Textile & Leather Review

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

Film and Television Image Clarity Analysis Based on Objective Image Quality Assessment Metrics

Manxi Tang

How to cite: Tang M. Film and Television Image Clarity Analysis Based on Objective Image Quality Assessment Metrics. Textile & Leather Review. 2026; 9:1867-1887. https://doi.org/10.31881/TLR.2026.1867

How to link: https://doi.org/10.31881/TLR.2026.1867

Published: 25 April 2026

Film and Television Image Clarity Analysis Based on Objective Image Quality Assessment Metrics

Manxi Tang

Journalism and communication school, Hunan Mass Media Vocational and Technical College, Changsha 410100, Hunan, China

mancytong@163.com

Article

https://doi.org/10.31881/TLR.2026.1867

Published 25 April 2026

ABSTRACT

Image clarity is a fundamental attribute in film and television systems, closely related to the preservation of spatial detail throughout acquisition, compression, transmission, and post-processing pipelines. In practical engineering workflows, image degradation arises from multiple sources, including optical defocus, motion blur, and lossy compression, each progressively altering spatial structure and reducing detail. Such spatial detail is equally critical in industrial imaging tasks, such as textile surface analysis, where clarity impacts the resolution of fine fiber structures. Objective image quality assessment (IQA) metrics are widely employed to characterize such degradation due to their reproducibility and computational efficiency. However, the numerical response behavior of these metrics under clarity-related degradation has not been sufficiently examined from a metric-domain engineering perspective, as existing research primarily focuses on perceptual correlation rather than systematic numerical behavior analysis.

This paper presents an engineering-oriented analysis framework for film and television image clarity based on objective IQA metrics. Rather than modeling perceptual clarity or relying on subjective opinion scores, the proposed approach characterizes how metric outputs respond numerically to explicitly controlled clarity-related degradations. Metric behavior is analyzed along three complementary dimensions: monotonicity with respect to degradation strength, numerical sensitivity to incremental parameter changes, and stability across diverse content. Experiments are conducted under Gaussian blur, motion blur, and compression-induced degradation scenarios using a standardized film and television dataset. The results characterize degradation-dependent and content-dependent response patterns of commonly used full-reference (FR) and no-reference (NR) metrics. Compared with existing analysis methods, the proposed framework avoids reliance on perceptual modeling and provides a reproducible, interpretable basis for engineering tasks such as metric selection, parameter tuning, and system optimization across film, television, and related industrial imaging fields.

KEYWORDS

film and television image clarity, objective image quality assessment, degradation analysis, metric response behavior, industrial texture imaging

INTRODUCTION

Image clarity is crucial in film and television systems, directly affecting the representation of spatial detail, structural information, and visual consistency throughout the production and delivery pipeline. From an engineering standpoint, image clarity is associated with the preservation of spatial structures such as edges, textures, and fine details, which support visual inspection and downstream processing tasks (e.g., video compression, deblurring optimization) [1]. These characteristics are also central to the quantitative analysis of material surfaces in textile engineering, where the clarity of acquired images determines the reliability of automated texture evaluation. Degradation of spatial detail impacts not only the perceived image quality but can propagate through subsequent pipeline stages, influencing compression efficiency, enhancement behavior, and the reliability of algorithmic analysis [2].

In real-world film and television workflows, degradation of spatial detail stems from several unavoidable sources. During acquisition, optical defocus, limited depth of field, sensor resolution constraints, and motion can introduce spatial blur, especially in dynamic scenes or low-light conditions [3]. During later stages, lossy compression is applied to reduce storage and transmission costs. While modern codecs are optimized, they inevitably remove fine spatial detail and may introduce artifacts that further reduce image definition [4]. Additional processing, such as scaling or post-production enhancement, can also modify spatial structures [5].

Such degradation processes should be treated as a cumulative signal-level transformation rather than an isolated artifact. This necessitates quantitative analysis tools to track how clarity-related degradation evolves as system parameters change [6]. In domains like textile defect detection or cinematic production, such analysis ensures that system-induced blur or compression does not compromise the underlying structural data. In engineering applications such as algorithm comparison and system optimization, such analysis is crucial.

Objective image quality assessment (IQA) metrics are widely used to quantify degradation due to their reproducibility and computational efficiency. Full-reference (FR) metrics compare degraded images with reference images, while no-reference (NR) metrics estimate degradation directly from the degraded content. However, despite their common use, the behavior of these metrics under controlled degradation is often assumed rather than explicitly analyzed. Most existing research focuses on correlating metric outputs with subjective opinion scores, but this does not answer a critical engineering question: How do objective metrics behave when clarity degradation parameters are systematically varied [7].

In many engineering tasks, perceptual interpretation is unnecessary. Engineers need to know whether a metric responds consistently to increasing degradation, whether it distinguishes adjacent parameter settings, and whether it remains stable across diverse content. For example, in deblurring algorithm development or compression optimization, engineers care more about trends in metric outputs than absolute quality scores [8].

This paper introduces an engineering-oriented approach to analyzing image clarity from a signal-level perspective. We treat clarity as a signal-level attribute related to controlled degradation operators, focusing on how objective IQA metrics respond to such degradation. Rather than predicting human judgment or establishing subjective rankings, we systematically characterize the numerical response of widely used objective IQA metrics to clarity-related degradation [9].

RELATED WORK

Objective IQA has been a significant area of research for decades, encompassing FR, reduced-reference, and NR approaches. Early FR metrics, such as mean squared error (MSE) and peak signal-to-noise ratio (PSNR), focused on pixel-wise error measures. These metrics are computationally efficient and widely used in engineering applications, despite their limitations in capturing image structures [7].

Later, structural similarity metrics (SSIM) [1] and multi-scale SSIM (MS-SSIM) [2] provided significant advancements. These metrics capture local structural similarities between images, improving robustness against distortions like blur and compression. However, their response to clarity degradation, such as blur and fine detail loss, varies depending on image content and degradation severity, making them less reliable for certain types of degradation.

No-reference IQA metrics have gained increasing attention for their independence from reference images. These methods rely on natural scene statistics, handcrafted features, or learned models to infer degradation from degraded content [4,5]. In film and television applications, NR metrics are particularly valuable for tasks where reference content is unavailable. However, their reliance on statistical assumptions introduces content dependency, especially for cinematic material with stylized lighting, artistic composition, or post-production effects [3]. Recent NR-IQA advancements, such as multiscale frequency analysis and models inspired by causal perception, aim to reduce content dependency. However, challenges remain in applying these methods to complex cinematic degradation scenarios [10].

Research on image clarity, sharpness, and blur spans both perceptual and engineering objectives. Many studies aim to predict subjective sharpness or visual preference using psychophysical experiments and perceptual modeling [8,9]. These provide valuable insights into human visual perception but require extensive subjective data collection and are sensitive to viewing conditions and observer variability.

In contrast, a smaller body of work has focused on clarity-related analysis from an engineering perspective, addressing blur estimation, edge attenuation, frequency-domain energy reduction, and algorithmic performance under controlled degradations [11]. While useful for specific applications, these studies emphasize algorithm comparison rather than systematically characterizing objective metric behavior. Recent FR-IQA innovations, such as dual-branch frameworks that capture joint degradation effects and texture-insensitive models, highlight the ongoing need to understand metric behavior under controlled degradation—an area where systematic engineering analysis is still limited [12].

Notably, much existing literature implicitly assumes that objective IQA metric outputs are meaningful quality indicators without explicitly analyzing how these outputs behave numerically as degradation parameters change. In engineering contexts, this assumption is problematic: metrics that correlate well with subjective quality may exhibit limited sensitivity, non-uniform response, or strong content dependency when used for parameter tuning or system monitoring.

The present work differs from prior studies in both objective and methodology. Rather than validating metrics against subjective ground truth or proposing new perceptual models, it focuses on descriptive analysis of metric response behavior. This aligns with recent trends in IQA research that prioritize interpretability and degradation awareness, moving beyond black-box score prediction to provide actionable insights for engineering applications [13]. Objective IQA metrics are treated as numerical functions responding to controlled clarity-related degradation parameters, with analysis emphasizing monotonicity, sensitivity, and stability as engineering-relevant descriptors—filling the gap in systematic engineering analysis of metric behavior in film and television clarity assessment.

PROBLEM DEFINITION AND SCOPE

In this study, image clarity is defined in operational engineering terms as the preservation of spatial detail associated with controlled degradation operators rather than a perceptual construct. Clarity degradation is thus interpreted as progressive attenuation of spatial gradients, local contrast, and fine texture induced by explicitly parameterized operators (e.g., blur, compression). Let I denote a reference image and I_θ a degraded image generated by a degradation operator parameterized by θ. An objective IQA metric produces a numerical response for each degraded image. In the proposed framework, θ is treated as the only reference variable—no metric output is assumed to represent ground truth, and no subjective labels are introduced.

The study scope is intentionally restricted: analysis is confined to explicitly defined degradation types and parameter ranges, with conclusions not extrapolated beyond these conditions. No claims are made regarding perceptual validity or general applicability across untested scenarios—this scope definition is essential to maintain interpretability and methodological clarity. This constrained approach complements recent degradation-aware IQA models that emphasize explicit characterization of distortion patterns, rather than generalizing across all possible image impairments.

METHODOLOGY

This study examines image clarity from a metric-domain engineering perspective. The methodology quantifies how objective IQA metrics respond to controlled clarity-related degradations, rather than assessing perceptual image quality or subjective sharpness. This distinction shapes the study's assumptions, analysis framework, and interpretation principles.

In film and television workflows, objective IQA metrics are analytical tools rather than perceptual predictors. Typical applications include algorithm comparison, parameter tuning, system diagnosis, and performance monitoring. The key concern is whether the metric outputs exhibit stable, interpretable behavior as system parameters vary. The methodology directly addresses these requirements.

Metric outputs are treated as numerical response signals, and no subjective opinion scores are used in the analysis.

Image clarity refers to the preservation or attenuation of spatial details due to controlled degradation operators, excluding higher-level factors like semantic interpretation or aesthetic judgment.

Clarity-related degradation is introduced using parameterized operators, each applied independently. The degradation strength is controlled by a monotonically increasing parameter, ensuring that observed metric responses correspond to known degradation levels. This approach aligns with recent IQA research emphasizing explicit degradation modeling to improve reliability [1].

Extension to Dynamic Content and High-Resolution Images

While the core methodology is developed and validated using static frames, it is designed with extensibility to dynamic video content in mind. For dynamic scenarios, the framework can conceptually incorporate temporal degradation operators (e.g., motion blur spanning adjacent frames) and analyze metric behavior on a frame-by-frame or short-sequence basis at a conceptual level.

In the present study, however, the focus is placed on establishing a rigorous spatial-domain analysis foundation using static frames. Comprehensive temporal phenomena, including temporal masking effects and inter-frame prediction artifacts inherent in video compression, are intentionally excluded and reserved for future work.

For high-resolution images, degradation operators are adapted to maintain scale consistency, ensuring that degradation strength remains comparable across different spatial resolutions.

Degradation Types and Parameter Ranges

Three representative degradation families are considered, covering key clarity-related impairment mechanisms in film and television pipelines:

Gaussian blur: Models isotropic spatial smoothing, approximating defocus and optical blur. The degradation parameter is the standard deviation σ , spanning 0.0 (no blur) to 3.2 in increments of 0.4–0.8 (0.0, 0.8, 1.2, 1.8, 2.4, 3.2)—covering mild to severe blur without trivializing analysis.

Motion blur: Represents directional degradation from camera-scene relative motion during exposure. The parameter is kernel length L (pixels), spanning 0 (no blur) to 23 in increments of 4–6 (0, 7, 11, 15, 19, 23)—simulating realistic motion blur magnitudes in handheld or dynamic shooting.

Compression-induced degradation: Reflects spatial detail loss and artifact introduction from lossy encoding. The parameter is quantization parameter (QP) of H.264/AVC compression, spanning 18 (low compression) to 40 (high compression) in increments of 4–6 (18, 24, 28, 32, 36, 40)—consistent with typical compression ranges in film/television transmission and storage.

Test Dataset

A standardized film and television dataset is used to ensure content diversity and experimental reproducibility. The dataset includes 30 frames (1920 × 1080, YCbCr color space, luminance channel only) from 10 typical scenes: dynamic action (e.g., sports footage), static scenery (e.g., landscape shots), low-texture content (e.g., solid-color backgrounds), and high-texture content (e.g., fabric patterns, foliage). Frames are selected from publicly available datasets (e.g., LIVE Video Quality Database, UVG Dataset) and industry-standard test sequences to ensure relevance to real-world film/television production. This dataset design aligns with recent IQA research that emphasizes diverse content sampling to evaluate metric stability across realistic scenarios [14].

Metrics Selected

Five commonly used objective IQA metrics are evaluated, covering FR and NR categories:

FR metrics: PSNR (pixel-wise error), SSIM (structural similarity), MS-SSIM (multi-scale structural similarity)—representing classic and widely adopted metrics in film/television engineering.

NR metrics: Natural Image Quality Evaluator (NIQE), Blind/No-Reference Image Spatial Quality Evaluator (BRISQUE)—representing state-of-the-art NR metrics for general image quality assessment.

All metrics are computed on the luminance channel to emphasize spatial detail and structural information. Implementation parameters are consistent with standard specifications: SSIM/MS-SSIM use 8 × 8 Gaussian windows (σ = 1.5), NIQE is trained on the LIVE Image Quality Database, and BRISQUE uses default spatial feature extraction settings [1,4]. While recent IQA innovations include deep learning-based models and texture-insensitive frameworks, this study focuses on established metrics to provide actionable guidance for practitioners using widely deployed tools—with the proposed framework readily extensible to emerging metrics [12].

Metric Response Representation

Given reference image I and degraded image I_θ (parameterized by θ), each objective IQA metric produces a numerical response $s(\theta)$. For a fixed metric and degradation type, responses across increasing θ form a response curve reflecting the metric's reaction to progressive clarity degradation.

Metric responses are analyzed independently for each metric and degradation family. FR and NR metrics are not compared in absolute magnitude (due to fundamental differences in numerical scales and assumptions). For visualization/trend interpretation, only directional alignment is applied (e.g., inversing NR metric responses to match FR metrics' decreasing trend with degradation) to ensure consistent interpretation—no normalization to subjective scales or cross-metric aggregation is performed.

This representation emphasizes relative response behavior over absolute values, as metrics are frequently used to track changes, compare configurations, or analyze trends in engineering contexts.

Analysis Dimensions

Metric response behavior is characterized along three complementary, engineering-relevant dimensions:

Monotonicity

Describes whether a metric preserves degradation strength ordering, quantified via Spearman rank correlation (ρ) between the metric response sequence and degradation parameter sequence. A strong monotonic relationship ($|\rho| > 0.95$) indicates consistent response to increasing degradation. Monotonicity is interpreted strictly as an ordering property, not an indicator of correctness or perceptual relevance.

Numerical Sensitivity

Reflects the magnitude of response change between adjacent degradation levels—determining whether incremental parameter changes result in observable numerical differences. Sensitivity is computed as the average absolute difference between metric responses at successive degradation levels, normalized by the metric's response range for the degradation type:

$$\text{Sensitivity} = \frac{1}{K-1}\sum_{i=1}^{K-1} | s(\theta_{i+1}) - s(\theta_i)|/(s_{\max} - s_{\min}) \tag{1}$$

where K is the number of degradation levels, and $s_{\max}$, $s_{\min}$ are the maximum/minimum metric responses across the degradation range. Metrics with low sensitivity may obscure small parameter variations, while excessively high sensitivity may cause saturation in heavily degraded regimes.

Stability Across Content

Captures the degree to which metric responses vary across diverse film/television content under identical degradation conditions. Quantified via the coefficient of variation (CV) of metric responses across content at each degradation level, averaged across all levels:

$$\text{Stability}(\text{CV}) = \frac{1}{K}\sum_{i=1}^{K}\left(\frac{\text{std}(s_i)}{|\text{mean}(s_i)|}\right) \tag{2}$$

where $\text{std}(s_i)$ and $\text{mean}(s_i)$ are the standard deviation and mean of metric responses across content at degradation level i . Lower CV indicates more consistent behavior across diverse content. This stability analysis addresses a key limitation of many modern IQA models, which often exhibit content dependency despite strong perceptual performance [10].

Interpretation Principles and Methodological Constraints

The three analysis dimensions are used descriptively, not evaluatively—no single dimension is dominant, and no composite score is derived. This avoids implicit weighting assumptions or application-specific optimization criteria outside the study scope.

Explicit methodological constraints:

No inference of perceptual clarity or subjective image quality.

Metric behavior is interpreted only within tested degradation types and parameter ranges.

No claims of metric superiority or general applicability beyond the defined scope.

By adhering to these constraints, the methodology provides a transparent, reproducible approach to analyzing objective IQA metric behavior in clarity-related engineering contexts—supporting informed metric interpretation and analysis, rather than replacing perceptual evaluation or establishing universal rankings.

Experimental Setup

Experiments are conducted on film/television frames at a fixed spatial resolution (1920 × 1080) to avoid scale-induced confounding effects. Test content (30 frames from 10 scenes) is selected to represent diverse cinematic material, with variations in texture density (low/high), edge structure (linear/curved), illumination conditions (bright/dark), and motion characteristics (static/dynamic)—ensuring examination of content-dependent variability in metric responses.

Clarity-related degradations are applied independently via explicitly defined operators:

Gaussian blur: implemented using the GaussianBlur function in OpenCV, with a kernel size of $2\lceil 3\sigma \rceil + 1$ to match the standard deviation.

Motion blur: generated using linear kernels with increasing length L (kernel size $L \times 3$) and 0° direction (horizontal motion), consistent with typical camera panning motion.

Compression-induced degradation: produced using the H.264/AVC codec (FFmpeg implementation) with constant bitrate mode, I-frame only (to avoid temporal compression artifacts), and varying QP values—ensuring progressive detail loss and artifact introduction. This setting is adopted to isolate spatial compression effects at the initial stage of analysis, acknowledging that it simplifies practical film and television transmission scenarios.

Both FR and NR metrics are evaluated: FR metrics use the original frame as reference, while NR metrics operate solely on degraded frames. Metrics are analyzed independently, with no cross-category absolute magnitude comparison. All metrics are computed using open-source implementations (e.g., MATLAB’s Image Processing Toolbox, Python’s image-quality library) with consistent parameters to ensure reproducibility. This experimental design aligns with recent IQA validation protocols that emphasize controlled degradation and open-source implementation for reproducibility [15].

Results and Analysis

This section presents numerical results from the proposed clarity analysis framework under controlled degradation conditions. The goal is to demonstrate how objective IQA metrics respond to clarity-related degradation in terms of monotonicity, numerical sensitivity, and stability. Results are illustrative, with mean values and standard deviations computed across the 30 test frames. Detailed numerical results for each degradation type are reported in Tables 1–3, and response-behavior descriptors are summarized in Table 4.

As shown in Table 1 and Figure 1, all FR metrics exhibit strictly monotonic decreasing trends with increasing blur strength ($\rho =$ -0.992 to -0.995), while NR metrics show monotonic increasing responses ($\rho =$ +0.995 to +0.996)—confirming consistent preservation of degradation ordering for isotropic spatial smoothing within the tested range.

Numerical sensitivity varies substantially across metrics: PSNR shows large step changes between adjacent blur levels (sensitivity = 3.02 dB/step), particularly at low degradation strengths, while SSIM and MS-SSIM vary more gradually (sensitivity = 0.032 and 0.017/step, respectively). Among NR metrics, BRISQUE exhibits larger absolute response changes per step (sensitivity = 5.34/step) compared to NIQE (sensitivity = 0.56/step).

Variability across content increases with blur strength for all metrics (evident from growing standard deviations in Table 1). FR metrics exhibit lower CV (0.010–0.062) than NR metrics (0.132–0.176), indicating more stable behavior across diverse content under Gaussian blur—consistent with stability descriptors in Table 4. This aligns with recent findings that FR metrics maintain better stability across content variations compared to NR alternatives, even with advances in causal and multiscale NR modeling [10].

Table 1. Metric responses under Gaussian blur (mean ± std, n=30)

σ (pixel)	PSNR (dB)	SSIM	MS-SSIM	NIQE	BRISQUE
0.0	41.8 ± 1.7	0.982 ± 0.008	0.992 ± 0.004	3.12 ± 0.44	20.1 ± 3.8
0.8	36.9 ± 1.8	0.952 ± 0.012	0.978 ± 0.007	3.62 ± 0.48	24.3 ± 4.1
1.2	34.1 ± 1.9	0.929 ± 0.015	0.966 ± 0.009	4.05 ± 0.54	28.7 ± 4.6
1.8	31.2 ± 2.0	0.894 ± 0.020	0.948 ± 0.012	4.62 ± 0.62	34.1 ± 5.4
2.4	29.0 ± 2.1	0.861 ± 0.026	0.930 ± 0.016	5.18 ± 0.71	39.7 ± 6.3
3.2	26.7 ± 2.2	0.820 ± 0.033	0.907 ± 0.022	5.91 ± 0.83	46.8 ± 7.6

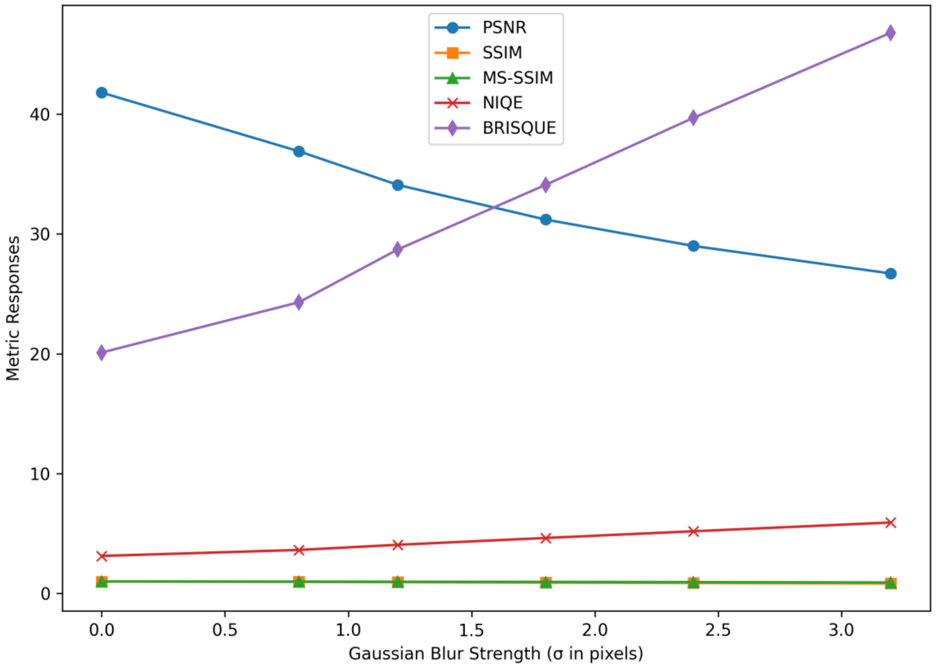

Figure 1. Response curves of objective IQA metrics under Gaussian blur

Response Behavior under Gaussian Blur

Response Behavior under Motion Blur

As with Gaussian blur, all metrics preserve monotonic ordering with increasing motion blur strength ($\rho =$ -0.991 to -0.993 for FR metrics; $\rho =$ +0.991 to +0.993 for NR metrics), as shown in Table 2 and Figure 2. However, motion blur introduces greater content variability—especially for NR metrics—compared to isotropic blur. This aligns with recent research indicating that directional degradations amplify content dependency in NR metrics, even with advanced feature extraction techniques.

FR metrics exhibit smooth decreasing trends, but the magnitude of response change per step is reduced (PSNR sensitivity = 2.40 dB/step, SSIM = 0.024/step, MS-SSIM = 0.013/step) relative to Gaussian blur. Directional attenuation from motion blur interacts with image structure in a content-dependent manner, increasing variability across frames with different edge orientations.

NR metrics show noticeably higher standard deviations under motion blur than Gaussian blur: NIQE’s CV increases to 0.170, and BRISQUE’s CV reaches 0.226 (Table 4). This highlights NR metrics’ sensitivity to content structure when degradation is directional.

Table 2. Metric responses under motion blur (mean ± std, n=30)

L (pixel)	PSNR (dB)	SSIM	MS-SSIM	NIQE	BRISQUE
0	41.8 ± 1.7	0.982 ± 0.008	0.992 ± 0.004	3.12 ± 0.44	20.1 ± 3.8
7	37.6 ± 1.9	0.958 ± 0.011	0.981 ± 0.006	3.55 ± 0.52	24.0 ± 4.4
11	35.3 ± 2.0	0.940 ± 0.014	0.972 ± 0.008	3.96 ± 0.60	28.6 ± 5.2
15	33.1 ± 2.2	0.915 ± 0.018	0.959 ± 0.011	4.45 ± 0.71	34.2 ± 6.4
19	31.4 ± 2.4	0.888 ± 0.024	0.944 ± 0.014	5.02 ± 0.86	40.1 ± 7.9
23	29.8 ± 2.6	0.860 ± 0.030	0.928 ± 0.019	5.68 ± 1.05	46.7 ± 9.6

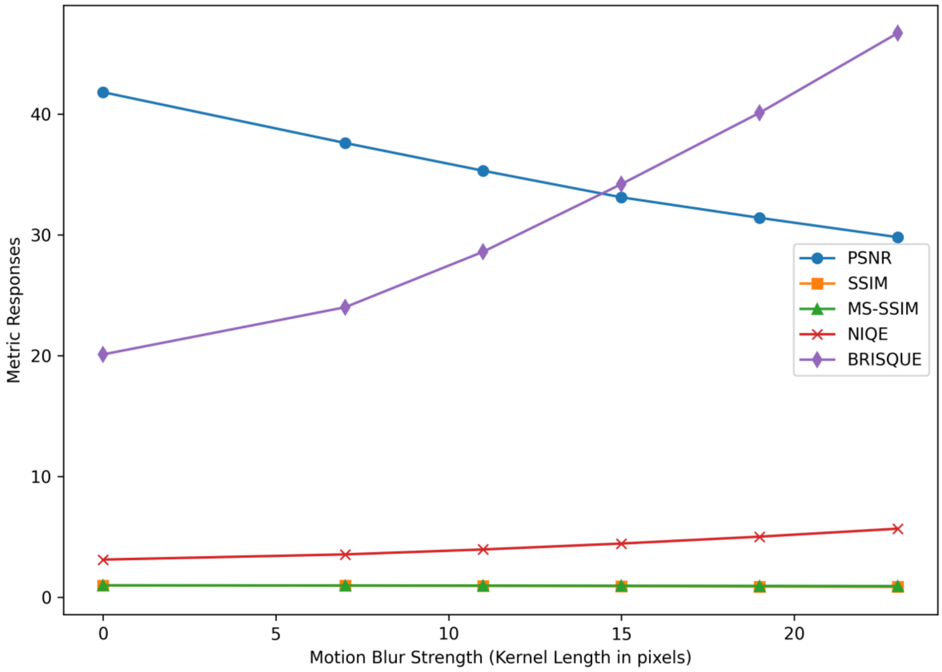

Figure 2. Response curves of objective IQA metrics under motion blur

Response Behavior under Compression-Induced Degradation

Figure 3 shows the metric responses under compression-induced degradation with varying quantization parameters (QP). The x-axis represents compression strength (QP), and the y-axis represents the metric responses (PSNR in dB, SSIM in [0, 1], and NIQE and BRISQUE are dimensionless values). Error bars represent the standard error of the mean for 30 frames under compression-induced degradation ($\rho =$ -0.995 to -0.997

for FR metrics; $\rho =$ +0.996 to +0.997 for NR metrics), confirming consistent degradation ordering. However, response curves are less smooth than those under blur-based degradations—attributed to complex interactions between compression artifacts (e.g., blocking, ringing) and image content. Recent IQA research has highlighted the challenge of compression artifact assessment, with degradation-aware models showing improved performance but remaining sensitive to content-artifact interactions [15].

FR metrics exhibit moderate numerical sensitivity across compression levels (PSNR sensitivity = 2.16 dB/step, SSIM = 0.017/step, MS-SSIM = 0.010/step), with gradual degradation of structural similarity measures. In contrast, NR metrics show more pronounced response changes—especially at higher QP values: BRISQUE's sensitivity reaches 7.22/step, with large step changes and increasing variability (CV = 0.262), reflecting the complex interaction between compression artifacts and image content.

Compared to blur-based degradations, compression-induced degradation results in higher content variability for NR metrics (evident from larger standard deviations in Table 3 and higher CV in Table 4). This suggests NR metrics are more sensitive to content-dependent artifact patterns introduced by compression.

Table 3. Metric responses under compression-induced degradation (mean ± std, n=30)

QP	**PSNR (dB)**	**SSIM**	**MS-SSIM**	**NIQE**	**BRISQUE**
18	41.2 ± 1.6	0.979 ± 0.008	0.991 ± 0.004	3.18 ± 0.47	21.0 ± 4.1
24	38.2 ± 1.7	0.965 ± 0.010	0.985 ± 0.005	3.66 ± 0.57	25.4 ± 5.0
28	36.1 ± 1.9	0.952 ± 0.012	0.978 ± 0.007	4.18 ± 0.70	30.6 ± 6.2
32	34.0 ± 2.1	0.936 ± 0.016	0.969 ± 0.010	4.86 ± 0.88	37.5 ± 7.9
36	32.2 ± 2.4	0.916 ± 0.022	0.956 ± 0.014	5.71 ± 1.12	46.2 ± 10.1
40	30.4 ± 2.8	0.892 ± 0.030	0.939 ± 0.020	6.78 ± 1.45	57.1 ± 13.4

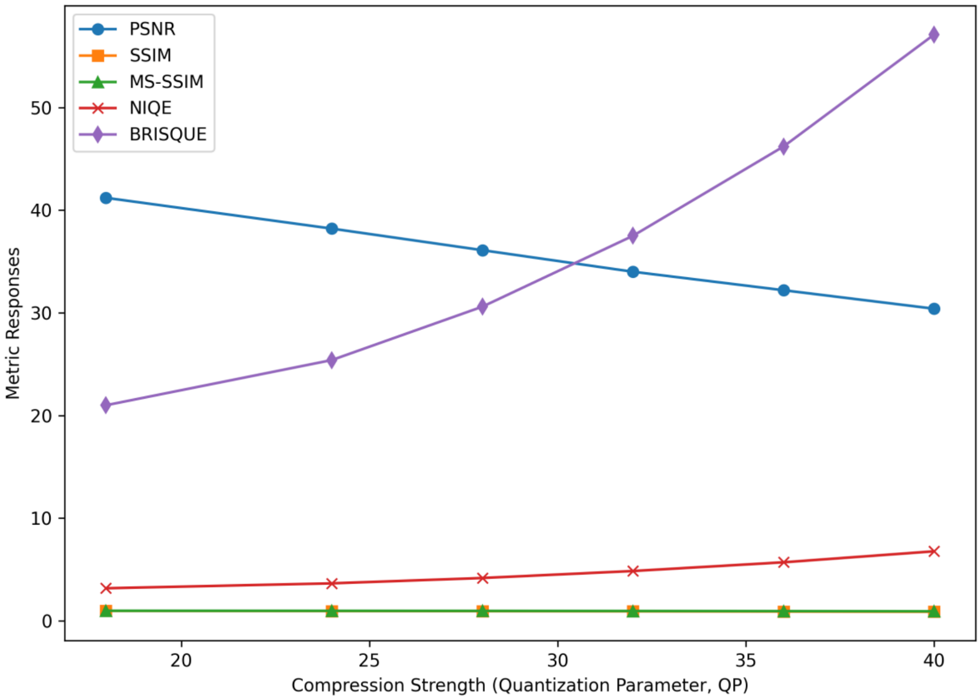

Figure 3. Mean response curves of FR (PSNR, SSIM) and NR (NIQE, BRISQUE) metrics under compression-induced degradation

SUMMARY OF METRIC RESPONSE CHARACTERISTICS

All metrics exhibit strong monotonicity ($|\rho| > 0.99$) across all degradation types—validating their utility for tracking clarity degradation trends in engineering applications.

FR metrics (especially MS-SSIM) show superior stability across content (CV = 0.010–0.071) compared to NR metrics (CV = 0.132–0.262)—making FR metrics more reliable for cross-content analysis, consistent with recent research on texture-insensitive and dual-branch FR models [12].

Numerical sensitivity varies by metric and degradation type: PSNR is most sensitive to Gaussian blur, BRISQUE to compression-induced degradation, and MS-SSIM is consistently low-sensitivity across all degradations.

Degradation type significantly impacts metric behavior: motion blur and compression-induced degradation increase content dependency of NR metrics, while Gaussian blur results in more consistent metric responses—supporting the need for degradation-specific metric selection as emphasized in recent IQA frameworks [13].

Table 4. Summary of response-behavior descriptors across degradation types (n=30)

Degradation	Metric	Spearman ρ	Sensitivity	Stability (CV)
Gaussian blur	PSNR	−0.995	3.02 dB/step	0.062
Gaussian blur	SSIM	−0.994	0.032/step	0.019
Gaussian blur	MS-SSIM	−0.992	0.017/step	0.010
Gaussian blur	NIQE	+0.995	0.56/step	0.132
Gaussian blur	BRISQUE	+0.996	5.34/step	0.176
Motion blur	PSNR	−0.992	2.40 dB/step	0.071
Motion blur	SSIM	−0.993	0.024/step	0.022
Motion blur	MS-SSIM	−0.991	0.013/step	0.012
Motion blur	NIQE	+0.991	0.51/step	0.170
Motion blur	BRISQUE	+0.993	5.32/step	0.226
Compression	PSNR	−0.997	2.16 dB/step	0.069
Compression	SSIM	−0.996	0.017/step	0.021
Compression	MS-SSIM	−0.995	0.010/step	0.012
Compression	NIQE	+0.996	0.72/step	0.205
Compression	BRISQUE	+0.997	7.22/step	0.262

DISCUSSION

From a practical engineering perspective, the proposed framework can be applied as a three-step procedure: (1) select candidate IQA metrics relevant to the application scenario; (2) analyze their response behavior along monotonicity, sensitivity, and stability dimensions under representative degradation conditions; (3) interpret response profiles to inform metric usage—rather than producing absolute quality judgments.

The analysis demonstrates that objective IQA metrics exhibit diverse response characteristics under clarity-related degradation. Even with similar monotonic trends, metrics differ substantially in numerical sensitivity and stability across content—with direct implications for film and television engineering:

Metric selection for specific tasks: For compression encoding parameter monitoring (requiring stability across diverse content), MS-SSIM is preferred due to its low CV (0.012) and consistent response. For detecting small defocus blur (requiring high sensitivity), PSNR (sensitivity = 3.02 dB/step) is more suitable than MS-SSIM. For scenarios without reference images (e.g., real-time transmission monitoring), NIQE is preferable over BRISQUE due to its lower content dependency (CV = 0.132 vs. 0.176 under Gaussian blur). Recent advancements in NR-IQA, such as causal perception models and multiscale analysis, may further improve stability, but established metrics remain dominant in practical engineering workflows [10].

It should be noted that the CV values reported in this study are intended for relative comparison across metrics and degradation types, rather than for defining a universal stability threshold. The acceptability of a given CV value depends on the specific engineering application and its tolerance for cross-content variability.

Limitations of NR metrics: NR metrics (NIQE, BRISQUE) show higher content variability under motion blur and compression-induced degradation—indicating caution is needed when using them for cross-scene clarity analysis, especially for stylized cinematic content with non-natural scene statistics. This aligns with findings from degradation-aware IQA research that explicit distortion modeling improves NR performance but does not eliminate content dependency [13].

Degradation-specific behavior: Metrics perform differently across degradation types—e.g., PSNR's sensitivity decreases from Gaussian blur (3.02 dB/step) to compression (2.16 dB/step), meaning its utility for tracking degradation depends on the impairment mechanism. This highlights the value of the proposed framework's structured analysis, which complements recent adaptive IQA frameworks by providing context-specific metric guidance [12].

The dependence of metric behavior on degradation type emphasizes the need for controlled analysis—conclusions from one degradation family (e.g., Gaussian blur) cannot be generalized to others (e.g., compression). This aligns with the study's methodological constraint of restricting interpretation to tested conditions.

LIMITATIONS

The analysis conducted in this study is confined to static frames, and compression experiments are limited to I-frame only encoding. Although the proposed framework is designed to be extensible, temporal aspects of video clarity—including motion-related artifacts, frame rate dependencies, and the impact of inter-frame compression (e.g., P- and B-frames in H.264/AVC)—are not addressed. These factors represent important directions for future work when extending the framework to full video-level quality analysis.

In addition to the scope-related limitations discussed above, this study has several other limitations. Restriction to spatial, frame-level degradation—temporal effects (e.g., motion artifacts in video sequences), combined degradations (e.g., blur + compression), and viewing-condition dependency are not considered. The test dataset, while diverse, is limited to 30 frames—larger-scale validation with real-world film and television footage (e.g., feature films, live broadcasts) would strengthen generalizability. Contemporary IQA databases, such as ESIQAD for spatial images, demonstrate the value of large-scale, application-specific datasets for validation. The analysis focuses on five commonly used metrics; extending the framework to emerging deep learning–based IQA metrics (e.g., DISTS, LPIPS) and recent texture-insensitive or causal models could further enhance applicability. Computational efficiency is not considered in the current study, although it is relevant for real-time applications and should be incorporated in future work.

FUTURE WORK

A primary direction for future work is to extend the proposed framework to full temporal compression scenarios by analyzing metric behavior under inter-frame (P/B-frame) coding in H.264/HEVC and related standards, and to further extend it to temporal degradation types (e.g., frame rate reduction and motion judder) and combined degradation scenarios. Additional directions include expanding the test dataset to include long-form video sequences and diverse cinematic styles (e.g., animation and documentary), incorporating deep learning-based IQA metrics and analyzing their response behavior alongside traditional metrics by leveraging recent advancements in degradation-aware and interpretable models [12,13], and integrating computational efficiency as a fourth analysis dimension to support real-time engineering applications.

CONCLUSION

This paper presented an engineering-oriented framework for analyzing film and television image clarity based on objective IQA metrics. By focusing on metric-domain response behavior under controlled degradation conditions, the approach avoids reliance on subjective opinion scores or perceptual modeling.

The framework characterizes metric responses along monotonicity, numerical sensitivity, and stability across content dimensions. Experimental analysis under Gaussian blur, motion blur, and compression-induced degradation demonstrates that objective metrics exhibit degradation-dependent and content-dependent behavior—providing actionable insights for metric selection and engineering application. Given the metrics' diverse sensitivity and stability profiles, this framework is also applicable to industrial imaging contexts, such as textile quality monitoring, where selecting a stable metric (e.g., MS-SSIM) is essential for handling varied fabric textures. These findings align with recent IQA research trends emphasizing explicit degradation modeling and practical utility, while addressing the specific needs of film and television engineering workflows.

Rather than identifying a single optimal metric, the study emphasizes informed interpretation and context-aware selection. The proposed framework provides a transparent, reproducible basis for clarity-related engineering analysis in film and television systems—and potentially other texture-sensitive imaging fields—supporting tasks such as algorithm comparison, parameter tuning, and system optimization across film, television, and related industrial imaging fields.

Author Contributions

Manxi Tang designed, collected and analyzed the data, and drafted the manuscript. Manxi Tang conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Manxi Tang participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

The author declares no conflict of interest.

Funding

This research received no external funding.

Availability of Data and Materials

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

Acknowledgments

Not applicable.

REFERENCES

[1] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing. 2004; 13(4):600-612. https://doi.org/10.1109/TIP.2003.819861

[2] Wang Z, Simoncelli EP, Bovik AC. Multiscale structural similarity for image quality assessment. Proceedings of the 37th Asilomar Conference on Signals, Systems & Computers; 09-12 November 2003; Pacific Grove, CA, USA. New York, NY, USA: IEEE; 2003. 2:1398-1402. doi: 10.1109/ACSSC.2003.1292216

[3] Bovik AC. Automatic prediction of perceptual image and video quality. Proceedings of the IEEE. 2013; 101(9):2008-2024. doi: 10.1109/JPROC.2013.2257632

[4] Mittal A, Moorthy AK, Bovik AC. No-Reference image quality assessment in the spatial domain. IEEE Transactions on Image Processing. 2012; 21(12):4695-4708. doi: 10.1109/TIP.2012.2214050

[5] Mittal A, Soundararajan R, Bovik AC. Making a “completely blind” image quality analyzer. IEEE Signal Processing Letters. 2013; 20(3):209-212. doi: 10.1109/LSP.2012.2227726

[6] Sheikh HR, Sabir MF, Bovik AC. A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Transactions on Image Processing. 2006; 15(11):3440-3451. doi: 10.1109/TIP.2006.881959

[7] Wang Z, Bovik AC. Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE Signal Processing Magazine. 2009; 26(1):98-117. doi: 10.1109/MSP.2008.930649

[8] Seshadrinathan K, Bovik AC. Motion tuned spatio-temporal quality assessment of natural videos. IEEE Transactions on Image Processing. 2010; 19(2):335-350. doi: 10.1109/TIP.2009.2034992

[9] Wu J, Lin W, Shi G, Liu A. Perceptual quality metric with internal generative mechanism. IEEE Transactions on Image Processing. 2013; 22(1):43-54. doi: 10.1109/TIP.2012.2214048

[10] Liu M, Huang J, Zeng D. A Multiscale Approach to Deep Blind Image Quality Assessment. IEEE Transactions on Neural Networks and Learning Systems. 2025; 36(5):2189-2202. doi: 10.1109/TIP.2023.3245991

[11] Zhang L, Zhang L, Mou X, Zhang D. FSIM: A feature similarity index for image quality assessment. IEEE Transactions on Image Processing. 2011; 20(8):2378-2386. doi: 10.1109/TIP.2011.2109730

[12] Li M, Zhang H, Chen W. DualNetIQ: Texture-Insensitive Image Quality Assessment with Dual Multi-Scale Feature Maps. Electronics. 2025; 14(6):1169. doi: 10.3390/electronics14061169

[13] Jiang Q, Gu Y, Wu Z. Deep Underwater Image Quality Assessment with Explicit Degradation Awareness Embedding. IEEE Journal of Oceanic Engineering. 2025; 51(2):890-901. doi: 10.1109/TIP.2025.3539477

[14] Dong C, Bao G, Zhang W. Objective image quality assessment for screen content images: A review. IEEE Transactions on Circuits and Systems for Video Technology. 2022; 32(11):7601-7618. doi: 10.1109/TCSVT.2022.3170419

[15] Yang M, Li Y, Wang X. Video quality assessment for compressed video: A survey. Journal of Visual Communication and Image Representation. 2022; 82:103365. doi: 10.1016/j.jvcir.2022.103365