Vision based controller design with the application to a 3D overhead crane system

Lun-Hui Lee¹, Chung-Hao Huang¹, Sung-Chih Ku² and Cheng-Yuan Chang²*
¹Institute of Nuclear Energy Research in Taiwan, Taiwan
²Chung Yuan Christian University, Taiwan
*Corresponding Author: ccy@cycu.edu.tw

Abstract— This proposed work designs controller using vision feedback for an overhead crane. This approach searches the image features and computes the useful vision based information in terms of several tracking areas block in each frame. According to the lightest or darkest point in the tracking area of a dynamic object; this work determines the trolley position and load swing for controller design. Experimental results verify effectiveness of the proposed work.

Keywords: overhead crane, vision feedback, controller design.

I. Introduction

A crane, consisting of a trolley and a flexible wire to hang the load, has the merit to move heavy cargoes rapidly. Fast and smooth handling of a crane system requires a smart and precise control of crane motion. There were many researchers developed approaches to control the crane systems based on varieties of optimal based methods using the Lagrange crane model [1]-[4]. These researches applied complex dynamic equations to design the controllers which were not easy to be implemented in industry. Besides, the input shaping method, focused on restraining the load swing but did not solve the positioning control problem, was also addressed [5]. Other researches also applied intelligent based control schemes for crane because of their model-free properties [6]-[8]. However, those schemes always took time to stable condition.

In addition to these, some researchers implemented image sensing to measure the swing angle of an overhead crane. Two charge coupled devices (CCDs) were utilized to measure the angle of the swing in [9]; besides, there is researcher using a camera installed on the trolley to measure the swing angle of a crane. A built-in video tracker was also addressed in [10]. In general, these visual tracking methods are complex and the high speed cameras or CCD are necessary. In [11], an adaptive fuzzy sliding mode control (AFSMC) for a two dimensional crane based on visual feedback was developed. But the depth information in an image did not be considered in their work.

This paper uses visual based approach to control the crane transporting the load rapidly and smoothly. A tag is also used to show the depth information of an image. Besides, for saving the computing efforts, we focused only on position and tracking areas and ignored the rest image areas. So the total processing areas are much less than original images, therefore, image processor can retrieve feedback information in real-time. Also, there is not necessary to use model matching algorithm to verify features in an image, but checking maximum or minimum grey value in tracking area instead to saving the computing effort. Therefore, the proposed contactless sensing technique is free from sensor allocation, electrical noise and other problems. Experiments on the regulation of the overhead crane were completed to verify the effectiveness.

II. AFSMC with Vision Feedback Method

This work uses vision from camera replacing encoders as a contactless feedback sensor of AFSMC for the crane systems. The control block diagram is shown in Fig. 1. Two cameras are set to measure along XZ and YZ planes. One computer acts as an image processor to calculate the position of trolley \( p = (x, y) \), rope length \( l \) and payload swing angle \( \theta = (\theta_x, \theta_y) \) from the images and the other computer derives the proper control powers \( u \). The image captured by the XZ camera is shown in Fig. 2 with two green tracking blocks and one red positioning block. Each block has a tracking point helping to decide the crane information, including the trolley and payload position and swing angle. Since the background is black in the image, we can only search the lightest pixel from left to right at the top of the 2D image to find the tracking point on the wire. Also, the thickness of the wire rope is the same from the top to the end; therefore, we can use the same method to find the lower tracking point (green) too.

The depth information of a 3D object is not easy to obtain by 2D picture because of the effects of the scaled size and height, aerial perspective, linear perspective and texture gradient. Therefore, in Fig. 2, we used a fixed marker on the end truck in the overhead crane system to decide the depth information and use a LED in red helping to obtain the depth information. Therefore, we can also search the largest pixel value from left to right to find the depth information of trolley on X-axis instead of using image model matching method. Similar ways can be used to capture the image information from YZ camera. So, only the tracking and positioning blocks areas have to be processed instead of the whole image. Searching lightest pixel instead of general model matching method also shortens the computation time.

The proposed manuscript set the tracking points at the centers of respective tracking blocks. Also, three tracking blocks will move according to the motions of corresponding tracking points to assure the points are at their centers. The sizes of tracking blocks are based on the moving speed of the tracking points. Besides, the upper tracking point and positioning point in Fig. 2 only move along horizontal axis; so the heights of the respective tracking blocks area can be set...
one pixel. However, the lower tracking block in Fig. 2 needs higher height for the load swing.

The initial position is set at the original point of the xyz-axes. When an image plane \((ip)\) is taken, the positioning blocks areas \((pa)\) and tracking blocks areas \((ta)\) are used to search the corresponding markers or tracking points. After the markers and tracking points are found, these points are set as the next centers of the positioning and tracking blocks areas, respectively.

Since the positioning marker point of XZ plane is a white point with one pixel height and of YZ plane is set a black point; therefore, the XZ plane positioning marker point \((p_{m_{xz}})\) and the YZ plane positioning marker point \((p_{m_{yz}})\) can be expressed as

\[
p_{m_{xz},t} = \max \left( ip_{xz,t} \right) \quad \text{and} \quad p_{m_{yz},t} = \min \left( ip_{yz,t} \right),
\]

\((t)\) denotes the time. Equation (1) shows the XZ plane positioning marker point is set to the maximum grey value of the image plane at time \(t\) with the positioning blocks areas at time \((t-1)\). Also, the YZ plane positioning marker point is set to the minimum grey value of the image plane with the positioning blocks areas at time \((t-1)\). The positioning marker at the left and right limits of the moving tracks along XZ planes are set \(pl_{lxz}\) and \(pl_{rxz}\). Besides, there is no fixed depth target on our 3D crane in YZ image; hence, different position of positioning marker \(p_{m_{xz},t}\) will correspond to different pair of left and right limits of crane motion in image along YZ plane. Thus, setting the left and right limits of crane motion along YZ plane to be \(pl_{lxz}\) and \(pl_{rxz}\), respectively.

Therefore, the actual position \(pma_{sz}\) and \(pma_{sz}\) could be obtained by

\[
pma_{sz} = \frac{pm_{sz} - pl_{sz}}{pl_{sz} - pl_{sz}}, \quad pm_{sz} = \frac{pm_{sz} - pl_{sz}}{pl_{sz} - pl_{sz}}.
\]

Both the upper tracking points \(utp_{xz,t}\) and \(utp_{yz,t}\) only move in the horizontal tracks, so we have

\[
up_{xz,t} = \max \left( ip_{xz,t} \right), \quad utp_{yz,t} = \max \left( ip_{yz,t} \right),
\]

\((t)\) where \(uta\) represents the upper tracking area. Besides, the lower tracking points \(lp_{xz,t}\) on XZ and YZ planes also move in two dimensional space because of load swing, so

\[
lp_{xz,t} = \max \left( ip_{xz,t} \right), \quad lp_{yz,t} = \max \left( ip_{yz,t} \right),
\]

where \(lta\) denotes the lower tracking area. These procedures also determine the other tracking or positioning blocks areas, so the tracking blocks follow the tracking points efficiently. The purpose to track these points is to find two points on the wire rope, and further calculates the load swing angle. A vertical line is drawn from upper tracking point \((x_{up}, y_{up})\) \((*\) denotes \(x\) or \(y)\) as the adjacent side of a right-angled triangle, a horizontal line is drawn from lower tracking point \((x_{lp}, y_{lp})\) as the opposite side, and wire rope could be the hypotenuse, therefore the swing angle \(\theta\) could be calculated by simple computation. Also, the length of wire rope can be obtained by

\[
l = \sqrt{(x_{yz,lp} - x_{yz,up})^2 + (y_{yz,lp} - y_{yz,up})^2} \cdot \sec (\theta).\]

Therefore, the proposed vision based tracking method could take place of encoder to offer all the information of 3D crane system as a feedback.

In Fig. 3, we used both vision and encoder based methods to test the tracking performance. We slightly forced the payload to swing along XZ- and YZ-planes by hand and the results are shown in Figs. 3(a) and 3(b) respectively. In Fig. 3(a), the green line is the tracking result by image and the blue line is by encoder. One can find the blue line decay to zero quickly; however, the image based result can track the swing correctly. Similar results are also shown in Fig. 3(b). These differences may be aroused by the interior friction inside the encoders. So, when the swing is small, the encoders cannot measure the load swing efficiently but vision based method still track the swing precisely.

Figure 4 displays the proposed control system. Let the desired trolley position and swing angle are \(g_p = [g_x, g_y]^{T}\) and \(g_d = [g_{dx}, g_{dy}]^{T} = [0, 0]^{T}\); therefore, the position error and the swing angle error is \(e_p = [e_{px}, e_{py}]^{T}\) and \(e_{\theta} = [e_{\theta x}, e_{\theta y}]^{T}\), respectively. The control force \(u_p = [u_{px}, u_{py}]^{T}\) is used to drive the trolley along X- and Y-directions. By AFSMC, two sliding functions for position and swing angle are

\[
s_1 = \dot{e}_p + c_1 e_p; \quad s_2 = \dot{e}_\theta + c_2 e_{\theta},\]

\((c_1, c_2)\) are real positive definite matrices defining the slopes of the sliding surface. The notation \(D\) in Fig. 4 represents the differentiation. Besides, we set \(c_1 = c_1 + w_c c_1^a\), where \(c_1^a\) is the base matrix of \(c_1\), \(w_2\) is the weighting matrix and the correction term \(c_2^a\) is a fuzzy adaptive matrix determined by

\[
R_i: \frac{c_2}{c_2^*} = \mu_i \left( F^* \right), \quad c_2 = \sum c_i^a \cdot \mu_i \left( F^* \right),\]

where \(\mu_i\) denotes X or Y, \(F^*\) is the \(i\)th input fuzzy set with the membership function \(\mu_i \left( F^* \right) = e^{-\frac{1}{2} \left( \frac{F - F_i}{F_i^a} \right)^2}\), and \(c_2^a\) is the output fuzzy singleton and \(i = 1, \ldots, I\). After defuzzification by gravity method, the output signal \(c_2^a\) is

\[
c_2^a = \frac{c_2^a}{\sum c_i^a \cdot \mu_i \left( F^* \right)}\]

So, by normalizing \(c_2^a\) to \([0, 1]\), the range of \(c_2\) is

\([c_2^* - c_2^a + w_2]\). Also, \(c_2\) will increase to make AFSMC go toward sliding surface with higher \(e_{\theta}\). Thus, a composite sliding function is \(s = s_1 + c_2\), where \(c = c^b + wc^b\), \(c^b\) is
base matrix, \( w \) is the weighting matrix, and \( c_j' \) is fuzzy adaptive matrix determined by

\[
R_j: \text{If } |x_p| = F_j, \text{ Then } c_j' = c_j'. \quad (9)
\]

In Eq.(9), \( R_j \) is \( j^{th} \) fuzzy rule ( \( j = 1, \cdots, J \) ), \( F_j \) is \( j^{th} \) input fuzzy set with the membership function \( \mu_{F_j}(|e_p|) = e^{-\frac{|e_p|}{\sigma_j}} \), and \( c_j' \) is the output fuzzy singleton. Thus, the output signal \( c_j' \) will be

\[
c_j' = \frac{\sum_{j=1}^{J} c_{j} \mu_{F_j}(|e_p|)}{\sum_{j=1}^{J} \mu_{F_j}(|e_p|)}. \quad (10)
\]

The slope \( c \) is \( [c^a, c^b + w] \) with \( c^b \) is normalized to \([0, 1]\). The parameter \( c \) increases to make AFSMC go toward the sliding surface or decreases to concentrate on positioning.

At last, the fuzzy based sliding function \( s \) determines the power of motors:

\[
R_j: \text{If } s_j = F_j, \text{ Then } u_j = u_j'. \quad (11)
\]

After defuzzification with gravity method, the signals \( u_j' \) to control the motors are

\[
u_j' = \frac{\sum_{j=1}^{J} u_{j} \mu_{F_j}(s_j)}{\sum_{j=1}^{J} \mu_{F_j}(s_j)}. \quad (12)
\]

### III. Experimental Results

For visual tracking, two CCD cameras with the maximum video rate 30 fps and resolution 720x480 pixels were installed. In this work, the sizes of one positioning and two tracking areas are 21x1, 41x1 and 61x31 pixels for the image of XZ plane from top to bottom in Fig. 2, respectively. A set of 2500 areas are 21x1, 41x1 and 61x31 pixels for the image of XZ-plane. The payload is 0.5kg and the rope length of the modeled crane system in the experiment is 70cm.

In Figs. 5-6, the results obtained using CCD camera and encoders with AFSMC were compared and analyzed. The control target is set at the coordinate (59.5,33.6), where is 70% of the maximum traveling distance from the initial point. Figure 5 shows the AFSMC results using encoders’ feedback; meanwhile, two CCD cameras were also used to observe the performance by encoders. Figures 5(a)-5(b) show the trolley traveling along X- and Y-axes and Figs. 5(c) and 5(d) draws the load swing along XZ- and YZ-planes. The BLUE and GREEN lines are by encoders and by CCD cameras, respectively. The results in Figs. 5(a)-5(c) represent similar results and also show encoders can track the trolley position well. However, in Fig. 5(d), when the measurement of the load swing by encoders is close to zero; actually, there still has \( \pm 1 \) load swing. This figure shows encoders cannot measure the correct load swing angles so the control performance is not good enough.

Figure 6 draws the AFSMC results using visual based control method by CCD cameras; meanwhile, using encoder sensors to observe the control performance by CCD cameras. Figures 6(a)-6(b) show the trolley traveling along X- and Y-axes and Figs. 6(c) and 6(d) draws the load swing along XZ- and YZ-planes. The BLUE and GREEN lines present the measurement with encoder and observation by CCD camera, respectively. Both the performances of trolley position tracking by using encoder or visual feedback are similar, shown in Figs. 5(a), 5(b) and Figs. 6(a) and 6(b). The figures clearly demonstrate that the AFSMC with visual feedback can suppress the load swing well. However, when compared Figs. 6(c) and 6(d) with Figs. 5(c) and 5(d), it is clear that the visual feedback performs better than those achieved by encoder feedback. The enhancement is confirmed.

### IV. Conclusions

The presented work developed a vision based scheme to measure the trolley position and load swing information. The AFSMC is applied to derive the control. The proposed method does not require the system model to develop the controller and adjust the sliding mode function slopes. Instead of the complex model matching method, only examining the grey values in some tracking areas is necessary in the proposed work to find the tracking point powerfully, saving the computing effort. The proposed visual tracking method can suppress the load swing well and is sensitive to small angles that cannot be detected using an encoder, shown by experimental results.

### References


---

**Fig. 1:** The 3D overhead crane system with image feedback.

**Fig. 2:** Concepts of positioning and tracking blocks, example of x-z plane.

**Fig. 3:** Tracking method illustration.

**Fig. 4:** Flow chart of the AFSMC controller.
Fig. 5: The AFSMC results by using encoder feedback. (a) X-axis position, (b) Y-axis position, (c) $\theta_x$, (d) $\theta_y$.

Fig. 6: The AFSMC results by using visual feedback. (a) X-axis position, (b) Y-axis position, (c) $\theta_x$, (d) $\theta_y$. 