

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie  
Wydział Informatyki

Pomiar charakterystyk czułości na kontrast w  
peryferyjnych obszarach widzenia oraz ich  
wykorzystanie do przyspieszania syntezy obrazów

Michał Chwesiuk

Promotor: dr hab. inż. Radosław Mantiuk, prof. ZUT

SZCZECIN 2019



# Spis treści

|  |    |
|--|----|
| <b>Wprowadzenie</b> . . . . .  | 5  |
| Definicja problemu . . . . .   | 5  |
| Teza i cele pracy . . . . .  | 6  |
| Metodologia badań . . . . .  | 7  |
| Organizacja treści rozprawy . . . . .  | 7  |
| <b>1. Wybrane cechy układu wzrokowego człowieka</b> . . . . .  | 9  |
| 1.1. Budowa i działanie układu wzrokowego człowieka . . . . .  | 9  |
| 1.2. Widzenie centralne i peryferyjne . . . . .  | 11 |
| 1.2.1. Ruchy gałek ocznych . . . . .   | 13 |
| 1.3. Charakterystyki czułości na kontrast . . . . .  | 15 |
| Podsumowanie rozdziału . . . . .   | 18 |
| <b>2. Model czułości na kontrast dla peryferyjnego obszaru widzenia</b> . . . . .                            | 19 |
| 2.1. Eksperymentalny pomiar czułości na kontrast . . . . .   | 20 |
| 2.1.1. Bodźce wzrokowe stosowane w eksperymencie . . . . .   | 22 |
| 2.1.2. Procedura eksperymentu . . . . .  | 29 |
| 2.1.3. Aparatura . . . . .   | 32 |
| 2.1.4. Uczestnicy eksperymentu . . . . .   | 35 |
| 2.2. Rezultaty eksperymentu . . . . .  | 35 |
| 2.2.1. Funkcja czułości na kontrast dla achromatycznego bodźca . . . . .                                     | 35 |
| 2.2.2. Funkcja czułości na kontrast dla achromatycznego bodźca w obszarze peryferyjnym . . . . .             | 36 |
| 2.2.3. Funkcja czułości na kontrast dla chromatycznego i mieszanego bodźca w obszarze peryferyjnym . . . . . | 40 |
| 2.2.4. Wyniki porównania metodologii śledzenia wzroku i migającego bodźca . . . . .                          | 42 |
| 2.3. Model czułości na kontrast dla widzenia peryferyjnego . . . . .   | 44 |
| 2.4. Dynamika percepcji dla zmiany kierunku widzenia . . . . .   | 50 |
| 2.4.1. Eksperyment percepcyjny . . . . .   | 50 |
| 2.4.2. Rezultaty eksperymentu . . . . .  | 56 |
| 2.4.3. Wskazania dla aplikacji . . . . .   | 57 |
| 2.4.4. Eye tracker o wysokiej częstotliwości . . . . .   | 58 |
| Podsumowanie rozdziału . . . . .   | 60 |

---

|   |           |
|---|-----------|
| <b>3. Aplikacje modelu czułości na kontrast w grafice komputerowej . . . . .</b>    | <b>61</b> |
| 3.1. Niejednorodne próbkowanie obrazu w metodzie śledzenia promieni . . . . .       | 61        |
| 3.1.1. Kierunkowa metoda śledzenia promieni . . . . .                               | 62        |
| 3.1.2. Rekonstrukcja próbek obrazu . . . . .  | 65        |
| 3.1.3. Pomiar szybkości renderingu . . . . .  | 67        |
| 3.1.4. Ocena jakości renderingu . . . . .   | 69        |
| 3.2. Wyświetlacz stereoskopowy z akomodacją . . . . .                               | 71        |
| 3.2.1. Rodzaje wyświetlaczy zawierających wskazówki akomodacyjne . . . . .          | 72        |
| 3.2.2. Algorytmy dekompozycji obrazu na wirtualne płaszczyzny . . . . .             | 74        |
| 3.2.3. Budowa wielopłaszczyznowego wyświetlacza light field'ów . . . . .            | 75        |
| 3.2.4. Pomiar wydajności wielopłaszczyznowego wyświetlacza light field'ów . . . . . | 80        |
| 3.2.5. Ocena jakości hybrydowej metody dekompozycji . . . . .                       | 81        |
| Podsumowanie rozdziału . . . . .  | 83        |
| <b>Podsumowanie . . . . .</b>   | <b>85</b> |
| Wnioski końcowe . . . . .   | 86        |
| Odniesienie do tezy . . . . .   | 87        |
| Dalsze prace . . . . .  | 87        |
| <b>Spis rysunków . . . . .</b>  | <b>89</b> |
| <b>Spis tabel . . . . .</b>   | <b>91</b> |
| <b>Publikacje własne . . . . .</b>  | <b>93</b> |
| <b>Literatura . . . . .</b>   | <b>95</b> |

# Wprowadzenie

Algorytmy komputerowe syntezujące obraz są kosztowne obliczeniowo i wymagają znacznych mocy obliczeniowych procesorów. W grafice komputerowej czasu rzeczywistego czas na proces syntezy jest ograniczony, co sprawia, że zastosowanie technik zmniejszających wymagany nakład obliczeniowy jest kluczowe. Jednym ze sposobów przyspieszania syntezy obrazów jest wykorzystanie *technik percepcyjnych*, w których dostosowuje się algorytm renderujący obraz do możliwości percepcyjnych *układu wzrokowy człowieka*.

Układ wzrokowy wybiórczo interpretuje informację o obserwowanej scenie. Odebrana przez komórki światłoczułe oka informacja przekazywana jest do kory wzrokowej w celu stworzenia obrazu otoczenia. Gęstość komórek światłoczułych uzależniona jest od odległości od punktu skupienia wzroku, a wraz z odległością kątową liczba neuronów w korze wzrokowej przeznaczonych na przetwarzanie informacji maleje. Z tego wynika malejąca rozdzielczość widzianego obrazu dla rosnących kątów patrzenia. Systemy komputerowe syntezujące obraz mogą zredukować próbkowanie sceny dla peryferyjnego obszaru widzenia, nie wpływając na jego odbiór przez człowieka. Inną cechą układu wzrokowego jest ograniczona *czułość na kontrast*, czyli czułość na różnicę w jasnościach i kolorach. Układ wzrokowy nie dostrzega informacji o małym kontraście. Próg dostrzegalnego kontrastu jest zależny od takich czynników jak jasność i koloru bodźca, jego częstotliwość przestrzenna oraz odległość kątowa bodźca od punktu skupienia wzroku. Jeżeli obserwator nie jest w stanie dostrzec małych kontrastów, można w procesie syntezy pominąć odpowiadające im obliczenia i w ten sposób przyspieszyć generowanie obrazu.

## Definicja problemu

Postęp technologiczny w budowie procesorów graficznych umożliwia stosowanie do syntezy obrazów złożonych modeli obliczania oświetlenia, lecz jednocześnie algorytmy renderujące grafikę wykonują wiele nadmiarowych obliczeń. Przykładem może być próbkowanie całego obrazu z jednakową częstotliwością przestrzenną, niebiorące pod uwagę faktu, że w peryferyjnych obszarach widzenia większość informacji nie będzie widziana przez człowieka. Kolejnym problemem jest określenie granic zwiększania złożoności renderingu. Jeżeli człowiek nie jest w stanie dostrzec różnicy pomiędzy obrazami wyświe-

tlonymi na wyświetlaczach 4K i 8K, to algorytm renderingu powinien ograniczyć się do generowania obrazu o rozdzielczości 4K, a zaoszczędzone zasoby procesora przeznaczyć np. na dokładniejsze obliczenie interakcji światła z materiałem obiektów. Możliwości percepcyjne układu wzrokowego człowieka stają się więc wyznacznikiem kierunków postępu technologicznego w grafice komputerowej.

W zakresie tej pracy jest zbadanie wybranych elementów układu wzrokowego człowieka, które zdaniem autora mają największy wpływ na zagadnienie przyspieszania renderingu czasu rzeczywistego. Zbadana została czułość układu wzrokowego człowieka na kontrast dla zakresu widzenia obejmującego również widzenie peryferyjne. Progi widzenia kontrastów zmierzono zarówno dla bodźca achromatycznego, jak i dla wybranych osi koloru. Model czułości na kontrast opracowany na podstawie wyników pomiarów wykorzystany został do przyspieszania syntezy obrazów w kierunkowym algorytmie metody śledzenia promieni. Zbadano również graniczną czułość układu wzrokowego człowieka na zmianę dokładności modelu. W peryferyjnym obszarze widzenia renderowany był model o zmniejszonej dokładności. Zmierzono, w jakim czasie musi nastąpić przerysowanie tego modelu na docelową dokładność, aby obserwator, przechodząc z widzenia peryferyjnego na skupienie wzroku na modelu, nie zauważył różnicy w jego jakości.

## Teza i cele pracy

Odnosząc się do opisanego wyżej problemu, w rozprawie postawiono tezę:

*Uwzględnienie progów widzialności informacji dla peryferyjnych obszarów widzenia umożliwia zmniejszenie częstotliwości próbkowania a co za tym idzie przyspieszenie syntezy obrazów.*

Za główny cel pracy postawiono pomiar charakterystyk układu wzrokowego człowieka w kontekście dostrzegalnych progów kontrastów bodźców wzrokowych, uwzględniając kierunkowość widzenia oraz kolor, a następnie opracowanie analitycznego modelu, wskazującego prawdopodobieństwo spostrzeżenia bodźca o zadanym kontraście.

Do szczegółowych celów pracy zalicza się:

- Opracowanie środowiska eksperymentalnego do pomiaru progu czułości na kontrast dla widzenia peryferyjnego. Opracowanie oraz implementacja eye trackera o częstotliwości zapewniającej precyzyjną rejestrację ruchów sakadycznych.
- Opracowanie i wykonanie eksperymentu mierzącego czułość na kontrast dla bodźców chromatycznych i achromatycznych w peryferyjnych obszarach widzenia (zob. Rozdział 2.1).

- Analiza wyników (zob. Rozdział 2.2) i opracowanie modeli (zob. Rozdział 2.3) czułości na kontrast dla układu wzrokowego człowieka w zależności od częstotliwości sygnału, typu bodźca i kąta widzenia.
- Opracowanie i wykonanie eksperymentu mierzącego opóźnienie percepcji obiektów spowodowane zmianą z widzenia peryferyjnego na centralne (zob. Rozdział 2.4).
- Aplikacja opracowanych modeli w algorytmach syntezy obrazów uwzględniającym widzenie kierunkowe (zob. Rozdział 3.1). Wykorzystanie modeli w wyświetlaczach stereoskopowych z akomodacją (zob. Rozdział 3.2).

## Metodologia badań

Praca ma charakter eksperymentalno-analityczny. Z uwagi na przedmiot badań, podstawową metodą badawczą jest prowadzenie psychofizycznych eksperymentów percepcyjnych mierzących zdolność do wykrywania bodźców wzrokowych o ustalonej charakterystyce. Uzyskane wyniki eksperymentów są poddawane analizie, mającej na celu opisanie natury wykrywania kontrastów i uzależnienie wartości progowych od częstotliwości bodźca, jego koloru oraz kierunku widzenia. Na podstawie uzyskanych wyników proponowany jest model analityczny wskazujący prawdopodobieństwo wykrycia bodźca wzrokowego, umożliwiając wykorzystanie zmierzonych cech w algorytmach komputerowej syntezy obrazów.

## Organizacja treści rozprawy

Na rozprawę składają się trzy główne rozdziały.

Rozdział 1 zawiera opis wybranych cechy układu wzrokowego. Zamieszczony został proces przetwarzania informacji zawartej w falach światła padających do gałki ocznej i jej syntezy do obrazu w świadomości obserwatora. Przedstawiona została budowa gałki ocznej oraz kory wzrokowej w kontekście przetwarzania informacji. Uwzględniony został opis rozmieszczenia komórek światłoczułych na siatkówce oka, który wskazuje na podział przestrzeni docelowego obrazu na obszary widzenia centralnego i peryferyjnego. Omówione zostało także zagadnienie ruchu gałek ocznych, opisując poszczególne typy ruchów. Rozdział zawiera również definicję terminu *kontrast* w kontekście bodźców wzrokowych oraz wyjaśnia cechę układu wzrokowego dotyczącą wykrywalnej wartości progowej kontrastu.

Rozdział 2 jest poświęcony przeprowadzonym eksperymentom percepcyjnym mającym na celu pomiar czułości układu wzrokowego na kontrast. W tym celu zostało przygotowane środowisko eksperymentalne, które umożliwiło dokonanie pomiarów. Zmie-

rzony zostały progi dostrzegalnego kontrastu dla bodźców wzrokowych achromatycznych, charakteryzujących się zmianą wyłącznie w wartości luminancji, chromatycznych o występujących w nich zmianach koloru oraz mieszanych, w których zmiana dotyczyła luminancji oraz koloru. Pomiary dokonano dla bodźców wzrokowych o pewnym zakresie jego częstotliwości przestrzennej w obszarze widzenia centralnego oraz kilku odległościach kątowych od punktu skupienia wzroku zawierających się w obszarze widzenia peryferyjnego. Na podstawie uzyskanych pomiarów w eksperymencie, zasugerowany został analityczny model wskazujący prawdopodobieństwo wykrycia określonego bodźca wzrokowego we wskazanym obszarze w polu widzenia obserwatora. W rozdziale omówiona została także kwestia dynamiki percepcji dla zmiany kierunku patrzenia z obszaru peryferyjnego na centralny oraz związany z nią eksperyment percepcyjny, gdzie sprawdzone zostały określone opóźnienia systemu syntezy obrazu w celu odnalezienia akceptowalnego interwału czasowego.

Rozdział 3 opisuje aplikacje zmierzonych cech psychofizycznych układu wzrokowego w algorytmach komputerowych kierunkowej syntezy obrazów w czasie rzeczywistym. W pierwszej aplikacji opisano sposób uzależnienia częstotliwości próbkowania sceny od odległości danego obszaru od punktu skupienia wzroku. W obszarach widzenia peryferyjnego częstotliwość ta została zredukowana, co przyspieszyło proces syntezy obrazu metodą śledzenia promieni. Druga aplikacja wykorzystuje informacje o kierunku patrzenia obserwatora do renderingu obrazów przeznaczonych do wyświetlenia na wyświetlaczu stereoskopowym z akomodacją. Opisany projekt jest wstępnym etapem budowy systemu, którego niezbędnym elementem będzie kierunkowy rendering obrazów wykorzystujący zmierzone charakterystyki czułości na kontrast.



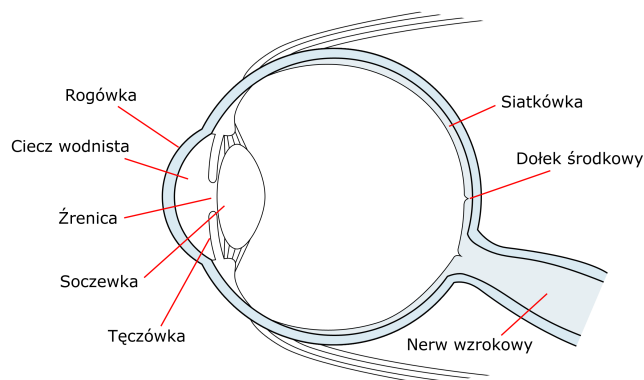
# 1. Wybrane cechy układu wzrokowego człowieka

W rozdziale przedstawiono budowę układu wzrokowego człowieka oraz proces przetwarzania informacji zawartej w falach światła wpadających do gałki ocznej. Omówiony został także podział na obszary zawarte w obserwowalnym polu widzenia oraz charakterystyka czułości na fale świetlne w tych obszarach.

## 1.1. Budowa i działanie układu wzrokowego człowieka

Układ wzrokowy człowieka (ang. *Human Visual System - HVS*) jest częścią ośrodkowego układu nerwowego odpowiedzialnego za wykrywanie i interpretację informacji, zawierającej się w świetle widzialnym, celem stworzenia w świadomości obserwatora obrazu otaczającego środowiska. Układ ten dzieli się na część mechaniczną i nerwową. Część mechaniczna składa się z pełniącej funkcję kamery gałki ocznej, której kształt przypomina kulę o średnicy około 24 mm (Rys. 1.1). Fale światła padające na gałkę oczną w pierwszej kolejności są załamywane do postaci obrazu przez rogówkę - wypukłą, półprzezroczystą kopułę stykającą się z przezroczystym płynem zwanym ciecżą wodnistą. Załamane fale przechodzą przez ten płyn i trafiają do źrenicy - kolistego otworu w tęczówce, którego średnica zmienia się w zakresie od 2 do 8 mm, kontrolując ilość przechodzącego światła. Zbyt mała lub duża jego ilość może uniemożliwić wytworzenie ostrego obrazu, dlatego HVS zmienia średnicę źrenicy zależnie od poziomu oświetlenia obserwowanej sceny. Mechanizm ten nazywany jest adaptacją. Inny mechanizm to akomodacja oka, która jest związana z soczewką - przezroczystym, elastycznym narządem w gałce ocznej, posiadającym zdolność do zmiany swojego kształtu w celu dostosowania ostrości widzenia do obiektów znajdujących się w różnej odległości. Zmiana kształtu powoduje zmianę ogniskowej soczewki i jej zdolności do załamывania światła - wypukły kształt pozwala na obserwację obiektów umiejscowionych bliżej, natomiast płaski umożliwia widzenie w oddali.

Dalsza droga fal światła wiedzie przez część neuronową HVS. Światło skupione przez soczewkę pada na siatkówkę - wielowarstwową błonę przylegającą do ściany oka. Zawiera ona wrażliwe na światło elementy, tzw. fotoreceptory, zdolne do fototransdukcji, czyli przekształcenia światła na sygnały elektryczne. Fotoreceptory, znajdujące się na siatkówce oka, dzielą się pręciki (ang. *rods*) i czopki (ang. *cones*). Pierwsze z nich,



Rysunek 1.1. Uproszczony schemat budowy gałki ocznej człowieka ze wskazanymi najważniejszymi jej elementami.

dzięki dużej wrażliwości na światło, odpowiedzialne są za widzenie skotopowe, czyli w warunkach słabego oświetlenia. Wyróżniają się także szybką reakcją na światło przez co ułatwiają wykrywanie ruchu. Pręciki nie rozróżniają długości fal absorbowanego światła, co uniemożliwia rozróżnianie barw. Czopki odpowiedzialne są za widzenie fotopowe, czyli w warunkach dobrego oświetlenia. Są one mniej wrażliwe na światło niż pręciki, przez co wymagają go więcej, żeby wygenerować impulsy elektryczne. Czopki podzielone są na trzy rodzaje - L, M oraz S, dzięki czemu umożliwiają rozróżnienie barw. Każda grupa odpowiedzialna jest za absorbowanie fal światła o innej długości, kolejno 563 nm, 534 nm i 420 nm (odpowiadająca barwom czerwonej, zielonej i niebieskiej). Na siatkówce oka znajduje się o wiele więcej pręcików (ok. 120 milionów) niż czopków (20 razy mniej - ok. 6 milionów) [90], mają też różne rozłożenie na siatkówce, co zostanie opisane w Rozdziale 1.2.

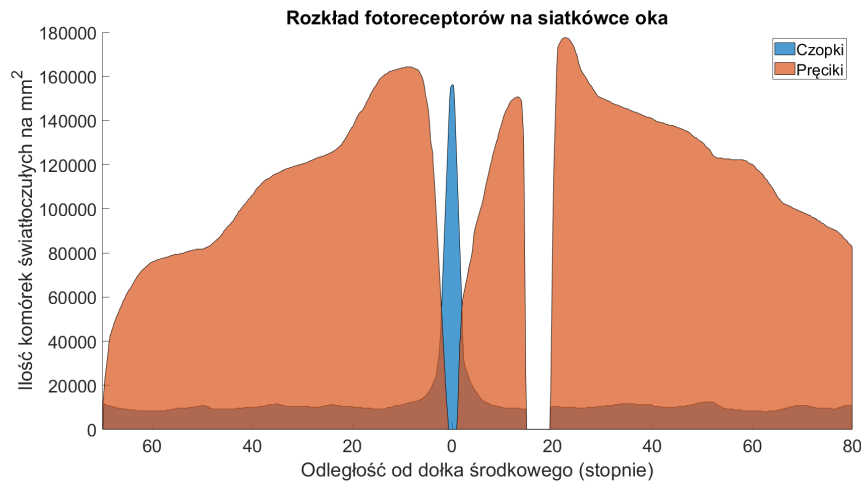
Wytworzone impulsy elektryczne przekazywane są przez nerw wzrokowy do kanału optycznego, gdzie włókna nerwowe rozpowszechniają je głównie do Ciała Kolankowatego Boczego (ang. *Lateral Geniculate Nucleus* - LGN). LGN rozprowadza impulsy do różnych regionów Kory Wzrokowej (ang. *Visual Cortex*). Najpierw informacja trafia do Pierwotnej Kory Mózgowej (ang. *Primary Visual Cortex* - V1), aby następnie była przesyłana przez hierarchię korową, która zawiera w sobie obszary V2, V3, V4, V5 i V6. W V1 ma miejsce przetwarzanie informacji zawartej w impulsach elektrycznych, celem stworzenia organizacji przestrzennej obrazu. Odbywa się to dzięki wyznaczaniu różnic w natężeniu sygnału zawartego w impulsach, przy czym operacja ta różni się w czasie od otrzymania impulsów. Na początku (około 40 ms od otrzymania informacji) poszczególne neurony V1 dzielą informacje zawartą w impulsach na małe zestawy, oddalone od siebie małym dystansem przestrzennym, poszukując lokalnych różnic w luminancji i kolorze. Wraz z czasem impulsy elektryczne łączone są w grupy oddalone od siebie większym dystansem przestrzennym, aby po czasie 100 ms wyznaczona została różnica globalna w rozumieniu obserwowanej sceny [47]. Należy zwrócić uwagę na fakt, że informacja wyjściowa z V1 nie

jest zakodowana do kategorii obrazu przestrzennego, bądź optycznego, a sygnału do różnicowego, stąd można ją porównać do operacji wykrywania krawędzi. V2 odgrywa podobną rolę co V1, z tą różnicą, że bierze pod uwagę sygnały z obu oczu i porównuje je w celu określenia odległości obiektów od obserwatora rozróżniając te, które są w położeniu aktualnej akomodacji oka. V3 natomiast ma za zadanie rozpoznanie ruchu obiektów, w tym ich kierunki i szybkości. Informacje głównie z V1, lecz także z V2, V3 i LGN są używane przez V4 w celu rozpoznawania kształtów obiektów obserwowanej sceny. Regiony V5, V6 są odpowiedzialne za analizę ruchu w obserwowanej scenie. V5 ma podobne zadanie co poprzednie regiony, lecz integruje także ruch lokalny obiektu wraz z ruchem globalnym, podczas gdy V6 analizuje sygnał pochodzący głównie z V1, ale także V5, modelując ruch obiektów względem tła. Dalsze, niewymienione tutaj regiony, są odpowiedzialne za rozpoznawanie złożonych kształtów, przedmiotów i twarzy, opierając się i modyfikując wspomnienia, jednak wszystkie opierają się na informacji przetworzonej przez omówione regiony [69].

## 1.2. Widzenie centralne i peryferyjne

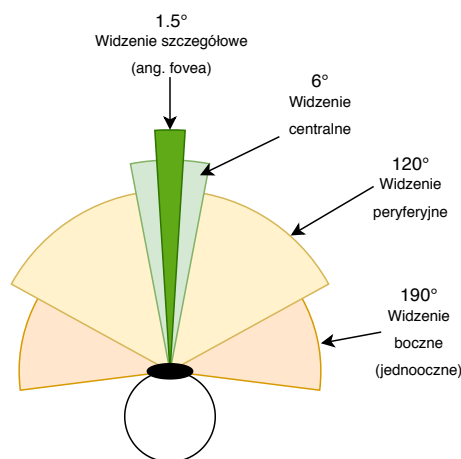
Rozmieszczenie fotoreceptorów na siatkówce oka jest niejednorodne [107]. Największe zagęszczenie czopków, fotoreceptorów odpowiedzialnych za widzenie barwne w warunkach fotopowych, występuje w dołku środkowym oka (ang. *fovea*). Jest to obszar leżący w obrębie środkowej części siatkówki, nazywanej plamką żółtą z racji zawierającego w sobie żółtego pigmentu. Rozmieszczenie to sprawia, że zdolność do postrzegalności bodźców wzrokowych zawartych w obserwowanej scenie jest najwyraźniejsza w tym obszarze, a jego wielkość przyjmuje się na około  $1^\circ$  pola widzenia [17]. Wraz ze wzrostem odległości kątowej od dołka środkowego, zwanej inaczej ekscentrycznością (ang. *eccentricity*), gęstość czopków maleje (Rys. 1.2). Do ekscentryczności około  $3^\circ$  od dołka środkowego gałki ocznej przyjmuje się za obszar widzenia centralnego (ang. *foveal vision*), zawierającego wystarczającą liczbę czopków, aby wygenerowane przez nie impulsy elektryczne zawierały szczegółowe informacje o bodźcach wzrokowych zawartych w obserwowanym otoczeniu.

Wraz z oddalaniem się od obszaru widzenia centralnego pojawia się tzw. widzenie peryferyjne (ang. *peripheral vision*). Charakteryzuje się drastycznie mniejszym zagęszczeniem czopków, które ustępują miejsca pręcikom. Ogranicza to ilość informacji o obserwowanej scenie w warunkach fotopowych, przez co zmniejsza się zdolność HVS do wykrywania szczegółowych informacji o bodźcach wzrokowych, tj. kolorów czy kształtów. Co więcej, liczba neuronów w korze wzrokowej, odpowiedzialnych za przetwarzanie informacji wygenerowanej przez fotoreceptory w obszarze peryferyjnym jest mniejsza, co pogarsza jakość wygenerowanego obrazu w obszarze peryferyjnym [95]. Obszar peryferyjny ma za zadanie modelowanie tła obserwowanego otoczenia, a także odtwarzanie informacji w ta-



Rysunek 1.2. Zagęszczenie fotoreceptorów na siatkówce oka. Obszar, w którym występuje brak fotoreceptorów, nazywany jest plamką ślepą.

ki sposób, aby możliwe było rozpoznanie znanych kształtów bez konieczności skupiania na nich wzroku. Oprócz tego, z uwagi na różnice fizjologiczne i anatomiczne w korze mózgowej oraz większe zagęszczenie pręcików, w których proces fototransdukcji zachodzi szybciej, HVS posiada zwiększoną zdolność do wykrywania zmian i ruchu w obszarze peryferyjnym [80]. Cecha ta pełni funkcję ostrzegawczą, zwracając uwagę obserwatora na potencjalne niebezpieczeństwa znajdujące się poza obszarem uwagi wzrokowej. Region widzenia peryferyjnego szacuje się na od ok.  $6^\circ$  do  $190^\circ$ - $220^\circ$  [96], przy czym, od wychylenia  $60^\circ$  względem dołka centralnego występuje widzenie jednooczne (Rys. 1.3).



Rysunek 1.3. Pole widzenia układu wzrokowego człowieka i jego podział na poszczególne regiony.

### 1.2.1. Ruchy gałek ocznych

Gdy obiekt zainteresowania obserwatora znajduje się w obszarze peryferyjnym pola widzenia, naturalną reakcją jest zmiana kierunku patrzenia w celu przeniesienia pozycji obiektu do obszaru widzenia centralnego. W tym celu wykorzystywane są mięśnie okomotoryczne, które zmieniają orientację gałki ocznej. Gdy obiekt jest zbyt duży lub jest zbyt blisko, aby zmieścił się w obszarze centralnym, kierunek patrzenia przenoszony jest wielokrotnie pomiędzy regionami zainteresowania. Ten mechanizm ma na celu gromadzenie informacji o detalach i integracji danych w jeden spójny obraz [78].

#### Fiksacje

Fiksacją gałek nazywamy sytuację, gdy na siatkówkę oka rzutowana jest informacja o obiekcie stacjonarnym względem obserwatora. Możliwa jest wtedy projekcja obiektu w stałym położeniu siatkówki oka, która maksymalizuje zdolność do postrzegania zawartych w nim detali. Łączny czas trwania fiksacji to 90% czasu obserwowania statycznej sceny [24]. Może być ona utożsamiana ze stabilizacją orientacji gałki ocznej, lecz w trakcie trwania fiksacji (150-600 ms) występują mimowolne ruchy gałki ocznej, zwane drzeniami (ang. *tremor*), dryfami (ang. *drift*) oraz ruchami mikrosakadycznymi (ang. *microsaccades*). Dwa pierwsze wymienione ruchy tłumaczone są jako próba rekompensowania szumu, powstałego podczas ruchu gałki ocznej, celem ustabilizowania jej do żądanej orientacji. Pomiar amplitud ruchów drzenia i dryfów wykazują wartości poniżej  $0.01^\circ$ , co w przybliżeniu jest równe wielkości jednej komórki światłoczułej. Nie są one zauważalne przez człowieka z uwagi na ich częstotliwość rzędu 100-150 Hz [64]. Aby obiekt był wyraźnie widoczny jego prędkość przemieszczania się musi być mniejsza niż 5% [51].

Występowanie ruchów mikrosakadycznych zapobiega powidokom, które są wynikiem adaptacji fotoreceptorów do nieruchomego obrazu. Ich brak sprawiałby, że obserwowany przez dłuższy czas obraz stałby się niewidoczny [22]. Innym przeznaczeniem ruchów mikrosakadycznych jest korekta położenia rzutu obrazu na siatkówce, który w wyniku drzeń i dryfów oddalił się od dołka środkowego gałki ocznej [16]. Ruchy mikrosakadyczne przenoszą obraz o odległość kilkuset fotoreceptorów, ok.  $\frac{1}{3}^\circ$  [51], a czas ich trwania wynosi maksymalnie ok. 25 ms [64].

#### Sakady i płynny poscig

Sakadą (ang. *saccade*) nazywamy ruch gałek ocznych występujący podczas zmiany jej orientacji [41]. Ich zasięg jest w zakresie możliwych kierunków widzenia, a prędkość może osiągnąć wartość rzędu 500% i jest zależna od wielkości zmian kierunku patrzenia [51]. Czas inicjacji ruchów sakadycznych trwa 200 ms, podczas gdy same trwają od 20 do 200 ms. W sytuacji nagłego pojawienia się bodźców wzrokowych w obszarze peryferyjnym, czas inicjacji może zostać skrócony o połowę [27]. W trakcie trwania ruchów sakadycznych

występuje chwilowa ślepotą percepcji wzrokowej spowodowana nieodbieraniem żadnych bodźców wzrokowych [35]. Interpretowanie informacji o płynnie poruszającym się obiekcie nie odbywa się za pomocą sakad, a dzięki adaptacji aparatu motoryczno-wizyjnego do przemieszczania się obiektu, przez co jego pozycja rzutowanego obrazu na siatkówce się nie zmienia. Występuje wtedy wrażenie poruszającego się tła, które w rzeczywistości jest statyczne. W sytuacji, gdy szybkość przemieszczających się obiektów jest mniejsza niż  $5^\circ$ , możliwe jest wystąpienie zjawiska fiksacji.

Jeśli ruchome obiekty sceny charakteryzują się odmienną luminancją, bądź występuje zjawisko migotania (ang. *flickering*) - dynamicznej zmiany jasności obiektu, obiekty te mogą być obserwowane dzięki ruchowi gałek ocznych zwanemu płynnym pościgiem (ang. *smooth pursuit*) [23]. Ruch ten podzielony jest na etapy przypominające fiksacje gałek ocznych rozgraniczone ruchami sakadycznymi, w trakcie których obiekty rzutowane są na część siatkówki zapewniającą najlepsze widzenie w danym momencie [100]. Po intencji śledzenia obiektu następuje opóźnienie płynnego pościgu, które dla obiektów o szybkości powyżej  $5\%$  wynosi 100 ms i wzrasta wraz ze zmniejszeniem prędkości obiektu. Gdy prędkość przekracza  $30\%$  lub gdy bodziec mniej wyróżnia się na tle innych elementów obserwowanej sceny, następują krótkie sekwencje płynnego pościgu i fiksacji, rozdzielone sakadami, zmniejszając tym samym zdolności spostrzegania szczegółów przez HVS [93].

## Konwergencja

Konwergencja (ang. *vergence*) jest ruchem oczu związanym z percepcją głębi i akomodacją oka. Polega na skorelowaniu ruchu obu gałek ocznych, w celu zredukowania zjawiska podwójnego widzenia zwanego diplopią. W zależności od odległości przedmiotu, na którym skupiona jest uwaga obserwatora gałki oczne muszą się obrócić, aby wektory kierunków patrzenia były skierowane ku temu obiektowi. Gałki oczne obracają się ku sobie, gdy w danym momencie obiekt zainteresowania jest bliżej niż akomodacja oka, w przeciwnym razie oddalają się od siebie. Prędkość konwergencji wynosi około  $25\%$ .

## Odruch przedsionkowo-oczny

Odruch przedsionkowo-oczny (ang. *vestibulo-ocular reflex*) pozwala na stabilizację obrazu rzutowanego na siatkówkę podczas ruchu głowy. Polega on na zmianie orientacji gałki ocznej w kierunku przeciwnym do obrotu i przesunięcia głowy, dzięki czemu obiekt zainteresowania uwagi wzrokowej obserwatora nie zmienia swojej pozycji w perspektywie jego pola widzenia. Czas odruchu przedsionkowo-ocznego wynosi mniej niż 10 ms, dzięki czemu uzyskujemy komfortowe widzenie po wystąpieniu ruchu głowy [6]. W przeciwnym wypadku wygenerowane obrazy sprawiałyby wrażenie fotografii wykonanej drżącą ręką.

### 1.3. Charakterystyki czułości na kontrast

Impulsy, wygenerowane dzięki fototransdukcji zachodzącej w fotoreceptorach, są przetwarzane w celu wytworzenia obrazu otaczającego środowiska w świadomości obserwatora. Proces przetwarzania jest podzielony na etapy, w których znaczącą rolę odgrywa wyszukiwanie różnic w luminancji i barwie impulsów znajdujących się w regionach oddalonych o różnej odległości, co przypomina operacje wykrywania krawędzi. Ze względu na to porównanie, ważne jest omówienie terminu Kontrastu (ang. *Contrast*), oznaczającego różnicę w charakterystyce danego obiektu względem innych występujących w tle, która sprawia, że jest on rozpoznawalny.

#### Definicja kontrastu

Istnieje wiele definicji kontrastu, z czego każda jest zależna od sytuacji, w której jest stosowana. Wielokrotnie komplikuje to rozwiązywanie problemów, związanych z cechami charakterystycznymi obiektów oraz utrudnia porównanie wyników, opublikowanych w pracach naukowych. Ogólną ideę kontrastu symbolizuje poniższa proporcja:

$$C = \frac{\Delta I}{\bar{I}},$$

gdzie  $\Delta I$  oznacza amplitudę zmian wartości w danej charakterystyce obiektu, natomiast  $\bar{I}$  jest wartością tej charakterystyki należącą do tła. Gdy wartość kontrastu  $C$  jest równa zero, co następuje dla braku zmian w charakterystyce obiektu, z perspektywy odbiorcy sygnał nie istnieje, więc nie ma on swojej reprezentacji np. na obrazie. Należy zwrócić uwagę na fakt, iż istotność amplitudy jest zależna od wartości tła. Jeśli średnia luminancja jest wysoka, mała amplituda nie wpłynie na końcową wartość kontrastu, podczas gdy ta sama różnica może powodować duży kontrast dla niższej wartości średniej luminancji, zgodnie z prawem Webera-Fechnera [18].

Dla obiektu, będącego bodźcem wzrokowym achromatycznym, za wartość jego charakterystyki przyjmuje się luminancję, dzięki czemu uzyskujemy formułę Webera [26]:

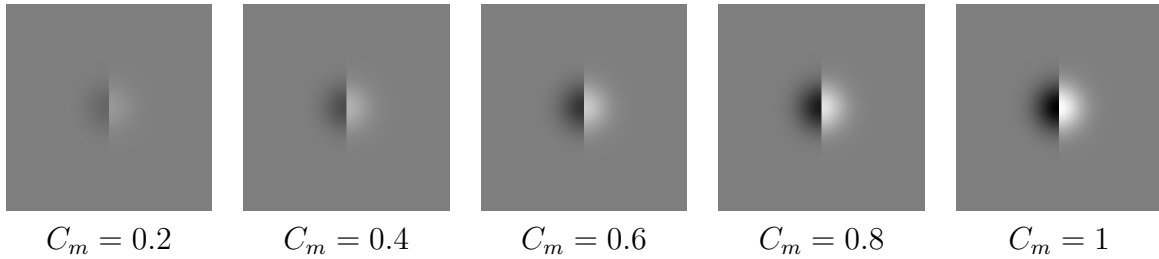
$$C_w = \frac{\Delta L}{L_b},$$

gdzie  $\Delta L$  to amplituda zmian luminancji bodźca wzrokowego, natomiast  $L_b$  to luminancja tła. Formuła ta nazywana jest także frakcją Webera. Kontrast Webera jest najczęściej używany, gdy na jednolitym tle występują małe obiekty, wyróżniające się inną wartością luminancji [83].

Drugim sposobem określenia kontrastu achromatycznego bodźca wzrokowego jest zastosowanie kontrastu Michelsona, inaczej nazywanego widzialnością (ang. *visibility*) [67], określanego formułą:

$$C_m = \frac{L_{max} - L_{min}}{L_{max} + L_{min}},$$

gdzie  $L_{min}$  i  $L_{max}$  oznaczają minimalną i maksymalną luminancję bodźca wzrokowego. Kontrast Michelsona stosowany jest, gdy bodziec wzrokowy składa się z obszarów, które w równym stopniu zawierają niską i wysoką wartość luminancji, jak sygnał sinusoidalny lub funkcja progowa (Rys. 1.4).



Rysunek 1.4. Przykład achromatycznego bodźca wzrokowego w postaci funkcji progowej (ang. *step function*) o różnych wartościach kontrastu Michelsona  $C_m$ .

Różnica między dwoma przedstawionymi definicjami kontrastu jest zauważalna, gdy przedstawimy kontrast Michelsona w podobnej formie co kontrast Webera:

$$C_m = \frac{\Delta L}{L + \Delta L},$$

gdzie  $\Delta L = \frac{L_{max} - L_{min}}{2}$ , natomiast  $L = L_{min}$ . Wartości przedstawionych definicji kontrastu nie pokrywają się, a także nie dzielą ze sobą zakresu przyjmowanych wartości. Kontrast Michelsona wyraża się w wartościach z przedziału  $[0, 1]$ , podczas gdy wartości kontrastu Webera zawierają się w zakresie  $[-1, \infty]$ .

Dla chromatycznych bodźców wzrokowych, w których zamiast zmian w wartościach luminancji występuje zmiana barwy, stosowany się kontrast Średniej Kwadratowej (ang. *Root Mean Square*) opisanej formułą:

$$C_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{\Delta I_i}{\bar{I}_i} \right)^2},$$

gdzie  $N$  jest to liczba kanałów kolorów, które bierzemy pod uwagę,  $\Delta I_i$  to amplituda zmian wartości koloru zawartego w  $i$ -tym kanale bodźca, natomiast  $\bar{I}_i$  oznacza kolor tła  $i$ -tym kanale. Formuła ta jest rozszerzeniem kontrastu Webera, który uzależnia swoją wartość od wielu cech, w odróżnieniu od standardowej formuły, która opierała się wyłącznie na luminancji.



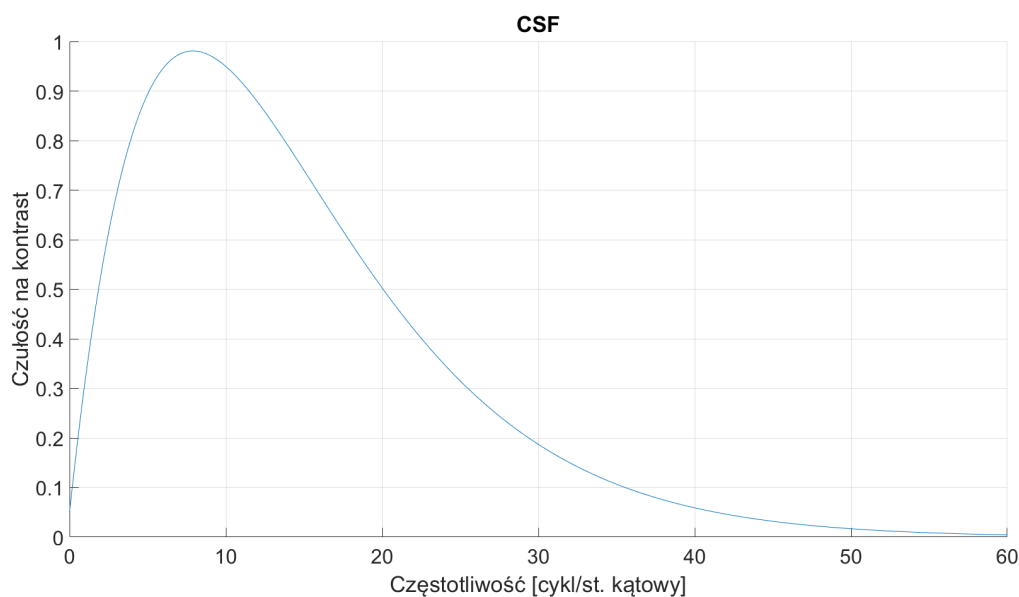
### Funkcja czułość na kontrast

Obiekty w obserwowanej scenie, można rozłożyć na bodźce wzrokowe, których suma tworzy reprezentację tego obiektu w obrazie wygenerowanym przez HVS. Układ wzrokowy człowieka posiada ograniczoną zdolność do wykrywania bodźców wzrokowych rzutowanych na siatkówkę oka. Jeśli wartość kontrastu tego bodźca jest zbyt niska, nie będzie on widoczny dla obserwatora. Podczas gdy kontrast wyrażony formułą Michelsona jest równy zero, jasny staje się fakt, iż bodziec nie jest widoczny ze względu na brak swojej reprezentacji z perspektywy obserwatora. Istnieją jednak pewne wartości progowe kontrastu, które są większe niż zero, ale poniżej których bodziec wzrokowy staje się niemożliwy do spostrzeżenia przez HVS. Wartość progowa kontrastu  $C_{thr}$  może zostać wykorzystana do wyznaczenia czułości na kontrast (ang. *contrast sensitivity*). Wartość czułości na kontrast  $S$  definiuje się jak odwrotność proggu kontrastu:

$$S = \frac{1}{C_{thr}}$$

Dobrze znaną charakterystyką w obszarze badań nad grafiką komputerową jest funkcja czułości na kontrast (ang. *contrast sensitivity function* - *CSF*), która wskazuje zależność między zdolnością HVS do wykrycia bodźca wzrokowego, a jego częstotliwością wyrażaną w cyklach na stopień kątowy widzenia (ang. *cycles per degree* - *cpd*) w perspektywie obserwatora sceny. Charakterystyka ta była mierzona w licznych badaniach naukowych nad zdolnościami ludzkiej percepcji [62], można więc powiedzieć że istnieje powszechnie zaakceptowana formuła tej zależności (Rys. 1.5). Wskazuje się, że HVS posiada największą wartość czułości na kontrast dla bodźca o częstotliwości od 2 do 4 *cpd* [7], zmniejszając zdolność do wykrywania bodźców o częstotliwości oddalonej od tego zakresu. Zanik czułości dla wyższych częstotliwości wynika z ograniczeń aparatu widzenia człowieka. W tej sytuacji obserwowane bodźce wzrokowe zawierają się w zbyt małym obszarze obrazu rzutowanego na siatkówce by możliwe było ich wykrycie. Całkowity zanik zdolności do postrzegania bodźców wzrokowych występuje dla częstotliwości równej 60 *cpd* uznawaną za rozdzielczość HVS. Zmniejszona zdolność do wykrywania bodźców w niskich częstotliwościach tłumaczona jest uznaniem takiej informacji przez HVS za tło, odrzucając częściowo informację o nim w procesie przetwarzania impulsów elektrycznych w korze wzrokowej, aby informacja zawierająca się w wyższych częstotliwościach była bardziej eksponowana [106] w wygenerowanym obrazie.

Charakterystyka czułości na kontrast HVS różni się dla bodźców chromatycznych. Zmienia się także pod wpływem warunków, w jakich jest on obserwowany, tj. luminancja tła i akomodacja oka. Czułość na kontrast zależna jest także od regionu w polu widzenia, w jakim bodziec wzrokowy rzutowany jest na siatkówkę. Mierzenie czułości na kontrast ma wiele zastosowań, np. w medycynie, gdzie pomiar charakterystyki CSF pacjentów, dający



Rysunek 1.5. Przykład funkcji czułości na kontrast bodźców achromatycznych. Czułość wyrażona jest jako odwrotność kontrastu Michelsona.

inne wyniki od przyjętej normy, pozwala wykryć niektóre choroby [20]. Pomiar CSF ma także szerokie zastosowanie w grafice komputerowej, co będzie tematem Rozdziału 3 tej pracy.

## Podsumowanie rozdziału

Charakterystyka HVS czułości na kontrast bodźców wzrokowych sprawia, że obserwator nie jest w stanie wykryć wszystkich detali obiektów w postrzeganym otoczeniu. Próg dostrzegalnego kontrastu jest zależny od częstotliwości i koloru bodźca wzrokowego oraz obszaru, na który jest on rzutowany na siatkówkę oka. Algorytmy komputerowe, syntezujące obrazy wirtualnych scen zawierające dane o kierunku patrzenia obserwatora, mogą wyznaczać obszary widzenia centralnego i peryferyjnego na docelowym obrazie. Umożliwia to uzależnienie częstotliwości próbkowania sceny od odległości poszczególnych fragmentów od punktu skupienia wzroku obserwatora, co pozwala zmniejszyć liczbę próbek bez pogorszenia odbioru percepcyjnego wynikowego obrazu, przyspieszając czas jego syntezy. Najpierw należy jednak wyznaczyć model czułości na kontrast HVS, uzależniający progi widzialnego kontrastu od obszaru padania bodźca wzrokowego na siatkówkę, co wymaga pomiarów czułości na kontrast dla wyznaczonych parametrów bodźca wzrokowego, które zostały wykonane drogą eksperymentu percepcyjnego i opisane w Rozdziale 2 tej pracy.

## 2. Model czułości na kontrast dla peryferyjnego obszaru widzenia

W rozdziale omówiono sposób modelowania charakterystyki czułości na kontrast bodźców wzrokowych w peryferyjnych obszarach widzenia oraz pomiar szybkości percepcji wzrokowej podczas zmiany kierunku patrzenia.

Układ wzrokowy człowieka ma ograniczone możliwości do dostrzegania bodźców wzrokowych, zawartych w obserwowanej scenie. Zdolność ta utożsamiana jest z czułością na kontrast bodźców, której pomiar był celem opracowanego eksperymentu percepcyjnego (Rozdz. 2.1). Uczestnikom eksperymentu zaprezentowane zostały bodźce wzrokowe (Rozdz. 2.1.1) o ustalonej charakterystyce, w celu ustalenia progu kontrastu umożliwiającego dostrzeżenie go przez człowieka. W procedurze eksperymentu została wykorzystana funkcja psychometryczna, która wskazywała kolejne wartości kontrastu (Rozdz. 2.1.2). Aby bodźce były w odpowiedni sposób przedstawione obserwatorowi, konieczna była wcześniejsza kalibracja sprzętu (Rozdz. 2.1.3). Na podstawie wyników eksperymentu (Rozdz. 2.2) oszacowana została charakterystyka HVS czułości na kontrast dla achromatycznych, chromatycznych oraz mieszanych bodźców wzrokowych o wybranych częstotliwościach w peryferyjnych obszarach widzenia. Na podstawie zebranych danych zbudowany został analityczny model, wskazujący prawdopodobieństwo wykrycia bodźca wzrokowego przez HVS (Rozdz. 2.3). Treść wskazanych części rozdziału opiera się na badaniach zrealizowanych w ramach autorskich publikacji [W2, W3].

Końcowa część rozdziału dotyczy zagadnienia dynamiki ludzkiej percepcji podczas zmiany kierunku patrzenia. Gdy obiekt zainteresowania obserwatora znajduje się w obszarze widzenia peryferyjnego, naturalną reakcją jest zmiana kierunku patrzenia w celu zmiany położenia rzutu obiektu na siatkówkę gałki ocznej na obszar widzenia centralnego. Gdy system renderingu opiera się o kierunek patrzenia, opóźnienie wyświetlania kolejnych klatek względem zmiany punktu skupienia wzroku nie może być zbyt duże (Rozdz. 2.4). W celu zmierzenia tych progów stworzony został system renderingu korzystający z techniki Poziomów Szczegółowości, który posłużył do przeprowadzenia eksperymentu percepcyjnego (Rozdz. 2.4.1). Testując różne opóźnienia systemu renderingu, zostały zebrane wyniki (Rozdz. 2.4.2), które pozwoliły wysunąć wnioski wskazujące akceptowalne opóźnienie systemów renderingu opartych o kierunkowość widzenia (Rozdz. 2.4.3). Wyniki tej części pracy zostały opisane w publikacji naukowej [W1].

## 2.1. Eksperymentalny pomiar czułości na kontrast

Funkcja czułości na kontrast (CSF) wyraża zdolność HVS do wykrycia achromatycznych bodźców wzrokowych zawartych w obserwowanej scenie. Próg dostrzegalnego kontrastu bodźca uzależniany jest przez CSF od jego częstotliwości, ale jest on zależny również od luminancji adaptacji gałki ocznej, odległości bodźca od aktualnej akomodacji oka, bądź obszaru siatkówki, na który obraz bodźca jest rzutowany przez rogówkę i soczewkę. Czułość na kontrast różni się także dla bodźców chromatycznych. Progi kontrastów były mierzone w poprzednich badaniach nad percepcją człowieka [70, 85, 102, 109], z których należy zwrócić uwagę na pracę Bartena [7], gdzie przedstawiony został przegląd i porównanie pomiarów wielu autorów. Współcześnie Kim [43] zmierzył progi kontrastu dla zakresu luminancji tła od 0.02 do 200  $\text{cd}/\text{m}^2$ , bodźców wzrokowych o częstotliwościach od 0.125 do 16  $\text{cpd}$  i kierunkach koloru zgodnych z zestawem danych ColorFest [109], przy czym pomiar ten odbył się dla obszarów widzenia centralnego.

Pomiar funkcji czułości na kontrast w obszarach widzenia peryferyjnego (ang. *gaze dependent contrast sensitivity function - GDSCF*) odbywa się dzięki wyświetlaniu bodźca wzrokowego wyświetlonego w zadanej odległości kątowej od dołka środkowego gałki ocznej obserwatora przez ustalony czas (nazywamy go wtedy migającym bodźcem). Taki zabieg ma na celu ograniczenie możliwości obserwatora do zmiany kierunku patrzenia w trakcie wyświetlania bodźca, i tym samym błędny pomiar czułości na kontrast dla widzenia centralnego zamiast peryferyjnego. Robson i Graham [89] do pomiarów GDSCF wyświetlali obserwatorom bodziec wzrokowy w postaci czterech cykli poziomej funkcji progowej przez 100 milisekund. Cannon [11] zastosował pionową siatkę sinusoidalną po prawej stronie od punktu skupienia wzroku obserwatora na 2 sekundy, w których zawierało się 350 milisekund na zapoczątkowanie i wygaszanie bodźca wzrokowego. Thomas [97] używał siatki wyświetlanej przez 1 sekundę, zarówno z nagłym rozpoczęciem i zakończeniem wyświetlaniem bodźca, jak i z jego stopniowym wzbudzeniem i wygaszaniem. Pointer and Hess [86] przedstawiał zorientowaną poziomo sinusoidę wewnątrz obwiedni Gaussa. Bodziec był wyświetlony przez 250 milisekund z Gaussowskim rozkładem czasowym wzbudzenia i wygaszania.

Głównym powodem pomiarów progów wykrywalnych kontrastów przez HVS chromatycznych bodźców wzrokowych była chęć zbadania mechanizmów, jakie zachodzą osobno w czopkach i pręcikach [5, 9, 30, 74, 75, 77]. Mullen [73] przeprowadził pomiary GDSCF dla bodźców chromatycznych, aby odpowiedzieć na pytanie, czy istnieją dowody, że obszar widzenia centralnego jest bardziej dostosowany do wykrywania barw niż kontrastu luminancji. Spadek czułości na kontrast w całym polu widzenia dla czerwono-zielonych bodźców wzrokowych został porównany z achromatycznymi dla częstotliwości 2  $\text{cpd}$  na dołku środkowym oka i odległości 10° oraz 18° od niego. Anderson [5] zmierzył GDSCF w zakresie ekscentryczności od 0° do 55° wykorzystując czerwono-zieloną sinusoidalną siatkę.

Na podstawie wyników tych badań wysunięto wniosek, że wraz z odległością kątową od dołka centralnego gałki ocznej, czułość na bodziec chromatyczny spada bardziej niż w przypadku achromatycznego, co sugeruje istnienie dodatkowych postreceptorowych ograniczeń wykrywania kolorów w obszarze widzenia peryferyjnego. Nagy i Wolf [75] zmierzili progi dyskryminacji przez HVS bodźców o czterech kolorach z czerwono-zielonej części widzialnego spektrum światła w trzech miejscach na tzw. siatkówce nosowej (w odległościach  $5^\circ$ ,  $20^\circ$  oraz  $40^\circ$ ). Nagy i Doyal [74] zmierzili progi rozróżniania kolorów dla czerwono-zielonych bodźców wzrokowych o zmiennym rozmiarze dla odległości kątowych  $10^\circ$  oraz  $25^\circ$ . Zastosowano pięć okrągłych bodźców wzrokowych o różnych średnicach. Dla odległości  $10^\circ$  wynosiły one  $0.5^\circ$ ,  $1^\circ$ ,  $2^\circ$ ,  $4^\circ$  i  $8^\circ$ , natomiast dla  $25^\circ$  były równe  $1^\circ$ ,  $2^\circ$ ,  $4^\circ$ ,  $8^\circ$  i  $16^\circ$ . Luminancje wyświetlonych bodźców utrzymywano na stałym poziomie  $17 \text{ cd/m}^2$ . Osiągnięte rezultaty wskazują na podobieństwo progów wykrywania kontrastu, gdy pole bodźca wyświetlonego w obszarze peryferyjnym jest duże oraz gdy bodziec w obszarze centralnym ma małe pole. Mullen, w badaniach z 2002 roku [71], zmierzyl czułość na kontrast dla sygnałów sinusoidalnych w dwóch układach kolorów - czerwono-zielonym oraz żółto-niebieskim. Bodźce były stopniowo wyświetlane i wygaszane w czasie oraz tłumione wraz z odległością od ich środka. Progi wykrywania ich kontrastu zostały zmierzone dla odległości od  $0^\circ$  do  $25^\circ$  od dołka środkowego gałki ocznej. Wyniki wskazały, że utrata czułości na kontrast bodźca żółtego-niebieskiego jest mniejsza, podobna do bodźca achromatycznego, w przeciwieństwie do gwałtownego spadku dla bodźca czerwono-zielonego. W badaniach Mullena z 2005 [72] pomiary wskazały na spadek czułości na kontrast dla bodźca czerwono-zielonego do odległości ok.  $20\text{-}30^\circ$ , po czym zdolność do jego wykrywania zanika. Hansen [33] zmierzyl czułość na bodźce chromatyczne dla dalekich odległości kątowych od dołka środkowego gałki ocznej, sięgających  $50^\circ$  i wykrył, że mimo dużego spadku wykrywalności przy takich odległościach nadal jest możliwe ich spostrzeżenie. Diez [21] zmierzyl czułości na bodźce wzrokowe w 21 miejscach pola widzenia. Vatson i Crognale [101] w swoich badaniach sugerują, że wzrost rozmiaru bodźca chromatycznego w obszarze peryferyjnym nie zmienia się, gdy bodziec odbierany jest przez czopki typu L i M, w przeciwieństwie do czopków typu S, dla których postrzegalny próg kontrastu się zwiększa.

Sposób naszego pomiaru GDSCF jest inspirowany badaniami Peli [84], w których za próg kontrastu została uznana wartość, pozwalająca na rozróżnienie orientacji poziomej i pionowej bodźca wzrokowego w postaci sygnału sinusoidalnego z wygaszanego filtrem Gaussa (Wzór Gabora, Rozdz. 2.1.1). Te pomiary odbywały się w obszarze widzenia centralnego oraz peryferyjnego w odległościach  $2.5^\circ$ ,  $5.1^\circ$ ,  $10.3^\circ$  oraz  $22.8^\circ$ . Dla każdej odległości progi kontrastu zostały zmierzone przy częstotliwości bodźców równych 1, 2, 4, 8 oraz 16 *cpd*. Wyświetlony bodziec wzrokowy składa się z 4 cykli, przy czym dobrze

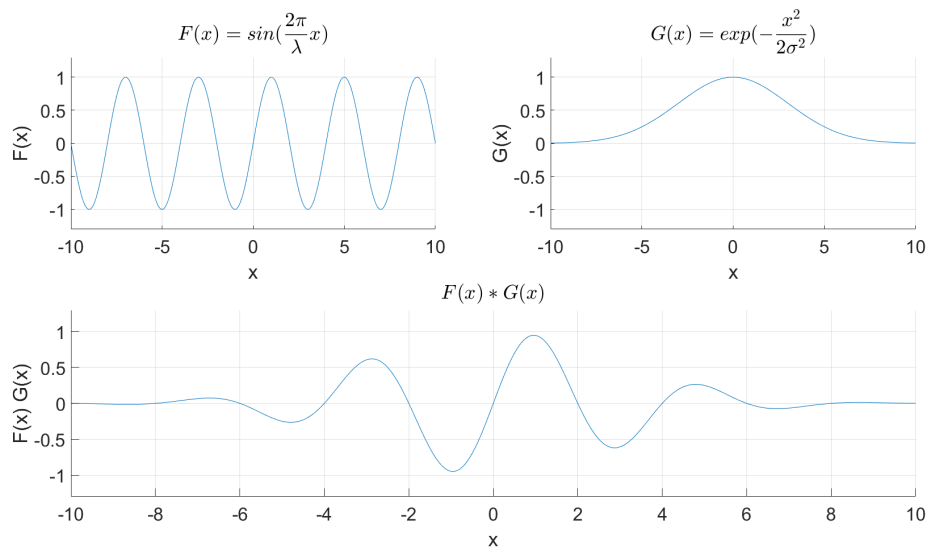
widoczne były dwa. Luminancja tła wynosiła  $37.5 \text{ cd/m}^2$ , a bodziec był wyświetlony przez 0,5 sekundy, z nagłym wyświetleniem i zniknięciem.

W naszych badaniach zajęliśmy się pomiarami progów wykrywalnych kontrastów siedmiu bodźców wzrokowych. Pierwszy z nich jest bodźcem achromatycznym, składającym się jedynie ze zmian wartości luminancji. Kolejne dwa są bodźcami chromatycznymi o kierunkach kolorów zgodnymi z zestawem danych ColorFest, w którym znajdują się kierunki czerwono-zielone i żółto-niebieskie. Pozostałe cztery bodźce, typu mieszanego łączą właściwości achromatycznych i chromatycznych, są następujących kolorów: różowo-zielono-cyjanowy, zielono-jasnoróżowy, ciemnoczerwono-cyjanowy i zielono żółty-niebiesko-fioletowy. Pomiar czułości na kontrast odbył się dla obszaru widzenia centralnego oraz peryferyjnego w odległościach  $5^\circ$ ,  $10^\circ$ ,  $15^\circ$ ,  $20^\circ$  oraz  $27^\circ$  od dołka środkowego oka. Dla każdej odległości kątowej pomiar czułości na kontrast bodźców achromatycznych odbył się dla ich częstotliwości równych 0.125, 0.25, 0.5, 1, 2, 4, 8 i 16 *cpd*, natomiast dla chromatycznych i mieszanych odbył się dla 2 *cpd*. Średnica bodźców wzrokowych była zależna od kontrastu i wynosiła maksymalnie  $8^\circ$ . Nasze badania cechują się zastosowaniem metodologii, która w odróżnieniu od wyświetlania bodźca wzrokowego przez określony czas, przedstawia je w sposób ciągły. Kontrola nad poprawnym obszarem jego padania na siatkówkę gałki ocznej sprawowana jest przez wykorzystanie eye trackera śledzącego kierunek patrzenia obserwatora. Luminancja tła podczas pomiarów wynosiła  $50 \text{ cd/m}^2$  dla bodźców achromatycznych i mieszanych, natomiast dla chromatycznych  $23 \text{ cd/m}^2$  oraz  $7 \text{ cd/m}^2$ , odpowiednio dla czerwono-zielonego i żółto-niebieskiego.

### 2.1.1. Bodźce wzrokowe stosowane w eksperymencie

Dowolny sygnał wzrokowy można rozłożyć na szereg sygnałów sinusoidalnych (analiza fourierowska), co powoduje, że siatka sinusoidalna może być stosowana jako bodziec wzrokowy w badaniach nad charakterystyką HVS pod względem czułości na kontrast. Podstawowy sygnał sinusoidalny ma charakter nieskończony, jej wartość waha się od wokół punktu równowagi o zadanej stałej amplitudzie, niezależnie od odległości od początku sygnału. Ta natura uniemożliwia wyświetlenie bodźca wyłącznie w obszarze peryferyjnym, dlatego w celu ograniczenia obszaru sygnał sinusoidalny multiplikowany jest przez filtr Gaussa [84, 86], dzięki czemu w rezultacie otrzymuje się Wzór Gabora (ang. *Gabor Pattern*). Filtrowanie sygnału dobrze obrazowane jest przez przekrój poprzeczny funkcji (Rys. 2.1), który wyraźnie wskazuje zanik amplitudy wraz z oddalaniem się od jej centrum.

Wzór Gabora w grafice komputerowej stosowany jest jako filtr liniowy wykorzystywany na obrazach celem ich analizy. W dziedzinie cyfrowego przetwarzania obrazu posłużył on do ekstrakcji określonych cech, co umożliwiło identyfikację słowa pisanego w dokumencie wielojęzycznym [81], oraz wyodrębnianie ze złożonych obrazów obszarów zawierających tylko tekst [88], czy też do rozpoznawania wyrazu twarzy [58]. Filtr Gabora



Rysunek 2.1. Przykład sinusoidalnej funkcji  $F(x)$  o długości fali ( $\lambda$ ) równej 4 i amplitudzie 1 oraz filtra gaussa  $G(x)$  o odchyleniu standardowym ( $\sigma$ ) równym 3. Iloczyn dwóch funkcji daje w wyniku Wzór Gabora.

ma również szerokie zastosowanie do analizy wzorców, w medycynie posłużył do badania rozkładu kierunkowości beleczki kostnej wewnątrz tkanki gąbczastej w kręgosłupie [29]. Wzór Gabora opisany jest wzorem:

$$\begin{aligned}
 F(x, y, ch_{rgb}) &= C \cdot 0.5 \cdot \text{wave}(x') \cdot \text{gaussian}(x', y') \cdot ch_{rgb} + ch_b, \\
 \text{wave}(x) &= \sin\left(\frac{2\pi}{\lambda}x + \psi\right), \\
 \text{gaussian}(x, y) &= \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right), \\
 x' &= x \cdot \cos(\theta) + y \cdot \sin(\theta), \\
 y' &= -x \cdot \sin(\theta) + y \cdot \cos(\theta),
 \end{aligned}$$

gdzie  $(x, y)$  jest położeniem piksela względem środka generowanego wzoru (punkt  $(0, 0)$  jest początkiem sygnału). Funkcja  $\text{wave}(x)$  jest sygnałem sinusoidalnym o częstotliwości równej odwrotności  $\lambda$  i przesunięciu fazowym  $\psi$ , natomiast funkcja  $\text{gaussian}(x, y)$  generuje obwiednie Gaussa o odchyleniu standardowym  $\sigma$  oraz parametrze  $\gamma$  odpowiedzialnym za wskazanie poziomu symetryczności kształtu filtra. Parametry  $ch_{rgb}$  oznaczają kolor bodźca,  $ch_b$  jest kolorem tła, a  $\theta$  zmienia orientację wygenerowanego wzoru, który domyślnie jest pionowy.  $C$  oznacza kontrast wyrażony jako kontrast Michelsona.

Do wygenerowania bodźców wzrokowych, celem wykorzystania ich do pomiaru progów kontrastu, użyte zostały procedury zawarte w Psychtoolbox, zestawie narzędzi do środowiska Matlab [10]. Dzięki wykorzystaniu tekstury proceduralnej, programów cieniujących (ang. *shaderów*) i procesora graficznego, czas generowania bodźca wzrokowego

został zminimalizowany, co pozwoliło na efektywne przeprowadzenie eksperymentu percepcyjnego. Procedury *CreateProceduralGabor()* i *Screen('DrawTexture', ...)* zawarte w zestawie narzędzi pozwalają na dostosowanie parametrów wzoru Gabora do potrzeb eksperymentu, spośród których przesunięcie fazowe sinusoidy  $\psi$  zostało ustawione na  $0^\circ$ , a symetryczność filtra gaussowskiego  $\gamma$  równe 1 w celu zachowania owalnego kształtu bodźca wzrokowego. Odchylenie standardowe obwiedni Gaussa  $\sigma$  zostało ustawione na 70, aby wygenerowany bodziec był zauważalny przy niskich częstotliwościach. Parametr  $\theta$  był ustawiany na 0, w celu ustawienia domyślnej pionowej orientacji bodźca, bądź na wartość 180, gdy była konieczność wyświetlenia wzoru Gabora w orientacji poziomej. Parametr  $C$  przekazywany do wzoru pozwalał ustawić kontrast Michelsona bodźca wzrokowego, który był ustawiany na różne wartości z przedziału  $[0, 1]$ , zależnie od stanu procedury eksperymentu (wartość 0 oznacza brak bodźca, natomiast 1 maksymalny kontrast możliwy do wyświetlenia). Wartości pozostałych dwóch parametrów,  $ch_b$  oraz  $ch_{rgb}$ , oznaczających kolor tła i bodźca, zostały ustawione zależnie od wariancji kolorystycznej bodźca, dla którego wyszukiwany był próg dostrzegalnego kontrastu.

W przeprowadzonym eksperymencie percepcyjnym zostały wyszukane progi dostrzegalnych kontrastów siedmiu bodźców wzrokowych, które można podzielić na trzy kategorie - achromatyczne, chromatyczne i mieszane. Achromatyczny wzór Gabora, nazywany w naszych badaniach A1, składa się wyłącznie z luminancji. Luminancja tła podczas pomiaru progów dla A1 ustawiana jest na  $50 \text{ cd/m}^2$ , natomiast kolory poszczególnych pikseli zawierających się w bodźcu, są ustawiane w odcieniach szarości. Gdy wzór Gabora jest równy 0, piksel ustawiany jest na kolor czarny, natomiast dla wartości 1 - biały. Podobna zasada dotyczy pozostałych bodźców wzrokowych, gdzie dla wartości równej 0 ustawiany jest tzw. kolor początkowy, a dla 1 kolor końcowy. Wartość parametru wzoru  $ch_b$  oznaczającego luminancję tła ustawiona została na 0.5, natomiast kolor bodźca  $ch_{rgb}$  na  $[1, 1, 1]$ . Chromatyczne bodźce wzrokowe, oznaczane przez nas C2 i C3, zmieniają swoje wartości koloru zgodnie z kierunkami kolorów zgodnymi ze wspomnianym wcześniej zestawem danych ColorFest [109]. Ich kolory to odpowiednio czerwono-zielony i żółto-niebieski. Chromatyczne bodźce wzrokowe nie zmieniają swojej wartości luminancji, która jest równa luminancji tła i wynosi  $23 \text{ cd/m}^2$  dla C2 i  $7 \text{ cd/m}^2$  dla C3. Wartości te wynikają z ograniczeń wykorzystanego w naszych badaniach wyświetlacza. Należy wspomnieć, że chromatyczne bodźce wzrokowe, w odróżnieniu od pozostałych, zostały wygenerowane za pomocą autorskiej procedury zaimplementowanej w środowisku Matlab, która generowała wzór Gabora w taki sam sposób, co wcześniej opisana metoda. Wynikało to z potrzeby zachowania stałego poziomu luminancji wymaganego do wygenerowania konsekwentnie chromatycznego bodźca. Wymagało to dostarczenia do procedury tworzenia wzoru Gabora charakterystyki wykorzystanego w eksperymencie wyświetlacza (Rozdz. 2.1.3), na co nie pozwalała procedura zawarta w zestawie narzędzi Psychtoolbox.



Pozostałe cztery bodźce wzrokowe, oznaczane jako M4, M5, M6 i M7, są typu mieszanego. Łączą one charakterystykę bodźców achromatycznych, jak i chromatycznych, zmieniając swoje wartości luminancji i koloru przy zmianie kontrastu. Ich kierunki kolorów podążają za szeroko stosowaną przestrzenią barw sRGB, modyfikując wyłącznie poszczególne ich kanały. W przypadku mieszanych bodźców wzrokowych M4, M5, M6 i M7 parametr koloru tła  $ch_b$  ustawiony był na 0.5, natomiast kolor wzoru Gabora  $ch_{rgb}$  wynosiły odpowiednio  $[1, 0, 0]$ ,  $[1, 0, 1]$ ,  $[0, 1, 1]$  oraz  $[0, 0, 1]$ . Te parametry wskazują, które kanały, pierwszy R (czerwony), drugi G (zielony), czy trzeci B (niebieski) były modyfikowane dla poszczególnych bodźców wzrokowych. Dla przykładu,  $ch_{rgb}$  dla bodźca wzrokowego M4 wynosiło  $[1, 0, 0]$ . Jedyne wskazują modyfikowane kanały, więc w tym przypadku zmieniał się wyłącznie kanał czerwony. Wartości pikseli we wskazanych kanałach składających się na bodziec wzrokowy były zależne od kontrastu, a dla  $C_m = 1$  te wartości wynosiły  $[0, 0.5, 0.5]$  i  $[1, 0.5, 0.5]$ . W przypadku innego bodźca wzrokowego M6, dla którego parametr  $ch_{rgb}$  wynosił  $[0, 1, 1]$ , modyfikowane były kanały G i B. Motywacją za pomiarem progów kontrastu w obszarze widzenia peryferyjnego bodźców mieszanych w naszych badaniach jest sprawdzenie charakterystyki HVS na czułość kontrastów bodźców wzrokowych, zmieniających swoje wartości wyłącznie w poszczególnych kanałach przestrzeni barw sRGB, która jest stosowana w systemach syntezy obrazów wirtualnych scen. Takie podejście może w przyszłości przysłużyć do wykorzystania pomiarów GDSCF w takich systemach. Korzystając wyłącznie z pomiaru progów kontrastu bodźców wzrokowych achromatycznych i chromatycznych konieczne jest konwertowanie wartości wygenerowanych pikseli na luminancję i chrominancję przez system, co wymaga znajomości charakterystyki stosowanych wyświetlaczy.

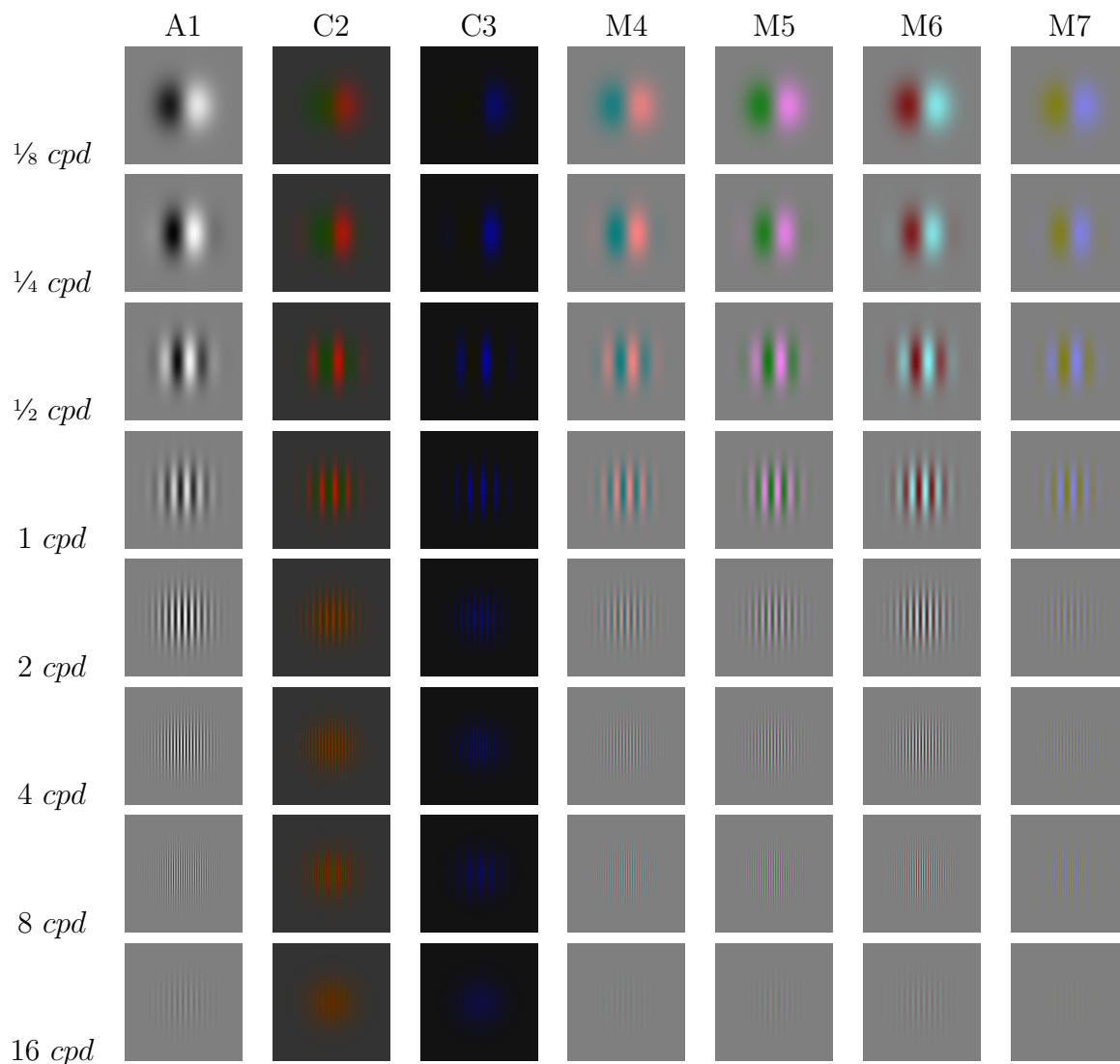
Długość fali sinusoidalnej  $\lambda$  został ustawiony w taki sposób, aby wygenerowane bodźce wzrokowe były o częstotliwości 0.125, 0.25, 0.5, 1, 2, 4, 8 i 16 *cpd*. Wszystkie typy bodźców wzrokowych wraz z wybranymi wartościami częstotliwości zostały przedstawione na Rys. 2.1.1.

### Kontrast bodźca achromatycznego

Z powodu różnej natury typów bodźców wzrokowych, konieczne jest omówienie definicji ich kontrastu. W przypadku achromatycznego bodźca wzrokowego A1 zastosowany został kontrast Webera, który dla potrzeb eksperymentu został zmodyfikowany do poniższej formuły:

$$C_{achrom} = \frac{L_{max} - L_{min}}{L_b},$$

gdzie  $L_{min}$  i  $L_{max}$  to minimalna i maksymalna luminancja bodźca, natomiast  $L_b$  to kolor tła. Należy wspomnieć, że kolor tła nie wynosił dokładnie tyle, ile zamierzono, czyli



Rysunek 2.2. Wygenerowane bodźce w postaci wzoru Gabora w wariacie achromatycznym (A1), chromatycznych (C2-C3), oraz mieszanych (M4-M7). Kontrast bodźców wynosi  $C_M = 0.7$ . Źródło własne [W3] (Rys. 3).

$50 \text{ cd/m}^2$ . Taki sam problem dotyczy luminancji wyświetlonych bodźców wzrokowych, która nie była zgodna z przekazanym parametrem  $C$  do procedury generującej ich obraz. Ten fakt wynikał z niedoskonałości wyświetlacza, która zrekompensowana została pomiarami spektrometrem (co będzie tematem dalszej części pracy). Pomiar luminancji koloru tła dał w wyniku  $48,37 \text{ cd/m}^2$  (pomiar spektrometrem zwrócił wynik  $Y_{xy}=[48.37, 0.3136, 0.3182]$ ).

### Kontrast bodźca chromatycznego

W przypadku chromatycznych bodźców wzrokowych C2 oraz C3, które składają się z więcej niż jednego kanału koloru, uzasadnione było zastosowanie kontrastu Średniej Kwadratowej. Wyliczony on został na podstawie różnic wartości luminancji i chrominancji

bodźców oraz tła w przestrzeni kolorów LMS, której poszczególne kanały odpowiadają wrażliwości na daną barwę trzech rodzajów czopków na siatkówce:

$$C_{chrom} = \frac{1}{\sqrt{3}} \sqrt{\left(\frac{\Delta L}{L_b}\right)^2 + \left(\frac{\Delta M}{M_b}\right)^2 + \left(\frac{\Delta S}{S_b}\right)^2},$$

gdzie  $\Delta L$ ,  $\Delta M$  i  $\Delta S$  oznaczają amplitudę zmian cech wyświetlonego bodźca, natomiast  $L_b$ ,  $M_b$  i  $S_b$  wartości cech tła w przestrzeni kolorów LMS. Cechy we wspomnianej przestrzeni uzyskaliśmy dzięki konwersji wartości  $Y_{xy}$  zmierzonych przez spektrometr. Wpierw skonwertowane one zostały do przestrzeni XYZ poniższą formułą:

$$\begin{aligned} X &= \frac{Y_x}{y}, \\ Y &= Y, \\ Z &= \frac{(1-x-y)Y}{y}, \end{aligned}$$

gdzie  $y \neq 0$ . w przeciwnym wypadku  $X = Y = Z = 0$ . Kolejnym krokiem było użycie równania liniowego stosowanego w badaniach nad poszukiwaniem progów widzialnych kontrastów [43], za pomocą których komponenty koloru są konwertowane z przestrzeni XYZ do LMS. Operacja ta opisana jest poniższą formułą:

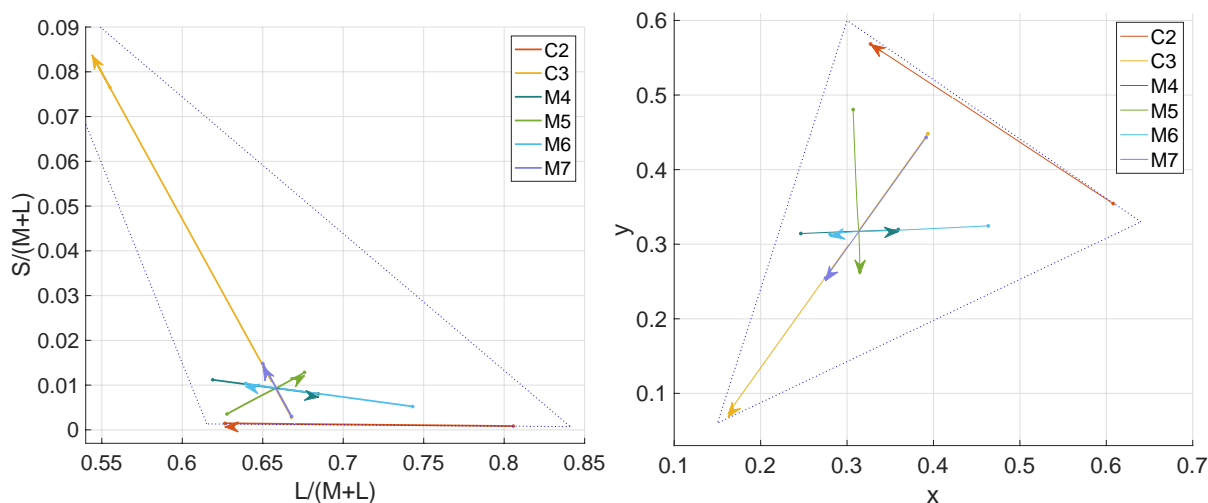
$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.15514 & 0.54312 & -0.03286 \\ -0.15514 & 0.45684 & 0.03286 \\ 0 & 0 & -0.00801 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Należy wspomnieć, że osiągnięte wartości LMS odpowiadają ilości światła odbieranego przez poszczególne typy czopków. Nie istnieje formuła, która pozwoli dokładnie oszacować tę wartość, lecz wskazane przybliżenie jest wystarczające, aby wykorzystać je do wyliczenia wartości kontrastu chromatycznych bodźców wzrokowych.

### Kontrast bodźca mieszanego

Kontrast mieszanych bodźców wzrokowych M4, M5, M6 i M7 został wyliczony tą samą procedurą co w przypadku bodźców chromatycznych. Ich wartości koloru w przestrzeni  $Y_{xy}$  zostały zmierzone spektrometrem, aby zostać skonwertowane do LMS, a sam kontrast był wyliczony formułą średniej kwadratowej, wykorzystując wartości cech bodźca i tła w tej przestrzeni.

W tabeli 2.1 zamieszczono pomiary wartości  $Y_{xy}$  i konwersję na LMS wszystkich typów bodźców wzrokowych wykorzystanych w naszych badaniach. Pomiar odbył się dla wartości kontrastu  $C_m = 0.8$ . Na Rys. 2.3 natomiast znajduje się graficzna interpretacja kierunku kolorów chromatycznych i mieszanych bodźców wzrokowych.



Rysunek 2.3. Kierunki kolorów dla chromatycznych i mieszanych bodźców z Tab. 2.1 wykreślone na schemacie Boyntona-MacLeoda (po lewej) i w przestrzeni chromatyczności CIE 1931  $xy$  (po prawej). Linie przerywane przedstawiają gamę kolorów możliwą od odwzorowania przez wyświetlacz, zwroty kierunków są oznaczone strzałkami. Należy zwrócić uwagę na pominiętą wartość luminancji dla bodźców M4-M7. Źródło własne [W3] (Rys. 1).

|    | Kolor początkowy<br>Kolor końcowy | Y     | x    | y    | L     | M     | S    | $\Delta Y$<br>$ch_{rgb}$ |
|----|-----------------------------------|-------|------|------|-------|-------|------|--------------------------|
| A1 | Czarny                            | 9.62  | 0.31 | 0.32 | 6.33  | 3.29  | 0.09 | 75.45                    |
|    | Biały                             | 85.06 | 0.31 | 0.32 | 55.97 | 29.09 | 0.80 | [1,1,1]                  |
| C2 | Zielony                           | 20.55 | 0.61 | 0.35 | 16.56 | 3.99  | 0.02 | 0.45                     |
|    | Czerwony                          | 20.10 | 0.33 | 0.57 | 12.59 | 7.51  | 0.03 | $ch_{rgb}(C2)$           |
| C3 | Żółty                             | 6.91  | 0.39 | 0.45 | 4.62  | 2.30  | 0.02 | 0.26                     |
|    | Niebieski                         | 6.65  | 0.17 | 0.08 | 3.69  | 2.96  | 0.51 | $ch_{rgb}(C3)$           |
| M4 | Różowy                            | 39.60 | 0.25 | 0.31 | 24.51 | 15.09 | 0.44 | 16.91                    |
|    | Zielono-cyjanowy                  | 56.51 | 0.36 | 0.32 | 38.69 | 17.83 | 0.45 | [1,0,0]                  |
| M5 | Zielony                           | 36.64 | 0.31 | 0.48 | 23.00 | 13.64 | 0.13 | 22.57                    |
|    | Jasno-różowy                      | 59.21 | 0.31 | 0.26 | 40.02 | 19.19 | 0.76 | [1,0,1]                  |
| M6 | Ciemno-czerwony                   | 18.24 | 0.46 | 0.32 | 13.56 | 4.69  | 0.10 | 58.99                    |
|    | Cyjanowy                          | 77.23 | 0.28 | 0.31 | 49.38 | 27.85 | 0.80 | [0,1,1]                  |
| M7 | Zielono-żółty                     | 45.17 | 0.39 | 0.44 | 30.17 | 15.00 | 0.13 | 6.31                     |
|    | Niebiesko-fioletowy               | 51.48 | 0.28 | 0.25 | 33.46 | 18.01 | 0.76 | [0,0,1]                  |

Tabela 2.1. Wartości  $Yxy$  i LMS punktów końcowych kolorów (górna część jako kolor początkowy, dolna jako końcowy).  $\Delta Y$  oznacza różnicę luminancji między punktami końcowymi. Wartość  $ch_{rgb}$  oznacza wykorzystany do wygenerowania bodźca wektor.  $ch_{rgb}(C2) = [0, 2080; -0.6772; 0]$ ,  $ch_{rgb}(C3) = [-0, 2080; -0, 6772; 0, 0797]$ . Źródło własne [W3] (Tabela. 1).

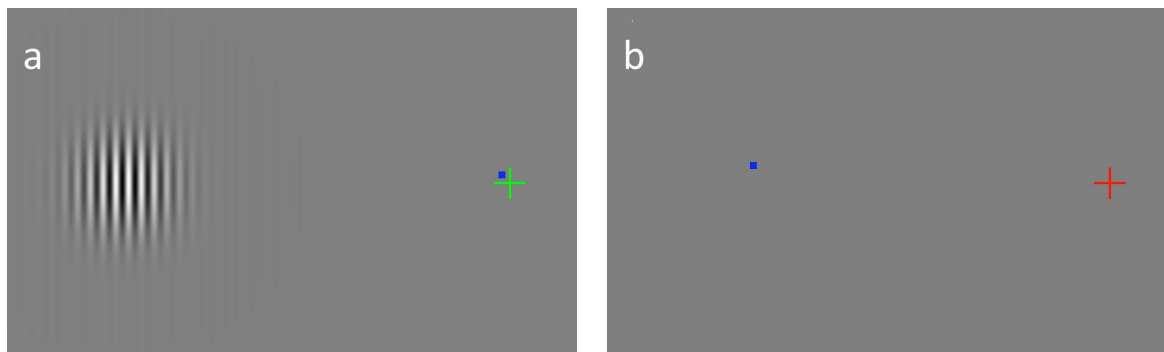
### 2.1.2. Procedura eksperymentu

Eksperyment percepcyjny odbywał się na przygotowanym stanowisku komputerowym wyposażonym w podpórkę na podbródek (ang. *chin rest*), która miała za zadanie ustabilizować głowę obserwatora. Wyświetlacz, służący do prezentowania bodźców wzrokowych obserwatorowi, został umieszczony w odległości 90 cm od jego oczu, co przekładało się ok. 57 pikseli wyświetlacza na stopień kątowy widzenia. Pod wyświetlaczem umieszczony został mobilny eye tracker, który miał na celu śledzenie wzroku obserwatora, celem oszacowania kierunku jego patrzenia. Na stanowisku znajdowała się także klawiatura, która pozwalała na interakcję ze strony uczestników eksperymentu.

Stosowana przez nas metodologia pomiarów, przedstawiona w autorskiej publikacji [W2], polegała na kontrolowaniu obszaru padania obrazu bodźca wzrokowego na siatkówkę poprzez wykorzystanie eye trackera. Metodologia ta różni się od poprzednich badań związanych z pomiarem czułości na kontrast, gdzie bodziec wyświetlany był przez ograniczony czas, aby obserwator nie zdążył zmienić kierunku patrzenia [84], tym samym powodując błędny pomiar czułości w obszarze widzenia centralnego zamiast peryferyjnego. Eye tracker monitorował punkt patrzenia obserwatora. Gdy wpatrywał się on we wskazany obszar, na ustalonej odległości kątowej od niego pojawiał się bodziec wzrokowy, natomiast gdy zmienił on swój kierunek patrzenia, co jest naturalną reakcją na pojawienie się wartego uwagi obiektu w obszarze peryferyjnym (Rozdz. 2.4), był on wygaszany, a odpowiedź obserwatora uznawana za nieważną.

Autorska procedura zaimplementowana w środowisku Matlab wyświetlała szare tło o określonej luminancji, na które nałożony został znacznik w postaci krzyżyka po prawej stronie ekranu (Rys. 2.4(b)). Uczestnicy eksperymentu byli proszeni o nieustanne wpatrywanie się w ten znacznik. Wyświetlacz został przesunięty w lewo, aby obserwowany znacznik znajdował się na wprost do obserwatora. Gdy eye tracker wykrył poprawny kierunek patrzenia, po lewej stronie w określonej odległości wyświetlony został odpowiedni bodziec wzrokowy (Rys. 2.4(a)). Przedstawiony on był w losowej orientacji, poziomej, bądź pionowej. Zadaniem obserwatora było rozpoznać kierunek wyświetlonej sinusoidalnej siatki i udzielić odpowiedzi przez naciśnięcie odpowiadających przycisków na klawiaturze, strzałka w górę bądź dół, w przypadku orientacji pionowej, natomiast strzałka w lewo lub prawo dla poziomej. Gdy wzrok obserwatora oddalił się od wskazanego punktu o  $2^\circ$  widzenia, bodziec był wygaszany (Rys. 2.4(b)), a eksperyment był wstrzymywany do czasu przywrócenia poprawnego kierunku patrzenia obserwatora. Kiedy to się stało, orientacja bodźca była ponownie losowana i po czasie 0.5 sekundy został on wyświetlony (Rys. 2.4(c)). Po udzieleniu przez obserwatora odpowiedzi, dotyczącej orientacji bodźca wzrokowego, następowała 200 milisekundowa przerwa, w trakcie której na miejsce bodźca wyświetlony został szum otoczony filtrem Gaussowskim, mający na celu zminimalizowanie efektu tzw. powidoku, czyli naświetlenia fotoreceptorów i wywołanie złudnego wrażenia

ciągłego istnienia bodźca wzrokowego, co powodowało błędne postrzeganie jego orientacji w kolejnych próbach. Po tej przerwie procedura była powtarzana dla innego kontrastu bodźca, do czasu odnalezienia wartości progowej.



(a) Obserwator patrzy na wyznaczony obszar, bodziec jest wyświetlany w ustalonej odległości. (b) Kierunek patrzenia znacznie oddalił się od ustalonego punktu, bodziec znika.

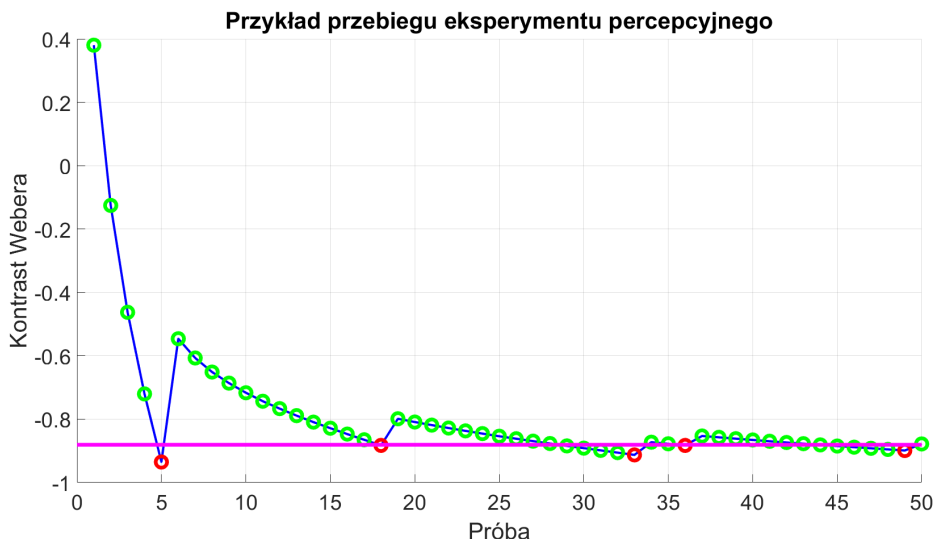


(c) Wzrok obserwatora wraca na wyznaczone miejsce. Bodziec pojawia się ponownie, ale z losowo dobraną orientacją.

Rysunek 2.4. Zrzuty ekranu z eksperymentu percepcyjnego. Zielony, bądź czerwony krzyżyk wskazuje pożądany kierunek patrzenia. Niebieski kwadrat przedstawia pozycję wzroku zarejestrowaną przez eye tracker. Źródło własne [W2] (Rys. 2).

Kolejne wartości kontrastu bodźców wzrokowych, prezentowanych obserwatorom w eksperymencie, dobierane były na podstawie bayesowskiej adaptacyjnej metody psychometrycznej QUEST [103]. Procedura ta stworzona została z myślą o wyszukiwaniu wartości progowych, uwzględniając rozkład odpowiedzi uczestników eksperymentu oraz kształt funkcji psychometrycznej [63]. QUEST adaptacyjne określał wartość kontrastu wyświetlonego bodźca dla danej próby na podstawie odpowiedzi obserwatora w próbach poprzednich, obniżając ją po prawidłowym określeniu orientacji bodźca, zwiększając natomiast gdy podana zostanie niewłaściwa odpowiedź. Z każdą kolejną udzieloną odpowiedzią, przedział wartości, w którym szacowana jest ta progowa się zmniejszała. Gdy odchylenie standardowe estymowanego przedziału wartości było nieistotne statystycznie, za wartość progową brano była średnia z końców tego zakresu. Procedura QUEST bierze pod uwagę specyfikę metodologii wymuszonego wyboru (ang. *force choice*) spośród dwóch

odpowiedzi stosowanej w eksperymencie percepcyjnym, gdzie przy braku świadomości obserwatora o orientacji bodźca istnieje 50% prawdopodobieństwa na poprawną odpowiedź. W naszym eksperymencie przeszukano wartość progową widzialnego kontrastu za pomocą jednostek  $\log_{10}(c)$ . Maksymalna liczba prób na pomiar progu widzialnego kontrastu dla danego bodźca wzrokowego i obszaru w polu widzenia wynosiła 50. Przykładowy dobór wartości kontrastu w eksperymencie zaprezentowano na Rys. 2.5.



Rysunek 2.5. Przykładowy przebieg pomiaru czułości na kontrast z wykorzystaniem procedury QUEST. Wykres przedstawia wyświetlone wartości kontrastu Webera bodźca w kolejnych próbach. Zielone kropki oznaczają prawidłowe odpowiedzi, natomiast czerwone złe. Fioletowa pozioma linia oznacza końcową wartość braną za próg kontrastu.

Pomiar progów kontrastu rozpoczęty został poprzez oszacowanie charakterystyk CSF w obszarze centralnym, gdzie czułość na bodźce uzależniona jest od ich częstotliwości bodźca. Osiągnięte wyniki posłużyły nam do weryfikacji, czy nasze oprogramowanie i środowisko zostały poprawnie przygotowane. Należało także zweryfikować metodologię, która dotyczyła wykorzystania eye trackera do kontroli kierunku patrzenia obserwatora. W tym celu przeprowadzony został dodatkowy pomiar funkcji GDCSF, w trakcie którego bodziec nadal był wygaszany po przesunięciu kierunku patrzenia obserwatora od wskazanego obszaru, lecz przywrócenie bodźca nie wiązało się z ponownym doborem jego orientacji, co pozwoliło spoglądać na bodziec dla mniejszych od wskazanych odległości kątowych od obszaru widzenia centralnego. Właściwy eksperyment składał się z pomiaru czułości na kontrast w obszarze peryferyjnym dla odległości  $5^\circ$ ,  $10^\circ$ ,  $15^\circ$ ,  $20^\circ$ , oraz  $27^\circ$ . Dla każdego obszaru pola widzenia pomiar odbył się dla siedmiu bodźców wzrokowych. Widzialność bodźca achromatycznego A1 została zmierzona dla częstotliwości 0.125, 0.25, 0.5, 1, 2, 4, 8 oraz 16 *cpd*, natomiast dla pozostałych bodźców, chromatycznych i mieszanych, pomiar odbył się wyłącznie dla częstotliwości 2 *cpd*.

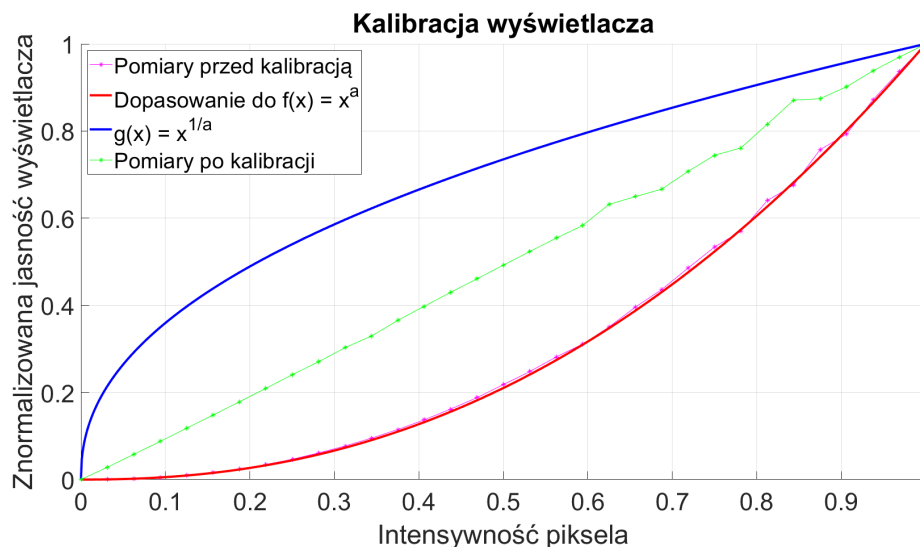
### 2.1.3. Aparatura

Zastosowana w naszych badaniach metodologia polegała na wykorzystaniu eye trackera, urządzenia służącego do śledzenia kierunku patrzenia wzroku obserwatora [24]. Idea jego działania polega na zastosowaniu kamery z filtrem podczerwieni, która śledzi pozycję źrenicy obserwatora oraz odbicia światła podczerwonego, rzucanego przez urządzenie od rogówki. Na podstawie wskazanych punktów możliwe jest oszacowanie kierunku patrzenia. Do poprawnego funkcjonowania eye trackera wymagany jest proces kalibracji, w której wyliczane są współczynniki służące do mapowania położenia źrenicy i odbicia rogówkowego do współrzędnych ekranu. W celu zapobiegnięcia niezamierzonym ruchom oczu w kierunku bodźca podczas pomiaru progów widzialnego kontrastu bodźców wzrokowych w obszarze widzenia peryferyjnego, wykorzystany został Eye Tribe - eye tracker działający w częstotliwości 60 Hz. Urządzenie było stawiane pod wyświetlaczem, co umożliwiało obserwowanie przez kamerę gałek ocznych uczestników eksperymentu, które były ustabilizowane przez podpórkę na podbródek, co zwiększyło precyzję wyliczeń kierunku patrzenia. Precyzja okulografa Eye Tribe wynosi  $0.5^\circ$  [79], co wystarczało do przeprowadzonego eksperymentu, w którym próg dopuszczalnego odchylenia od pożądanego kierunku patrzenia ustawiony był na  $2^\circ$ . Kalibracja eye trackera odbywała się przed każdą sesją eksperymentalną i polegała na obserwowaniu przez uczestników dziewięciu wyświetlanych punktów na wyświetlaczu. Proces ten trwał około 30 sekund i był obsługiwany przez dostarczone do sprzętu oprogramowanie.

Eksperyment percepcyjny przeprowadzony został z wykorzystaniem wyświetlacza SONY PVM-A250 PRIMASTER EL. Jest to wysokiej jakości wyświetlacz OLED o rozdzielczości 1920 x 1080 pikseli. Zapewnia on dobre odwzorowanie luminancji i chrominancji dzięki 10-bitowej głębi kolorów, która była niezbędna w eksperymencie do wyświetlenia bodźców o niskiej wartości kontrastu. Przed przeprowadzeniem eksperymentu należało przeprowadzić na nim proces kalibracji, aby luminancja i chrominancja była wyświetlona w sposób liniowy wraz ze wzrostem intensywności piksela. Współczesne systemy komputerowe systemów wyświetlają obrazy po zastosowaniu korekcji gamma, w celu optymalizacji wykorzystania bitów kodowania obrazu, wykorzystując nieliniową charakterystykę postrzegania luminancji i kolorów HVS [87], który jest bardziej wrażliwy na względne różnice w mniejszych wartościach luminancji (prawo Webera). Sama kalibracja polegała na zmierzeniu poziomu luminancji wyświetlacza, gdy ten przedstawiał szare tło o zadanej wartości piksela nazywanej intensywnością. W ten sposób zmierzaliśmy charakterystykę luminancji wyświetlacza zależną od zmiany intensywności piksela, która została dopasowana do funkcji  $F(x) = x^a$ . Funkcja do niej odwrotna  $G(x) = x^{\frac{1}{a}}$  posłużyła do wygenerowania wartości nowej funkcji gamma nakładanej przez oprogramowanie w procedurze eksperymentu. Po tej operacji dokonano kolejnego pomiaru luminancji, która wskazała lepsze, lecz nie idealne przedstawienie jasności tła w zależności od wartości piksela.



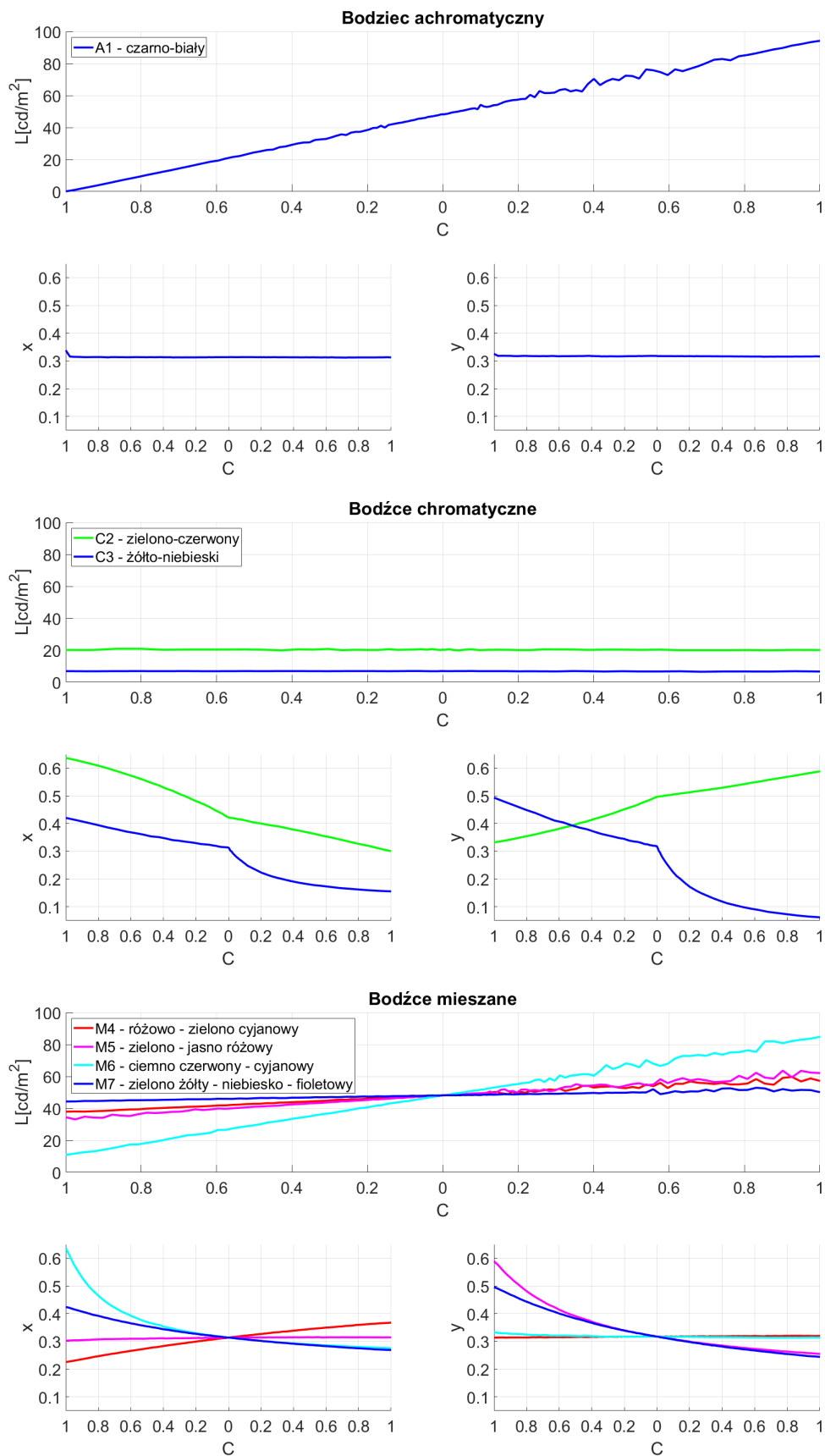
Pomiary, dokonane w przedstawionym procesie, przedstawiono na Rys. 2.6. Do kalibracji wykorzystaliśmy chromometru Minolta CS-100A. Profil kolorów został ustawiony na sRGB, natomiast wartości kontrastu bodźców wzrokowych dobierano, unikając poziomów luminancji poniżej  $1 \text{ cd/m}^2$  i powyżej  $100 \text{ cd/m}^2$ , przy których kalibracja wyświetlacza była nierzetelna.



Rysunek 2.6. Pomiar jasności pikseli służący do kalibracji wyświetlacza. Różowe punkty oznaczają zmierzone wartości przed kalibracją, natomiast zielone przedstawiają punkty po kalibracji. Czerwona i niebieska linia oznacza kolejno funkcje  $F(x)$  i  $G(x)$ .

Brak idealnej liniowej zależności między luminancją a intensywnością piksela, nie sprawił błędów w odwzorowaniu kontrastu wyświetlonych bodźców. Wynikało to z faktu, iż każdy z siedmiu wariacji kolorystycznych wzoru Gabora został zmierzony spektrometrem, dzięki czemu wyznaczono wartości cech bodźców w przestrzeni kolorów Yxy, w zależności od jego kontrastu (pomiar przedstawiony na Rys. 2.7). Wyniki tych pomiarów zostały dopasowane do wielomianu, który posłużył do wyliczenia kontrastu wyświetlonego bodźca wzrokowego na podstawie parametru  $C_m$  oraz kierunku koloru. Pomiar bodźców odbył się przy użyciu spektrometru Specbos 1211 i potwierdzone zostały za pomocą chromometru Minolta CS-100A.

Oprogramowanie eksperymentu zostało zaimplementowane w środowisku Matlab wykorzystując zestaw narzędzi Psychtoolbox, który umożliwiał wyświetlanie m.in. bodźców wzrokowych dzięki wbudowanemu potokowi renderowania. Użycie eye trackera, który był obsługiwany przez dostarczone od producenta oprogramowanie, komunikował się z naszymi procedurami przez protokół TCP/IP. Komputer, na którym uruchomione zostało oprogramowanie, wyposażony był w kartę graficzną Nvidia Quadro 3000, która oferowała wystarczającą moc obliczeniową do przeprowadzenia procedur oraz, co najważniejsze, umożliwiła wyświetlanie obrazu z 10-bitową głębią kolorów.



Rysunek 2.7. Wynik pomiaru spektrometrem wartości luminancji i chrominacji bodźców wzrokowych wykorzystanych w eksperymencie percepcyjnym.

### 2.1.4. Uczestnicy eksperymentu

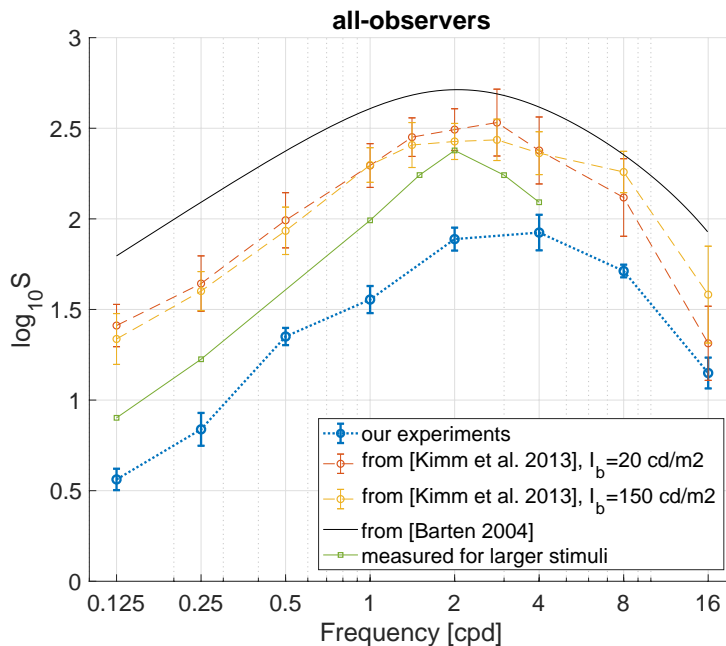
Eksperyment został wykonany na grupie 5 mężczyzn, których wiek wynosił od 21 do 47 lat. Ochotnicy oświadczyli, że posiadają zdrowy bądź skorygowany układ wzrokowy. Byli oni świadomi, na czym polega eksperyment i co powinni robić, nie mieli natomiast świadomości o działaniu zastosowanych procedur. Nie byli także informowani, czy orientacja uległa zmianie po niezamierzonym przesunięciu kierunku widzenia. Sesja eksperymentalna była podzielona na części i rozłożona w czasie. Pojedynczy pomiar czułości dla danej częstotliwości, obszaru widzenia oraz koloru trwał około 3 minut. Pomiedzy zmianami odległości kątovej wyświetlonego bodźca wzrokowego występowała minutowa przerwa, a podczas zmiany częstotliwości bądź koloru wymagane było 10 minut odpoczynku, z uwagi na męczącą naturę eksperymentu. Wszystkie sesje przeprowadzono w ciągu 5-8 tygodni w różnych odstępach czasu, w zależności od zmęczenia zgłoszonego przez uczestników.

## 2.2. Rezultaty eksperymentu

W tym podrozdziale zaprezentowane zostały wyniki pomiarów czułości na kontrast w eksperymencie percepcyjnym. Podzielono je na trzy kategorie. W pierwszej kolejności pomiar progów kontrastu przeprowadzono dla bodźca achromatycznego A1 w obszarze widzenia centralnego o częstotliwościach 0.125, 0.25, 0.5, 1, 2, 4, 8 oraz 16 *cpd* (Rozdz. 2.2.1). Następnie zmierzona została czułość na bodziec achromatyczny A1 dla tego samego zakresu częstotliwości, lecz w obszarze widzenia peryferyjnego w odległościach 5°, 10°, 15°, 20° oraz 27° (Rozdz. 2.2.2). W kolejnej części omówione zostały pomiary czułości na kontrast dla pozostałych bodźców wzrokowych (chromatycznych C2 i C3, oraz mieszanych M4, M5, M6 i M7) dla wspomnianych wcześniej odległości kątowych i w obszarze widzenia centralnego, lecz wyłącznie dla częstotliwości 2*cpd* (Rozdz. 2.2.3). W końcowej części rozdziału przedstawione zostały wyniki eksperymentu, wskazującego poprawność wykorzystanej w pomiarach metodologii (Rozdz. 2.2.4).

### 2.2.1. Funkcja czułości na kontrast dla achromatycznego bodźca

W celu sprawdzenia, czy nasz eksperyment został przygotowany poprawnie, wykonaliśmy pomiary czułości kontrastu dla bodźca achromatycznego A1 w obszarze widzenia centralnego i porównaliśmy je z pomiarami wykonanymi w literaturze. Na Rys. 2.8 niebieską linią oznaczono uśrednione wyniki przeprowadzonego przez nas eksperymentu mierzącego progi kontrastu bodźców o częstotliwościach 0.125, 0.25, 0.5, 1, 2, 4, 8 oraz 16 *cpd*. Na wykresie można zauważyć, że zmierzona przez nas charakterystyka CSF ma kształt zgodny z pomiarami uzyskanymi w literaturze, lecz osiągnęliśmy szczytową czułość na kontrast dla częstotliwości 4 *cpd*, podczas gdy Barten [7] oraz Kim [43] otrzymali

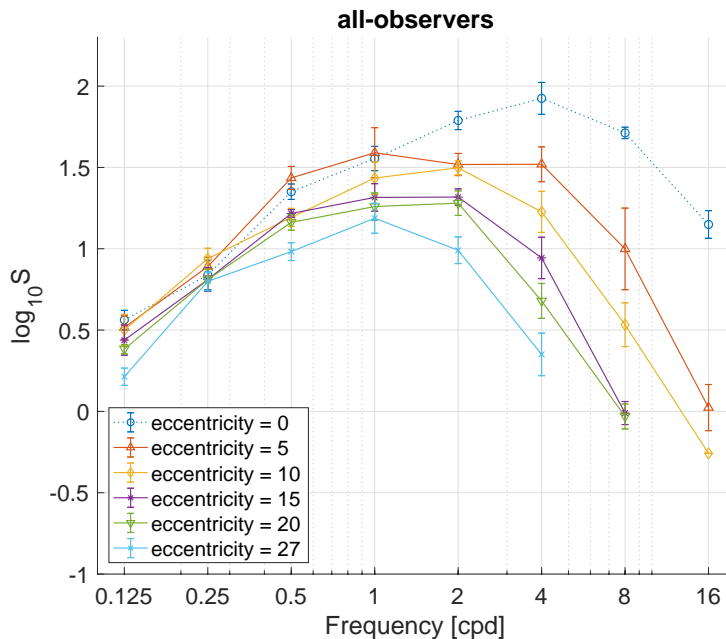


Rysunek 2.8. Wyniki pomiaru czułości na kontrast dla bodźca achromatycznego A1. Przedstawione zostały uśrednione wartości pomiarów wśród wszystkich uczestników eksperymentu przeprowadzonych w naszych badaniach oraz tych zawartych w literaturze. Pionowe linie przy pomiarach oznaczają błąd standardowy średniej. Źródło własne [W3] (Rys. 4).

najwyższą czułość dla częstotliwości równej 2 *cpd*. Ich pomiary zostały także zamieszczone na Rys. 2.8. Bezpośrednie porównanie wskazuje zaniżoną czułość na kontrast, uzyskaną w naszych pomiarach, co wyjaśnialiśmy różną wielkością bodźców wzrokowych w obu eksperymentach. Chcąc udowodnić tę obserwację, przeprowadziliśmy dodatkowe pomiary przy użyciu dwukrotnie większego rozmiaru bodźca (odchylenie standardowe obwiedni Gaussa  $\sigma$  zostało zmienione z wartości 70 na 140), których wyniki zostały przedstawione na wykresie w postaci zielonej linii. Zmierzona czułość, w tym przypadku, była podobna do pomiarów wykonanych w badaniach Kim. W dalszych pomiarach zachowaliśmy jednak ten sam rozmiar bodźca, ponieważ jego zwiększenie mogło powodować błędne pomiary dla obszarów widzenia peryferyjnego, w którym ważne jest pozycjonowanie bodźca wzrokowego w ustalonej odległości.

### 2.2.2. Funkcja czułości na kontrast dla achromatycznego bodźca w obszarze peryferyjnym

Kolejne serie pomiarów dotyczyły czułości na kontrast dla bodźca achromatycznego A1 w obszarze widzenia peryferyjnego w odległościach 5°, 10°, 15°, 20° i 27°. Uśrednione wyniki tych pomiarów zaprezentowane zostały na Rys. 2.9. Dla częstotliwości równych 0.125 i 0.25 *cpd* czułość na kontrast jest zbliżona do zmierzonej czułości w obszarze centralnym, natomiast dla wyższych częstotliwości czułość zaczyna się zmniejszać, przy czym



Rysunek 2.9. Pomiary czułości na kontrast dla achromatycznego bodźca wzrokowego A1 w peryferyjnych obszarach widzenia. Pionowe linie przy pomiarach oznaczają błąd standardowy średniej. Linia przerywana pokazuje pomiary dla obszaru widzenia centralnego. Źródło własne [W3] (Rys. 5).

siła spadku zwiększa się wraz z częstotliwością. Dla bodźca wzrokowego o częstotliwości 8 *cpd* różnica w czułości kontrastu między widzeniem centralnym, a peryferyjnym w odległości 20° zbliża się do dwóch rzędów wielkości skali logarytmicznej. Wartości progowe kontrastu dla częstotliwości 8 *cpd* i odległości 27° oraz częstotliwości 16 *cpd* w odległościach równej 20° i wyższych nie zostały zmierzone, ponieważ obserwatorzy nie byli w stanie zidentyfikować orientacji bodźca, nawet dla maksymalnej wartości kontrastu możliwej do wyświetlenia w eksperymencie.

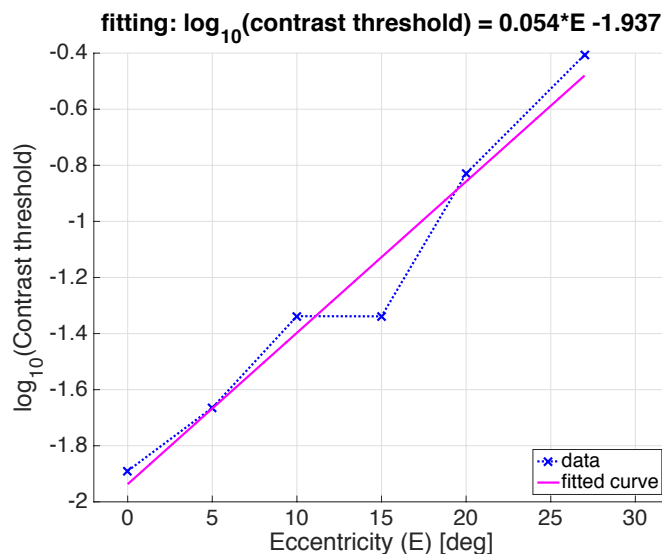
W celu zweryfikowania poprawności pomiarów GDCSF dla bodźca achromatycznego w częstotliwości 2 *cpd*, dopasowaliśmy logarytmy osiągniętych progów kontrastu do równania liniowego:

$$\log_{10}(C_{thr}) = m * E + b,$$

gdzie  $\log_{10}(C_{thr})$ , to wartość progowa kontrastu bodźca w skali logarytmicznej,  $E$  oznacza odległość od dołka środkowego oka, natomiast  $m$  i  $b$  to parametry równania. Metoda najmniejszych kwadratów wskazała na równanie liniowe w postaci:

$$\log_{10}(C_{thr}) = 0.054 * E - 1.937,$$

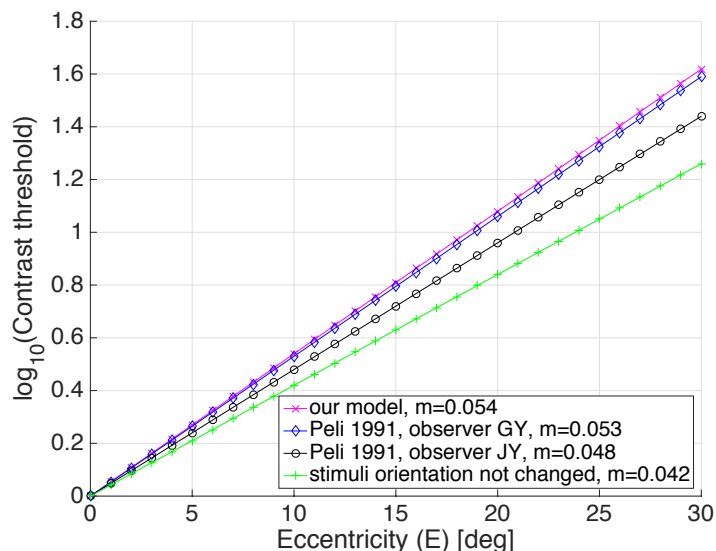
którego graficzną reprezentację przedstawiono na Rys. 2.10. Weryfikacja poprawności modelu jest problematycznym zadaniem, gdyż bezpośrednie porównanie z wynikami



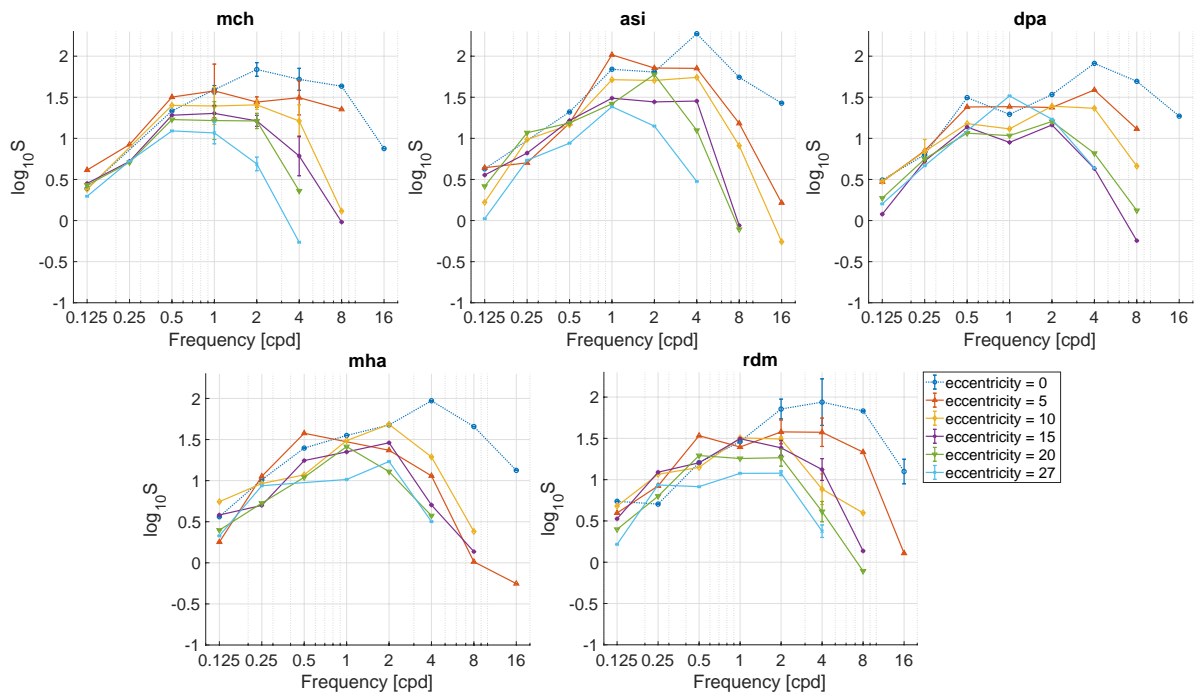
Rysunek 2.10. Uśrednione progi kontrastu uzyskane w pomiarach czułości na bodziec A1 przedstawione w skali logarytmicznej. Fioletowa linia pokazuje dopasowanie zmierzonych danych do równania liniowego. Źródło własne [W2] (Rys. 5).

pomiarów badań w tej dziedzinie wykazuje różnice, wynikające z różnych warunków eksperymentu, m.in. czas wyświetlenia bodźca [84]. Różnica ta objawia się w wartości progowej kontrastu dla widzenia centralnego (odległość  $0^\circ$ ), która w przedstawionym równaniu utożsamiana jest z parametrem  $b$ . Drugi parametr  $m$  wskazuje na zależność liniową pomiędzy progiem kontrastu, a odległością kątową. Parametr ten porównaliśmy z pomiarami Peli (Rys. 2.11), które traktowaliśmy jako referencyjne ze względu ich zgodność z innymi badaniami. Jak widać na wykresie, nasze dopasowanie parametru  $m$  do równania odpowiada pomiarom Peli dla obserwatora GY. Różnica między współczynnikami  $m$  wynosi 0.001, a biorąc pod uwagę fakt, że różnica średnich wartości dla obserwatorów GY i JY wynosi 0.0035, można wnioskować o poprawności naszych pomiarów GDCSF.

Pomiary czułości na kontrast dla achromatycznego bodźca wzrokowego wykazują, że szczytową zdolność rozpoznawania wzorców HVS osiąga dla bodźca wzrokowego o częstotliwości od 2 do 4 *cpd*. Model Bartena oraz Kima wskazują na szczytową czułość dla 2 *cpd*, z tego powodu więc pomiary dla chromatycznych oraz mieszanych bodźców wzrokowych odbyły się wyłącznie dla tej częstotliwości.



Rysunek 2.11. Porównanie dopasowania pomiarów GDSCF dla bodźca achromatycznego o częstotliwości 2 *cpd* do równania liniowego. Linie o kolorach czarnym i niebieskim wskazują pomiary Peli [84], natomiast fioletowa linia wskazuje progi kontrastu uzyskane w naszych badaniach. Zielona linia wskazuje dopasowanie do równania wyniki pomiarów bez zmiany orientacji bodźca po zmianie kierunku patrzenia obserwatora. Źródło własne [W2] (Rys. 6).

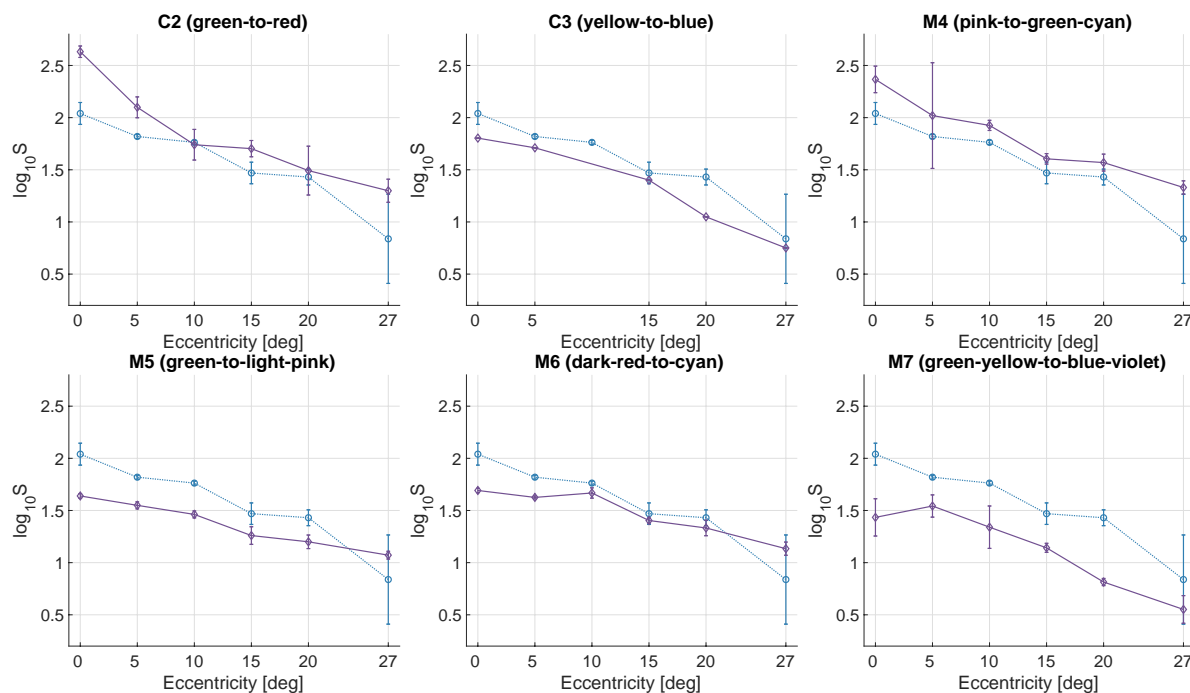


Rysunek 2.12. Pomiary czułości na kontrast dla achromatycznego bodźca wzrokowego A1 dla poszczególnych obserwatorów (oznaczonych jako mch, asi, dpa, mha i rdm). Źródło własne [W3] (Rys. 6).

Wyniki pomiarów GDCSF achromatycznego bodźca wzrokowego A1 dla poszczególnych obserwatorów przedstawiono na Rys. 2.12. Wykresy wykazują względną spójność pomiarów u uczestników eksperymentu, istnieją jednak nieregularności w kształtach uzyskanych krzywych, szczególnie widocznych przy większych odległości kątowych. Jak wspomniano wcześniej, obserwatorzy mieli problem ze spostrzeżeniem bodźców wzrokowych o częstotliwościach 8 i 16 *cpd* w odległościach 20° oraz 27°. W niektórych przypadkach można zauważyć, że progi kontrastu dla różnych odległości oscylują względem siebie, co może być spowodowane błędami pomiaru. Pomimo tych nieregularności zmierzonych dla poszczególnych obserwatorów, uśrednione wyniki konsekwentnie pokazują, że czułość na kontrast spada wraz z odległością od obszaru widzenia centralnego.

### 2.2.3. Funkcja czułości na kontrast dla chromatycznego i mieszanego bodźca w obszarze peryferyjnym

Na Rys 2.13 przedstawiono wyniki pomiarów czułości na kontrast dla chromatycznych oraz mieszanych bodźców wzrokowych. Uzyskane progi dotyczyły częstotliwości 2 *cpd* bodźca oraz obszaru widzenia centralnego i peryferyjnego w odległościach 5°, 10°, 15°, 20° i 27°.

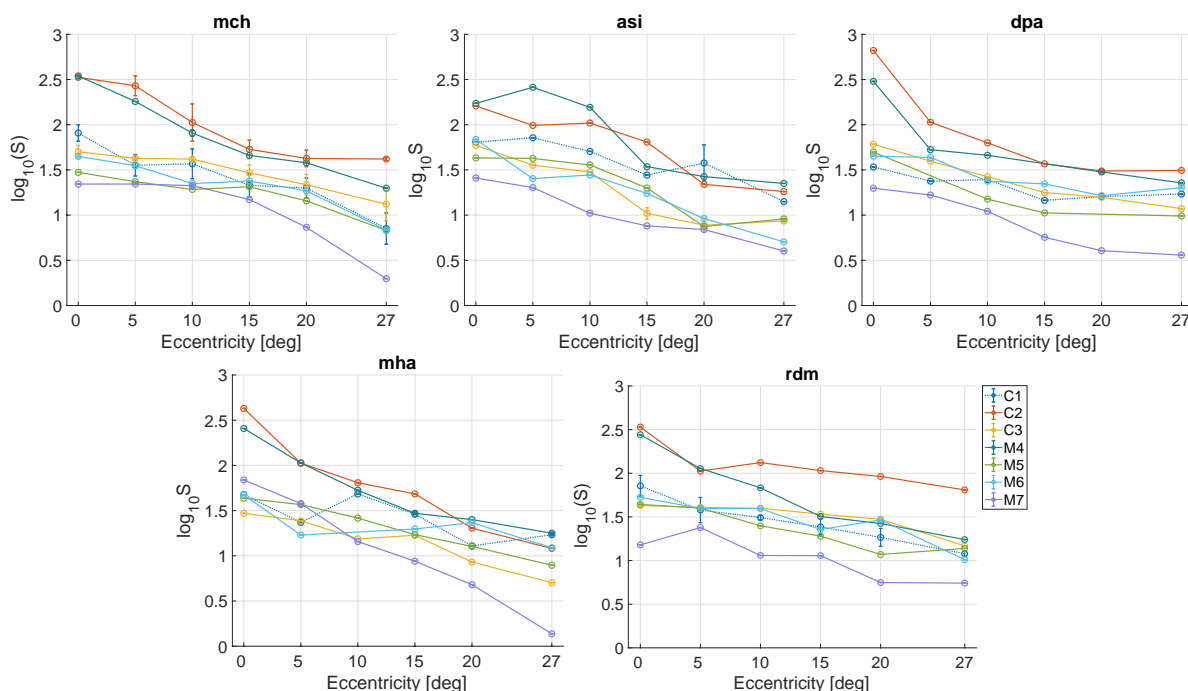


Rysunek 2.13. Uśrednione wyniki pomiarów czułości na kontrast dla chromatycznych (C2, C3) i mieszanych (M4-M7) bodźców wzrokowych uzyskane w eksperymencie percepcyjnym. Fioletowa linia oznacza wyniki dla poszczególnego bodźca, natomiast niebieska wskazuje pomiar A1. Pionowe linie przy pomiarach oznaczają błąd standardowy średniej. Źródło własne [W3] (Rys. 7).



Jak można zobaczyć na Rys. 2.13, dla wszystkich kierunków kolorów wykresy wykazują spadek czułości na kontrast wraz z oddalaniem się od obszaru widzenia centralnego, co jest zgodne z poprzednimi badaniami. Spadek ten jest porównywalny z pomiarami progów kontrastu dla bodźca achromatycznego. Pomiar dla C2 pokazuje wyższą czułość w porównaniu do A1. W obszarze widzenia centralnego  $\log_{10}S$  jest zbliżone do 2.6, co jest porównywalne z wartością 2.4 uzyskaną przez Kim [43] (Rys. 2.8). Czułość na kontrast dla bodźca C3 jest mniejsza niż C2 i A1.

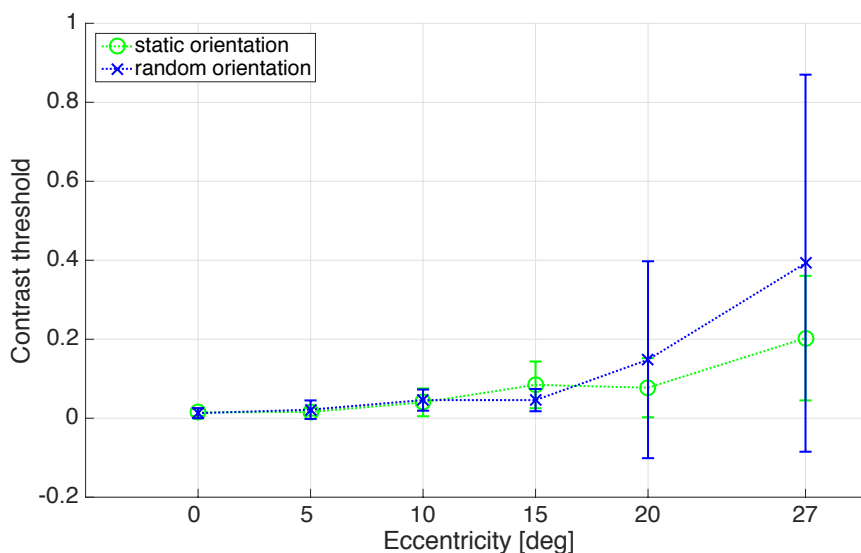
Czułość na kontrast mieszanego bodźca wzrokowego M7 jest niska, co można wytłumaczyć jego słabą zmianą luminancji ( $\Delta Y = 6.31$ ) i podobnym kierunkiem koloru co C3, ale zawartym w obszarze niskiej czułości (Rys. 2.3). Zbieżność kierunków występuje także dla M4 i M6, lecz czułość na niej znacznie się różni z powodu dużej zmiany luminancji dla M6 ( $\Delta Y = 58.99$ ). Czułość na kontrast HVS dla bodźców mieszanych wskazuje wyższą wartość dla kierunków bliższych C2 (np. M4), w szczególności, gdy amplituda wahań luminancji bodźca jest słaba. W przypadku podobieństwa do kierunku C3, wykazywana jest niższa czułość na kontrast (np. M7), natomiast zwiększenie wahań luminancji zbliża osiągnięte progi kontrastu do A1. Na Rys. 2.14 przedstawiono wyniki pomiarów czułości kontrastów bodźców C2, C3 i M4-M7 dla poszczególnych obserwatorów. Wskazują one spadek czułości na kontrast wraz z oddalaniem się od obszaru widzenia centralnego.



Rysunek 2.14. Pomiary czułości na kontrast chromatycznych (C2, C3) i mieszanych (M4-M7) bodźców wzrokowych o częstotliwości 2 *cpd* dla poszczególnych obserwatorów. Niebieska linia wskazuje pomiary dla bodźca achromatycznego A1. Źródło własne [W3] (Rys. 8).

### 2.2.4. Wyniki porównania metodologii śledzenia wzroku i migającego bodźca

Na Rys. 2.15 zestawiono ze sobą wyniki pomiarów charakterystyki GDCSF i pomiary uzyskane w dodatkowym eksperymencie percepcyjnym, w którym orientacja bodźca nie zmieniała się po niezamierzonej zmianie kierunku patrzenia. Miało to na celu zwerifikowanie metodologii wykorzystania eye trackera do kontroli obszaru padania obrazu bodźca wzrokowego na siatkówkę oka. Niebieski wykres przedstawia wyniki uzyskane dzięki zastosowaniu metodologii, natomiast zielony pokazuje pomiar w zmodyfikowanym eksperymencie. Jak można zauważyć, wartości progów kontrastu dla odległości kątowych  $20^\circ$  i  $27^\circ$  są niższe niż wartości zmierzone w poprzednim eksperymencie. Na podstawie tej obserwacji można wnioskować, że obserwator był w stanie podejrzeć bodziec wzrokowy w mniejszej odległości kątowej niż zamierzony. Wartości pomiarów uzyskanych w porównywanych metodologiach przedstawione zostały w Tabeli 2.2.

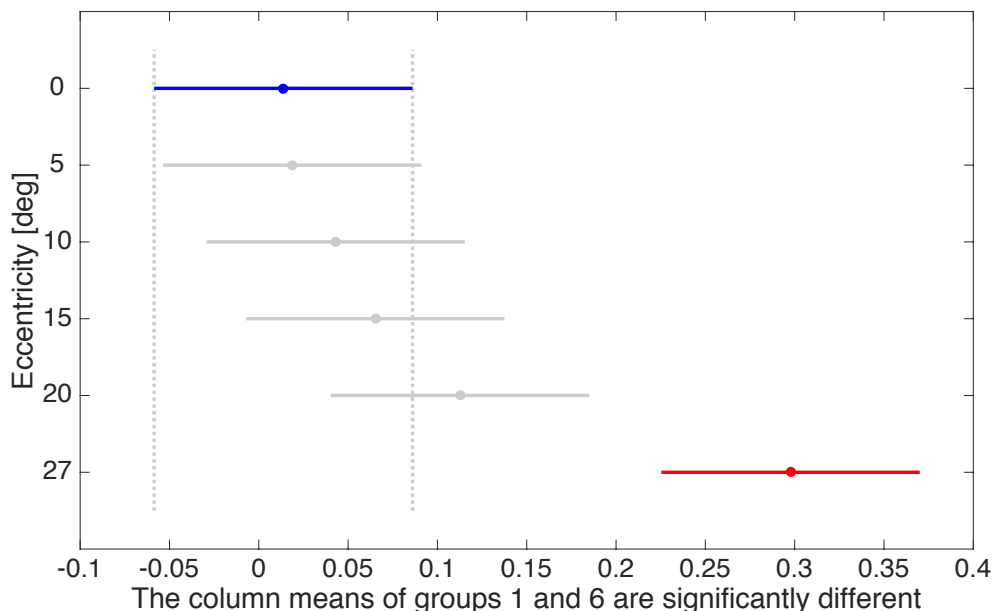


Rysunek 2.15. Progi widzialnego kontrastu dla achromatycznego bodźca wzrokowego o częstotliwości  $2\text{ cpd}$  uzyskane w porównywalnych metodologiach. Pionowe linie przy pomiarach oznaczają błąd standardowy średniej. Źródło własne [W2] (Rys. 3).

Różnica statystyczna między dwoma metodologiami zmierzona została poprzez dwukierunkową analizę wariancji (ANOVA). Zmienną zależną były zmierzone progi kontrastu, natomiast zmiennymi niezależnymi był typ eksperymentu (orientacja statyczna lub losowa) i odległość kątowa od obszaru widzenia centralnego. ANOVA ujawnia istotność statystyczną odległości kątowej ( $p = 0, F = 9.77$ ), rodzaju eksperymentu ( $p = 0.0012, F = 4.57$ ) oraz interakcję pomiędzy odległością kątową i typem eksperymentu ( $p = 0.0207, F = 1.98$ ). Na Rys. 2.16 przedstawiono przedziały ufności dla zmierzonych odległości kątowych. Wykres pokazuje, że wyniki osiągnięte w eksperymentach są znacząco różne dla odległości  $27^\circ$  od obszaru widzenia centralnego.

| Odległość kątowna [°]                                  | 0      | 5     | 10    | 15     | 20    | 27    |
|--|--------|-------|-------|--------|-------|-------|
| Losowa orientacja bodźca wzrokowego [ $\log_{10}$ ]    |        |       |       |        |       |       |
| Próg wykrywalnego kontrastu                            | 0.013  | 0.022 | 0.046 | 0.046  | 0.148 | 0.393 |
| Odchylenie standardowe                                 | 0.013  | 0.023 | 0.027 | 0.028  | 0.249 | 0.477 |
| Wartość minimalna                                      | 0.002  | 0.005 | 0.012 | 0.028  | 0.041 | 0.049 |
| Wartość maksymalna                                     | 0.038  | 0.071 | 0.082 | 0.107  | 0.712 | 1.162 |
| Statyczna orientacja bodźca wzrokowego [ $\log_{10}$ ] |        |       |       |        |       |       |
| Próg wykrywalnego kontrastu                            | 0.016  | 0.014 | 0.036 | 0.079  | 0.074 | 0.181 |
| Odchylenie standardowe                                 | 0.009  | 0.009 | 0.035 | 0.059  | 0.075 | 0.159 |
| Wartość minimalna                                      | 0.005  | 0.004 | 0.009 | 0.02   | 0.022 | 0.044 |
| Wartość maksymalna                                     | 0.028  | 0.034 | 0.115 | 0.159  | 0.235 | 0.455 |
| Różnica wykrytych progów [ $\log_{10}$ ]               | -0.003 | 0.008 | 0.010 | -0.033 | 0.074 | 0.212 |

Tabela 2.2. Wyniki dwóch metodologii pomiarów progów wykrywalnych kontrastów. Źródło własne [W2] (Tab. 1).



Rysunek 2.16. Wyniki porównania metodologii eksperymentu ze statyczną i losowo modyfikowaną orientacją bodźca. ANOVA ujawnia istotność statystyczną różnicy między nimi. Średni próg kontrastu dla każdej odległości kątownej jest reprezentowany przez kropkę, a przedział ufności jest reprezentowany przez linię. Wykryte progi dla różnych odległości są statystycznie różne, gdy ich interwały są rozłączne. Źródło własne [W2] (Rys. 4).

### 2.3. Model czułości na kontrast dla widzenia peryferyjnego

W algorytmach grafiki komputerowej czasu rzeczywistego opartych na kierunkowości widzenia należy tak uzależnić częstotliwość próbkowania sceny, aby pogorszenie jakości nie wpłynęło na percepcyjny odbiór syntezy obrazu przez ludzkiego obserwatora. W tym celu warto wyznaczyć pewną miarę, która wyznaczy wykrywalność wskazanych danych bodźców wzrokowych przez HVS, biorąc pod uwagę ich cechy, tj. luminancja, chrominancja i częstotliwość bodźca oraz odległość rzutu jego obrazu na siatkówce gałki ocznej od obszaru widzenia centralnego. Denes [19] przedstawia taki model jako prawdopodobieństwo wykrycia bodźców wzrokowych, który znajduje się w danym obszarze obrazu. Służy to do sprawdzenia, czy redukcja głębokości kanałów w obszarach obrazu, która zawiera obszary o niskiej częstotliwości tekstur, powoduje powstanie artefaktów pasmowania (ang. *banding artefacts*). Model ten bierze pod uwagę częstotliwość i kontrast bodźca, który rozbijany jest na trzy składowe: jedną achromatyczną (luminancję) i dwie chromatyczne - czerwonozieloną i żółtoniebieską. Brana jest również pod uwagę luminancja tła.

Na podstawie tego modelu zaproponowano w naszych badaniach model czułości na kontrast, który wykazuje prawdopodobieństwo wykrycia bodźca wzrokowego w obrazie, biorąc pod uwagę jego kontrast i odległość od obszaru widzenia centralnego. W odróżnieniu od oryginalnej koncepcji nie bierze on pod uwagę luminancji tła. Co więcej, brane są pod uwagę wyłącznie bodźce o częstotliwości 2 *cpd*, gdyż dla tej wartości zmierzona została charakterystyka czułości na kontrast bodźców achromatycznych i chromatycznych w obszarach widzenia peryferyjnego.

**Procedura** Wyznaczenie prawdopodobieństwa  $P$  danego bodźca wzrokowego podzielone jest na etapy. Procedura rozpoczyna się od wyliczenia wartości kolorów w przestrzeni LMS znajdujących się w dwóch skrajnie oddalonych fragmentach bodźca, które początkowo przedstawione są w przestrzeni XYZ. Konwersja wykonywana jest za pomocą równania liniowego opisanego poniższą formułą.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.15514 & 0.54312 & -0.03286 \\ -0.15514 & 0.45684 & 0.03286 \\ 0 & 0 & -0.00801 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Tak jak wspomniano we wcześniejszych fragmentach tej pracy, przestrzeń LMS wskazuje w przybliżeniu odbierane wartości koloru przez poszczególne typy czopków znajdujących się na siatkówce oka. Wartości te są następnie przekształcane do kierunków kolorów, które są wykorzystywane w pomiarach czułości na kontrast: jeden achromatyczny (od czarnego do białego) i dwa chromatyczne - od czerwonego do zielonego i od żółtego do

niebieskiego. Dokładne przekształcenie wartości odbieranych przez czopki nie jest znane, lecz ten aspekt nie jest kluczowy do wyznaczenia modelu. W tym celu używana jest jedna z najprostszych, a także powszechnie używanych formuł:

$$\begin{bmatrix} A \\ R \\ Y \end{bmatrix} = \begin{bmatrix} L + M \\ L - M \\ S \end{bmatrix},$$

gdzie  $A$  to cecha achromatyczna (luminancja),  $R$  odpowiada części czerwono-zielonej, a  $Y$  żółto-niebieskiej.

W rezultacie opisanych do tej pory operacji otrzymuje się dwa wektory wartości koloru bodźca:  $[A_1, R_1, Y_1]$  i  $[A_2, R_2, Y_2]$ . Na ich podstawie liczony zostaje kontrast poszczególnych wartości składowych. Do tego stosowana jest formuła z modelu Denes'a, na którym przedstawiona koncepcja jest oparta.

$$C_A = \frac{|A_2 - A_1|}{A_1}, C_R = \frac{R_2 - R_1}{\alpha|R_1| + (1 - \alpha)A_1}, C_Y = \frac{|Y_2 - Y_1|}{\alpha|Y_1| + (1 - \alpha)A_1},$$

gdzie  $\alpha$  jest zmienną z przedziału  $[0, 1]$ , którą należy zoptymalizować pod kątem otrzymanych wartości progowych kontrastu uzyskanych w eksperymencie percepcyjnym.

Uzyskane wartości kontrastu należy podzielić przez zmierzoną czułość bodźców na kontrast, które odpowiadają adekwatnym składowym, znajdującym się w przestrzeni ARY, uzyskując w ten sposób znormalizowany kontrast. Wyliczone to jest poniższą formułą:

$$c_k = \frac{C_k}{CSF(ch_{rgb}, \delta)},$$

gdzie  $C_k$  to kontrast wyliczony we wcześniejszym kroku, a  $CSF$  to wartość progu kontrastu dla bodźca o kolorze  $ch_{rgb}$  i odległości kątowej  $\delta$  od obszaru widzenia centralnego.

Następnie znormalizowane kontrasty są przekształcane w prawdopodobieństwo za pomocą funkcji psychometrycznej:

$$P_k = 1 - \exp(\ln(0.5)c_k^\beta),$$

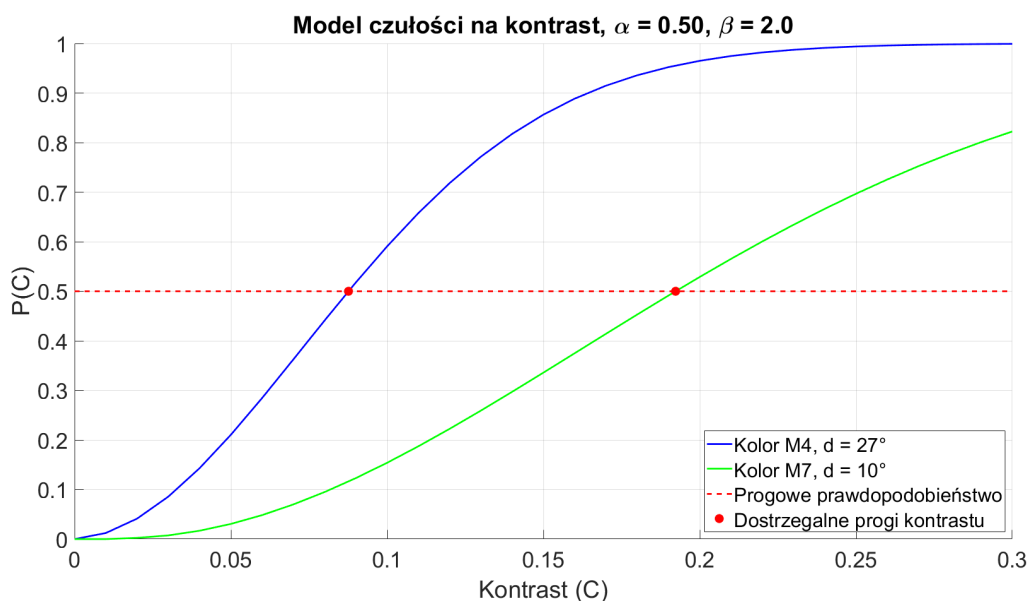
gdzie  $\beta$  to nachylenie funkcji psychometrycznej. Standardowo za tę wartość przyjmuje się  $\beta = 3.5$ , jednak w proponowanym modelu przyjęto  $\beta = 2$ , co dało lepsze dopasowanie do zmierzonych w eksperymencie wartości progowych kontrastu.

Na koniec łączy się uzyskane wartości  $P_k$  dla wszystkich składników kontrastu bodźca za pomocą sumowania prawdopodobieństw:

$$P = 1 - \prod_k (1 - P_k)$$

Uzyskana wartość wskazuje na tendencje HVS do wykrycia bodźca wzrokowego. Jeśli wynosi ona poniżej  $P = 0.5$  bodziec wzrokowy jest niewidoczny dla danej odległości kątownej od obszaru widzenia centralnego.

**Przykład** Wykres prawdopodobieństwa wykrycia danego bodźca wzrokowego przedstawiono jako funkcję jego kontrastu na Rys. 2.17. Ma ona kształt funkcji psychometrycznej. Dla zerowego kontrastu funkcja ta ma wartość zero. Wraz ze wzrostem kontrastu bodźca, funkcja zwiększa swoją wartość z pewnym nachyleniem uzależnionym od parametru  $\beta$ , aż w końcu uzyskuje wartość jeden. Z perspektywy opracowanego modelu czułości na kontrast, ważnym punktem w funkcji psychometrycznej jest  $P(C) = 0.5$ , która wskazuje dostrzegalny próg kontrastu.



Rysunek 2.17. Przykład modelu jako prawdopodobieństwo wykrycia danego bodźca wzrokowego w funkcji jego kontrastu. Czerwona przerywana linia przedstawia wartość prawdopodobieństwa tożsamą z progową wartością dostrzegalnego kontrastu.

**Optymalizacja modelu** Procedura wyznaczania progów widzialnych kontrastów, wymaga dopasowania wartości do uzyskanych w eksperymencie wyników. Pozwala na to dostosowanie parametru  $\alpha$  w modelu, który odpowiedzialny jest za chromatyczne składowe w kontraście. Dopasowanie to może się odbywać poprzez minimalizację funkcji celu, która

w tym przypadku będzie różnicą między uzyskanymi progami kontrastu w pomiarach, a tymi wyliczonymi za pomocą modelu.

Proces dopasowania modelu do wyników odbył się za pomocą metody Nelder-Meada (sympleksowej metody spadku, ang. *downhill simplex method*) [46]. Zmienną niezależną w procesie optymalizacji była zmienna  $\alpha$ , natomiast funkcją celu był błąd średniokwadratowy między progami uzyskanymi w modelu i pomiarach. Opisany jest on formułą:

$$D = \frac{1}{n} \sum_i^n (C_m - C_p)^2,$$

gdzie  $C_m$  oznacza próg dostrzegalnego kontrastu wyliczony w modelu, natomiast  $C_p$  to kontrast uzyskany w eksperymencie percepcyjnym. Błąd średniokwadratowy wyznaczany był dla progów dostrzegalności czterech mieszanych bodźców wzrokowych - M4, M5, M6 i M7 przy odległościach kątowych od obszaru widzenia centralnego równych  $0^\circ$ ,  $5^\circ$ ,  $10^\circ$ ,  $15^\circ$ ,  $20^\circ$  oraz  $27^\circ$  i częstotliwości  $2 \text{ cpd}$ . Pomiarów progów dla bodźców achromatycznych i chromatycznych posłużyły w modelu jako funkcje CSF użyte w formule wyznaczającej znormalizowany kontrast. Bodziec A1 stosowany został w składowej achromatycznej, natomiast C2 i C3 do składowych chromatycznych, kolejno czerwono-zielonej oraz żółto-niebieskiej.

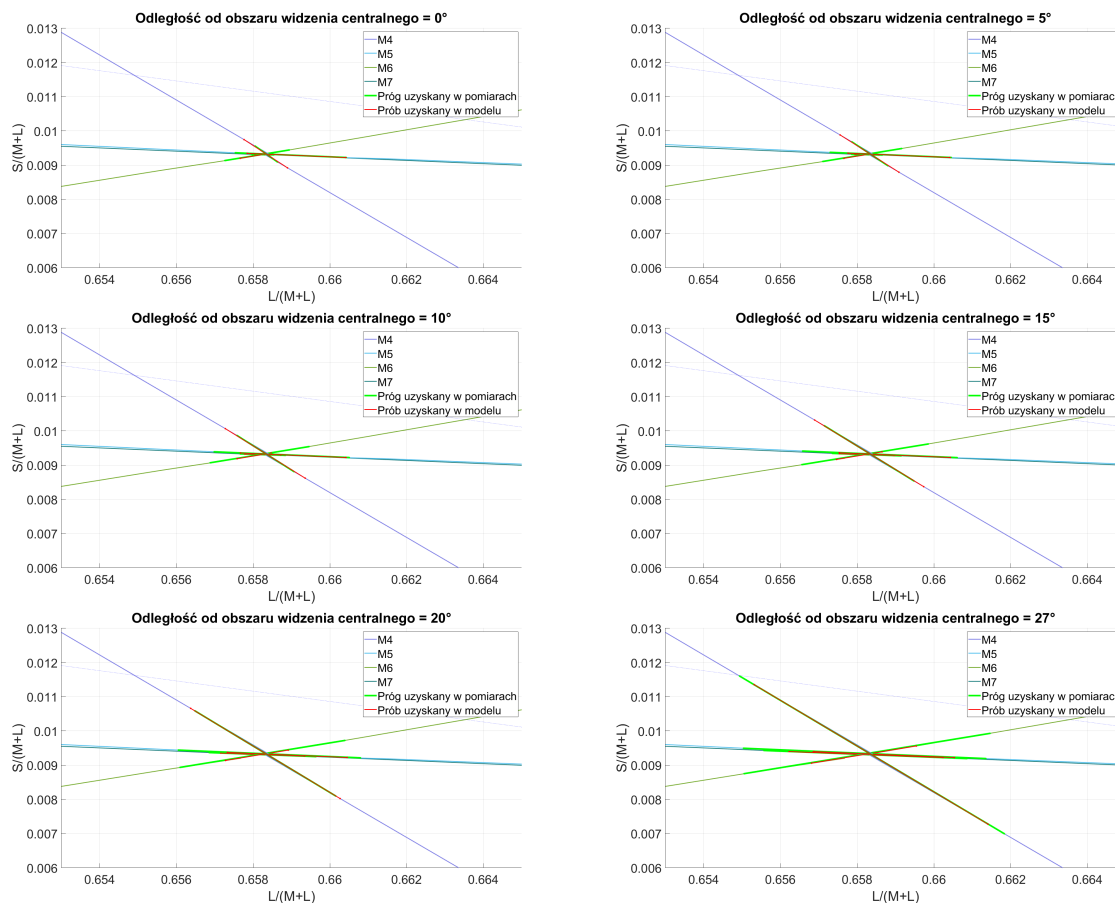
**Rezultaty** Optymalizacja modelu wskazała, że najlepsze dopasowanie do uzyskanych pomiarów występuje dla parametru  $\alpha = 0.96$ , dla którego błąd średniokwadratowy modelu wynosi ok. 0.00098. Wyniki dopasowania wraz z wartościami progowymi kontrastu przedstawione zostały w Tab. 2.3.

| Odległość kątowa  | $0^\circ$ | $5^\circ$ | $10^\circ$ | $15^\circ$ | $20^\circ$ | $27^\circ$ |
|---|-----------|-----------|------------|------------|------------|------------|
| Progi dostrzegalnych kontrastów wyliczone za pomocą modelu            |           |           |            |            |            |            |
| Bodziec M4  | 0.0028    | 0.0054    | 0.0094     | 0.0156     | 0.0253     | 0.0486     |
| Bodziec M5  | 0.0010    | 0.0015    | 0.0058     | 0.0124     | 0.0224     | 0.0460     |
| Bodziec M6  | 0.0125    | 0.0148    | 0.0174     | 0.0213     | 0.0270     | 0.0403     |
| Bodziec M7  | 0.0586    | 0.0774    | 0.1031     | 0.1381     | 0.1858     | 0.2830     |
| Progi dostrzegalnych kontrastów uzyskane w eksperymencie percepcyjnym |           |           |            |            |            |            |
| Bodziec M4  | 0.0048    | 0.0078    | 0.0126     | 0.0204     | 0.0330     | 0.0648     |
| Bodziec M5  | 0.0234    | 0.0316    | 0.0428     | 0.0578     | 0.0782     | 0.1192     |
| Bodziec M6  | 0.0212    | 0.0279    | 0.0368     | 0.0485     | 0.0639     | 0.0939     |
| Bodziec M7  | 0.0323    | 0.0494    | 0.0754     | 0.1151     | 0.1758     | 0.3180     |
| Błąd średniokwadratowy  | 0.0003    | 0.0005    | 0.0006     | 0.0008     | 0.0011     | 0.0025     |

Tabela 2.3. Porównanie progów dostrzegalnych kontrastów zmierzone w eksperymencie percepcyjnym i wyliczone za pomocą modelu.

Graficzna interpretacja uzyskanych progów widocznych kontrastów uzyskana za pomocą przedstawionego modelu znajduje się na Rys. 2.18. Na wykresie widoczna jest

niedokładność dopasowania osiągniętych progów widocznych kontrastów do tych zmierzonych. Widoczne jest to najbardziej w przypadku mieszanego bodźca wzrokowego M5. Widoczny jest jednak wzrost progu kontrastu wraz z oddalaniem się od obszaru widzenia centralnego. Stopnie wzrostu progu kontrastu są zgodne dla modelu i pomiarów. Można także zauważyć tendencję do zaniżonych progów kontrastu wyliczonych w modelu w porównaniu ze zmierzonymi, z wyjątkiem bodźca M7. Ta cecha modelu może posłużyć jako bezpieczny margines dostrzegalnych wartości kontrastu w systemach renderingu.

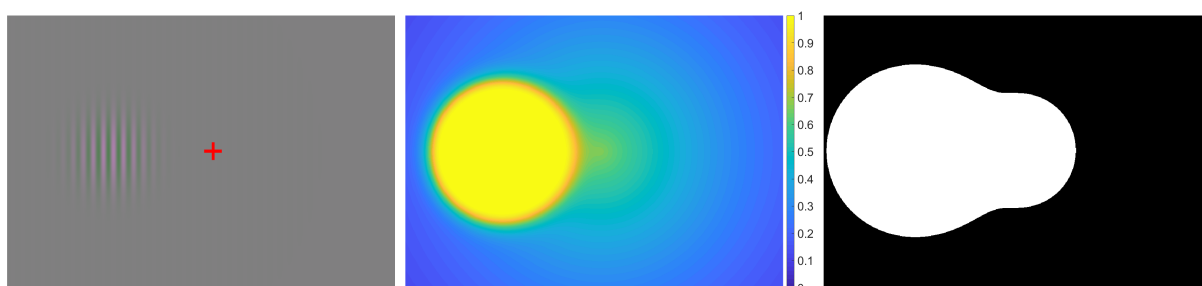


Rysunek 2.18. Graficzna prezentacja progów dostrzegalnych kontrastów w postaci schematu Boyntona-MacLeoda. Na kierunkach koloru bodźców M4-M7 zieloną linią oznaczono wartości zmierzone w eksperymencie percepcyjnym, natomiast czerwona linia wskazuje wartości wyliczone za pomocą modelu.

Dzięki przedstawionemu modelowi można wyznaczyć mapę prawdopodobieństwa, która umożliwi dostrzeżenie bodźca wzrokowego w danym obszarze. Przykład takich map zaprezentowany został na Rys. 2.19. W pierwszej kolejności obraz należy rozłożyć na bodźce wzrokowe o różnych częstotliwościach. Z powodu wykonanych pomiarów wyłącznie dla  $2\text{ cpd}$ , przykład ograniczony jest do takiego bodźca. Następnie dla każdej częstotliwości i obszaru wyznaczany jest kontrast Michelsona, który wykorzystany jest do wyliczenia



prawdopodobieństwa  $P$ . W kolejnym kroku następuje progowanie prawdopodobieństwa przez wartość 0.5, powyżej której bodźce wzrokowe są dostrzegalne.



(a) Bodziec M5,  $f = 2cpd$ ,  $C_m = 0.35$ ,  $d = 5^\circ$ .



(b) Bodziec M7,  $f = 2cpd$ ,  $C_m = 0.11$ ,  $d = 10^\circ$ .

Rysunek 2.19. Przykłady map prawdopodobieństwa modelu czułości na kontrast. Po lewej przedstawiony jest bodziec wraz z punktem skupienia wzroku obserwatora zaznaczony jako czerwony znacznik. Na środku znajduje się mapa prawdopodobieństwa wykrycia danego bodźca, po prawej zaś progowana mapa według wartości  $P = 0.5$ , poniżej której piksele zostały zaznaczone na czarno, a powyżej na biało.

Przedstawiony model analityczny zdolności HVS do postrzegania bodźców wzrokowych jest koncepcją, która może mieć znaczące zastosowanie w przyszłych badaniach nad systemami renderingu, uwzględniających kierunek patrzenia obserwatora. Rozwiązanie to należy w przyszłości zweryfikować poprzez porównanie wartości modelu ze zmierzonymi progami dostrzegalnego kontrastu dla większej liczby kierunków koloru, bardziej różniących się od siebie. Większa liczba pomiarów może także zwiększyć stopień dopasowania modelu do pomiarów. Kolejnym kierunkiem badań nad opisanym modelem jest wykorzystanie go w systemie renderingu czasu rzeczywistego. Model ten może posłużyć do adaptacyjnego wyznaczania maski próbkowania w algorytmie śledzenia promieni, bądź do wyznaczenia mniej widocznych obszarów w trójwymiarowej siatce obiektu i redukcji liczby wielokątów w niej zawartej.

## 2.4. Dynamika percepcji dla zmiany kierunku widzenia

Układ wzrokowy człowieka oprócz wykrywania zmian luminancji i kolorów statycznych bodźców wzrokowych jest także w stanie wykryć zmianę wartości ich cech, które występują po czasie. Fale światła odbite od obserwowanej sceny są rzutowane na siatkówkę, w której to fotoreceptory pobierają je w sposób okresowy. Informacja ta jest integrowana do postaci obrazu, stąd wrażenie, że obserwowane przez nas obiekty wydają się stabilne lub płynnie przemieszczają się po obserwowanej scenie. Ponieważ do zbierania i przetwarzania informacji potrzebny jest skończony czas, istnieją ograniczenia w reagowaniu naszego systemu wizualnego na tempo zmian [42]. Zdolność układu wzrokowego człowieka do postrzegania bodźca przerywanego, tj. przełączającego się pomiędzy stanem włączonym i wyłączonym, jest zależna od tempa zmian stanu. Gdy tempo to będzie wolne, obserwator zauważy zaprezentowany mu bodziec, lecz wraz ze wzrostem częstotliwości zmian wywołane zostanie wrażenie migotania (ang. *flickering*) bodźca. Powyżej określonej wartości krytycznej, zwanej krytyczną częstotliwością migotania, migotanie ustaje.

System renderingu czasu rzeczywistego, który oparty jest na kierunkowości patrzenia powinien generować i prezentować obserwatorowi kolejne klatki obrazu szybciej niż wynosi krytyczna wartość migotania. Gdy bodziec zawarty w scenie, którego parametry są zależne od kierunku patrzenia, zawiera się w obszarze peryferyjnym i wzbudzi uwagę obserwatora, może spowodować zmianę jego kierunku patrzenia. Powoduje to konieczność przerysowania całego obrazu, uwzględniając zmianę parametrów, a tym samym nową reprezentację graficzną bodźca. Wykrycie zmiany kierunku patrzenia przez system renderingu i wygenerowanie klatki obrazu zależnej od tego kierunku powoduje opóźnienie czasu działania systemu względem układu wzrokowego człowieka. Jeśli ten odstęp czasu będzie zbyt długi, obserwator dostrzeże w obszarze centralnym swojego widzenia elementy sceny będące w reprezentacji graficznej przeznaczone dla obszaru peryferyjnego, co spowoduje pogorszenie jakości odbioru sceny przez obserwatora. Akceptowalnym opóźnieniem systemu renderingu czasu rzeczywistego opartego na kierunku patrzenia, można nazwać takie, w którym przerysowanie to nie będzie dostrzegalne przez obserwatora.

### 2.4.1. Eksperyment percepcyjny

**Poziomy szczegółowości obiektów.** W celu pomiaru akceptowalnego opóźnienia został zaimplementowany system renderujący używający technikę *Poziomu Szczegółowości* (ang. *Level of Detail - LOD*) trójwymiarowych obiektów sceny. Polega to na zawarciu wewnątrz algorytmu renderującego szeregu różnych reprezentacji siatek geometrycznych trójwymiarowych obiektów, które mają zostać wyrenderowane. Każda dodatkowa siatka niebędąca oryginalną jest jej uproszczoną wersją zbudowaną ze zredukowanej liczby wielokątów, posiadająca ogólne cechy obiektu, który ma reprezentować, ale z wyciętymi szcze-

głównymi elementami, w które był oryginalnie wyposażony. Siatki te są stosowane, gdy chce się zredukować ilość obliczeń potrzebnych do przetworzenia geometrii, tym samym zredukować czas potrzebny do jego wyrenderowania [14]. Są one powszechnie stosowane w grach komputerowych, gdzie stosowany poziom geometrii obiektów sceny jest zależny od dystansu, który dzieli go od kamery. Gdy obiekt jest umiejscowiony na dalszym planie, zastosowanie uproszczonej geometrii obiektów nie wpłynie na jego percepcyjny odbiór, gdyż wycięte elementy szczególne w geometrii obiektu nie byłyby dostrzegalne, gdy cały obiekt zajmuje mały obszar w wygenerowanym obrazie. Technika ta jest używana w naszych badaniach, lecz poziom uproszczenia zależny jest od kierunku patrzenia.

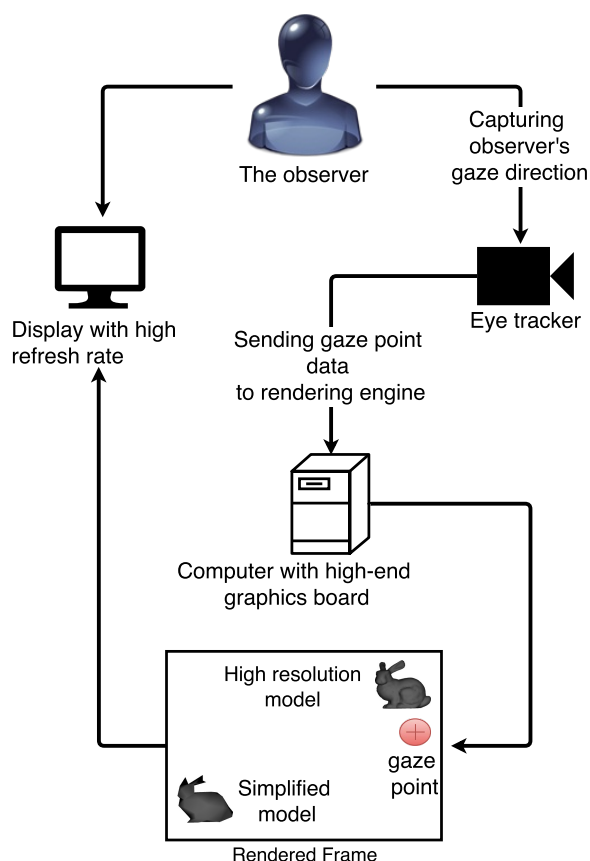
### System renderingu kierunkowego

Na podstawie zaimplementowanego systemu renderingu kierunkowego z LOD (schemat na Rys. 2.20) przeprowadzony został eksperyment percepcyjny, w którym zaprezentowano uczestnikom opisaną technikę. Mieli oni za zadanie obserwować wyświetlone na scenie obiekty, których reprezentacja geometryczna jest zależna od ich kierunku patrzenia. W trakcie trwania eksperymentu uczestnicy wielokrotnie byli proszeni o zmianę kierunku patrzenia w taki sposób, aby wyświetlone obiekty przemieściły się z obszaru peryferyjnego na obszar centralny i ocenili czy są w stanie dostrzec uproszczenie geometrii stosowane w związku z położeniem ich w obszarze centralnym. Dzięki zastosowaniu różnych wartości czasowych opóźnienia systemu renderingu, możliwe było zbadanie, jak bardzo zmiana geometrii jest dostrzegalna przez układ wzrokowy człowieka wraz ze zmianą tej wartości czasowej.

Eye tracker zintegrowany w system rejestruje kierunek patrzenia obserwatora wyliczając punkt, na który patrzy w przestrzeni wyświetlanego obrazu. Punkt ten jest wysyłany do komputera, który ma za zadanie wygenerować klatkę obrazu sceny. Zawierające się w niej obiekty posiadają kilka reprezentacji geometrii o różnej złożoności. Wybór siatki obiektów użytych do renderingu jest zależny od punktu patrzenia obserwatora. Gdy obiekt jest zawarty w obszarze widzenia centralnego, wybierana jest oryginalna geometria, natomiast gdy zawiera się w obszarze widzenia peryferyjnego, stosowana jest siatka o ograniczonej liczbie trójkątów. Wygenerowany obraz jest wysyłany do wyświetlacza, który prezentuje go obserwatorowi.

### Procedura

Eksperyment odbywał się na przystosowanym do systemu renderingu stanowisku wyposażonym w podpórkę na podbródek, eye tracker i wyświetlacz (Rys 2.21). W trakcie trwania eksperymentu zostały wyłączone wszystkie zewnętrzne źródła światła, dzięki czemu oświetlenie pochodziło wyłącznie z wyświetlacza. Uczestnicy eksperymentu byli proszeni o umieszczenie głowy w wyznaczonym na nią miejscu, co pozwalało na ustabilizowanie głowy w celu zwiększenia skuteczności działania eye trackera, a także ustawienie



Rysunek 2.20. Schemat systemu renderingu kierunkowego z techniką LOD. Źródło własne [W1] (Rys. 2).

oczu obserwatora w stałej ustalonej odległości 75 cm od wyświetlacza. Przed samym eksperymentem następowała 9-punktowa kalibracja eye trackera, która jest wymagana do poprawnego działania urządzenia. Procedura ta trwała około 20 sekund i wiązała się z obserwacją znaczników wyświetlanych w różnych obszarach wyświetlacza. Dzięki przetworzeniu dostarczonych danych w trakcie kalibracji, oprogramowanie eye trackera mogło wyliczyć pozycję wzroku obserwatora w przestrzeni wyświetlacza podczas trwania eksperymentu.

Główny etap eksperymentu rozpoczynał się od wyświetlenia czerwonego krzyżyka przedstawionego na szarym tle na poziomie 18% maksymalnej luminancji wyświetlacza. Obserwator był poproszony, aby spojrział na znacznik (Rys 2.22(a)), po czym gdy eye tracker wykrył, że patrzy on na krzyżyk, po czasie 500 ms w ustalonej odległości kątowej, 10°, 25° lub 35° wyświetlone zostały dwa obiekty. Jeden z wyświetlonych obiektów, zwany referencyjnym, został wyrenderowany wykorzystując oryginalną siatkę zachowującą wszystkie szczegóły obiektu. Z kolei drugi obiekt wygenerowany został przy wykorzystaniu siatki z uproszczoną geometrią o ograniczonej liczbie wielokątów, gdzie poziom uproszcze-



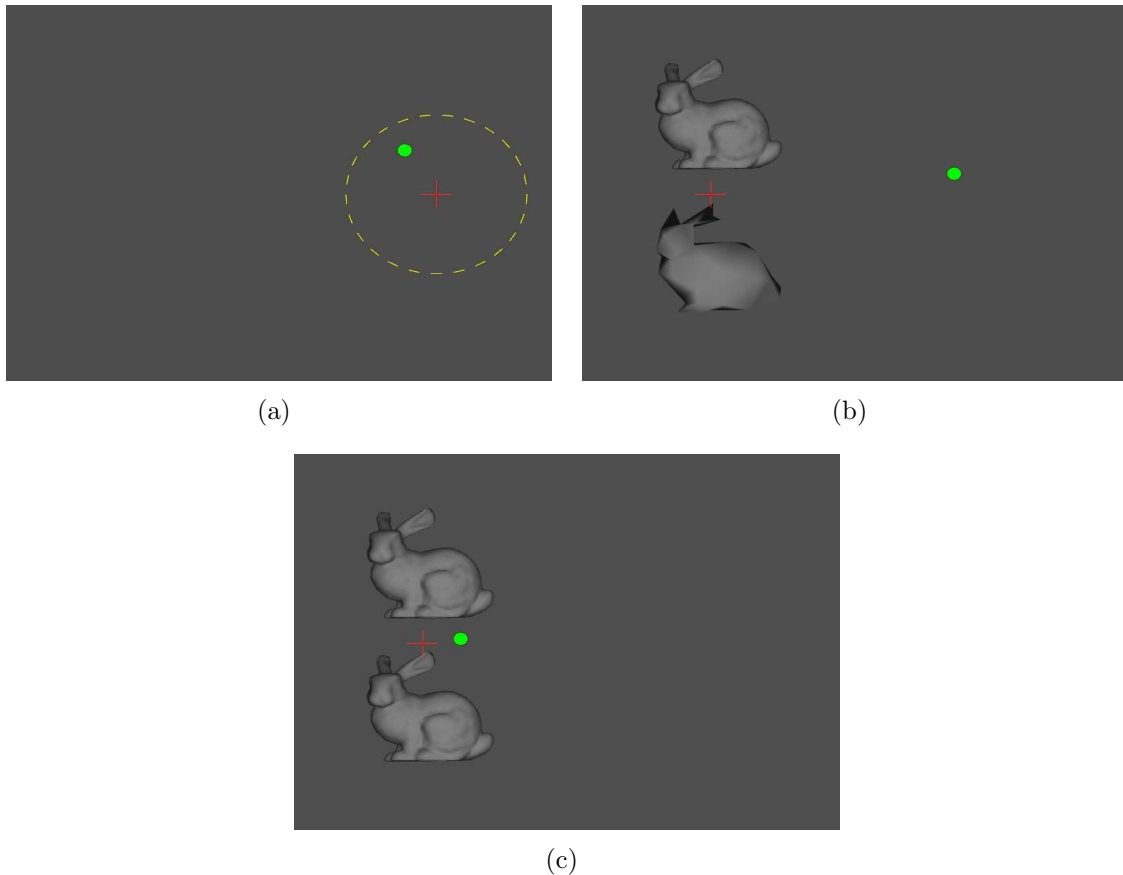
Rysunek 2.21. Stanowisko komputerowe, na którym odbywał się pomiar akceptowalnego opóźnienia systemu renderingu uwzględniającego kierunek patrzenia. W trakcie trwania eksperymentu oświetlenie stanowiska było wyłączane, jedynym źródłem światła był obraz pochodzący z wyświetlacza.

nia był zależny od odległości kątowej między ustawionym znacznikiem gdzie docelowo obserwator powinien skupić swój wzrok, a wyświetlonymi obiektami. Dwa obiekty były wyświetlone jeden pod drugim (Rys 2.22(b)), a kolejność wyświetlania tych obiektów była losowana na początku eksperymentu. Ostatnim etapem badania było poproszenie obserwatora o zmianę kierunku patrzenia na wyświetlone obiekty i wskazanie za pomocą odpowiedniego przycisku na klawiaturze (strzałka w górę/doł), który z dwóch obiektów posiada uproszczoną geometrię. W momencie, gdy eye tracker wykrył zmianę kierunku patrzenia do momentu, gdy pozycja jego wzroku w przestrzeni ekranu była oddalona o  $4^\circ$  od wskazanego wcześniej znacznika, obiekt uproszczony został przerysowany, wykorzystując referencyjną siatkę, sprawiając, że obydwa obiekty posiadały tę samą geometrię (Rys 2.22(c)).

Procedura ta została powtórzona po 30 razy dla trzech odległości kątowych oraz czterech częstotliwości odświeżania wyświetlacza równych 30 Hz, 60 Hz, 120 Hz i 144 Hz. Odpowiada to opóźnieniom 33 ms, 16 ms, 8 ms i 7 ms. Za każdym razem kolejność wyświetlania obiektu referencyjnego oraz uproszczonego była losowana.

### **Bodźce wzrokowe**

W eksperymencie użyty został obiekt Stanford Bunny (Rys. 2.23), który jest powszechnie używanym modelem w badaniach naukowych związanych z grafiką komputerową [55]. Za referencyjną geometrię siatki przyjęta została taka, która zawiera 4000 trójkątów. Decyzja ta wynika z faktu, iż obiekty stosowane w dzisiejszych standardach grafiki czasu rzeczywistego zawierają podobną ilość wielokątów, oraz to, że przedstawia ona obiekt w formie, w której nadal widoczne są jego cechy i szczegóły. Siatki geometrii obiektów uproszczonych zostały wygenerowane używając algorytmu Quadric Edge Col-

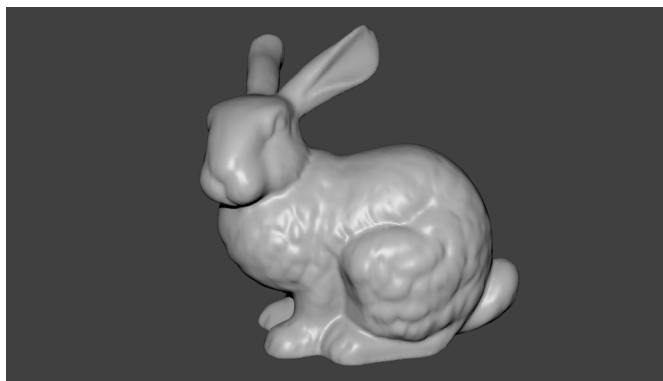


Rysunek 2.22. Przykładowa procedura eksperymentu mierzącego akceptowalne opóźnienie systemu renderingu opartego o kierunek patrzenia. Zielona kropka przedstawia położenie punkt skupienia wzroku obserwatora.

lapse Decimation w oprogramowaniu MeshLab [13]. Łącznie zostało wygenerowanych 40 geometrii, z każdą iteracją redukując ilość trójkątów o 100.

Złożoność uproszczonych siatek obiektów, które były użyte w omawianym eksperymencie, zostały dobrane za pomocą dodatkowego eksperymentu percepcyjnego, który został przeprowadzony przed pomiarem akceptowalnego opóźnienia systemu. Procedura ta była podobna do wspomnianego wcześniej eksperymentu z tą różnicą, że obserwator miał za zadanie przez cały czas jego trwania patrzeć na znacznik znajdujący się w ustalonej odległości od obiektów. W momencie zmiany kierunku patrzenia obserwatora na obiekty, znikaly one do momentu, gdy wzrok obserwatora nie skupił się ponownie na tym znaczniku. Zadaniem uczestnika było wskazanie, który obiekt (górny czy dolny) zbudowany jest z uproszczonej geometrii.

Eksperyment ten został powtórzony po 10 razy dla geometrii złożonych od 100 do 2500 trójkątów oraz odległości kątowych równych  $10^\circ$ ,  $20^\circ$  i  $35^\circ$ . Na podstawie uzyskanych rezultatów (Rys. 2.24) zostały dobrane reprezentacje uproszczonych siatek użytych w docelowym eksperymencie, różne dla trzech odległości kątowych, kolejno 1000, 1600 i 2000 trójkątów (Rys. 2.25).



Rysunek 2.23. Stanford Bunny jest to model trójwymiarowy opracowany przez skan Grega Turka i Marca Levoya w roku 1994 na Uniwersytecie Stanforda. Wygenerowany został za pomocą skanu figurki ceramicznej królika w celu przetestowania metod skanowania fizycznych obiektów [98]. Oryginalnie zeskanowany model zawiera 69541 trójkątów.

### Aparatura

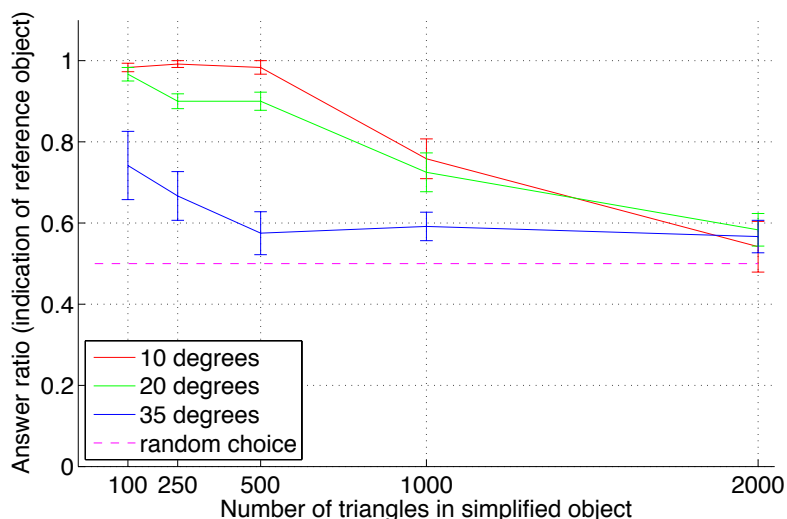
Stworzony na potrzeby eksperymentu system renderingu został wyposażony w Mirametrix S2 - przenośny eye tracker położony pod wyświetlaczem, który działa z częstotliwością 60 Hz. Urządzenie wymagało kalibracji przed każdą sesją, w której było używane. Po poprawnym przeprowadzeniu tego procesu oprogramowanie dołączone do S2 wysyłało pozycję punktu patrzenia w przestrzeni wyświetlacza, na którym przeprowadzona była kalibracja, korzystając z protokołu TCP/IP (Rys 2.26). Podczas przeprowadzonych sesji eksperymentów udało nam się uzyskać dokładność bliską 1 stopnia kątownego patrzenia, co było wystarczające dla naszych badań.

W celu wyrenderowania obiektów z użyciem techniki LOD zostało zaimplementowane oprogramowanie przy użyciu języka programowania C++ oraz niskopoziomowego API OpenGL wraz z rozszerzeniami GLEW i GLFW. Pozwoliło to na osiągnięcie czasu generowania obrazu sceny zawierającej obiekty składające się z ponad 4000 trójkątów w czasie krótszym niż 1 ms. Oprogramowanie wspierało oświetlenie Phong'a oraz 16-próbkowy Multi-Sampling Anti-Aliasing (16x MSAA). Oprogramowanie to było uruchomione na komputerze wyposażonym w kartę graficzną NVIDIA 780 GTX.

Do wyświetlenia wygenerowanego obrazu posłużył 22-calowy ASUS ROG SWIFT PG278Q LCD, wyświetlacz przeznaczony dla graczy, który wyświetla obraz w wystarczająco dla naszych badań dobrej jakości o rozdzielczości 2560 x 1440 pikseli. Wyświetlacz ten wyposażony jest w technologię G-Sync, która w trakcie naszych badań gwarantowała, że częstotliwość odświeżania jest stała.

### Uczestnicy eksperymentu

Eksperyment został wykonany na grupie 10 ochotników, których wiek wynosił od 20 do 23 lat. Wśród nich znajdowało się 8 mężczyzn i 2 kobiety. Ochotnicy oświadczy-

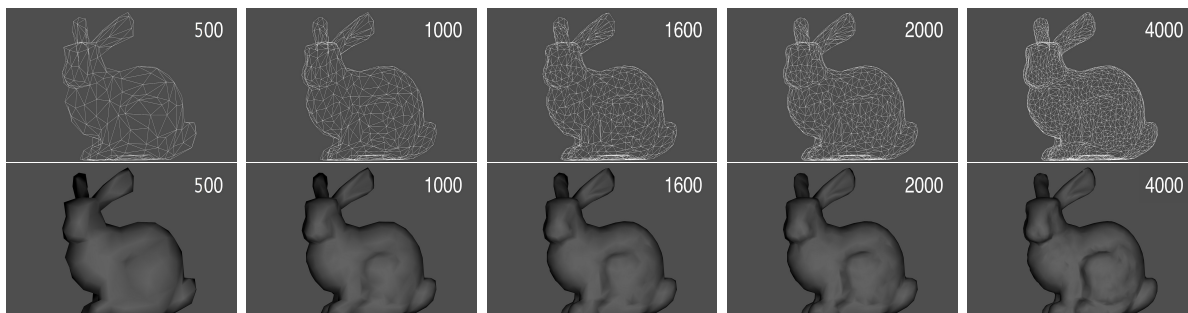


Rysunek 2.24. Wyniki pomiaru zdolności obserwatorów do rozróżnienia obiektu referencyjnego od uproszczonych jako funkcja znormalizowanego współczynnika poprawnych odpowiedzi do liczby trójkątów uproszczonej siatki. Różowa przerywana linia wskazuje sytuację, gdy liczba poprawnych odpowiedzi jest równa odpowiedziom błędnym, co wskazuje na niemożliwość rozróżnienia dwóch geometrii przez obserwatora dla danej odległości kątowej.

li, że posiadają zdrowy bądź skorygowany układ wzrokowy. Uczestnicy byli świadomi co powinni zrobić, ale nie mieli pojęcia na temat celu eksperymentu. Średni czas sesji eksperymentalnej wynosił około 12 minut.

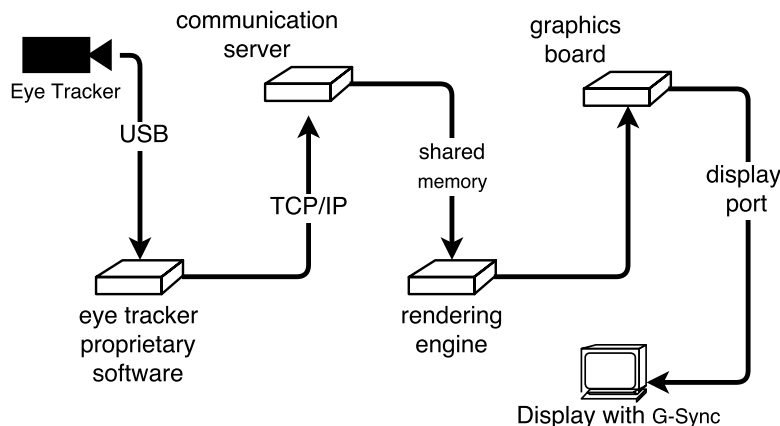
#### 2.4.2. Rezultaty eksperymentu

Wyniki eksperymentu przedstawiono na Rys 2.27. Wykres pokazuje znormalizowany stosunek prawidłowych odpowiedzi (poprawne wskazania na obiektach uproszczonych) do liczby prób w funkcji szybkości wyświetlania klatek. Wartość 0.5 jest równoważna losowemu wyborowi, tj. wskazuje na niezdolność do rozróżnienia modelu referencyjnego i modeli uproszczonych.



Rysunek 2.25. Siatki obiektów użytych w pomiarach opóźnienia percepcji wzroku. Na górze wersje siatkowe - zawierające tylko wierzchołki trójkątów, z których zbudowany jest model, na dole wersja cieniowana - wyświetlona obserwatorowi z wypełnionymi trójkątami i zastosowanemu modelowi oświetlenia Phong'a [8].





Rysunek 2.26. Schemat budowy systemu renderingu LOD uwzględniającego kierunkowość widzenia. Poprzez interfejs USB 3.0 do komputera, na którym uruchomione zostało oprogramowanie renderujące, podłączony był Eye tracker. Karta graficzna wysyłała do wyświetlacza wygenerowany obraz korzystając z interfejsu Display Port. Źródło własne [W1] (Rys. 3).

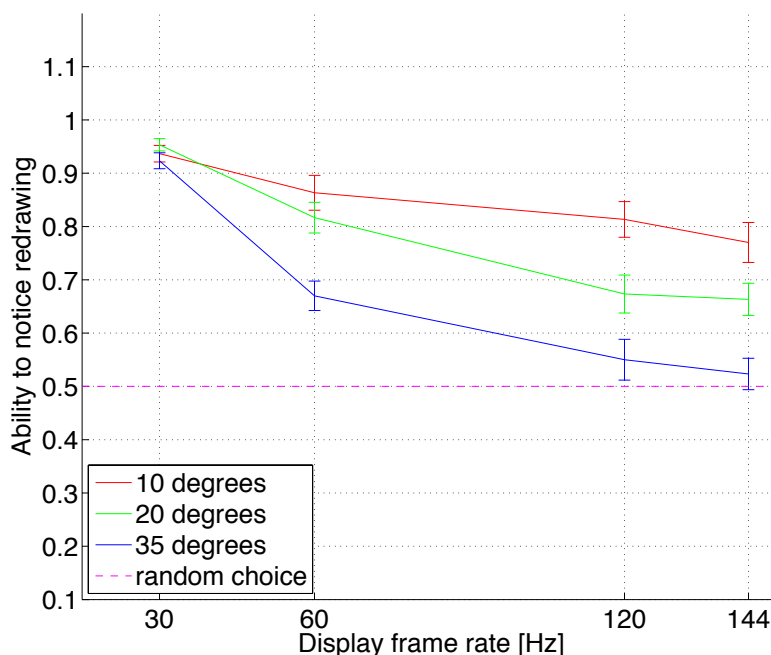
W naszym badaniach tylko dla częstotliwości odświeżania ekranu równej 144 Hz oraz odległości kątowej  $35^\circ$  wyniki są bliskie wskazania na niezdolność do obserwacji przerysowania. Należy zwrócić uwagę na fakt, że wartość ta jest lekko powyżej 0.5. To wskazuje na to, że w niektórych przypadkach wystąpiło zauważalne przerysowanie. Dla innych częstotliwości odświeżania wyświetlacza oraz odległości kątowych opóźnienie systemu było zbyt długie, aby zapewnić niezauważalną zmianę LOD. Szczególnie w przypadku mniejszych kątów widzenia przerysowanie jest wyraźnie widoczne.

Należy zwrócić uwagę na fakt, że wyniki wskazują tendencje do redukcji wykrywalności przerysowania obiektu wraz ze zwiększeniem częstotliwości odświeżania wyświetlacza.

### 2.4.3. Wskazania dla aplikacji

Mimo, że dla odległości kątowej  $35^\circ$ , opóźnienie systemu renderingu jest bliskie akceptacji dla częstotliwości odświeżania wyświetlacza równej 144 Hz, co przekłada się na czas 7 ms, jednak do tego czasu należy dodać czas działania eye trackera. Wykorzystane przez nas urządzenie działa z częstotliwością 60 Hz, co dodaje 17 ms opóźnienia systemu, które łącznie wynosi 24 ms. Na dzień dzisiejszy istnieją szybsze wyświetlacze oraz eye trackery, co pozwala na przeprowadzenie dodatkowych badań z użyciem nowego sprzętu i odnalezienie akceptowalnego opóźnienia systemu renderingu, także dla mniejszych odległości kątowych, na co wskazuje tendencja udowodniona w badaniach.

Trzeba także zwrócić uwagę na fakt, że eksperyment był przeprowadzany w konserwatywnych warunkach. Uczestnicy eksperymentu, będąc zapytany o jakość działania systemu, komentowali fakt, że częściej widzieli proces przerysowania obiektu, czyli efekt migotania bodźca, a nie uproszczoną geometrię jednego z nich. W sytuacji wykorzystania



Rysunek 2.27. Wyniki eksperymentu percepcyjnego dotyczącego pomiaru akceptowalnego opóźnienia renderingu uwzględniającego kierunkowość widzenia. Różowa przerywana linia, wyrażająca wartość 0.5, wskazuje sytuację, gdy użytkownik nie był w stanie dostrzec przerysowania geometrii podczas zmiany kierunku patrzenia. Źródło własne [W1] (Rys. 8).

omawianego systemu w złożonym środowisku grafiki czasu rzeczywistego, takim jak gry komputerowe (gdzie sceny w niej zawarte pełne są innych obiektów, które są statycznych, bądź też dynamicznie przemieszczających się), zdolność do rozpoznania modyfikacji geometrii obiektów podczas zmiany kierunku patrzenia zostałaaby przytłumiona przez inne bodźce.

#### 2.4.4. Eye tracker o wysokiej częstotliwości

Znaczącym czynnikiem opóźnienia systemów renderingu opartych na kierunkowości widzenia jest szybkość przechwytywania kierunku patrzenia obserwatora przez eye tracker. Zbyt niska częstotliwość funkcjonowania urządzenia powoduje wykrywalną desynchronizację pomiędzy aktualnym kierunkiem patrzenia, a zawartością wygenerowanego obrazu. Idealnym rozwiązaniem jest sprzężenie wykrywania punktu skupienia wzroku z algorytmem renderingu w taki sposób, aby przed rozpoczęciem pobierania próbek punkt skupienia był wyliczany przez eye tracker. Takie rozwiązania są stosowane wyłącznie w specjalistycznych środowiskach, lecz wraz ze wzrostem popularności renderingu uwzględniającego kierunek patrzenia rośnie zapotrzebowanie na takie systemy na rynku konsumenckim.

W ramach badań naszego zespołu magister inżynier Krzysztof Wolski opracował system śledzenia oczu oparty na dostępnej na rynku kamerze mającej na sobie filtr

podczerwieni oraz autorskim oprogramowaniem, który korzysta z algorytmu Cross Spread [108]. Zdjęcie opracowanego eye trackera przedstawione zostało na Rys. 2.28. Urządzenie, postawione pod wyświetlaczem, na podstawie pobranych klatek zawierających ludzkie oko dostarcza do oprogramowania pozycję źrenicy, która służy do wyliczenia pozycji skupienia wzroku w przestrzeni wyświetlacza.

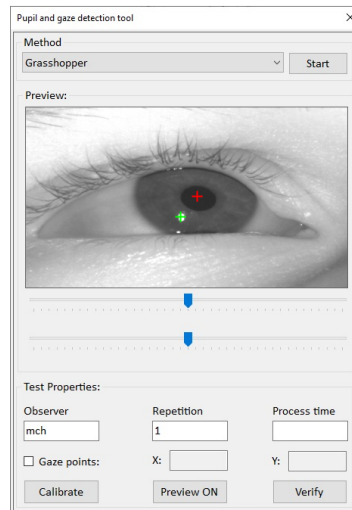


Rysunek 2.28. Eye tracker działający w wysokiej częstotliwości. Wydrukowana obudowa została stworzona pod konkretny model kamery. Kamera posiada uniwersalny otwór z gwintem, co pozwala ustabilizować ją na dowolnym statywie.

Algorytm Cross Spread zapoczątkowany jest wstępnym progowaniem uzyskanego obraz do postaci binarnej przy wykorzystaniu wartości wyznaczonej ręcznie w oprogramowaniu. Następnie na obrazie binarnym stosowany jest filtr medianowy w celu zredukowania szumu wynikającego z progowania. Pierwszym krokiem samego algorytmu jest wyszukanie granicznych punktów źrenicy. W tym celu wewnątrz obszaru zawierającego czarne piksele wybierany jest punkt, z którego tworzone są promienie w czterech ortogonalnych kierunkach. Są one śledzone do granicznego czarnego piksela wzdłuż jego drogi. Następnie wybierany jest punkt w pozycji będący średnią współrzędnych punktów granicznych. Odcięta i rzędna nowego punktu początkowego są obliczane przez uśrednienie współrzędnych odpowiednio poziomych i pionowych punktów granicznych. Następnie te procedury są powtarzane ponownie do momentu, aż pozycja punktu się nie zmienia.

Omawiany eye tracker bazuje na kamerze Point Grey Grasshopper3 4.1 MP Mono USB3 Vision, która jest w stanie przechwytywać obraz w rozdzielczości 2048 x 2048. W celu zwiększenia częstotliwości pobierania kolejnych próbek, rozdzielczość została zredukowana do 640 x 480 pikseli, co umożliwiło osiągnięcie częstotliwości przechwytywania równej 225 klatek na sekundę. Za pomocą autorskiego oprogramowania następowało wyszukiwanie środka źrenicy i odbicia rogówkowego algorytmem. Program został stworzony

w języku programowania C++ (Rys. 2.29), natomiast implementacja algorytmu Cross Spread, który wykonywany był na procesorze graficznym, przy wykorzystaniu biblioteki OpenCL. W ramach projektu badawczego eye tracker został przez nas rozbudowany o oświetlenie podczerwone. Składało się ono z diod LED wlutowanych na autorskiej płytce PCB. Obudowa do płytki wykonana została technologią druku 3D.



Rysunek 2.29. Oprogramowanie systemu eye trackingu. Suwaki na środku służą do ręcznego ustawienia wartości, według której obraz będzie progowany w celu zastosowania algorytmu Cross Spread.

## Podsumowanie rozdziału

Przygotowane środowisko eksperymentalne pozwoliło uzyskać pomiary progów widzialnego kontrastu bodźców wzrokowych HVS. Wyniki te posłużyły do zbudowania analitycznego modelu wskazującego prawdopodobieństwo wykrycia bodźca przez obserwatora. Sam model ukierunkowany jest na uzależnienie widzialności bodźców od obszaru jego padania na siatkówkę oka, co wymaga, aby system syntezy obrazu, który będzie korzystał z zaprezentowanego rozwiązania, miał informację o punkcie skupienia wzroku obserwatora. W kolejnym rozdziale będą zaprezentowane rozwiązania wykorzystania uzyskanych pomiarów do algorytmów komputerowych grafiki czasu rzeczywistego.

Pomiary opóźnienia systemu renderingu opartego o kierunkowość widzenia wskazały dla dużych odległości kątowych bliskie akceptowalnego interwału czasowemu. Pozwala to na ukrycie pogorszenia jakości obrazu w obszarze widzenia peryferyjnego podczas zmiany kierunku patrzenia obserwatora. Uzyskany wynik wskazuje opóźnienie możliwe do osiągnięcia przez współczesne systemy renderingu czasu rzeczywistego.

## **3. Aplikacje modelu czułości na kontrast w grafice komputerowej**

W rozdziale omówione zostały przykłady zastosowania zmierzonych charakterystyk czułości na kontrast układu wzrokowego człowieka w obszarach widzenia peryferyjnego. Rozdział podzielony został na dwie części, w których każda przedstawia inne rozwiązanie. W pierwszej kolejności tematem będzie algorytm ray tracingu, w którym gęstość próbkowania sceny uzależniona jest od zmierzonych wartości CSF (Rozdz. 3.1). Oprogramowanie realizujące algorytm generuje obraz, który wyświetlony w środowisku wirtualnej rzeczywistości korzysta z cech HVS (Rozdz. 3.1.1), dzięki czemu możliwe było uzyskanie korzyści w postaci redukcji czasu syntezy (Rozdz. 3.1.3) bez widocznej degradacji jakości obrazu przez obserwatorów (Rozdz. 3.1.4). Treść tej części rozdziału opiera się na badaniach zrealizowanych w ramach współautorskiej publikacji [W6].

Druga część rozdziału skupia się na light field'owym wyświetlaczu wielopłaszczyznowym, który uwzględnia akomodację oka (Rozdz. 3.2). Zaproponowana została metoda hybrydowa dekompozycji obrazu na poszczególne płaszczyzny, która korzysta z cech HVS (Rozdz. 3.2.2) w celu doboru odpowiedniej metody, co umożliwiło uzyskanie przyspieszenia syntezy obrazu (Rozdz. 3.2.4) przy właściwym jego wyświetlaniu (Rozdz. 3.2.5). Ta część pracy opiera się na badaniach zrealizowanych w ramach współautorskiej publikacji [W7].

### **3.1. Niejednorodne próbkowanie obrazu w metodzie śledzenia promieni**

Synteza i wyświetlanie obrazów wirtualnych scen zawiera się w znaczącej części badań, które związane są z grafiką komputerową. Procesem generowania takich obrazów nazywa się renderingiem, który można podzielić na dwie kategorie. Synteza obrazów realistycznych stara się odwzorować wirtualne sceny w jak najlepszej jakości, nie kierując się szybkością generowania kolejnych klatek. Druga kategoria, która jest tematem tego rozdziału, jest rendering czasu rzeczywistego. W tym przypadku jakość wynikowych obrazów ustępuje szybkości ich generowania. Często wiąże się to z potrzebą interakcji z wirtualnym otoczeniem i natychmiastową odpowiedzią od systemu renderingu, co wymusza generowanie klatek obrazu w czasie kilku milisekund.

**Ray Tracing.** Powszechnym algorytmem syntezującym obrazy wirtualnej sceny jest *metoda śledzenia promieni* (ang. *ray tracing*). Polega ona na prześledzeniu ścieżki promieni światła przechodzących przez poszczególne piksele docelowego obrazu, odbijających się od obiektów wirtualnej sceny, dochodzących do źródeł światła i symulując zachowanie fotonów w rzeczywistym świecie. Ta technika jest w stanie wygenerować obrazy o wysokim stopniu realizmu, zwykle wyższym niż w przypadku typowych metod renderowania czasu rzeczywistego, jak np. wypełniania wielokątów (ang. *scan line algorithm*). Algorytm ten charakteryzuje się jednak dużym kosztem obliczeniowym, co przekłada się na czas potrzebny do wygenerowania klatki obrazu. Chcąc użyć tej techniki w grafice komputerowej czasu rzeczywistego, czas renderingu pojedynczej klatki musi wynosić kilka milisekund, co w klasycznym podejściu metody śledzenia promieni może być nieosiągalne i wymaga stosowania pewnych uproszczeń. Jednym z nich jest redukcja liczby śledzonych promieni pierwotnych. Promienie te śledzone są od kamery i są odbijane od obiektów sceny. Czas renderingu jest zależny od całkowitej liczby śledzonych promieni.

**Wirtualna rzeczywistość.** Wyświetlacze zakładane na głowę (ang. *head-mounted displays - HMD*), które służą do przedstawienia obserwatorowi środowiska Wirtualnej Rzeczywistości (ang. *Virtual Reality - VR Headsets*), stają się coraz bardziej popularne. Sprawia to, że rozwój grafiki komputerowej czasu rzeczywistego wiąże się także z badaniami nad tą technologią. Hełmy te zawierają w sobie dwa wyświetlacze, jeden na każde oko, które prezentują obserwatorowi obraz wirtualnej sceny w sposób jak najbardziej immersyjny, sprawiając wrażenie obecności w niej. Wrażenie to jest zwiększane poprzez śledzenie głowy, żyroskop i kontrolery ruchowe. Dzisiejsze systemy VR osiągają rozdzielczość około 6 *cpd*, co wciąż jest dalekie od rozdzielczości układu wzrokowego człowieka wynoszącego 60 *cpd*. Środowisko VR wymaga częstszego wyświetlania kolejnych klatek obrazu, niż standardowe systemy grafiki czasu rzeczywistego. Zbyt niska częstotliwość odświeżania powoduje dyskomfort użytkowników, stąd wymóg generowania około 90 obrazów na sekundę.

### 3.1.1. Kierunkowa metoda śledzenia promieni

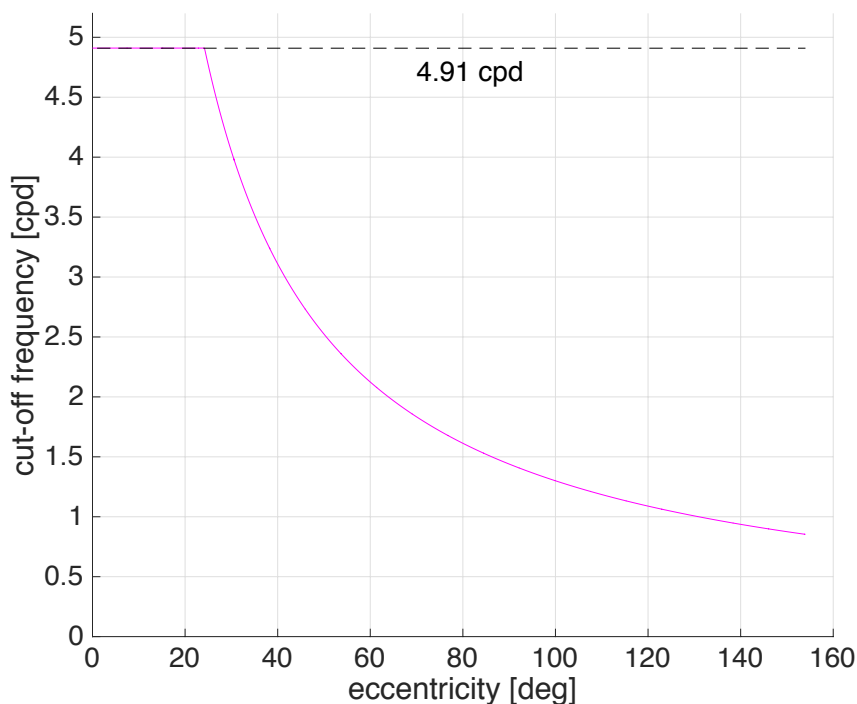
Gęstość próbkowania sceny w algorytmie śledzenia promieni, wpływa na jakość obrazu wynikowego. Domyślna liczba śledzonych promieni pierwotnych wynosi jeden na piksel obrazu. Dla rozdzielczości 4K, będącej standardem w grafice komputerowej, wartość ta sięga 9 milionów, ale nawet taka liczba promieni może powodować zjawisko aliasingu. Najbardziej popularna technika anty-aliasingu polega na zwiększeniu liczba promieni pierwotnych przechodzących przez piksele obrazu, co dodatkowo zwiększa zapotrzebowanie na moc obliczeniową komputera.

Charakterystyka algorytmu ray tracingu nie wymusza jednak rozkładu próbkowania przez co nie jest konieczne wyliczenia koloru każdego piksela w obrazie. Cecha ta może zostać wykorzystana do uzależnienia kierunków promieni pierwotnych od kierunku patrze-

nia obserwatora sceny. Takie rozwiązanie zastosowano w pracy Levoy'a i Whitaker'a [54], gdzie rozkłady promieni śledzonych przez płaszczyznę obrazu oraz rozkład próbek wzdłuż każdego promienia jest zależny właśnie od kierunku patrzenia obserwatora. W rezultacie rozdzielczość renderowanych obrazów zmienia się lokalnie wraz ze zmianą kierunku patrzenia obserwatora. Rozwiązanie opiera się na wykorzystaniu mipmap. Najpierw renderowany jest obraz o niskiej rozdzielczości, po czym następuje zwiększenie gęstości próbek dla mniejszego obszaru obrazu. W innym rozwiązaniu Günter [32] zaproponował silnik renderujący, który generuje trzy obrazy o niskiej rozdzielczości odpowiadające różnym obszarom widzenia. Następnie były one powiększane i łączone z nieskalowanym obrazem obszaru otaczającego punkt patrzenia. W ten sposób liczbę przetworzonych pikseli można zmniejszyć od 10 do 15 razy zapewniając jednocześnie niewidoczne dla obserwatora pogorszenie jakości obrazu wynikowego. Inną technikę zaproponował Stengel [94]. Polega ona na redukcji złożoności obliczeń w technice odroczonego cieniowania (ang. *deferred shading technique*) [2]. Próbkowanie sceny jest stałe dla całego obrazu, ale algorytm cieniowania jest uproszczony dla pikseli znajdujących się w obszarze peryferyjnym. Według autorów technika ta skraca czas cieniowania nawet o 80%. Kontrolowanie szczegółowości obliczeń związanych z oświetleniem zależnie od kierunku patrzenia zaproponowane zostało również przez Vaidyanathana [99]. Patney [82] wprowadził postprocesowe podbicie kontrastu w obszarze peryferyjnym, aby zmniejszyć poczucie widzenia tunelowego i zredukować liczbę próbek. W badaniach zauważono, że ludzie tolerowali nawet dwukrotnie większy promień rozmycia, zanim wykryli różnice w obrazie w stosunku do obrazu referencyjnego. Zaproponowano także nowatorski algorytm wielowarstwowego czasowego antyaliasingu uwzględniającego ruchy sakadyczne oka. Prostą implementację ray tracersa opartego o kierunkowość widzenia zaprezentowano w badaniach Siekawy [92], gdzie rozkład próbkowania promieni pierwotnych opiera się o charakterystykę czułości na kontrast. Podobne założenie można znaleźć w pracy Fujita i Harada [28]. Weier [104] zaproponował rendering oparty na kierunkowości widzenia, w którym algorytm ray tracingu korzystał z promieni prześlędzonych w poprzednich klatkach obrazu.

Zaprezentowane w pracy rozwiązanie sprowadza rekonstrukcję obrazu do interpolacji barycentrycznej. Korzystając z cech algorytmu ray tracingu, który pozwala na niejednorodny rozkład próbek sceny, można selektywnie dobrać punkty reprezentowane przez piksele obrazu, przez które promienie pierwotne będą przechodzić. Zagęszczenie punktów można uzależnić od ich odległości względem środka siatki, zgodnie z charakterystyką czułości na kontrast. Jak wyjaśniono w pracy Loschky'iego [56] uzyskane progi kontrastu można wyrazić jako funkcję rozdzielczości układu wzrokowego, która jest zależna od odległości dołka środkowego oka, tym samym wyznacza pewną częstotliwość przestrzenną próbek, powyżej której odbiór percepcyjny obserwowanego obrazu nie będzie rósł (Rys. 3.1), dzięki czemu można ją zmniejszyć do tej wartości progowej. Ten próg

odcinania częstotliwości można zdefiniować jako maksymalną liczbę próbek widoczną dla człowieka przy danej odległości od obszaru widzenia centralnego.

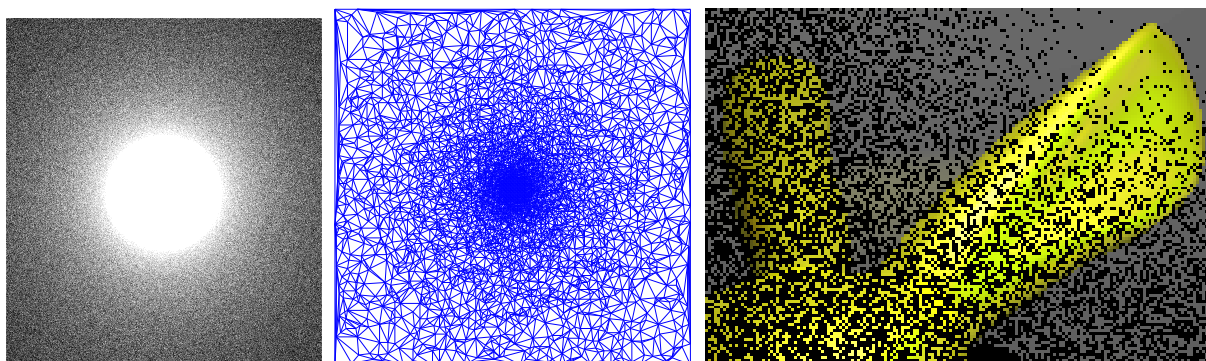


Rysunek 3.1. Częstotliwość odcinania próbkowania sceny zgodna z funkcją czułości na kontrast, jako funkcja odległości od obszaru widzenia centralnego. Fioletowa linia przedstawia zmierzoną czułość na kontrast, pozioma przerywana linia natomiast maksymalną częstotliwość próbkowania wyświetlacza VR (Hełm HTC Vive). Źródło własne [W6] (Rys. 2).

W celu wskazania pikseli, przez które śledzone będą promienie pierwotne, generowana jest maska bitowa. Początkowo maska ta złożona jest z wartości równych zero (symbolizowanych czarnymi pikselami), po czym współrzędne trójkątów siatki są rzutowane na maskę, ustawiając jej wartości w rzutowanych pozycjach na jeden (symbolizowanych białymi pikselami). Wykorzystując wskazane punkty, wykonywany jest algorytm ich triangulacji, w celu wygenerowania dwuwymiarowej siatki złożonej z trójkątów, których wierzchołki znajdują się na białych pikselach maski. Wymiary maski bitowej są dwukrotnie większe niż rozdzielczość w pikselach docelowego obrazu uzyskanego w procedurze renderowania. W algorytmie ray tracingu wybierany jest fragment siatki zależny od kierunku patrzenia obserwatora w taki sposób, aby największe zagęszczenie punktów znajdowało się w punkcie patrzenia obserwatora. Promienie pierwotne generowane są w taki sposób, aby przechodziły przez piksele, które znajdują się w białych punktach obranego fragmentu maski. Dzięki śledzeniu tych promieni, wyliczając na podstawie ich przecięć z obiektami sceny i wyliczeniu równania oświetlenia, wygenerowano obraz, który w białych miejscach w masce zawierał informacje zgodne ze zwykłą procedurą ray tracingu, natomiast pozostałe obszary oznaczają brak próbek sceny. Przykład siatki trójkątów, która jest zależna



od CSF, a także wygenerowanej na jej podstawie maski bitowej oraz próbkowaną scenę przedstawiono w Rys. 3.2.



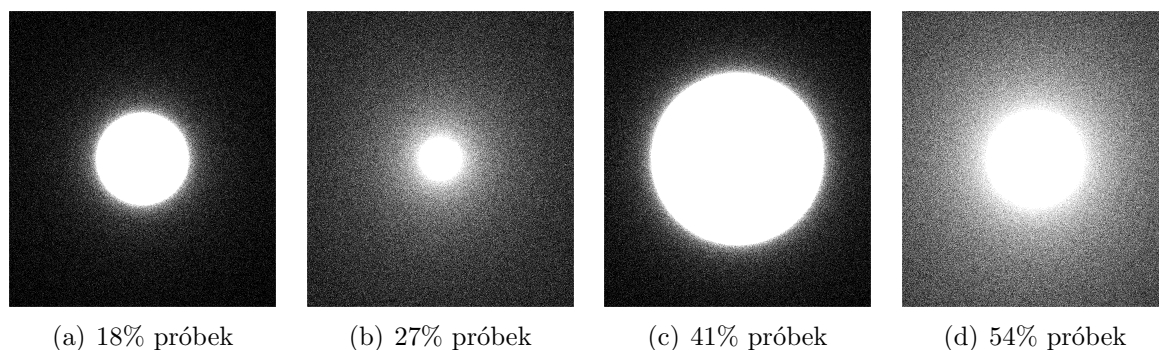
Rysunek 3.2. Po lewej przedstawiona jest maska próbkowania. Białe punkty oznaczają piksele, przez które będą przechodzić promienie pierwotne. Na środku zaprezentowana jest siatka użyta w procesie renderingu do rekonstrukcji obrazu na podstawie wybiórczych próbek sceny (dla przejrzystości liczba wierzchołków została zredukowana do wartości poniżej siedmiu tysięcy). Po prawej znajduje się przykład obrazu RGB składającego się z próbek sceny wybranych przy użyciu maski (czarne punkty wskazują piksele, przez które promienie pierwotne nie były śledzone). Źródło własne [W6] (Rys. 3).

### 3.1.2. Rekonstrukcja próbek obrazu

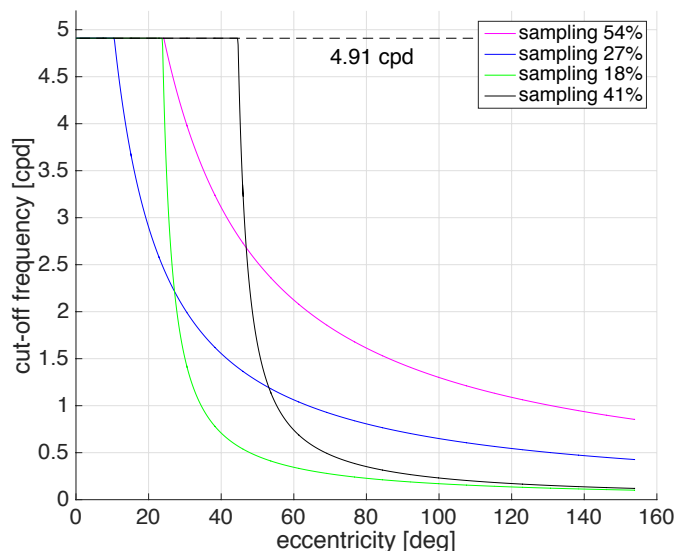
Próbki sceny uzyskane ze śledzenia promieni konwertowane są do przestrzeni kartezjańskiej, dzięki czemu można je wyświetlić na monitorze. Piksele niezawierające próbek będą rekonstruowane dzięki triangulacji informacji oraz renderowane za pomocą standardowej techniki syntezy obrazu (tzw. *scan line* [91]). Maska bitowa zawierająca rozkład pikseli konwertowana jest do siatki trójkątów za pomocą Triangulacji Delaunay. Każda próbka na mapie staje się wierzchołkiem siatki (Rys. 3.2 (środek)). Ray tracer śledzi promienie przechodzące przez wierzchołki siatki i przechowuje kolory odpowiednich pikseli. Pozostałym pikselom obrazu niebędącym w pozycji żadnego wierzchołka przypisywany jest kolor interpolowany barycentrycznie na podstawie wartości przechowywanych we wierzchołkach odpowiadającego im trójkąta. Testowane były również bardziej złożone techniki rekonstrukcji: technika push-pull, która była wprowadzona przez Gortlera [31] oraz mapy komórek [91], nie poprawiają one znacząco jakości obrazu, a wymagają większej mocy obliczeniowej.

W projekcie wykorzystane zostały cztery maski o różnym rozłożeniu pikseli (Rys. 3.3). Maska zawierająca 54% białych pikseli (Rys. 3.3(d)) jest zgodna ze zmierzoną charakterystyką czułości na kontrast HVS. Pozostałe maski zawierają 18% (Rys. 3.3(a)), 27% (Rys. 3.3(b)) i 41% (Rys. 3.3(c)) liczby próbek odpowiadających klasycznemu renderingowi. Te dodatkowe maski służą do przetestowania zależności między częstotliwością próbkowania, a szybkością procesu syntezy obrazu. Proces rekonstrukcji zależy bowiem

nie tylko od ilości próbek, ale ich rozłożenia. Dodatkowe maski próbkowania posłużą także do przetestowania jakości obrazu po dodatkowej redukcji liczby próbek. Teoretycznie ta redukcja powinna być niezauważalna dla maski 54%, natomiast dla pozostałych widoczna. Częstotliwości odcinania wszystkich czterech, użytych w tej pracy masek można zaobserwować na Rys. 3.4. Widać na nim, że maski oprócz liczby próbek różnią się także zależnością między odległością kątową od obszaru widzenia centralnego, a częstotliwością próbkowania.



Rysunek 3.3. Maski próbkowania obrazu w kierunkowym śledzeniu promieni. Procentowa wartość pod poszczególnymi rysunkami wskazuje na liczbę białych pikseli, tym samym liczbą śledzonych promieni pierwotnych, w stosunku do wszystkich pikseli docelowego obrazu. Ostatnia maska po prawej (54%) jest zgodna ze zmierzoną charakterystyką czułości na kontrast układu wzrokowego człowieka w obszarach peryferyjnych. Źródło własne [W6] (Rys. 5).

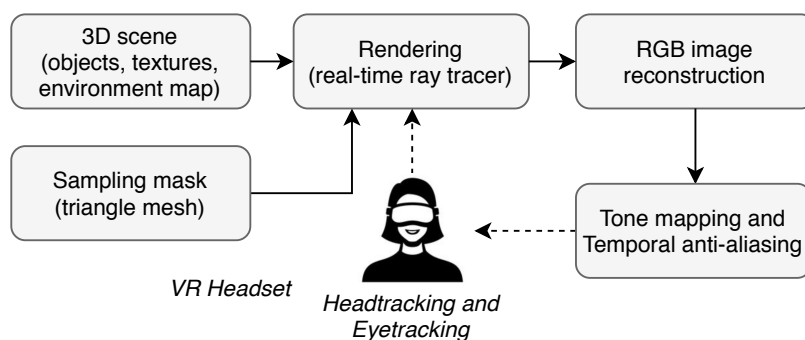


Rysunek 3.4. Częstotliwości odcinania próbkowania sceny jako funkcja odległości kątowej od obszaru widzenia centralnego. Zielona, niebieska, czarna i fioletowa linia przedstawia założoną czułość na kontrast obserwatora dla masek zawierających kolejno 18%, 27%, 41% i 54% próbek. Pozioma przerywana linia przedstawia maksymalną częstotliwość próbkowania wyświetlacza VR (Helm HTC Vive). Źródło własne [W6] (Rys. 6).

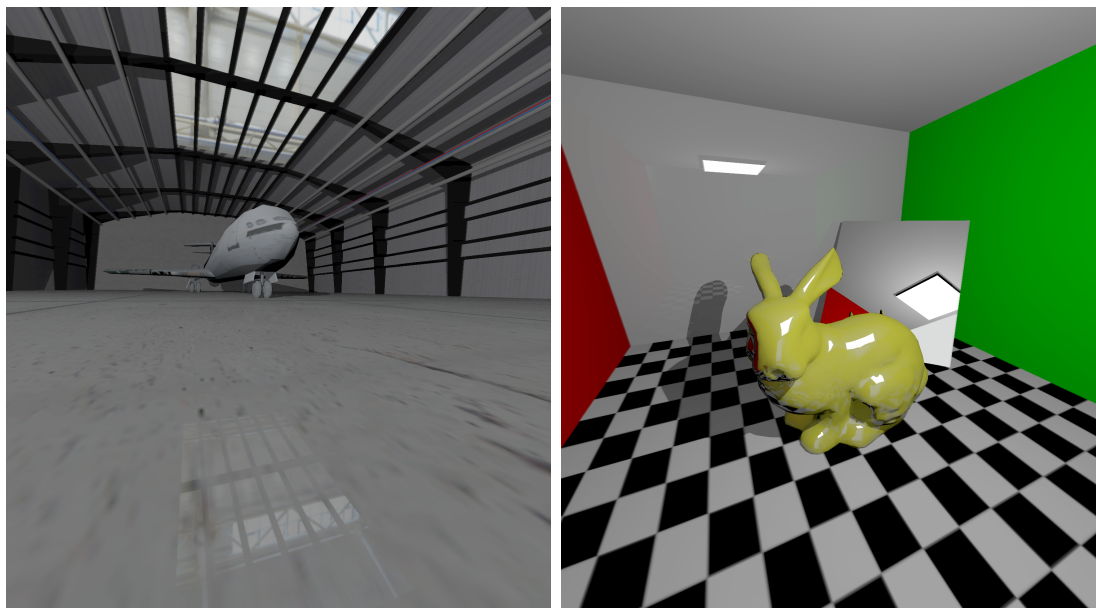
### 3.1.3. Pomiar szybkości renderingu

W celu przetestowania opisanego rozwiązania stworzony został system syntezy obrazu, który zawiera implementację metody śledzenia promieni w środowisku wirtualnej rzeczywistości. Ray Tracing umożliwia niejednorodną dystrybucją śledzonych promieni pierwotnych uzależnioną od kierunku patrzenia obserwatora (Rys. 3.5). Informacja o punkcie skupienia wzroku była rejestrowana przez eye tracker zintegrowany z hełmem wirtualnej rzeczywistości. Promienie pierwotne były generowane w ten sposób, aby przechodziły przez piksele zawarte w masce bitowej. Fragment maski dobierany był na podstawie kierunku patrzenia obserwatora. Proces renderingu został zrównoleglony, co umożliwiło syntezę obrazu w czasie rzeczywistym. Po wyszukaniu przecięć i odbić oraz policzeniu równania oświetlenia, obraz poddawany był procesowi rekonstrukcji poprzez wykorzystanie siatki trójkątów. Zastosowanie temporalnego anti-aliasingu miało na celu zredukowanie szumów wynikających z wybiórczego dobierania próbek.

Głównym założeniem stworzonego systemu był rendering obrazów w czasie rzeczywistym. W tym celu zaimplementowany został autorski algorytm śledzenia promieni wykonywany na procesorze graficznym. Program generował niejednorodnie promienie pierwotne na podstawie dostarczonej maski bitowej i kierunku patrzenia obserwatora, który podlegał ciągłym zmianom podczas używania hełmu VR. Sam algorytm został zaimplementowany w OpenCL, dzięki czemu generowanie promieni pierwotnych i równanie oświetlenia było wyliczane na karcie graficznej. Za śledzenie promieni, a także stworzenie hierarchicznego podziału przestrzeni wirtualnej sceny odpowiedzialna była biblioteka Radeon Rays. Zebrane próbki sceny skonwertowane zostały do obrazu RGB podczas fazy rekonstrukcji zaimplementowanej w API OpenGL. Końcowy obraz wyświetlony został w hełmie wirtualnej rzeczywistości za pomocą biblioteki OpenVR. W celu przyspieszenia renderingu wyliczane zostało wyłącznie jedno odbicie promieni. Do potoku renderowania dodany został algorytm temporalnego antyaliasingu (ang. *temporal anti-aliasing* - TAA), który redukował problem migotania obrazu podczas wyświetlania animacji.



Rysunek 3.5. Schemat architektury systemu renderingu uwzględniającego kierunkowość widzenia. Źródło własne [W6] (Rys. 1).

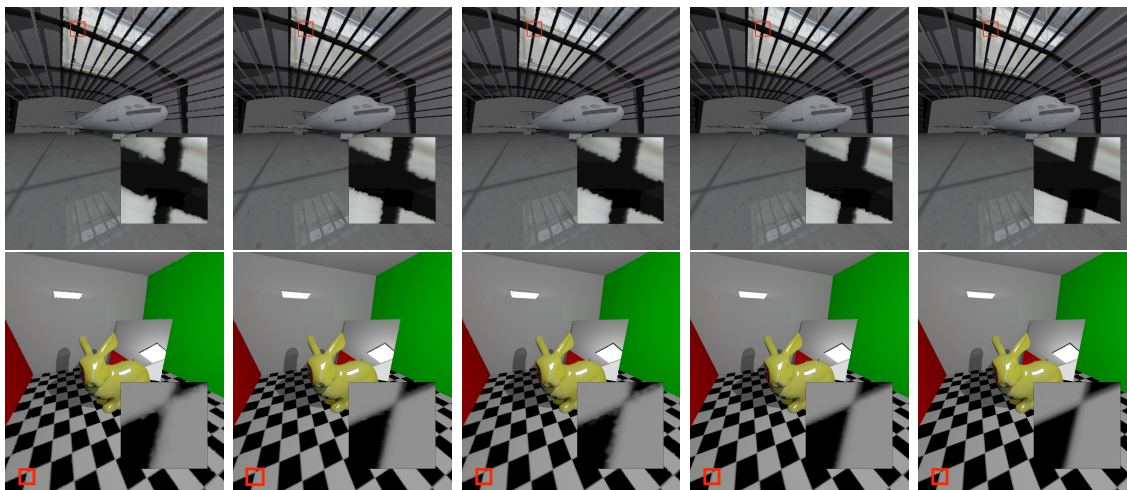


Rysunek 3.6. Przykład renderingu scen Air Shed (po lewej) oraz Bunny Box (po prawej). Źródło własne [W6] (Rys. 4).

Stworzony system renderingu został uruchomiony na komputerze wyposażonym w kartę graficzną NVIDIA Geforce GTX 1080. Podłączony do niego hełm HTC Vive VR wyświetlał obraz o rozdzielczości 1512 x 1680 pikseli dla każdego oka. Szybkość renderingu zmierzona została dla dwóch przykładowych scen (Rys. 3.6). Pierwsza z nich (nazwana Bunny Box) składa się z pięciu ścian oraz obiektu Stanford Bunny, natomiast w drugiej scenie (nazwanej Air Shed) znajdował się samolot zbudowany z około 1500 trójkątów oraz ściany pozorujących wnętrze hangaru.

W ramach testów wydajnościowych ray tracera porównano szybkości renderingu obrazów dla jednorodnego próbkowania pikseli oraz dla niejednorodnych rozkładów śledzonych promieni pierwotnych. Wykorzystano maskę wskazującą 54% pikseli w obrazie, a także maski zawierające 41%, 27% i 18%. Dodatkowa redukcja gęstości próbkowania sceny miała na celu sprawdzenie zależności między ilością promieni pierwotnych, a szybkością renderingu i jakością uzyskanych obrazów. Uzyskane obrazy prezentowane są na Rys. 3.7. Powiększone fragmenty pokazują uproszczenia będące wynikiem zredukowanego próbkowania w obszarze widzenia peryferyjnego.

W tabeli 3.1 przedstawione zostały uzyskane czasy renderingu dla różnych masek. Przy jednorodnym próbkowaniu sceny czas renderingu scen Air Shed i Bunny Box wynosił odpowiednio 66.0 i 49.5 ms. Dla maski zawierającej 54% pikseli czasy te wynosiły 29.8 i 24.1 ms, co daje ponad dwukrotne przyspieszenie renderingu. Wraz z redukcją liczby próbek czas renderingu maleje. Wyjątkiem jest maska zawierająca 41% pikseli obrazu. Najprawdopodobniej wynika to z charakterystycznej dystrybucji próbek, które w tej masce są w większości skupione wokół punktu patrzenia obserwatora, podczas gdy



Rysunek 3.7. Przykłady renderingu scen przy użyciu różnych masek próbkowania (od lewej 18%, 27%, 41%, 54% i 100% pikseli). Źródło własne [W6] (Rys.7).

dla dalszych obszarów ich liczebność jest bardzo mała. Uzyskany pomiar wskazuje na zasadnicze znaczenie obszaru widzenia centralnego.

| Scena     | Maska próbkowania | Czas renderingu [ms] | Uzyskane przyspieszenie |
|-----------|-------------------|----------------------|-------------------------|
| Air Shed  | 18%               | 18.1                 | 3.7x                    |
|           | 27%               | 23.7                 | 2.8x                    |
|           | 41%               | 37.5                 | 1.8x                    |
|           | 54%               | 29.8                 | 2.2x                    |
|           | 100%              | 66.0                 | 1.0x                    |
| Bunny Box | 18%               | 15.0                 | 3.3x                    |
|           | 27%               | 17.9                 | 2.8x                    |
|           | 41%               | 27.6                 | 1.8x                    |
|           | 54%               | 24.1                 | 2.1x                    |
|           | 100%              | 49.5                 | 1.0x                    |

Tabela 3.1. Wyniki pomiaru czasu syntezy obrazów dla pełnej rozdzielczości oraz niejednorodnego próbkowania. Po prawej przedstawione zostały uzyskane przyspieszenie w porównaniu renderingiem dla pełnej rozdzielczości. Źródło własne [W7] (Tab. 1).

#### 3.1.4. Ocena jakości renderingu

W celu zapewnienia poprawnego działania systemu renderingu opartego o kierunkowość widzenia należy zweryfikować, czy pogorszenie jakości obrazu wynikowego w obszarze widzenia peryferyjnego nie wpłynie na jego odbiór przez człowieka. W tym celu został przeprowadzony eksperyment percepcyjny, w którym obserwatorzy porównywali zaprezentowane zostały obrazy wyświetlone w pełnej rozdzielczości z obrazami wygenerowanymi za pomocą zredukowanej liczby próbek. Celem było odnalezienie największej

redukcji gęstości próbkowania, które nie spowoduje zauważalnego dla obserwatora pogorszenia jakości obrazu. Eksperyment przeprowadzany został z wykorzystaniem gogli HTC Vive, na których wyświetlone zostały obrazy zawierające sceny Air Shed oraz Bunny Box. W trakcie trwania eksperymentu obrazy renderowane były ze zmieniającymi się maskami próbkowania pikseli.

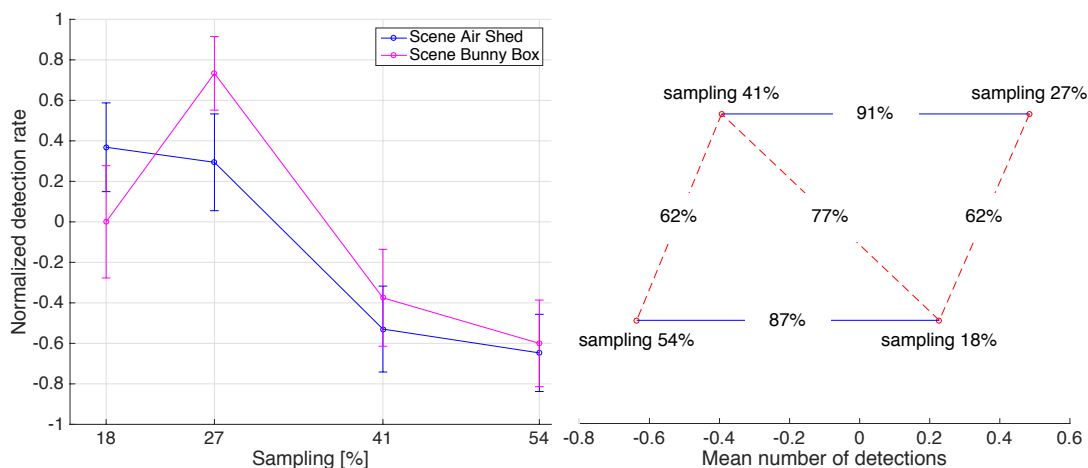
**Procedura** Uczestnicy eksperymentu poproszeni byli o założenie hełmu wirtualnej rzeczywistości i swobodne rozejrzenie się po wyświetlonej na nim scenie. Na początku eksperymentu przedstawiony został obraz wygenerowany w pełnej rozdzielczości. Po 20 sekundach obraz ten uległ zmianie na wygenerowany za pomocą losowo wybranej maski próbkowania. Maską była zmieniana wielokrotnie w 5-sekundowych odstępach w trakcie sesji trwającej łącznie 180 sekund. Zadaniem obserwatora było naciskanie przycisku myszki, gdy zauważy on/ona zmianę jakości wyświetlanego obrazu spowodowaną zmianą maski próbkowania.

**Uczestnicy eksperymentu** Eksperyment przeprowadzony został dla 6 osób (w wieku od 20 do 24 lat, 4 mężczyzn i 2 kobiety). Każdy z uczestników zadeklarował poprawny lub skorygowaną ostrość widzenia. Ochotnicy byli instruowani jak poprawnie wykonać eksperyment oraz na co zwracać uwagę, lecz nie znali celu prowadzonych badań.

**Wyniki eksperymentu** Graficzna interpretacja uzyskanych rezultatów eksperymentu znajduje się na Rys. 3.8. Przedstawione tam zostały znormalizowane częstotliwości wykrywania zmiany jakości obrazu dla każdej maski próbkowania. Częstotliwość równa wartości jeden równoważna jest ze stuprocentową skutecznością wykrywania pogorszenia jakości przy użyciu danej maski. Natomiast wartość zero oznacza sytuację, gdy użytkownik zauważył różnicę w 50% przypadków. Z kolei wartość -1 wskazuje na brak zdolności do wykrycia zmiany jakości podczas użycia danej maski.

Na wykresie przedstawiającym znormalizowaną częstotliwość wykrywania zmiany jakości (Rys. 3.8, po lewej) można zauważyć, że dla masek 18% i 27% częstotliwość wykrywania różnic jakości dla obu scen wynosi powyżej zera, co świadczy o dostrzegalności różnic w jakości. Dla masek 41% i 54% wartość ta znajdowała się poniżej progu dla obu scen, co z kolei świadczy o niedostrzegalności obniżenia jakości. Należy zwrócić uwagę na fakt, że eksperyment był przeprowadzany dla założeń konserwatywnych, gdyż znacznie łatwiej jest dostrzec pogorszenie jakości obrazu podczas zmiany maski.

Na podstawie rezultatów przetestowana została również istotność statystyczna uzyskanych wyników przy użyciu wielokrotnego porównywania próbek, które identyfikują statystyczną różnicę w testach rankingowych [63] (Rys. 3.8, po prawej). Maski uporządkowane są zgodnie z częstością wykrywania, przy najniższej wykrywalności po lewej stronie. Wartości procentowe wskazują prawdopodobieństwo, że obserwator wybierze próbkowanie



Rysunek 3.8. Wyniki eksperymentu percepcyjnego dotyczącego jakości renderingu uwzględniającego kierunkowość widzenia. Po lewej znajduje się znormalizowana częstotliwość wykrywania zmian jakości przy wykorzystaniu danej maski próbkowania (pionowe linie oznaczają odchylenie standardowe). Po prawej znajduje się wykres rankingowy ilustrujący istotność statystyczną wyników uzyskanych w eksperymencie. Źródło własne [W6] (Rys.8).

po prawej stronie jako gorsze, niż próbkowanie po lewej. Jeśli linia łącząca dwie próbki jest czerwona i przerywana, oznacza to, że nie ma statystycznej różnicy między tą parą próbek. Prawdopodobieństwa bliskie 50% zwykle powodują brak istotności statystycznej. Dla większych prawdopodobieństw przerywana linia zastępowana jest przez niebieską linię. Maski zawierająca 54% próbek wskazuje na znacząco lepsze wyniki (mniejsza częstotliwość wykrywania) od maski 18%, natomiast próbkowanie 41% jest lepsze od próbkowania 27%. Przeprowadzone eksperymenty pokazują, że zmniejszenie gęstości próbkowania dla peryferyjnego obszaru obrazu jest ledwo zauważalne, szczególnie w przypadku maski próbkowania zawierającej 54% próbek.

### 3.2. Wyświetlacz stereoskopowy z akomodacją

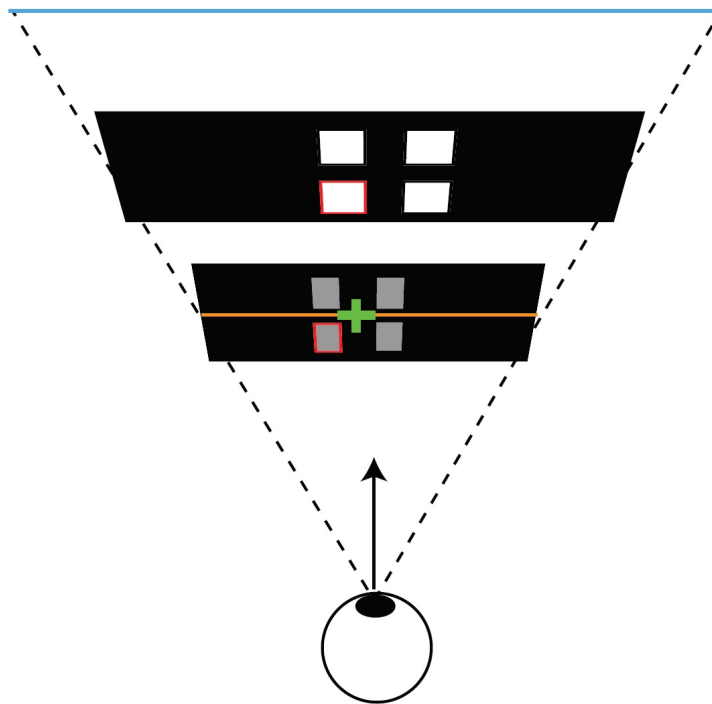
W ostatnich latach wyświetlacze HMD przedstawiające VR i AR (ang. *augmented reality* - rozszerzona rzeczywistość) rosną na popularności i znajdują zastosowanie w różnych dziedzinach, tj. gry, filmy, medycyna czy symulacje. Charakteryzują się one wykorzystaniem techniki stereoskopii (ang. *stereoscopy*) polegającej na przedstawieniu sceny w takiej postaci, że pojawia wrażenie widzenia dwuocznego. Wymaga to przedstawienia układowi wzrokowemu człowieka dwóch osobnych obrazów, po jeden na każde oko. Obrazy takie, oprócz odwzorowania kształtu i koloru, przedstawiają także wzajemne zależności przestrzenne obiektów, tj. ich odległość od obserwatora i głębokość sceny. Odwzorowanie tych zależności możliwe jest dzięki wykorzystaniu konwergencji i akomodacji gałek ocznych, co wymaga zawarcia dodatkowych informacji w wyświetlanym obrazie, które są pewnego rodzaju wskazówkami dla układu wzrokowego, umożliwiające odtworzenie tych

zależności w korze wzrokowej. Jest to niezwykle trudne zadanie z uwagi na kompromis między potrzebami, jakie taki system musi spełniać. Tyczy się to ograniczeń związanych z elementami optycznymi w projekcie sprzętu i dotyczą m.in. szerokość pola widzenia (ang. *field of view - FOV*) oraz rozdzielczość wyświetlacza [38]. Oprócz wyzwań, które niesie ze sobą sprzęt sam w sobie, ważne jest także generowanie treści, które również stanowi wyzwanie. Synteza klatek na wyświetlaczu stereoskopowym wymaga znacznie większej mocy obliczeniowej w porównaniu do standardowych obrazów dwuwymiarowych, które zawierać muszą wskazówki umożliwiające odtworzenie głębi ostrości i wywołanie akomodacji oka. Ponadto projekt potoku renderowania musi być zgodny z wykorzystaną architekturą wyświetlacza [45].

### 3.2.1. Rodzaje wyświetlaczy zawierających wskazówki akomodacyjne

Jedną z architektur wyświetlacza stereoskopowego, obsługującego wskazówki akomodacyjne, jest Wyświetlacz Wielopłaszczyznowy (ang. *Multi-layer display*). Przedstawia on scenę poprzez wyświetlanie kilku obrazów oddalonych od obserwatora w różnych odległościach, z których każdy przedstawia równe bodźce wzrokowe zawarte w scenie (Rys. 3.9). Obserwator widzi obraz sceny, który jest wynikiem złożenia wyświetlanych obrazów, co umożliwia wywołanie wrażenia głębi obiektów. Architekturę takich wyświetlaczy można podzielić na dwie kategorie. Systemy oparte na multipleksowaniu czasowym z przełączanymi soczewkami [37, 57] charakteryzują się małym rozmiarem, lecz do poprawnego działania wymagają wysokiej częstotliwości odświeżania wyświetlacza oraz sprzętu optycznego umożliwiającego dynamiczne dostrajanie ostrości soczewek. Drugim typem architektury wyświetlaczy wielopłaszczyznowych są wyświetlacze korzystające z dzielników wiązki (ang. *beam splitters*) oraz wielu fizycznych wyświetlaczy [1, 59]. Są one większych rozmiarów z uwagi na elementy układu optycznego, lecz oferują także szerszy FOV, co sprawiło, że w niniejszych badaniach to ta architektura była wybrana. Wymogiem dla poprawnego działania obu architektur jest śledzenie oczu obserwatora, gdyż układ elementów wyświetlacza i tym samym wygenerowane obrazy są dostosowane do ustalonego kierunku patrzenia. Niedawno opracowane zostały wyświetlacze o powierzchni ogniskowej, które reprezentują obraz trójwymiarowy w sposób ciągły [65], co wyeliminowało potrzebę śledzenia wzroku, aby móc wygenerować obraz w pojedynczej płaszczyźnie. Rozwiązanie to jednak jest wymagające obliczeniowo oraz wymaga drogich modulatorów fazy światła (ang. *light phase modulator*). Innym podejściem, aby uniknąć śledzenia wzroku, jest przeprowadzenie optymalizacji zawartości sceny dla danych kierunków patrzenia, lecz wymaga to dużej mocy obliczeniowej i precyzyjnej kalibracji osi obrotu gałek ocznych obserwatora [50]. Ponieważ niniejsze badania wiążą się z renderingiem uwzględniającym kierunek widzenia, wykorzystuje się tutaj śledzenie kierunku patrzenia, aby uzyskać poprawne przedstawienie obrazu na wyświetlaczu.





Rysunek 3.9. Rysunek przedstawiający koncepcję wyświetlaczy wielowarstwowych. Obserwator patrzy na obraz będący sumą obrazów umieszczonymi na różnych płaszczyznach, umiejscowionych w ustalonych odległościach.

Podczas dekompozycji obrazu na wirtualnej płaszczyźnie występuje problem niedoboru informacji na krawędziach obiektów, co może zostać rozwiązane w Wielopłaszczyznowym Wyświetlaczu Light Field'ów (ang. *Multi-layer light field display*) posiadającym dodatkową informację przestrzenną o pikselach obrazu. Termin Light Field (w wolnym tłumaczeniu *pole światła*) oznacza pięciowymiarową funkcję wektorową, opisującą intensywność lub chromatyczność światła, które jest obserwowane z każdej pozycji i kierunku w przestrzeni 3-wymiarowej. Przechwytywanie light field'ów odbywa się poprzez uzyskiwanie obrazu danej sceny z różnych pozycji i kierunków kamery [53]. Przykładem może być macierz aparatów, które robią zdjęcie tej samej sceny i są oddalone od siebie o ustalony dystans. Na podstawie zestawu uzyskanych obrazów, możliwe jest przetwarzanie uzyskanej informacji w sposób niemożliwy przy użyciu pojedynczego obrazu, tj. wizualizacja sceny z nieprzechwyczonej pozycji, bądź kierunku. W ostatnich badaniach, na podstawie mikrosoczewek, zaprezentowany został wyświetlacz light field'ów, który przedstawia obiekty wirtualnej sceny o zmiennej ostrości [39, 48]. W wyświetlaczach tych występuje jednak kompromis między polem widzenia, a rozdzielczością wyświetlacza. Wyświetlacze light field'ów oparte na architekturze wielopłaszczyznowej przedstawione zostały jako wydajne rozwiązania, które zapewniają poprawne odwzorowanie ostrości obiektów[40, 60, 68].

Architektura zawarta w pracy, opiera się o addytywny wyświetlacz light field'ów, który umożliwi poprawne zawarcie wskazówek akomodacyjnych w obrazie [68].

Z innych wyświetlaczy należy wspomnieć o wyświetlaczach holograficznych, które przedstawiają obraz scen zapewniając dokładne wskazówki, dotyczące ostrości obiektów [110], lecz ograniczony rozmiar piksela oraz mała rozdzielczość cyfrowych modulatorów falowych, co narzuca znaczny kompromis między rozmiarem wyświetlacza, a FOV [61]. Kolejną wartą wspomnienia architekturą jest dynamiczna zmiana głębokości płaszczyzn obrazów 2D, która jest przedstawiana obserwatorowi przy pomocy urządzeń umożliwiających dostrojenie ostrości [3, 25]. To rozwiązanie pozwala obserwować wirtualne sceny z poprawnymi wskazówkami akomodacyjnymi. Charakteryzuje się to jednak problemami, które są związane z opóźnieniem zmiany ostrości. Konflikt między wergencją, a akomodacją rozwiązywany jest głównie za pomocą generowania pełnych wskazówek. Może zostać również rozwiązany dzięki wyświetlaniu serii obrazów o zmiennej ogniskowej [44], lecz to wymusza kompromis między rozdzielczością wyświetlacza, a zakresem odtwarzanej ogniskowej. Ostatnio wykazano również, że prawidłowe renderowanie aberracji chromatycznej może skutecznie wywoływać akomodację bez zmiany ostrości elementów optycznych [12].

### 3.2.2. Algorytmy dekompozycji obrazu na wirtualne płaszczyzny

W wyświetlaczach wielopłaszczyznowych dekompozycja obrazu na poszczególne płaszczyzny odbywa się według reguły liniowego mieszania (ang. *linear blending* - LB), która przypisuje wartości pikseli proporcjonalne do odległości między punktem docelowym, a płaszczyznami wyświetlania [1]. Ta metoda skutecznie stwarza wrażenie akomodacji, lecz ze względu na proste rozważanie obrazu i mapy głębokości, granice okluzji oraz powierzchnie inne niż lambertiańskie nie są renderowane w sposób prawidłowy. Chcąc prawidłowo wygenerować sceny, które nie zawierają artefaktów graficznych, zaproponowana została optymalizacja siatkówkowa (ang. *retinal optimization* - RO) [66, 76], która optymalizuje stosy ogniskowych utożsamiane z light field'ami. Zawiera ona wiele obrazów wygenerowanych przy wykorzystaniu soczewki kamery o różnej ogniskowej [52]. Wyeliminowanie za pomocą RO artefaktów renderingu pozwala na wykorzystanie optymalizacji generowania light field'ów w wyświetlaczach wielopłaszczyznowych.

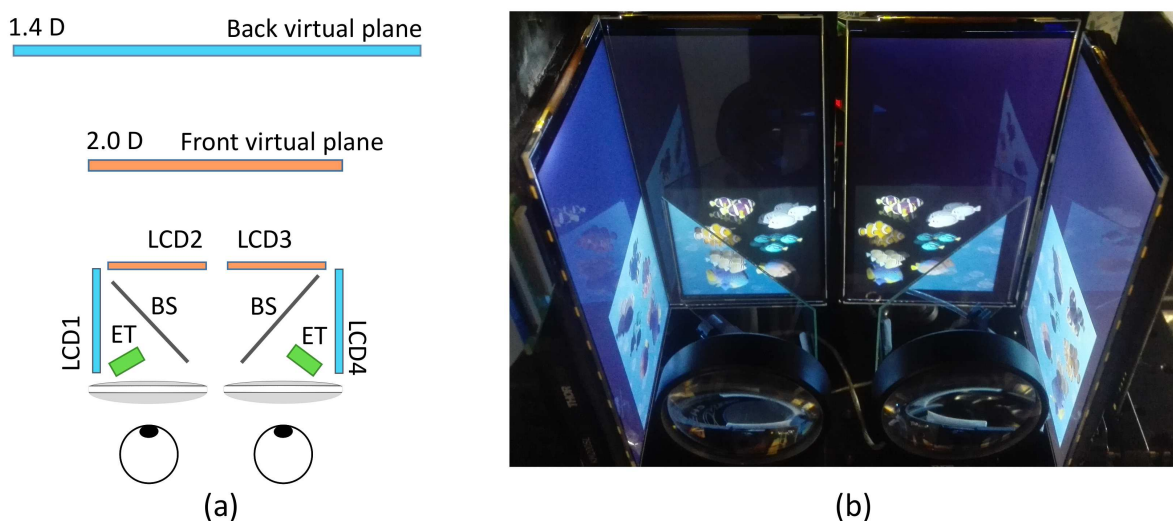
W wyświetlaczach wielopłaszczyznowych, wyświetlających light field'y, sceny składają się z kilku płaszczyzn dwuwymiarowych obrazów. W wyświetlaczach multiplikatywnych, gdzie wartości pikseli wyświetlanych na różnych płaszczyznach są ze sobą mnożone, uzyskując przy tym końcową wartość piksela, optymalizacja opisywana jest w postaci tensora i rozwiązana za pomocą różnych algorytmów faktoryzacji [40, 105]. Rozwiązania dla addytywnych wyświetlaczy light field'ów, czyli takich, gdzie wartości pikseli obrazów we wszystkich płaszczyznach są sumowane w celu uzyskania koloru końcowego piksela, bazują na polaryzacji wyświetlaczy LCD [49] lub niekoherentnym sumowaniu intensywności

pikseli odbitych od holograficznych elementów optycznych wyświetlacza [50]. W przypadku tych architektur problem optymalizacji sprowadza się do minimalizacji błędu liniowej metody najmniejszych kwadratów (ang. *linear least-squares error*) oznaczanej jako *LFS*. Problem ten jest rozwiązywany poprzez wykorzystanie techniki synchronicznej rekonstrukcji algebraicznej (ang. *simultaneous algebraic reconstruction technique*) oznaczanej jako SART, która może zostać rozwiązana w czasie rzeczywistym [4], bądź wykorzystana jako metoda wyznaczenia regionu zaufania [15], która wykorzystywana jest offline'owo. Generowanie obrazów light field'ów za pomocą LFS wymaga dużej mocy obliczeniowych, a osiągnięcie wydajności umożliwiającej wykorzystanie jej w czasie rzeczywistym, jest możliwe tylko poprzez redukcję liczby iteracji [40]. Możliwe jest także adaptacyjne dobieranie gęstości próbkowania w celu przyspieszenia procesu renderingu, lecz takie rozwiązanie zostało przedstawione wyłącznie w offline'owych przykładach wykorzystania SART [34]. Zaprezentowane w niniejszej pracy rozwiązanie polega na metodzie hybrydowej, która wykorzystuje metodę LB do syntezy całego obrazu sceny i selektywne renderowanie za pomocą LFS, co pozwala wyeliminować koszt obliczeniowy podczas generowania docelowych obrazów zawierających light field oraz ich dekompozycję do danych płaszczyzn. Ponadto zmodyfikowany algorytm SART jest dostosowany do uruchomienia na procesorze graficznym, co umożliwia rendering w czasie rzeczywistym z dobrą jakością wynikowych obrazów. Co więcej, implementacje algorytmów, przy wykorzystaniu zaproponowanej w niniejszej pracy metody opartej o LFS wykazują lepszą wydajność niż RO, pozwalając przy tym na uwarunkowanie próbkowania od kierunku patrzenia, w celu zmniejszenia kosztów obliczeniowych bez redukcji jakości odbieranego obrazu przez obserwatora.

### 3.2.3. Budowa wielopłaszczyznowego wyświetlacza light field'ów

#### Układ optyczny

W celu implementacji metody hybrydowej zbudowany został wielopłaszczyznowy wyświetlacz VR (Rys. 3.10). Dla każdego oka obraz był wyświetlany na dwóch wyświetlaczach LCD Topfoison TF60010A, o rozdzielczości 2560 x 1440 pikseli, w którym wyświetlany obraz sceny był łączony rozdzielaczem wiązki (Edmund Optics 64-408) i powiększany za pomocą soczewki achromatycznej (Thorlabs AC508-080-A). Do wyświetlacza zamontowany został eye tracker Pupil Labs, który był umiejscowiony pod soczewkami. Cały układ optyczny znajdował się na liniowym stoliku optycznym w celu dostosowania odległości pomiędzy poszczególnymi elementami układu. Odległość dioptrii, dla przedniej i tylnej wirtualnej płaszczyzny, ustawiona została na wartości 2.0D i 1.4D. Obraz znajdował się na fragmencie wyświetlacza obejmującym obszar 1200 x 1200 pikseli, co w wyniku dało FOV równy 40° i rozdzielczość przestrzenną 15 *cpd*. Osiągnięte parametry były znacznie wyższe, niż w dotychczas wytworzonych wyświetlaczach light field'ów wyświetlających VR



Rysunek 3.10. Po lewej (a) schemat wielopłaszczyznowego wyświetlacza light field'ów. BS oznacza rozdzielacz wiązek światła, natomiast ET to eye tracker. Po prawej (b) zdjęcie wyświetlacza. Źródło własne [W7] (Rys. 10).

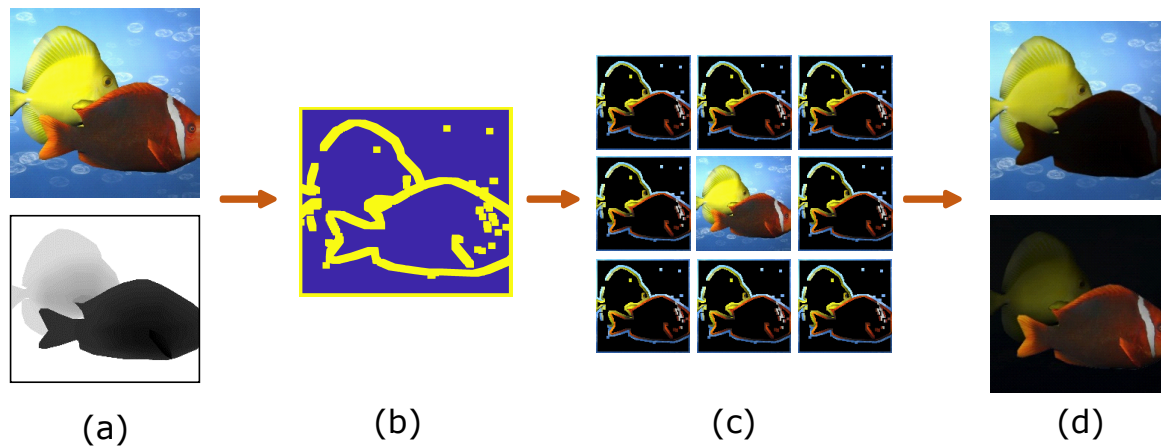
i AR, które rzadko przekraczają 10 *cpd*, co znacząco ogranicza rendering uwzględniający kierunek patrzenia [40, 50, 66].

### Rendering i dekompozycja obrazu

Wykorzystany w wyświetlaczu potok renderowania podzielić można na cztery etapy (Rys 3.11). W pierwszej kolejności renderowany jest obraz sceny z centralnego punktu widzenia oraz jej mapa głębokości (a). Następnie, na podstawie mapy głębokości i kontrastu Michelsona, dla obiektów na scenie obliczana jest maska LFS (b). Próbkki zawarte w masce są renderowane, w kolejnym etapie potoku, do dodatkowych obrazów sceny, przechwyconych ze zmienionego punktu widzenia (c). W końcowym etapie łączy się uzyskane obrazy za pomocą reguły LB i iteracyjną dekompozycją za pomocą SART (d), uzyskując dwa obrazy umieszczone na danej odległości w wirtualnej płaszczyźnie.

Domyślnie, potok renderowania - w celu wyświetlenia go - zawierałby syntezę dziewięciu obrazów o rozdzielczości 1200 x 1200 pikseli, co wiąże się z dużym kosztem obliczeniowym. Rendering sceny odbywa się przy użyciu algorytmu śledzenia promieni z próbkowaniem jednego promienia na każdy piksel obrazu, co wiąże się z jeszcze większym kosztem obliczeniowym. W celu zredukowania liczby śledzonych promieni i tym samym skrócić czas renderingu, w pierwszej kolejności renderowany jest jeden obraz, aby na podstawie informacji w nim zawartych dobierane są próbki, które są wymagane do przesłania promieni z innego punktu startowego. Na redukcję czasu renderingu wpływa także fakt, że implementacja SART opracowana została na procesorze graficznym przy wykorzystaniu CUDA, dla której zaobserwowano dziesięciokrotne przyspieszenie, względem rozwiązania wykorzystującego programy cieniujące w OpenGL. Algorytm ray tracingu w

danym potoku renderowania zaimplementowany został za pomocą Nvidia OptiX, który pozwalał na selektywne dobieranie próbek w obrazie przy minimalnym obciążeniu. System był uruchomiony na komputerze stacjonarnym wyposażonym w procesor Xeon 3.60 Ghz, 32.0 GB pamięci RAM oraz kartę graficzną Nvidia GTX 1080 Ti.



Rysunek 3.11. Etapy potoku renderowania wielopłaszczyznowego wyświetlacza light field'ów.

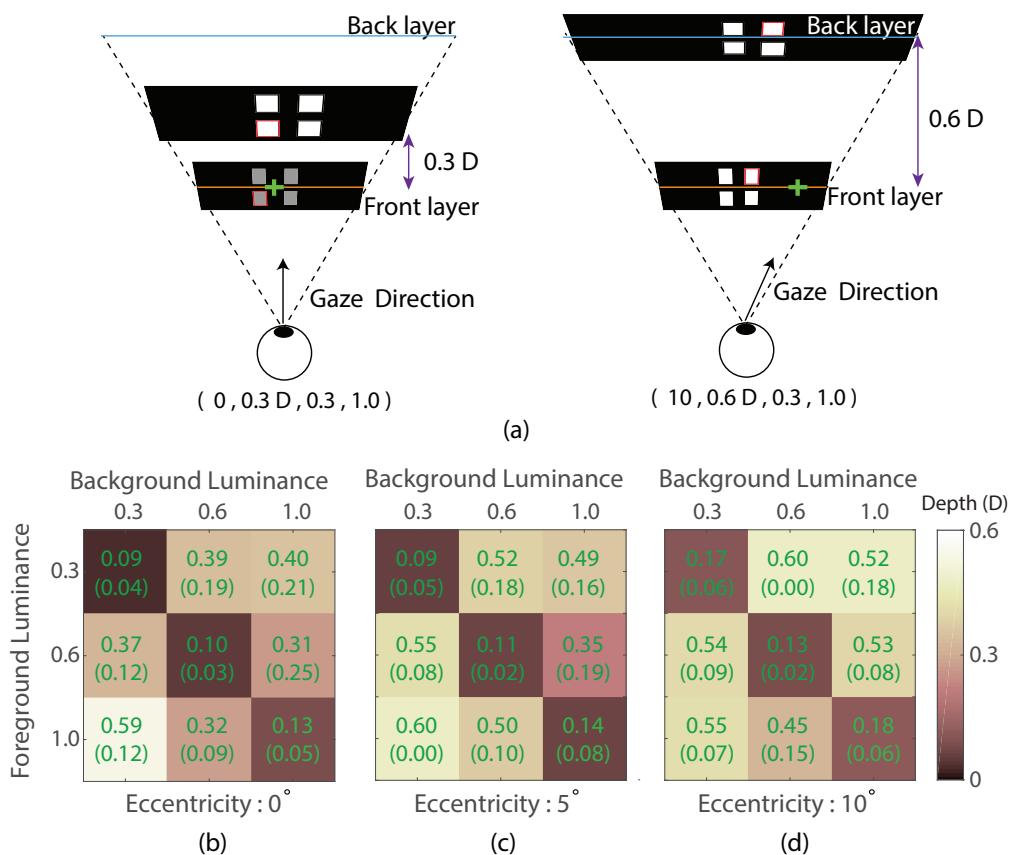
### Maska próbkowania

Metoda hybrydowa bazuje na doborze metody dekompozycji między regułą LB, a LFS. Stworzony wyświetlacz posłużył do wykonania wstępnego eksperymentu, który miał na celu porównać dwie metody pozyskiwania obrazu. Na wyświetlaczu znajdowały się dwie pary bodźców w postaci fragmentów sceny. Jedna para zawierała dwa takie same obrazy wygenerowane, przy wykorzystaniu LFS, druga natomiast składała się z obrazu wygenerowanego poprzez LFS i LB. Kolejność par była losowa, z czego jedna była wyświetlona na górze, druga na dole. Eksperyment miał charakter wymuszonego wyboru spośród dwóch możliwych odpowiedzi. Uczestnicy mieli za zadanie wskazać parę obrazów, pomiędzy którymi widoczna jest różnica. Sceny zawarte w obrazach charakteryzowały się różnym kontrastem Michelsona, głębokością zawartych w niej obiektów oraz częstotliwością kątową od 6 do 15 *cpd*. W eksperymencie brało udział 5 osób, które były instruowane, co mają wykonywać podczas eksperymentu, ale nie miały pojęcia jaki jest cel badań. Oświadczyły one także poprawność, bądź korekcję ich układu wzrokowego. Na podstawie proporcji poprawnych odpowiedzi do wszystkich próbek oraz funkcji psychometrycznej, uzyskane zostały zależności między widocznością artefaktów renderingu, a kontrastem Michelsona (obiektów sceny), ich różną głębokością i częstotliwością kątową.

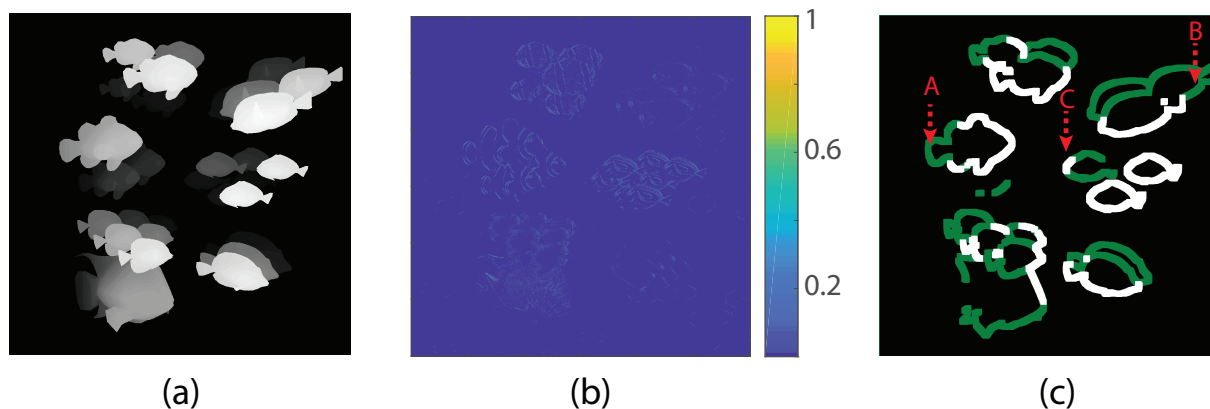
Wspomniany eksperyment wskazał, że reguła LB generowała obraz o odbieranej przez obserwatorów jakości, która jest zbliżona do LFS dla płaskich, oteksturowanych powierzchni, co wskazały także poprzednie badania [76]. Różnica była widoczna dla gra-

nic okluzji obiektów, gdzie reguła LB miała za mało informacji o scenie, aby prawidłowo dokonać dekompozycji obrazu. Widoczność artefaktów renderingu była zależna od głębokości obiektów, luminancji tła i odległości kątowej od obszaru widzenia centralnego. W celu pomiaru tej zależności wykonany został kolejny, podobny do poprzedniego, eksperyment percepcyjny. Jego celem było odnalezienie progu różnicy odległości i głębokości, dla których reguła LB generuje poprawne wyniki. Ten próg uzależniony został od kontrastu luminancji sceny oraz odległości kątowej. W eksperymencie została wykorzystana procedura QUEST [103], która dzięki wykorzystaniu funkcji psychometrycznej, wspomagała dobrać próg odległości, który znajdował się pomiędzy 0.05D, a 0.6D z krokiem równym 0.05D. Wartość ta oznaczała odległość między dwoma obiektami sceny, statycznym przednim oraz tylnym, który był przesuwany w celu dostosowania różnicy głębokości. Pomiar dla różnych odległości kątowych od obszaru widzenia centralnego, był możliwy dzięki wykorzystaniu eye trackera, który weryfikował, czy obserwator skupił swój wzrok na wyświetlonym znaczniku. W sytuacji wykrycia zmiany kierunku patrzenia, bodźce były ukrywane do czasu korekacji kierunku. Bodźce łącznie zajmowały powierzchnię równą  $3^\circ$  pola widzenia. W celu uniknięcia degradacji jakości wyświetlanych obiektów, obrazy były wyświetlane w centralnej części wyświetlacza, a marker oznaczający docelowy kierunek patrzenia znajdował się po lewej stronie.

Na podstawie uzyskanych wyników eksperymentu (Rys. 3.12) zaprojektowana została reguła wyboru pomiędzy LFS, a LB dla poszczególnych pikseli obrazu jako funkcja kontrastu Michelsona, odległości kątowej od obszaru widzenia centralnego oraz głębokości obiektów. Dla pikseli obrazu w płaszczyznach wyświetlacza, przedniej i tylnej, wyznaczany jest kontrast Michelsona oraz różnica głębokości między aktualnym pikselem, a sąsiednimi w obszarze kilku pikseli. Jeśli odległość ta jest większa, niż wartość dopasowana na podstawie wyników eksperymentu to stosowany jest LFS. Chociaż wymagane są dalsze prace dotyczące reguły dobierania metod dekompozycji, tj. zastosowanie charakterystyk czułości na kontrast HVS w obszarach peryferyjnych, przeprowadzone badania już na tym etapie wskazują przyspieszenie procesu renderingu w wielowarstwowych wyświetlaczach light field'ów, wykorzystujących przedstawioną metodę hybrydową. Na Rys. 3.13 znajduje się przykład maski próbkowania dla LFS. Z mapy głębokości (Rys. 3.13(a)) i kontrastu Michelsona (Rys. 3.13(b)) generowana jest maska wskazująca piksele, dla których konieczne jest zastosowanie LFS (Rys. 3.13(c)). W przedstawionym przykładzie kierunek patrzenia skierowany był ku środkowi wyświetlacza. Przykłady A i B wskazują na wyeliminowane przypadkowej różnicy między LB i LFS z powodu dużej odległości kątowej od obszaru widzenia centralnego, gdzie zdolność do spostrzeżenia różnic jest mniejsza. Przykład C wskazuje konieczność wykorzystania LFS, mimo niskiej różnicy głębokości obiektów, co wynika z niskiego kontrastu, pomiędzy pobliskimi krawędziami.



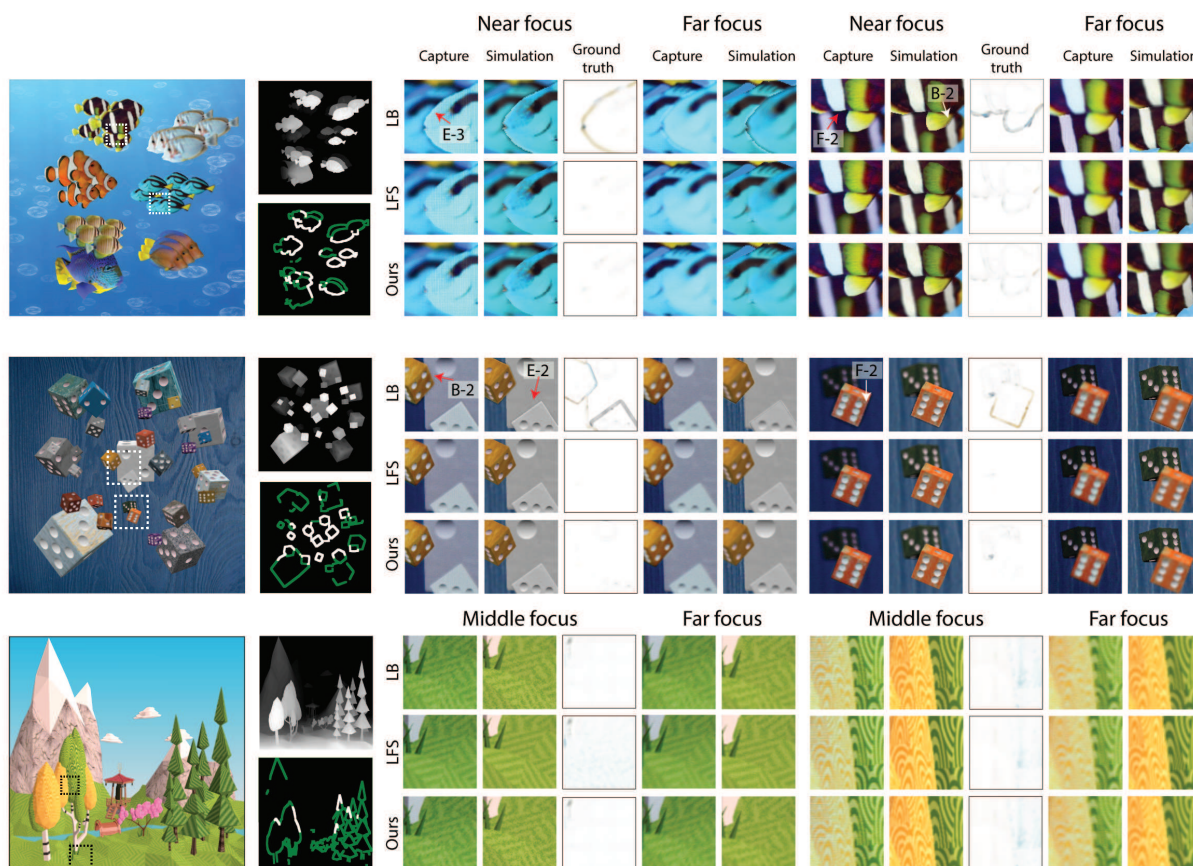
Rysunek 3.12. Wynik pomiaru zależności między luminancją sceny i odległość kątową od obszaru widzenia centralnego a dostrzegalną różnicą głębokości. Źródło własne [W7] (Rys. 5).



Rysunek 3.13. Przykład maski próbkowania LFS. (a) Mapa głębokości. (b) Mapa kontrastów Michelsona. (c) Wygenerowana maska. Białe obszary na masce wskazują konieczność użycia LFS dla kierunku patrzenia skierowanego ku środkowi wyświetlacza. Zielone obszary maski zawierają punkty, w których artefakty renderingu są niewidoczne dla dalszych odległości kątowych od obszaru widzenia centralnego. Źródło własne [W7] (Rys. 9).

### 3.2.4. Pomiar wydajności wielopłaszczyznowego wyświetlacza light field'ów

Aby zbadać możliwości zbudowanego wielowarstwowego wyświetlacza light field'ów, przygotowano zostały trzy sceny - Fish, Dice i Forest. Sceny te przedstawione zostały na Rys. 3.14. . W pierwszej kolumnie znajduje się obraz sceny przed dekompozycją. W kolejnej kolumnie, na górze, znajduje się mapa głębokości, na dole natomiast maska LFS (białe obszary wskazują konieczność użycia LFS, zielone obszary to punkty, w których różnica między LB, a LFS jest niewidoczna dla dalszych odległości kątowych od obszaru widzenia centralnego). Dla każdej sceny przedstawione zostały fragmenty obrazu wygenerowane każdą metodą. W wierszach od góry znajdują się kolejno: reguła LB, LFS oraz metoda hybrydowa. Obrazy wyświetlane w przedniej wirtualnej płaszczyźnie (ang. *near focus*), znajdujące się w kolumnach 3-5 oraz 8-10, pochodzą z kolejno przechwyconego obrazu z wyświetlacza, symulacji komputerowej oraz mapy SSIM (podobieństwa strukturalnego, ang. *structural similarity*) [36], która przedstawia różnicę między symulacją, a algorytmami wykorzystanymi w poprzednich pracach. Obrazy wyświetlane w dalszej wirtualnej płaszczyźnie znajdują się w kolumnach 6-7 oraz 11-12 (ang. *far focus*).



Rysunek 3.14. Sceny wykorzystane do pomiaru wydajności wielopłaszczyznowego wyświetlacza light field'ów oraz jakości wyświetlanych przez niego obrazów. Od góry sceny: Fish, Dice i Forest. Źródło własne [W7] (Rys. 12).



Dla trzech scen, zmierzony został czas renderingu przy trybie wyświetlania jednocznego, a wyniki zostały przedstawione w Tab. 3.2. Dekompozycja wykonywana została z liczbą iteracji równą 10. Wyniki wskazują na zależność między złożonością sceny, a czasem renderingu i dekompozycji. Co więcej, wykorzystana tutaj procedura generowania obrazu wielopłaszczyznowego jest wydajniejsza, niż rendering przy wykorzystaniu LFS dla pełnego obrazu, co wskazuje na korzyść płynącą z redukcji ilości próbek generowanych tą metodą na rzecz reguły LB. Wynika to z tego, że dla LFS próbki sceny brane są z dziewięciu różnych pozycji kamery, natomiast dla LB próbka brana jest wyłącznie z widoku centralnego. Pomiar wskazuje także, że czas dekompozycji jest zależny wyłącznie od wypełnienia procentowego maski. Szybkość renderingu, jako częstotliwość generowania kolejnych klatek, wynosi 84 Hz (4.25), 60 Hz (4.06) i 103 Hz (4.50) kolejno dla scen Fish, Dice i Forest. Wartości w nawiasach oznaczają multiplikatywny zysk prędkości w stosunku do pełnego renderowania za pomocą LFS, po odjęciu stałego kosztu generowania obrazu z centralnego widoku. W warunkach wyświetlania sceny dla obu oczu, całkowity czas renderingu należy zwiększyć 2-krotnie.

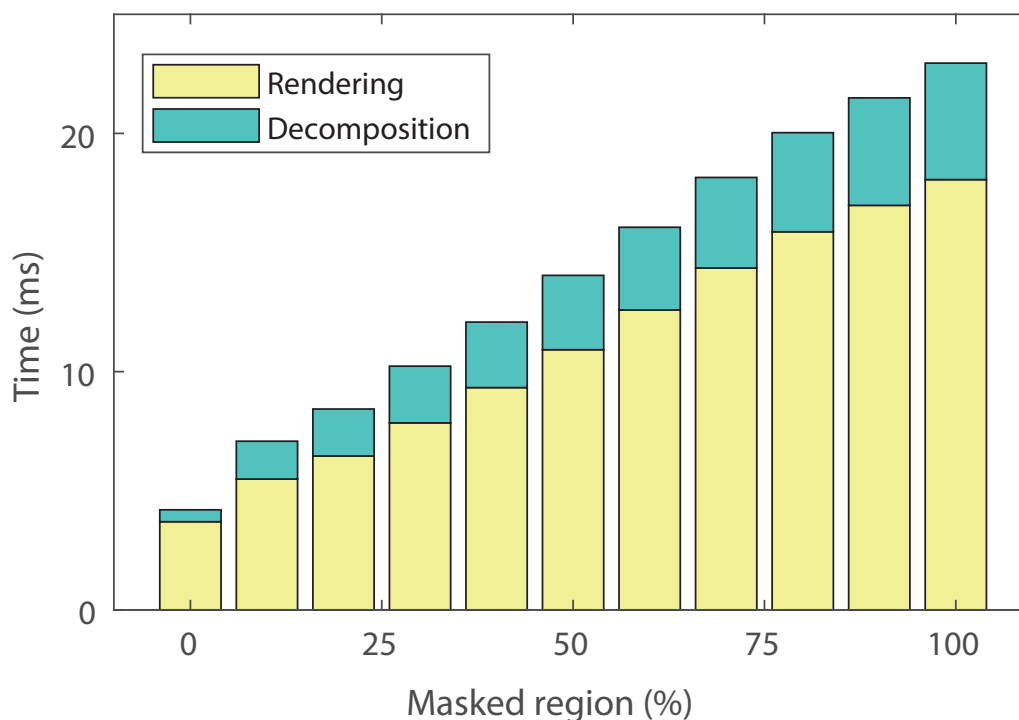
| Scena  | Ilość wielokątów | Maska [%] | Czas Renderingu [ms] | Czas Dekompozycji [ms] |
|--------|------------------|-----------|----------------------|------------------------|
| Fish   | 20498            | 7.3       | 9.26 (27.48)         | 2.57 (4.11)            |
| Dice   | 569810           | 6.5       | 14.11 (47.08)        | 2.44 (4.12)            |
| Forest | 16924            | 1.8       | 7.29 (28.31)         | 2.35 (4.19)            |

Tabela 3.2. Wyniki pomiaru wydajności metody hybrydowej dla różnych scen. Wartości w nawiasach wskazują czasy pełnego renderingu LFS. Źródło własne [W7] (Tab. 1).

W celu zbadania wpływu proporcji wypełnienia maski na wydajność potoku renderowania, zmierzone zostały czasy renderingu i dekompozycji dla różnych gęstości próbek zawartych w masce. Maski zawierały w sobie piksele dobrane losowo, a ich wypełnienie wynosiło ustalony procent. Wyniki pomiaru umieszczone zostały na Rys. 3.15. Zerowy procent próbek oznacza rendering wyłącznie przy wykorzystaniu reguły LB, natomiast 100% to pełen rendering LFS. Całkowity czas potoku rośnie liniowo wraz ze wzrostem zamaskowanego obszaru, co sugeruje, że obciążenie generowane przez selektywne renderowanie jest minimalne.

### 3.2.5. Ocena jakości hybrydowej metody dekompozycji

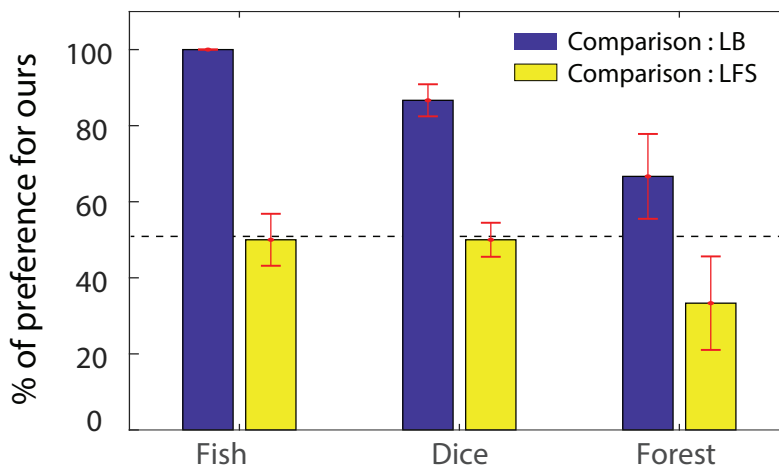
Jakość wyświetlanego obrazu zmierzona została w eksperymencie percepcyjnym. Porównane w nim zostały pary metod: metoda hybrydowa i reguła LB oraz metoda hybrydowa i LFS. Pomiar odbył się dla trzech statycznych scen wyświetlonych dla obu oczu. W celu zapobiegnięcia wpływu czasu renderingu na uzyskane wyniki, obrazy sceny były przygotowane przed procedurą eksperymentu. Ustalenie kierunku patrzenia obserwatora



Rysunek 3.15. Czasy renderingu i dekompozycji dla różnych proporcji wypełnienia maski. Żółtym kolorem oznaczony został czas renderingu, zielonym natomiast czas dekompozycji. Źródło własne [W7] (Rys. 11).

polegało na wyświetleniu znacznika i instruowanie uczestników eksperymentu, aby skupili na nim swój wzrok, co weryfikowane było przez eye tracker. W każdej próbie użytkownicy proszeni byli o wybranie sceny zapewniającej lepszą jakość spośród dwóch wyświetlonych obrazów. Dla każdej sceny pomiar odbył się dla pięciu różnych odległości kątowych od obszaru widzenia centralnego. W eksperymencie wzięło udział sześć osób. Wyniki eksperymentu przedstawione zostały na Rys. 3.16.

Na podstawie wyników eksperymentu można zaobserwować wysoką preferencję do wyboru metody hybrydowej przedstawionej w tej pracy nad regułą LB. Ten wybór można wytłumaczyć lepszą rekonstrukcją krawędzi w proponowanej metodzie. Dla sceny Forest różnica między LB, a prezentowanym algorytmem maleje, co może być skutkiem cech sceny, która składa się głównie z oteksturowanych obszarów o małym zagęszczeniu okluzji obiektów. W porównaniu z LFS, ilość odpowiedzi wskazujących na proponowaną metodę jest bliska 50% dla scen Fish i Dice, co wskazuje, że obserwatorzy byli obojętni w wyborze między dwoma metodami. Biorąc pod uwagę fakt, że te dwie sceny zawierają wiele okluzji obiektów, wyniki te sugerują, że wykorzystywany algorytm maskowania działa prawidłowo. Dla sceny Forest, LFS był bardziej preferowanym algorytmem, co może być spowodowane zawartością sceny, w której znajdują się obiekty o wysokiej częstotliwości. Uczestnicy eksperymentu twierdzili, że w tym przypadku preferują rozmytą teksturę generowaną przez LFS, niż wykorzystanie proponowanej w tej pracy metody.



Rysunek 3.16. Wyniki eksperymentu mierzącego jakość metody hybrydowej. Niebieskie i żółte słupki wskazują procent odpowiedzi uczestników wskazujących lepszą jakość obrazów renderowanych przy wykorzystaniu naszej metody w porównaniu z kolejno regułą LB i LFS. Czerwone słupki wskazują błąd standardowy. Źródło własne [W7] (Rys. 13).

Ze względu na to, że rekonstrukcja tekstur o wysokiej częstotliwości wymaga precyzyjnego wyrównania pikseli, małe ruchy źrenicy i głowy mogą prowadzić do postrzegania tych cech jako informacji zawierającej szum. Należy także wspomnieć, że proponowany układ optyczny charakteryzuje się niską jakością wyświetlanych obrazów w zewnętrznych obszarach, co spowodowane jest oparciem architektury systemu na soczewkach powiększających. Zmniejszenie degradacji obrazu może zostać rozwiązane poprzez wykorzystanie wyświetlacza wielopłaszczyznowego z holograficznymi elementami optycznymi [50].

## Podsumowanie rozdziału

W rozdziale zaproponowano techniki, które można wykorzystać przy badaniu czułości na kontrast układu wzrokowego człowieka w obszarach widzenia peryferyjnego. Obydwie metody polegały na wykorzystaniu tej cechy do redukcji liczby generowanych próbek sceny, w celu syntezy końcowego obrazu. Redukcja ta wiązała się ze zmniejszeniem jakości wyświetlanego obrazu, lecz w sposób nie pogarszający percepcyjnego odbioru sceny przez obserwatora. Głównie wiązało się to ze zmniejszeniem częstotliwości próbkowania w obszarze widzenia peryferyjnego.

W metodzie śledzenia promieni o niejednorodnym próbkowaniu, dystrybucja promieni pierwotnych uzależniona jest od odległości kątowej i obszaru widzenia centralnego. Pomiar wydajności systemu renderingu wskazał znaczącą redukcję czasu wymaganego do syntezy klatki obrazu, natomiast eksperyment percepcyjny wskazał brak dostrzegalnej redukcji jakości. Problemem w tym rozwiązaniu okazał się czasowy aliasing, który związany

może być z ruchami gałek ocznych, co ma związek z fiksacją i może być rozwiązany przez technikę wygładzania krawędzi, która korzysta z tych danych.

Wielopłaszczyznowy wyświetlacz light field'ów, korzystający z metody hybrydowej, który łączy regułę LB i LFS, wskazuje ciekawe pole do badań nad wykorzystaniem cech HVS w wyświetlaczach VR i AR. Z uwagi na rosnącą popularność tych wyświetlaczy oraz wymagania związane z osiągnięciem dopuszczalnego progu szybkości syntezy kolejnych próbek obrazu, które wymagane są do poprawnego użytkowania, przyspieszenie tych systemów przy użyciu zmierzonych charakterystyk może stać się integralną częścią takich systemów w przyszłości.

## Podsumowanie

W pracy zostało zaprezentowane wykorzystanie ograniczonej czułości na kontrast układu wzrokowego człowieka w algorytmach syntezy obrazu. Zmierzone progi kontrastów wskazują na zmniejszoną dostrzegalność bodźców wzrokowych w obszarze widzenia peryferyjnego. Badania, które umożliwiły zbadanie progów kontrastów bodźców chromatycznych, achromatycznych i mieszanych pozwoliły zbudować model analityczny, dzięki któremu możliwe było wskazanie dostrzegalnych bodźców we wskazanym obrazie. Model ten umożliwia implementację systemu renderingu uzależnionego od kierunku patrzenia.

W celu zastosowania czułości na kontrast w systemie renderingu musi on przechwytywać punkt skupienia wzroku obserwatora. Czas pomiędzy otrzymaniem tej informacji a synteza klatki obrazu musi być na tyle krótki, aby podczas zmiany kierunku patrzenia obserwator nie zauważył redukcji jakości obrazu w obszarach widzenia peryferyjnego. Przeprowadzony eksperyment percepcyjny wykazał, że całkowite opóźnienie klasycznego systemu renderingu jest zbyt duże, aby ukryć zmianę LOD obiektów. Jednak dla obiektu położonego w odległości  $35^\circ$  od obszaru widzenia centralnego i opóźnienia równemu 24 ms, przerysowanie obiektów było niezauważalne.

Jako przykład zastosowania charakterystyk czułości na kontrast w grafice komputerowej czasu rzeczywistego, zaprezentowany został algorytm śledzenia promieni w środowisku VR, który uzależnia rozkład próbkowania sceny od kierunku patrzenia obserwatora. Mniejsze zagęszczenie śledzonych promieni w obszarze widzenia peryferyjnego skraca czas renderowania i ma znikomy wpływ na odbieraną jakość obrazu, szczególnie dla maski próbkowania zgodnej ze zmierzonymi progami kontrastu. Drugim przykładem zastosowania cech układu wzrokowego człowieka była metoda hybrydowa dekompozycji obrazu w wyświetlaczu wielopłaszczyznowym light field'ów. Metoda ta dobierała mało kosztowne obliczeniowo reguły liniowego mieszania oraz syntezy light field'ów, które wymagały większych nakładów obliczeniowych. Dobór reguły uzależniony był od kierunku patrzenia obserwatora, ale również zawartości sceny, w której dla oteksturowanych powierzchni możliwe było dobranie dowolnej reguły, podczas gdy okluzja obiektów wymuszała dobór droższej obliczeniowo metody. Zastosowana strategia doboru metody umożliwiła redukcję czasu renderingu i dekompozycji, a tym samym umożliwienie syntezy obrazu na tym wyświetlaczu w czasie rzeczywistym.

## Wnioski końcowe

Na podstawie opisanych w rozprawie prac badawczych, wyciągnięto następujące wnioski:

- Stworzone środowisko, wyspecjalizowane do pomiaru czułości na kontrast, które składało się ze skalibrowanego wyświetlacza i eye trackera, umożliwiło wykonanie eksperymentu percepcyjnego, który wskazał dostrzegalne progi kontrastu, które z kolei były dostrzegalne przez układ wzrokowy człowieka w obszarach widzenia peryferyjnego (Rozdział.2.1).
- Uzyskane progi dostrzegalnych kontrastów achromatycznych, chromatycznych i mieszanych, wskazują na mniejszą zdolność do ich spostrzeżenia przez układ wzrokowy człowieka w obszarach widzenia peryferyjnego (Rozdział 2.2.2, opublikowane w [W3]).
- Obrona metodologia, którą wykorzystano podczas pomiarów dostrzegalnych progów kontrastów w obszarach widzenia peryferyjnego, oparta została na śledzeniu wzroku obserwatora. Daje to wyniki zgodne z poprzednimi technikami pomiaru, w których wyświetlany był migający bodziec wzrokowy (Rozdział 2.2.4, opublikowane w [W2]).
- Zaproponowany model analityczny, który ukazuje dostrzegalne przez układ wzrokowy człowieka bodźce, umożliwia wskazanie obszarów obrazu, w których można zredukować gęstość próbkowania bez pogorszenia jakości (Rozdział 2.3).
- Systemy renderingu, który jest oparty na kierunku patrzenia, musi charakteryzować się niskim, zmierzonym opóźnieniem wyświetlania kolejnych klatek, względem zmiany kierunku patrzenia. Ma to na celu uniknięcie dostrzeżenia redukcji jakości obrazu w obszarach widzenia peryferyjnego (Rozdział 2.4, opublikowane w [W1]).
- Zmierzone progi dostrzegalnych kontrastów umożliwiły redukcję gęstości próbkowania sceny w obszarach widzenia peryferyjnego, co skutkowało przyspieszeniem renderingu, bez widocznej redukcji jakości odbieranego obrazu przez obserwatora (Rozdział 3.1, opublikowane w [W6]).
- Zastosowanie charakterystyk czułości na kontrast może przyspieszyć algorytm dekompozycji obrazu. Przedstawienia go na wyświetlaczu stereoskopowym, obsługującym wskazówki o akomodacji oczu, umożliwia rendering w czasie rzeczywistym (Rozdział. 3.2, opublikowane w [W7]).

## Odniesienie do tezy

Na początku rozprawy postawiono tezę:

*Uwzględnienie progów widzialności informacji dla peryferyjnych obszarów widzenia umożliwia zmniejszenie częstotliwości próbkowania a co za tym idzie przyspieszenie syntezy obrazów.*

W toku przeprowadzonych badań wykazano możliwość wykorzystania zmierzonych progów kontrastu, które są dostrzegalne przez układ wzrokowy człowieka, w algorytmach grafiki czasu rzeczywistego. Malejąca wraz z odległością od obszaru widzenia centralnego czułość sprawia, że szczegóły obiektów, zawierające się w tych obszarach obrazu wirtualnej sceny, są niezauważalne dla obserwatora. Z tego wywnioskować można, że informacja wysokoczęstotliwościowa jest niedostrzegalna w obszarze widzenia peryferyjnego.

Unikanie renderingu przy niedostrzegalnej informacji odbywa się poprzez redukcję częstotliwości próbkowania w danych obszarach wynikowego obrazu. Gęstość tych próbek maleje zgodnie ze zmierzoną charakterystyką czułości na kontrast, czyli wraz ze wzrostem odległości od punktu skupienia wzroku obserwatora, co sprawia, że czas syntezy obrazu sceny jest redukowany, przy czym redukcja ta nie wpływa na jego odbiór przez percepcję człowieka.

W ramach rozprawy potwierdzona została wcześniej postawiona teza.

## Dalsze prace

Pomiary czułości na kontrast wykonane w niniejsze pracy skupiły się wyłącznie na ograniczonym zakresie odległości kątowych, częstotliwości bodźca oraz luminancji tła. Do odkrycia pełnej zależności pomiędzy dostrzegalnymi progami kontrastów a wspomnianymi parametrami bodźca wzrokowego. Jest to kluczowe szczególnie dla zaproponowanego modelu analitycznego, który wskazuje prawdopodobieństwo wykrycia bodźca wzrokowego. W tej chwili jest on ograniczony do bodźców o częstotliwości 2 *cpd*, co sprawia, że nie wykorzystuje on w pełni możliwości redukcji częstotliwości próbkowania, gdyż dla innych częstotliwości czułość na kontrast jest jeszcze niższa. Rozszerzenie modelu wymaga większej ilości pomiarów czułości na kontrast, głównie dla większego zakresu częstotliwości, lecz kluczowy jest także większy zakres luminancji tła. Optymalizacja modelu może także wymagać większej liczby bodźców wzrokowych, o różnych kierunkach koloru.

Zmierzone opóźnienie systemu renderingu jest zbyt duże, aby ukryć zmianę geometrii obiektów sceny, które są oddalone o małą odległość od obszaru widzenia centralnego. Dalsze prace wymagają zmniejszenia tej wartości czasowej, czego przykładem może być zastosowanie eye tracker, przechwytyującego kierunek patrzenia w częstotliwości, której

wartość jest bliższa dzisiejszym wyświetlaczom. Przeprowadzone badania wymagają także weryfikacji w rozbudowanym systemie grafiki komputerowej czasu rzeczywistego, którym mogą być gry komputerowe.

W zaprezentowanym algorytmie śledzenia promieni o niejednorodnym rozkładzie próbek występuje znaczący problem migotania w obszarze widzenia peryferyjnego. Powoduje go aliasing, co jest zależne od wykorzystania próbek o mniejszej gęstości. Problem ten należy rozwiązać poprzez stosowanie metody anty-aliasingu, która jest oparta na śledzeniu kierunku patrzenia obserwatora, szczególnie uwzględniając ruchy związaną z fiksacją gałek ocznych.

Algorytm hybrydowej dekompozycji obrazu na wirtualne płaszczyzny może zostać wykorzystany w multiplikatywnym wyświetlaczu wielopłaszczyznowym, w przeciwieństwie do zaprezentowanego zastosowania w wyświetlaczu addytywnym. Ciekawym tematem badań w tym kierunku może być fakt, iż główne artefakty graficzne występujące w regule liniowego mieszania, wynikają z natury wyświetlaczy addytywnych. Zastosowanie zaprezentowanej metody w wyświetlaczu multiplikatywnym może doprowadzić do interesujących spostrzeżeń na polu poprawnej dekompozycji obrazu sceny, celem wyświetlenia go na wirtualnych płaszczyznach.



# Spis rysunków

|       |  |    |
|-------|--|----|
| 1.1.  | Schemat budowy gałki ocznej człowieka . . . . .  | 10 |
| 1.2.  | Zagęszczenie fotoreceptorów na siatkówce oka . . . . .   | 12 |
| 1.3.  | Podział pola widzenia układu wzrokowego na obszary . . . . .   | 12 |
| 1.4.  | Przykład kontrastu jako achromatyczna funkcja progowa . . . . .  | 16 |
| 1.5.  | Funkcja czułości na kontrast układu wzrokowego człowieka . . . . .   | 18 |
| 2.1.  | Jednowymiarowy Wzór Gabora . . . . .   | 23 |
| 2.2.  | Bodźce wzrokowe wykorzystane w pomiarach czułości na kontrast . . . . .  | 26 |
| 2.3.  | Kierunki kolorów bodźców wzrokowych . . . . .  | 28 |
| 2.4.  | Procedura eksperymentu percepcyjnego mierzącego czułość na kontrast . . . . .  | 30 |
| 2.5.  | Przebieg procedury QUEST . . . . .   | 31 |
| 2.6.  | Kalibracja wyświetlacza . . . . .  | 33 |
| 2.7.  | Pomiar luminancji i chrominancji bodźców wzrokowych . . . . .  | 34 |
| 2.8.  | Wyniki pomiaru czułości na kontrast dla bodźca achromatycznego w obszarze<br>widzenia centralnego . . . . .              | 36 |
| 2.9.  | Wyniki pomiaru czułości na kontrast dla bodźca achromatycznego w obszarze<br>widzenia peryferyjnego . . . . .            | 37 |
| 2.10. | Dopasowanie progów kontrastu w przestrzeni logarytmicznej do równania liniowego  | 38 |
| 2.11. | Porównanie uzyskanego dopasowania do poprzednich prac . . . . .  | 39 |
| 2.12. | Uzyskane pomiary czułości na kontrast achromatycznego bodźca wzrokowego dla<br>poszczególnych obserwatorów . . . . .     | 39 |
| 2.13. | Wyniki pomiaru czułości na kontrast dla bodźców chromatycznych i mieszanych . .  | 40 |
| 2.14. | Wyniki pomiaru czułości na kontrast dla bodźców chromatycznych i mieszanych<br>dla poszczególnych obserwatorów . . . . . | 41 |
| 2.15. | Wyniki pomiarów przy wykorzystaniu metodologii losowej i statycznej orientacji<br>bodźca . . . . .                       | 42 |
| 2.16. | Porównanie metodologii losowej i statycznej orientacji bodźca . . . . .  | 43 |
| 2.17. | Funkcja psychometryczna w modelu analitycznym czułości na kontrast . . . . .   | 46 |
| 2.18. | Schemat Boyntona-MacLeoda dostrzegalnych progów kontrastu uzyskana w modelu<br>analitycznym . . . . .                    | 48 |
| 2.19. | Przykłady map prawdopodobieństwa uzyskana w analitycznym modelu czułości na<br>kontrast . . . . .                        | 49 |

|  |    |
|--|----|
| 2.20. Schemat systemu renderingu kierunkowego z techniką LOD . . . . .   | 52 |
| 2.21. Stanowisko eksperymentalne . . . . .   | 53 |
| 2.22. Procedura eksperymentu mierzącego akceptowalnego opóźnienie systemu renderingu opartego o kierunek patrzenia . . . . .     | 54 |
| 2.23. Stanford Bunny . . . . .   | 55 |
| 2.24. Wyniki pomiaru zdolności do rozróżnienia obiektu o niższym LOD i referencyjnego .  | 56 |
| 2.25. Obiekty wykorzystane do pomiaru akceptowalnego opóźnienia systemu renderingu opartego o kierunek patrzenia . . . . .       | 56 |
| 2.26. Schemat budowy systemu renderingu LOD uwzględniającego kierunkowość widzenia   | 57 |
| 2.27. Wyniki eksperymentu mierzącego akceptowalne opóźnienie systemu renderingu uwzględniającego kierunkowość widzenia . . . . . | 58 |
| 2.28. Zdjęcie eye trackera o wysokiej częstotliwości działania . . . . .   | 59 |
| 2.29. Oprogramowanie systemu eye trackingu . . . . .   | 60 |
| 3.1. Częstotliwość odcinania próbkowania sceny zgodna z funkcją czułości na kontrast .   | 64 |
| 3.2. Maski, siatka i przykład renderingu algorytmu śledzenia promieni o niejednorodnym próbkowaniu . . . . .                     | 65 |
| 3.3. Maski próbkowania obrazu w kierunkowym śledzeniu promieni . . . . .   | 66 |
| 3.4. Częstotliwości odcinania próbkowania masek wykorzystanych w algorytmie śledzenia promieni . . . . .                         | 66 |
| 3.5. Schemat architektury systemu renderingu uwzględniającego kierunkowość widzenia .  | 67 |
| 3.6. Sceny Air Shed i Bunny Box . . . . .  | 68 |
| 3.7. Przykłady renderingu scen przy użyciu masek próbkowania . . . . .   | 69 |
| 3.8. Wyniki eksperymentu percepcyjnego mierzącego jakość renderingu uwzględniającego kierunkowość widzenia . . . . .             | 71 |
| 3.9. Koncepcja wyświetlaczy wielowarstwowych . . . . .   | 73 |
| 3.10. Schemat i zdjęcie wielopłaszczyznowego wyświetlacza light field'ów . . . . .   | 76 |
| 3.11. Etapy potoku renderowania wielopłaszczyznowego wyświetlacza light field'ów. . . .  | 77 |
| 3.12. Wynik pomiaru dostrzegalnej różnicy głębokości obiektów sceny . . . . .  | 79 |
| 3.13. Przykład maski próbkowania LFS . . . . .   | 79 |
| 3.14. Sceny wyświetlone w wyświetlaczu wielopłaszczyznowym light field'ów . . . . .  | 80 |
| 3.15. Czasy renderingu i dekompozycji dla różnych proporcji wypełnienia maski próbkowania. . . . .                               | 82 |
| 3.16. Wyniki eksperymentu mierzącego jakość metody hybrydowej . . . . .  | 83 |

## Spis tabel

|      |   |    |
|------|---|----|
| 2.1. | Wartości $Y_{xy}$ i LMS bodźców wzrokowych . . . . .  | 28 |
| 2.2. | Wyniki pomiarów czułości na kontrast dla bodźca wzrokowego o losowej i statycznej orientacji . . . . .    | 43 |
| 2.3. | Porównanie progów dostrzegalnego kontrastu uzyskanego w modelu i pomiarach . .                            | 47 |
| 3.1. | Wyniki pomiaru czasu syntezy obrazów dla pełnej rozdzielczości oraz niejednorodnego próbkowania . . . . . | 69 |
| 3.2. | Wyniki pomiaru wydajności metody hybrydowej w wyświetlaczu wielopłaszczyznowym light field'ów . . . . .   | 81 |



## Publikacje własne

- [W1] Michał Chwesiuk, Radosław Mantiuk. Acceptable System Latency for Gaze-Dependent Level of Detail Rendering. *Proceedings of Eurographics, Posters, 2016*. Lisbon, Portugal.
- [W2] Michał Chwesiuk, Radosław Mantiuk. Measurements of Contrast Detection Thresholds for Peripheral Vision Using Non-flashing Stimuli. *Proceedings of the 9th KES International Conference on Intelligent Decision Technologies Part II, 2017*, pp. 258-267. Vilamoura, Portugal.
- [W3] Michał Chwesiuk, Radosław Mantiuk. Measurements of Contrast Thresholds for Peripheral Vision. *Proceedings of the ACM Symposium on Applied Perception 2019*. Barcelona, Spain.
- [W4] Mariusz Pasiński, Michał Chwesiuk. Towards Interactive Rendering for Lighting Design. *Proceedings of Central European Seminar on Computer Graphics, 2017*. Smolenice, Slovakia.
- [W5] Sumit Shekhar, Shida Kunz Beigpour, Matthias Ziegler, Michał Chwesiuk, Dawid Pałeń, Karol Myszkowski, Joachim Keinert, Radosław Mantiuk, Piotr Didyk. Light-Field Intrinsic Dataset. *British Machine Vision Conference, 2018*. Newcastle, England.
- [W6] Adam Siekawa, Michał Chwesiuk, Radosław Mantiuk, Rafał Piórkowski. Foveated Ray Tracing for VR Headsets. *Proceedings of the 25th International Conference on Multimedia Modeling Part I, 2019*, pp. 106-117. Thessaloniki, Greece.
- [W7] Hyeonseung Yu, Mojtaba Bemana, Marek Wernikowski, Michał Chwesiuk, Okan Tarahan Tursun, Gurprit Singh, Karol Myszkowski, Radosław Mantiuk, Hans-Peter Seidel, Piotr Didyk. A Perception-driven Hybrid Decomposition for Multi-layer Accommodative Displays. *IEEE Transactions on Visualization and Computer Graphics 25, no. 5 (2019)*, pp. 1940-1950. Osaka, Japan.



# Literatura

- [1] Kurt Akeley, Simon J Watt, Ahna Reza Girshick, Martin S Banks. A stereo display prototype with multiple focal distances. *ACM transactions on graphics (TOG)*, wolumen 23, strony 804–813. ACM, 2004.
- [2] Tomas Akenine-Moller, Eric Haines, Naty Hoffman. *Real-time rendering*. AK Peters/CRC Press, 2019.
- [3] Kaan Akşit, Ward Lopes, Jonghyun Kim, Peter Shirley, David Luebke. Near-eye varifocal augmented reality display using see-through screens. *ACM Transactions on Graphics (TOG)*, 36(6):189, 2017.
- [4] Anders H Andersen, Avinash C Kak. Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm. *Ultrasonic imaging*, 6(1):81–94, 1984.
- [5] Stephen J Anderson, Kathy T Mullen, Robert F Hess. Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors. *The Journal of Physiology*, 442(1):47–64, 1991.
- [6] ST Aw, T Haslwanter, GM Halmagyi, IS Curthoys, RA Yavor, MJ Todd. Three-dimensional vector analysis of the human vestibuloocular reflex in response to high-acceleration head rotations. i. responses in normal subjects. *Journal of neurophysiology*, 76(6):4009–4020, 1996.
- [7] Peter GJ Barten. *Contrast sensitivity of the human eye and its effects on image quality*, wolumen 21. Spie optical engineering press Bellingham, WA, 1999.
- [8] Gary Bishop, David M Weimer. Fast phong shading. *ACM SIGGRAPH Computer Graphics*, wolumen 20, strony 103–106. ACM, 1986.
- [9] Robert M Boynton, Naotake Kambe. Chromatic difference steps of moderate size measured along theoretically critical axes. *Color Research & Application*, 5(1):13–23, 1980.
- [10] David H Brainard. The psychophysics toolbox. *Spatial vision*, 10:433–436, 1997.
- [11] Mark W Cannon. Perceived contrast in the fovea and periphery. *JOSA A*, 2(10):1760–1768, 1985.
- [12] Steven A Cholewiak, Gordon D Love, Pratul P Srinivasan, Ren Ng, Martin S Banks. Chromablur: Rendering chromatic eye aberration improves accommodation and realism. *ACM Transactions on Graphics (TOG)*, 36(6):210, 2017.
- [13] Paolo Cignoni, Marco Callieri, Massimiliano Corsini, Matteo Dellepiane, Fabio Ganovelli,

- Guido Ranzuglia. Meshlab: an open-source mesh processing tool. *Eurographics Italian chapter conference*, wolumen 2008, strony 129–136, 2008.
- [14] James H Clark. Hierarchical geometric models for visible surface algorithms. *Communications of the ACM*, 19(10):547–554, 1976.
- [15] Thomas F Coleman, Yuying Li. A reflective newton method for minimizing a quadratic function subject to bounds on some of the variables. *SIAM Journal on Optimization*, 6(4):1040–1058, 1996.
- [16] Tom N Cornsweet. Determination of the stimuli for involuntary drifts and saccadic eye movements. *JOSA*, 46(11):987–993, 1956.
- [17] Christine A Curcio, Kimberly A Allen, Kenneth R Sloan, Connie L Lerea, James B Hurley, Ingrid B Klock, Ann H Milam. Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *Journal of Comparative Neurology*, 312(4):610–624, 1991.
- [18] Stanislas Dehaene. The neural basis of the weber–fechner law: a logarithmic mental number line. *Trends in cognitive sciences*, 7(4):145–147, 2003.
- [19] Gyorgy Denes, George Ash, Huameng Fang, Rafał K Mantiuk. A visual model for predicting chromatic banding artifacts. 2019.
- [20] Russell L DeValois, Karen K DeValois. *Spatial vision*, wolumen 14. Oxford university press, 1990.
- [21] MA Díez-Ajenjo, P Capilla, MJ Luque. Red-green vs. blue-yellow spatio-temporal contrast sensitivity across the visual field. *Journal of Modern Optics*, 58(19-20):1736–1748, 2011.
- [22] RW Ditchburn, BL Ginsborg. Vision with a stabilized retinal image. *Nature*, 170(4314):36, 1952.
- [23] Raymond Dodge. Five types of eye movement in the horizontal meridian plane of the field of regard. *American journal of physiology-legacy content*, 8(4):307–329, 1903.
- [24] Andrew T Duchowski. Eye tracking methodology. *Theory and practice*, 328(614):2–3, 2007.
- [25] David Dunn, Cary Tippets, Kent Torell, Petr Kellnhofer, Kaan Akşit, Piotr Didyk, Karol Myszkowski, David Luebke, Henry Fuchs. Wide field of view varifocal near-eye display using see-through deformable membrane mirrors. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1322–1331, 2017.
- [26] Gustav Theodor Fechner, Davis H Howes, Edwin Garrigues Boring. *Elements of psychophysics*, wolumen 1. Holt, Rinehart and Winston New York, 1966.
- [27] Burkhard Fischer, E Ramsperger. Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, 57(1):191–195, 1984.
- [28] Masahiro Fujita, Takahiro Harada. Foveated real-time ray tracing for virtual reality headset. *Light Transport Entertainment Research*, 2014.
- [29] Catherine M Gdyczynski, Amir Manbachi, SayedMasoud Hashemi, Bahman Lashkari, Richard S C Cobbold. On estimating the directionality distribution in pedicle trabecular bone from micro-ct images. *Physiological measurement*, 35(12):2415, 2014.
- [30] James Gordon, Israel Abramov. Color vision in the peripheral retina. ii. hue and saturation. *JOSA*, 67(2):202–207, 1977.



- 
- [31] Steven J Gortler, Radek Grzeszczuk, Richard Szeliski, Michael F Cohen. The lumigraph. *Siggraph*, wolumen 96, strony 43–54, 1996.
- [32] Brian Guenter, Mark Finch, Steven Drucker, Desney Tan, John Snyder. Foveated 3d graphics. *ACM Transactions on Graphics (TOG)*, 31(6):164, 2012.
- [33] Thorsten Hansen, Lars Pracejus, Karl R Gegenfurtner. Color perception in the intermediate periphery of the visual field. *Journal of Vision*, 9(4):26–26, 2009.
- [34] Felix Heide, Gordon Wetzstein, Ramesh Raskar, Wolfgang Heidrich. Adaptive image synthesis for compressive displays. *ACM Transactions on Graphics (TOG)*, 32:132, 2013.
- [35] John M Henderson, Andrew Hollingworth. Global transsaccadic change blindness during scene perception. *Psychological science*, 14(5):493–497, 2003.
- [36] Alain Hore, Djemel Ziou. Image quality metrics: Psnr vs. ssim. *2010 20th International Conference on Pattern Recognition*, strony 2366–2369. IEEE, 2010.
- [37] Xinda Hu, Hong Hua. High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics. *Optics express*, 22(11):13896–13903, 2014.
- [38] Hong Hua. Enabling focus cues in head-mounted displays. *Proceedings of the IEEE*, 105(5):805–824, 2017.
- [39] Hong Hua, Bahram Javidi. A 3d integral imaging optical see-through head-mounted display. *Optics express*, 22(11):13484–13491, 2014.
- [40] Fu-Chung Huang, David P Luebke, Gordon Wetzstein. The light field stereoscope. *SIGGRAPH Emerging Technologies*, strony 24–1, 2015.
- [41] Emile Javal. Essai sur la physiologie de la lecture. *Annales d’Oculistique*, 80:61–73, 1878.
- [42] Michael Kalloniatis, Charles Luu. *Webvision: The Organization of the Retina and Visual System*. University of Utah Health Sciences Center, webvision.med.utah.edu, 1995.
- [43] Kil Joong Kim, Rafal Mantiuk, Kyoung Ho Lee. Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance. *Human Vision and Electronic Imaging XVIII*, wolumen 8651, strona 86511A. International Society for Optics and Photonics, 2013.
- [44] Robert Konrad, Nitish Padmanaban, Keenan Molner, Emily A Cooper, Gordon Wetzstein. Accommodation-invariant computational near-eye displays. *ACM Transactions on Graphics (TOG)*, 36(4):88, 2017.
- [45] Gregory Kramida. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE transactions on visualization and computer graphics*, 22(7):1912–1931, 2015.
- [46] Jeffrey C Lagarias, James A Reeds, Margaret H Wright, Paul E Wright. Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on optimization*, 9(1):112–147, 1998.
- [47] Victor AF Lamme, Pieter R Roelfsema. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences*, 23(11):571–579, 2000.
- [48] Douglas Lanman, David Luebke. Near-eye light field displays. *ACM Transactions on Graphics (TOG)*, 32(6):220, 2013.
- [49] Douglas Lanman, Gordon Wetzstein, Matthew Hirsch, Wolfgang Heidrich, Ramesh Ra-

- skar. Polarization fields: dynamic light field display using multi-layer lcds. *ACM Transactions on Graphics (TOG)*, wolumen 30, strona 186. ACM, 2011.
- [50] Seungjae Lee, Jaebum Cho, Byounghyo Lee, Youngjin Jo, Changwon Jang, Dongyeon Kim, Byoungho Lee. Foveated retinal optimization for see-through near-eye multi-layer displays. *IEEE Access*, 6:2170–2180, 2017.
- [51] R John Leigh, David S Zee. *The neurology of eye movements*. OUP USA, 2015.
- [52] Anat Levin, Fredo Durand. Linear view synthesis using a dimensionality gap light field prior. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, strony 1831–1838. IEEE, 2010.
- [53] Marc Levoy, Pat Hanrahan. Light field rendering. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, strony 31–42. ACM, 1996.
- [54] Marc Levoy, Ross Whitaker. Gaze-directed volume rendering. *ACM SIGGRAPH Computer Graphics*, wolumen 24, strony 217–223. ACM, 1990.
- [55] Peter Lindstrom, Greg Turk. Fast and memory efficient polygonal simplification. *IEEE Visualization Proceedings*, strony 27–8, 1998.
- [56] Lester Loschky, George McConkie, Jian Yang, Michael Miller. The limits of visual resolution in natural scene viewing. *Visual Cognition*, 12(6):1057–1092, 2005.
- [57] Gordon D Love, David M Hoffman, Philip JW Hands, James Gao, Andrew K Kirby, Martin S Banks. High-speed switchable lens enables the development of a volumetric stereoscopic display. *Optics express*, 17(18):15716–15725, 2009.
- [58] Michael Lyons, Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba. Coding facial expressions with gabor wavelets. *Proceedings Third IEEE international conference on automatic face and gesture recognition*, strony 200–205. IEEE, 1998.
- [59] Kevin J MacKenzie, David M Hoffman, Simon J Watt. Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *Journal of vision*, 10:22–22, 2010.
- [60] Andrew Maimone, Henry Fuchs. Computational augmented reality eyeglasses. *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, strony 29–38. IEEE, 2013.
- [61] Andrew Maimone, Andreas Georgiou, Joel S Kollin. Holographic near-eye displays for virtual and augmented reality. *ACM Transactions on Graphics (TOG)*, 36(4):85, 2017.
- [62] James Mannos, David Sakrison. The effects of a visual fidelity criterion of the encoding of images. *IEEE transactions on Information Theory*, 20(4):525–536, 1974.
- [63] Rafał K Mantiuk, Anna Tomaszewska, Radosław Mantiuk. Comparison of four subjective methods for image quality assessment. *Computer graphics forum*, wolumen 31, strony 2478–2491. Wiley Online Library, 2012.
- [64] Susana Martinez-Conde, Stephen L Macknik, David H Hubel. The role of fixational eye movements in visual perception. *Nature reviews neuroscience*, 5(3):229, 2004.
- [65] Nathan Matsuda, Alexander Fix, Douglas Lanman. Focal surface displays. *ACM Transactions on Graphics (TOG)*, 36(4):86, 2017.
- [66] Olivier Mercier, Yusufu Sulai, Kevin Mackenzie, Marina Zannoli, James Hillis, Derek Now-

- rouzezahrai, Douglas Lanman. Fast gaze-contingent optimal decompositions for multifocal displays. *ACM Transactions on Graphics (TOG)*, 36(6):237, 2017.
- [67] Albert Abraham Michelson. *Studies in Optics*. University of Chicago science series. University of Chicago Press, 1927.
- [68] Seokil Moon, Chang-Kun Lee, Dukho Lee, Changwon Jang, ByoungHo Lee. Layered display with accommodation cue using scattering polarizers. *IEEE Journal of Selected Topics in Signal Processing*, 11(7):1223–1231, 2017.
- [69] May-Britt Moser, Edvard I Moser. Functional differentiation in the hippocampus. *Hippocampus*, 8(6):608–619, 1998.
- [70] Kathy T Mullen. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *The Journal of physiology*, 359(1):381–400, 1985.
- [71] Kathy T Mullen, i in. Differential distributions of red-green and blue-yellow cone opponency across the visual field. *Visual neuroscience*, 19(1):109–118, 2002.
- [72] Kathy T Mullen, Masato Sakurai, William Chu. Does l/m cone opponency disappear in human periphery? *Perception*, 34(8):951–959, 2005.
- [73] KT Mullen. Colour vision as a post-receptoral specialization of the central visual field. *Vision research*, 31(1):119–130, 1991.
- [74] Allen L Nagy, Jeffrey A Doyal. Red-green color discrimination as a function of stimulus field size in peripheral vision. *JOSA A*, 10(6):1147–1156, 1993.
- [75] Allen L Nagy, Steven Wolf. Red-green color discrimination in peripheral vision. *Vision research*, 33(2):235–242, 1993.
- [76] Rahul Narain, Rachel A Albert, Abdullah Bulbul, Gregory J Ward, Martin S Banks, James F O’Brien. Optimal presentation of imagery with focus cues on multi-plane displays. *ACM Transactions on Graphics (TOG)*, 34(4):59, 2015.
- [77] Cornelis Noorlander, Jan J Koenderink, Ron J Den Olden, B Wigbold Edens. Sensitivity to spatiotemporal colour contrast in the peripheral visual field. *Vision Research*, 23(1):1–11, 1983.
- [78] David Noton, Lawrence Stark. Eye movements and visual perception. *Scientific American*, 224(6):34–43, 1971.
- [79] Kristien Ooms, Lien Dupont, Lieselot Lapon, Stanislav Popelka. Accuracy and precision of fixation locations recorded with the low-cost eye tribe tracker in different experimental setups. *Journal of eye movement research*, 8(1), 2015.
- [80] SM Palmer, MGP Rosa. A distinct anatomical network of cortical areas for analysis of motion in far peripheral vision. *European Journal of Neuroscience*, 24(8):2389–2405, 2006.
- [81] Peeta Basa Pati, AG Ramakrishnan. Word level multi-script identification. *Pattern Recognition Letters*, 29(9):1218–1229, 2008.
- [82] Anjul Patney, Joohwan Kim, Marco Salvi, Anton Kaplanyan, Chris Wyman, Nir Benty, Aaron Lefohn, David Luebke. Perceptually-based foveated virtual reality. *ACM SIGGRAPH 2016 Emerging Technologies*, strona 17. ACM, 2016.
- [83] Eli Peli. Contrast in complex images. *JOSA A*, 7(10):2032–2040, 1990.

- 
- [84] Eli Peli, Jian Yang, Robert B Goldstein. Image invariance with changes in size: The role of peripheral contrast thresholds. *JOSA A*, 8(11):1762–1774, 1991.
- [85] Denis G Pelli, Peter Bex. Measuring contrast sensitivity. *Vision research*, 90:10–14, 2013.
- [86] JS Pointer, RF Hess. The contrast sensitivity gradient across the human visual field: With emphasis on the low spatial frequency range. *Vision research*, 29(9):1133–1151, 1989.
- [87] Charles Poynton. *Digital video and HD: Algorithms and Interfaces*. Elsevier, 2012.
- [88] S Sabari Raju, Peeta Basa Pati, AG Ramakrishnan. Gabor filter based block energy analysis for text extraction from digital document images. *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, strony 233–243. IEEE, 2004.
- [89] JG Robson, Norma Graham. Probability summation and regional variation in contrast sensitivity across the visual field. *Vision research*, 21(3):409–418, 1981.
- [90] Daniel L Schacter, Daniel T Gilbert, Daniel M Wegner. Psychology, 2011.
- [91] Adam Siekawa. Image reconstruction from spatially non-uniform samples. *Proceedings of CESC 2017: The 21th Central European Seminar on Computer Graphics (non-peer-reviewed)*, 2017.
- [92] Adam Siekawa, Supervised Radoslaw Mantiuk. Gaze-dependent ray tracing. *Proceedings of Central European Seminar on Computer Graphics (non-peer-reviewed)*. Citeseer, 2014.
- [93] Miriam Spering, Dirk Kerzel, Doris I Braun, Michael J Hawken, Karl R Gegenfurtner. Effects of contrast on smooth pursuit eye movements. *Journal of Vision*, 5(5):6–6, 2005.
- [94] Michael Stengel, Marcus Magnor. Gaze-contingent computational displays: Boosting perceptual fidelity. *IEEE Signal Processing Magazine*, 33:139–148, 2016.
- [95] Hans Strasburger, Ingo Rentschler, Martin Jüttner. Peripheral vision and pattern recognition: A review. *Journal of vision*, 11(5):13–13, 2011.
- [96] Martin Szinte, Patrick Cavanagh. Apparent motion from outside the visual field, retinotopic cortices may register extra-retinal positions. *PloS one*, 7(10):e47386, 2012.
- [97] James P Thomas. Effect of eccentricity on the relationship between detection and identification. *JOSA A*, 4(8):1599–1605, 1987.
- [98] Greg Turk, Marc Levoy. Zippered polygon meshes from range images. *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '94*, strony 311–318, New York, NY, USA, 1994. ACM.
- [99] Karthik Vaidyanathan, Marco Salvi, Robert Toth, Tim Foley, Tomas Akenine-Möller, Jim Nilsson, Jacob Munkberg, Jon Hasselgren, Masamichi Sugihara, Petrik Clarberg, i in. Coarse pixel shading. *Proceedings of High Performance Graphics*, strony 9–18. Eurographics Association, 2014.
- [100] Matteo Valsecchi, Karl R Gegenfurtner, Alexander C Schütz. Saccadic and smooth-pursuit eye movements during reading of drifting texts. *Journal of vision*, 13(10):8–8, 2013.
- [101] John E Vanston, Michael A Crognale. Effects of eccentricity on color contrast. *JOSA A*, 35(4):B122–B129, 2018.
- [102] Andrew B Watson. Visual detection of spatial contrast patterns: Evaluation of five simple models. *Optics Express*, 6(1):12–33, 2000.

- 
- [103] Andrew B Watson, Denis G Pelli. Quest: A bayesian adaptive psychometric method. *Perception & psychophysics*, 33:113–120, 1983.
- [104] Martin Weier, Michael Stengel, Thorsten Roth, Piotr Didyk, Elmar Eisemann, Martin Eisemann, Steve Grogorick, André Hinkenjann, Ernst Kruijff, Marcus Magnor, i in. Perception-driven accelerated rendering. *Computer Graphics Forum*, wolumen 36, strony 611–643. Wiley Online Library, 2017.
- [105] Gordon Wetzstein, Douglas Lanman, Wolfgang Heidrich, Ramesh Raskar. Layered 3d: tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Transactions on Graphics (ToG)*, wolumen 30, strona 95. ACM, 2011.
- [106] Paul Whittle, PDC Challands. The effect of background luminance on the brightness of flashes. *Vision research*, 9(9):1095–1110, 1969.
- [107] Samuel J Williamson, Herman Z Cummins. Light and color in nature and art. *Light and Color in Nature and Art, by Samuel J. Williamson, Herman Z. Cummins, pp. 512. ISBN 0-471-08374-7. Wiley-VCH, February 1983.*, strona 512, 1983.
- [108] Krzysztof Wolski, Radosław Mantiuk. Cross spread pupil tracking technique. *Journal of Electronic Imaging*, 25(6):063012, 2016.
- [109] Sophie M Wuerger, Andrew B Watson, Albert J Ahumada Jr. Towards a spatio-chromatic standard observer for detection. *Human Vision and Electronic Imaging VII*, wolumen 4662, strony 159–172. International Society for Optics and Photonics, 2002.
- [110] Han-Ju Yeom, Hee-Jae Kim, Seong-Bok Kim, HuiJun Zhang, BoNi Li, Yeong-Min Ji, Sang-Hoo Kim, Jae-Hyeung Park. 3d holographic head mounted display using holographic optical elements with astigmatism aberration compensation. *Optics express*, 23(25):32025–32034, 2015.