compliance with IMWUT's policy on Ethics

(Review of original submission) Review

This work presents a novel user authentication system leveraging natural head tremors as biometric signatures. The technical approach involves three main components: noise reduction, event detection, and feature extraction. For noise reduction, the system filters out motion artifacts and noise (walking, breathing, etc) using PSD analysis, employs GA-based MODWT denoising. The event detection leverages three-stage sliding window approach (with thresholds) to precisely localize each tremor event. Then, the feature extraction captures both muscular contraction features through frequency response and muscular endurance characteristics through temporal consistency, and uses Siamese network for feature reconstruction to handle behavioral inconsistencies. Their evaluations with 30 subjects (10 for initial

and 20 for evaluation) demonstrate robust and consistent performance across both standalone and mobile VR headsets. Also, it showed resistance to blind and impersonation attacks, while maintaining reasonable computational delays.

Strengths:

I think this is an intriguing idea and one that certainly merits some attention and exploration. Head tremors seem like unique biometric signature that could be captured by readily available IMUs, making this useful solution for VR security.

-- The proposed idea of using head tremor looks novel and highly effective.
-- Evaluation result looks promising (97.22% BAC) and thorough, including user studies that indicates strong acceptance of the system's usability.
-- The detailed steps of the overall protocol is well organized, and are well explained including reasonable justification for choosing each signal processing algorithms.
-- The length of the manuscripts looks reasonable.

Weaknesses:

I am not fully convinced that the system, as currently implemented is particularly secure and usable as the author suggests. I lay my concerns below, and hope the authors can address them in the future revisions.

--I was quite interested about the overall idea when I read the Section 1 and 2. The authors emphasize "natural" and "unobtrusive" aspect of the head tremor which can be inevitably detected. There appears to be a disconnect between this framing and the actual implementation. In the evaluation section (4.2), the subjects were told to deliberately perform the head tremor rather than leveraging naturally occurring movements. The authors should clearly specify that the head tremor must be actively performed by the user. The authors should clearly state in the introduction that their system requires active user participation --users must consciously generate head tremors to authenticate.

--The paper's noise reduction approach in Section 3.2, using sliding window-based PSD to differentiate between human motion and head tremors, requires further validation in realistic VR usage scenarios. The current evaluation only examines different head postures in relatively stationary conditions. However, VR applications often involve dynamic movements like walking, exercise, or interactive gameplay and it's important to assess how effectively the noise reduction and localization portion of

the protocol performs under these more challenging conditions. The authors should expand their evaluation to include common VR activities that involve full body motion to demonstrate whether the system can maintain its claimed performance (reducing noise and localizing head tremor).

--The evaluation of consistency across four weeks is valuable, but how does authentication process vary during extended VR sessions? For instance, when users wear HMDs for prolonged periods (over 20~30 minutes), their neck muscles naturally adapt to the headset's weight and positioning. This physiological adaptation can alter the characteristics of head tremors. The authors should investigate how their system performs at different points during longer VR sessions. For example, comparing success rates at the beginning, middle, and end of a 1 hour session.This would provide important insights into the system's reliability during typical VR usage patterns

--The current evaluation focuses only on single-device authentication. The authors should investigate whether head tremor biometric templates can be effectively transferred between different VR devices. Factors like sensor specs, headset weight, ergonomics, and strap designs may affect device-head coupling. Understanding this when enrolling on one device and authenticating on another would provide practical insights in multi-device environments.

--The security analysis needs to consider more sophisticated scenarios beyond blind and impersonation attack. Particularly concerning is video-based injection attacks: an attacker could record the victim's head tremors using camera, then use video processing techniques to reconstruct the corresponding IMU signals. Because the threat model already assumes physical access to the device, injecting such reconstructed signals is very feasible. The authors should evaluate the system's resilience and consider implementing appropriate countermeasures which provides more realistic security evaluation.

--The paper is missing a critical discussion of privacy implications and user perceptions regarding using physiological data for authentication. The user study addresses usability aspects like comfort and ease of use, but it overlooks important privacy considerations. Users may have concerns about their head movement patterns being captured, stored, and processed as biometric data because these could potentially reveal information about their physical condition.

--The paper needs more clear comparison against other VR authentication methods. While achieving 97% BAC is promising, without explicitly benchmarking against other biometric approaches, it's hard to understand the real advantages.

(Review of original submission) Recommendation(s) to 1AE

Major revisions (enumerated in a subsequent field) are required for this to be publishable

(Review of original submission) Major / minor revisions (recommendation to 1AE)

The authors should make several changes including: 1) revise the framing to clearly state that head tremors must be performed, 2) consider more realistic and sophisticated attack scenarios, 3) address further practical considerations such as privacy implications and cross-device scenarios and 4) evaluation of localization and noise reduction portion of the pipeline. The length is reasonable at the current state.

(Review of original submission) Confidence

Very confident - I am knowledgeable in the area

Return to submission and reviews