October | 2026
Vol. 135 No. 6

MI

MOTION IMAGING JOURNAL

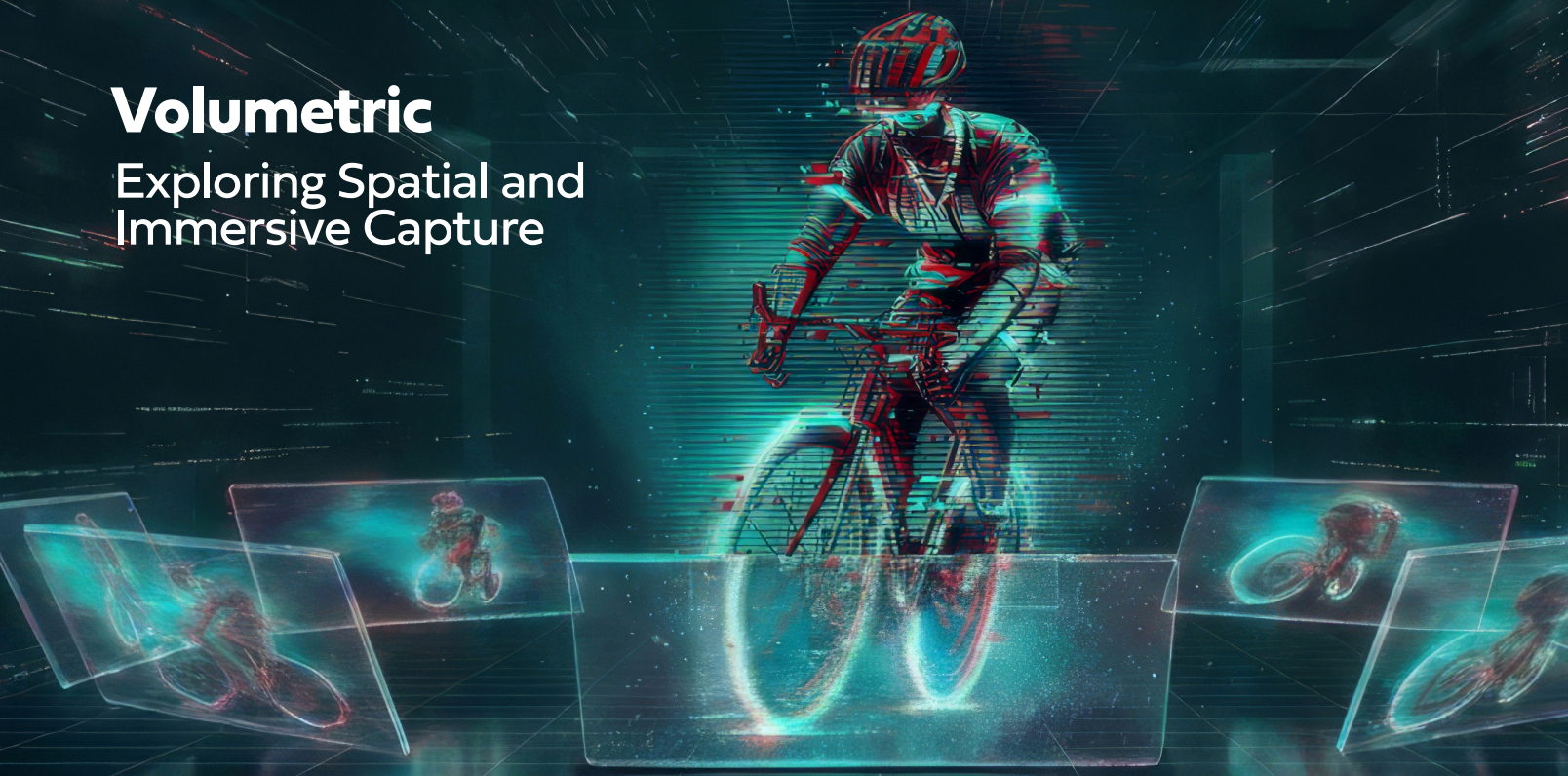
SMPTE

INSIDE:

- Compact 8K × 8K Camera; Imaging the Immersive Era
- Leveraging Private 5G for Live TV Production: A Case Study from the 2025 World University Winter Games
- Contextual AR in the Wild: Enhancing Live Cycling Broadcasts
- Optimizing 180° Stereoscopic Cinematography for the Apple Vision Pro
- Spectral Rendering Techniques to Improve Color Matching in Virtual Production
- Agfacolor at 90: The Race for Color and the Technology that Endured

Volumetric

Exploring Spatial and Immersive Capture



COVERING EMERGING TECHNOLOGIES FOR THE GLOBAL MEDIA COMMUNITY

Fix it in Live.

When live feeds arrive with different timing, formats, or signal characteristics, they require normalization in real time. FS8 handles the synchronization and processing; OG-FiDO-4 carries SDI over fiber.



FS8

Hybrid SDI/IP Media Processor

The new FS8 frame-sync supports multi-channel processing with SDI and SMPTE ST 2110 connectivity. Advanced features include an easy-to-configure user interface, HDR conversions, IP gateway, Dante audio, low-latency processing, and much more.



OG-FiDO-4

4-Channel openGear SDI/LC Fiber Converters

The OG-FiDO-4 family of SDI to/from fiber openGear cards is a new flexible architecture with 9 variations, including 12G single mode, 3G single mode, and 3G multi-mode LC fiber solutions.

www.aja.com/whats-new

AJA®

SMPTE OFFICERS

OFFICERS

PRESIDENT, 2025-26
Richard Welsh

PAST PRESIDENT, 2025-26
Renard Jenkins
I2A2 Technologies, Studios & Labs

EXECUTIVE VICE PRESIDENT, 2025-26
Eric Gsell
Dolby Laboratories

STANDARDS VICE PRESIDENT, 2026-2027
Raymond Yeung
Amazon Studios

EDUCATION VICE PRESIDENT, 2025-26
Polly Hickling
Atos

MEMBERSHIP VICE PRESIDENT, 2026–27
Chris Lapp
Cisco

FINANCE VICE PRESIDENT, 2026–27
Troy English
Ross Video, Ltd.

SECRETARY/TREASURER, 2026-27
Lisa Hobbs
Mediakind

EXECUTIVE DIRECTOR
Sally-Ann D'Amato
SMPTE Headquarters

DIRECTORS

STANDARDS
Steve LLamb

Florian Schleich

Frederick G. Walls

EDUCATION
Wim Van Roy

Carys Hughes

MEMBERSHIP
Jason Pepino

John Shike

Zandra Clarke

GOVERNORS

ASIA/PACIFIC REGION
Paul Whybrow, 2026–27
Varda Creative Leadership

Tony Ngai, 2025-26
Society of Motion Imaging Ltd.

CANADIAN REGION
Tony Meerakker, 2026–27
Meer Tech Systems

Jonathan Jobin, 2025-26
Grass Valley

USA - CENTRAL REGION
William T. Hayes, 2025-26

USA - EASTERN REGION
Sara Seidel, 2026–27
Riedel Communications

Dover Mundt, 2025-26
Riedel Communications

USA - HOLLYWOOD REGION
Allan Schollnick, 2025-26
AJS Media, Inc.

Marc Zorn, 2025-26

EMEA, CENTRAL & SOUTH AMERICA REGION
Fernando Bittencourt, 2025-26
FB Consultant

Dagmar Driesnack, 2026–27
Appear

USA - NEW YORK REGION
Jeff Cohen, 2025-26
Meta Martis

USA - SOUTHERN REGION
Andre Normil, 2026–27

U.K. REGION
Chris Johns, 2025-26
Sky UK

USA - WESTERN REGION
Jeffrey F. Way, 2025-26
Open Drives

SECTION CHAIRS

ATLANTA
Zandra Clarke
Warner Bros. Discovery

CHICAGO
Chris Gazdic
Wesco

DETROIT
Timothy O'Brien
Center of Academic Innovation

FLORIDA
Anne-Laure Treny
NBC Universal

HOLLYWOOD
Juan Reyes
Tech Align Group

HONG KONG
Peter Wong
A&S Broadcast Limited

INDIA
Ujwal Nirgudkar

ITALY
No Section Chair

MONTREAL/QUEBEC
Denis Bonneau
Technologie Optic.ca

NEW ENGLAND
Martin P. Feldman
Unique Scientific, Inc.

NEW YORK
John Gallagher
BMCC/CUNY

OCEANIA
Evdoxia Brownlow
Tabcorp

OHIO
Jeffrey Johnson
Great Lakes Science Center

PHILADELPHIA
Mark Mullen
Millersville University

PACIFIC/NORTHWEST
No Section Chair

POLAND
Jaroslaw Gretkiewicz
Dolby Poland Sp. zo.o

ROCKY MOUNTAIN
Sean Richardson
Dolby Laboratories

RUSSIA
Suspended

SACRAMENTO
Ian Godfrey
TSL Products

SAN FRANCISCO
Shane Ruggieri
Dolby

TEXAS
Trevor Van Dahm
MELISSA ISD

TORONTO
Roy Folkman
Cinesys

UNITED KINGDOM
John Ellerton
BT Media & Broadcast

WASHINGTON, DC
Jonothan Solomon
TVU Networks

STUDENT CHAPTERS BY SECTION

ATLANTA
Clark Atlanta University
Atlanta, GA

Southern University and A&M-College
Baton Rouge, LA

AUSTRALIA
Auckland University of Technology
Auckland, New Zealand

TAFE NSW St. Leonards Film and TV
NSW, Australia

DALLAS
Baylor University
Waco, TX

CHICAGO
Purdue University
West Lafayette, IN

FLORIDA
Full Sail University
Winterpark, FL

University of Central Florida
Orlando, FL

GERMANY
Universities of Applied Sciences Germany (HdM / HSHL / HSRM)

- Hochschule der Medien | Jan Fröhlich
- Hamm-Lippstadt University of Applied Sciences | Stefan Albertz
- RheinMain University of Applied Sciences | Wolfgang Ruppel

HOLLYWOOD
American Film Institute
Hollywood, CA

California State University Northridge
Los Angeles, CA

California State University Long Beach
Long Beach, CA

California State University Los Angeles
Los Angeles, CA

Citrus College
Glendora, CA

Chapman University
Orange, CA

Los Angeles City College
Los Angeles, CA

Media Arts Collaborative Charter School
Albuquerque, NM

Pasadena City College
Pasadena, CA

USC University of Southern California
Los Angeles, CA

University of Nevada, Las Vegas
Las Vegas, NV

HONG KONG
Hong Kong Design Institute (HKDI)
Hong Kong

Hong Kong Institute of Vocational Education (IVE)
New Territories, Hong Kong

INDIA
Whistling Woods International
Mumbai, India

Indian Institute of Creative Technologies
Mumbai, India

NEW ENGLAND
Fitchburg State University
Fitchburg, MA

Stonehill College
Easton, MA

NEW YORK
Barnard College
New York, NY

Borough of Manhattan Community College
New York, NY

Bronx Community College
Bronx, NY

Kingsborough Community College
Brooklyn, NY

New York City College of Technology
Brooklyn, NY

Rochester Institute of Technology
Rochester, NY

Rowan University
Glassboro, NJ

William Patterson University
Wayne, NJ

OTHER INTERNATIONAL
Thomas More BE
Belgium

PHILADELPHIA
Millersville University
Millersville, PA

SAN FRANCISCO
De Anza College
Cupertino, CA

City College of San Francisco & San Francisco State University
San Francisco, CA

Vallejo Student Chapter
Vallejo, San Francisco

TEXAS
University of North Texas
Denton, TX

UNITED KINGDOM
Birmingham City University
Birmingham, U.K.

Ravensbourne
London, U.K.

Southampton Solent University
Southampton, Hampshire, U.K.

University of Salford
Manchester, U.K.

University of Surrey
Guildford, U.K.

WASHINGTON, DC
George Mason University
Fairfax, VA

Loyola University
Baltimore, MD

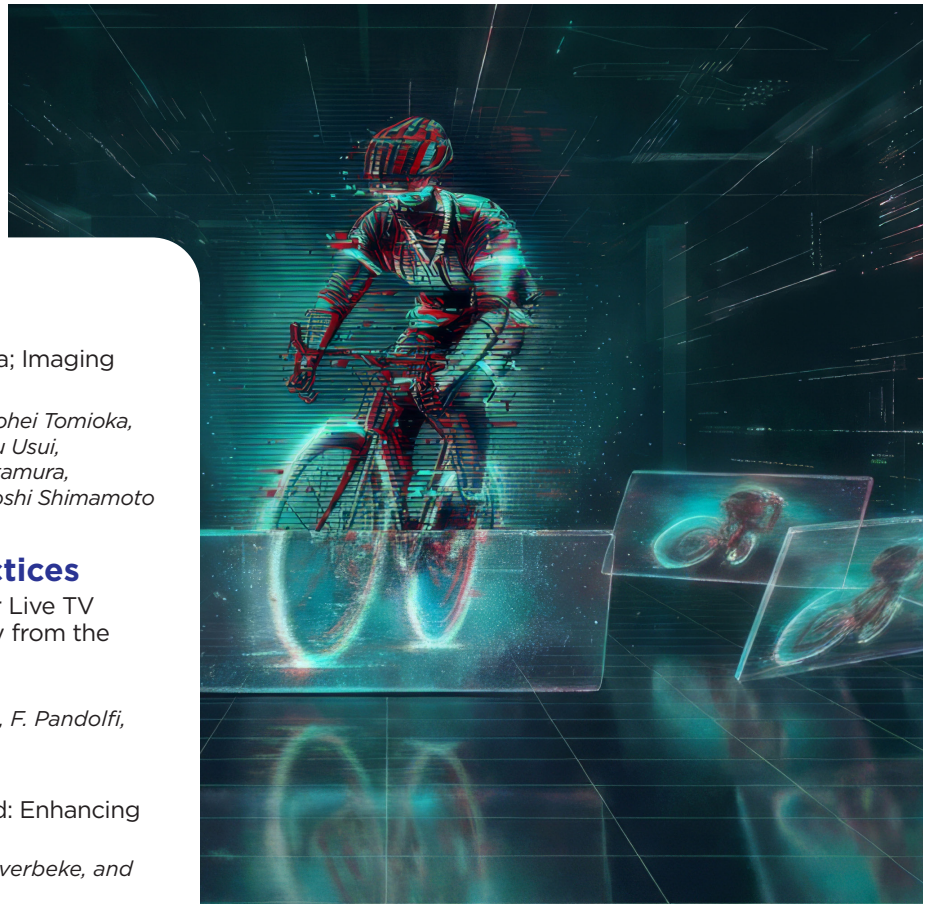
Montgomery College
Rockville, MD

BOARD OF EDITORS

Jaclyn Pytlarz, Chair
Ioan Allen
Harvey Arnold
Brian Bedard
John Belton
Steve Bilow
V. Michael Bove
Wayne Bretl
John Brooks
Dominic Case
Curtis Chan
Greg Coppa
Don Craig
Hamidreza Damghani
Michael A. Dolan
Thomas Edwards
John Ferder
Norbert Gerfelder
Brad Gilmer
Keith Graham
Eric Gsell
Randall Hoffner
Tomlinson Holman
Fred Huffman
Kim Iannone
Devanshi Kotak
Scott Kramer
Sara Kudrle
W. J. Kumeling
Michael Liebman
Peter Ludé
Ian MacSpadden
Andrew H. Maltz
Tets Maniwa
Sean McCarthy
Catherine Meininger
Teju Mulagada
Pavel Novikov
James E. O'Neal
Mark O'Thomas
Karl Paulsen
Kylee Peña
Glen Pensinger
Liz Pieri
Charles Poynton
Adam Schadle
Mark Schubin
Gavin Schutz
Thomas A. Scott
Neil Shaw
James Snyder
Steve Storozum
Stephen A. Stough
Simon Thompson
Paul Thurston
Vibhoothi
J.B Watson
William Y. Zou

MOTION IMAGING JOURNAL

VOLUME 135 | NUMBER 6 | OCTOBER 2026



Technical Papers

16 Compact 8K × 8K Camera; Imaging the Immersive Era
Kodai Kikuchi, Akira Honji, Kohei Tomioka, Tetsuya Hayashida, Takenobu Usui, Toshie Hiroshima, Kazuya Kitamura, Takayuki Yamashita, and Hiroshi Shimamoto

Applications/Practices

23 Leveraging Private 5G for Live TV Production: A Case Study from the 2025 World University Winter Games
G. Stante, A. Lucco Castello, F. Pandolfi, and L. Vignaroli

31 Contextual AR in the Wild: Enhancing Live Cycling Broadcasts
W. Clinckemaillie, J. Vanhaeverbeke, and S. Verstockt

38 Optimizing 180° Stereoscopic Cinematography for the Apple Vision Pro
Stanley Lin, Nathan Fairchild, Spencer Lin, Akaash Tumuluri, Jordan Halsey, and Scott Fisher

Off Topic

49 Spectral Rendering Techniques to Improve Color Matching in Virtual Production
Vlad Simion, David Long, Flip Phillips, and Kris Murray

60 Agfacolor at 90: The Race for Color and the Technology that Endured
Michael Zink

1545-0279/26©2026SMPTE

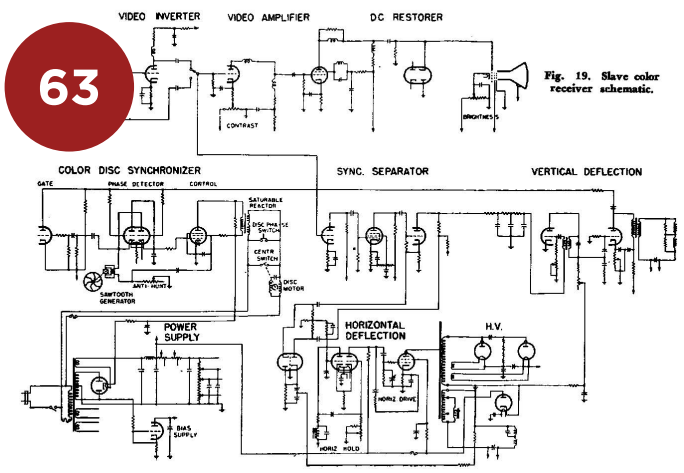


Fig. 19. Slave color receiver schematic.

DEPARTMENTS

1 SMPTE Officers
4 Corporate Members
6 From the Top: Protecting Our Purpose, Building Our Future: Why SMPTE Matters
12 SMPTE Fellows and Awards Honorees
14 Spotlight: Wim Van Roy
63 Retrospective: Historical Briefs from Past Issues



The Society of Motion Picture and Television Engineers, Inc., 50 Main Street, Suite 1000, White Plains NY, 10606

EXECUTIVE
Polly Hickling
Education Vice President

Jaclyn Pytlarz
Chair/Board of Editors

John Belton
Thomas Edwards
Catherine Meininger
James E. O'Neal
Associate Editors

Sally-Ann D'Amato
Executive Director

JOURNAL STAFF
Maja Davidovic
Director of Education
education@smpte.org

Dianne Purrier
Managing Editor, Digital and Print Publications
journal@smpte.org

Anthony Catalano
Executive Account Manager, Advertising Sales & Sponsorships
sales@smpte.org

Jeffrey Victor
Advertising Sales North America
jvictor@smpte.org

SMPTE Motion Imaging Journal (ISSN: 1545-0279) published 7 times per year by the Society of Motion Picture and Television Engineers, Inc., 50 Main Street, Suite 1000, White Plains NY, 10606. Periodicals postage paid at WhitePlains, NY, and at additional mailing office.

Copyright © 2026 by the Society of Motion Picture and Television Engineers, Inc. No part of this issue may be reproduced without written permission of the publisher. Statements and opinions expressed in articles or editorials in the SMPTE Motion Imaging Journal are expressions of contributors and do not necessarily represent the policies or opinions of the Society. A complete list of SMPTE membership types and benefits is available at www.smpte.org. Advertisements appearing in this publication are the sole responsibility of the advertiser. Subscription rates: SMPTE Members, $35 USD, included in membership dues; print subscription: $1,000 (shipping included); SMPTE owns the copyrights to all published material, unless otherwise stated. *SMPTE Motion Imaging Journal* is registered with Canada Post: CPM# 41003019.

POSTMASTER: Send change of address to *SMPTE Motion Imaging Journal*, 50 Main Street, Suite 1000, White Plains NY, 10606. All rights reserved.
PRINTED IN THE USA

DOI: 10.5594/JMI.2026/EJMX6968
Date of publication: 1 October 2026

CORPORATE MEMBERS

DIAMOND LEVEL

Apple
Amazon AWS
Blackmagic Design, Inc.
CBS/Paramount Pictures
Disney/ABC/ESPN
Dolby Laboratories
Fox Corporation
Google
Ross Video
Sony Electronics, Inc.
Telstra Corporation

GOLD LEVEL

Adobe Systems
AVID Technology, Inc.
Bloomberg
British Telecommunications, PLC
Deluxe
Diversified
NBC Universal
Rohde & Schwarz GmbH & Co. KG
Sudwestrundfunk/ARD

SILVER LEVEL

AI Media
AJA Video Systems Inc.
AMD
Canon USA, Inc
Dell
European Broadcasting Union
Fuse Technical Group
Globo
IMG Media Ltd.
Interdigital Communications
Library of Congress
Novastar
Oracle
Panasonic Corporation
Red Digital Camera
The Studio - B&H

BRONZE LEVEL

Academy of Motion Picture Arts & Sciences
Appear AS
Applied Electronics Ltd.
Arista Networks
Arqiva Ltd.
Belden, Inc.
Bridge Technologies
Canare
Canon, Inc.
Carl Zeiss AG
CBC Radio Canada
Channel 4 Television
Christie Digital Systems
Cisco
Colorfront
disguise
Evertz
Fraunhofer
Grass Valley, Inc.
Imagine Communications
Intel Corporation
Key Code Media
Koninklijke Philips NV
Leader Electronics Corporation
Luminex Network Intelligence
Matrox Graphics Inc.
Media Links Co., Ltd.
Meinberg-Funkuhren GmbH & Co.
Meta Platforms
MLB Advanced Media
Motion Picture Solutions
NEC Corporation
NEP Group
Net Insight
Nevion
NHK (Japan Broadcasting Corp.)
NFL Films
Nvidia
ORF - Austrian Broadcasting Corporation
Pixelogic
Quortex
Riedel Communications
SiriusXM Radio, Inc.
Tag VS
Telestream, Inc.
Tunnel Post
Universal Pictures
WT Vision
Xperi
Zixi

SMALL BUSINESS LEVEL

Adeas, B.V.
Alphacord Media Group
Analog Way
ARRI Inc.
Astrodesign, Inc.
ATEME SA
ATOS
Barco
BBC Future Media
BFE Studio and Medien Systeme GmbH
Boland Communications
Brompton Technology
Castlabs GmbH
Chesapeake Systems
CineCert
Cinergy GmbH
Cobalt Digital
ComScore Inc.
CST (Comission Superiere Technique de l'image et du son)
Dalet
DekTec America
Deltacast.tv
Eikon Group Co.
Glookast
Imagica Entertainment Media Services, Inc.
InSync Technology Ltd.
Intelligent Wave Inc.
Internet Initiative Japan
IntoPIX
LAWO
Maxson Network Limited
MediaTek Inc.
Megapixel VR
Muybridge
Nemal Electronics Intl. Inc.
Netgear AV
NTT Network Innovation Labs
Plus 24
Portrait Displays
Qencode
Qualabs
Qvest GmbH
RAVN Networks
Seiko Epson Corp.
Sencore, Inc.
6P Color Inc.
Sonic Origin
Starfish Technologies
Sutro Tower, Inc.
Tajimi Electronics Co., Ltd.
Tamura Corporation
Techex Ltd.
Tedial
Telos Alliance
Utah Scientific
Video Clarity

INDUSTRY PARTNERS

IBC
ITU-R
NAB
NISO

CONSULTANT LEVEL

Camplex
China Research Institute of Film Science and Technology
Decimator Design Pty. Ltd.
TSL Professional Products Ltd.

1545-0279/26©2026SMPTE

yellobrik®

like nothing else



- 12G SDI
- 3G SDI
- 1.5 SDI
- 10GbE Ethernet
- 1GbE Ethernet
- Serial RS 232/422
- GPI
- MADI
- Analog Audio
- HDMI
- Analog Sync
- Composite Video
- AES Audio

Designed to support modern AV fiber transport workflows

From simple point to point applications to fully integrated CWDM systems with up to 18 channels over one bi-directional fiber link

Simple, clean, professional system installations

- 19” rack frame holds up to 14 fiber yellobriks
- Central power - redundant protection
- Matched CWDM multiplexers
- Front-access fiber connections in recessed racks



Not sure where to start?

Let us help configure a system for you.

info@lynx-usa.com

www.yellobrik.com

LYNXTechnik USA®

GERMAN QUALITY SINCE 2003

www.lynx-usa.com
661-251-8600

Protecting Our Purpose, Building Our Future: Why SMPTE Matters

BY SALLY ANN D'AMATO



In this month's Journal message, I want to do something a little different and maybe even a bit radical — I want to level with you.

I'm a big believer in honesty and transparency, and I've tried to bring both to my role as Executive Director. From the beginning, I've made it a priority to be candid with our Board, our members, our partners, and anyone else I speak with about SMPTE—not only about what we are accomplishing, but also about the challenges we face.

So, if you've heard the whispers or conjecture about SMPTE—that we're in danger of becoming irrelevant, or that we're facing financial challenges—I'm here to tell you there is truth behind those concerns.

Our industry is going through enormous change. Budgets are tight. Companies are consolidating and reevaluating where they invest. And, of course, this trickles down to nonprofit organizations like SMPTE. As an association, we rely primarily on membership dues, sponsorships, and philanthropic support to fund our work. Those are often among the first expenses to be cut from budgets in difficult times.

But I don't believe the answer is to retreat. Organizations don't remain relevant by protecting the way things have always been done. They remain relevant by protecting their purpose while changing how they fulfill it. SMPTE has to adapt or be left behind.



And when I become overwhelmed by the challenges and the road ahead, I remember that my longtime vocal coach used to tell me that a house isn't built all at once—it's built brick by brick. Sometimes I need to remind myself to stop and take stock of just how much adapting we've already done.

In just a couple of short years, we've replaced a member portal with limited functionality with a new system that provides a better member experience, and which continues to evolve. We've launched new training courses and workshops and, with generous support from our partners, created a practical lab for hands-on SMPTE ST 2110 learning.

We've opened access to the entire SMPTE Standards Library, with support from our Diamond members and donors, making the technical foundation our industry relies on more accessible to everyone.

1545-0279/26©2026SMPTE

We've revised our Corporate Membership levels and benefits and introduced our Corporate Member Engagement Specialist program to provide more personalized, dedicated support to our sustaining members.

We've launched SMPTE's Substack, creating a more accessible and conversational channel for sharing technical knowledge, industry perspectives, and Society content.

We've strengthened our partnerships and collaborations with allied industry organizations, recognizing that the challenges facing our industry are too big for any one organization to tackle alone.

We've launched SMPTE's first major fundraising campaign in a decade. We are working to build a culture of philanthropy that can sustain the Society for years to come, including developing a planned giving program.

And yes, we've introduced the SMPTE Color Bears—because sometimes we need to remind ourselves that it's okay to have a little fun!

As I look ahead to all we'd like to accomplish over the next three to five years, I am genuinely excited by the possibilities. Our strategic priorities include creating new pathways and tools for standards development; expanding accessible, practical education and professional development for learners at every stage of their careers; building a robust digital knowledge platform; strengthening mentorship and community; establishing SMPTE as a leader in trusted AI for media and entertainment; and investing in the organizational capacity necessary to make all of that possible.

And then I remember one of my catchphrases: "Anything is possible... with money."

SMPTE exists to support this industry through standards, education, and community. But today, we need the industry to support SMPTE—not simply to keep doing what we have always done, but to evolve to meet what comes next.

In these uncertain times, I've been thinking about the "what ifs." What if SMPTE closed its doors? What if SMPTE could no longer provide conferences, the *Motion Imaging Journal*, courses and training, or even the infrastructure that supports standards development?

Dr. George W. Arnold, former ISO Vice President–Policy, said: "Standards provide the 'invisible fabric' that allows modern society to function." For more than a century, SMPTE has helped weave that invisible fabric for media and entertainment. When standards work, most people never think about them. The picture appears. The sound plays. Systems connect. Content moves around the world. Technologies created by different companies simply work together. That is the point. But what happens when that framework isn't there?

What happens when proprietary technologies can't work together? What happens when there is no neutral place for competitors to come together to solve shared technical problems?

And SMPTE's role extends beyond standards. What happens to the communities that have been built around this Society? What happens when those professional networks aren't there to connect and support our industry? What happens when the mentorship, education, and outreach that helped generations of people find their way into careers in media technology are no longer there to make those often-invisible career paths visible?

Those questions are bigger than SMPTE as an organization. They are questions about what kind of industry we want to build for the future.

So I reached out to our members and asked them to tell me their stories. How did they become involved with SMPTE? How have their careers been impacted? What might have been different if SMPTE hadn't been there? And perhaps most importantly: Why does SMPTE matter—and why is it worth supporting?

These are their stories.



Fernando Bittencourt, Past President of SET (Brazilian Society of TV Engineers)
Brazil

A Lifelong Connection

I have been a member of SMPTE for nearly 50 years. Through this organization, I have learned about the latest advances in television, film, and, more recently, media as a whole.

Driving Technological Progress

SMPTE has played a decisive role in the industry's transition, from analog to digital, composite video to component video, and SDTV to HDTV and beyond. More recently, from SDI to IP and from SDR to HDR.

Throughout these changes, SMPTE has set the standards that guided the industry and educated the market through technical papers, its journal, conferences, and study groups.

Long live SMPTE!



Laquie TN Campbell, M&E and Emerging Tech Strategist
United States (New Mexico)
SMPTE gave me more than a professional network. It provided me with a community where authentic relationships grew into opportunities, leadership, and a broader sense of where I belonged in this industry. Over time, I went from learning from people who welcomed me into those rooms to helping shape educational programming and making a point to be that same welcoming presence for someone walking into the room for the first time. I want new faces to know there is a place for them in media technology and to see people they can relate to already participating, leading, and helping shape its future. Without SMPTE, my career might have been narrower, but more importantly, I would have missed the opportunity to help widen that path for others.



Jeff Cohen, Partner, Jeff at Cohen Reps
United States (New York)
I've been a member of SMPTE for my entire working life, and I honestly can't imagine my career without it.

SMPTE turned my working life into a profession, not just a succession of jobs. Over more than 40 years, it has given me far more than technical knowledge. It has given me a community of colleagues, mentors, and lifelong friends.

The technology and the industry have changed enormously, but SMPTE has been a constant, giving me the knowledge and perspective to keep up with those changes, while staying connected to the people who make this industry so special.

Looking back, so many of the people, experiences, and opportunities that have meant the most to me have come through SMPTE. I'm incredibly grateful to be a part of this community for more than 40 years.



Dagmar Driesnack, Senior Solutions Engineer at Appear
Germany
When I joined the media and broadcast industry many years ago as a student/young engineer, SMPTE was always something out of reach for me—geographically (as I'm based in Germany) and also knowledge-wise. I never imagined that, with time and the help of my colleagues, I would become involved in SMPTE's work, let alone serve as a Member of the Board! It's still unbelievable. SMPTE is not unreachable—ask questions, join discussions, and be brave. SMPTE members are willing to help, united by a shared goal of moving the media industry forward for 110 years. Every (even silent) voice is needed. Share ideas, share your knowledge, and be part of SMPTE.



Eric Gsell, Staff Engineer at Dolby Laboratories
United States (Hollywood, CA)
SMPTE is so many things to me: a platform to learn about new and complex technologies; a place to understand the intricacies of a Standard; a network of colleagues and friends all working towards the same goals and dreams. I wouldn't be where I am today without the support and help of SMPTE.



William T. Hayes, Media Consultant & Mentor
United States (Iowa)
In the early 1980s, I was a radio engineer when a crazy station owner with a construction permit to build a UHF television station in Hawaii hired me to do the job. I was just as crazy and said yes. I knew I had to get up to speed on television technology in a hurry and joined SMPTE to learn about television, initially from the *Journal* but also by attending SMPTE conferences, where I met industry giants who willingly shared their knowledge, guidance, and friendship as I grew. Now, some 45 years later, I cannot imagine my career or the industry without SMPTE providing education, standards, and stability amid the chaos of media businesses.

1545-0279/26©2026SMPTE



Tony Ngai, Founder and Director at Society of Motion Imaging, Ltd.
Hong Kong
I joined SMPTE in 1974 as a student member in Hollywood. The first SMPTE meeting was in a Hollywood studio, and I immediately felt that the SMPTE community was full of passion, professionalism, kindness, and willingness to share, even with me, who was only a very junior student member at age 24. Now I am one of the founding members of the Hong Kong Section and have been with SMPTE for 52 years already and going strong; the same feeling is there and best described by SMPTE's Slogan, "We ARE ONE"!



Allan Schollnick, Business Development Executive at Hollywood Film Company
United States (Hollywood, CA)
SMPTE Standards have been part of my career for as long as I can remember. In 1982, a computer with 8-inch floppy disks, made by RA Gray Labs—the same company that made timecode generators and readers—provided automation for the variety of projection devices at the Strassenburgh Planetarium in Rochester, NY.

At Lorimar Technical Service near LAX, the QC, audio, and telecine suites followed SMPTE specifications for maintaining 18% gray ambient light around the display monitor to balance the light and reduce eye fatigue.

Later, as Quality Control Manager at The Walt Disney Company, I was tasked with writing video mastering specs. Industry engineers recommended I find the SMPTE RS-170A, which became the basic standard for horizontal and vertical blanking burst, front and back porch, where vertical line captions fall, VITS and VIRS fundamental elements that helped drive the formats and standards we have today.

The DCP is a great example: without it, every projector manufacturer could have created its own format. SMPTE standards helped make DCP a common theater standard, much as film had been.



Sara Seidel, Marketing Manager, the Americas at Riedel Communications
United States (New Jersey)
SMPTE has played a defining role in my professional journey. It took an introverted software engineer with a passion for technology and a few ideas and gave her a platform to share them—through technical papers, presentations, standards work, and conversations with others who shared that same passion.

More importantly, SMPTE gave me a community. It connected me with colleagues, mentors, and friends who challenged my thinking, encouraged me to contribute, and opened doors to opportunities I never could have imagined. Through SMPTE, I found opportunities not only to participate in our industry, but to lead, contribute to its technical direction, and help shape its future.

Over the years, SMPTE became much more than a professional organization to me. It became a community, a source of inspiration, and ultimately, a family. It helped give me the confidence to find my voice and the opportunity to use it in service of our industry.

I would not be the engineer, technologist, leader, or advocate I am today without SMPTE. Much of what I have accomplished in this industry began with the opportunities SMPTE gave me—and I am deeply grateful for that.



John Shike, Key Account Manager at Advanced Systems Group, LLC (ASG)
United States (San Francisco)
My introduction to the depth and breadth of video technology came at the SMPTE Annual Conference at the Hilton in New York. In the years that followed, as Head of Marketing for the Americas at Snell & Wilcox, I made it a priority for our company to exhibit at the SMPTE Fall Conference each year.

Today, as Director of Membership at SMPTE, I am delighted to support SMPTE Student Chapters and help students explore the important role standards and technology play in the creative world of media.



Rachel Romero, Technical Account Manager at TAG Video Systems
United States (Washington, DC)

Growing up as the firstborn in a Spanish household, my life had a clear path: study hard, go to college, make money, start a family. That's it. I never, ever expected to become an engineer. Maybe a doctor, scientist, or even the next Indiana Jones. Never an engineer, though. Why? Because the "engineers" in school were the boys in the AV club. The girls were the "talent" in front of the camera. It was never presented as an option to me.

So, I did what I had always done so well: be a good student. That somehow landed me in LA for the American Society of Cinematographers Master Class and, as they say, Bob's your uncle. I was completely hooked. The camaraderie, the science, the engineering! I was beyond curious about how it all worked to create the magic I saw onscreen. But how could I keep pursuing this passion when I couldn't afford to move to LA? Then, I learned about SMPTE.

Grade school teaches you the basics, college prepares you for a paycheck, but organizations like SMPTE? That is where I actually got to pursue my real interests. The SMPTE ST 2110 Boot Camp gave me the foundational knowledge I needed to jump right into the latest broadcast tech and gain access to amazing jobs. But the real magic of SMPTE? Learning side-by-side with industry leaders who took me under their wings and championed me. I remember thinking, Wait, is this normal? As a woman navigating a male-dominated field, that mentorship was everything I'd never had before.

Now, it's my turn to give back to this community. I want to keep these doors open and build that same bridge for the next generation of engineers, dreamers, and avid learners coming up behind us.



Paul Whybrow, SMPTE Regional Governor – Asia-Pacific
Australia

I may have only been involved with SMPTE in the last eight years of my career, but I know many for whom SMPTE has been in their blood for 40 years or more, guiding their technological, operational, and innovational journey from junior engineer through to CTO. The passion and collaboration around the globe to bring audiences the best viewing experience are at the heart of the organization. Being 110 years old is a mere steppingstone for SMPTE's next 50 years!

Conclusion

In this new era, I'm channeling my inner Mr. Rogers—or, more accurately, his mother—who famously reminded him in times of trouble to "look for the helpers."

SMPTE is full of helpers. People who give their time, talent, and treasure to this Society. People who serve as ambassadors for SMPTE and strive to pay forward the support, mentorship, and opportunities they received along their own journeys.

Now, I'm calling on all of our helpers.

There are so many ways to help strengthen SMPTE and ensure its future, and no act is too small. Make a personal contribution to our fundraising campaign. Explore whether your company has corporate giving opportunities. Talk to your organization about the value of Corporate Membership. Join the Standards Community. Bring a colleague or student to a Section meeting. Share information about the SMPTE Summit, our courses and training, or a local event with your network. Write a blog post. Get involved with your local Section or Student Chapter. Volunteer at an event. Share your story. Offer your skills, your ideas, your connections, or simply your voice.

It all matters.

For more than a century, people have stepped up to build this Society, support one another, and help move our industry forward. Now it's our turn to make sure that continues for generations to come.

Calling all helpers! Help us strengthen SMPTE and build what comes next.

Sign up to Help: https://www.surveymonkey.com/r/HELPSMPTE



SMPTE 110 YEARS OF INNOVATION 1916 – 2026

110 Campaign: https://www.smpte.org/110-anniversary-campaign

DOI: 10.5594/JMI.2026/VBUH3640
Date of publication: 1 October 2026

1545-0279/26©2026SMPTE

SMPTE
MEDIA
TECHNOLOGY
SUMMIT

Pasadena
2026

November 16-19
Pasadena Convention Center

Introducing the Keynote Speakers for the SMPTE Media Technology Business Forum

Monday, November 16, 2026

Executives from leading companies will participate in keynote conversations, panel discussions, and interactive roundtables focused on the business forces transforming the industry!

Monday Keynote Speakers



David Ross

David Ross, founder of Ross Video, shares leadership lessons and insights on navigating change in the media and technology industries.



David Bermbach

David Bermbach is an electronics engineer who studied at TU Munich and has spent the past 20 years shaping ARRI's technology and product landscape.



Adde Grandberg

Adde Granberg is a senior media and technology executive with more than 15 years of leadership experience at SVT, Sweden's national public service broadcaster.

Technical Conference Keynote Session

Tuesday, November 17, 2026



Don Pettit

Join NASA astronaut Don Pettit as he shares breathtaking images from the International Space Station, followed by a panel discussion on the Artemis II mission and imagery beyond low Earth orbit.

summit.smpte.org/2026

2026 SMPTE FELLOWS

"A Fellow of the Society is one who has by proficiency and contributions attained an outstanding rank among engineers or executives in the motion picture, television, or related industries."

Nine members will be inducted as SMPTE Fellows at the Fellows Luncheon on Tuesday, 17 November 2026.



Zandra Clarke
Broadcast Content Transmission Specialist,
Warner Bros. Discovery



Jimmy Fournier
Director General, Technologies,
National Film Board of Canada



Thomas Heritage
Senior R&D Engineer
British Broadcasting Corporation (BBC)



Harvey Landy
Director, Advanced Media Technology, *Office of the CTO, NBCUniversal*



Steve LLamb
VP, Technology Standards & Solutions,
Cinema Worldwide, Deluxe



Yukihiro Nishida
Chair, ARIB Standard Assembly
Unaffiliated



C.V. Rao
Chief Technology Officer,
Annapurna Studios Private Ltd.



Allan Schollnick
Business Development Consultant, *AJS Media, Inc.*



Marc Zorn
Cybersecurity and Media Production Technologist, Unaffiliated

2026 AWARDS AND HONOREES

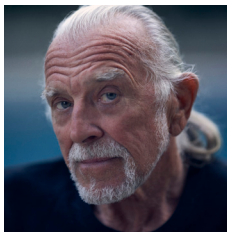
The MTS26 Gala will be held on Thursday, 19 November, 2026, featuring a reception and dinner at the Pasadena Convention Center, Pasadena, CA.

HONOR ROLL



Henry (Hank) Mahler
Associate Director, Advanced Technology *CBS Laboratory*
(posthumous)

HONORARY MEMBERSHIP



John Dykstra
ASC Visual Effects Designer,
John Dykstra Film Services

PROGRESS MEDAL AWARD



Robert P. Seidel
Operating Partner,
Lakewood Advisors

DAVID SARNOFF MEDAL



Jerry C. Whitaker
Consultant, Retired,
Independent/NA

LOUIS F. WOLF JR. MEMORIAL SCHOLARSHIP



Polly Chesnokova
Cinematography Fellow,
AFI Conservatory, Los Angeles

EXCELLENCE IN EDUCATION MEDAL



Anil Kokaram
Department Head and Chair, Electronic & Electrical Engineering, *Trinity College Dublin*

1545-0279/26©2026SMPTE

WORKFLOW SYSTEMS MEDAL



Bruce Devlin
CEO, Mr. MXF

JAMES A. LINDNER ARCHIVAL TECHNOLOGY MEDAL



Nick Bergh
Founder,
Endpoint Audio Labs, Inc.

EXCELLENCE IN STANDARDS AWARD



Wade L. Hanniball
Retired



Gerald (Jerry) Pierce
Former Chairman/Executive Director, *Inter-Society/ISDCF*
(posthumous)

JOURNAL AWARD



Charles Q. Robinson
Retired

JOURNAL CERTIFICATE OF MERIT

For the paper "*Evaluating the Energy Impact of Different Dimensions of UHD*"



Tania Pouli
Founder/ Board Member, *Vivid Manta, Greening of Streaming*



Khaoula Ounajim
R&D Engineer,
New Imaging Technologies



Daniel Menard
Professor, *INSA-Rennes and IETR Lab*

STUDENT PAPER AWARD



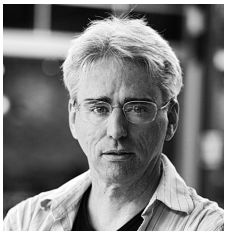
Miles Katz
Student, *Emerson College*

JOURNAL CERTIFICATE OF MERIT

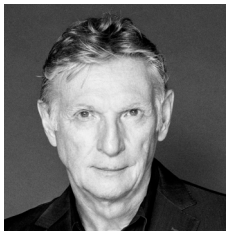
For the paper "*Color Management Principles for LED Panels in On-Set Virtual Production*"



David Long
Director, MAGIC Spell Studios,
Rochester Institute of Technology



Rod Bogart
Principal Color Scientist,
Epic Games



Spencer Stephens
SVP of Production Technology and Security, *MovieLabs*



Catherine Meininger
VP of Engineering,
Portrait Displays, Inc.



Joseph Goldstone
Image Science Engineer,
Independent Consultant



Tim S. Kang
Principal Engineer, Imaging Applications, *Aputure*



Kathryn Brillhart
Cinematographer and Virtual Production Supervisor



Jim Geduldick
Visual Effects/VP Supervisor & Cinematographer, *Spaceboy Labs*

DOI: 10.5594/JMI.2026/FCZG8370
Date of publication: 1 October 2026

SMPTE Spotlight:

Wim Van Roy

BY RUSSELL POOLE

CURRENT POSITION:
- SMPTE Education Director - 2025–2026
- Consultant - Educational Architect Lawo

PROFESSIONAL ORGANIZATIONS:
- SMPTE
- World Forum for Motor Museums

EDUCATION:
- Solvay Brussels School of Economics and Management
- LUCA School of Arts
- Thomas More University of Applied Sciences

When an industry transforms as quickly as media technology, continued education is imperative for maintaining an active workforce. For that education to be possible, instructors must be flexible, driven, and trusted in their field. Wim Van Roy has all of these qualities, which is why he's one of SMPTE's most valuable partners. An instructor for many of our SMPTE ST 2110 courses, Van Roy has 23 years of experience as a consultant running *Video Meets Light* in the media tech industry, providing services for companies like Lawo and EVS, and is also a Certified Professional in Training Management (CPTM™).

I WOULD LIKE MY LASTING IMPACT TO BE **A STRONGER CULTURE OF KNOWLEDGE SHARING—**ONE IN WHICH EXPERIENCED PROFESSIONALS ACTIVELY SUPPORT THE NEXT GENERATION AND LEARNERS CAN SEE CLEAR PATHWAYS INTO OUR INDUSTRY.

"I am particularly proud to have trained a global audience and to have helped people build confidence during major technological transitions," said Van Roy. "I hope to make high-quality technical education more accessible, practical and inclusive. I would like my lasting impact to be a stronger culture of knowledge sharing—one in which experienced professionals actively support the next generation and learners can see clear pathways into our industry."

While Van Roy's work at SMPTE has focused on education, he is an accomplished engineer with many impressive accolades. He has designed scalable education programs and supported organizations moving from traditional broadcast systems toward IP, cloud and software-defined workflows, working with organizations including WWE, Tomorrowland, Studio 100, NEP, EMG and Tennis Channel.

"My greatest influences have been the engineers, operators, educators and learners I have met throughout my career," said Van Roy. "I especially admire people who combine deep technical knowledge with humility and the ability to explain complex subjects clearly. Working alongside professionals in live production also taught me the importance of preparation, calm decision-making and genuine collaboration. My trainees continue to influence me as well: their questions constantly challenge me to make education clearer, more

1545-0279/26©2026SMPTE

practical, and more relevant."

Van Roy's SMPTE journey began and grew through professional development and training around IP networking and ST 2110, including the SMPTE Intensive Boot Camp community. It deepened when he became SMPTE Director of Education for 2025–2026, a role in which he continues to serve the organization through his commitment to continuing education and his outreach with fellow members.

"SMPTE has connected me with an international community committed to open standards, knowledge sharing and continuous learning," said Van Roy. "It has broadened my perspective and strengthened my belief that education is essential to helping our industry navigate change. I enjoy connecting the history of technology with the people and stories behind it."

Wim Van Roy is a valued SMPTE educator, engineer, and member of the industry. His interests are wide-ranging, and he serves as Chair of the World Forum for Motor Museums and as an ambassador for Autoworld and Mahymobiles in Belgium. His interest in technology serves him well as a learner and an educator, and is one of many reasons the future of the industry is secure.

DOI: 10.5594/JMI.2026/RFGZ9636
Date of publication: 1 October 2026



TECHNICAL PAPER

KEYWORDS 8K×8K CAMERA // IMMERSIVE VIDEO // MULTIPLATFORM CONTENT PRODUCTION // VR CONTENT EDITING SYSTEM

Compact 8K × 8K Camera; Imaging the Immersive Era

By Kodai Kikuchi, Akira Honji, Kohei Tomioka, Tetsuya Hayashida, Takenobu Usui, Toshie Hiroshima, Kazuya Kitamura, Takayuki Yamashita, and Hiroshi Shimamoto

Abstract

The demand for wide-field-of-view immersive video production is increasing. To address this, this work develops a compact 8K × 8K camera system. The camera incorporates a square-format image sensor with a compact 5/6-type footprint and ultra-high throughput (59-megapixel 8K × 8K resolution, 60 fps, 14-bit depth), achieving superior performance in high-mobility shooting scenarios and vertically extended image acquisition compared with recent large-format high-resolution sensors. The camera system employs a modular design in which the lightweight camera head is separated from the signal-processing unit by a fiber, enabling flexible deployment on gimbals, vehicles, and drones. Production trials were conducted using this system, including the parallel production of immersive VR content and 4K IMAX content derived from the same 180° fisheye 8K × 8K footage, and multi-region-of-interest (ROI) extraction for 2K broadcasting studio applications. These trials showed the system's effectiveness for multiplatform content production and helped simplify conventional broadcast workflows. The system not only broadens broadcast capabilities but also shows adaptability for the emerging immersive era.

Image acquisition technologies in broadcasting have traditionally evolved around horizontally oriented aspect ratios—most notably 16:9—supported by advances in frame rate and resolution, such as 2K, 4K, and 8K ultra-high definition television (UHDTV).[1] In recent years, however, the rapid diversification of contemporary media environments has increased the demand for wide-field-of-view (FoV) imaging, including 180° and 360° capture, for applications such as giant-screen venues (e.g., planetariums, flying theaters, and immersive museums), high-resolution head-mounted displays (HMDs), and emerging 3D reconstruction workflows such as Gaussian splatting (3DGS).[2-6] These use cases require imaging systems that can capture high-resolution, omnidirectional information—including from the vertical field—which conventional 16:9 or other rectangular-format sensors struggle to capture effectively.

Importantly, this technological shift is not peripheral to the broadcast industry. Wide-FoV acquisition offers several practical benefits for broadcasters:

- Multiplatform content deployment across broadcast, cinema, web (portrait orientation), VR, and large-screen exhibitions.
- production efficiency improvement through flexible view extraction from wide-FoV images.

Consequently, broadcast production environments require imaging systems that maintain traditional framing quality while supporting greater flexibility in viewpoint and spatial coverage. Accordingly, a novel camera system was designed with the following requirements:

- an image sensor capable of wide-FoV acquisition, including the vertical direction,
- a compact imaging platform suitable for agile deployment and field reporting, and
- a workflow compatible with existing 8K broadcast pipelines.

To achieve these goals, we designed a 7,680 × 7,680-pixel

1545-0279/26©2026SMPTE

ultra-high-resolution CMOS image sensor (referred to as an 8K × 8K image sensor) with a square (1:1) aspect ratio and incorporated it into a compact camera system. This paper presents an overview of the developed sensor and camera system followed by the results from multiple wide-FoV production trials conducted using the prototype.

8K × 8K Image Sensor

Figure 1 depicts the developed 8K × 8K image sensor, and **Table 1** summarizes its main specifications. Square aspect ratio imaging enables isotropic pixel sampling across both horizontal and vertical directions, allowing a single sensor to capture wide-FoV imagery in a single shot. This sensor employs the latest backside-illuminated stacked-pixel technology combined with a proprietary high-speed pixel readout circuit that incorporates folding column-parallel pixel wiring and a binary/gray code hybrid single-slope analog-to-digital converter (ADC).[7] This design provides an optimal balance between its approximately 50 Gbit/s ultra-high-throughput output (8K × 8K, 60 fps, 14-bit) and a compact physical footprint. The diagonal length of the sensor was 13.58 mm (5/6-type), which is significantly more compact than recent wide-aperture sensors. When the central 16:9 region (7,680 × 4,320 pixels) is cropped, the resulting optical size corresponds to a 2/3-type format, ensuring high compatibility with the B4-mount professional lenses commonly used in broadcasting production.

8K × 8K Camera System

The experimental camera system was developed by installing an 8K × 8 K image sensor, as shown in **Fig. 2.** The system adopts a 2-in-1 modular configuration in which the camera head is separated from the signal-processing unit. This architecture leverages the sensor's compactness, enabling an extremely small camera head while isolating the processing unit's power and weight to enhance operational flexibility. The two units are connected using a duplex single-mode optical fiber link capable of operating at distances exceeding 1 km. The entire system can be battery-powered, making it suitable for field production environments where AC power is difficult to secure, such as outdoor or remote locations.

The camera head, whose specifications are listed in **Table 2**, features a stick-type housing containing a densely integrated sensor board with an optical interface (I/F). This structure supports stable moving capture with common gimbals and shooting in confined spaces, which is infeasible using larger conventional UHD cameras. Additionally, by placing two heads with fisheye lenses in a cross-facing arrangement **(Fig. 3)**, 360° capture with minimized internodal distance can be achieved, which is advantageous for high-quality omnidirectional stitching. The default lens interface is a C-mount with support for B4, PL, and DSLR lenses via mount adapters.

As shown in **Table 3**, the signal processing unit performs real-time image adjustments equivalent to those of the broadcast camera control unit on the RAW output from the 8K × 8K sensor. The video is output via eight 12G-SDI channels in an 8K × 4K × 2 split configuration (upper and lower halves). In addition, the system enables 4K-resolution cropping and magnification from any region within the full 8K × 8K frame, aiding in precise focus adjustment, which is an essential requirement for ultra-high-resolution imaging.

TABLE 1. Specifications of the 8K × 8K Image Sensor.

Pixel number (active area)	7,680(H) × 7,680(V)
Pixel size	1.25 × 1.25 µm
Frame rate	60/1.001 fps
ADC resolution	14 bits
Shutter operation	Rolling
Power consumption	2.44 W

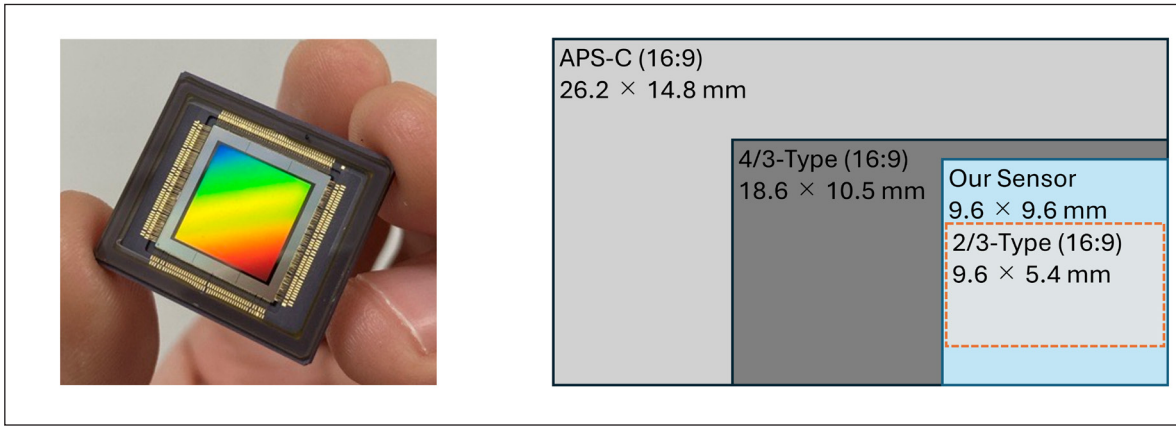


FIGURE 1. Appearance and optical size comparison of the 8K × 8K image sensor.

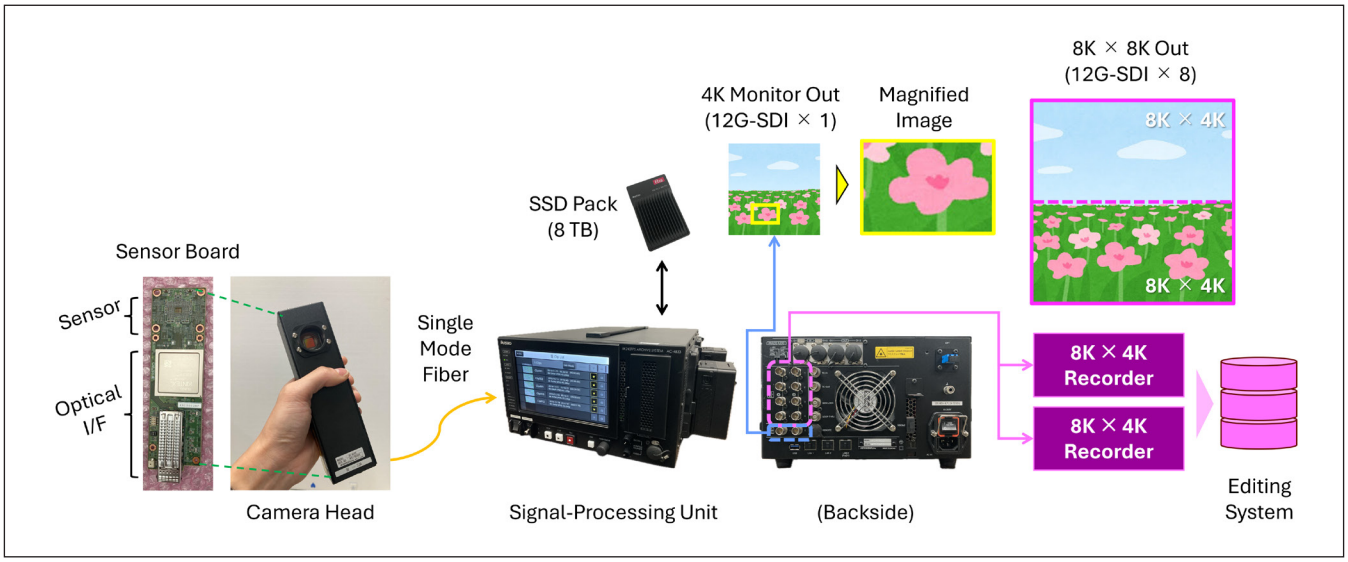


FIGURE 2. System configuration of the 8K × 8K camera.

TABLE 2. Specifications of the Camera Head.

Size/Weight	50 × 179 × 33mm/0.5kg
Lens Mount	C-Mount
Power Supply	USB-C 9V, 3A
Output I/F	QSFP28 100 Gbit/s

TABLE 3. Specifications of the Signal-Processing Unit.

Size/weight	317.5 × 160 × 390mm/10kg
Processing function	De-Bayer, Whit Balance, Gain, Gamma, Edge Enhancement
Recording Time	54 min (8-TB SSD Pack)
Power Supply	AC, V-Mount Battery
Output I/F	8K × 8K Line Output: 12G-SDI × 8 4K Monitor: 12G-SDI × 1



FIGURE 3. Example camera-head setups using a fisheye lens (Left: single-camera setup; right: dual-cross-facing setup for 360° capture).

For recording, the approximately 50 Gbit/s sensor output is encoded in real time into a proprietary JPEG XS-based compressed format (approximately 16 Gbit/s) and stored on a dedicated 8-TB SSD pack integrated into the signal-processing unit, providing up to 54 min of recording. The system allows real-time playback through the signal-processing unit for re-recording with two external 8K× 4K recorders, as shown on the right side of **Fig. 2**. This enables rapid transcoding into general editing codecs such as Apple ProRes (422 HQ), facilitating a practical workflow despite the large data volumes involved.

The camera performance was evaluated in a studio. According to the measurements conducted in compliance with ARIB TR-B45,[8] the camera achieved a sensitivity of F8 at 2000 lux, a signal-to-noise ratio (SNR) of 51 dB, and a dynamic range of 800%.

Figure 4 presents the resolution characteristics of the camera modulation transfer function (MTF) when the camera head was equipped with two different C-mount lenses (small lens: Edmund Optics HR 25 mm, fisheye lens: Entaniya HAL250 3.0 mm) with an F-number of 2.8. In both cases, a modulation of 20% or higher at 3200 television lines (TVL) was confirmed; for the fisheye lens, the value corresponds to the central region of the image.

Figure 5 shows an image captured using a fisheye lens (Entaniya HAL250 3.0 mm). The square-aspect-ratio sensor efficiently captured the full 180° circular fisheye projection at an 8K resolution. A comparison between a central cropped region and that from a commercial 8K 360° camera (dual-sensor configuration, 5.6K resolution per sensor, output available only in H.265 compression) reveals the superiority of the developed system in terms of higher spatial resolution and reduced compression distortion.

Content Production Experiments I: VR/Flat-Screen Parallel-Production

We conducted multiple production trials using the developed camera system. The following section presents several representative examples.

1545-0279/26©2026SMPTE

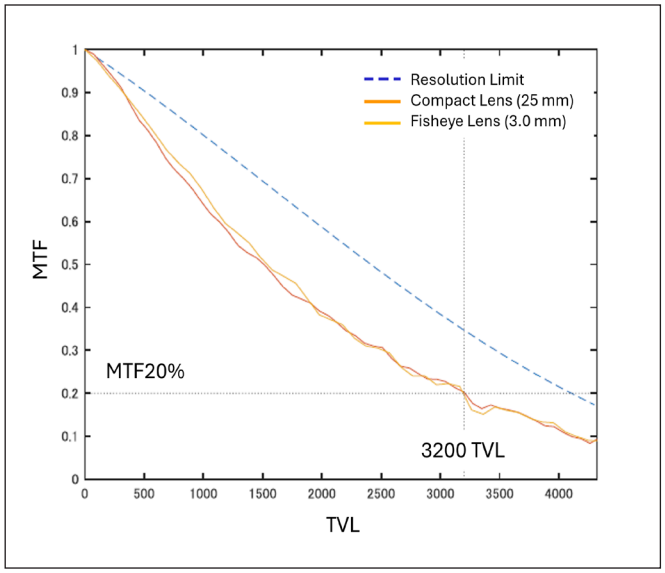


FIGURE 4. MTF measurement results of the camera using C-mount lenses (F2.8).

Figure 6 shows photographs of 8K × 8K shooting examples, in which a fisheye lens was used to obtain 180° images. To acquire smooth, high-quality moving shots with effective stabilization and horizon correction, various cinematography equipment was employed to leverage the compact and lightweight camera head:

- **Dolly shooting:** Simple dolly-style movements were executed using gimbals or stabilizers. Hand-pulled rickshaws were used on paved roads.
- **Vehicle shooting:** For more dynamic moving shots, the camera head was mounted at the tip of a filming arm attached to a vehicle, such as an electric bicycle or automobile, while the signal-processing unit was placed on a cargo rack or inside the vehicle.
- **Wired-drone shooting:** A drone equipped with a gimbal carried the camera head and was connected to a ground-based signal-processing unit via an optical cable. Mounting the camera in front of the drone enabled obstruction-free 180° FoV capture without propeller occlusion (shown in the lower-right image in **Fig. 6**).

A verification test evaluated a workflow in which a single 180° source was edited in parallel into two output formats: immersive VR and flat-screen content. In this test, dolly-shot footage captured around the Senso-ji Temple, an iconic landmark in Tokyo, Japan, served as the common source. The VR version targeted a monoscopic equirectangular projection (ERP, aspect ratio 2:1) format for HMDs (specifically Meta Quest 3), whereas the flat-screen version was edited in a 4K IMAX (aspect ratio 1.43:1) format intended for large-format cinema exhibitions.

For the VR content editing, the 8K × 8K fisheye footage was integrated and converted into a 15K × 8K ERP format. Editing was performed on the valid 180° ERP region using a nonlinear editor (DaVinci Resolve), including timeline assembly, color grading, and viewpoint adjustments (central position and horizontal alignment).

To streamline the HMD-based preview during editing, we developed a preview-assist system, as shown in **Fig. 7**, in which the preview PC receives the ERP output from the editing PC, maps it onto a virtual sphere, and displays arbitrary viewpoints on an HMD in real time. This eliminates the repetitive cycle of editing, exporting, and previewing, which has been a major burden in conventional VR editing workflows, allowing editors to directly check the appearance of HMDs during editing. Furthermore, separating the editing and preview PCs helps avoid performance degradation during viewpoint extraction and supports production workflows in which different operators handle editing and preview tasks.

Using this workflow, we generated 8K × 4K VR content (res-

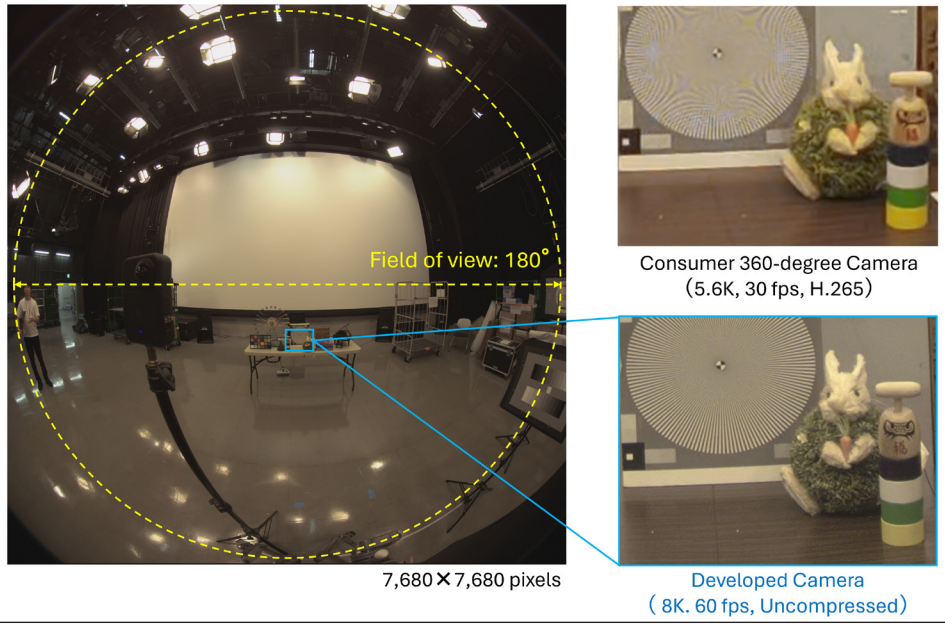


FIGURE 5. Example captured using a fisheye lens and comparison of cropped images.

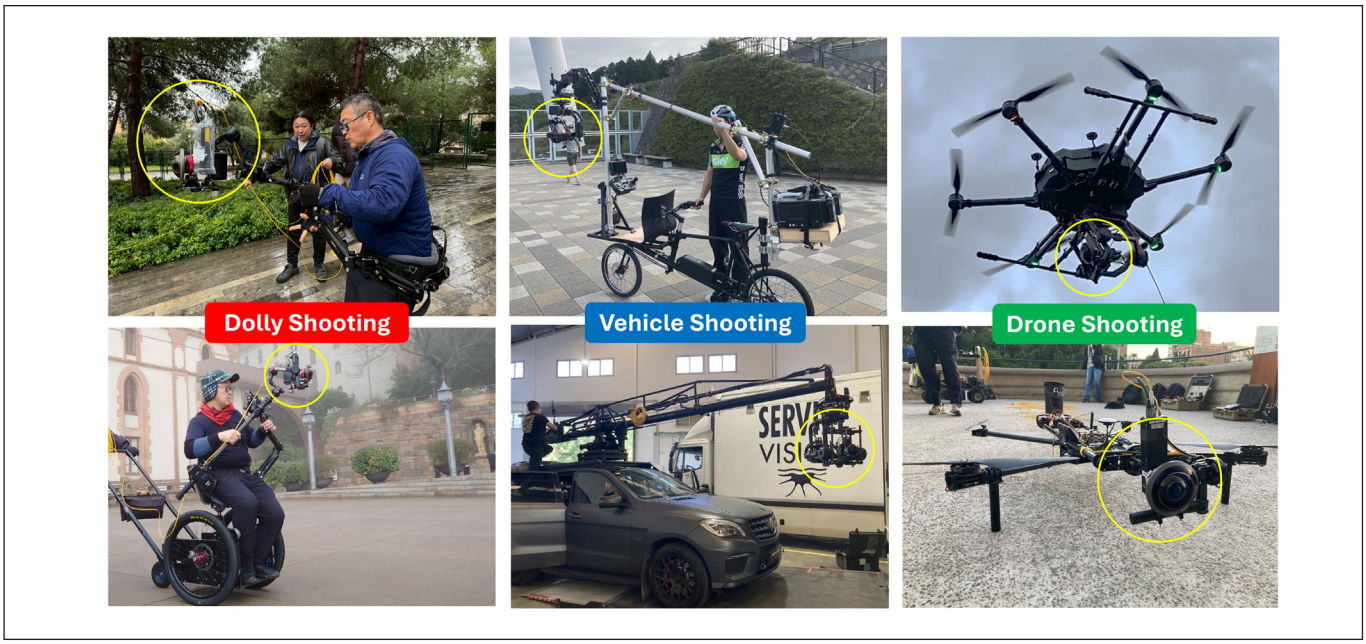


FIGURE 6. Examples of highly mobile wide-FoV shooting using cinematography equipment (circles indicate the 8K × 8K camera heads).



FIGURE 7. Photograph of the VR content editing system.

olution ceiling imposed by the HMD) in the style of a virtual tour.

For the 4K program production using the same source footage, we used reframing tools in DaVinci Resolve to extract arbitrary fields of view from the 180° footage, with keyframing to introduce pan, tilt, and zoom movements **(Fig. 8)**. The ultra-high-resolution acquisition provided by the 8K × 8K image sensor enabled the 4K-extracted footage to retain its visual detail for high-quality flat-screen presentation. This approach enables a wide variety of shot compositions from a single capture, a significant advantage when shooting opportunities are limited.

Through this workflow, we successfully produced VR and 4K IMAX content in parallel using identical 180° footage. The completed 4K IMAX piece was screened at the International Film Expo (GSCA Film Expo 2025, held in Los Angeles in March 2025), demonstrating the practical feasibility of multi-platform content generation within a unified broadcast production pipeline.

Content Production Experiments II: Multi-ROI Extraction

The developed camera system also shows strong potential for improving the efficiency of conventional broadcast production. **Figure 9** depicts a studio talk-show (for HDTV) production experiment in which the traditional workflow required multiple cameras to capture and switch close-up shots of individual cast members (in this case, five PTZ cameras were deployed in the studio set). In our trial, we installed an 8K × 8K camera head in the same set to capture a single, fixed wide-angle 8K studio shot. Multiple regions of interest (ROIs)—faces, group shots, etc.—were then extracted from this image, allowing flexible real-time switching using only one camera.

1545-0279/26©2026SMPTE

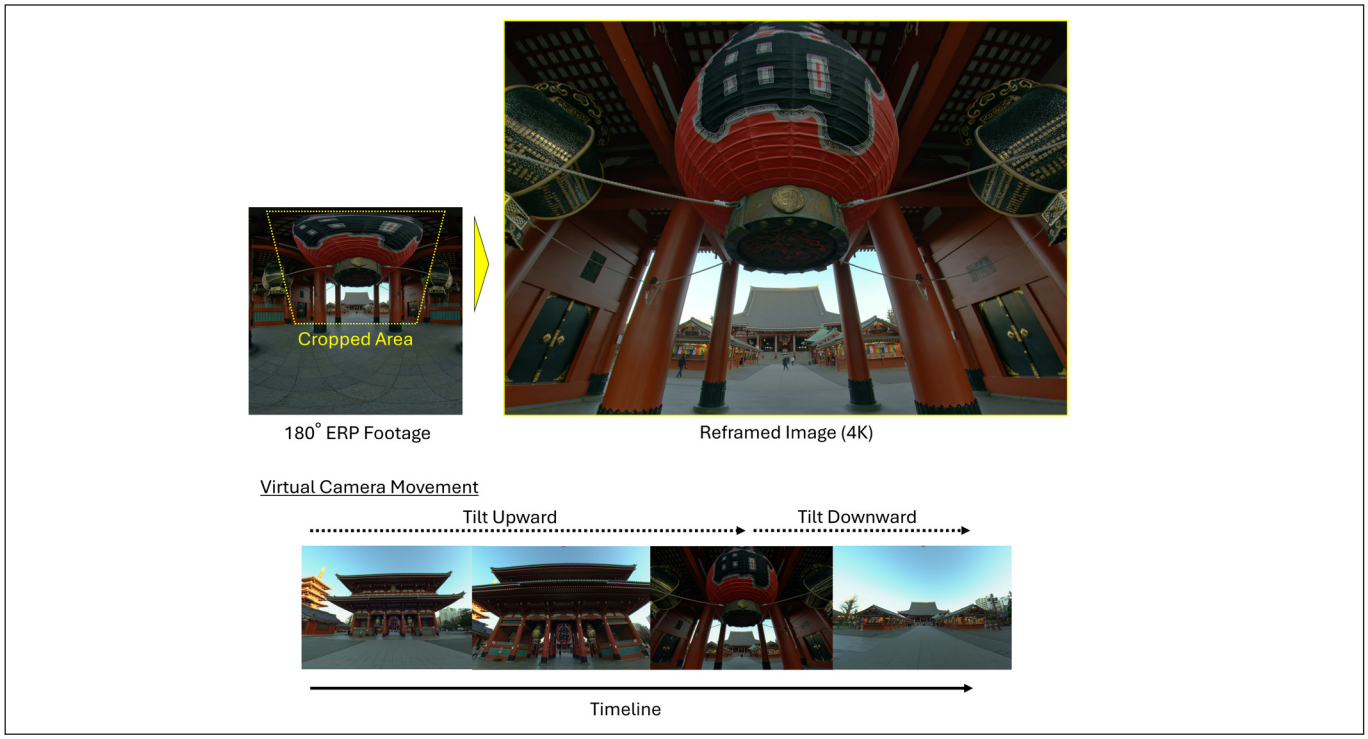


FIGURE 8. Example of 4K content production with reframing technique (Lower row: Example showing the added tilt motion when passing through the gate).

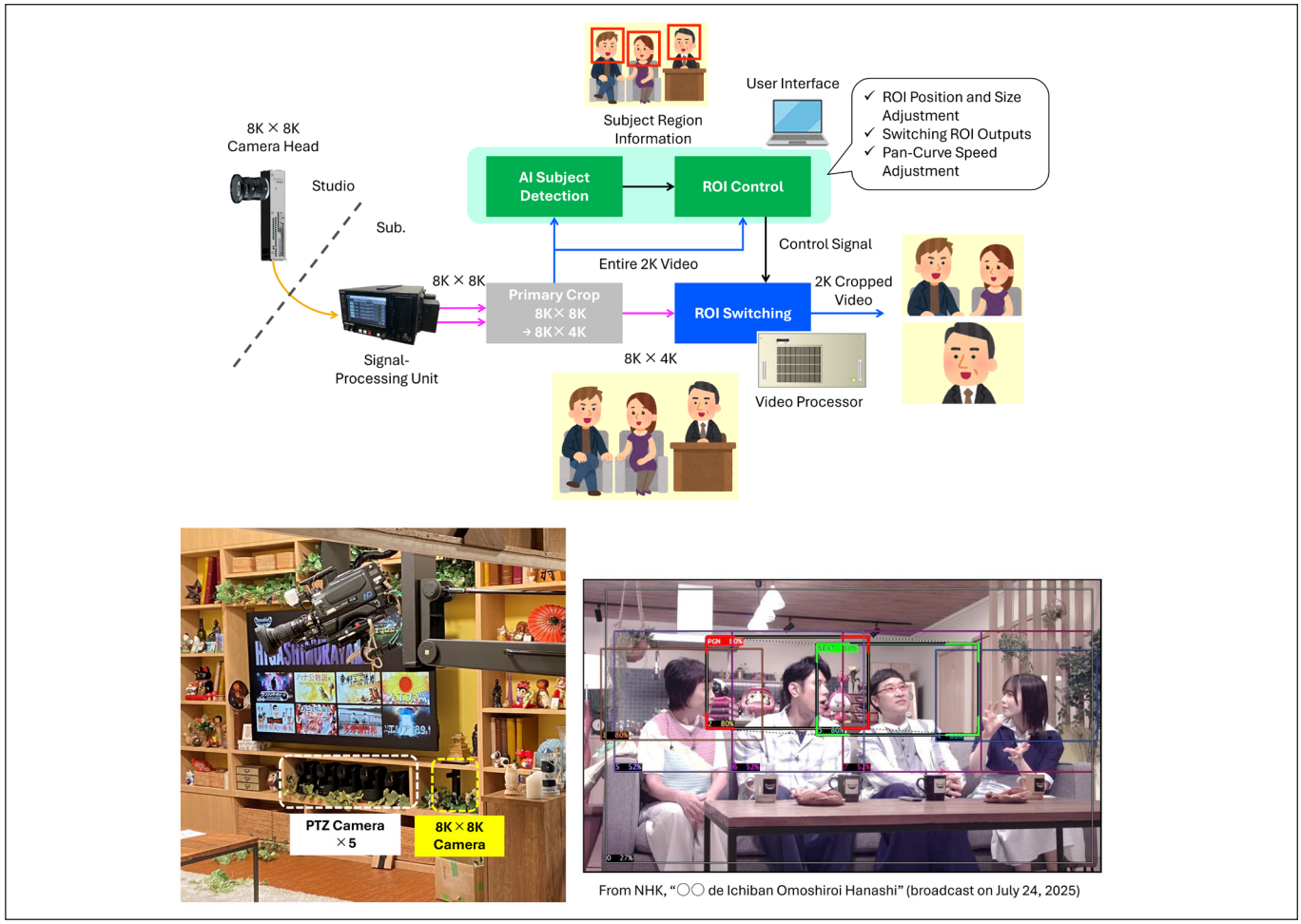


FIGURE 9. Multi-ROI production experiment for a studio talk show (Upper: system configuration; Lower left: camera system embedded in the studio set; Lower-right: ROI control interface showing the configured ROIs).

ROI extraction was performed using the developed HDTV multishot production system,[9] providing the following:

- simultaneous output of multiple ROIs from 8K video as virtual cameras,
- AI-driven person detection and ROI suggestion, and
- smooth virtual camera work generation and multi-shot switching.

The lower-right part of **Fig. 9** presents the captured image from the camera along with the defined ROIs, with rectangular frames indicating preset regions. In this experiment, five 2K virtual cameras were generated simultaneously, and inter-camera switching was performed using only a single 8K × 8K camera and a single operator, demonstrating the feasibility of reducing the number of cameras, simplifying the system configuration, and improving the overall production efficiency in studio-based shooting environments.

Conclusion

We introduced a compact 8K × 8K camera system optimized for wide-FoV, ultra-high-resolution immersive imaging and verified its capabilities through a series of production trials, including parallel editing of 4K IMAX and 8K VR content as well as multi-ROI switching operations. The results indicate that high-pixel-density, square-format sensors open up new possibilities for broadcast content production by enabling greater flexibility in multiplatform deployment and improved workflow efficiency.

To broaden the application range of the system, we are exploring highly mobile wide-FoV acquisition for immersive theaters, particularly flying-theater experiences, along with 360° capture using a dual-camera configuration to achieve greater freedom in omnidirectional coverage and the exploratory generation of high-quality 3DGS datasets from fisheye images to support more diverse reporting. Beyond broadcasting and entertainment, sensor architectures also show promise for adoption in fields such as medical imaging, security monitoring, and in-vehicle camera systems. Future studies should further explore these expanded application domains.

Acknowledgment

The 8K × 8K image sensor was developed in collaboration with the Sharp Semiconductor Innovation Corporation.

References

1. International Telecommunication Union-Radiocommunication (ITU-R), "Parameter values for ultra-high definition television systems for production and international programme exchange," ITU Recommendation BT.2020-2, Oct. 2015.
2. International Telecommunication Union-Radiocommunication (ITU-R), "Ultra-high definition 360-degree immersive video systems," ITU Recommendation BT.2123-0, Jul. 2019.
3. Peek Pro, "Immersive Experience Trends and Key Statistics in 2025," *Peek Pro Industry Blog*, 2025.
4. M. Wang et al., "VR content creation and exploration with deep learning: A survey," *Comput. Vis. Media*, 6(1):3–28, Mar. 2020.
5. B. Kerbl et al., "3D Gaussian Splatting for Real-Time Radiance Field Rendering," *ACM Trans. Graph.*, 42(4) 1, Jul. 2023.
6. C. Shin et al., "Seam360GS: Seamless 360° Gaussian Splatting from Real-World Omnidirectional Images," *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 28970–28979, Oct. 2025.
7. K. Kikuchi et al., "A 1.25 µm 59.3 Mpixel 60 fps CIS with 2 × 1 Multi-Directional Phase Detection Pixels," *Proc. IEEE Sens.*, Paper 6348, Sep. 2024.
8. ARIB TR-B45, "Measurement Methods for Sensitivity and Signal-to-Noise Ratio of Television Cameras," Ver. 1.1-E1, Apr. 2019.
9. International Telecommunication Union-Radiocommunication (ITU-R), "Experiences of Ultra-High Definition Television and High Dynamic Range Television," Report ITU-R BT.2556-0, 58–60, Sep. 2025.

About the Authors



Kodai Kikuchi has researched and developed ultra-high-definition camera and recording systems at NHK Science & Technology Research Laboratories (STRL) since 2013.



Akira Honji has researched and developed image sensors and 360-degree camera systems at NHK STRL since 2023.



Kohei Tomioka has researched and developed ultra-high-definition cameras and image sensors at NHK STRL since 2016.



Tetsuya Hayashida has researched imaging, camera, SHV, and 360-degree video technologies at NHK STRL since 1996.



Takenobu Usui has researched and developed information displays, image processing and advanced TV production systems at NHK STRL since 2007.



Toshie Hiroshima has researched and developed volumetric capture, super-resolution, and video production technologies at NHK STRL since 1999.



Kazuya Kitamura joined NHK in 2000 and has been engaged in research and development of image sensors and camera systems.



Takayuki Yamashita is an executive research engineer at NHK, leading next-generation imaging systems R&D and pioneering UHD-2/HDR broadcasting deployment and standardization.



Hiroshi Shimamoto has researched UHD-2 ultra-high-definition cameras, image sensors, and medical applications at NHK and NHK Foundation since 1993.

This paper first appeared in the Proceedings of the NAB 2026 Broadcast Engineering and Information Technology (BEIT) Conference. Reprinted here with permission of the author(s) and the National Association of Broadcasters, Washington, DC.

DOI: 10.5594/JMI.2026/ KAQU1510
Date of publication: 1 October 2026

1545-0279/26©2026SMPTE

KEYWORDS 5G // NPN // TV PRODUCTION // TV CONTRIBUTION // REMOTE PRODUCTION

Leveraging Private 5G for Live TV Production: A Case Study from the 2025 World University Winter Games

By G. Stante, A. Lucco Castello, F. Pandolfi, and L. Vignaroli

Abstract

The integration of 5G Non-Public Network (NPN) into event TV production can potentially offer significant benefits, particularly in complex or impractical environments for traditional cabled infrastructure. A live TV production during a global winter sports event presents these challenges. It therefore serves as a testbed to demonstrate how a private 5G network can simplify the deployment of broadcast production infrastructure. This paper presents the outcomes of the two technical trials integrating a private 5G network and a LEO satellite system (OneWeb) during the FISU World University Winter Games (WUG) 2025 held in the Piemonte region. The two trials focused on the following two events: the Opening Ceremony at Inalpi Arena in Turin and the Alpine Skiing competitions in Bardonecchia. Key performance metrics assessed include uplink throughput, uplink latency, video signal reliability, network coverage, spectrum interference, and terminal reconnection time when out of range. Additionally, the ease of installation, configuration, and usability of the private 5G system were also analyzed. This paper provides a comprehensive review of the entire experience. It highlights the advantages, limitations, and lessons learned, offering valuable insights into the future of TV production in challenging environments.

The use of private mobile networks in TV production represents a valuable opportunity to enhance the flexibility and efficiency of production infrastructure for live event coverage. Private 5G technology can offer significant advantages in situations where public networks are heavily congested due to large crowds, or in areas with limited or absent network coverage, such as certain urban or rural locations. As demonstrated during previous trials conducted around the globe,[1,2] private 5G networks can provide guaranteed transmission capacity, significantly reducing latency and jitter. In addition, strategically positioning antennas where needed can considerably improve network coverage, resulting in clear benefits to the quality of the transmitted video signal.

Based on previous activities,[3] which yielded promising results, RAI decided to leverage the Torino 2025—World University Winter Games[4] to conduct further field tests of the private 5G technology in a real-world sports broadcasting environment. This paper reports our first on-site experience using a standalone, non-public 5G network for live TV production, focusing on system deployment speed, ease of use, and network performance in terms of RF coverage and latency.

The Winter Universiade is one of the most important competitions for young athletes. It brings together student-athletes from around the world to compete across a wide range of winter sports. Held from January 13 – 23, the event showcased emerging talent in disciplines such as alpine skiing, ice hockey, figure skating, and more. Beyond the competition, the Universiade celebrated global university sports culture and international collaboration.

In this context, we deployed a compact private 5G system—commonly referred to as a "5G bubble"—integrating all the main functionalities: a 5G core network, a gNodeB (gNB) with two remote radio heads (RRHs) equipped with both di-

rectional and omnidirectional antennas, and a Low Earth Orbit (LEO) satellite system providing backhaul connectivity. This Proof of Concept (PoC) employed 100 MHz in the upper n77 band, allocated to RAI for experimental use by the Italian regulatory authority, the Ministry of Enterprises and Made in Italy.

The field trials took place in two challenging environments, specifically during the opening ceremony at Turin's Inalpi Arena and the skiing competitions at Bardonecchia ski resort, allowing us to assess 5G network performance in both indoor and outdoor scenarios. During the experimentation phase, we paid particular attention to evaluating the system's transportability and ease of installation.

In the following, this paper shares a detailed overview of the experience. It highlights the main benefits and drawbacks, challenges, and lessons learned, offering useful insights into how TV production could evolve through the use of a private 5G network.

5G Network Description

Among the various solutions available on the market, the private 5G network employed during these trials was provided by Neutral Wireless. Our 5G NPN setup consisted of a 1 rack-unit (RU) server hosting both the Core Network (CN) and the gNodeB (gNB). Each cell was equipped with a Remote Radio Head (RRH) for the RF signal processing, along with a dedicated antenna.

TABLE 1. 5G Network Settings.

5G Network Parameter	Setting Configured
Frequency Band	upper n77
Bandwidth	100 MHz
Frequency Range	3910 – 4010 MHz
TDD frame structure	DDSUUUUUUU (2:7 TDD ratio) DSUU (1:2 TDD ratio)
MIMO	DL: 2 x 2 MIMO UL: 1 x 2 RX Diversity

Table 1 reports the main 5G network parameters configured for the trial. RAI requested and obtained from the regulator (i.e., the Ministry of Enterprises and Made in Italy) a temporary license to use 100 MHz in the upper n77 band (i.e., 3.8 - 4.2 GHz) for experimental purposes. The frequency application procedure took nearly three months, since there is currently no standard procedure for this type of application. The requested frequency range (3910 – 4010 MHz) falls within a portion of the spectrum already used by RAI for the geographical distribution of television signals via point-to-point radio links to Digital Terrestrial Television (DTT). For this reason, particular care was taken to verify the absence of potential interference to and from the radio links at the event locations. Moreover, the regulator carried out additional interference simulations before issuing the authorization, with a positive outcome.

Compared with 4G, the new 5GNR (5G New Radio) defines many more frame structures that are allowed in the n77 band.[5] Since network traffic in TV production use cases is predominantly uplink, uplink-biased Time Division Duplexing (TDD) frame structures offer a significant advantage by enabling greater transmission capacity (i.e., uplink throughput). A possible TDD frame structure having these characteristics is the DDSUUUUUUU. This pattern uses two slots for downlink and 7 slots for uplink, providing a 2:7 (downlink: uplink) TDD ratio. As reported in Ref. 6, other frame structures also exist that are more suitable for video signal transmission, i.e., the DSUU. Its shorter structure helps to reduce latency and jitter, with only a slight trade-off in transmission capacity (due to a 1:2 TDD ratio). During setup, both configurations were tested, and because the available capacity was more than sufficient, the shorter structure was chosen for low-latency transmission. Additionally, the network was configured to support 2 × 2 MIMO (Multiple-Input Multiple-Output) in the downlink. In the uplink, since User Equipment (UE) capable of 2 × 2 MIMO is not yet available, receiver diversity (1 × 2) was configured to optimize signal reception.[7]

Two cells were deployed at both the Inalpi Arena and in Bardonecchia: one equipped with a directional antenna and the other with an omnidirectional antenna. The choice between the two depended on the specific area to be covered. To compensate for the low transmission power of UEs, high-gain antenna models were used. **Table 2** shows the antenna characteristics adopted for the two cells.

TABLE 2. 5G Network Cell Parameters.

Cell Parameters	Cell 1	Cell 2
Power	1W (30 dBm)	1W (30 dBm)
Antenna Type	Directional	Omni-Directional
Antenna Gain	18 dBi	12,8 dBi
Azimuth Beamwidth	65°	360°

Finally, potential concerns included the system's transportability and the time required on site for deployment and setup. Unexpectedly, the entire system proved highly portable: a small van was sufficient to transport all the hardware to the site (including the 5G system). As for deployment time, while there was an initial learning curve in becoming familiar with the hardware, the process was quicker than expected. Indeed, the whole 5G system, including the antennas, could be assembled on field in just half a day by a team of two people, significantly reducing deployment time and resulting in tangible cost savings. Moreover, for O&M (Operation & Maintenance), once the desired configuration is set, no additional support is needed to maintain system operability.

TV Production Setup

TV Production Architecture

Figure 1 illustrates the deployed television production architecture leveraging a 5G mobile network. It integrates multiple video transmission units, network switches, and production workstations.

On the left, the two separated 5G cells are shown, each equipped with its own Remote Radio Head (RRH). The two 5G antennas collect audio/video traffic from professional mobile video encoders and commercial smartphones. Video feeds are sent (as I, Q samples) via the RRH to the core 5G network through single-mode optical fiber. The I, Q samples are

1545-0279/26©2026SMPTE

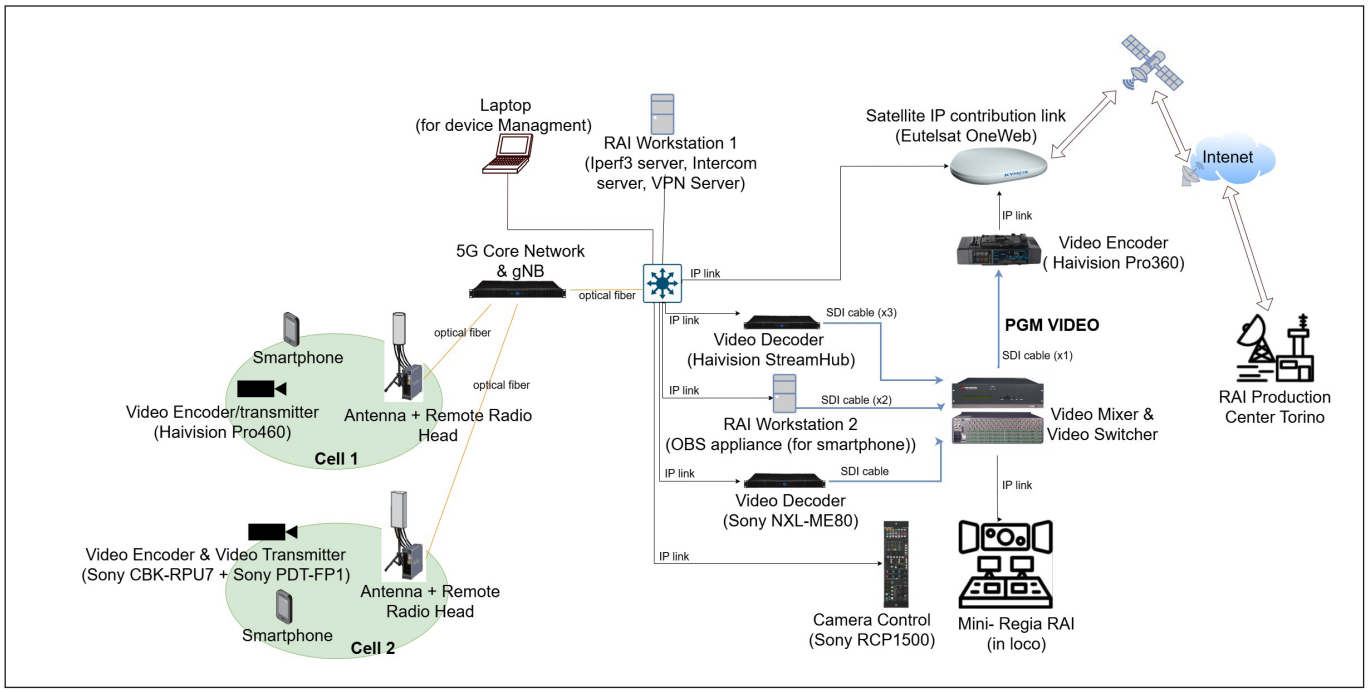


FIGURE 1. TV production setup.

then processed by the gNB, and the resulting IP packets are routed by the 5G core workstation. The IP video streams are then distributed within the TV compound via an Ethernet link through a network switch, configured with a star network topology.

To provide additional network services, two general-purpose workstations were integrated into the infrastructure. Workstation 1 was configured with an XCP-NG hypervisor to provide flexibility for installing application software, as needed. Several virtual machines were deployed on this hypervisor, each offering a different functionality: an Iperf3 server, an OpenVPN server, and an Intercom server. While the Iperf3 server was used to gather network statistics such as throughput, latency, and jitter, the OpenVPN server was set up to guarantee external network access (from the TV compound to UE connected to the 5G Radio Access Network) in order to enable TV operators to configure A/V equipment remotely. This feature was successfully tested in the lab but was not used during the trial, as the Haivision Pro460 encoder already supports it natively. As an intercom solution, we tested Mumble, which is a free VoIP application for low-latency voice communication over IP. It is an entirely software-based solution. The server component was hosted on a dedicated virtual machine within Workstation 1, while the client ran as an Android application installed on the 5G UEs. Although it technically works, the absence of a physical push-to-talk button is undoubtedly one disadvantage of this solution and the main reason why it was not adopted during the trial.

Workstation 2, instead, was equipped with a Blackmagic DeckLink 4K Extreme PCI Express card and was used for receiving video streams from 5G smartphones and transcoding them from the SRT (Secure Reliable Transport) protocol to SDI (Serial Digital Interface). All video signals, once decoded and converted to SDI, were mixed together through a video mixer available on site. The final output, i.e., the PGM video, was then encoded in H265 at 8 Mbit/s and sent via the Eutelsat OneWeb satellite system to the Rai Production Center in Turin.

TABLE 3. A/V Equipment Used.

Category	Device
Camera	Sony FX9 Sony 3500
Camera Control System	RCP-1500
Smartphones	Samsung S23 Apple Iphone 13 Apple Iphone 15
Mobile Video Encoder & 5G portable data transmitter	Haivision Pro460 (x3) Sony CBK-RPU7 Sony PDT-FP1
Video Decoder	Haivision Streamhub Sony NXL-ME80
LEO satellite contribution system	Eutelsat Oneweb
Video Mixer & Video Switcher	BlackMagic Atem 2 BlackMagic Smart Video Clean-Switch
Video Recoder & Video Player	BlackMagic HyperDeck Extreme 4K BlackMagic HyperDeck studio

Table 3 shows the hardware used during the trial. As a video encoder/decoder solution, we used both Haivision and Sony equipment. For both video chains, the perceived Quality of Experience was positive. The Haivision equipment consisted of three mobile Pro460 encoders and a Haivision Streamhub decoder. The units under test embedded 5G modems supporting natively 5G SA network and were capable of connecting successfully to the network. To minimize the time needed for the network search procedure, all unnecessary frequency bands except the one used (i.e., n77) were excluded.



FIGURE 2. Video feed sent via OneWeb.

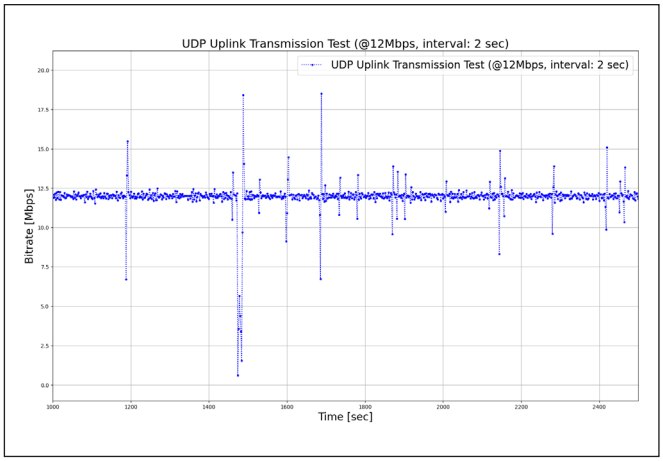


FIGURE 3. Sample of network behaviour during raw UDP transmission test (@12 Mbit/s) via LEO satellite systems.

The Sony video chain consisted of a portable encoder mounted directly on the back of the camera, a 5G portable data transmitter, a video decoder, and a camera control system. The specific unit used was the Sony PDT-FP1. It is a 5G portable data transmitter based on Android OS, shaped like a smartphone and developed for TV production needs. Both the camera control signal and video signals were received and transmitted through this 5G transmitter. Overall, the Sony system worked well, and the TV operators' feedback was positive. Despite adopting a 5G-based technical solution, TV operators perceived no difference compared to traditional SDI cameras. Video feed quality was good, as was the responsiveness of the remotely controlled cameras, thanks to the network's low delay.

TABLE 4. Encoder Settings.

	Haivision Encoder	Sony Encoder
Format	1080p50	1080p50
Bitrate	20 Mbit/s	20 Mbits/s
Transport Protocol	SST	Sony QoS
Latency Mode	ULL	ULL

As shown in **Table 4**, all video streams sent through the 5G network were in 1080p50 format, with a bit rate of 20 Mbit/s, in ultra-low-latency mode.

LEO Satellite Solution (OneWeb) for Uplink Video Contribution

After on-site video mixing operations, the finalized program feed was encoded in H.265 at 8 Mbit/s CBR (Constant Bit Rate) using a Haivision Pro360 unit and transmitted via Eutelsat OneWeb's satellite contribution link to the RAI Production Center in Torino. OneWeb's connectivity plan available to us provided uplink up to 20 Mbit/s in "best effort" mode and up to 4 Mbit/s of priority/guaranteed traffic. The transport protocol used was SST (Safe Streams Transport),[8] a proprietary protocol optimized for video transmission over mobile networks. **Figure 2** shows a screenshot of the video feed sent via the OneWeb solution.

The system's uplink throughput was sufficient for transmitting a single video feed, although sporadic visual artifacts were observed. Visual quality improved when the Pro360 end-to-end latency parameter was changed from 2 to 3 sec. Latency values higher than 3 sec were not tested in the field, as they are generally not accepted by the director's office during live TV production operations.

To better evaluate OneWeb's uplink performance, we conducted a dedicated test campaign. We collected the results and shared them with Eutelsat. Iperf3 software was used as a testbed tool to generate and collect IP traffic. The Iperf3 server was hosted on a virtual machine in an Azure data center located in Italy.

Figure 3 presents a sample from the test campaign, which is representative of the system's network behavior. This test involved a raw UDP/IP transmission at 12 Mbit/s CBR, with network statistics recorded every 2 sec.

As the graph indicates, the system can handle such uplink data rates. However, at random intervals, the bitrate drops drastically. Further analysis revealed that packets sent during these drops were not lost but buffered by the system and then delivered to the destination with high transmission delays, sometimes even several seconds, as shown by the presence of spikes occurring right after the drops. This intermittent behavior is typical of LEO satellite transmissions, which are significantly influenced by the relative movement between the satellite and the transmitting antenna. Due to this irregularity, for TV video contribution, it is highly recommended to bond the LEO satellite system with other backhaul solutions.

5G Network Deployment and Tests

Inalpi Arena

The first opportunity to test the private 5G network occurred during the opening ceremony of the Universiade, held on January 13, 2025, at the Inalpi Arena in Turin. This initial event proved especially valuable for assessing the solution's potential. The Inalpi building (also known as Pala Isozaki, named after the designer) is built largely of reinforced concrete and iron, both of which are challenging materials for radio propagation. Moreover, in addition to fixed structures, the arena contains many movable metallic structures, such as seat racks and lighting trusses.

It was decided to deploy two cells to optimize radio coverage: one inside and one outside the building. **Figure 4** shows

1545-0279/26©2026SMPTE

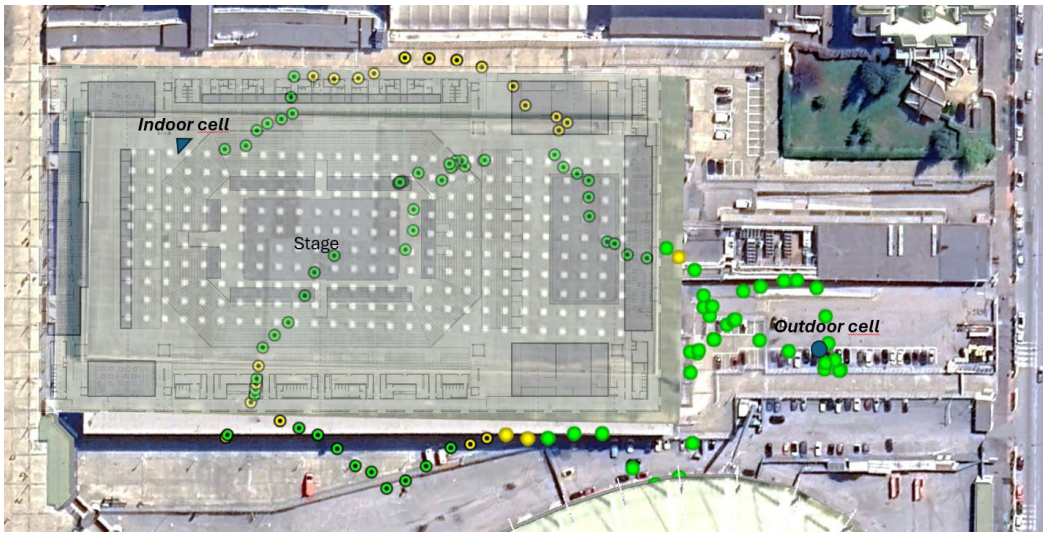


FIGURE 4. Positioning of the antennas inside Inalpi Arena and coverage measures. Dots with black square: Indoor cell; Full dots: Outdoor cell. (Quality of the Received Video: green=good; yellow=decent; red=bad/unusable for TV production).



FIGURE 5. Bottom cell; directive antenna.



FIGURE 6. Top cell; omnidirectional antenna.

the precise location, with the building's internal plan superimposed on the satellite image to better understand the antenna's position relative to the venue's architecture. For the indoor cell, a 2 × 2 sectorial antenna was used, placed in a corner of the structure at the mid-height of the audience bleachers. The intent was to use this cell to cover the camera movements around the stage and on the balconies. The second cell, an omnidirectional 2 × 2 antenna, was placed outside near the TV OB Van to support outside interviews, capture external footage of the venue, and perform handshake tests between the two cells.

The coverage results exceeded our expectations. In fact, with the aforementioned antenna positioning, we were able to receive video streams from two moving cameras, both in the stage area (center of the arena) and in the side corridors behind the seats. Furthermore, when exiting the building to the sides and back of the arena (the side opposite the indoor antenna), we verified that handover between the two cells did not cause any video interruptions or major artifacts.

In **Fig. 4**, you can see the coverage tests carried out inside and outside the structure. Unfortunately, GPS positioning on the phone is quite inaccurate indoors. However, we chose to report it for two main reasons: first, it still provides a general statistical indication of coverage; second, the handover between the two cells is clearly visible. In the image, the device is represented by full dots when connected to the external cell, and by dots with a black square when connected to the internal cell. The color of each dot (from green to red) indicates signal strength, which will be explained in detail in the following section.

Bardonecchia

In the same event, the private 5G network, with the same configuration, was tested in a second trial during the ski competition held in Bardonecchia, a ski resort near Turin. In this trial, the aim was to cover the main ski racing events (Super G and Alpine Combined) held at track 23. For camera planning, the goal was to cover both the starting and arrival areas of the race lane. This choice considered the presence of fixed cameras along the track. It gave mobile TV operators the flexibility to use wireless cameras to capture footage before the start and after each athlete's finish.

Figures 5 & 6 show the antenna positions. The cell with the directive antenna was placed near the finish line and



FIGURE 7. 5G Coverage test of lower cell (Quality of the Received Video: green=good; yellow=decent; red=bad/unusable for TV production).

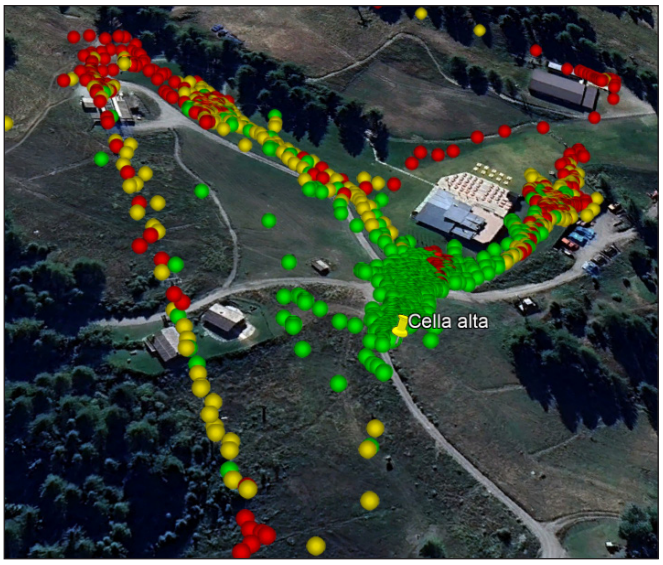


FIGURE 8. 5G Coverage test of upper cell (Quality of the Received Video: green=good; yellow=decent; red=bad/unusable for TV production).

next to the timekeeper's chalet. A 90° azimuth was sufficient to provide good coverage of the arrival and interview areas. In contrast, the omnidirectional antenna was placed at the start of the track because the desired coverage area spanned more than 180°, and it was possible to install the antenna near the middle of the useful area.

As at the previous location in Turin, we wanted to test the network coverage of the deployed system. In this case, the measurements were more accurate because, being outdoors, we could correlate signal strength with the exact GPS position with greater precision.

We measured signal strength in terms of Reference Signal Received Power (RSRP) using an Android application installed on the phones connected to the network. This measurement reflects downlink signal strength, while we are mainly interested in the uplink direction. Nevertheless, given a defined downlink transmission power and since downlink and uplink path loss are usually comparable, once we understood the minimum viable RSRP level to avoid artifacts, this measurement provides a good qualitative representation of practical coverage relative to the specific antenna positioning. In our case, with downlink transmission power set to 30 dBm, we found that to avoid artifacts and stream interruptions, the received signal had to be above –100 dBm RSRP. **Figures 7 & 8** show, in green, the points where the signal strength exceeds –90 dBm; in yellow, where it is between –100 and –90 dBm; and in red, where it is below –100 dBm.

As shown in **Figs. 7 & 8**, the two deployed cells were sufficient to cover the areas of interest. Each cell ensured a useful coverage radius of approximately 200 meters. Testing network coverage over longer distances was not possible because physical obstacles in the surrounding terrain prevented line-of-sight (LoS) with the receiving antenna.

To further assess network performance and define the system's potential, a network throughput analysis was conducted near the antenna with the network unloaded (i.e., with no other user equipments (UEs) connected). The measurements showed that, under optimal conditions and using the specific network configurations described earlier, uplink speeds of up to 440 Mbit/s could be achieved. This value is variable and depends on numerous factors, including the specific UE used for testing. Factors such as the gain of the receiving antennas, support for uplink MIMO, and the antenna's location and orientation within the UE all influence performance. **Figures 9 & 10** illustrate that, under identical conditions, some UE models can achieve better results than others. Overall, the results from all measurements confirm that the uplink transmission capacity of 5G private networks, configured as described above, is sufficient to handle the traffic demands of a medium-size TV production.

Figures 9 & 10 also show the measured 5G network latency and jitter values. The contribution of 5G network latency, around 15 ms, to the overall video delay is not considered significant, as it is generally at least an order of magnitude lower than the latency contribution of video encoding and decoding. Network jitter, on the other hand, is typically more impactful. In fact, with excessive network jitter, video receivers need large network buffers to ensure all IP packets are received in time and to prevent visual artifacts. In our case, both values were low enough that they did not require additional changes to the receiver hardware relative to the configured network parameter setup.

Finally, the end-to-end latency of the video stream was evaluated for both the Haivision and Sony encoding chains available on site. A timecode generator was used and overlaid on the source video, and an SDI monitor was placed downstream of the video mixer. Under this configuration, the minimum end-to-end delay—without introducing visual artifacts—was measured at 160 to 260 ms (8 to 13 video frames), depending on the encoding chain. These latency values are considered sufficiently low to allow typical TV production operations, particularly for use cases such as news gathering. It should be noted, however, that the measured values did not account for backhaul transmission

1545-0279/26©2026SMPTE

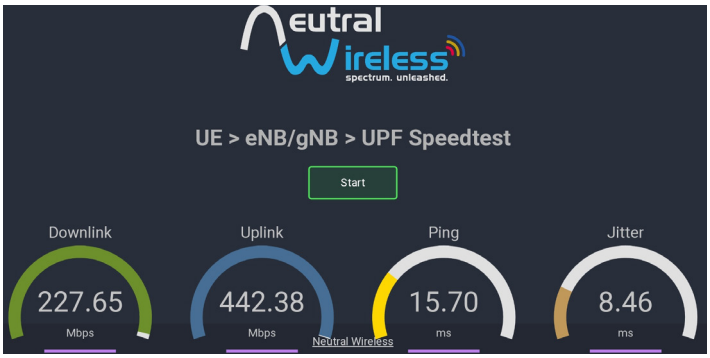


FIGURE 9. 5G Network uplink test with Samsung S23.

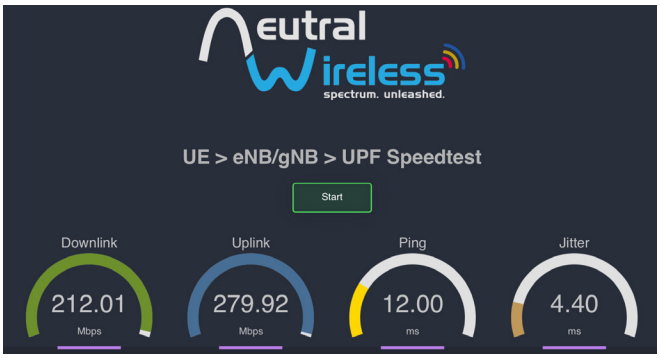


FIGURE 10. 5G Network uplink test with iPhone 15.

delay, which could be significant, especially when satellite links are used.

Interference Tests

On both occasions, Turin and Bardonecchia, interference tests were conducted to make sure we did not interfere with pre-existing public 5G systems and other RF transmissions. In fact, the frequencies granted by the Ministry for these trials (i.e. [3910-4010] MHz) were adjacent to those used in Italy for point-to-point radio links of digital terrestrial (DVB-T2) signal distribution.

Given the non-overlapping frequency bands and the large distance between the radio-link route and the experimental sites, we did not expect significant interference between the two systems, but it was a useful opportunity to verify that the system behaved as expected. As predicted, no interference was detected. In addition, since the frequencies used fell exactly between the transmission and reception frequencies used by Rai Way, they were asked to monitor in real time for any service interruptions or C/N decrease. As expected, they confirmed the total absence of interference on the radio links in question.

Conclusion

The 2025 Winter University Games, held in the Piedmont region of Italy, offered the RAI R&D department (RAI CRITS) an excellent opportunity to assess the potential of private 5G networks to support television production. The deployed 5G Standalone Non-Public Network (SNPN) was composed of two cells, each operating with 100 MHz of bandwidth in the upper n77 frequency band, made available through a temporary license granted by the Italian regulator. The user equipment included the Haivision Pro460 encoder, the Sony PDT-FP1, and several commercial smartphones. During the trial, both indoor and outdoor scenarios were tested.

As an indoor use case, 5G technology was tested to support a television production during the Olympic Opening Ceremony at the Inalpi Arena. The trial was considered successful. A single 5G cell, installed at the corner of the building, was sufficient to provide full network coverage across the entire indoor area. The received signal remained stable and clean as long as the user equipment maintained line-of-sight (LOS) with the receiving antenna. However, in non-line-of-sight (NLOS) conditions, signal stability decreased, and some video artifacts were observed on screen.

As an outdoor scenario, the 5G network was deployed at Bardonecchia Ski Resort to cover the starting area of the Olympic Track, the top of the mountain, and the arrival area near the TV compound. This scenario served as an ideal test bed for 5G technology, as deploying a traditional cabled TV production infrastructure in challenging terrain and environmental conditions is often highly complex or sometimes unfeasible. The 5G network proved to be very convenient, thanks to its portability and ease of installation and configuration. Indeed, deploying the 5G network was very fast; it took only half a day to set up, configure, and make it fully operational.

The network coverage was sufficient to support all camera operations around the cells. Uplink throughput was also remarkable: we measured up to 400 Mbit/s of uplink throughput in the proximity of antennas. Also, the feedback gathered from the operators and technicians was positive. Video streams from the professional encoders were stable and clean as long as they remained within the cell's useful coverage range. Depending on the hardware used, measured end-to-end video latency ranged from 160 to 260 ms, low enough to support typical TV production operations. The responsiveness of the remotely controlled cameras was also good, thanks to the 5G network's low latency (around 15 ms). In contrast, the measured video latency of smartphone feeds was around 1 sec, too high for TV operations.

However, despite these technological advantages, adoption of private 5G networks for TV production remains uncertain, as there is currently no established procedure for temporary frequency applications to regulators. For this reason, the experiment was also of interest to the regulator to gather information and feedback to develop suitable procedures.

Finally, the Eutelsat OneWåeb solution, based on LEO satellite constellations, was also tested during the PoC as an IP backhaul solution for transmitting video streams between the event venue and the RAI Production Center in Turin. The tests showed that, although this satellite system is not yet optimized for television use, setting an adequate network buffer (i.e., 2-3 sec) on the receiving video equipment made it possible to successfully transmit and receive a 1080p50 video stream at 8 Mbit/s.

Acknowledgments

The authors would like to thank their RAI CRITS colleagues A. Bertella, A. Cafasso, F. Debandi, A. Faga, R. Iacoviello, S. Mantovano, D. Milanesio, and G. Vitale for their technical contributions to the work. We also thank Samuel R. Yoffe and Kenneth W. Barlee (Neutral Wireless) for their on-site support and assistance, and Manuel Faccioli (Fondazione Ugo Bordoni) as activity supervisor.

Moreover, the authors would like to express their sincere gratitude to all the partners who contributed to the success of this project: Neutral Wireless, Eutelsat, Sony, EMG, One TV, CVE, the FISU Committee and the Municipality of Turin.

References

1. A. Stewart, "Sub-6 GHz Slicing on Public Networks—Field Results, Observations and Roadmap to Ubiquitous 5G Access for Professional Media," presented at IBC 2024, Amsterdam, The Netherlands, Sept. 2024.
2. European Broadcasting Union, *Tech Report 080 – EBU Members' Trials of 5G in Content Production and Contribution*, Mar. 2024.
3. G. Stante, D. Airola Gnota, F. Debandi, M. Serafini, L. Vignaroli, F. Graglia, G. Scozza, E. Mastromartino, M. Fiammengo, and L. Guida, "5G Jazz Jam Session – An Experiment in Containing Arts," presented at IBC 2023, Amsterdam, The Netherlands, Sept. 2023.
4. *Torino 2025 FISU World University Games – Media Guide*, White Paper, 2025.
5. M. B. Waddell, S. R. Yoffe, K. W. Barlee, D. G. Allan, I. Wagdin, M. R. Brew, D. Butler, and R. W. Stewart, "5G Standalone Non-Public Networks: Modernising Wireless Production," presented at IBC 2023, Amsterdam, The Netherlands, Sept. 2023.
6. M. B. Waddell, J. Kröger-Mayes, S. R. Yoffe, D. G. Allan, M. Brandstrup, J. Christoffersen, M. R. Brew, and R. W. Stewart, "Optimizing 5G for Low-Latency Broadcast Production," presented at IBC 2024, Amsterdam, The Netherlands, Sept. 2024.
7. S. R. Yoffe, D. G. Allan, K. W. Barlee, D. Anderson, M. R. Brew, R. W. Stewart, M. B. Waddell, J. Kröger-Mayes, and G. Stante, "Low-Latency Wireless Broadcast Production over 5G," Proc. *NAB Broadcast Engineering and Information Technology (BEIT) Conference*, Apr.2025.
8. Haivision, "SST Protocol,"Haivision. [Online]. Available: https://www.haivision.com/products/sst-protocol/

About the Authors



Giulio Stante, a research engineer at RAI, focuses on private 5G mobile networks and innovative IP network solutions for TV production.



Alessandro Lucco-Castello has been a computer engineer and RAI researcher since 2018, specializes in new television standards and convergence with mobile networks.



Federico Maria Pandolfi, R&D engineer at RAI, specializes in IP-based solutions for AV distribution over DVB, 5G, and Wi-Fi networks.



Luca Vignaroli (PhD), a RAI researcher since 1998, is an ICT and TV production expert focusing on 5G, DVB, and media-network convergence.

Presented at IBC 2025. Amsterdam, The Netherlands, 12-15 Sept. 2025. This paper is published here by kind permission of the IBC. *Copyright © IBC.*

DOI: 10.5594/JMI.2026/POET9825
Date of publication: 1 October 2026

Design your next products with DekTec



DTA-2172
2x 3G-SDI/ASI
Low profile
Genlock



DTA-2174B
4x 3G-SDI/ASI (1x 12G)
Single or quad-link 4K
Genlock



DTA-2178
8x 3G-SDI/ASI (2x 12G)
8x 12G-SDI with scaling
Genlock

DekTec
www.dektec.com
(303) 318-4298
infousa@dektec.com
Also available:
Satellite, QAM, DVB-T2, ATSC 3.0 receiver and modulator, and ASI I/O

1545-0279/26©2026SMPTE

KEYWORDS SPORTS BROADCASTING // COMPUTER VISION // AUGMENTED REALITY // CYCLIST RECOGNITION

Contextual AR in the Wild: Enhancing Live Cycling Broadcasts

By W. Clinckemaillie, J. Vanhaeverbeke, and S. Verstockt

Abstract

Advanced computer vision and machine learning are transforming sports broadcasting. While augmented reality has successfully enhanced fan engagement in controlled environments like soccer and basketball stadiums, its application in challenging outdoor sports like cycling is virtually non-existent. Over the past two years, we have focused on bridging this gap by developing a fully automated workflow to virtually enrich cyclists and other key objects in live race broadcasts. The first part of our work automates rider and team detection captured from diverse viewpoints. Combining modern object detection, Siamese networks for similarity learning, and advanced tracking offers a robust solution for cyclist identification and tracking. The second part of our work explores the use of 3D pose estimation of cyclists to create augmented reality visualizations. Finally, we also focus on the development of an automatic pipeline for POI detection, recognition, and visualization, which facilitates graphics generation, a process that was previously done manually.

"We need to find formats that make cycling broadcasts more popular," these are the words of Tomas Van Den Spiegel, CEO of Flanders Classics, leading up to the 2022 edition of the Ronde van Vlaanderen (Tour of Flanders), a major annual professional road cycling race held in Flanders, Belgium. Now, four years later, not much has changed, and others also confirm the need for more personalized and interactive digital sports experiences. Recent research also shows that the average viewer of traditional sports broadcasts is aging, and younger fans are dropping out. There is a need for innovation in sports storytelling, and recent developments in video enrichment and augmentation can contribute to this. The VENTOUX[i] project and its resulting tools are our answer to keeping sports fans engaged in a time when on-demand information has become mainstream, and viewers are looking for new forms of sports coverage.

VENTOUX—a Digital Transformation Media project funded by the Vlaams Departement Cultuur, Jeugd en Media (CJM—Flanders Department of Culture, Youth and Media)—focuses on making cycling broadcasts more interactive. Based on object-level metadata generated by artificial intelligence (AI), VENTOUX optimizes the searchability of the video, makes the video more interactive, and exploits the content more widely. This paper discusses detection, recognition, tracking, and enrichment of riders and points of interest (POIs) and demonstrates the potential and feasibility of our flexible methodology.

In this paper, augmented reality (AR) is used in the broadcast sense to denote computer-generated graphics composit-

[i] https://www.victoris.be/projects/ventoux/

ed into the video signal and spatially anchored to real-world subjects or scene locations using estimated geometry and tracking. Unlike interactive see-through or headset-based AR, the viewer watches a conventional broadcast stream augmented with scene-aligned graphics. We use the broader term on-screen graphical overlay when spatial anchoring is not required.

The paper synthesizes two research strands of the VENTOUX project. The cyclist recognition and AR visualization work summarizes research first reported by Clinckemaillie et al.[1] and subsequently extended in,[2] while the point-of-interest recognition and tracking work summarizes Vanhaeverbeke et al.[3] The present paper brings these two strands together from a broadcast engineering perspective.

Cyclist Recognition and AR Visualization

Model overview – cyclist pipeline. YOLOv11 detects cyclist upper bodies; BoT-SORT maintains rider tracks; a Siamese network performs one-shot team matching from jersey appearance; and TokenHMR estimates monocular 3D body pose/mesh information to derive overlay anchors. Each model is described in more detail at its first use below.

Cycling broadcasts present unique challenges for real-time athlete identification. Unlike stadium sports, cycling takes place in uncontrolled outdoor environments with dynamic camera angles, rapid motion, and visually similar uniforms. Producers often identify riders manually, relying on bib numbers or race context. However, in helicopter footage especially, bib numbers are usually not visible due to distance, angle, or compression. These limitations hinder the live presentation of rider information and the ability to extract structured data from footage after the race.

To address this, we present a modular cyclist recognition and visualization system that generates real-time metadata, supports augmented reality (AR) overlays, and enables richer post-race analysis. The pipeline is structured in five main stages: (1) cyclist detection and tracking, (2) team recognition, (3) 3D bounding box estimation, (4) temporal postprocessing, and (5) AR overlay placement. **Figure 1** provides an overview of the system. The architecture is modular by design: each component can be optimized or replaced independently depending on compute constraints or production requirements.

The first stage in the pipeline detects and tracks cyclists in the video frames. This provides the input regions for both team recognition and pose estimation. Detection focuses on each rider's upper body, as this region typically contains the most distinctive jersey features, such as logos, team colors, and sponsor patterns. To train and evaluate the detection models, we constructed a custom dataset[2] of over 20,000 annotated bounding boxes, collected from multiple cycling races. The annotations include cyclists and motorcycles to reduce false positives in dynamic race scenes. We trained and compared multiple versions of the YOLO (You Only Look Once) object detection architecture, including both YOLOv8 and YOLOv11, across different model sizes and input resolutions.[4] The goal was to find an optimal trade-off between detection accuracy and inference speed for integration into near-real-time broadcast workflows. Overall, the YOLOv11-large model at 1920 × 1080 resolution provided the most balanced result for our use case.

To maintain rider identity across frames, we use the BoT-SORT tracking algorithm, which combines motion prediction with lightweight appearance matching.[5] Compared to more complex tracking frameworks, BoT-SORT offers lower computational overhead, making it suitable for real-time use in the whole pipeline without sacrificing track consistency.

Because bib numbers are not consistently detectable, the system classifies cyclists at the team level by analyzing jersey appearance. While this approach does not yield individual identification, it is often sufficient in practice. In many race situations, such as breakaways or small front groups, typically only one rider per team is visible at a time. Combined with metadata from the broadcaster's camera feed, indicating which group or rider is being filmed, team recognition becomes a reliable and scalable method for supporting automated overlays and metadata enrichment.

To recognize teams by jersey appearance, we employ a Siamese neural network,[6] as illustrated in **Fig. 2**, trained with contrastive loss. This network processes pairs of images—one from the race footage and one from a known reference set—and learns to embed them so that visually similar jerseys are positioned close to each other in the feature space. In contrast, dissimilar jerseys are placed further apart. This form of metric learning enables the model to generalize to unseen teams without explicit class-based retraining. During inference, each detected torso crop is compared to a small set of reference images using cosine similarity, and the closest match is selected through a k-nearest neighbors (k-NN) classifier.

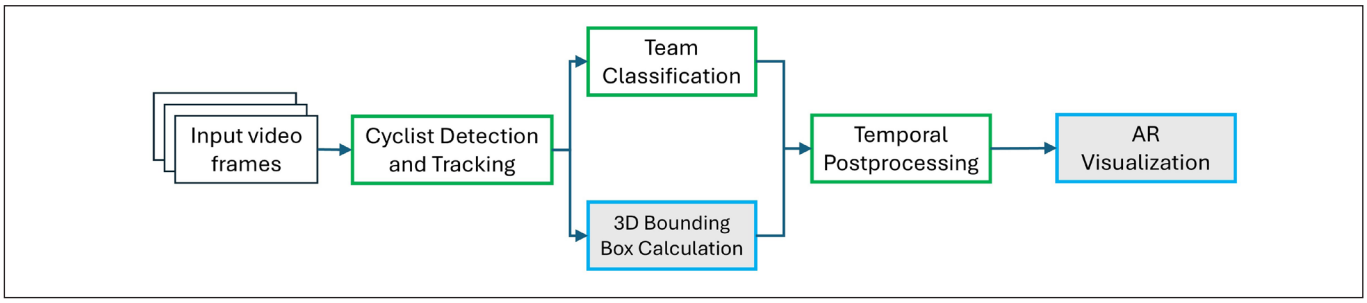


FIGURE 1. Overview of the proposed pipeline for cyclist recognition and AR visualization.

[2] https://universe.roboflow.com/ugent-hkozj/ventoux-chxhg

1545-0279/26©2026SMPTE

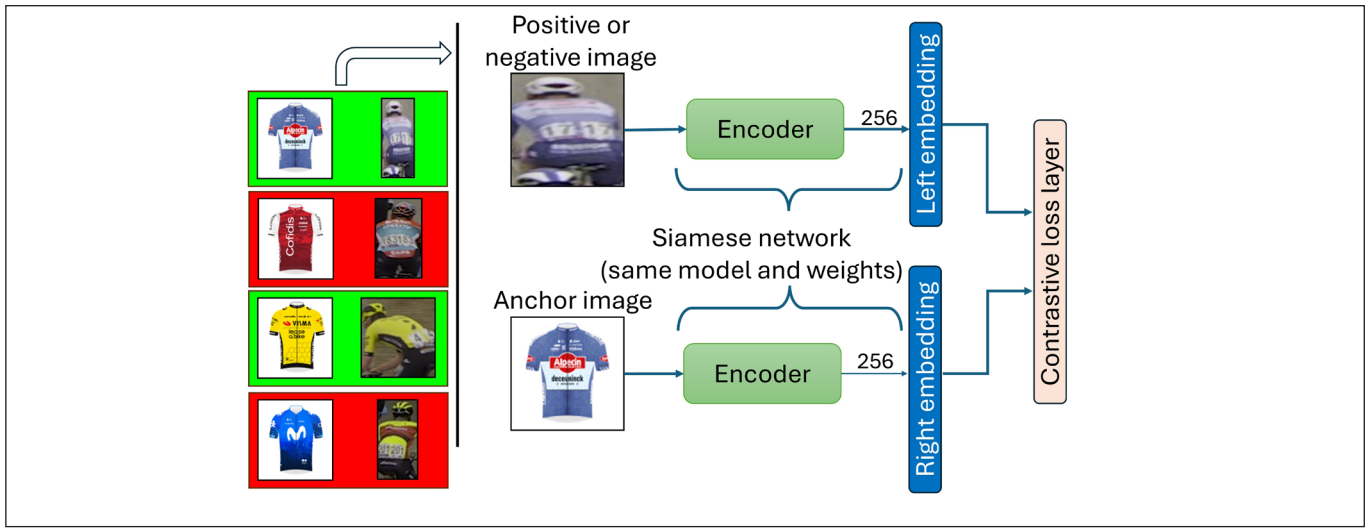


FIGURE 2. Siamese network architecture and examples of positive (green border) and negative (red border) pairs used in training.

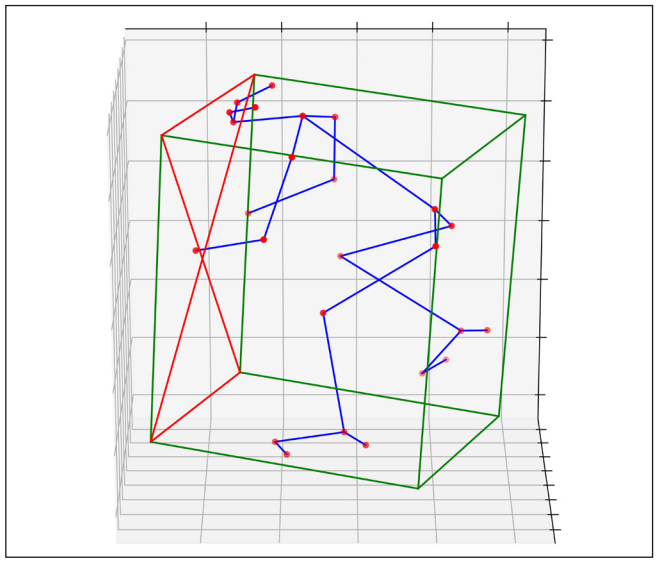


FIGURE 3. 3D-oriented bounding box.

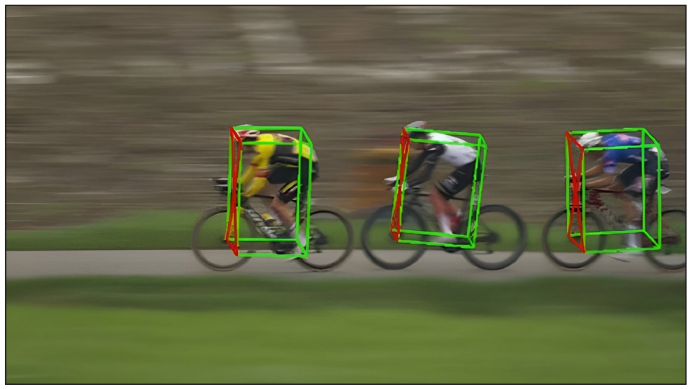


FIGURE 4. Projected 3D bounding boxes overlaid on cyclists.

To improve the model's ability to distinguish between teams with similar colors or patterns, we applied hard negative mining during training. This strategy involves selectively including challenging non-matching examples, such as teams with nearly identical base colors, to enhance the network's sensitivity to subtle differences. We trained the model on 24 professional teams, using approximately 80 cropped torso images per team. In validation, the model achieved 94% accuracy for teams seen during training. To evaluate its generalization capability, we tested the model on 15 unseen teams with new kits, achieving 85% accuracy. These results indicate that the system is robust to moderate variations in jersey design. This method allows new teams or special kits (for example, national champions) to be added at inference time without retraining. This flexibility is crucial in cycling and other sports, where kits often change between races or seasons. A small number of reference images is enough to integrate new teams into the system.

To accurately place augmented visual elements, we estimate each cyclist's 3D orientation and spatial extent using pose information. The goal is to construct a bounding box that captures the rider's position and heading in 3D space, allowing later overlay elements to remain adequately aligned as camera angles and motion change. The estimation is based on keypoints obtained from Human Mesh Recovery (using TokenHMR), which predicts a 3D body skeleton from a single image crop.[7] From this set of 3D points, a directional vector is calculated that approximates the rider's forward motion. Using this vector in combination with the spatial spread of the keypoints, a 3D-oriented bounding box is constructed around the predicted pose (see **Fig. 3**).

Since the keypoints are defined in 3D space, perspective projection into the 2D image frame is possible, as illustrated in **Fig. 4**. The projected 3D bounding box serves as a dynamic anchor for augmented elements, such as team labels or rider indicators, which benefit from orientation-aware placement. This improves overlay stability during movement and ensures visual coherence under varying camera angles.

To improve temporal consistency and visual stability, we apply lightweight postprocessing to detection, classification, and pose estimation outputs. Track IDs are only assigned after several consecutive detections, which helps suppress false positives and maintain continuity during short occlusions.

For team recognition, classification scores are smoothed over time using a moving average per track, reducing label flicker from single-frame errors. Missing 3D pose estimates are filled in by linear interpolation, ensuring overlays remain anchored even when keypoints are temporarily unavailable. Finally, jitter in the 3D bounding boxes is reduced through exponential smoothing, resulting in more stable and coherent AR overlays under motion.

The final stage of the pipeline integrates augmented visual elements into the video stream based on the computed 3D bounding boxes. These overlays help convey contextual information, such as rider identity or team affiliation, while maintaining spatial coherence with the moving cyclists. Multiple alignment strategies are possible for placing text or graphics in the frame. Each 3D bounding box yields four corner points defining a virtual text plane. Depending on the intended position—above the rider, in front on the ground, or perpendicular to the direction of motion—these points are projected onto the 2D frame using a perspective transformation. **Figure 5** shows three examples of such alignment strategies. In future implementations, the system could automatically choose the best alignment based on factors like rider direction, available space, and visibility, possibly aided by road segmentation.

Figure 6 illustrates overlays where name tags are dynamically attached to each cyclist based on the 3D bounding box projection. These tags move and rotate consistently with the rider's estimated pose and remain readable even when the camera moves or the cyclist changes orientation. To maintain visual clarity, overlays are limited to small groups, such as breakaways, where rider positions are less crowded and easier to distinguish.

Beyond text, the system also supports alternative AR visualizations (shown in **Fig. 7**). These include ground-level ellipses to highlight a rider's position and directional arrows that could indicate heading or wind direction, provided such metadata is available. These elements further enhance tactical insight for viewers without cluttering the screen.

Although the method supports dynamic overlay placement, achieving full real-time performance remains challenging. On an NVIDIA GTX 1080 Ti, pose estimation via TokenHMR requires ~80 ms per frame. On faster hardware such as the NVIDIA A100, this drops to ~35 ms, but real-time performance at broadcast frame rates still requires optimization. We apply frame skipping and interpolate intermediate poses from keyframes to bridge this gap. This reduces computational load and enables overlays with a fixed and predictable delay, typically just a few frames, which is acceptable in most broadcast settings. This trade-off ensures smoother overlays without requiring full-frame inference at all times.

This modular pipeline demonstrates how modern computer vision techniques can support semi-automated cyclist recognition and augmented visualizations in live cycling broadcasts. While real-time execution remains a technical challenge, the proposed methods enable flexible, scalable metadata enrichment with potential applications in both live and post-race production workflows.

Point of Interest Recognition and Tracking

Road cycling races not only highlight competition between elite cycling teams but also showcase regional landmarks. These points of interest (POIs) provide variety during broadcasts and promote local tourism, which has been an integral part of cycling coverage for many years.

Traditionally, when a POI appears in helicopter footage, its name is shown using a static on-screen overlay. While effective in the past, this method is now increasingly outdated for modern audiences, who expect more dynamic and interactive viewing experiences. Additionally, creating and managing these overlays demands manual effort from video editors during live events.

To address these challenges, this section introduces a fully automated method for recognizing POIs in real time during live broadcasts. This automation reduces the strain on video editors and enables continuous tracking of each identified POI while it remains in view. As a result, broadcasters can deliver more engaging and innovative visualizations to viewers.

The proposed methodology consists of five steps designed to be robust and flexible. This ensures the system can be extended to additional POIs and races without requiring large, labeled datasets for each new POI.

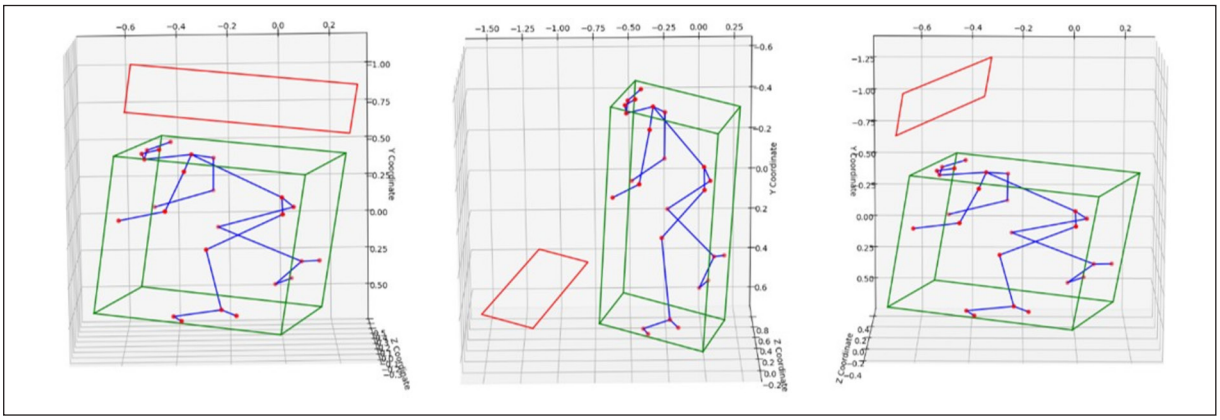


FIGURE 5. Several options for positioning the text area.

1545-0279/26©2026SMPTE



FIGURE 6. Examples of name tag AR overlays.

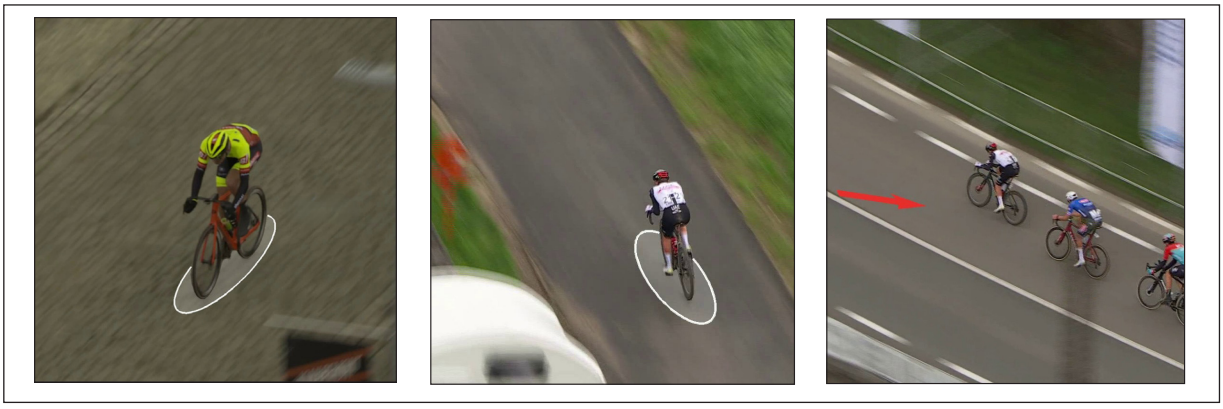


FIGURE 7. Examples of alternative AR visualizations: rider hotspot projected on the ground (left, center) and an arrow indicating wind direction (right).

Methodology

Model overview – ROI pipeline. UNISAL predicts visually salient image regions; SAM segments the candidate object; SuperPoint and SuperGlue match it against POI reference images; and MedianFlow or SAM 2 tracks the recognized POI across subsequent frames. Each model is described in more detail at its first use.

The first step involves verifying the helicopter's location to determine whether any points of interest are nearby. The haversine formula is used to calculate the distance between the helicopter's GPS position and the POIs stored in a database. If no POI is found within a 500-meter range, the POI recognition pipeline is paused and reattempted after a one-second interval. This approach ensures that subsequent, more computationally intensive steps are executed only when necessary.

When the helicopter is near one or more POIs, the pipeline advances to the saliency detection phase. Saliency detection identifies the most visually prominent areas within the captured images or video frames. For this, the UNISAL[8] model is utilized. Trained on eye-tracking data, UNISAL produces a heatmap that reflects human visual attention by highlighting the most engaging regions. Since POIs are intended to draw attention, they are expected to be prominent in the saliency map, as demonstrated in **Fig. 8**. This information allows subsequent steps to focus on these areas, thereby avoiding the need to analyze the entire image.

The resulting heatmap then informs the next stage, which involves identifying the most prominent object within the image. This is achieved by locating the points with the highest and lowest saliency on the heatmap, which are then used as prompts for the Segment Anything Model (SAM).[9] SAM can segment any object in an image based on a given prompt, such as input points, producing a segmentation mask. As illustrated in **Fig. 9**, this enables the effective isolation of regions of interest, such as a church labeled in blue, allowing the POI recognition to focus on the most prominent object.

In the fourth step, the identified salient object is further evaluated to determine if it effectively corresponds to a POI. The segmentation mask is used to crop the image around the salient object, which serves as the input for the POI rec-



FIGURE 8. Example visualization of the saliency heatmap predicted by the UNISAL model.

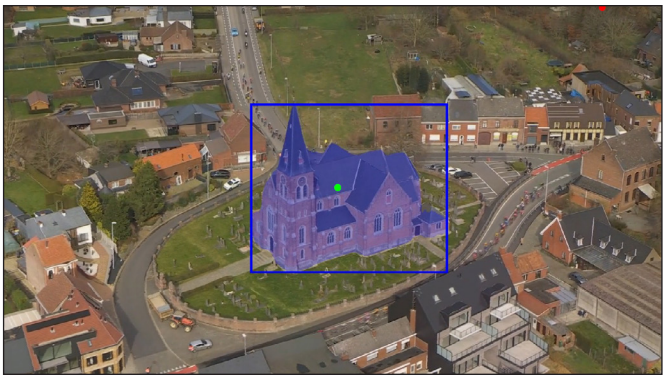


FIGURE 9. Example visualization of the predicted mask and derived bounding box by the Segment Anything Model (SAM) when prompted with the most and least salient point.

ognition process. To maintain flexibility and extensibility and support new POIs and races, the system uses an image-matching approach. The cropped image is compared to reference images of nearby POIs by extracting image keypoints and descriptors using the SuperPoint algorithm,[10] followed by matching with the SuperGlue model,[11] as illustrated in **Fig. 10**. When a high number of matches is found between the salient object and a reference image, it indicates that the point of interest represented by the reference image has been correctly recognized. This approach requires only about three reference images per point of interest, making it easily extensible to new POIs.

Once the POI is recognized, the final step is to track it throughout the live broadcast. Running the full recognition pipeline on every frame is computationally inefficient and would introduce delays in live coverage. To address this, an optical tracking algorithm is started after the initial recognition, enabling the system to follow the region of interest across subsequent frames. Various trackers, such as MedianFlow[12] and SAM 2,[13] have been tested and are available for use in the final implementation, enabling real-time operation during the live broadcast.

Live Integration and Visualization

This methodology is demonstrated through integration into a realistic broadcasting environment. The live helicopter video feed is delivered to a processing server via a Serial Digital Interface (SDI) connection, which is a standard in professional video production. Based on that, the server performs point-of-interest recognition and tracking as outlined in the methodology. Results are then transmitted to the visualization engine, such as Vizrt, using the User Datagram Protocol (UDP). When both the processing and visualization servers operate within the same local network, communication is both fast and reliable. This enables a decoupled system architecture, allowing the processing and visualization servers to function independently.

The visualization engine uses the live recognition and tracking data to trigger and display visualizations for viewers. Given the available tracking information, one visualization approach is to dynamically anchor POI information to the POI's on-screen location. This provides a more dynamic and visually appealing experience compared to traditional static overlays. This type of visualization is implemented in Vizrt to demonstrate the complete methodology, as shown in **Fig. 11**.

The methodology's performance is optimized for live broadcasting. The total recognition process takes approximately 250 milliseconds (ms) per check. Although this can be improved with more advanced hardware, it is already sufficient, as the system conducts these checks only once per second, which is adequate for the typical appearance rate of new POIs. The optical tracker must operate in real time to meet live broadcast demands. This is achievable depending on hardware specifications and the selected tracking algorithm. For instance, the MedianFlow tracker operates at around 10 milliseconds per frame, enabling smooth, real-time tracking during live coverage.

Conclusion

In conclusion, this work demonstrates how modern computer vision methods can enhance live cycling broadcasts through contextual augmented reality. We developed a mod-

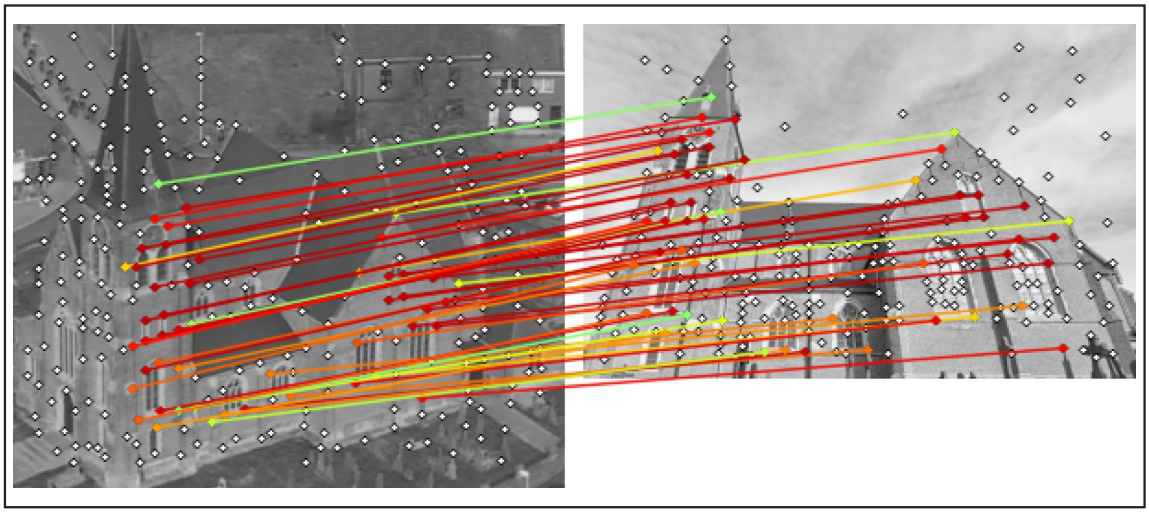


FIGURE 10. Example visualization of the keypoints and matches generated by the SuperGlue model between the most salient object in the helicopter view (left) and a POI reference image (right).

1545-0279/26©2026SMPTE



FIGURE 11. Example visualizations generated by Vizrt based on the POI recognition and tracking data.

ular pipeline for cyclist recognition and accurate overlay placement, combining detection, tracking, team classification, and 3D pose estimation into a system that supports live and post-race applications. Additionally, we introduced a real-time approach for recognizing and tracking points of interest, reducing manual workload and enabling more dynamic storytelling. Real-time performance remains a limitation of the current implementation. At 25 fps, a video frame lasts 40 ms, while TokenHMR-based pose estimation requires approximately 35-80 ms per frame depending on the hardware, and a full POI recognition pass takes approximately 250 ms. We therefore reduce the update rate of these computationally expensive stages: rider pose estimation is performed on subsampled frames and interpolated in between, while POIs are recognized periodically and then tracked in real time. For POI overlays, the visual impact is limited, since landmarks are static and the recognition latency delays only the moment an overlay appears rather than its position. For rider overlays, the impact is more visible: interpolation and temporal smoothing improve stability but may introduce a small delay or temporary misalignment during rapid motion.

The same methodology extends to other large-scale, non-stadium sports such as marathons and triathlons, where cameras are mobile, and athletes appear against uncontrolled backgrounds. POI recognition transfers directly to any aerial or mobile coverage with a geotagged feed and a landmark database, and appearance-based athlete recognition transfers wherever participants can be grouped into a small number of visually distinctive team or national apparel. Pose-based overlay placement transfers less readily: in cycling the bicycle constrains posture and yields a stable heading vector, whereas running and swimming articulate considerably more. Despite these differences, the proposed framework highlights the potential of context-aware augmented reality to make live coverage of large-scale outdoor sports more informative and engaging.

Acknowledgments

The authors thank their EMG, Wielerorakel, and IMEC colleagues for their contributions to this work. VENTOUX received funding from the Vlaams Departement Cultuur, Jeugd en Media (CJM – Flemish Department of Culture, Youth and Media) (call digital transformation Media - project number 94186); of course, the authors also want to thank them for their support. Finally, they would also like to thank the International Broadcasting Convention for permission to publish this paper.

References

1. W. Clinckemaillie, J. Vanhaeverbeke, M. Slembrouck, and S. Verstockt, "One-Shot Team Recognition and 3D Pose Estimation of Cyclists for Augmented Reality Visualization," Lect. Notes Comput. Sci., 15925:36–52, 2025.
2. W. Clinckemaillie, J. Vanhaeverbeke, M. Slembrouck, and S. Verstockt, "An end-to-end pipeline for team-aware, pose-aligned augmented reality in cycling broadcasts," Comput. Vis. Image Underst., 263:104602, Dec. 2025.
3. J. Vanhaeverbeke, R. Decorte, M. Slembrouck, S. Van Hoecke, and S. Verstockt, "Point of Interest Recognition and Tracking in Aerial Video during Live Cycling Broadcasts," Appl. Sci., 14(20):9246, Oct. 2024.
4. G. Jocher and J. Qiu, "Ultralytics YOLO11," Sept. 2024.
5. N. Aharon, R. Orfaig, and B. Z. Bobrovsky, "BoT-SORT: Robust Associations Multi-Pedestrian Tracking," arXiv, 2206.14651, Jun. 2022.
6. G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese Neural Networks for One-Shot Image Recognition," ICML Deep Learning Workshop, 2015.
7. S. K. Dwivedi, Y. Sun, P. Patel, Y. Feng, and M. J. Black, "TokenHMR: Advancing Human Mesh Recovery with a Tokenized Pose Representation," Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 1323–1333, Jun. 2024.
8. R. Droste, J. Jiao, and J. A. Noble, "Unified Image and Video Saliency Modeling," Lect. Notes Comput. Sci., 12350:419–435, Oct. 2020.
9. A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W. Y. Lo, et al., "Segment Anything," Proc. IEEE/CVF Int. Conf. Comput. Vis., 4015–4026, Oct. 2023.
10. D. DeTone, D. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-Supervised Interest Point Detection and Description," Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops, 224–236, Jun. 2018.
11. P. E. Sarlin, D. DeTone, D. Malisiewicz, and A. Rabinovich, "SuperGlue: Learning Feature Matching with Graph Neural Networks," Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 4938–4947, Jun. 2020.
12. Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-Learning-Detection," IEEE Trans. Pattern Anal. Mach. Intell., 34(7):1409–1422, Jul. 2012.
13. N. Ravi, V. Gabeur, Y. T. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson, et al., "SAM 2: Segment Anything in Images and Videos," arXiv, 2408.00714, Aug. 2024.

About the Authors

Winter Clinckemaillie is a PhD student at IDLab, Ghent University-imec, specializing in computer vision systems for sports, safety and AR applications.

Jelle Vanhaeverbeke is a senior postdoctoral researcher at IDLab, specializing in event detection, metadata enrichment, and computer vision.

Steven Verstockt leads IMPACT at IDLAB and researches multimedia and computer vision across sports, safety, and industrial applications.

Presented at IBC 2025. Amsterdam, The Netherlands, 12-15 Sept. 2025. This paper is published here by kind permission of the IBC. *Copyright © IBC.*

DOI: 10.5594/JMI.2026/HLGH4233
Date of publication: 1 October 2026

KEYWORDS VIRTUAL REALITY // ORTHOSTEREOSCOPIC IMAGING // VIRTUAL REALITY FILM PRODUCTION PIPELINE // CINEMATOGRAPHY

Optimizing 180° Stereoscopic Cinematography for the Apple Vision Pro

By Stanley Lin, Nathan Fairchild, Spencer Lin, Akaash Tumuluri, Jordan Halsey, and Scott Fisher

Abstract

The Apple Vision Pro establishes a new benchmark for image fidelity in head-mounted displays, creating both opportunities and challenges for cinematic VR filmmaking. Producing content that fully leverages its display capabilities requires reconsideration of image acquisition, post-production, and delivery workflows. This paper presents a replicable 180° orthostereoscopic imaging pipeline designed to balance accessibility, reliability, and cinematic quality while meeting the demands of next-generation virtual and mixed reality displays. Using our experimental VR short film *A Larger Living Space* as a case study, we outline an end-to-end workflow that integrates structured capture techniques, AI-based image enhancement, and scalable visual effects methods. Our pipeline emphasizes preservation of fine detail, reduction of noise and artifacts, and seamless integration of practical and virtual lighting within wide fields of view offered by dual-fisheye lenses available on the market. By uniting prosumer-grade tools with scalable post-production methods, our approach demonstrates that high-fidelity immersive video can be achieved outside of large-scale studio contexts. This work contributes a consolidated framework for immersive cinematography and provides a foundation for future exploration into higher bit depth capture, high frame rate imaging, and real-time virtual set integration for VR film productions.

The Apple Vision Pro (AVP) introduces an unprecedented standard of image fidelity in head-mounted displays (HMDs), enabling highly immersive VR cinematic experiences. However, leveraging its full potential requires rethinking both acquisition and post-production workflows for stereoscopic content. In this paper, we present a novel imaging and post-production pipeline that addresses the core challenges of 180° stereoscopic cinematography optimized for AVP delivery while using our original VR short film, *A Larger Living Space*, as a case study **(Fig. 1).** Our approach integrates prosumer-grade camera systems with AI-enhanced post-processing techniques to achieve high-fidelity immersive video. This approach can be broken down into four primary stages: image acquisition, stereoscopic preprocessing and VFX compositing, AI-assisted image restoration and upscaling, and Apple Vision Pro delivery.

We begin by outlining an optimized imaging pipeline built around the Canon R5 C and the RF 5.2mm f/2.8 Dual Fisheye lens, henceforth referred to as the Camera and the Lens, respectively. While this system offers a complete and accessible orthostereoscopic imaging camera package with a 190° angle of view, it also has significant limitations, including suboptimal sensor noise performance, low optical resolving power, and framing/lighting constraints inherent to ultra-wide-angle-of-view (AOV) capture. To address these limitations, we implemented structured image acquisition strategies, including a one-stop overexposure method to improve sensor noise performance, aperture selection to balance resolving power and light transmission, and a hybrid lighting design that leveraged virtual ceiling extensions, embedded LED practicals, and rigged overhead space lights.

1545-0279/26©2026SMPTE



FIGURE 1. Color-graded frame from *A Larger Living Space*, showcasing the fully denoised and upscaled footage with a computer-generated ceiling seamlessly noise-matched to the live-action plate (2025).

Crucially, we developed a scalable post-production pipeline designed to meet the display demands of the AVP. After stereoscopic reprojection, we processed the footage through two discrete AI-assisted restoration stages in Topaz Video AI: an initial denoising pass using Nyx, followed by a separate Proteus enhancement and 2× upscale to a 16384 × 8192 stereo equirectangular master, corresponding to 8192 × 8192 pixels per eye. This increased equatorial sampling density from approximately 22.8 to 45.5 pixels per degree while preserving the native 59.94p temporal cadence. The higher-resolution reconstruction was intended to increase sampling density and improve the reproduction of edges and fine textures during headset presentation rather than to recover optical information absent from the original capture. This departs from legacy post-processing techniques like total variation and wavelet denoising, which overly degrade detail critical to the immersive experience. Final masters were graded for P3-D65 primaries with an ST 2084 (PQ) transfer function and encoded as MV-HEVC for stereoscopic playback on Apple Vision Pro.

We also developed a visual effects (VFX) and compositing pipeline using tools such as Houdini, 3D Equalizer, and Nuke, relying on Pixar USD for render-agnostic scene assembly. Our virtual set extensions (particularly for ceiling replacement) enabled sophisticated lighting design within the extreme AOV of the Lens without compromising the frame. The completed stereo master was encoded in MV-HEVC with embedded spatial metadata for stereoscopic playback on AVP.

This research provides a blueprint for high-fidelity immersive film production using accessible equipment and offers a comprehensive, documented technical workflow tuned for the Apple Vision Pro. Beyond technical enhancements, our imaging pipeline contributes meaningfully to immersive storytelling by supporting more cinematic lighting, refined viewer gaze control through luminance contrast, and more naturalistic motion presentation.

Our work aims to democratize high-quality VR filmmaking by making high-fidelity cinematic results attainable for lower-budget productions. We hope to spark further exploration into 180° stereoscopic content design, enabling independent creators to meet emerging standards in XR media consumption. Future work will explore higher-bit-depth acquisition, 120p capture, and real-time motion-tracked VFX set extensions using camera telemetry data for accurate CG scene integration and single- or multi-axis camera movements.

Inspirations and Related Works

Recently, interest in stereoscopic VR capture systems has increased in both commercial and academic domains. Camera systems such as the KanDao QooCam EGO, Z CAM K1 Pro, and the upcoming Blackmagic URSA Cine Immersive each

target immersive imaging with varying trade-offs in resolution, affordability, and workflow maturity. The Blackmagic URSA Cine Immersive, in particular, is positioned as the most complete high-end solution: bundled with DaVinci Resolve Studio, it offers a highly integrated camera-to-headset workflow. However, this functionality is only accessible through the purchase of the camera itself, which retails for $29,995 USD,[1] effectively limiting its availability to creative teams with substantial budgets.

Prior academic work has largely focused on isolated aspects of immersive imaging rather than integrated workflows. Studies and articles have examined the relationship between frame rate and simulator sickness,[2] evaluated noise-reduction algorithms,[3] and benchmarked individual camera functionality.[4] Similarly, industry documentation from manufacturers provides specifications and operational guidelines but rarely addresses how to deploy these devices systematically to achieve cinematic imagery in VR contexts.

To verify this gap, we conducted a scoping review of immersive imaging literature between 2020–2025, searching multiple electronic databases with terms including stereoscopic video, immersive video, and VR cinematography. While we identified research on 360° video and volumetric capture pipelines, we found no studies that consolidated best practices into an end-to-end 180° orthostereoscopic imaging workflow suitable for modern head-mounted displays.

Our contribution addresses this absence. Unlike existing documentation, this work is specifically optimized for the Apple Vision Pro, whose 34 PPD acuity and high dynamic range (HDR) display impose stricter sharpness and noise thresholds than traditional cinema. The result is a practical framework enabling filmmakers and creatives to achieve high-fidelity immersive imagery from prosumer stereoscopic hardware at a fraction of the cost of high-end systems like the URSA Cine Immersive coupled with its immersive media post-production suite.

Imaging Hardware and Production Challenges

In VR cinematography, our camera system has limitations stemming from the Camera's sensor and processing characteristics and the Lens's optical design. The most significant production challenges included the Camera's suboptimal noise performance and limited dynamic range, the Lens's poor optical resolving power, and the Lens's extreme AOV, which introduced substantial framing and lighting constraints.

Camera System

We selected the Camera and Lens based on the following considerations:

- Native RF mount compatibility with full lens-to-body communication.
- Full-frame image format and 8192 × 4320 (8K) sensor resolution, providing sufficient image detail for further AI resolution upscaling and denoising.
- High frame rate (59.94p) support at full resolution and image format, ensuring smooth motion continuity for HMD playback.
- 190° AOV, producing a high degree of VR immersion when viewing footage in HMDs.

Sensor and Recording Limitations

The Camera's 8K, 36 × 24 mm full-frame sensor provides an acquisition resolution sufficient for subsequent AI-based upscaling and denoising processes. When paired with the dual fisheye Lens, the single 8K sensor is divided into two circular fisheye images corresponding to left- and right-eye views, resulting in an effective image diameter of approximately 3684 pixels per eye.[5] In comparison, the Apple Vision Pro (AVP) renders at 3660 × 3220 pixels per eye[6] with a 100° AOV,[7] substantially narrower than the 190° capture AOV of our imaging system. Because the AVP displays only a 100° subset of the original 190° capture at any given moment, the effective resolution per displayed pixel is reduced relative to our full-frame acquisition. This spatial "cropping" effect necessitates further post-production enhancement, including AI-driven upscaling, to ensure that the displayed imagery fully exploits the native per-eye resolution of the AVP and maintains perceptual fidelity within the HMD.

Following stereoscopic reprojection, each 180° equirectangular eye view measured 4096 × 4096 pixels, corresponding to approximately 22.8 pixels per degree at the equator. The subsequent 2× enhancement and upscale in Topaz Video AI produced an 8192 × 8192 image for each eye, or a combined stereo raster of 16384 × 8192 pixels, increasing equatorial sampling density to approximately 45.5 pixels per degree. This increase in sampling density provides a denser reconstructed raster for subsequent compression and high-resolution headset presentation.

High recording frame rates are essential for preserving motion continuity in VR, thereby reducing user discomfort and maintaining perceptual immersion. The AVP supports playback at frame rate multiples of 24p and 30p up to 100p, enabling judder-free rendering when source material is properly matched. Conventional cinematic frame rates such as 24p or 25p, while aesthetically established in traditional filmmaking, have been demonstrated to be inadequate in VR contexts, where sub-60p acquisition may introduce perceptible judder and simulator sickness (SS).[2] For this reason, principal photography was conducted at the Camera's maximum supported frame rate of 59.94p (60p) at full sensor resolution, recorded in Canon 12-bit Cinema RAW Light.

Using the Camera's full 8K 60p recording capabilities requires Canon's 12-bit Cinema RAW Light format, which omits in-camera noise reduction. Consequently, RAW footage exhibited elevated sensor noise, particularly within shadow regions.[4] These limitations were compounded by the project's low-light, gritty apartment setting. In an effort to mitigate these issues, Canon Log 3 encoding was selected for its lower noise floor, improved shadow handling, and enhanced color retention relative to other log profiles.[8] Nevertheless, even at the Camera's native ISO settings of 800 and 3200, the resulting footage displayed significant noise in dark regions and susceptibility to highlight clipping, as illustrated in the left panel of **Fig. 2.**

1545-0279/26©2026SMPTE



FIGURE 2. Side-by-side comparison of unprocessed RAW footage (left) and AI-denoised, upscaled footage (right) (2025).

Optical Characteristics and Limitations of the Lens

The Lens's 190° AOV and 60 mm interaxial distance produce a stereoscopic pair whose capture baseline approximates common human interpupillary distances. In this paper, we use orthostereoscopic imaging in its established stereoscopic-display sense: capture and presentation geometry intended to preserve naturalistic scale, perspective, and binocular depth relationships.[9–11] The term has a longstanding lineage in stereoscopic cinema, interactive display, telepresence, and virtual-reality research. A human-scale interaxial distance alone does not guarantee perfect orthostereoscopy, however, as perceived stereo geometry also depends on lens projection, reprojection, display field of view, headset optics, and viewer IPD.

Optical deficiencies in the Lens further degraded image quality. Because we needed to maximize light transmission to combat sensor noise, our initial approach prioritized shooting at the Lens's widest aperture (f/2.8). However, testing revealed significant degradation in resolving power at this aperture, leading to unacceptable reductions in image sharpness.

Image Acquisition Optimizations

To address these challenges systematically, we structured our image acquisition pipeline to improve three critical areas of capture: exposure, optics, and lighting.

Exposure and Noise Mitigation Strategies

To compensate for the Camera's noise performance and limited dynamic range, we established a recording configuration optimized for color retention and noise reduction:

- Recording Format: Canon 12-bit Cinema RAW Light
- Resolution: 8192 × 4320
- ISO: 3200 (base 3200)
- Color Profile: Canon Log 3
- Frame Rate: 59.94p

Furthermore, we implemented a one-stop overexposure strategy[8] to retain shadow detail, reduce perceptible noise, and protect highlights. This approach involved systematically exposing 1 stop above our target exposure while monitoring footage through a custom pull-down lookup table (LUT) designed to simulate the correct exposure and intended final look in real time.

We standardized all shots at f/5.6, exposing for f/8 via a 1-stop overexposure method. This approach enabled the sensor to capture increased data in shadow regions while protecting highlights, effectively reducing noise accumulation in underexposed areas.

Optimizing Optical Performance of the Lens

Careful aperture selection was a critical factor in balancing optical resolving power against light transmission. Empirical testing demonstrated the following:

- f/2.8 exhibited significant optical softness, making it unsuitable despite its increased light transmission.
- f/7.1 yielded the highest resolving power, albeit still exhibiting suboptimal sharpness, but at the cost of reduced light transmission.
- f/5.6 provided an optimal compromise, offering an acceptable level of sharpness while securing an additional 2/3 stop of light transmission compared to f/7.1.

Lighting and Framing Adaptations for Ultra-Wide AOV

The ultra-wide AOV of the Lens meant traditional lighting setups had to be moved out of frame or hidden. This increased the distance between fixtures and the illuminated area and/or required smaller, more discrete fixtures. This reduced the total light reaching the sensor from the scene,



FIGURE 3. Frame grab illustrating the set constructed without a ceiling to facilitate overhead lighting (2025).



FIGURE 4. Lighting diagram illustrating a set layout with lighting fixture, diffusion, and bounce placements (2025).

thereby increasing sensor noise. To circumvent these limitations, we employed a hybrid practical-lighting and computer-generated production-design strategy.

Set Design for Optimal Fixture Placement

Our apartment set was designed with doorways and structural corners, allowing conventional off-axis lighting setups to create strong accents and hidden hair lights without compromising framing, as seen in **Fig. 3.** This layered lighting approach allowed us to preserve the scene's visual integrity while using luminance contrast to guide viewer attention to the areas of the 180° frame that matter most to the narrative. Lighting subtly directs the gaze without compromising the viewer's freedom of exploration. We balanced interactive storytelling with cinematic presentation, motivating the viewer to follow the narrative without overt guidance.

Virtual Set Extensions to Motivate Overhead Lighting

To prepare for virtual ceiling set extensions, we built our set on a soundstage without a ceiling, as shown in **Fig. 3.** This allowed us to rig fixtures on the soundstage's overhead lighting grid to provide ambient fill and hard spots into the set. We composited a virtual ceiling into the set in post-production, maintaining the illusion of an enclosed environment. Our

1545-0279/26©2026SMPTE

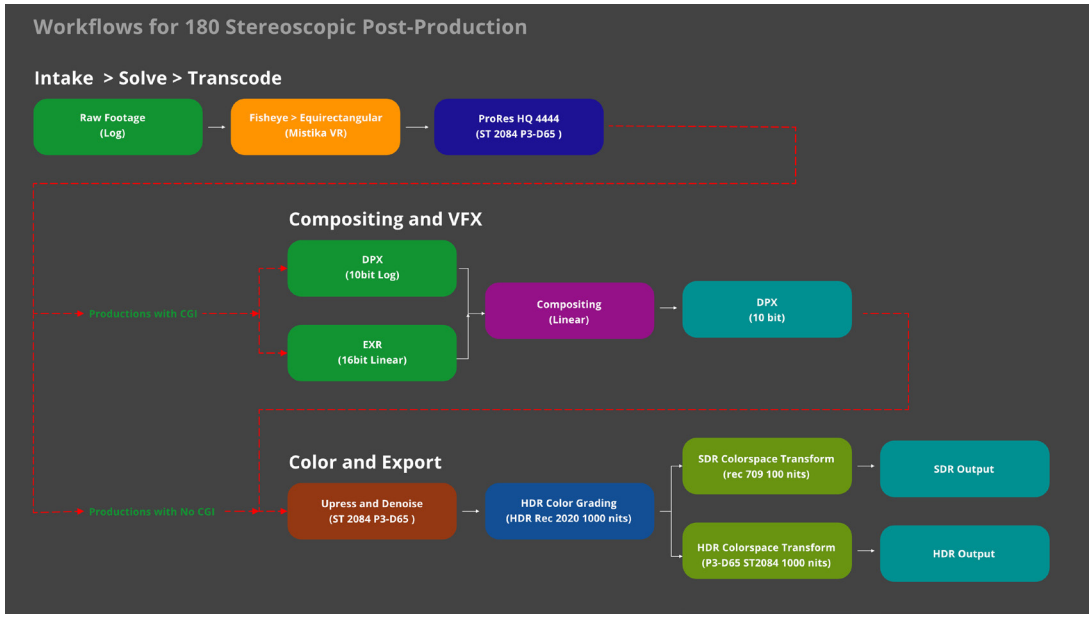


FIGURE 5. Our digital intermediate pipeline flow chart detailing workflows for both CGI and non-CGI projects (2025).

overhead lights were motivated by the house lights on the virtual ceiling, ensuring spatial consistency between practical and CG elements. This ceiling-less set design was our primary method for achieving our target exposure levels while hiding all large lighting fixtures.

Embedded Practical LED Fixtures

Discrete, high-output, high-color-quality LED fixtures were hidden within practical set decoration pieces such as lamps and magnetically mounted to the walls out of frame to provide subtle fill and hair lighting (refer to "B7C" and "MC Pro" in **Fig. 4**), further enhancing depth and separation between subjects and the background.

Post-Production Workflow

While our acquisition pipeline mitigates many imaging challenges, its most important role is to prepare our footage for post-production. RAW footage directly from the camera still exhibits noticeable noise and optical resolution deficiencies. The final refinement of our footage into a polished VR film is ultimately achieved through our post-production pipeline.

Digital Intermediate Pipeline

We imported Canon Cinema RAW Light footage into Mistika VR and decoded it using Canon Log 3 and Canon Cinema Gamut before stereoscopic reprojection. Following correction and reprojection, the imagery was exported using Rec. 2020 primaries with an ST 2084 (PQ) transfer function. This Rec. 2020/PQ interchange configuration was selected based on Canon workflow guidance to preserve the wide-gamut and HDR image information carried from the RAW source through subsequent post-production stages. P3 therefore describes a downstream mastering and delivery target rather than the native acquisition or intermediate color space.

For footage requiring no CGI integration, the Rec. 2020/PQ material was imported into DaVinci Resolve and transformed into DaVinci Wide Gamut with DaVinci Intermediate for color grading. Footage requiring CGI integration instead entered the ACES-managed compositing workflow described in subsequent sections before returning to the same grading and mastering pipeline. Final imagery was transformed to P3-D65 primaries with an ST 2084 (PQ) transfer function for Apple Vision Pro delivery.

When this production was completed, the required MV-HEVC export workflow was not yet available directly in DaVinci Resolve. We therefore converted the completed P3-D65/PQ stereo master externally using FFmpeg for MV-HEVC delivery. **Figure 5** illustrates both the CGI and non-CGI color-management workflows.

Strengths and Weaknesses of the Camera System for Our Post-Production Pipeline

Canon provides proprietary software for converting dual-fisheye imagery into stereoscopic 180° equirectangular format. However, in our testing, the software did not resolve imagery to an acceptable standard for research or production use. As a result, our workflow employed Mistika VR for this stage of the pipeline. Through camera-lens profiling conducted within Mistika VR, we identified two strengths and five primary weaknesses of our camera system:

- Camera system strengths:
  - ◊ Ability to shoot RAW at 60 frames per second.
  - ◊ A total capture resolution of 8192 × 4320.
- Camera system weaknesses:
  - ◊ Off-centered fisheye images projected from the lens onto the sensor, as seen in **Fig. 6**.
  - ◊ Low optical resolving power of the lens.
  - ◊ Unknown exact IPD.
  - ◊ Suboptimal noise performance, especially in low-light scenarios.

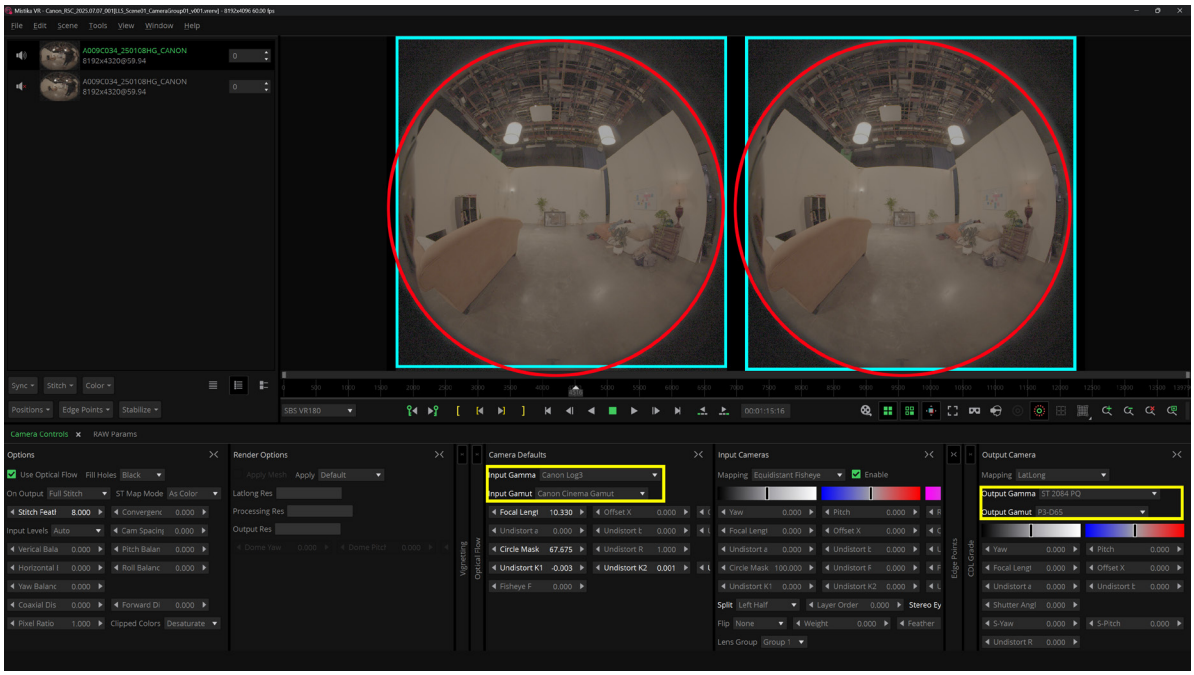


FIGURE 6. A diagram illustrating off-center fisheye images resulting from the camera and lens (2025).

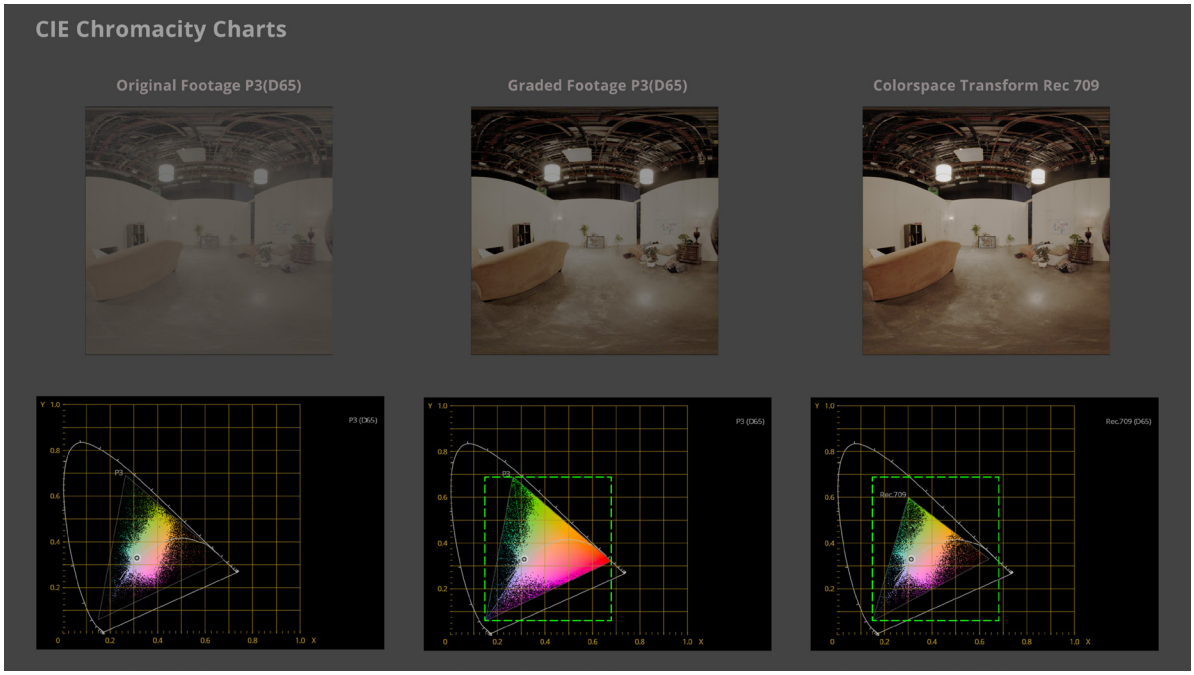


FIGURE 7. Comparison between ungraded footage, graded footage within the P3 color space, and a color space transform to Rec. 709 Gamma 2.4 (2025).

◊ Focus inconsistency between the two fisheye projections, resulting from each lens group needing to be independently focused before each shot.

Camera System Testing Methodology

Because of the Lens's extreme AOV and distortion characteristics, a standard lens grid calibration shoot was insufficient. Using the software 3D Equalizer, we determined that a more rigorous approach involving a detailed series of tests supplemented with high-precision photogrammetry was required.

Photogrammetry

Conventional grid-based calibration proved inadequate for the extreme distortion characteristics of the fisheye optics. To establish a more rigorous spatial reference, we combined 3D Equalizer lens and camera analysis with a high-precision photogrammetric reconstruction of the practical set. We incorporated custom marker sets into the RealityCapture reconstruction, while a Leica Disto site survey provided independently measured physical dimensions to constrain global scale and orientation. Integrating these surveyed measure-

1545-0279/26©2026SMPTE

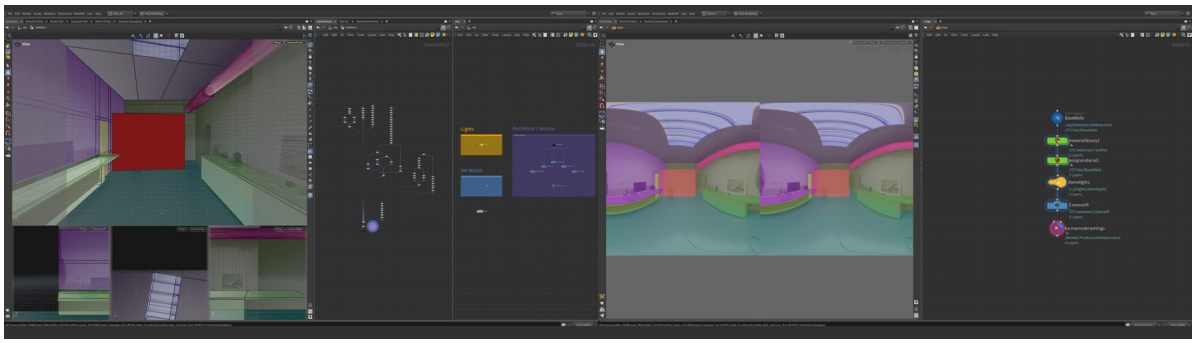


FIGURE 8. 3D viewport used for alignment and the stereo renders over the backplate (2025).

ments prevented the photogrammetric reconstruction from relying on image-derived scale alone.

The resulting photogrammetric geometry served as a spatial reference throughout camera validation and CGI integration. Initial base geometry derived from on-set measurements was aligned to the reconstructed environment and subsequently rendered against the stereo 180° equirectangular backplates. This provided visual feedback for evaluating correspondence between the known physical geometry, the photogrammetric reconstruction, and the stereoscopic plate. Once the scan had been oriented to the base geometry, no subsequent scale adjustment was applied. The calibrated environment was then carried forward into Houdini, where the modular set geometry and custom alignment tools described later in the Section, *A Larger Living Space—Pipeline Integration Example*, provided an additional means of validating geometric correspondence.

Footage Assets

To create a full lens profile for the Lens, we determined that we would shoot a series of tests to incorporate into 3D Equalizer. This would include a typical center grid lens shot with a static camera, a series of nodal camera pans for matchmoving, and a gimbaled handheld tracking shot. All of these would keep the main lens grid in full view.

Noise Testing

To evaluate noise performance, we tested multiple exposure strategies at both base ISO 800 and base ISO 3200. For each setting, exposures ranged from one stop under to two stops over. We ultimately observed the best noise performance at one- and two-stop overexposure.

Focus Testing

We shot a series of tests with incremental focus ranges starting at 9 in. and increasing a foot at a time to determine the best focal point for deep focus.

Footage Ingest, Transcode, Edit, and Output

We ingested the captured RAW footage through Mistika and split it into left- and right-eye streams. Using Mistika's lens guides, the off-center fisheyes were corrected, and an initial geometric solve was calculated. Following stereoscopic reprojection, the corrected imagery was exported using Rec. 2020 primaries with an ST 2084 (PQ) transfer function and subsequently imported into DaVinci Resolve. Within Resolve, the footage was transformed into DaVinci Wide Gamut and DaVinci Intermediate for grading before final transformation to P3-D65 primaries with an ST 2084 (PQ) transfer function for Apple Vision Pro delivery. A comparison between the graded P3 output and a Rec. 709 Gamma 2.4 transform is shown in **Fig. 7.**

CGI Integration Testing

Plates were processed for compositing as 10-bit DPX files. For scene alignment, we mapped the equirectangular plate to an inverted hemisphere oriented around the Z-Axis with five cameras locked at the zero position. We created five 90º rectilinear cameras oriented +X, -X, +Y, -Y, and +Z, with a square filmback (24mm x 24mm).

Camera Validation—Testing Geometry Alignment

We validated scene alignment using baseline measurements of camera height and on-set location data. We then aligned the initial geometry with the photogrammetry model.

To refine alignment, we rendered the base 3D geometry against the stereo backplates, as shown in **Fig. 8**. Comparing stereo 180° equirectangular renders provided feedback for iterative geometry refinement.

Once we refined the 3D base model, we imported and oriented the 3D photogrammetry scan, ensuring no scale adjustments, as shown in **Fig. 9.**

AI-Assisted Denoising and Upscaling

Following compositing, the completed stereo imagery was exported as ProRes 4444 XQ and processed in Topaz Video AI as two discrete restoration passes. A dedicated denoising pass using the Nyx model was performed first, followed by a separate enhancement and 2× upscaling pass using Proteus. The latter increased each 180° equirectangular eye view from 4096 × 4096 to 8192 × 8192 pixels, producing a final side-by-side stereo raster of 16384 × 8192 pixels. At the equator of the equirectangular projection, this increased sampling density from approximately 22.8 to 45.5 pixels per degree while preserving the original 59.94p temporal cadence.

Proteus was operated in manual parameter mode rather than Auto, allowing the Recover Detail, Sharpen, Dehalo, and Anti-Alias controls to be explicitly configured and held con-

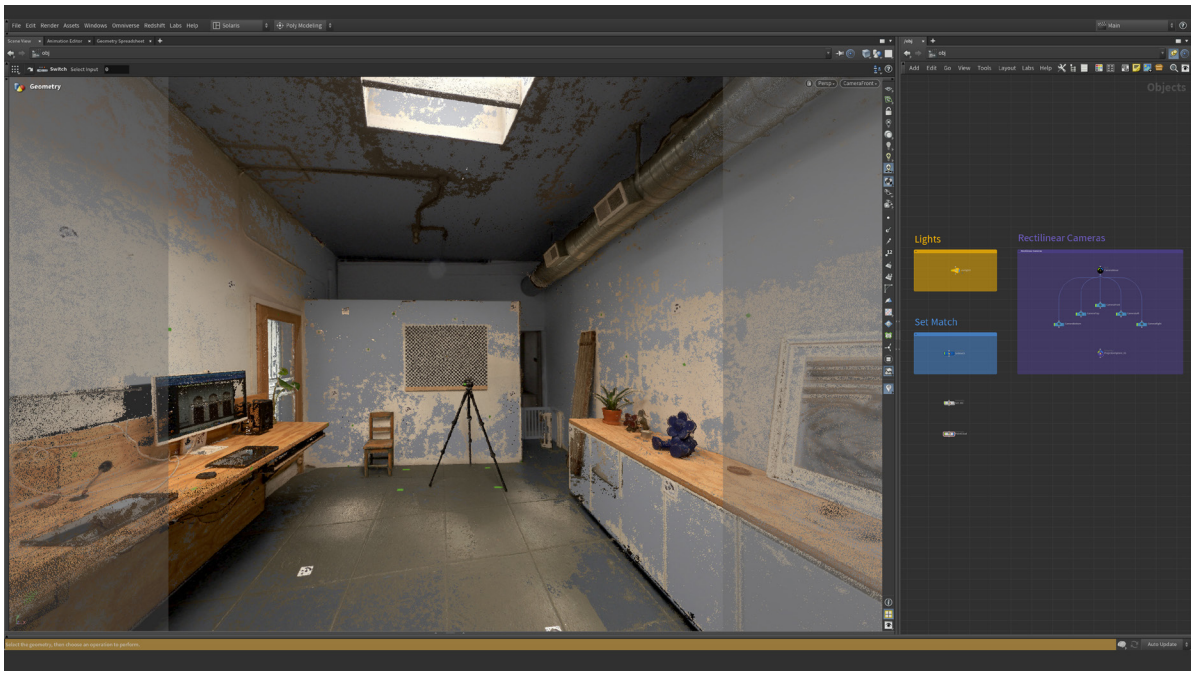


FIGURE 9. 3D viewport showing the aligned photogrammetry with the backplate (2025).

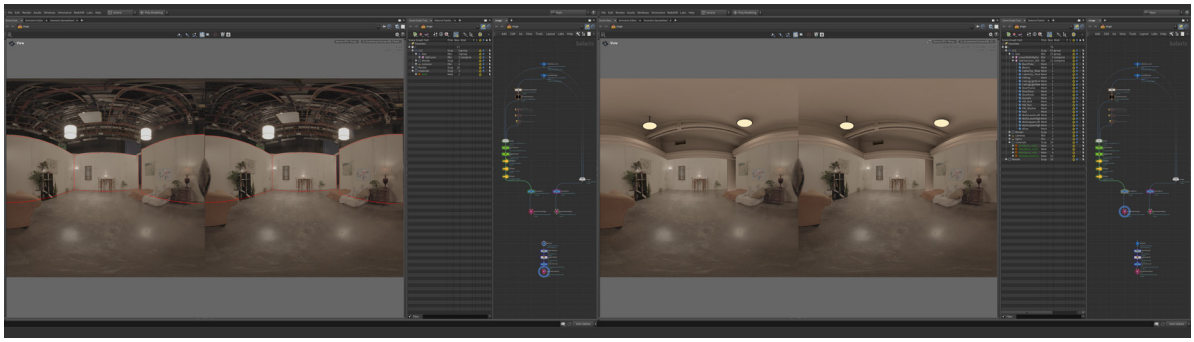


FIGURE 10. Renders showing the alignment tool and the render over the backplate (2025).

stant across the production. Auto mode's per-shot parameter estimation was rejected because shot-to-shot variation could produce a non-uniform rendition of detail and edge structure across cuts and, more critically for stereoscopic imagery, could result in divergent processing between corresponding left- and right-eye views. Maintaining a consistent parameter configuration across both eyes and throughout the production reduced the likelihood that the enhancement itself would introduce binocularly inconsistent high-frequency detail.

The ordering of the restoration stages was likewise deliberate. Denoising was performed before enhancement and upscaling because enlarging imagery before noise removal can promote stochastic sensor noise into more coherent spatial structures. A subsequent restoration or enhancement stage may then interpret and preserve those structures as image content rather than remove them as noise. Performing the Nyx denoising pass first therefore provided Proteus with a cleaner source image from which to reconstruct edges and fine textures.

Traditional mathematical denoising techniques, such as Total variation regularization and wavelet-based noise suppression, can reduce image noise at the expense of edge definition and high-frequency texture.[3] This trade-off becomes particularly visible during close HMD viewing. Neural video-denoising approaches can instead use spatial and temporal image structure when differentiating persistent scene detail from stochastic noise.[12] For this production, the objective of AI-assisted restoration was therefore not maximum noise removal, but sufficient suppression of sensor noise while retaining perceptually important fine structure for stereoscopic headset presentation.

The resulting 16384 × 8192 image should be understood as a higher-density reconstruction rather than as a native 16K optical capture. The resolving power of the original Camera and Lens remains the upper bound on captured scene information. The 2× upscale instead provides a denser sampling grid for subsequent compression and headset presentation without implying a corresponding doubling of captured optical resolution.

Apple Vision Pro Delivery and MV-HEVC Encoding

Following restoration, upscaling, and final color mastering, the completed stereoscopic master was encoded for Apple

1545-0279/26©2026SMPTE



FIGURE 11. Nuke compositing network (2025).

Vision Pro playback using Multiview High Efficiency Video Coding (MV-HEVC). Unlike conventional side-by-side stereo, in which both eye views occupy different portions of a single video frame, MV-HEVC represents the left- and right-eye images as separate layers within a single video track. The layers share a common frame size, allowing each eye to remain discrete during decoding while the codec exploits the substantial similarity between the two views to reduce redundant image information.[13]

When this production was completed, direct MV-HEVC export was not yet available in DaVinci Resolve. The final P3-D65/ST 2084 PQ stereo master was therefore converted externally using FFmpeg. The left- and right-eye imagery was encoded into the corresponding multiview layers, and the required spatial metadata was embedded in the resulting delivery file for correct stereoscopic interpretation on Apple Vision Pro. The completed master was subsequently reviewed natively in the headset to verify correct eye assignment, projection geometry, motion cadence, color reproduction, compression behavior, and overall binocular presentation.

A Larger Living Space—Pipeline Integration Example

To ensure our physical environment could be reproduced digitally, we built our practical apartment set using a modular flat system. All walls were fabricated from standardized 4 × 10 foot panels, which were subsequently replicated in a three-dimensional environment. These virtual modules served as a registration framework, aligned precisely with a base geometric model derived from on-set measurements. A custom Houdini Digital Asset (HDA) was developed to enforce consistent alignment across shots. The HDA also enabled the rendering of polylines over stereo equirectangular backplates, providing a mechanism to validate geometric correspondence and inter-pupillary distance (IPD) accuracy (alignment tool shown in **Fig. 10**).

The base geometry was used as a structural guide for set extensions. These extensions were first rendered as a quality-control pass against the original backplate, then used as matte cut-outs to support compositing operations. An occlusion pass was generated in parallel to improve the perceptual blending of digital extensions with the live-action plate. The set extensions were rendered within a Universal Scene Description (USD) pipeline, using Solaris and the Karma renderer with a MaterialX shading framework. The same USD-based workflow also facilitated the production of supplementary computer-generated imagery (CGI) in Unreal Engine, ensuring interoperability across platforms.

All 3D outputs were generated as multilayer, linear EXR files, incorporating the following render passes: beauty, combined albedo, combined reflection, combined refraction, emission, position, normals, UVs, and a Fresnel component. These passes provided the necessary information to support advanced compositing operations. To preserve the raw fidelity of the renders, no denoising was applied during the rendering process.

Compositing was completed in Nuke with a procedural network, shown in **Fig. 11**, that allows for work to be easily propagated across multiple shots. The 3D renders are grain- and grade-matched against the uncorrected footage, and grain is matched again in the compositing phase to ensure as close a match as possible before the denoise and uprez phase.

Although many components of this workflow can be generalized to other stereoscopic HMDs, several parameter choices were driven specifically by the Apple Vision Pro presentation. These include high-resolution reconstruction to compensate for the disparity between capture and display sampling den-

sity, 59.94 fps acquisition to maximize temporal resolution within the Camera's full-resolution recording limitations, HDR grading with a P3-D65/ST 2084 PQ delivery target, MV-HEVC stereoscopic encoding, and final quality-control evaluation within the headset itself. The resulting workflow should therefore be understood not simply as a general stereoscopic post-production pipeline, but as an accessible capture-to-display methodology whose image-processing and delivery decisions were evaluated against the perceptual demands and media architecture of Apple Vision Pro.

Future Work

Significant work remains in advancing techniques and technologies for immersive cinematography. Future research can investigate imaging and processing approaches that extend dynamic range through higher bit-depth acquisition (e.g., 16-bit RAW workflows). Such methods would improve the ability to capture low-light scenes without relying on deliberate overexposure followed by post-production enhancements. As stereoscopic camera systems continue to evolve, further research into higher frame rate capture (≥120p) will be important for minimizing motion discontinuity and reducing simulator sickness. In addition, integrating motion-tracked VFX set extensions with externally recorded camera telemetry offers a pathway to more accurate alignment of computer-generated elements with physical lighting fixtures and camera movement.

Conclusion

The insights gained from *A Larger Living Space* informed the development of a capture-to-delivery workflow optimized around the image-quality and playback requirements of Apple Vision Pro. Although many of its underlying methods can be generalized to other stereoscopic cameras and HMDs, specific acquisition, reconstruction, color-management, and encoding decisions were selected and evaluated against the characteristics of the Apple Vision Pro display and visionOS media pipeline. By combining established filmmaking practices with emerging VR and AI tools, we demonstrate a reproducible approach to high-fidelity 180° stereoscopic cinematography using comparatively accessible production hardware. Our acquisition approach (integrating deliberate overexposure, lensing trade-offs, and adapted lighting strategies) produced footage optimized for subsequent AI-driven denoising and 16K upscaling, allowing us to retain fine detail and minimize noise artifacts. Recording at higher frame rates further enhanced motion realism and viewer comfort. Overall, our findings suggest that immersive VR films can maintain the artistic intent and visual language of traditional cinema while meeting the technical requirements of next-generation HMDs, providing a practical foundation for both filmmakers and researchers advancing cinematic storytelling in VR.

References

1. B. Allan, "Blackmagic URSA Cine Immersive Camera—Measuring Resolution," *Newsshooter*, 2025. [Online]. Available: www.newsshooter.com.
2. J. Wang et al., "Effect of Frame Rate on User Experience, Performance, and Simulator Sickness in Virtual Reality," *IEEE Trans. Vis. Comput. Graph.*, 29(5):2478–2488, 2023. doi:10.1109/TVCG.2023.3247057.
3. Q. Yang et al., "Low Light Image Denoising Based on Poisson Noise Model and Weighted TV Regularization," *Proc. IEEE Int. Conf. Image Process.*, 3199–3203, 2018. doi:10.1109/ICIP.2018.8451840.
4. J. Behiri, "Canon EOS R5 C Lab Test—Rolling Shutter, Dynamic Range and Latitude," *CineD,* 2023. [Online]. Available: www.cined.com.
5. Canon Inc., "Canon RF 5.2mm F2.8 L Dual Fisheye," Canon Canada, 2025. [Online]. Available: www.canon.ca.
6. D. Heaney, "Apple Vision Pro Extended Teardown Reveals Its True Active Resolution," *UploadVR*, 2024. [Online]. Available: www.uploadvr.com.
7. Nivsus, "Apple Vision Pro FOV: Limitations and the Quest for Enhancement," *vskel*, 2024. [Online]. Available: vskel.com.
8. Canon Europe, "Enhance Filmmaking with Canon Log," Canon Europe, 2023. [Online]. Available: www.canon-europe.com.
9. M. T. Bolas and S. S. Fisher, "Head-Coupled Remote Stereoscopic Camera System for Telepresence Applications," *Proc. SPIE*, 1256:113–123, 1990. doi:10.1117/12.19896.
10. L. Lipton, *Foundations of the Stereoscopic Cinema: A Study in Depth*, Van Nostrand Reinhold: New York, 1982.
11. W. Robinett and J. P. Rolland, "A Computational Model for the Stereoscopic Optics of a Head-Mounted Display," *Presence*, 1(1):45–62, 1992. doi:10.1162/pres.1992.1.1.45.
12. M. Tassano, J. Delon, and T. Veit, "DVDNET: A Fast Network for Deep Video Denoising," *Proc. IEEE Int. Conf. Image Process.*, 1805–1809, 2019. doi:10.1109/ICIP.2019.8803136.
13. Apple Inc., "Converting Side-by-Side 3D Video to Multiview HEVC and Spatial Video," *Apple Developer Documentation.* [Online]. Available: developer.apple.com.

About the Authors



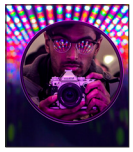
Stanley Lin is a creative technology researcher at USC Ganek Immersive Studio and Visiting Scholar at USC MEML.



Nathan Fairchild is an immersive filmmaker, currently working as assistant director of USC School of Cinematic Arts' Ganek Immersive Studio.



Spencer Lin is an AI & XR researcher focused on augmenting storytelling through AI characters in immersive mediums.



Akaash Tumuluri is a writer, director, and producer, currently serving as lead producer for the Ganek Immersive Studio at USC.



Jordan Reece Halsey is the director of USC's Ganek Immersive Studio and a media artist specializing in immersive technologies and simulation.



Scott S. Fisher is a professor at USC School of Cinematic Arts and Director of USC's Mobile & Environmental Media Lab.

Presented at the SMPTE 2025 Media Technology Summit, Pasadena, CA, 16-19 Oct. 2025.

Copyright © 2026 by SMPTE.

DOI: 10.5594/JMI.2026/YCNL3434
Date of publication: 1 October 2026

KEYWORDS VIRTUAL PRODUCTION // RENDERING // SPECTRAL RENDERING // FILMMAKING // SPECTRAL REDUCTION TECHNIQUES



FIGURE 1. Example of a VP shoot where there are real elements blended with virtual elements.

Spectral Rendering Techniques to Improve Color Matching in Virtual Production

By Vlad Simion, David Long, Flip Phillips, and Kris Murray

Abstract

As rendering engines become increasingly important in film and television, with their use in virtual production (VP), some underlying issues become more apparent. This paper aims to investigate how we can improve asset color matching of VP elements with real-life objects found on sets. Experiments were conducted in which objects were exposed to various types of lighting setups, and digital twins were rendered using both RGB methods and spectral methods, with data reduction techniques also employed. The renderings were then filmed, alongside their real-life counterparts. Color difference metrics were used to determine whether spectral rendering and data reduction techniques offered advantages over RGB renderings. The conclusion illustrates that spectral rendering offers advantages, including higher accuracy in rendering the colors of materials.

For many years, digitally rendered images have been utilized in a wide range of applications, including visual effects and virtual production. Virtual production (VP) is an emerging technology used in filmmaking where sets utilize digital backgrounds as a replacement for traditional green screen techniques.[1] Typically, projectors or LED panels are positioned behind the action and used to display virtual scenes. This allows actors to see and react to events in the scene and enables accurate lighting reflections. This also makes it possible to build smaller environments and extend them digitally. For example, creatives can extend bridges, walkways, or even entire worlds **(Fig. 1)**. However, extending real-world assets into the virtual world introduces challenges in colorimetric consistency. Since virtual and real assets behave differently, it is often insufficient to simply match digital lighting parameters (e.g., color temperature, gains), requiring time and effort to alter them manually.

This paper aims to address these challenges by investi-

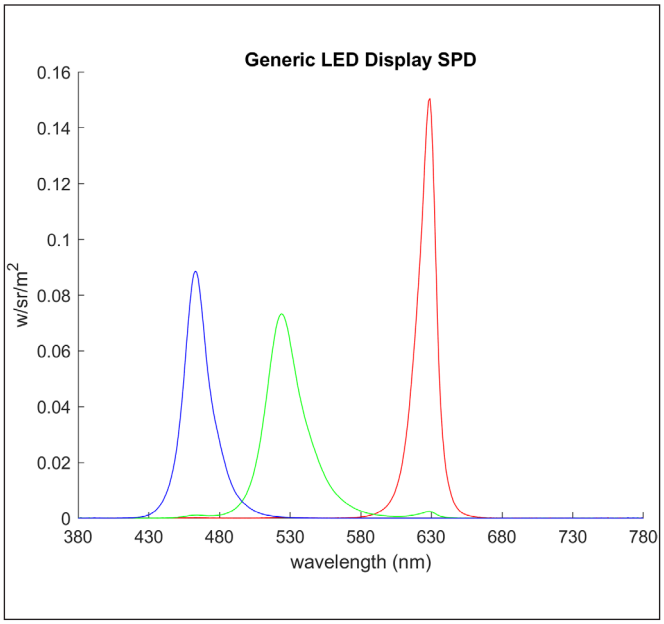


FIGURE 2. Generic LED display power distribution.

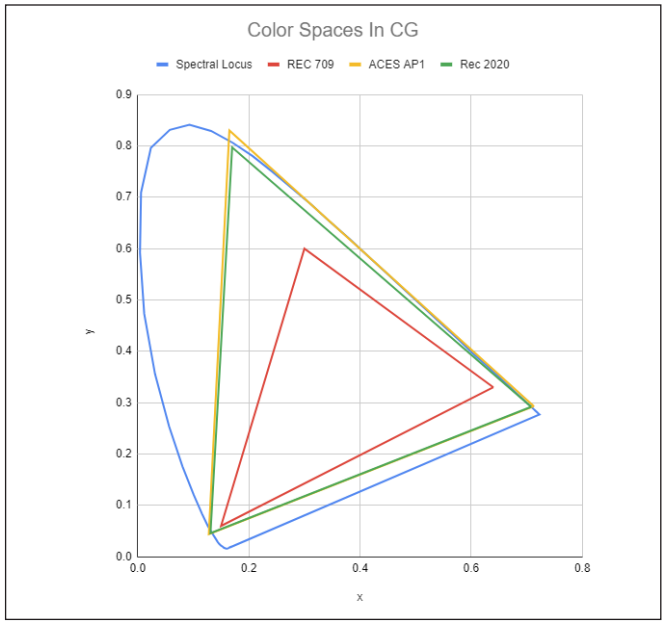


FIGURE 3. Color spaces used in computer graphics.

gating different color rendering methods. We will test how spectral rendering (full and abridged) compares to traditional RGB rendering within a practical environment that simulates virtual production workflows.

Most popular rendering engines used in the entertainment industry (Unreal, Unity, Blender) predominantly employ RGB rendering, where scalars are defined as Red, Green, and Blue encoding channels. Only engines that are used in more accurate or science-driven applications tend to be more spectrally accurate (e.g., Mitsuba or Manuka).[2,3] However, even as early as the late 1980s,[4-6] spectral rendering was being developed, but it has still not been widely adopted among popular rendering engines. This is because RGB is often sufficient to achieve results that artists like. Most modern rendering engines trace rays by computing their intersections with lights and objects in the scene and calculating RGB values that result from refraction, reflection, scattering, and transmission.[2] For example, rays are fired from the observer location, reflect off an object, and then hit a light source. A ray-object intersection is a "hit," and a "shading" operation is performed to determine the color produced at that intersection. Thus, a ray has an associated RGB triplet at each step throughout the rendering pipeline, which ultimately determines the color of the pixel that represents that ray. If the objects are measured, then these RGB values are also computed from an observer's perspective under a specific lighting condition, which makes them accurate only in certain environments. However, if the lighting in the rendering engine differs from the real-life measured lighting, or if the observer is swapped, then the RGB values should be recomputed from spectral data to ensure accuracy.

Real-world light is composed of photons within the visible spectrum, ranging from 380nm to 780nm. To quantify color vision, the CIE has adopted the XYZ color space. XYZ values contain information on how much light was detected across multiple wavelengths, thereby reducing the signal's dimensionality. They are obtained by integrating light seen by the eye using color-matching functions,[7] derived from the standard human observer as three integration curves. Not knowing the original spectra makes simulation challenging, as potential interactions between light and objects can be missed.

Because the eye has three types of retinal cells that integrate photons within a specific range, rather than determining their exact wavelength, stimuli with very different spectral compositions can integrate to identical XYZ values and be perceived identically by a common observer[7] (i.e., metamerism). A common type of metamerism failure, known as illuminant metamerism, occurs when two surfaces match under the same illuminant but no longer match when the illuminant is changed. Another example is observer metamerism, which arises from differences between observers. There are almost infinite ways of reproducing the same color, as spectra can be infinitely mixed to create identical XYZ values. Displays take advantage of this, as devices with different spectra can recreate the same colors.

Cameras, however, have different sensitivities to light than the human eye because the filters in front of their pixels block different wavelengths than the human visual system, at varying ratios. This discrepancy stems from historically available materials and design practices. There is also a different amount of crosstalk between the channels of a camera and the crosstalk of the eye's sensitivities. Paired with the absence of higher-order visual processing in the brain, cameras perceive the world differently than humans, leading to metamerism failures.

Most modern displays are LED-based and typically feature three diodes with different power distributions **(Fig. 2)**. The display uses a combination of gains for each LED to create a range of colors. Because a camera has different sensitivities than the human eye, it perceives LED displays very differently. Narrowband emitters are especially troublesome, as

1545-0279/26©2026SMPTE

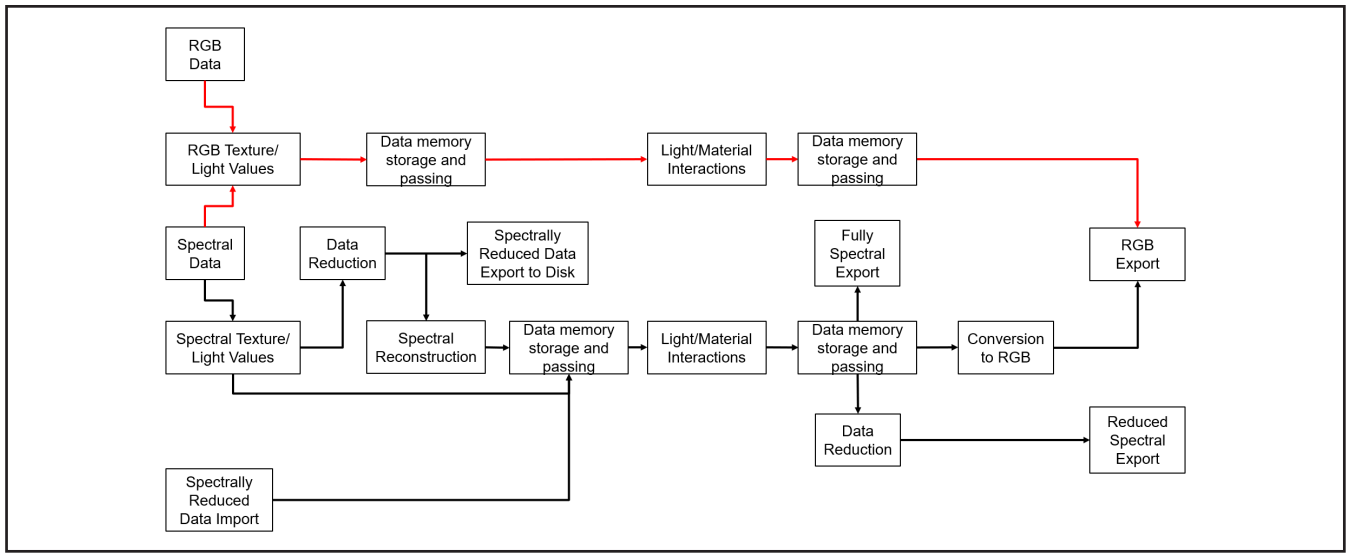


FIGURE 4. Rendering pipelines when using spectral rendering (black) versus pipelines that use traditional RGB rendering (red). The reduced spectral data can also be saved to storage for later recall, thereby skipping the analysis step and improving performance. The scene can also be exported as full-spectrum, reduced data, or converted to RGB.

even a small shift in the peak can cause large changes in the integrated color. This is a problem in current virtual production environments, where cameras are constantly pointed toward LED displays and often perceive both the real and virtual world in a way that differs significantly from the intended one. Simulating the real world using RGB values derived from XYZ values can further exacerbate this problem. This mismatch error is especially troublesome as it can break the illusion that creatives try to implement, as it makes the point where the two worlds collide visible. Currently, creatives attempt to conceal this intersection point using props (such as placing a plant that masks the blend point) or spend considerable time refining the virtual world to make the transition seamless and cohesive.

Background and Theory

Most popular rendering engines employ RGB rendering. We can convert XYZ stimulus values to estimate RGB values using a 3 x 3 matrix.[8] However, this conversion introduces a fundamental limitation: RGB values can represent only colors within the triangular region formed by those three primaries in the chromaticity diagram. This is because RGB color spaces are based on additive color mixing, where any representable color must be expressible as a weighted sum of the three primaries (see **Fig. 3)**. If any of the interactions of object and light would occur somewhere outside of the RGB space chosen, that interaction could not be represented.

We recognize that spectral information is crucial, and we have thus far observed how reducing it to RGB may be excessive. One goal of this research is to represent spectra using a limited number of terms, reducing both data and computational time compared to full spectral rendering while still offering benefits over traditional RGB techniques. The visible spectrum ranges from 380 to 780 nm, and most devices can measure spectra with precision every few nanometers; however, the data is reported as 400 data points. This may be taxing for each ray, but it can be condensed into fewer terms that describe the characteristics. The problem is twofold: information storage and interaction. The storage phase involves storing data about each asset's color and each ray. In contrast, the interaction phase is represented by computations that happen when a ray intersects an object. In the presented approach, light transport and storage could still occur by associating only a few values with a ray. In **Fig. 4**, we can see an example of the differences between spectral and traditional rendering pipelines. In traditional pipelines, we can only store and operate on RGB data. In spectral pipelines, several options are available, including using spectral data throughout the pipeline or applying data-reduction techniques to optimize the algorithm's memory footprint. To properly utilize the techniques, the full spectral data textures must be analyzed (or recalled from previous analyses). Ultimately, the engine can export either a fully spectral image, an image composed of reduced terms, or an RGB image converted from the spectral data.

As mentioned, carrying full spectral information is computationally expensive. However, most reflectance spectra can be represented with only five principal components with high accuracy.[7] Principal component analysis (PCA) is a linear dimensionality reduction technique that aims to preserve important information and represent it as a set of new variables, known as components. Analyzing the principal components of a data set provides a new perspective, as we are more interested in variance between samples, given that they share common traits. This approach may represent a spectrum of light (or an object's material properties) using only a few terms.[4] A simpler approach[9] could separate the spectrum into "bins" (equivalent to increasing the spectral bandpass) that contain information from multiple wavelengths. Similarly, a few-term method[5] could use each spectrum's polynomial representation in the transport

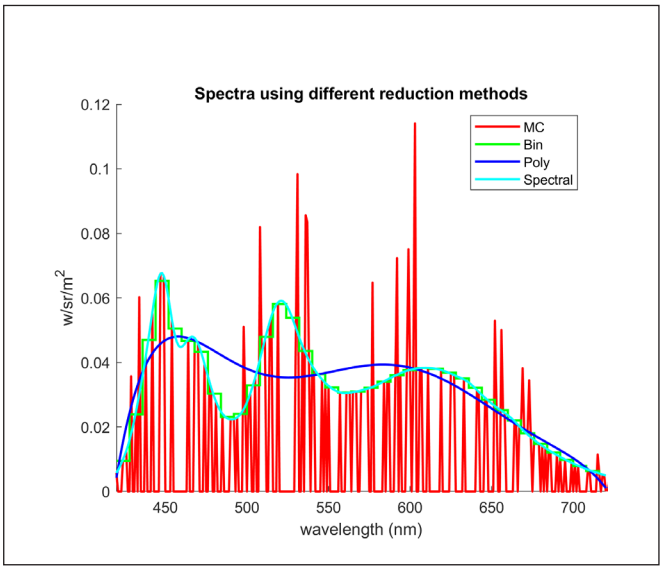


FIGURE 5. Different spectral reduction methods' effects on spectra.

and storage phases. The drawback is that when reconstructing the spectrum for the interactions, some loss of accuracy is inevitable. However, since most spectra are very smooth, this loss can be minimal. Lastly, the more common method is used,[3] whereby each ray is assigned a random wavelength through an MC simulation and traced through the scene. The drawback is that many rays are needed to ensure all characteristics influence the computation. **Figure 5** illustrates an example of a spectrum obtained through a MC simulation, an example of an original spectrum, a polynomial function that is a "best fit," and an example of how the spectrum could be binned. For this research, the studied methods are:

- Full Spectral Rendering—the ground truth, typically at 1nm intervals[10]
- PCA—Utilizing principal component analysis to reduce the size of the spectra[4,7]
- Binning—Splitting each spectrum into a smaller number of bins that contain the information of multiple wavelengths.[9]
- Monte Carlo (MC)—Randomly choosing wavelengths for each ray until a good level of quality is reached[3]
- Polynomial—Each spectrum is represented by several polynomial terms[5]
- RGB—Traditional RGB rendering using an RGB color space

Traditional RGB rendering pipelines store information about asset color in a texture map. In a graphics pipeline, there is a finite amount of textures and pixels that a computer can have readily available. Storing full spectral information means that a significantly smaller number of textures can be saved, as their size will increase. For some of these methods, the full spectrum must be computed and then inverted to enable interaction, such as in PCA and polynomial methods, which can increase computation time. The interactions occur in a space composed of vectors with 400 terms, allowing us to sample each wavelength in the visible spectrum and match the measurement device's reporting. Relying on spectral and color accuracy alone is insufficient to determine whether spectral rendering offers benefits that outweigh its drawbacks. At a glance, here are some disadvantages and advantages of each method and how they might be implemented in a rendering engine:

- Full Spectral Rendering—all spectra need to have full spectral information saved, and each time a shader operation executes, 400-value length arrays are executed.
- PCA—All spectra in the scene are analyzed to determine the principal components, eigenvalues, and eigenvectors. Each time a shading operation is executed, the spectrum must be rebuilt. Then, a new analysis needs to be performed to determine the new principal components, and all eigenvectors are saved to memory.

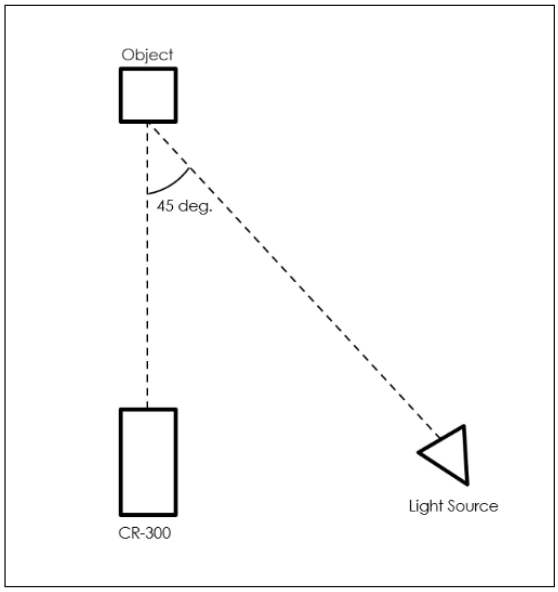


FIGURE 6. Measurement geometry



FIGURE 7. Textile samples (felts and shirts)(top), paint samples (middle), and color aid samples (bottom).

1545-0279/26©2026SMPTE

- Binning—This allows both texture and shading operations to run on smaller arrays, increasing efficiency, though at the expense of wavelength-domain precision and masking highly dynamic spectral profiles.
- Monte Carlo—Textures are still stored as full spectra, but each shading operation is executed on one (or a few) wavelengths per ray.
- Polynomial—Spectra are analyzed, and best-fit functions are determined. The coefficients are then stored. The spectra are then undone so that the shading operations can occur. The spectra must be re-analyzed after the shading operation to store new coefficients. This adds computation overhead for constant fitting and inspection.
- sRGB—Traditional RGB rendering with each texture being represented as three values, and each shading operation happens on these three values.

To convey color differences more easily, the CIE CAM 16[11] model was selected. This model converts XYZ values into Jab values, where J represents luminance and a-b represent color difference channels. The distance between points in the CIE CAM model is the color-difference metric used to measure each method's accuracy, as shown in Equation 1. Since LCH was developed with CAM02, we apply the same logic and derive a C for use in CAM16, as shown in Equation 1. However, since we may want to consider how other observers perceive these differences, a more observer-agnostic method should also be employed. Root Mean Square Error (RMSE) quantifies the difference between the actual and computed spectra across all wavelengths.

$$C = \sqrt{a^2 + b^2}$$

EQUATION 1. C from a,b in CIE CAM16.

It is also worth mentioning how color is managed in virtual production. This experiment will attempt to replicate a typical on-set experience, and one major issue in VP is color management. The primary goal of VP is to create sets that mimic reality and appear to the camera as if the action were taking place in a real-world environment. The Virtual Art Department goes to great lengths to ensure that the environments appear natural and that lighting, texture, and feel appear real. As discussed earlier, camera metamerism failure causes the camera to perceive the LED wall differently than the human eye would. To address this, the industry predominantly uses a method described by EPIC Games.[12] The goal is to ensure the final RGB values in the camera footage match the RGB values the VAD team observed during the design phase, which are presumed to accurately represent the real-life object they are trying to emulate. They propose a simple approach to characterizing the interaction between the camera and the LED and creating a correction matrix that gets applied to the LED display to ensure proper calibration.

Methodology

The following experiments were conducted to investigate the hypothesis that spectral rendering can enhance the matching of virtual production assets. The first experiment involved measuring the samples to ensure that our rendering engine is accurate. The second experiment involved filming the samples with a camera while they were displayed on screen, allowing us to observe the accuracy improvements of the renderings.

For the first experiment, we used a CR-300 spectroradiometer and a light source to measure the samples **(Fig. 6)**. In total, we selected 66 samples (excluding fluorescent samples): a Macbeth chart (24), 11 Textiles, 15 Paint Samples, and 15 Color Aid papers **(Fig. 7)**. They were each divided into separate groups, resulting in four distinct data sets. The measurement geometry was chosen to adhere to goniometric best practices[7] and measure only the diffuse reflectance of the samples, thereby eliminating the specular component (if any). We used four light sources:

- Creamsource Micro Bender—Broad Spectrum
- Creamsource Vortex V4—RGBWW Calibrated Mode—Broad Spectrum
- Creamsource Vortex V4—RGB Uncalibrated Mode—Narrow Band
- Quasar Rainbow RGB Only—Narrow Band

These lights were chosen because they are representative of current lighting fixtures found on movie sets. Additionally, the RGB lights **(Fig. 8)** utilize narrow-band emitters similar to those found in LED displays and are prone to metameric failures. In contrast, the Micro Bender and the RGBWW mode offer energy broadly across the spectrum. We set the lights to 6500 K, the default white point (D65) of sRGB and commonly referred to as the white point of daylight. To generate more data for our analysis, we also repeated the experiments with the lights set to 3200 K, a common temperature on sets. Because each light type has a different spectral composition, we balanced them by adjusting the settings until they produced the same x and y chromaticity values within their capabilities and within the two separate CCTs, mimicking typical set practices.

In total, each sample was measured eight times, once under each fixture and at each temperature. The data from the meter is reported in 1nm intervals. Reflectance spectra for the stimuli were determined by normalizing the full spectral

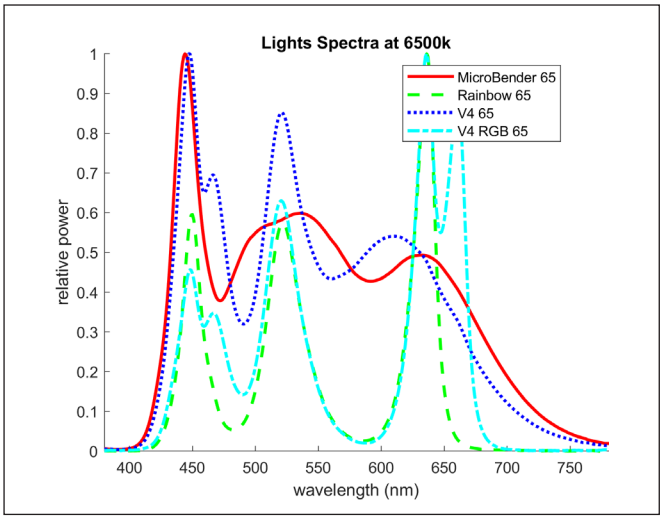


FIGURE 8. Peak normalized light spectra at 6500 K.

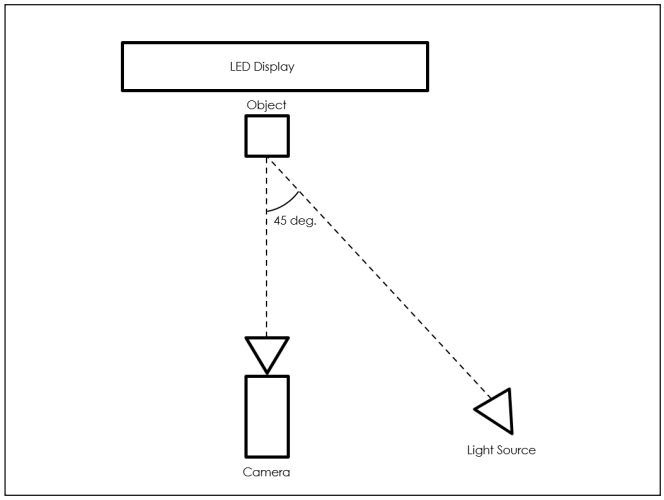


FIGURE 9. Filming geometry.

reflection from each sample under the broad-spectrum MicroBender and V4 lights to their respective sources, measured directly from a nominally perfect diffusing Spectralon reference. These measurements were performed with geometry matched as closely as possible to the photography to account for goniocolorimetric effects.

The render engine was then fed with reflectances of each sample, and each light-image combination was rendered. The engine was coded in Python and mimics a typical ray tracer. Inside the engine, the samples were illuminated by a light also positioned at the same distance and angle as in real life. To simplify the process, we used a diffuse illumination model, and all assets were evenly lit. For the spectral methods, we used the reflectances from the previous experiment and the light spectra of each lighting device. For the RGB renderings, we converted the reflectances to RGB values, and we used the RGB values for a 6500 K light.

These light values were obtained by converting the average x,y values of the lights to ACEScg RGB values. On startup, each reflectance and light source is converted to fewer terms based on the method (or kept as spectral data). Each time an interaction occurs, we recalculate the spectra and compute the interaction, except for the RGB renderings, where we compute the interaction using RGB terms. For PCA, this became increasingly costly, as the components had to be recalculated after each interaction.

To assess the advantages in virtual production, we conducted a second experiment. A RED Raptor camera was used to film both the rendered images and the real assets in all setups. The camera was kept at the same settings during takes, using the device's native settings where possible. A volume of ROE CB3 LED panels displayed the images. To ensure consistent virtual production practices, the LED display was calibrated to accurately represent colors to the cinema camera, as recommended by EPIC Games[12] and by Murray and Simion.[13] **Figure 9** shows the geometry used in filming. The digital images were scaled on the LED display so that the peak white patch matched real-life values. The footage was analyzed using Nuke, Python, and MATLAB, and the values of the samples were converted to CAM16 values using default viewing environments. These CAM16 values are scene-referred and will be used later to determine how closely the camera perceived the VP-filmed renderings and the real assets.

Results and discussion

First, we evaluated the number of terms used in each method (see **Table 1**). The number of terms determines the accuracy of the methods and also the storage capacity. Ultimately, for all our spectral calculations, we saved and computed 301 values. These values fall between 420 and 720 nm, as our light sources produced light only within this interval.

Simulations were conducted in MATLAB until we determined that we had used as few terms as possible without sacrificing accuracy. We targeted a deltaE CAM16 of 2.0 on methods that did not present an apparent asymptote, as the asymptote methods seemed to level off around that value. This deltaE CAM16 was calculated by comparing spectra before and after reconstructions using the chosen methods. It was gathered from synthesized spectra-to- colorimetry under prescribed ideal illumination and viewed by a (CIE 1931 2° observer). This allowed us to determine the ideal number of terms under many conditions, not just the ones studied. Asymptotes indicate diminishing returns when additional terms are added to some methods, such as Binning and Polynomial. The results of our testing are presented in **Table 2** and **Fig. 10.**

TABLE 1. Terms per pixel in each method used.

Method	RGB	Bin	MC	PCA	Poly	Spectral
Terms Per Pixel	3	38	150	6	6	301

TABLE 2. Mean and Max deltaE CAM16 for method testing showing the relationship between the number of terms and deltaE CAM16 between original spectrums and reconstructed spectrums using the chosen methods.

Method	Terms	Mean deltaE CAM16	Max deltaE CAM16
Bin	25	5.89	9.21
	30	3.56	5.39
	35	2.40	3.88
	40	1.75	2.55
	45	1.75	2.60
MC	100	7.12	10.55
	150	6.00	8.90
PCA	4	2.65	5.00
	5	2.14	3.58
	6	1.59	2.66
	7	1.08	2.38
	8	0.63	1.28
Poly	4	4.06	8.51
	5	2.44	5.77
	6	2.08	5.34
	7	1.26	2.52
	8	1.19	2.34

1545-0279/26©2026SMPTE

TABLE 3. Average and Max RMSE between methods within sample sets and collapsed across all sets. The comparison was made between the radiance measurement and the predicted radiance, with all data peak-normalized using a perfect reflective diffuser.

Method	Macbeth	Textiles	Paints	ColorAid	Collapsed
Average					
PCA	3.43E-03	3.42E-03	4.33E-03	3.90E-03	3.77E-03
Poly	4.08E-03	4.23E-03	3.93E-03	4.61E-03	4.21E-03
Bin	1.17E-05	5.37E-05	1.17E-05	1.30E-05	2.25E-05
Max					
PCA	2.65E-02	2.16E-02	2.57E-02	2.16E-02	2.65E-02
Poly	1.17E-02	3.12E-02	2.95E-02	2.91E-02	3.12E-02
Bin	3.45E-05	6.11E-04	9.20E-05	1.06E-04	6.11E-04

TABLE 4. Average and Max RMSE between methods within lighting devices and collapsed across all devices.

Method	MicroBen der32	MicroBen der65	Rainbo w32	Rainbo w65	V432	V465	V4RGB3 2	V4RGB6 5	Collapsed
Average									
PCA	1.84E-03	2.30E-03	5.57E-04	5.31E-04	2.75E-03	2.89E-03	1.02E-02	9.12E-03	3.77E-03
Poly	1.51E-03	2.20E-03	6.16E-04	5.84E-04	2.19E-03	2.50E-03	1.35E-02	1.06E-02	4.21E-03
Bin	6.01E-06	7.68E-06	1.69E-06	7.01E-05	8.37E-06	1.15E-05	3.88E-05	3.62E-05	2.25E-05
Max									
PCA	3.79E-03	8.12E-03	1.60E-03	1.53E-03	6.48E-03	1.05E-02	2.65E-02	2.37E-02	2.65E-02
Poly	3.93E-03	7.32E-03	1.68E-03	1.61E-03	5.92E-03	8.54E-03	3.53E-02	2.90E-02	3.53E-02
Bin	1.27E-05	1.82E-05	6.02E-06	4.95E-06	1.85E-05	3.64E-05	1.06E-04	9.12E-05	1.06E-04
LightType	Wide-band	Wide-band	Narrow-band	Narrow-band	Wide-band	Wide-band	Narrow-band	Narrow-band	

TABLE 5. deltaE CAM16 statistics for all color samples and data reduction methods.

Samples	RGB	Bin	MC	PCA	Poly	Spectral
Average						
Macbeth	5.610	5.632	6.559	5.712	6.424	5.507
Textiles	9.624	7.585	8.326	8.562	7.856	7.827
Paints	7.252	6.423	7.404	6.861	7.003	6.102
Color Aid	9.937	7.058	7.865	7.544	7.445	7.179
Max						
Macbeth	10.132	11.852	14.669	10.210	12.609	10.589
Textiles	17.935	16.950	17.026	16.647	16.928	16.908
Paints	15.401	16.608	17.490	12.988	14.036	14.158
Color Aid	18.644	15.953	16.256	15.816	16.563	16.213

The spectral distributions resulting from the renderings were also analyzed. The Monte Carlo (MC) method was excluded because it sometimes failed to choose a wavelength, rendering the RMSE values inaccurate. The Spectral method was excluded because it consistently yielded results within a very small error, likely due to engine precision, similar to the measured data. By inspecting the max and average RMSE values **(Tables 3–5)**, we can see that the binning method consistently outperformed PCA, which in turn outperformed the Polynomial method. With narrow-band emitters, PCA outperformed the Polynomial method.

Conversely, with broad-spectrum lights, the Polynomial method performed slightly better. This could be because polynomial methods struggle with spectra that have sharp peaks, as they require more terms to represent the values accurately. **Figure 11** illustrates various spectral reconstructions. One notable issue arises when using narrow-band emitters.

Some methods fail to preserve the peaks properly, most likely because they use fewer terms.

For the virtual production test, we reviewed the footage to determine if there were any improvements in accuracy. The RGB values of each patch were extracted and converted from camera-native values to linear ACEScg RGB values. The values were then converted to CIE CAM 16 values, and then the delta E between the real and VP values were calculated. The mean and standard deviation are presented in **Figs. 12 & 13.**

We can see that RGB consistently underperforms and has a higher deltaE CAM16 than any spectral method. As expected, the full spectral rendering is the most accurate, with the Binning method closely following. The polynomial method consistently underperforms. We should note that, when comparing the difference between the V4 set in narrow-band **(Figs. 14 & 15)** and broad-band modes

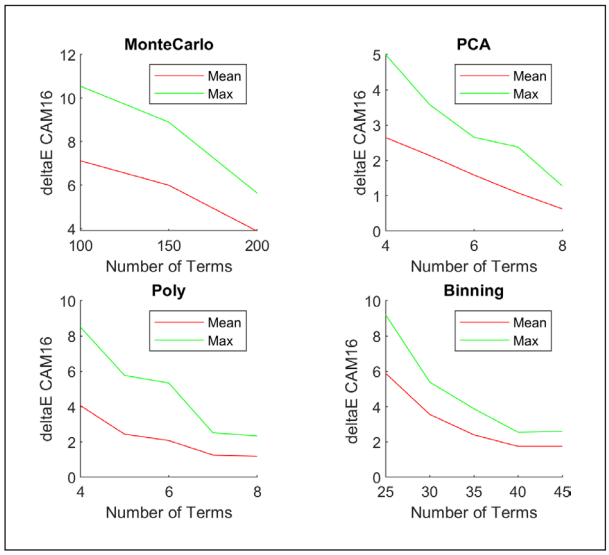


FIGURE 10. Number of terms affecting deltaE CAM16 plots, showing asymptotes that indicate diminishing returns when terms are added in certain methods, which influenced term decision.

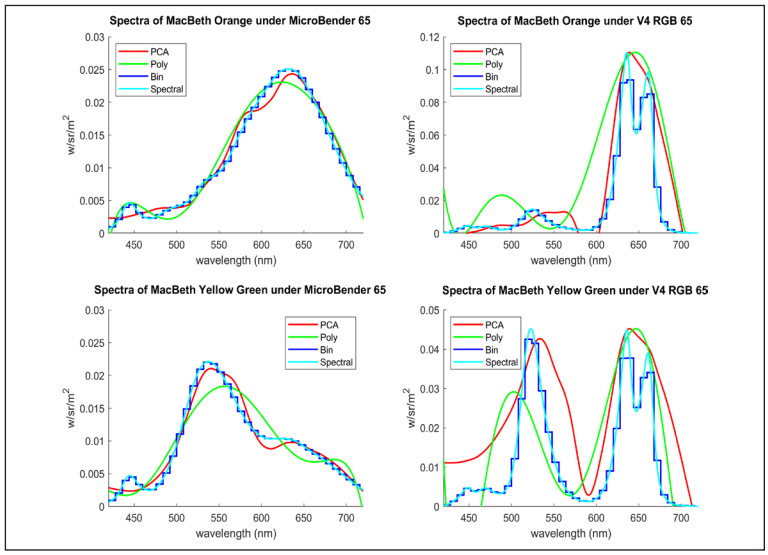


FIGURE 11. Different methods showing the reconstructed reflectance spectrum of Macbeth patches under different lighting conditions.

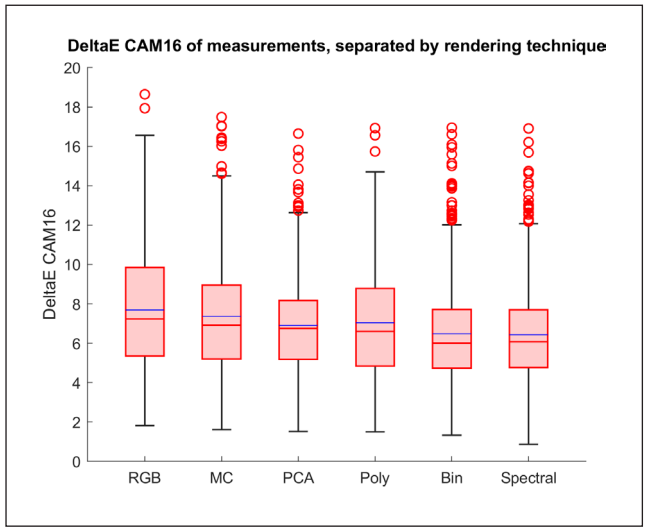


FIGURE 12. deltaE CAM16 data sorted by rendering methods with mean (blue), median (red), and lower and upper quartiles (red box), minimum and maximum data points not considered outliers (whiskers), and outliers (red circles).

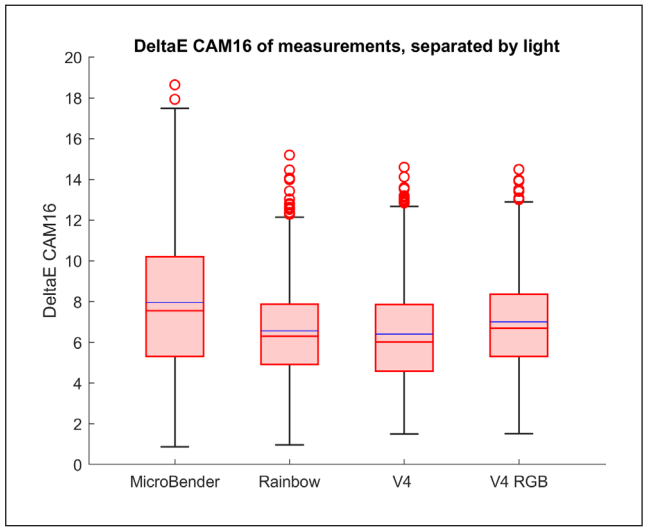


FIGURE 13. deltaE CAM16 data sorted by light device, with MicroBender and V4 being wide-band emitters and Rainbow and V4RGB being narrow-band emitters.

(Figs. 16 & 17), and when using PCA, deltaE CAM16 is higher in the narrow-band emitting mode. This could be because the narrow-band spectrum is much different from the smoother reflectance spectra, thus increasing the error, as we would need more eigenvectors to describe the behavior properly. The MC method also has a large deltaE CAM16 average, potentially due to noise, as the image seemed much less uniform.

When converting to CAM16 JAB, we used a Spectralon measurement for each light scenario as the scene illuminant value so data was not lost to clipping. These trends were observed across all datasets; therefore, we plotted only the Macbeth data at 6500 K for easier visualization.

For the visualization shown in **Figs. 14 & 16**, the dynamic range was matched during footage processing, as we noticed that in some cases, the grayscale of the charts was not evenly distributed across the camera's dynamic range. The CAM16 plots also confirm the dynamic range issue identified during image visualization, which can also be attributed to lighting uniformity. While the charts in real life are not evenly lit, as some of the lights are brighter toward the middle of the chart, even with diffusion, the LED wall does not suffer from this issue, and it remains evenly bright across the entire surface.

The yellow patch in **Fig. 14** shows this dynamic-range difference: it consistently looks darker than the surrounding patches, which could be due to luminance mapping from screen to camera. Interestingly, the ab color difference plots **(Figs. 15 & 17)** show no directional trends. This makes fixing the differences post-recording extremely difficult, as one

1545-0279/26©2026SMPTE

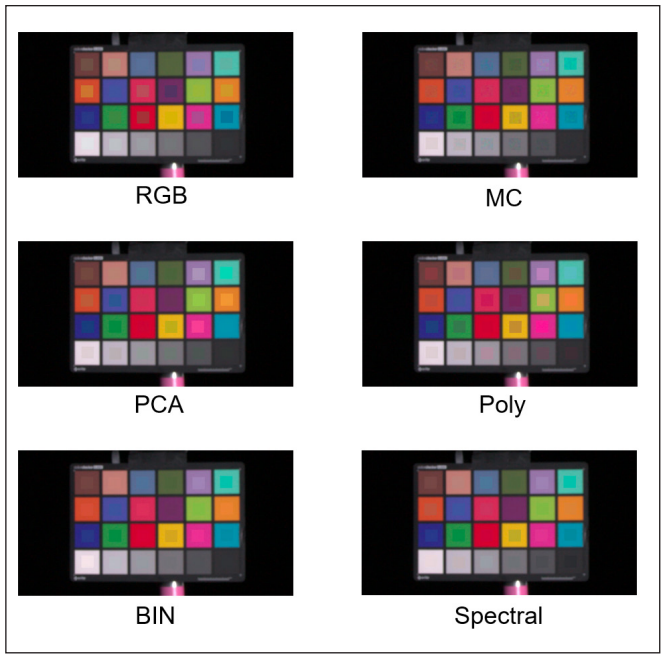


FIGURE 14. Example of the Macbeth chart illuminated by the narrow-band emitting V4RGB65 light source, after dynamic range matching. Outer patches are the real filmed samples, while the inner samples are the rendered, filmed versions.

would need to model the behavior and then create a proper fix. This can become very time-consuming, making it crucial and beneficial to address the behavior in the rendering.

Table 5 contains statistics gathered from the other samples. Further analyzing the data from the other samples reveals that they are similar trends of what methods perform better than others, and the lack of directionality of deltaE CAM16. Correlating this with previous RMSE data, the RMSE from the spectra accurately indicates which method will perform better in this experiment. Methods that do not preserve the spectrum well tend to underperform.

A few limitations influenced the data. First, the camera does not fully adhere to the Maxwell-Ives-Luther criterion, which means that the camera perceives stimuli differently than the human eye would. However, since this error applies to everything the camera perceives, and we are interested in differences between stimuli, it does not affect the conclusions. Our minimum observed average deltaE CAM16 is around 6. It might present itself as a large number; however, it originates from a multitude of sources that compound. Another factor is the accuracy of the calibration between the LED wall and the camera. The average deltaE CAM16 of the calibration is 1.63, with a maximum of 2.22. There may also be slight errors in our measurements due to variations in angles and light fall-off, as well as lens effects such as bloom or vignetting. To alleviate this, each sample was averaged over the entire area rather than selecting a single pixel. Other sources of error could also come from the LED panel drifting after being powered on, lights reacting to temperature differences after being powered on, or even the camera noise floor changing over time. All data from this paper is available upon request.

This research aims to inspire current industry trends to

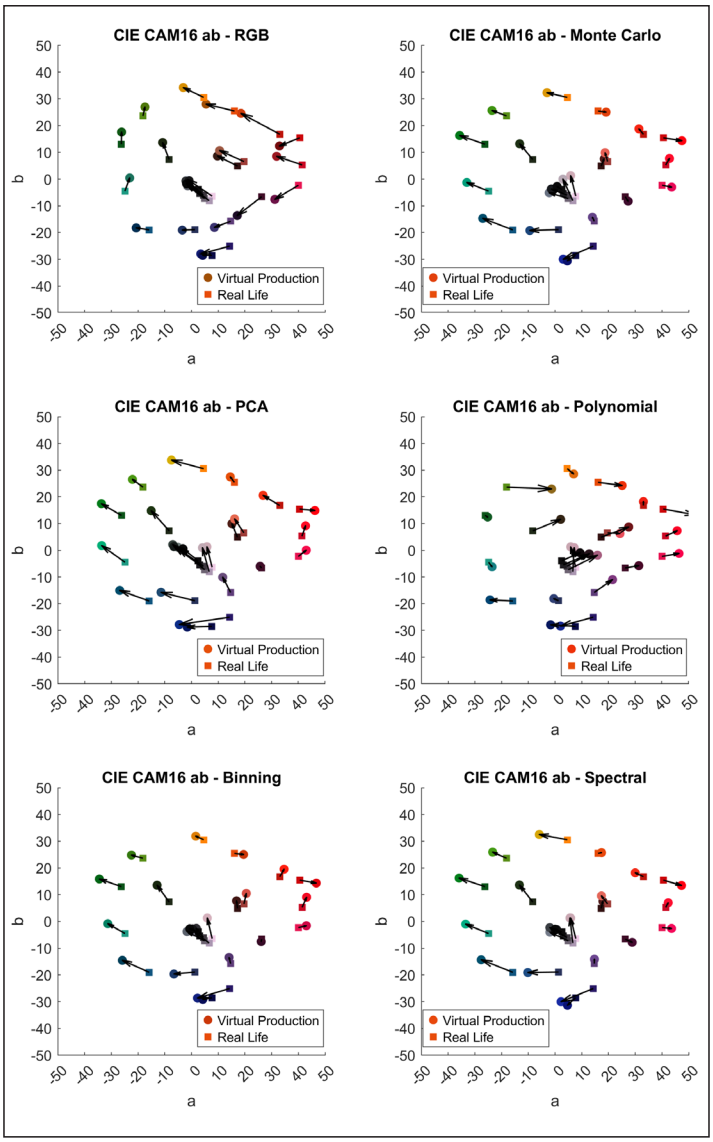


FIGURE 15. Differences in ab between real data and rendering data of the MacBeth from CAM16, as seen by the camera looking at the narrow band emitting V4 RGB 65. Each data point has the color of the patch it represents.

move toward a more spectrally accurate solution, which can alleviate color-matching issues between real and virtual assets. The paper demonstrated how pure spectral rendering increases accuracy; however, it is costly in both time and memory usage. Using methods that can compress spectral data helps reduce complexity, albeit at the cost of accuracy. Some of these methods performed as expected, with MC being largely a hit or miss, and PCA outperforming most methods but requiring more time. The polynomial method appears to be time-consuming (as it requires extensive fitting) and offers little benefit, as it sometimes creates impossible spectra that significantly increase the error. The binning method performed the best, and perhaps we can draw some inspiration from it. Further optimization of the number of terms could also yield more accurate results. There could even be a solution developed by combining different methods (for example,

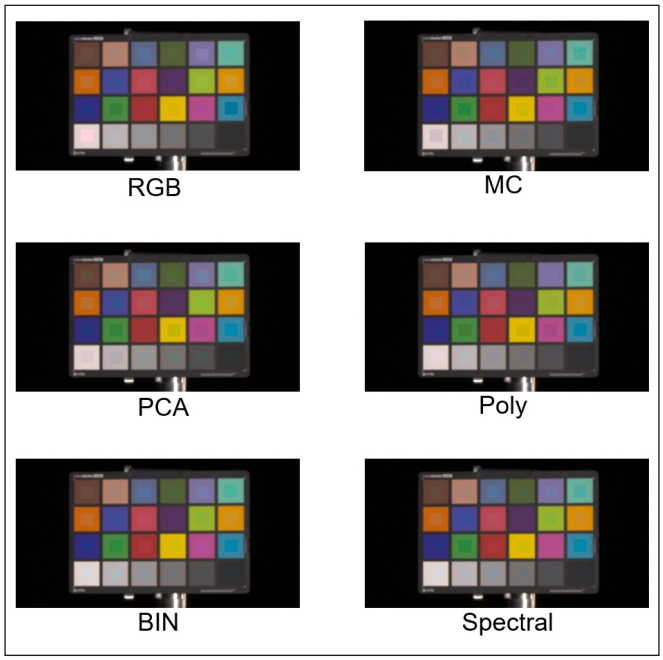


FIGURE 16. Example of the Macbeth chart illuminated by the wide band emitting V4 65 light source, after dynamic range matching. Outer patches are the real filmed samples, with inner samples being the rendered, filmed versions.

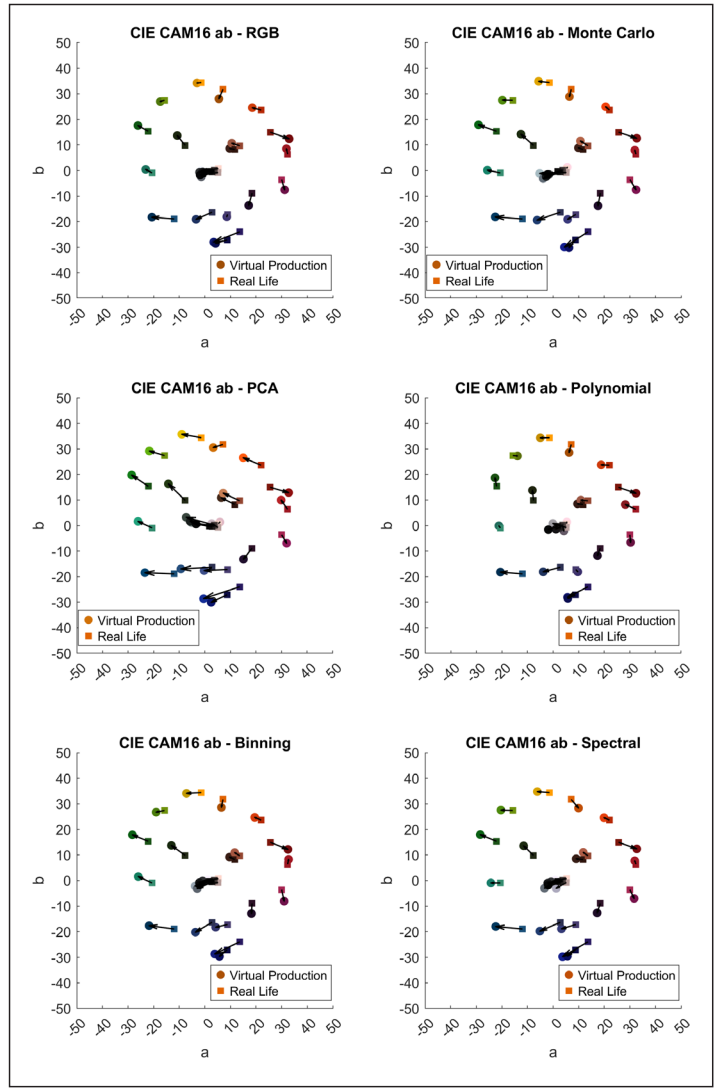


FIGURE 17. Differences in ab between real data and rendering data of the Macbeth from CAM16, as seen by the camera looking at the wide band emitting V4 65. Each data point has the color of the patch it represents.

using a more optimized PCA but performing the interactions at larger intervals, more similar to binning). In a future iteration of research on these topics, one could also include the spectra of the LED and the camera to create RGB values that can be used in VFX compositing and "fix" the camera-to-wall calibration spectrally. Lastly, optimizing computing methods could tip the scales in favor of one method over the other in terms of time and memory usage. We recommend that the industry investigate implementing these methods in order to improve light simulation in rendering engines. The industry could also explore more spectral reconstitution methods that may create synthetic spectra from current RGB assets, thereby aiding in color matching under various illuminants.

Conclusion

This research showed that rendering assets spectrally can provide significant improvements over traditional RGB rendering in terms of matching real and virtual objects in virtual production. Through a series of experiments involving various lighting conditions and data reduction techniques, along with full spectral rendering, we demonstrated measurable improvements over traditional RGB rendering. These findings suggest that adopting spectral rendering approaches significantly improves the realism and consistency of digitally created assets in VP. With technology improvements over the last years, it may be possible to implement these techniques effectively and bring their benefits into production environments.

Acknowledgments

The authors would like to acknowledge the RIT Munsell Color Science Laboratory, ROE, Megapixel VR, Creamsource, and Quasar for their support throughout this research. Flip Phillips supported, in part, by the National Science Foundation under Grant Award Number 2238180.

References

1. S. Zwerman and J. A. Okun, *The VES Handbook of Virtual Production*. New York, NY, USA: Routledge, 2024.
2. M. Pharr, W. Jakob, and G. Humphreys, *Physically Based Rendering: From Theory to Implementation*. Cambridge, MA: MIT Press, 2023.
3. L. Fascione, J. Hanika, M. Leone, M. Droske, J. Schwarzhaupt, T. Davidovič, A. Weidlich, and J. Meng, "Manuka: A Batch-shading Architecture for Spectral Path Tracing in Movie Production," *ACM Trans. Graph.,* 37 (3): 1–18, Aug. 2018. doi: 10.1145/3182161.
4. M. S. Peercy, "Linear Color Representations for Full-speed Spectral Rendering," *Proc. SIGGRAPH '93*, 1993, pp. 191–198, doi: 10.1145/166117.166142.
5. M. G. Raso and A. Fournier, "A Piecewise Polynomial Approach to Shading Using Spectral Distributions," *Proc. Graphics Interface '91*, Calgary, AB, Canada, pp. 40–46, June 3–7, 1991. doi: 10.20380/GI1991.06.

1545-0279/26©2026SMPTE

6. G. W. Meyer, "Wavelength Selection for Synthetic Image Generation," *Comput. Vis. Graph. Image Process.*, 41 (1):57–79, Jan. 1988. doi: 10.1016/0734-189X(88)90117-X.
7. R. S. Berns, F. W. Billmeyer, and M. Saltzman, *Billmeyer and Saltzman's Principles of Color Technology.* Hoboken, NJ: Wiley, 2019.
8. G. Wyszecki, *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Hoboken, NJ: Wiley, 2001.
9. C. C. Borel and R. F. Tuttle, "Multi- and Hyperspectral Scene Modeling," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVII*, 2011. doi: 10.1117/12.884585.
10. G. M. Johnson and M. D. Fairchild, "Full-Spectral Color Calculations in Realistic Image Synthesis," *IEEE Comput. Graph. Appl.*, 19 (4): 47–53, Jul./Aug. 1999. doi: 10.1109/38.773963.
11. C. Li, Z. Li, Z. Wang, Y. Xu, M. R. Luo, G. Cui, M. Melgosa, M. H. Brill, and M. Pointer, "Comprehensive Color Solutions: CAM16, CAT16, and CAM16-UCS," *Color Res. Appl.*, 42 (6): 703–718, Nov./Dec. 2017. doi: 10.1002/col.22131.
12. Epic Games, "Camera color calibration for in-camera VFX in Unreal Engine," *Unreal Engine 5.5 Documentation*, Epic Developer Community. [Online]. Available: https://dev.epicgames.com/documentation/en-us/unreal-engine/camera-color-calibration-for-in-camera-vfx-in-unreal-engine
13. K. Murray and V. Simion, "Color management in virtual production," *ACM SIGGRAPH 2023 Talks,* 2023, pp. 1–2, doi: 10.1145/3587421.3595445.

About the Authors



Vlad Simion is a color scientist who works at Seismiq, building color pipelines, imaging workflows, and color management software for virtual production, immersive, and emerging technologies.



David Long, PhD, is Director of MAGIC Spell Studios and Professor of Motion Picture Science at Rochester Institute of Technology.



Flip Phillips, PhD, is Professor of Motion Picture Science at Rochester Institute of Technology.



Kris Murray is CTO of 180 Studios and co-founder of Seismiq, building virtual production workflows and technology tools for filmmaking and creative teams.

Paper submitted for Journal first publication and presented at the SMPTE 2025 Media Technology Summit, Pasadena, CA, 16–19 Oct. 2025. *Copyright © 2026 by SMPTE.*

DOI: 10.5594/JMI.2026/JGAM2497
Date of publication: 1 October 2026



Agfacolor at 90: The Race for Color and the Technology That Endured

In October 1936, "AGFACOLOR Neu" was introduced as the world's first modern single-strip color film, triggering a "photographic" arms race during WWII. It endured long after and provided the foundation for many other international color film formats after the war; its specific color tones and palette gave many Eastern European films an iconic look. On this 90th anniversary, we should remember this important part of film history.

By Michael Zink

When Wilhelm Schneider took a "small detour into color photography,"[1] he did not know he was on the precipice of a world invention that would leave its mark on photography and cinematography to this day. Dr. Schneider worked at the Agfa Filmfabrik (film factory) in Wolfen, Germany, where he focused primarily on developing black-and-white film. However, in 1932, he noticed that colored residual images sometimes appeared during the development process. This discovery ultimately led to the invention of the first subtractive 3-color, chromogenic monopack reversal film, Agfacolor Neu, which celebrates the 90th anniversary of its debut on October 17, 1936.

The Original Work from Rudolf Fischer

The foundation of Schneider's work was based on research from about twenty years prior, when Dr. Rudolf Fischer discovered that "the addition of 'certain materials' to certain developing solutions [...] created a dye image."[2] He patented multiple inventions related to this discovery in 1911/12, describing how a single film could contain three light-sensitive layers, each containing a so-called "color coupler" that reacted with the developer to generate the desired colors. The goal was simple and appealing: an integrated monopack film in which three layers of color couplers were part of the emulsion to provide the chromogenic attributes. In other words, the colors would be generated in the film emulsion itself. If those layers were thin enough, the film could be run through existing cameras, allowing any user familiar with black-and-white photography to record color photographs.

However, Fischer's invention faced several challenges. Coupler instability could cause color shifts, and coupler migration could result in muddy-looking images. Additionally, reducing the film thickness to leverage existing equipment was not trivial either. In fact, Fischer solved none of these challenges, and his ideas did not materialize until Schneider picked up the work some 20 years later.

The Agfacolor Invention

Schneider and his team at Agfa found that "it was possible to anchor couplers within the proper emulsion layer by attaching long, fatty acids as side groups on the dye-coupler ring structures."[2] This discovery prevented the migration of the couplers into other emulsion layers, solving some of the biggest issues with Fischer's invention.[3] Scheider's invention was patented in 1935, describing improvements for dye diffusion, and it was the basis of the "Agfacolor Neu" film.

Furthermore, the researchers were able to "manufacture a three-layer product that was no thicker than black-and-white stocks, greatly reducing light scatter and greatly improving sharpness."[2] Another major advantage of these new color couplers was that a single developer could be used for all three color layers at the same time to generate the desired dyes. This proved to be a significant improvement from a process perspective.

The Kodachrome Approach

Earlier that year (1935), Kodak introduced its own integral 3-color system—Kodachrome—into the market, beating Agfa. However, that system did not solve the coupler-migration issue; instead, it used a workaround. The couplers were part of the developer solution, not the film itself. While this had some compelling benefits, such as a thin film with less light scattering, resulting in crisper images with vibrant color, the problem was the complexity of the film development process.

The Kodak process required multiple different machines

1545-0279/26©2026SMPTE

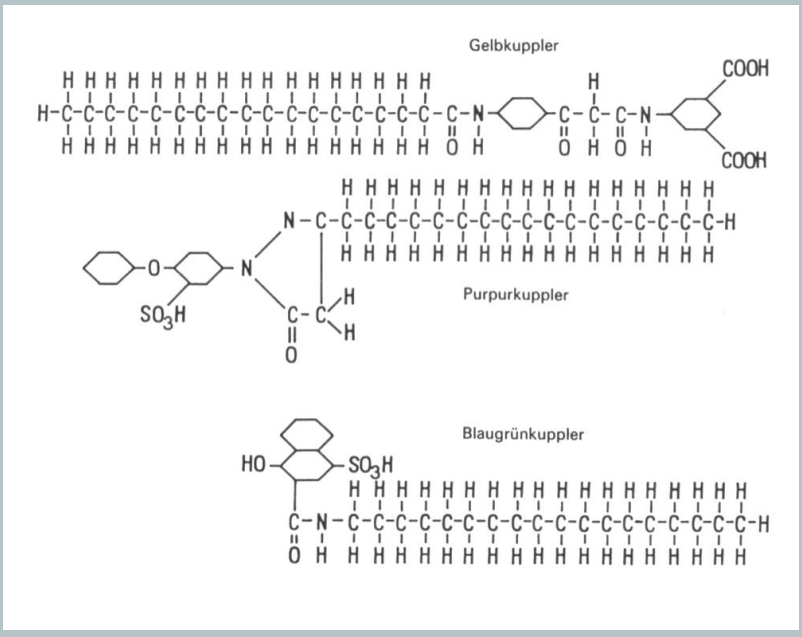


Figure from "film production," brochure of VEB Filmfabrik Wolfen (ORWO), GDR 1981.



Das kalte Herz (GDR 1950, Paul Verhoeven). *Credit: DEFA Foundation. Photographs of the Orwocolor print by Josephine Diecke, SNSF project Film Colors. Technologies, Cultures, Institutions.*

and chemical baths, and even after process improvements over the next few years, "Kodachrome processing consisted of at least 28 discrete steps," keeping it "enormously complex."[2]

By contrast, Agfacolor Neu had the color couplers in the film emulsion itself, which meant the film could be developed in only four baths. Coincidentally, that was the same number of baths required for black-and-white film, and obviously a huge reduction compared to the 28 steps with Kodachrome.

The "Arms Race" Between Rochester & Wolfen

Given that the breakthrough pursued by Schneider's team was interrupted by the launch of Kodachrome in April 1935, followed by the release of Agfacolor Neu in October 1936, the race between both sites—Kodak in Rochester, NY, and Agfa in Wolfen—intensified. The team at Wolfen grew to about 50 people, which was remarkable, because Scheider's department (the Technical-Scientific Laboratory) was not originally responsible for color film research. Instead, they were responsible for finding practical applications for work from another department (the Scientific Photo Laboratory). But given their success, Schneider received approval from his manager, Gustav Wilmanns, to formally pursue this research.[3]

The team size at Kodak likely grew to similar dimensions, and the battle expanded to other fronts as well. On the legal front, Kodak objected to the original patent underpinning the Agfacolor Neu process, but the lawsuit was ultimately settled in 1943. Then there was the geopolitical conflict between the two countries: World War II.

Immediately after the announcement of "Agfacolor Neu," German propaganda minister Joseph Göbbels pressured the team at Agfa to move forward with color film production. He understood the value of color images as psychologically, geopolitically, and strategically important.

In fact, in 1939, the decision was made to switch the German cinema production to the new Agfacolor process, as the production of the first full-feature film "Frauen sind doch bessere Diplomaten" ("Women Are Better Diplomats") got

underway. The production was anything but flawless, and Schneider, who was present for much of the filming, realized the film required four times as much lighting as black-and-white films. After many improvements to the film material and multiple reshoots, the film was released in October 1941, but the color reproduction was not fully satisfactory.[1]

Subsequent UFA productions achieved much better color reproduction, including "Die Goldene Stadt" ("Golden City"), which premiered in October 1942, followed by the large-scale production of "Münchhausen" ("The Adventures of Baron Munchhausen"), which premiered in March 1943.[1]

Post-War Situation

After World War II, the Agfacolor process and methods attracted worldwide interest. Kodak and many other companies sent specialists "for a debriefing of Agfa color chemists." Ultimately, "the US government seized the Agfa Corporation's chemical patents, and published them 'without restrictions or license' in 1945."[2]

As a result, many international photographic companies adopted or improved on the Agfacolor framework. This included notable names such as Fujicolor (Japan), Sovcolor (Soviet Union), Gevacolor (Belgium), and Ansco Color (U.S.), among others.

When Kodak ultimately released Eastmancolor in 1950, it paved the way for a single-strip color film to be used with much more simplicity in Hollywood productions, replacing the much more complicated Technicolor process.

To this day, it is not entirely clear whether the Kodak solution was a derivative of the Agfa solution, but most evidence suggests the similarities are merely correlation and that both systems were developed in parallel. Both Agfa and Kodak were under enormous government pressure during wartime, but "the war also ensured that the American and German corporations worked largely in isolation from each other."[2]

After the post-war division of Germany, the Agfa Filmfabrik in Wolfen was on Soviet territory. Agfa also opened a new plant in West Germany (in Leverkusen on British territory). For several years, both plants produced films under the same "Agfacolor" name, yet the processes were different derivatives. In 1964, the Filmfabrik in Wolfen changed its name to ORWO (which stands for ORiginal WOlfen) and rebranded future films as "Orwocolor."[4]

The Agfacolor Legacy

Agfacolor remained the foundational color system across the Soviet Union (under the brand name "Sovcolor") and the Eastern Bloc. East Germany's film production studio DEFA released its first film in Agfacolor in 1950: "Das kalte Herz" ("Heart of Stone"). It remains one of DEFA's most successful films, with almost 10 million tickets sold at release.
Between 1950 and 1964, DEFA produced more than 50 feature films in Agfacolor. After that, the brand disappeared when it was renamed Orwocolor.

In 1969 at the 41st Academy Awards, the film "War and Peace" won the Oscar® for Best Foreign Language Film. The 70mm epic film was primarily shot in Sovcolor, but problems with unstable Soviet film batches required filmmakers to turn to Wolfen, drawing on its expertise and using Orwo film stock to supplement production.

Orwocolor remained a major success in East Germany and across the Eastern Bloc until the Filmfabrik was liquidated in 1994, following German reunification. Even after its dissolution, however, the inventions developed by Schneider and his colleagues at Agfa endured and remain a critical part of film technology to this day.

References

1. R. Karlsch and P. W. Wagner, Die AGFA-ORWO-Story: Geschichte der Filmfabrik Wolfen und ihrer Nachfolger. Berlin, Germany: Verlag für Berlin-Brandenburg, 2010.
2. H. Heckman, "Undervalued stock: Eastman Colors innovation and diffusion, 1900–1957," Ph.D. dissertation, Univ. of Wisconsin–Madison, Madison, WI, USA, 2014. [Online]. Available: University of Wisconsin Digital Collections.
3. G. Koshofer, „Agfacolor für Fotografie und Kino: Eine Pionierleistung ‚Made in Germany,'" in UFA in Farbe: Technik, Politik und Starkult zwischen 1936 und 1945, F. Beyer, G. Koshofer, and M. Krüger, Eds. Munich, Germany: Collection Rolf Heyne, pp. 40–55, 2010.
4. J. Diecke, "Keeping your enemies closer: Strategies of knowledge transfer at the East German Filmfabrik Wolfen," in Global Film Color: The Monopack Revolution at Midcentury, S. Street and J. Yumibe, Eds. New Brunswick, NJ, USA: Rutgers Univ. Press, pp. 30–47, 2024.

Thank you to Michael Zink for contributing this historical overview, an important chapter in motion imaging history.

DOI: 10.5594/JMI.2026/SOVO9428
Date of publication: 1 October 2026

1545-0279/26©2026SMPTE

BY MICHAEL DOLAN



25 Years Ago in the Journal

The October 2001 *Journal* published in: "A Short History of Standardization in the SMPTE, by Gordon A. Chambers." "At the incorporation meeting in Washington, D.C. on 24 July 1916, Henry D. Hubbard of the Bureau of Standards spoke on the desirability of standards in the industry. The dimensions he proposed for what we now call 35mm film were exceedingly close to what is used today. Frame-Line Location...American War Standards/JAN...Pitch...Television...Video Magnetic Recording...International Standardization (and) Standards and Recommended Practices. The Society publishes annually a list of its Standards. In addition, related material not considered as meeting the criteria for standards is prepared and published by the Society as "Recommended Practices" and Engineering Guidelines.
For the full article, see: https://tinyurl.com/3krzht39

> "Large television centers have evolved over the years to include numerous production studios, telecine equipment, videotape machines, master control for audio and video, incoming and outgoing audio and video circuits, viewing rooms, equipment rooms and other technical services."

50 Years Ago in the Journal

The October 1976 *Journal* published in: "Industry News & Educational Activities," "Highlights of the Motion Picture Laboratories 19th Annual Seminar held 23-25 July in Memphis, Tenn.," included presentations by Garrett Brown, inventor of Steadicam, the spring-mounted camera stabilizer; King Vidor, veteran Hollywood director; Arthur Scott animation director of Hanna/Barbera, Hollywood; and other distinguished personalities in the motion picture world." Also included are: "Highlights of the Motion Picture Laboratories 19th Annual Seminar; The American Film Institute, The Canadian Broadcasting Corp, Free-loan individual titles; Unknown

In this column we provide interesting historical briefs from the Journal articles of days past. The purpose of this column is primarily entertainment, but we hope it will also stimulate your thinking and reflection on the Society's history, how far we have come in the industry, and (sometimes) how some things never change.

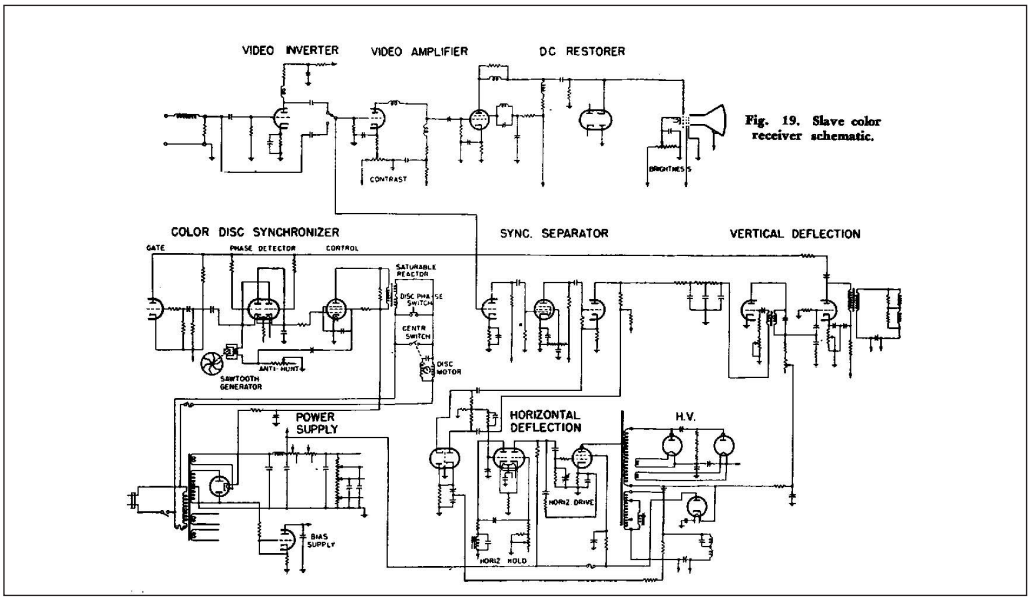


Fig. 19. Slave color receiver schematic.

Slave color receiver schematic (**Fig. 19** from *JSMPTE*, Oct. 1951, p. 359).

Risk; The Queen's Award to Industry; The Crisp-Matic; The Arriflex Company of America; Carter Equipment Co; Plastic Reel Corp, of America; Sony Corp. of America and Paramount Pictures Corp, a subsidiary of Gulf & Western Industries Inc; have formed Sony/Paramount Home Entertainment Center; The CATV Equipment & Installation Operation of GTE Sylvania;" and other. For the full article, see https://tinyurl.com/wbaymtes

75 Years Ago in the Journal

The October 1951 *Journal* published in: "Color Television – U.S.A. Standard" by P. C. Goldmark, J. W. Christensen, and J. J. Reeves," "Most of the fundamental data and early developmental stages of the field-sequential color-television system have been available to the engineering profession, partly from earlier publications and partly from material presented at the FCC hearing. Now that color television has attained the status of commercial operation, it seems advisable to fashion this paper in such a way that it is most useful to the studio and receiving equipment engineer. It is appropriate to begin with a recital of the official FCC color-television standards as they appeared in the Federal Register and to follow this with a brief discussion of their significance from a colorimetric point of view...5. Color transmission. The term "color transmission" means the transmission of color television signals which can be reproduced with different values of hue, saturation and luminance." For the full article, see: https://tinyurl.com/3c83nu6d

100 Years Ago in the Journal

The October 1926 (issue 27) published in: "The Effect of Motion Pictures on the Eyes" by Guy A. Henry. "A few years ago a public official in addressing a national welfare organization made the following startling prophecy: "Motion pictures will be extinct in ten years. The public realization that they are ruining eyesight will lead to a demand that they be abolished. Within ten years I predict that there will be no more motion picture shows in America. By that time they will be barred as a pernicious evil. They will be dropped by common consent for the common good as other useless things have been dropped in the past."..."Had this prediction been made in the early days of the moving picture when poor photography and faulty projection with distressing flicker prevailed, there would have been some grounds for it; but such wonderful changes have taken place in the production of motion pictures and in their showing that in recent years there has been no just cause for serious apprehension as to motion pictures harming the eyes, providing attention is given to certain conditions. Under favorable conditions moving pictures do not cause serious eye fatigue, but it must be borne in mind that several very important elements are necessary to make these "favorable conditions." For the full article, see: https://tinyurl.com/bp8k7cd8

DOI: 10.5594/JMI.2026/XDSA5574
Date of publication: 1 October 2026

1545-0279/26©2026SMPTE

STATEMENT OF OWNERSHIP, MANAGEMENT, AND CIRCULATION (REQUIRED BY 39 U.S.C.3685)

1. **Title of publication:**
   SMPTE Motion Imaging Journal

2 **Publication No. 285-200**

3. **Date of filing:**
   Oct. 1, 2026

4. **Frequency of issue:**
   Monthly

5. **No. of issues published annually:**
   7

6. **Annual subscription price:**
   $1,000

7. **Complete mailing address of known office of publication:**
   50 Main St., Suite 1000, White Plains
   (Westchester County), NY 10606-1953

8. **Complete address of the headquarters or general business offices of the publishers:**
   50 Main St., Suite 1000, White Plains
   (Westchester County), NY 10606

9. **Full names and complete addresses of publisher, editor, and managing editor:**
   Publisher: Society of Motion Picture and Television Engineers, Inc.
   50 Main St., Suite 1000, White Plains
   (Westchester County), NY 10606-1953

10. **Owner: Society of Motion Picture and Television Engineers, Inc.**
    50 Main St., Suite 1000, White Plains
    (Westchester County), NY 10606-1953

11. **Known bondholders, mortgages, and other security holders owning or holding 1% or more of total amount of bonds, mortgages, or other securities:**
    None

12. **The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes:**
    Have not changed during preceding 12 months

13. **Publication Title:** SMPTE **Motion Imaging Journal**

14. **Issue Date of Circulation Data:**
    October 2025 – September 2026

15. **Extent and nature of circulation:**

A. Total no. copies printed (net press run)	1,507*	1,054**
B. Paid and/or requested distribution:		
1. Sales through dealers and carriers, street vendors, and counter sales	---	---
2. Mail subscriptions	1,357*	929**
C. Total paid and/or requested circulation	1,357*	929**
D. Free distribution by mail, carrier, or other means:		
e. Samples, complimentary, and other free copies	48*	21**
F. Total distribution	1,405*	950**
G. Copies not distributed:		
1. Office use, leftover, unaccounted, spoiled after printing	102*	104*
2. Return from news agents	---	---
H. Total:	1,405*	950**

16. **I certify that the statements made by me above are correct and complete.**
Dianne Purrier, Managing Editor
Digital and Print Publications

* Average no. copies each issue during preceding 12 months

** Acutal no. copies of single issue published nearest filing date.

ONE CORE: SWITCH & ROUTE

ENX

Presented by Future at the NAB Show
BEST OF SHOW 2026
tvtech
WINNER

TVBEUROPE
BEST OF SHOW AT IBC 2025
WINNER

Evertz' ENX powers a unified media core that seamlessly connects live production, processing, and distribution across any infrastructure. Simplify operations, scale instantly, and deliver more personalized content from every live event.





www.evertz.com @EvertzTV @Evertz

evertz 60 YEARS 1966–2026