Research on Video Production Based on Digital Image Processing Technology

Yixin Pan, Jing Yang*, Fuhui Chen, Zhiguo Wang

Harbin Institute of Information Technology, Harbin, 150431, China

*Corresponding author: 365117735@qq.com

Abstract. Aiming at the problems of poor image quality, low efficiency and insufficient inter-frame coherence in traditional video production, this paper proposes a video production scheme based on digital image processing. A full-process processing system is constructed by integrating multiple technologies, including automatic color correction, optical flow inter-frame smoothing, intelligent matting and deep learning-based super-resolution reconstruction. Experiments show that the proposed scheme greatly improves image quality indicators and raises editing efficiency by 60%. It effectively optimizes visual effect and production efficiency, and can be applied to short videos, film and television post-production, teaching videos and other scenarios.

Keywords: Digital Image Processing; Video Production; Image Quality Enhancement; Video Editing; Super-resolution Reconstruction

1 Introduction

1.1 Research Background

The popularization of mobile networks and intelligent devices has boosted the rapid development of the video industry, making high definition, intelligence and high efficiency the mainstream development trend. Users demand higher video quality and production efficiency. Traditional video production relies on manual operation and fixed templates, often causing color distortion, detail loss, and slow processing[1].

Digital image processing can optimize images at the pixel level. However, existing studies mostly focus on single algorithm application, lacking an integrated full-process processing framework. Poor algorithm adaptability easily causes cascading errors and fails to meet the demands of large-scale, high-quality and high-efficiency video production[2].

1.2 Research Objectives and Contributions

Based on digital image processing technology, this paper constructs a standardized and automated video production workflow. A frame sequence processing pipeline is established to connect all production modules and ensure coordinated operation. Multiple

R. Zur et al. (eds.), *Proceedings of the 2026 5th International Conference on Art Design and Digital Technology (ADDT 2026)*, Advances in Computer Science Research 131,
https://doi.org/10.2991/978-94-6239-737-8_71

algorithms including color correction, matting, optical flow smoothing and super-resolution reconstruction are integrated to adaptively optimize image quality according to video content.

Combined with key frame extraction, motion detection and semantic analysis, intelligent editing and automatic effect adaptation are realized to reduce manual intervention. Comparative experiments combining objective indicators and subjective evaluation verify the feasibility of the scheme, and practical technical parameters and deployment suggestions are provided[3].

2 Related Work

2.1 Application of Digital Image Processing in Video Optimization

Deep learning has become an important direction in image denoising, enhancement, segmentation and reconstruction. Compared with traditional handcrafted processing, recent models can automatically learn multi-level features, but video-production deployment still requires pruning, quantization and acceleration to balance image quality and computation cost[4].

2.2 Video Key Frame and Inter-frame Processing Technology

Video is composed of continuous image sequences, and inter-frame temporal correlation determines playback smoothness. Key frame extraction, frame interpolation and motion compensation are core video processing tasks. Optical flow calculates pixel displacement for jitter elimination, slow-motion smoothing and frame rate improvement. However, blocking artifacts, occlusion and large displacement can still reduce stability. Recent flow-guided Transformer interpolation improves large-motion handling and detail preservation[5]. This paper combines multi-scale optical flow with edge-preserving filtering to reduce motion blur and artifacts.

2.3 Integration of Intelligent Technology and Video System

Image segmentation, object detection and semantic understanding promote intelligent matting, automatic target tracking and automatic video editing. Traditional green screen matting is limited by shooting conditions. Deep learning-based natural background matting has stronger universality, but edge accuracy and robustness still need improvement in hair, transparent objects and complex backgrounds[6].

3 System Architecture Design and Key Technologies

3.1 Overall Architecture Design

The system adopts a four-layer architecture. The input layer standardizes the format and parameters of imported materials. The core image processing layer undertakes image quality optimization. The video synthesis layer realizes frame arrangement and audio-video synchronization. The output encoding layer controls bit rate, evaluates image quality and exports final videos, realizing full-process integration of image processing and video production.The overall system architecture is shown in Figure 1.

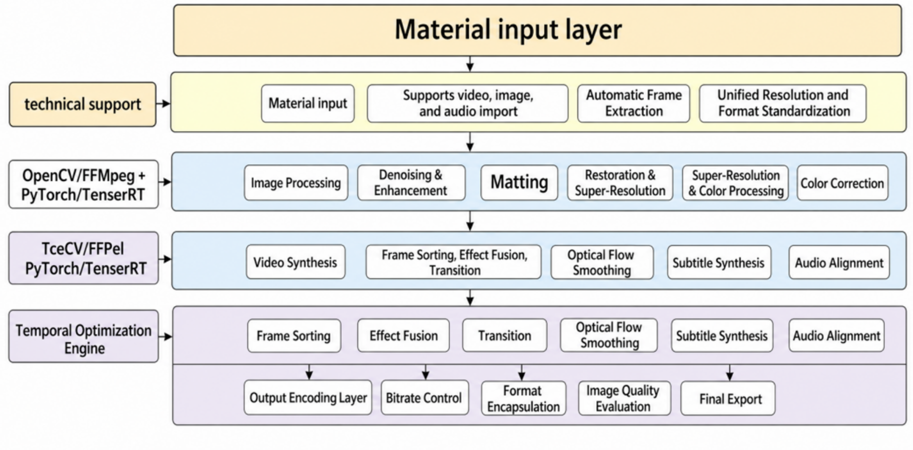

Fig. 1. Overall System Architecture Design

3.2 Key Technical Mechanism

With digital image processing as the core, the system constructs a stable video production mechanism around four major technical modules, including image preprocessing technology, image enhancement and image quality improvement technology, intelligent matting and synthesis technology, and video inter-frame processing technology. The architecture of the core technical mechanism is shown in Figure 2.

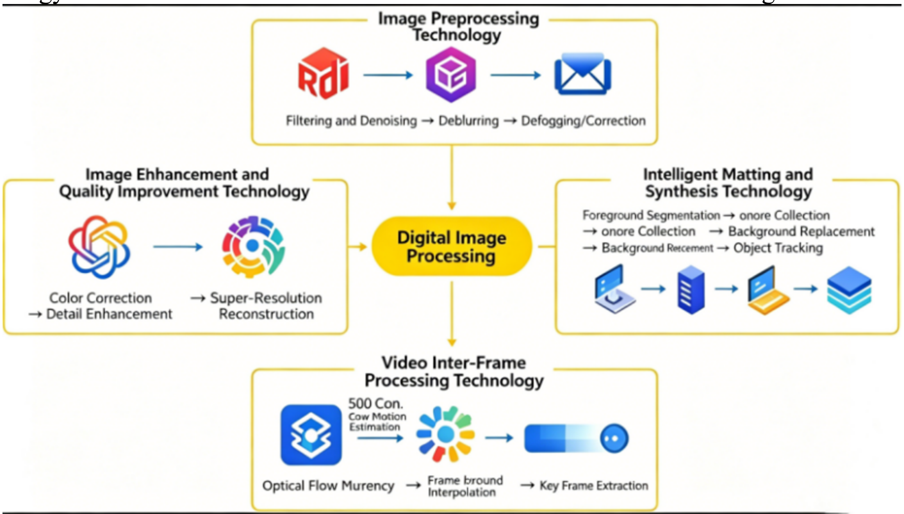

Fig. 2. Architecture of Core Technical Mechanism

4 System Implementation and Functional Modules

4.1 Core Functional Modules

The system covers the entire life cycle of video production with 9 core functional modules, as shown in Table 1.

Table 1. Core Functional Modules of the System.

No.	Core Modules	Module Description
1	Video Frame Extraction and Recombination	Extract image sequences by frame rate, supporting editing and reordering
2	Image Noise and Blur Removal	Remove noise, moiré patterns and compression artifacts, repair blurred pictures
3	Intelligent Color Correction	Automatically calibrate white balance, contrast and saturation to maintain natural tone
4	High-precision Intelligent Matting	Automatically segment foreground and optimize edges without green screen
5	Super-resolution Enhancement	Convert low-definition to high-definition, improve resolution and detail retention
6	Special Effect and Filter Synthesis	Realize filters, light and shadow, gradient special effects based on image processing
7	Inter-frame Smoothing and Frame Interpolation	Optimize video fluency and stability via optical flow frame interpolation
8	Subtitle and Image Synthesis	Realize text-to-image conversion, watermark and Logo fusion with translucent processing
9	Video Encoding and Output	Support H.264/H.265 encoding, multi-format export and image quality control

4.2 Algorithm Process Design

The standard production process is divided into four steps:First, import materials, unify format resolution and conduct frame extraction. Then perform frame-by-frame noise reduction, clarity optimization, color calibration and super-resolution restoration. Next, complete matting and special effect production, and optimize inter-frame fluency via optical flow algorithm. Finally, edit key frames, match audio mixing, and encode to export the final video.

4.3 Deployment and Operating Environment

The deployment and operating environment are configured in accordance with general system deployment specifications, as shown in Table 2.

Table 2. Deployment and Operating Environment.

Hardware/Software	Configuration Description
CPU	Intel i7 and above
GPU	NVIDIA GTX 1660 and above
Memory	16GB+
System	Windows 10 / Ubuntu 20.04
Dependency Libraries	OpenCV, PIL, PyTorch, FFmpeg

5 Experimental Design and Result Analysis

5.1 Experimental Environment

The experimental environment is configured as shown in Table 3.

Table 3. Experimental Environment

Hardware/Software	Configuration Description
Processor	Intel i9-12900K
Graphics Card	NVIDIA RTX 3090
Memory	32GB DDR4
Framework	OpenCV 4.8, PyTorch 2.0, FFmpeg 5.0

5.2 Evaluation Indicators

Objective Indicators: PSNR, SSIM, processing frame rate, matting accuracy. Subjective Indicators: Color naturalness, edge smoothness, fluency

5.3 Experimental Results and Analysis

1. Real-time Processing Performance: At 1080p resolution, the average processing frame rate exceeds 45 fps. At 4K resolution, the average frame rate remains about 30 fps because super-resolution reconstruction and optical-flow estimation increase computational overhead. With TensorRT acceleration, the 4K frame rate can reach 40 fps[5,7]. The acceleration configuration adopts FP16 mode, dynamic batch processing and a 4 GB workspace; the training settings are batch size 16, learning rate 0.0001, 100 epochs and Adam optimizer.

2. Intelligent Effect Accuracy: On the COCO-style test set, mAP@0.5 reaches 92.3%. On the public matting test set, the average Intersection over Union (mIoU) reaches 96.5%, and the segmentation accuracy of fine areas such as hair reaches 94.3%. With the AdaIN style transfer module, the average subjective naturalness score is 4.55/5 (91%)[6]. These results indicate improved single-frame quality, foreground extraction stability and visual coherence.

6 Conclusion

In conclusion, this study constructed a full-process video production framework centered on digital image processing and integrated automatic color correction, optical-flow-based inter-frame smoothing, intelligent matting, deep learning super-resolution and standardized encoding output. The workflow connects material import, frame extraction, enhancement, matting, effect synthesis, interpolation and export, reducing repeated manual adjustment and improving inter-module coordination. Experiments show that the system maintains more than 45 fps at 1080p and about 30 fps at 4K; with TensorRT FP16 acceleration, the 4K frame rate can reach 40 fps. The detection mAP@0.5 reaches 92.3%, the matting mIoU reaches 96.5%, the subjective naturalness score reaches 4.55/5, and editing efficiency is improved by 60%. These results demonstrate that the proposed method improves image clarity, color consistency, foreground extraction accuracy and temporal fluency, making it suitable for short-video production, film and television post-production and educational videos. The remaining limitation is that complex occlusion and very large motion may reduce edge stability, so future optimization should focus on lightweight deployment and robustness.

Acknowledgments

Thanks to Harbin Institute of Information Technology for providing the experimental platform, hardware equipment and technical support for this research, which makes the smooth development of this project possible.

References

1. Su Yuanting. Research on Video Production Technology Based on Digital Image Processing[J]. Digital Communication World, 2025(08):19-21. DOI:10.3969/J.ISSN.1672-7274.2025.08.007.
2. Zhou Jiawei. Research on Image and Video Recoloring Algorithm[D]. Qinghai University, 2024. DOI:10.27740/d.cnki.gqhdx.2024.000987.
3. Wang Lupei. Research on the Application of Digital Media Technology in Interactive International Chinese Teaching[D]. Xi'an Shiyou University, 2025. DOI:10.27400/d.cnki.gxasc.2025.000370.
4. Archana, R., Jeevaraj, P. S. E. Deep learning models for digital image processing: a review[J]. Artificial Intelligence Review, 2024, 57:11. DOI:10.1007/s10462-023-10631-z.
5. Gao, P., Tian, H., Qin, J. Video Frame Interpolation with Flow Transformer[C]. Proceedings of the 31st ACM International Conference on Multimedia, 2023:1933-1942. DOI:10.1145/3581783.3612440.
6. Huang, L., Liu, X., Wang, X., Li, J., Tan, B. Deep Learning Methods in Image Matting: A Survey[J]. Applied Sciences, 2023, 13(11):6512. DOI:10.3390/app13116512.
7. Liu, H., Ruan, Z., Zhao, P., Dong, C., Shang, F., Liu, Y., Yang, L., Timofte, R. Video super-resolution based on deep learning: a comprehensive survey[J]. Artificial Intelligence Review, 2022, 55:5981-6035. DOI:10.1007/s10462-022-10147-y.