Dissertation Abstract

Manual assembly training traditionally relies on experienced operators to guide trainees through task demonstrations, trials, evaluations, and discussions. This method, while effective, is limited by the availability of experts. Current virtual training systems (VTS) focus on delivering rich multimedia content for task demonstrations, reducing dependence on experts. However, these systems often lack automated, comprehensive, augmented feedback for trainees.

This research introduces EXAMINER (**EX**pert Independent Manual **A**sse**M**bly V**I**rtual Trai**NER**), a system that objectively evaluates and provides feedback on trainees' motor and cognitive skills in manual assembly tasks. By automating feedback, EXAMINER enhances training accessibility and reduces reliance on experts. This study explores the digitization of human skills, objective measurement techniques, and the integration of these elements into an effective training system. The proposed system is evaluated for its ability to deliver appropriate feedback based on trainee performance, aiming to improve training outcomes and adoption rates.

The resulting framework consists of the following components: skill digitization, skill comparison, feedback provider, and multimedia training material. The implementation focuses on the first three components, ensuring their seamless integration. The framework implementation utilized methodologies for skill digitization using a video camera, employing standard and contemporary techniques such as deep learning in computer vision for human pose estimation, recurrent neural networks for activity recognition, and computer vision for contextual sensing. Each underlying subcomponent shows promising performance.

The digitization process is critical because it is the foundation for subsequent skill analysis and comparison between trainees and experts. In analyzing these operations, the study takes a novel approach, employing algorithms such as edit distance and dynamic time warping to identify and quantify skill differences. This methodology enables a more in-depth understanding of manual assembly cognitive and motor skill differences.

Another contribution of this research is the introduction of the I-MA task data model. This model enhances the framework's adaptability across diverse training scenarios and revolutionizes how information is systematically organized and utilized within I-VTS. The modular design of the framework, emphasizing interconnected yet distinct components, significantly enhances system flexibility and scalability, catering to a wide range of training needs and environments.

In summary, this research offers a comprehensive, flexible, and efficient I-VTS framework, representing a significant leap forward in virtual training systems. The framework utilizes advanced digitization techniques, detailed skill analysis, and user-friendly augmented feedback to address current gaps in I-MA training and establish a new standard for future developments in the field.

Keywords: Deep learning, Computer vision, Manual assembly, Virtual training, Industry 4.0