Politecnico di Torino

Ingegneria del Cinema e dei Mezzi di Comunicazione

A.a. 2025/2026

Sessione di laurea Luglio 2026

Cinematografia Computazionale e Aberrazioni del Terzo Ordine: Mappatura Volumetrica del Frustum in Mitsuba 3.0 attraverso i Polinomi di Zernike.

Un approccio ibrido per il rendering cinematico di coma, astigmatismo ed aberrazione sferica.

Relatori:

Bartolomeo Montrucchio

Francesco Strada

Leonardo Vezzani

Candidati:

Gabriele Piergiovanni

"It's the kind of thing that makes you glad
you stopped and smelled the pine trees along the way,
you know?"

A Nonno Antonio.

Sommario

Questo elaborato illustra lo sviluppo, l'implementazione e la validazione di un sensore di rendering personalizzato all'interno del framework orientato alla ricerca Mitsuba 3. L'architettura proposta si basa su una mappatura volumetrica del cono di luce, strutturata attraverso l'interpolazione lineare di piani focali. Tali piani rappresentano le risposte all'impulso del sistema ottico, modellate analiticamente come combinazione di polinomi della serie di Zernike. Questo approccio è stato concepito per emulare, in modo fisicamente coerente, l'evoluzione tridimensionale delle aberrazioni ottiche lungo l'intero asse di profondità della scena. La validazione del sistema si è articolata in due fasi distinte. L'analisi quantitativa è stata condotta valutando la risoluzione ottica e il contrasto tramite il rendering di mire standardizzate ISO 12233 e lo studio della loro risposta al gradino per il calcolo della Modulation Transfer Function (MTF). Parallelamente, è stato misurato il differenziale (delta) della MTF su un dataset di scene fotorealistiche in computer grafica. Sul versante qualitativo, l'impatto percettivo del modello è stato indagato mediante la somministrazione di un questionario strutturato. Il test ha coinvolto gli utenti in un confronto visivo diretto tra le immagini generate dal tradizionale modello a lente sottile (thin-lens) nativo di Mitsuba e i rendering prodotti dal sensore proposto, con l'obiettivo di misurarne la validità estetica e il grado di realismo.

Ringraziamenti

Volevo innanzitutto esprimere la mia gratitudine al Professor Bartolomeo Montrucchio, relatore della tesi e al Professor Francesco Strada che si sono mostrati sempre disponibili sin dall'inizio di questi mesi di lavoro. Un ringraziamento sentito va inoltre a Leonardo Vezzani che mi ha seguito e aiutato nella fase più importante del mio percorso accademico.

Ringrazio i miei genitori Monica e Massimo, senza cui nulla di tutto questo sarebbe stato possibile, mio fratello Jacopo per avermi costantemente messo in dubbio permettendomi di essere una persona migliore una discussione a cena alla volta, grazie anche ai miei gatti perché chissà quanto sarebbe stato alto il mio cortisolo senza di loro.

Grazie agli amici che ci sono da sempre, come Ivan senza cui metà dei progetti universitari a cui ho preso parte non avrebbero avuto una voce e Mirko che, in veste di Dungeon Master, ha orchestrato una delle poche evasioni dalla realtà capaci di distrarmi in questi mesi di scrittura. Alice, Edoardo, Matteo, Miriam e Daniele, potrei *de facto* scrivere infinite cose su di voi, ma siccome Dani non vuole spoiler, mi limterò a ringraziarvi per esserci sempre stati per me. Ringrazio inoltre chi ha avuto l'infinita pazienza di sopportarmi in questi anni universitari: sia a Torino — Irene, Gaia, Paolo, Laura, Alberto, Ilaria e Michele — che fuori porta, come Giulia la mejor coinquilina de todas.

Ringrazio poi chiunque abbia condiviso la passione per il cinema e la sala con me, in particolare Giovanni, Francesco e Davide per avermi fatto compagnia in qualsivoglia proiezione, dal più putrido dei blockbuster fino al più supponente dei film d'autore. Impossibile poi non ringraziare i membri del team 2Hardware, persone estremamente talentuose e dotate di un'infinita pazienza, dote fondamentale quando si ha a che fare con persone disorganizzate come il sottoscritto. Se non siete stati menzionati espressamente in queste pagine, è solo perché, tenendo alla vita degli alberi, ho dovuto mettere un freno alla mia naturale tendenza a considerare amiche anche le pietre. In ogni caso: grazie, vi voglio bene.

Qualcuno diceva che siamo un mosaico di tutti quelli a cui abbiamo voluto bene, il mio è composto da una miriade di tasselli incredibili e spero che anche in quelli di chi ho conosciuto ce ne sia uno, da qualche parte, leggermente sghembo e decorato da un pattern di dubbio gusto.

Indice

Elenco delle figure VII

1 Introduzione 1

2 Stato dell'arte 4
2.1 Teoria ottica . . . 4
2.1.1 Ottica fisica e geometrica . . . 4
2.1.2 Il sistema ottico . . . 5
2.1.3 Sistemi LSI e formazione dell'immagine . . . 6
2.1.4 Point Spread Function . . . 7
2.1.5 Circolo di Confusione e Profondità di Campo . . . 8
2.1.6 Bokeh . . . 9
2.1.7 Aberrazioni ottiche . . . 10
2.1.8 Polinomi di Zernike . . . 16
2.2 Ray-Tracing . . . 19
2.2.1 Fondamenti di Ray-Tracing . . . 19
2.2.2 Rappresentazione del Bokeh in sistemi di Ray-Tracing . . . 21
2.2.3 Aberrazioni nel Ray-Tracing . . . 22
2.3 Obiettivi della Ricerca (Research Questions) . . . 23

3 Metodologia 25
3.1 Mitsuba 3.0 . . . 25
3.1.1 Integratori . . . 25
3.1.2 Campionatori . . . 26
3.1.3 Sensori . . . 26
3.2 Sviluppo Architetturale . . . 27
3.2.1 Disaccoppiamento Sampler-Sensore . . . 27
3.2.2 Mappa di profondità . . . 28
3.2.3 Approssimazione Discreta in Dominio di Fourier . . . 28
3.2.4 Modello Vettoriale Ibrido . . . 29
3.3 Architettura del Sistema . . . 29

3.4 Fase 1: Modellazione Ottica . . . 31
3.4.1 Il Modello Ottico Diretto . . . 31
3.4.2 Recupero di Fase . . . 35
3.4.3 Sintesi del Frustum Aberrato . . . 36
3.5 Fase 2: Rendering Ibrido . . . 37
3.6 Pipeline di rendering . . . 38
3.6.1 Lettura dell'XML . . . 38
3.6.2 Il Modulo Main . . . 39
3.6.3 Post-elaborazione . . . 40

4 Validazione Sperimentale 41
4.1 Test Analitico . . . 41
4.1.1 Test analitico: combinazioni di aberrazioni . . . 43
4.1.2 Test Analitico: MTF Globale su Scene Fotorealistiche . . . 47
4.2 Test Percettivo . . . 59

5 Risultati e discussione 60
5.1 Analisi MTF . . . 60
5.1.1 Analisi della Risposta Spaziale (Mira ISO) . . . 60
5.1.2 Stima della MTF su Scene Complesse . . . 65
5.2 Analisi dell'Esperienza Utente (Questionari) . . . 76
5.2.1 Risultati Statistici . . . 77
5.3 Discussione e Correlazione . . . 98
5.3.1 Analisi dei commenti . . . 98
5.3.2 Coerenza tra Metriche Oggettive e Soggettive . . . 99
5.3.3 Validazione dell'ipotesi di ricerca . . . 100

6 Conclusioni 101
6.1 Sintesi del lavoro svolto . . . 101
6.2 Risposte ai quesiti di ricerca . . . 101
6.3 Risultati della validazione . . . 102
6.4 Limiti dell'implementazione . . . 102
6.5 Sviluppi futuri . . . 103

Bibliografia 104

Elenco delle figure

2.1 Modello di proiezione prospettica di un sistema ottico. 6
2.2 Diagramma della risposta all'impulso di un sistema LSI. 7
2.3 Pattern di diffrazione generato da un'apertura circolare. 8
2.4 Effetto *bokeh* con soggetto in primo piano. 10
2.5 Lente ideale e aberrazione sferica a confronto. 12
2.6 Modello prospettico di una lente affetta da coma. 13
2.7 Modello prospettico di una lente affetta da astigmatismo. 14
2.8 Primi 21 polinomi di Zernike 16
2.9 Modello di formazione dell'immagine tramite ray-tracing. 19

3.1 Diagramma di flusso dell'architettura proposta. 30

4.1 Fotografia di una mira ottica ISO 12233 [25]. 42
4.2 Rendering ISO 12233: assenza di aberrazioni 43
4.3 Rendering ISO 12233: astigmatismo 44
4.4 Rendering ISO 12233: coma 44
4.5 Rendering ISO 12233: aberrazione sferica positiva 45
4.6 Rendering ISO 12233: aberrazione sferica negativa 45
4.7 Rendering ISO 12233: aberrazione sferica positiva (focus -1) 46
4.8 Rendering ISO 12233: aberrazione sferica negativa (focus -1) 46
4.9 Evoluzione della PSF al variare del defocus Z_4 48
4.10 Confronto visivo: Scena 01 - Cornell box. 50
4.11 Confronto visivo: Scena 02 - Cornell box glass. 51
4.12 Confronto visivo: Scena 03 - Cornell box water. 52
4.13 Confronto visivo: Scena 04 - Lego. 53
4.14 Confronto visivo: Scena 05 - Detective room. 54
4.15 Confronto visivo: Scena 06 - Perfume. 55
4.16 Confronto visivo: Scena 07 - Winter forest. 56
4.17 Confronto visivo: Scena 08 - Wooden pier. 57
4.18 Confronto visivo: Scena 09 - Car. 58

5.1 MTF ISO 12233 - senza aberrazioni. 61
5.2 MTF ISO 12233 - astigmatismo. 62
5.3 MTF ISO 12233 - coma. 62
5.4 MTF ISO 12233 - aberrazione sferica positiva. 63
5.5 MTF ISO 12233 - aberrazione sferica negativa. 63
5.6 MTF ISO 12233 - aberrazione sferica positiva (focus a -1). 64
5.7 MTF ISO 12233 - aberrazione sferica negativa (focus a -1). 64
5.8 Stima della MTF e delta - Cornell Box - file .exr 66
5.9 Stima della MTF e delta - Cornell Box - denoised .png 66
5.10 Stima della MTF e delta - Cornell Box Glass - file .exr 67
5.11 Stima della MTF e delta - Cornell Box Glass - denoised .png 67
5.12 Stima della MTF e delta - Cornell Box Water - file .exr 68
5.13 Stima della MTF e delta - Cornell Box Water - denoised .png . . . 68
5.14 Stima della MTF e delta - Lego - file .exr 69
5.15 Stima della MTF e delta - Lego - denoised .png 69
5.16 Stima della MTF e delta - Detective Room - file .exr 70
5.17 Stima della MTF e delta - Detective Room - denoised .png 70
5.18 Stima della MTF e delta - Perfume - file .exr 71
5.19 Stima della MTF e delta - Perfume - denoised .png 71
5.20 Stima della MTF e delta - Winter forest - file .exr 72
5.21 Stima della MTF e delta - Winter forest - denoised .png 72
5.22 Stima della MTF e delta - Wooden pier - file .exr 73
5.23 Stima della MTF e delta - Wooden pier - denoised .png 73
5.24 Stima della MTF e delta - Car - file .exr 74
5.25 Stima della MTF e delta - Car - denoised .png 74
5.26 Coppia *Cornell Box* — Likert competenza | età. 77
5.27 Coppia *Cornell Box* — preferenze A/B. 78
5.28 Coppia *Cornell Box Glass*— Likert competenza | età. 79
5.29 Coppia *Cornell Box Glass* — preferenze A/B. 80
5.30 Coppia *Cornell Box Water* — Likert competenza | età. 81
5.31 Coppia *Cornell Box Water* — preferenze A/B. 82
5.32 Coppia *Lego* — Likert competenza | età. 83
5.33 Coppia *Lego* — preferenze A/B. 84
5.34 Coppia *Detective room* — Likert competenza | età. 85
5.35 Coppia *Detective room* — preferenze A/B. 86
5.36 Coppia *Perfume* — Likert competenza | età. 87
5.37 Coppia *Perfume* — preferenze A/B. 88
5.38 Coppia *Winter forest* — Likert competenza | età. 89
5.39 Coppia *Winter forest* — preferenze A/B. 90
5.40 Coppia *Wooden pier* — Likert competenza | età. 91
5.41 Coppia *Wooden pier* — preferenze A/B. 92

5.42 Coppia *Car* — Likert competenza | età. 93
5.43 Coppia *Car* — preferenze A/B. 94
5.44 Media aggregata per competenza ed età. 95
5.45 Media aggregata — preferenze A/B. 96
5.46 Percezione complessiva dell'intero dataset. 97

Capitolo 1

Introduzione

Nel corso degli ultimi decenni le tecnologie di rendering fotorealistico basate sugli algoritmi di ray-tracing sono diventate sempre più raffinate in termini di fedeltà visiva, assumendo un ruolo cruciale in settori quali la produzione cinematografica e i videogiochi. In questi ambiti, l'obiettivo primario di questo tipo di sintesi d'immagine è quello di spingersi oltre la pura simulazione geometrica per emulare la resa organica e autentica della fotografia tradizionale. Nonostante ciò, le implementazioni standard della maggior parte di questi motori di rendering si basano sull'approssimazione matematica di sistemi ottici ideali, escludendo di fatto le sottili imperfezioni fisiche intrinseche a qualsiasi lente reale, restituendo allo spettatore immagini caratterizzate da una nitidezza clinica e innaturale.

Tali irregolarità derivano dalle proprietà geometriche delle lenti e prendono il nome di aberrazioni ottiche. Sebbene i moderni processi di produzione industriale mirino a minimizzare queste imperfezioni per approssimare il comportamento di una lente ideale, gli obiettivi fisici conservano quasi sempre delle distorsioni residue, particolarmente evidenti in dispositivi datati o in fotocamere di fascia economica. Allo stesso tempo, specifici difetti, come l'aberrazione sferica, responsabile di un caratteristico alone luminoso che avvolge l'intera immagine, possono essere ricercati intenzionalmente per ottenere peculiari effetti artistici. L'impatto visivo di queste anomalie risulta maggiormente marcato nelle aree fuori fuoco della scena, dove si manifesta il cosiddetto *bokeh*, ovvero la particolare resa estetica che le sorgenti luminose sfocate assumono in tali regioni.

In letteratura esistono vari approcci volti a riprodurre tali effetti attraverso la simulazione fisica di interi sistemi di lenti all'interno del motore di rendering. Tuttavia, l'applicabilità di questi metodi è fortemente limitata: i dati necessari per una modellazione accurata sono infatti reperibili solo per una ristretta cerchia di obiettivi, essendo i parametri costruttivi di frequente protetti dal segreto industriale

dei produttori.

L'obiettivo di questo lavoro è pertanto quello di implementare un sensore personalizzato all'interno del motore di rendering Mitsuba 3 [1], scelto in virtù della sua intrinseca flessibilità nella creazione di componenti architetturali. Invece di basare la simulazione sui parametri costruttivi interni della lente, il modello adotterà un approccio inverso basato sull'analisi della risposta all'impulso del sistema che definiamo come Point Spread Function (PSF). Sarà proprio la derivazione analitica di tale metrica, calcolata matematicamente a partire dalle deformazioni del fronte d'onda, a svincolarci dalle specifiche meccaniche del vetro, fornendo al sensore i dati necessari per simulare le aberrazioni ottiche in modo fisicamente accurato.

Il modello proposto non si limita a simulare l'impronta bidimensionale dell'aberrazione a una singola distanza focale. Al contrario, la pipeline prevede innanzitutto la generazione di molteplici PSF a diverse distanze di defocus, che vengono successivamente organizzate all'interno di uno stack volumetrico. Durante la fase di tracciamento, la pila di dati pre-calcolati viene interpolata linearmente per generare la geometria del frustum aberrato. Questo approccio permette di ricostruire in modo continuo l'intera evoluzione tridimensionale del difetto ottico al variare della messa a fuoco, definendo le traiettorie corrette per i raggi primari. La fedeltà visiva di tale simulazione poggia sulla descrizione del fronte d'onda tramite i polinomi di Zernike. Tali funzioni costituiscono infatti il fondamento analitico su cui vengono modellate le PSF sintetiche: l'architettura è progettata e implementata per elaborare direttamente i coefficienti associati alle aberrazioni maggiormente riscontrabili negli obiettivi fotografici e cinematografici reali, restituendo una simulazione scalabile e nativamente continua.

Il presente elaborato è strutturato come segue:

- **Capitolo 2 - Stato dell'arte:** introduce i fondamenti teorici dell'ottica geometrica e del ray-tracing, analizzando la classificazione delle aberrazioni ottiche e delineando i limiti degli approcci attualmente esistenti in letteratura per la simulazione fotorealistica del *bokeh*.
- **Capitolo 3 - Metodologia:** dettaglia l'architettura logica e matematica della soluzione proposta, dalla generazione dello stack di PSFs tramite i coefficienti di Zernike, fino all'implementazione del frustum aberrato e all'integrazione del sensore personalizzato all'interno del motore di rendering Mitsuba 3.
- **Capitolo 4 - Setup sperimentale e validazione:** descrive l'ambiente di test, la metodologia di campionamento e gli scenari di rendering adottati per valutare il modello. L'analisi si articola su due fronti, illustrando da una parte i criteri analitici utilizzati per verificare la robustezza e la correttezza fisica

del sistema e dall'altra l'impianto metodologico dei questionari soggettivi, impiegati per indagarne l'impatto percettivo sull'osservatore.

- **Capitolo 5 - Risultati e discussione:** presenta e dettaglia criticamente le metriche oggettive ottenute a partire dall'analisi degli output visivi, correlando l'andamento analitico delle curve definito dalla Modulation Transfer Function (MTF) con i riscontri percettivi emersi dai questionari. Il capitolo analizza i dati quantitativi derivati dalle scale Likert e dalle valutazioni comparative a coppie, integrandoli con l'analisi qualitativa dei commenti liberi. Questo approccio permette di evidenziare la coerenza tra le metriche matematiche e il responso visivo umano, giungendo così alla validazione finale dell'ipotesi di ricerca.
- **Capitolo 6 - Conclusioni:** sintetizza i traguardi raggiunti dal progetto, delineandone le limitazioni attuali e tracciando le possibili traiettorie per i futuri sviluppi di ricerca.

Capitolo 2

Stato dell'arte

Prima di procedere con l'analisi nel dettaglio del modello proposto, è però doveroso introdurre i fondamenti teorici su cui si basa la nostra logica di sviluppo. Il presente capitolo si articola a partire dalla descrizione dei principi di ottica necessari a comprendere il comportamento di una lente reale, passando poi all'analisi dell'architettura dei modelli di ray-tracing, per terminare con i quesiti relativi alla nostra ricerca.

2.1 Teoria ottica

Questa sezione è dedicata all'approfondimento dei modelli fisici che regolano la propagazione della luce e la formazione dell'immagine. Il percorso analitico ha inizio con la formalizzazione teorica di un sistema ottico ideale per giungere alla descrizione formale della Point Spread Function (PSF). Successivamente, l'attenzione si sposta sulla resa visiva della sfocatura, introducendo concetti chiave quali il Circle of Confusion (CoC) e il *bokeh*. Il discorso si allarga quindi all'analisi delle aberrazioni ottiche del terzo ordine, per concludersi infine con l'introduzione dei polinomi di Zernike e con l'analisi del loro ruolo nella rappresentazione matematica del fronte d'onda aberrato.

2.1.1 Ottica fisica e geometrica

La prima distinzione fondamentale da fare è quella tra ottica fisica e ottica geometrica: la prima osserva il comportamento della luce trattandola come un'onda elettromagnetica, mentre la seconda la modella come un fascio di raggi. Quest'ultima rappresenta un caso limite del modello fisico, fornendo un'approssimazione accurata nelle situazioni in cui le dimensioni dei corpi con cui la luce interagisce sono decisamente maggiori della lunghezza d'onda della luce stessa. Nel caso di lenti

fotografiche o cinematografiche è lecito adottare l'approssimazione geometrica; al contrario, in ambiti come la microscopia, è indispensabile ricorrere all'ottica fisica. Ai fini di questo lavoro, l'utilizzo di un motore di rendering basato sul ray-tracing rende vincolante e metodologicamente rigoroso il ricorso esclusivo all'astrazione geometrica [2, cap. 5].

2.1.2 Il sistema ottico

Nella sua forma più essenziale, un dispositivo fotografico reale può essere semplificato, considerandolo come composto solamente da due elementi principali, una lente (o un sistema di lenti) e un sensore.

Concentrando l'analisi sulla lente, possiamo identificarla come l'unione di due diottri rifrangenti di cui almeno uno curvo, realizzati con materiali trasparenti quali vetro o plastica. Quando almeno uno dei due lati del disco di vetro è convesso, il fascio di raggi che lo attraversa parallelamente al suo asse verrà focalizzato, ergo verrà fatto convergere, in un punto ad una determinata distanza rispetto alla lente, detto fuoco. Il raggio di luce devia rispetto alla sua originale traiettoria perché sta attraversando due mezzi trasparenti con diverso indice di rifrazione; ogni mezzo trasparente come l'aria, l'acqua o il vetro ne ha uno differente. Si tratta di una misura atta a stabilire quanto diminuisce la velocità della propagazione della luce in un mezzo diverso dal vuoto, ed è definita dalla seguente relazione:

$$n = \frac{c}{v} \tag{2.1}$$

Dove n è l'indice di rifrazione, c è la velocità della luce nel vuoto e v è la velocità della luce mentre sta attraversando il mezzo preso in esame. La deviazione dell'angolo con cui viaggia il raggio di luce è regolata dalla legge di Snell-Descartes, la quale identifica il rapporto fondamentale [2, cap. 4]:

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \tag{2.2}$$

dove n_1 e n_2 rappresentano gli indici di rifrazione dei due mezzi attraversati, mentre θ_1 e θ_2 indicano rispettivamente l'angolo di incidenza e l'angolo di rifrazione, entrambi misurati rispetto alla normale alla superficie di separazione tra i due mezzi. Oggi, la maggior parte delle fotocamere ha sistemi ottici costituiti da insiemi di lenti, che definiamo elementi, orientate lungo un asse comune, spesso divise in gruppi da due o tre, alternate da diaframmi, sezioni circolari opache adibite a limitare il diametro del fascio di luce. La presenza di più lenti con curvature diverse della superficie è inoltre necessaria per la correzione delle aberrazioni ottiche, su cui entreremo in seguito nel dettaglio. Tuttavia, nello studio dell'ottica geometrica di base, la notevole complessità strutturale di un simile apparato reale viene

comunemente astratta e ricondotta a un equivalente teorico basilare: il modello a lente sottile (o di proiezione prospettica), come quello illustrato in Figura 2.1.

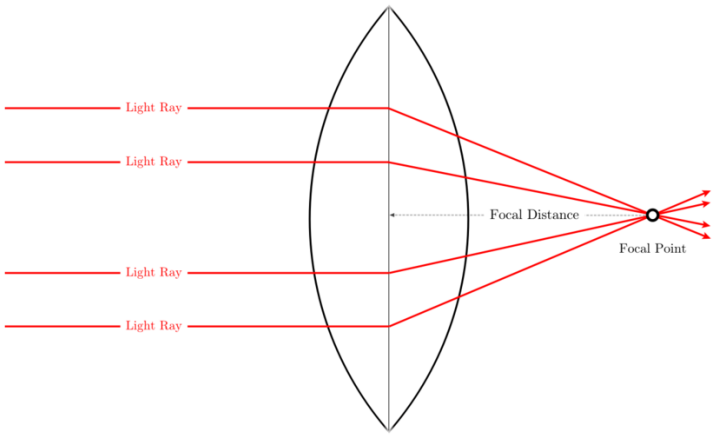


Figura 2.1: Modello geometrico di una lente convergente. Il diagramma illustra la rifrazione dei raggi parassiali che attraversano il disco di vetro per convergere sul piano focale. La distanza tra il centro della lente e il punto di convergenza definisce la lunghezza focale f [3].

2.1.3 Sistemi LSI e formazione dell'immagine

L'apparato ottico appena descritto è di fatto l'elemento fondamentale che permette di convertire una scena reale in un'immagine. Dal punto di vista teorico, esso prende il nome di sistema di imaging e può essere descritto come un operatore Lineare Spazio-Invariante (LSI) che acquisisce il campo luminoso in ingresso ed elabora i valori di intensità e colore per ciascuna coordinata spaziale, restituendo il risultato finale. Ernst Abbe fu il primo a fornire una descrizione formale di questo processo, definendolo come la risposta del sistema a onde luminose non coerenti provenienti dalla scena catturata da esso. Questo dettaglio, ovvero l'assenza di coerenza di fase tipica delle sorgenti naturali e delle sorgenti artificiali a incandescenza, è fondamentale per il nostro contesto: nella fotografia tradizionale e nella cinematografia la luce è spazialmente incoerente. Ciò implica che il sistema ottico somma linearmente le singole intensità della luce e non le ampiezze d'onda complesse. Di conseguenza, l'immagine finale può essere matematicamente descritta

come la convoluzione tra l'intensità luminosa della scena proiettata sul piano focale e la risposta all'impulso del sistema stesso [4].

Nonostante il modello di simulazione al centro di questo lavoro operi nel dominio dell'ottica geometrica tramite il tracciamento dei raggi, richiamare questo approccio analitico è un passaggio necessario. Infatti, è proprio modellando la risposta del sistema LSI a un singolo impulso luminoso che possiamo definire l'impronta e la forma della sfocatura della nostra lente virtuale, concetto cardine su cui si basa l'intera logica del nostro sensore.

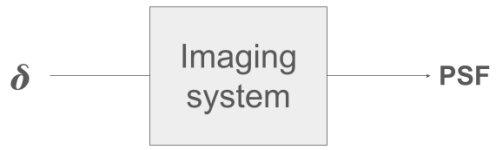


Figura 2.2: Diagramma della risposta all'impulso di un sistema LSI.

2.1.4 Point Spread Function

Come anticipato dallo schema in Figura 2.2, per tradurre analiticamente la formazione dell'immagine è necessario definire la Point Spread Function (PSF), ovvero la risposta del sistema ottico a una singola sorgente di luce puntiforme. Riprendendo la modellazione del sistema di imaging come operatore LSI, la PSF ne costituisce l'equivalente matematico della risposta all'impulso nel dominio spaziale. Nel mondo reale, la PSF può essere osservata fotografando una stella (che possiamo approssimare come fonte di luce puntiforme di intensità elevata): il risultato è tipicamente un pattern di diffrazione circolare che fornisce informazioni cruciali sulle caratteristiche e sui difetti del sistema stesso.

Anche in assenza di imperfezioni fisiche del vetro, una lente circolare non farà mai convergere la luce in un punto infinitamente piccolo a causa del limite fisico imposto dalla diffrazione. La regione centrale luminosa di questa PSF ideale prende il nome di disco di Airy. Assumendo di proiettare l'immagine sul piano focale, il

raggio di tale disco (misurato dal centro geometrico fino al primo anello di buio) è pari a:

$$r = 1.22\frac{\lambda f}{D_0} \tag{2.3}$$

dove λ è la lunghezza d'onda della luce, D_0 è il diametro dell'apertura della lente e f è la lunghezza focale. All'interno di questo perimetro si concentra circa l'84% dell'energia luminosa totale [2, cap. 10].

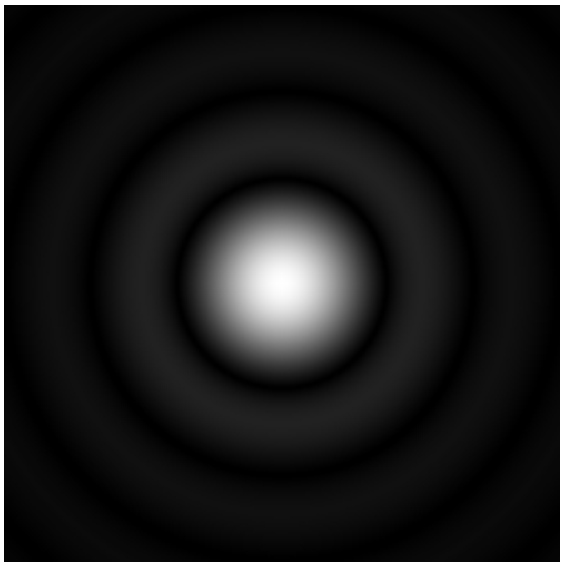


Figura 2.3: Pattern di diffrazione generato da un'apertura circolare (disco di Airy).

2.1.5 Circolo di Confusione e Profondità di Campo

Se la PSF descrive il limite fisico della messa a fuoco, il Circolo di Confusione CoC ne descrive l'espansione geometrica quando ci si allontana dal piano focale. Nella realtà, i raggi di luce provenienti da un punto non a fuoco si proiettano sul sensore formando uno spot luminoso più ampio della PSF di base.

La dimensione minima del CoC, raggiunta esattamente sul piano di fuoco, prende il nome di Circolo di minima Confusione (Circle of least Confusion) ed è il parametro cardine per definire la profondità di campo, ovvero l'area tridimensionale all'interno della quale un oggetto viene percepito visivamente come nitido dall'osservatore umano. Tale profondità può essere calcolata tramite la seguente formula approssimativa, che si avvale proprio del diametro del CoC:

$$T \approx \frac{2u^2NC}{f^2} \tag{2.4}$$

dove T è la profondità di campo, u è la distanza dell'oggetto a fuoco, N è il numero f-stop, C è il diametro del circolo di confusione e f è la lunghezza focale della lente [5, cap. 22].

2.1.6 Bokeh

Avendo approfondito i concetti di Circle of Confusion e di Point Spread Function, risulta doveroso introdurre il termine *bokeh*, elemento centrale per la comprensione della logica di rendering proposta. Di origine nipponica, la parola si riferisce alla geometria e alla distribuzione di intensità che una sorgente luminosa puntiforme assume quando si trova fuori fuoco; tale fenomeno è intrinsecamente legato alle caratteristiche della PSF e del circolo di confusione.

Il *bokeh* si manifesta tipicamente attraverso l'impiego di profondità di campo ridotte e permette di attenuare la complessità percepita delle aree sfocate in modo esteticamente gradevole. Per questo motivo, esso rappresenta una risorsa compositiva preziosa in ambito fotografico e cinematografico, utile a isolare visivamente il soggetto a fuoco e a guidare l'attenzione dello spettatore [6].

In condizioni ideali, una lente perfetta dotata di apertura circolare produce un *bokeh* anch'esso circolare e caratterizzato da un'intensità uniforme. Esistono tuttavia notevoli deviazioni da questo comportamento: l'impiego di ottiche specifiche, quali le lenti anamorfiche, genera una proiezione ellittica a causa dell'asimmetria costruttiva del sistema, mentre la presenza di aberrazioni o imperfezioni fisiche nel vetro può alterarne significativamente sia i contorni che la distribuzione luminosa rispetto al modello teorico standard.

Figura 2.4: Dimostrazione dell'effetto *bokeh* sullo sfondo con soggetto principale posizionato in primo piano. Si noti la forma e la morbidezza dei punti di luce sfocati. (Scatto dell'autore. Obiettivo 135mm, f/5.6, 1/20s, ISO 100).

2.1.7 Aberrazioni ottiche

Le lenti, per via della loro forma e a causa di eventuali difetti sulla loro superficie, possono dare vita a immagini che soffrono delle cosiddette aberrazioni ottiche. Tali anomalie si manifestano come sfocature o distorsioni percepibili all'interno dell'immagine e sono causate dalla deviazione dei raggi luminosi rispetto alla loro traiettoria geometrica ideale [2, cap. 6].

Lenti a diffrazione limitata

L'obiettivo dei produttori di lenti è quello di progettare i cosiddetti sistemi a diffrazione limitata (*diffraction-limited*), ovvero ottiche in cui le aberrazioni sono minimizzate al punto da rendere la diffrazione l'unico limite fisico alla risoluzione dell'immagine.

Nel contesto dell'ottica geometrica, tali lenti mirano ad approssimare il più possibile il comportamento di un sistema ottico ideale, il quale è definito nel momento in cui le tre condizioni di Maxwell vengono soddisfatte [5, cap. 11]:

- Tutti i raggi provenienti da un singolo punto dell'oggetto convergono in un unico punto nel piano immagine.

- Un piano oggetto ortogonale all'asse ottico viene riprodotto su un piano immagine anch'esso ortogonale.

- La forma dell'oggetto e dell'immagine che lo riproduce devono essere geometricamente identiche, a meno di un fattore di ingrandimento costante.

Nella realtà costruttiva, nessuna lente fisica è in grado di soddisfare simultaneamente e in modo perfetto queste tre condizioni per ogni punto dello spazio; le deviazioni dei raggi luminosi rispetto a questo modello teorico costituiscono, appunto, le aberrazioni ottiche.

Aberrazioni del terzo ordine

Espandendo l'approssimazione trigonometrica della legge di Snell fino al terzo ordine, è possibile classificare i principali scostamenti dei raggi rispetto al modello parassiale ideale. In un sistema ottico reale, questi difetti si dividono in cinque aberrazioni monocromatiche geometriche, note come aberrazioni di von Seidel, a cui si aggiungono due aberrazioni cromatiche, derivanti invece dalla dispersione della luce attraverso il vetro [5, cap. 11].

Nelle lenti impiegate in ambito fotografico e cinematografico, le eventuali aberrazioni residue visibili nell'immagine appartengono tipicamente a questo terzo ordine. Al contrario, in campi che richiedono risoluzioni estreme, come la microscopia o l'astronomia, diviene necessario modellare e compensare anche le aberrazioni geometriche di ordini superiori.

Le sette aberrazioni principali si classificano dunque in:

- **Aberrazioni monocromatiche (di Seidel):**
  - Aberrazione sferica
  - Coma
  - Astigmatismo
  - Curvatura di campo (o di Petzval)
  - Distorsione
- **Aberrazioni cromatiche:**
  - Aberrazione cromatica assiale
  - Aberrazione cromatica laterale

Aberrazione sferica

L'aberrazione sferica è un difetto che si presenta quando il rapporto tra l'apertura della lente e la sua lunghezza focale risulta particolarmente elevato. In presenza di questa anomalia, il fascio di raggi rifratti non va a convergere in un singolo punto ideale. Al contrario, i raggi che attraversano la porzione periferica della lente (raggi marginali) deviano maggiormente e convergono sull'asse ottico a una distanza inferiore rispetto ai raggi che attraversano la zona centrale (raggi parassiali). Questa discrepanza genera una CoC non uniforme. In base alla posizione del piano focale rispetto ai due fuochi, la PSF assumerà conformazioni differenti: se il piano intercetta il fuoco marginale si otterrà un nucleo denso circondato da un alone; se intercetta il fuoco parassiale, si formerà un disco sfocato con un anello luminoso sul perimetro dello stesso. Tale comportamento definisce l'aberrazione sferica positiva, caratteristica delle lenti convesse. Al contrario, le lenti concave presentano un'aberrazione sferica negativa, in cui il fuoco marginale si proietta oltre quello parassiale [5, § 11.3]. Visivamente, questa aberrazione conferisce all'immagine una morbidezza diffusa. Sebbene in alcuni specifici contesti fotografici e cinematografici questo effetto possa essere ricercato per ragioni puramente estetiche, la progettazione ottica tradizionale mira alla sua eliminazione. L'approccio pratico più immediato per mitigare il difetto consiste nella chiusura del diaframma, un'operazione che blocca fisicamente l'ingresso dei raggi marginali. A livello ingegneristico e costruttivo, invece, il problema viene compensato alla radice accoppiando elementi con aberrazioni di segno opposto, come l'alternanza strutturale di lenti convesse e concave all'interno dello stesso gruppo ottico.

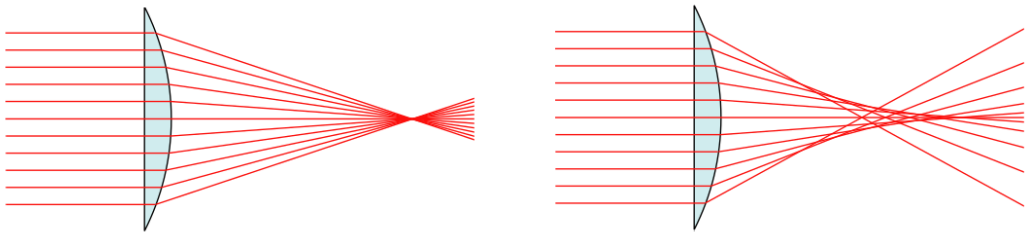


Figura 2.5: Modello di proiezione prospettica della formazione di un'immagine senza aberrazioni vs. in presenza di aberrazione sferica positiva [7].

Coma

Il coma è un'aberrazione ottica che si manifesta quando i raggi luminosi provengono da un punto oggetto situato fuori dall'asse ottico principale, presentandosi anche per deviazioni angolari minime. L'entità di questo difetto aumenta proporzionalmente alla distanza del punto dall'asse e risulta particolarmente pronunciata quando il

sistema lavora con ampie aperture del diaframma. In sostanza, come illustrato in Figura 2.6, il difetto è causato da una variazione dell'ingrandimento trasversale: i raggi che attraversano le porzioni marginali della lente convergono in punti focali sfalsati lateralmente rispetto a quelli formati dai raggi che ne attraversano il centro. Il risultato visivo sul piano sensore è che il punto luminoso non viene riprodotto in modo uniforme, bensì assume una caratteristica forma a cometa (da cui deriva il nome dell'aberrazione). Di conseguenza, il *bokeh* di una lente affetta da coma perde la sua classica regolarità circolare: la macchia di sfocatura si deforma, presentando un nucleo luminoso denso affiancato da una coda più ampia e sfumata, la cui direzione varia a seconda della posizione del punto rispetto all'asse ottico. Proprio per via di queste sue caratteristiche, il difetto risulta notoriamente critico in ambito astronomico più che cinematografico: le stelle, essendo sorgenti luminose puntiformi ad alto contrasto spesso collocate ai bordi del campo visivo, agiscono come perfetti rivelatori, rendendo la distorsione asimmetrica immediatamente e spiccatamente percepibile [5, § 11.4].

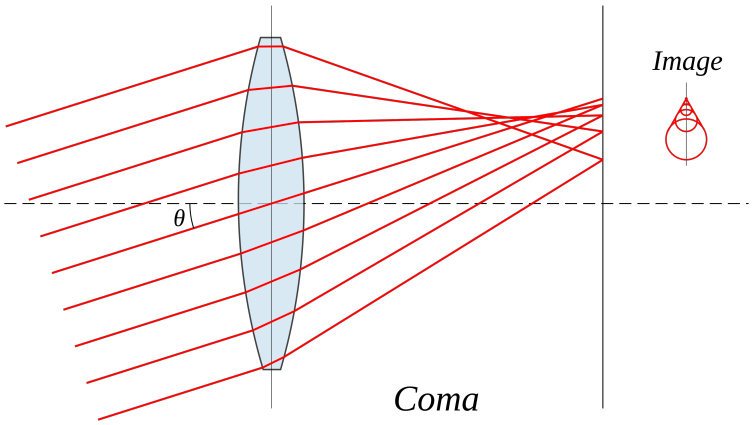


Figura 2.6: Modello prospettico della formazione di un'immagine in una lente affetta da coma [8].

Astigmatismo

L'astigmatismo è un'aberrazione che si manifesta quando un fascio di luce, proveniente da un punto fuori asse, attraversa una lente che presenta una inclinazione o una forma non perfettamente uniforme rispetto all'angolo di incidenza del raggio. Per questo motivo, i raggi che viaggiano su un piano (il piano tangenziale) vedono una curvatura diversa rispetto ai raggi che viaggiano sul piano ortogonale (il piano sagittale). Il risultato è che questi due gruppi di raggi non convergono insieme

nello stesso punto, ma creano due linee focali perpendicolari separate nello spazio. Poiché la luce converge prima su un piano e poi sull'altro, il fascio di raggi, nello spazio compreso tra queste due linee, non può formare un punto rotondo: sarà costretto a schiacciarsi, assumendo una forma ellittica che cambia orientamento man mano che ci si sposta in profondità. Un aspetto fondamentale che distingue l'astigmatismo dalle altre aberrazioni analizzate è proprio questa sua peculiare evoluzione tridimensionale: la distorsione non è statica, ma subisce una trasformazione geometrica continua, passando da una linea all'altra attraverso una forma ellittica, man mano che ci si sposta in profondità lungo l'asse di propagazione [5, § 11.5].

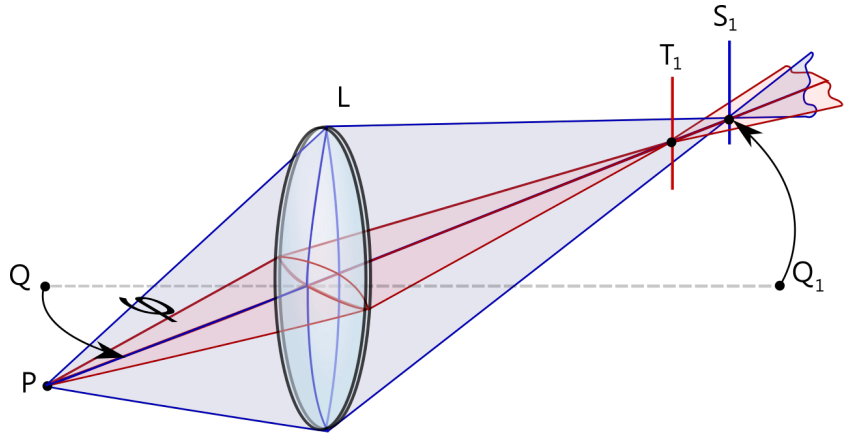


Figura 2.7: Modello prospettico della formazione di un'immagine in una lente affetta da astigmatismo [9].

Curvatura di campo

La curvatura di campo (o di Petzval) si manifesta quando un oggetto piano, posizionato perpendicolarmente all'asse ottico, non può essere messo a fuoco nella sua interezza sul piano del sensore. Invece di proiettarsi su un piano, l'immagine nitida si forma su una superficie curva: di conseguenza, se si mette a fuoco il centro dell'immagine, i bordi risulteranno sfocati, e viceversa. Questo fenomeno è causato dal fatto che il percorso ottico dei raggi che attraversano la zona centrale della lente differisce da quello dei raggi che attraversano le zone periferiche, portando alla focalizzazione su piani distinti. Tale aberrazione rappresenta un chiaro esempio di violazione della seconda condizione di Maxwell, poiché il piano oggetto ortogonale non viene riprodotto su un piano immagine coerente [5, § 11.6].

Distorsione

A differenza delle aberrazioni precedentemente analizzate, la distorsione non compromette la nitidezza dell'immagine, ma ne altera la geometria. Si tratta di un difetto

per cui l'ingrandimento trasversale del sistema varia in funzione della distanza dall'asse ottico. Questo effetto è particolarmente evidente nelle lenti grandangolari, dove le linee rette che non passano per il centro dell'immagine appaiono visibilmente arcuate. La tipologia di distorsione è legata allo schema ottico: nelle lenti grandangolari è tipica la distorsione "a barile" (*barrel distortion*), in cui le linee si incurvano verso l'esterno; al contrario, in molti teleobiettivi si manifesta una distorsione "a cuscinetto" (*pincushion distortion*), in cui le linee vengono attratte verso il centro [5, § 11.7].

Aberrazione cromatica assiale

Uscendo dal mondo delle aberrazioni monocromatiche, passiamo a quelle che coinvolgono le lunghezze d'onda della luce. L'aberrazione cromatica assiale (o longitudinale) si verifica quando le diverse lunghezze d'onda dello spettro visibile non convergono nello stesso punto lungo l'asse ottico. Poiché l'indice di rifrazione del vetro varia in base al colore (dispersione), la lente focalizza i raggi blu più vicino alla lente rispetto ai raggi rossi. Questo comporta che, se si mette a fuoco un oggetto su un piano, le componenti cromatiche che focalizzano altrove risulteranno sfocate, creando tipici aloni colorati attorno ai contorni degli oggetti. A differenza di quella laterale, questo difetto è intrinsecamente legato alla profondità di campo: la sua visibilità dipende fortemente dall'apertura del diaframma, poiché restringendolo si riduce il fascio di luce aberrato, mitigando parzialmente il fenomeno [5, § 13.1].

Aberrazione cromatica laterale

L'aberrazione cromatica laterale (o trasversale) si manifesta perpendicolarmente all'asse ottico. Anche in questo caso la causa risiede nella dispersione cromatica del vetro, ma l'effetto si traduce in una variazione dell'ingrandimento dell'immagine al variare della lunghezza d'onda: le diverse componenti cromatiche vengono rifratte con angoli leggermente differenti, producendo proiezioni di dimensioni diverse per ogni canale. Il risultato è la comparsa di frange colorate (tipicamente magenta e verde) lungo i bordi degli oggetti ad alto contrasto, soprattutto nelle zone periferiche, dove l'effetto è più marcato mentre tende ad annullarsi sull'asse. A differenza di quella assiale, la distorsione laterale non è influenzata dalla chiusura del diaframma; la correzione avviene principalmente in fase di progettazione mediante l'uso di elementi ottici apocromatici o, in post-produzione, tramite algoritmi che riscalano i canali cromatici in modo differenziato [5, § 13.2].

2.1.8 Polinomi di Zernike

I polinomi di Zernike, che devono il nome al fisico olandese Frits Zernike, costituiscono una sequenza di funzioni ortogonali continue sul disco unitario. Proprio questa peculiare proprietà matematica ne ha favorito la massiccia adozione in svariati ambiti dell'ottica.

Al giorno d'oggi infatti, sia in contesti quali astronomia e microscopia, sia in ambiti come cinematografia e fotografia, le lenti che compongono i sistemi di imaging utilizzati sono solitamente circolari; ciò rende perfetti questi polinomi per la descrizione del fronte d'onda di una lente reale e, di conseguenza, per la descrizione delle aberrazioni ottiche di cui tale lente può soffrire.

Ovviamente, di questi polinomi se ne può fare un'espansione in serie, e ogni termine di questa espansione è utile a rappresentare una specifica aberrazione ottica. È importante notare però che sviluppi in serie complessi, oltre il decimo termine, sono necessari in contesti quali la sopracitata microscopia, mentre in contesti fotografici le aberrazioni che possono manifestarsi sono rappresentabili con i primi termini di questa espansione [10].

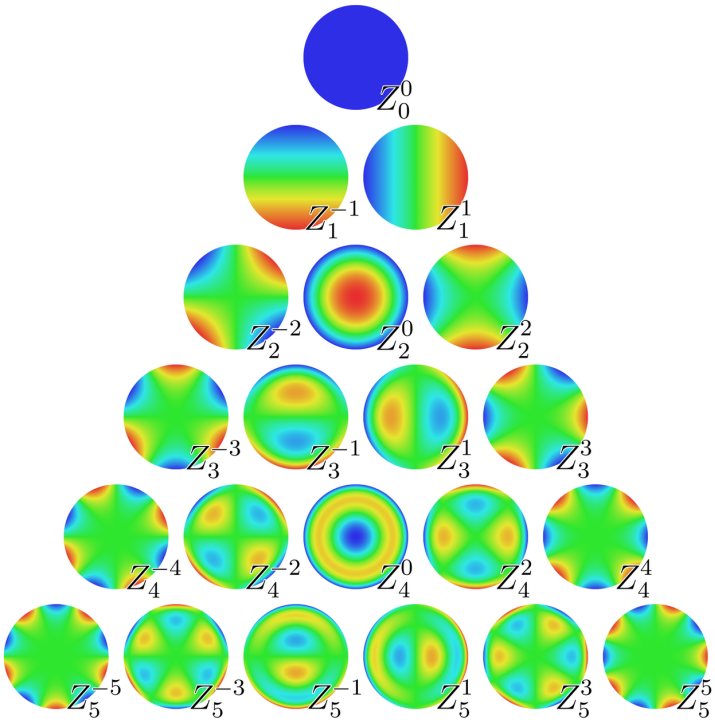


Figura 2.8: Rappresentazione dei primi 21 polinomi di Zernike [11].

In generale, una funzione atta alla descrizione di un fronte d'onda in coordinate polari, che possiamo identificare come $W(\rho, \theta)$, può essere espansa in una serie di polinomi ortonormali sulla superficie del disco unitario secondo la seguente formula:

$$W(\rho, \theta) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} C_n^m Z_n^m(\rho, \theta) \tag{2.5}$$

Dove C_n^m è il coefficiente rappresentante l'ampiezza di ciascun polinomio Z_n^m, mentre gli indici n e m servono rispettivamente a descrivere l'ordine radiale del polinomio e la frequenza angolare. Il polinomio Z_n^m può essere descritto in due modi, a seconda della sua simmetria rispetto all'asse ottico. I polinomi pari sono descritti da:

$$Z_n^m(\rho, \theta) = R_n^m(\rho) \cos(m\theta) \quad m \geq 0 \tag{2.6}$$

Mentre i polinomi dispari sono descritti da:

$$Z_n^{-m}(\rho, \theta) = R_n^m(\rho) \sin(m\theta) \quad m > 0 \tag{2.7}$$

Dove ρ è la distanza radiale normalizzata compresa tra zero e uno, θ è l'angolo polare, mentre $R_n^m(\rho)$ è il polinomio radiale di Zernike, che a sua volta è definito come:

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} \rho^{n-2k} \tag{2.8}$$

Esiste una notazione alternativa per identificare i polinomi di Zernike che mappa gli indici n e m ad un singolo indice j. In questo modo l'identificazione di un polinomio risulta più immediata; tale notazione è definita dagli indici sequenziali di Noll [12] e mappa i due indici sopracitati secondo la seguente regola:

$$j = \frac{n(n+1)}{2} + |m| + \begin{cases} 0, & m > 0 \wedge n \equiv \{0, 1\} \\ 0, & m < 0 \wedge n \equiv \{2, 3\} \\ 1, & m \geq 0 \wedge n \equiv \{2, 3\} \\ 1, & m \leq 0 \wedge n \equiv \{0, 1\} \end{cases} \tag{2.9}$$

Alcune delle aberrazioni citate precedentemente possono essere descritte attraverso le funzioni di ordine inferiore appartenenti a questa serie, come riportato di seguito:

- Defocus ($j = 4$):

$$Z_2^0(\rho, \theta) = \sqrt{3}(2\rho^2 - 1) \tag{2.10}$$

- Astigmatismo - componente obliqua ($j = 5$):

$$Z_2^{-2}(\rho, \theta) = \sqrt{6}\rho^2 \sin(2\theta) \tag{2.11}$$

- Astigmatismo - componente verticale ($j = 6$):

$$Z_2^2(\rho, \theta) = \sqrt{6}\rho^2 \cos(2\theta) \tag{2.12}$$

- Coma - componente verticale ($j = 7$):

$$Z_3^{-1}(\rho, \theta) = \sqrt{8}(3\rho^3 - 2\rho) \sin(\theta) \tag{2.13}$$

- Coma - componente orizzontale ($j = 8$):

$$Z_3^1(\rho, \theta) = \sqrt{8}(3\rho^3 - 2\rho) \cos(\theta) \tag{2.14}$$

- Aberrazione sferica ($j = 11$):

$$Z_4^0(\rho, \theta) = \sqrt{5}(6\rho^4 - 6\rho^2 + 1) \tag{2.15}$$

È infine importante fare una considerazione riguardo alla relazione matematica e ottica tra l'aberrazione sferica e il defocus, facilmente osservabile attraverso queste funzioni. Analizzandone la formula, si nota come esso presenti una totale simmetria circolare (non essendo influenzato dall'angolo θ), una proprietà condivisa con l'aberrazione sferica. Ciò che rende l'aberrazione sferica intimamente legata al defocus è la presenza, al suo interno, della componente $-\rho^2$: essendo di segno negativo, questa agisce in direzione opposta rispetto alla componente pura di defocus, permettendo a queste due aberrazioni di compensarsi parzialmente a vicenda nello spazio reale.

Avendo definito un modello matematico rigoroso per descrivere il fronte d'onda attraverso i polinomi di Zernike, possediamo ora gli strumenti teorici necessari per parametrizzare e quantificare i difetti di una lente fisica. La vera sfida, tuttavia, consiste nel tradurre questa astrazione analitica in un modello computazionale in grado di simulare fisicamente il percorso della luce e la conseguente formazione dell'immagine. È proprio in questo contesto di sintesi visiva che entrano in gioco i motori di rendering e gli algoritmi di ray-tracing, i quali rappresentano il fulcro della prossima sezione e l'infrastruttura fondamentale su cui si basa il nostro lavoro.

2.2 Ray-Tracing

Questa sezione analizza le fondamenta teoriche e l'evoluzione storica degli algoritmi di ray-tracing. Il percorso si snoda attraverso una sintesi dell'evoluzione tecnologica di tali algoritmi, per poi approfondire lo stato dell'arte relativo alla resa del *bokeh* nei sistemi di ray-tracing, concludendosi infine nella trattazione delle sfide tecniche legate alla modellazione delle aberrazioni ottiche in queste architetture.

2.2.1 Fondamenti di Ray-Tracing

Il ray-tracing è una tecnica di sintesi delle immagini basata sulla simulazione del comportamento della luce secondo le leggi dell'ottica geometrica. Nel contesto della computer grafica, gli algoritmi standard appartenenti a questa famiglia sono utilizzati per generare immagini fotorealistiche simulando il percorso dei raggi luminosi che attraversano una scena tridimensionale fino a raggiungere la camera virtuale.

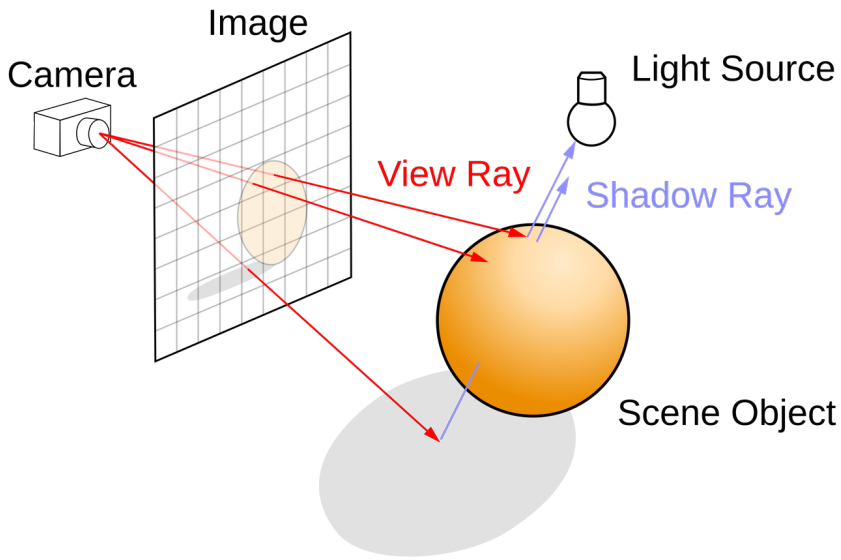


Figura 2.9: Modello semplificato del processo di formazione di un'immagine tramite ray-tracing [13].

La prima iterazione concettuale di questa tecnologia risale al 1968, quando Arthur Appel presentò un pionieristico algoritmo di ray-casting [14]. Pur trattandosi di un approccio basilare rispetto agli standard moderni, fu il primo a risolvere una limitazione critica dei sistemi grafici dell'epoca (basati su rasterizzazione e scanline rendering): la capacità di riprodurre fedelmente i volumi tramite la corretta proiezione delle ombre. Il sistema generava una griglia di punti nello

spazio, che andava a coincidere con il piano immagine, e tracciava un singolo raggio dall'osservatore attraverso ogni punto della griglia fino all'intersezione con le geometrie della scena. Valutando l'inclinazione della superficie intersecata rispetto alla sorgente luminosa, il sistema era in grado di calcolare ombreggiature e chiaroscuri.

Gli algoritmi successivi basati sul ray-casting, come il celebre modello di illuminazione locale presentato da Bui Tuong Phong nel 1975 [15], pur apportando notevoli migliorie visive, modellavano la luce in modo fortemente semplificato, ignorando fenomeni di illuminazione globale come riflessioni inter-oggetto o rifrazioni complesse. Un punto di svolta si ebbe nel 1979, quando Turner Whitted propose un algoritmo in grado di simulare i sopracitati fenomeni, introducendo il ray-tracing ricorsivo [16]. In questo modello, all'impatto di un raggio primario con una superficie, l'algoritmo generava molteplici raggi secondari: un raggio riflesso, un raggio trasmesso (in caso di mezzi dielettrici e trasparenti come vetro o acqua) e i cosiddetti shadow rays puntati verso le sorgenti luminose per verificare la visibilità diretta. I raggi secondari potevano a loro volta collidere con altre geometrie, generando una struttura computazionale definita "albero dei raggi" (*ray tree*). Tuttavia, un tale numero di ricorsioni introduceva una complessità computazionale che non poteva essere ignorata. Per aggirare questo problema, Whitted introdusse l'uso dei bounding volumes, volumi geometrici semplificati che racchiudono geometrie complesse, permettendo di scartare a priori le intersezioni non necessarie e ottimizzando drasticamente le prestazioni.

Un ulteriore, fondamentale passo avanti si verificò nel 1984, quando Robert L. Cook, Thomas Porter e Loren Carpenter presentarono il Ray-Tracing Distribuito (Distributed Ray-Tracing) [17]. Fino a quel momento, le immagini generate con il ray-tracing soffrivano di una perfezione algoritmica innaturale e di pesanti artefatti di aliasing. L'approccio distribuito risolse queste criticità introducendo il campionamento stocasticobasato sul lancio di molteplici raggi per ogni singolo pixel con distribuzioni spaziali e temporali leggermente variate. Questa innovazione permise la simulazione fisicamente accurata di ombre morbide (soft shadows), riflessioni glossy, motion blur e, fattore centrale per il presente lavoro, la profondità di campo (*Depth of Field*). Il modello concettuale a foro stenopeico (*pinhole camera*) venne definitivamente superato: i raggi non convergevano più in un singolo punto ideale, ma venivano campionati sulla superficie di una lente virtuale, riproducendo il meccanismo di messa a fuoco di un sistema ottico reale.

Direzionalità del Ray Tracing

A livello architetturale, il calcolo del trasporto di luce può essere affrontato secondo due direttrici principali:

- **Ray-Tracing Diretto (Forward Ray-Tracing):** Simula la propagazione fisica reale, proiettando i raggi di luce dalle sorgenti emissive verso la scena. Dopo molteplici interazioni e deflessioni basate sui materiali (BRDF), solo una minuscola frazione di questi raggi intercetterà effettivamente il piano immagine. Pur essendo fisicamente ineccepibile, risulta estremamente inefficiente dal punto di vista computazionale, sprecando immense risorse per calcolare percorsi luminosi che non contribuiranno mai all'immagine finale.

- **Ray-Tracing Inverso (Backward Ray-Tracing):** Ribalta il paradigma, facendo partire i raggi dal sensore (o dall'osservatore) per farli viaggiare a ritroso attraverso il piano immagine verso la scena, terminando il percorso solo quando colpiscono una sorgente luminosa. L'algoritmo di Whitted [16] si fonda su questo principio, che garantisce che ogni percorso calcolato contribuisca all'immagine generata. Per via della sua efficienza superiore rispetto al ray-tracing diretto, il ray-tracing inverso è lo standard *de facto* su cui si basano quasi tutti i moderni motori di rendering.

2.2.2 Rappresentazione del Bokeh in sistemi di Ray-Tracing

Prima di affrontare le problematiche relative alla simulazione delle imperfezioni ottiche, è fondamentale comprendere come la profondità di campo e l'effetto *bokeh* di base vengano generati all'interno di un motore di rendering tradizionale. Nei sistemi standard, infatti, l'ottica della fotocamera virtuale è quasi universalmente approssimata utilizzando il modello matematico della Lente Sottile (*Thin Lens*).

Tale modello considera la lente come un elemento idealmente privo di spessore, basando il calcolo del trasporto luminoso su un'astrazione geometrica perfetta. A livello algoritmico, il sampler seleziona una coordinata sul piano del sensore e un punto stocastico sulla superficie dell'apertura circolare virtuale. Il raggio originato da questi due punti viene tracciato e proiettato attraverso il piano di fuoco per determinare le intersezioni con la scena tridimensionale.

Se l'oggetto intersecato giace esattamente sul piano focale, tutti i raggi campionati sull'apertura e passanti per quel punto convergeranno in un singolo pixel sul sensore, producendo un'immagine nitida. Viceversa, qualora l'oggetto si trovi al di fuori di tale piano, i raggi intersecheranno punti spaziali differenti; di conseguenza,

la radianza raccolta verrà distribuita su un'area più estesa del piano immagine, dando geometricamente origine al Circle of Confusion (CoC).

Proprio per via della natura puramente teorica della *Thin Lens*, i raggi luminosi non subiscono alcuna distorsione locale attraversando l'apertura. Di conseguenza, il *bokeh* generato da questo approccio si presenterà come un cerchio geometricamente perfetto, caratterizzato da una distribuzione di intensità rigorosamente costante (fatta eccezione per eventuali cadute di luce dovute a vignettatura meccanica). Pur essendo un risultato matematicamente ineccepibile per un sistema ideale, questa omogeneità assoluta lo priva delle naturali asimmetrie tipiche delle lenti fisiche.

2.2.3 Aberrazioni nel Ray-Tracing

Nonostante il modello a lente sottile sia efficiente dal punto di vista computazionale, esso genera immagini otticamente ineccepibili, che come già accennato sono prive del carattere organico che contraddistingue fotografia e cinematografia reale. Per quanto si tenda ad associare in modo intuitivo l'assenza di difetti al massimo grado di realismo, nell'ambito della sintesi di immagini un *bokeh* perfettamente omogeneo e perfetto dal punto di vista geometrico finisce per svelare la natura sintetica del render, risultando paradossalmente meno credibile all'occhio umano rispetto a una sfocatura fisicamente imperfetta.

L'introduzione delle aberrazioni ottiche all'interno degli algoritmi di ray-tracing è un tema dibattuto nella letteratura di settore. Un approccio fisicamente rigoroso per superare i limiti della lente sottile è il ricorso al modello a "Lente Spessa" (*Thick Lens Model* o *Lens Prescription*), come esplorato ad esempio da Wu et al. [18]. In questo scenario l'algoritmo abbandona l'astrazione bidimensionale per simulare il reale passaggio dei raggi attraverso un gruppo ottico composto da molteplici elementi in vetro. Riproducendo accuratamente lo spessore, i raggi di curvatura e gli specifici indici di rifrazione di ciascuna lente, le aberrazioni — quali aberrazione sferica, coma e astigmatismo — emergono in modo del tutto naturale come conseguenza diretta delle leggi della rifrazione.

Nonostante l'eccellente fedeltà visiva prodotta, questo approccio porta con sé due criticità fondamentali. La prima è di natura prestazionale: il calcolo di complesse intersezioni raggio-superficie per molteplici elementi curvi, ripetuto per ogni singolo raggio campionato, incrementa drasticamente i tempi di rendering. La seconda, ancor più vincolante, è di natura pratica: le *lens prescriptions* (ovvero le esatte specifiche costruttive geometriche) costituiscono segreti industriali rigorosamente protetti dai produttori di ottiche. L'adozione di questo modello vincola pertanto la simulazione all'impiego di brevetti storici scaduti o design fortemente semplificati, precludendo di fatto la riproduzione fedele delle moderne lenti cinematografiche.

2.3 Obiettivi della Ricerca (Research Questions)

Il presente lavoro di tesi si propone come diretta evoluzione e ottimizzazione di una precedente iterazione algoritmica esplorativa. In tale approccio, la simulazione delle aberrazioni ottiche nel ray-tracer veniva affidata alla costruzione di due Look-Up Tables (LUT) tridimensionali: una adibita a mappare le deviazioni direzionali del frustum e l'altra utilizzata per forzare la corretta distribuzione dell'intensità luminosa nel campionatore. Queste strutture dati venivano generate interpolando un set discreto di PSF pre-catturate a vari livelli di defocus.

Sebbene visivamente promettente, questo primo approccio presentava due criticità fondamentali: l'enorme onere computazionale e di memoria richiesto per generare e stoccare matrici 3D sufficientemente dense, e la rigidità intrinseca di un modello basato su dati campionati discretamente. L'uso di LUT volumetriche limitava la capacità del sistema di generalizzare in modo fluido e continuo l'evoluzione del *bokeh* nello spazio intermedio tra due campioni noti.

Per superare queste limitazioni algoritmiche, l'obiettivo primario di questo lavoro è quello di concepire un'architettura computazionale elegante, in grado di bypassare la dipendenza da informazioni costruttive inaccessibili sfruttando esclusivamente l'analisi delle risposte all'impulso del sistema. Nello specifico, la ricerca intende esplorare e rispondere alle seguenti domande (*Research Questions*):

- **RQ1 (Analisi e Phase Retrieval):** È possibile estrarre in modo robusto le aberrazioni latenti di un sistema ottico (sotto forma di coefficienti di Zernike) operando unicamente in modalità inversa, ovvero analizzando uno stack di risposte all'impulso (PSF) simulate e affette da defocus variabile, aggirando così il limite delle *lens prescriptions* proprietarie?
- **RQ2 (Modellazione Vettoriale Continua):** In che modo si può superare il limite della discretizzazione spaziale tipico delle LUT tridimensionali, trasformando l'errore del fronte d'onda in un modello vettoriale continuo (Frustum Aberrato) capace di descrivere matematicamente l'esatta traiettoria (pendenza e intercetta) di ogni singolo raggio luminoso che attraversa la pupilla?
- **RQ3 (Sintesi e Rendering):** È possibile integrare questo nuovo modello vettoriale continuo all'interno dell'architettura stocastica di un motore di ray-tracing avanzato (Mitsuba 3) affinché l'intensità e la forma organica del *bokeh* emergano in modo del tutto naturale? L'obiettivo è dimostrare che, deformando dinamicamente i raggi (warping), non è più necessario forzare artificialmente la luminosità, poiché le caustiche ottiche si manifesteranno come conseguenza diretta dell'addensamento geometrico dei raggi.

Rispondendo a queste tre domande, il presente lavoro mira a creare una pipeline end-to-end che, partendo dalle specifiche aberrazioni definite dall'utente, conduca alla generazione di render virtuali dotati di un *bokeh* fisicamente equivalente al profilo ottico sintetizzato, coniugando per la prima volta l'assoluta fedeltà ottica con prestazioni computazionali adatte ad ambienti di produzione.

Capitolo 3

Metodologia

3.1 Mitsuba 3.0

L'ambiente scelto per lo sviluppo dell'algoritmo è Mitsuba (nella sua versione 3.0), un motore di rendering dedicato alla simulazione fisicamente accurata del trasporto di luce diretto e inverso. La scelta è ricaduta su questo framework per via della sua elevata flessibilità architetturale e dell'ampio margine di personalizzazione garantito all'utente. Mitsuba permette infatti di implementare e modificare agevolmente i vari elementi tipici di un sistema di sintesi di immagini basato sul ray-tracing [1].

Una delle caratteristiche peculiari che si è rivelata determinante per la sua adozione, oltre alla suddetta modularità, è il sistema di varianti basato sul compilatore Just-In-Time (JIT) Dr.Jit [19]. Questa architettura permette di eseguire il rendering sfruttando l'accelerazione hardware tramite backend LLVM (per l'esecuzione ottimizzata su CPU) o CUDA (su GPU NVIDIA), garantendo al contempo la possibilità di calcolare gradienti per il rendering differenziabile. Inoltre, il sistema consente di operare nativamente nello spazio colore RGB o nel dominio spettrale, a seconda del livello di accuratezza fisica richiesto dalla simulazione.

L'architettura di Mitsuba, così come quella dei motori di rendering in generale, può essere suddivisa in tre componenti fondamentali che, nel nostro specifico contesto di sviluppo, possono essere implementati in modo indipendente e modulare:

3.1.1 Integratori

Gli integratori rappresentano il nucleo computazionale del motore di rendering: hanno il compito di risolvere matematicamente l'equazione del trasporto luminoso (la cosiddetta rendering equation) per descrivere la traiettoria e l'accumulo della luce all'interno della scena. Diversi motori implementano svariate tipologie di integratori,

ciascuno ottimizzato per calcolare la propagazione della luce con approcci differenti. Un classico path tracer, ad esempio, stima l'illuminazione globale tracciando stocasticamente i percorsi dei raggi in modo indiretto, esplorando la scena a partire dal sensore fino a raggiungere le sorgenti emissive. Altri integratori, come il volumetric path tracer (`volpath`), lavorano invece computando in modo coerente la volumetria in presenza di nebbie. Questa variante estende le capacità del sistema risolvendo la più generale Equazione del Trasporto Radiativo (RTE): ciò permette di simulare non solo le riflessioni sulle geometrie solide, ma anche le complesse interazioni fisiche della luce — quali assorbimento, emissione e scattering (dispersione) — all'interno di mezzi continui, come foschia, fluidi o vetro.

3.1.2 Campionatori

Il campionatore (o sampler) è il componente che si occupa della generazione dei punti di campionamento nello spazio multidimensionale. Il suo compito è decidere l'origine e la distribuzione statistica dei raggi che verranno proiettati nella scena per essere poi valutati dall'integratore. Si tratta di un elemento critico per l'efficienza dell'intero sistema: l'obiettivo primario dei moderni algoritmi di campionamento è produrre un'immagine di alta qualità utilizzando il minor numero di campioni possibile, minimizzando così la varianza stocastica e i conseguenti artefatti visivi. Un sampler comune è l'independent sampler, che genera campioni in modo completamente casuale, ma esistono anche approcci più sofisticati come il multijittered sampler [20] che permette di ottenere una distribuzione più uniforme dei campioni utilizzandone un numero inferiore rispetto a metodi tradizionali.

3.1.3 Sensori

All'interno dell'architettura di Mitsuba, il termine "Sensore" non indica solamente il componente fisico che cattura la luce, ma rappresenta il modo in cui il software approssima la camera all'interno dell'ambiente virtuale. Il suo ruolo va quindi oltre la semplice raccolta dell'energia luminosa. A partire dalla posizione, dall'orientamento e dal modello ottico selezionato (ad esempio la lente sottile), il sensore costruisce materialmente, per ogni campione generato sull'apertura, il relativo frustum di raggi, direzionando ciascuno di essi nello spazio. Va da sé che questo componente si rivelerà di fondamentale importanza all'interno della nostra architettura: avendo il controllo assoluto sulla direzione dei raggi primari, ci permetterà di deviarne matematicamente la traiettoria per simulare le aberrazioni ottiche prima ancora che avvenga l'intersezione con la scena.

3.2 Sviluppo Architetturale

Per poter arrivare all'architettura definitiva del sensore si è passati attraverso un processo di sviluppo iterativo, resosi necessario per superare le complessità legate alla simulazione volumetrica della luce. Il limite computazionale e teorico più stringente non risiedeva nella rappresentazione coerente delle aberrazioni su un singolo piano focale, bensì nella modellazione matematicamente corretta della loro evoluzione al variare del defocus, ovvero nello spazio tridimensionale.

Considerando una tipica scena caratterizzata da una profondità di campo ridotta, un oggetto posto sul piano focale presenterà una Point Spread Function (PSF) compatta, dettata dal limite di diffrazione e dalle aberrazioni residue. Al contrario, un elemento in background esibirà un *bokeh* fortemente espanso. La morfologia di quest'ultimo non è una semplice scalatura della PSF a fuoco, ma il risultato di un'evoluzione caustica complessa. Riprodurre questa coerenza spaziale ha richiesto l'esplorazione, e il conseguente superamento, di diverse architetture intermedie.

3.2.1 Disaccoppiamento Sampler-Sensore

Nelle fasi preliminari della ricerca, il sistema è stato strutturato provando a separare le responsabilità della simulazione tra due componenti distinti del motore di rendering: il campionatore (sampler) e il sensore. L'informazione relativa alla deviazione dei raggi, calcolata a partire dal gradiente dei coefficienti di Zernike, veniva passata al sensore sotto forma di mappa angolare. Parallelamente, la distribuzione spaziale dei campioni sull'apertura veniva affidata a un sampler personalizzato, istruito per generare punti seguendo statisticamente la morfologia della PSF.

Questo approccio ha presentato rapidamente due limiti evidenti:

- **Troncamento delle code di diffrazione:** Il calcolo della deviazione angolare basato esclusivamente sui gradienti analitici del fronte d'onda è vincolato all'approssimazione dell'ottica geometrica. Le derivate locali dei polinomi di Zernike non producono angoli di inclinazione sufficientemente severi da proiettare i raggi nelle estese e deboli aree periferiche della PSF. Ciò causava la mancata simulazione dei fenomeni ondulatori e la conseguente perdita delle delicate code di diffrazione tipiche di aberrazioni come il coma. Come vedremo, questa criticità verrà aggirata nell'architettura finale estraendo le traiettorie (slope e intercept) direttamente dai piani generati tramite Fast Fourier Transform (FFT), la quale conserva intrinsecamente l'informazione ondulatoria necessaria a riprodurre tali componenti periferiche.

- **Incoerenza spaziale:** Dividere la simulazione tra sampler e sensore rompeva l'integrità del frustum ottico. In ottica fisica, la traiettoria del raggio non può essere gestita in fasi separate; l'origine e la deviazione angolare devono essere valutate congiuntamente per mantenere la coerenza volumetrica.

3.2.2 Mappa di profondità

Per ovviare al problema della coerenza spaziale, una seconda iterazione ha esplorato l'utilizzo di una mappa di profondità (Z-Buffer). L'idea alla base era quella di effettuare un pre-pass di rendering per estrarre la distanza esatta di ogni geometria dalla camera, fornendo al sensore la cognizione topologica della scena. In base alla profondità letta dalla Depth Map, il sistema avrebbe applicato dinamicamente la PSF corrispondente a quello specifico livello di defocus.

Nonostante la validità concettuale, l'approccio depth-aware è stato scartato a causa di gravi imprecisioni intrinseche sui bordi degli oggetti (edge artifacts). Il *bokeh*, per sua natura ottica, comporta che la luce proveniente dagli oggetti in background invada parzialmente le silhouette degli oggetti in primo piano sfocati (light bleeding). La Depth Map, operando come una proiezione bidimensionale rigida in spazio schermo, restituiva valori di profondità discreti e netti ai confini delle geometrie. Questo causava la rottura della sfocatura sui contorni, generando aloni irrealistici e artefatti di aliasing che distruggevano l'illusione di un volume ottico continuo.

3.2.3 Approssimazione Discreta in Dominio di Fourier

Accantonato l'approccio basato su proiezione bidimensionale per tornare a un tracciamento puramente geometrico, si è tentato di spostare il carico computazionale interamente sulla fase offline, pre-calcolando la FFT dei polinomi di Zernike ritagliati sull'apertura circolare.

Sebbene questo metodo garantisse una rappresentazione eccellente della PSF sul singolo piano, si è rivelato inefficace nella gestione della profondità. Il frustum di raggi generato tendeva a comportarsi come una banale proiezione planare, limitandosi a scalare dimensionalmente la PSF a fuoco nelle zone defocalizzate. Questo comportamento si discosta dalla realtà fisica: l'aberrazione sferica, ad esempio, genera morfologie drasticamente differenti tra la zona intrafocale (centro luminoso), il fuoco esatto e la zona extrafocale (effetto ad anello). Dinamiche non lineari che una semplice scalatura statica non può riprodurre.

3.2.4 Modello Vettoriale Ibrido

Il superamento di tutte le criticità precedenti ha portato alla formulazione dell'architettura finale, concepita fondendo la precisione geometrica dei gradienti analitici con la coerenza distributiva dei piani FFT, operando integralmente nello spazio tridimensionale nativo del ray-tracing.

La soluzione si concretizza nell'abbandono dell'approccio statistico in favore di un tracciamento vettoriale puro. L'algoritmo definitivo utilizza le derivate esatte di Zernike per calcolare la deviazione angolare nel punto di origine sull'apertura, ma si affida ai piani generati tramite FFT per interpolare matematicamente l'intersezione di queste traiettorie lungo l'asse longitudinale. Estrapolando l'evoluzione dei raggi attraverso molteplici piani di defocus pre-calcolati, il sistema esegue un fitting lineare che restituisce due mappe vettoriali (pendenze e intercette). Questo modello garantisce che i raggi all'interno del frustum siano distribuiti in modo da ricostruire organicamente, senza frammentazioni né artefatti sui bordi, la corretta evoluzione volumetrica della PSF a qualsiasi distanza dalla camera.

3.3 Architettura del Sistema

La pipeline di elaborazione proposta in questo lavoro di tesi è stata concepita secondo un'architettura che si divide in due fasi principali. Queste si integrano sequenzialmente per tradurre le proprietà fisiche di una lente aberrata in mappe vettoriali parametriche, le quali vengono poi caricate e integrate all'interno dell'equazione del rendering di Mitsuba. Il flusso di lavoro si articola come in Figura 3.1:

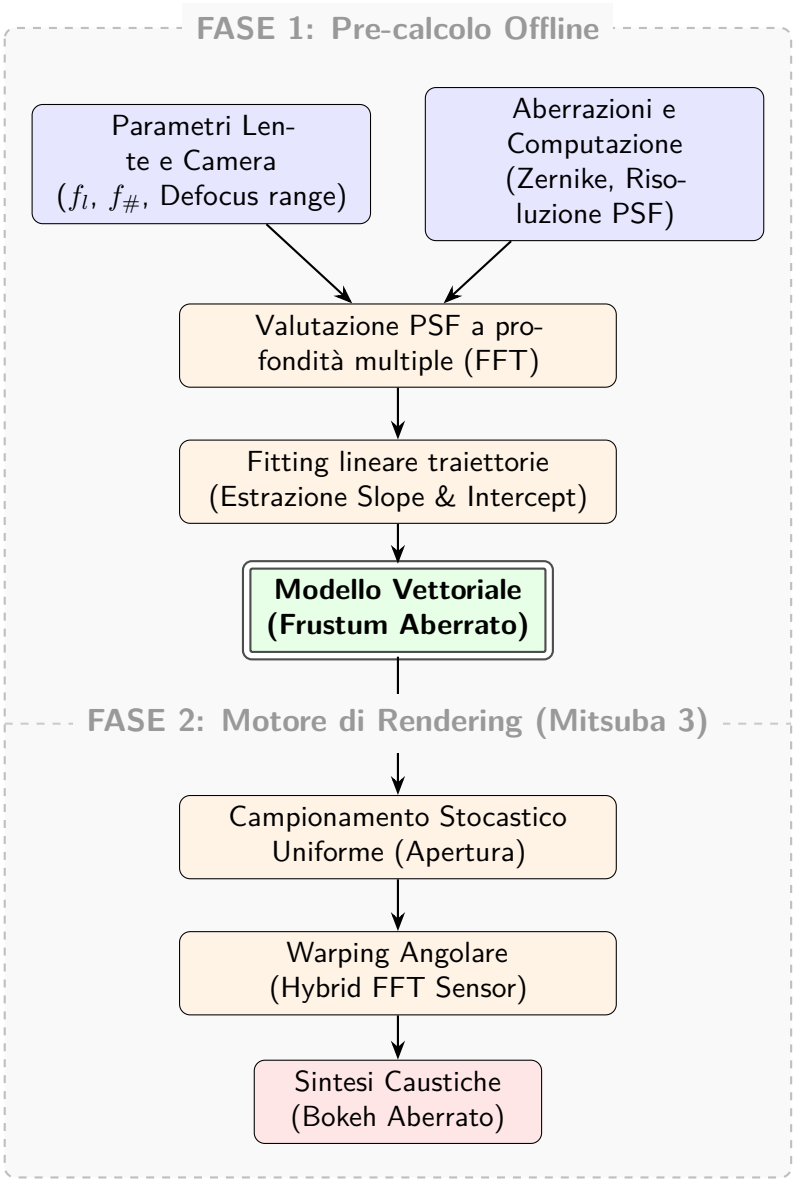


Figura 3.1: Diagramma di flusso dell'architettura proposta: dalla modellazione ottica offline alla sintesi volumetrica integrata nel path tracer.

3.4 Fase 1: Modellazione Ottica

L'architettura del sistema offline viene inizializzata dalle direttive fornite dall'utente tramite l'interfaccia a riga di comando del modulo `PSF_analysis_main.py`. Nello specifico, in fase di inizializzazione l'applicativo riceve in input le proprietà fisiche della lente virtuale (lunghezza focale ed f-stop), il set di coefficienti di Zernike che si desidera simulare e il relativo range di defocus.

Acquisiti questi parametri, la pipeline si articola in due macro-fasi logiche e interdipendenti: la definizione di un modello ottico diretto (Forward Model) e la risoluzione del problema inverso (Phase Retrieval). Il primo implementa le leggi matematiche necessarie a sintetizzare la risposta impulsiva del sistema; il secondo sfrutta iterativamente questo stesso modello per compiere il percorso inverso, deducendo le aberrazioni latenti a partire dai dati appena generati, i quali vengono manipolati per simulare a tutti gli effetti delle reali misurazioni empiriche.

3.4.1 Il Modello Ottico Diretto

Il primo stadio di questa sezione prevede l'inizializzazione dello spazio di lavoro e la discretizzazione dei parametri fisici necessari a simulare la risposta impulsiva del sistema. Per garantire la massima chiarezza e isolare la componente matematica dal flusso del testo, il processo è stato segmentato nei seguenti passaggi logici:

1. Inizializzazione dello Spazio Computazionale

Il primo passaggio dell'algoritmo prevede la definizione della griglia bidimensionale necessaria per il campionamento della FFT. La scelta della risoluzione spaziale (N_{res}) rappresenta un parametro critico dell'architettura: essa definisce direttamente il livello di dettaglio con cui viene ricostruito il difetto ottico. Un campionamento insufficiente comporterebbe infatti la perdita delle micro-strutture della PSF, degradandone significativamente la qualità visiva. Di conseguenza, N_{res} deve essere sufficientemente elevato da preservare l'integrità del segnale ottico, mantenendo al contempo un costo computazionale, in termini di tempo e di memoria, compatibile con la generazione di un intero stack volumetrico.

Per determinare la configurazione più adeguata, l'algoritmo è stato testato valutando tre diversi ordini di grandezza della griglia (Tabella 3.1). Sulla base dei dati emersi, le tre configurazioni si distinguono per i seguenti profili di utilizzo:

- $N_{\text{res}} = 256 \times 256$ pixel: presenta i tempi di calcolo e l'occupazione di memoria più contenuti; tuttavia, la ridotta densità di campionamento penalizza la

definizione spaziale della PSF, compromettendo la fedeltà ottica e il generale realismo del volume simulato.

- $N_{\text{res}} = 512 \times 512$ pixel: offre il compromesso ottimale tra il campionamento spaziale e le performance del sistema; è pertanto la risoluzione adottata per la validazione e il rendering finale.
- $N_{\text{res}} = 1024 \times 1024$ pixel: comporta un severo incremento dei tempi di elaborazione e del peso su disco, imponendo un carico computazionale che risulta ingiustificato ai fini dell'implementazione pratica.

Risoluzione (N_{res})	**Tempo di Calcolo (s)**	**Peso del Volume (MB)**
256×256	133.32 s	4.92 Mb
512×512	1110.80 s	16.57 Mb
1024×1024	4153.80 s	52.44 Mb

Tabella 3.1: Impatto della risoluzione spaziale sui tempi totali di ottimizzazione (phase retrieval) e sull'occupazione di memoria del volume PSF esportato.

2. Discretizzazione del Dominio Spaziale

Una volta definita la griglia, il secondo passaggio consiste nel determinare l'equivalente fisico del singolo pixel (Δx_{fft}) calcolato nel dominio di Fourier, proiettandolo sul piano del sensore. Questo calcolo rappresenta il collegamento tra l'astrazione matematica della trasformata e la scala metrica della fotocamera virtuale, permettendo di mappare i risultati dell'ottica ondulatoria nello spazio tridimensionale di Mitsuba.

Variabili e costanti ottiche:

Lunghezza d'onda di riferimento: $\lambda = 550$ nm (corrispondente al picco di sensibilità del sistema visivo umano).

Rapporto pupillare (pupil ratio): $r_p = 0.25$ (parametro introdotto per prevenire fenomeni di aliasing numerico ai bordi della finestra computazionale).

Apertura relativa (f-stop): $f_{\#}$.

Formulazione matematica:

$$\Delta x_{\text{fft}} = r_p \cdot \frac{\lambda}{10^6} \cdot f_{\#} \tag{3.1}$$

3. Calcolo dell'Apertura Angolare di Diffusione

Ottenuta la dimensione dei pixel, il sistema procede calcolando l'estensione lineare complessiva dell'area di campionamento simulata e la conseguente apertura angolare massima (θ_{psf}) del cono di diffusione della luce. Questo parametro è di cruciale importanza per l'architettura del sistema: definisce infatti il limite geometrico invalicabile del frustum ottico. È questo l'angolo che informerà il sensore personalizzato in fase di tracciamento, stabilendo l'ampiezza massima entro la quale i raggi primari potranno essere deviati.

Componenti e variabili di calcolo:

Larghezza fisica totale dell'area di campionamento: $W_{\text{psf}} = N_{\text{res}} \cdot \Delta x_{\text{fft}}$.

Lunghezza focale della lente: f_l.

Formulazione matematica:

$$\theta_{\text{psf}} = 2 \arctan\left(\frac{W_{\text{psf}}/2}{f_l}\right) \tag{3.2}$$

4. Parametrizzazione del Fronte d'Onda

Il quarto passaggio è dedicato alla modellazione dell'errore descritto dal fronte d'onda sull'apertura del sistema, espresso tramite i polinomi di Zernike. In questa fase, l'implementazione computazionale della funzione `fast_zernike_gradient` all'interno del motore di calcolo è strutturata per processare analiticamente un set completo di aberrazioni primarie, garantendo una deviazione geometrica dei raggi estremamente fedele alla realtà fisica.

Parametri della base di Zernike:

Coordinate normalizzate sull' apertura: (u, v) vincolate alla condizione $u^2 + v^2 \leq 1$.

Coefficienti di ampiezza dell'aberrazione: c_j.

Termini attivamente processati dai gradienti (ray deviation):

- Defocus (Z_4)
- Astigmatismo (Z_5, Z_6)
- Coma (Z_7, Z_8)

– Aberrazione Sferica Primaria (Z_{11})

Formulazione matematica:

$$W(u,v) = \sum_j c_j Z_j(u,v) \tag{3.3}$$

5. Costruzione della Pupil Function

Definite le aberrazioni, l'algoritmo procede con la costruzione della cosiddetta pupil function. In termini fisici, questa entità matematica descrive lo stato esatto dell'onda luminosa nell'istante in cui attraversa la lente, accorpando in un'unica mappa complessa due informazioni cruciali: la trasmissione fisica e la distorsione ottica. Da un lato, la forma geometrica del diaframma stabilisce quanta luce riesce a passare; dall'altro, i polinomi di Zernike determinano come i fronti d'onda vengono alterati (ritardati o anticipati) a causa dei difetti della lente.

Componenti della funzione:

Ampiezza — $A(u,v)$: Rappresenta l'apertura fisica del sistema. Funziona come una maschera binaria dove la luce passa o viene bloccata. I margini di questa maschera vengono intenzionalmente sfumati (tramite filtri di smorzamento) per evitare che un confine digitale troppo netto generi artefatti a onde durante i successivi calcoli in frequenza.

Fase — $e^{i\,2\pi\,W(u,v)}$: Converte l'errore del fronte d'onda calcolato in precedenza (W) in un vero e proprio sfasamento, introducendo matematicamente la natura ondulatoria della luce all'interno del calcolo.

Formulazione matematica:

$$P(u,v) = A(u,v)\, e^{i\,2\pi\,W(u,v)} \tag{3.4}$$

6. Sintesi della Risposta Impulsiva (PSF)

L'ultima fase del processo offline consiste nel calcolo dell'impronta bidimensionale finale dell'intensità luminosa, espressa come il quadrato del modulo della Trasformata di Fourier della pupil function. Questo calcolo produce la distribuzione energetica concreta che andrà a popolare lo stack volumetrico, costituendo la mappa su cui il

sensore baserà la stocastica o la correzione geometrica dei raggi durante la fase di rendering.

Operatori computazionali:

Operatore Trasformata Rapida di Fourier: $\mathcal{F}$.

Formulazione matematica:

$$\mathrm{PSF}(x, y) = |\mathcal{F}\{P(u, v)\}|^2 \tag{3.5}$$

Il flusso fin qui descritto definisce il modello ottico diretto. Questa architettura matematica produce le mappe energetiche di base che, venendo scalate spazialmente proprio attraverso l'apertura angolare massima (θ_{psf}) definita precedentemente, traducono le intensità adimensionali di Fourier nei limiti geometrici concreti del frustum per il motore di rendering. Oltre a generare questo asset finale per Mitsuba, il modello costituisce il cuore operativo della fase di calibrazione: l'algoritmo di recupero di fase, descritto nella sezione successiva, sfrutterà queste stesse equazioni per generare iterativamente migliaia di PSF di prova, al fine di confrontarle con i dati di riferimento e dedurre le aberrazioni latenti da iniettare nel sensore.

3.4.2 Recupero di Fase

Come introdotto all'inizio della sezione, l'algoritmo di calibrazione opera su sequenze di target sintetici generati dal Forward Model. Per trasformare queste simulazioni ideali in misurazioni empiriche verosimili su cui testare il sistema, ai dati viene deliberatamente iniettato un rumore gaussiano ($\mathcal{N}(0, \sigma^2)$), il quale simula il rumore termico e di lettura tipico di un sensore CMOS reale. Disporre di un ground truth rigoroso, ma fisicamente degradato, si rivela un passaggio essenziale per validare la robustezza dell'ottimizzatore.

L'estrazione dei coefficienti di Zernike latenti, disponendo esclusivamente delle misurazioni di intensità (le PSF), costituisce un classico problema di phase retrieval [21]. Poiché si tratta di un problema inverso mal posto, in cui la minimizzazione dell'errore è ostacolata da un dominio non convesso e ricco di minimi locali, l'architettura risolutiva è stata divisa in due stadi.

1. Recupero di Fase Globale

L'algoritmo estrae tre campioni chiave dallo stack di immagini (focus negativo, fuoco e focus positivo) al fine di stimare le aberrazioni statiche del sistema. Questa

valutazione si rende necessaria per superare l'ambiguità intrinseca del problema inverso: se si operasse singolarmente sui vari livelli di profondità, l'ottimizzatore potrebbe compensare l'errore trovando combinazioni di coefficienti sempre diverse ma capaci di generare PSF morfologicamente identiche. Tale instabilità numerica porterebbe a fluttuazioni irrealistiche delle aberrazioni lungo l'asse ottico, distruggendo la coerenza volumetrica della simulazione. Ancorando la ricerca a tre campioni spaziati, il sistema forza l'algoritmo a identificare un unico set coerente di aberrazioni statiche.

Per evitare la stagnazione in questo complesso spazio delle soluzioni, viene impiegato un approccio multi-start: l'ottimizzatore viene inizializzato in parallelo partendo da una griglia di guess iniziali differenti, sparpagliate lungo i domini di defocus, coma e sferica. Dal punto di vista implementativo, la risoluzione è affidata alla libreria open-source SciPy [22], sfruttando la funzione `scipy.optimize.minimize`. L'algoritmo di minimizzazione prescelto è l'L-BFGS-B (Limited-memory Broyden–Fletcher–Goldfarb–Shanno con bounds): la presenza di vincoli di confine (bounds) si rivela essenziale per impedire che i polinomi divergano verso valori fisicamente irrealistici durante la ricerca.

2. Recupero di Fase Locale

Identificato il set di parametri globali ottimali, il sistema procede raffinando i coefficienti iterativamente per ogni singolo piano acquisito. La loss function utilizzata dall'ottimizzatore non valuta la differenza diretta tra le intensità lineari, bensì l'Errore Quadratico Medio tra le radici quadrate delle intensità normalizzate:

$$E = \sum_{x,y} \left(\sqrt{\max(I_{\text{sim}}(x,y), 0)} - \sqrt{\max(I_{\text{target}}(x,y), 0)} \right)^2 \tag{3.6}$$

Questa compressione matematica del range dinamico si rivela cruciale: impedisce al denso picco di energia centrale della PSF di dominare incontrastato il calcolo del gradiente, costringendo di fatto l'ottimizzatore a valutare e fittare anche i deboli anelli periferici generati dall'aberrazione sferica o dal coma.

3.4.3 Sintesi del Frustum Aberrato

Una volta ottenuti i coefficienti best-fit per l'intero stack, l'approccio si distacca radicalmente dai metodi classici di post-elaborazione in spazio immagine. Invece di campionare l'output raster della FFT, il sistema ricostruisce analiticamente la deviazione vettoriale dei raggi a partire dalla geometria del fronte d'onda.

Sfruttando la linearità dell'operatore derivata, il codice calcola il gradiente spaziale differenziando direttamente i polinomi di Zernike. A livello pratico, questo permette di ricavare la deviazione dei raggi in modo analitico ed esatto: ad esempio, la componente orizzontale del gradiente causata dall'astigmatismo verticale (Z_6) risulta direttamente proporzionale alla coordinata u sulla pupilla.

Questo campo vettoriale, che descrive la pendenza locale dell'onda luminosa, viene poi scalato secondo i parametri fisici del sistema per calcolare l'effettivo spostamento trasversale del raggio. Per estendere questo calcolo alle tre dimensioni e ricostruire l'intero frustum, il sistema valuta la mappa vettoriale a diverse profondità focali. Ciò si ottiene campionando iterativamente un set discreto di coefficienti per il solo termine di defocus (Z_4). Di conseguenza, per ogni coordinata di partenza (u, v) sull'apertura, l'algoritmo calcola una sequenza di scostamenti discreti: unendo questi punti lungo l'asse ottico, si traccia l'effettiva traiettoria spaziale della luce.

Infine, per garantire una valutazione continua ed efficiente in fase di rendering, su ciascun set di coordinate spaziali viene applicato un algoritmo di regressione lineare ai minimi quadrati. Il risultato è un modello parametrico estremamente leggero: per ogni campione della pupilla, vengono immagazzinati unicamente una pendenza (slope, m) e un'intercetta (intercept, q). Questo dataset sintetizza l'intera geometria del cono di luce aberrato, fornendo al sensore l'equazione esatta per deviare la traiettoria di qualsiasi raggio primario senza dover ricorrere a costosi calcoli ondulatori a runtime.

3.5 Fase 2: Rendering Ibrido

L'ultimo tassello metodologico prevede l'iniezione del frustum vettoriale precalcolato all'interno del processo Monte Carlo in Mitsuba 3, tramite un'architettura personalizzata denominata Hybrid FFT Sensor.

Invece di ricorrere al classico importance sampling basato su CDF per alterare la densità dei raggi in partenza, l'approccio proposto implementa un modello che opera come segue:

1. **Campionamento Uniforme (Bassa Varianza):** Per calcolare l'illuminazione di un pixel sul piano immagine, i raggi primari vengono campionati stocasticamente, ma in modo del tutto uniforme, sulla superficie circolare dell'apertura ottica. Questo garantisce che il peso statistico (weight) del raggio rimanga unitario, stabilizzando fortemente l'equazione di rendering.

2. **Deformazione Angolare Dinamica (Warping):** L'alterazione dell'energia viene demandata interamente al dominio spaziale. Per ogni raggio generato, il sensore interroga il modello parametrico del frustum (caricato direttamente in memoria GPU per azzerare i colli di bottiglia). Anziché limitarsi a interrogare una singola mappa bidimensionale, il sistema immagazzina i coefficienti lineari (slope e intercept) che definiscono analiticamente l'evoluzione volumetrica del fascio di luce nello spazio tridimensionale. La direzione parassiale ideale $\mathbf{d}_{\text{ideal}}$ viene alterata sommando la deviazione trasversale, calcolata istantaneamente valutando l'equazione della retta alla profondità desiderata:

$$\Delta(z) = \text{slope} \cdot z + \text{intercept} \tag{3.7}$$

dove z rappresenta lo scostamento longitudinale rispetto al piano focale.

Questa soluzione ha il vantaggio di conservare l'energia e di fare emergere in modo organico il fenomeno ottico nella sua componente volumetrica. La complessa distribuzione luminosa del *bokeh* aberrato non viene sovrapposta artificialmente come un filtro bidimensionale, ma emerge naturalmente per via puramente geometrica come una vera e propria caustica ottica: l'accumulo fisico delle traiettorie dei raggi deviati dalle loro corrette formule polinomiali produce i netti picchi di luminosità e le delicate sfumature sui bordi del Circolo di Confusione nell'immagine renderizzata.

3.6 Pipeline di rendering

Per validare empiricamente l'architettura del sensore personalizzato e gestire in modo scalabile la mole di simulazioni richieste, è stata sviluppata un'infrastruttura software modulare scritta in Python. Questo strato applicativo ha il compito di orchestrare la comunicazione tra i dati pre-calcolati (le PSF), la geometria della scena e il motore di rendering Mitsuba 3, automatizzando il processo dalla configurazione della scena fino al denoising dell'output finale.

L'infrastruttura si divide in tre moduli principali: l'interprete dei file di scena (`blender_XML_reader.py`), il gestore delle interfacce utente (`render_options.py`) e l'eseguibile principale per il controllo del flusso (`main.py`).

3.6.1 Lettura dell'XML

Le scene tridimensionali utilizzate come target visivo vengono modellate in Blender ed esportate nativamente nel formato XML supportato da Mitsuba. Tuttavia, il formato di esportazione standard genera file contenenti definizioni fisiche adatte esclusivamente ai sensori integrati nel framework. Per consentire l'utilizzo del

sensore ibrido senza costringere l'utente a lunghe modifiche manuali, è stato implementato un parser dedicato.

Lo script `blender_XML_reader.py` analizza e manipola l'albero XML (ElementTree) prima di inizializzare il rendering. Le operazioni principali eseguite includono:

- **Rimozione dei conflitti ottici:** Tag incompatibili o hardcoded originariamente da Blender (come vecchie implementazioni della lunghezza focale o distanze di clip) vengono disinnescati.
- **Iniezione parametrica:** Lo script legge il file `.npz` contenente le direttive spaziali del frustum vettoriale (pre-calcolato nella Fase 1 e inietta dinamicamente la mappa della PSF, le coordinate e le dimensioni fisiche dell'apertura direttamente all'interno della configurazione del sensore.
- **Gestione dinamica delle risorse (Resolver):** Il parser effettua una validazione preventiva dei percorsi relativi di texture e mesh geometriche, aggiungendole dinamicamente al `file_resolver` del thread di Mitsuba. Questo assicura che il motore non fallisca il caricamento di risorse esterne, un problema comune nell'importazione di XML complessi.

Al termine dell'elaborazione, il sistema genera un file temporaneo validato che Mitsuba può caricare e compilare in totale sicurezza, garantendo che le informazioni relative all'estensione del *bokeh* vengano applicate come Single Source of Truth.

3.6.2 Il Modulo Main

Lo script `main.py` funge da orchestratore della simulazione, sfruttando l'interfaccia a riga di comando per raccogliere le specifiche dall'utente (scelta della scena, modalità di campionamento e numero di Sample Per Pixel).

Il motore viene inizializzato sfruttando l'architettura Just-In-Time (JIT) di Mitsuba 3. Per garantire l'accelerazione hardware necessaria al calcolo intensivo del tracciamento vettoriale del *bokeh*, il sistema è configurato per compilare sulla variante `llvm_ad_rgb`, che sfrutta il framework LLVM per parallelizzare efficientemente l'esecuzione sulla CPU e garantire la compatibilità con la generazione stocastica progressiva.

Una delle implementazioni architetturali più rilevanti risiede nel **Rendering Progressivo in Chunk**. Poiché simulare un *bokeh* fotorealistico basato su traiettorie aberrate richiede solitamente alti valori di Samples Per Pixel (SPP) per convergere

eliminando la varianza, processare l'intera immagine in un'unica iterazione potrebbe causare colli di bottiglia o saturare la cache della memoria.

Per ovviare a ciò, il processo di accumulo (`progressive_render`) è stato segmentato:

1. Il target totale di SPP viene frammentato in chunk gestibili (es. 256 SPP alla volta).
2. Per ogni iterazione, l'integratore di percorso (`path integrator`) simula il lotto di campioni sfruttando il sensore ibrido.
3. L'immagine parziale viene miscelata dinamicamente con i risultati delle iterazioni precedenti, forzando la valutazione del grafo computazionale (`dr.eval()`) e sincronizzando il thread per svuotare la memoria temporanea.

3.6.3 Post-elaborazione

L'ultimo passaggio di questa infrastruttura è relativo alla pulizia e l'esportazione dell'immagine renderizzata. Al termine del tracciamento progressivo, qualora l'output presenti ancora del rumore stocastico residuo tipico dell'approccio Monte Carlo, il risultato viene convertito in un array continuo e passato al motore Intel Open Image Denoise (OIDN) [23]. Il file finale viene quindi esportato parallelamente in formato lineare a 32-bit (`exr`), essenziale per preservare l'intera gamma dinamica del calcolo fisico, e in formato sRGB (`png`) per l'ispezione visiva immediata.

Con quest'ultima fase di salvataggio si chiude la catena computazionale del sistema. L'intera metodologia fin qui descritta, dalla discretizzazione del fronte d'onda, al recupero di fase, fino all'iniezione nel motore di tracciamento, dimostra come sia possibile colmare il divario tra l'ottica ondulatoria e la computer grafica, ottenendo una simulazione volumetrica delle aberrazioni che risulta al contempo fisicamente coerente, del tutto automatizzata e computazionalmente scalabile.

Il codice sorgente alla base dell'architettura proposta, comprensivo degli script di automazione e dell'implementazione del sensore personalizzato, è liberamente consultabile presso il repository GitHub dedicato[1].

[1] `https://github.com/GabrielePiergiovanni/Mitsuba_custom_sampler`

Capitolo 4

Validazione Sperimentale

Per poter valutare l'efficacia del sensore basato sul frustum aberrato, la fase di validazione è stata divisa in due macro-aree: una analitica e una percettiva. I test analitici hanno coinvolto sia lo studio della risposta al gradino su una mira ISO (International Standard Organization) standard, sia l'estrazione della Modulation Transfer Function (MTF) globale da un set di 9 scene complesse renderizzate con una specifica combinazione di coefficienti di Zernike. Questo stesso set di immagini fotorealistiche è stato successivamente impiegato per la validazione percettiva tramite la somministrazione di questionari mirati.

4.1 Test Analitico

Una parte dei test oggettivi, come menzionato poco sopra, ha coinvolto l'utilizzo di una mira ISO [24]. Le mire ISO sono dei bersagli ottici standardizzati, solitamente usati per la validazione delle prestazioni di lenti e sistemi di imaging reali attraverso l'analisi della loro Spatial Frequency Response (SFR) ovvero la risposta spaziale in frequenza. La mira impiegata è la ISO 12233, di cui viene analizzato in particolare l'elemento *slanted-edge*. Si tratta di un rettangolo inclinato di 5.71 gradi rispetto alla griglia dei pixel, una geometria che consente di estrarre la SFR e la MTF. Tale analisi permette di calcolare la risposta al gradino, descrivendo il comportamento del sistema in presenza di una brusca transizione di contrasto spaziale (dal nero profondo al bianco puro). L'inclinazione specifica di questo bordo non è casuale: essendo disallineato rispetto alla griglia ortogonale dei pixel del sensore, il bordo interseca ogni riga di pixel in una posizione sub-pixel leggermente diversa. Questo consente di effettuare un sovra-campionamento (*oversampling*) della transizione, costruendo una Edge Spread Function (ESF) con precisione sub-pixel e aggirando i limiti di *aliasing* imposti dalla risoluzione del sensore.

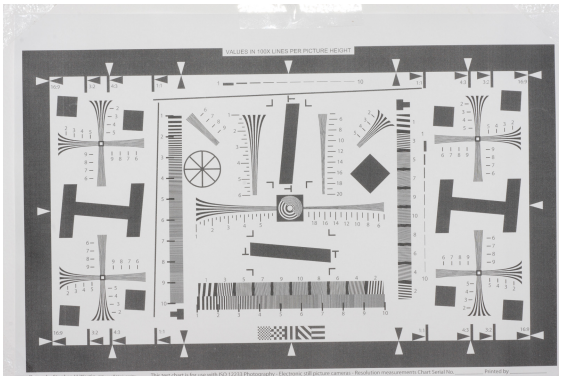


Figura 4.1: Fotografia di una mira ottica ISO 12233 [25].

Calcolando la derivata spaziale dell'ESF si ottiene la Line Spread Function (LSF), la cui Trasformata di Fourier monodimensionale genera infine la curva MTF. Nel contesto di questo lavoro, renderizzare la mira ISO all'interno di Mitsuba 3 ha permesso di trattare il motore di sintesi al pari di un sistema ottico reale: valutando la MTF sui bordi inclinati dell'immagine sintetizzata, è stato possibile misurare matematicamente la caduta di risoluzione, la diffusione del blur e il degrado del micro-contrasto indotti dall'iniezione del frustum aberrato, validando oggettivamente l'accuratezza fisica del modello vettoriale proposto.

Mitsuba 3.0 non supporta nativamente file vettoriali in formato `.svg`. Importare la mira come immagine raster `.png` avrebbe introdotto artefatti di campionamento sui bordi netti: a differenza di un formato vettoriale, che descrive la geometria in termini di primitive matematiche indipendenti dalla risoluzione, un formato raster codifica l'immagine come una griglia finita di valori discreti, degradando inevitabilmente per aliasing le transizioni di contrasto che costituiscono l'elemento critico per il calcolo della ESF e della MTF. La pipeline adottata è stata pertanto quella di importare il file vettoriale in Blender (versione 4.5.5 [26]). Sul software di grafica 3D il flusso di lavoro si è articolato come segue: è stato innanzitutto risolto il problema di *Z-fighting* che si è presentato, essendo il file in un formato `.svg` diviso in livelli posizionati nello stesso piano dello spazio. Dopo aver spostato i blocchi dallo sfondo di qualche millimetro, per evitare la proiezione di ombre sul bianco da parte dei materiali, è stato applicato a quest'ultimo un materiale emissivo. Da Blender l'esportazione è stata effettuata tramite l'addon mitsuba-blender exporter [27] come file `xml`. Il file, assieme alle meshes è stato poi importato all'interno dell'ambiente di lavoro per passare allo step successivo.

Al fine di garantire tempi di calcolo compatibili con la mole di dati generata, tutte le simulazioni descritte in questo capitolo sono state eseguite sfruttando l'accelerazione hardware fornita da una CPU AMD Ryzen 7 5700X a 8 core (3.40

GHz), avvalendosi del framework LLVM per la parallelizzazione intensiva sui core a disposizione, e supportata da 32 GB di memoria RAM.

4.1.1 Test analitico: combinazioni di aberrazioni

Tabella 4.1: Supporto delle singole aberrazioni ottiche in relazione ai diversi modelli di rendering.

Tipo di test	**Astigmatismo Obliquo** ($j = 5$)	**Astigmatismo Verticale** ($j = 6$)	**Coma Orizzontale** ($j = 7$)	**Coma Verticale** ($j = 8$)	**Sferica** ($j = 11$)
Astigmatismo	0.25	0.25	0	0	0
Coma	0	0	0.25	0.25	0
Sferica positiva	0	0	0	0	0.75
Sferica negativa	0	0	0	0	-0.75

Nello specifico della mira ISO, il rendering è stato calcolato a 4096 *sample per pixel* (SPP) avvalendosi dell'algoritmo di *Path Tracing* supportato da campionamento *multijitter*, garantendo così la totale assenza di rumore stocastico sui bordi analizzati. A seguire, si riportano i risultati visivi ottenuti, mentre i grafici della risposta al gradino verranno dettagliati nel prossimo capitolo:

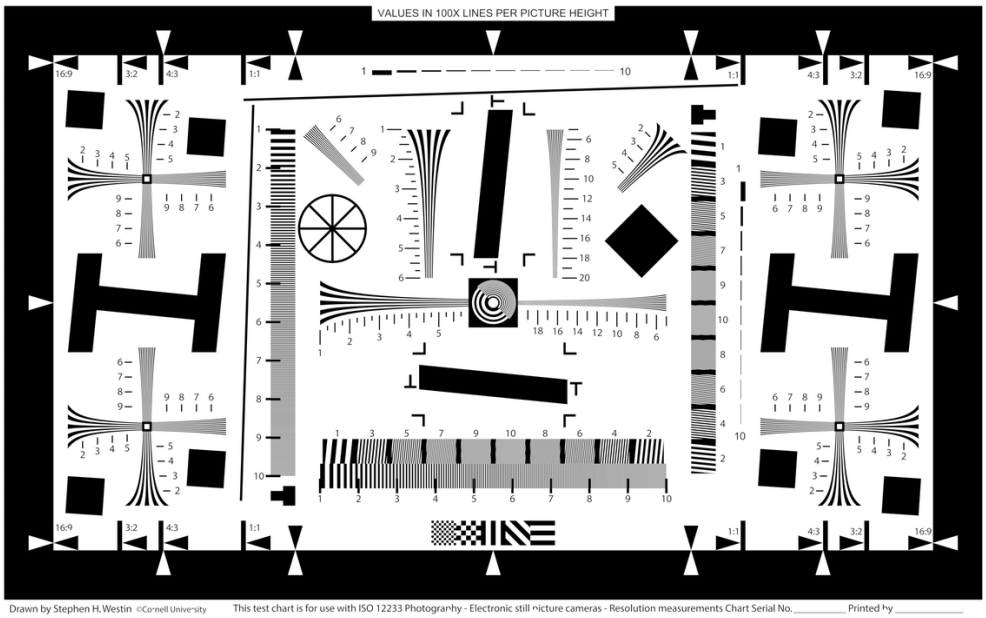


Figura 4.2: Rendering ISO 12233: assenza di aberrazioni

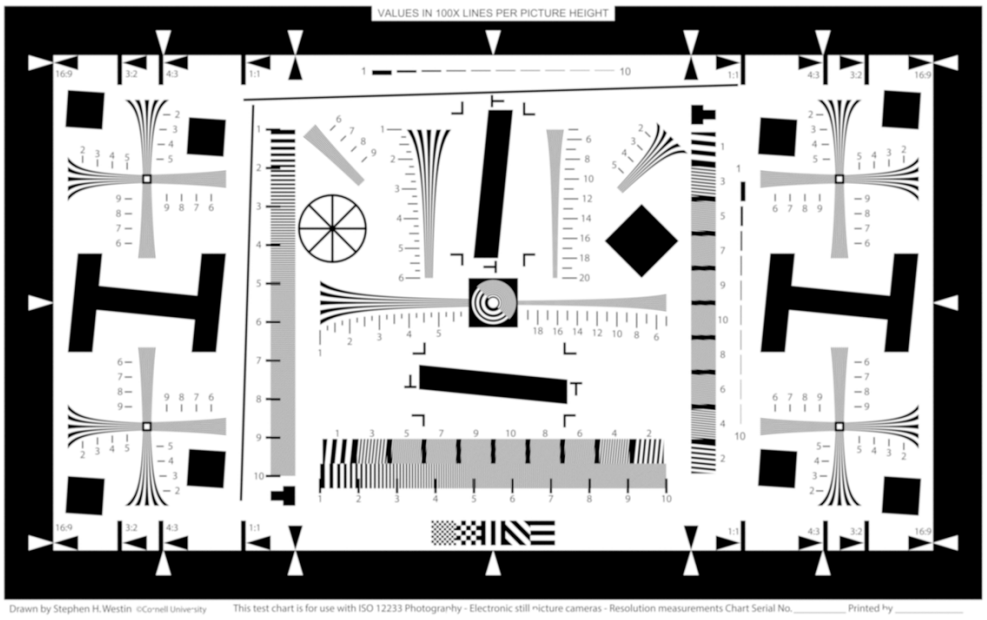


Figura 4.3: Rendering ISO 12233: astigmatismo

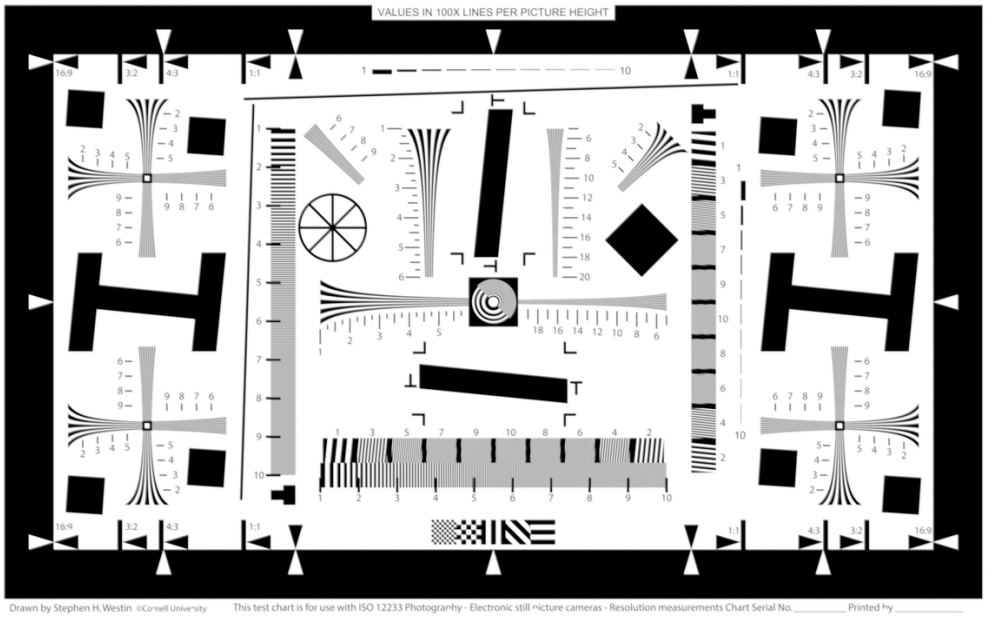


Figura 4.4: Rendering ISO 12233: coma

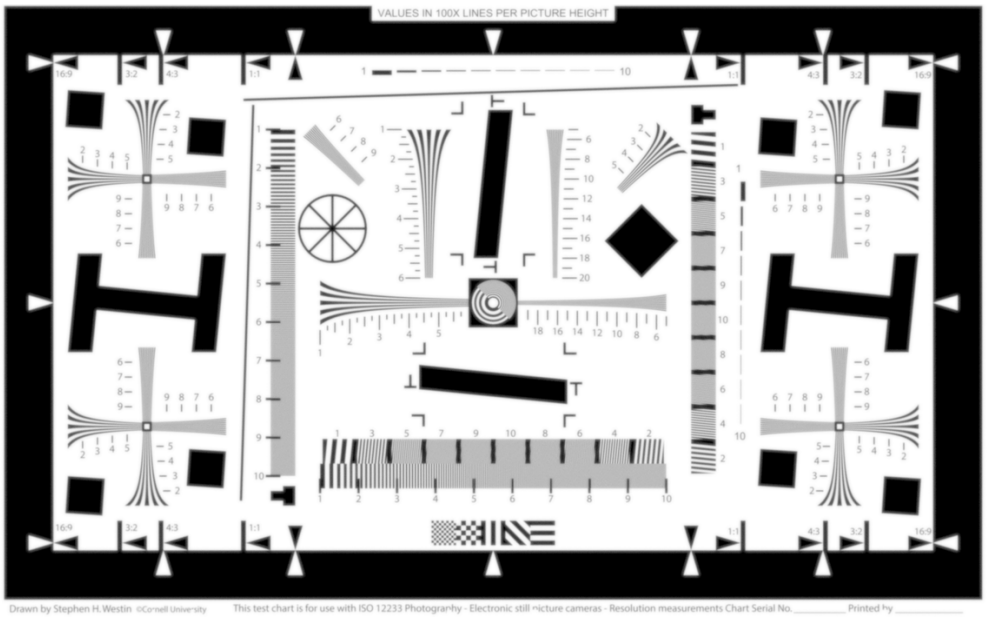


Figura 4.5: Rendering ISO 12233: aberrazione sferica positiva

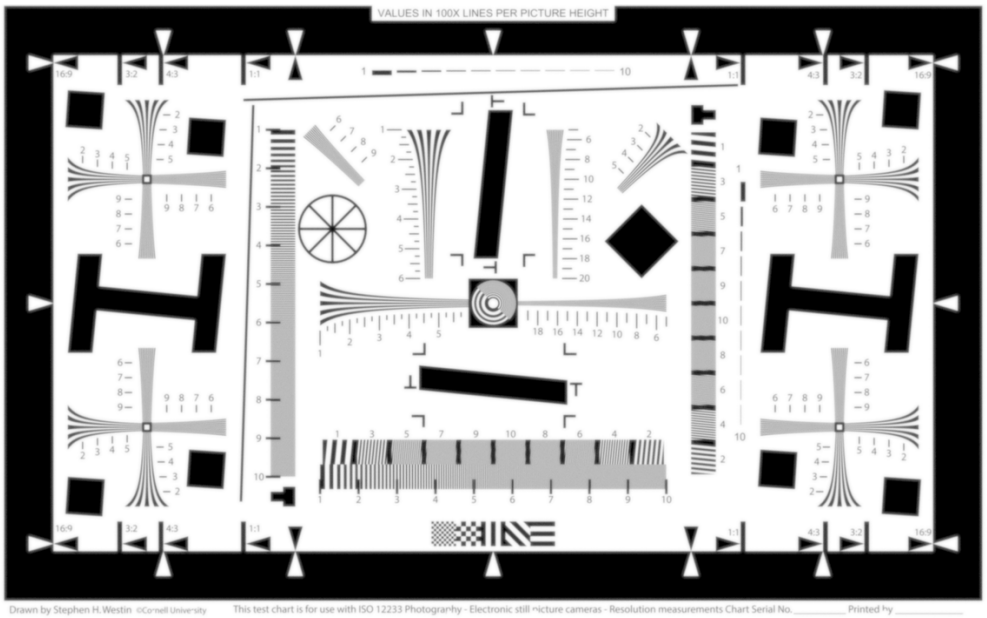


Figura 4.6: Rendering ISO 12233: aberrazione sferica negativa

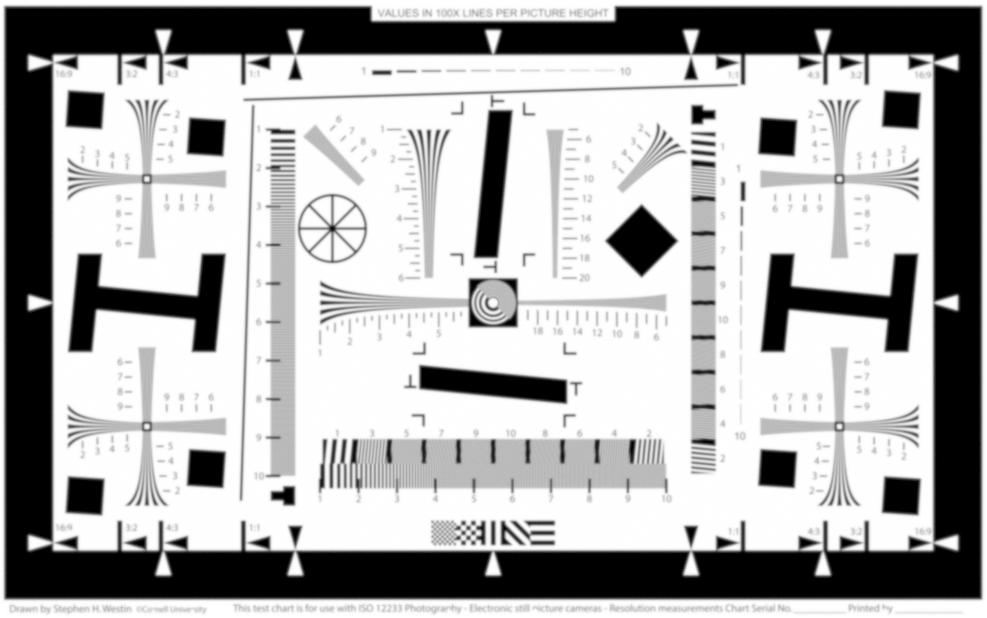


Figura 4.7: Rendering ISO 12233: aberrazione sferica positiva (focus -1)

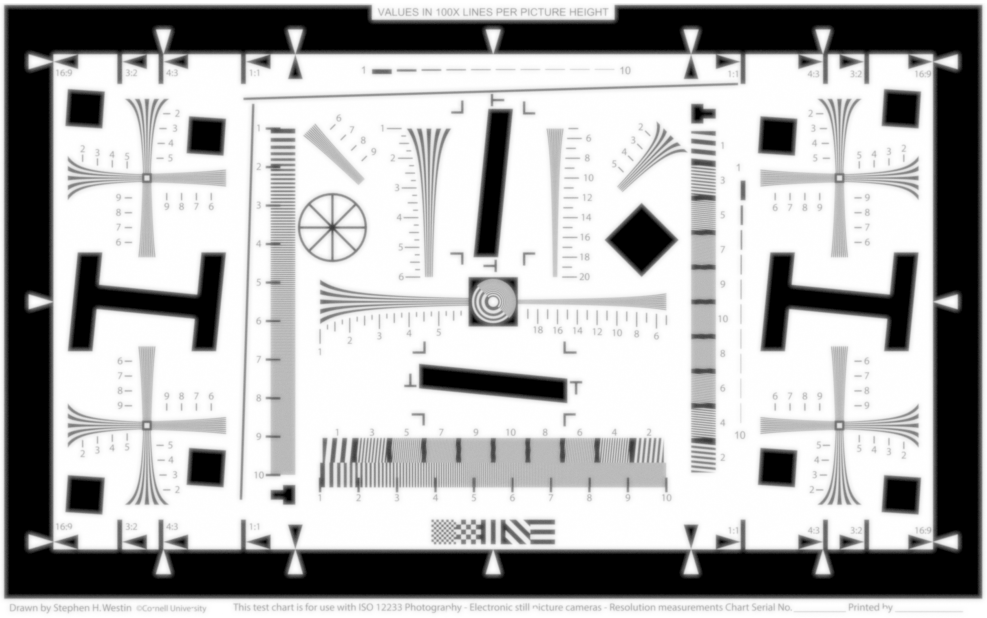


Figura 4.8: Rendering ISO 12233: aberrazione sferica negativa (focus -1)

4.1.2 Test Analitico: MTF Globale su Scene Fotorealistiche

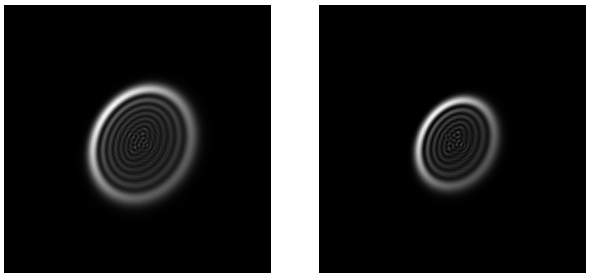
La seconda parte dei test oggettivi si ricollega direttamente alla fase di validazione percettiva descritta nella sezione successiva: le scene utilizzate per l'estrazione della MTF globale, infatti, costituiscono il medesimo dataset visivo sottoposto al campione di utenti tramite i questionari.

Le scene, presentate di seguito assieme alla Point Spread Function generata per costruire il frustum aberrato, provengono in parte da archivi consolidati in ambito accademico, come il database di Benedikt Bitterli [28], e in parte dalla libreria Open Source BlenderKit [29]. Analogamente a quanto eseguito per le mire ISO, è stato necessario implementare una pipeline di conversione dedicata per asset, texture e materiali, affinché i dati geometrici e radiometrici potessero essere correttamente esportati in formato `xml` e interpretati dal motore di rendering personalizzato.

Al fine di misurare l'impatto visivo delle aberrazioni sulla percezione dei dettagli fini, è stato generato un set di 9 scene fotorealistiche. Per queste simulazioni si è scelto di non replicare l'impronta di un singolo obiettivo commerciale, bensì di definire un profilo ottico empirico concepito come un vero e proprio *stress test* per l'architettura di rendering. I seguenti pesi dei coefficienti di Zernike sono stati calibrati con il preciso intento di attivare simultaneamente tutte le aberrazioni primarie supportate, generando un fronte d'onda complesso e asimmetrico la cui entità fosse chiaramente percettibile dall'osservatore, ma tale da non compromettere la leggibilità strutturale della scena:

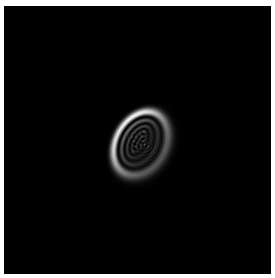
- Z_5 (Astigmatismo obliquo): 0.25
- Z_6 (Astigmatismo verticale): 0.15
- Z_7 (Coma verticale): 0.10
- Z_8 (Coma orizzontale): 0.10
- Z_{11} (Aberrazione sferica): 0.15

Tali valori sono stati mantenuti costanti, facendo variare unicamente il coefficiente di defocus (Z_4) attraverso una discretizzazione del dominio della profondità compresa tra -2.5 e 2.5, al fine di mappare l'evoluzione morfologica del *bokeh* in base alla distanza dal piano focale.

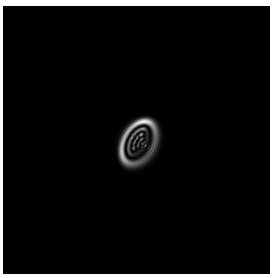


$Z_4 = -2.5$

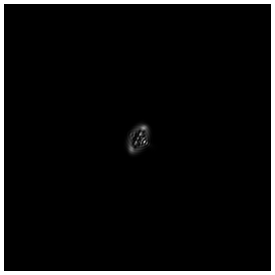
$Z_4 = -2.0$



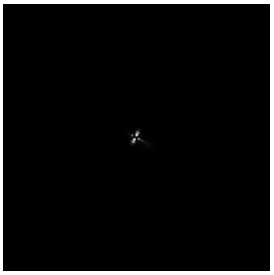
$Z_4 = -1.5$



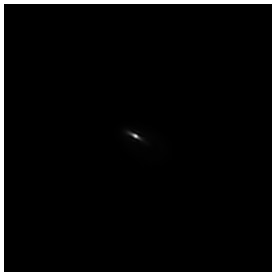
$Z_4 = -1.0$



$Z_4 = -0.5$



$Z_4 = 0.0$



$Z_4 = 0.5$

Figura 4.9: Evoluzione della PSF al variare del defocus Z_4

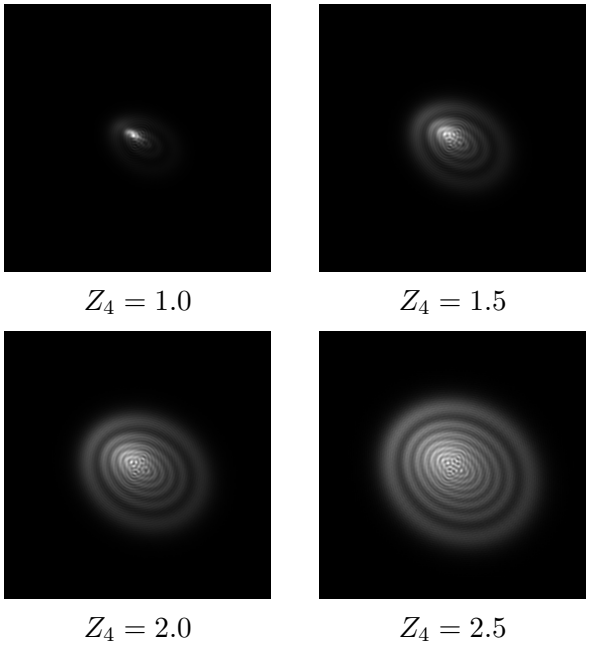


Figura 4.9: Evoluzione della PSF al variare del defocus Z_4.

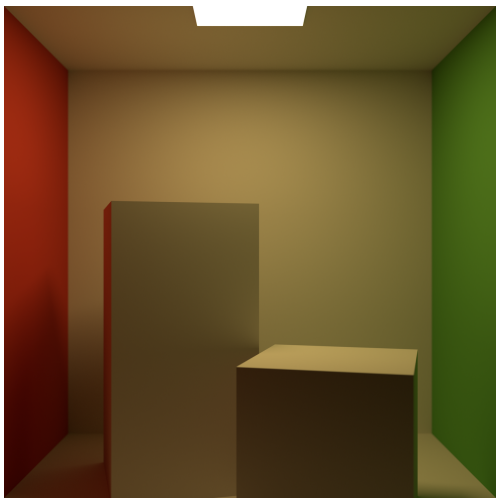
Metodologia di Estrazione della MTF Efficace (PSD)

A differenza dell'analisi condotta sulla mira ISO 12233, progettata per valutare la risposta del sensore su un piano bidimensionale ideale, la misurazione del micro-contrasto su scene fotorealistiche richiede un approccio in grado di tenere conto della natura volumetrica dell'illuminazione, delle caustiche e della complessa geometria delle mesh. Per questo motivo, l'analisi matematica di questi 9 render non si basa sull'identificazione di un singolo gradino di contrasto, ma sull'elaborazione dell'intera Densità Spettrale di Potenza (*Power Spectral Density*, PSD) dell'immagine.

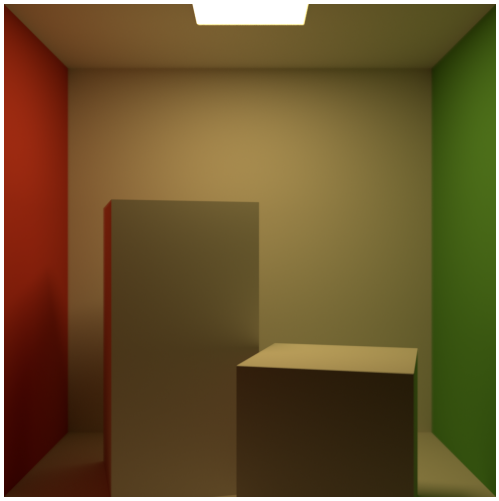
Confrontando l'energia spettrale delle frequenze dell'immagine di riferimento (generata con il modello *Thin Lens*) con quella dell'immagine in esame (generata dal sensore ibrido), è possibile estrapolare il differenziale spettrografico. Questo metodo di stima analitica permette di ricavare una MTF globale del sistema. In questo modo si quantifica oggettivamente la perdita di risoluzione e la dispersione energetica alle alte frequenze spaziali causate unicamente dalle complesse traiettorie vettoriali del *bokeh* aberrato, e non da semplici sfocature gaussiane.

Tabella 4.2: Caratteristiche: Scena 01 - Cornell box [30]

Caratteristiche	Scena standard utilizzata per valutazione di sistemi di rendering. Utile nel nostro contesto per osservare il comportamento in presenza di superfici nette.		
Tempo Thinlens	624.62 s	**Tempo Hybrid FFT**	625.57 s
Integrator	Path Tracer	**Samples (SPP)**	1024
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	2.76 m	**Apertura**	50 mm



(a) Thinlens (Denoised)

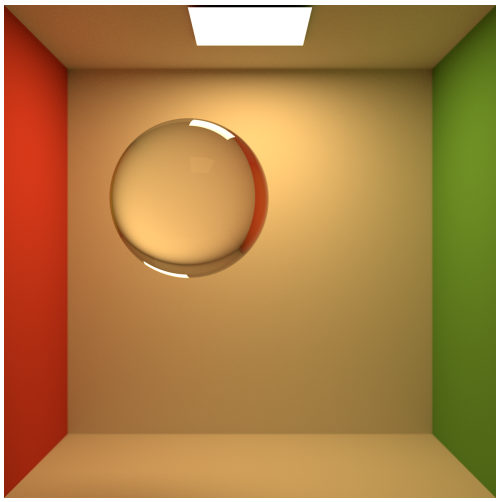


(b) Aberrated (Denoised)

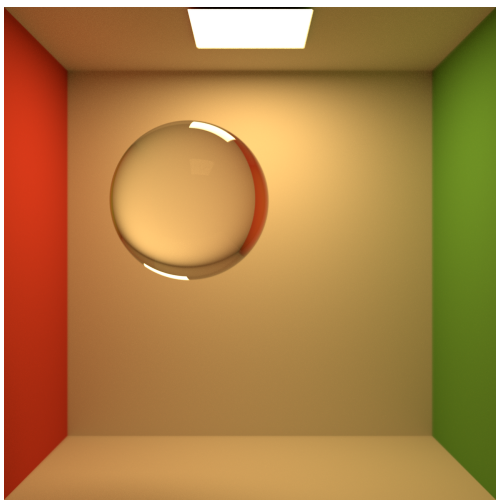
Figura 4.10: Confronto visivo: Scena 01 - Cornell box.

Tabella 4.3: Caratteristiche: Scena 02 - Cornell box glass [31]

Caratteristiche	Selezionata per osservare il comportamento dell'algoritmo in presenza di materiali quali il vetro.		
Tempo Thinlens	745.47 s	**Tempo Hybrid FFT**	759.53 s
Integrator	Path Tracer	**Samples (SPP)**	1024
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	2.76 m	**Apertura**	50 mm



(a) Thinlens

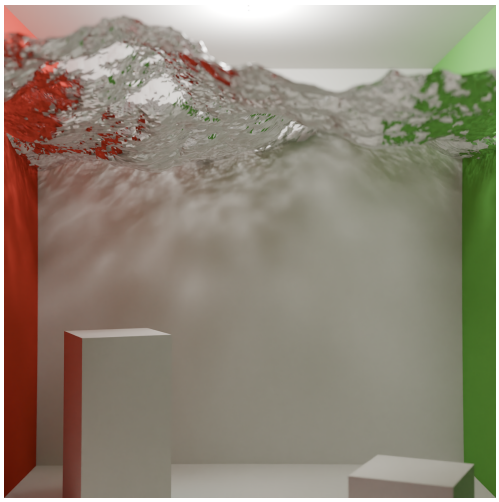


(b) Aberrated

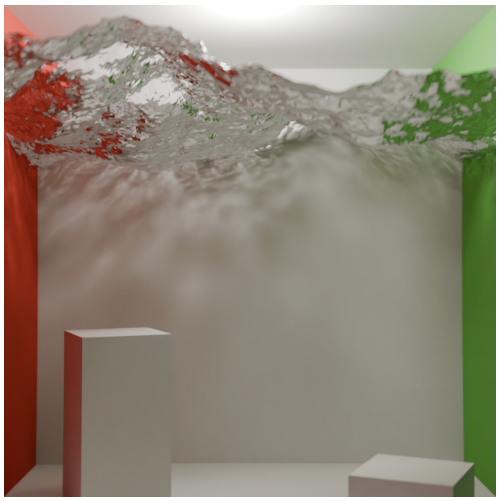
Figura 4.11: Confronto visivo: Scena 02 - Cornell box glass.

Tabella 4.4: Caratteristiche: Scena 03 - Cornell box water [32]

Caratteristiche	Scena selezionata per osservare il comportamento del frustum in presenza di caustiche.		
Tempo Thinlens	939.79 s	**Tempo Hybrid FFT**	968.90 s
Integrator	Path Tracer	**Samples (SPP)**	1024
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	2.76 m	**Apertura**	50 mm



(a) Thinlens



(b) Aberrated

Figura 4.12: Confronto visivo: Scena 03 - Cornell box water.

Tabella 4.5: Caratteristiche: Scena 04 - Lego [33]

Caratteristiche	In questo caso la scena è stata selezionata per valutare il comportamento dell'algoritmo in presenza di materiali non organici molto dettagliati.		
Tempo Thinlens	190.25 s	**Tempo Hybrid FFT**	241.20 s
Integrator	Path Tracer	**Samples (SPP)**	1024
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	1.625 m	**Apertura**	50 mm

(a) Thinlens

(b) Aberrated

Figura 4.13: Confronto visivo: Scena 04 - Lego.

Tabella 4.6: Caratteristiche: Scena 05 - Detective room [34]

Caratteristiche	La scena è stata selezionata per via di una varietà di texture interessanti, in particolare testi che presentano un contrasto netto tra bianco e nero.		
Tempo Thinlens	929.66 s	**Tempo Hybrid FFT**	988.42 s
Integrator	Path Tracer	**Samples (SPP)**	256
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	6 m	**Apertura**	50 mm

(a) Thinlens

(b) Aberrated

Figura 4.14: Confronto visivo: Scena 05 - Detective room.

Tabella 4.7: Caratteristiche: Scena 06 - Perfume [35]

Caratteristiche	In questo caso la selezione è ricaduta sull'immagine per poter osservare il rendering dei riflessi e materiali trasmissivi.		
Tempo Thinlens	728.36 s	**Tempo Hybrid FFT**	745.73 s
Integrator	Path Tracer	**Samples (SPP)**	1024
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	8.5 m	**Apertura**	50 mm

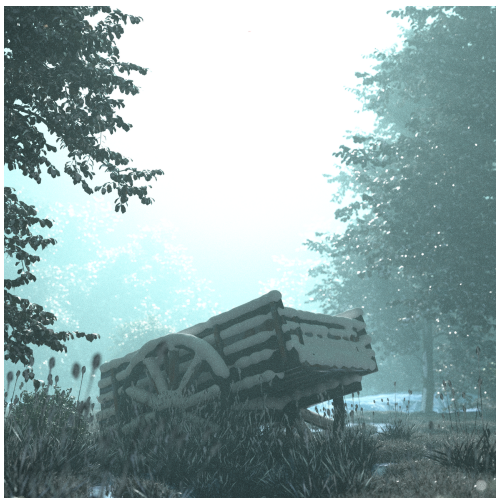
(a) Thinlens

(b) Aberrated

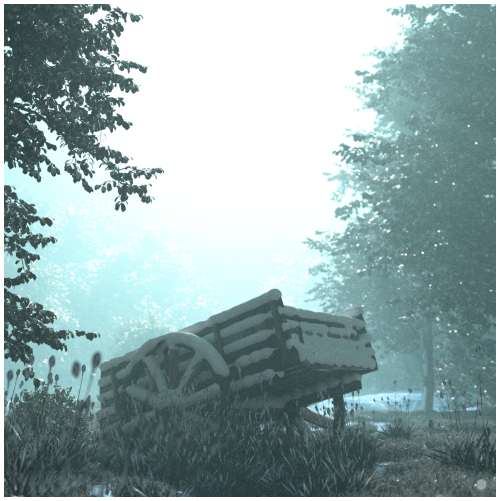
Figura 4.15: Confronto visivo: Scena 06 - Perfume.

Tabella 4.8: Caratteristiche: Scena 07 - Winter Forest [36]

Caratteristiche	Scena renderizzata attraverso un integratore volumetrico, in modo da poter osservare il comportamento del sensore custom in presenza di un volume permeante la scena.		
Tempo Thinlens	3516.33 s	**Tempo Hybrid FFT**	3563.20 s
Integrator	Volpathmis	**Samples (SPP)**	512
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	10 m	**Apertura**	50 mm



(a) Thinlens



(b) Aberrated

Figura 4.16: Confronto visivo: Scena 07 - Winter forest.

Tabella 4.9: Caratteristiche: Scena 08 - Wooden Pier [37]

Caratteristiche	Scena selezionata per la presenza di acqua e materiali organici complessi.		
Tempo Thinlens	1140.71 s	**Tempo Hybrid FFT**	1111.02 s
Integrator	Path tracer	**Samples (SPP)**	256
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	20 m	**Apertura**	50 mm

(a) Thinlens

(b) Aberrated

Figura 4.17: Confronto visivo: Scena 08 - Wooden pier.

Tabella 4.10: Caratteristiche: Scena 09 - Car [38]

Caratteristiche	Scena selezionata per la presenza di materiali riflettenti e metallici.		
Tempo Thinlens	830.00 s	**Tempo Hybrid FFT**	897.00 s
Integrator	Path tracer	**Samples (SPP)**	512
f-stops	1.4	**Sampler**	Multijitter
Distanza Fuoco	6 m	**Apertura**	50 mm

(a) Thinlens

(b) Aberrated

Figura 4.18: Confronto visivo: Scena 09 - Car.

4.2 Test Percettivo

Al fine di raccogliere valutazioni soggettive sull'impatto visivo delle aberrazioni, è stato realizzato un questionario somministrabile da browser, accessibile tramite una pagina GitHub Pages. La scelta di uno strumento custom al posto di soluzioni preconfezionate come Google Forms è nata dalla necessità di controllare l'ordine di presentazione degli stimoli: qualora tutte le coppie di immagini fossero proposte nella stessa sequenza e con la stessa disposizione sinistra/destra a ogni compilazione, il partecipante potrebbe inferire un pattern sistematico, compromettendo la validità delle risposte. Il questionario implementa pertanto due livelli di randomizzazione per sessione: l'ordine in cui le nove coppie di immagini vengono presentate è permutato casualmente, e per ciascuna coppia la disposizione sinistra/destra delle due immagini (aberrata e standard) viene anch'essa invertita con probabilità uniforme. Le risposte vengono salvate mappando la scelta dell'utente dall'etichetta visiva (Immagine 1 / Immagine 2) all'identità interna (A/B), correggendo il segno dei valori Likert nelle coppie invertite prima della trasmissione. Per prevenire doppi invii, viene generato un identificatore di sessione anonimo tramite `crypto.randomUUID()` e un segnale di dispositivo ottenuto come hash SHA-256 di attributi non personali (piattaforma, lingua, risoluzione schermo, fuso orario, numero di core logici); la combinazione dei due viene memorizzata in `localStorage` e confrontata all'avvio della pagina. Le risposte vengono trasmesse tramite richiesta POST a un endpoint Google Apps Script e archiviate in un foglio Google Sheets.

Il questionario è strutturato in tre fasi. Nella prima fase vengono raccolti i dati demografici del partecipante (età) e il livello di competenza auto-dichiarato su una scala a cinque livelli in ciascuno dei seguenti ambiti: modellazione 3D, VFX, rendering, fotografia e ottica. Nella seconda fase vengono presentate le nove coppie di immagini; per ciascuna coppia il partecipante completa una tabella Likert a cinque colonne su sei dimensioni di giudizio (qualità dei dettagli, somiglianza con una fotografia reale, qualità dell'area a fuoco, sensazione di artificialità, qualità dell'area fuori fuoco, realismo complessivo), e risponde a due domande a scelta forzata riguardanti la preferenza estetica e il fotorealismo percepito. Nella terza e ultima fase vengono raccolte impressioni globali sull'insieme delle immagini, una valutazione numerica della capacità del partecipante di distinguere i dettagli fini, e un paio di campi facoltativi per commenti liberi generali e opinioni legate alla possibilità che queste immagini potessero essere frutto di AI generativa.

Capitolo 5

Risultati e discussione

In questo capitolo vengono presentati e analizzati i risultati della sperimentazione, valutando l'efficacia del sensore ibrido proposto rispetto ai modelli tradizionali a lente sottile. L'analisi si articola su due fronti: una validazione tecnica, focalizzata sulla quantificazione della risoluzione tramite curve MTF e sulla misurazione delle performance computazionali (eseguite su architettura AMD Ryzen 7 5700X), e una validazione percettiva, basata sull'elaborazione statistica dei questionari soggettivi raccolti. L'obiettivo è determinare se la rappresentazione volumetrica delle aberrazioni tramite il frustum proposto garantisca una resa fotorealistica superiore senza introdurre colli di bottiglia critici nella pipeline di rendering.

5.1 Analisi MTF

Questa sezione è dedicata alla verifica quantitativa della nitidezza e del micro-contrasto spaziale. Attraverso il confronto tra le immagini generate dal sensore ibrido e quelle ottenute tramite il modello ideale *Thin Lens*, viene estrapolato il differenziale spettrografico del sistema. Tale analisi permette di isolare l'impatto delle traiettorie vettoriali del *bokeh* sulla dispersione energetica alle alte frequenze, distinguendo in modo oggettivo il degrado ottico indotto dalle aberrazioni reali rispetto ai limiti intrinseci di un sistema privo di imperfezioni.

5.1.1 Analisi della Risposta Spaziale (Mira ISO)

I grafici seguenti illustrano l'andamento della Modulation Transfer Function (MTF) estratta dalla risposta al gradino sulla mira ISO 12233. Osservando le curve, emerge in modo evidente come l'iniezione del frustum aberrato provochi una drastica e rapida caduta del contrasto alle alte frequenze spaziali in tutte le configurazioni

testate, distaccandosi nettamente dalla curva ideale di riferimento del modello *Thin Lens*.

Un aspetto di particolare interesse analitico riguarda il comportamento dell'aberrazione sferica (positiva e negativa). Sul piano di messa a fuoco esatto, i due profili di errore restituiscono un degrado della risoluzione pressoché identico, rendendoli di fatto indistinguibili in termini di pura nitidezza bidimensionale. Tuttavia, questa simmetria si rompe non appena si introduce una componente di *defocus*: spostando il piano focale avvicinandosi alla camera (nella cosiddetta zona intrafocale), la risposta al gradino dei due profili diverge radicalmente. Questa variazione conferma l'accuratezza del modello proposto nel riprodurre la complessa natura tridimensionale e asimmetrica dell'aberrazione sferica all'interno del volume ottico.

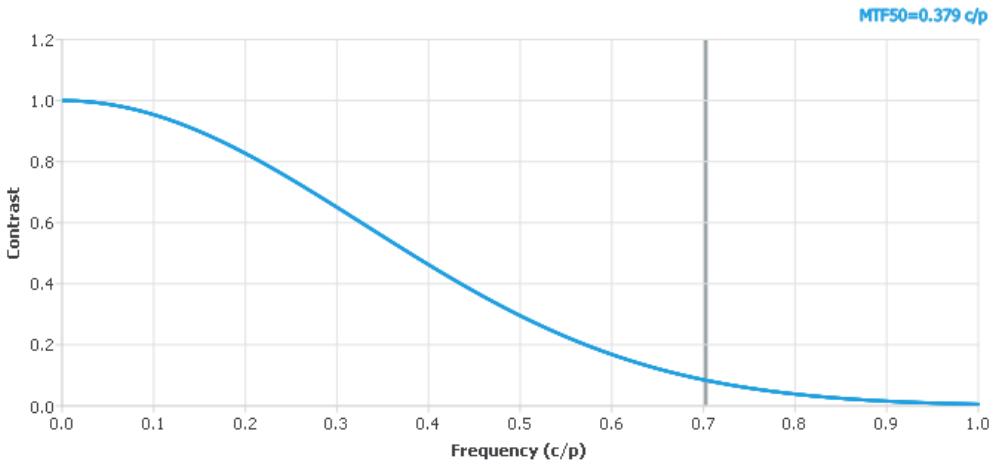


Figura 5.1: MTF ISO 12233 - senza aberrazioni.

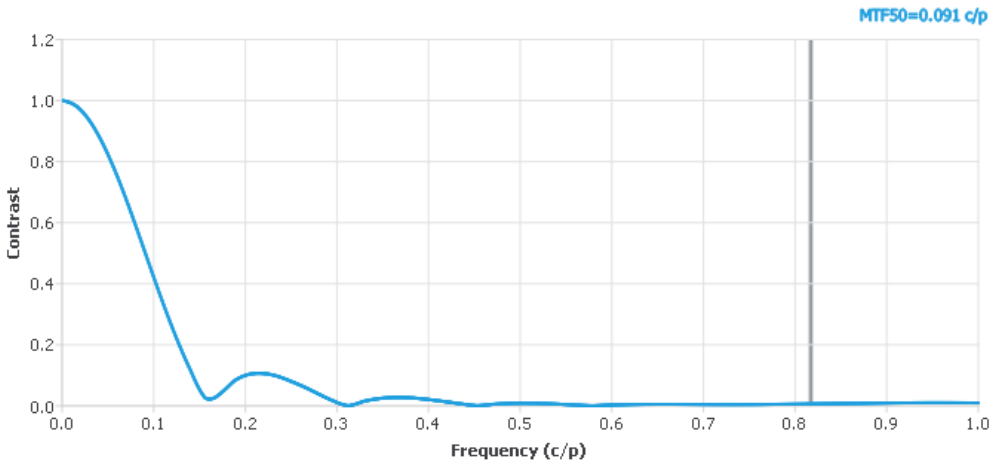


Figura 5.2: MTF ISO 12233 - astigmatismo.

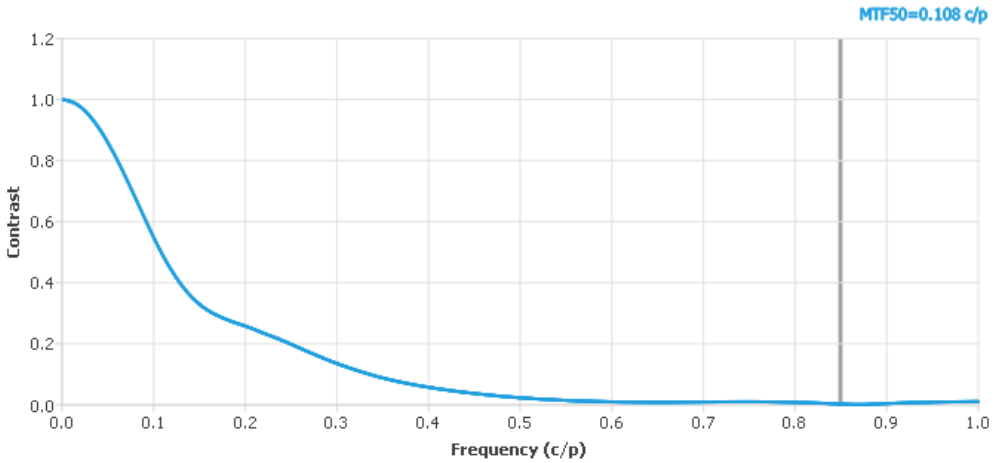


Figura 5.3: MTF ISO 12233 - coma.

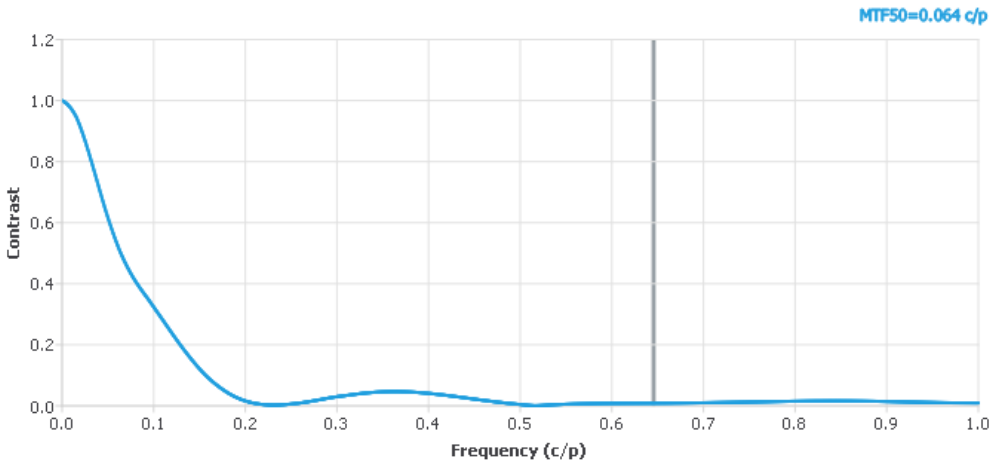


Figura 5.4: MTF ISO 12233 - aberrazione sferica positiva.

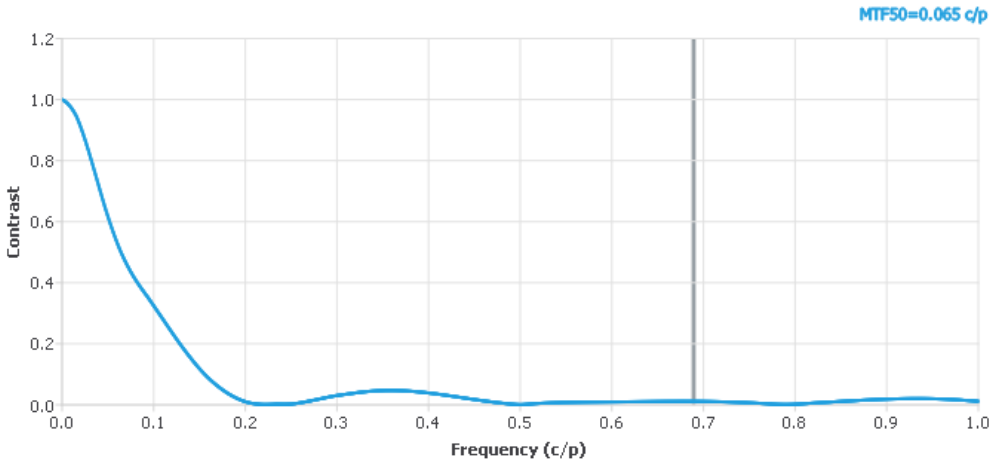


Figura 5.5: MTF ISO 12233 - aberrazione sferica negativa.

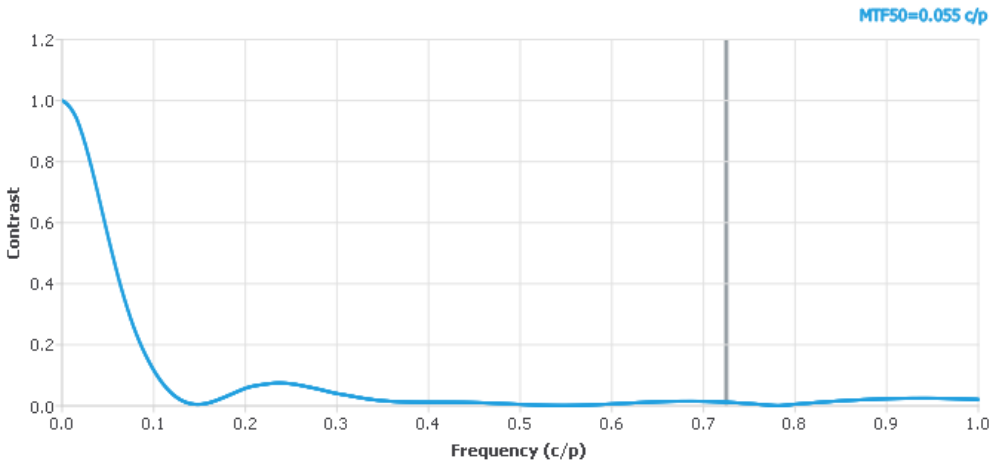


Figura 5.6: MTF ISO 12233 - aberrazione sferica positiva (focus a -1).

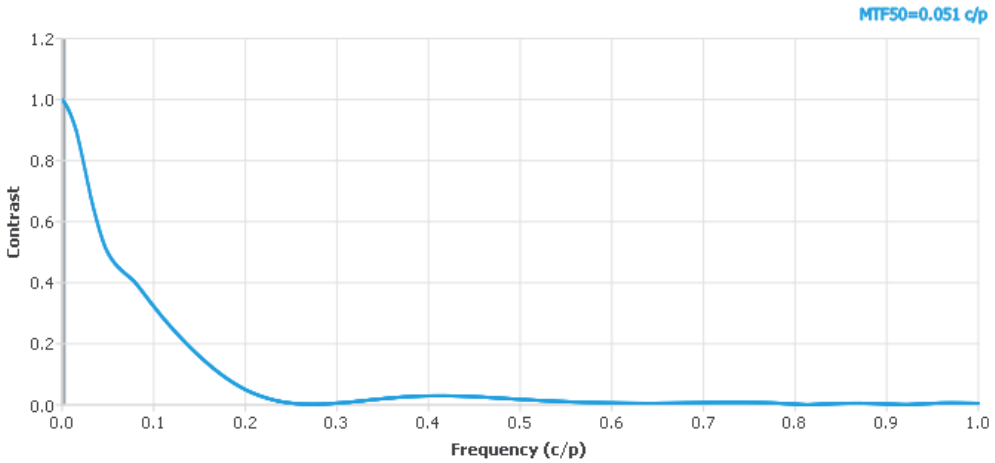


Figura 5.7: MTF ISO 12233 - aberrazione sferica negativa (focus a -1).

5.1.2 Stima della MTF su Scene Complesse

Questa sezione presenta l'estrazione della MTF estesa all'intero set di scene fotorealistiche. L'analisi è stata condotta tramite uno script dedicato che computa i profili di risposta spaziale, li sovrappone e ne estrapola il differenziale (*Delta MTF*).

Per ciascuna delle nove scene testate, la misurazione è stata eseguita su due differenti stadi di output: il file di calcolo lineare a 32-bit (`.exr`) e la sua controparte finale (`.png`) sottoposta al filtro *Open Image Denoise* (OIDN). La necessità di questo duplice approccio è di natura metodologica: l'obiettivo è quantificare l'impatto del processo di riduzione del rumore sulla fedeltà ottica dell'immagine. Mentre il formato `.exr` conserva l'informazione radiometrica pura generata dall'algoritmo di tracciamento, che tiene conto della varianza stocastica, l'algoritmo di *denoising* basato su reti neurali rischia intrinsecamente di levigare i micro-dettagli, scambiando le alte frequenze spaziali delle aberrazioni o delle texture per artefatti da Monte Carlo.

Il confronto diretto tra le due funzioni MTF permette, dunque, di isolare l'eventuale perdita di risoluzione introdotta dalla pipeline di post-elaborazione. Per esaltare visivamente le deviazioni e il decadimento del micro-contrasto, la scala delle ordinate all'interno dei grafici è stata impostata in formato logaritmico.

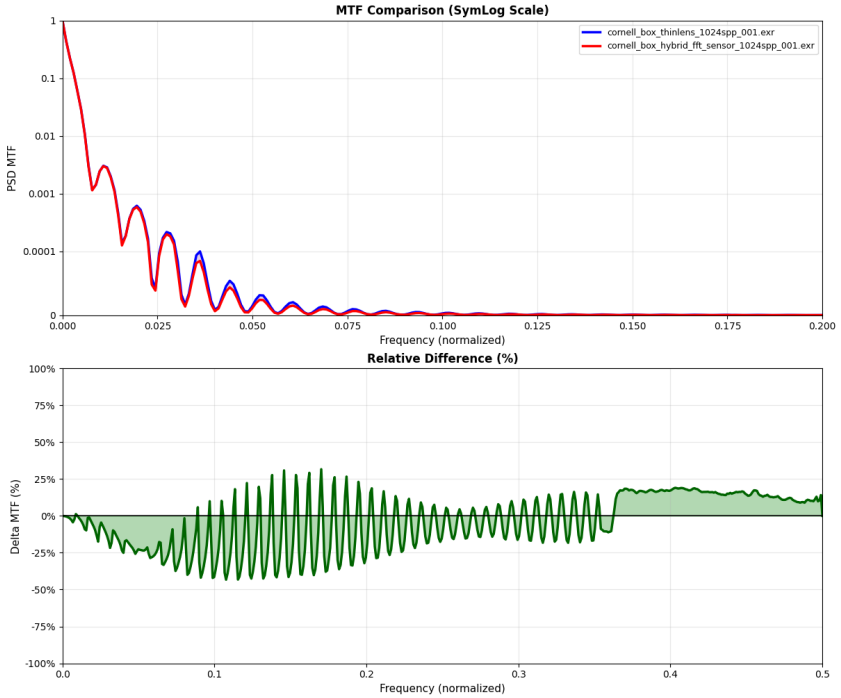


Figura 5.8: Stima della MTF e delta - Cornell Box - file .exr

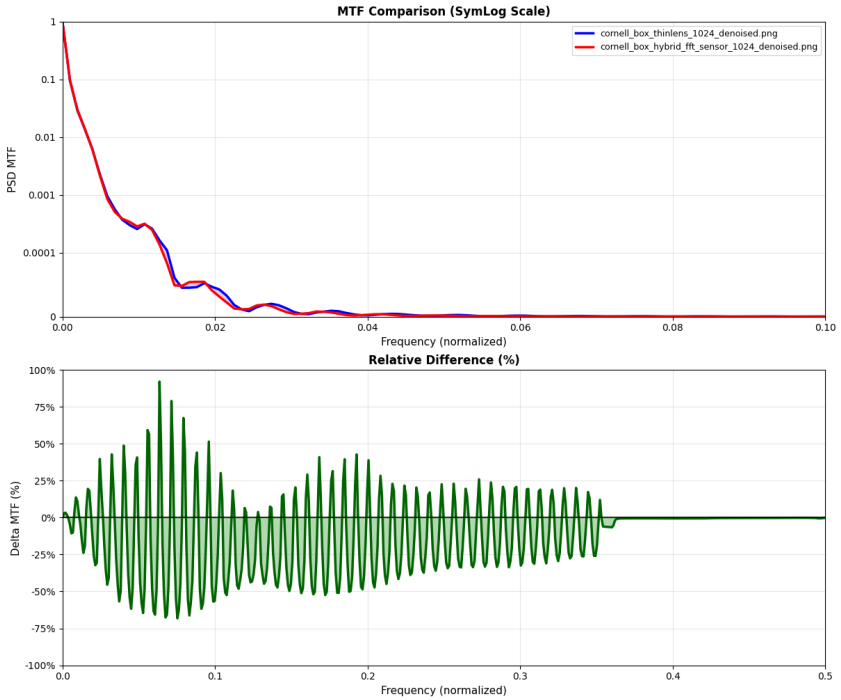


Figura 5.9: Stima della MTF e delta - Cornell Box - denoised .png

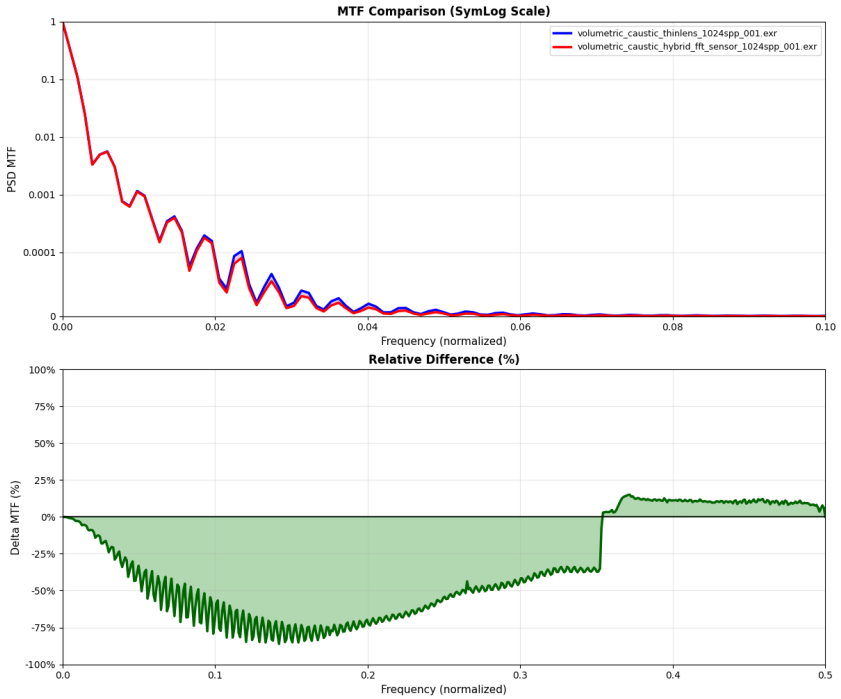


Figura 5.10: Stima della MTF e delta - Cornell Box Glass - file .exr

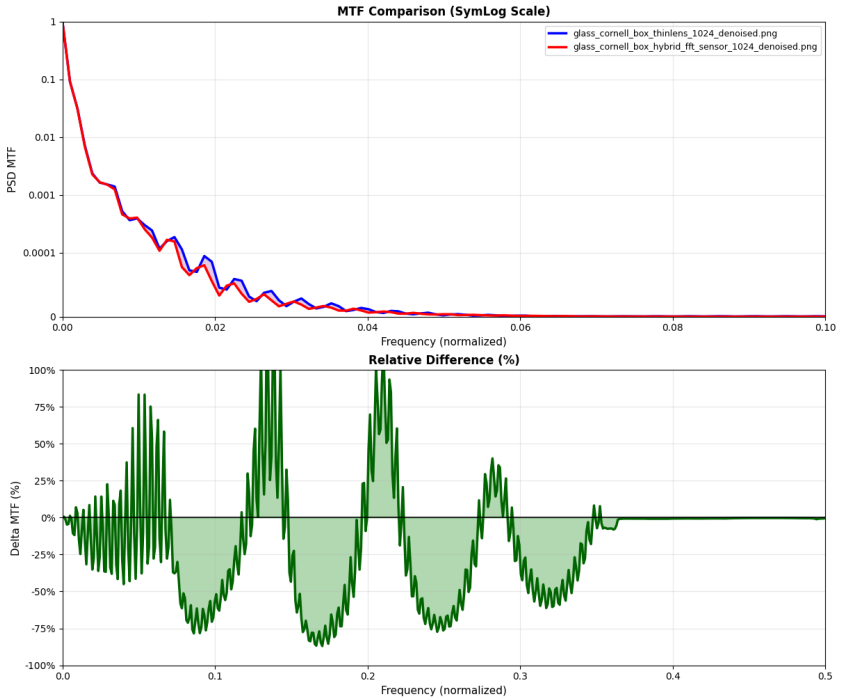


Figura 5.11: Stima della MTF e delta - Cornell Box Glass - denoised .png

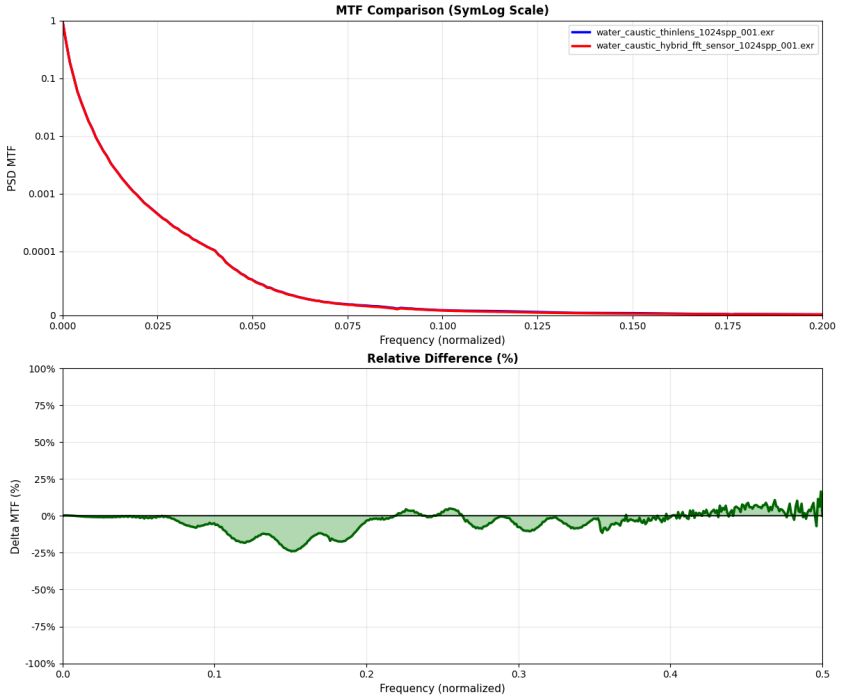


Figura 5.12: Stima della MTF e delta - Cornell Box Water - file .exr

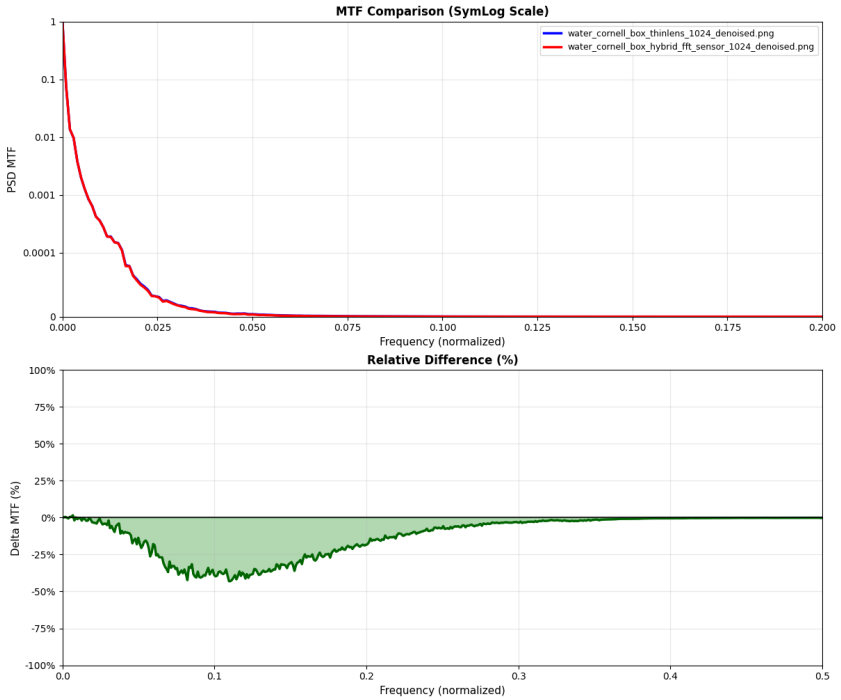


Figura 5.13: Stima della MTF e delta - Cornell Box Water - denoised .png

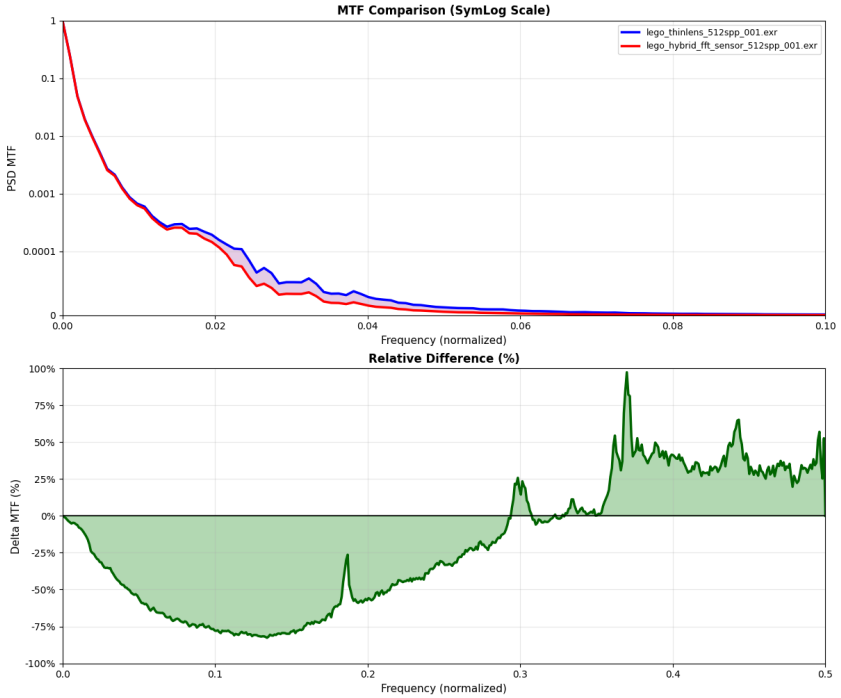


Figura 5.14: Stima della MTF e delta - Lego - file .exr

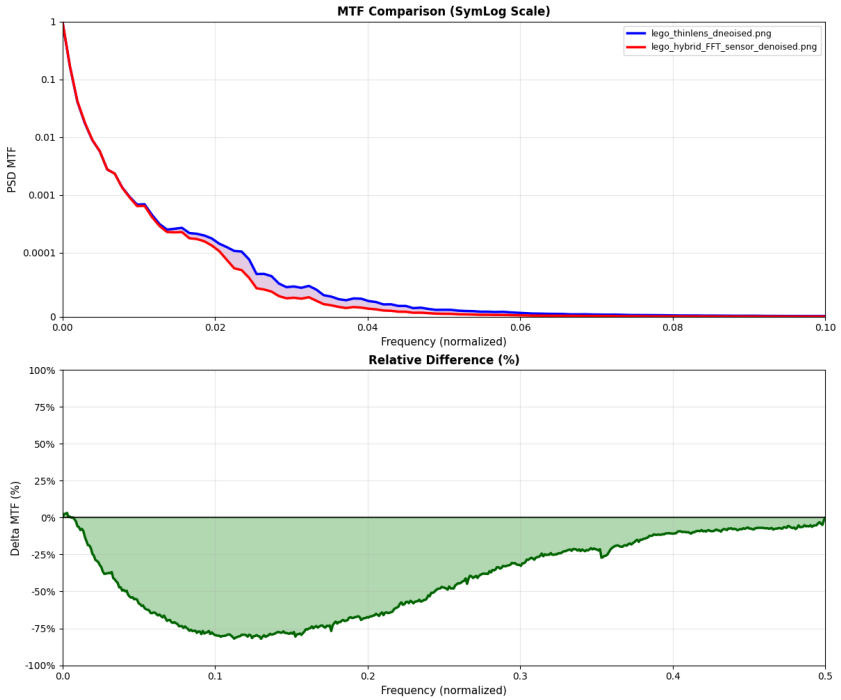


Figura 5.15: Stima della MTF e delta - Lego - denoised .png

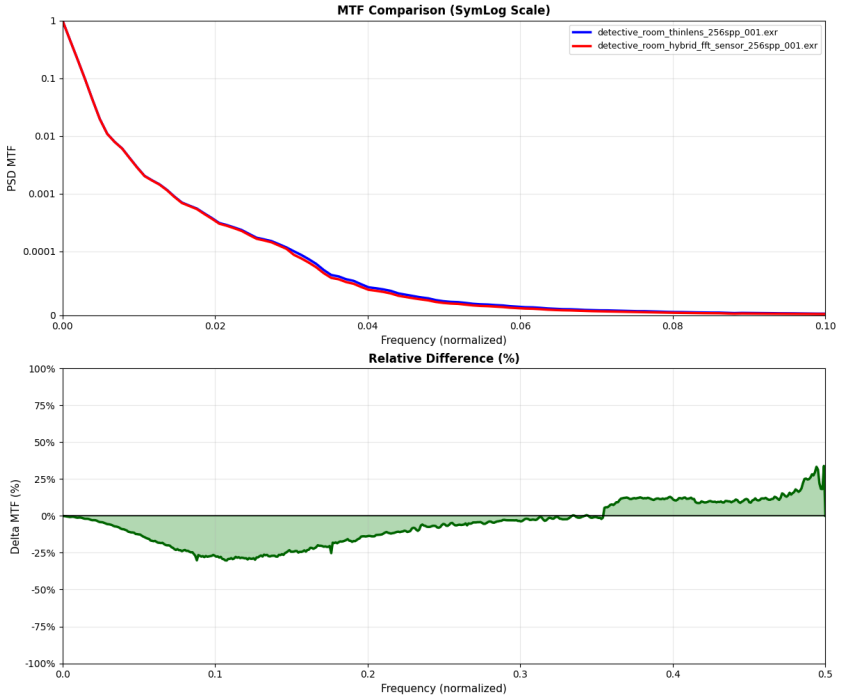


Figura 5.16: Stima della MTF e delta - Detective Room - file .exr

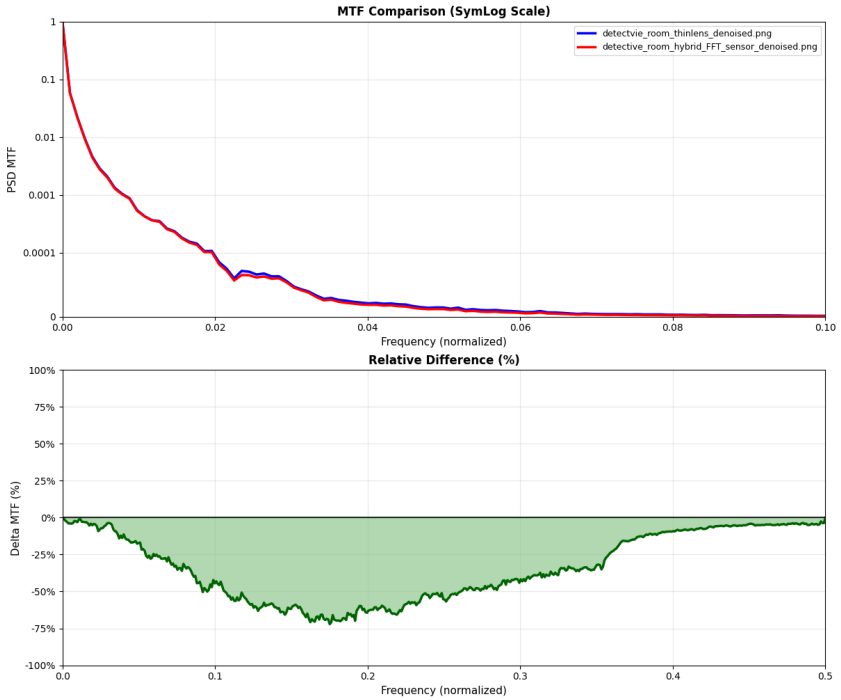


Figura 5.17: Stima della MTF e delta - Detective Room - denoised .png

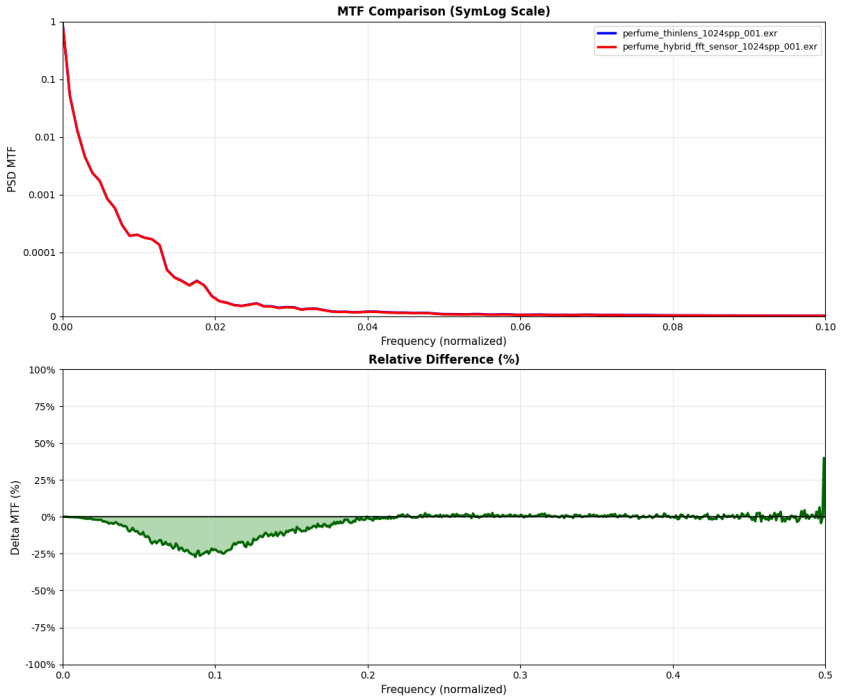


Figura 5.18: Stima della MTF e delta - Perfume - file .exr

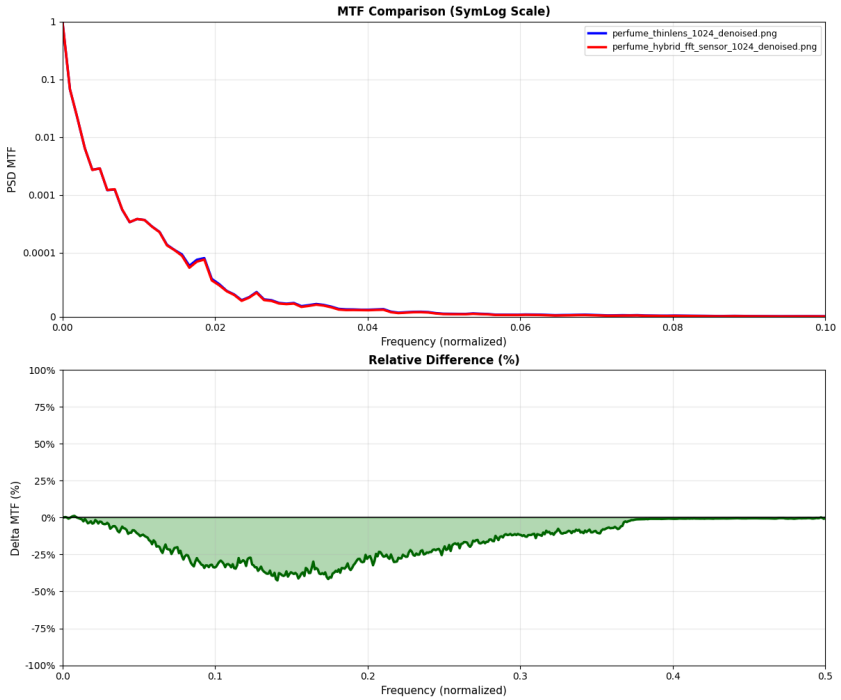


Figura 5.19: Stima della MTF e delta - Perfume - denoised .png

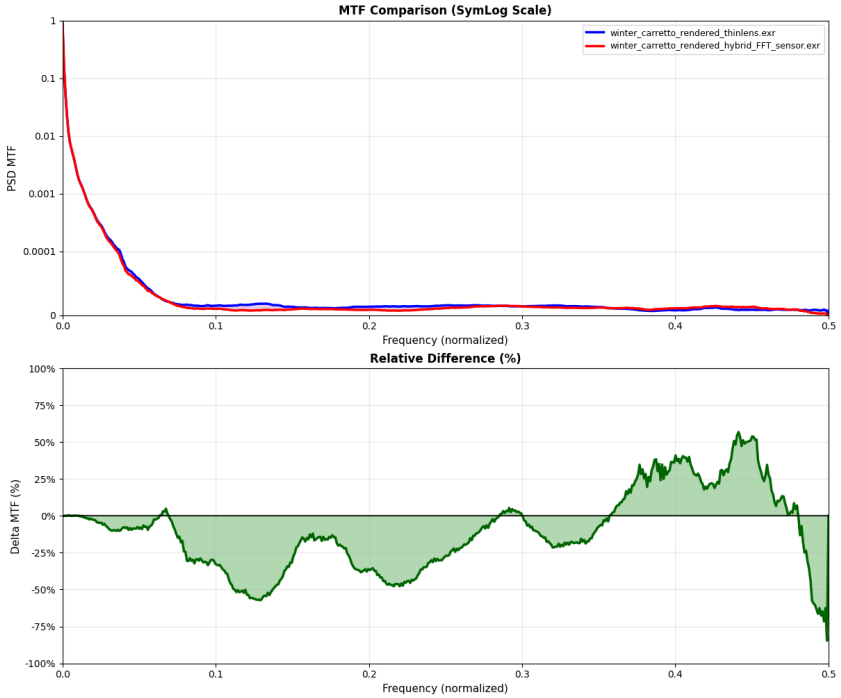


Figura 5.20: Stima della MTF e delta - Winter forest - file .exr

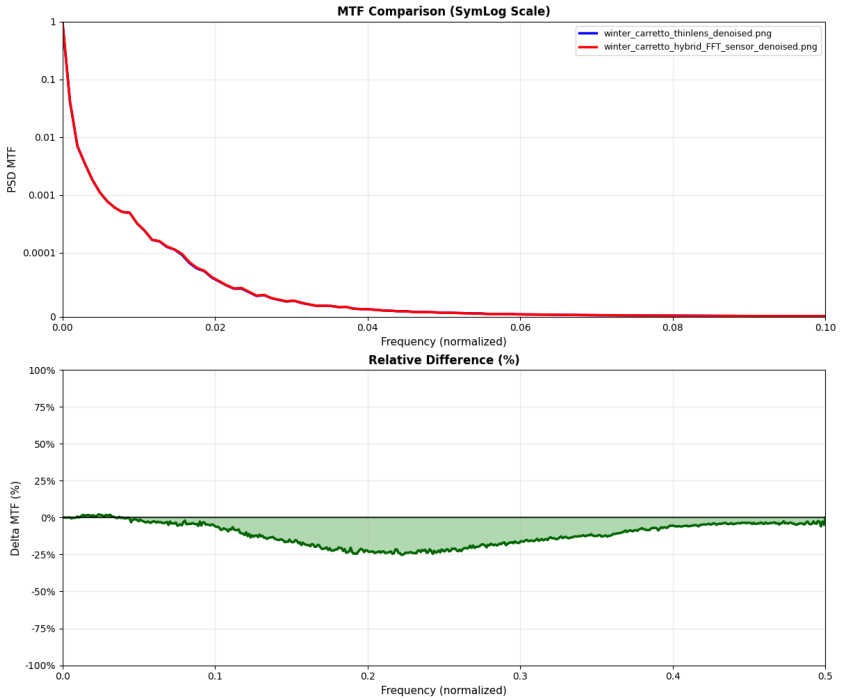


Figura 5.21: Stima della MTF e delta - Winter forest - denoised .png

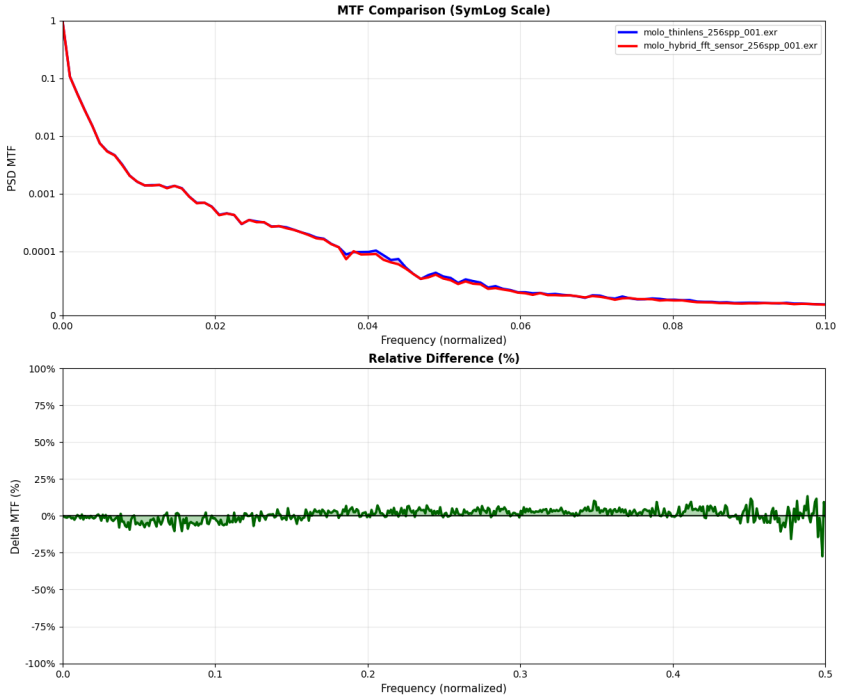


Figura 5.22: Stima della MTF e delta - Wooden pier - file .exr

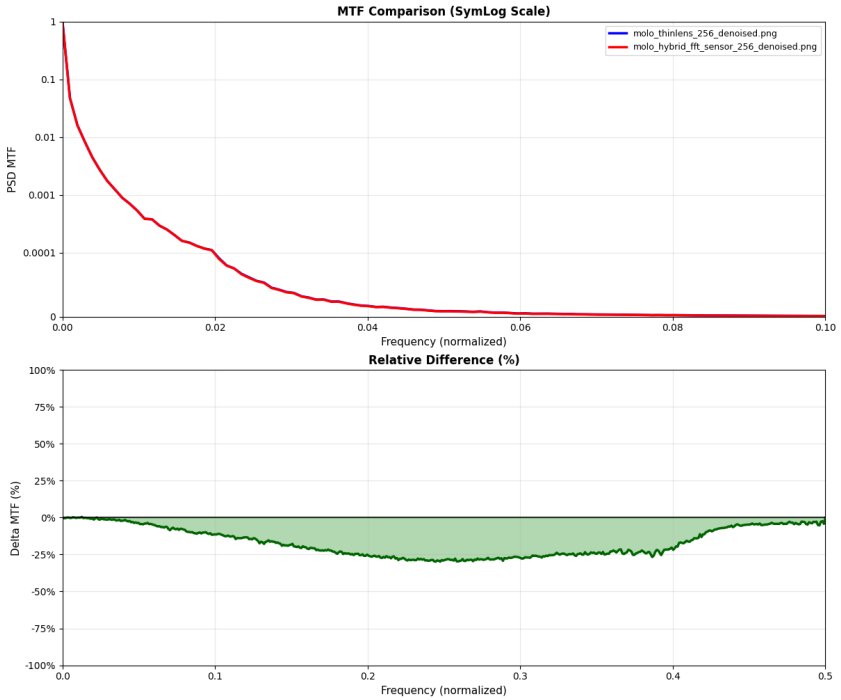


Figura 5.23: Stima della MTF e delta - Wooden pier - denoised .png

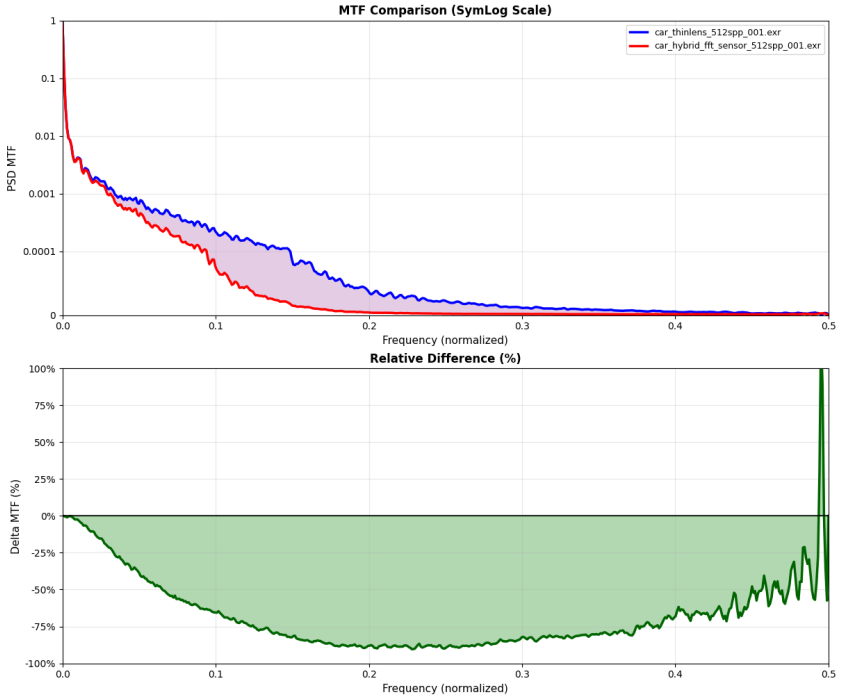


Figura 5.24: Stima della MTF e delta - Car - file .exr

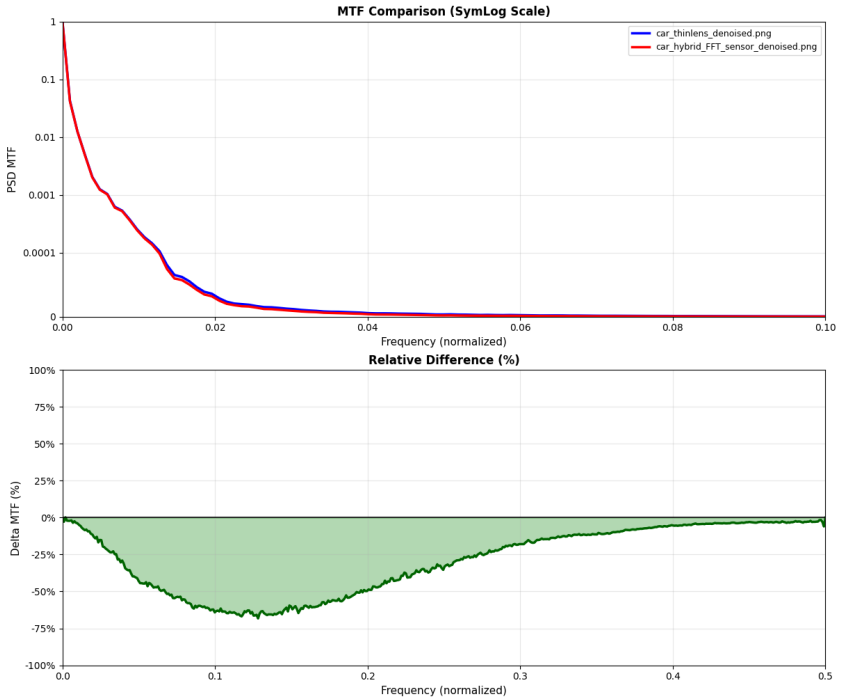


Figura 5.25: Stima della MTF e delta - Car - denoised .png

Analisi Comparativa dei Differenziali

Dall'analisi complessiva dei differenziali, è fondamentale premettere una considerazione sulla scala logaritmica adottata per la densità spettrale: sebbene in alcuni grafici il Delta percentuale raggiunga picchi notevoli, nella maggior parte delle casistiche questi scostamenti si verificano su magnitudini energetiche estremamente ridotte. Di conseguenza, le discrepanze tra l'output lineare e quello filtrato si traducono in alterazioni numericamente marginali sul segnale complessivo, confinate a variazioni microscopiche del micro-contrasto che non compromettono la struttura globale della scena.

Un esempio di questo comportamento si osserva in *Detective Room* (Figure 5.16 e 5.17) e *Perfume* (Figure 5.18 e 5.19). Nel primo caso, la fitta presenza di texture ad alta frequenza porta l'algoritmo di *denoising* ad agire come un filtro passa-basso, attenuando analiticamente le frequenze associate alla porosità delle superfici in penombra. Nel caso del profumo, il filtro tende a omogeneizzare i gradienti luminosi e le micro-riflessioni sull'acqua. Tuttavia, pur registrando una flessione percentuale evidente nel grafico del Delta, l'impatto complessivo sulla conservazione del segnale rimane trascurabile, proprio a causa dei valori spettrali assoluti minimi su cui si verifica la correzione.

Al contrario, l'intervento del *denoiser* produce alterazioni spettrali più marcate in scenari caratterizzati da oggetti voluminosi in forte primo piano, come *Lego* (Figure 5.14 e 5.15) e *Car* (Figure 5.24 e 5.25). A causa della ridotta distanza dalla camera, la geometria del *frustum* di tracciamento risulta particolarmente ripida nella sua discesa, generando una dispersione vettoriale (e quindi un *bokeh*) fisicamente molto pronunciata e densa di energia. In questi contesti, pur preservando intatta la leggibilità visiva macroscopica delle aberrazioni, l'azione di OIDN si rivela maggiormente invasiva sul piano strettamente analitico: la curva della MTF nel file `.png` tende ad appiattirsi alle alte frequenze, smussando i picchi di energia terminali. Questo fenomeno si traduce in una fisiologica flessione della micro-risoluzione oggettiva; l'algoritmo, nel processo di soppressione della varianza stocastica, finisce per attenuare anche quelle frequenze necessarie a ricostruire matematicamente la rugosità della plastica del modellino e i dettagli speculari più minuti sulla vernice metallizzata dell'automobile.

Infine, l'analisi conferma l'esistenza di scenari in cui l'intervento del filtro si rivela radiometricamente conservativo, come per il *Wooden Pier* (Figure 5.22 e 5.23). In assenza di forti sfuocati in primo piano o di complessi fenomeni di rifrazione, il differenziale si mantiene stabile e prossimo allo zero, permettendo all'algoritmo di ricostruire i campioni mancanti senza alterare in modo significativo la struttura spettrale delle aberrazioni calcolate dal sensore.

5.2 Analisi dell'Esperienza Utente (Questionari)

Di seguito vengono presentati i risultati del questionario di percezione soggettiva, condotto su un campione di 113 rispondenti attraverso il confronto di nove coppie di immagini. Per ciascuna coppia sono riportati due grafici, per un totale di diciotto figure, seguite da una media aggregata delle valutazioni e da un quadro d'insieme sull'intero campione.

La prima categoria di grafici (una per coppia) mostra i punteggi medi sui sei criteri percettivi valutati (dettagli fini, fedeltà fotografica, sensazione di artificialità, qualità dell'area a fuoco e di quella sfocata, realismo complessivo), su una scala Likert da -2 (forte preferenza per il render standard) a $+2$ (forte preferenza per il render aberrato). I punteggi sono suddivisi in sei pannelli: cinque riportano la scomposizione per livello di competenza dichiarata in altrettante aree (modellazione 3D, VFX, rendering, fotografia, ottica), il sesto la scomposizione per fascia d'età. Nel pannello legato alla media aggregata le barre d'errore rappresentano l'errore standard della media. La seconda categoria di grafici (una per coppia) riporta le preferenze espresse in termini di gradevolezza e fotorealismo percepito, questa volta come scelta binaria tra le due immagini. I due pannelli superiori mostrano la percentuale di preferenza per l'immagine standard o aberrata in funzione della fascia d'età; i due pannelli inferiori la stessa informazione in funzione del livello di competenza dichiarato in ciascuna delle cinque aree, con il colore che identifica l'immagine preferita e l'opacità il livello di competenza.

Chiude la sezione un quadro d'insieme sull'intero campione, che mette in relazione le risposte alla domanda "hai notato differenze apprezzabili?" con età, competenza e dispositivo utilizzato per la compilazione, un profilo percettivo medio aggregato sulle nove immagini, e un confronto della capacità di discriminazione percettiva tra dispositivo desktop e mobile.

5.2.1 Risultati Statistici

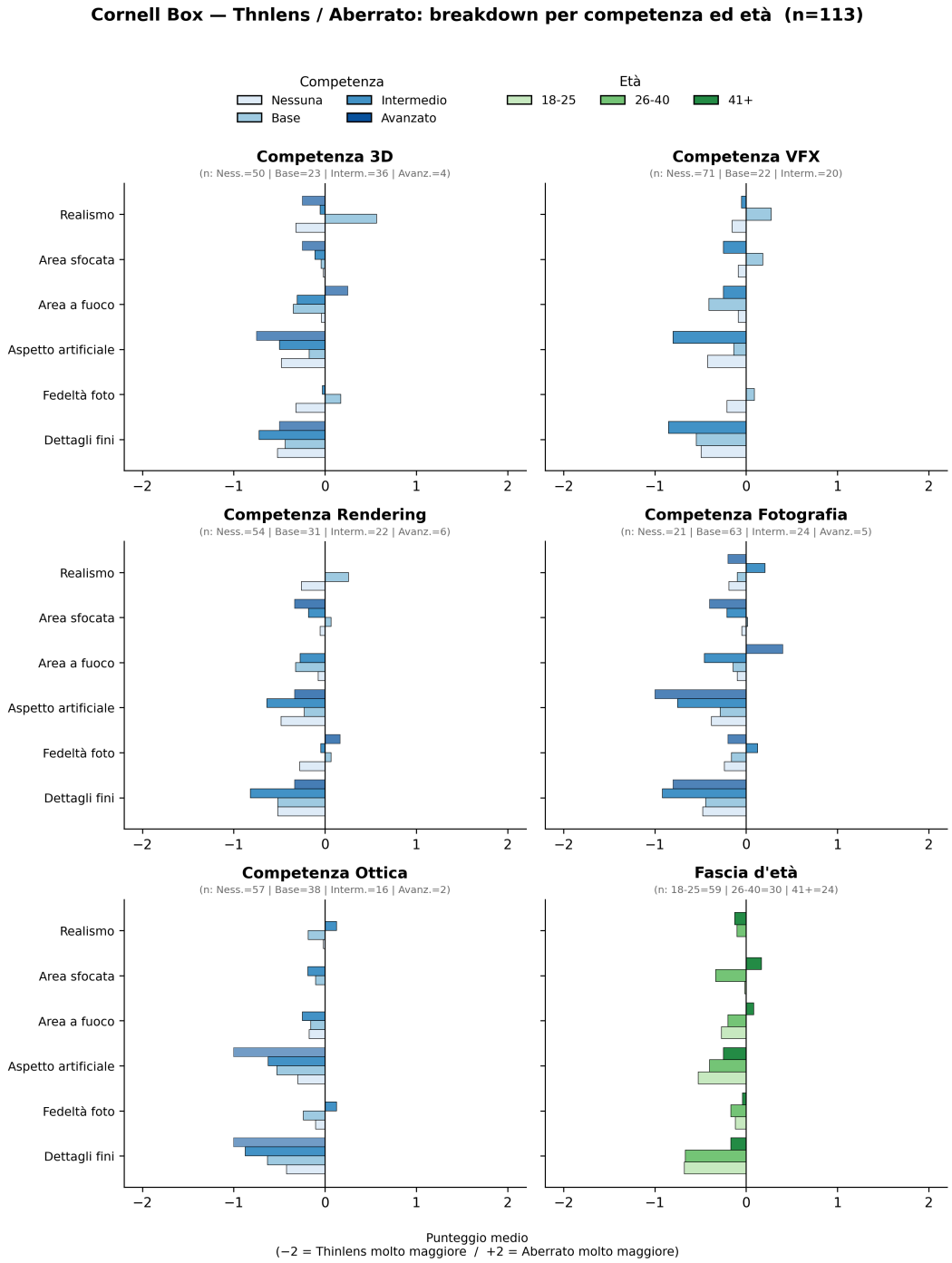


Figura 5.26: Coppia *Cornell Box* — Likert competenza | età.

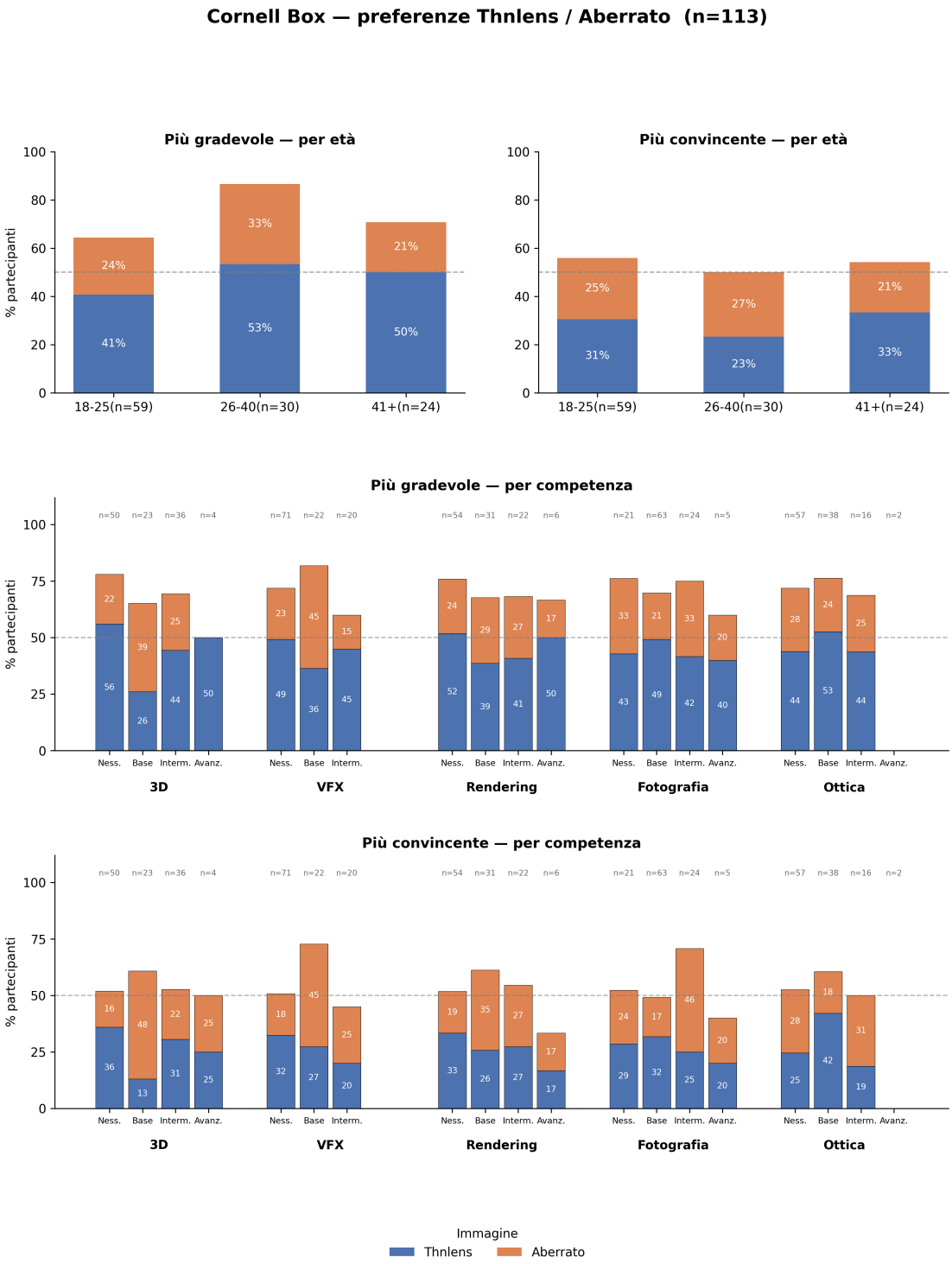


Figura 5.27: Coppia *Cornell Box* — preferenze A/B.

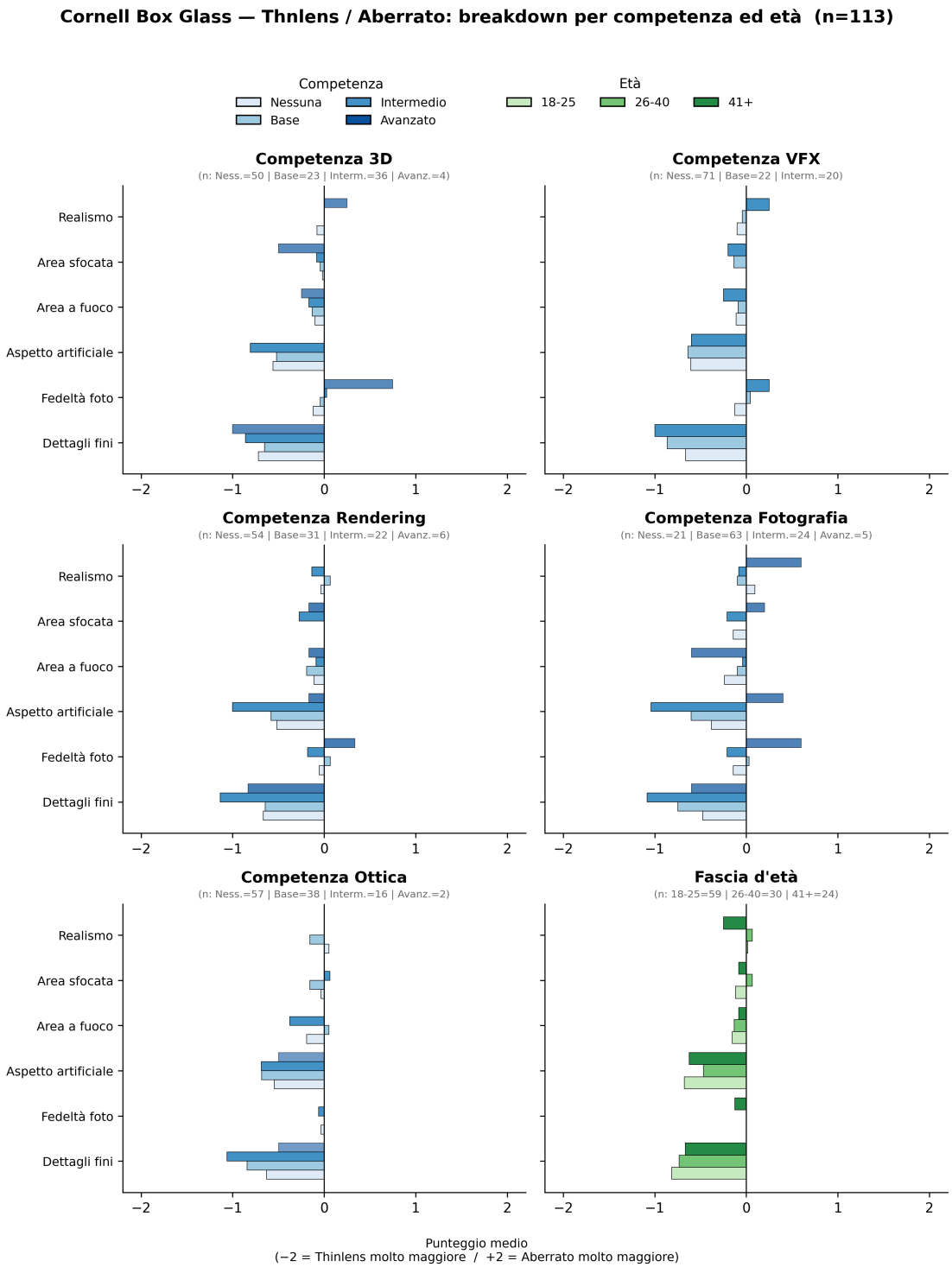


Figura 5.28: Coppia *Cornell Box Glass*— Likert competenza | età.

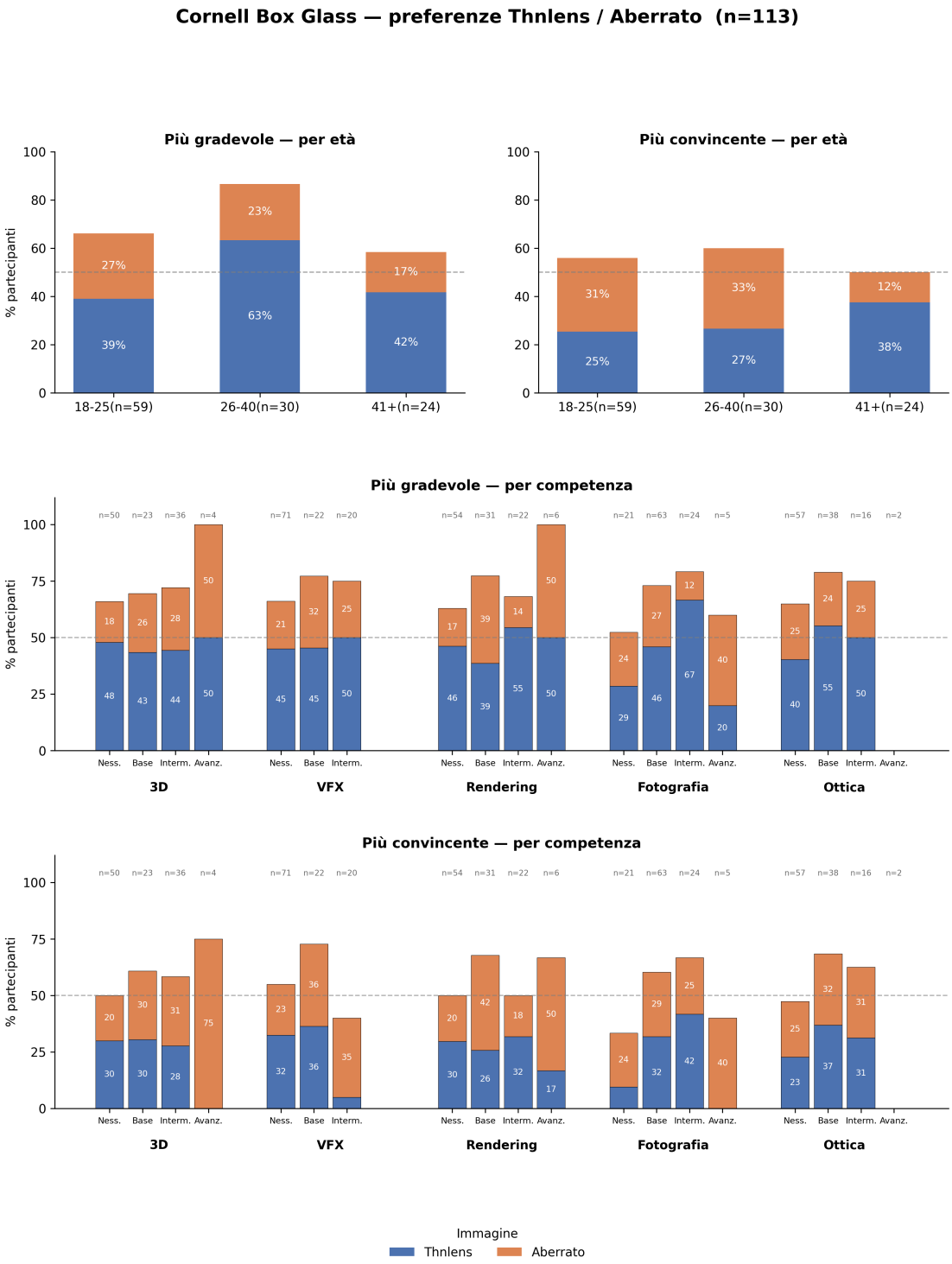


Figura 5.29: Coppia *Cornell Box Glass* — preferenze A/B.

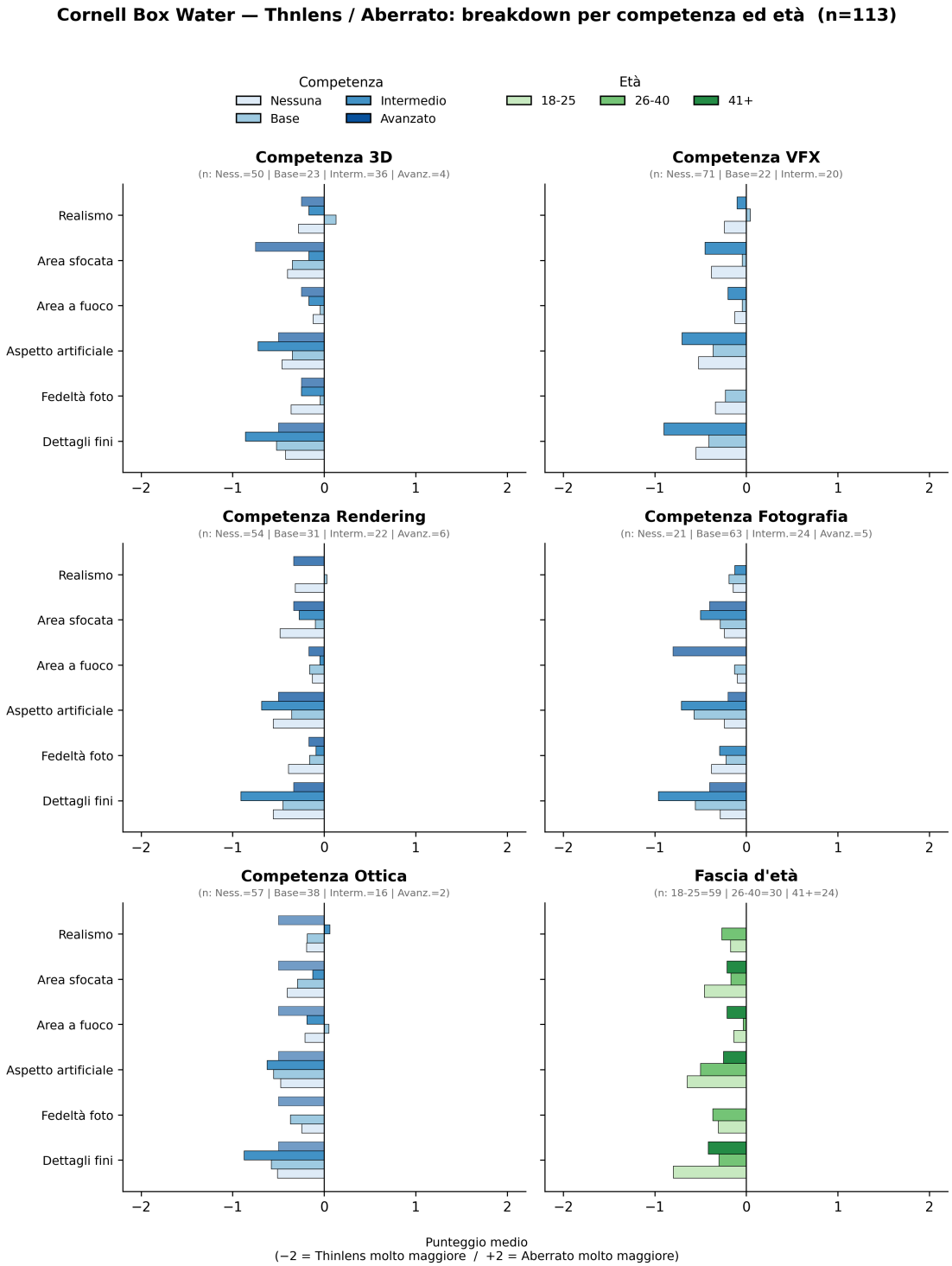


Figura 5.30: Coppia *Cornell Box Water* — Likert competenza | età.

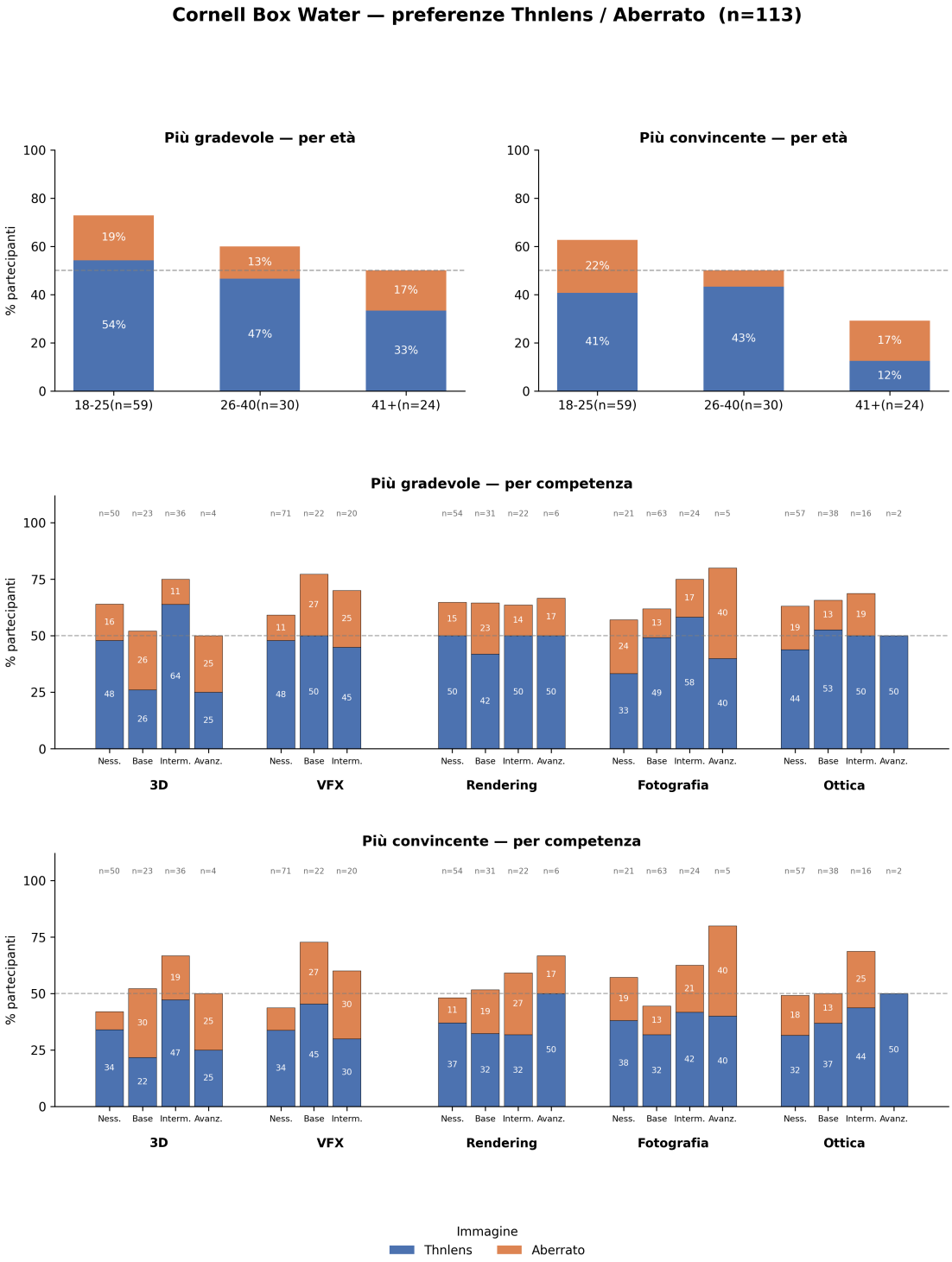


Figura 5.31: Coppia *Cornell Box Water* — preferenze A/B.

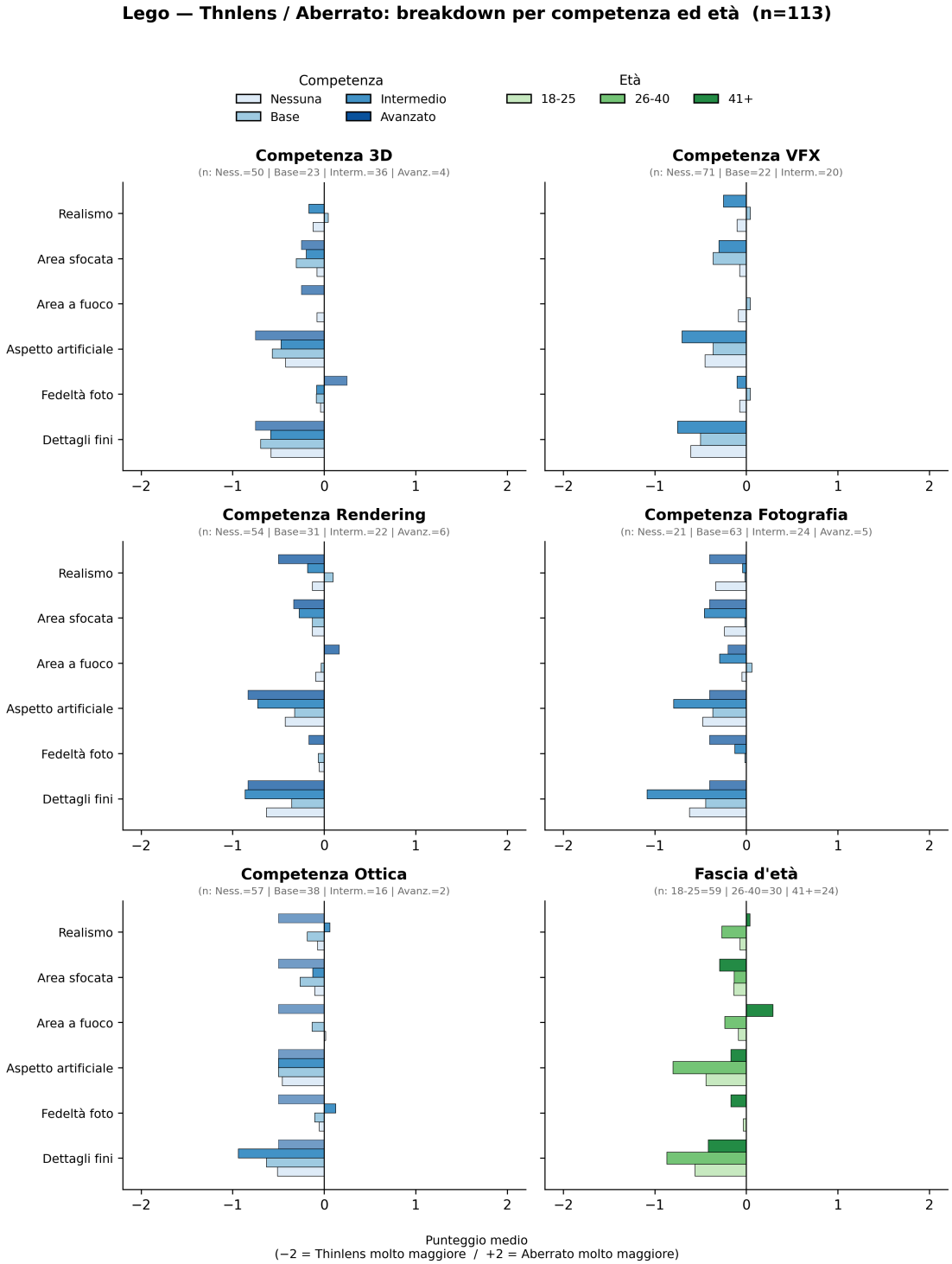


Figura 5.32: Coppia *Lego* — Likert competenza | età.

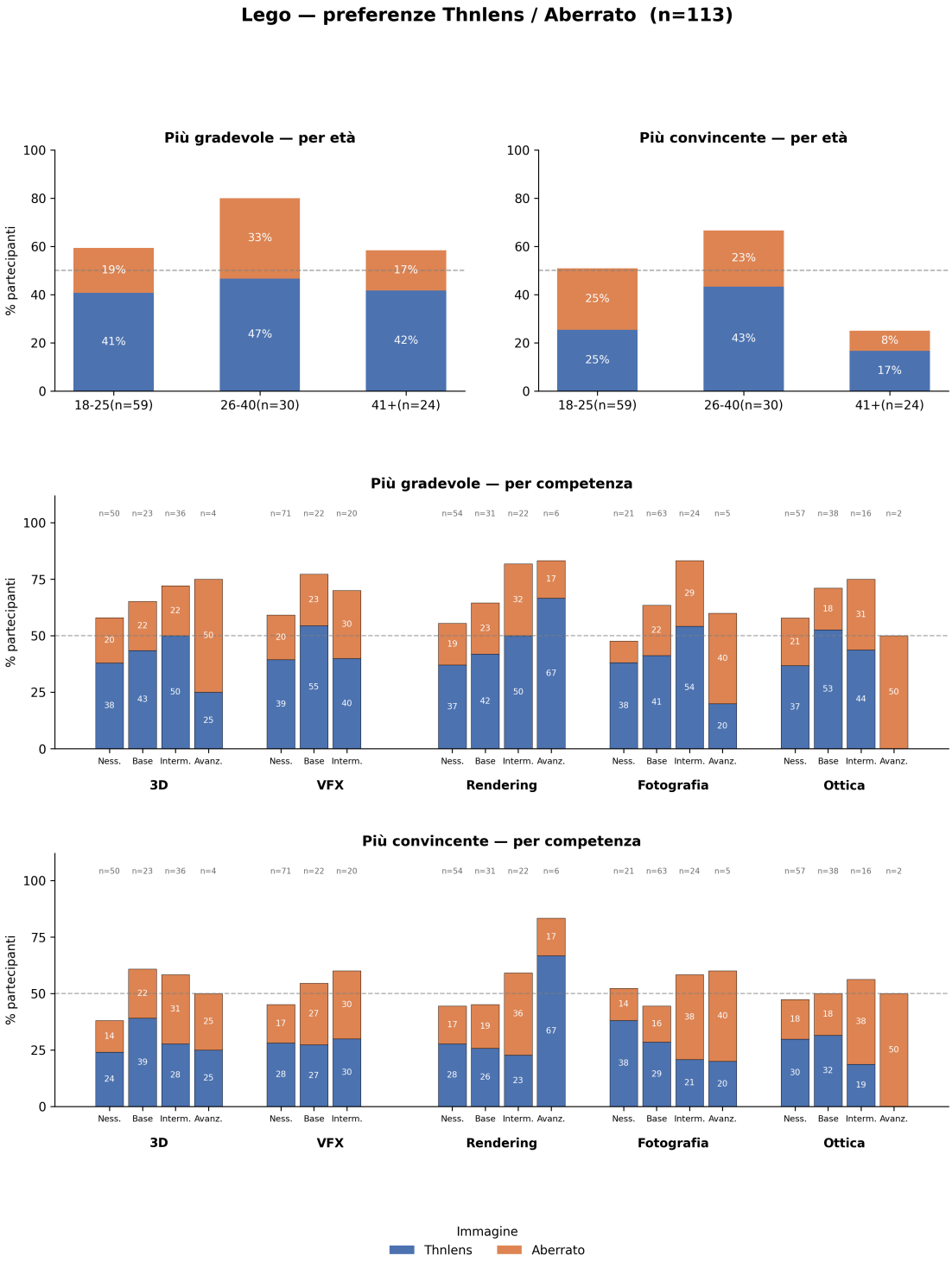


Figura 5.33: Coppia *Lego* — preferenze A/B.

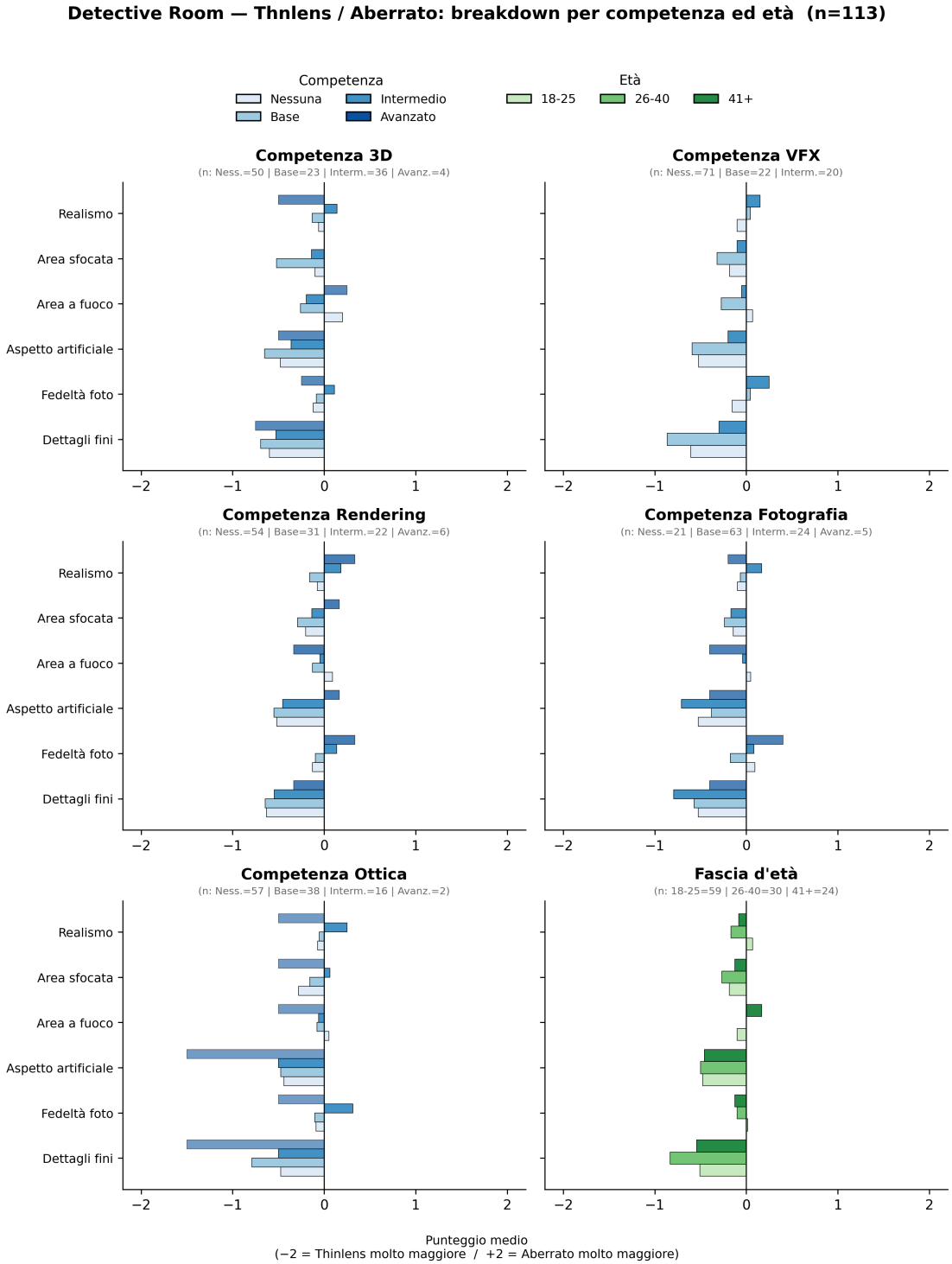


Figura 5.34: Coppia *Detective room* — Likert competenza | età.

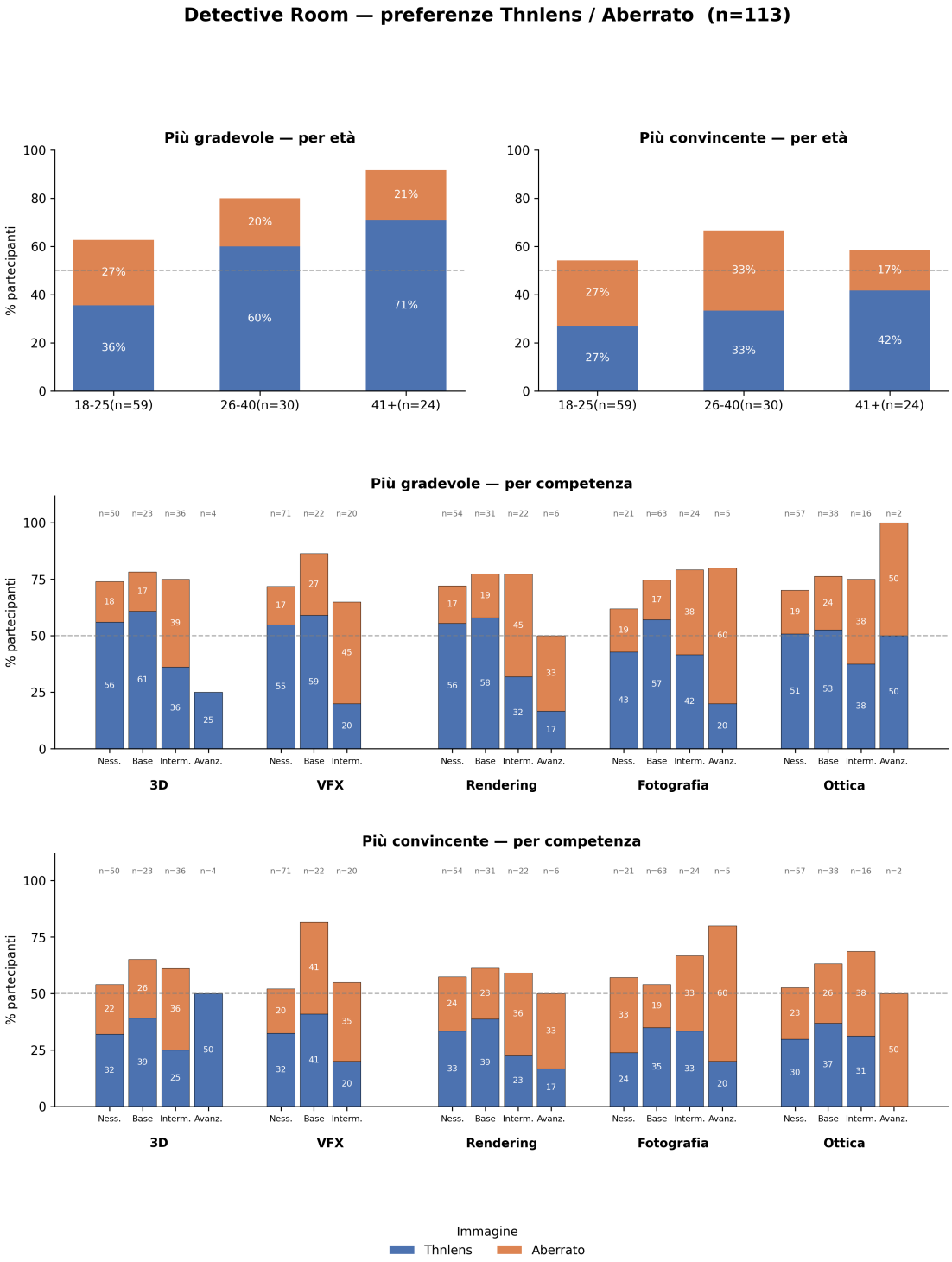


Figura 5.35: Coppia *Detective room* — preferenze A/B.

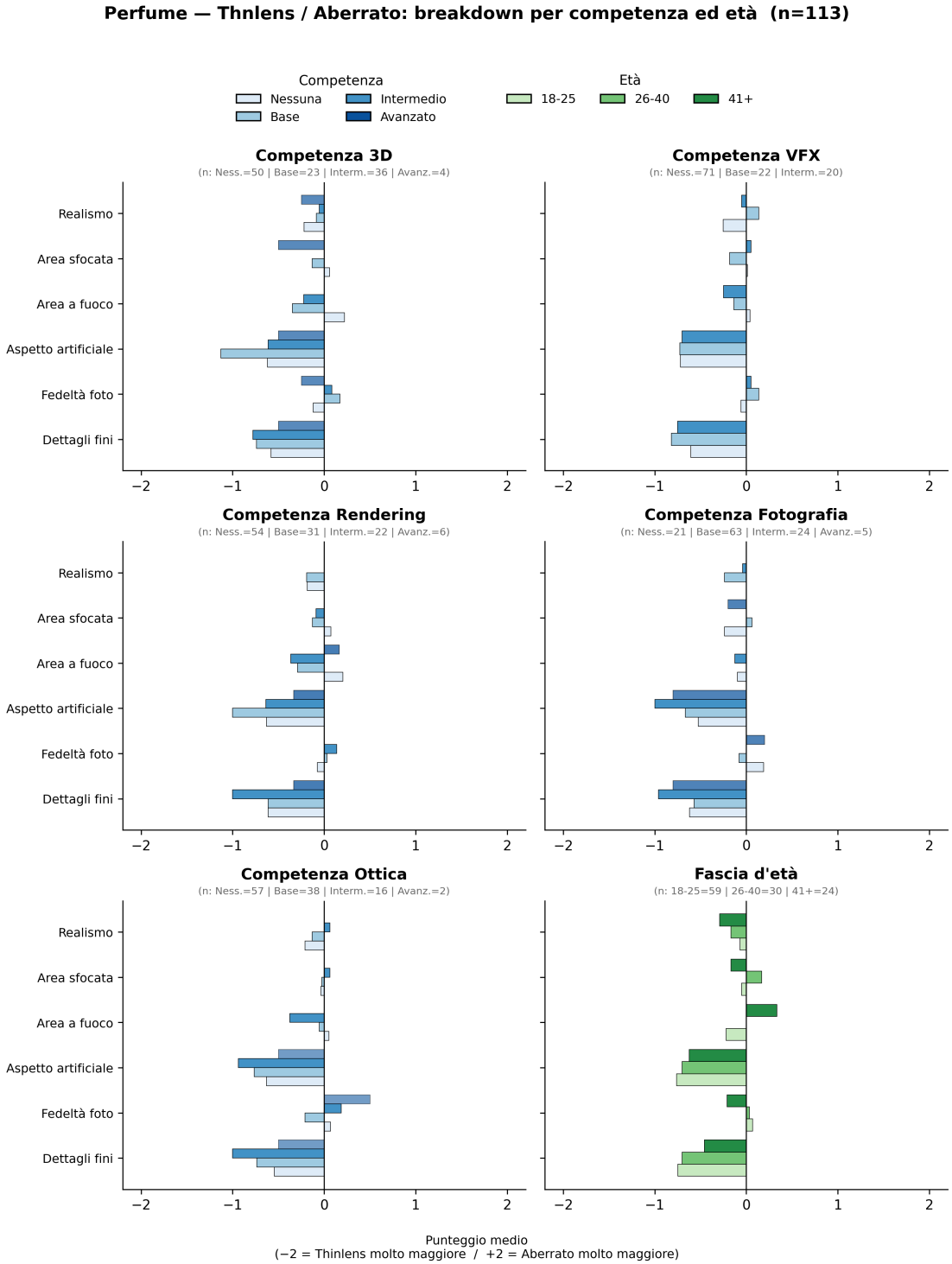


Figura 5.36: Coppia *Perfume* — Likert competenza | età.

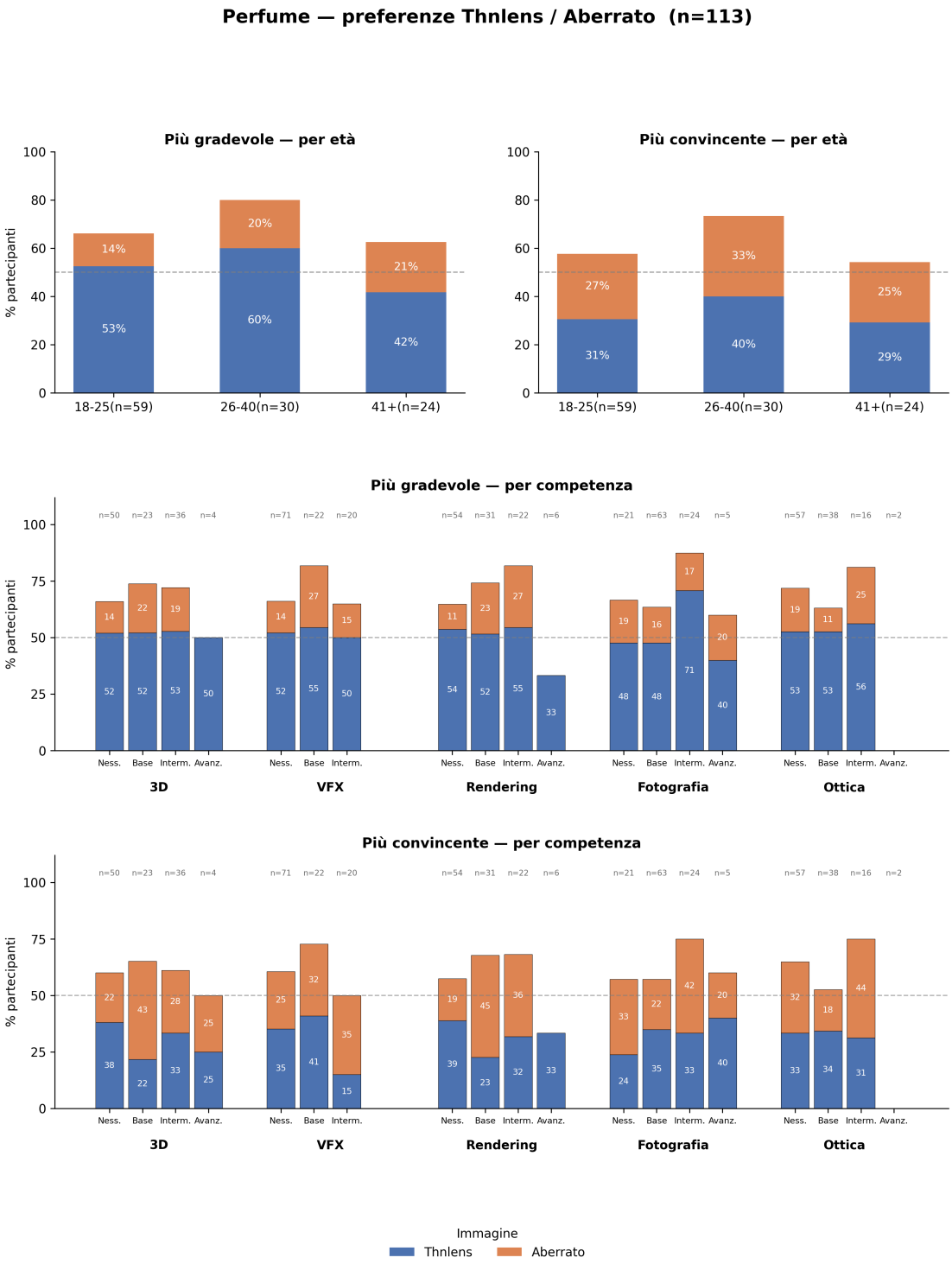


Figura 5.37: Coppia *Perfume* — preferenze A/B.

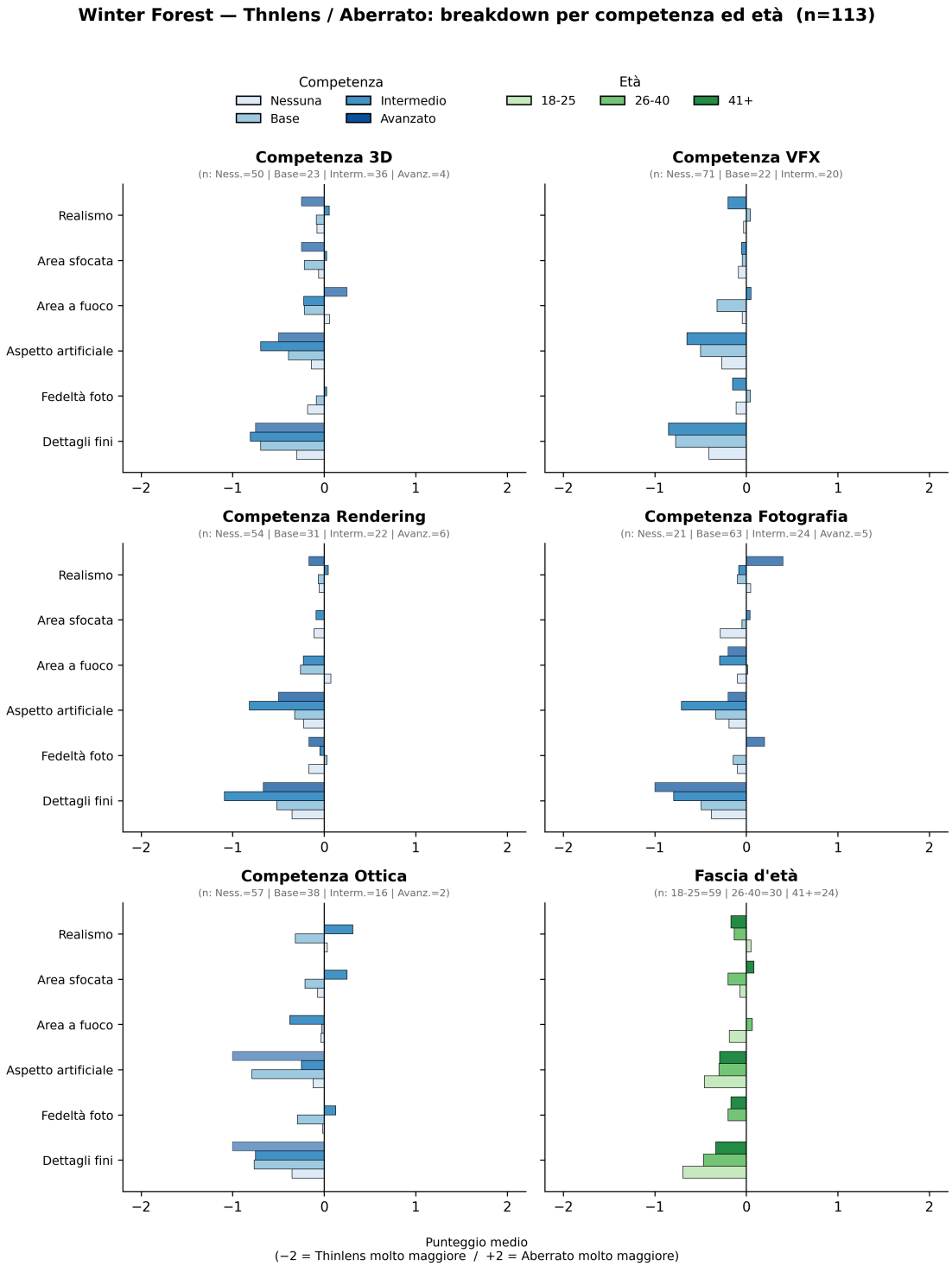


Figura 5.38: Coppia *Winter forest* — Likert competenza | età.

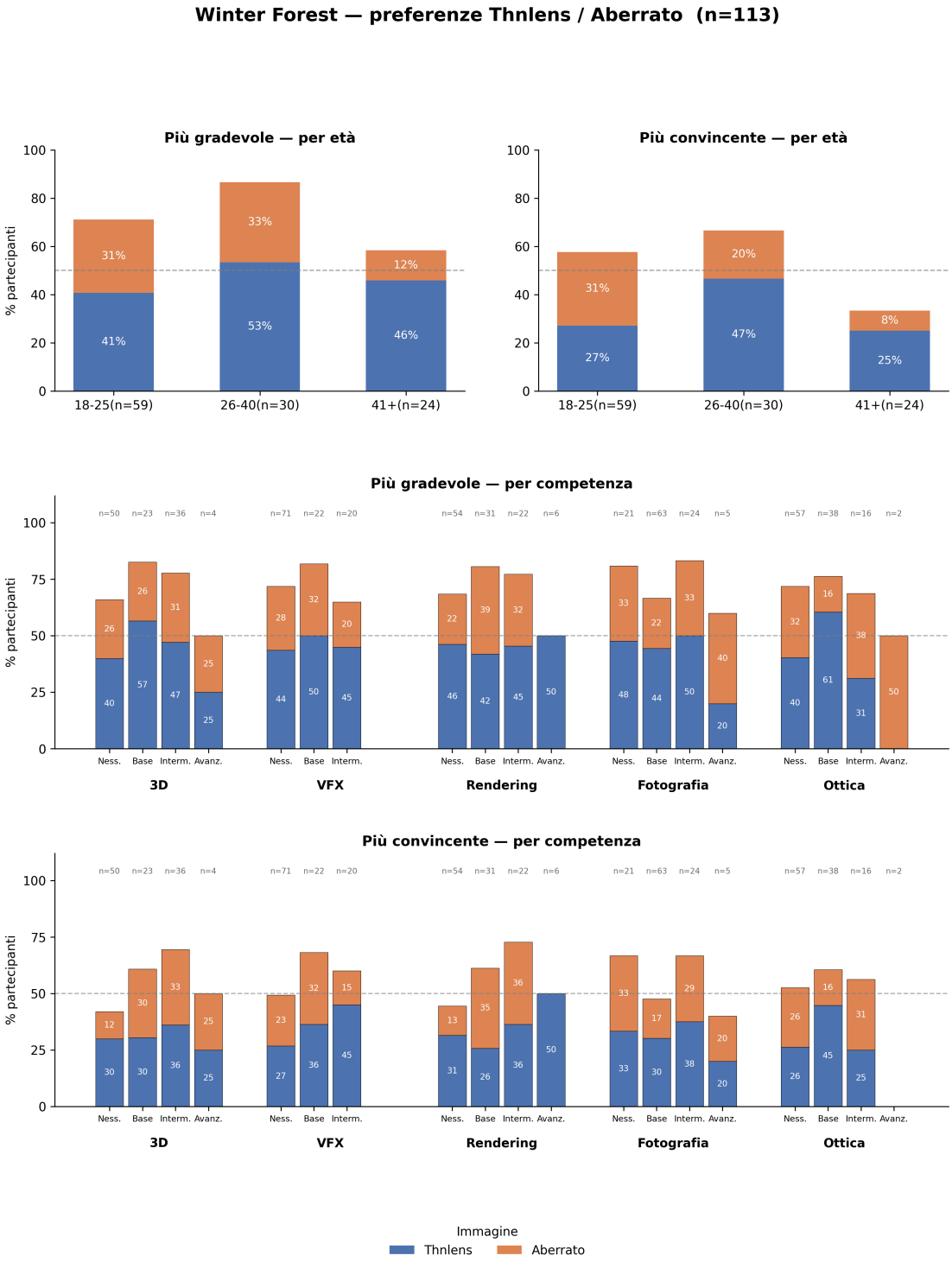


Figura 5.39: Coppia *Winter forest* — preferenze A/B.

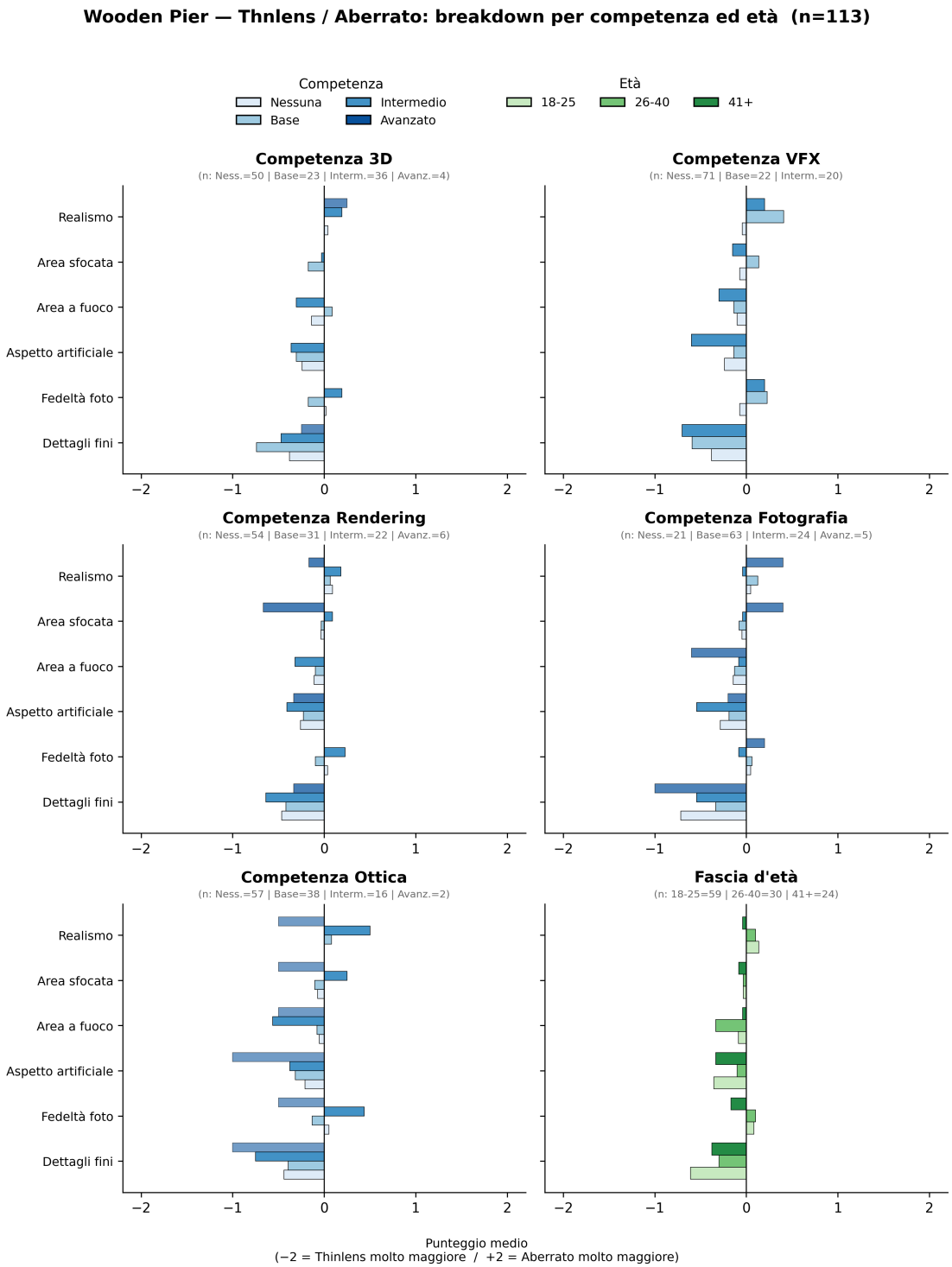


Figura 5.40: Coppia *Wooden pier* — Likert competenza | età.

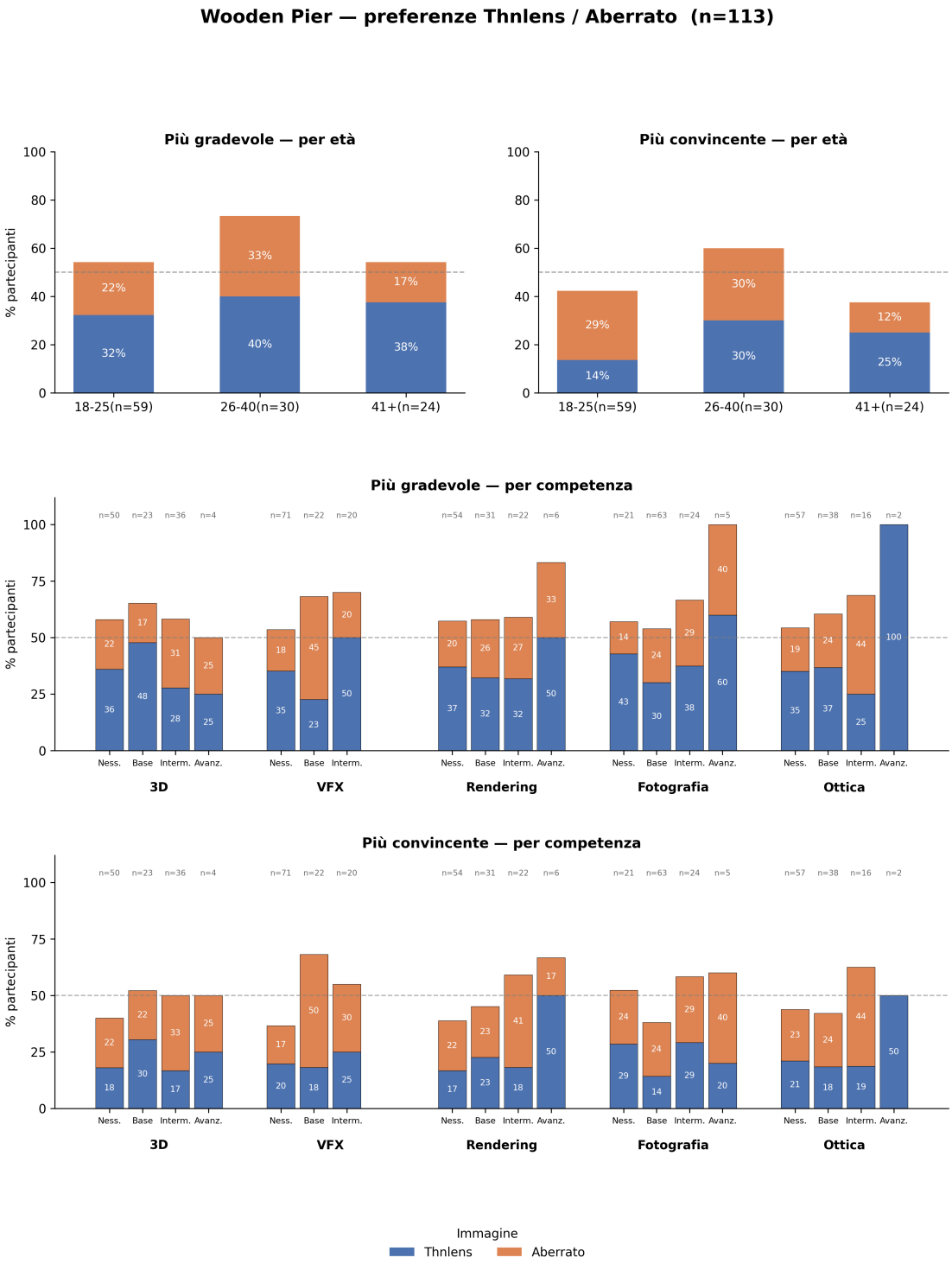


Figura 5.41: Coppia *Wooden pier* — preferenze A/B.

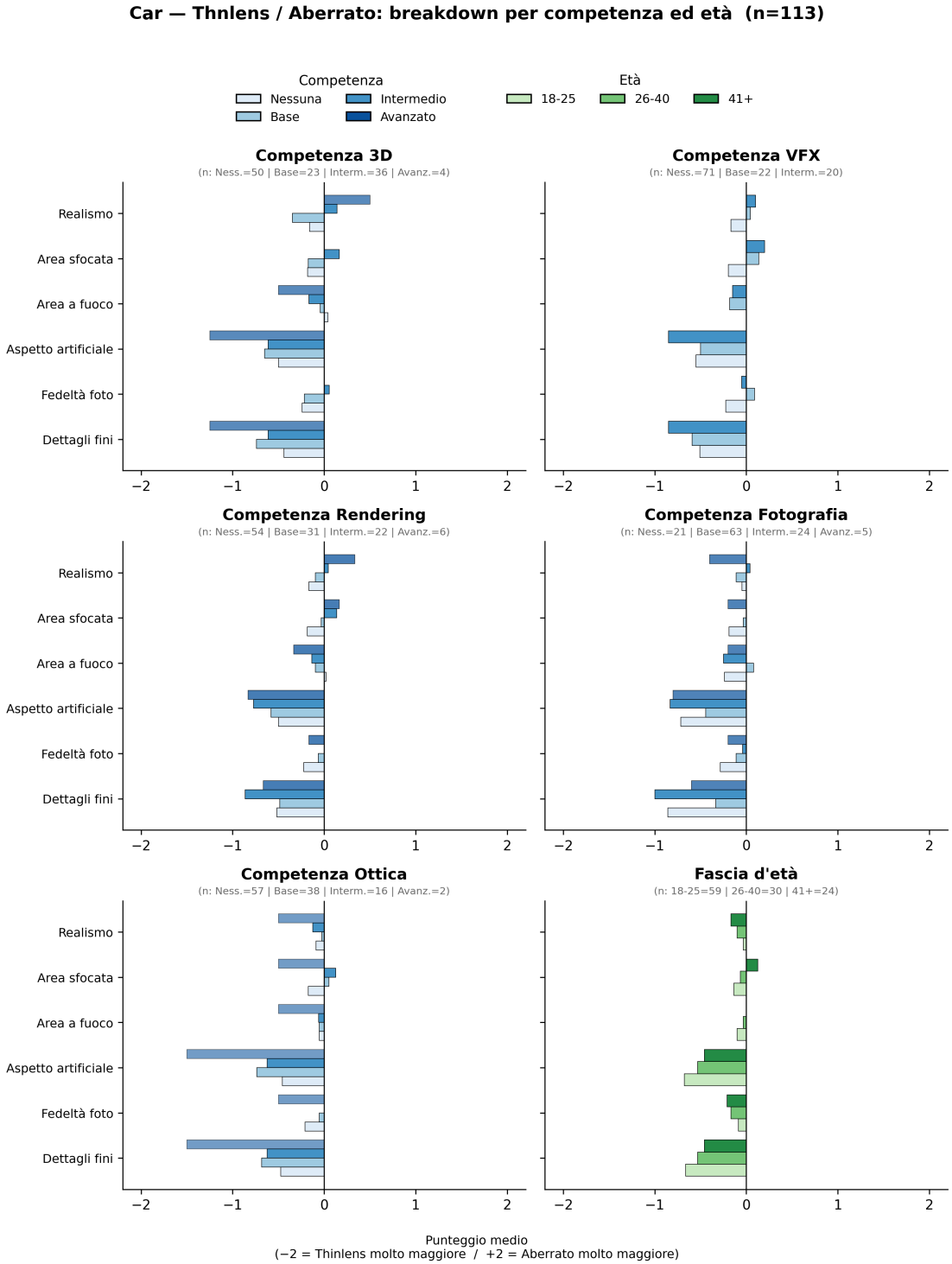


Figura 5.42: Coppia *Car* — Likert competenza | età.

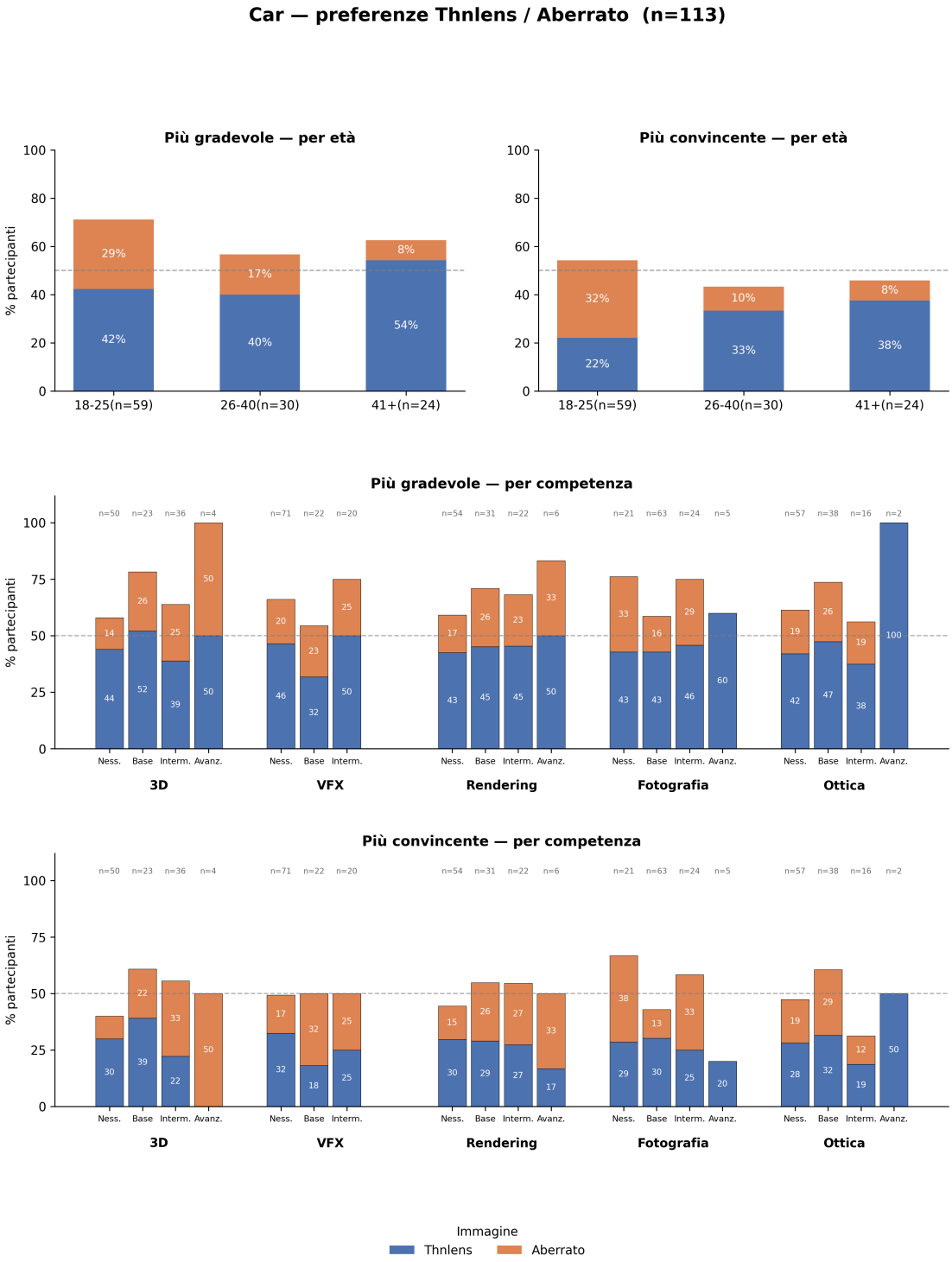


Figura 5.43: Coppia *Car* — preferenze A/B.

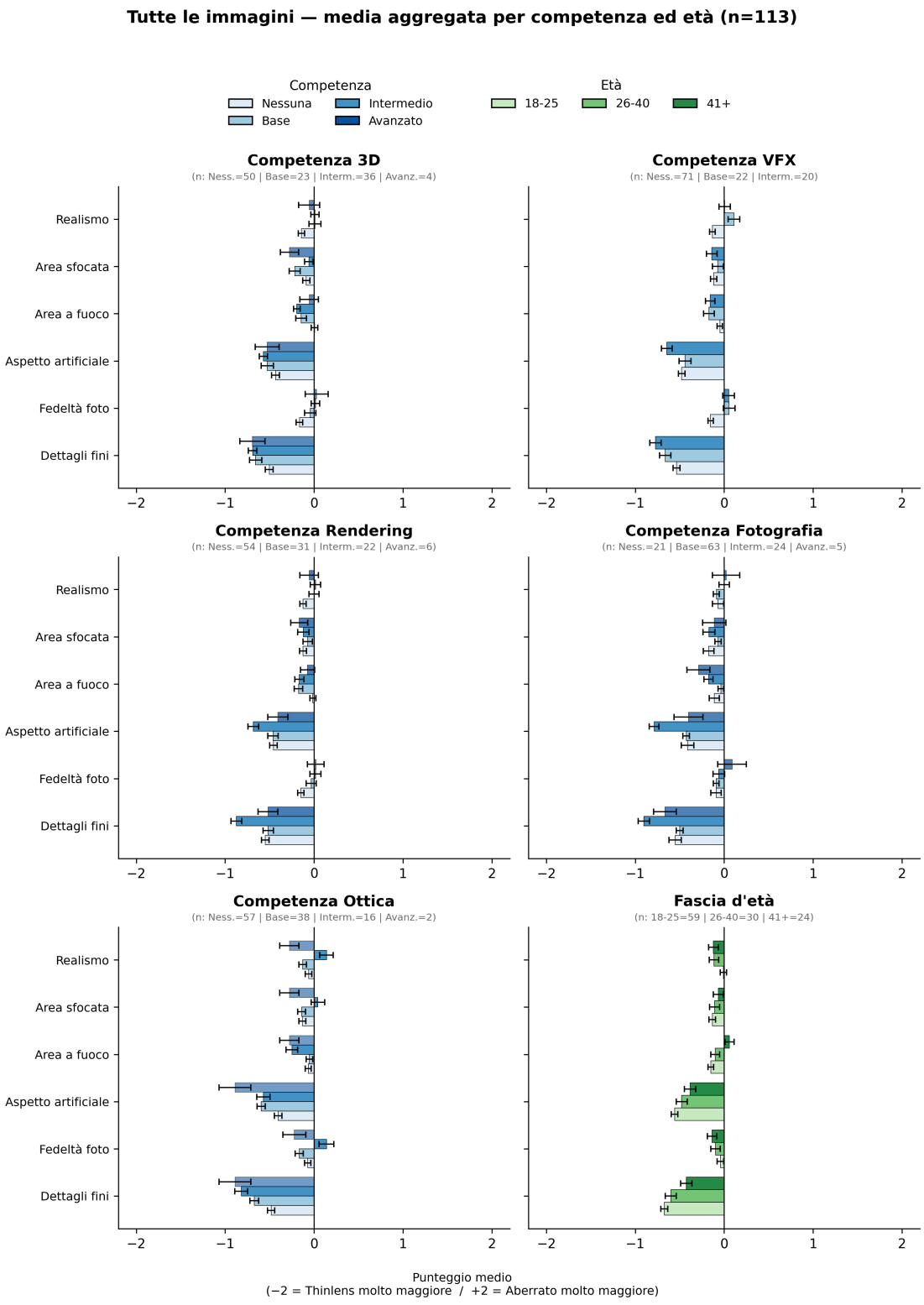


Figura 5.44: Media aggregata per competenza ed età.

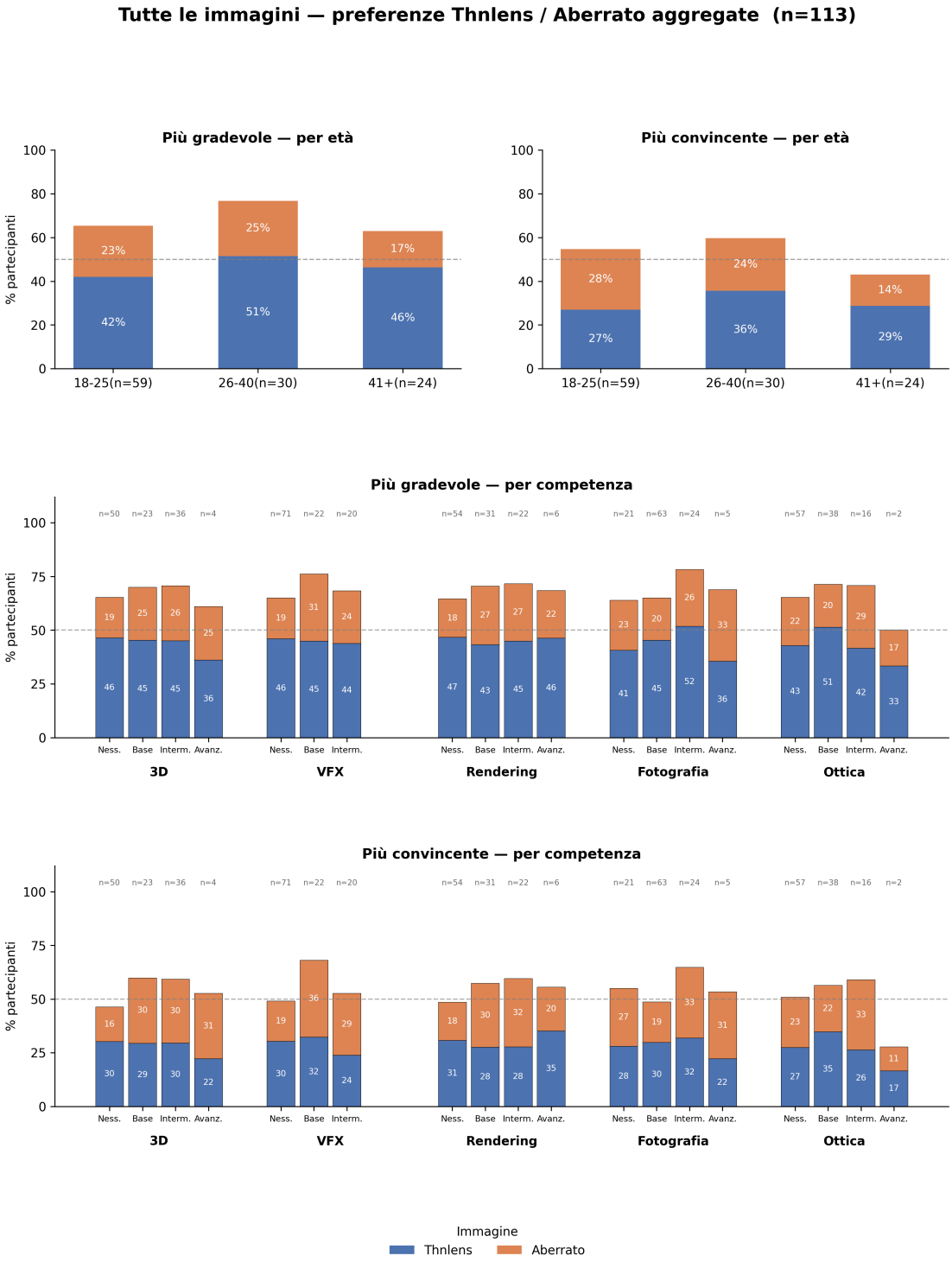


Figura 5.45: Media aggregata — preferenze A/B.

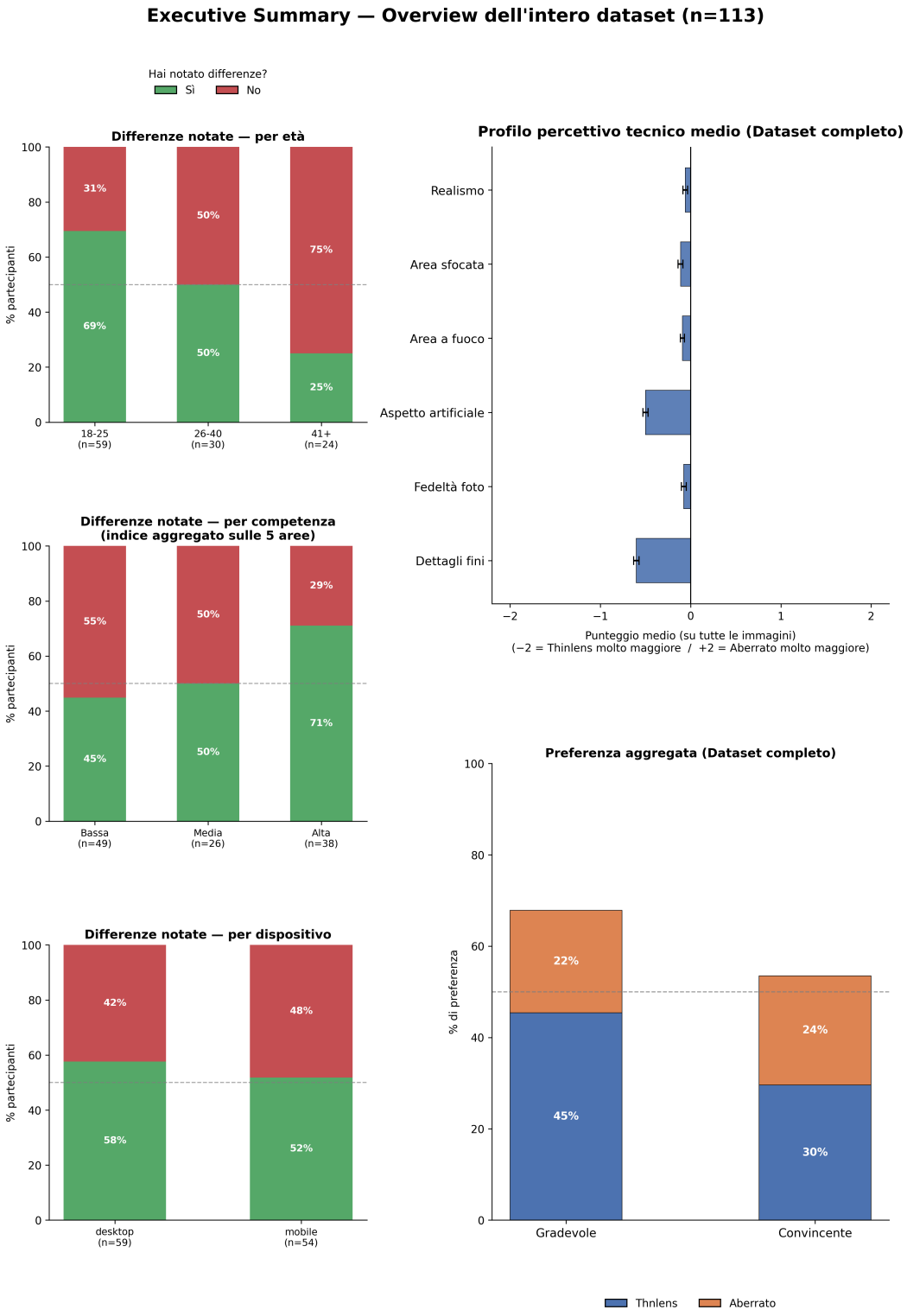


Figura 5.46: Percezione complessiva dell'intero dataset.

5.3 Discussione e Correlazione

Dall'analisi dei risultati aggregati, misurati su scala Likert, emerge una tendenza chiara e sostanzialmente trasversale, indipendente dal livello di esperienza o dall'età anagrafica dei partecipanti. Per quanto concerne la leggibilità dei dettagli fini, il campione indica una percezione di maggiore nitidezza e definizione nelle immagini generate con l'algoritmo standard (Thinlens). Il punteggio medio si attesta infatti intorno al valore di -1 (su una scala in cui -2 indica una forte preferenza per il modello standard). Questo risultato, lungi dal rappresentare una criticità, conferma a livello percettivo l'efficacia del modello vettoriale proposto: l'iniezione coerente di aberrazioni ottiche nel *frustum* introduce fisiologicamente un ammorbidimento dei micro-dettagli, un fenomeno aderente ai cali di risoluzione documentabili nei grafici MTF delle lenti reali.

Un andamento analogo, seppur di entità lievemente inferiore, si riscontra nel parametro relativo alla sensazione di artificialità. Gli utenti percepiscono i render generati con il modello standard come maggiormente artificiali rispetto alle controparti aberrate. Questo dato riveste un'importanza cruciale per la ricerca, poiché fornisce una risposta affermativa al quesito centrale del lavoro: l'introduzione di imperfezioni ottiche fisicamente accurate incrementa l'organicità e la credibilità visiva dell'immagine, rendendola percettivamente più verosimile.

Parallelamente, si osserva un interessante dualismo tra estetica e fotorealismo. Se da un lato l'assenza di difetti rende le immagini standard mediamente più gradevoli alla vista, dall'altro lato i giudizi su quale versione risulti complessivamente più convincente o realistica sono marcatamente più eterogenei. Infine, disaggregando i punteggi, si nota come la sensibilità a queste discrepanze percettive tenda ad assottigliarsi all'aumentare dell'età dei rispondenti, mentre cresca in modo direttamente proporzionale al livello di competenza tecnica dichiarata.

5.3.1 Analisi dei commenti

L'analisi dei commenti liberi mostra diversi pattern coerenti con l'effetto atteso dell'iniezione delle aberrazioni ottiche. Il più netto riguarda la distanza di messa a fuoco del soggetto: nelle scene con soggetto in primo piano — l'automobile e l'escavatore Lego, entrambe a DOF stretto — l'aberrazione viene percepita in modo netto e descritta esplicitamente in termini di fuoco ("leggero fuori fuoco per simulare la non perfezione delle lenti reali", "una a fuoco e una fuori fuoco"), con un rispondente che arriva persino a spiegarla come "upscaling" — una lettura tecnicamente sbagliata ma comunque segno che l'effetto è abbastanza forte da richiedere una spiegazione.

Nel pontile, a distanza di cattura maggiore e DOF nativamente più ampio, lo stesso tipo di iniezione viene invece descritta come marginale o assente, con le due immagini di una coppia giudicate "indistinguibili". Oltre a questo, emergono nei commenti diversi artefatti specifici legati a come l'aberrazione si manifesta otticamente: un rispondente descrive un "effetto di diffusione sulle luci" che produce un "effetto di doppio bordo" attorno a bordi e specular highlights molto luminosi, lo stesso effetto che sul Lego arriva a "mascherare quasi i pochi poligoni nei fori dei singoli pezzi".

Lo stesso rispondente nota anche una riduzione del color banding nelle ombre e una minore visibilità delle riflessioni interne del vetro nella foto del profumo nelle versioni con aberrazione applicata, mentre un altro rispondente segnala una dominante cromatica giallognola percepita in modo consistente su una delle due versioni delle Cornell Box colorate. Questi stessi indizi ottici si riflettono, in negativo, nei giudizi sull'origine AI delle immagini: nitidezza eccessiva e dettagli fini "troppo perfetti" — contorni e ombre "troppo netti per non essere AI", o uno scattering nel vetro giudicato "troppo elevato" — vengono letti sistematicamente come indizio di artificiosità, suggerendo che l'iniezione delle aberrazioni riduca anche il sospetto di generazione sintetica, non solo il punteggio di realismo esplicitamente richiesto dal questionario.

5.3.2 Coerenza tra Metriche Oggettive e Soggettive

L'impianto analitico finora delineato trova una solida validazione incrociando i dati oggettivi, derivati dalle curve MTF calcolate sugli output finali presentati nel questionario (ovvero le immagini sottoposte a *denoising*), con i riscontri percettivi del campione statistico. La divergenza più significativa tra le funzioni di trasferimento della modulazione si registra sistematicamente nel dominio delle alte frequenze spaziali, ovvero l'area spettrale responsabile della risoluzione dei micro-dettagli, della texture e della nitidezza dei bordi.

Questa flessione analitica si correla perfettamente con i risultati quantitativi della scala Likert: la percezione di una minore leggibilità dei dettagli fini nelle immagini aberrate è la diretta traduzione visiva e cognitiva dell'appiattimento della curva MTF registrato nei file esportati. Il decadimento energetico alle alte frequenze, consolidato nell'immagine finale dall'interazione tra la dispersione vettoriale del raggio e l'algoritmo di ricostruzione spaziale, viene decodificato dal sistema visivo umano esattamente come quella sensazione di morbidezza o fuoco imperfetto evidenziata nei commenti liberi.

Tale coerenza tra matematica e percezione si manifesta in modo inequivocabile anche su base contestuale. Nelle scene in cui l'analisi dell'immagine finale

mostrava le alterazioni spettrali più marcate a causa di un *frustum* molto ripido, come in *Lego* e *Car*, gli utenti hanno individuato la variazione in modo esplicito, razionalizzandola come un effetto di lente fisica. Viceversa, dove la curva MTF dell'output filtrato certificava un comportamento strutturalmente invariato rispetto al modello standard, come nel caso del *Wooden Pier*, il campione ha descritto l'impatto dell'aberrazione come trascurabile, arrivando a giudicare le coppie di immagini come "indistinguibili". Questo parallelismo attesta in via definitiva come il calcolo vettoriale delle aberrazioni si traduca in uno stimolo ottico finale fedelmente mappabile sulle metriche percettive umane.

5.3.3 Validazione dell'ipotesi di ricerca

L'insieme dei risultati ottenuti conferma pienamente l'ipotesi alla base di questo studio: l'eccessiva perfezione geometrica del modello a lente sottile rende le immagini innaturalmente sintetiche, ponendo un limite intrinseco al fotorealismo. Come dimostrato, questo ostacolo può essere superato iniettando i difetti ottici a monte del processo di sintesi, integrandoli direttamente nel calcolo vettoriale dei raggi primari.

Sul versante oggettivo, le misurazioni radiometriche e l'analisi spettrale delle curve MTF hanno dimostrato che il modello di *frustum* ibrido, basato sull'estrapolazione tridimensionale dei polinomi di Zernike, è in grado di simulare la reale dispersione energetica del fronte d'onda. Questo approccio genera caustiche volumetriche fisicamente coerenti senza ricorrere a post-elaborazioni bidimensionali in spazio schermo. Il sistema ha inoltre dimostrato una notevole robustezza all'interno di una pipeline di produzione moderna, tollerando fisiologicamente l'intervento degli algoritmi di *denoising* basati su intelligenza artificiale.

Sul versante soggettivo, il responso del campione umano ha certificato il successo percettivo dell'operazione. La ricerca ha evidenziato come il sistema visivo sia intrinsecamente addestrato a decodificare le aberrazioni ottiche come sinonimo di autenticità fotografica. Accettando una misurabile, ma coerente, perdita di microrisoluzione sulle alte frequenze, il modello aberrato riesce a spezzare la sterilità visiva della CGI tradizionale. L'inclusione di difetti ottici strutturati — come coma, astigmatismo e aberrazione sferica — conferisce all'immagine quella matericità indispensabile per colmare, in ultima analisi, il divario percettivo tra sintesi virtuale e ripresa dal vero.

Capitolo 6

Conclusioni

6.1 Sintesi del lavoro svolto

In conclusione, il presente lavoro ha affrontato il problema della simulazione volumetricamente accurata delle aberrazioni del fronte d'onda all'interno del frustum di un sensore fotografico in un ambiente virtuale. Per superare le limitazioni imposte dai modelli tradizionali, è stato strutturato un framework innovativo articolato in due fasi distinte. La prima consiste nel calcolo offline di un volume tridimensionale, generato tramite interpolazione lineare di inclinazioni (slopes) e intercette (intercepts) ricavate a partire dalla FFT di svariate Point Spread Function (PSF) simulate entro un determinato range di defocus prestabilito. La seconda fase ha previsto l'integrazione di tale mappa volumetrica all'interno di un sensore personalizzato, progettato per sfruttare questi dati al fine di deviare fisicamente i raggi luminosi durante il processo di rendering.

6.2 Risposte ai quesiti di ricerca

Le risposte ai quesiti su cui questo lavoro di tesi ha permesso di fare luce, presentati nel capitolo 2 e alla base dello sviluppo di questo progetto, vengono presentate a seguire:

- **RQ1 (Analisi e Phase Retrieval):**
  La ricerca ha dimostrato la possibilità di operare in maniera indipendente rispetto alle *lens prescriptions* proprietarie. Il modello conferma infatti la validità di un approccio inverso basato sull'analisi dei dati: a partire da una singola risposta all'impulso (PSF), è possibile risalire matematicamente ai parametri che ne determinano la morfologia. Sebbene, come già descritto, si operi al momento solo attraverso la costruzione di risposte all'impulso

sintetiche, si è dimostrata la fattibilità dell'estrazione di tali informazioni, ponendo una solida base per una futura implementazione volta all'analisi e alla simulazione di PSF ricavate da fotografie reali.

- **RQ2 (Modellazione Vettoriale Continua):**
  Il limite della discretizzazione spaziale, intrinseco all'approccio iniziale basato sulle LUT tridimensionali, è stato superato. L'adozione di una mappa di raggi, le cui traiettorie sono interpolate su piani ricavati tramite FFT, ha permesso di definire una mappa vettoriale continua, capace di parametrizzare volumetricamente il frustum desiderato in ogni sua coordinata.

- **RQ3 (Sintesi e Rendering):**
  Infine, è stata ampiamente dimostrata la possibilità di integrare tale volume all'interno di un sensore personalizzato. Quest'ultimo è in grado di leggere la struttura dati e utilizzarla come mappa direzionale per deviare fisicamente la traiettoria dei raggi al momento della loro generazione, consentendone il naturale addensamento e la relativa dispersione spaziale.

6.3 Risultati della validazione

L'analisi dell'implementazione e i conseguenti test di validazione consentono di asserire che l'algoritmo proposto riesce a simulare accuratamente l'evoluzione volumetrica delle aberrazioni, mantenendo un onere computazionale in fase di tracciamento del tutto paragonabile a quello dei modelli tradizionali. A livello temporale, l'unico sovraccarico è dettato dalla generazione preliminare dei piani di PSF, la quale, a seconda della risoluzione selezionata, può richiedere svariati minuti. Tuttavia, l'approccio data-driven adottato si rivela estremamente efficiente: qualora si desideri iniettare una specifica combinazione di aberrazioni, tale processo di calcolo preliminare viene effettuato una singola volta. La mappa volumetrica così ottenuta risulta in seguito dinamicamente compatibile con qualsivoglia combinazione di apertura e lunghezza focale impostata in fase di render.

6.4 Limiti dell'implementazione

Nonostante i risultati ottenuti siano promettenti, il modello proposto presenta alcune limitazioni strutturali che delineano chiari percorsi per la ricerca futura. Attualmente, il framework opera esclusivamente tramite l'iniezione sintetica e manuale dei difetti ottici. Inoltre, sebbene l'algoritmo elabori coerentemente le aberrazioni primarie implementate, integrando con precisione difetti quali coma e aberrazione sferica, permane un limite intrinseco legato allo strumento matematico impiegato: i polinomi di Zernike non sono strutturalmente sufficienti per modellare

aberrazioni del terzo ordine quali la curvatura di campo, la distorsione e i difetti cromatici. Per iniettare e simulare tali artefatti fisici, si renderà necessario il ricorso a una strategia differente.

6.5 Sviluppi futuri

In conclusione, i successivi passi di questa ricerca potrebbero esplorare l'estensione dell'architettura per supportare un ventaglio più ampio di aberrazioni, espandendo la base di calcolo attuale. Al contempo, per esplorare le potenzialità del modello oltre i vincoli legati a Mitsuba, si potrebbe ipotizzare la conversione dell'algoritmo in un modulo aggiuntivo per suite di produzione standard, come Blender. Tuttavia, l'orizzonte più ambizioso per questo progetto risiede nella possibilità di emanciparsi dall'inserimento di parametri sintetici. Sarebbe infatti di grande interesse teorico riuscire a strutturare un modulo capace di analizzare scatti fotografici reali, estraendone empiricamente la PSF per tradurla automaticamente nel dominio di Zernike. Qualora realizzato, un simile approccio sancirebbe la genesi di un vero e proprio gemello digitale dell'ottica, chiudendo in via definitiva il divario tecnologico tra la realtà fisica della lente e la sua controparte virtuale.

Bibliografia

[1] Wenzel Jakob et al. *Mitsuba 3 renderer*. Ver. 3.0.1. 2022. URL: https://mitsuba-renderer.org (cit. alle pp. 2, 25).

[2] Eugene Hecht. *Optics*. 5ª ed. Pearson, 2018 (cit. alle pp. 5, 8, 10).

[3] Wikimedia Commons. *File:Convergent lens diagram with focus point.png — Wikimedia Commons, the free media repository*. [Online; accessed 11-July-2026]. 2026. URL: https://commons.wikimedia.org/w/index.php?title=File:Convergent_lens_diagram_with_focus_point.png&oldid=1169609740 (cit. a p. 6).

[4] Joseph W. Goodman. *Introduction to Fourier Optics*. 3rd. Roberts e Company Publishers, 2005. ISBN: 978-0974707723 (cit. a p. 7).

[5] Sidney F. Ray. *Applied Photographic Optics: Lenses and Optical Systems for Photography, Film, Video, Electronic and Digital Imaging*. 3rd. Focal Press, 2002. ISBN: 978-0240515403 (cit. alle pp. 9–15).

[6] Hubert H Nasse. *Depth of field and bokeh*. Rapp. tecn. Carl Zeiss Camera Lens Division, 2010 (cit. a p. 9).

[7] Wikimedia Commons. *File:Spherical aberration 2.svg — Wikimedia Commons, the free media repository*. 2021. URL: https://commons.wikimedia.org/w/index.php?title=File:Spherical_aberration_2.svg&oldid=577091619 (cit. a p. 12).

[8] Wikimedia Commons. *File:Lens-coma.svg — Wikimedia Commons, the free media repository*. 2022. URL: https://commons.wikimedia.org/w/index.php?title=File:Lens-coma.svg&oldid=711622646 (cit. a p. 13).

[9] Wikimedia Commons. *File:Astigmatism.svg — Wikimedia Commons, the free media repository*. 2023. URL: https://commons.wikimedia.org/w/index.php?title=File:Astigmatism.svg&oldid=805655712 (cit. a p. 14).

[10] Vasudevan Lakshminarayanan e Andre Fleck. «Zernike polynomials: a guide». In: *Journal of Modern Optics* 58.18 (2011), pp. 1678–1678. DOI: 10.1080/09500340.2011.633763 (cit. a p. 16).

[11] Wikimedia Commons. *File:Zernike polynomials2.png — Wikimedia Commons, the free media repository*. 2025. URL: %5Curl%7Bhttps://commons.wikimedia.org/w/index.php?title=File:Zernike_polynomials2.png&oldid=1124015489%7D (cit. a p. 16).

[12] Robert J. Noll. «Zernike polynomials and atmospheric turbulence». In: *J. Opt. Soc. Am.* 66.3 (mar. 1976), pp. 207–211. DOI: 10.1364/JOSA.66.000207. URL: https://opg.optica.org/abstract.cfm?URI=josa-66-3-207 (cit. a p. 17).

[13] Henrik. *Ray-Tracing diagram*. Licenza: CC BY-SA 4.0, https://creativecommons.org/licenses/by-sa/4.0. Wikimedia Commons. 2026. URL: https://commons.wikimedia.org/... (cit. a p. 19).

[14] Arthur Appel. «Some techniques for shading machine renderings of solids». In: *Proceedings of the April 30–May 2, 1968, spring joint computer conference* (1968), pp. 37–45. DOI: 10.1145/1468075.1468082 (cit. a p. 19).

[15] Bui Tuong Phong. «Illumination for Computer Generated Pictures». In: *Communications of the ACM* 18.6 (1975), pp. 311–317. URL: https://cir.nii.ac.jp/crid/1573105974986537984 (cit. a p. 20).

[16] Turner Whitted. «An improved illumination model for shaded display». In: *Communications of the ACM* 23.6 (1980), pp. 343–349. DOI: 10.1145/800249.807419 (cit. alle pp. 20, 21).

[17] Robert L Cook, Thomas Porter e Loren Carpenter. «Distributed ray tracing». In: *ACM SIGGRAPH computer graphics* 18.3 (1984), pp. 137–145 (cit. a p. 20).

[18] Jiaze Wu, Changwen Zheng, Xiaohui Hu, Yang Wang e Liqiang Zhang. «Realistic rendering of bokeh effect based on optical aberrations». In: *The Visual Computer* 26 (2010), pp. 555–563. DOI: 10.1007/s00371-010-0459-5 (cit. a p. 22).

[19] Wenzel Jakob, Sébastien Speierer, Nicolas Roussel e Delio Vicini. «DR.JIT: a just-in-time compiler for differentiable rendering». In: *ACM Transactions on Graphics* 41.4 (2022), pp. 1–19. ISSN: 1557-7368. DOI: 10.1145/3528223.3530099. URL: http://dx.doi.org/10.1145/3528223.3530099 (cit. a p. 25).

[20] Andrew Kensler. «Correlated Multi-Jittered Sampling». en. In: (2013) (cit. a p. 26).

[21] J. R. Fienup. «Phase retrieval algorithms: a comparison». In: *Appl. Opt.* 21.15 (ago. 1982), pp. 2758–2769. DOI: 10.1364/AO.21.002758. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-21-15-2758 (cit. a p. 35).

[22] Pauli Virtanen et al. «SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python». In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2 (cit. a p. 36).

[23] Attila T. Áfra. *Intel® Open Image Denoise*. https://www.openimagedenoise.org. 2026 (cit. a p. 40).

[24] Imatest LLC. *ISO 12233 — Resolution and SFR*. 2024. URL: https://www.imatest.com/imaging/iso-12233/ (visitato il giorno 06/06/2026) (cit. a p. 41).

[25] Wikimedia Commons. *File:Sharpness chart02 f22.jpg — Wikimedia Commons, the free media repository*. 2025. URL: %5Curl%7Bhttps://commons.wikimedia.org/w/index.php?title=File:Sharpness_chart02_f22.jpg&oldid=1047089369%7D (cit. a p. 42).

[26] Blender Foundation. *Blender - a 3D modeling and rendering package*. Version 4.5.5. 2025. URL: https://www.blender.org (cit. a p. 42).

[27] Baptiste Nicolet e contributors. *Mitsuba Blender Add-on*. Licenza: BSD 3-Clause. 2022. URL: https://github.com/mitsuba-renderer/mitsuba-blender (cit. a p. 42).

[28] Benedikt Bitterli. *Rendering resources*. 2016. URL: https://benedikt-bitterli.me/resources/ (cit. a p. 47).

[29] *BlenderKit Database*. https://www.blenderkit.com/ (cit. a p. 47).

[30] Cornell University Program of Computer Graphics. *The Cornell Box*. https://benedikt-bitterli.me/resources/. Modello 3D di riferimento per test di radiosity e rendering. Convertito e ottimizzato da Benedikt Bitterli. 1984 (cit. a p. 50).

[31] Cornell University e Benedikt Bitterli. *Cornell Box with Glass Elements*. https://benedikt-bitterli.me/resources/. Variazione della scena originale per l'analisi di materiali dielettrici e riflessioni complesse. 2016 (cit. a p. 51).

[32] Cornell University e Benedikt Bitterli. *Water Cornell Box*. https://benedikt-bitterli.me/resources/. Variazione della scena originale con volume d'acqua per lo studio dell'accumulo di caustiche. 2016 (cit. a p. 52).

[33] Heinzelnisse. *Lego Digger*. https://benedikt-bitterli.me/resources/. Modello 3D ad alta densità geometrica ottimizzato per il rendering. Archivio Rendering Resources. 2016 (cit. a p. 53).

[34] Free poly. *Room with characteristics*. https://www.blenderkit.com/asset-gallery-detail/e2ced268-fe10-4c2e-bfba-ce5fd6377313/ (cit. a p. 54).

[35] Srabon Arafat. *Perfume Bottle Scene*. https://www.blenderkit.com/asset-gallery-detail/9542bb6e-2547-43da-9349-d739256fef2b/ (cit. a p. 55).

[36] sueti begum. *Winter forest volumetric scene*. `https://www.blenderkit.com/asset-gallery-detail/c2d22f8b-2844-4b0f-be79-f11be5a365fc/` (cit. a p. 56).

[37] Jernej Kalin. *Pier on the sea scene*. `https://www.blenderkit.com/asset-gallery-detail/098efcbe-ae8d-4765-9643-acb470a9e63a/` (cit. a p. 57).

[38] Polara 3D. *Car in highway Scene*. `https://www.blenderkit.com/asset-gallery-detail/53c5f63d-f238-4965-b7fd-a6d8a2682af4/` (cit. a p. 58).