# 3D Human Hand Posture Reconstruction Using a Single 2D Image

**Mahdi Vaezi**                                                                    m.vaezi7777@gmail.com
*Faculty of Electrical & Computer Engineering*
*Science and Research Branch of Azad University*
*Tehran, Iran*


**Dr. Mohammad Ali Nekouie**                                       Manekoui@eetd.kntu.ac.ir
*Faculty of Electrical & Computer Engineering*
*K. N. T. University of Technology*
*Tehran, Iran*

---

## Abstract

Passive sensing of the 3D geometric posture of the human hand has been studied extensively over the past decade. However, these research efforts have been hampered by the computational complexity caused by inverse kinematics and 3D reconstruction. In this paper, our objective focuses on 3D hand posture estimation based on a single 2D image with aim of robotic applications. We introduce the human hand model with 27 degrees of freedom (DOFs) and analyze some of its constraints to reduce the DOFs without any significant degradation of performance. A novel algorithm to estimate the 3D hand posture from eight 2D projected feature points is proposed. Experimental results using real images confirm that our algorithm gives good estimates of the 3D hand pose.

**Keywords:** 3D Hand Posture Estimation; Model-based Approach; Gesture Recognition; Human- Computer Interface; Machine Vision, Color Image.

---

## 1. INTRODUCTION

Hand posture analysis is an interesting research field and has received much attention in recent years. It is the pivotal context of particular applications such as gesture recognition [1-4], human computer interaction (HCI) [5], sign language recognition (SLR) [6-8], virtual reality (VR), computer graphic animation (CGA) and medical studies. General solutions for posture analysis are divided into two categories: One attempt is to use mechanical devices, such as glove-based devices, to directly measure hand joint angles and spatial positions[21]. The other attempt uses computer vision-based techniques. Although the former can give real-time processing and reliable information, it requires the user to wear a cumbersome device and generally carry a load of cables that connect the device to a computer. All these requirements make the sensing of natural hand motion difficult. On the other hand, the latter is suitable for hand posture estimation since vision is a non-invasive way of sensing. Vision-based approaches can be classified into two types:

Appearance and three-dimensional(3D) model-based approach. The appearance-based methods are mainly based on the visual image model and use the image templates to describe the postures. The gestures are modeled by relating the appearance of any gesture to the appearance of the set of predefined, template gestures. Starner et al. [6] Use silhouette moments as the features to analyze the American Sign Language (ASL). In their research project, ªReal-time American Sign Language Recognition Using Desktop and Wearable Computer Based Videoº, they present two real-time hidden Markov model-based systems for recognizing sentence-level continuous ASL using a single camera to track the user's unadorned hands. The major advantage of this approach is the simplicity of their parameter computation. However, the loss of precise spatial information makes them less suitable for manipulative hand posture analysis. Since appearance-based methods are sensitive to viewpoint changes and cannot provide precise spatial information, it is less suited for manipulative and interactive applications, Smagt in [21] presents a new method to provide computer vision processes based data by

Mahdi Vaezi, Mohammad Ali Nekouie

utilizing and MRI device, although the results are much more precise but this method requires an expensive and voluminous MRI device presented in field. Conventional model-based methods are mainly used in two areas: 3D hand tracking and 3D hand posture estimation. Hand tracking is to locally track and estimate the positions of joints and tips of the hand in the image sequence. By analysis of static and dynamic motions of the human hand, Lee and Knuii [9] and Pitarch in [22] and Cobos in [24] present some constraints on the joints and use them to simulate the human hand in real images. In the experiments, some of them used markers to identify the fingertips. On the basis of Lee's contribution, Lien et al. [10] proposed a fast hand model fitting method for the tracking of hand motion. Although they improve the performance of the tracking algorithm, the computation of inverse kinematics is still required as it is contributed in [22] and [21], Rehg [11] described Digital Eyes for a real-time hand tracking system, in which the articulated motion of fingers was recognized as a 3D mouse by using a hand model having 27 degrees of freedoms (DOFs). This approach was based on the assumption that the positions of fingertips in the human hand, relative to the palm, are almost always sufficient to differentiate a finite number of different gestures. The hand gesture was estimated by a non-linear least squares method that minimizes the residual distances in finger links and tips of the model hand and those of the observed hand.

Shimada et al. [4] present a method to track the pose (joint angles) of a moving hand and refine the 3D shape (widths and lengths) of the given hand model from a monocular image sequence. First, the algorithm uses the silhouette features and motion prediction to obtain the approximated 3D shape. Then, with inequality constraints, they refine the estimation by the extended Kalman filter (EKF). Without the motion information, some research efforts have concentrated on 3D hand posture estimation. In the study of Chang [12], a prototype system for estimating the position and orientation of a human hand as well as the joint angles of the thumb and the fingers from a single image is developed. The hand pose is estimated by using sparse range data generated by laser beams and by using the generalized Hough transform, almost like the MRI method in [23]. Possible configurations for the fingers and the thumb are generated by the inverse kinematic technique. Although the above algorithms have promising results, posture estimation is not yet advanced enough to provide a flexible and reliable performance for potential applications. The estimation of kinematic parameters from the detected features is a complex and cumbersome task. They face the following problems: First, the articulated mechanism of the human hand, which involves high DOF, is more difficult to analyze than a single rigid object: its state space is larger and its appearance is more complicated. Second, model-based methods always involve finding the inverse kinematics, which are in general ill posed. It is obviously a task of computational complexity to estimate these kinematic parameters from the detected features. Third, previous methods on 3D require the use of multiple cameras, which not only is resource consuming, but also needs some form of 3D reconstruction that itself is computationally intense. Finally, it should be pointed out that the knowledge of exact hand posture parameters seems unnecessary for the recognition of communicative gestures.

In this paper, the goal of our work is to avoid the complex computation of inverse kinematics and 3D reconstruction; that is, without using 3D information, we propose a new approach to estimate the 3D hand posture from a single two-dimensional (2D) image. Preliminary results can be found in Refs. [13,14] , which deals only with finger posture. This paper extends the idea further to compute the 3D posture for the entire hand. First, we analyze the human hand model with 27 DOFs and its constraints. The constraints play an important role in our study, which help us to reduce 27 to 12 DOFs without significant degradation of performance. Using the hand model and its constraints, we develop an algorithm to estimate the 3D hand posture by using seven feature points to retrieve the 3D hand posture. The seven feature points are the point of wrist, the tips of the fingers and thumb, and the metacarpophalangeal joints for the thumb. We use color markers to identify these seven points and retrieve the approximate posture of the hand.

In the experiments, two feature extraction methods are utilized: one for model building by generating the real length of hand's bones and the other for on-line hand posture estimation. In extracting the parameters for the hand model, a higher degree of accuracy in detecting the feature points is necessary. In this regard, the feature points are extracted from the out-stretched hand with Color markers, placed on the necessary positions of the hand.

This paper is organized as follows: Section 2 discusses the hand model and its constraints. Section 3 presents the methodology to estimate the hand posture.

## 2. HAND MODEL AND ITS CONSTRAINTS
Lee and Knuii [9] and others in[22] and [24] defined a hand model with 27 DOFs The joints of the human hand are classified into three kinds: flexion, directive or spherical joints, which consist of one DOFs

(extension/flexion), two DOFs (one for extension/ flexion and one for adduction/abduction) and three DOFs (rotation), respectively, (see Fig. 1). For each finger, there are four DOFs described by u1 -u4 :
The thumb has five DOFs:
 Including the six DOFs for the translation and rotation of the wrist, the model has 27 DOFs.
Conventional models of the human hand are lacking in constraints. It limits their usefulness in computer vision and animation. The lack of constraints leads to unnatural model behavior. On the other hand, because the movements of the fingers are inter-dependent in the human hand and in order to reduce search space of matching, constraints are essential to further realize the hand motion.

## 2.1. CONSTRAINT 1:
This constraint is proposed by Rijpkema [15]. The angles of D (Distal) joints and P (Proximal) joints are dependent in middle finger.

$$\theta_4 = \frac{2}{3}\theta_3$$

(1)

## 2.2. CONSTRAINT 2:
This constraints are proposed in [16],[15],[24] and also other references. The joint angles of P and M joints of four fingers have a dependency represented by the following equation:

$$\theta_1 = k\theta_3 \qquad 0 \le k \le 1/2 \qquad (2)$$

## 2.3. CONSTRAINT 3:
Five points (Wrist joint, Metacarpalangeal joint, Proximal joint, Distal joint and Tip represented by W, M, P, D and T, respectively) of each finger are coplanar. We define this plane as the `finger plane'. According to Constraint 4, we omit the adduction/abduction DOF for the M joint, so that M, P, D joints of four fingers are all extension/flexion joints.

## 2.4. CONSTRAINT 4:
The joint angles of I and M of the thumb have a dependency represented by the following equation:

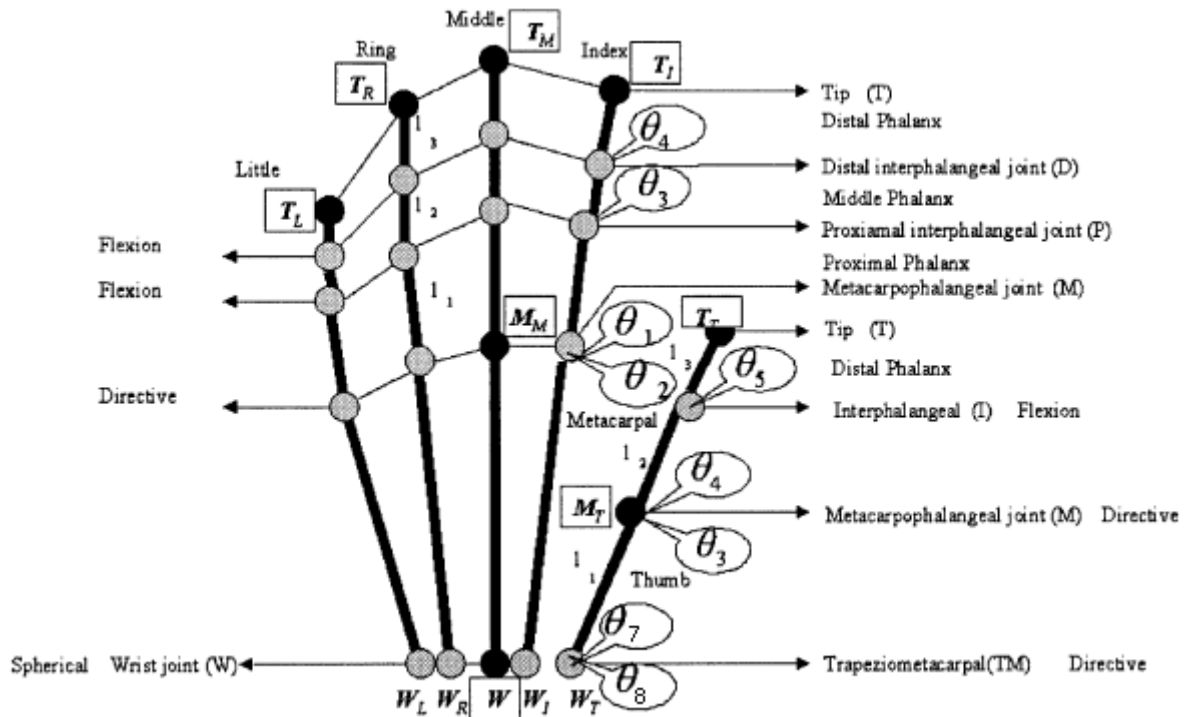$$\theta_5 = a\theta_4 \qquad a \ge 0 \qquad (3)$$

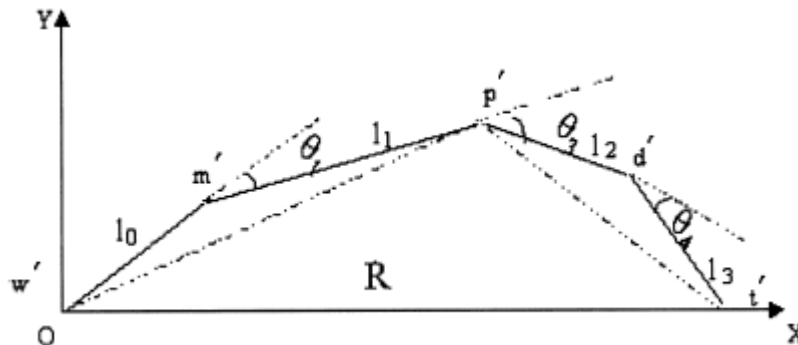**FIGURE 1:** Twenty- seven DOFs human hand model.



**FIGURE 2:** Pentagon geometry of a finger

## 2.5. CONSTRAINT 5
Four points (Trapeziometacarpal joint, Metacarpalangeal joint, Interpalangeal joint and thumb tip represented by TM, M, I and T, respectively) of the thumb are co planar. We define this plane as the `thumb plane'.

## 2.6. CONSTRAINT 6
The palm is assumed not to become hollow in a non prehensile configuration. We assume that the M joint of each finger are located in the plane called `palm plane' and this plane is perpendicular to the `finger plane' of the middle finger. In general, three points are necessary to align the palm with the images. There are also some other constraints that must be considered.

## 2.7. CONSTRAINT 7
Thumb plane makes an angle $\theta 8$ with palm plane, representing $\theta 2$ in fig. 1.

## 2.8. CONSTRAINT 8

$\theta3$ in figure 1 has a negligible amount of movement, more relation between angles in thumb finger could be obtained, [15]:

$$\theta7 = \frac{\pi}{2} - 1.2 * \theta8 \qquad (4)$$

$$\theta4 = \pi - \left(\frac{7}{5}\right) * \theta7 \qquad (5)$$

$$\theta5 = 2 * \theta4 \qquad (6)$$

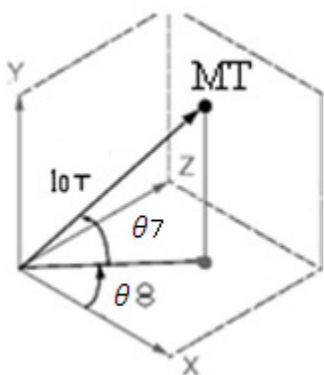Fig. 3a, b show details to better understand thumb's geometry.
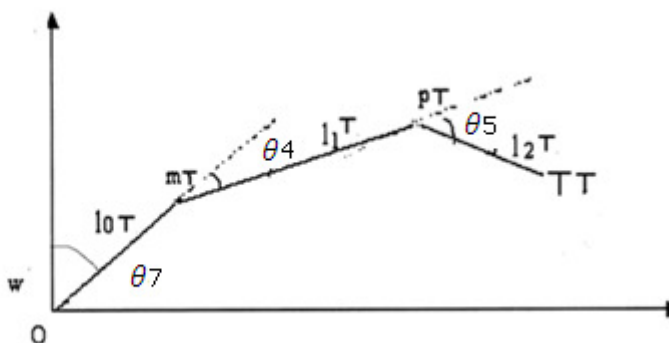


**FIGURE 3:a.** geometry of WT joint.      **FIGURE 3.b.** Thumb geometry.

## 3. HAND POSTURE ESTIMATION

Now the geometric characteristics of human hand must be analyzed with neutralizing color markers on the hand so the required information could be obtained from a 2D image, we assume that the image is aligned in the x-z plane and Y axis is the depth which must be calculated. First, the wrist line could be used to estimate its rotation around X axis, this rotational degree is also needed to compensate its side effect on the observed distance between MT and a fixe point on the line from W to Mm which is used to measure the thumb's second rotation $\theta8$, this angle could also be measured by simpler methods like what will be used on the four fingers but because of thumb's special geometry described method might bring more accurate results. Then we begin by the four fingers, simple relations from figure 2 could be obtained which makes R (distance between each fingers tip and wrist point W) a function of $\theta3$:

$$(w'p')^2 = l_0^2 + l_1^2 - 2l_0l_1 \cos(\pi - \theta_1) \tag{7}$$

$$(p't')^2 = l_2^2 + l_3^2 - 2l_2l_3 \cos(\pi - \theta_4) \tag{8}$$

$$\cos \angle m'p'w' = \frac{l_1^2 + (w'p')^2 - l_0^2}{2l_1(w'p')} \tag{9}$$

$$\cos \angle d'p't' = \frac{l_2^2 + (p't')^2 - l_3^3}{2l_2(p't')} \tag{10}$$

$$\angle w'p't' = \pi - \theta_3 - \angle m'p'w' - \angle d'p't' \tag{11}$$

$$R^2 = (w'p')^2 + (p't')^2 - 2(w'p')(p't')\cos\angle w'p't' \tag{12}$$

When processing the default picture of user's hand all hands measures are obtained simply, but when it is a object pictures turn the fingers posture in unknown and only available measure is R, but since it is very difficult to calculate $\theta 3$directly from equation 14 so the inverse method must be neutralized, it means that for each of the four fingers all amount of R with respect to changes in $\theta 3$ are calculated after default picture is processed, then when it is object pictures turn degree of $\theta 3$ is obtained comparing the new R with the amounts those are recorded in R-$\theta 3$ table or profile, in practice it takes a 90×4 matrix to save the table and those Ri are chosen which have less difference with obtained R from average variance of each column. Then if there is any rotation in wrist around X axis it reduces the length of the wrist's marker line with respect to its length in default picture. so the amount of this rotation ($\theta$) could be calculated simply, the thumb's rotational movement $\theta 8$ must be obtained with same method but what about the unpredicted change in related marker line's length which comes from the wrist's rotation? Let's take a look to figure 4.
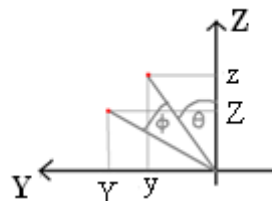


**FIGURE 4:** unknown angle calculation.

The wrist rotation makes side effects on thumb's measured distances on the hand, if $\theta$ is wrist rotation's degree and $\theta+\Phi$ is obtained rotation from the object picture for thumb the $\Phi$ is obtainable as $\theta 8$ and also side effects on other measured amounts are compensable through some simple triangular relations; according to described constraints other angles of rotation in fingers are relevant DOFs and could be calculated using mentioned relations. Thumb's more accurate geometric analyses and measuring wrist's rotation around X axis are two of the three main distinctions with other similar works like [2(5)], the other important matter is to calculate the angle between four fingers on the palm plane which are simply neglected in [21], and these angles are easily acquirable neutralizing simple triangular formulas.

$$\theta_2 = \sin^{-1}((2 * \sqrt{s * (s - Ri) * (s - R(i+1)) * (s - D)})/(Ri * (Ri + 1))) \qquad (13)$$

$$s = 0.5 * (Ri + R(i+1) + D) \qquad (14)$$

Now that all of the hand's variables are calculable so it is time to put them in practice, the following stages must be accomplished in order to generate all of the hand's geometric data:

1- Processing a default image of user's hand; in this stage real data like finger link's length and tip's distance from wrist must be generated and stored, the relations between this lengths are described in many references.

2- Generating R($\theta$3) curve (or table) as explained in relation (13).

3- Now the subject image of the user's hand is putted in program.

4- First the wrist roll rotation's amount is get comparing wrist line length with its length in default image; this amount of rotation is also required to compensate the other distances variation caused by perspective effect.

5- Calculating $\theta$8 (fig. 3) with the same bases as shown in figure 4.

6- Crossing measured distance between each finger's tip and wrist's center point with respected R($\theta$3) curve gives the current $\theta$3 angle.

7- The middle angles between two neighbor fingers is calculable if their tips distance with wrist (R) and two tips distance with each other is potted in equations (14) and (15).

8- The angle $\theta$8 is calculable with the same method mentioned in stage 7, in the thumbs natural varying rang [22] this angle is proportionate to the distance between MT and W (wrist line's centroid) points on hand , a linear relation can be assumed with its factor obtained in exercise.

9- Thumb's three remaining angles could be estimated using relations (6) and (7) and (8).

10- Now all of the hand's geometric data are generated and its model could be reconstructed in a robotic hand or drawn in a graphic application or paraphrased in a human- computer interface application.

## 4. RESULTS
A spherical coordinate based voluminous MATLAB program is created as mfile and many tradition vision and image processing methods are executed, following are object images and their containing hand's posture estimation are rearranged in neighboring figures.
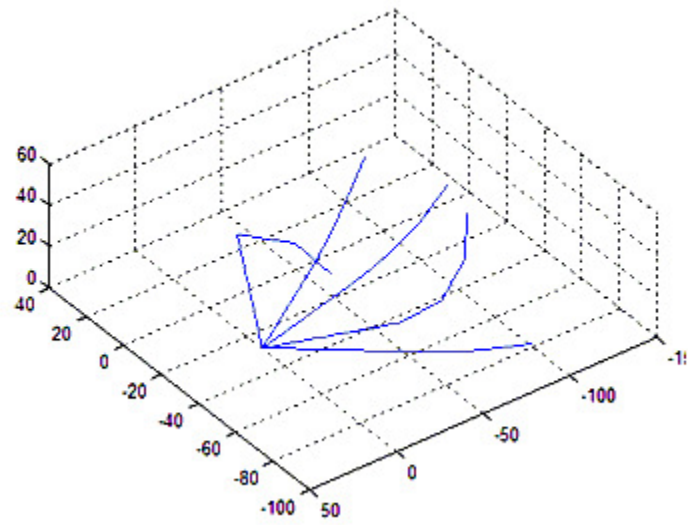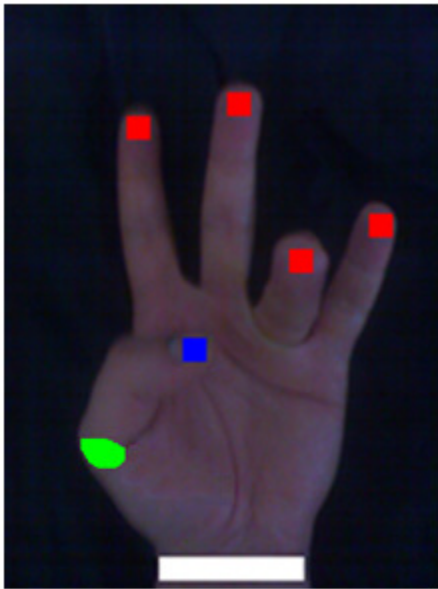
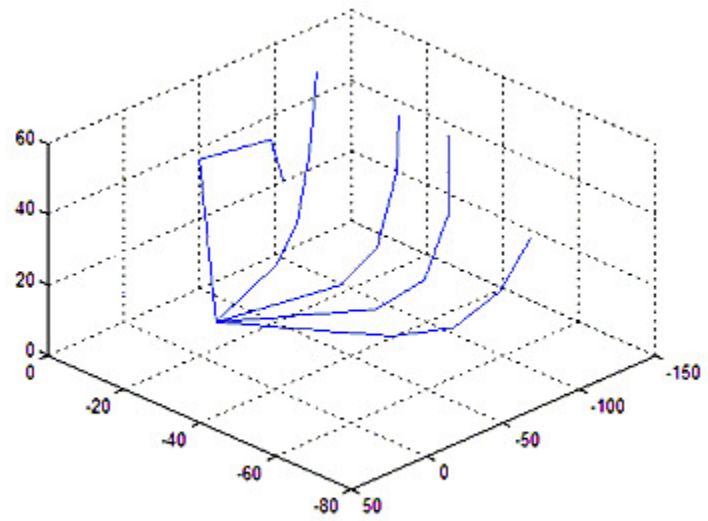**FIGURE 5:a.** original image of hand with color marks      **FIGURE 5:b.** reconstructed posture.
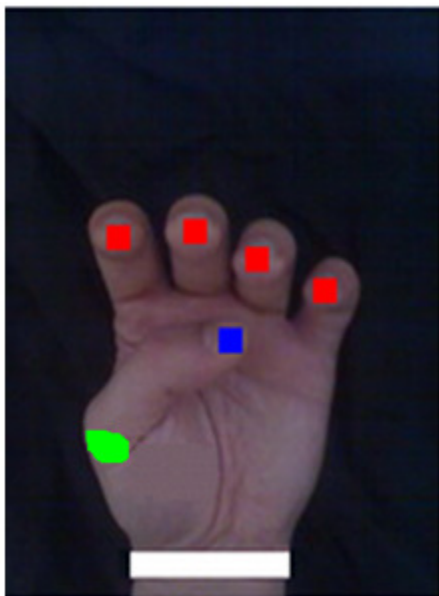


**FIGURE 6:a.** original image of hand with color marks      **FIGURE 6.b.** reconstructed posture.
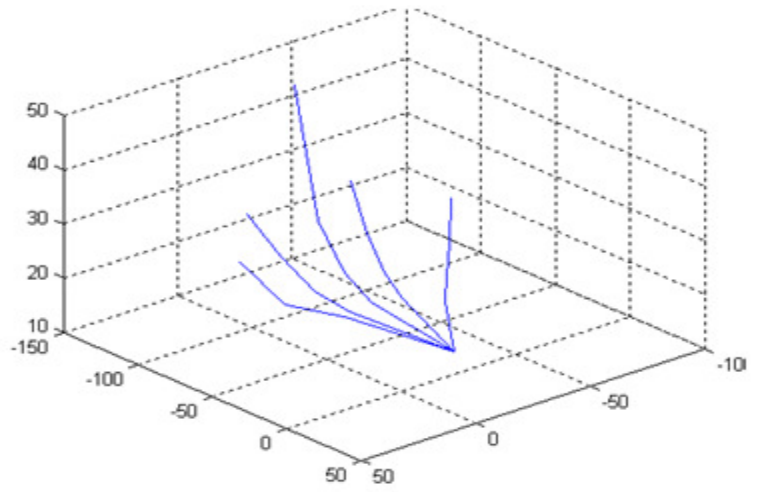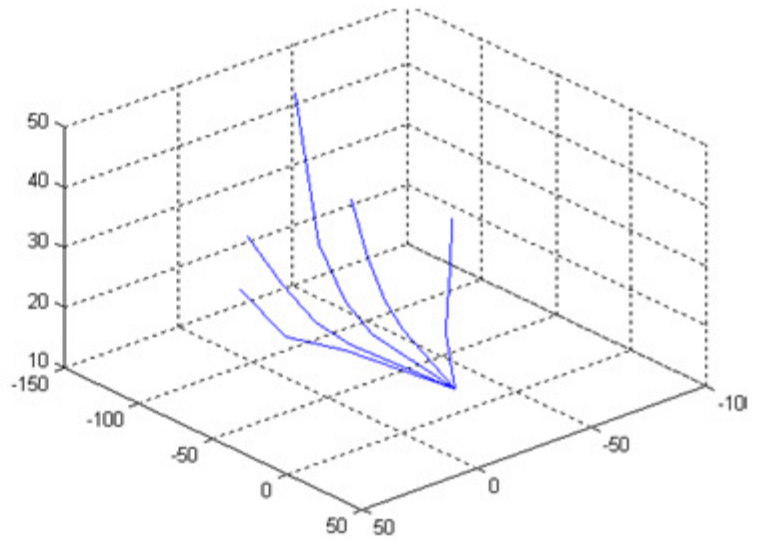
fig. 7-a , 7-b

**FIGURE 7.a.** original image of hand with color marks    **FIGURE 7.b.** reconstructed posture.



**FIGURE 8.a.** original image of hand with color marks    **FIGURE 8.b.** reconstructed posture.
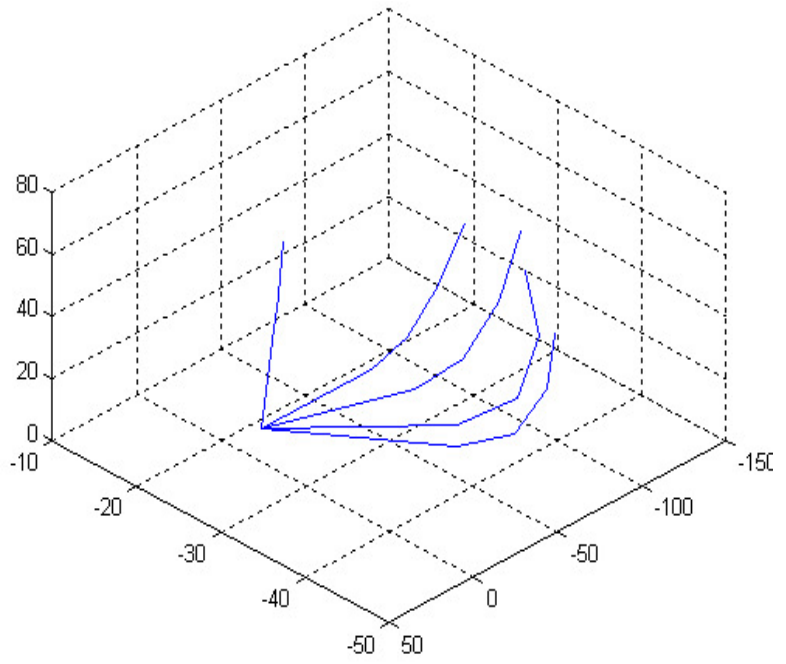
Mahdi Vaezi, Mohammad Ali Nekouie
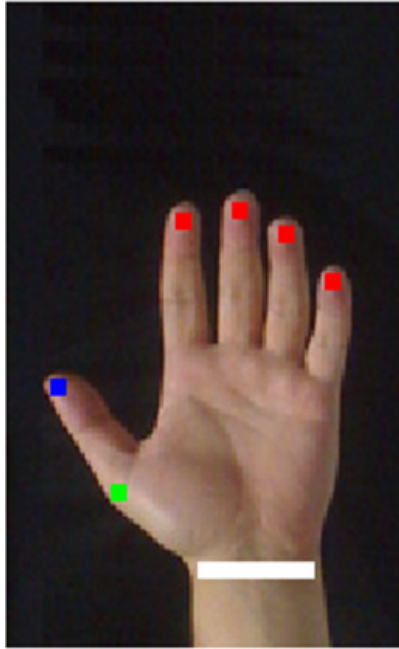
**FIGURE 9.a.** original image of hand with color marks       **FIGURE 9.b.** reconstructed posture.
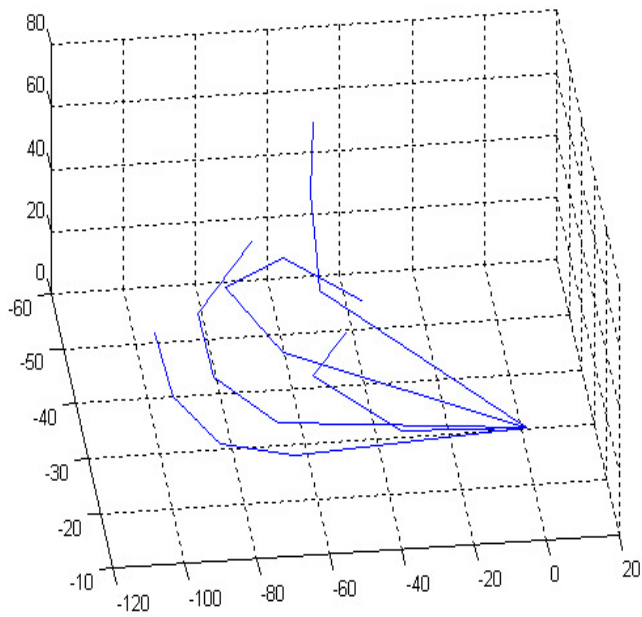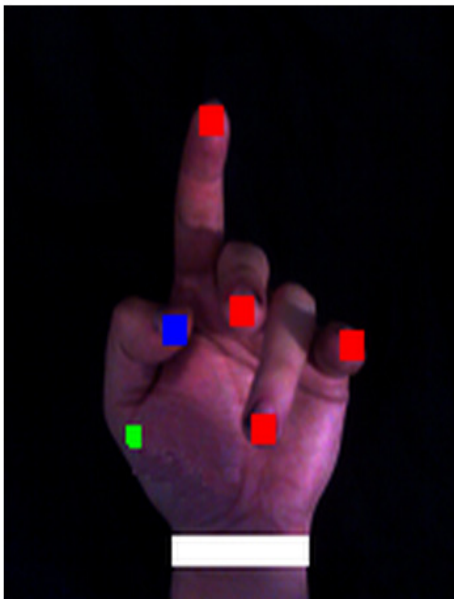


**FIGURE 10.a.** original image of hand with color marks       **FIGURE 10.b.** reconstructed posture.

## 5. CONCLUSION

We proposed an algorithm to estimate the 3D hand posture  using a 2D image, this is a promising method because it avoids 3D or tradition kinematic complexity and all required motion information should be easily obtained from image sequences, the only drawback point that might be noted is the necessity of color markers but it is obvious that alternating this method would require neutralizing complex and voluminous mathematic algorithms like couture fitting or texture processing [12] those will naught the profitability of work, further improvement could be obtained by  proposing a better method to measure thumb's rotation and perhaps utilizing a neural network to reduce the errors might be exist but since the only available way to measure the possible errors is subjective  method this is not  possible now, the presented work is executed with aim of robotics and human- computer interface project's image processing and data generating segments. As further work's direction, a method to vary the acquired factors in equations should be presented so the presented algorithm could be utilized for types of human hand which may be different in details without human operator's adjustments, for example by a neural network's management; this will require a device to measure the hands variables via a more precise way, like a glove as is described in [21], so the network training's necessary data could be obtained and the current method results could be assessed and  corrected so at the end this method's efficiency  improved as a individual procedure. In compare to similar works the following advantages are mentionable:

1- Only one color camera is required, unlike what are shown in [21] (instrumental glove in addition of vision devices) or in [23] (utilizing a whole MRI device) and many other works, this is very reasonable that this is the only requirement.

2- Number of calculated DOFs are increased by four concurrent with decrease in number of marked point on hand like what is presented in [20], even the represented algorithm to compute the variables is much simpler,  especially in obtaining thumb's angles, and the necessary machine vision and image processing methods are not very voluminous or complicated ones; these all matters make the current manner very suitable for planting on embedded systems.

3- By generating various properties in hand gestures now gestures recognition ideas can be easily achieved through assuming some tolerances in angles to recognize kinds of gestures, useful in intelligent interface systems and robotics assembly lines, etc.

## 6. REFERENCES

[1] J.Triesch, C.Vonder Malsburg. "A Gesture Interface for Human-Robot Interaction". The Proceedings of FG'98, Nara, Japan: 2-12, 1998.

[2] A.F.Bobick, A.D.Wilson. "A State-Based Approach to the Representation And Recognition Of Gesture". IEEE Transactions on Pattern Analysis And Machine Intelligence 19 (12), 1997.

[3] T.J.Darrel, A.Pentland. "Attention-Driven Expression and Gesture Analysis In An Interactive Environment". First International Conference on Automatic Face and Gesture Recognition, 1995.

[4] N.Shimada, Y.Shirai, J.Miura. "Hand Gesture Estimation and Model Refinement Using Monocular Camera Ambiguity Limitation By Inequality Constraints". Second International Conference On Automatic Face And Gesture Recognition, IEEE Computer Society Press, Los Alamitos, CA, 1998.

[5] V.I.Pavlovic, R.Sharma, T.S.Huang. "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review". IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (7), 1997.

Mahdi Vaezi, Mohammad Ali Nekouie

[6] T.E.Starner. "Unencumbered Virtual Environments". PhD Thesis, MIT Media Arts and Sciences Section, USA, 1993.

[7] J.S.Kim, W.Jang, Z.Bien. "A Dynamic Gesture Recognition System for the Korean Sign Language (KSL)". IEEE Transactions On System, Man, And Cybernetics 26 (2), 1996.

[8] S.Tamura, S.Kawasaki. "Recognition Of Sign Language Motion Images". Pattern Recognition 21 (4): 343-353, 1998.

[9] J.Lee, T.L.Knuii. "Model-Based Analysis Of Hand Posture". IEEE Computer Graphics and Application: 77-86, 1995.

[10] C.C.Lien, C.L.Huang. "Model-Based Articulated Hand Motion Tracking for Gesture Recognition". Image And Vision Computing 16: 121-134, 1998.

[11] J.M.Rehg, T.Kanade. "Visual Tracking Of High DOF Articulated Structures: An Application To Human Hand Tracking". ECCV94: B35-B46, 1994.

[12] C.C.Chang, W.H.Tsai. "Model-Based Analysis Of Hand Gestures From Single Images Without Using Marked Gloves Or Attaching Marks On Hands", Fourth Asian Conference On Computer Vision, 2000.

[13] C.S.Chua, H.Y.Guan, Y.K.Ho. "Model-Based Finger Posture Estimation". Fourth Asian Conference on Computer Vision, Taipei, Taiwan, 2000.

[14] H.Y.Guan, C.S.Chua, Y.K.Ho. "Hand Posture Estimation from 2D Monocular Image". Second International Conference On 3-D Digital Imaging And Modeling, Ottawa, Canada, 1999.

[15] H.Rijpkema, M.Girard. "Computer Animation Of Knowledge-Based Human Grasping". IEEE Computer Graphics And Application 25 (4): 339-348, 1991.

[16] J.M.Kuch, T.S.Huang. "Vision Based Hand Modeling And Tracking For Virtual Teleconferencing And Telecollaboration", ICCV95: 666-671, 1995.

[17] J.Lee, T.L.Kunii. "Constraint-Based Hand Animation, Models and Techniques in Computer Animation". IEEE Computer Graphics and Application, 1993.

[18] A.Hubertus, G.Medioni. "Detection of Intensity Changes with Sub Pixel Accuracy Using Laplacian-Gaussian Masks". IEEE Transactions on Pattern Analysis And Machine Intelligence 8 (5), 1986.

[19] F.Mokhtarian, A.Mackworth. "Scale-Based Description And Recognition Of Planar Curves And Two Dimensional Shapes". IEEE Transactions On Pattern Analysis And Machine Intelligence 8 (1), 1986.

[20] Ho,Yeong Khein. "Model-Based 3D Hand Posture Estimation From A Single 2D Image". Msc Thesis, Nanyang Technological University In Signgapor, 2001.

[21] M.Veber, T.Bajd, M. Munih. "Assessing Joint Angles In Human Hand Via Optical Tracking Device And Calibrating Instrumented Glove". Springer Science+Business Media: 1-13, 2007.

[22] E.P.Pitarch. "Virtual Human Hand: Grasping Strategy And Simulation". Phd Thesis, Universitat Polit`Ecnica De Catalunya, 2007.

[23] P.Smagt, G.Stillfried. "Using MRI Data to Compute A Hand Kinematic Model". Institute Of Robotics and Mechatronics, German Aerospace Center: 1-10, 2010.

[24] S.Cobos. "Efficient Human Hand Kinematics for Manipulation Tasks". IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008.