

Programmable Vision: Intrinsic Visual Attributes-Based Vision Modulation

Qing Zhang

ABSTRACT

Analog glasses are a common and effective vision augmentation technology, enabling wearers to regain their clear vision. Meanwhile, many researchers have also explored if we can use digital technology in a similar fashion to augment our visual perception. For example, augmented reality headsets and comparable technology focus on enhancing human vision with computational assistance.

This dissertation explores the concept and the benefits of programming human vision by treating our intrinsic visual attributes as *built-in* functions or methods. To approach this concept, we create a simple variant of optical see-through glass(es), which enable four studies that illustrate the potential and benefits of modulating vision based on leveraging intrinsic visual attributes.

These studies examine the feasibility of conveying directional cues by subtly triggering motion perception, dynamically adjusting peripheral vision, synchronizing eye movements of two parties, and simulating low vision. The studies reveal that programmable vision based on intrinsic visual attributes is a direct vision-computer interaction approach and has several benefits: (1) computational efficiency, (2) instant real-world adaptability, (3) direct and unobtrusive visual information delivery, (4) low-resolution requirements, (5) optical glasses-like visual field.

INTRODUCTION

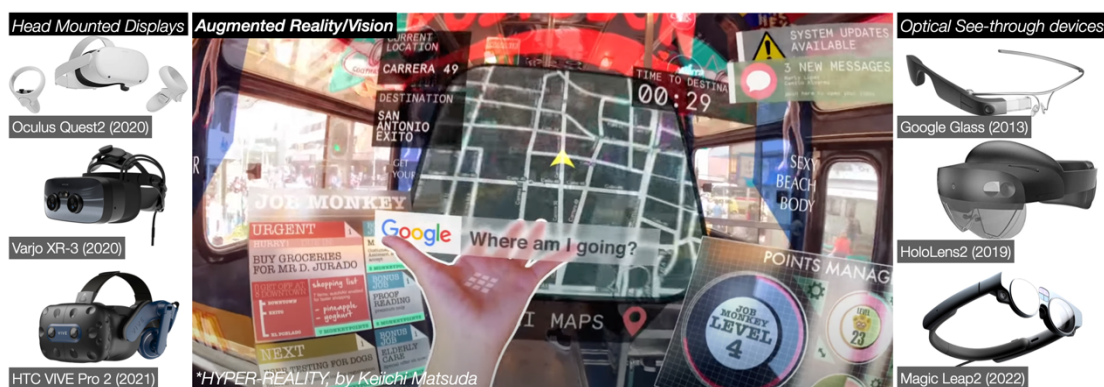


Figure 1. An example of Augmented Reality/Vision. The scene in the middle is a frame originally from *HYPER-REALITY* created by Keiichi Matsuda. The leftmost column is the mainstream head-mounted display often used to modulate vision. The rightmost column is the optical see-through device intensively used in vision modulation.

Being put in a dark environment, or losing one's vision (temporarily or permanently), make one understand how much we rely on seeing with our eyes. As vision perception is one of our most dominant senses, many researchers attempt to enhance and augment this sense.

Analog vision augmentations, e.g., binoculars and the microscope, led to a new understanding of our world. Then, people could see previously invisible worlds. These new augmented vision capabilities led to novel insights, e.g., geographic data collection and astronomical observations in the case of binoculars and germ theory in the case of the microscope. More recently, researchers have attempted to digitally enhance our vision using virtual/augmented/mixed reality glasses.

Just like the Figure 1. depicts, a common vision of augmented reality or augmented vision is adding additional virtual contents overlapping/overlaying with the real world. The other way to modulate vision is to digitize the real world as a two-dimensional video and then substitute the real world with the digitized counterpart. In short, both mainstream approaches target modifying the real world, focusing on indirect ways to interact with our visual system. The challenges those conventional approaches face: are (1) indirect interaction with vision, (2) heavy computational power consumption, (3) complex calibration algorithm, (4) high cost, (5) unrealistic/unintuitive digital substitution, and (6) complex optical design.

Let's dive into one of those challenges, indirect interaction with vision. The scene shown in Figure 1. presents a navigation example. Yet, if we want to arrive at a certain destination, why do we need first to wear a bulky Head-mounted display and then open up a graphic interface? Not only that, but we still need to wait for the application to load, type in our destination, and wait for the answer to appear. Let's assume that we are lucky enough, the device is computationally powerful enough, and we already see some map-like content displayed in front of our eyes. Unluckily, we still need to understand the map and pay a tremendous amount of attention to prevent we misread the map or missing the upcoming auditory reminders.

In most cases, arriving at a certain destination only requires a series of few directional decisions. It means the series of operations with the aforementioned computational device is redundant or can be further simplified. Additionally, the conventional ways to substitute the real world with a digitized one or add virtual content overlaying with real-world objects will more or less occlude the visual information behind the computationally generated visual content. Such methods can be fatal in some daily living scenarios, for instance, when a fast-moving car blocked from the virtual content is approaching us.

To enable the human vision to better fit our rapidly changing world, some researchers argue humans should become cyborgs. We should combine our bodies with computational assistance, as this may overcome the limitations of our bodies' abilities to cope with the complexity of the external world. In such a context, biological methods may permanently alter our bodies, and certain

surgeries can physically modify our bodies to adapt to a limited set of tasks. However, no one can guarantee safety, especially regarding negative long-term effects of those cybernetic enhancements.

On the other hand, not implanting devices or biologically altering our bodies, researchers created external devices that augment/ compensate for limitations in our current visual system by applying computational assistance.

In contrast, Programmable Vision here aims to directly interact with our visual information processing mechanisms without or with a less secondary agency between the user and the real world.

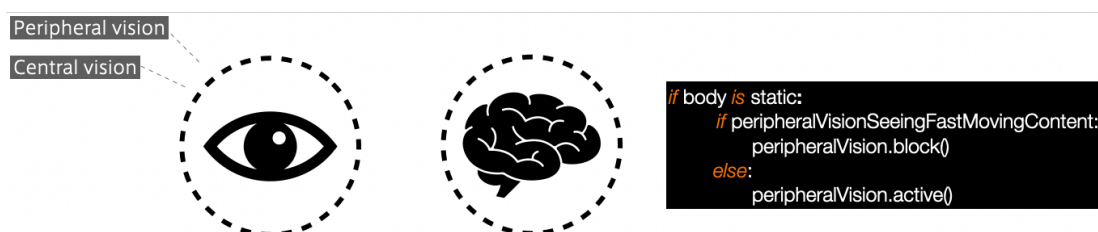


Figure 2. The basic idea of Programmable Vision leveraging the intrinsic visual attributes as built-in functions and or methods, just like how we create a computer program to perform a specific task.

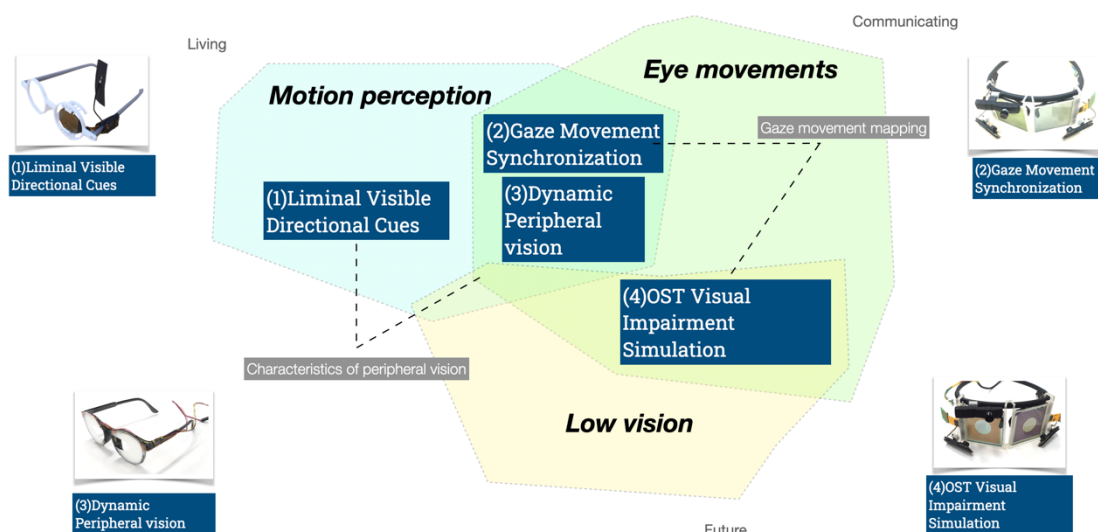


Figure 3. The overview of all four works in this dissertation.

Programmable Human Vision Based on Intrinsic Visual Attributes

In terms of programming human vision, human intrinsic visual attributes and their quantifiable adjustable features suggest the potential that we can manipulate them, similar to how we design computer programs.

To perform a specific task proficiently and minimize potential mistakes, we build computer programs to help us. If we know a certain task's workflow, we can use code (conventional programming) to construct and optimize the work procedures. Meanwhile, to address problems with more complex structures and larger scales, we can also take advantage of machine learning to approach the desired solutions heuristically.

On the other hand, findings in cognitive science regarding how the human brain functions indicate the possibility and potential that we can also program our visual perception by treating the intrinsic visual attributes as the variables to program human visual perception favorably. See Figure 2.

Research Questions

Therefore, the main research question here is **can we program/modulate visual perception by leveraging intrinsic human visual attributes?** To validate and implement Programmable Vision and to address this question, this dissertation contains four studies exploring the intrinsic visual attributes of directly triggering motion perception, dynamically adjusting peripheral vision, directing eye movements, and simulating visual impairment. We rely on motion perception to interact with the real world. Eye movements are the starter of most attempts to observe and interact with the surroundings, and it is also a rich information resource in terms of communication. Low vision, on the other hand, is an unavoidable trend that, sooner or later, we need to face. Altogether, addressing these three parts enables us to examine our concept's validity from various perspectives.

Then the research question is divided into specific sub-research questions explored over Chapter 4 to Chapter 7. They are (1) Can dynamically adjustable peripheral vision reduce typical motion sickness symptoms without impacting immersion in both VR and the real world? (2) Can we convey directional cues by directly triggering motion perception? (3) Can we synchronize two parties' gaze movements in both synchronous and asynchronous scenarios by modulating saliency? (4) Can we effectively simulate both central and peripheral vision loss and increase the users' awareness of visual impairment?

Contributions

This thesis examines an initial exploration of the concept ***Programmable Vision***. It systematically evaluates how human intrinsic visual attributes can be used to program visual perception in specific scenarios.

In the following, we present the main contributions of this dissertation: (1) Concept of “Programmable Vision” together with an initial classification of visual attributes and their

capabilities to be used for programming visual perception with minimal liminality. In comparison, most existing vision augmentation research relies on substituting the real world with digitized images (e.g., virtual reality, video see-through augmented reality, etc.). Modulate visual perception based on intrinsic visual attributes is rarely explored. (2) Leveraging the characteristics of peripheral vision as intrinsic attributes, we present a prototype of smart glasses that can dynamically occlude a user's peripheral vision. In the user study, we confirmed that our method could significantly reduce typical motion sickness symptoms (difficulty concentrating, nausea, and head feeling heavy) in a VR environment (n=19) while having little impact on immersion, as well as in a daily commute (n=2). (3) We present a method that can convey directional cues with liminal visibility and high accuracy. Our user study (n=14) confirmed that we could deliver visual directional cues without distracting or covering the participants' visual field. (4) Our gaze movement synchronization study (n=6) proposes a wearable method that synchronizes gaze movements between two parties, such as a leader and a follower, for both real-time or offline replay. (5) Our method in visual impairment simulation is the first lightweight optical-see-through-based and real-time eye-tracking integrated visual impairment simulation solution. Its user study (n=14) confirmed that our method could enable users to experience different types of visual impairments in everyday activities unconstrainedly without significant obstructions to one's field of vision or the need for pre-defined virtual scenarios. Additionally, our method significantly reduced the users' peripheral vision and visual acuity. Altogether, the user study of our approach increased the users' awareness of visual impairments without triggering typical motion sickness symptoms such as headache and visual fatigue.

Structure of This Dissertation

This dissertation contains nine chapters. The first chapter introduced the core concept, what programmable vision is, what those intrinsic visual attributes are, and the background behind this concept. A comparison between programmable human vision and the related visual perception modulation research was briefly reviewed. Contributions that distinguish this dissertation from current works are described.

After this introduction Chapter 1, we begin with a literature review (Chapter 2) that introduces the intrinsic visual attributes and the prior works conducted to modulate visual perception.

Chapter 3 presents the approach of modulating visual perception while leveraging intrinsic visual attributes. Chapter 4 goes over the attempt to reduce motion sickness symptoms in both virtual reality and real-world scenarios while taking advantage of the characteristics of peripheral vision and central vision. Chapter 5 introduces the concept of conveying distinguishable directional cues with liminal visibility, followed by two tasks to evaluate its feasibility. Chapter 6 explored the typical eye movements and proposed a method synchronizing two parties' gaze movements. Chapter 7 investigates visual impairments and engineers an optical see-through-based prototype that can simulate both central vision loss and peripheral vision loss without banning access to the

real world. Chapter 8 and Chapter 9 summarizes our research and further discuss the concept of programmable vision with benefits and limitations.

Chapter	Paper Title	Authors	Conference
1&2	Programmable Peripheral Vision: augment/reshape human visual perception	Qing Zhang.	CHI'22
4	Tunnel Vision – Dynamic Peripheral Vision Blocking Glasses for Reducing Motion Sickness Symptoms	Qing Zhang , Hiroo Yamamura, Holger Baldauf, Dingding Zheng, Kanyu Chen, Junichi Yamaoka, Kai Kunze.	ISWC'21
6	GazeSync: Eye Movement Transfer Using an Optical Eye Tracker and Monochrome Liquid Crystal Displays	Qing Zhang , Yifei Huang, George Chernyshov, Juling Li, Yun Suen Pai, and Kai Kunze.	IUI'22
7	Seeing our Blind Spots: Smart Glasses-based Simulation to Increase Design Students' Awareness of Visual Impairment	Qing Zhang , Giulia Barbareschi, Yifei Huang, Juling Li, Yun Suen Pai, Jamie A Ward, and Kai Kunze.	UIST'22
7	Experience Visual Impairment via Optical See-through Smart Glasses	Qing Zhang , Xiongqi Wang, Thad Starner, Yifei Huang, George Chernyshov, Giulia Barbaresch, Yun Suen Pai, Jing Huang, Junichi Yamaoka, Jamie A Ward, and Kai Kunze.	UIST'22

Table 1. Related publications included in this dissertation.

APPROACH: Programmable Vision

In the context of augmenting vision, prior studies (reviewed in Chapter 2) have explored vision augmentation from various aspects, such as substituting real-world scenery with two-dimensional videos, generating virtual objects that overlap with real-world environment, leveraging computationally adjustable materials to substitute optical lenses, and adding seamless visual information via heads up displays.

However, equipping human vision with external functions more or less inevitably leads to a bulky design (e.g., sophisticated optical design), heavy demand for computational resources (Augmented Reality, Diminished Reality), and additional cognitive processing (distinguishing and comprehending the displayed content). Although the computational devices will become more powerful and lighter over time, it is likely to require a considerable amount of time, especially since the current computer chip's manufacturing process will touch its physical limitation in a couple of years.

Aiming to directly interact/communicate with our visual information processing, we propose Programmable Vision, which modulates vision by manipulating the different characteristics of intrinsic visual attributes. See Figure 4. The keywords on the left hand are some of the intrinsic visual attributes related to visual information processing and the physiological structures of the human visual system based on findings in psychology, physiology, and neuroscience (see Chapter 2). The basic idea of Programmable Vision is conveyed in the pseudo-code with the black background, see Figure 4; the intrinsic visual attributes are treated as *built-in* functions and methods, similar to how we create a computer program to perform a specific task.

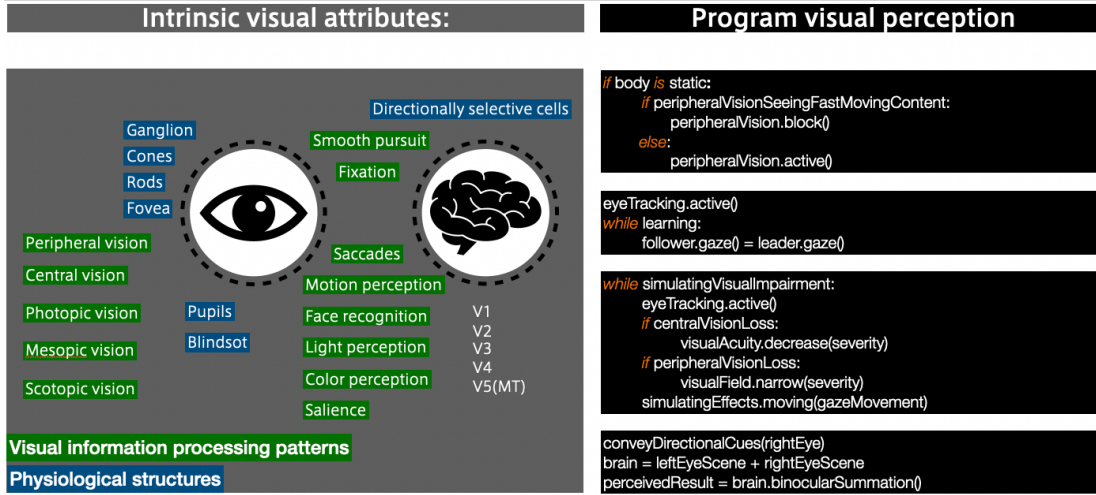


Figure 4. The concept of Programmable Vision.

Comparison to Alternative Approaches

To distinguish the difference and novelty of Programmable Vision with the related works, we compared the basic concepts of Programmable Vision, video see-through-based approach, and optical see-through-based approach. See Figure 5. Programmable Vision aims to interact and directly augment visual information processing while keeping the real-world scenery unmodulated. On the other hand, video see-through-based methods substitute the real world with digitized video plus computational image modulation, which bans access to real-world scenery. Optical see-through-based methods generate virtual content overlaying with real-world objects. In such cases, if the generated content is too vivid and salient, we sacrifice the visual information of the real world hidden behind the generated content. Additionally, both OST and VST-based methods heavily rely on computational resources. Besides their high cost, the mainstream and the latest high-end products can only provide about a 110-degree field of view, which is far narrower than the human visual field.

To also point out the key features of Programmable Vision with OST and VST-based methods, the comparison is shown in Figure 6. We summarize the key distinguishing feature of Programmable Vision as *directer interaction/communication* with our visual perception. The corresponding benefits of Programmable Vision integrated approach are low resolution constrain, computational

efficiency, low battery consumption, wider field of view, low cost, real-world adaptability, and low weight.

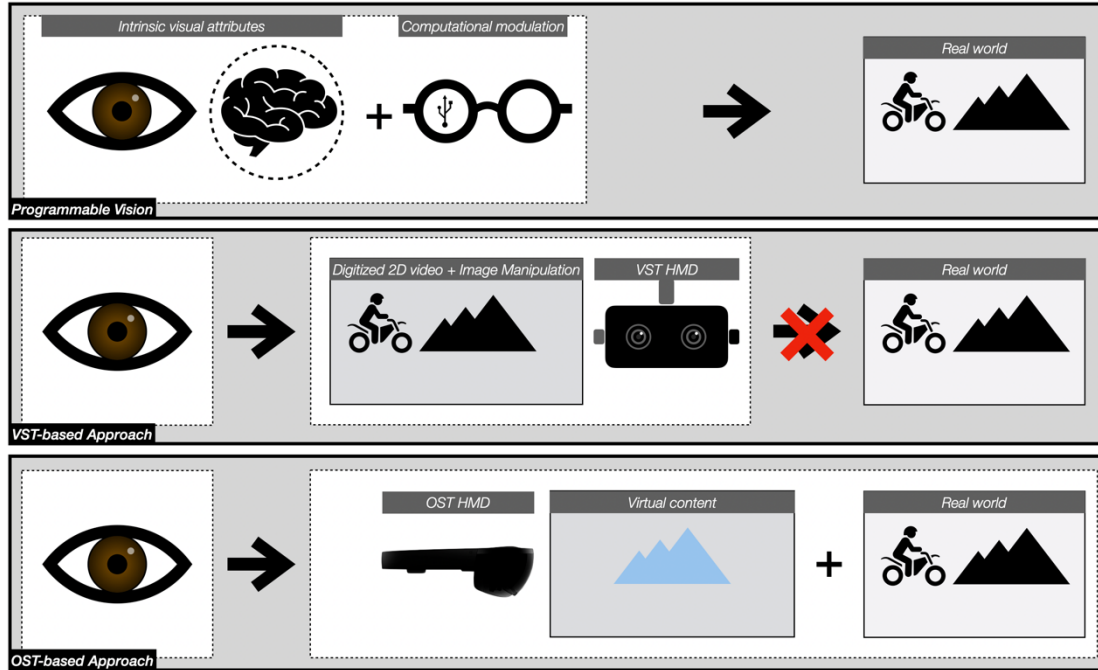


Figure 5. Programmable Vision (uppermost row) aims to interact with visual perception directly while keeping real-world scenes unmodulated. VST (video see-through) based approach substitutes the real world with a digitized two-dimensional video and bans access to the real-world scene. The OST (optical see-through) based approach adds virtual visual content to overlay the real-world scenery.

To sum up, Programmable Vision is a concept to treat intrinsic visual attributes as the *built-in* functions and methods in visual information processing. In this chapter, we introduced Programmable Vision and the background regarding how to modulate vision. We also compared the difference that how Programmable Vision, VST and OST-based approach interact with human vision as well as the real-world scenery. To further point out the key features of Programmable Vision and the related VST & OST based methods, we listed the difference in the Figure 6.

In the following Chapters from 4 to 7, we conducted four feasibility studies on triggering motion perception subtly, synchronizing eye movements, dynamically adjusting peripheral vision, and simulating visual impairment. The reasons behind this scope are we rely on motion perception to interact with the real world; eye movement is a starter of an attempt to observe the world and is also a rich information resource in terms of communication. Visual impairment is an inevitable trend that sooner or later we are likely to face. Altogether, addressing these three parts enable us to examine the validity of Programmable Vision from various perspectives.

Optical see-through AR	Video see-through AR	Programmable Vision
Adding luminous contents	Substituting the real world by digitized 2-dimensional video	Director communication with visual perception
Cognitive process needed	Cognitive process needed	
Limited contrast	Limited Resolution	No resolution constrain
Limited resolution	Computational heavy	Computational efficiency
Limited field of view	Limited field of view	Original/wider field of view
High cost	High cost	Low cost
Battery life	Battery life	Low Battery Consumption
Latency	Latency	Real world adaptability
Bulky	Bulky	Lightweight
	Sensory conflicts	
Key features comparison of Programmable Vision		

Figure 6. Comparison of Programmable Vision and OST, VST-based methods.

DISCUSSION: Towards Programmable Vision

In this chapter, we summarize and discuss the benefits of programming visual sensation/perception and leveraging intrinsic visual attributes in terms of optimizing visual tasks. We conclude with an outlook for the field and discuss limitations and challenges.

Benefits of Intrinsic Visual Attributes-based Visual Perception Modulation

Overall, our feasibility studies over four individual research confirmed that the concept of **Programmable Vision** could be an inspiring work concerning an intuitive human-computer interface and computationally efficient human assistance.

The benefits of this study are summarized and listed below as three subsections: (1) Low computational power requirements, (2) Instant real world adaptable, (3) Direct and unobtrusive Vision-Computer Interaction, (4) Low-resolution requirements, (5) Optical glasses-like field of view.

Limitations and Open Challenges

This dissertation provides some initial examples of programmable vision. However, for the area to develop, we must understand the vision system on a deeper level in terms of optics, cognitive psychology, neuroscience, and computer simulation. Collaborations across these domains will expand the possible topics programmable vision can address and lead to more compact and practical hardware.

More immediately, several optimizations are possible for the devices created in this dissertation. For example, the current transparent liquid crystal displays need to be replaced with ones with higher refresh rates and higher contrast. The eye-tracking system used requires significant power and extra hardware. A more compact or integrated eye tracker would increase portability. Finally, although we considered social acceptability when designing our prototypes (such as deciding not to include outward-facing cameras), our devices' appearance is still unusual when compared to devices typically worn. Future work may consider redesigning for a more conventional appearance.

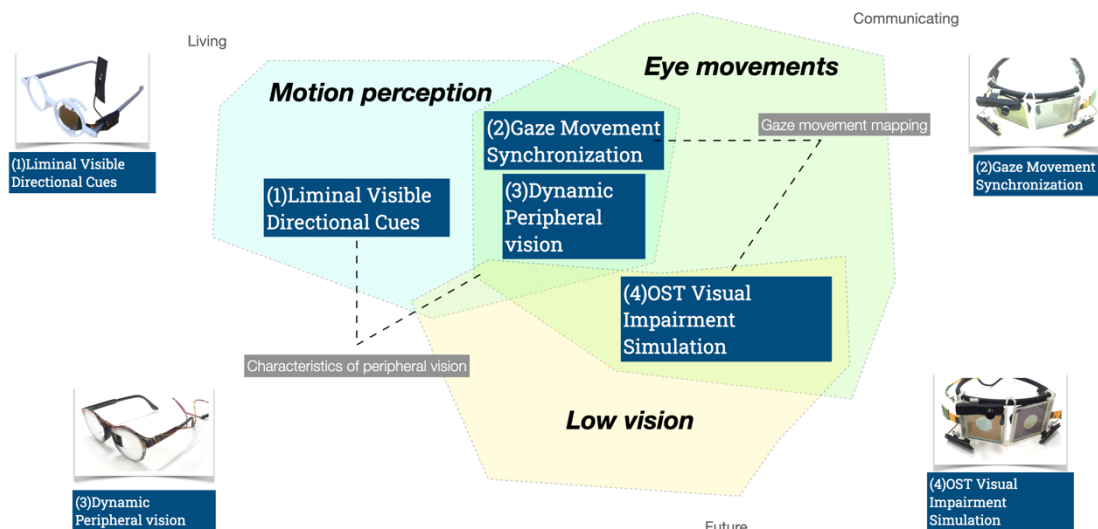


Figure 7. The overview of our approach. We conducted four research to validate Programmable Vision while considering conveying directional cues via subtly triggering motion perception, dynamically adjusting peripheral vision, synchronizing eye movements of two parties, and simulating low vision.

CONCLUSION

This dissertation explores the concept of programmable human vision by treating our intrinsic visual attributes as built-in functions or methods. To approach this concept, we created a simple variant of optical see-through glass(es), which enabled four studies that illustrate the potential and benefits of modulating vision based on leveraging intrinsic visual attributes.

Over four individual research studies, we have provided initial evidence that the concept of **Programmable Vision** can modulate visual experience and perception by leveraging intrinsic visual attributes. The user studies suggest that Programmable Vision can be a direct approach to interact with visual information processing with a simplified interface and reliable performance. Exploring Programmable Vision fully will require a deep interdisciplinary understanding of the vision system over optics, cognitive psychology, neuroscience, and computer science. Comprehensive collaboration among the aforementioned domains is likely to lower the difficulty of implementing Programmable Vision devices and make them more feasible for the general public.