Textile & Leather Review

ISSN 2623-6281 | www.tlr-journal.com | 10.31881/TLR

Augmented Reality-Based Personalized Calibration for Color Vision Deficiency: A Framework for Enhanced Human-Robot Visual Perception

Xinghong Hu, Wuyao Shen, Elaine Yuen Ying To, Shaoying Tan, Taizhi Wang, Yue Guo, Zhenyu Xiao

How to cite: Hu X, Shen W, To E, Tan S, Wang T, Guo Y, Xiao Z. Augmented Reality-Based Personalized Calibration for Color Vision Deficiency: A Framework for Enhanced Human-Robot Visual Perception. Textile & Leather Review. 2026; 9:2466-2483. https://doi.org/10.31881/TLR.2026.2466

How to link: https://doi.org/10.31881/TLR.2026.2466

Published: 25 April 2026

Augmented Reality-Based Personalized Calibration for Color Vision Deficiency: A Framework for Enhanced Human-Robot Visual Perception

Xinghong Hu[1], Wuyao Shen[2], Elaine Yuen Ying To[3,4], Shaoying Tan[3,4,5], Taizhi Wang[2], Yue Guo[6], Zhenyu Xiao[6]*

[1]Shenzhen Polytechnic University, Shenzhen, Guangdong, 518055, China
[2]BINOVIZ LIMITED, Hong Kong SAR, 999077, China
[3]School of Optometry, The Hong Kong Polytechnic University, Hong Kong SAR, 999077, China
[4]Research Center for SHARP Vision (RCSV), The Hong Kong Polytechnic University, Hong Kong SAR, 999077, China
[5]Center for Eye and Vision Research (CEVR), 17W Hong Kong Science Park, Hong Kong SAR, 999077, China
[6]Tulingjishi (ShenZhen) Technology Company Limited, Shenzhen, Guangdong, 518083, China
*zhenyuxiao@outlook.com

Ar icle
https://doi.org/10.31881/TLR.2026.2466
Published 25 April 2026

ABSTRACT

This paper introduces an interactive calibration system for color vision assessment utilizing augmented reality (AR) glasses, and investigates its potential relevance as a perceptual calibration stage for human–robot systems. The system re-imagines the traditional Farnsworth-Munsell 100 Hue (FM 100 Hue) test —a benchmark in the textile industry for evaluating the color discrimination of quality control personnel— as an interactive AR-based task, enabling portable color discrimination assessment under ordinary ambient lighting conditions. We quantitatively evaluate the correlation between this AR-implemented assessment and the standard physical test through a comparative user study. Results demonstrate a high degree of diagnostic consistency, indicating that the AR platform can reproduce key characteristics of the conventional assessment while reducing part of the variability associated with display and ambient conditions. Beyond clinical assessment, this work positions AR glasses as a bidirectional perceptual interface, enabling individualized modeling of human color perception for robotic systems. The proposed method facilitates robust, color-aware perception, decision-making, and task execution in human-centered robotics, with potential applications in automated textile quality assurance and assistive robotics. This calibration framework is a promising step toward achieving seamless and adaptive visual perception alignment between humans and humanoid or other robotic platforms.

KEYWORDS

color vision deficiency, augmented reality, human-robot interaction, fm 100 hue test, textile color discrimination

INTRODUCTION

Color vision deficiency (CVD), commonly referred to as color blindness, affects approximately 5% of the global population and represents one of the most prevalent inherited visual anomalies [1]. Individuals with CVD experience altered hue perception and reduced discrimination along specific color confusion axes, which can lead to errors in tasks that rely heavily on color cues, such as interpreting visual indicators, recognizing objects, or responding to warning signals [2,3]. In textile manufacturing and quality inspection, accurate color perception is essential because slight color differences can affect the acceptability, consistency, and commercial value of dyed fabrics and finished products. Traditional shade assessment and manual color matching often depend on trained human inspectors, making color vision assessment an important component of industrial quality control workflows. While CVD has been extensively studied in clinical and vision science contexts, its implications for intelligent systems and robotic applications have received comparatively limited attention.

In robotic and autonomous systems, color perception plays a critical role in perception, decision-making, and interaction. Modern robots rely on color information for object detection, semantic scene understanding, status signaling, and safety-related communication. In many practical deployments—including teleoperation, shared autonomy, and collaborative robotics—humans remain integral components of the control loop. Human operators frequently rely on color-coded visual feedback provided by robotic systems, such as AR overlays, graphical user interfaces, or head-mounted displays. When an operator exhibits atypical color perception, a perceptual mismatch may arise between the robot's sensor-based representation of the environment and the human's interpretation of color-coded information. Such mismatches can reduce task efficiency, increase cognitive load, and, in safety-critical scenarios, lead to operational risks.

Despite rapid advances in robotic perception and vision algorithms, most systems implicitly assume a nominal human observer with normal color vision. Individual variability in human color perception—particularly due to CVD—is rarely modeled or calibrated in human–robot systems. This oversight becomes increasingly problematic as augmented reality (AR) technologies are integrated into robotic interfaces. AR glasses are now widely explored as human–robot interfaces for teleoperation, collaborative robotics, and assistive applications, where virtual color cues are superimposed onto the real world to guide human decision-making. Without individualized perceptual calibration, these AR-mediated cues may fail to convey intended information accurately to all users.

From a clinical and perceptual standpoint, the Farnsworth–Munsell 100 Hue (FM 100 Hue) test is widely regarded as a gold-standard method for assessing fine-grained color discrimination ability. This test is a critical benchmark in textile laboratories for certifying the visual performance of human color matchers. The test evaluates a subject's ability to order colored caps along a continuous hue circle and yields quantitative error scores that characterize both the severity and type of color vision deficiency. However, conventional FM 100 Hue testing requires controlled lighting conditions, physical test materials, and expert supervision, which limit its accessibility and scalability beyond clinical environments. These constraints make it impractical for integration into robotic systems that operate in diverse, real-world settings.

Augmented reality provides a promising pathway to bridge this gap. By presenting calibrated virtual color stimuli through head-mounted displays, AR systems can deliver standardized perceptual tests while maintaining flexibility, portability, and robustness to environmental variations. More importantly, AR enables the seamless integration of perceptual assessment into wearable human–robot interfaces. This opens the possibility of treating color vision calibration not merely as a clinical diagnostic procedure, but as a foundational component of perception alignment in human–robot systems.

In this work, we propose an interactive AR-based implementation of the FM 100 Hue test designed for wearable AR glasses. The system transforms the classical hue-sorting task into a gamified AR experience that can be self-administered by users in non-clinical environments. By validating the AR-based test against the physical FM 100 Hue test through a within-subject study, we demonstrate that the AR implementation can reproduce key characteristics of the conventional assessment. Beyond validation, we frame the proposed system as a perceptual calibration module for human–robot interaction.

The contributions of this paper are fourfold. First, we design a wearable AR-based version of the FM 100 Hue test that preserves diagnostic fidelity while improving accessibility and user engagement. Second, we establish a physical-to-virtual color calibration pipeline that aligns AR-rendered stimuli with standardized color references, reducing environmental and device-induced variability. Third, we experimentally validate the consistency between AR-based and traditional FM 100 Hue testing through quantitative analysis. Finally, we demonstrate how AR-based color vision calibration can serve as a foundational step toward perception-aware human–robot interfaces, supporting adaptive visualization and robust color-dependent interaction in robotic systems.

RELATED WORK

Modeling Color Vision Deficiency

Accurate modeling of CVD perception is essential for personalized color assessment and compensation. Existing approaches generally focus on identifying the type and severity of CVD and then representing the resulting perceptual distortion through computational color-transformation models.

Typically, assessment methods for CVD can be divided into three main types: general tests, computer-based tests, and genetic tests. General tests include traditional pseudoisochromatic plate tests [4] and hue tests represented by methods such as FM100-Hue [5,6]. In hue tests, subjects need to sort multiple groups of color chips to form a seamless color gradient. By analyzing the error rate of sorting, the subject's color discrimination ability and possible color vision defects can be evaluated. The Cambridge Color Test [7] requires subjects to observe a 'C' letter surrounded by a gray background on a display, and the chroma of the letter is continuously adjusted until the subject can just no longer detect the opening of the 'C'. Although the Cambridge Color Test provides quantitative results for the subject's color discrimination ability, it cannot distinguish between color blindness and severe color weakness. Genetic tests are usually costly and require specialized equipment [8].

H. Brettel et al. [9] laid the foundation for modeling CVD perception through projection transformations in the LMS color space. Existing methods for simulating CVD vision are mostly based on the color mapping model proposed by Machado et al. [10], where the colors observed by individuals with CVD can be obtained through linear transformation of the actual colors. Xu et al. [8] used different transformation parameters to map color chips that are easily confused by CVD into the color vision deficiency space, establishing the relationship between parameters and test results by collecting color chip arrangement test results from individuals with normal color vision. Shen et al. [11] presented subjects with two color patches side by side, allowing them to adjust one until the color difference becomes indistinguishable, to obtain confusing color pairs and solve for the transformation parameters.

However, these methods predominantly rely on controlled laboratory setups or constrained display conditions, limiting both ecological validity and practical deployment in real-world contexts. To address these deficiencies, this paper aims to develop an interactive calibration system leveraging AR technology to overcome the limitations of existing approaches. By transforming traditional color vision tests into gamified AR experiences, our method enables personalized assessment within an immersive environment, which not only preserves diagnostic precision but also bridges the gap between laboratory testing and natural visual experiences, paving the way for adaptive color compensation tailored to each individual's perception.

Color Vision Testing — Digital/Clinical Validation

Once perceptual differences are modeled, the corresponding assessment method must also be validated. The evolution of color vision testing [12] has progressed from traditional clinical methods, which are cumbersome and equipment-dependent [13], to electronic tests that are influenced by environmental lighting [14].

Based on interviews with 23 colorblind individuals, Enrico et al. [15] validate the reliability of the video system in helping users distinguish colors they typically cannot perceive. S. Ghose et al. [16] validated the consistency between computer-based FM100 and manual tests. A. Fanlo-Zarazaga et al. [17] found that the DIVE Color Test showed excellent agreement with the Ishihara test and strong correlation with the FM100 Hue test, offering faster, more accurate, and repeatable detection of red-green color vision deficiency. R. Trukša et al. [18] indicated that the FM100 hue test effectively identifies moderate to severe red-green color vision deficiencies, but may fail to detect mild deficiencies, with user testing revealing some misclassifications between different deficiencies. In addition, some researchers defined the CVD simulator [9], simulating the vision of CVD to validate the performance of algorithm.

Despite these advances, conventional screen-based tests are still constrained by monitor color gamut, viewing geometry, and ambient illumination. These limitations motivate the use of AR as a wearable testing medium. AR can present calibrated visual stimuli within the user's field of view while improving portability and preserving better control over stimulus presentation than unconstrained screen-based setups. Building on prior digital validation work, our study investigates whether an AR-based FM100 Hue implementation can provide consistent assessment results in a practical wearable format.

AR and Wearable Color Vision Systems

In a clinical pilot [19] with 24 CVD subjects, a wearable AR device significantly improved color vision—raising Ishihara scores from 5.8 ± 3.0 to 14.8 ± 5.0 ($p = 0.03$)—with 50% subsequently passing as normal. Motivated by these findings, many AR devices for CVD are introduced [20].

Enrico et al. [12] developed Chroma, a wearable augmented-reality system based on Google Glass, helping colorblind individuals by filtering real-time scenes based on their type of color blindness, but the types of degrees are limited. Shota et al. [21] proposes an AR glasses-based system for CVD support that uses a deep learning model to automatically assess meat cooking levels and provides real-time feedback to CVD individuals. The system enhances cooking accuracy, as confirmed by both quantitative and subjective evaluations. Tang et al. [22]introduced ALCC-Glasses, an optical see-through head-mounted display that modulates light chroma in real-time to compensate for CVD, improving color discriminability, but relying on users' self-awareness, limiting its applicability for those with mild or moderate anomalous trichromacy. Keresteš et al. [23] highlights that customized image adaptation, tailored to the user's CVD severity, provides a better solution than generic daltonization methods. Melillo et al. [19] validates AR device that significantly improves color vision in individuals with CVD, as shown by enhanced Ishihara Vision Test scores and 50% of subjects achieving normal vision results, confirming the effectiveness of AR technology for color vision correction. Qin et al. [24] developed Hue4U, a personalized, real-time color correction system for CVD individuals using Meta Quest headsets, which adapts to users without prior medical diagnosis and significantly improves color discrimination.

Previous AR applications have mainly focused on real-time color enhancement or color labeling for CVD users, with digital color vision tests getting popular [25]. Despite these advancements, most AR applications have focused on real-time color enhancement or labeling, rather than personalized calibration—a key step for accurately modeling individual color perception characteristics. This gap is significant, as generic solutions may not fully accommodate the variability in how different individuals experience CVD. Our work aims to fill this gap by introducing a gamified AR calibration process that generates a user-specific color projection matrix, enabling more precise and adaptive color compensation. This personalized approach can improve the overall user experience and effectiveness, particularly for those with less severe forms of CVD.

From a robotics perspective, these AR-based color vision systems can be interpreted as perceptual augmentation modules within human–robot interaction pipelines. In teleoperation, collaborative robotics, and assistive robotic systems, AR overlays are increasingly used to convey robot states, task constraints, and environmental semantics to human operators. However, without individualized perceptual calibration, such visual cues may be suboptimal or misleading for users with atypical color perception. Our work addresses this gap by providing a standardized, robot-compatible calibration mechanism that aligns human color perception with machine-generated visual cues. In the present paper, we focus on the calibration and validation of this perceptual module itself, robot-in-the-loop behavioral evaluation remains future work.

SYSTEM OVERVIEW

The proposed system performs color vision calibration through an AR-based implementation of FM 100 Hue test. The process consists of three stages: (1) calibration setup and color alignment, (2) AR interactive testing across four hue subsets, and (3) validation against the physical FM 100 Hue results (Figure 1).

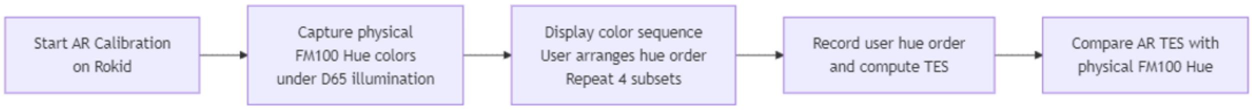


Figure 1. System Workflow

The calibration begins with capturing the RGB values of the physical FM 100 Hue color caps under standardized D65 illumination for color alignment on the AR display. The user then performs the hue-sorting task within the AR environment across four subsets, each representing a quarter of the full hue circle. The system records the user's final hue order for each subset, computes the Total Error Score (TES), and compares the aggregated AR results with the physical FM 100 Hue outcomes to validate consistency and calibration reliability.

Within a robotic context, the proposed system can be regarded as a perceptual calibration module that precedes or complements robot vision and AR-based human–robot interfaces, ensuring that color-coded information presented by robotic systems is perceptually meaningful to individual users.

AR GAME DESIGN

The FM 100 Hue test is a standard color discrimination assessment using 85 color caps arranged in hue order. We replicate this test in AR, displayed as an interactive sequence of color icons on Rokid glasses. The game integrates both pure hue discs and real-world icons (e.g., traffic signs, fruits) to enhance engagement. The final sequence deviation from correct hue order is analyzed to infer hue confusion axes and severity (Figure 2).

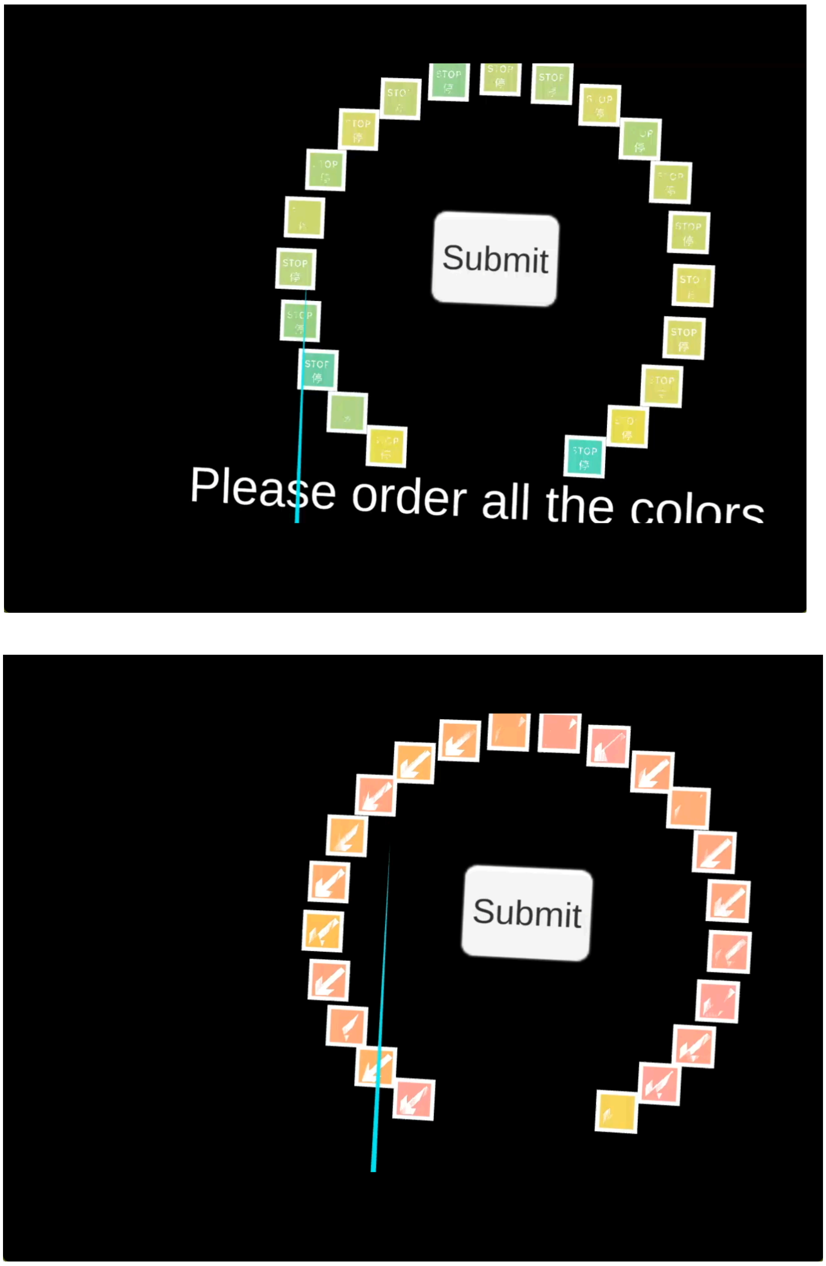


Figure 2. FM 100 Hue Sorting Game

To reduce environmental influence, the in-game colors are sampled directly from physical FM 100 Hue color caps using the Rokid camera’s RGB readings. Measured color deviation <2% after gamma correction under 6500K illumination. These RGB values serve as calibrated inputs, ensuring consistency between real-world and AR-displayed colors.

IMPLEMENTATION AND CALIBRATION STRATEGY

AR Device Calibration and Color Alignment

Since AR display spectra differ from physical environments, accurate color alignment is critical.

Our strategy:

1.Capture RGB values of physical FM 100 Hue color caps using the Rokid camera under neutral illumination;

2.Use these captured RGBs as in-game color references;

3.Display colors on AR glasses and measure rendered RGB via internal sensors, adjusting via gamma correction to minimize deviation.

This physical-to-AR alignment ensures consistency between real and virtual color stimuli, reducing ambient light effects without requiring full spectral calibration.

Software Implementation

• AR Hardware: We utilize Rokid AR smart glasses paired with an Android smartphone (Rokid Air Pro glasses and an Android cell phone as computing unit). The glasses provide a stereoscopic see-through display and 3DoF head tracking. Users wear the AR headset and see virtual test objects (colored tiles) overlaid on their real environment. The smartphone serves as the computing platform running the Unity application and also provides network connectivity for potential cloud-based analysis. The AR display has a resolution and field of view sufficient to show all test items in a comfortable view, and the test was conducted under ordinary indoor lighting conditions

• Unity Application (FM100 Test): The core of the system is developed in Unity 2021 using the Rokid XR Unity SDK for AR support. The application is structured into multiple scenes and managers for modularity. Persistent managers (e.g., GameManager, RandomManager) handle test logic and randomization, while Unity's EventSystem manages user interactions via a pointer controlled by an external Android device. A pointer (an on-screen reticle or raycast) in the AR environment is controlled by the orientation of the Android device via its motion sensors. By tapping the screen of the Android device while the pointer is over a selectable object, the user can "click" it. In the sorting game, this mechanism is used to pick up and drop virtual color tiles. We opted for a select-and-swap interaction: the user selects one tile, then selects a second tile, and the two swap positions. This was found more manageable with pointer input than a drag-and-drop approach in 3DoF conditions. Visual feedback (like highlighting the tile under the pointer and a laser pointer beam) helps the user target items. The application logs all user moves and the final sequence for scoring.

• Color Sample Rendering: Each of the 85 FM100 hues is represented as a virtual colored cube or card in Unity. We derived the RGB color values for these objects from known CIE tristimulus values of the physical FM100 caps under standard D65 illumination. By using Unity's unlit shader materials and a linear color space in rendering, we ensure that the displayed colors are not altered by shading or lighting effects. Only a neutral lighting is applied for the UI elements to avoid color shifts on the hue stimuli. The AR content is rendered with a transparent background so that real world is visible. To preserve stimulus clarity, we used controlled rendering settings and calibrated color values. This careful color management is critical for test fidelity, as discussed in Section Color Rendering Pipeline.

• Data Logging and Networking: The system records detailed data for each session in JSON format stored locally on the smartphone. Logged data include the computed error score for each cap and the overall TES, timing per move, The architecture is designed to be extensible to transmit results to a remote server or electronic medical record system for applications. In our current implementation, data is analyzed on-device, but for more computationally intensive AI tasks (e.g., deep neural network inference or cloud-based anomaly analysis), the system can offload data to a cloud service and retrieve results asynchronously.

RESULTS

We evaluated system through a user study comparing the AR FM100 test to the traditional test.

User Study: AR-FM100 vs. Manual FM100

Study Design: We recruited n = 52 participants (40 with normal color vision, 12 with known CVD) through the optometry clinic. Among the 52 recruited patients, we excluded data from 5 patients who showed significant differences between the AR test and the physical test. The following calculations are based on the test results of these 47 patients. Ages ranged from 14 to 37 and each CVD subject was age-matched with approximately three normal controls to mitigate any age effects on the FM100. Each participant took both the manual FM100 test and the AR-based FM100 test in a randomized order (to counterbalance any learning or fatigue effects). A rest period of 10 minutes was given between tests. The manual test was administered under standardized clinical illumination per the test instructions. The AR test was conducted in the same room under ordinary indoor lighting conditions. We ensured the subject's visual acuity was corrected to normal and the AR headset was adjusted for each user.

Metrics Collected: For each test (AR and manual), we recorded the Total Error Score (TES). We also measured the completion time for each test. After both tests, participants filled out a short usability questionnaire rating their experience (e.g., ease of use, comfort, preference for one format or the other). For analysis of diagnostic performance, we designated a TES threshold to classify "CVD vs Normal" for each test. Based on literature, a TES above 100 is often indicative of a color vision defect for FM100, but we refined this threshold using our normal group's distribution (we ended up using >72 as a threshold for abnormal, which yielded the best separation for the manual test in our sample).

To quantify how closely each patient's AR test score matches the physical test score, we computed various statistics on the paired data. Here are the key metrics and their results (calculated excluding five extreme outliers where |AR−physical| > 80 points for more representative agreement):

Pearson correlation coefficient (r): Measures the linear association between AR and physical scores. Using the formula $r = \frac{\sum(AR_i - \overline{AR})(Phys_i - \overline{Phys})}{\sqrt{\sum(AR_i - \overline{AR})^2 \sum(Phys_i - \overline{Phys})^2}}$, we obtained r≈0.87. This strong positive correlation indicates that higher physical scores tend to correspond to higher AR scores. The coefficient of determination is $R^2=r^2\approx0.76$, meaning about 76% of the variance in AR scores is explained by the physical scores.

Mean difference ("bias"): Compute the average of the pointwise differences (AR – Physical). We found mean(AR−Phys) ≈ +21.0 points. This means AR results are on average about 21 points higher than the physical scores (after outlier exclusion) , indicating a systematic positive bias of the AR test relative to the manual test. Although this offset does not eliminate the strong association between the two modalities, it suggests that raw AR TES values should not be interpreted as directly interchangeable with physical FM 100 Hue TES values, particularly near threshold-based diagnostic boundaries.

Standard deviation of differences (SD): Compute SD of (AR−Physical) deviations. We found SD ≈ 30.8 points. This measures the spread of the AR-Phys differences around the mean bias. A lower SD implies more consistent agreement.

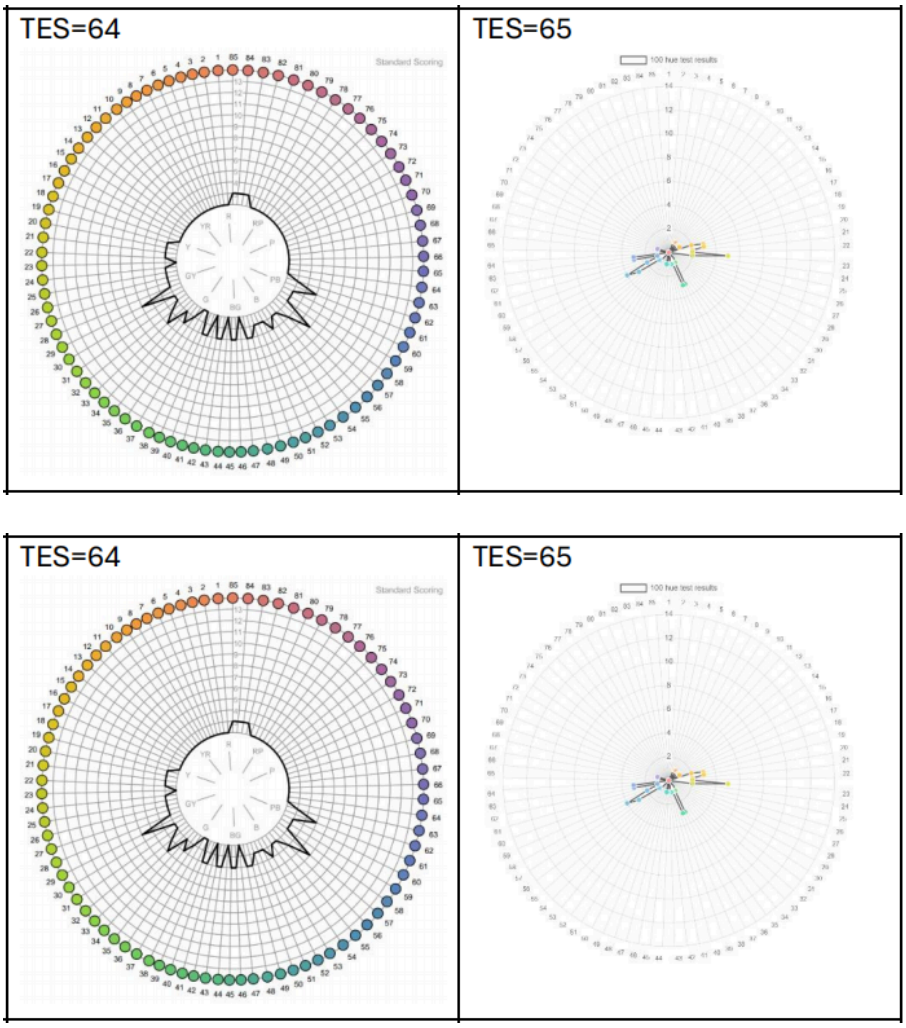
Bland–Altman (limits of agreement): Using the mean ±1.96·SD of the differences, the 95% limits of agreement are approximately −39.4 to +81.4 points. This interval shows that almost all (≈95%) of AR scores lie within about ±60 points of the physical scores (accounting for the +21-point bias). The relatively narrow spread (±~60) indicates reasonable agreement given the score scale. Taken together, these findings indicate strong relative agreement between the AR and physical tests, while also showing that score-level substitution without calibration would be inappropriate.

Linear regression fit (AR vs. Physical): Fitting AR = m·Physical + b gave a slope m≈1.00 and intercept b≈20.9. This regression line (R^2≈0.76) is very close to the identity line (slope=1, intercept=0), indicating that, on average, AR scores increase almost one-for-one with physical scores. The near-unity slope and small intercept suggest minimal scaling bias between the two tests.

Experiments on all participants show that

5 subjects (≈11.6%) showed absolute differences exceeding 80 points, suggesting potential difficulties in operating AR devices.

Based on the judging rule of FM-100 test , the AR test correctly classified 40 out of 47 subjects (accuracy = 85.1%) (Figure 3).



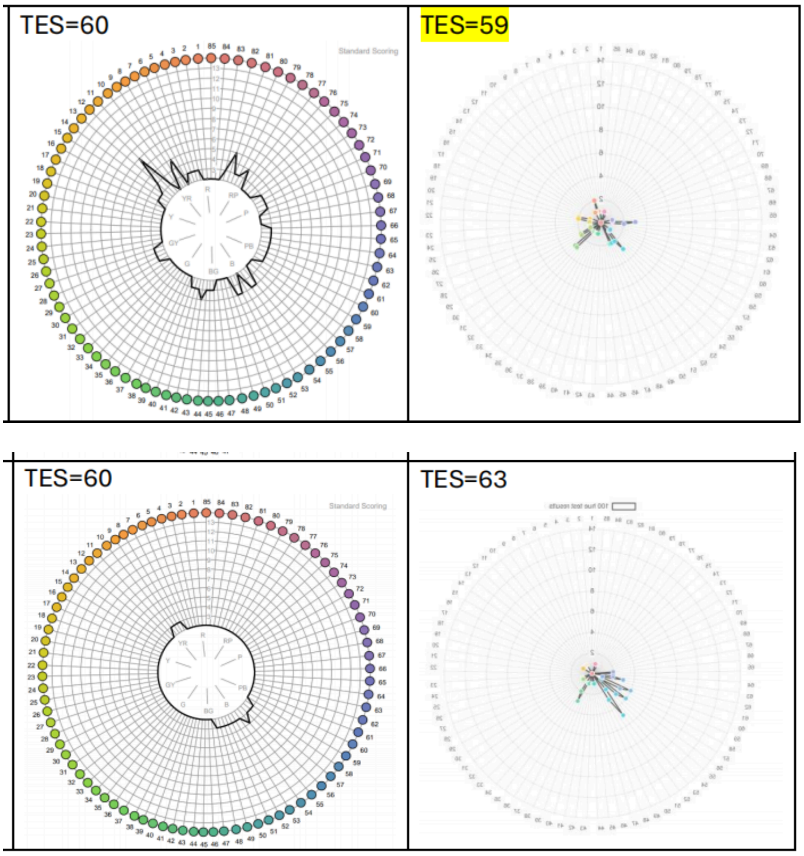


Figure 3. FM-100 physical test results (left) and AR test result (right) graph and TES score

DISCUSSION

The experimental results demonstrate that the proposed AR-based FM 100 Hue test shows strong relative agreement with the traditional physical test, supporting its feasibility as a standardized wearable platform for color vision assessment. The strong correlation between AR-based and manual test scores indicates that calibrated AR displays can preserve inter-subject differences in color discrimination performance. At the same time, the observed positive bias indicates that the present implementation should not yet be treated as a raw-score-equivalent replacement for the physical FM 100 Hue test, especially when decisions depend on threshold-based classification. Moreover, participant feedback suggests that the gamified AR format improves comfort and engagement, which is particularly important for repeated or long-term use in real-world systems. Beyond validation, the proposed system has important implications for robotic and intelligent systems that rely on human–machine collaboration. We emphasize that the present study does not include a robot-in-the-loop validation task, therefore, the robotic use cases discussed here should be interpreted as application scenarios enabled by the calibration output. From a perception alignment perspective, the AR-based FM 100

Hue test provides quantitative, user-specific information about color discrimination ability. Unlike categorical labels such as “normal” or “color deficient,” the resulting error patterns and scores capture continuous variations in perceptual sensitivity across the hue spectrum. This information can be directly incorporated into human–robot interfaces, enabling robotic systems to adapt how color-coded information is presented to individual users.

In teleoperated and shared-autonomy robotic systems, human operators often rely on visual overlays to interpret robot state, environment semantics, and task constraints. Color is frequently used as a compact and intuitive encoding channel for such information. However, without accounting for individual perceptual differences, color-based cues may be ambiguous or misleading for operators with CVD. By integrating AR-based color vision calibration into the interface pipeline, robotic systems can dynamically adjust color mappings, enhance discriminability along confusion axes, or select alternative visual encodings that better align with the operator’s perceptual capabilities. This perception-aware adaptation has the potential to improve task performance, reduce cognitive workload, and enhance operational safety.

In collaborative and assistive robotics, where robots operate in close proximity to humans, effective communication between human and robot is essential. AR glasses are increasingly explored as lightweight, hands-free interfaces that convey robot intentions, alerts, or guidance cues. The proposed calibration framework enables such AR-mediated communication to be inclusive and user-aware, ensuring that visual cues remain interpretable across a diverse population with varying color perception. In this sense, the system supports a shift from generic human–robot interfaces toward personalized and adaptive interaction paradigms.

Several limitations and future directions remain. The current implementation relies primarily on RGB-space calibration. While this approach is sufficient to achieve strong agreement with the physical FM 100 Hue test, more accurate modeling of human color perception may require calibration in perceptually uniform color spaces or device-specific spectral characterization. Future work will explore extending the framework to LMS-based models and integrating spectral measurements to further improve perceptual fidelity. Additionally, longitudinal studies are needed to evaluate test–retest reliability and user adaptation effects across different AR devices and robotic application scenarios.

Overall, this work demonstrates that AR-based color vision calibration is not only clinically meaningful but also highly relevant to robotics and human–robot interaction. By treating human color perception as a calibratable component of the perception pipeline, the proposed approach contributes to the broader goal of aligning human and machine perception, enabling more reliable, adaptive, and inclusive robotic systems.

CONCLUSION

We presented an AR–based interactive calibration system that implements the classical FM 100 Hue test on wearable AR glasses. Through gamified interaction and color-accurate rendering, the system enables users to perform standardized color discrimination testing in an environment-robust and portable manner. Comparative experiments between the AR-based and physical FM 100 Hue tests confirmed high consistency in diagnostic outcomes, validating the reliability of the AR method for individual color vision calibration. This level of agreement suggests that this proposed system could provide a portable and robust tool for monitoring the color vision of workers in textile manufacturing and quality inspection. More broadly, wearable AR devices can serve not only as practical platforms for color vision assessment but also as promising calibration tools for future personalized color adaptation in AR and human–robot interface settings.

Author Contributions

Conceptualization, Zhenyu Xiao and Xinghong Hu; methodology, Xinghong Hu; software, Taizhi Wang; validation, Shaoying Tan and Wuyao Shen; formal analysis, Elaine Yuen Ying To; resources, Zhenyu Xiao and Wuyao Shen; writing—original draft preparation, Sichuang Xu; writing—review and editing, Zhenyu Xiao, Xinghong Hu, and Wuyao Shen; visualization, Taizhi Wang; supervision, Zhenyu Xiao; project administration, Zhenyu Xiao; funding acquisition, Xinghong Hu. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This research was supported by the Shenzhen Science and Technology Program (Grant No. RCBS20221008093249079), the Shenzhen Polytechnic University-level Scientific Research Project (No. 6022310013K), and the Shenzhen Polytechnic High-level Talent Research Start-up Project (No. 6022310050k).

Data Sharing Agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

The study was approved by the Ethics Committee of Shenzhen Polytechnic University, and was conducted in accordance with the ethical principles of the Declaration of Helsinki. Informed consent was obtained from all individual participants involved in the study.

Acknowledgements

Not applicable.

REFERENCES

[1] Ribeiro MG, Gomes AJP. Recoloring algorithms for colorblind people: A survey. ACM Computing Surveys. 2019;52(4):72. doi: 10.1145/3329118

[2] Angerbauer K, Rodrigues N, Cutura R, Öney S, Pathmanathan N, Morariu C, Weiskopf D, Sedlmair M. Accessibility for color vision deficiencies: Challenges and findings of a large scale study on paper figures. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; 29 Apr-5 May 2022; New Orleans, LA, USA. New York, NY: Association for Computing Machinery; 2022. Article 134, 1-23. doi: 10.1145/3491102.3502133

[3] International Commission on Illumination. CIE 240:2020 - Enhancement of Images for Colour-Deficient Observers. Vienna: CIE Central Bureau; 2020. 63 p. doi: 10.25039/TR.240.2020

[4] Honson VJ, Dain SJ. Performance of the standard pseudoisochromatic plate test. American Journal of Optometry and Physiological Optics. 1988;65(7):561-570. doi: 10.1097/00006324-198807000-00006

[5] Farnsworth D. The Farnsworth-Munsell 100-hue and dichotomous tests for color vision. Journal of the Optical Society of America. 1943;33(10):568-578. doi: 10.1364/JOSA.33.000568

[6] Farnsworth D. The Farnsworth-Munsell 100-Hue Test for the Examination of Color Discrimination: Manual. Baltimore, MD: Munsell Color Company, Inc.; 1957.

[7] Regan BC, Reffin JP, Mollon JD. Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. Vision Research. 1994;34(10):1279-1299. doi: 10.1016/0042-6989(94)90203-8

[8] Xu L, Li Q, Li Q, Liu X, Xu Q, Luo MR. Personalized image enhancement method for color deficient observers. Optics Express. 2022;30(8):13079-13094. doi: 10.1364/OE.450808

[9] Brettel H, Viénot F, Mollon JD. Computerized simulation of color appearance for dichromats. Journal of the Optical Society of America A: Optics, Image Science, and Vision. 1997;14(10):2647-2655. doi: 10.1364/JOSAA.14.002647

[10] Machado GM, Oliveira MM, Fernandes LAF. A physiologically-based model for simulation of color vision deficiency. IEEE Transactions on Visualization and Computer Graphics. 2009;15(6):1291-1298. doi: 10.1109/TVCG.2009.113

[11] Shen W, Mao X, Hu X, Wong TT. Seamless visual sharing with color vision deficiencies. ACM Transactions on Graphics. 2016;35(4):70. doi: 10.1145/2897824.2925878

[12] French A, Rose K, Thompson K, Cornell E. The evolution of colour vision testing. Australian Orthoptic Journal. 2008;40(2):7-15.

[13] Dain SJ. Clinical colour vision tests. Clinical and Experimental Optometry. 2004;87(4-5):276-293. doi: 10.1111/j.1444-0938.2004.tb05057.x

[14] Baraas RC, Foster DH, Amano K, Nascimento SMC. Anomalous trichromats' judgments of surface color in natural scenes under different daylights. Visual Neuroscience. 2006;23(3-4):629-635. doi: 10.1017/S0952523806233297

[15] Tanuwidjaja E, Huynh D, Koa K, Nguyen C, Shao C, Torbett P, Emmenegger C, Weibel N. Chroma: A wearable augmented-reality solution for color blindness. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; 13-17 Sep 2014; Seattle, WA, USA. New York, NY: Association for Computing Machinery; 2014. p. 799-810. doi: 10.1145/2632048.2632091

[16] Ghose S, Parmar T, Dada T, Vanathi M, Sharma S. A new computer-based Farnsworth Munsell 100-hue test for evaluation of color vision. International Ophthalmology. 2014;34(4):747-751. doi: 10.1007/s10792-013-9865-9

[17] Fanlo-Zarazaga A, Echevarría JI, Pinilla J, Alejandre A, Pérez-Roche T, Gutiérrez D, Ortín M, Pueyo V. Validation of a New Digital and Automated Color Perception Test. Diagnostics. 2024;14(4):396. doi: 10.3390/diagnostics14040396

[18] Trukša R, Fomins S, Jansone-Langina Z, Tenisa L. Colour Vision Changes across Lifespan: Insights from FM100 and CAD Tests. Vision. 2024;8(3):53. doi: 10.3390/vision8030053

[19] Melillo P, Riccio D, Di Perna L, Sanniti Di Baja G, De Nino M, Rossi S, Testa F, Simonelli F, Frucci M. Wearable improved vision system for color vision deficiency correction. IEEE Journal of Translational Engineering in Health and Medicine. 2017;5:3800107. doi: 10.1109/JTEHM.2017.2679746

[20] Sutton J, Langlotz T, Plopski A. Seeing Colours: Addressing Colour Vision Deficiency with Vision Augmentations using Computational Glasses. ACM Transactions on Computer-Human Interaction. 2022;29(3):26. doi: 10.1145/3486899

[21] Chiba S, Zhu Z, Inoue D, Mao X. Deep learning and augmented-reality glasses based meat cooking support for color vision disorder compensation. In: 2023 Nicograph International (NicoInt); 30 Jun-2 Jul 2023; Sapporo, Japan. Piscataway, NJ: IEEE; 2023. p. 62-67. doi: 10.1109/NICOINT59725.2023.00019

[22] Tang Y, Zhu Z, Toyoura M, Go K, Kashiwagi K, Fujishiro I, Mao X. ALCC-Glasses: Arriving light chroma controllable optical see-through head-mounted display system for color vision deficiency compensation. Applied Sciences. 2020;10(7):2381. doi: 10.3390/app10072381

[23] Milić Keresteš N, Đurđević S, Novaković D, Zarić M, Kašiković N, Dedijer S, Vladić G. Customized Daltonization: Adaptation of different image types for observers with different severities of color vision deficiencies. Universal Access in the Information Society. 2023;22(2):351-367. doi: 10.1007/s10209-021-00847-7

[24] Qin J, Checherin S, Li Y, van der Zwaag B-J, Durmaz-Incel O. Hue4U: Real-Time Personalized Color Correction in Augmented Reality [preprint]. arXiv. 2025. doi: 10.48550/arXiv.2509.06776

[25] Murphy RA. Comparing Color Vision Testing Using the Farnsworth-Munsell 100-Hue, Ishihara Compatible, and Digital TCV Software [thesis]. Forest Grove, OR: Pacific University College of Optometry; 2015.