Human Motion Analysis Using Image Processing Technology

Wen-Yen Wu and Jia-Yu Wu
Department of Industrial Management
I-Shou University
Kaohsiung, Taiwan
Corresponding Author: wywu@isu.edu.tw

***Abstract*—In this paper, the use of machine vision to analyze human motion is mainly divided into the following steps: (1) image capture, (2) image pre-processing, (3) marker detection, (4) marker classification, (5) lumbar spine pressure calculation. Then use the self-developed judgment rule to judge the position of the marked point. Corresponding to the coordinate information of the correct position marking point, as well as inputting the subject's height, weight, and lifting weight, it can be brought into a simplified biomechanical model, and the pressure on the lumbar spine can be estimated to provide operators with potential injuries improve posture or lift loads.**

Keywords— Biomechanics, Anthropometry, Machine vision, Digital image process.

I. Introduction

The motivation of this study is the design and implementation of the artificial lifting system. The purpose is to explore the practicability of the methods mentioned in this paper in artificial lifting and to propose simple and effective treatment methods, hoping to achieve the best results with the least cost. The ultimate goal is to develop a system that can measure the strength of the lumbar spine in real time, and can make a preliminary judgment on the posture of manual lifting movements on site, so as to help the lifting personnel's physical safety.

When material handling is moving towards automation, many tasks need to be handled by hand, and it also happens in daily life, which can be called manual material handling (MMH). This is the main cause of excessive musculoskeletal injuries caused by manual weightlifting, so many scholars are studying this range.

There are four types of safety limits used to define human material handling: (1) epidemiological, (2) biomechanical, (3) physiological, and (4) psychophysical. In the general workplace, long-term manual handling is not common. This study will study the pressure on the intervertebral disc, so it will be studied using a biomechanical method.

The National Institute for Occupational Safety and Health (NIOSH) [6] proposed the MMH equation in 1981 based on biomechanics, physiology, and psychophysics. It was revised again in 1991. Although this equation cannot calculate various formulas, it can be adjusted according to this formula.

This paper builds a static human motion system based on manual material handling, including sensors, signal sources, and image processing programs that we handle ourselves. We use a digital camera as the sensor and a ping pong ball as the source. The application of machine vision in human motion recognition can be divided into the following steps: (1) image capture, (2) image preprocessing, (3) detection mark, (4) frame mark recognition and classification. The above process can be used for biomechanical calculations. This research aims to establish a human motion analysis system.

The method adopted are as follows. First, literature and information related to human motion are collected, and after evaluating each possible technique, efficient image preprocessing and position detection are established. We then use our image processing program to find the marker we want to detect and detect the location. With a simple and easy algorithmic method and a high-precision system, faster human motion analysis is expected.

II. Human Motion Analysis

Lifting is usually the beginning or final action in material handling, and lifting is usually the bottleneck in the handling process, because it is a small number of tasks that must be completed by manpower. If the personnel's posture or weight is not properly carried, it will often affect the health of the personnel. There are three kinds of moments that are most important when people perform lifting movements. The first one is the saggittal moment, which is also the main part discussed in this paper. When a person lifts an object placed on the ground, the person must overcome the moment (the product of gravity and the arm of force) in the direction of the central plane (the plane perpendicular to the face and body). When the mass of the object is greater, and the distance between the person and the object The farther the distance, the greater the force required; the second is the lateral moment, when the weight is moved from the left foot to the right foot, the downward component moment of the lateral plane (a plane parallel to the body), and vice versa The same is true; the third is the torsional moment, which is what the waist bears when it must turn during the lifting process.

The static human posture image captured by the digital camera is transmitted to the detection system for detection. The detection system can be divided into four parts: image preprocessing, detection, classification and calculation of intervertebral strength. We can calculate the weight of the elevator through this system. In the process of image preprocessing, color segmentation and noise reduction are performed on the image. The posture of the human body is clearly displayed. We then use a contour tracking method to locate the marked region we wish to detect and detect the location. After classification, we need to input information such as height, weight, and weight of the lifting source for biomechanical calculations. Finally, we use a series of judgment rules to calculate the lifting capacity. Figure 1 shows the flowchart of the proposed method.

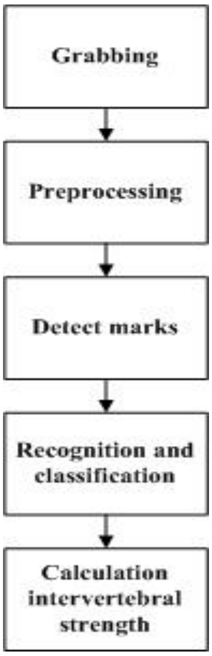

Fig. 1. Flowchart of the proposed method.

Due to the evolution of technology, the functions of digital equipment are becoming more and more powerful. Therefore, in capturing images, this article uses a digital camera as a device for capturing images, and fixes it on a tripod for shooting. When shooting, the distance from the image to be tested is about 3 to 4 meters, and the optical zoom feature provided by the digital camera is used to control the image to be tested to fall within a fixed range. The camera settings are: the size of each image is 640 x 480 pixels, the white background part is used for white balance setting, the sensitivity is set to ISO 50, the image format is Jpeg, the image quality is the best, and the rest uses the preset settings of the camera itself value. When the image is input to the computer, it will be converted to 24bit BMP format for storage to facilitate subsequent processing. The test subjects who were photographed stood in front of a simple background and faced the digital camera with the right sagittal plane, as shown in Figure 2. Since it is easier to observe the moving action from the sagittal plane of the human body, the action of lifting materials was photographed in this way. Input the captured image files into the computer for further processing.

Fig. 2. Sagittal images of static human motion

In this paper, the detection objects are markers attached to the main joints of the human body. The image preprocessing method used in the system is described below.

In the previous sections, we used table tennis with strong color features as markers. To obtain features, the RGB image will be transformed into HSV space [4]. The HSV space contains Hue, Saturation, and Value.

$$H' = \cos^{-1}\left\{\frac{0.5[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right\}$$

$$H = \begin{cases} H', & if\ B \le G \\ 360 - H', & otherwise \end{cases} \quad (1)$$

$$S = \frac{\mathrm{Max}(R,G,B) - \mathrm{Min}(R,G,B)}{\mathrm{Max}(R,G,B)} \quad (2)$$

$$V = \frac{\mathrm{Max}(R,G,B)}{255} \quad (3)$$

In the above equations, the ranges of HSV [3] are 0° to 360°, 0 to 1, and 0 to 1, respectively. Color quantization divides the HSV color model into eight regions, each representing red, yellow, green, cyan, blue, magenta, black, and white.

The logo of table tennis has only one color, that is orange. Orange is red in HSV space, so we extract red. The system will save the color through thresholding. On the basis of matching the scope of research needs. We just found out that the H value is equivalent to the ping-pong color gamut. When the H value is between 5 and 15 degrees. This results in a sharper image. Figure 3 shows the accumulation of H values and the range required for the study. We just found out that the H value is equivalent to the ping-pong color gamut. When the H value is between 5 and 15 degrees. This results in a sharper image. Figure 3 shows the accumulation and pursuit of H values.

Fig. 3. H value accumulating and pursuing

Although the marker points attached to the surface of the human body have strong color information, they are also susceptible to interference from other objects in the complex background, and the RGB signal image is less resistant to the impact of brightness changes. The retained table tennis image will be applied. The RGB signal distribution on the Line Profile is shown in Figure 4.

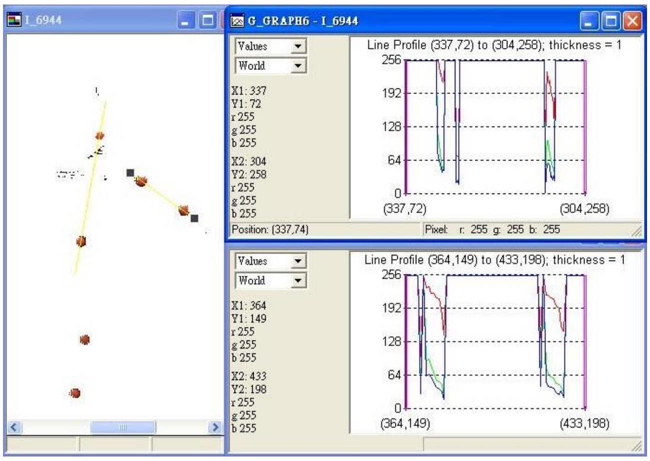

Fig. 4. RGB value distribution of table tennis at different positions.

We adopt the morphological method to reduce the noise, because it is a better method to reduce random noise interference in this paper.

When shooting outdoors, the image usually has noise, which should be light, shadow, and dust. To filter out noise, morphological operations are employed after image thresholding. The erosion operation can effectively reduce the outline of the object, and at the same time can filter out some small noises. We then use the dilation operation to recover the contours reduced by the erosion operation. By combining erosion and dilation operations, noise can be filtered out [4].

In this section, we use the contour tracking method [5] to search for all markers in an image. Because the pixels in a digital image are discrete data. Let P be the (x, y) pixel to be tested, and there will be 8 pixels around it. These 8 pixels are P0(x+1,y), P1(x+1,y-1), P2(x,y-1), P3(x-1,y+1), P4(x - 1, y-1), P5 (x-1, y), P6 (x, y+1) and P7 (x+1, y+1). There is an 8-adjacency relationship between P and 8 pixels. Figure 5 shows the 8-adjacency relationship.

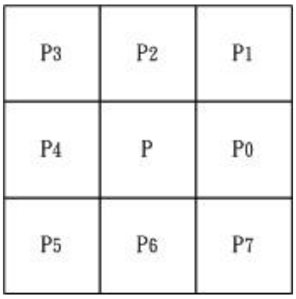

Fig. 5. 8-adjacency

We can use the relationship mentioned above to define a rule as follows.

Rule 1: Let f1 and f2 be two gray value elements in the image and have 8 adjacencies. It must exist between two elements if f1 – f2 = ± 255.

Using the above regular features, we can define a continuous operation to detect object contours. The contour tracking method is divided into three steps. First shift from left to right, search from top to bottom in the image, and find out the starting point S of the marked contour according to rule one. Secondly, look up the other eight elements clockwise with S, and find the first edge in Pn that conforms to rule 1. And change the gray value f of Pn to K, K∈[0, 255]. Finally, if Pn is not S, return to above step, otherwise stop.

According to the first step, we can identify the marker we want to detect and detect the position. Then we'll go to each marker to correct the joints. There are six main joints in the sagittal plane of the human body: shoulder, elbow, wrist, hip, knee, and ankle. Arrange numbers 1 to 6 in sequence. Figure 6 shows the main authentication procedure.

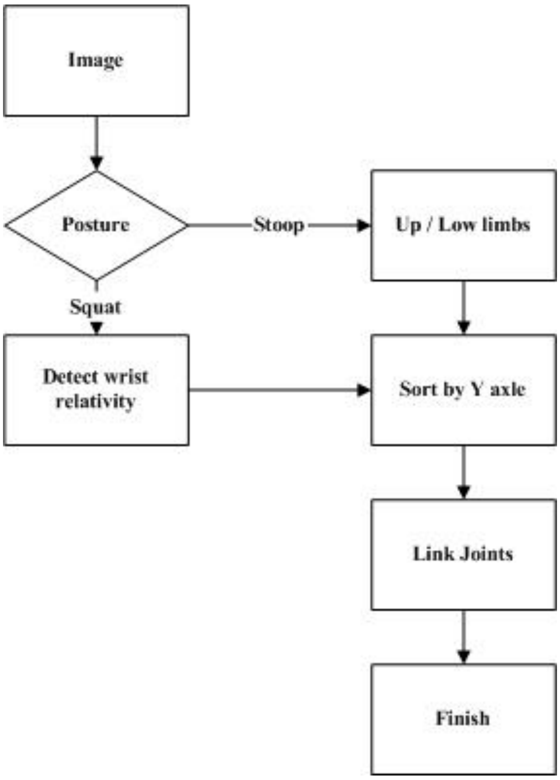

Fig. 6. The main discrimination procedure

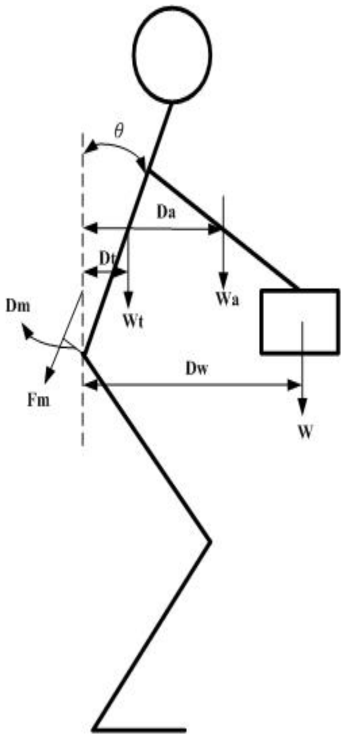

Fig.7. Biomechanical lifting model

This study utilized the biomechanical model described by Chaffin [1]. This program can input data, including height, weight and weight of the lifting source, and calculate the strength of the lumbar spine (L5/S1) at that time. Figure 7 shows the biomechanical lift model.

Entered values (height, weight, weight from source of lift) are not available for biomechanical lift mode. So we use the percent weight distribution stated by Webb [7] and each position of the body center of mass brought by Dempster [2] to help bring into equation (4). The strength Fc defense of the lumbar spine is:

$$F_c = F_m + (W + W_a + W_t)\cos\theta \quad (4)$$

$$F_m = (W_t * D_t + W_a * D_a + W * D_w) / D_m \quad (5)$$

Where F_m represents the strength of the erector spinae, which can be calculated by formula (5). Among them, W, W_a, and Wt represent the strength of lifting source, upper body and upper body except upper body, respectively. D_t, D_a, and Dw represent the relative moment distances from L5/S1 to W_t, Wa, and W, respectively.

III. Experimental Results

The lifting posture analysis is mainly to assist the lifting personnel to judge the posture used and the weight of the lifting object, whether the pressure on the lumbar spine is caused by injury, and to provide the operator with the possibility of improving the posture or lifting load that may cause injury. The main contribution of this paper is to propose a method for measuring the force on the lumbar spine of lifting personnel. First, the CCD optical device is used to capture the marker points attached to the surface of the human body, and then the image processing program developed by itself is used to convert the input image into an RGB image. The signal is converted into an HSV image, which can reduce the influence of light changes and highlight the color of the ping-pong ball markers. Then, a morphological operation of erosion and dilation is used to eliminate the noise of non-marked points, and then contour tracking is used to detect the position of marked points. After a series of mathematical calculations, the relative position of the ping-pong ball attached to the joint points is calculated, and the judgment rule is defined based on the self-observation of the relationship between the joint positions to judge the position of the marked point. Corresponding to the coordinate information of the correct position marking point, as well as inputting the subject's height, weight, and lifting weight, it can be brought into a simplified biomechanical model to find out the pressure on the lumbar spine.

In this experiment, 58 images were examined, which were captured against simple backgrounds. The size of each image is 640 × 480 pixels. We will illustrate how to use flow with the following example.

The input image is converted through the threshold between the HSV three signals set in this article, and the image will leave eight main colors. The converted image is shown in Figure 8.

When the image maintains the color range of the ping-pong ball and reduces most of the noise, as shown in Figure 9. We will then identify and classify the markers we wish to detect. Figure 10 shows the properly linked markup. Finally, we can calculate the strength of the lumbar spine.

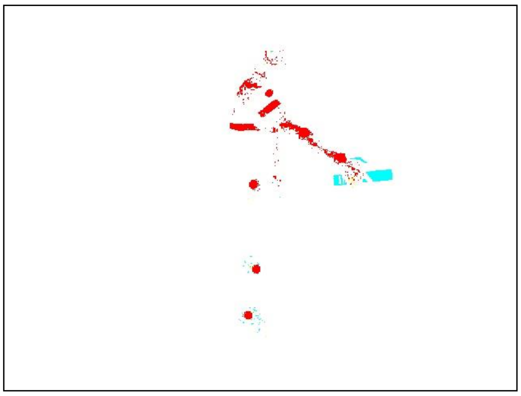

Fig. 8. Image after color separation

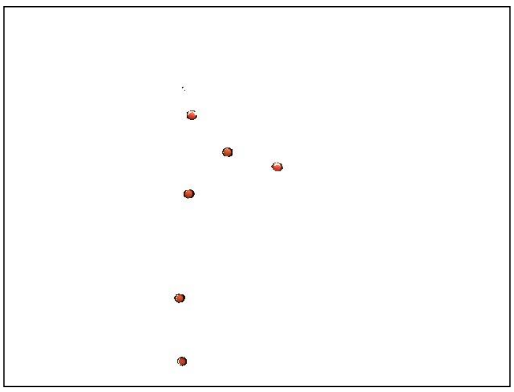

Fig. 9. Image after noise reduction

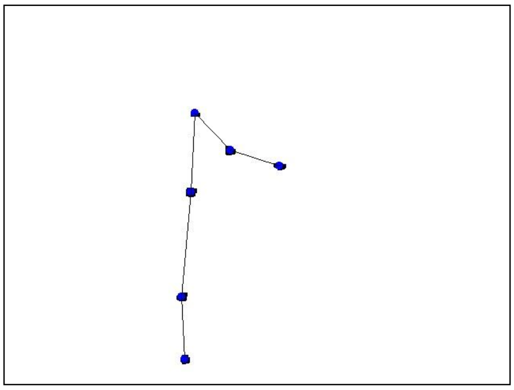

Fig. 10 Image after marks linking

In 54 of the total 58 images, the strength of the lumbar spine could be successfully calculated in the system. So the system has a 93% success rate.

First of all, in terms of marker detection rate, among the seven failures, six of them were due to situations that the system could not handle during processing, and the remaining one was due to the wrong number of detected markers. The most likely reason that the system cannot handle is that there are too many unpredictable noises in the image, and the noise area is not large, which cannot be effectively eliminated in the pre-processing, so that the system makes mistakes. The situation where the number of detections is incorrect is that the skin color of the subject is close to the color of the marking point, and the area is not small, so that it is misjudged as a marking point during processing.

IV. Conclusions

In this paper, we used several simple processes to detect marks and calculate the strength of the lumbar vertebrae. First, we use the HSV space to extract the color features. Second, the morphology method is used to reduce noise effectively. Third, the contour tracking method is used to search all of the marks in the images. Finally, we can calculate the strength of the lumbar vertebrae by the purpose of biomechanical model. To increase the accuracy of recognize marks, the method to classify in the future needs improvement again. On the other hand, we hope to reduce process time if software operation can be optimized. After solving these questions, we hope to set up a useful human motion analysis system.

References

[1] Chaffin, D.B. and Andersson, G.B.J. (1991). Occupational biomechanics (2nd ed), John Wiley & Sons, pp.215-239.

[2] Dempster, W.T. (1955). Space requirement of the seated operator, Teach. Report, Wright-Patterson AFB, WADC.

[3] Foley, J.D., Dam, A.V., Feiner S.K. and Hughes, J.F. (1993). Computer graphics: principles and practice, USA.

[4] Gonzalez, R.C. and Woods, R.E. (2001). Digital image processing (2nd ed), New Jersey: Prentice Hall.

[5] Jau, M.K. and Wu, W.Y. (2004). A genetic algorithm for significant point detection, Proceedings of the 2004 Twelfth National Conference on Fuzzy Theory and Its Applications (CFSAT 2004), I-Lan, Taiwan, November.

[6] National Institute for Occupational Safety and Health (1981). A work practices guide for manual lifting, Teach. Report No. 81-122, U.S. Dept. of Health and Human Services (NIOSH). Cincinnati, OH.

[7] Webb Associates (1978). Anthropometric source book, Vol. I, NASA 1024, National Aeronautics and Space Administration, Washington D.C., IV-1-IV76.