A Motion Capture System Framework for the Study of Human Manufacturing Repetitive Motions

Francis Mendez  
*Texas State University - San Marcos, fm16@txstate.edu*

David Wierschem  
*Texas State University - San Marcos, dw50@txstate.edu*

Jesus Jimenez  
*Texas State University - San Marcos, Jesus.Jimenez@txstate.edu*

Follow this and additional works at: [https://digitalcommons.georgiasouthern.edu/pmhr_2018](https://digitalcommons.georgiasouthern.edu/pmhr_2018)  
Part of the [Industrial Engineering Commons](https://digitalcommons.georgiasouthern.edu/indeng), [Operational Research Commons](https://digitalcommons.georgiasouthern.edu/operres), and the [Operations and Supply Chain Management Commons](https://digitalcommons.georgiasouthern.edu/opsmgmt)

**Recommended Citation**  

This research paper is brought to you for free and open access by the Progress in Material Handling Research at Digital Commons@Georgia Southern. It has been accepted for inclusion in 15th IMHRC Proceedings (Savannah, Georgia. USA – 2018) by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.
A Motion Capture System Framework for the Study of Human Manufacturing Repetitive Motions

Francis Mendez  
Department of Computer Information Systems and Quantitative Methods  
Texas State University  
San Marcos, TX USA  
fm16@txstate.edu

David Wierschem  
Department of Computer Information Systems and Quantitative Methods  
Texas State University  
San Marcos, TX USA  
dw50@txstate.edu

Jesus Jimenez  
Ingram School of Engineering  
Texas State University  
San Marcos, TX USA  
jesus.jimenez@txstate.edu

Abstract—Our research proposes a framework to obtain and analyze real time data concerning the dynamic and natural motion of individuals in manufacturing-like processes that involve human labor. The framework that we propose consists of five main components: a tracking system, a system of sensors, a processor that collects time series data, data processing and discovery module, and an alert or reporting component. Using motion capture cameras, data is collected on a variety of human subjects performing simulated labor-intensive manufacturing operations. This data is analyzed for identification of actual and optimal activity motions. This project has significant potential impact for contribution and advancement of the material handling, logistics, and supply chain industries. This simulation process will enable a company to modify human operations that are non-value-added from the activity process.

Keywords—smart manufacturing, analytics, motion capture

I. INTRODUCTION

This manuscript proposes and implements a framework to study human motion that relies on modern approaches to data collection and data analysis. The main objective of the study is to improve workers’ safety and efficiency within a manufacturing environment.

At the beginning of the 20th Century, Frederick Winslow Taylor introduced a scientific approach to management. Within this context, “scientific” means based on experimentation rather than on tradition or opinions. Taylor’s approach is known as “time and motion study” and consists of breaking down any physical work task into elementary motions. The objective was to seek the “one best way” to perform a given movement [1][2][3].

During this period, in which the majority of manufacturing processes were based on human activity, efforts for improvement included assembly line process specialization. This application of scientific management was successful in increasing the industrial output of the United States [4].

Taylor prepared the path for industrial acceptance of John Dewey’s theory that man is merely a machine. Increasing efficiency further has been achieved by replacing human workers with automation (i.e., robotics). However, there still remain manufacturing processes which are, and always will be, best performed by human workers [4].

For this reason, time and motion studies are still of great relevance in multiple fields [5][6][7]; and modern tools have been devised to make its implementation more comprehensive, efficient and precise [8][9][10].

II. RELATED WORK

A. Technologies

Today a host of technologies exist to support continued efforts for analyzing human motions. Because of the reduced portion of the manufacturing process conducted by humans, most of these technologies are being applied to non-manufacturing areas such as: ergonomics, computer gaming, entertainment (movie animation), posture analysis, and gait analysis.

One of the more advanced technologies is that of motion capture. Motion capture is defined as the process of recording a live motion event and translating it into mathematically-useable signals [11]. There are a variety of specific technologies that meet this definition including: magnetic, mechanic, optical, acoustic, and inertial.

Magnetic motion capture systems place sensors on the body to measure low-frequency magnetic fields generated by a transmitter source. A transmitter source consists of three perpendicular coils that emit a magnetic field when a current is applied. The current sent to these coils generates three mutually perpendicular fields during each measurement cycle. The 3D sensors measure the strength of those fields. Sensors and source are connected to a processor that calculates position and orientation of each sensor. One advantage of magnetic systems is that they do not suffer from line-of-sight issues. A disadvantage of magnetic systems is that magnetic fields decrease in power rapidly as the distance from the generating source increases and so they can easily be disturbed by magnetic materials.

Mechanical trackers utilize rigid or flexible goniometers. Goniometers are angle measuring devices. These goniometers are placed in general correspondence to the joints of the user. They provide joint angle data to kinematic algorithms used to determine body posture. There are some disadvantages in using mechanical systems. For instance, it is difficult to align the goniometer with body joints. Also, positioning the goniometers
on soft tissue can be problematic. Goniometers placed on soft tissue allows their position relative to the body to change as motion occurs.

Acoustic tracking systems can determine position through time-of-flight of ultrasonic pulses and triangulation or phase coherence. Transmitters can either be placed on a body segment or fixed in the measurement volume. The main disadvantage of acoustic tracking systems is that the physics of sound limit the accuracy, update rate and range. In addition, a clear line-of-sight must be maintained and reflections of sound can also interfere.

Inertial sensors use the property of bodies to maintain constant translational and rotational velocity, unless disturbed by forces or torques, respectively. Miniaturized and micromachined sensors make practical inertial tracking; in particular, rate sensors and accelerometers. Rate gyroscopes measures angular velocity, and if integrated over time provides the change in angle with respect to an initially known angle. Accelerometers measure acceleration, including gravitational acceleration. Velocity and position can also be determined. A disadvantage of inertial systems is that noise and bias errors, associated with small and inexpensive sensors, make it impractical to track orientation and position for long time periods if no compensation is applied.

Optical sensing encompasses a large and varying collection of technologies. Image-based systems determine position by using multiple cameras to track predetermined points (markers) on the subject’s body segments, aligned with specific bony landmarks. The position is estimated through the use of multiple 2D images of the working volume. Stereometric techniques correlate common tracking points on the tracked objects in each image and use this information along with knowledge concerning the relationship between each of the images and camera parameters to calculate position.

The markers can either be passive (reflective) or active (light emitting). Reflective systems use infrared (IR) LED’s mounted around the camera lens, along with IR pass filters placed over the camera lens and measure the light reflected from the retroreflective markers. It is the most flexible and common method used in the industry.

Optical systems based on pulsed-LED’s measure the infrared light emitted by the LED’s placed on the body segments. This technique uses LED markers connected by wires to the motion capture suit. A battery or charger pack must also be worn by the subject. Also, camera tracking of natural objects without the aid of markers is possible, but in general less accurate (i.e., Kinect). It is largely based on computer vision techniques for pattern recognition and often requires high computational resources.

Structured light systems use lasers or beamed light to create a plane of light that is swept across the image. They are more appropriate for mapping applications than dynamic tracking of human body motion. The main disadvantages of optical systems is that they suffer from occlusion (line-of-sight) problems whenever a required light path is blocked and interference from other light sources or reflections may also be a problem.

B. Human Conditions

The human body is a complex combination of rigid and soft components that both allows for work to be accomplished through motion and flexibility to allow for infinite variations of movement. While these characteristics make the human body extremely adaptable and malleable it makes measurement and evaluation of motions extremely difficult.

C. Body dynamic simulations

Rigid body dynamics: rigid-body dynamics is a subtopic of classical mechanics involving the use of Newton's laws of motion to solve for the motion of rigid bodies moving in 1D, 2D, or 3D space. We may think of a rigid body as a distributed mass, that is, a mass that has length, area, and/or volume rather than occupying only a single point in space. Rigid body models have application in stiff strings (modeling them as disks of mass interconnected by ideal springs), rigid bridges, resonator braces, and so on. Rigid bodies do not change, that is, the relative distance of two points on the object is fixed [12].

Soft body dynamics: soft body dynamics is a field of computer graphics that focuses on visually realistic physical simulations of the motion and properties of deformable objects (or soft bodies). The applications are mostly in video games and film. Unlike in simulation of rigid bodies, the shape of soft bodies can change, meaning that the relative distance of two points on the object is not fixed. While the relative distances of points are not fixed, the body is expected to retain its shape to some degree (unlike a fluid). The scope of soft body dynamics is quite broad, including simulation of soft organic materials such as muscle, fat, hair and vegetation, as well as other deformable materials such as clothing and fabric. Generally, these methods only provide visually plausible emulations rather than accurate scientific/engineering simulations, though there is some crossover with scientific methods, particularly in the case of finite element simulations. (http://joeyfladderak.com/lets-talk-physics-soft-body-dynamics/)

D. Skeletal motion vs topical (skin) motion

The majority of motion research focuses on the skeletal movements of a soft body, attempting to minimize the variability of movement associated with skin deformation. The ancillary motion between the motion of the rigid bone and the soft body skin and muscles is referred to as the “skin motion artifact”. Significant literature exists for identifying ways for capturing the variability in motion attributable to the skin motion artifact to remove it from analysis. The purpose of this practice is to better identify bone position and orientation and improve joint kinematic estimates. However, the applied workplace application of this research is more interested in the human motions as they actually occur, including the associated skin artefacts.
III. DESIGN AND IMPLEMENTATION

The proposed framework consists of a motion capture (MoCap) environment and data collection system, a data preprocessing and storage module, a data-intensive analytics module, and a reporting module. Figure 1 shows the system configuration. Figure 2 shows the proposed framework.

The system is further broken down into sub-processes as displayed below in Figure 2.

A. MoCap environment and data collection

The MoCap environment is setup to simulate a manufacturing work space that requires human participation in an assembly process. Human subjects (workers) are outfitted with special biometric capturing clothing and reflective markers.

The data capturing environment consists of six Qualisys Oqus infrared cameras and one Opus 210c video camera positioned around a simulated manufacturing operation board. The infrared cameras are positioned relative to the human subject to maximize the potential that all motion markers are visually accessible by at least three cameras at all times. They are set to record human motions as a series of real time x-y-z coordinates as a function of time (>1000 frames per second) associated with the positions of the reflective markers. The investigators are unaware of any other M&MH research efforts with this level of accuracy.

In addition to the subject and movement data, other documented experimental factors include those associated with the task itself including: motion distance and direction, number of repetitions, and instance duration.

Finally, biomedical data is also collected by special biomedical clothing worn by the subject. This clothing captures information such as heart rate, HRV (to estimate stress and fatigue), heart rate recovery, breathing rate (RPM), minute ventilation (L/min), activity intensity, and peak acceleration.

In total, three types of data are collected: descriptive, biomedical, and motion. Once collected the data must be formatted and labeled to be streamed to the analytical module. The descriptive and biomedical data are coded into an excel spreadsheet. The motion data requires further processing. Using the MoCap manager software, the experiment team must review the motion capture data by ensuring proper association.
of coordinate data with the correct marker. Once completed, the software builds a computer model (i.e., virtual model) of the operator performing the defined moves. Upon completion of this step, the experiment team must verify and validate the motion capture data by comparing the virtual model of the moves to the actual moves obtained from video recordings. This step helps to identify and mark spurious data events such as sneezes, scratches, or other motions outside of those expected.

B. Data Cleaning and Transformation

Once collected, the data is saved in a .C3D format. The C3D format is a public domain, standard binary file format used to record synchronized 3D and analog data. It is supported by all major 3D Motion Capture System manufacturers, and other companies in the Biomechanics, Motion Capture, and Animation Industries [13]. The file is capable of being manipulated using the Python package c3d 0.3.0 [14] and R [15].

In order to analyze multiple repetitions, individual motion segments must be identified and their corresponding time series must be aligned.

Therefore, the successful analysis of multiple repetitions of a movement depends on identifying those movements, very clearly, by a segmentation process; and the posterior alignment of those segments by an alignment process. The segmentation of the time series corresponding to each movement depends on the identification of the initial and destination points of each movement. The alignment of these time series, depends on the segmentation of the movements.

C. Segmentation Process

The data collected is composed of a repeating group of time series each composed of point coordinates (spatial data) along a predefined path with identified start and stop locations. Spatial data is collected for a group of individual sensors, over time (i.e., coordinates $X_{kt}, Y_{kt}, Z_{kt}$; where $k$ is the sensor identifier and $t$ is time).

For any single worker, the data is collected as a continuum of multiple repetitions (i.e., $X_{kr}, Y_{kr}, Z_{kr}$; where $r$ is the repetition identifier). In order to prepare the data for the analysis phase, it is necessary to separate and identify the individual motion repetitions and segments of the repeated movements. In this way, the data is segmented into unique iterations that can be analyzed by the machine learning algorithms.

The segmentation process depends on the identification of each movement segment composed of the start location (arrow tail) and the destination location (arrowhead). These points are clearly defined in the description for each movement (see Figure 4). The movement starts at time $t_{ini}$ and ends at $t_{final}$, for each segment. Thus, each segment is composed of a series of point coordinates over a measured period of time.

Thus, data collection then results in a number of point motion segments where each segment may differ from other segments by time, and/or point positions.

Identification of a segment requires first identifying the start and end points for each segment. As the motions are completed by humans there is not a singular static point that can be identified as either a starting point or an ending point.

In general, a starting point is defined as that location from which directed motion to achieve an objective begins. In the case of this research, it occurs at the point where the worker initiates the use of the tool. This is captured by a distinct change in direction within a specified origination area (defined by its proximity to the origination point). Further, the destination point is also captured by a distinct change in direction occurring within a specified destination area. A distinct change in direction is defined by the specific motion being measured.

Once the segment origination and destination points have been identified the segment can be identified. Once all segments have been identified it then becomes necessary to align each segment.

First a base segment must be selected and it’s motion duration determined. Once selected, each subsequent segment must be aligned to the base segment motion duration for purposes of comparison and analysis.

To align the subsequent segments, select a segment. Identify the origination ($t_{ini}$) and destination ($t_{final}$) times. Let the difference between the two times be $\Delta_t$.

When multiple repetitions of the same movement are generated, most likely it is the case that $\Delta_{t;r=i} \neq \Delta_{t;r=j}$; where $i \neq j$. That is, given two repetitions, most likely, the times elapsed between the beginning of the movement and the end of the movement are not the same for the two repetitions. Therefore, there is no one-to-one correspondence between the number of data points for the time series corresponding to repetitions $r=1$ and $r=2$.

We use dynamic time warping to adjust subsequent segments to the base segment allowing for comparative analysis of the motions.

Once all segments have been aligned, techniques for outlier identification are used to identify anomalies in the data. Once identified, the anomalies are examined and corrected.

IV. Data Analysis

The main pursuit of the data analysis phase is to reveal patterns of physical human motion, that could reveal behavior and habits, that take place during the performance of manufacturing tasks. Pattern discovery techniques that are applicable to time series are most relevant. Data visualization techniques are used to allow human experts to study these motions, in the search for such patterns.

A hypothesis associated with this framework is that the study of repetitive motion patterns could help discover behavior that would result in an improvement of worker safety and productivity.

Being able to analyze motions from individual workers also creates various possibilities. For instance, the ability to observe
the variability of how an individual worker behaves in the workplace, may allow for the improvement of the work environment (i.e., workstation design), or motion of the task (i.e., resting times). The possibility of capturing and analyzing a worker’s pattern of motions also provides for the possibility of comparing the motions of any group of two or more workers performing the same task. The results from such analysis could be used to identify and improve task specific motions resulting in safer and more productive motion patterns (i.e., training, practices).

A. Descriptive

The data will be analyzed using a variety of machine learning and statistical data analysis tools (e.g., time series, regression, induction trees, random forests) to identify characteristics and patterns of the motions.

In preliminary research, the time series for the position of one reflector on a subjects’ dominant arm was described using time series plots. Figure 5 presents nine iterations (in different colors) of the behavior of this one reflector on the right hand of an “operator” where the motion consisted of the subject using their right hand, carried a tool from the right side to left side of the work environment. The figure below only displays what corresponds to the horizontal displacement of the hand (the x coordinate).

![Figure 5: Nine iterations of hand horizontal movement](image)

This type of plot (Figure 5) shows the degree of variability in the motion, and at which point in space-time the variability occurs. For instance, it can be noticed, in Figure 5, that there is less variability at the beginning of the motion than at the end of the motion.

Using time series matching techniques (i.e., dynamic time warping), we can determine the degree of similarity between these nine movements. For instance, Table 1 shows that, on the one hand, iterations 4 and 5 (i.e., it4 and it5) are motions that were executed very similarly (i.e., DTW distance of 3491.214) relative to the others; on the other hand, iterations 3 and 9 (i.e., it3 and it9) appear to be very dissimilar (i.e., DTW distance of 12,685.604).

### Table 1: Dynamic time warping distance

<table>
<thead>
<tr>
<th>Iteration</th>
<th>DTW Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>it3</td>
<td>12,685.604</td>
</tr>
<tr>
<td>it4</td>
<td>3491.214</td>
</tr>
<tr>
<td>it5</td>
<td>3491.214</td>
</tr>
<tr>
<td>it6</td>
<td>3491.214</td>
</tr>
<tr>
<td>it7</td>
<td>3491.214</td>
</tr>
<tr>
<td>it8</td>
<td>3491.214</td>
</tr>
<tr>
<td>it9</td>
<td>12,685.604</td>
</tr>
</tbody>
</table>

B. Prescriptive

Predictive techniques, such as time series forecasting, OLS and other smoothing techniques, will be used to generalize the conduct of any given motions. These generalizations of movements will be used to create simulations.

Constraints of motion will be incorporated using prescriptive techniques. These techniques will be used to define motions that minimize or avoid the execution of harmful motions.

These results, together with the generalizations from predictive information will be used to create simulations of the actions of laborers in the work environment to be applied activities to improve work practices, task design, and training of the labor force.

V. Case Study

To illustrate the process described a case of a repetitive right-arm movement, from left to right, for a right-handed person is shown. The following analysis involves the movements of the markers attached to the right hand (one marker), the right wrist (two markers) and the right elbow (one marker). The objective is to show how standard statistical process control tools may be used to evaluate the variation in motions of a worker.

Besides tracking the motion of parts of the human body, along the right-to-left movement in this case; it is also the intention to track the movement of parts of the body in relation to other parts of the body. Figure 6 shows the dynamics between the right-elbow marker and the outer right-wrist marker.

At the beginning of the motion, the wrist is located to the left of the elbow. As the motion progresses, the wrist moves to the right of the elbow. This pattern reveals, among other things, the point in time in which, characteristically, the wrist and the elbow coincide along the x-axis of movement. These types of patterns could assist in identifying cases of overextension.
Figure 6: Four Iterations of Right-Wrist/Right-Elbow Markers Along the X-Axis

Figure 7: Four Iterations of Right-Wrist/Right-Elbow Markers Along the Z-Axis

Figure 8: Comparison of Iterations of Elbow Markers Along the Z-Axis

Figure 9: Comparison of Iterations of Elbow Markers Along the X-Axis

A comparison between consecutive iterations is represented in Figure 8. The inspection of consecutive iterations reveals how similar the motions are between them.

Another objective is to monitor the changes in motions, perhaps the degradation of it, as time passes (Figure 9).

Table 2: DTW Distances for Iterations of Hand Motion

Since each iteration of motion generates a multidimensional time series (i.e., [x,y,z] coordinates), Dynamic Time Warping (DTW) Distance is used to compare the different iterations.
(Shokoohi-Yekta, et al., 2017; Müller, 2007). Table 2 shows the DTW distances for nine iterations of hand motion. These distances seem to be distributed Gaussian (Figure 10); perhaps implying the existence of a “typical” distance from the first iteration (h.it2).

Figure 10: Distribution of the DTW Distances Between Iterations – Hand Motion

Given that the distribution of the distances from the first iteration (h.it2) seems to be Gaussian, the data was analyzed using a control chart. The experiment shows that the deviation between iteration 2 and iteration 3 is atypically large; as are the distances between iteration 2 and iterations 4 and 6, which are atypically close to iteration 2.

Figure 11: Control Chart for Right-Hand Motion

Table 3 and Figure 11 show DTW distances between the first iteration (e.it.2) and the other eight iterations of elbow motion.

This time, the distribution of DTW seems to depart from a Gaussian distribution.

Table 3: DTW Distances for Iterations of Elbow

<table>
<thead>
<tr>
<th>Iteration</th>
<th>e.it.2</th>
<th>e.it.3</th>
<th>e.it.4</th>
<th>e.it.5</th>
<th>e.it.6</th>
<th>e.it.7</th>
<th>e.it.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iteration</td>
<td>4394.528</td>
<td>3236.597</td>
<td>2193.053</td>
<td>1106.945</td>
<td>5248.491</td>
<td>3737.613</td>
<td>4745.895</td>
</tr>
<tr>
<td>Iteration</td>
<td>3854.342</td>
<td>3858.802</td>
<td>3206.383</td>
<td>5359.429</td>
<td>5929.963</td>
<td>6910.205</td>
<td>4654.082</td>
</tr>
</tbody>
</table>

The control chart in Figure 12 reveals that, most iterations are relatively close to the first one, with the exception of iterations 3 and 7.

Figure 12: Control Chart for Right Elbow

REFERENCES


