

Image Dehazing System Using Dark Channel Prior and Deep Learning Approaches

Abstract

This thesis presents a comprehensive study and implementation of image dehazing techniques to address the challenges posed by haze, fog, and smoke in digital images. Haze significantly degrades image quality by reducing contrast and color fidelity, which adversely affects various computer vision applications including surveillance, autonomous driving, and remote sensing. This research explores two primary approaches to image dehazing: a physics-based method utilizing the Dark Channel Prior (DCP) with several enhancements, and a learning-based method employing a U-Net convolutional neural network architecture. The physics-based approach incorporates guided filtering for transmission map refinement and adaptive contrast enhancement techniques to overcome limitations of the original DCP algorithm. The deep learning approach leverages a dataset of paired hazy and clear images to train a model capable of directly mapping hazy inputs to their dehazed counterparts. Both methods are implemented in a user-friendly application with a modern graphical interface, allowing for real-time processing of both still images and video content. Extensive experimental results demonstrate that our enhanced DCP method achieves superior performance compared to the original algorithm, particularly in preserving edge details and natural colors. The deep learning model shows promising results with the potential for further improvement through additional training data and architectural refinements. This research contributes to the field by providing a comparative analysis of traditional and learning-based dehazing techniques, along with an accessible tool for practical applications in various domains including environmental monitoring, transportation safety, and computer vision preprocessing.

Table of Contents

- [List of Figures](#)
- [List of Tables](#)
- [List of Abbreviations](#)
- [Chapter 1: Introduction](#)
 - [1.1 Background](#)
 - [1.2 Problem Statement](#)
 - [1.3 Research Significance](#)
 - [1.4 Thesis Structure](#)
- [Chapter 2: Literature Review, Motivation, and Objective](#)
 - [2.1 Image Formation Model](#)
 - [2.2 Traditional Dehazing Methods](#)
 - [2.2.1 Dark Channel Prior](#)
 - [2.2.2 Color Attenuation Prior](#)
 - [2.2.3 Non-local Prior](#)
 - [2.2.4 Contrast Enhancement-Based Methods](#)
 - [2.2.5 Fusion-Based Methods](#)
 - [2.3 Learning-Based Dehazing Methods](#)
 - [2.3.1 CNN-Based Methods](#)
 - [2.3.2 GAN-Based Methods](#)
 - [2.3.3 Transformer-Based Methods](#)
 - [2.3.4 Physics-Guided Learning Methods](#)
 - [2.3.5 Unsupervised and Self-Supervised Methods](#)
 - [2.4 Evaluation Metrics](#)

- [2.4.1 Full-Reference Metrics](#)
- [2.4.2 No-Reference Metrics](#)
- [2.4.3 Human Visual Perception](#)
- [2.5 Challenges and Limitations](#)
 - [2.5.1 Scene Diversity and Complexity](#)
 - [2.5.2 Color Distortion and Artifacts](#)
 - [2.5.3 Computational Efficiency](#)
 - [2.5.4 Training Data Scarcity](#)
 - [2.5.5 Domain Gap and Generalization](#)
- [2.6 Research Motivation](#)
- [2.7 Research Objectives](#)
- [Chapter 3: Proposed System](#)
 - [3.1 System Overview](#)
 - [3.1.1 System Architecture](#)
 - [3.1.2 Component Interaction](#)
 - [3.1.3 Implementation Environment](#)
 - [3.2 Enhanced Dark Channel Prior Method](#)
 - [3.2.1 Dark Channel Computation](#)
 - [3.2.2 Atmospheric Light Estimation](#)
 - [3.2.3 Transmission Map Estimation](#)
 - [3.2.4 Guided Filtering for Transmission Refinement](#)
 - [3.2.5 Scene Radiance Recovery](#)
 - [3.2.6 Adaptive Contrast Enhancement](#)
 - [3.2.7 Sky Region Detection and Processing](#)
 - [3.3 Deep Learning-Based Dehazing Method](#)
 - [3.3.1 U-Net Architecture](#)
 - [3.3.2 Dataset Preparation](#)
 - [3.3.3 Data Augmentation](#)
 - [3.3.4 Training Process](#)
 - [3.3.5 Model Optimization](#)
 - [3.3.6 Loss Function Design](#)
 - [3.3.7 Inference Pipeline](#)
 - [3.4 Video Dehazing Extension](#)
 - [3.4.1 Frame Extraction and Processing](#)
 - [3.4.2 Temporal Consistency Preservation](#)
 - [3.4.3 Video Reconstruction](#)
 - [3.4.4 Performance Optimization](#)
 - [3.5 Graphical User Interface](#)
 - [3.5.1 Design Principles](#)
 - [3.5.2 Implementation Details](#)
 - [3.5.3 User Interaction Flow](#)
 - [3.5.4 Visualization Features](#)
 - [3.5.5 Parameter Control Panel](#)
- [Chapter 4: Results and Discussion](#)
 - [4.1 Experimental Setup](#)
 - [4.1.1 Hardware and Software Configuration](#)

- [4.1.2 Dataset Description](#)
 - [4.1.3 Evaluation Methodology](#)
 - [4.1.4 Implementation Details](#)
- [4.2 Qualitative Analysis](#)
 - [4.2.1 Visual Comparison of Dehazing Methods](#)
 - [4.2.2 Case Studies](#)
 - [4.2.3 Visual Assessment by Expert Panel](#)
- [4.3 Quantitative Analysis](#)
 - [4.3.1 Image Quality Metrics](#)
 - [4.3.2 Computational Efficiency](#)
 - [4.3.3 Statistical Significance of Results](#)
- [4.4 Ablation Studies](#)
 - [4.4.1 Impact of Parameter Selection](#)
 - [4.4.2 Contribution of Individual Components](#)
 - [4.4.3 Network Architecture Variations](#)
- [4.5 Comparative Analysis](#)
 - [4.5.1 Comparison with State-of-the-Art Methods](#)
 - [4.5.2 Strengths and Limitations](#)
 - [4.5.3 Performance Across Different Haze Densities](#)
 - [4.5.4 Performance on Various Scene Types](#)
- [4.6 User Experience Evaluation](#)
 - [4.6.1 Survey Design and Methodology](#)
 - [4.6.2 Usability Assessment Results](#)
 - [4.6.3 User Feedback Analysis](#)
- [Chapter 5: Conclusion and Future Scope](#)
 - [5.1 Summary of Contributions](#)
 - [5.2 Limitations of the Current Work](#)
 - [5.3 Future Research Directions](#)
 - [5.3.1 Algorithm Improvements](#)
 - [5.3.2 Application Extensions](#)
 - [5.3.3 Integration with Other Computer Vision Tasks](#)
 - [5.3.4 Hardware Acceleration and Optimization](#)
 - [5.3.5 Multi-Modal Dehazing Approaches](#)
 - [5.4 Concluding Remarks](#)
- [References](#)
- [Appendices](#)
 - [Appendix A: Code Implementation](#)
 - [Appendix B: Additional Experimental Results](#)
 - [Appendix C: User Manual](#)
 - [Appendix D: Dataset Details](#)
 - [Appendix E: Mathematical Derivations](#)

List of Figures

1. Fig. 1.1: Examples of hazy images and their impact on visibility

2. Fig. 1.2: Applications affected by haze in images
3. Fig. 1.3: Comparison of clear and hazy images in autonomous driving scenarios
4. Fig. 1.4: Impact of haze on surveillance system effectiveness
5. Fig. 2.1: Atmospheric scattering model illustration
6. Fig. 2.2: Dark channel prior examples
7. Fig. 2.3: Comparison of traditional dehazing methods
8. Fig. 2.4: Evolution of learning-based dehazing approaches
9. Fig. 2.5: Contrast enhancement-based dehazing results
10. Fig. 2.6: Fusion-based dehazing framework
11. Fig. 2.7: CNN architecture evolution for image dehazing
12. Fig. 2.8: GAN-based dehazing model structure
13. Fig. 2.9: Transformer architecture for image dehazing
14. Fig. 2.10: Physics-guided neural network approach
15. Fig. 2.11: Unsupervised dehazing framework
16. Fig. 3.1: System architecture overview
17. Fig. 3.2: Component interaction diagram
18. Fig. 3.3: Dark channel computation process
19. Fig. 3.4: Atmospheric light estimation visualization
20. Fig. 3.5: Transmission map before and after refinement
21. Fig. 3.6: Guided filtering process illustration
22. Fig. 3.7: Contrast enhancement effect demonstration
23. Fig. 3.8: Sky region detection and processing
24. Fig. 3.9: U-Net architecture for dehazing
25. Fig. 3.10: Dataset preparation pipeline
26. Fig. 3.11: Data augmentation examples
27. Fig. 3.12: Training loss curves
28. Fig. 3.13: Model optimization process
29. Fig. 3.14: Loss function comparison
30. Fig. 3.15: Inference pipeline diagram
31. Fig. 3.16: Video dehazing framework
32. Fig. 3.17: Temporal consistency preservation
33. Fig. 3.18: GUI design and components
34. Fig. 3.19: User interaction workflow
35. Fig. 3.20: Visualization features demonstration
36. Fig. 4.1: Visual comparison on synthetic hazy images
37. Fig. 4.2: Visual comparison on real-world hazy images
38. Fig. 4.3: Case study: Extreme haze conditions
39. Fig. 4.4: Case study: Night-time hazy scenes
40. Fig. 4.5: Case study: Urban environments with complex structures
41. Fig. 4.6: Expert visual assessment results
42. Fig. 4.7: Performance metrics comparison chart
43. Fig. 4.8: Computation time comparison
44. Fig. 4.9: Statistical significance analysis
45. Fig. 4.10: Effect of patch size parameter
46. Fig. 4.11: Effect of omega parameter
47. Fig. 4.12: Component contribution analysis
48. Fig. 4.13: Network architecture variations comparison
49. Fig. 4.14: Comparison with state-of-the-art methods
50. Fig. 4.15: Performance across different haze densities
51. Fig. 4.16: Performance on various scene types
52. Fig. 4.17: User experience survey results
53. Fig. 5.1: Future research directions overview

List of Tables

1. Table 2.1: Summary of traditional dehazing methods
2. Table 2.2: Summary of learning-based dehazing methods
3. Table 2.3: Comparison of evaluation metrics
4. Table 2.4: Challenges in image dehazing and proposed solutions
5. Table 3.1: System components and their functions
6. Table 3.2: Enhanced DCP algorithm parameters
7. Table 3.3: U-Net architecture details
8. Table 3.4: Dataset statistics
9. Table 3.5: Data augmentation techniques
10. Table 3.6: Training hyperparameters
11. Table 3.7: Loss function components
12. Table 3.8: Video processing parameters
13. Table 3.9: GUI components and their functionality
14. Table 4.1: Hardware specifications
15. Table 4.2: Software dependencies
16. Table 4.3: Synthetic dataset characteristics
17. Table 4.4: Real-world dataset characteristics
18. Table 4.5: Evaluation metrics and their significance
19. Table 4.6: Quantitative results on synthetic dataset
20. Table 4.7: Quantitative results on real-world dataset
21. Table 4.8: Computation time for different image resolutions
22. Table 4.9: Statistical analysis of performance metrics
23. Table 4.10: Ablation study results
24. Table 4.11: Network architecture variations performance
25. Table 4.12: Comparison with state-of-the-art methods
26. Table 4.13: Performance across different haze densities
27. Table 4.14: Performance on various scene types
28. Table 4.15: User experience survey results
29. Table 4.16: Expert feedback summary
30. Table 5.1: Summary of contributions
31. Table 5.2: Limitations and potential solutions
32. Table 5.3: Future research directions

List of Abbreviations

- **CLAHE**: Contrast Limited Adaptive Histogram Equalization
- **CNN**: Convolutional Neural Network
- **DCP**: Dark Channel Prior
- **FADE**: Fog Aware Density Evaluator
- **FCN**: Fully Convolutional Network
- **GAN**: Generative Adversarial Network
- **GPU**: Graphics Processing Unit
- **GUI**: Graphical User Interface
- **HDR**: High Dynamic Range
- **HVS**: Human Visual System
- **IoU**: Intersection over Union
- **LIDAR**: Light Detection and Ranging
- **MAE**: Mean Absolute Error
- **MSE**: Mean Squared Error
- **NIQE**: Natural Image Quality Evaluator

- **NLP**: Non-Local Prior
- **PSNR**: Peak Signal-to-Noise Ratio
- **RGB**: Red, Green, Blue
- **ROI**: Region of Interest
- **SSIM**: Structural Similarity Index Measure
- **SVD**: Singular Value Decomposition
- **UI**: User Interface
- **UX**: User Experience
- **VGG**: Visual Geometry Group (neural network)

Chapter 1: Introduction

1.1 Background

Digital images have become an integral part of modern society, serving critical roles in various applications including surveillance, autonomous driving, remote sensing, and medical imaging. However, the quality of these images can be significantly degraded by atmospheric conditions such as haze, fog, and smoke. These phenomena introduce a semi-transparent layer that reduces contrast, fades colors, and obscures details in the captured scenes.

Haze formation is primarily caused by atmospheric particles that absorb and scatter light rays traveling from the scene to the camera. This scattering process follows the atmospheric scattering model, which describes how light interacts with particles in the atmosphere. The visual degradation becomes more severe as the distance between the camera and the scene increases, creating a depth-dependent effect that makes distant objects appear more hazy than closer ones.

The presence of haze in images poses significant challenges for both human perception and computer vision systems. For humans, hazy images provide less visual information, making it difficult to identify objects and assess scenes accurately. For computer vision algorithms, haze reduces the effectiveness of feature extraction, object detection, and scene understanding processes, potentially leading to system failures in critical applications like autonomous driving or surveillance.

The impact of haze on computer vision systems is particularly concerning in safety-critical applications. For instance, in autonomous driving, haze can significantly reduce the visibility range of both cameras and LIDAR sensors, making it difficult to detect pedestrians, vehicles, and road signs at a safe distance. Similarly, in surveillance systems, haze can obscure important details such as facial features or license plate numbers, rendering the surveillance footage less useful for security and forensic purposes.

In remote sensing and satellite imaging, haze can obscure ground features and affect the accuracy of land cover classification, vegetation monitoring, and urban planning applications. In medical imaging, while not affected by atmospheric haze, similar degradation effects can occur due to tissue properties, reducing the diagnostic value of the images.

The economic impact of haze-related visibility issues is substantial. According to transportation safety studies, reduced visibility conditions contribute to thousands of accidents annually, resulting in significant human and economic costs. In the aviation industry, haze-related flight delays and cancellations cost millions of dollars each year. For outdoor surveillance systems, the effectiveness of security infrastructure worth billions of dollars is compromised during hazy conditions.

Consequently, image dehazing has emerged as an important research area in computer vision and image processing. The goal of image dehazing is to recover a clear, haze-free image from its hazy counterpart, effectively reversing the degradation caused by atmospheric conditions. This process involves estimating and removing the haze component while preserving the original scene details and natural appearance.

Over the past two decades, numerous approaches have been proposed for image dehazing, ranging from traditional physics-based methods to modern learning-based techniques. Physics-based methods typically rely on the atmospheric scattering model and various priors or assumptions about the statistical properties of haze-free images. Learning-based methods, on the other hand, leverage deep neural networks to learn the mapping from hazy to clear images directly from data, without explicitly modeling the physical process of haze formation.

Despite significant progress, image dehazing remains a challenging problem due to its ill-posed nature and the complexity of real-world haze distributions. The development of effective, efficient, and robust dehazing algorithms continues to be an active area of research, with important implications for various applications that rely on high-quality visual information.

1.2 Problem Statement

Despite significant advances in image dehazing techniques over the past decade, several challenges remain unresolved. The primary challenge lies in accurately estimating the atmospheric light and transmission map, which are essential components of the atmospheric scattering model used for haze removal. Traditional methods often rely on prior assumptions about the statistical properties of haze-free images, which may not hold true for all scenes and conditions.

The Dark Channel Prior (DCP), proposed by He et al., has been widely adopted due to its effectiveness and simplicity. However, it suffers from several limitations, including color distortion in sky regions, halo artifacts around depth discontinuities, and high computational complexity. While various improvements have been proposed to address these issues, a comprehensive solution that maintains a balance between effectiveness and efficiency is still needed.

More recently, learning-based approaches using deep neural networks have shown promising results by directly learning the mapping from hazy to clear images. However, these methods typically require large amounts of training data consisting of paired hazy and clear images, which are difficult to obtain in real-world scenarios. Additionally, models trained on synthetic data often struggle to generalize to real-world hazy conditions due to the domain gap between synthetic and real haze distributions.

Furthermore, most existing dehazing methods focus on processing still images, with limited attention given to video dehazing, which introduces additional challenges related to temporal consistency and real-time processing requirements. The development of a comprehensive dehazing system that can effectively process both images and videos while maintaining visual quality and computational efficiency remains an open research problem.

The specific problems addressed in this thesis can be summarized as follows:

1. **Limitations of Traditional Methods:** Existing physics-based methods like DCP often make simplifying assumptions that do not hold for all scenes, leading to artifacts and quality issues in the dehazed images. These methods also tend to be computationally expensive, limiting their practical applicability.
2. **Data Requirements for Learning-Based Methods:** Deep learning approaches require large amounts of paired training data (hazy and clear image pairs), which are difficult to collect in real-world settings. Creating synthetic haze often does not accurately represent the complexity of natural haze distributions.
3. **Generalization to Diverse Scenes:** Both traditional and learning-based methods often struggle with diverse scene types, such as images containing large sky regions, night-time scenes, or scenes with non-uniform haze distribution.
4. **Computational Efficiency vs. Quality Trade-off:** There is a persistent trade-off between dehazing quality and computational efficiency. High-quality dehazing often requires complex algorithms or deep networks that are computationally intensive, making real-time processing challenging.
5. **Temporal Consistency in Videos:** Applying image dehazing algorithms to video frames independently often results in temporal inconsistencies, such as flickering or jittering effects, which degrade the visual quality of the dehazed video.
6. **Usability and Accessibility:** Many existing dehazing algorithms are implemented as research prototypes without user-friendly interfaces, making them inaccessible to non-expert users who could benefit from dehazing technology.
7. **Objective Evaluation:** The lack of standardized evaluation protocols and the subjective nature of image quality assessment make it difficult to fairly compare different dehazing methods and measure progress in the field.

Addressing these challenges requires a multifaceted approach that combines the strengths of traditional physics-based methods and modern learning-based techniques, while also considering practical aspects such as computational efficiency, user interface design, and evaluation methodology.

1.3 Research Significance

This research addresses the aforementioned challenges by developing an enhanced image dehazing system that combines the strengths of traditional physics-based methods and modern learning-based approaches. The significance of this work lies in several aspects:

1. **Improved Visual Quality:** By enhancing the Dark Channel Prior method with guided filtering for transmission map refinement and adaptive contrast enhancement techniques, our approach achieves superior dehazing results with reduced artifacts and better preservation of natural colors and details. This improvement in visual quality directly benefits applications that rely on high-quality images, such as surveillance, remote sensing, and medical imaging.

2. **Complementary Approaches:** The integration of both physics-based and learning-based methods provides a comprehensive solution that can leverage the interpretability and theoretical foundation of traditional approaches while benefiting from the data-driven capabilities of deep learning models. This hybrid approach offers greater robustness across diverse scenes and haze conditions compared to either approach alone.
3. **Practical Applicability:** The development of a user-friendly graphical interface makes advanced dehazing techniques accessible to non-expert users, enabling practical applications in various domains including photography enhancement, surveillance system improvement, and pre-processing for other computer vision tasks. This bridges the gap between academic research and real-world deployment of dehazing technology.
4. **Video Processing Capability:** The extension of our dehazing methods to video content addresses the growing need for temporal processing in applications such as autonomous driving, drone surveillance, and outdoor monitoring systems. By ensuring temporal consistency in the dehazed video, our system provides a more complete solution for real-world scenarios where video is the primary data format.
5. **Educational Value:** The comprehensive analysis and comparison of different dehazing techniques provide valuable insights for researchers and practitioners in the field, contributing to a better understanding of the strengths and limitations of various approaches. This knowledge can guide future research and development efforts in image dehazing.
6. **Benchmark for Evaluation:** Our extensive experimental evaluation using both synthetic and real-world datasets establishes a benchmark for assessing the performance of dehazing algorithms. The combination of quantitative metrics and qualitative assessments provides a more comprehensive evaluation framework that can be adopted by the research community.
7. **Cross-Domain Applications:** The improved image quality resulting from our dehazing system can benefit a wide range of downstream applications beyond the immediate scope of dehazing. These include object detection and recognition, scene understanding, image segmentation, and 3D reconstruction, all of which rely on clear, high-quality images as input.
8. **Economic and Safety Impact:** By improving visibility in hazy conditions, our research has potential economic and safety implications. Enhanced visibility in transportation systems can reduce accidents and associated costs. Improved surveillance in adverse weather conditions can enhance security effectiveness. Better remote sensing imagery can lead to more accurate environmental monitoring and disaster response.

By addressing these aspects, this research contributes to advancing the state of the art in image dehazing and provides practical solutions for real-world applications affected by atmospheric degradation. The combination of theoretical advancements, algorithmic improvements, and practical implementation makes this work relevant to both the academic community and industry practitioners dealing with image quality issues in hazy conditions.

1.4 Thesis Structure

The remainder of this thesis is organized as follows:

Chapter 2: Literature Review, Motivation, and Objective provides a comprehensive review of existing image dehazing methods, including traditional approaches based on physical models and learning-based methods using deep neural networks. It also discusses evaluation metrics, challenges, and limitations in the field, leading to the motivation and specific objectives of this research.

The literature review begins with an explanation of the atmospheric scattering model that forms the theoretical foundation for most dehazing methods. It then categorizes and analyzes traditional dehazing approaches, including the Dark Channel Prior, Color Attenuation Prior, and Non-local Prior, highlighting their strengths and limitations. The review continues with an examination of learning-based methods, covering CNN-based approaches, GAN-based methods, and emerging transformer architectures.

The chapter also discusses evaluation methodologies, including full-reference metrics like PSNR and SSIM, no-reference metrics like NIQE and FADE, and the importance of human visual assessment. It identifies key challenges in the field, such as scene diversity, color distortion, computational efficiency, and the scarcity of training data. Based on this analysis, the chapter presents the motivation for this research and outlines specific objectives aimed at addressing the identified gaps and limitations.

Chapter 3: Proposed System presents the detailed design and implementation of our image dehazing system. It describes the enhanced Dark Channel Prior method with guided filtering and contrast enhancement, the U-Net-based deep learning approach, the video dehazing extension, and the graphical user interface development.

The chapter begins with an overview of the system architecture, explaining how different components interact to provide a comprehensive dehazing solution. It then delves into the technical details of the enhanced DCP method, including dark channel computation, atmospheric light estimation, transmission map refinement using guided filtering, scene radiance recovery, and adaptive contrast enhancement. The implementation details, including parameter selection and optimization strategies, are provided with supporting code snippets and illustrations.

The chapter continues with a description of the deep learning-based approach, detailing the U-Net architecture, dataset preparation, data augmentation techniques, training process, and model optimization. It explains how the video dehazing extension handles temporal consistency while maintaining computational efficiency. Finally, it presents the design principles and implementation details of the graphical user interface, focusing on user experience, visualization features, and interaction flow.

Chapter 4: Results and Discussion evaluates the performance of the proposed system through extensive experiments on both synthetic and real-world datasets. It includes qualitative and quantitative analyses, ablation studies to assess the contribution of individual components, and comparisons with state-of-the-art methods.

The chapter begins by describing the experimental setup, including hardware and software configurations, dataset characteristics, and evaluation methodology. It presents a qualitative analysis through visual comparisons and case studies, highlighting the performance of our methods in different scenarios such as extreme haze conditions and night-time scenes. The quantitative analysis includes performance metrics such as PSNR, SSIM, and computation time, with statistical analysis to establish the significance of the results.

The chapter includes ablation studies that investigate the impact of different parameters and the contribution of individual components to the overall performance. It provides a comprehensive comparison with state-of-the-art methods, analyzing the strengths and limitations of our approaches in relation to existing techniques. The chapter concludes with a user experience evaluation that assesses the practical usability of our system through surveys and feedback analysis.

Chapter 5: Conclusion and Future Scope summarizes the main contributions of this research, acknowledges its limitations, and suggests directions for future work to further advance image dehazing techniques and their applications.

The chapter begins with a concise summary of the key contributions, highlighting the improvements achieved in visual quality, computational efficiency, and practical applicability. It acknowledges the limitations of the current work, including challenges related to extreme conditions, parameter sensitivity, and the need for more diverse training data. The chapter then explores future research directions, including algorithmic improvements, application extensions, integration with other computer vision tasks, hardware acceleration, and multi-modal approaches. It concludes with final remarks on the significance of image dehazing in the broader context of computer vision and its potential impact on various applications.

The thesis concludes with a comprehensive list of references and appendices containing additional technical details, code implementations, supplementary experimental results, user manual, dataset details, and mathematical derivations. These appendices provide valuable resources for researchers and practitioners interested in implementing or extending our dehazing system.

Chapter 2: Literature Review, Motivation, and Objective

2.1 Image Formation Model

To understand image dehazing algorithms, it is essential to first comprehend how haze affects images. The widely accepted atmospheric scattering model, proposed by Koschmieder (1924) and later refined by Narasimhan and Nayar (2002), describes the formation of a hazy image $I(x)$ at pixel position x can be expressed as:

$$I(x) = J(x)t(x) + A(1 - t(x))$$

where:

- $J(x)$ where:
- $J(x)$ is the scene radiance (haze-free image) that we aim to recover
- A is the global atmospheric light
- $t(x)$ is the transmission map, representing the portion of light that reaches the camera without being scattered

The transmission $t(x)$ is related to the scene depth $d(x)$ by:

$$t(x) = e^{-\beta d(x)}$$

where β is the atmospheric scattering coefficient.

This model indicates that a hazy image consists of two components: the attenuated scene radiance ($J(x)t(x)$) and the atmospheric light contribution ($A(1 - t(x))$). As the scene depth increases, the transmission decreases exponentially, causing distant objects to be more affected by haze.

The physical basis of this model lies in the interaction between light and atmospheric particles. When light travels through the atmosphere, it encounters various particles such as water droplets, dust, and pollutants. These particles cause two primary effects: attenuation and airlight.

Attenuation refers to the weakening of light from the scene to the camera due to absorption and scattering. This is represented by the term $J(x)t(x)$ in the model. The transmission $t(x)$ quantifies how much of the original scene radiance reaches the camera without being scattered away from the line of sight.

Airlight, represented by $A(1 - t(x))$, is the atmospheric light scattered towards the camera by the same particles. This scattered light adds a haze veil to the image, reducing contrast and saturation. The atmospheric light A is often assumed to be the color of the most haze-opaque region in the image, typically corresponding to the sky or distant objects completely obscured by haze.

The goal of image dehazing is to estimate the atmospheric light A and the transmission map $t(x)$ from the hazy image $I(x)$, and then recover the scene radiance $J(x)$ using the inverse of the atmospheric scattering model:

$$J(x) = (I(x) - A) / \max(t(x), t_0) + A$$

where t_0 is a small constant (typically 0.1) to prevent division by zero and to avoid noise amplification in regions with very low transmission.

However, this is an ill-posed problem since both A and $t(x)$ are unknown, and there are infinitely many combinations of J , A , and t that can produce the same hazy image I . Therefore, additional priors or constraints are needed to make the problem solvable.

The atmospheric scattering model makes several assumptions that may not always hold in real-world scenarios:

1. **Homogeneous Atmosphere:** The model assumes that atmospheric conditions are uniform throughout the scene, which may not be true for scenes with non-uniform haze distribution.
2. **Single Scattering:** The model considers only first-order scattering and ignores multiple scattering events, which become significant in dense haze conditions.
3. **Constant Atmospheric Light:** The model assumes a global atmospheric light value, whereas in reality, the atmospheric light may vary spatially, especially in scenes with multiple light sources.
4. **Lambertian Surfaces:** The model assumes that scene surfaces are Lambertian (perfectly diffuse), which is not always the case for specular or transparent objects.

Despite these limitations, the atmospheric scattering model provides a solid theoretical foundation for most dehazing algorithms and has proven effective in practice for a wide range of scenes and conditions.

2.2 Traditional Dehazing Methods

2.2.1 Dark Channel Prior

The Dark Channel Prior (DCP), proposed by He et al. (2009), is one of the most influential traditional dehazing methods. It is based on the observation that in most non-sky regions of haze-free outdoor images, at least one color channel has very low intensity at some pixels. This observation leads to the definition of the dark channel of an image J as:

$$J^{\text{dark}}(x) = \min_{y \in \Omega(x)} (\min_{c \in \{r, g, b\}} J^c(y))$$

where $\Omega(x)$ is a local patch centered at x , and J^c is a color channel of J .

The dark channel prior states that for haze-free outdoor images, the dark channel tends to have very low values (close to zero) except in sky regions. This prior is based on the observation that natural outdoor scenes contain abundant shadows, dark objects, or surfaces with low reflectance in at least one color channel.

The statistical basis for this prior was established through an extensive study of outdoor haze-free images. He et al. analyzed over 3,000 outdoor images and found that in more than 75% of the non-sky pixels, the dark channel value was below 25 (out of 255), confirming the validity of this prior for a wide range of natural scenes.

Based on this prior, He et al. derived a method to estimate the transmission map:

$$t(x) = 1 - \omega \min_{y \in \Omega(x)} (\min_{c \in \{r, g, b\}} (I^c(y)/A^c))$$

where ω ($0 < \omega \leq 1$) is a parameter to control the amount of haze removal, typically set to 0.95 to preserve a small amount of haze for distant objects, maintaining depth perception.

The atmospheric light A is estimated by finding the brightest pixels in the dark channel. Specifically, the method selects the top 0.1% brightest pixels in the dark channel and then picks the pixel with the highest intensity in the original image among these candidates.

The DCP algorithm can be summarized in the following steps:

1. Compute the dark channel of the hazy image
2. Estimate the atmospheric light A using the dark channel
3. Estimate the transmission map $t(x)$ using the dark channel prior
4. Refine the transmission map using soft matting or guided filtering
5. Recover the scene radiance $J(x)$ using the atmospheric scattering model

The DCP method has several advantages:

1. **Effectiveness:** It can effectively remove haze from a wide range of outdoor images without requiring multiple images or additional information.
2. **Simplicity:** The algorithm is relatively simple to implement and understand, making it accessible for various applications.
3. **Physical Basis:** It is grounded in the atmospheric scattering model, providing a physical interpretation of the dehazing process.

However, the DCP method also has several limitations:

1. **Sky Regions:** The dark channel prior does not hold for sky regions, leading to color distortions and over-enhancement in these areas.
2. **Halo Artifacts:** The patch-based operations can create halo artifacts around depth discontinuities, especially when using large patch sizes.
3. **Computational Complexity:** The original implementation using soft matting for transmission refinement is computationally expensive, although this has been addressed in later work using guided filtering.
4. **Parameter Sensitivity:** The performance of the algorithm is sensitive to the patch size and the ω parameter, requiring careful tuning for optimal results.
5. **Color Distortion:** In some cases, the method can produce color distortions, particularly in regions with bright colors or when the atmospheric light is incorrectly estimated.

Despite these limitations, the DCP method has been widely adopted and has inspired numerous improvements and extensions, making it a cornerstone in the field of image dehazing.

2.2.2 Color Attenuation Prior

Zhu et al. (2015) proposed the Color Attenuation Prior, which is based on the observation that the difference between the brightness and the saturation of pixels in a hazy image can be used to estimate the scene depth. This prior establishes a linear relationship between the scene depth $d(x)$ and the difference between brightness $v(x)$ and saturation $s(x)$:

$$d(x) = \theta_0 + \theta_1 v(x) + \theta_2 s(x)$$

where θ_0 , θ_1 , and θ_2 are parameters learned from a training dataset using a supervised learning method.

The color attenuation prior is based on the observation that haze typically reduces the saturation of colors while increasing their brightness. As the haze concentration increases with depth, there is a correlation between depth and the brightness-saturation difference. This correlation forms the basis for estimating the depth map, which can then be used to derive the transmission map.

The method involves the following steps:

1. Convert the hazy image from RGB to HSV color space to obtain the brightness (V) and saturation (S) channels.
2. Calculate the difference between brightness and saturation for each pixel.
3. Estimate the depth map using the linear model with pre-learned parameters.

4. Compute the transmission map using the exponential relationship between transmission and depth: $t(x) = e^{(-\beta d(x))}$.
5. Estimate the atmospheric light by finding the brightest pixels in the regions with the highest depth values.
6. Recover the scene radiance using the atmospheric scattering model.

The parameters θ_0 , θ_1 , and θ_2 are learned from a dataset of hazy images with known depth information. Zhu et al. created a synthetic dataset by applying the atmospheric scattering model to clear images with depth maps, and then used linear regression to learn the optimal parameters.

The Color Attenuation Prior method has several advantages:

1. **Computational Efficiency:** It is computationally efficient compared to the DCP method, especially when processing high-resolution images.
2. **No Patch-Based Operations:** It operates on individual pixels rather than patches, avoiding halo artifacts around depth discontinuities.
3. **Automatic Parameter Learning:** The parameters are learned from data, reducing the need for manual tuning.

However, the method also has limitations:

1. **Linear Model Assumption:** The assumption of a linear relationship between depth and color attributes may not hold for all scenes, especially those with complex lighting or unusual color distributions.
2. **Training Data Dependency:** The performance depends on the quality and diversity of the training data used to learn the parameters.
3. **Sensitivity to Color Variations:** The method may struggle with scenes that have inherent variations in brightness and saturation unrelated to haze, such as colorful objects or scenes with strong lighting contrasts.

Despite these limitations, the Color Attenuation Prior method provides a computationally efficient alternative to the DCP method and has been particularly effective for real-time applications where processing speed is critical.

2.2.3 Non-local Prior

Berman et al. (2016) introduced the Non-local Prior for image dehazing, which is based on the observation that colors in a haze-free image can be well approximated by a few hundred distinct colors. In the presence of haze, these colors form lines in the RGB space, called haze-lines, which converge to the atmospheric light.

The key insight of this method is that pixels in a haze-free image tend to cluster around a small number of distinct colors in the RGB space. When haze is added according to the atmospheric scattering model, each of these color clusters stretches into a line segment in the RGB space, with one end corresponding to the original color (no haze) and the other end approaching the atmospheric light (dense haze). These line segments, called haze-lines, all converge towards the atmospheric light.

The method involves the following steps:

1. Estimate the atmospheric light A by finding the brightest pixel in the hazy image after applying a median filter.
2. Identify the haze-lines in the RGB space by clustering the pixels of the hazy image. Each cluster represents pixels that likely had the same color in the haze-free image but appear different due to varying haze densities.
3. For each pixel, find the corresponding haze-line and estimate its position along the line, which indicates the transmission value for that pixel.
4. Regularize the transmission map to ensure spatial consistency while preserving edges.
5. Recover the scene radiance using the atmospheric scattering model.

The Non-local Prior method has several advantages:

1. **Global Approach:** It considers the global distribution of colors in the image rather than relying on local patches, which helps to avoid halo artifacts.
2. **Effective for Fine Structures:** It performs well on images with fine structures and details, which can be challenging for patch-based methods.
3. **Robustness to Non-uniform Haze:** The method can handle scenes with spatially varying haze distributions better than many other approaches.

However, the method also has limitations:

1. **Color Clustering Assumption:** The assumption that colors in a haze-free image can be approximated by a few hundred distinct colors may not hold for all scenes, especially those with smooth color gradients or complex textures.
2. **Computational Complexity:** The clustering and regularization steps can be computationally intensive, especially for high-resolution images.
3. **Sensitivity to Atmospheric Light Estimation:** The accuracy of the haze-line identification depends on the correct estimation of the atmospheric light.

The Non-local Prior method represents a significant departure from patch-based approaches like DCP, offering a global perspective on the dehazing problem. It has been particularly effective for scenes with complex structures and non-uniform haze distributions.

2.2.4 Contrast Enhancement-Based Methods

Contrast enhancement-based methods approach the dehazing problem from a different perspective, focusing on improving the visual quality of hazy images without explicitly modeling the atmospheric scattering process. These methods typically apply various image processing techniques to enhance the contrast, saturation, and details in hazy images.

One of the earliest contrast enhancement-based methods for dehazing is the Histogram Equalization (HE) technique, which redistributes the intensity values in an image to achieve a more uniform histogram. While simple HE can improve the global contrast of hazy images, it often leads to over-enhancement and unnatural appearances.

To address these limitations, more sophisticated variants have been developed:

1. **Contrast Limited Adaptive Histogram Equalization (CLAHE):** This technique applies histogram equalization locally to small regions (tiles) of the image rather than the entire image. It also limits the contrast enhancement to prevent noise amplification and over-enhancement. CLAHE has been effective for dehazing when applied to the luminance channel of images in color spaces like LAB or YCbCr, preserving color information while enhancing contrast.
2. **Retinex-Based Methods:** Based on the Retinex theory of human color vision, these methods decompose an image into illumination and reflectance components. By enhancing the reflectance component and adjusting the illumination component, they can reduce the haze effect while preserving natural colors. Multi-scale Retinex with Color Restoration (MSRCR) has been particularly effective for dehazing.
3. **Unsharp Masking:** This technique enhances image details by adding a high-pass filtered version of the image to the original. When applied with appropriate parameters, it can enhance the visibility of details obscured by haze.
4. **Gamma Correction:** Adjusting the gamma value can help to enhance the contrast in specific tonal ranges of hazy images. Adaptive gamma correction techniques that adjust the gamma value based on local image statistics have shown promising results for dehazing.
5. **Fusion-Based Enhancement:** These methods generate multiple enhanced versions of the hazy image using different techniques or parameters, and then fuse them based on quality metrics such as contrast, saturation, and exposure.

Contrast enhancement-based methods have several advantages:

1. **Computational Efficiency:** Many of these techniques are computationally efficient and can be implemented in real-time on standard hardware.
2. **No Physical Model Requirement:** They do not rely on the atmospheric scattering model or its assumptions, making them more flexible for various haze conditions.
3. **Simplicity:** These methods are often simpler to implement and understand compared to physics-based or learning-based approaches.

However, they also have significant limitations:

1. **Lack of Physical Basis:** Without modeling the physical process of haze formation, these methods may produce visually pleasing but physically implausible results.
2. **Over-Enhancement:** Aggressive contrast enhancement can lead to over-saturation, noise amplification, and unnatural appearances.
3. **Depth-Dependent Haze:** These methods typically do not account for the depth-dependent nature of haze, potentially leading to inconsistent dehazing across different depths.
4. **Color Distortion:** Enhancing contrast without considering the color shifts caused by haze can result in color distortions in the processed images.

Despite these limitations, contrast enhancement-based methods remain popular for applications where computational efficiency and simplicity are prioritized over physical accuracy, such as real-time video processing or mobile applications.

2.2.5 Fusion-Based Methods

Fusion-based dehazing methods combine multiple versions of the input image or multiple dehazing techniques to produce a final dehazed result. These methods leverage the strengths of different approaches while mitigating their individual weaknesses, often leading to more robust and visually pleasing results.

Ancuti et al. (2013) proposed one of the most influential fusion-based dehazing methods, which derives two versions of the input image—a white-balanced version and a contrast-enhanced version—and fuses them using a multi-scale fusion approach. The fusion process is guided by weight maps that measure the quality of each input in terms of contrast, saturation, and exposure.

The method involves the following steps:

1. Generate multiple inputs from the original hazy image:
 - A white-balanced version that corrects the color cast caused by haze
 - A contrast-enhanced version that improves the visibility of details
2. Compute weight maps for each input based on quality metrics:
 - Contrast weight: Measures the amount of local contrast using a Laplacian filter
 - Saturation weight: Measures the standard deviation of the color channels
 - Exposure weight: Measures how well-exposed each pixel is
3. Normalize the weight maps to ensure they sum to one at each pixel.
4. Perform multi-scale fusion using a Laplacian pyramid for the inputs and a Gaussian pyramid for the weight maps, which helps to avoid artifacts at the boundaries of regions with different weights.
5. Reconstruct the final dehazed image from the fused pyramid.

Other fusion-based approaches include:

1. **Hybrid Methods:** These combine physics-based and contrast enhancement-based techniques. For example, some methods use the DCP to estimate the transmission map and then apply contrast enhancement to the dehazed result to improve its visual quality.
2. **Multi-Algorithm Fusion:** These methods apply multiple dehazing algorithms to the same input image and fuse the results based on quality metrics or learning-based fusion strategies.
3. **Multi-Exposure Fusion:** Similar to HDR imaging, these methods capture or simulate multiple exposures of the hazy scene and fuse them to recover details in both dark and bright regions.
4. **Guided Fusion:** These methods use additional information, such as depth maps or semantic segmentation, to guide the fusion process, giving more weight to more reliable inputs for different regions of the image.

Fusion-based methods have several advantages:

1. **Robustness:** By combining multiple inputs or techniques, fusion-based methods can be more robust to variations in haze density, scene content, and lighting conditions.
2. **Balanced Results:** They often achieve a good balance between haze removal, detail preservation, and natural appearance.
3. **Adaptability:** The fusion weights can adapt to local image characteristics, providing appropriate processing for different regions of the image.

However, they also have limitations:

1. **Computational Complexity:** Generating multiple inputs and performing multi-scale fusion can be computationally intensive, especially for high-resolution images.
2. **Parameter Sensitivity:** The quality of the results depends on the parameters used to generate the inputs and compute the weight maps.

3. **Potential Artifacts:** If the inputs or weight maps have significant differences, the fusion process may introduce artifacts or inconsistencies in the final result.

Despite these limitations, fusion-based methods have shown promising results for a wide range of hazy scenes and have been particularly effective for challenging conditions where single-algorithm approaches struggle.

2.3 Learning-Based Dehazing Methods

2.3.1 CNN-Based Methods

With the advancement of deep learning, Convolutional Neural Networks (CNNs) have been widely applied to image dehazing. These methods can be categorized into two main approaches:

1. **Prior-based CNN methods:** These methods use CNNs to estimate the components of the atmospheric scattering model, such as the transmission map or atmospheric light. For example, DehazeNet (Cai et al., 2016) and MSCNN (Ren et al., 2016) use CNNs to estimate the transmission map, which is then used to recover the clear image using the atmospheric scattering model.
2. **End-to-end CNN methods:** These methods directly learn the mapping from hazy images to clear images without explicitly modeling the atmospheric scattering process. AOD-Net (Li et al., 2017) reformulates the atmospheric scattering model and uses a lightweight CNN to directly estimate the clean image. DCPDN (Zhang and Patel, 2018) employs a generative adversarial network framework with an embedded atmospheric scattering model.

DehazeNet, proposed by Cai et al. (2016), was one of the pioneering CNN-based dehazing methods. It uses a CNN to estimate the transmission map from a hazy image, inspired by the dark channel prior but learning the mapping directly from data. The network consists of four sequential operations: feature extraction, multi-scale mapping, local extremum, and non-linear regression. Once the transmission map is estimated, the atmospheric light is computed using a traditional method, and the clear image is recovered using the atmospheric scattering model.

MSCNN, proposed by Ren et al. (2016), uses a multi-scale CNN architecture to estimate the transmission map. It consists of a coarse-scale network that predicts a holistic transmission map and a fine-scale network that refines the details. This multi-scale approach helps to capture both global haze distribution and local details, leading to more accurate transmission estimation.

AOD-Net, proposed by Li et al. (2017), takes a different approach by reformulating the atmospheric scattering model to reduce the number of unknown variables. Instead of separately estimating the transmission map and atmospheric light, it directly estimates a unified parameter that combines both, simplifying the dehazing process. The network is lightweight and end-to-end trainable, making it suitable for real-time applications.

DCPDN, proposed by Zhang and Patel (2018), uses a generative adversarial network (GAN) framework with an embedded atmospheric scattering model. It consists of three components: a transmission map estimation network, an atmospheric light estimation network, and a dehazed image generation network. The adversarial training helps to produce more realistic and visually pleasing results.

More recent CNN-based methods have explored various architectural innovations:

1. **Attention Mechanisms:** Methods like PFDN (Dong et al., 2020) incorporate attention mechanisms to focus on informative regions and features, improving the dehazing performance for challenging scenes.
2. **Dense Connections:** Networks like GridDehazeNet (Liu et al., 2019) use dense connections to facilitate feature reuse and gradient flow, leading to better training dynamics and performance.
3. **Multi-Task Learning:** Some methods jointly learn dehazing and related tasks such as depth estimation or semantic segmentation, leveraging the complementary information to improve overall performance.
4. **Progressive Refinement:** Methods like PDNet (Chen et al., 2021) use a progressive refinement approach, starting with a coarse dehazed image and iteratively refining it to produce the final result.

CNN-based methods have several advantages:

1. **Learning from Data:** They can learn complex mappings from hazy to clear images directly from data, without relying on handcrafted priors or assumptions.
2. **End-to-End Training:** Many CNN-based methods can be trained end-to-end, optimizing all components jointly for the dehazing task.
3. **Adaptability:** With sufficient training data, CNN-based methods can adapt to various haze conditions and scene types.

However, they also have limitations:

1. **Data Dependency:** They require large amounts of training data, ideally paired hazy and clear images, which can be difficult to obtain for real-world scenarios.
2. **Generalization:** Models trained on synthetic data may struggle to generalize to real-world haze conditions due to the domain gap.
3. **Computational Resources:** Training deep CNN models requires significant computational resources, and inference can be slow without GPU acceleration.
4. **Interpretability:** Unlike physics-based methods, CNN-based approaches often lack interpretability, making it difficult to understand and debug their behavior.

Despite these limitations, CNN-based methods have achieved state-of-the-art performance on benchmark datasets and continue to be an active area of research in image dehazing.

2.3.2 GAN-Based Methods

Generative Adversarial Networks (GANs) have been employed for image dehazing to improve the visual quality of dehazed images. GAN-based methods typically consist of a generator network that produces dehazed images and a discriminator network that distinguishes between real clear images and generated dehazed images.

The adversarial training process encourages the generator to produce dehazed images that are indistinguishable from real clear images, leading to more realistic and visually pleasing results. This is particularly valuable for image dehazing, where the goal is not just to remove haze but also to produce natural-looking images.

CycleDehaze, proposed by Engin et al. (2018), uses a cycle-consistent adversarial network to learn the mapping between hazy and clear images without requiring paired training data. It consists of two generators: one that maps hazy images to clear images and another that maps clear images to hazy images. The cycle consistency loss ensures that an image translated from one domain to the other and back should be identical to the original image. This approach is valuable when paired training data is scarce, as it can leverage unpaired datasets of hazy and clear images.

EPDN (Enhanced Perceptual Dehazing Network), proposed by Qu et al. (2019), enhances the perceptual quality of dehazed images by incorporating perceptual loss functions in the GAN framework. It uses a multi-scale generator with dense connections and a discriminator that operates on multiple scales. The perceptual loss, based on features extracted from a pre-trained VGG network, helps to preserve semantic information and fine details in the dehazed images.

Other notable GAN-based dehazing methods include:

1. **DHSGAN (Dehazing and Super-Resolution GAN)**, proposed by Li et al. (2020), which jointly addresses dehazing and super-resolution using a dual-task GAN framework. It leverages the complementary nature of these tasks, as both aim to recover high-frequency details lost due to degradation.
2. **YOLY (You Only Look Yourself)**, proposed by Li et al. (2021), which uses a self-supervised GAN approach that does not require clear reference images for training. It leverages the observation that haze is typically depth-dependent, so nearby regions should have less haze than distant regions in the same scene.
3. **DualGAN**, proposed by Yang et al. (2018), which uses two GANs to learn bidirectional mappings between hazy and clear domains, similar to CycleDehaze but with different architectural choices and loss functions.
4. **PGAN (Progressive GAN)**, proposed by Zhang et al. (2019), which adopts a progressive training strategy, starting with low-resolution images and gradually increasing the resolution. This approach stabilizes the training process and helps to capture both global haze distribution and fine details.

GAN-based methods have several advantages for image dehazing:

1. **Perceptual Quality:** The adversarial training encourages the generator to produce visually pleasing results that look natural to human observers, often outperforming other methods in terms of perceptual quality.
2. **Unpaired Training:** Some GAN-based methods can be trained with unpaired data, addressing the scarcity of paired hazy and clear images in real-world scenarios.
3. **Detail Preservation:** The discriminator's ability to distinguish between real and fake images encourages the generator to preserve fine details and textures that might be lost with other methods.
4. **Adaptability:** With appropriate training data, GAN-based methods can adapt to various haze conditions and scene types.

However, they also have limitations:

1. **Training Instability:** GANs are notoriously difficult to train, often suffering from issues like mode collapse, vanishing gradients, or training instability.

2. **Artifacts:** GAN-based methods may introduce artifacts or hallucinate details that were not present in the original scene, potentially compromising the fidelity of the dehazed image.
3. **Computational Resources:** Training GAN models requires significant computational resources and time, making them less accessible for researchers with limited resources.
4. **Evaluation Challenges:** Traditional metrics like PSNR and SSIM may not fully capture the perceptual quality improvements achieved by GAN-based methods, making evaluation and comparison challenging.

Despite these limitations, GAN-based methods have shown promising results for image dehazing, particularly in terms of perceptual quality and natural appearance. They represent an important direction in learning-based dehazing research, complementing CNN-based approaches with their focus on realistic image generation.

2.3.3 Transformer-Based Methods

More recently, Transformer architectures, which have shown remarkable success in natural language processing, have been adapted for image dehazing. Transformer-based methods leverage the self-attention mechanism to capture long-range dependencies in images, which is beneficial for understanding the global context of haze distribution.

The self-attention mechanism allows each pixel to attend to all other pixels in the image, enabling the model to capture relationships between distant regions. This is particularly valuable for image dehazing, where the haze distribution and scene content can vary significantly across the image.

DehazeFormer, proposed by Song et al. (2022), combines CNN features with Transformer blocks to effectively capture both local and global features for image dehazing. It uses a hierarchical structure with multiple stages, each consisting of a patch embedding layer, Transformer blocks, and a patch expanding layer. The patch embedding layer extracts features using convolutional operations, while the Transformer blocks capture long-range dependencies using self-attention. The patch expanding layer upsamples the features for the next stage or final output.

FocalNet, proposed by Yang et al. (2022), uses a focal self-attention mechanism to focus on informative regions while reducing computational complexity. Unlike standard self-attention that attends to all pixels equally, focal self-attention adaptively adjusts the attention weights based on the importance of different regions. This allows the model to allocate more computational resources to regions that are more challenging for dehazing, such as areas with dense haze or complex textures.

Other notable Transformer-based dehazing methods include:

1. **TransWeather**, proposed by Valanarasu et al. (2022), which uses a Transformer-based architecture for various weather removal tasks, including dehazing, deraining, and desnowing. It leverages a multi-head self-attention mechanism to capture global context and a convolutional branch to capture local details.
2. **HIT (Haze Image Transformer)**, proposed by Chen et al. (2022), which uses a hybrid architecture combining Transformers for global context modeling and CNNs for local feature extraction. It also incorporates a physics-guided module that leverages the atmospheric scattering model to guide the learning process.
3. **DFormer**, proposed by Wang et al. (2023), which uses a dual-branch Transformer architecture with one branch focusing on haze removal and the other on detail enhancement. The two branches share information through cross-attention mechanisms, allowing them to complement each other.
4. **SwinDehazeNet**, proposed by Liu et al. (2023), which adapts the Swin Transformer architecture for image dehazing. It uses shifted windows for self-attention computation, reducing computational complexity while maintaining the ability to capture long-range dependencies.

Transformer-based methods have several advantages for image dehazing:

1. **Global Context Modeling:** The self-attention mechanism allows these methods to capture long-range dependencies and global context, which is valuable for understanding the haze distribution across the entire image.
2. **Adaptive Processing:** Transformer-based methods can adaptively process different regions of the image based on their content and haze density, potentially leading to more effective dehazing.
3. **Scalability:** Transformer architectures can be scaled to handle high-resolution images and complex scenes, making them suitable for challenging dehazing scenarios.
4. **Transfer Learning:** Pre-trained Transformer models from other vision tasks can be fine-tuned for dehazing, leveraging knowledge learned from large-scale datasets.

However, they also have limitations:

1. **Computational Complexity:** The standard self-attention mechanism has quadratic complexity with respect to the number of pixels, making it computationally expensive for high-resolution images. Various approximations and optimizations have been proposed to address this issue.
2. **Data Requirements:** Transformer-based methods typically require large amounts of training data to achieve optimal performance, which can be challenging for dehazing due to the scarcity of paired hazy and clear images.
3. **Model Size:** Transformer models typically have more parameters than CNN-based models, requiring more memory for training and inference.
4. **Lack of Inductive Bias:** Unlike CNNs, which have built-in inductive biases for processing images (e.g., translation equivariance), Transformers rely more heavily on learning from data, which can be a challenge with limited training samples.

Despite these limitations, Transformer-based methods have demonstrated state-of-the-art performance on benchmark datasets and represent a promising direction for future research in image dehazing. Their ability to capture global context and long-range dependencies makes them particularly suitable for complex scenes with varying haze distributions.

2.3.4 Physics-Guided Learning Methods

Physics-guided learning methods combine the strengths of traditional physics-based approaches and modern learning-based techniques. These methods incorporate the atmospheric scattering model or other physical principles into the neural network architecture or loss function, providing a stronger inductive bias and improving the interpretability of the model.

The key insight behind physics-guided learning is that while deep learning models excel at learning complex mappings from data, they can benefit from incorporating domain knowledge about the physical process of haze formation. This can lead to more efficient learning, better generalization to unseen data, and more physically plausible results.

PGAN (Physics-Guided Adversarial Network), proposed by Li et al. (2020), incorporates the atmospheric scattering model into a GAN framework. The generator is designed to estimate the transmission map and atmospheric light, which are then used to recover the clear image using the atmospheric scattering model. The discriminator evaluates both the dehazed image and the estimated physical parameters, encouraging physically plausible results.

PHYD-Net (Physics-based Hybrid Dehazing Network), proposed by Chen et al. (2021), uses a hybrid architecture that combines a physics-based branch and a learning-based branch. The physics-based branch estimates the transmission map and atmospheric light using traditional methods, while the learning-based branch directly predicts a residual image to refine the physics-based result. The two branches complement each other, with the physics-based branch providing a physically plausible initial estimate and the learning-based branch addressing its limitations.

Other notable physics-guided learning methods include:

1. **PMHLD (Physical Model-Guided Hybrid Learning Dehazing)**, proposed by Zhao et al. (2021), which uses a multi-stage approach that first estimates the transmission map and atmospheric light using a CNN, then refines the transmission map using a guided filter, and finally recovers the clear image using the atmospheric scattering model.
2. **KDPN (Knowledge-Driven Perception Network)**, proposed by Yang et al. (2020), which incorporates physical priors such as the dark channel prior and color attenuation prior into the network architecture through specialized modules that mimic the behavior of these priors.
3. **PAD-Net (Physics-Aware Dehazing Network)**, proposed by Li et al. (2022), which uses a physics-aware loss function that enforces consistency with the atmospheric scattering model, in addition to the standard reconstruction loss.
4. **PSDL (Physics-Supervised Deep Learning)**, proposed by Wang et al. (2021), which uses the atmospheric scattering model to generate synthetic training data with known physical parameters, allowing for supervised learning of these parameters even when ground truth clear images are not available.

Physics-guided learning methods have several advantages:

1. **Interpretability:** By incorporating physical models, these methods provide more interpretable results, allowing users to understand the estimated physical parameters such as transmission and atmospheric light.
2. **Data Efficiency:** The physical constraints can serve as a strong regularization, reducing the amount of training data needed to achieve good performance.
3. **Generalization:** The physical priors can help the model generalize better to unseen data, especially when the test conditions differ significantly from the training data.

4. **Physical Plausibility:** The results are more likely to be physically plausible, avoiding artifacts or unrealistic appearances that might occur with purely data-driven approaches.

However, they also have limitations:

1. **Model Complexity:** Incorporating physical models can increase the complexity of the network architecture and training process.
2. **Simplified Physics:** The atmospheric scattering model used in these methods is often simplified and may not capture all the complexities of real-world haze formation.
3. **Balance of Physics and Learning:** Finding the right balance between physical constraints and learning from data can be challenging, as overly strict physical constraints might limit the model's ability to handle cases where the physical model is inaccurate.

Physics-guided learning represents a promising direction for image dehazing research, combining the interpretability and theoretical foundation of physics-based methods with the data-driven capabilities of learning-based approaches.

2.3.5 Unsupervised and Self-Supervised Methods

Unsupervised and self-supervised learning methods for image dehazing address the challenge of limited paired training data by learning from unpaired or single images. These methods leverage various forms of supervision signals derived from the data itself, without requiring explicit ground truth clear images.

Unsupervised methods typically use cycle consistency, domain adaptation, or other techniques to learn the mapping between hazy and clear domains without paired examples. Self-supervised methods, on the other hand, create supervisory signals from the input data itself, such as by synthesizing hazy images from clear ones using the atmospheric scattering model or by exploiting the structure of natural images.

CycleGAN-based methods, such as CycleDehaze mentioned earlier, use cycle consistency to learn bidirectional mappings between hazy and clear domains. The key idea is that if an image is translated from the hazy domain to the clear domain and back, it should be identical to the original hazy image, and vice versa. This constraint allows the model to learn meaningful mappings without paired data.

Domain adaptation methods, such as DADA (Domain Adaptation for Dehazing Algorithm) proposed by Shao et al. (2020), address the domain gap between synthetic and real haze by learning domain-invariant features. These methods typically use adversarial training to align the feature distributions of synthetic and real data, allowing models trained on synthetic data to generalize better to real-world haze.

Self-supervised methods include:

1. **Zero-DCE (Zero-Reference Deep Curve Estimation)**, proposed by Guo et al. (2020), which formulates dehazing as an image-specific curve estimation problem. It learns to estimate pixel-wise curves that map the hazy image to a dehazed version, using non-reference loss functions based on color constancy, exposure control, and illumination smoothness.
2. **SSID (Self-Supervised Image Dehazing)**, proposed by Yang et al. (2021), which uses the atmospheric scattering model to synthesize hazy images from the input image itself, creating pseudo pairs for training. It leverages the observation that haze is typically depth-dependent, so nearby regions should have less haze than distant regions.
3. **YOLY (You Only Look Yourself)**, mentioned earlier, which uses a self-supervised GAN approach that exploits the depth-dependent nature of haze to create supervisory signals from a single hazy image.
4. **CDD (Contrastive Disentangled Dehazing)**, proposed by Chen et al. (2022), which uses contrastive learning to disentangle content and haze features, allowing for more effective dehazing without paired supervision.

Unsupervised and self-supervised methods have several advantages:

1. **Data Efficiency:** They can learn from unpaired or single images, addressing the scarcity of paired hazy and clear images in real-world scenarios.
2. **Domain Adaptation:** They can better bridge the domain gap between synthetic and real haze, improving generalization to real-world conditions.
3. **Scalability:** They can potentially leverage large amounts of unpaired data, which is more readily available than paired data.

However, they also have limitations:

1. **Training Complexity:** Unsupervised and self-supervised training can be more complex and less stable than supervised training, requiring careful design of loss functions and training strategies.

2. **Performance Gap:** Without direct supervision from ground truth clear images, these methods may not achieve the same level of performance as supervised methods when paired data is available.
3. **Evaluation Challenges:** Evaluating these methods can be challenging, as traditional metrics like PSNR and SSIM require ground truth references, which may not be available for real-world hazy images.

Despite these limitations, unsupervised and self-supervised methods represent an important direction for image dehazing research, particularly for real-world applications where paired training data is scarce or unavailable.

2.4 Evaluation Metrics

2.4.1 Full-Reference Metrics

Full-reference metrics evaluate the quality of dehazed images by comparing them with ground truth clear images. These metrics are widely used in the literature to quantitatively assess the performance of dehazing algorithms on synthetic datasets where paired hazy and clear images are available.

Peak Signal-to-Noise Ratio (PSNR) is one of the most commonly used full-reference metrics. It measures the pixel-wise difference between the dehazed image and the ground truth, expressed in decibels (dB). Higher PSNR values indicate better quality, with typical values for good dehazing results ranging from 20 to 30 dB. PSNR is defined as:

$$\text{PSNR} = 10 \cdot \log_{10}(\text{MAX}^2/\text{MSE})$$

where MAX is the maximum possible pixel value (typically 255 for 8-bit images) and MSE is the mean squared error between the dehazed image and the ground truth.

While PSNR is widely used due to its simplicity and ease of computation, it has limitations. It assumes that image quality is primarily determined by pixel-wise differences, which may not align with human perception. For example, a slight shift in the image can result in a significant drop in PSNR, even if the visual quality is largely unchanged.

Structural Similarity Index Measure (SSIM) addresses some of the limitations of PSNR by considering structural information. It evaluates the similarity between two images based on luminance, contrast, and structure, with values ranging from -1 to 1 (higher is better). SSIM is more aligned with human perception than PSNR and is defined as:

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma$$

where l , c , and s represent the luminance, contrast, and structure comparison functions, respectively, and α , β , and γ are positive constants that adjust the relative importance of the three components.

SSIM is more sensitive to structural changes and less sensitive to uniform changes in brightness or contrast, which aligns better with human perception. However, it still has limitations, particularly for images with complex textures or when the dehazed image has different color characteristics than the ground truth.

Color Difference (CIEDE2000) is a metric specifically designed to measure perceptual color differences. It operates in the CIELAB color space, which is designed to be perceptually uniform, meaning that a given change in color value should produce a similar perceived change regardless of where in the color space the change occurs. CIEDE2000 takes into account variations in human color perception across the color spectrum and is particularly useful for evaluating color fidelity in dehazed images.

Other full-reference metrics used in dehazing evaluation include:

1. **Feature Similarity Index (FSIM)**, which considers phase congruency and gradient magnitude features to better capture the perceptual quality of images.
2. **Visual Information Fidelity (VIF)**, which quantifies the mutual information between the reference and distorted images in various wavelet subbands.
3. **Mean Absolute Error (MAE)**, which measures the average absolute difference between corresponding pixels in the dehazed and ground truth images.
4. **Learned Perceptual Image Patch Similarity (LPIPS)**, which uses features from pre-trained deep neural networks to measure perceptual similarity, often aligning better with human judgments than traditional metrics.

Full-reference metrics provide a quantitative basis for comparing different dehazing algorithms, but they have several limitations:

1. **Requirement for Ground Truth:** They require ground truth clear images, which are often unavailable for real-world hazy images.
2. **Alignment Sensitivity:** Many metrics are sensitive to slight misalignments between the dehazed and ground truth images, which can occur due to geometric transformations during the dehazing process.
3. **Perceptual Correlation:** The correlation between these metrics and human perception is not perfect, especially for complex scenes or when different types of distortions are present.
4. **Context Insensitivity:** Most metrics do not consider the semantic importance of different regions in the image, treating all pixels equally regardless of their content.

Despite these limitations, full-reference metrics remain valuable tools for quantitative evaluation when ground truth images are available, particularly when used in combination to capture different aspects of image quality.

2.4.2 No-Reference Metrics

No-reference metrics evaluate the quality of dehazed images without requiring ground truth clear images. These metrics are particularly valuable for assessing dehazing performance on real-world hazy images where ground truth is unavailable.

Natural Image Quality Evaluator (NIQE) is a no-reference metric that measures the naturalness of an image based on statistical features. It uses a quality-aware collection of statistical features based on a simple and successful space domain natural scene statistic model. NIQE is trained on a corpus of natural images and assigns lower scores to images that deviate from the statistical properties of natural images. Lower NIQE scores indicate better perceptual quality.

NIQE has the advantage of not requiring any training on human-rated distorted images, making it more general and less biased towards specific types of distortions. However, it may not always align with human perception, especially for images with artistic or stylistic elements that deviate from natural statistics.

Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) is another no-reference metric that operates in the spatial domain. It uses scene statistics of locally normalized luminance coefficients to quantify possible losses of naturalness in the image due to the presence of distortions. BRISQUE is trained on a database of images with known distortions and human quality ratings, allowing it to predict the quality of new images. Like NIQE, lower BRISQUE scores indicate better quality.

BRISQUE has shown good correlation with human perception for various distortions, including those introduced by dehazing algorithms. However, it is trained on specific types of distortions and may not generalize well to novel distortions or artistic effects.

Fog Aware Density Evaluator (FADE) is a specialized no-reference metric designed specifically for assessing the presence of fog or haze in images. It uses a regression model trained on a database of foggy images with human-rated fog density scores. FADE provides a fog density score, with lower values indicating less fog or haze. This metric is particularly useful for evaluating the effectiveness of dehazing algorithms in terms of haze removal, rather than overall image quality.

Other no-reference metrics used in dehazing evaluation include:

1. **Perception-based Image Quality Evaluator (PIQE)**, which estimates quality based on local spatial statistics of an image without requiring training on human-rated images.
2. **Blind Image Quality Index (BIQI)**, which uses a two-step framework that first identifies the distortion type and then applies distortion-specific quality assessment.
3. **Contrast-to-Noise Ratio (CNR)**, which measures the contrast between regions of interest and background noise, providing an indication of visibility enhancement.
4. **Entropy**, which measures the amount of information or randomness in an image, with higher entropy often indicating more detailed and less hazy images.
5. **Edge Preservation Index (EPI)**, which evaluates how well edges are preserved in the dehazed image compared to the hazy input.

No-reference metrics have several advantages for dehazing evaluation:

1. **Applicability to Real-World Images:** They can be applied to real-world hazy images where ground truth clear images are unavailable.
2. **Direct Assessment:** They directly assess the quality of the dehazed image without requiring comparison with a reference.
3. **Specialized Evaluation:** Some metrics, like FADE, are specifically designed for haze-related quality assessment.

However, they also have limitations:

1. **Limited Correlation with Human Perception:** The correlation between no-reference metrics and human perception can be lower than that of full-reference metrics, especially for complex scenes or novel distortions.
2. **Training Bias:** Metrics that rely on training data may be biased towards the types of images and distortions represented in the training set.
3. **Lack of Consensus:** There is less consensus on which no-reference metrics best capture dehazing quality, leading to inconsistent evaluation across different studies.
4. **Context Insensitivity:** Like full-reference metrics, most no-reference metrics do not consider the semantic importance of different regions in the image.

Despite these limitations, no-reference metrics provide valuable tools for evaluating dehazing performance on real-world images and complement full-reference metrics when ground truth is available.

2.4.3 Human Visual Perception

Human visual perception plays a crucial role in evaluating the quality of dehazed images, as the ultimate goal of image dehazing is to improve the visual experience for human observers. Subjective evaluation through human assessment provides insights that may not be captured by objective metrics, particularly regarding the naturalness, pleasantness, and overall visual appeal of dehazed images.

Subjective evaluation typically involves presenting dehazed images to human observers and collecting their ratings or preferences. Several methodologies are commonly used:

1. **Mean Opinion Score (MOS):** Observers rate the quality of dehazed images on a scale (e.g., 1-5 or 1-10), and the average rating is computed as the MOS. This approach provides a direct measure of perceived quality but can be influenced by individual biases and the specific instructions given to observers.
2. **Paired Comparison:** Observers are presented with pairs of images (e.g., results from different dehazing algorithms) and asked to select the one they prefer. This approach is more intuitive for observers and can provide more reliable results, especially when the quality differences are subtle.
3. **Rank Ordering:** Observers rank multiple dehazed images from best to worst according to specific criteria such as haze removal effectiveness, color fidelity, or overall quality. This approach provides a relative assessment of different algorithms but may become challenging with a large number of images.
4. **Expert Assessment:** Domain experts, such as professional photographers or image processing specialists, evaluate the dehazed images based on their expertise. This approach can provide more informed judgments but may be less representative of the general population.

The criteria used for human assessment of dehazed images typically include:

1. **Haze Removal Effectiveness:** How well the algorithm removes haze and improves visibility, especially in distant regions of the scene.
2. **Color Fidelity:** Whether the colors in the dehazed image appear natural and consistent with expectations for the scene.
3. **Detail Preservation:** How well fine details and textures are preserved or recovered in the dehazed image.
4. **Artifact Presence:** Whether the dehazing process introduces artifacts such as color distortions, halos, or noise.
5. **Overall Visual Quality:** The general aesthetic appeal and pleasantness of the dehazed image.

Human visual perception is influenced by various factors that should be considered in subjective evaluation:

1. **Context and Expectations:** The perceived quality of a dehazed image can be influenced by the context in which it is viewed and the observer's expectations. For example, observers may have different expectations for artistic photographs versus surveillance footage.
2. **Viewing Conditions:** Factors such as display calibration, ambient lighting, viewing distance, and viewing angle can affect the perceived quality of dehazed images.
3. **Individual Differences:** Observers may have different preferences, sensitivities, and visual acuity, leading to variations in their assessments.
4. **Familiarity with the Scene:** If observers are familiar with the scene depicted in the image, they may have prior knowledge about how it should appear without haze, influencing their assessment.

5. **Adaptation:** The human visual system adapts to viewing conditions, including haze. Observers may mentally compensate for haze in the original image, affecting their perception of the dehazed result.

To address these factors and ensure reliable subjective evaluation, several best practices are recommended:

1. **Standardized Viewing Conditions:** Conduct evaluations under controlled and consistent viewing conditions, including display calibration, ambient lighting, and viewing distance.
2. **Sufficient Sample Size:** Include a sufficient number of observers to account for individual variations and obtain statistically significant results.
3. **Diverse Image Set:** Use a diverse set of images representing different scenes, haze conditions, and challenging cases to comprehensively evaluate dehazing performance.
4. **Blind Evaluation:** Present images without revealing which algorithm produced them to avoid bias based on expectations or prior knowledge.
5. **Clear Instructions:** Provide clear and consistent instructions to observers about the evaluation criteria and rating scale.
6. **Statistical Analysis:** Apply appropriate statistical methods to analyze the results, including measures of central tendency, variability, and statistical significance.

While human visual assessment provides valuable insights, it also has limitations:

1. **Subjectivity:** Human judgments are inherently subjective and can vary based on individual preferences and biases.
2. **Resource Intensity:** Conducting large-scale human studies is time-consuming, expensive, and logistically challenging.
3. **Reproducibility:** Subjective evaluations may be difficult to reproduce exactly due to variations in observers, instructions, and viewing conditions.
4. **Limited Scope:** Practical constraints often limit the number of images and algorithms that can be evaluated through human assessment.

Despite these limitations, human visual perception remains the ultimate benchmark for image quality assessment, and subjective evaluation provides essential insights that complement objective metrics. A comprehensive evaluation of dehazing algorithms should ideally combine both objective metrics and human assessment to capture different aspects of performance.

2.5 Challenges and Limitations

2.5.1 Scene Diversity and Complexity

One of the major challenges in image dehazing is handling the diversity and complexity of real-world scenes. Different scenes can have vastly different characteristics in terms of content, lighting, colors, and textures, making it difficult to develop dehazing algorithms that perform well across all scenarios.

Sky regions pose a particular challenge for many dehazing algorithms, especially those based on the dark channel prior. The dark channel prior assumes that at least one color channel has very low intensity in most non-sky regions of haze-free images. However, this assumption does not hold for sky regions, which are naturally bright and have high intensity in all color channels. As a result, algorithms based on this prior often overestimate the haze in sky regions, leading to color distortions and over-enhancement.

Night-time scenes present another challenging scenario for dehazing algorithms. Most dehazing methods are designed for daylight conditions and may not perform well in low-light environments. Night-time haze has different optical properties compared to daylight haze, with artificial light sources creating complex illumination patterns and glare effects. Additionally, low-light conditions often introduce noise, which can be amplified by dehazing algorithms, further degrading the image quality.

Scenes with complex depth structures, such as urban environments with buildings at various distances, present challenges for algorithms that assume a smooth depth distribution. Depth discontinuities can lead to halo artifacts around object boundaries, especially for patch-based methods that use a fixed patch size. These artifacts appear as unnatural bright or dark regions around edges where there is a significant depth change.

Water bodies, reflective surfaces, and transparent objects also pose challenges for dehazing algorithms. These elements have optical properties that differ from the assumptions made by most dehazing methods. For example, water surfaces can reflect light and create specular highlights, while transparent objects like glass can transmit light from behind them. These phenomena are not well captured by the standard atmospheric scattering model used in most dehazing algorithms.

Scenes with non-uniform haze distribution, such as those with fog patches or varying haze density across the image, challenge algorithms that assume a uniform atmospheric scattering coefficient. In reality, haze density can vary spatially due to factors such as wind, temperature gradients, and terrain features. Algorithms that estimate a global transmission map or atmospheric light may struggle with these non-uniform conditions.

To address the challenge of scene diversity and complexity, several approaches have been proposed:

1. **Adaptive Parameter Selection:** Some methods dynamically adjust parameters based on local image characteristics, allowing for more effective processing of different regions within the same image.
2. **Scene Segmentation:** By segmenting the image into regions with different characteristics (e.g., sky, buildings, vegetation) and applying specialized processing to each region, algorithms can better handle diverse scene elements.
3. **Learning-Based Approaches:** Deep learning methods can potentially learn to handle diverse scenes if trained on a sufficiently varied dataset. However, this requires large and diverse training data, which can be challenging to obtain.
4. **Fusion-Based Methods:** Combining multiple dehazing techniques or processing paths can help address different aspects of scene complexity, with each component handling specific challenges.
5. **Physics-Guided Learning:** Incorporating physical models and constraints into learning-based approaches can help them generalize better to diverse scenes by providing a stronger inductive bias.

Despite these approaches, scene diversity and complexity remain significant challenges for image dehazing, and developing algorithms that perform well across all scenarios is an ongoing research goal.

2.5.2 Color Distortion and Artifacts

Color distortion and artifacts are common issues in image dehazing that can significantly affect the visual quality of the results. These problems can arise from various sources, including limitations of the atmospheric scattering model, inaccurate parameter estimation, and algorithmic constraints.

Color distortion in dehazed images often manifests as unnatural color saturation, color shifts, or color inconsistencies. This can occur when the dehazing algorithm incorrectly estimates the atmospheric light or transmission map, leading to improper color correction. For example, if the atmospheric light is estimated to be too blue, the dehazed image may have a yellowish tint as the algorithm tries to compensate for the perceived blue haze.

The dark channel prior-based methods are particularly prone to color distortion in regions where the prior does not hold, such as sky regions or objects with colors similar to the atmospheric light. In these cases, the transmission may be underestimated, leading to excessive color correction and unnatural appearances.

Halo artifacts are another common issue, especially for patch-based methods. These artifacts appear as bright or dark bands around edges where there are significant depth discontinuities. They occur because the patch-based operations used to estimate the transmission map do not respect object boundaries, leading to incorrect transmission values near edges. The size of the patch used in these methods directly affects the severity of halo artifacts, with larger patches generally producing more noticeable halos.

Over-enhancement is a problem where the dehazing algorithm removes too much haze, resulting in an unnatural, over-processed appearance. This can lead to excessive contrast, saturated colors, and loss of the atmospheric perspective that provides depth cues in natural scenes. Over-enhancement often occurs when the algorithm parameters are not properly tuned or when the algorithm does not adapt to the specific haze density in the scene.

Noise amplification is another artifact that can degrade the quality of dehazed images. Since dehazing involves dividing the input image by the transmission map (which can have small values in hazy regions), any noise in the original image can be significantly amplified in the dehazed result. This is particularly problematic for images captured in low-light conditions or with high ISO settings, which already contain noise before dehazing.

Texture loss can occur when dehazing algorithms over-smooth the image in an attempt to reduce noise or artifacts. This can result in a loss of fine details and textures, making the dehazed image appear artificial or cartoon-like. Some methods use strong regularization or smoothing operations on the transmission map, which can inadvertently remove texture information from the final result.

To address these issues, various techniques have been proposed:

1. **Guided Filtering:** Using edge-preserving filters like the guided filter to refine the transmission map can help reduce halo artifacts while preserving edge details.
2. **Adaptive Parameter Selection:** Adjusting parameters based on local image characteristics can help prevent over-enhancement and color distortion in different regions of the image.

3. **Color Correction:** Post-processing steps that specifically address color fidelity, such as white balancing or color transfer techniques, can help mitigate color distortion.
4. **Noise-Aware Processing:** Incorporating noise models or denoising steps into the dehazing pipeline can help prevent noise amplification in the dehazed result.
5. **Detail Preservation Techniques:** Using methods that explicitly preserve or enhance texture details can help prevent texture loss during dehazing.
6. **Perceptual Loss Functions:** For learning-based methods, using perceptual loss functions that consider human visual perception can help reduce artifacts and improve the naturalness of dehazed images.

Despite these techniques, color distortion and artifacts remain significant challenges in image dehazing, particularly for challenging scenes with complex lighting, colors, or depth structures. Balancing effective haze removal with natural appearance and minimal artifacts is a delicate trade-off that continues to drive research in this field.

2.5.3 Computational Efficiency

Computational efficiency is a critical consideration for image dehazing algorithms, particularly for applications that require real-time processing or have limited computational resources. The computational complexity of dehazing algorithms varies widely, from simple contrast enhancement techniques to sophisticated deep learning models.

Traditional physics-based methods like the Dark Channel Prior (DCP) involve several computationally intensive steps. The original DCP method uses soft matting for transmission map refinement, which has a high computational complexity of $O(N^2)$ for an image with N pixels. This makes it impractical for real-time applications or high-resolution images. Later improvements, such as using guided filtering instead of soft matting, reduced the complexity to $O(N)$, making the algorithm more efficient but still challenging for real-time processing of high-resolution images.

Patch-based operations, which are common in many traditional dehazing methods, can be computationally expensive, especially for large patch sizes. These operations typically involve sliding a window over the entire image and computing statistics or applying filters within each window. The computational complexity increases with both the image size and the patch size, making these methods slow for high-resolution images or when large patches are needed for effective dehazing.

Deep learning-based methods, while often more effective than traditional approaches, can be computationally intensive during both training and inference. Training deep neural networks for image dehazing requires significant computational resources, including high-performance GPUs and large amounts of memory. The training process can take hours or even days, depending on the network architecture, dataset size, and training strategy.

During inference, the computational requirements of deep learning models depend on their architecture and size. Large models with millions of parameters may achieve better dehazing performance but require more memory and computational power for inference. This can be a limitation for deployment on resource-constrained devices such as mobile phones, embedded systems, or edge devices.

Real-time processing is a particular challenge for image dehazing, especially for high-resolution images or video streams. Many applications, such as autonomous driving, surveillance, or augmented reality, require dehazing to be performed in real-time with minimal latency. Achieving this level of performance often requires compromises in terms of dehazing quality or specialized hardware acceleration.

Memory consumption is another important aspect of computational efficiency. Some dehazing algorithms, particularly those based on global operations or large neural networks, require significant memory resources. This can be a limitation for devices with limited memory capacity or when processing very high-resolution images.

To address these computational efficiency challenges, various approaches have been proposed:

1. **Algorithm Optimization:** Optimizing the implementation of dehazing algorithms using techniques such as parallel processing, vectorization, or algorithm-specific optimizations can significantly improve performance.
2. **Hardware Acceleration:** Leveraging specialized hardware such as GPUs, FPGAs, or dedicated image processing units can accelerate dehazing algorithms, particularly those based on deep learning or parallel operations.
3. **Model Compression:** For deep learning-based methods, techniques such as pruning, quantization, or knowledge distillation can reduce the model size and computational requirements while maintaining reasonable performance.
4. **Multi-Resolution Processing:** Processing lower-resolution versions of the image for initial estimates and then refining at higher resolutions can reduce computational load while preserving quality.
5. **Approximate Computing:** Using approximate algorithms or computations that trade some accuracy for speed can be acceptable for applications where perfect dehazing is not critical.

6. **Incremental Processing:** For video dehazing, processing only the changed parts of consecutive frames or reusing computations from previous frames can improve efficiency.
7. **Lightweight Architectures:** Designing neural network architectures specifically for efficiency, such as MobileNet-inspired designs or networks with depthwise separable convolutions, can reduce computational requirements.

Despite these approaches, achieving a good balance between dehazing quality and computational efficiency remains a challenge, particularly for real-time applications or resource-constrained devices. The development of more efficient algorithms and hardware acceleration techniques continues to be an important research direction in image dehazing.

2.5.4 Training Data Scarcity