International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Media Controller with FIR Filter

Chaitanya Singh[1], Tejashwini Rath[2], Sneha Singh[3], Dr. Swati Agrawal [4]

[4] (Associate Professor)
Department of electronics and telecommunication Bhilai
Institute of technology Durg (C.G.), India

ABSTRACT :

This research introduces a software-driven media control system that integrates real-time hand gesture recognition with Finite Impulse Response (FIR) audio filtering for enhanced multimedia interaction. The system uses computer vision techniques, leveraging OpenCV and MediaPipe libraries, to detect and interpret hand gestures for touchless media control. Simultaneously, a Python-implemented FIR filter processes audio to reduce noise and improve clarity. The modular, multithreaded architecture supports seamless integration of gesture control and audio processing, offering a flexible alternative to hardware-based solutions. This approach has potential applications in smart homes, assistive technologies, and accessibility systems.

Keywords: Hand Gesture Recognition, FIR Filter, Computer Vision, Audio Signal Processing, Human-Computer Interaction, MediaPipe, OpenCV, Real-Time Processing, Touchless Interface, Digital Signal Processing

1. Introduction

The field of Human-Computer Interaction (HCI) continues to evolve towards more intuitive, touchless interfaces. With increasing demand in smart environments, virtual reality, and assistive applications, systems that combine gesture recognition and audio enhancement are gain relevance Conventional media control relies on physical input, which may be inconvenient or inaccessible for certain users. While gesture-based systems exist, many lack integration with real-time audio enhancement, reducing their utility in multimedia contexts.

This study proposes a unified system that combines computer vision-based hand gesture recognition and FIR audio filtering. Its software-based design provides cost-effective, adaptable, and scalable interaction, operable on standard consumer hardware.

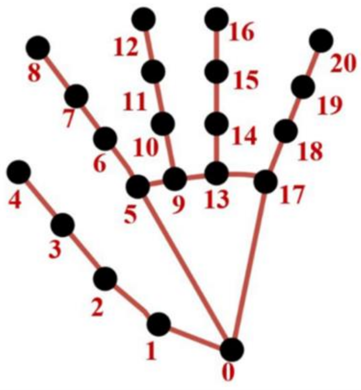
2. Methodology

2.1 System Architecture

The system architecture includes four key components: Hand Gesture Recognition Module, Gesture-to-Command Mapping System, Audio Processing Pipeline, and Integration Controller. This modular design allows independent development and efficient communication between components.

2.2 Hand Gesture Recognition Implementation

The gesture recognition module uses MediaPipe's hand tracking model to identify 21 hand landmarks from video captured at 30fps. The process involves frame preprocessing, landmark extraction, gesture classification through rule-based analysis, and temporal filtering to ensure gesture stability before command execution.
Recognized Gestures Include:
- Play/Pause: Closed fist
- Volume Up: fingers raised
- Volume Down: fingers lowered
-NextTrack:Right swipe
- Previous Track: Left swipe



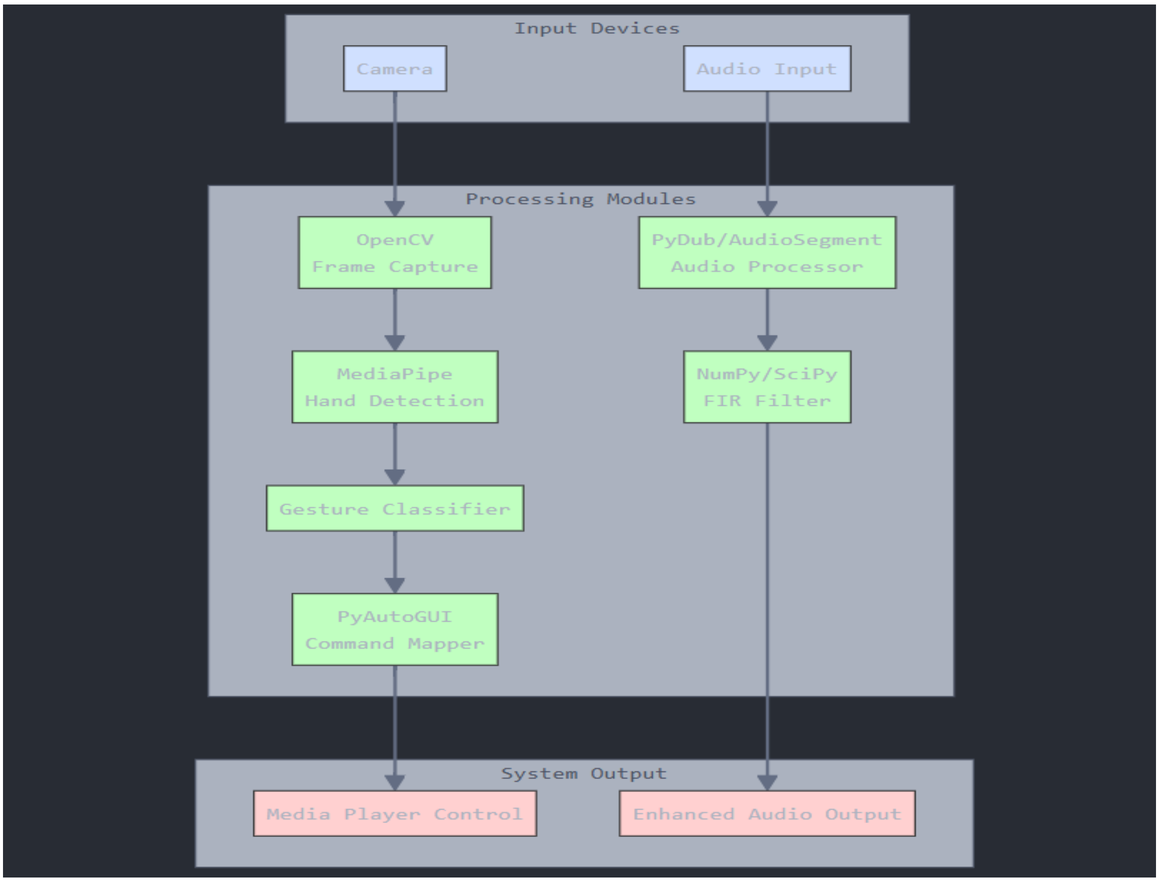
2.3 FIR Filter Design and Implementation

The audio module implements a low-pass FIR filter using the windowing method. Audio is captured via PyAudio, processed into arrays, filtered using linear convolution with designed coefficients, and then output in real-time. The filter is designed with 101 taps and implemented using SciPy's signal processing functions.

2.4 System Integration and Synchronization

Integration is achieved through Python's multithreading, enabling parallel execution of gesture and audio modules without performance degradation.

2.5 System Block Diagram



2.6 Performance Optimization Strategies

To maintain real-time processing, the following strategies are implemented:

- Efficient buffer management
- Frame rate adjustment
- Resource usage monitoring
- Exception handling for robustness

3. Experimental Results and Discussion

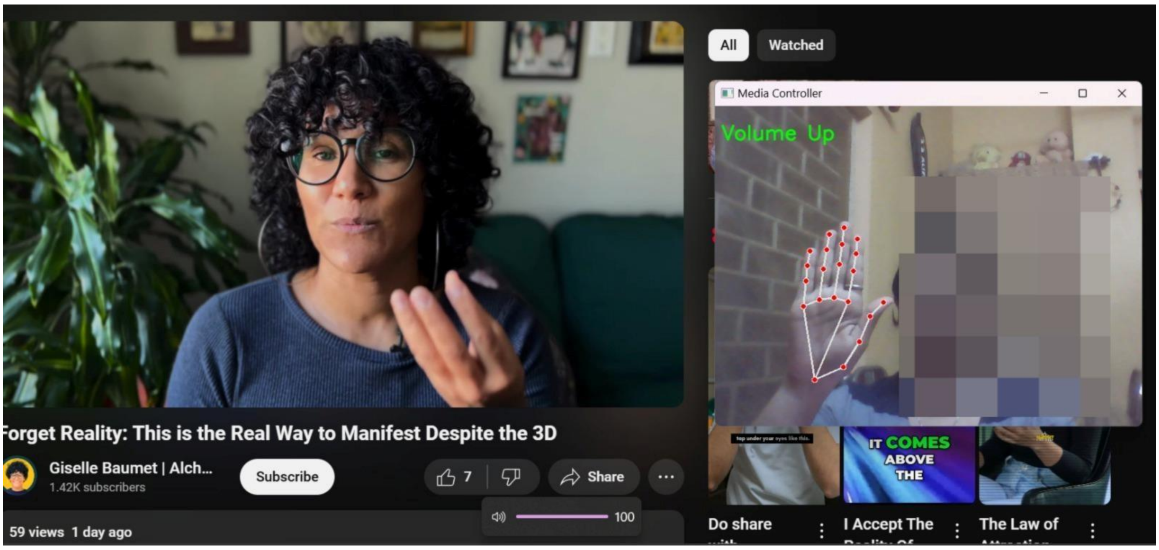
3.1 Experimental Setup

Testing was conducted on a standard laptop with an R5 processor, integrated webcam, and 8GB RAM, under different lighting and audio conditions.

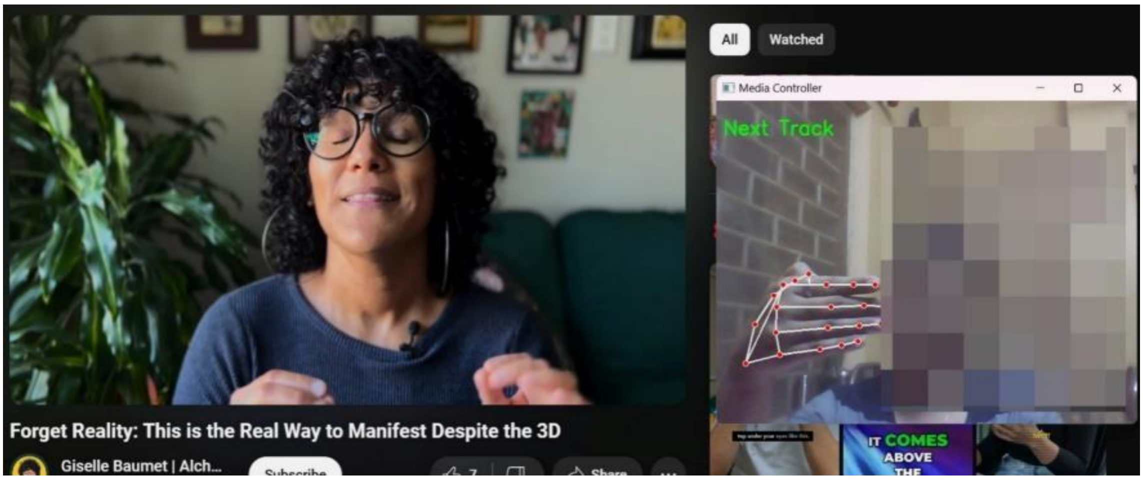
3.2 Gesture Recognition Observations

The system was able to recognize gestures under varied environmental scenarios. Observations indicated reliable detection in most practical use cases, with occasional challenges in extremely low-light or rapidly changing motion environments.

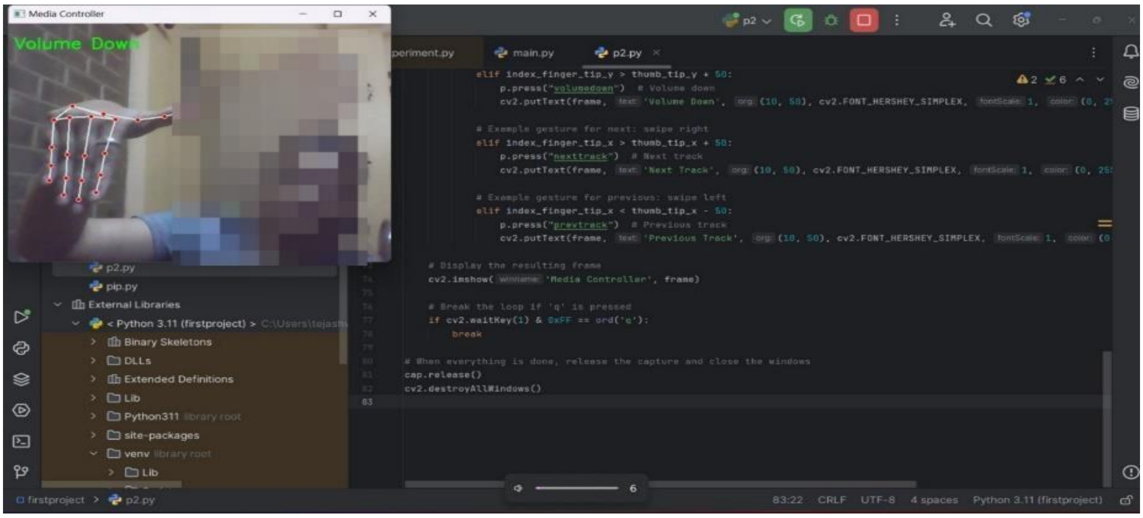
1. The hand is in a position that indicates an upward gesture . The system emulates an "up" action, typically used for volume increase or navigating upwards in a media control context.



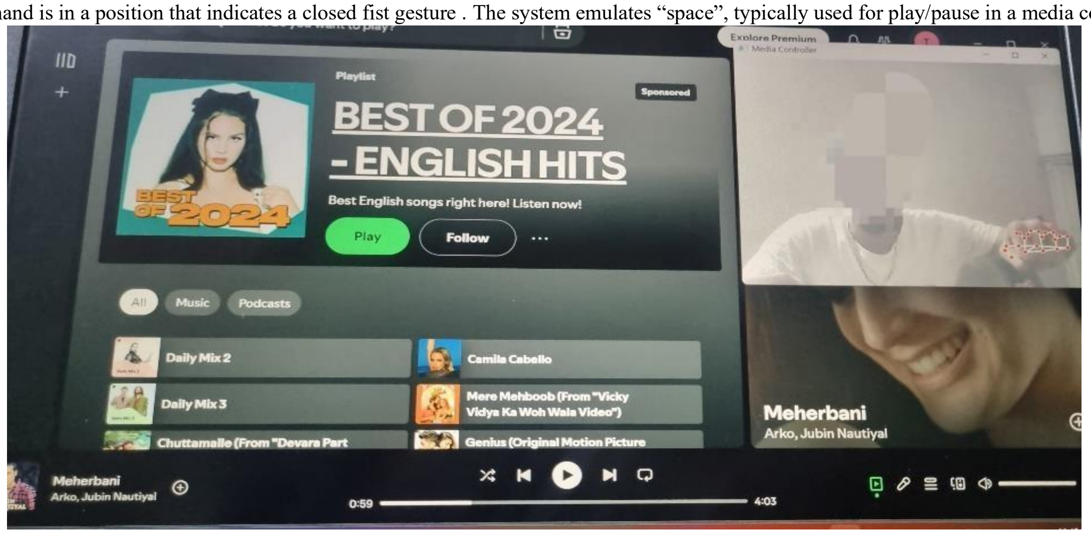
2. The hand is in a position that indicates a right ward gesture . The system emulates a “next” action, typically used for playing next track navigating media control context



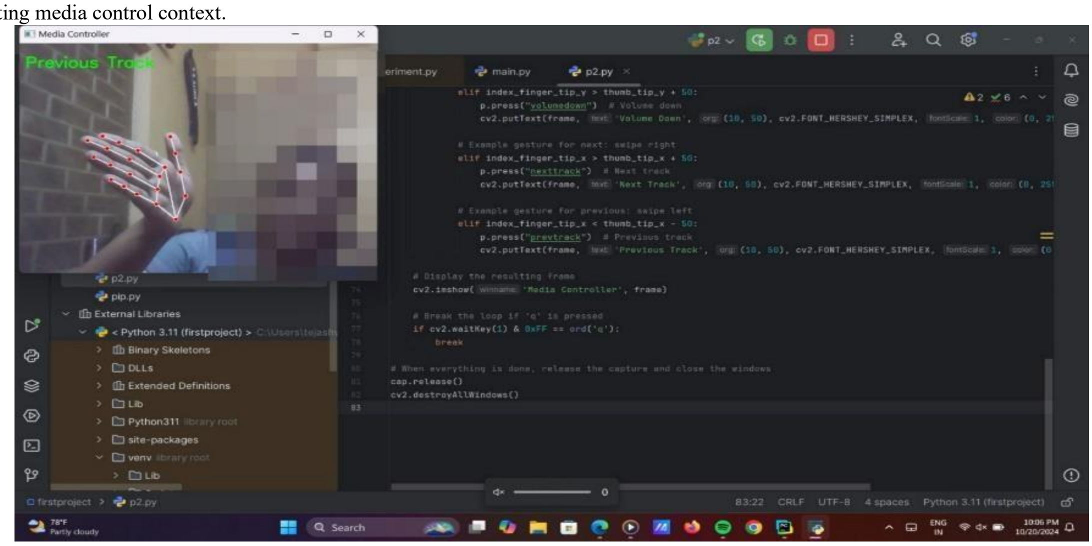
3. The hand is in a position that indicates a downward gesture . The system emulates a "down" action, typically used for volume decrease or navigating downwards in a media control context.



4. The hand is in a position that indicates a closed fist gesture . The system emulates “space”, typically used for play/pause in a media control context.

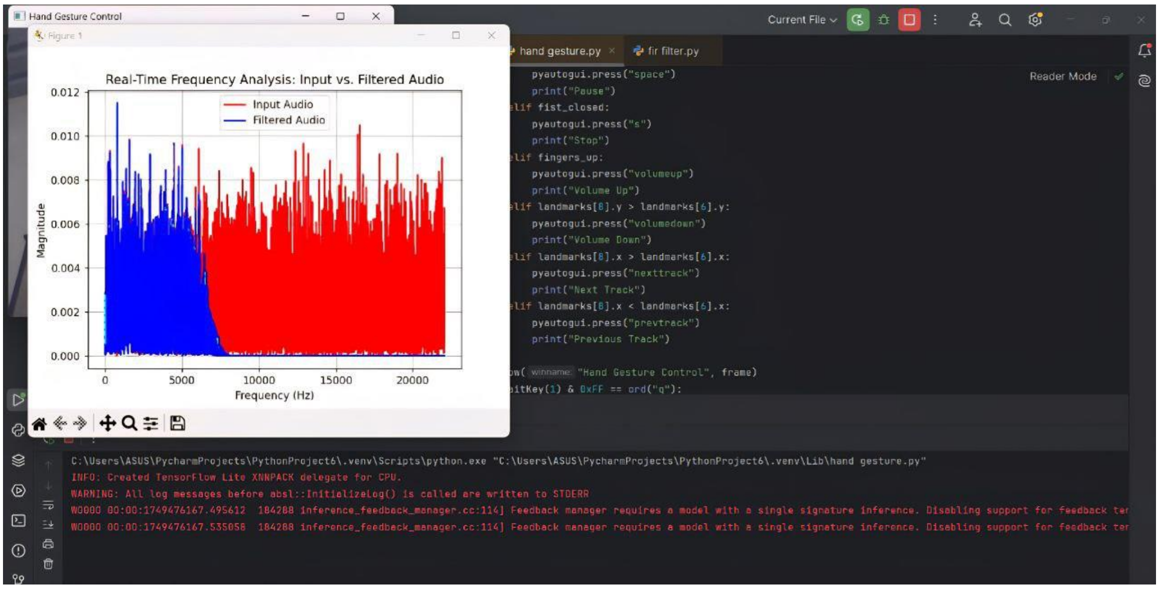


5. The hand is in a position that indicates a left ward gesture . The system emulates a “previous” action, typically used for playing previous track navigating media control context.



3.3 Audio Processing Observations

The FIR filter provided smoother audio output and reduced background noise. The filter maintained consistent performance during real-time processing.



Conclusion

The project demonstrates the technical feasibility of integrating hand gesture recognition and audio filtering in a real-time media control application. The multithreaded implementation ensures efficient use of computing resources.

Future Work

Potential future enhancements include:

1. IoT and Smart Device Integration- Integrating the system with IoT ecosystems can enable gesture-based control of smart home appliances, enhancing convenience and automation.
2. Multi-Modal Interaction- Expanding the system to include voice commands or facial expression detection can create a richer and more intuitive user interaction experience.
3. Machine Learning Integration-Incorporating deep learning models can enhance gesture recognition accuracy and allow personalized gesture sets tailored to individual users.

REFERENCES

1. **Rakh, A., Dherange, V., Shelke, V., Londhe, S., & Akhare, P.**
   "Real-Time Volume Control Using OpenCV", *Journal of Transactions in Systems Engineering*, Vol 2 No 2, pp 256–264 (2024). DOI: 10.15157/JTSE.2024.2.2.256-264 (mdpi.com, journals.tultech.eu)
2. **Patil, V., Sutar, S., Ghadge, S., & Palkar, S.**
   "Gesture Recognition for Media Interaction: A Streamlit Implementation with OpenCV and MediaPipe", *IJRASET*, Vol 7 Issue 4 (2023). DOI: 10.22214/ijraset.2023.55775 (ijraset.com)
3. **Sruthi, S. & Swetha, S.**
   "Hand Gesture Controlled Presentation using OpenCV and MediaPipe", *International Journal of Engineering Technology and Management Sciences*, Vol 7 Issue 4, pp 338–342 (2023). DOI: 10.46647/ijetms.2023.v07i04.046 (ijetms.in)
4. **Budiman, S. N., Lestanti, S., Evvandri, S. M., & Putri, R. K.**
   "Finger Gesture Recognition to Control Volume on Computer using OpenCV and MediaPipe", *Antivirus: Jurnal Ilmiah Teknik Informatika*, Vol 16 Issue 2 (2024). DOI: 10.35457/antivirus.v16i2.2508 (ejournal.unisbablitar.ac.id)
5. **Uludağli, M.Ç. & Acartürk, C.**
   "Intelligent Gesture Recognition Based on Screen Reflectance Multi-Band Spectral Features", *Sensors*, Vol 24, 5519 (2024). DOI: cross-reference on MDPI (mdpi.com)
6. **Yadav, K. & Bhattacharya, J.**
   "Real-Time Hand Gesture Detection and Recognition for Human Computer Interaction", in *Intelligent Systems Technologies and Applications*, Vol 384, pp 559–567 (2016, still relevant benchmark). DOI: 10.1007/978-3-319-23036-8_49 (link.springer.com)

7. **Karsh, R. et al.**
"mIV3Net: modified inception V3 network for hand gesture recognition", *Multimedia Tools and Applications*, Vol 183, pp 10587–10613 (2024). DOI: 10.1007/s11042-023-15865-1 (dl.acm.org)
8. **Chelali, F. et al.**
"Textural feature descriptors for a static and dynamic hand gesture recognition system", *Multimedia Tools and Applications*, Vol 183, pp 8165–8187 (2024). DOI: 10.1007/s11042-023-15410-0 (dl.acm.org)
9. **Gomaa, A. Z. et al.**
"SynthoGestures: A Novel Framework for Synthetic Dynamic Hand Gesture Generation for Driving Scenarios", Adjunct Proc. of ACM UIST '23, pp 1–3 (2023). DOI: 10.1145/3586182.3616635 (dl.acm.org)
10. **Sung, G., Sokal, K., Uboweja, E., Bazarevsky, V., Baccash, J., Bazavan, E., Chang, C.-L., & Grundmann, M.**
"On-device Real-time Hand Gesture Recognition", arXiv (2021). DOI: 10.48550/arXiv.2111.00038 (arxiv.org)
11. **Li, Z., Liang, C., Wang, Y., Qin, Y., Yu, C., Yan, Y., Fan, M., & Shi, Y.**
"Enabling Voice-Accompanying Hand-to-Face Gesture Recognition with Cross-Device Sensing", arXiv (2023). DOI: 10.48550/arXiv.2303.10441 (arxiv.org)
12. **Sen, A., Mishra, T. K., & Dash, R.**
"Deep learning based Hand gesture recognition system and design of a Human-Machine Interface", arXiv (2022). DOI: 10.48550/arXiv.2207.03112 (arxiv.org)
13. **Zhou, B., Aiskovich, M., & Guven, S.**
"Acoustic Sensing-based Hand Gesture Detection for Wearable Device Interaction", arXiv (2021). DOI: 10.48550/arXiv.2112.05986 (arxiv.org)
14. **Ibrahim, A., El-Refai, A., Ahmed, S., Aboul-Ela, M., Eraqi, H. M., & Moustafa, M.**
"Pervasive Hand Gesture Recognition for Smartphones using Non-audible Sound and Deep Learning", arXiv (2021). DOI: 10.48550/arXiv.2108.02148 (arxiv.org)
15. **Binh, N. D.**
"Sound Waves Gesture Recognition for Human-Computer Interaction", in *Context-Aware Systems and Applications*, LNCS Vol 165, pp 41–50 (2016). DOI: 10.1007/978-3-319-29236-6_5 (link.springer.com)