Procedia Manufacturing
Volume 5, 2016, Pages 106–121
44th Proceedings of the North American Manufacturing Research Institution of SME http://www.sme.org/namrc

Recognizing Gaze-Motor Behavioral Patterns in Manual Grinding Tasks

Gregory Bales[1], Jayanti Das[1], Barbara Linke[1], and Zhaodan Kong[1*]

[1]The Department of Mechanical and Aerospace Engineering, University of California, Davis, Davis, CA, USA
glbales, jydas, bslinke, zdkong@ucdavis.edu

Abstract
This paper reports our progress in developing techniques for "parsing" raw gaze and force data from manual grinding tasks into a principled model. A grinding task, though simple, requires the practitioner to combine elements from the large repertoire of her skillset. Based on the joint, gaze, and force data collected from a series of experiments, and by extending existing scanpath methods, we develop a visualization method called Gaze-Motor Space-Time Cube (GMSTC), which can help us gain insight into the joint gaze-motor routine existing in complex manual tasks. For instance, there exists a strong correlation between the spectra of a subject's fixation and force distributions. Such insight might be hard to extract through an examination of either the gaze or the force data separately. Furthermore, by comparing data obtained from operators with different levels of skill, we are able to quantitatively describe characteristics of human manual skill. For instance, we find that an experienced subject exhibits longer fixation durations and smaller fixation variations than an intermediate one. A detailed understanding of gaze-motor behavior broadens our knowledge of how a manual task is executed. Our results help to provide this extra insight, and have implications in the way in which knowledge and manual expertise is transferred from one generation of practitioners to the next.

Keywords: Human factors, man-machine interactions, data visualization, attention, grinding

1 Introduction

This paper studies human manual skills involved in manufacturing tasks. Despite the current proliferation of highly automated factories, humans are still an integral part of most manufacturing processes, not only as the supervisors of machines, but also as the performers of manual tasks, such as fine polishing. The skills involved in these manual tasks are largely procedural rather than declarative, meaning that they cannot be easily articulated by the individuals (Goldstein, 2014). A lack of understanding of these manual skills may prolong the transfer of this knowledge from generation to generation. It may also impede the development of novel manufacturing themes, such as cloud

* Corresponding author.

 Selection and peer-review under responsibility of the Scientific Programme Committee of NAMRI/SME

doi:10.1016/j.promfg.2016.08.011

manufacturing (Lee et al., 2015) and human-robot collaborations (Wojtara et al., 2009). To elaborate on the last point, in order to establish an effective dialog within a human-robot team, the robot must be able to understand the intention of the human by inferring from her behavior. In this paper, we focuses on analyzing the visual-attention-motor behavior in the context of manual grinding tasks (Odum et al., 2014), (Klocke et al., 2011). Our study, on the one hand, improves our understanding of complex manual skills that are beyond our everyday activities and, on the other hand, enables models of human skills which are crucial for building future industrial robots that can understand human intentions.

Visual attention is a remarkable human capability of reducing the huge amount of visual data entering our eyes into a manageable level. Given the nonuniform distribution of photoreceptors in our retina (with a high-resolution central fovea and a low-resolution periphery), the vision system cannot reconstruct a detailed 3D model of the scene. Instead, visual attention guides our eyes to important parts of the scene and puts the object of interests in the fovea (Borji and Itti, 2013). In the past few decades, a considerable amount of experimental and computational research has been conducted to understand and model visual attention process (Borji and Itti, 2013), (Rayner, 2009). Aside from being an interesting scientific problem itself, visual attention also attracts significant interest from engineering fields especially but not limited to robotics (Begum and Karray, 2011), (Siagian and Itti, 2009), computer vision (Guo and Zhang, 2010), and human-robot collaborations (Bauer et al., 2008), (Wojtara et al., 2009).

Visual attention can be roughly divided into two categories, bottom-up attention and top-down attention (Borji and Itti, 2013), (Posner, 2011). Top-down or voluntary attention is our ability to intentionally attend to something. It is a goal-driven process based on aspects such as tasks, knowledge, expectations, and memory. In a manual grinding scenario, for instance, it may involve finding a sample. Bottom-up or reflexive attention is a stimulus-driven process in which a salient sensory event captures our attention. Such an event might be a crack appearing on the surface of the grinding sample. A popular bottom-up attention model is the saliency map model, based on intrinsic image features such as color, orientation, and intensity (Itti et al., 1998), (Borji and Itti, 2013). However, it has been argued and demonstrated that at least with purposeful actions, top-down attention rather than bottom-up attention plays a much dominant role (Land, 2009), (Hayhoe and Ballard, 2005), (Borji et al., 2014). Some interesting observations from the studies of top-down attention in natural behavior, such as making a cup of tea (Land, 2009) and making a sandwich (Borji et al., 2014), include: the first movement of the hand toward the object of interest occur before the first saccade; at the end of each object-related action, the eyes move to the next object before the manipulation is over. In this paper, we focus on top-down attention involved in manual manufacturing tasks.

Gaze, a coordinated motion of the eyes and the head, is often used in research as a proxy for visual attention[†]. The majority of studies are concerned with two types of gaze movements, saccades, fast ballistic movements in which the gaze moves rapidly from one location to another, and fixations, the period during which the gaze is stationary and useful information is collected (Duchowski, 2007), (Cristino et al., 2010). Sequences of time-ordered gaze movements composed of fixations and saccades, called scanpaths, have been used for visualizing and analyzing gaze since the early 1970s (Noton and Stark, 1971), (Duchowski et al., 2010). Scanpaths facilitate the quantitative comparison of gaze data collected from different subjects and trials. One algorithm that can be used for such comparisons is based on a metric called Levenshtein distance, more commonly referred to as string editing distance. The string edit algorithm (SEA) compare scanpaths by solving an optimization problem with unit costs assigned to three different character operations: deletion, insertion, and substitution (Privitera and Stark, 2000), (Duchowski et al., 2010). Some issues with SEA are that it cannot differentiate between close and distance regions and it does not take into account fixation

[†] It should be pointed out that gaze, which can be tracked by an eye-tracker, corresponds to the overt movements of the eyes, not the covert movements of visual attention. Thus, a very important assumption that is usually accepted in visual attention research is that the attention is linked to the gaze direction, even though this might not be always true (Duchowski, 2007).